A Review of Integrated Analysis of Production-Distribution Systems

Ana Maria Sarmiento, Rakesh Nagi
Department of Industrial Engineering, 342 Bell Hall
State University of New York at Buffalo, Buffalo, NY 14260

Abstract

This paper reviews recent work on integrated analysis of production-distribution systems, and identifies important areas where further research is needed. By integrated analysis we understand analysis performed on models that integrate decisions of different production and distribution functions for a simultaneous optimization. We review work that explicitly considers the transportation system in the analysis, since we are interested in the following questions: (i) How have logistics aspects been included in the integrated analysis? and (ii) What competitive advantages, if any, have been obtained from the integration of the distribution function to other production functions within a company and among different companies? In our review we also mention whether the work has been done at the strategic level, i.e. if it concerns the design of the distribution system, or at the tactical level, i.e. if it concerns optimization problems for which the characteristics of the distribution system are provided.

Keywords: Production, Distribution, Inventory, Routing, Logistics.

*Email: as1@eng.buffalo.edu
†To whom correspondence should be addressed. Email: nagi@eng.buffalo.edu
1 Introduction

The characteristics of today’s competitive environment, such as the speed with which products are designed, manufactured and distributed, as well as the need for higher efficiency and lower operational costs, are forcing companies to continuously search for ways to improve their operations. Optimization models and algorithms, decision support systems and computerized analysis tools are examples of approaches taken by companies in an attempt to improve their operational performance and remain competitive under the threat of increasing competition.

Recently, a new approach to the analysis of production and distribution operations has been identified, which has proven to be of significant relevance to companies that have adopted it. This approach is based on the integration of decisions of different functions (e.g. supply process, distribution, inventory management, production planning, facilities location, etc.) into a single optimization model. The problem of simultaneously considering the characteristics and requirements of different functions to perform an overall optimization has attracted the attention of researchers in recent years and some models have been proposed in this direction. The basic idea behind these models is to simultaneously optimize decision variables of different functions that have traditionally been optimized sequentially, in the sense that the optimized output of one stage becomes the input to the other (first setting inventory levels and then scheduling distribution, for instance). However, a unified body of literature that deals comprehensively with these types of integrated analyses does not exist yet.

The objective of this paper is to review existing literature on integrated analyses of production and distribution functions and to identify areas where further research is needed. We wish to focus on models that consider the transportation system explicitly since our main interest is to concentrate on the following points: (i) How have logistics aspects been included in the integrated analysis? and (ii) What competitive advantages have been obtained from the integration of the distribution function to other production functions within a company and among different companies? We are also interested in identifying work that has been done at the strategic level, i.e. concerning the design of the distribution system, and work done at the tactical level, i.e. concerning optimization problems for which the characteristics of the distribution system are provided. Considering the relevance of logistics’ costs in the overall operational costs, we believe that by focusing on the study of the relationship between distribution and other functions in a company, new fruitful opportunities can be identified. Given the focus of our paper, we do not present the review of the work done on integrated analyses which do not consider the transportation system explicitly, see for instance Pyke and Cohen (1990, 1993, 1994) and Zipkin (1986).
The remainder of the paper is organized as follows. Section 2 presents the review of work done on integrated analyses of production, distribution and inventory planning. The Inventory/Routing Problem is presented in Section 3 along with a comparison of this problem to the work reviewed in Section 2. Section 4 presents directions for further research and our conclusions are given in Section 5.

2 Integrated Analysis of Production, Distribution and Inventory Planning

In a fairly recent review paper, Bhatnagar et al. (1993) address the issue of coordination in organizations. They identify two levels on it, coordination between functions, which they call the General Coordination problem, and coordination within the same function at different echelons in an organization, called the Multi-Plant Coordination problem. The focus of their work is on the Multi-Plant Coordination problem, whereas ours is on the General Coordination one. In spite of the focus of their work, the authors present a good categorization and some literature review for the general coordination problem. Within this problem they distinguish three categories that represent the integration of decision making pertaining to: (1) supply and production planning, (2) production and distribution planning, and (3) inventory and distribution planning. Thomas and Griffin (1996) present a review on the coordination of functions in these three areas and list some topics for future research. We adopt the last two categories in our work since they most typically consider the distribution function. The principal difference between our work and that of Thomas and Griffin (1996) is that we focus on research that explicitly considers the transportation system.

It is not an easy task to classify the existing work on integrated analysis, mainly because of two reasons: (1) there are a wide variety of assumptions and considerations that can be made when proposing models for analysis, and (2) the literature on the field is not extensive and a unified body does not exist. A classification based on the objective function cannot be done because most, if not all, models consider the minimization of costs as their objective function. In this work we attempt a classification based on the type of decisions to be taken in the model, e.g., production, distribution, inventory management, etc., and on the number of locations per echelon in the model. We also differentiate the work for which an expedited transportation mode is included from that which only includes a regular transportation mode.

The problem presented by the analysis of inventory-production-distribution systems is so complex that optimal solutions are very hard to obtain. Within this problem, different considerations and levels of analyses have been proposed along with heuristic solution approaches. Cohen and Lee
(1988) present a strategic model structure and a hierarchical decomposition approach. The scope of their work is to analyze interactions between functions in a complete supply chain network. To model these interactions they consider four submodules where each represents a part of the overall supply chain: (1) material control, (2) production control, (3) finished goods stockpile, and (4) distribution network control. Stochastic considerations are incorporated in the submodules and relevant costs for set-up, inventory holding and shortage are considered. In the hierarchical decomposition, each submodule is optimized in a given sequence, subject to some service level target defined for that submodule, and the output of a submodule solution is used as the input data to all other subproblems.

