BİLKENT ÜNİVERSİTESİ
ENDÜSTRİ MÜHENDİSLİĞİ BÖLÜMÜ

ÜNİVERSİTE - SANAYİ/İŞ DÜNYASI İŞBİRLİĞİ PROJELERİ 2020

Derleyenler
Prof. Dr. Nesim Erkip
Prof. Dr. Savaş Dayanık
Yeşim Gülseren
Dr. Emre Uzun

ISBN: 978-605-9788-37-3
BASKI: Meteksan Matbaacılık, Temmuz 2020
YAYINCI SERTİFİKA NO: 27028
MATBAA SERTİFİKA NO: 13563

Düzenleme Kurulu

Prof. Dr. Nesim Erkip	*Bilkent Üniversitesi*
Prof. Dr. Savaş Dayanık	*Bilkent Üniversitesi*
Dr. Emre Uzun	*Bilkent Üniversitesi*
Yeşim Gülseren	*Bilkent Üniversitesi*
Ece Çiğdem	*Bilkent Üniversitesi*
Ege Bilaloğlu	*Bilkent Üniversitesi*
Berk Şahin	*Bilkent Üniversitesi*
Serkan Turhan	*Bilkent Üniversitesi*
İsmail Burak Taş	*Bilkent Üniversitesi*

Değerlendirme Kurulu

Prof. Dr. M. Selim Aktürk
Dr. Öğr. Üyesi Çağın Ararat
Dr. Öğr. Üyesi Özlem Çavuş İyigün
Prof. Dr. Savaş Dayanık
Prof. Dr. Nesim Erkip
Dr. Öğr. Üyesi Özlem Karsu
Prof. Dr. Bahar Yetiş Kara
Prof. Dr. Oya Ekin Karaşan
Dr. Öğr. Üyesi Ayşe Selin Kocaman
Dr. Öğr. Üyesi Çiğdem Gizem Körpeoğlu
Prof. Dr. Mustafa Pınar
Doç. Dr. Alper Şen
Dr. Öğr. Üyesi Firdevs Ulus
Dr. Emre Uzun

İÇİNDEKİLER

Bugüne kadar bu programa katılarak katkıda bulunan kuruluşlar:

2019-2020 Döneminde Projelerimize Katkıda Bulunan Kişilere Teşekkür Ederiz...

Bilkent Üniversitesi Endüstri Mühendisliği Bölümü'nden

Prof. Dr. M. Selim Aktürk
Dr. Öğr. Üyesi Çağın Ararat
Dr. Öğr. Üyesi Özlem Çavuş İyigün
Prof. Dr. Savaş Dayanık
Prof. Dr. Nesim Erkip
Dr. Öğr. Üyesi Özlem Karsu
Prof. Dr. Bahar Yetiş Kara
Prof. Dr. Oya Ekin Karaşan
Dr. Öğr. Üyesi Ayşe Selin Kocaman
Dr. Öğr. Üyesi Çiğdem Gizem Körpeoğlu
Prof. Dr. Mustafa Ç. Pınar
Doç. Dr. Alper Şen
Dr. Öğr. Üyesi Firdevs Ulus
Dr. Emre Uzun

Bilkent Üniversitesi Mühendislik Fakültesi
Üniversite-Sanayi İşbirliği Koordinatörü

Yeşim Gülseren

Bilkent Üniversitesi Endüstri Mühendisliği Bölümü
İdari Asistanı

Aslı Koca

İş Dünyasından Değerli Temsilciler

A101 - Yeni Mağazacılık	
Gürhan	Özesenli
Akbank T.A.Ş.	
Ege	Çağlar
Burcu	Alpul
Sibel	Yılmaz
Hakan	Şahin
Arçelik A.Ş. Bulaşık Makinesi İşletmesi	
Tanzer	Tuncalp
Tamer	Yıldırım
Serdar	Çevik
Ödül	Binbaşıoğlu
Arçelik A.Ş.	**Pişirici Cihazlar İşletmesi**
Altuğ	Hocaoğlu
Murat	Canlıoğlu
Burhan	Bahçecioğlu
Arçelik A.Ş. Buzdolabı İşletmesi	
Çiğdem	Koç
Sinem	Duman
Kerem	Alanlı
Arçelik A.Ş. Garage	
İbrahim	Güzel
Başkent Enerjisa	
Efe Can	Okumuş
Nisan Yagmur	Akkuş
Olcay	Cezayir
Ozge	Koç

BSH Ev Aletleri San. Ve Tic. A.Ş.

Onur	Durak
Esra	Gökçe

Brisa

Mustafa	Tacettin
Sinan	Daglioğlu

DHL

Serkan	Kabalı
Niyazi Omer	Usta

Dönmez Debriyaj San. Ve Tic. A.Ş.

Sabri	Tüysüz

Durukan Şekerleme San. Tic. A.Ş.

Gonca	Altuntaş
Tugce	Uzun
Hamdullah	Kaya

ETİ Gıda San.ve Tic. A.Ş.

Tolga	Köken
Birtan	Pala
Nusret Bora	Alptekin
Esra	Sancaklı

Erkunt Traktör A.Ş.

Dilara	Baykal Bilici
Bahar	Aydin Baydik

Havelsan A.Ş.

Deniz	Taşkın
Onder	Altıntaş

ING Bank

Zeynep	Gümüş

IBM

Emrah	Zarifoglu
Ilyas	Iyoob

Unilever Türkiye

Tuğçe	Serez
Kübra	Gürsoy Aslan
Furkan	Güçlü

ÖNSÖZ

Bu kitap, 2019-2020 öğretim yılında Bilkent Üniversitesi Endüstri Mühendisliği Bölümü tarafından gerçekleştirilen Üniversite-Sanayi İşbirliği Bitirme Projeleri özetlerini kapsamaktadır. Programımız 26 yıl önce sistem tasarımı derslerinin sanayi projelerine dönüştürülmesi ile başlamıştır. Bu süre içerisinde farklı sektör ve büyüklükte 107 şirketle toplam 491 proje gerçekleştirilmiştir.

Endüstri Mühendisliği Bölümü dördüncü sınıf öğrencilerinden oluşan proje ekipleri, akademik ve iş dünyasından danışmanların gözetiminde firmanın gündemine girmiş olan ve çözüm bekleyen gerçek problemlerini çözmektedirler. Yapılan projeler sonucunda ortaya çıkan ürün, yöntem veya hizmet, ilgili firmaya önemli yarar ve katma değer sağlamaktadır.

Endüstri Mühendisliği Proje Fuarı ve Yarışması, 2002-2003 yılında yapılan projelerin ilgili tüm firma, kuruluş ve üniversitelerle paylaşılması, iş dünyasının seçkin kuruluşlarının birbirleriyle ve üniversite ile olan etkileşiminin artırılması ve öğrencilerimizin iş hayatına daha donanımlı hazırlanmasını sağlamak amacıyla başlatılmıştır. Her yıl sistematik ve etkin bir şekilde yapılan bu çalışmaların daha kalıcı ve yaygın olarak paylaşılması amacıyla da "Endüstri Projeleri" kitabı serisi hazırlanmış ve bu dönemde gerçekleştirilen projeler gizlilik ilkesine bağlı kalınarak özet halinde sizlere sunulmuştur.

Kitaba girecek olan projelerin seçim aşamasında desteklerini esirgemeyen "Değerlendirme Kurulu"muza, fuar ve yarışma jürimizde görev alan Orhan Dağlıoğlugil (A101 – Yeni Mağazacılık), Barış Karakullukçu (Türk Telekom), Evren Cantürk (Akbank), Özgü Kokal (McKinsey & Company) ve Dr. Özlem Karsu 'ya (Bilkent Üniversitesi) teşekkür ederiz.

Prof. Dr. Nesim Erkip
Prof. Dr. Savaş Dayanık
Dr. Emre Uzun

Bilkent Üniversitesi Endüstri Mühendisliği Bölümü
Sistem Tasarımı Dersi Koordinatörleri

Bilkent Üniversitesi Endüstri Mühendisliği Bölümü Başkanı'ndan

Bilkent Üniversitesi Endüstri Mühendisliği Bölümü olarak öğrencilerimizin teknolojik ve sosyal değişikliklere uyum sağlayabilen, yaşam boyu öğrenmeyi hedefleyen ve sorgulayan iyi bir endüstri mühendisi olarak mezun olmasını amaçlamaktayız. Karmaşık sistemler ve problemlere bir bütün olarak bakabilme ve analitik düşünebilme yeteneğinin kazandırılması, eğitim programımızın en önemli amaçlarındandır. Bu doğrultuda Bölümümüz 2007 yılında Accreditation Board for Engineering and Technology (ABET) adlı bağımsız kuruluş tarafından eğitim kalitesini belgeleyen tam akreditasyonu Türkiye'de ilk alan mühendislik bölümüdür.

Eğitimde dünya çapında kalite standartlarını kullanan Bilkent Üniversitesi Endüstri Mühendisliği Bölümü, Üniversite-Sanayi İşbirliği adı altında ülkemizde örnek gösterilen programını 25 yıldır başarılı bir şekilde uygulamaktadır. Bu programın ana hedefi mezuniyet aşamasındaki öğrencilerimize kapsamlı ve derinlikli bir mesleki deneyim kazandırmaktır. Bu kapsamda 6-7 kişilik proje ekipleri, akademik ve iş dünyasından danışmanların gözetiminde firmanın gündemine girmiş olan ve çözüm bekleyen gerçek problemlerini çözmektedirler.

Bu yıl 18.sini düzenlediğimiz Endüstri Mühendisliği Proje Fuarı ve Yarışması'nda 21 proje yer almaktadır. Bu organizasyonda bütün bir yıl boyunca projeleri üzerinde özveri ile çalışan öğrencilerimizin çalışmaları sergilenmekte ve projelerine ait sunumlar yapılmaktadır. Öğrencilerimizi bu vesile ile kutluyor ve programa büyük katkıları olan tüm firma yetkililerine ve danışmanlarımıza teşekkür ediyorum.

Ayrıca bütün bu süreç boyunca yoğun ve özverili çalışmalarıyla programın hedeflerine uygun şekilde yürümesi için büyük çaba gösteren program koordinatörleri Prof. Dr. Nesim Erkip, Prof. Dr. Savaş Dayanık, Dr. Emre Uzun hocalarımıza, Üniversite-Sanayi İşbirliği Mezuniyet Projeleri Koordinatörümüz Yeşim Gülseren'e, asistanlarımız, Ece Çiğdem, Serkan Turhan, Berk Şahin, İsmail Burak Taş, Ege Bilaloğlu'na ve emeği geçen herkese çok teşekkür ediyorum.

Saygılarımla,

Prof. Dr. M. Selim Aktürk
Endüstri Mühendisliği Bölüm Başkanı

FİRMALARDAN GELEN TEŞEKKÜR MEKTUPLARI

ARÇELİK A.Ş. GARAGE'dan,

Dayanıklı tüketim ve tüketici elektroniği sektörlerinde üretim, pazarlama ve satış sonrası destek hizmetleri ile faaliyet gösteren Arçelik, 1955 yılında kurulmuş bir şirkettir. Arçelik bugün; dünya çapında 30.000'in üzerinde çalışanı, Türkiye, Romanya, Rusya, Çin, Güney Afrika, Tayland, Pakistan, Hindistan ve Bangladeş'de olmak üzere 9 ülkede, 23 üretim tesisi, 35 satış ve pazarlama ofisi, 13 markası (Arçelik, Beko, Grundig, Blomberg, Elektrabregenz, Arctic, Leisure, Flavel, Defy, Altus, Dawlance , Singer ve Voltas-Beko) ile 146 ülkede ürün ve hizmet sunmaktadır.

Ar-Ge birimini 1991 yılında oluşturan Arçelik bugün Türkiye ve dünyada 18 Ar-Ge merkezinde 1600'ün üzerinde Ar-Ge personeli ile kendi patentli teknolojilerini geliştirerek global pazarlarda rekabet edebilir bir güce ulaşmıştır. Kurulduğu ilk yıllardan bu yana işbirliklerinin öneminin farkında olarak üniversiteler ile farklı işbirliği süreçleri oluşturulmuştır. Lisans tez çalışmaları üniversite - sanayi işbirlikleri süreçlerinin içerisinde çok önemli bir halkayı oluşturmaktadır. Bu çalışmalar ile sanayinin tecrübe ettiği gerçek sorunlara çözümler bulunabilmekte, yenilikçi ürün ve süreçler geliştirilmesine katkı sağlanabilmektedir. Bununla birlikte mühendis adaylarımızın sanayi tecrübesini yaşayarak mezun olmalarına da katkı sağlamaktadır. Tamamlanan projeler ile edinilen bilgi ve tecrübenin mühendis adaylarına ileride yapacakları çalışmalarda yol gösterici nitelikte faydalar sağlayacağını öngörüyoruz.

Bu anlamda Bilkent Üniversitesi'nin Sanayi Odaklı Bitirme Projeleri kapsamında yürütülen çalışmaları çok değerli buluyoruz. Bilkent Üniversitesi Rektörlüğü ve Mühendislik Fakültesi yönetici ve akademisyenlerimize, üniversite – sanayi işbirliği yürütüğümüz projelerde hedeflenen çıktı ve kazanımlara ulaşmamızda bizlere destek olan değerli mühendis adaylarımıza ve süreç içerisinde değerli katkıları ile projelerin uygulanabilirliğine yönelik geri bildirimleri sağlayarak proje ekibimizi başarılı sonuçlara yönlendiren Endüstri Mühendisliği ve Elektrik – Elektronik Mühendisliği Departmanlarındaki değerli akademisyenlerimize teşekkür ederiz.

Evrim ÖZGÜL
Arçelik A.Ş.
Global Ar-Ge Teşvikleri ve Üniversite – Sanayi İlişkileri Yöneticisi

Arçelik A.Ş. Bulaşık Makinası İşletmesi'nden

1993 yılında Sincan Organize Sanayi Bölgesi'nde yıllık 200.000 adetle başlayan yolculuğumuza 2020 yılında 2,5 milyon bulaşık makinesi hedefiyle devam ediyoruz.

Bugün geldiğimiz bu noktada dünya pazarının %10 nundan fazlasını tek çatıda üretebilen bir işletme olmanın gururunu yaşarken , diğer taraftanda bu ölçeğin getirmiş olduğu değişkenliği yönetmede her gün daha çok zorlanır hale gelmiş durumdayız.

Böyle bir ortamda **Taktiksel Üretim Planlama için Karar Destek Sistemi** projemizle hem daha hızlı , keskin ve adaptif planlar yapabilmeyi hem de bu konuda farklı gözlerin sürecimize desteklerini hedefleyerek projemize adım attık.

Tüm planların COVİD19 kapsamında revise olduğu bu yıl projemizi zamanında eksiksiz kapatabilmenin başarısını yaşadık.
Üniversite sanayi işbirliği proje zincirimizin son halkası olan bu projede :

- Akademik katkıları için Sayın Nesip ERKİP hocamızın şahsında Bilkent Endüstri Mühendisliği nin değerli akademik ekibine ve
- Projeyi başarıyla tamamlayan EKİP 17'nin değerli meslekdaşlarımıza
- Her yıl olduğu gibi bu yılda üniversite Sanayi işbirliği alanında bizleri bir araya getiren Mezuniyet Projeler Koordinatörlüğü'ne

teşekkür ediyoruz.
İlerki dönemlerde sağlıkla karantinasız bir araya gelerek başka projeşlerde buluşabilme ümidiyle…

N. Tanzer Tunçalp

Arçelik A.Ş.
Bulaşık Makinesi İşletmesi
Üretim Planlama Yöneticisi

Arçelik A.Ş. Pişirici Cihazlar İşletmesi'nden,

Şirketimiz Arçelik Dünya çapında 30,000 çalışanı ile 11'den fazla markayla 145'ten fazla ülkede ürün ve hizmet sunmaktadır. Türkiye'de ankastre ve solo beyaz eşya pazarında 1. olan şirketimiz Avrupa'da da beyaz eşya pazarında 2. ve solo beyaz eşya pazarında 1. olarak ülkemizi gururla temsil etmektedir.

Pişirici Cihazlar İşletmesi olarak fabrikamızda solo fırın, ankastre fırın, setüstü fırın, ocak ve davlumbaz üretmekteyiz. Yıllık 4 milyonu aşan üretim adediyle müşterilerimize daha ucuz ve daha kaliteli ürünler sunmayı hedeflemekteyiz.

İşletme olarak son dönemlerde üretim adetlerindeki artışlara cevap vermek ve Üniversite Sanayi İşbirliği'ne katkı vermek amacıyla 2019 Eylül ayında Bilkent Üniversitesi Endüstri Mühendisliği Bölümü'nden 7 öğrenci arkadaşımızla başladığımız Paralel Makine Sistemleri için İş Ataması Yoluyla Kapasite Belirlenmesi Projesi'nde beklentilerimizin üzerine çıkmış bulunmaktayız.

Bu projeyle beraber artan üretim adetlerine daha iyi cevap verebileceğiz ve mekanik üretim birimimizdeki üretim maliyetinde ortalama %8.11'lik düşüş görmekteyiz. Ek fayda olarak bir paket halinde oluşturulan karar destek sistemi yapacağımız yeni yatırımlarda kapasite analizlerimizi daha kolay yapmamızı sağlayacak.

Emeği büyük olan ve pandemi süresince dikkatlerinde en ufak bozulma olmayan proje ekibimize teşekkür borçluyuz. Proje süresince hem proje ekibine hem de bize yol gösteren Bilkent Üniversitesi Endüstri Mühendisliği Akademik kadrosuna ve Bilkent Üniversitesi Mezuniyet Projeleri Koordinatörlüğü'ne teşekkürlerimizi sunarız.

Altuğ Hocaoğlu

Arçelik Pişirici Cihazlar İşletmesi Üretim Mühendisliği Yöneticisi

Brisa A.Ş'den,

Bridgestone Corporation ve Sabancı Holding iştiraki olarak, Bridgestone ve Lassa markalı lastiklerle Türkiye lastik sektörünün lideriyiz. 1974'ten beri İzmit'teki ve 2018'den bu yana Aksaray'daki akıllı teknolojilerle donatılmış fabrikamızla Türkiye'nin büyümesine katkı sağlıyoruz.

Bilkent Üniversitesi, güçlü akademik kadrosuyla Türkiye'nin önde gelen ve başarılarıyla gururu olan bir üniversitemiz. Bu yıl biz de Bilkent Üniversitesi Endüstri Mühendisliği son sınıf öğrencilerinin çok uzun zamandır Türkiye'nin önde gelen şirketleriyle yürüttükleri projelerden birinde yer almaktan büyük mutluluk duyuyoruz.

Brisa olarak, Bilkent Üniversitesi'nde yetişen öğrencilerin ve danışman hocalarıyla birlikte, şirketimizde karıştırma prosesindeki olası problemleri önceden tahminleyen bir modeli geliştirdik. Bu projenin fabrikamızın operasyonel verimliliğine katkı sağlamakla beraber, bir üniversite-sanayi işbirliğinin katkısı ve potansiyelini görmek açılarından da bize çok faydalı oldu. Öğrencilerimizin neler yapabileceği konusunda fikir sahibi olduk ve bundan sonraki yıllarda yapılabilecek işbirlikleri konusunda da alternatif projeleri şimdiden düşünmeye başladık.

Bu sene yürüttüğümüz projedeki katkılarından ve özverili çalışmalarından dolayı, Bilkent Üniversitesi Endüstri Mühendisliği Bölümü'ne, değerli öğrencileri ve akademik danışmanlarına, Üniversite-Sanayi İşbirliği Koordinatörlerine teşekkürlerimizi sunarım.

Mustafa Tacettin
Aksaray Fabrika Direktörü

Durukan Şekerleme A.Ş'den,

durukan

16.04.2020

BİLKENT ÜNİVERSİTESİ ENDÜSTRİ MÜHENDİSLİĞİ BÖLÜMÜ'NE

ANKARA

Üniversiteniz tarafından koordine edilen sanayi odaklı bitirme projeleri kapsamında Ekim 2019'da başlamış olduğumuz "Talep Tahmini ve Stok Eniyilemesi için Karar Destek Sistemi Tasarımı" projesinde 15.04.2020 proje ekip üyeleri tarafından gerçekleştirilen sunumda gelinmiş olan nokta ve ortaya konan iş çıktıları bizi çok mutlu etti.

Bu çalışmanın gerçekleştirilmesindeki destek ve katkılarından dolayı proje koordinatörü Sn Emre Uzun'a, akademik danışman Sn. Emre Nadar'a teşekkürlerimizi sunarız.

Özverili çalışmaları ile bu projeye imza atmış olan İhsan Doğramacı Bilkent Üniversitesi Endüstri Mühendisliği öğrencileri Sn. Osman AKGÜN, Sn. Umay ATAY, Sn. Bilge BALCILAR, Sn. Özge ÇALIŞKAN, Sn. Emre Can DURMUŞ, Sn. Şimal İrem GÖKALP ve Sn. Neslişah ÖZKAN'a teşekkür ederiz.

Uğur DURUKAN
Yönetim Kurulu Başkanı

ANKARA
A.S.O. 1. Organize Sanayi Bölgesi Dağıstan Caddesi No:11 06935
Sincan - Ankara / TÜRKİYE Tel: +90 (312) 268 71 71 Fax: +90 (312) 268 73 73
Ankara Ticaret Sicili Müdürlüğü - Ticaret Sicil Numarası: 65768

İSTANBUL ŞUBE
Yeşilköy Mahallesi Atatürk Caddesi No:10 İDTM A2 Blok Kat: 14/424
34149 Bakırköy - İstanbul / TÜRKİYE Tel: +90 (212) 578 37 56 Fax: +90 (212) 578 73 73
İstanbul Ticaret Sicili Müdürlüğü – Ticaret Sicil Numarası: 965749

Sincan Vergi Dairesi – 320 000 2083
Durukan Şekerleme San. Ve Tic. A.Ş.
E-mail: info@durukan.com.tr Web: http://www.durukan.com.tr

E-Ticaret Deposunda Sipariş Toplama, Dağıtma ve Paketleme Süreçlerinin Optimizasyonu

A101 Yeni Mağazacılık A.Ş.

Proje Ekibi

Yağız Dilber, Büşra Güray, Ege Semih Kozoğlu, Batur Kutlay, Cemre Polat, Mehmet Burak Sayıcı, Görkem Tüzel

Şirket Danışmanı
Gürhan Özesenli
Tedarik Zinciri Direktörü

Akademik Danışman
Dr. Öğr. Üyesi Ayşe Selin Kocaman
Endüstri Mühendisliği Bölümü

ÖZET

A101 Yeni Mağazacılık A.Ş. e-ticaret operasyonları için yeni bir depoya sahiptir. Üç ana süreç olan toplama, dağıtma ve paketleme süregelen sistemde hedeflenen verimde uygulanamamaktadır. Değişken talepler, agresif kampanya stratejileri, çalışanların değişken performansları darboğazlara sebep olmaktadır. Proje amacı, bahsedilen üç süreç için dengeli ve verimli çalışan planlaması oluşturan bir destek aracı geliştirmektir. Çözüm yaklaşımında fazla talep modeli ve denge modeli olmak üzere iki farklı matematiksel model geliştirilmiştir. Günlük siparişlerdeki ürün sayısına göre çalıştırılacak model belirlenir. Fazla talep modeli, her vardiya için karşılanamamış talebi enazlamayı; denge modeli, farklı performans gruplarındaki çalışanlar arası varyansı enazlamayı amaçlar. İki modeli de içeren karar destek aracı Python kullanılarak geliştirilmiştir.

Anahtar Kelimeler: Depo operasyonları, iş gücü dağılımı, karar destek aracı

Optimization of Order Picking, Sorting and Packaging Processes in A101 E-Commerce Warehouse

1. General Information About the Company

A101 Yeni Mağazacılık A.Ş. started operating in Turkey in 2008. By 2020, the company has 9000 stores in 81 cities of Turkey and the distribution of goods is maintained with 46 warehouses. 1250 SKUs are available in A101 stores. In 2018 the company started its operations on e-commerce channel. On a101.com.tr an average of 3600 SKUs is on sale each day.

2. System Analysis

2.1. Problem Definition

Although aggressive campaigning strategy can be effective for increasing market share and sales volumes, the company struggles to establish the warehouse management due to high variability in demand. During the campaign times, satisfying the service level becomes an issue due to high demand. The main problem is to satisfy excess demand as soon as possible with correct workforce allocation. Those campaigns load stress on the workforce and create excess traffic in the warehouse. Shipment up to 2 days is normal in warehouse. However, this excess demand hinders the delivery time. Lack of arrangement blocks detecting bottlenecks and decreases reaction time, so unbalanced workforce occurs, resulting in low worker utilization and significant number of customer complaints.

2.2. Objective & Deliverables

In the project, there are two different mathematical models. Excess demand model is executed if the total expected demand obtained in a campaign cannot be handled by the warehouse. Objective of excess demand model is minimization of excess demand in the system. Balance model is executed in case the total expected demand is handled by warehouse. Objective of balance model is minimizing variances of the workloads for each worker level. Additionally, to implement randomness in the system, another model is coded in Arena Simulation. The simulation model mimics the warehouse operations with assigned workers while giving the completion time of orders. The designed models will be operated throughout campaign days and will provide shift-based information. Main deliverable of the system is a decision support tool which will suggest a worker allocation according to worker performance levels.

3. Literature Review

According to the article written by Postema (2017), a heuristic approach for the management of capacity and order is proposed. Several researches are studied in the article regarding the routing problem of warehouses. Aykin (2000) analysed approaches regarding shift scheduling problem and one of them is the set covering model approach. For the excess demand model, the goal of set covering model is used and it is extended according to warehouse's constraints. Set covering model's main approach is minimizing number of workers by satisfying needed number of workers for each shift. Thus, the extension includes the addition of minimizing the excess demand as a goal to excess demand model. For the balance model in the project, Dewi and Sestina's (2015) workforce scheduling approach is used. In the article, objective is to minimize the sum of variances of workloads of each level of worker. In the study, the deviations are equalized to difference between average workload and assigned worker's workload. For our project, the objective is extended and written according to variances of each worker level in the balance model. A, B and C determines the levels of performances as high, average and low respectively. These levels will be further explained in the following parts.

4. Methodology

4.1. Inputs and Outputs of the System

Our model takes various inputs such as the calculated demand, the number of the workstations and worker's performances. The output of the system gives worker allocation with the time needed to satisfy the demand., The interface of the tool demonstrates the assigned schedule and the time interval needed for finishing the demand with allocated workforce.

4.2. Solution Approach

For the two different mathematical models, decision regarding when to use is given by considering the demand value causing bottleneck. Demand levels are determined by using clustered campaign data. A threshold value which is the demand amount causing bottleneck is calculated considering the maximum capacity of warehouse when all stations are used with maximum performance per shift. In a shift, it is found that maximum 9480 units of demand can be satisfied. This value is used to decide and execute the appropriate model. If the given demand is bigger than threshold value, the excess model is executed. Otherwise, the balance model will be executed. Since the goals are different, different KPIs are considered. These are demand satisfaction in excess demand model and minimizing total variance in balance model. This heuristic approach is demonstrated in Figure 1 below. The models within the heuristic will be furtherly explained in the following sections.

One of the most important aspects about schedules is the classification of worker performances which are A, B and C. These are based on time studies done in line with Pareto principles. The workers are clustered into 3 groups. Group A represents the top 10% performance segment (completion time). Group B and C represent the following 80% and the bottom 10% respectively.

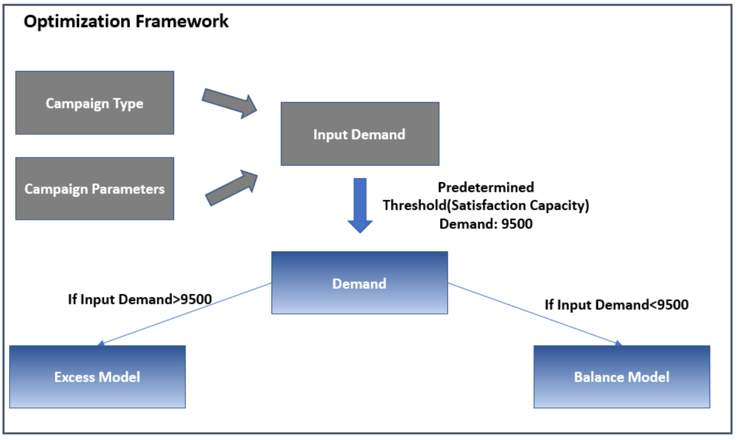


Figure 1. Explanation of the models' execution.

4.2.1. Excess Demand Model (Model 1)

The excess demand exists when default workforce allocation cannot satisfy the input demand. Main purpose of this model is to properly allocate workforce classes (A-B-C). The model's objective function provides us to minimize excess demand in each shift of the week. In the Appendix A, you can see the mathematical model. It includes several constraints. Firstly, constraints' purposes are demand satisfaction by defining excess demand. Initialization of the system constraints are added. The second and third shifts for all days (1, 2...6) and provide excess demand variable and demand satisfaction are showed. The first shifts for the days 2...6 so that 3rd shift's excess demand of the day i can be transferred to 1st shift of day (i+1) are showed. Feasibility constraints for the limited number of workstations and number of workers are added. Non-negativity of variables regarding excess demand and number of workers are also added.

Verification of the model is coded in Python. Number of workers in three activities are recorded for different scenarios with different demands. Warehouse engineer Can Sevilmiş confirmed the scenarios, he agreed that the number of workers in three activities for different demands is realistic and similar to his previous daily plans. Besides providing the schedule for the warehouse, the excess demand for each activity of each shift is also verified and preliminary results are obtained.

After the verification process, a schedule for the warehouse is found in Python to validate the model. Then, with the help of the found schedule, an extra sorter to the warehouse is added so the sorter capacity which was 4 before is changed to 5. Since the KPI in excess demand model is lead time of the demand satisfaction, a significant effect can be clearly observed. By adding the extra sorter, lead time is decreased by approximately two shifts compared to the system with four sorters. Figure 2 shows comparison of the number of shifts which are needed to handle excess demand during campaign. This shows that with one extra sorting station, in average, %30 of the lead time would be decreased. This result is validated by the company data provided.

Demand	W/4 Sorter	W/5 Sorter
3000	3	1
4000	5	3
4500	6	4
5000	6	5
5500	7	5
6000	8	6
6500	9	6
7000	9	7

Figure 2. Comparison of number of shifts required to satisfy the specified demand for 4 sorters and 5 sorters

4.2.2. Balance Model (Model 2)

Theoretically, the minimum completion time can be achieved by assigning the best workers (Level A) since they are more experienced and faster than the rest. Assigning all jobs to the same group of people will violate the fairness among workers in real life circumstances. In the Appendix B, you can see the mathematical model. The objective of the model is to minimize the variance between clustered workforce. The most prior constraint is demand satisfaction. Other constraints are limiting to A, B & C level workers since warehouse has limited number of workers (feasibility constraints). The mean workload for level A, B & C workers (Z_A, Z_B, Z_C) is calculated. How much load each group level deviated from the average workload is calculated lastly.

This model is verified on Python. A comparison is made between “below bottleneck case” and “at the bottleneck case”. Figure 3. shows that the allocation of the workers according to 9491 and 6500 units of demand. The allocations show that this model’s results fit with the real time results of the warehouse. In figure 3, worker allocations of each workstation with the worker levels for different demands is provided.

```
Demand :  9491 products, assignments are:
SortA: [3.0]
SortB: [1.0]
SortC: [0.0]
---
PackA: [0.0]
PackB: [11.0]
PackC: [1.0]
---
PickA: [0.0]
PickB: [4.0]
PickC: [3.0]
```

```
Demand :  6500 products, assignments are:
SortA: [0.0]
SortB: [3.0]
SortC: [0.0]
---
PackA: [1.0]
PackB: [6.0]
PackC: [0.0]
---
PickA: [0.0]
PickB: [0.0]
PickC: [5.0]
```

Figure 3. Allocation comparison of two different demand units (at & below the bottleneck value)

In order to perform the validation process of the worker allocation balancing model (Model 2); it is compared with existing warehouse allocation which ignores the workload balance. The goal is to minimize the total variances between each worker level A, B and C. The model is only executed for the demand values under the amount causing bottleneck since excess demand model is used for demand above it. For both models, deviations for all worker combinations of each demand are determined according to average workload calculated for each worker and workload per worker of each level. Regarding the KPI which is how much the total variance is minimized, it is concluded that balance model brings 80% improvement approximately.

4.2.3. Simulation Model

A simulation model in Arena is developed to see the performance of the warehouse in a random environment. Flow of the orders and operations are implemented in the Arena. Three processes, creation of orders, number of workers in each performance level and their speed, batch size of the picking cart, completion time of the demand satisfaction are considered. The advantage of the simulation model is that the random distributions are used for the operations. It is also used for investigation of batch size's effect on completion time of demand satisfaction.

Process Analyzer tool is used to test the validity of the model with different scenarios. While running the model, the inputs (demand, number of workers, batch size) being used are used confirmed by the warehouse representative. It is expected that as the number of workers decreases, time of the demand satisfaction increases. The model is verified by comparing Scenario 1 with other scenarios which can be seen at Figure 4.

	Scenario Properties				Controls			Response
	S	Name	Program File	Reps	Resource PickA	Resource PackA	Resource SortA	ttime
1		Resources initial	2 : A101	3	1.0000	1.0000	1.0000	6.365
2		Picker Resource increases	2 : A101	3	2.0000	1.0000	1.0000	6.252
3		Packer Resource increases	2 : A101	3	1.0000	2.0000	1.0000	6.037
4		Sorter Resource increases	2 : A101	3	1.0000	1.0000	2.0000	6.343
5		Picker Resource increases more	2 : A101	3	3.0000	1.0000	1.0000	6.343
6		Packer Resource increases more	2 : A101	3	1.0000	3.0000	1.0000	5.599
7		Sorter Resource increases more	2 : A101	3	1.0000	1.0000	3.0000	6.347
8		Picker Resource decreases	2 : A101	3	0.0000	1.0000	1.0000	6.442
9		Sorter Resource decreases	2 : A101	3	1.0000	1.0000	0.0000	6.432
10		Packer Resource decreases	2 : A101	3	1.0000	0.0000	1.0000	6.671

Figure 4. Scenarios for Verification of Simulation Model

Process Analyzer tool is used to test the validity of the model with 50 different scenarios which can be observed in the warehouse conditions. In these scenarios, for each activity, number of people is randomly assigned. Same demand satisfaction is ensured in different lead times. The average lead time of 50 scenario is 6.630 hours. For this model, lead time of demand satisfaction is chosen as key performance indicator (KPI). Thus, when the scheduled solution is compared with the Process Analyzer (PAN) results, there is 4% decrease in the lead time so resource allocation makes the warehouse operate more efficiently.

4.3. User Interface

The tool is utilized using Python. Python user interface is coded by calling Tkinter Library. User interface includes three section. First one is for planning workers schedule. At planning section, inputs are campaign demand, maximum number of A-B-C level workers for each station, workers' performance and average product amount in each order. With that parameters, user interface gives opportunity to access Excess Demand model and Balance model. 7-days workers schedule is given as a result for each station. Second section includes image for insight about main campaign's demand. Third section is for picking time prediction with number of travelled address in one picking tour. The visuals of the interface for each section can be seen in the Appendix C.

5. Benefits to the Company and Implementation Plan

The main benefit of the project is a parametric user interface which is a standardized tool that can be used at any warehouse circumstances and address any needs. The user can change;

- Picker, sorter and packager performances
- Maximum number of pickers, sorters and packagers
- Number of A, B and C level of workers
- Campaign demand
- Average product number in one order
- Average demand for each day of the week

Other benefits related with the models and user interface are;

- Ensuring the fairness between the workers by trying to assign workers according to their performance ratings based on the balance model
- Satisfying the campaign demand in minimum number of shifts by assigning workers according to their performance ratings based on the excess model
- Faster scheduling for worker allocation which was a manual process before

As of May 1st, a meeting was conducted with company to decide on implementation plan due to recent changes in warehouse environment and pandemic situation. The updates are;

- Warehouse moved to a new area and 7-days operation started
- Lack of data in the new warehouse environment leads to inconvenience about test drive
- Moving procedure is not finished yet and operations continue in both warehouses
- No picking and sorting algorithm is defined for the new warehouse yet so work flow is changed in the new warehouse

- Implementation is not feasible in short period due to issues company deals with

According to these updates, the long-term implementation plan is aligned between all stakeholders. This implementation plan includes finalization of decision support tool, delivery of final decision support tool to the company with user manual guide and detailed training video which are completed and delivered as of May 18th. Warehouse coordinator Can Sevilmiş and industrial advisor Gürhan Özesenli concluded that the tool is feasible. The company will collect data in the new warehouse environment for more accurate test phase of the tool in July.

REFERENCES

Aykin, T. (2000). A comparative evaluation of modeling approaches to the labor shifht scheduling problem. *European Journal of Operational Research 125*(2), 381-397.

Deloitte. (2019). *2019 Global Impact Report* [PDF file]. Retrived from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Consumer-Business/cons-global-powers-retailing-2019.pdf

Dewi, D. S., &Septiana, T. (2015). Workforce scheduling considering physical and mental workload: a case study of domestic freight forwarding. *Procedia Manufacturing*, *4*, 445-453.

Postema, J. T. (2017). Metaheuristics for Order Batching in Ecommerce Warehouses.

APPENDICES

Appendix A

The Excess Demand Model (Model 1)

Indices

- j represents level of works which are A, B, C
- i represents type of workers which are picker, sorter, packager and 1, 2, 3 respectively
- k represents shifts 1,2,3,4
- d represents days of the weeks which are 1,2,3,4,5,6,7 and 1 represents campaign starting day

Parameters

$CAP_{i\,j}$: Capacity of j level worker at activity i

$D_{d\,k}$: Demand of day d at shift k

Decision Variables

$X_{d\,k\,i\,j}$: Number of j level workers at day d, shift k for activity i

$E_{d\,k\,i}$: Excess demand at day d, shift k, activity i

Model

$$\min \quad \sum_d \sum_k \sum_i E_{d\,k\,i} \tag{1}$$

subject to

$$\sum_{j=1}^{3} X_{1\,1\,i\,j}\, CAP_{1\,j} + E_{1\,1\,i} = D_{1\,1} \qquad \forall i \tag{2}$$

$$\sum_{j=1}^{3} X_{d\,k\,i\,j}\, CAP_{i\,j} + E_{d\,k\,i} = D_{d\,k} + E_{d(k-1)\,i} \qquad \forall i \text{ and } d = 1..7 \quad k = 2,3,4 \tag{3}$$

$$\sum_{j=1}^{3} X_{d\,1\,i\,j}\, CAP_{i\,j} + E_{d\,1\,i} = D_{d\,1} + E_{(d-1)\,4\,i} \qquad \forall i \text{ for } d = 2..7 \tag{4}$$

$$\sum_{k=1}^{3} \sum_{i=1}^{3} X_{d\,i\,k\,1} \le \#\text{ of A level workers} \qquad \forall d \tag{5}$$

$$\sum_{k=1}^{3} \sum_{i=1}^{3} X_{d\,i\,k\,2} \le \#\text{ of B level workers} \qquad \forall d \tag{6}$$

$$\sum_{k=1}^{3} \sum_{i=1}^{3} X_{di\,k\,3} \le \#\text{ of C level workers} \qquad \forall d \tag{7}$$

$$\sum_{j=1}^{3} X_{d\,2\,k\,j} \le 4 \qquad \forall d, \forall k \tag{8}$$

$$\sum_{j=1}^{3} X_{d\,3\,k\,j} \le 12 \qquad \forall d, \forall k \tag{9}$$

$$X_{d\,k\,i\,j} \text{ integer} \tag{10}$$

$$E_{d\,k\,i} \ge 0 \tag{11}$$

Appendix B

The Balance Model (Model2)

Indices

- j represents level of works which are A, B, C
- i represents type of workers which are picker, sorter, packager and 1, 2, 3 respectively

Parameters

W: Demand unit.
CAP_{ji}: Capacity of j level worker in activity i
Z_j: Average workload of j level worker in every activity i.
V_J: Variance of workload of j level worker.
d: Mean workload for all workers

Decision Variable

X_{ji}: Number of workers at j level in activity i

Model

minimize $v_A+v_B+v_C$ (1)

subject to

$$X_{ji} \leq \frac{W}{Cap_{ji}} \quad \forall i, \forall j \quad (2)$$

$$\sum_{j=1}^{3} \text{Capji.Xji} \geq \text{W} \quad \forall i \quad (3)$$

$$\sum_{i=1}^{3} X_{Ai} \leq 5 \quad (4)$$

$$\sum_{i=1}^{3} X_{Bi} \leq 13 \quad (5)$$

$$\sum_{i=1}^{3} X_{Ci} \leq 5 \quad (6)$$

$$Z_A=\frac{X_{A1}.Cap_{A1}+X_{A2}.Cap_{A2}+X_{A3}.Cap_{A3}}{\sum X_{Ai}} \quad (7)$$

$$Z_B=\frac{X_{B1}.Cap_{B1}+X_{B2}.Cap_{B2}+X_{B3}.Cap_{B3}}{\sum X_{Bi}} \quad (8)$$

$$Z_C=\frac{X_{C1}.Cap_{C1}+X_{C2}.Cap_{C2}+X_{C3}.Cap_{C3}}{\sum X_{Ci}} \quad (9)$$

$$d= (Z_A +Z_B +Z_C)/3 \quad (10)$$

$$v_A =(Z_A - d)^2 \quad (11)$$

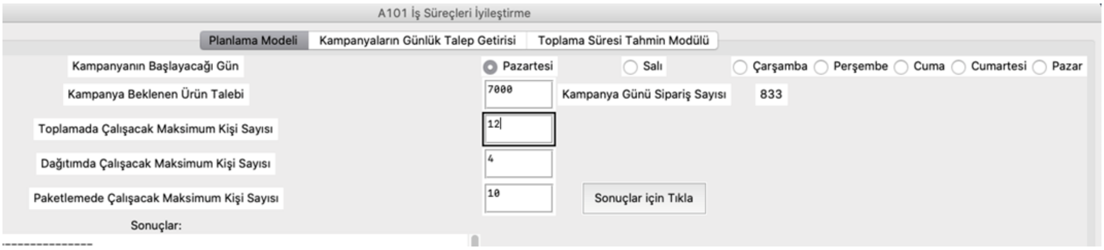
$$v_B= (Z_B - d)^2 \quad (12)$$

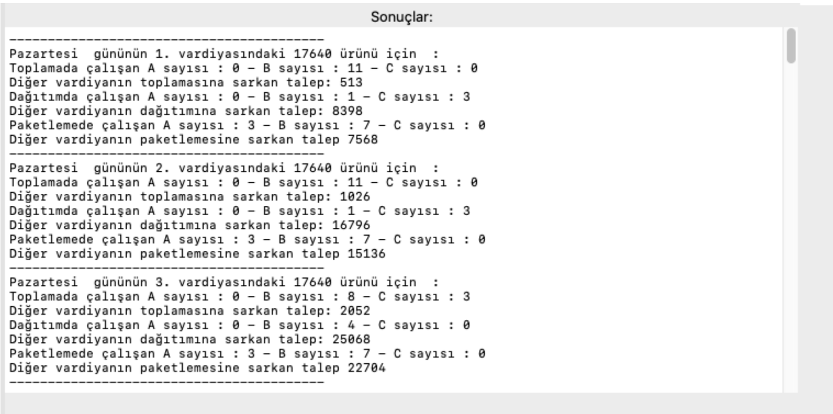
$$v_C= (Z_C - d)^2 \quad (13)$$

X_{ji} integer (14)

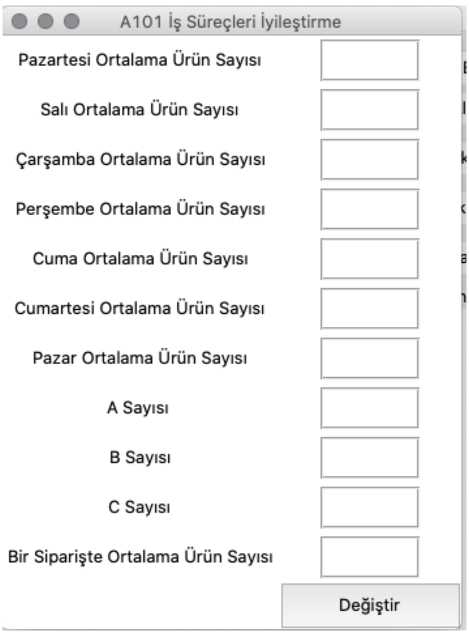
Appendix C

First Section

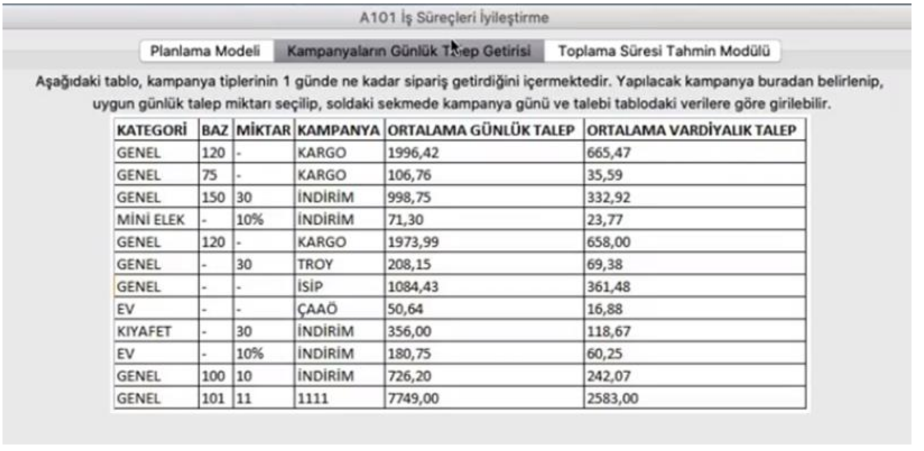




Parameters section;



Second Section



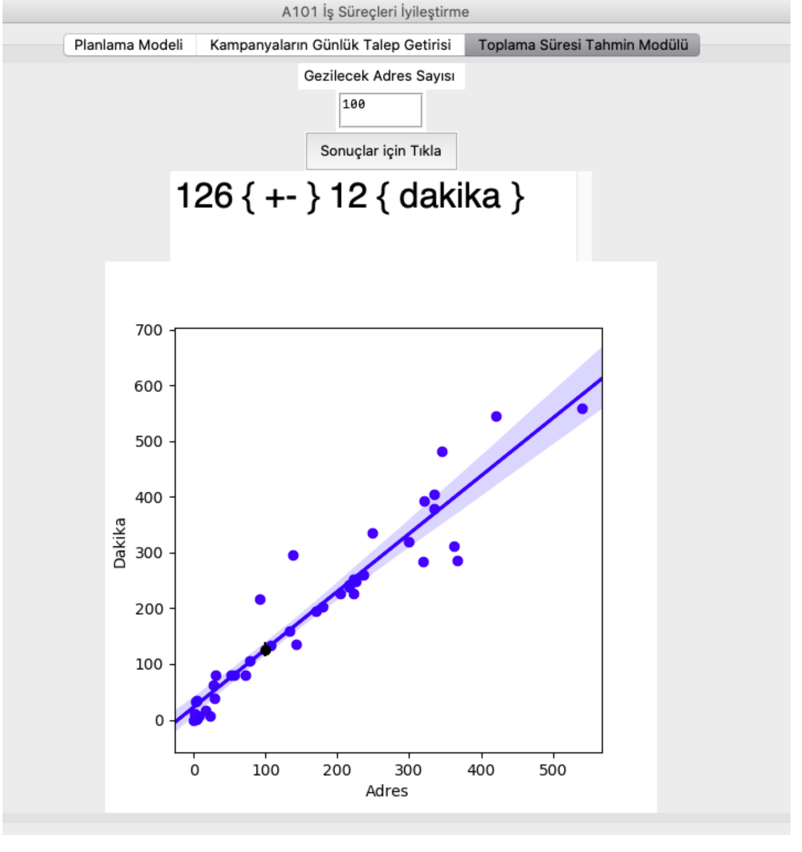
A101 İş Süreçleri İyileştirme

Planlama Modeli | Kampanyaların Günlük Talep Getirisi | Toplama Süresi Tahmin Modülü

Aşağıdaki tablo, kampanya tiplerinin 1 günde ne kadar sipariş getirdiğini içermektedir. Yapılacak kampanya buradan belirlenip, uygun günlük talep miktarı seçilip, soldaki sekmede kampanya günü ve talebi tablodaki verilere göre girilebilir.

KATEGORİ	BAZ	MİKTAR	KAMPANYA	ORTALAMA GÜNLÜK TALEP	ORTALAMA VARDİYALIK TALEP
GENEL	120	-	KARGO	1996,42	665,47
GENEL	75	-	KARGO	106,76	35,59
GENEL	150	30	İNDİRİM	998,75	332,92
MİNİ ELEK	-	10%	İNDİRİM	71,30	23,77
GENEL	120	-	KARGO	1973,99	658,00
GENEL	-	30	TROY	208,15	69,38
GENEL	-	-	İSİP	1084,43	361,48
EV	-	-	ÇAAÖ	50,64	16,88
KIYAFET	-	30	İNDİRİM	356,00	118,67
EV	-	10%	İNDİRİM	180,75	60,25
GENEL	100	10	İNDİRİM	726,20	242,07
GENEL	101	11	1111	7749,00	2583,00

Third Section



Operasyon Merkezi Proaktif İşgücü Dağıtım Sistemi

Akbank

Proje Ekibi

Çağla Aközlü, Halil İbrahim Kavak, Mert Kır, Taha Can Mercan, Can Özkan, İdil Yaşar

Şirket Danışmanı
Zeynep Burcu Tarhan Alpul
Ödeme & Kredi Operasyonları Müdürü

Akademik Danışman
Prof. Dr. Oya Ekin Karaşan
Endüstri Mühendisliği Bölümü

ÖZET

Akbank Operasyon Merkezi'nin kullandığı Operaktif Sistemi, şubelerden ve müşteri hizmetleri servisinden gelen işlemleri etkin çalışan tahsisi ile ayarlamada hedefe ulaşmak için yoğun efor sarf edilmektedir. Önceden karar verilen SLA adı altındaki "işin söz verilen yapılma süresi" limitleri aşılmakta ve günün genel akışı olumsuz yönde etkilenmektedir. İşlemlerin tamamlanma sürelerini dengelemek için bir kullanıcı tahsis modeli geliştirilmiş ve gelebilecek işlem sayısı ve çalışan sayısı göz önünde bulundurularak optimize atamalar yapılmıştır. Geçmiş yıllardan belli günlerdeki veriler analizlenmiş ve gerçekleşmiş atamalar model sonuçları ile karşılaştırılmıştır. Kullanıcı tahsis modeli ile güncel gelen verilerin bir arayüzde birleştirilmesi planlanmış ve kullanıcıya karar destek sistemi oluşturulmuştur.

Anahtar Kelimeler: Karar Destek Sistemi, SLA Dengesi, Proaktif Sistem, Atama

System Design to Proactively Allocate the Workforce of Operations Center

1. Company Information

Akbank was founded as a privately owned commercial bank in Adana on January 30, 1948. Established originally with the core objective of providing funding to local cotton growers, the Bank opened its first branch in Istanbul's Sirkeci district on July 14, 1950. In 1954, after relocating its headquarters to Istanbul, the Bank rapidly expanded its branch network and automated all its banking operations by 1963.

Initially offered to the public in 1990, Akbank stock began trading in international markets and as an American Depository Receipt (ADR) after its second public offering in 1998.

Akbank has grouped its digital banking services, which it pioneered in Turkey, under Akbank Mobile and Internet, providing solutions to clients' financial needs and delivering service at the most convenient points with an excellent customer experience. In a world of fast advancing technology and increasingly demanding customers, Akbank strives to satisfy client needs without time or physical location limitations while pioneering technological innovations both in the sector and in Turkey.

Akbank Banking Center was inaugurated in 2010. Equipped with state-of-the-art technology, the Center is continuously boosting Akbank's productivity and service quality. In 2017, broke ground for Akbank Data and Life Center and continued its investments throughout 2018. The Center became operational in 2019. The complex comprises of Akbank Data Center, which constitutes the core of Akbank's entire technology infrastructure, and Akbank Life Center, which offers social services to 3,000 Akbank employees. The Center is Akbank's largest one-off investment to date.

With a strong and extensive domestic distribution network of 758 branches staffed by more than 12 thousand employees, Akbank operates from its headquarters in Istanbul and 21 regional directorates across Turkey. In addition to providing services at branch locations, its traditional delivery channel, Akbank also serves via the Akbank Internet Branches, more than 5,100 ATMs, Akbank Mobile, the Call Center, and more than 610 thousand POS terminals as well as other high-tech channels.

2. System Analysis and Problem Definition

2.1. Problem Definition and Scope of the Project

In the current framework of Akbank, after receiving the transactions from customers via its branches or Customer Service Operations (MHS) if these transactions need further operation they are sent to Akbank's currently used system, Operaktif. Different types of transactions have specific planned completion times, which are decided by Akbank. These completion times are called Service Level Agreements (SLA) and are considered while assigning employees to jobs. If the jobs can not be completed within the determined SLA times, it also affects other transactions and causes delays in the system. In addition, the transactions are queued according to the due dates in Operaktif's "flexible assignment pool".

When the arrival transaction density increases at a certain amount, the managers intervene in the system to handle it. The managers change the job-employee type assignment constraints of the system called "diploma prioritization" in Akbank's parlance. They intervene with the set of skills of employees such as decreasing the number of capable jobs of a worker for a certain period. However, such manual changes, diploma prioritizations, do not have any technical basis, they are completely run by the experiences of the managers. Therefore, it can cause problems in the following periods of the day. Furthermore, Akbank sometimes uses a button called "panic button" which allows to prioritize diplomas of multiple employees at the same time with one action. When the panic button is pressed, in a moment, employees' diploma types are reduced to one diploma and it is used for EFT jobs usually.

Since Operaktif cannot efficiently track the completed and waiting jobs in the system, panic button usage and prioritization decisions of workers according to their skill sets can be delayed. The purpose of the project for each transaction group is to develop an automated employee assignment system for each time interval that takes into account employees' skills and balances SLA exceedances. For each predetermined period in a day, by using received data from Akbank, the number of jobs arrived in a period were analysed and the number of available employees were also determined for each period.

After the data analysis part, a mathematical allocation model was created as a major model of the study. The allocation model takes forecasted demand and unfinished jobs from the previous period as input, mainly. According to the inputs, employees whose diploma will be prioritized are determined for each period, separately.

2.2. Analysis of the System and Data Interpretation

For the data analysis, the maximum number of transactions arrived at OM that occurs on the first, 15th and last day of each month were analysed after interviewing with Akbank managers. Additionally, to observe the volume of transactions in the first workday of the new year, data for the 2nd of January 2019, was also requested from the company to compare with other days. In the requested data; how much time has passed after the transaction arrived, its processing time and which employee completed the job were examined. In addition, the number of diploma (skill set of an employee) types and arrived job types can be observed as well.

On the other and, to develop and run the mathematical models, a forecast to find the number of each job type that can arrive per period of each day was made. The aim of the forecasting is to have an idea on frequency of incoming jobs in a certain period of time. According to the data, it is observed that 41,50% of the jobs come in between 09:00- 12:00 and 26,22% of total jobs come in between 14:00-16:00. The reason behind this might be the habits of the business world. Because of the dramatic increase in the number of cumulative pending jobs in the operation center system as shown in Appendix A, the focus was worker allocation periodically to reduce overtime and the number of users that have their diploma sets modified at any period during the day. Additionally, keeping the "SLA-meeting" percentages at an equal level across job types which have assigned SLA duration were aimed.

3.System Design

3.1. Proposed Approach

Two main areas constitute the base of the system that will serve Akbank. These can be found in literature as allocation and scheduling problems. Allocation is determined as the basis of the project since it is more useful to assign job types to employees within the workday in several critical time zones. Scheduling was discussed to determine an upper bound for the room for improvement and based on what would happen if we would know a specific day's transactions and their arrival times. Since the assumptions in the article serve the same purposes as our model "Parallel machine scheduling with additional resources: Notation, classification, models and solution methods" by Edis et al. was taken as the primary source for the scheduling discussion (2013).

The idea of including both scheduling and allocation models comes from the article "An exact algorithm for a workforce allocation problem with application to an analysis of cross-training policies" by Michael J. Brusco (2008). The hierarchy of workforce planning in this paper is presented as (i) planning, (ii) scheduling, which is assignment of employees to daily shift patterns and (iii) allocation, which is within and between shift adjustments of labor across departments to more effectively meet changing demand. While the scheduling enables better analysis, allocation gives comprehensive results and is the main model that is utilized.

3.2. Mathematical Model

Since the goal of the project is to minimize the overtime by considering the balance in exceeding SLA times for each transaction group, it is determined to use an allocation model. The employees with the right diploma type or in other words with the suitable skill set can be assigned to the right transaction group by using the allocation model. Therefore, it is logical to use the allocation model because this way it can be easier to aggregate jobs that belong to the same diploma types and distinguish whether they have SLA restriction or not. The model was coded in ILOG and approved with past data.

3.2.1. Allocation Model

The model is going to be used for creating a proactive system to keep the SLA meeting percentage at an equal level by assigning the employees with right skill sets and prioritizing diplomas of employees. To be able to do this an allocation model is created which is resolved at the beginning of each period in a day by taking information about available worker set, capability matrices (diploma type matrices) for major and minor skills of available employees, number of forecasted jobs for the current and upcoming periods, average processing and waiting times for each job type compared to determined SLAs and the period length as inputs. The output of the system will include assignments of determined employees to specific diploma types within the next period as well as the rest of the day. So the main idea is to aggregate jobs according to their diploma types and also under these diplomas they are grouped according to their SLA status. After the aggregation process a forecast based on the same day's previous data is made to see how many jobs arrive at a certain period on a certain day for instance to see how many jobs arrived under Krediler-2 with an SLA of 20 minutes on the second period of the following Monday. Based on this data allocation model that is created to have the minimum number of people prioritized and minimum number of remaining jobs at the end of each period by considering the SLA balance will give which employees to be assigned to a certain diploma type. Details of the model like parameters, such as the ones that ensure SLA balance and constraints regarding capabilities of employees, processing times and leftover jobs can be found at Appendix B. While running the model the following assumptions were made; system was considered to be discrete instead of a continuous one, average processing times were calculated via being weighted on skill set basis, jobs that are not completed at the end of the period will continue to be processed by the same employee in the next period and workdays are assumed to start at 09:00 and finish at 20:00. After running the model based on these assumptions it is found out that it takes 5 minutes to get a result on ILOG and the total average rate of 32% improvement in leftover jobs was derived from the model.

3.3 Validation of the Results

The old data from the company have been used to validate the results. These data belong to 3 different days which are Fridays of the three different months. The reason behind this is, when data analyzed it has been found out that Fridays are the busiest ones in terms of transactions except for special days like bairams. Three different months have been selected to be able to make better analysis and validate results. The first result belongs to April 12th of 2019 and according to the analysis, the average workload for the end of each period is 230 minutes but with the allocation model that has been created this amount is reduced to 178 minutes. Also the number of diplomas prioritized is reduced to 81 from 322. As a result, the improvement rate calculated at 22% compared to Akbank's real data. Comparisons of the real data versus the one with the allocation model can be found at Appendix C. Similar comparisons were made for June 21st of 2019 and May 10th of 2019 and the improvement rates were 56.36% and 21.8% respectively compared to real data. Thus, based on the previous data average improvement rate of 32% derived from the model and improvements were ensured via analyzing performance of employees and reducing period durations.

4. Success of the Developed System

4.1 Implementation Stages

Because of the high confidentiality of the data, it is decided to implement the model on a system called secure ftp which is used to transfer data in the most secure way. The main reason for this is most of the inputs are obtained from Akbank's current Operaktif System like data of the remaining jobs from the previous period. And because there is crucial information about the customers, there is no room for security threads. To solve this problem inputs such as available worker set, capability matrices (diploma type matrices) for major and minor skills of available employees in the flexible assignment pool transferred as an Excel file through secure ftp. After getting the inputs allocation model was run on IBM ILOG CPLEX and outputs of the model transferred to an Excel file. All of the outputs regarding which employee is going to be assigned to which transaction group was stored in that file with a user friendly design of the spreadsheet to deliver the output to the user in the most readable way. Delivery of this Excel file which includes outputs is also transferred to the user through secure ftp as well. This process of transferring both inputs and outputs repeated for each period.

The flow diagram about how the system works is shown in Appendix D. The diagram works at the beginning of each period and provides results. The system gets the real daily data from Operaktif and converts into parameters in Excel. The parameters are entered to the proposed system and become input in the model encoded in ILOG. The result is analyzed by the algorithm and sent to the user interface as a diploma prioritization suggestion to the managers. The final decisions are made by the manager and transmits the decisions to Operaktif.

4.2 Contributions to the Company

Since the managers intervene in the allocation of employees by prioritizing the set of skills during peak hours, the implementation is handled manually. With the integrated system, suggestions were given to managers about the assignment of employees to the job types by considering the forecasts based on past data. During the peak hours, there is no need for manual intervention since the precautions occur on the screen at the beginning of each period before these hours.

With the proposed model, SLA balances were preserved and the transaction completion times did not exceed enough to affect other jobs. For instance, for April 12th of 2019 the number of transaction that exceeds SLA times decreased to 2 from 8 and difference between SLA exceedance times of diplomas with and without the model can be seen in Appendix E. In addition, the overtime of the workers that affects the company in many aspects decreased slightly. With piling up less jobs at the end of each period the workload will be kept on a desirable level and even if employees have to work overtime they will face less amount of workload which can lead to an increase in employee satisfaction. As a result, the improvement rate for the leftover jobs for April 12th of 2019 was calculated at 22% compared to Akbank's real data. In addition, the total average rate of 32% improvement in remaining jobs for the analyzed three days was derived from the model.

4.3 Results

If the proposed decision making tool based on the allocation model is used, the decision process would become more time efficient and will give better results in terms of reducing the number of remaining jobs and SLA balance. This is because the proposed system considers the whole day while giving recommendations to assign workers. Previously these decisions were made empirically according to the amount of jobs coming through and decisions were not based on mathematical models or analyzing data but based on the course of events and experience. Decision process of employee assignment to grouped jobs is expected to be automated with the help of the proposed system however it would be preferable to have a manager which gives the last call about the decisions because the manager knows the working environment better such as the performance differences of employees on certain job types. As a possible improvement a performance coefficient matrix for employees can be created based on performance measures made by the company at the end of each month. This way the model can give recommendations about assigning specifically which employee to a job group and the model will become a more precise one.

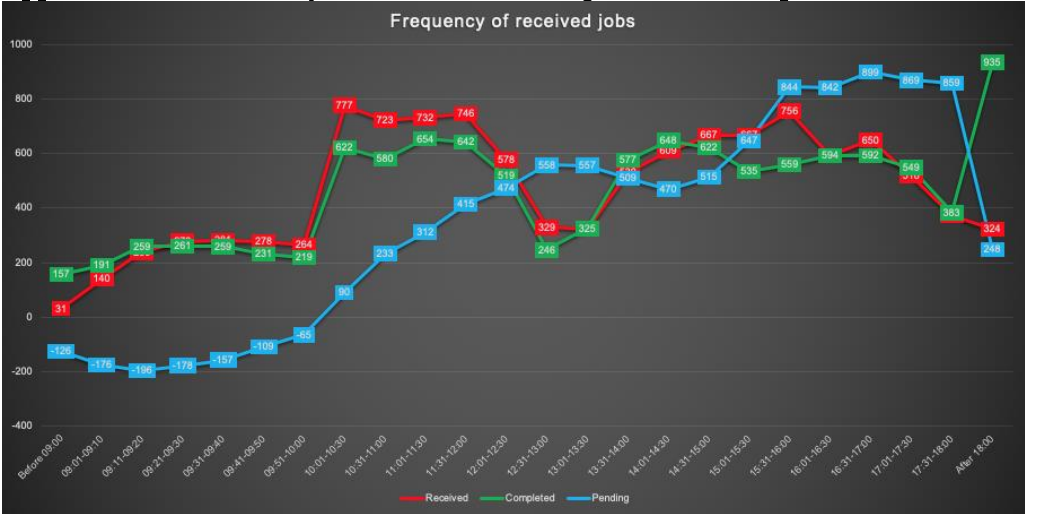
REFERENCES

[1] E. B. Edis, C. Oguz, I. Ozkarahan. "Parallel machine scheduling with additional resources: Notation, classification, models and solution methods". *European Journal of Operational Research*, vol. 230, no. 3, pp. 449-463, 2013.

[2] M. J. Brusco, "An exact algorithm for a workforce allocation problem with application to an analysis of cross-training policies," *IIE Transactions*, vol. 40, no. 5, pp. 495–508, 2008.

APPENDIX

Appendix A: Number of Cumulative Pending Jobs in the Operation Center



Appendix B: Allocation Model

Parameters:

$$M_{ij} = \begin{cases} 1 \text{ if employee } i \text{ has major skill for job type } j \text{ (has major diploma for job type } j) \\ 0 \text{ otherwise} \end{cases}$$

$$S_{ij} = \begin{cases} 1 \text{ if employee } i \text{ has minor skill for job type } j \text{ (has minor diploma for job type } j) \\ 0 \text{ otherwise} \end{cases}$$

p_j : processing time of job type j

d_{jt} : number of jobs received in job type j in the beginning of period t

dur_t : duration of period t (e.g. 30 minutes, 1 hour)

$perf_t$: performance constant in period t

Decision variable:

x_{ijt} : if worker i is assigned to work on job with a major skill type j in period t

y_{ijt} : if worker i is assigned to work on job with a minor skill type j in period t

r_{jt} : remaining amount of jobs or job type j at the end of period t

cp_{jt} : amount of jobs of job type j completed in period t

$MaxSup$: maximum amount of workers assigned to any job type with a minor skill in any period

$$Min \sum_{j=1}^{J} \sum_{t=1}^{T} r_{jt} * p_{jt} * \frac{\frac{dur_1}{SLA_j}}{\sum_a \frac{dur_1}{SLA_a}} + (0{,}00001) * \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{t=1}^{T} y_{ijt}$$

$$x_{ijt} \le M_{ij} * w_{it} \quad \forall i, j, t$$

$$y_{ijt} \le S_{ij} * w_{it} \quad \forall i, j, t$$

$$\sum_j x_{ijt} \le 99999 * \left(1 - \sum_j y_{ijt}\right) \quad \forall i, t$$

$$\sum_j y_{ijt} \le 1 \quad \forall i, t$$

$$k_{ijt} = M_{ij} * \left(\frac{d_{jt} * p_j}{(\sum_a M_{ia} * d_{at} * p_j) + 0{,}00001}\right) * \left(\frac{\frac{dur_t}{SLA_j}}{(\sum_b M_{ib} * \frac{dur_t}{SLA_b}) + 0{,}00001}\right) \quad \forall i, j, t$$

$$l_{ijt} = \left(\frac{k_{ijt}}{(\sum_a k_{ijt}) + 0{,}00001)}\right) \quad \forall i, j, t$$

$$com_{jt} \le \frac{dur_t}{p_j} \sum_i (l_{ijt} * x_{ijt} + y_{ijt}) \quad \forall j, t$$

$$com_{jt} \le d_{jt} + r_{j(t-1)} \quad \forall j, t$$

Appendix C: Comparison of Real Data vs Model on April 12th of 2019

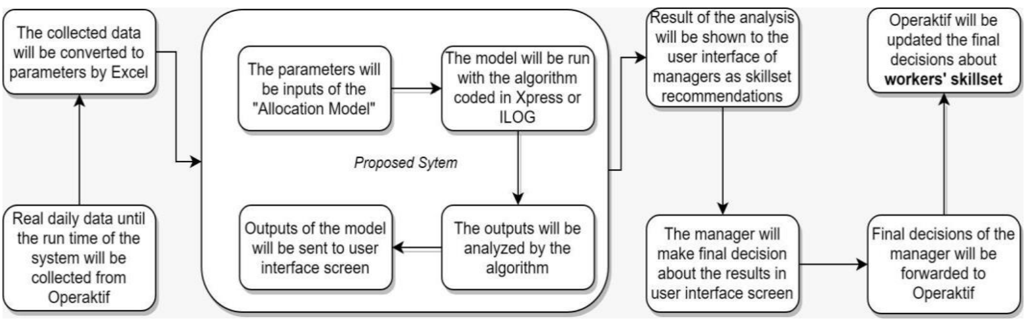
Remaining workload for the end of each period in seconds on April 12nd

	Periods										
Diplomas	1	2	3	4	5	6	7	8	9	10	11
Gayrinakdi A	0	0	0	67	0	335	134	0	201	0	0
Gayrinakdi B	0	938	402	871	670	670	1206	1206	2479	871	0
Gayrinakdi C	0	1673	1195	1673	8126	2629	6931	4302	12667	3346	956
Krediler 1A	0	350	420	420	280	1680	140	1190	280	1400	70
Krediler 1B	819	1701	4221	4032	1323	378	1134	2772	4914	630	189
Krediler 2A	0	0	44	44	0	0	88	0	0	0	0
Krediler 2B	0	0	1235	1482	988	247	1729	1482	988	1235	0
Krediler 3A	0	0	0	0	1404	0	702	0	0	234	0
Krediler 3B	0	2408	6020	3440	8256	4472	15308	11352	7912	5676	344
Krediler 4	0	273	1274	1001	455	182	1183	1001	364	273	273

Remaining workload for the end of each period in seconds in the designed system with ILOG.

	Periods										
Diplomas	1	2	3	4	5	6	7	8	9	10	11
Gayrinakdi A	0	0	0	0	0	8	8	0	0	0	0
Gayrinakdi B	0	0	0	0	0	0	0	0	0	0	0
Gayrinakdi C	0	1298	2877	2652	2498	4320	4127	4425	3932	0	0
Krediler 1A	410	0	0	0	1083	2496	569	0	0	0	0
Krediler 1B	0	0	0	0	0	0	0	0	0	0	0
Krediler 2A	0	0	0	0	0	34	34	34	34	34	34
Krediler 2B	0	1963	3241	2798	2704	2683	2642	2706	2677	2677	2677
Krediler 3A	0	0	0	1315	1315	1968	1968	1968	612	612	612
Krediler 3B	3463	5708	5679	4635	4314	4050	4353	4317	3385	2500	2500
Krediler 4	602	159	0	0	149	1193	2110	262	0	0	0

Appendix D: Flow Diagram About The Proposed System



Appendix E: Comparison of the model and real day's values in terms of SLA limits for April 12th of 2019

SLA output averages for daily in minutes in the designed system with ILOG.

Proposed ILOG System (12.04.2019)		
Diploma	SLA Limits (min)	Exceeding (min)
Gayrinakdi B	20	0
Gayrinakdi C	40	0
İhracat Alış B	15	0
Krediler 1B	20	0
Krediler 2B	20	86,59135779
Krediler 3B	20	0
Ödemeler 1B	20	0
Ödemeler 3B	90	181,4726842
YR Transfer-1	40	0
YR Transfer-2	40	-
YR Transfer-3	40	-

SLA output averages for daily in minutes on April 12th

Akbank's System (12.04.2019)		
Diploma	SLA Limits (min)	Exceeding (min)
Gayrinakdi B	20	14,89539749
Gayrinakdi C	40	1,83908046
İhracat Alış B	15	44,96119017
Krediler 1B	20	19,73584906
Krediler 2B	20	17,37704918
Krediler 3B	20	21,27041742
Ödemeler 1B	20	27,01492537
Ödemeler 3B	90	0
YR Transfer-1	40	10,66898349
YR Transfer-2	40	-
YR Transfer-3	40	-

Taktiksel Üretim Planlama için Karar Destek Sistemi

Arçelik A.Ş. Bulaşık Makinesi İşletmesi

Proje Ekibi

Göksen Akgül, Yaprak Aymergen, Turan Azizov, Rahime Beyza Koçak, Güliz Başak Özkan, Ceylin İzgü Sever, Sevde Sena Zeybek

Şirket Danışmanları

N. Tanzer Tunçalp
Tamer Yıldırım
Ödül Binbaşıoğlu

Akademik Danışman

Prof. Dr. Nesim K. Erkip
Endüstri Mühendisliği Bölümü

ÖZET

Arçelik, mevcut durumda üretim sisteminin kontrolü için SAP'yi kullanmaktadır. Hammaddeler yan endüstrilerden günlük olarak ve farklı partiler halinde tedarik edildiğinden, her ürün tipi için farklı kısıtlamalar ve sınırlamalar ortaya çıkmaktadır. Güncel durumda fabrikadaki montaj hatlarının üretim planı her ay için manuel olarak hazırlanmakta ve bu süreç yaklaşık olarak 2 iş günü sürmektedir. Planlayıcı manuel işlemler sırasında her ürün için tek seferde birden fazla kısıtlamayı dikkate alamayacağı için bazı üretim planlarının hazırlanma sürecinde bir bütün olarak değerlendirilmesi zorlaşmaktadır. Projenin amacı, kullanıcının daha kısa sürede her talep için mevcut olan tüm kısıtlamaları dikkate alan bir üretim planı hazırlamasına olanak tanıyacak, ve hazırlanan planı geliştirmeye yönelik geri bildirimler sunacak bir karar destek sistemi sunmaktır. Karar destek sistemi, kullanıcıların aylık üretim planını verilen kısıtlara göre en uygun biçimde hazırlamasını sağlayacaktır.

