PREDICTING PFAS EXPOSURE RISKS FROM RURAL PRIVATE WELLS USING INTEGRATED MECHANISTIC AND MACHINE-LEARNED BAYESIAN NETWORK MODELS

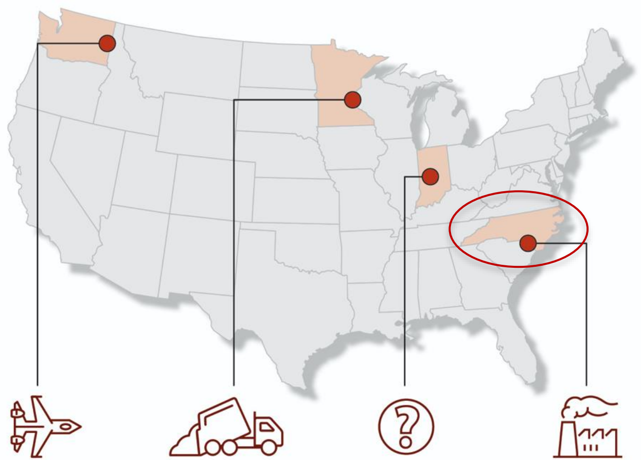
Thursday, April 11, 2024, BayesiaLab Spring Conference, Cincinnati, OH

NCSU: **Hana C Long** (presenter), Jackie McDonald Gibson, Krishnamohan Ganta
RTI: Rohit Warrier, Riley Mulhern, Jennifer Hoponick Redmon, Ted Lillys

Overall Project

NC STATE UNIVERSITY · RTI INTERNATIONAL · INDIANA UNIVERSITY · EMORY UNIVERSITY

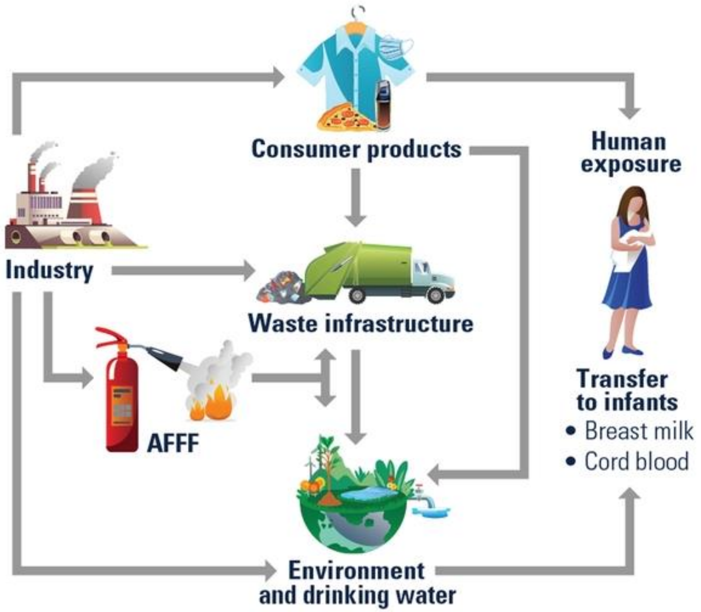
1. Build computational models to predict PFAS risk in private wells
   - Mechanistic fate/transport model
   - Machine-learned Bayesian Network
   - Integrated mechanistic/MLBN model
   - Model validation
2. Conduct citizen-science well monitoring campaign
3. Develop user-friendly risk map



	Airway Heights, Spokane County, WA	Washington County, MN	Monroe County, IN	Robeson County, NC
Community Partner(s)	• Washington Depatment of Health • Spokane Regional Health District	• Minnesota Department of Health	• Monroe County Health Department	• Robeson County Health Department • North Carolina Division of Public Health • American Indian Mothers
Major Potential PFAS Source(s)	• Fairchild Air Force Base	• Waste disposal sites used by the 3M Company	• No known point sources; potential diffuse sources	• Chemours Fayetteville Works • Land application of biosolids
Population Relying on Private Wells	59,506	47,417	7,380	40,250

PFAS Exposure

- Persistent in body (especially legacy/long-chain PFAS)
- Increases risk of cancer, infertility, liver damage, obesity, affects birthweight, child development, immune function, cholesterol, thyroid function
- Most Americans exposed through drinking water, most water treatment doesn't remove legacy PFAS (EPA 2021)
- PFAS detectable in blood of most Americans (Lewis et al. 2015)



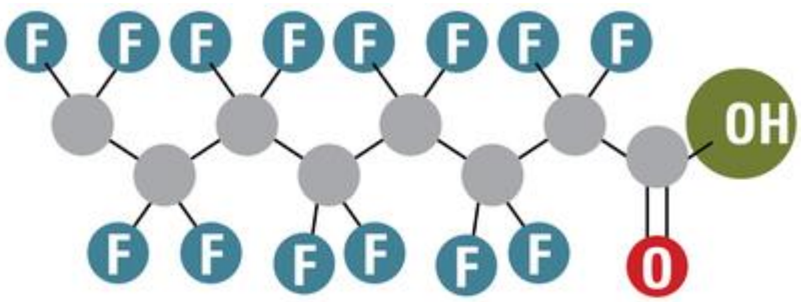
USGS 2022

https://epi.dph.ncdhhs.gov/oee/pfas/PFAS_Factsheet.pdf
https://genxstudy.ncsu.edu/our-findings/
https://content.ces.ncsu.edu/Guide-to-Understanding-and-Addressing-PFAS-in-our-communities

PFAS Types

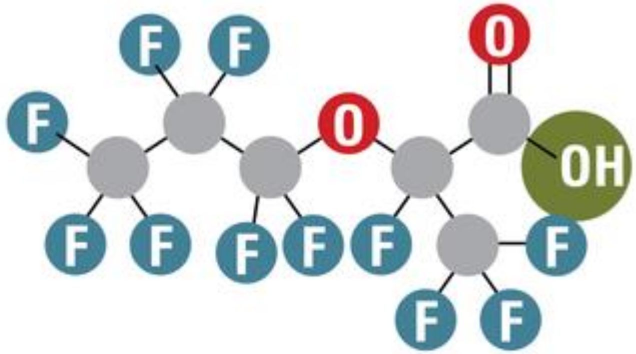
Long-chain/Legacy PFAS
- Includes PFOA, PFOS
- Predominant until 2000

Short-chain
- Includes GenX



PFOA (C8)

