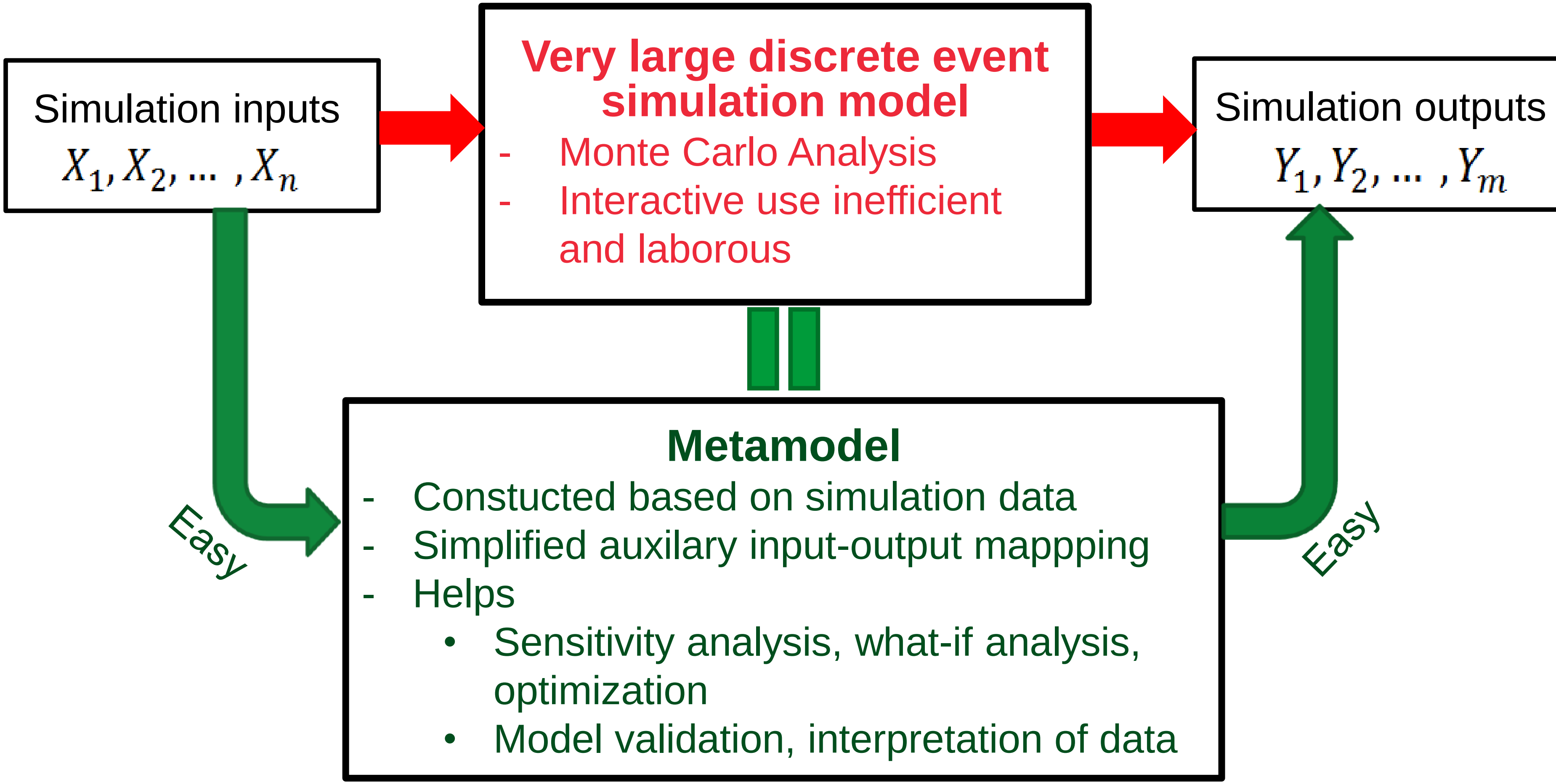
# Simulation Metamodeling using Bayesian Networks

Jirka Poropudas, Jouni Pousi, and Kai Virtanen

Aalto University
School of Science
Systems Analysis Laboratory
sal.aalto.fi
kai.virtanen@aalto.fi

# Simulation metamodeling

## Complex and stochastic dynamic system



# Increasing complexity of models increases the need for metamodeling

- Several existing approaches
  - Regression models, kriging models, splines, games, neural networks

  => Mappings from inputs to **the expected values** of outputs

  => LOSS OF INFORMATION!

