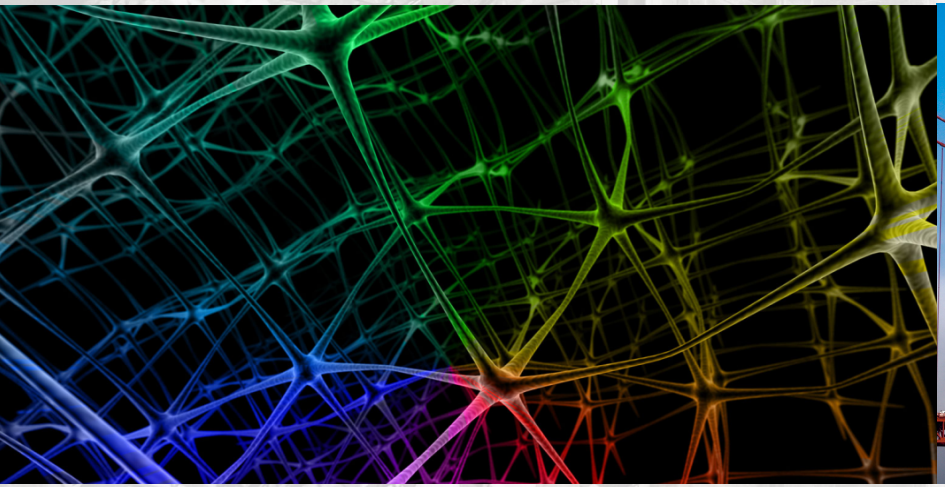
UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Transportation network planning using fast computational models

Hadi Meidani

Acknowledgment: Neelotpal Shukla, Mohammad Nabian

Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign



Uncertainty Quantification Group
at the University of Illinois at Urbana-Champaign

Two network planning problems

- **Network planning under uncertain Origin-Destination demands**
- Planning for resilient transportation networks

Motivation

- Policy makers and urban planners need estimates of traffic volumes on roads in order to make effective decisions
- Inaccurate estimations can lead to huge financial burdens
  - Overdesigned roads are expensive to construct
  - Underdesigned roads have to be expanded at a later stage – even more expensive
- Traffic volume forecasts are often wrong – 25% of road projects have difference between actual and forecasted traffics of more than ±40%[1]
- Trip generation and land use development responsible for more than 50% of incorrect forecasts in road projects[1,2]
- Major reason – forecasts based on static point-estimates of travel demand between urban zones
- Methods to incorporate uncertainty such as Monte Carlo integration require multiple model solves – effective but not efficient due to computation cost

[1]Flyvbjerg, Bent, Mette K. Skamris Holm, and Søren L. Buhl. "How (in) accurate are demand forecasts in public works projects?: The case of transportation." *Journal of the American Planning Association* 71.2 (2005): 131-146.

[2]Zhao, Yong, and Kara Maria Kockelman. "The propagation of uncertainty through travel demand models: an exploratory analysis." *The Annals of regional science* 36.1 (2002): 145-163.

Process of Traffic Volume Estimation[1]

Trip generation
- Convert socioeconomic factors to demand values
- Cross-classification or regression
- Biggest source of uncertainty – poor sampling, data aggregation, etc.[2]

Trip distribution
- Convert trip attraction-production to OD trip matrix
- Uncertainties in parameters of 'gravity model' – most commonly used – minor uncertainties

Mode split
- Assignment of traffic demand over multiple travel modes
- Not relevant in case of unimodal networks

Traffic assignment
- Assignment of traffic to edges in network
- User-equilibrium traffic assignment – no user has benefit from changing route choice

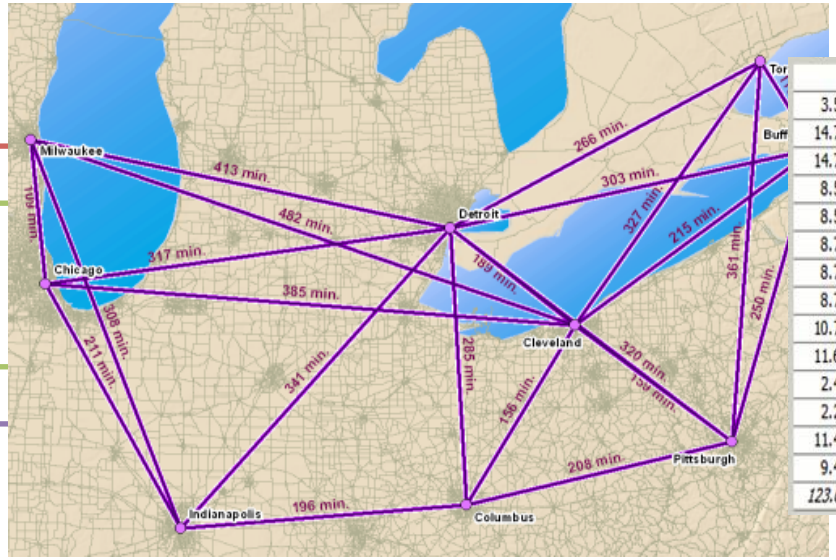
[1]McNally, Michael G. "The four-step model." *Handbook of Transport Modelling: 2nd Edition*. Emerald Group Publishing Limited, 2007. 35-53.
[2]Zhao, Yong, and Kara Maria Kockelman. "The propagation of uncertainty through travel demand models: an exploratory analysis." *The Annals of regional science* 36.1 (2002): 145-163.

Process of Traffic Volume Estimation[1]

Trip generation

Trip distribution

Mode split



	9.5	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	1	0.9	3.5	3.3
3.5		3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
14.7	17.9		7.4	7.4	7.4	7.4	7.4	7.4	6.7	6.7	2	2	6.7	6.7
14.7	17.9	7.4		7.4	7.4	7.4	7.4	7.4	6.7	6.7	2	2	6.7	6.7
8.5	10.7	4.6	4.6		4.6	5	5	4	3.9	3.9	1.3	1.4	4.1	4.1
8.5	10.7	3.9	3.9	3.9		3.9	3.9	3.9	3.9	3.9	1.3	1.4	4.1	4.1
8.7	17.9	4.1	4.1	4.1	4.1		4.1	4.1	4.1	4.1	1.3	1.4	4.1	4.1
8.7	17.9	5	5	5	5	5		4.1	4.1	4.1	1.3	1.4	4.1	4.1
8.7	11	5	5	5	5	5	5		4.1	4.1	1.3	1.4	4.1	4.1
10.7	13.8	4	4	4	4	4	4	4		5	1.3	1.3	4.1	4.7
11.6	15.6	2.1	2.1	2.1	2.1	2.1	2.1	2.1	5		1.3	1.4	4.4	12
2.4	3.3	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.4	1.4		1.6	5.6	0.8
2.2	3	0.8	0.8	0.8	0.8	0.8	0.8	0.8	1	1	1.3		0.8	0.8
11.4	14.5	2	2	2	2	2	2	2	4.5	4.5	1.4	1		6.1
9.4	17.2	2.7	2.7	2.7	2.7	2.7	2.7	2.7	6.1	6.1	1.4	1	4.7	
123.6	*180.9*	*49.4*	*49.4*	*52.2*	*52.9*	*53*	*52.2*	*50.3*	*58.4*	*58.5*	*21.6*	*21.7*	*60.6*	*65.3*

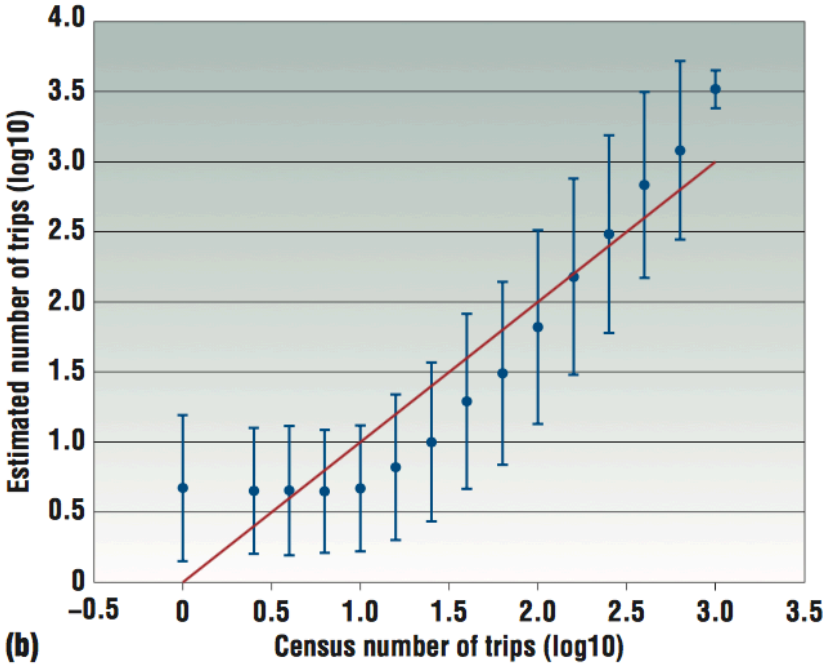
Traffic assignment

- Assignment of traffic to edges in network
- User-equilibrium traffic assignment – no user has benefit from changing route choice

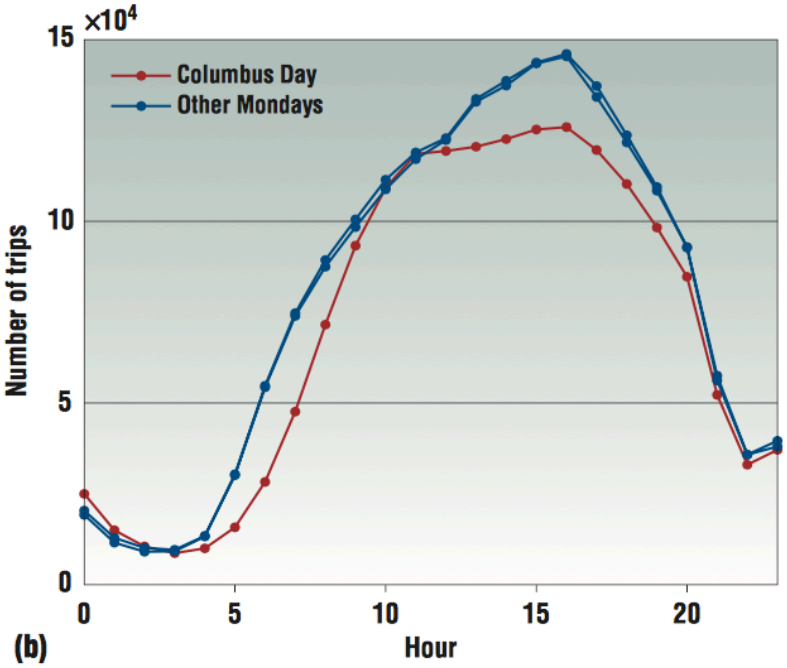
[1]McNally, Michael G. "The four-step model." *Handbook of Transport Modelling: 2nd Edition*. Emerald Group Publishing Limited, 2007. 35-53.