Xu et al. 2021



GenX

Xu et al. 2021

PFAS Regulation in NC

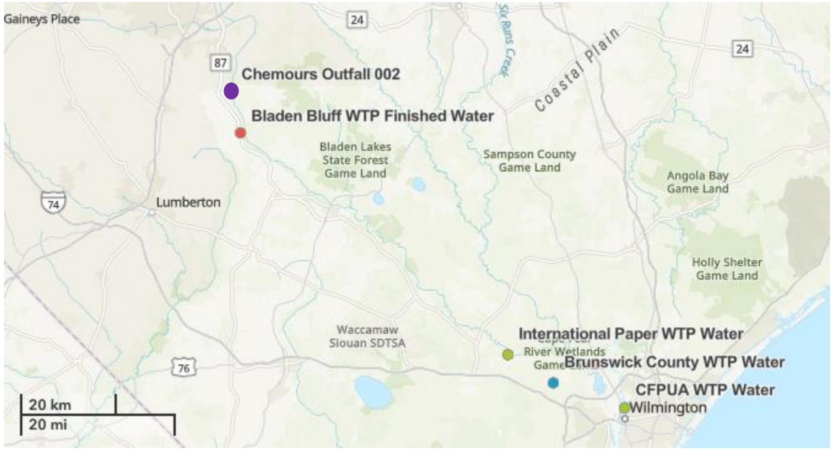
- 2017 NCDEQ established provisional health goal of 140 ppt for GenX
- 2023 EPA proposed MCL NPDWR for 6 PFAS (finalized in 2024, +3 years to meet MCLs)
  - MCLs are 4 ppt for PFOA and PFOS, hazard index for 4 others including GenX

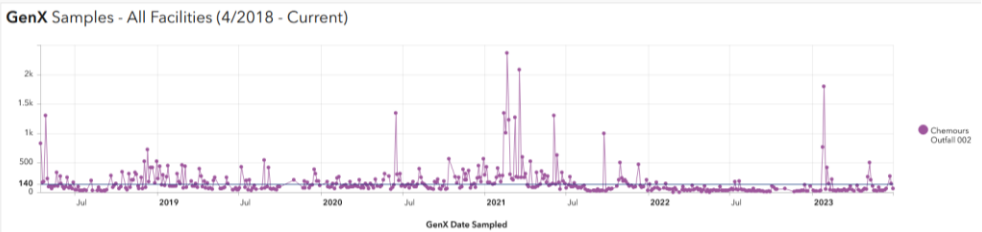
$$\text{Hazard Index} = \left(\frac{[\text{GenX}_{\text{water}}]}{[10\ \text{ppt}]}\right) + \left(\frac{[\text{PFBS}_{\text{water}}]}{[2000\ \text{ppt}]}\right) + \left(\frac{[\text{PFNA}_{\text{water}}]}{[10\ \text{ppt}]}\right) + \left(\frac{[\text{PFHxS}_{\text{water}}]}{[9.0\ \text{ppt}]}\right) \leq 1$$

https://www.epa.gov/system/files/documents/2023-03/How%20do%20I%20calculate%20the%20Hazard%20Index._3.14.23.pdf

Chemours Fayetteville Works Facility

- 1980 production began at Chemours (then DuPont)
- 2009 Chemours replaced PFOA with GenX
- Drinking water in Cape Fear River Basin affected by both groundwater transport and surface water (downstream) and air deposition (upstream)





https://epi.dph.ncdhhs.gov/oee/a_z/pfas.html
https://factor.niehs.nih.gov/2019/3/feature/2-feature-pfas

Challenges of Modeling PFAS in Groundwater

- PFAS properties – high number of chemicals, different adsorptions, transport interactions, hydrophobic/-philic
- Difficult to get data on soil measurements
- Difficult to predict fluxes through saturated, vadose zones
- Complex transport, irregular occurrence patterns
- Accuracy of transport principles at very low levels (ppt)
- Time to model/calibrate fate & transport of multiple chemicals

Simon et al. 2019
Roostaei et al. 2021

Bayesian Network

Coined in 1985

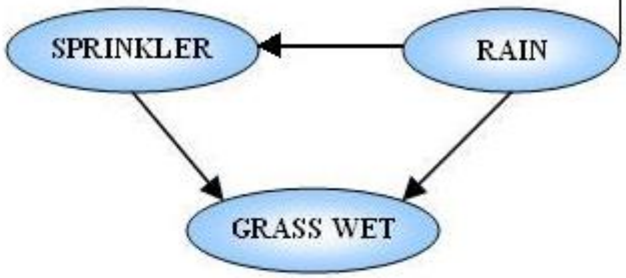
Based on Bayesian statistics (how good is a model given assumed evidence)

Components to a Bayesian network:

- Directed acyclic graph
- Conditional probability distributions

RAIN	SPRINKLER T	SPRINKLER F
F	0.4	0.6
T	0.01	0.99

RAIN T	RAIN F
0.2	0.8



SPRINKLER	RAIN	GRASS WET T	GRASS WET F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Machine-Learned Bayesian Network

Two components to a Bayesian network:

- Directed acyclic graph
- Conditional probability distributions

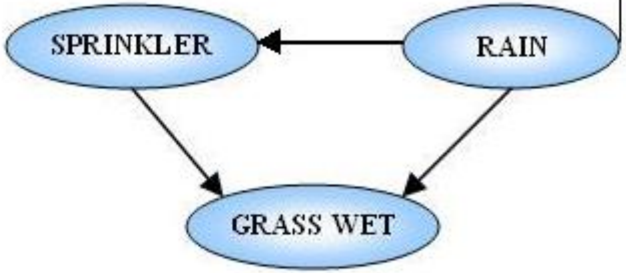
are “learned” from data.

Model complexity vs accuracy optimized by minimizing “min description length” (MDL) score:

$$MDL = \alpha \times [\textit{bits to store model}] + [\textit{bits to store data given model}]$$

