Production Planning & Control
The Management of Operations
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713737146

Rolling horizon heuristics for production planning and set-up scheduling with backlogs and error-prone demand forecasts

Alistair R. Clark [a]
[a] University of the West of England, Faculty of Computing, Engineering and Mathematical Sciences. Bristol BS16 1QY. UK

To cite this Article: Alistair R. Clark , 'Rolling horizon heuristics for production planning and set-up scheduling with backlogs and error-prone demand forecasts', Production Planning & Control, 16:1, 81 - 97
To link to this article: DOI: 10.1080/09537280412331286565
URL: http://dx.doi.org/10.1080/09537280412331286565

Production Planning & Control,
Vol. 16, No. 1, January 2005, 81–97

Rolling horizon heuristics for production planning and set-up scheduling with backlogs and error-prone demand forecasts

ALISTAIR R. CLARK

University of the West of England, Faculty of Computing, Engineering and Mathematical Sciences, Bristol BS16 1QY, UK

Three families of models and fast heuristic methods are developed for identifying a production plan and immediate set-up schedule for a manufacturing line with changeover times. The initial model is exact, but is optimally solvable only for very short planning horizons. A second modelling and solution method optimizes production one period at a time, first in a backward pass to identify target inventory levels, and then in a forward pass to build up these target inventories. Finally, a third method plans set-ups and lots on a period-by-period basis, estimating the capacity usage of future set-ups. All three methods are first tested under static conditions and then on a rolling horizon basis with differing degrees of demand forecast accuracy, tightness of capacity and length of horizon. Computational experiments confirm that even under great forecasting uncertainty the planning horizon should extend beyond the time at which the horizon is rolled forward and the forecasts updated. Tests also show that the degree of capacity tightness and horizon length affects which approximate models and methods are most successful. The degree of forecast error appears to have limited impact on the planning horizon to be used and relative performance of the models.

Keywords: Production planning; Set-ups; Demand forecasts; Rolling horizons; Heuristics

1. Introduction

The motivation for this paper arose from a production planning and scheduling problem encountered in the canning of liquid products at a drinks manufacturer. The challenge was to optimize the fulfilment of customer orders, taking into account forecast demand, the capacity of the canning line and the impact of product changeover (set-up) times.

As the forecasts of individual customer orders and product-specific demand at the drinks manufacturer were prone to error, it was not clear that it was worthwhile to carry out scheduling for more than a week in advance. However, it was useful to identify a provisional production plan for several weeks ahead in order to have advance warning of possible production backlogs and be able to act accordingly. Consequently, the proposed planning and scheduling system had the following outputs:

(1) A firm production schedule for week 1, detailing which products should be produced on the canning line, their lot sizes and in what sequence.
(2) A provisional plan showing the lot size of product set up in each of the subsequent weeks of the planning horizon. Many products would not be produced every week if scarce capacity was to be used efficiently.

If a large number of products had to be sequenced in week 1, then a mathematical algorithm would almost certainly identify a more efficient production sequence

Email: Alistair.Clark@uwe.ac.uk

Production Planning & Control
ISSN 0953–7287 print/ISSN 1366–5871 online
http://www.tandf.co.uk/journals
DOI: 10.1080/09537280412331286565

than a human scheduler. However, the straightforward nature of product changeover times meant that the sequencing of products was well handled by the manufacturer's experienced scheduler. Consequently the mathematical planning and scheduling model that was developed to decide outputs (1) and (2) above did not seek to sequence the set-ups scheduled for week 1. A simplified version of the model developed for the drinks manufacturer is explained in section 3 and then used to test the fast solution approaches proposed in sections 4 and 5.

A schedule that is optimal for forecast demand over a given horizon will almost certainly be sub-optimal when implemented for the actually occurring demand. Furthermore, only the schedule's production decisions relating to the first planning period would be implemented, after which the horizon would be rolled forward and a new schedule calculated once more with updated demand, inventory and capacity information.

This raises the question as to how far ahead does production need to be planned in order to optimize, over the long term, the actually implemented schedules? The answer is not obvious. If forecasts are generally reliable, then it could well pay to schedule medium-term production in detail in order to provide more accurate production targets for the immediate period to be scheduled. However, if the forecasts are of poor quality, then medium-term scheduling might be unnecessary. It could be a waste of computational resources and effort in period 1 to decide optimal set-up allocations in, say, period 3 if the demand forecasts will have changed by then. So to what extent and how should set-ups, production quantities and their impact in period 3 be modelled so that this impact is taken account of in period 1, the (only) actually implemented period of the model output? Too crude a modelling representation in period 3 could lead to a bad period 1 decision if a forecast surge in demand in period 3 beyond immediately available capacity necessitates production ahead of time. If the surge fails to occur, then the period 1 decision of the crude model might be quite reasonable. the crude model could be just as effective as an exact model, but much more efficient (or even computationally viable) in terms of the computing effort invested.

Accordingly, the objective of this paper is to explore and compare the quality and computing times of several alternative models and associated solution approaches, particularly when implemented on a rolling horizon basis with forecasts of demand, and then assess the tradeoffs between solution computing effort (i.e. input) and the impact of planning and scheduling decisions (output). The paper will use a simplified (but widely encountered) production planning and scheduling problem, distilled from that encountered at the drinks manufacturer, to gain insights into the research question posed above and offer guidance through computational testing.

2. Review of previous research

Closely related to the forecasting of demand is the 'nervousness' that generally occurs when lot-sizing models are applied using a rolling horizon. The concern here is that the demand in the new period(s) added as the horizon rolls forward can cause changes to already planned (but unimplemented) production quantities, thereby necessitating more set-ups than would have occurred with a static solution calculated over the entirety of the rolled horizons. Such nervousness occurs even if demand is perfectly forecast. A further complication is the different behaviour of models in static and rolling horizon environments. Kazan *et al.* (2000) observed that simple heuristic approaches, for example that by Silver and Meal (1973), often outperform statically optimal approaches such as the Wagner–Whitin algorithm Wagner and Whitin (1958) when applied on a rolling horizon basis.

The issue of how to minimize the impact of nervousness has been extensively researched by many investigators (Kropp *et al.* 1983, Blackburn *et al.* 1986, Maesand Van Wassenhove 1986, Baker 1993, Kadipasaoglu and Sridharan 1997, Kazan *et al.* 2000). A comprehensive study of dynamic lot-sizing rules, including nervousness and planning horizon lengths, but excluding demand forecast error, was recently carried out by Simpson (2001). The ability of certain simple lot-sizing rules to dampen nervousness was also researched by Ho (2002). Generally, as one would expect, researchers have found that the longer the planning horizon the better the rolling horizon performance of static models (Baker and Peterson 1979). When demand is perfectly known in advance, nervousness is reported to be accentuated by short planning horizons (Sridharan and Berry 1990).

Almost all of the above studies do not consider the impact of nervousness-induced set-ups on available production capacity through, for example, increased total set-up time. Maes and Van Wassenhove (1986) compared several capacitated lot-sizing heuristics under rolling horizon conditions, warning that tight capacity may cause extra problems as 'small lot-sizes may be planned just to satisfy strict capacity constraints whereas in reality a certain flexibility (e.g. overtime, backlogging) exists'. The model presented below in section 3 permits backlogging so as to allow the batching of small lot sizes if this decreases backlogs overall.

Kimms (1998) suggested several stability measures for capacitated rolling-horizon production planning and developed an effective single-item interactive method for reducing nervousness. He commented that 'if capacities are scarce, we do expect that finding a robust plan is easier than in uncapacitated cases, because the degree of freedom for rescheduling is smaller. This may turn out to be false, if even finding a feasible plan is a non-trivial task. This for instance is the case with positive setup times.' As the model below in section 3 permits backlogging, finding a feasible plan is generally not problematic, but the degree of freedom for rescheduling will be greater than if backlogs were prohibited.

Drexl and Kimms (1997) noted that little research has been carried out into capacitated lot-sizing on a rolling horizon basis. Exceptions include Dimitriadis *et al.* (1997) who, on a different problem to that treated in this paper, developed aggregate models to provide estimates of future capacity requirements and demonstrated their effectiveness. In this paper, we will evaluate and compare both static and rolling horizon performance of several finite capacity lot-sizing solution methods that take set-up times into account. The relative effect of nervousness will be evaluated by the impact of additional set-ups in decreasing capacity available for lot production. Scarcity of such capacity will be felt through the inability to meet demand and hence in backlogs, the principal criterion in the objective function of the model in section 3 below.

Forecasts of demand are rarely free of error. Not surprisingly, and in line with intuition, research confirms the negative impact of forecast errors. De Bodt and Van Wassenhove (1983) showed that even small errors have a large effect on costs. Wemmerlöv (1985) found that as forecast errors increase, so do inventory and nervousness costs while the average order cycle decreases. Zhao and Lee (1993) confirmed that forecasting errors significantly increase total costs and schedule instability, reducing service levels and, in later research (Zhao *et al.* 1995), found that they often have a large impact on the relative performance of lot-sizing rules. Xie *et al.* (2003) concluded that tightness of capacity and length of planning horizon, both of which are experimental factors in the tests of section 6, significantly affected the impact of master production schedule (MPS) freezing parameters on cost, stability and service in multi-item capacitated systems under demand uncertainty. Significantly, both De Bodt and Van Wassenhove (1983) and Fildes and Kingsman (1997) confirm that improved forecasting of demand is more important for performance than the choice of lot-sizing rule, an issue that is explored in the computational tests of section 6.

