The Production Planning Problem: Clearing Functions, Variable Lead Times, Delay Equations and Partial Differential Equations

D. Armbruster

Abstract Determining the production rate of a factory as a function of current and previous states is at the heart of the production planning problem. Different approaches to this problem presented in this book are reviewed and their relationship is discussed. Necessary conditions for the success of a clearing function as a quasi steady approximation are presented and more sophisticated approaches allowing the prediction of outflow in transient situations are discussed. Open loop solutions to the deterministic production problem are introduced and promising new research directions are outlined.

Keywords Supply chains · Production planning problem · Conservation laws · Clearing functions

1 Introduction

The production planning problem, the starts into a factory that generate a desired production profile in the future, is either explicitly or implicitly a major theme of almost half of the chapters in this book. Aouam and Uzsoy [1] have production planning in the title, Lefeber [15] deals with the issue in the context of the reference tracking problem using Model Predictive Control, Göttlich et al. [11] change the control variable from production starts to outing probabilities and then try to match a particular output pattern, Braun and Schwartz [7] assume a model for a production planning problem and deal with the nervousness of the scheduling algorithm, Perdaen et al. [20] measured the success of controlling a reentrant manufacturing line through the Push-Pull-Point via its missed production targets and Ringhofer [21] uses traffic

D. Armbruster (✉)
School of Mathematical and Statistical Sciences,
Arizona State University, Tempe, AZ 85287-1804, USA
e-mail: armbruster@asu.edu

D. Armbruster and K. G. Kempf (eds.), *Decision Policies for Production Networks*,
DOI: 10.1007/978-0-85729-644-3_12,

type models based on hyperbolic partial differential equations (PDEs) to study the production planning problem with priority rules.
The common problem of determining the output of a production unit (machine, factory, supply chain node) is dealt with at very different levels of sophistication. This paper will connect some of these approaches and determine their applicability and their approximation errors. Finally we discuss some of the practical questions of choosing an appropriate clearing function mode for a production planning problem and identify the open questions associated with those problems in general and the clearing function approach in particular.

2 Clearing Function Models

Typically, all production units are stochastic and hence the production process is a stochastic process. As a result, the mathematical model that comes closest to reality is a discrete event simulation model. However, even if every production detail is modeled, the characterization of the stochastic processes involved is non-trivial generating another *model* of reality. Given the fact that the stochastic processes are not well understood and given that they are very time consuming to simulate, the need for aggregate models is generally accepted and this need drives the discussion. Hence we will discuss deterministic models that, one hopes, represent *average* behavior in some sense.
Depending on the perspective of the author the number of items a production unit produces is either characterized by a flux (or outflux), typically denoted by $F(t)$ and defined as the rate of production as a function of time, or by the number of units produced in a time interval n (shift, day, etc.) often called X_n. Given that the production unit has a finite production rate μ or capacity C the simplest constraint is to require that

$$\begin{aligned} X_n &= C_n, \quad \text{or} \\ F(t) &= \mu. \end{aligned} \tag{1}$$

This constraint is only true for a system that is overloaded and hence produces at constant average production rate μ. At the same time, due to the stochastic nature of the production process, a truly overloaded system leads to increasing work in progress (wip), making this an unrealistic assumption for the characterization of the production process for any significant length of time.

If arrival rates are less than maximal production rates, mass-balance equations model the time evolution of an inventory I:

$$\begin{aligned} I_{n+1} &= I_n + R_n - X_n, \quad \text{or} \\ \frac{\mathrm{d}I}{\mathrm{d}t}(t) &= \lambda(t) - F(t), \end{aligned} \tag{2}$$

where R_n are the starts in the time interval n (Aouam and Uzsoy [1] use the term release rates) and $\lambda(t)$ and $F(t)$ are the start rate and the outflux, respectively. Notice that $R_n = \int_{t_{n-1}}^{t_n} \lambda(s)\mathrm{d}s$. As only $\lambda(t)$, (or R_n) is controlled, the system is not defined without a model for the outflux $F(t)$ (or X_n). This is where the clearing function first introduced by Karmarkar [14] comes in. The clearing function is a state equation that defines the outflux F as a function of the wip W in steady state, i.e.

