Aalto University
School of Science

# Delays

# Delays

- Nothing happens instantly, everything takes time
  - Deciding what to do
  - Implementing the actions
  - Actions taking effect
  - Observing the new system state
- Modeller must understand
  - How delays behave
  - How delays are described
  - How to choose between different delay models
  - How to estimate delay duration

Aalto University

# Delays

- Delay is a process where output lags behind the input


- Delays are critical source of dynamics in all systems
  - Some delays create instability and oscillation
  - Other delays may filter out unwanted noise and help seeing the actual trend
- Delays include at least one storage

# Delays – examples

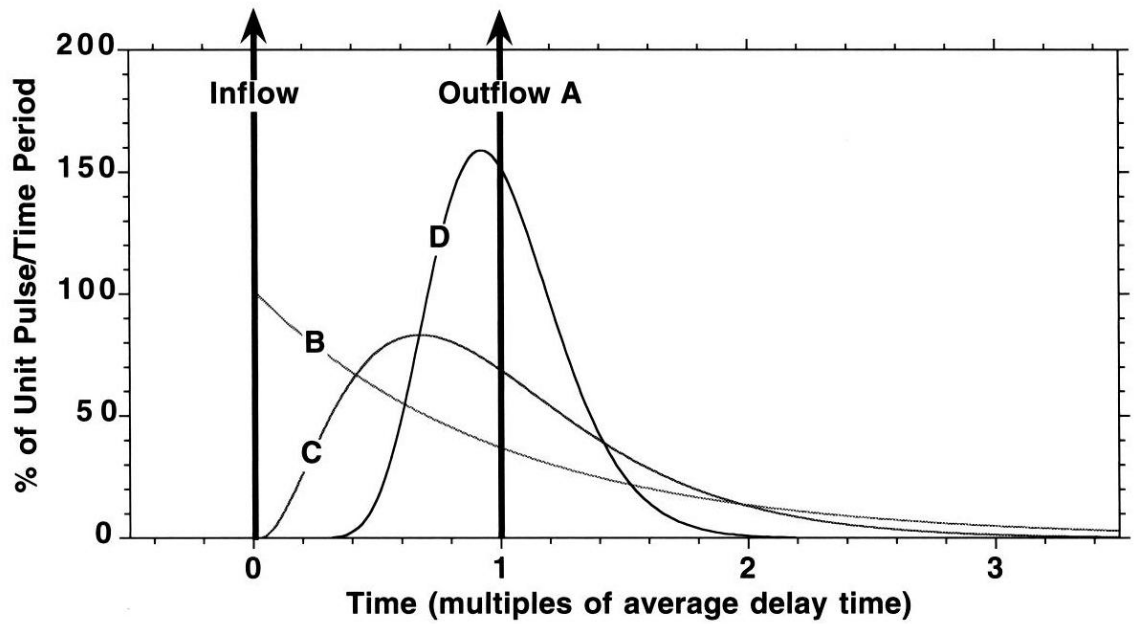
- How long does it take for mailed letters to arrive?
- How long do patients stay in hospital?
- How fast can a factory increase capacity?
- How fast can agricultural production be doubled?
- How long does it take for economic forecasters to revise their estimates of inflation?

# Delay types

- Material delays
  - Material balance is maintained: material is preserved, does not disappear or appear from nowhere
  - Delay must contain at least one storage because output is different from input (delayed)
  - Example: sending letters
  - Also energy and money is preserved
- Information delays
  - Information can be increased or disappear
  - Updating beliefs and mental models can be slow
  - Example: expected demand of product based on orders

# Material delay

- Constant delay
  - FIFO (First In, First Out), e.g. conveyor belt
- Variable delay
  - LIFO (Last In, First Out)
  - Delay distribution
  - Order of items may change
  - All distributions have same average delay (1)



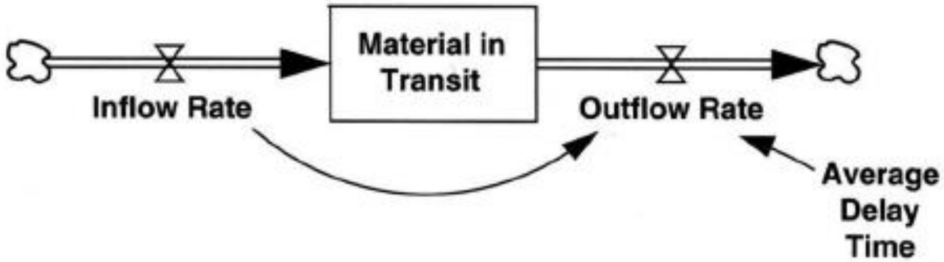
# Material delay

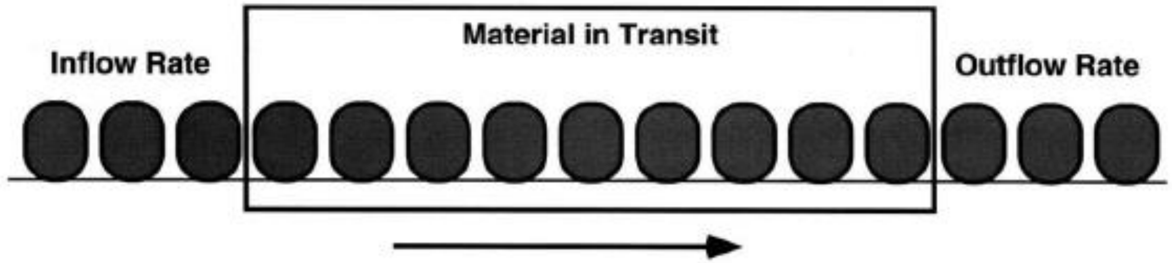
- Example: Cashier queue in market
  - Input: number of customers joining queue
  - Output: number of customers that exit market
  - Delay depends on
    - Number of bought items
    - Speed of cashiers
  - In case of multiple queues, order of customers may change because one queue may be faster than another