RAIN	SPRINKLER T	SPRINKLER F
F	0.4	0.6
T	0.01	0.99

RAIN T	RAIN F
0.2	0.8



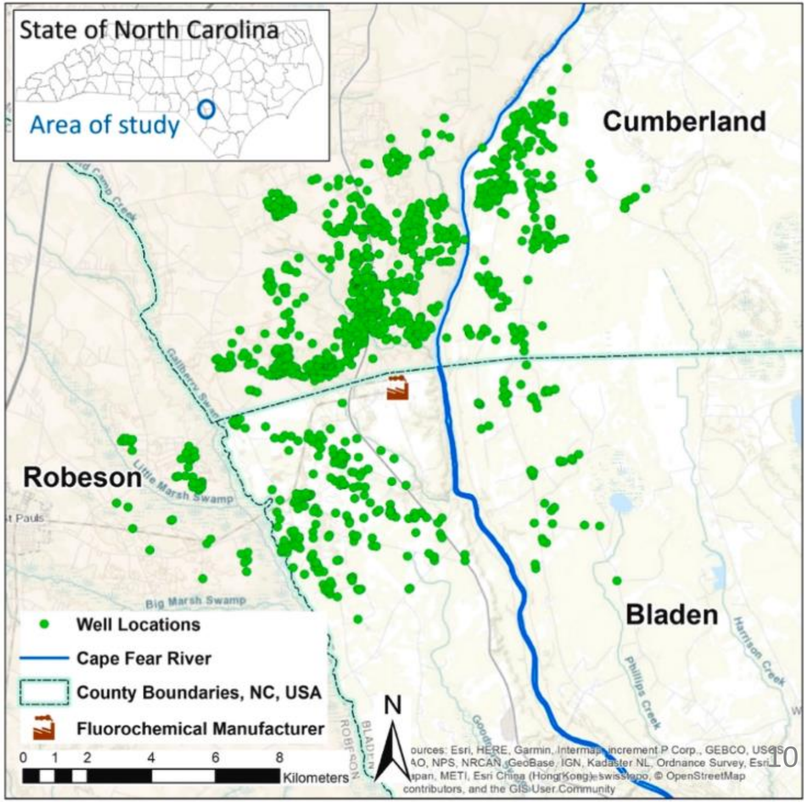
SPRINKLER	RAIN	GRASS WET T	GRASS WET F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Previous Work

NC STATE UNIVERSITY

INDIANA UNIVERSITY

- Dataset curation



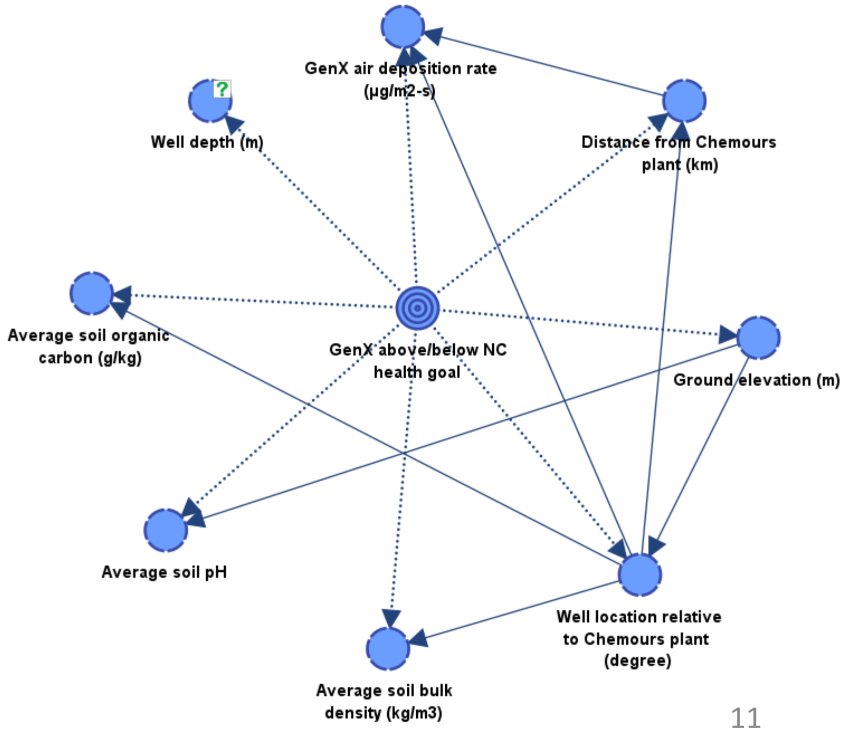
Roostaei et al. 2021

Previous Work

NC STATE UNIVERSITY

INDIANA UNIVERSITY

- Dataset curation
- Train MLBN with same inputs as mechanistic GW flow model
- Model validation
- Spatial risk maps



Roostaei et al. 2021

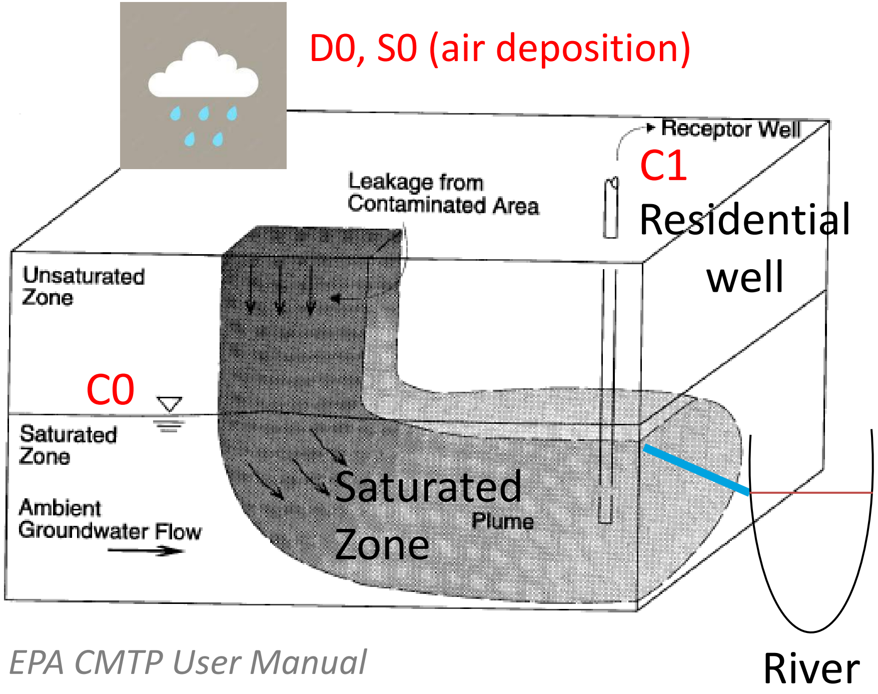
Models compared

Goal: compare predictive performance of models with low- to high-mechanistic modeler effort:

Models compared labeled by input effort:	GW model inputs	GW flow model outputs	FT model output conc.
Low-effort BN (Roostaei et al. 2021)	✓		
Medium-effort BN	✓	✓	
High-effort BN	✓	✓	✓
Mechanistic FT model	✓	✓	✓
Mechanistic modeling effort required:		**1 mo.**	**1 mo.**