$$F = \Phi(W). \tag{3}$$

The functional form of the clearing function Φ has been determined in many different ways: Measured in real factories, modeled via an $M/M/1$ queue, modeled after the fundamental diagram of a traffic model [17] etc. (see e.g. [1, 11]),

$$\begin{aligned} \Phi &= \frac{\mu W}{1+W} && M/M/1 \\ \Phi &= \mu W - W^2 && \text{fundamental diagram of traffic.} \end{aligned} \tag{4}$$

Lefeber [15] defines the clearing function by its functional inverse, i.e. the wip as a function of the flux, approximated as an $M/G/1$ queue:

$$W = \frac{c_B^2 + c_E^2}{2} \frac{r^2}{1-r} + r \tag{5}$$

where r is the utilization of the machine $r = \frac{\lambda}{\mu}$ and c_B and c_E are the coefficient of variation of the arrival and machine departure processes. Both Aouam [1] and Lefeber [15] notice that the clearing function can be approximated by piecewise linear functions, making the production planning problem an Integer-LP optimization problem.

Even at this low level of approximation there is a basic inconsistency: the clearing function is supposed to describe the outflux *in steady state* as a function of wip level. However, the clearing function is used with a wip level that is a function of time and is updated constantly to determine the outflux as a function of time. Hence, by making the outflux follow instantaneously any change in the wip level, the fundamental assumption is that the wip level changes slowly relative to the damping time of the underlying stochastic process. Therefore the fundamental assumption that justifies the use of a clearing function is that by the time the wip-level has reached a new state, the stochastic process determining the outflux is back in steady state. As a result the outflux is never in transient and always characterized by its steady state behavior. This is known as the *quasi-steady assumption* or the *adiabatic model*.

The quasi-steady assumption poses a major problem for the applicability of any type of clearing function approach. Since almost no research in production planning is concerned with the specific nature of the stochastic process, there are no good estimates to my knowledge about the damping time of the stochastic processes. In fact, even the concept is ill-defined without discussing the timescales and magnitudes of the stochastic disturbances. One way presumably would be analogous to Aouam and Uzsoy's [1] formulation of their ZOIP algorithm in assuming that there is the everyday stochasticity (in their case restricted to the demand variability) which would be captured by the clearing function and there are extraordinary events that require extraordinary measures for which the clearing function approach is not well suited. I would consider operator overload, operator negligence and scheduled machine maintenance to be part of the everyday stochasticity. For semiconductor production lines the time that e.g. a scheduled machine shutdown would be felt could be described as of an order less than the cycle time. The stochastic damping time would therefore be of the order of a day or two. Hence, to stay with the semiconductor production model, ramping start-ups by 20% over a weekly schedule would not violate the quasi-steady assumption of a clearing function model but ramping up within a day would.

3 Dealing with Delays

All clearing function approaches so far have considered the wip at the current time interval as the independent variable determining the outflux at the current time. As most production is not started and completed within a day and no production process is instantaneous this is in general not a good model. This is especially true for cycle times that are long relative to the planning period since parts that have just entered the production process will not be involved in determining the current outflux, unless the factory is reentrant or some other special circumstances apply. There are two approaches that cover the delayed response of a production unit in this book (but see also Hackman and Leachman [12] who have a detailed discussion of delayed timing issues for linear models of production systems): Effective processing times (EPT) and partial differential equation models.

3.1 The Effective Processing Time Approach

The effective processing time t_e is the mean time that a part needs to get through a stochastic processing unit, without considering the waiting time [13, 15]. Hence for a single machine it can be considered as $\frac{1}{\mu}$, with μ the average machine processing rate. By focussing on the start rates of the machines the delay experienced by a part will be fixed, independent of the buffer length. We define u_{up} to be the uptake rate of the machine immediately upstream and t_e its effective processing time. Calling the u_d the uptake rate of the machine immediately downstream of an inventory $I(t)$ (buffer), its time evolution can be written as

$$\frac{\mathrm{d}I}{\mathrm{d}t} = u_{\mathrm{up}}(t - t_e) - u_d(t). \tag{6}$$

By using an ordinary differential equation for the time evolution of the inventory, parts are losing their identity and, without additional modeling, the cycle time through a factory cannot be recovered from this model unless queuing is minimal. However, by using the functional inverse of the clearing function (Eq. 5) to bound the uptake of a machine, we can approximate the overall production rate of a production line rather accurately. Notice though that using the clearing function (or its functional inverse) still makes the effective process time model a quasi-steady state model with all the problems discussed before. In particular, the production rates of a machine are based on the average behavior of the machine and arrival processes and hence fast changing transients may not be resolved properly.

