Analysis of Performance Approximations for Queueing Networks with Non-Homogeneous Arrival Processes

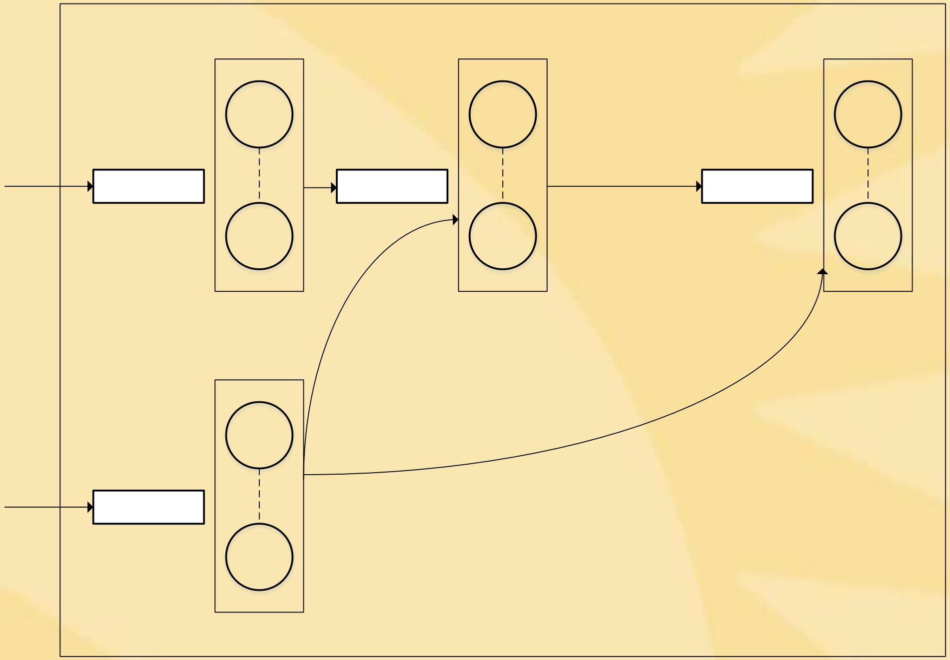
Ronald G. Askin Faculty of Industrial Engineering
Girish Jampani Hanumantha (Ph.D. Student)
School of Computing, Informatics, and Decision Systems Engineering
Arizona State University
Tempe, AZ 85721 USA
ron.askin@asu.edu
gjampani@asu.edu

Dynamic Queueing Network Analysis
Outline

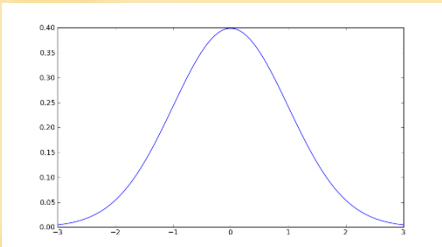
- Introduction
- Problem Definition
- Impact of Non-stationarity
- Literature Review
- Research Methodology
- Results
- Conclusions and Future Research

Introduction

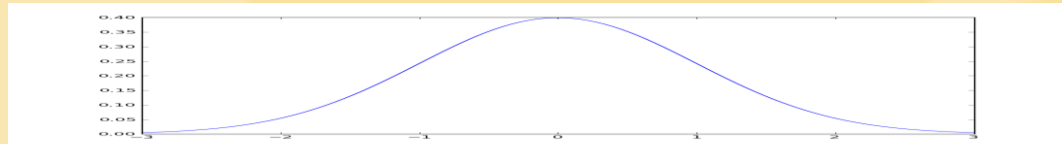
- The manufacturing facility as a dynamic queueing network



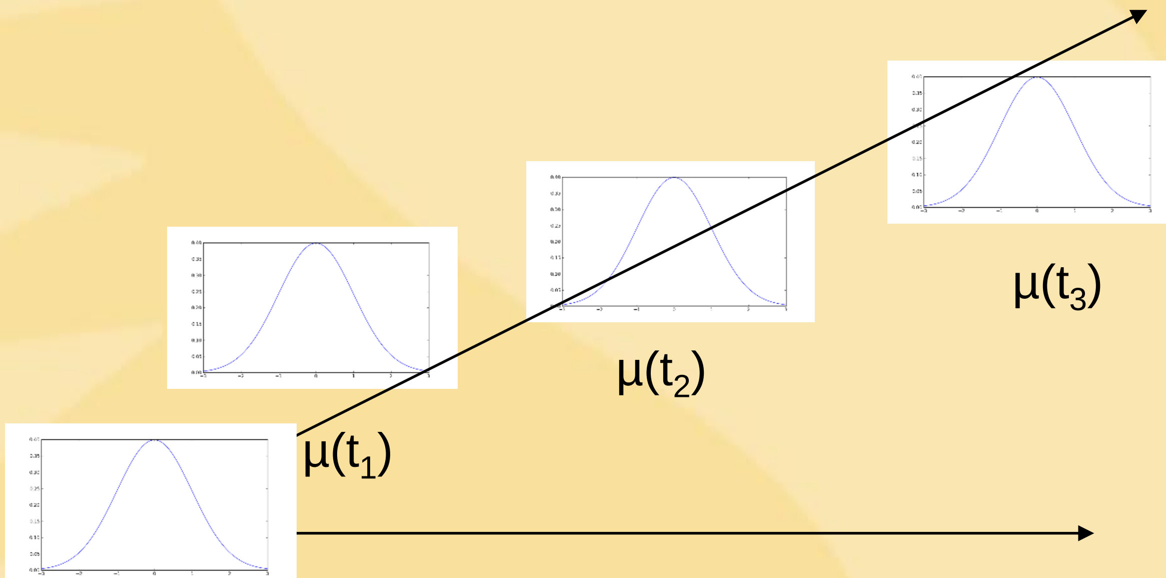
Randomness, Certainty, Variability



μ
Random



$f(\mu)$
Uncertain (Bayesian)

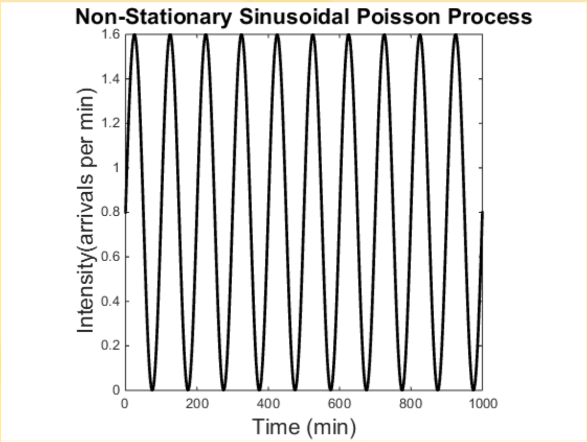


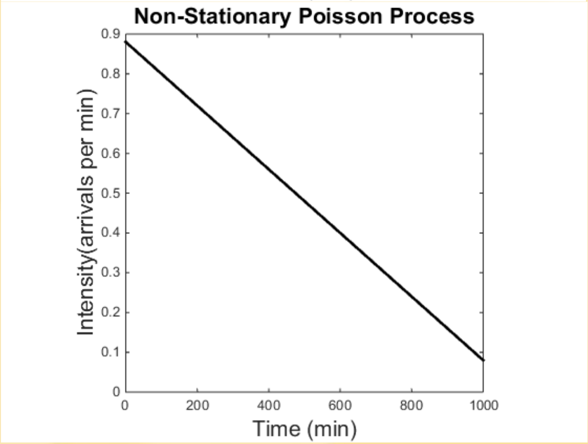
Variable
(Dynamic)

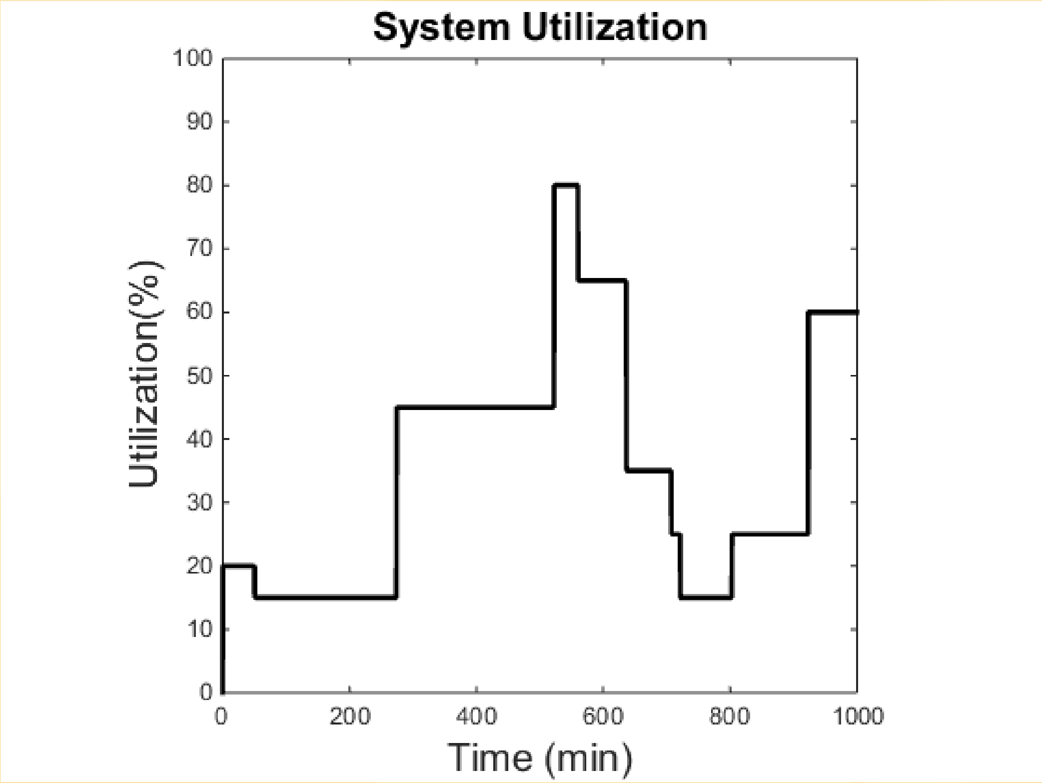
$\mu(t_0)$

Introduction

- Often systems may not adhere to product form model assumptions







Introduction

- Rough cut planning and scenario analysis under dynamic conditions call for efficient algorithms.

- Research question : How do we leverage available efficient algorithms for steady-state analysis to develop reliable approximations for first (and maybe higher) order estimates of performance measures under dynamic conditions?

Problem Definition

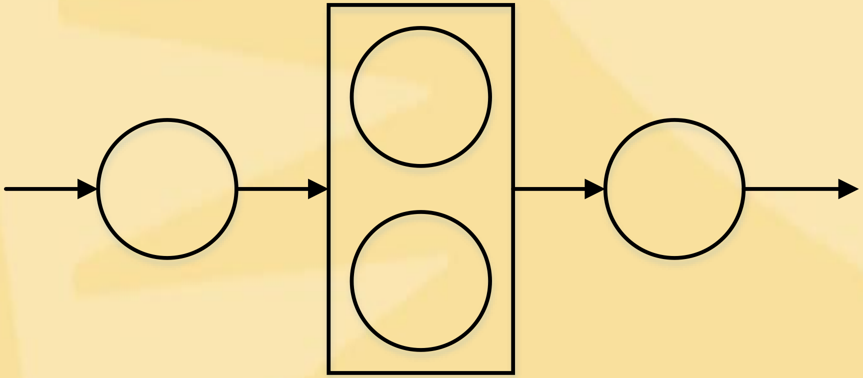
- Consider an interconnected system of servers
- Non-stationary arrivals, Markovian service processes (possibly intermittent or non-homogenous) and FCFS queueing discipline.
- L workstations and R routing chains (customer classes/part types).
- Infinite queue capacity and no migration.
- External and internal arrivals governed by an irreducible stochastic routing matrix, P_r. (i.e. process plans by part type)
- State of the system can be described by state vector

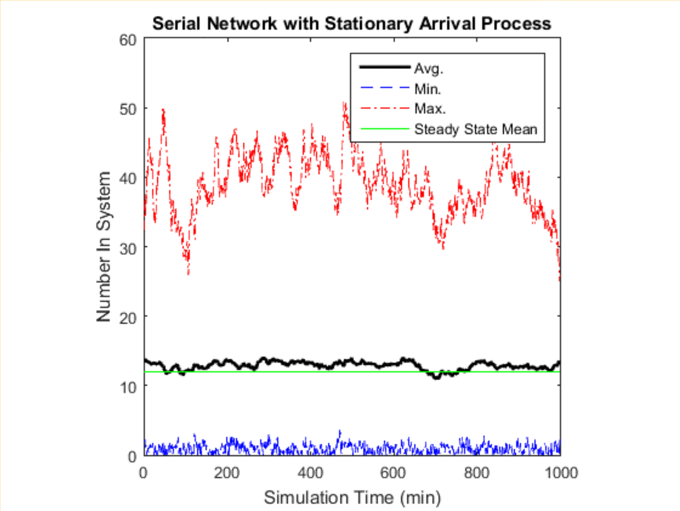
$$\mathbf{K} = (\mathbf{k}_1, \ldots, \mathbf{k}_R) \text{ where } \mathbf{k}_r = (k_{r,1}, \ldots, k_{r,L})$$

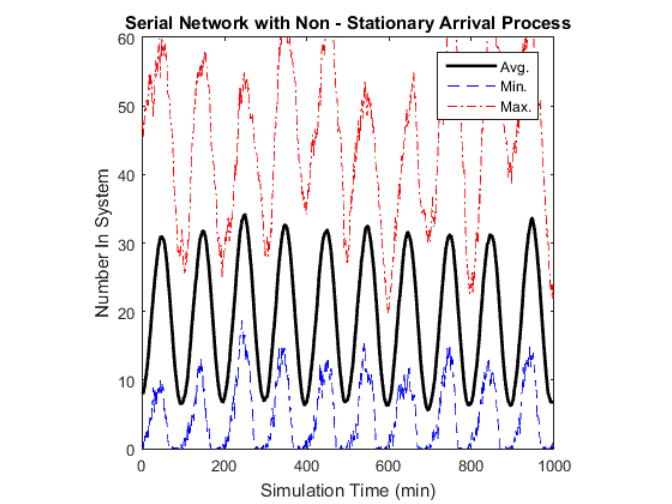
where $k_{r,l}$: number of class r jobs at workstation l

Impact of Non-stationarity

- Example 1: Three stage serial production line.



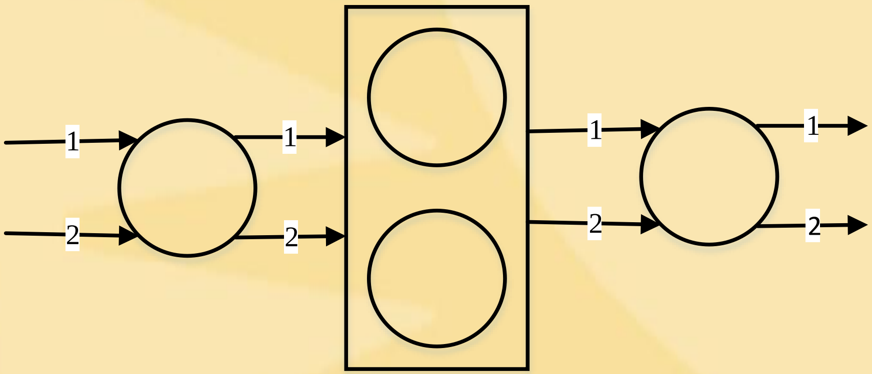




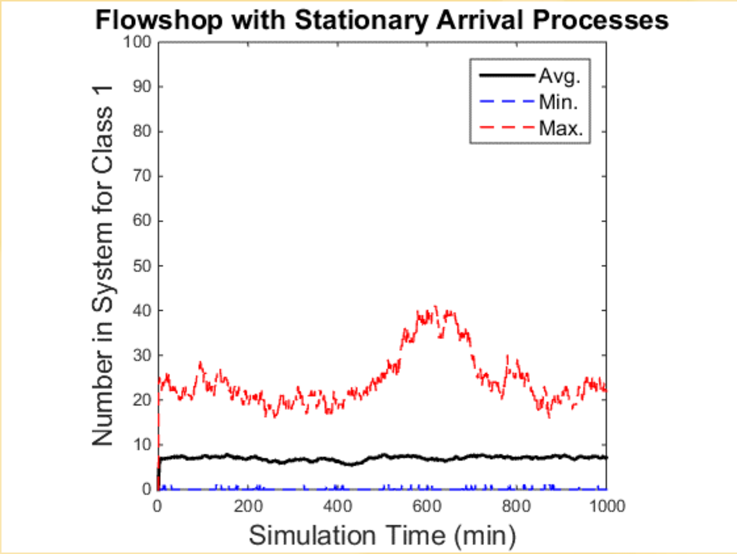
Sinusoidal Arrival Process

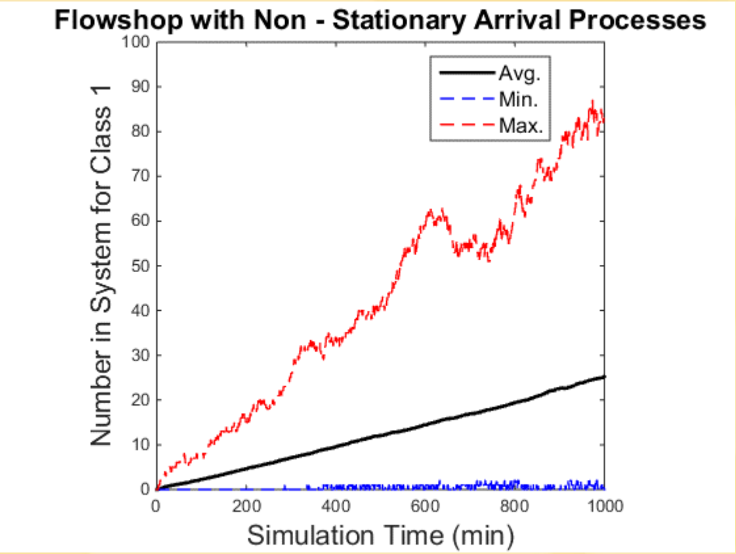
Impact of Non-stationarity

- Example 2: Three stage flow shop.



Arrival Process : Complementary Linear Non – Homogenous Poisson Processes





Literature Review : Nonstationary Queues

Some general results e. g. Heyman and Whitt (1984): dynamic steady state for $M_t/G/c$
Three broad categories of approximations

- Systems Approximations (period by period stationary approximation)
  - Green and Kolesar (1991): Pointwise stationary approximation (PSA)
  - Green et al. (2001): Stationary independent period-by-period (SIPP) approximation
  - Stolletz (2008): Stationary backlog carryover (SBC) approximation
- Numerical Approximations (Simplification assumptions and numerically compute)
  - Rothkopf and Oren (1979): Closure approximation
  - Grassmann (1977): Randomization method
- Process Approximations (Limiting heavy traffic fluid/diffusion methods)
  - Mandelbaum and Massey (1995)

Wang et al. (1996): Pointwise stationary fluid flow approximation(PSFFA)

Ingolfsson et al. (2007): Experimental comparison of seven service-level approximations for nonstationary queues.

Literature Review : Nonstationary Queueing Networks

Duda (1986): Parametric decomposition based on transient analysis of *GI/GI/1* queue

Massey and Whit (1993): Networks of infinite-servers queues with nonstationary Poisson input.

Malone (1995): Decomposition approximation for open networks with nonstationary input.

Mandelbaum and Massey (1995): Fluid and diffusion limits for large scale Markovian service networks

Whitt (1999): Generalized Jackson network based approximation framework for time-dependent Markovian networks-simplifying system of ODEs

Liu and Whitt (2013): Analysis of networks of time-varying many-server fluid queues.

Motivation

- Research focus so far has been on
  - Nonstationary queues
  - Queuing networks under stationary arrival assumptions or special conditions

- Need efficient algorithms for QN's under dynamic conditions.

- Incorporating network structure enables
  - Understanding evolution of congestion at different points
  - Modelling class priorities

- Focus is on first-order estimates of system performance.

System Approximation Research Methodology

- Total observation window T is broken down into a finite number of time epochs of equal length t_s, $s=1,..,T/t_s$.
- System dynamics are studied through snapshots of system performance tracked for each time epoch.
- Snapshots are a weighted combination of steady state performance metrics for two closed queueing networks, with the floor and ceiling levels of WIP (Basic Closed Model).

Research Methodology

- Steady state estimates are precomputed using an exact Mean Value Analysis (MVA) algorithm (see Reiser and Lavenberg (1980)).

- Characteristics of MVA
  - Provides steady state estimates for all intermediate levels of WIP(provides performance for all values of $k = 1,..,K_r$)
  - Complexity : $O(\prod_{r=1}^{R} K_r)$, where K_r is number of jobs of class ‘r’.

Research Methodology

- Step 1. Solve closed network, C($\mathbf{N}^*$) using the MVA algorithm. The state vector $\mathbf{N}^*$ is set to a sufficiently large value.

- Step 2. Initialize i=0, t=0 and the vector of total jobs in each routing chain for initial conditons, $\mathbf{N_0} = (n_1(0), n_2(0), \ldots, n_R(0))$ Set $X_r(0) = X_r^c(\mathbf{N}_0)$

Research Methodology

- Step 3. Set $i = i + 1$, $t = t + t_s$.

  Update mean number in routing chain $r = 1, \ldots, R$ as

$$n_r(t) = \max(n_r(t - t_s) + t_s \sum_{j=1}^{L} \lambda_{rj}(t - t_s, t) - t_s \cdot X_r(t - t_s), 0)$$

(Starting + Arrivals – Completions; or 0)

  If $t < T$ go to step 4 else STOP.

Research Methodology

- Step 4. For each $r = 1, \ldots, R$

  a) Update throughput rate as

$$X_r(t) = (1-\alpha_r) X_r^c(\mathbf{N}_l) + \alpha_r X_r^c(\mathbf{N}_u)$$, where

$$\mathbf{N}_l = (\lfloor n_1(t) \rfloor, \lfloor n_2(t) \rfloor, \ldots, \lfloor n_R(t) \rfloor)$$,

$$\mathbf{N}_u = (\lceil n_1(t) \rceil, \lceil n_2(t) \rceil, \ldots, \lceil n_R(t) \rceil)$$, and $\alpha_r = n_r(t) - \lfloor n_r(t) \rfloor$

b) Update mean total time in system as

$$W_r(t) = \frac{n_r(t)}{X_r(t)}$$ (Little's law for each routing chain)

c) Update Cumulative production as

$$\chi_r(t) = \chi_r(t-t_s) + \min(X_r(t) \cdot t_s, (n_r(t-t_s) + t_s \sum_{j=1}^{L} \lambda_{rj}(t-t_s, t)))$$

- Step 5. Go to step 3.

Computational Analysis

- Method pieces together snapshots of how a stationary system would perform at each time step with the given WIP level.

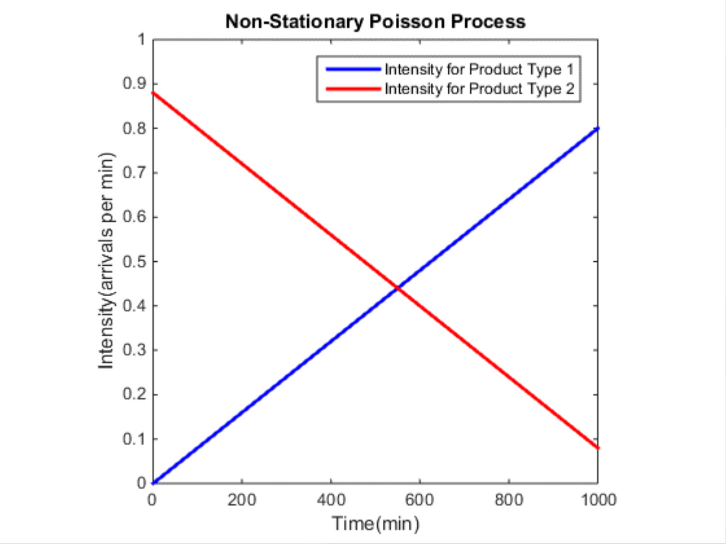
- Approximation MVA assumes distribution of jobs across workstations not representative of actual distribution for non-homogenous process.