# Pipeline delay

- Pipeline preserves order of items in queue (FIFO)

$$Outflow(t) = Inflow(t - D)$$

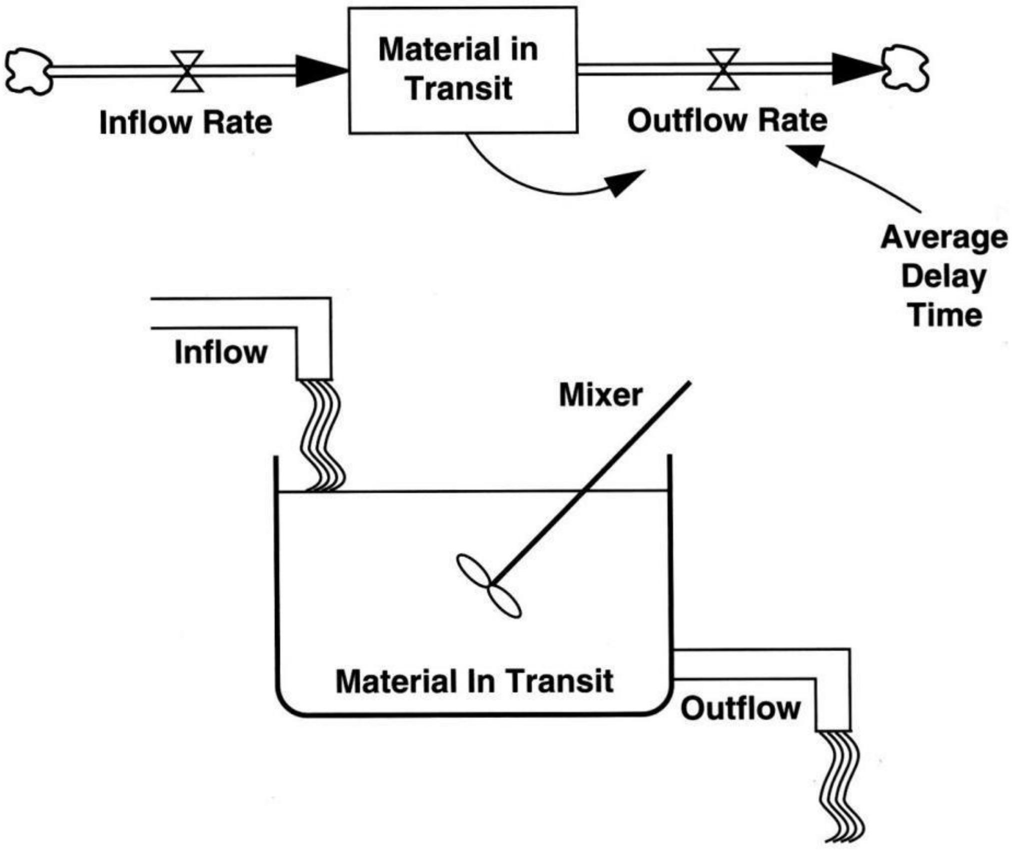



Material in Transit(t) = INTEGRAL(Inflow(t) – Outflow(t), Material in Transit(0))
Outflow(t) = Inflow(t – Average Delay Time)

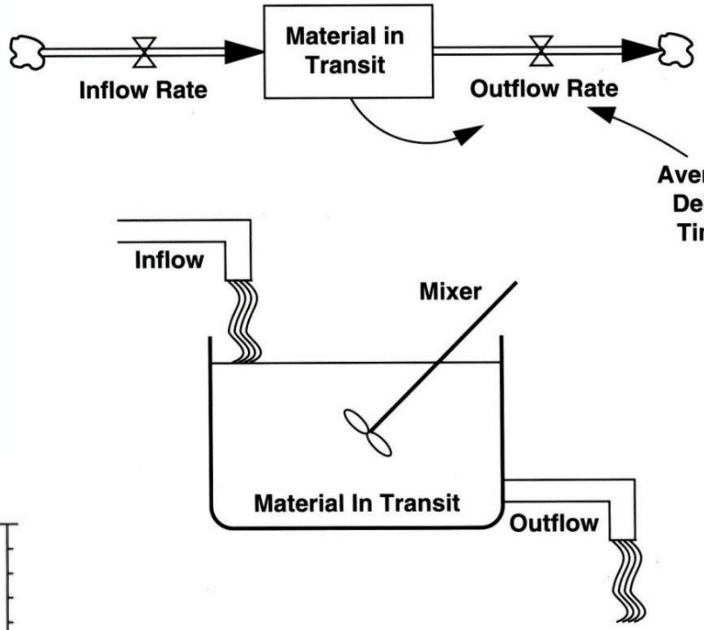
# First order material delay

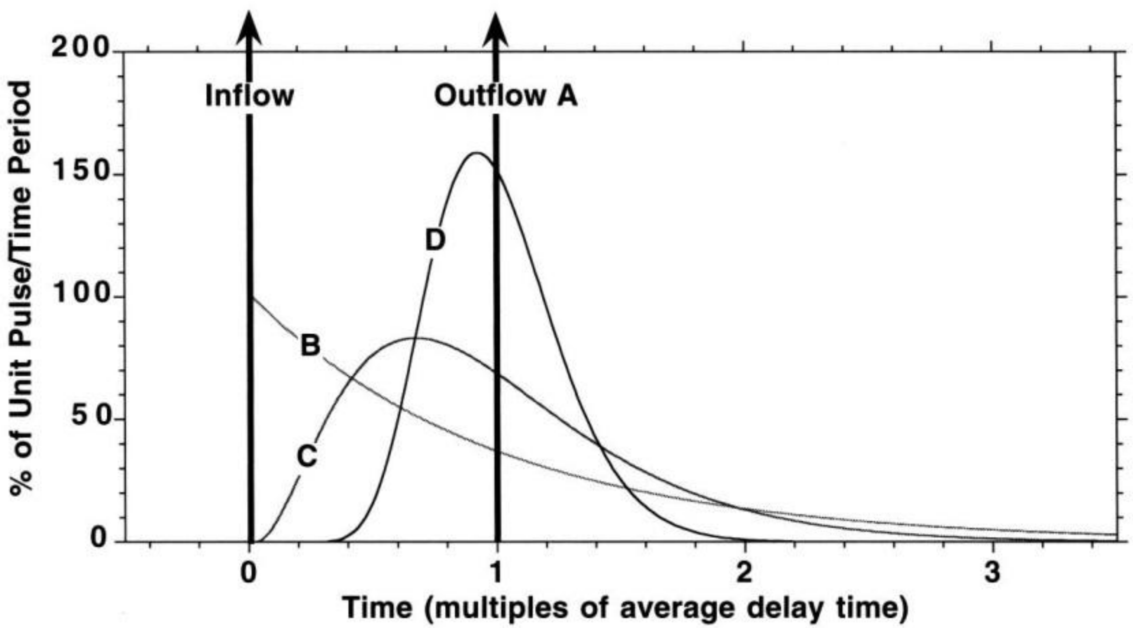
- Outflow is proportional to stock of material in transit
  - Outflow(t) = Stock(t)/D
  - D = average delay time
- Stock is perfectly mixed
  - all items have same probability of exit, independent on their arrival time
  - Memory-less process



