Learning the Kalman Update End-to-End for Hydrological Forecasting

From Noise-Covariance Optimization to Neural Gain Learning

A unified differentiable Kalman framework for learning hydrological state updates

Yiqun Sun, Xin Tian, Qiongfang Li, Zheng Fang, Sida Liu, Hamid Moradkhani

Hohai University | Deltares | IGRAC | Wageningen University & Research | University of Alabama

EGU General Assembly 2026 | Vienna, Austria

The Problem and the Gap

Classical Kalman filtering

UKF, EnKF

✓ **Sequential predict-update loop**
Online correction at every time step

✗ **Q/R have no ground truth in hydrology**
Hand-tuned, not reproducible

Recent deep-learning DA

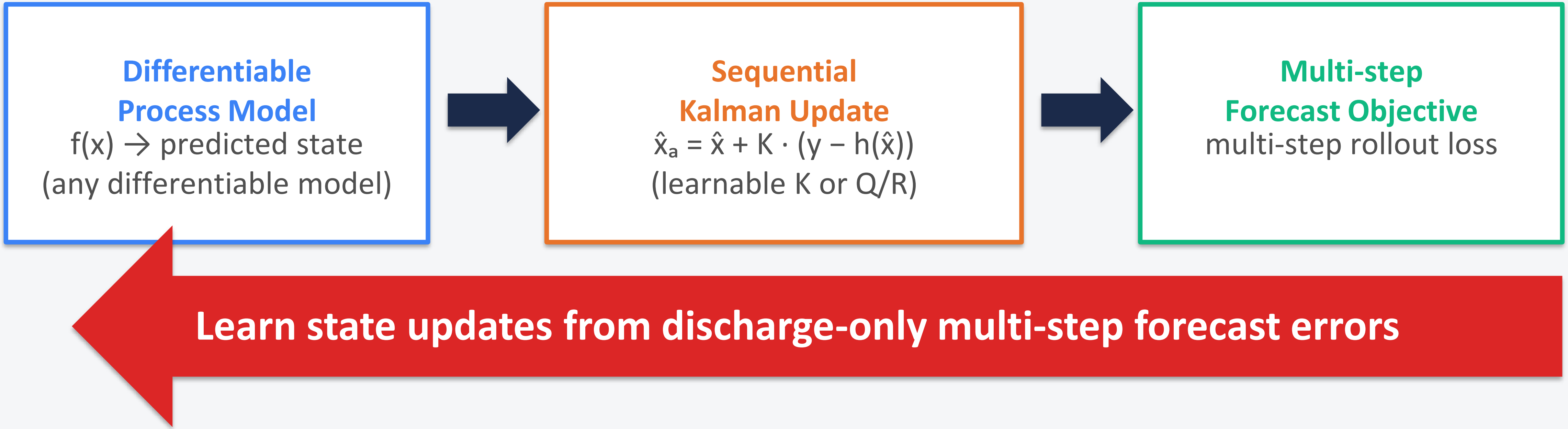
variational, autoregressive, ...

✓ **Learns the update from data**
No more hand-tuning

✗ **Abandons the sequential loop**
Offline — loses online correction

Can we keep the sequential Kalman loop AND learn the update from data?

Our Proposal: A Unified Differentiable Kalman Framework



- After each update, train with a 12-hour discharge forecast loss
- Self-supervised: discharge observations only — no state labels
- Extensible to differentiable Kalman-type filters

Two Learnable Update Paradigms Within the Framework

Learned-noise UKF

Indirect update learning

- Retain the classical covariance-based UKF update
- Learn diag(Q) + R via backpropagation
- **6 parameters only**
- Interpretable Q/R parameters