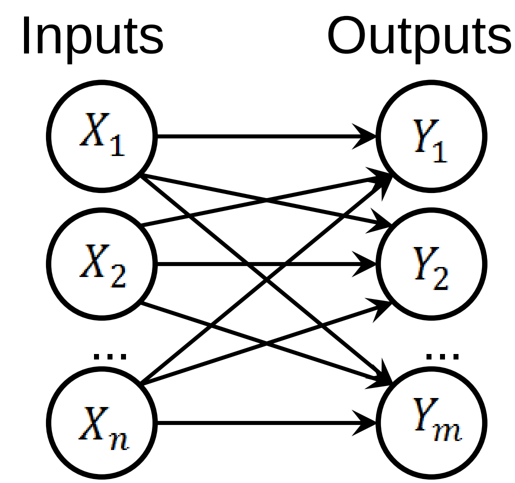
**New features allowed by Bayesian networks**

Joint probability distribution of inputs and outputs

=>

"Mappings between input and output distributions"

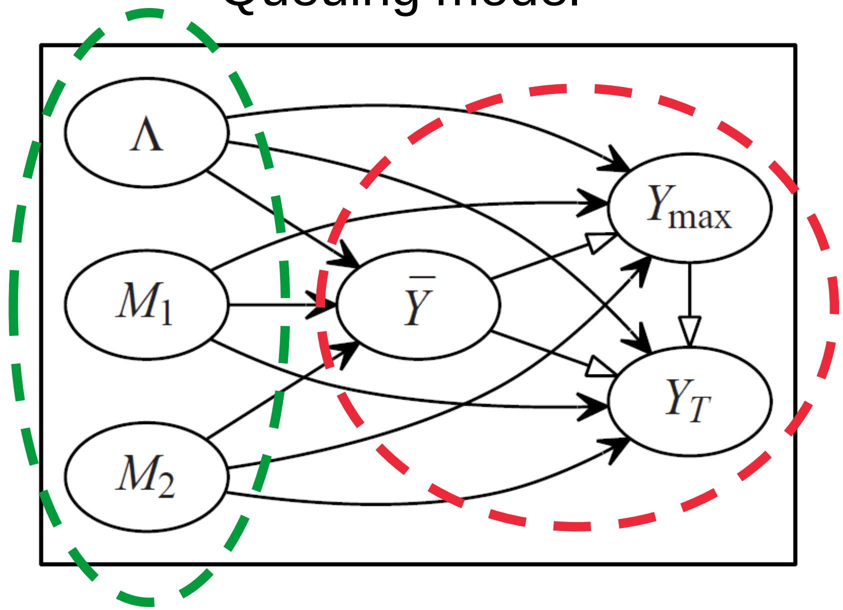
- Probabilistic dependencies between variables
- Efficient calculation of conditional probability distributions => Versatile what-if analyses
- Random inputs reflect input uncertainty



# Bayesian networks (BNs) as simulation metamodels

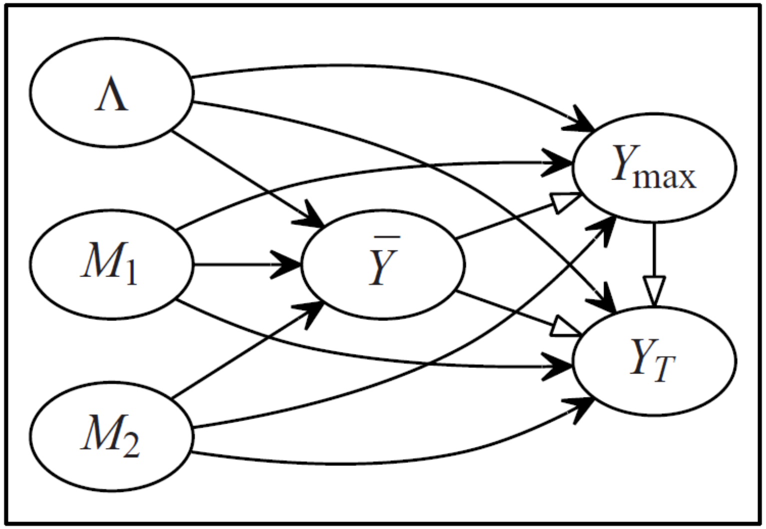
- Joint probability distribution of random variables
- Simulation inputs and outputs
  - Nodes
  - Discrete random variables
- Dependencies
  - Arcs
  - Conditional probabilities
- Available algorithms
  - Construction of BNs
  - Calculation of conditional probability distributions

