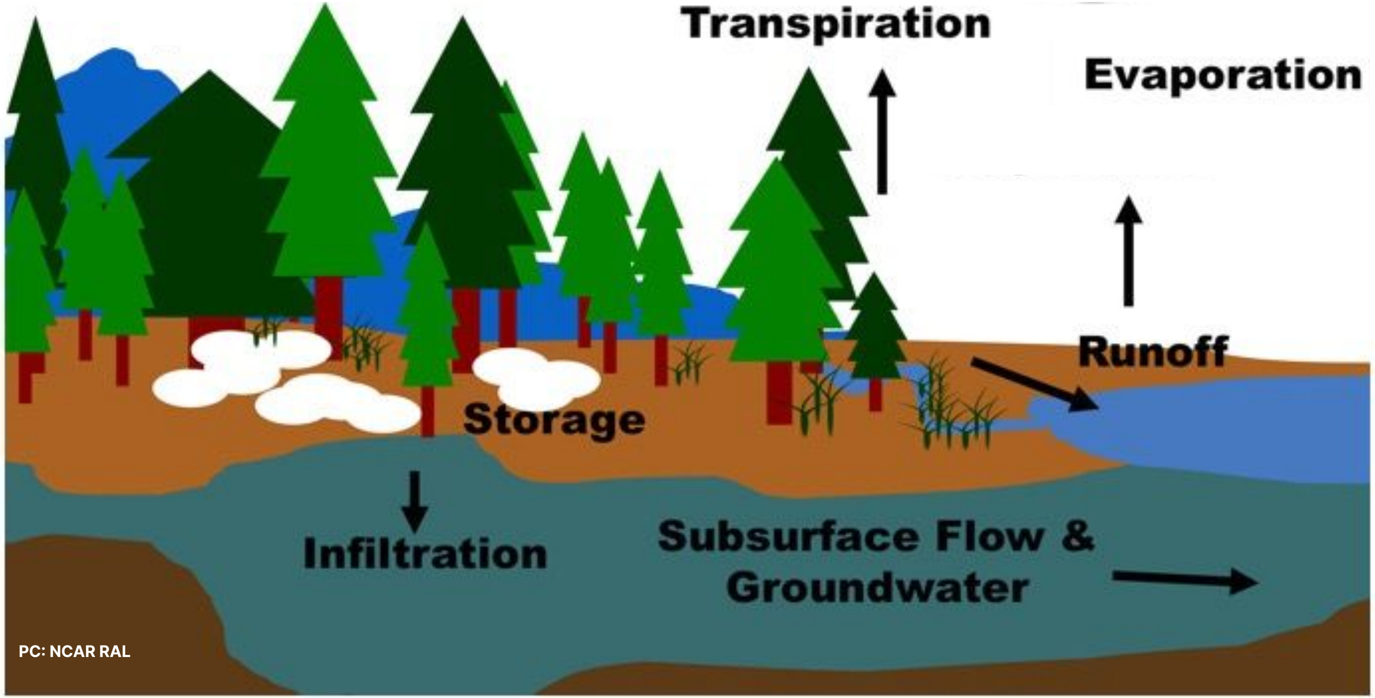
Towards physics-based machine learning for land surface modeling:

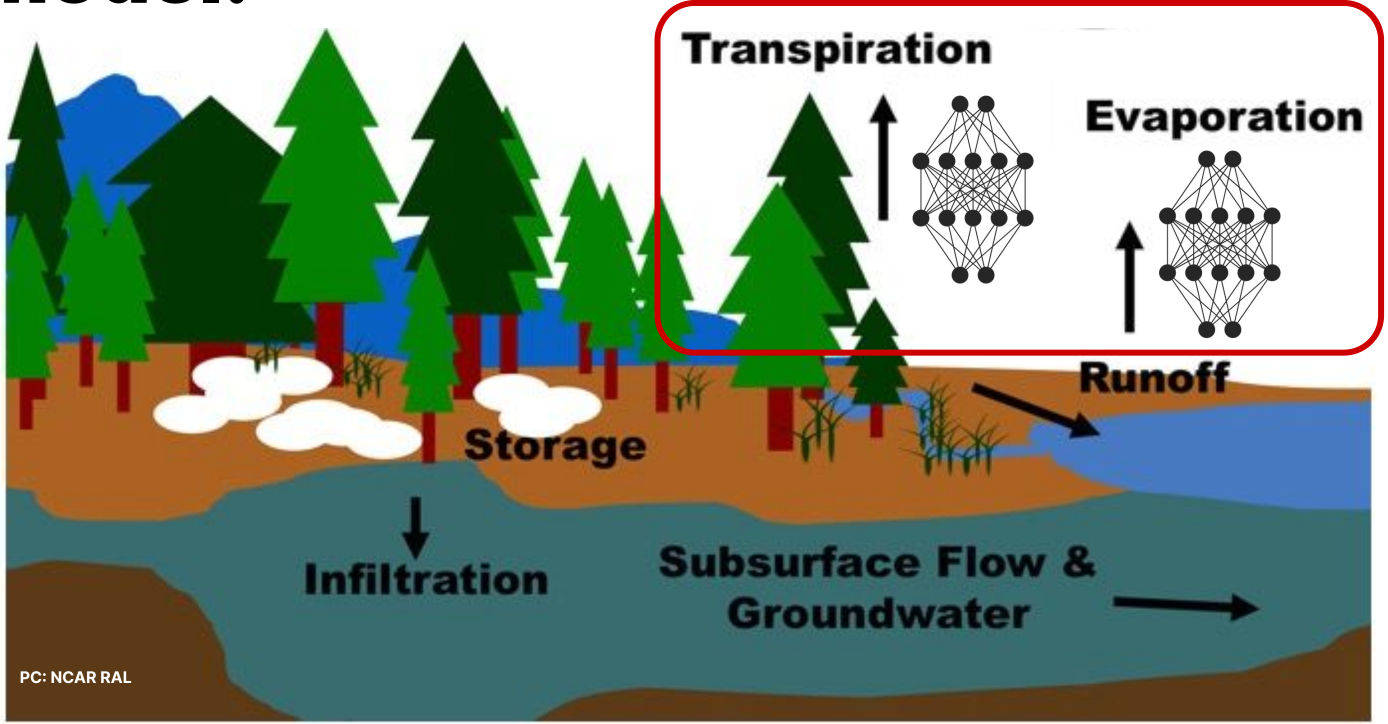
The case of land-atmosphere interactions.

Andrew Bennett
andrbenn@email.arizona.edu
March 29, 2022



Transpiration
Evaporation
Runoff
Storage
Infiltration
Subsurface Flow &
Groundwater
PC: NCAR RAL

The main idea:
Put the neural network *inside* of the hydrologic model!



We reframe things in terms of energy fluxes

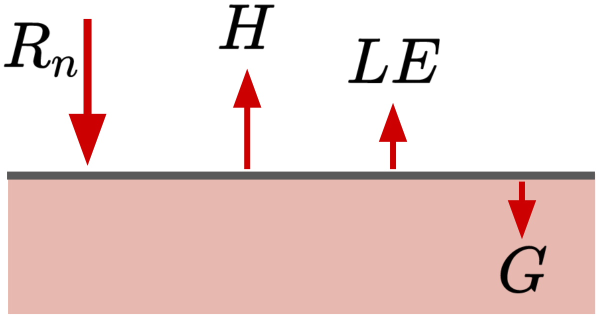
Net radiation

Sensible heat
"Conductive transport"

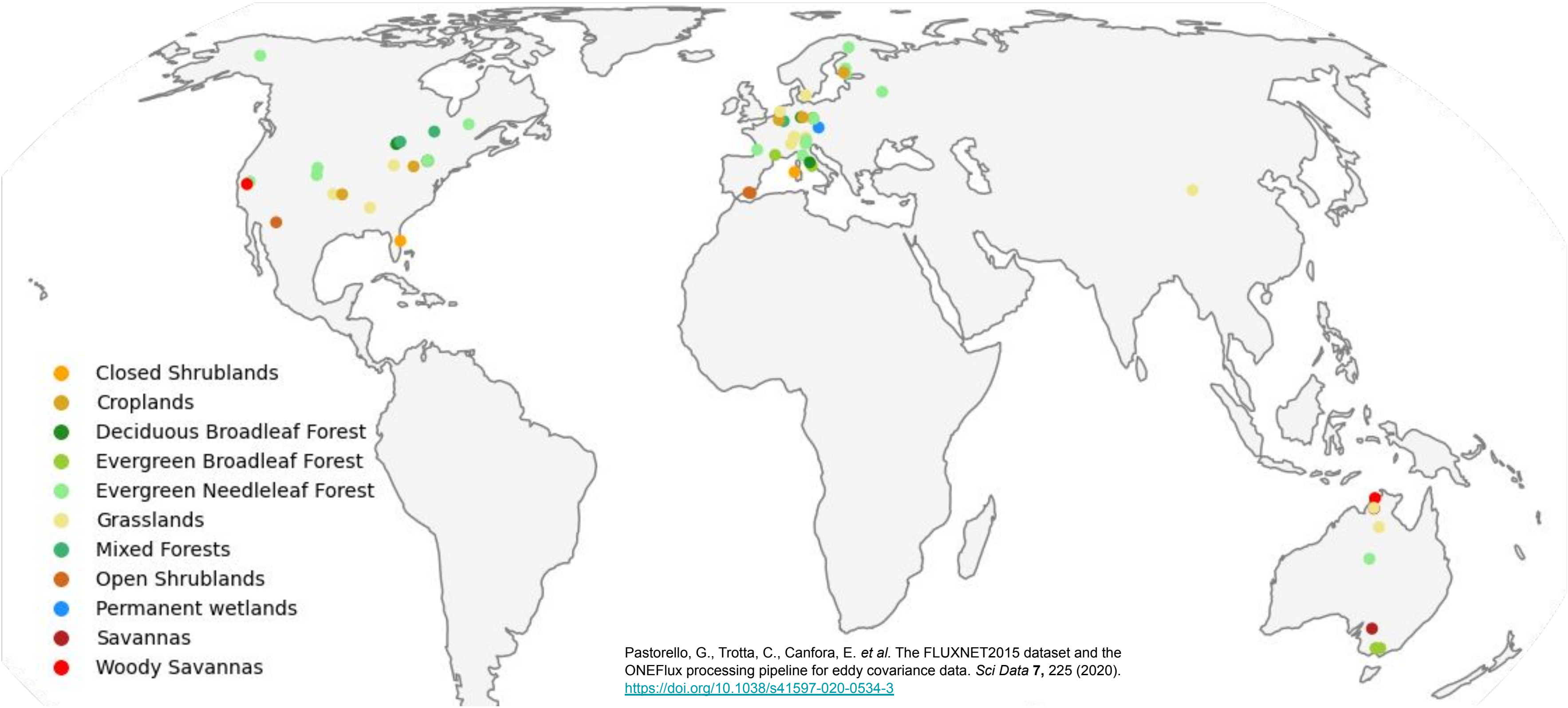
Latent heat
"Evaporative transport"