The relationship between the effective processing time approach and clearing functions in fact is complicated and not completely understood. In particular, a constant effective processing time is not equivalent to a linear clearing function. A clearing function describes the interaction between the stochastic processes that describe the machine availability and the stochastic processes that describe the product availability whereas the effective processing time focusses on the machine availability alone, making it necessary to develop a model for the uptake of the machine again. Lefeber [15] uses the clearing function as a bound for this uptake model but one could imagine more sophisticated approaches.

3.2 Transport Equations

Considering a factory as a pipe and parts flowing through the factory as a fluid, we can describe the transport through the factory via standard transport equations studied extensively in fluid mechanics. In contrast to fluid mechanics, the spatial variable defining the transport direction is not given by physical space but rather by the degree of completion of the part, or the stage of the production. Calling $x \in [0, 1]$ the degree of completion, $\rho(x, t)$ describes the density of parts at stage x at time t. If the fluid moves with a velocity field $v(x, t)$ then the flux is described as $F(x, t) = v(x, t)\rho(x, t)$. Mass conservation then is given by the partial differential equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial F}{\partial x} = 0. \tag{7}$$

Since $v(x, t) \geq 0$ the fluid moves from left to right, allowing a boundary condition to be imposed at $x = 0$. Typically the boundary condition is $F(0, t) = \lambda(t)$, i.e. the local flux at zero is the arrival rate of the parts into the factory. Together with an initial wip profile $\rho(x, 0) = \rho_0(x)$ this sets up a well defined hyperbolic problem. Notice that we are describing a flow that is continuous in its parts and continuous in its spatial direction. This should be distinguished from the so-called fluid equation models of queueing theory [6] which are continuous in its parts but describe a flow through a finite and distinct number of queues, leading to a set of Ordinary Differential Equations (ODEs).

To clarify issues, let us examine the solutions to the mass conservation Eq. (7) for a constant velocity $v(x,t) = c$. In this case the transport equation becomes a linear first-order wave equation with the solution

$$\rho(x,t) = \begin{cases} \rho_0(x - ct) & \text{for } t < \frac{x}{c} \\ \frac{\lambda(t-\frac{x}{c})}{c} & \text{for } t > \frac{x}{c}. \end{cases} \tag{8}$$

Integrating Eq. (7) over x and defining the total wip $W(t) = \int_0^1 \rho(x,t)\mathrm{d}x$ we get

$$\frac{\mathrm{d}W}{\mathrm{d}t} = F(0,t) - F(1,t) = \lambda(t) - \lambda\left(t - \frac{1}{c}\right). \tag{9}$$

Comparing the inventory $I(t)$ in the EPT approach and the wip $W(t)$ in the PDE approach, we see that the two delay Eqs. (6) and (9) only differ in their accounting of the parts—the EPT approach counts them *after* the machine whereas the PDE approach counts it *in* the machine.

When we integrate the transport equation over the completion space all wip is aggregated into one variable, the total wip, and any uneven distribution of the wip is lost. Hence again, cycle time of an individual part cannot be recovered in the delay equation model nor are we resolving short term fluctuations of the outflux.

In contrast, in the PDE model, we can clearly follow the transport of any local wip portion given by $\rho(x,t)\mathrm{d}x$ over time through the factory. Hence, if the observation time interval Δt and the cycle time τ satisfy $\tau \gg \Delta t$, a PDE model (or its discretization) is the only one that allows us to follow the flow of parts through the production unit. For cycle time of the order of the observation times, the clearing functions based on the total wip are appropriate. This observation is independent of the velocity model that is used to describe the flow through the factory, i.e. independent of the type of clearing function that is used.