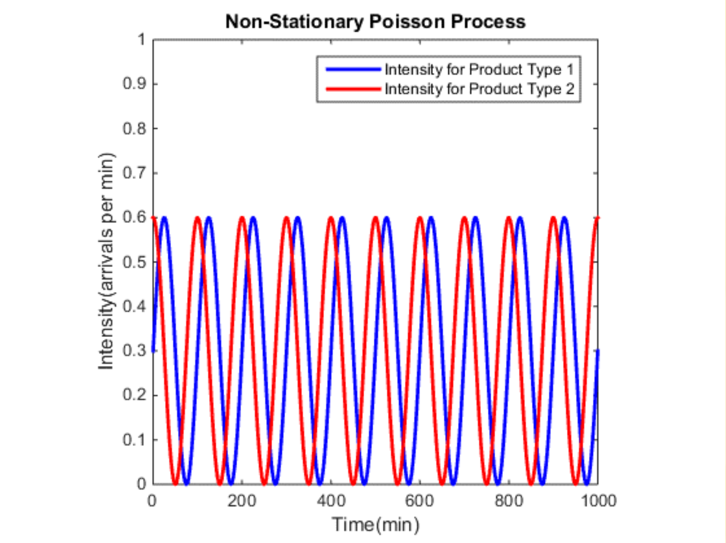
- Approximation was applied to a simple serial system and a jobshop with four classes and sixteen workstations.

- First order moments of interest were compared against simulation results.

Experiments

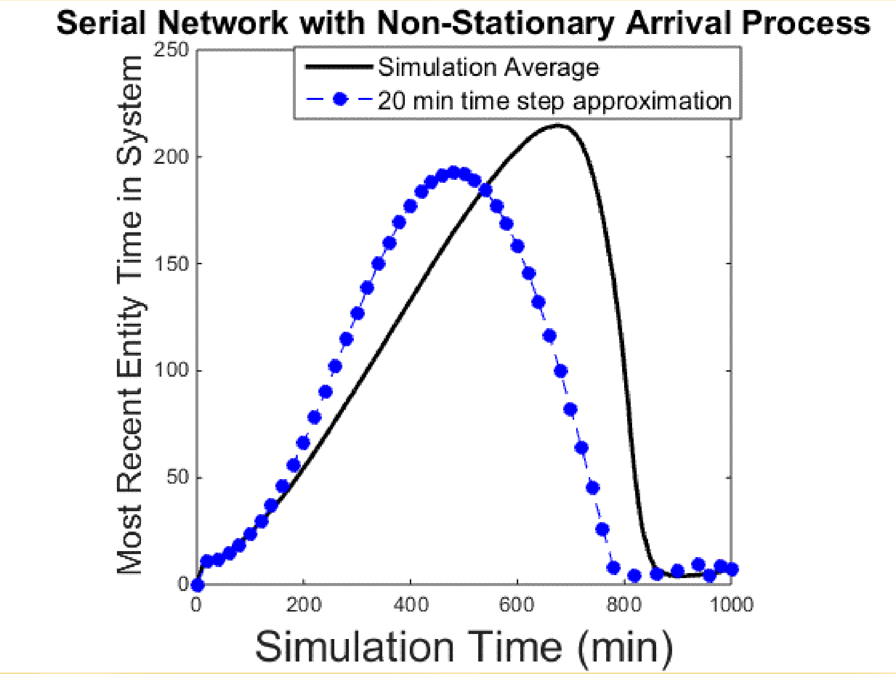
- Models
  - 1a.
    - Type: Three workstation serial line
    - Arrival Process: Non – homogenous Sinusoidal Poisson Process with $2\pi = 1000\,\text{min}$.
  - 1b.
    - Type: Three workstation serial line
    - Arrival Process: Non – homogenous Sinusoidal Poisson Process with $2\pi = 100\,\text{min}$.
  - 2.
    - Type: Two class three workstation flow shop
    - Arrival Processes: Complementary non-homogenous linear Poisson process
  - 3.
    - Type: Two class three workstation flow shop
    - Arrival Processes: Complementary non-Homogenous sinusoidal Poisson process.

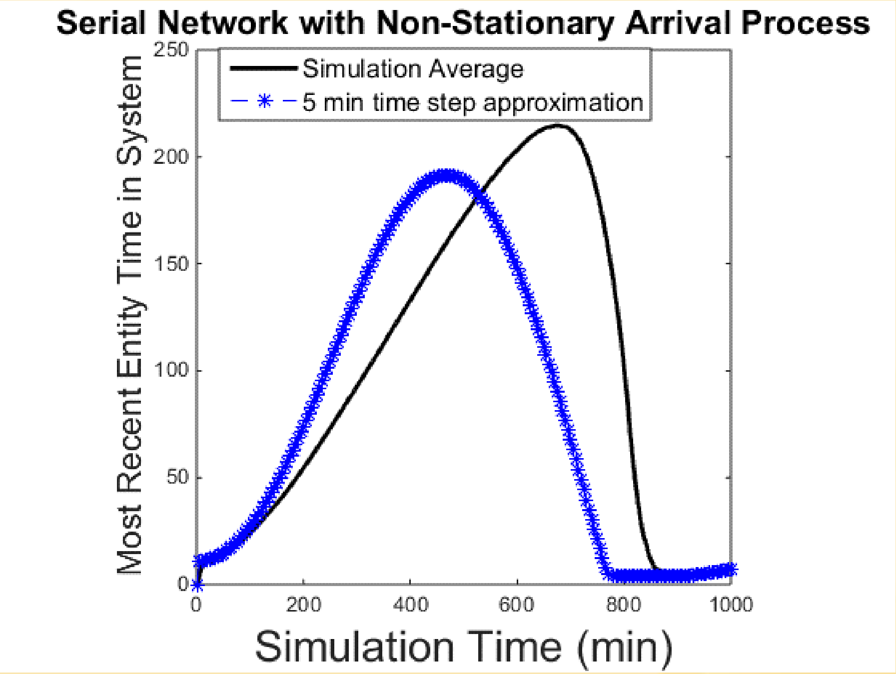




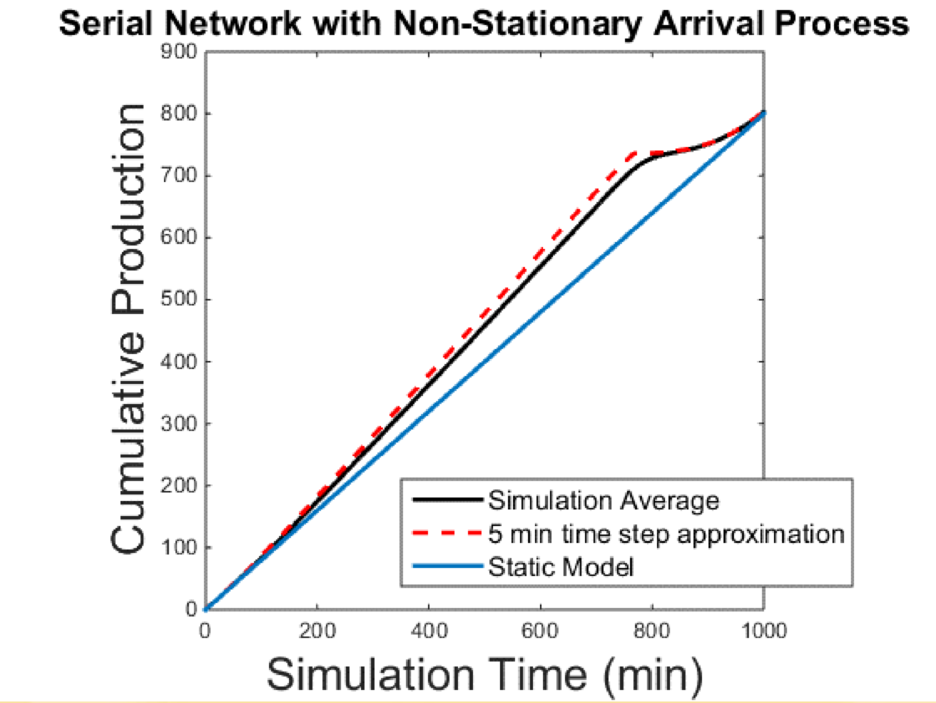
Model 1a :Time in system

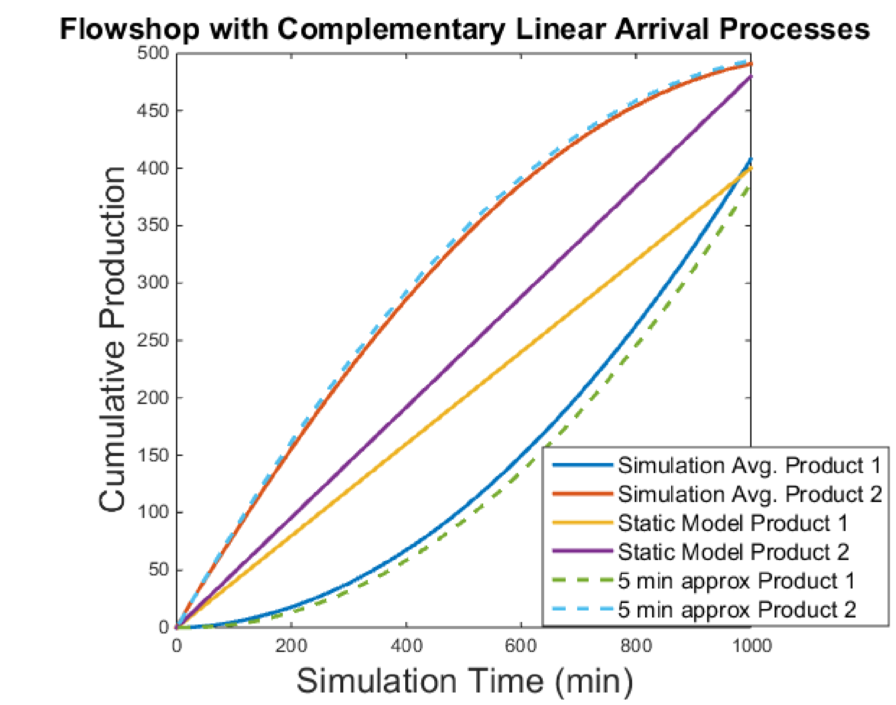
$$2\pi = 1000\,\text{min}.$$





Cumulative production results





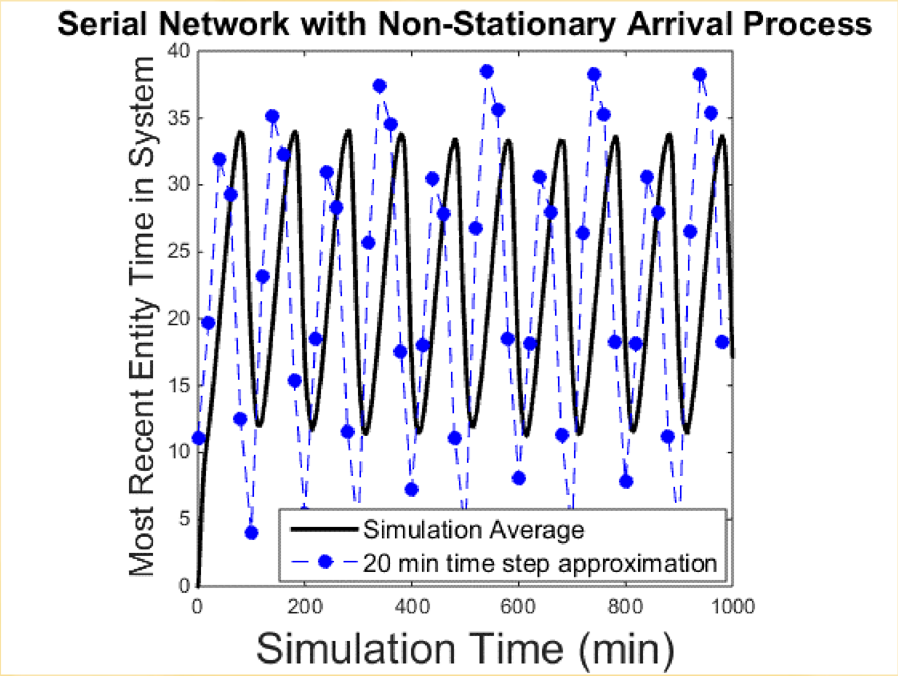
$$2\pi = 1000\,\text{min}.$$

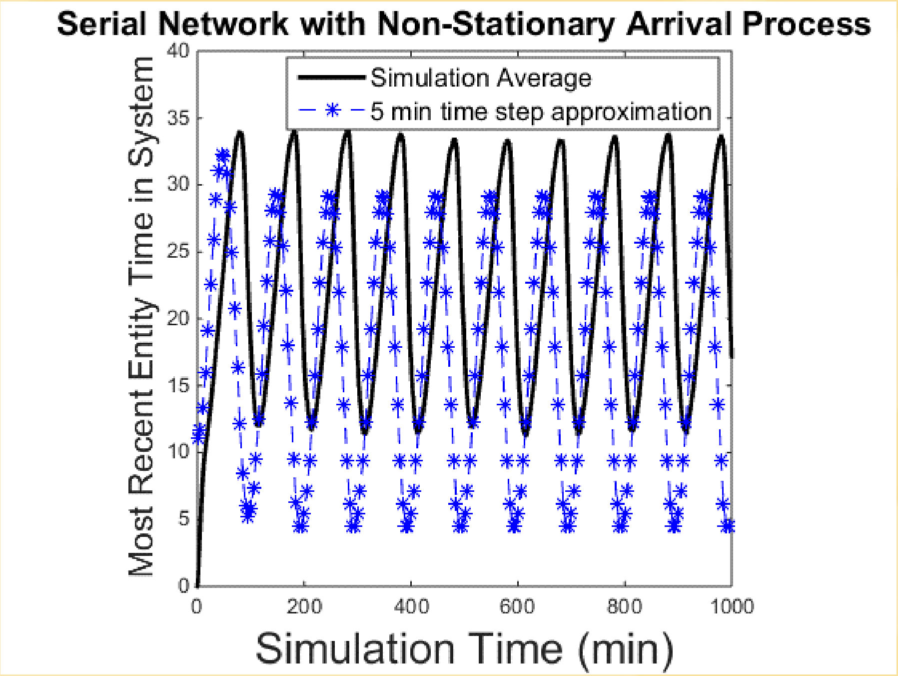
Model 1a

Model 2

Model 1b: Time in system

$$2\pi = 100\,\text{min}.$$

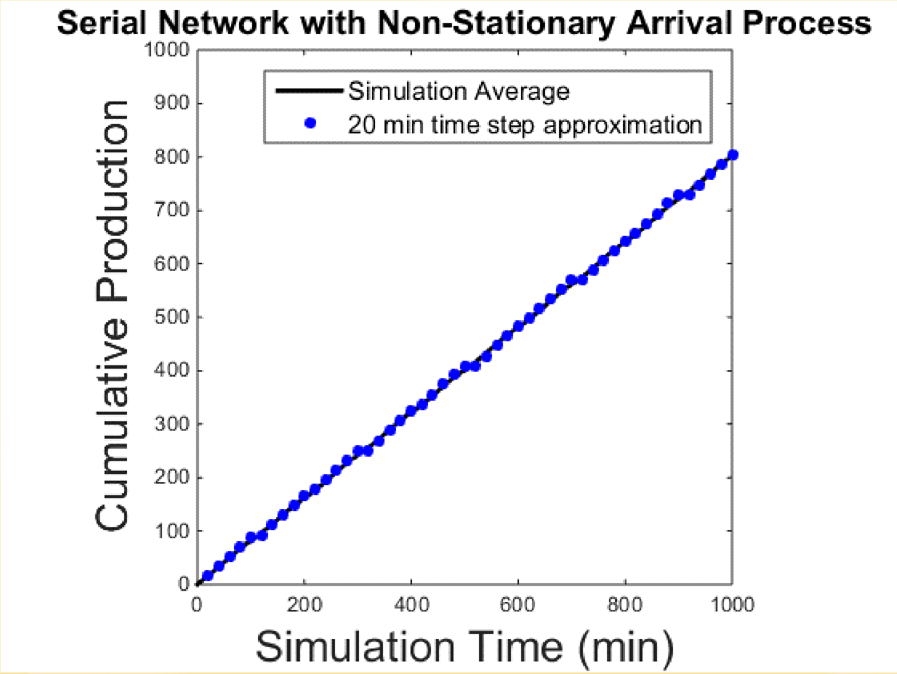


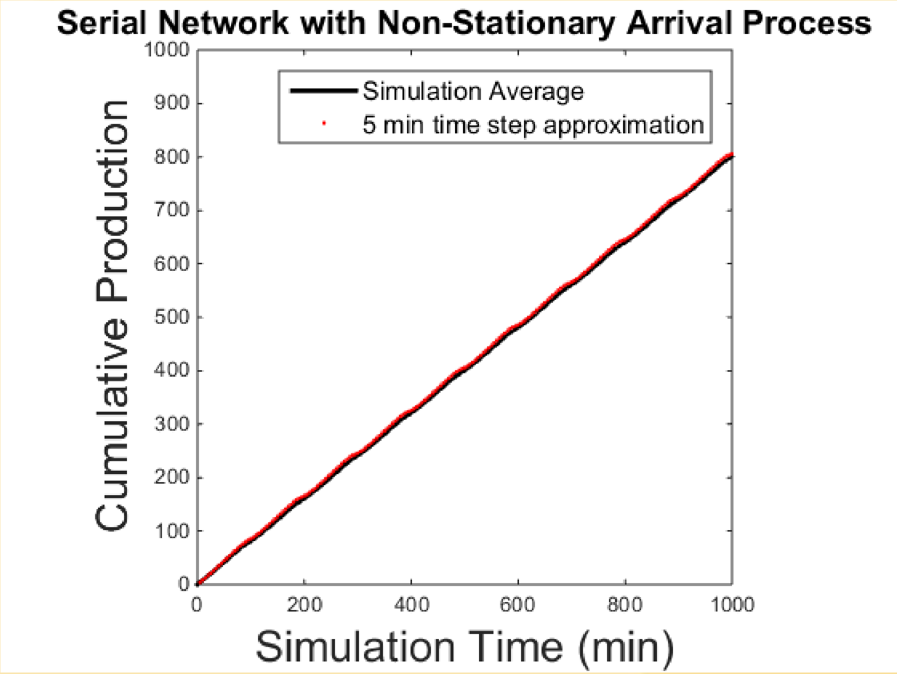


Note: Static model gives a constant time in system of 15.56 mins

Model 1b: Cumulative production

$$2\pi = 100\,\text{min}.$$

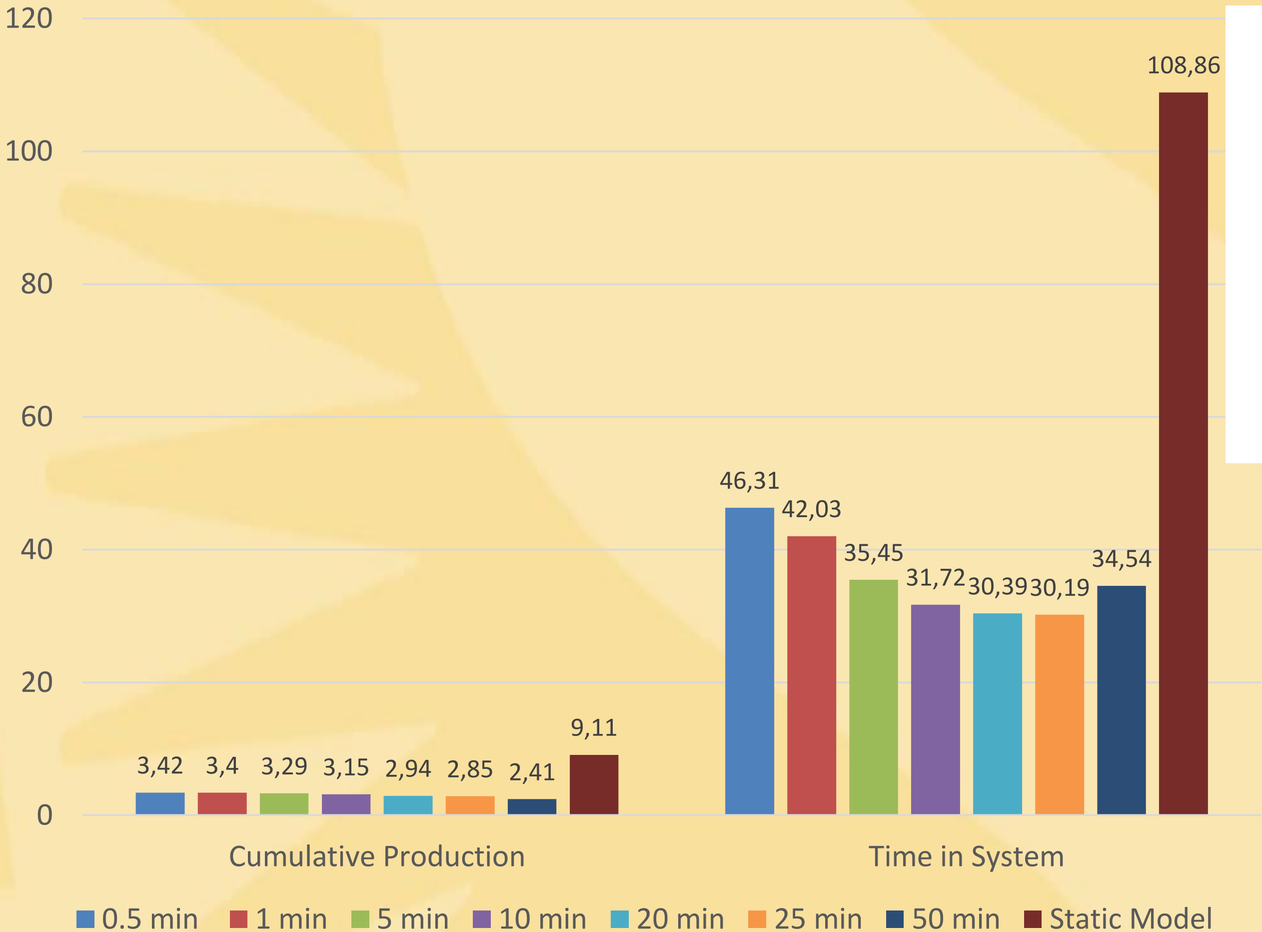




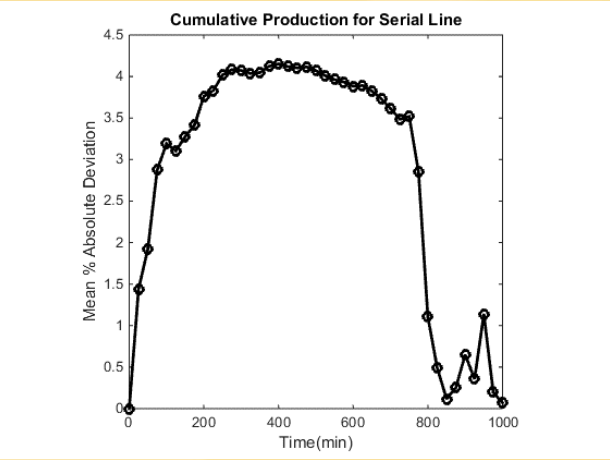
Model 1b: Absolute Deviation

$$2\pi = 100\,\text{min}.$$

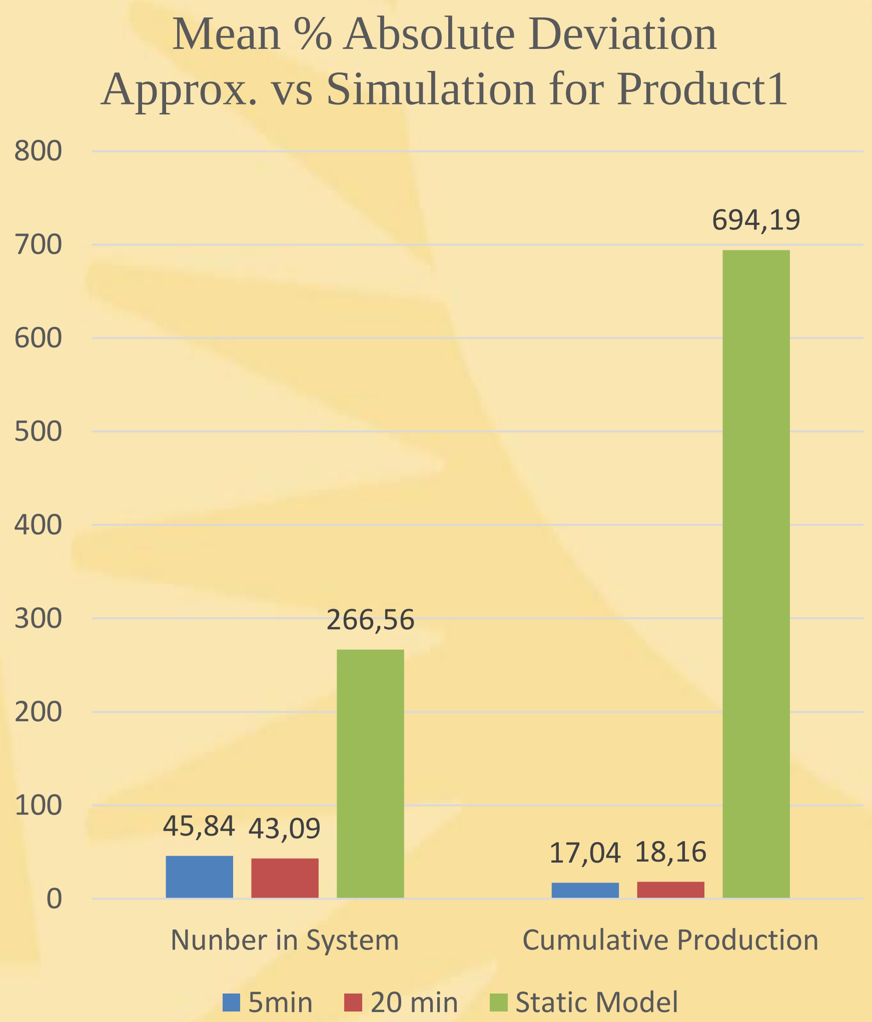
Mean % Absolute Deviation Approx. vs Simulation

