Planning order releases for networks of work centers: Queueing-theoretical foundation of meta-models of material flows

Hubert Missbauer

Department of Information Systems,
Production and Logistics Management

University of Innsbruck

A-6020 Innsbruck, Austria

Contents

Characteristics of the planning and control problem under consideration

- Discrete manufacturing
  Examples: mechanical engineering, production of electronic devices, furniture, etc.
- Usually a multi-stage production; each stage is a network of work centers
- Stock keeping units (SKUs; components or final products) are produced in lot sizes >= 1
- Complexity leads to a hierarchical manufacturing planning and control system (see next slide)
- =>The network of work centers constitutes a *production unit*:
  - receives production targets from top level
  - is autonomous with respect to detailed planning (scheduling, sequencing)
- Top (goods flow control) level determines the targets for the production units: order release function (see next slide)

Hierarchical structure of a typical manufacturing planning and control system

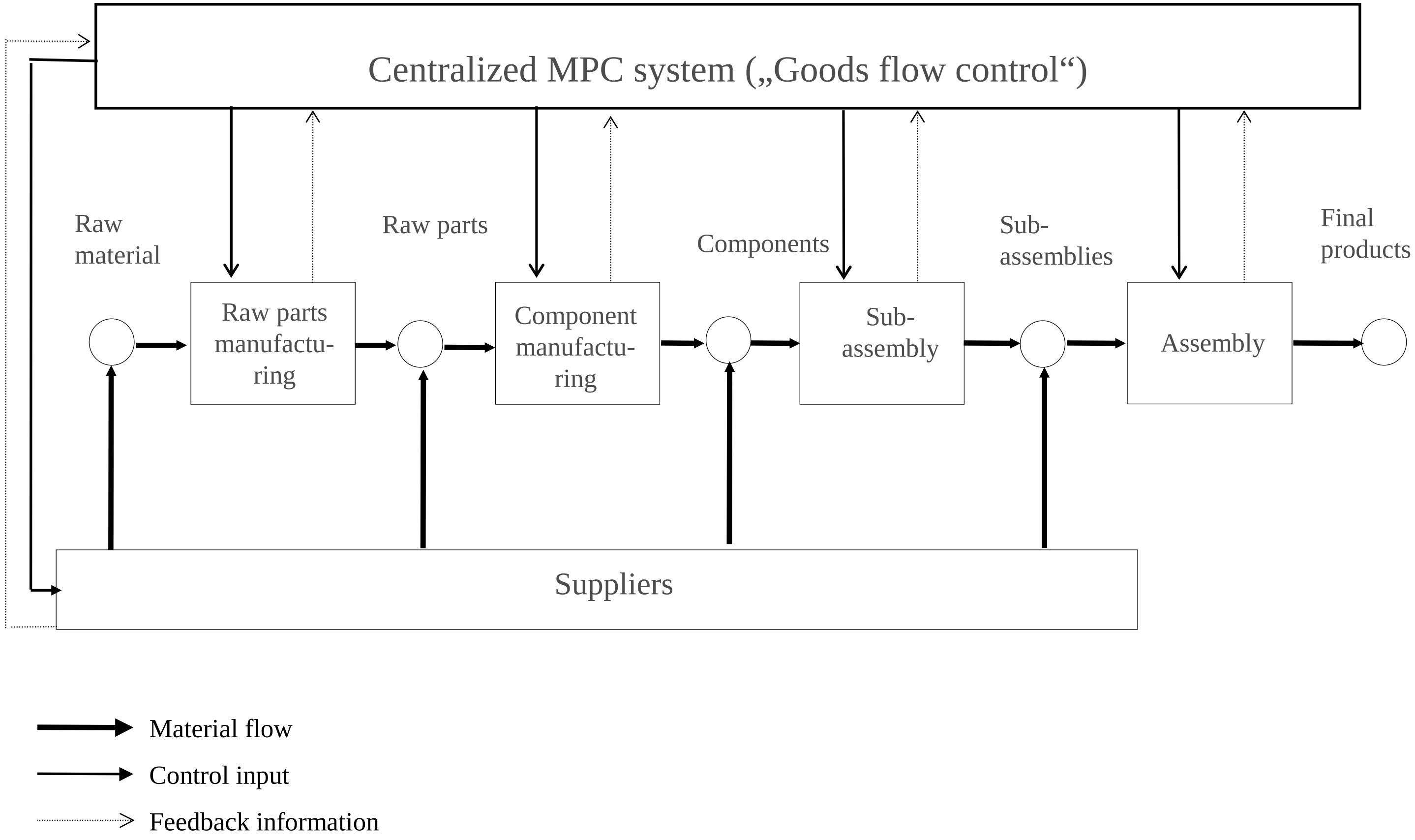
Top level ("Goods flow control"):

Planning and control of the material flow through the entire logistic chain, including capacity planning, at an appropriate level of aggregation.

Base level ("Production unit control"):

Detailed scheduling of the orders within the production units, usually performed at the shop floor level and for each production unit separately.

Structure of the hierarchical MPC system under consideration



Hierarchical structure of a typical manufacturing planning and control system

Top level ("Goods flow control"):

Planning and control of the material flow through the entire logistic chain, including capacity planning, at an appropriate level of aggregation.

Base level ("Production unit control"):

Detailed scheduling of the orders within the production units, usually performed at the shop floor level and for each production unit separately.

Interface:

- Order release decisions
- Foreknowledge of flow times – lead times for order release planning

Decision problem of determining

- Order releases over time,
- Required output over time
- Load-dependent lead times over time

Simultaneously

(Alternative: e.g., Selçuk 2007)

Order release function: a practical example

- Production of optical storage media (CDs, DVDs);
  ~300.000 discs/day
- Three-stage production: (1) producing the discs, (2) printing,
  (3) packing
- Make-to-order manufacturer, 24 hours a day, 7 days a week
- Flow time per stage ~ 1 day, operation times ~ 1-2 hours
  => 90-95% of the flow time is queueing time!
- High short-term fluctuations of the demand;
  Non-replacing machines in stages 1 (production) and 2 (printing)
  => temporary bottlenecks!

⇒ large number of small orders makes a hierarchical MPC system favorable (fluid approximation)!

⇒ Importance of the order release function!

