Continuum models and their validation for semiconductor production

Dieter Armbruster
Department of Mathematics,
Arizona State University
& Department of Mechanical Engineering
Eindhoven University of Technology
E-mail: armbruster@asu.edu

June 20, 2006

Collaborators

Christian Ringhofer, Mathematics ASU
Karl Kempf, Decision Science Intel
Tae-Chang Jo, now Inha University South Korea
Dan Marthaler, now Northrop Grumman Aerospace Corporation
Erjen Lefeber, Emiel v. d. Rijt, Systems Engineering, Eindhoven Technical University
Partial support by NSF grants DMS 0204543 and from Intel Corporation is gratefully acknowledged

Overview:

1. Introduction
2. Adiabatic (quasi-static) models
   - validation against χ simulations
3. Advection diffusion equations
   - validation against factory data
   - validation against χ simulations
4. Policies, bottlenecks and degree of re-entrant behavior

1) Introduction

Example: Chip production in semiconductor manufacturing.
Factory investment several billions of $$.
Issues:
1.) Hardware: how many machines, topology of production flow
2.) Software: starts policies, dispatch policies, production mix
Idea: Generate a faithful representation of the factory and do simulation experiments using *Discrete Event Simulations*, e.g. χ (TU Eindhoven)
Problem: Simulation of production flows with stochastic demand and stochastic production processes requires Monte Carlo Simulations

It is not scalable.

2) Continuum Models of production flows

Fundamental Idea:
Model high volume, many stages, production via a continuum.
Basic variable:
product density (mass density) $\rho(\mathbf{x}, \mathbf{t})$.
x- is the position in the production process, $x \in [0, 1]$.- degree of completion- stage of production
Note: For a re-entrant process a machine corresponds to many positions x.

Mass conservation and state equations

Quasi-stationary model (adiabatic model): Mass conservation and state equation

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial F}{\partial x} &= 0 \\ F &= \rho v_{eq} \end{aligned}$$

Typical models for the equilibrium velocity v_{eq} (state equation) are

$$v_{LW}(\rho) = v_0(1 - \frac{\rho}{\rho_c}), \quad (1)$$

$$v_Q(\rho) = \frac{v_0}{1 + \frac{L(\rho)}{L_c}}, \quad (2)$$

$$v_{eq}(\rho) = \Phi(L), \quad (3)$$

with L the total load (Work in progress, WIP) given as

$$L(\rho) = \int_0^1 \rho(x,t)dx. \quad (4)$$

Φ maybe determined experimentally or theoretically.

Note:
The equilibrium velocity is closely related to the concept of a clearing function $\Omega(Wip)$ (Karmarkar, Uszoy).
Ω - Throughput as a function of Wip.
used locally to characterize a single machine or a group of machines
used globally to characterize a larger production unit via an equivalent queue.
Typical models:

$$\begin{aligned}
\Omega_c(Wip) &= min\{v_0 Wip, \mu_0\} \\
\Omega_Q(Wip) &= \frac{v_0 Wip}{1 + Wip} \\
\Omega_K(Wip) &= \alpha(1 - exp(-\beta Wip))
\end{aligned}$$

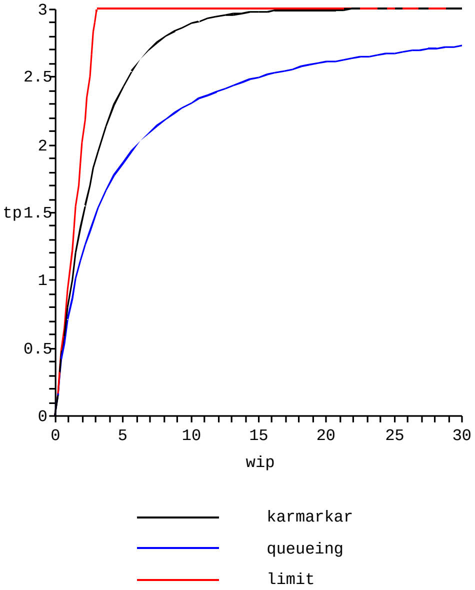


Figure: 1. Three different clearing functions

Validation through χ-simulations

Simulate a network in χ, consisting of 5 machines, re-entrant with 4 production loops.
Run experiment for different influx levels
Generate average cycle time, average WIP, and average throughput.
Figure 2 shows resulting state equation.