3.2.1 Clearing Functions for PDEs

Since the Karmarkar clearing function $F = \frac{\mu W}{1+W}$ [14] is only a good approximation if the cycle time is of the order of the observation time, a clearing function describing the flux in a spatially extended partial differential equation should depend on the local variable x. In particular $F(1,t)$ should depend on the density $\rho(1,t)$. We have argued in [3] that for a strongly re-entrant flow with FIFO dispatching rules the velocity should be uniform over the total completion space and hence

$$F(x,t) = v(W(t))\rho(x,t) = \frac{\mu}{1+W}\rho(x,t) \tag{10}$$

would be a good flux function.

For acyclic flows (linear production lines) heuristic discussions lead to space dependent clearing functions given the local production rate at stage x either just as a function of the local density e.g.

$$F(x,t) = \frac{\mu}{1+\rho(x,t)}\rho(x,t) \tag{11}$$

or, as in Ringhofer's chapter [21] as a linear interpolation between the flux expected for the whole factory and the flux expected at the very last machine.

$$F(x,t) = \frac{1-x}{\tau\left(\int_0^1 \rho(z,t)\mathrm{d}z - x\tau_0\right)}\rho(x,t), \tag{12}$$

where $\tau(x,t)$ is the time to completion of a production sitting at stage x at time t.

Another heuristic model that develops a clearing function for a linear production line with finite buffers has recently been developed by Armbruster et al. [4],

$$F(x,t) := \begin{cases} \frac{\mu\rho}{1+\rho+k\rho(1-x)} & \text{for } \rho < M \\ 0 & \text{for } \rho \geq M, \end{cases} \tag{13}$$

where k is an adjustable constant and M is the maximal buffer space. Experiments (Goossens [10]) that shut down the last machine in the factory and subsequently restart the whole factory with full buffers show a cascading collapse of production traveling upstream in the factory and, once the last machine has been repaired, a slower recovery to steady state. PDE simulations using the flux (Eq. 13) show good, though not perfect, agreement with the discrete event simulations.

4 Transient Clearing Functions

We have seen that, using any type of clearing function model, whether in a discrete mass balance equation describing inventories or in a continuous flow model which is characterized as a hyperbolic PDE, the model assumes that the local production rate instantaneously adjusts to the one given by the equilibrium relationship between flux and wip described by the clearing function. Recently Missbauer [19] has studied the issue of clearing functions for systems that are not in steady state. He considers a simple M/M/1 queue with a production rate of $\mu = 1$ and studies the expected output $E[X]$ over five time units as a function of the expected load $E[L]$ at the end of the five time units, depending on the initial wip w_0 and the arrival rate $\lambda(t)$, i.e.

$$E[L] = w_0 + \int_0^5 \lambda(s)\mathrm{d}s. \tag{14}$$

He argues that clearing functions that describe such transient behavior should not just depend on the total load of the system but on three variables:

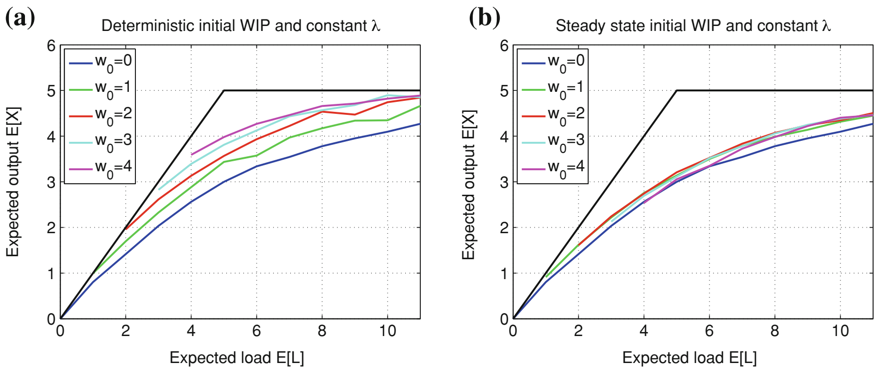


Fig. 1 Simulation of an $M/M/1$ system for different initial wip and load, 10^3 simulations per data point. **a** Deterministic initial wip, **b** Initial wips sampled from steady state distributions

