The impact of macroeconomic leading indicators for tactical sales forecasting on SKU inventory management

(presented at the 7^{th} IESM Conference, October 11–13, 2017, Saarbrücken, Germany)

Yves R. Sagaert*, Stijn De Vuyst*, Nikolaos Kourentzes† and El-Houssaine Aghezzaf*
*Department of Industrial Management
Ghent University, Technologiepark 903, 9052 BE, Zwijnaarde
Email: yves.sagaert@ugent.be
†Department of Management Science
Lancaster University Management School, LA1 4YW UK, Lancaster

***Abstract*—An accurate sales forecasting has indispensable effects on the supply chain management as this input is essential in the decision making process. Macroeconomic leading indicators can provide early indications of global changing economic dynamics. By including this external information, the global tactical sales forecasting can be improved. This paper wants to quantify the impact on inventory level, where decisions are typically taken on an individual product base. For this, the high-level forecast needs to be disaggregated to the product level. Techniques that make use of the hierarchical structure present can benefit from pooling individual forecasts on different hierarchical levels. We propose an empirical technique to reconcile the forecast distributions of different aggregation levels in a hierarchical structure. We focus on the first and second moment of the forecasting distribution, the mean and variance. We evaluate our proposed method on inventory and service level via inventory simulations.**

I. Introduction

Recent research has shown that macroeconomic leading indicators have the potential to detect early warning signals from economic activity. Furthermore, [13] shows that this information can improve tactical sales forecasting. The tactical level is typically formulated on a global scale, where the economic dynamics are expected as well. The sales forecasts steer the business decision process on global budget, global capacity pooling, as well as on production planning and inventory management. The latter typically happens on an operational level. In the studied case in this research, inventory management is done on production plant level, which are geographically distributed. Furthermore, inventory decisions are taken on Stock-Keeping-Units (SKU) level, as each product or material has an individual stock.

The methodology proposed in [12] avoids to forecast the macroeconomic indicators individually by shifting these indicators in time prior to insert them in the sales forecasting model. As the forecast model uses then the already realised values of the indicators, this would bring less uncertainty to the calculated point forecast. As a consequence, the prediction interval around the point forecast is more narrow on a tactical level. In this research, we investigate whether the formulation of more narrow prediction intervals around the point forecast on global level can be beneficial for the lower levels of forecasting. Our paper aims to link the tactical forecast to its impact on inventory level. This is in line with the literature in sales forecasting that argues for evaluation of forecast models where they are used [2], [9], [17]. As the forecasting models are used for inventory management, the evaluation should also occur on inventory performance. However, inventory performance is largely impacted by the shape of the demand distribution [19]. Here, inventory simulation models are proposed to quantify the impact of different forecasting models [15].

In order to quantify the impact on inventory, it is necessary to address the first and second moment of the forecast distribution. This means that both the mean and the variance of the high-level forecast need to be disaggregated to the lower levels. This is in contrast with traditional techniques for disaggregating, as these focus mainly on the calculated mean of the forecast [5]. Here, we use hierarchical forecasting via reconciling to link all levels of the company structure. This approach was first proposed by [7], who found that reconciliation can improve the forecast accuracy on several levels of the hierarchical structure. Here, all the time series are forecasted independently at all levels of the hierarchy. Next, these forecasts are combined and reconciled using a regression model. This approach allows to combine extrapolations of patterns in historical sales on lower levels with forecasts on highest level that are augmented with exogenous indicators. In this way, information of all levels is addressed.

We propose to reconcile the uncertainties on different hierarchical levels through an empirical method. By reconciling the prediction intervals across different hierarchical levels, we investigate whether more narrow prediction intervals on high level can result in lower stock in inventory management on SKU level. For this, the reconciled prediction intervals on SKU level are used as input for an inventory simulation with real case data.