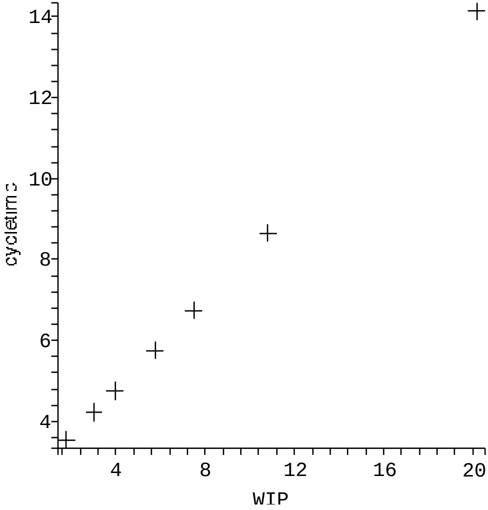


Figure: 2. Seven datapoints for a state equation describing the relationship between cycle time and WIP.

Usage:

1. Determine steady state parameters through inter/extrapolation
2. Use as state equation for transient behavior

Figures 3, 4 show two experiments

1. Transition from a steady state with 75% utilization to 85% utilization with pull policy
2. Transition from a steady state with 67% utilization to 85% utilization with push policy

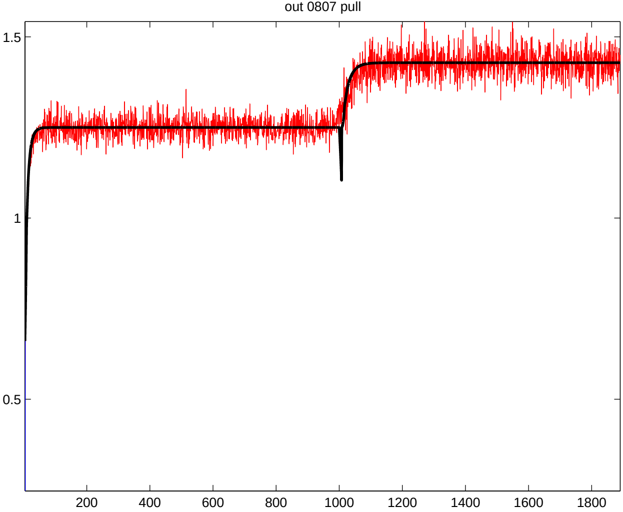


Figure: 3. Output for a pull policy

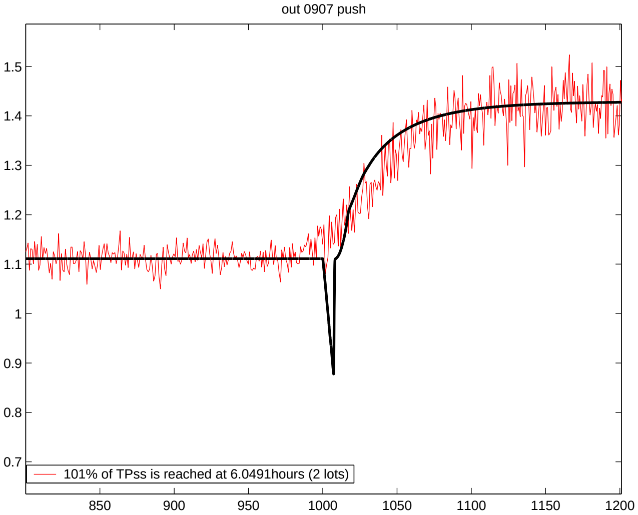


Figure: 4. Output for a push policy, zoomed

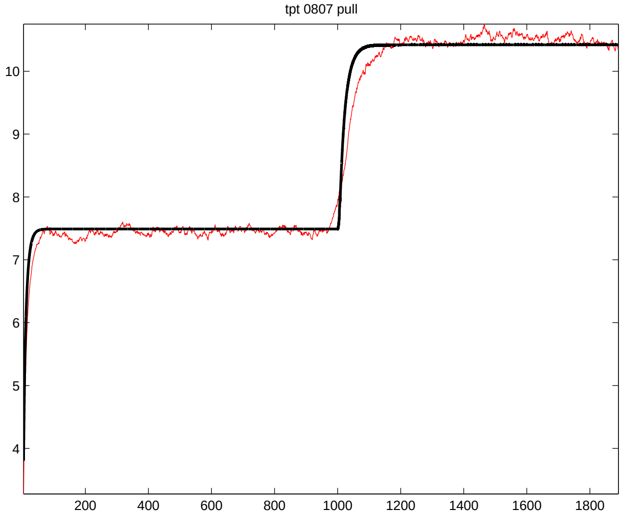


Figure: 5 Cycle time

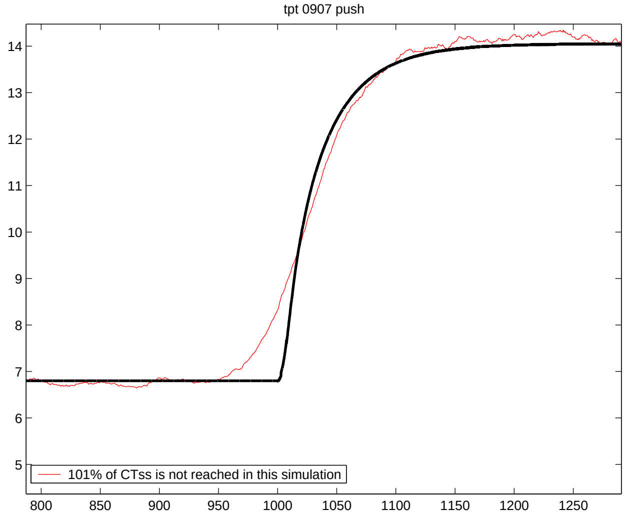


Figure: 6 Cycle time

Notice:
Inverse response in Figures 3 and 4
Due to global velocity **model**: Increase in influx leads to a change in total WIP
Result: infinite wave speed and decrease in velocity and hence decrease of output before the increase arrives at the end of the factory.
