Measuring the effectiveness of allocation algorithms by means of simulation modelling

Elmien Thom

Thesis presented in partial fulfilment of the requirements for the degree of
Master of Commerce
in the Faculty of Economic and Management Sciences at Stellenbosch University

Supervisor: Prof SE Visagie

March 2016

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2016

Abstract

The allocation of stock to stores is one of the most important processes in the management of a retail chain. In the clothing industry, allocation decisions include, amongst other, the determination of the number of each size (for example small, medium and large) to send to each store. A case study of this problem in Pep Stores Ltd. (PEP), a major retailer in South Africa, is considered.

In PEP, products are ordered from factories about seven months before they are available in the stores. They are then shipped to the distribution centra, after which they are distributed per road to the stores. Before the products are ordered, preliminary allocation decisions are made. Once the stock arrives at the distribution centra, decisions about the allocation of products and sizes to the stores are finalised. Allocation decisions are adjusted throughout the season as more recent sales data become available.

In this thesis, simulation models are developed to compare four allocation methods in terms of total expected sales, shortages and surpluses. The algorithms include PEP's current algorithm, an existing algorithm that minimises the expected number of weeks that shortages and surpluses occur at stores, a new algorithm with the objective to maximise expected sales, and a relaxation of the new algorithm.

The simulation models are developed according to two modelling approaches. Each approach is applied to Summer and Winter products, resulting in four simulation models. The two simulation approaches deliver similar results for both Summer and Winter products, namely that all four allocation methods are approximately equally effective.

Opsomming

Die toewysing van voorraad na winkels is een van die belangrikste prosesse in die bestuur van 'n kettingwinkel. In die klere-industrie behels toewysingsbesluite onder andere die bepaling van hoeveelhede van elke grootte (byvoorbeeld klein, medium en groot) wat aan elke winkel gestuur moet word. 'n Gevallestudie van hierdie probleem in Pep Stores Bpk. (PEP), een van die vernaamste kleinhandelaars in Suid-Afrika, word in hierdie projek beskou.

In PEP word produkte sowat sewe maande voordat dit in die winkels beskikbaar is, by fabrieke bestel. Vanaf die fabrieke word die produkte na distribusiesentra verskeep, vanwaar dit per pad na die onderskeie winkels versprei word. Voordat die produkte bestel word, word voorlopige toewysingsbesluite geneem. Wanneer die voorraad by die distribusiesentra aankom, word besluite in verband met die toewysing van produkte en groottes aan winkels gefinaliseer. Toewysingsbesluite word gedurig aangepas deur die seisoen soos meer onlangse verkoopsdata beskikbaar word.

In hierdie tesis word simulasiemodelle ontwikkel om vier toewysingsmetodes in terme van totale verwagte verkope, tekorte en surplusse te vergelyk. Die algoritmes sluit PEP se huidige algoritme in, asook 'n bestaande algoritme wat die verwagte aantal weke tekorte en surplusse wat by winkels voorkom, minimeer, 'n nuwe algoritme met die doel om verwagte verkope te maksimeer, en 'n verslapping van die nuwe algoritme.

Die simulasiemodelle word volgens twee modelleringsbenaderings ontwikkel. Elke benadering word op Somer- en Winterprodukte toegepas, sodat daar vier simulasiemodelle ontstaan. Die twee simulasiebenaderings lewer soortgelyke resultate vir Somer- en Winterprodukte, naamlik dat al vier toewysingsmetodes ongeveer ewe effektief is.

Acknowledgements

I would like to thank the following people and organisations for their contribution to the project.

- Prof SE Visagie, my supervisor, for his guidance, advice, support and neverending enthusiasm throughout the project.
- Prof JH Nel, dr CG Jacobs, prof M Kidd and prof SJ Steel for their assistance and insight.
- The IDEE group, especially Emma Gibson, for their help and ideas.
- Erik van Tonder, my fiancé, for his help, love and encouragement.
- My family, for their loving support, and for putting up with my moods!
- Liezl van Eck, PEP employee, for her friendly and enthusiastic assistance with data and information.
- The Harry Crossley foundation, Stellenbosch University and PEP for the generous financial support towards the project.

Soli Deo Gloria.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Retailing is defined as "business activities involved in selling goods and services to consumers for their personal, family, or household use" [8]. A business establishment or a firm involved in retailing is called a retailer [32, 41, 53]. Often, retailers consist of many different stores in which products are sold to customers. These retailers are called retail chains.

One of the most important decisions for a retail chain is how to allocate stock from distribution centra to stores. A fashion chain has the additional problem of allocating the correct number of each size (for example small, medium and large) to each store. A case study of this problem within the context of Pep Stores Ltd. (PEP) [52] is considered in this thesis. In the following sections, an explanation of where this problem fits in the supply chain of a retailer is given, as well as a description of the problem within PEP.

1.1 The retail supply chain

The supply chain of a retailer consists of all activities associated with the processing of raw materials into finished products, as well as supplying products to customers. These activities include, among others, the management of demand and supply, the sourcing of raw materials and parts, the manufacturing and/or assembly of products, the storage and control of inventory, the placement and management of orders, the distribution of products and the delivery of products to customers. The information systems that are used to monitor all these activities is another important element of the supply chain [43, 56].

There are four important role players in the processing of products to their final form as supplied to the customers: the suppliers, the producers, the distributors and the retailers [63]. The flow of products through these four role players is represented in Figure 1.1.

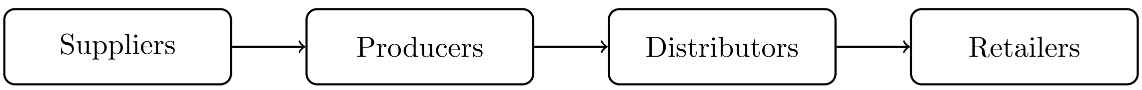


Figure 1.1: A schematic representation of the flow of products through the supply chain.

The suppliers, who are at the root of the supply chain, are responsible for supplying raw materials and parts to the producers. The producers then process the raw materials or assemble the parts into the finished products that are sold to customers. Then, the stock is sent in bulk to the distributors, who distribute the products to the retail stores. They typically make use of a distribution centre (or distribution centra), which is a warehouse in which stock is stored, managed and reorganised before it is sent to the retailers. The retailers sell the received products directly to the buyers in the retail outlets [6, 19, 63].

1.2 The distribution network of a retail chain

The distribution network of a retail chain refers to the flow of products in the last three phases of the supply chain in Figure 1.1. In Figure 1.2, a schematic representation of the organisation of a distribution network of a typical retail chain is given.

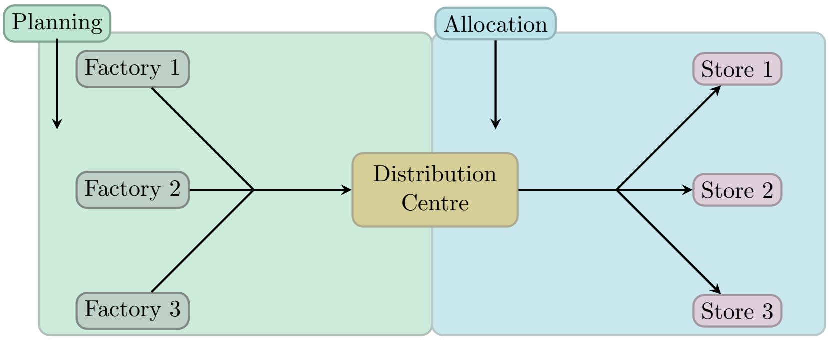


Figure 1.2: A schematic representation of the most important elements in the distribution network of a typical retail chain, with the underlying processes of planning and allocation.

The manufacturing or assembly of products usually takes place in factories, after which the products are sent to the distributors. The distributors then process the stock in a distribution centre (distribution centra), from where it is allocated to the stores, where the retailers sell it to the customers.

Underlying the distribution network are two processes: the planning process and the allocation process. Decisions made during the planning process influence the part of the distribution network from when orders are placed until the finished products arrive at the distribution centre, as shown in Figure 1.2. Decisions made during the allocation process influence the part of the distribution network from when the finished products arrive at the distribution centre until they are available in the stores.

The planning process includes decisions like which products to order, as well as the order quantity and order frequency of products. An importent component of the process is assortment planning, which includes decisions concerning the properties of products. In the clothing industry, this includes decisions about how many and which products to include in the product line, how many and which styles to buy, how many and which product sizes (for example small, medium and large) to buy and how to manage the inventory levels of the product lines, styles and sizes [44, 57].

The last step in the planning process is to place orders [33]. Usually, it takes a few months for orders to arrive at the distribution centre, after which the allocation process starts. During the allocation process, decisions are made about the allocation of stock to stores.

Allocation can be done using either a push or a pull system. In a push system, decisions about how many of each product to send to each store are made on a central level for all the retail outlets. In a pull system, which is a decentralised approach, allocation is made based on requests from the store managers [21].

During the planning phase, orders are placed according to preliminary allocation decisions. These decisions are made using forecasts based on the sales data from previous seasons that are available at the time. When the stock arrives, the allocation decisions are finalised. For a push system, the demand forecasts may be updated as the season progresses and more recent sales data of the current season become available. Allocations are then made based on the updated forecasts as well as the stock received. For a pull system, allocations are based on requests from the store managers, together with the stock received.

1.3 The size-mix problem

In literature, there are some references to the determination of *size-mixes* as a part of assortment planning. This entails decisions about how many different sizes are ordered and how many units of each size are ordered. These decisions are based on the expected demand of each size. Demand is typically forecasted by using historical sales data [65].

During the planning process, orders are placed based on preliminary size-mix allocations to stores. During the season, size allocation decisions may be updated for each new order. In the case of a push system, the updated size allocation decisions are based on the adjustment of demand forecasts as more sales data become available. For a pull system, store managers' requests may change throughout the season, which also has an effect on allocation decisions.

1.4 Measuring an assortment or allocation model's effectiveness

After developing a new (size-mix) assortment planning or allocation model, it is important to measure its effectiveness and compare it to existing methods (if there are any). Ideally, the expected effectiveness should be determined to obtain confidence in the model before implementation. Otherwise, the model could be implemented in practice to observe its actual effectiveness. This is usually done by implementing the model for a test group of stores and comparing results to that of a control group, for which the old system is used. However, this method is expensive and risky, and direct comparison is not possible, as two methods cannot be implemented at the same store at once.

1.5 PEP

PEP is a filial of the South African company Pepkor [51]. PEP sells, among others, clothing and shoes, cellular products and homeware. The first PEP store was opened in 1965 in the Northern Cape, and since then the company has grown to become the largest single brand retailer in Africa. PEP has more than 1800 stores in Southern Africa, and has more than 15 000 employees [52].

1.5.1 The distribution network in PEP

The flow of products in the distribution network in PEP, as well as the time frames in which they take place, are given in Figure 1.3. About 6 to 10 months before PEP's products are available in the stores, orders are placed at the factories, where they are manufactured. The factories are mainly situated in the Far East. After manufacturing, the products are shipped to harbours in Cape Town and Durban. The products are then transported via road to one of PEP's three distribution centra. PEP's two largest distribution centra, where about 90% of their products are processed, are in Durban and Johannesburg. There is also a small distribution centre in Cape Town. From the distribution centra they are transported via road to 17 hubs and then to the stores. The distribution process takes about 2 weeks.

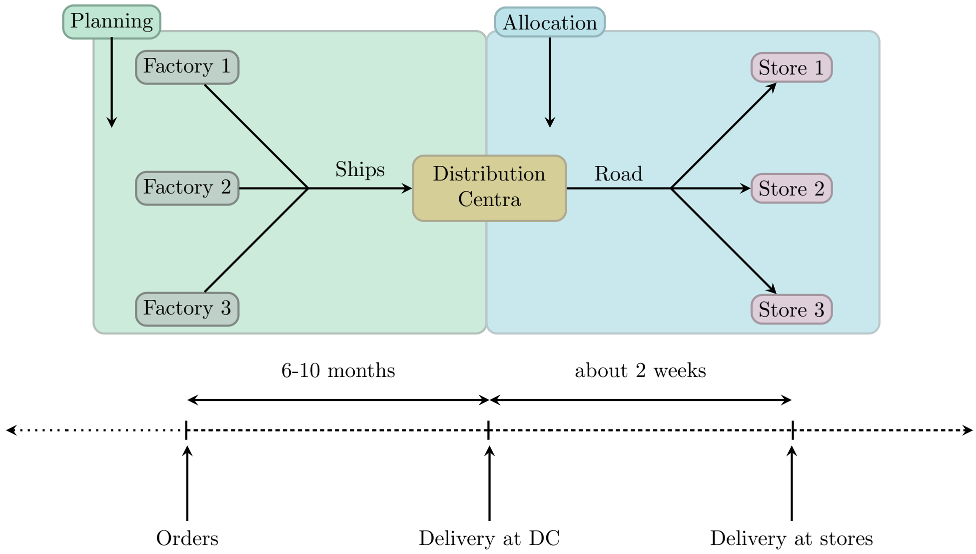


Figure 1.3: A schematic representation of the most important elements in PEP's distribution network with the underlying processes of planning and allocation.

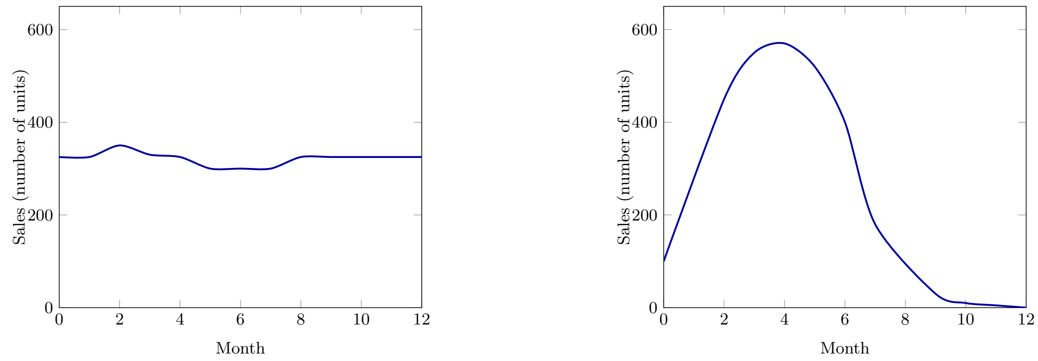
In PEP, planning and allocation are done on a central level for all stores. Irrespective of decisions like order quantities and frequency, preliminary allocation decisions are already made during the planning process. These include decisions about how many units of each product to send to each store, and, in the case of fashion products, how many units of each size to send to each store. During the allocation process, which is done through a push system, the initial planning is adjusted when making final size-mix allocation decisions. For these adjustments, the initial planning is considered, as well as the forecasted future demand at each store for each size, which can now be done more accurately with more recent sales data. The initial planning is re-adjusted for every new order that arrives in the distribution centra.

1.5.2 Product structure in PEP

PEP distinguishes between two types of products. Firstly, there are the products with a more or less constant demand throughout the whole year. Underwear, for example, falls in this category. The graph in Figure 1.4(a) contains a representation of the possible sales for this type of product.

The second type of product's demand is of a seasonal nature. Typically, fashion items like summer (or winter) clothing that peak in summer (or winter) months, fall in this category. The graph in Figure 1.4(b) contains a representation of the possible sales for this type of product. In this study, products of the second type are considered.

Products are further classified according to subgroups or -classes, which can be subdivided into different styles. Formal, long-sleeved shirts may form part of one group and formal, threequarter-sleeved shirts may form part of another group. Different coloured shirts in the same group are classified as different styles. In other words, a red formal shirt with long sleeves is classified as one style, and a green formal shirt with long sleeves in the same cut is classified as another style. Each order contains one style consisting of different sizes.



(a) Sales for a non-seasonal product. **(b)** Sales for a seasonal product.

Figure 1.4: Possible sales over time for (a) a non-seasonal product and (b) a seasonal product.

1.5.3 Adjustments during the allocation phase

A representation of the possible size profiles for three successive styles of a specific subclass (say long-sleeved shirts) in a specific size (say mediums) for a specific store, is given in Figure 1.5. Suppose the first style is red shirts, the second blue shirts and the third green shirts. When the red shirts' order arrives at the distribution centre, sales data from a similar product of the previous season are used to make adjustments to the initial planning when determining how many red mediums to send to this store. When the blue shirts' order arrives at the distribution centre, partial sales of the red medium shirts are already known. This can be used to make adjustments in the planned quantities when determining how many blue medium shirts to send to this store. In the same way, the sales of red and blue shirts may be used to make adjustments in the initial planning when allocation decisions for the green shirts are made.

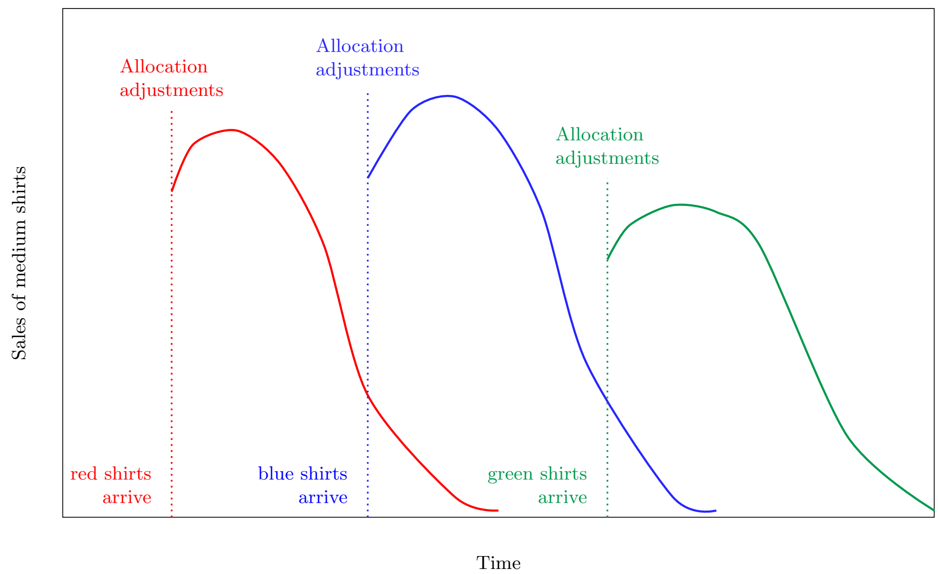


Figure 1.5: A schematic representation of possible sales of medium shirts for the successive styles in a store over time.

Suppose, as an example, the sales of the red medium shirts at a certain store were better than expected. When the blue shirts' order arrives at the distribution centre, more blue medium shirts will be sent to the store than initially planned. Based on the sales of red medium shirts at other stores, less blue medium shirts are sent to (an)other store(s), as the order was placed months earlier based on the initial planning and thus the total number of blue medium shirts available for allocation is fixed.

1.6 Problem statement

Two problems are being investigated in this thesis. Problem I is the size-mix allocation problem, and more specifically, the adjustment of size-mix allocation decisions as the season progresses, within the context of PEP. Problem II concerns the validation of algorithms developed to solve Problem I.

1.7 Thesis objectives

The thesis problems will be addressed by pursuing the following objectives.

Objective I

a To describe the problem of allocation adjustment decisions in relation with supply chain management and the distribution network of a retail chain;

b To explain the context of the problem within PEP Stores;

Objective II

a To describe existing literature on size-mix allocation and related problems;

b To investigate effectiveness measures of assortment and allocation models applied in literature;

Objective III

a To collect and analyse relevant data to solve allocation adjustment decisions and measure the effectiveness of size-mix allocation models;

b To describe, validate and clean the collected data;

Objective IV

a To describe existing algorithms that will be tested by means of an effectiveness measure;

b To develop and describe new allocation adjustment methods to compare against existing algorithms by means of the identified effectiveness measure;

Objective V

a To develop simulation models to measure the effectiveness of allocation algorithms;

b To test the validity and accuracy of the simulation models;

Objective VI

a To use the newly developed simulation models to measure the relative effectiveness of different allocation methods;

b To make recommendations based on results, discuss ideas for future research and provide a summary of the study.

1.8 Structure and layout of thesis

The remainder of the thesis will be structured as follows. In Chapter 2, literature related to the study is discussed. Data that were received from PEP are discussed in Chapter 3. The allocation algorithms are discussed in Chapter 4. In Chapter 5, simulation models are developed to simulate sales of Summer products for the purpose of comparing allocation algorithms. Similar models for Winter products are developed in Chapter 6. Results for the comparison of allocation algorithms are provided in Chapter 7. Finally, in Chapter 8, recommendations are made, ideas for future research are discussed and the work completed in the study is summarised.

CHAPTER 2

Literature review

This study involves a few broad topics from literature. The size-mix allocation problem is the main related topic to this thesis. This and other related problems are discussed in §2.1. In §2.2, the methods that were used to measure the effectiveness of the models developed in §2.1 are covered. The study also involves the development of a simulation model, which in turn involves the estimation of demand as well as forecasting methods. Simulation and related topics are discussed in §2.3.

2.1 The size-mix allocation and related problems

The size-mix allocation and other problems relating to it occur during the planning and allocation phases of the distribution network of a fashion retailer. During the planning phase, size-mix ordering decisions have to be made for the company as a whole. These decisions form part of a bigger planning process, namely assortment planning. The size-mix allocation problem, which is the main problem addressed in this thesis, is a follow-up of the size-mix ordering problem. During allocation, ordering decisions have already been made, and a fixed size-mix has to be broken down into smaller size-mixes for each store. The size-mix allocation problem is a special case of the general allocation of stock to the stores of a retail chain.

The size-mix allocation problem itself is not very abundant in literature, but there do exist many publications on the related problems, namely the size-mix ordering and general allocation problems. In §2.1.1 and §2.1.2, a brief overview is given on these related problems. In §2.1.3, the publications that could be found regarding the size-mix allocation problem are discussed.

2.1.1 The size-mix ordering problem

Determining the size-mix that should be ordered for the entire company is part of assortment planning, which falls within the planning process in the distribution network of a retailer. During assortment planning, it is attempted to maintain a balance between variety, depth and service level. Variety planning entails the planning of the number of product categories that are supplied to the consumer, depth refer to the planning of the number of stock keeping units that are supplied, and service level concerns the number of individual items of a specific stock keeping unit supplied to each store [44]. A stock keeping unit is a unique item and is indicated with a series of letters and/or numbers so that the item can be uniquely identified according to the properties of the item [68].

The size-mix ordering problem forms part of the depth and service level decisions. Decisions about the number of units that are offered for sale are part of the depth decisions, because they involve decisions about the number of stock keeping units that are supplied for sale. Decisions about the number of units of each size that are ordered and therefore offered for sale, form part of the service level decisions, because they involve planning about the number of items in each stock keeping unit.

Different variations of this problem have been addressed in literature. Silver and Kelle [65], for example, developed a model to determine the number of units of each size held in inventory given a restricted budget and the objective to minimise the expected number of units short. Robb [58] solved the same problem by using a Markov process to model how an individual's size changes over time. He compared three methods, including the method developed by Silver and Kelle, a new method and a benchmark method. Gaul *et al.* [25], Kießling *et al.* [35] and Kurz *et al.* [37] considered the size-mix ordering problem when products are ordered in pre-packs, each consisting of a specific size-mix. Gaul *et al.* developed an integer problem as well as a heuristic approach to the problem. Kießling *et al.* [35] expanded the model by Gaul *et al.* by taking markdowns into account and developed a stochastic mixed integer problem to solve it. Kurz *et al.* [37] developed a heuristic method that works according to the principle of ordering more of a size that normally sells out quickly at a store than one that takes longer to sell out.

The size-mix ordering problem arise during the planning phase in the distribution network of a retail chain, and is usually solved months before the size-mix allocation problem. The size-mix ordering problem is solved using only historical sales information, while the size-mix allocation problem may be solved using new sales information that becomes available as the season progresses. When allocation takes place, the size-mix ordering problem has already been solved, which means that the amount of stock that has to be allocated is fixed. The size-mix ordering problem precedes the size-mix allocation problem and the two problems are therefore related, but the methods used when solving the size-mix ordering problem cannot be applied directly to the size-mix allocation problem.

2.1.2 The general allocation problem

The general allocation problem involves the allocation of stock to stores from a distribution centre or warehouse, where products do not necessarily consist of different sizes. This problem have been well researched, and many different approaches exist in literature. This section provides a brief overview of the most important publications.

McGavin *et al.* [46] considered the allocation of stock from a warehouse to N identical stores, and developed an allocation policy that takes place in two intervals, with the objective to minimise shortages at each store. Hill [29] compared four pull allocation policies that increase in complexity. The simplest policy allocates stock to stores in the same sequence in which the store orders are processed, and the most complex policy is a method based on the probability of a shortage at a store. Axsäter *et al.* [5] assumed that shortages can be backordered, and developed an allocation formulation to minimise holding cost plus ordering cost. They solved it with a heuristic similar to the two-interval approach by McGavin *et al* [46].

All these authors solved the problem for a pull system, which is more commonly used than the push system [29]. This means that demand is assumed to be a random variable. The problem in this study concerns a push system, and PEP assumes demand is deterministic and known in advance. Therefore these allocation methods could not be considered to aid in the solution to the problem addressed in this thesis.

2.1.3 The size-mix allocation problem

Only two studies concerning the size-mix allocation problem could be found in literature. One is a recent study by Caro *et al.* [12, 13] about a size-mix allocation problem in the well-kown international fashion company Zara [75], which has more than 1500 stores. The other is a study by Thom *et al.* [69] that addressed the same problem as the one considered in this thesis.

Caro *et al.* [12, 13] used operations research techniques to solve Zara's size-mix allocation problem. The allocation process in Zara takes place from a distribution centre, where stock is processed and sent to the stores. Caro *et al.* formulated a mixed integer programming problem, where total sales are maximised subject to stock constraints. Forecasts of future sales, inventory levels of each size in the warehouse and decisions about the size-mix made during the planning process, are used as inputs to the model. Forecasts are done using historical data and requests by store managers. Results show a 3 to 4% improvement in sales from the previous system, which only took into account the requests of store managers.

An important aspect of Zara's problem is that less important sizes (for example extra small and extra large) are removed from the shelves when the important sizes (for example small, medium and large) of a product are sold out. The problem addressed in this thesis does not have that quality.

In the article by Thom *et al.* [69], four size-mix allocation models were developed. These models were tested using data sets provided by PEP. The models follow a goal programming, mixed integer approach and minimise the number of weeks' shortages and surpluses, subject to stock and integrality constraints. There are also bounds on the number of units of each size that may be allocated to each store, based on requirements of PEP. Two of the models are exact approaches, and the other two are heuristics developed in order to decrease computational time. Results show that the newly developed models improve by about 14% on PEP's current method in terms of the objectives set in the goal programming formulation.

2.2 Measuring the effectiveness of models

In order to gain confidence in the allocation and assortment models discussed in §2.1, the authors had to measure their effectiveness. The methods that were used to measure the effectiveness of these models are discussed in this section.

One method is to implement the model in practice, usually for a limited range of products, to observe its actual effects. This method was followed by Kurz *et al.* [37], who conducted a real-world blind study to compare the newly developed size-mix ordering heuristic with the old system. The new system was applied to 10 stores, and 10 stores for which the old system was still in place, were used as a control group. The consistency of supply with demand was measured for both groups in order to compare the two systems. Kießling *et al.* [35] conducted a field study to compare their size-mix ordering model's results to sales from the same commodity group that took place in a previous year when the old system was used. Caro *et al.* [12, 13] performed a real-world pilot study to measure the improvement in performance brought about by the implementation of the new allocation method. Like Kurz *et al.*, Caro *et al.* also implemented the new method for a test group of stores and compared results with a control group, for which the old system was used.

A second method is to compare the expected effectiveness of a model to a benchmark method, or, in the case of a heuristic, to optimality. This is usually done by calculating some objective function value, for example the expected sales or the expected number of shortages and/or surpluses. Silver & Kelle [65] tested the effectiveness of their size-mix planning model by comparing the expected number of shortages (the objective function value) to that of a simple benchmark method. Robb [58] used the same measure as Silver & Kelle. Gaul *et al.* [25] measured the effectiveness of their heuristic by calculating the difference between the optimal objective function values of their heuristic approach and the exact approach. Thom *et al.* [69] compared their size-mix allocation models with PEP's current method by using six different measures of the expected number of weeks of understocks and overstocks.

The third method is to simulate demand and compare the sales generated by different allocation models. Hill [29] used simulation to compare the four allocation methods developed in the article with regards to customer service and total system stock. Demand at each store was assumed to follow a Poisson distribution with a mean value of 6. The heuristics developed by McGavin *et al.* [46] to solve a size-mix planning problem were tested by simulating pseudo-random gamma demands. The gamma demand distribution was selected based on demand properties associated with the problem under consideration. Axsäter *et al.* [5] also tested the effectiveness of their warehouse replenishment by simulation. Demand was generated for 68 test problems from the normal distribution in some cases and the negative binomial distribution in other cases, depending on the distribution of the historical data of the test instances.

In the context of this thesis, real world experiments will typically only cover one or two small subclasses in order to minimise PEP's risk. Results may differ for different subclasses; therefore, a method that can accommodate more subclasses would be more suitable. This method is also rather time consuming, as a whole season has to pass before it is possible to see the full impact of an allocation model. Another disadvantage is that different methods cannot be directly compared to one another, as it is impossible to implement different methods at the same stores at the same time. On the other hand, the expected effectiveness of an allocation model within PEP's context may be an inaccurate indication of the resulting number of unit sales. Furthermore, forecasts made by PEP may be inaccurate, so that the calculated expected effectiveness is not a true representation of reality. This method may be used to give an initial indication of performance, but another method is necessary to obtain more certainty. A simulation method is therefore the most appropriate method for PEP. A discussion on simulation and related topics follows in the next section.

2.3 Simulation and related topics

Simulation is a technique where the operation of a real-world system is imitated. Simulation usually involves a simulation model, which consists of a set of assumptions about the operation of the system. These assumptions are in the form of mathematical or logical relationships [34, 74].

A system is defined as "a collection of entities that act and interact toward the accomplishment of some logical end" [62]. For example, if weekly sales for a particular product are simulated, the system may consist of the stores where the sales take place, the products that are sold and the customers that buy the products [34, 74].

It is often desirable to describe the state of a system. The state of a system can be defined as "the collection of variables necessary to describe the status of the system at any given time" [62]. In the sales example, the state variables are the opening stock, the demand and the closing stock in a particular week [34, 74].

A system can be classified as a discrete or continuous system. In a discrete system, the state variables only change at discrete points in time; in a continous system, the state variables change continuously over time [34, 74]. Weekly sales may be modelled as a discrete system by simulating weekly demand and stock levels. Then the state of the system changes once a week.

A system may be modelled by means of a stochastic or a deterministic simulation model. A stochastic simulation model is a model that contains one or more random elements; a deterministic simulation model is one that contains no random elements. Stochastic simulation where the state of a system changes at discrete points in time, is called discrete event simulation [34, 74]. These models usually involve the generation of random variables from a statistical distribution.

In order to simulate sales, it is necessary to be able to estimate the parameters of a demand distribution based on historical sales data. In §2.3.1, statistical procedures for estimating demand distributions from sales data are discussed, and in §2.3.2, forecasting approaches for the estimation of demand parameters are discussed. Another technique that will be used as part of the simulation model is Monte Carlo sampling, which will be explained in §2.3.3.

2.3.1 Statistical procedures to estimate demand distributions from sales data

In literature, parameters for demand distributions have frequently been derived from sales data using statistical methods developed for the estimation of distribution parameters from censored data. Maximum-likelihood estimators (MLE) or similar approaches are most often used.

Conrad [17], Nahmias [49], Anupindi *et al.* [4] and Stefanescu [66] used maximum likelihood estimation to estimate the parameters of different demand distributions when only sales data are available. Conrad [17] studied a newsvendor type problem and estimated the mean of a Poisson demand distribution. Nahmias [49] used a normal distribution to model demand and compared the MLE method to a best unbiased estimator approach and a new estimation method derived in the article. Anupindi *et al.* [4] assumed a Poisson arrival process and, in addition to lost sales, also incorporated the possiblity of product substitution. Their model was tested using sales data for vendor machine products. Stefanescu [66] modelled demand with the multivariate normal distribution and used the Estimation-Maximization algorithm first proposed by Dempster *et al.* [20] to determine the optimal demand parameters.

Hill [30] developed an approach to estimate demand parameters based on data obtained from point-of-sales scanning systems. Assuming that customer arrival rates follow a Poisson distribution, their approach was to estimate customer arrival rates and the moments of customer order size in order to eventually determine the parameters of any demand distribution used by the modeller. Agrawal & Smith [1] developed a new method for the estimation of demand parameters when demand follows a negative binomial distribution. Lin [42] also assumed negative binomial demand and developed an estimation method where demand parameters are updated throughout the season as sales data become available. Lau and Lau [40] developed an approach for the estimation of demand distributions for a newsvendor type product when only sales data are available. Lariviere and Porteus [39] discussed the estimation of demand parameters from censored sales following a newsvendor distribution. They used a Bayesian approach where demand parameters are frequently updated as more sales information becomes available. Conlon and Mortimer [16] developed a method to estimate demand parameters using the Expectation-Maximization algorithm. Their method is applicable for the estimation of demand when availability is reviewed periodically.

These methods are, however, not practical in the context of this study, because they require enough historical data to be able to estimate a statistical distribution. The data sets provided by PEP for testing purposes typically only have three years of historical data. Each week has its own distribution so that, in most cases, only three data points are available to estimate a distribution from. There are also important factors other than historical sales that have an impact on demand and that have to be incorporated into the model.

2.3.2 Estimation of demand by forecasting future sales

Another possible method of demand estimation is to make use of an underlying forecasting method in order to generate demand parameters. This approach was followed in an article by Wecker [72], where estimation of demand was based on an autoregressive model. A process is developed to estimate true demand from sales data with the assumption that the error terms follow a normal distribution with mean 0 and known variance σ^2. However, for the problem in this thesis, the true population variances are not known, and there are too few data points to obtain good estimations.

There are many possible forecasting methods that could be used as a basis for demand parameter estimation. Forecasting methods can be either qualitative or quantitative. A purely qualitative forecast is based only on the judgement of the forecaster, requiring no statistical analysis or manipulation of historical data. A purely quantitative forecast requires no judgement but is only based on statistical manipulation of historical data [28]. The data that were received from PEP for testing purposes possess a clear pattern that makes it ideal for statistical analysis; therefore, only quantitative methods were considered.

Quantitative methods can be further divided into extrapolation (or time series) methods and causal methods. Extrapolation methods analyse the underlying pattern in historical data and extrapolate the pattern in order to generate future forecasts. Traditional extrapolation methods include naïve forecasting, moving average methods, simple exponential smoothing, Holt's exponential smoothing method, Winter's exponential smoothing method, time-series decomposition and autoregressive integrated moving average (ARIMA) models. Causal methods attempt to find the factors that caused the patterns in historical sales data in order to forecast future values. Traditional causal methods include simple and multiple linear regression [73, 74].