Ground heat flux

$$R_n = H + LE + G$$

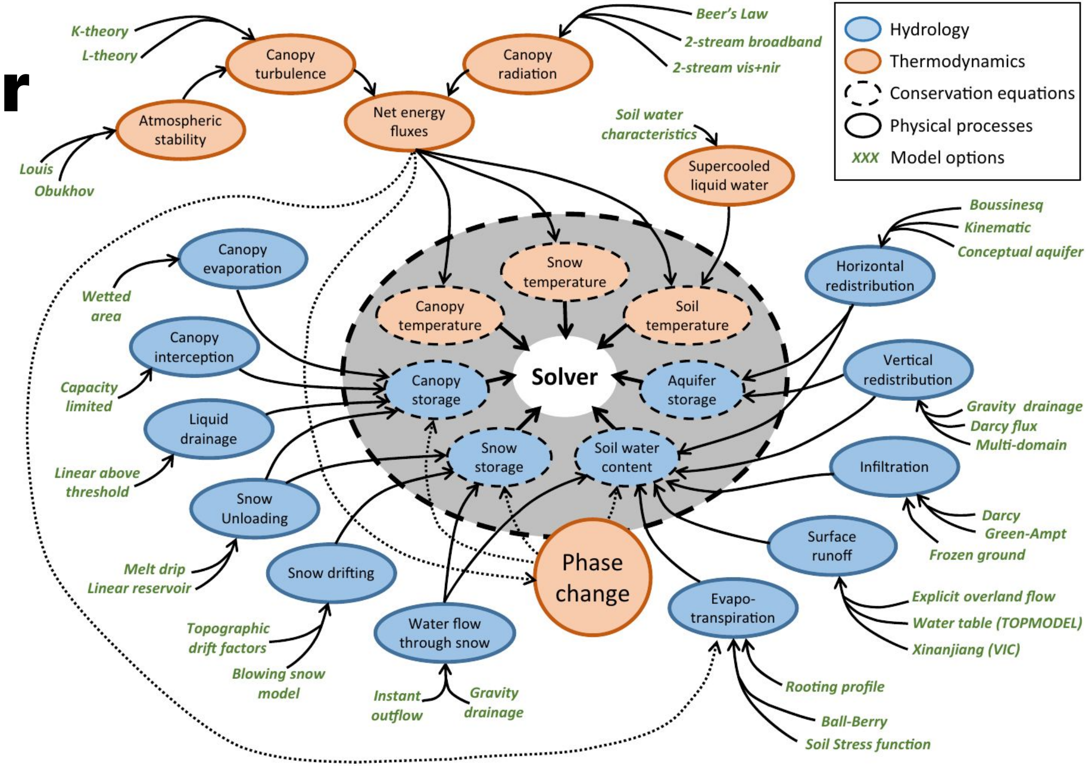


We gathered data from 60 FluxNet sites, totalling over 500 site-years of half-hourly data



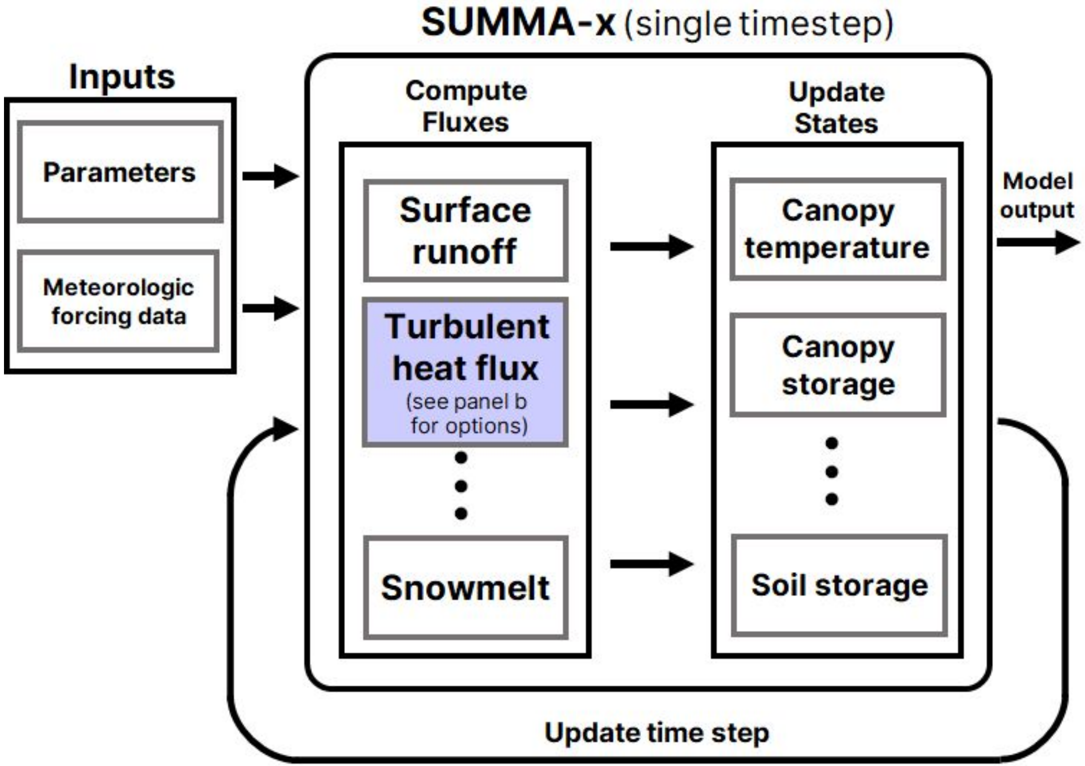
Pastorello, G., Trotta, C., Canfora, E. *et al.* The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci Data* **7,** 225 (2020). https://doi.org/10.1038/s41597-020-0534-3

We used the SUMMA hydrologic modeling framework for all of our configurations

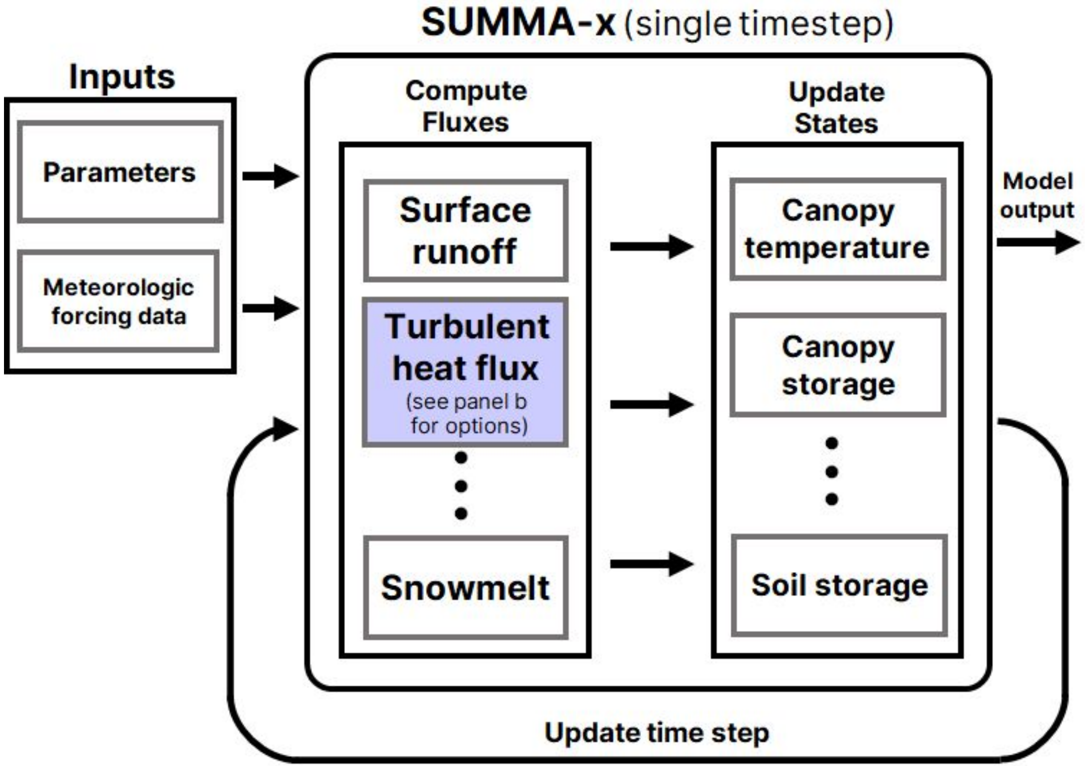


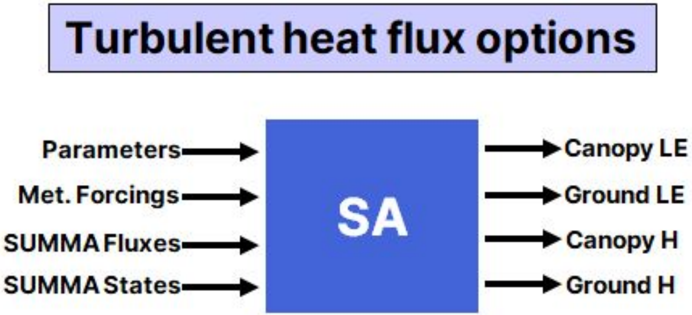
Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., et al. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling concept. Water Resources Research, 51, 1–17. https://doi.org/10.1002/2015WR017200.A