Queuing model



# Construction of BN Metamodels

1) Selection of variables
   - Simulation inputs and outputs
2) Collecting simulation data
   - All input combinations are simulated
3) Determination of network structure
   - Initial structure:
   - Dependencies found in data:
4) Estimation of probabilities
   - Conditional probability distributions for outputs
   - Input uncertainty
5) Validation
   - Comparison with independent simulation data

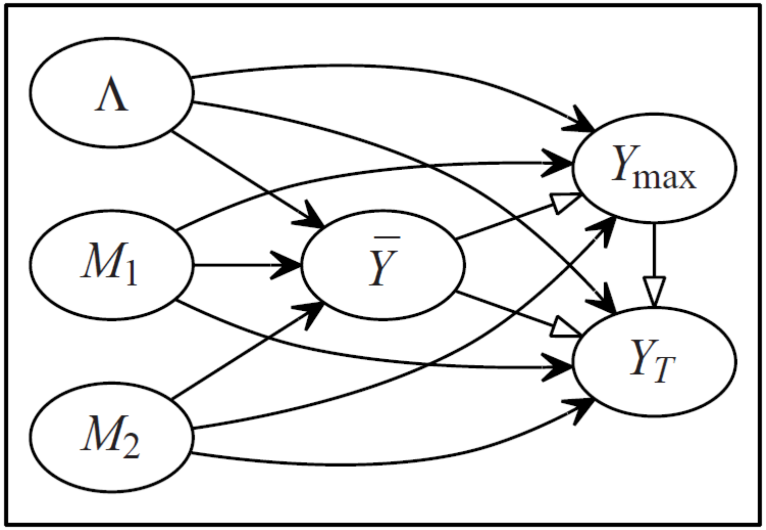


# Utilization of BN Metamodels

- What-if analysis

  Values of some variables fixed =>

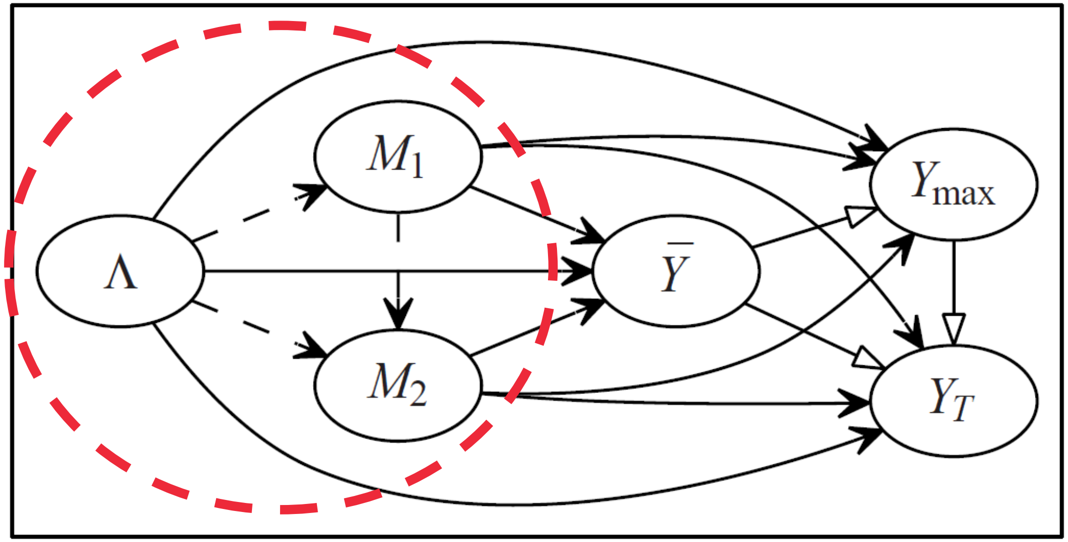
  Conditional probability distributions for other variables updated
- Applications of what-if analyses
  – Effect of input uncertainty on outputs
  – Dependence between inputs and outputs
  – Dependence between outputs
  – Inverse reasoning