Reality: re-entrant production has this effect but much less pronounced.

3) Advection diffusion equations

Including variance in stochastic models typically introduces diffusion. We expect that the basic mass conservation model becomes

$$\begin{aligned} \frac{\partial \rho}{\partial t} + v_{eq}\frac{\partial \rho}{\partial x} &= D\frac{\partial^2 \rho}{\partial x^2} \quad (5) \\ v_{eq}(t) &= \Phi(W(\rho(x,t))) \quad (6) \\ W(t) &= \int_0^1 \rho(x,t)dx \quad (7) \end{aligned}$$

a) Real Factory Data

Data analysis of sanitized data of a real INTEL factory for about 3 months production.

Details:

- 920 lots
- time log in and out at all machines
- identify 8 approximately equally spaced machines
- Determine time-in at all 8 machines
- interpolate paths
- generate histograms at different times.
- Fit the histograms to the explicit solution of the advection-diffusion equation by a least square fit for the diffusion coefficient.

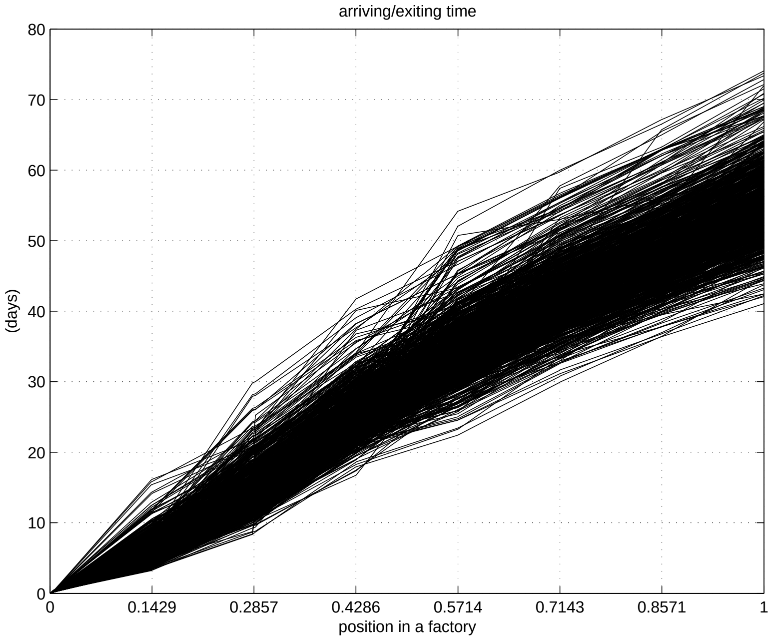
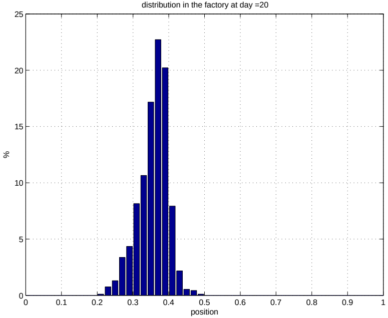
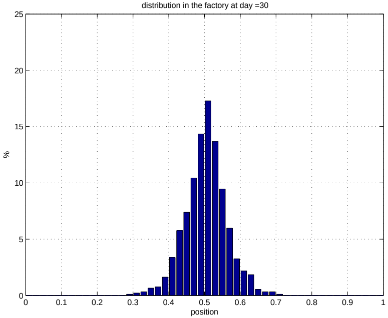