[2]Zhao, Yong, and Kara Maria Kockelman. "The propagation of uncertainty through travel demand models: an exploratory analysis." *The Annals of regional science* 36.1 (2002): 145-163.

Uncertainty in OD matrix



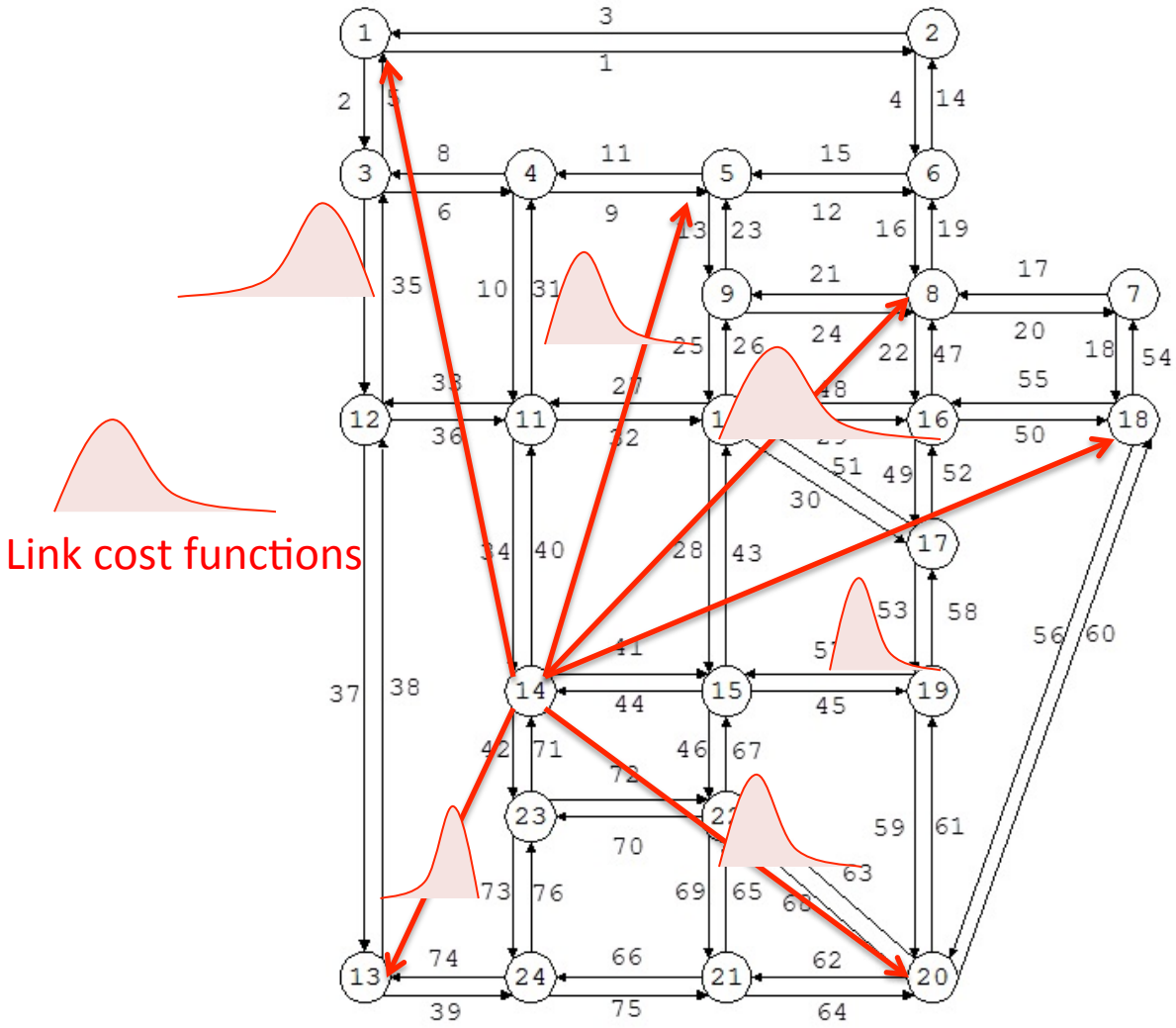
Comparison between mobile phone and census OD estimates



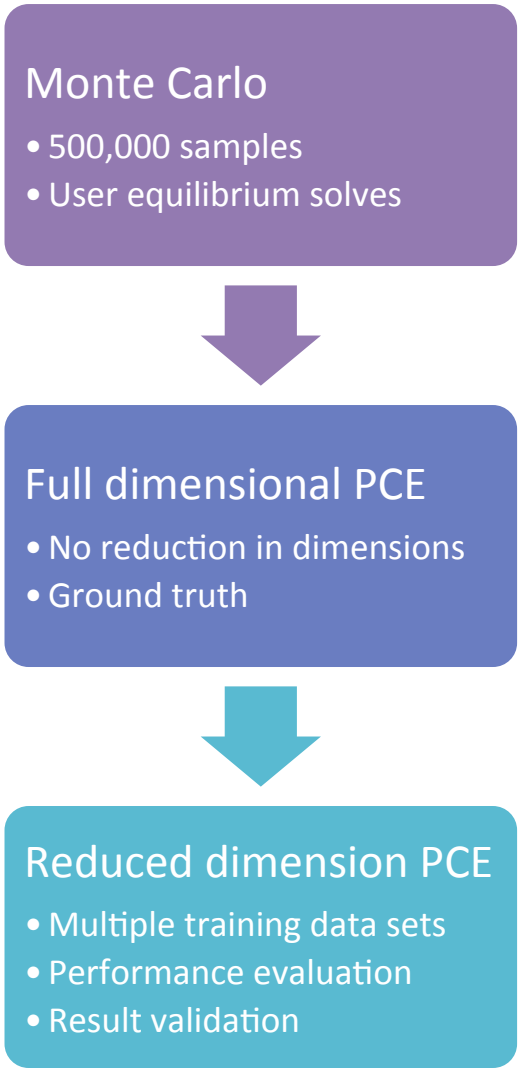
Comparison between trips in a regular day and a special day

[Calabrese, Francesco, et al. *IEEE Pervasive Computing* 99 (2011).]

Uncertainties in user equilibrium problem



Fast stochastic approach

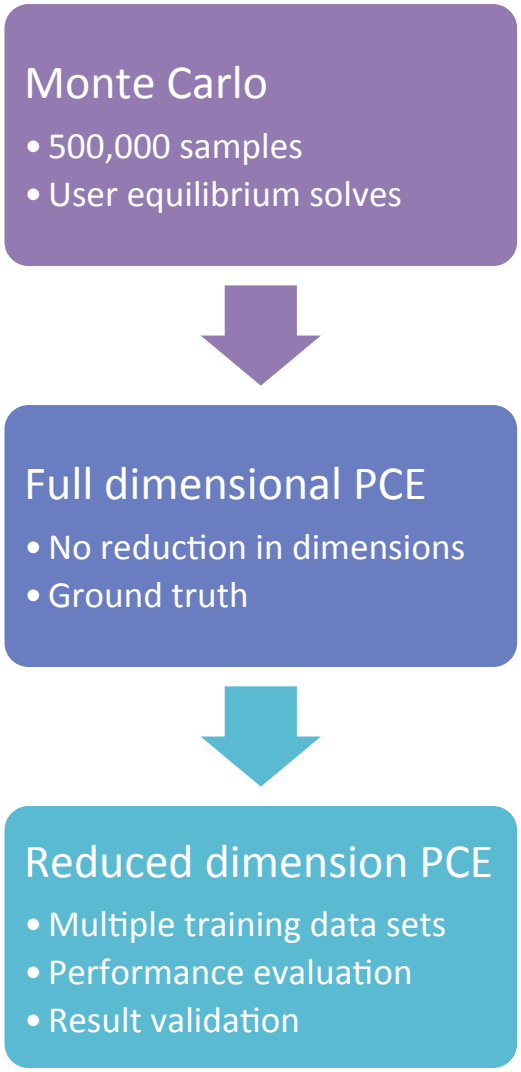


User Equilibrium

- Uncertain traffic demand modeled as uniform random variables with 50% variation about nominal value
- Variational inequality method[1]
- Iterative procedure creates constrained quadratic programs – can be solved using any of the available solvers[2]
- Network flow conservation constraints
- Number of iterations known a priori based on tolerance value – 1 in this case
- Symmetric positive definite matrix required as input – traffic cost coefficients from edges

[1]Dafermos, Stella. "Traffic equilibrium and variational inequalities." *Transportation science* 14.1 (1980): 42-54.
[2]Coleman, T.F. and Y. Li, "A Reflective Newton Method for Minimizing a Quadratic Function Subject to Bounds on Some of the Variables," *SIAM Journal on Optimization*, Vol. 6, Number 4, pp. 1040–1058, 1996.

Fast stochastic approach



PCE surrogate[1,2]

- High dimensional problem – second order selected
- Measurement matrix populated for all multi-indices for each sample set
- Least squares method used to solve for coefficients
- Coefficient amplitude plotted against coefficient index in order to analyze sparsity and identify important coefficients – amplitude higher than set threshold 10^{-2}
- Dimensions associated with important coefficients obtained and used as ground truth

[1]Xiu, Dongbin, and George Em Karniadakis. "The Wiener--Askey polynomial chaos for stochastic differential equations." *SIAM journal on scientific computing* 24.2 (2002): 619-644.
[2]Ghanem, Roger G., and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Courier Corporation, 2003.

