Decision Support Systems 64 (2014) 43–56

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

An expanded database structure for a class of multi-period, stochastic mathematical programming models for process industries

Narain Gupta [a,*], Goutam Dutta [b], Robert Fourer [c]

[a] *Management Development Institute, Mehrauli Road, Sukhrali, Gurgaon 122007, India*
[b] *Indian Institute of Management, Ahmedabad 380015, India*
[c] *AMPL Optimization, USA*

ARTICLE INFO

Article history:
Received 28 May 2013
Received in revised form 9 March 2014
Accepted 18 April 2014
Available online 30 April 2014

Keywords:
Decision support system
Process industries
Optimization
Stochastic programming (SLP)
Database structure
Management science

ABSTRACT

We introduce a multiple scenario, multiple period, optimization-based decision support system (DSS) for strategic planning in a process industry. The DSS is based on a two stage stochastic linear program (SLP) with recourse for strategic planning. The model can be used with little or no knowledge of Management Sciences. The model maximizes the expected contribution (to profit), subject to constraints of material balance, facility capacity, facility input, facility output, inventory balance constraints, and additional constraints for non-anticipativity. We describe the database structure for a SLP based DSS in contrast to the deterministic linear programming (LP) based DSS. In the second part of this paper, we compare a completely relational database structure with a hierarchical one using multiple criteria. We demonstrate that by using completely relational databases, the efficiency of model generation can be improved by 60% compared to hierarchical databases.

1. Introduction and motivation

We introduce a user friendly, model data independent, model solver independent, stochastic optimization based DSS for strategic planning in a process industry. This research is an extension of an earlier work by Dutta [12], and Dutta & Fourer [14,24] where a multi-period optimization based DSS was developed for process industries. Fourer [24], in his seminal work, showed that the fundamental principles of relational database construction could be used to represent a linear program. This work was carried out for a single period deterministic optimization. Dutta [12] and Dutta & Fourer [13,14] extended the research of single period planning to multiple period planning. These applications ranged from a steel company in India [17,19], a steel company in North America [12,13], to a pharmaceutical company in Western India [15] and even further to an aluminum company [16] in Eastern India. The DSS customized for the integrated steel plant in North-America, demonstrated a potential impact of 16–17% increase in the bottom line of the company [13].

In the first part of the paper, we discuss the design and development of a multiple period, multiple scenario DSS. While several researchers [4,22,27,28] have done work in the application of stochastic optimization and a set of researchers [11,18,20] has worked on the need for user friendly DSS, this is probably the first attempt that tries to integrate these two concepts. Here, we attempt to address the following seven questions in detail:

1. How is the database structure of a SLP model different from a deterministic LP model?
2. How are the diagnostic rules in the database structure of a multi-scenario DSS and model different from a multi-period deterministic model?
3. What are the key features of a multi-period, multi-scenario optimization based DSS?
4. How are multi-dimensional data values reported, with scenarios as an additional dimension?
5. In what way can the optimal results be represented in a multi-period, multi-scenario optimization based DSS?
   In the second part of this paper, we compare the performance of hierarchical and relational structures on a deterministic DSS. A single period optimization based DSS by Fourer [24] and a multiple period optimization based DSS by Dutta and Fourer [12,14] compare hierarchical and relational databases. In this paper, we develop a DSS with a completely relational structure and compare the performance with Dutta & Fourer's [14] partially relational database structure. We ask two additional questions that were not addressed in the earlier research.
6. What is the performance of the relational database structure compared to the hierarchical database structure in a multi-period planning model?

* Corresponding author. Tel.: +91 124 4560000, +91 9818037324.
E-mail addresses: narain.gupta@mdi.ac.in (N. Gupta), goutam@iimahd.ernet.in (G. Dutta), 4er@ampl.com (R. Fourer).

http://dx.doi.org/10.1016/j.dss.2014.04.003
0167-9236/

7. What are the issues related to database structure design for linear programming when we have a choice of different variations of relational and hierarchical databases?

This research is further extended to design and develop a completely relational database structure for a generic, multi-period, multi-scenario, optimization-based DSS. We study the reasoning behind the increased difficulty in the design and implementation of the DSS and discuss how to expand the DSS, to include multiple scenarios.

1.1. Outline of the paper

The paper introduces database structure and SLP in Section 1. In Section 2, we review the literature on database optimization interface, SLP, and its application in industries. In Section 3, we address how the database structure of an SLP model is different compared to the deterministic model (question 1 of Section 1). Section 4 addresses the question of how the diagnostic rules of multi-period multi-scenario DSS are different from the rules of a multi-period single scenario DSS. We also address the question about the additional rules implemented in the DSS (question 2 of Section 1). This section also describes the key steps of optimization, and the important features of the DSS, followed by issues related to data storage, retrieval, loading, and update (questions 3 and 5 of Section 1). In Section 5, we address the basic features of the DSS with respect to data reporting, updating, and creating multi-dimensional included layouts (question 4 of Section 1). In Section 6 of the paper, we address the sixth and seventh questions about how a completely relational database structure performs compared to a partially relational and partially hierarchical one. We conclude the paper in Section 7 with the future scope of the research, and the list of references. The Appendix A describes the mathematical formulation of the SLP.