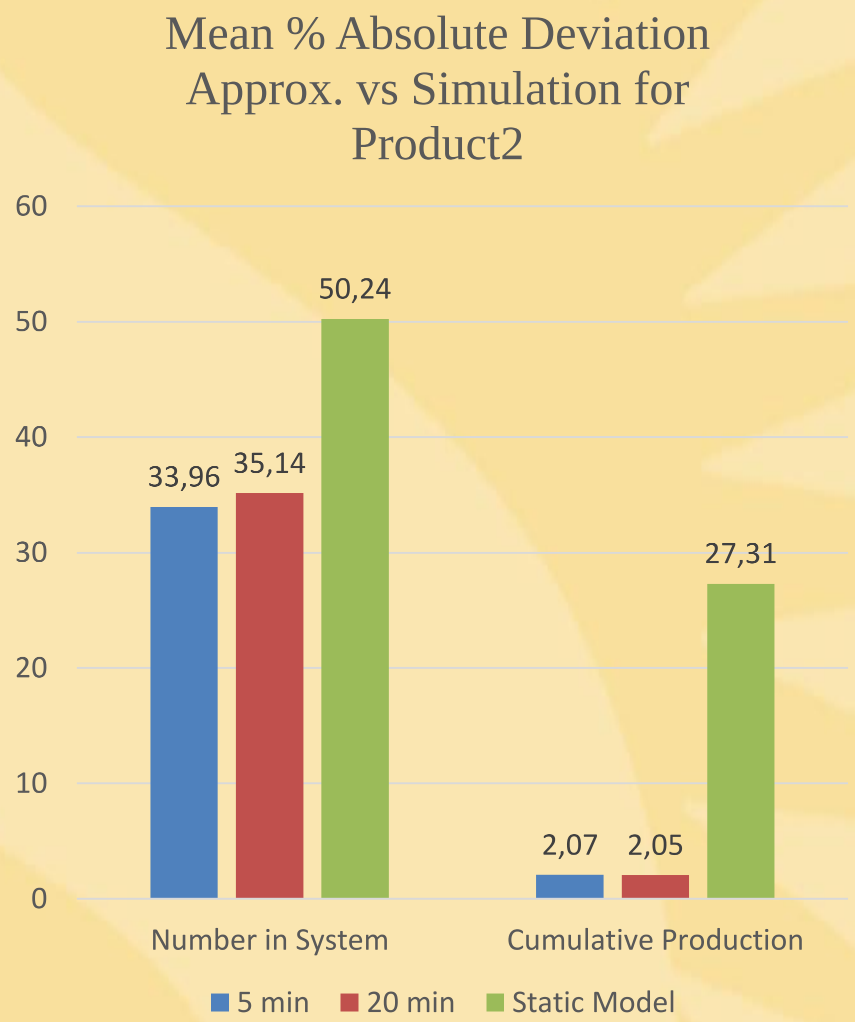


Absolute Deviation

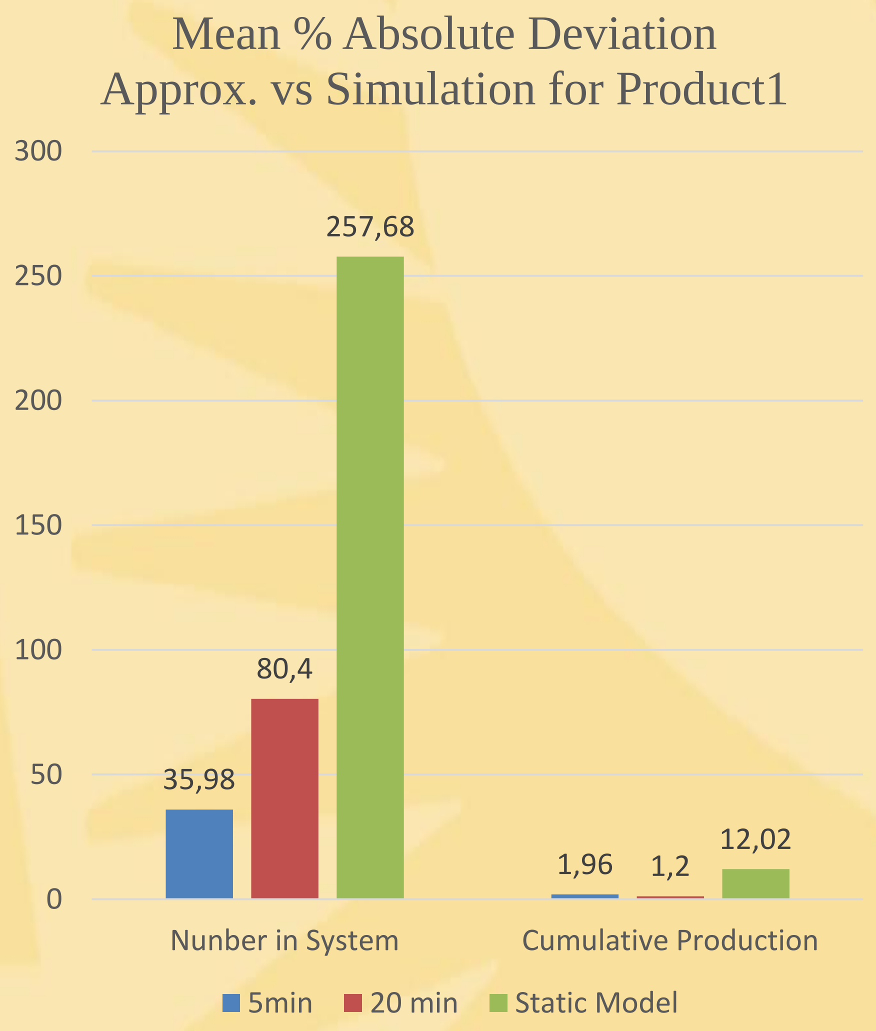


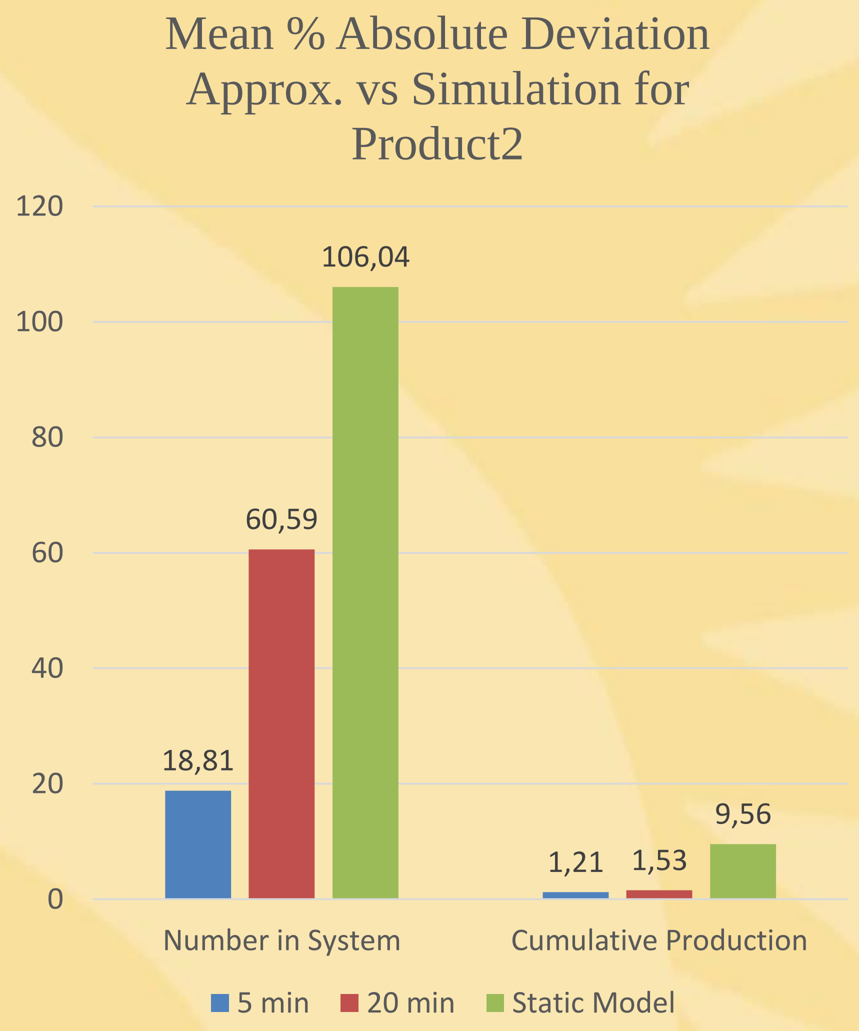
Model 2: Absolute Deviation





Model 3: Absolute Deviation





Enhancements to Base model : Motivation

- Implicit Dynamic distribution of jobs ignored in previous model.
- Steady state assumptions don't hold under transient conditions.
- Instantaneous throughput rates of individual workstations need not be identical.

Route Workstation Based Throughput Model (RWBTM)

- Step 1

Use MVA to solve the closed queueing network $C(\mathbf{N}^*)$ ($\mathbf{N}^*$ sufficiently large).

- Step 2

Initialize $i = 0,\ t = 0$.
Set $\mathbf{N_0} = (n_1(0),\ \ldots,\ n_R(0))$ to match an equivalent open queueing network.
Set $X_r(0) = X_r{}^c(\mathbf{N_0})$ for each $r = 1, \ldots, R$.
Set $n_{rl}(0) = n^c_{rl}(\mathbf{N_0}),\ \gamma_r l(0) = 0,\ \psi_{rl}(0) = 0,\ X_{rl}(0) = \nu_{rl} X^c_r(\mathbf{N_0})$ for each $r = 1, \ldots, R$ and each $l \in S(r)$.

- Step 3

Set $t = t + t_s$

Route Workstation Based Throughput Model (RWBTM)

- Step 4

For each $r = 1, \ldots, R$ and each $l \in S(r)$

1. Compute mean number of jobs of routing chain r leaving workstation l at time t

$$\psi_{rl}(t) = min(n_{rl}(t - t_s) + t_s \lambda_{rl}(t - t_s, t), t_s X_{rl}(t - t_s))$$

2. Compute total internal arrivals of routing chain r at workstation l at time t

$$\gamma_{rl}(t) = \sum_{j=1}^{L} \psi_{rj}(t) p_{jl}^{r}$$

Route Workstation Based Throughput Model (RWBTM)

- Step 5

Update cumulative production at time t for each $r = 1, ..., R$

$$\chi_r(t) = \chi_r(t - t_s) + \sum_{l=1}^{L} \left[\left(1 - \sum_{j=1}^{L} p_{lj}^{r} \right) \psi_{rl}(t) \right]$$

Route Workstation Based Throughput Model (RWBTM)

- Step 6

For each $r = 1, \ldots, R$ and $l \in S(r)$

1. Update mean number in routing chain r at workstation l as

$$n_{rl}(t) = n_{rl}(t - t_s) + t_s \lambda_{rl}(t - t_s, t) + \gamma_{rl}(t) - \psi_{rl}(t) \tag{1}$$

2. Obtain closest CQN, $\boldsymbol{N_l^*(t)}$ for each workstation l as

$$\boldsymbol{N_l^*}(t) = \left\{ \boldsymbol{N} : \ \boldsymbol{n_l^*(t)} = min_{\boldsymbol{N} \in \boldsymbol{N^*}} \left\| \boldsymbol{n}_l(t) - \boldsymbol{n}_l^c(\boldsymbol{N}) \right\|_2 \right\} \tag{2}$$

3. Update throughput rate for routing chain r at workstation l as

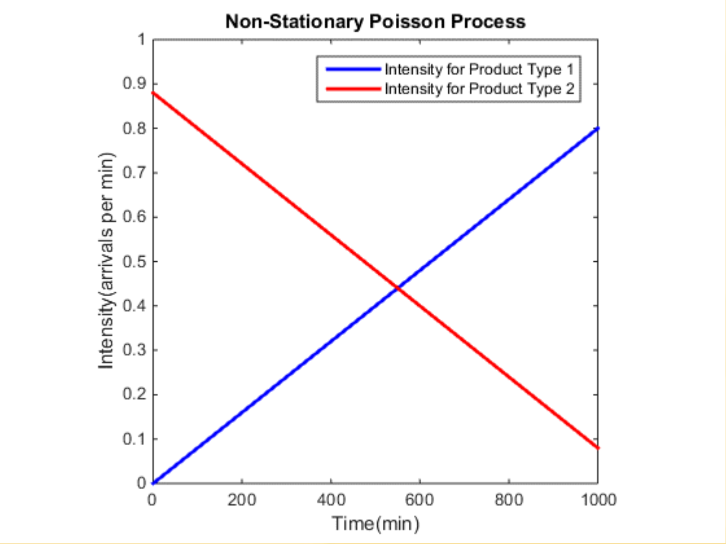
$$X_{rl}(t) = \nu_{rl} X_r^c(\boldsymbol{N_l^*}(t)) \tag{3}$$

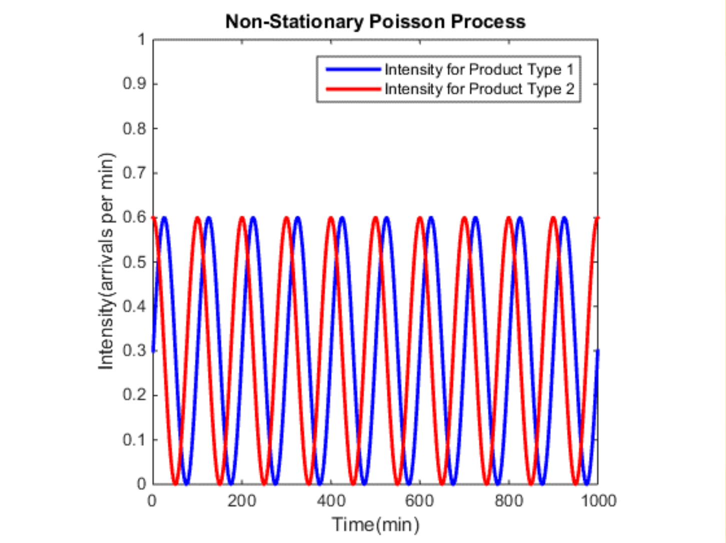
- Step 7

If $t < T$ GO TO 3. Else STOP.

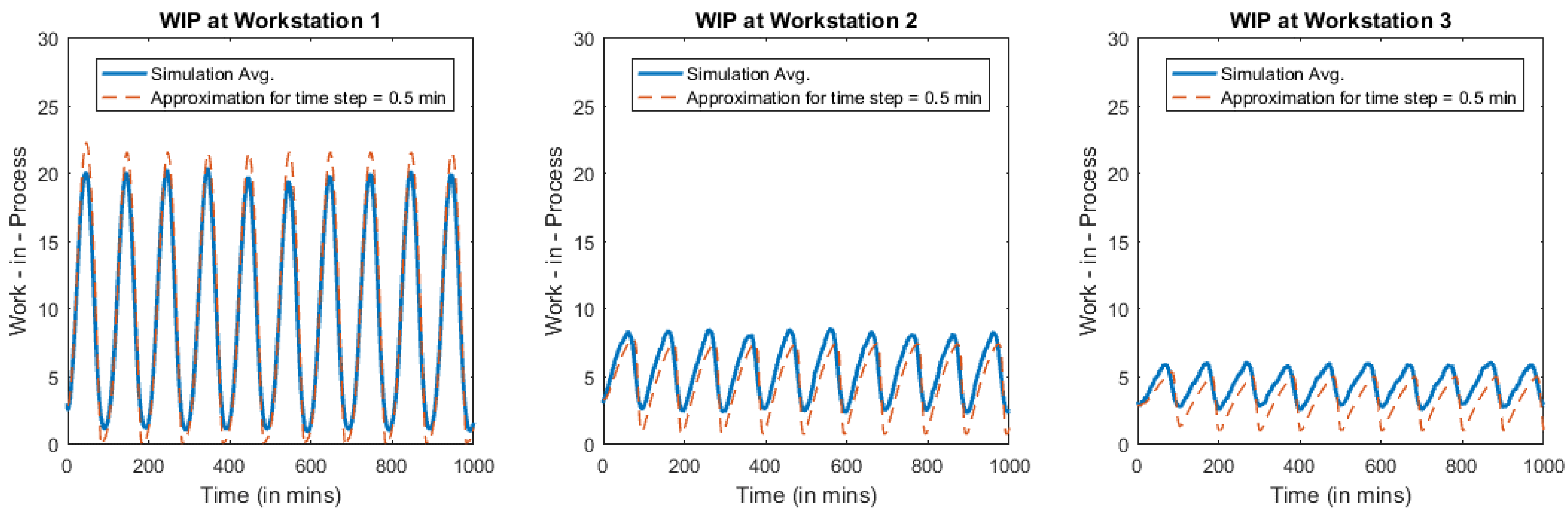
Experiments

- Models
  - 1a.
    - Type: Three workstation serial line
    - Arrival Process: Non – homogenous Sinusoidal Poisson Process with $2\pi = 1000\,\text{min}$.
  - 1b.
    - Type: Three workstation serial line
    - Arrival Process: Non – homogenous Sinusoidal Poisson Process with $2\pi = 100\,\text{min}$.
  - 2.
    - Type: Two class three workstation flow shop
    - Arrival Processes: Complementary non-homogenous linear Poisson process
  - 3.
    - Type: Two class three workstation flow shop
    - Arrival Processes: Complementary non-Homogenous sinusoidal Poisson process.

