ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Aslı BORU İPEK

ANALYSIS OF INVENTORY ROUTING PROBLEM WITH SIMULATION AND BIG DATA APPROACH

DEPARTMENT OF INDUSTRIAL ENGINEERING

ADANA, 2020

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

ANALYSIS OF INVENTORY ROUTING PROBLEM WITH SIMULATION AND BIG DATA APPROACH

Aslı BORU İPEK

PhD THESIS

DEPARTMENT OF INDUSTRIAL ENGINEERING

We certify that the thesis titled above was reviewed and approved for the award of the degree of the Philosophy of Doctorate by the board of jury on 14/01/2020.

Prof. Dr. Rızvan EROL
SUPERVISOR

Assoc. Prof. Dr. Tolunay GÖÇKEN
MEMBER

Assist. Prof. Dr. Ebru YILMAZ
MEMBER

Assist. Prof. Dr. Yusuf KUVVETLİ
MEMBER

Assist. Prof. Dr. Ayşe Tuğba DOSDOĞRU
MEMBER

PhD Thesis is written at the Department of Industrial Engineering of Institute of Natural and Applied Sciences of Çukurova University.
Registration Number:

Prof. Dr. Mustafa GÖK
Director
Institute of Natural and Applied Sciences

This study supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK), 2211-C PhD scholarship program.

This dissertation is dedicated to my father.

ÖZ

DOKTORA TEZİ

ENVANTER ROTALAMA PROBLEMİNİN SİMÜLASYON VE BÜYÜK VERİ YAKLAŞIMI İLE ANALİZİ

Aslı BORU İPEK

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Prof. Dr. Rızvan EROL
II. Danışman : Doç. Dr. Mustafa GÖÇKEN
Yıl: 2020, Sayfa: 123
Jüri : Prof. Dr. Rızvan EROL
: Doç. Dr. Tolunay GÖÇKEN
: Dr. Öğr. Üyesi Ebru YILMAZ
: Dr. Öğr. Üyesi Yusuf KUVVETLİ
: Dr. Öğr. Üyesi Ayşe Tuğba DOSDOĞRU

Tedarik zincirlerinde envanter ve rotalama, şirketin hayatta kalması ve sürdürülebilirliği için anahtar unsurlardır. Uygun olmayan envanter kontrolü ve rotalama şirketin karlılığına ve pazar rekabetçiliğine zarar verecektir. Bu nedenle, envanter rotalama probleminin (IRP) çözümü, her bir tedarik zinciri üyesi için en uygun envanter seviyesini sağlamalı ve ayrıca periyotlar boyunca rotalama maliyetini en aza indiren etkili bir rota sağlamalıdır. Bu noktada, melez metodoloji, IRP'deki değişikliklerin riskini ve maliyetini en aza indirgeyebilmek için araştırmacıya imkan sağlamaktadır. Bu çalışmada IRP'yi çözmek için iki bölüm içeren yeni melez metodolojiler geliştirilmiştir. İlk bölümde, aralıklı, düzensiz ve çok düzensiz talebi içeren talep modelleri için üç farklı rotalama stratejisinin değerlendirildiği IRP'yi çözmek için simülasyon optimizasyonu ve yapay zeka tabanlı simülasyonu içeren yeni bir melez yöntem oluşturulmuştur. İkinci bölümde, müşteri değerlendirmelerini anlamlı bilgilere dönüştürmek ve değer yaratmak için büyük veri ve duygu analizi kullanılmıştır. Daha sonra, IRP'yi analiz etmek için yapay zeka ve simülasyon optimizasyonu kullanılmıştır. Melez metodolojiler tarafından elde edilen sayısal sonuçlar, dinamik ve stokastik IRP'nin, metot uygun şekilde kurulduğunda başarıyla çözülebileceğini göstermiştir.

Anahtar Kelimeler: Tedarik Zinciri Yönetimi, Envanter Rotalama Problemi, Simülasyon, Simülasyon Optimizasyonu, Büyük Veri

ABSTRACT

PhD THESIS

ANALYSIS OF INVENTORY ROUTING PROBLEM WITH SIMULATION AND BIG DATA APPROACH

Aslı BORU İPEK

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING

Supervisor	: Prof. Dr. Rızvan EROL
II. Supervisor	: Assoc. Prof. Dr. Mustafa GÖÇKEN
	Year: 2020, Pages: 123
Jury	: Prof. Dr. Rızvan EROL
	: Assoc. Prof. Dr. Tolunay GÖÇKEN
	: Assist. Prof. Dr. Ebru YILMAZ
	: Assist. Prof. Dr. Yusuf KUVVETLİ
	: Assist. Prof. Dr. Ayşe Tuğba DOSDOĞRU

In supply chain environments, inventory and routing are key elements to company's survival and sustainability. Unsatisfactory inventory control and routing will hurt company profitability and market competitiveness. Therefore, the solution of inventory routing problem (IRP) should provide optimal inventory level for each supply chain member and also obtain efficient route that minimizes the routing cost over periods. At this point, hybrid methodology can empower analyst to minimize the risk and cost of the changes in the IRP. In this paper, new hybrid methodologies that include two parts are developed to solve the IRP. In the first part, a new hybrid method including simulation optimization and artificial intelligence based simulation is created to solve the IRP in which three different routing strategies are evaluated for uneven demand patterns including intermittent, erratic, and lumpy demand. In the second part, big data and sentiment analysis are used to transform customer reviews into meaningful information and to create value. Then, artificial intelligence and simulation optimization are employed to analyze the IRP. Numerical results proposed by the hybrid methodologies showed that dynamic and stochastic IRP can be successfully solved when the method is properly set up.

Keywords: Supply Chain Management, Inventory Routing Problem, Simulation, Simulation Optimization, Big Data

GENİŞLETİLMİŞ ÖZET

Giderek karmaşıklaşan endüstriyel ortamlarda müşterilerin artan beklentilerini karşılamak için işletmelerin kendilerini sürekli geliştirmeleri gerekmektedir. Ayrıca, işletmelerin hızla gelişen ve değişen rekabet ortamında ayakta kalabilmesi için rakipleriyle rekabet edebilmesi ve pazar paylarını genişletebilmesi de önemlidir. Bu kapsamda gerçek yaşamla tutarlı ve geçerli ekonomik koşulları dikkate alarak yeni modellerin geliştirilmesi işletmelerin kârını büyük ölçüde arttıracaktır.

Günümüz piyasa koşullarında maliyetleri minimize ederek kârını artırmak isteyen işletmeler etkin ve verimli envanter kontrol politikası uygulaması gerekmektedir. Envanterler, işletme sermayelerinin büyük bir kısmını içermekte ve kullanılamadığı takdirde ciddi bir kazanç ve fırsat kaybına neden olmaktadır. İşletmelerin envanter politikasını en düşük maliyetle yürütebilmesi siparişlerin uygun zaman ve miktarda verilmesine bağlıdır. İşletmeler genellikle maliyeti en aza indirmek için envanter seviyesini azaltmaya çalışmaktadır ve bu nedenle genellikle uygulamada eksiklikler gözlemlenmektedir. Bu noktada, doğru envanter kontrol politikasının kullanılması, envanter eksikliğin azaltılmasının en etkili ve verimli yoludur. Envanter kontrol politikalarında temel soru, tedarik zinciri üyelerinde siparişin boyutunun ve zamanlamasının belirlenmesidir. Tedarik zinciri üyelerinde bulunan her bir ürünün envanter seviyesinin eksiklik veya fazlalık olmadan kesin olarak belirlenmesi ve toplam tedarik zinciri maliyetinin en aza indirilmesi tedarik zincirinin performansını büyük ölçüde etkilemektedir. Uygun envanter seviyelerinin bulunması satış kaybını önlemek için gereklidir. Satış kaybını önlemek için fazladan envanter tutma depolama maliyetlerini artıracak ve işletme karını düşürecektir. Toplam tedarik zinciri maliyetini en aza indirmek için her bir tedarik zinciri üyesinin envanteri kontrol edilmelidir. Her şirket farklı süreçlere ve farklı stok biçimlerine sahip olabilir. Bu nedenle, tüm tedarik zincirinde envanter kontrolünün iyi bir şekilde yapılması etkili bir sistem geliştirmek için gereklidir. Ayrıca, müşteri taleplerinin tamamıyla karşılanması, maliyetinin uygun olması ve teslimatların zamanında yapılması için envanter ve rotalamanın eş zamanlı yürütülmesi de büyük önem arz etmektedir. Sistematik bir yapısı olan envanter ve rotalama problemi sürekli gelişme gösteren bir problemdir. Envanter ve rotalama probleminde, tedarik zinciri üyelerinin envanter düzeyleri kontrol edilerek, tedarik zincirinde envanter yönetimi ve araç rotalama konusunda ortaya çıkan problemlerin eş zamanlı olarak çözüme ulaştırılması sağlanır. Envanter ve rotalama problemlerinde işletmeler, müşterilere en iyi hizmeti verebilmek için araçlarını nasıl yönlendireceklerine ve bu yerlerin her birinde ne kadar envanter tutacaklarına karar verirler.

Literatürde envanter ve rotalama problemi ile ilgili yapılan çalışmalarda genellikle modelde bulunan değişkenlerin bir kısmının sabit veya biliniyor olduğu varsayılarak basitleştirmeler yapılmaktadır. Hâlbuki gerçek hayatta işletmelerin böyle bir sistem ile yönetilmesi mümkün değildir. Bu nedenle, stokastik parametreler kullanılarak dinamik çevre koşullarının modele eklenmesi gerekmektedir. Örneğin, stokastik envanter ve rotalama problemi, müşterilerin taleplerini kesin olarak bilemez ama olasılıksal olarak ele alabilir. Dinamik ve stokastik envanter ve rotalama probleminde ise, sistem oluşturulurken kullanılan parametrelere karar vermede oldukça zorlanılır. Ayrıca, yeni kararlar alındığında sistem yeniden revize edilir. Dolayısıyla dinamik ve stokastik modellerde, bir çözüm stratejisi tasarlanırken her adımın ayrıntılı olarak düşünülmesi gerekmektedir. Elde edilen sonuçların uygulamada etkin olabilmesi için iş dünyasının çevresini tanımlayan temel unsurları dikkate alması gereklidir. Mesela, belirsizlik envanter ve rotalama probleminde karmaşıklık yaratan bir faktör olmasına rağmen mevcut bulunan sistemlerde nadiren hesaba katılmaktadır. Birçok çalışmada envanter ve rotalama problemi alanında belirsizliğin etkin şekilde ele alınmamasından kaynaklanan sorunların giderek arttığı ve kullanılan sistemlerle çözümünün pek mümkün olmadığı vurgulanmaktadır. Yöneticiler ve araştırmacılar karşılaştıkları bu tür zorluklarla başa çıkmak için yeni yöntemler geliştirirken bir yandan da bilinen yöntemleri melez olarak kullanarak yöntemlerin performansını arttırmaya çalışmaktadırlar.

Günümüzde, tüm faaliyetler, günler veya daha kısa zaman aralıklarında gerçekleşmektedir. Dolayısıyla, tedarik zincirinde kararların verilmesi hızlı ve esnek bir modelleme yaklaşımı gerektirir. Bu konuda başarısını kanıtlayan yöntemlerden biri simülasyondur. Bilgisayarların kullanım alanının artmasıyla birlikte hızla gelişen bir karar verme tekniği olan simülasyon günümüzde birçok farklı alanda da kullanılmaktadır. Envanter yönetimi de bu alanların içinde yer almaktadır. Günümüz iş dünyasında tedarik zinciri üyelerinin rekabet avantajı sağlayabilmesi için sistem içinde oluşan aksaklıkları en kısa sürede, en doğru ve en etkin bir şekilde bulması gerekmektedir. Burada simülasyon sistem performansının dinamik çevre koşullarında birçok değişkene bağımlı olarak nasıl sonuçlar verdiğini gözlemlememize olanak sağlamaktadır. Ayrıca, mevcut bulunan bir sistemi analiz etmek veya gerçek hayatta oluşabilecek sonuçlara en yakın sonuçları elde edebilmek için kullanılabilecek en iyi çözüm tekniklerinden biridir. Simülasyon, rotalama ve envanter kontrol politikasında sistemin işlemesi için sistemin davranışlarının analiz edilmesinde veya değişiklik stratejilerinin değerlendirilmesinde kullanılabilir. Özellikle, alınacak stratejik kararların tedarik zinciri üyeleri üzerindeki etkilerini karşılaştırarak analiz etmek ve en iyi sonuç veren stratejiyi belirlemek için simülasyondan yararlanan yöneticiler önemli bir zaman ve maliyet külfetinden kurtulmaktadır. Diğer yandan not etmek gerekir ki simülasyon tek başına problemleri çözme yeteneğine sahip değildir, sadece problemin açık olarak tanımlanmasına ve sayısal olarak alternatif çözümlerin değerlendirilmesine olanak sağlamaktadır. Bu nedenle, tedarik zinciri yapısına simülasyon ve optimizasyonun entegre edilmesi gerekmektedir. Simülasyon optimizasyonu modern hesaplama gücü ile karmaşık tedarik zinciri problemlerini modellemede oldukça iyi sonuçlar vermektedir. Sistemin bileşenleri ve sistem bileşenlerinin sınırları arasında hiçbir kısıtlama yoktur.

Çalışmada envanter ve rotalama problemini çözmek için yeni melez metodolojiler geliştirilmiştir. İlk bölümde, simülasyon optimizasyonu ve yapay zeka tabanlı simülasyonu içeren yeni bir melez yöntem tasarlanmıştır. İkinci bölümde, müşteri değerlendirmelerini anlamlı bilgilere dönüştürmek ve değer yaratmak için büyük veri ve duygu analizi kullanılmıştır. Daha sonra, envanter ve rotalama problemini analiz etmek için yapay zeka ve simülasyon optimizasyonu kullanılmıştır. Her işletmeye kolayca adapte edilebilecek olan yöntemler, en uygun rota ve optimum envanter kontrol parametrelerinin tedarik zincirinin performansını artıracak şekilde belirlenmesi amacıyla kullanılmaktadır. İşletmelerin envanter yönetimi ve rotalama ile ilgili karşılaşabileceği problemlere cevap bulabilmeleri ve değişen çevre koşullarında bile işletmelerin oluşturulan sistemi kolayca kendi bünyelerine adapte edebilmeleri tasarlanan modeller ile sağlanabilmektedir.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Prof. Dr. Rızvan EROL and Assoc. Prof. Dr. Mustafa GÖÇKEN for his devoted guidance, suggestions and support.

I am greatly indebted to Assist. Prof. Dr. Ayşe Tuğba DOSDOĞRU for generously sharing her vast knowledge and offering help whenever needed.

I am also grateful to Assoc. Prof. Dr. Tolunay GÖÇKEN, Assist. Prof. Dr. Ebru YILMAZ and Assist. Prof. Dr. Yusuf KUVVETLİ for accepting to read and review this thesis and for their invaluable suggestions.

I would like to thank to my family, for their invaluable support. I would like to express my special thanks to my husband and my mother who accompanied me whenever I needed. My father always supported me during all steps I took.

I would also like to express my deepest gratitude to Adana Alparslan Türkeş Science and Technology University Industrial Engineering Department for the warm and stimulating environment.

I gratefully acknowledge the Scientific and Technological Research Council of Turkey (TÜBİTAK) for the 2211-C PhD scholarship program.

TABLE OF CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

1. INTRODUCTION

The term Supply-Chain Management (SCM) includes the broad range of activities, processes, and other resources until the end product is delivered to the customer. It ensures a bird's-eye view of the whole supply chain (Mercado, 2008). The supply chain contains the suppliers, manufacturers, warehouses, retailers, transporters, and even customers themselves. Successful SCM requires various decisions concerned with the information flow, products, and funds. These supply chain decisions can be classified into three phases including strategy, planning, and operation according to the time frame and the frequency of each decision.

In strategy phase, a company decides how to set up the supply chain in the next few years. At this point, the marketing and pricing plans are taken into account for product. Company also decides supply chain configuration, resource allocation, and performing processes. Furthermore, this phase includes whether to the location and capacities of supply chain members. Transportation mode and information system are also important to improve the supply chain performance during this phase.

In planning phase, the time horizon is a quarter to a year for decisions. Fixed supply chain configuration that is defined in supply chain strategy phase is used. This configuration creates constraints within which planning should be carried out. The planning phase in companies starts with demand forecasting. This phase includes decision related to supply, manufacturing, and subcontracting. Inventory control policies, market size, and price promotions are also considered during this phase.

In operation phase, the time frame is considered as weekly or daily for decisions. Note that fixed supply chain configuration and specified planning policies were used at this phase. This phase includes decision related to individual customer orders such as inventory allocation, production, shipping mode and shipment, vehicle routing, and replenishment orders. Due to the short term, demand information includes less uncertainty in operational decisions.

More detailed information related to decision phases in a supply chain can be found in Chopra and Meindl (2007).

Companies implement SCM in order to maintain their competitive advantage and to meet the customer demands. It can also be said that SCM provides a significant advantage to stay ahead of their international competition. The SCM includes all processes in the supply chain. Companies should organize activities to meet customer orders quickly and smoothly. Required products and materials should be provided for production. Inventory management is one of the most important processes to determine the inventory levels of all the necessary materials. Transportation and routing are also so important to deliver the finished products from the supply chain members to the customers.

1.1. Inventory Control System

Customers have the detailed information about products, quality, and service. In the past, customers have generally gone to a local retailer to buy the products but now Web can be used to buy the products that are provided by any company in the world. In today's competitive world, companies should raise their attention to customers. They must satisfy customer needs at profit to remain viable in the sector. Companies can simply keep prices down or other ways can be used to improve customer satisfaction. In this case, inventory management is important to satisfy the incoming customer orders. Companies can hold less stock to reduce costs but this situation can cause more frequent stockout. Due to the continuing trend on customer satisfaction, we have to balance the holding cost and stockout cost.

The inventory management varies according to industries and organization. If there is an improper inventory management system, most operations are generally impossible. Stocks influence the lead times and availability of materials.

In addition, they affect the financial performance. Moreover, stocks provide more efficient and productive operations. To completely describe an inventory system, following issues should be analyzed (Liu and Esogbue, 1999).

Demand: Inventory problem is encountered because of the time sensitive demands. In general, demand cannot be controlled directly. The reason is that it generally depends on outside requirements. When exact information is available about the demand, it is known as deterministic demand. Also, inventory system such as Wilson economic lot size model is denoted as a deterministic system. When the demand is uncertain, it can be defined by some probability distribution. This situation is known as stochastic demand. The resultant inventory system is denoted as stochastic system. However, the source of uncertainty in demand is not so much because of randomness but ambiguity, vagueness, imprecision, and etc., then it is denoted as fuzzy demand and the resultant inventory system is denoted as a fuzzy system.

Order: The order quantity is always controllable. Nevertheless, lead time generally occurs in the ordering process. It is the length of time between placing an order and its actual addition to stock. In classic approach, lead time is assumed small enough and therefore its value is taken zero. Nevertheless, this assumption might not be valid in a number of situations. In general, very long lead time is involved and therefore, it is important factor in inventory management.

Cost: In inventory control system, three types of costs are usually utilized. First type of cost is the ordering cost that can be considered as the cost of ordering products from outside. Second type is holding cost that can be considered as the cost of carrying inventories over a certain period of time. Final type is stockout cost that can be considered as the cost incurred as a result of shortages.

Constraints: Inventory control system can include various constraints such as capacity constraints, scheduling constraints, time constraints, and etc. Potential constraints must be taken into account to determine operating doctrines.

Operating Doctrine: In supply chain, optimal order quantity should be found to minimize the total inventory cost. It is determined that when one cost decreases (or increases), the other costs may increase (or decrease). Therefore, the ordering cost, holding cost, and shortage cost should be balanced. The operating doctrine is the policy to follow in optimal inventory management considering the foregoing five issues.

More detailed information can be found in Liu and Esogbue (1999).

1.2. Vehicle Routing

Transportation can be defined as moving inventory from point to point. It highly affects the responsiveness and efficiency of supply chain. Faster transportation can allow a supply chain member to be more responsive; however, it decreases its efficiency. Although strategy targets can vary from company to company, transportation is important to satisfy the customer needs in today's competitive environment (Chopra and Meindl, 2007). Transportation costs contain a main part of logistics costs. Previous studies showed that about 5% to 6% of retail product price includes the transportation costs. This share increases up to 30% in some product like foodstuffs. For this reason, controlling the transportation costs is critical for companies. The common problems with transportation generally occur in transportation network, vehicles' capacity, vehicle routing, and meeting customer needs (Hajghasem and Shojaie, 2016). Among them, vehicle routing directly affects the distribution speed, cost, and benefit. Determining proper route can increase customer satisfaction, improve service level, and decrease the cost. Various algorithms have been created in literature and many software packages are developed to facilitate the solving of Vehicle Routing Problem (VRP). The VRP has known as one of the successful field in operational research. Basically, VRP can be defined as the problem of identifying least-cost routes from one place to another place.

Real world problems include a large number of variants and therefore, many ways are tried to extend the classical VRP. Until recently, various surveys and bibliographies have indeed been published on the VRP. For example, Gendreau, Laporte, and Séguin (1996) summarized the previous studies considering stochastic VRP (SVRP). In SVRP, some elements of the problem such as demands and travel times are random.

Toth and Vigo (2002) presented the static and deterministic capacitated VRP (CVRP). In the study, the case with symmetric and asymmetric cost matrices was considered for CVRP. Montoya-Torres et al. (2015) presented the studies related to the VRP with multiple depots. In addition, single objective function and multiple objective functions were analyzed in the VRP. Solution methodologies were divided into three main parts including exact method, heuristic or meta-heuristic. The analysis results showed that the most widely used meta-heuristic algorithms are genetic algorithms, tabu search, and simulated annealing. Koç, Bektas, Jabali, and Laporte (2016) analyzed the fleet size and mix VRP, the heterogeneous fixed fleet VRP, and the fleet size and mix VRP with time windows. The study demonstrated a detail analysis of the previous studies and comparatively evaluated the performance of the meta-heuristics. Braekers, Ramaekers, and Nieuwenhuyse (2016) presented the VRP variants that contain real-world constraints and assumptions that increase the applicability of model in practice. It is determined that meta-heuristics are most widely used method while real-time solution methods and simulation are employed less often.

Cordeau, Gendreau, Laporte, Potvin, and Semet (2002) presented the four attributes including accuracy, speed, simplicity, and flexibility for the VRP. They concentrated on the Clarke and Wright algorithm, the Sweep algorithm, and the Fisher and Jaikumar algorithm. In addition, Taburoute, Taillard, adaptive memory procedure, granular tabu search algorithm, and unified tabu search algorithm were assessed for the VRP. Cordeau, Laporte, Savelsbergh, and Vigo (2007) analyzed the VRP considering the exact algorithms, heuristics, and meta-heuristics. The classical VRP heuristics were described under route construction methods, two-phase methods, and route improvement methods. VRP meta-heuristics were divided into three classes including local search, population search, and learning mechanisms. In the study, stochastic VRP, VRP with time windows, and IRP were also taken into account. Prodhon and Prins (2016) presented the variants of the VRP. In addition, basic heuristics, local search procedures, and meta-heuristics were analyzed for the VRP. Furthermore, an approach was given based on splitting giant tours.

1.3. Inventory Routing Problem

The distribution process is directly affected by the transportation and inventories management. However, transportation and inventory management are generally considered independently in modeling approaches. Thus, their mutual impact was usually neglected. To analyze the interrelationship between inventory management and transportation, these two activities can be considered simultaneously. This is known as an IRP. In this problem, the objective is to minimize the total cost of inventory management and vehicle routing. Due to the problem complexity, an optimal solution is so far unreachable for each problem characteristic. The complexity of problem is related to simultaneous consideration of inventory management and routing (Popović, Vidović, and Radivojević, 2012).

IRP can include the large variety of aspects and assumptions. This makes the IRP very hard to classify. For example, Kleywegt, Nori, and Savelsbergh (2002) only considered the IRP with direct deliveries in which each vehicle visits only one customer. They developed a dynamic programming based approximation methods. In the study, combination of common random numbers and orthogonal arrays were used to reduce the required sample size. Homogeneous vehicles with known capacity were used to distribute a product from supplier to customers. Unsatisfied demand was considered as lost sales. The main aim was to maximize the expected discounted value while choosing a distribution policy.

Andersson, Hoff, Christiansen, Hasle, and Løkketangen (2010) classified the IRP in order to discuss the similarities and differences between the proposed systems as follows.

Time: When time is taken into account to classify the IRP literature, it is seen that there are three different types expressing the planning periods used. Instant is the first type in which the planning horizon is too short, the maximum one visit per customer is required.

Finite is the second type. When a customer needs more than one visit, it can be considered a finite problem. If there is a finite and natural end to the horizon, a fixed horizon can be used.

Infinite is the third type. When analyzing a problem related to the infinite planning horizon, distribution strategies are taken as the decisions.