2. Literature review

An LP can be represented in several ways depending on the need of the user and the system. Earlier work [11,25] summarizes the common methods of LP representation schemes in practice. The LP representation is approached in the literature as *Structured Modeling* [1–3], *Graph-Grammar* [5–7], *Netform* [10], and *Block Schematic Diagram* [20]. The matrix generator (MG) form of an LP representation, which is a translation form, has been attempted by several authors. For instance, DATAFORM (Ketron Inc., 1975), and UIMP [18] are common matrix generators. Research on mathematical optimization [25] recognizes that it is difficult to develop an LP representation form which can be commonly understood by modelers, computers, and practitioners.

To manage the sales, considering production, sales, and foreign exchange, a modeling environment has been proposed by Combos et al. [8] in their paper. This modeling environment was based on data warehousing and knowledge discovery in databases. Rudberg & Thulin [21] described how the Advance Planning System (APS) can act as an enabler in the adaptation of logistics and supply chain principles. The APS described by them is a decision support system for supply chain planners. They used case study findings to present the power of decision support. Recent research [23] uses a simulation modeling approach in sales and operations planning for decision support. Valente [28] presented a simulation and optimization based decision support system to address uncertainty in the model parameters. They addressed the different complexities involved in modeling uncertainty. Most of the existing studies talk about supply chain planning and decision making, using decision support tools.

A comparative study [11] of LP representation schemes such as MG, algebraic languages, and block-schematic languages, reports that very few LP representation schemes have been implemented in commercial software systems. Other approaches are still at the prototype stage of development. A study on relational databases [24] in the context of an LP formulation visualizes the subset of the Cartesian product of sets of the LP as a relation in the mathematical sense.

From the database viewpoint, the structure of an LP can be represented by two sets of entities, variables and constraints, along with a many-to-many relationship that records which variables are associated with which constraint. To represent non-linear models, stochastic models, and other types of models, the model schema can be expanded. The tables for the coefficients store only the non-zero elements in the array representation. Thus, the representation conforms closely to the MPS format used for input by most of the optimizers. Fourer [24] emphasizes how the development of a database structure with a direct relation to the variables and the constraints of a large scale mathematical programming can lead to a user friendly DSS.

Fourer [24] describes the algebraic formulation of a single period deterministic model and the corresponding database structure. Further extensions by Dutta and Fourer [12–15] present a multi period deterministic model and a hierarchical database structure, and compare it to a partially relational database structure. In this research, we compare and contrast the design and performance of a completely relational data structure and a partially relational and partially hierarchical data structure of a multi period deterministic planning model. This research also emphasizes how the uncertainty in the parameters can be captured in a multi scenario planning model. We design and develop a completely relational database structure for the multi scenario planning model and discuss the increase in design and computational complexity which has been caused due to the inclusion of uncertainty in the parameters. In this research, we also address uncertainty in the model parameters, and extend the multi-period optimization model [12] to develop and implement a two stage SLP in a DSS. The study, being generic, is capable of modeling uncertainty in any parameter in the optimization model.

The size of the optimization model increases significantly with the increase in the number of uncertain parameters in the model. The efficiency of optimization model generation and the performance of the database structure primarily depend on the size of the model. Considering this, we realize the need to study a completely relational database structure in the context of an SLP. A review of literature also reveals that not much work has been published on relational databases for an SLP which can demonstrate the implementation of a user-friendly DSS. The research primarily focuses on the design and development of a completely relational database structure in the context of a class of multiple scenario mathematical optimization models, and compares it to the performance of a hierarchical database structure.

2.1. Two-stage stochastic linear program (SLP) with recourse

The two stages of the stochastic program are defined by a set of decisions made in those stages. The decisions made in the first stage are the decisions which are implemented before the realization of the randomness in the system. The second stage decisions are the ones which are implemented after the realization of the randomness. The decisions made in the first stage are non-anticipative in nature, and do not depend on the outcome of the randomness. The focus of the SLP is to rectify the decision taken for the first stage, well in advance, such that the solution remains the same regardless of the outcome of the random realization. To simplify the understanding of the two-stage SLP with recourse, we discuss an example, which is the first SLP with recourse, formulated by Ferguson & Dantzig [4]. The term recourse is defined by Fragniere [9] as the decision variables adapting to the different outcomes of the random parameters at each time period. In a stochastic program with recourse, the response of the randomness of the model is corrected as a part of the model. We introduce SLP using the deterministic equivalent linear program developed by Ferguson & Dantzig [4]. It is a generalized two-stage program:

c_1 The cost vector of the first stage
c_2 The cost vector of the second stage