Order release planning based on stochastic models of manufacturing systems

Hubert Missbauer

Department of Information Systems, Production and Logistics Management

University of Innsbruck

6020 Innsbruck, Austria

Contents

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

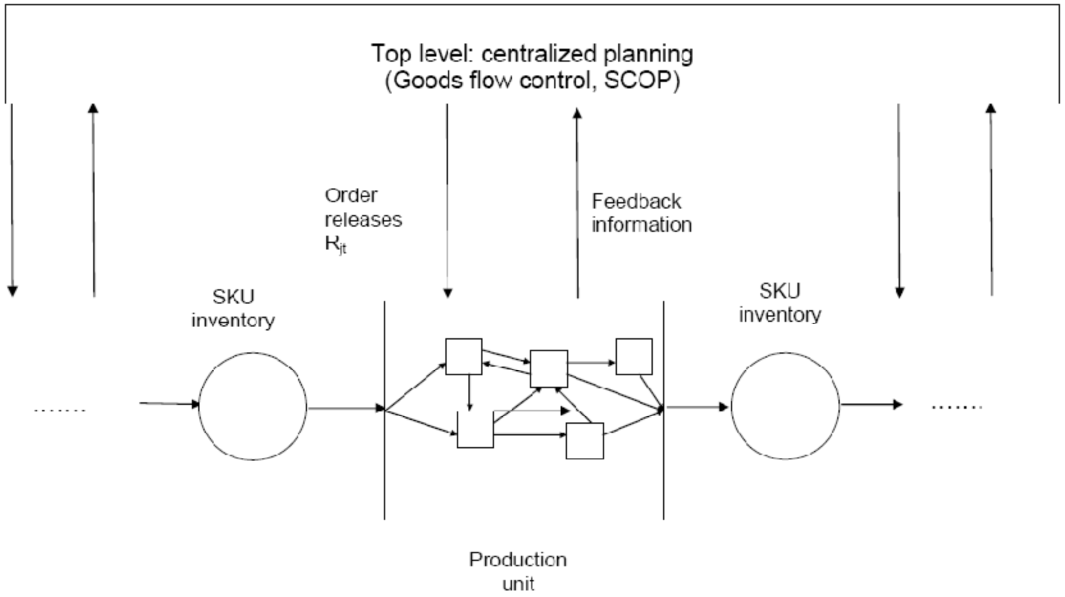
Decision problem of determining

- Order releases over time
- Output over time
- Load-dependent lead times over time

=>"Dynamic production function" (Hackman 2008):

$$R = (R_{jt}; \forall j, t) \xrightarrow{f} X = (X_{jt}; \forall j, t)$$

Order release planning in hierarchical MPC systems – conceptual model



- Production units are in a transient state
- Analytical description is hardly possible (e.g., „human aspects of scheduling“)
- Multiple products => discrete service time distribution
- Arrival process is determined by the goods flow control (SCOP) level.
  Possible properties: Once per period, autocorrelation, similar release intervals per product

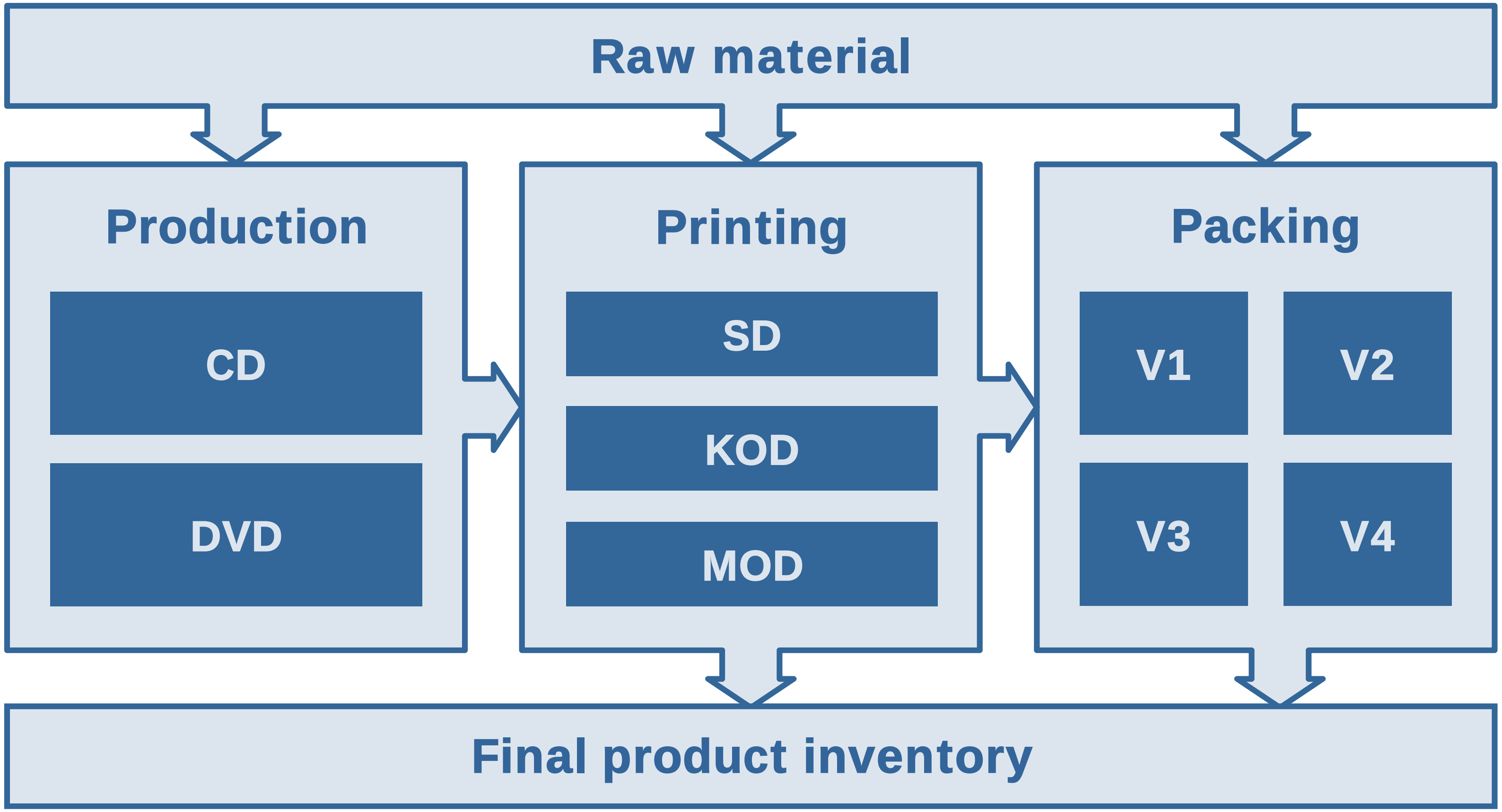
Order release function: a practical example