# First order material delay

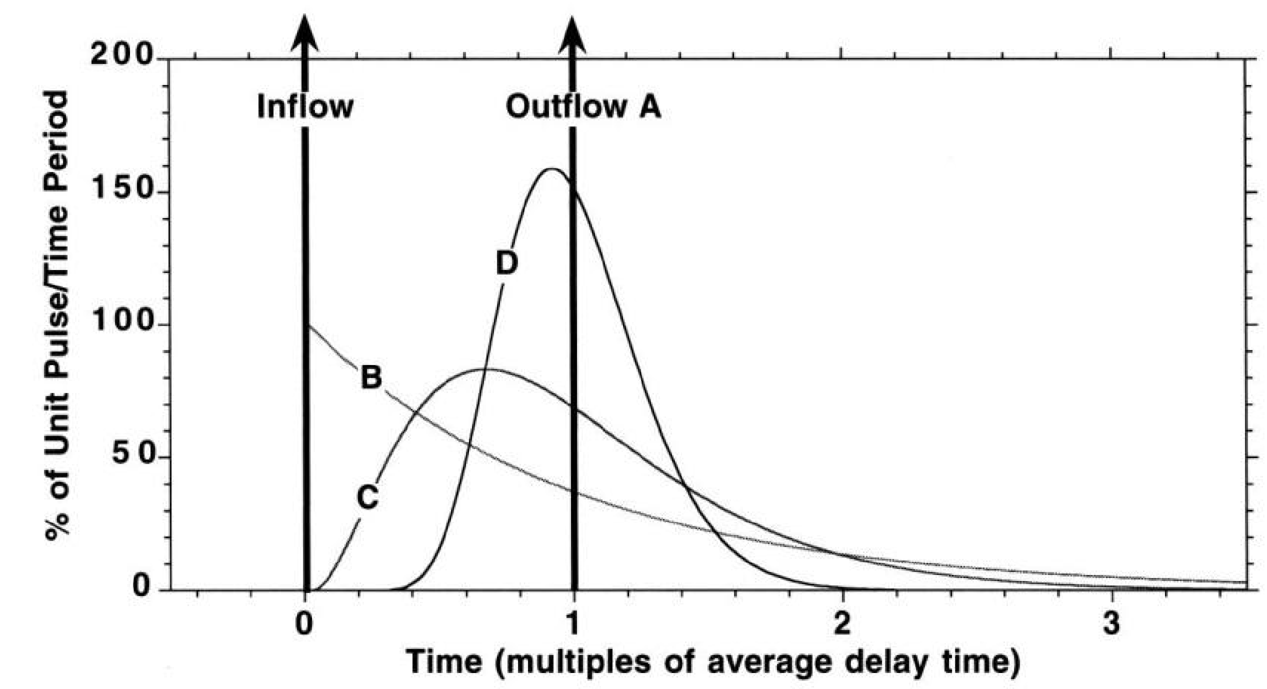
- The model is linear first-order negative feedback system
  - Exponential distribution (B) is the response to unit impulse function (Dirac's delta δ)



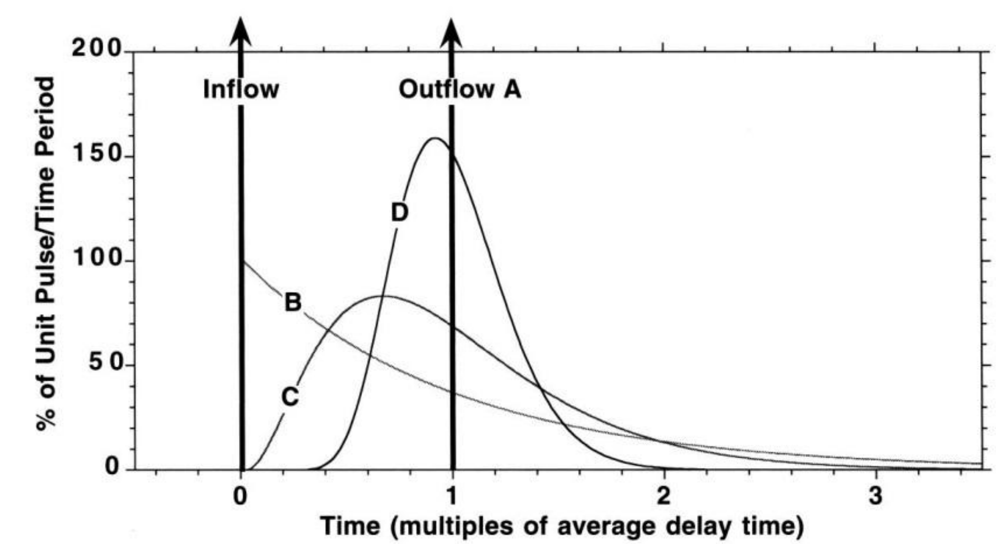
# Higher order material delays

- Pipeline delay (A) with FIFO discipline is good model e.g. for assembly lines
- First order delay process (B) with perfect mixing is suitable model for chemical and heat diffusion in biological systems
- Between these extremes are 2., 3., … order systems (C,D)