Fast stochastic approach

Monte Carlo
- 500,000 samples
- User equilibrium solves

Full dimensional PCE
- No reduction in dimensions
- Ground truth

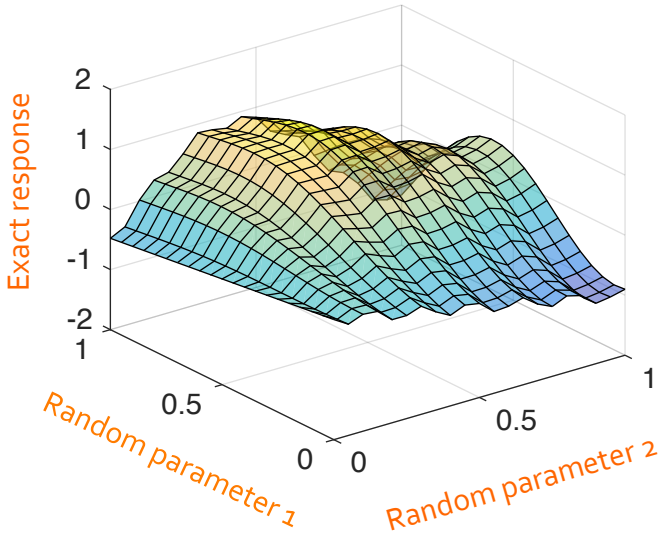
Reduced dimension PCE
- Multiple training data sets
- Performance evaluation
- Result validation

Dimension Reduction

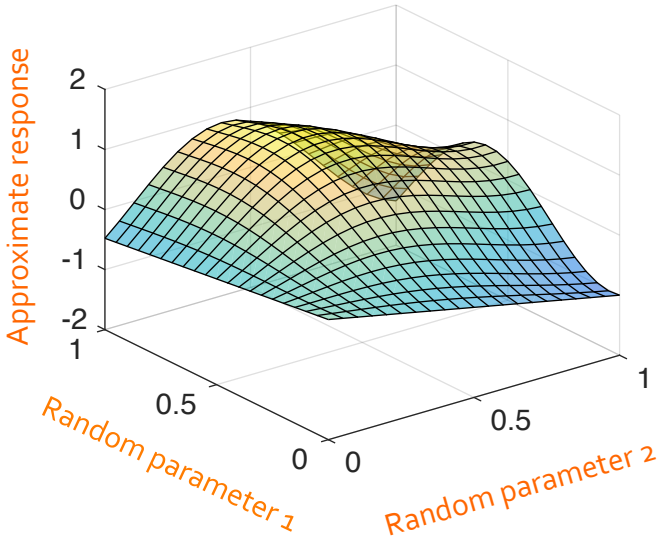
- Greedy algorithm for dimension reduction[1]
- Starts with 1 dimension and order 2 – identifies most important dimension based on sparsity of coefficient vector
- Iteratively adds dimension or order till no further improvements can be made i.e. error tolerance reached – 10^{-2} and 10^{-3} tested
- 3 different sizes for training data – 100,300 and 500
- 5 different sampling locations for each size
- Results evaluated using cross validation – L2 norm of normalized vector

[1]Alemazkoor, Negin, and Hadi Meidani. "Divide and conquer: An incremental sparsity promoting compressive sampling approach for polynomial chaos expansions." *Computer Methods in Applied Mechanics and Engineering* 318 (2017): 937-956.

Fast surrogates

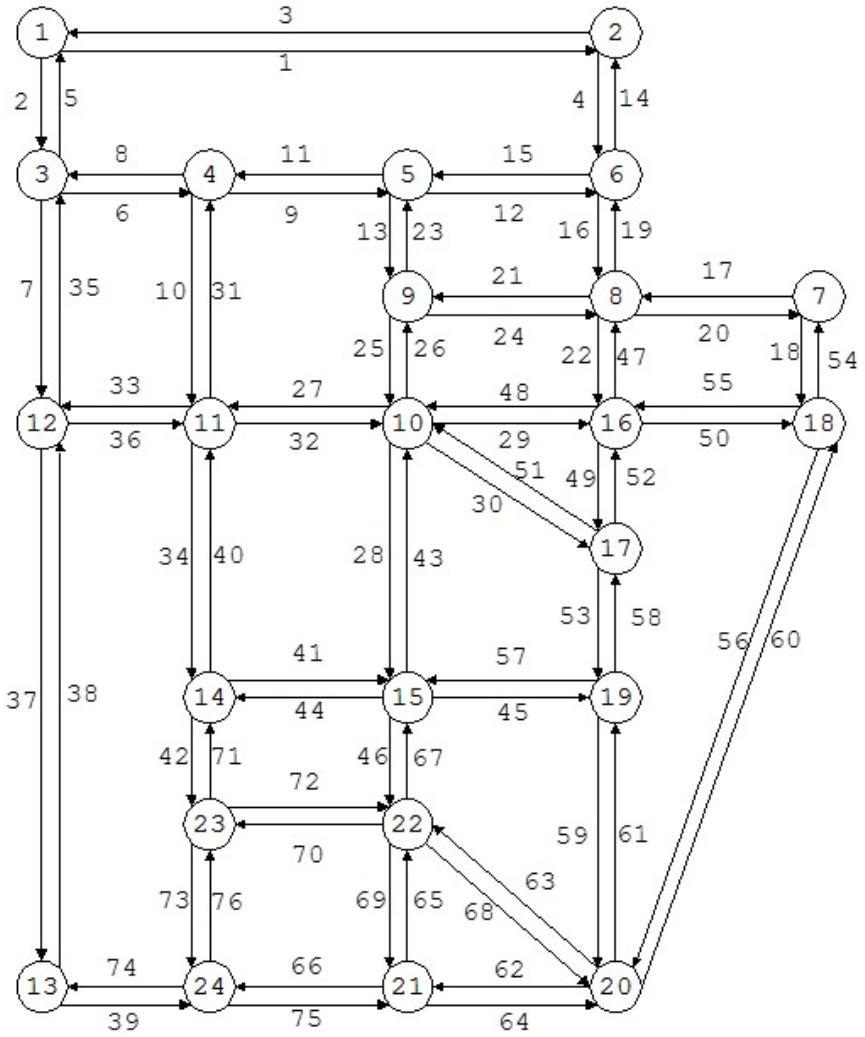


Response from expensive solver



Cheap (analytical) response surface

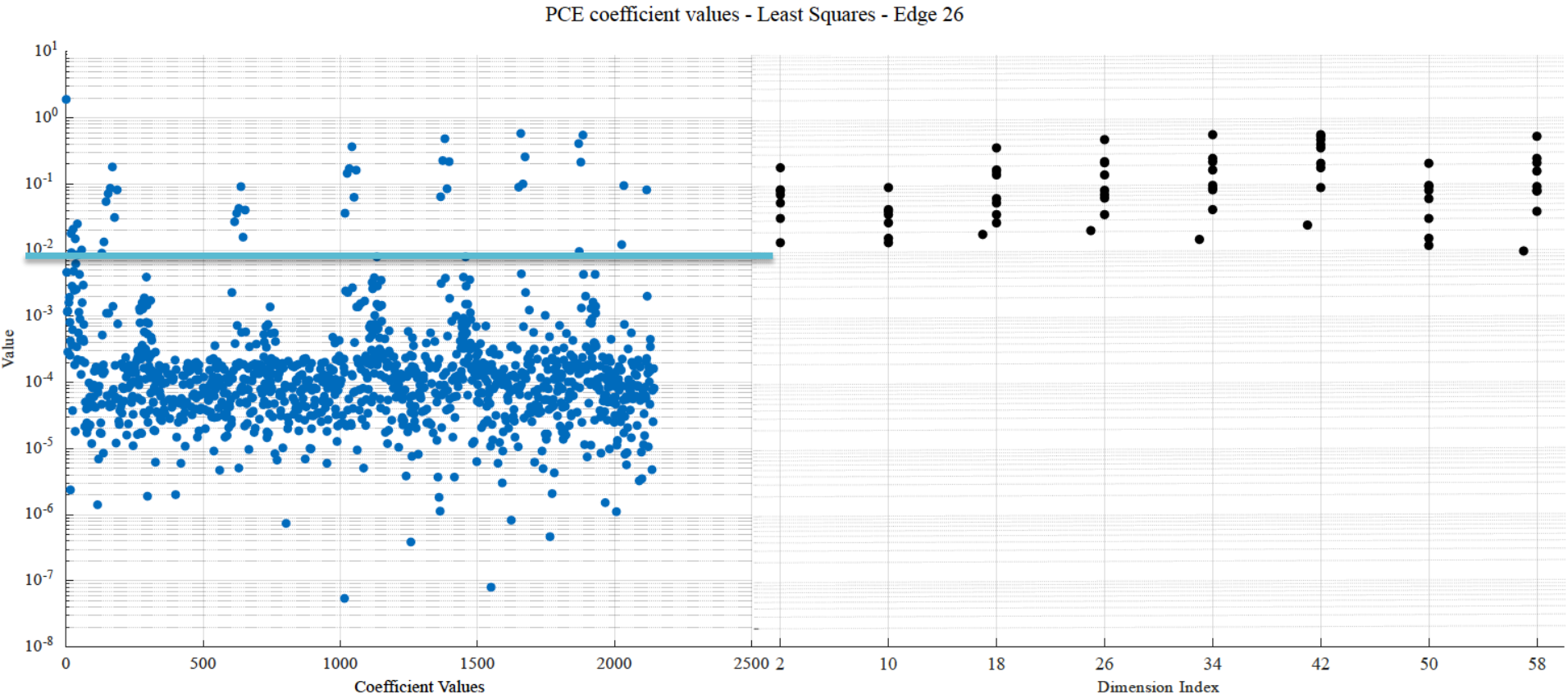
Problem setup



- Benchmark transportation network – Sioux Falls
  - 24 nodes (also centroids)
  - 76 directed edges
- 11% uncertainty in network
  - Uncertain OD pairs: 64
  - Total OD pairs: 576
- Edge analyzed: 26
  - Origin node: 10 (uncertain)
  - Terminal node: 9 (uncertain)
- Intuitively, traffic volume on edge 26 should not be affected by all the 64 OD pairs
- OD pairs with node 10 as origin should have the highest impact on edge 26 traffic