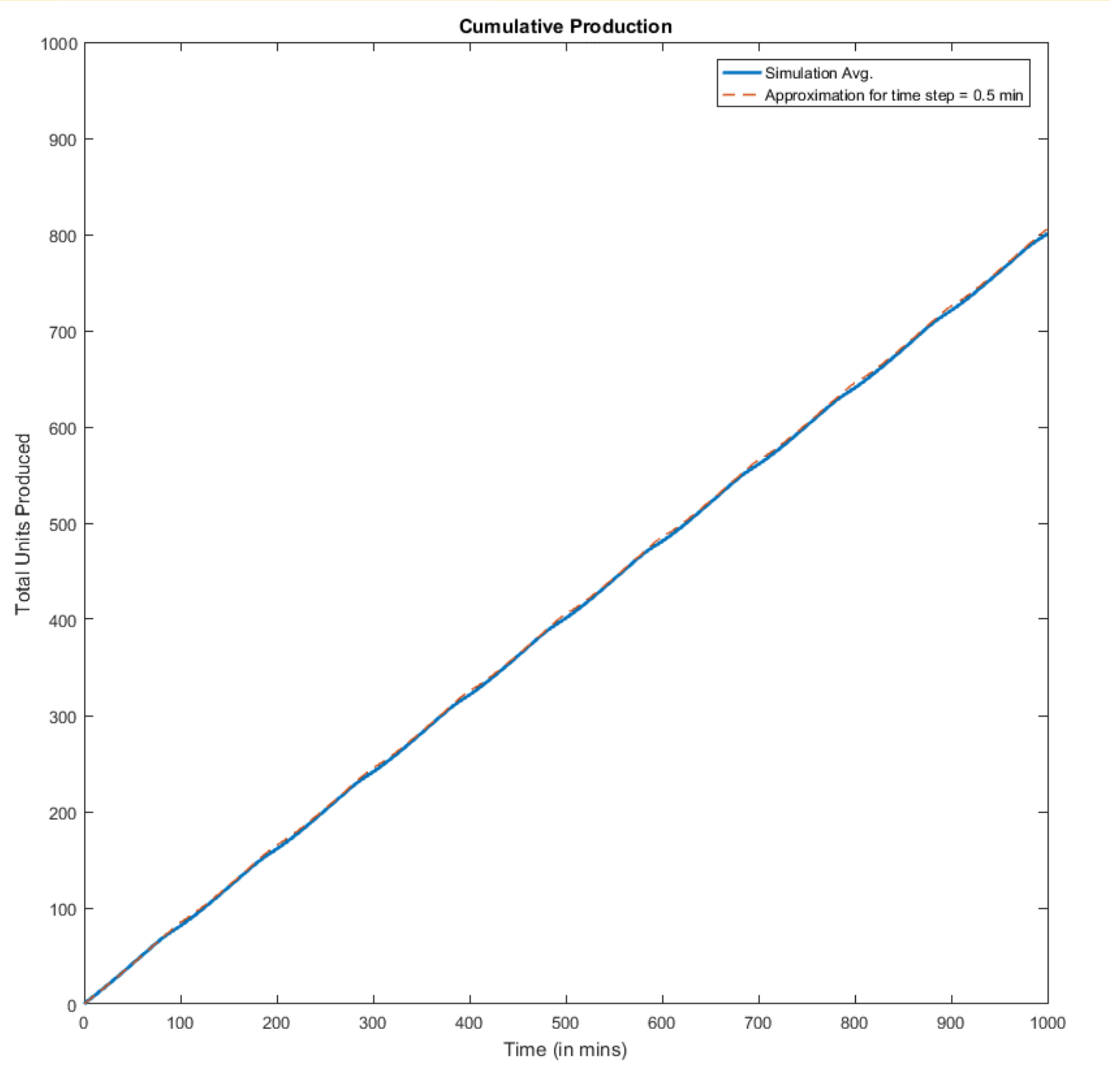




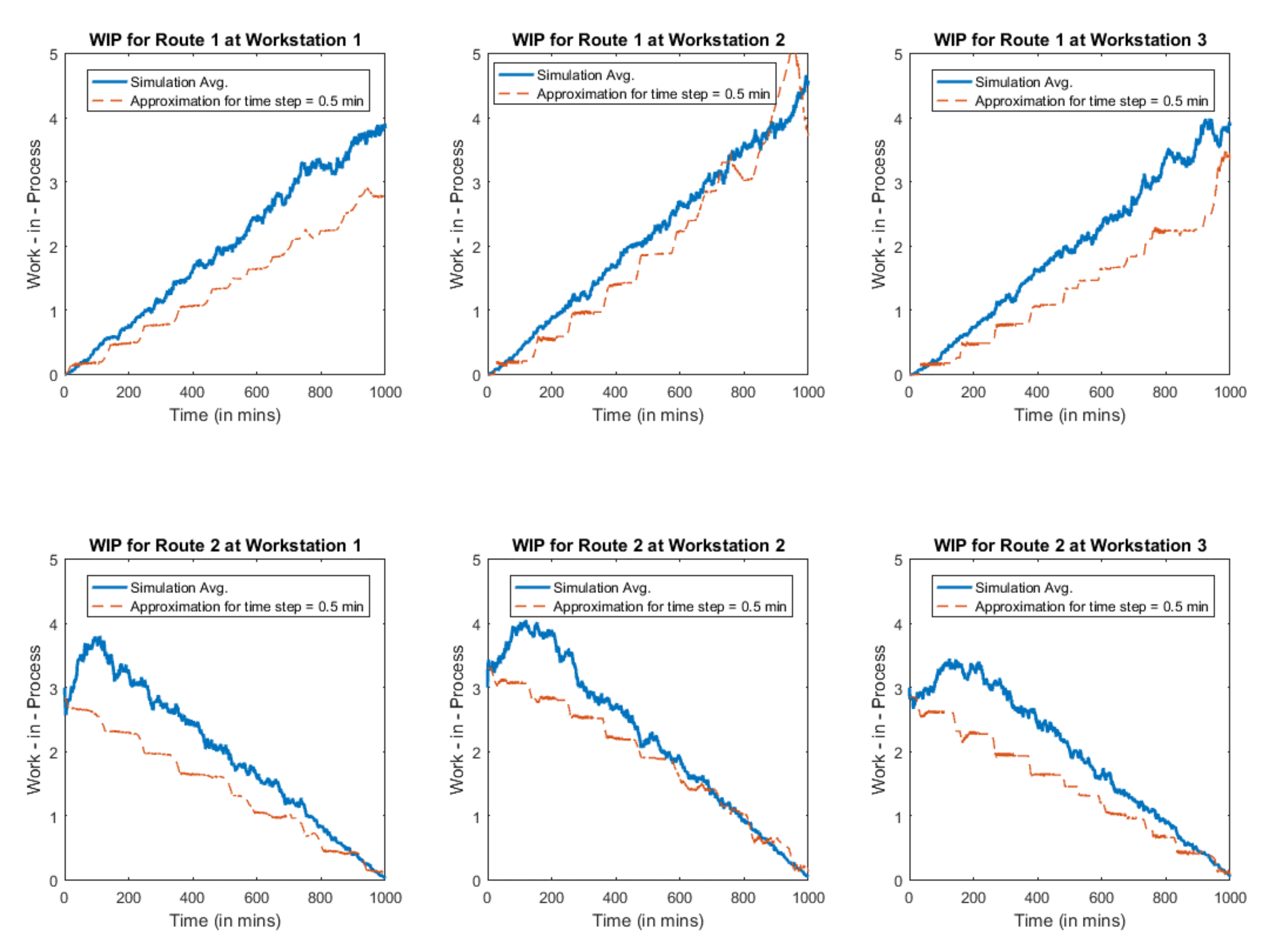
RWBTM: Model 1b WIP



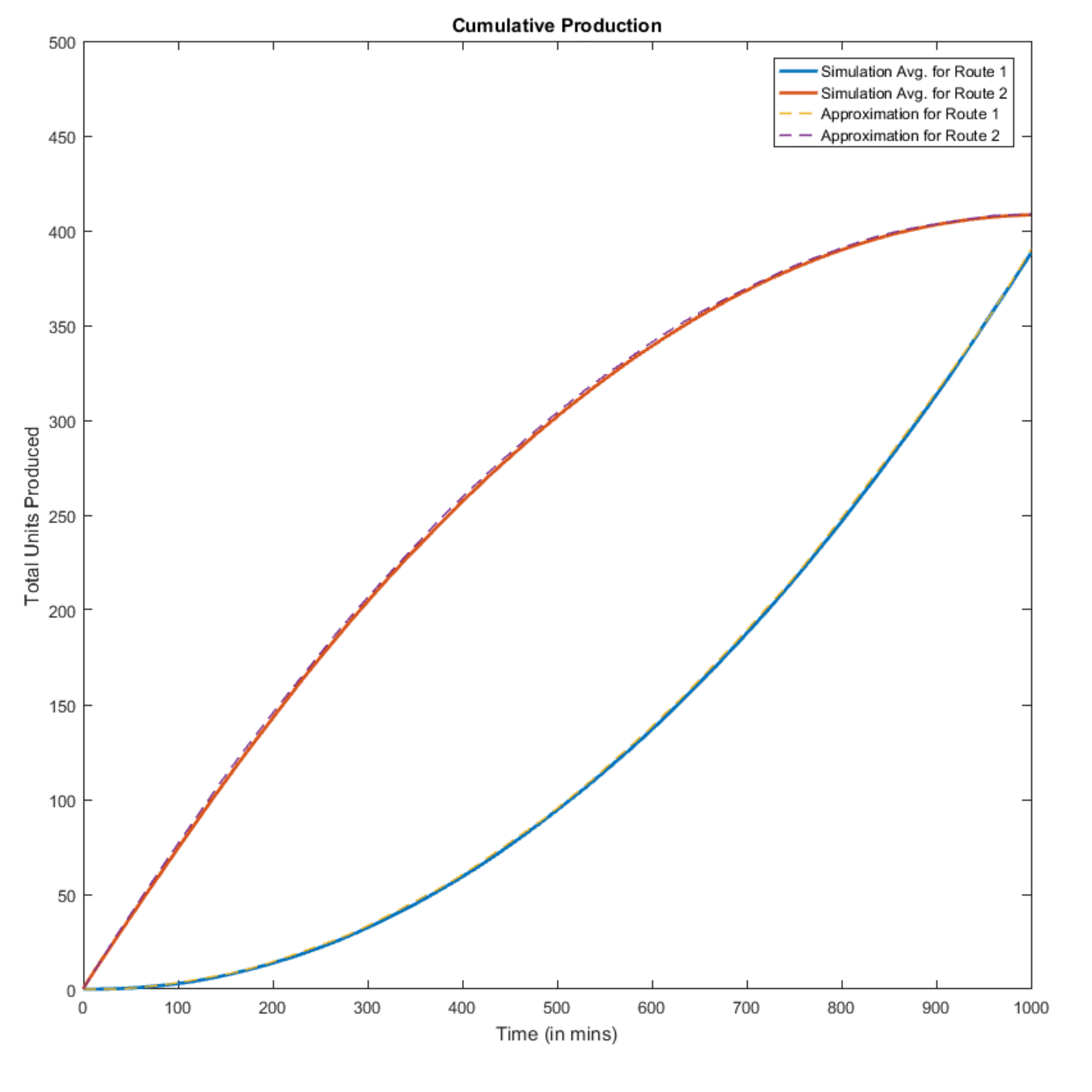
Model 1b: Cumulative production



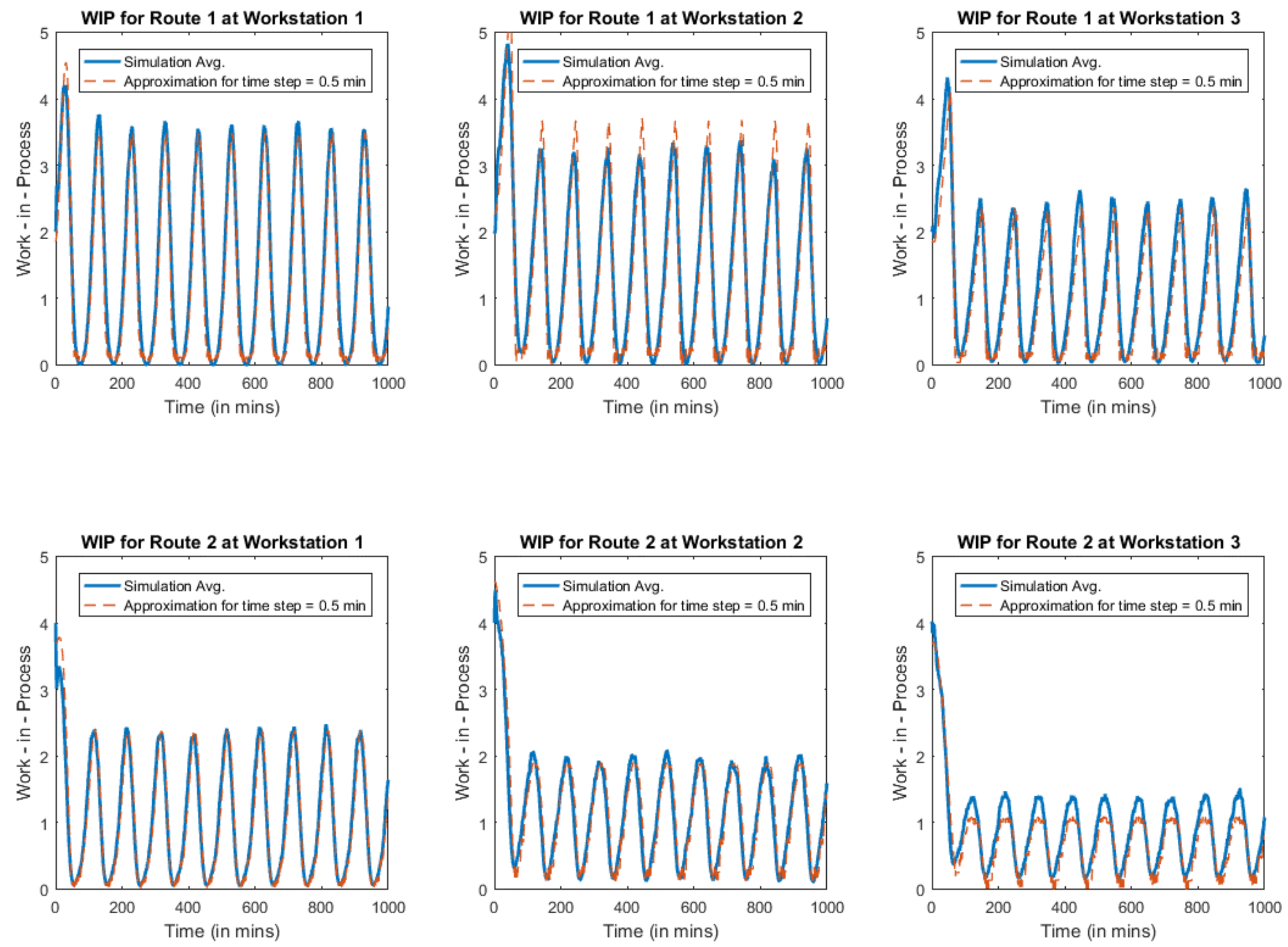
Model 2: Work - in - Process



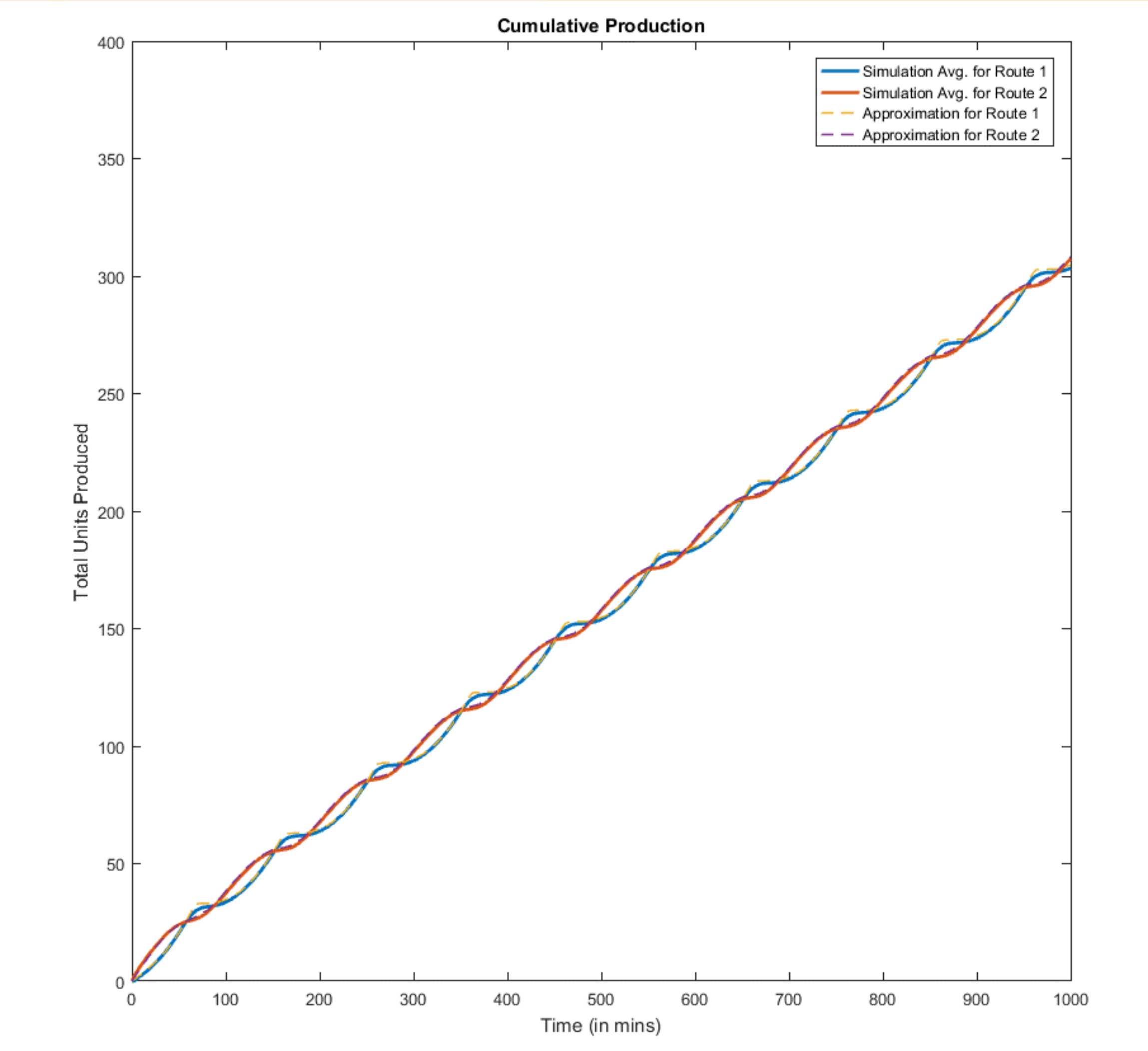
Model 2: Cumulative production



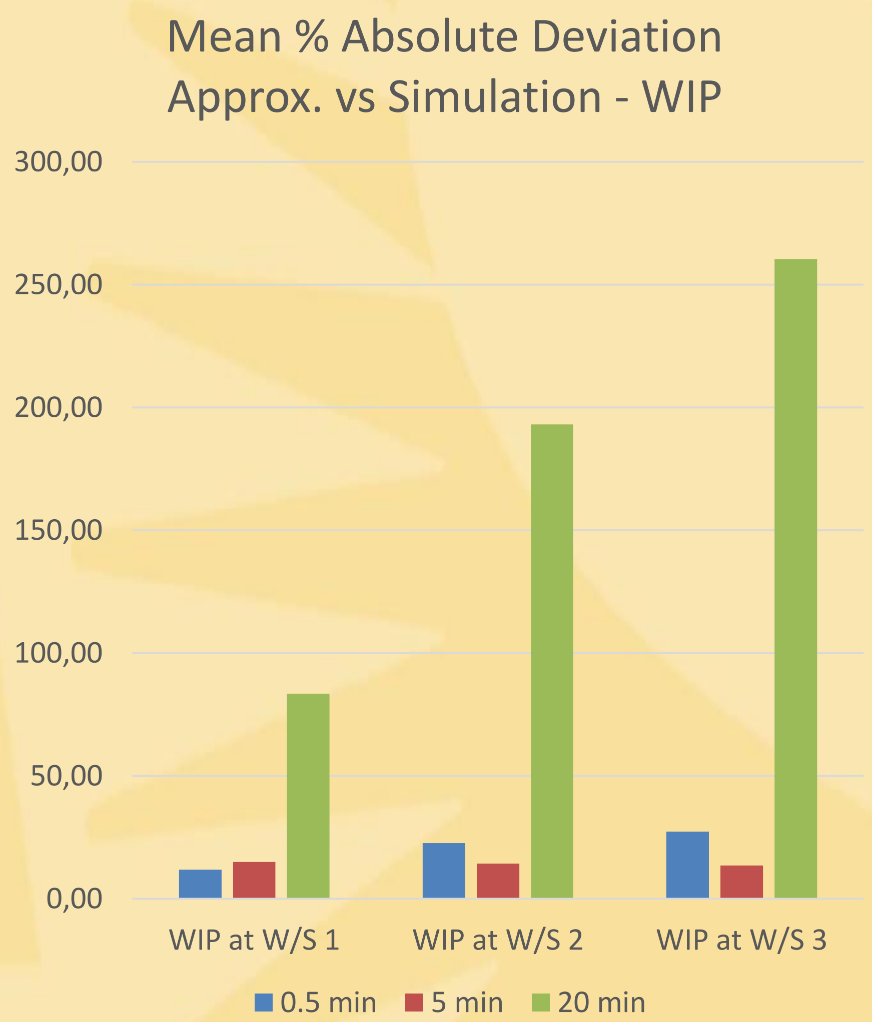
Model 3: Work - in - Process

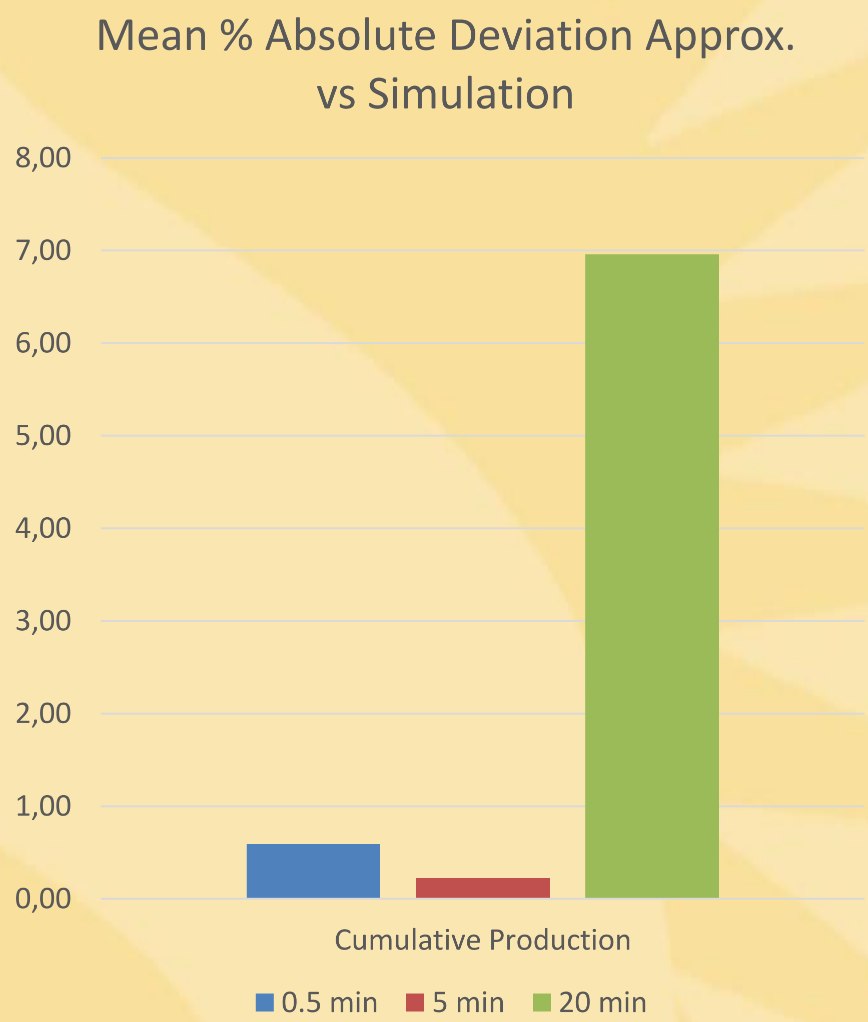


Model 3: Cumulative production

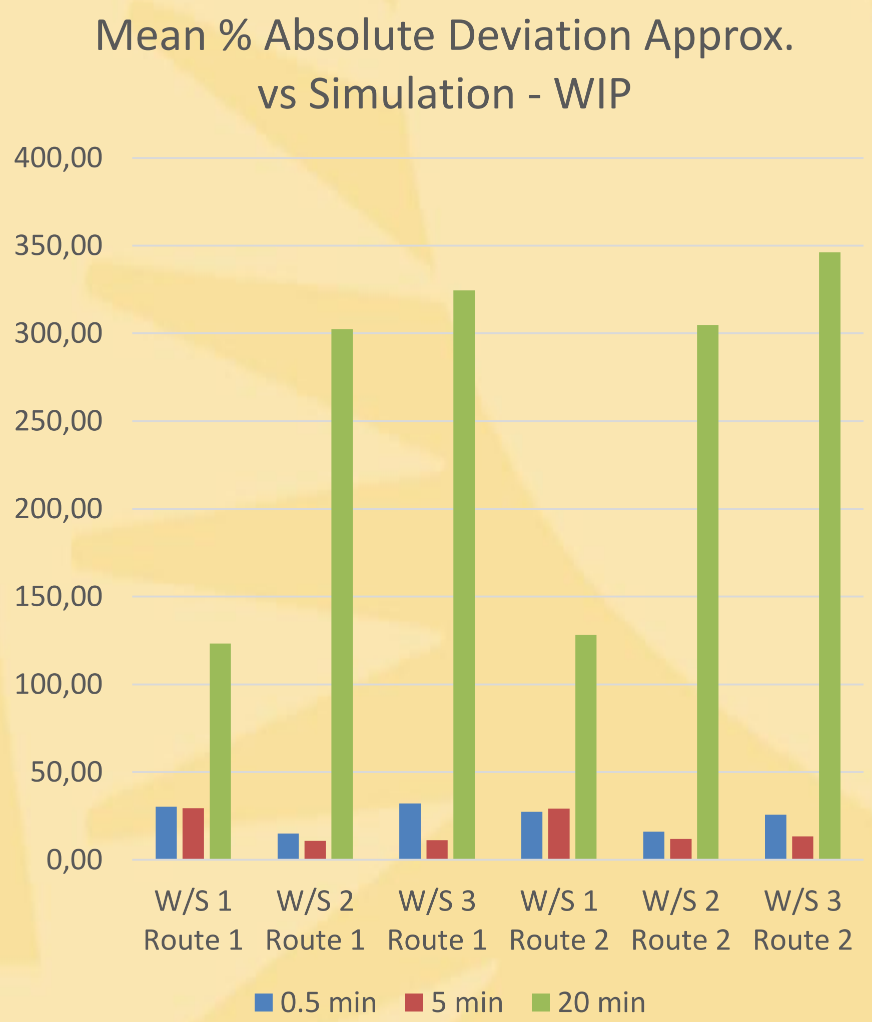


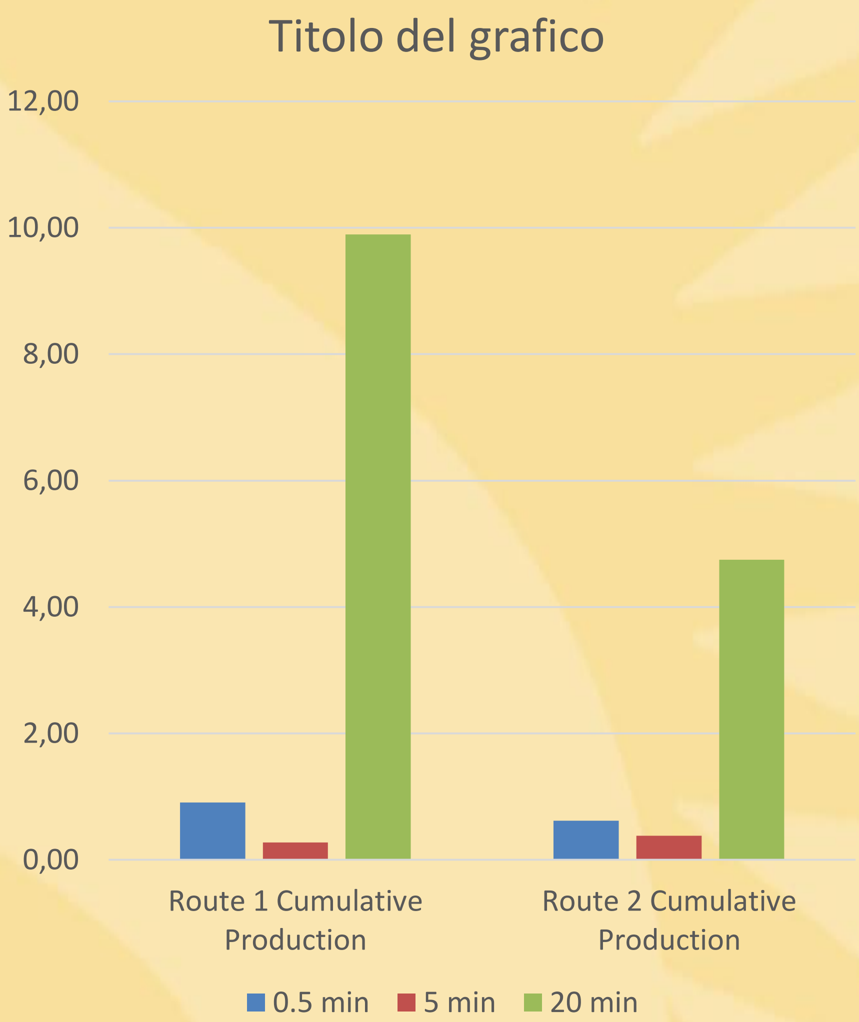
Model 1a: Absolute Deviation



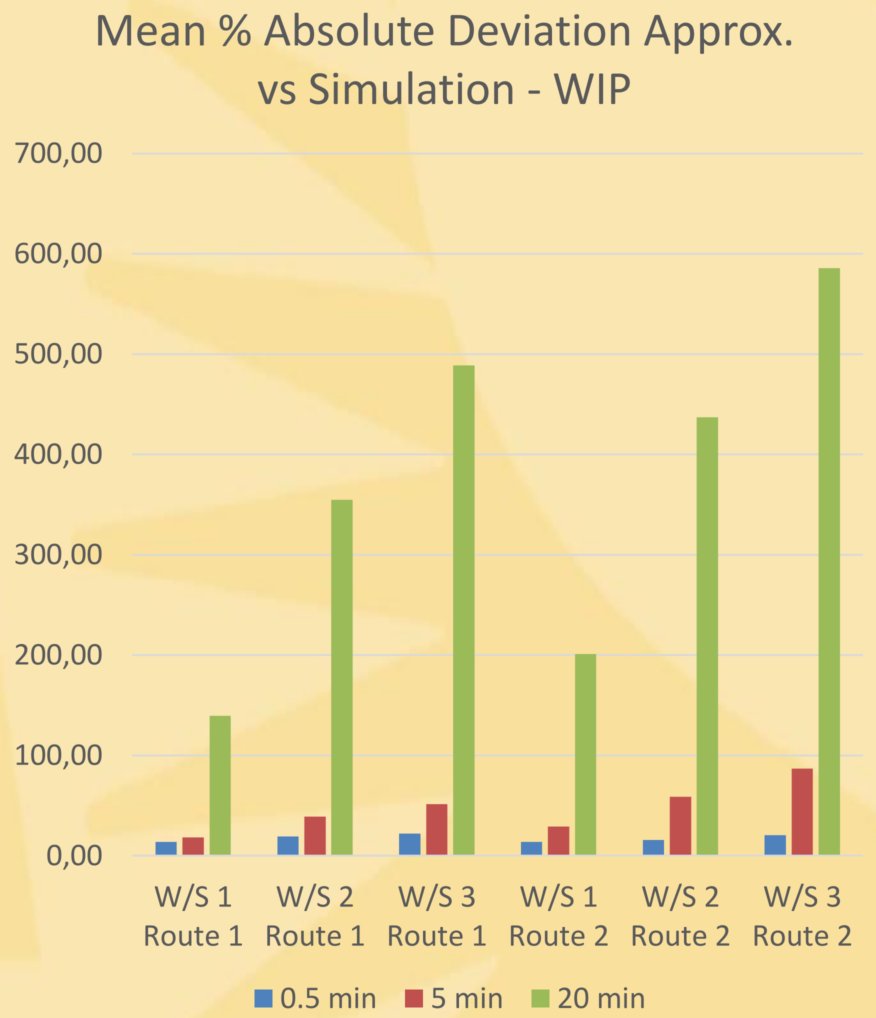


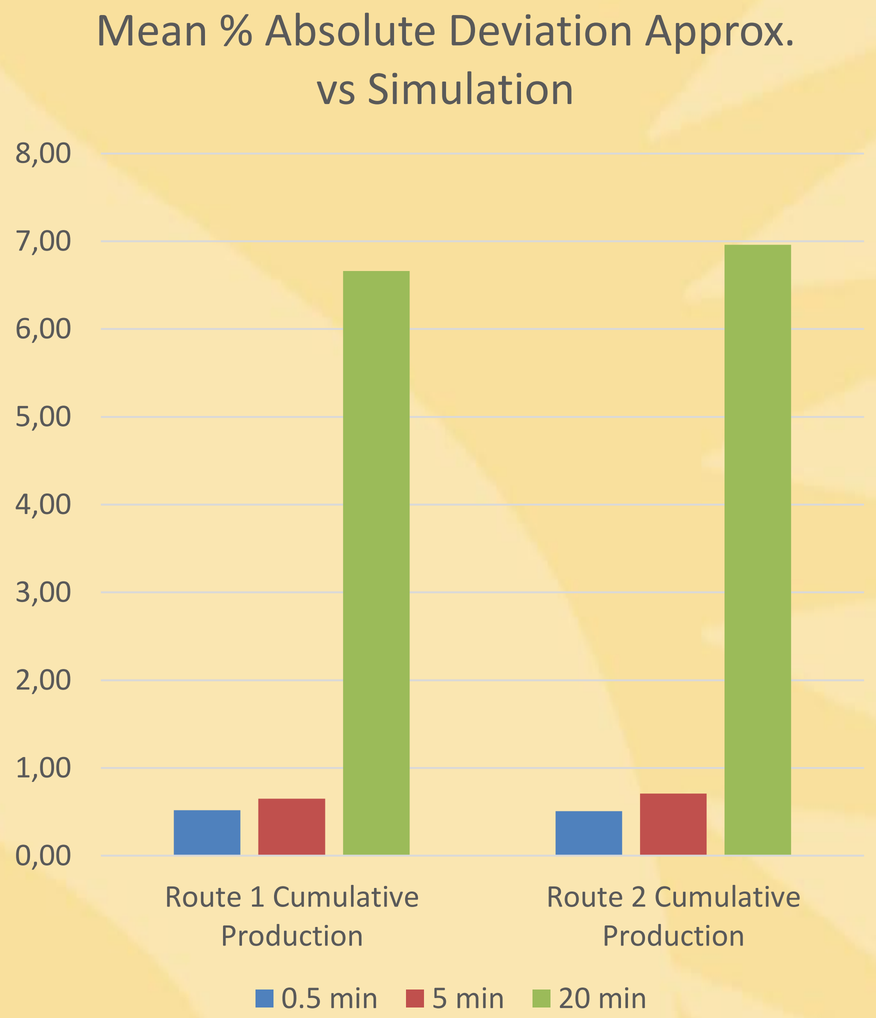
Model 2: Absolute Deviation





Model 3: Absolute Deviation





Open Network Based Throughout Model(ONBTM)

$\lambda_{ij}^r(t),\ r \in \{0,\ldots,R\},\ i \in \{0,\ldots,L\},\ j \in \{1,\ldots,L\}$: rate of class r arrivals at station j from station i(station 0 for external arrivals).

- Step 1

Initialize n_{rl}.

- Step 2

Update throughput for class r at workstation l for each $r = 1, \ldots, R$ and $l \in S(r) \setminus \{0\}$

$$X_{rl}(t) = min\left[\left(\frac{n_{rl}(t)}{\sum_{p=1}^{R} n_{pl}(t) + 1}\right)\mu_{rl}, \left(\frac{n_{rl}(t) + \frac{t_s \lambda_{0l}^r(t)}{2}}{t_s}\right)\right]$$

Min of (Effective allocated production resource; Available WIP)

Open Network Based Throughout Model(ONBTM)