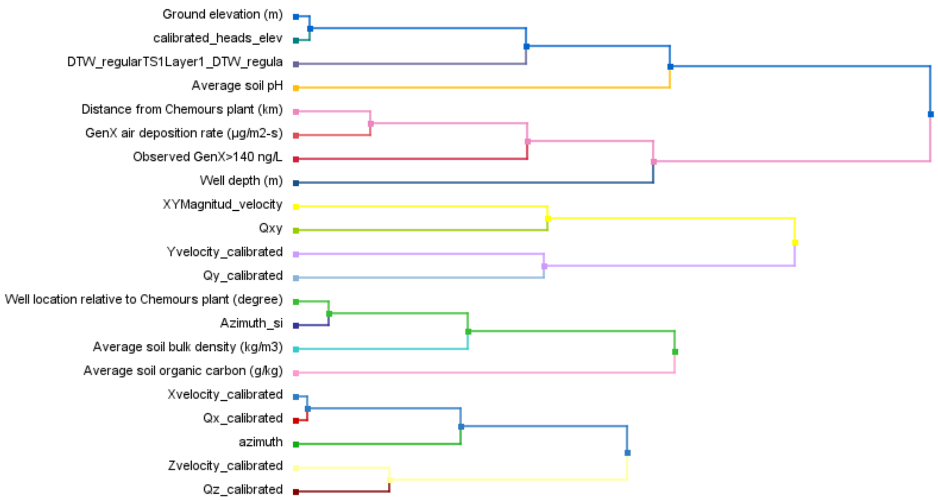
Mechanistic Model

- Modflow 2005 & MT3D engines in Groundwater Vistas Model parameterized using data from NC DEQ, Chemours reports
- Model C_1 (concentration at wells) using only parameters used in MLBN:
  - steady state/long-term transport (saturated zone)
  - empirical relationship between concentration at water table C_0 and modeled air deposition rates



EPA CMTP User Manual

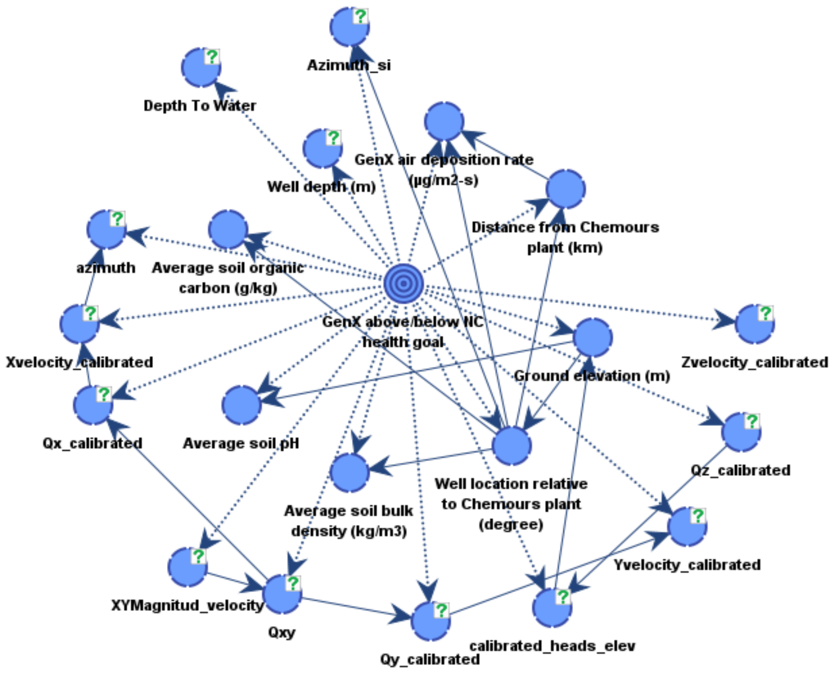
MLBN Model Development



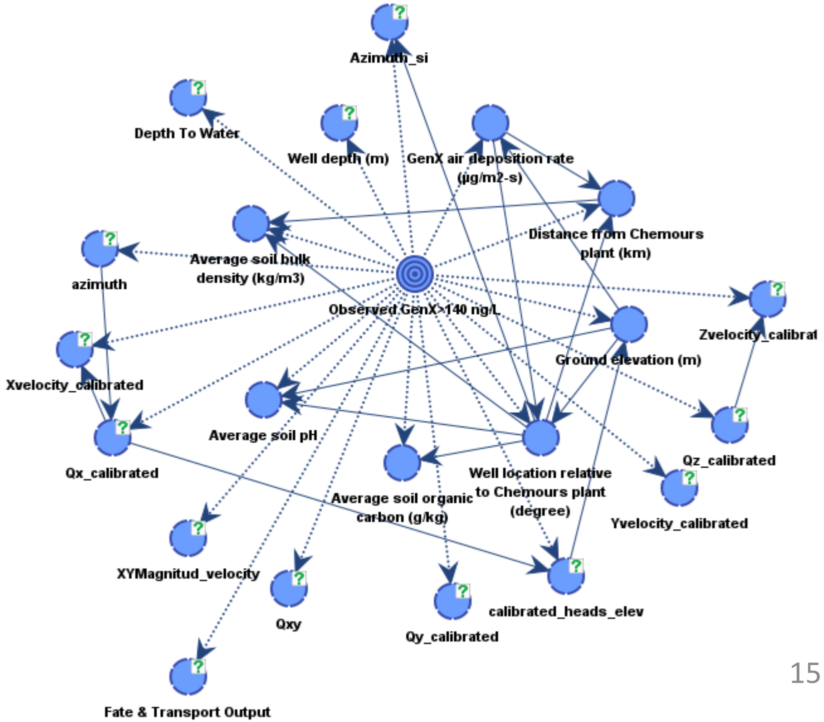
- Imputation of missing data
- Discretize variables
- Supervised learning
- Structural coefficient analysis
- Adjustment of included variables, discretization
- Supervised learning
- Cross-validation analysis

MLBN Model Development

Medium-effort model

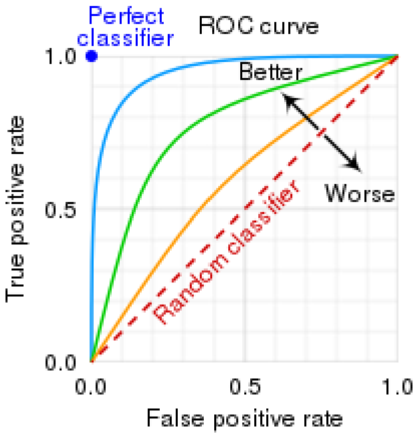


High-effort model



Model Performance Comparisons

- Area under receiver operating characteristic curve (AU-ROC)