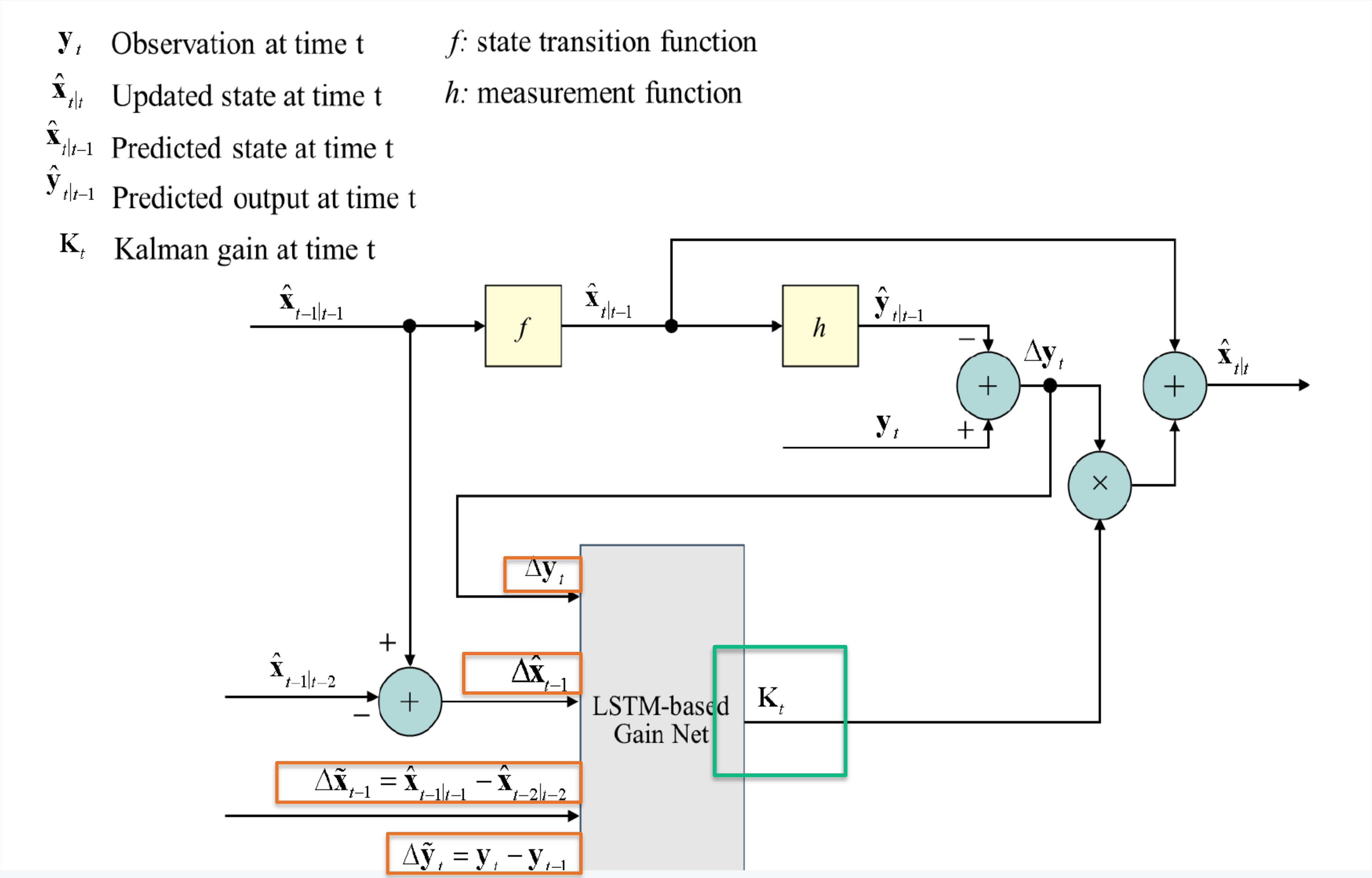
LSTM-NGF

Direct update learning

- LSTM replaces covariance-based gain computation entirely
- Input: innovation + state evolution
- **~49,000 parameters**
- Time-varying, context-aware gain

Same framework · Same process model · Same objective · **Only the update-learning strategy differs**

LSTM-NGF: How the Neural Gain Is Computed



Key idea

LSTM learns K from innovation sequences directly

Inputs

- Innovation
- Observation diff
- Forward state evolution
- Update difference

Why it matters

Time-varying gain can help when covariance updates weaken

Controlled Comparison Within the Proposed Framework

Study Area

Regge Catchment

The Netherlands · 957 km² · Lowland

Data

Hourly P, ETpot, Q (1999–2011)

12 years · ~105,000 time steps

Process Model

WALRUS (5 state variables, 1 output)

Experiment Design

Data Split

70 / 15 / 15 (train / val / test)

Chronological (2009–2011 = test)