Model configurations

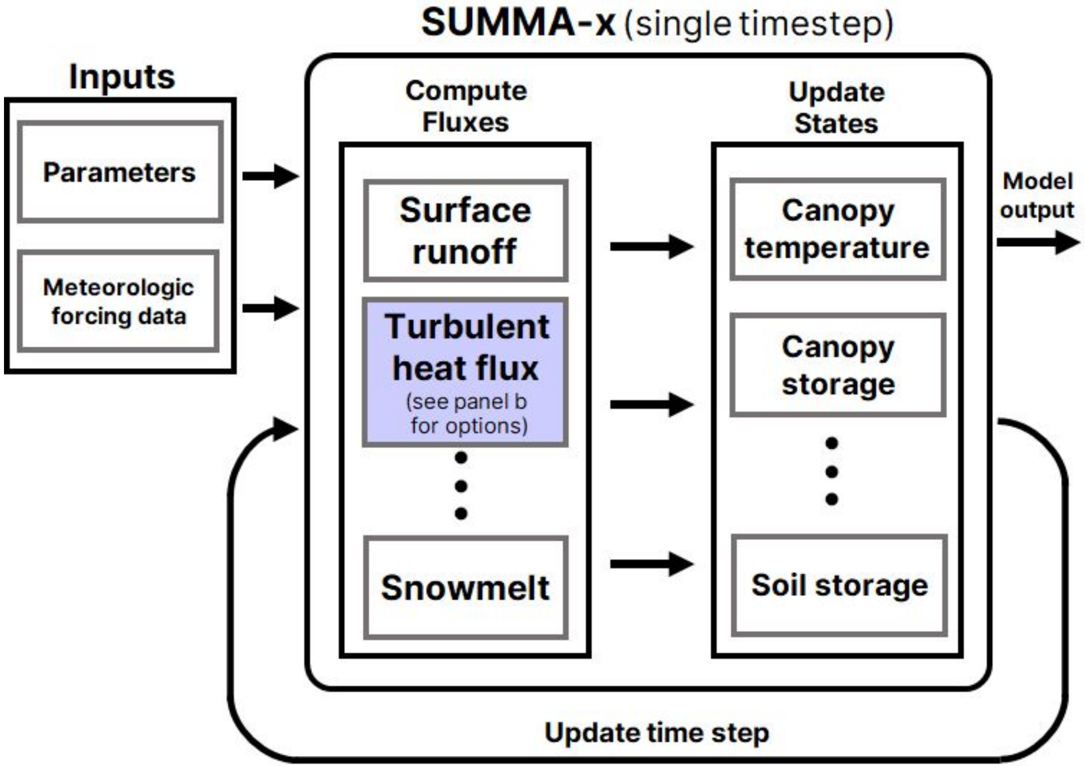


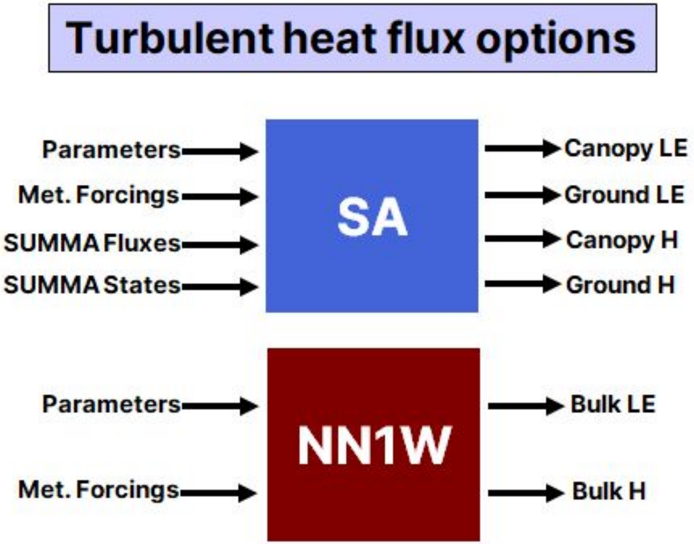
Model configurations



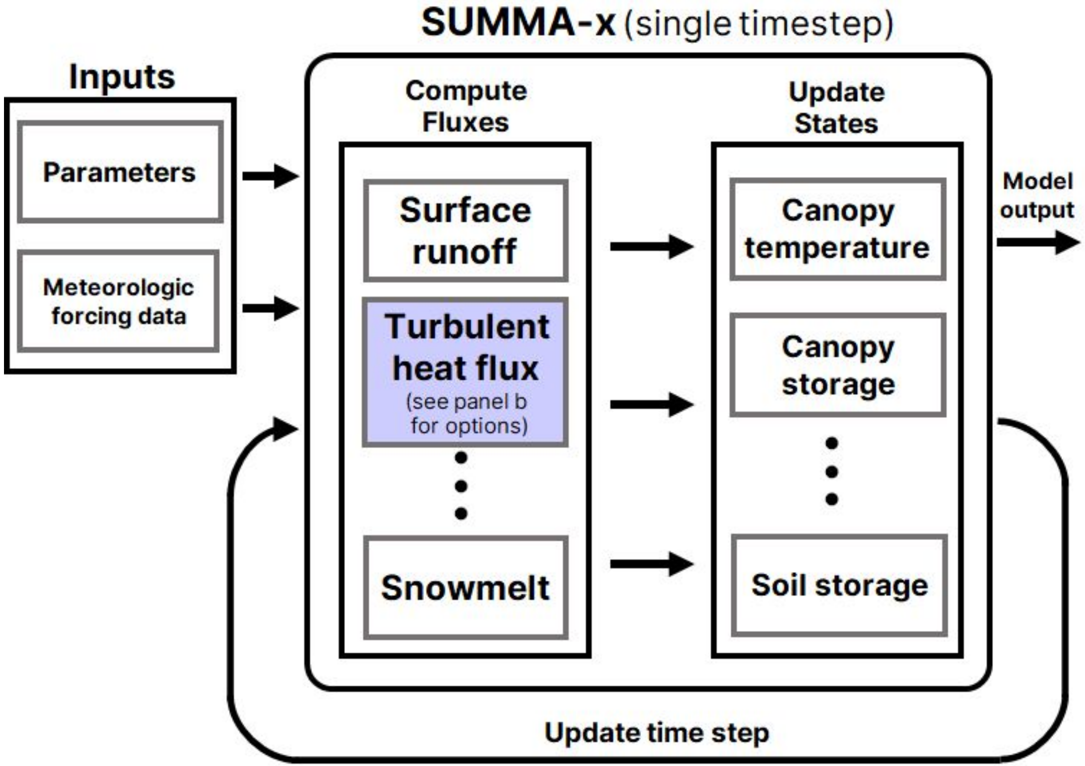


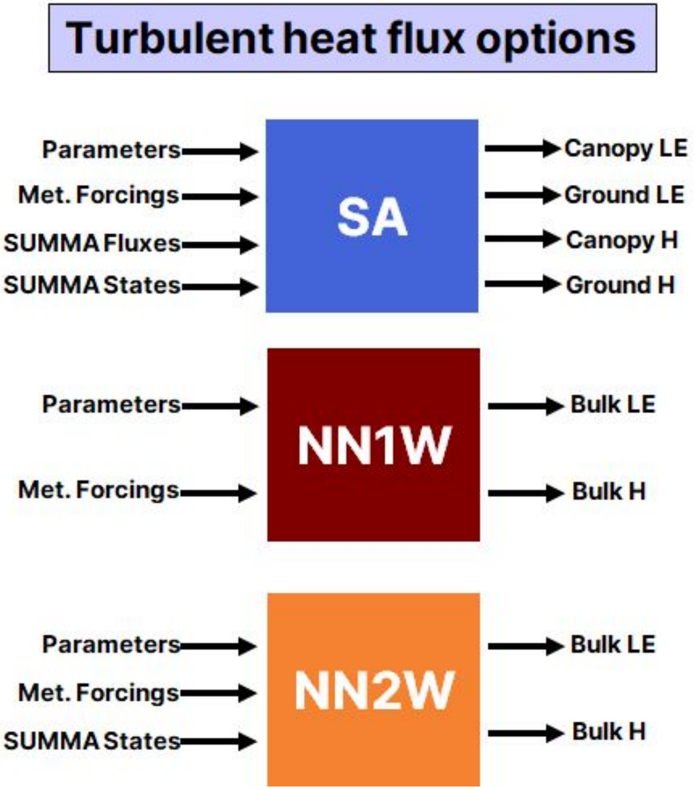
Model configurations



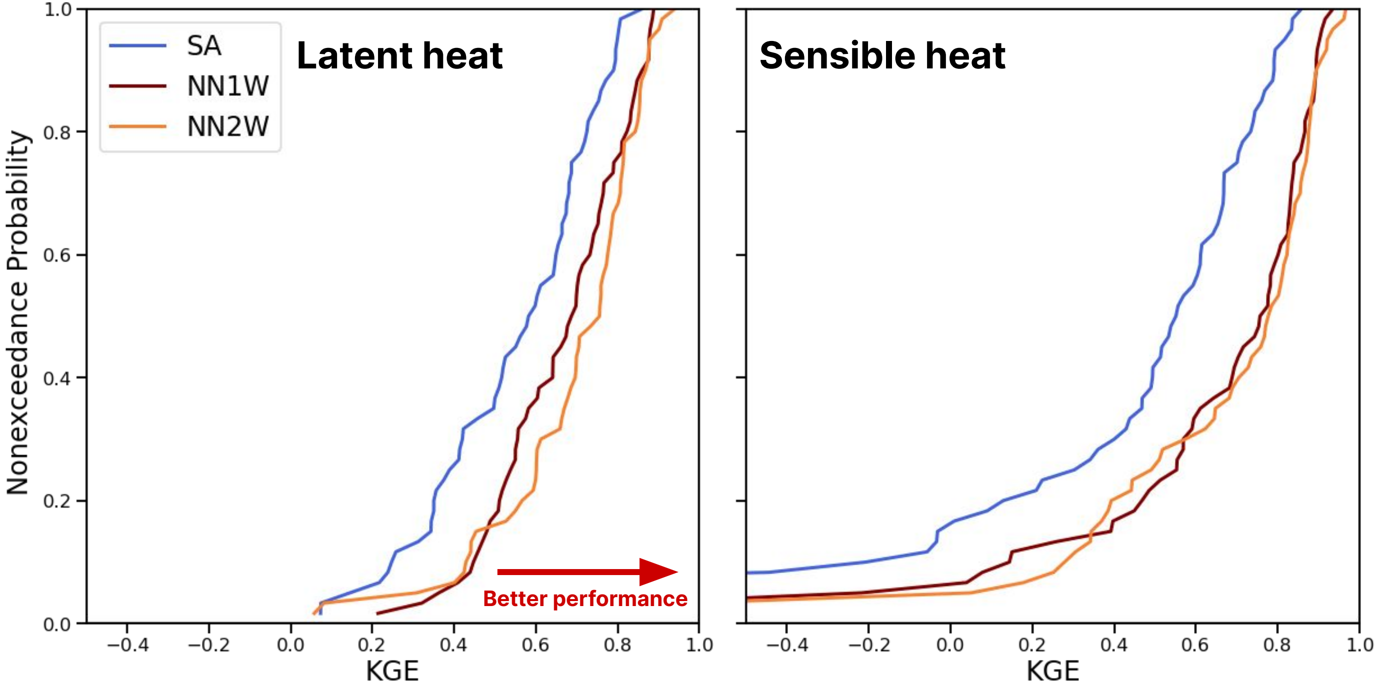


Model configurations

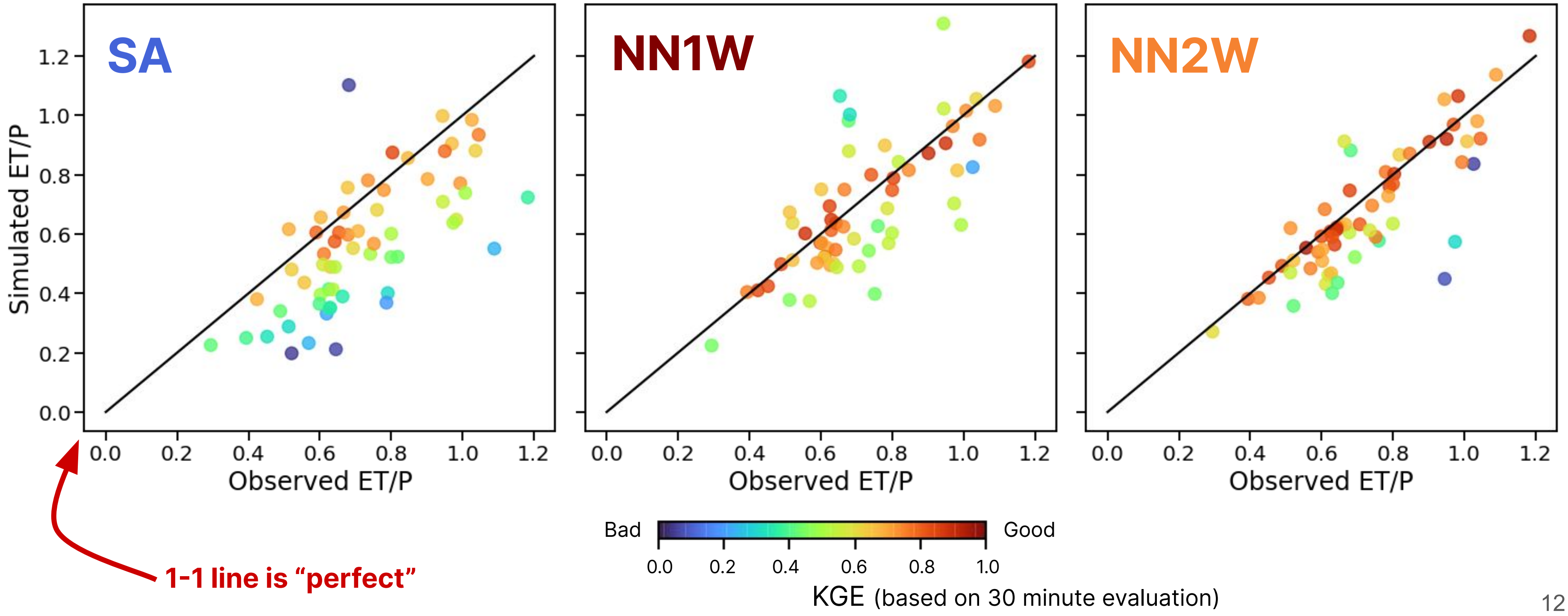




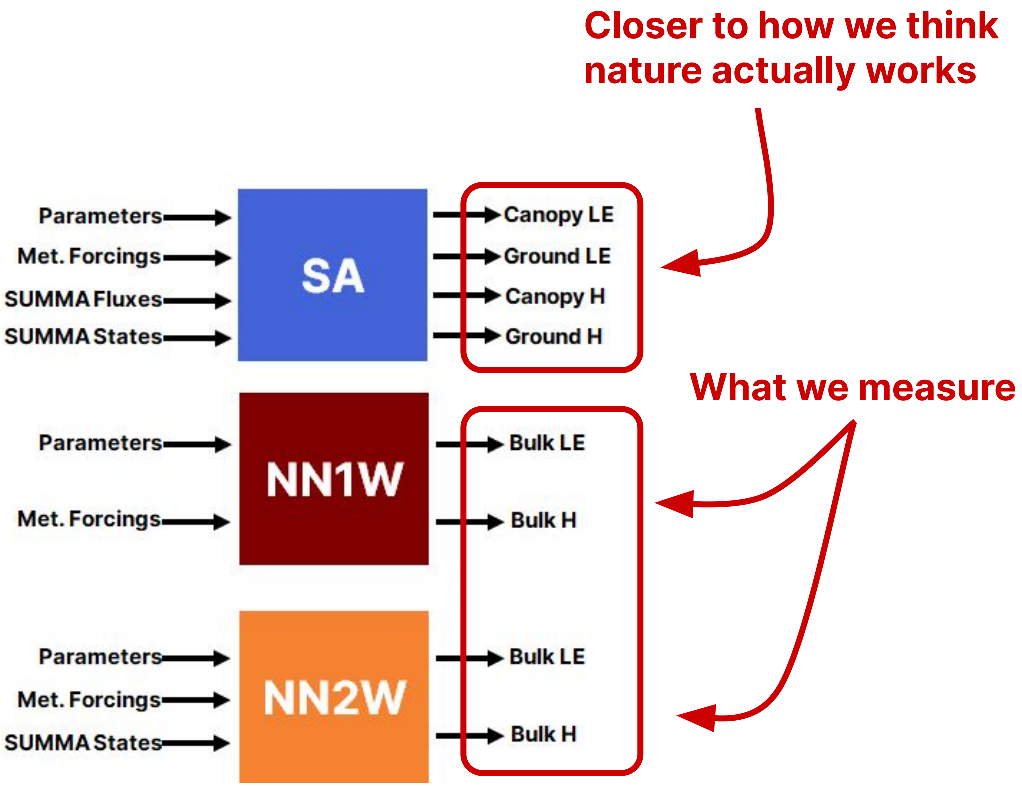