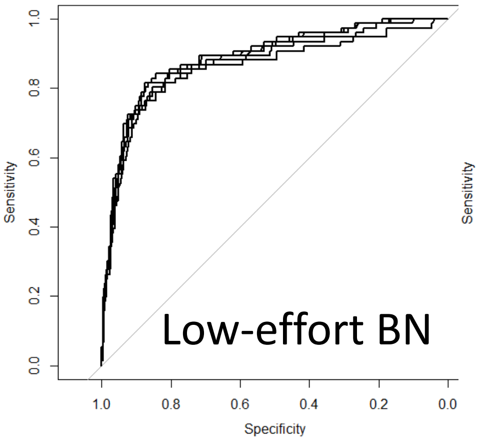


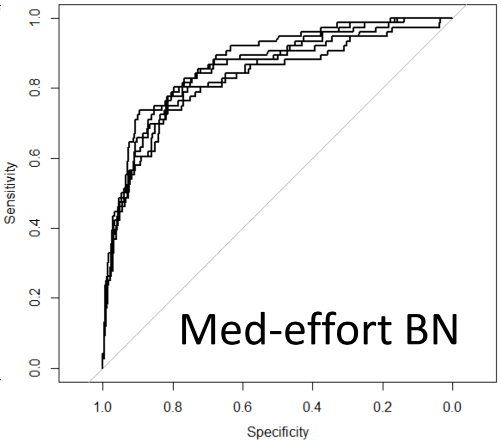
Model Performance Comparison (AU-ROC)

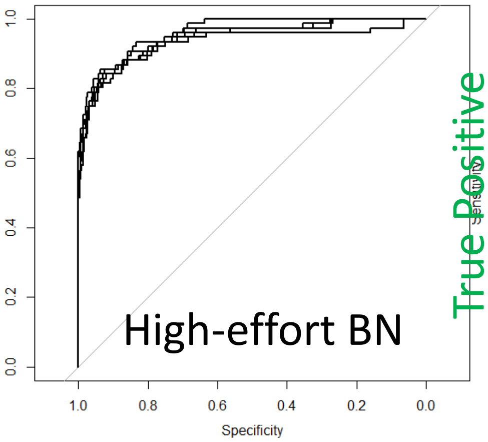
5-fold cross-validation (n=5), wells with depths (n=424):

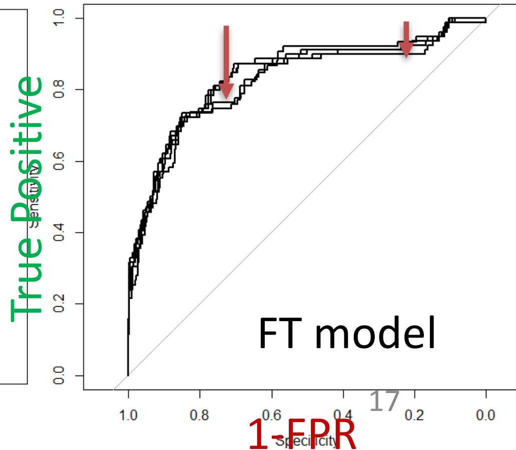
5 CV sets' Average ± StdDev (final model)

Low-effort BN	Med-effort BN	High-effort BN	FT model
0.854±0.010 (0.905)	0.845±0.0123 (0.908)	0.868±0.0061 (0.916)	0.832±0.0055 (0.803)





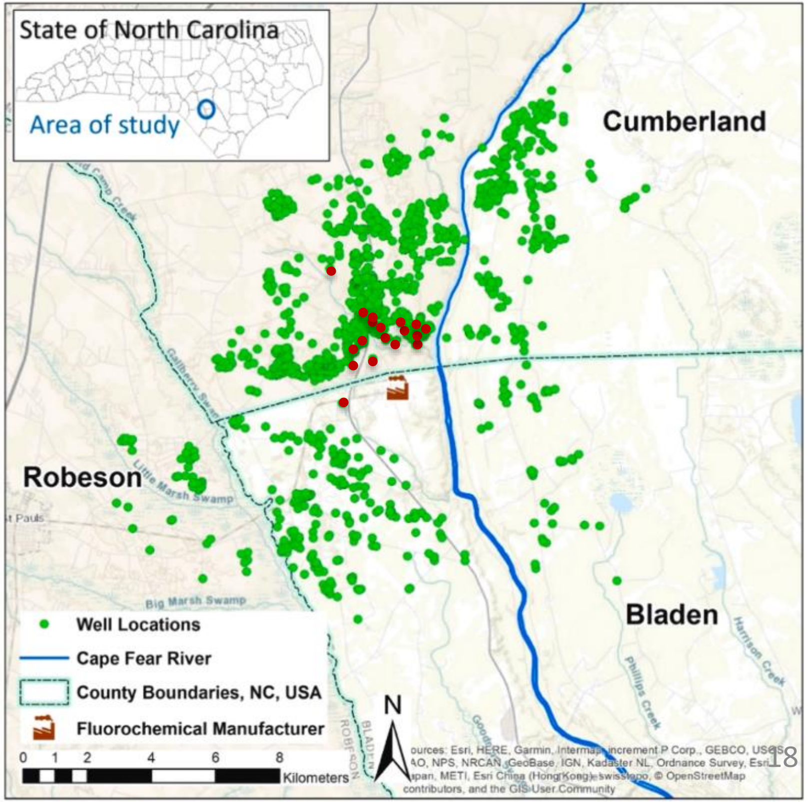




Low detection rate vs low identification accuracy

- If only 10 % of wells are contaminated, a model can have 90 % accuracy and never detect any positives (low power)

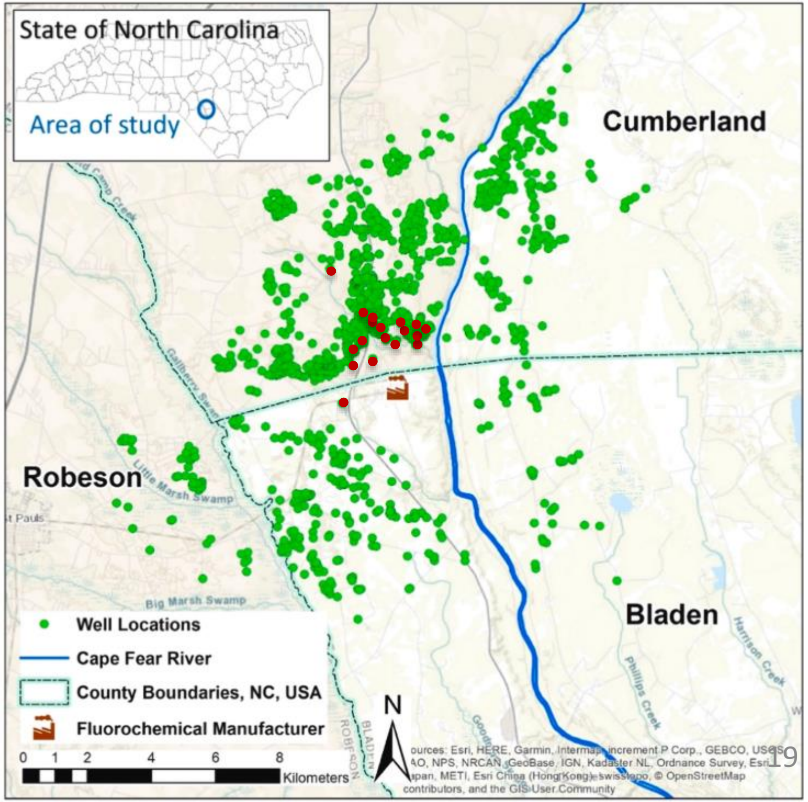
	Low-effort BN (DT=50%)	FT Model (DT=140ppt)
TPR	58-61%	22-32%
TNR	91-92%	98-100%
FPR	8-9%	0-2%
FNR	39-42%	65-75%
Precision	59-62%	80-96%
Accuracy	85-86%	85-87%