The purpose of the framework is to predict the impact on performance, of alternative manufacturing strategies, and to develop an analytically based methodology to answer the following questions: (1) how can production and distribution control policies be coordinated to achieve synergies in performance, and (2) how do service level requirements for material input, work-in-process and finished goods availability affect costs, lead times and flexibility? This framework represents an important piece in the analysis of interactions in the supply chain, since it accounts for the linkages in performance measures between the four functions considered.

From a much more simplified perspective, Mak and Wong (1995) propose the use of a genetic algorithm to solve the inventory-production-distribution problem. Their model consists of three echelons composed of several suppliers, one manufacturing plant and several retailers respectively. Their interest is to simultaneously obtain optimal stock levels, production quantities and transportation quantities so as to minimize total system costs. These costs are inventory holding, shortage, manufacturing and transportation costs. They formulate the problem as an integer program, but a number of simplifying assumptions, restricts its applicability in practical situations. In specific, the model assumes delivery costs known and fixed for every period, direct shipments between all locations and weight limits for transporting products and materials between every pair of locations, in every period.

Cohen and Lee (1989) develop a model that supports resource deployment decisions in a global manufacturing and distribution network. The decisions considered involve both, the design of the international network and the management of material flow within the network. The network considers raw materials suppliers, manufacturing plants, distribution channels, warehousing locations and customers’ geographical dispersion. The problem is formulated as a mixed integer, non-linear program whose objective is to maximize after-tax profits in all countries in which the firm operates. Costs considered are: variable and fixed for procurement, production, distribution and transportation as well as tariffs, duties and transfer pricing. The model is a useful tool for the evaluation of global
manufacturing strategy alternatives. For an extensive review on strategic production-distribution models in a global supply environment, the reader is referred to the work of Vidal and Goetschalckx (1997).

Substantial savings have been achieved by companies that applied an integrated analysis to their operations and developed decision support tools that accounted for this integration. Three such cases are presented by Blumenfeld et al. (1987), King and Love (1980) and Martin et al. (1993).

Blumenfeld et al. (1987) developed a decision support tool for the analysis of the logistics operations at General Motors, that identified a logistics cost savings opportunity of $2.9 Million per year. While keeping the analysis and models as simple as possible, the authors developed a tool that allowed the Delco Electronics Division to examine the impact on total corporate cost due to different shipping strategies for its products. The authors recognized that the minimization of total network cost required the simultaneous determination of optimal routes and shipment sizes and they focussed on analyzing the trade-offs between inventory and transportation costs. A model for the analysis of these trade-offs was included in the decision support tool along with a solution technique to determine the minimum cost for the network under consideration. Results obtained from the research done at each stage of the project are reported in a series of papers (Blumenfeld et al. 1985a, 1985b; Burns et al. 1985).

King and Love (1980) present a case study of a system implemented by Kelly-Springfield, a major tire manufacturer with four factories and nine major distribution centers located throughout the United States. The system coordinates sales forecast, inventory control, production planning and distribution planning. The rapid proliferation of products and the characteristics of the economy at the time the system was developed (inflation, energy shortages, cutbacks in customer spending, record interest rates, etc.) forced the company to improve the efficiency of its operations in order to remain in the market place, in a time when tire manufacturing plants were reporting losses or closing operations. The system is composed by four submodules (production, inventory control, distribution and forecasting), each of which obtains input information from its preceding stage and processes this information to deliver an optimal output to its succeeding stage. The use of feedback loops between production and inventory control and between distribution and inventory control assures the interaction between these functions. However, the optimization of the parameters of interest (reorder points, lot sizes, shipment sizes, etc.) is not done simultaneously but sequentially by the different functions. The implementation of this system resulted in an increase in customer service level and a reduction in inventory levels which represented substantial savings for the company.

Martin et al. (1993) present the development of a system called FLAGPOL (FLAt Glass Products
Optimization modeL) for the Flat Glass Products Group of the Libbey-Owens-Ford Glass company. FLAGPOL is a linear programming model that includes decision variables for production (levels of production), inventory and distribution (to customers and interplant shipments). At the time FLAGPOL was developed, the Flat Glass Products group (FGP) consisted of four manufacturing plants that served approximately 300 customers and produced above 200 product types. The group of people involved in the goal setting for the model included corporate staff members from finance, marketing, management information systems, materials management, transportation, production planning and representatives from the plants, such as managers of production scheduling and cost analysts. One of the goals of the model was to optimize production, inventory and distribution in the multi-plant system based on a 12-month planning horizon. The model was originally conceived to be of tactical and operational scope, however, the implementation of the model proved it to be very valuable to strategic decisions as well.

By providing the ability to plan on a system-wide basis rather than by plants in isolation, as was previously done, the use of FLAGPOL has resulted in substantial benefits to Libbey-Owens-Ford. The authors report annual savings estimated at over $ 2.0 Million. Some of the sources of these savings are the realignments in the assignment of customers to plants, the justification of the investment in rail-car capacity for interplant shipments and the modification in production schedules.

Figures 1, 2 and 3 present block representations of the classification done for models that consider two echelons in their analyses. The dashed blocks represent the echelons in the model and the decisions to be taken at each echelon are represented by the solid line blocks contained in them. The following subsections review the work done for each of the problems presented in Figures 1, 2 and 3 respectively. In each subsection we present separately the models that have considered a finite horizon in their analysis and those that consider an infinite horizon/steady state.

2.1 Distribution-Inventory

This section addresses the models that consider warehousing/distribution as the first echelon and retailer or end customer as the second echelon. In the following, more than one location in either echelon are presented first and the models that consider single locations at each echelon are presented thereafter.