Interestingly, Zhao and Lee (1993) discovered that prolongation of the planning horizon can actually worsen material requirements planning (MRP) performance when demand is uncertain but result in better performance when demand is free of forecast error. Similarly, Sridharan and Berry (1990) also found that while a large planning horizon decreases costs when demand is deterministic, costs tend to rise as demand forecast error increases. This important effect of forecast error is investigated in section 6.

This paper builds upon previous research by the author. Motivated by industrial problems encountered as a consultant, Clark (1998) developed a fast myopic rule-based heuristic for rolling-horizon lot-sizing and scheduling on a set of parallel machines with sequence-dependent set-up times. The heuristic produced schedules that were robust to demand forecast errors. However, the model assumed just a single set-up at the beginning of each planning time period. Clark and Clark (2000) extended this work, modelling multiple set-ups per period and demonstrated the potential of using fast approximate models on a rolling horizon basis. The poor quality of myopic models was partly avoided by considering forecasts of future demand, but assumed they were error-free. The current paper further develops similar models and their solution methods, and tests them under a wider variety of capacity and forecast accuracies. Clark (2003) developed approximate static MIP models and heuristic methods to assist in identifying a capacity feasible MPS for large product structures in MRP systems with sequence-dependent set-up times. However, that paper used a narrower range of solution methods than the current paper, did not investigate rolling horizon use of the models and worked with perfectly-known demand.

The model developed in section 3 below is an extension of the general lotsizing and scheduling problem (GLSP) of Fleischman and Meyr (1997). The GLSP schedules and sizes lots of multiple products on a single finite capacity machine, permitting multiple set-ups in each single 'large-bucket' time period. This paper extends the GLSP to include product-group set-up times and backlogs. The review on capacitated lot sizing by Karimia *et al.* (2003) notes that there is very little literature on problems with backlogging and multi-group joint set-ups, both features of the model in section 3. The review's authors highlight the development of heuristics with reasonable speed and solution quality for this kind of model as an interesting research area. Thecurrent article's solution methods in subsections 4.2 and 4.3 can be seen as a contribution in that direction, although their period-by-period MIP approach is generally applicable to a wide variety of lot-sizing models.

3. The planning and scheduling model

Consider a manufacturing line that processes P distinct products p from G groups g. Production is planned on a rolling horizon basis for each of T consecutive planning periods t, but demand often overshoots capacity and so backlogs must be allowed for. A generally applicable mixed integer programming (MIP) formulation of the problem is:

Model *PPS* (*production planning and scheduling*)

$$\min \sum_{p,t} \left(h_p I_{pt}^+ + I_{pt}^- \right) \tag{1}$$

such that

$$I_{p,t-1}^+ - I_{p,t-1}^- + x_{pt} - I_{pt}^+ + I_{pt}^- = d_{pt} \quad \forall\, p, t \tag{2}$$

$$\sum_p a x_{pt} + s\left(\sum_p y_{pt} - 1\right) + e\left(\sum_l z_{gt} - 1\right) \le B_t \quad \forall\, t \tag{3}$$

$$y_{pt} \le z_{gt} \quad \forall\, g, p \in P(g), t \tag{4}$$

$$x_{pt} \le M_{pt}\, y_{pt} \quad \forall\, p, t \tag{5}$$

$$x_{pt}, I_{pt}^+, I_{pt}^- \ge 0 \quad \forall\, p, t \tag{6}$$

$$y_{pt} = 0 \text{ or } 1 \quad \forall\, p, t \tag{7}$$

$$z_{gt} = 0 \text{ or } 1 \quad \forall\, g, t \tag{8}$$

where the symbol $\forall$ means 'for all'. The decision variables are:

- x_{pt} Quantity (lot size) of product p produced in period t (≥ 0).
- I_{pt}^+ Inventory of product p at the end of period t (≥ 0).
- I_{pt}^- Backlog of demand for product p at the end of period t (≥ 0).
- y_{pt} Binary variable taking value 1 if product p is produced in period t, and 0 otherwise.
- z_{gt} Binary variable taking value 1 if a product from group g is produced in period t, and 0 otherwise.

The parameter and data inputs are:

- h_p Penalty for holding one unit of inventory of product p from one period to the next.
- d_{pt} Forecast demand for product p at the end of week t.
- I_{p0}^+ Current inventory of product p at the start of period 1.
- I_{p0}^- Current backlog of demand for product p at the start of period 1.
- B_t Available production time on the manufacturing line in period t.
- a Manufacturing line time required to produce one unit of any product, excluding changeovers.
- s Manufacturing line set-up time needed to changeover between products if no change of group is involved.
- e Extra manufacturing line set-up time needed in the changeover between product groups if a change of group is involved.
- $P(g)$ Set of products in group g.
- M_{pt} Upper bound on x_{pt}, calculated as $\min(B_t/a,\ \sum_{\tau=1}^{T} d_{p\tau})$.

In model *PPS*, the planning horizon is T periods while the scheduling horizon is 1 period, the length of time before demand forecasts are updated on a rolling horizon basis. The objective function minimizes a weighted sum of backlogsr and inventory quantities, the former generally having much more weight than the latter (i.e. $0 < h_p \ll 1$) so as to allow, but strongly discourage, planned backlogs rather than outrightly prohibiting them. The dimension of the backlog and inventory penalties can be specific costs, but are more likely to be cost-proxies that reflect the relative undesirability of backlogs and inventory of each product. If desired, backlogs can be strictly forbidden by fixing all the I_{pt}^- values to be zero, but this is not helpful when demand exceeds capacity as model *PPS* will not be able to identify even a feasible plan. A scheduling person needs to know by how much and in which planning periods demand will be backlogged, so as to be forewarned and take appropriate action. Note that the objective function (1) is able to measure the capacity impact of 'nervousness' through its effect principally on backlogs and, to a lesser extent, on inventories.

The first constraints (2) balance production, inventories and backlogs with demand. The inventory I_{pt}^+ and backlog I_{pt}^- cannot both be positive for a given pair (p, t), an occurrence that is prevented by both having positive coefficients in the objective function (1). Many products have current inventory I_{p0}^+ (or backlog I_{p0}^-) which, in a data preparation phase, is discounted from (added to) demand in the first period(s), resulting in zero current inventory and what is known as *effective* demand, calculated as follows:

For $t = 1$ *to* T do {
  Let $d_{pt} = \max(d_{pt} - I_{p0}, 0)$; Let $I_{p0} = \max(I_{p0} - d_{pt}, 0)$;
  If $I_{p0} = 0$ then stop;
} (end of t loop)
If $I_{p0} > 0$ then do not plan for product p.

where $I_{p0} = I_{p0}^{+} - I_{p0}^{-}$ is the current inventory (positive or negative) of product p.

Constraints (3) ensure that production and set-ups take place within the available time. The use of the different s and e set-up times reflect the common reality that a changeover between products from different groups takes longer than one within the same group. For example, on the canning line that motivated this research, consecutive products may be filled from the same liquid (i.e. group), making possible a half-hour set-up, or from different liquids, necessitating a one-hour set-up. The -1 terms in (3) reflect the fact that in many plants, and certainly in the one that inspired this research, the first product of a period is either a continuation from the previous period or a new set-up carried out in the idle time between consecutive periods. If this is not the case, then the -1 terms may be taken out of expression (3) without altering the generality of the methods developed and evaluated in this paper. Constraints (4) and (5) ensure that if a product is processed then both it and its group need to be set up.

Safety stocks reduce production backlogs of demand in the presence of demand forecast errors (Bertrand *et al.* 1990), but are excluded from the model so that the impact of the alternative heuristics methods in reducing backlogs can be strictly evaluated. Generally speaking, the smaller the backlogs, the less the level of safety stocks that is needed. Specifically, future research, outside the scope of the current paper, can investigate which of the heuristic methods performs best in reducing backlogs at various levels of safety stock.

Maes and Van Wassenhove (1986) point out that capacitated lot-sizing models such as *PPS* are powerful and very flexible (i.e. one can add additional constraints such as maximum inventory or overtime availability), but slow (or impossible) to solve if the problem instance is very large. In fact, model *PPS* is NP-hard (Papadimitriou 1994) and so the time needed to solve it optimally explodes exponentially as the problem size increases. In addition, as already discussed, exact optimization is illusory when demand forecasts are often revised and upsets such as machine failure are common, necessitating frequent replanning. For such operational use, the model needs to be solved heuristically if an optimal solution is not quickly available. Bearing this in mind, the next section develops several fast solution approaches.

4. Static heuristic solution approaches

Maes and Van Wassenhove (1986) argue from a practical point of view for simple but very fast algorithms, a perspective with which this paper wholly agrees in developing the methods in subsections 4.2 and 4.3 below. They also advise the use of a routine that looks ahead in time to facilitate future capacity feasibility, an approach we also follow. In this spirit, several alternative heuristic solution approaches are now developed and evaluated, first for a static fixed-horizon application in this section, and then in a rolling horizon environment in section 5 with both perfect and imperfect forecasts of demand.

4.1 *Default branch-and-cut search*

The first solution approach to assess must be the 'lazy' default – simply let Cplex 7.1, as an industrial strength branch-and-cut MIP solver, try to find a good solution within a pre-specified amount of search time. The tests in subsection 4.4 below showed that this approach could obtain good results within a minute or less, but often did not make the best use of limited computing time. As expected, explicitly prohibiting backlogs resulted in infeasible solutions for the test instances where demand was too much for the cumulative capacity available, creating unavoidable backlogs.