- 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-identical machines in stages 1 (production) and 2 (printing)
  => temporary bottlenecks => problems with utilization, due-date performance!

Goals of the project

- Reduction of flow times (flexibility!)
- => Improved coordination of manufacturing stages

Optical storage media: the production process



Approaches for order release planning for the production units

(Review: Missbauer / Uzsoy, in press)

Rule-based:

- Traditional order release mechanisms, e.g., LUMS release mechanism (Hendry/Kingsman), load-oriented order release (Wiendahl 1995)
- Many other variants of a „basic release procedure“ (Land 2004)

Optimization-based:

- Traditional fixed lead time models (Hackman/Leachman 1989)
- Input/Output control models
- Load-dependent lead time models (Voß/Woodruff 2006)
- Iterative approaches (Hung/ Leachman 1996)
- Clearing function models (Missbauer 2002, Asmundsson et al. 2006)

Aggregate model of a production unit

(Zäpfel / Missbauer 1993, based on Karmarkar 1989)

$$\sum_t h_t W_t + \sum_t H_t I_t \quad \rightarrow \text{ Min!}$$

Subject to

$$W_t = W_{t-1} + R_t - X_t \qquad t = 1, \ldots, T$$

$$I_t = I_{t-1} + X_t - D_t \qquad t = 1, \ldots, T$$

$$X_t \leq a_n + b_n \left(W_{t-1} + R_t \right) \qquad t = 1, \ldots, T; n = 1, \ldots, N$$

$$I_t, R_t, W_t, X_t \geq 0 \qquad t = 1, \ldots, T$$

Symbols (all measured in hours of work):

W_t	WIP at the end of period t
R_t	Work released in period t
X_t	Output (actual production) in period t
I_t	Finished goods inventory at the end of period t
D_t	Demand in period t (parameter)

Generic model of a production unit for aggregate order release planning

- Input: Demand for product groups j in periods t (D_{jt}).
- Result: Released work for all product groups j and periods t (R_{jt}); work-in-process (WIP) for all work centers m (W_{jmt}), output (X_{jmt}), final product inventory (I_{jt}).

Structure of the model:

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

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

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

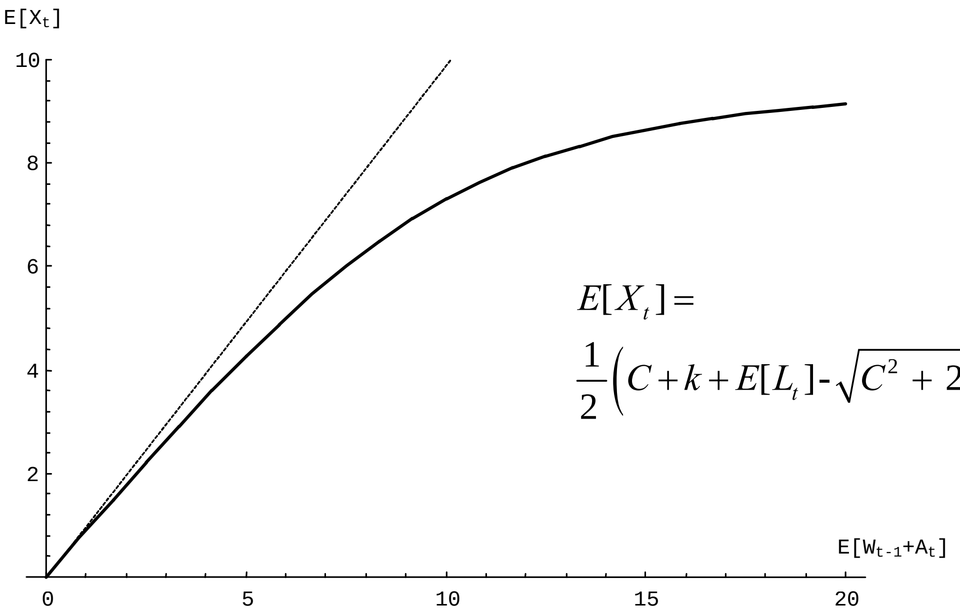
$$\sum_{j=1}^{J}\sum_{m=1}^{M}\sum_{t=1}^{T} \hat{W}_{jmt}\cdot h_{jmt} + \sum_{j=1}^{J}\sum_{t=1}^{T} \hat{I}_{jt}\cdot l_{jt} \rightarrow Min!$$