2.1.1 Single Supply and Multiple Demand Locations

The problem presented by a system with one depot (warehouse) that supplies multiple geographically scattered customers (retailers) in the context of integrated analysis has been analyzed by
Federgruen and Zipkin (1984), Federgruen et al. (1986), Burns et al. (1985), Anily and Federgruen (1990), Viswanathan and Mathur (1997) and Chandra (1993). The problem is quite recent and not much literature is available on it. This problem is sometimes confused by researchers with the Inventory/Routing Problem which is presented in Section 3. A discussion of our view of the differences between these two problems is given at the end of that section. The “one-depot, multiple-retailers” problem is a tactical problem that considers a depot that allocates a product (or products) to several retailers in such a way that overall costs are minimized. These costs generally include holding and shortage at retailers (shortage costs in cases that consider stochastic demand at retailers, Federgruen and Zipkin, 1984; Federgruen et al., 1986) and transportation costs which can include fixed and variable costs. The decision variables of interest in the problem are shipment sizes and delivery routes. Given that the formulation of the problem is NP-hard, heuristic solutions have been developed. The general approach taken to the solution of the problem is to analyze the case of direct shipments between depot and retailers and use this analysis as a base for the development of algorithms for the case when delivery routes are to be determined as well. With the exception of Anily and Federgruen and Viswanathan and Mathur which only present the routing case, all other authors present both cases: (i) direct shipments, in which case the decision variables of interest are the shipment sizes only, and (ii) shipments through routes, in which case delivery routes are to be determined along with shipment sizes.

**Finite Horizon Models**

Federgruen and Zipkin (1984) consider a one-period problem in which the amount of product at the depot is limited. Their work was one of the first (if not the first) to integrate the problems of product allocation and vehicle routing into a single model. They propose a heuristic solution based on the decomposition of the main problem into a non-linear inventory allocation subproblem and a number of Traveling Salesman subproblems (one for each vehicle considered). Federgruen et al. (1986) extend this work to the case in which the product considered is perishable. A perishable product is one that has a determined usage life span, after which it has to be discarded at a given cost. The problem is also a one-period problem, and the product in the system is classified into two age classes, “old” units which are the ones that will perish in the present period and “fresh” units which are those that are at least one period away from their perish dates. To reduce the number of out-of-date units, a number of distribution policies can be considered in practice. The model by Federgruen et al. (1986) accounts for at least the following distribution policies: (1) A rotation
policy that removes all “still usable” product from the individual locations’ inventories at the end of every period and returns it to the depot for redistribution, together with the fresh quantity, (2) A retention policy that maintains product received by each location at that location until it is used or outdated, and (3) A combination of (1) and (2). The problem of obtaining shipment sizes and delivery routes is an extension of the work by Federgruen and Zipkin (1984) to two product classes in the system. The solution approach used is the same as in Federgruen and Zipkin (1984) with the variation that the inventory allocation subproblem accounts for two product classes. To solve this sub-problem the authors use a Lagrangean dualization approach. Their work accounts for different cost parameters at different locations.

Both of these papers present a comparison of the integrated to the sequential approach. Federgruen and Zipkin (1984) show that about 6-7% savings can be achieved by using the combined approach and Federgruen et al. (1986) show that travel costs are substantially less using the combined approach and find that for most instances of the problem considered, the delivery requirements can be met with one vehicle less than those required by the sequential approach.

**Infinite Horizon/Steady State Models**

Burns et al. (1985) consider an infinite horizon problem and develop an analytical method to minimize overall costs. They are interested in comparing the cost performance of the two distribution strategies: direct shipping and peddling. Rather than considering specific locations for each customer in a detailed network, the authors consider the density of customers in a given region, and find optimal region sizes as well.

Inventory costs are included in the objective function in an aggregate form, i.e. the cost of holding inventory at the depot, on transit and at the retailers is obtained by approximating the time that the product spends in the system and multiplying it by an interest factor. For direct shipping they obtain an EOQ type of solution by trading off inventory and transportation costs. If the shipment size obtained from the solution is greater than the capacity of the truck (which is given) then a full-truck load is scheduled. For the peddling strategy a full-truck load is the optimal shipment size. The analysis presented in this paper provides guidelines for the distribution problem rather than precise answers to it, given the number of simplifying assumptions and heuristic derivations.

Anily and Federgruen (1990) consider a problem very similar in structure to that of Burns et al. (1985). They derive upper and lower bounds for the system-wide long-run average costs. They show that under weak probabilistic conditions these bounds are asymptotically optimal. They develop
a solution procedure whose computational requirements grow roughly linearly with the number of locations considered.

Viswanathan and Mathur (1997) consider the same problem as Anily and Federgruen (1990) with the generalization of multiple products in the system. They develop a heuristic based on a joint replenishment problem to obtain a stationary nested joint replenishment policy (SNJRP). They consider vehicles with limited capacity and present computational results comparing the performance of the proposed heuristic with the heuristics proposed by Anily and Federgruen (1990), for the case of a single product. Their results show that the SNJRP policy performs better in the majority of cases in terms of cost. The authors report that no other heuristic was known to handle multiple products, therefore a comparison was not possible for problems that considered more than one product in the system.

2.1.2 Single Supply and Single Demand Locations

Infinite Horizon/Steady State Models

Blumenfeld et al. (1985b) analyze the trade-offs between safety stock at the second echelon and expediting costs at the first one. Expediting is a function of the order-up-to level and cycle length; for this reason, their analysis focuses on these two variables. Expedited shipments are considered to have a zero traveling time and to contain enough material to last until the next ordering epoch. These assumptions are taken to simplify the model so that the trade-offs can be better visualized. The work considers an unconstrained regular transportation fleet and there is only one trip per cycle.

No explicit allocation of inventory costs is done, so following the approach of Burns et al. (1985) and Blumenfeld et al. (1985a) the average time that the inventory spends in the system (at first echelon, in transit and at second echelon) is calculated and an interest rate is applied to this quantity. The authors calculate the probability that an order needs to be expedited, which includes the uncertainty due to fluctuations in material consumption and in travel time. Their objective is to find the optimal inventory level and cycle length that will trade-off expediting.

