

# **A decision support system for forecasting and optimal procurement**

Neven Miculnic

April 28, 2016

---

## ABSTRACT

---

Optimal procurement in most industries involve forecasting of two quantities: prices of raw materials and customer's demand. The aim of this work is to integrate forecasts into production planning models, with the aim of minimizing overall procurement, holding and production costs under demand satisfaction constraints. The decision support system should allow the decision maker to integrate qualitative, unstructured information through a simple interface, for scenario selection and solution refinement

---

## PROBLEM DEFINITION

---

### 1.1 INTRODUCTION

This problem is variation of lot sizing problem in stochastic setting. Variation is as follows. We have stochastic prediction for raw material supply cost  $\mathbf{s}$  at future time moment  $t$ . The *factory* converts all bought raw materials at time  $t$  and stores them into a warehouse. Number of discrete time moments is denoted by  $n$ .

One time moment storage costs fixed amount  $h$ , that is the holding cost.

Likewise demand for the product is equally stochastic, denoted as  $\mathbf{d}$ . In case there's no products in storage to satisfy demand we pay backlogging cost denoted as  $b$ .

Our aim is to optimize our procurement policy by varying  $\mathbf{x}$ , that is product amount we buy at time moment  $t$ . However we are constrained by the maximum raw materials we can buy each day by  $\mathbf{x}_{\max}$ .

### 1.2 FORMAL PROBLEM DEFINITION

#### 1.2.1 *Definitions*

Every input value is known beforehand, and it's assumed for random variables their distributions are known. Following is deterministic problem variant, and in subsequent chapters randomness and uncertainty is embedded into problem. Font convention:

$x, y, z$	variables and constants
$\mathbf{x}, \mathbf{y}, \mathbf{z}$	vector variables and constants
$X, Y, Z$	random variables
$\mathbf{X}, \mathbf{Y}, \mathbf{Z}$	random vectors
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}$	probability distributions
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	matrices, by context differentiated from random vectors

Now following is formal problem description for deterministic variant:

$\mathbf{s} = [s_1, s_2, \dots, s_n]^\top$	Supply cost vector
$\mathbf{x} = [x_1, x_2, \dots, x_n]^\top$	Procurement quantity vector
$\mathbf{x}_{\max}$	Procurement quantity limits vector
$\mathbf{d} = [d_1, d_2, \dots, d_n]^\top$	Demand random vector, $Y_t$ is a random variable
$b$	backlogging cost
$h$	holding cost

### 1.2.2 Variables

$\mathbf{x}$  is our decision variable, as described previously.

### 1.2.3 Constraints

$$\begin{aligned}\mathbf{x} &\in \mathbb{N}_0^n \\ \mathbf{x} &\leq \mathbf{x}_{\max} \\ \sum x &= \sum d\end{aligned}$$

## 1.3 OBJECTIVE FUNCTION

First let's define auxiliary matrix:

**Definition 1.3.1.**  $\mathbf{C}$  matrix defines cost for satisfying demand with specific raw supply material purchase date. It's element  $c_{ij}$  equals:

$$c_{ij} = \begin{cases} b(i-j) + s_i & j < i \\ h(j-i) + s_i & j \geq i \end{cases}$$

That is using purchases raw materials at  $i$  to satisfy demand at time moment  $j$  incurs cost  $c_{ij}$ .

**Definition 1.3.2.**  $c(t)$  defines total speeding we pay at time  $t$ .

$$c(t) = \sum_i x_i c_{it} \quad (1)$$

**Definition 1.3.3.**  $f(\mathbf{x})$  is objective function for this problem. Our aim is to minimize it.

$$f(\mathbf{x}) = \sum_t c(t) \quad (2)$$

---

## DETERMINISTIC APPROACH

---

### 2.1 SOLUTION

Problem as defined in 1 can be reduced to transportation problem<sup>1</sup> since we have cost matrix  $\mathbf{C}$  defining “transportation” costs associated with each possible assignment option. For successful reduction we only need adding dummy source or destination as described in<sup>2</sup>

### 2.2 VARIANTS

Problem as defined previously could seem rather simplistic and not allowing useful extensions users might want, such as current storage amount and similar. In following few subsections most useful extensions are described.