Tasks of the order release function

- Adjusting WIP - flow time and output (utilization) over time, considering the highly non-linear relationship and non-stationarity
- Load levelling over time
- Load balancing among work centers

Load levelling and load balancing require deferring and pulling forward of orders.

Desired properties of the order release model for one production unit

- Order release volumes and output over time are essential decision variables.
- Modelling the non-stationary behaviour of production unit is necessary.
- Should be based on microscopic theory of discrete material flows through networks of work centers.
  => based on theory of transient queuing networks.
- Informal shop floor control rules make analytical description of arrival and departure processes virtually impossible.
  => Meta-modelling approach, based on theory of transient queuing networks.
- Special emphasis on non-linear relationships between WIP, flow time and output.

Generic model of a production unit for aggregate order release planning

Problem

- Input: Demand for (possibly aggregate) products j (D_{jt}).
- Result: WIP for all (aggregate) products j and work centres m (W_{jmt}), released work (R_{jt}), output (X_{jmt}), final product inventory (I_{jt}).

Structure of the model

$$\underbrace{W_{jmt}}_{\text{WIP}} = W_{j,m,t-1} + \underbrace{A_{jmt}}_{\text{Input}} - \underbrace{X_{jmt}}_{\text{Output}} \qquad \forall j,m,t$$

$$\underbrace{A_{jmt}}_{\text{Input}} = \underbrace{\sum_{i=1}^{M}\sum_{\tau=0}^{\infty} X_{j,i,t-\tau}\,\tilde{p}_{jim}\,z_{jim\tau}}_{\text{Work arriving from other w.c.'s}} + \underbrace{\sum_{\tau=0}^{\infty} R_{j,t-\tau}\,\tilde{p}_{j0m}\,z_{j0m\tau}}_{\text{Work arriving from release}} \qquad \forall j,m,t$$

$$\underbrace{I_{jt}}_{\text{Fin.prod.inv.}} = I_{j,t-1} + \underbrace{\sum_{m=1}^{M}\sum_{\tau=0}^{\infty} X_{j,m,t-\tau}\,\tilde{p}_{jm0}\,z_{jm0\tau}}_{\text{PU output}} - \underbrace{D_{jt}}_{\text{Demand}} \qquad \forall j,t$$

$$\sum_{j=1}^{J}\sum_{m=1}^{M}\sum_{t=1}^{T}\hat{W}_{jmt}\cdot h_{jmt} + \sum_{j=1}^{J}\sum_{t=1}^{T}\hat{I}_{jt}\cdot l_{jt} \to Min!$$

Nonlinear, saturating Clearing Function: The basic idea

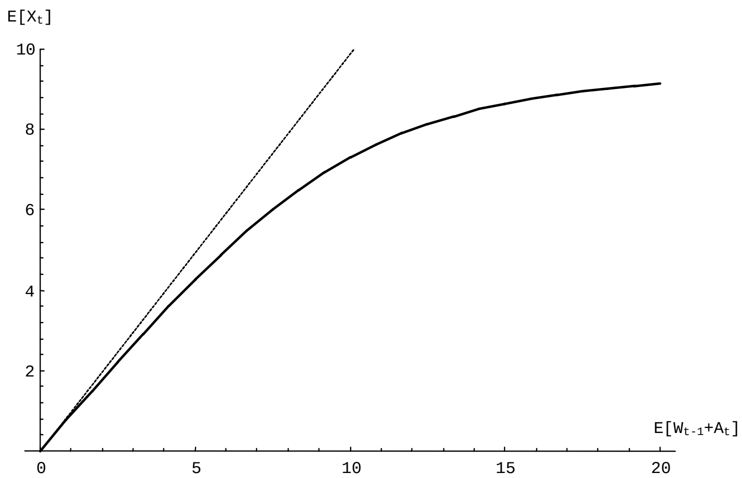
- Basic definition of a clearing function

  $X_{it} <= f_i(\text{WIP measure}_t)$

- Average WIP as WIP measure leads to multiple optima and oscillating release volumes.
  More appropriate: average WIP over periods (Asmundsson et al. 2002); total available work (Karmarkar 1989, Missbauer 2002)

- Saturating clearing function with load (available work) as WIP measure:

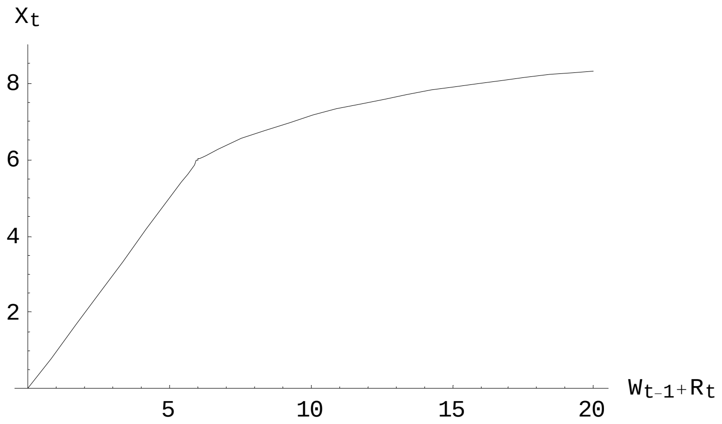
  $X_{it} <= f_i(W_{i,t-1}+A_{it};\, C_{it})$



Nonlinear, saturating Clearing Function: Some details

- Karmarkar (1989) formulation for a single work center:

$$X_t = f(W_{t-1}, R_t, C_t) = \min\left[C_t \frac{W_{t-1} + R_t}{W_{t-1} + R_t + k}; W_{t-1} + R_t\right]$$