Quantitative traditional methods that are recognised in literature as appropriate for sales forecasting are Winter's exponential smoothing method, multiple regression, time series decomposition and ARIMA. This is because sales data often exhibit strong seasonal patterns [3, 14, 73]. This is also true for PEP's fashion sales data. Out of these methods, multiple regression is the only causal one. Analysing the data received from PEP, it is clear that there are important factors other than the underlying patterns in historical data that influence sales. In particular, the number of units of stock that are sent to the stores have an impact on sales. Because the study is about the effect of allocation decisions, it is important to be able to predict sales for different amounts of stock sent to stores. Therefore, multiple regression is the most suited traditional technique that could be considered.

A new causal technique that has recently become popular in the sales forecasting literature is artificial neural networks (ANNs) based modelling [67]. The technique involves the modelling of mathematical relationships among variables by attempting to replicate processes in the human brain and nervous system. ANNs have the ability to continuously learn about these relationships by analysing historical data [28].

Many recent studies have compared ANNs to traditional methods, with mixed results [14]. Although in some cases, ANNs outperformed traditional methods [3, 14, 23, 54], there were other studies where the performance of ANNs were similar but no better than the traditional methods [15], and studies where traditional methods performed better than ANNs [11, 18, 50]. Even though ANN models can sometimes be very effective, Alon *et al.* [3] remark that they may not be ideal for companies to implement, as they require special software and expertise and are computationally expensive. With a view to possible implementation purposes, it was decided not to use ANN models in this study.

Following the above discussion, multiple regression seems like the logical model to base the simulation on. The technique as well as its applications in literature are discussed in the following sections.

Multiple regression

Regression analysis is the study of the mathematical relationship between a variable called the dependent variable and one or more variables called the independent or explanatory variables. The mathematical relationship is used to predict the mean or average value of the dependent variable when the values of the independent variables are known. Regression analysis with only one independent variable is called simple linear regression. If there are more than one independent variable, the term "multiple regression" is used [27, 73, 74].

Let Y represent the value of the dependent variable, $\hat{Y}$ the predicted value of the dependent variable and X_i the value of the i^{th} independent variable. Then the population multiple regression equation is given by

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_k X_k + \epsilon,$$

where β_0 is the intercept, β_i are the slopes associated with X_i for all i, and $\epsilon = Y - \hat{Y}$ is the population error term. The error term should follow a normal distribution with mean 0.

The values for the slopes β_i are usually estimated from sample data. The estimates for β_i are represented by $\hat{\beta}_i$ for all i. Now $\hat{Y}$ can be estimated by the regression line

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots \hat{\beta}_k X_k.$$

Let $\mathcal{J} = \{1, 2, \dots, j, \dots, J\}$ be a set of observations. Then the estimates $\hat{\beta}_i$ may be estimated by minimising the sum of the squared errors for all observations in set $\mathcal{J}$, in other words, by minimising

$$\begin{aligned}\sum_{j \in \mathcal{J}} \epsilon_j^2 &= \sum_{j \in \mathcal{J}} (Y_j - \hat{Y}_j)^2 \\ &= \sum_{j \in \mathcal{J}} (Y_j - \hat{\beta}_0 - \hat{\beta}_1 X_{1j} - \hat{\beta}_2 X_{2j} - \dots \hat{\beta}_k X_{kj}),\end{aligned}$$

where ϵ_j is the error of the j^{th} observation, Y_j is the j^{th} dependent variable, $\hat{Y}_j$ the j^{th} predicted value and X_{ij} the value of the i^{th} independent variable for the j^{th} observation.

The accuracy of the regression model can be determined by the coefficient of determination, R^2, which measures how well the regression line fits the data. An R^2 value close to 1 indicates a good fit. In multiple regression, however, the value of R^2 may become deceiving, especially in the comparison of two regression models with a different number of independent variables. The value of R^2 tends to increase as more independent variables are added to the regression equation, even though the model does not necessarily become more accurate. The adjusted R^2 value adjusts the value of R^2 by taking into account the number of independent variables, and should therefore be inspected together with the R^2 value in the case of multiple regression [74].

The suitability of the independent variables should also be validated by testing the hypothesis

$$H_0 : \beta_i = 0, \text{ and}$$

$$H_a : \beta_i \neq 0.$$

For each independent variable i, H_0 is the null hypothesis and H_a the alternative hypothesis [74]. If β_i is 0, it means that the i^{th} independent variable has no influence on the dependent variable when used in conjunction with the other variables; therefore, if H_0 is rejected, it means that the independent variable has a significant explanatory effect on the dependent variable. The test statistic for each independent variable i is given by

$$t = \frac{\hat{\beta}_i}{\text{StdErr}(\hat{\beta}_i)},$$

where $\text{StdErr}(\hat{\beta}_i)$ is the standard error of $\hat{\beta}_i$. The null hypothesis H_0 is rejected if $|t| \geq t_{(\frac{\alpha}{2}, n-k-1)}$, where α is the significance level, n the number of observations and k the number of independent variables.

The joint explanatory power of the independent variables can be tested by means of the F hypothesis test, given by

$$\begin{aligned} &H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0, \text{ and} \\ &H_a : \text{at least one } \beta_i \neq 0. \end{aligned}$$

The F statistic, accompanied by a corresponding p value, is usually provided by computer software.

Assumptions of multiple regression

Regression modelling is based on a set of assumptions that must hold for the model to be valid. The key assumptions of multiple linear regression are the following [27].

1. The regression model is linear in the parameters.
2. The error terms of the regression are homoscedastic.
   In other words, there is no heteroscedasticity in the error terms, or the variance of the error terms is constant over different values of the independent variables.
3. The error terms of the regression are normally distributed with a mean value of zero.
4. There is no autocorrelation in the error terms.
   This implies that no error term corresponding to one observation is influenced by an error term corresponding to another observation. In other words, there is no positive or negative correlation between any two residuals corresponding to different observations.
5. There is no multicollinearity in the independent variables.
   This means that there is no linear relationship between two different independent variables.

Applications of multiple regression in literature

Since the 1960s, when regression was first applied in the retail industry, it has become a popular sales forecasting tool, especially for segmented market appeals like clothing retailers, restaurants, book shops and jewellers [60]. Even with the evolution of promising new forecasting methods that have in some cases proved to outperform traditional methods [3, 14], regression remains a widely used technique for sales forecasting [45].

In literature, regression is used both as an analysis tool and for forecasting in the sales industry. Gaur *et al.* [26], for example, developed a regression-based method to determine whether financial indicators influence retail sales. The dependent variable for their regression is the total sales of a retailer, and the independent variables are sales forecasts generated by equity analysts, the term of the forecast, and the return on an aggregate financial market index over the term of the forecast. The forecast is also frequently updated by the latest financial indicator information. They conclude that financial indicators are in fact statistically significant explanatory variables, and that they can improve on forecasts made by equity analysts by including financial indicators in their model.

Lam *et al.* [38] developed a log-linear regression model that forecasts sales in order to determine the optimal number of hourly staff members. They forecasted hourly sales and used store traffic and the number of staff members at each hour as independent variables. This enabled them to analyse the effect that the number of staff members have on total sales and ultimately on the gross profit net of staff cost, which they aimed to maximise.

Forst [24] forecasted weekly sales of a small restaurant near Marquette University in Milwaukee, Wisconsin. They compared seven multiple regression models and nine ARIMA models to one another in order to find the best forecasting method. They found that the model with the best performance was a multiple regression model with a dummy variable indicating the week number, and a lag variable representing sales in the previous week.

2.3.3 Monte Carlo sampling

Monte Carlo sampling is the procedure of selecting a point from a set so that each point in the set has a specified probability to be selected. In other words, if the set is defined as $\mathcal{I} = \{1, 2, \ldots, i, \ldots, I\}$, each point i has a probability p_i associated with it, where $\sum_{i \in \mathcal{I}} p_i = 1$. Selection is done in such a manner that point i is selected with probability p_i. If sampling is repeated several times with replacement, point i's frequency of occurrence should make out approximately $p_i \times 100\%$ of all selected points [48, 74].

There exists a number of Monte Carlo sampling techniques. The one which will be used in this thesis is called roulette-wheel selection, which follows the analogy of a roulette game. An imaginary roulette wheel consists of I compartments, and the area of each compartment i is proportional to the probability p_i. The roulette wheel is spun, and the compartment in which the point falls, is selected. Mathematically, the cumulative sum of the probabilities is calculated, resulting in a set of i numbers in the range $(0, 1]$, say $q_1, q_2, \ldots, q_i, \ldots, q_I$. Then $q_1 = p_1$, $q_2 = q_1 + p_2, \ldots, q_i = q_{i-1} + p_i, \ldots q_I = q_{I-1} + p_I = 1$. The range $(0, 1]$ is segmented into the intervals $(0, q_1], (q_1, q_2], \ldots, (q_{i-1}, q_i], \ldots, (q_{I-1}, q_I]$, corresponding to points $1, 2, \ldots, i, \ldots, I$. A uniform random number is generated, and the point corresponding to the interval in which the random number falls, is selected. The uniform random number is often generated by computer software, in which case pseudo-randomness is used [22, 74].

CHAPTER 3

Data and data handling

This chapter provides an overview of the data that were received from PEP for testing purposes, as well as a description of the handling of the data during model building. The data can be divided into two groups: sales data and allocation data. The sales data are discussed in §3.1 and the allocation data in §3.2.

3.1 Sales data

Four data sets containing sales information were supplied by PEP, each associated with a different subclass. Two subclasses are from Summer products and two from Winter products. These data sets were used to build models te simulate future demand.

Table 3.1 provides a summary of the properties of these data sets. Each subclass has a unique number and description. The season in which the sales took place is noted because the sales characteristics differ depending on the time of year. The last column supplies the years for which sales data were available. In each case, all the available years were used when building simulation models. The last year was used as a hold-out set in each case so that the accuracy of the models could be verified against actual sales. At least four years of data were available for each subclass, so that at least three years could be used as historical data when building the simulation models.

Subclass no	Subclass description	Season	Available years
A_S	Ladies fancy sandals	Summer	2010–2014
B_S	Men fancy sandals	Summer	2011–2014
A_W	Teenage girls fancy slippers	Winter	2011–2014
B_W	Ladies spun poly jackets	Winter	2011–2014

Table 3.1: Properties of the sales data received from PEP.

In PEP, sales are recorded every Saturday, which is considered to be the last day of the week. For every Saturday, the corresponding number of units of each size in each style that were sold at each store during that week is supplied. The data also contain the opening stock, inflows in number of units (in other words, the number of units of stock that was received by the store) and closing stock for each style in each size at every store in every week.

Summer sales usually start in about the 30^{th} week of the year, where Sunday is seen as the first day of the week. Week 1 begins on the first Sunday of the year, so that, if the 1^{st} of January is not a Sunday, the first Saturday of the year is assigned a week number of 0. This means that the 30$^{\text{th}}$ week of the year is either in the last week of July or in the first week of August. Sales continue until early in the next year, for about 26 weeks. Winter sales usually start in the fifth week of the year, which is in the first or second week of February, and also continue for about 26 weeks.

3.1.1 Data cleaning

All data sets were cleaned so that every season contains 26 weeks of sales for every size and store. For Summer data sets, the period of 26 weeks starts in the 30$^{\text{th}}$ week of the year and for Winter data sets in the fifth week of the year. In some cases, one or two units of sales were recorded before or after this period, but these weeks were omitted from the simulation models so that each season has the same number of weeks. If sales were not recorded for all 26 weeks, the missing weeks were inserted, with a value of 0 for sales and inflows. Ensuring that each season consists of the same number of weeks made the simulation results more accurate.

New stores for which no historical data exists, were not considered, because there is no reasonable way to simulate future demand if there is no history to base it on. Only stores for which at least one year of historical data exists in all sizes, were included. Stores that closed down before the year for which the forecasts are made, were also omitted, because the sales generated by the models were compared to actual data for verification purposes. Only the stores that occur in allocation data sets were considered, because allocation data are needed to compare allocation algorithms.

3.1.2 Calculation of demand

It is a well established fact in literature that demand cannot simply be assumed to be equal to sales in the case of a stockout [72]. In this case study, only three or at most four years of historical data were available. Each week, size and store is associated with a unique distribution, so that there were typically only three data points for the estimation of distributions. Therefore, there were not enough data available to use the statistical procedures described in §2.3.1. A simpler procedure was developed based on the assumption that, after stock has been sold out, demand decreases more or less linearly during the following weeks. The procedure requires the following parameters: define the set $\mathcal{K} = \{1, 2, \ldots, k, \ldots, K\}$ as the set of weeks in a season, $\mathcal{T} = \{1, 2, \ldots, t, \ldots, T\}$ as the set of stores included in the data set and $\mathcal{S} = \{1, 2, \ldots, s \ldots, S\}$ as the set of sizes in the data set. Let

- o_{stk} be the opening stock in size s at store t in week k as calculated by PEP,
- c_{stk} be the closing stock in size s at store t in week k as calculated by PEP,
- ℓ_{stk} be the number of units of inflows in size s at store t in week k as recorded by PEP, and
- a_{stk} be the sales in size s at store t in week k as recorded by PEP.

Define the following variables. Let

- z_{stk} be a variable indicating whether the first stock of the season in size s has arrived at store t,
- v_{stk} be the stock available to be sold in size s at store t in week k,
- u_{stk} be a variable indicating whether a stockout occurs in size s at store t in week k,
- f_{stk} be the number of weeks of stockouts left in size s at store t in week k (including week k),
- n_{st} be the number of units of estimated demand for the next few weeks in the case of a stockout in size s at store t, and
- d_{stk} be the estimated demand in size s at store t in week k.

The procedure is performed for each week in a season, for all sizes and stores, and is given in Algorithm 1.

Algorithm 1: Algorithm to calculate demand

1 **for** $k \in \mathcal{K}$ **do**
2   **for** $t \in \mathcal{T}$ **do**
3   **for** $s \in \mathcal{S}$ **do**
4   **if** The first stock of the season has arrived **then**
5   z_{stk} = True
6
7   **else**
8   z_{stk} = False
9   **end**
10             $v_{stk} = o_{stk} + \ell_{stk}$
11   **if** $v_{stk} \leq 0$ and z_{stk} **then**
12   u_{stk}= True
13   **end**
14   **if** u_{stk} **then**
15   **if** $u_{st,k-1}$= False **then**
16   f_{stk} = min(number of weeks before the end of the season, number of weeks before new stock arrives, 3)
17
18   **else**
19                     $f_{stk} = \max(f_{st,k-1} - 1, 0)$
20   **end**
21   **end**
22   **if** $f_{stk} > 0$ **then**
23   **if** u_{k-1} = False **then**
24                     $n_{st} = \lceil f_{stk} \times \text{average}(a_{k-1}, a_{k-2}, a_{k-3}))/2 \rceil$
25   **else**
26   n_{st}=$n_{st} - d_{st,k-1}$
27   **end**
28
29                 $d_{stk} = \min\left[\max\left\{\text{round}\left(\frac{n_{st} \times f_{stk}}{(f_{stk}(f_{stk}+1))/2}\right), 1\right\}, n_{st}\right]$
30   **else**
31                 $d_{stk} = a_{stk}$
32   **end**
33   **end**
34   **end**
35 **end**

In lines 4–9 of the algorithm, a test is performed to establish whether the first stock of the season has arrived. Before this condition holds, demand is assumed to be 0, or equivalently, equal to sales. For each week, it is assumed that the opening stock plus the inflows is available to be sold during that week. This is indicated in line 10. A stockout is recorded when the available stock during a week is less than or equal to 0, as indicated in lines 11–13.

In lines 14–21, a number is assigned to each week indicating how many weeks, including that week, are left in which demand has to be estimated. In the case of a stockout, a non-negative demand is estimated for the next three weeks, or until either new stock arrives or the season ends.

The formulas in lines 24 and 29 determine the number of demand units allocated for each of the following weeks in the case of a stockout. These formulas ensure that demand gradually dies out from the average of the previous three weeks to zero. Without rounding, these formulas would ensure an exactly linear decline from the average of the previous three weeks' sales to zero. However, demand is required to be integer and therefore rounding is necessary.

In line 24, the total number of units of estimated demand for the next few weeks in the case of a stockout, is determined. The ceiling of the average is taken so that at least one unit of demand is estimated. In line 26, the variable n_{st} is updated by subtracting the demand that was estimated during the previous week.

Finally, demand is estimated in lines 29 and 31. In the case of a stockout, demand is estimated according to the formula in line 29; if there is no stockout during the specific week, demand is equal to sales (line 31). At most n_{st} is estimated during a week. If more than 0 units still have to be estimated, at least 1 unit is estimated. If more than 1 unit is estimated, the formula is based on the number of units of demand that still has to be estimated and the number of weeks that are left in which demand has to be estimated. Demand is then rounded to the nearest integer.

3.1.3 Data validation

Data sets were validated to ensure reliability. Calculations were performed to verify the correctness of opening stock and closing stock. Opening stock during week i should equal closing stock in week $i-1$, and closing stock in week i should equal the sum of opening stock and inflows in week i minus sales in week i. In some cases, slight errors occurred in data recordings and calculations. When these errors were corrected by PEP, it sometimes resulted in negative values for the opening stock, inflows, closing stock and/or sales. However, these errors typically occur less than 1% of the time and have a negligible effect on results. Therefore, negative values were left as is.

In some of the data sets considered during experiments, there are one or two sizes for which sales data are incomplete. This means that no sales were recorded for the size during at least one of the historical seasons. Because the historical data is already limited, all historical seasons are necessary to accurately simulate future sales. Therefore only sizes with complete data for all available years were included in experiments. For Subclasses $\mathrm{A_S}$, $\mathrm{A_W}$ and $\mathrm{B_W}$, six sizes were kept after omitting sizes for which data were incomplete, and for Subclass $\mathrm{B_S}$, five sizes were kept.

There are very few outliers in the data, and only extreme outliers were adjusted. Extreme outliers occurred in the demand of Subclass $\mathrm{A_S}$, in the week ending on the 24$^{\text{th}}$ of December 2011. Demand during this week in all sizes and most stores were disproportionally high in comparison with demand during other years during the corresponding week, and caused inaccurate results during experiments.

A graphical display of the weekly demand for Subclass $\mathrm{A_S}$ on a company level, summed over all sizes and stores, is given in Figure 3.1. The disproportionally high demand during December 2011 is very clear. This phenomenon was not present in other data sets.

After discussion with PEP, it was assumed that these demand values were outliers, and they were adjusted during model building by using the average demand during the corresponding week of 2010, 2012 and 2013. Company level weekly demand, after adjusting the outlier, is given in Figure 3.2.

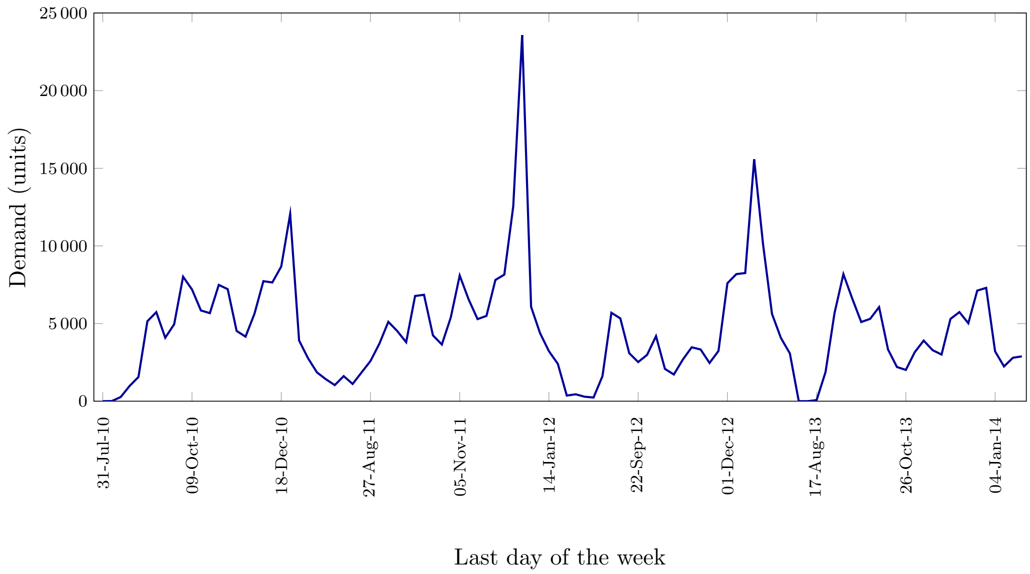


Figure 3.1: Weekly demand on a company level for Subclass A_S, for the years 2010–2013.

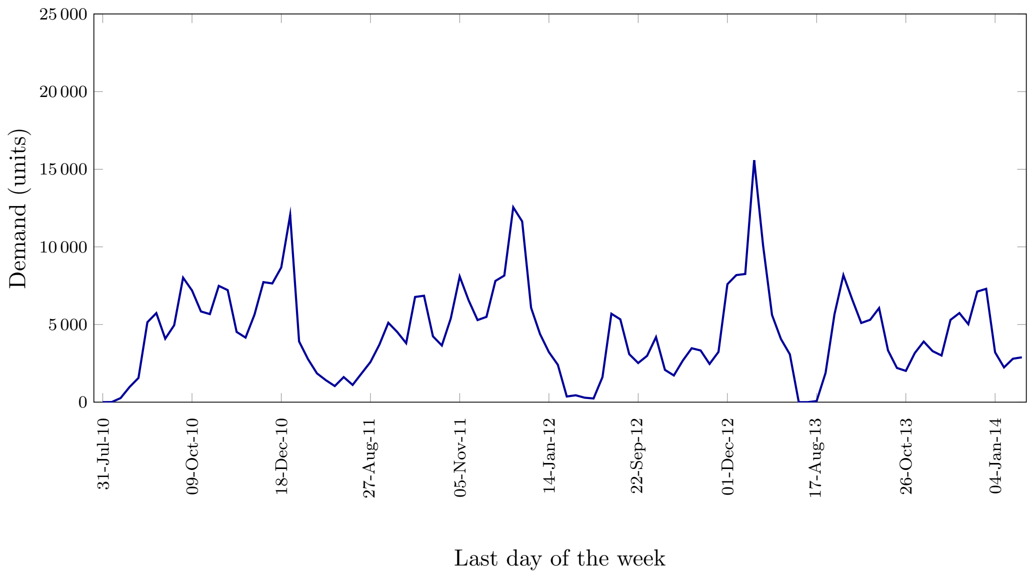


Figure 3.2: Weekly demand for Subclass A_S, for the years 2010–2013, after adjusting the outlier.

3.2 Allocation data

In PEP, allocation adjustment decisions are made for each style from a specific subclass. Allocation algorithms were performed for the last year of the subclasses in Table 3.1, and data sets associated with each style from these subclasses were received from PEP.

Each data set associated with a certain style contains the unique store numbers of the stores that sell items of that style. For each store, the expected demand as determined by PEP is given, as well as the expected rate of sales. Both of these quantities are given in number of units per week. The expected demand is calculated so that the total expected demand at all stores is equal to the number of units that were ordered.

Each store is classified according to the size of the store in terms of sales, and receives a number according to this classification. In PEP, these numbers are referred to as grade numbers. Upper and lower bounds, referred to as "grade minima" and "grade maxima", are associated with each grade number. Allocations to stores that are linked to a specific grade number have to fall within the grade minimum and maximum. These bounds are based on the expected sales at all stores that have the same grade number. The objective of the grade minima is to ensure that all stores receive a certain minimum number of each style in order to at least keep a minimum footprint of stock over all stores. The grade maxima prevent the allocation planners from sending too much stock of a certain style to one store. This ensures that stores have a larger variety. The bounds also prevent a scenario in which all stock is sent to the best performing, usually larger, stores at the expense of other, poorer performing stores.

Stores are also grouped in clusters according to sales properties that are similar to other stores in the cluster. Each cluster has a size profile associated with it. The size profile represents the expected spread of sales over the sizes and is given in the form of a percentage per size.

3.2.1 Data validation

Allocation data were also validated. Missing data occur in some cases, and data were adjusted to complete the data sets. However, missing data occur infrequently enough so that results are not influenced significantly.

In one or two of the data sets, some of the rate of sales data are missing. In these cases, the rate of sales were approximated by calculating the average ratio of rate of sales to demand for the other stores. Outliers, defined as values that are more than two standard deviations away from the average, were excluded in these calculations. The decision to calculate missing rates of sales in this manner was made in collaboration with PEP.

If the data do not contain a size profile for a certain store, the size profile according to which orders are made was used as a proxy for that size profile. PEP refers to this profile as the "company profile", and it is determined by the historical sales of all the stores of the company.

For some styles, no allocation data were available. If the allocation data for a specific style were not available, actual allocations were used in the place of solutions that would have been generated by the allocation algorithms. For these styles, it was assumed that all algorithms arrived at the same allocation solutions.

CHAPTER 4

Allocation algorithms

This chapter provides a description of four allocation methods that will be compared to one another by means of simulation. In §4.1, the algorithm that is currently used by PEP is discussed. In §4.2 follows an overview of the second algorithm that is being tested. The formulation for this algorithm was done in an article by Thom *et al.* [69]. A new algorithm is developed in this thesis, of which the formulation is given in §4.3. A fourth method which involves the relaxation of the new algorithm is discussed in §4.4.

4.1 Allocation algorithm 1: PEP's algorithm

PEP currently solves the problem by means of a computer automated heuristic. The heuristic ensures a feasible solution while attempting to satisfy demand as far as possible. Feasibility requires that the total allocated stock adds up to the ordered stock, that grade minima and maxima are adhered to and that all allocation quantities are integer and non-negative.

To arrive at an initial solution, PEP uses the expected demand per store as well as the size profiles associated with each store to calculate the expected demand per size, per store. This initial solution may be infeasible, and the rest of the algorithm ensures that constraints are met as far as possible.

The first step of the algorithm is to round all initial allocations to the nearest integer. Then, units are added to and subtracted from stores until all grade minimum and maximum constraints are satisfied. If there is no feasible solution, the lower and/or upper bound that is exceeded is relaxed. In these cases, a lower bound is relaxed to the minimum of all grade minima, and an upper bound is relaxed to the maximum of all grade maxima. The next step is to add units to and subtract units from sizes until, for each size, the number of units allocated to the size adds up to the number of units that were ordered.

4.2 Algorithm algorithm 2: Thom *et al.*'s formulation

In this thesis, simulation is used to measure the effectiveness of one of Thom *et al.*'s exact approaches, namely Model 2, as Model 1 delivered poor results. In all experiments, Thom *et al.*'s Model 2 took less than 10 seconds to solve, so that it was not necessary to consider their heuristic approaches.

Thom *et al.*'s Model 2 follows a goal programming approach. The target for each size at each store is the expected demand calculated by using PEP's expected demand and size profile for the store. These targets are the same as the initial solution in Algorithm 1. Underachievement on the target is equivalent to expected shortages and overachievement is equivalent to expected surpluses. The sum of the expected shortages and surpluses, measured in number of weeks' stock, is minimised. Expected unit shortages and surpluses are divided by the expected weekly rate of sales of the corresponding size at the corresponding store to convert the units to number of weeks. The reasoning behind the method is that one unit of stock makes a bigger difference at a store that, for example, sells 10 units in a season than a store that sells 500 units in a season. A weight α, usually larger than 0.5, is associated with the minimisation of shortages, so that a weight of $1 - \alpha$ is assigned to the minimisation of surpluses. The minimisation of shortages is regarded more important than the minimisation of surpluses, because shortages lead to lost sales and dissatisfied customers, whereas surplus stock may still be sold out at a discount at the end of the season.

The constraints include Algorithm 1's constraints, with added constraints ensuring that the allocations for each size and store remain close to the targets. This is necessary because the sum of the deviations may be minimised by very small deviations for some allocations that are traded off for large deviations for other. The added constraints prevent this scenario.

The effectiveness of the algorithms developed by Thom *et al.* [69] were only measured by the number of weeks that shortages and surpluses were expected to occur. If the data that were used to calculate the effectiveness measures are inaccurate, these results may be deceiving. The expected effect of the allocation algorithm on total sales was also not considered. A more detailed description of Algorithms 1 and 2 can be found in Thom *et al.* [69].

4.3 Allocation algorithm 3: Maximising expected sales

The algorithm discussed in this section is a mixed integer programming problem. It is an adaptation of Algorithm 2. The objective of Algorithm 3 is to maximise the expected total sales, measured in number of units. Algorithm 2 was adapted in this manner because it is uncertain whether the objective function of Algorithm 2 is efficient in increasing the total sales obtained by the current system (Algorithm 1).

4.3.1 Assumptions

The following assumptions had to be made with regards to the data received from PEP when modelling and testing the algorithm in this section. The assumptions are the same as those made by Thom *et al.* [69].

1. The expected demand at each store as determined by PEP, is a good approximation for future demand.
   This implies that demand is deterministic and known in advance.

2. The size profiles as determined by PEP, are good approximations to the actual spread of sales over the sizes.
   Size profiles are used to determine the expected demand as well as the expected rate of sales for each size at each store.

3. The rate of sales at each store is constant over time and approximately equal to the rate of sales provided by PEP.

The rate of sales determined by PEP are used in the algorithm as the expected future rates of sales.

4. Allocations are done as if there is no initial stock of the specific style in a store at the beginning of a season.
   The data for the initial stock of a style are only available on a store level and not on a size level. In other words, the total initial stock at a store is available, but the sizes of the initial stock are not known. However, if the data were available, the formulation could easily be adjusted to incorporate initial stock without altering the complexity of the formulation.

4.3.2 Formulation

Let $\mathcal{T} = \{1, 2, \ldots, t, \ldots, T\}$ be the set of all stores and $\mathcal{S} = \{1, 2, \ldots, s, \ldots, S\}$ be the set of all sizes of the style under consideration. The following parameters are used in the algorithm. Let

- d_t be the expected demand at store t,
- d_{ts} be the expected demand for size s at store t, as calculated using the size profile of store t,
- b_s be the total number of units of size $s$that were ordered,
- r_t be the expected number of units that will be sold per week at store t,
- r_{ts} be the expected number of units of size s that will be sold per week at store t,
- g_t be the minimum number of units that may be sent to store t according to PEP's grade minimum requirements,
- h_t be the maximum number of units that may be sent to store t according to PEP's grade maximum requirements,
- m_t be the maximum deviation from d_t for PEP's algorithm, measured in number of weeks' stock, and let
- m_{ts} be the maximum deviation from d_{ts} for PEP's algorithm, measured in number of weeks' stock.

Define the following variables. Let

- x_{ts} be the number of units of size s that are sent to store t, and let
- y_{ts} be the expected number of sales of size s at store t.

The mathematical formulation of this model is given by

$$\text{maximise } z = \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}} y_{ts} \tag{4.1}$$

subject to

$$d_{ts} \geq y_{ts}, \qquad t \in \mathcal{T},\ s \in \mathcal{S} \tag{4.2}$$

$$x_{ts} \geq y_{ts}, \qquad t \in \mathcal{T},\ s \in \mathcal{S} \tag{4.3}$$

$$\sum_{t\in\mathcal{T}} x_{ts} = b_s, \qquad s \in \mathcal{S} \tag{4.4}$$

$$\sum_{s\in\mathcal{S}}\sum_{t\in\mathcal{T}} x_{ts} = \sum_{s\in\mathcal{S}} b_s \tag{4.5}$$

$$\max(g_t, d_t - r_t m_t) \leq \sum_{s\in\mathcal{S}} x_{ts} \leq \min(h_t, d_t + r_t m_t), \qquad t \in \mathcal{T} \tag{4.6}$$

$$d_{ts} - r_{ts}m_{ts} \leq \; x_{ts} \leq d_{ts} + r_{ts}m_{ts}, \qquad t \in \mathcal{T},\; s \in \mathcal{S} \qquad (4.7)$$

$$x_{ts} \;\in \mathbb{Z}^{+}, \qquad t \in \mathcal{T},\; s \in \mathcal{S}, \qquad (4.8)$$

where $\mathbb{Z}^{+}$ is the set of non-negative integers.

The expected number of sales for a specific size at a specific store is the minimum of the corresponding allocation and expected demand. Adding the expected sales over all sizes and stores determines the expected total sales. This quantity is maximised in objective function (4.1). The expected number of sales for every size at every store are calculated in constraint sets (4.2) and (4.3).

Constraint set (4.4) ensures that the total number of units of a size that are allocated equals the total number of units of that size that were ordered. Constraint set (4.5) ensures that the total number of units that are allocated over all sizes and stores add up to the total number of units of all sizes and stores that were ordered.

Constraint sets (4.6) and (4.7) ensure that the allowed number of weeks of shortages and surpluses stay below a certain maximum on a store and a size level. In both cases, the deviation is restricted by the maximum deviation in number of weeks' stock that PEP's algorithm achieved, as PEP's algorithm is known to supply a feasible solution. Constraint set (4.6) also includes PEP's grade minimum and maximum requirements.

4.4 Allocation Algorithm 3′

Allocation Algorithm 3′ is an adaptation of Algorithm 3 where the bounds in constraint set (4.6) are relaxed to include only PEP's grade minimum and maximum requirements, and constraint set (4.7) is omitted. While the objective of Algorithm 3 is not to minimise number of week's shortages and surpluses, these constraints ensure that the shortages and surpluses are kept below a certain maximum. Algorithm 3 therefore attempts to balance the objective of maximising sales and minimising shortages and surpluses in number of weeks' stock, whereas Algorithm 2 only minimises shortages and surpluses and Algorithm 3′ only maximises sales.

CHAPTER 5

Models for Summer subclasses

In this chapter, two simulation models, Simulation Model S_1 and Simulation Model S_2, are developed for Summer subclasses. Sales are simulated on a subclass level, which means that different styles in the same subclass are handled as one product. This was done on recommendation by PEP [71]. Each simulation model involves the simulation of sales for one particular subclass. Subclass A_S (ladies fancy sandals) is used as a training set in the development of both models.

The system that is simulated by the two models is described in §5.1. Assumptions that were made when building and implementing the models are given in §5.2, and the development of the models is discussed in §5.3–§5.4. In §5.5, the two models are compared with regards to accuracy. A third model that was experimented with is discussed in §5.6.

5.1 The simulation system

The system that is simulated by each model consists of the stock of the product that is being sold (for example, ladies fancy sandals), including all the sizes of the product, as well as the stores at which the product is sold, and the customers that buy the products. According to the operation of the simulation system, one unit of a certain size of the product that is being simulated is sold at a certain store if (a) the unit is demanded by a customer at the store and (b) there is stock available in the specific size at the specific store.

Thus, for each week, it is necessary to determine the demand for each size at each store, as well as the available stock in each size at each store. The opening stock, demand and closing stock, which are the state variables of the system, can then be determined. The state variables change every week, so that the system is classified as a discreet system.

Customer demand for both models is generated by means of random variable generation. Therefore the simulation models are stochastic models. The mean for the random variable is determined with regression forecasting. The two models follow different approaches in the generation of demand. For the first approach, a product with all its sizes together is handled as one entity, and for the second approach, each size is handled as a separate entity.

In the first approach, total customer demand for the product, summed over all sizes and all stores, is generated for each week. Thus, a random variable is created for each week. The mean of the variable associated with a certain week is the expected total customer demand during that week. The expected mean demand is determined through a regression model which forecasts weekly demand. The total demand is then scaled down to a store level by means of roulette-wheel sampling, and the store demand is further scaled down to a size level, also through roulette-wheel sampling.