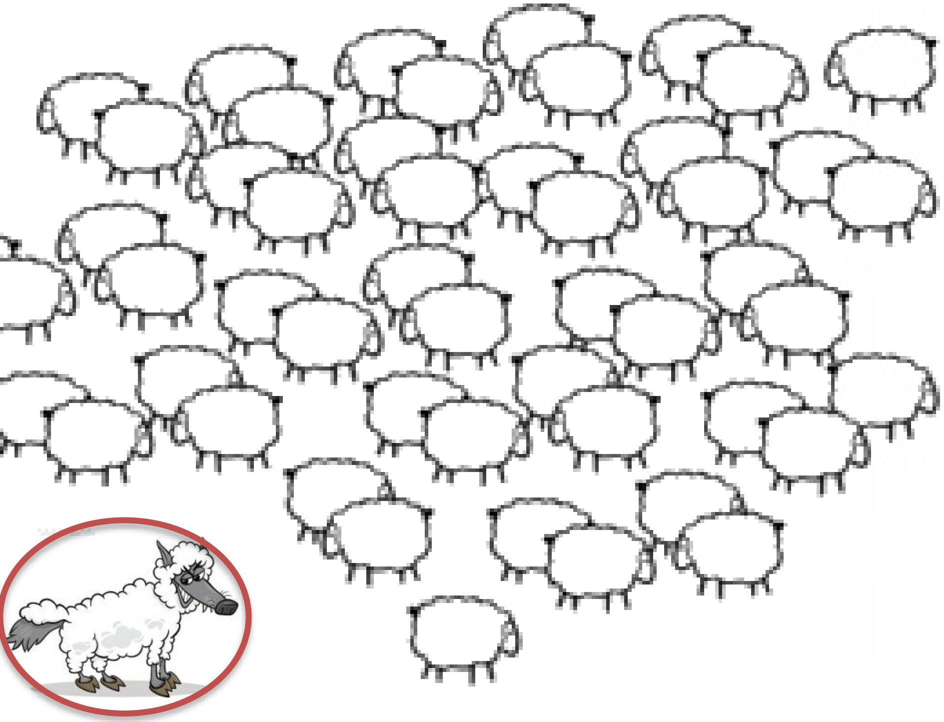
Low detection rate vs low identification accuracy

- CV (non-stratified k-folds) made FT model TPR and FNR worse

	Low-effort BN (DT=50%)	FT Model (DT=140ppt)
TPR	58-61% (68%)	22-32% (46%)
TNR	91-92% (92%)	98-100% (95%)
FPR	8-9% (7%)	0-2% (5%)
FNR	39-42% (32%)	65-75% (54%)
Precision	59-62% (67%)	80-96% (67%)
Accuracy	85-86% (87%)	85-87% (86%)



Model Performance Comparisons



- F1 score
  - Balances true positive rate and positive predictive value
- F2 score
  - Better score reduces false negatives (Type II Error)

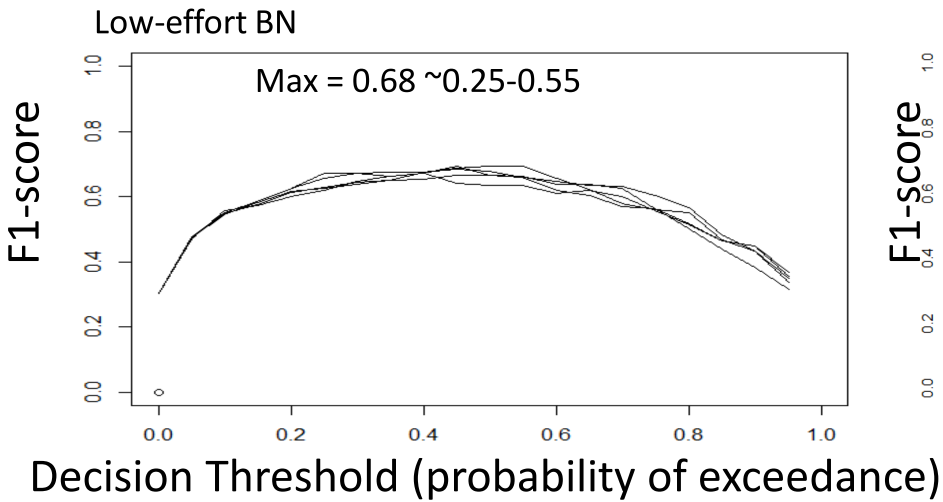
Low wolf detection rate?

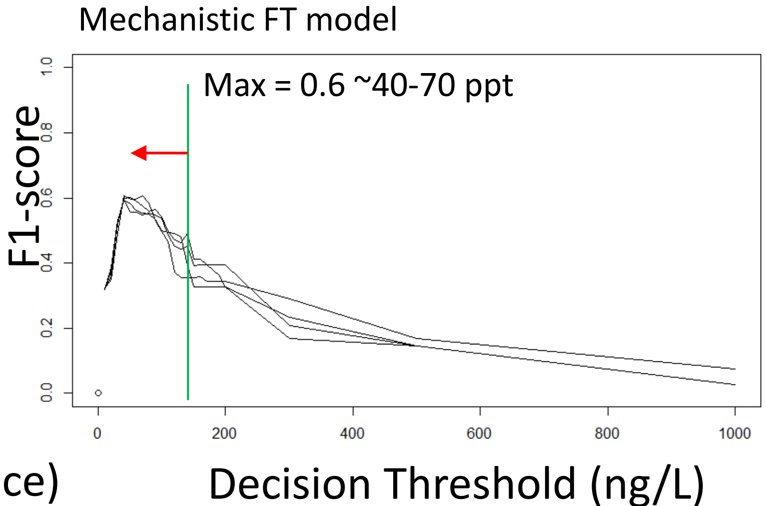
Low wolf identification accuracy?