Both neural network parameterizations outperformed the standalone model, for both latent and sensible heat



Inclusion of soil states in NN2W improves long-term water balance over NN1W



One of the major shortcomings is a mismatch between process fidelity and the observed data for training

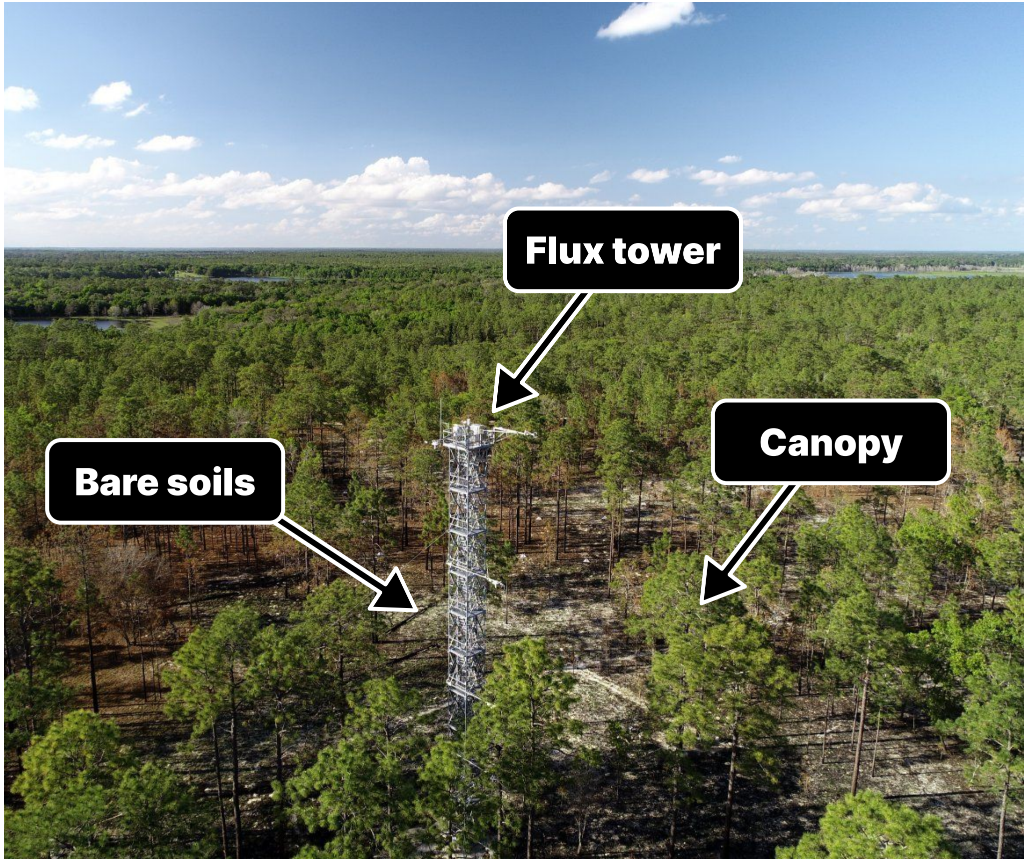


The land surface is heterogeneous!