Demand: On the one hand, there are cases in which the demand is considered deterministic. On the other hand, there are cases where stochasticity is an important component and integrated into the solution framework. The proposed method is called stochastic if it contains uncertainty about demand and it is called deterministic if it does not contain uncertainty about demand.

Topology: One to one, one to many, and many to many can be used to classify problems. The many-to-one case is included in the one-to-many case. In addition, it can easily be transformed to a one to-many topology.

Routing: The routing component of the IRP can be considered as the VRP. In literature, the warehouse for the VRP is generally defined as the central place where the goods are distributed and all the routes start and end. The VRP can be defined as the distribution of products from one or several warehouses in the logistics system and the collection of products from customers.

Vehicle fleet: The fleet used to distribute or collect products can be classified by composition and size. In a homogeneous fleet, all vehicles have the same characteristics such as speed, fixed cost, and variable cost. In the heterogeneous fleet, some or all of the features of the vehicles are different.

Inventory: Basically, two inventory control systems are available. The first is the periodic review policy, where the inventory level is revised at periodic times and the order quantity is variable. According to this approach, the period between replenishment orders is fixed since orders are given periodically. The second option is a continuous review policy where the time between orders is variable. The selected inventory control policy differs according to the business. It should be ensured that the proposed models can be easily adapted to dynamic environmental conditions. Each company can create different systems by using inventory control methods and techniques depending on their inventory policy, production type, financial position, and other factors (Ada, 2010).

1.4. Problem Definition

To remain competitive in today's business world, companies should cope with the increasing customer expectations and with the growing markets. Products should be obtained at the right place, at right quantity, and at right time. Furthermore, companies have to consider risks formed by the sudden fluctuations in global and local economies. Thus, the success measures of the companies include many different factors such as higher quality, lower costs, shorter lead time, and etc. At this point, SCM has become an important necessity for companies (Gumus, Guneri, and Keles, 2009). Successful SCM requires many decisions including the flow of information, product, and funds (Chopra and Meindl, 2007). In this paper, we focused on the inventory control and routing that are the core topics in SCM. The performance of supply chain is closely linked to the performance of inventory control and routing. In addition, simultaneous consideration of inventory control and routing offers a tremendous opportunity for companies to gain competitive advantage in supply chain. The simultaneous consideration of inventory control and routing is known as the IRP. In IRP, customer demands are met while satisfying the objectives related to the inventory control and the routing. In this case, determining the customer demand is a challenging task for supply chain to eliminate the uncertainty and make the supply chain stable. Inaccurate demand forecasting can increase the total supply chain cost. Therefore, demand forecasting is so critical for any supply chain to make the correct decisions and to achieve the benefits in this regularly changing business scenario (Bhadouria and Jayant, 2017). Accurate forecasting of demand under uncertain environment also improves supply chain activities (Jaipuria and Mahapatra, 2014). In literature, many researches are presented to predict the customer demand based on casual models such as regression and time-series models such as moving-average (Efendigil, Önüt, and Kahraman, 2009). Although these models perform well, some limitations are available. For example, accurate prediction is generally guaranteed when a large amount of data is used in these models. Therefore, there is a need to develop a more efficient and precise model for forecasting the demand. At this point, artificial intelligence (AI) techniques can be employed for demand forecasting.

AI is interested in intelligent behavior in artifacts. At this point, the perception, learning, reasoning, communicating, and acting in complex environments can be considered as an intelligent behavior (Nilsson, 1998). Today's complex environment needs intelligent behavior in systems to combine knowledge, methodologies, and procedures from various sources. AI adapts itself and learns to do better problem solving in stochastic and dynamic environments.

To increase the performance of AI technique, many methodologies that find actual and potential use within simulation environments were developed (Doukidis and Angelides, 1994). A combination of AI and simulation improves the functionality of the simulation due to the realistic perspective. Furthermore, it provides intelligent and accurate system for decision-making process in SCM.

IRP includes parameters and constraints related to the routing and the inventory. In there, vendor managed inventory control policy is generally applied in supply chains. The simultaneous controlling of inventory and routing has a huge impact on companies since it provides faster and more efficient ways to improve the supply chain performance. In IRP, the supply chain can be managed as an interdependent and interconnected structure. At this point, simulation ensures the flexibility to model supply chain under the interdependent and interconnected structure. Furthermore, it provides the essential level of realism. Simulation allows user to change parameters easily within the model (Cope, Fayez, Mollaghasemi, and Kaylani, 2007). In addition to simulation, optimization models allow researchers to determine the best possible alternatives (Göçken, Dosdoğru, Boru, and Geyik, 2017). Hence, integrating simulation and optimization into the IRP ensures a remarkable solution model. It also provides a structured approach to the system of design and configuration when analytical expressions are unavailable for input/output relationships (Ammeri, Hachicha, Chabchoub, and Masmoudi, 2011). Hence, simulation optimization (SO) is an exciting and fast developing area for both research and practice in supply chains. In this paper, SO and AI based simulation are firstly developed to solve the IRP under different demand patterns including intermittent, erratic, and lumpy. In proposed method, SO is used to conduct "what-if" scenario analysis considering the occurrence of unplanned events in supply chain. Then, AI is employed to forecast demand and simulation embraces proposed supply chain with their complicated and nonlinear relationships. To analyze proposed methods with a greater level of detail, various performance measures such as totally met order quantity (TMOQ), totally lost order quantity (TLOQ), partially lost order quantity (PLOQ), number of totally met order (NTMO), number of totally lost order (NTLO), and number of partially lost order (NPLO) are used.

In today's competitive world, customers review offer new opportunities for companies to improve the supply chain. In addition, it enables companies to gain feedback to forecast the sales. Therefore, we used the sentiment analysis to offer a new methodology for the IRP. Firstly, SentiStrength is used to calculate the score for customer reviews. After, Artificial Neural Network (ANN) was integrated with the output of SentiStrength and demands of same type of product to solve the IRP.

After that the output of ANN is fed into nondominated sorting genetic algorithm II (NSGA-II) based SO.

1.5. Research Objectives and Contributions

This research is focused on developing new methodologies to solve the IRP. It includes two parts. SO and AI based simulation are employed in the first part while SO and big data analysis are employed in the second part.

The main contribution of the first part can be summarized as follows: (i) the application of a NSGA-II to optimize the parameters of IRP considering two objective functions; (ii) the application of a Genetic Algorithm (GA) to determine the de-noising degree, the number of neurons in hidden layers, and training algorithm; (iii) the implementation of de-noising with Maximal Overlap Discrete Wavelet Transform (MODWT) to improve the customer demand data; (iv) the application of a Genetic Algorithm based Artificial Neural Network (GA-ANN) to forecast the uneven customer demand patterns. Integration of NSGA-II based SO and GA-ANN based simulation is necessary to provide a better coordination level in the supply chain. This integration directly affects the synchronization and overall supply chain performance.

The main contribution of the second part can be summarized as follows: (i) the implementation of a big data analysis; (ii) the application of SentiStrength to calculate the sentiment score of a customer review; (iii) the application of the ANN to forecast the customer demand; (iv) the application of the NSGA-II based SO to solve the IRP.

1.6. Overview

The remainder of this thesis is created as follows. Chapter 2 discusses the relevant contributions from the literature. In particular, three main research areas are reviewed: (1) inventory routing problems; (2) demand forecasting; and (3) sentiment analysis and big data. Chapter 3 describes the proposed methodology that includes the details of both the optimization phase and the simulation phase. In Chapter 4, a detailed analysis of proposed systems is given. Finally, Chapter 5 provides concluding remarks for the results obtained in this research.

2. LITERATURE REVIEW

Over the years, several computing methods are used synergistically rather than exclusively in order to improve the supply chain performance. The construction of complementary hybrid methods provides a powerful solution methodology in real-world problems. In this paper, hybrid inventory and routing based methods are firstly analyzed. Then, demand forecasting that has generated an increasing interest nowadays because of the arising of competition is taken into account. Finally, big data and sentiment analysis are investigated.

2.1. Solution Methods for Inventory Routing Problem

2.1.1. Exact Methods

Mathematical determination of the exact solution of the IRP is a very complex task. Campbell, Clarke, Kleywegt, and Savelsbergh (1998) presented a two phase algorithm for the IRP. Firstly, algorithm was used to determine when and how much to transfer to each customer. An integer programming approach was the heart of the first phase. In the second phase, a set of delivery routes were determined. Note that small set of routes were used in order to make integer program computationally tractable. In the study, the stochastic IRP was also modeled as a discrete time Markov decision process.

Campbell and Savelsbergh (2004) presented two phase methodology in which integer programming was used in the first phase to define a high-level base plan and second phase considered vehicle routes and schedules. In the study, homogeneous vehicles were used and the usage was deterministic. Also, customers were classified as critical customers and impending customers. Zhong and Aghezzaf (2011) considered the single-vehicle cyclic IRP whose aim was to determine an optimal distribution plan on a cyclical basis considering the inventory at retailers. In the paper, a hybrid global optimization method was proposed to solve the problem to global optimality. The proposed method was compared with standard steepest descent hybrid algorithm.

Coelho and Laporte (2013a) used a branch-and-cut algorithm to solve the multi-product IRP. They considered deterministic version of the problem and no stockout was allowed to occur. Consistency features including driver partial consistency and visit spacing were applied in a multi-vehicle system. Coelho and Laporte (2013b) firstly solved heterogeneous instances of the multi-vehicle IRP using branch-and-cut algorithm. Instances were solved with flexibility and consistency features. The proposed method was also used to solve the classical IRP that considered a homogeneous or a heterogeneous fleet of vehicles under different inventory policies. Roldán, Basagoiti, and Coelho (2016) proposed an algorithm that decomposed the problem into smaller parts. The first part determined which customers to be visited in each period. The second part used three different customer selection methods to determine how much to transfer to each customer in each period. Final part was to create vehicle routes using specialized exact algorithm. Nikolakopoulos and Ganas (2017) presented a centralized economic model predictive control strategy that used a multi-period deterministic model. A rolling horizon was used to solve the dynamic and stochastic IRP in which homogeneous multi-vehicle was used. No split delivery and no backlogging were permitted.

Iassinovskaia, Limbourg, and Riane (2017) used the classical branch-and-cut CPLEX solver to solve the mixed integer linear program (MILP) on small-scale instances of pickup and delivery IRP within time windows. In addition, clustering heuristic was employed before executing the branch-and-cut algorithm. In the study, homogeneous fleet of vehicles was used for two echelon supply chain including a main producer and multiple customers. Cheng, Yang, Qi, and Rousseau (2017) used the branch-and-cut algorithm to demonstrate the difference between objective functions and the importance of the heterogeneous fleet for green inventory routing problem. In proposed algorithm, they first added user cuts to the model at the search tree's root node. Then, they used a commercial solver to solve linear programming relaxation problems at each node. Lefever, Aghezzaf, Hadj-Hamou, and Penz (2018) proposed the branch-and-cut algorithm for the IRP with transshipment and investigated four different aspects of the problem including valid inequalities, improved bounds, reformulation, and variable eliminations. In the paper, the maximum level was used as inventory control policy. There was no limit on the number of transshipments and transshipments were made by direct shipping.

Archetti, Christiansen, and Speranza (2018) developed a mathematical programming model and solved with the branch-and-cut method. In proposed IRP, single-commodity and single-vehicle were considered with pickups and deliveries. The solution was obtained considering a decentralized policy in which each customer implemented the (s, S) policy. In addition, the impact of the valid inequalities and vehicle capacity were assessed considering the solution quality and the problem difficulty. Morales, Franco, and Mendez-Giraldo (2018) developed a dynamic programming approach to solve the IRP with time windows constraints and disruptions over the network. In addition, the dynamism of the nodes was added by using the deterministic demand. In the study, different policies combining disruption and dynamism of the nodes were tested.

2.1.2. Fuzzy Based Methods

Fuzzy based methods can be used to model systems that are difficult to define precisely. The uncertainties in the supply chain can be represented by fuzzy logic. Xie and Hu (2009) solved the IRP in emergency logistics with fuzzy demands and split delivery using heuristic algorithm. In the study, a fuzzy demand was converted to a deterministic demand. Homogenous vehicles were used and no backorder was allowed.

Niakan and Rahimi (2015) developed a two-phase approach which included auxiliary crisp multi-objective model and fuzzy solution approach to solve the multi-product IRP. A smoothing approach was applied to forecast the demand for each period. In addition, the seasonal requirement factor was defined as a coefficient due to seasonal variations in demand for some products. In the study, demand, transportation, and shortage cost were considered to be fuzzy. Xiao and Rao (2016) presented an adaptive fuzzy logic-based genetic approach to solve the multi-period IRP with time window constraints. In proposed supply chain, supplier provided multiple products and out of stock inventory was not allowed. The demand of each customer was known.

Rahimi, Baboli, and Rekik (2017) integrated the service level in the IRPs in which certain parameters such as demand, variable transportation costs, and vehicle speed were assumed uncertain and modeled as a fuzzy possibilistic distribution. In addition, the environmental footprint was taken into account considering the greenhouse gas emission. A heterogeneous fleet of vehicles was used with different sizes and different energy technologies. Hu, Li, and Li (2018) proposed a credibilistic goal programming model within the framework of credibility theory to handle the fuzzy IRP. In the study, hazardous materials were used considering fuzzy demands in supply chain that composed of a single supplier, multiple manufacturers, and retailers. The supply chain system included inventory risk and transportation risk. Also, the transportation risk measurement considered the change of the vehicle loading. Li, Chu, Feng, Chu, and Zhou (2018) solved a bi-objective food production IRP using constraint-based two-phase iterative heuristic and a fuzzy logic method. The aim was to minimize the total cost and maximize the average food quality. The customer demand at each retailer was assumed to be deterministic and time varying.

2.1.3. Heuristics and Meta-heuristics Methods

Heuristics and meta-heuristics methods provide powerful and flexible search methodologies. They produce good-quality solutions in reasonable computation time. The aim of these methods is to investigate the solution space effectively and to provide solutions that are close to the optimal.

Gaur and Fisher (2004) presented a heuristic that includes the randomized sequential matching algorithm and randomized splitting of clusters for the periodic IRP. Proposed system was included inventory routing module, truck assignment module, and workload balancing module. In the paper, heterogeneous fleet of trucks and single homogeneous product were used. Popović et al. (2012) created a variable neighborhood search heuristic. The proposed method was compared with the MILP model and the deterministic ‘‘compartment transfer’’ heuristic. In the study, fuel was transported by homogeneous multi-compartment vehicles that have unlimited size. It was also assumed that one vehicle can visit up to three places per route. Coelho, Cordeau, and Laporte (2012) introduced the concept of transshipment within the IRP. In the study, the order-up-to policy or the maximum level policy was used as a replenishment policy and tested with/without transshipment. To solve the problem, an adaptive large neighborhood search was employed. In addition, the impact of individual operators was assessed in proposed heuristic. In the study, stockout was not allowed and single vehicle was used.

Dubedout, Dejax, Neagu, and Yeung (2012) presented three different heuristics to solve the rich IRP problem. The first heuristic was based on the basic deterministic greedy and local search heuristic. In addition, the single start Greedy Randomized Adaptative Search Procedure (GRASP) and the multi-start GRASP methodology were created to improve the solutions quality of “local search”. Proposed methods were tested using 16 different instances adapted from real life data. The results showed that GRASP methodologies can yield better result than the original heuristic.

Qin, Miao, Ruan, and Zhang (2014) presented a new method for the periodic IRP. The inventory problem was analyzed by a local search method. The tabu search method was employed to routing problem. In the study, identical vehicles were used in proposed supply chain network that included a supplier and retailers. Archetti and Speranza (2016) used the heuristic method and compared the retailer-managed inventory, vendor managed inventory, and adjusted vendor managed inventory considering total cost, inventory cost, transportation cost, maximum number of vehicles, average vehicle load, and the number of routes. In the study, homogeneous capacitated vehicles were used and no stockout occurred.

Hasni, Toumi, Jarboui, and Mjirda (2017) presented a general variable neighborhood search algorithm for a multi-product version of the IRP. In the study, they defined six neighborhood structures. In addition, the proposed method was compared with the branch & cut algorithm. Montagné, Gamache, and Gendreau (2019) presented a constructive greedy heuristic based on the sequential resolution of shortest paths in directed acyclic graphs. The performance of the heuristic was compared with respect to the company's current solution. The results showed that greedy heuristic was able to collect more wasted oil, by consuming less fuel, and with fewer trucks.

Meta-heuristics are widely used to solve the IRP. Liao, Li, and Wu (2013) presented two models about supply chain operation. First model was the regular chain mode in which supply chain members belong to the same enterprise. Second model was the franchise chain mode in which supply chain members do not belong to the same enterprise. In the study, agriculture products deteriorated at constant deterioration rate and retailers' ordering lead time was assumed as zero. Improved GA was developed to analyze the considered problem. Shukla, Tiwaric, and Ceglarek (2013) introduced algorithm portfolios based on evolutionary algorithms including GA, age genetic algorithm, memetic algorithm, sexual genetic algorithm, and GA with chromosome differentiation. All evolutionary algorithms were tuned within their parameter limits to initiate the experiments. Different combinations of mutation and crossover probabilities were tried to determine the best parameter settings. In the study, set of IRP instances were generated considering number of sales-points, number of homogeneous vehicles in the fleet, and different demand rates.

Shaabani and Kamalabadi (2016) used a CPLEX solver, branch-and-cut, and population-based simulated annealing (PBSA) to solve the IRP for a bi-level supply chain. The parameters of PBSA were empirically defined based on the minimal cost criterion after solving various test problems. In the study, proposed system used perishable multi-products that have a fixed lifetime. No stockout was allowed and split deliveries were not permitted. Dabiri, Tarokh, and Alinaghian (2017) presented a new transport cost calculation pattern. In the study, homogeneous fleet was used and shortages were not allowed. In addition, the daily customer demands were assumed as predetermined. The multi-objective particle swarm optimization and augmented ε-constraint method were compared to demonstrate the practical utilization of the proposed methodology.

Azadeh, Elahi, Farahani, and Nasirian (2017) used the GA to solve the IRP for transshipment of perishable product. In the paper, GA parameters were determined based on the Taguchi method. It was assumed that backlogging was not permitted and only one vehicle was used by supplier. To meet customer demands, the maximum level policy was used. Gholamian and Heydari (2017) proposed a hybrid method in which simulated annealing was used to analyze the location routing model and GA was employed to solve the inventory problem. The parameters of method were adjusted based on the Taguchi method. In the study, (S-1, S) inventory policy was considered with stochastic demand and stochastic lead time. It was assumed that manufacturer has no limit in manufacturing the product.

Rayat, Musavi, and Bozorgi-Amiri (2017) developed a reliable distribution network considering multi-product and multi-period supply chain. In the study, the risk of probabilistic facility disruption was considered and continuous review inventory control policy was applied with partial backordering. Multi-objective simulated annealing was employed to resolve the problem. The results of proposed method were also compared with exact solution method and three other meta-heuristics. Timajchi, Al-e-Hashem, and Rekik (2019) presented a transshipment enabled IRP considering the minimization of the accident risk and loss. In the study, hybrid GA was developed to solve the multi-objective problem. They performed three sensitivity analyses including transshipment option, deterioration rate, and unit transportation costs. In addition, the effect of the backorder ratio was analyzed.

Jafarkhan and Yaghoubi (2018) presented a robust IRP with the two flexibility instruments of transshipment and substitution. In the study, efficient solution method was developed and its results were compared with adaptive large neighborhood search heuristic algorithm. Under supply and demand uncertainties, sensitivity analyses were also employed considering the different states of transshipment and substitution. Ramadhan, Imran, and Rizana (2018) proposed the record-to-record travel algorithm to analyze the IRP. In the study, multi-products were used and transported with capacitated homogenous vehicle. The fleet length limit and maximum number of stop were considered as constraints. It was assumed that no stockout was occurred. Balamurugan, Karunamoorthy, Arunkumar, and Santhosh (2018) proposed an artificial immune system algorithm to solve the IRP while minimizing the amount of CO_2 emission based on the total travel distance. In the paper, the demands of the production units were known and each production unit demand was fulfilled directly from the suppliers using homogeneous vehicles. Wang, Tao, and Shi (2018) proposed an improved adaptive GA integrated with greedy algorithm to analyze the low-carbon IRP. The proposed model considered carbon emissions in the optimization. The model was evaluated according to the results of the different maximum load capacity and carbon tax values.

2.1.4. Simulation Based Methods

Simulation can be used to dynamically perform inventory control and routing decisions. It provides modeling and artificially reproducing complex IRP systems in a natural way. Reiman, Rubio, and Wein (1999) presented the operational aspects of the IRP in a dynamic stochastic setting. Fixed routing and dynamic routing were analyzed as two types of the IRP. In the study, a single capacitated vehicle was used to serve a single warehouse and a finite set of retailers. Unsatisfied demand was backordered. The aim was to minimize costs including transportation costs, holding cost, and backordering cost. Four simulation experiments were performed to evaluate various aspects of system performance.

Sayarshad and Gao (2018) presented a dynamic inventory routing and pricing method. In the study, non-myopic strategy was proposed under the elastic demand. To calculate social welfare, customer waiting, inventory holding, delivery, and lost-sales costs were considered. Furthermore, simulation was employed to evaluate the model in terms of fleet design, tour length, and inventory costs. Mes, Schutten, and Rivera (2014) used the SO to solve dynamic and stochastic IRP with order-up-to level inventory policy using back-ordering and homogeneous vehicles. In the study, optimal learning techniques were utilized to tune the parameters. In addition, record observations and waste deposits were analyzed using simulation.

Lee, Cho, and Jung (2014) used simulation to analyze the performance of the proposed system including the vehicle routing and the inventory management. The results showed that the proposed system demonstrated a better performance than the existing system when the transportation cost, the transportation efficiency, and the stock level were taken into account. Juan, Grasman, Cáceres-Cruz, and Bektaş (2014) developed a hybrid solution model combining Monte Carlo simulation with heuristics to solve a single-period IRP. In the study, homogeneous vehicles were used to transport products. Supply chain included a single distribution depot and multiple retail centers with stochastic demands. In addition, different refill policies were used to determine the optimal inventory level. Jarugumilli, Grasman, and Ramakrishnan (2006) formulated the sequential decision problem as a stochastic dynamic program. Modified A* algorithm and simulation were employed to manage inventory routing. In the study, periodic review inventory control system was utilized in proposed supply chain including a single warehouse and multiple retailers.

Abdollahi, Arvan, Omidvar, and Ameri (2014) used a new SO procedure to IRP that considered one or a combination of some risk factors such as disruption risks, stock-out risks, and cost risk. In the study, the maximum level was used as the inventory policy and unmet demand was backordered. Cáceres-Cruz, Juan, Bektas, Grasman, and Faulin (2012) used simulation with a multi-start asymmetric randomization to analyze a single-period IRP under stochastic demands. Five different service policies were applied for each customer. Gruler et al. (2020) used simheuristic algorithm that considered random realizations of customer demands. In addition, proposed algorithm integrated Monte Carlo simulation into a variable neighborhood search meta-heuristic to solve the multi-period IRP that consists of three different stages. The paper extended the work of Juan et al. (2014).

2.2. Demand Forecasting in Supply Chain

Accurate demand forecasting provides valuable information about local and global markets. Companies can make informed decisions about the markets potential. In addition, demand forecasting gives information about business growth and pricing strategies. It helps companies reduce risks involved in business activities and cope with uncontrollable and competitive forces.

In literature, different demand patterns are used. We summarized some of papers associated with forecasting of different demand patterns. Silver, Ho, and Deemer (1971) developed a practical procedure for controlling an item having an erratic demand pattern of a particular nature, namely transactions with a geometric probability distribution. Pierre, Vannieuwenhuyse, Dominanta, and Dessel (2003) presented a dynamic variant of the ABC storage policy for today's erratic demand environments. Zhang, Bao, and Zhang (2010) presented a “decomposition-and-ensemble” principle for erratic demand forecasting. Support vector machine (SVM) was used to model and formulate the erratic demand series. Onyeocha, Khoury, and Geraghty (2015) investigated a four-product five-stage system characterized by erratic demand and dynamic priority control of inventory release from finite capacity buffers based on demand performance, minimum run/batch quantity, a weekday product family priority rule, and finite capacity buffers. Molina, Ponte, Parreño, Fuente, and Costas (2016) showed the comparison of an autoregressive integrated moving average (ARIMA) models and ANNs for forecasting the demand medicines with an erratic nature. In this study, the demand of two drugs in a public hospital was analyzed. Jiang, Liu, Huang, and Yuan (2017) proposed an adaptive autoregressive SVM model that was tested by erratic demands of heavy truck spare parts.