Anahtar Kelimeler: Üretim Planlama, Karar Destek, Kısıt, Zaman, Matematiksel Modelleme

Decision Support System for Tactical Production Planning

1. Company Information

Arçelik was founded in 1955 and entered the Turkish white goods sector. Arçelik is specialized in the consumer goods industry and has offered its services and products in more than 145 countries around the globe. It operates with more than 30000 employees and 18 facilities in 7 countries. It has 11 brands including Beko, Blomberg, Grundig, and so on. The company's only dishwasher plant is located in Ankara with more than 1400 employees. This factory supplies 10 percent of the world's annual dishwasher demand.

2. System Analysis

2.1. Current System Information

The amounts of the dishwasher plant's monthly demands vary depending on the seasons. The company divides its planning horizon into two stages during a given year: cold and warm periods. The cold period usually lasts during autumn and winter. The warm period starts around April and continues throughout Summer. Usually, the demand is higher during the warm periods compared to the one in cold periods.

The dishwasher plant operates 6 days per week in 2 shifts a day producing 8000 production units per day during cold periods, and in 3 shifts a day producing 12,000 production units per day during warm periods. The production process involves 3 or 4 assembly lines depending on the demand amounts. The production system of the dishwasher plant relies mainly on Make to Stock (MTS) and Make to Order (MTO) production systems. The production planning department prepares a monthly production plan according to the incoming demand amounts. This approach allows the company to handle the unpredictable and periodical changes such as unevenly distributed monthly demands and machine breakdowns.

2.2. Current System Analysis

The monthly assembly line production plan has been done manually up until now and this process starts at each third week of the month to be prepared for the following month. It takes around two working days. The planning stage involves the contributions of the side industries from which the raw materials have been supplied. In the planning process, it is assumed that the raw materials from the side industries will be available upon the time as it is promised. After completing the tentative monthly plan, the planner goes through the entire schedule to aggregate lots below a certain threshold. However, deviations can occur, and this results in the revisions during the operating assembly line production plan. Everyday revision meetings are organized to address these issues. According to the decisions made in these meetings, the action plan is initiated, and the production plan is reorganized for the given day. Therefore, the finalized plan of the given month consists of many iterations during that month.

3. Problem Definition and Scope

3.1. Indications

During the manual planning process, certain predetermined restrictions or constraints are considered in an order determined by the planner. It is sometimes impossible to schedule the production of a certain demand at a promised date as that demand may be subject to multiple constraints at a time. In a situation like this, some of the demands are back-ordered and cannot be supplied at the given deadline. This is because it is extremely difficult to consider multiple constraints simultaneously in manual operations. The planner achieves a feasible plan at the end of the planning phase; however, multiple backordered products remain present in the plan. Even if a planner achieves a schedule without any backorders, better feasible solutions may exist. Thus, manual planning is unable to uncover alternative solutions since it cannot consider the other possibilities beforehand.

3.2. Problem Definition

The manual production process is one of the main concerns of the plant as it is not efficient and time-wasting. Also, the current planning management is not dynamic since the planner does not have the time to prepare multiple assembly line production plans and compare them. There are multiple constraints to be considered in the planning stage regarding assembly lines, products, customer-related priorities, and so on.

The project scope is determined to address the issues with the help of mathematical models, linear programming (LP), and user interface. This tool aims to develop a better monthly assembly line production plan in terms of feasibility and the backorder amount in a noticeably short time. The system is prepared in a way that every input can be entered dynamically therefore, the decision support system will be able to provide different plans according to the given inputs in a shorter amount of time. This will help the user to compare different scenarios and handle situations with little damage to the original plan. The objective of the system is to provide three assembly line production plans with zero backorders.

4. Planned Approach and Proposed System

4.1. Literature Review

Many academic papers and textbooks have been scanned and analyzed. During the modeling stage, “Operations Research: Applications and Algorithms” (2009) book by Winston has helped us develop the mathematical (Linear Programming) model. Moreover, “Production and Operations Analysis” (Nahmias, 2005) have had an immense contribution in developing the constraints involving inventory, backorder, and material amounts. Another book that was used in model development is "Strategic Decision Making: Multi-objective Decision Analysis with Spreadsheets" (2009) by Kirkwood. Several chapters have been studied from the book and information used in the modeling stage with multiple objectives. An important part of the literature review is Lexicographical Preference Models (Kirkwood, 2009) which includes a set of preferences. Since the system requires many objective functions, one of them should be preferred over the other objective function. For instance, minimizing the total number of backorders has the highest priority as an objective function. The received solution of the first objective function is considered as an input for the second prior objective function, maximizing the outputs which are produced at the preferred assembly line.

We implemented the idea of the lexicographical approach in objective functions as giving a priority order. For instance, minimizing the total number of backorders has the highest priority as an objective function. The received solution of the first objective function is considered as an input of the second prior objective function, maximizing the outputs which are produced at the preferred assembly line.

4.2. Mathematical Model

We have prepared a mathematical model with the given constraints for the problem:

- Assembly lines' capacity limits
- Special product and/or material constraints
- Most preferred and undesired assembly line assignments for each SKU
- The priority of customers and/or demands
- The need for dynamically changing the periods that constraints are affecting the production

From parameters, demand is expected to change in each period for each product, and the upper limits for all the constraints can be changed dynamically. Since we apply the lexicographical preference modeling, our mathematical model contains two models.

The first objective is to minimize the total number of backorders while placing the production on the earliest shift possible. We have given different penalty values of backorder for 15 different priority types specified by the company. Model 1 prioritizes the production of products that have a higher backorder penalty value. For maximizing the production at earlier periods all shifts t are assigned a weight has a value of $0.001*t$. We multiply these weights with the products that are produced in that shift.

The second objective is to minimize the idle capacity for each assembly line. While moving onto minimizing the second objective function, we use the output of the first model as an input for the second model. This input is added to

Model 2 as a constraint which implies that the total number of backorders should be less than or equal to the value that we found in Model 1.

The mathematical model is in Appendix A. The models have *(number of shifts)*(number of demands)*(5)* decision variables.

4.3. Solution Approach

Our main algorithm is based on a Linear Programming model and the dual problem of the LP. The production planner runs the algorithm one week before the production month. Thus, these scenarios could be classified as proactive scenarios. The first step of the algorithm that is proposed to the company includes taking the demand data for next month as an input, solving the LP, returning feasible production schedules for three assembly lines as output. Before running the algorithm, the planner can also enable and disable all the constraints and specify the periods that constraints will be effective.

This assembly line production plan is implemented if there are not any changes that the planner would like to make. The criteria of an acceptable output could be the backorder amount and some other features that the planner specified.

If the outcome is not satisfactory according to the planner, i.e. the company cannot satisfy all the demands until their delivery dates in the given month, the algorithm returns some alternative options based on constraints that affect the system the most to decrease the backorder amount. In addition to the constraints, periods that constraints are effective and the products that use the largest proportion of the limits of the constraints are provided. This evaluation is possible with the help of the dual problem of the model. The system identifies the constraints with the largest dual variable value. Then identified dual variables are clustered if they are of the same type of constraints and effective on consecutive periods. After the constraints and the periods are identified, the system also lists three products that have the largest usage of these constraints. Listed products are the ones that have the largest *Backorder Penalty*Production Amount* values.

The planner may decide to make changes in the input for the next iterations, which are the number of shifts, the number of products being produced per shift, upper limits of the material constraints according to the given information. Then the system will solve the LP for another iteration and will give new feedback on a different plan prepared according to the updated input. This process can be repeated until the output is acceptable and implementable. When the plan is acceptable, the system will solve the mathematical model with an additional objective, which is to assign the products to their most preferred assembly lines.

In addition to the production plan, the program provides a report in which it generates reports of line utilization rates, percentage of demands that have been satisfied, and the total production amount. Also, the system reports the number of days that each demand is delayed and the total number of such backorders and the percentage of material constraints' usage per shift/day in separate Excel sheets. Additionally, it is possible to compare the results of the schedules after several iterations. The compare segment of the system gives the difference between line utilization rates and production amounts of two different plans.

Under the scenarios in which there is a disruption to the current system such as additional or canceled demands, quality control issues, delays in material delivery, or machine breakdowns, a need for revision occurs. Those scenarios could be classified as reactive scenarios. The user can select the days following the disruption for planning purposes, alter the input data according to the period they want to plan, and change the constraints accordingly. The system will give the same outputs as the previous scenario for the user's consideration, with the difference of time frame. A flowchart of the algorithm is given in Figure 1.

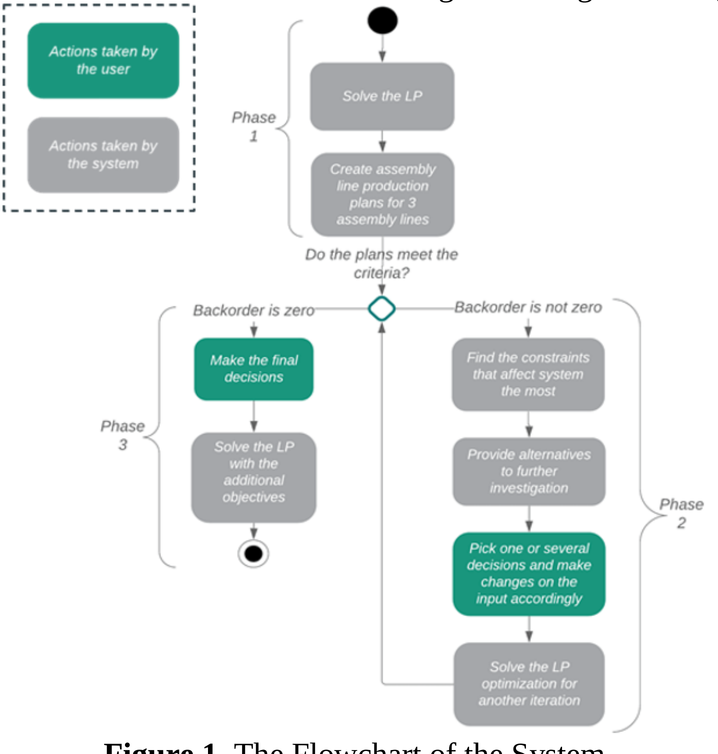


Figure 1. The Flowchart of the System

4.4. User Interface

To ease the use of our algorithm in the current system, a user-friendly user interface with multiple tabs has been developed. In the first tab, the user can select the planning year and month, add, or remove the workdays from the month, select the demand data from their computer and specify whether they want to use the lot aggregation option. In the second tab, the user can adjust the assembly line capacities for any given day. They can enter multiple *production rates per shift* and the *number of shifts per day* values for each of the assembly lines. In the third tab, the user can add or remove the constraints in the system. By double-clicking on the constraints, they can adjust the upper and lower limits of them. Apart from editing, the user can also set the dates for a given constraint to be active only on the specified dates. Then, the user runs the system by clicking on the specified button. The output is stored as an excel file and the constraints which can be further investigated according to the solution and to the will of the user are stored in the fourth tab of the user interface. The user can perform multiple iterations with the system and compare each output in the fifth tab of the user interface. The design of the user interface can be found in Appendix B.

4.5. Performance Measures and Deliverables

To validate the system, several Key Performance Indicators (KPI) are used:

- Amount of backorder reduction at the end of each iteration.
- Amount of demand postponed and backorder ratio.
- Comparison of backorder minimization between the first and the last iterations.
- Priority weighted backorder amount (It is desired to keep as little backorder as possible for the 1st Priority products. Therefore, the percentage of backorder according to the priority types is valuable information for the user.)
- The runtime of each iteration and the total runtime of the completion of all the iterations.

These KPIs are for the in-system validation process (self-validation) which will measure the performance and consistency of our algorithm. Additionally, to compare the system's output with the plan that has been prepared by the Production Planning Engineer manually, we have written another Python code that operates on the assembly line production plan of the Arçelik. Code reports:

- Utilization rate per day
- Amount of backorder and the number of days that products have been delayed
- The difference of the production amounts per SKU, between Tactical Decision Support System's output and Arçelik Engineer's manually prepared plan

5. Validation

With the help of the performance measures, we can compare the system's output with the company's manually done schedules that have been prepared in the previous months. Comparing the two plans with the same performance indicators will allow us to see the differences in results explicitly.

Test runs and iterations have been conducted according to the largest dual variables in LP solution for real data of November 2019 and March 2020. The first iteration is excluding product constraints 1 by 1. Results have shown that the most notable improvement on the objective value has been reported in the case where the “Beko Üst Table (Hot Melt)” constraint is excluded for both demand data sets. Since this constraint had the largest dual variable value in LP, it verified the effectiveness of dual values. Another iteration is to add 1 extra week prior to the month. Results have shown that adding extra shifts has greatly decreased the objective values of the plans.

In addition, we have modified the real-life demand data and increased the total demand amount to a number that will exceed the total capacity of the factory regardless of the workdays and number of shifts. After examining the percentages of backorder weighted by the priorities of the demands, the program indeed decides to produce the products that have the biggest priority and delays the other ones.

We have also compared the production planner’s production plan with the system’s output production plan. We had two data files for this stage for November 2019: the initial demand list and the actual production data. Compared to the initial demand list, the actual production data is finalized, and it includes additional demands, cancellations, and other disruptions. However, each individual change and its effects on the system are not available. Therefore, to make the comparison mentioned, data that has a greater amount of production is taken into consideration, and adjustments on the initial demand list are made accordingly.

Initially, we have derived the percentages of total delayed demand amounts that could not be satisfied until the customers’ desired due date and the total number of demand requests that were delayed. Delay time was considered in terms of days. These calculations were made for our LP’s optimization, the lot aggregated version of our plan, and the production planner’s actual plan. To validate our system's prioritization of high-level priority products, we have also considered the priority state of the demands delayed and calculated each percentage for 15 different priority types.

As for the comparison between our lot-aggregated plan and not-aggregated plan, we have observed that the percentages of total delayed demand amounts which cannot be satisfied with the customers’ desired due date. In consideration of our assumption of assigning the priority and the due dates of the aggregated demands based on the demand that has the highest priority level and the earliest due date, the number of high prioritized demands increased in the lot-aggregated plan. The increase in the percentages of delayed demands for the lot-aggregated plan is an expected result.

In the finalized plan, we had the production amounts in batches, and we have divided these batches into demands. In Appendix C, the percentage of the delays of the higher priority level products in the finalized plan is larger than our system output's results with not lot-aggregated demand data and lot-aggregated demand data. Also, the backorder percentage of the demands which are included in priority 2 and priority 3 in the finalized plan, is less than the lot-aggregated and non-lot-aggregated plan. Besides, the percentage of delayed incidents with higher priority levels in the finalized plan is higher than our system output's results for both non-aggregated and aggregated demand data which can be seen in Appendix D. The delay percentages of priority 2 which is not considered as a high priority in the finalized plan is lower than our plan. Therefore, the system has successfully reduced the backorder amounts and the number of delayed incidents of products with higher priority levels. The improvement percentages for demands with a high priority are higher with a positive value and negative values have been observed for the lower priority demands, which represent that the production planner's finalized plan has done better in comparison to ours, this is an expected result since these demands' productions are not prioritized in our system. The improvement percentages can be seen in Appendix E.

In March 2020, the production rate of assembly lines has been changed in consequence of the pandemic. Line rates and production amounts have been decreased accordingly, especially in weeks three and four. Therefore, line rates have been adjusted to make a fair comparison. March 2020 validation consists of the same steps such as calculation methods of November 2019 validation had. Also, the results of the comparison between the planner's production plan for March data and our production plan is quite like comparison results for November 2019 plans. The production percentage of the high priority demands have been increased in our output plan.

6. Implementation and Integration to the Company

Due to the current events, we are unable to meet with the company representatives face to face. Therefore, the decision support system has been introduced to the production planning department representatives in long-duration sessions. In these sessions, we have run the algorithm with the inputs provided by the planner. This process allowed the planner to observe the performance of the tool and understand the steps. After the algorithm finished, the output was analyzed by the planner and valuable feedback was given accordingly.

The tool will be delivered to Arçelik as soon as both sides agree on the additional features of the product. We plan on delivering it as a standalone program in .exe file format so that the user can access the product without downloading an external module. There is a handbook with the instructions necessary to run the algorithm. This will allow the other users to understand the mechanism and run the algorithm easily.

7. Benefits of the Project

The short-term outcome of this project is the minimization of the time spent on preparing a monthly production plan. Our system provides assembly line production plans within 15 to 30 minutes depending on the number of monthly demands given as input. This plan also considers all the production and capacity constraints the system has simultaneously, so the user does not have to worry about overlooking a critical constraint.

The program also addresses the fact that the user was not able to see the most challenging system constraints and demand entries while planning and taking action accordingly. With the help of the mathematical model and the dual problem of the model, these constraints are easily identifiable, and the user interface creates reports that are easy for the user to analyze and understand.

For the long-term, since the monthly chart is the result of an optimization program, we anticipate that there will be fewer revisions in the monthly production plans. Furthermore, we believe that production planning engineers can identify potential disruptions in production and take an active role in changing them with the help of our program. Additionally, since the system provides the constraints that have the biggest effect on the backorders, if a certain constraint is reported repeatedly, the user can identify the production system's bottleneck after several months. Likewise, if the same periods are given by the system in the constraint reports, this means the demands are accumulated on the same days/weeks each month. Such information can be utilized and used on strategic level decisions of the production facility.

REFERENCES

Arçelik, "Arçelik A.Ş. Hakkında: Kurumsal Tanıtım". Accessed: 01 May 2020. http://www.arcelikas.com/sayfa/72/Arcelik_Kurumsal_Tanitim

Kirkwood, C. W. 1996. Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets. Duxbury Press, Belmont, CA, 1997, ISBN 0-534-51692-0.

Nahmias, S. 2005. Production and operations analysis. Boston: McGraw-Hill, Irwin.

Winston, W. L. 2009. Operations research: applications and algorithms. Belmont, CA: Brooks/Cole, Cengage Learning.

APPENDIX

Appendix A: The Mathematical Model

Sets:

I: set of products, i=1,...,I

J: set of materials, j=1,...,J

T: set of shifts, t=1,...,T

T_u: set of shifts specified by the user, t=1,...,T_u

K_s: set of production capacity constraints per shift, ks=1,2,3

K_s^u: subset of production capacity constraints per shift, ks=1,2,3

K_d : set of production capacity constraints per day, kd=1,...,5

K_d^u: subset of production capacity constraints per day, kd=1,...,5

M: set of assembly lines, m=1,2,3

D: set of last shifts of the days, t=3, 6, 9, ..., T

Parameters:

D_{it}: Demand for product i in shift t

A_{ij}: The number of material j required to produce one product i

Cm_j: The maximum amount of material j that can be stocked (raw material inventory capacity)

Cp_i: Desired number of product i that can be stocked in inventory

U_m: production capacity of assembly line m per shift (limit of production capacity of different assembly lines)

c_i: Cost of holding one product i in the inventory for one period

p_i: Penalty point of backordering one product i for one period

$E_{i(kd)}$: 1 if constraint kd is present for product i, 0 otherwise

$E_{i(ks)}$: 1 if constraint ks is present for product i, 0 otherwise

F_{kdt}: The upper limit of constraint kd in period t

F_{kst}: The upper limit of constraint ks in period t

R_t: An array of t numbers in increasing order starting with 0.01*1, 0.01*2, ..., 0.01*t

Decision Variables:

P_{imt}: the number of product i produced in assembly line m by the end of shift t

B_{it}: Number of backorders from product i at the end of shift t

G_{jt}: the number of material j purchased in shift t

Im_{jt}: the number of material j in the inventory at the end of shift t

Ip_{it}: the number of product i in the inventory at the end of shift t

$z_1 = \sum_{i,t} B_{it} * p_i + \sum_{i,m,t} P_{imt} * R_t$

$z_2 = \sum_{m,t}(1 - (\sum_i P_{imt})/U_m)$

Model 1:

$$\text{Min } z_1 \tag{1}$$

Subject To

$$\sum_{i\in I, m\in M} A_{ij} * P_{imt} + Im_{jt} = G_{jt} + Im_{j(t-1)} \qquad \forall j \in J \text{ and } \forall t \in T \tag{2}$$

$$\sum_{m\in M} P_{imt} + Ip_{i(t-1)} + B_{it} = D_{it} + Ip_{it} + B_{i(t-1)} \qquad \forall i \in I \text{ and } \forall t \in T \tag{3}$$

$$Im_{jt} \le Cm_j \qquad \forall j \in J \text{ and } \forall t \in T \tag{4}$$

$$\sum_{i\in I} P_{imt} \le U_m \qquad \forall t \in T \text{ and } \forall m \in M \tag{5}$$

$$\sum_{i\in I, m\in M} E_{i(ks)} * P_{imt} \le F_{kst} \qquad \forall\ ks \in K_s^u \text{ and } \forall t \in T_u \tag{6}$$

$$\sum_{i\in I, m\in M} E_{i(ks)} * P_{imt} \ge FL_{ks} \qquad \forall t \in T_u,\ ks=3 \tag{7}$$

$$\sum_{i\in I, m\in M} (E_{i(kd)} * P_{im(t-1)} + E_{i(kd)} * P_{im(t-2)} + E_{i(kd)} * P_{imt}) \le F_{kdt} \qquad \forall\ kd \in K_d^u \text{ and } \forall t \in D \tag{8}$$

$$P_{imt}, B_{it}, G_{jt}, Im_{jt}, Ip_{it} \ge 0 \qquad \forall j \in J, i \in I, t \in T, m \in M \tag{9}$$

Model 2:

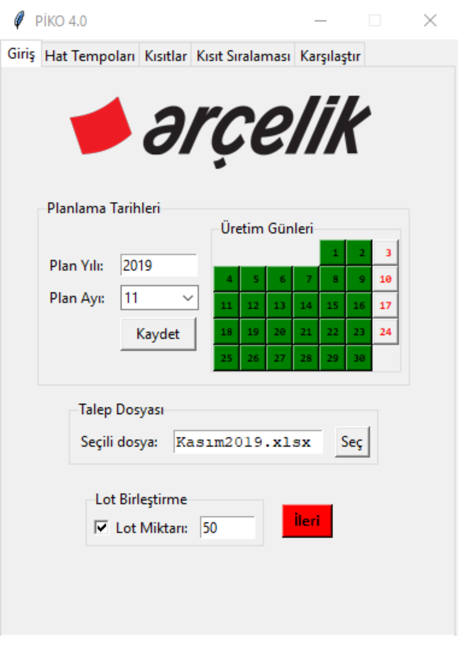
$$\text{Min } z_2 \tag{10}$$

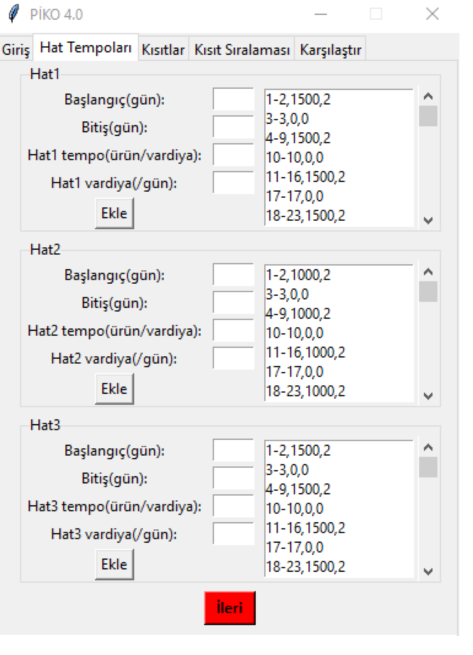
Subject To

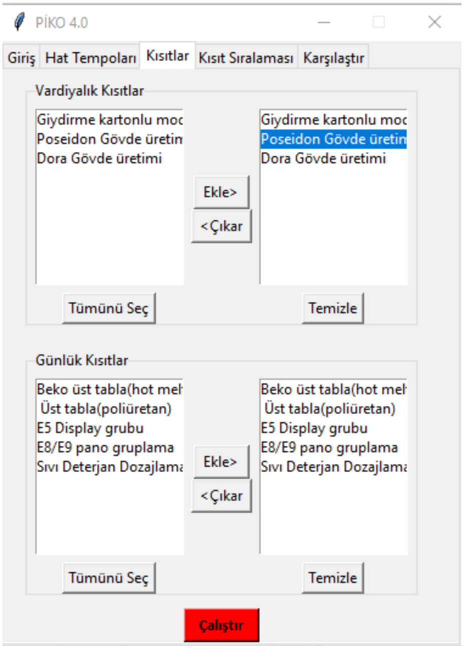
$$\sum_{i,t} B_{it} * p_i + \sum_{i,m,t} P_{imt} * R_t \le z_1^* \qquad \forall i \in I \text{ and } \forall t \in T \tag{11}$$

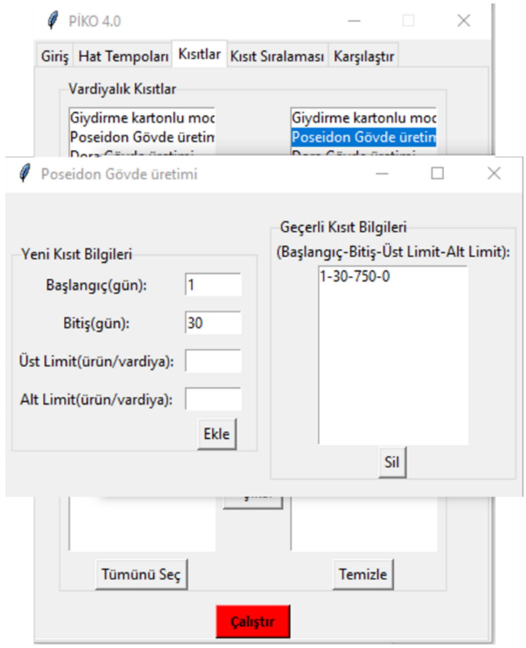
(2) – (9)

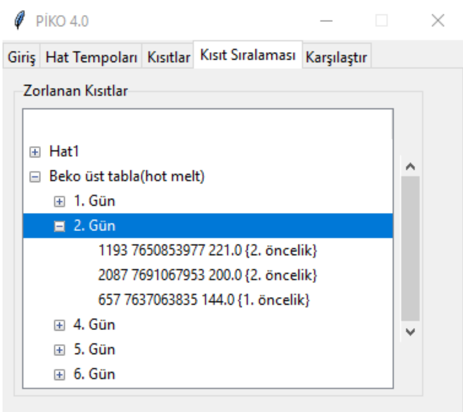
Appendix B: User Interface

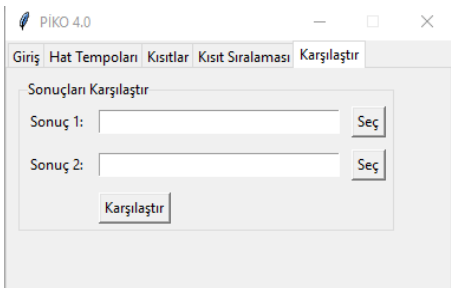





PİKO 4.0
Giriş
Hat Tempoları
Kısıtlar
Kısıt Sıralaması
Karşılaştır
Vardiyalık Kısıtlar
Giydirme kartonlu moc
Poseidon Gövde üretin
Dora Gövde üretimi
Ekle>
<Çıkar
Tümünü Seç
Temizle
Günlük Kısıtlar
Beko üst tabla(hot mel
Üst tabla(poliüretan)
E5 Display grubu
E8/E9 pano gruplama
Sıvı Deterjan Dozajlama
Çalıştır


PİKO 4.0
Giriş
Hat Tempoları
Kısıtlar
Kısıt Sıralaması
Karşılaştır
Vardiyalık Kısıtlar
Giydirme kartonlu moc
Poseidon Gövde üretin
Poseidon Gövde üretimi
Yeni Kısıt Bilgileri
Başlangıç(gün):
1
Bitiş(gün):
30
Üst Limit(ürün/vardiya):
Alt Limit(ürün/vardiya):
Ekle
Geçerli Kısıt Bilgileri
(Başlangıç-Bitiş-Üst Limit-Alt Limit):
1-30-750-0
Sil
Tümünü Seç
Temizle
Çalıştır


PİKO 4.0
Giriş
Hat Tempoları
Kısıtlar
Kısıt Sıralaması
Karşılaştır
Zorlanan Kısıtlar
Hat1
Beko üst tabla(hot melt)
1. Gün
2. Gün
1193 7650853977 221.0 {2. öncelik}
2087 7691067953 200.0 {2. öncelik}
657 7637063835 144.0 {1. öncelik}
4. Gün
5. Gün
6. Gün


PİKO 4.0
Giriş
Hat Tempoları
Kısıtlar
Kısıt Sıralaması
Karşılaştır
Sonuçları Karşılaştır
Sonuç 1:
Seç
Sonuç 2:
Seç
Karşılaştır

Appendix C: Percentages of Demand Amounts Delayed

(Total Number of Demands Delayed/Total Number of Demands)*100					
				Algorithm Output	
Priority	Demand Type	Amount	Finalized Plan	Not Agg. Plan	Agg. Plan
1.öncelik	SLF	40737	60.1	20.2	26.4
1.öncelik	SLA	10580	53.3	12.0	36.4
1.öncelik	SLB	-	-	-	-
1.öncelik	SLC	-	-	-	-
1.öncelik	FCS	2343	39.1	83.8	96.5
2.öncelik	SLF	26477	75.6	67.8	87.0
2.öncelik	SLA	28597	68.3	78.3	84.9
2.öncelik	SLB	50	100.0	100.0	100.0
2.öncelik	SLC	2652	51.0	93.9	100.0
2.öncelik	FCS	1	100.0	0.0	-
3.öncelik	SLF	-	-	-	-
3.öncelik	SLA	48722	71.5	45.4	52.7
3.öncelik	SLB	-	-	-	-
3.öncelik	SLC	36	94.4	100.0	-
3.öncelik	FCS	33104	0.0	51.2	56.7

Appendix D: Percentages of Demand Incidents Delayed

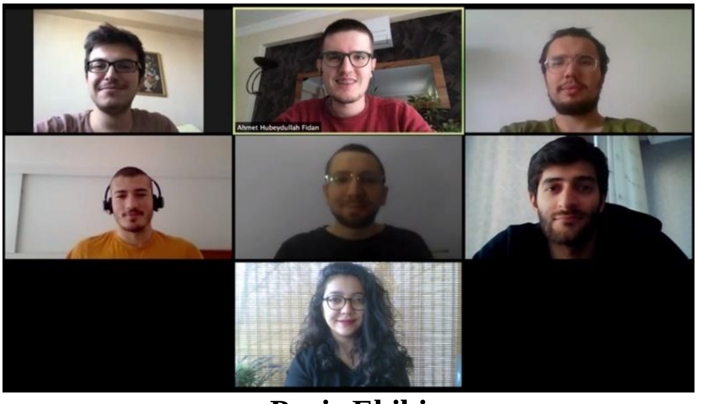
(Total Number of Incidents Delayed/Total Number of Incidents)*100					
				Algorithm Output	
Priority	Demand Type	Amount	Finalized Plan	Not Agg. Plan	Agg. Plan
1.öncelik	SLF	531	13.7	4.0	8.5
1.öncelik	SLA	155	3.8	1.1	2.6
1.öncelik	SLB	-	-	-	-
1.öncelik	SLC	-	-	-	-
1.öncelik	FCS	40	0.8	1.1	1.0
2.öncelik	SLF	647	22.4	18.4	21.2
2.öncelik	SLA	266	9.4	9.2	8.8
2.öncelik	SLB	-	0.0	0.0	0.1
2.öncelik	SLC	12	0.3	0.5	0.7
2.öncelik	FCS	-	-	-	-
3.öncelik	SLF	-	-	-	-
3.öncelik	SLA	448	13.9	8.5	10.7
3.öncelik	SLB	-	-	-	-
3.öncelik	SLC	36	1.4	1.5	0.0
3.öncelik	FCS	237	0.0	4.5	3.8

Appendix E: Improvement Percentages Table

		Improvement of Demand Delayed		Improvement of Number of Incidents	
Priority	Demand Type	Not Agg Plan	Lot Agg Plan	Not Agg Plan	Lot Agg Plan
Priority 1	SLF	40%	34%	10%	5%
Priority 1	SLA	41%	17%	3%	1%
Priority 1	SLB	0%	0%	0%	0%
Priority 1	SLC	0%	0%	0%	0%
Priority 1	FCS	-45%	-57%	0%	0%
Priority 2	SLF	8%	-11%	4%	1%
Priority 2	SLA	-10%	-17%	0%	1%
Priority 2	SLB	0%	0%	0%	0%
Priority 2	SLC	-43%	-49%	0%	0%
Priority 2	FCS	100%	100%	0%	0%
Priority 3	SLF	0%	0%	0%	0%
Priority 3	SLA	26%	19%	5%	3%
Priority 3	SLB	0%	0%	0%	0%
Priority 3	SLC	-6%	94%	0%	1%
Priority 3	FCS	-51%	-57%	-5%	-4%

Depo İyileştirme Çalışması

Arçelik A.Ş. Bulaşık Makinesi İşletmesi



Proje Ekibi

Ebrar Samed Akdağ, Cemil Akgün, Sima Alkan, Ahmet Emin Babutcu, Jafar Bakhshaliyev, Yavuz Batı, Ahmet Hubeydullah Fidan

Şirket Danışmanı	**Akademik Danışman**
Serdar Çevik	Prof. Dr. Nesim K. Erkip
Endüstri Mühendisi Takım Lideri	Endüstri Mühendisliği Bölümü

ÖZET

Bu rapor, yüksek yoğunluklu trafik yaratan malzeme akışları göz önünde bulundurularak, Ankara Arçelik Bulaşık Makinesi İşletmesi deposunun yerleşim sorunu için bir öneri içermektedir. Üretimle ilgili olan ve olmayan çeşitli malzemeler ve bunların etkisiz depolama ve alan tahsisi uzun malzeme taşıma sürelerine yol açar. Proje, verileri ve gözlemleri kullanarak geri bildirim mekanizması ile birbirini besleyen iki aşamalı çözüm modeli üzerinden ilerlemiştir. Bu aşamalarda kullanılan yöntemler; forkliftlerle alınan toplam mesafenin iyileştirilmesi için yardımcı algoritmaların kullanımı ve yenilikçi önerilerle depo sisteminin kullanılan alan ve iş güvenliği odaklı iyileştirilmesidir. Sunulan alternatifler farklı senaryolarda ve farklı akış trafiği seviyelerinde test edilip çalışmanın uygulanabilir olmasına imkan vermiştir. Çalışma, alternatif yerleşim planı ile depodaki trafik ve malzeme taşıma sürelerinde iyileştirme sağlayan bir düzen sunmayı amaçlamaktadır.

Anahtar Kelimeler: İyileştirme, Süreç Analizi, Yerleşim Düzeni, Malzeme Taşıma

Material Warehouse Improvement Study

1. System Description

Arçelik Ankara Dishwasher factory, which is the subject of our project, adopts a production rate according to the fluctuating demand throughout the year. The factory operates 3 shifts during the seasons with high demand, and 2 shifts in low demand season where 4.000 machines can be produced per shift. The cycle time for a dishwasher is 7 seconds. Currently, Arçelik produces 2,5 million dishwashers per year.

The study is conducted for the material warehouse of the factory. The warehouse has 10 specified areas with the total area of 3746 m². The given layout by the company is adjusted with adding the areas of each warehouse location and with the accurate area shapes by the AutoCAD program which can be seen in Figure 1 below,

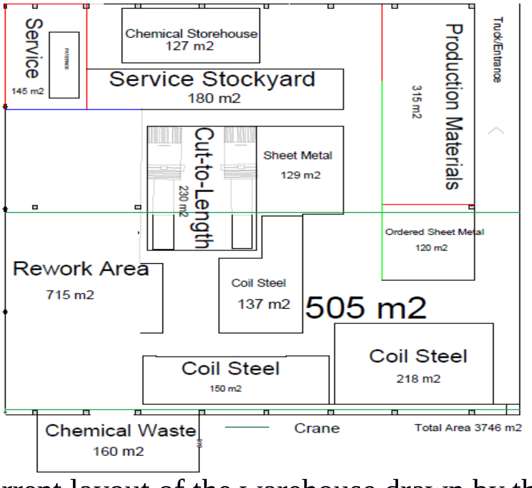


Figure 1. Current layout of the warehouse drawn by the AutoCAD (Çevik, Re: Bilkent - Arçelik Proje)

The interdepartmental flow consists of forklift movements as well as crane movements. The relations among the departments can be seen in Appendix A. Materials stored in the warehouse are from different types. Some materials are directly related to the production rate, while some are affected by the production in an indirect way and some materials are not related to the production. Therefore, we grouped materials according to their relevance to production. It can be seen in Appendix B.

2. Problem Definition

Problems of the warehouse can be classified with three main points as occupational health and safety (OHS) issues, inefficient utilization of the warehouse areas, and time waste due to material handling. These problems threaten the processes in the warehouse since, for future plans, Arçelik aims to increase their production rate to 3 million dishwashers per year.

Storage of materials is not executed by a systematic regulation which causes inefficient utilization of warehouse areas and OHS issues. For instance, cardboards and hinges that should be stored in Production Materials Stockyard are sometimes stored in Sheet Metal Stockyard. Furthermore, some stacking methods constitute serious OHS problems. As an instance, 3-4 dishwashers in Rework Area are put on top of each other without any safety.

Allocations of the warehouse areas with specific job descriptions were not determined by considering efficiency in terms of total forklift movement. Therefore, forklift flows take inefficiently long time which leads to a waste of time. To sum up, the project focuses on OHS issues, the inefficient utilization of the warehouse areas, and the time waste.

During the company meeting, constraints decided as; Crane, Coil Steel Stockyard, and Cut-to-Length Lines are fixed and Cut-to-Length Lines must be close to Coil Steel and Sheet Metal Stockyard. Moreover, Service Office has a fixed underground system and Service Area stockyard has a fixed second floor and required aisle width is 3,3 meters for the forklift and pedestrian paths (Akın).

Furthermore, it is stated that Arçelik does not prefer Service Area to be changed due to the high-priced system (Akın). However, the project considers flexibility on solution methods free from the constraints due to company incentive to see unconstrained solutions along with the constraint implementations (Akın).

3. Objectives and the Scope of Project

The scope of the project is mainly demonstrated under three topics: high traffic of material flow, the underutilized capacity of the warehouse, and time waste due to inefficient allocation of the areas. The objectives of the project are the following: decreasing material handling time, to increasing utilization of warehouse capacity so that warehouse could gain free spaces and providing efficient warehouse layout in order to deal with increased production capacity in the future.

We plan to reach our objectives by adopting a system with two stages that are explained in detail below. As a result, we offer different alternative layouts and area-specific improvements in order to provide flexibility to choose the best possible option for Arçelik.

4. Methodology

Warehouse improvement study is divided into 2 stages for a better focus. These are comprehensive layout improvements and improvements within the areas. These stages work simultaneously and update each other. For the Stage 1 layout suggestions, two different methodologies are used.

4.1. Stage 1: Comprehensive Layout Improvement Study

Yearly forklift flow of the warehouse areas is given in Appendix C which is formed by the data and the observations. Then, flows are ranked from highest to lowest which can be seen in Appendix D. During the allocation decisions, the defined forklift and pedestrian paths in the warehouse are regulated considering the OHS issues. According to the floor lines observed in the production plant and the fact that a forklift width is one meter, forklift paths should have 2.5 meters width and the pedestrian paths should be 0.8 meters long with the outline of yellow lines. Moreover, blind spot mirrors on the corners of the forklift paths must be implemented in order to eliminate a significant OHS problem since there are multiple forklifts operating in the warehouse simultaneously.

4.1.1 CRAFT & Excel Methodology

CRAFT, proposed by Buffa et al. (137), is one of the oldest algorithms in the literature, used in layout problems. CRAFT is an improvement type needs an initial layout, a from-to chart of flows among departments and cost per unit distance traveled. In our case, initial layout is the current layout of the warehouse, we have a from-to chart as stated above. However, we do not use cost value because our objective is to minimize the total distance traveled (TDT) in the warehouse. Excluding the cost factor doesn't change the algorithm in an adverse way. CRAFT Algorithm steps can be listed as follows;

- CRAFT uses center points of departments and rectilinear distances between departments to calculate the TDT.
- It proceeds by checking pairwise exchange of departments that are either next to each other or with the same area.
- After calculating the effect of each possible pairwise exchange, CRAFT chooses the alternative with the most reduction in TDT.
- The algorithm then finds new center points of departments and re-calculate the total distance accurately.
- CRAFT iterates until no further improvement can be made.

We utilize Excel to find the effects of each pairwise exchange. However, we don't consider all possible exchanges since there are some strict limitations. Coil Steel Stockyard and Cut-to-Length Lines must be in the range of crane to be able to operate properly, because crane provides the flow between these two departments. After finding the best possible exchange in Excel, we use online drawing tool ("Draw io") to implement iterations. We propose three layouts following this methodology which can be seen in Figure 5 in Appendix E and in Figure 4. The reason for getting three alternatives is that the assumptions we made before each run of CRAFT. One assumption is that Service Area can be freely relocated. The other one is that Service Office is fixed but shelves can be relocated. The last one is that neither Service Office nor shelves can be moved.

4.1.2. MATLAB & AutoCAD Methodology

The aim is to consider minimum forklift flow therefore, minimum traffic in the warehouse to assign 9 warehouse areas to 9 empty locations to be able to generate new layouts by utilizing and redefining the Norén and Eriksson (2) minimum material handling heuristic. Also, distance and area comparisons are calculated by utilizing the Excel and AutoCAD to see the benefit of changing the current layout with the new layout scenarios. In terms of OHS, defined forklift & pedestrian paths are determined.

After the data analysis, MATLAB implements the minimum forklift flow allocation by the given forklift flows and calculates the rectilinear distances from Entrance/Truck to assign high flow areas closer. Since this is an unconstrained and initial trial, we made iterations between areas by the given constraints.

Therefore, the assignments are done by the iterations with the AutoCAD to consider area-specific requirements in Stage 2. After the high-density flow and close distance to Entrance/Truck assignment layout generated by the MATLAB, the most beneficial and feasible iterations are implemented in 2 alternative ways which can be seen in Figure 6 in Appendix E.

4.2. Stage 2: Improvement Study within the Warehouse Areas

Stage 2 improvement study provides specific solutions for each area. In terms of OHS, dangerously placed materials are considered. The product flows and the current capacity in each area is analyzed if there is a need for a better storage system that the proposed system in Stage 1 cannot satisfy the efficiency of the warehouse. Since the total production is planned to increase, the main problem for all departments is that the inefficient utilization of areas. If capacity of the warehouse areas becomes inadequate, proposed system provides free spaces to make sure warehouse capacity can meet area needs. After the meetings with Serdar Çevik, Coil Steel Stockyard, and Rework Area additional improvement implementations are validated therefore, Stage 2 improvements focus on these two areas (Çevik).

4.2.1. Coil Steel Stockyard Analysis

Coil Steel Stockyard storage capacity is approximately sufficient for 28 days in the current production rate. During the observations, it is seen that the storage method is the most important problem in this area. The coil steels are damaged due to the placement on the iron fixture. Several damages lead to the disposal of a few wraps of the coil steels. The company faces loss of money due to this damage. Therefore, coil chocks are considered better for the storage since a lot of local and global companies suggest different types of coil chocks and guarantees coil steels would not be damaged at all. The total cost of the chocks is approximately 233.000 TL, although it really depends upon the company ("KLP Lankhorst Coil Storage Systems"). The chocks have an extreme long life ("Roll Blocks"). We need to find the expected time in which the benefit of chocks exceeds the total cost. To find the benefit of the coil chocks, calculations are done by considering the damage cost by calculation with the proportion of the length, flow of the coil steels, and the monthly percentage of the coil steels which are damaged. Benefits start after a month and sensitivity analysis is done for different percentage of increases in the flow of the coil steels. The sensitivity analysis also confirms that the return on investment is maximum 2 years.

4.2.2. Rework Area Analysis

Dishwashers in Rework Area are irregularly stacked on the floor which leads to OHS issues and to inefficiency in terms of storage. Sometimes dishwashers are stored at Chemical Waste area since the capacity is inadequate. Also, they are usually stacked in 3 or 4 layers on top of each other without extra caution which leads to falling hazard. Therefore, a rack-allocation system is proposed for the dishwashers so that the space can be utilized much more efficiently. There are 8 scrap containers which allocates approximately 25 m² (2,5 m x 1,35 m for each container). Since each containers' height is 1,5 meters, there is a flexibility of implementing a rack system above to store dishwashers. A dishwasher occupies 0,423m² space with foam covering it. 2-layered rack system is proposed considering safety issue and forklift ability. In one compartment, there can be stored two dishwashers side-by-side, 2 dishwashers on top of each other, and 3 dishwashers back-to-back. Thus, 12 dishwashers can be stored in a compartment There will be 16 compartments in total, which means we can store up to 192 dishwashers in the rack system. The system costs 27.650₺ ("Ertaş"). Therefore, this solution provides 81m² (0,423*192) gain in terms of space, and also eliminates the safety issues caused by the stacking method. Also, this improvement provides a safer storage system for the dishwashers.

Dishwashers are stored with foam coatings, and foams are thrown away before the dishwasher is reworked. They are stored in a container which is carried outside by a forklift. The foams occupy too much space due to their volumes and container gets full quite fast which leads to a higher forklift traffic. Therefore, we propose a foam grinder machine implementation which will reduce forklift traffic by 10 moves a day. We observe that each move takes 5 minutes. Thus, we can save up to 250 hours of material handling per year. The machine costs around 29.500 TL ("Başkent Freze").

5. Scenario Analysis and Comparison for the Alternative Layouts

Performance metrics that are used to compare current layout with the new suggested layouts are improvement in forklift flow and area measurements. We obtained five different layout. The main reason for this variability is different assumptions during each run of the system. For instance, Arçelik is not fully open to change in Service Area as mentioned above. Thus, we assume that Service Area is fixed in some alternatives. Such assumptions led to different alternatives, which is good since we can offer Arçelik to choose as regards to their preference.

- **Scenario 1: Annual Flow & Center of Mass**
  Considers today's annual forklift flows and coordinates of center of mass of departments as center points. Thus, this scenario evaluates our alternatives if they are implemented in today's conditions.
- **Scenario 2: Peak Flow & Center of Mass**
  Arçelik plans to increase its yearly production rate from 2,5 to 3 million as mentioned in above sections. Hence, we tested our model in terms of peak flows while still using center of mass coordinates of departments.
- **Scenario 3: Peak Flow & Rational Mass of Centers**
  In this scenario, we determined the new center points by taking into account only the parts of the areas that a forklift can enter. Thus, we achieved a more realistic scenario. We also used peak flows for this scenario.
- **Scenario 4: Peak Flow & Stage 2 Implementations**
  This scenario tests whether stage 2 suggestions that we propose are really effective. In addition to that, peak flows and more realistic center points, just like in Scenario 3, are used in this scenario.

Table 1. Comparison table for alternative layouts 1,2 and 3

	LAYOUTS								
	Alternative-1			Alternative-2			Alternative-3		
	Absolute Distance	Improvement (%)	Free Space	Absolute Distance	Improvement (%)	Free Space	Absolute Distance	Improvement (%)	Free Space
Scenario-1: Current	926.496,00	9,56%	196,72 m²	126.376,13	1,30%	196,72 m²	739.132,56	7,63%	318 m²
Scenario-2: Peak Flows	725.016,96	6,78%	196,72 m²	-103.758,28	-0,97%	196,72 m²	520.877,43	4,87%	318 m²
Scenario-3: Rational	1.889.190,00	19,50%	196,72 m²	1.080.449,93	11,15%	196,72 m²	871.924,32	9,00%	318 m²
Scenario-4: Stage2 Imp.	2.165.310,00	22,35%	271,72 m²	1.255.409,93	12,96%	271,72 m²	1.069.420,32	11,04%	393 m²

Table 2. Comparison table for alternative layouts 4 and 5

	LAYOUTS					
	Alternative 4			Alternative 5		
	Absolute Distance	Improvement (%)	Free Space	Absolute Distance	Improvement (%)	Free Space
Scenario 1: Current System	43.596,00	0,80%	302 m²	162.077,15	2,90%	316 m²
Scenario 2: Peak Flows	52.315,20	0,77%	302 m²	194.492,58	2,88%	316 m²
Scenario 3: Rational Centers	522.185,32	9,46%	302 m²	784.603,17	13,42%	316 m²
Scenario 4: Stage 2 Improvement	43.596,00	0,80%	254 m²	162.077,15	2,90%	268 m²

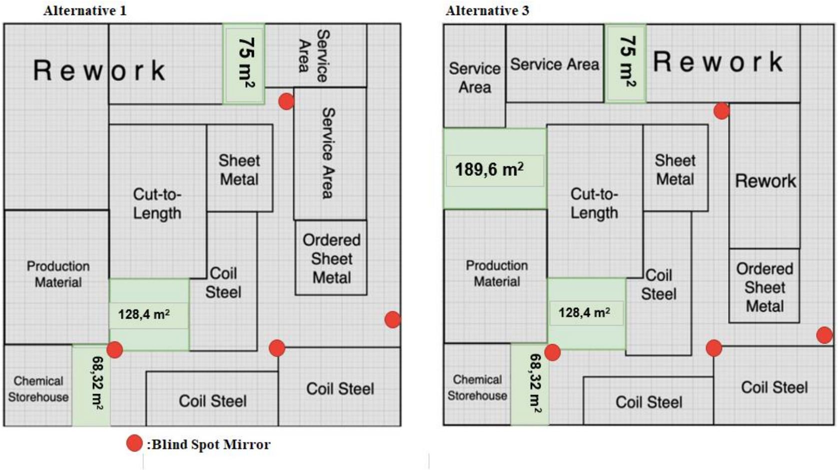


Figure 2. Optimal Layout Suggestions

6. Selection of the Best Layouts and Contribution to the Company

After the meetings with Serdar Çevik, solution proposals of the optimal layouts as Alternative 1 or Alternative 3, blind spot mirror implementations, additional improvement studies in the Coil Steel Stockyard and Rework Area are presented to him. He expressed his satisfaction with the solutions, and stated that they will be considered by the Arçelik (Çevik). Implementation costs in terms of Stage 1 and Stage 2 improvements are also indicated as:

- 4 Blind Spot Mirrors: 2.000 ₺ **(**“İlgi Trafik”)
- Rubber Chocks for Coil Steel Storage: 233.000 ₺(Akkaya)
- Foam Grinder for Rework Area: 29.500 ₺(“BaşkentFreze”)
- Rack System for Rework Area: 27.650 ₺ (“Ertaş”)

Alternative layouts are analyzed in terms of TDT and free space as performance metrics as in the Table 1, Table 2 and, Appendix F. Also, all alternatives adapted required aisle width as 3,3 meters. When we consider all the alternatives, Alternative 1 and 3 gain much more in terms of absolute distance and free space in each scenario. The remaining alternatives are leading to loss or lesser gain in some of the scenarios (alternative 2 for scenario 2 as an instance) in Table 1 and 2. Therefore, the project focuses on Alternatives 1 and 3 for the implementation. Their contribution to the company by the free space gained and less forklift traffic by lesser distances traveled can be found in Table 1. While the most optimal layout can be achieved by the Alternative 1, if the company chooses Alternative 3 to consider Service Area constraint since it is costly to move, they will still gain free spaces and decrease the TDT even though they are less than the results of Alternative 1. With the suggested system, Arçelik can reduce yearly forklift flow by up to 22,35%, and gain free space up to 393 m² at the same time.

7. Conclusion

Since the company aims to increase their production, the warehouse is improved with the suggestions in Stage 1 and Stage 2 to satisfy the peak flows. Different solutions are proposed to the company for the different scenarios to make the analysis flexible. Flexible methodologies provide a system that can be arranged according to the production growth rate. Since the warehouse's current situation causes OHS problems and time waste due to inefficient area allocation, when the production rate increases, warehouse will face more problems and inadequacy in terms of capacity. The project proposes alternative solutions and improvements to eliminate warehouse issues and increase efficiency with the current and possible future peak demand scenarios.

8. Action Plan

The implementation process takes place step by step. Since there are only 2 shifts in the low demand season, third shift time plan as well as Sunday in which production stops, can be used to change the layout. Also, the area between Coil Steel Stockyard and Ordered Sheet Metal, and Truck area are empty and sufficient to use as transportation break locations for the materials.

8.1. Implementation of the Alternative Layout 1

- **Step 1:** Since it's low season, inventory level of production materials might be low which might allow us to squeeze it from top in order to open space for Service Area Office. If it doesn't, area in front of Chemicals Storehouse or the empty area between coil steels and ordered sheet metals, as aforementioned above, can be used for storage. Then, the space freed up can be filled with Service Area Office.
- **Step 2:** With the reallocation of Service Area Office, 145 m² will be freed up. Fixtures and some testing equipment of the Rework Area can be moved over there. Also, by moving up to storage part of the Rework Area, enough space can be opened up to shift Chemical Storehouse to its desired location in the suggested layout.
- **Step 3:** We can shift Chemical Storehouse to the free space gained in the Rework Area. The freed-up space from Chemical Storehouse can be filled with the containers as well as some storage sections of the Rework Area. With the mentioned shift, we can again move Rework Area up and gain free space between Rework Area and newly shifted Chemicals Storehouse. Therefore, we can shift the Production Materials Stockyard at that location.
- **Step 4:** Production Materials Stockyard can be moved up to the space between Rework Area and Chemical Storehouse.
- **Step 5:** Free space gained by moving Production Materials can be filled with Service Area shelfs.
- **Step 6:** The remaining part of Rework Area can be moved up to the space opened up by Service shelfs.

8.2. Implementation of the Alternative Layout 3

- **Step 1:** Due to our observations, in the low demand season, Production Materials occupies approximately 30% (almost 100 m²) of its area. Therefore, starting by the transporting Rework Area materials to empty areas in the Production Materials is reasonable.
- **Step 2:** While the containers and dishwashers which are located at the bottom half of Rework Area are transported to Production Materials' empty areas, cardboards and hinges will be transported to the Rework Area's emptied middle area simultaneously.
- **Step 3:** Chemical Storehouse will be located at the bottom of the free space gained by the Rework Area's transfer.
- **Step 4:** Remaining racks and materials of Rework Area will be transported to emptied Chemical Storehouse area.

REFERENCES

Akın, Can. Interview. Conducted by Sima Alkan, 27 Nov 2019.

"Başkent Freze Strafor Kırma Makinası" *Başkent Freze*, http://www.baskentfreze.com/?newUrun=1&Id=1192176&CatId=bs544200&Fstate=&/Strofar-K%C4%B1rma-Makinesi-130-Serisi. Accessed at 4 May 2020.

Buffa, Elwood S., et al. "Allocating Facilities with CRAFT." *Harvard Business Review*, Vol. 42 Issue 2, Mar. 1964, p136-158, http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=ce3bdd35-0a3d-48a3-9405-21d2e9dee8c4%40pdc-v-sessmgr04

Çevik, Serdar. "Re: Bilkent - Arçelik Proje." Received by Cemil Akgün, 27 Mar 2020. Email Interview.

Çevik, Serdar. Interview. Conducted by Ahmet H. Fidan, 15 April 2020.

"Ertaş Raf Çelik Raf Sistemleri" *Ertaş Raf*, http://www.ertasraf.com/lorem-services. Accessed 15 May 2020.

"Flowchart Maker & Online Diagram Software." *Draw io*, https://app.diagrams.net/

"İlgi Trafik Market" *İlgi Trafik*, https://www.ilgitrafik.com/trafik-ayna-fiyatlari-otopark-aynasi/. Accessed at 15 May 2020.

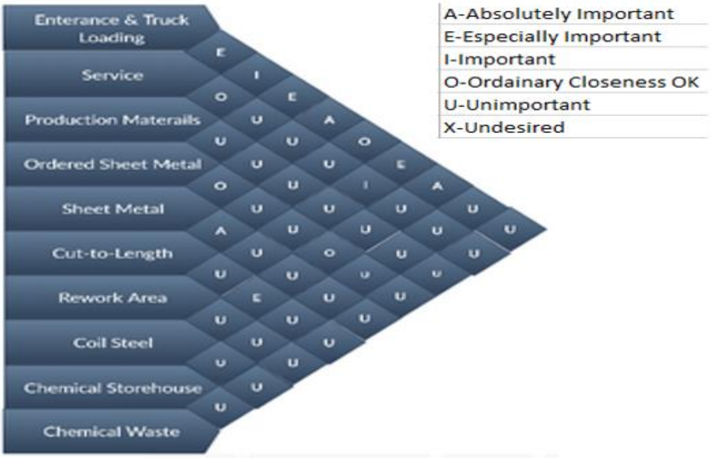
"KLP® Rollblocks" *KLP LANKHORST COIL AND PIPE STORAGE SYSTEMS*, 11 June 2019, http://www.potenza.com.tr/en/urun/demir-celik/klp-lankhorst-coil-and-pipe-storage-systems

"Roll blocks" *Lankhorst Mouldings*, 2020, https://www.coilstorage.com/en/rollblock

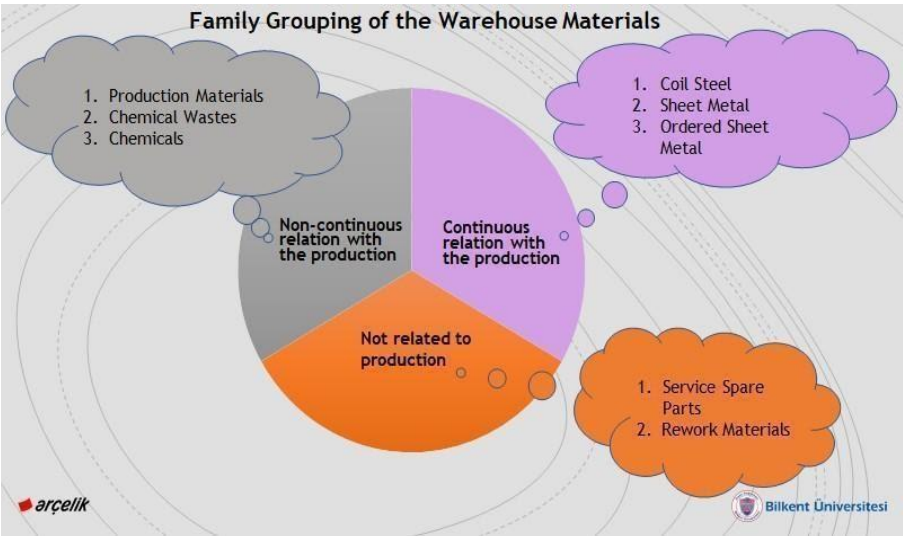
Norén, P., & Eriksson, J. (2011). *A heuristic algorithm for space allocation in a pallet storage warehouse.* Lund University Publications.

APPENDICES

Appendix A: Relationship Chart of the Warehouse Areas



Appendix B: Family Grouping of the Warehouse Materials



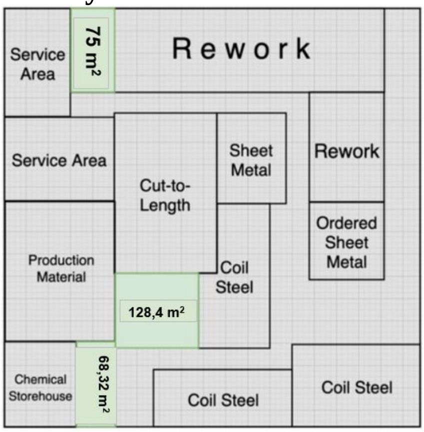
Appendix C: From–to Charts of the Warehouse Flows

					Yearly Forklift Flow							
	Location Name	1	2	3	4	5	6	7	8	9	10	11
1	Entrance/Exit	0	0	9000	4320	0	18000	3600	5400	9720	720	360
2	Truck Parking	0	0	3600	0	0	0	5400	4320	11880	360	360
3	Production Materials Stockyard	7200	3600	0	0	0	0	0	0	0	0	0
4	Coil Steel Stockyard	4320	0	0	0	0	0	0	0	0	0	0
5	Cut-to-Length Lines	0	0	0	0	0	22680	0	0	0	0	0
6	Sheet Metal Stockyard	18000	0	0	0	22680	0	0	0	0	0	0
7	Ordered Sheet Metal Stockyard	3600	10,8	0	0	0	0	0	0	0	0	0
8	Service Area and Stockyard	5400	23760	0	0	0	0	0	0	0	0	0
9	Rework Area	12600	7200	0	0	0	0	0	0	0	0	0
10	Chemical Waste Area	720	360	0	0	0	0	0	0	0	0	0
11	Chemicals Storehouse	360	360	0	0	0	0	0	0	0	0	0

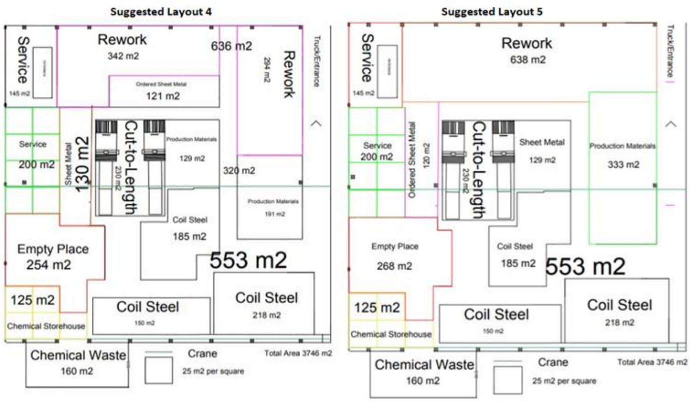
Appendix D: Yearly Forklift Flow of Departments

Location Name	Yearly Flow Rank (Highest to Lowest)
Sheet Metal Stockyard	58.680,00
Rework Area	41.400,00
Service Area and Stockyard	38.880,00
Production Materials Stockyard	23.400,00
Cut-to-Length Lines	22.680,00
Ordered Sheet Metal Stockyard	12.610,80
Coil Steel Stockyard	8.640,00
Chemical Waste Area	2.160,00
Chemicals Storehouse	1.440,00

Appendix E: Alternative Layouts

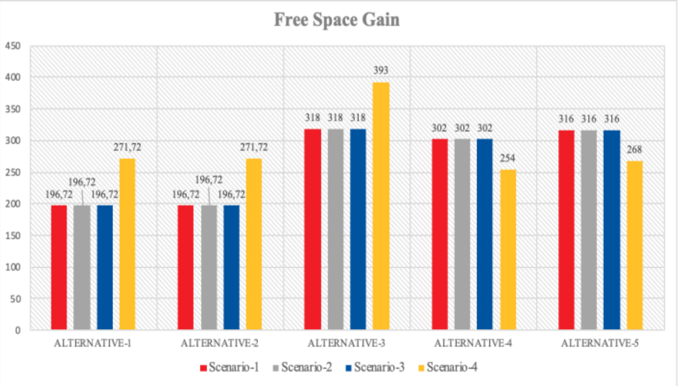


Alternative Layout 2 by CRAFT & Excel methodology

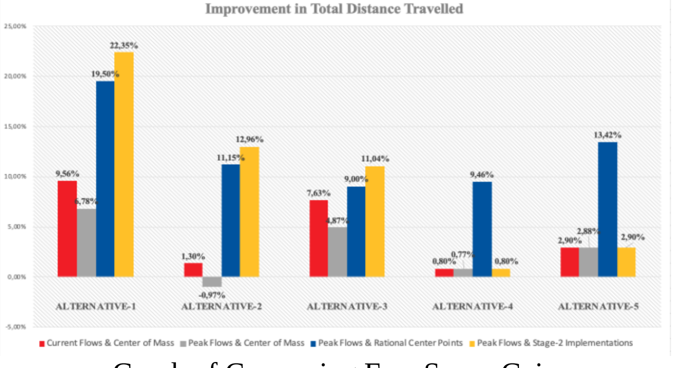


Alternative Layout 4 and 5 by MATLAB & AutoCAD Methodology

Appendix F: Analogy of the Alternatives



Graph of Comparing Improvements in Total Distance Traveled



Graph of Comparing Free Space Gain

Belirsiz Talep Tahminleri Altında İthal Kritik Malzeme İhtiyaç Planlaması

Arçelik A.Ş. Buzdolabı İşletmesi

Proje Ekibi
Büşra Bayrak, Muzaffer Eralp Çalhan, Azra Çağlasu Gürkan, Ebru Tarhan, Zeynep Uçar, Utku Yalınbaş, Gökhan Yetişir

Şirket Danışmanı
Sinem Duman
Üretim Planlama Mühendisi

Akademik Danışman
Dr. Öğr. Üyesi. Özlem Karsu
Endüstri Mühendisliği Bölümü

ÖZET

Buzdolabı ürünlerinde satış taleplerinin dinamikliği, talep tahminlerinde dalgalanmalara neden olmaktadır. Talep tahminlerindeki belirsizlikler ithal malzemelerin tedariğini zorlaştırmakta ve ani gerçekleşen taleplerin karşılanma oranlarını düşürmektedir. Taleplerin karşılanma oranlarındaki düşüşler şirketin elde edeceği olası kâr miktarında kayıplara neden olmaktadır. Bu projenin amacı, talep tahminlerindeki sapmaları ve malzemelerin temin sürelerini göz önünde bulundurarak malzemelerin stok seviyelerini ve taleplerin karşılanma oranlarının iyileştirecek bir karar destek sistemi oluşturmaktır. Geliştirilen karar destek sisteminden alınan sonuçlar ile mevcut sistemin kıyaslanması sonucunda önerilen sistemin stok seviyelerinde %14 iyileştirme sağladığı gözlenmiştir. Bu raporda, önerilen sistemin çalışma prensibi ve şirkete olası faydaları açıklanmaktadır. Oluşturulan sistemin yürürlüğe koyulma süreci hakkında bilgi verilmiştir.

Anahtar Kelimeler: sipariş sistemi, envanter politikası, güvenlik stoku, karar destek sistemi

Imported Critical Material Requirement Planning Under Uncertain Demand Forecasts

1. General Information About Company and Its Services

1.1. Company Information

Arçelik refrigerator plant has 1800 end products that are classified in terms of width, color, energy, and volume. Because Arçelik is a global company and it has a wide range of customers, there is high variability among products. This leads to an even higher variety among materials. Therefore, it is important for the company to make accurate decisions to fulfill the demand while considering the high holding cost.

1.2. System Analysis

In the current system, the ordering process works based on the information gained by Total Visibility Reports (TVR) which were prepared by the Supply Chain Department. These reports contain the demand data of upcoming months and forecast data for each Stock Keeping Unit (SKU). By using these reports Production Planning Department prepares the production plan. Bill of materials (BOM) data and production plan is used to determine the material required for each SKU. Using this information, the time and amount of orders for materials are determined manually and based on experiences.

1.3. Problem Definition

The main problem in Arçelik is the lack of a decision support system that provides the order amount and order time for each specified material systematically. The material orders are placed 6 months before the production of the corresponding SKU. Therefore, the variation between forecast and actual demand reaches 35 percent based on the TVR reports. Due to the unsteady market demand and variation between forecast and actual demands, the order amount cannot be determined accurately. That is why Arçelik encounters an excessive amount of inventory to meet customer demand.

2. Proposed System Development

A proposed system which aims to determine proper order amount and order time for each material for 4 material groups is designed. A heuristic approach is generated by inspiring Material Requirements Planning (MRP). The flow chart of the system can be seen in Figure 1.

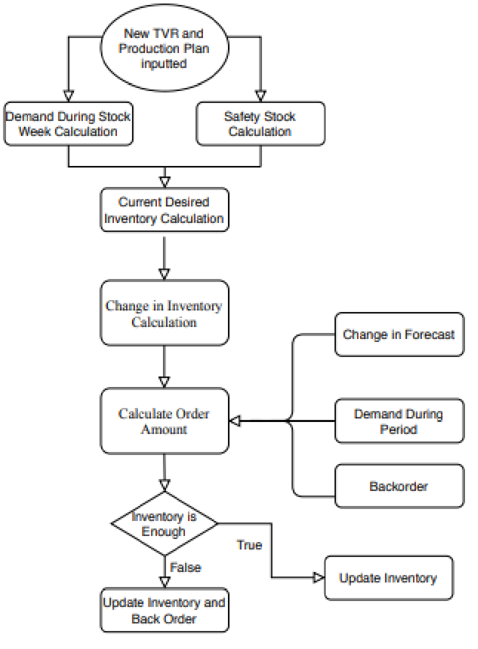


Figure 1. Flow Chart of the System

The scope of the project contains only assigning proper order amounts, that is why, system considers forecasts and production plans to be 100% accurate and it takes Total Visibility Report (TVR) and Production Plans as inputs. Stock-day which is the number of days the company can continue the production without replenishment is an important measure for the company Therefore, the system takes stock-day as an input and calculates demand during stock-day to keep enough inventory.

The first aim of the company is meeting demand, that is why, safety stock amount is calculated. The project focuses on the safety stock calculation and 4 variants of the proposed system, each with different safety stock calculations, can be found in Figure 2.

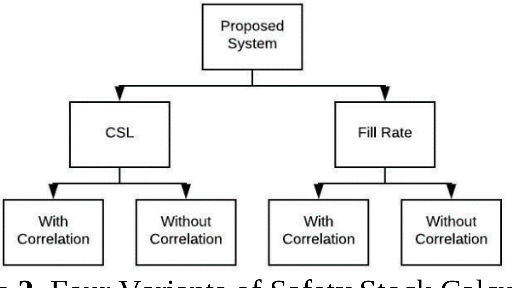


Figure 2. Four Variants of Safety Stock Calculation

The goal of forecasting is to predict the systematic component and estimate the random component. The random component is usually estimated as the standard deviation of forecast error. Even though standard deviation of demand is not necessarily the same as the forecast error, safety inventory calculations could be based on forecast error (Chopra and Meindl, 2000). The uncertainty of supply is not a concern for this project since lead times are given and assumed to be constant. The last factor that has an impact on safety stock level is the desired level of end product availability.

To evaluate safety stock level, following equation is used (Chopra and Meindl, 2000).

$$ss = F^{-1}(CSL) \times \sigma_L$$

$$= F^{-1}(CSL) \times \sqrt{L} \times \sigma_D)$$

$$= NORMSINV(CSL) \times \sqrt{L} \times \sigma_D$$

where ss represents safety stock, σ_D is the standard deviation of demand per period and σ_L is the standard deviation of demand during lead time.

Since considering fill rate instead of cycle service level provides an opportunity to determine desired service level for the company more easily and decreases inventory level, another system is designed with safety stock level considering fill rate. The following formula calculates the difference between safety stock, which is calculated with CSL and safety stock, which is calculated with fill rate. To find the safety stock with fill rate, this difference is added to the safety stock with CSL (Silver et al, 1998).

$$G_u(z)(t) = OrderAmount(t-L) * (1-FR)/\sigma_D$$
$$k = \sqrt{ln(25/G_u(z)(t)^2)}$$
$$Z = (a_0 + a_1 * k + a_2 * k^2 + a_3 * k^3)/(b_0 + b_1 * k + b_2 * k^2 + b_3 * k^3)$$
$$ss_{FR} = ss_{CSL} - Z * \sigma_D$$

a_0	a_1	a_2	a_3	b_0	b_1	b_2	b_3	b_4
-5.3925569	5.6211054	-3.883683	1.0897299	1	-7.249648×10^{-1}	5.073266×10^{-1}	6.6913686×10^{-2}	-3.291291×10^{-3}

where $G_u(z)$ is the expected value shortage per order cycle and FR represents fill rate.