# Example: Queuing model

Single queue with Poisson arrivals and
two servers with exponential service times

DEPENDENCE
BETWEEN
INPUTS

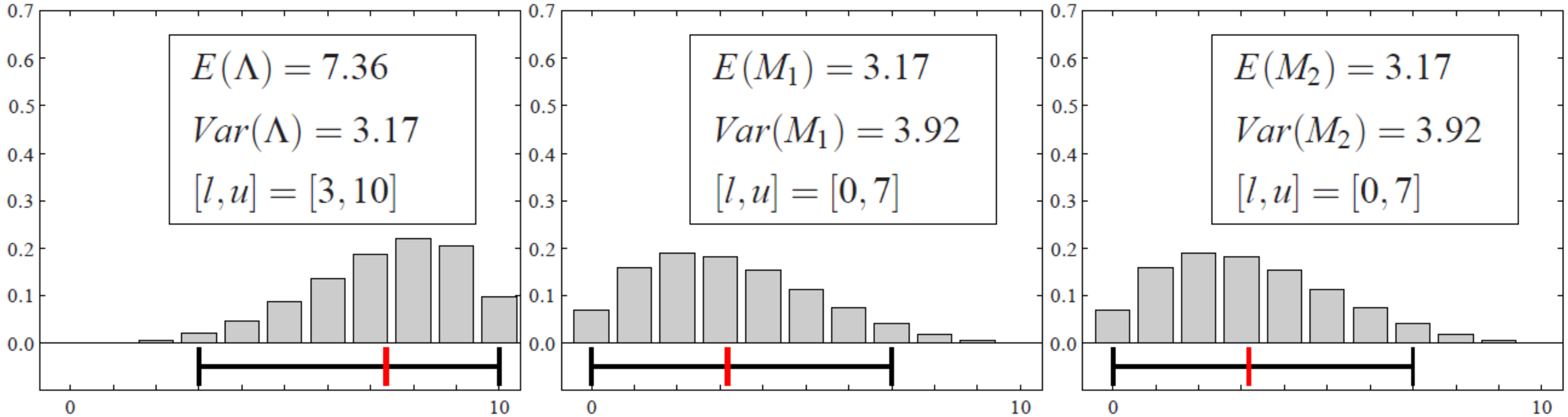



- Simulation inputs: $\Lambda, M_1, M_2$ (Arrival intensity; Service intensities)
- Simulation outputs: $\bar{Y}, Y_{max}, Y_T$ (Average and maximum number of customers; Number of customers in the end of simulation)

# Effect of Input Uncertainty...

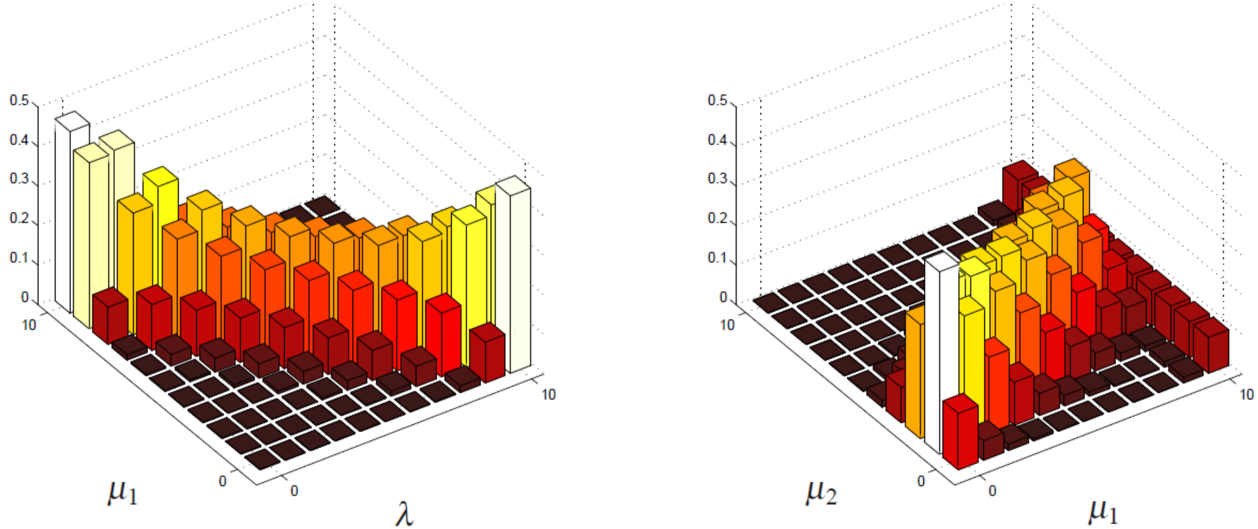
- Inputs considered as dependent random variables
  - ”Prior uncertainty” assessed by subject matter expert