This method was tested with search time limits of 10 seconds and 1 minute, denoted Exact10secs and Exact1min respectively.

4.2 *Backward–Forward method*

The T periods in model *PPS* could each be optimized separately were it not for the inventories I_{pt}^{+} and backlogs I_{pt}^{-} spanning more than one period in constraints (2). If we knew or could intelligently estimate their values beforehand, then model *PPS* would decompose into T single-period models that can each be solved very quickly for the size of problem under consideration.

Thus the second solution approach optimizes production one period at a time, first in a backward pass to identify the target inventory levels S_{pt} that are necessary to fulfil demand after period t, and then in a forward period-by-period pass that plans production to build up these target inventories.

First, the Backward pass is carried out as follows:

Take out the back-order variable I_{pt}^{-} from the model.
Let $S_{pT} = 0 \quad \forall\, p$.
For $t = T$ down to 1 do {
  Fix $I_{pt}^{+} = S_{pt}$ as data $\forall\, p$, and free the $I_{p,t-1}^{+}$
    as variables $\forall\, p$.
  Solve the model just for the $I_{p,t-1}^{+}$ and
    period t variables.
  Let $S_{p,t-1} = I_{p,t-1}^{+}\ \forall\, p$.
} (end of t loop)

The target inventory levels S_{p0} for period zero provide indicative minimum levels for the current inventories I^+_{p0} needed to satisfy demand over the T-week planning horizon without backlogs. Assuming that effective demand has already been calculated so that current inventories and back-orders have been zeroed, then an indicator of insufficient capacity to meet overall demand will be the value $\sum_p S_{p0}$.

Then the Forward pass is executed:

Restore the backorder variables I^-_{pt} to the model.
For $t = 1$ to T do {
  Inflate demand by the target: Let $d_{pt} = d_{pt} + S_{pt}\ \forall p$.
  Solve the model for just the period t variables.
  Restore demand: Let $d_{pt} = d_{pt} - S_{pt}\ \forall p$.
  Recalculate and fix inventories $I^+_{p,t-1}$ and back orders $I^-_{p,t-1}$ as data $\forall p$.
} (end of t loop).

This simple Backward–Forward method (denoted *BF*) is fast as the MIPs are small, and can result in a good plan as shown in the computational tests in subsection 4.4 below.

4.3 *Forward pass with linear set-up approximations*

The third and final approach also determines set-ups period by period, but in a single forward pass. Initially, the binary y and z variables representing the set-ups for period 2 onwards are eliminated and compensated for by

(1) either subtracting from the capacity B_t the estimated total time spent on set-ups, resulting in reduced capacity B^*_t in period 2 onwards, denoted model $PPS^{B^*}_\tau$,
(2) or increasing the value of the unit production time from a to a^* in those periods, denoted model $PPS^{a^*}_\tau$.

After solving and fixing the set-ups in period 1, each model is reformulated so that now the set-ups for period 2 are to be determined, with either the reduced capacity B^*_t or the increased unit production time a^* compensating for the lack of explicit representation of set-up times in period 3 onwards. The model is solved afresh, the set-ups in period 2 are fixed, and so on, with $T-2$ further models being solved, until the set-ups for periods $1, \dots, T$ have all been decided.

Used in this manner, each model determines the set-ups in period τ and has the same objective function (1) as model PPS. However, constraints (3) to (5) now apply only to periods 1 to τ, with the values of the set-up variables y_{pt} and z_{gt} having been fixed for $t = 1, \dots, \tau-1$ by previous applications of the model in the forward pass. The constraints for period $\tau+1$ onwards are:

$$I^+_{p,t-1} - I^-_{p,t-1} + x_{pt} - I^+_{pt} + I^-_{pt} = d_{pt} \text{ (unchanged)} \quad \forall\, p,\ t \ge \tau+1 \tag{9}$$

$$PPS^{B^*}_\tau:\quad \sum_p a\, x_{pt} \le B^*_t \quad \text{or}$$

$$PPS^{a^*}_\tau:\quad \sum_p a^*\, x_{pt} \le B_t \quad \forall\ t \ge \tau+1 \tag{10}$$

The number of binary variables in model PPS^{B*}_τ or PPS^{a*}_τ is the same as that for each of the $2T$ MIPs in the *BF* method, but there are more continuous variables, so the solution time of each of the T constituent MIPs will be longer. However, only a single (forward) pass is needed.

How much should the value of the reduced capacity B^*_t or the inflated production time a^* be determined? One possibility is to do nothing and simply let $B^*_t = B_t$ for all t (equivalently let $a^* = \text{a}$). This method is denoted $B^* = B$. Three other possibilities are also considered and tested:

4.3.1 Use Backward–Forward solution.

By constraint (3), B^*_t and a^* should satisfy

$$B_t - B^*_t = s\left(\sum_p y_{pt} - 1\right) + e\left(\sum_g z_{gt} - 1\right) \quad \forall\ t \tag{11}$$

and

$$\sum_p a^* x_{pt} = \sum_p a x_{pt} + s\left(\sum_p y_{pt} - 1\right) + e\left(\sum_g z_{gt} - 1\right) \quad \forall\ t \tag{12}$$

respectively. Thus reasonable values of B^*_t could be calculated from the quickly-obtained results x^{BF}, y^{BF} and z^{BF} of the Backward-then-Forward pass:

$$B^*_t = B_t - s\left(\sum_{pt} y^{BF}_{pt}/T - 1\right) - e\left(\sum_{gt} z^{BF}_{gt}/T - 1\right) \quad \forall\ t \tag{13}$$

averaging set-up usage over all periods. Similarly a^* could be estimated as

$$a^* = a + \frac{s\left(\sum_{p,t} y_{pt}^{BF} - T\right) + e\left(\sum_{g,t} z_{gt}^{BF} - T\right)}{\sum_{p,t} x_{pt}^{BF}} \tag{14}$$

These two methods are denoted B^*BF and a^*BF respectively.

4.3.2 Assume demands and capacity are roughly matched.

Alternatively, assuming that

(1) the total production $\sum_t x_{pt}$ of a product p equals its total effective demand $\sum_t d_{pt}$, which is a reasonable assumption unless capacity is very tight and not enough to meet overall demand,
(2) the canning line is using all available capacity time on set-ups and production, which is also a reasonable assumption unless capacity is loose, so that the right-hand side of constraint (3) is equal to B_t,

then the following formula for B_t^* can be derived:

$$B_t^* = B_t - \frac{\sum_t B_t - a\sum_{p,t} d_{pt}}{T} \quad \text{in model } PPS_\tau^{B^*} \tag{15}$$

and similarly for a^*:

$$a^* = \frac{\sum_t B_t}{\sum_{p,t} d_{pt}} \text{ in model } PPS_\tau^{a^*} \tag{16}$$

These two methods are denoted B^*Formula and a^*Formula respectively. By assumptions 1 and 2 above, we *a priori* expect them to perform best when capacity is moderately tight, as verified by the tests in subsections 4.4 and 6.2 below.

4.3.3 Solve a fast MIP to estimate B_t^* and a^*.

A third alternative is to formulate a representative single-period model that captures the general relationship between the tightness B_t of capacity and the frequency (y, z) and size (s, e) of set-ups. Mean per-period values of demand and capacity are inputs to the following MIP whose solution is then used to estimate B_t^* and a^*:

Model *EBA* (*Estimate B_t^* and a^**):

$$\min \sum_p g_p \bar{I}_p^- \tag{17}$$

such that

$$\bar{x}_p + \bar{I}_p^- = \bar{d}_p \quad \forall\, p \tag{18}$$

$$\sum_p a\,\bar{x}_p + s\left(\sum_p \bar{y}_p - 1\right) + e\left(\sum_l \bar{z}_g - 1\right) \le \bar{B} \tag{19}$$

$$\bar{y}_p \le \bar{z}_g \quad \forall\, g, p \in P(g) \tag{20}$$

$$\bar{x}_p \le \bar{M}_p\,\bar{y}_p \quad \forall\, p \tag{21}$$

where the decision variables are:

$$\bar{y}_p = \begin{cases} 1 & \text{if product } p \text{ is produced;} \\ 0 & \text{otherwise.} \end{cases}$$

$$\bar{z}_g = \begin{cases} 1 & \text{if a product from group } g \text{ is produced;} \\ 0 & \text{otherwise.} \end{cases}$$

$\bar{x}_p$ = Mean quantity of product p produced.

$\bar{I}_p^-$ = Resulting mean back orders of product p.

The data provided are $\bar{d}_p$, the mean demand for product p over periods $1, \ldots, T$ after discounting initial stocks I_{p0}^+ and back orders I_{p0}^- and $\bar{B}$, the mean available production time per period. The upper bound $\bar{M}_p$ on $\bar{x}_p$ is calculated as $\min(\bar{B}/a, \bar{d}_p)$. Model EBA has $P+G$ binary variables and, like the one-period models in the Backward-then-Forward pass, is quickly solvable to optimality. Thus values of B_t^* could be estimated from the $\bar{x}$, $\bar{y}$ and $\bar{z}$ of the model solution:

$$B_t^* = B_t - s\left(\sum_p \bar{y}_p - 1\right) - e\left(\sum_g \bar{z}_g - 1\right) \quad \forall\, t \tag{22}$$

Similarly a^* could be estimated as

$$a^* = a + \frac{s\left(\sum_p \bar{y}_p - 1\right) + e\left(\sum_g \bar{z}_g - 1\right)}{\sum_p \bar{x}_p} \tag{23}$$

These two methods are denoted B^*EBA and a^*EBA respectively.