Then both systems are designed by considering correlation factor for forecast errors of each material. In this case, the standard deviation of material i is calculated by considering the correlation between the forecast errors of each SKU (j, k, …) which contains material i. The following formula is used to calculate the standard deviation of material i by considering correlation factor (Baker, 1985).

$$s_i = k_\beta \sqrt{\sum_{j=1}^{n} a_{ij}\sigma_j^2 + \sum_{j=2}^{n}\sum_{k=1}^{j-1} 2\,\rho_{jk}\, a_{ij}a_{ik}\sigma_j\sigma_k}$$

By using demand during stock-day and safety stock level, current desired inventory level is calculated. This value may differ from the current inventory level because of some unexpected situations. That is why, the system calculates the difference between current inventory level and current desired inventory level and it reflects this difference to the order amount. If current inventory level is greater than the desired inventory level, order amount is decreased by this difference; if current inventory level is less than the desired level, order amount is increased.

Every week a new TVR is created which means that the forecasts may change every week. To work with the most accurate forecast, system considers the difference between the forecast of previous week and current week. If the forecast is increased, the order amount is affected positively; if the forecast of previous week is larger than the forecast made this week for the same period, the order amount is decreased by the difference.

In order not to miss the demand, the system is willing to hold the amount of the demand of the period after lead time as inventory. That is why, this amount is added to the order amount.

Backorder levels are the last component of the order amounts. To meet the unmet demand as soon as possible, backorders are added to the order amounts.

In the final stage, inventory and backorder level are updated. If inventory for each material of an SKU is enough to cover actual demand derived from Production Plan, then the system updates the inventory by decreasing by actual demand amount. If inventory level of even one material of the SKU is not enough to cover actual demand, system decreases inventory levels by the amount which is possible to be produced. Then, it writes unsatisfied amount as back order.

3. Validation of the System

To validate the system constructed, a simulation is made for the 2019 data. Since the maximum leadtime is 23 weeks, the simulation started from the 24th week and continued until the 45th week. The system was analyzed by changing the parameters for the safety stock calculation and four different scenarios are obtained. To begin with, the 4 variants of the proposed system that can be seen in Figure 3 were compared with respect to inventory and backorder. The ultimate goal was to choose the one which gives the best results and to compare those results with Arçelik data. Consequently, the best results were obtained from the system with the one which considers fillrate instead of CSL, and ignores the correlation.

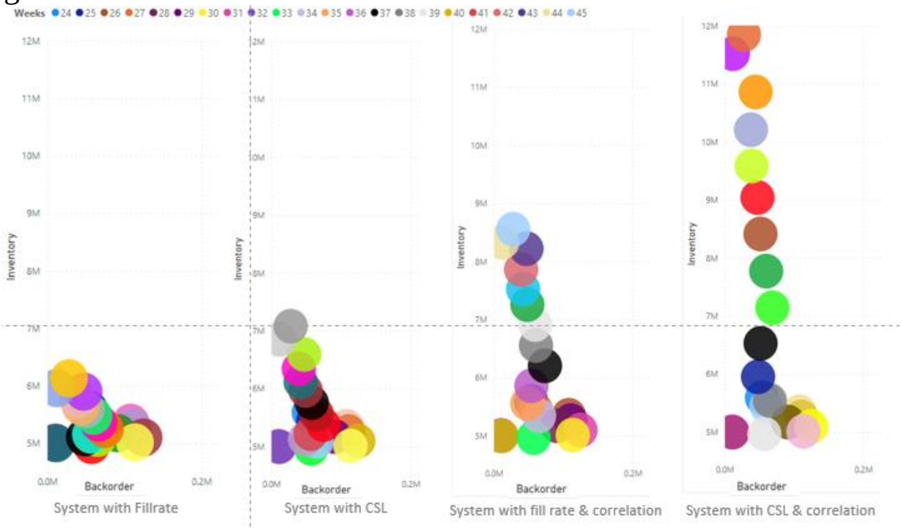


Figure 3. Different Scenario Analysis

Figure 3 indicates that this system overall was better than the other 3 as the inventory levels are much better compared to the others, while the backorder levels change between 10,180 and 124,417 for all variants. It can be seen that the highest inventory level observed in this system was 6,134,458 and it increases as we get to the last scenario. In the last system the highest inventory level is equal to 11,853,826 which is much higher than the proposed system.

The results from the selected system were further compared to results from Arçelik data. In Figure 4, while all the circles on the right represent the data from Arçelik, on the left side our results can be seen. To illustrate, in week 30, our system's inventory level was 5,018,131 and Arçelik's was 8,572,045. Furthermore, in the same week, while we had 113,214 backorders, Arçelik had 663,766. Similar comparisons can be made for each week and it can be said that our system gives better results in terms of both inventory levels and backorder levels in each week. In addition, the average inventory for Arçelik is equal to 8,642,876 and for our system it is equal to 5,384,463. For backorder levels, it is calculated that Arçelik's average backorder is 803,540 and ours is 60,182.

4. Implementation

To implement the system, a user interface is designed on Microsoft Excel. The main page shown in Figure 5 welcomes user with two buttons. Once the user clicks the button "Run the System", the page shown in Figure 6 appears. To run the system, user is asked to input TVR, production plan, desired service level and desired stock and click on the "Run" button.

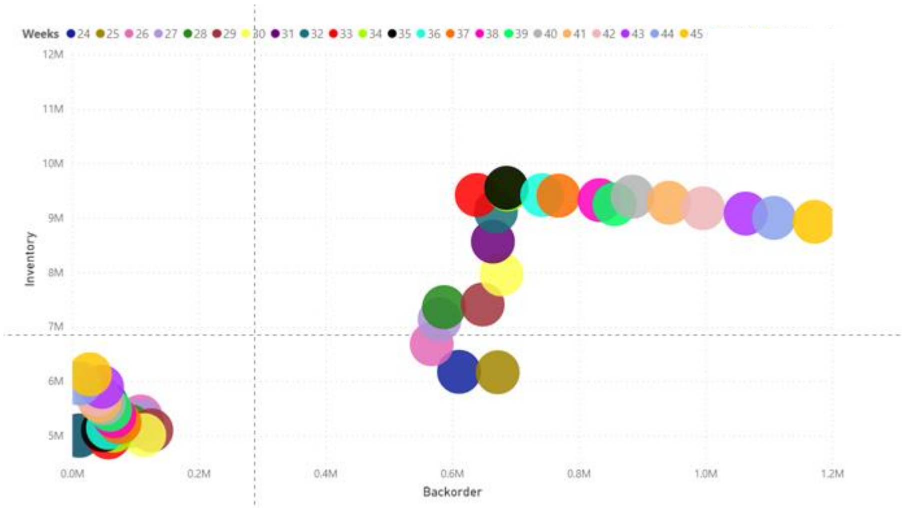


Figure 4. Comparison of Selected System and Current System

The necessary order amount to be placed is available on the sheet "Order_amount". User may also track the current inventory and backorder amounts on "Inventory" and "Backorder" sheets, respectively. Users are also enabled make changes on Bill of Materials (BOM). Once "Change BOM" button on the main page is clicked, the page shown in Figure 7 appears. On this page, user may either add a new SKU or a new material to the system by the corresponding buttons.

5. Benefits to the Company

Benefits of the proposed system can be explained in terms of inventory level and backorder level. The inventory level of Arçelik's system and the proposed system are compared, and it is observed that the proposed system leads to keep less inventory than the current system in Arçelik. The highest inventory level of the proposed system is 6,134,458 in week 45 where the inventory level of the current system is 8,934,559 in week 45. Moreover, the maximum inventory level that the current system reaches is 9,558,708 in week 35, which is approximately 1.56 times more than the highest inventory level of the proposed system. Therefore, the inventory level of Arçelik is said to be improved by 14% on average by keeping safety stock and considering fill rate.

Another benefit that the proposed system provides is the reduction in backorder levels. The reduction might be an important improvement for the company, since its main aim is to satisfy the demand. As it can be observed from the Figure 4, the backorder level of the proposed system is lower than the current system in each week. On average, the backorder level of the proposed system is 83,106 while it is 561,530 in the current system. Hence, the backorder level of the company is claimed to be reduced by the proposed system.

As taking the benefits into account, the company might keep less inventory in order to alleviate the effects of the unsteady market demand and mismatch between forecast and actual demands. Additionally, the proposed system keeps safety stock in order not to be backordered. Therefore, decrease in the backorder level may indicate a benefit which avoids the unsatisfied demands.

REFERENCES

Chopra, Sunil, and Peter, Meindl. *Supply Chain Management*: Strategy, Planning and Operation. Pearson, 2000, p. 326-372. *EBSCOhost*, doi:10.1016/0272-6963(85)90031-2.

Kenneth R. Baker, "Safety Stocks and Component Commonality." *Journal of Operations Management*, vol. 6, no. 1, Nov. 1985, p. 3–22.

Silver, Edward A., et al. *Inventory Management and Production Planning and Scheduling*. 3rd ed., Wiley, 1998. p.268.

APPENDIX

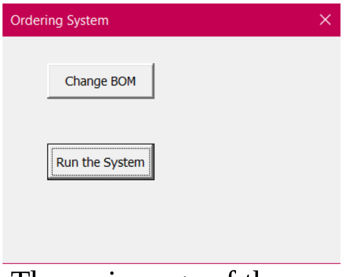


Figure 5. The main page of the user interface

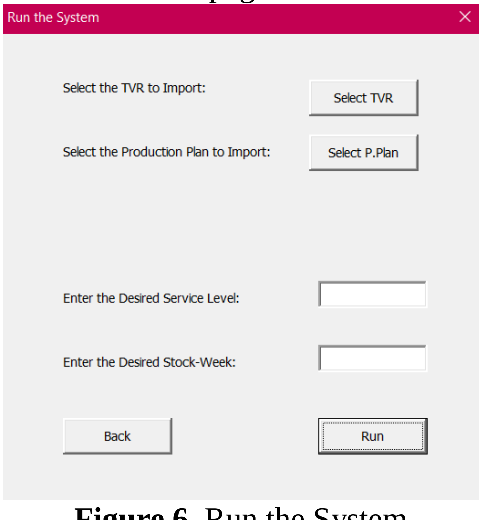


Figure 6. Run the System

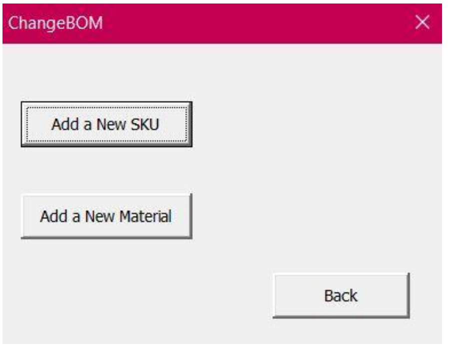


Figure 7. ChangeBom

Eklemeli İmalat Merkezlerinde Karar Destek Sistemi ile Teslim Süresi Tahminleme Metodu

Arçelik A.Ş. Garage

Proje Ekibi

Göktürk Balıkcı, Zeynep Gürakın, Yusuf Karasayar, M. Emir Kayrak, Mustafa Özdemir, Pınar Şener, Deniz Taşkeser

Şirket Danışmanı
İbrahim Güzel
Kıdemli Uzman

Akademik Danışman
Doç. Dr. Yiğit Karpat
Endüstri Mühendisliği Bölümü

ÖZET

Arçelik Merkez AR-GE'ye bağlı Arçelik Garage & Maker Lab (AG), prototip taleplerinin net teslim süresini verememekte, yeni gelen talepleri ve önceliklendirmeleri göz önüne alarak süre optimizasyonunu otomatik yapamamaktadır. Projenin amacı, gelen talepleri uygun iş paketlerine yerleştirip tahmini teslim süresini müşteriye aktaracak bir karar destek sisteminin oluşturulmasıdır. Bu amaç doğrultusunda, AG'de çalışan uzmanların zımni bilgileri, veri analizleri ve durum sınıflandırmaları aracılığıyla açıklanabilir ve kullanılabilir hale getirilmiştir. Proje ürünü, Python arayüzü kullanılarak oluşturulan karar destek sistemidir.

Anahtar Kelimeler 3D Yazıcı, Ardıl İşlem, Eklemeli İmalat

Delivery Time Estimation Method Through a Decision Support System for Additive Manufacturing Hub

1. Company & System Description

Arçelik was founded in 1955 at Sütlüce, İstanbul as it headquarters. Company initiated its production activities in 1959 and within 5 years, it has 5 production facilities in different locations. Arçelik A.Ş. now has 30.000 employees all around the world, has 18 production facilities in 7 countries. Arçelik Garage is initiated by Arçelik and supports start-ups and facilities with prototyping. Arçelik Garage has lots of high-technology equipment that enables advanced production techniques for additive manufacturing (AM) which are SLS, SLA, etc. Having these technologies and machines make Arçelik Garage one of the largest rapid prototyping (RP) centers in Turkey.

2. System Analysis & Problem Definition

2.1. System Analysis

In current system of Arçelik Garage, desired prototype's 3D models arrive Arçelik Garage's E-Lab system. The one who orders, has to register a project number, a due date, the intended use of the prototype, the number of parts needed, the master model and 3D format file of the part. SLS technology is under consideration for this project. Specialists in Arçelik Garage accept or reject the order by considering the part geometry, whether they can produce the order and this procedure is called "work assignment". If an order is accepted, combining with some other orders previously accepted or will accepted later, nesting operations can be done. Nesting is done automatically by a software and a work package is formed after nesting operation. Work packages may consist of different orders from different customers and individual parts in the work package will be produced together at once in the same machine. Then, work package formed is ready to load to the appropriate 3D printing machine. The technology and machine needed to produce the orders are decided by customers, but it can be changed by specialists in Arçelik Garage after negotiation with customer, because preferred technology for the order may not be suitable for the specifications of that order. After the production of work package, it needs to be cooled down to be able to start post-processing. As Gibson et al. (2016) argues that almost each part needs manual interventions to end up with exact geometry and surface quality, these interventions are called "post-process". Required post-process operations may differ from part's geometry and from the technology chosen. The orders are ready to be packaged and shipped to the customers after post-processing operations accomplished.

2.2. Problem Definition

The main problem that Arçelik Garage faces with is they cannot give a true estimation of delivery time to their customers for their orders. The reason behind this is they do not know some durations in the order of operations.

Decision of accepting or rejecting an order is done by Arçelik Garage's specialists but this is not a process taking too much time to decide, as it was said so. The time between order arrival and work assignment is highly variant; after order arrival, sometimes work assignment is done on the same day, sometimes after 1 day but sometimes after 7 days. This variation is highly dependent on the unavailability of the machine and waiting orders. Therefore, this delay affects the estimation of the time of delivery. In addition to that, after completion of work assignment, the duration of forming a work package is again cannot be estimated. Deciding which individual orders will be in the same work package is a tacit knowledge and decision is given by the specialists in Arçelik Garage, according to their prior backgrounds. Delay here occurs because of the backlogs.

Another issue that causes unknown delivery time is the post-process operations. Majority of the post- processing elements are manual, depending on human labor. Some printing technologies require less post processing operations, some requires additional ones. The time it takes obviously related to the specifications of the work package. It may be dependent on volume of the work package, or surface area, or another metric of the work package. If all these elements points are considered, Arçelik Garage cannot give an estimation of delivery time for the products ordered.

2.3. The objectives & scope of the project

The scope of the project is to create a decision support system (DSS) to Arçelik Garage. The project enables Arçelik Garage to give a delivery time to its customers' orders. This DSS will be 3- Point DSS giving 3 feedbacks will throughout the process.

First support point is after the arrival of demand, giving customer a feedback about a rough date of delivery. Second support point is before the work assignment, giving feedback to both Arçelik Garage experts and customer; since order is accepted and ready to be go into a work package, the delivery time to the customer will be more accurate compared to the first support point. Third support is giving feedback to the Arçelik Garage expert, related to post- process operations. Printing time, cooling time and post-process time will be showed in total to expert and this will be the most precise time. By utilizing this output, facility can have standardized decision-making process and a systematic, dynamic and a sustainable system. Providing the total delivery time systematically will remarkably increase the customer satisfaction. Additionally, the system will be more reliable and less dependent on the personal judgement of employees since it relies on a systematic approach and relevant algorithms for the estimating delivery times. By knowing the true delivery times of orders, Arçelik Garage can schedule their machines easier thus can broaden their manufacturing power with the help of the system.

As a part of delivery time estimation, we provide a work package estimation formulation to experts that they can use to get higher efficiency in terms of capacity allocation. With the regression analysis we put on use, the relations between part features and printing time are well-defined and this will contribute to the evaluation ability of the engineers.

3. Literature Review

3.1. Literature review for MTO companies dealing with due dates

Since the main aim of the project is to create a decision support system that shows the delivery times, the approaches of the large scale make-to-order (MTO) manufacturing companies are investigated. Park et al. (1999) argue that heuristic delivery date decision algorithm (HDDDA) is used to determine a feasible delivery date with the current capacity. According to Hendry et al. (1998), the workload control (WLC) concepts are applied to address decision levels in MTO firms, the job entry-level and job release level, which helps us to consider this perspective while designing our delivery estimation algorithm.

3.2. Literature review for order review/release analysis

To estimate the delivery times, workflow in the shop floor is also important. Therefore, Shop Floor Control (SFC) and Order Review/release (ORR) activities were also investigated. As Melnyk and Ragatz (1989) say, ORR capsulate all the operations starting from the arrival time of an order until the time that order is in process in the shop floor. According to Sabuncuoglu and Karapınar (1999), objective of the ORR is improving the overall performance of system and it is used in workload control systems (WLC) and helps to determine which orders can be released to the shop floor and the time that these orders should be released to the shop floor. If

In Arçelik Garage's case, methods that based on calculated release times are more appropriate. Idea behind this is releasing the orders (jobs) at predetermined times based on an estimation of flow times in the system; trying to control the workload of shop floor. Since Arçelik Garage is dealing with the problem of giving a due date to its customers, these systems were useful for us to follow.

4. Proposed System & Methodology

The main aim of this project is converting the tacit knowledge of the workers in Arçelik Garage to an explicit knowledge. Although the workers in Arçelik Garage have unclear estimations about the delivery times, they do not have standardized and easily accessed datasets about duration of the work package's operations. The expected outcome of the project is a decision support system that shows the delivery times in every process of the printing workflow and offers work package formations. Delivery time estimations of demands change according to the locations of demands in the printing process timeline. Besides that, some of the components affecting delivery time like printing duration is known for all the cases. Yet, other components of the delivery time vary like cooling, post-process times before the creation of actual work package, etc. Unknown variables like the time between demand arrival and work assignment will be estimated by statistical analysis, forecasting and utilize from the expert opinion.

The decision support system will be accessed through a computer software with a user-friendly interface that easy to manage, Python tkinter. Our forward-looking expectation is decreasing delivery time of the products with decreasing the waiting time in the system.

4.1. Formulation of the System

General formula for delivery time calculation is developed as follows:

$$\alpha = \left[x + \left(\sum_{n=0}^{N} (M_0 + M_{n1} + M_2 + p_n) \right) + d \right] \times (1 + S)$$

N: Number of estimated work packages (backlog)

α: Estimated total delivery time

x: Remaining time of current printing process or time until end of the shift

M_0 : Preparation time of printer (average value)

M_{n1} : Printing time of predicted nth work package n=0, 1, 2, …, N

M_2 : Cooling time of printed work package (a function of printing time)

p_n: Post-process time of predicted work package n. n=0, 1, 2, …, N

d: Delay (day factor – hour factor), S: Safety factor (%10 of remaing. Sum.)

4.1.1. Work Package Estimation Method

When the system has too much backlog and printers operate simultaneously, we need to form predicted work packages to arrange process flow. These work packages will be called “predicted work packages” as we estimate them by our algorithm not by tacit knowledge. Predicted work packages will be formed according to load factor, which means the summation of x-y-z dimensions of each part. Load factor has a maximum point that we should stop putting the orders to the same work package. According to the past data given, maximum x-y-z multiplication of a work package is known and therefore the correlation between x-y-z dimension multiplications of work packages and sum of x-y-z dimension multiplications of the parts are found by regression model in R by using linear model. Therefore, we can put limit on the sum of x-y-z multiplication per part in terms of work package x-y-z multiplication. By this mean, we reveal the load factor to form work packages. Independent variable is work packages x-y-z dimensions’ multiplication. Dependent variable is sum of

each part's x-y-z dimensions' multiplications in a work package. Fit line can be seen in Appendix A. Below is regression formula:
(Sum of parts x-y-z) = -5.849e+06 + 1.022*(Package x-y-z)
Adjusted R-Squared Value: 0.5939 and p-value: 0.000761

4.1.2. Printing Time Estimation

To give a total delivery time, printing time of the predicted work packages needed to be known. After trying lots of explanatory variables, the ones giving the most powerful correlation are being found. Fit line can be seen in Appendix B. Below is regression formula: (Printing Time) = 2.942e+00 + 1.165e-05*(Work Package Volume) + 2.116e-05*(# of Parts) + 5.402e-02*(Sum of Surface Area of Parts) + -8.779e-12* (Work Package Volume*Sum of Surface Area of Parts)
Adjusted R-squared Value: 0.5693 and p-value: 2.2e-16

Hence, printing time was estimated for predicted work packages that are formed according to the regression model. For verification purpose, we allocate 10 of work packages to test our proposed system and 87 of them to develop the algorithm. Comparison of the results can be seen in Appendix C.

4.1.3. Post-Process Time Estimation

Post-process time of each estimated work package was calculated via regression model that we have found in R. Explanatory variable is machining time of that work package for the time estimation of post-process. When we formed this regression model, we have used the tacit knowledge of the expert. Fit line can be seen in Appendix D. Below is regression formula: (Post-Process Time) = -0.081228 + 0.093686*(Printing Time)
Adjusted R-squared value: 0.919 and p-value: 2.2e-16

Similarly, for verification purpose, we allocate 10 of work packages to test our proposed system and 87 of them to develop the algorithm. Comparison of the results can be seen in Appendix E.

4.2. Cases for Delivery Time Estimation

The decision support system considers the current state of the machine and current workload. There are three different cases according to current state of the machine and current workload of work packages. They are:

- Machine is idle.
- Machine is working, demand is not enough to create a work package.
- Machine is working, demand is enough to create more than a work package.

Each case has its own components which will be explained later in this report but there exist common components in each case. There is a calendar factor in common, called "day factor". It stands for the time duration that will be added to the total estimated time according to ending time of the current printing operation or estimated ending time of the next printing operation. Day factor stands for the operations that weekend intervenes the operations. For instance,

the printing has ended at the end of the shift on Thursday, the remaining operations will take Friday and the end product can be send on Monday. These 2 days after operations will be added as day factor component.
Another common component in calendar factor is the safety factor. It is determined as 10% of the sum of prior operations' durations. It is added by considering the effect of human factor and occasional situations such as machine malfunctions and leisure days.

4.2.1. Case 1

It's an idle stage of machine. When the machine is not working and the demand is not enough to create a work package, there is either no demand or the demand load is not enough to run the machine to its full capacity. If the desired capacity of the machine cannot be reached, then the current demand forms a work package. A shift including lunch break is 9 hours and printing operation does not finish in less than 8 hours and cooling takes at least half of the printing time. These two incidents takes at least 12 hours so they have to wait until the end of the shift even if the full capacity is reached, considering that until the end of the shift, there may be prioritized new demands arriving which ends up with forming a different work package. The total estimated delivery time is the summation of the components that are shown in Figure 1.

Rem. time until printing starts (end of shift-arrival time) + Printing Op. (case1, mean of the printing times) + Delay (day fact. + safety fact.)

Figure 1. Components of estimated delivery time, case 1

"The remaining time until printing starts" is the first component, because the printing never starts before the end of the shift in this case. Also, the total daily demand usually does not exceed the capacity of the machine until the end of the shift according to the engineers in AG. In fact, it rarely exceeds the half of the capacity. Thus, average of the printing times of the past operations with less than or equal to half capacity average is taken. Lastly, day factor and safety factor are added.

4.2.2. Case 2

In this case, the machine is working, and the current workload is not enough to create a work package. The total estimated time is the summation of the components that are shown in Figure 2.

Rem. time until printing ends (end of printing-arrival time) + Op. Time of incompl wp(case2, mean of printg.time) + Delay (hour&day fact. + safety fact.)

Figure 2. components of estimated delivery time, case 2

The remaining time until the end of the printing is the first component.
Then, in case 2, average of the printing times of the past operations is taken because there are times that the next work package uses either full capacity or almost zero capacity. The capacity usage can differenciate between these two

edges. As another difference from case 1, there is another calendar factor called hour factor. This delay mainly stems from the time that should be waited between two printing operations, preparation time. This time period usually results with 1 day delay to start the printing operations. Lastly, day factor and safety factor are added.

4.2.3. Case 3

In this case, the machine is working and the current workload creates more than one work package. The total estimated time is the summation of the components that are shown in Figure 3.

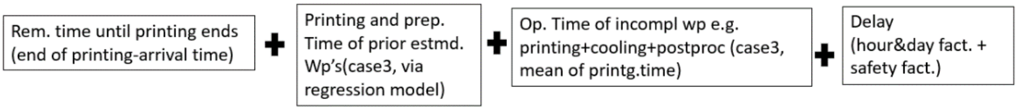


Figure 3. Components of estimated delivery time, case 3

In case 3, for the operation time of incomplete work package that arriving demand belongs to, the time until the starting time of printing of the arriving demand is more than one printing time. So, there is a long time period until the printing operation of the last incomplete work package. In this time period, AG usually receives a large number of orders. Thus, the average of the printing times of the past operations with bigger than or equal to half capacity is taken. The regression model for work package estimation is utilized and according to the estimated work package features, the regression model for printing time estimation is utilized and this time period is added. Lastly, day factor, hour factor and safety factor are added.

4.3. Inputs and Outputs of the Proposed System

The system needs to take the inputs given below from the user to estimate the total delivery time of a product:

- Demand number
- Demand's volume
- Demand's x,y,z dimensions (cm)
- Time left until the current printing operation ends
- Priority level of the order

The more accurate those inputs are the more accurate the system will respond. All the inputs above are going to be considered while generating the outputs. The outputs of the system are:

- Estimated delivery date of the order
- Estimated work packages
- Estimated post-process times

4.4. Outcome and Benefits to the Company

The first support point takes place between the demand arrival time and package creation, and the aim of this point is giving a rough delivery time. We aimed to develop an algorithm that calculates the delivery time with analyse of the component of the delivery time independently and adding the variance effect. Arçelik Garage employees should enter the demands to the software when the demand arrives. Machine capacity for SLS technology is related with sum of (Volume / Surface Area) for each product. Since details of the work package uncertainty are considerably high in this stage, average printing and post-process time is used. Remaining printing time of work package that works in progress is known in this stage, it is used as an input for the system. In the second support point, the advice of potential work packages is given to the expert before forming actual work package. Most precise delivery time is given in the third support point since there will not be unknown delay times here. Since the third support will be given after the work package creation, printing and post-process times will be input for system from the Magics output.

Python software is chosen to develop the algorithm and user interface. The specifications of the product such as id, volume, axis, and priority are taken as an input and are stored with the usage of dictionary method in Python. Volume of the products stored in an array list. However, in the work package selection process, if a product does not fit a work package because of volume limitation, it rolls another work package array list. Most appropriate way to choose work package is using the total work package volume but separate volume of the products is used for the work package selection due to the limitations. However, when the high priority product arrives, software accommodates the product in the very first work package even the next work package is full. After the selection process, if first work package exceeds the load factor, the last product that arrive just before the high priority product arrival is being put in the second work package. The design of the user interface of the decision support points can be found in Appendix F.

The decision support system gives an estimation of the total delivery time as output and by utilizing this, facility can make more accurate production planning and scheduling and the customer can rely on the estimated time to provide service or do modifications accordingly. Providing the total delivery time systematically will remarkably increase the customer satisfaction. As a part of delivery time estimation, we provide a work package estimation formulation to Arçelik Garage which they can use to acquire higher efficiency in terms of capacity allocation. With the regression analysis we put on use, the relations between part features and printing time are well-defined and this will contribute to the evaluation ability of the engineers at Arçelik Garage.

REFERENCES

Arçelik, "Arçelik A.Ş Tarihçe". 19 Ekim 2019 tarihinde erişildi. http://www.arcelikas.com/sayfa/76/Tarihce. [Accessed: 19-Oct-2019]

Gibson, Ian, et al. "Additive Manufacturing Technologies - 3D Printing, Rapid Prototyping, and Direct Digital Manufacturing: Ian Gibson." Springer, Springer-Verlag New York, www.springer.com/gp/book/9781493921126.

Hendry, L.c., et al. "The Effect of Workload Control (WLC) on Performance in Make-to-Order Companies." *Journal of Operations Management*, vol. 16, no. 1, 1998, pp. 63–75., doi:10.1016/s0272-6963(97)00011-9.

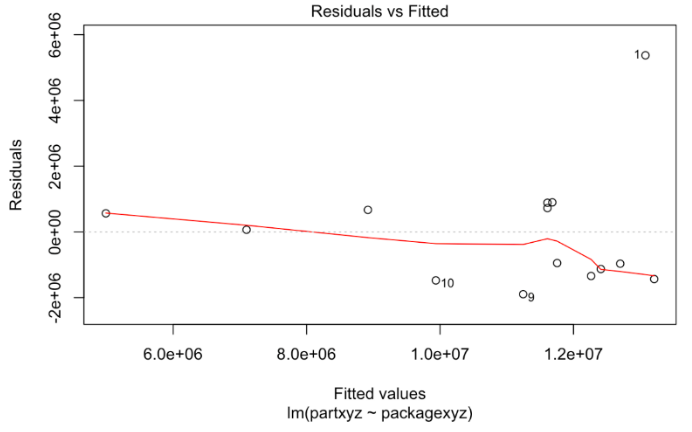
Melnyk, Steven A., and Gary L. Ragatz. "Order Review/Release: Research Issues and Perspectives." *International Journal of Production Research*, vol. 27, no. 7, 1989, pp. 1081–1096., doi:10.1080/00207548908942609.

Park, Changkyu, et al. "Delivery Date Decision Support System for the Large Scale Make-to-Order Manufacturing Companies: A Korean Electric Motor Company Case." *Production Planning & Control*, vol. 10, no. 6, 1999, pp. 585–597., doi:10.1080/095372899232885.

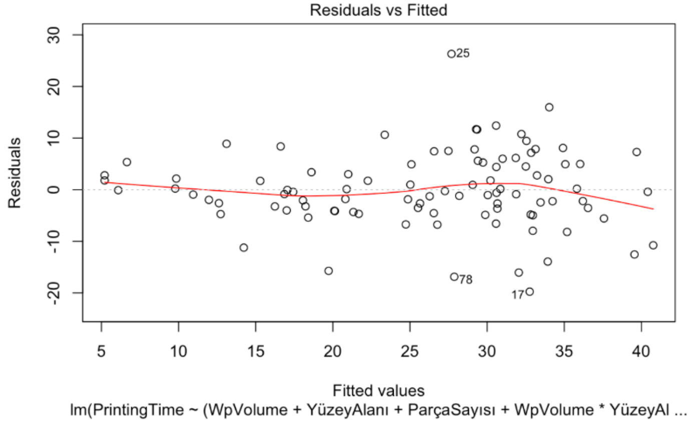
Sabuncuoglu, I., and H.y. Karapınar. "Analysis of Order Review/Release Problems in Production Systems." *International Journal of Production Economics*, vol. 62, no. 3, 1999, pp. 259–279., doi:10.1016/s0925-5273(98)00248-5.

APPENDIX

Appendix A: The Fit Line and Residuals for Load Factor

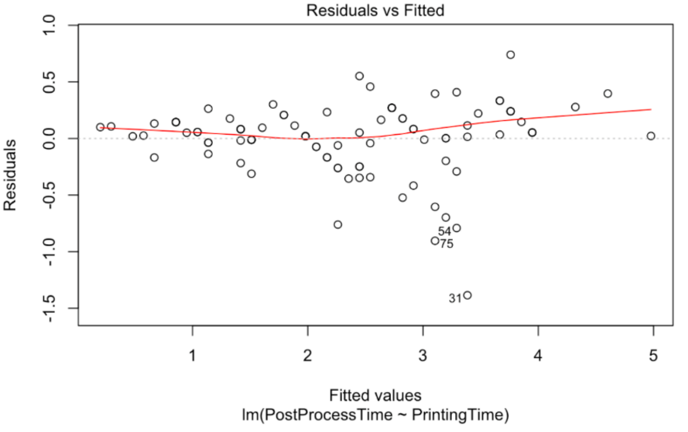


Appendix B: The Fit Line and Residuals for Printing Time Estimation

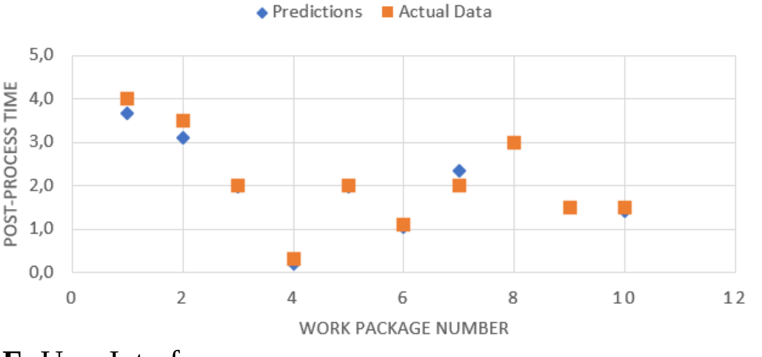


Appendix C: Verification for Printing Time

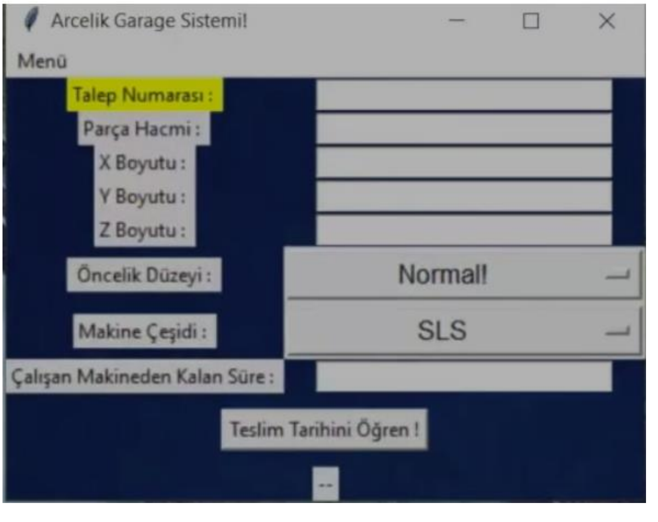
Appendix D: The Fit Line and Residuals for Post-Process Time



Appendix E: Verification for Printing Time



Appendix F: User Interface



Paralel Makine Sistemleri İçin İş Ataması Yoluyla Kapasite Belirlenmesi

Arçelik A.Ş. Pişirici Cihazlar İşletmesi

Proje Ekibi

Burak Altıntaş, Ali Batuhan Karahan, Saad Malik, Didem Onuk, Özge Özkul, Başak Uzun, Müge Vatansever

Şirket Danışmanı
Murat Çanlıoğlu
Endüstri Mühendisliği Takım Lideri

Akademik Danışman
Dr. Öğr. Üyesi Firdevs Ulus
Endüstri Mühendisliği Bölümü

ÖZET

Bu projenin amacı, Arçelik Pişirici Cihazlar İşletmesi Mekanik Atölyesindeki transfer presi makinelerinin üretim planı gereksinimleri için uygun maliyetli, sistematik makine-iş ataması yapan ve üretim planındaki kapasite aşımlarını görüntüleyen bir sistem tasarlamaktır. İşletmedeki talep artışı sistemde kapasite sorunlarına yol açmaktadır. Proje kapsamında toplam üretim maliyetini enazlamayı amaçlayarak makine-iş ataması yapan bir matematiksel model geliştirilmiştir. Geliştirilen sezgisel yöntem ile de matematiksel model sonucunda boş kalan makinelere işler adetlere bölünerek atanmıştır. Geçmiş veriler ile model çalıştırılmış ve daha etkin makine tahsisi sağlanmıştır. Oluşturulan model kullanıcı dostu arayüz tasarımına entegre edilerek karar destek sistemi oluşturulmuştur.

Anahtar Kelimeler: Makine iş ataması, Maliyet enazlama, Kapasite yönetimi

Job Allocation in Parallel Machine Environment For Capacity Determination

1. Company Description

Arçelik A.Ş which is a part of the Koç Holding Consumer Durables Group was founded in 1955 in Istanbul. The majority of the company is owned by Koç Holding. Arçelik A.Ş is a market leader in the home appliances industry and has 20 production facilities and over 30,000 employees in 8 countries.

2. System Analysis and Problem Definition

2.1. System Analysis

The project that our team is responsible for revolves around the Arçelik Cooking Appliances Plant (CAP) Mechanic Shop. In this part of the CAP, there are 21 workstations and 4 different machine types: cavity line (6), transfer press (9), roll forming (2) and punta (4). The main focus of this project is transfer press machines. These machines are specialized for various types of pressing operations. For some production operations, these machines can be substituted for each other.

Currently, initial machine work allocation is made by planners according to the production plan in an Excel-based system, therefore, the allocation is not systematic. The production plan is scheduled as 3 shifts each day and the factory uses 6 days of the week as regular production. The seventh day of the week is used as overtime. If there is an infeasible machine work schedule, the production department allocates machines and jobs manually. Reallocation is generated according to their experience, machine parts incidence matrices and capacity utilization of machines. Also, the overall equipment effectiveness (OEE) of each machine is a significant concern for machine job assignments. The OEE at Arçelik is defined as the proportion of machine capacity which is expected to be utilized within a time period. These values depend on the performances of the machines in the previous month, and they change every month. The OEE values utilized in this project are directly provided by Arçelik.

2.2. Problem Definition and Scope of the Project

Capacity overflows frequently occur in the Mechanic Shop with an increase in the annual production of Ovens and Hobs. There are 3 root causes of this problem. The first reason is inefficient machine-job allocation. Another reason is that job assignments are not systematic. In the current system, the assignments depends on operators' experience, hence, it may cause unbalanced capacity utilization. The last reason is the inability to detect overflows in advance and low monitorization of capacity.

The objective of this project is to design an efficient, flexible and customizable decision support system for Arçelik CAP. This system provides a cost-effective and systematic machine-job allocation by using the production plan. Also, the system allows the company to monitor the capacity utilization of transfer press machines and the assigned jobs in each machine. These machine job allocations, capacity utilization, and unassigned jobs can be seen through a user-interface. The interface which is integrated with the mathematical model can show capacity assignments upon regular updates to the user.

3. Proposed System

3.1. Mathematical Model

The mathematical model finds the minimum cost allocations of the jobs to the machines. In particular, it determines machine job allocations for both normal and overtime horizon, and the unassigned jobs. This model is solved by Excel OpenSolver. There are two sets in the model, set of jobs and machines. The model requires the following inputs; the production plan, current capacity utilization of machines, setup times, electricity consumption values, required the number of labors for each machine, setup costs, electricity costs, labor costs, and the cost of unassigned jobs. These inputs are used as parameters of the mathematical model. The objective of the model is minimizing the total production cost. There are 4 cost components; total electricity cost, labor cost, setup cost and unassigned cost. Also, there are 8 constraints. The first two constraints (1, 2) ensure that the assignments are made by considering the capabilities of the machines, that is, job i is assigned to machine j only if it is capable of producing job i. The third constraint (3) ensures that all of the demanded jobs are produced, either at the normal horizon or at the overtime horizon or unassigned for next week. The fourth and the fifth constraints (4, 5) check that the run time of each machine does not exceed the time limit both in the normal horizon and overtime horizon. The sixth and the seventh constraints (6, 7) are used to calculate the setup count. There are two important assumptions for this calculation. The first one is that for each job change, setup is required in each machine. The second one is that at the start of the week, setup is required for each machine. The last constraint (8) is about decision variables for being binary or integer, and nonnegative. Parameters, decision variables, constraints and the objective function of the model can be found in Appendix A.

3.2. Heuristics

When the mathematical model is solved, there may be some unassigned jobs in the system even though there is enough capacity in the machines as can be seen in Appendix B. The reason behind this is that the mathematical model cannot divide jobs quantities. To handle this situation, pre-processing and post-processing approaches are developed by Excel VBA. Both of these approaches are tried to deal with unassigned jobs and some trade-offs are considered in terms of time and cost. The pre-processing approach divides jobs according to their machine hour before running the model. These divided jobs are added to the production plan with its same material id as a different job. On the other hand, the post-processing approach divides unassigned jobs according to their quantities after running the model. Toledo and Armentano (2011) propose a heuristic approach for distributing excess capacity. The idea is moving production from machine to machine. Those divided unassigned jobs are reassigned to the machines which have the available capacity.

After executing these approaches for different data sets, the run time of the mathematical model differs significantly as can be seen in the Appendix C. Additionally, in the pre-processing approach, the divided jobs are considered as different jobs by the system and unnecessary additional setup cost is added to the objective function. Thus, we decided to implement the post-processing algorithm in our system.

4. Validation

After the post-processing algorithm is implemented and attached to the mathematical model, some extreme cases are considered for the validation of the system.

In the first extreme case, all the quantities are multiplied by 100 to foresee the situation of the extreme increase in the quantities of the production plan. In this situation, as expected, the number of unassigned items and the total production cost increases accordingly.

In the second extreme case, all the quantities are divided by 100 and an extreme decrease in the production is considered. In this situation, as expected, unassigned items and the total production cost decreases.

The third extreme case is increasing the labor costs. If a machine needs less number of laborers to work, their utilization level increases when we increase the labor cost. This increase is also affected by the alternativeness relation of the machines. If there is an alternative machine, which needs less number of laborers, its utilization level remains the same. As some of the utilization levels increase, the total production cost increases for this extreme case.

The last extreme case is increasing the electricity costs. In this case, the utilization of the machines which consume more electricity decreases.

As a result, with all of these extreme cases, both the mathematical model and the post-processing algorithm are verified for the credibility of the system. After the verification of the mathematical model and heuristic approach, all of these cases are run again for the user form to test the applicability and sustainability of the overall system.

5. Integration and Implementation

In order to ensure the applicability of the project and to make the company benefit from the system in a convenient way, the design is made using Excel. The user interface of the system allows users to easily add data to the system, apply the mathematical model, and analyze the results. The system has a capacity of 12 machines and the input/output data will be in weekly terms as preferred by Arçelik Mechanic Shop.

On the start page, the system will take weekly inputs from the user which can be seen in Appendix D. After the inputs, the user can run the mathematical model. Also, six operations are provided in the system for mechanic shop workers. These operations are; allocation of the unassigned jobs, adding a new machine, removing a machine, assigning jobs to the particular machine, changing machine parameters, and changing production time for normal and overtime horizon. The list of operations is in Appendix E.

On each page of the system, there is a help button to help the user about processes. Also, the system provides reports about the unassigned jobs list, capacity utilization, and the remaining capacity of the machines, the total cost of the production plan, and assigned jobs list on each machine to deepen the analysis of the production plan.

6. Benefits to the Company

Through the proposed decision support system, six main benefits are offered to the company. These are; user-friendly interface, machine-job allocation, monitoring allocated/unassigned jobs, monitoring capacity usage, financial analysis, horizon flexibility.

The system is presented and utilized through the user interface. The user-friendly system provides a cost-effective and systematic machine-job allocation by using the weekly production plan. The decision support system provides reports for analysis and it also allows control of allocated and unassigned jobs. If it is needed, the user can assign the jobs manually before the model runs. Also, after the model run and unassigned jobs are shown through the user interface, these unassigned jobs can be analysed and some of them can be pre-assigned according to their allocation priorities.

Another important benefit of the system is the financial analysis. The system is flexible to be used in the case of new machinery investment or machine removal. Financial analysis can be done with the help of machine add/remove features. Cost values can be analysed by trying different machine add/remove scenarios. The analysis also can be made also by trying the forecasted demands. If a production plan is prepared according to the forecasted demands, this plan can be tried in the system and the cost results can be analysed regarding the plan.

A flexible and customizable system is aimed at this project. Therefore, the decision support system provides a horizon flexibility feature. With the help of this feature, the user can alter the lengths of both normal and overtime horizons. This is a very useful feature, especially in national and religious holidays. For example, in the case of holidays, the normal weekly horizons of 6 regular, 1 overtime days can change.

After proposed decision support is applied to the Arçelik CAP, the time spent for the manual assignments and calculations will be decreased considerably. Also, it is expected to observe a reduction in the calculation errors since the human factor is decreased with the automation of the system.

7. Conclusion

At the end of the project, a more organized, automated, cost-effective system is proposed to the company. The proposed system is run for the previous production plans from different months provided by the company. According to 18-week of data for different periods (January - February - August - September 2019), at most 24.7%, at least 1.72% improvement achieved based on the overall cost objectives. On average, 10.71% improvement on labour cost, 8.54% improvement on electricity cost, and 5.65% improvement on setup cost are achieved, and overall 8.11% on average improvement is obtained based on these 18-week data. The average cost improvement per 18-week data can be seen in Appendix F. Also, there is a trade-off between the time and cost of the project. Since overall production planning procedures run through Excel, the mathematical model is also solved by Excel OpenSolver for compatibility with the company. The system is run for one week's production plan and if the period increases to one month, the solution time increases to 35 minutes which is too much for the company. In contrast, a more advanced solver program like CPLEX saves the company's time and may give broader analysis. When the same monthly production plan, which was solved by Excel OpenSolver, solved by CPLEX, run time decreased to 4 minutes. Though OpenSolver is free to download and CPLEX may be a costly solution, there may be possible solvers not that expensive and can decrease solution at the same time. At this stage, the company should consider possible solver programs for this trade-off between cost, time and ability to broader analysis.

REFERENCES

Arçelik Corporate Overview, www.arcelikas.com/page/72/Corporate_Overview.

Toledo, Franklina Maria Bragion, and Vinícius Amaral Armentano. "A Lagrangian-Based Heuristic for the Capacitated Lot-Sizing Problem in Parallel Machines." *European Journal of Operational Research*, North-Holland, 13 Jan. 2011, www.sciencedirect.com/science/article/pii/S0377221705005321.

APPENDICES
APPENDIX A- Mathematical Model

Aggregate Model

Sets:
I: $1,2,\ldots,i,\ldots I$: *The job set*
J: $1,2,\ldots,j,\ldots J$: *The machine set*

Parameters:
m_{ij}: $\{1\ if\ job\ i\ can\ be\ manufactured\ by\ machine\ j;\ 0\ o.w.\}$
nh: $Normal\ horizon\ of\ production\ (6\ days \times 3\ shifts \times 7.5\ hour)$
oth: $Overtime\ horizon\ of\ production\ (1\ day \times 3\ shifts \times 7.5\ hour)$
Q_i: $Weekly\ demand\ of\ job\ i$
s_j: $The\ setup\ time\ incurred\ for\ the\ initialization\ of\ manufacturing\ of\ machine\ j\ (h)$
cs: $Cost\ per\ setup\ (TL)$
t_{ij}: $The\ time\ required\ to\ manufacture\ a\ unit\ of\ job\ i\ by\ machine\ j\ (h)$
e: $Electric\ cost\ (\frac{TL}{Kwh})$
c_j: $Electiricity\ consumption\ of\ machine\ j\ (Kwh)$
l: $Labor\ cost\ (\frac{TL}{h})$
n_j: $Number\ of\ required\ labors\ for\ machine\ j$
OEE_j: $Overall\ Equipment\ Effectiveness\ of\ machine\ j$
pcost: $Penalty\ cost\ for\ per\ backlogged\ job$

Decision Variables:
x_{ij}: $\{1\ if\ job\ i\ is\ allocated\ to\ machine\ j\ at\ normal\ horizon;\ 0\ o.w.\}$
$x_{o_{ij}}$: $\{1\ if\ job\ i\ is\ allocated\ to\ machine\ j\ at\ overtime\ horizon;\ 0\ o.w.\}$
b_i: $\{1\ if\ job\ i\ unassigned;\ 0\ o.w.\}$
sc_{nj}: $Setup\ count\ for\ machine\ j\ at\ normal\ horizon$
sc_{oj}: $Setup\ count\ for\ machine\ j\ at\ overtime\ horizon$

Mathematical Model:

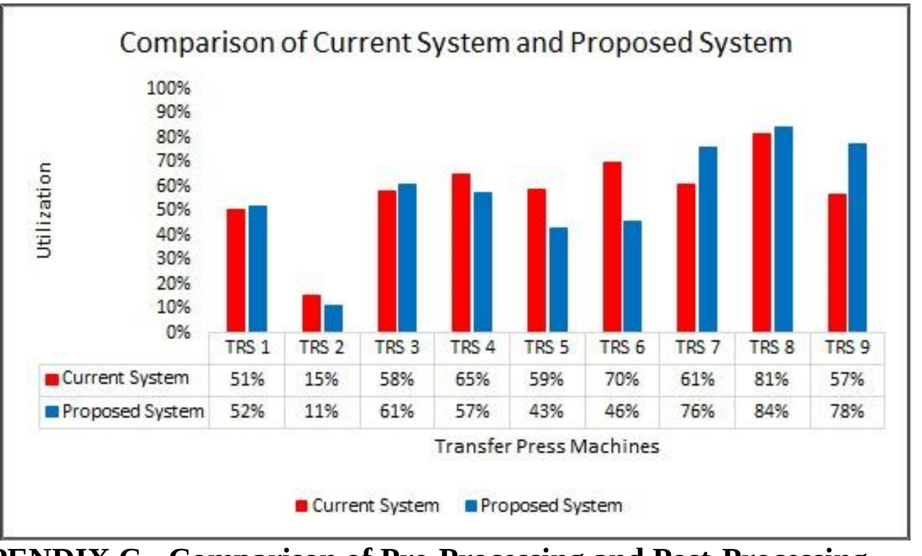
minimize

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\left(x_{ij}+x_{o_{ij}}\right)\times t_{ij}\times c_j\times e \qquad \textit{(Electricity Cost)}$$
$$+\sum_{i=1}^{I}\sum_{j=1}^{J}\left(x_{ij}\times l+x_{o_{ij}}\times l\times(1.5)\right)\times n_j\times t_{ij} \qquad \textit{(Labor Cost)}$$
$$+\sum_{j=1}^{J}(sc_{nj}+sc_{oj})\times cs \qquad \textit{(Setup Cost)}$$
$$+\sum_{i=1}^{I} b_i\times Q_i\times pcost \qquad \textit{(Cost of Unassigned Jobs)}$$

subject to

$$x_{ij} \leq m_{ij} \qquad \forall i \in I,\ j \in J \qquad (1)$$
$$x_{o_{ij}} \leq m_{ij} \qquad \forall i \in I,\ j \in J \qquad (2)$$
$$\textstyle\sum_{j=1}^{J} x_{ij} + x_{o_{ij}} + b_i = 1 \qquad \forall i \in I \qquad (3)$$
$$\textstyle\sum_{i=1}^{I} x_{ij} \times t_{ij} \times Q_i + sc_{nj} \times s_j \leq nh \times OEE_j \qquad \forall j \in J \qquad (4)$$
$$\textstyle\sum_{i=1}^{I} x_{o_{ij}} \times t_{ij} \times Q_i + sc_{oj} \times s_j \leq oth \times OEE_j \qquad \forall j \in J \qquad (5)$$
$$\textstyle\sum_{i=1}^{I} x_{ij} \leq sc_{nj} \qquad \forall j \in J \qquad (6)$$
$$\textstyle\sum_{i=1}^{I} x_{o_{ij}} \leq sc_{oj} \qquad \forall j \in J \qquad (7)$$
$$x_{ij}, x_{o_{ij}}, b_i \in \{0,1\}, sc_{nj}, sc_{oj} \geq 0 \text{ and Integer} \qquad \forall i \in I,\ j \in J \qquad (8)$$

APPENDIX B - Machine Utilization Rates



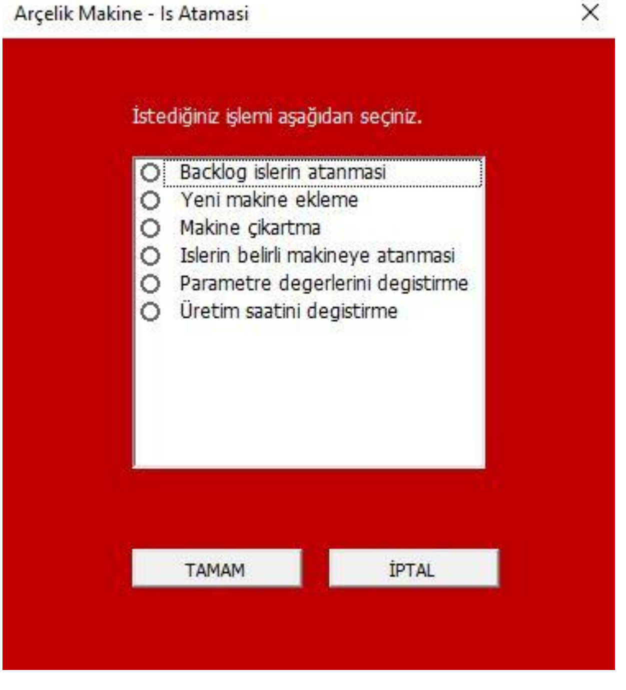
APPENDIX C - Comparison of Pre-Processing and Post-Processing Algorithms with Different Data Sets

	1st data set	2nd data set	3rd data set	4th data set
Number of Jobs	231	237	226	239
Run Time of Mathematical Model	06’ 45’’	07’ 23’’	07’ 36’’	07’ 20’’
Objective Function (TL)	222,089*	42,852*	66,824*	69,053*
Number of Jobs after Pre-Processing	1,086	1,054	1,118	1,100
Run Time of Math. Model with Pre-Processing (hour)	>> 2	>> 2	>> 2	>> 2
Run Time of Math. Model with Post-Processing	07’ 13’’	08’ 12’’	08’ 25’’	08’ 08’’

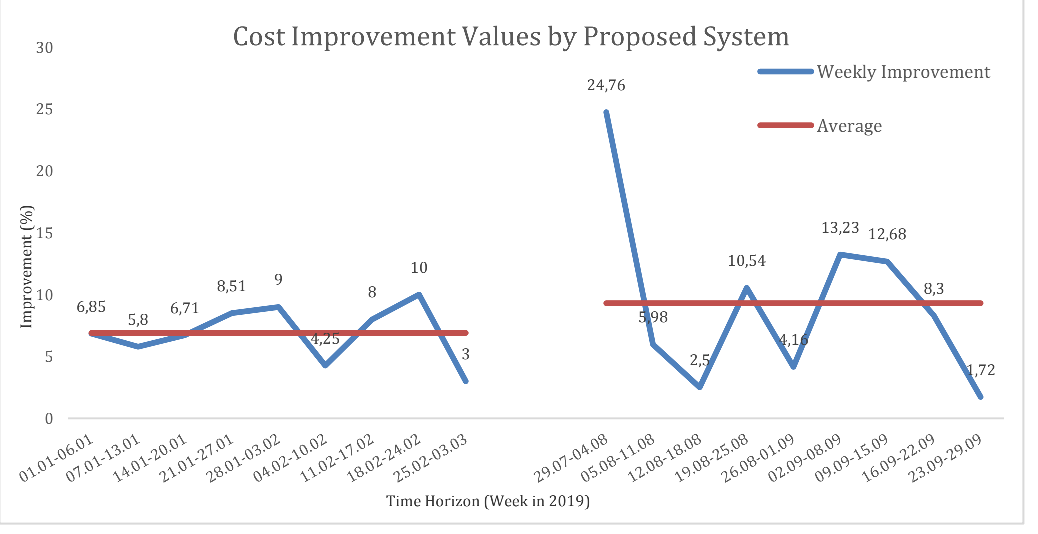
*Objective functions are indexed to the average production cost of 4 datasets according to company request.

APPENDIX D - Start Page of the Decision Support System

APPENDIX E - List of the Operations in Decision Support System

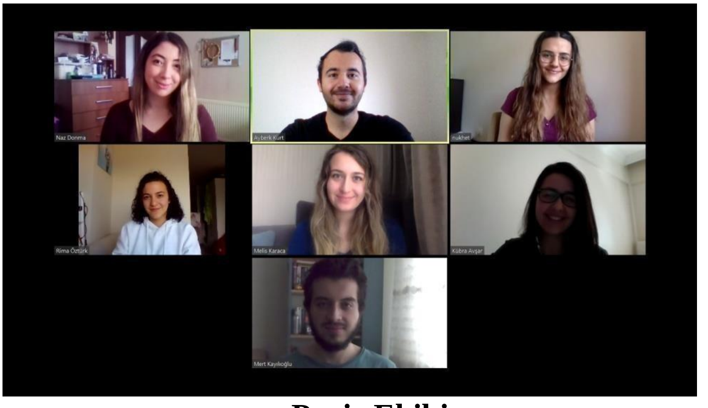


APPENDIX F – Average Cost Improvement per Week



Sürekli Üretimde Benzer Özellik Gösteren Karışımları Bularak Ön Uyarı Sistemi Oluşturma

Brisa A.Ş.



Proje Ekibi

Nükhet Akgün, Kübra Avşar, Fatma Naz Donma, Melis Karaca, Mert Kayılıoğlu, Ayberk Kurt, Rima Öztürk

Şirket Danışmanı

Mustafa Tacettin

Brisa Endüstri Mühendisliği ve Bilgi Sistemleri Müdürü

Akademik Danışman

Dr. Öğr. Üyesi. Çağın Ararat

Endüstri Mühendisliği Bölümü

ÖZET

Bu projede, lastik üretim sürecinde benzer davranışları tespit ederek karışması normalden uzun zaman alan bileşiklerle ilgili bir erken uyarı mekanizması oluşturulması hedeflenmektedir. Bu sayede üretim kapasitesinin uzun vadede%1 artırılması beklenmektedir. Şirketin genel bilgileri verildikten sonra mevcut üretim sistemi ve karıştırma prosedürü açıklanmaktadır. Mevcut sistemin analizi ile, benzer özellik gösteren karışımlar tespit edilmektedir. Elde edilen benzerlik kurallarına göre, karışım süresinde uzama olması öngörülen ürünler için mevcut reaktif uyarı sistemi yerine proaktif bir uyarı sistemi oluşturularak sistemden alınan uyarılara göre ilgili departmanın önlem alması sağlanmaktadır. Rapor, sorun ve olası çözüm yöntemleri hakkında literatür araştırması sonucunda bulunan çözümün nasıl uygulandığı, yeni sistemin getirdiği faydalar ve sistemin doğrulanması hakkında bilgi vermektedir.

Anahtar Kelimeler Erken Uyarı Sistemi, Apriori Algoritması, Veri Madenciliği, Sürekli Üretim

Creating a Proactive Warning System by Detecting Similar Behaviours of Compounds in the Continuous Production System

1. Introduction

Lassa Lastik Sanayi ve Ticaret A.Ş. was established in 1974 with the enterprise of Sabancı Holding and its partners. Brisa produces tires with a continuous manufacturing system. Materials that are stocked in the raw material warehouse follow these stages: mixer, stock preparation, tire building, curing, final finish, and they are sent to the tire warehouse. Tires are produced in mass which means that make-to-stock manufacturing is conducted.

1.1 Problem Definition

A proactive warning system is planned to be created by detecting the compounds behaving in a similar way. The current warning system informs the technology department each morning with the information of the previous working day about the long lasting mixing processes of compounds from the standard. Since the technology department is informed about these time prolonging of compounds of the previous day, they would not have sufficient time to take cautions. Moreover, this current system does not consider the possibility of similar behaviours among compounds. If a new system is created by identifying these similar compounds, when a time prolonging occurs in a specific product code, then the technology department would predict that the corresponding compound's mixing time could also prolong. For example, in case product A and product B behave similarly, if product A takes longer to mix, then the warning system would warn the corresponding department that mixing of product B could take longer.

1.2 Objectives and Scope of the Project

Since a proactive warning system gives the company time to intervene in compounds that may experience time prolonging in mixing time, the decrease in the number of compounds which excess the standard mixing time, and the cycle time in production will be considered as the main performance measures. Our deliverables for the project are to develop the warning system, integrate all the common compound sets into the system and verify the model of the system in the company. In case of detection of compounds exceeding the standard time, the system is expected to issue a warning for other compounds having similar properties with this compound in the production plan.

2. Model Development

2.1Analysis of Data

In the factory, two mixers perform mixing procedure starting from the entrance of the raw materials and ending with the mixed compounds to be sent to the cooling zone for further usage in tire forming steps. Dumping process changes based on the compounds' features that are identified by their product codes, version numbers, and issue numbers.

Version number changes when the prescription of compounds changes whereas issue number changes when the value of a parameter such as temperature or time changes. Together with version and issue number, a unique code for each product is used in the classification of products. This classification aids finding the minimum and maximum temperature and time levels that compounds should be dumped which refers to the ending of the mixing procedure of compounds. These threshold values are used to determine the behaviour of compounds. If the compounds are dumped within the maximum and minimum levels, then this process is called as a *normal dumping process* whereas if they are dumped before they reach their minimum temperature or minimum time, this process is called as an *abnormal dumping process.*

2.2 Apriori Algorithm

To detect similarities among compounds, association rule mining, which is a concept of machine learning, is used. As one of the sub-categories of association rule mining, the apriori algorithm is the main algorithm to perform in this project since it focuses on identifying the relationship between data sets. In our project, the data set would be the compounds that are being mixed, and after running the algorithm we would expect to find similarity rules. There are three measures of association rule mining that are used in similarity detection. (Han, Kamber and Pei, 2012). These are *support*, *confidence*, and *lift*.

The *support* is the portion of instances that contain all items, compounds in our case, in the antecedent and consequent parts of the rule (Han, Kamber and Pei, 2012):

The *confidence*, on the other hand, can be defined as the probability of finding

$$Support(X \Rightarrow Y) = P(X \cap Y)$$

the consequent part of the rule in instances under the condition that these

$$Confidence(X \Rightarrow Y) = P(Y|X) = \frac{P(X \cap Y)}{P(X)}$$

instances also include the antecedent part (Han, Kamber and Pei, 2012):

As the last measure, *lift* gives information about the increase in the probability of the consequent part given the antecedent part (Han, Kamber and Pei, 2012):

A meaningful rule should have a lift value greater than one which means when the consequent part occurs, it is more likely that the antecedent occurs

$$Lift(X \Rightarrow Y) = \frac{Confidence(X \Rightarrow Y)}{P(X)} = \frac{P(X \cap Y)}{P(X).P(Y)}$$

(positive association). Association rules are required to satisfy a user-specified minimum support and a user-specified minimum confidence at the same time (Han, Kamber and Pei, 2012).

2.3 Time, Temperature and Abnormality Buckets

To detect the compounds with similar characteristics, as the first step of association rule mining, compounds are clustered into time-based buckets which consist of 5-day intervals. For temperature and time parameters, buckets are formed by taking the average of dumping times or dumping temperatures of a specific product for 5- day periods. For abnormality, buckets are formed with binary variables indicating whether the compound is dumped abnormally or not. The aim is to identify the similarity in the products by analysing them in terms of the increase or decrease in time/temperature, and the existence of the abnormality between consecutive buckets. In this way, while conducting the apriori algorithm, a correlation can be determined precisely according to the change in the behaviour of compounds shifting from one bucket to another.

To implement the bucket formation process into a code, the R language is used. Before reading the data in our code, outliers are eliminated to improve clarity in the project result by using Excel. Outlier data are eliminated by choosing a threshold value of 480 minutes corresponding to one shift. If the mixing time of a process is longer than the threshold value, it is identified as an outlier and eliminated.

2.4 Association Rule Mining

After the buckets are formed, to perform association rule mining, products are clustered according to the increase or decrease in the corresponding parameter between consecutive buckets. For instance, to analyse the increase in the time parameter, at first, a transactional data matrix is formed. In this matrix, see Appendix A, the products are placed by analysing how they increase through time. In Appendix A, products under column 12 correspond to the products whose collect time differences increase when going from the first bucket to the second bucket. The matrix is formed by using the same logic. By observing this increase, we expect to find a pattern of the compounds to detect the similarities. This methodology is repeated for the temperature and abnormality parameters.

As a next step of the algorithm, frequencies of each item in the matrix are analysed and items are eliminated by comparing those frequencies with a pre-defined minimum support value. This minimum support value is used to eliminate the compounds who have lower frequency than this value. To choose an appropriate support threshold, frequency bar plots are used. For instance, Appendix B is an example of a frequency bar plot for analysing the time increment.

According to Appendix B, the support parameter of 0.1 is chosen to be used for running the algorithm, so that the algorithm searches rules among products that have an equal or larger relative frequency higher than 0.1. When the support parameter is chosen greater than 0.1 such as 0.15, the algorithm eliminates more products to search rules so that the number of rules decreases whereas if it is lower than 0.1, then the number of rules increases.

Once the support level is determined, to run the algorithm, the confidence level is entered, and the similarity rules are detected by using the same logic for each time/temperature incremental/decremental transaction matrices. However, abnormality-based rules are formed by looking at whether the product performs abnormal behaviour or not within the given time buckets. Since we have five different types of rules such as incremental time, decremental time, incremental temperature, decremental temperature, and abnormality; and the number of rules found diverges for each, how many of the
output rules to use should be determined. Therefore, rules are ranked by their lift values, and the first twenty rules are taken.

3. Implementation

To create a warning system, the Excel data are analysed which was provided by the company as an input. In this file, mixing times and mixing temperatures of compounds can be observed. After eliminating outliers from this data, time and temperature-based time buckets are created, and association rule mining is applied by RStudio. With the help of this process, the apriori algorithm is applied in one step. The compounds that behave similarly are obtained with this algorithm. After detecting similarities, these outputs are transported to the system of the company by using SQL database. The reason why we used RStudio and SQL is that the company database system is supported by Oracle and these programming languages conform with the company's database. In the implementation process, parameters of apriori algorithm are fixed with experimental design. Taguchi's orthogonal arrays from experimental design are used to decide which alternative combination leads to the best output results in fixing the parameters process. Different possible parameter values were tried along with two different levels for each parameter. Since there are three different parameters to be fixed, confidence, support, and bucket size, and two different levels were used, $L_4(2^3)$ orthogonal array was chosen.

The rules obtained by the R code are recorded to the database and it will be updated as new inputs enter the system. When there is a time-prolonging of a compound, SQL system draws the similar compounds from the list and sends an e-mail to the related department. These e-mails are daily and the related department is warned before the production starts. Example of these e-mails can be seen in the Appendix C.

4. Validation

After the implementation of the warning system to Brisa's database, daily emails are sent to the production department in order to warn the possible delays in the mixing time of compounds. From the 27th of February until the 26th of March, those emails were collected, including warnings, and analysed them to see how accurate the proactive warnings work. For this one-month period, 88 warnings were received in total. For each warning, three conditions are specified as right, wrong, or inconsiderable for the validation. After the warning came for a specific compound, the first mixed date of that compound was found and checked whether the mixing time of that compound specifies these conditions. If there was a delay in the mixing time of the compound, the warning corresponding to that specific compound was labelled as right. If there was not any delay, the warning for that specific compound was labelled as wrong.

In some cases, the same warnings were received in the following days. As it can be seen in the Table 1, the warning for compound 7, based on compound 6, was received both in 12/03/2020 and 15/03/2020. If compound 7 was not mixed between 12/03/2020 and 15/03/2020, the situation of the warning sent in March 12 was specified as inconsiderable. Therefore, the validation is continued with the warning received in March 15.

Table 1: Sample e-mail warnings

Date	**Prolonged**	**Expected to Prolong**
12.Mar	Compound 6	Compound 7
15.Mar	Compound 6	Compound 7

After applying the same procedure; right, wrong, or inconsiderable conditions were specified for all 88 warnings in the one-month implementation period, the number of compounds whose mixing time prolonged was 48. The number of compounds whose mixing time did not prolong was 3. The number of compounds which were not mixed until the next same warning was received was 37. Therefore, the accuracy of the warning system is %94 since 48 warnings were true out of 51 considerable warnings.

5. Benefits and Improvements

Benefits of the new system can be thought as long term and short-term benefits. The main goal of the project is a 1% increase in the production capacity. Although it is possible to observe the capacity increase only in the long term, short-term benefits can be observed to measure the improvements of the project.

The most obvious measurement would be the number of warnings after implementing this system. For this reason, a measurement to be made in the number of warnings received has a great importance in terms of observing the improvement provided by the project. As an expectation, this number will decrease day by day with the help of the proactive warning system. The reason is that when a warning is sent for a specific compound and the precaution is taken, then the possible problems would be solved for the similar compounds. Since possible time-prolonging is prevented, new warnings would not be sent for these similar compounds.

To understand the improvement of the system, it was assumed that the effect of a warning lasts three days. Accordingly, in case of an X→Y warning ,the time prolonging of the product Y was observed for the following three days. If the necessary precautions were taken considering the X→Y warning, no Y→Z warning would be received for the following three days. Since hte possible time prolonging foreseen for the product Y would be prevented with the help of precautions taken, and no warning would be received that stems from product Y.

With this measurement method, warnings that would not be seen in the system were detected in the case of taking precautions and it was assumed that those warnings would not come out and were removed from the list of received warnings. In this case, the total number of warnings in the list would reduce from 88 to 45, as it can be seen in Appendix D. It was determined that the system would provide 51% improvement if all necessary precautions are taken.

6. Conclusion

At the beginning of the project, the expectation of the company was to create a warning system to detect the similar compounds and take precautions according to these results in order to increase the capacity. With the help of the Excel and RStudio, five different lists of the similarity rules, which are time increment, time decrement, temperature increment, temperature decrement and abnormality rules, were detected. After the visit to the company in Aksaray, the comparison between rules and the real data was done and the time increment list was decided to be used. The system, which gives the rules as output, was integrated into the system of the company with the help of our industrial advisor. Warnings were started to be sent to the related departments by an e-mail.

With the new warning system, the validation is approximately 94% and 51% improvement is expected to be provided. According to the meetings with the industrial advisor, it can be concluded that the project has met the expectations of the company.

6.1 Future Work Suggestions

The warning system model finally proposes the implementation of cross- rule prioritization found. According to this suggestion, based on the previous three days, the historical warnings of the two mixtures are taken into consideration and the system gives a warning. As seen in Appendix E, the system has detected that products 8 and 2 behave similarly in the last three days and recommended that precautions should be taken as a priority. According to e-mail, if the relationship of mixtures has been observed in three times in the last three days, the warning is red and very high priority. The relationship is yellow and medium priority if it has been observed once or twice, and blue if it has never been observed, and it is of low priority.