In the second approach, total customer demand for a certain size, summed over all stores, is generated for each week. A random variable is therefore created for each week in each size. The means for the random variables are again determined by regression forecasting. A separate regression model is associated with each size, and forecasts weekly demand for the subclass in the particular size. For each size, the weekly demand is then scaled down to a size level by means of roulette-wheel sampling.

Weekly inflows in each size at each store is given as input to the model. The inflows depend on the specific allocation algorithm that was used. The model then keeps track of the available stock by taking into account the inflows and sales of each week. Sales take place when demand is present and stock is available.

Note that the average demands over many simulation runs will converge to the regression values. However, the stochastic generation of demand is necessary because shortages and surpluses have to be simulated. An above-average demand will lead to an increased number of stockouts, and a below-average demand will lead to a high level of surpluses, so that the average sales will not necessarily be the same when the regression values (or averages) have been used for demand instead of stochastic demand around the averages.

5.2 Assumptions

The following assumptions were made when building and testing the simulation models.

1. Future demand can be derived from historical sales data.
   The simulation of future demand is based on a regression forecast which is derived from historical data.

2. In the case of a stockout, demand decreases more or less linearly to zero during the next three weeks.
   It is a well established fact in literature that demand has to be estimated in the case of a stockout and cannot simply be assumed to be equal to sales [72]. This simple method to estimate demand from sales data was decided on because it was the best available approach, since there are not enough data available to use the statistical procedures as described in Chapter 2 that are normally used in literature.

3. Availability has an influence on demand.
   This means that demand for a product increases if more units of the product are available in the store. This assumption was made in collaboration with PEP employees, who have observed this phenomenon in the past.

4. Demand follows a Poisson or normal distribution.
   Because there are in most cases only three years of historical data available for testing purposes, and each week follows a different distribution, it was not possible to estimate a distribution from the data. The most commonly used distributions in the case of demand are the Poisson and normal distributions [1]. Therefore, when no distribution can be determined, it is fair to assume one of these two distributions.

5. Future store demand occurs in the same ratio as historical demand.
The probability of demand occurring at a certain store is derived from the historical contribution of a store's demand to the total demand of all the stores in the company.

6. Future size demand at each store occurs in the same ratio as historical demand.
The probability of demand occurring in a certain size is derived from the historical contribution of a size's demand to the total demand of all sizes.

7. Total weekly demand on a size level follows the same pattern as total weekly demand on a company level. The same applies to weekly sales.
"Total weekly demand on a size level" refers to weekly demand for a size, summed over all stores, and "total weekly demand on a company level" refers to weekly demand for the entire company, summed over all sizes and stores. Correlation coefficients were calculated between weekly company level demand and each size's weekly demand for Subclass A_S to test whether this assumption is reasonable. All correlation coefficients are larger than 0.96. A graphical representation of weekly demand for Subclass A_S on a company and size level in Figure 5.1 further confirms the high correlation between company level demand and size level demand. Figure 5.1(a) contains weekly demand on a company level and Figure 5.1(b) contains weekly demand for the same subclass in three different sizes. Weekly demand for other sizes have a similar pattern, and weekly sales have a similar pattern to weekly demand.

5.3 Simulation Model S_1

In this section, Simulation Model S_1 is developed. Simulation Model S_1 follows the first modelling approach, where weekly Poisson demand, summed over all sizes and stores, is generated on a company level. The section provides a description of the simulation as well as the regression model on which weekly demand is based.

5.3.1 The simulation code

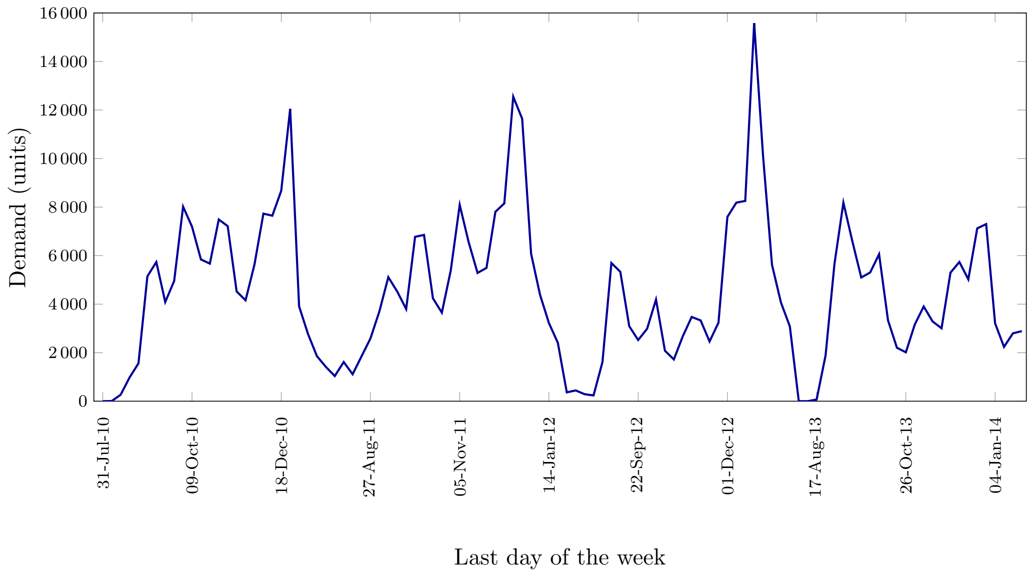
Let $\mathcal{K} = \{1, 2, \ldots, k, \ldots, K\}$ be the set of all weeks in the season, $\mathcal{T} = \{1, 2, \ldots, t, \ldots, T\}$ be the set of all stores included in the model and $\mathcal{S} = \{1, 2, \ldots, s, \ldots, S\}$ be the set of all sizes included in the model. Define the following variables. Let

$\hat{Y}_k$ be the expected number of units demand in week k,
d_k be the simulated demand in week k,
d_{stk} be the simulated demand in size s at store t in week k,
a_{stk} be the simulated sales in size s at store t in week k,
ℓ_{stk} be the unit inflow of size s at store t in week k,
o_{st} be the stock on hand in size s at store t,
p_t be the average historical proportion of store t's demand to total demand in a year, where $\sum_{t\in\mathcal{T}} p_t = 1$, and let
q_{st} be the historical proportion of demand in size s at store t to total demand at store t, where $\sum_{s\in\mathcal{S}} q_{st} = 1$ for all $t \in \mathcal{T}$.

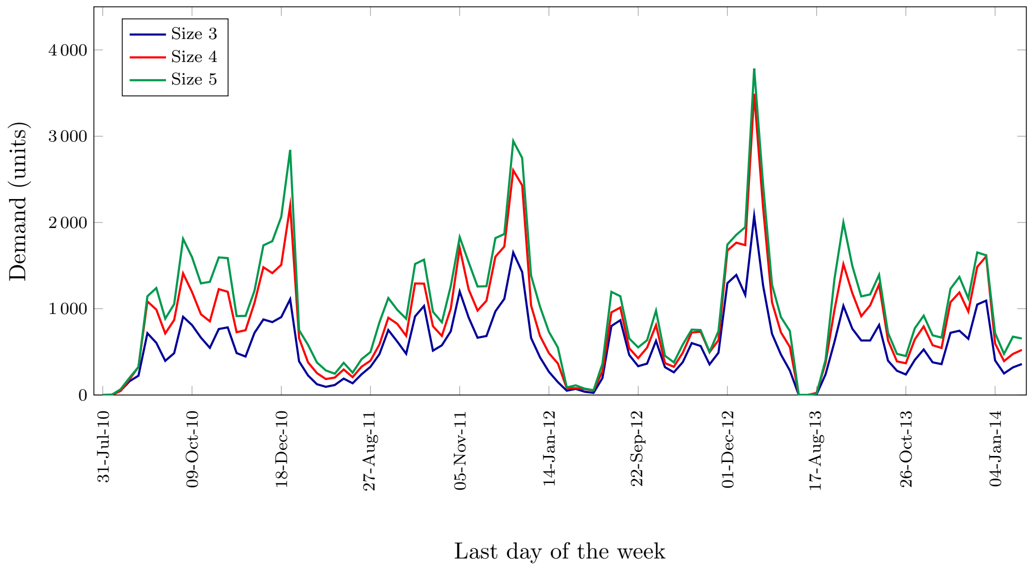
In lines 1–5, the stock on hand in all sizes at all stores is set to 0. This implies that the opening stock at all stores is 0 at the beginning of the season. In lines 7–8, demand for week k is

Algorithm 2: Pseudo code for Simulation Model S_1

for $t = 1$ to T **do**
    for $s = 1$ to S **do**
        $o_{ts} = 0$
    end
end
for $k = 1$ to K **do**
    Determine $\hat{Y}_k$ by means of a regression equation.
    Generate demand d_k, where $d_k \sim \text{Poisson}(\hat{Y}_k)$.
    for $t = 1$ to T **do**
        for $s = 1$ to S **do**
            $d_{tsk} = 0$
            $a_{tsk} = 0$
            $o_{ts} = o_{ts} + \ell_{tsk}$
        end
    end
    for $i = 1$ *to* d_k **do**
        Use roulette-wheel selection to select a store, $\tilde{t}$, from the set $\mathcal{T}$, where store t is selected with probability $wp_t d_k + (1-w)\left(\sum_{s \in \mathcal{S}} o_{ts}\right)^2$.
        Use roulette-wheel selection to select a size, $\tilde{s}$, from the set $\mathcal{S}$, where store s is selected with probability $q_{s\tilde{t}}$.
        $d_{\tilde{t}\tilde{s}k} = d_{\tilde{t}\tilde{s}k} + 1$
        if $o_{\tilde{t}\tilde{s}} > 0$ **then**
            $a_{\tilde{t}\tilde{s}k} = a_{\tilde{t}\tilde{s}k} + 1$
            $o_{\tilde{t}\tilde{s}} = o_{\tilde{t}\tilde{s}} - 1$
        end
    end
end



(a) Weekly demand for Subclass A_S on a company level.



(b) Weekly demand for Subclass A_S in Sizes 3, 4 and 5.

Figure 5.1: Weekly demand for Subclass A_S (a) on a company level and (b) on a size level.

generated on a company level by using a regression forecasted value as the mean value for a random variable. In some cases, a negative regression value may be possible. In these cases, the mean value is assumed to be 0. However, this did not happen very often in experiments.

According to Assumption 4, demand follows a Poisson or normal distribution. Experiments were conducted using both the Poisson and the normal distribution to generate demand. In experiments with the normal distribution, the demand that was generated was rounded to the nearest integer, as demand cannot be a fraction. Because there are not enough years of historical data available, the parameters for a normal distribution cannot be deduced from the data. The normal distribution is commonly used in the generation of demand [1]; therefore it is a fair assumption that demand follows a normal distribution. However, difficulty was experienced in determining the standard deviation for the normal distribution. The standard deviation cannot be estimated from historical demand, because there are too few data points. One idea was to use the standard deviation obtained from the regression equation, but then the assumption is made that the standard deviation of demand is constant throughout the season. But the fact that this is not true is precisely the reason for the heteroscedasticity in initial experiments. Another idea that was considered is to use the square rooted mean and standard deviation values and square it afterwards, but this results in a chi-square instead of a normal distribution. Another problem with the normal distribution is that negative values are sometimes simulated, and a negative demand is not possible. The best approach to handle that is to make negative values 0; however, this can potentially result in a skew distribution which is not normal any more. Therefore, the Poisson distribution is used in the generation of demand in the final model. The Poisson distribution has only one parameter and is more suitable than the normal distribution because it is a discrete distribution.

In lines 9–15, demand and sales for week k in all sizes and at all stores are initialised with a value of 0, and the stock on hand is updated with the inflows for the week. The inflows depend on the particular allocation method that is being used, so that different allocation methods can be compared to one another by changing the inflows in the simulation model and testing the effect on sales.

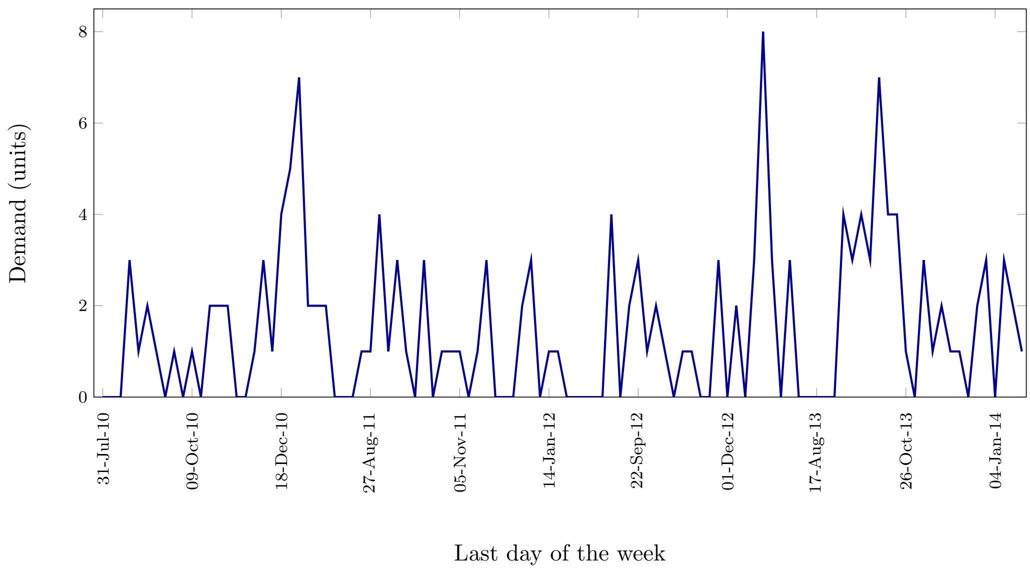
In lines 16–24, each unit of demand is first assigned to a store (line 17) and then to a size (line 18) at that store, after which sales take place if stock is available (lines 20–23). The probability of a given unit of demand occurring at store t is based on the historical proportion of demand at that store to total demand, as well as availability. A weight w is associated with the historical proportion of demand and a weight $(1 - w)$ with availability. The formula is derived using Assumption 1 of the simulation model which states that future demand can be derived from historical sales, Assumption 3 which states that availability has an influence on demand, as well as the regression equation. The second term is squared, because in the regression equation, there is a quadratic relationship between Y_{k-1}, which represents historical demand in the regression equation, and L_k, which has an impact on availability. The probability of demand in a certain size occurring at a specific store is not influenced by availability, because availability does not influence a particular person's dress or shoe size.

The assigned demand is updated in line 19. If one unit of demand in size s was assigned to store t and stock is available, one unit of sales in size s at store t is realised in week k (line 21). If stock is not available, a stockout takes place. If a sale took place, the stock on hand is updated in line 22.

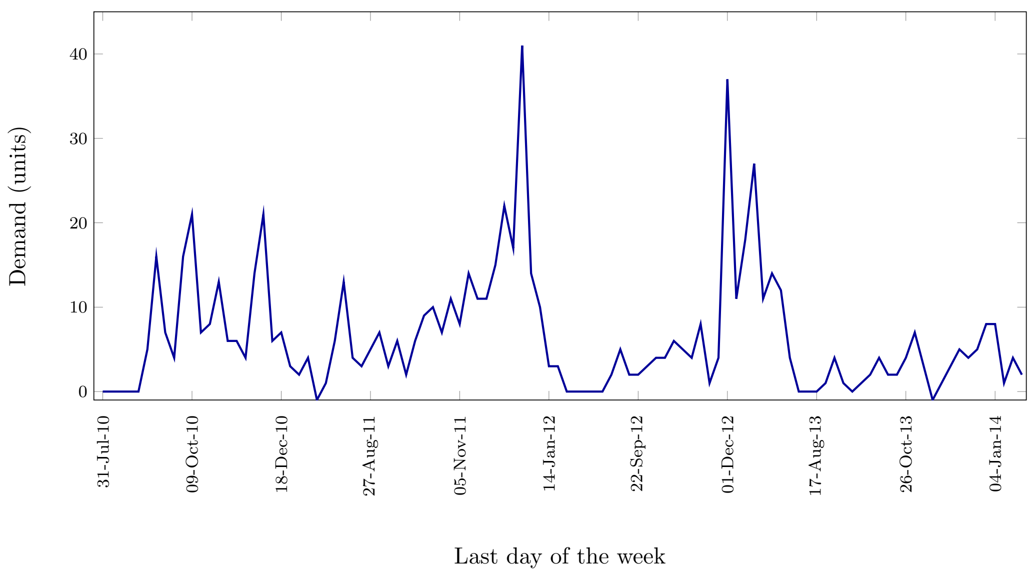
5.3.2 Development of the regression model

In this section, more detail is given on the regression equation that is used in line 7 of the pseudocode. The regression is based on historical demand, which is determined by means of Algorithm 1.

The weekly demand for Subclass $\mathrm{A_S}$ at two stores, one in Hlatikulu and one in Rylands, Cape Town, for the years 2010–2013, is given in Figure 5.2. The demand is summed over all sizes in both cases. The store in Hlatikulu has an average demand of less than 2 units per week for this subclass, and no pattern can be distinguished in the weekly demand. About 21% of all stores that were considered in experiments with Subclass $\mathrm{A_S}$ has an average demand per year less than or equal to the demand in Hlatikulu. The demand for the store in Rylands exhibits more of a pattern, but even for this store the best adjusted R^2 value that could be achieved was 0.53, and the total forecasted demand overestimated the actual total by almost 88%. About 90% of stores in Subclass A_S has an average demand that is less than or equal to the demand in Rylands. It was concluded that regression on a store level is not practical.



(a) Weekly demand at Hlatikulu for Subclass A_S.



(b) Weekly demand at Rylands, Cape Town, for Subclass A_S.

Figure 5.2: Weekly demand for Subclass A_S at a store in (a) Hlatikulu and (b) Rylands, Cape Town, for the years 2010–2013.

Therefore, it was decided to perform the regression and simulation on a company level, and divide the total demand among stores afterwards. The weekly demand on a company level for Subclass A_S for the years 2010–2013 is given in Figure 5.3. The outlier that was described in §3.1.3 was adjusted on a company level by using the average demand during the corresponding week of 2010, 2012 and 2013, rounded to the nearest integer.

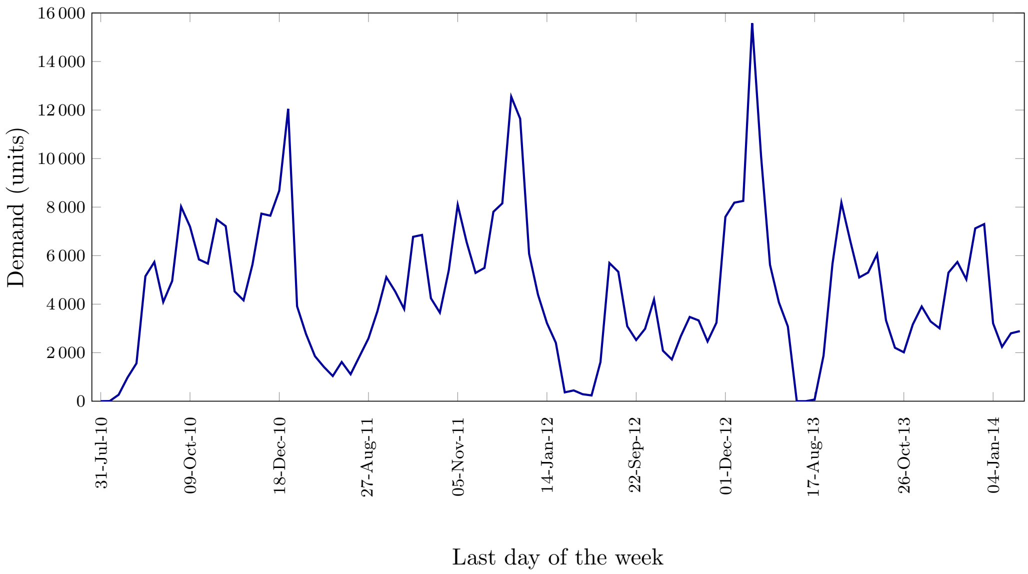


Figure 5.3: Weekly demand on a company level for Subclass A_S, for the years 2010–2013.

As is often the case with sales or demand data, a clear seasonal pattern is discerned. The pattern repeats itself every 26 weeks (the length of a season). The seasonality confirms the suitability of a multiple regression model with seasonal variables. Peaks are present at around the end of every month, and larger peaks are observed close to Christmas.

Based on the pattern of the demand time series, PEP's input and many experiments, the following variables are defined for inclusion in the model. Let $\mathcal{K} = \{1, 2, \ldots, k, \ldots, K\}$ be the set of weeks in a season, and let

Y_k be the total demand in week k,
L_k be the total unit inflow during week k,
W_k be week k's week number of the year, where Sunday is considered the first day of the week,

$$E_k = \begin{cases} 1 & \text{if the last day of week } k \text{ is after the } 29^{\text{th}} \text{ or before the } 11^{\text{th}} \text{ of a month, excluding} \\ & \text{the end of December and the beginning of January,} \\ 0 & \text{otherwise, and let} \end{cases}$$

$$C_k = \begin{cases} 1 & \text{if the last day of week } k \text{ falls in the interval from the } 17^{\text{th}} \text{ to the } 30^{\text{th}} \text{ of December,} \\ 0 & \text{otherwise.} \end{cases}$$

Demand in week k, Y_k, is the dependent variable, and $\hat{Y}_k$ is the forecasted value for Y_k. The inclusion of independent variables was finalised in experiments where a significance level of 0.05 was used.

It is a well-known belief in PEP that availability has an influence on demand. This was confirmed by the fact that L_k is a significant explanatory variables for Y_k. A variable indicating stock on hand was also experimented with. The variable was significant, but its p value was higher than the p value of L_k. As stock on hand is a function of inflows, using both variables leads to multicollinearity; therefore, only L_k was used.

The variables Y_{k-1} and L_{k-1} were included in the model because, during initial experiments, positive autocorrelation was present. After including Y_{k-1} and L_{k-1}, the autocorrelation was removed. Because the weeks in between seasons were not included in the model, Y_{k-1} and L_{k-1} were assumed to be 0 during the first week of each season. When forecasting for a whole season in advance, the true value for Y_{k-1} is not yet known, and the forecasted value, $\hat{Y}_{k-1}$ is used as a proxy for the actual value. The variable Y_{k-26} is also a significant explanatory variable of demand, but it was decided not to include this variable, because a whole season's data will be lost, and there are already very little historical years of data available.

The variables E_k, C_k and W_k were defined based on the pattern of the data. The purpose of the dummy variable E_k is to handle the small peaks, which normally falls between the end of a month and the beginning of the next month, after most of PEP's customers have received their salaries and wages. The exact definition of E_k was settled by inspecting the data and by experimenting. Unlike other months, demand does not go up near the end of December, because most customers have little spending money left after Christmas time. The dummy variable C_k is defined to include two weeks around Christmas, when demand usually peaks because customers buy Christmas presents. Variable W_k was included in the model to handle further seasonal patterns in the data. The value of W_k ranges from 0 tot 53. Week 1 begins on the first Sunday of the year, so that, if the 1$^{\text{st}}$ of January is not a Sunday, the first Saturday of the year is assigned a week number of 0.

Separate dummy variables for each month's small peak were experimented with, but these variables were less significant than the one dummy variable that was settled on. Explanatory variables indicating the month in which the week falls were also tried. The number of the month have been tried, as well as monthly dummy variables. Month and week numbers were not significant when they were both included in the model, and week number had a lower p value; therefore week number was included and month number not. When monthly dummy variables were included, the signs of the coefficients in the regression equation suggested that demand is higher in January than in other months, which is clearly not true. This anomaly is an indication of multicollinearity; therefore, monthly dummy variables were not used.

In other experiments, data were grouped in month or two-week periods instead of week periods, but the results for this approach were poorer than that of the final model. Initially, all weeks for which sales were recorded were included in the model, without ensuring that each season had exactly 26 weeks. It was also considered to include weeks between seasons in the model, with 0 sales and inflows during these weeks. All of these experiments resulted in autocorrelation, and were therefore not continued.

With reference to the article by Gaur *et al.* [26], financial indicators were also included during experiments. JSE price and index values, the GDP per capita, oil prices, CPI index values, disposable income of households and the repo rate were experimented with. None of the financial indicators proved to be significant explanatory variables of demand.

Having established the variables, $\hat{\beta}_1, \hat{\beta}_2 \ldots \hat{\beta}_6$ were estimated in the following regression model

$$\hat{Y}_k = \hat{\beta}_1 Y_{k-1} + \hat{\beta}_2 L_k + \hat{\beta}_3 L_{k-1} + \hat{\beta}_4 W_k + \hat{\beta}_5 E_k + \hat{\beta}_6 C_k. \tag{5.1}$$

The intercept was initially included in the model, but the p value was higher than 0.05, indicating that it is not statistically significant. The intercept was also not significant in the other Summer data set that was received for experiments, so it was assumed that the intercept is generally not significant for Summer products. This is a reasonable assumption, because it makes sense that, if the values of all variables are 0, demand is 0. In other words, if demand during the previous week was 0, there were no inflows during this or the previous week, if the week number is 0 and the week is not near the end of a month and it is not Christmas, demand is 0. Therefore no intercept was included in the model. The resulting model is given by

$$\hat{Y}_k = 0.57 Y_{k-1} + 0.06 L_k + 0.05 L_{k-1} + 17.33 W_k + 1382.60 E_k + 3918.70 C_k. \tag{5.2}$$

Intuitively, the regression equation makes sense from analysing historical data. When a positive demand occurred during the previous week, demand is more likely to be higher during this week. Inflows during this and the previous week increases demand during this week, because a higher availability increases demand (Assumption 3). As explained before, it is known that demand is higher during the Christmas season as well as near the end of the month. The Christmas peak is higher than the other peaks; therefore it makes sense that the coefficient of C_k is higher than the coefficient of E_k.

However, a plot of the residuals against predicted values for regression (5.2), shown in Figure 5.4, indicate that the size of the error increases as the predicted value increases. This is a clear indication of heteroscedasticity, which causes regression results to be unreliable.

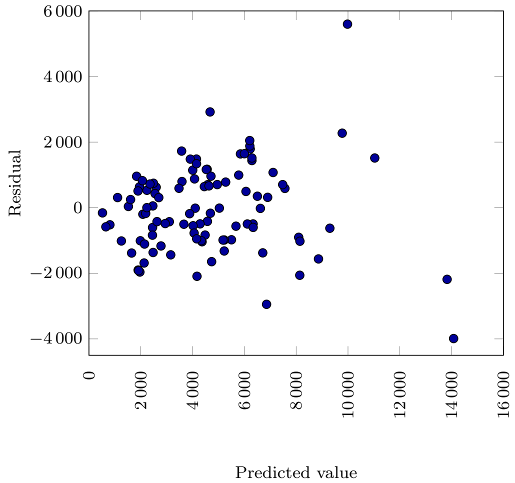
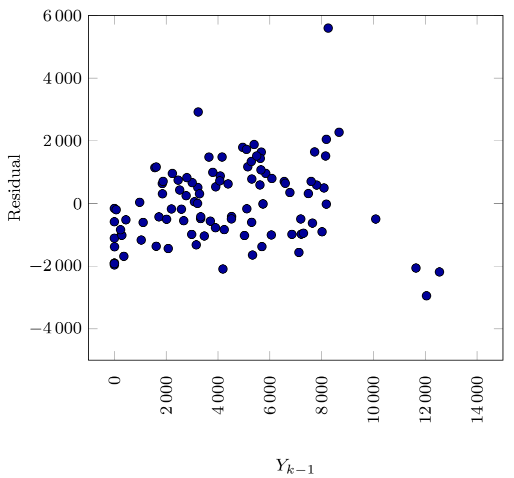


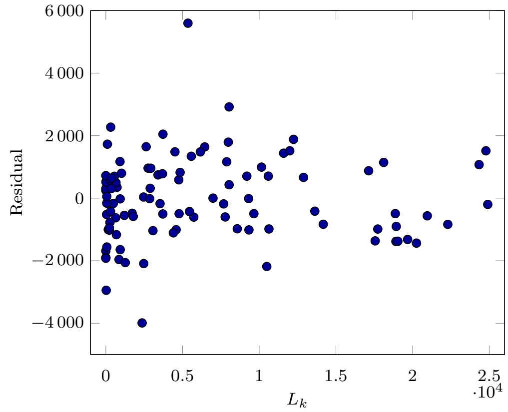
Figure 5.4: Residuals against predicted values for regression (5.2).

In Figure 5.5, residuals are plotted against lagged demand, inflows and lagged inflows to find a possible reason for the heteroscedasticity. The plot in Figure 5.5(a) indicates that the size of residuals increases as lagged demand increases. This effect is not seen in the other two plots. It is concluded that lagged demand is the cause of the error.

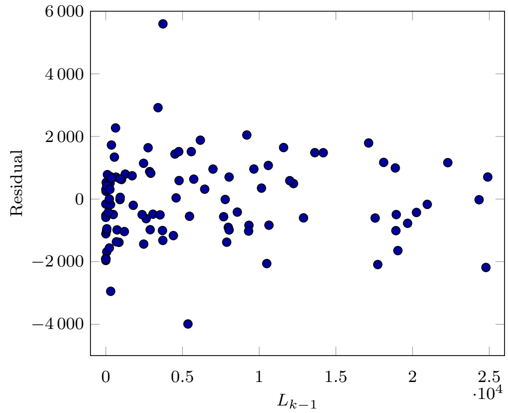


(a) Residuals against Y_{k-1}

In an attempt to remove the heteroscedasticity from the data, $\sqrt{\hat{Y}_k}$ and $\sqrt{Y_{k-1}}$ are used in the



(b) Residuals against L_k



(c) Residuals against L_{k-1}

Figure 5.5: Residuals plotted against (a) lagged demand, (b) inflows and (c) lagged inflows for regression (5.2).

place of $\hat{Y}_k$ and $\hat{Y}_{k-1}$. Because a 0 demand is possible, using the natural log of $\hat{Y}_k$ and $\hat{Y}_{k-1}$, which would be another typical solution for removing heteroscedasticity, would lead to undefined values.

The new regression model for Subclass $\mathrm{A_S}$ is given by

$$\sqrt{\hat{Y}_k} = 0.73\sqrt{Y_{k-1}} + 0.00057L_k + 0.00041L_{k-1} + 0.21W_k + 9.53E_k + 17.19C_k. \tag{5.3}$$

If the residuals are again plotted against predicted values, the graph in Figure 5.6 is obtained. This time, there is no pattern that indicates the presence of heteroscedasticity.

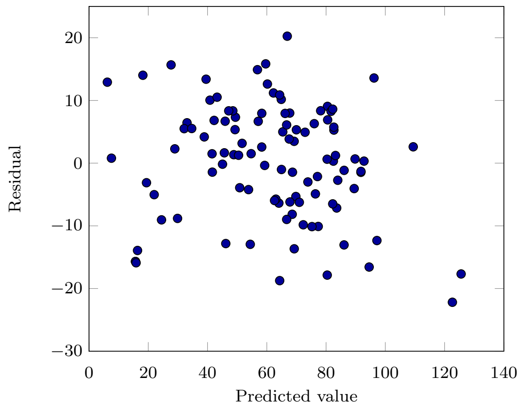


Figure 5.6: Residuals against predicted values for regression (5.3).

The new regression equation also makes intuitive sense for the same reasons as regression (5.2). The R^2 value is 0.981 and the adjusted R^2 value is 0.979, indicating a very good fit. The explanatory power of each variable was also established by hypothesis testing. The t test values associated with each independent variable as well as the accompanying p values are given in Table 5.1.

Variable	t value	p value
$\sqrt{Y_{k-1}}$	24.06	< 0.0001
L_k	3.97	0.0001
L_{k-1}	2.80	0.0062
W_k	3.33	0.0012
E_k	4.37	< 0.0001
C_k	4.40	< 0.0001

Table 5.1: The t test value and the accompanying p value for each independent variable in regression (5.3).

The p values indicate that all variables are significant explanatory variables at a significance level of 0.05. The p value for the F test, obtained from SAS 9.3 [61], is smaller than 0.0001; therefore, the null hypothesis that all regression coefficients are 0 is rejected and it is concluded that the model as a whole is significant in explaining demand.

The results so far indicate that regression (5.3) is suitable to forecast demand for Subclass $\mathrm{A_S}$. However, the validity of the model must be determined by testing that the assumptions of regression hold.

5.3.3 Validation and accuracy of the regression model

Firstly, the coefficients in regression equation (5.3) are constants, so the regression is linear in parameters and Assumption 1 holds. Secondly, a modified version of the Breusch and Pagan [9] test by Koenker [36] was performed to formally test for homoscedasticity. For this test, the null hypothesis of homoscedasticity is tested against the alternative hypothesis of heteroscedasticity. The p value for the hypothesis of the Breusch-Pagan test for regression model (5.3), including all variables, is given by 0.06. The value was obtained from SAS [61]. Therefore, at a significance level of 0.05, H_0 is not rejected, and it can be assumed that the error terms are homoscedastic.

To verify that residuals are normally distributed as assumed (Assumption 3), four tests for normality were performed on the residuals of regression (5.3), namely the Shapiro-Wilk test, the Kolmogorov-Smirnov test, the Cramér-von Mises test and the Anderson-Darling test. In each test, the null hypothesis (H_0) is that residuals follow a normal distribution. The results, which were obtained from SAS [61], are given in Table 5.2. The last column indicates whether H_0 is rejected at a significance level of $\alpha = 0.05$.

Test	Statistic	p value	Reject H_0?
Shapiro-Wilk	$W = 0.99$	0.48	No
Kolmogorov-Smirnov	$D = 0.07$	> 0.15	No
Cramér-von Mises	$W^2 = 0.07$	0.25	No
Anderson-Darling	$A^2 = 0.47$	> 0.25	No

Table 5.2: Results for normality tests on the residuals of regression (5.3). A significance level of $\alpha = 0.05$ is used.

In each case, the null hypothesis of normality is not rejected at a significance level of 0.05, so that it can be assumed that errors follow a normal distribution. The reported values for the mean and standard deviations of the errors are 0.4 and 9.41, respectively. The mean value is very close to 0, so it may be assumed that error terms are normally distributed with a mean value of 0.

Assumption 4 of no autocorrelation in the residuals was first tested by means of the Durbin-Watson test, but the result was indecisive. The Durbin-Watson value is 1.63, which is between the lower (1.58) and upper (1.78) Durbin-Watson bounds for 104 observations and 6 degrees of freedom. Therefore, the Runs test, also known as the Geary test, was performed. The null hypothesis of this test is no autocorrelation. Performing the Runs test for regression (5.3) yields a p value of 0.1. The null hypothesis of no autocorrelation is not rejected and it is reasonable to assume that there is no autocorrelation in the regression model.

Finally, Assumption 5, namely that there is no multicollinearity in the independent variables, must hold. One indication of multicollinearity is when coefficient signs do not make intuitive sense. It has already been established that the coefficients in regression (5.3) make sense. Another indication of multicollinearity is when the regression model has a high R^2 value and the null hypothesis of the F test is rejected, but none or very few of the dependent variables are statistically significant explanatory variables. This is not the case either, as all independent variables are significant explanatory variables of demand. Inspecting the pairwise correlation coefficients among independent variables is another method of detecting possible multicollinearity. Usually, if the absolute value of a correlation coefficient between two coefficients is higher than 0.8, multicollinearity is regarded a problem. The pairwise correlation coefficients of the independent variables in regression (5.3) are given in Table 5.3.