Nonlinear, saturating Clearing function

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

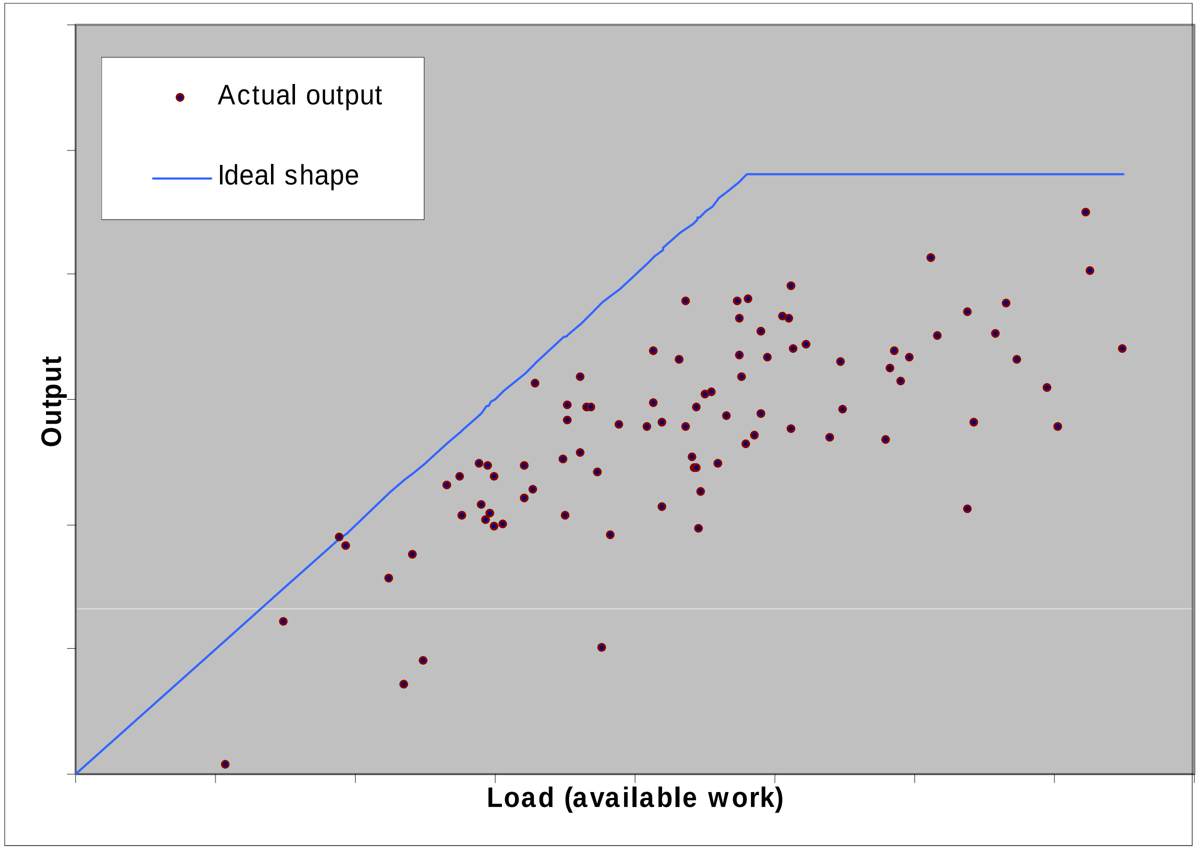
$$X_{it} = f_i(W_{i,t\text{-}1} + A_{it};\ C_{it})$$

- Shape of the function can be derived, e.g., from M/G/1 queueing model (Missbauer 2002):



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

Data for empirical estimation of clearing function parameters
Days of one month, manufacturing of optical storage media



Queueing-theoretical analysis of clearing functions – main insights

- Theorem 1

  If the model of a 1-dimensional clearing function is accepted, the usual procedure for parameter estimation of clearing functions (least squares regression) leads to a biased result. Clearing function is too „optimistic“.

  The bias can be corrected in a well-defined way.