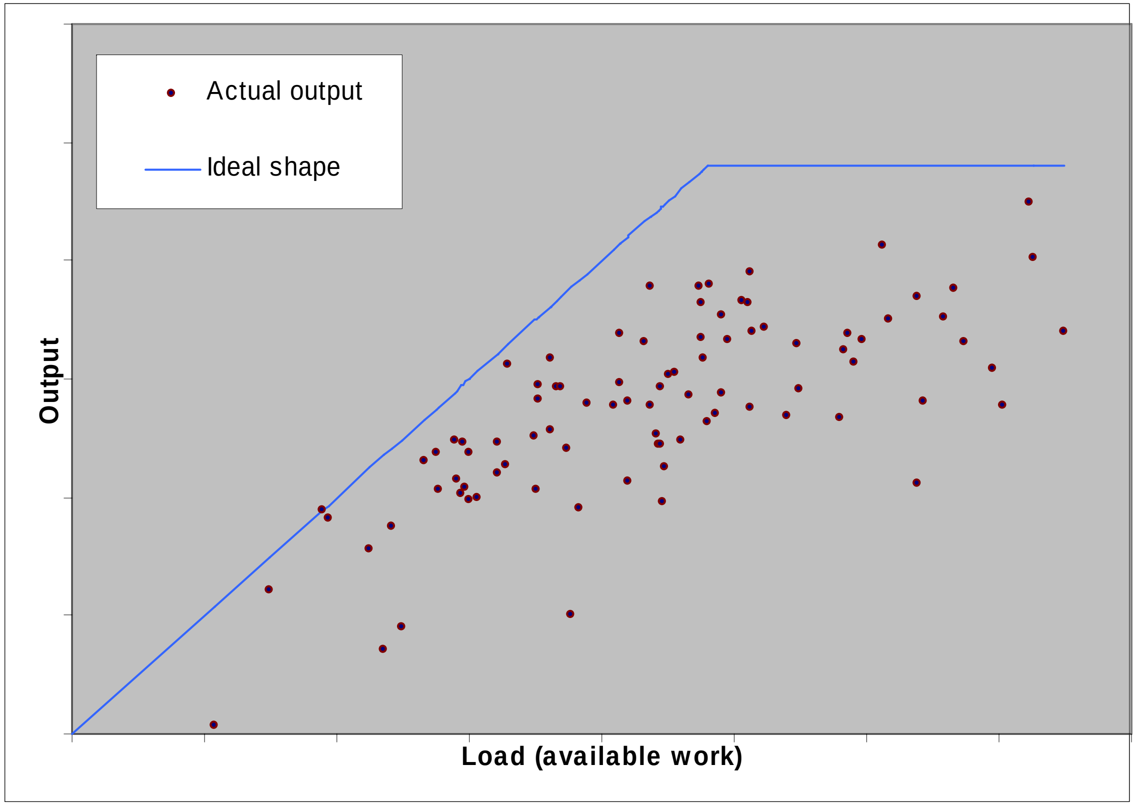


- Function derived from M/G/1 model (Missbauer 1998, 2002):

$$E(X_t) = \frac{1}{2}\left[C + k + E(L_t) - \sqrt{C^2 + 2Ck + k^2 - 2C\,E(L_t) + 2k\,E(L_t) + E(L_t)^2}\right]$$

$$\text{M/G/1:} \quad k = \frac{\mu\sigma^2}{2} + \frac{1}{2\mu}$$

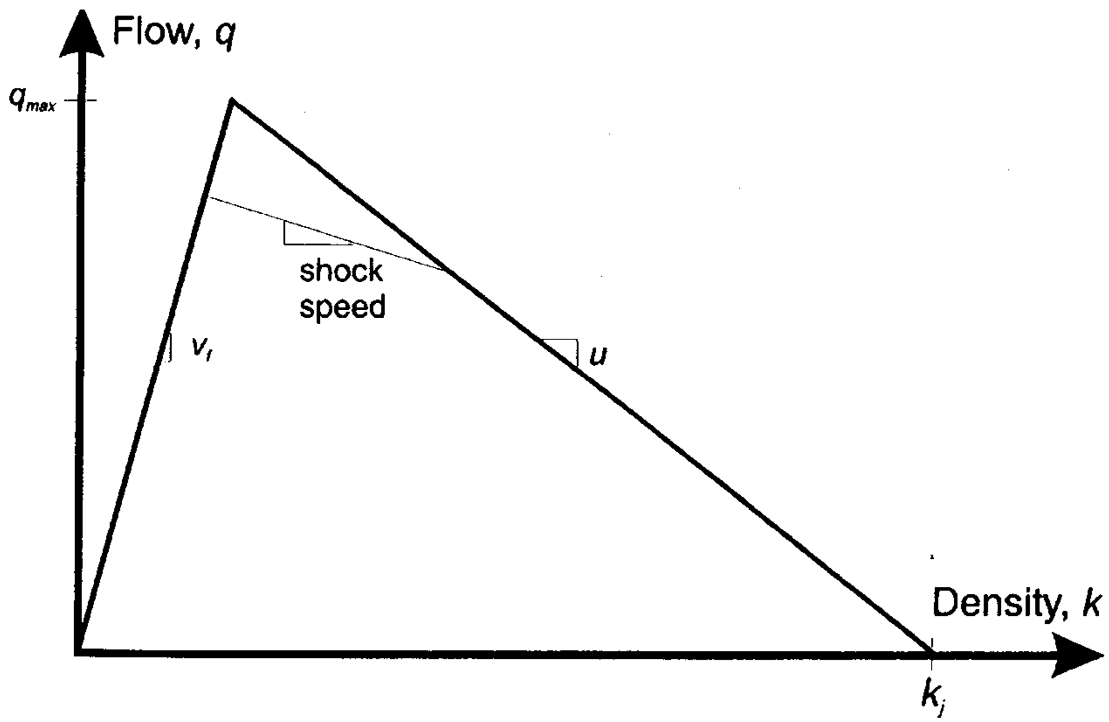
Data for empirical estimation of clearing function parameters