4.4 *Static computational tests*

The methods of subsections 4.1 to 4.3 are now tested to compare their solution quality and computational run times. All tests were carried out on a Sun Enterprise 450 workstation with a 400 MHz Ultrasparc processor and 1.5 Gb of RAM, using the mathematical programming language AMPL (Fourer *et al.* 2003) for all modelling and basic calculations, calling Cplex 7.1 (ILOG 2001) to solve the MIPs by a branch-and-cut algorithm.

The test data were based on a generalization of the $P = 41$ products from $G = 14$ groups on a single manufacturing line that normally operated for $B_t = \bar{B} = 76$ hours per period (week) t. The demand profiles

d_{pt} were based on actual data supplied by the manufacturer, suitably modified for experimental purposes in order to assess how well the solution approaches developed in this paper performed under different conditions (such as tightness of capacity, and varying accuracy of demand forecasts).

The production rate, excluding changeovers, was 1,000 units an hour, i.e. $a = 0.001$ hours/unit. The set-up time to change between products in the same group was $s = 0.5$ hours while the extra setup time involving a change of group was $e = 0.5$ hours also. These values are approximate and simplified for research purposes. However, it must be pointed out that the production line was not perfectly under control and that in practice the actual values of a, s and e varied over a range of $\pm 20\%$. This paper will not explore the implications when production parameter values are not well known and also vary, but it should be clear that under such conditions, even with perfect demand forecasts, exact optimization of model *PPS* is again illusory, generally resulting in sub-optimal schedules.

The inventory penalty was set to $h_p = 0.001$ so as to give overwhelming priority to the minimization of backlogs over inventory. For situations where this priority is not paramount and tradeoffs between backlogs and inventories are interesting, the model *PPS* could be bi-objective and the efficient (Pareto) frontier of non-dominated solutions explored to provide the human scheduler with a choice of tradeoffs (Ergott and Gandibleux 2002, Good win Wright 2004).

The mean demand $\bar{d}$ per product per period was set to

$$\frac{\bar{B} - n(P-1)s - n(G-1)e}{Pa}$$

so that capacity was either 'very tight' (i.e. $n = 0 \Rightarrow \bar{d} = 1,854$, allowing for no set-ups at all, inevitably causing backlogs), 'moderately tight' ($n = 0.5 \Rightarrow \bar{d} = 1,530$, allowing for a given product to be set-up every second period, probably causing some backlogs), or 'loose' ($n = 1 \Rightarrow \bar{d} = 1,207$, allowing for a given product to be set-up in every period, so that any backlog will be due to the variability of demand around the product's mean). The computational tests below will show that tightness of capacity has a strong effect on the relative performance of the solution methods.

The mean demand $\bar{d}_p$ of a product p was in proportion to that encountered at the manufacturer where some products had high demand while certain others were specialist low-volume brands. The demand profile d_{pt} for each product p varied normally around its mean $\bar{d}_p$ with coefficient of variation of 0.25 so that 95.5% of demands d_{pt} are within the interval $[0.5\bar{d}_p, 1.5\bar{d}_p]$ with none negative. The initial inventories I^+_{p0} and backlogs I^-_{p0} were set to zero, noting that this would not be the case in a 'steady state' under tight capacity.

One hundred different demand sets $\{d_{pt} \mid p = 1, \ldots, 41; t = 1, \ldots, 10\}$ were randomly generated, each of which could then be proportionally scaled for tightness of capacity as defined above. Each demand set was solved and evaluated for all 300 combinations of solution method (10 levels), tightness of capacity (three levels) and planning horizon ($T = 1, 2, \ldots, 10$). The emphasis of the computational tests was on the relative quality of *fast* solution methods, i.e. that would execute in a minute or less. All methods gave identical results for $T = 1$, of course.

Figure 1 summarizes the mean solution quality, as given by objective function (1), of all 10 methods over $T = 1, 2, \ldots, 10$ for each of the three levels of tightness of capacity. For moderately and very tight capacity, the solution values increase as the horizon T increases because there are more periods over which capacity struggles to cope with demand. For loose capacity, the solution values for the best-performing methods generally level off as T increases because capacity is easily able to meet demand in the later periods. For loose capacity, most of the backlog values were zero, but some were positive due to the absence of initial inventory to meet any above-average demand in the first period. For moderately tight capacity, some backlogs and stock were unavoidable. For very tight capacity, the backlog values were inevitably very large.

For each of the 27 combinations of capacity tightness and horizon $T = 2, 3, \ldots, 10$, the Friedman non-parametric chi-squared statistical test (Lowry 2003) or evaluating multiple treatments was applied on the same group of individuals (i.e. the 100 random demand sets). The null hypothesis was H_0: the median values of the solution objective (1) are identical for the 10 methods. In all 27 tests the solution method treatment was highly significant with p-value <0.001, resulting in acceptance of the alternative hypothesis H_1: the median objective values of the 10 solution methods are *not* all identical.

Observe from figure 1 that, for loose capacity, the default branch-and-cut Exact method of subsection 4.1 was competitive for all values of the planning horizon T when 1 minute of computing time was allowed. If only 10 seconds is allowed, then for $T = 4$ onwards it was bettered by other methods, particularly for larger values of T. In fact, for $T = 5$ onwards, method Backward–Forward (*BF*) was not only the most competitive, but also the fastest (taking from 1.5 seconds elapsed time at $T = 2$ to just 4.2 seconds at $T = 10$). Method $B^* = B$ performed almost as well, but took about twice as long when $T = 2$, rising to nine times as long when

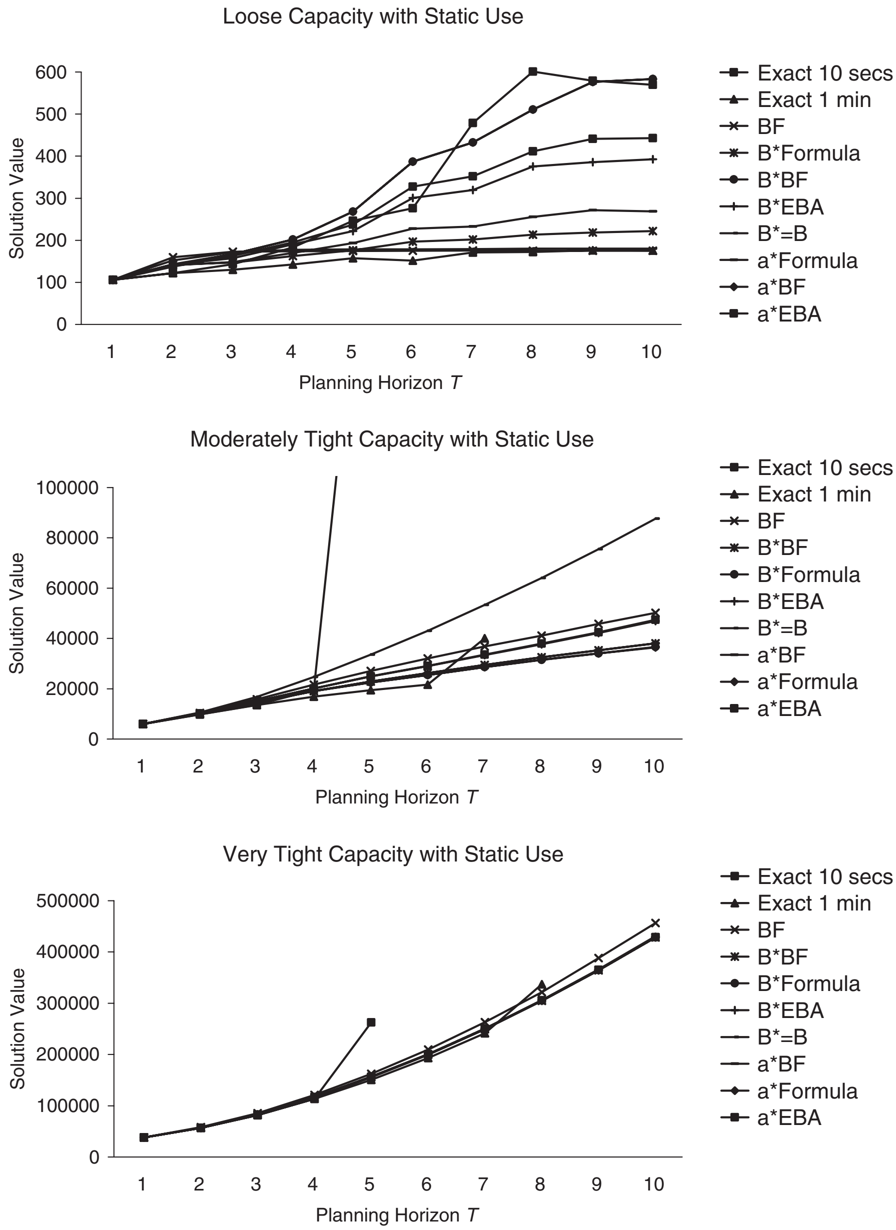

Figure 1. Static solution quality.

$T = 10$. In the best two methods (Exact1min and *BF*), 55 of the 100 instances had zero backlogs. Note the reasonable performance of the Formula method for estimating a^* and B^*, and then the poor performance of the BF and EBA methods for doing so. Observe also that each B^* method generally did better than the equivalent a^* method.