No correlation coefficient between two different variables is bigger than 0.8 or smaller than -0.8. Therefore there is no indication that multicollinearity plays a significant role, and it may

	$\sqrt{Y_{k-1}}$	L_k	L_{k-1}	W_k	E_k	C_k
$\sqrt{Y_{k-1}}$	1	-0.041	0.228	0.198	-0.171	0.355
L_k	-0.041	1	0.085	0.222	0.232	-0.028
L_{k-1}	0.228	0.085	1	0.277	0.028	0.011
W_k	0.198	0.222	0.277	1	0.188	0.300
E_k	-0.171	0.232	0.028	0.188	1	-0.201
C_k	0.355	-0.028	0.011	0.300	-0.201	1

Table 5.3: Pairwise correlation coefficients of independent variables in regression (5.3).

be assumed that there is no multicollinearity among independent variables.

Regression (5.3) is therefore valid, as all the assumptions of regression hold. The t and F tests indicate that all independent variables are significant, and the high R^2 and adjusted R^2 values indicate a good fit. A graph of the time series of the fit and forecast against actual values can be found in Figure 5.7.

It is clear from the graph that the values of the fit and the forecast are close to the actual values, although the regression model overestimates demand a little in 2014. Actual total demand in 2014 was 90 855, and the total predicted by the regression model is 93 106.5. This amounts to a 2.48% overestimation in total demand. This phenomenon is due to an overestimation in demand during the first week. The second week's demand was also overestimated because the first week's forecast was used as a proxy for the lag variable of demand. The overestimation in the second week's demand in turn led to an overestimation in the third week, and so on.

Demand is overestimated during the first week as well as the following four weeks, because demand during these weeks were relatively lower than during the corresponding weeks in historical years. Another possible reason for the overestimation is that inflows were relatively high during these weeks, but actual demand was not as high as would have been expected given the high inflows. The assumption made during regression modelling is that each extra unit of inflows increases demand, but in reality, this is only true to a certain extent. Demand is also overestimated during the last two weeks of November, because actual demand during those weeks was low relative to demand during the corresponding weeks in other years. Demand is underestimated during the last two weeks of the season because demand during these weeks was relatively high compared to the corresponding weeks in previous years.

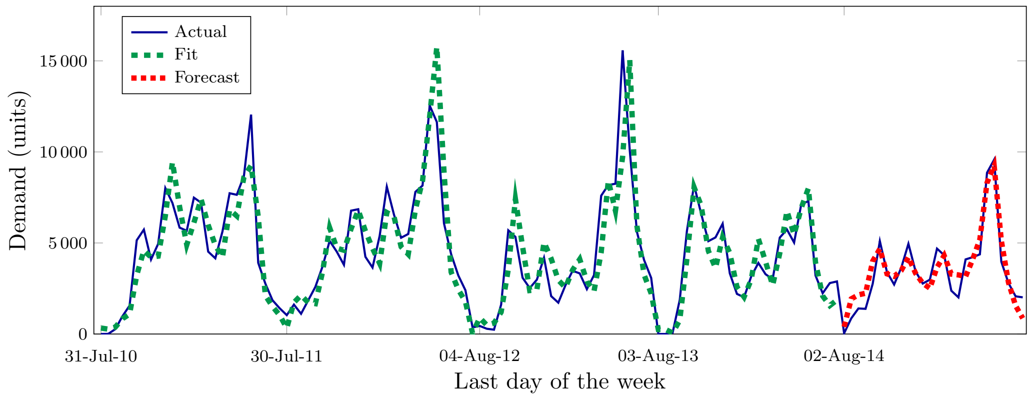


Figure 5.7: Graphical display of the fit and forecast of regression (5.3) in years 2010–2014.

5.3.4 The necessary number of simulation replications

The random variation in stochastic simulation models is usually handled by replicating the simulation experiment a number of times and reporting on the average results. It is important to determine the number of replications necessary to obtain certainty in the reliability of results [10].

A formula that is often used to find the number of necessary replications [2, 10, 70] is based on a t confidence interval for the estimate of the mean μ from m initial replications. Given m initial replications, let

$N(m)$ be the number of replications required,
$\bar{X}(m)$ be the estimate of the real mean μ,
$S(m)$ be the estimate of the real standard deviation σ,
α be the level of significance used,
ϵ be the allowable percentage error of the estimate $\bar{X}(m)$, where $\epsilon = \left|\bar{X}(m) - \mu\right| / |\mu|$, and let
$t_{m-1,1-\alpha/2}$ be the critical value of the two-tailed t-distribution at a significance level of α, given $m-1$ degrees of freedom.

Then $N(m)$, the number of replications required, is given by

$$N(m) = \left(\frac{S(m)t_{m-1,1-\alpha/2}}{\bar{X}(m)\epsilon}\right)^2. \tag{5.4}$$

In order to use the t confidence interval, the results from the simulation replications (in this case, total sales) is required to be independent, identically and normally distributed [10]. As the total sales generated by one replication does not influence the total sales generated by any other replication, it is safe to assume that the simulation replications are independent. The Shapiro-Wilk, Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling tests were performed on 100 values of total sales from Data set $\mathrm{A_S}$, each value generated by a different replication of Simulation Model $\mathrm{S_1}$. The simulation was implemented in Python 3.3 [55]. In these experiments, actual inflows as obtained from the real data were used, and the value of w in the calculation of store probabilities was chosen as $w = 0.99$. The results as obtained from SAS [61] are given in Table 5.4. A significance level of $\alpha = 0.05$ is used to decide whether to reject the null hypothesis.

Test	Statistic	p value	Reject H_0?
Shapiro-Wilk	$W = 0.99$	0.24	No
Kolmogorov-Smirnov	$D = 0.050$	> 0.15	No
Cramér-von Mises	$W^2 = 0.037$	> 0.25	No
Anderson-Darling	$A^2 = 0.33$	> 0.25	No

Table 5.4: Results for normality tests on total sales from 100 simulation replications. A significance level of $\alpha = 0.05$ is used.

In each case, the null hypothesis of normality is not rejected, and it may be assumed that simulated sales are normally distributed. Formula (5.4) may therefore be used.

An initial number of $m = 10$ replications of Simulation Model $\mathrm{S_1}$ were performed on Subclass $\mathrm{A_S}$ in order to find the number of replications that should be performed for sufficient reliability. The mean value for total sales in the experiment was $\bar{X}(m) = 90\,248.6$ and the standard deviation $S(m) = 272.66$. The allowable percentage error was chosen as 0.01 and a significance level $\alpha = 0.05$ was used. The critical t value $t_{9,0.975} = 2.262$. Substituting these values into formula (5.4) yields $N(m) = 0.47 \approx 1$ simulation replication. Thus, if at least one simulation run is performed, there is a 95% certainty of an error of less than 0.01. As solution time is not an important factor, it was decided to perform ten simulation replications in further experiments. A precision of 0.0022 (about 200 units) can then be obtained at a significance level of 0.05.

5.3.5 Verification and validation of the simulation model

The process of verification and validation (VV) is essential to be able to have confidence in the results of a simulation model. Verification involves ensuring that the conceptual model has been correctly computerised. Validation is performed to ensure that the conceptual model is sufficiently accurate for its purpose. According to Robinson [59], there are four important VV techniques, namely conceptual model validation, data validation, verification and white box validation, and black box validation [59].

During conceptual model validation, the modeller ensures that the model possesses sufficient detail to meet the objectives of the simulation study, and that assumptions are correct [59]. The assumptions of the simulation model in §5.2 have been made in collaboration with employees from PEP, who has an in-depth knowledge of the real system, and by analysing data from the real system. This ensured that the conceptual model does possess the necessary detail to represent reality satisfactorily, and that assumptions are realistic. An important part of conceptual model validation was also done in §5.3.3 when the regression model was validated by testing that regression assumptions hold.

Data validation means to validate that the data used in model building, validation and experimentation are sufficiently accurate [59]. Real world data from PEP have been used during model building and in all experiments. There are some slight errors in the data that have already been discussed, but they are small enough to have no effect on results. During data handling, adjustments have been made following the best available approaches in order to represent reality as closely as possible.

While verification ensures that the conceptual model has been accurately computerised, white box validation ensures that every part of the model is accurate compared to the real world. Both of these processes are performed continuously during model building [59]. The code for Simulation Model S_1 has continuously been checked for possible errors during model building as a first step in model verification, and no errors have been found. Stock keeping for each individual store has been verified and calculations have been done correctly. In one validation experiment, disproportionally small demands were artificially generated, and the model responded as expected: stock in each size at each store accumulated into a large number by the end of the season. Similarly, disproportionally large demands were generated, and stores ran out of stock as expected. Another important experiment involved an allocation system where stock was deliberately sent to stores with a low demand at the expense of stores with a high demand. This caused sales to decrease significantly, as expected.

Black box validation is performed to determine whether the model as a whole is an accurate representation of the real world. This usually involves statistical tests to measure how closely the simulation model resemble the real world and is only done once the model has been completed [59]. Sales for Subclass A_S as generated by Simulation Model S_1 were compared to real world sales using intraclass correlation coefficients.

The intraclass correlation coefficient (ICC) is a statistic that can be used when two or more measures called judges, are applied to assess or score different objects of measurements called targets. The ICC indicates the correlation among the scores of the targets by the different judges [47, 64]. In this case, there are two judges: the real system and the simulation model. The objects are sales observations.

Bartko [7] introduced a notation system where judges are represented by columns from the set $\{1, 2, \ldots, j, \ldots, k\}$ and targets are represented by rows from the set $\{1, 2, \ldots, i, \ldots, n\}$. Then x_{ij} represents the score of target i given by judge j. The notation is illustrated in Table 5.5.

Target	Judge					
	1	2	$\cdots$	j	$\cdots$	k
1	x_{11}	x_{12}	$\cdots$	x_{1j}	$\cdots$	x_{1k}
2	x_{21}	x_{22}	$\cdots$	x_{2j}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	x_{i1}	x_{i2}	$\cdots$	x_{ij}	$\cdots$	x_{ik}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	x_{n1}	x_{n2}	$\cdots$	x_{nj}	$\cdots$	x_{nk}

Table 5.5: ICC table notation for target scores given by judges.

Different forms of the ICC exist, each based on an additive variance model. The one used in this study is based on an analysis of variance (ANOVA) model. The underlying ANOVA model can be either one-way or two-way. A one-way model is used if each target is rated by a different set of judges, and a two-way model is used if each target is rated by the same judges. In this case, actual sales are compared to simulated sales for each sales observation; therefore, a two-way model is used [47, 64].

The model can also be either random or mixed: random if a sample of judges was selected and mixed if the judges are the only judges of interest [47, 64]. There could potentially exist other models to simulate sales; therefore, the sample model is appropriate.

A two-way random ICC is based on the ANOVA model given by

$$x_{ij} = \mu + r_i + c_j + a_{ij} + e_{ij}, \tag{5.5}$$

where

- μ is the population mean for all observations,
- r_i is an IID normally distributed variable with mean $\mu = 0$ and variance σ_r^2 representing effects caused by the i^{th} target,
- c_j is an IID normally distributed variable with mean $\mu = 0$ and variance σ_c^2 representing effects caused by the j^{th} judge,
- a_{ij} is an IID normally distributed variable with mean $\mu = 0$ and variance σ_a^2 representing effects caused by the interaction between the i^{th} target and the j^{th} judge, and
- e_{ij} is an IID normally distributed variable with mean $\mu = 0$ and variance σ_e^2 representing effects caused by the residuals.

A two-way random ICC can be classified as a single score ICC or an average score ICC. In the case of a single score, the correlation among the scores by different judges for each separate observation is measured; in the case of an average score, the correlation among the average scores for observations by the k judges is given. A single score measure thus indicates the reliability of a single observation, and an average score measure indicates the reliability of the observations on average. Furthermore, any one of these ICCs can measure the degree of absolute agreement among target scores or the degree of consistency among target scores. The difference between the two is that the consistency measure ignores the variance of effects caused by the judges. What matters for the consistency measure is that the relationship between the targets as scored by one judge should be similar to the relationship between the targets as scored by another judge, even if one judge scored higher on average. The different two-way random ICC forms are defined as follows [47, 64].

ICC(A,1) represents the two-way random single score ICC statistic measuring the degree of absolute agreement among target scores,

ICC(C,1) represents the two-way random single score ICC statistic measuring the degree of consistency among target scores,

ICC(A,k) represents the two-way random average score ICC statistic measuring the degree of absolute agreement among target scores, and

ICC(C,k) represents the two-way random average score ICC statistic measuring the degree of consistency among target scores.

The formulas for the four ICC models are given by

$$\begin{aligned}
\mathrm{ICC(A,1)} &= \frac{\sigma_r^2}{\sigma_r^2+\sigma_c^2+\sigma_a^2+\sigma_e^2}, \\
\mathrm{ICC(C,1)} &= \frac{\sigma_r^2}{\sigma_r^2+\sigma_a^2+\sigma_e^2}, \\
\mathrm{ICC}(\mathrm{A},k) &= \frac{\sigma_r^2}{\sigma_r^2+(\sigma_c^2+\sigma_a^2+\sigma_e^2)/k}, \quad \text{and} \\
\mathrm{ICC}(\mathrm{C},k) &= \frac{\sigma_r^2}{\sigma_r^2+(\sigma_a^2+\sigma_e^2)/k}.
\end{aligned}$$

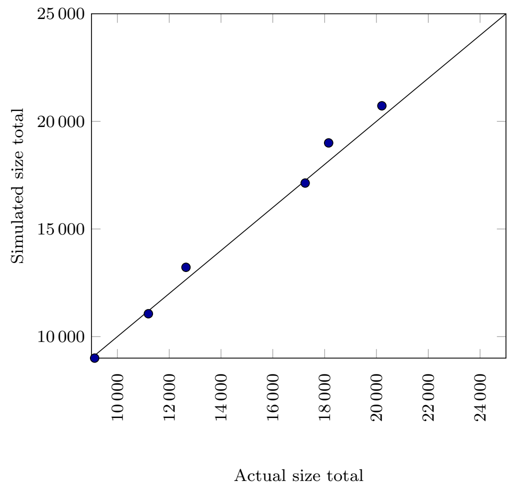
The ICC statistics for Simulation Model S_1 as applied to Subclass S_A are given in Table 5.6. Inflows that actually took place in 2014 were used during simulation experiments. A value of $w = 0.99$ was used. The reported results are for the total average sales of ten replications, compared to actual sales. As it is important to obtain accurate results on a size, week and store level, simulated size totals, week totals and store totals were compared to the actual totals. All ICC values are close to 1, indicating a very high correlation between actual and simulated sales.

ICC statistic	Totals grouped by	ICC value
	size	0.995
ICC(A,1)	week	0.936
	store	0.957
	size	0.995
ICC(C,1)	week	0.934
	store	0.957
	size	0.997
ICC(A,k)	week	0.967
	store	0.978
	size	0.998
ICC(C,k)	week	0.966
	store	0.978

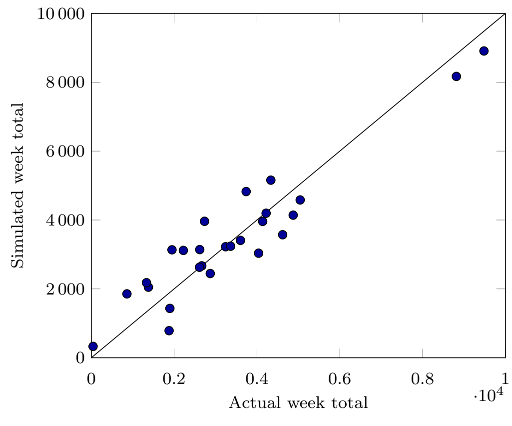
Table 5.6: ICC values for Simulation Model S_1 applied to Subclass A_S. Simulated size totals, week totals and store totals were compared to actual totals. A value of $w = 0.99$ was used.

Scatter plots of simulated sales totals against actual sales totals on a size, week and store level, are given in Figure 5.8. In each case, the line $y = x$ is supplied to indicate the correlation between simulated and actual sales.

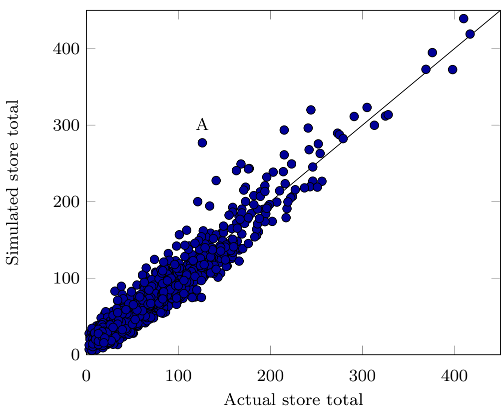
The points in Figures 5.8(a) and Figure 5.8(b) are evenly spread around the line $y = x$. This indicates that simulated size and week totals are close to actual values, with no outliers. In Figure 5.8(c), most of the points are in the bottom left-hand part of the graph, because most the stores had a demand of less than 200 units in 2014. The point labelled 'A' is an outlier: simulated sales overestimated actual sales by a disproportional amount. This point represents the sales of a store in Rehoboth in Namibia. Demand for this store was overestimated in the simulation model because this store's average historical proportion of total demand in a year was 0.0048, while the actual proportion in 2014 was only 0.0014. There are also a few other points close to point A that are tending away from the line $y = x$ for the same reason.



(a) Correlation for size totals.



(b) Correlation for week totals.



(c) Correlation for store totals.

Figure 5.8: Scatter plots indicating the correlation between simulated and actual (a) size totals, (b) week totals and (c) store totals, for Subclass A_S. Simulation Model S_1 was used to simulate sales. The data point labelled 'A' is an outlier.

Excepting a few disproportional errors on a store level, the ICC measures and correlation plots indicate that the sales predicted by the simulation model is sufficiently close to actual sales. This completes the black box validation.

5.3.6 Graphical results

In this section, graphical results are given to illustrate the accuracy of simulated sales. Actual inflows were used in all simulations.

Figure 5.9 contains a graphical representation of the results for one simulation replication of Simulation Model S_1, applied to Subclass A_S. Sales were summed to a company level for every week. The graph illustrates that weekly results are very accurate on a company level. Sales are overestimated during the first few weeks and the last two weeks in November, and underestimated during the last two weeks, due to inaccurate estimates of demand by the regression model.

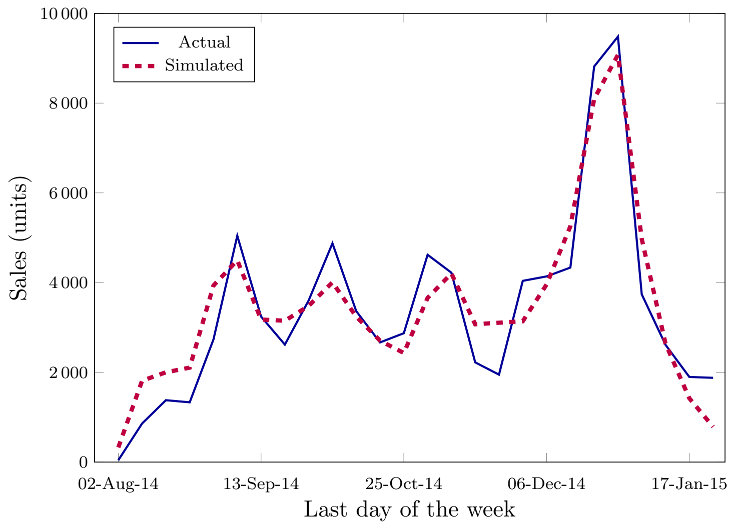


Figure 5.9: Weekly simulated sales for Subclass A_S on a company level, compared to actual sales. Simulation Model S_1 was used to simulate sales.

In Figure 5.10, a graphical representation is given of results for one simulation replication of the same simulation model and subclass on a size level. The size that was used for this particular graph is Size 3, but results for other sizes are similar. As can be seen from the graph, results are still very accurate, but less accurate than on a company level. Over- and underestimations occur at the same places as on a company level. The Christmas peak is also underestimated because actual demand in Size 3 during Christmas time in 2014 made up a higher percentage of total demand than during historical years, and total demand during Christmas time was already slightly underestimated.

Figure 5.11 contains weekly results for one simulation replication of the same model and subclass for Size 3 on a store level. Figure 5.11(a) contains the results for a store in Simunye, Swaziland. This store only sold three Size 3 units during the whole Summer of 2014. Because these sales took place during random weeks in the season, it is not possible to correctly predict the weeks during which sales occur. Sales are not, for example, higher during Christmas time or at the end of a month, as is the case with company level sales. However, the total number of sales for the season was predicted correctly by the simulation model. Judging by historical data, about 55% of stores in this subclass sells the same number of units or less for Size 3 in a season, and sales at these stores make up about 22% of total Size 3 sales.

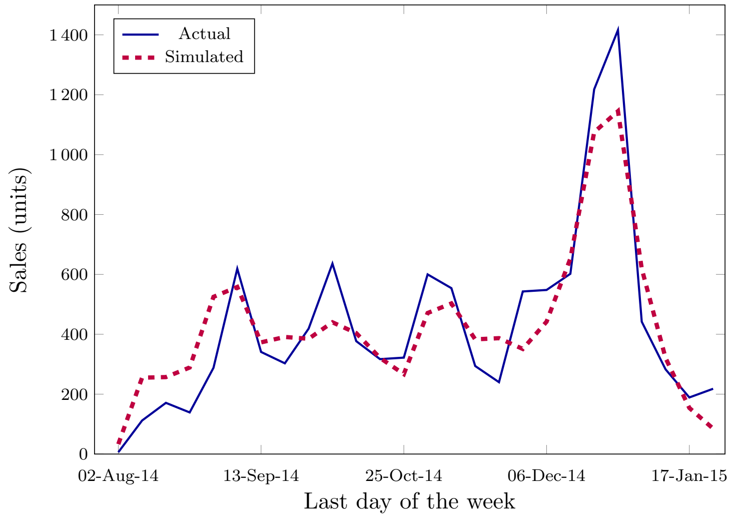
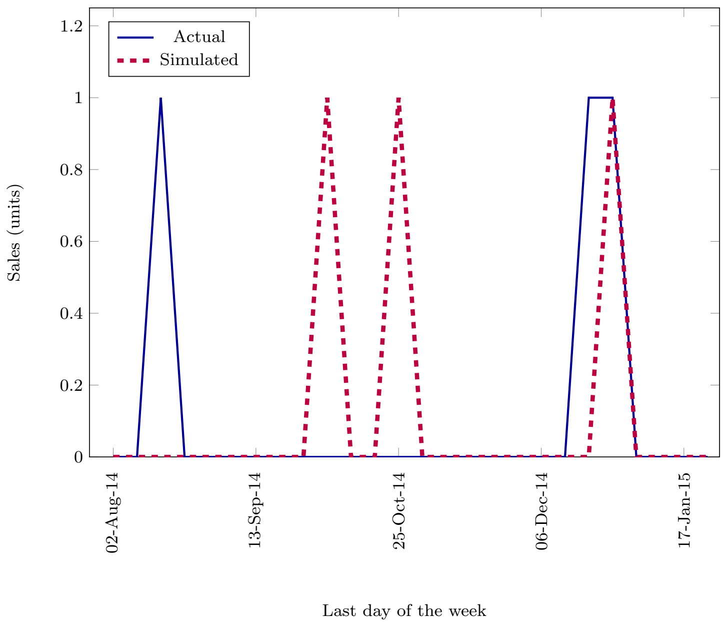


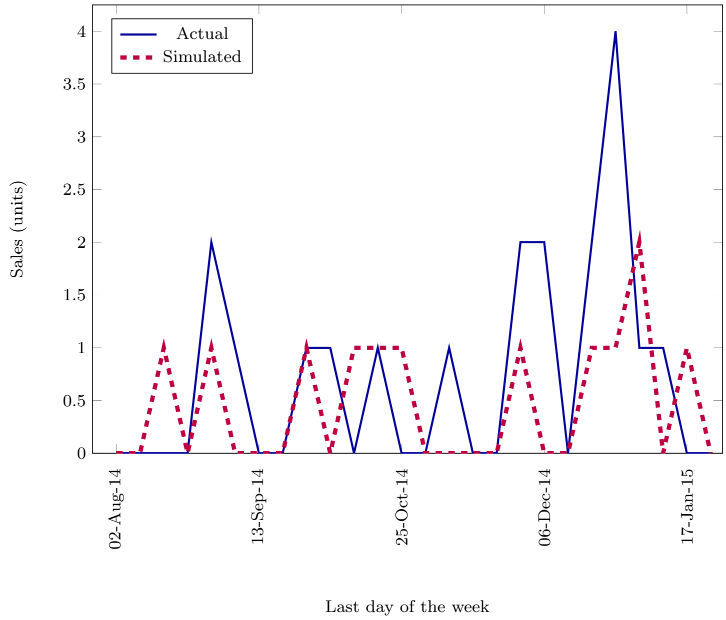
Figure 5.10: Weekly simulated sales in Size 3 for Subclass A_S, compared to actual sales. Simulation Model S_1 was used to simulate sales.

Results for a store in Aberdeen, which can be seen as a medium store in terms of unit sales, are given in Figure 5.11(b). For these and similar stores, simulation results are more accurate than for small stores like the one in Simunye, because there is more of a pattern in unit sales which is similar to the pattern on a group level. About 13% of stores in this subclass sell at least as many Size 3 units per season as the one in Aberdeen, and sales at these stores make up about 44% of total Size 3 sales.

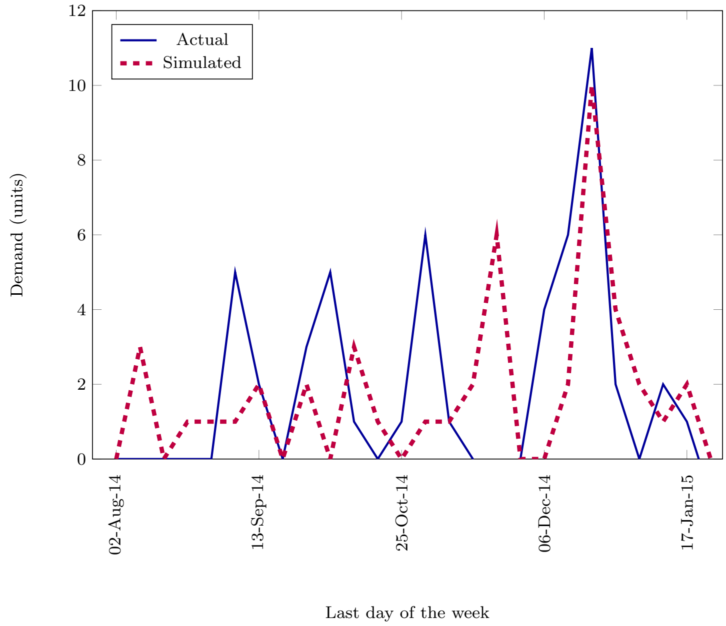
Finally, the results for a store in Parow Centre are given in Figure 5.11(c). The store in Parow Centre can be seen as a large store in terms of unit sales. Results for the store in Parow Centre are even more accurate, because the pattern for this store is clearer than for the one in Aberdeen. Only about 1% of stores in this subclass sell at least as many Size 3 units as Parow Centre in a season, and sales at these stores make up about 9% of total Size 3 sales for the subclass.



(a) Simulated sales in Size 3 at Simunye, Swaziland, compared to actual sales.



(b) Simulated sales in Size 3 at Aberdeen, compared to actual sales.



(c) Simulated sales in Size 3 at Parow Centre, compared to actual sales.

Figure 5.11: Simulated sales in Size 3 compared to actual sales for Subclass A_S at a store in (a) Simunye (a small store), (b) Aberdeen (a medium store) and (c) Parow Centre (a large store). Simulation Model S_1 was used to simulate sales.

5.3.7 Sensitivity of simulation output to changes in w

Ten replications of the simulation model were performed for different values of w. The values of w in experiments are based on the ratio of the coefficients of $\sqrt{Y_{k-1}}$ and L_k, in other words $\hat{\beta}_1 : \hat{\beta}_2 = 0.73 : 0.00057 = 0.99922 : 0.00078$. The value w represents the weight of historical sales in the calculation of the store probability p_t. Judging by the regression equation, a value of more or less 0.999 is appropriate.

The ICCs and total simulated sales for different values of w are given in Table 5.7. Results are reported for the total average sales of ten replications, compared to actual sales, where totals are grouped by size, week and store.

Measure	Totals grouped by	w=0.8	w=0.98	w=0.985	w=0.99	w=0.995	w=0.999
Total sales		90590.7	90515.1	90465.2	90139.4	89728.5	87587.9
	size	0.993	0.993	0.993	0.995	0.995	0.997
ICC(A,1)	week	0.933	0.935	0.936	0.936	0.937	0.935
	store	0.959	0.958	0.957	0.957	0.953	0.931
	size	0.995	0.995	0.995	0.995	0.995	0.997
ICC(C,1)	week	0.932	0.933	0.936	0.934	0.935	0.933
	store	0.959	0.958	0.958	0.957	0.953	0.931
	size	0.996	0.997	0.997	0.997	0.997	0.998
ICC(A,k)	week	0.965	0.967	0.967	0.967	0.967	0.967
	store	0.979	0.979	0.978	0.978	0.976	0.964
	size	0.997	0.997	0.997	0.998	0.998	0.998
ICC(C,k)	week	0.965	0.966	0.967	0.966	0.966	0.965
	store	0.979	0.979	0.978	0.978	0.976	0.964

Table 5.7: Total sales and ICC statistics for the average of ten simulation replications of Simulation Model S_1, for different values of w. Subclass A_S was used during experiments. Simulated size, week and store totals were compared to actual totals.

For lower values of w, sales are higher, because historical sales carry a smaller weight in store demand simulation and availability plays a bigger role. The higher the weight of availability in store demand simulation, the more demand occurs at stores where stock is available, instead of where it historically occurred.

The ICC values for the week and size totals are almost the same for each w; however, for store totals, the ICC values decrease between $w = 0.995$ and $w = 0.999$, and there is a decrease of 2% in total sales. The single measure ICC values for the store totals are slightly higher for $w = 0.99$ than for $w = 0.995$, as well as total sales. For values of w less than or equal to 0.99, the ICC values are almost identical, and sales are also very close together, even if w is set much lower at 0.8. A possible reason for the insensitivity of ICC values for the value of w is that expected store demand was considered during actual stock allocation. PEP's expected store demand is based on the spread of sales in historical data, which is similar to how store demand generation is done in the simulation model. Therefore stock is available in stores where demand was historically higher, so that, in most cases, the same stores are selected during roulette wheel selection irrespective of whether they are selected according to historical sales or availability.

Because the ICC statistics are similar for $w = 0.99$ and lower, and a value close to 0.999 is desirable, $w = 0.99$ was used for all further experiments with data set A_S. The total sales generated with $w = 0.99$, averaged over 10 simulation replications, are 90 139.4 units. Actual total sales are 88 581. The final simulation overestimates sales by about 2%, but this is in line with the overestimation of demand in the regression model.

An accurate value for w may be determined by further experiments; however, this will not significantly change the outcome. The variation of w in experiments proved that total sales and ICC values are not very sensitive to changes in w.

5.3.8 Other approaches to the simulation model

Many different approaches were tried before arriving at the final simulation model. Amongst other, the square root sign in the regression equation caused great difficulty. Another problem was finding a way to generate demand on a store level. Some of the less fruitful experiments in the model building process, together with the reasons for their poor performance, are discussed in this section.

In initial experiments, the square root of simulated demand values was used as a proxy for $\sqrt{Y_{k-1}}$ in the regression instead of $\sqrt{\hat{Y}_{k-1}}$. This caused final total sales to overestimate actual sales by more than the regression. The reason for this phenomenon is that the squared value $\left(\sqrt{\hat{Y}_k}\right)^2 = \hat{Y}_k$ is used as the mean value λ in the Poisson distribution to generate demand. When the square root of the Poisson generated demand is taken, the square root of expected demand is overestimated on average, and the error becomes larger as the weeks progress, because the previous week's demand is used in the calculation of expected demand for this week.

Another approach for generating store level demand was also considered. The idea was to find an expected value for each store by scaling down the company level regression coefficients to a store level. Then Poisson or normal demand is generated for each size and store. The scaling down of the regression coefficients would have been an easy task if the square root signs were not present in the regression equation. Without the square root sign, the coefficients could be scaled down proportionally to each store's historical contribution to demand. Then the final expected demands for stores would add up to the expected demand on a company level. However, the initial calculations for each store demand have to be squared to obtain the actual expected demand, and the squared values do not add up to the squared value of the regression equation. The total demand on a company level could not merely be divided proportionally between stores, because it is known that availability also influences demand on a store level. No solution could be found to handle the problems with this approach.

5.4 Simulation Model S_2

In this section, Simulation Model S_2 is developed. This model follows the second approach, where weekly demand, summed over all stores, is generated separately for each size. The model is an adaptation of Simulation Model S_1, and its development is based on Assumption 7 in §5.2, namely that demand for a size in a store follows the same pattern as demand for that size on a company level. The implication is that the same principles that were applied to Simulation Model S_1 on a company level can be applied to Simulation Model S_2 on a size level and thus each size will have its own regression/simulation model.

5.4.1 The simulation code

The same sets that were defined for Simulation Model S_1 also apply to Simulation Model S_2. The following variables are defined in addition to those defined for Simulation Model S_1. Let

$\hat{Y}_{sk}$ be the expected number of units demand in size s during week k,
d_{sk} be the simulated demand for size s in week k, and let
p_{st} be the average historical proportion of store t's demand in size s to total demand in size s in a season, with $\sum_{t\in\mathcal{T}} p_{st} = 1$ for all sizes in $\mathcal{S}$.

Algorithm 3: Pseudo code of simulation model S_2

1 **for** $s = 1$ to S **do**
2   **for** $t = 1$ to T **do**
3   **for** $s = 1$ to S **do**
4             $o_{ts} = 0$
5   **end**
6   **end**
7   **for** $k = 1$ to K **do**
8   Determine $\hat{Y}_{sk}$ by means of a regression equation.
9   Generate demand d_{sk}, where $d_{sk} \sim \text{Poisson}(\hat{Y}_{sk})$
10   **for** $t = 1$ to T **do**
11   **for** $s = 1$ to S **do**
12                 $d_{tsk} = 0$
13                 $a_{tsk} = 0$
14                 $o_{ts} = o_{ts} + \ell_{tsk}$
15   **end**
16   **end**
17   **for** $i = 1$ to d_k **do**
18   Use roulette-wheel selection to select a store, $\tilde{t}$, from the set $\mathcal{T}$, where store t is selected with probability $wp_{st}d_{sk} + (1-w)\left(\sum_{s\in\mathcal{S}} o_{ts}\right)^2$.
19             $d_{\tilde{t}sk} = d_{\tilde{t}sk} + 1$
20   **if** $o_{\tilde{t}s} > 0$ **then**
21                 $a_{\tilde{t}sk} = a_{\tilde{t}sk} + 1$
22                 $o_{\tilde{t}s} = o_{\tilde{t}s} - 1$
23   **end**
24
25   **end**
26   **end**
27 **end**

Algorithm 3 follows the same principles as Algorithm 2, except that the simulation is done for each size separately, so this algorithm has an extra *for*-loop starting in line 1. Line 18 in Algorithm 2 is omitted, because each size is handled separately, so store demand is already on a size level. An implication of the manner in which this simulation is done is that size demand also depends on availability. This implicit assumption may not be fully realistic.