Which is worse in environmental engineering?

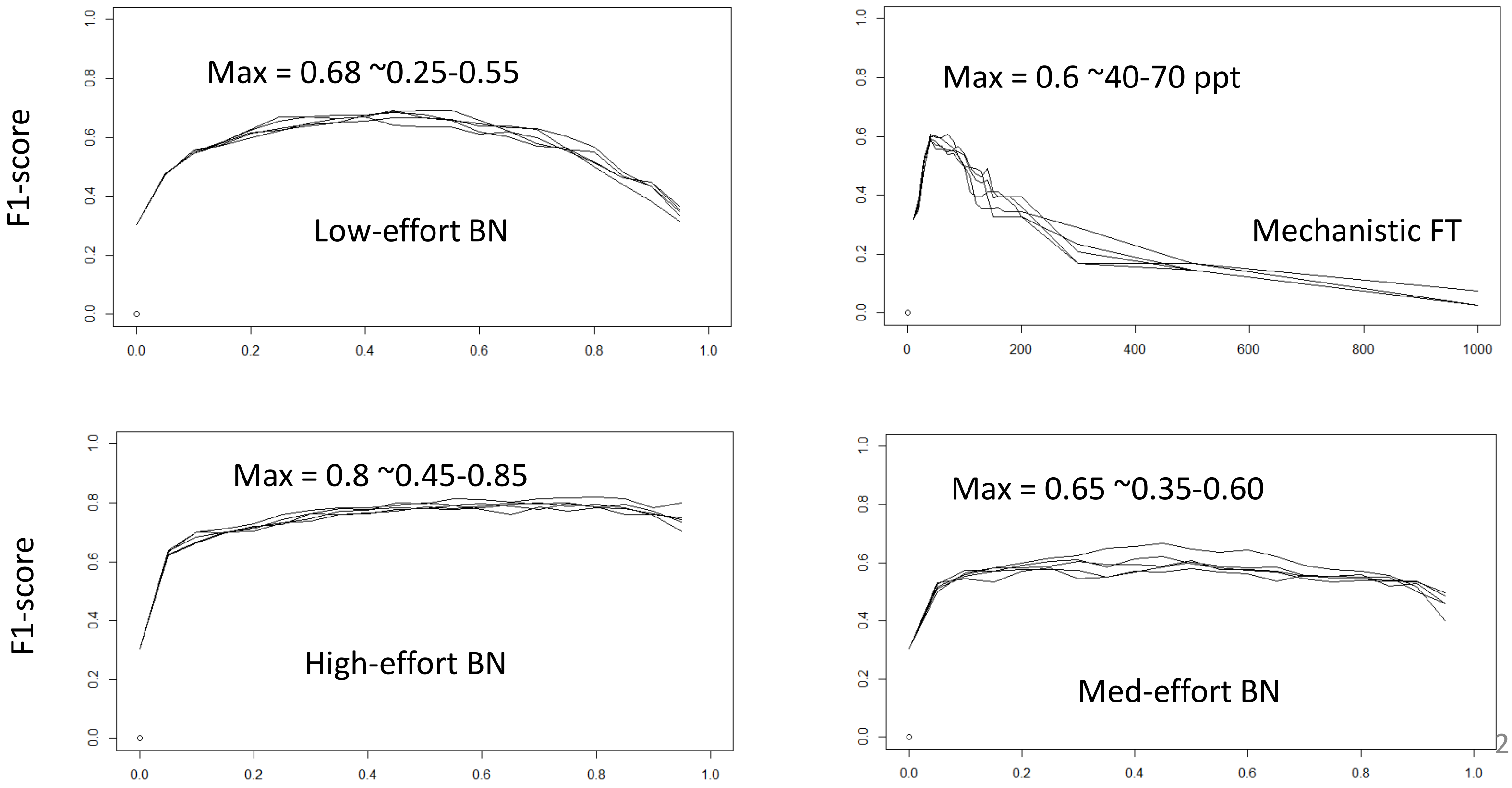
Model Performance Comparison (F1-score)

- BN models & Mechanistic FT model have similar accuracy (86%)
- Mechanistic FT model had high PPV but low TPR (tradeoff in model performance/risk of FP and FN based on decision threshold (posterior probability for BN, concentration for mechanistic–baked into mechanistic model calibration)

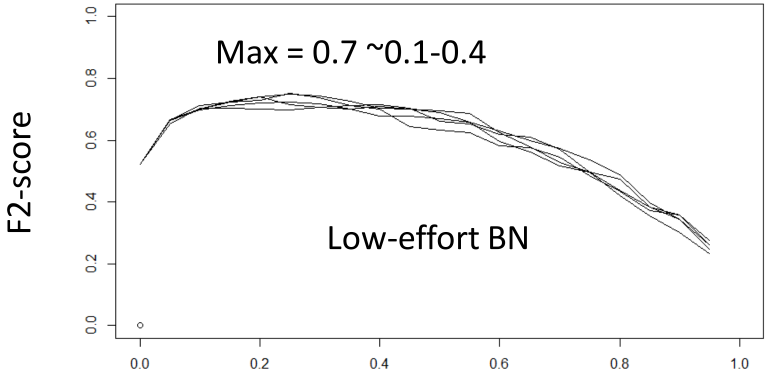


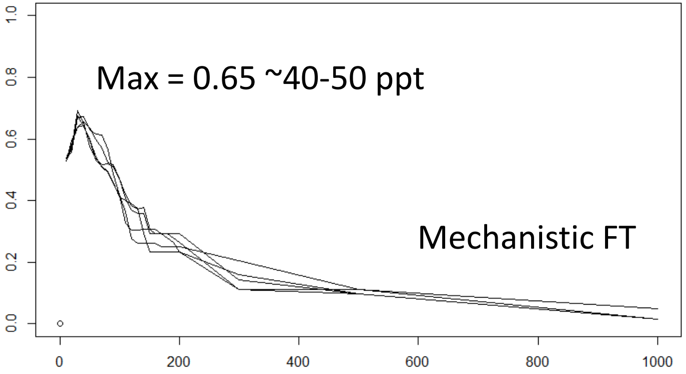


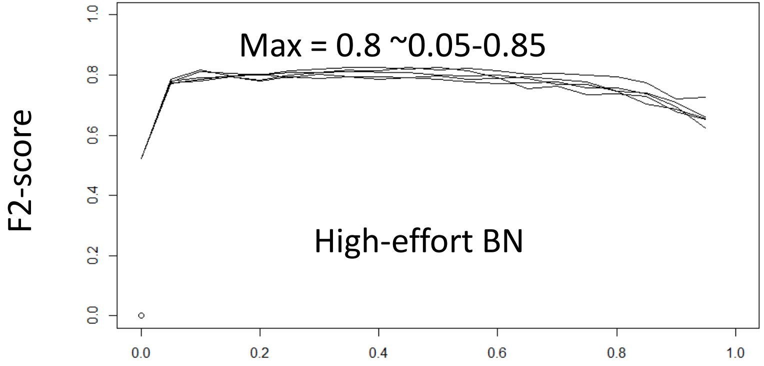
Model Performance Comparison (F1 scores)