MARGINAL DISTRIBUTIONS



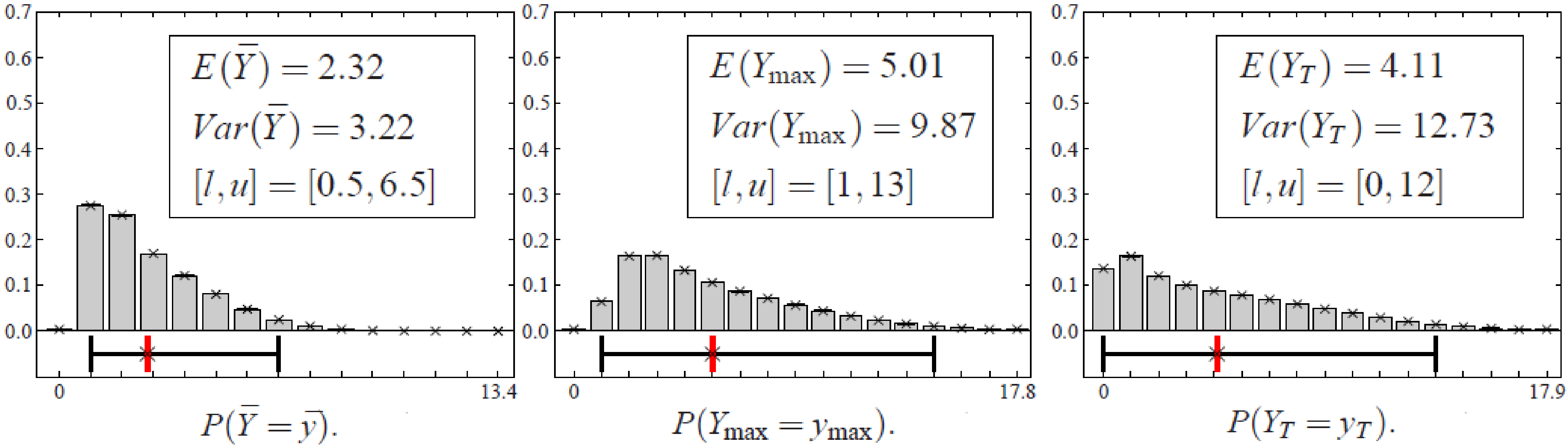

(a) $P(\Lambda = \lambda)$. (b) $P(M_1 = \mu_1)$. (c) $P(M_2 = \mu_2)$.

CONDITIONAL DISTRIBUTIONS




(a) $P(M_1 = \mu_1 | \Lambda = \lambda)$. (b) $P(M_2 = \mu_2 | \Lambda = 8, M_1 = \mu_1)$.