Days of one month, printing of optical storage media

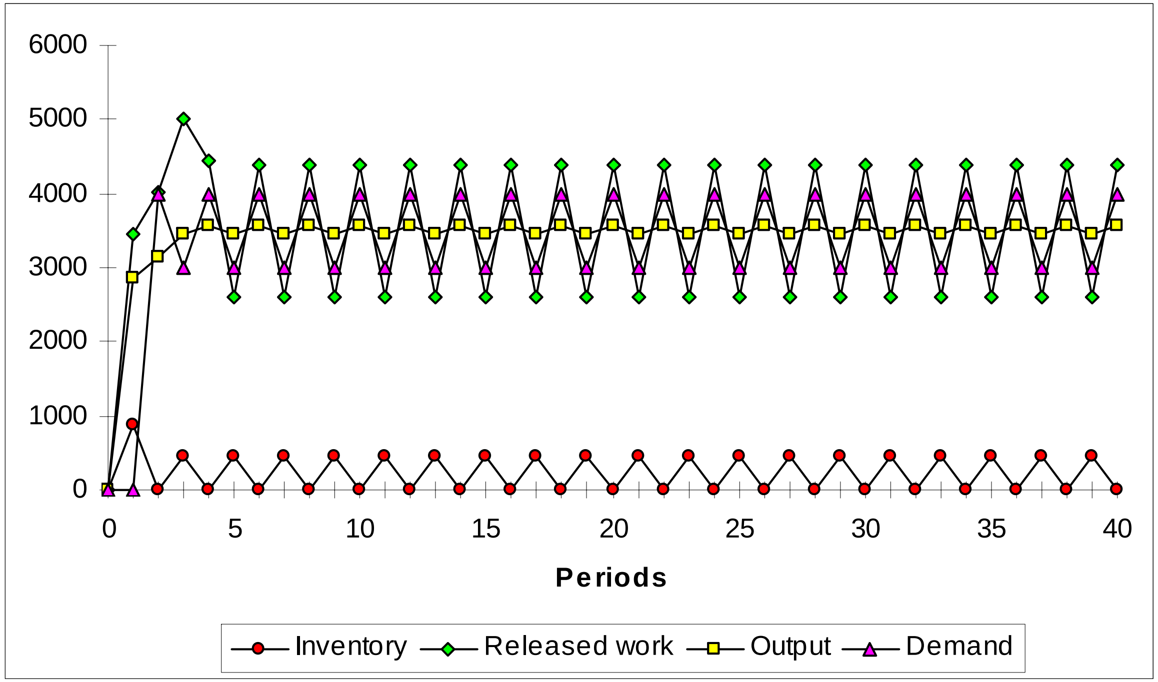
Output decreasing with WIP increase: Flow-density-relation for a traffic link

Source: Cassidy 2003, p. 183

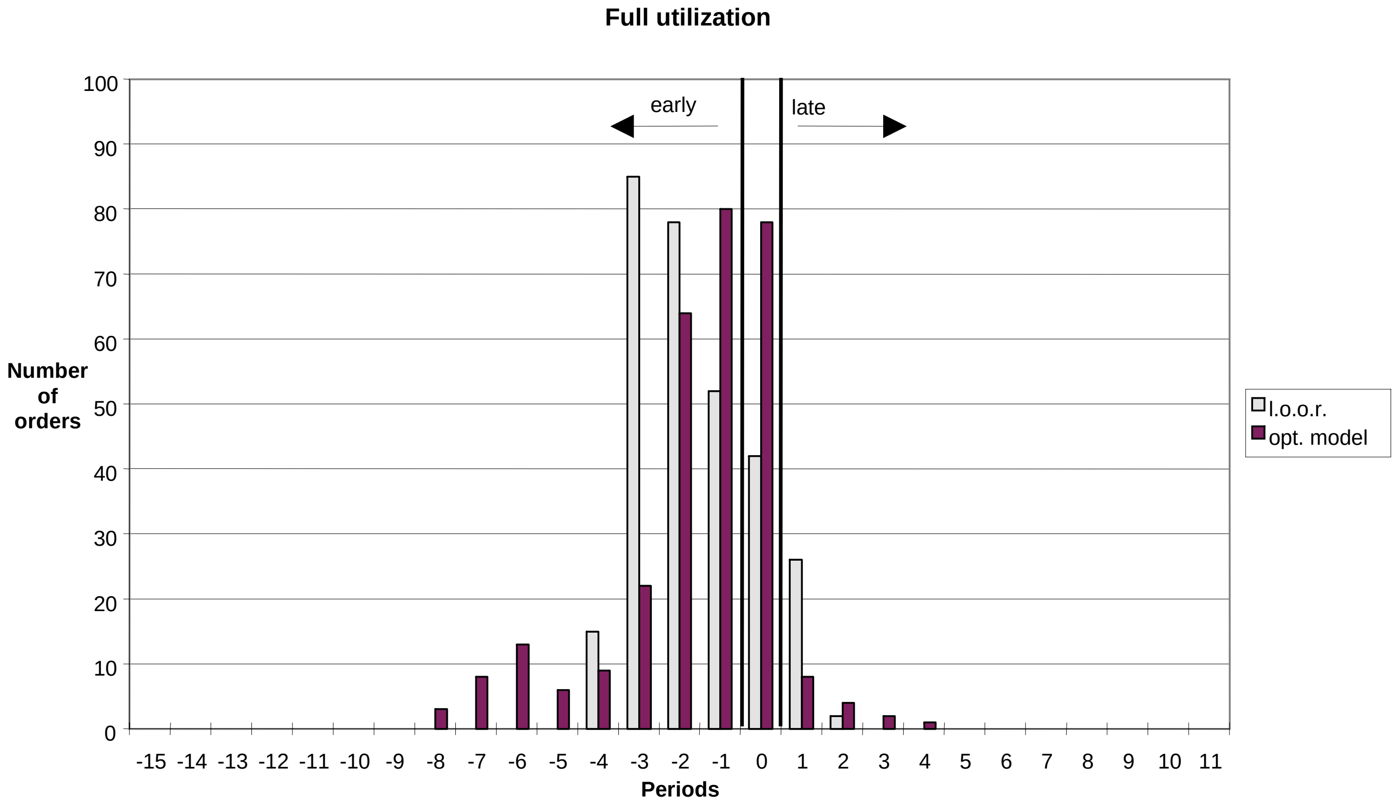


Dynamic behaviour of Clearing Function models: Optimization result, 1 work center

Source: Missbauer 1998



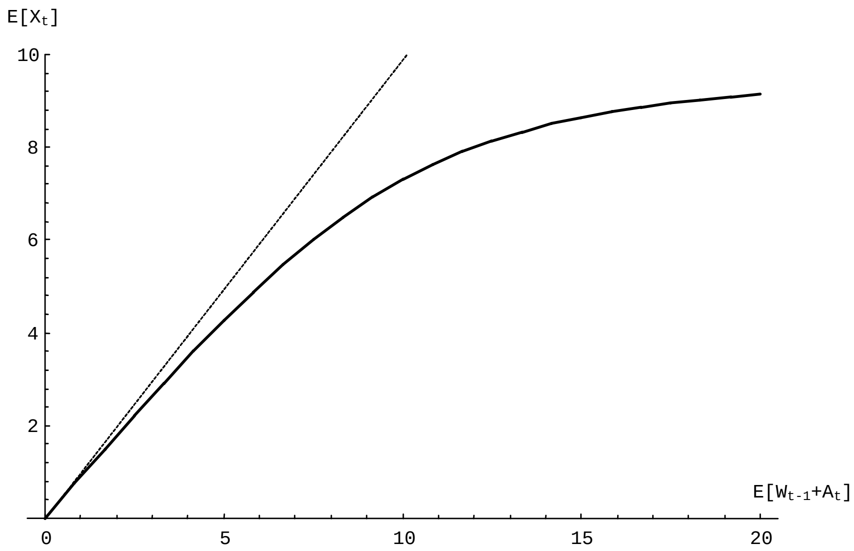
Dynamic behaviour of Clearing Function models: Simulation result, due-date deviation, job shop