5.4.2 Regression models

Demand for Simulation Model S_2 is also based on regression as for Simulation Model S_1. A different regression model is implemented for each size as indicated in line 8 of the pseudo-code.

Because it is assumed that demand on a size level follows the same pattern as demand on a company level, the same explanatory variables that were used for Simulation Model S_1 are used for this model, but adapted for size level demand. In addition to the sets and variables that are already defined, let

Y_{sk} be the total demand for size s during week k, and let
L_{sk} be the total unit inflow of size s during week k.

The variables Y_{sk} are the dependent variables and $\hat{Y}_{sk}$ are the forecasted values for Y_{sk} for all sizes s in the set $\mathcal{S}$. Data set $\mathrm{A_S}$ has six sizes ranging from Size 3 to Size 8; therefore, six regression equations were determined. The equations are given by

$$\sqrt{\hat{Y}_{3k}} = 0.70\sqrt{Y_{3,k-1}} + 0.002L_{3k} + 0.0013L_{3,k-1} + 0.08W_k + 4.27E_k + 6.20C_k, \tag{5.6}$$

$$\sqrt{\hat{Y}_{4k}} = 0.70\sqrt{Y_{4,k-1}} + 0.0014L_{4k} + 0.00079L_{4,k-1} + 0.11W_k + 4.69E_k + 8.69C_k, \tag{5.7}$$

$$\sqrt{\hat{Y}_{5k}} = 0.72\sqrt{Y_{5,k-1}} + 0.0012L_{5k} + 0.00085L_{5,k-1} + 0.10W_k + 4.89E_k + 8.80C_k, \tag{5.8}$$

$$\sqrt{\hat{Y}_{6k}} = 0.74\sqrt{Y_{6,k-1}} + 0.0011L_{6k} + 0.00096L_{6,k-1} + 0.091W_k + 3.82E_k + 7.29C_k, \tag{5.9}$$

$$\sqrt{\hat{Y}_{7k}} = 0.76\sqrt{Y_{7,k-1}} + 0.0013L_{7k} + 0.0011L_{7,k-1} + 0.067W_k + 2.87E_k + 5.81C_k, \tag{5.10}$$

and

$$\sqrt{\hat{Y}_{8k}} = 0.77\sqrt{Y_{8,k-1}} + 0.0015L_{8k} + 0.0015L_{8,k-1} + 0.051M_{8k} + 2.46E_{8k} + 4.20C_{8k}. \tag{5.11}$$

The signs and sizes of coefficients make intuitive sense for the same reasons as regression equation (5.3) of Simulation Model $\mathrm{S_1}$ for the same subclass. All R^2 and adjusted R^2 values are above 0.98. Assumption 7 implies that these models are valid because regression (5.3) is valid. It is therefore not necessary to repeat validity tests.

Figure 5.12 contains a graphical representation of the regression fit and forecast of regression (5.6), the regression model for the demand of Subclass $\mathrm{S_A}$ in Size 3. Results for other sizes are similar. The same adjustment that was made on a company level to demand during the week ending in the 24$^{\text{th}}$ of December was also made on a size level.

The graph indicates that the regression fit and forecast for demand closely resembles actual demand. The under- and overestimations during 2014 are similar to the company level model, as expected. A difference between this regression and the company level regression is that the Christmas peak in 2014 is underestimated for this size, probably due to a low Christmas peak in 2010 and 2013 for Size 3.

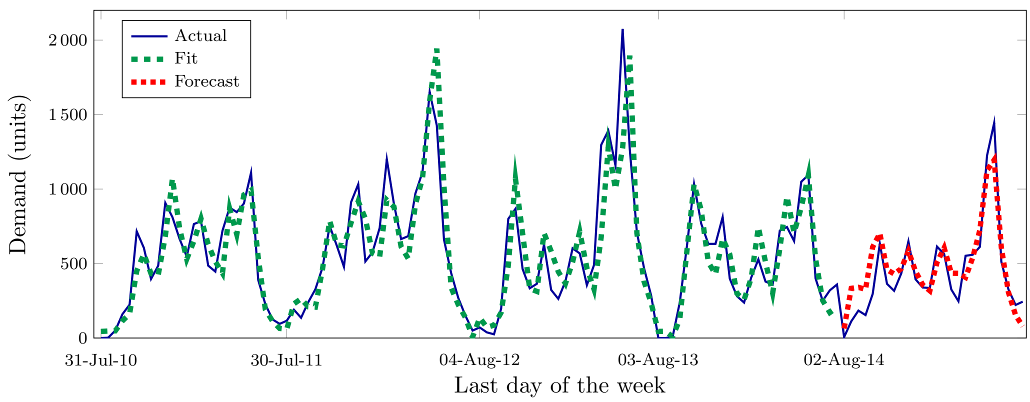


Figure 5.12: Graphical display of the fit and forecast of regression (5.6) (the regression for Size 3) in years 2010–2014. The fit and forecast for regressions (5.7)–(5.11) have similar graphical appearances.

5.4.3 The necessary number of simulation replications

The necessary number of simulation replications was determined by formula (5.4) which was explained in §5.3.5. Because total sales simulated by Simulation Model S_1 are normally distributed, it may, by Assumption 7, be assumed that total sales for each size simulated by Simulation Model S_2 are also normally distributed. If the parameters of two normal distributions are added together, a normal distribution is obtained; therefore, if the total sales for each size is normally distributed, total sales for all sizes added together is also normally distributed. Therefore formula (5.4) may be used for Simulation Model S_2.

Again, 10 initial replications were performed. The mean value for total sales is $\bar{X}(m) = 90\,192.80$ and the standard deviation is $S(m) = 195.78$. The allowable percentage deviation was chosen as 0.01 and a significance level of 0.05 was used. The critical t value $t_{9,0.0975} = 2.262$. Substituting these values into formula (5.4) yields $N(m) = 0.24 \approx 1$ simulation replication. Again, it was decided to perform ten runs to obtain more precision, in this case 0.0016 (about 145 units).

5.4.4 Validation and verification of the simulation model

The first three steps in the VV process is already completed. Simulation Model S_2 is conceptually valid for the same reasons that Simulation Model S_1 is conceptually valid. The same data set, namely the data set for Subclass S_A, was used during the building of Simulation Model S_2 that was used during model building for Simulation Model S_1; therefore, data are valid. The same white box validation experiments were performed during the building of Simulation Model S_2 as for Simulation Model S_1, and no logic errors were found.

The ICC statistics for this model were calculated as a form of black box validation and can be found in Table 5.8. A value of $w = 0.99$ was used for this model as well.

ICC statistic	Totals grouped by	ICC value
ICC(A,1)	size	0.994
	week	0.938
	store	0.944
ICC(C,1)	size	0.994
	week	0.937
	store	0.944
ICC(A,k)	size	0.997
	week	0.968
	store	0.971
ICC(C,k)	size	0.997
	week	0.967
	store	0.971

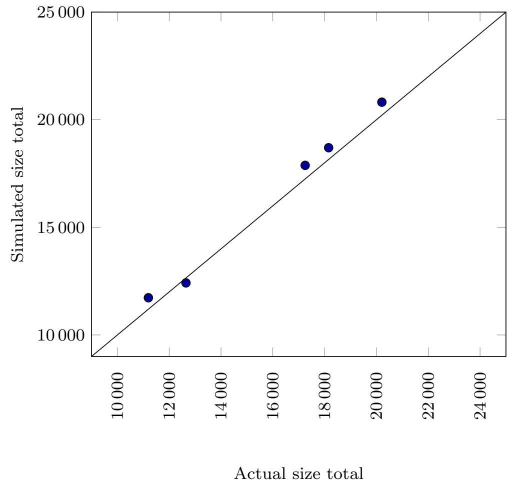
Table 5.8: ICC values for Simulation Model S_2 applied to Subclass A_S. Simulated size totals, week totals and store totals are compared to actual totals. A value of $w = 0.99$ was used.

All ICC values are very high, indicating that the simulated sales are very close to actual sales. Figure 5.13 contains scatter plots of total sales generated by the simulation model against actual values, together with the line $y = x$, indicating the correlation between simulated and actual values.

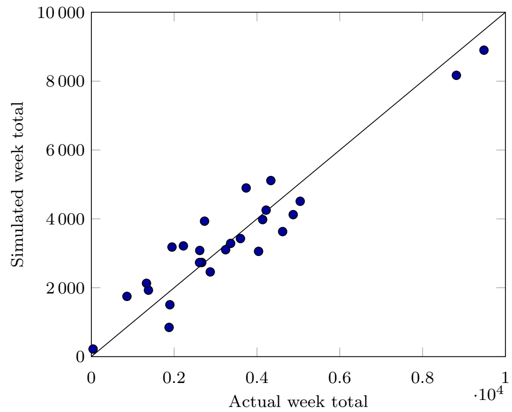
The plots in Figure 5.13(a) and 5.13(b) indicating the correlation of size and week totals, are similar to the plots in Figure 5.8(a) and 5.8(b), the corresponding plots for Simulation Model S_1. Points are again spread evenly around the line $y = x$ in both cases. Outlier 'A' that occurred in

Figure 5.8(c) also occurs in Figure 5.13(c), together with a few other outliers. All the outliers represent stores for which demand was overestimated by the simulation model for the same reason that point A was overestimated: the actual demand proportion for those stores in 2014 was significantly lower than in historical years for all sizes. The same issue occurred for these stores in Simulation Model S_1 as well, but the effect is enhanced in this model because store demand is generated on a size level. The overestimation of demand in these stores occurred for every size, and the average overestimation per size is greater than the overestimation on a company level in Simulation Model S_1.

Again, the ICC measures and correlation plots indicate that the sales predicted by the Simulation Model S_2 are reliable, excepting the few outliers. This concludes the black box validation.



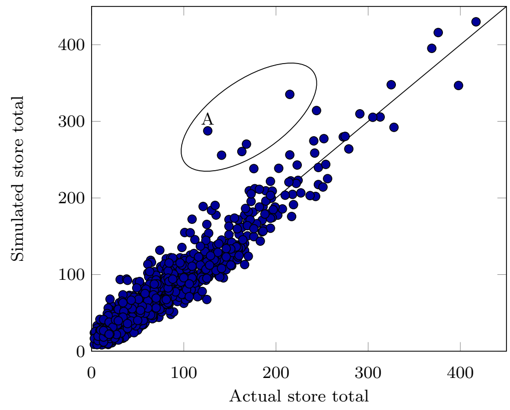
(a) Correlation for size totals



(b) Correlation for week totals.

5.4.5 Graphical results

Figure 5.14 contains a graphical representation of the weekly company level sales generated by one simulation replication of Simulation Model S_2, applied to Subclass A_S, together with actual weekly sales. Actual inflows as obtained from the real data were used in the simulation. This graph is very similar to the one for Simulation Model S_1, and also illustrates that sales are simulated very accurately on a company level.



(c) Correlation for store totals.

Figure 5.13: Scatter plots indicating the correlation between simulated and actual (a) size totals, (b) week totals and (c) store totals, for Subclass A_S. Simulation Model S_2 was used to simulate sales.

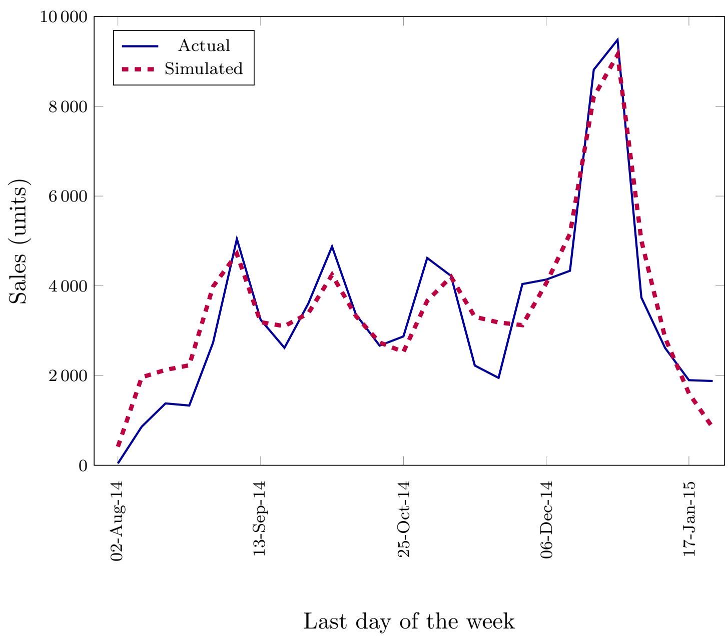


Figure 5.14: Weekly simulated sales for Subclass A_S, compared to actual sales. Simulation Model S_2 was used to simulate sales.

In Figure 5.15, a graphical representation is given of results for one simulation replication of the same simulation model and subclass, for Size 3. The results for other sizes are similar. As for Simulation Model S_1, results are still very accurate, but less accurate than on a company level.

Figure 5.16 contains weekly results for one simulation replication of the same model and subclass for Size 3 on a store level. Simulated and actual sales are plotted for the same three stores as for Simulation Model S_1. Results are very similar to the results for Simulation Model S_1, and it is again observed that simulated sales for larger stores that sell more units per season are more accurate than for small stores.

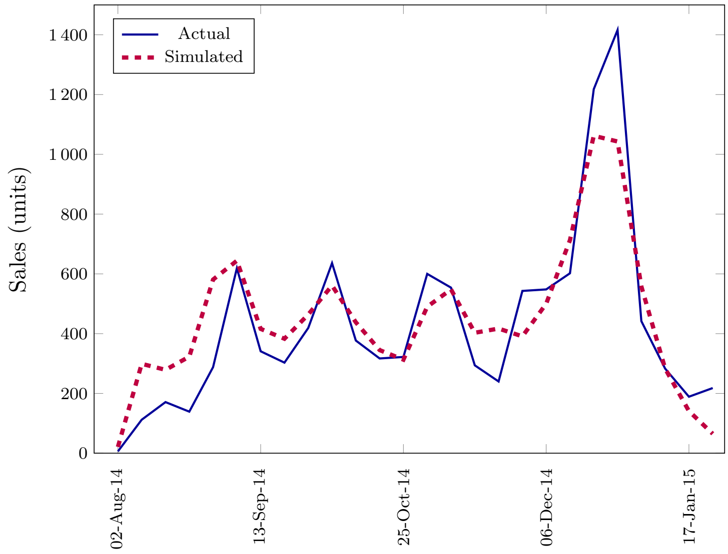
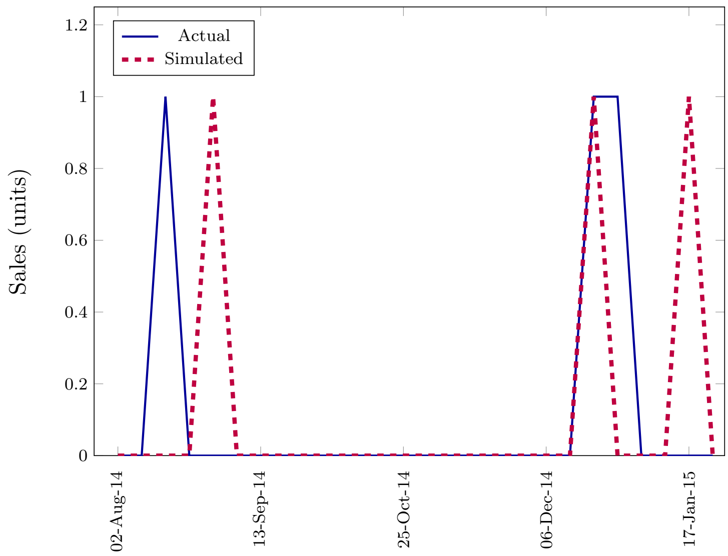


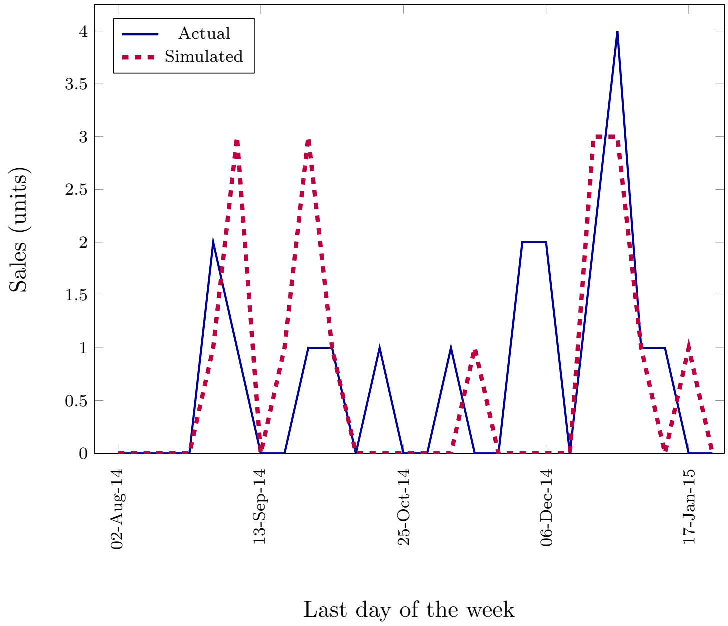
Figure 5.15: Weekly simulated sales in Size 3 for Subclass A_S, compared to actual sales. Simulation Model S_2 was used to simulate sales.



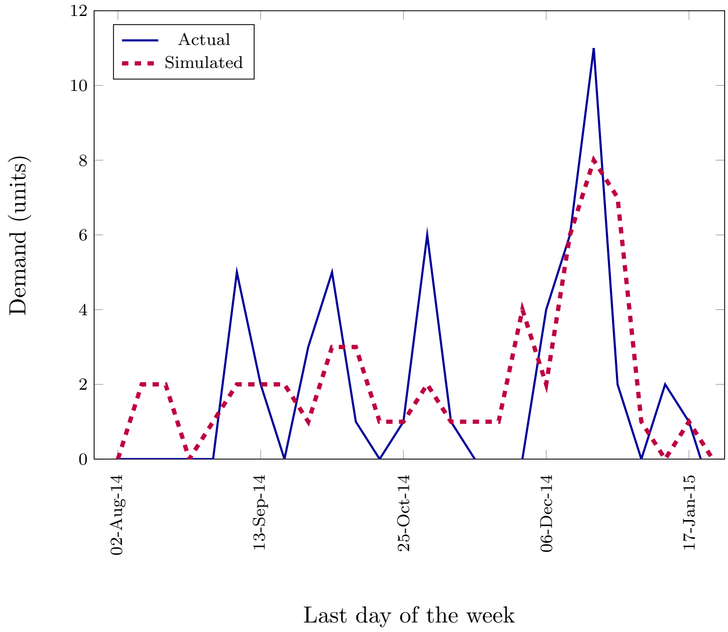
(a) Simulated sales in Size 3 at Simunye, Swaziland, compared to actual sales.

5.5 Comparison of Simulation Model S_1 and Simulation Model S_2

In Table 5.9, Simulation Model S_1 and Simulation Model S_2 are compared to one another with respect to the ICC statistics for the sales generated by the models, as well as total sales. The total simulated sales are very similar for the two models. The actual number of units sold is 88591, so Simulation Model S_1's total is slightly closer to actual sales than Simulation Model S_2's total. The ICC statistics for size and week totals are also very similar for the two models. Simulation Model S_2 performs slightly better than Simulation Model S_1 for week totals, but not significantly. The performance of Simulation Model S_1 for store totals is about 1% better than Simulation



(b) Simulated sales in Size 3 at Aberdeen, compared to actual sales.



(c) Simulated sales in Size 3 at Parow Centre, compared to actual sales.

Figure 5.16: Simulated sales in Size 3 for Subclass A_S by Simulation Model S_2 compared to actual sales at a store in (a) Simunye, Swaziland (b) Aberdeen and (c) Parow Centre.

Model S_2. Both models are very reliable, because their ICC values are very close to 1.

A comparison of the average total weekly sales (on a company level) generated by 10 simulation runs of Simulation Model S_1 and Simulation Model S_2, together with actual sales, can be found in Figure 5.17. The difference between the company level weekly sales is so small that it can hardly be observed in the graph.

Measure	Totals grouped by	Simulation Model S_1	Simulation Model S_2
Total sales		90139.4	90186.9
	size	0.995	0.993
ICC(A,1)	week	0.936	0.940
	store	0.957	0.944
	size	0.995	0.994
ICC(C,1)	week	0.934	0.938
	store	0.957	0.944
	size	0.997	0.997
ICC(A,k)	week	0.967	0.969
	store	0.978	0.971
	size	0.998	0.997
ICC(C,k)	week	0.966	0.968
	store	0.978	0.971

Table 5.9: A comparison of Simulation Model S_1 and Simulation Model S_2.

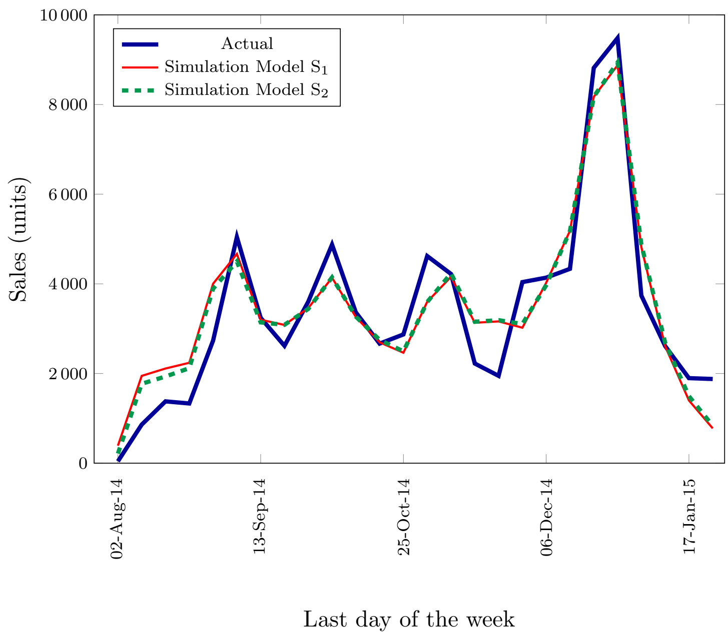


Figure 5.17: Weekly simulated sales for Subclass A_S by Simulation Model S_1 and Simulation Model S_1, together with actual sales.

Figure 5.18 contains a graphical representation of weekly sales on a size level generated by Simulation Model S_1 and Simulation Model S_2, as well as actual sales. The particular results are for Size 3 of Subclass A_S, but the results for other sizes are similar. The difference between the two models is still small, although slightly bigger than on a company level.

It is apparent from the ICC measures as well as the two graphs that sales generated by Simulation Model S_1 and Simulation Model S_2 for this data set are very similar. Because both models are trustworthy and one is not significantly more accurate than the other, experimentation with both models were continued.

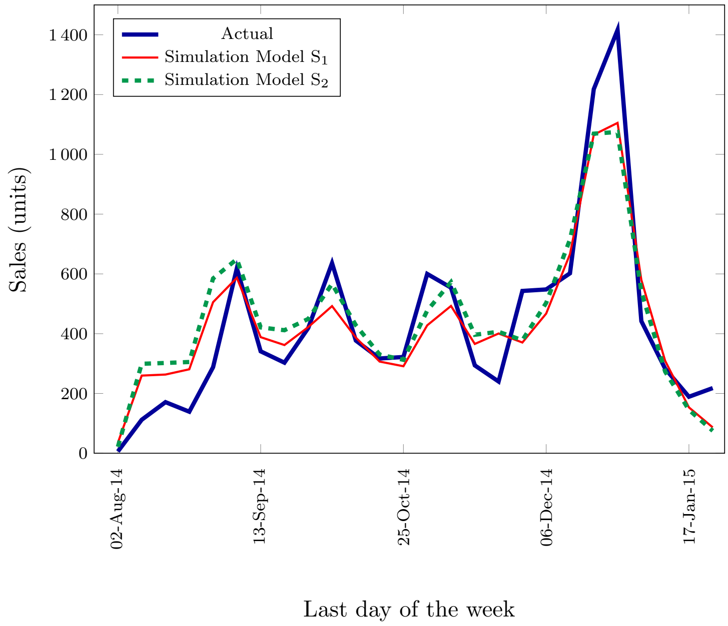


Figure 5.18: Weekly simulated sales in Size 3 for Subclass A_S by Simulation Model S_1 and Simulation Model S_2, together with actual sales.

5.6 Style level model

A third model, where sales were simulated on a style level, was also implemented during experimentation. This model followed the same modelling approach as Simulation Model S_1. However, this model's ICC values for week totals were only about 0.8, and the ICC values for size and store totals were less than or equal to that of Simulation Model S_1 and Simulation Model S_2. The lower ICC values for week totals is a result of the underlying regression models being less accurate (their R^2 values were lower). The results when comparing allocation algorithms to one another using this model did not differ significantly from the results for Simulation Model S_1 and Simulation Model S_2. This confirmed the recommendation of PEP [71] that different styles from the same subclass may be handled as one entity. Therefore, no further experiments for this model were conducted.

CHAPTER 6

Models for Winter subclasses

This chapter involves the development of simulation models for Winter subclasses. Again, two models are developed, following the same modelling approaches as Simulation Model S_1 and S_2. The only difference in the models for Winter subclasses is in the regression, where different explanatory variables are used. The system modelled by Simulation Model W_1 and Simulation Model W_2 have the same operation and elements as the system modelled by Simulation Model S_1 and Simulation Model S_2, which was described in Chapter 5. The same assumptions that were given in §5.2 also apply to this model. The development of Simulation Model W_1 and W_2 is discussed in §6.1 and §6.2, respectively. This is followed by a comparison of the models in §6.3.

6.1 Simulation Model W_1

Simulation Model W_1, like Simulation Model S_1, is based on an underlying multiple regression model which forecasts weekly demand for the entire company. The simulation code for Simulation Model W_1 is the same as the code for Simulation Model S_1 discussed in §5.3.1. Subclass A_W (teenage girls fancy slippers) was used as a training set when building this model.

6.1.1 Development of the regression model

The weekly demand on a company level for Subclass A_W is given in Figure 6.1. A clear seasonal pattern is discerned, which is roughly the same for every year. As in Subclass A_S, a peak occurs near the end of every month. This subclass does not have an exceptionally high peak similar to the peak during Christmas for Subclass A_S, but demand during April is higher than other months, which could be as a result of Easter falling within this time.

Based on the pattern of the data, PEP's input and experiments, the following variables are defined for inclusion in the model. Let $\mathcal{K} = \{1, 2, \ldots, k, \ldots, K\}$ be the set of weeks in a season, and let

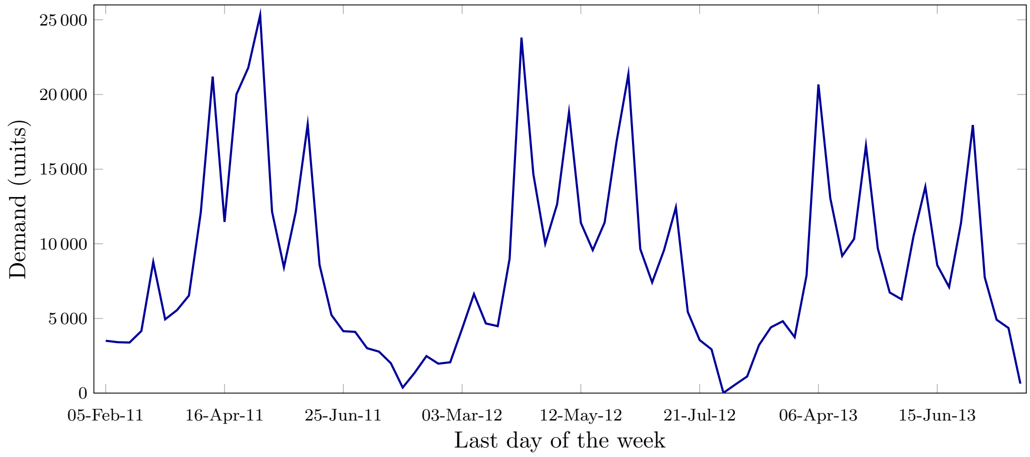


Figure 6.1: Weekly demand on a company level for Subclass $\mathrm{A_W}$, for the years 2011–2013.

Y_k be the total demand in week k

L_k be the total unit inflow during week k

$$E_k = \begin{cases} 1 \text{ if the last day of week } k \text{ is after the } 29^{\text{th}} \text{ or before the } 11^{\text{th}} \text{ of a month} \\ 0 \text{ otherwise,} \end{cases}$$

$$F_k = \begin{cases} 1 \text{ if week } k \text{ falls in February} \\ 0 \text{ otherwise,} \end{cases}$$

$$M_k = \begin{cases} 1 \text{ if week } k \text{ falls in March} \\ 0 \text{ otherwise,} \end{cases}$$

$$A_k = \begin{cases} 1 \text{ if week } k \text{ falls in May} \\ 0 \text{ otherwise,} \end{cases}$$

$$J_k = \begin{cases} 1 \text{ if week } k \text{ falls in June} \\ 0 \text{ otherwise.} \end{cases}$$

$$U_k = \begin{cases} 1 \text{ if week } k \text{ falls in July} \\ 0 \text{ otherwise, and let} \end{cases}$$

$$G_k = \begin{cases} 1 \text{ if week } k \text{ falls in Aug} \\ 0 \text{ otherwise.} \end{cases}$$

Once again, demand in week k, Y_k, is the dependent variable, and $\hat{Y}_k$ is the forecasted value for Y_k. The inclusion of independent variables was finalised in experiments where a significance level of 0.05 was used.

As with Summer products, inflows have an impact on demand, and the variable L_k was again included in the model. The variable E_k as well as the monthly dummy variables were included to handle the seasonal patterns in the data. The variable E_k is defined in the same way as for the Summer data set, and proved to be significant. For the monthly dummy variables, April is used as the reference month, because demand is highest during April.

As for Subclass $\mathrm{A_S}$, the variables Y_{k-1} and L_{k-1} were included to handle the presence of positive autocorrelation in the residuals. After including these lag variables autocorrelation was no longer present.

The number in which a week falls was also included during experiments, because it was a significant variable for the Summer data set, but this variable was not significant for Subclass $\mathrm{A_W}$. An extra dummy variable indicating Easter time was also tried, but it was not a significant explanatory variable of demand. The higher peak during April which could be as a result of Easter is already explained by the monthly dummy variables. The same financial indicators that were tested for Summer products were also tested for Subclass A_W, but once again, these variables were not significant at a significance level of 0.05.

Having established the independent variables, $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta_{10}}$ were estimated by

$$\sqrt{\hat{Y}_k} = \hat{\beta}_0 + \hat{\beta}_1\sqrt{Y_{k-1}} + \hat{\beta}_2 L_k + \hat{\beta}_3 L_{k-1} + \hat{\beta}_4 E_k + \hat{\beta}_5 F_k + \hat{\beta}_6 M_k + \hat{\beta}_7 P_k + \hat{\beta}_8 A_k + \hat{\beta}_9 J_k + \hat{\beta}_{10} U_k. \quad (6.1)$$

The square root sign has been included in the model to handle possible heteroscedasticity, which was an important factor during experimentation with Subclass A_W. Although heteroscedasticity is not present in this data set without the square root sign, the model has to be applicable to other data sets as well. Simulation Model W_1 is mainly based on this data set, but the data set for Subclass B_W was also considered, and without the square root sign, heteroscedasticity is a serious problem.

After estimating $\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta_{10}}$, the resulting regression equation is given by

$$\begin{aligned}\sqrt{\hat{Y}_k} = 39.99 + 0.57\sqrt{Y_{k-1}} + 0.00036L_k + 0.00029L_{k-1} + 24.42E_k - 27.66F_k - 24.06M_k \\ - 10.42A_k - 11.86J_k - 26.46U_k - 76.91G_k. \quad (6.2)\end{aligned}$$

The R^2 value for this regression model is 0.85 and the adjusted R^2 value is 0.82, indicating a good fit. The intercept is included in the model because its p value indicates that it is highly significant in the regression equation.

The regression equation makes intuitive sense. If all variables are 0, the week falls in the month of April. Despite there also being no demand during the previous week and no inflows during this or the previous week, demand is still expected to be positive, because demand during April is generally quite high in historical years, especially 2011.

Similar to Simulation Model S_1, a higher demand during the previous week indicates that demand is likely to be higher during this week. Inflows during this and the previous week also increase demand during this week. Demand is exceptionally high after the end of the month when customers receive salaries and wages; therefore, the high coefficient of E_k is reasonable. The coefficients of the dummy variables indicate that demand during April is indeed highest, as anticipated after studying the graphical display of weekly demand. Demand during May is second highest, followed by June, March, July, February and lastly August. This pattern is clearly observed in Figure 6.1.

The p values of the variables in regression equation (6.2) are given in Table 6.1. Most variables are highly significant. The p value for the variable L_{k-1} is slightly higher than 0.05 and therefore L_{k-1} is not significant at a significance level of 0.05. In this data set, the lag variable $\sqrt{Y_{k-1}}$ alone is sufficient to get rid of autocorrelation; however, L_{k-1} was kept for the sake of applicability to other data sets for which autocorrelation could be a bigger problem. The inclusion or exclusion of the variable does not make much of a difference in this data set.

The p value of A_k is also higher than 0.05, indicating that the demand during May is not significantly different from demand during April; however, it may be significantly different from demand during other months.

The F test was also performed to test the joint explanatory power of the variables. The p value for the F test, obtained from SAS [61], is smaller than 0.0001; therefore, the null hypothesis

Variable	t value	p value
Intercept	3.84	0.0003
$\sqrt{Y_{k-1}}$	7.32	< 0.0001
L_k	2.28	0.026
L_{k-1}	1.96	0.054
E_k	6.61	< 0.0001
F_k	-3.1	0.0028
M_k	-3.58	0.0006
A_k	-1.77	0.081
J_k	-2.06	0.043
U_k	-3.95	0.0002
G_k	-4.91	< 0.0001

Table 6.1: The t and p values for regression (6.2).

that all regression coefficients are 0 is rejected and it is concluded that the model as a whole is significant in explaining demand.

So far, the regression equation is intuitively reasonable and measures indicate a good fit. In the next section, the regression assumptions are formally tested for this model.

6.1.2 Validation and accuracy of the regression model

Firstly, regression equation (6.2) is linear in parameters, as all regression coefficients are constants. Therefore Assumption 1 that the model is linear in the parameters holds.

The modified Breusch-Pagan test was performed to verify that homoscedasticity holds. The p value for the test is given by 0.49, so that the null hypothesis of homoscedasticity is not rejected at a significance level of 0.05. It is therefore reasonable to assume homoscedasticity (Assumption 2).

To test that the residuals are normally distributed, the test statistics for the Shapiro-Wilk, Kolmogorov-Smirnov, Cramér-von Mises and Anderson-Darling tests were calculated for the residuals. The test statistics, together with the accompanying p values, are given in Table 6.2. The last column indicates whether H_0 is rejected at a significance level of 0.05.

