

CAUSAL MODELING OF SOIL PROCESSES FOR IMPROVED GENERALIZATION

Somya Sharma^a, Swati Sharma^{*,b}, Andy Neal^c, Sara Malvar^d, Eduardo Rodrigues^d, John Crawford^e, Emre Kiciman^b, Ranveer Chandra^b

^a University of Minnesota Twin Cities

^b Microsoft Research, Redmond

^c Rothamsted Research, UK

^d Microsoft Research, Brazil

^e University of Glasgow

*Corresponding author: swatisharma@microsoft.com

GOAL: SOIL CARBON PREDICTION

Causal graph
learning



Learning causal
relations among
different soil
processes.

Generalizability



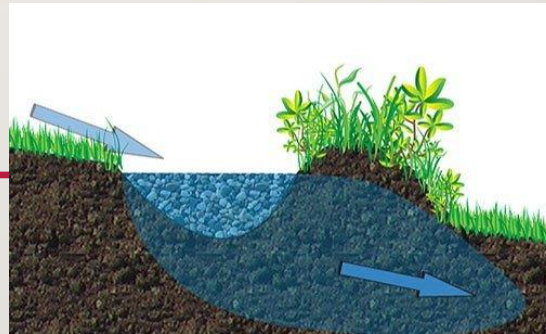
Improved soil carbon
prediction across
different scenarios