  Proof: Missbauer 2010

- Theorem 2

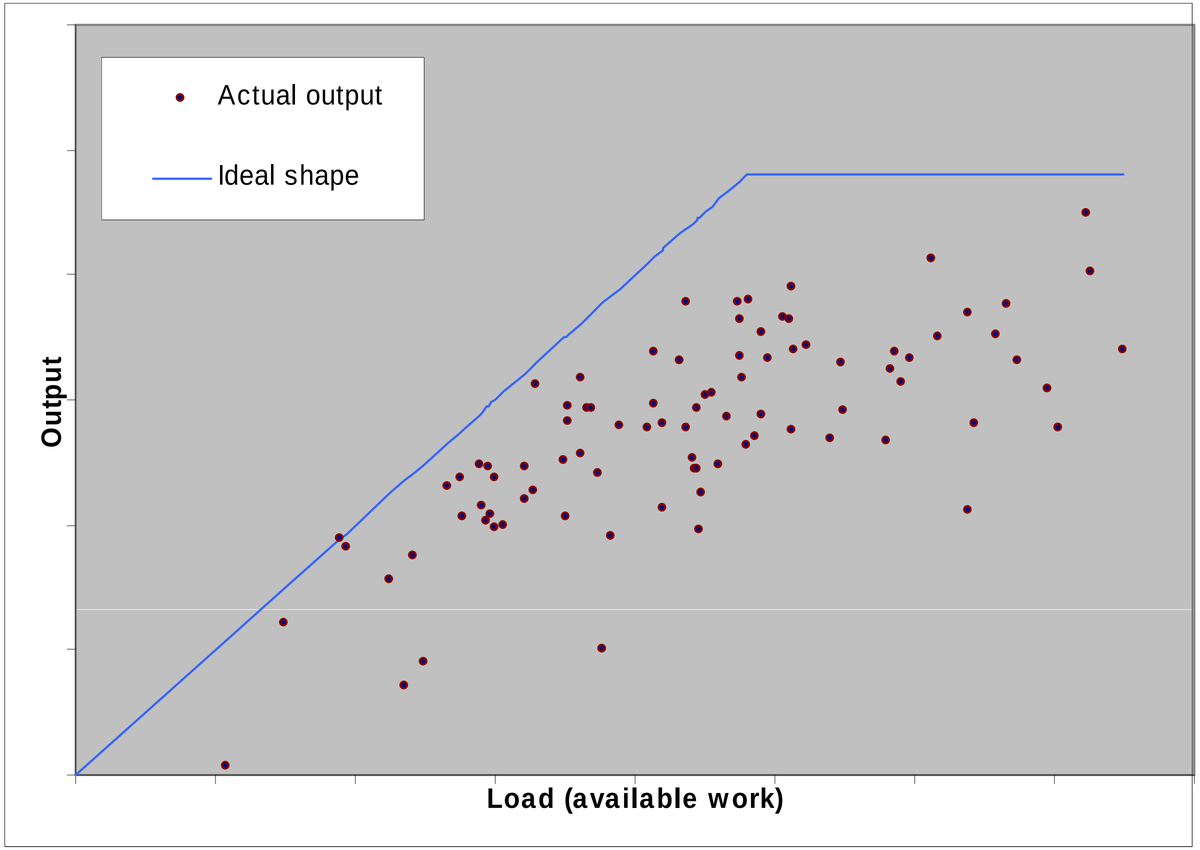
  Order release models aim at optimizing transient states.

  In the transient state: The output X_{it} must be modelled as a function of at least three independent variables:

  - Expected initial work-in-process level ($W_{i,t-1}$),
  - Expected input during the period (A_{it}),
  - Variance of the work-in-process level at the beginning of the period.

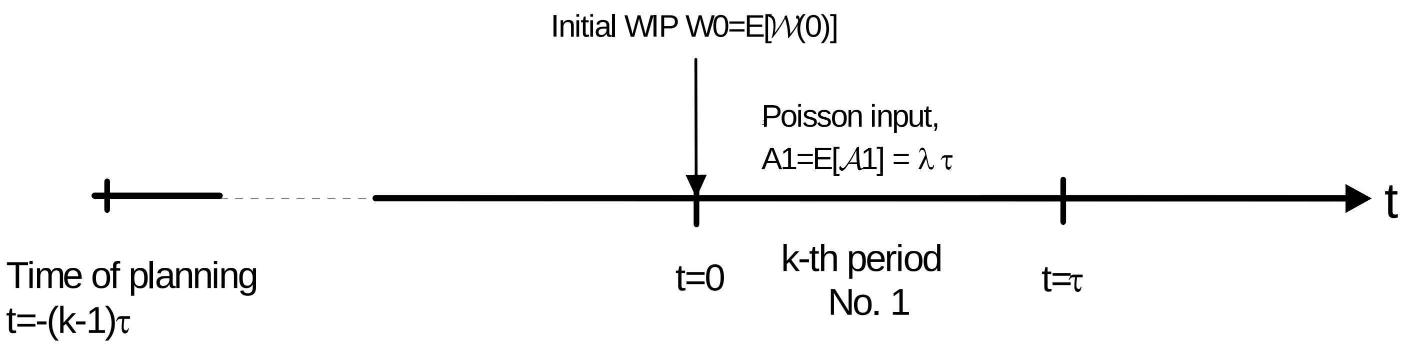
Data for empirical estimation of clearing function parameters

Days of one month, manufacturing of optical storage media



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

One work center, index i is omitted



For given model variables before the period under consideration:

- W_0 [time units] is the planned value, interpreted as Expected value E[$\mathcal{W}_0$]
- A_1 [jobs=expected time units; μ=1] is the planned value, interpreted as Expected value E[$\mathcal{A}_1$]

Formulation of the clearing function 1/3

$p_{rn}(t)$ Conditional probability of having n customers in the system at time t given r customers in the system at time 0 (Cohen 1969).