In contrast, when capacity was moderately tight, method $B^* = B$ (followed by *BF*) was the least competitive for all values of T. However, methods B^*BF and a^*BF performed the best for $T \geq 7$, but took from 40 to 75 seconds. The exact method was the most competitive for $T \leq 6$ allowing 1 minute of computing time, but only for $T \leq 4$ if just 10 seconds was permitted. In fact, both exact methods were unable to find feasible solutions for $T \geq 8$ and 6 respectively. The Exact10secs solution value for $T = 5$ was so large that it was plotted as way off the top of the graph in figure 1. Overall the B^*Formula and a^*Formula methods performed competitively for all values of T, taking from 6 to 50 seconds of elapsed computing time. Observe also that each B^* method generally did the same as the equivalent a^* method.

When capacity was very tight, all methods were overwhelmed by the lack of capacity, building up cumulative backlogs. However, note that the exact methods were the best for small T, but once more were unable to identify feasible solutions in less than 10 seconds (1 minute) from $T = 6$ ($T = 9$) upwards. Observe that the BF solution was noticeably the worst for all values of T. Generally, the B^* and a^* methods (including $B^* = B$) all performed similarly and well, and certainly so for $T \geq 7$ taking from 7 to just 11 seconds of elapsed computing time. Again, each B^* method generally did the same as the equivalent a^* method.

To conclude overall, the Exact 1 minute method is the best when capacity is loose, but otherwise only for $T \leq 6$ or 7, However, for longer planning horizons, the Exact methods are unable to even identify feasible solutions when capacity is moderately or very tight, precisely the circumstances when a planning model is most needed. In this situation, the a^*/B^* BF/Formula methods are the best. The very fast B. method performs reasonably well when capacity is loose or moderately tight, but is the worst when very tight.

Since each B^* method generally performed the same as or somewhat better than the equivalent a^* method, results from now on in this paper will be presented only for the B^* methods in order to avoid over-cluttered figures. The rolling horizon tests below did include all the a^* methods, but their performance was again the generally same or a little worse than the equivalent B^* method.

5. Solution approaches for rolling horizons

If demand forecasts remain unchanged between one rolling horizon and the next, then the production plan for the first horizon's periods 2 to T should change only a little by the reapplication of model PPS, becoming the plan for the next horizon's periods 1 to $T-1$. The addition of period T will cause some changes as fresh needs or opportunities for set-ups are implemented by the model. If demand forecasts are substantially modified, then the new horizon's plan for periods 1 to $T-1$ should be expected to change significantly.

How can the methods developed for static use in section 4 be adapted for rolling horizons? Several possibilities are now considered.

5.1 *The exact model PPS*

Model PPS, if used on a rolling horizon basis, still has to be solved in its entirety, but, as in its static use, each Cplex branch-and-cut search will need to be terminated after a limited amount of computing time, say, 10 seconds or 1 minute, as in subsection 4.4.

Only the first period's schedule is actually implemented – the other periods are present in the model just to take account of forecast future demand. If the forecasts are error-prone, then it may be effective to weight earlier periods (particularly the first) more heavily in the objective function, for example with a geometric decay:

$$\min \sum_{p,t} t^{-\alpha\beta}\left(I_{pt}^{-} + h_p I_{pt}^{+}\right) \qquad (24)$$

where $\alpha \geq 0$ is a measure of the error of the demand forecasts (as explained below in section 6) and $\beta \geq 0$ is a time-period weighting exponent. The less reliable the forecasts, then the greater the value of α and so the less future periods are relatively weighted in objective function (24). If $\beta = 0$ then all periods are equally weighted, but if $\beta = 20$ and $\alpha = 0.05$ or 0.10 (values used in the tests of section 6), then the weighting is non-trivial, as shown in figure 2.

5.2 *The backward-then-forward pass*

Only the first period's plan is actually implemented, and so the forward pass of subsection 4.2 can be curtailed after period 1, enabling some economy of calculation. As each model optimizes over just a single period, it makes no difference to weight the objective function as in expression (24) above.

5.3 *Forward pass with linear set-up approximations*

The case made in subsection 5.1 for weighting the objective function (24) also applies here.

The first of the T single-period optimizations of model $PPS_{\tau}^{B^*}$ provides a period 1 solution that has implicitly

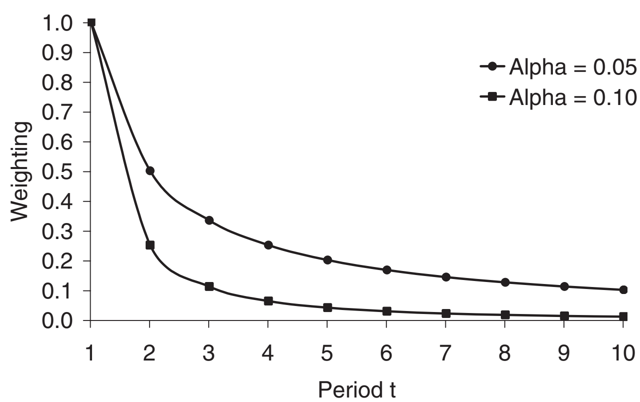

Figure 2. Time period weights when $\beta = 20$ and $\alpha = 0.05$ or 0.10.

taken into account and fedback future capacity requirements through the use of the reduced capacity B_t^*. This period 1 solution could be used on a rolling horizon basis, thus saving approximately a T-fold reduction of computing effort by not solving model $PPS_\tau^{B^*}$ T times. However, although the period 1 set-ups (y_{p1}, z_{p1}) are fixed in the subsequent $T-1$ single-period optimizations of $PPS_\tau^{B^*}$, the period 1 lot-sizes x_{p1} are not fixed. It is possible that they could change as a consequence of explicitly taking into account the capacity requirements of future set-ups in the remaining $T-1$ single-period optimizations of $PPS_\tau^{B^*}$. Thus all of the T single-period set-up optimizations of model $PPS_\tau^{B^*}$ in subsection 4.3 should be carried out, as was done in the computational tests of section 6 below.

6. Rolling horizon computational tests

6.1 *Experimental design*

The quality of a set of implemented rolling-horizon decisions is evaluated by comparing their outcome to that resulting from an alternative set of decisions. In the tests below, the outcome is the eventual result, i.e. the value of the objective expression (25):

$$\sum_p \sum_{\theta=1}^{10} \left(h_p I_{p\theta}^+ + I_{p\theta}^-\right) \tag{25}$$

calculated from the implemented production decisions and the actual demand that occurs as the horizon rolls forward a total of 10 times, large enough to enable valid comparisons to be made between the experimental factors. The final periods of the last $T-1$ rolling horizon models lie beyond period 10 and so use forecasts of demand whose actual values are never eventually known to the model and user.

Five factors were taken into account in the statistical experimental design to compare the rolling horizon outcomes, namely:

(1) the length T of the model planning horizon;
(2) the solution method;
(3) the degree α of demand forecast error;
(4) the use, or not, of weights $t^{-\alpha\beta}$ in objective function (24);
(5) the tightness of production capacity.

We assume that as the lead-time from the forecast to the actual demand event decreases, accuracy will improve with the incorporation of new market and customer information into the forecast. As mentioned in subsection 5.1, a parameter $\alpha \geq 0$ is used to quantify the degree of error in the demand forecasts, where $\alpha = 0$ corresponds to perfect demand forecasts. The base value V_T for first forecast F_T of demand for a given product at the end of the planning horizon T periods ahead is simulated as the eventual actual demand value V_0, multiplied by unity plus a random element that is proportional to T, α and the standardized normally distributed random variable r:

$$V_T = \max\{0, V_0(1 + T\alpha r)\} \tag{26}$$

Thus, the larger the value of α, the further the first forecast's base value V_T is likely to be from the eventual value V_0. The values V_T and V_0 are the same when $\alpha = 0$.

As the planning horizon rolls forward, the updated forecast F_t of demand for the same product in period $t\,(=T, T-1, \ldots, 1, 0)$ is simulated as the interpolated base value $V_t = V_0 + (t/T)(V_T - V_0)$ at fraction t/T of the distance from V_0 to V_T which is then, similarly to expression (26), multiplied by unity plus a random element that is proportional to the period t, α and the standard normal variable r:

$$F_t = \max\{0, V_t(1 + t\alpha r)\} \quad t = T,\ T-1, \ldots, 1, 0 \tag{27}$$

The value V_t assures the convergence of the forecast F_t to the eventual actual demand V_0, as illustrated in figure 3. The larger the value of α, the further the forecast F_t will be from V_t. When $\alpha = 0$, the forecast is perfect: $F_t = V_t = V_0$.

This method of simulating forecasts is simple and intuitively sensible – it not only incorporates a random element, but also recognizes that accuracy tends to improve as the lead-time between the forecast and the actual demand diminishes. It enables us to test how the methods in section 5 perform under varying degrees of forecast accuracy. It should be emphasized, however, that expressions (26) and (27) are not intended to reflect how people actually make forecasts, an active subject of ongoing research (Goodwin 2002). In addition, note that confirmed orders can implicitly be considered as that part of the demand quantity which is taken as almost certain and below which the forecast is highly unlikely to fall. For example, in figure 3 demand never falls below 50 units. The forecast uncertainty is within a range from 50 units upwards.