Willemain, Smart, and Schwarz (2004) created a new bootstrapping approach to forecast the intermittent demands over a fixed lead time. Shale, Boylan, and Johnston (2006) presented the improvement in forecasting performance for intermittent demand items. Hua, Zhang, Yang, and Tan (2007) developed a new approach for forecasting the intermittent demand of spare parts. Two types of performance measures were also presented to assess forecasting methods. Teunter and Duncan (2009) presented a method for forecasting intermittent demand. The results showed that Croston's method and bootstrapping outperformed moving average and single exponential smoothing. Kocer (2013) proposed a modified Markov chain model to forecast intermittent demand. Kourentzes (2013) proposed the ANN method for intermittent demand time series forecasting. Proposed method allowed interactions between the interdemand intervals and the demand of intermittent items. Prestwich, Rossi, Tarim, and Hnich (2014) used several intermittent demand patterns and proposed several new error measures with almost no infinities, and with correct forecaster ranking. Ramaekers and Janssens (2014) used the SO to develop a framework for inventory management decision support of intermittent demand. In the study, the simulation model that was developed in Microsoft Excel spreadsheets aimed to determine the optimal inventory system. A demand forecasting and an order decision were made at each review-time. Lei, Li, and Tan (2016) presented a new forecasting algorithm using the material intermittent demand data. Then, the results were compared with the exponential smoothing. Cheng, Chiang, and Chen (2016) employed the ARIMA and Croston’s method for forecasting the intermittent demand of medical supplies. Jung, Park, Kim, and Kim (2017) presented a new bootstrap method that takes into consideration the unique characteristics of intermittent demand to improve forecasting performance. They deployed a simple experiment that utilized artificial data to compare the results of suggested method and the conventional Markov bootstrap method. Dastidar (2017) developed the method to improve forecasting intermittent long-tail demand for a global power-generation business. Lolli et al. (2017) used the ANNs trained by back-propagation and extreme learning machines. In addition, different input patterns and architectures were used to compare the forecasting results of proposed ANN methods for intermittent demand. Li and Lim (2018) proposed a greedy aggregation–decomposition method in order to forecast the intermittent demand. Kaya and Turkyilmaz (2018) employed ANNs, support vector regressions, and decision tree techniques. In the paper, the intermittent package of R software was employed to produce artificial demand data.

Verganti (1997) investigated the demand management mechanisms by means of quantitative analyses. The ability of order overplanning was also explored to cope with uncertain lumpiness. Bartezzaghi, Verganti, and Zotteri (1999) analyzed the behavior of forecasting techniques, especially exponentially weighted moving average, early sales, and order overplanning, under the lumpy demand pattern. Feng, Yamashiro, and Zhang (2003) developed an optimal batch size and raw material ordering policy. In the study, a two-stage production system was developed with a fixed-interval and lumpy demand delivery system for minimizing total production cost. Gutierrez, Solis, and Bendore (2004) presented a lumpiness factor and a coefficient of skewness as two measures to characterize the lumpy demand. Exponential smoothing, ANNs, and Croston’s method were utilized to actual lumpy demand data of an electronics distributor. Miragliotta and Staudacher (2004) presented an advanced lumpiness management approach to manage the uncertain lumpy demand. Hollier, Mak, and Yiu (2005) presented the characteristics of inventory system for lumpy demand. In the study, a modified (s, S) model was used with a maximum issue quantity restriction and a critical inventory position. Amin-Naseri and Tabar (2008) used the recurrent ANN for lumpy demand forecasting of spare parts. Croston's method and Syntetos & Boylan approximation were also used to evaluate the proposed method. Huang and Ye (2010) analyzed the lumpy demand at the suppliers and the shipment size constraint under vendor managed inventory systems. Mukhopadhyay, Solis, and Gutierrez (2012) developed an optimally weighted moving average method and an intelligent pattern-seeking method to forecast lumpy demand. Berling and Marklund (2013) presented an approximation model to optimize reorder points in one-warehouse N-retailer inventory systems under highly variable lumpy demand pattern.

Demand forecasting is very important, both in academic and business practice. It has bothered managers and scholars for a long time. The forecasting of demand can be considered as the primary revenue source for the company since all departments set themselves with respect to results obtained from demand forecasting. Traditional forecasting methods can be successfully used when the demand of item is smooth and continuous. However, forecasting is hardly difficult with traditional methods when the demand of item has changing values (Kaya and Turkyilmaz, 2018). Furthermore, traditional methods may not capture the nonlinear pattern in demand data. At this point, ANN modeling is a logical choice to handle these limitations (Amin-Naseri and Tabar, 2008). ANNs can capture interactions between the non-zero demand and the inter-arrival rate of demand events (Kourentzes, 2013).

2.3. Big Data and Sentiment Analysis

In recent years, social media users have grown exponentially. This growth has evoked researchers and practitioners on how to analyze and employ review creatively. People share their sentiments toward products or services with other users by means of reviews. Hence, the future sales performance of products has significantly influenced by the quality of the reviews and sentiments stated in the reviews (Balaji, 2015). Therefore, the customers' review can be integrated with predictive models since they can have a critical influence on sales. On the other hand, review or sentiment is not taken into account by the traditional models.

We firstly analyzed the studies related to sentiment analysis and sales. For example, Yu, Liu, Huang, and An (2012) proposed an autoregressive sentiment aware (ARSA) model and also incorporated the quality factor in ARSA for product sales prediction. Furthermore, probabilistic model was used to analyze sentiments in reviews. The pattern of reviews and its relationship were investigated using a real example from the movie sector. The results demonstrated that proposed methods give a remarkable prediction performance when compared with the autoregressive model. Ahn and Spangler (2014) analyzed the performance of sales forecasting considering social media-based features. ARIMA was used to model monthly automobile sales. In the study, a semantic dictionary-based approach was used to calculate the sentiment score.

Gaikar and Marakarkandy (2015) determined the factors responsible for purchase intention of the product as a Hindi movie. In the study, tweets were extracted about the movie from twitter using an Application Programming Interface. Sentiment analysis tool was then employed to determine positive and negative tweets. The impact of valence and volume of online reviews were analyzed by means of sentiment analysis on attitude and purchase intention of a product. Schneider and Gupta (2016) used a random projections model to reduce the dimensionality of the bag-of-words in the sales forecasting model. In addition, SVM was employed to compare the results of the proposed model.

Fan, Che, and Chen (2017) combined the Bass and Norton model with the extracted sentiment index for product sales forecasting. Naïve Bayes, SVM, and k-nearest neighbors were utilized to classify online reviews. To evaluate the performance of the proposed model, real-world automotive industry data was employed. Lau, Zhang, and Xu (2018) presented big data analytics methodology that included a parallel aspect-oriented sentiment analysis. In addition, a parallel co-evolutionary extreme learning machine was applied to sentiment based sales forecasting. An aspect-oriented sentiment analysis method was employed to mine context-sensitive sentiments.

Pai and Liu (2018) used both time series models and multivariate regression models to forecast monthly total vehicle sales. Two categories of time series data were utilized to forecast the sales. First, historical monthly total vehicle sales were used as the dataset while second category included sentiment scores of tweets and two stock market values. In the study, deseasonalizing procedures were also applied to forecast the sales.

Sentiment analysis is one of the fastest growing research areas in computer science, making it challenge to keep track of all the activities in the area. In recent years, sentiment analysis has shifted from analyzing online product reviews to social media texts from Twitter and Facebook. Many topics beyond product reviews like stock markets, elections, disasters, medicine, software development and cyberbullying extend the utilization of sentiment analysis.

Zainuddin and Selamat (2014) used the SVM to classify the datasets as positives or negatives. Three different weighting scheme including the entire corpus, binary occurrence, and term occurrence were employed to generate the word vectors. The results demonstrated that chi-squared statistics can help to reduce the dimensionality and the noise in the text. Jadav and Vaghela (2016) used the optimized SVM in which the value of kernel hyper parameters was changed. To compare the performance of the proposed optimized SVM, Naïve Bayes, and SVM were also employed. In the study, unstructured movie review was firstly converted into structured form and then score was calculated to be given as input for proposed methods.

Dey, Chakraborty, Biswas, Bose, and Tiwari (2016) used two supervised machine learning algorithms including Naïve Bayes and k-nearest neighbor. In the paper, the movie reviews and hotel reviews were utilized in proposed algorithms. The results of the study showed that both algorithms can be used to classify the movie reviews and hotel reviews. Kumar, Desai, and Majumdar (2016) classified the review as positive and negative review by means of Naïve Bayes classifier, Logistic Regression, and SentiWordNet algorithm. In the study, Amazon website was used to obtain the reviews. Swain and Cao (2019) analyzed the effect of social media content produced by supply chain members on supply chain performance. The sentiment analysis was applied by means of the Naïve Bayes to test proposed hypotheses. The findings showed that social media content can be helpful for information sharing and collaboration within the supply chain.

Raghuwanshi and Pawar (2017) used the linear and probabilistic classification models. In the study, the performance of each classification model was compared. One of the open source machine learning software package in Python that is scikit–learn was used in the study as software. Singla, Randhawa, and Jain (2017) used unstructured data of mobile phone reviews that have been extracted from Amazon. In the study, Naïve Bayes, SVM, and decision tree were used as classification models. In addition, undersampling technique was applied for treating imbalanced data. Khader, Awajan, and Al-Naymat (2018) reviewed the MapReduce based sentiment analysis techniques. In the study, subjectivity/objectivity, polarity, sentiment analysis levels, features selection, and sentiment classification were explained related to sentiment analysis. The characteristics of big data create new challenges into the sentiment analysis. Several researches proposed to enhance the efficiency of sentiment analysis extraction using big data frameworks (Khader et al., 2018).

While much research work has been devoted to supply chain management and demand forecast, research on designing big data analytics methodologies to enhance sales forecasting is seldom reported in existing literature. The big data of consumer-contributed product comments on online social media provide management with unprecedented opportunities to leverage collective consumer intelligence for enhancing supply chain management in general (Lau et al., 2018). Tiwari, Wee, and Daryanto (2018) presented methodologies and applications on the supply chain that utilized big data analytics for decision-making. In the study, the viewpoint of future big data application was also presented in supply chain management.

Arunachalam, Kumar, and Kawalek (2018) presented a systematic literature review to analyze big data analytics in supply chain. In the study, journal papers and maturity models were determined using appropriate key words. The capabilities of big data analytics were defined and the challenges and issues of practicing big data analytics were addressed. Mishra, Gunasekaran, Papadopoulos, and Childe (2018) used two steps to perform data analysis for big data and supply chain management. Firstly, BibExcel software was used to implement bibliometric analysis. Then, Gephi was employed to conduct network analysis.

Although big data and applications are of great interest to both academicians and practitioners, the number of studies on big data analysis and supply chain management is still scarce. Wamba and Akter (2015) presented a detailed literature review on big data analysis and supply chain management. There are still many unanswered questions about how big data should be analyzed for supply chain management. Brinch, Stentoft, and Jensen (2017) discussed the terminology of big data and its applications in procurement, manufacturing, service, logistics, planning, and recycling processes. Big data provides an integrated view of operational performance and customer interaction and provides real-time information that helps you make critical decisions. Wang and Alexander (2015) presented the applications and opportunities of big data in supply chain management and business administration. In addition, they discussed general challenges related to supply chain management and big data. Groves, Collins, Gini, and Ketter (2014) highlighted the advantages of using a competitive simulation environment. Also, key performance indicators that are specifically relevant to autonomous supply chain processes were presented. Robak, Franczyk, and Robak (2014) presented the utilization of big data in conjunction with data science and predictive analysis for logistics and supply chains. Rozados and Tjahjono (2014) discussed the big data and supply chain. As a result of the study, it has been found that the sources of supply chain management are generally unstructured formats that are difficult to analyze with traditional information technology tools. Chae (2015) proposed the framework to extract relevant intelligence in supply chain. Twitter Analytic that focuses on different dimensions in the analysis of Twitter data was used to analyze the supply chain. The results showed that wording in the entire tweets was similar to that of hashtags. Wang, Gunasekaran, Ngai, and Papadopoulos (2016) presented big data analytics and classified the applications of big data related to logistics and supply chain management. Biswas and Sen (2016) presented a big data system for data analytics and management. The role of supply chain analytics in SCM was also presented. In the study, the contextual significance of analytics was emphasized. Hofmann (2017) incorporated big data applications in supply chain decisions. In the study, bullwhip simulations were investigated to reduce inefficiencies in the supply chain. Supply chain consisted of a single retailer and a single manufacturer. Simple exponential smoothing was used to forecast demand. Kaur and Singh (2018) used t-test to conduct for statistical significance between heuristic and optimal solutions. In the study, deterministic demands were considered under multi-period supply chain which consisted of multi-supplier and multi-carrier. In addition, late deliveries and shortages were not allowed. Andersson and Jonsson (2018) identified how the demand planning process was impacted by proposed interventions. Demand planning and demand plan performance are critical for aftermarket performance. Therefore, the effect of causal-based aftermarket demand planning was also presented. Zhan and Tan (2020) presented the analytic infrastructure to support firms in harvesting big data and expand their competence sets. Five main stages were used to develop the proposed analytic infrastructure including data capture and management; data cleaning and integration; data analytics; competence set analysis–deduction graph; information interpretation and decision making.

3. MATERIAL AND METHOD

3.1. Material

The data that used in this paper is taken from Chen, Sain, and Guo (2012). In the study, online retailer is a small business and a relatively new entrant to the online retail sector. The retail company was set up in 1981. It is a UK-based and registered non-store business with some 80 members of staff. For years in the past, direct mailing catalogues were used in the company. Also, orders were taken over phone calls. Then, web site was created and company shifted completely to the Web. The company collected a huge amount of data about many customers from all parts of the United Kingdom and Europe. In addition, the company utilizes Amazon.co.uk to market and sells its products. The customer transaction dataset held by the merchant has 11 variables and it contains all the transactions occurring in years 2010 and 2011. Variables in the customer transaction dataset includes Invoice, StockCode, Description, Quantity, Price, InvoiceDate, Address Line 1, Address Line 2, Address Line 3, PostCode, and Country.

In this study, data is adapted considering three different uneven demand patterns. In erratic demand, average inter-demand interval (ADI) is less than 1.32 and squared coefficient of variation (CV^2) is greater than 0.49. The mean of customer order quantity in erratic demand is 12 units. In intermittent demand, ADI is greater than 1.32 and CV^2 is less than 0.49. The mean of customer order quantity in intermittent demand is 368 units. In lumpy demand, ADI is greater than 1.32, CV^2 is greater than 0.49. The mean of customer order quantity in lumpy demand is 368 units.

In order to ensure the volume and variety of the big data, Amazon reviews are used as customer reviews. All customer reviews are filtered out and then the text of customer reviews for one year are stored as dataset. For companies, customer reviews are important since they provide a clearer understanding of customers' demand and interests. Furthermore, they basically provide other customers to leave their feedback on product. Customers cannot physically inspect the product in shopping online. At this point, customer reviews help customers understand almost every detail of the product.

3.2. Simulation Optimization

SO can be described as the optimization of performance metrics based on outputs that are obtained from stochastic simulations. Even though various SO based articles are available in literature, SO is still in its infancy (Ghiani, Legato, Musmanno, and Vocaturo, 2004). SO has contributed to enable improved and more focused supply chain decisions. It also allows manager to extract the most meaningful value from the supply chain. The design of interaction is crucial in SO. The possibilities of integrating simulation and optimization are so vast. Therefore, we should have good overview related to the different approaches. The analysis of advantages and limitations is important to determine the appropriate approaches in SO (Figueira and Almada-Lobo, 2014).

Figueira and Almada-Lobo (2014) classified SO considering key dimensions: simulation purpose, hierarchical structure, search method, and search scheme. The interaction between optimization and simulation was addressed in simulation purpose and hierarchical structure. The search algorithm design was explained in search method and search scheme. In addition to these four dimensions, solution evaluation, analytical model enhancement, and solution generation approaches were used to categorize SO methods. Solution evaluation approaches can be considered as the optimization of a simulation model. In solution evaluation approaches, solution performance is evaluated by simulation. In solution generation approaches, analytical models can be formulated and solved. Also, their solutions are simulated to calculate all the variables of interest. In solution generation approaches, some variables are computed by simulation that is the part of the whole solution generation. Analytical model enhancement approaches are employed to improve the analytical model using the simulation results. Detail about the classification of SO methods can be found in Figueira and Almada-Lobo (2014).

In Carson and Maria (1997), SO methods classified into gradient based search methods, stochastic optimization, response surface methodology, heuristic methods, A-teams, and statistical methods. Gradient based search methods estimate response function gradient to evaluate the shape of the objective function. In there, deterministic mathematical programming is employed. In stochastic optimization, a local optimum is found for an objective function. The value of objective function is not known analytically but it can be estimated. In response surface methodology, a series of regression models is fitted to the output variable of a simulation model and the results of regression function are optimized. Heuristic methods are widely used in SO methods. Many of heuristic methods balance exploration with exploitation. In addition, they provide efficient global search strategies. A-teams is a process that includes the integrating of various problem solving strategies. Statistical methods use statistical analysis that includes data collection, data analysis, underlying probability distributions, the relationship among the alternatives, the performance measure of interest, and etc (Carson and Maria, 1997).

In SO, different combinations of values are tried to provide the most desirable output for variables (Ammeri et al., 2011). In proposed SO, the loss function ($L(\theta)$) is minimized.

$$L(\theta) = E[Q(\theta, V)] \qquad \theta \in \Theta \tag{3.1}$$

where V denotes the amalgamation of the random effects in the simulation. θ represents the vector of terms being optimized. Θ is the domain of allowable values for vector θ. Q is a sample realization of the loss function calculated from running the simulation (Spall, 2003). SO generates solutions from the search space. SO is terminated until some stopping criteria are satisfied.

3.2.1. Simulation

Simulation simulates the operations of real world systems. It generates an artificial history related to the behavior of the system over time. In this manner, system can be observed and “what if” questions about the system can be answered (Chen and Lee, 2011). Although different categorization of simulations is available in literature, it can be categorized as static vs. dynamic; continuous-change vs. discrete change; and deterministic vs. stochastic (Kelton, Smith, Sturrock, Göçken, and Dosdoğru, 2015). In literature, different types of simulation are used in supply chain management. Tarokh and Golkar (2006) classified supply chain simulation methods into four types: spreadsheet simulation, system dynamics, discrete-event dynamic systems simulation, and business games.

1) Spreadsheet Simulation: Spreadsheet software increases the popularity of corporate modeling. Managers can easily use this type of simulation. However, spreadsheet simulation is static model. Therefore, defining time based parameters is difficult. In this type of simulation, file formats are standardized. The model is created by means of spreadsheet that also does the sampling and performs the computations. The table structure in a spreadsheet obtains the user to organize the computations and results (Tarokh and Golkar, 2006).

2) System Dynamics: System dynamics is continuous model. To imitate interactions between diverse parts of the system, differential equations are used. In this type of simulation, the system structure and the governing rules are constant (Seleim, Azab, and AlGeddawy, 2012).

The feedback principle plays critical role in system dynamics. The managerial target is compared with its realization. If it involves any undesirable deviation, corrective action should be taken. System dynamics is developed as an influence diagram. A change of strategy can be expressed as a change in this picture. Alternatively, only the target value of the current performance metric can be changed (Kleijnen and Smits, 2003).

3) Discrete-Event Dynamic Systems Simulation: Discrete-event dynamic systems simulation is more comprehensive than spreadsheet simulation and system dynamics. It denotes individual events such as customer arrivals. Furthermore, discrete-event dynamic systems simulation includes uncertainties related to proposed system such as random breakdown, random arrivals, and etc (Kleijnen and Smits, 2003). In proposed model, discrete entities vary state as events take place in the simulation. The state of the model varies only when those events take place (Tarokh and Golkar, 2006).

4) Business Games: Technological and economic processes can be easily simulated. On the other hand, human behavior modeling is difficult. Hence, user should operate within the simulated ‘world’. At this point, business games can be used. In literature, strategic and operational games are used as two subtypes of business games. Details about the business games can be found in Kleijnen and Smits (2003).

3.2.1.1. Simulation and Modeling

Model formation is to create a structure that captures interesting or noteworthy features and processes of the research subject. A model is a simplified definition of the system via a computerized representation that allows us to conduct virtual experiments on the behavior of a system. Basically, computer simulation is the imitation of some processes that determine the characteristics that need to be considered such as the generated output and the interactions with the environment. A system model is an abstract expression of the reality except for some details of environment. The model, which is operated by the simulation mechanism, contains some procedure, rule, equation or constraint sequences according to the input and output behavior. Simulation modeling is an extremely broad topic that can cover almost all disciplines (Cros et al., 2006). In addition, simulation is remarkable tool to visualize, understate, and analyze the system. On the other hand, modeling is one of the most important parts of a simulation study. The following steps are used in simulation modeling (Maria, 1997).

Step 1. The system problems are identified and requirements are set for a proposed system.

Step 2. The problem is formulated. The objective of the study, constraints, performance measures, and other specific parameters are determined. The time frame of the model is also defined.

Step 3. Data is collected related to existing system. The sources of randomness are identified. Suitable input probability distribution is selected for stochastic input variables. At this point, Software packages such as ExpertFit and BestFit can be used for distribution fitting. They also provide goodness-of-fit tests and parameter estimation.

Step 4. A model is formulated. Schematics and network diagrams of the system are developed. The conceptual models are translated into simulation software in acceptable form. Verification is used to evaluate the execution model.

Step 5. The model is validated. The performance of model is analyzed under known conditions. Statistical inference tests and confidence assessment are performed.

Step 6. The model is documented considering input variables, objectives, and assumptions.

Step 7. The suitable experimental design is selected and applied.

Step 8. The experimental conditions are created for runs.

Step 9. Simulation runs are performed in accordance with steps 7-8 above.

Step 10. The results of simulation are presented and interpreted. Numerical estimates of the desired performance measure are computed for each configuration of interest. Hypotheses related to system performance are tested. The graphical analysis of the output data is constructed.

Step 11. Further recommendations are made about system.

The benefits of simulation modeling and analysis can be summarized as follows (Maria, 1997):

- It provides a better understanding of the system and allows to observe the system's operation in detail.
- Hypotheses are tested considering system feasibility.
- Users can compress time or expand time according to the properties of considered phenomena.
- The effects of changes on the system can be studied without disrupting the real system.
- New or unknown situations can be experimented.
- The "driving" variables and the inter-relationships among them are identified.
- Bottlenecks are identified.
- Various performance measures are used to analyze the system structure.
- A system approach is employed to solve problem.
- Robust well designed systems are developed and system development time is reduced (Maria, 1997).

3.2.1.2. Simulation Software (SIMIO)

Simio is designed to simplify the simulation modeling. Simio provides shift from the process orientation to an object orientation. It enables user to easily employ the object orientation. In addition, execution can be efficiently made. Note that multiple modeling paradigms that include an event orientation and process orientation can also be supported by Simio. Furthermore, both discrete and continuous systems are fully supported by Simio (Pegden, 2007).

The modeling framework in Simio is based on intelligent objects that are created by user and can be reused in various modeling projects. Libraries are used to store the objects. In addition, it can be easily shared. Pre-built objects may be preferred by beginning modelers that can also easily create their own intelligent objects. The objects are combined to create a model. Simio model resembles the real system. Every model is automatically creating block that may be employed in creating higher level models (Pegden, 2007).

The properties of Simio are summarized as follows (Pegden and Sturrock, 2010).

1. Its framework is a graphical object-oriented modeling framework.
2. It assists 3D animation as a natural part of the modeling process.
3. It allows objects to be built. New objects with complex behavior are also created.
4. The multiple modeling paradigms, discrete and continuous systems are assisted by Simio. In addition, it provides an event, object, process, and agent modeling view.
5. It enables specialized properties to directly assist applications.

The same basic principles are used in Simio and object oriented programming languages. On the other hand, these principles in Simio are implemented within a modeling framework and not a programming framework (Pegden, 2007). Simio enables additional capabilities for advanced users. Some programming background such as Visual Basic, C++, and C# can be used to increase the performance of Simio. For example, specific add-ons can be added according to the requirement. A new dynamic selection rules can be added. To design and run experiments, custom algorithms can be added. To import and export tables, custom routines are added. Custom steps and elements are added to improve the capabilities (Kelton et al., 2015).

3.2.2. Optimization Methods

Optimization can be broadly defined as the process of determining the values to be taken by decision variables in order to optimize the value of a defined objective function under certain constraints in a system. Thus, it is the process of determining the input(s) in order to obtain the desired output. In the study, GA and NSGA-II were used as an optimization algorithm.

GA is a population based stochastic algorithm. Optimization with the GA is very useful in many areas. GA is a research technique that aims at the regulation of self-learning and decision-making systems by using random research methods in the light of evolutionary approach principles (Elmas, 2016). GA method is effective in optimizing the supply chain management (Othman et al., 2016).