Test	Statistic	p value	Reject H_0?
Shapiro-Wilk	$W = 0.983$	0.403	No
Kolmogorov-Smirnov	$D = 0.0457$	> 0.150	No
Cramér-von Mises	$W^2 = 0.0244$	> 0.250	No
Anderson-Darling	$A^2 = 0.238$	> 0.250	No

Table 6.2: Results for normality tests on the residuals of regression (6.2). A significance level of $\alpha = 0.05$ is used.

In all cases, the p values are higher than 0.05 and the null hypothesis of normality is not rejected. The reported mean value of the residuals is 0 and the standard deviation is 13.39. Therefore, Assumption 3 holds.

The Durbin-Watson value for the regression is given by 2.01. The Durbin-Watson upper and lower bounds for 78 observations and 10 degrees of freedom are 1.39 and 1.9. As 2.01 falls in the interval $(1.9, 4 - 1.9) = (1.9, 2.1)$, the null hypotheses of no positive or negative autocorrelation are not rejected. The p value for the Runs test is 0.65, so that the null hypothesis of no autocorrelation is not rejected at a significance level of 0.05. This confirms that no autocorrelation

	$\sqrt{Y_{k-1}}$	L_k	L_{k-1}	E_k	F_k	M_k	A_k	J_k	U_k	G_k
$\sqrt{Y_{k-1}}$	1.00	−0.12	0.13	−0.16	−0.61	−0.29	0.33	0.24	−0.02	−0.08
L_k	−0.12	1.00	0.12	0.13	0.09	0.27	0.08	−0.10	−0.30	−0.10
L_{k-1}	0.13	0.12	1.00	−0.02	−0.06	0.17	0.05	−0.04	−0.26	−0.10
E_k	−0.16	0.13	−0.02	1.00	−0.08	0.09	−0.03	0.00	−0.03	0.18
F_k	−0.61	0.09	−0.06	−0.08	1.00	−0.19	−0.17	−0.18	−0.17	−0.06
M_k	−0.29	0.27	0.17	0.09	−0.19	1.00	−0.21	−0.22	−0.21	−0.07
A_k	0.33	0.08	0.05	−0.03	−0.17	−0.21	1.00	−0.20	−0.20	−0.06
J_k	0.24	−0.10	−0.04	0.00	−0.18	−0.22	−0.20	1.00	−0.20	−0.06
U_k	−0.02	−0.30	−0.26	−0.03	−0.17	−0.21	−0.20	−0.20	1.00	−0.06
G_k	−0.08	−0.10	−0.10	0.18	−0.06	−0.07	−0.06	−0.06	−0.06	1.00

Table 6.3: Pairwise correlation coefficients between independent variables in regression (6.2).

in the error terms may be assumed.

Assumption 5 states that there is no multicollinearity between independent variables. One indication of multicollinearity is if the signs of the coefficients in the regression equation do not make sense. It has previously been established that the regression equation does make intuitive sense. High R^2 and adjusted R^2 values, together with a highly significant F value but few significant explanatory variables is another indication of multicollinearity. This is not the case, as most of the explanatory variables are significant at a significance level of 0.05.

Lastly, the pairwise correlation coefficients between variables are inspected to detect the possible presence of multicollinearity. These values can be found in Table 6.3. As no correlation coefficient between two different variables is higher than 0.8 or lower than −0.8, there is no indication that multicollinearity plays a significant role. The assumption of no multicollinearity is therefore reasonable.

A graphical representation of the fit and forecast of regression (6.2) can be found in Figure 6.2. The fit is very accurate. Although the largest part of the forecast overestimates actual demand, the pattern of the forecast is relatively accurate, except that the peak at the beginning of July is severely underestimated in 2014. This is probably because the corresponding peak is exceptionally low in 2011, which significantly lowered the coefficient for the July dummy variable. A possible reason for the large overestimation in actual demand is that the average demand in 2014 was lower than during other years. The effect is enhanced by the fact that $\hat{Y}_{k-1}$ is used as a proxy for Y_{k-1} because Y_{k-1} is not known in advance. Total demand in 2014 is overestimated by 27%. However, it is more important that the pattern of demand is estimated accurately rather than that demand is estimated on the right level. The comparison of total sales generated by different allocation models will not be affected by the initial overestimation of demand.

6.1.3 The number of simulation replications

The number of simulation replications that should be performed was determined by formula (5.4). Because Simulation Model W_1 is the same as Simulation Model S_1 except for the regression equation, it is safe to assume that total sales generated by this model is also normally distributed and that the formula may be used.

After 10 initial replications of the model, a mean value of $\bar{X}(m) = 15\,4547.2$ and a standard deviation of $S(m) = 168.93$ were obtained. The allowable percentage deviation was chosen as 0.01 and a significance level of 0.05 was used. The critical t value $t_{9,0.0975} = 2.262$. Substituting these values in equation (5.4) yields $N(m) = 0.06 \approx 1$ simulation run. It was decided to perform ten runs for more precision. For ten replications a precision of 0.0008 (about 125 units) is

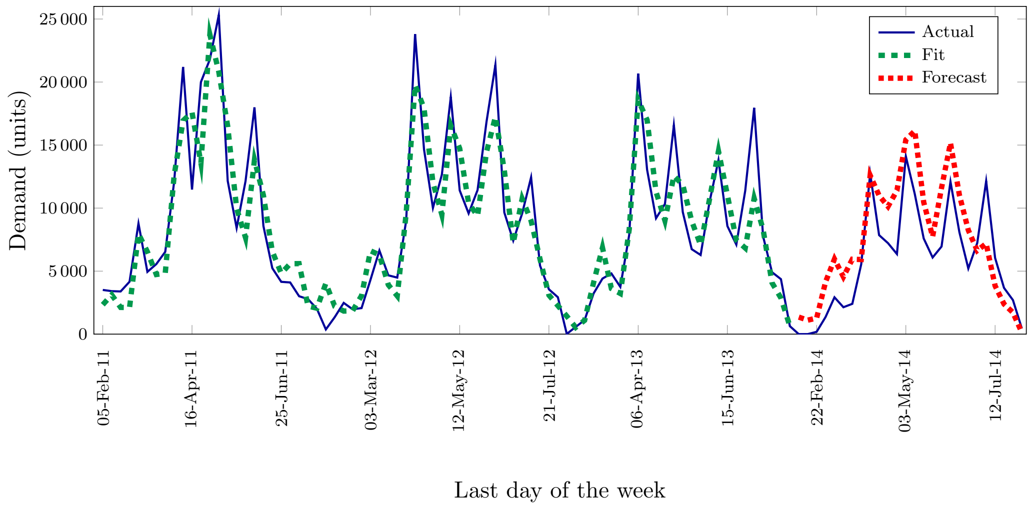


Figure 6.2: Graphical display of the fit and forecast of regression (6.2) in years 2011–2014.

obtained.

6.1.4 Validation and verification of the simulation model

The first three steps in the VV process were not repeated for this model, as Simulation Model W_1 is identical to Simulation Model S_1 except for the regression model, which has already been validated.

The ICC statistics for this model were calculated as a form of black box validation. The statistics can be found in Table 6.4.

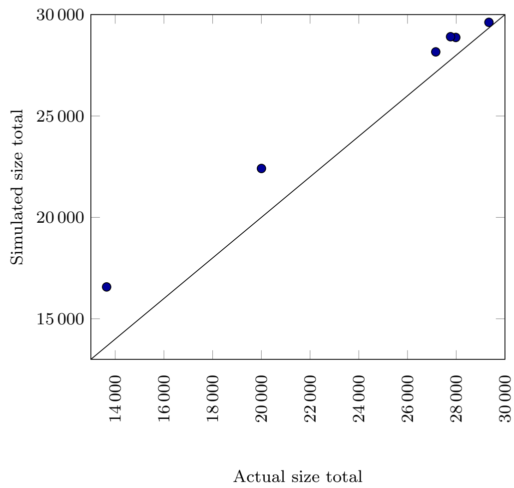
ICC statistic	Totals grouped by	ICC value
	size	0.956
ICC(A,1)	week	0.827
	store	0.993
	size	0.984
ICC(C,1)	week	0.824
	store	0.994
	size	0.977
ICC(A,k)	week	0.905
	store	0.997
	size	0.992
ICC(C,k)	week	0.903
	store	0.997

Table 6.4: ICC values for Simulation Model W_1 applied to Subclass A_W. Simulated size totals, week totals and store totals were compared to actual totals. A value of $w = 0.99$ was used.

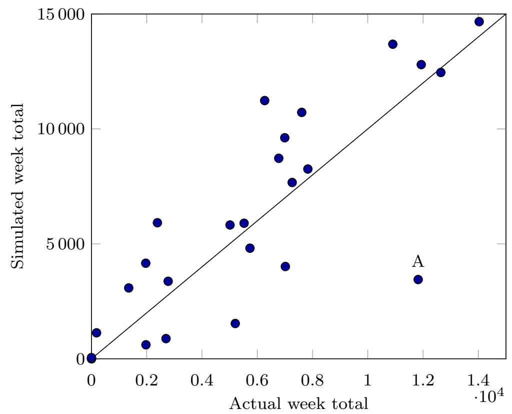
All ICC values are above 0.9, except for the single measures of the week totals. This could be expected, because the R^2 values of the regression equation are lower than for Subclass A_S and they are below 0.9. However, the ICC values are still above 0.8 and it can be concluded that the model is an accurate representation of reality.

The total simulated sales is 154 546.3, which is about 6% above the actual total sales (145 904). Sales are overestimated because demand was overestimated by the regression model. However, stockouts generated by the simulation model to some extent corrected the large overestimation, because there was not enough stock available to satisfy the high demand generated by the model.

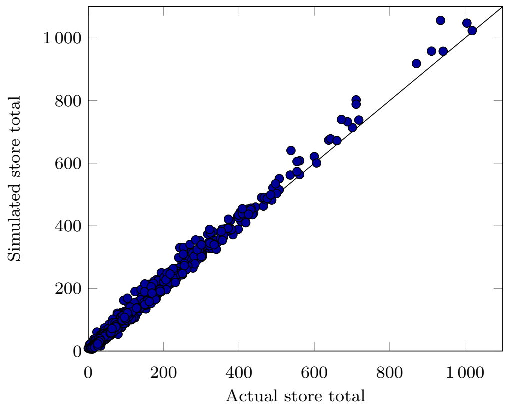
Scatter plots of simulated sales totals against actual sales totals on a size, week and store level, are given in Figure 6.3. In each case, the line $y = x$ is supplied to indicate the correlation between simulated and actual sales. The data points in Figure 6.3(a), representing the correlation between actual and simulated size totals, are close to the line $y = x$, but they are all above the line because each size was overestimated due to the overestimation of demand by the regression model. The data points in Figure 6.3(b), representing the correlation between actual and simulated week totals, are spread relatively evenly about the line $y = x$, although more points are above the line than below the line. This is once again due to the overestimation of demand by the regression model. The point labelled 'A' is disproportionally low relative to the other points. This point represents sales during the first week in July, for which demand was severely underestimated by the regression equation, which in turn lead to an underestimation in sales by the simulation model. The data points in Figure 6.3(c), representing the correlation between actual and simulated store totals, are very close to the line $y = x$, but most of the points are overestimated as in the other two plots. Most of the points are in the bottom left-hand section of the graph, because most stores sold less than 600 units throughout the season.



(a) Correlation for size totals



(b) Correlation for week totals.



(c) Correlation for store totals.

Figure 6.3: Scatter plots indicating the correlation between simulated and actual (a) size totals, (b) week totals and (c) store totals, for Subclass A_W. Simulation Model W_1 was used to simulate sales. The point labelled 'A' is an outlier.

6.1.5 Graphical results

A graphical representation of weekly sales on a company level generated by one simulation replication, compared to actual sales, is given in Figure 5.9. The sales pattern is similar to the regression forecast, although the overestimation is smaller, as already discussed. The graph confirms the high correlation between actual and simulated sales indicated by the ICC values in Table 6.4. The large underestimation during July, which is due to the underestimation in demand by the regression model, explains the lower ICC statistics on a week level.

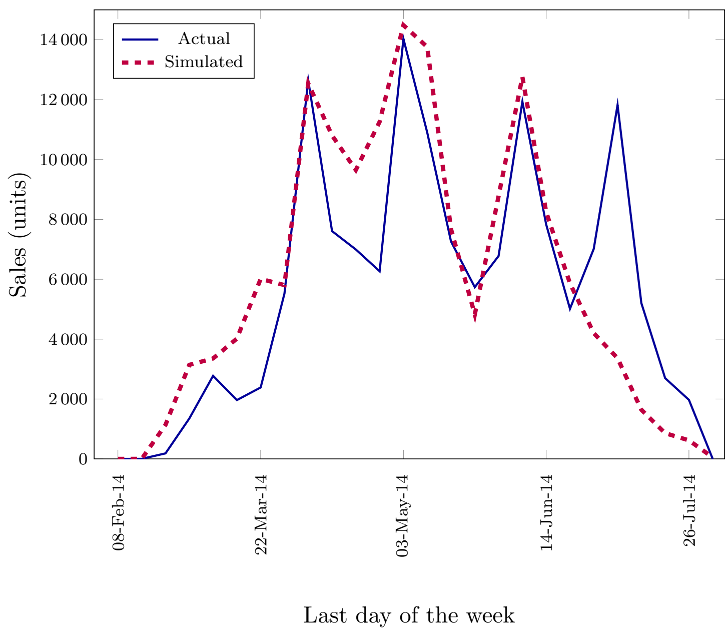


Figure 6.4: Weekly simulated sales for Subclass A_W by Simulation Model W_1, compared to actual sales.

Figure 6.5 contains the weekly sales for Size 7 as generated by one simulation replication of Simulation Model W_1, compared to actual sales. As with Simulation Model S_1, results on a size level are similar but less accurate than on a company level. Results for other sizes are similar, although some overestimations are slightly bigger and others slightly smaller.

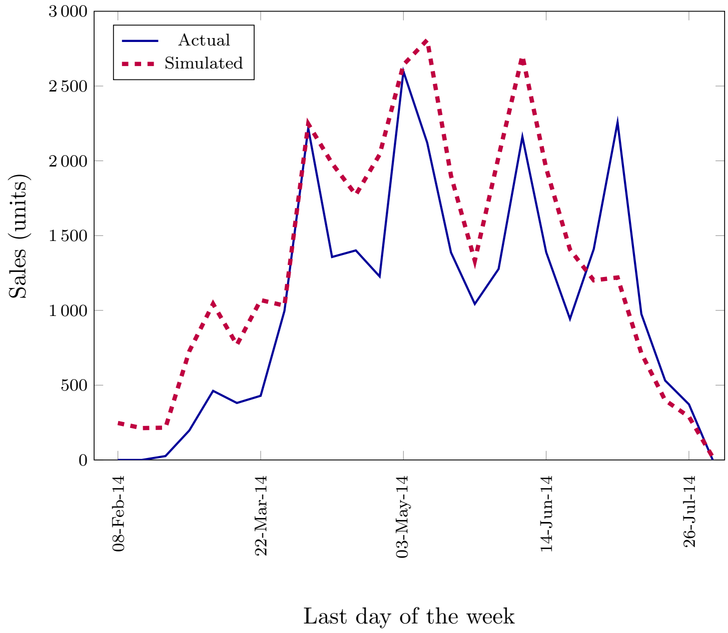
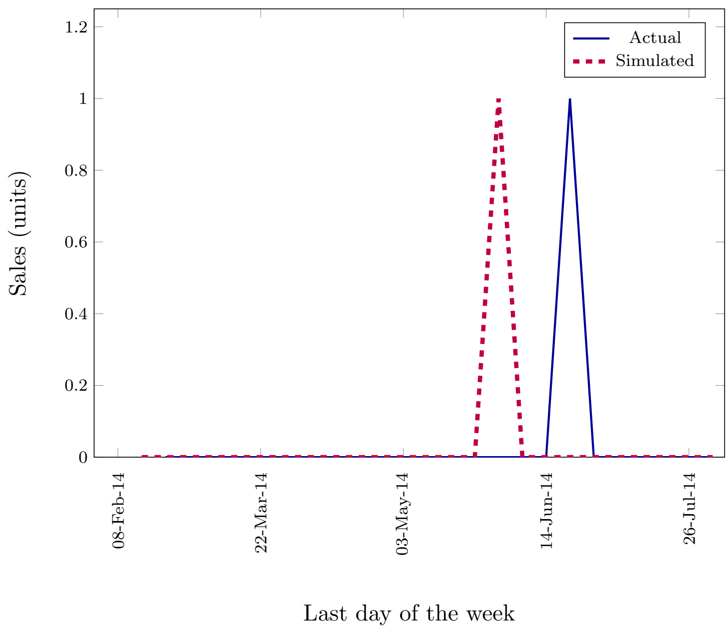
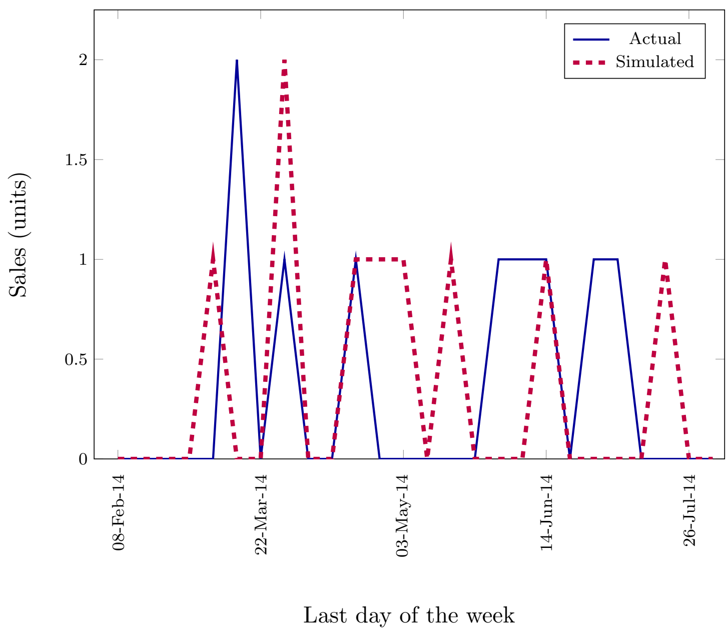


Figure 6.5: Weekly simulated sales in Size 7 for Subclass A_W, compared to actual sales. Simulation Model W_1 was used to simulate sales.

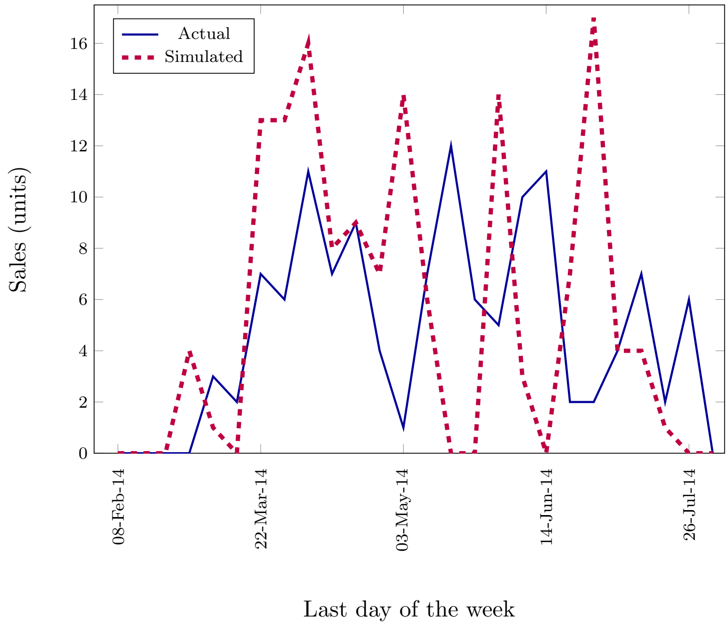
Store level results in Size 3 for one simulation replication can be found in Figure 6.6. Figure 6.6(a) contains the results for a store in Simunye, which in reality sold only one Size 3 unit during the season. Similar to Simulation Model S_1, the model is not able to predict the time at which sales took place, but it correctly predicted only one sale during the season. Figure 6.6(a) shows the results for a store in Aberdeen, which can be regarded as a medium store with regards to number of Size 3 sales during a season. As before, results are slightly more accurate for a medium store. In Figure 6.6(c), results for a store in Musina are given. Although this store sold the most Size 3 units for the subclass, the simulation is not really more accurate than for the store in Aberdeen. This implies that the simulation for this subclass is less accurate in predicting the time of year at which store level sales take place.



(a) Simulated sales in Size 3 at Simunye, Swaziland, compared to actual sales.



(b) Simulated sales in Size 3 at Aberdeen, compared to actual sales.



(c) Simulated sales in Size 3 at Musina, compared to actual sales.

Figure 6.6: Simulated sales in Size 3 for Subclass A_W compared to actual sales at a store in (a) Simunye, (b) Aberdeen and (c) Musina. Simulation Model W_1 was used to simulate sales.

6.1.6 The value of w

A value of $w = 0.99$ was used during all experiments, because the ratio $\hat{\beta}_1 : \hat{\beta}_2 = 0.57 : 0.00036 \approx 0.999 : 0.001$, which is similar to the Summer dataset. Simulation Model W_1 is the same as Simulation Model S_1, except that a different regression equation was used; therefore, the same value of w was used for this data set. Again, the most accurate value of w can be determined by further experiments.

6.2 Simulation Model W_2

In this section, Simulation Model W_2 is developed. Similar to Simulation Model S_2, Simulation Model W_2 is an adaptation of Simulation Model W_1. Simulation Model W_2 uses the same code as Simulation Model S_2 which was described in §5.4.1. Graphical representations of the simulations as compared to actual sales for this model are almost identical to that of Simulation Model W_1 and are thus not shown or discussed.

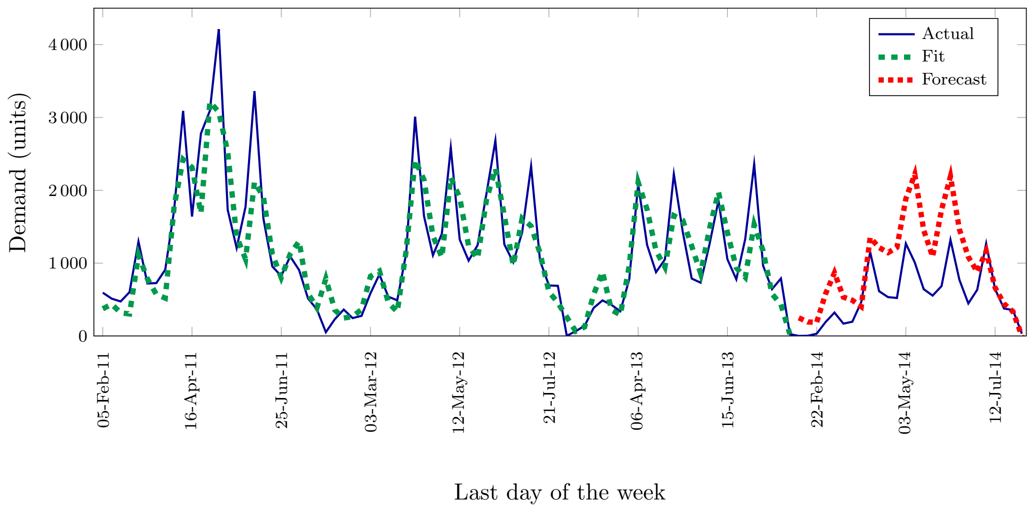
6.2.1 Regression models

As with Simulation Model S_2, a different regression model is implemented for each size, and Simulation Model W_2 uses the same explanatory variables as Simulation Model W_1, but adapted to a size level. All variables for the regression models of Simulation Model W_2 are already defined. As before, the variables Y_{sk} are the dependant variables and $\hat{Y}_{sk}$ are the forecasted values for Y_{sk} for all sizes s in the set $\mathcal{S}$. Subclass A_W contains six sizes ranging from Size 3 to Size 8; therefore, six regression equations were determined. The equations are given by

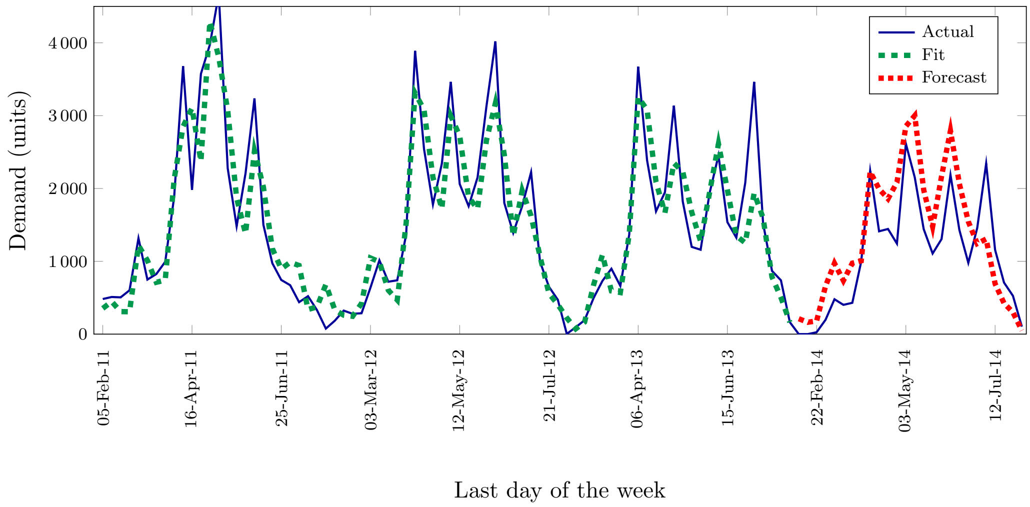
$$\begin{aligned}
\sqrt{Y_{3,k}} &= 13.96 + 0.57\sqrt{Y_{3,k-1}} + 0.00066L_{3,k} + 0.00038L_{3,k-1} + 10.83E_k - 9.01F_k - 8.96M_k \\
&\quad - 2.51A_k - 3.13J_k - 7.51U_k - 35.77G_k, && (6.3) \\
\sqrt{Y_{4,k}} &= 15.58 + 0.57\sqrt{Y_{4,k-1}} + 0.00077L_{4,k} + 0.00055L_{4,k-1} + 10.83E_k - 10.09F_k - 9.28M_k \\
&\quad - 3.52A_k - 4.38J_k - 9.58U_k - 36.22G_k, && (6.4) \\
\sqrt{Y_{5,k}} &= 16.38 + 0.57\sqrt{Y_{5,k-1}} + 0.0010L_{5,k} + 0.00086L_{5,k-1} + 9.82E_k - 10.83F_k - 9.13M_k \\
&\quad - 4.84A_k - 5.26J_k - 11.52U_k - 30.51G_k, && (6.5) \\
\sqrt{Y_{6,k}} &= 16.67 + 0.58\sqrt{Y_{6,k-1}} + 0.00096L_{6,k} + 0.00077L_{6,k-1} + 9.75E_k - 11.78F_k - 9.88M_k \\
&\quad - 4.86A_k - 5.38J_k - 11.71U_k - 30.17G_k, && (6.6) \\
\sqrt{Y_{7,k}} &= 16.93 + 0.58\sqrt{Y_{7,k-1}} + 0.00080L_{7,k} + 0.00074L_{7,k-1} + 10.21E_k - 12.59F_k - 10.88M_k \\
&\quad - 4.29A_k - 5.09J_k - 11.56U_k - 30.08G_k, \text{ and} && (6.7) \\
\sqrt{Y_{8,k}} &= 16.58 + 0.59\sqrt{Y_{8,k-1}} + 0.0011L_{8,k} + 0.00095L_{8,k-1} + 8.55E_k - 12.08F_k - 10.11M_k \\
&\quad - 5.24A_k - 5.46J_k - 12.30U_k - 27.26G_k. && (6.8)
\end{aligned}$$

The signs and sizes of coefficients make intuitive sense for the same reason as regression equation (6.2) of Simulation Model W_1 for the same subclass. All R^2 values are between 0.8 and 0.88, and all adjusted R^2 values are between 0.77 and 0.86, indicating a good fit. Assumption 7 in §5.2 implies that these models are valid because regression model (6.2) is valid. It is therefore not necessary to validate regression assumptions again.

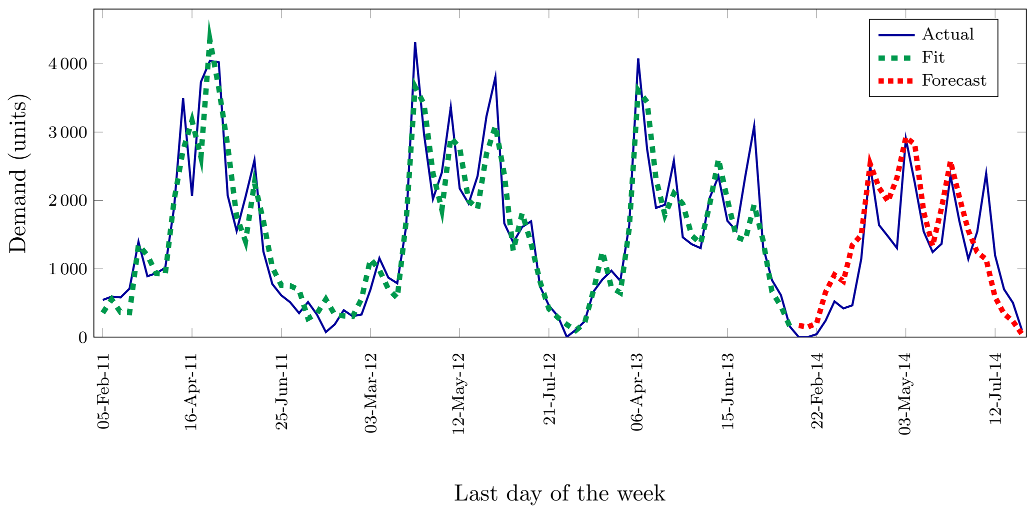
Figure 6.7 contains the regression fit and forecast for Size 3, 7 and 8 of Subclass A_W. The patterns for all sizes are similar to the company level regression of Simulation Model W_1, but the size of the overestimation differ. The overestimation of the total sales in 2014 in Size 3 is 76%, the overestimation for Size 7 is 23% and the overestimation for Size 8 is 15%. On average over all sizes, total sales are overestimated by 28%.



(a) Graphical display of the fit and forecast of regression (6.3) (Size 3) in years 2011–2014.



(b) Graphical display of the fit and forecast of regression (6.7) (Size 7) in years 2011–2014.



(c) Graphical display of the fit and forecast of regression (6.8) (Size 8) in years 2011-2014.

Figure 6.7: Graphical display of the fit and forecast for (a) Size 3, (b) Size 7 and (c) Size 8 of Subclass A_W, in years 2011–2014.

6.2.2 The number of simulation replications

The number of simulation replications that should be performed were determined by formula (5.4). Because Simulation Model W_2 is the same as Simulation Model S_2 except for the regression equations, it may be assumed that total sales generated by this model is also normally distributed and that the formula may be used.

After 10 initial replications of the model, a mean value of $\bar{X}(m) = 155\,161.2$ and a standard deviation of $S(m) = 103.68$ were obtained. The allowable percentage deviation was chosen as 0.01 and a significance level of 0.05 was used. The critical t value $t_{9,0.0975} = 2.262$. Substituting these values in formula (5.4) yields $N(m) = 0.023 \approx 1$ simulation run. It was decided to perform ten runs for more precision. For ten replications a precision of 0.0005 (about 80 units) is obtained.

6.2.3 Validation and verification of the simulation model

The first three steps in the VV process were not repeated for this model, as Simulation Model W_2 is identical to Simulation Model S_2 except for the regression models, which have already been validated. It is assumed that Simulation Model W_2 is valid because Simulation Model S_2 is valid.

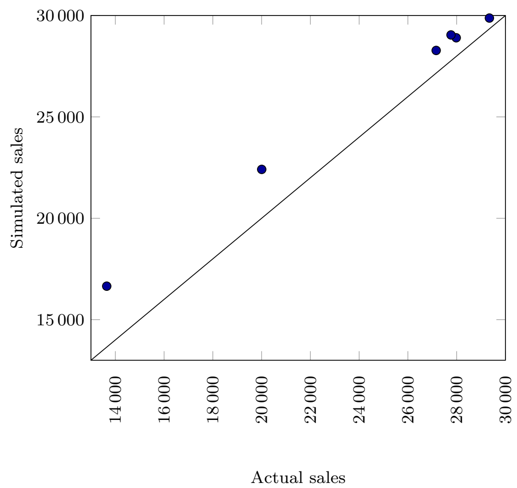
The ICC statistics for this model were calculated as a form of black box validation. The statistics are given in Table 6.4.

ICC statistic	Totals grouped by	ICC value
ICC(A,1)	size	0.954
	week	0.826
	store	0.992
ICC(C,1)	size	0.986
	week	0.823
	store	0.994
ICC(A,k)	size	0.976
	week	0.905
	store	0.996
ICC(C,k)	size	0.993
	week	0.903
	store	0.997

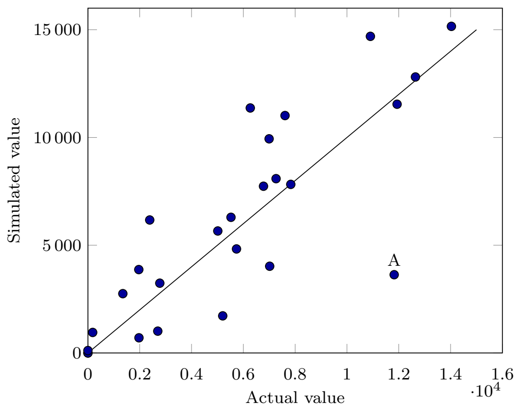
Table 6.5: ICC values for Simulation Model W_2 applied to Subclass A_W. Simulated size totals, week totals and store totals were compared to actual totals. A value of $w = 0.99$ was used.

As for Simulation Model W_1, all ICC values are above 0.9, except for the single measures of the week totals. Again, this is as expected because the R^2 and adjusted R^2 values are around 0.8. However, a correlation of more than 0.8 is sufficiently high to conclude that the model is an accurate representation of reality. The total simulated sales is 155 161.2, which is also about 6% above the actual total sales (145 904).

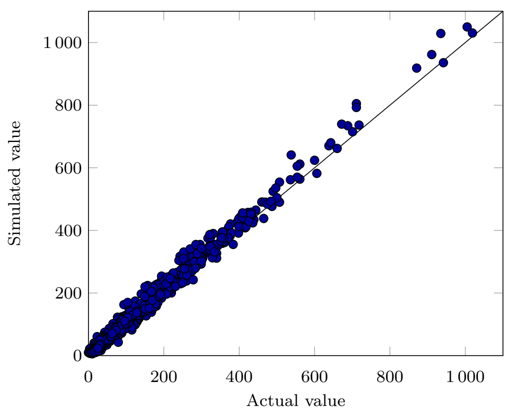
Scatter plots of the correlation between simulated and actual size totals, week totals and store totals, together with the line $y = x$, are given in Figure 6.3. The plots are almost identical to the corresponding plots for Simulation Model W_1. The same outlier occurs in Figure 6.8(b) for the same reason. No other outliers are detected.



(a) Correlation for size totals



(b) Correlation for week totals.



(a) Correlation for store totals.

Figure 6.9: Scatter plots indicating the correlation between simulated and actual (a) size totals, (b) week totals and (c) store totals, for Subclass A_W. Simulation Model W_2 was used to simulate sales.