REFERENCES

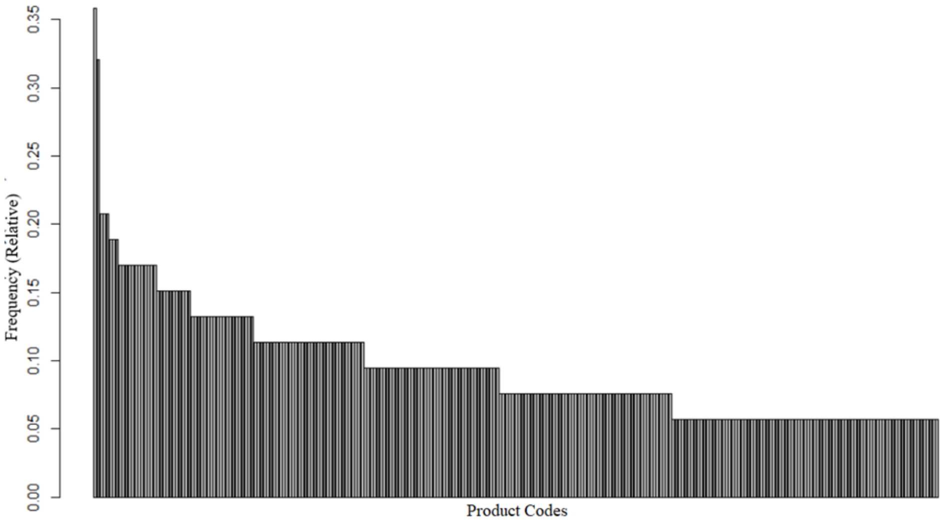
Han, J., Kamber, M., and Pei, J. (2012). Data mining: concepts and techniques. Amsterdam: Elsevier

APPENDICIES

Appendix A: Compounds Increasing Between Two Consecutive Buckets

From 1st to 2nd Bucket: 12	**From 2nd to 3rd Bucket: 23**	**From 3rd to 4th Bucket: 34**	**From 4th to 5th Bucket: 45**	**From 5th to 6th Bucket: 56**
Product 1	Product 2	Product 2	Product 8	Product 9
Product 14	Product 5	Product 4	Product 13	Product 18
Product 15	Product 6	Product 9	Product 14	Product 25
Product 16	Product 7	Product 12	Product 15	Product 56
Product 17	Product 9	Product 32	Product 16	Product 61
Product 18	Product 13	Product 39	Product 2	Product 63
Product 21	Product 14	Product 45	Product 22	Product 64
Product 22	Product 19	Product 55	Product 23	Product 81
Product 23	Product 26	Product 61	Product 24	Product 82
Product 24	Product 29	Product 62	Product 26	Product 89
Product 25	Product 31	Product 83	Product 29	Product 94
Product 27	Product 33	Product 84	Product 52	
Product 34	Product 36	Product 85	Product 57	
Product 35	Product 41	Product 86	Product 62	
Product 39	Product 51	Product 87	Product 77	
Product 44	Product 69	Product 88	Product 9	
Product 45	Product 7	Product 89	Product 91	
Product 46	Product 71		Product 92	
Product 47	Product 72		Product 93	

Appendix B: Item Frequency Bar Plot for Time Increment

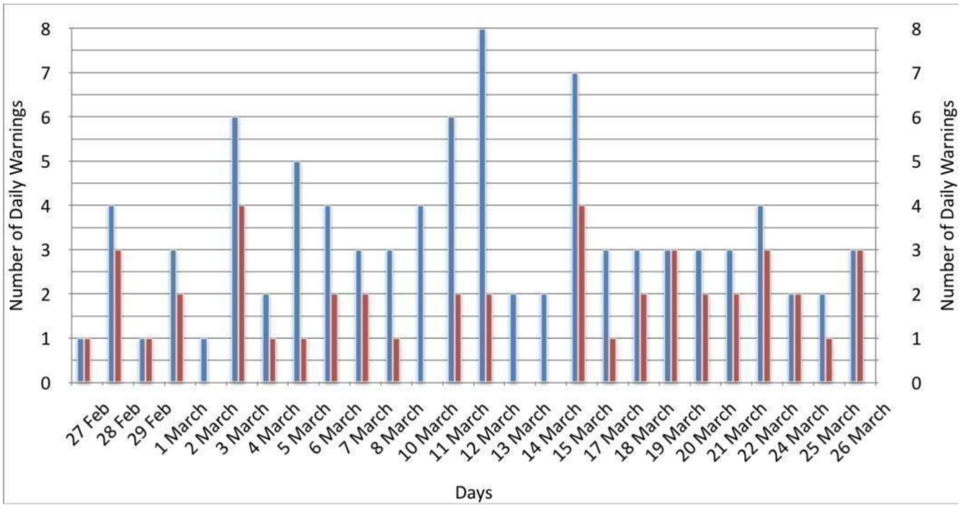


Appendix C: A Sample E-mail Warning

From: <sa-email-shopfloor@brisa.com.tr>
Date: Fri, Mar 6, 2020 at 8:01 AM
Subject: Benzer Karisimlar
To: < @brisa.com.tr>, < @brisa.com.tr>, < @gmail.com>

Mikser Problem Potansiyeli Olan Kod Listesi		
;	Uzayan Karışımlar	Uzaması Beklenen Karışımlar
1	[illegible] Product 1	[illegible] Product 2
2	[illegible] Product 3	[illegible] Product 4
3	[illegible]	[illegible]
4	[illegible]	[illegible]

Appendix D: Daily Theoretical Number of Warnings with Improvements



Appendix E: A Sample E-mail with Priority Labels

sa-email-shopfloor@brisa.com.tr

to s.daglioglu, M.TACETTIN, me

Turkish > English Translate message

Mikser Problem Potansiyeli Olan Kod Listesi		
;	**Karisim1**	**Karisim2**
1	Product 8	Product 2
2	Product 9	Product 10
3	Product 3	Product 4
4	Product 5	Product 2
5	Product 6	Product 7
6	Product 2	Product 8
7	Product 1	Product 2

Kontrol Panel Hazırlama Bölümü İş Çizilgeleme Karar Destek Sistemi

B/S/H Ev Aletleri Sanayi ve Ticaret A.Ş.

Proje Ekibi
Ali Eren Sürmeli, Bartu Özcan, Mustafa Emin Tos, Ömer Doğan, Pınar Dökücü, Sebiha Rozerin Alp, Tuna Darcan

Şirket Danışmanı
Onur Durak,
Üretim Mühendisliği

Akademik Danışman
Prof. Dr. Mustafa Çelebi Pınar,
Endüstri Mühendisliği Bölümü

ÖZET

Bu projenin amacı, BSH Bulaşık Makinesi Fabrikası içerisindeki Tasarım Atölyesi'nde üretilen kontrol panelleri için iş çizelgelemesi yapan bir karar destek sistemi geliştirilmesidir. Mevcut sistemde çizelgelemenin el yordamıyla yapılmasından kaynaklı servis seviyelerin düşüş ve ana hat üretiminin durması problemleri bulunmaktadır. Problemin büyüklüğü göz önünde bulundurularak matematiksel modelin kabul edilebilir süre içerisinde sonuç vermemesinden dolayı sezgisel model geliştirilmiştir. Problemimiz için en uygun sezgisel yöntem olan Genetik Algoritma kullanılmıştır. Geçmiş veriler üzerinden şirkete danışılarak sistemin doğrulaması gerçekleştirilmiştir, %51,7 iyileştirme sağlanmıştır. Sezgisel model kullanıcı dostu arayüze entegre edilmiştir.

Anahtar Kelimeler: Esnek Akış Tipi Çizelgeleme, Genetik Algoritma, Karar Destek Aracı

Control Panel Decision Support System for Scheduling

1. General Information

The majority of the company which operates in Turkey is owned by BSH Hausgeräte GmbH, a very small share of the company belongs to other partners. BSH Ev Aletleri San. ve Tic. AŞ. announces that %99,96 of total share of the company is owned by BSH Hausgeräte GmbH and the remaining %0,04 belongs to other shareholders. At the BSH's Çerkezköy location has 6 different factories including Cooling 1, Cooling 2, Cooking, Laundry, Dishwasher and Plastic. Our project will be carried out with the dishwasher factory. BSH Çerkezköy Plant employs around 8,000 people, 2086 of which are white collar and 5758 of which are blue collar.

2. System Analysis & Problem Definition

2.1 System Analysis

In Design Shop, there are three stages. First one is Printing Process which has four machines. Second stage is Milling Process which has three machines and third stage is Press Process which has one machine. Each machine in each stage work parallel and they are not identical because of their different technologies. There are 452 variants of control panels which are produced in Design Shop. Each job can be grouped as family based on their model; each model has different process time whereas each family have different set up time. In set up time, transition between families is considered. 95% of production process takes place in stage 1 and 2 where stage 3 only consists of 5% of production. By the request of the company, we will not consider improving stage 3, only working on stage 1 and 2. BSH currently schedules these control panels manually and they demand an automatic decision support tool.

2.2 Problem Definition, Scope and Needs Assessment

BSH does not have an automatic system for scheduling production of control panels. Foreman of the workers receives three days demand from main production line, and he arranges sequence of production accordingly. This manual system has deficiencies such as dependency on human experience and open to human error. BSH indicates that absence of foreman causes confusion in scheduling and design shop team cannot finish their duties. It also creates mismatching in main production line. All jobs have different set-up and process time. Manual scheduling cannot arrange best schedule and interrupt demand responsiveness. These time differences can confuse operators and manual scheduling becomes harder. As a result, delay of producing control panels occurs or producing panels before due date causes unnecessary stock. These deficiencies cause two issues:

1. Decreasing in service level
2. Increasing over time cost

The solution for releasing this system from human dependency is developing an automatic system. For this reason, our project aims to provide a decision support tool which will give feasible solution for three days demand in Design Shop area. This tool will also consider adding new machines to the system. When an admin user defines a new machine to the system, it will give a production schedule considering new machine.

3. Solution Approach

In this project, major constraint was determined sequence of jobs. Some models have jobs that should go through a specific order. As an example, in one model, panel must go milling and then printing because its milled holes must be printed too.

Critical Assumptions: Due to company demand, OEE rate (Overall Equipment Effectiveness) should be use rather than using setup time. OEE consists availability, performance, quality. Setup time takes place in performance criteria. So, using OEE covers more and embraces a lot of scenarios such as power loss. Program does not include a distinction between shifts or days. It depends on the number of jobs that have been assigned. As an example, it schedules as many jobs that have been assigned to the input system. Based on data, it is observed that largest workload of three days was 96. So, system is limited by 100 jobs, but it can be changeable by the request of the company. System can increase the capacity of machines and processes by adding manually. It understands files in a specific format and automatically identifies the number of machines, jobs and process. Currently maximum number of machines is 7 whereas maximum number of process is 3.

3.1 Mathematical Model

As a starting point, we adapted our project to the “hybrid flexible flow shop scheduling problem” by R. Ruiz, F. S. Şerifoğlu, and T. Urlings (2008). Initial model includes sequence-dependent setup times, eligible machine sets and precedence relationships as these constraints were realized during the factory visit. Then with several visits to company and communication with industrial advisor, we finalized our mixed-integer model. Instead of sequence-dependent setup times, we gathered the information on job families and we used family-dependent setup times since setup times differ between job families, not between each job. Along with the family-dependent setup times, families of the jobs are included.

Mathematical model uses start times, finish times and completion times of the jobs as variables to determine the maximum completion time of jobs, in other words makespan. Also, for each position in machines in the schedule, we used a binary variable to determine whether a specific job is processed at that position of a machine. Processing times, setup times and precedence relationships were used as parameters in our mathematical model. Objective

function is defined as minimizing the makespan. Completion time of a job is determined by start time and finish time. Start time and finish time of a job is determined by processing time and family-dependent setup time. (Appendix 1)
IBM ILOG CPLEX is used to solve the problem. However, since it is a commercial software that needs to be purchased by the company and it does not give desirable solutions in an acceptable time due to the complexity of the problem, we decided to develop a heuristic approach and not to use IBM ILOG CPLEX in this project.

3.2 Heuristic Approach

Mixed-integer programming did not give optimal results in reasonable time. So, metaheuristics such as simulated annealing, tabu search and genetic algorithm were tried. With the help of literature review we discovered that genetic algorithm gives effective solutions by examining the article of F. Pezzellaa, G. Morgantia, and G. Ciaschettib (2007). Genetic algorithm enables a decision system in complicated situations stands on changes in human genes. At this point, complexity of scheduling that cannot be answered with using mathematical model became suitable for genetic algorithm. We started adapting our model to genetic algorithm. For coding process, we evaluated alternative coding languages and made trials. Firstly, we started to code with Java, then realized it would be much easier for us to use C# library packages. In addition, the ease of applicability and designing on Visual Studio was crucial about progressing with C#. With the help of design interface of Visual Studio, we created a user-friendly interface.

3.3 Designing Algorithm

Genetic algorithm relies on chromosomes which are main elements of creating procedures of scheduling. Machine and job schedules are chromosomes for our problem. Chromosomes have a length multiplication of number of jobs and number of processes. Machine schedule demonstrates the machines in which the processes are processed. An example of machine schedule is below.

```
macSchedule:

1 2 2 2 0 0 1 2 0 1
```

Figure 1. Example of Machine Schedule

Here, since we have 2 processes, first position implies the first process of the first job is processed in machine 2, second process of the first job is processed in machine 3, first process of the second job performed in machine 3, etc.

Job schedule is a little bit more complicated. It works together with the machine schedule.

```
jobSchedule:

4 2 1 1 0 2 0 3 4 3
```

Figure 2. Example of Job Schedule

We have a class named Schedule. While initializing a Schedule class variable, we declare a job schedule, machine schedule and number of machines, jobs and processes. We can initialize manually or by importing, it will initialize automatically. Collection of these schedules creates our population. For each schedule, we calculate makespan and idle times. Makespan of each schedule gives fitness score, which is the indicator in this algorithm. Lowest fitness score is the best fit value.

First, data in .xml format should be imported. Therefore, program automatically creates a matrix with processing times of the jobs. Then, initial population of schedules is created. As mentioned before, each schedule consists of job and machine schedules. Length of each chromosome is multiplication of number of jobs and processes. After that, genetic algorithm iterations start.

First phase is selection. We decided to use tournament selection method because tournament selection performs better than roulette wheel selection according to J. Zhong, X. Hu, M. Gu and J. Zhang (2005). Tournament selection is performed as following. We randomly pick five schedules and sort in ascending order. Then, we select two schedules with the lowest fit values as mother and father schedules.

After the selection of mother and father, we move on to crossover phase. Crossover is performed uniformly to eliminate illegal offspring. Single point and two-point crossover create illegal offspring. First, we create clones of mother and father. Then, for each position in job and machine schedule of mother, algorithm randomly change the value with the same position of father and eventually, it creates a child. Then, we do the reverse as father-mother to create another child.

Then, we perform exchange values mutation. For each child, we take two randomly selected positions in machine and job schedules and exchange the values of these two to create a schedule.

After the mutation phase, these two children go into the population. Since the population is updated, we check whether the best fitness score is updated. If these children schedules have a better fitness score than the current best, we update the best score. If none of the children schedules update the current best score for three hundred thousand iterations, algorithm refreshes last one million schedules. Genetic algorithm runs until the population reaches ten million chromosomes.

3.4. Decision Support System

Our support system consists of two main stages. The first stage is a MS Excel file runs a VBA code to process demand data that is input in line with the database which consists processing times and information about specific models that required certain flow. Output of this file is a XML file which is required format for our support tool.

In the user interface, the section that is planned to use most frequently is "Data Retrieval". When "Import XML" is clicked, it will retrieve the processing times of the jobs that will be processed. The location of the .xml file can be changed by clicking "…" button. After the process is finished, "Export as XML" is used to export.

If the processing times matrix is needed to be changed manually, then "Specify Details Manually" section should be used. Number of jobs, processes and machines can be changed manually with the limit on jobs is 100, process is 3 and machine is 7. When "Create" button is clicked, empty matrix is created. Values can be entered or changed on that matrix.

After the matrix is ready, "Run" button should be clicked. After pressing the button, matrix disappears, and a Gantt chart appears below the matrix. On the bottom, it gives the current makespan, total idle time, current progress percentage and the time passed until the current best solution is found. Process can be stopped by clicking "Stop" anytime or it stops automatically when the progress is 100%. You may see picture of our decision support tool in Appendix 3.

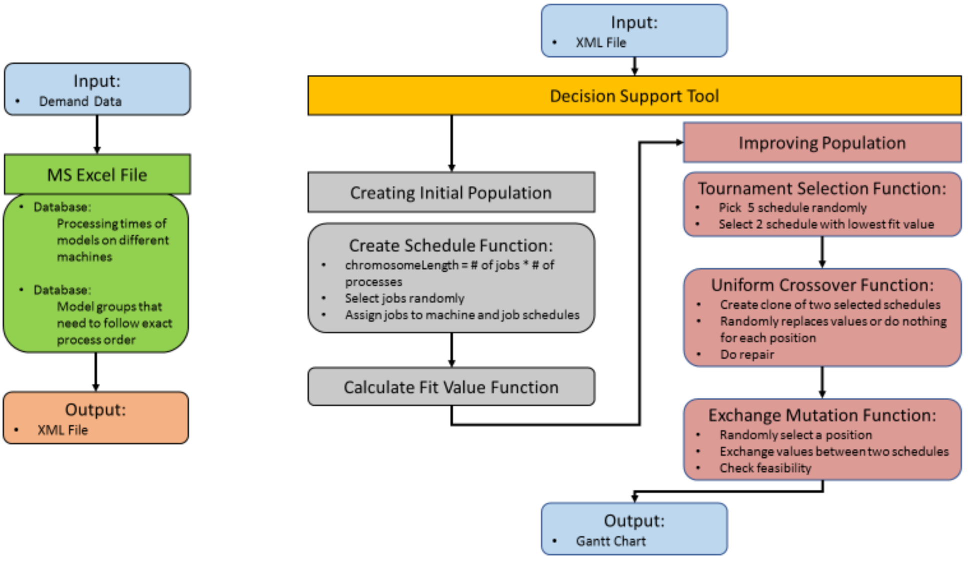


Figure 3. Working Principible of Decision Support Tool

4. Validation

We choose the busiest three days demand data in order to test. This was 7,8 and 9 November and had a total of 96 demand. We multiplied batch size with time and import these values to our program for every machine. It took 10 minutes for us to receive results. Program reached best solution in 4th minute. The makespan value was 3455 minutes which is 57.5 hours. We divide this value with 0.8 because of OEE value determined by the company, we got 71.9 hours. Company works three shifts, so our value just covers the 72 hours' time limit. Therefore, validation is executed. This result shows that our program can manage most crowded demand in desirable time. (see in Appendix 2)

5. Company Benefits and Progress Plan

5.1 Company Benefits

Decision support tool will provide an automatic scheduling system for Design Shop area. When the company advisor enters the demand, the tool will give a production schedule in minutes. Therefore, there will be shorter planning periods. Besides, as there will be a certain production schedule, workers will not need to leave their workstations and check the next job. These will increase the time efficiency in the system. As there will be a particular production schedule, there will be no need to foreman of workers and any problem caused by absence of this worker will be eliminated, as well. This tool will also consider adding new machines to the system. When an admin user defines a new machine to the system, it will give a production schedule considering new machines.

Using 4-9th May Demand Data, a detailed comparison is made between decision support tool and current system. Decision support tool gives a solution in 12 minutes for 17588 products. Current system finishes planned products in 17 shifts which corresponds to 8155 minutes. Decision support tool finds a solution with makespan of 3935 minutes. Based on these results, we observed 51,7 % improvement.

5.2 Implementation

Due to directions of BSH, user friendly, easy to learn and future oriented decision support system is generated. It is designed as responsive to increase in model portfolio and increase in production capacity which require adding new machine to the system. When we consider features of the decision support tool, this automatic system can be used in Design Shop area instead of current manual scheduling.

Our system contains two main documents; an Excel which contains VBA code and Genetic Algorithm based .exe file which coded in C#. It is not feasible for us to implement decision support tool in the factory because of COVID-19. So, system will be delivered to Company Advisor including user guideline in 27th May for real time testing. After trial, final product will be delivered according to feedback of Company Advisor.

REFERENCES

M. Pinedo, Scheduling: theory, algoritms, and system. Upper Saddle River: Prentice Hall, 2002.

R. Ruiz, F. S. Şerifoğlu, and T. Urlings, "Modeling realistic hybrid flexible flowshop scheduling problems," Computers & Operations Research, vol. 35, no. 4, pp. 1151–1175, 2008.

E. Akyol, T. Saraç, "A Mix Integer Programming Model for Parallel Machine Scheduling Problem: Using Shared Resource (in Turkish)" Gazi Üniversitesi Fen Bilimleri Dergisi, Part C, pp.109-126, 2017.
F. Pezzella, G. Morgantia, and G. Ciaschettib, "A genetic algorithm for the Flexible Job-shop Scheduling Problem," Computers & Operations Research, pp. 3202–3212, Mar. 2007
J. Zhong, X. Hu, M. Gu and J. Zhang, *Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms*, 2005.

APPENDIX

Appendix 1. Mixed-Integer Model

Parameters:

P_{jlk}: *Processing time of job j at machine l at stage k*

h_{jlk}: *Setup time of the first job j at machine l at stage k*

S_{ablk}: *Sequence dependent setup time for jobs in Family b where previous job is in Family a at machine l at stage k*

F_j: *Family of job j*

E_j: *Eligible machine set of job j*

$Prec_k$: *Jobs that should be processed at stage k before other stages*

Decision Variables:

$$X_{jylk}: \begin{cases} 1, & \text{if job } j \text{ is processed at sequence position } y \text{ at machine } l \text{ at stage } k \\ 0, & \text{otherwise} \end{cases}$$

C_{jk}: *Completion time of job j at stage k*

$start_{jk}$: *Starting time of job j at stage k*

C_{max}: *Maximum completion time (makespan)*

Model:

$$\min \quad C_{max}$$

$$s.t. \quad C_{jk} + M * \left(1 - X_{jylk}\right) \geq h_{jlk} + P_{jlk}, \qquad \forall\, i, k, y = 1, l \in E_j$$

$$C_{jk} - M * \left(1 - X_{jylk}\right) \leq h_{jlk} + P_{jlk}, \qquad \forall\, i, k, y = 1, l \in E_j$$

$$C_{jk} + M * \left(2 - X_{jylk} - X_{i(y-1)lk}\right) \geq C_{ik} + P_{jlk} + S_{ablk}, \qquad \forall\, i, k, y > 1, l \in E_i \cap E_j, \quad i \neq j, a \in F_i, b \in F_j$$

$$C_{jk} - M * \left(2 - X_{jylk} - X_{i(y-1)lk}\right) \leq C_{ik} + P_{jlk} + S_{ablk}, \qquad \forall\, i, k, y > 1, l \in E_i \cap E_j, \quad i \neq j, a \in F_i, b \in F_j$$

$$C_{max} \geq C_{jk}, \forall j, k$$

$$\sum_j X_{jylk} \leq 1, \forall y, k, l \in E_j$$

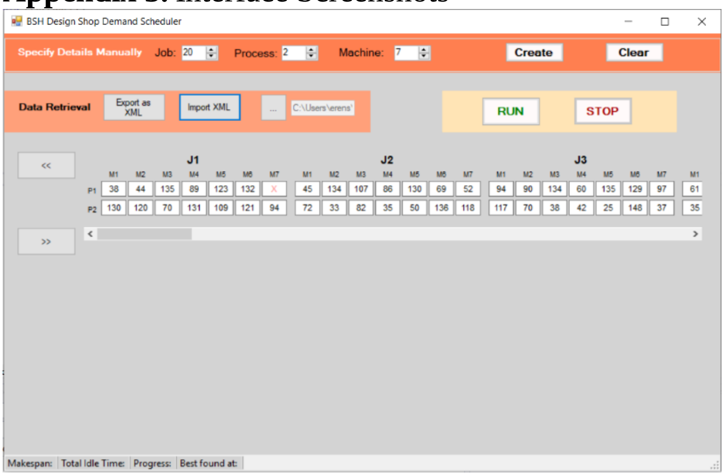
$$\sum_{y}\sum_{l\in E_j} X_{jylk} = 1, \forall j,k$$
$$\sum_{j} X_{jylk} - \sum_{i} X_{i(y-1)lk} \leq 0\ , \forall i,j,y,k\ , y > 1, l \in E_i \cap E_j, i \neq j$$

$$start_{jk} \geq C_{ik} - M * \left(2 - X_{jylk} - X_{i(y-1)lk}\right), \forall i,j,y,k,y > 1, l \in E_i \cap E_j, i \neq j$$

$$start_{jk} \leq C_{ik} + M * (2 - X_{jylk} - X_{i(y-1)lk}), \forall i,j,y,k,y > 1, l \in E_i \cap E_j, i \neq j$$
$$start_{jk} \leq M * \left(1 - X_{jylk}\right), \forall i,j,y,k,y = 1, l \in E_j$$
$$start_{jk} \geq -M * \left(1 - X_{jylk}\right), \forall i,j,y,k,y = 1, l \in E_j$$
$$start_{j2} \geq C_{j1}, j \in Prec_1$$
$$start_{j1} \geq C_{j2}, j \in Prec_2$$
$$C_{jk} \geq 0, \forall j,k$$
$$start_{jk} \geq 0, \forall j,k$$

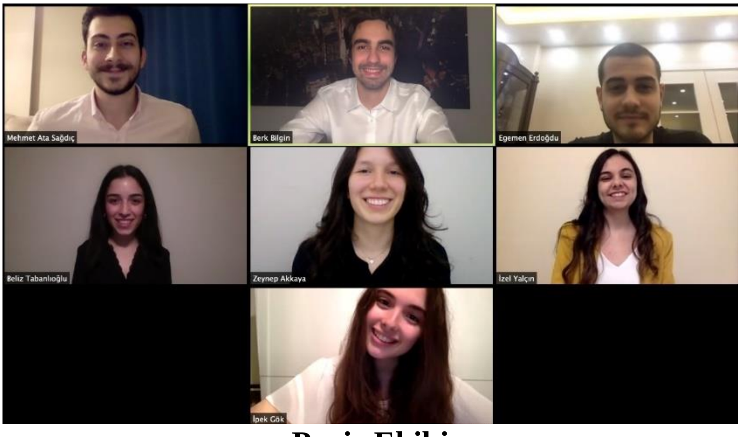
Appendix 2. Real Data Test Screenshots

Appendix 3. Interface Screenshots



Taktiksel Sevkiyat Lojistiği Planlaması

DHL Lider Lojistik Ortağı



Proje Ekibi

Zeynep Akkaya, Berk Bilgin, Egemen Erdoğdu, İpek Gök,
Mehmet Ata Sağdıç, Beliz Tabanlıoğlu, İzel Yalçın

Şirket Danışmanı
Serkan Kabalı
Bölgesel Değer Zinciri Müdürü
Niyazi Usta
Değer Yaratma Uzmanı

Akademik Danışman
Prof. Dr. Bahar Yetiş Kara
Endüstri Mühendisliği Bölümü

ÖZET

DHL Lider Lojistik Ortağı, müşterilerinden biri olan Mattel'in Avrupa'daki oyuncak dağıtım ağını yönetmekten sorumludur. Bu projenin amacı farklı teslimat noktalarını belirli kısıtlar altında gruplayarak Milkrun rotaları oluşturulmasını mümkün kılmak ve planlamaya bütünsel bir bakış açısı kazandırmaktır. Proje kapsamında toplam taşıma maliyetini enazlamak için kurulmuş bir matematiksel model ve hızlı sonuç geliştirmesi için tasarlanmış bir sezgisel yöntem bulunmaktadır. Oluşturulan çözüm metodu hem taktiksel hem operasyonel kısıtları barındıran geniş bir yelpazeyi yönetebilmektedir. Sezgisel yöntem şirketin eski taşıma verileri kullanılarak doğrulanmıştır. Geliştirilen karar destek sistemi, kullanıcı dostu bir arayüz ile şirketin kullanımına sunulmuştur. Elde edilen sistem sonucunda taşıma maliyetlerinde günlük ortalama 7.56% azalma sağlandığı görülmüştür.

Anahtar Kelimeler: Rota Eniyilemesi, Milkrun, Zaman Penceresi

Tactical Outbound Logistics Planning

1. Company Information

DHL Lead Logistics Partner (DHL LLP) manages supply chain networks and operations while they value design, management, operations and continuous improvement throughout the process (2020). DHL partners up with its clients and stakeholders to fully determine the current and future supply chain needs. These needs include compliance, performance management, and network optimization. One of the many clients of DHL LLP is Mattel, which is one of the biggest multinational toy manufacturing company that produces many well-known brands such as Hot Wheels, Barbie, Fisher Price and Polly Pocket dolls.

2. System Analysis

2.1. Current Distribution Process

In order to manage its huge distribution network within Europe, Mattel is in cooperation with DHL LLP. The finished products are stored in Mattel's six main distribution centers (DCs) in Europe that are located in Czech Republic, France, Greece, Netherlands, England and Turkey.

The distribution process is triggered when Mattel transmits orders to Oracle Transportation Management system (OTM) (Usta 2019). OTM makes a pallet estimation and transportation plan. Then, prepared goods are loaded into trucks and delivered to customers.

Delivery methods of DHL LLP are divided as indirect and direct deliveries as seen in Figure 1. Currently, selection of the delivery method is based on volume constraints and this selection process is incapable of minimizing the total transportation cost.

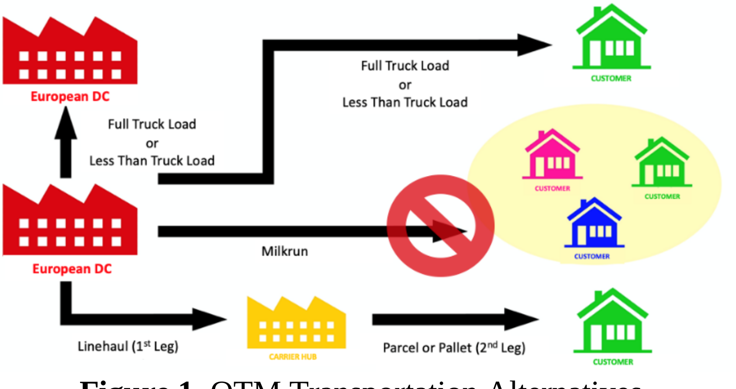


Figure 1. OTM Transportation Alternatives

For DHL LLP and Mattel business partnership, indirect deliveries are composed of two steps: linehaul and pallet transportation. Also, there are three types of transportation modes for direct deliveries: full truckload (FTL), less than truckload (LTL) and Milkrun routes. Although OTM has the ability to build

Milkrun routes, this option is deactivated. DHL LLP cannot utilize this feature due to several reasons. First, because of the high number of destination points i.e. 9000, OTM cannot consider the time windows and geographic proximity at the same time. Second, the planning time of DHL LLP is limited and complex structure of Milkrun routes increases the time OTM takes to make a transportation plan. Third, there are risks associated with the concept of time window due to OTM's incapability of scheduling time buffers between destination points. Note that, a time window is the time slot from the earliest to the latest possible delivery date of each shipment. The time windows are spread into varying time horizons, which may extend up to several months.

2.2. Cost Structure

DHL LLP has a particularly challenging and different cost structure. The costing mechanism varies depending on the transportation mode, volume, carrier, customer location, scheduled appointments and number of shipments within a truck.

Carrier agreements for direct deliveries are made in a way that LTL cost, FTL cost and stop-and-go costs differ based on carriers and customer location zip codes. For FTL shipments, transportation cost of the truck is FTL cost for the selected carrier. For LTL shipments, transportation cost is based on the number of pallets per truck. For a Milkrun route, regardless of the number of pallets exist in the truck, FTL cost of the first stop is charged. Also, for each additional stop, a stop-and-go cost is incurred.

There are two different cost scenarios for indirect deliveries that change with respect to countries. Firstly, if the cost of linehaul equals to zero, cost of the second step is charged. Secondly, if both linehaul and the second step costs are different than zero, then linehaul cost equals to FTL cost and the second step costs are calculated based on the number of pallets.

Some shipments may require appointments within their designated time windows. These appointments are scheduled either hourly-based or daily-based. A booking cost is incurred to the trucks containing at least one shipment with appointment. The booking cost changes with respect to locations of the customers and carriers.

3. Problem Definition

There are two main symptoms that are in the scope of this project. Firstly, while OTM assigns order delivery dates, it does not take into consideration the Milkrun opportunities for orders from different customers with intersecting time windows of delivery dates. Intersecting time windows are the time slots in which the earliest and latest possible delivery dates of different orders intersect. Moreover, although orders are inputted to the system at the same time, they cannot be consolidated. So, it can be said that orders are treated separately. Due to the high number of destination points, it is hard for OTM to consider all customer locations at the same time. As a result, the second symptom is that not only time windows but also customer locations are not considered while determining the transportation modes. Based on these symptoms, the problem can be defined as OTM's considering the orders separately and not consolidating LTL shipments to form Milkrun routes.

In short, this project deals with outbound logistics route optimization from Mattel's European distribution centers to toy retailers. The most significant goal is to provide a holistic view on the overall distribution network as to benefit from cost improvements.

4. System Design

4.1. Mathematical Model

The problem cannot be directly defined as either Standard Vehicle Routing (VRP) problem or Generalized Bin Packing Problem (GBBP). Considering that the cost structure is not based on kilometers/mileage, the problem differs from VRP. In GBBP, all the shipments should be delivered, and the shipment volume should not exceed the truck capacity (Ongarj et al., 2013). Although GBBP explains the basis of the problem, it is not enough to use solely. The reason is that in GBBP, items have a relative cost based on their volumes while the cost structure of the project is different. Furthermore, GBBP does not take into consideration carrier assignments and time windows (Baldi et al. 2018).

Tactical Outbound Logistics Planning problem is considered on a basis of two integrated decisions. First decision is selecting which of the FTL and LTL transportation methods to be used for sending shipments to customers. Second decision is deciding the planned delivery time for each shipment.

In order to utilize the data provided by DHL LLP, pre-process is performed. The input for the mathematical model is constructed as follows. Time windows are set in a matrix form that represents the deliverability of each shipment for each day before clustering. Carriers are allowed to go to specific zip codes that stand for different regions. Hence, zip codes and carriers are matched with a cover matrix for further iterations. Different cost scenarios for each shipment together with its FTL and LTL costs, booking costs and stop-and-go costs are constructed for accelerating the process. In order to have multiple shipments with appointments in a single truck, a cover matrix that represents the combinable shipments is constructed regarding their scheduled appointments and given time buffers. Since there are shipments with hourly-based appointments, this matrix allows the model to be constructed on a daily basis instead of hourly basis. Also, a binary cover matrix for Milkrun coverage area constraint is generated. Milkrun coverage area represents the possible destination points that are in a provided radius from each shipment location as seen in Figure 2.

Figure 2. Milkrun Coverage Area with 50-kilometer radius

Mathematical model takes the cover matrix for Milkrun coverage area and zip code-carrier match as input. Also, time windows, appointment information and volume of shipments are given as input to model. Cost scenarios and carrier-shipment location suitability are utilized. Moreover, shipments with appointment which can be in the same truck are given as input, so that appointment-required shipments can be delivered as the first and the second stops of the Milkrun route, as the company specified. As output, model assigns carriers and transportation modes to trucks with their determined first stops. It also determines whether a shipment with an appointment can be bundled with another shipment that also has an appointment.

Objective function is built as minimizing total transportation cost. Sum of booking costs, stop-and-go costs in Milkrun vehicles, LTL costs and first stops' FTL costs form the total transportation cost.

The mathematical model consists of set of constraints each of which defines the certain characteristic of the system. The constraints are classified as four main groups which are customer-based constraints, transportation mode constraints, capacity constraints and Milkrun routing constraints.

The mathematical model was solved using Gurobi through CVX in MATLAB using the data sent by DHL LLP and covering the dates of 01.02.2019-11.03.2020. The obtained results were shared with the company and the applicability of the results was confirmed.

Mathematical model, required sets, parameters and decision variables can be found in Appendix A.

4.2. Heuristic Approach

Since permanent licensing of the optimization tools would incur incremental costs and it is not possible to use a free solver tool for such a complicated system, a heuristic approach is designed. Additionally, the constructed mathematical model was limited in terms of the dynamism that is the constantly changing environment of the logistic system. For that manner, several algorithms such as Genetic Algorithm, Tabu Search Algorithm, and Nearest Neighbor Algorithm are evaluated. However, considering the structure of the problem, these algorithms do not directly address the problem. So, to catch up with this dynamism and provide applicable solutions in the limited time frame of the planning, a unique construction heuristic based on implicit enumeration with early detections is designed (Jambekar et al. 1977).

This heuristic approach consists of three main components, which are combination construction, feasibility checks and finalizing the solution. The daily OTM output, the longitude and latitude information of the customers and the pallet rates are inputted to the heuristic and converted into their applicable formats for further use.

To elaborate on the dynamism of the problem, user input, such as maximum number of shipments in one truck, maximum number of shipments which require an appointment in one truck, Milkrun radius, truck capacity, minimum and maximum time buffers and the time limit for the overall planning are defined.

Considering the input, carrier-customer pairings based on zip codes are done and customer-wise Milkrun cover matrix is constructed, in which the customers are paired with each other in accordance to their distance in between. Using these, in the combination construction part of the heuristic approach, all the possible combinations are created. Note that, a combination indicates a route with its carrier, cost and possible delivery dates.

From these combinations, the feasibility checks given in Figure 3 are conducted to ensure that the heuristic yields an eventual feasible solution.

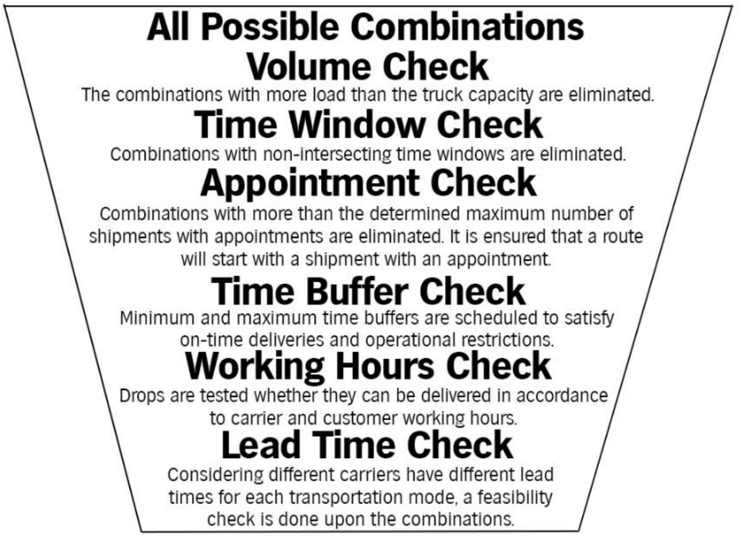


Figure 3. Feasibility Checks Throughout the Algorithm

With the combinations that are feasible up to this point, two suboptimal decisions are made. First one being, if one route has multiple carrier alternatives, the alternatives with the higher costs are eliminated. The reason is that even though they are feasible, they are dominated by the one alternative with the lowest transportation cost. Secondly, if a route is indeed a Milkrun, the equivalent LTL costs of the customers constructing that Milkrun route are summed. If that sum is lower than the cost of the Milkrun route, this means that although the Milkrun option is feasible, it is not advantageous and therefore it is eliminated. These suboptimal decisions aim to make early detections in the combination pool in order to fasten the system by bounding some solutions that are feasible, but anticipated to be poor.

The withstanding combinations are then gathered to form a solution where all the customer shipments are present. This gathering is done upon choosing random indexes from these combinations, until a unique solution is obtained where all the demand is satisfied. Since the algorithm is capable of providing unique solutions quickly and the problem requires a more integral perspective, splitting possible solution space into parts and solving it by using Branch and Bound type of algorithms are expected to curb the performance of the heuristic approach. Throughout the process, the total cost of the solution is updated if it is indeed the lowest one until that time. This procedure continues up to the given time limit that guarantees finding a cost-effective feasible solution.

An addition apart from the scope of this project is the insertion of the linehaul shipments, which is the first step of indirect deliveries. Linehauls are costly transportation modes, where several shipments are combined in one truck to be transferred to carrier hub. Assigning these shipments to the already existing routes will eliminate the costly linehaul transportation. In order to seize such opportunities, the solutions obtained above are utilized.

First of all, a check whether the combinations are feasible in terms of maximum number of shipments is done. Then it is tested whether the first stops are in Milkrun area of the hubs. If a route has a close first stop to the hub and it has not reached the maximum number of shipments yet, it is applicable to go through this addition to the heuristic. From this, hub, linehaul shipments and the first stop pairings are done. All the possible combinations in accordance to the remaining volume are created. The costs of these newly formed combinations are updated. The heuristic approach checks if the linehaul can be eliminated. If it is not possible to eliminate the linehaul, the combinations are reinstated. So, in a way, this selection process works on all or none basis. If indeed the linehaul could be eliminated, the routes are finalized. Overall, the heuristic yields the solution with the lowest cost founded with the given parameters.

In order to incorporate the upcoming changes that may be done by DHL, the designed heuristic is coded using a free computing language, which is Python. The heuristic approach will be operated on a daily basis.

5. Verification and Validation

The performance of heuristic approach is verified with the mathematical model. The features of the heuristic approach which are not included in the mathematical model are turned off and the same input is applied to both mathematical model and the heuristic approach. To evaluate the performance of the heuristic approach, it was tested with big combination pools. For the mathematical model, the number of shipments in each test cannot exceed 20 due to computer performance issues and time limitations. However, the number of possible combinations in these tests go up to 20,000, while in any of the tests conducted with real data, this number did not exceed 5,000. The reason is to observe the deviation from the optimal solution for the worst-case scenarios by pushing the limits of the heuristic approach in a very limited time. 10 runs are made for each input and the average results are taken.

The results demonstrated that even with such big combination pools, outcome of the heuristic approach deviates from the optimal solution by 2.49% on average with 1.97 standard deviation. The summary table is available in Appendix B. Verification results conclude that the heuristic approach yields significant and close results to the optimal solution in an applicable time.

For validation purposes, old routing plans of the company and routing plans of the heuristic approach are compared by choosing OTM output from 12 different days including 384 individual shipments. Based on the OTM output, new routes are created by the heuristic method. It is observed that the heuristic approach found the same results as the company at worst. In 87.5% of the tests, the heuristic approach achieved a better solution regarding the total transportation cost. New routing plans, calculated costs, and improvement performances of these 12 different tests are shared with the company. In conclusion, the applicability of the suggested plans is confirmed by DHL LLP.

6. User Interface

To implement the proposed approach, a user interface is created using Python for the use of planners of DHL LLP. OTM output, longitude-latitude data, carrier data and the cost structure are taken as inputs. The first step is importing them into the decision support system, which is done on the Home page as shown in Appendix C.

If users do not want to include specific source location IDs, zip codes or carriers to the transportation plan, they can exclude them on the New Plan page, which is provided in Appendix D. Planning constraints may vary in different countries and in time, possible changes might occur in the planning process. Therefore, parameters such as the maximum number of stops used in the algorithm are allowed to be determined by the user. The user can also decide whether to include hub transportation into planning process by clicking Hub checkbox. Additionally, since DHL LLP has a limited planning time, users are able to set a time limit. By the time the time limit is reached, the best solution that is achieved until that time is provided as the final solution. Allowing the users to determine parameters, decision support system will be long-lasting for the company.

After all the selections are made, user clicks the Solve button and the algorithm starts to operate. User can follow the cost improvement and the time elapsed on the screen. When it is completed, the report of the transportation plan is displayed on the Report page. In this report, as given in Appendix E, shipment-carrier-truck assignments are presented with related shipment information including the planned delivery time. The first stop of a truck is presented as the first line of that truck. Savings per truck and the total saving are demonstrated by taking the difference between cost of OTM and the decision support system, which allows users to compare the results of the decision support system with the results of OTM. It is told by the company that in some cases, some shipments might be found risky to be included in a Milkrun route. Therefore, on the Report page, users are enabled to deselect those shipments by considering the trade-off between the risk and the saving. If some shipments are deselected, refresh option directly overwrites new transportation plans and displays the updated costs and savings. Undo and reset options are also provided for a dynamic use. After completing the analysis on this page, users are allowed to export the report as an Excel file. Since the merge shipments are entered back into OTM, the Excel file is designed in the most suitable format for OTM in order to ease the entering process. Additionally, an info button is provided for each page to support the user related to the use of the interface. User interface was introduced to the company with a meeting.

7. Project Contributions

The primary contribution of this project is to reduce the total transportation cost of the company by enhancing the current decision-making process and providing the company with a holistic perspective. The main performance measures were selected as the total transportation cost and the time elapsed to make the transportation plan each day. After completing 12 different tests with real data, the improvement percentages are found as up to 20.61%, and 7.56% on average. This implies that the decision support system provides a cost-effective plan in a reasonable time. The performance comparison table is provided in Appendix F. Since users are enabled to select several input parameters, the decision support system allows them to respond to changing conditions dynamically. With this, sustainability of the system is ensured. Moreover, consolidating the shipments with overlapping time windows, the number of required trucks in the delivery process is reduced. This reduction would allow the company to have a more environment-friendly system in the long run. Additionally, when the delivery dates are determined, decision support system selects the latest intersecting day. With this, user is able to input these shipments to heuristic until the delivery date. This provides the chance to merge these shipments with the new ones in the upcoming days. Finally, in this project, the need for manual interventions is eliminated. As a result, workflow process is standardized, and the workforce is utilized better.

As previously mentioned, the main objective of this project was stated by the company as to reduce the transportation cost by seizing Milkrun opportunities. Therefore, as a result of the project, their expectations are satisfied. In their evaluation, the company emphasized that the time buffer parameter is substantially beneficial in ensuring the time window satisfaction of customers. It eliminates the risk of violating the time windows due to uncertainties in unloading times. Also, hub extension is appreciated since it was out of the project scope and it may lead to many cost improvement opportunities. These benefits are highlighted since they are not included in OTM. Since the decision support system is designed to include as many planning constraints provided as possible, applicability of the project is approved. As for the user interface, it is stated that ease of use is ensured due to the user-friendly design of the decision support system. To provide comprehensive explanations, a handbook is delivered to the company.

During the first meetings, the company did not find it necessary to make some parameters user-defined. Instead, they wanted them to be fixed. However, after demonstrating the results of the parameter relaxation tests, the company acknowledged the importance of parameter selection and required those parameters to be user-defined. Moreover, pre-defined limitations are reconsidered and unnecessary limitations are eliminated from the overall system.

Although the heuristic approach was developed based on France data, since many planning constraints are included, the conducted tests demonstrated that the heuristic approach is also applicable for other countries in Europe in which Mattel's transportation network is managed.

Implementation of the decision support system is approved by senior management. An executive presentation is requested to introduce the decision support system to company employees. The company stated that after the necessary adjustments, the decision support system designed for Mattel deliveries may also be implemented to other companies whose transportation network management is under the responsibility of DHL LLP. The decision support system is delivered to the company with a meeting on 24.04.2020. Based on the feedbacks, necessary improvements were carried out.

REFERENCES

Baldi, Mauro Maria, et al. "A Generalized Bin Packing Problem for Parcel Delivery in Last-Mile Logistics." European Journal of Operational

Research, vol. 274, no. 3, 2019, pp. 990–999., doi:10.1016/j.ejor.2018.10.056.
Jambekar, Anil B., and David I. Steinberg. "An Implicit Enumeration Algorithm for the All Integer Programming Problem." Computers & Mathematics with Applications, vol. 4, no. 1, 1978, pp. 15–31., doi:10.1016/0898-1221(78)90004-4.
"Lead Logistics Provider LLP: DHL." DHL, 2020, www.dhl.com/tr-en/home/all-products-and-solutions/logistics-solutions-and-special-expertise/lead-logistics-provider-llp.html.
Ongarj, L., & Ongkunaruk, P. (2013). An integer programming for a bin packing problem with time windows: A case study of a Thai seasoning company. 2013 10th International Conference on Service Systems and Service Management. https://doi.org/10.1109/icsssm.2013.6602522
Usta Niyazi Ömer. "Mattel Carrier Handbook." *Carrier Handbook*, 2019

APPENDIX

Appendix A. Mathematical model

J : Shipment Set , C : Carrier Set , T : Time Interval
K = $\| J \|$

Decision Variables

$FTL_{jckt} = \begin{cases} 1, & \text{if shipment j is in cluster k assigned to carrier c at time t} \\ 0, & \text{otherwise} \end{cases}$

$LTL_{jct} = \begin{cases} 1, & \text{if shipment j is assigned to carrier c at time t} \\ 0, & \text{otherwise} \end{cases}$

$B_{jckt} = \begin{cases} 1, & \text{if first stop of cluster k assigned to carrier c is shipment j at time t} \\ 0, & \text{otherwise} \end{cases}$

$Z_{ij} = \begin{cases} 1, & \text{if shipment i with appointment is bundled with shipment j with appointment} \\ 0, & \text{otherwise} \end{cases}$

Parameters

d = given Milkrun radius in kilometers
v = maximum truck capacity in pallets
s = maximum number of shipments in a truck

$D_{ij} = \begin{cases} 1, & \text{if the distance between location of shipment i and j is smaller than d} \\ 0, & \text{otherwise} \end{cases}$

$K_{jc} = \begin{cases} 1, & \text{if carrier c can deliver shipment j (based on zipcodes)} \\ 0, & \text{otherwise} \end{cases}$

E_j = Earliest possible delivery date of shipment j
L_j = Latest possible delivery date of shipment j
V_j = Volume of shipment j based on number of pallets

$R_{jct} = \begin{cases} 1, & \text{if carrier c can deliver shipment j with LTL lead time at time t} \\ 0, & \text{otherwise} \end{cases}$

$S_{jct} = \begin{cases} 1, & \text{if carrier c can deliver shipment j with FTL lead time at time t} \\ 0, & \text{otherwise} \end{cases}$

$A_j = \begin{cases} 1, & \text{if shipment j has an appointment} \\ 0, & \text{otherwise} \end{cases}$

$H_{ij} = \begin{cases} 1, & \text{if shipment i with appointment can be bundled with shipment j and i is prior to j} \\ 0, & \text{otherwise} \end{cases}$

C_{jc}^1:LTL cost of delivery of shipment j made by carrier c
C_{jc}^2:FTL cost of delivery of shipment j made by carrier c
P_c:stop&go cost of carrier c

Mathematical Model

$$min\ (\sum_{c=1}^{C}(\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{t=1}^{T}(FTL_{jckt}-B_{jckt})P_c)+\sum_{c=1}^{C}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{t=1}^{T}(B_{jckt}C_{jc}^{2})+\sum_{c=1}^{C}\sum_{j=1}^{J}\sum_{t=1}^{T}(LTL_{jct}C_{jc}^{1})$$

$$(1)\quad B_{jckt} \le FTL_{jckt}\ ;\ \forall j \in J, \forall c \in C, \forall k \in K, \forall t \in T$$

$$(2)\quad \sum_{c=1}^{C}\sum_{t=1}^{T}(\sum_{k=1}^{K} FTL_{jckt} + LTL_{jct}) = 1\ ;\ \forall j \in J$$

$$(3)\quad \sum_{k=1}^{K} FTL_{jckt} \le S_{jct}\ ;\ \forall j \in J, \forall c \in C, \forall t \in T$$

$$(4)\quad LTL_{jct} \le R_{jct}\ ;\ \forall j \in J, \forall c \in C, \forall t \in T$$

$$(5)\quad \sum_{t=1}^{T} FTL_{jckt} \le 1\ ;\ \forall j \in J, \forall c \in C, \forall k \in K$$

$$(6)\quad \sum_{j=1}^{J} FTL_{jckt}V_j \le v\ ;\ \forall c \in C, \forall k \in K, \forall t \in T$$

$$(7)\quad \sum_{j=1}^{J} FTL_{jckt} \le s\ ;\ \forall t \in T, \forall c \in C, \forall k \in K$$

$$(8)\quad \sum_{k=1}^{K} B_{jckt} \le K_{jc}\ ;\ \forall j \in J, \forall c \in C, \forall t \in T$$

$$(9)\quad \sum_{j=1}^{J} FTL_{jckt}A_j \le 2\ ;\ \forall c \in C, \forall k \in K, \forall t \in T$$

$$(10)\quad \sum_{c=1}^{C}\sum_{t=1}^{T}\sum_{k=1}^{K} B_{jckt} \ge \sum_{c=1}^{C}\sum_{t=1}^{T}\sum_{k=1}^{K} FTL_{jckt}A_j - \sum_{i=1}^{I} Z_{ij}\ ;\ \forall j \in J$$

$$(11)\quad FTL_{ickt} + FTL_{jckt} \le D_{ij} + 1\ ;\ \forall i,j \in J, \forall c \in C, \forall k \in K, \forall t \in T,\ i < j$$

$$(12)\quad \frac{\sum_{j=1}^{J} FTL_{jckt}}{J} \le \sum_{j=1}^{J} B_{jckt}\ ;\ \forall c \in C, \forall k \in K, \forall t \in T$$

$$(13)\quad \sum_{j=1}^{J} B_{jckt} \le 1\ ;\ \forall c \in C, \forall k \in K, \forall t \in T$$

$$(14)\quad Z_{ij} \le H_{ij}\ ;\ \forall i,j \in J$$

$$(15)\quad Z_{ij} \le \sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{t=1}^{T} FTL_{jckt};\ \forall i,j \in J$$

$$(16)\quad Z_{ij} \le \sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{t=1}^{T} FTL_{ickt};\ \forall i,j \in J$$

$$(17)\quad \sum_{i=1}^{I} Z_{ij} \le 1;\ \forall j \in J$$

$$(18)\quad \sum_{j=1}^{J} Z_{ij} \le 1;\ \forall i \in J$$

$$(19)\quad FTL_{jckt} \in \{0,1\};\ \forall j \in J, \forall c \in C, \forall k \in K, \forall t \in T$$

$$(20)\quad LTL_{jct} \in \{0,1\};\ \forall j \in J, \forall c \in C, \forall t \in T$$

$$(21)\quad B_{jckt} \in \{0,1\};\ \forall j \in J, \forall c \in C, \forall k \in K, \forall t \in T$$

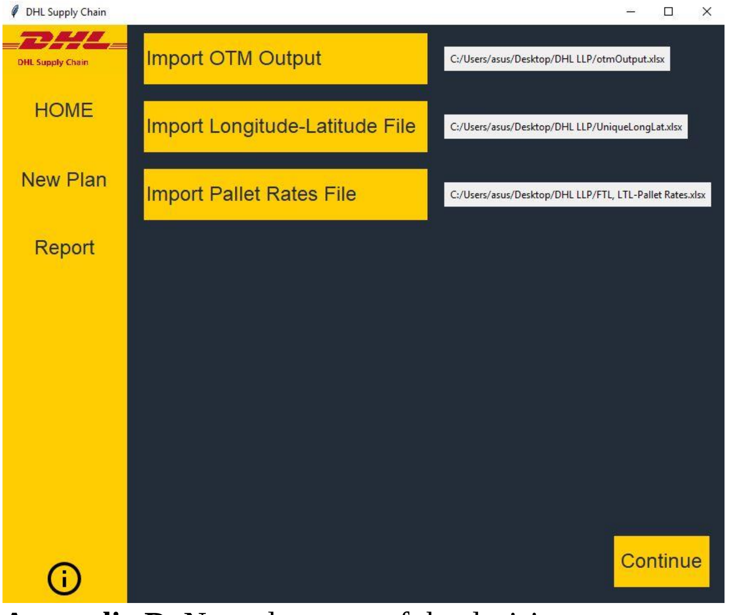
$$(22)\quad Z_{ij} \in \{0,1\};\ \forall i \in J, \forall j \in J$$

Appendix B. Verification of the Heuristic Approach

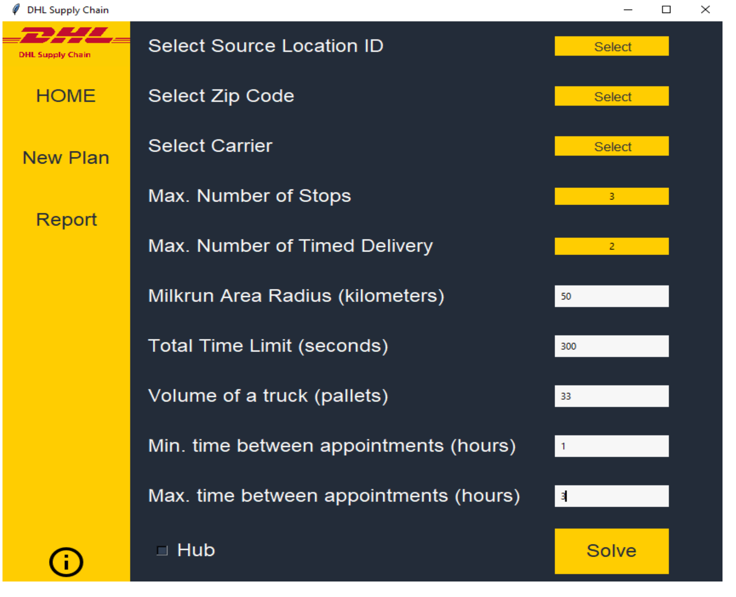
Shipment Tests									
Carrier Count	Shipment Count	# of Combinations Pre-eliminations	# of Combinations Post-elimination	# of Unique Solutions in Average	Heuristic Approach Cost	Heuristic Approach Run Time (seconds)	Mathematical Model Cost	Mathematical Model Run Time (seconds)	Deviation Percentage
1	5	85	25	520	€ 2,921.10	30.39	€ 2,921.10	27	0.00%
	8	100	33	13,384	€ 4,908.57	45.02	€ 4,907.46	88	0.02%
2	12	2196	47	52,222	€ 10,416.02	103.32	€ 10,328.14	307.49	0.85%
	15	4423	140	82,787	€ 12,292.33	162.39	€ 11,777.65	611.527	4.37%
3	17	8840	160	105,104	€ 12,813.98	300.01	€ 12,338.83	1439.591	3.85%
	20	15019	677	99,223	€ 14,046.02	300.01	€ 13,328.85	2519.826	5.38%

Carrier Tests									
Shipment Count	Carrier Count	# of Combinations Pre-eliminations	# of Combinations Post-elimination	# of Unique Solutions in Average	Heuristic Approach Cost	Heuristic Approach Run Time (seconds)	Mathematical Model Cost	Mathematical Model Run Time (seconds)	Deviation Percentage
8	1	100	33	13,384	€ 4,908.57	45.02	€ 4,907.46	88	0.02%
	2	800	35	21,318	€ 4,414.36	58.82	€ 4,323.36	129	2.10%
15	2	4423	140	82,787	€ 12,292.33	162.39	€ 11,777.65	611.527	4.37%
	3	6366	144	88,615	€ 11,728.21	186.77	€ 11,289.07	986.546	3.89%
20	3	15019	677	99,223	€ 14,046.02	300.01	€ 13,328.85	2519.826	5.38%
	4	20050	677	93,447	€ 13,030.35	300.01	€ 12,782.37	4251.183	1.94%

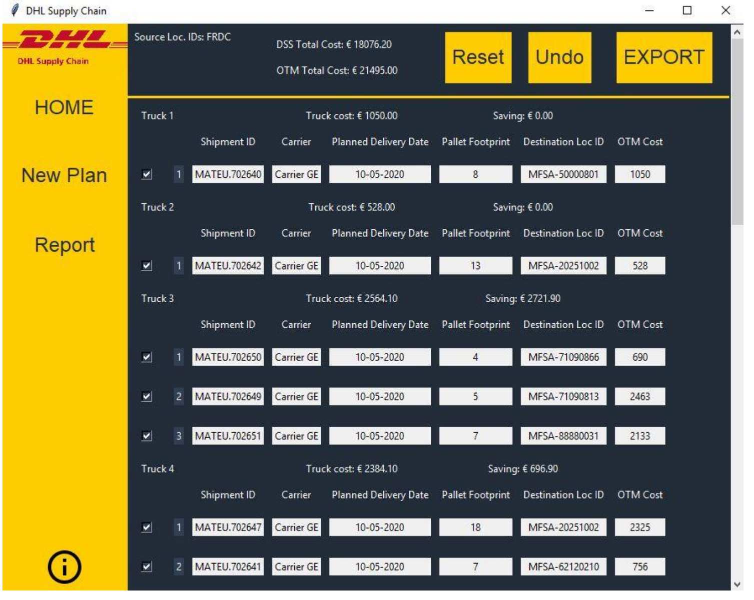
Appendix C. Home page of the decision support system



Appendix D. New plan page of the decision support system



Appendix E. Report page of the decision support system

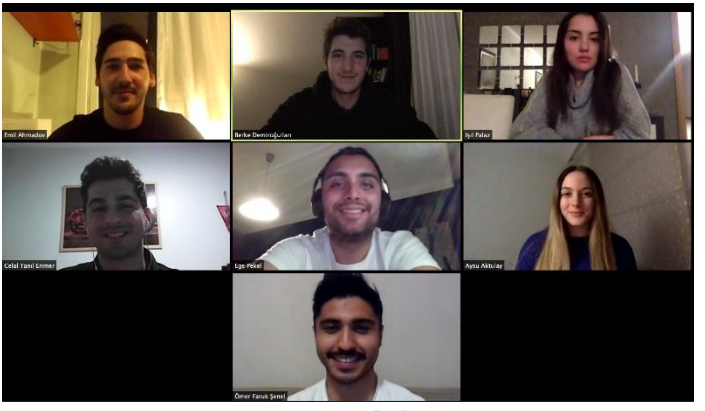


Appendix F. Performance comparison of the heuristic approach and OTM

Test Number	Shipment Count	OTM Cost	Heuristic Approach Cost	Improvement Percentage
1	24	€ 2,372.44	€ 2,372.44	0.00%
2	13	€ 5,506.67	€ 5,013.33	8.96%
3	33	€ 52,519.56	€ 41,696.64	20.61%
4	41	€ 45,477.51	€ 44,573.92	1.99%
5	59	€ 49,453.44	€ 47,319.68	4.31%
6	24	€ 23,050.74	€ 23,050.74	0.00%
7	57	€ 63,467.18	€ 61,504.43	3.09%
8	12	€ 12,915.39	€ 11,353.62	12.09%
9	42	€ 53,903.88	€ 52,515.16	2.58%
10	20	€ 21,612.00	€ 18,193.20	15.82%
11	17	€ 21,771.00	€ 20,281.00	6.84%
12	42	€ 51,555.31	€ 44,090.64	14.5%

Talep Tahmini Karar Destek Sistemi

Dönmez Debriyaj Sanayi ve Ticaret A.Ş.



Proje Ekibi

Emil Ahmadov, Aysu Aktulay, Hayati Berke Demiroğulları, Celal Tanıl Erimer, Işıl Palaz, Ege Pekel, Ömer Faruk Şenel

Şirket Danışmanı
Sabri Tüysüz
Planlama ve Lojistik Müdürü

Akademik Danışman
Dr. Emre Uzun
Endüstri Mühendisliği Bölümü

ÖZET

Dönmez Debriyaj A.Ş.'nin mevcut talep tahmin sistemi ayda bir çalıştırılmaktadır ve gelen siparişlere reaksiyon süresi yaklaşık 10 haftadır. Bu projenin amacı mevcut sistemin iyileştirilerek tahminlerin daha fazla geçmiş veri odaklı hale getirilmesidir. Bu bağlamda R ve R Shiny yazılım dili kullanılarak tamamen şirketin ihtiyaçlarına yönelik bir karar destek sistemi geliştirilmiştir. Bu yazılım şirketin geçmiş verilerini kullanarak 500'den fazla son ürün için kullanıcının seçmiş olduğu parametrelere göre dinamik talep tahminlemesi yapmaktadır. Kurulan sistem yardımıyla örnek çalışmalar yapılmış, bu çalışmalar sonucunda şirketin tahminlerinde yaptığı hata payı %50 oranında iyileştirilmiştir ve bu iyileşmelerin şirketin envanter harcamalarını %44 oranında azaltacağı öngörülmüştür.

Anahtar Kelimeler: Talep Tahmini, Aftermarket, Karar Destek Sistemi

Forecasting Decision Support System for Dönmez Debriyaj A.Ş

1. Company Information

Dönmez Debriyaj Sanayi ve Ticaret A.Ş. (DD) was founded in 1986 in İzmir, Turkey to produce clutch, clutch cover assemblies and release bearings for heavy-duty commercial vehicles. DD was producing clutches for Mercedes in its early years but now, the company has become the largest original equipment manufacturer (OEM) and aftermarket supplier of clutch in Turkey. Company offers 1166 end products to their customers and exports products to over 80 countries.

2.System Analysis

Current system relies highly on the intuitions and incentives of the workers. Forecasting process starts with the field reports from the sales people of the company. Sales managers take the field reports and start to analyze this data. These field reports contain the information about the customers of the retailers and their expectation of the demand for the next period. Other than the field reports, the sales manager also views some important key performance indicators for the market, such as mega projects, number of registered heavy vehicles inside the country, expected market share of the company and elections. After considering all these factors, the sales manager gives a prediction for the next period. Intuitive methods result in less forecast accuracy. In the figure below, the difference between forecasts and realized demand can be seen. Mean Absolute Error for product “125 801” is 163.6 units.

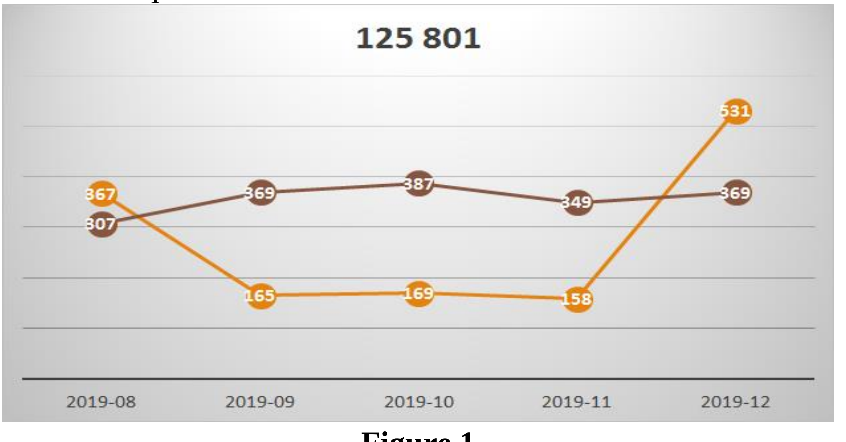


Figure 1

3. Problem Definition

DD produces different types of products for the automotive aftermarket. Company use intuitive methods or simple forecasting methods, such as moving average, for forecasting demand amounts, which cannot fully comprehend the factors affecting demand. These intuitive methods include a method which is dependent on field reports given by retailers and sales departments in the domestic market. These reports are analyzed and future forecasts are made accordingly. Obtained results are revised by the sales department's intuitions. For foreign market, the same intuitive method is used. However, foreign market sales are dependent on the exact demand quantities of foreign customers. In general, the sales department revises the forecast reports with respect to DD's annual production planning budget and market dynamics .Considering the aftermarket industry's dynamic and unsteady structure, current forecasting methods are not comprehensive enough to cover all the factors affecting demand. This root cause induces several results such as prolonged response times to demand, excessive inventory etc.

4.Proposed System

In order to solve this problem, a solution methodology was proposed. This solution method is based on the improving forecast accuracy of the company. In this section the proposed system will be explained in detail. Before explaining the proposed methodology, it is important to explain the methods used in the approach. For this purpose, firstly, literature survey done will be explained. Explanation of the system will start with required inputs for the system. Secondly, how these inputs are used and how they are processed will be stated. After these stages, outputs of the system will be explained and discussed.

4.1. Literature Survey

Forecasting methods that are used in this project are explained below:

- Multiple regression methods are used to analyze and interpret the relation between predictors and response of a data. This method is used to understand if there exists a relation between parameters and data as it is hypothesized.
- Generalized Linear Model (GLM) is a flexible generalization of ordinary linear regression. In GLM, each outcome of the dependent variables are assumed to be generated from a particular probability distribution such as; normal, binomial, Poisson and gamma.
- ARIMA which is also called Box-Jenkins is a method to forecast the possible demand data which is regardless if the data has a trend or seasonal pattern through time. The progress of this method is to make the dataset look like stationary by transforming data. After some tests with the data, it makes the forecast for the next period whether it is stationary or not.
- Exponential Smoothing Methods:
- Single exponential smoothing is a method that is used for forecasting the time series when there is no existence of trend or seasonal factors. This method is implemented by considering only the actual demand and forecasting values of the previous period.

- Double exponential smoothing is a method which is used for the dataset possibly having a trend through the periods. This method includes a level component and a trend component in each period and also includes two variables that are updated each period to catch the correct forecast value.
- Triple exponential smoothing method, also known as Holt-Winters method is used when the data set has both trend and seasonality. Triple exponential smoothing is concerned on three aspects of the time series: a typical value (average), a slope (trend) over time and a cyclical repeating pattern (seasonality).
- Negative Binomial Regression is a type of generalized linear model is used for over-dispersed count data, which is when the conditional variance exceeds the conditional mean. It has an extra parameter to model the over-dispersion. If the conditional distribution of the outcome variable is over-dispersed, the confidence intervals for the Negative binomial regression are likely to be narrower as compared to those from a Poisson regression model.
- Poisson Regression is used to model response variables (Y-values) that are counts. It tells you which explanatory variables have a statistically significant effect on the response variable. In other words, it tells you which X-values work on the Y-value.

4.2. System Inputs

We proposed a forecast methodology which is built on the order and sales data of each product. Training and testing period of these data are projected to be flexible and set as needed. Other than the sales and orders data, our system is planned to have the data for some important factors which can affect the sales amount. These important factors are decided through meetings with sales managers of the company. Some of these inputs are historical discount rates, working days in each month, consumer confidence index and etc. Firstly user should choose if the forecasted product is sold in domestic market or export market and then choose the product ID. System will divide data into two pieces (by input taken from the user); data that will be used to train the model (which will be referred to as "training data" after this point) and data that will be used to test the model (which will be referred to as "testing data" onwards). Training data is analyzed and the suitable method among; Exponential Smoothing Method and ARIMA model are fitted to the testing data. These models are used in order to predict the sales amount for the testing period. Mean absolute error is found for each model.

Next step is to analyze the training data and data for the important factors and seek a relation between the sales amount and factors. This will be done by using Linear Regression model. These factors are tested to decide the significance where the null hypothesis is that the factors are not significant. Insignificant factors will be eliminated from the regression model. Final regression model is used to make predictions for the testing period. Again, mean absolute error is found for each products' testing period. Afterwards another model is tried which is Generalized Linear Model (GLM) for Poisson family. Generalized Linear Model for Poisson family also uses training data to fit a model for the sales data and identifies the relationships between factors and sales data. Best GLM model is found and is tested through the testing data. Mean absolute error is also calculated for GLM model.

4.3. Data Analysis

The forecasting system logic is designed to analyze any kind of time-series data with five forecast methods. All products in the company are time series and 72 months of order data is used. For any time-series item, historical data of the product is divided into two parts; training period data and test period data. Objective of this division is to train the model in training period data and test the models in test period data. In order to determine the duration of train/test periods, several experiments are conducted. As a result, test period is selected as 68 months and train period is selected as 6 months.

In the training period, five forecast methods are used to analyze the time series product. These forecasting models are Exponential Smoothing methods, ARIMA, Linear Regression model, Generalized Linear Model for Poisson Distribution and Generalized Linear Model for Negative Binomial Distribution. Then, these methods are applied in the test period. Each of these forecasting methods have subcategories which are used for different type of data. All possible models of forecast methods are evaluated during this period. For example, there are specific Exponential Smoothing method for data involving seasonality or trend. In the next step, best Exponential Smoothing method, best ARIMA and best fitted regression models for the data are chosen. In this phase, factors that are not considered important for the demand amount are neglected from the regression models and best regression models are found by this methodology. Also, coefficients for each factor is determined and best fitted model is established. After the best sub model of each forecasting method is found, these models are used to forecast for the duration of testing period. Mean Absolute Error (MAE) for each model is found in testing period. Next step after finding MAE of each model is to compare all the methods' error values. Model with the lowest MAE is chosen as best method to do the forecast for the forecasting horizon.

4.4. Outputs

After forecasts are done for the forecasting period, outputs are displayed. These outputs are diverse and gives insights about the models and forecasts. The outputs show best-fit models and MAE of forecast methods, forecast values for test period made by using best fitted sub models of Exponential Smoothing, ARIMA, GLM for Poisson, GLM for Negative Binomial and Linear Regression methods along with the future forecast values.

Flow chart of the proposed system is given in *Appendix A*.

5. Benchmark

From the very beginning of the project, all the comparisons and results are done by the measure of MAE. Additionally, RMSE is used to measure the effects of outliers. In the below part, all of the work of our group is summarized in terms of a "Weighted Average" measure that is developed by our group. This mainly measures the error corresponds to the unit mean.

In this case, the weighted average for MAE and RMSE is calculated for benchmarking (*Equation 1*). This method calculates the sum product of all items' average and their error values and divides this number with the sum of all averages. It can be interpreted as an average error made for one unit of product. In MAE, our model yields 26.52 which is 20.85 lower than the company. In RMSE, our model yields 27.50 and it is 28.53 lower than the company. Results can be seen in *Table 1*.

$$Weighted\ Average = \frac{\sum_i \quad Monthly\ Average\ Sales(i) * MAE(i)}{\sum \quad Monthly\ Average\ Sales} \quad (Equation\ 1)$$

Table 1: Weighted Average Comparison

Weighted Average	MAE	RMSE
Our Model	26.52	27.50
Company	47.37	56.03
Difference	20.85	28.53

As a result of the comparison between MAE and RMSE, since the difference increases as going through from MAE to RMSE, it can be said that the outliers have important effects on the system. If our model ended up with a larger RMSE result which is worse, we would say that outliers do not have an important impact on the forecasts and proposed methods cannot predict the future orders as good as expected. As a result, since our group is better 78% on average for all products, outliers are not common causes that they may not be included in our model especially in the regression models.

6.User Interface and Application

The application is built by integration of RStudio and RShiny according to the expectations of the company. Snapshot of our initial user interface can be seen in Figure 2.

Figure 2: Snapshot of the User Interface System

Our interface has two panels; sidebar panel and main panel. Sidebar panel is on the left-hand side in *Figure 1* and it prompts user to initialize the system by asking certain inputs. Firstly, user selects which market to forecast in the very first section. Based on this selection, Product ID's are updated in the second section. Then, user selects the Product ID and future forecasting period along with the training period. All these selections are passed to the server code as an input.