Complexity of the surrogate

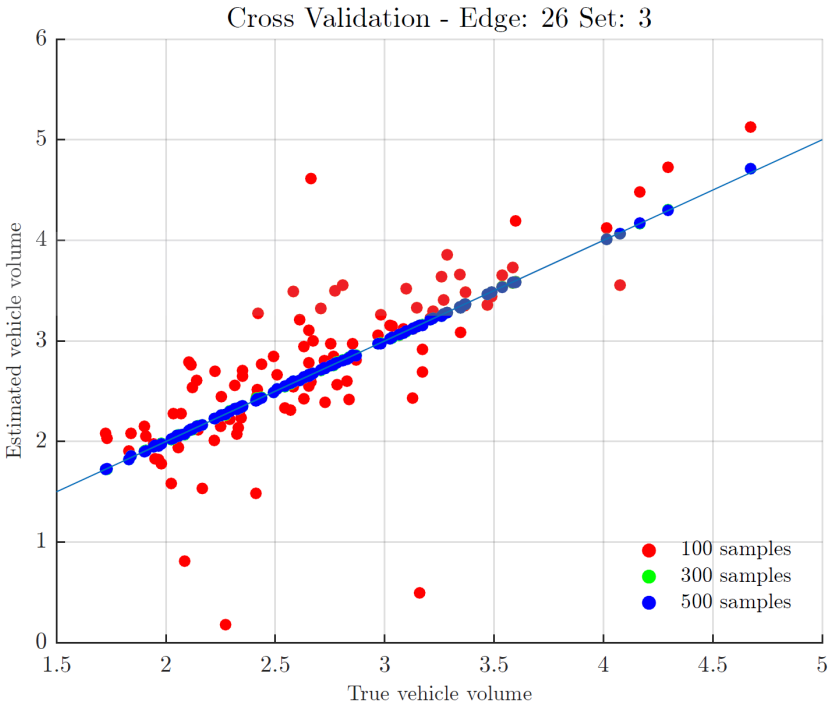
10^{-2} used as amplitude threshold



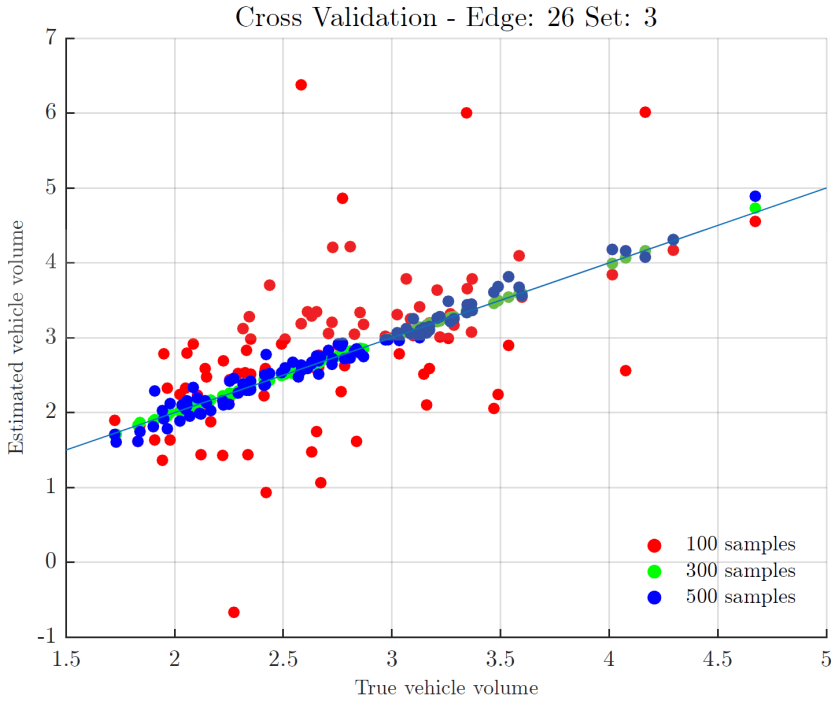
Important dimensions i.e. OD pairs that are influential

Cross Validation – Reduced Dimension PCE

Threshold: 10^{-2}



Threshold: 10^{-3}



- Testing data sample size: 100
- All 8 important dimensions captured by both surrogates
- Similar performance by surrogates constructed using 300 and 500 samples
- Better performance by surrogates constructed using higher error threshold.

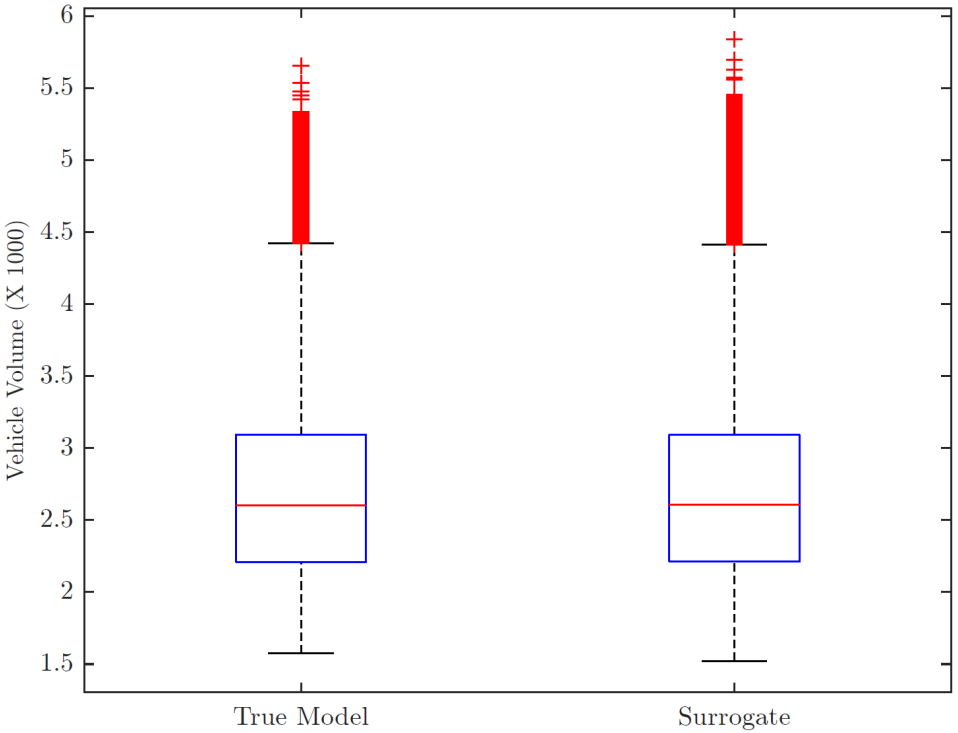
Statistics

Based on surrogate created using 300 samples – 2nd best L2 score of 0.0298

Surrogate using 500 samples had score 0.0287 – a marginal increase considering the 67%
increase in number of samples

True model − actual results
Surrogate − 500,000 points evaluated using surrogate

Statistic	True Model	Surrogate
Mean	2.6943	2.6937
Variance	0.3849	0.3847
Median	2.6040	2.6076



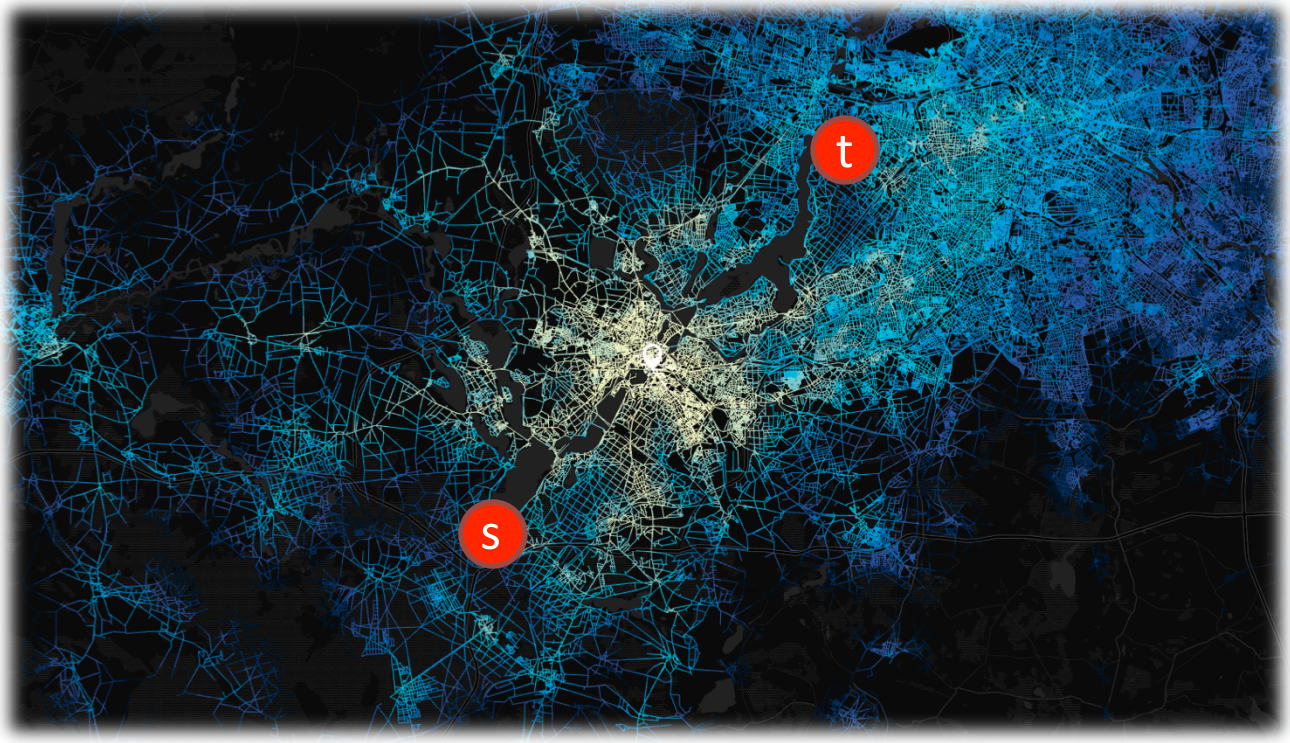
Discussion and conclusion

- Accurate surrogates can be constructed using few data samples using dimension reduction technique
  - Accuracy validated through cross-validation procedure and comparison of statistical values
  - Both point estimates and complete distribution accurately captured
  - Evaluated at large number of data points to effectively characterize flow within network under uncertain traffic demand scenarios
- Technique applicable because of intuition – flow on a certain edge only depends upon the uncertainty in OD pairs the demand between which is routed through the edge in consideration
  - Edge-specific traffic volume estimation can be formulated as a sparse problem – superficially high dimensional but outcomes depends on a subset of dimensions
- Surrogates constructed using low error threshold tend to have higher order whereas surrogates of higher error threshold capture additional dimensions but remain low order
  - Dimensions capture more information than order – based on evidence
- It can significantly speed up network design problems

Two network planning problems

- Network planning under uncertain Origin-Destination demands
- **Planning for resilient transportation networks**

Two-terminal Reliability [1-5]



Quantity of interest: Probability that nodes s and t are connected after the given earthquake.

Why is this important?

The two-terminal connectivity is relevant when, for instance, the accessibility from a major attraction point to a major hospital, or from a feedstock to demand zones, is to be maintained.

