Do internals of neural networks make sense in the context of hydrology

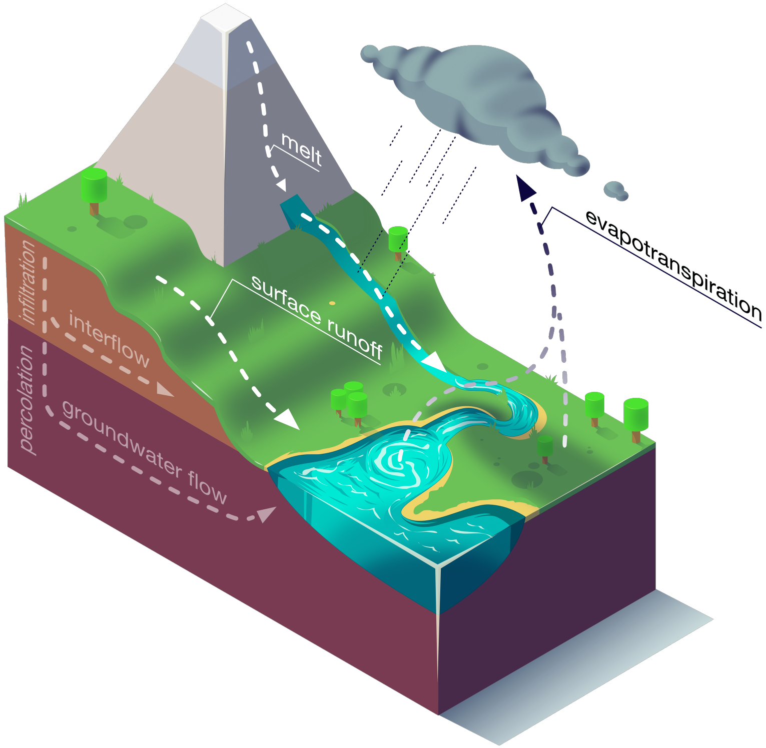
Frederik Kratzert[1], M. Herrnegger[2], D. Klotz[1], S. Hochreiter[1], G. Klambauer[1]

JOHANNES KEPLER UNIVERSITY LINZ

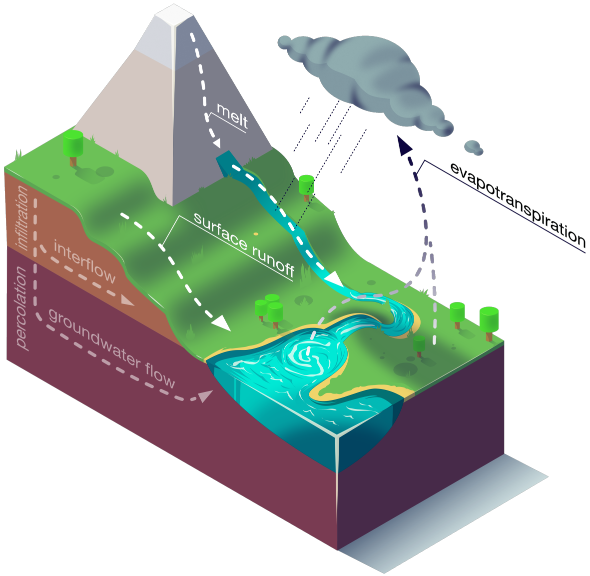
[1]LIT AI Lab & Institute for Machine Learning
Johannes Kepler University
Linz, Austria

[2]Institute for Hydrology and Water Management
University of Natural Resources and Life Sciences
Vienna, Austria

Rainfall-runoff modeling with LSTMs

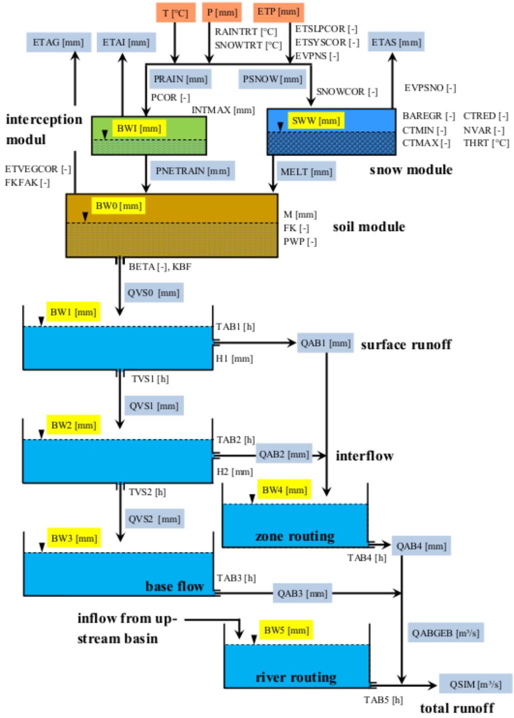


Classical modeling approach



State-(space)-model formulation:

$$\boldsymbol{S}_t = f(\boldsymbol{I}_t, \boldsymbol{S}_{t-1}, \Theta_i)$$
$$\boldsymbol{O}_t = f(\boldsymbol{I}_t, \boldsymbol{S}_t, \Theta_j)$$



COSERO model, taken from official handbook

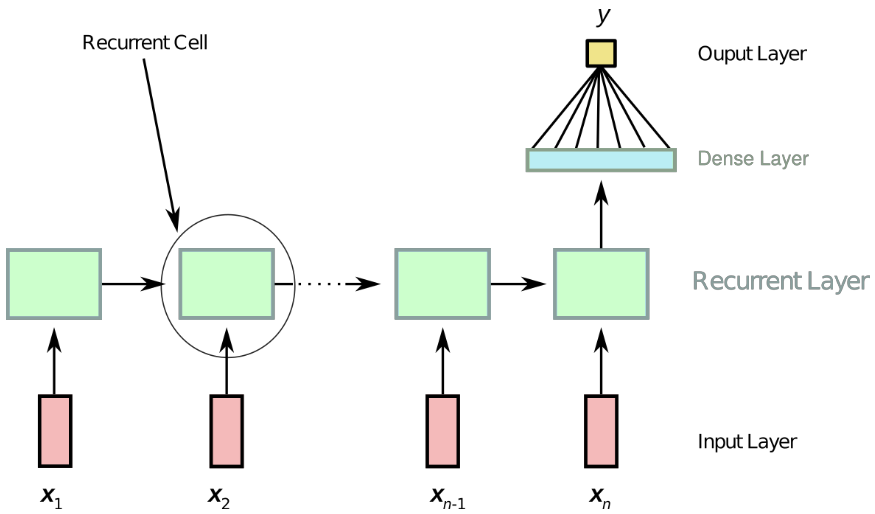
Long Short-Term Memory (LSTM)

- Recurrent neural network with explicit memory [1]
- State of the art for many applications
  - Used e.g. in Siri, Google Assistant, Alexa
- Works very well for rainfall-runoff modeling [2]

[1] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[2] Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, Hydrol. Earth Syst. Sci., 22, 6005-6022.

Long Short-Term Memory (LSTM)

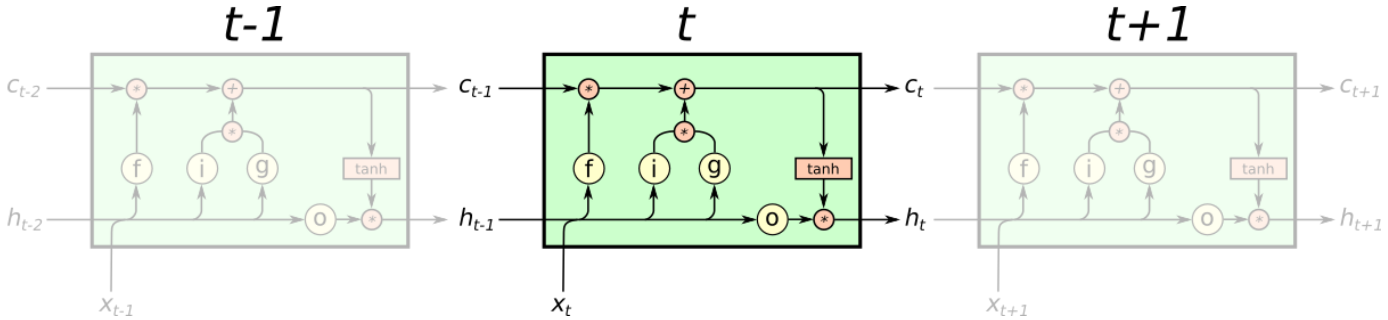


Hyd. State-(space)-model formulation:

$$\boldsymbol{S}_t = f(\boldsymbol{I}_t, \boldsymbol{S}_{t-1}, \Theta_i)$$
$$\boldsymbol{O}_t = f(\boldsymbol{I}_t, \boldsymbol{S}_t, \Theta_j)$$