In Just-In-Time environments, frequent, small shipments are usually required between suppliers and manufacturers and many times transportation contracts for specific volumes are made to assure proper supply. In such situations emergency shipments are contracted by suppliers whenever customer’s demand presents a sudden increase and a higher amount of product is required at the customer’s location. The problem faced by the logistic manager is to decide the number of vehicles required to make the shipments in such a way that a balance between spare capacity in contracted ve-
hicles and use of emergency shipments is achieved. Yano and Gerchak (1989) analyze this problem and present a solution methodology to simultaneously determine safety stock level at the location in the second echelon (customer), number of vehicles required for regular delivery and time between shipments in such a way that overall operational costs are minimized. These costs include inventory holding and shortage costs at the customer’s location and emergency and regular transportation costs. The integrated analysis showed that the time between shipments is smaller than the value obtained from the appropriate EOQ computation.

In our opinion, two important contributions of this work are the following. First, it deals with the transportation system at a strategic level, i.e. it determines the optimal size for the regular fleet and establishes the time between shipments. Second, it shows that full-truck loads and one-vehicle per shipment are not always optimal, this is due to the fact that depending on demand variability it can be more profitable to have spare space in the regular vehicles, even if more than one truck is needed per trip.

Originally motivated by the study of an IBM operation in Italy, Speranza and Ukovich (1994) develop a model to simultaneously minimize inventory and transportation costs in a system that considers shipments of multiple products between two locations. In our opinion, two important contributions of this work are the following. First, the system is analyzed under the consideration that shipments can only take place at discrete epochs in time. Second, it explicitly considers individual products for allocation into trucks as opposed to considering an aggregation of the multiple products.

Using a small example, the authors show that the results obtained by Blumenfeld et al. (1985a) are not applicable for the case in which discrete shipping times are considered and that transportation costs increase considerably (not off-setting the reduction in inventory costs) when the solution obtained is rounded up to obtain integrality.

As in Yano and Gerchak (1989) they find that shipping less-than-full truck loads may lead to improved policies. They consider the following two cases for shipping frequencies: (1) a product is assigned a single frequency, and (2) a product can be assigned to more than one frequency (order splitting). For consolidation, they consider the following two options: (1) only products shipped at the same frequency may share the same truck, and (2) all products whose shipments happen to be simultaneous may share the same truck. The combination of frequency assignment and consolidation strategy results in four problems that are analyzed by the authors. The distribution system is also viewed from a strategic point of view; the problem determines the number of trucks to be used and allocates different products to trucks.
2.2 Inventory-Distribution-Inventory

Finite Horizon Models

Chandra (1993) analyzes the “one-depot, multiple-retailers” problem with two new considerations. The product ordering cost at the depot is included in the model, and customers face dynamic demand. He solves the problem with the use of an iterative approximate algorithm, which starts with an initial feasible solution to the sequential problem. It then evaluates how the depot ordering decisions are affected if the delivery schedules for customers are changed. The change that leads to the greatest reduction in overall costs is adopted and the process is repeated until there is no further gain from coordinating the two decisions. The algorithm estimates the benefits of integration in terms of cost reduction over the case when the depot and customer decisions are made independently. The author presents an experimental study which investigates the effect of coordination between depot’s ordering policy and its distribution schedules. Results show that costs savings from the integrated approach over the sequential one range between 3 to 11%.

Infinite Horizon/Steady State Models

Ernst and Pyke (1993) extend the model of Yano and Gerchak (1989) to include the consideration of the inventory level at the first echelon. They analyze two forms of transportation costs, linear and concave forms. The interest of this work is to simultaneously obtain the base stock levels at both echelons, the optimal vehicle capacity and shipping frequency. The two papers consider emergency shipments to be unconstrained in capacity and availability and assume a close coordination between the two echelons.

2.3 Production-Inventory-Distribution-Inventory

Finite Horizon Models

Chandra and Fisher (1994) combine the production scheduling and vehicle routing problems to investigate the value of coordination between these functions with the use of a computational study. To our knowledge this is the first work that analyzes the integration of these functions. They consider a two-level system with one manufacturing plant which has a finished goods stockpile in the first echelon, and several retailers in the second echelon.

An analysis of different scenarios in the system is conducted. These scenarios are obtained by the variation of the following parameters: length of planning horizon, number of products and
number of retailers, and set-up, inventory holding and vehicle travel costs. The authors compare
the integrated approach to the sequential optimization of the problem and find that a reduction in
operational costs can be achieved, ranging between 3-20%, if the integrated analysis is preferred.

A very important observation made by the authors is that, depending on the system parameters,
the analysis of coordinated functions may or may not be worth the effort of integrated analysis. This
observation was made earlier by Benjamin (1989), but the authors also present conditions under
which coordinating efforts are most beneficial and show that under the right conditions the value of
coordination can be extremely high.

Haq et al. (1991) consider a three echelon system with one production facility, several warehouses
and several retailers. A multi-stage model is used for the production facility. They formulate the
integrated problem as a mixed integer program whose objective is to determine the production and
distribution batch sizes and the inventory levels at all the locations, in such a way that total system
cost is minimized. This total cost considers production, set-up and recycling costs at production
stages, distribution cost, which is consider as a linear function, inventory holding costs at all echelons
and backlogging costs at retailers.