Clearing Function as a steady-state property

- Saturating clearing function with load (available work) as WIP measure:

  $X_{it} <= f_i(W_{i,t-1}+A_{it};\ C_{it})$



Main insights of the research

- Theorem 1

  Clearing function is a steady-state property of queueing systems

  => Inconsistent with order release models.

- Theorem 2

  In the transient state: Output X_{it} must be modelled as a function of at least three independent variables: Planned initial WIP ($W_{i,t-1}$), Planned Input (A_{it}), Variance of initial WIP.

- Theorem 3

  Usual procedure for parameter estimation of clearing functions (regression) leads to a biased result. Clearing function is too „optimistic“.

- Theorem 4

  The bias is well-defined: Regression function („empirical clearing function“) can be corrected in a well-defined way to yield the clearing function for order release planning.

Transient Clearing Function: output estimation from the "estimated past"

- Clearing Function is a steady-state property
  => Logical contradiction when applied for order release planning

- Extension: two-dimensional Clearing Function (Andersson et al. 1981)

$$X_t = f(W_{t-1}, A_t)$$

- Transient Clearing Function: Output estimation from the "estimated past". General model:

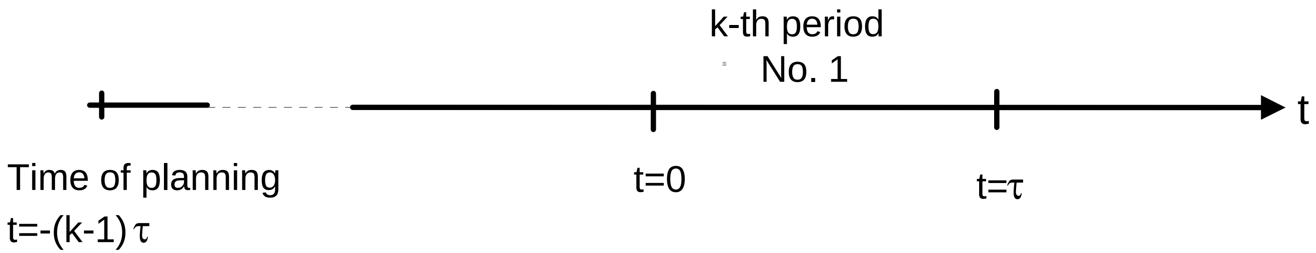
$$X_{jmt} \leq f_{jm}(A_{\hat{j}m\tau}, X_{\hat{j}m\tau}[\hat{j}=1,...,J;\tau=1,...,t-1]; A_{\hat{j}mt}[\hat{j}=1,...,J]) \qquad \forall j,m,t$$

with A_{jmt} the input in period t

- Research question: optimal number and definition of independent variables for the output estimation!

Transient M/M/1 model in the context of the order release planning model

One work center, index i is omitted



For given model variables before period under consideration:

- W_0 is planned WIP, interpreted as Expected value $E[W_0]$
- A_1 is planned input, interpreted as Expected value $E[A_1]$
- For exposition, W_0 and A_1 are measured as number of orders. $W_0 = Ls_0$

Formulation of the clearing function

$p_{ij}(t)$ conditional probability of *j* customers in the system at time *t* given *i* customers in the system at time 0 for given arrival and service rate (for the transient M/M/1 model, see Cohen 1969)

Expected output of the period k=1 for deterministic initial WIP W_0:

$$E[X_1] = \tau - \int_{t=0}^{\tau} p_{W_0 0}(t)\,dt$$

For stochastic initial WIP W_0 with distribution $p_n(0)$:

$$p_0(t) = \sum_{i=0}^{\infty} p_i(0).p_{i0}(t) \qquad 0 < t \le \tau$$

Expected output, measured in units of time:

$$E\left[X_1\right] = \tau - \int_{t=0}^{\tau} p_0(t)dt = \tau - \int_{t=0}^{\tau} \sum_{i=0}^{\infty} p_i(0).p_{i0}(t)\,dt$$