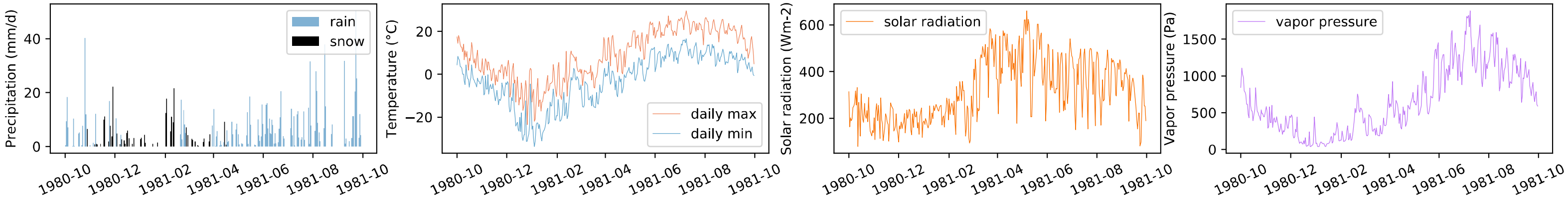
Analogous for the LSTM:

$$\{\boldsymbol{c}_t, \boldsymbol{h}_t\} = f_{\mathrm{LSTM}}(\boldsymbol{x}_t, \boldsymbol{h}_{t-1}, \boldsymbol{c}_{t-1}, \Theta_k)$$
$$y_t = f_{\mathrm{Dense}}(\boldsymbol{h}_t, \Theta_l)$$



Data + Setup

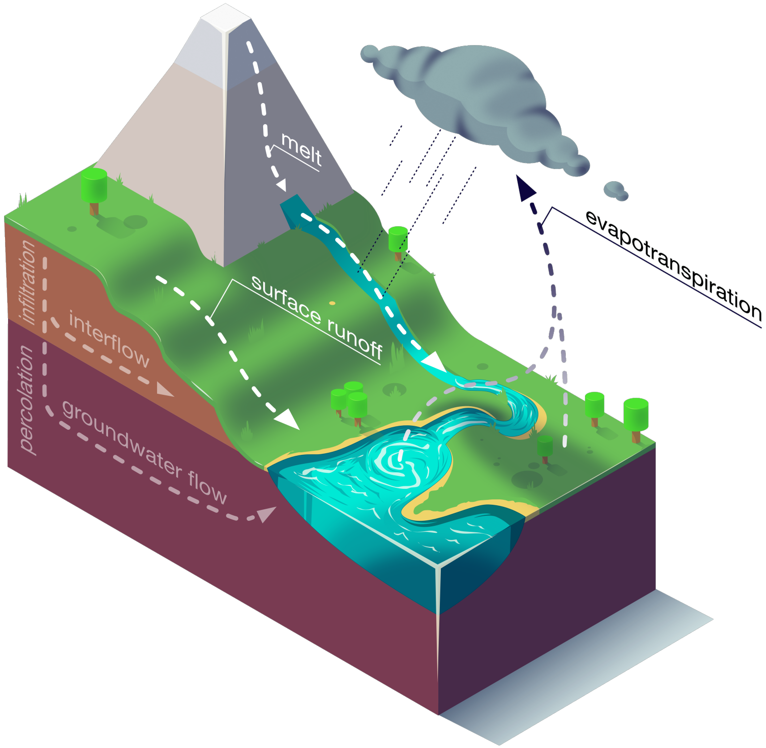
- 47 humid & snow-influenced basins [3]
- 20 yrs for training, 5 yrs for validation, remaining ~7 yrs for final evaluation
- Task: predict discharge of single day from previous 365 daily meteorological observations