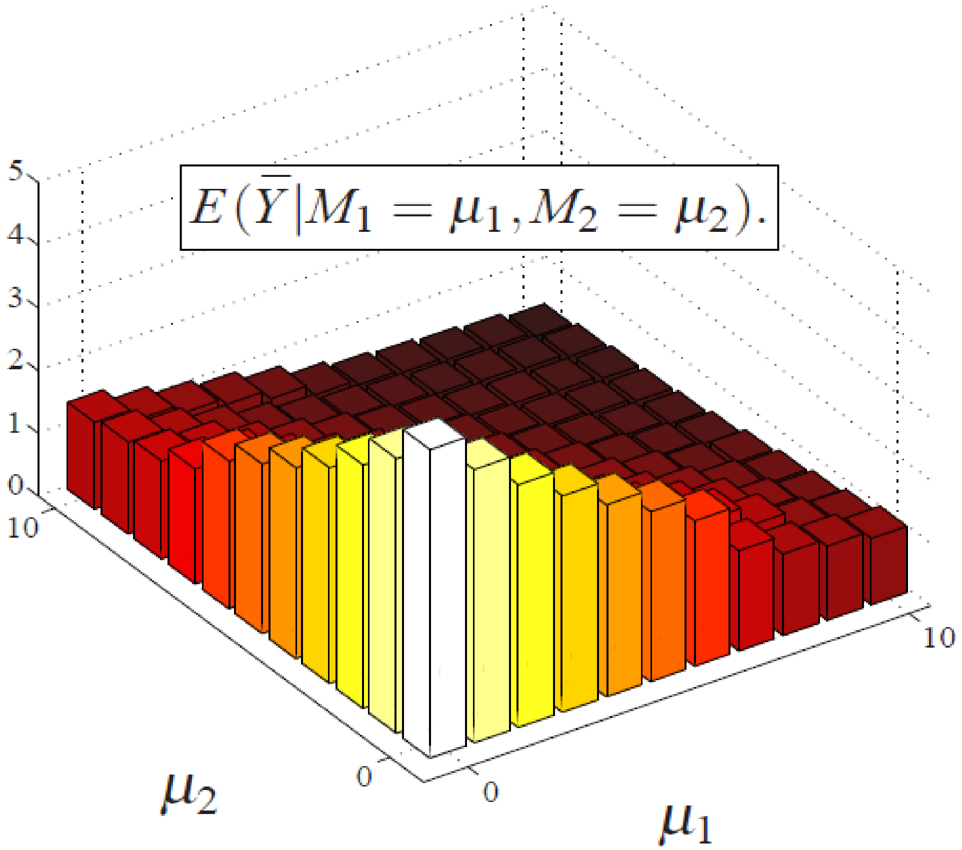
# ... on Distributions of Outputs




- Probability distributions include the effect of input uncertainty
- Enables calculation of descriptive statistics such as expected values, variances, and quantiles

# Dependence between Inputs and Outputs

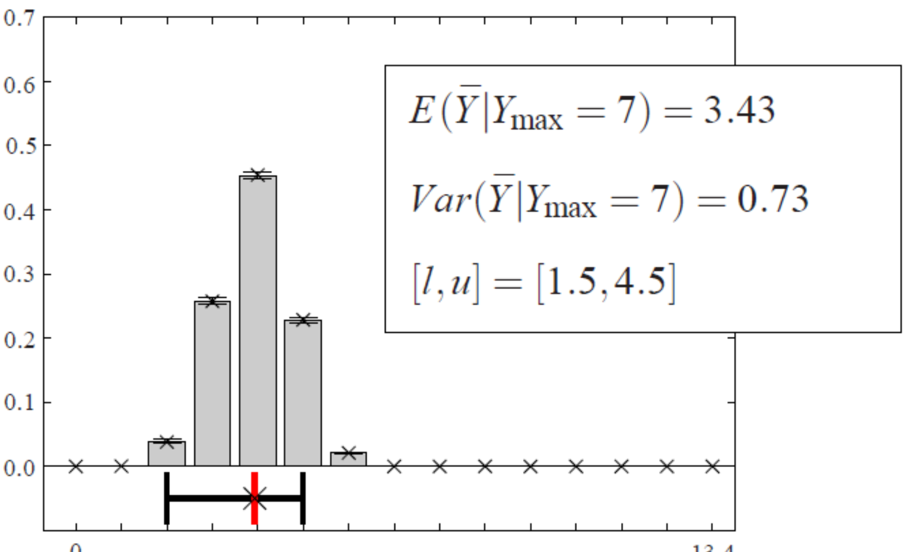
- Studied by calculating conditional probability distributions of outputs for fixed values of inputs