Figure: 7. Paths of all 920 lots


distribution in the factory at day =20
%
position
0
5
10
15
20
25
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


distribution in the factory at day =30
%
position
0
5
10
15
20
25
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

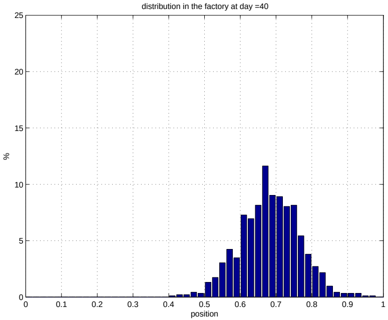


Figure: 8 Histograms at t=20, 30 and 40

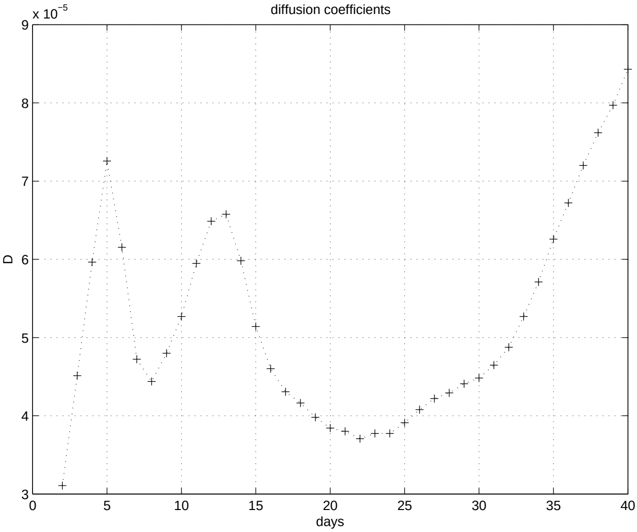


Figure: 9. Diffusion coefficient

Result:

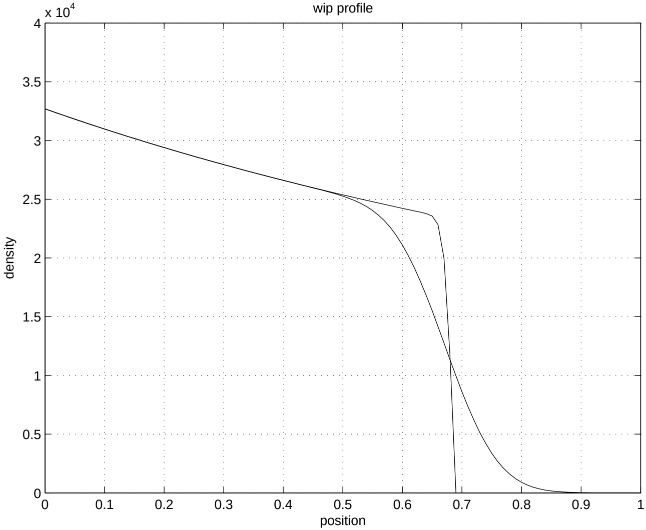


Figure: 10 WIP profile after a step up in influx, with and without diffusion

b) χ **Simulations** - Emiel v.d. Rijt

100 identical machines
characterized through mean process time t_e and squared coefficient of variation c_e^2.
Arrival process: average interarrival time t_a and its squared coefficient of variation c_a^2
Goal: Determine the dependence of the WIP profile on the ratios