Server code runs according to the system flow and applies the best forecast method. Server is programmed to return 11 outputs to the user interface. These outputs are shown in the tabs in the main panel of the user interface which is placed in the right-hand side on *Figure1*.

7.Integration of Proposed Solutions

Company's expectation from this project is to have a decision support system which will be used for forecasting. Current forecasting method of the company is based on heuristic methods and is highly dependent on intuitive decisions. Hence, this project developed and delivered a decision support system for forecasting based on past data and parameters of the company.

The implementation was done in two phases. Integration was supposed to take place in the company facilities however, due to the extraordinary circumstances, implementation phases were done by a video conference call with IA and team members. In that meeting, initial version of the system was introduced to IA. Firstly, R Studio and R applications were installed to the company. Then, user manual, system design, code block, interpretation of input tabs and output tabs were explained. After the first phase, company was able to test the system in the provided user interface and they gave feedbacks to our team which are included in the final version.

Second phase included the final version of the system. Based on the feedbacks, plotting feature was redesigned, monthly inputs were made more flexible with the integration of sales/order option and some cosmetic changes were made such as colorization and fonts.

8.Outcomes of the Project

Project will have a lot of benefits for the company in future. Major benefit of the project will be accurate forecasts. Accurate forecasts will cause a better understanding of the customer demand. Any kind of pattern in demand is comprehended by the system. Better prediction of demand will increase the company's profit in the future and market share. Better accuracy will also prevent the loss of customers. Additionally, enhancement of forecasting system will improve the production planning. Company will be able to better plan its production process. Required amount of raw materials which will be ordered and kept in the inventory, will decrease. Also, excessive amounts of production will be prevented, which will consequently cause a decrease in inventory levels.

8.Conclusion

Main aim of this project was to provide a tool that can estimate the future orders and sales of the customers in Turkey and abroad basically. In this way, company had some primary expectations, for instance, to have better forecasts comparing to the sales-persons' foresights or field reports. Thus, after providing the tool that our group created, they are now able to do forecasting relying on the mathematical models that consider affects which are parameters, seasonality, trend and other effects that have an impact on the order and sales data. At this point, our group was able to forecast 70%-80% better than the company. Hence, primary expectation of a company is satisfied. In addition to the primary expectation, company stated some other concerns regarding to the tool, especially for the interface from our group. These concerns were especially about combining the sales and order data manually considering a specific period. That means, company thought that even though the tool that we provided can mostly fill the primary concern of the company, since there are some factors that the tool cannot take into account, we are demanded to modify the interface which also provides a customization service to the company. As a result, after providing this service, now, company is able to include their foresights into the model by choosing which period is an order period or a sales period.

Provided system can be made compatible with Dönmez Debriyaj A.Ş.'s current ERP software in future plans. This compatibility could increase efficiency in forecasting. Moreover, this software can be very useful with the inventory planning for the future.

REFERENCES

Breitschwerdt, D. , Cornet, A. , Kempf, S. , Schmidt, L. (2017, June). "The changing aftermarket game–and how automotive suppliers can benefit from arising opportunities", McKinsey&Company.

Hyndman, R. J., & Athanasopoulos, G. (2018). "Forecasting: principles and practice", Melbourne: OTexts.

İstanbulluoğlu, O. (2019). "Otomotiv Satış Sonrası Pazarı'na Genel Bakış", AutoWorld Dergisi.

Nwogu, E. C., Iwueze, I. Z., Nlebedim, V. U. (2016,January 11). "Some Tests for Seasonality in Time Series Data", Retrieved from http://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=2030&context=jmasm.

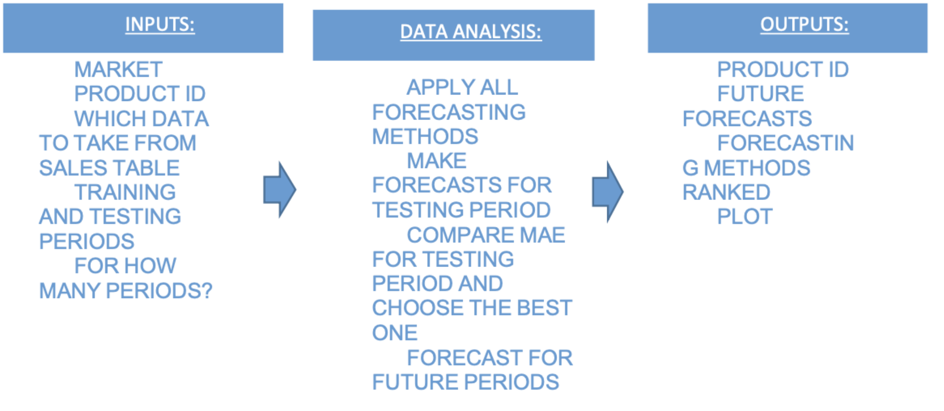
Ollech, D. (2019, February 6). Retrieved from https://cran.r-project.org/web/packages/seastests/vignettes/seastests-vignette.html.

Shcherbakov, Maxim & Brebels, Adriaan & Shcherbakova, N.L. & Tyukov, Anton & Janovsky, T.A. & Kamaev, V.A. (2013). "A survey of forecast error measures", World Applied Sciences Journal, 24. 171-176, 10.5829/idosi.wasj.2013.24.itmies.80032.

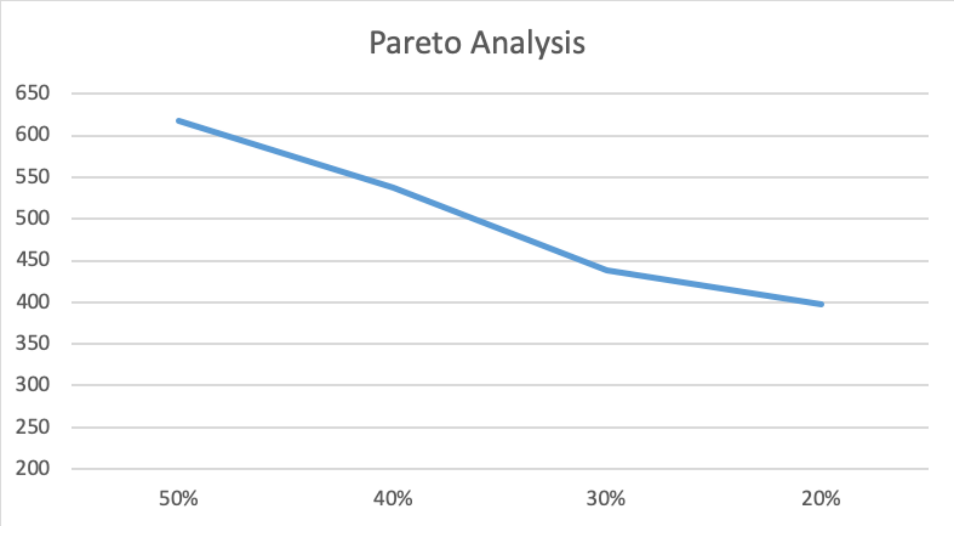
Unkel, S., Farrington, C. P., Garthwaite, P. H., Robertson, C., & Andrews, N. (2011). "Statistical methods for the prospective detection of infectious disease outbreaks: a review", Journal of the Royal Statistical Society: Series A (Statistics in Society), 175(1), 49–82. doi: 10.1111/j.1467-985x.2011.00714.x

APPENDIX

Appendix A: Flow Chart of the Proposed System



Appendix B: Pareto Analysis



Hammadde Kategorizasyonu, Talep Tahmini ve Stok Eniyilemesi için Karar Destek Sistemi

Durukan Şekerleme Sanayi ve Ticaret A.Ş.

Proje Ekibi

Osman Akgün, Umay Atay, Bilge Balcılar, Özge Çalışkan, Emre Can Durmuş, Şimal İrem Gökalp, Neslişah Özkan

Şirket Danışmanı
Gonca Altıntaş
Tedarik Zinciri Müdürü

Akademik Danışman
Dr. Öğr. Üyesi Emre Nadar
Endüstri Mühendisliği Bölümü

ÖZET

Projenin amacı stok maliyetlerini düşürerek mevcut hammadde envanter yönetimini geliştirmektir. Bu projede ele alınan hammadeler durağan tüketim verisi, belirli sabit termin süresi ve sıfır sabit sipariş maliyeti gibi özelliklere sahiptir. Durukan Şekerleme'nin sağladığı tüketim verilerine R Studio'da altı farklı talep tahmin yöntemi uygulanmış; A, B ve C kategorileri için RMSE değerlerine göre en uygun yöntemler seçilmiştir. Bu hammaddelerin periyodik envanter kontrolü için dengeli temel stok, bağımsız temel stok ve miyopik temel stok politikalarının kullanımı değerlendirilmiştir. Uygulanan sezgisel yöntemlerin performansı, dinamik programlamayla hesaplanan en düşük envanter maliyetleriyle kıyaslanarak ölçülmüştür. Projemiz sonucunda, şirkete talep tahmini ve önerilen şipariş miktarlarını sunan dinamik ve güncellenebilir bir karar destek sistemi Excel VBA kullanılarak hazırlanmıştır.

Anahtar Kelimeler: Talep tahminleme, karar destek sistemi, hammadde stok yönetimi

Decision Support System for Raw Material Categorization, Demand Forecasting and Stock Optimization

1. System Description

1.1. General Information about the Company

Durukan Confectionery was established in 1993. The owners of the company are Uğur Durukan, Mahmut Nedim Durukan, and Ertuğrul Durukan. The market share of the company in the domestic market is approximately 60%. The company also sells its own branded products in international markets such as the USA, European Countries, Middle East and North Africa Region, and the Far East Countries.

1.2. Manufacturing System

Durukan's manufacturing system can be examined in two separate lines based on the product's segments: dragee chewy and compressed candy production. Their manufacturing system mainly follows the same processes; compounding, baking, mixing, kneading, smashing, formalizing, placing stick (for lollipops), and packaging respectively. Because of the ongoing chemical reactions, mechanical treatments, and heat treatments, the manufacturing type is continuous.

2. Problem Analysis

2.1. Analysis of the Current System

The company uses an experience-based stock policy and determines fixed safety stock quantity for each raw material. Moreover, the supply chain department forecasts the demand by taking the average of the last 7 months' consumption quantities of each raw material. To avoid stock-out, they determine the raw material order amounts by taking the maximum of the monthly average consumption and the amount that will be required to satisfy the end-product orders.

For the most critical raw materials, "Kristal Şeker", "Glukoz Şurubu", and "Sitrik Asit", the company runs their MRP system twice a month and according to the results, they update the order requests and production plan for these raw materials if necessary. For the other raw materials, the company reviews the inventory levels monthly and gives orders respectively.

The lead times of the raw materials are deterministic as 7, 15, 21, and 60 days for different raw materials. The shelf life of each raw material is at least two years. Since the company does not store the raw materials in their inventory for more than a year, the materials do not face deterioration.

The manufacturing of the sugar includes a common initial step independent from the type of candy. The company states that "Kristal Şeker", "Glukoz Şurubu" and "Sitrik Asit" are common raw materials and their consumptions are dependent on each other.

2.2. Problem Definition

The current raw material stock policy contains two sub-problems which are inaccurate forecasting and inefficient stock policy. The main reason for the forecasting problem is that the demand coming from the international market is highly uncertain. Furthermore, the company states that the accuracy of the forecast made by the sales department is approximately 40% in terms of MAPE. Due to the unreliable forecasts and uncertain demands, the company holds high amounts of safety stock which is costly. The second problem is that the company does not rely on a structural policy since they use an experience-based inventory policy method, which is not sustainable and practical for the company in the long-term.

3. Project Approach

Considering the problems mentioned above, we aimed to increase the accuracy of the raw material demand forecasting by proposing new methods and to decrease the inventory cost by implementing different stock policies. In this project, instead of examining all 1072 different raw materials, we worked with consumption data of 23 different raw materials that reflect all characteristics. In this way, we were able to provide a sustainable Decision Support System (DSS) that provides the opportunity for the company to implement a more cost-efficient inventory management system on all raw materials.

4. Literature Review

We adopted the test procedures in Zivot and Wang (2003) to make statistical inferences about trends and seasonalities of raw materials. We evaluated the use of different forecast methods: moving average, exponential smoothing, and Holt's method (Nahmias and Olsen, 2015). We also reviewed more complex forecast methods such as ARIMA (NIST Handbook, 2003), neural network, and naive method (Hyndman and Athanasopoulos, 2015). We proposed the use of the base-stock replenishment policy for raw materials with independent demands (Hariga and Ben-Daya, 1999). Since the company stated that some raw materials might have correlated demands, we also reviewed coordinated replenishment policies. In this regard, we evaluated the use of the balanced base-stock policy that was introduced by Karaarslan et al. (2013) as a heuristic approach for assemble-to-order inventory systems.

5. Methodology

Our methodology includes seven major steps: "Data Analysis", "Correlation Analysis", "ABC Analysis", "Forecast Methods Implementations", "Stock Policy Implementations", "Comparison of Stock Policies", and "Optimality Gap Analysis by Dynamic Programming".

5.1. Data Analysis

In order to analyze the seasonality and trend of each item, we applied three tests, which are Augmented Dickey-Fuller, Kwiatkowski-Phillips-Schmidt-Shin, and Phillips-Perron, using R Studio. For most of the items, results of the tests were inconsistent. This means that applying only these tests were not adequate for our project. In order to increase the reliability of our data analysis, we applied a linear regression model to each item and selected the confidence level as 95%. We found most of the p-values of items are larger than 0.05. For the items having p-values less than 0.05, we calculated the confidence intervals under the same significance level and observed that the intervals are very close to zero. Therefore, we could not directly conclude that the consumption data of the raw materials is nonstationary.

5.2. Correlation Analysis

The company stated that most of the end-products approximately contain three main ingredients: 45% “Kristal Şeker”, 45% “Glukoz Şurubu”, and 4% “Sitrik Asit". We investigated whether there exists any significant correlation between these three main ingredients. Hence, we measured the linear correlation among the variables by calculating the pairwise Pearson correlation coefficient in R Studio. According to our hypothesis tests, we rejected the null hypothesis under a 95% confidence level and concluded that there exists a significant and positive correlation. Therefore, the stock policies developed for assemble-to-order systems seem to be reasonable options for these raw materials.

5.3. ABC Analysis

We conducted an ABC analysis to determine the importance of each raw material according to their consumption amounts and unit costs. The calculations showed that “Kristal Şeker” has a 40% cumulative contribution among 23 raw materials, thus it is categorized as category A. In other words, its contribution is significantly higher than the other raw materials. Since this did not reflect a completely realistic case due to the high consumption level of “Kristal Şeker” according to the company, we conducted the ABC analysis by omitting “Kristal Şeker”. See Appendix A for category A items.

5.4. Forecast Methods Implementations

We applied six different forecast methods: moving average, exponential smoothing, Holt’s method, ARIMA, naive method, and neural network. We implemented the moving average method, the exponential smoothing method, and Holt’s method in Excel. In the moving average method, to find the most appropriate value for the number of observations used for the average calculation, we tried six different values: 2,3,4,5,6,7. Then, we calculated the errors using MAD, MSE, and RMSE values. In the exponential smoothing method, to find the optimal smoothing constant value, we solved an optimization problem which minimizes the MSE value of the method by Excel Solver. In the

Holt's method, we calculated the optimal constant values using the same approach that we applied in the exponential smoothing method.

In addition to these methods, we applied ARIMA, naive method, and neural network in R Studio, since applying them in Excel is more complicated. We started with loading the package "forecast" in R Studio. Also, we created the models for each forecast method by using R Studio's specified functions. However, since neural network assumes that the data has a normal distribution, we applied the Box-Cox transformation to normalize the data for the possibility of any opposite case. Lastly, we calculated the MSE and RMSE values of forecasts.

After implementing all forecast methods, we realized that implementing and comparing forecast methods in different applications are not very user-friendly. Hence, we decided to apply all forecast methods in R. Then, using the R Shiny package, we designed a user-friendly forecast application, completing the first part of our DSS.

We divided our monthly consumption data of raw materials into two sets: training and testing data sets. In general practice, the 80% of the data is allocated to the training set and the remaining 20% forms the testing set (Bell, 2014). Therefore, we chose the last 9 data points (January-September 2019) of the 44 data points as our testing set and the earlier 35 data points as our training set. For each raw material, we computed the RMSE values of all the aforementioned forecast methods. Since calculating RMSE values for selected 23 items takes 80 minutes and this solution time will be significantly larger for 1072 materials, we decided to generalize forecast methods for raw materials in categories B and C via the scalarization method of Kirkwood (1997). If a forecast method is more accurate for a raw material, we assigned a higher value between 0 and 1. Then, for categories B and C, we summed up these scaled values for each forecast method. Finally, we selected the forecast method with the highest total scaled value as the generalized method for categories B and C. Results of this process can be seen in Appendix A. Even though this process may lead to some reduction in accuracy, it helps us to decrease the total solution time from 80 minutes to 20 minutes for 23 items. Although we generalized the forecast method, there is still a significant improvement on the accuracy of the company's forecast method. These improvements can be seen in the Performance Analysis section.

5.5. Stock Policy Implementations

Since we could not reach any strong conclusions regarding some characteristics of the raw materials from our data analysis, we decided to evaluate the use of several different stock policies: the independent base-stock policy that is reasonable for items with independent stationary demands, the myopic base-stock policy that is reasonable for items with independent non-stationary demands, and the balanced base-stock policy that is reasonable for items with correlated demands. Taking into account company officials' opinions, we decided to implement weekly-review inventory control policies for category A items and monthly-review policies for the other items.

5.5.1. Independent Base-Stock Policy (S-1, S)

We assumed that our demand over review period plus lead time (T+L) is normally distributed with mean $\mu_{(T+L)}$ and standard deviation $\sigma_{(T+L)}$ since we assumed weekly and monthly consumption data as normally distributed. Also, we defined type 1 service level α as 0.9 and z-value as safety factor (z). As the last step of the policy, we calculated the base-stock level (S) with $S = \mu_{(T+L)} + z * \sigma_{(T+L)}$ formulation.

After completing all the required steps, we applied this policy to make a comparison with the company's current policy. An example of this implementation for "Kristal Şeker" can be seen in Appendix B. As it is seen in this appendix, since this material is in category A, the independent base-stock policy is evaluated as weekly. While we used 196 actual past consumption data for weekly review, we used 44 actual past consumption data for monthly review. Both actual past consumption data begin from 2016. The mean consumption and its standard deviation in our project are calculated by using these actual past consumption data based on the review type. In this appendix, there are two columns for order quantity which are named as "Calculated Order Quantity" and "Recommended Order Quantity". "Recommended Order Quantity" in the table is determined by considering the minimum order quantity and the incremental quantity enforced by the supplier.

5.5.2. Balanced Base-Stock Policy

This method is applicable for two items with potentially correlated demands: "Kristal Şeker" and "Glukoz Şurubu". We selected these two items according to their importance among the three items with potentially correlated demands. This policy prioritizes the components according to their unit costs. It assumes that if the unit cost of the first component (h_1) is greater than or equal to the unit cost of the second component (h_2) and that the lead time plus review period length of the first component (L_1+ R_1) is greater than or equal to the lead time plus review period length of the second component (L_2+ R_2). This assumption is valid in our case. We selected "Kristal Şeker" as component 1 with $h_1 = 3.62$, $L_1 = 1$, and $R_1 = 1$; "Glukoz Şurubu" ($h_2 = 2.2$), $L_2 = 1$, and $R_2 = 1$ (the units of L and R are weeks). We found the optimal base stock level (S_1) by using the formula below suggested by Karaarslan et al. (2013). Where p is the penalty cost and $F_k(x)$ is the cumulative distribution function of demand over k periods. Since the penalty cost is not measured by the company, by choosing the target service level as 90%.

$$\frac{1}{R_1}\sum_{k=0}^{R_1-1} F_k(S_1) = \frac{p}{p+\ h_1 + h_2} \qquad ,where\ \ p = \frac{\gamma(h_1+h_2)}{(1-\gamma)}$$

The computation of all these steps can be seen in Appendix C. The results of this policy can be seen in Appendix D. We observed that backorders arise in many periods when this policy is implemented. Therefore, we constructed a new heuristic policy that prioritizes the components according to their mean demand values in order to better reflect our desire to fulfill the demand on time as often as possible. See Appendix E for the results. We observed that the results are improved by the new heuristic strategy. Therefore, we decided to apply this heuristic strategy in our project instead of the classic balanced base-stock policy.

We compared this method with the independent base-stock policy for these two items. Although the balanced base-stock policy provides lower holding costs for these materials, it tends to lead to stock-outs. This result indicates that the correlation between these two items is not significant enough to justify the use of the balanced base-stock policy. Hence, we decided to eliminate this policy from consideration in implementation.

5.5.3. Myopic Base Stock Policy

Some raw materials seem to have non-stationary demand (although not significant) according to our data analysis. We evaluated the use of the myopic base-stock policy as a heuristic approach for such items by considering our forecasted demand values and calculated safety stock. The suggested order quantity calculations are done by the formulation below.

If $\quad I_t + OO_t > SS + \sum_{t=1}^{2*L} F_t$

Then $\quad O_t = 0$

Else $\quad O_t = \sum_{t=1}^{2*L} F_t + SS - I_t - OO_t$

I_t: On-hand inventory level at period t

OO_t: On-order inventory level at period t

SS: Calculated safety stock (= z * σ_{T+L})

F_t: Forecasted value for period t

O_t: Suggested order quantity at period t

L: Lead time

5.6. Comparison of Stock Policies

We formulated a multi-objective decision-making problem to make a comparison between the above stock policies. To solve this problem, we used the weighted scalarization method where we assigned values between 0 and 1 for each stock policy and each raw material by taking into account the holding cost and service level realizations. The calculation for category A can be seen in Appendix F. For categories B and C, we also calculated the total scaled values across raw materials for each stock policy in order to generalize them. Finally, we selected the stock policy which has the highest scaled value. See Appendix G for our policy recommendations.

5.7. Optimality Gap Analysis by Dynamic Programming

We developed a dynamic programming (DP) algorithm for the two critical raw materials: "Kristal Şeker" and "Glukoz Şurubu". In this model, we considered the stages as weeks and states as inventory positions (on-hand inventory plus on-order inventory minus backorders). We restricted our analysis to discrete states and decision spaces by discretizing the normal distribution of the demand. The expected cost in each state consists of the holding cost, the order cost, and the backorder cost. The backorder cost must be included in the recursion to force the system to keep inventory. Since the backorder cost is not measured in the current system, we selected the unit backorder cost as 5 TL which is the minimum required cost to avoid the stock-out. The optimal policy found with the DP algorithm specifies the optimal order quantity in each state of each stage. The DP recursion can be seen in Figure 1. It takes hours to compute the optimal policy (avg. 230 min) for an item considering 52 weeks, whereas we can compute our heuristics within seconds. Based on this reason, we used the DP algorithm only to examine optimality gaps for proposed heuristics policies of two critical items. The results of this examination can be seen in the Performance Analysis section.

t: *week number*, $\quad t \in Z$
i: *inventory position*, $\quad i \in Z$
o: *order amount*, $\quad o \in N$
d_t: *demand in week t*, $\quad d_t \sim N(\mu, \sigma)$
D_t: *demand over lead time*
b: *unit backorder cost*
h: *unit holding cost*
u: *unit order cost*
L: *lead time of an item (weekly)*
$R_t(i)^*$: *optimal cost function*

$$R_t(i)^* = \min\{E[\,b * \max\{D_t - i - o, 0\} + u * o + h * \max\{i + o - D_t, 0\} + R_{t+1}(i + o - d_t)]\}$$

$$\text{where} \quad D_t = \sum_{x=t}^{t+L} d_x\,, \qquad R_{53}(i) = 0, \qquad o \geq 0, \qquad D_t \geq 0$$

Figure 1. The DP recursion

6. Decision Support System (DSS)

As the main outcome of this project, we supplied a DSS to the company. This tool provides the forecast values and recommended order quantities for the raw materials to the user. For these outputs, it uses R chunks and Excel VBA Modules. The flowchart of the DSS can be seen in Figure 2.

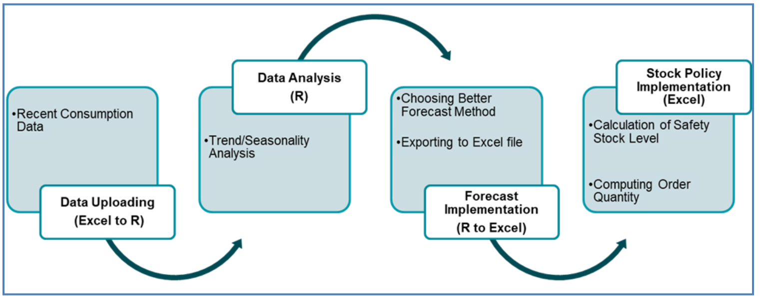


Figure 2. Flowchart of the DSS

The interface of this tool is designed according to the company's expectations. See Appendix H for the main menu of the DSS. In order to create a user-friendly tool, we placed the buttons into the program to ease the access to R Shiny, constructing a bridge between R Shiny and Excel VBA. Thanks to this feature, the user can easily forecast the consumption data. The user interface of R Shiny application can be seen in Appendix I. The use of the DSS is explained in detail to the company by User Manual.

The sustainability of this DSS is crucial for the long-term success of the project. For this reason, we created shortcuts to update the whole system. This update should be applied annually or when a new raw material is added to the system or subtracted from the system.

7. Performance Analysis

The main objectives of this project are to improve the accuracy of the forecasts and to reduce the inventory costs. The comparison of RMSE values between the current forecast method of the company and our proposed forecast methods for each raw material for these 9 months can be seen in Appendix J. This table shows that we could reduce the RMSE values of all raw materials, excluding "AF Jox Yassı", on average 23% for 9 months period. The six review periods comparison between the current inventory management system and our proposed stock policies for each category can be seen in Appendix K. This table shows that we could significantly reduce the total inventory holding costs, on average 32%, without sacrificing much the service level considering six review periods comparison. Using the DP algorithm, we observed that the costs of our proposed policy for "Kristal Şeker" and "Glukoz Şurubu" are only 3.14% and 7.82% larger than the optimal cost respectively. See Appendix L and M for these calculations. In conclusion, we achieved our key goals in forecast and stock policy implementations.

REFERENCES

Bell, Jason. *Machine Learning: Hands-on for Developers and Technical Professionals*. John Wiley & Sons, Inc., 2014.

Hariga, Moncer, and Mohamed Ben-Daya. *"Some Stochastic Inventory Models with Deterministic Variable Lead Time."* European Journal of Operational

Research, vol. 113, no. 1, 1999, pp. 42–51., doi:10.1016/s0377-2217(97)00441-4.
Kirkwood, Craig W. *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets*. Brooks/Cole, Cengage, 1997.
Nahmias, Steven, and Tava Olsen. *Production and Operations Analysis: Strategy, Quality, Analytics, Application*. Waveland Press, Inc., 2015.
NIST-International Sematech e-Handbook: *NIST Handbook 151. Dataplot: NIST Handbook 148*. National Institute of Standards and Technology, 2003.
R. J. Hyndman and G. Athanasopoulos. *Forecasting principles and practice: a comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples*. USA: O texts, 2015.
"Who We Are: Durukan Confectionery Inc." *Who We Are | Durukan Confectionery Inc.*, www.durukan.com.tr/en/who-we-are.
X., Zhang A. *Optimal Order-Up-To Policies in an Assemble-To-Order System With a Single Product*. 1995.
Zivot, Eric, and Jiahui Wang. *Modeling Financial Time Series with S-Plus.* Springer, 2003

APPENDIX

Appendix A. Decided Forecast Methods

Category A	Forecast Method	Category B	Forecast Method	Category C	Forecast Method
Kristal Şeker	Neural Network	For All Raw Materials	Exponential Smoothing	For All Raw Materials	Moving Average
Glukoz Şurubu	Neural Network				
Koli Rocco Noname	Exponential Smoothing				
Sitrik Asit	Exponential Smoothing				
Glucidex	Exponential Smoothing				

Appendix B. Independent Base-Stock Policy for "Kristal Şeker"

Review Period (week)		T =	1				
Service Level		α =	0.9	z value =	1.28		
Weekly Demand			Lead Time (week)		Demand Over Review Period and Lead Time		
Mean (μ) =		36,499	Mean (μ) =	1	Mean (μ) =	72,999	
Std. Dev. (σ) =		18,437	Std. Dev. (σ) =	0	Std. Dev. (σ) =	26,074	
Order Up To Level (S)=		106,413					
Weeks	Dates	Beginning Inventory	S	Calculated Order Quantity	Suggested Order Quantity	Actual Consumption	Ending Inventory
1	(Oct 7-13 , 2019)	78,945	106,413	27,468	27,500	39,483	**39,462**
2	(Oct 14-20 , 2019)	66,962	106,413	39,451	39,500	16,150	**50,812**
3	(Oct 21-27 , 2019)	90,312	106,413	16,101	16,150	27,172	**63,140**
4	(Oct 28- Nov 3 , 2019)	79,290	106,413	27,124	27,150	35,432	**43,858**
5	(Nov 4-10 , 2019)	71,008	106,413	35,405	35,450	40,252	**30,756**
6	(Nov 11-17 , 2019)	66,206	106,413	40,208	40,250	51,333	**14,873**
7	(Nov 18-24 , 2019)	55,123	106,413	51,290	51,300	34,682	**20,441**
8	(Nov 25- Dec 1 , 2019)	71,741	106,413	34,673	34,700	29,112	**42,629**

Appendix C. Balanced Base-Stock Policy Calculations

Items	Unit Cost (h)	Review Period (R)(Days)	Lead Time (L)(Days)	Uncertainty Period (Δ = R + L)	Target Service Level (Type 1)	Penalty Cost
Kristal Şeker	3.62	1	1	2	□ = 0.9	□(h1+h2)/(1-□) = 52.38
Glukoz Şurubu	2.2	1	1	2		

Appendix D. The Results of the Balanced Base-Stock Policy

Kristal Şeker

Weeks	Dates	Beginning Inventory	S	Calculated Order Quantity	Suggested Order Quantity	Actual Consumption	Ending Inventory
1	(Oct 7-13 , 2019)	78,945	60,098	0	0	39,482	**39,462**
2	(Oct 14-20 , 2019)	39,462	60,098	20,636	20,650	16,150	**23,312**
3	(Oct 21-27 , 2019)	43,962	60,098	16,136	16,150	27,172	**16,790**
4	(Oct 28- Nov 3 , 2019)	32,940	60,098	27,159	27,200	35,432	**-2,492**
5	(Nov 4-10 , 2019)	24,708	60,098	35,390	35,400	40,252	**-15,544**
6	(Nov 11-17 , 2019)	19,856	60,098	40,243	40,250	51,333	**-31,477**
7	(Nov 18-24 , 2019)	8,773	60,098	51,325	51,350	34,682	**-25,909**
8	(Nov 25- Dec 1 , 2019)	25,441	60,098	34,658	34,700	29,112	**-3,671**

Glukoz Şurubu

Weeks	Dates	Beginning Inventory	S	Calculated Order Quantity	Suggested Order Quantity	Actual Consumption	Ending Inventory
1	(Oct 7-13 , 2019)	82,188	60,098	0	0	2,515	**79,673**
2	(Oct 14-20 , 2019)	79,673	60,098	0	0	77,180	**2,493**
3	(Oct 21-27 , 2019)	2,493	60,098	57,605	75,000	59,809	**-57,316**
4	(Oct 28- Nov 3 , 2019)	17,684	60,098	42,414	50,000	17,171	**513**
5	(Nov 4-10 , 2019)	50,513	60,098	9,585	25,000	30,577	**19,936**
6	(Nov 11-17 , 2019)	44,936	60,098	15,162	25,000	22,020	**22,916**
7	(Nov 18-24 , 2019)	47,916	60,098	12,182	25,000	48,055	**-138**
8	(Nov 25- Dec 1 , 2019)	24,862	60,098	35,237	50,000	41,124	**-16,262**

Appendix E. The Results of the Heuristic Balanced Base-Stock Policy

Kristal Şeker

Weeks	Dates	Beginning Inventory	S	Calculated Order Quantity	Suggested Order Quantity	Actual Consumption	Ending Inventory
1	(Oct 7-13 , 2019)	78,945	75,670	0	0	39,483	39,462
2	(Oct 14-20 , 2019)	39,462	75,670	36,208	36,250	16,150	23,312
3	(Oct 21-27 , 2019)	59,562	75,670	16,108	16,150	27,172	32,390
4	(Oct 28- Nov 3 , 2019)	48,540	75,670	27,130	27,150	35,432	13,108
5	(Nov 4-10 , 2019)	40,258	75,670	35,412	35,450	40,252	6
6	(Nov 11-17 , 2019)	35,456	75,670	40,214	40,250	51,333	-15,877
7	(Nov 18-24 , 2019)	24,373	75,670	51,297	51,300	34,682	-10,309
8	(Nov 25- Dec 1 , 2019)	40,991	75,670	34,679	34,700	29,112	11,879

Glukoz Şurubu

Weeks	Dates	Beginning Inventory	S	Calculated Order Quantity	Suggested Order Quantity	Actual Consumption	Ending Inventory
1	(Oct 7-13 , 2019)	82,188	75,670	0	0	2,515	79,673
2	(Oct 14-20 , 2019)	79,673	75,670	0	0	77,180	2,493
3	(Oct 21-27 , 2019)	2,493	75,670	73,177	75,000	59,809	-57,316
4	(Oct 28- Nov 3 , 2019)	17,684	75,670	57,986	75,000	17,171	513
5	(Nov 4-10 , 2019)	75,513	75,670	157	25,000	30,577	44,936
6	(Nov 11-17 , 2019)	69,936	75,670	5,734	25,000	22,020	47,916
7	(Nov 18-24 , 2019)	72,916	75,670	2,754	25,000	48,055	24,862
8	(Nov 25- Dec 1 , 2019)	49,862	75,670	25,808	50,000	41,124	8,738

Appendix F. Calculations of the $P_{i,j}$ Values for Category A Raw Materials

Category A Items	Independent Base-Stock Policy				Myopic Base-Stock Policy			
	HC(i,j)	(α)i,j	(S/M)i,j	P(i,j)	HC(i,j)	(α)i,j	(S/M)i,j	P(i,j)
KRISTAL ŞEKER	51,865	0.6667	0.0668	0.4	59,989	1	0	0.6
GLUKOZ ŞURUBU	1,272	0.6667	0.2020	0.4	21,907	1	0	0.6
KOLI ROCCO NONAME	1,845	1	0	1	3,131	1	0	0.6
SITRIK ASIT	645	1	0	1	1,475	1	0	0.6
GLUCIDEX	3,492	1	0	1	5,972	1	0	0.6

Appendix G. Suggested Stock Policies

Category A	Stock Policy	Category B	Stock Policy	Category C	Stock Policy
Kristal Şeker	Myopic Base-Stock	For All Raw Materials	Independent Base-Stock	For All Raw Materials	Independent Base-Stock
Glukoz Şurubu	Myopic Base-Stock				
Koli Rocco Noname	Independent Base-Stock				
Sitrik Asit	Independent Base-Stock				
Glucidex	Independent Base-Stock				

Appendix H. The Main Menu of DSS

Appendix I. User Interface of Forecast System

Appendix J. Performance Analysis of Forecast Method

ITEMS	RMSE Values		
	Company's Current Forecast Method	Selected Forecast Method	Change
KRISTAL SEKER	113,349	96,764	-14.63%
GLUKOZ ŞURUBU	157,823	127,806	-19.02%
SITRIK ASIT	2,839	2,316	-18.42%
GLUCIDEX	2,344	1,781	-24.02%
AROMA CILEK BELL	135	102	-24.44%
AROMA KOLA SILESIA	34	24	-29.41%
AROMA ENKAP SEFTALI	21	15	-28.57%
BOYA RED	133	118	-11.28%
BOYA TURMERIC	23	17	-26.09%
TABAKA BASKILI PVC U POPS	13,482	12,079	-10.41%
KAPAK PIRAMIT SARI 4.0MM	90,668	67,329	-25.74%
AF(25CPP-103)	2,272	1,198	-47.27%
KOLI ROCCO NONAME CYL	27,683	25,969	-6.19%
AF U SOURISH	27	16	-40.74%
KUTU ROCCO YASSI	50,942	43,148	-15.30%
AF ROCCO YASSI	264	238	-9.85%
AF JOX YASSI	98	100	2.04%
KOLI U TOO NONAME	2,330	2,193	-5.88%
AF UTOO FRUITY	622	548	-11.90%
AF JOX BNBN CLASSIC	103	93	-9.71%
AF U XL MULT	105	17	-83.81%
AF ROCCO KLASIK	432	304	-29.63%
AF U FRESH PEPPERMINT	501	307	-38.72%
		AVERAGE	-23.00%

Appendix K. Performance Analysis of the Stock Policy

Categories	Company's Current Stock Policy	Our Stock Policy	Change Percentage
	Total Holding Cost	Total Holding Cost	
Category A	134,046	87,878	-34.44%
Category B	17,826	12,050	-32.40%
Category C	14,294	10,117	-29.22%

Appendix L. Comparison of DP for "Kristal Şeker"

Kristal Şeker					
	Holding Cost =	0.6516	Backorder Cost =	5	
Dates (in weeks)	Actual Consumption	Optimal Policy (DP)		Myopic Base-Stock Policy	
		Order Amount	Ending Inventory	Order Amount	Ending Inventory
(Oct 21-27 , 2019)	27,172	16,000	66,139	19,200	67,539
(Oct 28- Nov 3 , 2019)	35,431	27,000	46,708	21,100	51,308
(Nov 4-10 , 2019)	40,252	26,000	33,455	25,400	32,155
(Nov 11-17 , 2019)	51,332	40,000	8,123	40,650	6223
(Nov 18-24 , 2019)	34,682	50,000	13,440	56,250	12,190
(Nov 25- Dec 1 , 2019)	29,112	24,000	34,328	800	39,328
Holding Cost		131,751.10 TL		136,019.09 TL	
Backorder Cost		0		0	
Total Inventory Cost		131,751.11 TL		136,019.09 TL	

Appendix M. Comparison of DP for "Glukoz Şurubu"

Glukoz Şurubu					
	Holding Cost=	0.4	Backorder Cost=	5	
Dates (in weeks)	Actual Consumption	Optimal Policy (DP)		Myopic Base-Stock Policy	
		Order Amount	Ending Inventory	Order Amount	Ending Inventory
(Oct 21-27 , 2019)	59,808	75,000	17,684	75,000	17,684
(Oct 28- Nov 3 , 2019)	17,171	0	75,513	50,000	75,513
(Nov 4-10 , 2019)	30,576	50,000	44,936	0	94,936
(Nov 11-17 , 2019)	22,020	25,000	72,916	25,000	72,916
(Nov 18-24 , 2019)	48,054	50,000	49,861	25,000	49,861
(Nov 25- Dec 1 , 2019)	41,124	25,000	58,737	25,000	33,737
Holding Cost		127,859.83 TL		137,859.83 TL	
Backorder Cost		0		0	
Total Inventory Cost		127,859.83 TL		137,859.83 TL	

Müşteri Operasyonları için Dinamik Rotalama

Enerjisa Başkent Elektrik Dağıtım A.Ş.

Proje Ekibi

Berkay Çuhadar, Şebnem Manolya Demir, Mert Köksal, Ömer Can Serin, William Baris Stout, Muhammed Furkan Yıldırım, Burcu Yiğit

Şirket Danışmanı
Efe Can Okumuş
İş Uygulamaları Uzmanı

Akademik Danışman
Prof. Dr. Bahar Yetiş Kara
Endüstri Mühendisliği Bölümü

ÖZET

Enerjisa Başkent Elektrik Dağıtım A.Ş., iletim hatlarından bölgelerdeki son tüketicilere elektrik dağıtımını sağlamak için gerekli olan endeks okuma, yeni bağlantı, bağlantı kesme-açma, sayaç operasyonları ve ticari kayıp süreçleri gibi şebeke ve müşteri ile ilgili faaliyetlerden sorumludur. Alanda, farklı yetkinliklere sahip ekipler tarafından gerçekleştirilen ve bölgeye göre ayrılan çok sayıda, farklı ve müşteri ile ilgili operasyonlar bulunmaktadır. Bu proje, yasal ve operasyonel kısıtlamaları göz önünde bulundurarak, sadece günlük görevleri önceliklendirmekle kalmayıp aynı zamanda sahada katedilen toplam mesafeyi en aza indirerek saha planlama ekibine karar desteği sağlamayı amaçlamaktadır. Proje, bu bağlamda, gün içinde gelen talepler için dinamik yönlendirme sağlayan bir karar destek sistemi tasarlamayı içerir.

Anahtar Kelimeler: rota optimizasyonu, araç rotalama problemi, dinamik modelleme, öncelikli talep.

Dynamic Routing for Enerjisa Customer Operations

1. System Description

1.1. Company Information

Enerjisa, with more than 9,000 employees, is one of the top energy companies in Turkey. Founded in 1996, the company currently has approximately 8 million contracts, serving for more than 21 million users on its network of about 380,000 kilometers. Through its sub-companies, Enerjisa delivers service to 14 cities in 3 regions.

1.2. Operational Analysis

Enerjisa has mainly two departments for its customer operations: Commercial Loss and Customer Technical Services. While these departments have several units, the units of operations in the scope of this project are theft-illegal (responsible for detecting and then disconnecting electricity theft), disconnecting/connecting (responsible for disconnecting or connecting electricity due to various reasons), and meter operations (responsible for maintenance, inspection, and replacement of new meters due to various reason). Considering that these three units have sub-units, these three types of operations are carried out by specialized teams of minimum two employees via car.

In the current system of Enerjisa, the flow of orders is as follows:

1. The order notifications come up through multiple channels such as the "186" report line, the customer technical services, the detections and the reports from the field and the theft-loss tracking system.
2. The orders are created at the operation centers via the notifications.
3. Through the SAP system, the coordinators analyze, list, and then assign the orders to the field teams according to their locations. (The assignments are made by clustering the orders which are relatively close to each other.)
4. The field teams track the orders from their tablets online.

In addition to the regular flow of orders, there are also some important aspects to be considered. First, the field teams determine their own route of tasks (according to orders) manually, according to their experiences as soon as orders are assigned. When an order comes within the day, they are dynamically assigned to the teams in the field. For instance on a given day, at the Şereflikoçhisar operation center, only 45% of the orders arrive before the beginning of the working hours while at the Çubuk operation center, 90% of the orders arrive within the day. After that, the teams need to re-schedule their own routes manually again. The past data analysis provides information, on the numbers of field teams at an operation center and daily jobs of the field teams. Excel has been used to analyze the date and important findings which are relevant have been found. First, it has been found that disconnecting operations within the day are turned into connecting operations with 12%. To add on, the total number of the teams at each operation center has been obtained with their capabilities on operations. Furthermore, the past data gives an insight on each operation center's total number of orders and the types of the orders found by filtering the total daily operations. This, combined with the total number of teams, gives information about the scale of the problem for each operation center; which is such an important factor to be considered while structuring a solution approach.

1.3. Problem Definition

There are three main symptoms of the current routing of teams in Enerjisa. First symptom is about manual routing and experience-based intuition. This creates backtracking in the system, resulting in unnecessary movement, and consumes more time than necessary. Secondly, the costs of delays are not minimized efficiently in the current routing system, resulting in a group of teams not finishing tasks in given time limits. If there is a delay in connecting a connection, there is a penalty cost; and in the current system, this penalty cost is relatively high. Therefore, the system should have least amount of delays. In addition, several symptoms and constraints have been noted via making a field trip by the project team to observe daily operations of the teams in the field. The significant points from the field trip and the observations of the system are used to design the proposed approach.

The purpose of this project is to construct a system that can determine daily vehicle routes at each operation center. The vehicles have different competency levels and the order assignments should be made accordingly. Some of the orders have higher prioritization due to penalty costs mentioned above, which is necessary to be taken into consideration. Lastly, as the order list is not finalized before the working hours start and there are orders coming within the day, the routes should be adjusted dynamically with the addition of these new orders.

The problem is a version of Vehicle Routing Problem (VRP) model since all the demand points are to be visited with a particular number of vehicles and the aim is to minimize the total distance travelled. Since the teams have a planned period of working hours and all the demand points are visited in that period, the model is a Distance Constrained Vehicle Routing Problem (DCVRP). As an extension of the DCVRP model, a multi-objective decision making model may be utilized to minimize the distance and prioritize certain customers. Hence, the literature review is analyzed mainly in four parts (Appendix A).

2. Proposed Approaches

2.1. Exact Model

The nature of the problem is multi-objective: First objective is to minimize the total distance travelled. The second objective is to minimize the priority value, which is priority factor of demand points multiplied by their completion times. Based on the literature review results, the DCVRP model, which minimizes the total distance, was converted to a multi-criteria model by adding the priority objective.

The epsilon algorithm is applied to the multi criteria model by taking the priority function as a constraint and leaving the total distance function in the objective to be minimized. In this case, the aim was to find the non-dominated points. Those points are labeled as non-dominated when there cannot be any other point that is better in one objective and at least equal in the other objective. The non-dominated points found are demonstrated on the Pareto curve.

$$\alpha \cdot d + (1 - \alpha) \cdot p$$

The teams have different competency levels; every order cannot be met by every team. Incompetent teams should not be assigned to do the corresponding jobs. To avoid assigning tasks to incompetent teams, two additional constraints are added to the model. The final mathematical model has been coded in Java language and library of CPLEX is called to be used in solving the model.

Figure 1: Competency-Constrained Model

Parameters

$T : Total \quad service \quad time \quad for \quad each \quad vehicle$

$d_{ij} : Distacnce \quad between \quad node \quad i \quad and \quad j \quad i, j \in N$

$s_i : Service \quad time \quad of \quad node \quad i \quad i \in N$

$p_i : Priorty \quad of \quad node \quad i \quad i \in N$

$$z_{ik} = \begin{cases} 1 & if \quad team \quad k \quad is \quad capable \quad of \quad doing \quad the \quad task \quad in \quad i \quad \forall i \in N \quad \forall k \in K \\ 0 & otherwise \end{cases}$$

Decision Variables

$$X_{ijk} = \begin{cases} 1 & if \quad team \quad k \quad moves \quad from \quad i \quad to \quad j \quad \forall i, j \in N \quad \forall k \in K \\ 0 & otherwise \end{cases}$$

$t_i : Time \quad that \quad the \quad vehicle \quad leaves \quad node \quad i \quad \forall i \in N$

Model

$$min \quad (f_1(p,t), f_2(X,d))$$

$$f_1(p,t) = \sum_{i \in N} p_i \cdot t_i$$

$$f_2(X,d) = \sum_{i \in N} \sum_{j \in N} X_{ij} \cdot d_{ij}$$

$s.t.$

$$\sum_{j \in N} \sum_{k \in K} X_{ijk} = 1 \quad \forall i \in N \setminus \{1\}$$

$$\sum_{j \in N} \sum_{k \in K} X_{jik} = 1 \quad \forall i \in N \setminus \{1\}$$

$$\sum_{j \in N} \sum_{k \in K} X_{1jk} = m$$

$$\sum_{j \in N} \sum_{k \in K} X_{j1k} = m$$

$$t_i - t_j + T \cdot X_{ijk} \leq T - d_{ij} \cdot X_{ijk} \quad \forall i \in N, \quad \forall j \in N \setminus \{1\}, \quad \forall k \in K$$

$$0 \leq t_i \leq T - d_{i1} \cdot X_{i1k} \quad \forall i \in N, \quad \forall k \in K$$

$$\sum_{j \in N} X_{ijk} \leq z_{ik} \cdot M \quad \forall i \in N, \quad \forall k \in K$$

$$\sum_{j \in N} X_{jik} \leq z_{ik} \cdot M \quad \forall i \in N, \quad \forall k \in K$$

$$X_{ijk} \in \{0,1\} \quad \forall i,j \in N, \quad \forall k \in K$$

Moreover, realized route of an operation center in Şereflikoçhisar district is compared with the model results to see the performance of the solution. The results of Şereflikoçhisar region, have been generated using this model, with different alpha values. (Appendix B).

To implement this model to the current system of Enerjisa, a free-solver research was conducted. PuLP has been selected as it has been the most appropriate free solver among the candidates. It was seen that neither PuLP nor CPLEX was capable enough for large operation centers in a desired amount of time. Those regions' outputs in CPLEX have been tabled (Appendix C).

It is seen from this table that such regions would take waiting time of more than three hours and still will not be able to find the optimal value with 0% gap. This is simply not acceptable because these models should be solved every day to obtain the daily routes. The issue occurred by the solving speed/capacity of PuLP and CPLEX indicates that a faster and reliable solution is in need and this could be achieved by a heuristic approach.

2.2. Heuristic Algorithm

According to the project plan, constructing and coding a heuristic approach has been an alternative solution approach if a free solver could not be utilized. The decision support system is based on the heuristic algorithm which is generated in Java. Before applying the heuristic algorithm, the distance matrices of daily operations are created Python via taking the input of each demand point's coordinates. Both the service time and priority data are calculated in the algorithm using predetermined (by the company) service times and priority coefficients for each order type before the heuristic algorithm starts. Then, the following procedure is applied:

1) Field teams are clustered according to their competency levels (Appendix D).
2) The heuristic procedure is applied for each cluster hierarchically starting from the most competent teams (type 7) to the least competent ones (types 1-2-3), in order to make use of the capabilities of more competent teams.
3) For the competency level at hand, a "node list" is created by selecting the tasks that the teams at that level can handle from the set of "unassigned" nodes.
4) With all these tasks in the "task list", the Travelling Salesman Problem (TSP) is solved by applying the Nearest Neighbor (NN) method.
5) The nodes in the TSP are assigned to the teams until the team's capacity is full. This is applied for all teams in the that competence level.
6) The tasks assigned to the teams are marked "assigned" and the others are marked as "unassigned".
7) For possible improvements in travel times, the 2-Opt method is applied to each team's route.
8) Priority-based 2-Opt improvement is applied to take priorities into consideration in the routes (Appendix E).
    a. It starts from the last node (D) of the first route. This node is compared with each other of the remaining nodes of that route. (starting from the first node (A)).
    b. If this last node D's priority is larger than the compared node A, a "new route" is constructed by reversing the portion of the route between these two nodes.
    c. For this new route, the ratio: change in distance divided by the change in priority, is calculated.
    d. If this ratio is smaller than a predetermined alpha value, the route is replaced with the "new route".
    e. This is applied for each node.

The reason for starting the assignments from the most competent team (type 7), is to avoid the visitation of the same node with different teams in the same day. A different type of order on the same location is entered in the system with the same coordinate, twice. As nearest neighbor algorithm is used to construct the route, the nodes are routed after one another with a travel time of zero. If the algorithm would have started from the least capable teams, teams capable of doing one task, these two orders would have been in the different "order list" and be assigned to different teams. But as the algorithm starts with the most capable team, it is more likely that these are assigned to the same team.

Alpha value is the tradeoff level that can be increased from the distance for a unit improvement in the priority value. A sensitivity analysis has been conducted to guide the company in the alpha value decision. The company will determine their own alpha values while implementing the heuristic approach into their own operations because at some operation centers, priority has more value than distance and therefore the alpha value is implemented as a parameter. The heuristic results for an arbitrary alpha value 0.0005, has been tabled (Appendix F).

The algorithm determines the closest node that is assigned to a team with a competency level that can handle the new task, from the list of uncompleted tasks. Once this closest node is determined, the new node is assigned to the same vehicle and placed to be visited right after the "closest node". The algorithm also updates the schedule of this vehicle. Also, before running this algorithm, the information on distance is updated with the addition of new orders. As discussed with the company, the tasks arriving within the day are mostly desired to be completed within the day, and the company stated that they would prefer to handle them in a cost-effective way instead of scheduling them to be completed after that day's task list. Re-running the algorithm was also an option discussed earlier, however that approach changes the vehicles' task list significantly, therefore yields to longer distances travelled, backtracking and creates a less stable system.

2.3. Performance Analysis

By looking at these results, it is seen from the CPU (speed), the algorithm generated can tackle larger size problems in few seconds, which was the goal behind generating such heuristic algorithm. The result comparison for the heuristic algorithm and the exact model is tabled (Appendix G). It is beneficial to compare the heuristic results with the realized results of Enerjisa as well to see the difference (Appendix H). Not only the heuristics algorithm has higher CPU speed, but also it gives remarkably better results than the current system of Enerjisa. To add on, the routes are created for Şereflikoçhisar region by the heuristic algorithm, exact model, and the realized route of Enerjisa (Appendix I).

The heuristic approach is seen to be remarkably faster than the exact model approach while it can solve larger size operation centers as well. It is known that no technical difficulties resulting in not obtaining the daily vehicle route overnight will be encountered on the heuristic approach, it is reliable to give a reasonable solution in comparison with the current system of Enerjisa. As it is required to run the decision support system over night and use the dynamism algorithm within the day, a faster decision support system will be more reliable for bigger operation centers. Therefore, the heuristic algorithm generated forms the basis of the decision support system.

3. Decision Support System

The decision support system that has been provided to the company should be user friendly; therefore, the heuristic algorithm is compressed in an executable java file, which takes only txt files that obtained from an Excel using Macros and Python. For the heuristic to run accordingly, the Excel file should contain all information to be an input to the algorithm. These are total number of orders, number of teams for each competency level, working hours of each team, and alpha value. The information is structure in this Excel file in a certain manner that is used by the company and exported into txt files using Excel Macros. The data file containing priorities and service times, are generated automatically by the Java code to be later used by the algorithm itself. The distance data is created in Python and kept in a txt file which is also used in the algorithm.

The decision support system consists of two executable Java files: Commodity and Dynamic. By running the Commodity executable file, the daily route for each team is obtained. The output gives the sequence of orders that each team should have. This is generated by the heuristic algorithm and provides the completion times for each assigned order (Appendix J).

To make the decision support system dynamic, the input Excel file needs to be updated. The updated input file should have the new order’s order code and coordinates. Moreover, parameters such as number of new orders and number of finished jobs for each team should be stated for the dynamic use of the algorithm. After the input Excel file is updated, rest of the input files, such as the distance data and priorities, are updated automatically by running the executable file Dynamic. The distance data should also be updated including the coordinates of the new orders. Using the heuristic algorithm, the Dynamic executable file will provide the updated routes, including the newly added ones throughout the day, for each team (Appendix K).

4. Project Achievements

The performance measure of this project is the total distance travelled by all vehicles each day. The heuristic algorithm uses nearest neighbor algorithm to shorten the distance and 2-Opt algorithm is applied to each vehicle’s route to further decrease the distance. The priority algorithm creates a balance between the need to prioritize certain tasks and the distance traded-off.

The dynamic structure reduces backtracking as it assigns the new order based on the orders that have not been completed yet, the future orders, so the teams do not necessarily go back to the orders they have visited before. The new order is placed after the closest order to minimize the distance to be travelled with the addition of their new order. There are also some elements in a parametric structure such as the number of daily orders, the number of teams, and the working hours, in the decision support system, and these variables are to be input by the company daily. This way, the number of daily orders adapts to the number of orders of the operation center that changes every day while the number of teams can be used in every operation center. In addition, team occupancy can be adjusted by the working hours given for each team every day.

By determining the routes of field teams according to the output of the decision support system using different values of parameters, alpha and working hours each day, and by analyzing the realized route, company will set the appropriate values for these parameters. Once these test runs are complete and the routes of the field teams are realized according to the decision support system's output, the integration process will start. This process includes the integration of the decision support system to the computers of the coordinators for each operation center and the field team's tablets. Also, it is in company's plans to obtain the Google API system to be used in the distance information gathering process. To ease the implementation process, the decision support system has been transformed into a user-friendly version.

In conclusion, the algorithm of the decision support system speeds up the system by giving the daily route in minutes. Besides, the decisions made by the algorithm from inputs to outputs provide automation. In the end, as the objective of the algorithm is to minimize the total distance travelled by the field teams, the decision support system is able to determine the routes in a way that performs better in terms of this performance measure (total distance travelled) than the method that was used by Enerjisa. While doing this, it was crucial to reduce the possible delays on the highly prioritized orders. The decision support system has a priority algorithm embedded to make sure that this prioritization is conducted in a standardized manner. All these result in a decision support system that takes all the necessities into account.

References

[1] E. M. T. O., A. H. E. Z., and M. G. E., "Literature Review On The Vehicle Routing Problem In The Green Transportation Context," *Luna Azul*, no. 42, pp. 362–387, 2015. [Online]. Available: ScienceDirect, https://www.sciencedirect.com. [Accessed: Dec 28, 2019].

[2] G. Laporte, “The Vehicle Routing Problem: An overview of exact and approximate algorithms,” *European Journal of Operational Research,* vol. 59, pp. 345-358, 1992. [Online]. Available: ScienceDirect, https://www.sciencedirect.com. [Accessed: Oct 19, 2019].

[3] J. F. Cordeau, et al., “A guide to vehicle routing heuristics,” *Journal of the Operational Research Society*, vol. 53, issue 5, pp. 512-522, December 2017. [Online]. Available: Taylor & Francis Online, https://www.tandofline.com. [Accessed: Oct 20, 2019].

[4] Z. Yang, et al., “Ant Based Solver for Dynamic Vehicle Routing Problem with Time Windows and Multiple Priorities,” *IEEE Congress on Evolutionary Computation*, pp. 25-28, May 2015. [Online]. Available: ResearchGate, https://www.researchgate.net. [Accessed: Oct 19, 2019].

[5] S. L. Smith, “Dynamic Vehicle Routing with Priority Classes of Stochastic Demands,” *SIAM Journal on Control and Optimization*, vol. 48, issue 5, pp. 3224-3245, 2010. [Online]. Available: DSpace@MIT, https://www.dspace.mit.edu. [Accessed: Oct 20, 2019].

[6] H. N. Psaraftis, “Dynamic Vehicle Routing Problems,” *Elsevier Science Publishers B.V.*, 1988. [Online]. Available: MarTrans, https://www.martrans.org. [Accessed: Oct 20, 2019].

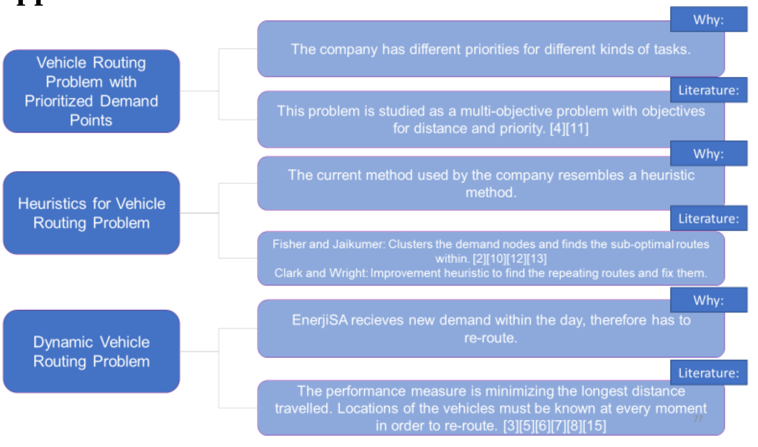
[7] A. Larsen, et al., “Partially dynamic vehicle routing—models and algorithms,” *Journal of the Operational Research Society*, vol. 53, issue 6, pp. 637-646, June 2002. [Online]. Available: SpringerLink, https://www.link.springer.com. [Accessed: Oct 19, 2019].

[8] V. Pillac, et al., “A review of dynamic vehicle routing problems,” *European Journal of Operational Research*, vol. 225, pp. 1-11, 2013. [Online]. Available: Elsevier, https://www.elsevier.com. [Accessed: Oct 20, 2019].

[9] G. Wang and Y. Yu, "An Improved Ant Colony Algorithm for VRP Problem," 2010 Third International Symposium on Intelligent Information Technology and Security Informatics, Jinggangshan, 2010, pp. 129-133.

APPENDICES

Appendix A. Literature Review



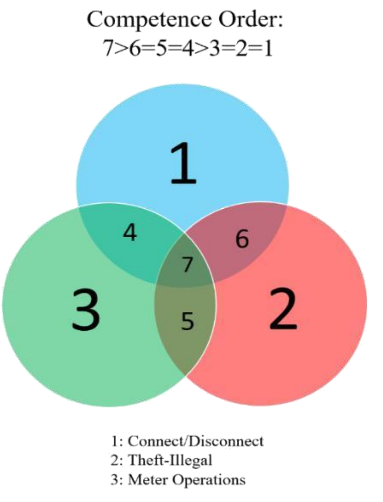
Appendix B. Exact Model's Performance for Different Alpha Values

	Realized Route of Enerjisa	Exact Model (α=1)	Exact Model (α=0)
Total Distance (km)	186.5	146.8	170.2
Task Time (min)	89.0	89.0	89.0
Travel Time (min)	194.0	130.0	172.0

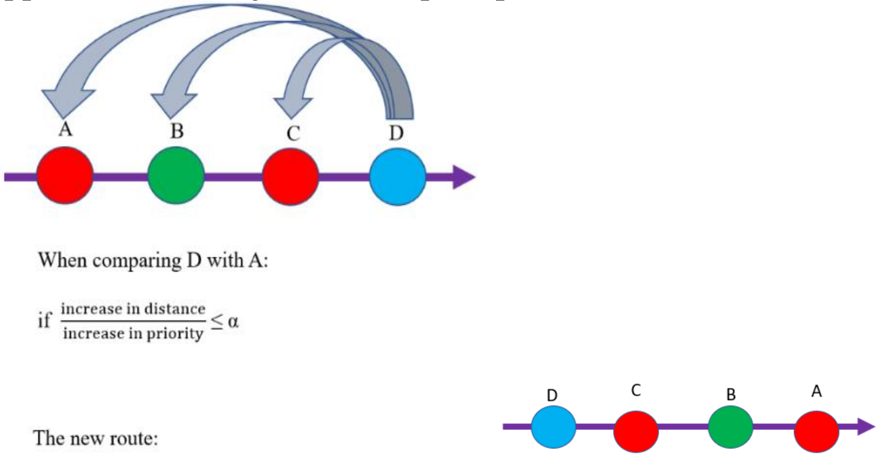
Appendix C. Exact Model's Results

Operation Center	Size	Number of Teams	Objective Value	Gap	CPU (speed)
Şereflikoçhisar	20	3	77.39 mins	0%	~5 sec
Şile	32	2	230.66 mins	6.27%	Over 3 hours
Düziçi	37	5	140 mins	5.45%	Over 3 hours
Kahramankazan	54	5	252.9 mins	2.77%	Over 3 hours

Appendix D. Competence Levels



Appendix E. Priority-Based 2-Opt Improvement



Appendix F. Heuristic Algorithm's Results

Operation Center	Size	Number of Teams	Objective Value (mins)	CPU (Speed)
Şereflikoçhisar	20	3	140	2 seconds
Şile	32	2	222.106	6 seconds
Düziçi	37	5	153.4	4 seconds
Kahramankazan	54	5	216.003	11 seconds

Appendix G. Heuristic Algorithm's Performance

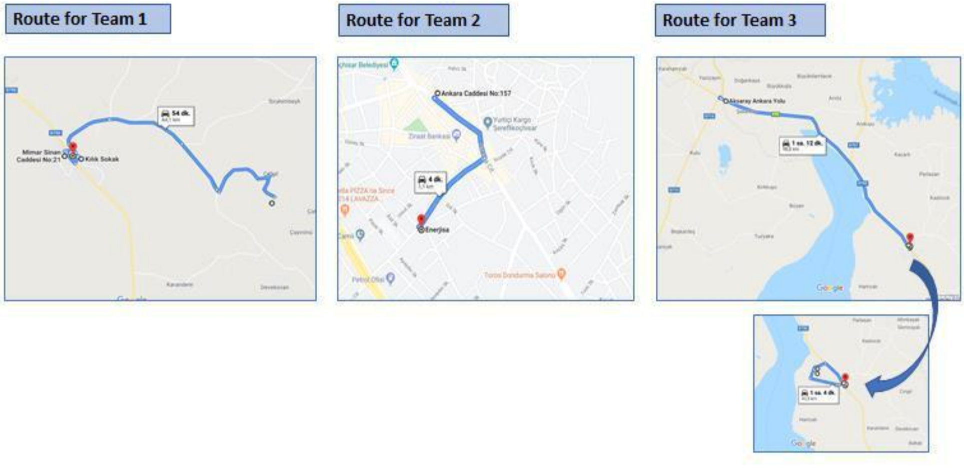
Op. Center	Size	# of Teams	Ex Model Obj. [X]	Exact Model Gap	Exact CPU (Speed)	Heuristic Obj. (mins) [Y]	Heuristic CPU (Speed)	Gap [(Y-X)/Y]
Ş. koçhisar	20	3	77.39	0%	~5 sec	140	2 sec	44.721%
Şile	32	2	230.66	6.27%	Over 3 hours	222.106	6 sec	-3.851%
Düziçi	37	5	140	5.45%	Over 3 hours	153.4	4 sec	8.735%
K.kazan	54	5	252.9	2.77%	Over 3 hours	216.003	11 sec	-17.08%

Appendix H. Decision Support System's Performance

Operation Center	Size	Number of Teams	Realized Enerjisa Results (mins)	Heuristic Obj. (mins)	Difference
Şereflikoçhisar	20	3	194	140	-38.571%
Şile	32	2	268	222.106	-20.663%
Düziçi	37	5	193	153.4	-25.815%
Kahramankazan	54	5	352	216.003	-62.960%

Appendix I. Route Maps

Realized Routes of Enerjisa



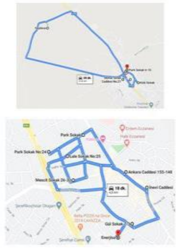
Exact Model's Routes

Route for Team 1

Route for Team 2

Heuristic Algorithm's Routes

Route for Team 1



Route for Team 2

Appendix J. Completion Times of Assigned Orders

output - Notepad

File Edit Format View Help

```
Order Code      Departure Time  Car
000658575227    8:25            30420075
000658591383    8:31            30420075
000658575273    9:00            30420075
000658577507    9:19            30420075
000658576951    9:23            30420075
000658577839    9:28            30420075
000658591458    9:33            30420075
000658578070    9:37            30420075
000658587850    9:46            30420075
000658576374    9:50            30420075
000658584382    10:51           30420075
000658594508    10:58           30420075
000658577568    11:02           30420075
000658592840    11:07           30420075
000658577296    11:11           30420075
000658591545    11:17           30420075
000658577944    11:21           30420075
000658576714    11:25           30420075
000658575494    11:30           30420075
000658577583    11:36           30420075
000658575600    11:40           30420075
000658575818    11:44           30420075
000658581055    11:49           30420075
000658584304    11:54           30420075
000658577312    11:58           30420075
000658599005    13:03           30420075
000658577517    13:07           30420075
000658583449    13:13           30420075
000658583234    13:18           30420075
000658577740    13:22           30420075
000658589037    13:27           30420075
000658577409    13:31           30420075
000658577102    13:35           30420075
000658575360    13:39           30420075
000658584475    13:45           30420075
000658576969    13:49           30420075
```

Appendix K. Route Update with Added Nodes

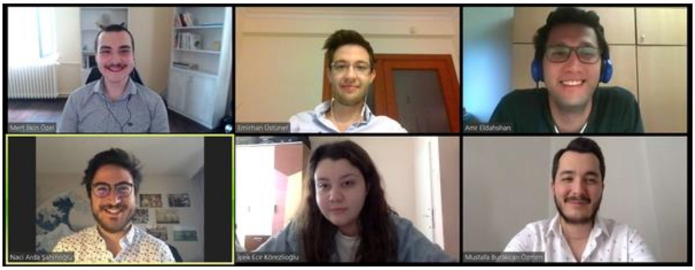
dynamicoutput - Notepad

File Edit Format View Help

```
-------------------   ADDED NODES  ----------------------
Order Code      Car             Sequence
000658599429    30420075        4
000658599617    30420075        5
000658599665    30420071        3
000658600743    30420075        20
000658600939    30420071        59
Order Code      Departure Time  Car
000658577839    10:25           30420075
000658591458    10:31           30420075
000658578070    11:00           30420075
000658587850    11:19           30420075
000658599429    11:35           30420075
000658599617    11:39           30420075
000658576374    11:55           30420075
000658584382    11:59           30420075
000658594508    12:08           30420075
000658577568    12:12           30420075
000658592840    13:13           30420075
```

Depo İçi Raf Adresleme Sisteminin Geliştirilmesi

Erkunt Traktör Sanayi A.Ş.



Proje Ekibi
Amr Eldashan, İpek Ece Körezlioğlu Mert İlkin Özel, Burakcan Özmen, Arda Şahinoğlu, Emirhan Üstüner

Şirket Danışmanı
Dilara Baykal Bilici
Kurumsal Performans Uzman Mühendisi

Akademik Danışman
Doç. Dr. Arnab Basu
Endüstri Mühendisliği Bölümü

ÖZET

Bu projede Erkunt Traktör A.Ş.nin depo için raf adresleme sisteminin geliştirilmesi üzerine çeşitli iyileştirilmelerde bulunulmuştur. Toplayıcılar, depo içinde kendilerine verilen listelere göre SPS, Kanban ve Jundate olmak üzere 3 farklı kategoride ürün toplamaktadır. Toplayıcıların yürüme mesafelerinin minimizasyonu için ürünlerin kullanım frekansına göre yeni bir depo içi raf adresleme algoritması geliştirilmiştir. Ürünlerin raflara atanma önceliği, kullanım sıklığına göre belirlenmektedir ve parça boyutları dikkate alındığından, algoritma Sırt Çantası problemini temel almaktadır. Model, Excel VBA ile uygulamaya geçirilmiştir.

Anahtar Kelimeler: dinamik raf adresleme, depo iyileştirme, sipariş toplama, yürüme mesafesi minimizasyonu

Warehouse Placement Improvement Algorithm

1.General Information about Company

Erkunt Traktör was established in September 2003 by Zeynep and Tuna Armağan using only indigenous capital and it has a market share of approximately 30% in Turkey. Erkunt Tractor was built on an area of 45.000 m2 with 14.500 m2 being indoors with 4 branches and 1 main factory which is located in Sincan, Ankara. They have become the 3rd largest in terms of production in 6 years within the tractor sector because of their fuel efficiency and 3 year warranty. Erkunt Traktör has recently made a partnership in 2017 with an Indian company called Mahindra. With this new partnership they have started producing customized tractors for the first time in Turkey and launched a new export brand named ArmaTrac which is now being exported to more than 17 countries overseas.

2.Problem Definition and System Analysis

2.1. System Analysis

During production in order to feed the assembly line, products need to be collected from the warehouse but since tractor production requires so many part with various functions and sizes, Erkunt Traktör divides their order picking lists by 3 categories. The materials are separated according to their characteristics and tracked by specific material handling systems. Erkunt classifies the material handling using 3 methods:

- **SPS:** The parts of the product, which are specific to that model of tractor. SPS- addressed parts and assembly line move at the same speed. SPS parts are usually large in size and expensive. One can roughly state that SPS system parts are the ones which make all the differences among different models of the products.
- **Jundate:**Parts that are addressed with the Jundate feeding system are usually those which commonly exist in all the products (vital parts), such as: fuel tanks, oil filters etc. Jundate delivers parts that do not move at the same speed as the assembly line. Some parts can be slower or faster depending on the demand feeding rate(s). When it’s time to assemble, the part(s) would be taken and assembled with the accumulated product.
- **Kanban:** It delivers parts that are usually smaller in comparison to the other systems however, Kanban materials are ordered in larger quantities and come in a wide range of varieties, handling Kanban material becomes more complex due to the variations. They are triggered by the line worker(s) and 1000-1500 parts per tractor can be fed by the Kanban system.

The picking lists of the parts assigned to the mentioned methods are created daily and distributed to the operators who pick the parts, then transferred to the assembly line. Each operator has a picking cart, which has the capacity of 8 different sections that parts are categorized based on their assembly target. According to the data our team acquired recently, there are 8 different categories of SPS carts and there are 2 different types of Jundate carts operated by the warehouse operators. The scope of the project is limited with the inner section of warehouse.

2.2. Problem Definition

One of the most crucial problems that Erkunt faces is the walking distance inside the warehouse. Workers first receive a list of items necessary for the assembly line to be delivered to the factory. Workers start from the entrance point in the aisle, walk inside the warehouse according to an optimal routing algorithm, pick up the item(s) required then drop them off at the drop off point and repeat. The walking distance is known as traveling time. The distance walked can vary from list to list and item to item. This is because the items stored inside the shelves can change. Items ordered frequently together should be placed next to each other. Two different factors affect the walking distance like shelf location and item location. The company explicitly asks to focus the item relocation in order to save costs and it is asked to minimize the sum of all the trips for the whole year. This can be achieved by using a dynamic slotting algorithm that assigns a specific location address to every part in the warehouse. When commonly used items are placed close enough, workers will walk less distances.