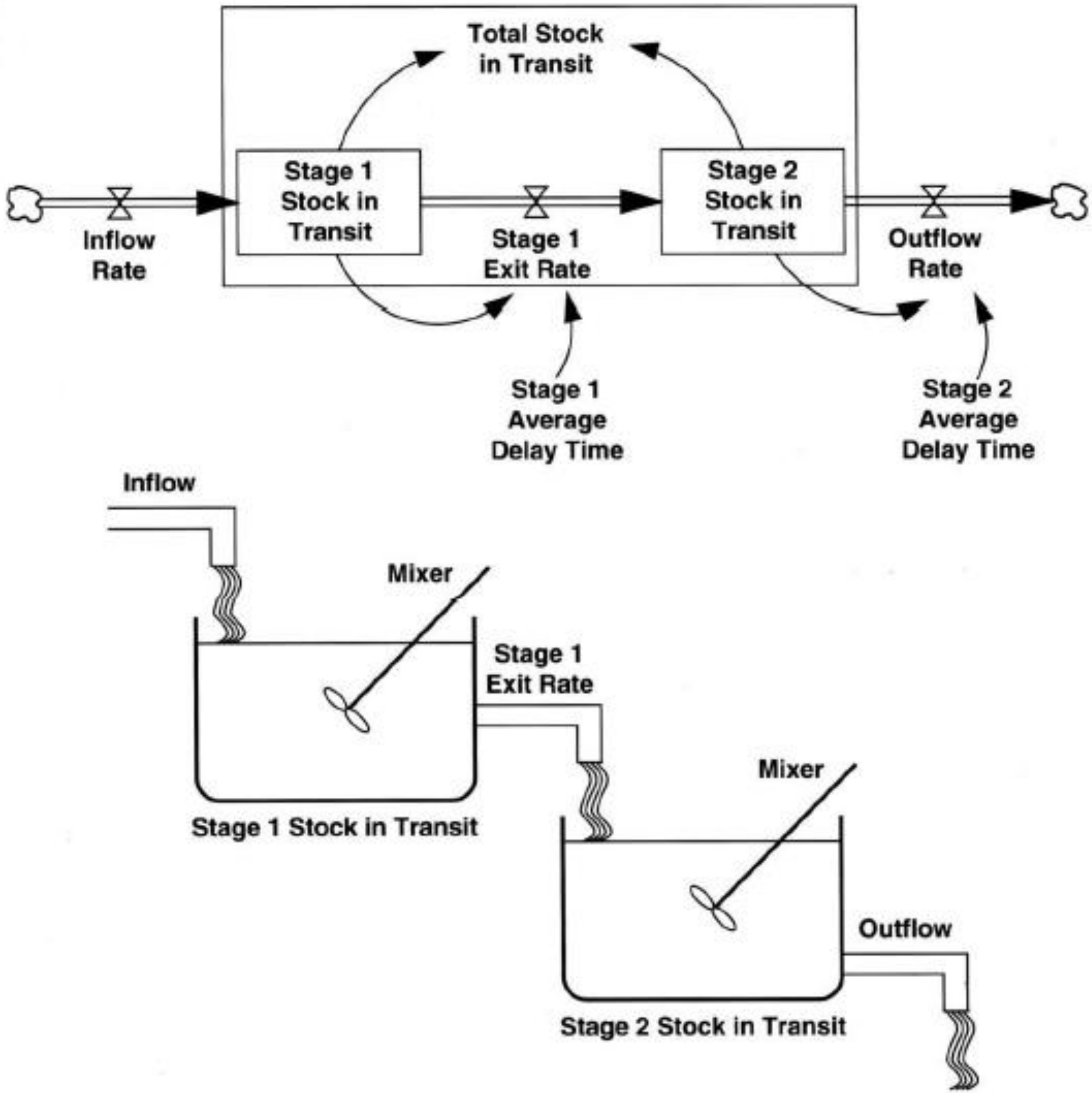


# Higher order material delays

- Example: mailing 1000 letters
  - Letters do not arrive at same time, so pipeline delay is wrong model
  - Letters are not delivered immediately, so 1. order delay is wrong
  - There is partial mixing, but not complete
    - This happens when delay is caused by multiple processing phases with some mixing