Flux towers measure bulk fluxes

But we want to model the various components

Without fancy techniques supervised machine learning can only learn bulk fluxes from observations then



So, we've got tradeoffs

	Process based model	Machine learned model
Superior performance	✗	✓
Process fidelity	✓	✗


Why don't we have both?

So, we've got tradeoffs

	Process based model	Machine learned model
Superior performance	✗	✓
Process fidelity	✓	✓

Why don't "physics" based models perform well?

These bulk transfer equations are very common in hydrologic and land surface modeling:

- Andreadis et al., 2009
- Bonan, 1991
- Inclan and Forkel, 1995
- Sellers et al., 1986
- Mahat et al., 2013
- Clark et al., 2015
- ...

$$\left.\begin{aligned} Q_h^{veg} &= -\rho_{air} c_p C_h^{veg} \left(T^{veg} - T^{cas} \right) \\ Q_h^{sfc} &= -\rho_{air} c_p C_h^{sfc} \left(T^{sfc} - T^{cas} \right) \end{aligned}\right\} \textbf{Sensible heat fluxes}$$

$$\left.\begin{aligned} Q_{evap}^{veg} &= -\frac{L_{vap}\rho_{air}\varepsilon}{P_{air}} C_{evap}^{veg} \left[e_{sat}\left(T^{veg} \right) - e^{cas} \right] \\ Q_{trans}^{veg} &= -\frac{L_{vap}\rho_{air}\varepsilon}{P_{air}} C_{trans}^{veg} \left[e_{sat}\left(T^{veg} \right) - e^{cas} \right] \\ Q_l^{sfc} &= -\frac{L_{vap}\rho_{air}\varepsilon}{P_{air}} C_w^{sfc} \left[\phi_{hum}^{sfc} e_{sat}\left(T^{sfc} \right) - e^{cas} \right] \end{aligned}\right\} \textbf{Latent heat fluxes}$$

Why don't "physics" based models perform well?

These consist of three main parts

1. Constants & parameters
2. Temperature or moisture gradients
3. Conductance terms

$$Q_h^{veg} = -\rho_{air} c_p C_h^{veg} \left(T^{veg} - T^{cas}\right)$$

$$Q_h^{sfc} = -\rho_{air} c_p C_h^{sfc} \left(T^{sfc} - T^{cas}\right)$$

Sensible heat fluxes

$$Q_{evap}^{veg} = -\frac{L_{vap}\rho_{air}\varepsilon}{P_{air}} C_{evap}^{veg} \left[e_{sat}\left(T^{veg}\right) - e^{cas}\right]$$

$$Q_{trans}^{veg} = -\frac{L_{vap}\rho_{air}\varepsilon}{P_{air}} C_{trans}^{veg} \left[e_{sat}\left(T^{veg}\right) - e^{cas}\right]$$

$$Q_l^{sfc} = -\frac{L_{vap}\rho_{air}\varepsilon}{P_{air}} C_w^{sfc} \left[\phi_{hum}^{sfc} e_{sat}\left(T^{sfc}\right) - e^{cas}\right]$$

Latent heat fluxes

Why don’t “physics” based models perform well?

These consist of three main parts

1. Constants & parameters
2. Temperature or moisture gradients
3. Conductance terms

$$\theta \bullet C \bullet \Delta$$

Why don't "physics" based models perform well?

These consist of three main parts

1. Constants & parameters
2. Temperature or moisture gradients
3. Conductance terms

I'm going to argue these are either:

1. Pretty well known
2. Parts of other processes

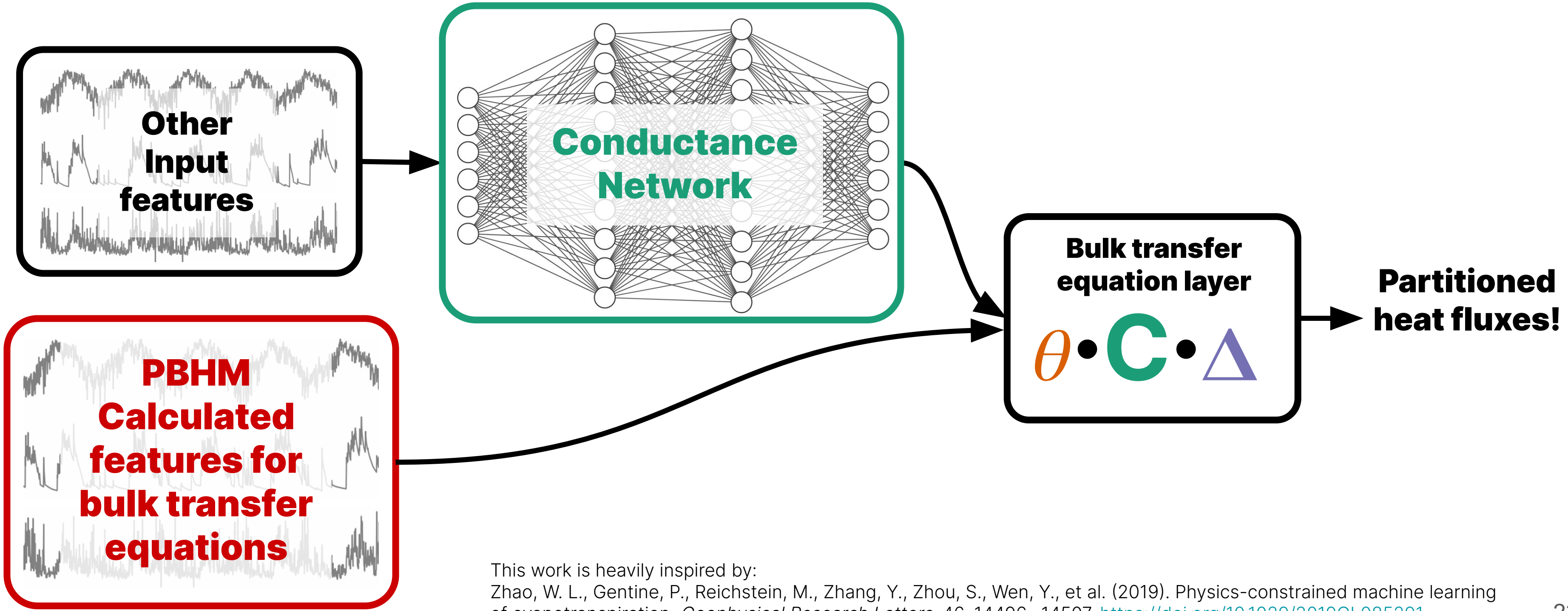
Why don't "physics" based models perform well?

These consist of three main parts

1. Constants & parameters
2. Temperature or moisture gradients
3. $\boxed{\text{Conductance terms}}$

And likewise, this is where the model uncertainty really is...

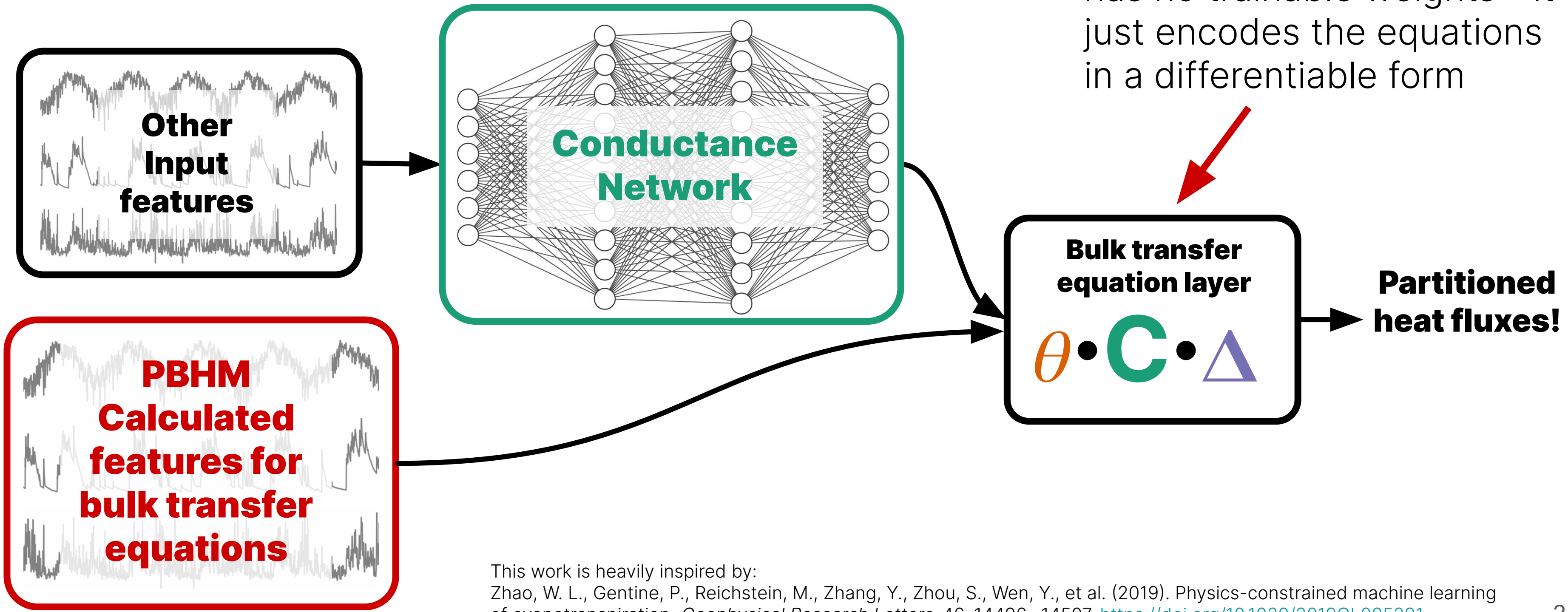
The hybrid neural network architecture



This work is heavily inspired by:
Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., et al. (2019). Physics-constrained machine learning of evapotranspiration. *Geophysical Research Letters*, 46, 14496– 14507. https://doi.org/10.1029/2019GL085291