Similarly to the static tests of subsection 4.4, 25 different eventual demand sets $\{F_{0pt} \mid p = 1, \ldots, 41;\ t = 1, \ldots, 19\}$ were randomly generated and proportionally scaled for tightness of capacity. Each of these 25 demand sets was solved on a rolling horizon basis for 10 periods, with demand forecasts updated as described above, and evaluated for all $9 \times 10 \times 3 \times 3 = 810$ combinations of planning horizon $(T = 2, \ldots, 10)$, solution method (10 levels), tightness of capacity (three levels), and forecast error α $(=0,\ 0.05$ and $0.10)$. To evaluate the impact of a non-trivial weight $\beta = 20$ in the objective function (24), all methods were then

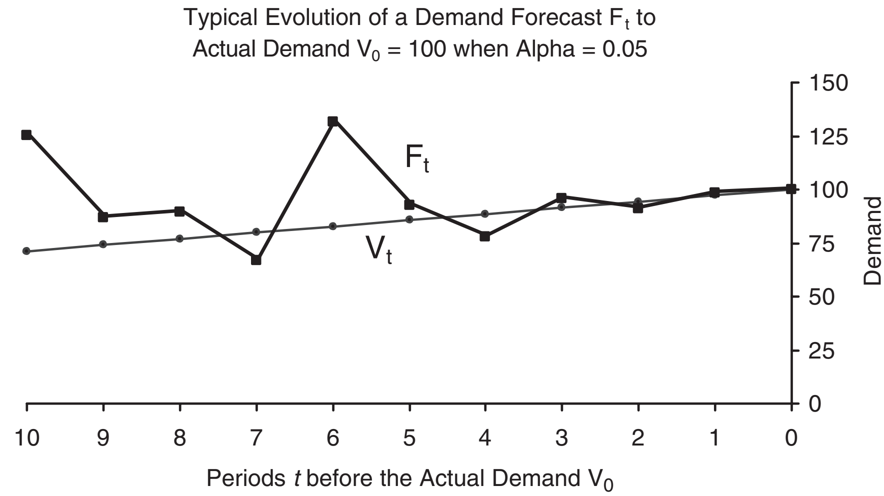

Figure 3. Typical evolution of a demand forecast.

tested again on the same 25 demand and forecast sets, for $\alpha = 0.05$ and 0.10 and all three levels of tightness of capacity. The one-period horizon $T = 1$, whose solution for a given capacity tightness does not depend on the values of α and β nor on the solution method used, was also evaluated as a benchmark for the $T = 2, \ldots, 10$ tests.

6.2 *Experimental results*

Figure 4 to 6 summarize the mean quality of the solutions, i.e. the value of the objective expression (25): $\sum_p \sum_{\theta=1}^{10} \left(h_p I_{p\theta}^+ + I_{p\theta}^-\right)$ calculated from the implemented production decisions and the actual demand that occurs as the horizon rolls forward a total of 10 times. As such, the values can only be conceptually compared with figure 1's static $T = 10$ values to which they are broadly similar in the equivalent case of perfect forecasts.

Although some of the detail is not visible, figure 4 to 6 nicely illustrate the range of results and major patterns. Relevant details for specific solution methods are brought out in the discussion below. To answer the question posed in section 1, looking at the $T = 1$ results for the better performing methods, it is clear overall that production does need to be scheduled ahead of period 1 in order to optimize, over the long term, the actually implemented schedules. In other words, on a rolling horizon basis, it is worthwhile to schedule ahead, with the possible exception of tight capacity under very poor forecasts of demand.

The Friedman test was again applied on the 25 random demand sets for each of the combinations of capacity tightness, horizon T, forecasts error α and weight β specified just above in subsection 6.1. The null hypothesis was H_0: the median values of objective function (1), calculated from the implemented production decisions and the actual demand as the horizon rolls forward a total of 10 times, are identical for the 10 methods. In all tests for $t = 2, 3, \ldots, 10$ the solution method treatment was highly significant with p-value <0.001, resulting in acceptance of the alternative hypothesis H_1: the median solution values of all 10 methods are *not* all identical. For example, in the last graph in figure 5 (moderately tight capacity for $\alpha = 0.10$ and $\beta = 20$) where the solutions for different methods at $T = 3$ appear to be very close to each other, the Friedman test gave a p-value less than 0.001.

Figure 4 shows that, with perfect forecasts, the Exact1min method continues to be the best when capacity is loose, closely followed by the *BF* and then $B^* = B$ methods, as in the static case. Note that for these three methods the length of the planning horizon T makes little difference to the quality of the solution when forecasts are perfect, but certainly does for the other methods. In fact, counter-intuitively, the longer the planning horizon, the worse the quality of the already inferior B^* BF/Formula/EBA solutions. As forecasts worsen, the performance of all methods deteriorates. However, the Exact1min method still continues to be the best, closely followed by the $B^* = B$ and *BF* methods. Observe how all but one of the other methods, including Exact10secs, perform very badly (going up and off the graph) when the forecasts are very poor ($\alpha = 0.10$). Note that all methods perform relatively well at $T = 3$ when forecasts are poor, i.e. scheduling three periods ahead is worthwhile, even when only the immediate period's schedule is actually implemented.

Figure 5 shows that for moderately tight capacity, the Exact1min method also continues to be the best with

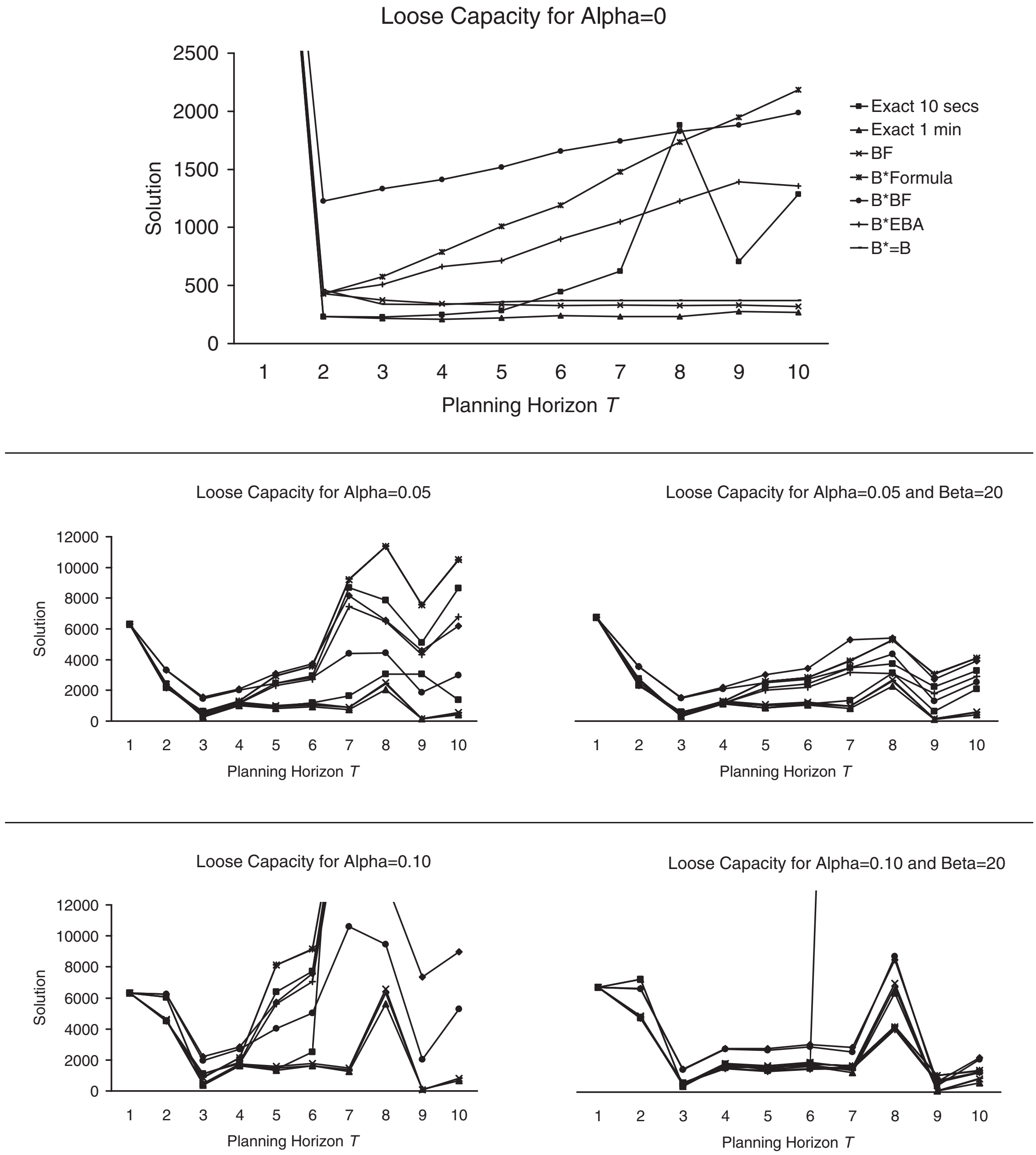

Figure 4. Rolling horizon solutions for loose capacity.

perfect forecasts for $T \leq 7$, closely followed by the B^*Formula method, as in the static case. The next best are the $B^* = B$ and BF methods, somewhat in contrast to their worst performance in the static case. Observe again that for these five methods the length of the planning horizon T makes little difference to the quality of the solution, but certainly does for the B^*EBA method, getting worse as the horizon lengthens. As forecasts become more error-prone, the performance of all methods deteriorates overall, although the Exact1min method still continues to be the best, closely followed broadly identically by the B^*Formula, $B^* = B$ and B. methods, i.e. the quality ranking is roughly maintained as forecasts worsen. Note how the best solutions are obtained with a planning horizon T between 3 and 5 and when $T = 9$, i.e. scheduling three to five periods ahead is worthwhile when forecasts are poor, even when just implementing period 1's schedule.