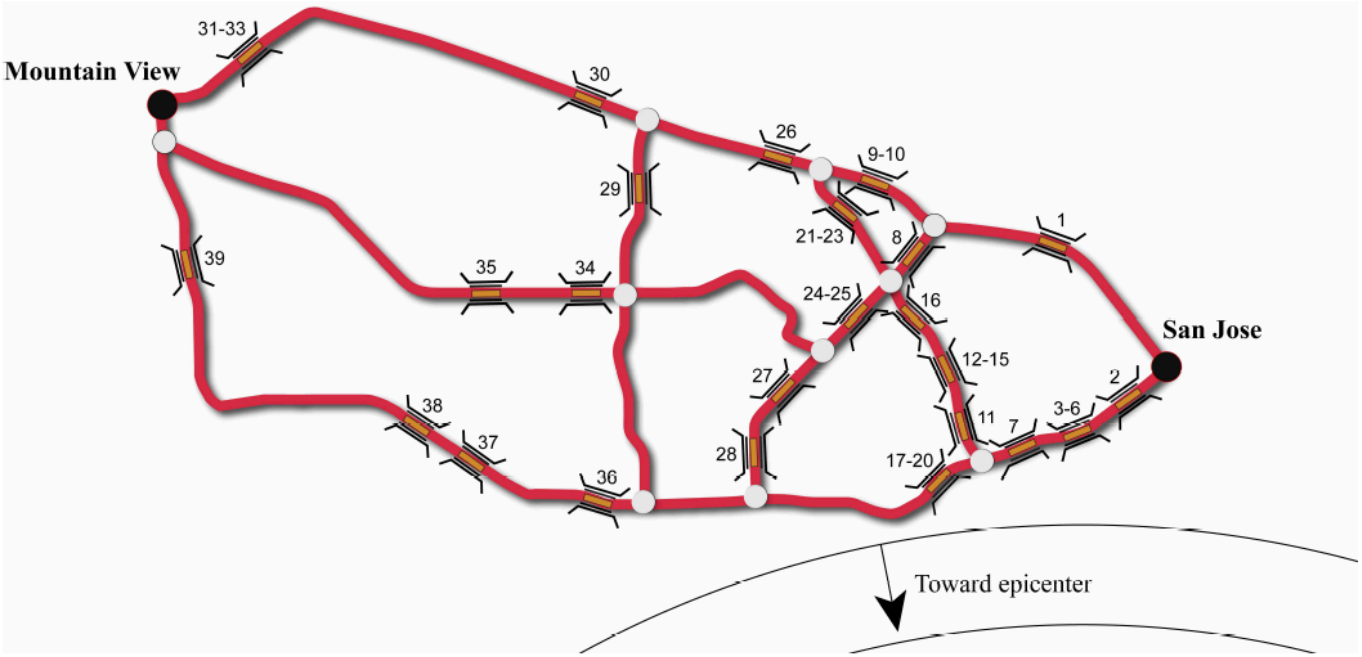
Standard Procedure for Two-Terminal Connectivity Evaluation

1. Find the component-level response of the system subject to an earthquake.
   - The component level response is bridge failure probabilities.
   - Involves ground motion intensity prediction and bridge fragility analysis, which are computationally inexpensive [6].

2. Calculate the system-level response based on the component-level response.
   - Computationally expensive.
   - Is usually performed using Monte Carlo Simulation.
   - In each MCS, a network realization is generated (i.e. bridges will randomly fail or survive), and then the network connectivity is evaluated (i.e. 0 or 1), USING Depth-First Search.
   - The process is repeated until convergence is achieved.

Expensive!

Case study: San Jose to Mountain View network

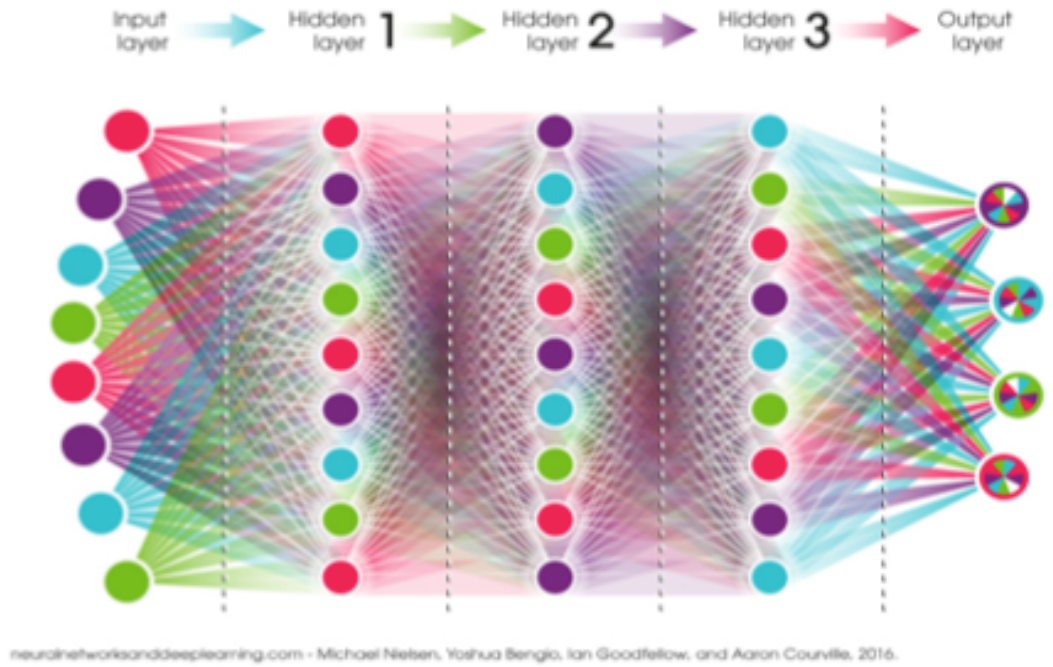


Goal: Quantify the relationship between output (i.e., connectivity) and inputs (i.e., failure probability, network topology).

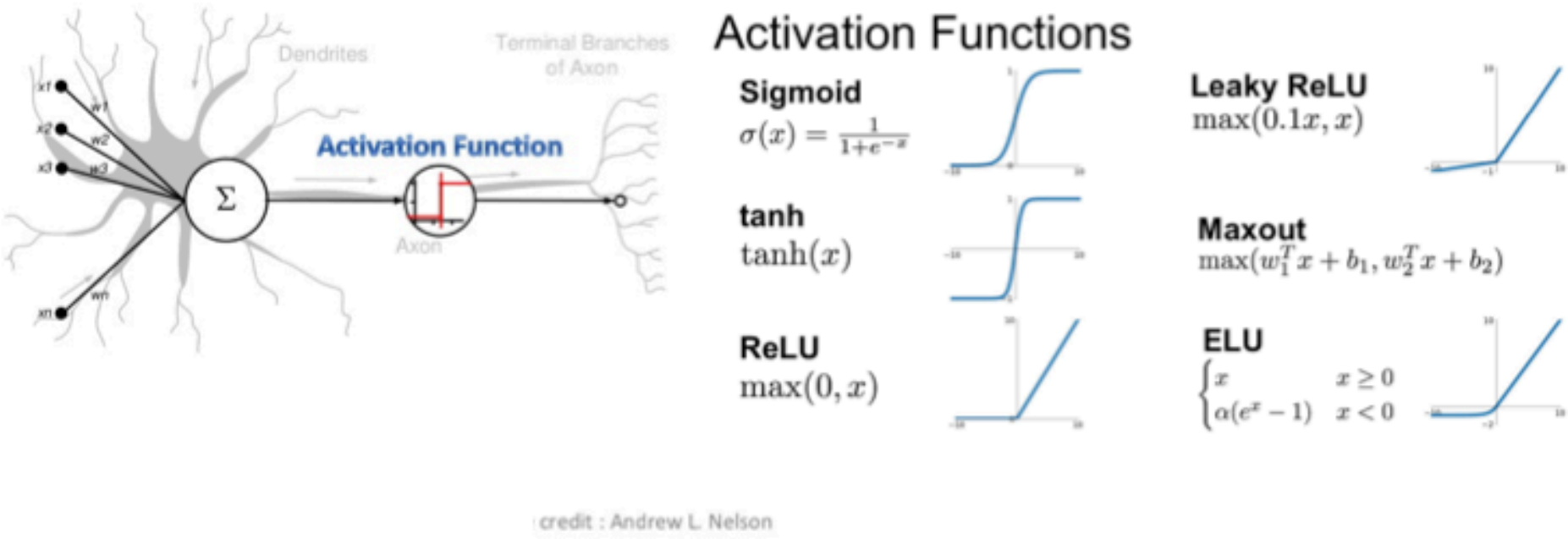
Challenge: Computation cost is too large/impossible to handle if Monte Carlo Simulation is used.

Methodology: Using DNN instead of Monte Carlo Simulation to alleviate computational burden.

What is a Deep Neural Network?



A deep neural network consists of one input, one output and multiple fully-connected hidden layers. Each layer is represented as a series of neurons and progressively extracts higher and higher-level features of input until the final layer makes a decision about what the input shows. The more layers the network has, the higher- level features it will learn [7,8].