$$\frac{c_a^2}{c_e^2} \quad \text{and} \quad u = \frac{t_e}{t_a}$$

Simulations - Wip profiles:

utilization $u = 0.75$,

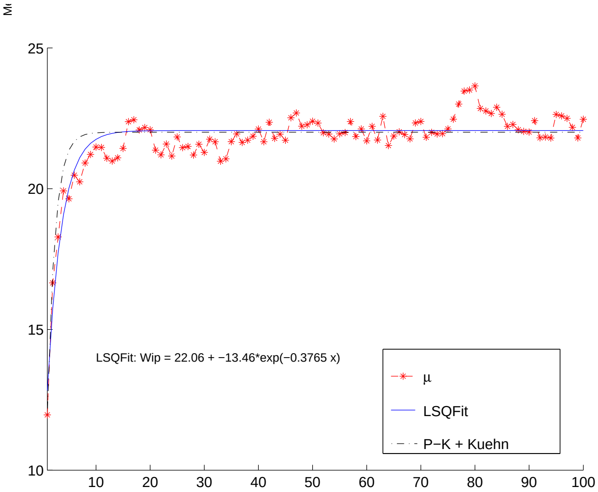


Figure: 11 WIP profile for $\frac{c_a^2}{c_e^2} = \frac{1}{9}$

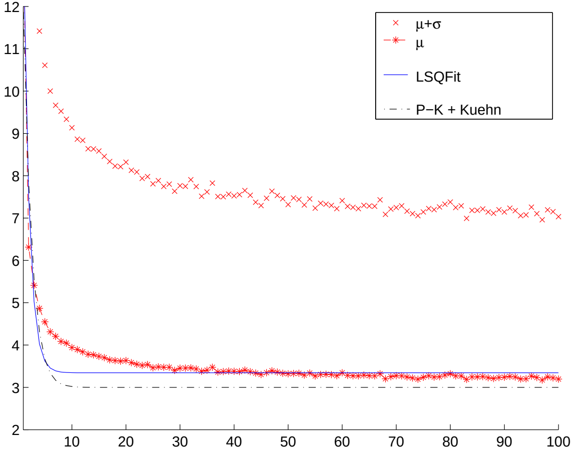


Figure: 12 WIP profile for $\frac{c_a^2}{c_e^2} = \frac{9}{1}$

Boundary layer is utilization dependent: $u = 0.99$

The BL has reduced to one machine.

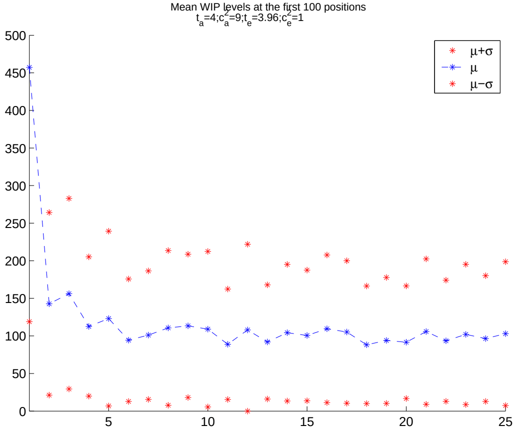


Figure: 13 WIP profile for $\frac{c_a^2}{c_e^2} = \frac{9}{1}$

PDE model

Consider the steady state of the advection diffusion equation:

$$\rho v_{eq} + D\frac{d\rho}{dx} = \lambda \tag{8}$$

with boundary condition $\rho(0)$ e.g. given by

$$\rho(0) = \frac{1}{2}(c_a^2 + c_e^2)\frac{u^2}{1-u} + u$$

The solution to Eq.(8) is

$$\rho(x) = \frac{\lambda}{v_{eq}} + (\rho_0 - \frac{\lambda}{v_{eq}})e^{-\frac{v_{eq}}{D}x} \tag{9}$$