IMPORTANCE OF SOIL CARBON

Soil health



Co-benefits



Improve
agricultural
productivity

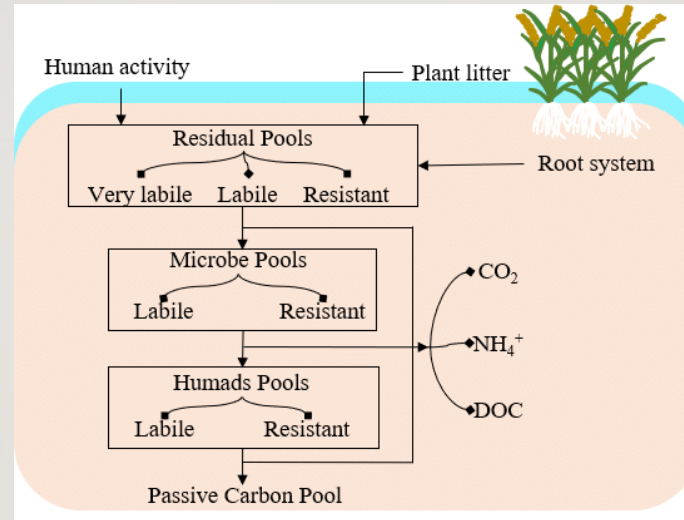


Enhance
climate
resilience

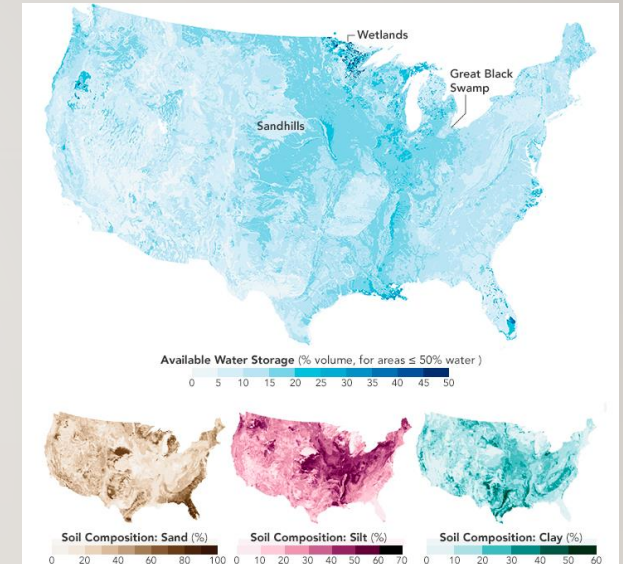
CHALLENGES



**Complex “living”
medium**

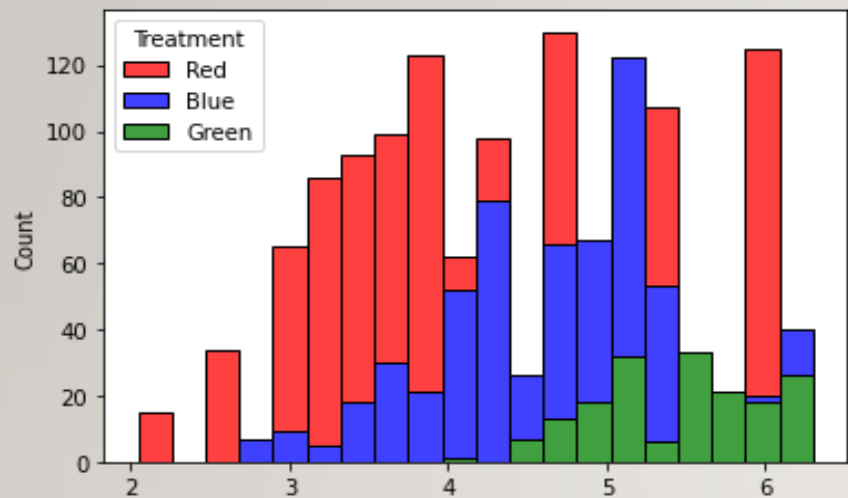


No universal model

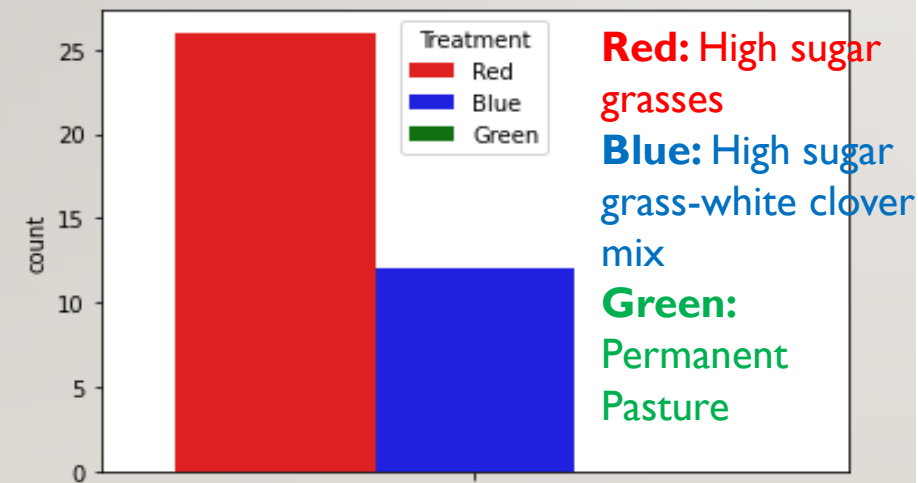


Different environments

SOIL CARBON PREDICTION FOR NORTHWYKE DATA*



Soil organic carbon distribution



Plough event distribution

Train: **Red, Blue** treatments
Test: **Green** treatments

*<https://www.rothamsted.ac.uk/north-wyke-farm-platform>

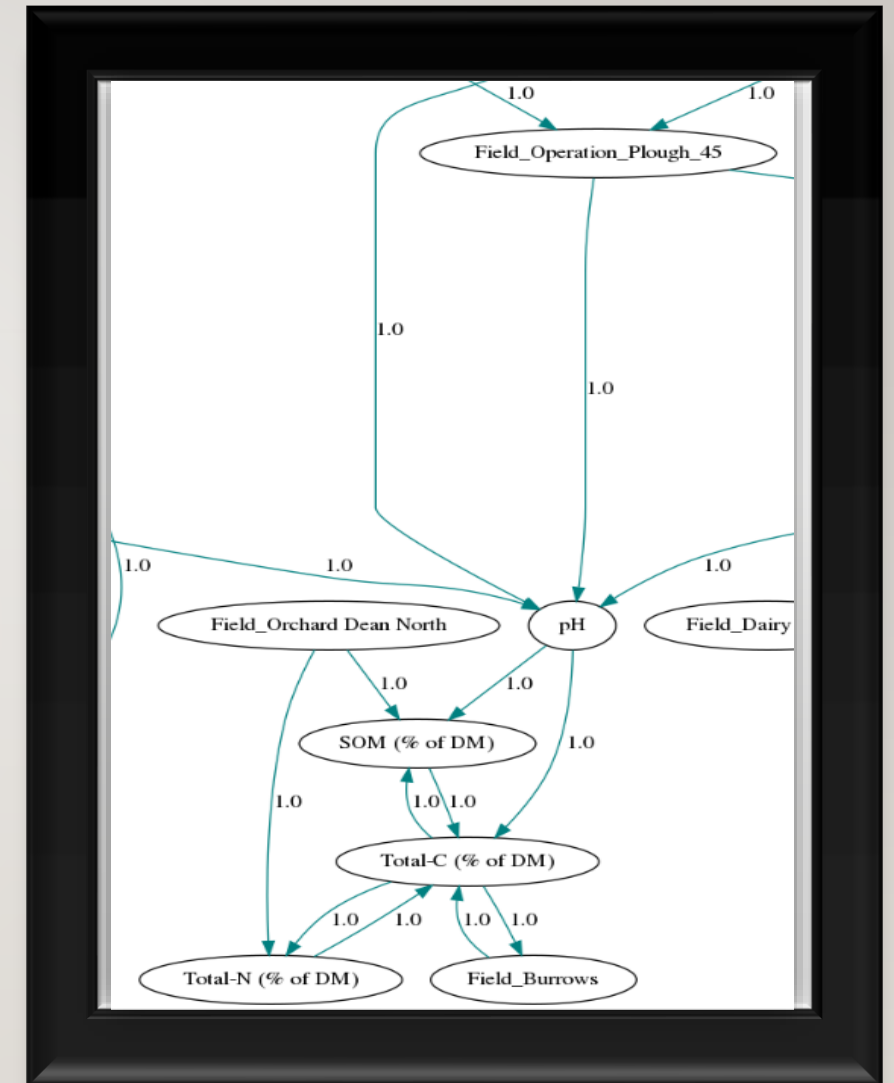
	Model	Causal	MSE	MAE
PC + GraphSAGE			0.002	1.0116
GES + GraphSAGE			0.0248	0.9193
GIES + GraphSAGE		Yes	0.7803	0.7904
PC + ECC MPNN		Yes	0.2816	0.5258
GES + ECC MPNN		Yes	0.2864	0.5302
GIES + ECC MPNN		Yes	0.2951	0.5385
Random Edges + GraphSAGE		No		
XGBoost		No		
MLP		No		
Random Forest		No	2.7996	1.6263

Causal ML approaches generalize better than ML approaches

ML approaches do not work well on OOD test data

CAUSAL GRAPHS DISCOVER SOME WELL-KNOWN LINKS

- ❑ Relationship between carbon and nitrogen in soil
- ❑ Tilling/ploughing changes pH of the soil that affects soil organic carbon and total carbon, nitrogen



THANK YOU

*Corresponding author: swatisharma@microsoft.com