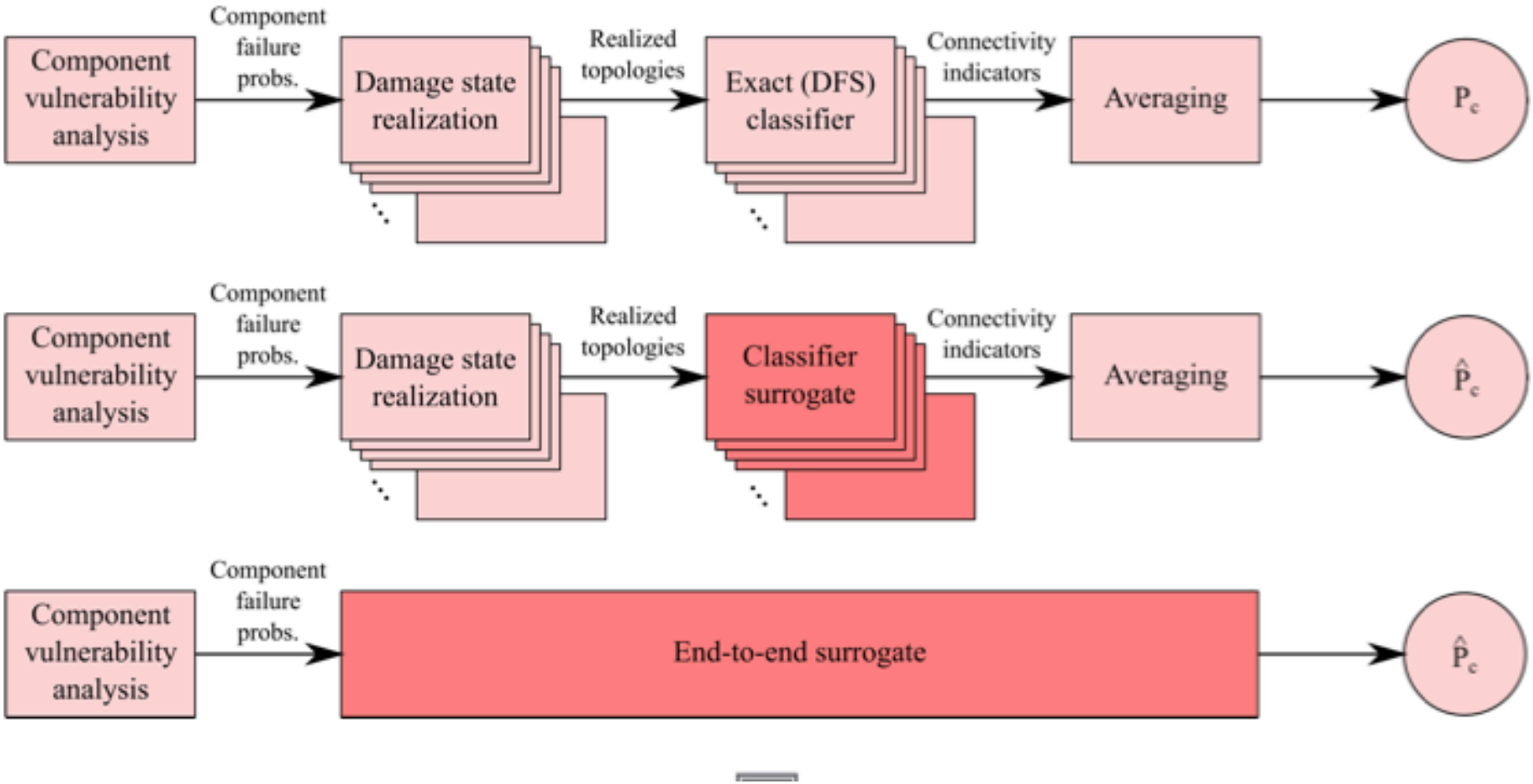


credit : Andrew L. Nelson

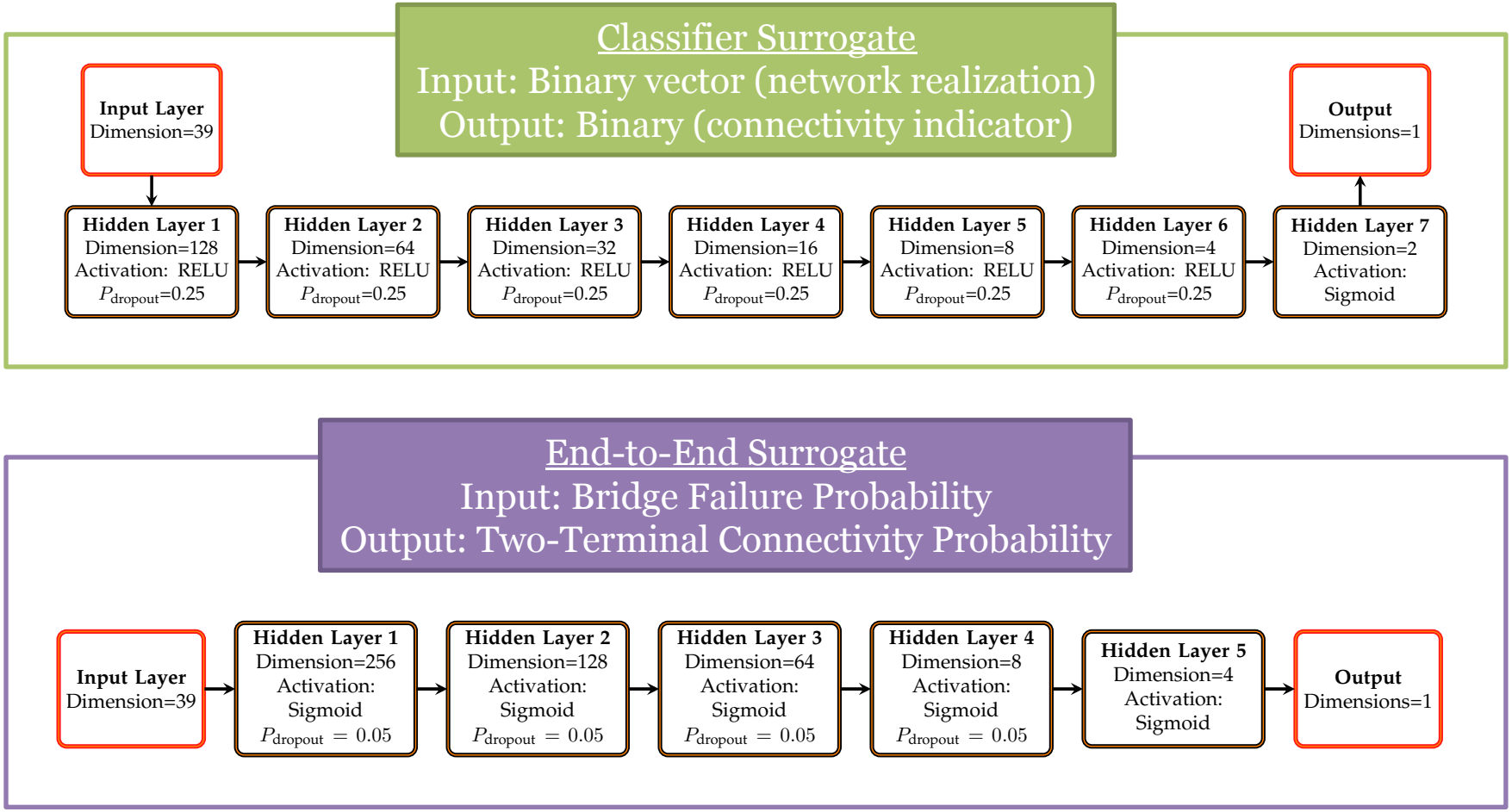
An artificial neuron contains a nonlinear activation function and has several incoming and outgoing weighted connections.

Neurons are trained to filter and detect specific features or patterns by receiving weighted input, transforming it with the activation function and passing it to the outgoing connections.

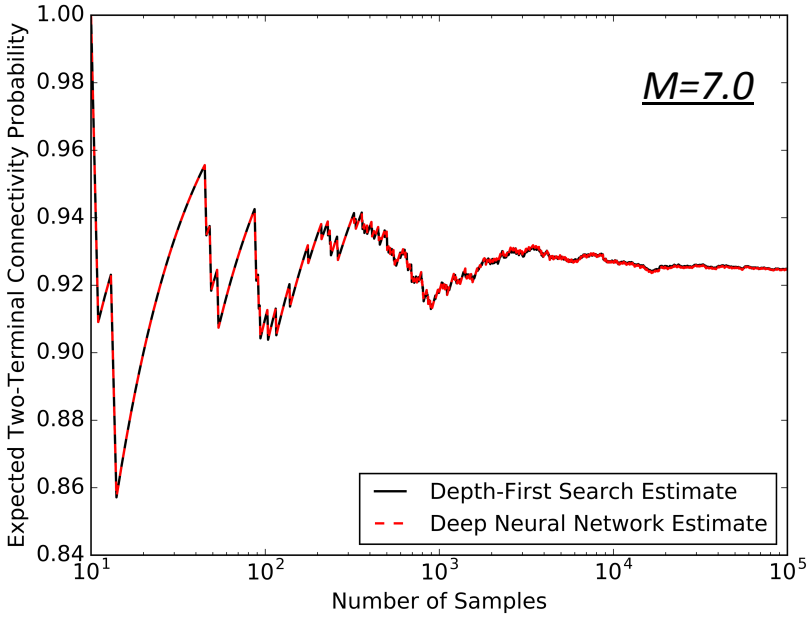
Workflow for calculation of the expected two-terminal connectivity