- Step 3

For each $r = 1, ..., R$ and $l \in S(r) \setminus \{0\}$ (0 is external)

Update arrival rates of class r arriving at workstation l from workstation k

$$\lambda^r_{kl}(t) = p^r_{kl} X_{rk}(t),\ \ k \in S(r) \setminus \{0\}$$

- Step 4

Update WIP for each $r = 1, ..., R$ and $l \in S(r) \setminus \{0\}$

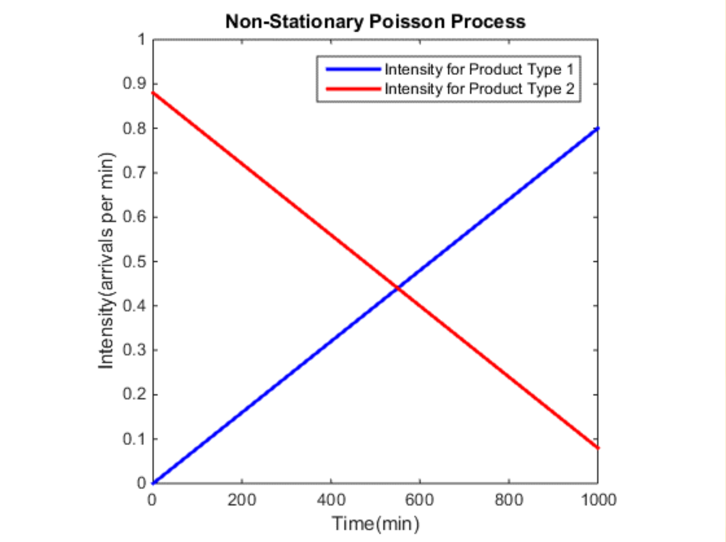
$$n_{rl}(t + t_s) = max\left[n_{rl}(t) + t_s\left(\sum_{j=0}^{L} \lambda^r_{jl}(t) - \sum_{j=0}^{L} \lambda^r_{lj}(t)\right), 0\right]$$

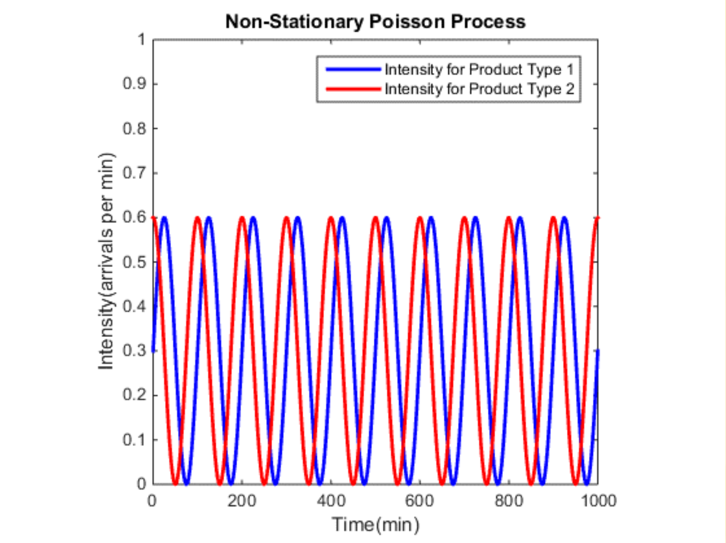
- Step 5

Set $t = t + t_s$. If $t < T$ GO TO 2. Else STOP.

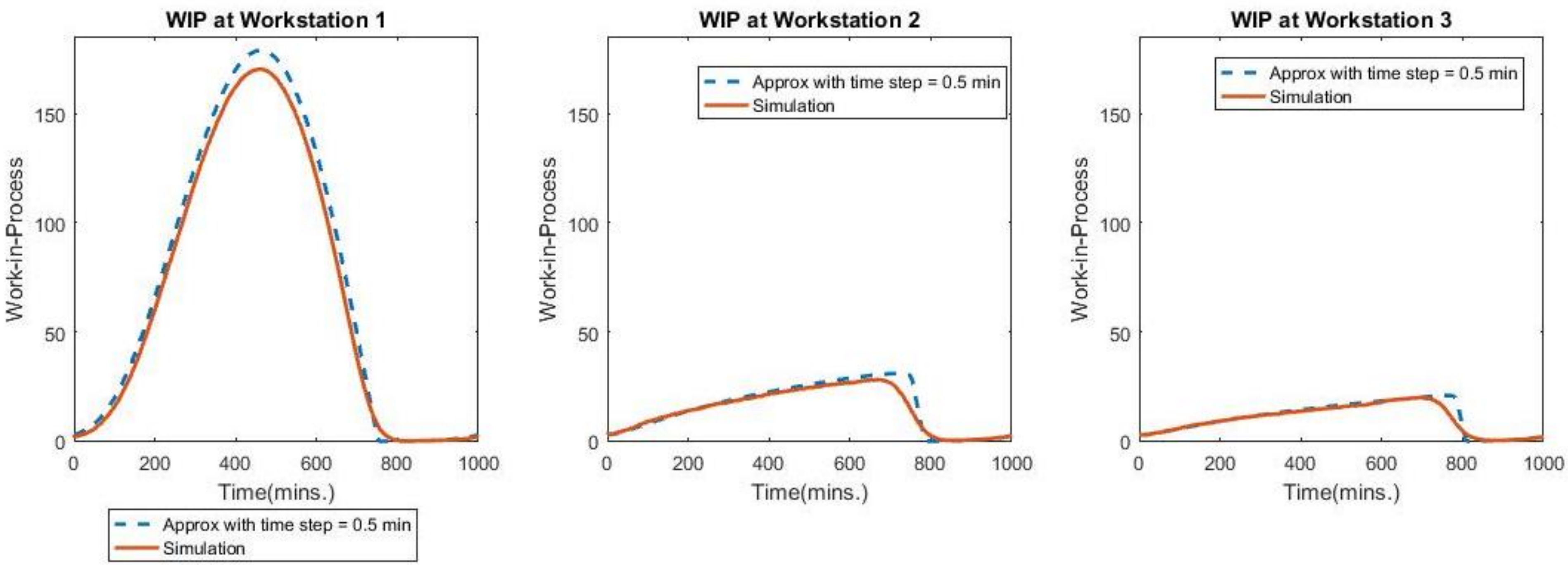
Experiments

- Models
  - 1a.
    - Type: Three workstation serial line
    - Arrival Process: Non – homogenous Sinusoidal Poisson Process with $2\pi = 1000\,\text{min}$.
  - 1b.
    - Type: Three workstation serial line
    - Arrival Process: Non – homogenous Sinusoidal Poisson Process with $2\pi = 100\,\text{min}$.
  - 2.
    - Type: Two class three workstation flow shop
    - Arrival Processes: Complementary non-homogenous linear Poisson process
  - 3.
    - Type: Two class three workstation flow shop
    - Arrival Processes: Complementary non-Homogenous sinusoidal Poisson process.

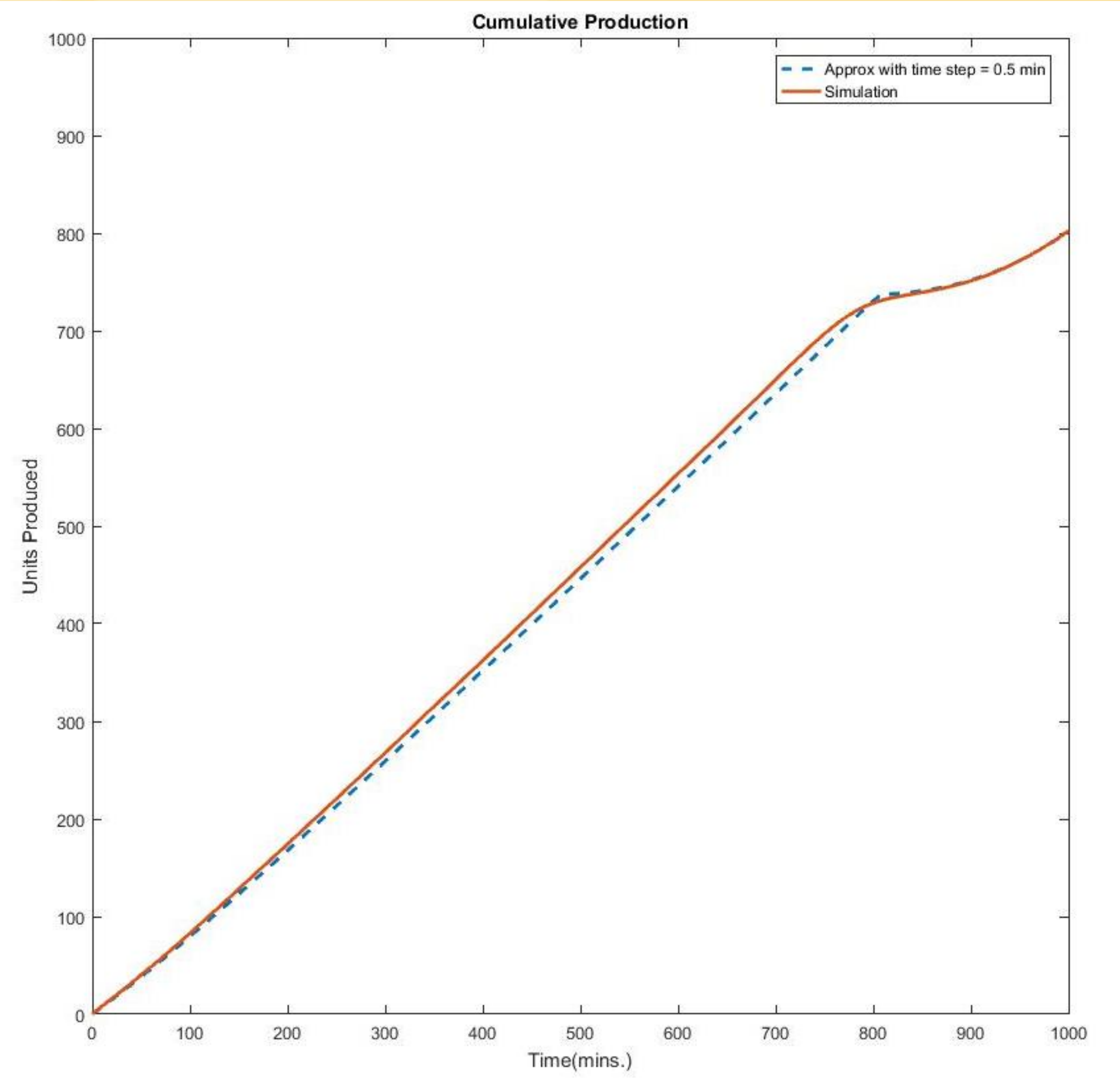




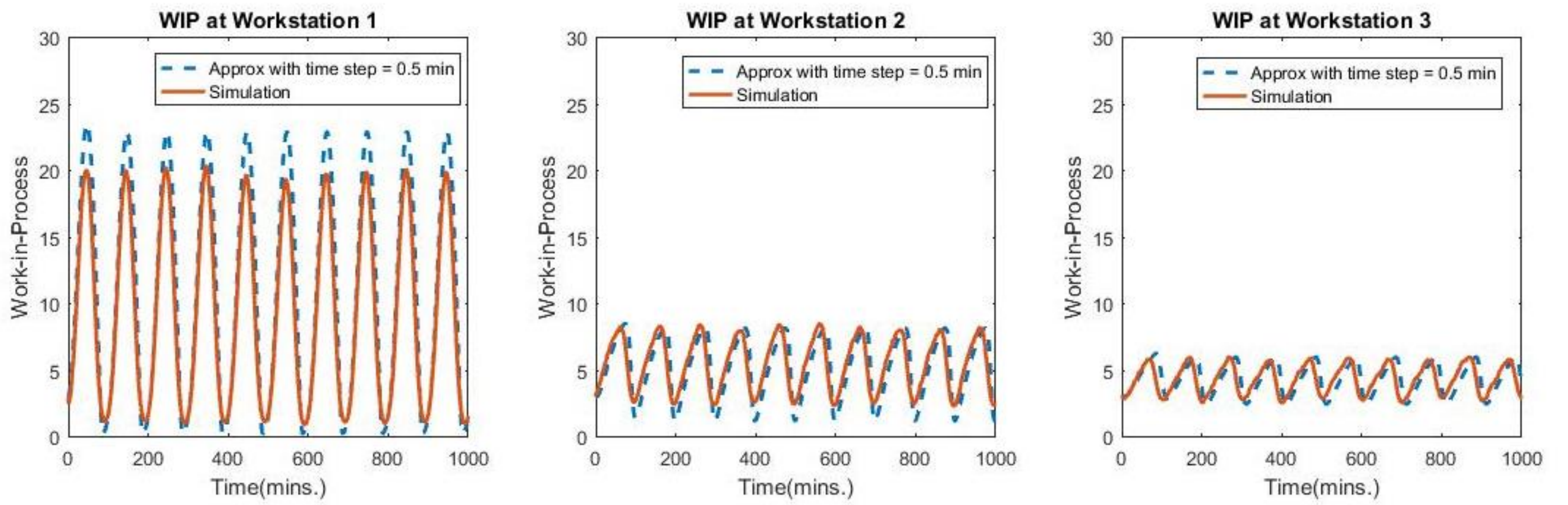
ONBTM Results: Model 1a Work - in - Process