$$q = 2.44 \frac{mm}{d}$$

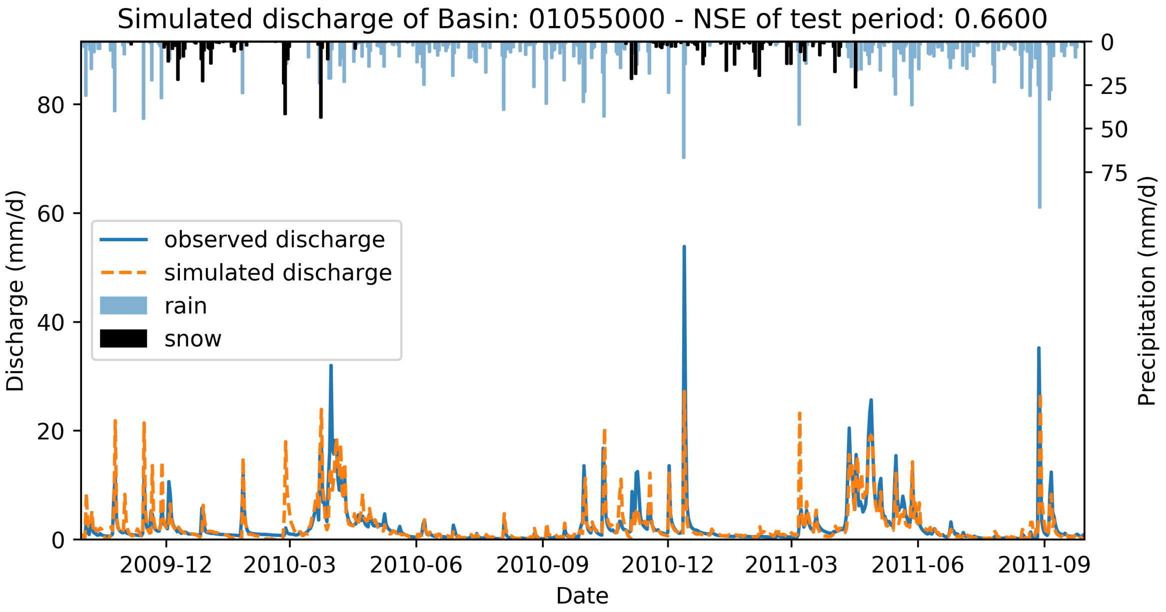
[3] A. Newman; K. Sampson; M. P. Clark; A. Bock; R. J. Viger; D. Blodgett, 2014. A large-sample watershed-scale hydrometeorological dataset for the contiguous USA. Boulder, CO: UCAR/NCAR.