Ishii et al. (1988) also consider a three echelon system with one manufacturer, one wholesaler
and one retailer respectively. They assume a Pull system in which products have short life cycles
since product models change frequently in the market place. Their objective is to minimize situations
of dead stock (obsolete products) and shortages. They consider the presence of two types of products
in the system, new products and products that are in the final stage of their life cycle. The variables
of interest are the transportation ordering levels, the stock levels and the production ordering level
that minimize the dead stock inventory at the retailer. They propose an algorithm for the solution of
the problem and present a numerical example to illustrate the algorithm.

Infinite Horizon/Steady State Models

Blumenfeld et al. (1985a) and Benjamin (1989) consider several locations on each echelon
for the integrated analysis of production, inventory and transportation from a tactical perspective
and present formulations of deterministic models. Blumenfeld et al. are interested in analyzing
the existing trade-offs between transportation, inventory holding and production set-up costs in the
network. The authors analyze the cases of direct shipping between nodes in the echelons, shipping
through a consolidation terminal and a combination of both, and obtain shipment sizes that trade-off
these costs. They present an interesting approach to the analysis of more complex networks. The
approach is based on the subdivision of the original network into subnetworks and on the application of results obtained for the cases of one supply point shipping to one demand point and one supply point shipping to many demand points.

Blumenfeld et al. are not concerned with the specification of the distribution system (capacity and number of vehicles) but on obtaining the value of the shipment size (cycle length) that trades-off the respective costs. It does not consider explicitly the different inventory costs (at plants, in transit and at retailers) but combines them into a single aggregate one. A number of simplifying assumptions are made in the model (e.g. deterministic demand, specified vehicle capacity) however, these simplifications are justified by the objective of the paper which is to provide insight into the trade-offs existing among the considered costs.

Benjamin (1989) considers the simultaneous optimization of the production lot size problem, the transportation problem and the economic order quantity problem. He accounts for supply constraints and explicitly considers inventory costs; his interest is to find optimal production sizes for supply points and order quantities for demand points. The model assumes an unconstrained transportation system and direct shipments between each node, therefore, no routing decisions have to be taken. Although this work considers multiple products, no product to truck allocation decisions are made. Benjamin presents a comparison between the simultaneous and the separate optimization approaches and finds that the magnitude of the advantage of optimizing the problem simultaneously, depends on the relative size of setup and holding costs at each of the supply and demand points. This is an interesting remark since it highlights the fact that simultaneous optimization is not always better.

Chien (1993) analyzes the case of direct shipments between a single supply location and a single demand point. Demand for the product follows a known probability distribution. The transportation cost is fixed per shipment and the truck has limited capacity. The objective is to maximize the expected profit of the operation by determining a joint optimal production rate and shipment size. Expression for production costs, per-unit inventory carrying costs (at plant and at retailer), transportation costs, shortage penalty cost, regular revenue and salvage revenue are obtained as functions of the demand density. The problem is solved by an iterative procedure which yields solutions that are within 0.2-3.8% of optimality in terms of expected profits.

3 Inventory/Routing Problem

The Inventory/Routing Problem (IRP) typically considers a distribution firm which operates a central depot that supplies a large set of geographically scattered customers. Products distributed are usually industrial gases, heating oil, fuel or other products that can be stored in containers. Customers have
containers of a given capacity to store the product which they consume on a daily basis. The central depot is responsible for maintaining an adequate level of supply in the customer’s tanks and most of the times is liable for shortages. The distribution operates as a Push-system since the distribution firm decides the delivery schedule entirely by itself based traditionally, on customer’s demand estimates and recently on actual inventory levels, provided by electronic data exchange systems between customers and the central depot. Emergency shipments are scheduled whenever shortages occur to bring the on-hand inventory at customer’s location to an acceptable level, however, given the higher cost of emergency shipments with respect to regular ones, shortages are sought to be avoided whenever possible.

The problem faced by the management of the depot is to efficiently construct routes for the distribution vehicles that minimize operational costs in such a way that shortages are avoided while unnecessary delivers are not scheduled. That is, customers with high on-hand inventory levels should not be visited since in many cases, the firm is paid for the amount of product delivered. Considering that customers face a stochastic demand pattern, safety stock levels have to be calculated for each customer, so that replenishment decisions can be taken on time. The IRP is a large and complex logistic problem that combines a temporal element (the time at which replenishments are done) with a spatial element (the routing of vehicles). The time of replenishment to each customer is determined by the depot based on a customer selection process. The selection of customers determines the requirements for the routing of vehicles. Given that customer selection and vehicle routing are interdependent issues, at an operational level, the delivery system requires the following type of decisions: (i) Determination of customers to be included for delivery, (ii) Assignment of delivery vehicles to customers, and (iii) Vehicle routing and scheduling. Some of the parameters that can be taken into account explicitly in an IRP formulation are for example, customer’s inventory holding capacity, product depletion statistics, re-order rules, and inventory and stockout costs, as well as transportation costs.

Previous to the use of electronic data interchange technologies (EDI), most Inventory/Routing Problems considered customer demands as random variables which became known at the time the delivery vehicle arrived at the customer’s location. This implied that accurate forecasting methods should be used to determine the product level at every customer’s tank to efficiently determine which customers to visit. It also implied that the amount of product contained in the delivery truck might not be sufficient to complete the planned route. The wide spread use of data exchange technologies, however, will eliminate the need to forecast customer’s demand. The information about the amount of inventory at each customer’s location will be available to the decision maker at the time routes
are planned for vehicles, this fact will make the determination of the customers to be delivered in a
given day and the amount of product to be delivered a much easier job than it has been in the past,
by simplifying the Inventory/Routing problem greatly.