II. Literature review

There are different strategies to handle time series with several levels of aggregation [10]. The most simple approach ignores the structure completely and generates a direct forecast. [4] gives an overview of direct versus aggregation derived forecasts. He concludes that forecasting via aggregation results in superior forecast performances in a majority of the research, even if an information loss may exist by summing product time series. Forecasting approaches that account for this hierarchical structure can be divided in four categories. First, top-down forecasting calculates a direct forecast at the top level time series and generates lower level forecasts by dividing the top level forecast proportional to the lower level sales volume. Alternatively, the volume of the sales forecast on the lower level can be used to disaggregate the top level forecast. Second, a sales forecast is formulated for each of the lowest level time series and the top level forecast is obtained by simply summing these forecasts. This is known as bottom-up forecasting. There is clear trade-off between these two approaches, noted by [18]. Top series may be easier to forecast, but more univariate patterns can be captured on low level series. Third, a combination of both top-down and bottom-up can be used, sometimes called middle-out. This approach is often used in practice [5]. Here, a direct forecast is calculated on an intermediate level, and higher and lower levels are calculated using aggregation and disaggregation. Fourth, combinatorial forecasting creates a direct forecast at all levels of the hierarchy and combines them. [1] proposes a new approach for this combination process, by reconciling the individually formulated forecasts using a regression model. This combination process allows for to alter the individual forecasts on all levels of the hierarchical structure. Alternatively, the forecasts on all hierarchical contribute to the final forecast. Furthermore, forecast reconciliation is flexible, as it can be used across multiple hierarchies, such as product, geographic or customer hierarchies [14].

Most of the literature focusses on the point forecasts in these hierarchical time series. However, [7] notes the theoretical formulation to formulate prediction intervals of the reconciled forecasts and defers the optimisation of this formula to future work. In this paper, we propose an approach to reconcile the empirical prediction intervals across the hierarchical structure in a similar matter as the reconciliation takes place for the point forecast.

III. Methodology

In our methodology, the exogenous macroeconomic indicators are identified on a high aggregation level of the sales, as we do not expect an individual product to have the same dynamics as the global or national economy. However, the impact of the forecasting models happens on the product level, in terms of the inventory control for this product. Therefore it is essential that the tactical forecasts on high level are linked to the lower levels of the hierarchical structure. Furthermore, the top levels do not exhibit any low level historical pattern, such as product-specific seasonality. This type of information will be captured by low level models. As both levels can add information to the final forecast, hierarchical reconciliation is used as an approach to combine the different hierarchical levels. Next, the performance on the product level is evaluated via production simulation and inventory performance.

A. *Macroeconomic indicators selection and forecasting model*

On a global level, the sales forecast is modelled using the external information of macroeconomic indicators. These are selected using the methodology described by [13] that is based on the Least Absolute Shrinkage and Selection Operator (LASSO) [16]. For the aggregated time series on top level, univariate information such as seasonality and auto-regressive process are combined with external indicators. The relevant indicators are selected in a fully automatic way, simultaneously with their leading effect. This means that the optimal shift in time is determined for each indicator separately. Furthermore, the used framework is set up to formulate unconditional forecasts. By design, the indicators are only used when their leading effect is higher than or equal to the forecast horizon. While this complicates the computation as for each forecast horizon an individual LASSO model needs to be formulated, it makes the application in practice extremely relevant.

The amount of potential indicators p is very large, while the number of historical data point n is typically limited, creating a $p \gg n$ problem. This problem expands even further as for each indicator, the optimal leading effect needs to be addressed as well. [13] shows that LASSO in this context was capable to retain useful and relevant indicators automatically.

The prediction interval around the point forecasts is formulated empirically for the h-step ahead forecasts by $\hat{Y}_h \pm k\hat{\sigma}_h$, where k is the appropriate percentile point of the standard normal distribution. Here, we assume that the error has a Gaussian distribution with zero mean and standard deviation σ. The estimation $\hat{\sigma}_h$ is done via the Root Mean Squared Error (RMSE) on the historical sales. For the univariate models, we re-estimate the RMSE for each forecast horizon, while for the LASSO model a different model is formulated for each forecast horizon already.

B. *Hierarchical reconciliation*

Once the individual forecasts are formulated, these can be combined across the hierarchical structure. Naturally, bottom-up aggregation cannot benefit from the leading indicator forecasts as this approach generates higher time series forecasts by summing up the lower levels. [14] notes that traditional techniques for aggregation often have weak performance. For example, top-down disaggregation yields often good forecasts at higher aggregation levels, but worse forecast performance at lower levels. In this research, we combine the forecasts on different aggregation levels, via hierarchical reconciliation. Using a linear regression on the different individual time series, we obtain the coherent reconciled forecasts for the product hierarchy. The weights for this linear combination can be derived from the structure of the hierarchy, as discussed by [7]. This work argues that the optimal combination is independent of the data, and that the weight of an aggregated forecast is set equal to the amount of lower level series.