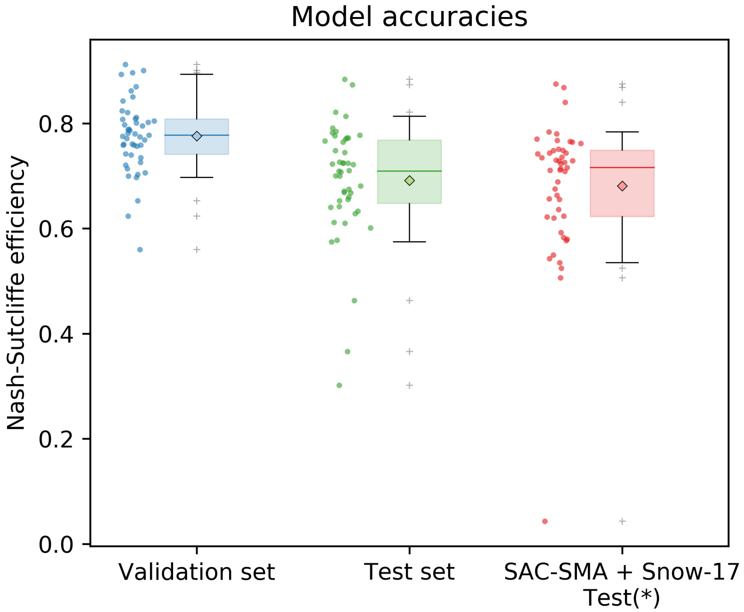
Results and network inspection



Model accuracy

- Comparable performance to SAC-SMA + Snow-17

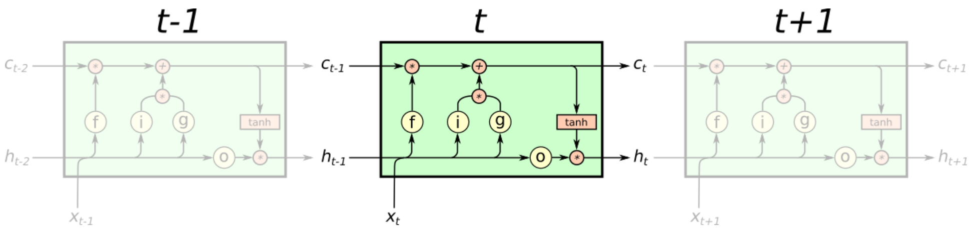




(*) covers not the exact same period.
Only meant for rough comparison

First glimpse into the LSTM

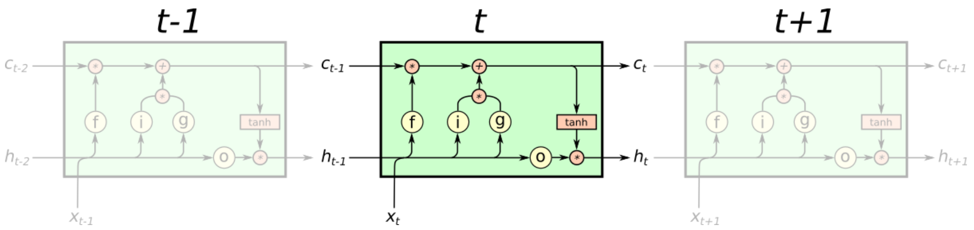
Found on the last days before submitting HESS paper

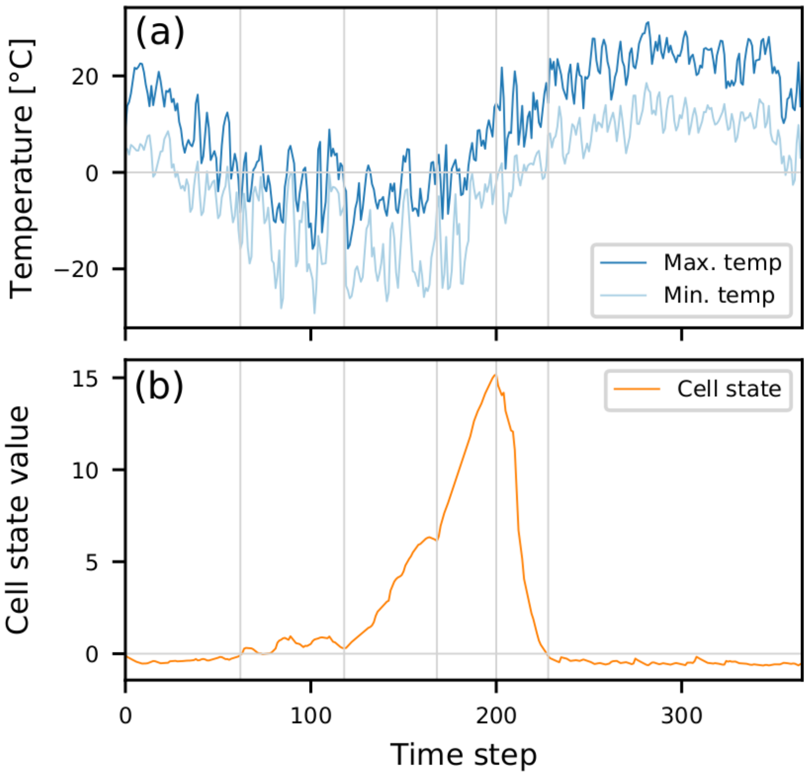


First glimpse into the LSTM

Found on the last days before submitting HESS paper

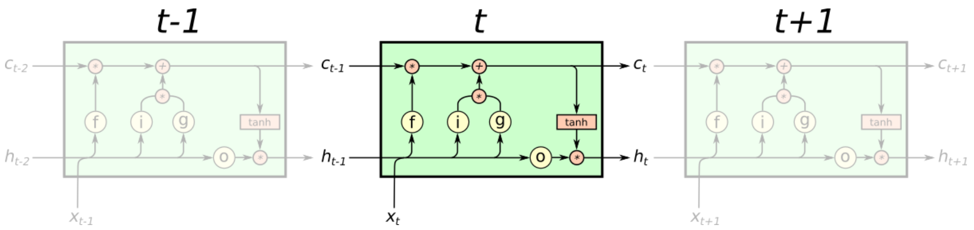
➔ Coincidence?