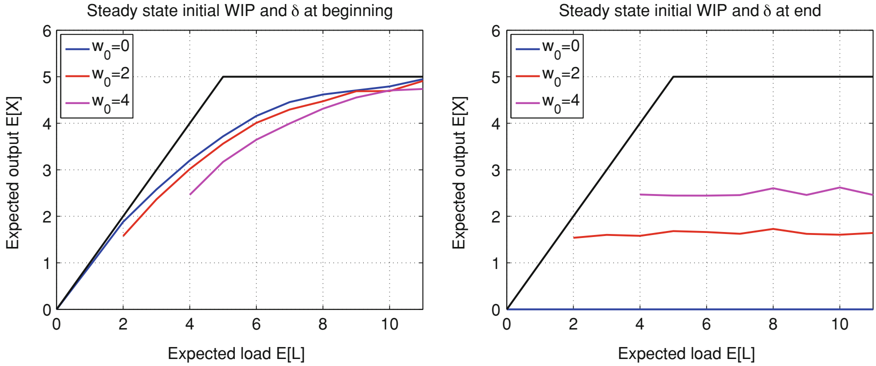
- the expected initial wip level,
- the expected input during the period,
- the probability distribution of the wip level at the beginning of the period.

Fonteijn [9] extended Missbauer's study. Figure 1a shows that with a deterministic intial wip and a constant rate of influx $\lambda = \frac{E[L]-w_0}{5}$, the output over five time periods depends crucially on the initial wip as shown before by Missbauer.

However, if the initial wip is randomly distributed corresponding to an expected value of the initial wip of w_0, the curves describing the mean outflux behavior (averaged over the initial wip distribution) as a function of the expected load all pretty much collapse into each other. Figure 1b shows that the differences between different mean initial wips become very small.

The assertion that the *total* expected input during the observation period determines the outflux can be shown to be wrong by looking at the outflux for the same total input, distributed differently over time: Fig. 2a shows the clearing function for an experiment where the necessary influx to generate the expected load over the time period of five time units is generated at the *beginning* of the time period. As a result, the initial wip is instantaneously increased and hence the outflux is higher than in the case of a constant influx in time as shown in Fig. 1b. Figure 2b shows the clearing function for an experiment where the necessary influx to generate the expected load over the time period of five time units is generated at the *end* of the time period. As a result, none of the influx will come out of the factory within the time period and the outflux is only determined by the initial wip.

We can conclude from these four figures that a steady state-based clearing function will be a reasonably good description of the outflux, if the system is increased from an average of 20% of production capacity to an average of 90% of production capacity with a constant ramp within five cycle times. If ramping is done much faster or if the system is prepared in a particular initial state, the initial condition matters and so does the timing of the ramp.



(a) Total influx at $t = 0$ **(b)** Total influx at $t = 5$

Fig. 2 Clearing functions for an $M/M/1$ queue with different average initial wips and (**a**) all influx arriving at the beginning of the observation interval and (**b**) at the end of the observation interval

Another approach to go beyond the quasi-static models based on regular clearing functions follows from Armbruster et al. [2]. They have developed a hierarchical set of moment equations that models the time dependent behavior of the flux and higher order moments like the variance, etc. The hierarchy of moments generate an infinite set of hyperbolic equations which by itself is not of practical use. The standard approach to such hierarchies is to truncate them at some level via a *moment closure* that defines a higher-order moment whose time dependence is not resolved any more by a relationship to lower-order moments. The simplest such closure is the clearing function, defining the flux in terms of the density. The next more sophisticated approach leads to a system of two PDEs where the flux becomes a dynamic variable. In [2] the following system of two partial differential equations is derived:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial F}{\partial x} &= 0 \\ \frac{\partial F}{\partial t} + \frac{\partial v(x,t) F}{\partial x} &= 0, \end{aligned} \tag{15}$$

where again $F(x,t) = v(x,t)\rho(x,t)$. Intuitively, the second equation describes the fact that a perturbation in the production rate, e.g. a region of high production rate, travels downstream with a velocity v . Treating again a factory as a single $M/M/1$ queue, the flux at the beginning of the factory is given as

$$F(0,t) = \rho(0,t)\frac{v_0}{1+W}. \tag{16}$$

This model performs better than any other model for Missbauer's test cases [19] but the errors are still significant for highly variable inputs [9].