GA takes the process of living things in nature. GA is based on adaptation of new individuals born from mother and father individuals to the conditions. In addition, it based on the survival of their lives. New individuals may have either good genes taken from their parents or bad genes. In this case, individuals with bad genes will not survive. One of the key features of the GA is that it searches for a solution on a group and therefore selects the best from a large number of solutions (Elmas, 2016). The working principle of the GA is summarized as follows (Elmas, 2016).

1. It starts by determining the number of individuals in the population. There is no definite number for this number of individuals.
2. The function that finds how good the chromosome is called fitness function. The fitness function decodes the chromosomes and makes them the parameters of the problem. According to these parameters, fitness values are calculated and chromosome fitness is found. The success of the GA depends on the efficiency and precision of this function.

3. The chromosomes are evaluated according to the chromosome fitness. Selection methods such as roulette wheel selection and tournament selection are used to make this selection.
4. Crossover and mutation operators are considered as key operator for the GA. The crossover is simply the displacement of the determined parts between the two individuals. Mutation is defined as the external replacement of a part of the same individual. The low mutation probability may cause some features to be lost in the population. This is an obstacle to find the best solution. The possibility of a high mutation probability may result in disruption of the solutions. Therefore, the probability of changing for mutation probability is generally selected in the range of 0.1% to 15%. To ensure gene diversity, the possibility of crossover probability is generally selected in the range of 60% to 90%.
5. As a result of these changes, the information found in the chromosome structures must be the same as the information in the first generation. In order to do this, after the application of the crossover and mutation operators, the repair operator may be required depending on the type of problem. By applying the repair operator, it is ensured that the current chromosome information is preserved.
6. By removing old chromosomes, a new fixed population size is provided.
7. The best individual in the current solution set is transferred to the next population.
8. The success of the new population is found.
9. The GA is run repeatedly until the determined cycle or stop criteria is met.
10. As a result of the operation of the GA, the best individual is taken as a solution (Elmas, 2016).

For each supply chain member in proposed model, the chromosome of GA is considered to be composed of four parts: (1) the number of neurons in first hidden layer; (2) the number of neurons in second hidden layer; (3) training algorithm; and (4) the de-noising degree. In proposed GA, population size is 20, crossover probability is 0.8, and mutation probability is 0.05. In GA, tournament selection is used as selection method. GA is coded by using C# (Visual Studio Community 2017).

NSGA-II introduced a fast nondominated sorting procedure with $O(MN^2)$ computational complexity, an elitist-preserving approach, and a parameterless niching operator for diversity preservation (crowded comparison operator) (Sarkar and Modak, 2005). In NSGA-II, the population is initialized considering the problem range and constraint. Non dominated sorting is applied. The value of crowding distance is calculated. Genetic operators are applied to individuals. Finally, offspring population and current generation population are combined to create the new generation. Details about NSGA-II can be found in Yusoff, Ngadiman, and Zain (2011).

For each supply chain member in proposed model, the chromosome of NSGA-II is considered to be composed of three parts: (1) the initial inventory; (2) the reorder point; and (3) the order-up-to level. In our proposed NSGA-II, population size is 20, crossover probability is 0.8, and mutation probability is 0.05. In NSGA-II, tournament selection is used as selection method. NSGA-II is coded by using C# (Visual Studio Community 2017).

3.3. AI and AI Types

AI is related to intelligent behavior in artifacts. Intelligent behavior includes perception and reasoning in complex environments. Moreover, it learns, communicates, and acts in complex environments (Nilsson, 1998). AI can be successfully applied in many sectors. For example, manufacturing sector involves soft production, intelligent design, agile production, virtual production, intelligent workflow, and etc. The application of AI considerably increases production efficiency.

Wirth (2018) presented the different form of AI including narrow AI, hybrid AI, and strong AI.

Narrow AI: Weak or narrow AI is created to a specific problem or task. It can be re-trained and/or modified to deal with other challenges. Narrow AI based system lacks the flexibility of human intelligence. Although Narrow AI based system falls short in terms of the scope of components, they may be very robust in its domain (Wirth, 2018).

Strong AI: Strong AI can also be denoted as Artificial General Intelligence (AGI) and full AI. AGI based system is as flexible and robust as human intelligence. It is created to a specific problem or task. AGI has not been obtained (yet) (Wirth, 2018).

Hybrid AI: Hybrid AI can be considered between the narrow AI and AGI. Due to the new extremely powerful AI systems, narrow AI grows rapidly. To cope with new challenges, multiple narrow AI solutions are also blended. Multiple narrow AI is still not AGI. However, it is more powerful than narrow AI (Wirth, 2018).

Wong, Guo, and Leung (2013) used two categories to classify AI techniques. First category is symbolic AI that focuses on development of knowledge-based systems. Second category is computational intelligence that focuses on development of a set of nature-inspired computational approaches (Wong et al., 2013). Acquisition, accumulation, representation, and use of knowledge are important to construct a knowledge-based system. In this system, three core components are available from the perspective of the end user. First, knowledge base includes highly specialized and problem-related knowledge. Second component is the knowledge inference mechanism that ensures solution recommendations for users. Final component is the user interface that bridges the gap between end users and the system (Wong et al., 2013).

Computational intelligence includes evolutionary computations, ANNs, and fuzzy logic systems. Evolutionary computation is mainly inspired by optimumseeking mechanisms from the real world. In evolutionary optimization, few or no assumptions are used to solve the problem. Due to the large solution spaces and randomness, evolutionary computation is remarkable method for complex optimization problems. ANN is a computational model inspired by research into biological neural networks. ANN includes a number of interconnected neurons. Fuzzy logic serves mainly as a system of concepts, principles, and methods for handling modes of reasoning with imprecise information (Wong et al., 2013).

3.3.1. Artificial Neural Network

ANNs are inspired by the human brain. They use features similar to some organizational principles of the human brain. They are parallel and distributed information processing structures that consisting of process elements each having their own memory. ANNs are self-learning devices that do not require a programmer's traditional abilities. In addition to learning, these networks have the ability to memorize and create relationships between information. ANNs are used successfully in processing uncertain, noisy, and incomplete information. ANNs are composed of a large number of interconnected elements. ANNs, like the human brain, have the ability to learn, remember, and generalize (Elmas, 2016). The basic ANN includes following elements (Elmas, 2016).

Inputs: Inputs bring the information from the environment to the neuron. The inputs may come from the previous neurons or from the outside world.

Weights: Weights are the appropriate coefficients that determine the effect of the inputs on the artificial neuron. Each input has its own weight. The large value of a weight means that it is strongly attached or important to the artificial neuron. On the other hand, small weight value means that it is weak or not important to the artificial neuron.

Summation function: In the summation function, each weight is multiplied by specified inputs and then threshold value is added.

Activity function: The result of the summation function is passed through the activity function and then transmitted to the output. The purpose of the activity function is to allow the output of the summation function to change. The neuron does not produce output when it is below the threshold level of the activity function. The neuron produces output when it is above the threshold level of the activity function.

Output function: The output is where the result of activity function is sent to the outside world or to other neuron.

All ANNs are derived from this basic structure. Differences in this structure provide different classification of ANNs. Details about the ANN can be found in Elmas (2016).

Simple ANN can be mathematically expressed considering the multilayer feedforward training network with one hidden layer is applied, the net input to unit k in the hidden layer (Zain, Haron, Qasem, and Sharif, 2012).

$$net_{hidden} = \sum_{j=1}^{J} C_{j,k} i_j + \theta_k \tag{3.2}$$

where J is the number of nodes of the input layer, $C_{j,k}$ is the weight between the input nodes and hidden nodes, i_j is the value of the inputs and θ_k is the biases on the hidden nodes. Consequently, the net input to unit z in the output layer is expressed in following equation.

$$net_{output} = \sum_{k=1}^{K} D_{k,z} h_k + \emptyset_z \tag{3.3}$$

where K denotes the number of nodes in the hidden layer, $D_{k,z}$ represents the weight between hidden and output nodes, h_k denotes the value of the output for hidden nodes, and $\emptyset_z$ denotes the biases on the output nodes.

The output for hidden nodes can be given as,

$$h_k = f(net_{hidden}) \tag{3.4}$$

In addition, the output for output nodes can be given as,

$$o_z = f\left(net_{output}\right) = R_a \tag{3.5}$$

where f is the transfer function to predict R_a value (Zain et al., 2012). Note that given mathematical formulation includes three layers in ANN (one input layer, one hidden layer, and output layer) but it can be easily revised for proposed ANN structure.

3.4. Proposed SO and AI Based Methodology

Demand forecasting, which has generated an increasing interest nowadays because of an increase in competition, is considered in supply chain. Demand forecasting is the main requirement for the competitive survival of business. It can also provide an insight into the expansion of decisions and capital investment.

Syntetos, Boylan, and Croston (2005) proposed a theoretically coherent scheme for categorizing demand into smooth, erratic, intermittent or lumpy. In this categorization scheme, the ADI and the CV^2 of demand are compared with cutoffs of 1.32 for ADI and 0.49 for CV^2. Details about the categorization can be found in Syntetos et al. (2005).

ADI: The average interval between two demands of the same ith item, usually expressed in periods as follows:

$$ADI_i = \frac{\sum_{n=1}^{N_{pi}} t_i^n}{N_{p_i}} \tag{3.6}$$

i denotes the item number, i=1, …I. N_{pi} represents the number of non-zero demand time interval of the ith. t_i^n represents the nth time interval of non-zero demand of the ith item.

CV= The standard deviation of the demand of the ith item, divided by the average demand d_i.

$$CV_i = \frac{\sqrt{\frac{\sum_{n=1}^{Np_i}\left(d_i^n - d_i\right)^2}{Np_i}}}{d_i} \tag{3.7}$$

where d_i denotes the average non-zero demand of the ith item and d_i^n represents the demand value at time n of the ith item. Details can be found in Costantino, Gravio, Patriarca, and Petrella (2018).

Proposed system including Phase 1 and Phase 2 allows controlling continuous changes of the method representing the current state of the inventory and routing system. In Phase 1, SO is used to optimize the initial inventory, reorder point, and order-up-to level in two echelon supply chain where five Distribution Center (DC)s and one supplier are available. Integrating optimization and simulation is highly demanding method in problem solving due to the need for computational power, even for today's standards. The interactions and uncertainties related to the proposed system can be easily captured by SO. However, the design of SO is crucial. The possibilities of combining simulation and optimization are so vast and hence a good overview of methods is necessary. In this study, NSGA-II is used as optimization method. Simulation model is developed by using Simio (Version: 7.121.12363). Processor is Intel (R) Core (TM) i7-8700K CPU @ 3.70 GHz and system type is 64 bit operating system.

In SO methods, total supply chain cost (TSCC) is minimized while the average service level is maximized. In proposed methods, each objective function along the Pareto-front is only improved by degrading the other objective function.

In this case, none of the Pareto-optimal solutions is exactly better. The general structure of Phase 1 is depicted in Figure 3.1. Note that the SO method runs during the half year in Phase 1. Then, the AI and data driven simulation in Phase 2 run for the next half year.

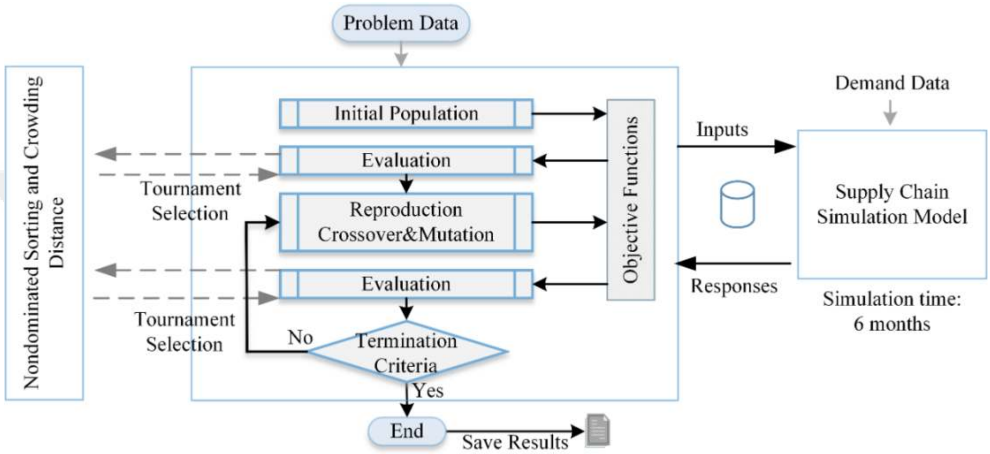

Figure 3.1. The general structure of Phase 1

In Phase 2, de-noising with MODWT is firstly used to improve the data quality. The outliers and systematic noises are identified to improve the demand data. After the preprocessing of data, AI and data driven simulation are used in the supply chain system. AI is utilized at each replenishment time to provide accurate demand forecasting. Forecasted demand is used to calculate the reorder point and order-up-to-level at each replenishment time (R) for each supply chain member. Note that only R is considered as a fixed value and assumed to be 5 days. In each replenishment time, the demand of supply chain members is forecasted by GA-ANN. The procedure of GA is initialized by defining chromosome structure to represent a set of parameters. Then, selection, crossover, and mutation operators are repeatedly applied to form new chromosomes in GA. The general structure of Phase 2 is depicted in Figure 3.2.

In this paper, ANN is selected as an AI method. Proposed ANN consists of four layers. First layer is the input layer that connects to the input variables. Second and third layer are called hidden layers that are between the input and output layers. Thus, we used two hidden layer. The last layer is the output layer that connects to the output variables. Information is transmitted through the connections between layers.

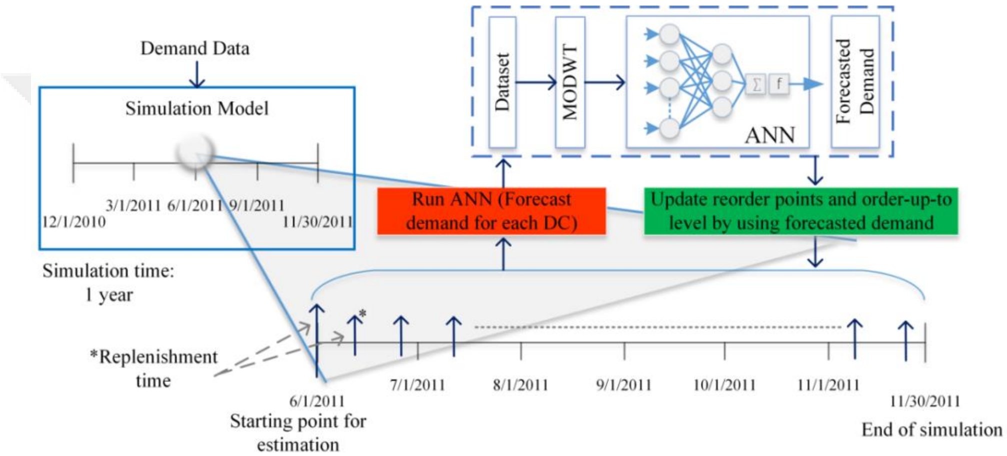

Figure 3.2. The general structure of Phase 2

In ANN, performance directly depends on the configuration of parameters in layers. The determination of ANN architecture is very crucial to improve the method performance. In literature, no certain method is available to perform well for the all types of problem. Therefore, researchers are tried to find systematic way for the ANN method development. In this paper, GA optimized the de-noising degree and the parameters of ANN including the number of neurons in hidden layers and training algorithm. Note that log-sigmoid is utilized as activation function in proposed ANN method. In ANN, mean absolute percentage error (MAPE) is utilized to evaluate the forecasting performance. The MAPE is calculated as follows:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{e_i}{a_i}\right| \tag{3.8}$$

where $e_i = f_i - a_i$, f_i is the forecasted demand, a_i is the actual demand.

3.4.1. Inventory Control Policy

Inventory control policies can differ in size and complexity. The selected inventory control policy varies according to the types of business structures. In this paper, a periodic (R, s, S) inventory control policy is considered. The inventory level is controlled at equal intervals of time R. If the inventory level falls below a reorder point (s), the supply chain member is replenished up to an order-up-to level (S). If the inventory level is more than the reorder point at time R, no replenishment order is placed.

In Phase 1, the initial inventory, reorder point, and order-up-to level are optimized by SO. In Phase 2, reorder point (s_c) and order-up-to level (S_c) are calculated at each replenishment time for each supply chain member using the output of GA-ANN. Taylor III (2013) presented the formula for reorder point with variable demand and variable lead time as follows.

$$s = \bar{d}\bar{L} + Z\sqrt{\sigma_d^2\bar{L} + \sigma_L^2\bar{d}^2} \tag{3.9}$$

$\bar{d}$ denotes the average daily demand, $\bar{L}$ represents the average lead time, $\sqrt{\sigma_d^2\bar{L} + \sigma_L^2\bar{d}^2}$ denotes standard deviation of demand during lead time, and Z is the number of standard deviations corresponding to the service level probability. This formula is adapted and the average daily demand is changed with the forecasted demand (d_f, obtained via GA-ANN) as follows. The 95% service level ($Z = 1.645$) is used for reorder point (s_c).

$$s_c = d_f\bar{L} + 1.645\sqrt{\sigma_d^2\bar{L} + \sigma_L^2 {d_f}^2} \tag{3.10}$$

Using calculated reorder point, order-up-to level (S_c) is computed as follows:

$$S_c = s_c + 5 \times \bar{d} \tag{3.11}$$

3.4.2. Routing Strategies

The various routing policies are used to achieve a variety of supply chain objectives ranging from low cost to high responsiveness. At this point, a wrong policy can hurt the customer service level while increasing the cost. Therefore, many different policies are developed to provide ongoing and consistent customer satisfaction while increasing profit. In literature, the strategies are generally represented as simple policies. On the other hand, strategy based policies enable us practical application on IRP side. Therefore, we used three different strategies.

Strategy 1 assigns route for vehicle considering TSCC. Under intense competition, marginal profit is becoming thinner and thinner in recent years, so companies should reduce total cost to overcome today's management challenges. The costs incurred in the IRP can play a major role while deciding on which IRP model to use in supply chain. Therefore, the largest TSCC is assigned first in Strategy 1.

Strategy 2 assigns route for vehicle considering least inventory first principle. The least inventory first principle is defined as the policy under which the vehicle goes next to the not-yet-visited supply chain member with the smallest inventory position. At the beginning of each replenishment time, the DCs' orders are ordered from smallest to largest by considering DCs' inventory levels. This means the order of DC with the smallest inventory level has the highest delivery priority.

In Strategy 3, vehicle starts its route according to predetermined priority rule. Priority order, from highest to lowest, in proposed methods is given as DC1, DC2, DC3, DC4, and DC5, respectively. Note that the vehicle considers the shortest paths when more than one route is available between these destinations for all strategies considered.

In this study, order splitting is not allowed. Vehicle is loaded and routed considering same strategy based policies. Thus, vehicle starts and continues its route according to loaded order. The vehicle considers the shortest paths while returning back to the Supplier. We used a single uncapacitated vehicle and vehicle speed is considered to be uniform (80, 90) meters per hour. In all strategies, each DC is visited at most once in each review period. Vehicle load changes during transportation. Quantity to be delivered to each DC is determined according to the value of replenishment order. Routing based cost ($V(R, s, S)$) is calculated using following equation.

$$V(R, s, S) = \sum_{d=1}^{number\ of\ total\ deliveries} \tau + \gamma_d + \sum_{i \epsilon S\{d\}} (\vartheta_i + \alpha Dist_{ij}) \qquad (3.12)$$

τ is the capital cost per vehicle, γ_d is the fixed cost of initiating delivery, and ϑ_i is fixed cost of each customer stop. Transportation cost per unit distance is denoted as α. The Supplier to the DC_i round trip distance is represented as $Dist_{ij}$ (i denotes the number of DC in the system, $i = 1, \ldots, I$ and j represents the Supplier). Finally, $S\{d\}$ denotes the set of DCs that is to be delivered at delivery d. Thus, $i \epsilon S\{d\}$ denotes each DC that is to be delivered at delivery d.

3.4.3. Performance Measurements

In order to analyze the performance of the methods, we used various measures including cost based analysis, quantity based analysis, lead time based analysis, routing based analysis, and average service level. In cost based analysis, the expected total cost for all supply chain members is the sum of the inventory based cost ($IC(R,s,S)$) and routing based cost ($V(R,s,S)$).

$$TSCC = IC(R,s,S) + V(R,s,S) \tag{3.13}$$

$$IC(R,s,S) = \sum_{n=1}^{Periods\ Considered}\{\sum_{i=1}^{I} h_i X_{in}^{+} + I\{X_{in} \leq s\}\ (k_i X_{in}^{-} + p_i P_i + c_i + O_i)\} \tag{3.14}$$

where X_{in}^{-} is the unmet customer order quantity of DC_i over period n and X_{in}^{+} is the remaining inventory quantity of DC_i over period n. h_i is the inventory holding cost rate of DC_i for each unit of inventory. k_i is specified as the lost sales cost rate of DC_i for each unit of stockout. p_i is the processing cost of DC_i. P_i is the processing time of DC_i. The order cost per use of DC_i is represented as c_i which is the cost charged for any order of DC_i irrespective of the time spent in there. The order processing cost of DC_i is represented as O_i. Note that order processing cost includes cost per use and order processing cost rate, which is proportional to the order processing time. The cost parameters in proposed supply chain are given in Table 3.1.

Table 3.1. The cost parameters in proposed supply chain

DCs Related Cost Parameters ($)	**Vehicle Related Cost Parameters ($)**
Average Holding Cost: Uniform (2,5) (h_i)	Capital Cost per Vehicle: 2000 (τ)
Lost Sales Cost: Uniform (50, 100) (k_i)	Fixed Cost of Initiating Delivery: 100 (γ_d)
Processing Cost: Uniform (5,10) (p_i)	Fixed Cost of Customer Stop: 100 (ϑ_i)
Order Cost Per Use: Uniform (50,100) (c_i)	Transportation Cost per Unit Distance: 0.05 (α)
Order Processing Cost Rate: Uniform (2,5)	
Cost Per Use: Uniform (5,10)	

In routing based analysis, the space utilization per delivery $(S\rho)$ is calculated using following formula.

$$S\rho = \frac{1}{T}\int_0^T Q(t)dt \qquad (3.15)$$

T is the total transportation time of vehicle loaded with orders per each delivery, and $Q(t)$ denotes the quantity of transporting orders at time t during a delivery. Transportation time per each delivery is the time difference between loading time of products on vehicle at supplier and unloaded time of the last order on the vehicle to be delivered. Delivery time is the total transportation time that is the time to transport products from supplier to DCs and back to supplier.

Finally, the average service level is calculated as follows:

$$Average\ service\ level\ = \frac{\sum_{a=0}^{Per\ Arrival} min(1,\frac{Current\ Inventory\ Level}{Incoming\ Order\ Quantity})}{Total\ Number\ of\ Incoming\ Orders} \qquad (3.16)$$

3.5. Big Data

There is still debate about when a data set will be considered as big data. However, there is still no discussion of five attributes related to big data. These are variety, velocity, volume, verification, and value. Variety represents that each technology can produce its own data. The data in big data databases are not written in the same code (non–unicode). Velocity refers to the increasing speed of the big data generation. Volume indicates that the amount of data collected increases at any time. Verification means that the data should be monitored considering the security level and should be visible to the right people. Value refers to the creation of a positive value for the organization after data production and processing (Işıklı, 2014). In addition to this, Wang and Alexander (2015) presented variability as an attribute related to big data. Thus, they described big data characteristics using six attributes related to big data.

There are two common data sources that grouped under big data. First data source is the shared data through automation and organization accession. This includes e-mails, mainframe logs, blocks, Adobe PDF documents, business processes, and other structured, unstructured, or semi-structured data available within the organization. Second data source includes data outside the organization. This includes non-public free information and subscription-based information. In addition, social media sites, freely shared product history by competitors, organizational hierarchies of corporate customers, helpful tips provided by third parties, and customer complaints can be included (Sathi, 2012). In addition, social media websites and job posting websites include a huge amount of data that are very useful for economic evaluation and community development. They provide the emotions and interests of people associated with web communities. The collected data from the web is seen as a remarkable source of data processing. Companies must use the data they collect to create value for their customers and employees.

3.6. Sentiment Analysis

A sentiment can be described as a view or opinion that is stated. It is a feeling of someone that includes either textual or verbal form. The sentiment can also be specified as a personal positive or negative feeling. The extraction of information from a piece of textual form is defined as sentiment analysis (Tyagi and Sharma, 2018).

Different algorithms can be utilized for sentiment classification. We summarized the sentiment classifiers in this section. Details about the sentiment classification techniques can be found in Medhat, Hassan, and Korashy (2014).

3.6.1. Machine Learning Approach

3.6.1.1. Supervised Learning

The supervised learning methods rely on the existence of labeled training documents. In literature, various supervised classifiers are used in sentiment analysis (Medhat et al., 2014). The some of them are briefly presented as follows.

Decision Tree Classifiers

A hierarchical form is used for the training data space. To partition the data, a condition on the attribute value is employed to partition the data. The partition of the data space is done recursively. This continues until the leaf nodes contain certain minimum numbers of records (Jain and Dandannavar, 2016).