$$p_{rn}(t) = (1-\rho)\rho^n U(1-\rho) + \rho^{\frac{1}{2}(n-r)} e^{-(1+\rho)t/\beta} I_{n-r}\left(2\frac{t}{\beta}\sqrt{\rho}\right)$$

$$-\rho^{\frac{1}{2}(n-r)} \int_t^\infty e^{-(1+\rho)\tau/\beta} \left\{ \begin{array}{l} I_{r+n}\left(2\frac{\tau}{\beta}\sqrt{\rho}\right) - 2\rho^{\frac{1}{2}} I_{r+n+1}\left(2\frac{\tau}{\beta}\sqrt{\rho}\right) + \\ \rho I_{r+n+2}\left(2\frac{\tau}{\beta}\sqrt{\rho}\right) \end{array} \right\} \frac{d\tau}{\beta}, \qquad t \geq 0.$$

$I_j(x)$ .... modified Bessel function of the first kind

$$U(t) = \begin{cases} 0, & t < 0 \\ 1/2, & t = 0 \\ 1 & t > 0 \end{cases}$$

Formulation of the clearing function 2/3

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

$$E[X_1]^{detW0} = \begin{cases} W_0 + E[X_1^{\text{Rest}}] & \text{for } W_0 < \tau \\ \tau & \text{for } W_0 \geq \tau \end{cases}$$

Probability distribution of the number of jobs in system after initial WIP is processed ($p_n(W_0)$) for $\mu=1$:

$$p_n(W_0) = \frac{(\rho W_0)^n}{n!} e^{-\rho W}$$

Probability that the server is idle after $t=W_0$:

$$p_0(t) = \sum_{i=0}^{\infty} p_i(W_0).p_{i0}(t - W_0) \qquad W_0 < t \leq \tau$$

Hence: $$E[X_1^{\text{Rest}}] = \tau - W_0 - \int_{t=W_0}^{\tau} p_0(t)\,dt$$

Formulation of the clearing function 3/3

Expected output of the period k=1 for stochastic initial WIP with continuous distribution with p.d.f. $f_{W0}(w)$:

$$E[X_1]^{stochW0} = \int_{w=0}^{\infty} f_{W0}(w) E[X_1]^{detW0} \qquad \text{(note the } \delta\text{(t)!)}$$

Expected load (available work) in the period:

$$E\left[L_1\right] = E\left[W_0\right] + \lambda.\tau$$

Clearing function – numerical example

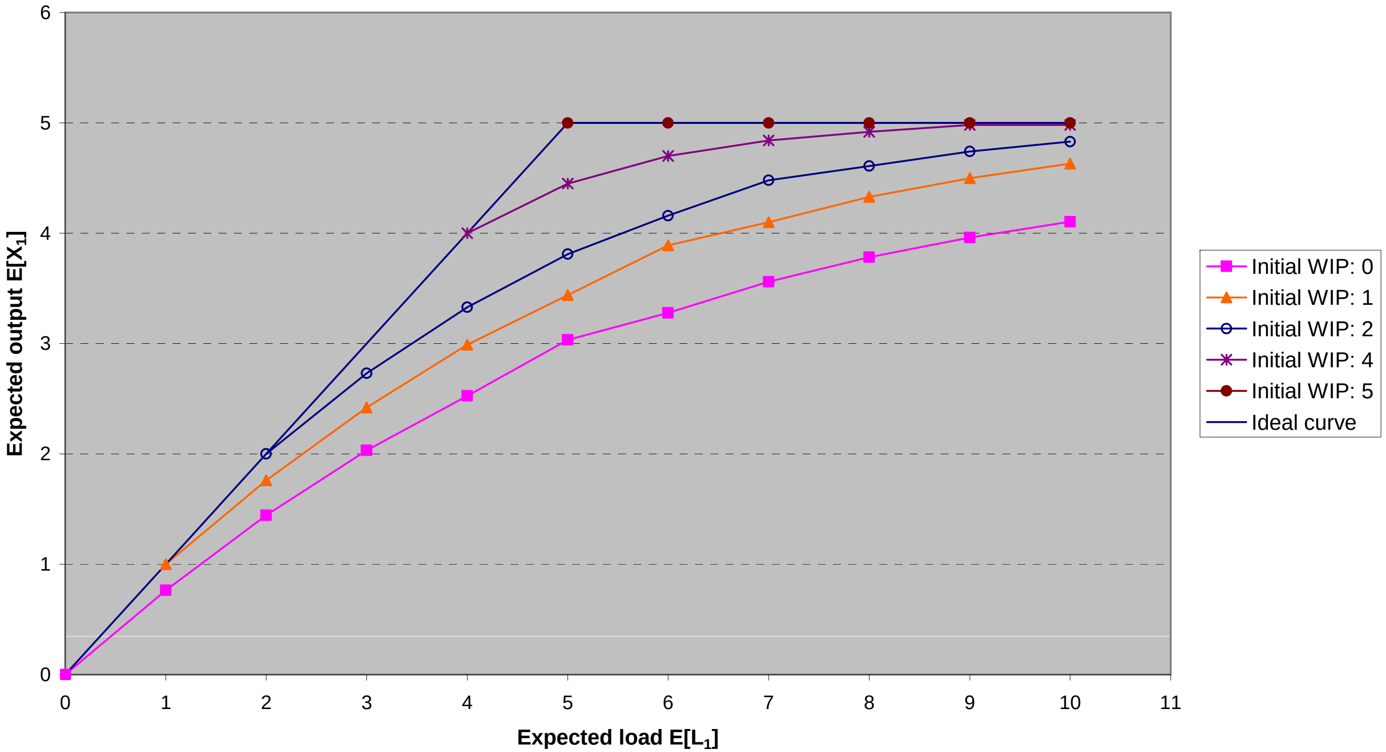
- Methodology of the proof: Numerical counterexample for the hypothesis: „The accuracy of a 1-dimensional (traditional) clearing function

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

is sufficient".