5 Solving the Production Planning Problem Using PDE Models

In [18] La Marca et al. determine the start rate as a function of time that minimizes the mismatch between a desired production rate over a given time interval and the actual production rate according to the solution of a PDE model of a production flow. Specifically they define the cost function

$$J(\rho, \lambda) := \frac{1}{2}\int_0^{\tau} (\mathrm{d}(t) - F(1,t))^2 \, \mathrm{d}t, \tag{17}$$

where $\mathrm{d}(t)$ is the instantaneous demand rate and $F(1,t) = \rho(1,t)v(1,t)$ is the instantaneous outflux. Minimizing the cost functional $J(\rho, \lambda)$ over all possible influx functions $\lambda(t)$, subject to the PDE–dynamics introduced previously, i.e.

$$\begin{aligned}
&\min_{\lambda(t)} J(\rho, \lambda) \text{ subject to} \\
&\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial}{\partial x}(F(x,t)) = 0 \\
&\lambda(t) = v(\rho)\rho(0,t) \\
&\rho_0(x) = \rho(x,0) \\
&F(x,t) = \frac{\mu\rho(x,t)}{1 + \int_0^1 \rho(s,t)\mathrm{d}s}
\end{aligned} \tag{18}$$

solves the production planning problem over the time horizon τ. The method is based on the formal adjoint method for constrained optimization, incorporating the hyperbolic PDE as a constraint of a nonlinear optimization problem.

Figure 3 shows the optimal influx for a sinusoidally varying demand. Notice that the nonlinear dependence of the lead time on the wip in the system generates a sawtooth-like form of the optimal influx. In contrast, a constant lead time would have lead to an input function that has the exact same functional form as the demand, just phase-shifted.

6 Conclusion and Open Problems

6.1 Practical Considerations

The production planning problem can be considered a prototype problem of multiscale modeling. We have seen that, depending on the modeling context, different time scales become relevant. At the same time different approximations are appropriate for those different time scales. It is my contention that the current literature on production planning and clearing function does not pay enough attention to the time scale issue and how the purpose and usage of the model chooses its time scale and model sophistication.

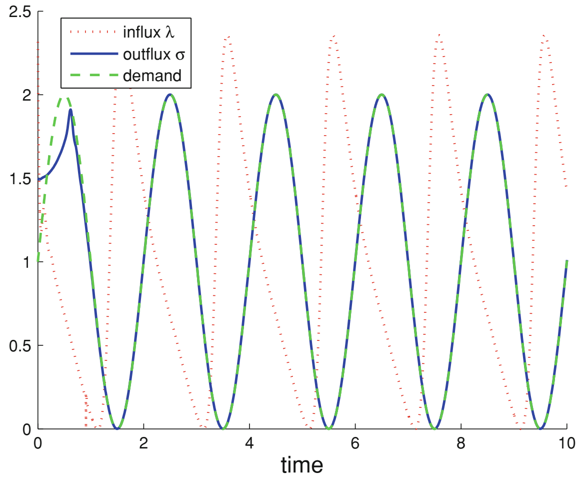


Fig. 3 A sinusoidally oscillating demand function (from [18])

A clearing function approach (and all derivatives discussed here) is based on a model that assumes that the production rate is described by a steady state relationship to the wip. Since the processes that are described here are stochastic processes, the definition of the stochastic process becomes very important. In particular, a steady state that allows us to determine a clearing function is defined by a stationary process, where the average quantities (e.g. wip and outflux) over the the relevant time interval do not differ in a relevant way from the long term averages. The concept of ordinary (captured in the average description) versus extraordinary events (not captured) becomes crucial. No extraordinary event can be incorporated into a clearing function approach although extraordinary events may define initial conditions and recovery from the initial conditions may be modeled by a clearing function approach. Specifically, a disaster like the Japan earthquake in 2011 cannot be modeled in an aggregate description of a supply chain but modeling the recovery from the disaster could be attempted.