C. *Inventory simulation*

The designed simulation allows to simulate the production and inventory build-up, based on the sales forecasts and their prediction intervals. The product forecasts with their respective prediction intervals are an central input in this. These are the reconciled forecasts at the lowest level of the hierarchical structure. Here, we assume all products to be produced via a Make-to-Stock policy. In the inventory simulation is as follows: at the end of each month, the actual realised sales is taken into account, and the remaining stock is reviewed. The inventory is controlled by how much of each product should be produced in the upcoming month that is planned. To approximate the real world inventory management, the simulation enforces a production stand-off of 5 months. This condition makes that the planned amount to produce in the month after the production stand-off is equal to the total sales in the next 6 months, reduced with the production of the next 5 months and the current inventory of that product.

To be more specific, we considering one product time series at time t. Let I_t be the inventory at the end of period t, P_t the amount of items produced in a month t, and $\hat{Y}_{t+h}$ the forecasted demand for horizon $t+h$. The inventory position at time $t+h$ can then be estimated as

$$\hat{I}_{t+h} = I_t + \sum_{i=1}^{h} (P_{t+i} - \hat{Y}_{t+i}) . \tag{1}$$

The demand at $t+h$ will be satisfied entirely whenever

$$I_{t+h-1} + P_{t+h} > Y_{t+h} . \tag{2}$$

If we choose to produce P_{t+h} items at time $t+h$, the probability that the demand in period $t+h$ is not entirely satisfied can be estimated as

$$\hat{\alpha}(P_{t+h}) = \text{Prob}\left[\hat{I}_{t+h-1} + P_{t+h} < \hat{Y}_{t+h}\right] \tag{3}$$

Given the (continuous) distribution function

$$F^{*}_{t+h}(x) = \text{Prob}\left[\sum_{i=1}^{h} \hat{Y}_{t+i} \leq x\right] , \tag{4}$$

of the summed forecasted demand and a certain acceptable shortage probability α_{t+h} for period $t+h$, the production quantity P_{t+h} in that period can be allocated as

$$P_{t+h} = F^{*-1}_{t+h}(1-\alpha_{t+h}) - I_t - \sum_{i=1}^{h-1} P_{t+i} , \tag{5}$$

The production is determined in the simulation via (5). This equation takes into account the inventory on-hand and the allocated production up to $h-1$ (due to the production stand-off). Furthermore, it accounts via (4) for the estimated demand up to period h and the related shortage probability, shown in (3), over the period 1 to h. As a consequence, the simulation allows for a certain acceptable shortage, which is evaluated by the fill rate at product level.

IV. Performance measures

In order to evaluate the effect of the empirical formulated prediction intervals in the hierarchical structure of the sales, we will analyse the forecast results at each stage of our methodology.

We compare the forecasting performance from an accuracy standpoint using the Mean Absolute Percentage Error (MAPE) in (6), as is widely used in practice and relevant to the case company.

$$MAPE_h = \frac{1}{n} \sum_{t=1}^{n} \frac{\left|Y_{t+h} - \hat{Y}_{t+h}\right|}{Y_{t+h}}, \tag{6}$$

where Y_{t+h} is the actual product demand and $\hat{Y}_{t+h}$ is the forecast for time period $t+h$.

The uncertainty around the point forecast is evaluated via $\hat{\sigma}_h$, as this is the key element of the prediction interval. We define $AvgRel\hat{\sigma}_h$ as:

$$AvgRel\hat{\sigma}_h = \sqrt[m]{\prod_{p=1}^{m} \left(\frac{\hat{\sigma}^A_h}{\hat{\sigma}^B_h}\right)}, \tag{7}$$

where $\hat{\sigma}^A_h$ refers to the evaluated forecast via hierarchical reconciliation and $\hat{\sigma}^B_h$ to the benchmark model on low level with purely historical data. In literature, the $\hat{\sigma}_h$ is traditionally obtained by multiplying the estimated demand variance by the length of the forecast horizon. [11] notes that this approach is flawed, as forecast errors for different periods of the forecast horizon are positively correlated. In contrast to the traditional methods, we determine the $\hat{\sigma}_h$ on all levels of the hierarchical structure empirically.