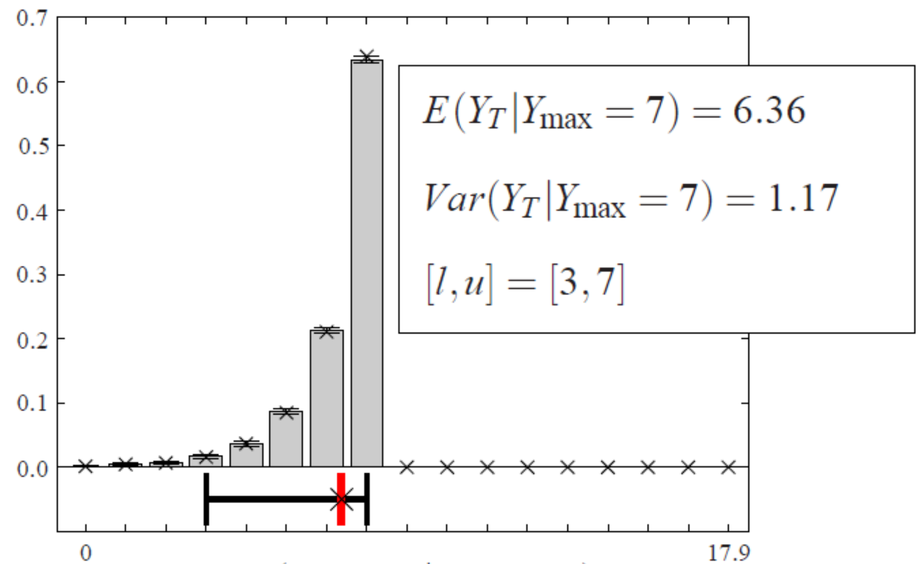

**Only the conditional expected value is presented**

# Dependence between Outputs

- Studied by calculating conditional probability distributions for fixed values of outputs (for example: $Y_{max} = 7$)




$P(\bar{Y}=\bar{y}|Y_{\max}=7).$



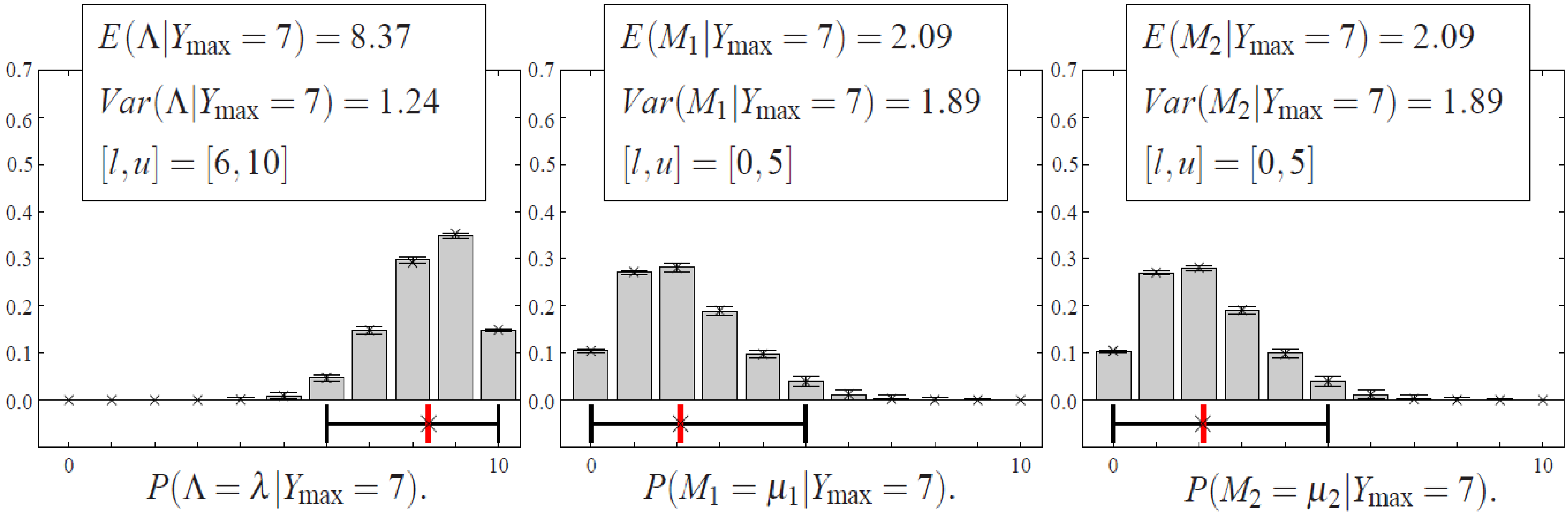

$P(Y_T=y_T|Y_{\max}=7).$

Aalto University
School of Science

Systems
Analysis Laboratory

# Inverse reasoning

- Probability distributions of inputs updated conditional on fixed values of outputs (for example: $Y_{max} = 7$ )
  - "Posterior uncertainty" (cf. "prior uncertainty" related to inputs)