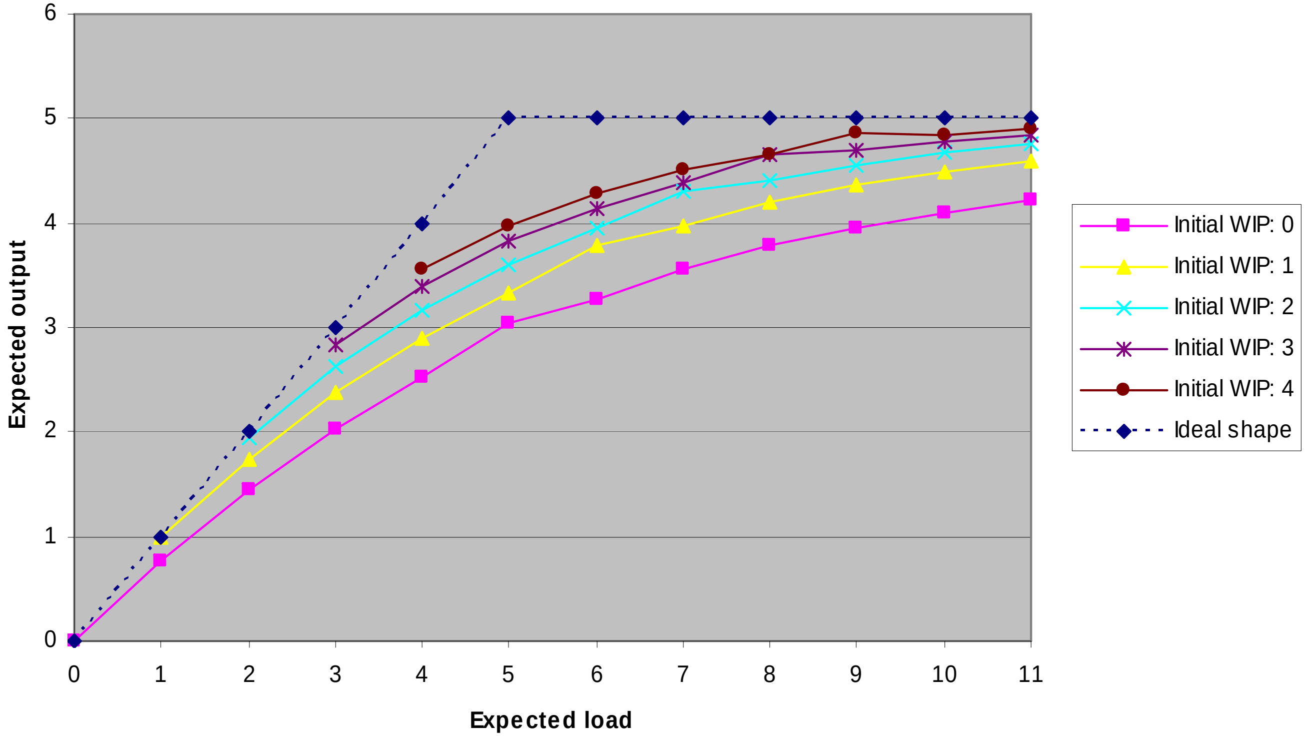
Clearing function – numerical example

- Method of the proof: Numerical counterexample for the hypothesis: „One-dimensional (traditional) clearing function or two-dimensional clearing function $X_{it}=f_i(W_{i,t-1}; A_{it})$ is sufficient“.
- Period length τ=5, service rate μ=1, input rate λ.
- Expected load (available work) in the period with workload measured in orders:

$$E\left[L_1\right]=E\left[Ls_0\right]+\lambda.\tau$$

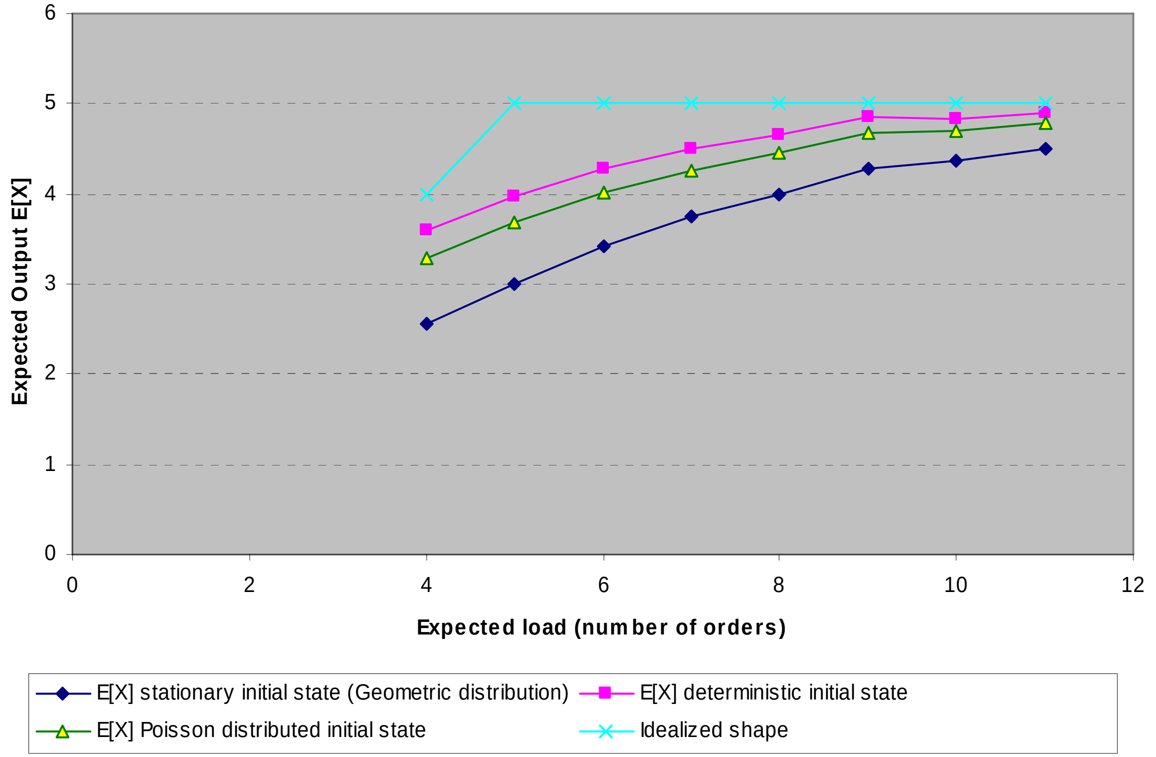
- Research questions:
  - Impact of the composition of the expected load given a one-dimensional clearing function
  - Impact of the distribution of the initial WIP Ls_0:
    - Deterministic (first period)
    - Geometric distribution (steady-state)
    - Poisson distribution (Pulse input with given mean)

Clearing function – impact of composition of the expected load



Clearing functions for period 1 for different deterministic initial WIP
Period length τ=5, service rate μ=1, input rate λ.

Clearing function – impact of distribution of initial WIP



Clearing functions for period 1 for different distributions of the initial WIP

Period length τ=5, service rate μ=1, E[Ls_0]=4, input rate λ.

Conclusions so far

- Theorem 1

  Clearing function is a steady-state property of queueing systems

  => Inconsistent with order release models.

  Proven!

- Theorem 2

  Transient state: Output X_{it} must be modelled as a function of at least three independent variables: Planned initial WIP ($W_{i,t-1}$), Planned Input (A_{it}), Variance of initial WIP.