Finally, we display the effect of the proposed method and the univariate benchmark on inventory management via an inventory simulation. To judge how well a forecasting model performs on product j, we use average on-hand inventory $\bar{I}^{+(j)}_{\alpha}$ and the fill rate $\overline{FR}^{(j)}_{\alpha}$ (FR) achieved over the test periods as performance metrics. Here α is a tuning parameter in the decision of the production quantity. The fill rate refers to the fraction of demanded items which can be obtained immediately from stock without backordering. To assess inventory performance, a weighted average is taken across the different products:

$$\bar{I}^{+}_{\alpha} = \sum_{j=1}^{14} w_j \bar{I}^{+(j)} , \qquad \text{and} \qquad \overline{FR}_{\alpha} = \sum_{j=1}^{14} w_j FR^{(j)} , \tag{8}$$

with weights according to the actual demanded volumes per product j

$$w_j = \frac{\sum_{t=n+1}^{n+T} Y^{(j)}_t}{\sum_{j=1}^{m} \sum_{t=n+1}^{n+T} Y^{(j)}_t} . \tag{9}$$

V. Results

The dataset for the company case contains one business unit from a tire manufacturer for the period 2005–2015. This experiment considers three hierarchical levels from the business unit structure. The highest level is the overall sales for the business unit. The intermediate level is the plant level, which consists of 5 manufacturing plants. The lowest level differentiate all the customers that are served from this plant. The customers are Business-to-Business users and typically order only one material from an individual plant. The lowest level consists of 14 different time series of customer–product combination. In our experiment, we have omitted the customers who do not order frequently, as these customers result in intermittent demand patterns.

Figure 1 shows a generic hierarchical structure of three levels. The top level is the aggregated sales of the business unit. The intermediate level are the individual manufacturing plants and the lowest levels are the individual customers that are served from each plant. Customers who order from several plants are shown multiple times on the lowest level, as the

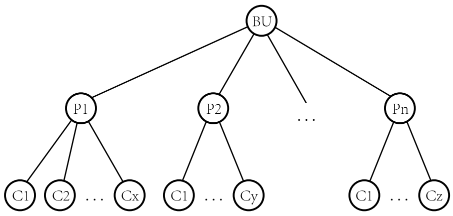

Fig. 1. Hierarchical structure that exists of three levels: Business Unit (BU), Manufacturing Plant (P) and Customer level (C)

time series on this level represent an unique combination of product and customer.

The models are trained in the period 2005:01-2012:12 and their performance is tested in the period 2013:01-2014:12 over a rolling origin experiment. The design of our experiment is as follows: the sales forecasts are generated for the next 6 months to allow to plan production. However, the planning decisions are made with a production stand-off of 5 months. The benchmark on the lowest level is exponential smoothing (ETS) as proposed by [6].

TABLE I. MAPE FOR FORECASTING SETUP

	1	2	3	4	5	6
ETS	31.2	33.2	33.6	33.9	33.9	37.4
Hierarchical	30.2	31.4	31.8	31.7	31.7	35.3

Table I compares the individual forecasts on each aggregation level with the reconciled forecasts. The MAPE for each horizon h is aggregated across forecasts from 13 origins and 14 products to form the MAPE displayed. The hierarchical reconciled forecasting method that makes use of external data outperforms the benchmark technique on the product level. Here, the forecast horizon $h = 6$ is of special interest for the inventory simulation. For this horizon, we see that the reconciled forecast outperforms the benchmark model substantially.

TABLE II. EVALUATING FORECAST UNCERTAINTY VIA $AvgRel\hat{\sigma}_h$

	1	2	3	4	5	6
Hierarchical	1.9	1.7	1.6	1.5	1.4	1.3
ETS	1	1	1	1	1	1

Table II exhibits the $AvgRel\hat{\sigma}_h$ for each forecast horizon, aggregated via the geometric mean across time series and rolling origins. The demand forecast on the top level that is formulated with macroeconomic indicators, has a more narrow prediction interval than the benchmark model ETS. This is intuitive as these formulated forecasts use realised values of the indicators. As a consequence, the empirical reconciled $AvgRel\hat{\sigma}_h$ is expected to decrease via the reconciliation. However, the results in table II suggest the opposite. A crucial point in the hierarchical reconciliation is the weighting scheme in the linear combination of the different aggregation levels. The modelling technique of LASSO differs substantially of the nature benchmark, which has an impact on traditional weighting scheme for hierarchical reconciliation as proposed by [7].