The Inventory/Routing Problem considers that shortage costs are incurred by the distribution firm
while holding costs (whatever their form) are the responsibility of the customer. Larson (1988) dis-
tinguishes two types of Inventory/Routing Problems, Strategic Inventory/Routing Problems (SIRP)
and Tactical Inventory/Routing Problems (TIRP). SIRP focuses on estimating the minimum size
(or cost) vehicle fleet required by the firm, over the long term, to serve its customers when only
the probability distribution for the per unit demand is known for each customer. TIRP deals with
routing an existing vehicle fleet to supply customers over the short term, whose actual demands for
replenishment can be estimated. According to Webb and Larson (1995), a major difference between
the strategic and tactical versions of the problem is that in solving the SIRP all possible realizations
of the TIRP must, at least implicitly, be considered simultaneously.

Bell et al. (1983) describe the development of a decision support system for the IRP at Air
Products and Chemicals Inc. This work was awarded the TIMS Practice Prize for 1983. The
system consists of several modules that include customer usage forecasting, a distance-network
and a shortest path algorithm, a mathematical optimization module to produce delivery routes and
an interface for possible manual modification of schedules and of operational parameters. The
optimization module uses a sophisticated Lagrangian relaxation algorithm to solve mixed integer
programs with up to 800,000 variables and 200,000 constraints to near optimality. The benefits
obtained from the implementation of the system include a significant reduction in operating costs
(over $2 million annually), an increase in vehicle productivity and a higher utilization of the firm’s
computer network.

Golden et al. (1984) developed a heuristic for the optimization of an integrated delivery planning
system for a large energy-products company that distributes liquid propane. The study was done
for a distribution district that serves approximately 3,000 customers. The purpose of the study was
to compare the distribution rule used by the company to the heuristic algorithm proposed by the
authors. The company’s rule for distribution was based on the re-supply point for each customer.
Based on historical data the company calculated an average consumption rate for each customer and
kept a record of the last replenishment date for each customer. This information was used to calculate
when each customer would hit the re-supply point and the next replenishment was scheduled based
on this information. The authors developed a heuristic that includes a customer selection algorithm
that is able to select the set of customers to be visited on each day in a cost-effective manner. The
heuristic also accounts for vehicle routing and trip to truck assignment. Each of these components was solved as effectively as possible and a simulation experiment was used to evaluate the integrated performance of the components. Results from the simulated comparison of the proposed heuristic to the distribution rule used by the company showed that the heuristic had a superior performance. The number of gallons/hour delivered was improved by 8.4%, the number of stockouts was reduced by 50% and total costs were reduced by 23%.

Trudeau and Dror (1992) develop an algorithmic procedure to solve the stochastic IRP which takes explicitly into account route direction, costs of stockouts, route failures and their interrelations. This study permits a better understanding of the interdependent factors that impact the efficiency of the delivery system. Many of the simplifying assumptions made in past studies are not made in this work, for instance, the authors do not consider the demand rate to be the expected value of the demand distribution for each customer as an estimated consumption rate, also, they do not rely on the artificial capacity device to account for route failures as done in the solution of the standard Vehicle Routing Problem (VRP) with route failures. The authors recognize the fact that a short term myopic approach has the tendency to defer as many deliveries as possible to later planning periods. To solve this problem, they include a temporal “cost” of selecting the customers for service and assigning them to specific days so that the long-time horizon objective is properly projected into the weekly tactical planning. A simulation experiment is presented to illustrate the interrelation between stockouts and route failures. Since no tractable exact solution methodology exists, the authors compare their results with the industry standards, showing that their procedure is far superior to industry practices.

Dror and Ball (1987) also recognize the fact that a short-term optimization approach has the tendency to postpone as many deliveries as possible to later periods, and present a procedure to convert the long-term problem into a single-period problem that can be solved with the use of standard routing algorithms. Their objective is to minimize annual costs subject to no customer shortages. The reduction procedure considers the definition of single-period costs that reflect long-term costs, the definition of safety stock level and a specification of the customer subset to be considered during a single period.

Considering a set of vehicle routes for the IRP, Dror and Levy (1986) develop a heuristic route improvement scheme that is capable of examining and operating in all the routes simultaneously. The scheme is based on a node interchange operation. They solve the annual IRP based on a sequence of consecutive weekly solutions. Two sets of customers are considered, those who need replenishment done in the present planning period and those who do not. A feasible solution to the
problem is obtained which (i) includes the scheduling of all those customers who require service in the planning period, and (ii) satisfies vehicle capacity constraints. After this feasible solution is found, an improvement scheme is applied to it to improve the quality of the solution. This scheme is designed to examine the given feasible solution and to search for favorable trade-offs when interchanging customers’ positions on a route and between different routes. The goal of the improvement stage is to reduce the total distribution cost for the planning period while maintaining route capacity constraints. Three routing improvement procedures are presented along with results of the comparison performed via computational experiments.

Chien et al. (1989) analyze the IRP when a limited amount of product is available at the depot. They assume that the entire demand of customers need not be satisfied at the time customers are visited by the delivery truck. Their objective is to maximize total profit, which is defined as the revenue obtained from units delivered minus fixed and variable routing costs and possible shortage costs incurred. They formulate the problem as a mixed integer program and develop a solution approach based on Lagrangean relaxation. Computational experience reported by the authors shows that the procedure is able to generate good quality solutions for several instances of the problem which are generated with a variety of revenue and cost structures, and vehicle availability and vehicle capacity combinations.

Larson (1988) presents the characteristics of an SIRP that was implemented into a decision support system developed to assist strategic decisions taken by the City of New York in the design of its new sludge transport and disposal system. The key decision outcome of the model is the least cost fleet size and fleet mix. While the system considers the assignment of customers to specific clusters, a basic assumption is that all replenishments are made on a single route visiting all customers in a cluster. This assumption caused some locations to be visited more frequently than required. Motivated by this inefficiency of the system, Webb and Larson (1995) consider a period/phase approach to solve the SIRP. They present the development of routing solutions based on the period and phase of replenishment of each customer, and develop a simple model for the tactical routing problems the fleet will eventually encounter. Estimates of the fleet size required are developed on the basis of these routing solutions. The period/phase approach can be generalized to take long-term operating costs into account.