According to the company officials interviewed, the problem occurs because of complexity, irregularity and unnecessary repetition of the usage of the paths of the materials while they are being transported to the assembly line(s). Throughout our visits to the company, the engineers told that their former warehouse layout was based on the quantity of the parts that were used. Specifically, commonly used parts were placed closer to the aisles in order to pick them in a quicker way. However, there were also rarely used parts that were used in some models which were a minority. When those rarely used parts came into the system, their former system broke up and the newcomer parts were placed, randomly. Since their system became inefficient, the operators work in the warehouse were having an unoptimized route to pick up the materials. All shelves which workers can reach by their hands are almost full. Thus, if new parts come to this warehouse to store, we have to put these items to shelves which workers cannot reach and they require the assistance of stairs which costs even more time loss.

As a given constraint, order picking methodology can not be changed and the warehouse shelf capacities are full so buying extra shelves can not be offered. Also changing the layout is not preferred by the company. Hence the main objective of minimizing the total walking distance can be achieved by reallocating the items in warehouse.

2.3. Scope and Objectives of the Project

One of the main objectives for the company and for us is the applicability of our suggested model in the client's production system. Company demands a system that would result in an outcome which would decrease the number of workers steps in warehouse. Erkunt Traktör requested a model that inputs the required parts and model of the tractor. This model will allocate the shelves where the parts are located and the workers / operators will collect the given parts accordingly.

The main objectives of this project are the following:

- Initially, arranging the shelves by changing the material location in the warehouse to decrease the workers steps therein.
- If this arrangement can be applied by the company and beneficial to them, then developing a model to allocate shelves according to their annual demand forecast. Thus, this model can be used at the beginning of every year by Erkunt.
- Developing and thinking about dynamic shelf area where parts are located randomly in the same warehouse.
- Focusing on the removal of unnecessary repetitions occurring at the warehouse while collecting parts and the frequencies of the necessary parts.
- The company expects to reflect our ideas through commonly used software(s) such as Microsoft Excel. Since they use daily constructed plans for their production, it needs to be as user-friendly as possible for them to intervene and take actions if a casualty or latency occurs.

3.Proposed System

The proposed system inputs are required to use the complete pivoted MRP list of yearly production planning and the parts, the complete Bill of Materials (BOM) requirements of the products, the complete list of SPS cars and SPS categories, a pick frequency table of part constructed with the combination of BOM and annual production plan. Then, the system output will be the aisle and shelf ID's where part is located. In our proposed system, the main idea is that the materials are sorted in each category in descending order according to their picking frequencies. The distance from each cell to the exit of the corridor are collected and we sorted these distances in ascending order. In this manner these matched materials with high pick frequencies to cells with low distances. The distance has been calculated using the current layout in the warehouse. When an order list arrives, the distance will be based on the number of items and their locations. In the warehouse, Manhattan distance is used to scale the distances between points. Manhattan Distance is the distance between two points measured along axes at right angles, in a plane with p_1 at (x_1, y_1) and p_2 at (x_2,

y_2), it is $|x_1$-x_2 | + $|y_1$- $y_2|$ (Black 37). In order to implement this system and get results for "Dynamic Slotting Optimization algorithm" Microsoft Excel is defined to be the best software considering the compatibility with existing company operations structure. This algorithm uses the correlation of items and their frequency as input in order to find the optimal spot for each item.

3.1. Algorithm: Dynamic Slotting Optimization

Multiple items in an order are picked from their perspective locations in one tour that begins at the entrance door, visits these locations in a specified sequence and returns to this door and the tour is arranged such that its length (total travel distance) is minimized. It's more meaningful to place items that have a high turnover in prime locations near to the output door with the correlated strategy of "Order Oriented" (Mantel et al 308). Let us have set A that includes all the parts in the warehouse. We are trying to assign these parts to available slots under the restriction of some set of rules. A figure for slots can be found in Appendix A. The output that will be taken from this heuristic will be the location of the parts. For this algorithm we have N slots in a single zone and M zones in a warehouse. Also we define $C_{i,j}$ and f_i as $C_{i,j}$ given to be correlational weight between part i and j and f_i given to be picking frequency for part i. For a given part set, $A = 1,2,3 \ldots K$ the procedure is as follows:

Step 1: Initialize the zone (in our case the aisle) and the slot index. Let slot 1 be the current slot and the slot index becomes $n = 1$. Let zone 1 be the current zone and the index is $m = 1$.

Step 2: For current zone $n = 1$ choose an unassigned part from the set A, with the highest usage frequency. Let us define a set $A_0 = i$, maximize f_i , i∈ A. Select the part with the highest frequency from the set A_0 and place it in the current index. Let the selected part i as the current part and update $n = n + 1$. Delete the selected part i from A_0 and set A.

Step 3: Select all unassigned parts that have positive correlation with part i and form a set by using them. Let us call that set A_1= j, C_i , $j > 0$, $i, j \in A$. Assign parts with the highest correlation weight nearest to part i and place the part that has the lowest correlation with part i to the furthest slot (the placement strategy for the parts are represented in Appendix A). The parts are placed in an S-shaped pattern and we start placing from the bottom. If the number of correlated parts exceeds the number of slots in a certain cell, move to the next cell. Delete all assigned parts from the A_0 and set A. The remaining parts that have zero correlation with part i forms a set called A_0= j, C_j, j = 0, $i, j \in A$. Continue until A_1= ∅ and update slot index as $n = n + 1$ total number correlated parts with part i.

Step 4: If A_1= ∅ and A_2= ∅, it means that there is no correlated part left for part i so we need to choose another part. To choose another part go back to Step 3 but only if $n < N$. Now, if $n = N$ it means that zone m does not have any available slots left, update $m = m + 1$ and go to Step 2. If $A_1 = \emptyset$ and $A_2 = \emptyset$ or $m = M$, it means that all parts are assigned, so go to Step 5.

Step 5: Stop.

4. Integration and Implementation

The Warehouse Placement Improvement System was created by integrating the heuristic model into the interface used in Excel VBA. VBA language is preferred for users to easily transfer data to the system and to work in harmony with other programs used by the company. The user interface consists of two parts where login and annual production inputs will be made. These interfaces are shown in Appendix B. After logging in, the user is expected to enter the tractor codes and tractor numbers that they will produce in line with the company's annual production plans. After this data is entered into the system, the heuristic model we create finds the Bill of Materials required for the production of each tractor and calculates the frequencies according to the capacity of these parts to be transported in SPS cars. These frequencies are placed in the frequency table according to SPS categories. A proximity score is assigned to the pieces according to their frequency. For each SPS category, the parts with high frequency are lined up closest to the corridor exit. Warehouse Placement Improvement System takes into account the database consisting of four main parts while doing this operation. In this database, vehicle codes, product receipts, the amount of parts required for the vehicles planned and the width of the parts are kept. Appendix C shows these database parts. This database also lets the company to include new arriving parts from new suppliers. Recorded parts are shown to the user with sub module database. See Appendix D. Thus, there is space for 300+ new items to be injected in the current database. Warehouse Shelf Addressing System provides a detailed inspection and inspection table with the code, width, cell and shelf numbers of the parts, and a new layout plan with part codes and new location codes. This table is shown in Appendix E.

5.Validation

We performed hypothesis testing to compare the means of the sample data and from the means have an understanding on the whole population. Hence, we selected sample of lists that are representative of the whole population and calculated the walking distances before and after the algorithm has been used. All SPS lists that are give to the order pickers by the production planning department are the subsets of major lists divided and labeled considering the name of on going production process. So, in order to cover all possible SPS lists we take the largest three SPS lists used in the company which are; “Transmission”, “Pre-Painting” and “Post-Painting”. As shown in Table 1, each item in the lists is inspected individually and percentage difference in the walking distance of the part from door to the part location before and after algorithm run is gathered.

First, by using t-test we make sure that our results are in the desired direction. We defined our null hypothesis (t_0) as the desired decrease in the total walking distance which is 3%. Using historical walking distance of each item and new walking distance after the run of Dynamic Slotting Algorithm, we are able to compare the change in the cumulative walking distance of chosen three list sample.

Transmisyon				Boya Öncesi				Boya Sonrası			
Parça ID	Eski Lokasyon - Çıkış	Yeni Lokasyon - Çıkış	Yüzdelik Değişim	Parça ID	Eski Lokasyon - Çıkış	Yeni Lokasyon - Çıkış	Yüzdelik Değişim	Parça ID	Eski Lokasyon - Çıkış	Yeni Lokasyon - Çıkış	Yüzdelik Değişim
100575	62	6.60	89%	100326	35	19	45%	100503	31.80	19	40%
100583	94	6.60	93%	100330	28	19	30%	100531	28.60	19	33%
100861	83	6.60	92%	100648	33	7	80%	100537	28.20	19	32%
100895	65	6.60	90%	100657	47	7	86%	100562	17.40	7	62%
101269	87	6.60	92%	100812	33	7	80%	100563	17.40	7	62%
101486	58	6.60	89%	100814	52	7	87%	100614	17.40	7	62%
101487	58	6.60	89%	100833	46	7	86%	100828	6.60	7	0%
101753	65	6.60	90%	100834	28	7	76%	100829	14.20	7	54%
101851	69	6.60	90%	100835	36	7	82%	100852	6.60	7	0%
102171	87	6.60	92%	100845	33	7	80%	100858	6.60	7	0%
102198	90	6.60	93%	100938	45	7	85%	100862	24.60	7	73%
102290	94	6.60	93%	100944	48	7	86%	100868	24.60	7	73%
102386	87	6.60	92%	100946	45	7	85%	100869	24.60	7	73%
102443	79	6.60	92%	100949	46	7	86%	100881	21.40	7	69%
102472	83	6.60	92%	100951	47	7	86%	101005	10.60	7	38%
102498	72	6.60	91%	101006	45	7	85%	101268	14.20	7	54%
102536	69	6.60	90%	101007	41	7	84%	101328	21.40	7	69%
102595	72	6.60	91%	101037	28	7	76%	101374	10.60	7	38%
102611	79	6.60	92%	101080	58	7	89%	101459	10.60	7	38%
102612	80	6.60	92%	101127	38	7	83%	101507	10.20	7	35%
102726	69	6.60	90%	101128	38	7	83%	101508	10.20	7	35%

Table 1: Walking distance database consisting three lists and percentage differences

Input Data Set	Historical Output	Algorithm Output	Difference
Transmission	31635.33	11858.87	19776.46
Pre-Painting	15070.61	8695.61	6375
Post-Painting	10282.21	12562.41	-2280.2

Table 2: Comparison of the total walking distance of each order picking list before and after the algorithm run used in t-test.

$$\bar{d} = 7957.09$$

$$Sd^2 = 123501311.30$$

$$t_0 = 1.24$$

$$t_{\alpha/2,n-1} = 6.21$$

$$t_{\alpha/2,n-1} < |t_0|$$

As a result of the t-test, we conclude that, we can not reject our null hypothesis which is the verification of the decrease in the walking distance by 3%. In addition to the t-test, we calculated the average decrease in the total walking distance by using moving averages method. Percentage changes in each test is given in Table 2. Mean decrease of each SPS list is calculated and with those averages, the mean decrease in the total walking distance is calculated as 11.53%.

Transmission	61.80%
Pre-Painting	33.94%
Post Painting	-61.17%
Overall	11.527%

Table 3: Percentage change in each SPS list with overall average decrease

6. Contribution to The Company

The results showed that total walking distance of workers in the warehouse has been decreased by 11.53% on the basis of three major order picking lists. This decrease can be interpreted as a decrease in the man-hours needed for warehouse to function properly. Therefore, after the implementation of The Warehouse Placement Improvement System, newly gained man-hours from the warehouse can be used for other operations in the company where it is needed. Considering time study conducted in Fall 2019, we obtained 54 minute reduction in transmission, 210 minute in pre-painting and 69 minute in post painting order picking operations. Since creation of The Warehouse Placement Improvement System did not require any investments, the company gained additional man-hours without spending any money. Because of the COVID-19, the company is currently non-operating and cost benefit analysis will be conducted by the company after the new schedule of 2020 is finalized. In addition, the decreased total walking distance also mean potentially, there will be less accidents in the warehouse since flow across the the warehouse will be smoother with lesser traffic.

7. Conclusion

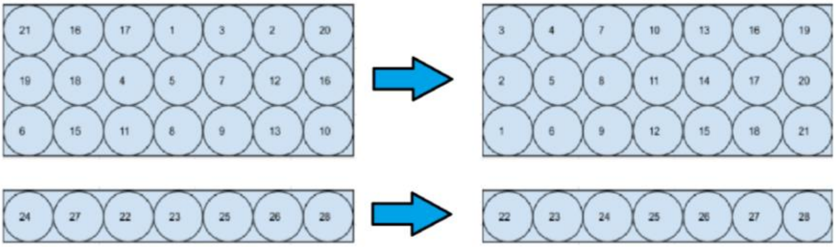
Given in Section 6, the decrease in the operating hours of SPS list order picking operations is verified by the company. Our team successfully met Erkunt Traktör's expectations, which will contribute to the operating efficiency of the warehouse, thereby helping improve the company's overall efficiency. Even though the application our solution cannot met with real time data, results are verified and the decrease of 11.53% is accepted as the successful result of the project. In addition, in Summer 2020, algorithm domain will be expanded to cover 10 more SPS lists that will be renewed according to the new list ingredients since most of suppliers stopped the part traffic due to coronavirus break.

REFERENCES

Paul E. Black. (1998). *Dictionary of Algorithms and Data Structures* [Online]. Available: xlinux.nist.gov/dads/HTML/manhattanDistance.html. [Accessed: 09-May-2020].

Mantel et al. (2007). *Order oriented slotting: A new assignment strategy for warehouses*. European Journal of Industrial Engineering. 1. 301-316. 10.1504/EJIE.2007.014689.

APPENDIX A



*Item number 1 is the item with highest frequency. Other numbers are the numbers with highest correlation with item number 1.
**Heavy items are sorted with themselves since only the first shelf can contain heavy items.

APPENDIX B

Please fill in the list of vehicle codes and their corresponding quantity.

Vehicle Code	Quantity
A05023PT103A/S2/6,5-16x13,6-28	30
A05043PM023A/S01/280/70-R20x380/70-R28	25
A12544KT0440D/+L2/360/70-R28x420/85-R38	15
E08044KT033A/MS1/11,2-24x16,9-30	40

103463
A05023KT023A/SS01/7,5-16x14,9-30
A05023PT023A/SS00/7,5-16x380/85-R28
A05023PT103A/S2/6,5-16x13,6-28
A05023PT113A/S1/5,50-16x13,6-24
A05043KT023A/SS01/9,5-24x14,9-30
A05043PM023A/S01/280/70-R20x380/70-R28
A05043PT103A/S1/280/70-R18x13,6-28

APPENDIX C

Araç Kodu
103463
A05023KT023A/SS01/7,5-16x14,9-30
A05023PT023A/SS00/7,5-16x380/85-R28
A05023PT103A/S2/6,5-16x13,6-28
A05023PT113A/S1/5,50-16x13,6-24
A05043KT023A/SS01/9,5-24x14,9-30
A05043PM023A/S01/280/70-R20x380/70-R28
A05043PT103A/S1/280/70-R18x13,6-28
A05043PT113A/S1/240/70-R16x13,6-24
A05043PT113A/S1/260/70-R16x13,6-24
A05543PT103A/S1/280/70-R18x13,6-28
A05543PT113A/S1/240/70-R16x13,6-24

Item Code	Width (CM)
100325	90
100326	90
100330	30
100359	35
100453	90
100503	45
100531	50
100531	90
100537	90
100562	40
100563	90

APPENDIX D

Filtering Criteria			Filtering Output	Formula
Araç Kodu	Quantity		Kalem Kodu	Quantity
A05043KT023A/SS01/9,5-24x14,9-30	>0		100445	25
A11544KT0440D/+S1/360/70-R28x420/85-R38	>0		100928	25
E07544KT033A/+MCS1/280/85-R24x420/85-R30	>0		100929	25
E09044KT033A/+ARM2/320/85-R24x460/85-R30	>0		100930	25
F07524PT0300/S2/7,5-16x13,6-38	>0		103030	25
E08043PM023A/S2/280/70-R20x380/70-R28	>0		103127	25
			103137	25
			103150	25
			103151	25
			103160	25
			103182	25
			103197	25

APPENDIX E

HP	Location	HO	Location	HN	Location	HM	Location
100562	MLZD.HP.01.00.00	109116	MLZD.HO.01.00.00	103520	MLZD.HN.01.00.00	E060012322821	MLZD.HM.01.00.00
103127	MLZD.HP.01.00.00	109117	MLZD.HO.01.00.00	103521	MLZD.HN.01.00.00	E060012322841	MLZD.HM.01.00.00
103137	MLZD.HP.01.00.00	109118	MLZD.HO.01.01.00	103792	MLZD.HN.01.00.00	E060012322871	MLZD.HM.01.00.00
103150	MLZD.HP.01.01.00	109146	MLZD.HO.01.01.00	104184	MLZD.HN.01.00.00	E060012717521	MLZD.HM.01.00.00
103182	MLZD.HP.01.01.00	109166	MLZD.HO.01.01.00	104200	MLZD.HN.01.01.00	E060013111341	MLZD.HM.01.01.00
103197	MLZD.HP.01.02.00	109176	MLZD.HO.01.01.00	104525	MLZD.HN.01.01.00	E060013116221	MLZD.HM.01.01.00
103198	MLZD.HP.01.02.00	109260	MLZD.HO.01.02.00	106200	MLZD.HN.01.01.00	E060013248341	MLZD.HM.01.02.00
E060013111351	MLZD.HP.01.03.00	109262	MLZD.HO.01.02.00	106201	MLZD.HN.01.01.00	E060013248351	MLZD.HM.01.02.00
E060053143561	MLZD.HP.01.03.00	109298	MLZD.HO.01.02.00	106202	MLZD.HN.01.02.00	E060013400051	MLZD.HM.01.02.00
E060053148071	MLZD.HP.02.00.00	109301	MLZD.HO.01.02.00	106825	MLZD.HN.01.02.00	E060013400091	MLZD.HM.01.02.00
E060053148081	MLZD.HP.02.00.00	109307	MLZD.HO.01.02.00	106869	MLZD.HN.01.02.00	E060013415131	MLZD.HM.01.02.00
100928	MLZD.HP.02.00.00	109311	MLZD.HO.01.03.00	106899	MLZD.HN.01.03.00	E060015335671	MLZD.HM.01.03.00
106252	MLZD.HP.02.01.00	109316	MLZD.HO.01.03.00	106974	MLZD.HN.01.03.00	E060015516641	MLZD.HM.01.03.00
100503	MLZD.HP.02.01.00	109318	MLZD.HO.01.03.00	107237	MLZD.HN.01.03.00	E060016835631	MLZD.HM.01.03.00
103151	MLZD.HP.02.01.00	109327	MLZD.HO.01.03.00	107248	MLZD.HN.01.03.00	E060016835681	MLZD.HM.02.00.00
103160	MLZD.HP.02.02.00	109344	MLZD.HO.02.00.00	107249	MLZD.HN.02.00.00	E060017312931	MLZD.HM.02.00.00
103595	MLZD.HP.02.02.00	109352	MLZD.HO.02.00.00	107307	MLZD.HN.02.00.00	E060017312941	MLZD.HM.02.00.00
105505	MLZD.HP.02.02.00	109382	MLZD.HO.02.00.00	107317	MLZD.HN.02.00.00	E060017313451	MLZD.HM.02.01.00
105506	MLZD.HP.02.03.00	109383	MLZD.HO.02.00.00	107409	MLZD.HN.02.00.00	E060017313971	MLZD.HM.02.01.00

Satış Noktalarında Kalite Kontrol için Örneklem Stratejileri

ETİ Gıda Sanayi ve Ticaret A.Ş.

Proje Ekibi
Bora Baş
Yusuf Duyar
Özge Garip
Eylül Koşok
Tunç Taner
Berkay Taşçı

Şirket Danışmanı
Birtan Pala
Kalite Güvence Müdürü
Bora Alptekin
Kalite Güvence Uzmanı

Akademik Danışman
Prof. Dr. Savaş Dayanık
Endüstri Mühendisliği Bölümü

ÖZET

ETİ Gıda Sanayi ve Ticaret A.Ş. bugüne kadar yalnızca fabrika ortamında gerçekleştirdiği ürün kalite kontrol sürecini satış noktalarına taşımak istemektedir. Bu uygulama ile daha kapsamlı ve doğru bir saha bilgisine ulaşarak, ürünlerde daha detaylı hata analizi yapabilmek amaçlanmıştır. Projenin amacı ürünlerin mevcut satış hacmi, şikâyet âdeti ve fabrika içi kalite kontrol verilerini kullanarak, bu uygulamayı verimli bir ürün seçimi ile gerçekleştirmektir. Bu veriler kıstas alınarak kurulan sistem ile eldeki iş gücünü bir yıllık bir süreç içerisinde en doğru şehirlere, ürünlere ve kontrol adetlerine dağıtmak amaçlanmıştır. Excel üzerinden tasarlanan bir arayüz ile sisteme ayrılacak işgücü, uygulanmak istenen şehirler ve yıl içerisinde uygulamanın gerçekleştirilme sıklığı gibi değişkenler şirketin kontrolüne bırakılmıştır.

Anahtar Kelimeler: Kalite kontrol, keşifsel veri analizi, ürün örneklemesi, hata tespiti

Sampling Strategies for Quality Assurance at Points of Sales

1.General System Description

1.1.Company Information

ETİ was founded in 1961 as a single production facility by Firuz Kanatlı in Eskişehir. Today, ETİ has seven production facilities that are spread over an area of 430,000 square meters, more than 7,000 employees, 190 domestic and overseas distributors and exports to 40 countries on five continents. It is a leading company in the food sector in terms of product diversity where it owns over 220 brands. It's production volume varies between 850 and a thousand tones. With such a high production volume, quality standards of the products are remarkably important for the company. Currently ETİ has production facilities with most of those in Turkey and one in Romania. ETİ is producing consumer goods in 9 categories such as biscuits, cakes, frozen foods etc. According to the results of Fortune Turkey magazine 2018, ETİ is the 72th most company with total 3.169.705.374 TL net sales.

1.2.Quality Control System

ETİ divides their quality inspections into two categories; quantitative and qualitative inspections. Quantitative inspections of the products are mainly done in laboratory environments and evaluated based on numerical results. The size, radius or length of the products are also considered as quantitative measures. On the other hand, qualitative inspections are made upon qualitative traits of the products outside the laboratory environment. Tests are conducted by an operator where the items are checked according to their sensory traits such as favor, color or texture. There are three different stages of quality control inspections in the facility. Every quality control inspection is done according to the "Quality Control Parameters" and each parameter denotes a quality defect. These quality control parameters are determined by experts and every parameter has a numeric score based on their importance. ETİ currently conducts their inspection over 180 different parameters to determine a product's adequacy in terms of the company's quality standards.

First stage of quality control starts with the analysis of the raw materials in the laboratory environment. Second stage of the quality control process takes place at the production stage while the product is being processed on assembly lines by operators. Third and last stage of the quality control in the facility takes place after the production stage is completed. This stage is called the "Auditing Stage" and it is very similar to the second stage except it is a more detailed inspection. In this stage, one box from each production batch is randomly selected and both quantitative and qualitative inspections are carried for one last time to determine the product's final score before reaching the consumer. As each defect parameter has its own score, if any product in box receives a total score higher than the company's critical level, the whole batch is blockaded and will never be sent to the consumers. There are four operators in total that are responsible for the auditing stage across all manufacturing plants. Auditing is the most significant and critical stage for the company as products that pass through this stage are sent directly to the consumers.

2.Project Description

2.1.Problem Definition

In ETİ company a new point-of-sales quality control process is planned to be designed in order to solve and prevent some problems. The current quality control process is restricted to be done only in company as during the production and after the production. However, after the finished products leave the company in order to be put into the market, some new defects occur. The reasons for new defects and the impact of defects couldn't be detected. Only consumer complaints are providing various information about the sources of defects. In appendix A and B it is observed that some consumer complaints are due to post-production defects. Any defects that occurred after the product was dispatched from the factory can be considered as a post-production defect. Examples of post-production defects can be thought as inappropriate conditions of supermarkets storage areas or being damaged during the transportation of products. The reason for this type of defects can only be investigated at the point of sales for future improvement.

The point-of-sales quality control process not only detects the current post-production defects but also detects the possible consumer complaints. During the problem definition, the point of the sales quality control process should not be done always in the same periods with the same parameters. One of the main characteristics of the problem is the variance between audit reports and consumer complaints that occurs because of some main reasons such as post-production errors. Then the after-production quality maintenance can be the bottleneck in improvement of the overall quality for that product.

The problem focused on this project has two root causes. First one is generalization of defective or non-defective items by checking only one package from the whole batch and the other cause is post-production errors. The company currently does not have a quality control mechanism to survey the products after leaving the company. Moreover, batch size for each production changes according to the demand rate of the customers and production ability of the factory. Since one parcel is chosen for audit and result of this parcel according to the product defect score is generalized for the whole batch. This generalization is not accurate for all products in that batch and also there could be some different defects rather than inspected parcels. Planned point of sale quality control takes a step at this problem and gives further information about the reasons for defective products and helps ETİ to take action in order to increase their brand value and consumer satisfaction.

2.2.Project Aim

Overall objective of this project is to create the optimal point-of-sale quality control plan for the company that will minimize the variance between the auditing reports and the consumer complaint reports in order to develop their quality control process in the right direction by taking some actions about pre-production and post-production reasons by having more accurate results in their quality control stage. By applying this quality control plan, the company will have further knowledge on the problems that they could not detect at the auditing process and the reasons behind errors caused by outer effects such as retailer, transportation process or storage conditions causing consumer complaints can be obtained more precisely. It is expected that eliminating the factors causing post-production errors will reduce the variance between the audit reports and consumer complaint reports. The plan developed at the end of the project will produce an exact schedule of the point- of-sale quality control system. There will be some alternative plans for them to choose according to their budget and goal. Although it is an exact plan with directions, it will be dynamic. After its first implementation, it will have the flexibility to be modified with changing workforce constraints, changing demand/sales rate of products or varying consumer complaints which differ between areas. Complaint rates, sales rates and audit reports are considered during development of POS quality control plan. This leads ETİ to be able to see products and parameters having high complaints and high audit points, although they have high sales rates. Thus, another benefit of this plan is determining the most problematic products having the highest possibility of being complained about.

3.0. Solution Approach

3.1.Exploratory Data Analysis

For the years 2018 and 2019, auditing, complaint and sales data provided by the company were analyzed to have a better understanding of the situation. First of all, customer complaints are analyzed to see the most problematic products. The top 10 highest complaint received products in 2018 and 2019 can be seen in Appendix A and B. As it can be seen, the complaints from the first three products did not change and most of the products are the same between the two tables. This implies that managing the quality of these products are critical for improving customer satisfaction. The table in Appendix C shows the average defective points of the first ten products that received the highest average points in the auditing stage in a descending order. Second column shows the total number of passes in the auditing stage and is proportional to its sale volume. There is no product listed higher than point 40, as the blockaded products are kept out of the comparison as they will never meet the consumer. The products that cover the 80 percent of the sales volume are listed in Appendix D. 1/3 of the brands cover up the 80 percent of the sales. The figure in Appendix E shows the complaint, audit and sales percentages of each product plot in a 3D plot. Auditing percentages for each product are calculated by dividing the sum of the points received by a specific product by the sum of the points received by all products. Similarly, complaint and sales percentages are calculated as the proportion of each product's value divided by the total number. Products that have their names written in the plot are the ones that received higher than the average complaint and auditing scores among all of the products. Before developing the models, exploratory data analysis enabled us to see the most problematic products and have a better understanding of the situation.

3.2.Model

This model takes four parameters as input which are calculated from the audit, sales and complaint data. Then, sums them together to obtain weights for each product. Other parameters are the total number of workers that will do the operation, cities that will be checked, threshold values for the calculated sample sizes, period the plan will be utilized. After calculating the weights of importance, the model constructs a plan and outputs sample sizes to be checked for each product in the given cities, by distributing the total workforce to the products' sample sizes.

Audit parameter is the percentage of the defect points received by that product among the total defect points given in that year. Complaint parameter is the percentage of a product's complaints among all of the complaints received. Sales parameter is the percentage of a product's sales volume. Last parameter is called the consistency parameter which is the value of the coefficient of variation obtained from the statistical model approach above. The reason for that is if only the mean value of the auditing points is used, products that have lower mean and higher variation can be disregarded. As the objective is to gain information about the products' performances' at the point of sales, products that are unstable should also be given importance. All of the parameters are scaled into percentages to perform convenient calculations. As the data for each product brand differs across cities, these percentages are calculated separately for each city with their corresponding data. Therefore, there are 45*x*6 = 270 product-city couples which have different parameters.

Summation of these 4 parameters I_{ij} yields a total weight for each product-city couple that corresponds to its importance. However, during the summation, the model adds another parameter which is the absolute value of difference between audit and complaint parameter. The model favors the complaint parameter or the audit parameter based on which is higher. The reason for that is to be able to reflect the maximum value among the audit and complaint data. If the complaint percentages are higher than the average defect points percentage, the model replaces the audit percentage with the complaint percentage as the complaint percentage is higher and poses more risk.

S_{ij} *: Percentage of sales of product i sold in city j among all products sold in selected cities*

C_{ij} *: Percentage of complaints received from product i in city j among all complaints received from products in selected cities*

A_i *: Percentage of audit points received from product i among all products*

$|C_{ij} - A_i|$ *:Absolute value of the difference between complaint and audit for product i sold in city j*

Y_i *: Percentage of standard deviation/mean of audit points received from product i*

I_{ij} *: Total importance weight of product i in city j*

$$I_{ij} = C_{ij} + S_{ij} + A_i + |C_{ij} - A_i| + Y_i$$

Then I_{ij} is assigned to each product-city couple and the total workforce is distributed according to this number. However, after making this calculation some sample sizes are calculated really low such as two or three. This issue is solved by an input value called threshold parameter. In order not to decelerate the catch up time of the defective products, upper and lower threshold values of the user choice are entered which makes the model output sample sizes only between upper and lower threshold values. Products with sample sizes calculated lower or higher than the specified threshold values are rounded to zero and these excess sample sizes are assigned to other products. For example, if there are products with the sample sizes both above the upper and below the lower threshold, sample sizes that remain above the upper threshold value are distributed to the sample sizes of the products that stay below the lower threshold. If the products in the plan remain only above the upper or only below the lower threshold, the model again distributes the excess sample sizes to the other products proportional to their I_{ij} values.

Next, to calculate the total workforce, total number of available quality control workers for the pos quality check processes which are entered by the user multiplied by the total available time for this process. Then, dividing this number by the required time to check one product gives the total number of sample sizes that can be checked.

$$Total\ Sample\ Size = \frac{Number\ of\ Workers\ x\ Number\ Of\ Days\ x\ Work\ Hours\ in\ Units\ Period}{Required\ Time\ To\ Check\ One\ Product}$$

However, some products can never pass the lower threshold value and not be checked over long periods of time. To overcome this issue, the model increases the I_{ij} of these products proportional to the number of periods that product remains unchecked.

$$I_{ij} = \frac{I_{ij} + number\ of\ periods\ that\ product\ is\ not\ checked}{\sum_i \sum_j I_{ij}}$$

Reason for making this calculation is to increase the importance of a product in the following month, in order to make it exceed the lower threshold value and be included in the plan. As every product that is not checked is interpreted as losing information about its performance at the point of sales. Therefore, product's gain importance in the period that they are remaining unchecked.

In the application progress, after the initial run of the system, the model introduces a new parameter to the system which is named as field audit and denoted by F_{ij}. This parameter covers the results obtained during the implementation of this quality control plan. Calculation of this parameter is obtained by dividing the total defective products found, by the total sample sizes checked for each product in the plan. Then, this percentage is divided by the sum of these percentages of all the products. This is done to bring field audit results to the same scale with other parameters. This parameter is added to the model shown as below.

D_{ij} : *Number of total defective products found from product i in city j*

K_{ij} : *Number of product i inspected in city j*

$F_{ij} = \frac{\frac{D_{ij}}{K_{ij}}}{\sum_i \sum_j \frac{D_{ij}}{K_{ij}}}$ and

$I'_{ij} = C_{ij} + S_{ij} + A_i + |C_{ij} - A_i| + Y_i + F_{ij}$

Also in case of a new product launch, a plan that includes this product can be constructed as well. If a new product is added, it is constantly checked throughout the plan with the sample size equal to the lower threshold value. This applies for the initial year of the new launch as the product has no complaint or audit data to derive other parameters. After the audit, sales and complaint data is obtained, new products can take place in the model fully.

3.4.User Interface

A user interface is developed to make the utilization of the model easier and more user friendly. As the algorithms and data processing parts of the model take place in Excel, the interface is also decided to be developed in the same environment. User interface asks for 4 inputs from the user; workforce as the number of quality control operators, chosen cities, desired periods and the threshold values as a constraint. Users can choose the desired period that the plan will be constructed upon as either monthly, quarterly or semi-annually although the default structure of the model is constructed using annual data. Furthermore, it is possible for the user to add every city of their choice or new products in the future, if they feel the need for deepening the investigation in specific locations. After the inputs are entered and the model is run, interface provides the sample sizes for each product in the results part where the company managers now can easily manage further organizations.

4.Assessment and Validation

After the model gives outputs of sample sizes, there is a need to validate the model and observe its performance. The output of the model provides a schedule of products that will be controlled. However, the output of the system will be the number of defectives caught for each controlled product. This data will be recorded and used to keep in track and compare the defective rate from the field with the current known rate by the factory data. Since the output plan will be a whole new system for the company, there can be no comparison in terms of the coverage or the product choice. It will be the first quality control in the field, so the focus on assessing the system will be the failure rate. Although the new failure rates will be obtained from the field once the point of sale quality control plan is implemented, there is no past data for it to be made a comparison. Until the system completes a year and starts giving healthy results, the assessment should be made with a failure rate defined by the past data. To define the past failure rate for each product with the current data inside the company, three different approaches are made using complaint & Sales data, by blockaded product percentage and by average defect points. Among the approaches, sales data was the best option. Failure rate is calculated by taking the complaint amount of the product, multiplying it with 10000(the assumption of ETİ that only 1 consumer out of 10000 reports a faced defect) and dividing it to the total sales.

4.1.Simulation

Under normal circumstances, the system would be implemented and the pilot study on a small scale would be done. Model output would be tested in the cities and defective products that are caught would be recorded to evaluate the model's performance. However, due to the extraordinary conditions the system could not be implemented in any scale and no pilot study was performed. Therefore, a simulation of the system is made using R software to evaluate the model's performance.

By multiplying the failure rates determined for every product with the total sales volume of that product, the expected amount of total defective products in the market is found. Then, this amount of numbers between 0 and “total sale number of the product”, are generated. These numbers will represent the place of defective products and serve as an index for them. If there is an index smaller than the controlled amount determined by the model, that defect will be considered caught.

This method is just to replace the market inspection process and obtain the outputs that it would give. Under normal conditions, it won’t be needed and the caught failures data will be recorded manually. For now, the data from the simulation will be used to continue the validation and observe the model’s performance.

4.2.Quality Control Process Chart

To track the failure rate for consecutive periods, the best option is to use a quality control chart. In every period (month quarter or semiannual) that the model works and determines the sample size inspection, there will be new failure rates found for that sample with the observed defects.

As discussed in the beginning of the chapter, it is determined that the previous year's failure rate based on complaint and sales data is the base value that the assessment will be made on. That rate will be considered as the ideal rate that is expected to be seen on the field since it is based on the official reported failures of the previous year. There are two things to observe about a product's failure rate. First one is the deviation of the rate in a single period. It is obvious that the yearly rate is not valid for every period because the distribution of failures may be heterogeneous through a year. Although it is possible to have a huge deviation, it needs a certain limit that is not to be passed for a period. To track this, the most suitable option is a P-Chart for the failure rates with an upper limit of 3 times the standard deviation, and no lower limit. The center line will be the ideal rate expected at the end of the year. Every separate point on the chart is the failure rate for its period and a point above the upper limit indicates an unexpected problem for that period itself. However, this is not enough since a deviation of 3 times the standard deviation may be acceptable for a period but it is certainly a problem for a year time. If all the periods are just slightly under the upper limit and conclude on a real close value to it, then the P-chart will detect no problem where there is actually a huge one. Since the nature of a P-chart is to control every point separately it is not enough. This is why a cumulative-sum graph is also needed. The second thing to observe in the quality chart is the deviation of the failure rate at the end of the year, more specifically, at the end of every period by taking into account all the previous periods. A CUSUM graph updates its location with the failure rate of every period and each point of the graph gives a weighted cumulative result. Every period has the same weight regardless of the inspection amount. A rate below the ideal one declines the graph, while a rate above the ideal one raises the graph. The slope depends on how big the difference between the rates is. The graph of cu-sum allows detecting every little shift during the whole horizon, and catches if there is a huge deviation until a certain point in the entire year. However, its position at the periods near the end is mostly used, since it can normally be highly deviated in the beginning but needs to be closer to the center line as more periods (in other words weight) pass. It makes sense that when there is more proportion of the whole data, it is expected to have a closer position to the yearly rate. For this system, it is not a huge problem if it is below but far from the center, since a lower failure rate is not critical but if it is way higher than the ideal rate at any point near the end, it is a problem.

An example of this mixed quality control chart of Crax Adana is shown in Figure 2 where the brackets under the failure rate points additionally show the inspected amount and the caught defectives. The periods in this schedule are months. First of all, none of the points in the p-chart are above the upper limit so no periodic problem is detected. More specifically, for the first six months there are no defects caught so the failure rate is zero. Since the weight of each period is the same, the first 6 periods make the same decrease in the graph. For the seventh period, the rate has again a negative difference with the ideal rate, so it goes downwards too, but with a smaller slope since it is not as far as the previous ones. The last period has a higher rate than the expected, so it gives a positive inclination to the graph. However, the CUSUM graph at almost every point is way below the expected rate, so it can be said there is no problem at the end.

For the current situation, the charts for all of the products gives consistent results with no significant problem detected. This is because the ideal failure rate from the previous year is currently based on complaints and sales data, where also the simulation is also based on the same data. Once the point of sale quality control plan is implemented and manually recorded field data replaces the simulation, there will be a difference to allow comparison. Still, it won't be a completely healthy assessment. After the plan works for a year and starts using its own previous data instead of complaints and sales as the ideal rate, the quality control chart will reach its full meaning.

Appendix

Appendix A

Table 1. Top 10 most complained products and their causes in 2018

MARKA	ŞİKAYET KONUSU	ŞİKAYET KAYNAĞI	n
Burçak	Yabancı Madde	Üretim Kaynaklı	51
Crax	Yabancı Madde	Üretim Kaynaklı	49
Eti Çikolata	Üretim Sonrası Hatalar	Depolama - Satış	46
Form	Yabancı Madde	Üretim Kaynaklı	36
Hoşbeş	Yabancı Madde	Üretim Kaynaklı	29
Karam	Üretim Sonrası Hatalar	Depolama - Satış	26
Gong	Yabancı Madde	Üretim Kaynaklı	22
Popkek	Ambalaj Bozukluğu	Tasarım - Ambalaj	18
Puf	Yabancı Madde	Tedarikçi Kaynaklı	17
Cicibebe	Yabancı Madde	Üretim Kaynaklı	16

Appendix B

Table 2. Top 10 most complained products and their causes in 2019

MARKA	ŞİKAYET KONUSU	ŞİKAYET KAYNAĞI	n
Burçak	Yabancı Madde	Üretim Kaynaklı	50
Crax	Yabancı Madde	Üretim Kaynaklı	36
Eti Çikolata	Üretim Sonrası Hatalar	Depolama - Satış	29
Karam	Üretim Sonrası Hatalar	Depolama - Satış	26
Cicibebe	Organoleptik Konular	Tüketici Algısı	23
Gong	Yabancı Madde	Üretim Kaynaklı	22
Süt Burger	Mikrobiyal Bozulma	Depolama - Satış	22
Form	Yabancı Madde	Üretim Kaynaklı	17
Hoşbeş	Yabancı Madde	Üretim Kaynaklı	17
Lifalif	Yabancı Madde	Üretim Kaynaklı	16

Appendix C

Table 3. Highest point receiving products in the auditing stage in 2018

Brand	Total_pass	Avg_Defect_Points
Benimo	763	26.77
Tutku	1161	24.61
Kremalı	227	24.41
Popkek	1804	23.68
Karam	1002	23.46
Kombo	163	22.75
Hoşbeş	2251	22.7
Eti Kek	34	22
Bidolu	115	21.91
Burçak	2616	21.79

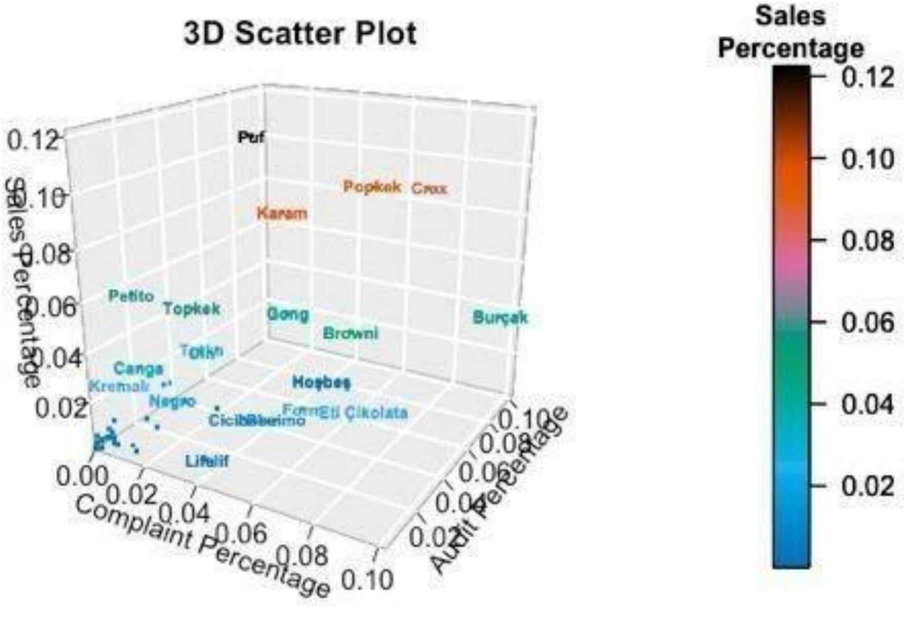
Appendix D

Table 4. Products that cover 80 % of the sales volume in 2018

Brand	Sales Percentage	Sales Coverage
Puf	12.194%	12.194%
Popkek	9.609%	21.803%
Karam	9.189%	30.993%
Crax	8.514%	39.506%
Petito	7.721%	47.227%
Topkek	5.774%	53.001%
Browni	4.765%	57.766%
Burçak	4.411%	62.176%
Gong	4.029%	66.206%
Cin	3.555%	69.760%
Canga	3.348%	73.108%
Tutku	2.542%	75.650%
Eti Çikolata	2.529%	78.178%
Kremalı	2.496%	80.675%

Appendix E

Figure 1. 3D plot with audit, complaint and sales percentages in 2018



Appendix F

Table 6: Estimated parameters for the prior distributions using the auditing data in 2018

Brand	Tot_audit_pts	Tot_audit_pass	Std	Pi	alpha	beta
Popkek	38626	1809	0.2486	0.5338	1.616	1.411
Browni	26712	1538	0.2495	0.4342	1.279	1.667
Crax	64880	3586	0.2416	0.4523	1.467	1.777
Burçak	41102	2493	0.2142	0.4122	1.765	2.517
Hoşbeş	44750	2220	0.2151	0.5039	2.22	2.185
Cin	12112	853	0.182	0.355	2.099	3.814
Cicibebe	18506	1226	0.1995	0.3774	1.851	3.055
Puf	12918	792	0.258	0.4078	1.072	1.557
Tutku	20062	1085	0.2219	0.4623	1.872	2.178
Topkek	9618	510	0.2491	0.4715	1.421	1.593

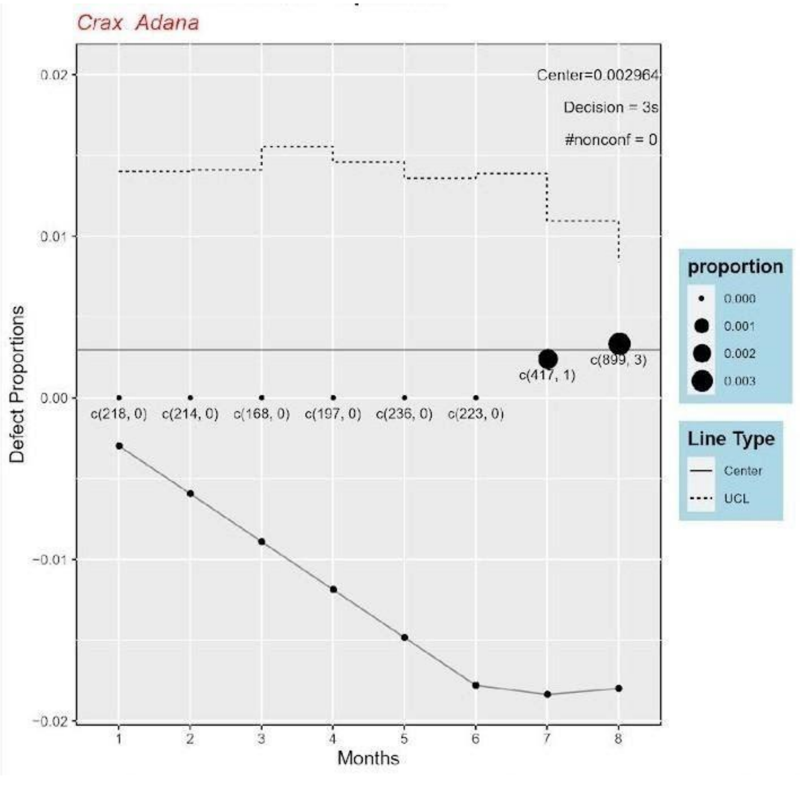
Appendix G

Table 7: Estimated parameters for the posterior distribution of defective probabilities

Brand	Mean	alpha	beta	mean_pos	alpha_pos	beta_pos
Popkek	0.5896	1.685	1.173	0.009494	40002	4173317
Browni	0.476	1.214	1.336	0.008931	220001	24412434
Crax	0.4781	1.49	1.626	0.05963	960001	15138849
Burçak	0.5447	1.849	1.546	0.005892	60002	10123423
Hoşbeş	0.5674	1.998	1.523	0.02135	1340002	61415669
Cin	0.468	1.854	2.108	0.008023	50002	6182070
Cicibebe	0.4112	1.908	2.732	0.03426	1990002	56094794
Puf	0.421	1.135	1.561	0.00567	250001	43838703
Tutku	0.6153	1.466	0.9163	0.06514	710001	10189488
Topkek	0.5221	1.412	1.293	0.01047	490001	46325005

Appendix H

Figure 2: P and Cusum chart of Crax Adana



Veri Zarflama Analizi ile Stok Tutma Birimi Portföyü Eniyilemesi

Eti Gıda Sanayi ve Ticaret A.Ş.

Proje Ekibi

Mert Cüceloğlu, Damla Kahya, Defne Kavas,
Ece Kavukcu, Besime Eda Öksüz, Berkay Özşeker, Didar Uslu

Şirket Danışmanı	**Akademik Danışman**
Tolga Köken	Prof. Dr. Savaş Dayanık
Üretim Planlama Müdürü	Endüstri Mühendisliği Bölümü

ÖZET

Eti'nin ürün portföyünde 2000'i aşkın karar verme birimi (KVB) bulunmaktadır. Ürün portföyü düzenlemelerine dair karar mekanizmasında eniyilemeye yönelik bir sistem bulunmamaktadır. Bu projenin amacı mevcut karar verme birimlerinin, sürdürülebilir bir sistem dahilinde performans takibinin yapılmasını ve ürün çıkarma, ürün ekleme gibi önem arz eden kararların bu sistem dahilinde verilmesini sağlamaktır. Bu amaç doğrultusunda, öncelikli olarak bu projenin paydaşları ile görüşülerek, karar verme birimleri için öne çıkan değerlendirme ölçütleri belirlenmiştir. Bu ölçütler bazında ürün grupları oluşturulmuş ve karar verme birimlerinin performansları kendi ürün grupları içerisinde değerlendirilmiştir. Projenin çıktıları, Eti'ye gösterge paneli ile sunulmuştur.

Anahtar Kelimeler: karar verme birimi, veri zarflama analizi, parametre, gösterge paneli

SKU Portfolio Optimization via Data Envelopment Analysis

1. System Description and Analysis

1.1. Company Description

Eti Gıda Sanayi ve Ticaret A.Ş. is a fast-moving consumer goods company which was founded in 1962 by Firuz Kanatlı with the establishment of its first factory in Eskişehir. It was started as a sole proprietorship, but then has transformed into a family business. The wide range of product portfolio mainly consists of crackers, biscuits, cakes, wafers, functional products, breakfast products, frozen goods and chocolate. The production of those products are handled in eight different production facilities (Eti, 1). In 1978, ETİ expanded its operations by establishing "ETİ Makine Sanayi ve Ticaret A.Ş." with the aim of fulfilling the equipment and machinery requirements of the food production facilities. This installment has provided ETİ with a solid R&D facility in which integrated technology is used to design and manufacture machinery and equipment that are used in every step in these automated production plants from dough preparation to assembly (Eti, 2).

1.2 Analysis of the Current System

Eti, organizes its portfolio with stock keeping units (SKU), which are unique identification codes that describe every distinct product that is produced in the company. Elementally, any minor or major distinction in the raw material-to-shelf period results constitutes a new SKU, reaching the total number of 2000 SKUs in the portfolio. This high number of SKUs in the company's portfolio, causes operational complexity and difficulty in performance tracking. Since Eti produces fast moving consumer goods, the market is highly volatile and dynamic, causing SKU's performance in the market to change rapidly. Currently a long-tail graph which shows the annual margin that represents what percentage of sales has turned into profit for each SKU is used to best describe the current SKU portfolio of Eti and can be seen in Appendix 1.

With the fast-changing customer demands, being able to provide different product types increases customer satisfaction, but it also proportionally makes the long-tail part of the graph even longer. At this point, the presence of the prune non-margin SKUs, which represent approximately 20% of the sales margin fitting a near-perfect Pareto distribution, creates a need for revision or delisting up to a certain level since they directly increase setup time, processing time, cost, labor or capacity consumption. Under these circumstances, it is vital for ETİ to track the performance of each SKU in terms of profitability, operational complexity and contribution to branding. Overall, Eti is in search of an optimality-based and standardized decision-making system, which will allow continuous performance assessment.

2. Scope of the Project

2.1. Problem Definition

The ever-growing size of the SKU portfolio, increases the operational complexity and results in a great level of difficulty in monitoring performance. However, closely monitoring each SKU is vital especially in a rapidly changing industry. Therefore, need for an improvement in the current SKU tracking system can be predicated on the below mentioned problems that were detected upon initial analysis.

Firstly, an optimality-based SKU performance tracking system that would enable Eti to gain quick and efficient insights on each SKU is necessary. The current approach allows SKUs that might need reevaluation or replenishment to be overseen. Furthermore, the lack of a continuous reviewing system results in a long delay until an issue is recognized, which relatively costs the company a loss in production time, profits and overall efficiency. Another issue that requires an immediate amendment is the lack of continuity in the reviewing process. Even when an SKU is scrutinized, an objective and technical method to quantitatively assess the overall performance remains necessary. Usually a series of discussions amongst the decision makers based on estimates constitutes the final decision on whether an improvement or delisting will be made. Finally, based on the interviews conducted with each department, it is clear that each of these departments evaluate the performance of an SKU with regards to very distinct parameters. While each of them is effectual on the performance on different levels, the current system reflects these different points of view to a limited extent.

2.2. Objectives of the Project and Deliverables

The main objective of the project is to provide a comprehensive and dynamic SKU evaluation mechanism which takes into account all stakeholders. Hence the first aim of this project is to rank and evaluate the performances of SKUs with respect to each other. Another objective is to be able to assess how SKUs performances change with respect to time. This is important as seasonality and the volatility within the market affect SKUs performances over time. Finally, another aim of the project is to assess how a new developed SKU will perform within the portfolio. When all objectives are considered, it was decided that the expected deliverable of the project is a dashboard which allows an ease of access to comprehensive information. This dashboard is expected to present an SKU with respect to its cluster by assessing its historic and projected performance in order to evaluate the overall success but also be able to take precautions in the cases of possible “red flags”. Another expected feature of the dashboard is the ability to evaluate listing decisions. This evaluation process includes making estimations about the performance measures of the new SKUs, evaluating their projected performance and making comparisons with other SKUs in the same segment. This will emphasize how much incremental value the new SKU is estimated to bring in.

3. Solution Methodology

This section will be discussing the methodology followed throughout the project by introducing models and algorithms utilized. At this stage firstly the dataset was prepared for analysis with elimination of age effect, imputation and clustering. Then, analysis is conducted via Data Envelopment Analysis.

3.1. Elimination of Age Effect

The dataset provided by Eti involves values of 9 decision-making criteria of 487 SKUs over 22 months. These criteria are: sales in kilograms, sales in TL, EBITDA, contribution margin in TL, refund/return rate, capacity by kg/h, productivity in man-hour/tons, waste/defect rate and number of complaints. The dataset provides comprehensive information, however as the market dynamics, and internal processes of SKUs change with respect to time; the expected values of decision-making criteria of SKUs change according to the age of the SKU. In this context, the raw data introduced by Eti is processed to eliminate the age effect of the SKU.

The first step in the elimination of the age effect on input data involves plotting the SKUs' age versus the parameter values for every parameter. The logarithm of the parameter values is used in the plots to alter the effect of outliers. For instance, as can be seen in Appendix 2, the EBITDA values on the logarithm base of two are plotted against the age of the SKUs to observe the variation.

After the initial observation, a local regression model is constructed on the plot to find the expected parameter values for different ages of SKUs, in this case EBITDA values for a given age as seen in the graph to the right in Figure 1. The local variation amount that is the span of the fit is optimized using the function in R. The points that are under the fitted line performed poorly for the expectation for that age and the point above belong to SKUs that performed better than the expectation. After the model is constructed, the residuals of the 22 data points that belong to different ages of an SKU are summed up for each SKU to find its overall performance compared to the expectation at different ages. The sum of residuals of the SKU's data points, that is the distance between the fit for the age is used as the processed data in the succeeding methods. The models are confirmed to be accurate by the industrial advisor as the R-square value of all models are above 0.92.

Thirdly, to detect if there is a necessary distinction between the SKUs in terms of the parameter values that suggests a requirement for multiple models to determine the expected values, the density plot of the monthly average values for each SKU is plotted. In the case of EBITDA as the parameter, depicted in the left graph in Figure 1, there is a breakeven point that separates the SKUs in terms of EBITDA performance that is the logarithm of monthly average EBITDA of

260 so the SKUs that have a logarithm of monthly average EBITDA higher than 260 are used to build the first model and the SKUs below are used in the second local regression model.

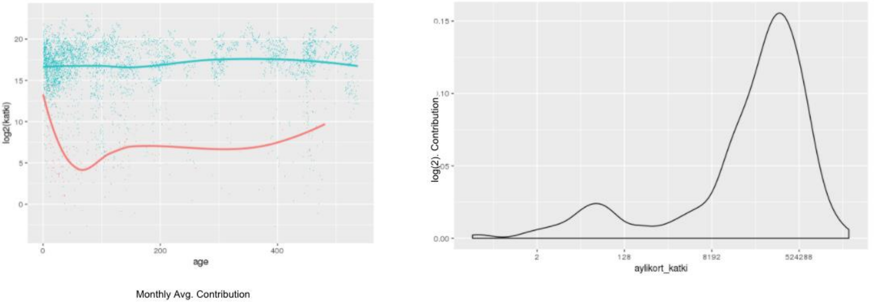


Figure 1. Graphical representation of elimination of age factor.

This analysis incorporated with the calculation that can be found in Appendix 3 is performed in order to comprehend the meaning of this data in practice. Specifically, the historic data is fitted to a regression that displays the trend in past sales data, in other words, how much the sales volume increases or decreases. This allows Eti to interpret sales data trends and forms an additional output and benefit of the project.

3.2. Imputation of the Data

The main challenge faced at the initial stage of the project was the dataset being incomplete, resulting in a pervasive problem. 18% of the dataset consisted of N/A data. When the N/A variables were checked, it was observed that some of the N/A variables are existent due to the lifecycle of the product such as delisting or listing. Nonetheless, some of the N/A variables occurred due to mistakes or errors. Since the labels or any information to identify the actual product were not provided, clusters were created for SKUs with an unsupervised approach, which will be discussed further. Due to the nature of unsupervised approaches, N/A features and variables were very crucial.

In order to eliminate this problem from creating bias, Random Forest Algorithm was used for imputation. The Random Forest Algorithm is a form of classification and regression trees, in which predictive models are created that recursively subdivide the data based on values of the predictor variables. The main reasoning behind the superiority of this model for this case specifically is that it does not rely on distributional assumptions and allows nonlinear relations and interactions. The validation for this methodology will be discussed further.

3.3. Clustering of SKUs

Clustering the SKUs is significant to avoid evaluation of the performance of an SKU by comparing it to an irrelevant SKU in terms of product category, position in the market, generated revenue etc. In addition, the method prevents the evaluation of all SKU performances with the same parameters which is not realistic. A parameter that is effective on determining the performance of an SKU may yield no information for others. Two methods are used for clustering.

3.3.1. K-means Approach

The initial method utilized is k-means, which is used to find groups in the data, in which the number of groups present are shown with the variable *K*. The algorithm assigns each data point to one of *K* groups based on the parameters given. Data points are clustered based on similarity in values of these parameters. The appropriate K value is found by trial-and-error of certain numbers (Carreira and Miguel 2). The elbow method within the k-means approach is used in order to decide on the range in which the number of clusters should be. Elbow method looks for the percentage of variance explained for each number. For this purpose, an elbow graph is drawn in order to assess the appropriate range for the number of clusters by assessing SSE (sum of squared error) values for each number of clusters, and this graph can be seen in Appendix 4. As it can be seen, the range for the number of clusters is found to be between 6 and 15.

3.3.2. Gaussian Mixture Models

There are certain shortcomings of the k-means method, as it does not take into account variance and it only provides which cluster the data points belong to. In addition, the method does not provide information as to what the level of agreement between the cluster and the point is. An alternative to this approach is the Gaussian mixture model, which compensates the shortcomings of k-means. This method can be described as very similar to k-means, but first of all it accounts for variance and secondly, it provides the probabilities of data points belonging to each of the clusters. Different number of clusters and model methods are tested according to the Bayesian inference criterion, and optimal model is returned as a result of this algorithm. According to the algorithm, VEV (Ellipsoidal Distributed, Variable Volumed, Equal Shaped, Variable Oriented) Model selected as the superior model with 8 clusters. For evaluating SKU's these 8 clusters were considered as a base point.

3.4. Evaluation of Performance Measuring Methods

The main objective of this project is to evaluate each SKUs' performance with respect to each other, hence this brings significant importance to deciding on the appropriate methodology to base the performance measurement on. In this project, the main methodology used is Data Envelopment Analysis.

3.5. Data Envelopment Analysis

Data envelopment analysis is a non-parametric performance measurement technique for benchmarking groups of entities. It is very common in Operations Research and economics for the estimation of production frontiers. Decision-making units (DMUs) form the basis of this algorithm as each of these units are evaluated as part of a group that utilizes inputs to produce outputs. For this project, SKUs are considered to be DMUs. The result of this evaluation provides an efficiency score between 0 and 1 for each unit and this value represents the degree of efficiency (Ersoy 3). This approach observes the minimum number of inputs necessary to produce maximum number of outputs which is a valuable output to be presented to Eti. The model also decides on the efficiency limit for decision-making units and the SKUs below this limit are considered to be inefficient.

There are many advantages to using DEA. Firstly, it allows the usage of different metrics. As the parameters in the project have different metrics, this is crucial. In addition, the model takes the existence of a different parameter that affects both inputs and outputs into account. Moreover, the model takes the time factor into consideration, this is beneficial for the historical performance analysis to be made.

3.5.1. Types of Data Envelopment Analysis Models

In the literature reviews conducted, there exist numerous approaches to construct the DEA model. Two major models that are evaluated to be used for the extent of this research are CCR (Charnes, Cooper, Rhodes) and BCC (Banker, Charnes, Cooper) models.

3.5.1.1. The CCR Model

The CCR model can be interpreted based on the inputs or outputs, where the main objective is to make an overall evaluation of the total efficiency and predict which decision-making units are inefficient. This model searches for the optimal combination of the inputs that would yield the most efficient set of outputs. The CCR model is constructed in two ways: the ratio CCR model and the weighted CCR model. The weighted model can be described as the linearized version of the ratio model, meaning that it is derived and simplified from the ratio model. The model can be seen in detail in Appendix 5.

3.5.1.2. The BCC Model

The previously explained CCR model calculates the efficiency scores of the decision-making units under the assumption of constant returns to scale. The difference between the BCC and CCR model is that BCC aims to acquire the technical efficiency scores which represent how efficiently the resources are utilized. Thus, this model interprets the source of inefficiency, which is a useful out for the project. This model’s allowance to variability enables the usage of different parameters or evaluation of parameters if the decision-making units require different handling. The model can be seen in Appendix 6.

3.5.2. Reference Sets and Improvement Percentages

In Data Envelopment Analysis, as the concept is reliant on finding efficiencies of decision-making units with respect to others within the portfolio, it is important to define benchmarks in order to create an evaluation basis. For this purpose, reference sets are defined. These reference sets can be described as groups of efficient SKUs which create a basis for defining the efficiency score of an SKU. A reference unit is found by projecting a decision-making unit radially to the efficient surface (Malik 5). As a result of this method, it is possible to observe which SKUs construct the reference set of an SKU and at what percentages they are contributing to conducting the reference set.

In addition, improvement percentages required for each SKU to reach an efficiency score of 1, can also be found using this algorithm. Thus, reference sets are used and the result is communicated by indicating at what percentages of each output and input parameter should be improved. For calculating the improvement percentages, a simple calculation is made:

- *Potential Improvement Percentage = (Target Input-Observed Input) / Observed Input*

3.6. Improvements to the DEA Model

The Data Envelopment Analysis Model is comprehensive and to-the-point in terms of evaluating how SKUs are ranked when compared to other SKUs in the product portfolio. However, there are certain shortcomings. In order to overcome these shortcomings, two main replenishments have been made.

3.6.1. Super-Efficiency Method

The DEA model gives degree of the efficiencies of decision making units which are SKUs in this project. After obtaining efficiency scores of each SKUs which are between 0-1, it was observed that there was an excessive number of SKUs ranked 1. Hence, a separate method was needed to rank the SKUs which have an efficiency score of 1 in order to interpret their efficiencies when compared to each other. In this case, the super-efficiency as a ranking methodology is used for differentiating the performance of extreme-efficient SKUs.

3.6.2. Malmquist Productivity Index (MPI)

Malmquist productivity index is a method to measure of the levels and efficiency change of production function. It measures the performance improvement of a system by comparing different time of periods. In addition, Malmquist index is measured as the product of catch - up or recovery and frontier - shift or innovation terms (Dorri and Mohsen 6). After obtaining Malmquist productivity index, the company is planned a perspective about each SKUs performance situation in certain time periods.

3.6.3. Bootstrap Method and Confidence Intervals

Bootstrap is the procedure of drawing with replacement from a model, mimicking the data creation process of the actual model, and generating multiple predictions that can be used for statistical inference (Kao and Shiang 3). In this project, it is used for sensitivity analysis in order to test hypotheses. It generates a new efficiency score for each SKU by taking random SKU's out of the data and generates a more conservative frontier line for efficiency. Therefore, the difference between the new efficiency scores according to the new generated lines and CCR scores of each SKU are examined. By resampling, the randomness of the model is reflected to the deviation of the CCR efficiency scores. Then, bootstrap confidence intervals are created by resampling and this interval is observed to be wider when the deviation is higher. As a result, the variance is examined and the hypothesis testing is done.

4. Validation

Validating methodologies and solution approaches in this project was crucial at every step as no prior optimality-based system existed regarding SKU performance tracking. Hence, comparing the developed system with the previous one was not possible. Every action taken was discussed with this projects' Industrial Advisor Tolga Köken and his approval was given. At the initial stages of the project, a kick-off meeting was done on the 11th of October at Eti Headquarters in Eskişehir in order to discuss the scope of the project and comprehend the problem.

After this, a list of parameters was developed and discussed with Tolga Köken. Five meetings were done with marketing, production planning, finance and sales departments in ETİ in order to validate which parameters contributed to determining the performance of an SKU. After this stage, validation was also required for the imputation stage. For this purpose, pre-existent data for certain SKUs were removed from the dataset and these data were imputed using the Random Forest Algorithm we used for the N/A data. Then, the actual values and imputed values were compared and were found to be significantly close.

Previously, SKUs to be delisted or improved were not decided according to an optimality-based system. As the system provided with this project gives a concrete and definite view and score regarding performance, the main method of validation lies in discussing the scoring of the SKUs and the results the system provides one-by-one with Eti officials. As discussed with Eti, the best way to conclude this is to assess whether or not previously delisted SKUs within the portfolio will be detected as "red flags" by the developed system. For this purpose, there are deliberately added SKUs in the dataset which were already delisted. Furthermore, it is possible to differentiate the delisted SKUs from the rest of the product portfolio as their product age remains constant after a period. Hence, every delisted SKU within the dataset was checked and validated to be detected as "red flags" by the system. Clustering was also validated using a similar procedure. As the dataset provided was masked and manipulated, there was no indication as to which product an SKU was. Therefore, it wasn't possible to assess whether the clusters held certain characteristics such as a common product category. At this stage, the SKUs in the same cluster were communicated to Tolga Köken and it was found to be reasonable and meaningful for any two SKUs to be in the same cluster.

The above-mentioned validation stages for imputation, clustering and data envelopment analysis were conducted on 17 January 2020 at Eti Headquarters in Eskişehir. The additions and improvements made according to the feedback received after this stage have been communicated to Mr. Köken via phone or e-mail due to the pandemic.

5. Deliverables and Benefits to the Company

5.1. Benefits to the Company

The main aim of this project is to create a tool for the decision makers and engineers at ETİ to systematically and continuously evaluate the performances of the SKUs in their product portfolio. Prior to the development of this model, the stakeholders evaluated performances according to know-hows and this evaluation was not optimality based. On the other hand, this tool enables the company to monitor the trends, symptoms and critical situations in the SKU's performance in any interval desired, eliminating the subjectivity and discontinuity. Due to the input obtained from the company advisors, it usually takes about a week to gather the necessary data for the SKU's performance and a meeting to discuss and analyze the health of each SKU. The output of this project is expected to decrease these efforts and time required to make these assessments by providing a dynamic dashboard to quickly examine them statistically and graphically.

Furthermore, this analysis provides the engineers the ability to quantify the required improvements in under-performing SKUs. In the light of the outputs of the DEA model, the percentage improvements needed to be made for each parameter in order to advance the SKU's performance up to the efficient frontier in its cluster are estimated and displayed. Consequently, the improvement decisions will be able to be made based on factual and pre-computed outcomes in hand.

5.2. Dashboard and Implementation

As previously mentioned, a dashboard is the main deliverable which the outputs are contacted to Eti through. The dashboard is designed to be user-friendly, accessible and comprehensive. The dashboard consists of multiple screens and each screen provides different sets of information to the company and also eases access to desired information. The screen of the dashboard is as follows: overview screen, SKU performance screen, product ranking screen, performance improvement screen, product performance forecast screen.

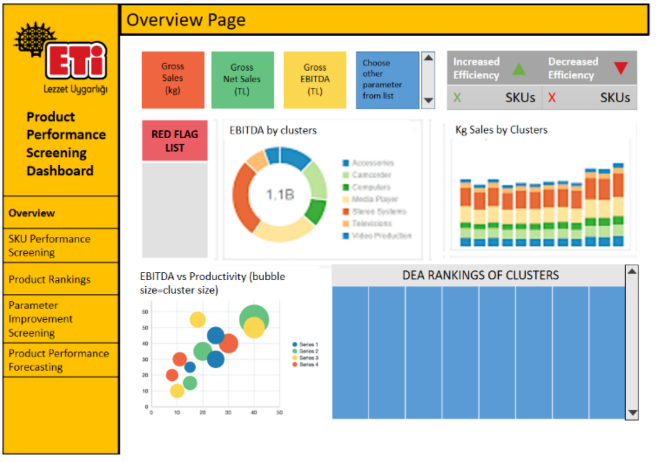


Figure 1: Display of the dashboard (overview screen)

The dashboard is created using RShiny that is within the free software R. Different versions of the software R can be downloaded to any computer for free and changes in the algorithms can be made dynamically by the company in the future. In addition, the dashboard is accessible and easy to use for any employee that is involved in the process.

REFERENCES

"ETİ Gıda Sanayi Ve Ticaret A.Ş." ETÄ°, www.etietieti.com/eti-gida-sanayi-ve-ticaret.

Carreira-Perpinan, Miguel A. "Generalised Blurring Mean-Shift Algorithms for Nonparametric Clustering." 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008, doi:10.1109/cvpr.2008.4587435

Ersoy, Nazlı. "Banka Etkinliklerinin Veri Zarflama Analizi Ile Değerlendirilmesi Finans Ekonomi Ve Sosyal Araştırmalar Dergisi, 2018, doi:10.29106/fesa.382748.

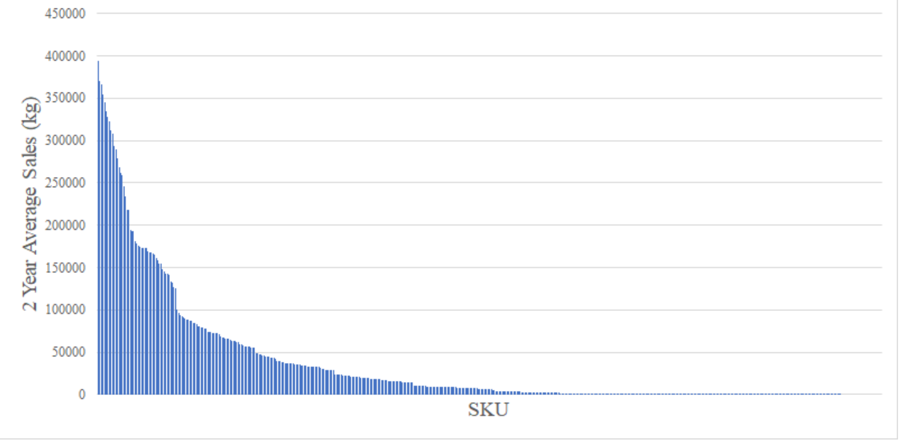
Malik, Meilisa, et al. "Data Envelopment Analysis (DEA) Model in Operation Management." IOP Conference Series: Materials Science and Engineering, vol. 300, 2018, p. 012008., doi:10.1088/1757-899x/300/1/012008.

Dorri, Negar, and Mohsen Rostamy Malkhalifeh. "Malmquist Productivity Index and Decision Science, vol. 2017, no. 1, 2017, pp. 7–13., doi:10.5899/2017/dea-00143

Kao, Chiang, and Shiang-Tai Liu. "Data Envelopment Analysis With Missing Data." Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis, pp. 291–304., doi:10.1007/978-0-387-71607-7_16.

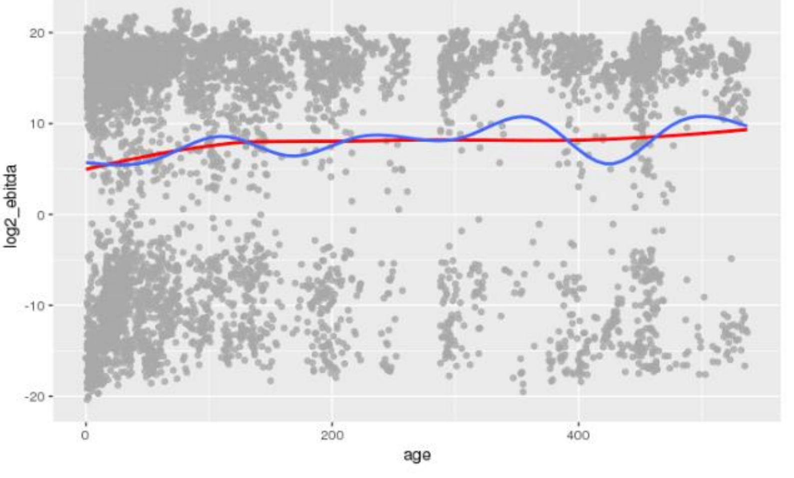
APPENDIX

Appendix 1: Average Sales of SKUs



Average sales of Eti's 487 SKUs over 2-years time are shown.