# Second order material delay

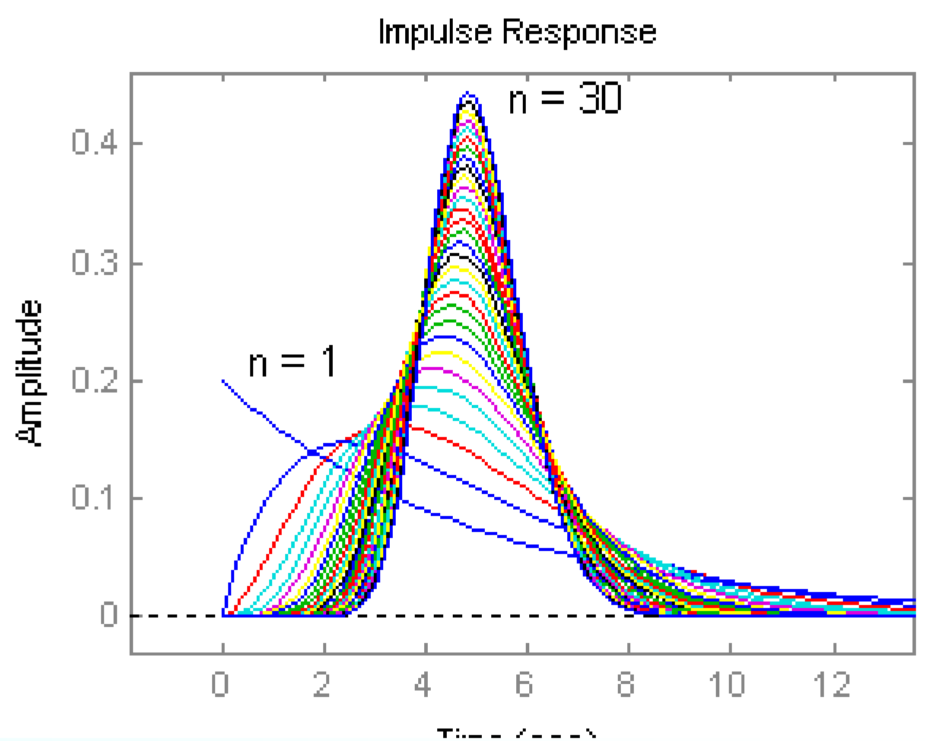



Stage 1 Exit Rate = Stage 1 Stock in Transit/Stage 1 Average Delay Time

Outflow Rate = Stage 2 Stock in Transit/Stage 2 Average Delay Time

# Higher order material delays

- Order n delay is formed by combining n first order delays with delay D/n in sequence
  - The higher the order, the less mixing will occur and the less variance in output
  - Infinite order gives pipeline delay



# Higher order material delays – Erlang distribution

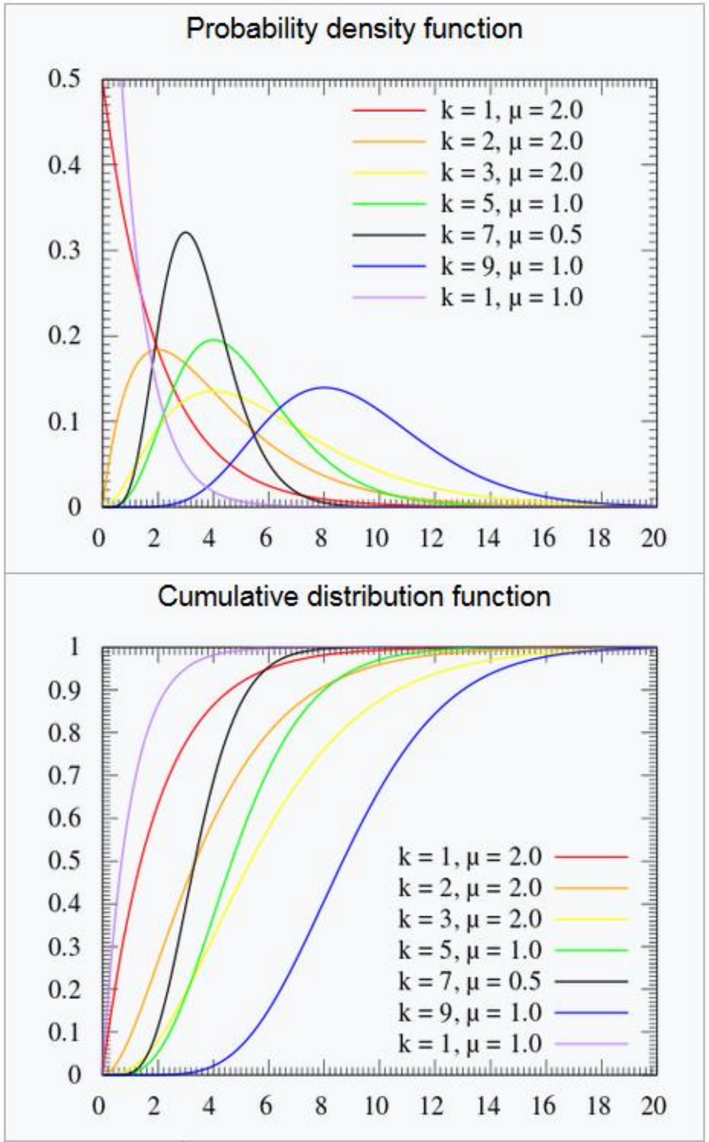
- The delay of order k follows the *Erlang distribution*

$$f(x;k,\lambda)=\frac{\lambda^k x^{k-1}e^{-\lambda x}}{(k-1)!}\quad \text{for } x,\lambda\geq 0$$

$$F(x;k,\lambda)=1-\sum_{n=0}^{k-1}\frac{1}{n!}e^{-\lambda x}(\lambda x)^n$$

  - Mean = k/λ
  - Variance = k/λ²

- Extended for non-integer k by the *gamma distribution*



# How much is in the delay? Little's Law (John Little, MIT)

- Assume Pipeline delay system in equilibrium
  - Input I, output O and delay time D: O(t) = I(t-D)
  - Initial stock = 0
  - First D time units input is I and output is 0
  - After D time units system reaches equilibrium with I=O and stock in transit **S=DI**
- Assume first order delay
  - Outflow O = S/D
  - Due to equilibrium I=O ⇒ **S=DI**
- Same is true for any delay process in equilibrium

# Little’s Law - Example

- Construction delays in electric utility industry, 1970’s
  - Lead time for constructing new plants was 5 years
  - Average service life was 20 years
    - 1/20 of capacity expires yearly
  - Therefore, company with 10 GW capacity needs to build 0.5 GW new capacity yearly to replace old plants
  - With 5 year construction delay, company needs to be constructing 5*0.5=2.5 GW to replace expiring plants
  - But, construction times doubled in the 1970’s $\Rightarrow$ work in progress (WIP) grew to 5 GW (50% of capacity)
  - Furthermore, energy consumption was increasing 7% yearly
  - $\Rightarrow$ WIP should be 17.4 GW (174% of capacity)!