Other work related to IRP is presented in Dror (1983), Dror et al. (1986) and Assad et al. (1982). Inasmuch as IRP modeling is relatively new, it is not possible to say that a general standard formulation exists yet. Further research is needed to characterize the optimal solution methodology of different problem instances and to reduce the computational complexity of the problems, specially
for large ones. As noted by Golden and Assad (1986) research is needed to identify the appropriate strategy for incorporating long-run inventory-related costs and short-term routing costs into the same model. Also, it is necessary to include electronic data interchange considerations between customers and the central depot, in the light of EDI advancements and other technologies alike. Further work on strategic IRP is also needed, Larson (1988) identifies several areas for further research in this respect. He proposes the consideration of transshipment points, their number and optimal locations, and also, the consideration of strategic decisions other than fleet size and mix (e.g., customer inventory holding capacities). Ball (1988) describes application environments in which Inventory/Routing problems arise, presents formulations for several versions of the problem and reviews some solution procedures. He also identifies areas where further research is needed.

We believe that there is a significant difference between Inventory/Routing Problems and the integrated analysis of the “one-depot, multiple-retailers” problem, presented in Section 2. The aim of the work presented in Section 2 is to simultaneously optimize problems that have traditionally been treated sequentially, and to coordinate the performance of functions at different echelons to achieve a “global” optimization of the system. The benefits obtained from this integrated effort are shared among the functions involved in the optimization process. The Inventory/Routing problem, on the other hand, looks for the optimization of the depot’s operation only, not accounting for the cost performance of its customers. Although this problem integrates inventory replenishment considerations to the vehicle routing problem into a single model, the routing decisions depend on the selection of customers, and the selection of customers in turn depends on the amount of product a vehicle (or vehicles) can deliver. This interdependency between customer selection and vehicle routes is the core of the IRP and is not a problem that has been solved sequentially in the past. Practitioners have accounted for this interrelation in one way or another when deciding how to allocate product and how to route their vehicles. The main difference between these two problems therefore, lies in the formulation of the objective function. The “one-depot, multiple-retailers” problem includes the retailer’s cost in the minimization function while the IRP looks only at the depot’s costs.

4 Directions for Further Research

A recent study done by the Global Logistics Research Team at Michigan State University (1995), on the logistics aspects of several world class companies, shows that the management of these enterprises recognize the fact that the ability to visualize and develop cooperative relationships with other firms throughout the supply chain is critical to world class logistics performance. The study also
shows that leading organizations are developing increased unification among supply chain partners. This increased unification has several implications. Information is more freely shared between partners in the supply chain and between functions within the same company. This facilitates the development of models for a global analysis of operational performance. Managers in the physical distribution system can take advantage of this new information-exchange capability to obtain (i) a broader vision of the logistics system and of its relation to other production functions, and (ii) an enhanced understanding of the impact that decisions taken in the logistics system have on the performance of other functions within the system, such as inventory control, production planning, etc. This enlarged vision of the distribution system can also assist managers in the development of performance improvement programs that could render significant benefits to the organization. However, in spite of the awareness and willingness for cooperative efforts, there seems to be a lack of tools and methods that can help managers in the analysis of the integrated systems (of which their companies are part of) in this emerging collaborative environment.

The models reviewed in this paper represent a significant advance in the integrated analysis of production-distribution systems. However, research in this area is still relatively fragmented, showing many gaps that further research must fill. The full potential of the integrated analysis has not been completely explored yet.

As mentioned before, we differentiate two broad areas in the integrated analysis of production-distribution systems research: (i) Inventory/Distribution and (ii) Production/Distribution. In our categorization we consider the integration of two functions at a time only, because the majority of the research has focussed on this direction.

The objective of Inventory/Distribution problems has commonly been the minimization of total cost, which includes inventory costs (holding and shortage) at both, supply and demand points and transportation costs. The decision variables in these models are usually shipment sizes, inventory levels and optimal routes when applicable. Some models do not consider routing of vehicles but direct shipments between locations, while others explicitly account for vehicle routing. Most of the models assume an unconstrained transportation system.

Further research in the Inventory/Distribution problem could take into consideration more complex networks for analysis. Algorithms that explicitly consider the location of several customers and depots, as well as the routing of vehicles and inventory levels setting are needed for more complete dynamic scenarios. While optimal solutions are very hard to obtain, heuristic procedures could be developed to obtain approximate solutions for this complex problem. Validation methods would as well be required. Research is also needed for the explicit consideration of multiple products in
the system, work in the line of Speranza and Ukovich (1994) that combines the product allocation into trucks to the Inventory/Distribution problem would constitute a significant contribution to the field. The analysis of different instances of the Inventory/Distribution problem under stochastic demand considerations is still a largely open research area. Further research is also needed for Inventory/Distribution scenarios that acknowledge the existence of emergency shipments, and consider a constrained transportation system.

As stated by Ballou (1992), the integration of inventory control and physical distribution brings much closer ties to the production/distribution function in many firms, such that in the future we may see production and logistics merging much closer in concept and practice. Production and distribution can be completely decoupled if there is a sufficiently large inventory between them, however, the trend towards reduced inventory levels will cause the two functions to have a closer interaction. As companies move towards this reduction in inventory levels, models for the Production/Distribution problem will become more important. Further, due to the fact that to this date very little work exists in this direction (Chandra and Fisher 1994), the analysis of Production/Distribution systems still remains as a widely open area for research. Some directions for future research are, the determination of the minimum safety stock required between the two functions to guarantee a reliable operation of the system in terms of a desired service level, the analytical formulation and solution of the problem, and in general, the study of the interrelation of the production and distribution functions.