Appendix 2: Logarithm of EBITDA Input versus Age



The EBITDA values on the logarithm base of two are plotted against the age of the SKUs to observe the variation.

Appendix 3: Calculations for Age Elimination

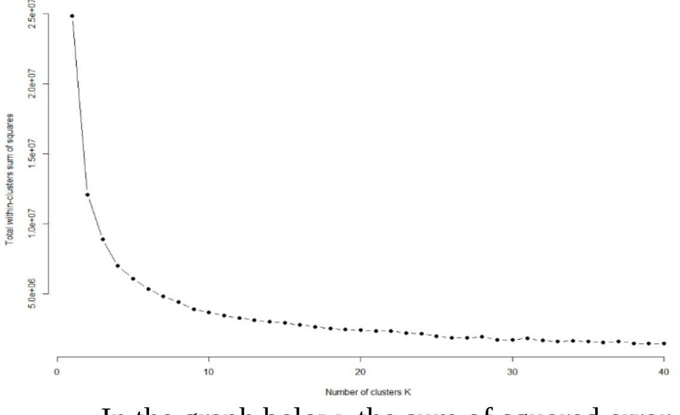
$$\ln s_{t+1} = a + b.\,age_{t+1}$$
$$\ln s_t = a + b.\,age_t$$
$$-----------$$

$$\Delta \ln s = b.\,\Delta age$$

$$\Delta \ln s = \ln\left(\frac{s_{t+1}}{s_t}\right) = b.\,\Delta age$$
$$\frac{s_{t+1}}{s_t} = e^{b.\Delta age}$$
$$s_{t+1} = s_t.\,e^{b.\Delta age}$$

Here, y=a+b.age is the equation of the linear regression line and s is the sales data in a given month. b is obtained via the Excel Regression Model. The difference of these two volumes yields the growth in sales volume. Finally, the value eb.Δage outputs the growth rate of the product every month. For instance, if this value is found to be 1.002, then the product's sales are currently growing 0.2% percent every month, and 4.4% in the 22-month period.

Appendix 4: Elbow Graph for the Number of Clusters



In the graph below, the sum of squared error (SSE) values corresponding to number of clusters between 0 and 40 can be seen. It is understood that the most suitable range of clusters is [6-15].

Appendix 5: The CCR Model

$$\textit{maximize} \quad \frac{\sum_{r=1}^{s} u_{rk} Y_{rk}}{\sum_{i=1}^{m} v_{rk} X_{ik}}$$

$$\textit{subject to} \quad \frac{\sum_{r=1}^{s} u_{rk} Y_{rj}}{\sum_{i=1}^{m} v_{ik} X_{ij}} \leq 1$$

$$u_{rk} \geq 0 \quad v_{ik} \geq 0$$

Since the nature of linear programming suggests that the objective function can not be fractional, its denominator is equated to 1 and this equation is added as a constraint. The weighted CCR model can be seen below:

$$\textit{maximize} \quad \sum_{r=1}^{s} u_{rk} Y_{rk}$$

$$\textit{subject to} \quad \sum_{i=1}^{m} v_{ik} X_{ik} = 1$$

$$\sum_{r=1}^{s} u_{rk} Y_{rj} - \sum_{i=1}^{m} v_{ik} X_{ij} \leq 0$$

$$u_{rk} \geq \varepsilon \quad v_{ik} \geq \varepsilon$$

u_{rk}: The weight given to the r^{th} output by the kth decision variable
v_{ik}: The weight given to the i^{th} input by the kth decision variable k.
Y_{rk}: r^{th} output created by the k^{th} decision variable
X_{ik}: i^{th} input used by the k^{th} decision variable
Y_{rj} : r^{th} output created by the j^{th} decision variable
X_{ij}: i^{th} input used by the j^{th} decision variable
ε: Sufficiently small positive number (ex: 0.0000001)

Appendix 6: The BCC Model

maximize $$\frac{\sum_{r=1}^{s} u_{rk} Y_{rk-\mu_0}}{\sum_{i=1}^{m} v_{rk} X_{ik}}$$

subject to $$\frac{\sum_{r=1}^{s} u_{rk} Y_{rj} - \mu_0}{\sum_{i=1}^{m} v_{ik} X_{ij}} \leq 1$$

The model is simplified to be:

maximize $$\sum_{r=1}^{s} u_{rk} Y_{rk} - \mu_0$$

subject to $$\sum_{i=1}^{m} v_{ik} X_{ik} = 1$$

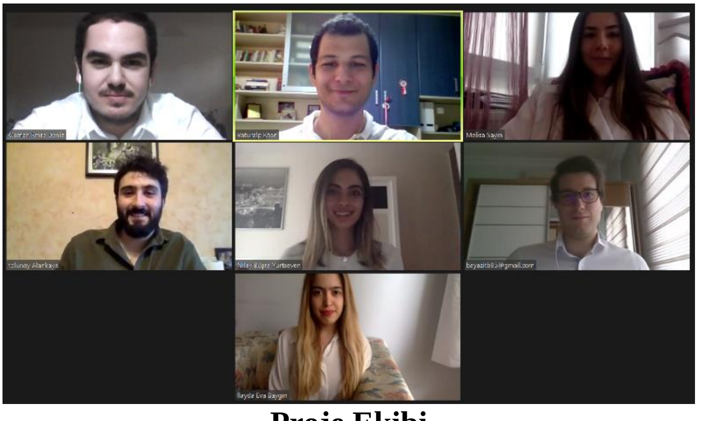
$$\left(\sum_{r=1}^{s} u_{rk} Y_{rj}\right) - \left(\sum_{i=1}^{m} v_{ik} X_{ij}\right) - \mu_0 \leq 0$$

$$u_{rk} \geq \varepsilon \quad v_{ik} \geq \varepsilon \quad \mu_0 : \textit{free}$$

where:
u_0: Variable on the returns to scale

Çalışan Servisleri İçin Çok Araçlı Rota Optimizasyonu

Havelsan A.Ş.



Proje Ekibi

Tolunay Alankaya, Bahadır Bayazıt, İlayda Eva Baygın, Osman Emre Deniz, Baturalp Köse, Melisa Sayın, Nilay Büşra Yurtseven

Şirket Danışmanı
Deniz Taşkın
Ar-Ge Mühendisi

Akademik Danışman
Prof. Dr. Oya Karaşan
Endüstri Mühendisliği Bölümü

ÖZET

Bu proje Havelsan çalışanlarının ulaşım sistemini geliştirmeyi amaçlamaktadır. Şirket çalışanları personel servisinde çok fazla zaman harcamaktadır. Proje kapsamında, rota bulmak için oluşturulan model ve bu modelin parametrelerinin nasıl elde edileceği hakkında bilgiler sağlanmıştır. Sezgisel algoritma oluşturulmuş ve ilgili parametreler sezgisel algoritmada kullanılmış, önerilen rotalar bulunmuştur. Önerilen rotalar görselleştirilmiş ve sonuçlar paylaşılmıştır.

Anahtar Kelimeler: Rota Planlaması, Seçmeli Araç Rotalama Problemi (SVRP), Tabu Arama Algoritması, Cplex Modeli.

Multi-Vehicle Route Optimization for Havelsan Employee Bus

1. Introduction

1.1 General Information about Company

Havelsan is a leading technology company which develops unique systems for military, public and private sector. Company offers intelligent solutions with today's latest technologies. In addition to the production of command control and defense technologies for our air forces and navy, Havelsan provides simulators with high domestic contribution for all kinds of land, marine and air platforms.

1.2 Problem Definition

The working hours of the company are from 07:30 to 16:30. Havelsan provides transportation service to employees with Çalıkıran Turizm’s buses. However, employees complain about the long travel times. For a workplace, the waste of time spent in bus to arrive at work or home results in a serious loss of productivity. When the existing system is analyzed, it is seen that the time spent on the road is more than expected. In the current system, the routes and stops are determined according to experience of the drivers. Also, there is no specific bus to which the personnel are registered.

Havelsan launched a project to design an application where workers can get information about bus and driver, when they will be picked off from stop they will wait. Moreover, the application should visualize the stops, routes, instant position of the buses on the map. Routes and stops could be changed when the new workers are assigned the system. According to employees’ residences, the optimal stops and routes should be determined by the application.

1.3 Current System and Assumptions

In the current system, the company has 50 buses in total: 40 of the buses has 19 capacity, 5 of them has 27 capacity and the remaining 5 buses has 16 capacity. For the purpose of simplicity and to enable buffer spaces in the buses, the 5 buses with 16 people capacity are assumed to be 19. Since we did not want to increase the number of buses and accordingly increase the cost, we assume the upper limit of bus number as 50 but also let the model to decrease the number according to different scenarios. In addition, we company has a restriction that the distance between each employee’s house and the bus stop that the employees are assigned cannot exceed 1 km.

We assume that the routes to be constructed are symmetrical, in other words the route that the bus uses in the morning while coming to Havelsan will be identical with the route that it uses in the afternoon while leaving from Havelsan. That is, we assume that all the routes are formed starting from Havelsan.

1.4 The Objectives and Scope of the Project

The design of a pick-up bus system is a complex dynamic system which includes decisions such as the location of employees, route and capacity of service vehicles and location of the stops. In the current system, stops are pre-determined by the contracted company and employees are assigned according to this. However, these variables are linked with each other such that they should be decided through a holistic perspective. With the solution we propose, we take into account the effects of these decisions on each other.

The aim of this project is minimizing the time spent on the bus. For that purpose, we tried to minimize the distance travelled by each bus. We have developed a Cplex model to find the optimal routes for each bus by minimizing the distance travelled. Since the company has no license for this tool, we have developed a heuristic algorithm by using Pycharm tool. In the heuristic algorithm that we propose, some of the bus stops are selected among possible bus stops locations which are determined according to the employees' locations by the algorithm which the company is provided to us. The routes are determined by taking into consideration of the bus stops. So, the route and stop locations of the service vehicles that are obtained as a result of the algorithm is the input to the personnel registration system which will be constructed by company.

2. Proposed System

2.1 Model

In literature, a lot of research has done to approach the current problem with the correct solution method. In the problem discussed, both the capacity of vehicles and the distance that each vehicle can travel are limited. Hence, only a subset of bus stop locations can be chosen. Therefore, Selective Vehicle Routing Problem (SVRP) is chosen to be used for the problem. The aim of SVRP is determining a feasible subset of stops that covers all demand while minimizing the total cost of travelling. In our case, distances are accepted as costs in the problem.

The employees that can be assigned to potential bus stop locations, which satisfy the maximum walking distance requirement, are given as a parameter to the model. According to this, the model chooses the optimal bus stop locations from the potential bus stop locations. As it can be seen in the Figure 1, the model chooses the stops that yield the optimal route plan and abort the others. Also, the model assigns the employees to specific bus stops. The first box shows the placement of stops and employees. Later, stops are determined by the system and employees assigned to these stops as shown in the second box. Lastly, model constructs an optimal route by using these stops. Unused stops are eliminated by the system. As a result, routes are determined according to the stops which are chosen by the model. Lastly, the model constructs an optimal route by using these stops.

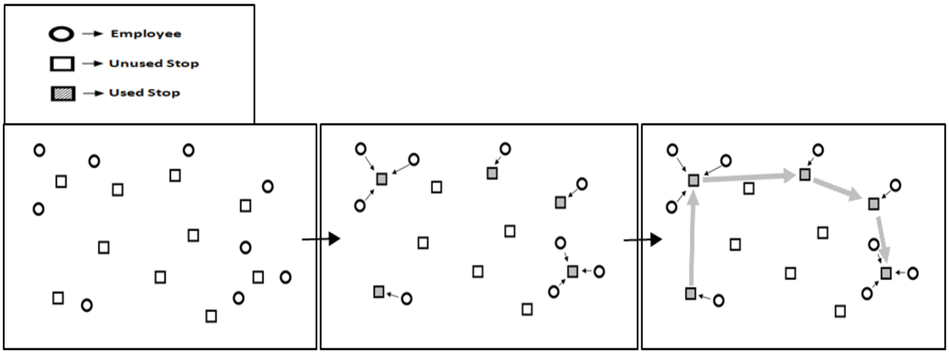


Figure 1. Iterations of SVRP

2.1.1 Sets and Notation of Problem

A directed graph G = (V, A) is constructed, where V = {1,...,n} is the set of n nodes and A is the set of arcs. Node 1 represents the depot (Havelsan) so the remaining node set V \{1} corresponds to the n-1 bus stops. There is also an employee set E = {1,...,e} which represents the employees.

Parameters

$bus\_capacity$ → gives the capacity of bus.

bus_cost → gives the cost of bus.

$cost_{ij}$ $\quad i, j \in V$ → gives the distance matrix between stops.

$dist_{ej}$ $\quad e \in E\ , j \in V$ → gives the distance matrix between employees and stops.

Decision Variables

π_j $\quad j \in V$ → total number of employees in a bus at node j.

$ss_{ij} = \begin{cases} 1 & if\ arc\ from\ i\ to\ j\ is\ used \\ & 0\ otherwise \end{cases}$ → (s)top to (s)top

$es_{ij} = \begin{cases} 1\ if employee\ e\ is\ assigned\ to\ stop\ j \\ \quad 0\ otherwise \end{cases}$ → (e)mployee to (s)top

2.1.2 Formulation

Minimize

$$\sum_{i \in V} \sum_{j \in V} cost_{ij}\ .\ ss_{ij} \ + \ \sum_{j \in V \backslash \{1\}} ss_{1j}.bus\_cost \tag{1}$$

Subject to

$$dist_{ej}.es_{ej} \leq 1 \quad \forall e \in E\ , \forall j \in V \tag{2}$$

$$es_{ej} \leq \sum_{i \in V} ss_{ij} \forall e \in E\ , \quad \forall j \in V, \quad i \neq j \tag{3}$$

$$\sum_{j \in V} es_{ej} = 1 \qquad \forall e \in E \tag{4}$$

$$\sum_{j \in V} ss_{ij} = \sum_{j \in V} ss_{ji} \forall i \in V\ , \quad i \neq j \tag{5}$$

$$\pi_j \geq \pi_i + \sum_{e \in E} es_{ei} - M.\left(1 - ss_{ij}\right) \forall i, j \in V\ , i \neq j \tag{6}$$

$$\pi_j \leq bus\_capacity \quad \forall j \in V \tag{7}$$

$$\pi_j \geq \sum_{e \in E} es_{ej} \qquad \forall j \in V \tag{8}$$

$$\pi_j \geq 0, ss_{ij},\ es_{ej} \in \{0,1\} \forall i, j, e, s \in V \tag{9}$$

(1) The objective function of the model is the minimum distance travelled by the buses in other words it is the sum of all arcs that are included in the tour while also minimizing the total bus number in the system.
(2) If employee e is assigned to bus stop s then the distance between employee e and bus stop s cannot exceed the maximum distance which is 1 km.
(3) If there is no arc that goes out from node s then this means that node s is not included into the tour so no employee can be assigned to bus stop s.
(4) Each and every employee must be assigned to exactly 1 bus stop.
(5) The number of incoming arcs to node j node must be equal to the number of outgoing arcs from node j where $j \in V$.
(6) Miller Tucker Zemlin subtour elimination constraints. This constraint uses the capacity limitations and also keeps track of the number of employees on the bus. If the bus goes from node i and to node j, when the bus arrives to bus stop j, the number of people in the bus must be greater than or equal to the number of people waiting in bus stop i plus the number of people in the bus when it arrives the bus stop i.
(7) Number of people in the bus when the bus arrives to bus stop j must be less than the capacity of the bus.
(8) This constraint enables us to get more sensible π values.
(9) Nonnegativity constraints

2.2 Heuristic Algorithm

The heuristic algorithm composed of two sub algorithms which are SVRP and Tabu Search. Both of the algorithms developed by using Phyton and for the Tabu Search, Google OR Tools are used.

2.2.1 SVRP Heuristic Algorithm

Similar to SVRP Cplex model, the SVRP heuristic algorithm chooses bus stops from the potential bus stop locations while trying to minimize the total distance travelled by each bus. Also, the routes are assumed to be symmetrical as in the Cplex model part.

We have defined a variable which is called as ratio. This variable enables buses going to the closest bus stop from their existing location while providing buses going to a bus stop with having more employees potentially assigned to that stop. Potentially assigned employees are basically sum of employees whose location is less than 1 km to that bus stop and 1 km limitation is determined by the company. Ratio is distance between current location of a bus and unvisited possible stop divided by sum of employees whose location is less than 1 km to the unvisited possible stop.

$$ratio = \frac{\text{distance between current location of a bus and unvisited possible stop}}{\text{sum of employees whose location is less than 1 km to the unvisited possible stop}}$$

The heuristic algorithm finds the routes for buses one by one. That is, it starts with the first bus and after it complete its route, second bus starts and follows the same steps. The algorithm stops when all employees are assigned to the bus stops which are chosen by the algorithm itself. At first, a bus with capacity of 19 starts from Havelsan and goes to a bus stop where the ratio of the bus stop is the minimum. After choosing the next bus stop, if the capacity of the bus is enough, all the employees who can potentially be assigned to that stop are assigned. Then, the necessary switches and changes are done in order not to re-assign those employees to another bus stop. The routes continue to be created according to the same steps until there is no employee left to assign. However, if the algorithm encounters a situation that the total number of employees who can potentially be assigned to that bus stop is greater than remaining bus capacity, then the 19 people capacity of the bus changes with 27. The aim of this approach is to make use of the buses that the company currently has and this change could be done 5 times as there are 5 buses with capacity of 27 in the system. If all of the 5 buses are used, then for the buses with capacity of 19, the algorithm will choose the employees who are closest to that stop and the number of employees that the bus can took will be equal to its remaining capacity. The employees who are not assigned to that route will be picked up by another bus in another route.

Another case that we might have is that, for example, a bus that visits Çayyolu and the remaining capacity of that bus is 4 in that case the algorithm tends to make the capacity full and may be the bus can go to Bağlıca just because of the capacity purposes. In order to prevent this, we assume that if the remaining capacity of a bus is lower than 5, which can later be changed, and next possible bus stop location’s distance is higher than 10km, then the bus will not go to that bus stop and the route for that bus will finish. This limitation also enables us to have some buffer spaces for employees.

2.2.2 Improvement Algorithm

The selected bus stops are extracted from the distance matrix and a new distance matrix with the corresponding demand matrix is created. This arrangement has done in order to combine the solution of SVRP heuristic with a metaheuristic search.

We have used a solving package named Google OR tools which enables us to use other search methods which are Tabu Search, Simulated Annealing and Guided Local Search. Time and iteration parameters can be arranged but other parameters are go through a learning process and automatically altered by the algorithm in the solving process. We chose to use Tabu Search Algorithm as our improvement algorithm since we tried the algorithms with various scenarios and it gave us the best solutions among other algorithms. These scenarios are made for 3 different capacities (20, 25 and 30) and 3 samples for each. The results are available in Appendix A, Table A. As it is seen, we tried with 190 employees and 50 bus stops which are generated randomly. In Table B, solutions for the various samples are shown for all the 3 algorithms. In Table C, the averages are seen for all the samples and it is obvious that Tabu Search Algorithm is the best among them. So, it is chosen as the improvement algorithm for our solution.

3. Validation and Performance Measurement

The company's current bus system information about routes, bus stop locations, employee addresses and number of employees in each station is unavailable to us. We can only reach the data about the total number of buses, their capacities and total number of employees. Hence we would not be able to use the real data that is necessary for testing our algorithm and comparing the results of the algorithm with the current system of the company. Rather than that, we build a similar system that the company is currently using. We have reached this similarity level by using various scenarios of bus stop locations and employee addresses, with different amounts, and we have used the exact capacity of buses. For instance, one of our scenarios is constructed with 700 employees and we have created 2100 potential stop locations. As a result, 37 feasible routes are obtained and 4 buses with capacity of 27 and 33 buses capacity of 19 is used. We have created scenarios with the given tool by the company. Also, the tool visualizes all points on the Ankara map so we can make inferences about our scenarios. Hence, we have created different scenarios of bus stop locations and employee addresses using the tool in order to use it as an input for the algorithm. Additionally, as the number of buses is determined by the algorithm, the number of buses can change according to the input and we have constructed our system according to company's existing buses and capacities. The visualization of the example and the output of it are available in Appendix B.

Our algorithm selects the bus stops that are nearest to employees while ensuring minimum distance for each route. That is, we used SVRP for the case and we also considered the comfort of employees. Since a heuristic algorithm cannot guarantee to an optimal solution, we made performance measurement by referencing to the Cplex model. As a result, we analyze the results of our heuristic by looking at the visualized routes' feasibility Then, we compared the results of our heuristic with the model of Cplex that finds the optimal route plan.

We found the routes for 10 different inputs with both model and heuristic. Those samples include 50 stops and 100 employees in order to get a reasonable solution. The capacity of a bus is 19 and if 100 employees are tried to assign those buses, it is obvious that 6 buses will be needed. The solutions that show routes and bus utilities are available in Appendix C, also the gaps (errors) table for all data is available in Appendix D. According to the gaps table in Appendix D, it is seen that the average error of the heuristic is calculated to be 11.10%.

4. Project Development and Implementation Plan

We have started to formulate our model by using Ilog CPLEX. After that, we have learned that the company has no license for this program therefore it was impossible for them to implement our model to their system. Although we have finalized our model on CPLEX, we decided on developing a heuristic algorithm by using Python for them to easily implement our algorithm to their system and CPLEX model is used to compare our results which we obtained from the heuristic approach.

Employees will be able to access the necessary information about company's transportation service through a user interface. The user interface is simply an information system for employees who use Havelsan's transportation service and employees log in to the interface with their id. This system will be designed by the IT department of the company. The input of this information system consists of the output of the heuristic algorithm and it is planned to show the following information, the license number of the employee's bus, the bus stop locations, the employees' assigned bus stop locations and available seats on any buses.

The passengers register to the system with their house locations and the system sends the possible bus stop locations and employee locations as input to the heuristic algorithm. After the execution of the algorithm the results can be seen in the admin panel and necessary information can be shared with the passengers. The chart that summarizes information about the development and implementation of the interface is given in Appendix E.

5. Conclusion

Having mentioned all the details of the project, we calculated the average error of the heuristic algorithm as 11.10%. As a result of the work done, buses will use the routes that are determined by our heuristic algorithm and Havelsan staff will not waste time reaching the company. Also, the employees will be informed through an interface called Admin Paneli that shows the stops to be used by the employees and the number of available seats in the buses.

In the study, the maps were not divided into regions, they were studied as a whole. As a future suggestion, the map can be divided by neighborhood. Possible stop points are selected for each region and routes covering those districts can be created. Another suggestion would be creating routes by dividing the map according to equal angles. In other words, by taking Havelsan as the center and dividing its surroundings by 360 degrees, it is possible to create routes between certain angles.

REFERENCES

Pop, P. C., Kara, I., & Horvat Marc, A. (2012). New mathematical models of the generalized vehicle routing problem and extensions. *Applied Mathematical Modelling. Simulation and Computation for Engineering*

and Environmental Systems, *36*(1), 97. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=msn&AN=MR2834995&site=eds-live
Valle, C. A., Martinez, L. C., da Cunha, A. S., & Mateus, G. R. (2010). Heuristic and exact algorithms for a min–max selective vehicle routing problem. *Computers and Operations Research*. https://doi.org/10.1016/j.cor.2010.10.010
Thampi and Khatter. Machine Learning and Metaheuristics Algorithms, and Applications. Springer Singapore, 2020.

APPENDICES

Appendix A. Comparison of the Results of Algorithms

Table 1. Parameters

	Capacity 1	Capacity 2	Capacity 3
Bus Capacity	20	25	30
Total Number of Employees	190	190	190
Total Number of Stops	**50**	**50**	**50**

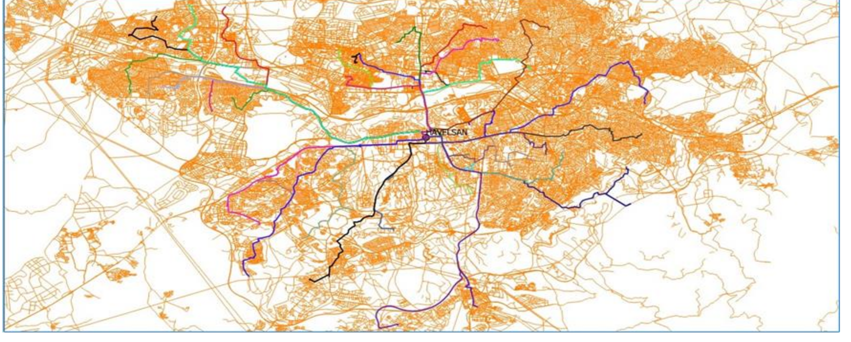
Table 2. Samples of Algorithms

Sample 1: Capacity: 20	Simulated Annealing	Guided Local Search	Tabu Search
Total Distance	472.04	453.24	451.61
Total Number of Bus	10	10	10
Sample 2: Capacity: 20	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	512.54	498.35	490.05
Total Number of Bus	10	10	10
Sample 3: Capacity: 20	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	437.58	411.8	401.29
Total Number of Bus	10	10	10
Sample 1: Capacity: 25	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	368.35	342.93	340.72
Total Number of Bus	8	8	8
Sample 2: Capacity: 25	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	406.15	389.74	388.69
Total Number of Bus	8	8	8
Sample 3: Capacity: 25	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	335.64	320.36	315.09
Total Number of Bus	8	8	8
Sample 1: Capacity: 30	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	358.3	346.94	341.74
Total Number of Bus	7	7	7
Sample 2: Capacity: 30	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	295.43	290.22	288.8
Total Number of Bus	7	7	7
Sample 3: Capacity: 30	**Simulated Annealing**	**Guided Local Search**	**Tabu Search**
Total Distance	315.12	309.86	307.36
Total Number of Bus	7	7	7

Table 3. Comparison of the Algorithms

Capacity: 20	Simulated Annealing	Guided Local Search	Tabu Search
Average Total Distance	474.05	454.46	447.65
Total Number of Bus	10	10	10
Capacity: 25	Simulated Annealing	Guided Local Search	Tabu Search
Average Total Distance	370.05	351.01	348.17
Total Number of Bus	8	8	8
Capacity: 30	Simulated Annealing	Guided Local Search	Tabu Search
Average Total Distance	322.95	315.67	312.63
Total Number of Bus	7	7	7

Appendix B. The Visualization and the Output of the Scenario



Routes	Distance Travelled (km)	Bus Capacity
Route 1	18,861	16
Route 2	15,818	27
Route 3	19,583	19
Route 4	19,512	18
Route 5	24,294	19
Route 6	15,532	25
Route 7	16,295	18
Route 8	13,091	19
Route 9	17,623	15
Route 10	17,642	19
Route 11	17,922	18
Route 12	17,807	26
Route 13	13,097	17
Route 14	17,556	18
Route 15	19,013	19
Route 16	17,839	27
Route 17	15,612	17
Route 18	17,801	18
Route 19	16,357	19
Route 20	18,996	19
Route 21	16,996	18
Route 22	15,513	17
Route 23	16,437	19
Route 24	24,745	18

Route 25	19,180	17
Route 26	16,533	19
Route 27	12,730	19
Route 28	19,386	18
Route 29	16,628	19
Route 30	26,328	17
Route 31	17,254	18
Route 32	18,908	19
Route 33	19,605	16
Route 34	18,691	17
Route 35	24,165	18
Route 36	15,184	16
Route 37	17,263	19
Total	665,794 km	700 employees

Appendix C. Outputs of CPLEX Model and Heuristic

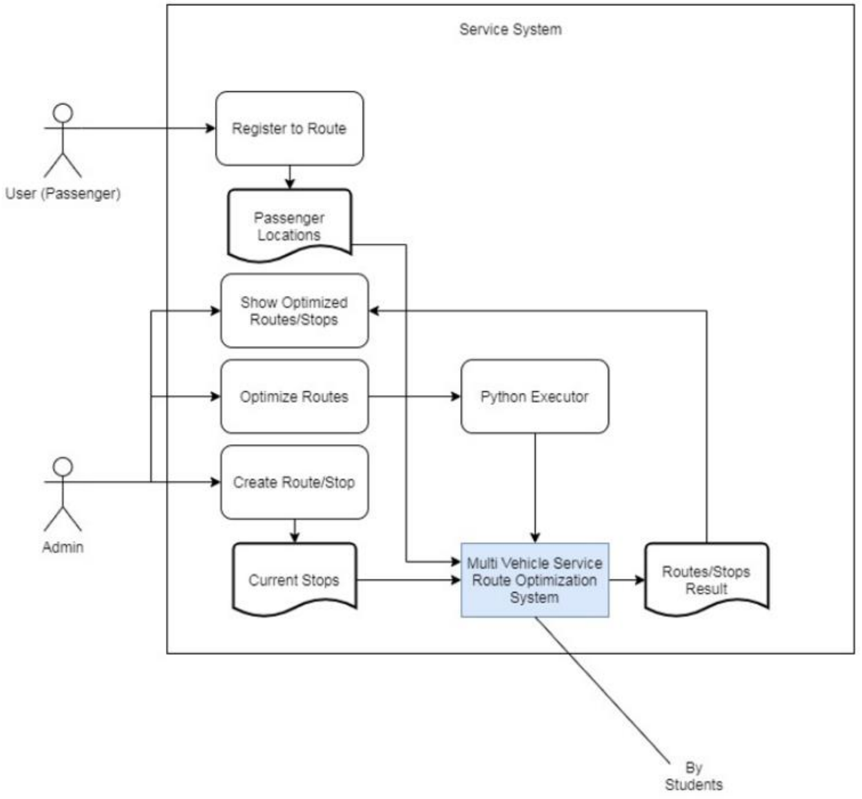
CPLEX MODEL			HEURISTIC		
Data 1	Route	Total Workers(#)	Data 1	Route	Total Workers(#)
Bus 1	1->3->40->16->1	16	Bus 1	1->8->12->14->1	19
Bus 2	1->12->14->20->1	19	Bus 2	1->4->11->15->3->2->1	19
Bus 3	1->25->1	9	Bus 3	1->7->9->5->13->1	15
Bus 4	1->27->41->1	19	Bus 4	1->6->1	18
Bus 5	1->31->1	16	Bus 5	1->10->1	19
Bus 6	1->49->33->1	17	Bus 6	1->15->1	9
Total Distance(km)		120.166171	Total Distance(km)		161.38752
CPLEX MODEL			HEURISTIC		
Data 2	Route	Total Workers(#)	Data 2	Route	Total Workers(#)
Bus 1	1->6->1	11	Bus 1	1->15->10->11->12->1	9
Bus 2	1->12->42->40->1	19	Bus 2	1->2->4->1	19
Bus 3	1->26->1	17	Bus 3	1->8->16->9->5->3->1	17
Bus 4	1->28->1	17	Bus 4	1->13->6->1	19
Bus 5	1->31->17->43->23->41->1	16	Bus 5	1->14->1	19
Bus 6	1->33->44->1	16	Bus 6	1->7->1	17
Total Distance(km)		121.991022	Total Distance(km)		127.908776
CPLEX MODEL			HEURISTIC		
Data 3	Route	Total Workers(#)	Data 3	Route	Total Workers(#)
Bus 1	1->4->13->8->11->2->1	19	Bus 1	1->11->42->1	18
Bus 2	1->12->6->1	19	Bus 2	1->22->1	14
Bus 3	1->3->7->1	16	Bus 3	1->24->5->1	17
Bus 4	1->10->1	15	Bus 4	1->26->25->27->1	18
Bus 5	1->9->1	19	Bus 5	1->31->1	15
Bus 6	1->5->1	12	Bus 6	1->33->1	18
Total Distance(km)		87.890597	Total Distance(km)		90.139404
CPLEX MODEL			HEURISTIC		
Data 4	Route	Total Workers(#)	Data 4	Route	Total Workers(#)
Bus 1	1->12->13->6->7->1	13	Bus 1	1->6->47->21->1	18
Bus 2	1->9->2->14->1	19	Bus 2	1->14->1	12
Bus 3	1->11->1	15	Bus 3	1->20->7->40->1	15
Bus 4	1->5->8->1	15	Bus 4	1->26->1	19
Bus 5	1->10->4->1	19	Bus 5	1->46->50->1	18
Bus 6	1->3->1	19	Bus 6	1->48->37->1	18
Total Distance(km)		125.62638	Total Distance(km)		170.04667
CPLEX MODEL			HEURISTIC		
Data 5	Route	Total Workers(#)	Data 5	Route	Total Workers(#)
Bus 1	1->2->25->1	16	Bus 1	1->13->15->8->9->1	19
Bus 2	1->26->1	15	Bus 2	1->7->6->14->10->3->5->1	19
Bus 3	1->33->5->1	19	Bus 3	1->11->1	17
Bus 4	1->45->22->14->1	19	Bus 4	1->12->1	9
Bus 5	1->46->1	10	Bus 5	1->2->1	17
Bus 6	1->50->42->39->13->29->1	19	Bus 6	1->4->1	19
Total Distance(km)		165.300762	Total Distance(km)		175.391282

CPLEX MODEL			HEURISTIC		
Data 6	Route	Total Workers(#)	Data 6	Route	Total Workers(#)
Bus 1	1->15->3->6->13->4->7->14->1	16	Bus 1	1->2->1	17
Bus 2	1->12->5->1	15	Bus 2	1->13->32->7->14->37->1	19
Bus 3	1->2->1	15	Bus 3	1->15->1	13
Bus 4	1->9->1	17	Bus 4	1->16->41->1	19
Bus 5	1->8->1	18	Bus 5	1->21->1	13
Bus 6	1->10->11->1	19	Bus 6	1->25->11->30->1	19
Total Distance(km)		108.761295	Total Distance(km)		109.250266
CPLEX MODEL			HEURISTIC		
Data 7	Route	Total Workers(#)	Data 7	Route	Total Workers(#)
Bus 1	1->7->1	16	Bus 1	1->15->1	15
Bus 2	1->14->42->1	19	Bus 2	1->14->11->1	18
Bus 3	1->16->1	10	Bus 3	1->16->3->8->9->5->7->4->1	16
Bus 4	1->28->9->1	17	Bus 4	1->10->1	18
Bus 5	1->34->11->2->1	19	Bus 5	1->13->1	19
Bus 6	1->47->22->32->46->1	19	Bus 6	1->6->2->12->11	14
Total Distance(km)		149.274428	Total Distance(km)		153.964365
CPLEX MODEL			HEURISTIC		
Data 8	Route	Total Workers(#)	Data 8	Route	Total Workers(#)
Bus 1	1->15->13->1	18	Bus 1	1->18->1	19
Bus 2	1->11->1	15	Bus 2	1->25->9->48->1	18
Bus 3	1->10->16->1	19	Bus 3	1->28->1	16
Bus 4	1->7->12->2->1	15	Bus 4	1->31->27->36->39->1	18
Bus 5	1->8->9->4->14->6->3->1	14	Bus 5	1->40->1	10
Bus 6	1->5->1	19	Bus 6	1->47->13->10->41->1	19
Total Distance(km)		127.936612	Total Distance(km)		165.236595
CPLEX MODEL			HEURISTIC		
Data 9	Route	Total Workers(#)	Data 9	Route	Total Workers(#)
Bus 1	1->12->10->6->1	15	Bus 1	1->4->1	9
Bus 2	1->11->7->1	19	Bus 2	1->11->1	17
Bus 3	1->9->1	12	Bus 3	1->16->39->9->5->1	19
Bus 4	1->8->5->2->15->13->14->1	16	Bus 4	1->18->1	17
Bus 5	1->4->1	19	Bus 5	1->40->26->1	19
Bus 6	1->3->1	19	Bus 6	1->45->27->12->1	19
Total Distance(km)		126.091517	Total Distance(km)		157.748068
CPLEX MODEL			HEURISTIC		
Data 10	Route	Total Workers(#)	Data 10	Route	Total Workers(#)
Bus 1	1->2->1	8	Bus 1	1->10->12->1	15
Bus 2	1->9->1	16	Bus 2	1->3->13->6->14->5->1	12
Bus 3	1->10->33->6->1	19	Bus 3	1->9->11->1	18
Bus 4	1->23->21->1	19	Bus 4	1->4->1	17
Bus 5	1->46->45->12->1	19	Bus 5	1->7->6->1	19
Bus 6	1->50->18->1	19	Bus 6	1->2->1	19
Total Distance(km)		137.261834	Total Distance(km)		137.974398

Appendix D. Gaps Table

Gap of Data 1	25.53%
Gap of data 2	4.63%
Gap of Data 3	2.49%
Gap of Data 4	26.00%
Gap of Data 5	5.75%
Gap of Data 6	0.43%
Gap of Data 7	3.05%
Gap of Data 8	22.57%
Gap of Data 9	20.07%
Gap of Data 10	0.52%
Average Gap	11.10%

Appendix E. Development and Implementation Plan of the System



Zaman Dizisi Talep Özellikleri Mühendisliği ve Tahmini

IBM

Proje Ekibi

Özgür Şafak Açıkalın, İzzet Egemen Elver, Selen Erkan, Başak Fiş, Berk Güney, Sudenur Soysal

Şirket Danışmanı

Ilyas Iyoob, IBM,
Chief Data Scientist ve Yapay Zeka Müdürü
Emrah Zarifoğlu, IBM,
Program Direktörü, Data Bilimi

Akademik Danışman

Dr. Öğr. Üyesi. Gizem Körpeoğlu
Endüstri Mühendisliği Bölümü

ÖZET

IBM, mevcut tahmin sistemlerinin yeterince iyi sonuç vermediğini belirtmektedir. Sorun, belirli veri kümelerindeki faktörlerin tahmin sonuçlarının doğruluğunu artıracak kadar açıklayıcı olmadığından kaynaklanmaktadır. Bu proje, çeşitli veri kümelerine otomatik algoritmalarla yeni öngörücü faktörler ekler ve tahmin performansını artırmayı amaçlar. Süreci otomatikleştirmek için bir kullanıcı arayüzü oluşturulmuştur.

Anahtar kelimeler: Tahmin, yapay zeka, veri bilimi, zaman serileri, özellik mühendisliği

1 Identification of the Engineering Problem

1.1 General information about the company

International Business Machines (IBM) was established in New York, USA in 1911 as the Computing- Tabulating- Recording Company (C-T-R). The general public holds a 41.9% stake at IBM. IBM has the highest market share in artificial intelligence applications according to International Data Corporations (IDC). Competitors of IBM are the major technology firms such as SAP, SAS, Google, Hewlett Packard (HP), Accenture, etc. IBM is a global brand that operates in more than 175 countries and it has approximately 356000 employees. Our project focuses on technology solutions, specifically analytics, artificial intelligence, cloud computing, and supply chain management.

1.2 Analysis and the interpretation of the data

We analyzed a detailed cloud service dataset by a world-wide known company to determine the scope of our project. The analysis on Gartner research shows that IBM's cloud services are falling back within its competitors in terms of completeness of vision and ability to execute. That is why IBM is categorized as niche players according to Gartner's (2019) analysis. Analysis denotes that IBM's services are not progressed with regards to hyper-scale architecture which means scaling appropriately as more demand is added to the system. This situation leads to problems with forecasting the demand accurately which is the scope of our project.

1.3 Problems and symptoms

Any kind of information and data are affected by many factors internally and externally. Determination of all the effective factors is nearly impossible due to the complexity of the environment. This is one of the main reasons for higher error levels than expected error levels on forecast results achieved by the current, traditional forecasting methods and predictive data analysis applications. One of the main roots of this problem, as also indicated by IBM, is existing limited and generic features of the data and insufficient usage of modern forecasting methods such as machine learning. Thus, the selection and elimination of useful features are crucial points of forecasting and our project. Therefore, our project focuses on feature engineering, feature selection and elimination methods, and creative and more accurate forecasting with lower error rates. Thus, the main problem is undescriptive and limited features. On the other hand, another problem is redundant features that are decreasing the time efficiency of forecasting. With our user interface we eliminate the redundant features in the input data which provides shorter run times.

2. Model Development and the Details of the Approach

2.1.1 Deliverables of the project and the system explanation

Initial outcomes of the project are detailed research on forecasting techniques, feature engineering, and machine learning techniques on time series forecasting analysis for the sake of both IBM and us to implement our project further. We offer a user interface that provides data preparation, feature generation and elimination, and model figuration by automation.

For example, the user working in the sales department of a company will upload the sales data of their top selling product as time series with just one column that denotes the date and will choose the sales quantity column to predict possible demand. Then, we create new features based on the original dataset. After this process, the necessary features are chosen automatically. New, creative, useful and statistical features that may be able to capture factors, seasonality, and trends better, so that are expected to increase forecasting accuracy will be modified to the input data automatically by our algorithm. If the user does not choose a specific method the system will apply all forecasting methods and utilize the one with the best results. If the data is numerical, we do regression and evaluate the results by R Square values and other error measurements. If one of the algorithms yields 3% higher R Square value, we select it as the best model. If not then we check for other error measures. For example, for the ice cream sales data the date will be indexed as year, month, day, time immediately. The new pre-generated features will be added to the dataset that was given by the user as an input. For example, the generic, statistical and normative features such as day, time, min-max, ispeakday, will be added to all the data that is given by the system but later will be eliminated by algorithms, automatically. Finally, testing or forecasting will be performed on the new merged and automatedly modified data set based on the user's choice. Details or other options that we provide with our user interface is fully explained on Section 3.3 The algorithm behind the user interface.

2.1.2 Resources that are utilized throughout the project

Open-source coding languages, R and Python, are utilized to reach a result on how these generated features are applicable on different datasets for both forecasting and various uses of data analysis. Java Script and Python are utilized for the user interface, and Python again, is used in order to implement the background programming of the interface.

SPSS Modeler, which is a product of IBM, is utilized for data selection on the forecasting of the specified dataset that the company provided. In the end our user interface is used for verification and validation.

2.1.3 How this project will eliminate the problems in the current system?

First, we generated new generic, statistical and conceptual features that would be beneficial for forecasting, to merge with the data sets that are given as an input by the user of our system. Then, we designed a UI system that follows the processes of feature generation, feature elimination and forecasting, automatedly. Our system design offers:

1) Automation of data analysis processes
2) Flexibility, user customization (in terms of methodology) during forecasting
3) User friendly UI design that merges three complex processes: feature generation, feature elimination and forecasting
4) Higher forecasting accuracy
5) Detailed output report for better understanding of the data for the company

We provide the most successful forecasting method for each time scale with a comparison of forecasting methods in terms of mean absolute percentage error (MAPE), mean squared error (MSE), and mean absolute error (MAE). MAPE expresses accuracy as a percentage of the error. MSE expresses average of a set of errors. MAE expresses accuracy in the same units as the data, which helps conceptualize the amount of error.

Our algorithm contains many mathematical functions and uses them to find mutual relations between output and input values. Therefore, there are some computational restrictions for these functions. For instance, data sets may contain 0 values in input variables. In this situation, log0 causes computation errors. To eliminate these problems, we added new mathematical conditions avoiding these problematic terms in our algorithm.

In the end, the user of our system (IBM) is able to reach better forecasting results and having better understanding of their data and variables, which are their main concern, with the help of our UI system that combines complex algorithms and feature based processes automatedly. Also, with the reports provided, the company is able to reach critical information about their correlation values, error terms, specific performance indicators, etc., in addition to forecasting, which help the company on their evaluation of the data.

3 Benefits to the company

3.1 Inputs and outputs of the proposed system

The main input of the system is the data to be analyzed. Users are also expected to specify what is the target feature, what will be predicted. Other than these, user will be expected to enter some specified inputs for the forecasting methods that cannot be done automatically. We implemented a system that will take default parameters for algorithms for some specifics and will allow the qualified users to intervene with the system. We developed our work on this issue on creating the user interface work package. The output of the system is the output reports for specified input variables and model analysis.

The output report provides storing and visualization of many information related to the input data. After the implementation of our code, new folders named pictures and heat maps will be created to store visualization outputs of the algorithms.

3.2 Major components

3.2.1 User Interface

The user interface is the main component of the system. This is the platform for communicating with the user and collecting information about the dataset. Algorithms behind the user interface are our other major components. The system provides forecast results based on the given data by using machine learning algorithms and traditional forecasting algorithms at the same time. Machine learning algorithms is used to learn the relation between features in the dataset and predict a forecasting result based on these relations. Our aim is to predict better results compared to traditional results. The user interface provides three classification methods: Decision Tree, Random Forest, Logistic Regression.

3.2.2 Provided regression and classification methods

Decision Tree: The decision tree learns recursively by splitting the dataset from the root onwards (node by node manner) according to the splitting metric at each decision node.

Random Forest: Random forests takes average of multiple decision trees, trained on different parts of the same training set, with the goal of reducing the variance.

Logistic Regression: The algorithm models the probability of a certain class or event existing such as pass/fail, win/lose or healthy/sick. A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variables.

Our user interface provides three regression methods: Linear Regression, MARS Regression, LASSO (Least Absolute Shrinkage and Selection Operator) Regression.

Linear Regression: Linear regression shows the relationship between two variables using a straight line. The algorithm attempts to draw a line that comes closest to the data by finding the slope and intercept that define the line and minimize regression errors.

MARS Regression: The algorithm automates variable selection, detection of interactions, and accounts for non-linearities.

LASSO Regression: The method calculates a penalty which affects the value of coefficients of regression. As penalty increases more coefficients becomes zero (ineffective). Current snapshot of the user interface can be seen on the Appendix, figure 4.

The forecasting methods that we provide results with our user interface are as follows:

Seasonal Autoregressive Integrated Moving-Average (SARIMA): The Seasonal Autoregressive Integrated Moving Average method predicts the next step in the sequence as a linear function of the differenced observations, errors, differenced seasonal observations, and seasonal errors at previous time steps.

Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors (SARIMAX): The Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors is an extension of the SARIMA model that also consists exogenous variables. Exogenous variables can be thought as parallel input sequences that have been observed at the same time steps as the original series.

3.3 The algorithm behind the user interface

The flowchart of the operations that is processed in the model interface background can be seen on Appendix, figure 5. As a starting point, we generate general features from the given dataset. After features are generated, we focused on the feature elimination algorithms and automation of all processes. The user interface which is the compact data testing and forecasting tool that we provide, gets the data from the user as the input and the user chooses the target column to be forecasted or tested and the date column. Then the user decides which option to choose from PCA and feature selection for feature modification. Then the algorithm performs data preparation, correcting and denoting missing values, removing duplicates, etc. Then, the user must decide whether to test the data or perform forecasting. If the user chooses testing, he/she has to input what percentage of the data will be used for testing-training. If the user chooses forecasting, he/she has to choose which index to start forecasting from. For visualization and splitting (test-train) purposes the user should input which index is the beginning of the testing or the forecasting should start. No matter if the data is utilized for testing or forecasting purposes, the user has to choose a regression algorithm if the data is continuous and has to choose a classification algorithm if the data is categorical. Then the program runs and provides three kind of output reports giving information about the process, the methods and the forecast: the summary report, extensive report and the technical report. Any information on the output report, such as fitting graphs, standard deviation, added features, chosen features, correlation values, etc., all of the output information, will be stored as JSON objects. If testing was performed by the algorithm, the system compares the real value of the test data and the forecasts, and draws a graph of them with respect to the time. If forecasting was performed, the system outputs forecast values.

4 Validation and Verification

For verification, we worked on the happiness score data of the countries. The initial dataset had few features such as life expectancy and family bond scores but as outside-in features, we added new features such as government spending, property rights index, region, labor and business freedom index. Then, we used SPSS Modeler's feature selection algorithm. the algorithm considers 19 of our 31 features as important. 17 features that are chosen are our newly added features. Also, we tried linear regression with both the original data and our new data. Results in Table 1 that is in Appendix show that the model with hour new features yields less mean absolute error.

We worked with a similar methodology for the Seatbelts dataset which is in the R library. We tried to see the effects of inside- out features by using the MARS algorithm. We created two models. The first one is the linear regression model with a step algorithm with the original dataset. After that, we tried to use the MARS algorithm. The results show that the model with the MARS algorithm has a higher R- squared value and less mean absolute error loss.

We have also calculated the correlation values of the combination of existing and added features and structured the correlation output as a heat- map for a better rational and visual understanding of the data, which supported the verification of our system design indirectly. The heat map shows the correlation between all of the features between themselves and shows the amount of the correlation between all of the features and the target column (will be taken as an input in our UI design). The correlation calculation outputs and the heat map, are also a part of the user interface output reports that we provide.

The comparison of the statistical measurements denotes mean absolute error, mean squared error for the prediction is lower when we add our mathematical data. Moreover, we implemented linear regression for the models with and without our added mathematical features, and by the ANOVA test, we observed that our model with newly added features is significantly better than the simple model based on the p value of the test. Also, Akaike's Information Criterion (AIC) comparison shows that, our model has lower AIC value which denotes that the model with our added features is significantly better than the previous one. Statistical measurements and most important variables are also identified. As statistical error measurements exhibit added features enhance the precision of our forecasts with linear regression.

4.1 Validation plan of the system

We validated our project with a contemporary problem: Covid-19. We have gathered data of number of people who have tested positive for Covid-19 around the world and predicted the numbers for the upcoming days for Turkey case. We have successfully forecasted them with minimal error rates.

In addition, we validated our project with the dataset that IBM send us. We have tested our project on the dataset and confirmed that our newly added features are selected by our feature selection algorithm for better prediction. Thus, we can confirm that with our new features added to their dataset predictions are more accurate and R squared values are higher with less error (MAE, MAPE etc.) rates. We have also made the project available in GitHub with a ReadMe document for their ease of use.

After our work on Covid-19 dataset we have decided to (with our instructor's guidance) validate our results by comparing with known cases from academic papers. From the "Time Series Analysis and Its Applications with R Examples" by Shumway, Robert&Stoffer and David, the dataset of diseases that includes the air pollution and temperature of air is selected to validate the results by comparing with an academic case. In our program with this dataset, we concluded with better results. Our best model comes from MARS regression with 0.806 R^2 value. In our program the values of AIC and BIC they are 2.97 and 2.98 respectively. In comparison their results are 4.72 and 4.77 which results worse. As it can be seen that our error level is much less from the authors' result. We have taken these results as a success and considered our project validated.

REFERENCES

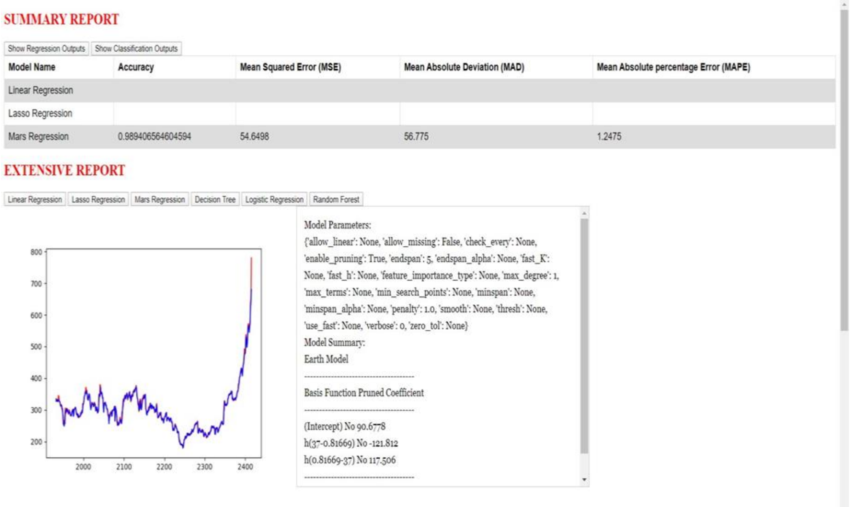
Bala, R., Gill, B., Smith, D., & Wright, D. (2019, July 16). "Magic Quadrant for Cloud Infrastructure as a Service" *Worldwide*. Retrieved from https://www.gartner.com/doc/reprints?id=1-1CMAPXNO&ct=190709&st=sb.

Brockwell, P. J., & Davis, R. A. (2016). "Introduction to time series and forecasting" *Springer*. Retrieved November 22, 2019, from http://garfield.library.upenn.edu/classics1989/A1989AV48500001.pdf

Chatfield, C. (2000). "Time-series forecasting" Retrieved November 23, 2019, from https://epdf.pub/time-series-forecasting.html.

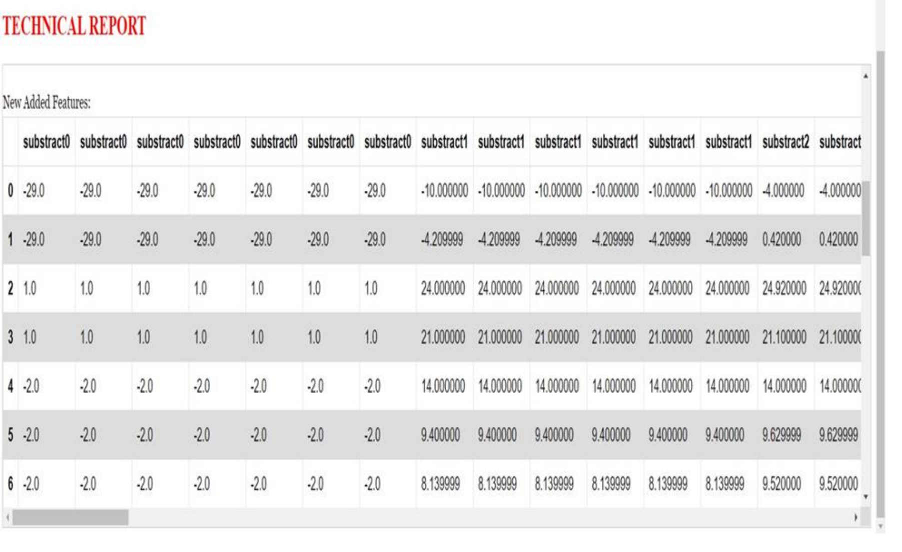
Shumway, Robert & Stoffer, David. (2011). "Time Series Analysis and Its Applications With R Examples" *Springer*. September 2016.

APPENDIX



Model Name	Accuracy	Mean Squared Error (MSE)	Mean Absolute Deviation (MAD)	Mean Absolute percentage Error (MAPE)
Linear Regression				
Lasso Regression				
Mars Regression	0.989406564604594	54.6498	56.775	1.2475

Figure 1: The Summary Output Report and the Extensive Output Report

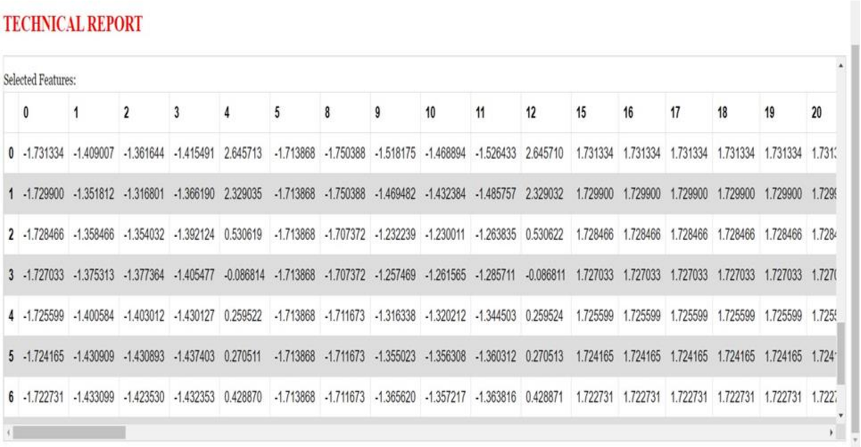


TECHNICAL REPORT

New Added Features:

	substract0	substract0	substract0	substract0	substract0	substract0	substract0	substract1	substract1	substract1	substract1	substract1	substract1	substract2	substract
0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-10.000000	-10.000000	-10.000000	-10.000000	-10.000000	-10.000000	-4.000000	-4.000000
1	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-4.209999	-4.209999	-4.209999	-4.209999	-4.209999	-4.209999	0.420000	0.420000
2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	24.000000	24.000000	24.000000	24.000000	24.000000	24.000000	24.920000	24.92000(
3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000	21.100000	21.10000(
4	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0	14.000000	14.000000	14.000000	14.000000	14.000000	14.000000	14.000000	14.00000(
5	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0	9.400000	9.400000	9.400000	9.400000	9.400000	9.400000	9.629999	9.629999
6	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0	8.139999	8.139999	8.139999	8.139999	8.139999	8.139999	9.520000	9.520000

Figure 2: The Technical Output Report – New Added Features



TECHNICAL REPORT

Selected Features:

	0	1	2	3	4	5	8	9	10	11	12	15	16	17	18	19	20
0	-1.731334	-1.409007	-1.361644	-1.415491	2.645713	-1.713868	-1.750388	-1.518175	-1.468894	-1.526433	2.645710	1.731334	1.731334	1.731334	1.731334	1.731334	1.731
1	-1.729900	-1.351812	-1.316801	-1.366190	2.329035	-1.713868	-1.750388	-1.469482	-1.432384	-1.485757	2.329032	1.729900	1.729900	1.729900	1.729900	1.729900	1.729
2	-1.728466	-1.358466	-1.354032	-1.392124	0.530619	-1.713868	-1.707372	-1.232239	-1.230011	-1.263835	0.530622	1.728466	1.728466	1.728466	1.728466	1.728466	1.728
3	-1.727033	-1.375313	-1.377364	-1.405477	-0.086814	-1.713868	-1.707372	-1.257469	-1.261565	-1.285711	-0.086811	1.727033	1.727033	1.727033	1.727033	1.727033	1.727
4	-1.725599	-1.400584	-1.403012	-1.430127	0.259522	-1.713868	-1.711673	-1.316338	-1.320212	-1.344503	0.259524	1.725599	1.725599	1.725599	1.725599	1.725599	1.725
5	-1.724165	-1.430909	-1.430893	-1.437403	0.270511	-1.713868	-1.711673	-1.355023	-1.356308	-1.360312	0.270513	1.724165	1.724165	1.724165	1.724165	1.724165	1.724
6	-1.722731	-1.433099	-1.423530	-1.432353	0.428870	-1.713868	-1.711673	-1.365620	-1.357217	-1.363816	0.428871	1.722731	1.722731	1.722731	1.722731	1.722731	1.722

Figure 3: The Technical Output Report – Selected Features

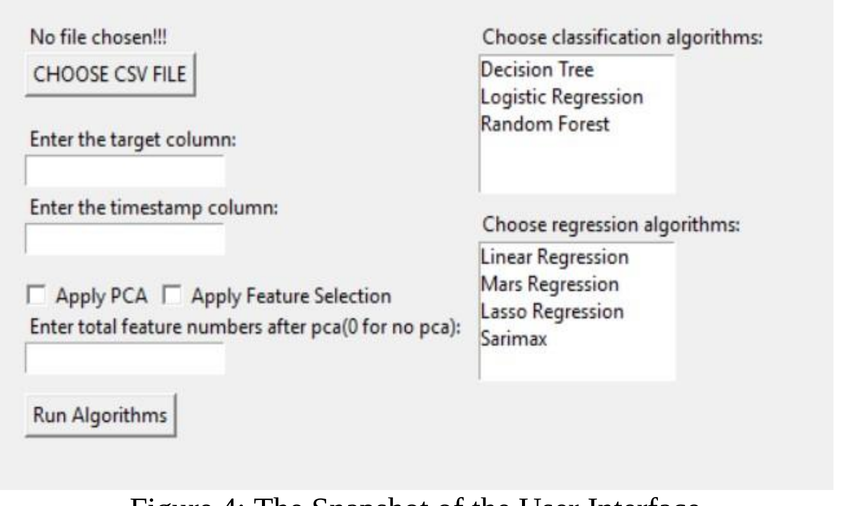


Figure 4: The Snapshot of the User Interface

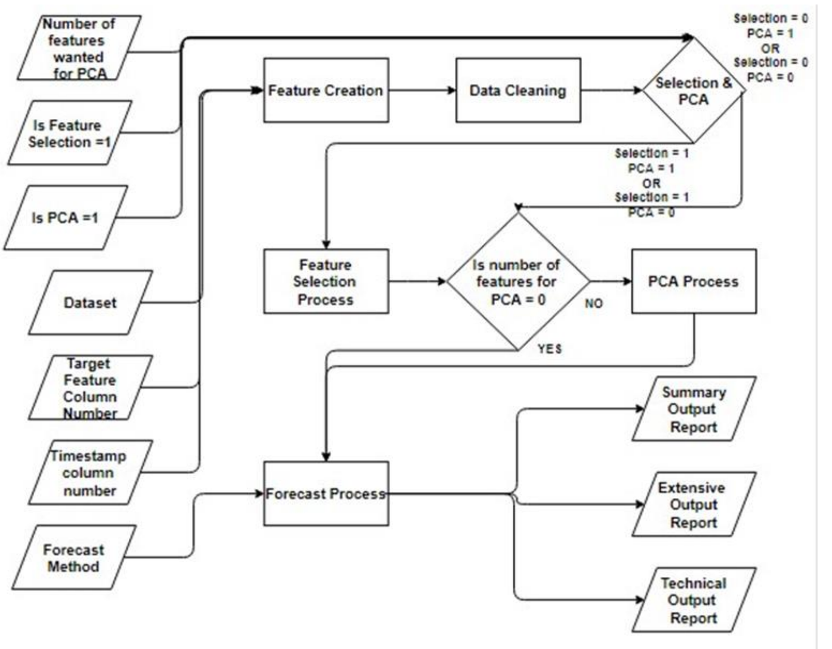
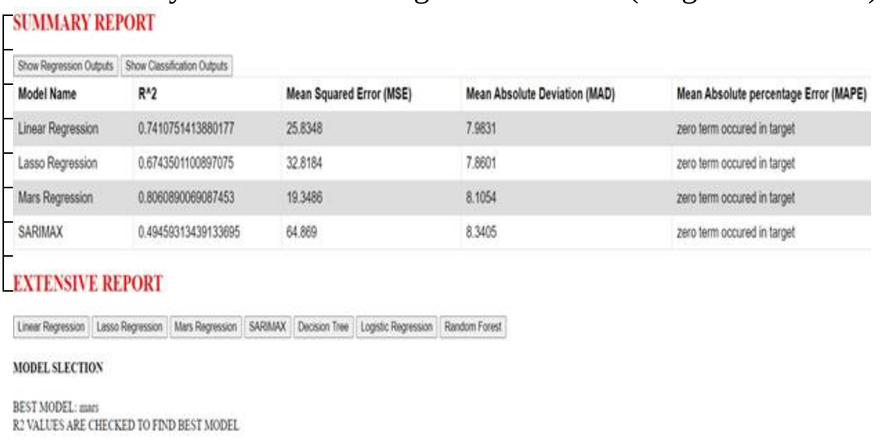


Figure 5: Flowchart of the operations

Original dataset	
Minimum Error	-1.771
Maximum Error	1.678
Mean Error	-0.065
Mean Absolute Error	0.563
Standard Deviation	0.685
Linear Correlation	0.829
Occurrences	74

Table 1: Analysis of the linear regression models (Original and New)



SUMMARY REPORT

Show Regression Outputs | Show Classification Outputs

Model Name	R^2	Mean Squared Error (MSE)	Mean Absolute Deviation (MAD)	Mean Absolute percentage Error (MAPE)
Linear Regression	0.7410751413880177	25.8348	7.9831	zero term occured in target
Lasso Regression	0.6743501100897075	32.8184	7.8601	zero term occured in target
Mars Regression	0.8060890069087453	19.3486	8.1054	zero term occured in target
SARIMAX	0.49459313439133695	64.869	8.3405	zero term occured in target

EXTENSIVE REPORT

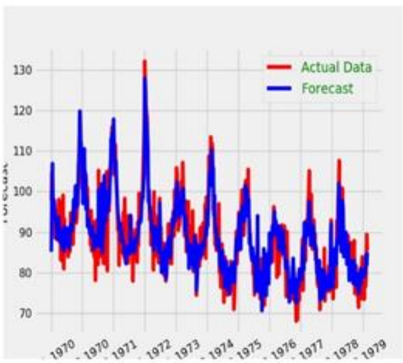
Linear Regression | Lasso Regression | Mars Regression | SARIMAX | Decision Tree | Logistic Regression | Random Forest

MODEL SLECTION

BEST MODEL: mars
R2 VALUES ARE CHECKED TO FIND BEST MODEL

Figure 6: The output of our program

$$M_t = \beta_0 + \beta_1 t + w_t \tag{2.18}$$
$$M_t = \beta_0 + \beta_1 t + \beta_2 (T_t - T_\cdot) + w_t \tag{2.19}$$
$$M_t = \beta_0 + \beta_1 t + \beta_2 (T_t - T_\cdot) + \beta_3 (T_t - T_\cdot)^2 + w_t \tag{2.20}$$
$$M_t = \beta_0 + \beta_1 t + \beta_2 (T_t - T_\cdot) + \beta_3 (T_t - T_\cdot)^2 + \beta_4 P_t + w_t \tag{2.21}$$



AIC: 2.970491747065617
BIC: 2.9871471858356102
SSE: 9829.065857758691
Explained Variance: 0.8061
Mean Squared Log Error: zero term occured in target
MAE: 3.4941
RMSE: 4.3987
OUTPUTS

	index	TRUE VALUES	FORECASTED VALUES
0	508	97.85	85.167756
1	509	104.64	97.796783
2	510	94.36	106.856950

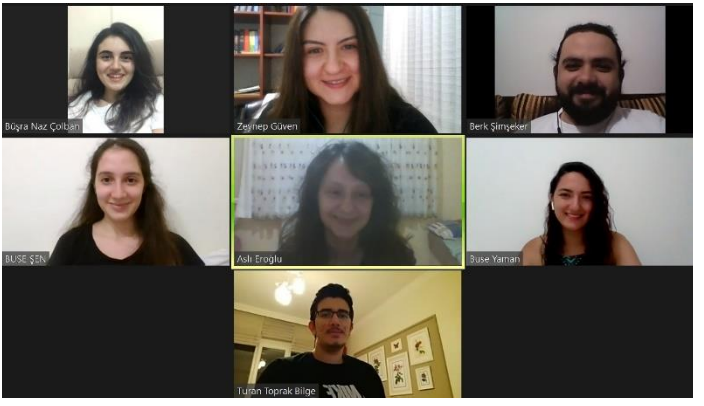
Table 2.2. Summary Statistics for Mortality Models

Model	k	SSE	df	MSE	R^2	AIC	BIC
(2.18)	2	40,020	506	79.0	.21	5.38	5.40
(2.19)	3	31,413	505	62.2	.38	5.14	5.17
(2.20)	4	27,985	504	55.5	.45	5.03	5.07
(2.21)	5	20,508	503	40.8	.60	4.72	4.77

Figure 7: The output of the reference book

Bir Şeker Üretim Şirketinde Network Tasarım ve Üretim Planlama Modellemesi

INFORMS Yarışma Projesi



Proje Ekibi
Turan Toprak Bilge, Büşra Naz Çolban, Aslı Eroğlu,
Zeynep Güven, Buse Şen, Berk Şimşeker, Buse Yaman

Akademik Danışman
Dr. Öğr. Üyesi. Özlem Çavuş
Endüstri Mühendisliği Bölümü

ÖZET

Bu proje kapsamında şeker üretim şirketi için, beş üretim tesisi arasında yapılması planlanan hammadde transfer miktarları belirlendi ve her üretim tesisi için altı aylık şeker talebini karşılamaya yönelik üretim planları hazırlandı. İlk aşama olarak üretim tesisleri arasındaki transfer miktarlarını ve her tesiste üretilecek şeker miktarını belirleyen bir optimizasyon modeli oluşturuldu. Bu optimizasyon modeli ile transfer ve üretim maliyetlerinin en aza indirilmesi amaçlandı. İkinci aşamada optimizasyon modeli, ilk aşamadan elde edilen transfer ve üretim miktarları kullanılarak, üretim planının çizelgelenmesini sağlayacak şekilde modifiye edildi. Son aşama olan simülasyon modeli ile makinelerin çalışma hızlarının rastgeleliğini de göz önünde bulundurarak optimizasyon modellerinin sonuçları değerlendirildi. Bu üç aşamanın sonunda altı ay içerisinde şeker talebinin en az maliyet ile karşılanmasını sağlayan üretim planı ve transfer miktarları elde edildi.

Anahtar Kelimeler Optimizasyon Modeli, Simülasyon Modeli, Üretim Planlama

Network Design and Production Planning Modelling in a Candy Manufacturing Company

1. System Analysis

1.1 Description of the Company

Jelly Bean Manufacturing Company is a fictitious company that is introduced in the problem statement of the 2020 Informs OR & Analytics Student Team Competition which is a competition conducted by Informs and sponsored by Bayer. However, this hypothetical company represents Bayer, and the given data is the real historical data of that company.

Jelly Bean Manufacturing is headquartered in Chicago and operates in 5 different facilities located in Green Bay (WI), Detroit (MI), Columbus (OH), Springfield (MO), and Omaha (NE). The company produces 4800 types of jelly beans that are in 40 different colors, 5 different sizes, 12 different flavors, and 2 packaging types. The production can be done in any of the facilities. All facilities have different production capacities since the processing rates of the machines, the number of drums and the number of machines in facilities differ.

1.2 Production System

The raw material jelly beans are already colored, so the coloring operation is not in the scope of the problem. Production of jelly beans consists of three main stages which are classification, flavoring, and packaging. Initially, raw materials are stored in Raw Material Inventory (RMI) drums and an RMI drum has only one color but different sizes of jelly beans, so the process begins with classifying the mixed jelly beans according to their sizes.

After classification, jelly beans are stored in pre-finish inventory (PFI) drums and sent to the pre-finish operation where the jelly beans are flavored with 12 different flavors according to demand. Flavored jelly beans are stored in the packaging inventory (PI) drums until the packaging process starts. Finally, jelly beans are packaged in bags or boxes according to the demand, and the production process finishes with the packaging of the jelly beans. The visualization of these processes can be seen in Figure 1.

2. Project Description

2.1 Problem Definition

The objectives of this project are to determine optimal raw material transfer amounts between facilities and internal work orders for all facilities so that the fulfilled demand for Halloween is maximized and the total cost is minimized. Cost includes the cost of transfer of raw materials between facilities and the cost of bag/box packaging costs for each facility. The deliverables of the project are the internal work orders for all facilities, the schedule of internal work orders, and the amount of jelly beans that should be transferred between facilities.

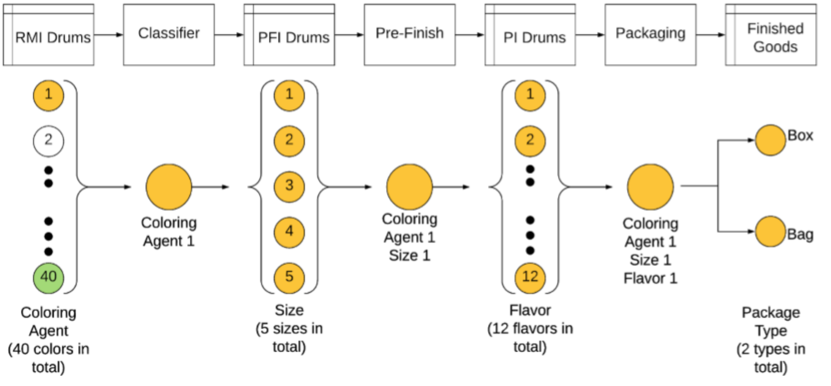


Figure 1: Process Flow

The number of machines, the number of drums, and drum capacities in each facility are different which makes the production capacities of facilities different. Additionally, processing rates of the machines depend on the type of jelly beans processed except for the classifier machine. Therefore, to use the production capacities of facilities in the most efficient way and not to exceed them due to this randomness, determining raw material transfer amounts between facilities and internal work orders assigned to the facilities is a significant decision to be made to achieve the objectives. The initial inventory level in each facility should be consistent with the production capacity of that facility and work orders assigned to that facility. To ensure this consistency, many constraints must be taken into account when deciding on the raw material transfer amounts and internal work orders. All decisions must be made before the production period since there is no chance to transfer raw materials between facilities and change the production plan once the production starts.