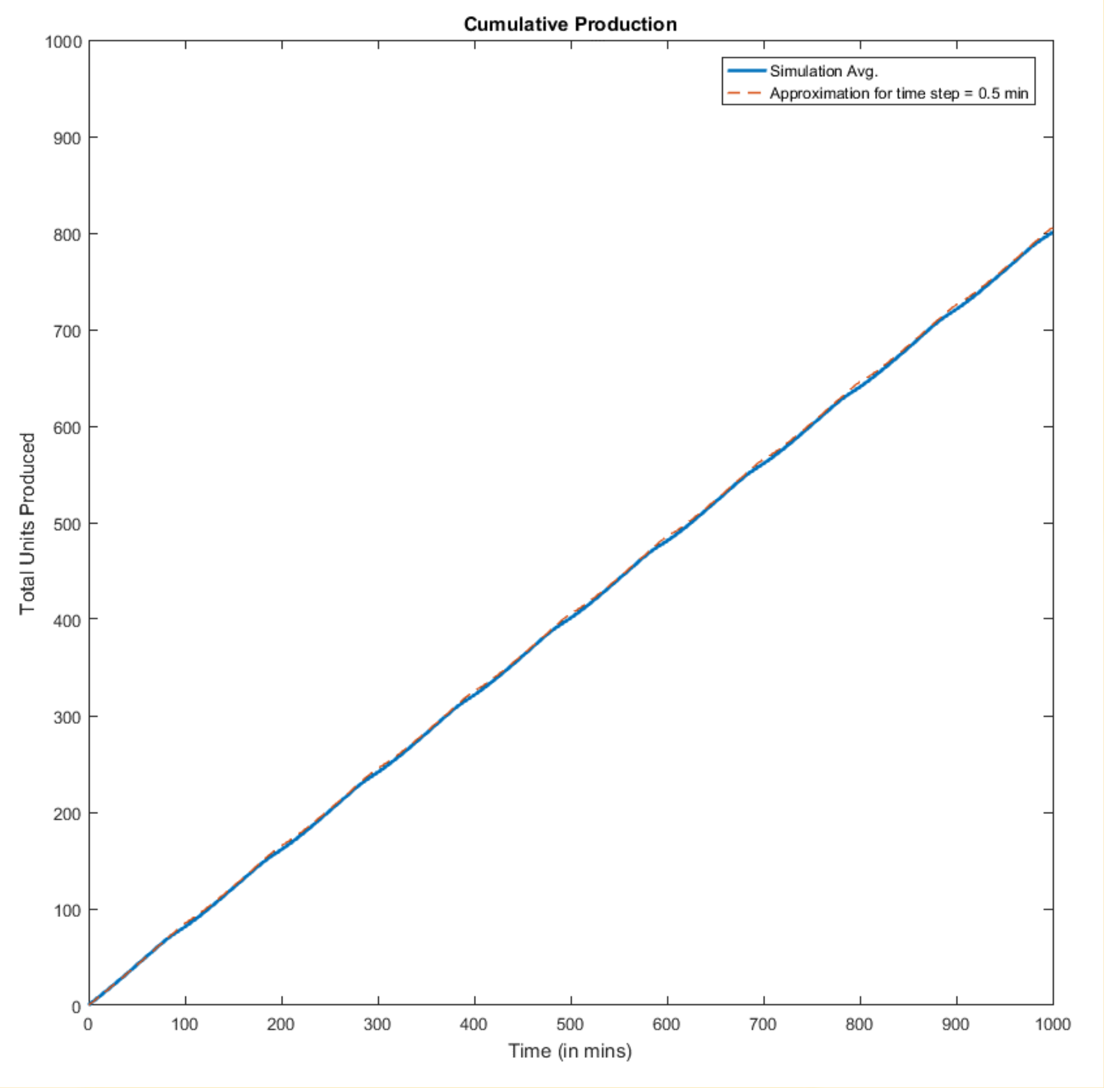
Model 1a: Cumulative production



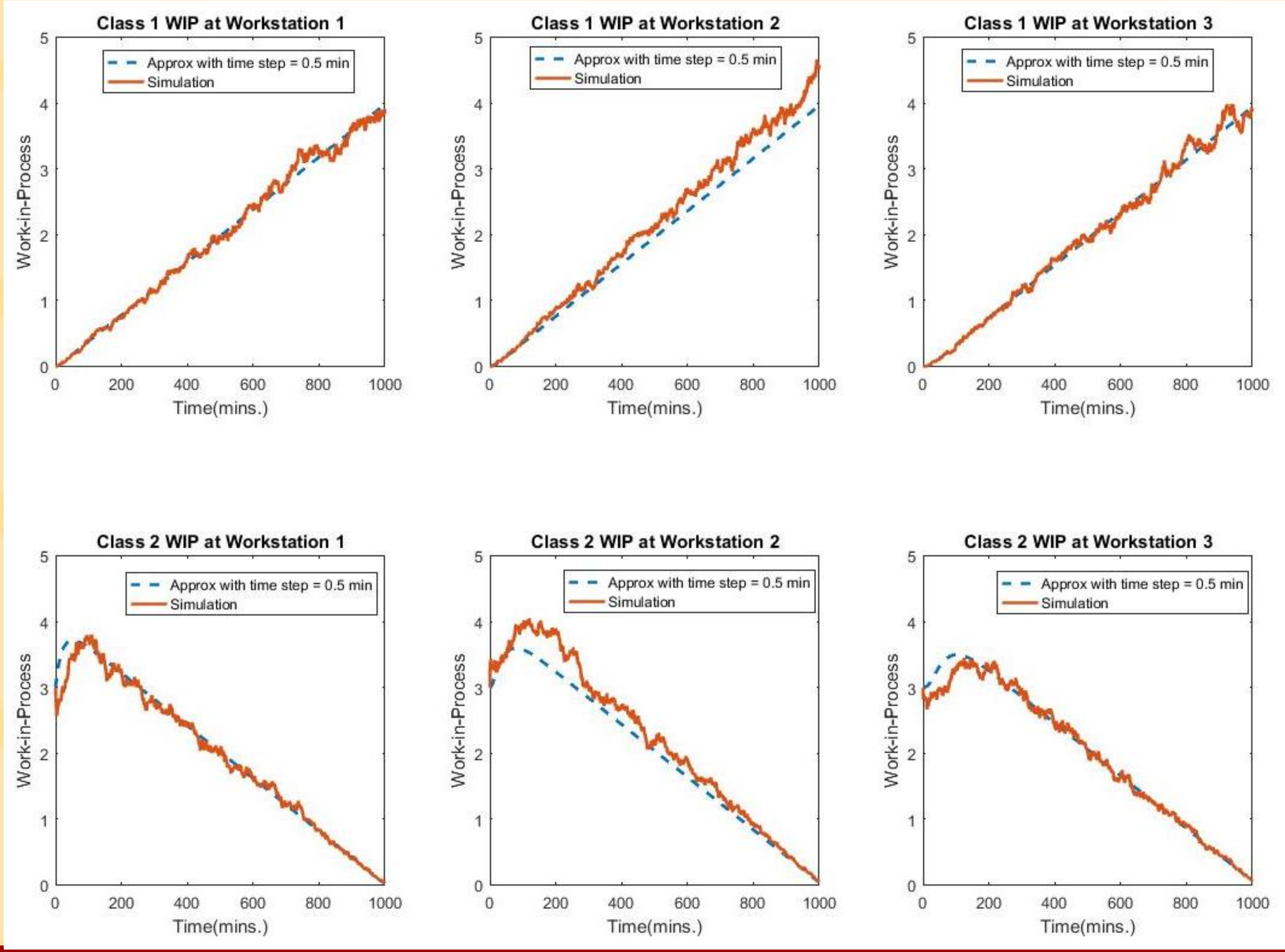
Model 1b :Work - in - Process



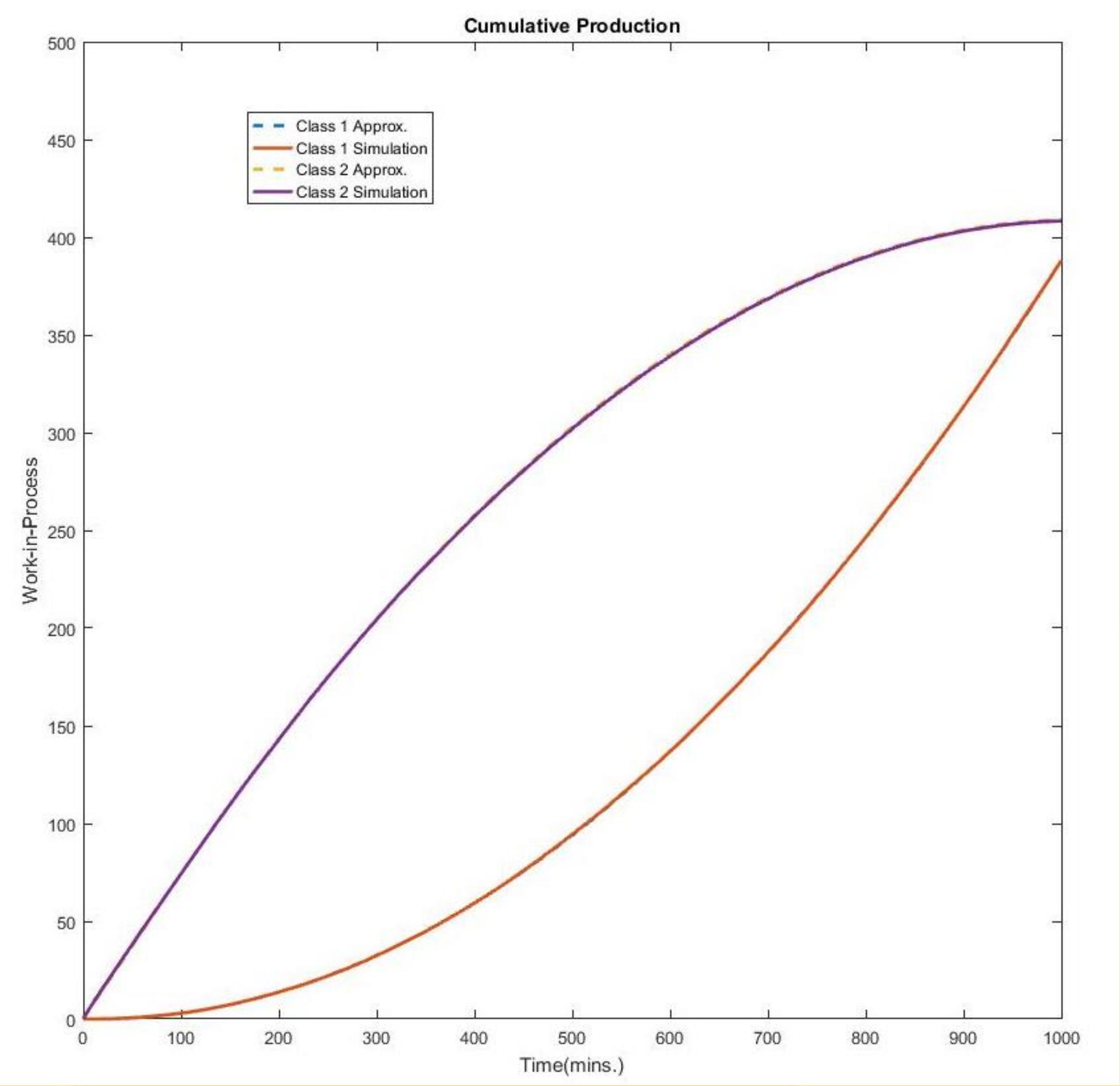
Model 1b: Cumulative production



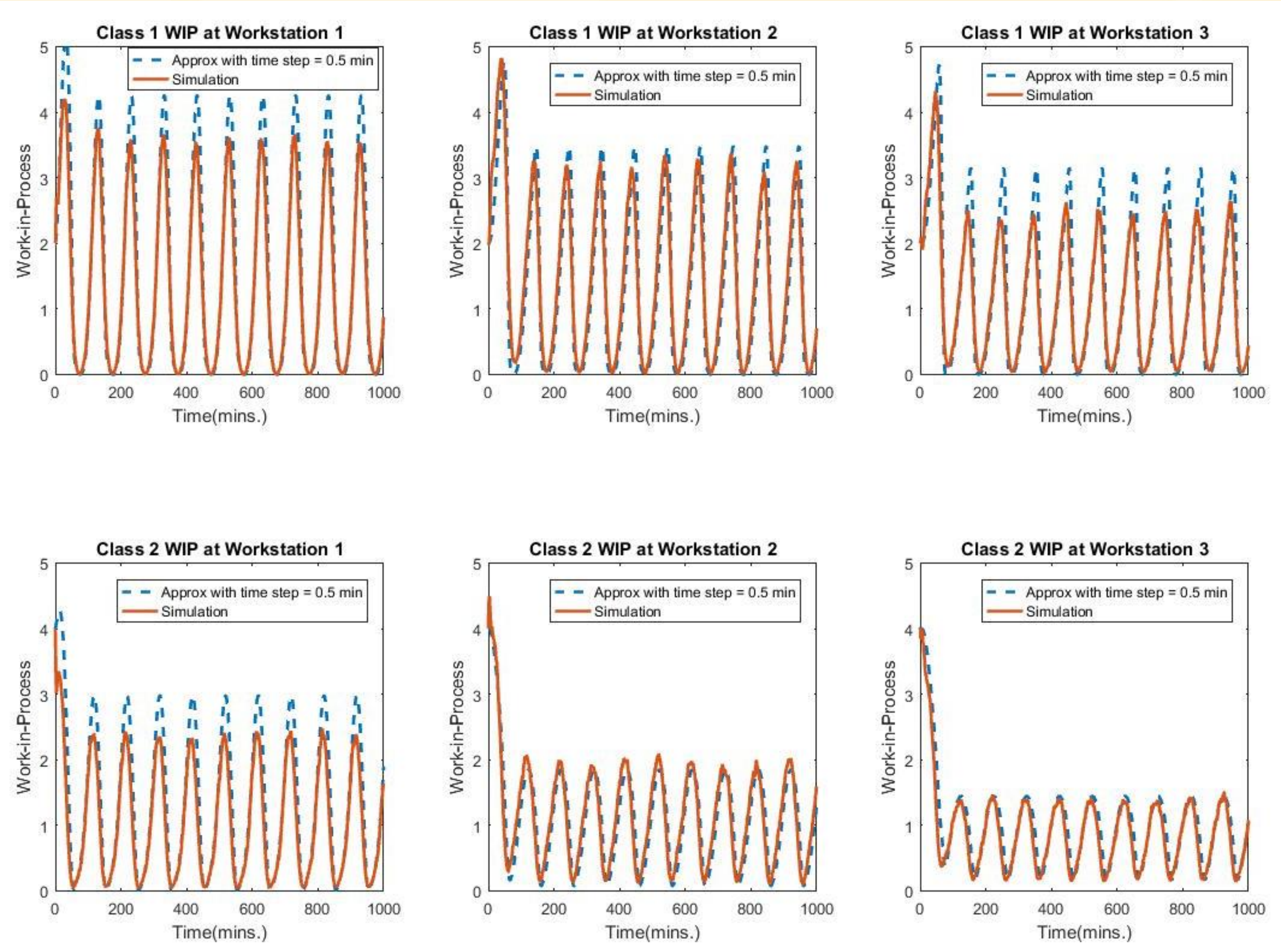
Model 2: Work - in - Process



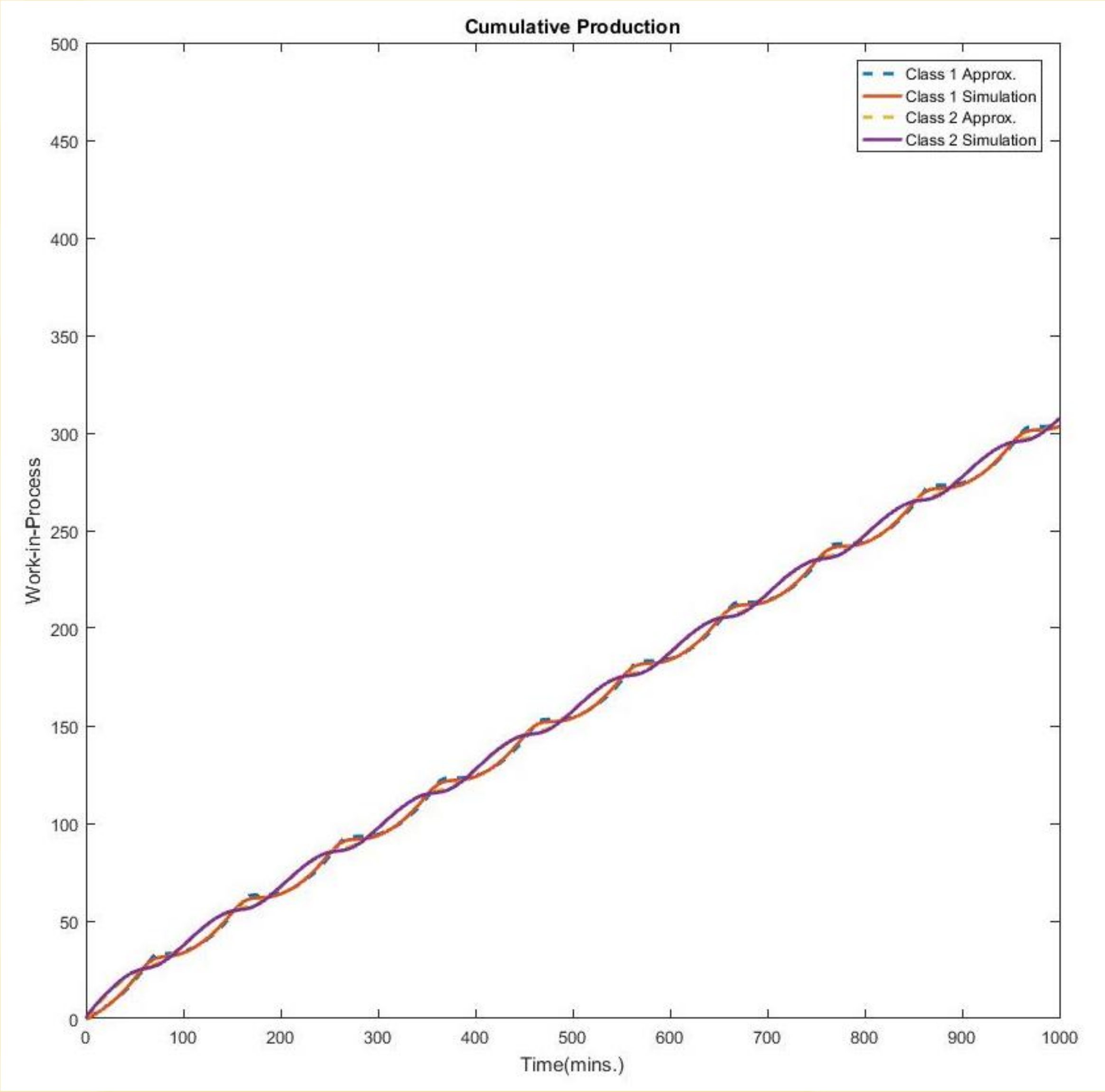
Model 2: Cumulative production



Model 3: Work - in - Process



Model 3: Cumulative production



Experiment: "Large" Jobshop Instance

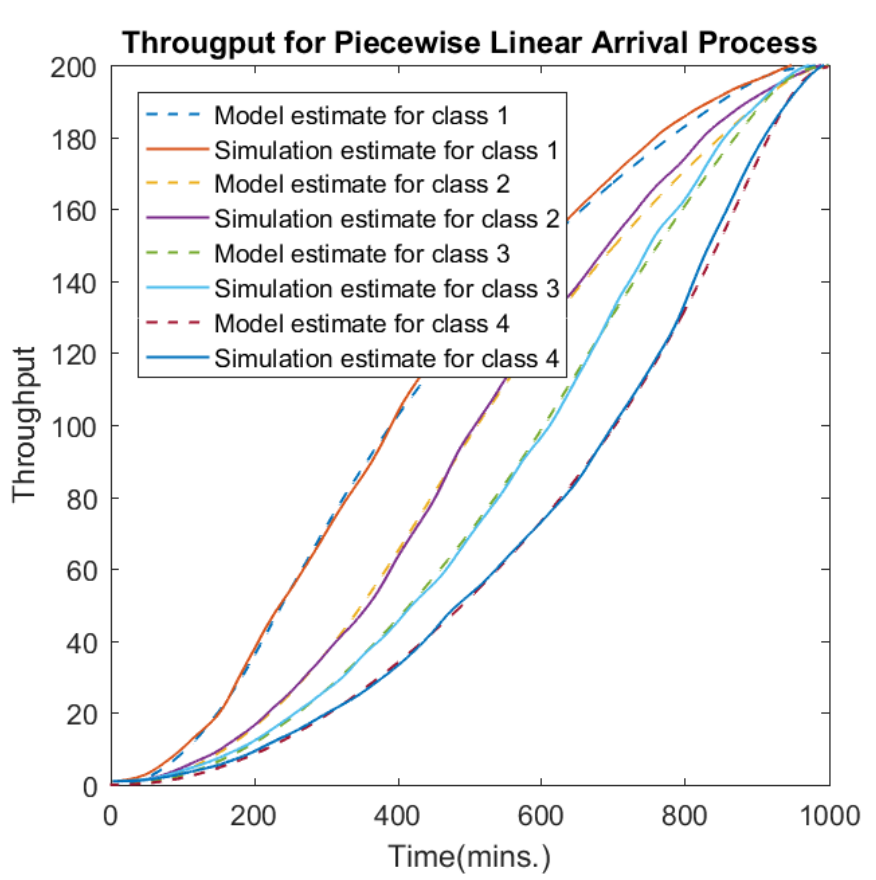
- Jobshop instance with four classes (1,..,4) and sixteen workstations (1,..16) (bottleneck: workstation 4)

Product	Routing
1	$4 \to 12 \to 2 \to 5 \to 13 \to 8$
2	$2 \to 3 \to 1 \to 4 \to 6 \to 7 \to 5 \to 9 \to 10 \to 11 \to 12 \to 13 \to 14 \to 15 \to 16$
3	$1 \to 8 \to 10 \to 9 \to 11 \to 13 \to 14 \to 4 \to 16$
4	$4 \to 5 \to 3 \to 8 \to 10 \to 6 \to 9 \to 12 \to 1 \to 15 \to 16$

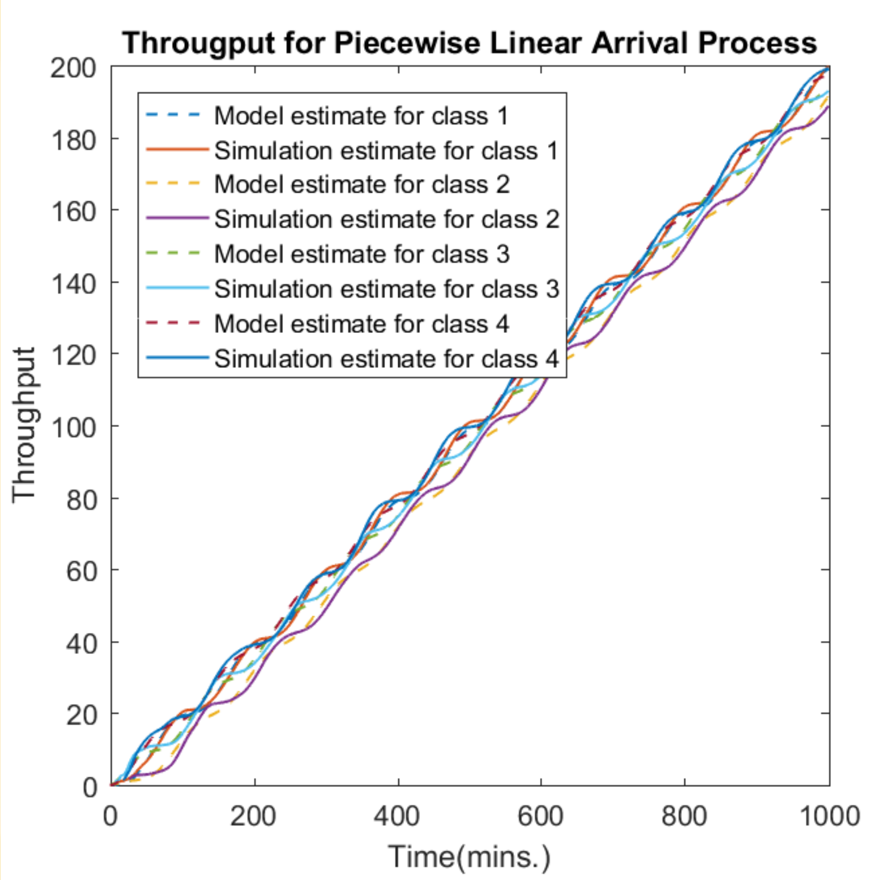
Table 1: Jobshop Routing

- Basic Model and ONBTM estimates compared against thousand simulation replications
- Two non-homogenous arrival patterns were investigated
  - Peak shifted triangular pattern with peak offset = 200 mins. (Pattern 2)
  - Phase shifted sinusoidal with $2\pi = 100\ mins$. (Pattern 1)