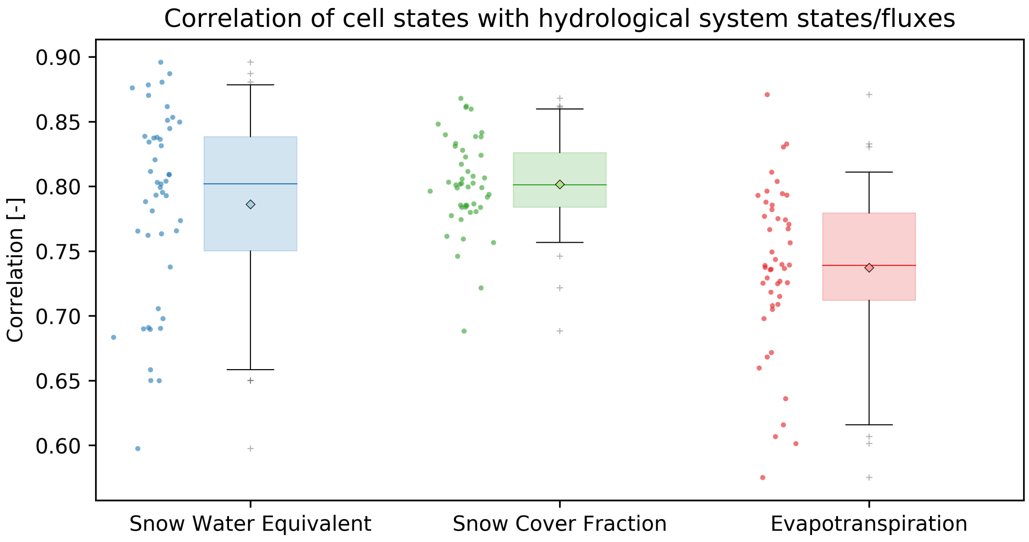




Correlations

- LSTM memory cell (c_t) & cell output (h_t) vs. hydrological states/fluxes
  - SCF + ET from MODIS
  - Modeled SWE from SAC-SMA/Snow-17

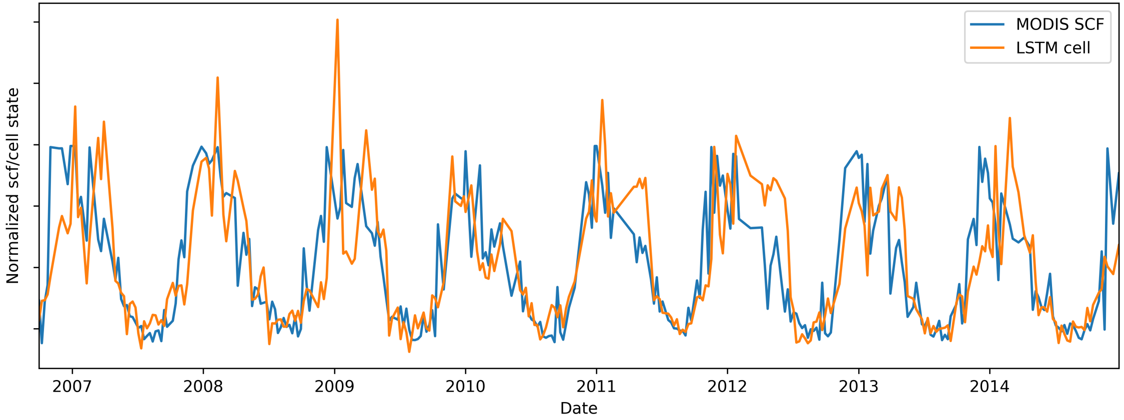




Some examples

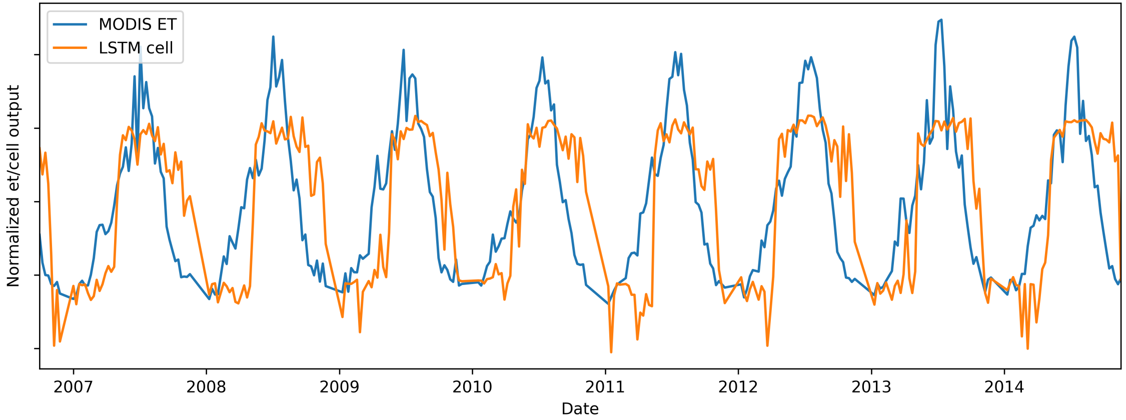
- **SCF:** Plot of basin most similar to median correlation