Notice:
with $\rho_{ss} = \frac{\lambda}{v_{eq}}$, e.g.

$$\rho_{ss} = c_e^2 \frac{u^2}{1-u} + u$$

we get that

- if $c_a^2 > c_e^2$ then $\rho''(x) > 0$,
- if $c_a^2 < c_e^2$ then $\rho''(x) < 0$,
- utilization dependence:
  Scaling argument: $D \approx \sigma^2, \ \sigma^2 = c v_{eq}^2$
  $\lim_{u \to 1} v_{eq} \to 0$
  Hence from $e^{-\frac{v_{eq}}{D} x}$ the boundary layer ξ becomes $\xi \approx (1-u)$.

4 Policies and bottlenecks

Assume: Every machine has a fixed deterministic maximal capacity, μ

Mass conservation then leads to

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial F}{\partial x} &= 0 \\ F &= \min\{\rho v_{eq}, \mu\} \\ v_{eq} &= \Phi(L) \end{aligned}$$

Hybrid model: v_{eq} addresses slow down due to ubiquitous stochasticity
Capacity limit: Addresses hard upper bound at bottleneck machines.

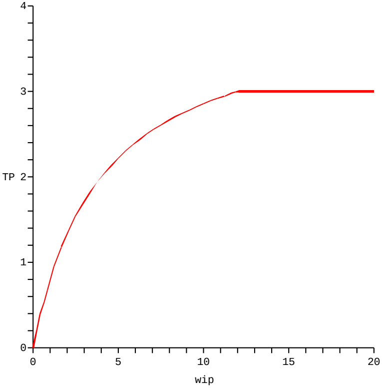


Figure: 14. Nonsmooth clearing function

This allows to address **dispatch policies:**
Consider N -stages x_i that are serviced at the same machine.
Assume the total desired flux $\Sigma_i^N \rho_i v_i$ at a machine is greater than the total capacity of that machine.
Then we have three options

- **Push** policy: Allocate capacity from front
- **Pull** policy: Allocate capacity from back
- **FIFO** policy: on average allocate μ/N to each stage

movies

Factory Master Equation for Linear Factory from First Principles

Goal:
Derive Factory Master Equation from 'First Principles' with methods from gas-dynamics.
General idea:
Boltzmann equation for the density $f(x, y, t)$ of a particle at position x with attribute y at time t:

$$\partial_t f + \partial_x[u(x, y, t)f] + \partial_y(E(x, y, t)f) = Q[f] \ , \qquad (10)$$

$\partial_y(Ef)$ models a continuous change in attribute,
$Q[f]$ models a random and discontinuous change in the attribute
Note that

$$\int Q[f](x, y, t) \ dy = 0, \ \forall f \ . \qquad (11)$$

Define part density $\rho(x,t)$ and flux density $F(x,t)$ by

$$\begin{aligned} \rho(x,t) &= \int f(x,y,t)\; dy, \\ F(x,t) &= \int u(x,y,t) f(x,y,t)\; dy \end{aligned}$$

integrating we get the conservation law

$$\partial_t \rho + \partial_x F = 0\ .$$

as the zero order moment equation.
Goal: Determine E and Q from detailed underlying kinetic behavior and extend to higher moments.

Conclusions

PDE models of production flows are highly effective simulation tools. They can be adjusted to the desired level of accuracy and modeling sophistication. In addition

- they have execution times in seconds
- they can be adjusted to include policies
- they can be adjusted to cover inhomogeneous production lines.
- they allow to simulate transient situations
- they can be linked to generate simulations for the whole supply chain
- they can be justified (in parts) from first principles

References:

- Dieter Armbruster, Daniel Marthaler, Christian Ringhofer, Karl Kempf, Tae- Chang Jo: A continuum model for a re-entrant factory, in print, Operations research 38 pages (2006)
- D. Armbruster, C. Ringhofer, Thermalized kinetic and fluid models for reentrant supply chains, SIAM J. on Multiscale modeling and Simulation, **3**(4), pp 782 - 800, (2005)
- D. Armbruster, P. Degond, C. Ringhofer: A Model for the Dynamics of large Queuing Networks and Supply Chains, SIAM J. Applied Mathematics **66**(3) pp. 896-920. (2006)