#### *Starting storage capacity*

in case we have already certain number of product in stock we can easily embed that knowledge into model by adding new supplier/node as new decision variable  $x_0$ . It's maximum,  $x_{\max 0}$  is equal to starting storage capacity, and  $c_{0j} = (j - 1)h$

#### *Ending storage requirement*

For example we'd like to have some extra product in stock by the end of analysis, and it's quite easy to accommodate such requirement. Simple add to  $d_n$  ending storage requirement, thus obtaining new  $d_n$ .

---

<sup>1</sup> Frederick S Hillier. *Introduction to Operations Research*. McGraw-Hill, 2014. ISBN: 0073523453.

<sup>2</sup> Ibid.

### *Allowing future backlogging*

In model as described, time stops at time moment  $n$ , however, in realistic scenario we're looking at only short time snapshot of ongoing process. To allow such future purchases to backlog at previous times, we can extend the model with  $m$  future moments:

$$d_{n+i} = 0$$

$x_{\max n+i} = y_i$  allowable backlogging from future supply purchases

**C** as previously defined, simply extended

**s** extended with future costs

for  $i$  in  $1, 2, \dots, m$ :

---

## DEEPER ANALYSIS

---

First we are going to analyse problem deeply without presuming any independence or probability distribution on random variables. Later in subsequent chapters we are going to focus more on where demand at time  $t$  has independent Gaussian distribution and mean.

### 3.1 COST FUNCTION

The minimizing function is:

$$\mathbf{E} [\mathbf{s}^T \mathbf{x} + \sum c(t)]$$

Due to linearity of expectation and  $x$  being variable it's equal to:

$$\mathbf{x} \mathbf{E} [\mathbf{s}^T] + \mathbf{E} [\sum c(t)]$$

Therefore only needed modeling information for supply cost is its expectation  $\mathbf{E} [\mathbf{s}]$ . The other part is more trickier since  $D_i$  and  $D_j$  aren't usually independent.

### 3.2 HANDLING DEMAND COST NON-LINEARITY

As we can see in equation 1 we have non-linearity depending whether we're satisfying all demand or are we backlogging demand at time  $t$ . Therefore here are two possible solutions for minimizing objective function 2.

#### *Simulation*

We generate multiple scenarios for demand vector,  $d$  according to probability distribution. For small  $n$  and relatively small number of outcomes in each ran-



dom variable we can exhaustedly model each scenario, scale it appropriately and feed to MIP solver<sup>1</sup>

### *Safety net approach*

Alternatively, we can artificially add new constraints and avoiding backlogging with arbitrary probably. This model assumes backlogging cost are significantly greater than storage cost, that is backlogging penalty is severe.

Thus we chose values arbitrary realizations of random variables  $D_t$  and add additional constraints of the form:

$$X_t \geq D_i \forall t \quad (3)$$

which reduces cost function 1 to:

$$c(t) = h(X_i - D_i) \quad X_i \geq D_i$$

and enables are faster solving approaches. Given parameters  $D_i$  chosen and their underlying distribution we derive non-increasing function  $p(t)$  which represent probability of not breaking newly introduced constraints 3

### *Min-cost max flow with uncertainty*

Since deterministic case can be modeled with min-cost max flow, we can investigate further in this directi

### *Stochastic Dynamic Programmnig*

TODO.. maybe, I've got recommendation by a friend, need to read some papers.

---

<sup>1</sup> There's a trick on using binary variable for discontinuity in cost function 1

---

## BIBLIOGRAPHY

---

Hillier, Frederick S. *Introduction to Operations Research*. McGraw-Hill, 2014. ISBN: 0073523453.