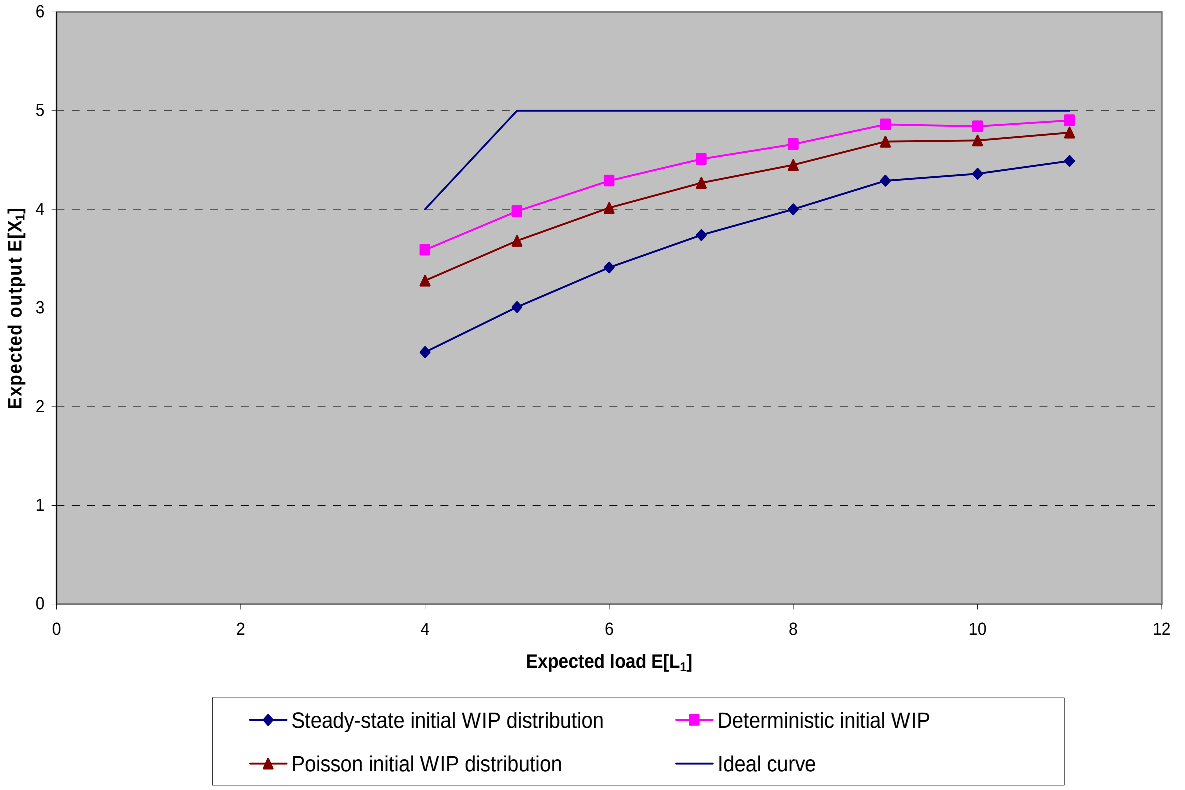
- Example: Period length $\tau=5$, service rate $\mu=1$, input rate λ.

Clearing function – impact of the composition of the expected load



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

Clearing function – impact of the discrete distribution of the initial WIP



Clearing functions for period 1 for different distributions of the initial WIP W_{t-1}.
$E[W_{t-1}]=4$ orders, period length $t=5$, service rate $\mu=1$, input rate λ.

Queueing-theoretical analysis of clearing functions – main insights

- Theorem 1

  If the model of a 1-dimensional clearing function is accepted, the usual procedure for parameter estimation of clearing functions (least squares regression) leads to a biased result. Clearing function is too „optimistic".

  The bias can be corrected in a well-defined way.

  Proof: Missbauer 2010

- Theorem 2 **proven**

  Order release models aim at optimizing transient states.

  In the transient state: The output X_{it} must be modelled as a function of at least three independent variables:

  - Expected initial work-in-process level ($W_{i,t-1}$),
  - Expected input during the period (A_{it}),
  - Variance of the work-in-process level at the beginning of the period.