Snow cover fraction (SCF) cell: corr=0.80, p=2.17e-66



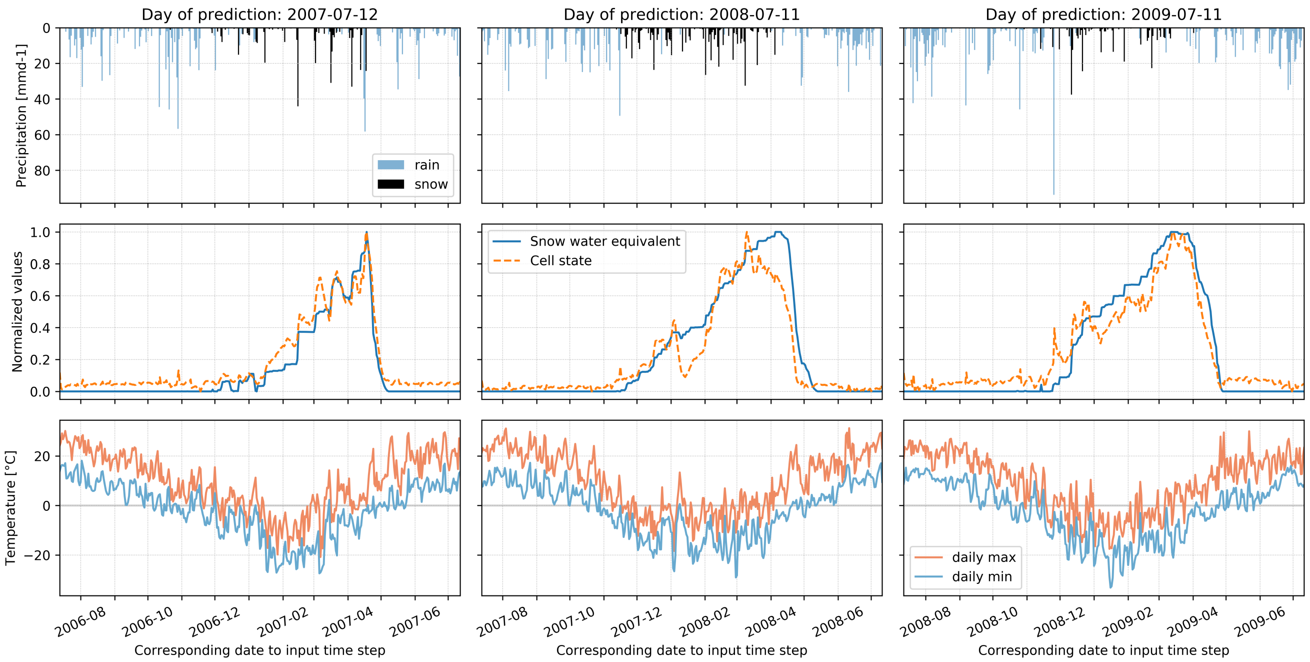
- **ET:** Plot of basin most similar to median correlation

Evapotranspiration (ET) cell: corr=0.73, p=8.21e-55



Snow cell

Remember: LSTM has never seen SWE data. Purely learned from rainfall-runoff task.



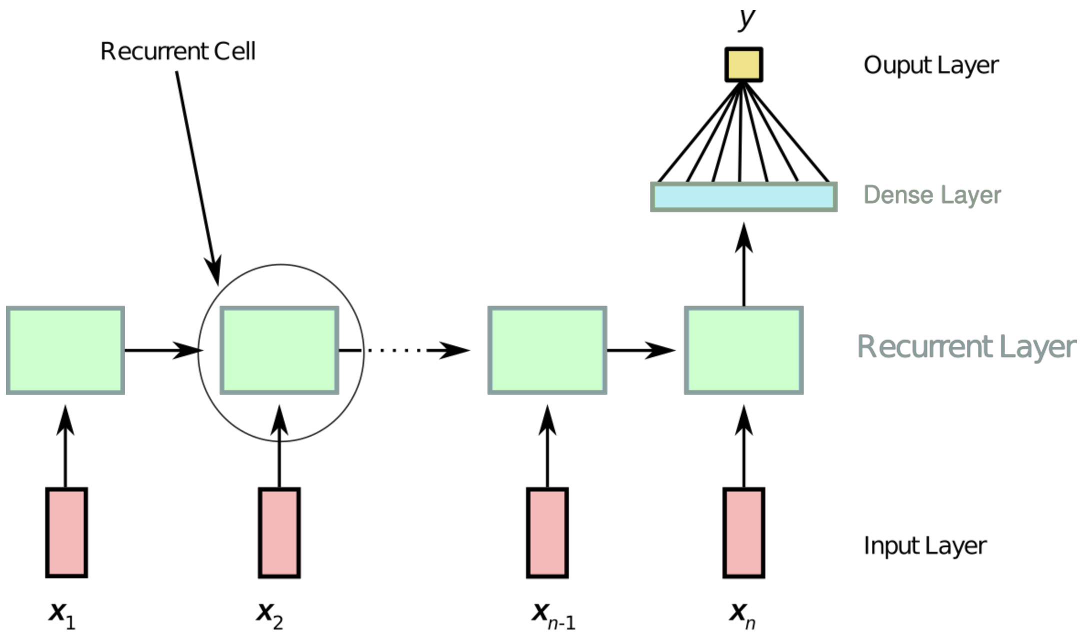
Questions?