The inventory simulation performs the final evaluation on product-customer level for the different forecasting models. Figure 2 exhibits the simulation results for one specific

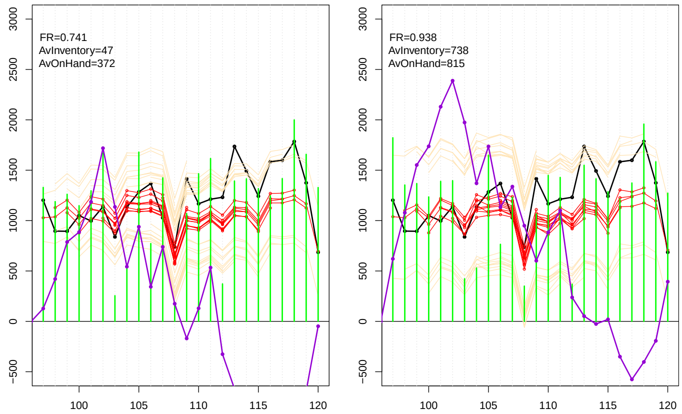

Fig. 2. Production simulation for ETS and LASSO over a rolling origin with the actual demand (black), the point forecasts (red), 20% and 80% quantiles of the forecast (orange), the production quantities (green) and the inventory position (purple). Production stand-off $h = 4$ and parameter $\alpha = 0.05$.

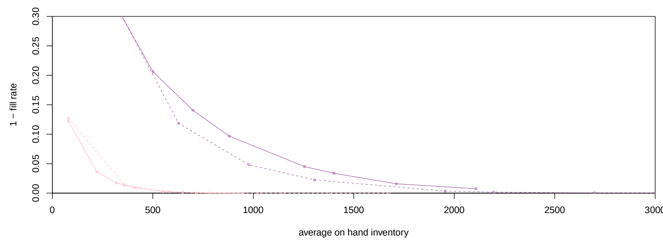

Fig. 3. Results of production simulation for LASSO (solid line) and ETS (dotted) for both $h = 0$ (light) and $h = 4$ (dark). On each line, the parameter α for the production allocation decreases from 0.5 (top left) to 0.001 (bottom right). The vertical axis is $1 - \overline{\text{FR}}_\alpha$, and on the horizontal axis is the average on-hand inventory $\bar{I}^+_\alpha$.

product. In this figure, the black line represents the actual demand. The green bars indicate the allocated production, and the purple line is the inventory position for this product. In red, the consecutive forecasts are shown, with their respective prediction intervals in orange.

Figure 3 shows the resulting performance curves from the inventory simulation. Given a desired fill rate, the average on-hand inventory is lower for the reconciled forecasts on the lowest levels, compared to univariate models. We can see that the hierarchical reconciliation method outperforms the benchmark technique for a given service level of the fill-rate.

VI. CONCLUSION

Incorporating exogenous information in top level sales forecasts can improve these forecasts. In our methodology, we select macroeconomic indicators and their leading effect fully automatic from a large pool of indicators. We found that the forecasting accuracy is improved and that the prediction intervals are more narrow on the top level. However, supply chain decision are often taken on a lower level of aggregation. It is interesting to quantify the difference of this exogenous information on an operational level, such as inventory management. For this, it is essential to link the top level to the lowest individual product level, as inventory stock is kept at this level. Dealing with hierarchical structures in sales forecasting is described in literature [5], [7], but is mainly focussed on the point forecast. For inventory decision, the uncertainty around this point forecast (prediction intervals) is equally important. In this paper, we propose a new empirical approach to reconcile prediction intervals in a hierarchical structure.

We evaluate our proposed approach on each stage of our methodology. The results of the reconciled forecasts and prediction intervals show the following: (i) the forecast accuracy on the product level improves by reconciling the different hierarchical levels. (ii) while we expect the prediction intervals to become more narrow by reconciling across different hierarchical levels this is not the case, and (iii) The inventory simulation exhibits lower stock levels for the same fill rate. Whereas the LASSO prediction intervals on the top level produce more narrow prediction intervals, this improvement does not translate to lower levels of aggregation. A key aspect here is that the weighting of reconciliation occurs via the MSE on the train sample. As the LASSO technique has a different cost function, it is questionable if this weighting scheme is appropriate. Investigating other weighting schemes for this purpose is put as further research.