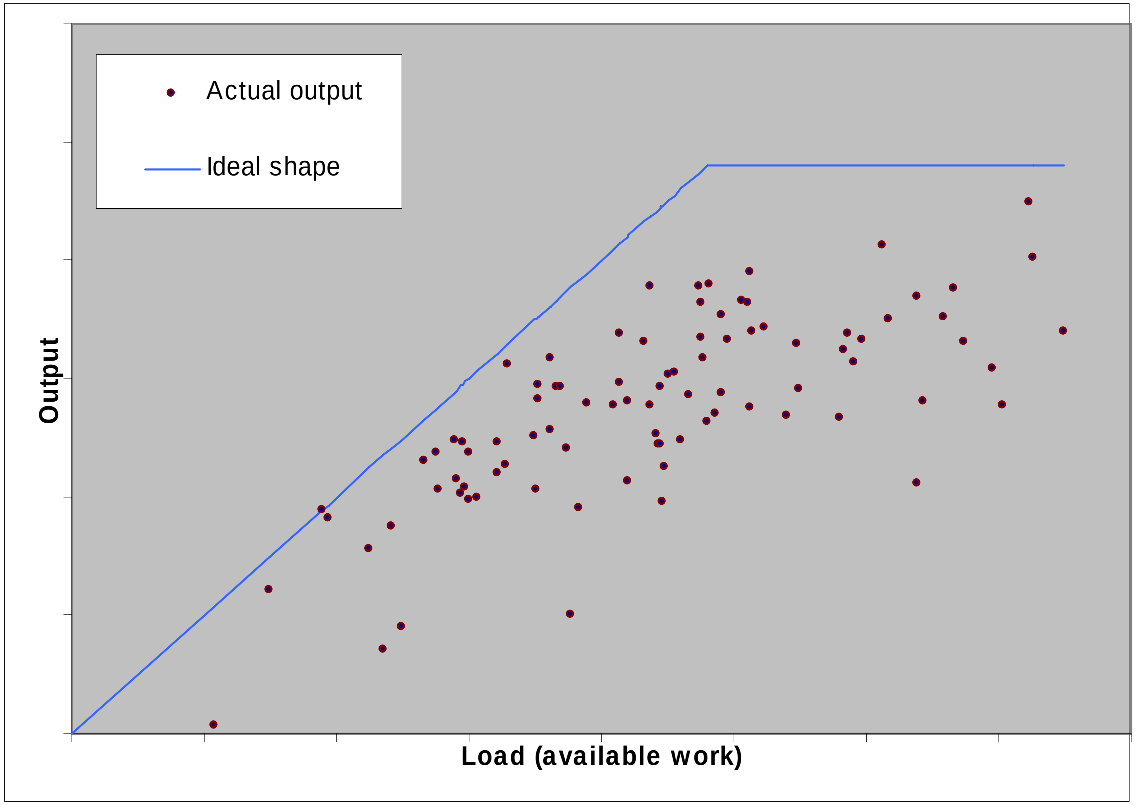
  Proven!

Remaining question:

- How to handle the (non-observable!) Variance of the initial WIP ($W_{i,t-1}$)?

=> First we analyze the empirical parameter estimation for the clearing function

Data for empirical estimation of clearing function parameters



Days of one month, printing of optical storage media

Empirical parameter estimation vs. order release planning

	Order release planning	**Empirical data**
Initial WIP W_{t-1}	stochastic	deterministic
Input A_t	stochastic	deterministic
Timing of input	stochastic	stochastic

=> Theorem 3

Usual procedure for parameter estimation of clearing functions (regression) leads to a biased result.

Proven! Clearing function is too „optimistic“.

Correction of the empirical clearing function $E[X_k]=f(W_{k-1}; A_k)$

Definition of the empirical clearing (=regression) function

$E[X_k \mid W_{k\text{-}1} = w;\ A_k = y]$ Conditional expectation of the output in period *k* conditioned the initial WIP $W_{k\text{-}1} = w$ and the input during the period *k* $A_k = y$.

Additional information:

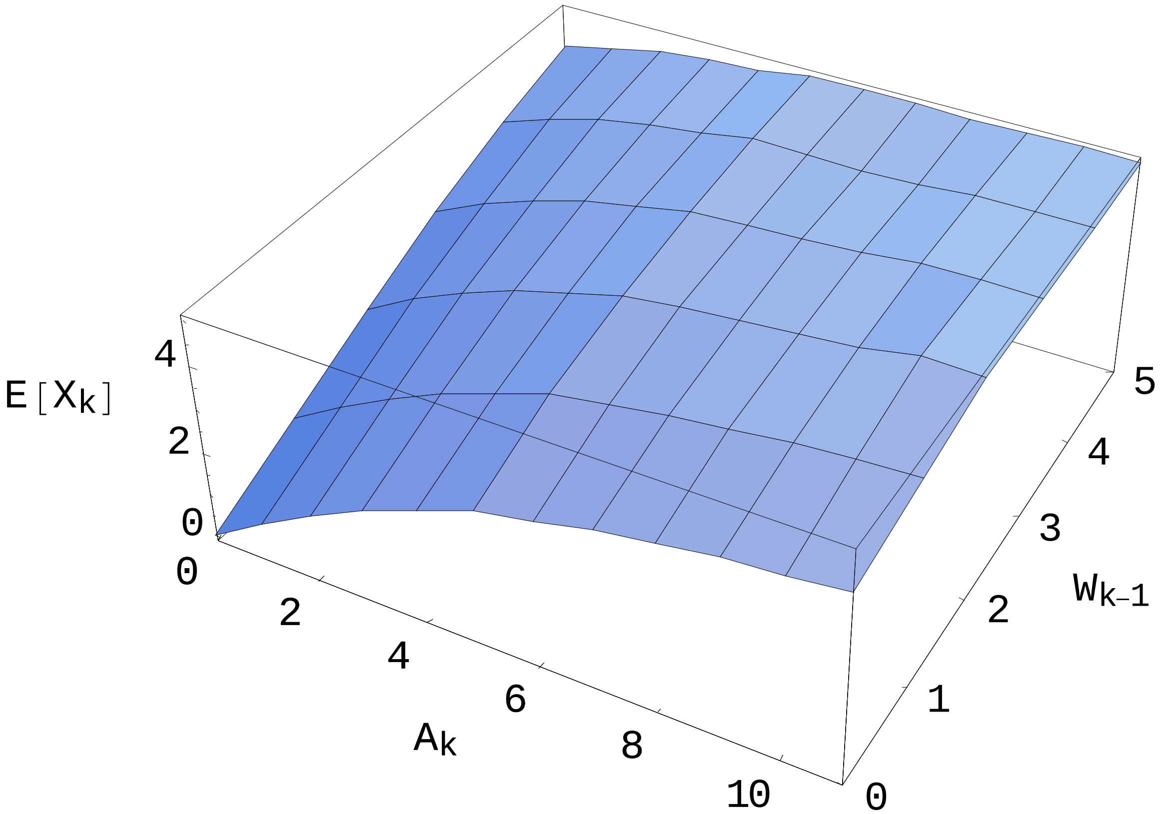
$f_{W_{k-1},A_k}(w,y)$ Joint distribution: Probability of initial WIP $W_{k\text{-}1} = w$ and input $A_k=y$.