The hybrid neural network architecture

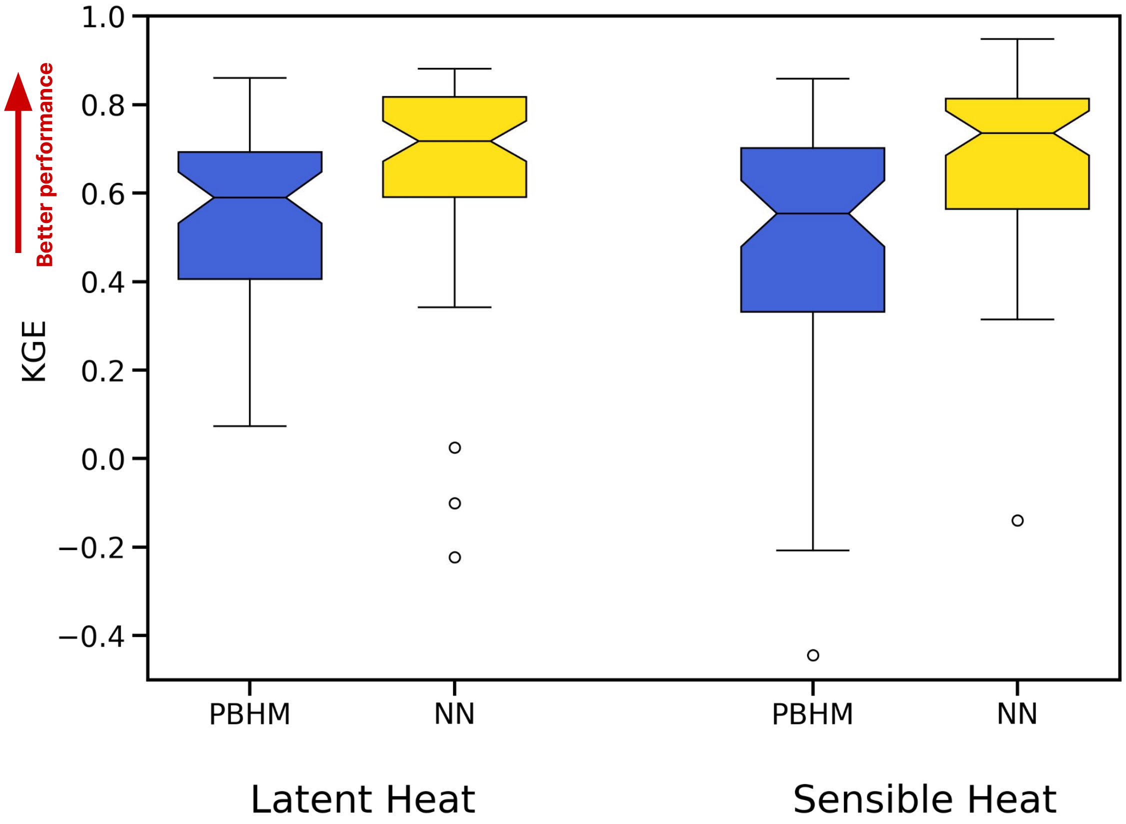
Technical note: This "layer" has no trainable weights - it just encodes the equations in a differentiable form



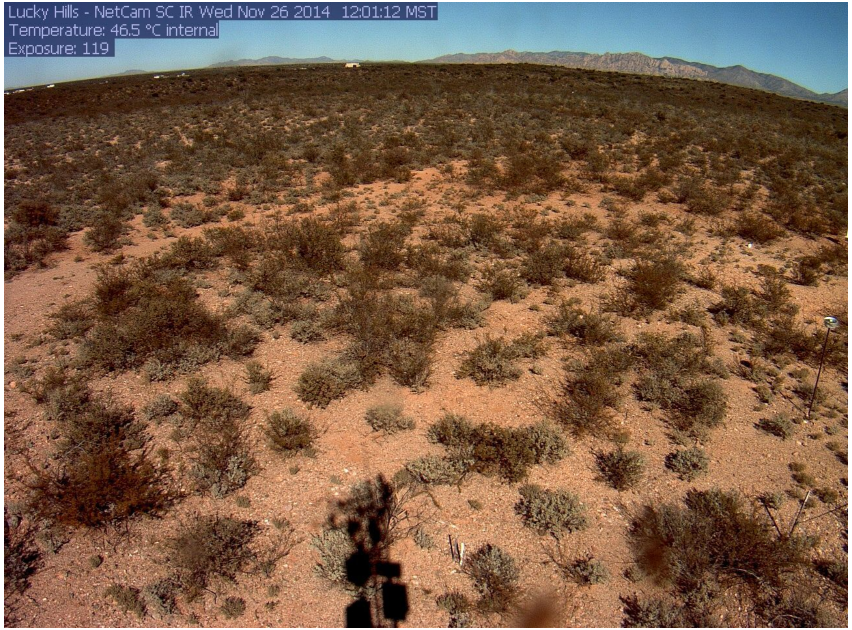
This work is heavily inspired by:
Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., et al. (2019). Physics-constrained machine learning of evapotranspiration. *Geophysical Research Letters*, 46, 14496– 14507. https://doi.org/10.1029/2019GL085291

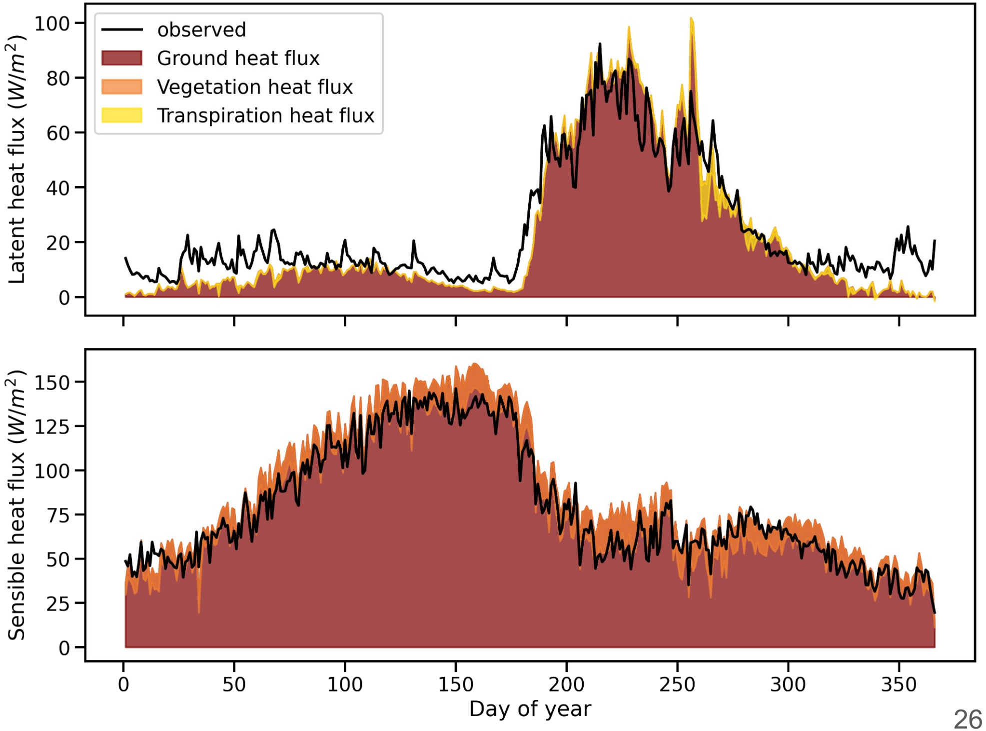
We're still able to outperform a calibrated PBHM using the same bulk transfer equations*