REFERENCES

[1] G. Athanasopoulos, R. A. Ahmed and R. J. Hyndman, *Hierarchical forecasts for Australian domestic tourism*, International Journal of Forecasting, Volume 25, Issue 1, 2009, Pages 146-166

[2] M. Babai, M. Ali, J. Boylan, and A. Syntetos, *Forecasting and inventory performance in a two-stage supply chain with ARIMA(0,1,1) demand: Theory and empirical analysis*, International Journal of Production Economics, Volume 143, Issue 2, 2013, Pages 463–471

[3] A. Chen, J. Blue, *Performance analysis of demand planning approaches for aggregating, forecasting and disaggregating interrelated demands*, International Journal of Production Economics, Volume 128, Issue 2, 2010, Pages 586-602

[4] G. Fliedner, *An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation*, Computers & Operations Research, Volume 26, Issue 10, 1999, Pages 1133–1149

[5] G. Fliedner, *Hierarchical forecasting: issues and use guidelines*, Industrial Management & Data Systems, Volume 101, Issue 1, 2001, Pages 5–12

[6] R. Hyndman, A. B. Koehler, J. K. Ord and R. D. Snyder, *Forecasting with exponential smoothing: the state space approach*, Springer Science & Business Media, 2008

[7] R. J. Hyndman, R. A. Ahmed, G. Athanasopoulos, and H. L. Shang, *Optimal combination forecasts for hierarchical time series*, Computational Statistics & Data Analysis, Volume 55, Issue 9, 2011, Pages 2579–2589

[8] A. Kerkknen, J. Korpela and J. Huiskonen, *Demand forecasting errors in industrial context: Measurement and impacts*, International Journal of Production Economics, Volume 118, Issue 1, March 2009, Pages 43–48, ISSN 0925-5273

[9] N. Kourentzes, *Intermittent demand forecasts with neural networks*, International Journal of Production Economics, Volume 143, Issue 1, 2013, Pages 198–206

[10] S. Moon, C. Hicks and A. Simpson, *The development of a hierarchical forecasting method for predicting spare parts demand in the South Korean Navy - A case study*, International Journal of Production Economics, Volume 140, Issue 2, 2012, Pages 794–802

[11] D. Prak, R. Teunter, A. Syntetos, *On the calculation of safety stocks when demand is forecasted*, European Journal of Operational Research, Volume 256, Issue 2, 2017, Pages 454–461

[12] Y. R. Sagaert, E.-H. Aghezzaf, N. Kourentzes and B. Desmet, *Tactical sales forecasting using a very large set of macroeconomic indicators*, European Journal of Operational Research, 2017

[13] Y. R. Sagaert, E.-H. Aghezzaf, N. Kourentzes and B. Desmet, *Temporal Big Data for Tactical Sales Forecasting in the Tire Industry*, Interfaces, 2017

[14] A. A. Syntetos, Z. Babai, J. E. Boylan, S. Kolassa, K. Nikolopoulos, *Supply chain forecasting: Theory, practice, their gap and the future*, European Journal of Operational Research, Volume 252, Issue 1, 2016, Pages 1-26

[15] L. Tiacci and S. Saetta, *An approach to evaluate the impact of interaction between demand forecasting method and stock control policy on the inventory system performances*, International Journal of Production Economics, Volume 118, Issue 1, March 2009, Pages 63–71

[16] R. Tibshirani, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society Series B (Methodological), 1996, Pages 267–288

[17] X. Wang and F. Petropoulos, *To select or to combine? The inventory performance of model and expert forecasts*, International Journal of Production Research, Volume 54, Issue 17, 2016, Pages 5271–5282

[18] G. Zotteri, M. Kalchschmidt and F. Caniato, *The impact of aggregation level on forecasting performance*, International Journal of Production Economics, Volume 93, 2005, Pages 479-491

[19] G. Zotteri *The impact of distributions of uncertain lumpy demand on inventories*, Production Planning & Control, Volume 11, Issue 1, 2000, Pages 32–43