When capacity is very tight with perfect forecasts, figure 6 shows that the Exact1min method also continues to be the best for $T \leq 8$, closely followed by the B^* Formula/EBA and $B^* = B$ methods, as in the static case. The next best is the B. method and by far the worst is the B^*BF method. Again, as forecast error

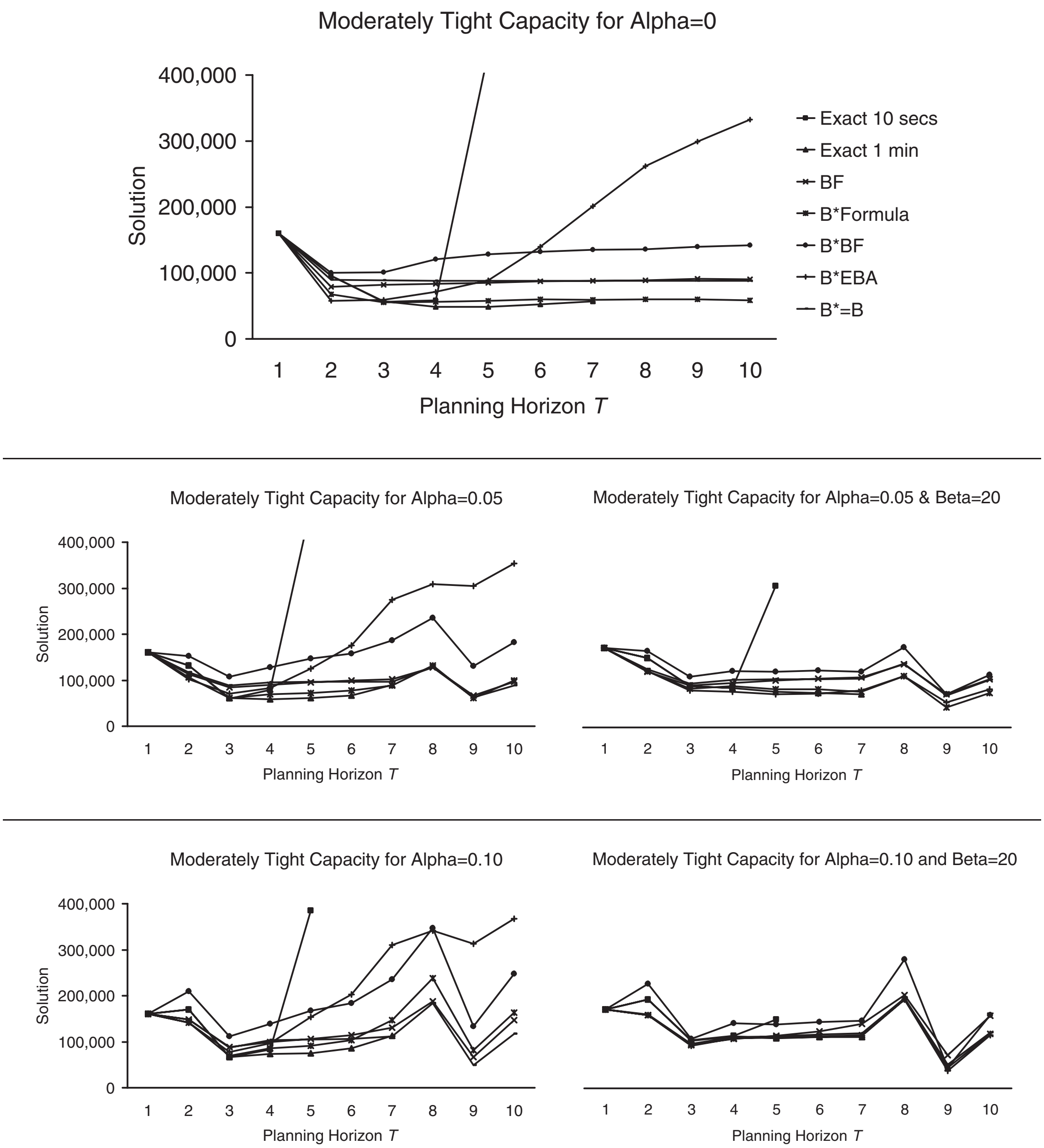

Figure 5. Rolling horizon solutions for moderately tight capacity.

increases, the performance of all methods deteriorates overall, with the Exact1min, B^* Formula/EBA and $B^* = B$ methods continuing to be the best. Note again how the best solutions are obtained with a planning horizon T between 3 and 6 and when $T = 9$, i.e. scheduling three to six periods ahead is worthwhile with poor forecasts, even when only the period 1's schedule is implemented.

For all tightnesses of capacity, the value $\beta = 20$ of the time-period weighting exponent in expression (24) has a slightly worsening effect on the best methods at the well-performing small values of the horizon T. However, as can be seen in figure 4 to 6, it has a dramatic improving impact on the methods whose $\beta = 0$ solution quality deteriorated with the lengthening of the horizon T. The value $\beta = 20$ enables these methods to heavily discount unreliable demand forecasts far in the future, permitting a radical improvement in their performance. Nevertheless, their solutions are still inferior to those of the best-performing methods identified in figures 4 to 6.

7. Conclusions

At the beginning of this paper, the question was raised as to how far ahead production needs to be scheduled in order to optimize, over the long term,

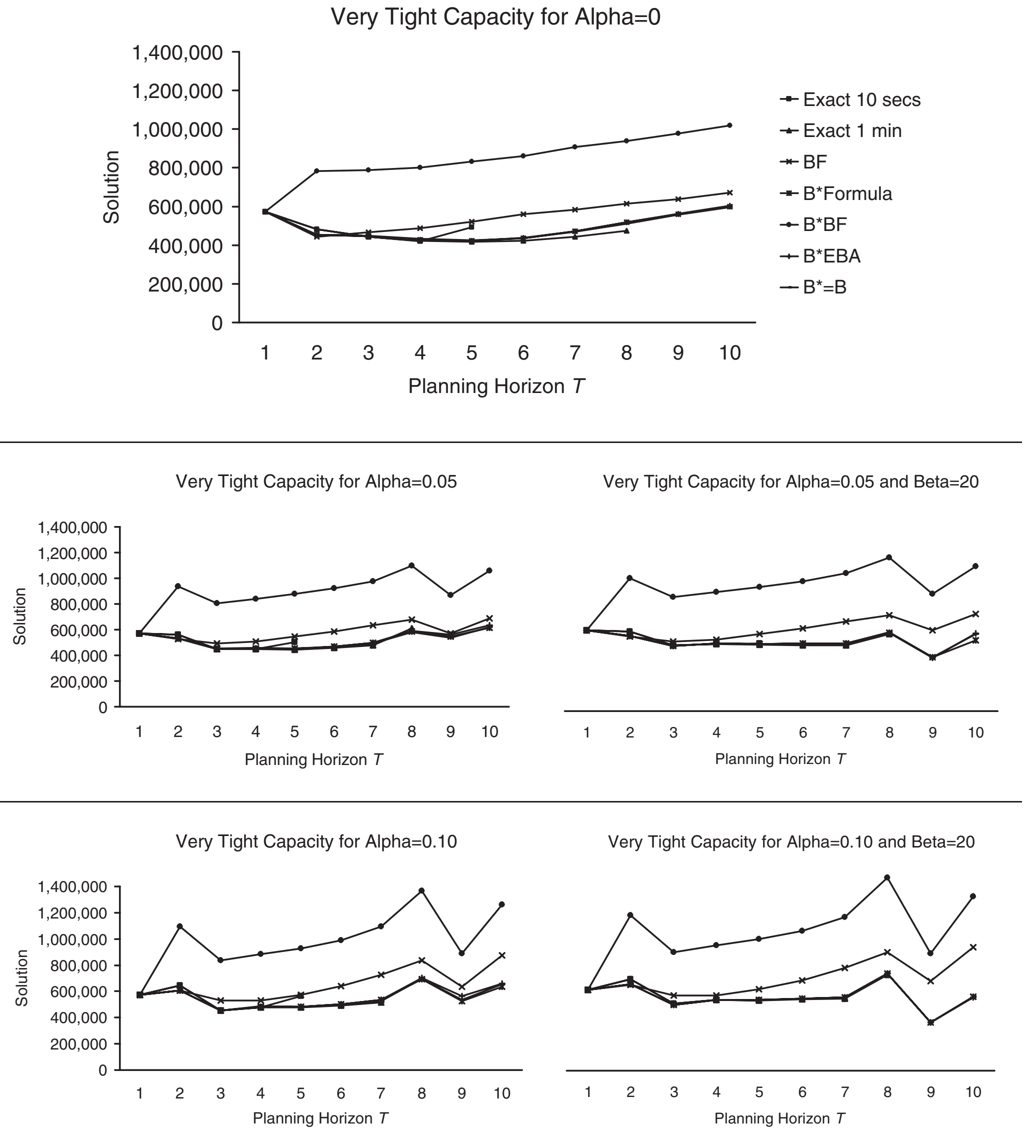

Figure 6. Rolling horizon solutions for very tight capacity.

the actually implemented schedules. The test results analysed above in subsection 6.2 show that, irrespective of the quality of demand forecasts and tightness of capacity, the scheduling horizon should extend to between 3 and 5 periods ahead. This is a useful result that, while strictly valid only for the system and data in hand, can be extrapolated from the canning line, a typical production system, as a guiding indicator for other systems.