Basic Model: Cumulative Production

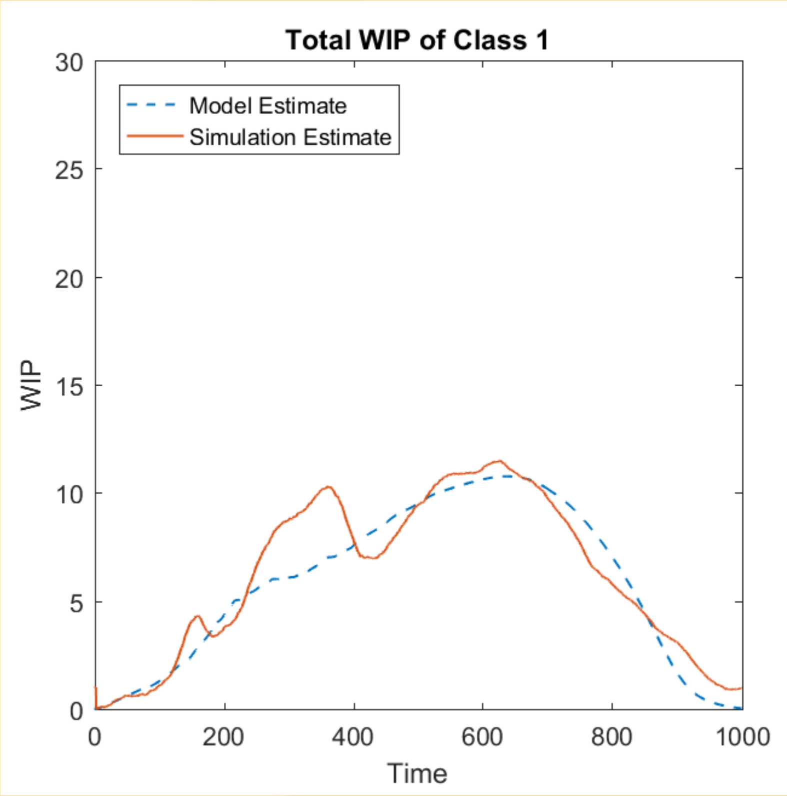


Pattern 1

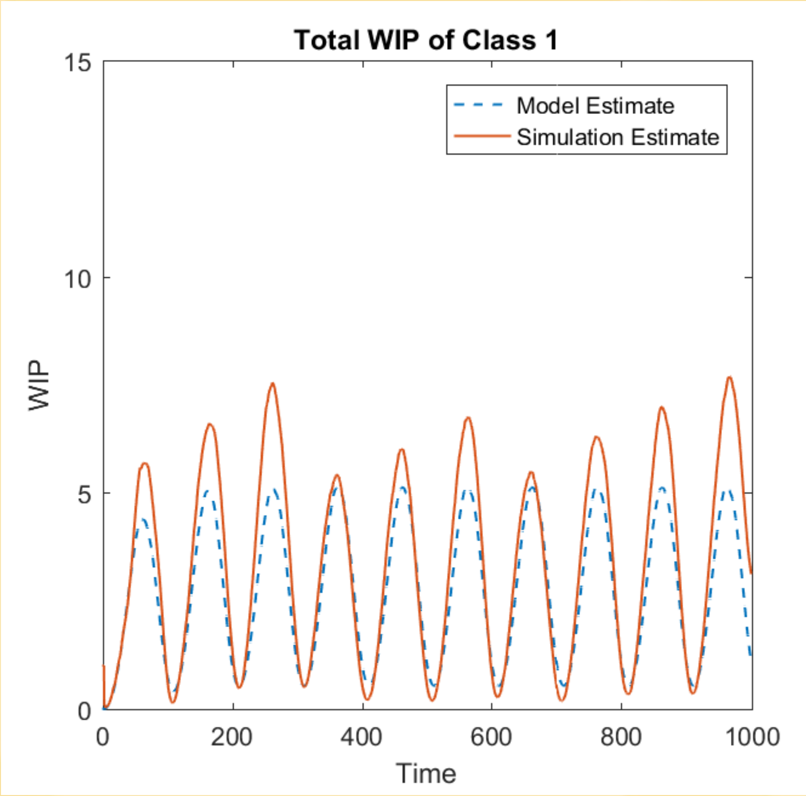


Pattern 2

Basic Model: Class 1 WIP

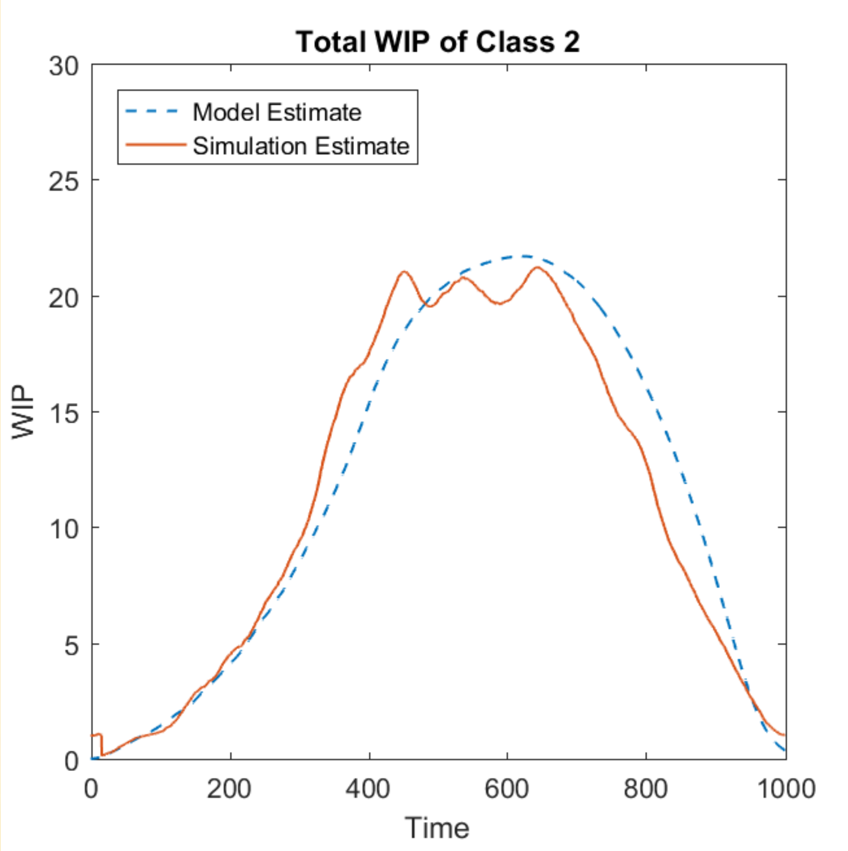


Pattern 1

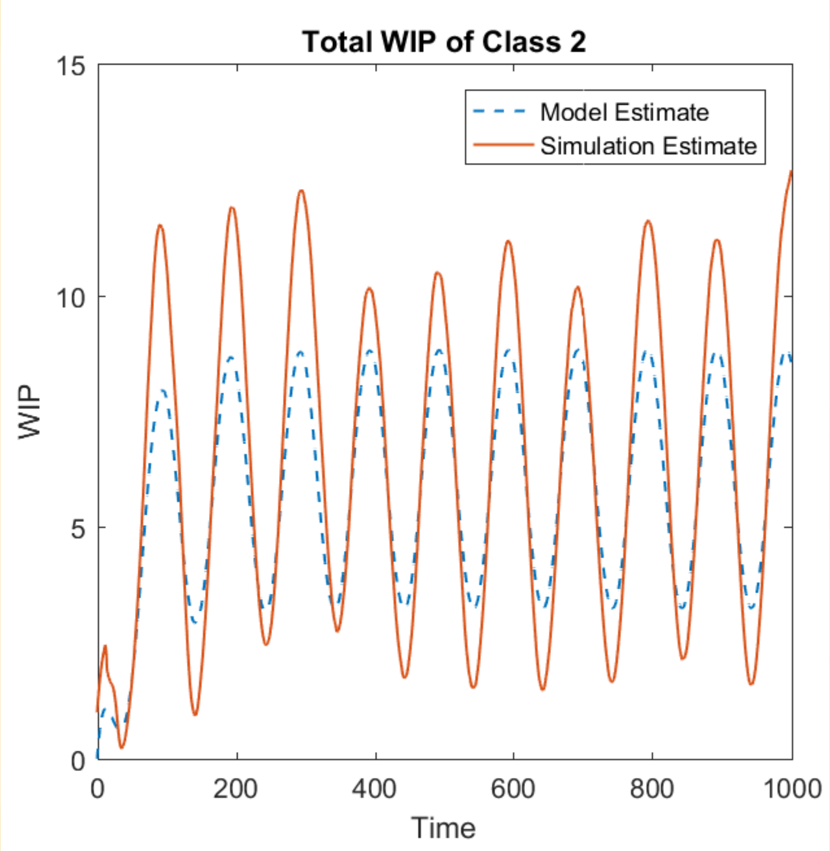


Pattern 2

Basic Model: Class 2 WIP

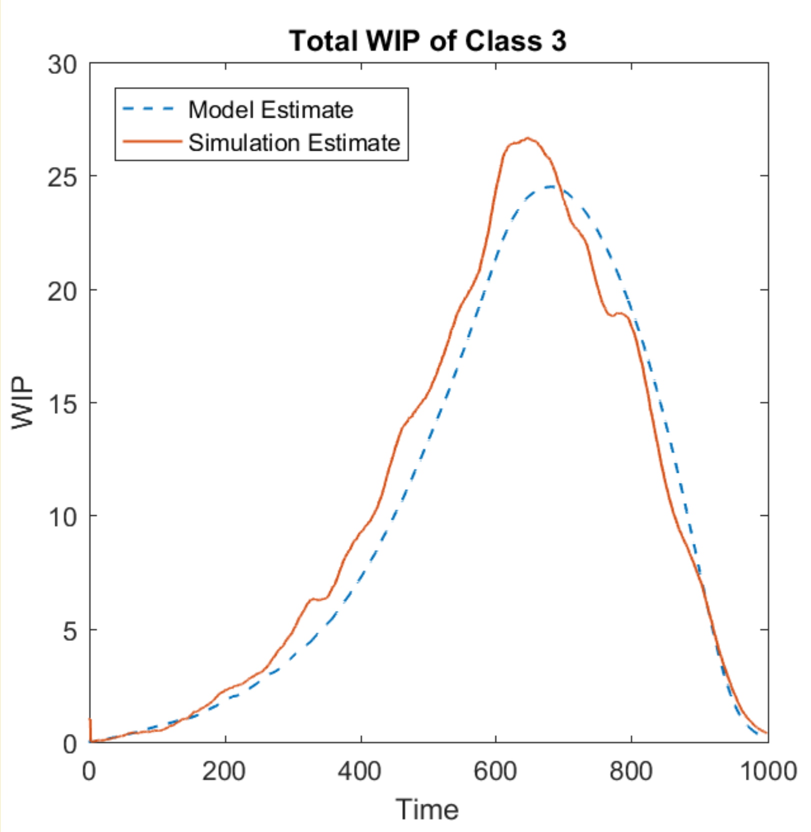


Pattern 1

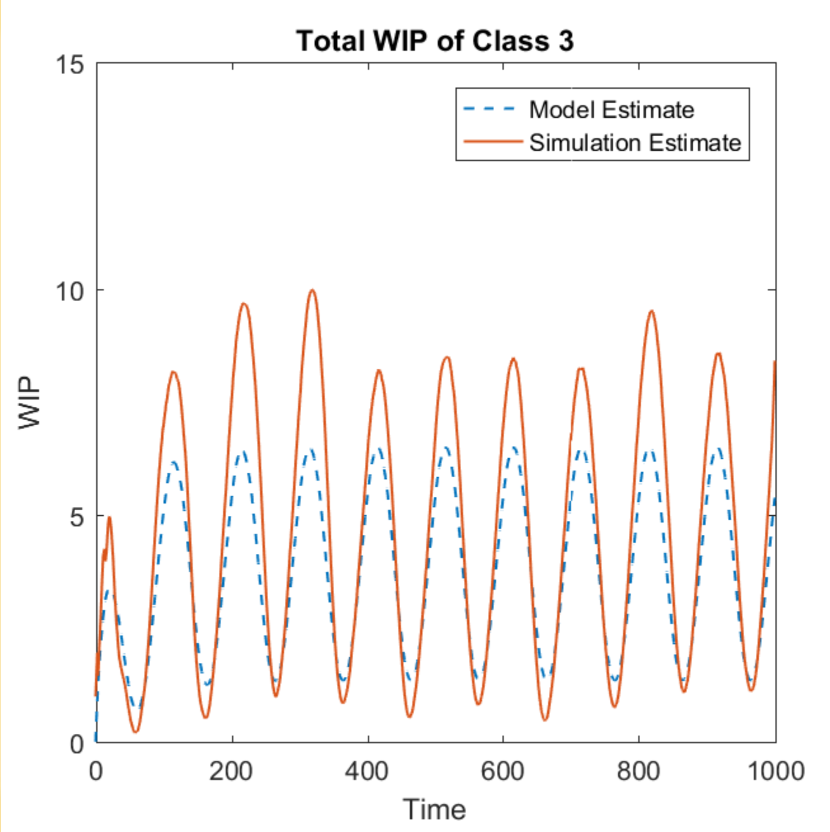


Pattern 2

Basic Model: Class 3 WIP

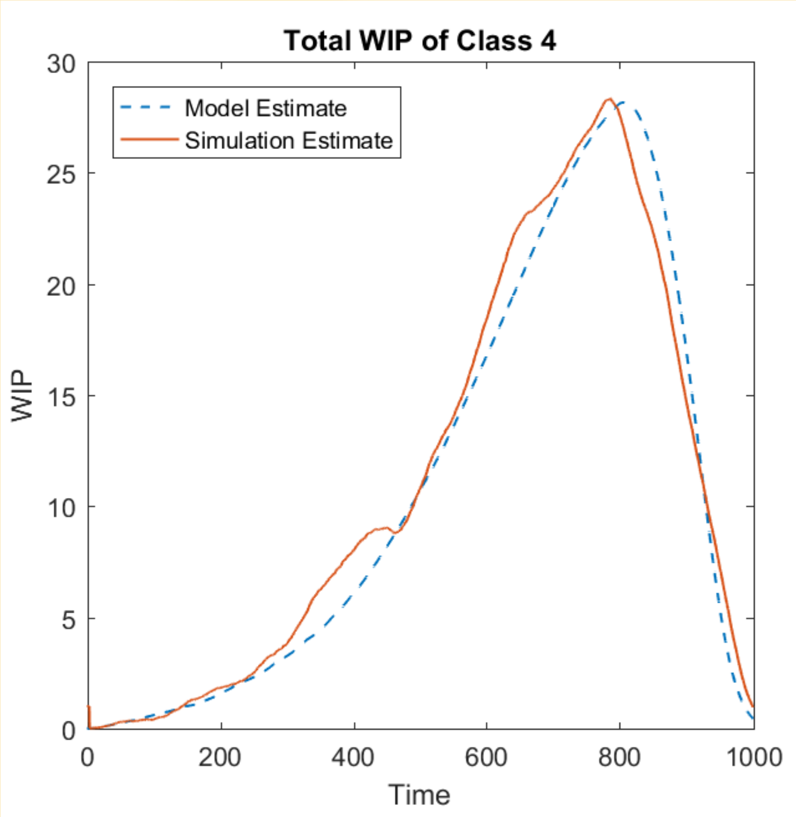


Pattern 1

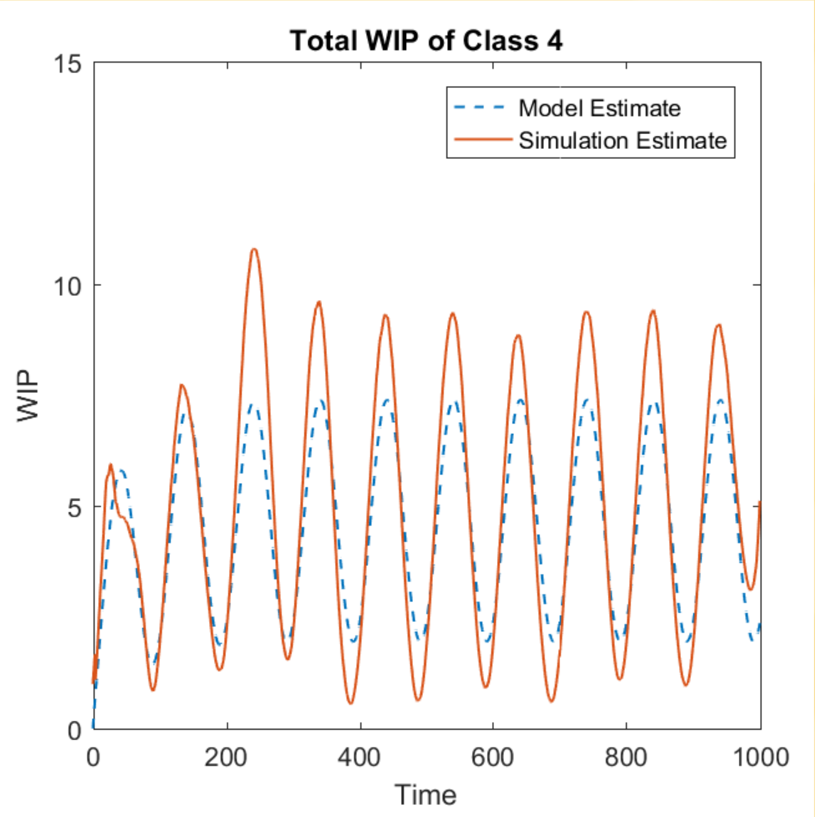


Pattern 2

Basic Model: Class 4 WIP

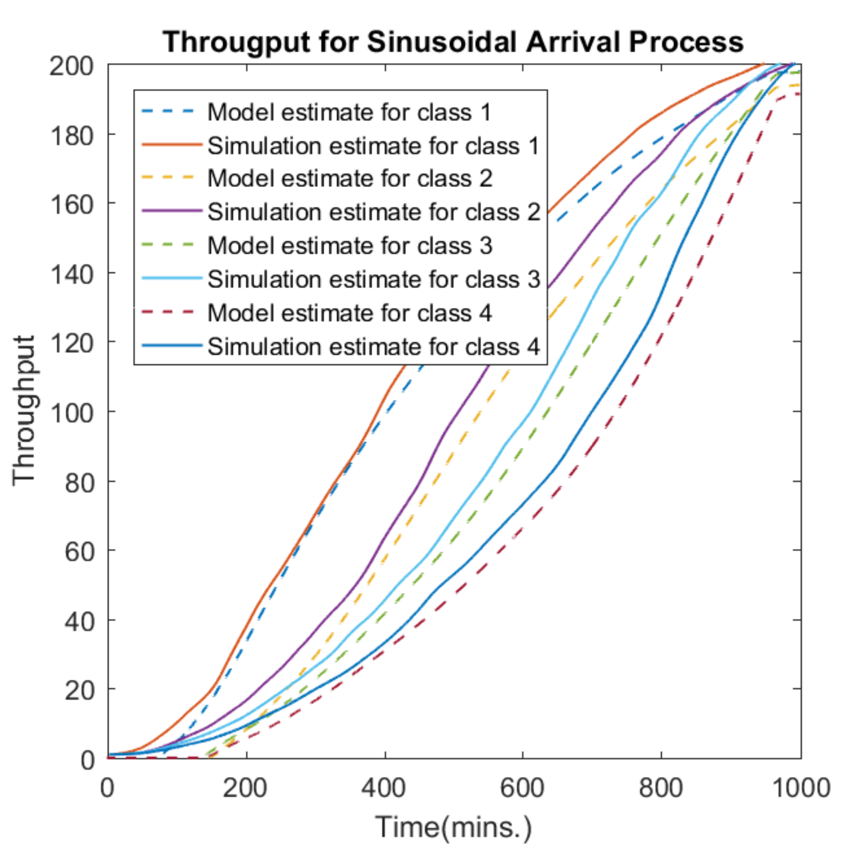


Pattern 1

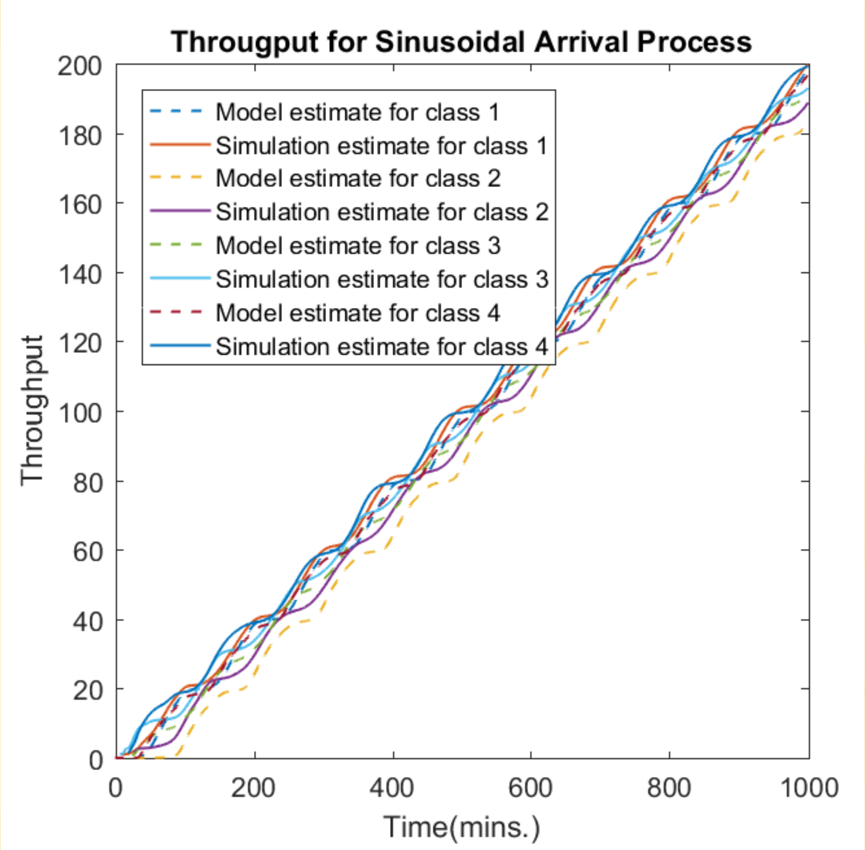


Pattern 2

ONBTM: Cumulative Production

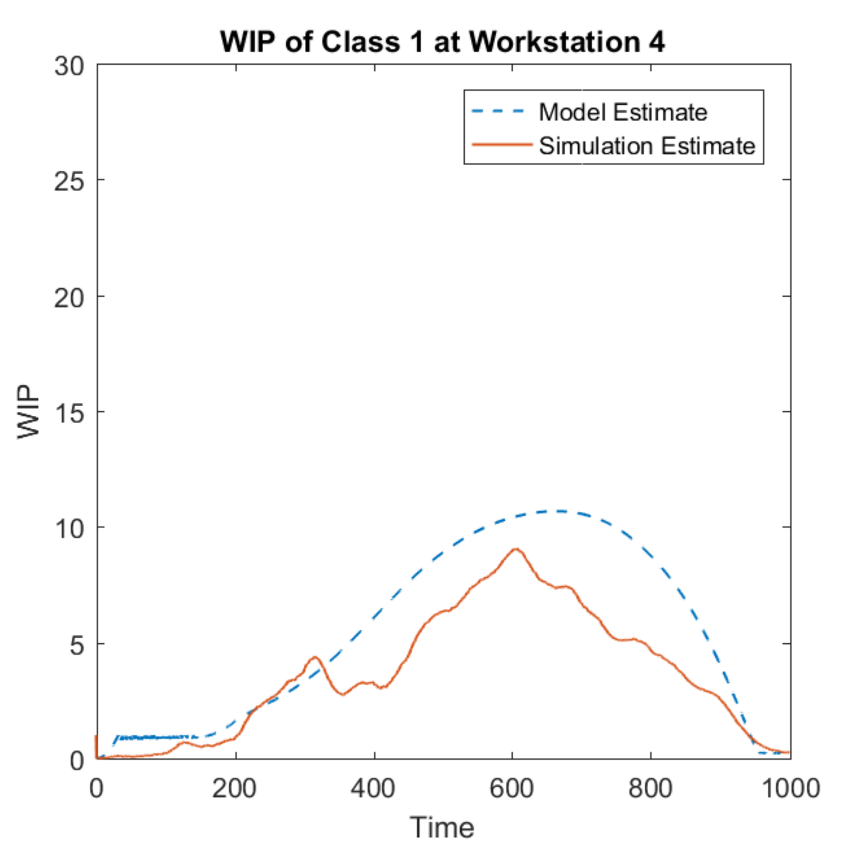


Pattern 1

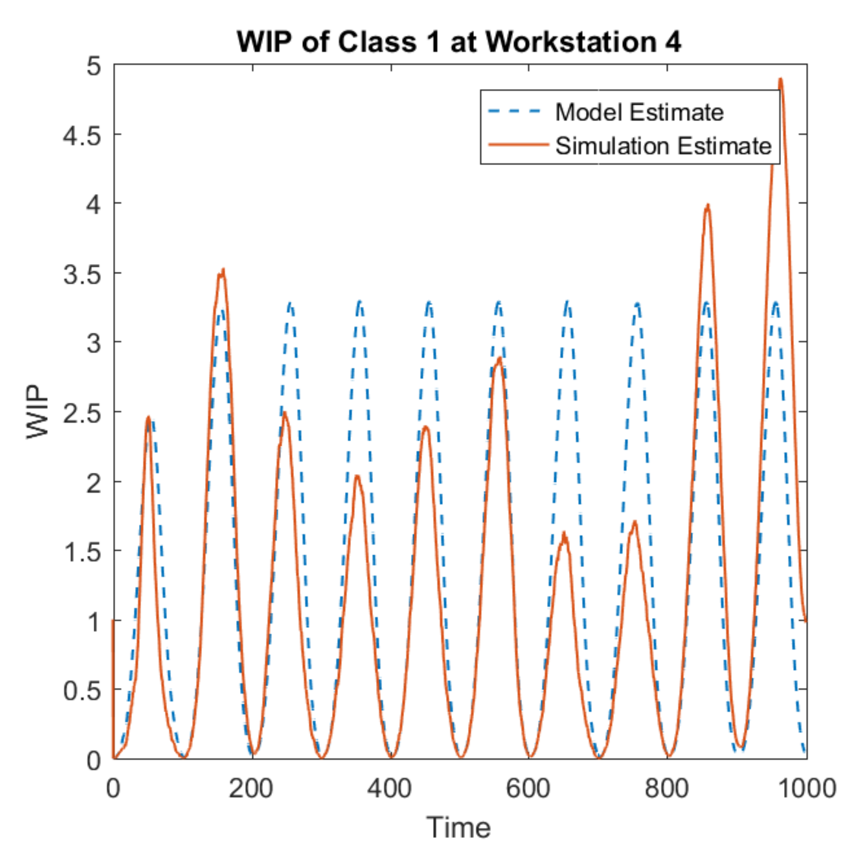


Pattern 2

ONBTM: Class 1 WIP at Workstation 4

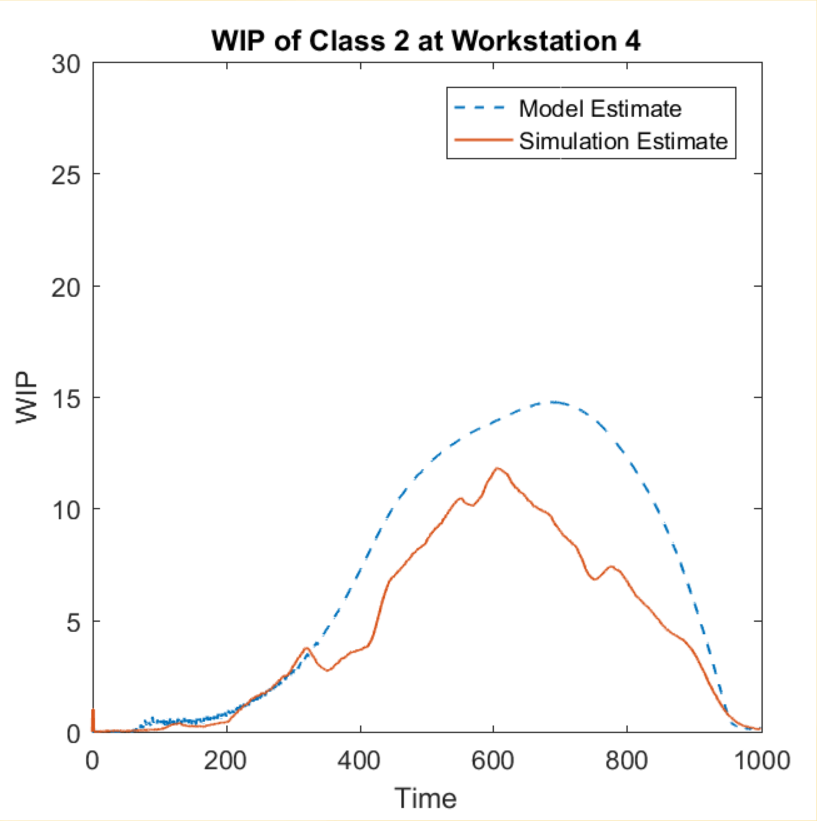


Pattern 1

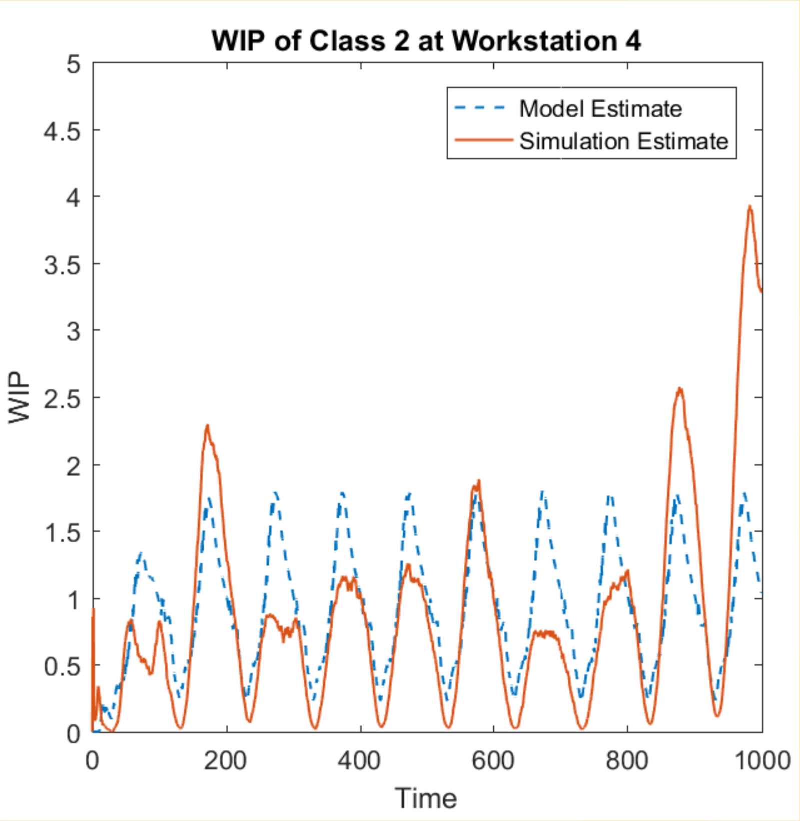


Pattern 2