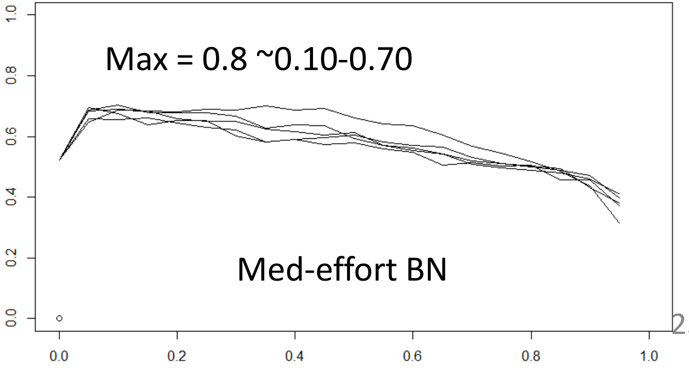


Model Performance Comparison (F2 scores)

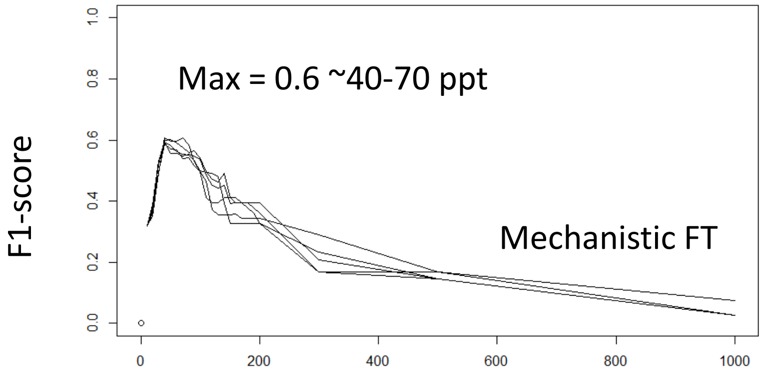


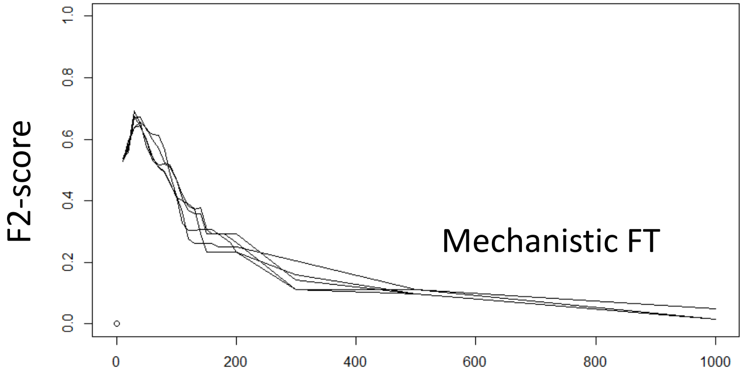


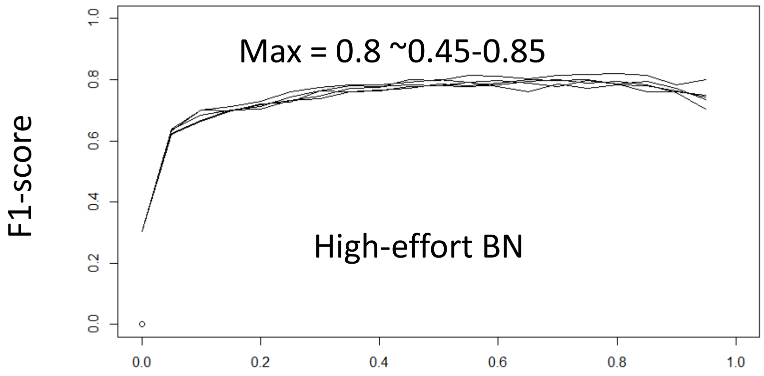


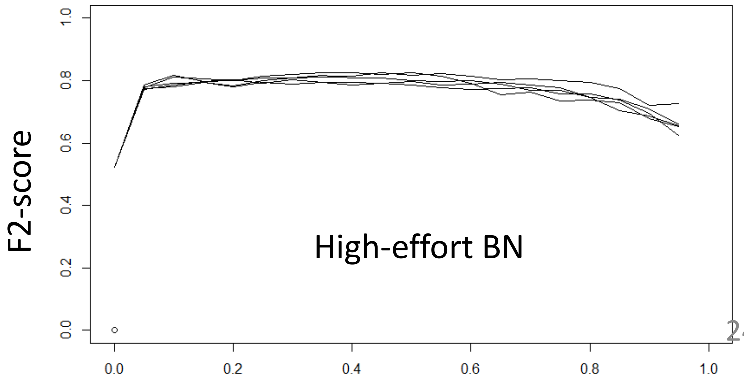


Model Performance Comparison (F2 scores)









Takeaways, Limitations

- MLBN model and mechanistic FT models have similar accuracy and AU-ROC metrics; integrated MLBN models make small gains in predictive power

- MLBNs appear more robust to different decision thresholds (for mechanistic model, decision threshold is baked into calibration)

- Mechanistic model is more susceptible to imbalanced datasets in CV

- Recalibrating mechanistic model to give more weight to high concentrations (improve TPR) would take a lot of additional effort, but incorporating this knowledge to improve performance in the hybrid BN is trivial

Importance, Applications

- In this particular case study area, low-effort MLBN by itself performs as well as more time-consuming mechanistic model (at the selected level of sophistication) – MLBN very promising for PFAS modeling

- Improving risk prediction and awareness, particularly for vulnerable communities/ private well owners is timely

Acknowledgements

Jacqueline MacDonald Gibson, PhD
Head of the Department of Civil, Construction, and Environmental Engineering, Professor

Modeling Team

NCSU

Hana C Long
Postdoc

Krishna Ganta
Student

RTI:
Rohit Warrier
Riley Mulhern
Jennifer Hoponick Redmon
Ted Lillys

QUESTIONS?