# Little’s Law - Example

- Longer construction times caused orders for power plants to surge in mid 1970
- Huge debts were taken to finance ever-greater stock of WIP
- Electric power rates were raised
- This caused power demand to fall
- Utilities were carrying debts for power plants they didn’t need
- A number of major utilities went bankrupt
- **Lesson:** Power plants with short planning, permitting, and construction times were a better investment even though their power production cost was higher

# Information delays

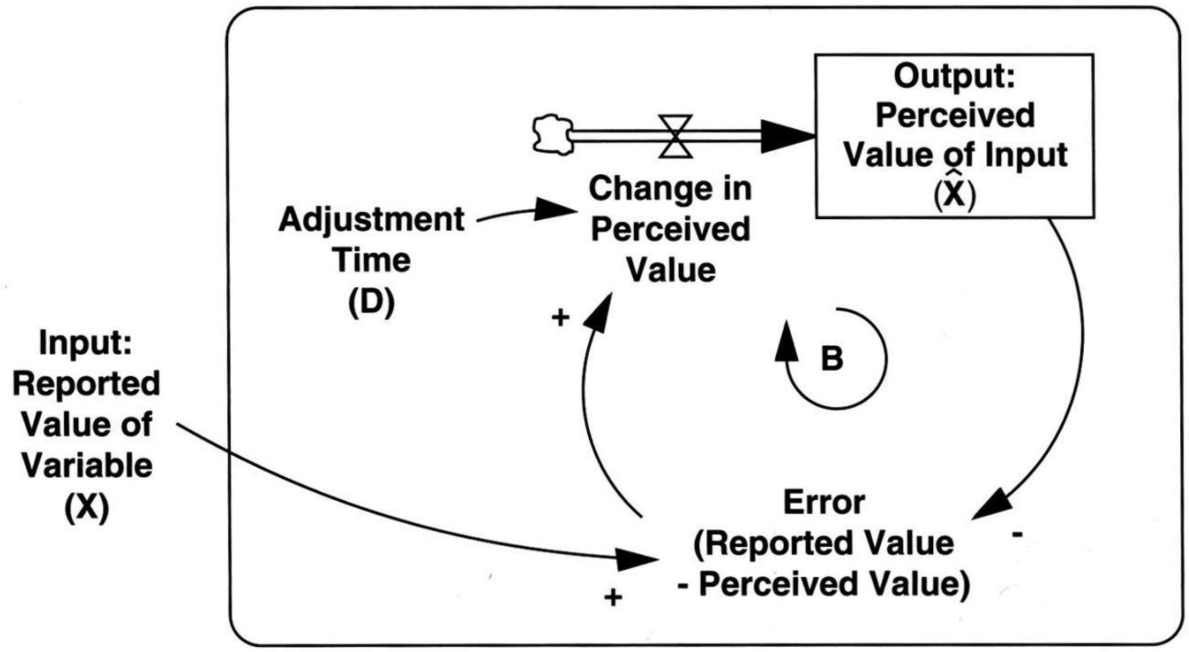
- Delays occur in channels of information feedback, measurement or perception of a variable, or in updating beliefs and forecasts
- Because information behaves differently from material (no law of material conservation), different model is needed

# Information delays – modeling perceptions

- Simplest and widely used model of belief adjustment is *exponential smoothing* or *adaptive expectations*.

$$\hat{X}(t) = \hat{X}(0) + \int_0^t C_v(\tau)\,d\tau$$

$$C_v(t) = \frac{X(t) - \hat{X}(t)}{D}$$

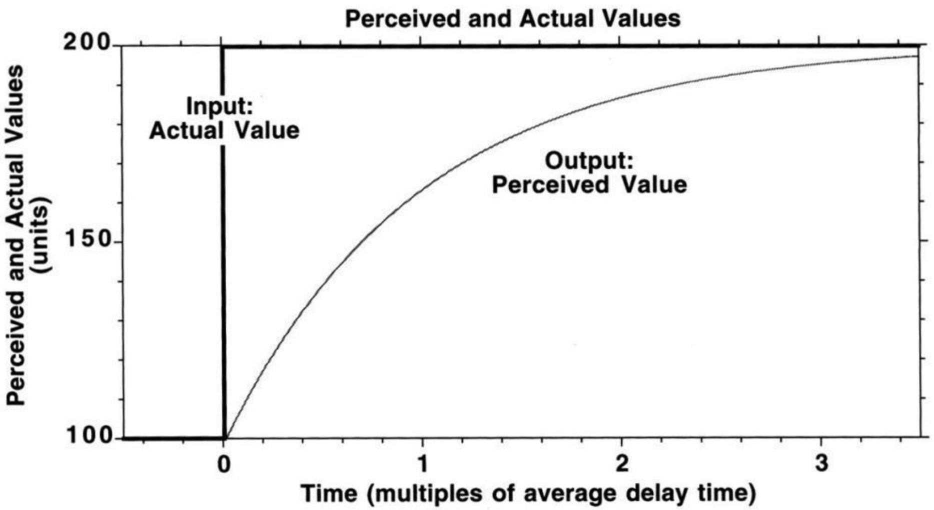


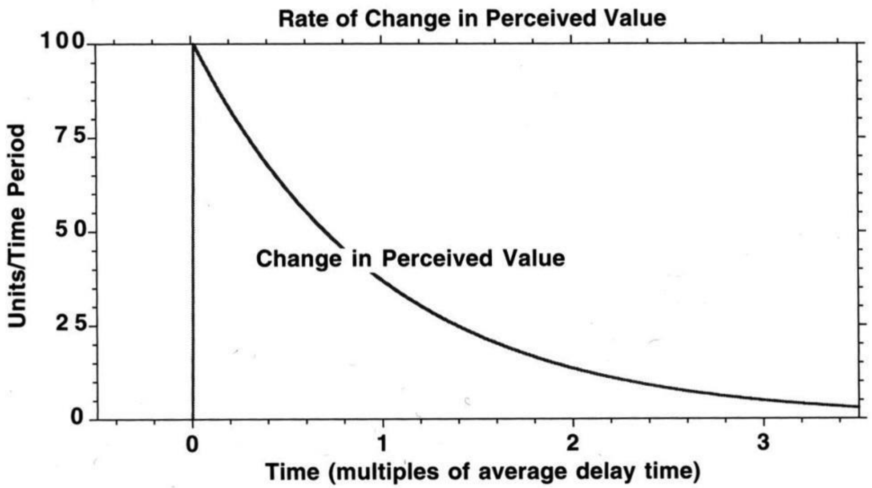

$\hat{X}$ = INTEGRAL(Change in Perceived Value, $\hat{X}(0)$)