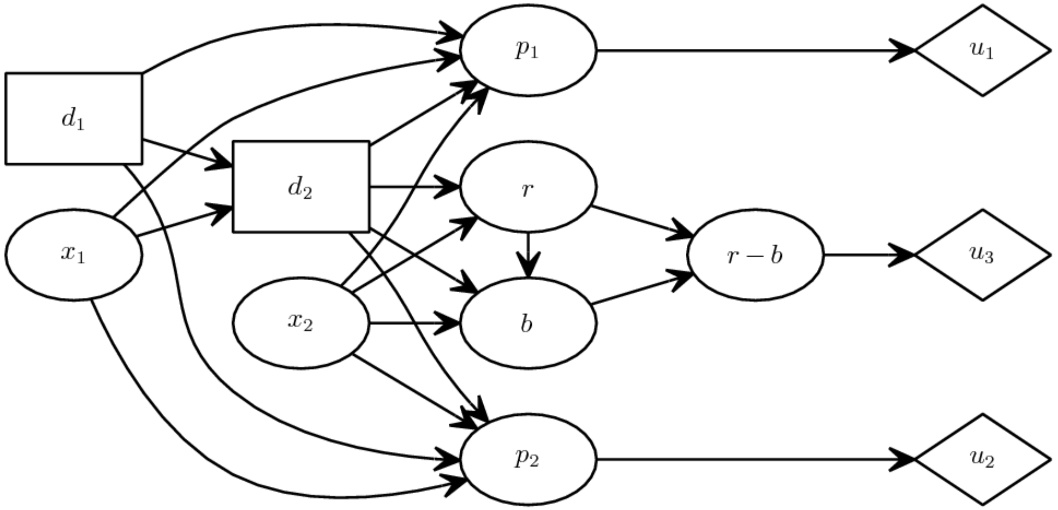
# Conclusion

- Simulation metamodeling benefits from BNs
  - Complete representation of probability distributions
    - No loss of information
      - cf. existing metamodels representing only expected values
  - New analysis capabilities
    - For example, input uncertainty and inverse reasoning
  - Available software with readily implemented algorithms
- Limitations
  - Construction necissitates large simulation data sets
  - Continuous variables have to be discretized

# Future research

- BN metamodeling
  - Continuous variables
    - Discretization
    - Interpolation
  - Error bounds
    - Bootstrapping
- Multi-criteria influence diagrams
  - Tool for decision support

Multi-Criteria Influence Diagram



# References


Friedman, L. W. 1996. *The simulation metamodel*. Norwell, MA: Kluwer Academic Publishers.

Henderson, S. G. 2003. Input model uncertainty: why do we care and what should we do about it?, *Winter Simulation Conference 2003*.

Kleijnen, J. P. C. 2008. *Design and analysis of simulation experiments.* New York, NY: Springer-Verlag.

Jensen, F. V., and T. D. Nielsen. 2007. *Bayesian networks and decision graphs*. New York, NY: Springer-Verlag.

Neapolitan, R. E. 2004. *Learning Bayesian Networks*. Upper Saddle River, NJ: Prentice Hall.

Poropudas, J., and K. Virtanen. 2007. Analyzing air combat simulation results with dynamic Bayesian networks, *Winter Simulation Conference 2007*.

Poropudas, J., and K. Virtanen. 2009. Influence diagrams in analysis of discrete event simulation data, *Winter Simulation Conference 2009*.

Poropudas, J., and K. Virtanen. 2010. Simulation metamodeling in continuous time using dynamic Bayesian networks, *Winter Simulation Conference 2010*.

Poropudas, J., and K. Virtanen. 2011. Simulation metamodeling with dynamic Bayesian networks, *European Journal of Operational Research*.

Poropudas, J., Pousi, J., and K. Virtanen. 2011. Multiple input and multiple output simulation metamodeling using Bayesian networks, *Winter Simulation Conference 2011*.

Poropudas, J. 2011. *Bayesian networks, Influence Diagrams, and Games in Simulation Metamodeling*, Doctoral dissertation, Aalto University School of Science.