Linear Classifiers

The output of the linear classifier is defined as $p = \bar{A}.\bar{X} + b$. The normalized document word frequency is $\bar{X} = \{x_1, \dots, x_n\}$. The vector of linear coefficients with the same dimensionality is represented as $\bar{A} = \{a_1, \dots, a_n\}$. Finally, b is a scalar (Medhat et al., 2014). ANN and SVM are two of the most famous linear classifiers. The ANN principle is explained in previous section.

Support Vector Machines: SVM is an example of a linear two-class classifier. A linear classifier is based on a linear discriminant function of the form

$$f(x) = w^T x + b \tag{3.17}$$

The vector w is known as the weight vector, and b is called the bias. To define a linear classifier, the dot product between two vectors is required. It is defined as $w^T x = \sum_i w_i x_i$. The notation x_i represents the ith vector in a dataset $\{(x_i, y_i)\}_{i=1}^{n}$, where y_i denotes the label associated with x_i. Details about the SVM can be found in Ben-Hur and Weston (2010).

Rule-based Classifiers (RBC)

The data space in RBC is modeled with a set of rules. In the RBC, the right hand side is the class label. On the other hand, the left hand side denotes a condition on the feature set expressed in disjunctive normal form. Details can be found in Medhat et al. (2014).

Probabilistic Classifiers

For classification, mixture models are employed in probabilistic classifiers. It is assumed that each class is a component of the mixture in mixture models. Each mixture component is a generative model. Hence, probabilistic classifiers are also called as generative classifiers (Medhat et al., 2014). In this paper, three of the most famous probabilistic classifiers are explained as follows.

Naïve Bayes: Naïve Bayes is a probabilistic classifier. It can be represented as follows:

$$\hat{c} = argmax_{c \in C} P(c|d) \tag{3.18}$$

where d denotes a document. Out of all classes $c \in C$ the classifier returns the class $\hat{c}$.

The intuition of Bayesian classification is to utilize Bayes' rule to transform Eq. 3.18 into other probabilities.

$$\hat{c} = argmax_{c \in C} P(c|d) = argmax_{c \in C} \frac{P(d|c)P(c)}{P(d)} \tag{3.19}$$

The most probable class $\hat{c}$ given some document d is computed by selecting the class. The class has the highest product of two probabilities:

$$\hat{c} = argmax_{c \in C} P(d|c)P(c) \tag{3.20}$$

where $P(c)$ denotes the prior probability of the class and $P(d|c)$ represents the likelihood of the document. Details about Naïve Bayes can be found in Jurafsky and Martin (2015).

Bayesian Network (BN): BN is considered as a complete model for the variables and their relationships. BN model is a directed acyclic graph. In the directed acyclic graph, nodes denote random variables and edges denote conditional dependencies (Medhat et al., 2014).

Maximum Entropy: Maximum entropy classifier can be defined as follows.

$$p(c|x) = \frac{exp\left(\sum_{i=1}^{N} w_i f_i(c,x)\right)}{\sum_{c' \in C} exp\left(\sum_{i=1}^{N} w_i f_i(c',x)\right)} \tag{3.21}$$

$f_i(c,x)$ denotes feature i for a particular class c for a given observation x. Z is the normalization factor. w_i represents the weight coefficient. The number of features is defined as N. Details about maximum entropy classifier can be found in Jurafsky and Martin (2015).

3.6.1.2. Unsupervised Learning (USL)

In USL, a large number of labeled training documents are needed to classify documents or texts. However, it is sometimes difficult to form these labeled training documents. At this point, the USL methods overcome these difficulties (Medhat et al., 2014).

3.6.2. Lexion-based Approach (LBA)

LBA is based on determining the opinion lexicon for computing the sentiment for a given text. Lexicon is built to determine the polarity of words or sentences. In this approach, the number of neutral, positive, and negative words is counted in the text (Jain and Dandannavar, 2016).

3.6.2.1. SentiStrength

SentiStrength is a lexicon-based classifier that utilizes additional linguistic information and rules to find sentiment strength in short informal English text. The SentiStrength output includes two integers for each text. One integer signifies the positive sentiment strength whose sentiment score is between +2 and +5 and the other signifies the negative sentiment strength whose sentiment score is ranging from -2 to -5 (Thelwall, Buckley, and Paltoglou, 2012). ±1 is used to signify the neutral words.

SentiStrength uses a dictionary of sentiment words with associated strength measures and exploits a range of recognized non-standard spellings and other common textual methods of expressing sentiment (Thelwall, Buckley, Paltoglou, Cai, and Kappas, 2010).

3.7. Proposed Big Data and Sentiment Analysis

Nowadays, people can easily access the information by a simple search on the Internet. The internet is used intensively in almost every sector such as health, entertainment, sports, art, and etc. They want to decide on their activities by learning about other people's feelings, experiences or views. In today's technology, people share their opinions, ideas, and feelings through social media sites, forums, and blogs. Especially with the widespread use of social media channels, people share a variety of information about their lives and emotions. Social media is a new alternative to communication tools such as newspapers, television, and radio. The popularity of social media is increasing day by day. Social media makes possible to reach the important findings without the need to make a survey to thousands of people. In addition, companies can reach people more easily by developing survey information management systems over the Internet in order to reduce the expenses for survey studies.

Online social media websites present the production, transmission, and consumption of fundamental information. To provide a connection between the consumers and the producers of information, social media can be successfully used. Today's world is connected to each other by means of the social network. For example, Twitter connects millions of people through that network. Social media is a great chance for companies to gain feedback related to product and insight in how to improve the product better. It also offers new opportunities for companies to directly interact with their customers (Gaikar and Marakarkandy, 2015). Especially, product reviews are one of the key sources for market research information of companies. Companies can learn about the market by monitoring reviews closely. On the other hand, the analysis of consumer-provided online reviews can be difficult due to the voluminous and dynamic structure. Most reviews include not only numerical ratings of product opinions but also textual content (Schneider and Gupta, 2016). Therefore, it is difficult to manually extract meaningful information from these data. In addition, it is hard for companies to summarize and analyze opinions about products. The internet environment contains a lot of information but surplus of information can make reaching the correct information complex and difficult. In addition, deciding about the sentiment of opinion is a challenging problem. For example, it is a troublesome process for a social media user to read and comment on hundreds of comments about a place such as restaurant. Thus, it may take a long time to evaluate people's feelings and thoughts in the absence of sentiment analysis. Sentiment analysis is the process of analyzing the data. It can be successfully used to analyze the online shopping dataset. Online website includes too much product reviews. Before buying a product, a customer read reviews about that product. The opinions of other people can be easily determined via sentiment analysis that is a type of natural language processing task. Sentiment analysis categorizes the views of people about a certain thing or topic into classes (Rani and Singh, 2017). In sentiment analysis, text polarity is generally categorized into three classes. First one is positive that holds a positive sentiment in the users' opinion. Second one is negative that holds a negative sentiment in the users' opinion. Finally, neutral does not hold a positive or negative sentiment in the users' opinion (Khader et al., 2018).

Big data analytics includes multimedia data analysis, structured data analysis, text data analysis, web data analysis, and mobile data analysis. Sentiment analysis is subclass of text data analysis and web data analysis (Edison and Aloysius, 2017). Sentiment analysis is employed to analyze customer reviews. At this point, big data is used to support long-term customer relationship and understand customers' life cycle and behavior in a more detailed way.

In this paper, sentiment analysis is used to offer a new methodology for the IRP. In Phase I, SentiStrength is used to calculate the score for customer reviews. After, ANN is integrated with the output of SentiStrength and the actual customer demands. In Phase II, the output of ANN (i.e., forecasted customer demands) is fed into NSGA-II based SO to solve the IRP. The general structure of the proposed methodology is given in Figure 3.3.

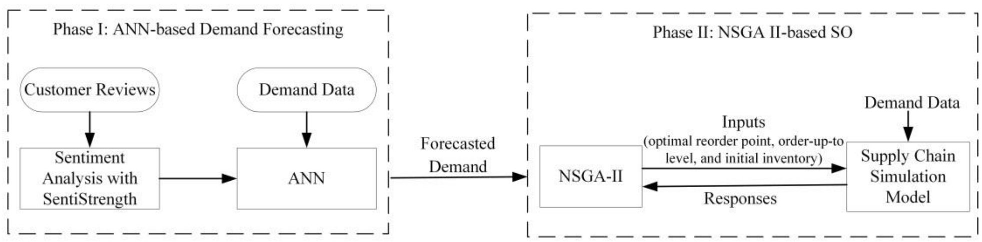

Figure 3.3. The general structure of the proposed methodology

In this paper, unstructured products information that is related to electronics customer reviews in Amazon is used. SentiStrength is subsequently used to extract sentiment strength from informal English text. It simultaneously extracts positive and negative sentiment strength from short informal electronic text. The sentiment score of customer reviews are calculated by sentiment analysis of reviews through SentiStrength. Note that sentiment analysis approach is lexicon and sentiment analysis levels are sentences. Also, sentiment analysis classification is polarity.

SentiStrength splits text into words. To determine the sentiment, each word is checked. The overall score of sentence is determined considering the highest positive and negative score for constituent words of sentence. When multiple sentences are available in the customer reviews, the maximum score of the each sentence is taken. The evaluation of sample sentences is given in Table 3.2. In this text, the maximum positive (ps) strength of each sentiment is 3 and the maximum negative (ng) strength of each sentiment is -1.

Along with reviews, customer demand data which is extracted from Chen et al. (2012) for the same period is stored, as well. After that ANN is employed to forecast the customer demand in supply chain.

Table 3.2. The evaluation of sample sentences

Customer Review	Ps	Ng	Emotion Rationale
What can I say, it's a Nikon lens cap. Good quality as expected. I love these caps with the spring action releases.	3	-1	• What can I say, it's a Nikon [proper noun] lens cap .[sentence: 1,-1] • Good[3] quality as expected .[sentence: 3,-1] • I love[3] these caps with the spring a [sentence: 3,-1]

To define the ANN, tf.keras is used. ANN splits dataset into training sets (80%) and testing set (20%). Note that dataset is normalized to [0, 1] using the following equation:

$$y = \frac{x - minimum}{maximum - minimum} \tag{3.22}$$

where x denotes the set of observed values. Minimum and maximum are the minimum and maximum values in the set of observed values, respectively. The structure of the ANN is given in Figure 3.4. ANN consists of three layers. First layer is the input layer. Second layer is the hidden layer that includes 20 hidden neurons. The last layer is the output layer.

In this paper, mean square error (MSE) is chosen as objective function in Keras. The MSE is calculated as follows:

$$\text{Mean square error} = \frac{1}{n}\sum_{i=1}^{n}(e_i)^2 \tag{3.23}$$

where $e_i = f_i - a_i$, f_i is the forecasted demand, a_i is the actual demand.

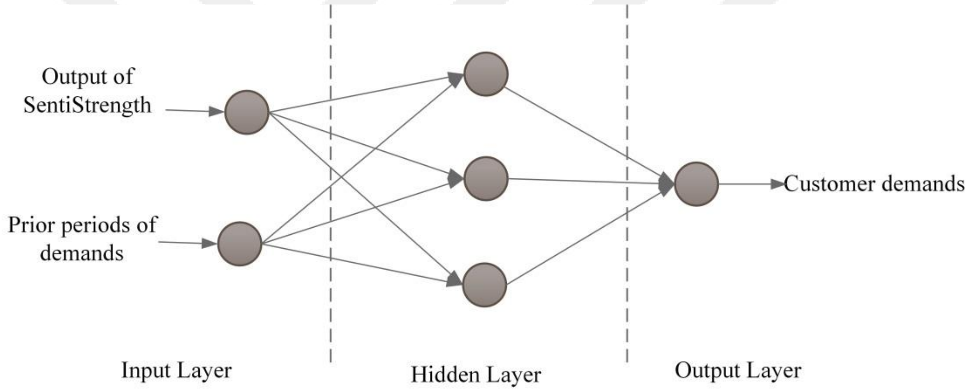

Figure 3.4. The structure of the ANN

For the optimizer, "adam" is selected. Rectified linear unit (ReLU) is utilized as activation function. Using forecasted customer demand, SO method is proposed to solve the IRP in proposed supply chain (Figure 3.5). Inventory levels of DCs are all monitored considering the periodic (R, s, S) inventory control policy. The idea is that every review period R units of time we control the inventory level. R is a fixed constant and assumed to be 5 days. At the beginning of each review period, the inventory level is replenished until the order-up-to level (S) whenever it decreases to a value smaller than or equal to the reorder level(s). If the inventory level is above s, nothing is done until at least the next review period. Note that SO model runs during three months. In this paper, supply chain includes five customers and five DCs. Each customer demand is satisfied by different DC.

The output of ANN is fed into the NSGA-II based SO to solve the IRP. NSGA-II is developed to assign new values for selected decision variables including reorder point, order-up-to level, and initial inventory that remain the same across the entire finite time horizon for customers. In each cycle, simulation output is returned to the NSGA-II. Then, NSGA-II once more tries to find better decision variables to increase model performance. Each DC has its own initial inventory, reorder point, and order-up-to level values, separately.

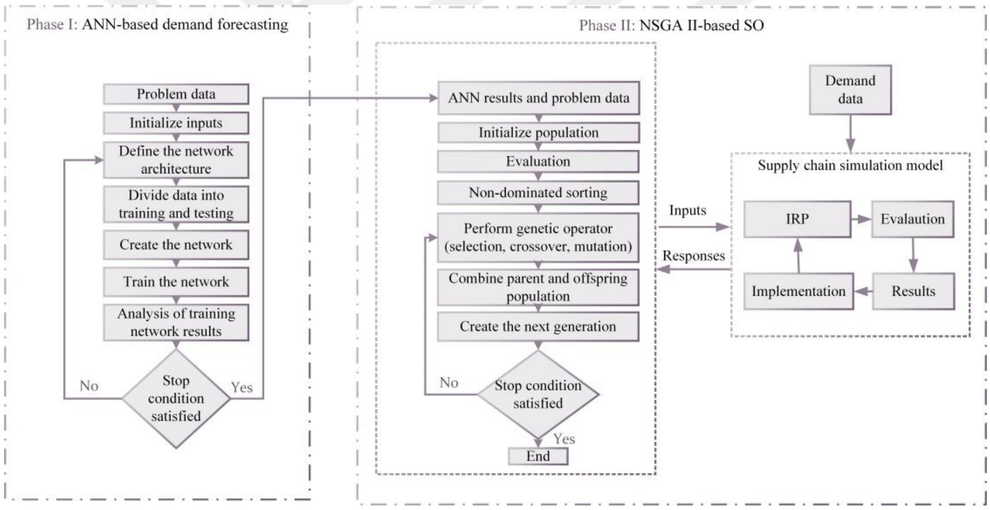

Figure 3.5. The proposed supply chain

4. RESEARCH FINDINGS AND DISCUSSION

4.1. Analysis of Proposed SO and AI based Methodology

Integrating optimization and simulation provides the high computational power to solve problems. SO methods can easily describe the interactions and uncertainties related to the system. However, the design of SO is crucial. A good overview of methods is necessary since the possibilities of combining simulation and optimization are so vast. In this study, NSGA-II is used as optimization method. In SO methods, TSCC is minimized while the average service level is maximized. Each objective function along the Pareto-front in proposed methods is improved by reducing the other objective function. At this point, one of the Pareto-optimal solutions can be considered as an acceptable solution since none of them is exactly better than the other solutions. In this study, the lowest cost solution is selected when Pareto-optimal solutions are found.

Simulation is the process where a real world problem is defined by a model. In simulation, main function is to determine how the system acts over periods. Until recently, various AI concepts were used to improve the performance of the simulation. To provide real world modeling, simulation and AI can be used. Doukidis and Angelides (1994) proposed a framework for the integration of simulation and AI. In addition, the fundamental similarities between them are demonstrated. They also presented that combining simulation and AI can be mutually useful. Rothenberg (1991) presented some of the major concepts of AI and demonstrated their applicability to simulation. The primary role of AI concepts on simulation is to provide additional features such as inferencing, search methods, reasoning, and representations (Rothenberg, 1991).

The performance of ANN directly depends on the parameters in layers. The determination of ANN architecture is very significant to increase the model performance. There is no certain methodology to perform well for the all types of problem.

Table 4.1. Optimized parameters of ANN

		Number of neuron in first hidden layer						Number of neuron in second hidden layer					
		Strategy 1		Strategy 2		Strategy 3		Strategy 1		Strategy 2		Strategy 3	
		min	max	min	max	min	max	min	max	min	max	min	max
Erratic demand	DC1	11	20	3	17	3	15	4	20	13	19	15	20
	DC2	3	19	8	15	5	17	18	20	18	20	19	20
	DC3	10	19	10	19	5	18	4	17	4	17	3	20
	DC4	5	17	6	19	4	14	7	20	5	18	17	19
	DC5	6	14	3	13	3	19	5	18	16	20	7	19
Intermittent demand	DC1	5	17	12	20	11	19	12	19	5	15	3	17
	DC2	8	19	3	18	13	20	5	20	6	20	6	19
	DC3	9	19	12	18	13	18	5	16	4	8	4	11
	DC4	12	20	12	18	11	20	4	15	5	10	4	11
	DC5	11	19	2	18	10	20	4	19	6	18	6	20
Lumpy demand	DC1	6	20	9	17	13	17	7	20	4	20	6	17
	DC2	13	19	13	19	13	19	5	18	5	18	5	18
	DC3	13	20	8	17	13	18	7	19	3	17	5	14
	DC4	9	19	15	20	14	20	5	17	3	10	3	11
	DC5	9	19	10	20	9	19	4	20	4	19	7	20

In this paper, the parameters of ANN including the number of neuron in hidden layers and training algorithm are optimized via GA. Note that, at first, pre-analysis is made to determine the upper bound of the number of neurons in the hidden layer. The model runs with different number of hidden neurons. It is determined that using more than 20 neurons did not increase performance at all. Therefore, the upper bound of the number of neurons in the hidden layers is fixed at 20 in GA. The minimum (min) and maximum (max) number of neuron in hidden layers varies according to the demand patterns and supply chain members (see Table 4.1). The values of the parameters determined by GA generally change according to the strategy type and supply chain member. The number of neurons in the first and second hidden layers is between 3 and 20 in erratic demand. The number of neurons in the first hidden layers is between 2 and 20 in intermittent demand while the number of neurons in the second hidden layers is between 3 and 20 in intermittent demand. The number of neurons in the first hidden layers is between 6 and 20 in lumpy demand while the number of neurons in the second hidden layers is between 3 and 20 in lumpy demand. GA is also employed to determine the de-noising degree whose values vary between 2 and 3.

Table 4.2. The initial pool of training algorithms

Index	Training Algorithm
1	Batch training with weight and bias learning rules
2	Broyden, Fletcher, Goldfarb, and Shanno Quasi-Newton method
3	Conjugate gradient backpropagation with Powell-Beale restarts
4	Conjugate gradient backpropagation with Fletcher-Reeves updates
5	Conjugate gradient backpropagation with Polak-Ribiere updates
6	Gradient descent
7	Gradient descent with adaptive learning rate
8	Gradient descent with momentum
9	Gradient descent momentum and an adaptive learning rate
10	Levenberg-Marquardt optimization
11	One step secant method
12	Resilient backpropagation algorithm
13	Scaled conjugate gradient method

Table 4.3. The descriptive statistics of reorder point in Phase 2

		Strategy 1				Strategy 2				Strategy 3			
		*RN	min	mean	max	RN	min	mean	max	RN	min	mean	max
Erratic demand	DC1	5	16	29	40	7	13	19	21	7	14	20	27
	DC2	4	33	44	56	3	34	41	45	3	35	41	45
	DC3	7	24	27	36	9	22	25	32	9	22	25	30
	DC4	6	36	52	82	6	33	46	66	5	37	46	65
	DC5	6	18	29	34	3	32	39	47	6	18	27	32
Intermittent demand	DC1	8	711	1334	1762	8	742	1555	2594	8	714	1281	1716
	DC2	13	597	1457	2167	12	572	1677	2636	12	622	1736	2643
	DC3	13	418	907	2328	12	534	1341	2099	15	529	1507	2905
	DC4	12	358	888	1400	10	441	1273	1825	13	348	945	1404
	DC5	13	714	1647	2291	16	825	1645	2271	14	705	1589	2284
Lumpy demand	DC1	7	543	731	1092	9	490	713	1372	8	543	828	2265
	DC2	11	644	917	1172	11	637	914	1157	11	715	922	1355
	DC3	8	237	653	1201	13	227	466	818	12	196	522	1036
	DC4	7	532	971	1610	9	401	605	831	9	424	728	1151
	DC5	9	276	502	767	9	341	665	1284	10	296	489	742

*RN represents the total replenishment numbers for each supply chain member at Phase 2.

Table 4.4. The descriptive statistics of order-up-to level in Phase 2

		Strategy 1				Strategy 2				Strategy 3			
		*RN	min	mean	max	RN	min	mean	max	RN	min	mean	max
Erratic demand	DC1	5	30	53	54	7	27	33	34	7	28	34	42
	DC2	4	50	60	72	3	51	57	62	3	52	58	61
	DC3	7	40	43	51	9	38	41	47	9	37	40	45
	DC4	6	58	72	101	6	52	65	85	5	56	66	83
	DC5	6	32	44	48	3	18	24	31	6	32	42	47
Intermittent demand	DC1	8	1177	1756	2178	8	1208	1977	3010	8	1180	1701	2131
	DC2	13	1097	1937	2642	12	1072	2156	3105	12	1122	2219	3114
	DC3	13	863	1306	2720	12	979	1744	2496	15	974	1904	3297
	DC4	12	660	1196	1706	10	743	1578	2149	13	348	945	1404
	DC5	13	1292	2103	2726	16	1403	2108	2761	14	1283	2041	2719
Lumpy demand	DC1	7	681	1025	1373	9	776	1011	1656	8	841	1126	2549
	DC2	11	918	1200	1448	11	911	1196	1433	11	989	1275	1631
	DC3	8	436	890	1429	13	426	685	1043	12	395	752	1295
	DC4	7	840	1274	1926	9	709	914	1143	9	732	1036	1442
	DC5	9	413	682	963	9	478	845	1484	10	433	671	938

*RN represents the total replenishment numbers for each supply chain member at Phase 2.

Training algorithm is also optimized by GA that selects the proper algorithm from initial pool as given in Table 4.2. In this paper, the conjugate gradient backpropagation with Powell-Beale restarts and one step secant method are generally selected to create the ANN for all strategies. Note that log-sigmoid is utilized as activation function and two hidden layers are used in proposed ANN methods. In ANN, MAPE is utilized to evaluate the prediction performance.

Inventory control is one of the most important aspects providing good customer service levels. Managers and researchers investigate how they can reduce inventory level without influencing customer service levels. An adequate level of inventory is necessary in order to be able to react to market demand and to minimize supply and demand imbalances in the supply chain. In this paper, the inventory level is controlled at each replenishment time. To provide optimal inventory level for each supply chain member, reorder point and order-up-to level are calculated at each replenishment time using forecasted customer demand. The descriptive statistics of reorder point and order-up-to level are given in Table 4.3-4.4 that includes minimum (min) and maximum (max) values. Remember that replenishment cycle is 5 days and the results of AI based simulation are collected during six months. The total replenishment number is different for each supply chain member and its value varies between 3 and 16 in Phase 2 (see Table 4.3-4.4). The vehicle visits at least 1 and at most 5 supply chain members per each delivery.

Average service level can be considered as one of the key performance indicators in supply chain. The high level of customer satisfaction can be ensured by improving the service level. Taking a glance at average service levels of strategies reveals that proposed method help properly control supply chain so that good customer service is maintained. The average service level of proposed methods is given in Table 4.5.

Table 4.5. Average service level for each strategy under different demand patterns

	Strategy 1	Strategy 2	Strategy 3
Erratic demand	1	1	1
Intermittent demand	0.9907	0.9981	0.9854
Lumpy demand	0.9969	0.9874	0.9955

Cost based analysis, order and quantity based analysis, lead time based analysis, and routing based analysis are also given as follows to compare the performance of routing strategies under different demand patterns.

4.1.1. Cost Component Analysis

In today's competitive world, the competition is not only between the companies in local market but also between companies in global market. To remain competitive and grow sales, supply chain should be controlled considering performance measurements. The criteria for the supply chain performance measurement can be different for companies since the participants and the structure of the network can vary according to the view of its own vision.

According to cost based analysis, the cost of order processing has the highest share among the cost components of erratic demand. The smallest share of cost component in erratic demand is either lost sales cost or processing cost. In proposed system, supply chain includes five members (DC1, DC2, DC3, DC4, and DC5). Thus, total cost component value is the sum of inventory based cost value for all DCs. When either total average holding cost or total processing cost is considered, Strategy 2 can be selected in erratic demand. On the other hand, Strategy 3 gives better results when the other total cost components are taken into account in erratic demand (Table 4.6).