The expected output during period *k* (clearing function) is:

$$E[X_k] = \sum_{w=0}^{\infty}\sum_{y=0}^{\infty} E[X_k \mid W_{k\text{-}1} = w;\ A_k = y]\ .f_{W_{k-1},A_k}(w,y)\ .$$

=> Theorem 4 proven!

Two-dimensional clearing function for the above M/M/1 system



Correction for Geometric distribution of W_{k-1}, Poisson distribution of A_k

	$A_k=0$	$A_k=1$	$A_k=2$	$A_k=3$	$A_k=4$	$A_k=5$	$A_k=6$	$A_k=7$	$A_k=8$	$A_k=9$	$A_k=10$
$W_{k-1}=0$	Indeterminate	0.0559	0.0595	0.0595	0.056	0.0828	0.0498	0.0471	0.0414	0.0345	0.0269
$W_{k-1}=1$	0.0732	0.0909	0.0967	0.0969	0.0927	0.111	0.0829	0.0776	0.0703	0.0624	0.0547
$W_{k-1}=2$	0.17	0.164	0.154	0.143	0.131	0.138	0.107	0.0967	0.0853	0.0744	0.0992
$W_{k-1}=3$	0.248	0.224	0.201	0.179	0.157	0.155	0.12	0.105	0.0905	0.098	0.0815
$W_{k-1}=4$	0.303	0.266	0.233	0.202	0.173	0.165	0.127	0.0998	0.0886	0.0892	0.0753
$W_{k-1}=5$	0.343	0.295	0.254	0.217	0.184	0.177	0.148	0.125	0.106	0.0904	0.0785

Relative difference, defined as:
(empirical value – corrected value)/empirical value

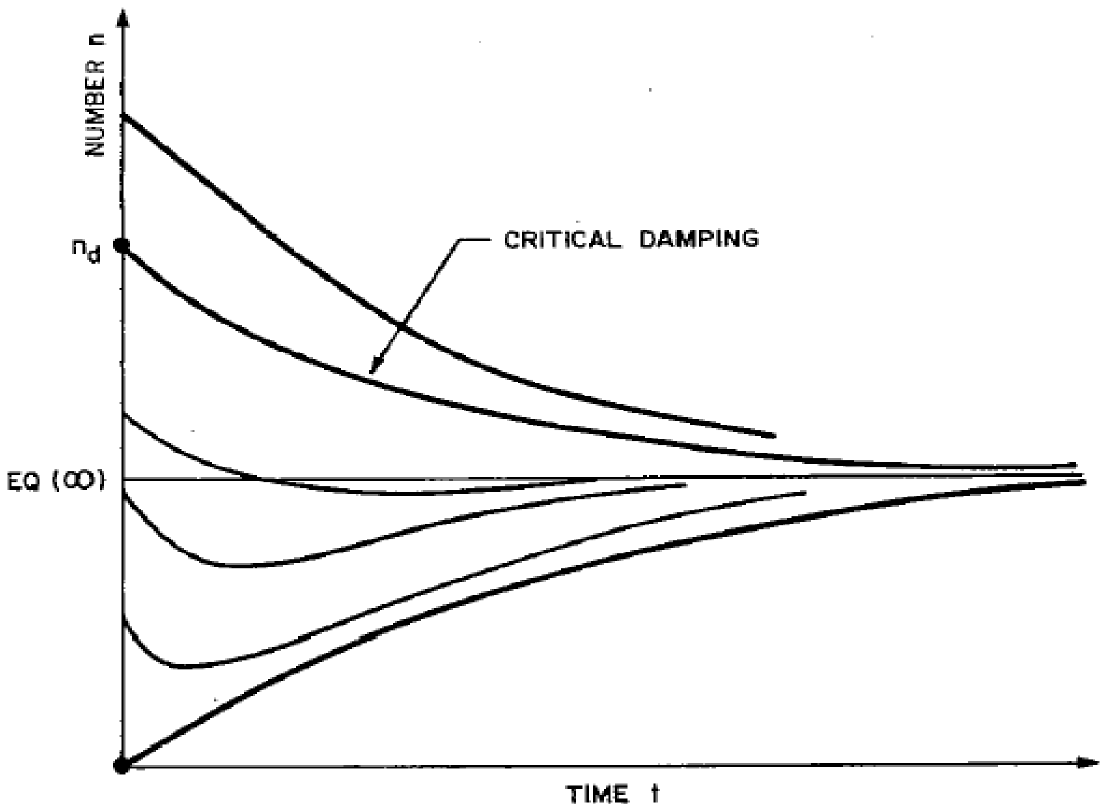
Estimating the joint distribution $f_{W_{k-1},A_k}(w,y)$

- Joint distribution $f_{W_{k-1},A_k}(w,y)$ is not observable!
- Planned values W_{k-1}, A_k can be interpreted as expected values.
- Assumption of stochastic independence is problematic, but may open a way.
- Abate and Whitt (papers 1987-1994): Approximations of the time-dependent values of the first and second moment of W(t) for engineering applications.

  Convergence to steady-state values can be approximated by a mixture of exponentials!

E[Ls(t)] as a function of t for different Ls(0)=n

Source: Abate and Whitt 1987, p. 43



Conclusions and research topics

- Clearing function models should be improved w.r.t. (1) transient effects and (2) parameter estimation.
- Desired result: Theoretically consistent way to handle load-dependent lead times in order release planning.
- A lot of research efforts remain to be done:
  - Formulation Var[WIP] over time
  - Formulation of the optimization model
  - Optimization technique
  - Extensive testing
- Linking mathematical analysis, empirical parameter setting and optimization is a big challenge.
- Still a lack of mathematical analysis and approximate models of material flows in stochastic systems.