Turning to the exact and heuristic methods used to solve the model, the tests show that the relative static performance of the methods generally tends to be reproduced in their rolling horizon use, which is good news for those researchers who have carried out tests only under static conditions. The one exception is the case of moderately tight capacity where the bad static performance of the $B^* = B$ and BF methods contrasts with their relative superiority over the very poor performance of the B^* BF/EBA methods.

Interestingly, the computational results do not mirror the MRP findings of Lee (1993) that prolongation of the planning horizon worsens performance when demand is uncertain, but produces better performance when demand is free of forecast error. However, the test results do support the finding of Sridharan and Berry (1990) that costs tend to rise as demand forecast error increases. But their other finding, namely, that a large planning horizon decreases costs when demand is deterministic, is not supported by the results for the data used.

In addition, as Zhao and Lee (1993) found, forecasting errors do significantly reduce service levels (as measured by backlogs). However, contrary to the results of Zhao *et al*. (1995) with various lot-sizing rules, figures 4 to 6 show that the degree of forecast error appears to have limited impact on the relative performance of the solution methods tested. Overall the Exact 1 minute method reigns supreme in terms of quality and robustness to forecast error, but is not viable for the longer planning horizons T when capacity is moderately or very tight. This is not really a disadvantage (at least for data parameters similar to those used in the tests) as the rolling horizon results show that best Exact1min solutions occur over shorter horizons when $3 \leq T \leq 6$. Nevertheless, a faster method that performs nearly as well on a rolling horizon basis under conditions of moderately or very tight capacity, and takes a fraction of the time, may well be preferable and more practicable if numerous schedules need to be generated and compared under many varying scenarios.

If there are more than the 41 products of this study, then the Exact 1 minute method could well be impracticble under loose capacity. In this case, the test results suggest that the quicker BF and $B^* = B$ methods (but not B^*Formula) would provide a good alternative, even with very poor forecasts of demand. Under moderately tight capacity, the results suggest the use of method $B^* = BF$, BF or B^*Formula. If capacity is very tight, then method $B^* = B$ or B^*Formula (but not BF) is indicated by the test results. All three methods will be relatively fast as their MIPs cover only one planning period. If the number of products is so large that the MIPs cannot be optimally solved in reasonable time, then limits can be imposed on each MIP's solution time, as in Clark (2003). Some indication of capacity tightness, such as that in subsection 4.4, needs to enter into the choice of fast solution method. However, in the absence of such capacity information, method $B^* = B$ would appear to be a robust choice that avoids bad performance. Simplicity wins, it seems.

Acknowledgements

I would like to thank the referees for their valuable and constructive feedback, resulting in a better paper.

References

Baker, K.R., Requirements planning. In *Handbooks in Operations Research and Management Science*, edited by S.C. Graves; Vol. 4, pp. 571–627, 1993 (Elsevier Science: New York).

Baker, K.R. and Peterson, D.W., An analytical framework for evaluating rolling horizon schedules. *Management Science*, 1979, **25**, 341–351.

Bertrand, J.W.M., Wortmann, J.C. and Wijingaard, J., *Production Control: A Structured and Design Oriented Approach*, 1990 (Elsevier: Amsterdam).

Blackburn, J.D., Kropp, D.H. and Millen, R.A., A comparison of strategies to dampen nervousness in MRP systems. *Management Science*, 1986, **32**(4), 413–429.

Clark, A.R., 1998, Batch sequencing and sizing with regular varying demand. *Production Planning & Control*, **9**(3), 260–266.

Clark, A.R., Optimization approximations for capacity constrained material requirements planning. *International Journal of Production Economics*, 2003, **84**(2), 115–131.

Clark, A.R. and Clark, S.J., Rolling-horizon lot-sizing when setup times are sequence-dependent. *International Journal of Production Research*, 2000, **38**(10), 2287–2308.

De Bodt, M.A. and Van Wassenhove, L.N., Cost increases due to demand uncertainty in MRP lot sizing. *Decision Sciences*, 1983, **13**(3), 345–362.

Dimitriadis, A.D., Shah, N. and Pantelides, C.C., Rtn-based rolling horizon algorithms for medium term scheduling of multipurpose plants. *Computers and Chemical Engineering*, 1997, **21**(Supplement), S1061–S1066.

Drexl, A. and Kimms, A., Lot sizing and scheduling – survey and extensions. *European Journal of Operational Research*, 1997, **99**, 221–235.

Ergott, M. and Gandibleux, X., Multiobjective combinatorial optimization – theory, methodology and applications. *International Series in Operational Research and Management Science*, 2002, **52**, 369–444.

Fildes, R. and Kingsman, B., Demand uncertainty and lot sizing in MRP systems: the value of forecasting. *Working paper No. MS 01/97*, Department of Management Science, Lancaster University, 1997.

Fleischman, B. and Meyr, H., The general lotsizing and scheduling problem. *OR Spektrum*, 1997, **19**(1), 11–21.

Fourer, R., Gay, D.M. and Kernighan, B.W., *AMPL – A Modeling Language for Mathematical Programming*, 2nd edn, 2003 (Duxbury Press/Brooks-Cole: Belmont, CA) (http://www.ampl.com/).

Goodwin, P., Integrating management judgment and statistical methods to improve short-term forecasts. *Omega*, 2002, **30**, 127–135.

Goodwin, P. and Wright, G., *Decision Analysis for Management Judgement*, 3rd edn, 2004 (Wiley: Chichester).

Ho, C.-J., Evaluating dampening effects of alternative lot-sizing rules to reduce MRP system nervousness. *International Journal of Production Research*, 2002, **40**(11), 2633–2652.

ILOG, *CPLEX 7.1 User's Manual*, ILOG S.A, BP 85, 9 Rue de Verdun, 94253 Gentilly Cedex, France, 2001 (www.cplex.com and www.iro.umontreal.ca/˜gendron/IFT6551/CPLEX).

Kadipasaoglu, S.N. and Sridharan, S.V., Measurement of instability in multi-level MRP systems. *International Journal of Production Research*, 1997, **35**, 713–737.

Karimia, B., Fatemi Ghomia, S.M.T. and Wilson, J.M., The capacitated lot sizing problem: a review of models and algorithms. *Omega*, 2003, **31**, 365–378.

Kazan, O., Nagi, R. and Rump, C.M., New lot-sizing formulations for less nervous production schedules. *Computers and Operations Research*, 2000, **27**, 1325–1345.

Kimms, A., Stability measures for rolling schedules with applications to capacity expansion planning, master production scheduling and lot sizing. *Omega*, 1998, **26**(3), 355–366.

Kropp, D.H., Carlson, R.C. and Jucker, J.V., Heuristic lot-sizing approaches for dealing with MRP systems nervousness. *Decision Sciences*, 1983, **14**, 156–165.

Lowry, R., Concepts and Applications of Inferential Statistics, 2003 (http://faculty.vassar.edu/lowry/webtext.html).

Maes, J. and Van Wassenhove, L.N., Multi item single level capacitated dynamic lotsizing heuristics: a computational comparison. Part ii: Rolling horizon. *IIE Transactions*, 1986, **18**, 124–129.

Papadimitriou, C.H., *Computational Complexity*, 1994 (Addison-Wesley: Reading MA).

Silver, E.A. and Meal, H.C., A heuristic for selecting lot size quantities for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment. *Production and Inventory Management*, 1973, **14**, 64–74.

Simpson, N. C., Questioning the relative virtues of dynamic lot sizing rules. *Computers and Operations Research*, 2001, **28**, 899–914.

Sridharan, V. and Berry, W.L., Freezing the master production schedule for make-to-stock products under demand uncertainty in a rolling horizon planning environment. *Decision Sciences*, 1990, **21**(1), 87–120.

Wagner, H.M. and Whitin, T.M., Dynamic version of the economic lot size model. *Management Science*, 1958, **5**, 89–96.

Wemmerlöv, U., Comments on 'Cost increases due to demand uncertainty in MRP lot-sizing': a verification of ordering probabilities. *Decision Sciences*, 1985, **16**(4), 410–419.

Xie, J., Zhao, X., Lee, T.S. and Goodale, J., Freezing the master production schedule under single resource constraint and demand uncertainty. *International Journal of Production Economics*, 2003, **83**, 65–84.

Zhao, X. and Lee, T.S., Freezing the master production schedule in multilevel material requirements planning systems under demand uncertainty. *Journal of Operations Management*, 1993, **11**, 185–205.

Zhao, X., Lee, T.S. and Goodale, J., Lot-sizing rules and freezing the master production schedule in MRP systems under demand uncertainty. *International Journal of Production Research*, 1995, **33**(8), 2241–2276.

Alistair Clark is a Principal Lecturer in Operational Research (OR) in the Faculty of Computing, Engineering and Mathematical Sciences at the University of the West of England, Bristol. He was formerly a Production Systems Specialist at UniSoma (an OR consulting and systems development company in Brazil), an academic adviser at IBM's Latin American Institute of Technology, an analyst in the UK government OR Service, a lecturer at several UK universities, and a VSO volunteer teacher in Wa, Ghana. Alistair has led production planning and scheduling consulting and systems development assignments for Polyenka, Tilibra and Thomson Consumer Electronics in Brazil, and Matthew Clark Brands in the UK, among other companies. He belongs to the UK Operational Research Society (ORS) on whose Council he represents the Western region of England. He is also the UK representative on the Council of EURO – the Association of European Operational Research Societies.