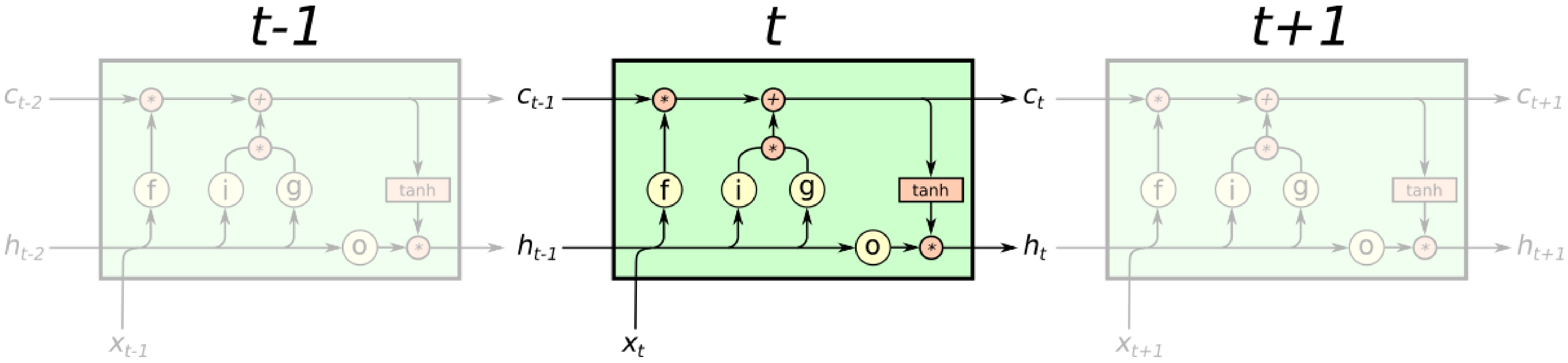
Slides + Code for reproducing results:
https://github.com/kratzert/AGU2018

@fkratzert

$$
\begin{aligned}
\mathbf{f}_t &= \sigma(\mathbf{W}_f \mathbf{h}_{t-1} + \mathbf{U}_f \mathbf{x}_t + \mathbf{b}_f) \\
\mathbf{i}_t &= \sigma(\mathbf{W}_i \mathbf{h}_{t-1} + \mathbf{U}_i \mathbf{x}_t + \mathbf{b}_i) \\
\mathbf{g}_t &= \tanh(\mathbf{W}_g \mathbf{h}_{t-1} + \mathbf{U}_g \mathbf{x}_t + \mathbf{b}_g) \\
\mathbf{o}_t &= \sigma(\mathbf{W}_o \mathbf{h}_{t-1} + \mathbf{U}_o \mathbf{x}_t + \mathbf{b}_o)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \mathbf{g}_t \\
\mathbf{h}_t &= \mathbf{o}_t \odot \tanh(\mathbf{c}_t)
\end{aligned}
$$





A glimpse into the Unobserved: Runoff simulation for ungauged catchments with LSTMs

Frederik Kratzert* kratzert@ml.jku.at
Daniel Klotz* klotz@ml.jku.at
Mathew Herrnegger † mathew.herrnegger@boku.ac.at
Sepp Hochreiter* hochreit@ml.jku.at

Workshop on Modeling and Decision-Making in the Spatiotemporal Domain, 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

https://openreview.net/pdf?id=Bylhm72oKX