The highest share of cost component in intermittent demand is generally average holding cost (Table 4.7). The lowest share of cost component in intermittent demand is either lost sales cost or processing cost.

Table 4.6. The cost component analysis of erratic demand

Erratic demand			Average Holding Cost	Order Cost Per Use	Lost Sales Cost	Order Processing Cost	Processing Cost
DC1	Strategy 1	Cost	1541	8589	185	11044	137
		Share	7	40	1	51	1
	Strategy 2	Cost	458	8262	1033	11432	163
		Share	2	39	5	54	1
	Strategy 3	Cost	741	8340	646	11453	197
		Share	3	39	3	54	1
DC2	Strategy 1	Cost	896	7155	5215	9739	143
		Share	4	31	23	42	1
	Strategy 2	Cost	1664	7095	3211	9442	132
		Share	8	33	15	44	1
	Strategy 3	Cost	2259	7194	673	9537	131
		Share	11	36	3	48	1
DC3	Strategy 1	Cost	1346	9043	3110	12379	194
		Share	5	35	12	47	1
	Strategy 2	Cost	761	8896	2102	12260	240
		Share	3	37	9	51	1
	Strategy 3	Cost	661	9182	1664	12423	224
		Share	3	38	7	51	1
DC4	Strategy 1	Cost	1610	8272	910	11250	183
		Share	7	37	4	51	1
	Strategy 2	Cost	1723	8479	1298	11032	145
		Share	8	37	6	49	1
	Strategy 3	Cost	1274	7887	1526	10836	108
		Share	6	36	7	50	1

Table 4.6. (Continued)

Erratic demand			**Average Holding Cost**	**Order Cost Per Use**	**Lost Sales Cost**	**Order Processing Cost**	**Processing Cost**
DC5	Strategy 1	Cost	827	5466	5997	7257	112
		Share	4	28	31	37	1
	Strategy 2	Cost	1090	5318	7050	7360	38
		Share	5	25	34	35	0.2
	Strategy 3	Cost	1072	5276	3866	7232	104
		Share	6	30	22	41	1

Table 4.7. The cost component analysis of intermittent demand

Intermittent demand			Average Holding Cost	Order Cost Per Use	Lost Sales Cost	Order Processing Cost	Processing Cost
DC1	Strategy 1	Cost	27124	8461	0	11176	1688
		Share	56	17	0	23	3
	Strategy 2	Cost	30438	8135	14080	11139	1899
		Share	46	12	21	17	3
	Strategy 3	Cost	23863	8102	0	11032	1731
		Share	53	18	0	25	4
DC2	Strategy 1	Cost	30024	8566	13385	11730	3050
		Share	45	13	20	18	5
	Strategy 2	Cost	29726	8677	43118	11685	3222
		Share	31	9	45	12	3
	Strategy 3	Cost	31060	8700	9000	11761	3127
		Share	49	14	14	18	5
DC3	Strategy 1	Cost	14968	9271	40135	12936	2645
		Share	19	12	50	16	3
	Strategy 2	Cost	24724	9363	30986	12610	2300
		Share	31	12	39	16	3
	Strategy 3	Cost	19830	9596	2047	12493	2721
		Share	42	21	4	27	6
DC4	Strategy 1	Cost	19016	10502	24700	13875	2190
		Share	27	15	35	20	3
	Strategy 2	Cost	23196	10502	47533	13857	2096
		Share	24	11	49	14	2
	Strategy 3	Cost	19184	10321	25976	14189	2237
		Share	27	14	36	20	3

Table 4.7. (Continued)

Intermittent demand			Average Holding Cost	Order Cost Per Use	Lost Sales Cost	Order Processing Cost	Processing Cost
DC5	Strategy 1	Cost	33735	12235	0	16892	3178
		Share	51	19	0	26	5
	Strategy 2	Cost	28012	12513	0	16938	3459
		Share	46	21	0	28	6
	Strategy 3	Cost	28479	12664	377	17069	3287
		Share	46	20	1	28	5

Table 4.8. The cost component analysis of lumpy demand

Lumpy demand			Average Holding Cost	Order Cost Per Use	Lost Sales Cost	Order Processing Cost	Processing Cost
DC1	Strategy 1	Cost	16259	6282	0	8959	1261
		Share	50	19	0	27	4
	Strategy 2	Cost	21204	6702	0	9162	1363
		Share	55	17	0	24	4
	Strategy 3	Cost	24010	6676	3362	8573	1583
		Share	54	15	8	19	4
DC2	Strategy 1	Cost	19103	6543	0	9067	918
		Share	54	18	0	25	3
	Strategy 2	Cost	20307	6403	0	8916	1040
		Share	55	17	0	24	3
	Strategy 3	Cost	23544	6646	0	9106	976
		Share	58	17	0	23	2
DC3	Strategy 1	Cost	13772	6843	90424	9446	1453
		Share	11	6	74	8	1
	Strategy 2	Cost	9545	7038	27122	9524	1508
		Share	17	13	50	17	3
	Strategy 3	Cost	9945	6970	54560	9342	1437
		Share	12	8	66	11	2
DC4	Strategy 1	Cost	16857	6631	1939	8921	1536
		Share	47	18	5	25	4
	Strategy 2	Cost	13130	6471	14442	8746	1364
		Share	30	15	33	20	3
	Strategy 3	Cost	15317	6654	21146	9090	1477
		Share	29	12	39	17	3

Table 4.8. (Continued)

Lumpy demand			Average Holding Cost	Order Cost Per Use	Lost Sales Cost	Order Processing Cost	Processing Cost
DC5	Strategy 1	Cost	9925	4596	43655	6259	829
		Share	15	7	67	10	1
	Strategy 2	Cost	13007	4687	50732	6263	927
		Share	17	6	67	8	1
	Strategy 3	Cost	11879	4584	36510	5897	958
		Share	20	8	61	10	2

Strategy 3 can be selected in intermittent demand when either total average holding cost or total lost sales cost is considered. In intermittent demand, Strategy 1 can be considered when either total order cost per use or total processing cost is taken into account. In intermittent demand, Strategy 2 gives better results when total order processing cost is considered. Determining the priority of performance measures is a main concern for supply chain managers and researchers to select the best strategy in supply chain.

The highest share of cost component in lumpy demand is average holding cost or lost sales cost (Table 4.8). The smallest share of cost component in lumpy demand is either lost sales cost or processing cost. Strategy 2 can be chosen in lumpy demand when total lost sales cost is taken into account. Strategy 3 gives better results in lumpy demand when total order processing cost is considered. In lumpy demand, Strategy 1 gives better results when other total cost components are considered.

When total routing cost is considered, Strategy 1 can be selected for lumpy demand patterns (Table 4.9). Strategy 2 gives better results for erratic demand pattern with respect to the total routing cost. Strategy 3 can be chosen for intermittent demand when total routing cost is taken into account.

Table 4.9. The total routing cost for each demand pattern and each strategy

	Erratic Demand	**Intermittent Demand**	**Lumpy Demand**
Strategy 1	129482	577463	349196
Strategy 2	129464	586486	383050
Strategy 3	131065	559061	358861

When the TSCC is considered, Strategy 2 can be selected for lumpy demand patterns (Table 4.10). Thus, the least inventory is controlled in supply chain to cope with the fluctuating nature of lumpy demand. Strategy 3 can be chosen for intermittent demand and erratic demand pattern when the TSCC is taken into account. According to the TSCC analysis, it can be said that selection highest priority in DCs gives satisfying cost results in intermittent demand and erratic demand pattern.

Table 4.10. The TSCC for each demand pattern and each strategy

	Erratic Demand	Intermittent Demand	Lumpy Demand
Strategy 1	242081	908945	640675
Strategy 2	240150	986694	632658
Strategy 3	235570	847906	639104

4.1.2. Quantity Based Analysis and Order Based Analysis

Quantity based analysis that collected during six months is given in Figure 4.1-4.6.

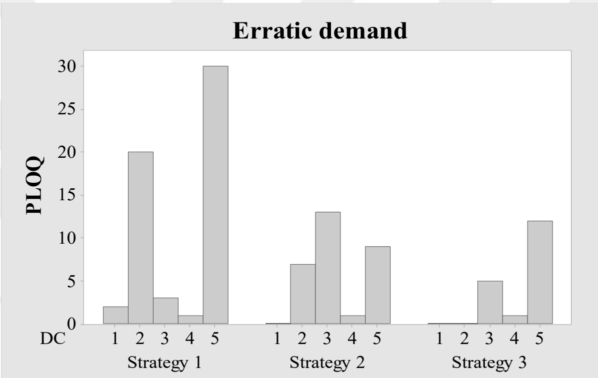

Figure 4.1. The analysis of PLOQ for erratic demand

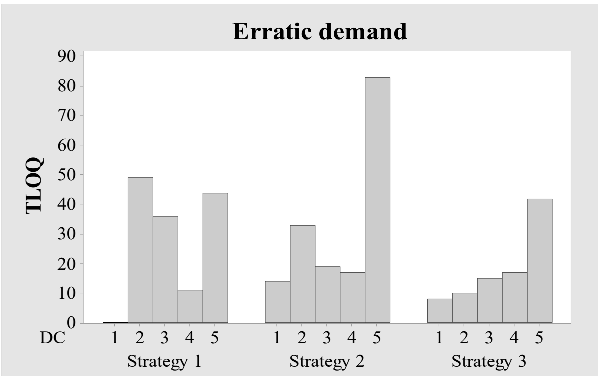

Figure 4.2. The analysis of TLOQ for erratic demand

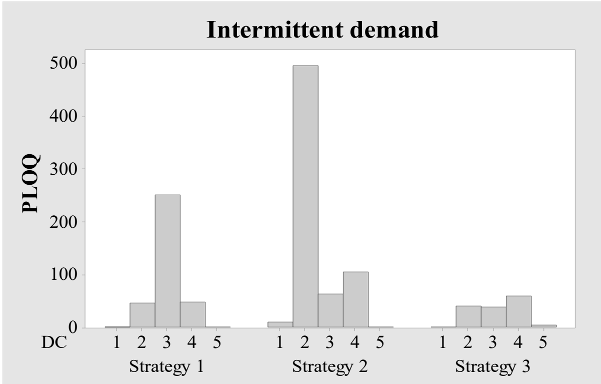

Figure 4.3. The analysis of PLOQ for intermittent demand

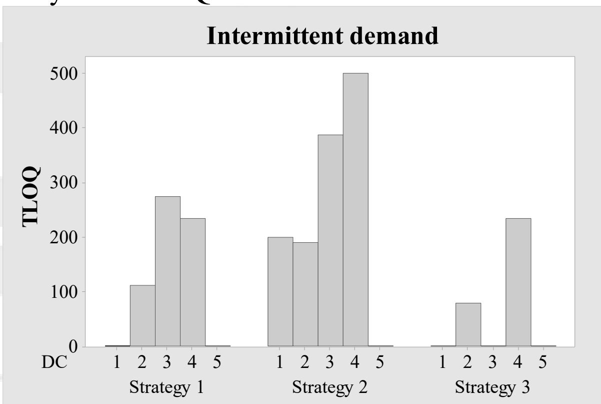

Figure 4.4. The analysis of TLOQ for intermittent demand

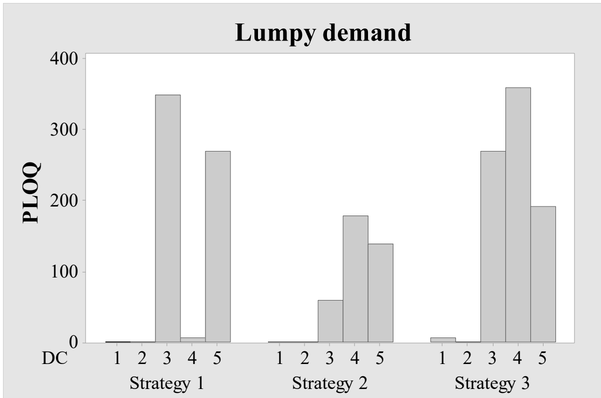

Figure 4.5. The analysis of PLOQ for lumpy demand

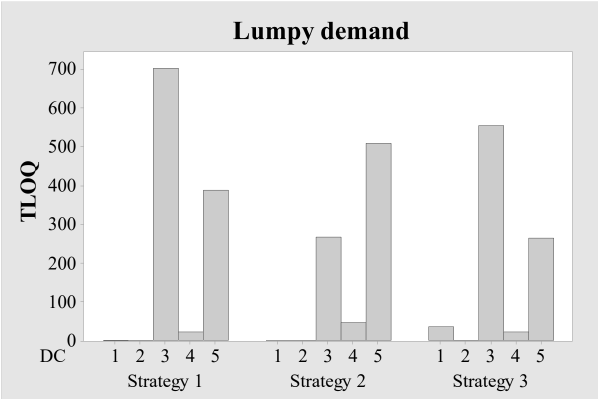

Figure 4.6. The analysis of TLOQ for lumpy demand

Remember that proposed supply chain includes five DCs. Therefore, we also evaluated the summation of the values of PLOQ and TLOQ for DCs in each demand pattern. According the results of total PLOQ and total TLOQ, Strategy 2 gives better results in lumpy demand. However, Strategy 3 gives better results in erratic demand and intermittent demand with respect to the results of total PLOQ and total TLOQ. The nature of lumpy demand causes difficulties in controlling inventory level, and therefore, focusing on the smallest inventory improves the PLOQ and the TLOQ. On the other hand, giving priority to DCs improves the lost order quantities in erratic demand and intermittent demand.

Order based analysis is given in Table 4.11. The NPLO and the NTLO include the values that collected during six months. In addition, we evaluated the summation of the values of NPLO and NTLO for DCs in each demand pattern since proposed supply chain includes five DCs.

In lumpy demand, Strategy 1 and Strategy 2 give the better results than Strategy 3 with respect to the total NPLO. In erratic demand, Strategy 3 can be chosen when the results of total NPLO are considered. In intermittent demand, both Strategy 1 and Strategy 3 give the better results than Strategy 2 with respect to the total NPLO.

Table 4.11. Order based analysis

		NPLO			NTLO		
		Strategy 1	**Strategy 2**	**Strategy 3**	**Strategy 1**	**Strategy 2**	**Strategy 3**
Erratic demand	DC1	1	0	0	0	9	6
	DC2	4	1	0	13	8	3
	DC3	1	2	1	3	2	2
	DC4	1	1	1	4	4	4
	DC5	3	2	2	14	17	13
Intermittent demand	DC1	0	1	0	0	2	0
	DC2	1	2	1	3	3	1
	DC3	2	2	1	5	7	0
	DC4	1	2	1	2	4	2
	DC5	0	0	1	0	0	0
Lumpy demand	DC1	0	0	1	0	0	3
	DC2	0	0	0	0	0	0
	DC3	2	2	2	7	5	6
	DC4	1	1	1	1	2	1
	DC5	2	2	2	3	4	2

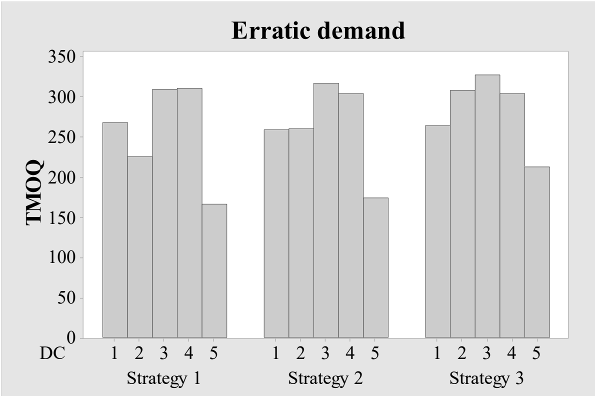

Figure 4.7. The analysis of TMOQ for erratic demand

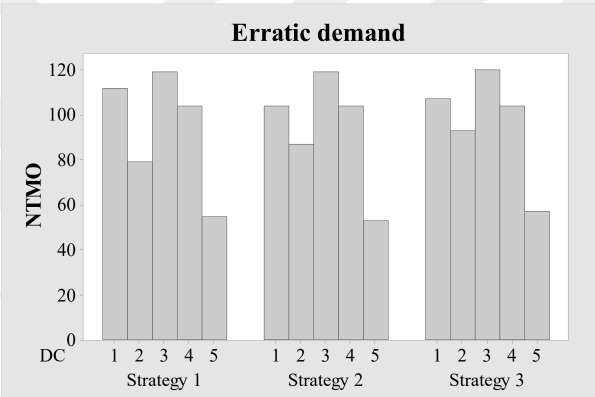

Figure 4.8. The analysis of NTMO for erratic demand

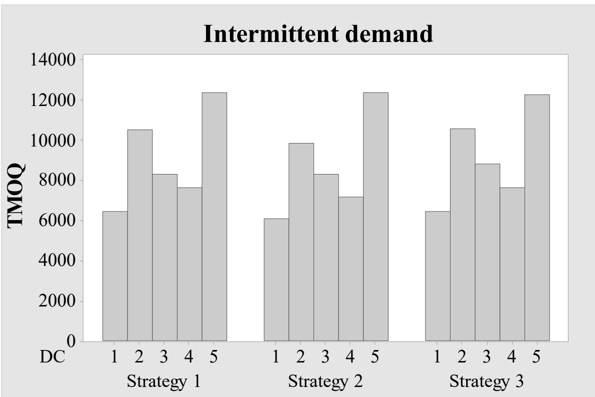

Figure 4.9. The analysis of TMOQ for intermittent demand

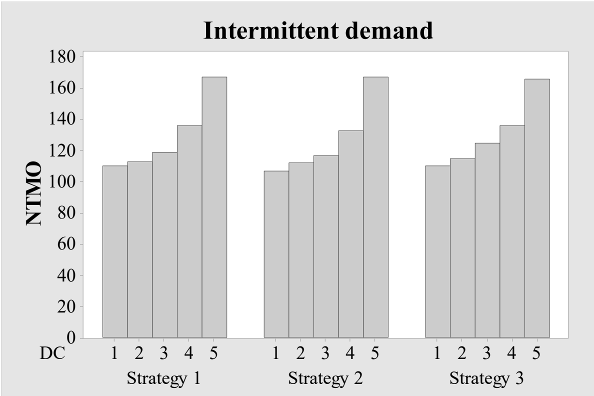

Figure 4.10. The analysis of NTMO for intermittent demand

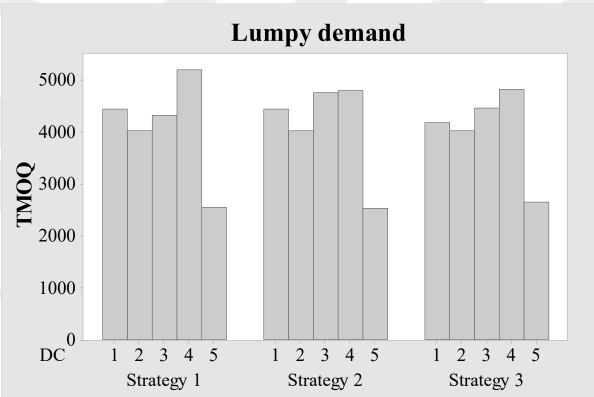

Figure 4.11. The analysis of TMOQ for lumpy demand

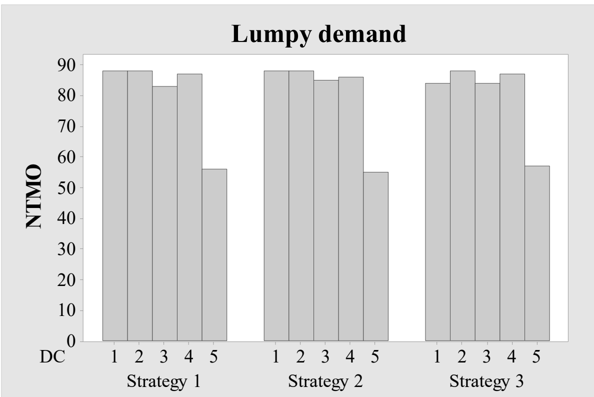

Figure 4.12. The analysis of NTMO for lumpy demand

In erratic demand and intermittent demand, Strategy 3 can be chosen when the results of total NTLO are considered. In lumpy demand, both Strategy 1 and Strategy 2 give the better results than Strategy 3 with respect to the total NTLO.

To provide cost effective system, finding optimal inventory is one of the key point. The excess inventory can cause pointless storage costs while inventory shortage yields to lost sales. The higher the inventory levels in DCs, the lower the chance of lost sales. However, excess holding cost can be incurred when customer order quantity can be lower than inventory level. Hence, determining exact inventory level at each DC in the IRP without shortages and excesses is a main concern for supply chain managers and researchers. They have dedicated themselves to search more robust model.

The analysis of the TMOQ and the NTMO whose values are collected during six months for demand patterns is given in Figure 4.7-4.12. The results vary according to the demand patterns and supply chain members. We also evaluated the summation of the values of TMOQ and NTMO for DCs in each demand pattern.

According the results of total TMOQ and NTMO, Strategy 2 gives better results in lumpy demand. However, Strategy 3 gives better results in erratic demand and intermittent demand with respect to the results of total TMOQ and NTMO.

4.1.3. Lead Time based Analysis

The average holding units during lead time is given in Table 4.12. The value of average holding units during lead time changes according to the supply chain members and strategy types. The minimum value of the average holding units during lead time is 0 units for all strategies. The range of average holding units during lead time is 47 units in erratic demand. In intermittent demand, the range of average holding units during lead time is 2445 units. The range of average holding units during lead time is 1141 units in lumpy demand.

Table 4.12. Descriptive statistics of average holding units during lead time

		Strategy 1			Strategy 2			Strategy 3		
		min	**mean**	**max**	**min**	**mean**	**max**	**min**	**mean**	**max**
Erratic demand	DC1	4	14	36	0	10	17	3	11	18
	DC2	0	6	29	0	6	18	0	11	27
	DC3	0	15	32	0	11	30	0	10	22
	DC4	0	21	47	0	25	45	0	19	32
	DC5	0	12	30	0	5	16	0	13	30
Intermittent demand	DC1	191	924	1505	71	1094	2420	179	878	1510
	DC2	17	1014	1804	0	1077	2445	40	1178	2276
	DC3	26	463	1130	0	738	1533	77	97	120
	DC4	0	596	1329	0	680	1202	0	651	1181
	DC5	187	1056	2077	336	1122	2013	121	1028	1892
Lumpy demand	DC1	146	432	953	142	522	999	120	456	984
	DC2	431	758	1067	222	713	1068	632	806	1100
	DC3	0	289	648	0	259	452	0	286	716
	DC4	4	542	878	0	394	661	0	424	1141
	DC5	0	299	578	0	415	936	0	309	636

Table 4.13. The lead time period length ratio to review period length

		Strategy 1			Strategy 2			Strategy 3		
		min	**mean**	**max**	**min**	**mean**	**max**	**min**	**mean**	**max**
Erratic demand	DC1	0.103	0.147	0.188	0.109	0.132	0.197	0.098	0.121	0.176
	DC2	0.150	0.161	0.188	0.153	0.169	0.194	0.141	0.165	0.191
	DC3	0.104	0.133	0.224	0.109	0.130	0.226	0.113	0.143	0.219
	DC4	0.110	0.163	0.249	0.105	0.132	0.188	0.110	0.149	0.254
	DC5	0.144	0.150	0.157	0.172	0.189	0.219	0.148	0.156	0.162
Intermittent demand	DC1	0.243	0.336	0.537	0.283	0.392	0.654	0.301	0.432	0.565
	DC2	0.265	0.368	0.525	0.294	0.417	0.580	0.339	0.489	0.696
	DC3	0.205	0.337	0.486	0.318	0.404	0.580	0.235	0.416	0.663
	DC4	0.243	0.318	0.400	0.212	0.429	0.783	0.207	0.327	0.519
	DC5	0.329	0.457	0.729	0.280	0.388	0.607	0.285	0.396	0.615
Lumpy demand	DC1	0.260	0.339	0.449	0.198	0.277	0.368	0.225	0.378	0.603
	DC2	0.218	0.269	0.340	0.234	0.278	0.419	0.228	0.353	0.531
	DC3	0.181	0.291	0.470	0.158	0.298	0.525	0.162	0.317	0.508
	DC4	0.274	0.379	0.539	0.195	0.294	0.524	0.192	0.311	0.450
	DC5	0.209	0.263	0.375	0.215	0.310	0.554	0.194	0.236	0.353

Table 4.14. The mean of order met probabilities during lead time

		Strategy 1	Strategy 2	Strategy 3
Erratic demand	DC1	0.590	1.000	1.000
	DC2	1.000	1.000	1.000
	DC3	1.000	1.000	1.000
	DC4	0.404	0.818	0.677
	DC5	1.000	1.000	1.000
Intermittent demand	DC1	1.000	0.794	1.000
	DC2	0.903	0.951	0.876
	DC3	0.735	0.783	0.988
	DC4	1.000	0.844	1.000
	DC5	1.000	1.000	0.998
Lumpy demand	DC1	1.000	1.000	0.899
	DC2	1.000	1.000	1.000
	DC3	0.665	0.739	0.763
	DC4	0.665	0.882	0.886
	DC5	0.748	0.612	0.692

Note that the range is the value that is the difference between the minimum value and the maximum value.