Change in Perceived Value = Error/D = $(X - \hat{X})/D$

# Information delays – modeling perceptions

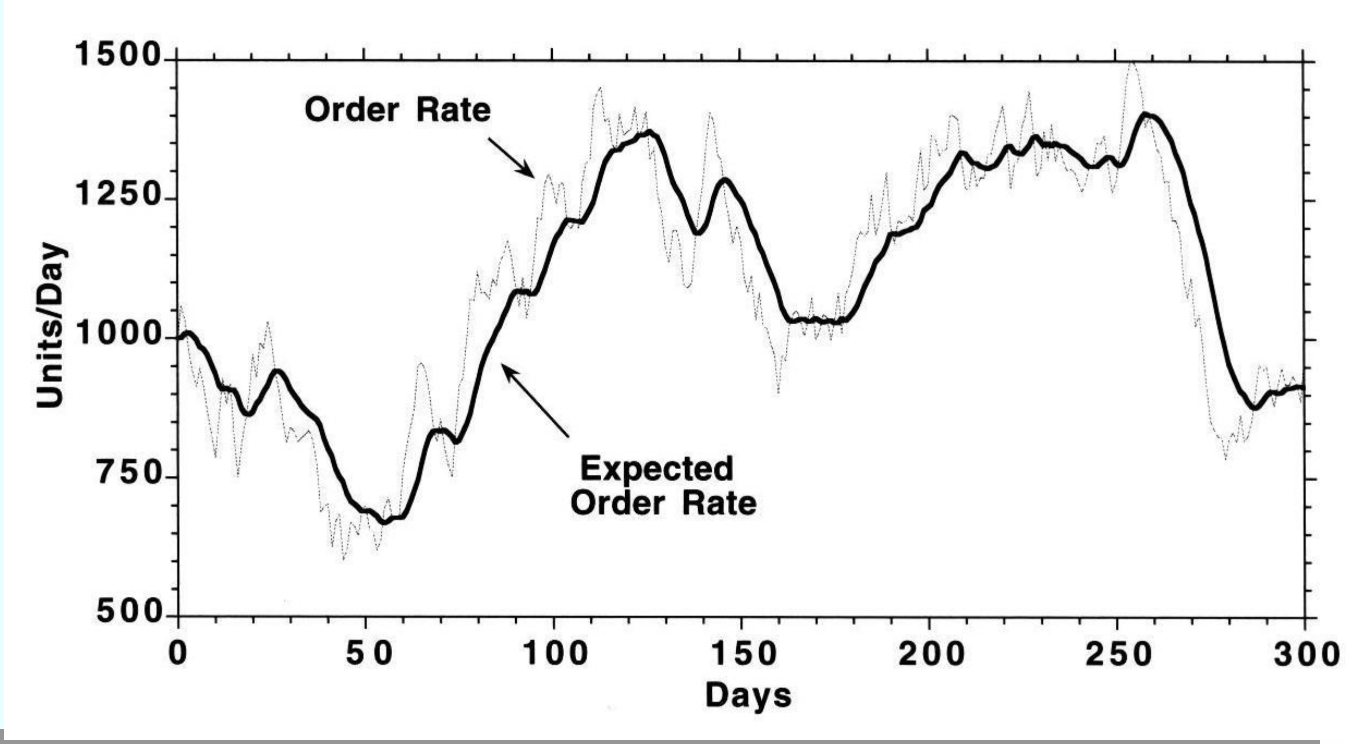
- The belief $\hat{X}$ is defined as a storage variable
  - The larger the difference between new information and current belief, the larger change in the state
  - Similar to first order negative feedback



# Information delays – modeling perceptions

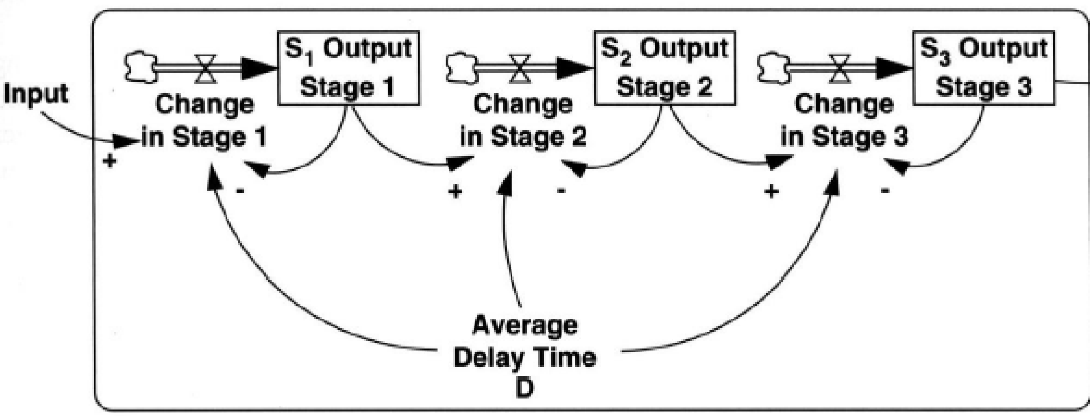
- Adaptive expectations smooth out short term noise
- Example: exponential smoothing of order rate with 7 day adjustment time



# Higher order information delays

- First order information delay causes immediate reaction to change in input
  - Often the input change must be permanent before beliefs change
  - This can be modelled by higher order information delays where reaction starts after some time
- Extreme case: pipeline delay
  - $\textit{Reported Value}(t) = \textit{Actual Value}(t - D)$

# Higher order information delays




Output = SMOOTH3(Input, D)

Output = $S_3$

$S_3$ = INTEGRAL(Change in Stage 3, $S_3(0)$)