**other pure ML based approaches outperform this but that's asking a different question*

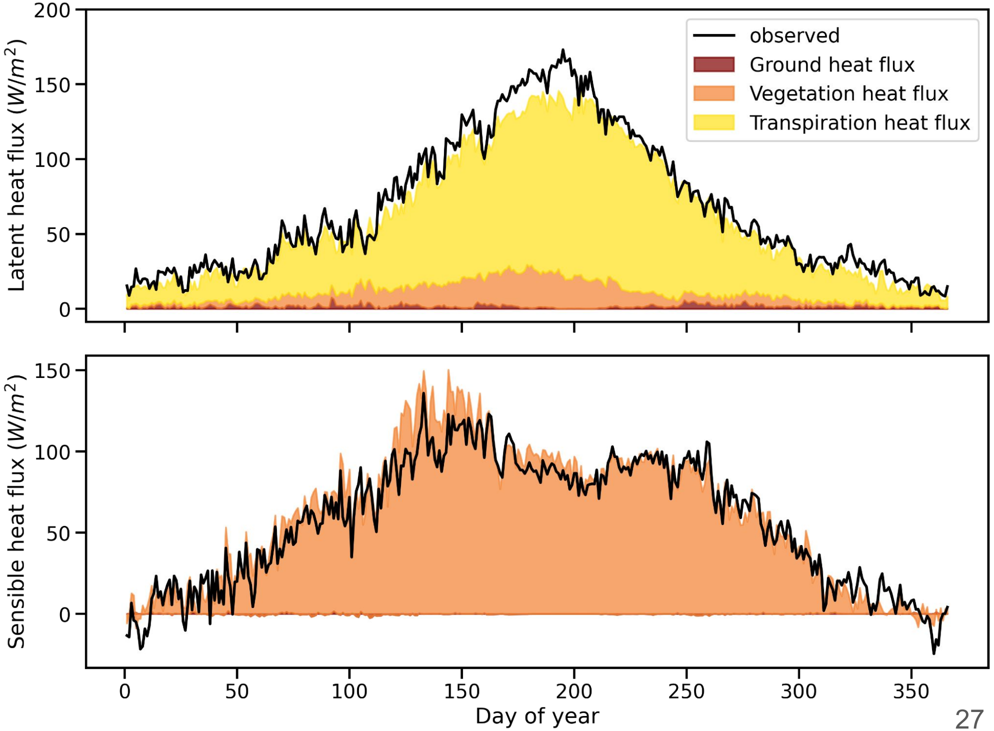


Walnut Gulch near Tombstone, AZ (US-Whs) shows ground component is largest

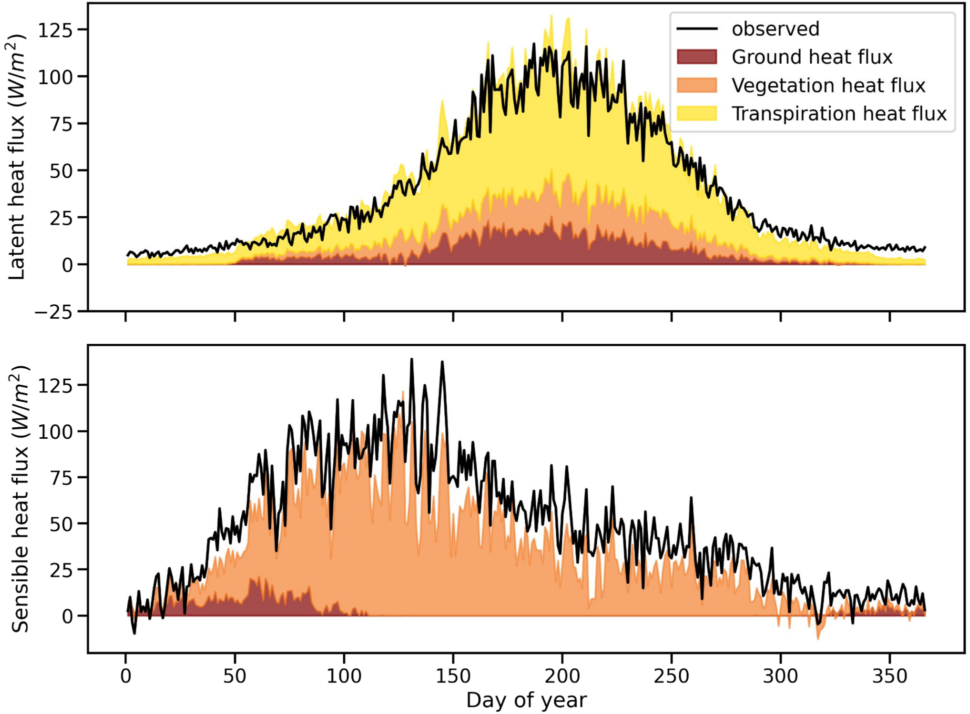




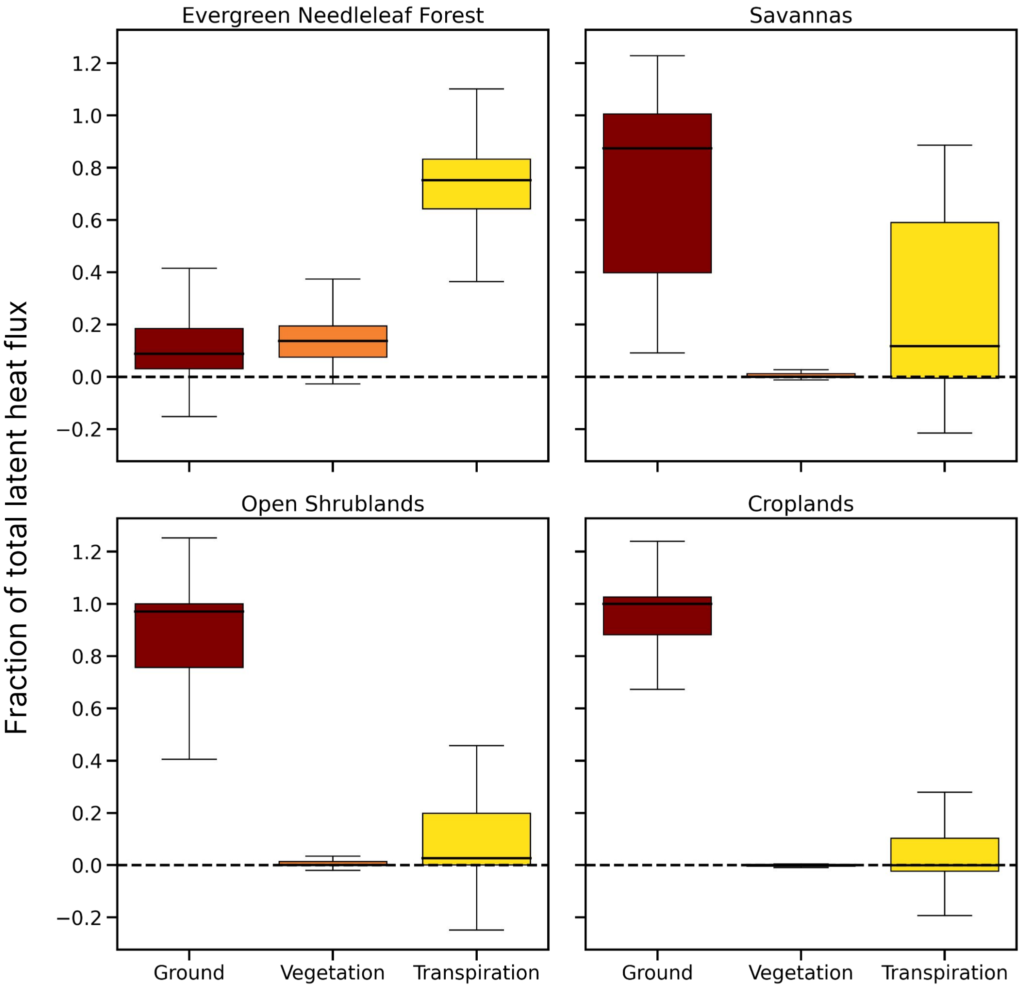
Blodgett Forest near Sacramento, CA (US-Blo) shows vegetation components are largest



Mixed forest near Vielsalm, Belgium shows a larger mixture between components

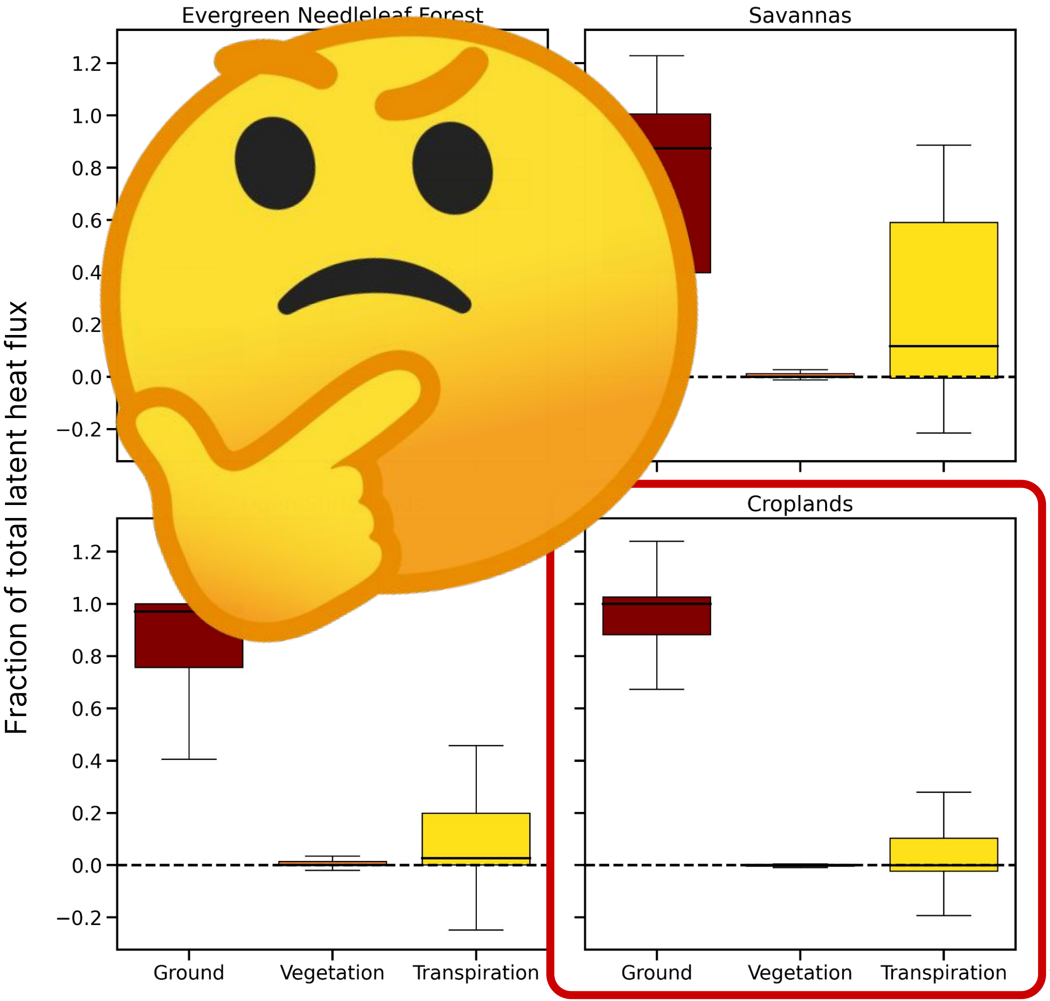


The overall partitioning matches physical intuition to a first order



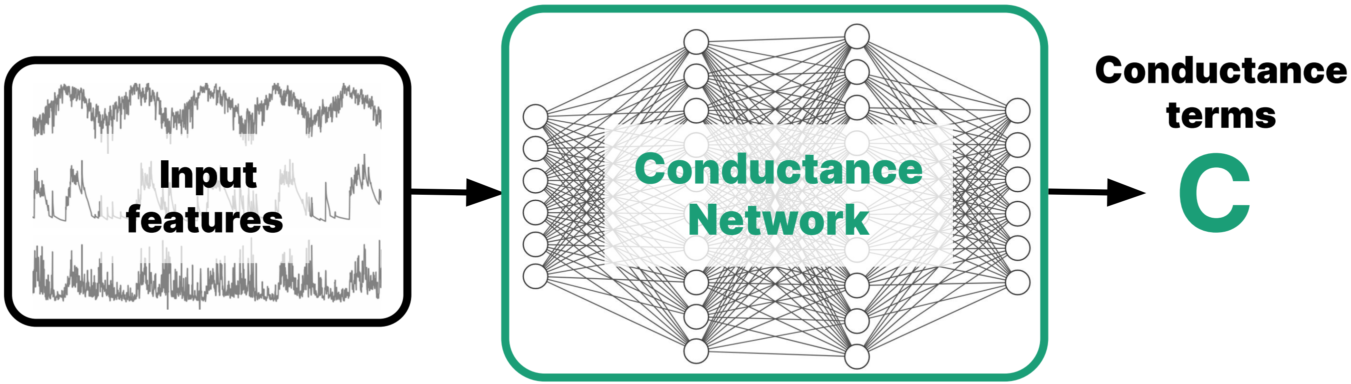
Note: values can be <0 and >1 because condensation exists