ONBTM: Class 2 WIP at Workstation 4

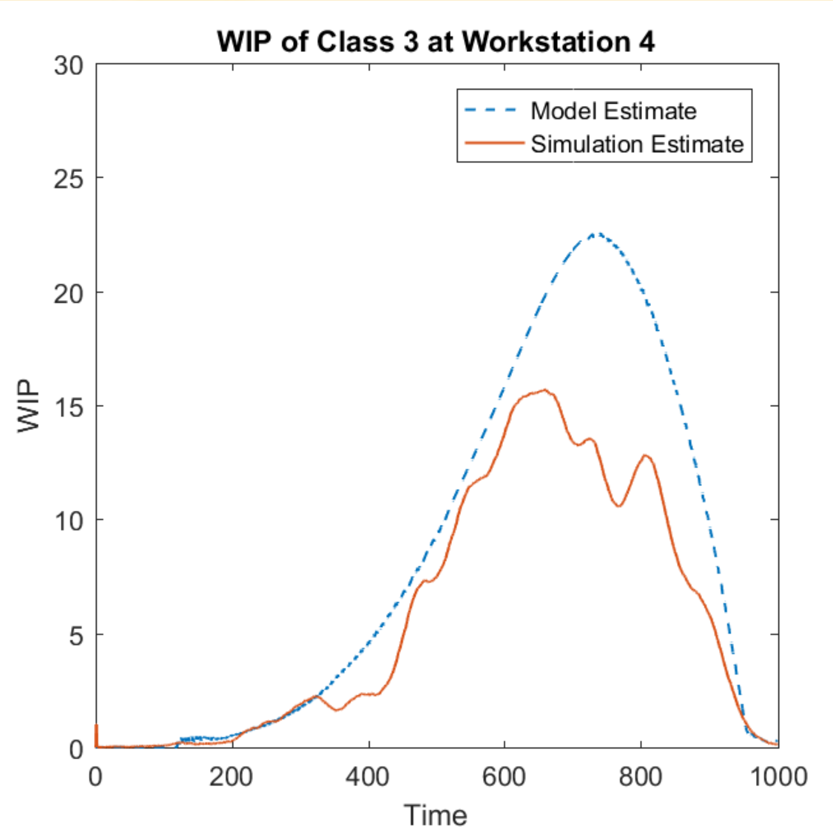


Pattern 1

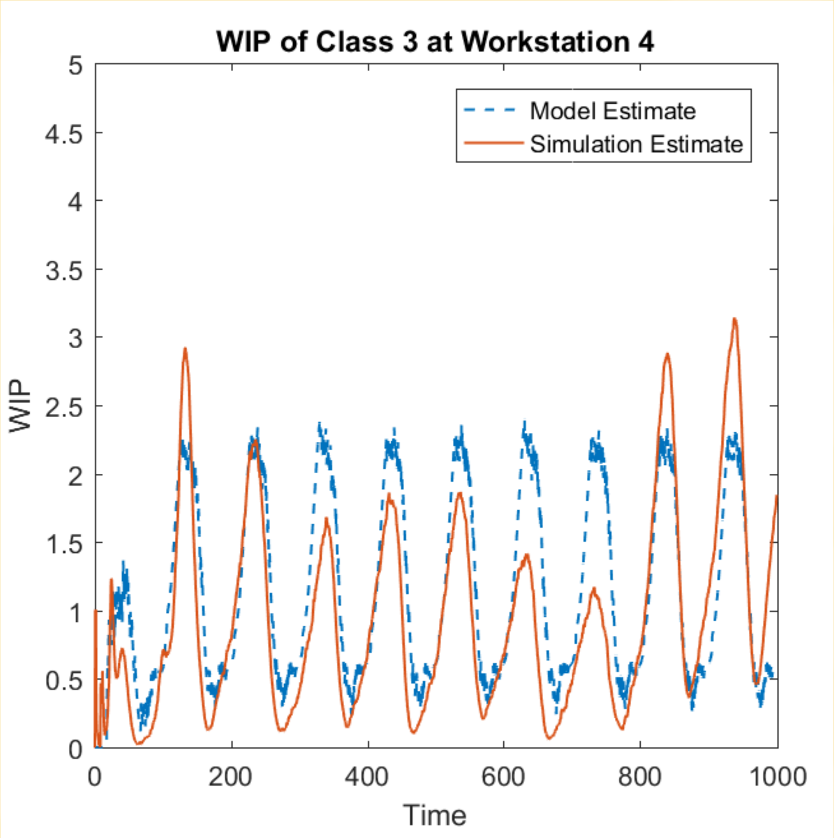


Pattern 2

ONBTM: Class 3 WIP at Workstation 4

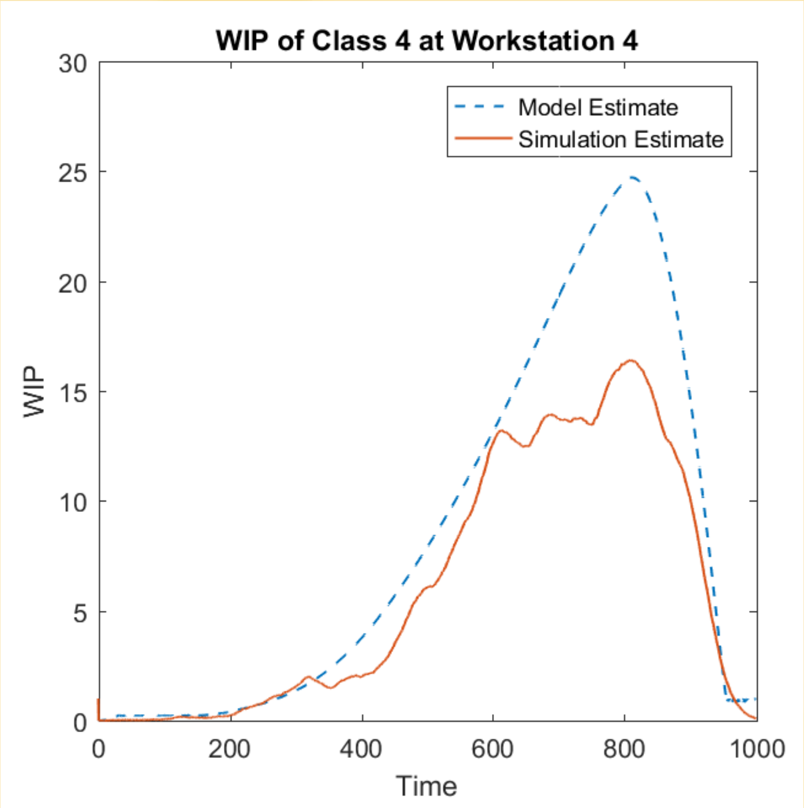


Pattern 1

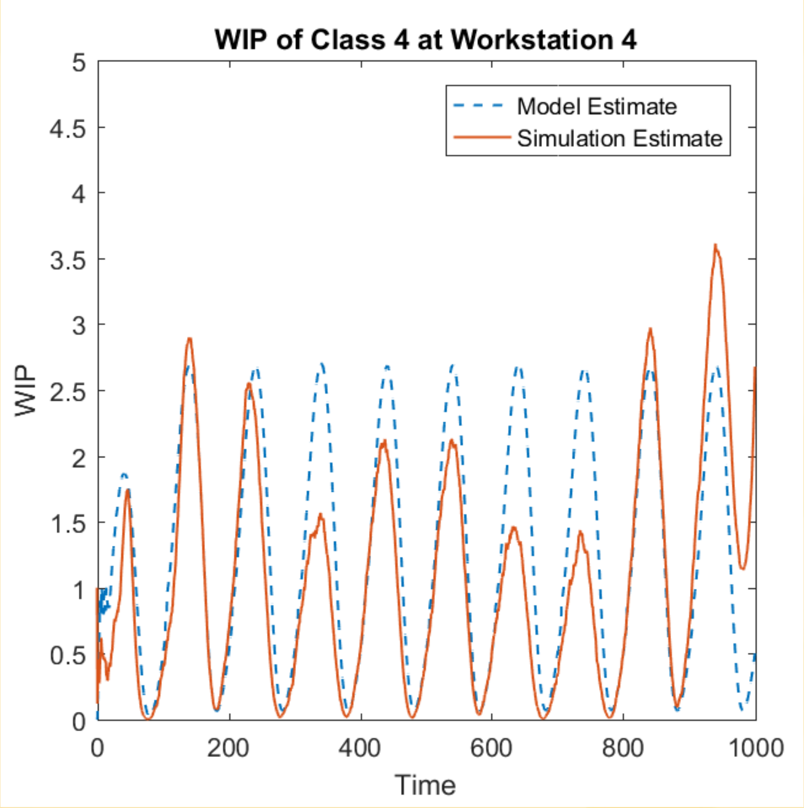


Pattern 2

ONBTM: Class 4 WIP at Workstation 4

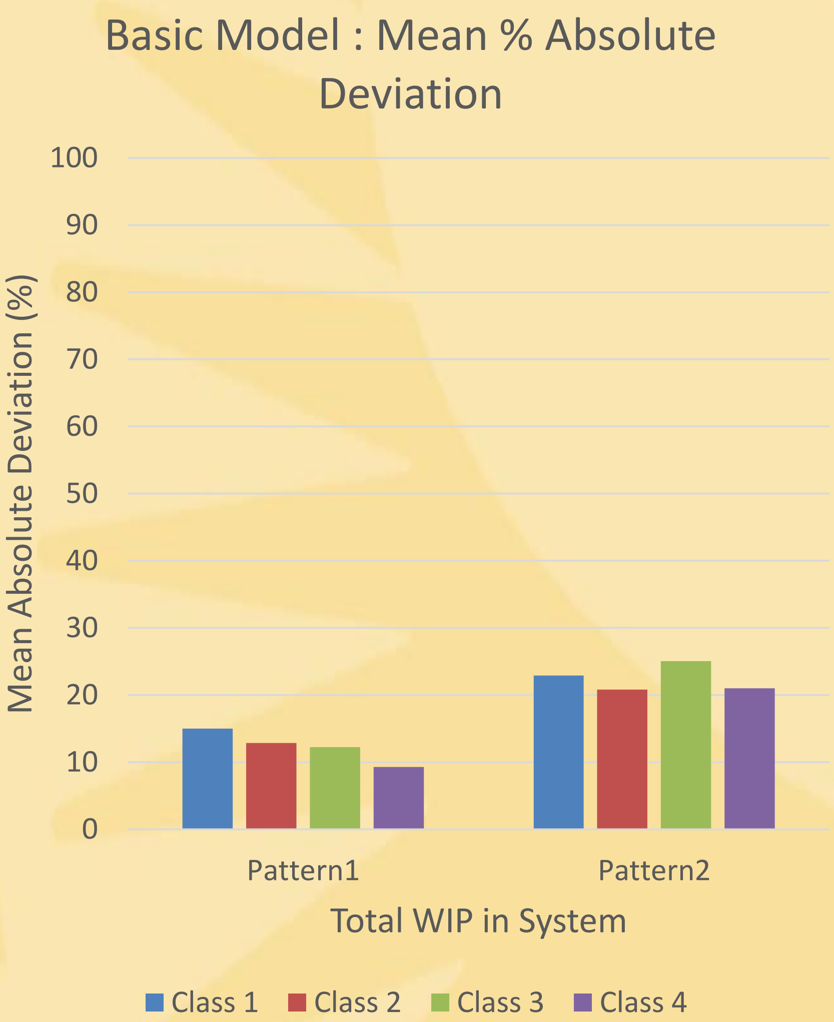


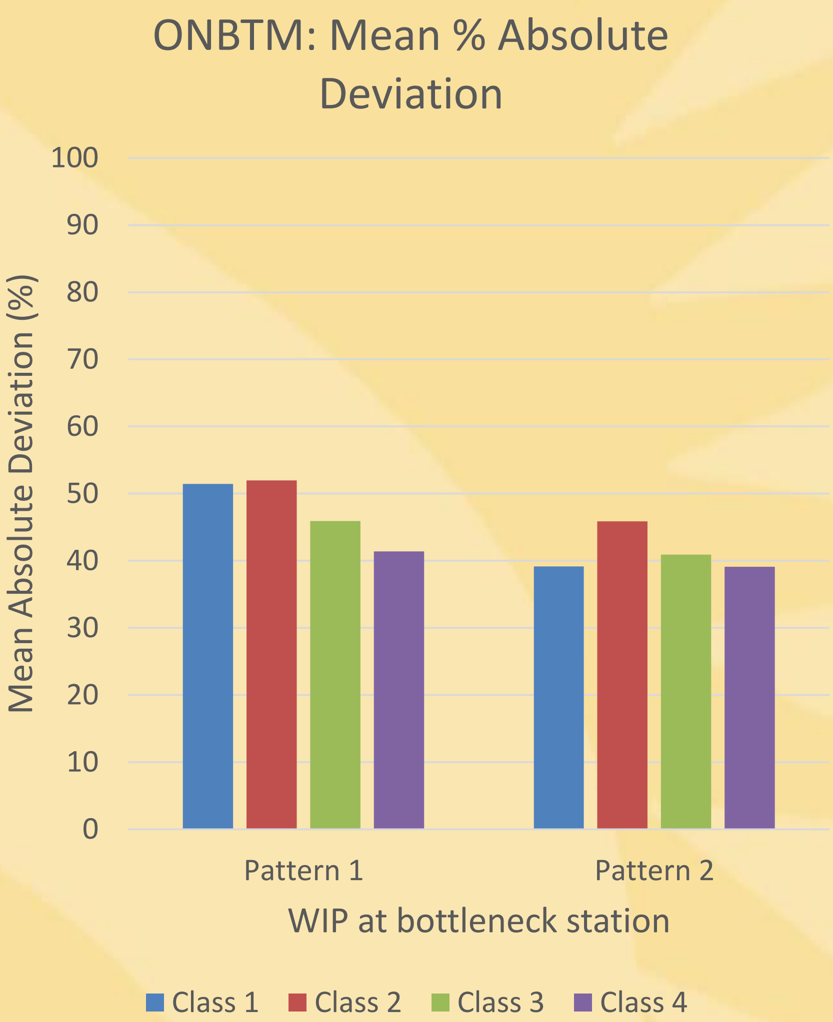
Pattern 1



Pattern 2

Large JobShop: Absolute Deviation





Extensions

- This work will be extended to study:
  - other non-stationary demand patterns.
  - part priorities
  - product mix changes.
  - buffer sizing for workstations.
  - effect of different scheduling disciplines.
  - capacity/workstation availability conditions.
  - More memory efficient (single stage) WS level closed approximations.

Conclusions

- All hope is not lost, we can tractably do rough cut estimation
- More work is needed

Thanks for Listening

Questions?