Hence any situation that allows the production system to adjust to its steady state *before* the outflux is measured can in principle be well approximated by a clearing function approach. Any situation that requests a higher temporal resolution for the outflux will be badly modeled by a clearing function and will need models that describe the time evolution of the wip in the factory and the flux together.

When a clearing function model is appropriate there are still other timescale considerations that influence the choice of a model: If the cycle time is much larger than the observation time interval a PDE model (Eq. 7) or an effective processing time model (Eq. 6) with a suitable delay are the only ones that properly model the flow of parts through the production unit. If the cycle time is of the order of the observation times, an instantaneous clearing function model (Eqs. 2 and 3) without additional delays based on the total wip is appropriate. If the cycle time is much shorter than the observation time, an instantaneous clearing function based on the wip at the moment of the change in influx is appropriate. For instance, if a change in influx is done at the beginning of the time interval, the outflux should be based on the wip at the end of the time interval. If the influx is changed at the end of the time interval, the outflux is based on the wip at the beginning, etc.

6.2 Continuous Versus Discrete Models

There are two ways to consider the relationships between the different approaches discussed here: Once could start with time intervals based on production time units (shifts, days, weekly schedules) and on production intervals based on machines and come to the iterative model shown in Eq. (2). Going to a continuum in time leads to ODE models which are called fluid models in the queuing theory context and are the bases of the effective processing time models (Eq. 6). Assuming a large number of machines allows us to go to the continuum in production space leading to the PDE models (Eq. 7). Alternatively one could consider the ODE model a discretization of the PDE in space and the discrete time model a subsequent discretization of the ODE model in time. Usually discretization of a PDE in space and time can be done on any scale and with many different schemes while the discretizations in Eqs. (2) and (6) are based on the granularity of the actual production process. However, for simulations of hyperbolic PDEs, the space and time discretization are not independent. In order to have a stable algorithm they have to satisfy a necessary condition known as the CFL condition [16]. Daganzo [8] has shown that the CFL condition is not just a numerical analysis issue but that it is equivalent in choosing order policies that prevent the bullwhip effect.

6.3 Further Work

The current understanding of the production planning problem and the promising approaches presented in this book generate a multitude of research problems, at least some of which are well within reach of current mathematical modeling and optimization techniques.

- As Aouam [1] noted, the production planning problem has really two parts, production planning under nonlinear lead times, which has been the topic for most of this chapter, and production planning under stochasticity leading to stochastic optimization. While La Marca et al. [18] conceptually solved the tracking problem for nonlinear lead times, their approach falls short of a usable and robust algorithm since it deals only with the open loop problem. Hence, any perturbation that disturbs demand or production during the planning time horizon will typically invalidate the optimal production plan calculated with La Marca's algorithm and hence will require a complete replanning. Model predictive control linking the tracking algorithm to a discrete event simulation that provides the reality against which the re-planning will have to occur is a promising direction. However other stochastic optimization approaches should also be tried for the PDE-based clearing function models.

At this point we also have a connection to the control theoretical approaches discussed by Braun and Schwartz [7]. Any closed loop feedback control system will have to deal with schedule nervousness and limit the amount of variations that are allowed for an optimal schedule. Inherently, the approaches in [7] do not care where the errors in the model come from—they could be coming from variations in the demand but they could also be coming from the linearization of a fundamentally nonlinear clearing function. As the modeling errors increase and the request for scheduling stability stays the same, at some time there will not be a feasible solution that is at the same time smooth enough and accurate enough. Whether real industrial problems can satisfy these constraints and whether general rules for the success of this approach can be developed are open problems.

Alternatively, limiting schedule changes in a closed loop version of La Marca's model could be done in much the same way as in [7] through frozen horizons, move suppressions and schedule change suppression. This would have the advantage of a much better—nonlinear—model but the disadvantage that the LP-based optimization tools would not suffice any more.

- Deriving clearing function models from first principles is an extremely hard problem as it adds another layer to the already very hard problem of the relationship between queuing systems and their fluid models. A fluid model as it is used in approximation theory for queuing theory treats the products arriving at a queue as a continuum flux leading to an ODE description for the average behavior of a queuing system. That problem is still not completely solved for multi-class queueing networks and arbitrary priority rules at the machines. However for single class queueing networks the equivalence between the fluid model and the long term average behavior is well understood (see [6] for a introduction into this subject).