6.3 Comparison of Simulation Model W_1 and Simulation Model W_2

In Table 6.6, Simulation Model W_1 and Simulation Model W_2 are compared to one another with respect to ICC statistics for the sales generated by the models, as well as total sales. Total sales generated by Simulation Model W_1 is a little closer to actual sales (145904) than sales generated by Simulation Model W_2. Both models are very reliable, because their ICC valus are very close to 1. The ICC statistics for the two models are almost identical, although in most cases the ICC statistics for Simulation Model W_1 are slightly higher than for Simulation Model W_2.

Measure	Totals grouped by	Simulation Model W_1	Simulation Model W_2
Total sales		154546.3	155161.2
	size	0.956	0.954
ICC(A,1)	week	0.827	0.826
	store	0.993	0.992
	size	0.984	0.986
ICC(C,1)	week	0.824	0.823
	store	0.994	0.994
	size	0.977	0.976
ICC(A,k)	week	0.905	0.905
	store	0.997	0.996
	size	0.992	0.993
ICC(C,k)	week	0.903	0.903
	store	0.997	0.997

Table 6.6: A comparison of Simulation Model W_1 and Simulation Model W_2.

A comparison of the average total weekly sales (on a company level) generated by 10 simulation runs of Simulation Model W_1 and Simulation Model W_2, together with actual sales, is given in Figure 6.10. Similar to Simulation Model S_1 and Simulation Model S_2, the company level weekly sales for Subclass A_W generated by Simulation Model W_1 and Simulation Model W_2 are very close together.

Figure 6.11 contains a graphical representation of weekly sales on a size level generated by Simulation Model W_1 and Simulation Model W_2, as well as actual sales. The particular results are for Size 5 of Subclass A_W. Similar to Simulation Model S_1 and Simulation Model S_2, the difference between the two models is slightly bigger than on a company level.

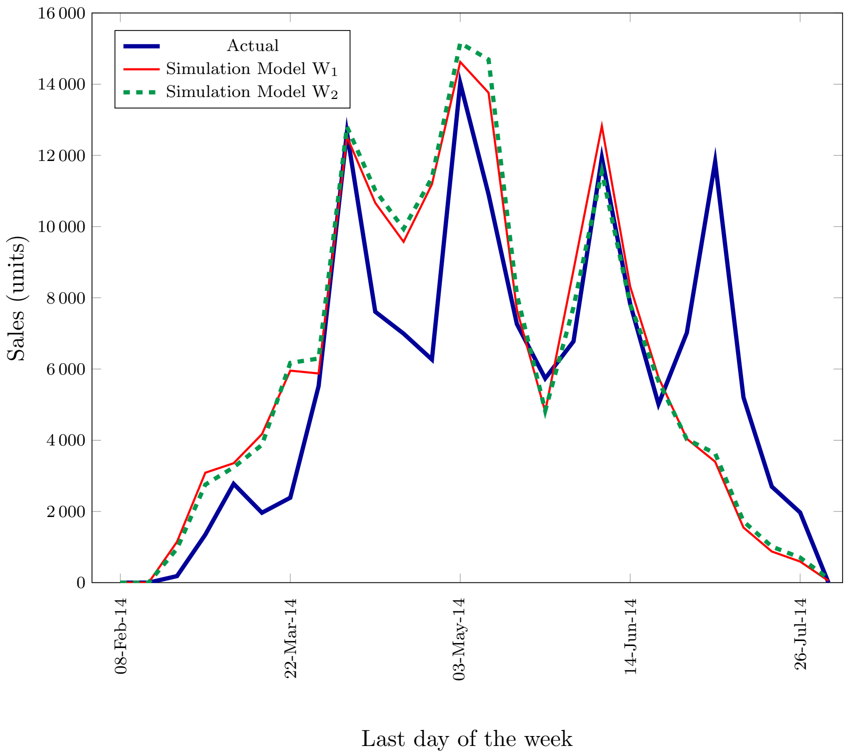


Figure 6.10: Weekly simulated sales for Subclass A_W by Simulation Model W_1 and Simulation Model W_2, compared to actual sales.

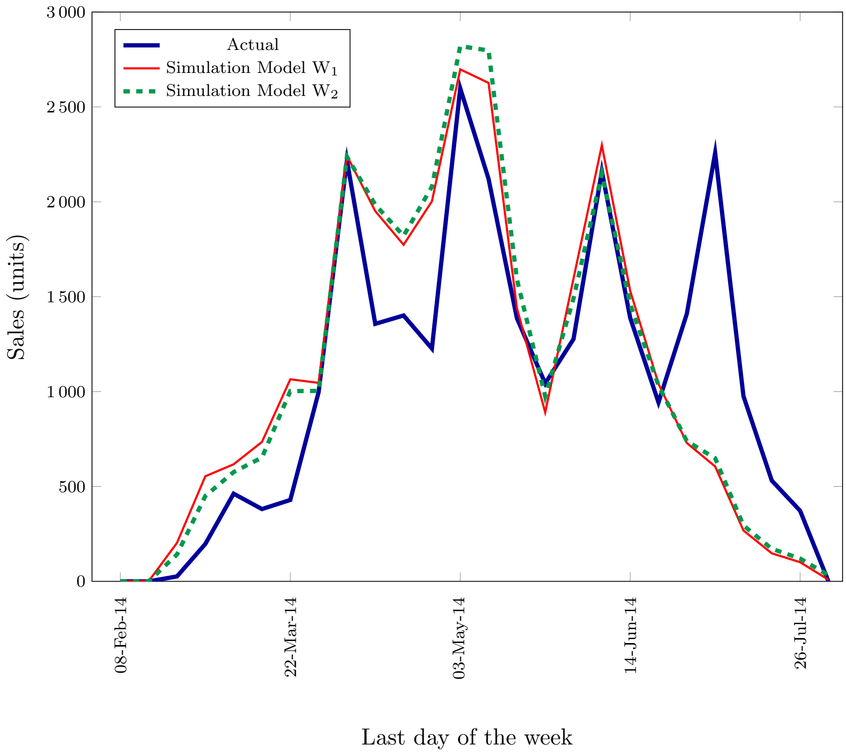


Figure 6.11: Weekly simulated sales in Size 5 for Subclass A_W by Simulation Model W_1 and Simulation Model W_2, compared to actual sales.

CHAPTER 7

Results

In this chapter, simulation models that were developed during previous chapters are applied to new data sets to compare allocation algorithms to one another. Algorithm 1 was implemented by PEP for all data sets described in Chapter 3, and the allocation adjustment solutions for all styles were supplied. The other two algorithms were implemented in CPLEX 12.5 [31]. For Algorithm 2, the value of α was varied for values larger than 0.5 (because the minimisation of shortages is more important than the minimisation of surpluses), and it was found that a value of 0.6 delivered reasonably good objective function values [69]. The determination of the most appropriate value for α may be achieved by further experimentation and is not part of the scope of this thesis. The value of α may also be adjusted by PEP.

7.1 Models for summer products

The simulation models that were developed with Subclass $\mathrm{A_S}$ as a training set were applied to Subclass $\mathrm{B_S}$. In this section, Simulation Model $\mathrm{S_1}$ and Simulation Model $\mathrm{S_2}$ are validated for the new data. In §7.1.1–7.1.2, the regression models are validated, and in §7.1.3, the final simulation results are validated against actual values when using actual 2014 inflows.

7.1.1 Regression for Simulation Model $\mathrm{S_1}$ for Subclass $\mathrm{B_S}$

The regression equation of Simulation Model $\mathrm{S_1}$ when applied to Subclass $\mathrm{B_S}$ is given by

$$\sqrt{Y_k} = 0.86\sqrt{Y_{k-1}} + 4.06E_k + 0.03W_k + 8.32C_k + 0.00072L_k + 0.00072L_{k-1}. \tag{7.1}$$

As with Subclass $\mathrm{A_S}$, the signs of all regression coefficients are positive, and the size of the coefficients relative to one another are similar. The equation therefore makes intuitive sense.

For each variable, a t test was performed to test its significance in explaining demand. The t test values for regression (7.1), together with their accompanying p values, are given in Table 7.1. The p value of W_k indicates that W_k is not a significant explanatory variable of demand for Subclass $\mathrm{B_S}$. As the model was developed using a different subclass, it is expected that some variables may not be significant. However, all other variables are highly significant, so as far as the explanatory power of variables is concerned, the model provides a good fit for this data set.

Variable	t value	p value
$\sqrt{Y_{k-1}}$	33.74	< 0.0001
L_k	5.06	< 0.0001
L_{k-1}	5.26	< 0.0001
E_k	4.38	< 0.0001
W_k	1.15	0.2549
C_k	5.02	< 0.0001

Table 7.1: The t test and the accompanying p values for the independent variables in regression (7.1).

The p value of the F test, obtained by SAS [61], is smaller than 0.0001, which implies that the joint explanatory power of the variables is highly significant. The value of $R^2 = 0.9866$ and the value of the adjusted $R^2 = 0.9855$, indicating a very good fit.

The coefficients are constants, so Assumption 1 of multiple linear regression, namely that the regression equation is linear in its parameters, holds. The Breusch-Pagan test was performed to test whether the homoscedasticity assumption of residuals (Assumption 2) holds. The p value reported by SAS [61] is 0.0776. Therefore the null hypothesis of homoscedasticity is not rejected at a significance level of 0.05, and homoscedasticity may be assumed.

Four normality tests were performed on residuals to test whether residuals are normally distributed with a mean value of 0. The test statistics and p values for the Shapiro-Wilk, Kolmogorov-Smirnov, Cramér von Mises and Anderson-Darling tests are given in Table 7.2. The last column indicates whether the null hypothesis of normality is rejected at a significance level of 0.05.

Test	Statistic	p value	Reject H_0?
Shapiro-Wilk	$W = 0.98$	0.29	No
Kolmogorov-Smirnov	$D = 0.072$	> 0.15	No
Cramér von Mises	$W^2 = 0.039$	> 0.25	No
Anderson-Darling	$A^2 = 0.27$	> 0.25	No

Table 7.2: Results for normality tests on the residuals of regression (7.1). A significance level of $\alpha = 0.05$ is used.

As all p values are higher than 0.05, all tests indicate that normality of residuals may be assumed. The reported mean and standard deviation are -0.05 and 3.38, respectively. As the mean is close to 0, Assumption 3 of multiple linear regression holds.

The Durbin-Watson test was conducted to test for autocorrelation in error terms. The test statistic, obtained from SAS [61], is 1.911. The Durbin-Watson upper and lower bounds for 78 observations and 6 degrees of freedom are 1.50 and 1.77. As 1.911 falls in the interval $(1.77, 4-1.77) = (1.77, 2.23)$, the null hypotheses of no positive and no negative autocorrelation are not rejected. The p value for the Runs test is 0.18, indicating that the null hypothesis of randomness in the error terms is not rejected at a significance level of 0.05. This confirms that no autocorrelation in the error terms may be assumed.

Finally, the assumption of multicollinearity (Assumption 5) is tested. There is no signal of multicollinearity in the information recorded so far. The signs and size of the regression coefficients are reasonable, and all but one of the p values are significant.

The pairwise correlation coefficients between variables are reported in Table 7.3. The absolute values of all correlation coefficients between different variables are lower than 0.8. Therefore it may be assumed that multicollinearity does not play a significant role.

	$\sqrt{Y_{k-1}}$	L_k	L_{k-1}	E_k	W_k	C_k
$\sqrt{Y_{k-1}}$	1	-0.25	-0.08	-0.20	0.08	0.33
L_k	-0.25	1	0.08	-0.18	0.03	-0.10
L_{k-1}	-0.08	0.08	1	0.07	0.05	-0.10
E_k	-0.20	-0.18	0.07	1	0.18	-0.20
W_k	0.08	0.03	0.05	0.18	1	0.30
C_k	0.33	-0.10	-0.10	-0.20	0.30	1

Table 7.3: The pairwise correlation coefficients of the variables in regression (7.1).

As all regression assumptions are satisfied, it is concluded that the regression model represented by equation (7.1) is valid. A graphical display of the fit and forecast for the years 2011-2014, compared to actual demand, is given in Figure 7.1.

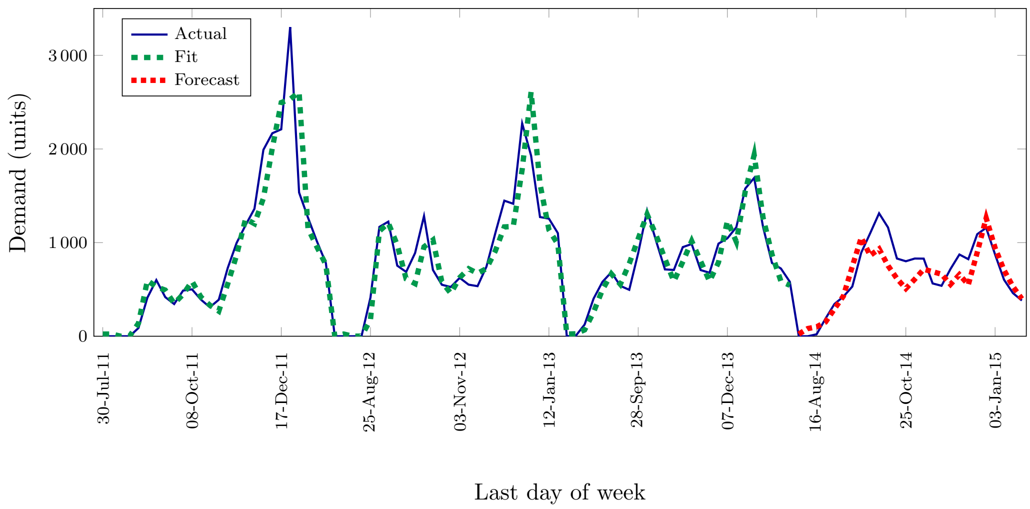


Figure 7.1: Graphical display of the fit and forecast of Regression 7.1 for the years years 2011–2014 for Subclass B_S. Simulation Model S_1 was used to simulate sales.

The graph indicates a very accurate fit. The forecast is also accurate, although there are a few slight mispredictions. The total of the forecast underestimates total actual demand by about 9%. One cause of the underestimation could be the low demand from August to October in 2011, which caused the corresponding weeks' demand to be underestimated during other years. In 2011–2013, the variable $\sqrt{Y_{k-1}}$ corrects for this underestimation to some extent, but during 2014, $\sqrt{Y_{k-1}}$ is not known and $\sqrt{\hat{Y}_{k-1}}$ is used as a proxy for $\sqrt{Y_{k-1}}$. From the end of November 2014 onwards, demand is also mostly underestimated, because very little new stock was sent and demand was higher than would have been expected. Old stock might have played a more significant role in this data set than the original training set.

The lower stock levels is also a possible reason for the lower peak in actual demand during Christmas time, which was predicted by the regression model. The only time during 2014 when the model significantly overestimates demand is in the second last week of November, possibly due to a higher demand during the corresponding week in previous years. Inflows during the previous week were also quite high (2 459), but demand did not increase as expected.

7.1.2 Regression for Simulation Model S_2 for Subclass B_S

Five regression equations were fitted for Simulation Model S_2, as Subclass B_S contains five sizes. The sizes range from 6 to 10. The regression equations for Simulation Model S_2 when applied to Subclass B_S are given by

$$\sqrt{Y_{6,k}} = 0.86\sqrt{Y_{6,k-1}} + 0.0017L_{6,k} + 0.0016L_{6,k-1} + 2.62E_k + 0.0047W_k + 4.03C_k, \quad (7.2)$$
$$\sqrt{Y_{7,k}} = 0.86\sqrt{Y_{7,k-1}} + 0.0013L_{7,k} + 0.0015L_{7,k-1} + 1.89E_k + 0.014W_k + 4.20C_k, \quad (7.3)$$
$$\sqrt{Y_{8,k}} = 0.86\sqrt{Y_{8,k-1}} + 0.0014L_{8,k} + 0.0014L_{8,k-1} + 1.94E_k + 0.017W_k + 4.33C_k, \quad (7.4)$$
$$\sqrt{Y_{9,k}} = 0.86\sqrt{Y_{9,k-1}} + 0.0017L_{9,k} + 0.0017L_{9,k-1} + 1.42E_k + 0.014W_k + 3.82C_k, \text{ and} \quad (7.5)$$
$$\sqrt{Y_{10,k}} = 0.88\sqrt{Y_{10,k-1}} + 0.0024L_{10,k} + 0.0020L_{10,k-1} + 0.99E_k + 0.012W_k + 2.01C_k. \quad (7.6)$$

Again, all coefficients are positive as in the company level model, and the relative size of the coefficients are similar to Subclass A_S. The equations therefore make intuitive sense. All R^2 and adjusted R^2 values are at least 0.98, indicating a very good fit. Assumption 7 implies that these models are valid because the company level model is valid, so the regression assumptions for these models are not tested again.

A graphical display of the fit and forecast for regression (7.4) (the regression for Size 8) in years 2011-2014 is given in Figure 7.2. The results for other sizes are similar. The pattern in the fit and forecast is very similar to that of the company level model: under- and overestimations occur at the same places for the same reasons.

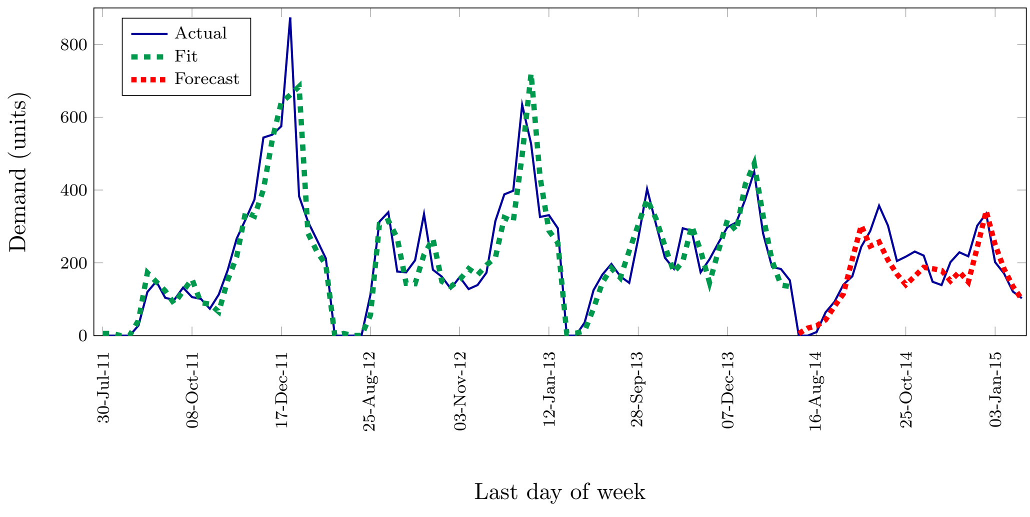


Figure 7.2: Graphical display of the fit and forecast of regression 7.4 (the regression model for Size 8) in years 2011-2014 for Subclass B_S.

7.1.3 Accuracy of Simulation Models S_1 and S_2 for Subclass B_S

After establishing that the regression models for Simulation Model S_1 and Simulation Model S_2 as applied to Subclass B_S are valid and accurate, the rest of the simulation was performed for both models. The accuracy of the two models can be verified by calculating the ICC statistics for sales generated by the models when using 2014's inflows, compared to actual sales. The section also contains a graphical display of the weekly sales simulated on a company level by both models, together with actual sales.

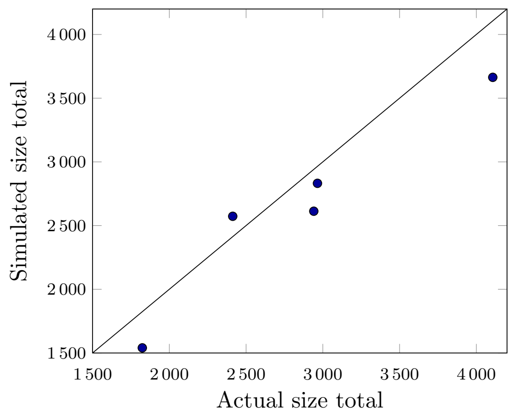
The ICC values and total sales generated for Subclass B_S by Simulation Model S_1 and Simulation Model S_2 are given in Table 7.4. Most ICC statistics are higher than 0.9, except for the single measure statistics for week totals. The slightly lower week totals is probably because of the large underestimation in demand by the regression models; however, the values are still very close to 0.9 and it is concluded that both models accurately represent reality.

For this subclass, Simulation Model S_2 is more accurate than Simulation Model S_1, because all ICC values for Simulation Model S_2 is greater than or equal to the corresponding values for Simulation Model S_1. The number of unit sales generated by Simulation Model S_2 is also slightly closer to the actual sales figure, namely 14 250. The difference between the two models is, however, not very big. Both models underestimate total sales, Simulation Model S_1 by about 7% and Simulation Model S_2 by about 6%.

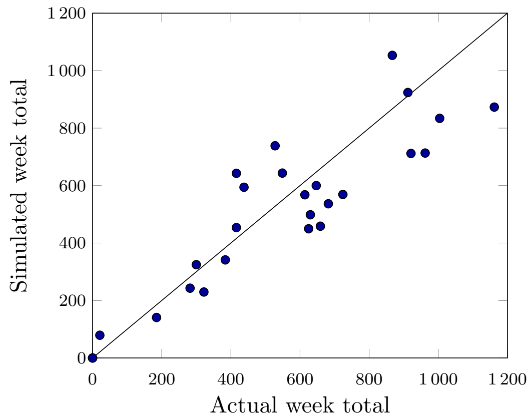
Measure	Totals grouped by	Simulation Model S_1	Simulation Model S_2
Total sales		13220.9	13391.9
ICC(A,1)	size	0.935	0.955
	week	0.884	0.893
	store	0.933	0.934
ICC(C,1)	size	0.958	0.970
	week	0.888	0.895
	store	0.936	0.936
ICC(A,k)	size	0.966	0.977
	week	0.938	0.944
	store	0.966	0.966
ICC(C,k)	size	0.978	0.985
	week	0.941	0.945
	store	0.967	0.967

Table 7.4: ICC values for Simulation Model S_1 and Simulation Model S_2 applied to Subclass B_S. Simulated size totals, week totals and store totals were compared to actual totals. A value of $w = 0.99$ was used.

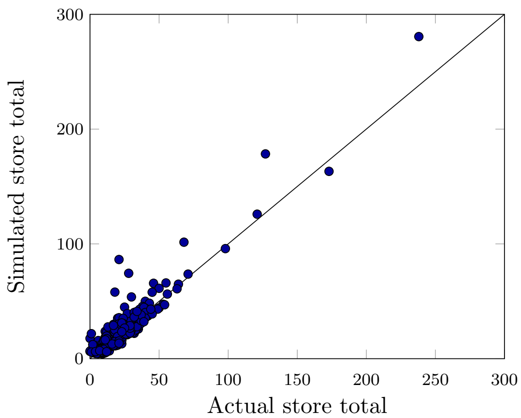
Figure 7.3 contains scatter plots indicating the correlation between simulated and actual size, week and store totals for Subclass B_S, where Simulation Model S_1 was used to generate sales. In all correlation plots, data points are spread evenly about the line $y = x$, although there are more points below the line because actual sales are underestimated. This is also true for store totals, although it is difficult to observe because so many points are clustered together in the lower left-hand corner. The cluster of points is due to the fact that most stores sell less than 100 units per season of Subclass B_S. The scatter plots for Simulation Model S_2 are almost identical and are therefore not displayed to avoid duplication.



(a) Correlation for size totals



(b) Correlation for week totals.



(c) Correlation for store totals.

Figure 7.3: Scatter plots indicating the correlation between simulated and actual (a) size totals, (b) week totals and (c) store totals for Subclass B_S. The simulation was performed using Simulation Model S_1.

In Figure 7.4, the total weekly sales on a company level generated by 10 simulation runs of each model is displayed against actual sales. Over- and underestimations are due to mispredictions of demand by the regression model. Other than that, the sales predicted by both models are accurate. The two models' predictions are also close together.

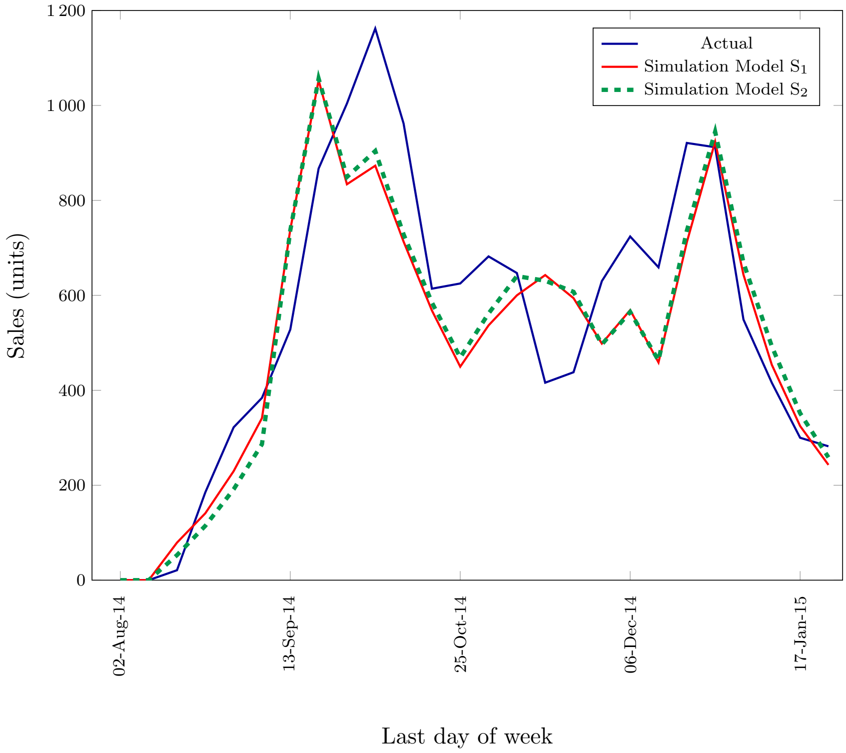


Figure 7.4: Average weekly simulated sales over 10 runs for Subclass $\mathrm{B_S}$ by Simulation Model $\mathrm{S_1}$ and Simulation Model $\mathrm{S_2}$, together with actual sales for Subclass $\mathrm{B_S}$.

7.2 Models for Winter products

The simulation models that were developed with Subclass $\mathrm{A_W}$ as a training set were applied to Subclass $\mathrm{B_W}$. In this section, Simulation Model $\mathrm{W_1}$ and Simulation Model $\mathrm{W_2}$ are validated for the new data. The regression models are validated in §7.2.1–7.2.2, and the final simulation results are validated against actual values when using actual 2014 inflows in §7.2.3.

7.2.1 Regression for Simulation Model $\mathrm{W_1}$ for Subclass $\mathrm{B_W}$

The regression equation of Simulation Model $\mathrm{W_1}$ when applied to Subclass $\mathrm{B_S}$ is given by

$$\begin{aligned}\sqrt{Y_k} = {} & 4.80 + 0.81\sqrt{Y_{k-1}} + 0.0011L_k + 0.0009L_{k-1} + 5.48E_k - 5.93F_k - 2.13M_k + \\ & - 1.76A_k + 0.29J_k - 6.92U_k - 17.79G_k. \end{aligned} \tag{7.7}$$

The signs of the regression coefficients are the same as for Subclass $\mathrm{A_W}$, except that J_k's coefficient is positive. This indicates that for this subclass, which is ladies' spun poly jackets, June's demand is on average slightly higher than April's demand. This may be due to the fact that it is usually colder during June than during April, so that the demand for warm jackets is higher. Subclass A_W is girls' shoes, for which the demand will not necessarily be higher during the colder month of June. The relative sizes of the coefficients are similar to the other data set, although all coefficients are smaller, because total demand for this product is lower than for the other product. The regression equation therefore makes intuitive sense.

For each variable, a t test was performed to test its significance in explaining demand. The t test values for regression (7.7), together with their accompanying p values, are given in Table 7.5. The intercept's p value indicates that it is not significant for this data set at a significance level of 0.05. The variables $\sqrt{Y_{k-1}}$, L_k, L_{k-1} and E_k are highly significant. The variables U_k and G_k are also highly significant, indicating the need for monthly dummy variables, as demand during July and August is significantly different from demand during April, the reference month. The dummy variables F_k, M_k and A_k have high p values, indicating that demand during February, March and May is not significantly different from demand during April, but they may differ from one another.

Variable	t value	p value
Intercept	1.56	0.1231
$\sqrt{Y_{k-1}}$	16.81	< 0.0001
L_k	5.09	< 0.0001
L_{k-1}	4.29	< 0.0001
E_k	4.04	0.0001
F_k	-1.76	0.0838
M_k	-0.83	0.4119
A_k	-0.73	0.4661
J_k	0.12	0.9021
U_k	-2.74	0.0078
G_k	-2.92	0.0048

Table 7.5: The t test and the accompanying p values for the independent variables in regression (7.1).

The joint explanatory power of the variables was tested with the F test. The p value of the F statistic is lower than 0.0001, indicating that collectively, the independent variables have significant explanatory power. The value of $R^2 = 0.947$ and the value of the adjusted $R^2 = 0.940$, indicating a very good fit.

The coefficients are constants, so Assumption 1 of multiple linear regression, namely that the regression equation is linear in its parameters, holds. The Breusch-Pagan test was performed to test whether the homoscedasticity assumption of residuals (Assumption 2) holds. The p value reported by SAS [61] is 0.0194. The null hypothesis of homoscedasticity is rejected at a significance level of 0.05. However, for the purposes of the model a significance level of 0.01, at which the null hypothesis is not rejected, is acceptable. White's test was also performed to obtain more certainty. The p value for White's test is 0.1037, which means that the null hypothesis of homoscedasticity in White's test is not rejected at a significance level of 0.05. Homoscedasticity is therefore assumed.

The Shapiro-Wilk, Kolmogorov-Smirnov, Cramér von Mises and Anderson-Darling tests for normality were performed on residuals to test whether residuals are normally distributed with a mean value of 0. The test statistics and p values for the tests are given in Table 7.6. The last column indicates whether the null hypothesis of normality is rejected at a significance level of 0.05.

All p values are higher than 0.05 and the null hypotheses of normality in the residuals are not rejected. The reported mean is 0 and the standard deviation is 5.25. It may therefore be

Test	Statistic	p value
Shapiro-Wilk	$W = 0.97$	0.079
Kolmogorov-Smirnov	$D = 0.09$	0.089
Cramér von Mises	$W^2 = 0.10$	0.13
Anderson-Darling	$A^2 = 0.67$	0.08

Table 7.6: Results for normality tests on the residuals of regression (7.7). A significance level of $\alpha = 0.05$ is used.

assumed that the residuals are normally distributed with a mean value of 0.

Finally, the assumption of multicollinearity (Assumption 5) is tested. There is no signal of multicollinearity in the information recorded so far. The signs and size of the regression coefficients are reasonable, and most of the p values are significant. The pairwise correlation coefficients between variables are reported in Table 7.7.

$\sqrt{Y_{k-1}}$	L_k	L_{k-1}	E_k	F_k	M_k	A_k	J_k	U_k	G_k
1	0.02	0.23	−0.08	−0.63	−0.32	0.34	0.35	0.10	−0.04
0.02	1	0.27	−0.05	−0.11	0.08	0.07	−0.09	−0.25	−0.08
0.23	0.27	1	0.02	−0.20	0.10	0.09	0.05	−0.24	−0.08
−0.08	−0.05	0.02	1	−0.08	0.09	−0.03	0.00	−0.03	0.18
−0.63	−0.11	−0.20	−0.08	1	−0.19	−0.17	−0.18	−0.17	−0.06
−0.32	0.08	0.10	0.09	−0.19	1	−0.21	−0.22	−0.21	−0.07
0.34	0.07	0.09	−0.03	−0.17	−0.21	1	−0.20	−0.20	−0.06
0.35	−0.09	0.05	0.00	−0.18	−0.22	−0.20	1	−0.20	−0.06
0.10	−0.25	−0.24	−0.03	−0.17	−0.21	−0.20	−0.20	1	−0.06
−0.04	−0.08	−0.08	0.18	−0.06	−0.07	−0.06	−0.06	−0.06	1

Table 7.7: The pairwise correlation coefficients of the variables in regression (7.7).

The absolute value of all correlation coefficients between different variables are lower than 0.8. Therefore it may be assumed that multicollinearity does not play a significant role.

As all regression assumptions are satisfied, it is concluded that the regression model represented by equation (7.7) is valid. A graphical display of the fit and forecast for the years 2011-2014, compared to actual demand, is given in Figure 7.5.

As in other data sets, the fit is accurate. The pattern of the forecast is also close to reality, except that actual demand is underestimated during March and July and overestimated during the first two weeks of June. Demand during March and July is possibly underestimated because demand during March and July 2014 was higher than average demand during the same months in historical years. Demand during the first two weeks in June is probably overestimated because of a large inflows during the first week in June which did not influence demand as much as predicted by the model. The regression's total demand overestimates actual demand by about 7%.