Very few researchers have analytically approached the problem presented by the integration of more than two functions. The benefits and difficulties that these kinds of problems present remain as open questions. A methodology to balance the complexity of this kind of problems with the applicability of the results obtained from them could be vital in the development of solution methodologies that remain tractable and at the same time suitable for practical application. Analytical formulation of problems that consider more than two functions and exact or approximate solution procedures are still needed.

Three aspects in the distribution system that have to be more widely considered in the integrated analysis are the consideration of non-linear functions to represent transportation costs, the presence of common carriers to perform regular or emergency shipments, and the variability in traveling time. While linear functions for the transportation costs make the analysis less complex, they most often, do not represent reality, thus the consideration of non-linear functions (often concave functions for the distribution problem) should be more widely included in research. The presence of common carriers provides an additional flexibility that should be accounted for in the tactical and strategic integrated models. Also, the investigation of alternative transportation modes is needed. Most of the
models reviewed, consider the transportation time in the distribution system to be fixed, models that include variability considerations in traveling time could be of use to investigate the suboptimality in solutions due to the time variability effect.

The review of the work done on integrated analysis makes evident that the consideration of two or more functions and their interrelations into a single model makes the optimization problem much harder to solve than the previous disjoint optimization problems. Comprehensive mathematical programming formulations which incorporate simultaneously all aspects of the integrated problems will, in general, contain too many variables and constraints to be easily solved using exact algorithms. Researchers should exercise increased creativity in the analysis of the integrated models and in the development of heuristic procedures capable of handling the bigger challenge of integrated analyses. As pointed out in the work by Benjamin (1989) and Chandra and Fisher (1994), the solution obtained from the optimization of the integrated problem is not always more attractive than the sequential optimization of the problem of each function. Depending on system parameters and characteristics, the benefits obtained from an integrated analysis might not offset the increased complexity of the problem and the greater solution effort. More research is needed to identify overall frameworks for which integrated analyses are beneficial and compensate the increased complexity of the problem. To undertake the optimization of integrated problems, researchers can make use of the increasing capabilities of computer technology. The investigation of ways to reduce the computational complexity of the problems, especially of large problems, could be of significant impact in the solution procedures. Simulation models that permit user interaction could also be of interest.

The integration of production and distribution brings forth issues pertaining to the scope that each part (production and distribution) will have. For instance, the definition of the set of decisions that each model will be allowed to make. For example, should the transportation time of goods be a parameter in the production model prescribed by the transportation model or should it be a variable? Another important issue to consider is the time grid for both functions. It is most likely that there will be a mismatch in terms of time unit between the production and distribution models. This issue must be resolved in order for the decision variables of the two models to interact.

The centralized optimization of integrated systems is based on the premise that information is more openly shared between functions on a timely fashion. Research to characterize information systems capable of supporting this interaction is needed. Technological advances, like Electronic Data Interchange (EDI) and Satellite-Location devices could play an important role on the configuration of the information system. Relevant issues about the use of these technologies in relation to
integrated analyses have to be identified and studied.

The integrated analysis of production-distribution systems has proven to be of significant benefit to companies that have applied it under adequate conditions. Substantial savings and efficiency improvement have been some of the results that the analysis of logistics integrated to other production functions have granted to some companies. Given the number of assumptions and constraints that the analysis of integrated functions can account for, a wide variety of problems and models can arise. The identification of the relevant instances of the integrated problem and their solutions, is a task that still needs to be undertaken. The solutions to these problems and their possible implementation into decision support systems can provide companies with valuable tools to gain competitive advantage as they move into higher collaborative and competitive environments.

5 Conclusions

In this paper we have reviewed recent work in the area of integrated analysis on systems that explicitly consider logistics integrated to other functions in production, and we have identified areas where further research is needed. A clear classification of the Inventory/Distribution and Production/Distribution problems is hard to make, given the diversity and number of assumptions that such problems can take into consideration. The Inventory/Routing problem, on the other hand, is somehow a better defined problem which has received increased attention in recent years.

The survey of research done on integrated analysis shows that, in some cases, the integration of the logistics function into the analysis of previously isolated production functions (e.g., inventory control, facilities location and production planning) has the potential of providing significant benefits to companies, in the form of costs savings and efficiency improvement. However, many aspects of the integrated analysis have not been covered yet, in particular, the characterization of systems for which integrated analyses are most beneficial. We believe that, given the relevance of logistics costs in overall operational costs, the integrated analysis of production/distribution systems can provide significant competitive advantage to companies that adopt it.

Acknowledgements

The authors gratefully acknowledge the valuable comments and remarks of two anonymous referees. Rakesh Nagi acknowledges the support of National Science Foundation under grant DMI-962409. Ana Maria acknowledges the Society Of Logistics Engineers (SOLE) for the doctoral dissertation proposal award of 1998.
References


**Figure 1: Classification of Distribution-Inventory Models**

In this figure, the classification of Distribution-Inventory models is depicted, showing how different supply and demand locations can be modeled with varying routing and transportation modes. The figure includes a breakdown of models based on the number of supply and demand locations and their corresponding decisions. Smaller details are labeled using **Further Research**, which indicates areas for additional study.

**Figure 2: Classification of Inventory-Distribution-Inventory Models**

This figure extends the previous classification by adding an additional layer that includes production decisions. It shows how production plants, wholesalers, and retailers interact within the inventory system. Similarly, **Further Research** markers are included to highlight potential areas for deeper exploration.

**Figure 3: Classification of Production-Inventory-Distribution-Inventory Models**

The final figure combines production, inventory, and distribution decisions into a comprehensive model. It illustrates how different elements within a supply chain can be modeled, with **Further Research** annotations pointing to areas needing further investigation.