3. Methodology Approach

Initially, to be able to use the data of random processing rates of flavoring and packaging machines in the models that are built, a detailed data analysis have been made. After the data analysis, the methodology approach to solve the problem is decided. This approach consists of two stages of optimization models and a simulation model for each facility. It is decided to use the means of processing rates in the optimization model and generate the random processing rates from fitted distributions in the simulation model. The approach can be seen in the flow chart provided in Figure 2.

In the first stage, an optimization model is designed to determine the transfer amounts and the internal work orders given to each facility. The objective is to minimize total transportation and production cost.

Besides the first optimization model, a second optimization model is designed to find the schedule of the internal work orders for each facility. The outputs of the optimization models, which are the scheduled internal work orders, are given as inputs to the simulation model.

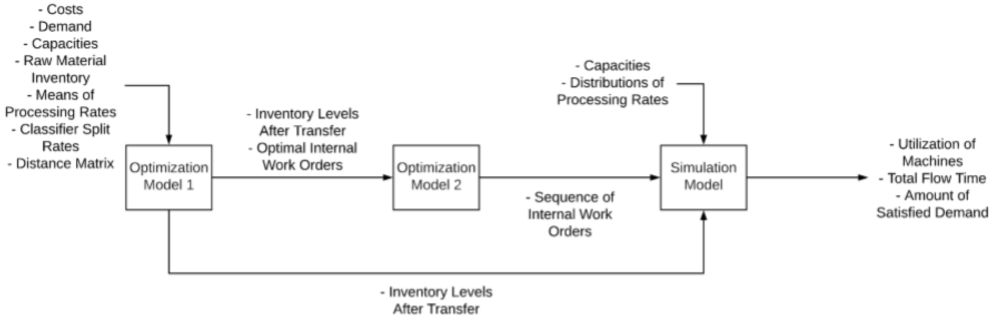


Figure 2: Approach Flow Chart

The simulation model constructed for each facility simulates the production systems of facilities by including all the constraints, assumptions, and randomness of processing rates.

3.1 Data Analysis

The company provides data for the following items: deterministic demand data for each product type, processing rate of the classifier machine for each site, classifier split percentages with respect to size and color, the distances between the facilities, initial inventory levels in each facility, number of drums and machines in each facility, capacity of different types of drums in each facility and historical data of processing rates of flavoring and packaging machines. Historical data of processing rates provided by the company depend on the type of jelly beans processed in the machine, for instance, the processing rates depend on the site-size-flavor combination for the pre-finish operation and the site-size-package type combination for the packaging operation. 1000 historical data are available for each combination. To be able to use the data of random processing rates in the models, distributions are fitted to these data. It is decided to use the expected values of these random processing rates in the optimization models and generate the random processing rates in the simulation model from fitted distributions. With the help of statistical analysis software Minitab, the historical data are analyzed to fit a distribution to each processing rate data. After the major distributions are fitted to the data, it is seen that most of the data are coming from normal distribution. However, due to long-tail of normal distribution, the data sampled from this distribution may include some negative or too large values. Therefore, truncated normal distributions are used in simulation models to avoid getting negative or too large processing rates. Besides, a few historical data are used as they are since none of the well-known distributions is a good fit because of low p-values. For these data sets, each of 1000 values is assumed to be observed with probability 1/1000. This means empirical discrete distribution is fitted to these data.

3.2 First Stage Optimization Model

3.2.1 Purpose and Assumptions of the First Optimization Model

The first optimization model aims to find the transfer amounts of jelly beans between facilities and internal work orders given to each facility. The objective of the model is to minimize total cost which consists of transportation and production costs. In this model, demand fulfillment is written as a constraint instead of writing it as a part of the objective function. It is assumed that demand will be fully satisfied after checking and verifying that the model gives a feasible solution when demand is forced to be satisfied completely. Additionally, production is not only made for satisfying the demand but also all the excess raw materials are processed in the given time period. All the excess raw materials are packaged in bags. Flavor assignments to the excess raw materials are made by considering the percentage of each flavor over the total demand of each color-size combination. Instead of assigning flavors by aiming the minimum flow time, these ratios are taken into account to have a meaningful stock of jelly beans by considering the possibility of selling these excess jelly beans in the future. The details of the optimization model can be found in Appendix A.

3.3 Second Stage Optimization Model

3.3.1 Purpose and Assumptions of the Second Stage Optimization Model

The aim of building a second optimization model is to find the schedule of the internal work orders. Even though with the given data and mean processing rates all the demand can be satisfied, in case of a scenario in which the demand cannot be fully satisfied, by scheduling the work orders the percentage of demand satisfied can be increased by decreasing the total flow time.

The objective of this model is to minimize the weighted sum of total cost and the total amount of unsatisfied demand, and the minimization of unsatisfied demand amount is prioritized by giving that part of the objective a larger weight. Unlike the first optimization model, demand fulfillment is written as a part of the objective function in this optimization model.

A new parameter β is defined in this model that represents the capacity usage of facilities. It is observed that when the value of β is less than or equal to 0.5, the model cannot satisfy all the demand. Therefore, by decreasing the value of β from 0.5 to 0.1 by small decrements, it is observed that when the capacity is decreased, which types of jelly beans are preferred to be produced first and which types to be produced last or not to be produced at all. Two new measures are defined in this model to schedule colors in facilities and flavors within colors and facilities. These measures are calculated by dividing the production amount of each type assigned to each facility by the optimal production amount of that type assigned to that facility. By decreasing the value of β, the values of these two measures are saved, and the average values of these measures are calculated. Finally, the average values are sorted from largest to smallest so that the schedule is determined.

In this model, the internal work orders and end inventory levels after transportation that are determined by the first stage optimization model are given as inputs. The details of the second optimization model can be found in Appendix B.

3.4 Simulation Model

3.4.1 Aims, Inputs and Outputs of the Simulation Model

There are two main aims of building simulation models for each facility. The first one is simulating the processes of facilities by including randomness of processing rates and all the assumptions and constraints in the problem definition. This is especially important to obtain realistic completion time results for each facility since in the optimization models most of the constraints related to drums and change-over-time in the flavoring machine, and randomness of processing rates are not included even though their effect on total completion time is not negligible. The second aim is evaluating the outputs of the optimization models by considering the performance measures such as total completion time, utilization of machines, therefore bottleneck operations, and the percentage of demand satisfied.

Inputs of the simulation models are internal work orders in schedule, the distributions fitted to processing rates of flavoring machines for each size-flavor combination, and the distributions fitted to processing rates of packaging machines for each size-packaging type combination. The internal work orders as the input of simulation models are determined by the first optimization model and they are scheduled according to the obtained schedule from the second optimization model. The fitted distributions to processing rates of machines are determined by data analysis. Additionally, the number of drums, their capacities, and the number of machines are other input data required. Since there is a lot of input data to be read by the simulation model, instead of entering these input data to Arena by hand, they are read from Excel files.

Outputs of the simulation models are total completion time, percentage of demand satisfied, utilization of machines, and total cost of production. These are also our performance measures to evaluate the results of optimization models.

3.4.2 Evaluating the Simulation Results

Since the pre-finish and packaging operation rates used in the simulation model are generated using random distributions, it is necessary to decide how many replications to be run to obtain more realistic average values of performance measures. To decide on the appropriate number of replications, the half-width values given by the Arena are examined for a different number of replications. It is decided that running 40 replications is sufficient to obtain meaningful results. Considering all these, simulation results show that all demand can be met in 6 months period and all excess raw material can be processed in this period for all facilities in all replications. In addition, comparing the results of the optimization and simulation models, it is observed that the flow time, bottleneck operations and cost values for each facility are consistent with each other.

4. Analytical Solutions and Results

4.1 Transfer Amounts of Raw Materials

According to the results of the first stage optimization model, transfers are done from Columbus, Detroit, and Springfield to Omaha. The transferred amount for those facilities is 500,000 lbs, which is the capacity of the trucks. Costs differ according to the distance between facilities that transfer is done. The transfer cost from Columbus to Omaha is $27,300, from Detroit to Omaha is $25,500 and from Springfield to Omaha is $22,050.

This result shows that Omaha is an efficiently working facility and it has a production capacity to produce much more than its initial raw material amounts.

4.2 Production Amounts, Days and Costs

Table 1 is filled using both the results of the first optimization model and the simulation model. Here, for the first and third columns, the results are taken from the first optimization model. The second column results are taken from the simulation model outputs.

Table 1. Facility-Based Production Results

Manufacturing Site	Total Production (lbs)	Total Number of Days to Complete Production	Total Production Cost ($)
Green Bay	2,360,152	63.99	2,407,355.04
Omaha	11,739,298	157.73	11,723,575.49
Springfield	3,487,516	130.08	3,823,259.41
Columbus	8,395,866	156.45	9,125,227.39
Detroit	10,355,579	168.79	10,872,998.08

4.3 Feedback Mechanism

Total flow times calculated by the optimization model are slightly less than the flow times from the simulation model since the simulation model considers the randomness and all the dynamic changes while the optimization model does not. Thus, although time limit constraints in the optimization model aim to restrict the processing times not to exceed the 6 months production period, we were able to restrict them with a 60-120 hours tolerance. This does not cause any trouble in our scenario as we can easily meet the demand within the given time. However, in a case where the demand cannot be satisfied in the given time, the optimization model may need a more realistic time limit. To provide the optimization model more meaningful time limits, a feedback mechanism is built between the optimization model and the simulation model.

First, we decrease our time limit to 1000 hours (new value of pt parameter - see Appendix A for the model) to create a case for which demand is not satisfied. The objective of the first stage optimization model is changed to the objective of the second stage optimization model, which includes the minimization of unsatisfied demand. Then, we run the optimization model, put its production plan to the simulation model, run the simulation model, and redetermine the time limits for each facility. Note that time limits are different for each facility now (pt is turned into pt_i). The new time limits (new pt_i's) are calculated by multiplying the desired time (1000) with the previous time limit and then dividing this result by the simulation completion time, which basically implies cross multiplication. We repeat this iteratively until we obtain a completion time of 1000 hours with a 5-hour tolerance level for each facility as a result of the simulation model. After 4 iterations, the desired values are achieved. Namely, in the case in which the demand cannot be met in the given time, time constraints in the optimization model are made more realistic by this feedback mechanism.

5. Conclusion

In conclusion, firstly the production plan that minimizes the total cost is determined by using two stages of optimization models. Simulation models that are built for each facility are separately run to simulate the production system by considering all constraints and assumptions of the production system. According to the results of the simulation models, it is shown that all facilities complete their production in the given production period. Therefore, in 6 months production period all the demand is satisfied and all the excess raw materials are processed by the determined production plan with the total cost of $ 38,027,315. Furthermore, for a case that facilities cannot complete the whole production in the given period, a feedback mechanism between the first stage of the optimization model and the simulation model is constructed. Hence, by improving the sensitivity of time limit constraints of the optimization model, the optimization model will be forced to obey the time limitations more strictly.

REFERENCES

[1] Karabakal, N., Gunal, A., & Ritchie, W. (2000). Supply-chain analysis at Volkswagen of America. ¨ Interfaces, 30(4), 46-55.
[2] Kelton, W. David, Randall P. Sadowski, and Nancy B. Zupick. Simulation with Arena. Singapore: McGraw-Hill Education, 2015.
[3] Liu, S., & Papageorgiou, L. G. (2013). Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry. Omega, 41(2), 369-382.
[4] Smith, Jeffrey, David Sturrock, and W. Kelton. Simio and Simulation: Modeling, Analysis, Applications. Sewickley, PA: Simio LLC, 2018.
[5] Tsiakis, P., & Papageorgiou, L. G. (2008). Optimal production allocation and distribution supply chain networks. International Journal of Production Economics, 111(2), 468-483.

APPENDIX A: First Stage Optimization Model

A.I. Indices and Sets

i, j: indices for manufacturing facilities
I: $\{1,...,5\}$: set of manufacturing facilities
k: index for color agents
K: $\{1,...,40\}$: set of color agents
s: index for size types
S: $\{1,...,5\}$: set of size types
f: index for flavor types
F: $\{1,...,12\}$: set of flavor types
l: index for package types
L: $\{1,2\}$: set of package types (bag and box respectively)

A.II. Parameters

c: cost of transportation ($3.5 per mile per 50,000 lbs of jelly bean)

I_{ik}^0: initial inventory level of color k at facility i, $i \in I, k \in K$

RMI_i: single RMI drum capacity at facility i, $i \in I$

n_i: number of RMI drums at facility i, $i \in I$

λ_{isl}: processing rate of packaging machine at facility i for the jelly beans of size s and package type l, $i \in I, s \in S, l \in L$

γ_i: processing rate of classifier at facility i (in pounds per hour), $i \in I$

θ_{isf}: processing rate of pre-finish operation for jelly beans of size s and flavor f at facility i (in pounds per hour), $i \in I, s \in S, f \in F$

ρ_{ks}: percentage split of jelly beans for color k and size s in the classifier process, $k \in K, s \in S$

c_{il}: production cost of one unit of finished pack for package type l at facility i, $i \in I, l \in L$

D_{ksfl}: demand for jelly bean in color k, size s, flavor f, and package type l (in pack units), $k \in K, s \in S, f \in F, l \in L$

ds_{ij}: distance from facility i to facility j (in miles), $i \in I, j \in I$

T_i: truck capacity of facility i (in pounds), $i \in I$

pt: total available production time excluding assumed maximum change-over time

npf_i: number of Pre-Finish machines at facility i, $i \in I$

np_{il}: number of packaging machines of package type l at facility i, $i \in I, l \in L$

d_l: pounds in one unit of package type l, $l \in L$

η_{ksf}: the ratio of demanded jelly beans of color k, size s and flavor f to the all demand, $k \in K, s \in S, f \in F$

A.III. Decision Variables

t_{ijk}: the amount of raw material of color k transported from facility i to facility j (in pounds), $i \in I, j \in I, k \in K$

p_{iksfl}: at the end of production period, the amount of produced jelly beans in color k, size s, flavor f and package type l at facility i (in pounds), $i \in I, k \in K, s \in S, f \in F, l \in L$

I_{ik}: inventory level of raw material (end inventory level) of color k at facility i after raw material transfer, $i \in I, k \in K$

x_{ik}: number of RMI drums used for color k at facility i, $i \in I, k \in K$

e_{iks}: the amount of excess jelly beans of color k, size s produced in facility i, $i \in I, k \in K, s \in S$

A.IV. Mathematical Model

$$\min \sum_{\substack{i\in I, j\in I,\\ k\in K, i\neq j}} c * ds_{ij} * \frac{t_{ijk}}{50000} + \sum_{\substack{i\in I, k\in K, s\in S,\\ f\in F, l\in L}} \frac{c_{il} * p_{iksfl}}{d_l} + \sum_{i\in I, k\in K, s\in S} \frac{c_{il} * e_{iks}}{d_1} \tag{1}$$

$$\text{subject to} \quad I_{ik} = I_{ik}^0 - \sum_{j\in I, i\neq j} t_{ijk} + \sum_{j\in I, i\neq j} t_{jik}, \quad \forall i \in I, k \in K \tag{2}$$

$$I_{ik} \leq x_{ik} * RMI_i, \quad \forall i \in I, k \in K \tag{3}$$

$$\sum_{k\in K} x_{ik} \leq n_i, \quad \forall i \in I \tag{4}$$

$$\sum_{f\in F, l\in L} p_{iksfl} + e_{iks} = \rho_{ks} * I_{ik}, \quad \forall i \in I, k \in K, s \in S \tag{5}$$

$$\sum_{\substack{k\in K, s\in S,\\ f\in F, l\in L}} p_{iksfl} + \sum_{k\in K, s\in S} e_{iks} \leq \gamma_i * pt, \quad \forall i \in I \tag{6}$$

$$\sum_{\substack{k\in K, s\in S,\\ f\in F, l\in L}} \frac{p_{iksfl}}{\theta_{isf} * npf_i} + \sum_{\substack{k\in K, s\in S,\\ f\in F}} \frac{e_{iks} * \eta_{ksf}}{\theta_{isf} * npf_i} \leq pt, \quad \forall i \in I \tag{7}$$

$$\sum_{\substack{k\in K, s\in S,\\ f\in F, l\in L}} \frac{p_{iksfl}}{\lambda_{isl} * np_{il}} + \sum_{k\in K, s\in S} \frac{e_{iks}}{\lambda_{is1} * np_{i1}} \leq pt, \quad \forall i \in I \tag{8}$$

$$\sum_{j\in I, k\in K, i\neq j} t_{ijk} \leq T_i, \quad \forall i \in I \tag{9}$$

$$\sum_{i\in I} p_{iksfl} = D_{ksfl} * d_l, \quad \forall k \in K, s \in S, f \in F, l \in L \tag{10}$$

$$I_{ik} \in \mathbb{R}^+, \quad \forall i \in I, k \in K \tag{11}$$

$$t_{ijk} \in \mathbb{R}^+, \quad \forall i \in I, j \in I, k \in K \tag{12}$$

$$p_{iksfl} \in \mathbb{R}^+, \quad \forall i \in I, k \in K, s \in S, f \in F, l \in L \tag{13}$$

$$e_{iks} \in \mathbb{R}^+, \quad \forall i \in I, k \in K, s \in S \tag{14}$$

$$x_{ik} \in \mathbb{Z}^+, \quad \forall i \in I, k \in K \tag{15}$$

APPENDIX B: Second Stage Optimization Model

Note that, here, only the parameters and variables different than the ones of the first optimization model are defined.

B.I. Parameters

$\hat{I}_{ik}$: the optimal inventory level of color k at facility i after raw material transfer, $i \in I, k \in K$ (optimal end inventory levels of the first optimization model)

w: weight

β: a multiplier to represent what percentage of the capacity is used for all machinery

η_{ksf}: the ratio of demanded jellybeans of color k, size s and flavor f to the all demand, $k \in K, s \in S, f \in F$

$\hat{p}_{iksfl}$: the optimal amount of jellybeans of color k, size s, flavor f, and package type l that are produced at facility i to meet the demand, $i \in I, k \in K, s \in S, f \in F, l \in L$ (optimal production amounts of the first optimization model)

B.II. Decision Variables

p_{iksfl}: at the end of production period, the amount of produced jelly bean in color k, size s, flavor f and package type l, at facility i (in pounds), $i \in I, k \in K, s \in S, f \in F, l \in L$

UD_{ksfl}: amount of unsatisfied demand of jelly beans of color k, size s, flavor f, and packaging type l, $k \in K, s \in S, f \in F, l \in L$

B.III. Mathematical Model

$$\min \quad (1-w) * \sum_{\substack{i \in I, k \in K, s \in S,\\ f \in F, l \in L}} \frac{c_{il} * p_{iksfl}}{d_l} + w * \sum_{\substack{k \in K, s \in S,\\ f \in F, l \in L}} UD_{ksfl} \tag{16}$$

$$\text{subject to} \sum_{f \in F, l \in L} p_{iksfl} \leq \rho_{ks} * I_{ik}, \quad \forall i \in I, k \in K, s \in S \tag{17}$$

$$\sum_{\substack{k \in K, s \in S,\\ f \in F, l \in L}} p_{iksfl} \leq \gamma_i * \beta * pt, \quad \forall i \in I \tag{18}$$

$$\sum_{\substack{k \in K, s \in S,\\ f \in F, l \in L}} \frac{p_{iksfl}}{\beta * \theta_{isf} * npf_i} \leq pt, \quad \forall i \in I \tag{19}$$

$$\sum_{\substack{k \in K, s \in S,\\ f \in F, l \in L}} \frac{p_{iksfl}}{\beta * \lambda_{isl} * np_{il}} \leq pt, \quad \forall i \in I \tag{20}$$

$$\sum_{i \in I} p_{iksfl} = D_{ksfl} * d_l - UD_{ksfl}, \quad \forall k \in K, s \in S, f \in F, l \in L \tag{21}$$

$$p_{iksfl} \leq \hat{p}_{iksfl}, \quad \forall i \in I, k \in K, s \in S, f \in F, l \in L \tag{22}$$

$$p_{iksfl} \in \mathbb{R}^+, \quad \forall i \in I, k \in K, s \in S, f \in F, l \in L \tag{23}$$

$$UD_{ksfl} \in \mathbb{R}^+, \quad \forall k \in K, s \in S, f \in F, l \in L \tag{24}$$

α_{ik}: the ratio of work order satisfied for color k at facility i

α_{ikf}: the ratio of work order satisfied for color k, flavor f at facility i

$$\alpha_{ik} = \sum_{\substack{s \in S, f \in F,\\ l \in L}} \frac{p_{iksfl}}{\hat{p}_{iksfl}}, \quad \forall i \in I, k \in K \tag{25}$$

$$\alpha_{ikf} = \sum_{s \in S, l \in L} \frac{p_{iksfl}}{\hat{p}_{iksfl}}, \quad \forall i \in I, k \in K, f \in F \tag{26}$$

Süreç Takip Modülü

ING Bank A.Ş.

Proje Ekibi

Yağmur Altıntığ, Gözde Yağmur Başar, Atahan Bayır, İbrahim Ethem Gürcan, Seyit Ulutaş, Damla Yalçın, Mehmet Cerir Yülek

Şirket Danışmanı
Zeynep Gümüş
ING Bank A.Ş. İnşaat Emlak Mimari Grubu Proje Müdürü

Akademik Danışman
Prof. Dr.Mehmet Selim Aktürk
Endüstri Mühendisliği Bölümü

ÖZET

Proje, ING Bank A.Ş. için şube tadilat süreçlerini takip edip bunu yöneten ve gecikmeleri önleyip fazladan kira ödenmesinin önüne geçen bir web tabanlı karar destek sistemi geliştirmeyi amaçlamaktadır. Süreç adımları analiz edilmiş ve yapılan veri analizi ile standart bir süreç planı oluşturulmuştur. Bu plana kritik yol yöntemi uygulayan sistem kullanıcıları düzenli olarak bilgilendirmektedir. Bu proje ile şube tadilat süreçlerinde ortalama altı günlük bir iyileştirme sağlanmıştır.

Anahtar Kelimeler: proje yönetimi, kritik yol yöntemi, takip modülü, süreç analizi, django

Project Tracking Module

1. Identification of Problem

The department that has been the focal point of our project is the construction and real estate department. They are responsible for keeping track and executing, renewal and construction projects of the bank's branches all across Turkey. Because the operations are based on construction the executions depend on meeting due dates for each step as it is the prerequisite of another step. With the data given from the bank, network schedules were analyzed and prerequisites were determined. After this, using past data from the bank processes that caused delays were found, these were processes that were being delayed because of "human factors". Due to the hardness of following every operation of each branch ING was having a hard time meeting due dates. This was in fact their observation as well that they had given at the beginning of the project. After confirming this with delay data it was clear that the main problem was not being able to follow the current situations of projects, due to the lack of a computerized system. The deliverable of our project is a process tracking module helping the department track the current states of ongoing projects, tracking due dates helping them meet requirements on time which would enable the constructions to end on time, the performance measure here is time gained by the company for each project.

2. Proposed System and Details of The Approach

The model has been designed to provide information about the projects and the activities of the braches to the company, as well as the warning system which is activated in the case of delays. In general, ING Bank conducts each of their operations manually without a system and it is resulted in the insufficient examination of the possible causes of delays and being less aware of the time component and relatively excessive renting cost. Therefore, it has been decided to create a platform, which will enable them to utilize in their process tracking and operation examinations.

According to the needs and requirements of the construction department, possible and suitable programs to fulfill the needs of the company have been researched and online platform has been chosen as the basis of the system. This platform has been conducted from the scratch only for ING Bank users and it provides access from different computers via their personal systems. Coding language was the second decision, which should have had to provide additional tools to develop new structures and to create the platform. After an examination, Python programming language has been decided as the basis of the system because of its wide usage and possible frameworks. As the next step, web framework has been chosen to be utilized in order to increase the user

attainability and accessibility. By doing that, users would be able to reach the platform from their computers by using the manual prepared for them.
Tools of Python script have been highly beneficial for the project, because of the large options it presents. As the further step, the type of web framework has needed to be selected. After some research, Django web framework of Python has been chosen to create the online platform, since it enables users to create Web applications for wide access. By using Django, login page, user creation, project and activity pages, and all the projections that are provided for the usage of construction department are conducted and coded.
Online platform enables users to see all the projects under the field of 'All Projects'. This part consists of 'Ongoing Projects', 'Uninitiated Projects' and 'Completed Projects' with respect to the situations of the activities. Additionally, users can obtain more information about each project by clicking on the project name, which consists of specific activities and their estimated start-finish times and their situations such as 'Completed', 'Ongoing', or 'Delayed'.
The users can also mark the situations of the projects. They can change their situations from 'Uninitiated' to 'Ongoing' or 'Ongoing' to 'Completed'. They can also start and finish each activity of the project and make the system to store their actual dates for further utilization. With the help of this system, users can have the historical data which keeps records about the projects as well as their actual start and finish times. Figures of the proposed system can be found in Appendix A.

System also presents the 'expected start' and 'expected finish' times of the projects and their activities, according to the durations and the precedence's of the activities that have been specified during the standardization process of the project. This calculation is operated by the Critical Path Method (CPM), which runs behind the system. With the help of this method, users can observe whether the activities are conducted in a designated time frame and each of the steps are completed according to the schedule. If there is any delay or the prescription, it gives warnings about the delays and makes it possible to observe the problematic steps or the reasons behind them by providing the data to compare with.

3. Validation of Proposed Approach

Main contribution of proposed Project Tracking Module is to determine possible delays in advance and simplify the project tracking process for employees with a computerized system. In order to measure the accuracy of this statement, data given from the company, which are schedules of previous projects, and the results provided from proposed system are compared. During the comparison, finish date of each project are compared with dates provided by the system. To make the comparison more accurate, actual durations of each activity is used while calculating finish dates by using proposed system. According to results, proposed system saves approximately 6 days per project. In addition to these benefits, proposed system can also contribute further improvements in system. Since delays occurred in each project are stored in system's database, they can be used to identify the main problems during construction processes. Results are available in Appendix B.

4. Integration

The system runs through a server and server uses distinctive computer languages such as Python, HTML, CSS, JavaScript, etc. These languages and Django package ought to be installed on the machine which runs the system. Installation of Django is crucial since the whole web framework is created by using it. It is completely different compared to the existing system which is based on mailing among colleagues. In order ING Bank to utilize it, whole project should be integrated into the ING Bank's network so that users can access from different computers. Moreover, since our system is based on a web framework, all data transformations will be conducted through a website which needs an Internet connection.

In order to inform ING about above points, a detailed user manual is prepared. In addition to detailed description of proposed system, installation of Python and Django is also explained in this manual. Additionally, information regarding user account that has been created by project group is also shared with ING. By following those provided information, ING Bank will be able to register and enter the proposed system.

5. Possible Contribution to the Company

Our decision support system will contribute to ING Bank by converting their manual system to an automated system. The system is intended to provide project management systems with a decision support system and to show all of the projects in one area to help keep track of the projects and manage them more efficiently. Their previous system has required manual Excel entries and communications through emails. Our decision support system will create a decrease in their mail traffic and their workload. In the event of a delay, the system will send a notification to inform the relevant employers. It will be easier to control the processes of the projects in detail and they can reach any information directly through the system with their username and password. This system enables users to be able to view all of the projects in one page and track their current status. System is able to detect the root cause of the problems and the stages that are involved. This module will provide a framework that will store all of the data of the projects' stages. This will provide benefit for the elimination of these problems and also to forecast and implement these particular work stages. Different departments, workers, subcontractors and beneficiaries could be incorporated into the program; therefore, the coherence between participants will increase. Finally, using a decision support system will reduce the additional cost of rent payments, and employee's will no longer invest their time into the manual management of the project process.

5. Conclusion

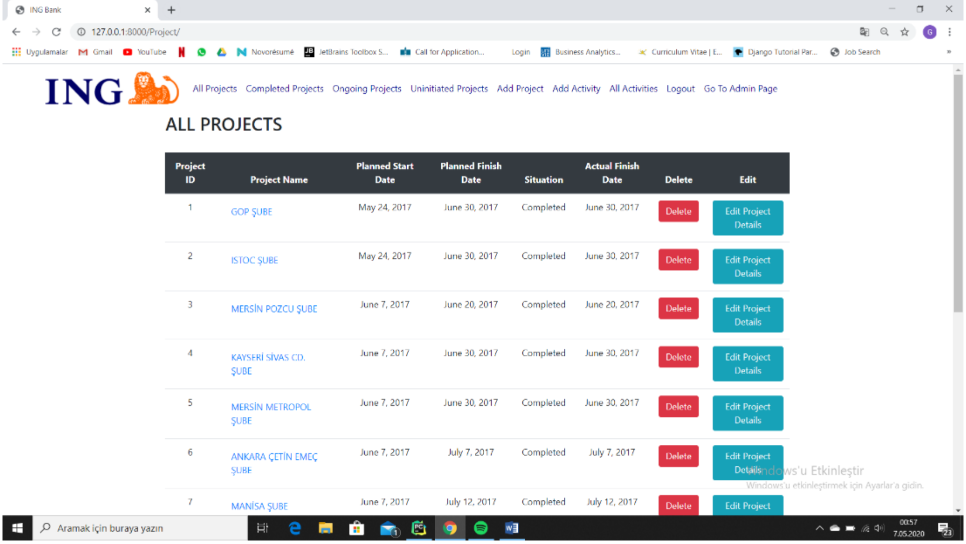
The decision-making support system and tracking module meets the ING Bank's construction department's needs with industrial engineering approach and it digitalizes it. Construction team was having difficulties with tracking the steps of the construction and because of having more than one project to track, delays occurred in the projects. The system eases the tracking process by using workflow diagram, which determined by the users and critical path method to follow the steps and due dates. The other important issue has been resolved by the system's increased awareness module, which warns each user respect to their department (electric, mechanic or architect) job stages. The system shortened each project's due date approximately 6 days in application. Module has been created compatible with servers. So, the system easily implemented into ING's servers and could easily reach by any users from regardless to location just by an internet connection. Furthermore, the frontend of the system is designed considering the users' needs, such as easy to use with user friendly interface. System creates a database for company and each step is saved according to their department and project. The data provide with a chance to a better data analysis and interpretation. It can also be used for performance measuring, possible unnecessary delay detection, resource allocation and possible loss detection from delays. System could use it to measure performance for the departments and subcontractors, when delays occurred. Furthermore, the module will be used for employ new participants, outsourcing some steps or reallocating resources when new projects added into the system or present projects finished with respect to collected data.

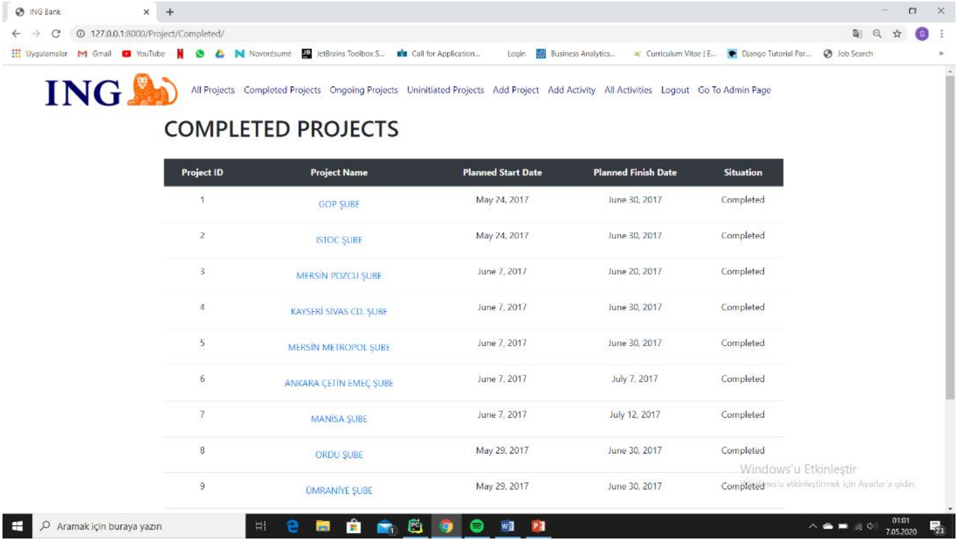
References

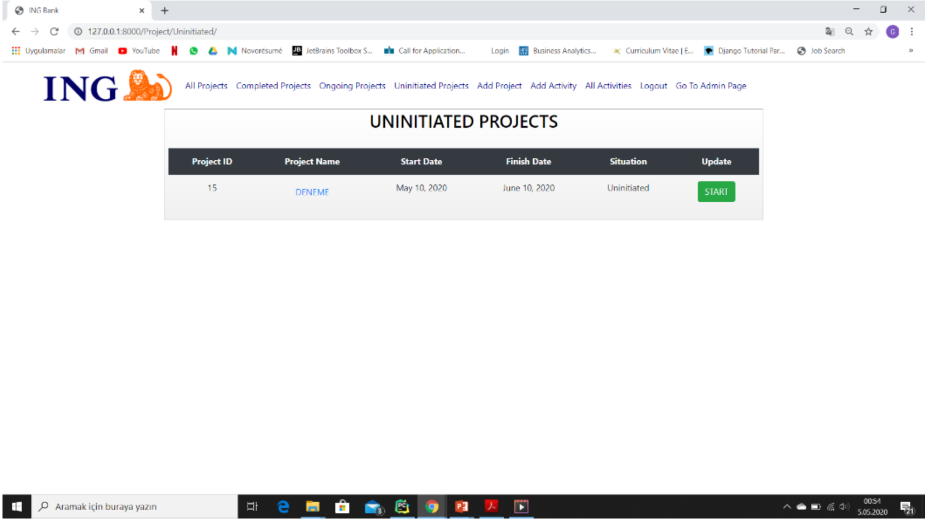
Forcier J, Bissex P. and Chun W.J. (2009) 'Python Web Development with Django'. Pearson Education, Inc. 380 pg. Developer's Library.

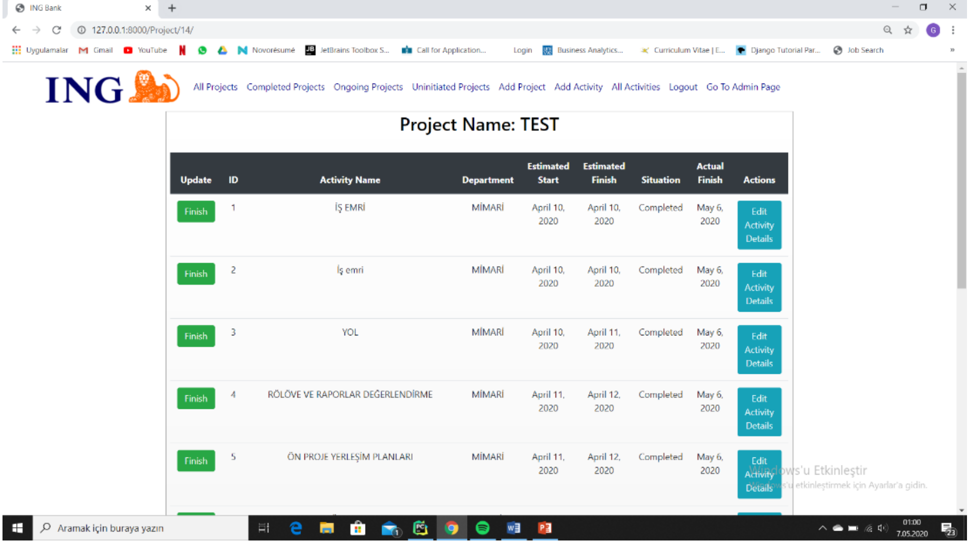
Appendix

Appendix A Proposed System






ING Bank
127.0.0.1:8000/Project/Uninitiated/
ING
All Projects Completed Projects Ongoing Projects Uninitiated Projects Add Project Add Activity All Activities Logout Go To Admin Page
UNINITIATED PROJECTS
Project ID Project Name Start Date Finish Date Situation Update
15 DENEME May 10, 2020 June 10, 2020 Uninitiated START
Aramak için buraya yazın

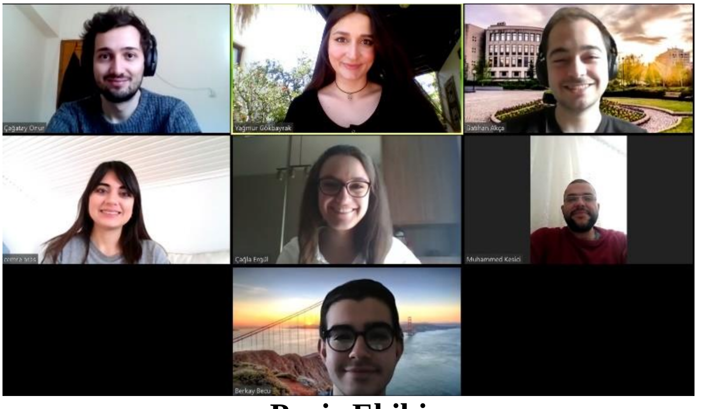

ING Bank
127.0.0.1:8000/Project/14/
ING
All Projects Completed Projects Ongoing Projects Uninitiated Projects Add Project Add Activity All Activities Logout Go To Admin Page
Project Name: TEST
Update ID Activity Name Department Estimated Start Estimated Finish Situation Actual Finish Actions
Finish 1 İŞ EMRİ MİMARİ April 10, 2020 April 10, 2020 Completed May 6, 2020 Edit Activity Details
Finish 2 İş emri MİMARİ April 10, 2020 April 10, 2020 Completed May 6, 2020 Edit Activity Details
Finish 3 YOL MİMARİ April 10, 2020 April 11, 2020 Completed May 6, 2020 Edit Activity Details
Finish 4 RÖLÖVE VE RAPORLAR DEĞERLENDİRME MİMARİ April 11, 2020 April 12, 2020 Completed May 6, 2020 Edit Activity Details
Finish 5 ÖN PROJE YERLEŞİM PLANLARI MİMARİ April 11, 2020 April 12, 2020 Completed May 6, 2020 Edit Activity Details
Aramak için buraya yazın

Appendix B Validation

Project Name	Actual Start Date	Actual Finish Date	Total Duration (days)	Provided Finish Date	Days Saved
GOP Şube	May 24, 2017	June 30, 2017	38	June 11, 2017	19
ISTOC Şube	May 24, 2017	June 30, 2017	38	June 18, 2017	12
Ordu Şube	May 29, 2017	June 30, 2017	33	June 23, 2017	7
Ümraniye Şube	May 29, 2017	June 30, 2017	33	June 23, 2017	7
Kayseri Sivas Cd. Şube	June 7, 2017	June 23, 2017	17	June 20, 2017	3
Mersin Metropol Şube	June 7, 2017	June 30, 2017	24	June 28, 2017	2
Ankara Çetin Emeç Şube	June 7, 2017	July 7, 2017	31	July 2, 2017	5
Manisa Şube	June 7, 2017	July 12, 2017	36	July 2, 2017	10
Nişantaşı Şube	June 8, 2017	July 10, 2017	33	July 3, 2017	7
Siirt Şube	July 24, 2019	September 9, 2019	48	August 18, 2019	9
Söke Şube	October 17, 2019	November 8, 2019	23	November 5, 2019	3
Sultanbeyli Şube	November 8, 2019	November 5, 2019	28	November 2, 2019	3

Servis Seviyesi Odaklı Stok Optimizasyonu

Unilever Türkiye



Proje Ekibi
Batıhan Akça, Cemre Aras, Berkay Becu, Çağla Ergül, Zeynep Yağmur Gökbayrak, Muhammed Kesici, Çağatay Onur

Şirket Danışmanı
Kübra Gürsoy Aslan
Talep Planlamacı

Akademik Danışman
Doç. Dr. Alper Şen
Endüstri Mühendisliği Bölümü

ÖZET

Projenin amacı, Unilever Food Solutions envanter yönetimini iyileştirecek bir karar destek sistemi geliştirmektir. Karar destek sisteminin amacı; toplam stok tutma maliyetini enazlarken hedef servis seviyesine ulaşmaktır. Sistemde UFS'in aylık talep tahminleri kullanarak bu tahminleri haftalara dağıtan bir talep tahmini dağıtım modeli geliştirilmiştir. Haftalık talep tahminleri ve tahmin hataları kullanılarak, ürün bazında ve toplam hizmet seviye kısıtlarını sağlayan ve envanter taşıma maliyetini enazlayan bir stok optimizasyon problem tanımlanmış ve bu problemin çözümü için bir tam sayılı doğrusal matematiksel programlama modeli geliştirilmiştir. Karar destek sistemi yardımıyla talep tahmininde %2.34'lük, sistem servis seviyesinde ortalama %12'lik bir iyileşme gözlenirken; ortalama stok seviyesinde %7'lik azalma görülmüştür.

Anahtar Kelimeler: Envanter yönetimi, sipariş karşılama performansı optimizasyonu, servis seviyesi

End-to-End Stock Optimization

1. General Information About the Company

Unilever was established in 1930 by the merger of Dutch margarine company Margarine Unie and British soap company Lever Brothers. Unilever supplies products in 190 countries with more than 400 brands. Today the company operates in Turkey with 8 factories, 5 warehouses, and 5000 employees. Unilever manufactures products in three main categories which are home care, food & refreshment, and beauty & personal care. The scope of the project is on Unilever Food Solutions (UFS) which operates in desserts, dressings, tea and savoury product categories.

2. Current System Description

Supply chain management in UFS starts with the customer service department, which aims to maximize customer satisfaction through understanding their demands. The sales department determines the sales targets based on the feedback that the customer service department gets from the customers. Based on these, demand planning department prepares forecasts and works on maximizing forecast accuracy. The demand planning department delivers the finalized forecasts to the supply planning department and negotiates on production quantities. Then, the supply planning department prepares the production plan, informs the sales department on final production quantities and initiates the procurement and the manufacturing processes.

The raw material planners and the procurement department manage the inventory in accordance with production plan and works towards minimizing excess inventory. UFS has two domestic production facilities at Çorlu and Çayırova. Each production facility ships its final goods to the associated warehouses daily. Final goods at the warehouses are eventually delivered to customers. The supply chain ends in the logistics department as the finished goods leave the warehouse to be delivered to the customer. Also, customer feedback regarding the products in the market are received. Based on this feedback, targets of the sales department are updated, and necessary adjustments are made in the production system.

3. Problem Definition

According to the company, the current forecasting model provides accurate results at category and subcategory levels, but accuracy at SKU level is not satisfactory. According to their metric, Weighted Average Method 1 (see Appendix B), weekly forecast error is calculated as 128.04% and monthly forecast error is found to be 90.8%. It is concluded that monthly forecasts are more accurate than weekly forecasts where disaggregating monthly forecasts into weekly forecasts increases inaccuracy. Thus, one problem is to reduce inaccuracy in forecast disaggregation.

As a result of this inaccuracy in demand forecasting, the company encounters with stock-out and excessive inventory problems in their finished good stocks. In case of stock-outs, UFS is forced to deliver materials with delays. Since customers are service platforms such as wholesalers and restaurants, backorders and service delays may result in the permanent loss of the customer. Thus, the fundamental problem is to achieve target service levels. However, increasing service level requires carrying higher levels of stocks. This would increase warehousing costs and ties up working capital. Considering Supply Chain Triangle embraced by Unilever, the main objective is to satisfy service level targets while balancing inventory holding costs.

4. Proposed Models

4.1. Forecast Disaggregation Model

As mentioned before, weekly forecast accuracy is not satisfactory for stock optimization model to perform effectively. Since monthly forecasts are more accurate, a supplementary model is constructed to disaggregate the monthly forecasts into weekly forecasts. Taking the historical data of monthly and weekly observed demand as input, the model aims to disaggregate the monthly forecast into weeks using the least squares approach. The model is as follows:

$$N: Set\ of\ weeks\ in\ a\ month$$
$$n: Set\ of\ instances$$
$$i: Order\ of\ the\ week\ in\ a\ particular\ month, i \in N$$
$$j: Number\ of\ the\ instance\ in\ the\ historical\ data, j \in n$$
$$W_{ij}: Weekly\ sales\ of\ i\ in\ j$$
$$MS_j: Monthly\ sales\ of\ j$$
$$b_i: Disaggregation\ coefficient\ of\ i$$
$$\min_{b_i} \sum_{j\in n} \sum_{i\in N} (W_{ij} - b_i \times MS_j)^2$$
$$subject\ to$$
$$\sum_{i\in N} b_i = 1$$
$$b_i \geq 0\ , i \in N$$

The problem can be solved analytically to obtain the coefficient in closed form as follows:

$$b_i^* = \frac{\sum_{j=1}^{n} MS_j W_{ij}}{\sum_{j=1}^{n} MS_j^2}\ \ \forall i$$

These coefficients can be calculated separately for different product groups. “Dressings” category shows a different demand distribution in split weeks than the other categories. Hence, the coefficients for "Dressings" category are calculated separately than the other categories. The length of the split week was considered to be the second criterion that affects the distribution. However, this criterion was later decided to be excluded since it only explains 14% of the variance. Lastly, the number of split weeks in a month is considered. A month can have one split week in the beginning or in the end, two split weeks both in the beginning and the end or it does not have any split week. It is observed that there is fewer demand in split weeks (between 9% and 15% of monthly demand on average of January 2018- September 2019 per SKU) so the number of split weeks also affects the distribution. For that reason, the calculation of coefficients are done separately for the months that have different number of split weeks.

4.2 Stock Optimization Model

The order placement flow of UFS is as the following:

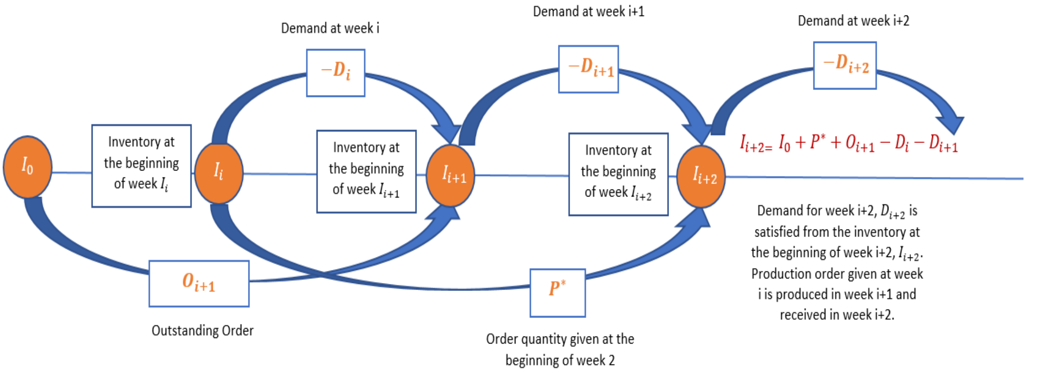


Figure 1. **Order Placement Flow**

The stock optimization model aims to minimize the total expected inventory cost while satisfying the overall system service level target as well as SKU-specific service level targets (if any) for each SKU. The decision variables are the production quantity of each SKU to be produced next week. It is important to note that the production order given in week *i* will be produced during week *i+1*, and this will be available (in stocks) in week *i+2*.

For product *j*, the inventory level at the beginning of week *i* is denoted by I^j_i. P^j is the production quantity for product *j*, the decision variable of the model (to be determined in week *i)* and O^j_{i+1} is the outstanding order to be received at the beginning of week *i*. This is the production order that was placed in the week *i-1*. I^j_{i+2} is the inventory at the beginning of week 2. Taking this flow into consideration, the stock optimization model is as follows:

Model Sets and Indices:

J: $Set\ of\ SKUs$

$j \in J$

Model Parameters:
S^j: *Lower limit for the service level of SKU j*
h^j: *Holding cost (per kg) of SKU j*
SL: *Lower limit for the overall (system) service level*
O^j: *Outstanding order of SKU j*
I^j: *Beginning inventory level of SKU j*
Decision Variables:
P^j: *Production order of SKU j*
Model:

$$\min_{P^j} \sum_{j \in J} h^j \times E\left[max\left(I_i^j + O_{i+1}^j + P^j - D_i^j - D_{i+1}^j - D_{i+2}^j, 0\right)\right]$$

$$subject\ to$$

$$\frac{\sum_j E\left[\min(I_i^j + O_{i+1}^j + P^j - D_i^j - D_{i+1}^j, D_{i+2}^j\ )\right]}{\sum_j E\left[D_{i+2}^j\right]} \geq SL$$

$$\frac{E\left[\min(I_i^j + O_{i+1}^j + P^j - D_i^j - D_{i+1}^j, D_{i+2}^j\ )\right]}{E\left[D_{i+2}^j\right]} \geq S^j \ , \quad \forall\, j \in J$$

$$P^j \geq 0 \ , \quad \forall\, j \in J$$

In order to have a consistent MOQ constraint among all SKUs, the "free" lot size is discretized with respect to the MOQ. More importantly, the non-linear model can be converted into a mixed integer model by the discretization. The discretization also makes the results of the model more applicable as production orders are given in a similar manner. Currently, the lot size is discretized by 50% of the MOQ upon the request of the company. After taking this constraint and discretization process into consideration, the production order of an SKU can be defined by the following:

$$Pj = \begin{cases} 0 \\ MOQj + a \times Lot\ Size^j \end{cases}$$

where the production order can have zero quantity, MOQ value or MOQ plus some multiple of the lot size. Finally, an integer decision variable k^j is defined which implies the possible production order quantities of SKU j with respect to the MOQ constraints:

$$k^j = \begin{cases} 1, & P^j = 0 \\ 2, & P^j = MOQ^j \\ a: a > 2, & P^j = MOQ + (a-2) \times Lot\ Size^j \end{cases}$$

By the normality assumption of the demand, 3-sigma approach is used in the sense that an upper bound is put so that the production order quantity satisfies the three-week demand (along with initial inventory and outstanding order) by 99.9% probability. By using this approach, the upper bound of the production quantity for SKU j can be defined by the following:

$$\grave{P}^j = \max(\sum_{i}^{i+2} F_i^j \times 3\sigma_{\cdot}^j - I_i^j - O_{i+1}^j, 0)$$

This definition can be alternatively, applying the MOQ constraints and corresponding integer decision variable, written as:

$$\grave{k}^j = \left\lceil \frac{max(\sum_{i}^{i+2} F_i^j \times 3\sigma^j - I_i^j - O_{i+1}^j - MOQ^j, 0)}{Lot\ Size^j} \right\rceil + 2$$

For a particular value of k^j, $k^j = a$ (which corresponds to $P^j = \overline{P^j{}_a}$), two parameters are defined:

$$Y_a^j = \frac{E\left[\min(I_i^j + O_{i+1}^j + \overline{P^j{}_a} - D_i^j - D_{i+1}^j, D_{i+2}^j)\right]}{E\left[D_{i+2}^j\right]}$$

$$Z_a^j = E\left[max\left(I_i^j + O_{i+1}^j + \overline{P^j{}_a} - D_i^j - D_{i+1}^j - D_{i+2}^j, 0\right)\right]$$

A new binary decision variable is defined:

$$x_a^j = \begin{cases} 1, & if\ \ k^j = a\ \left(or\ P^j = \overline{P^j{}_a}\right) \\ 0, & otherwise \end{cases}$$

The proposed model is converted into a binary programming model:

$$\min_{x_a^j} \sum_{j \in J} \sum_{a=1}^{\grave{k}^j} h^j \times Z_a^j \times x_a^j$$

$$subject\ to$$

$$\sum_{j \in J} \sum_{a=1}^{\grave{k}^j} Y_a^j \times x_a^j \geq SL$$

$$\sum_{a=1}^{\grave{k}^j} Y_a^j \times x_a^j \geq S^j\ , \quad \forall\, j \in J$$

$$\sum_{a=1}^{\grave{k}^j} x_a^j = 1\ \ \forall\, j \in J$$

$$x_a^j \in \{0,1\}, \qquad \forall\, j \in J, a = 1, \dots, \grave{k}^j$$

5. Validation of Models

5.1. Forecast Disaggregation Model

For validation of the model, we took monthly and weekly demands of 149 SKUs between February 2018 and September 2019. After computing the coefficients of the forecast disaggregation model, we disaggregated the monthly forecasts into weekly forecasts. To calculate forecast errors and compare the results with UFS' forecasts, different metrics are used (see Appendix A for the metric formulas). The weights are according to demand for each week. To average forecast error among the SKUs, we used two different methods. First method computes the arithmetic average of forecast errors of each SKU. Second method assigns weights according to total demand of each SKU. UFS uses weighted MAPE (Method 1). However, weighted MAPE (Method 2) is a better alternative since it reflects the significance of demand quantity in calculating average forecast error. According to Table 1, the forecast disaggregation model reduces the forecast errors which indicates better results in weekly forecasting.

Table 1. Unilever's Weekly Forecast Errors vs Project's Weekly Forecast Errors

	MAPE (Method 1)	MAPE (Method 2)	Weighted MAPE (Method 1)	Weighted MAPE (Method 2)	SMAPE (Method 1)	SMAPE (Method 2)	Weighted SMAPE (Method 1)	Weighted SMAPE (Method 2)
UFS	1412.50%	3091.59%	128.04%	118.95%	67.42%	68.91%	60.99%	58.23%
Project	1256.83%	2480.64%	124.60%	100.00%	64.13%	62.10%	59.11%	55.37%

5.2. Stock Optimization Model

The validation of this model involves achieving the overall service level target. Thus, we simulated the system with 149 SKUs by using realized weekly demand and monthly forecast data. Model parameters (coefficients of forecast disaggregation model and forecast error distribution parameters) are computed and the system is simulated with 2018 data by targeting 90% overall service level. 70% individual service level targets are assigned as our industrial advisor suggested. Since the company only provided the stock levels for the weeks at the end of each month, we could only compare the stock levels and realized overall service level at the end of each week. The results (see Appendix B) indicate that model was able to achieve 90% overall service level target for every week with an average of 94% while UFS achieved this target only at the ending week of November 2018 with an average of 72%. Comparing the average stock levels (per SKU), model gave higher stock levels except for the ending week of July with a weekly average of 6861.51 kg while the average for UFS is 5521.57 kg. Although the model gave 24% higher inventory, it was able to increase the overall service level by 32%. Considering the tradeoff between service and stock level, this result is expected. However, we need to show that the model gives lower stock levels for the same overall service level target to complete validation. Since overall realized service level of UFS is 72% on average, we run the simulation with the same parameters, only changing the overall service level target to 72%.The results (see Appendix C) indicates that that the model was able to achieve overall service level target (72%) for every week except the ending week of March (68%) and August (69%) with an average of 80%. However, there is also a sharp decrease in UFS service levels for the ending weeks of March and August, compared to previous ending weeks, which is due to underestimating monthly forecasts. As our forecast disaggregation model depends on monthly forecasts of UFS, the model failed to achieve service level targets in those weeks (with a relatively small margin). On the other hand, UFS achieved overall service level target only 6 ending weeks out of 11 with a 72% average. Comparing the average stock levels (per SKU), the model gave lower stock levels except for the ending week of June, August, September and December, with a weekly average of 5147.13 kg. However, in those weeks, there is a significant, more than 20%, difference in overall realized service level where UFS failed to achieve 72% target. UFS carried 5521.57 kg stock on average which is 7% higher than the model result.

In this simulation, the model was able to satisfy 72% service level target and dominate UFS average by 12%. Secondly, the model was able to decrease the average stock levels (per week, SKU) by 7% compared to UFS average. Since the model gave higher service level and lower stock level, it can be concluded that it provides valid results.

6. Implementation of the System under Different Scenarios

6.1. Different Service Level Targets

Currently, UFS determined the system service level requirement as 90%. However, they want to increase this target in the following years. The system was simulated with respect to 90%, 95% and 99% overall service level requirements. The model was able to achieve these targets in actual service levels. The model was also run for different individual service level targets, 50% and 70%. The model was able to satisfy these levels in actual service levels.

6.2. Non-Stationary Weekly Demands

We tested the stationarity assumption of weekly forecast errors by following two methods to compute them. Firstly, we fit a (normal) distribution to monthly forecast error of each SKU. Secondly, under the hypothesis that forecast errors depend on the demand quantities (thus, being non-stationary); we disaggregate the parameters of monthly forecast error distribution into weekly forecast error distribution by using the coefficients of forecast disaggregation model. According to the results, the model was able to satisfy the service level requirements under the assumption of stationarity. The other method gave ineffective results in the sense that the model failed to satisfy service level requirements.

6.3. Biased and Non-normal Weekly Forecast Error Distribution

After fitting a normal distribution to forecast errors, we conduct a hypothesis testing for μ (under the null hypothesis that mean parameter is 0) of each fitted distribution as well as chi-square goodness-of-fit test to see whether forecast error are coming from 0-mean normal distribution. The results show that forecast errors of 27 SKUs do not come from 0-mean normal distribution. However, non-normal error distribution did not impact the stock levels negatively since the model was able to satisfy individual service levels and reduce the stock levels in the simulation for 72% overall service level target.

6.4. Forecasting with Information from Customers

There are some cases in which demand planners get advance information on the actual customer orders and forecast the demand accurately for SKUs with a limited customer target. However, they were not able to provide these cases in the demand/forecast data they shared with us so we could not reflect this information to our model parameters. This is one of the reasons that causes higher error in our forecast model and higher variances in forecast error distribution. However, the interface allows entering forecast manually so model parameters would reflect the system more accurately as they use the interface.

6.5. Increasing the Problem Size

The model can run with 149 SKUs. However, the current problem size is bigger with 270 listed SKUs in 2019. The problem size can also increase because of lot size discretization and upper bounds of production quantities. We also tested for different problem sizes to observe if free solver, GNU Linear Programming Kit, can solve the binary programming model. We were able to run the model for 270 SKUs with 10% of lot size discretization, 50% lower and 99.9% upper bound for individual service level.

7. Interface

A user-friendly interface is constructed using the Shiny package of R (See Appendix E). The company will receive the outcomes of the models through this interface as Excel files. The coding in R includes 3 stages; design of a database, integration of stock optimization and forecast disaggregation model to the interface and additional features for viewing and updating the database and model parameters. There are many parameters and inputs of the models which would change since the set of SKUs and their demand, as well as the service level targets would change in time. Updating these parameters and providing the inputs to the models require the data to be stored in a structured way. Thus, a database is constructed in SQLite since it can be integrated to R (See Appendix D). The interface integrates proposed models into R and provides additional features including exporting and updating the outcomes of the model, updating sigma values and coefficients of forecast disaggregation model, adding/deleting SKU, deleting forecast and sales data in a given time period, entering new sales and forecast data, updating sales and forecast data, reviewing data regarding an SKU and supplementary features, updating holding cost and changing the sensitivity of the lot size.

8. Contributions to the Company

The expected benefits include an increase in the service levels and prevention of customer goodwill (and a decrease in market share as customers eventually defect to competition) in case of stock outs. In addition, the average stock levels are expected to be reduced while satisfying the service level requirements. By keeping lower stocks, the inventory holding costs will also be reduced. The inventory replenishment would be more frequent which indicates

effective inventory management. Avoiding excess inventory which ties up working capital would allow UFS to keep their assets in more liquid form.
Taking the simulation run for 149 SKU, 2018 year at 72% service level target, the average system service level of UFS in 2018, as benchmark:

- The project was able to satisfy 72% service level target in 9 out of 11 weeks (missing the target less than 1% in the other two weeks) while achieving 80% service level on average. UFS was able to satisfy the target in 6 out of 11 weeks while achieving a 72% average service level. The project was able to increase the average system service level by 12%.
- The project was able to decrease average stock by 7%.
- The average inventory holding cost per SKU is reduced by 9.4%.

9. Implementation

The entire solution and the interface were explained to our industrial advisors and they expressed that they had been satisfied with the outcome of the project. In addition, some "Help" buttons are added to the relevant modules of the interface to guide the company officials. This way, they can consult to these buttons to clarify any confusions. However due to the active pandemic, the complete implementation of the production decision suggestions could not be performed with the company. UFS's customer portfolio generally consists restaurants, hotels, cafes and with the closure of these places, production order of Unilever Food Solution has decreased considerably in recent months. In future, with normal production order levels, the implementation of proposed solution would be possible and more effective.

APPENDICES

APPENDIX A. Metrics for Forecast Error Accuracy Calculations

$$n: Total\ Week\ Numbers\ (February\ 2018 - September\ 2019)$$
$$N: Number\ of\ SKUs\ (149\ SKUs)$$
$$A_{ij}: Actual\ sales\ of\ SKU\ i\ in\ week\ j$$
$$F_{ij}: Forecast\ of\ SKU\ i\ in\ week\ j$$

1) MAPE (Method 1)

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\left|A_{ij} - F_{ij}\right|}{A_{ij}}\right)$$

2) MAPE (Method 2)

$$\sum_{i=1}^{N}\frac{Total\ sales\ of\ SKU\ i}{Total\ sales}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\left|A_{ij} - F_{ij}\right|}{A_{ij}}\right)$$

3) Weighted MAPE (Method 1)

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{n}\left|A_{ij} - F_{ij}\right|}{\sum_{j=1}^{n} A_{ij}}\right)$$

4) Weighted MAPE (Method 2)

$$\sum_{i=1}^{N}\frac{Total\ sales\ of\ SKU\ i}{Total\ sales}\left(\frac{\sum_{j=1}^{n}\left|A_{ij} - F_{ij}\right|}{\sum_{j=1}^{n} A_{ij}}\right)$$

5) SMAPE (Method 1)

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\left|A_{ij} - F_{ij}\right|}{(A_{ij} + F_{ij})}\right)$$

6) SMAPE (Method 2)

$$\sum_{i=1}^{N}\frac{Total\ sales\ of\ SKU\ i}{Total\ sales}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{\left|A_{ij} - F_{ij}\right|}{(A_{ij} + F_{ij})}\right)$$

7) Weighted SMAPE (Method 1)

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{n}\left|A_{ij} - F_{ij}\right|}{\sum_{j=1}^{n}(A_{ij} + F_{ij})}\right)$$

8) Weighted SMAPE (Method 2)

$$\sum_{i=1}^{N}\frac{Total\ sales\ of\ SKU\ i}{Total\ sales}\left(\frac{\sum_{j=1}^{n}\left|A_{ij} - F_{ij}\right|}{\sum_{j=1}^{n}(A_{ij} + F_{ij})}\right)$$

APPENDIX B. Results of the Validation for 90% Overall System Service Level Target

	Average Stock Level Per SKU (Project)	Average Stock Level Per SKU (UFS)	Overall Realized Service Level (Project)	Overall Realized Service Level (UFS)
February 2018	7216.27	5669.38	98%	81%
March 2018	6404.05	5381.01	90%	61%
April 2018	6403.65	5827.08	93%	85%
May 2018	6049.79	5858.18	90%	75%
June 2018	7624.90	5073.45	91%	50%
July 2018	6211.10	6446.61	92%	86%
August 2018	6533.17	4118.60	93%	44%
September 2018	8269.54	5230.06	98%	51%
October 2018	7597.10	5635.34	100%	87%
November 2018	7018.87	6218.92	95%	93%
December 2018	6148.12	5278.61	93%	71%
Weekly Average	6861.51	5521.57	94%	72%

Table 2. Simulation Results at 90% Overall Service Level Target

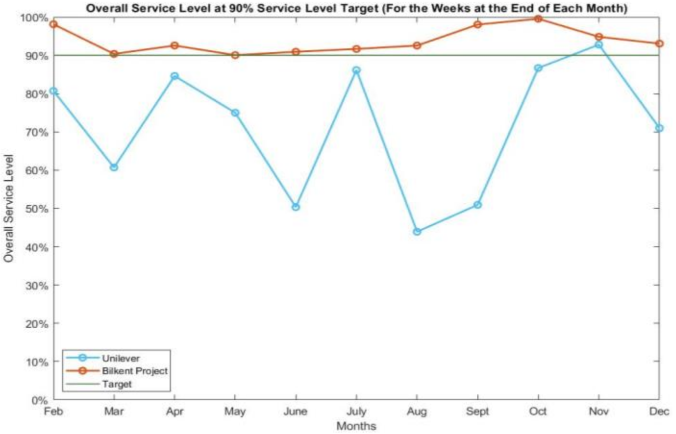


Figure 2. Overall Realized Service Levels at 90% Overall Service Level Target

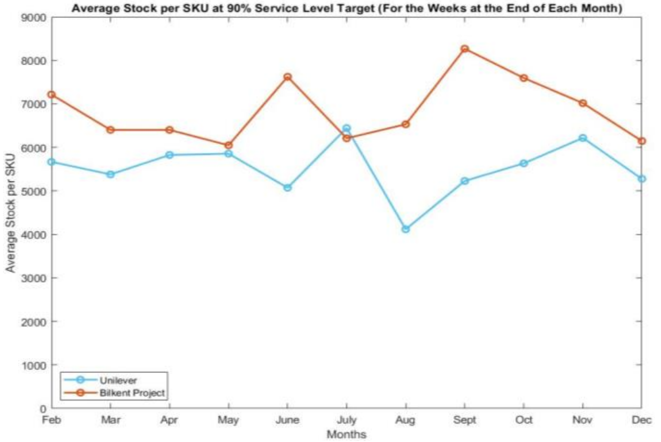


Figure 3. Average Stock per SKU at 90% Overall Service Level Target

APPENDIX C. Results of the Validation for 72% Overall System Service Level Target

	Average Stock Level Per SKU (Project)	Average Stock Level Per SKU (UFS)	Overall Realized Service Level (Project)	Overall Realized Service Level (UFS)
February 2018	4185.76	5669.38	79%	81%
March 2018	3681.67	5381.01	68%	61%
April 2018	5095.00	5827.08	79%	85%
May 2018	4767.30	5858.18	79%	75%
June 2018	6201.80	5073.45	76%	50%
July 2018	6146.48	6446.61	88%	86%
August 2018	5409.80	4118.60	69%	44%
September 2018	5689.56	5230.06	75%	51%
October 2018	5263.05	5635.34	87%	87%
November 2018	4279.73	6218.92	90%	93%
December 2018	5898.27	5278.61	90%	71%
Weekly Average	5147.13	5521.57	80%	72%

Table 3. Simulation Results at 72% Overall Service Level Target

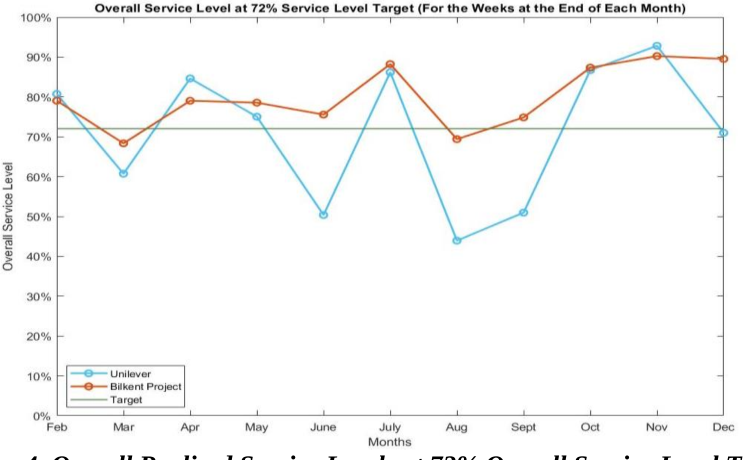


Figure 4. Overall Realized Service Levels at 72% Overall Service Level Target

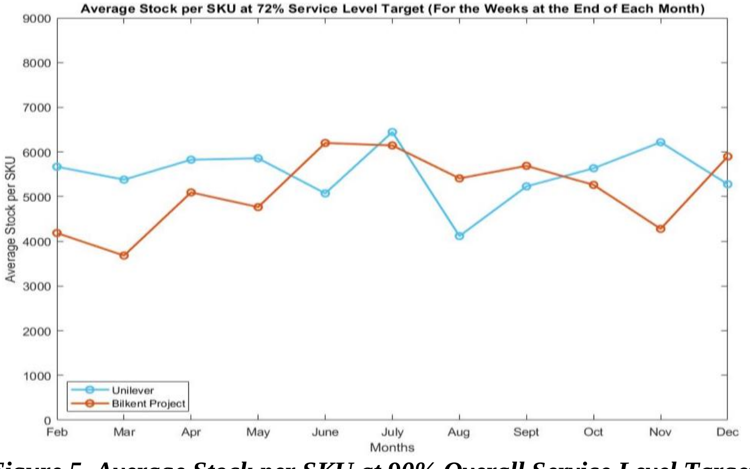
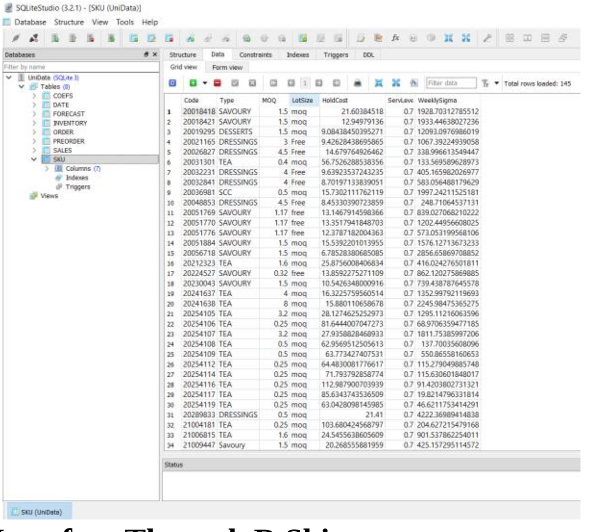


Figure 5. Average Stock per SKU at 90% Overall Service Level Target

APPENDIX D. Database in SQLite



	Code	Type	MOQ	LotSize	HoldCost	ServLeve	WeeklySigma
1	20018418	SAVOURY	1.5	moq	21.60384518	0.7	1928.70312785512
2	20018421	SAVOURY	1.5	moq	12.94979136	0.7	1933.44638027236
3	20019295	DESSERTS	1.5	moq	9.08438450395271	0.7	12093.0976986019
4	20021165	DRESSINGS	3	Free	9.42628438695865	0.7	1067.39224939058
5	20026827	DRESSINGS	4.5	Free	14.679764926462	0.7	338.996613549447
6	20031301	TEA	0.4	moq	56.7526288538356	0.7	133.569589628973
7	20032231	DRESSINGS	4	Free	9.63923537243235	0.7	405.165982026977
8	20032841	DRESSINGS	4	Free	8.70197133839051	0.7	583.056488179629
9	20036981	SCC	0.5	moq	15.7302111762119	0.7	1997.24211525181
10	20048853	DRESSINGS	4.5	Free	8.45330390723859	0.7	248.71064537131
11	20051769	SAVOURY	1.17	free	13.1467914598366	0.7	839.027068210222
12	20051770	SAVOURY	1.17	free	13.3517941848703	0.7	1202.44956608025
13	20051776	SAVOURY	1.17	free	12.3787182004363	0.7	573.053199568106
14	20051884	SAVOURY	1.5	moq	15.5392201013955	0.7	1576.12713673233
15	20056718	SAVOURY	1.5	moq	6.78528380685085	0.7	2856.65869708852
16	20212323	TEA	1.6	moq	25.8756008406834	0.7	416.024276501811
17	20224527	SAVOURY	0.32	free	13.8592275271109	0.7	862.120275869885
18	20230043	SAVOURY	1.5	moq	10.5426348000916	0.7	739.438787645578
19	20241637	TEA	4	moq	16.3225759560514	0.7	1352.99792119693
20	20241638	TEA	8	moq	15.880110658678	0.7	2245.98475365275
21	20254105	TEA	3.2	moq	28.1274625252973	0.7	1295.11216063596
22	20254106	TEA	0.25	moq	81.6444007047273	0.7	68.9706359477185
23	20254107	TEA	3.2	moq	27.9358828468933	0.7	1811.75385997206
24	20254108	TEA	0.5	moq	62.9569512505613	0.7	137.70035608096
25	20254109	TEA	0.5	moq	63.773427407531	0.7	550.86558160653
26	20254112	TEA	0.25	moq	64.4830081776617	0.7	115.279049885748
27	20254114	TEA	0.25	moq	71.793792858774	0.7	115.630601848017
28	20254116	TEA	0.25	moq	112.987900703939	0.7	91.4203802731321
29	20254117	TEA	0.25	moq	85.6343743536509	0.7	19.8214796331814
30	20254119	TEA	0.25	moq	63.0428098145985	0.7	46.6211753414291
31	20289833	DRESSINGS	0.5	moq	21.41	0.7	4222.36989414838
32	21004181	TEA	0.25	moq	103.680424568797	0.7	204.627215479168
33	21006815	TEA	1.6	moq	24.5455638605609	0.7	901.537862254011
34	21009447	Savoury	1.5	moq	20.268555881959	0.7	425.157295114572

APPENDIX E. Interface Through R Shiny