7.2.2 Regression for Simulation Model $\mathbf{W_2}$ for Subclass $\mathbf{B_S}$

Six regression equations were fitted for Simulation Model W_2, as Subclass B_S contains six sizes. The sizes range from 32 to 42 and increase in intervals of 2. The regression equations for Simulation Model W_2 when applied to Subclass B_S are given by

$$\sqrt{Y_{32,k}} = 1.50 + 0.76\sqrt{Y_{32,k-1}} + 0.0053L_{32,k} + 0.0041L_{32,k-1} + 1.95E_k - 1.90F_k - 0.42M_k +$$

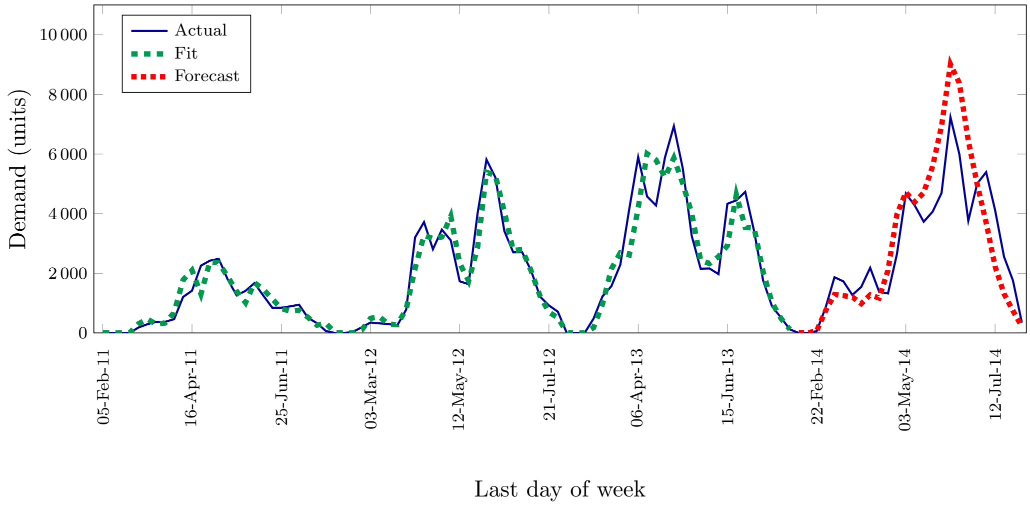


Figure 7.5: Graphical display of the fit and forecast of regression (7.7) in years 2011–2014 for Subclass B_W.

$$- 0.52A_k + 0.22J_k - 1.86U_k - 4.41G_k, \tag{7.8}$$

$$\sqrt{Y_{34,k}} = 1.73 + 0.79\sqrt{Y_{34,k-1}} + 0.0034L_{34,k} + 0.0027L_{34,k-1} + 2.09E_k - 2.17F_k - 0.91M_k - 0.42A_k + 0.35J_k - 2.41U_k - 6.02G_k, \tag{7.9}$$

$$\sqrt{Y_{36,k}} = 1.79 + 0.83\sqrt{Y_{36,k-1}} + 0.0020L_{36,k} + 0.0020L_{36,k-1} + 2.48E_k - 2.48F_k - 0.80M_k - 0.92A_k + 0.32J_k - 2.90U_k - 8.56G_k, \tag{7.10}$$

$$\sqrt{Y_{38,k}} = 2.27 + 0.82\sqrt{Y_{38,k-1}} + 0.0020L_{38,k} + 0.0018L_{38,k-1} + 2.43E_k - 2.91F_k - 1.41M_k - 0.93A_k + 0.22J_k - 3.06U_k - 8.74G_k \tag{7.11}$$

$$\sqrt{Y_{40,k}} = 2.20 + 0.83\sqrt{Y_{40,k-1}} + 0.0021L_{40,k} + 0.0019L_{40,k-1} + 2.19E_k - 2.63F_k - 0.97M_k - 0.77A_k - 0.12J_k - 3.43U_k - 7.81G_k, \text{ and} \tag{7.12}$$

$$\sqrt{Y_{42,k}} = 2.75 + 0.79\sqrt{Y_{42,k-1}} + 0.0024L_{42,k} + 0.0021L_{42,k-1} + 2.20E_k - 2.99F_k - 0.99M_k - 0.81A_k - 0.44J_k - 3.53U_k - 7.70G_k. \tag{7.13}$$

The regression coefficients have the same signs as the company level regression, and have more or less the same relative sizes. The equations therefore make intuitive sense. All R^2 values are between 0.94 and 0.95, and the adjusted R^2 values are between 0.93 and 0.94, indicating a very good fit. Assumption 7 implies that these models are valid because the company level model is valid, so the regression assumptions for these models are not tested again.

A graphical display of the fit and forecast for regression (7.9) (Size 34) in years 2011-2014 is given in Figure 7.6. The results for other sizes are similar. The pattern of the fit and forecast is very similar to that of the company level regression model. Under- and overestimations occur in the same places and for the same reasons.

7.2.3 Accuracy of Simulation Model W_1 and Simulation Model W_2 for Subclass B_W

After establishing that the regression models for Simulation Model W_1 and Simulation Model W_2 when applied to Subclass B_W are valid and accurate, the rest of the simulation was performed for both models. In this section, the accuracy of the two models are verified by calculating the ICC statistics for sales generated by the models when using 2014's inflows, compared to actual

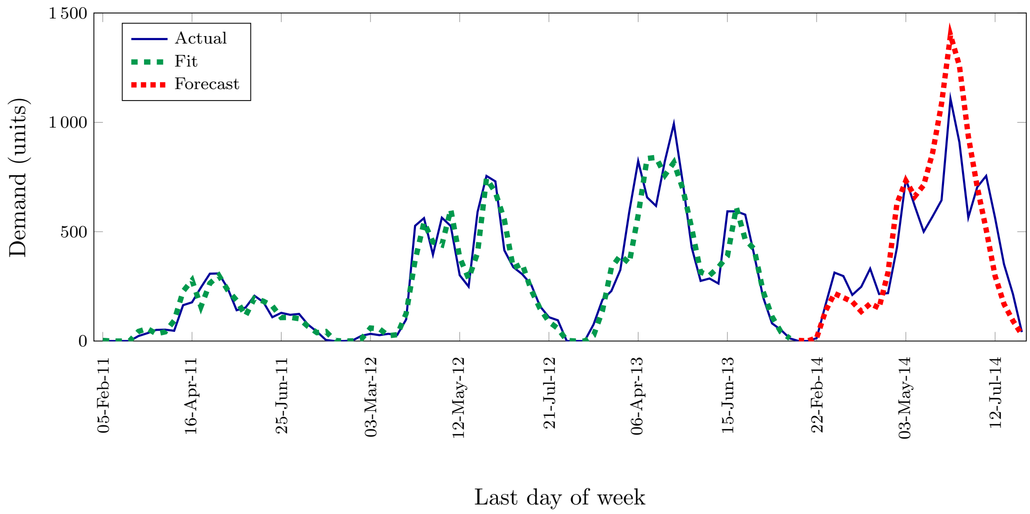


Figure 7.6: Graphical display of the fit and forecast of regression (7.9) (Size 34) in years 2011–2014, for Subclass A_W.

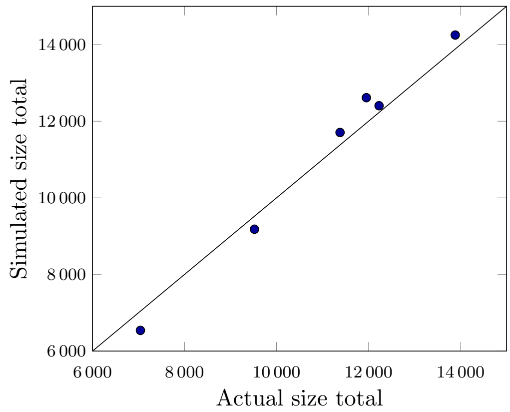
sales. A graphical display of the weekly sales simulated on a company level by both models, compared to actual sales, is also supplied.

The ICC values and total sales generated for Subclass B_W by Simulation Model W_1 and Simulation Model W_2 are given in Table 7.8. In both cases, the actual number of units of total sales (66 011) is overestimated. Simulation Model W_1 overestimates sales by about 1% and Simulation Model W_2 by about 2%. According to the ICC statistics, Simulation Model W_1 is more accurate for size and store totals than Simulation Model W_2, but Simulation Model W_2 is more accurate for week totals. Most of the correlation coefficients for both models is above 0.9, except the single measures for week totals. This may be due to the high overestimation of demand by the regression model which in turn led to an overestimation in total sales. It is concluded that both models are accurate representations of reality, and overall one is not more accurate than the other.

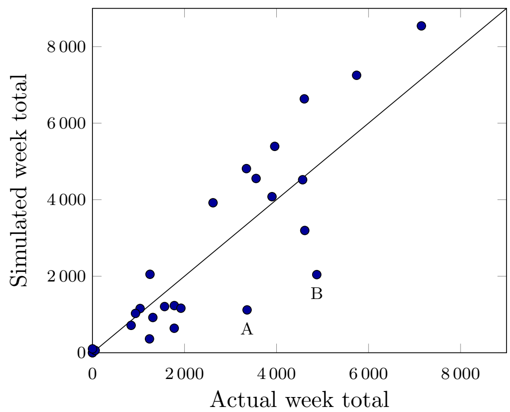
ICC statistic	Totals grouped by	Simulation Model W_1	Simulation Model W_2
Total sales		66697.6	67239.3
	size	0.987	0.984
ICC(A,1)	week	0.867	0.888
	store	0.978	0.971
	size	0.985	0.985
ICC(C,1)	week	0.862	0.884
	store	0.978	0.972
	size	0.993	0.992
ICC(A,k)	week	0.929	0.941
	store	0.989	0.986
	size	0.993	0.992
ICC(C,k)	week	0.926	0.939
	store	0.989	0.986

Table 7.8: ICC values and total sales generated for Subclass B_W by Simulation Model W_1 and Simulation Model W_2.

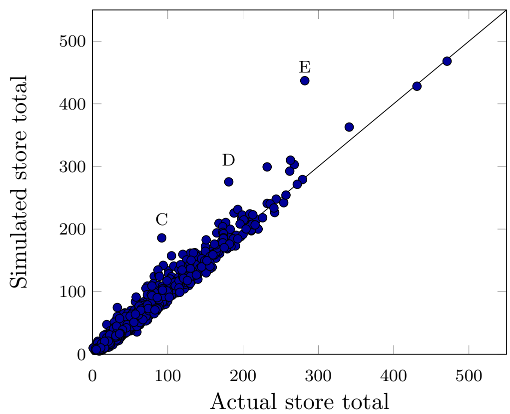
Figure 7.7 contains scatter plots indicating the correlation between simulated and actual size, week and store totals for Subclass B_W, where Simulation Model W_1 was used to generate sales. The scatter plots for Simulation Model W_2 are almost identical and are therefore not displayed to avoid duplication.



(a) Correlation for size totals



(b) Correlation for week totals.



(c) Correlation for store totals.

Figure 7.7: Scatter plots indicating the correlation between simulated and actual (a) size totals, (b) week totals and (c) store totals for Subclass B_W. Simulation Model W_1 was used to simulate sales

The data points representing size totals in Figure 7.7(a) are spread evenly about the line $y = x$. The data points representing week totals in Figure 7.7(b) are mostly spread evenly about the line, except for the points labelled 'A' and 'B'. These points represent the first two weeks of July, during which demand was underestimated, which also lead to an underestimation in sales.

For the store totals in Figure 7.7(c), most of the points are clustered in the bottom left-hand part of the graph, because most stores sold less than 300 units of this subclass in 2014. The points are mostly close to the line $y = x$, except for the points labelled 'C', 'D' and 'E'. Point C represents sales for a store in Rustenburg, for which the historical demand proportion was 0.003, while the demand proportion in 2014 was only 0.002. Similarly, point D represent sales for a store in Upington, for which the historical demand proportion was 0.002, while the demand proportion in 2014 was only 0.001. The higher historical demand proportion caused the model to overpredict actual demand, which lead to an overprediction in actual sales. Point E represents a store in Parow for which the historical demand proportion was about 0.004, but the proportion of inflows that was sent to the store was 0.006. The availability factor in the simulation model caused the model to predict a higher demand than what actually took place, which in turn lead to an overprediction in sales.

In Figure 7.8, the total weekly sales on a company level generated by 10 simulation runs of Simulation Model W_1 and Simulation Model W_2 are displayed against actual sales. Over- and underestimations are due to mispredictions of demand by the regression model. The lines representing the sales for Simulation Model W_1 and Simulation Model W_2 are very close together, and in both cases, the pattern accurately imitates reality.

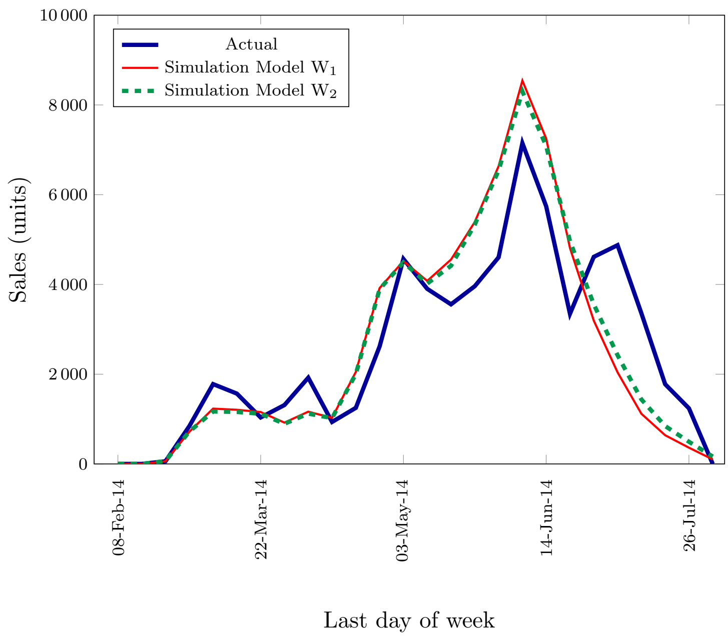


Figure 7.8: Weekly simulated sales for Subclass B_W by Simulation Model W_1 and Simulation Model W_2, compared to actual sales.

7.3 Comparison of allocation models by means of simulation

After validating that all four models generate sufficiently accurate results, they were used to compare allocation algorithms to one another. The allocation algorithms that were described in Chapter 4, as well as the relaxation of Algorithm 3 described in §4.4 (Algorithm 3′) were tested for Subclass B_S and B_W.

For each data set, each allocation algorithm was implemented and supplied as input to the simulation models when simulating sales. The same seeds were used when generating weekly demand during all experiments for the same data set.

The total number of unit sales, shortages and surpluses for Subclass B_S generated with each allocation algorithm by Simulation Model S_1 and Simulation Model S_2 are supplied in Table 7.9(a). Results are reported in number of units, averaged over 10 simulation replications. The same results are reported in terms of percentage improvement on PEP in Table 7.9(b).

	Measure	PEP	Algorithm 2	Algorithm 3	Algorithm 3′
Simulation Model S_1	sales	13166.9	13115.4	13183	13210.2
	shortages	2602.4	2653.9	2586.3	2559.1
	surpluses	5771.1	5822.6	5755	5727.8
Simulation Model S_2	sales	13340.6	13332.8	13327.4	13323.3
	shortages	2215.3	2223.1	2228.5	2232.6
	surpluses	5597.4	5605.2	5610.6	5614.7

(a) Total sales, shortages and surpluses in units.

	Measure	Algorithm 2	Algorithm 3	Algorithm 3′
Simulation Model S_1	sales	-0.39%	0.12%	0.33%
	shortages	-1.98%	0.62%	1.66%
	surpluses	-0.89%	0.28%	0.75%
Simulation Model S_2	sales	-0.06%	-0.10%	-0.13%
	shortages	-0.35%	-0.60%	-0.78%
	surpluses	-0.14%	-0.24%	-0.31%

(b) Average percentage improvement on PEP.

Table 7.9: Results for different allocation algorithms as generated by Simulation Model S_1 and Simulation Model S_2 for Subclass B_S in (a) number of units and (b) the average percentage improvement on PEP.

According to Simulation Model S_1, Algorithm 3′ performed the best, followed by Algorithm 3, followed by PEP, and Algorithm 2 performed the worst. However, the difference between the sales, shortages and surpluses of the best and worst performing algorithms is less than 100 units. The biggest percentage difference in terms of sales is less than 1% and in terms of shortages and surpluses less than 3%. In reality, Simulation Model S_1 predicted no significant difference between the different algorithms.

According to Simulation Model S_2, PEP performed best, followed by Algorithm 2, then Algorithm 3 and then Algorithm 3′. The order of performance is different than what was predicted by Simulation Model S_1. The difference between the best and worst performing algorithm is even smaller than for Simulation Model S_1. A possible reason for the smaller difference is that, in this model, availability on a size level also influences demand, so that units sell where they are available to a greater extent than for Simulation Model S_1. The difference between sales, shortages and surpluses generated by any two different algorithms is less than 20 units. The percentage difference between any two algorithms is less than 0.2% in terms of sales and less than 1% in terms of shortages and surpluses.

Table 7.10 contains the corresponding results for Subclass B_W. Results are very similar to the results for the Summer products, except that Algorithm 2's results are slightly better than PEP's results for both Simulation Model W_1 and Simulation Model W_2. Again, the differences between algorithms are not significantly large.

	Measure	PEP	Algorithm 2	Algorithm 3	Algorithm 3′
	sales	67083.20	67119.50	67201.20	67278.40
Simulation Model W_1	shortages	11242.80	11206.50	11124.80	11047.60
	surpluses	8534.80	8498.50	8416.80	8339.60
	sales	66078.56	66081.00	66056.22	66030.00
Simulation Model W_2	shortages	12731.33	12728.89	12753.67	12777.40
	surpluses	9539.44	9537.00	9561.78	9588.00

(a) Total sales, shortages and surpluses in units.

	Measure	Algorithm 2	Algorithm 3	Algorithm 3′
	sales	0.05%	0.18%	0.29%
Simulation Model W_1	shortages	0.32%	1.05%	1.74%
	surpluses	0.43%	1.38%	2.29%
	sales	0.00%	-0.03%	-0.07%
Simulation Model W_2	shortages	0.02%	-0.18%	-0.36%
	surpluses	0.03%	-0.23%	-0.51%

(b) Average percentage improvement on PEP.

Table 7.10: Results for different allocation algorithms as generated by Simulation Model W_1 and Simulation Model W_2 for Subclass B_W in (a) number of units and (b) the average percentage improvement on PEP.

Judging by both subclasses and both models, it is concluded that the new allocation algorithms do not increase or decrease total sales significantly compared to PEP's algorithm. According to the first modelling approach, Algorithm 3′ and Algorithm 3 improve slightly on PEP's algorithm, and Algorithm 2's results are sometimes slightly better and sometimes slightly worse than PEP. Algorithm 3′ performs the best of the four methods. The second modelling approach measures a very small difference between the four methods in all cases.

One possible reason for the little difference among the methods is that all four models aim for the expected demand as predicted by PEP. Within the constraints of integrality and received stock, there is not much scope for optimisation in terms of number of unit sales. The little optimisation that may be possible is counteracted by the fact that availability influences demand, which means that stock sells at the stores where they are available. This is, however, only true within bounds: when deliberately sending more stock to stores where demand is expected to be lower, sales decrease significantly.

Another reason may be that PEP has many small stores that sell very little units in each size per season. It is nearly impossible to accurately predict demand over time for these stores, because they do not follow a specific pattern. When moving units between stores with random demand, the sales that realise in the end do not change significantly.

CHAPTER 8

Conclusions

This chapter provides a summary of this study. In §8.1, a summary of the findings of the study is given. In §8.2, recommendations are made based on results, and ideas for future work are discusssed. The chapter is concluded with a summary of the thesis in §8.3.

8.1 Summary of findings

In this thesis, four size-mix allocation methods are compared to one another by simulating sales. The current method used by PEP, an existing method developed by Thom *et al.*, a new method and a relaxation of the new method are tested.

Two simulation modelling approaches are followed, and separate models are developed for Summer and Winter products. Both approaches are more or less equally accurate according to ICC measures. Analyses of visual displays lead to the conclusion that weekly sales on a company level are simulated very accurately by both models, as well as weekly sales on a size level. On a store level, the total sales for the season are very accurate, but in most cases the simulation models are not able to accurately predict the time of the season at which sales take place. This is because the demand for these stores is unpredictable and occur at random times during the season. However, the total sales per store for the season are simulated very accurately according to ICC measures.

The results for both modelling approaches indicate that there is no significant difference between the four allocation methods. A possible reason is that there is not much scope for optimisation given the constraints of the problem and the fact that all four models aim for the same expected demand. Because a higher availability increases demand, moving one or two units to another store does not significantly influences sales.

8.2 Recommendations and future work

Based on the results in Chapter 7, it is recommended that PEP implement Allocation Algorithm $3'$, as its performance is the best of the four models. However, the improvement is so small that PEP may continue using the current method if the implementation of a new method is regarded too expensive.

Research could be continued to further refine simulation models. The most appropriate value for the parameter w could be found through more experiments. The simulation models could also be applied to more data sets to test their ability to handle a range of different products. A few alterations may be made to regression models to more accurately simulate sales for a wider range of products.

It has been established in the study that small deviations from expected demand does not influence total sales, possibly because a higher availability at a store increases demand. Sensitivity analyses to determine how large a deviation would have a significant impact on total sales could be valuable to PEP.

Because small stores exhibit unpredictable demand, optimisation possibilities for the small stores are limited. An idea for future research is to group smaller stores together into one entity when allocating stock. The new entities will then have larger demand which is more predictable. This may allow more scope for optimisation.

Different approaches to simulation modelling may also be tried. Factors other than historical sales and availability could be incorporated to simulate store demand, for example the composition of the population near a store.

Either one of the two simulation approaches developed in this thesis could be used during further experimentation, as both models were found to accurately represent reality. PEP may also use either one of these models to test the expected effect when varying allocations or inputs to their allocation algorithm. The first modelling approach is recommended, because this approach assumes that availability of a product influences demand, but not availability of size. This is more accurate than assuming that a size's availability influences demand for the size, as a particular person's dress or shoe size cannot be influenced by availability.

8.3 Thesis summary and achievement of objectives

In Chapter 1 of this thesis, Objective I was achieved by providing a description of allocation decisions within the broader context of the supply chain and applied to PEP's situation. The scope and objectives of the thesis were also explained.

Chapter 2 contains a discussion of the existing literature relevant to the study. It was established that, for the purposes of this study, simulation modelling based on underlying regression is the most appropriate method to measure the effectiveness of an allocation algorithm. Hereby Objective II was reached.

Sales and allocation data that were received, as well as the handling of the data, were described in Chapter 3 in fulfilment of Objective III. Objective IV was achieved in Chapter 4, in which the allocation algorithms that were compared to one another were described.

In Chapters 5 and 6, four simulation models, based on underlying regression models, were developed in order to simulate sales when using a specific allocation algorithm. Two simulation approaches were followed and separate models were developed for Summer and Winter products. This was done in fulfilment of Objective V. The development as well as the validation, verification and accuracy of the models were described. All four models are valid and accurate representations of reality. The two models for Summer products were found to be more or less equally accurate, and the same applies to the two models for Winter products.

In Chapter 7, the models were applied to new data sets. Regression assumptions were verified and the accuracy of the models for the new data sets was discussed. Again, the accuracy of the two modelling approaches did not differ significantly. The allocation algorithms described in Chapter 4 were implemented for these data sets and compared to one another with regards to total unit sales, shortages and surpluses. Hereby Objective VI(a) was reached. The results for the four algorithms were very similar, and no algorithm could improve significantly on the current system in PEP.

Finally, in this chapter, recommendations were made and ideas for future research were discussed. A summary of the thesis was also provided. This concludes the fulfilment of Objective VI(b).

References

[1] Agrawal N & Smith S A, 1996. *Estimating negative binomial demand for retail inventory management with unobservable lost sales.* Naval Research Logistics (NRL), **43(6)**, pp. 839–861.

[2] Ahmed K I, 1999. *Modeling drivers' acceleration and lane changing behavior.* Ph.D. thesis, Massachusetts Institute of Technology, Cambridge (MA).

[3] Alon I, Qi M & Sadowski R J, 2001. *Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods.* Journal of Retailing and Consumer Services, **8(3)**, pp. 147–156.

[4] Anupindi R, Dada M & Gupta S, 1998. *Estimation of consumer demand with stock-out based substitution: An application to vending machine products.* Marketing Science, **17(4)**, pp. 406–423.

[5] Axsäter S, Marklund J & Silver E A, 2002. *Heuristic methods for centralized control of one-warehouse, N-retailer inventory systems.* Manufacturing & Service Operations Management, **4(1)**, pp. 75–97.

[6] Bartholdi III J J & Hackman S T, 2002. *Warehouse & distribution science: release 0.1.2.* [Online], [cited 1 October 2013], Available at: `http://www2.isye.gatech.edu/~spyros/courses/IE6202/Fall-2002/Bartholdi-Hackman.pdf`.

[7] Bartko J J & Carpenter W T, 1976. *On the methods and theory of reliability.* The Journal of nervous and mental disease, **163(5)**, pp. 307–317.

[8] Berman B & Joel R E, 2001. *Retail management.* Eighth edition. Prentice Hall, Upper Saddle River (NJ).

[9] Breusch T S & Pagan A R, 1979. *A simple test for heteroscedasticity and random coefficient variation.* Econometrica: Journal of the Econometric Society, pp. 1287–1294.

[10] Burghout W, 2004. *A note on the number of replication runs in stochastic traffic simulation models*, (Unpublished), Centre for Traffic Research, Stockholm.

[11] Callen J L, Kwan C C, Yip P C & Yuan Y, 1996. *Neural network forecasting of quarterly accounting earnings.* International Journal of Forecasting, **12(4)**, pp. 475–482.

[12] Caro F & Gallien J, 2010. *Inventory management of a fast-fashion retail network.* Operations Research, **58(2)**, pp. 257–273.

[13] Caro F, Gallien J, Díaz M, García J, Corredoira J M, Montes M, Ramos J A & Correa J, 2010. *Zara uses operations research to reengineer its global distribution process.* Interfaces, **40(1)**, pp. 71–84.

[14] Chu C W & Zhang G P, 2003. *A comparative study of linear and nonlinear models for aggregate retail sales forecasting.* International Journal of production economics, **86(3)**, pp. 217–231.

[15] Church K B & Curram S P, 1996. *Forecasting consumers' expenditure: A comparison between econometric and neural network models.* International Journal of Forecasting, **12(2)**, pp. 255–267.

[16] Conlon C T & Mortimer J H, 2008. *Demand estimation under incomplete product availability.* Technical report, National Bureau of Economic Research.

[17] Conrad S, 1976. *Sales data and the estimation of demand.* Operational Research Quarterly, **27(1)** pp. 123–127.

[18] Darbellay G A & Slama M, 2000. *Forecasting the short-term demand for electricity: Do neural networks stand a better chance?* International Journal of Forecasting, **16(1)**, pp. 71–83.

[19] De Villiers A P, 2011. *Minimising the total travel distance to pick orders on a unidertional picking line.* Masters thesis, Stellenbosch University, Stellenbosch, South Africa.

[20] Dempster A P, Laird N M & Rubin D B, 1977. *Maximum likelihood from incomplete data via the EM algorithm.* Journal of the royal statistical society, **39(1)**, pp. 1–38.

[21] Donofrio T, 2010. *Advanced Planning and Optimization Series.* [Online], [cited 23 October 2013] Available at: `http://risnews.edgl.com/Media/DocumentLibrary/RIS%20Plan%20Articles%20No%201-7%20Advanced%20Planning%20and%20Optimization%20Series%20030510.pdf`.

[22] Dréo J, Petrowski A, Siarry P & Taillard E, 2006. *Metaheuristics for hard optimization: methods and case studies.* Springer Science & Business Media, Berlin, Germany.

[23] Elkateb M M, Solaiman K & Al-Turki Y, 1998. *A comparative study of medium-weather-dependent load forecasting using enhanced artificial/fuzzy neural network and statistical techniques.* Neurocomputing, **23(1)**, pp. 3–13.

[24] Forst F G, 2011. *Forecasting restaurant sales using multiple regression and Box-Jenkins analysis.* Journal of Applied Business Research (JABR), **8(2)**, pp. 15–19.

[25] Gaul C, Kurz S & Rambau J, 2010. *On the lot-type design problem.* Optimization Methods & Software, **25(2)**, pp. 217–227.

[26] Gaur V, Osadchiy N & Seshadri S, 2008. *Sales Forecasting with Financial Indicators and Experts' Input.* Johnson School Research Paper Series, **22(5)**, pp. 1056-1076.

[27] Gujarati D, 2003. *Basic Econometrics*, Fourth Edition. The McGraw-Hill Companies, Inc., New York (NY).

[28] Hanke J E, Reitsch A G & Wichern D W, 2001. *Business forecasting.* Prentice Hall, Upper Saddle River (NJ).

[29] Hill R M, 1989. *Allocating warehouse stock in a retail chain.* Journal of the Operational Research Society, **40(11)**, pp. 983–992.

[30] Hill R M, 1992. *Parameter estimation and performance measurement in lost sales inventory systems.* International journal of production economics, **28(2)**, pp. 211–215.

[31] IBM ILOG CPLEX Optimization Studio, 2015. *CPLEX 12.5.* [Online], [cited 30 November 2015], Available at: `http://www-03.ibm.com/software/products/en/ibmilogcpleoptistud`.

[32] James D L, Bruce J W & Michael J E, 1981. *Retailing today.* Second edition. Harcourt Brace Jovanovich (NY).

[33] Kang K Y, 1999. *Development of an assortment-planning model for fashion-sensitive products.* Ph.D. thesis, Virginia Polytechnic Institute and State University, Blacksburg (VA).

[34] Kelton W D & Law A M, 2007. *Simulation modeling and analysis.* Fourth Edition. The McGraw-Hill Companies, Inc., New York (NY).

[35] Kiessling M, Kreisel T, Kurz S & Rambau J, 2014. *Evaluation of a new supply strategy based on stochastic programming for a fashion discounter.* arXiv preprint arXiv:1401.6394.

[36] Koenker R, 1981. *A note on studentizing a test for heteroscedasticity.* Journal of Econometrics, **17(1)**, pp. 107–112.

[37] Kurz S, Rambau J, Schlüchtermann J & Wolf R, 2008. *The top-dog index: A new measurement for the demand consistency of the size distribution in pre-pack orders for a fashion discounter with many small branches.* [Online], [cited 17 October 2013], Available at: `http://arxiv.org/pdf/0804.1412v1.pdf`.

[38] Lam S, Vandenbosch M & Pearce M, 1998. *Retail sales force scheduling based on store traffic forecasting.* Journal of Retailing, **74(1)**, pp. 61–88.

[39] Lariviere M A & Porteus E L, 1999. *Stalking information: Bayesian inventory management with unobserved lost sales.* Management Science, **45(3)**, pp. 346–363.

[40] Lau H S & Lau A H L, 1996. *Estimating the demand distributions of single-period items having frequent stockouts.* European Journal of Operational Research, **92(2)**, pp. 254–265.

[41] Lewison D, 1997. *Retailing.* Prentice Hall, Upper Saddle River (NJ).

[42] Lin K Y, 2006. *Dynamic pricing with real-time demand learning.* European Journal of Operational Research, **174(1)**, pp. 522–538.

[43] Lummus R R & Vokurka R J, 1999. *Defining supply chain management: a historical perspective and practical guidelines.* Industrial Management & Data Systems, **99(1)**, pp. 11–17.

[44] Mantrala M K, Levy M, Kahn B E, Fox E J, Gaidarev P, Dankworth B & Shah D, 2009. *Why is assortment planning so difficult for retailers? A framework and research agenda.* Journal of Retailing, **85(1)**, pp. 71–83.

[45] McCarthy T M, Davis D F, Golicic S L & Mentzer J T, 2006. *The evolution of sales forecasting management: a 20-year longitudinal study of forecasting practices.* Journal of Forecasting, **25(5)**, pp. 303–324.

[46] McGavin E J, Schwarz L B & Ward J E, 1993. *Two-interval inventory-allocation policies in a one-warehouse N-identical-retailer distribution system.* Management Science, **39(9)**, pp. 1092–1107.

[47] McGraw K O & Wong S P, 1996. *Forming inferences about some intraclass correlation coefficients.* Psychological methods, **1(1)**, p. 30.

[48] Mosegaard K & Tarantola A, 1995. *Monte Carlo sampling of solutions to inverse problems*. J. geophys. Res, **100(B7)**, pp. 12431–12447.

[49] Nahmias S, 1994. *Demand estimation in lost sales inventory systems*. Naval Research Logistics, **41(6)**, pp. 739–758.

[50] Ntungo C & Boyd M, 1998. *Commodity futures trading performance using neural network models versus ARIMA models*. Journal of Futures Markets, **18(8)**, pp. 965–983.

[51] Pep Stores Official Website, 2013. *Pep*. [Online], [cited 4 June 2013], Available at: http://www.pepkor.co.za/pep.html.

[52] Pep Stores Official Website, 2013. *Pep facts*. [Online], [cited 11 June 2013], Available at: http://www.pepstores.com/who-we-are/pep-facts.html.

@articlepeterson2002retailing, title=Retailing in the 21st century: reflections and prologue to research, author=Peterson, Robert A and Balasubramanian, Sridhar, journal=Journal of retailing, volume=78, number=1, pages=9–16, year=2002, publisher=Elsevier

[53] Peterson R A & Balasubramanian S, 2002. *Retailing in the 21st century: reflections and prologue to research*. Journal of retailing, **78(1)**, pp.9–16.

[54] Prybutok V R, Yi J & Mitchell D, 2000. *Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations*. European Journal of Operational Research, **122(1)**, pp. 31–40.

[55] Python Programming Language, 2015. *Python 3.3*. [Online], [cited 13 January 2015], Available at: https://www.python.org/.

[56] Quinn F J, 1997. *What's the Buzz?* Logistics Management, **36(2)**, pp. 43–7.

[57] Rajaram K, 2001. *Assortment planning in fashion retailing: methodology, application and analysis*. European Journal of Operational Research, **129(1)**, pp. 186–208.

[58] Robb D J, 1993. *Procurement of prescribed sizes: A Markovian approach*. European Journal of Operational Research, **71(1)**, pp. 110–119.

[59] Robinson S, 1997. *Simulation model verification and validation: increasing the users' confidence*. In *Proceedings of the 29th conference on Winter simulation*, pp. 53–59. Institute of Electrical and Electronics Engineers Computer Society.

[60] Rogers D, 1992. *A review of sales forecasting models most commonly applied in retail site evaluation*. International Journal of Retail & Distribution Management, **20(4)**.

[61] SAS, 2015. *Sas 9.3*. [Online], [cited 23 November 2015], Available at: http://www.sas.com/.

[62] Schmidt J W & Taylor R E, 1970. *Simulation and analysis of industrial systems*. Richard D.Irwin, Inc, Homewood (IL).

[63] Scott M, 2009. *Optimisation of a Single-aisle Picking Line*. Honours project, Stellenbosch University, Stellenbosch, South Africa.

[64] Shrout P E & Fleiss J L, 1979. *Intraclass correlations: uses in assessing rater reliability*. Psychological bulletin, **86(2)**, p. 420.

[65] Silver E A & Kelle P, 1989. *Stocking of prescribed sizes of an item*. Journal of the Operational Research Society, **40(8)**, pp. 719–727.

[66] STEFANESCU C, 2009. *Multivariate customer demand: modeling and estimation from censored sales.* [Online], [cited 14 December 2015], Available at: `http://ssrn.com/abstract=1334353`.

[67] SUN Z L, CHOI T M, AU K F & YU Y, 2008. *Sales forecasting using extreme learning machine with applications in fashion retailing.* Decision Support Systems, **46(1)**, pp. 411–419.

[68] TECH TERMS, 2013. *Definition of SKU.* [Online], [cited 23 October 2013], Available at: `http://www.techterms.com/definition/sku`.

[69] THOM E W, VISAGIE S E & MATTHEWS J, 2014. *Aanpassing van grootte-mengsels tydens voorraadtoewysing in 'n kettingwinkel.* Litnet Akademies: Natuurwetenskappe, **11(3)**, pp. 1–25.

[70] TOLEDO T, KOUTSOPOULOS H, DAVOL A, BEN-AKIVA M, BURGHOUT W, ANDRÉASSON I, JOHANSSON T & LUNDIN C, 2003. *Calibration and validation of microscopic traffic simulation tools: Stockholm case study.* Transportation Research Record: Journal of the Transportation Research Board, (1831), pp. 65–75.

[71] VAN ECK L, 2013. *Personal communication.* Contactable at: `liezlv@pepstores.com`.

[72] WECKER W E, 1978. *Predicting demand from sales data in the presence of stockouts.* Management Science, **24(10)**, pp. 1043–1054.

[73] WILSON J H, KEATING B & JOHN GALT SOLUTIONS, INC., 2009. *Business Forecasting with ForecastX^{TM}.* Sixth Edition. The McGraw-Hill Companies, Inc., New York (NY).

[74] WINSTON W, 2004. *Operations Research.* Fourth Edition. Brooks/Cole Cengage Learning, Belmont (TN).

[75] ZARA OFFICIAL WEBSITE, 2013. *Zara.* [Online], [cited 19 October 2013], Available at: `http://www.zara.com`.