  To derive a clearing function for a supply chain or a factory with a large number of machines requires the additional limit of infinitely many production steps modeled through a continuum motion along the completion line. No first principle theory dealing with the interplay of a large number of products going through a large number of production steps exists to my knowledge.
- It will be much easier to determine more sophisticated approaches for highly transient systems. While the studies of Fonteijn suggest that using two PDEs (Eq. 15) based on the multi-moment expansion is better than one, a complete study of the approximation errors associated with these equations has not been done. In particular, it is unclear for which acyclic production systems the closure (Eq. 16) is the correct one and what other closure options are available.
- Asmundsson [5] has discussed the production planning problem for production of more than one product type. The Asmundsson ACF approach is a rough way of doing this with big time buckets and potentially restrictive assumptions, but seems to work well in many cases as a practical approach. There is an obvious

relationship to Ringhofer's [21] service rule discussion for PDE models that has not been explored.

It is hoped that this book serves as an incentive for many researchers in applied mathematics, industrial engineering and operations research to study some of these fascinating problems.

Acknowledgements This research was supported in parts by a grant from the Stiftung Volkswagenwerk, by NSF grant DMS 1023101 and by the INTEL Research Council.

References

1. Aouam T, Uzsoy R, An exploratory analysis of production planning in the face of stochastic demand and workload-dependent lead times, this volume
2. Armbruster D, Marthaler D, Ringhofer C (2004) Kinetic and fluid model hierarchies for supply chains. SIAM Multiscale Model Simul 2(1):43–61
3. Armbruster D, Marthaler D, Ringhofer C, Kempf K, T-C Jo (2006) A continuum model for a re-entrant factory. Oper Res 54(5):933–950
4. Armbruster D, Göttlich S, Herty M (2011) A scalar conservation law with discontinuous flux for supply chains with finite buffers. SIAM J App Math 71(4):1070–1087
5. Asmundsson JM, Rardin RL et al (2009) Production planning models with resources subject to congestion. Naval Res Logist 56:142–157
6. Bramson M (2008) Stability of queueing networks. Lecture notes in mathematics, 1950. Springer, Berlin
7. Braun MW, Schwartz JD, A control theoretic evaluation of schedule nervousness suppression techniques for Master Production Scheduling, this volume
8. Daganzo CA (2003) Theory of supply chains. Springer, New York
9. Fonteijn J (2009) Analysis of clearing functions and transient PDE models.Technical report TU Eindhoven, SE 420613
10. Goossens P (2007) Modeling of manufacturing systems with finite buffer sizes using PDEs. Masters thesis, TU Eindhoven, Department of Mechanical Engineering, SE 420523
11. Göttlich S, Herty M, Ringhofer C, Optimal order and distribution strategies in production networks, this volume
12. Hackman T, Leachman RC (1989) A General framework for modeling production. Manag Sci 35(4):478–495
13. Hopp WJ, Spearman ML (2000) Factory physics, 2nd edn. Irwin/McGraw-Hill, New York
14. Karmarkar US (1989) Capacity loading and release planning with work-in-progress (wip) and lead-times. J Manuf Oper Manag 2:105–123
15. Lefeber E, Modeling and control of manufacturing systems, this volume
16. LeVeque RJ (2002) Finite volume methods for hyperbolic problems. Cambridge University Press, Cambridge
17. Lighthill MJ, Whitham GB (1955) On kinematic waves. II. A theory of traffic flow on long crowded roads. Proc Royal Soc Lond A(229):317–345
18. La Marca. M, Armbruster D, Herty M, Ringhofer C (2010) Control of continuum models of production systems. IEEE Trans Autom Control 55(11):2511–2526
19. Missbauer H (2010) Order release planning with clearing functions. A queueing-theoretical analysis of the clearing function concept. Int J Prod Econ. doi:10.1016/j.ijpe.2009.09.003
20. Perdaen D, Armbruster D, Kempf K, Lefeber E, Controlling a re-entrant manufacturing line via the pushpull point, this volume
21. Ringhofer C, Traffic flow models and service rules for complex production systems, this volume