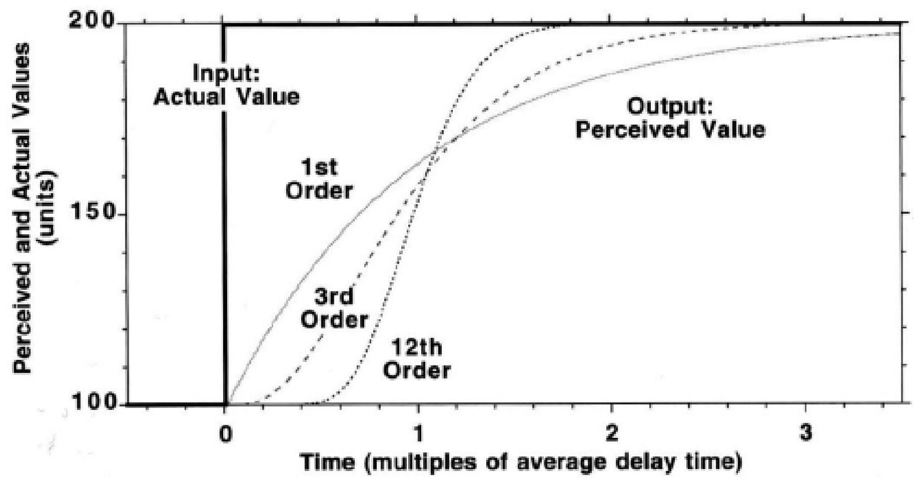
Change in Stage 3 = $(S_2 - S_3)/(D/3)$

$S_2$ = INTEGRAL(Change in Stage 2, $S_2(0)$)

Change in Stage 2 = $(S_1 - S_2)/(D/3)$

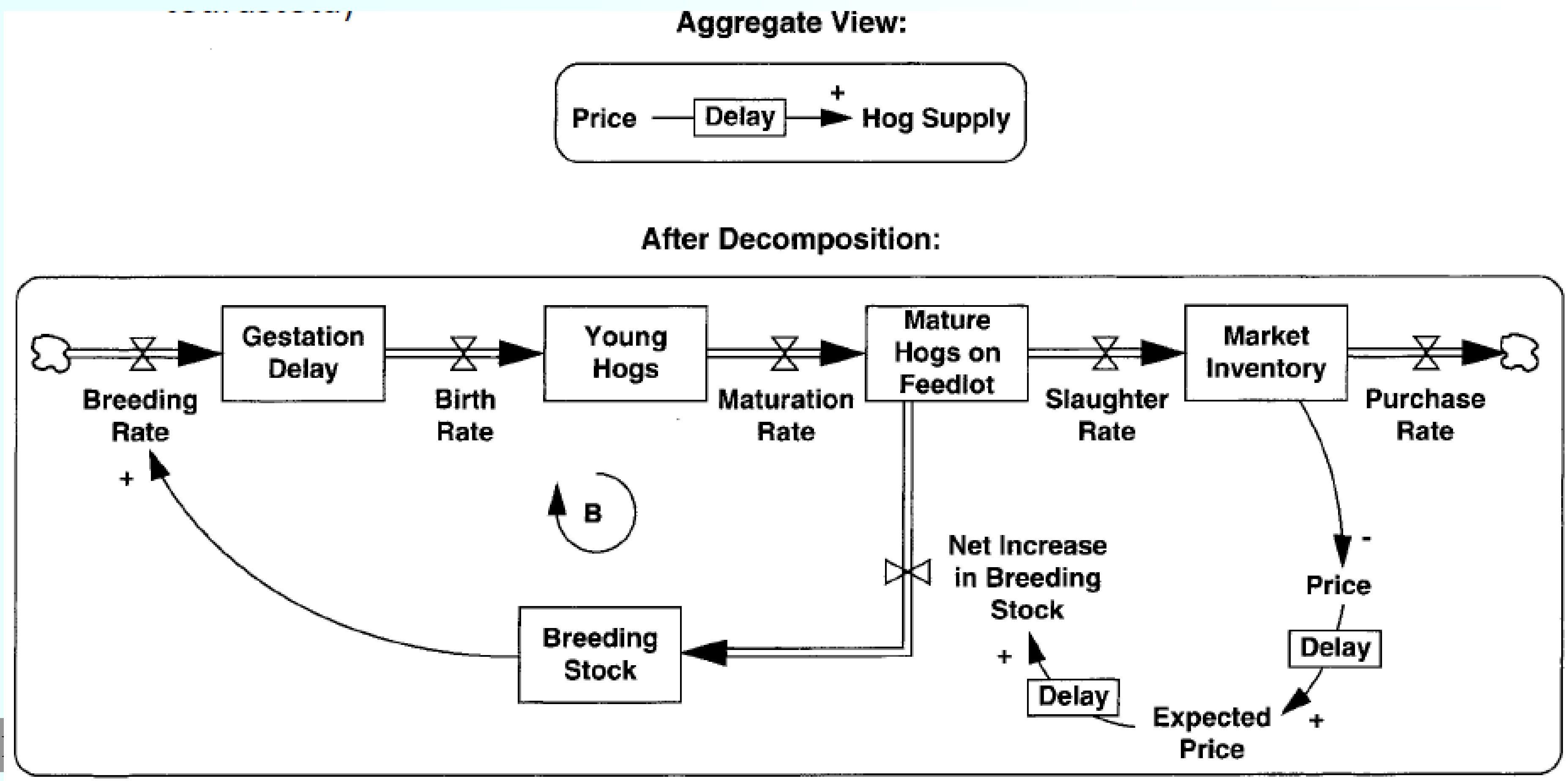
$S_1$ = INTEGRAL(Change in Stage 1, $S_1(0)$)

Change in Stage 1 = $(\text{Input} - S_1)/(D/3)$



# Third order information delay in producing pork

- Pork meat production has a 22 month average delay
  - System tends to oscillate with 4 year period



# Higher order information delays

- Often very high order delays are not needed
- Statistical methods can be used to judge order of delay
- Using pipe delay may slow down large simulation systems or cause inaccuracies
  - Unless delay is a multiple of time step
- Discrete event system simulation (DES) can handle pipe delays more efficiently