Adjusted coefficient of determination R^2 for different independent variables in a clearing function (Häussler 2009)

Independent variable(s); shape of regression fct.	Machine with high R^2 in the linear regression (non-bottleneck)	Machine with low R^2 in the linear regression (bottleneck)
$W_{t-1} + A_t$, linear	0.81	0.56
$W_{t-1} + A_t$, steady-state M/G/1	0.81	0.62
W_{t-1}; A_t, linear	0.81	0.58
W_{t-1}; A_t, quadratic	0.82	0.63
W_{t-2}; A_t; A_{t-1}, first-order polynomial	---	0.68
W_{t-1}; A_t; A_{t-1}, first-order polynomial	0.84	---

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

Source: Abate and Whitt 1987, p. 43

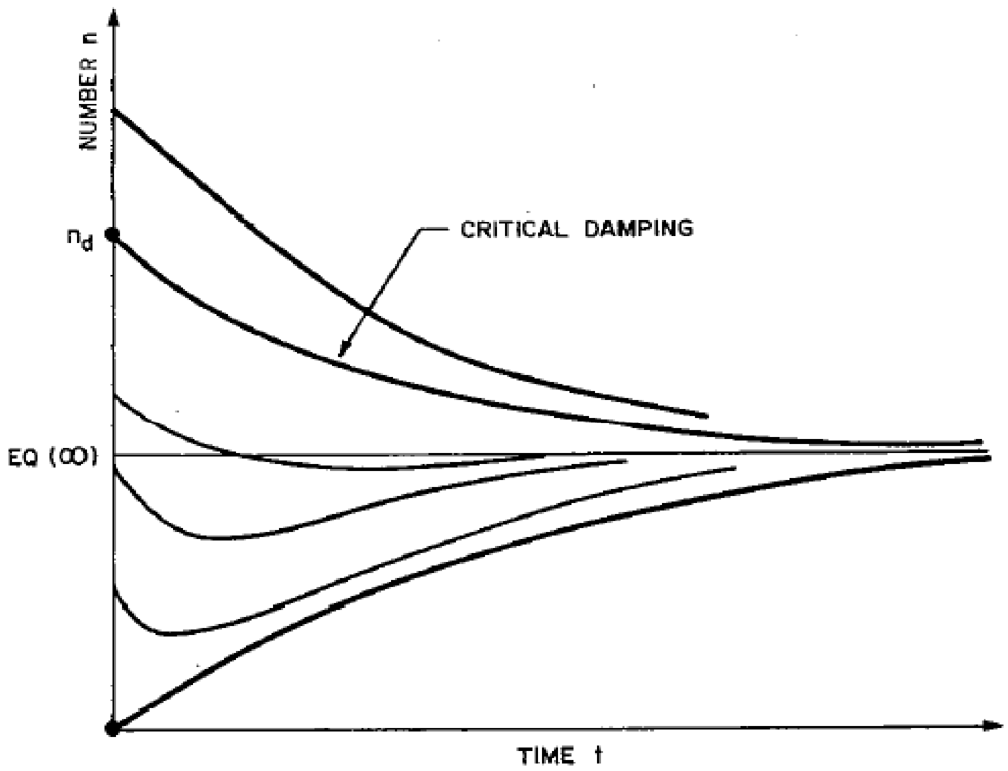


Fig. 1. $E(Q(t)\,|\,Q(0) = n)$ as a function of n and t.

Metamodel of time-dependent WIP – continuous time

If the time-dependent WIP at a work centre converges from a steady-state value W^{before} to the new finite steady-state value W^{after}, starting from time t_{d}, the following relationships hold:

$$W(t) = W^{before} + (W^{after} - W^{before})(1 - e^{-(t-t_d)/\beta_d}) \qquad \text{if } W^{after} > W^{before}$$

$$W(t) = W^{after} + (W^{before} - W^{after})e^{-(t-t_d)/\beta_u} \qquad \text{if } W^{after} < W^{before}$$

$$\frac{dW(t)}{dt} = A(t) - X(t)$$

$$\Rightarrow X(t) = A(t) - \frac{dW(t)}{dt} \quad ,$$

Metamodel of time-dependent WIP – discrete time

Substituting $W(t) = W_{mt}$, $W^{before} = W_{m,t-1}$, $t - t_d = 1$ (time unit is the period length), $W^{after} = E[W^{Amt}]$ yields:

$$W_{mt} = W_{m,t-1} + A_{mt} - C_{mt} \qquad \text{if } A_{mt} > C_{mt} \qquad (16)$$

$$W_{mt} = W_{m,t-1} + (E[W^{Amt}] - W_{m,t-1})(1 - e^{-1/\beta_d}) \qquad \text{if } E(W^{A_{mt}}) > W_{m,t-1} \qquad (17)$$

$$W_{mt} = E(W^{Amt}) + (W_{m,t-1} - E[W^{Amt}])e^{-1/\beta_u} \qquad \text{if } E(W^{A_{mt}}) < W_{m,t-1} \qquad (18)$$

$$X_{mt} = C_{mt} \qquad \text{if } A_{mt} > C_{mt} \qquad (21)$$

$$X_{mt} = A_{mt} - \left(E[W^{Amt}] - W_{m,t-1}\right)\left(1 - e^{-1/\beta_d}\right) \qquad \text{if } E(W^{A_{mt}}) > W_{m,t-1} \qquad (22)$$

$$X_{mt} = W_{m,t-1} + A_{mt} - E[W^{Amt}] + \left(E[W^{Amt}] - W_{m,t-1}\right)e^{-1/\beta_u} \qquad \text{if } E(W^{A_{mt}}) < W_{m,t-1} \qquad (23)$$

Metamodel of time-dependent WIP – variables

$$\phi_{m,t-1,t} = 1 \quad \text{if } E[W^{A_{mt}}] > W_{m,t-1}$$
$$= 0 \quad \text{if } E[W^{A_{mt}}] < W_{m,t-1}$$

$$\xi_{mt} = 1 \quad \text{if } A_{mt} < C_{mt}$$
$$= 0 \quad \text{if } A_{mt} > C_{mt}$$