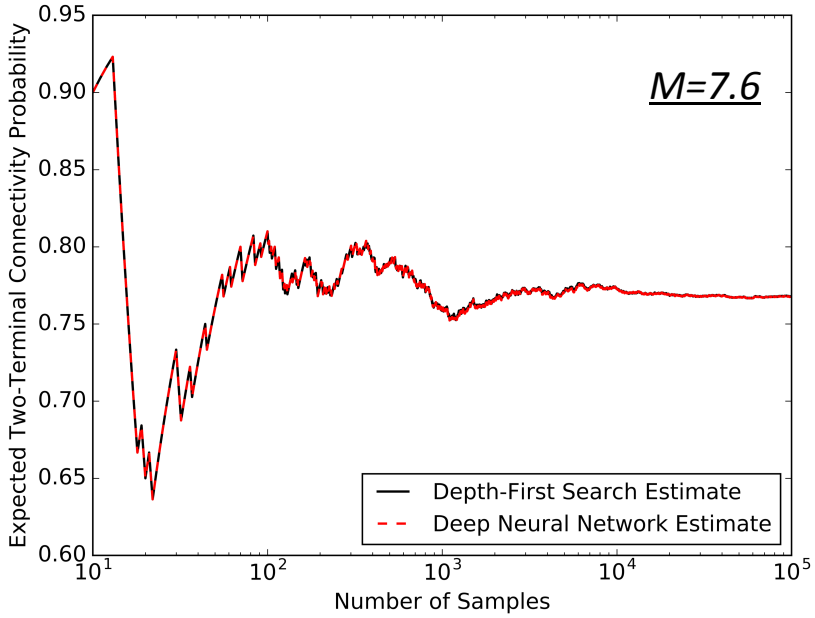


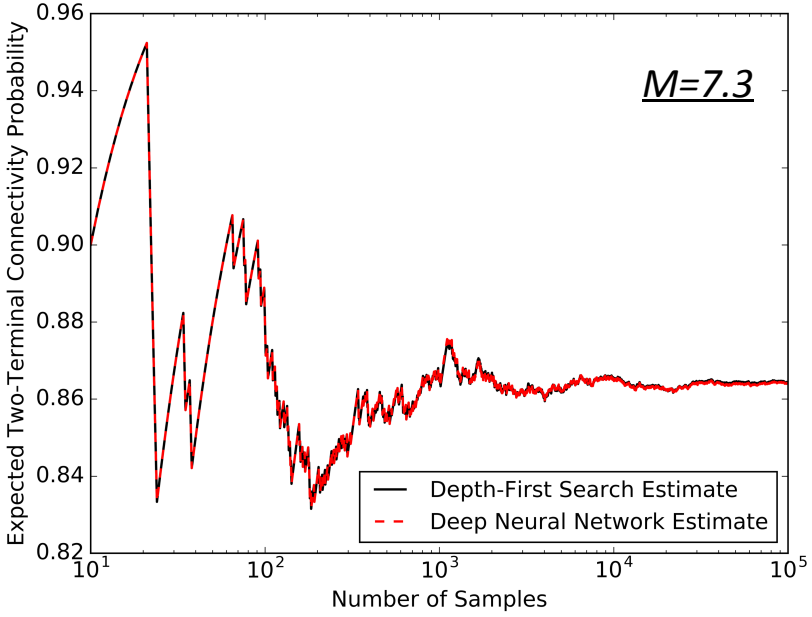
Two Types of Surrogates

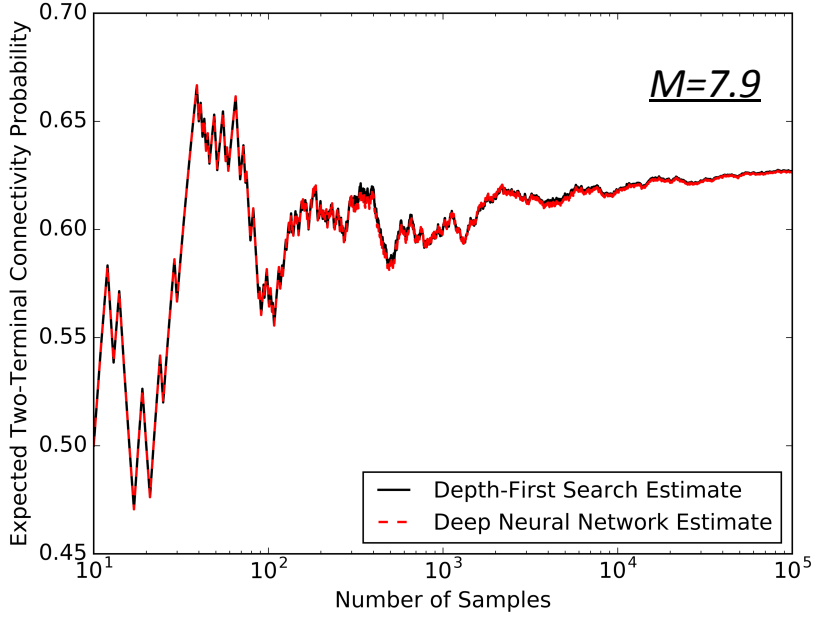


Classifier surrogate results for earthquakes with deterministic magnitude

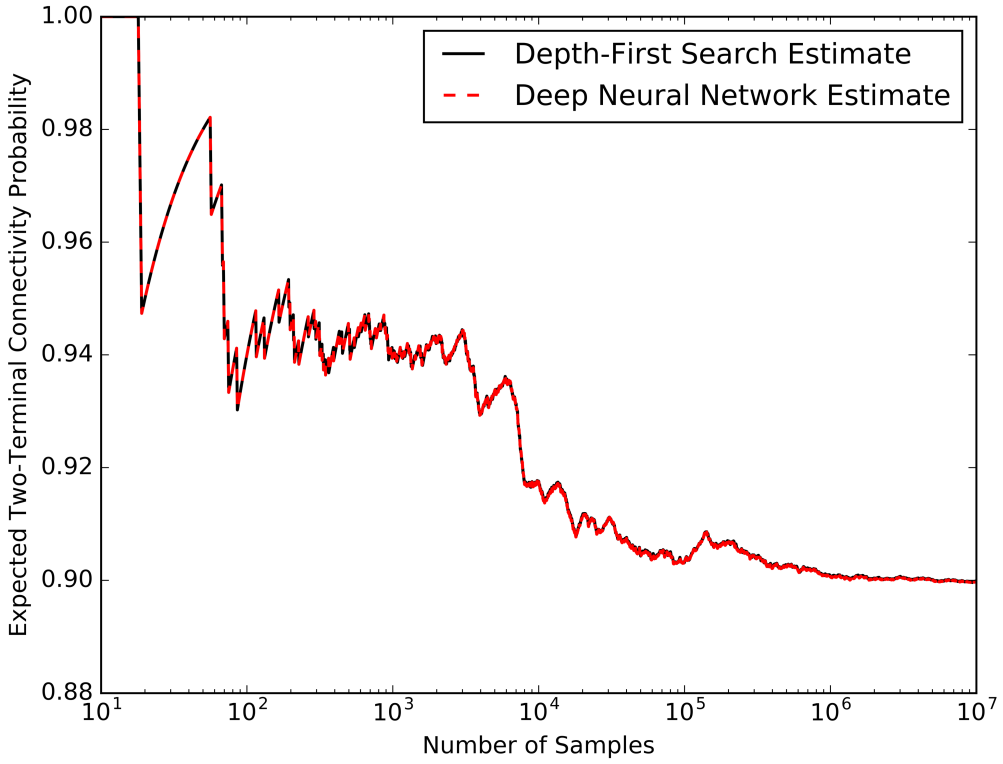






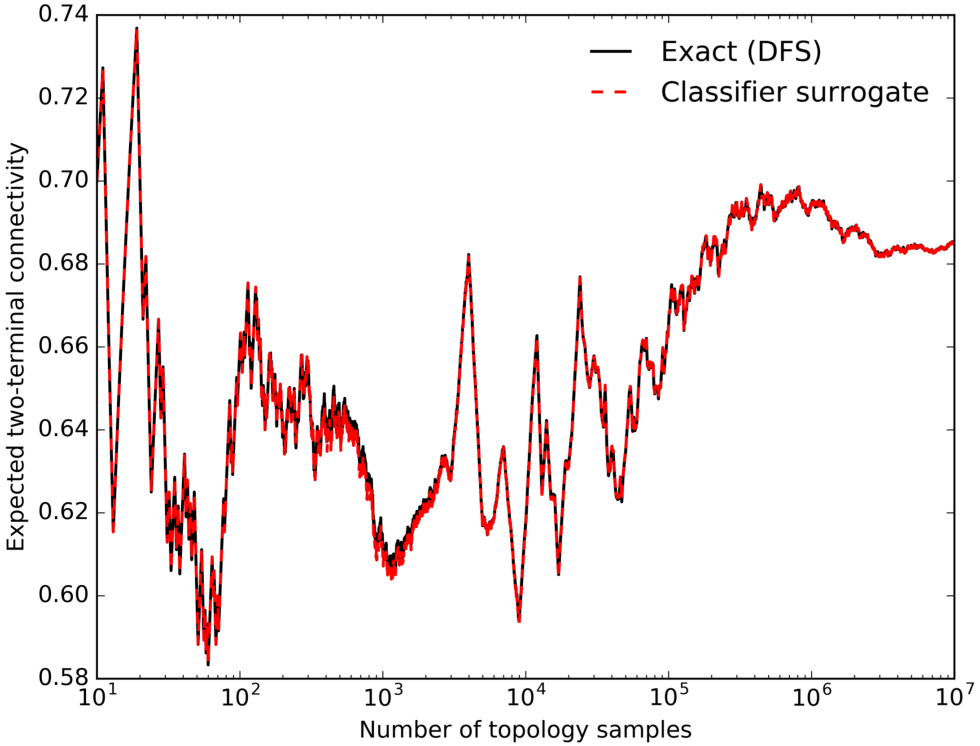


Classifier surrogate results for earthquakes with probabilistic magnitude (ranging between 6.9-7.4 Mw)



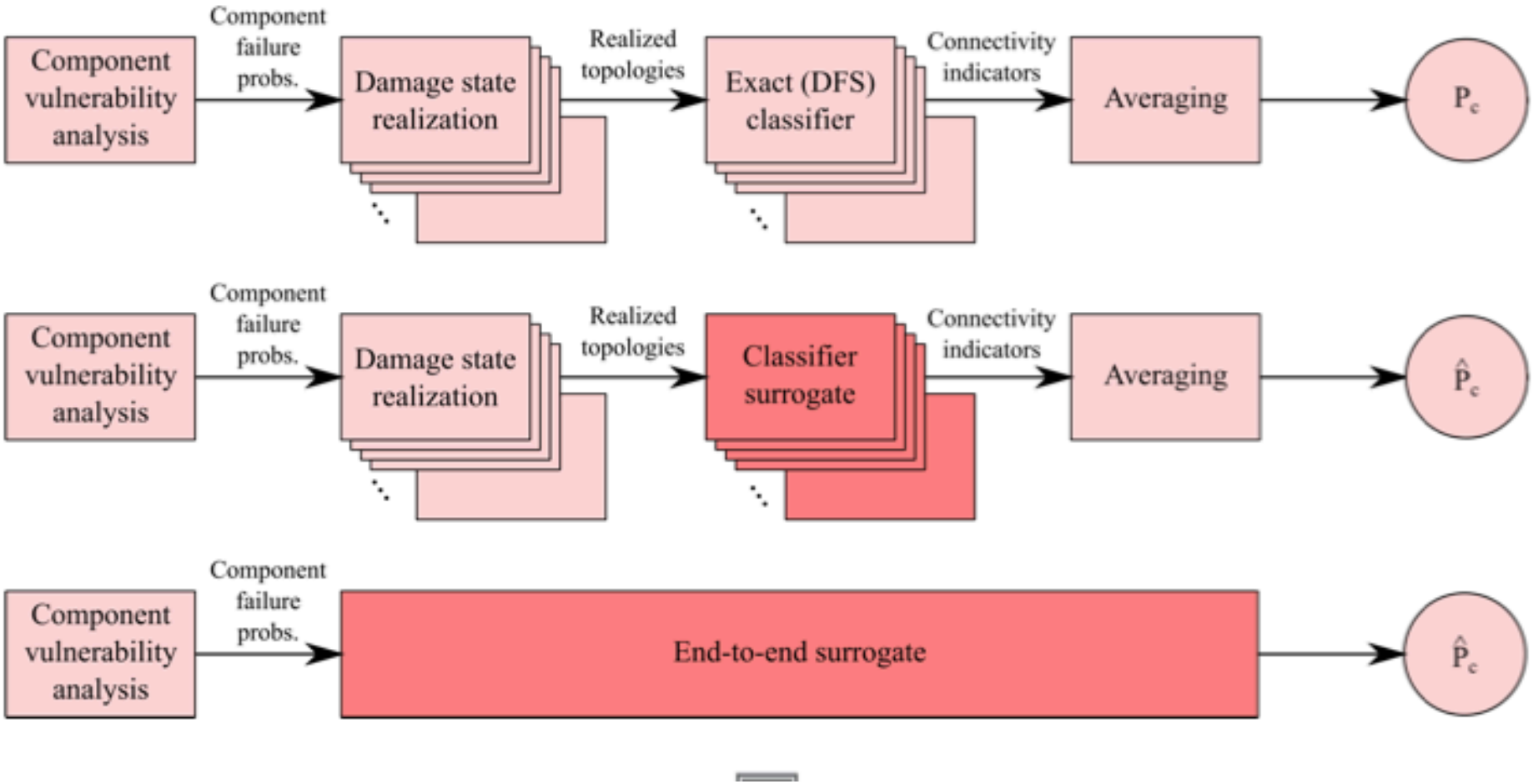
Standard Computational time: 647.83 s
Surrogate computational time: 98.33 s