The lead time period length ratio to review period length varies with respect the demand patterns and supply chain members. Its minimum (min), its mean, and its maximum (max) values are given in Table 4.13. In Strategy 1, lead time of supply chain member comprised minimum 10.3 percent of review period and maximum 72.9 percent of review period. In Strategy 2, lead time of supply chain member comprised minimum 10.5 percent of review period and maximum 78.3 percent of review period. In Strategy 3, lead time of supply chain member comprised minimum 9.8 percent of review period and maximum 69.6 percent of review period.

The mean of order met probabilities during lead time varies according to the supply chain members and strategy types (Table 4.14). According to the mean of order met probabilities during lead time, Strategy 2 can be chosen in erratic demand pattern. On the other hand, Strategy 3 can be selected for intermittent demand and lumpy demand when the mean of order met probabilities during lead time is taken into account.

4.1.4. Routing based Analysis

In an economic environment of fierce competition, routing is very challenging problem since real-life settings lead to a very large number of variants born of the requirement to manage a wide variety of characteristics and decisions to account for the particular supply chain member. Hence, we paid particular attention to the modeling phase aiming to propose an advanced modeling approach for obtaining a flexible, parametric, and time efficient method for an accurate capture of the supply-chain behavior. We address the IRP from the perspective of different demand patterns considering strategy based policies. Various analyses are employed to determine the better routing strategy for each demand pattern. Firstly, the total number of delivery for each strategy and each demand pattern is given in

Table 4.15 since all other routing analyses are made considering the total number of delivery.

Table 4.15. The total number of delivery for each strategy and each demand pattern

	Strategy 1	Strategy 2	Strategy 3
Erratic demand	23	22	26
Intermittent demand	34	30	29
Lumpy demand	22	26	27

The mean of transported quantity from supplier to DCs is given in Figure 4.13-4.15 for demand patterns. Vehicle transports different quantity for each supply chain member. Also, the mean of transported quantity varies according to the strategy type. The descriptive statistics of space utilization per delivery are given in Table 4.16. Note that mean represents the average of the space utilization per delivery for the DCs.

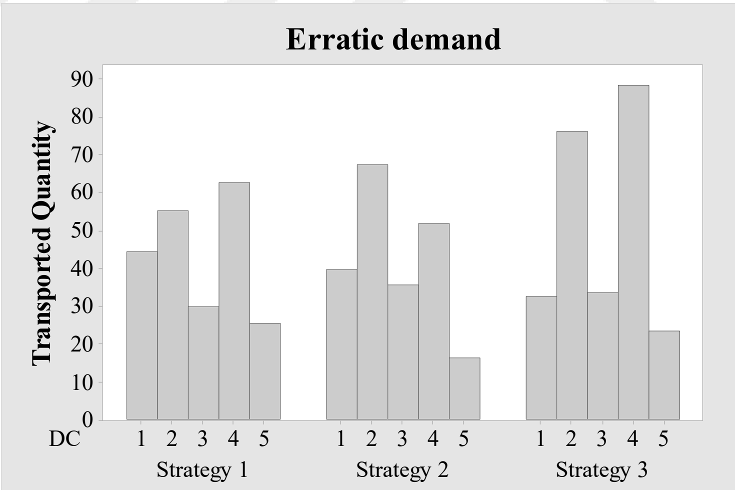

Figure 4.13. The mean of transported quantity from supplier to DCs in erratic demand

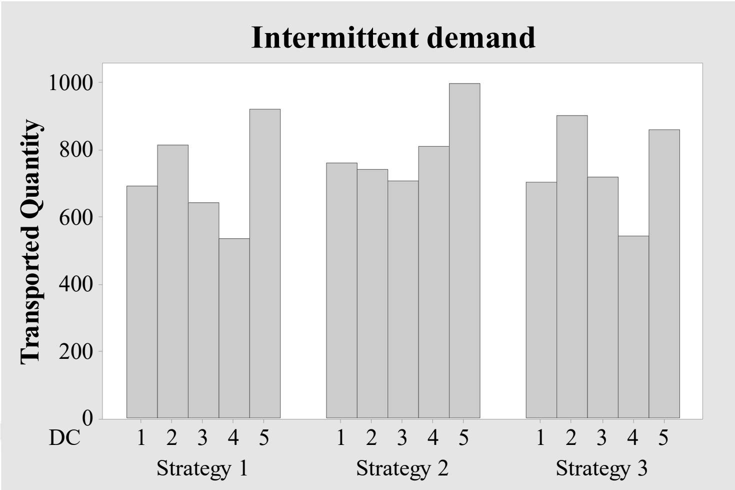

Figure 4.14. The mean of transported quantity from supplier to DCs in intermittent demand

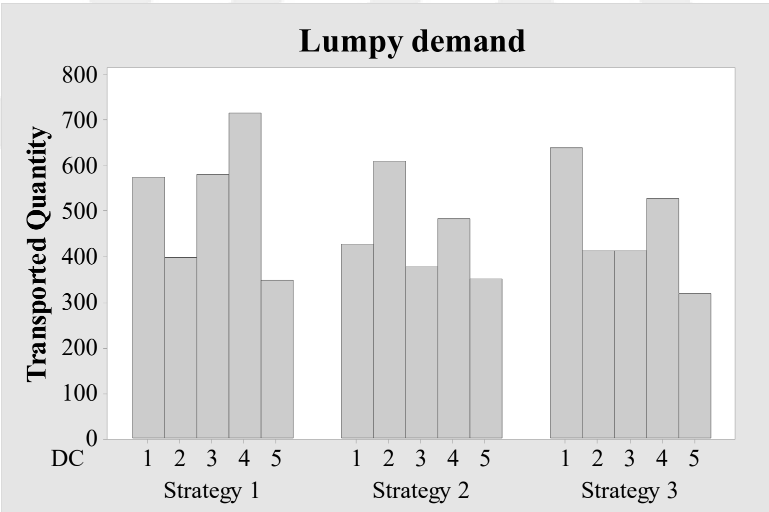

Figure 4.15. The mean of transported quantity from supplier to DCs in lumpy demand

The space utilization per delivery varies between 7 units and 1609 units for demand patterns in Strategy 1. The space utilization per delivery varies between 9 units and 2986 units for demand patterns in Strategy 2.

Table 4.16. The descriptive statistics of space utilization per delivery for strategies

Demand Pattern	Strategy 1			Strategy 2			Strategy 3		
	min	mean	max	min	mean	max	min	mean	max
Erratic demand	7	27	74	9	27	128	7	24	83
Intermittent demand	152	647	1609	180	828	2986	170	865	1874
Lumpy demand	105	486	1393	102	408	1130	91	431	1528

Table 4.17. Transportation time per each delivery

Demand Pattern	Strategy 1			Strategy 2			Strategy 3		
	min	mean	max	min	mean	max	min	mean	max
Erratic demand	9.02	13.46	24.65	9.03	11.97	21.52	9.13	11.95	21.74
Intermittent demand	8.99	17.31	32.27	9.24	18.68	39.48	9.70	22.69	47.52
Lumpy demand	9.06	16.06	33.95	9.13	17.51	45.85	9.10	18.75	49.03

Table 4.18. Total traveling time per delivery

Demand Pattern	Strategy 1			Strategy 2			Strategy 3		
	min	mean	max	min	mean	max	min	mean	max
Erratic demand	18.46	25.06	34.05	18.14	23.22	33.26	18.68	23.52	33.51
Intermittent demand	18.43	29.67	48.05	18.09	30.56	48.55	18.50	34.40	60.82
Lumpy demand	18.27	27.21	44.29	17.91	29.36	61.18	17.91	30.55	60.17

The space utilization per delivery varies between 7 units and 1874 units for demand patterns in Strategy 3. When the mean space utilization per delivery is taken into account in erratic demand, Strategy 1 and Strategy 2 give better results than Strategy 3. Strategy 1 can be chosen in lumpy demand with respect to the mean space utilization per delivery. In intermittent demand, Strategy 3 can be selected according to the mean space utilization per delivery.

The descriptive statistics of transportation and travelling is given in Table 4.17-4.18. When the mean of transportation time per each delivery and mean of total traveling time per delivery are taken into account in erratic demand, Strategy 2 and Strategy 3 give better results than Strategy 3. Strategy 1 can be chosen in lumpy demand and intermittent demand with respect to the mean of transportation time per each delivery and mean of total traveling time per delivery.

The solving of IRP is challenging task in a stochastic and/or dynamic environment. Differences occur when customer demand patterns are changes. Many different viewpoints must be taken into account by considering uncertainty and dynamic nature of the system. A fundamental challenge in stochastic environment is computability and tractability.

4.2. Analysis of Big Data and Sentiment Analysis

For supply chain system, data is the raw material and the data quality determines the position of an organization. Organizations should pay attention to generate right data for decision-making. Companies are struggling in handling the big data. They are surveying new methods to give valuable insights to industries. The big data makes these data valuable to analyze the supply chain. Many research and study tried to analyze, build perspective, and advance the opportunities for big data research and application in supply chain management.

In this study, electronics customer reviews in Amazon is used and SentiStrength is employed to extract sentiment score for each customer review. For example, one customer review is given as follows:

"I love these mice! I have one for my laptop and absolutely loved it. So I bought one for my husband. It works great with everything, a true PLUG-n-PLAY!! It even worked on an older model laptop he brought home for work. My only complaint is that I wish the usb fit more seamlessly into the mouse."

SentiStrength calculates the positive sentiment score whose value is 5 and negative sentiment score whose value is -2. Although different forms are available to calculate the final result score, the scale output form is used in this study. It presents the result as the summation of positive sentiment and negative sentiment. Thus, positive sentiment and negative sentiment are summed and the final score is determined as 3. Considering same principle, sentiment score is calculated for each review. Then, average sentiment score is calculated for each day. Along with reviews, customer demand data is fed into the ANN to forecast demand in supply chain. In ANN, overall dataset is divided into training dataset (80%) and testing dataset (20%).

The value of MSE for testing set is given in Table 4.19 for each customer. The output of ANN is fed into the NSGA-II based SO to solve the IRP. In proposed methodology, NSGA-II is used to determine the optimal reorder point, order-up-to level, and initial inventory (Table 4.20).

Table 4.19. The MSE value for each customer in ANN

Customers	MSE
Customer 1	0.01
Customer 2	0.05
Customer 3	0.11
Customer 4	0.05
Customer 5	0.03

For quantity based analysis and order based analysis, TMOQ and NTMO are calculated for each supply chain member. For example, TMOQ for DC1 in Table 4.21 is 2487 units that represent the total satisfied customer order quantity during the three months. Also, NTMO for DC1 is 17 orders that denote the total number of customer orders that satisfied during the three months. Although all customer demands are met in proposed method, satisfied order quantity varies according to the supply chain member (Table 4.21).

Table 4.20. The determined values of inventory control variables for each supply chain member

Supply chain member	Initial Inventory	Reorder Point	Order-up-to Level
DC1	603	403	1425
DC2	658	371	1340
DC3	1230	480	730
DC4	1230	480	699
DC5	1230	498	1164

Table 4.21. Quantity based analysis and order based analysis

Supply chain member	NTMO	TMOQ
DC1	17	2487
DC2	20	3284
DC3	5	816
DC4	13	1894
DC5	12	2224

Our proposed method adds a new dimension to the supply chain modeling approaches. It can empower analyst to minimize the risk and cost of the changes in the IRP. In proposed method, total routing cost is 90010.83. In addition, the other cost component analysis for each supply chain member is evaluated in Table 4.22. In general, the highest share of cost component is average holding cost while the smallest share of cost component is lost sales cost. In DC3, the smallest share of cost component is lost sales cost and processing cost.

Table 4.22. Cost component analysis for supply chain member

		Average Holding Cost	Order Cost Per Use	Lost Sales Cost	Order Processing Cost	Processing Cost
DC1	Cost ($)	5,545	1,283	0	1,684	564
	Share (%)	61	14	0	19	6
DC2	Cost ($)	4,818	1,510	0	2,058	953
	Share (%)	52	16	0	22	10
DC3	Cost ($)	7,937	364	0	495	0
	Share (%)	90	4	0	6	0
DC4	Cost ($)	5,062	976	0	1,303	309
	Share (%)	66	13	0	17	4
DC5	Cost ($)	6,381	917	0	1,220	430
	Share (%)	71	10	0	14	5

5. CONCLUSION AND RECOMMENDATIONS

In today's world, companies need to analyze the supply chain due to the increasing competition in local and global economies. Especially, integrating inventory and routing decisions has become an important necessity in supply chain. It is very critical for any company to cope with its competitors in changing business environment. However, there is no standard method to handle the IRP. A detailed literature review showed that big data analysis, simulation, optimization methods, and AI based methods play a significant role on the solving IRP. Therefore, new hybrid methods are developed to solve the stochastic and dynamic IRP. They add a new dimension to the supply chain modeling approaches. They also provide an attractive opportunity to represent the ambiguity in supply chain with real life uncertainty.

Social media users have been growing exponentially in recent years. This growth has evoked researchers to analyze the social media and customer sentiment because they offer significant opportunities to advance business intelligence in supply chain. The customers' review can be integrated with predictive models in supply chain. However, supply chain members are struggling in understanding the general sentiments in today's world. Therefore, various methods are used to give valuable insights in supply chain. Customer reviews include a lot of information about services and products. Positive comments can influence other customer' willingness to buy. On the other hand, negative comments can have a negative impact on the customers. The challenge for companies is to benefit from the customer reviews without becoming overwhelmed. This study introduces a new methodology that includes sentiment analysis and SO to solve the IRP.

5.1. Conclusion

Providing optimal routing and inventory control system is a crucial foundation for achieving both strategically and tactically success in supply chain.

At this point, it was deemed crucial to determine proper strategy for each demand pattern to satisfy objective function. The determining proper strategy makes a strategic difference to a company's ability to achieve the supply chain objectives.

This study is focused on creating new methodologies to solve the IRP. It includes two parts. SO and AI based simulation are used in the first part while SO and big data analysis are used in the second part.

In the first part, a new methodology that includes the integration of optimization and simulation and also AI and simulation yields a tremendous improvement in inventory control system and routing. In addition, it provides competitive advantage under dynamic and stochastic IRP.

In the second part, big data and sentiment analysis are employed to provide decision support during the determination of investments, marketing strategies, and products to be developed by companies. Big data analysis can offer unique information related to market trends, purchasing, and other business activities.

The main conclusions of the study can be summarized as follows:

i. The highest share of cost component in erratic demand is order processing cost. According to cost based analysis, either average holding cost or lost sales cost has the highest share among the cost components of intermittent demand and lumpy demand.

ii. Strategy 1 can be selected for lumpy demand patterns when total routing cost is considered. Strategy 2 gives better results for erratic demand pattern with respect to the total routing cost. When total routing cost is taken into account, Strategy 3 can be chosen for intermittent demand.

iii. Strategy 2 can be selected for lumpy demand patterns when the TSCC is considered. For intermittent demand and erratic demand pattern, Strategy 3 can be chosen when the TSCC is taken into account.

iv. Strategy 2 gives better results in lumpy demand when the results of total PLOQ and total TLOQ are taken into account. On the other hand, Strategy

3 gives better results with respect to the results of total PLOQ and total TLOQ in erratic demand and intermittent demand.

v. Strategy 2 gives better results in lumpy demand when the results of total TMOQ and total NTMO are taken into account. On the other hand, Strategy 3 gives better results with respect to the results of total TMOQ and total NTMO in erratic demand and intermittent demand.

vi. Strategy 2 can be chosen in erratic demand pattern according to the mean of order met probabilities during lead time. However, Strategy 3 can be selected for intermittent demand and lumpy demand according to the mean of order met probabilities during lead time.

vii. Strategy 1 and Strategy 2 give better results than Strategy 3 in erratic demand according to the mean space utilization per delivery. Strategy 1 can be chosen in lumpy demand according to the mean space utilization per delivery. In intermittent demand, Strategy 3 can be selected with respect to the mean space utilization per delivery.

viii. In general, the highest share of cost component is average holding cost while the smallest share of cost component is lost sales cost in the second part (includes big data analysis and NSGA-II based SO).

5.2. Recommendations

The proposed method can encourage further development of the hybrid SO and AI based simulations. The recommendations can be summarized as follows:

i. The parameters of the ANN can be determined by integrating other AI methods including symbolic AI and computational intelligence.

ii. Proposed SO can be compared with other types of SO including gradient based search methods, stochastic optimization, response surface methodology, heuristic methods, A-teams, and statistical methods.

iii. Machine learning approach (e.g. supervised learning and unsupervised learning) can be utilized to determine the sentiment score for customer reviews and to compare the results of SentiStrength.

iv. Different routing policies can be applied for proposed system. In addition, meta-heuristic and heuristic can be employed to determine the route.

v. The proposed methods can be applied to other problems, such as the dynamic vehicle routing problem, multi-echelon supply chain, and dual-channel supply chain.

REFERENCES

Abdollahi, M., Arvan, M., Omidvar, A., and Ameri, F., 2014. A simulation optimization approach to apply value at risk analysis on the inventory routing problem with backlogged demand. International Journal of Industrial Engineering Computations, 5(4), 603−620.

Ada E., 2010. Tedarik zincirinde toplam kalite yönetimi. Yüksek Lisans Tezi.

Ahn, H.-I., and Spangler, W. S., 2014. Sales prediction with social media analysis. In 2014 Annual SRII Global Conference, p. 213−222.

Amin-Naseri, M. R., and Tabar, B. R., 2008. Neural network approach to lumpy demand forecasting for spare parts in process industries. In 2008 International Conference on Computer and Communication Engineering, Kuala Lumpur, Malaysia, p. 1378−1382.

Ammeri, A., Hachicha, W., Chabchoub, H., and Masmoudi, F., 2011. A comprehensive litterature review of mono-objective simulation optimization methods. Advances in Production Engineering & Management, 6(4), 291–302.

Andersson, H., Hoff, A., Christiansen, M., Hasle, G., and Løkketangen, A., 2010. Industrial aspects and literature survey: Combined inventory management and routing. Computers & Operations Research, 37(9), 1515−1536.

Andersson, J., and Jonsson, P., 2018. Big data in spare parts supply chains: The potential of using product-in-use data in aftermarket demand planning. International Journal of Physical Distribution & Logistics Management, 48(5), 524−544.

Archetti, C., Christiansen, M., and Speranza, M. G., 2018. Inventory routing with pickups and deliveries. European Journal of Operational Research, 268(1), 314−324.

Archetti, C., and Speranza, M. G., 2016. The inventory routing problem: the value of integration. International Transactions in Operational Research, 23(3), 393−407.

Arunachalam, D., Kumar, N., and Kawalek, J. P., 2018. Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. Transportation Research Part E: Logistics and Transportation Review, 114, 416−436.

Azadeh, A., Elahi, S., Farahani, M. H., and Nasirian, B., 2017. A genetic algorithm−Taguchi based approach to inventory routing problem of a single perishable product with transshipment. Computers & Industrial Engineering, 104, 124−133.

Balaji, S. M. E., 2015. Predicting sentiment analysis from online reviews. International Research Journal of Engineering and Technology, 2(9), 528−531.

Balamurugan, T., Karunamoorthy, L., Arunkumar, N., and Santhosh, D., 2018. Optimization of inventory routing problem to minimize carbon dioxide emission. International Journal of Simulation Modelling, 17(1), 42−54.

Bartezzaghi, E., Verganti, R., and Zotteri, G., 1999. A simulation framework for forecasting uncertain lumpy demand. International Journal of Production Economics, 59(1−3), 499−510.

Bhadouria, S., and Jayant, A., 2017. Development of ANN models for demand forecasting. American Journal of Engineering Research, 6(12), 142–147.

Ben-Hur, A., and Weston, J., 2010. A user's guide to support vector machines. In Data Mining Techniques for the Life Sciences, Methods in Molecular Biology (Methods and Protocols), vol. 609, Humana Press, p. 223−239.

Berling, P., and Marklund, J., 2013. A model for heuristic coordination of real life distribution inventory systems with lumpy demand. European Journal of Operational Research, 230(3), 515−526.

Biswas, S., and Sen, J., 2016. A proposed architecture for big data driven supply chain analytics. SSRN Electronic Journal 13(3). doi: 10.2139/ssrn.2795906.

Braekers, K., Ramaekers, K., and Van Nieuwenhuyse, I., 2016. The vehicle routing problem: State of the art classification and review. Computers & Industrial Engineering, 99, 300-313.

Brinch, M., Stentoft, J., and Jensen, J. K., 2017. Big data and its applications in supply chain management: findings from a Delphi study. In Proceedings of the 50th Hawaii International Conference on System Sciences, p. 1351−1360.

Cáceres-Cruz, J., Juan, A. A., Bektas, T., Grasman, S. E., and Faulin, J., 2012. Combining Monte Carlo simulation with heuristics for solving the inventory routing problem with stochastic demands. In Proceedings of the 2012 Winter Simulation Conference, Berlin, p. 1−9.

Campbell, A., Clarke, L., Kleywegt, A., and Savelsbergh, M., 1998. The inventory routing problem. In Fleet Management and Logistics, Springer, Boston, MA, p. 95−113.

Campbell, A. M., and Savelsbergh, M. W. P., 2004. A decomposition approach for the inventory-routing problem. Transportation Science, 38(4), 488−502.

Carson, Y., and Maria, A., 1997. Simulation optimization: methods and applications. In Winter Simulation Conference Proceedings, Atlanta, GA, USA, p. 118−126.

Chae, B. K., 2015. Insights from hashtag# supplychain and twitter analytics: considering twitter and twitter data for supply chain practice and research. International Journal of Production Economics, 165, 247−259.

Chen, C.-H., and Lee, L. H., 2011. Stochastic simulation optimization: an optimal computing budget allocation. System Engineering and Operations Research-vol. 1, World scientific, p. 227.

Chen, D., Sain, S. L., and Guo, K., 2012. Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. Journal of Database Marketing & Customer Strategy Management, 19, 197–208.

Cheng, C.-Y., Chiang, K.-L., and Chen, M.-Y., 2016. Intermittent demand forecasting in a tertiary pediatric intensive care unit. Journal of Medical Systems, 40(10), 217.

Cheng, C., Yang, P., Qi, M., and Rousseau, L.-M., 2017. Modeling a green inventory routing problem with a heterogeneous fleet. Transportation Research Part E: Logistics and Transportation Review, 97, 97−112.

Chopra, S., and Meindl, P., 2007. Supply chain management, strategy, planning, and operation. Third ed. Upper Saddle River, N.J.: Pearson Prentice Hall, p. 536.

Coelho, L. C., Cordeau, J.-F., and Laporte, G., 2012. The inventory−routing problem with transshipment. Computers & Operations Research, 39(11), 2537−2548.

Coelho, L. C., and Laporte, G., 2013a. A branch−and−cut algorithm for the multi−product multi−vehicle inventory−routing problem. International Journal of Production Research, 51(23−24), 7156−7169.

Coelho, L. C., and Laporte, G., 2013b. The exact solution of several classes of inventory−routing problems. Computers & Operations Research, 40(2), 558−565.

Cope, D., Fayez, M. S., Mollaghasemi, M., and Kaylani, A., 2007. Supply chain simulation modeling made easy: an innovative approach. In Proceedings of the 2007 Winter Simulation Conference, p. 1887−1896.

Cordeau, J.-F., Gendreau, M., Laporte, G., Potvin, J.-Y., and Semet, F., 2002. A guide to vehicle routing heuristics. Journal of the Operational Research Society, 53(5), 512−522.

Cordeau, J.-F., Laporte, G., Savelsbergh, M. W. P., and Vigo, D. 2007. Vehicle routing. Handbooks in Operations Research and Management Science, 14, 367−428.

Costantino, F., Di Gravio, G., Patriarca, R., and Petrella, L. 2018. Spare parts management for irregular demand items. Omega, 81, 57−66.

Cros, M.-J., Garcia, F., Martin-Clouaire, R., and Rellier J.-P. Çevirmen: Yildiz, A. K., 2006. 3. Bölüm: Yöntemler, Algoritmalar ve Yazılım, 3.1 Modelleme ve Simülasyon. In CIGR Handbook of Agricultural Engineering Volume VI Information Technology, p. 118-133.

Dabiri, N., Tarokh, M. J., and Alinaghian, M., 2017. New mathematical model for the bi-objective inventory routing problem with a step cost function: A multi-objective particle swarm optimization solution approach. Applied Mathematical Modelling, 49, 302−318.