Metamodel of time-dependent WIP – output constraints

$$X_{mt} \le C_{mt} \qquad (24)$$

$$X_{mt} \le A_{mt} - \left(E([W^{A_{mt}}] - W_{m,t-1} \right) \left(1 - e^{-1/\beta_d} \right) + N(2 - \phi_{m,t-1,t} - \xi_{mt}) \qquad (25)$$

$$X_{mt} \le W_{m,t-1} + A_{mt} - E[W^{A_{mt}}] + \left(E[W^{A_{mt}}] - W_{m,t-1} \right) e^{-1/\beta_u} + N\phi_{m,t-1,t} \qquad (26)$$

$$\begin{aligned} E[W^{A_{mt}}] &< W_{m,t-1} + N\phi_{m,t-1,t} \\ E[W^{A_{mt}}] &> W_{m,t-1} - N(1 - \phi_{m,t-1,t}) \end{aligned} \qquad (27)$$

$$\begin{aligned} A_{mt} &< C_{mt} + N(1 - \xi_{mt}) \\ A_{mt} &> C_{mt} - N\xi_{mt} \end{aligned} \qquad (28)$$

$$\begin{aligned} E[W^{A_{mt}}] &\le f(A_{mt}) + N(1 - \xi_{mt}) \\ E[W^{A_{mt}}] &\ge f(A_{mt}) - N(1 - \xi_{mt}) \end{aligned} \qquad (29)$$

Research issues

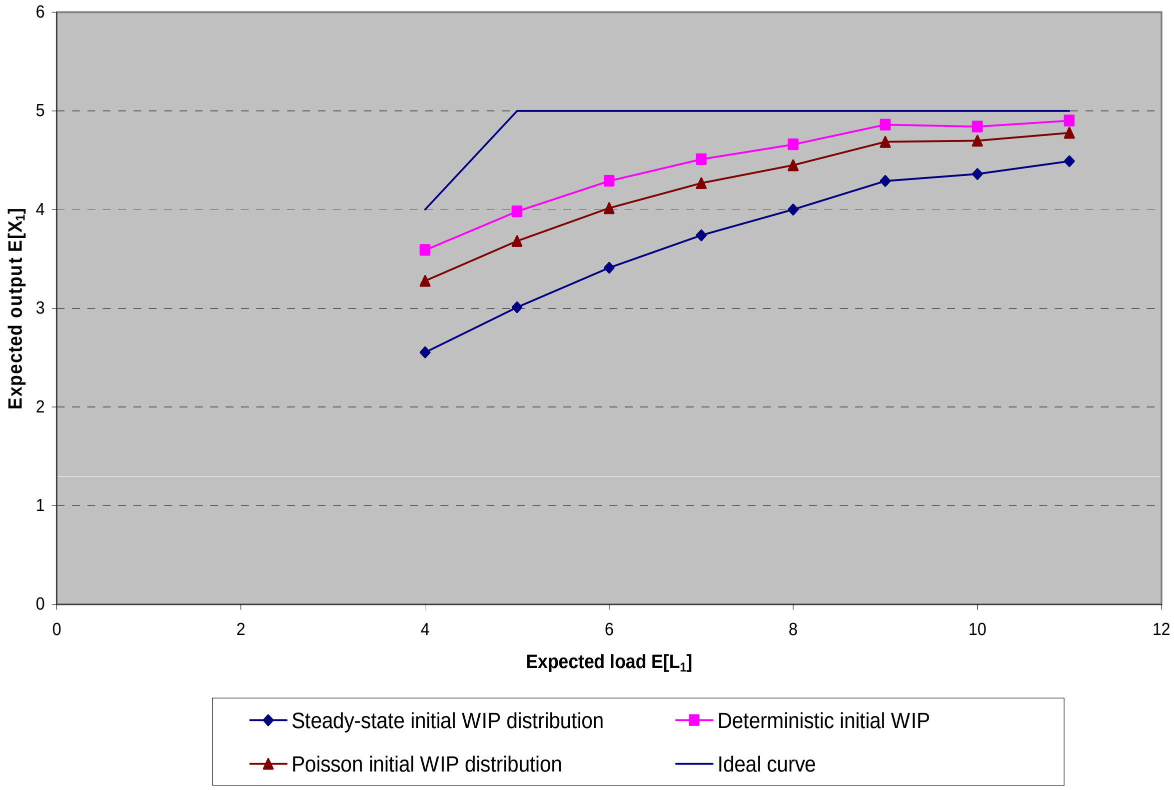
Technical issues

- Formulation of the models (Transient CF, WIP diffusion)
- Parameter setting of the models
- Optimization technique

Conceptual issues

- Unifying theory: Relationships between
  - Transient CF
  - WIP diffusion models
  - Flow time oriented models (e.g., Hung/Leachman)
- Rolling horizon planning – considering information gain over time

Clearing function – impact of the discrete distribution of the initial WIP



Clearing functions for period 1 for different distributions of the initial WIP W_{t-1}.
$E[W_{t-1}]=4$ orders, period length $t=5$, service rate $\mu=1$, input rate λ.

WIP over time with time-dependent 1-dimensional clearing function (optimization result)