The overall partitioning matches physical intuition to a first order, **mostly**

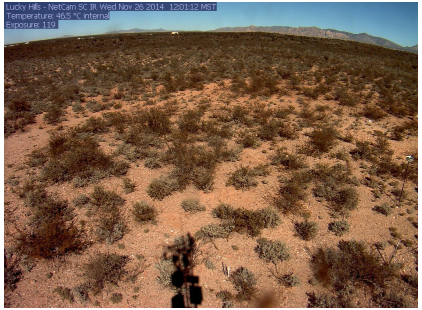


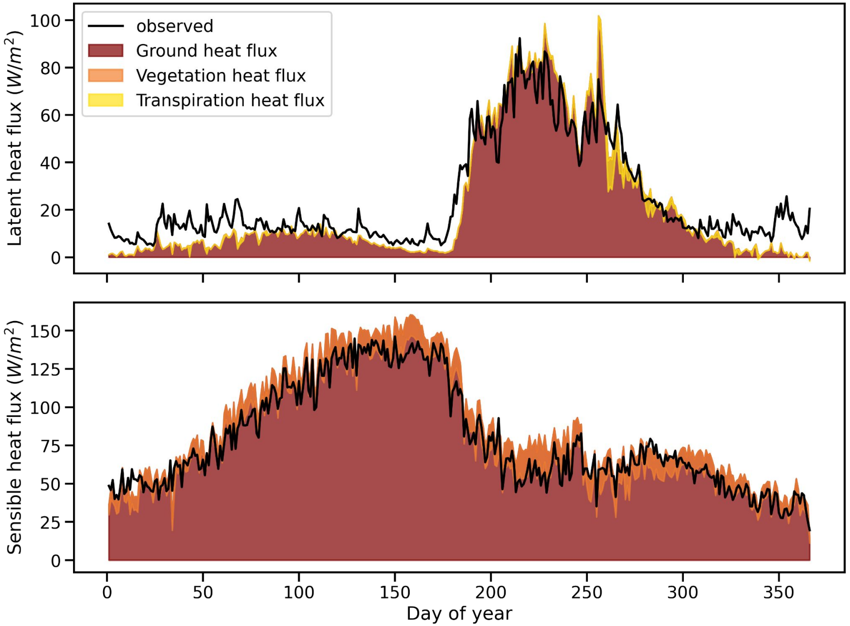
Note: values can be <0 and >1 because condensation exists

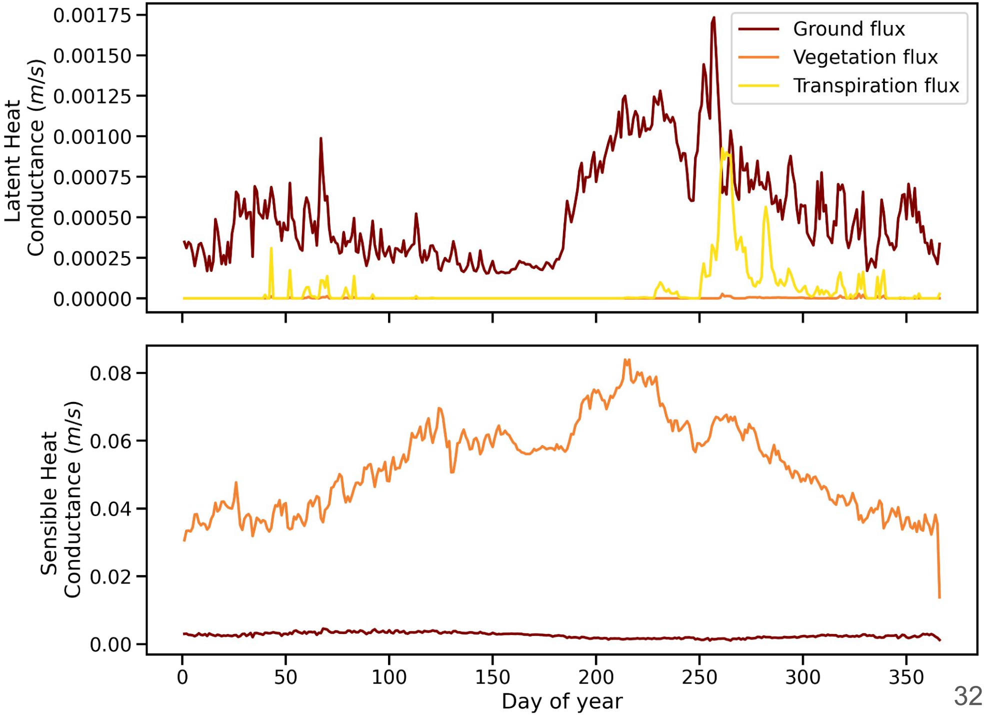
We can also truncate the network to analyze the conductances!



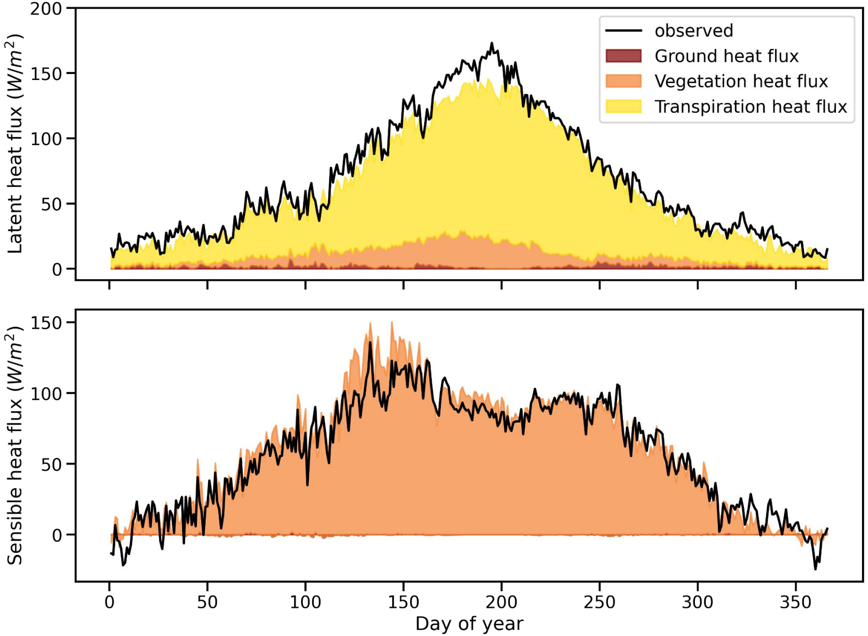
Comparing conductances shows the network learns nonlinear behavior

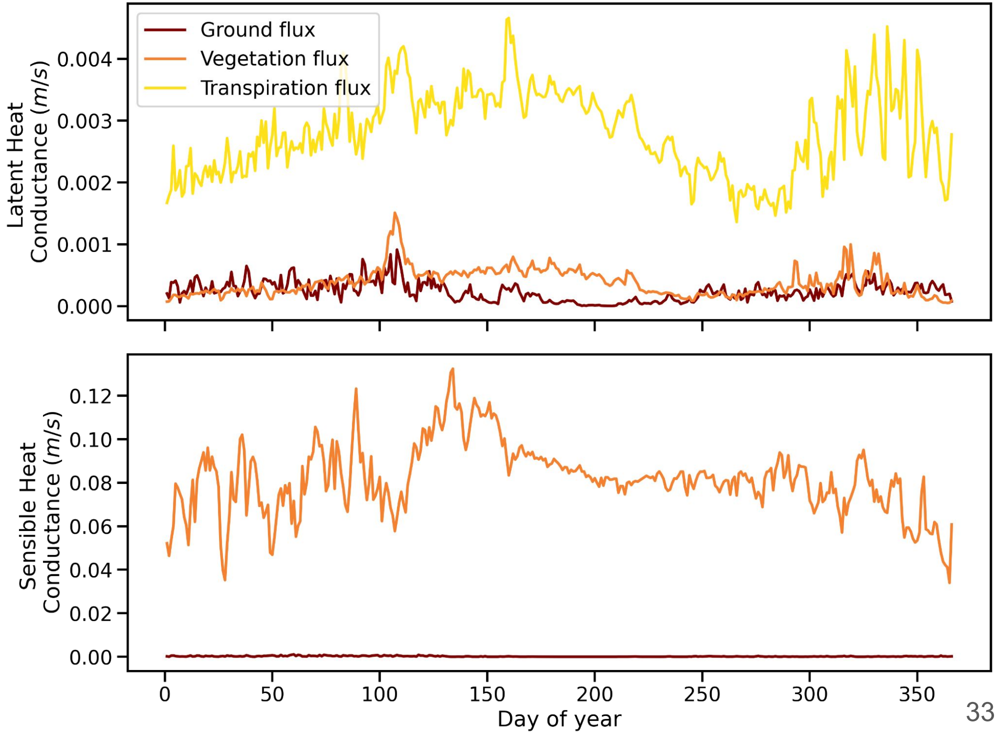






Conductances are not 1-1 with heat fluxes at Blodgett, CA





Wrapup and future work

We’ve quantified that a large amount of predictive performance is due to conductance terms

Need methods/data to better constrain the partitioning, particularly at sites with human interventions, like croplands

Coupling to the PBHM is still incomplete, but needed to analyze the effects on the full water cycle

Lots of validation work is still needed to build confidence

Thanks!
Questions?