Classifier surrogate results for earthquakes with probabilistic magnitude (ranging between 6.9-7.4 Mw) considering the GMPE uncertainties

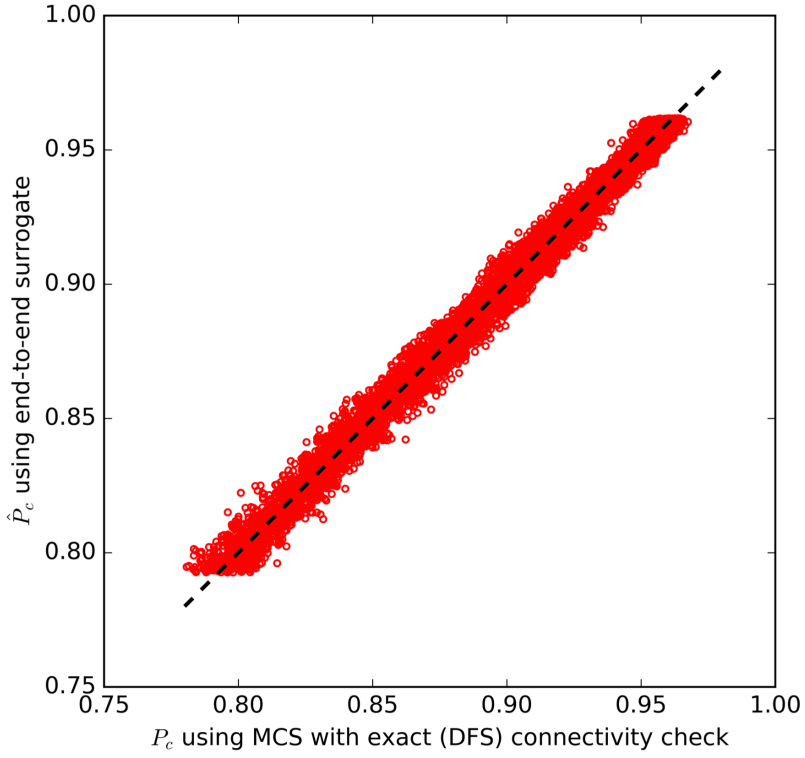


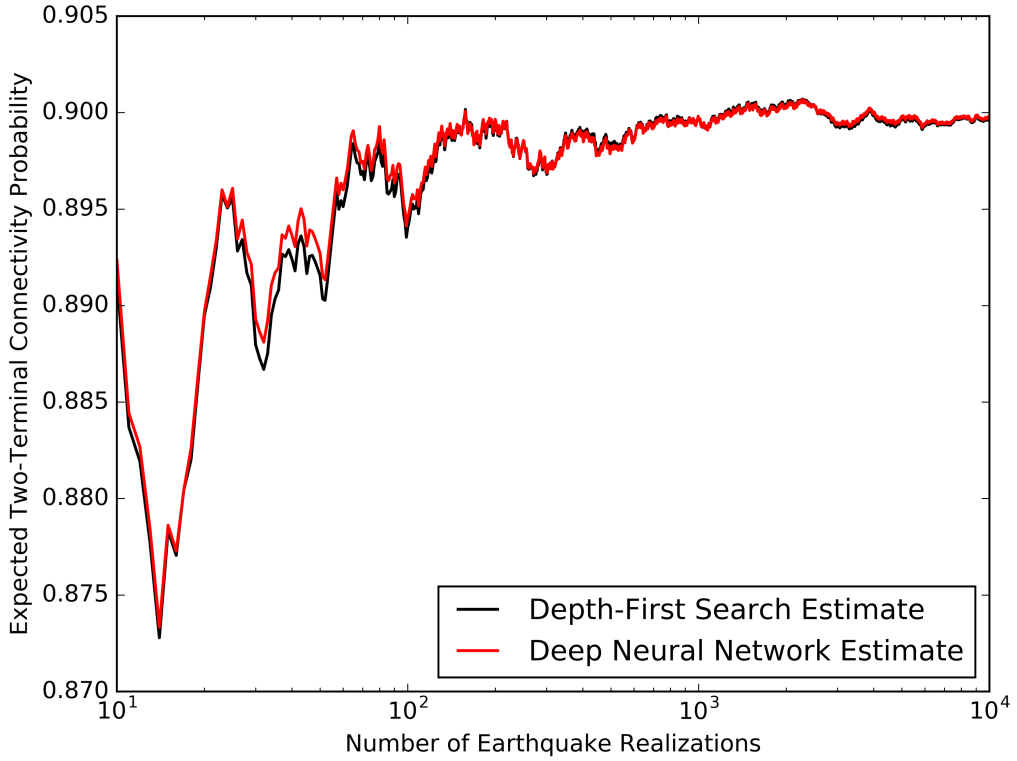
Standard Computational time: 647.83 s
Surrogate computational time: 98.33 s

Workflow for calculation of the expected two-terminal connectivity



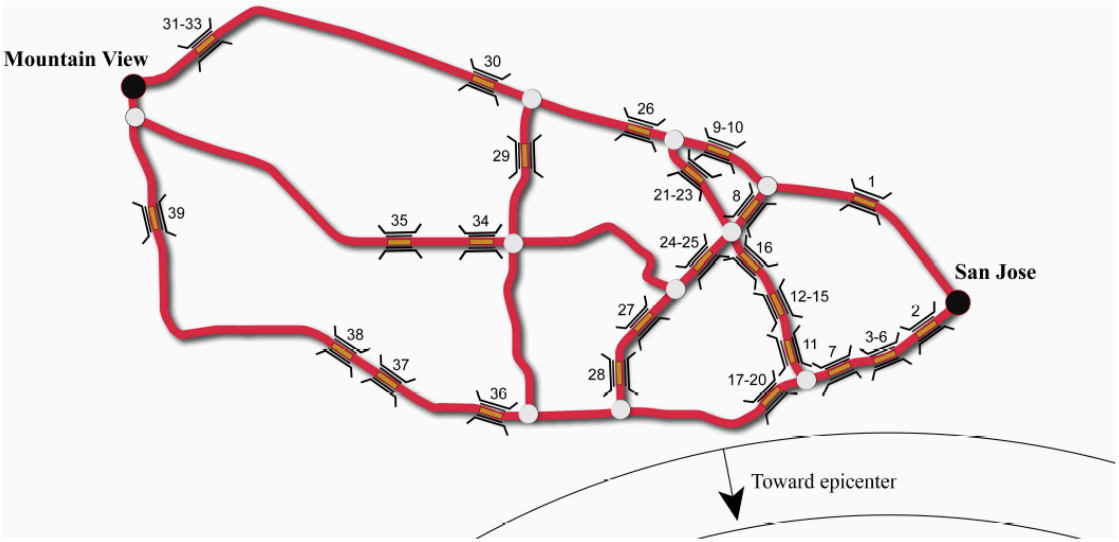
End-to-end surrogate results for earthquakes with probabilistic magnitude (ranging between 6.9-7.4 Mw)





Standard Computational time: 7882.74 s
Surrogate computational time: 0.71 s

One-at-a-time sensitivity analysis in network planning



Rank	Bridge ID	Improvement in connectivity (%) (DFS estimate)	Improvement in connectivity (%) (surrogate estimate)	DFS time (s)	Surrogate time (s)
1	0	8.22	8.19	8361.22	0.75
2	13	7.80	7.78	7855.76	0.72
3	25	2.17	2.14	8334.18	0.74
⋮	⋮	⋮	⋮	⋮	⋮
37	14	0.07	0.06	8089.66	0.75
38	10	0.01	0.01	8131.76	0.74
39	26	0.00	0.00	8079.06	0.75

Summary

Goal: Quantify the relationship between output (i.e., connectivity) and inputs (i.e., failure probability, network topology).

Challenge: Computation cost is too large/impossible to handle if Monte Carlo Simulation is used.

Methodology: Using gradient of DNN instead of Monte Carlo Simulation to alleviate computational burden.

Applications:

-- Test robustness of system in the presence of uncertainty.

-- Identify critical system components for upgrading/investment [9].

References

[1] Bocchini, P. and Frangopol, D. M. (2011). Generalized bridge network performance
analysis with correlation and time-variant reliability. *Structural Safety*, 33(2):155–164.

[2] Bocchini, P. and Frangopol, D. M. (2011). A stochastic computational framework for the joint transportation network fragility analysis and traffic flow distribution under extreme events. *Probabilistic Engineering Mechanics*, 26(2):182–193.

[3] Stern, R., Song, J., and Work, D. (2017). Accelerated Monte Carlo system reliability analysis through machine-learning-based surrogate models of network connectivity. *Reliability Engineering & System Safety*, 164:1–9.

[4] Zacks, S. (2012). *Introduction to reliability analysis: probability models and statistical methods*. Springer Science & Business Media.

[5] Faturechi, R. and Miller-Hooks, E. (2014). Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review. *Journal of infrastructure systems*, 21(1):04014025.

[6] Graizer, V. and Kalkan, E. (2016). Summary of the GK15 ground-motion prediction equation for horizontal pga and 5% damped psa from shallow crustal continental earth- quakes. *Bulletin of the Seismological Society of America*, 106(2):687–707.

[7] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT Press.

[8] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436– 444.

[9] Buckle, G., Friedland, I., Mander, J., Martin, G., Nutt, R., and Power, M. (2006). *Seismic retrofitting manual for highway structures*. MCEER, Buffalo.

Thank you!