Dastidar, A. G., 2017. Intermittent demand forecasting for long tail SKUs. In 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), Zeal Education Society, Pune, India, p. 334−339.

Dey, L., Chakraborty, S., Biswas, A., Bose, B., and Tiwari, S., 2016. Sentiment analysis of review datasets using Naive Bayes and K-NN classifier. International Journal of Information Engineering and Electronic Business, 4, 54–62.

Doukidis, G. I., and Angelides, M. C., 1994. A framework for integrating artificial intelligence and simulation. Artificial Intelligence Review, 8, 55–85.

Dubedout, H., Dejax, P., Neagu, N., and Yeung, T., 2012. A GRASP for real life inventory routing problem: application to bulk gas distribution. In 9th International Conference on Modeling, Optimization & Simulation, Bordeaux, France.

Edison, M., and Aloysius, A., 2017. Lexicon based acronyms and emoticons classification of sentiment analysis (SA) on big data. International Journal of Database Theory and Application, 10(7), 41−54.

Efendigil, T., Önüt, S., and Kahraman, C., 2009. A decision support system for demand forecasting with artificial neural networks and neuro−fuzzy models: A comparative analysis. Expert Systems with Applications, 36, 6697–6707.

Elmas, Ç., 2016. Yapay Zekâ Uygulamaları. Baskı, Ankara, Seçkin Yayıncılık, p. 479.

Fan, Z.-P., Che, Y.-J., and Chen, Z.-Y., 2017. Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. Journal of Business Research, 74, 90−100.

Feng, D.-Z., Yamashiro, M., and Zhang, L.-B., 2003. Optimal ordering and production policy for a two-stage production delivery system under lumpy demand. In Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, p. 833−837.

Figueira, G., and Almada-Lobo, B., 2014. Hybrid simulation-optimization methods: A taxonomy and discussion. Simulation Modelling Practice and Theory, 46, 118−134.

Gaikar, D., and Marakarkandy, B., 2015. Product sales prediction based on sentiment analysis using twitter data. International Journal of Computer Science and Information Technologies, 6(3), 2303−2313.

Gaur, V., and Fisher, M. L., 2004. A periodic inventory routing problem at a supermarket chain. Operations Research, 52(6), 813−822.

Gendreau, M., Laporte, G., and Séguin, R., 1996. Stochastic vehicle routing. European Journal of Operational Research, 88(1), 3−12.

Gholamian, M. R., and Heydari, M., 2017. An inventory model with METRIC approach in location−routing−inventory problem. Advances in Production Engineering & Management, 12(2), 115−126.

Ghiani, G., Legato, P., Musmanno, R., and Vocaturo, F., 2004. Optimization via simulation: solution concepts, algorithms, parallel computing strategies and commercial software. International Journal of Computing, 3(3), 7−12.

Göçken, M., Dosdoğru, A. T., Boru, A., and Geyik, F., 2017. Characterizing continuous (s, S) policy with supplier selection using simulation optimization. Simulation, 93(5), 379−396.

Groves, W., Collins, J., Gini, M., and Ketter, W., 2014. Agent−assisted supply chain management: Analysis and lessons learned. Decision Support Systems, 57, 274−284.

Gruler, A., Panadero, J., de Armas, J., Pérez, J. A. M., and Juan, A. A., 2020. A variable neighborhood search simheuristic for the multiperiod inventory routing problem with stochastic demands. International Transactions in Operational Research, 27(1), 314–335.

Gumus, A. T., Guneri, A. F., and Keles, S., 2009. Supply chain network design using an integrated neuro-fuzzy and MILP approach: A comparative design study. Expert Systems with Applications, 36, 12570–12577.

Gutierrez, R. S., Solis, A. O., and Bendore, N. R., 2004. Lumpy demand characterization and forecasting performance: an exploratory case study. http://utminers.utep.edu/solis/wdsi%202004,%202006%20&%202007/Lumpy%20demand%20forecasting%20WDSI%202004%20proceedings.pdf. (Access Date: 31 December 2019).

Hajghasem, M., and Shojaie, A. A., 2016. Optimal routing in supply chain aimed at minimizing vehicle cost and supply. Procedia Economics and Finance, 36, 353−362.

Hasni, S., Toumi, S., Jarboui, B., and Mjirda, A., 2017. GVNS based heuristic for solving the multi-product multi-vehicle inventory routing problem. Electronic Notes in Discrete Mathematics, 58, 71−78.

Hofmann, E., 2017. Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect. International Journal of Production Research, 55(17), 5108−5126.

Hollier, R. H., Mak, K. L., and Yiu, K. F. C., 2005. Optimal inventory control of lumpy demand items using (s, S) policies with a maximum issue quantity restriction and opportunistic replenishments. International Journal of Production Research, 43(23), 4929−4944.

Hu, H., Li, J., and Li, X., 2018. A credibilistic goal programming model for inventory routing problem with hazardous materials. Soft Computing, 22(17), 5803−5816.

Hua, Z. S., Zhang, B., Yang, J., and Tan, D. S., 2007. A new approach of forecasting intermittent demand for spare parts inventories in the process industries. Journal of the Operational Research Society, 58(1), 52−61.

Huang, B., and Ye, Z., 2010. The effects of lumpy demand and shipment size constraint: A response to "Revisit the note on supply chain integration in vendor-managed inventory". Decision Support Systems, 48(2), 421−425.

Iassinovskaia, G., Limbourg, S., and Riane, F., 2017. The inventory-routing problem of returnable transport items with time windows and simultaneous pickup and delivery in closed-loop supply chains. International Journal of Production Economics, 183, 570−582.

Işıklı, Ş., 2014. Büyük veri, epistemoloji ve etik tartışmalar. AJIT−e: Online Academic Journal of Information Technology, 5(17), 89−122.

Jadav, B. M., and Vaghela, V. B., 2016. Sentiment analysis using support vector machine based on feature selection and semantic analysis. International Journal of Computer Applications, 146(13), 26−30.

Jafarkhan, F., and Yaghoubi, S., 2018. An efficient solution method for the flexible and robust inventory-routing of red blood cells. Computers & Industrial Engineering, 117, 191−206.

Jain, A. P., and Dandannavar, P., 2016. Application of machine learning techniques to sentiment analysis. In 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Bangalore, p. 628−632.

Jaipuria, S., and Mahapatra, S. S., 2014. An improved demand forecasting method to reduce bullwhip effect in supply chains. Expert Systems with Applications, 41, 2395–2408.

Jarugumilli, S., Grasman, S. E., and Ramakrishnan, S., 2006. A simulation framework for real-time management and control of inventory routing decisions. In Proceedings of the 2006 Winter Simulation Conference, Monterey, CA, p. 1485–1492.

Jiang, P., Liu, X., Huang, Y., and Yuan, Y., 2017. An adaptive data-driven model for erratic demand forecasting. IFAC-PapersOnLine, 50(1), 9020–9025.

Juan, A. A., Grasman, S. E., Cáceres-Cruz, J., and Bektaş, T., 2014. A simheuristic algorithm for the single-period stochastic inventory-routing problem with stock-outs. Simulation Modelling Practice and Theory, 46, 40–52.

Jung, G., Park, J., Kim, Y., and Kim, Y. B., 2017. A modified bootstrap method for intermittent demand forecasting for rare spare parts. International Journal of Industrial Engineering, 24(2), 245–254.

Jurafsky, D., and Martin, J. H., 2015. Speech and language processing. Chapter 7: Classification: Naive Bayes, Logistic Regression, Sentiment. https://pdfs.semanticscholar.org/cbcb/1cf6d7d71e402c3ee742cb3700446fb05aaa.pdf (Access Date: 31 December 2019).

Kaur, H., and Singh, S. P., 2018. Heuristic modeling for sustainable procurement and logistics in a supply chain using big data. Computers & Operations Research, 98, 301–321.

Kaya, G. O., and Turkyilmaz, A., 2018. Intermittent demand forecasting using data mining techniques. Applied Computer Science, 14(2), 38–47.

Kelton, W. D., Smith, J. S., Sturrock, D. T., Göçken, M., and Dosdoğru, A. T., 2015. Simio & simulation: modeling, analysis, applications: Turkish Translation. 3rd edition, Simio LLC, p. 438.

Khader, M., Awajan, A., and Al-Naymat, G., 2018. Sentiment analysis based on MapReduce: A survey. In Proceedings of the 10th International Conference on Advances in Information Technology, p. 1–8.

Kleijnen, J. P. C., and Smits, M. T., 2003. Performance metrics in supply chain management. Journal of the Operational Research Society, 54(5), 507−514.

Kleywegt, A. J., Nori, V. S., and Savelsbergh, M. W. P., 2002. The stochastic inventory routing problem with direct deliveries. Transportation Science, 36(1), 94−118.

Kocer, U. U., 2013. Forecasting intermittent demand by Markov chain model. International Journal of Innovative Computing, Information and Control, 9(8), 3307−3318.

Koç, Ç., Bektaş, T., Jabali, O., and Laporte, G., 2016. Thirty years of heterogeneous vehicle routing. European Journal of Operational Research, 249(1), 1-21.

Kourentzes, N., 2013. Intermittent demand forecasts with neural networks. International Journal of Production Economics, 143(1), 198−206.

Kumar, K. L. S., Desai, J., and Majumdar, J., 2016. Opinion mining and sentiment analysis on online customer review. In 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, p. 1−4.

Lau, R. Y. K., Zhang, W., and Xu, W., 2018. Parallel aspect-oriented sentiment analysis for sales forecasting with big data. Production and Operations Management, 27(10), 1775−1794.

Lee, K., Cho, H., and Jung, M., 2014. Simultaneous control of vehicle routing and inventory for dynamic inbound supply chain. Computers in Industry, 65(6), 1001−1008.

Lefever, W., Aghezzaf, E.-H., Hadj-Hamou, K., and Penz, B., 2018. Analysis of an improved branch-and-cut formulation for the inventory-routing problem with transshipment. Computers & Operations Research, 98, 137−148.

Lei, M., Li, S., and Tan, Q., 2016. Intermittent demand forecasting with fuzzy markov chain and multi aggregation prediction algorithm. Journal of Intelligent & Fuzzy Systems, 31(6), 2911−2918.

Li, Y., Chu, F., Feng, C., Chu, C., and Zhou, M., 2018. Integrated production inventory routing planning for intelligent food logistics systems. IEEE Transactions on Intelligent Transportation Systems, 20(3), 867−878.

Li, C., and Lim, A., 2018. A greedy aggregation–decomposition method for intermittent demand forecasting in fashion retailing. European Journal of Operational Research, 269(3), 860−869.

Liao, L., Li, J., and Wu, Y., 2013. Modeling and optimization of inventory-distribution routing problem for agriculture products supply chain. Hindawi Publishing Corporation Discrete Dynamics in Nature and Society, Volume 2013, Article ID 409869, 12 pages.

Liu, B., and Esogbue, A. O., 1999. Decision criteria and optimal inventory processes. Springer Science & Business Media, p. 210.

Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T., and Gucci, S., 2017. Single-hidden layer neural networks for forecasting intermittent demand. International Journal of Production Economics, 183, 116−128.

Maria, A., 1997. Introduction to modeling and simulation. In Proceedings of the 1997 Winter Simulation Conference, p.7−13.

Medhat, W., Hassan, A., and Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093−1113.

Mercado, E. C., 2008. Hands-on inventory management. Auerbach Publications, Taylor & Francis Group, p. 110.

Mes, M., Schutten, M., and Rivera, A. P., 2014. Inventory routing for dynamic waste collection. Waste management, 34(9), 1564−1576.

Miragliotta, G., and Staudacher, A. P., 2004. Exploiting information sharing, stock management and capacity oversizing in the management of lumpy demand. International Journal of Production Research, 42(13), 2533−2554.

Mishra, D., Gunasekaran, A., Papadopoulos, T., and Childe, S. J., 2018. Big Data and supply chain management: a review and bibliometric analysis. Annals of Operations Research, 270, 313−336.

Molina, A., Ponte, B., Parreño, J., De la Fuente, D., and Costas, J., 2016. Forecasting erratic demand of medicines in a public hospital: A comparison of artificial neural networks and ARIMA models. In Proceedings on the International Conference on Artificial Intelligence (ICAI), p. 401−406.

Montagné, R., Gamache, M., and Gendreau, M., 2019. A shortest path-based algorithm for the inventory routing problem of waste vegetable oil collection. Journal of the Operational Research Society, 70(6), 986−997.

Montoya-Torres, J. R., Franco, J. L., Isaza, S. N., Jiménez, H. F., and Herazo-Padilla, N., 2015. A literature review on the vehicle routing problem with multiple depots. Computers & Industrial Engineering, 79, 115-129.

Morales, F., Franco, C., and Mendez-Giraldo, G., 2018. Dynamic inventory routing problem: Policies considering network disruptions. International Journal of Industrial Engineering Computations, 9(4), 523−534.

Mukhopadhyay, S., Solis, A. O., and Gutierrez, R. S., 2012. The accuracy of non-traditional versus traditional methods of forecasting lumpy demand. Journal of Forecasting, 31(8), 721−735.

Niakan, F., and Rahimi, M., 2015. A multi-objective healthcare inventory routing problem; a fuzzy possibilistic approach. Transportation Research Part E: Logistics and Transportation Review, 80, 74−94.

Nikolakopoulos, A., and Ganas, I., 2017. Economic model predictive inventory routing and control. Central European Journal of Operations Research, 25, 587−609.

Nilsson, N. J., 1998. Artificial intelligence: a new synthesis. Morgan Kaufmann, China Machine Press, p. 513.

Onyeocha, C. E., Khoury, J., and Geraghty, J., 2015. Evaluation of multi-product lean manufacturing systems with setup and erratic demand. Computers & Industrial Engineering, 87, 465−480.

Othman, S. N., Mustaffa, N. H., Radzi, N. H. M., Sallehuddin, R., and Bazin, N. E. N., 2016. A modelling of genetic algorithm for inventory routing problem simulation optimisation. International Journal of Supply Chain Management, 5(4), 43−51.

Pai, P.-F., and Liu, C.-H., 2018. Predicting vehicle sales by sentiment analysis of Twitter data and stock market values. IEEE Access, 6, 57655−57662.

Pegden, C. D., 2007. SIMIO: a new simulation system based on intelligent objects. In 2007 Winter Simulation Conference, Washington, DC, p. 2293−2300.

Pegden, C. D., and Sturrock, D. T., 2010. Introduction to Simio. In Proceedings of the 2010 Winter Simulation Conference, Baltimore, MD, p. 1−10.

Pierre, B., Vannieuwenhuyse, B., Dominanta, D., and Van Dessel, H., 2003. Dynamic ABC storage policy in erratic demand environments. Jurnal Teknik Industri, 5(1), 1−12.

Popović, D., Vidović, M., and Radivojević, G., 2012. Variable neighborhood search heuristic for the inventory routing problem in fuel delivery. Expert Systems with Applications, 39(18), 13390−13398.

Prestwich, S., Rossi, R., Tarim, S. A., and Hnich, B., 2014. Mean-based error measures for intermittent demand forecasting. International Journal of Production Research, 52(22), 6782−6791.

Prodhon, C., and Prins, C., 2016. Metaheuristics for vehicle routing problems. In Metaheuristics, 407−437.

Qin, L., Miao, L., Ruan, Q., and Zhang, Y., 2014. A local search method for periodic inventory routing problem. Expert Systems with Applications, 41(2), 765−778.

Raghuwanshi, A. S., and Pawar, S. K., 2017. Polarity classification of twitter data using sentiment analysis. International Journal on Recent and Innovation Trends in Computing and Communication, 5(6), 434−439.

Rahimi, M., Baboli, A., and Rekik, Y., 2017. Multi-objective inventory routing problem: A stochastic model to consider profit, service level and green criteria. Transportation Research Part E: Logistics and Transportation Review, 101, 59−83.

Ramadhan, F., Imran, A., and Rizana, A. F., 2018. A record-to-record travel algorithm for multi-product and multi-period inventory routing problem. In 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), p. 274−278.

Ramaekers, K., and Janssens, G. K., 2014. Optimal policies for demand forecasting and inventory management of goods with intermittent demand. Journal of Applied Operational Research, 6(2), 111–123.

Rani, S., and Singh, J., 2017. Sentiment analysis of tweets using support vector machine. International Journal of Computer Science and Mobile Applications, 5(10), 83−91.

Rayat, F., Musavi, M., and Bozorgi-Amiri, A., 2017. Bi-objective reliable location-inventory-routing problem with partial backordering under disruption risks: A modified AMOSA approach. Applied Soft Computing, 59, 622−643.

Reiman, M. I., Rubio, R., and Wein, L. M., 1999. Heavy traffic analysis of the dynamic stochastic inventory-routing problem. Transportation Science, 33(4), 361−380.

Robak, S., Franczyk, B., and Robak, M., 2014. Research problems associated with big data utilization in logistics and supply chains design and management. In Position papers of the 2014 Federated Conference on Computer Science and Information Systems, p. 245–249.

Roldán, R. F., Basagoiti, R., and Coelho, L. C., 2016. Robustness of inventory replenishment and customer selection policies for the dynamic and stochastic inventory-routing problem. Computers and Operations Research, 74, 14−20.

Rothenberg, J., 1991. Tutorial: artificial intelligence and simulation. In Proceedings of the 1991 Winter Simulation Conference, p. 218−222.

Rozados, I. V., and Tjahjono, B., 2014. Big data analytics in supply chain management: Trends and related research. In 6th International Conference on Operations and Supply Chain Management, Bali, doi: 10.13140/RG.2.1.4935.2563.

Sarkar, D., and Modak, J. M., 2005. Pareto-optimal solutions for multi-objective optimization of fed-batch bioreactors using nondominated sorting genetic algorithm. Chemical Engineering Science, 60, 481–492.

Sathi, A., 2012. Big Data analytics: disruptive technologies for changing the game. Mc Press, p.77.

Sayarshad, H. R., and Gao, H. O., 2018. A non-myopic dynamic inventory routing and pricing problem. Transportation Research Part E: Logistics and Transportation Review, 109, 83−98.

Schneider, M. J., and Gupta, S., 2016. Forecasting sales of new and existing products using consumer reviews: A random projections approach. International Journal of Forecasting, 32(2), 243−256.

Seleim, A., Azab, A., and AlGeddawy, T., 2012. Simulation methods for changeable manufacturing. Procedia CIRP, 3, 179−184.

Shaabani, H., and Kamalabadi, I. N., 2016. An efficient population-based simulated annealing algorithm for the multi-product multi-retailer perishable inventory routing problem. Computers & Industrial Engineering, 99, 189−201.

Shale, E. A., Boylan, J. E., and Johnston, F. R., 2006. Forecasting for intermittent demand: the estimation of an unbiased average. Journal of the Operational Research Society, 57, 588−592.

Shukla, N., Tiwaric, M. K., and Ceglarek, D., 2013. Genetic-algorithms-based algorithm portfolio for inventory routing problem with stochastic demand. International Journal of Production Research, 51(1), 118–137.

Silver, E. A., Ho, C.-M., and Deemer, R. L., 1971. Cost-minimizing inventory control of items having a special type of erratic demand pattern. INFOR: Information Systems and Operational Research, 9(3), 198−219.

Singla, Z., Randhawa, S., and Jain, S., 2017. Sentiment analysis of customer product reviews using machine learning. In 2017 International Conference on Intelligent Computing and Control (I2C2), Coimbatore, p. 1−5.

Spall, J. C., 2003. Introduction to stochastic search and optimization: estimation, simulation, and control. John Wiley & Sons, p. 595.

Swain, A. K., and Cao, R. Q., 2019. Using sentiment analysis to improve supply chain intelligence. Information Systems Frontiers, 21(2), 469–484.

Syntetos, A. A., Boylan, J. E., and Croston, J. D., 2005. On the categorization of demand patterns. Journal of the Operational Research Society, 56, 495–503.

Tarokh, M. J., and Golkar, M., 2006. Supply chain simulation methods. In 2006 IEEE International Conference on Service Operations and Logistics, and Informatics, Shanghai, p. 448−454.

Taylor III, B. W., 2013. Introduction to management science. Pearson Education, Inc.

Teunter, R. H., and Duncan, L., 2009. Forecasting intermittent demand: a comparative study. Journal of the Operational Research Society, 60(3), 321−329.

Thelwall, M., Buckley, K., and Paltoglou, G., 2012. Sentiment strength detection for the social Web. Journal of the American Society for Information Science and Technology, 63(1), 163−173.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A., 2010. Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 61(12), 2544–2558.

Timajchi, A., Al-e-Hashem, S. M. J. M., and Rekik, Y., 2019. Inventory routing problem for hazardous and deteriorating items in the presence of accident risk with transshipment option. International Journal of Production Economics, 209, 302−315.

Tiwari, S., Wee, H. M., and Daryanto, Y., 2018. Big data analytics in supply chain management between 2010 and 2016: Insights to industries. Computers & Industrial Engineering, 115, 319−330.

Toth, P., & Vigo, D., 2002. Models, relaxations and exact approaches for the capacitated vehicle routing problem. Discrete Applied Mathematics, 123(1-3), 487-512.

Tyagi, A., and Sharma, N., 2018. Sentiment analysis using logistic regression and effective word score heuristic. International Journal of Engineering & Technology, 7(2.24), 20−23.

Verganti, R., 1997. Order overplanning with uncertain lumpy demand: a simplified theory. International Journal of Production Research, 35(12), 3229−3248.

Wamba, S. F., and Akter, S., 2015. Big data analytics for supply chain management: A literature review and research agenda. In Workshop on Enterprise and Organizational Modeling and Simulation, p. 61−72.

Wang, G., Gunasekaran, A., Ngai, E. W. T., and Papadopoulos, T., 2016. Big data analytics in logistics and supply chain management: certain investigations for research and applications. International Journal of Production Economics, 176, 98−110.

Wang, L., and Alexander, C. A., 2015. Big data driven supply chain management and business administration. American Journal of Economics and Business Administration, 7(2), 60−67.

Wang, S., Tao, F., and Shi, Y., 2018. Optimization of inventory routing problem in refined oil logistics with the perspective of carbon tax. Energies, 11(6), 1437.

Willemain, T. R., Smart, C. N., and Schwarz, H. F., 2004. A new approach to forecasting intermittent demand for service parts inventories. International Journal of Forecasting, 20, 375–387.

Wirth, N., 2018. Hello marketing, what can artificial intelligence help you with?. International Journal of Market Research, 60(5), 435−438.

Wong, W. K., Guo, Z. X., and Leung, S. Y. S., 2013. Optimizing decision making in the apparel supply chain using artificial intelligence (AI): from production to retail. Woodhead Publishing Series in Textiles: Number 143, p. 231.

Xiao, N., and Rao, Y. L., 2016. Multi-product multi-period inventory routing optimization with time window constrains. International Journal of Simulation Modelling, 15(2), 352−364.

Xie, B., and Hu, X., 2009. A model and heuristic algorithm for inventory routing problem in emergency logistics. In ICCTP 2009: Critical Issues in Transportation Systems Planning, Development, and Management, p. 3147−3153.

Yu, X., Liu, Y., Huang, J. X., and An, A., 2012. Mining online reviews for predicting sales performance: A case study in the movie domain. IEEE Transactions on Knowledge and Data engineering, 24(4), 720−734.

Yusoff, Y., Ngadiman, M. S., and Syntetos, A. M., 2011. Overview of NSGA-II for optimizing machining process parameters. Procedia Engineering, 15, 3978–3983.

Zain, A. M., Haron, H., Qasem, S. N., and Sharif, S., 2012. Regression and ANN models for estimating minimum value of machining performance. Applied Mathematical Modelling, 36(4), 1477−1492.

Zainuddin, N., and Selamat, A., 2014. Sentiment analysis using support vector machine. In 2014 International Conference on Computer, Communications, and Control Technology (I4CT), p. 333−337.

Zhan, Y., and Tan, K. H., 2020. An analytic infrastructure for harvesting big data to enhance supply chain performance. European Journal of Operational Research, 281(3), 559−574.

Zhang, R., Bao, Y., and Zhang, J., 2010. Forecasting erratic demand by support vector machines with ensemble empirical mode decomposition. In The 3rd International Conference on Information Sciences and Interaction Sciences, Chengdu, p. 567−571.

Zhong, Y., and Aghezzaf, E.-H., 2011. Combining DC-programming and steepest-descent to solve the single-vehicle inventory routing problem. Computers & Industrial Engineering, 61(2), 313−321.

CURRICULUM VITAE

Aslı BORU İPEK was born on 24.08.1989. She completed her primary, secondary and high school education in Bursa. She received her Bachelor of Science in Textile Engineering and Industrial Engineering from the Gaziantep University. She is a PhD candidate and a Research Assistant in Adana Alparslan Türkeş Science and Technology University. She holds a MSc degree from Adana Science and Technology University and Çukurova University (Joint Master of Science Program). She has studied simulation optimization, meta-heuristics, and supply chain management. Her email is aboru@atu.edu.tr.