Training

800 h sliding windows + augmentation

Training horizon T = 12 h

Loss: composite multi-step objective (weighted MSE-based)

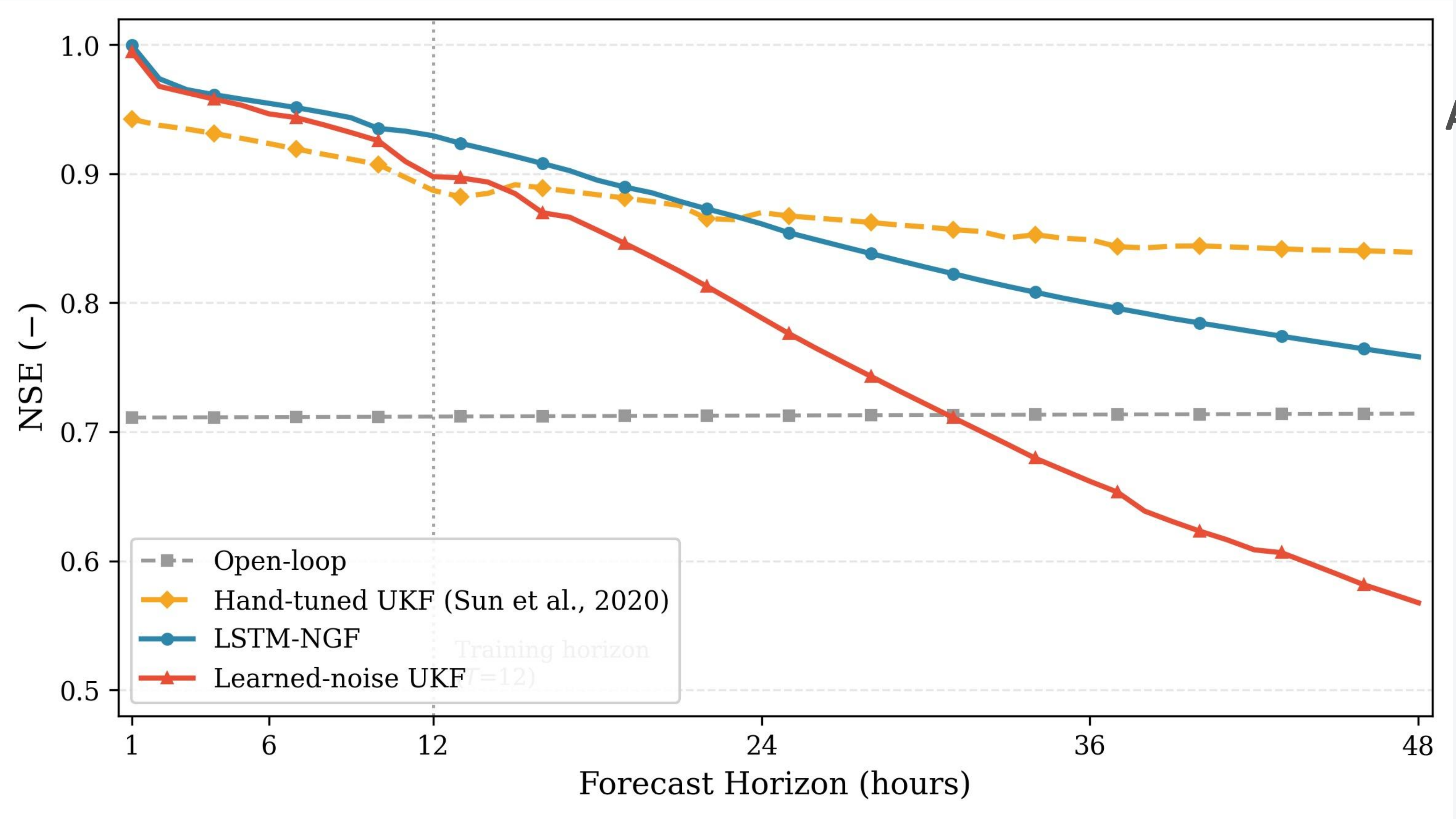
Evaluation

Dense rolling forecasts (main focus: H1–H24)

Training uses a 12-hour rollout objective; results beyond H12 therefore evaluate out-of-training-horizon generalization.

Same process model, same data split, same training objective — only the update-learning strategy changes.

Direct Gain Learning Gives the Best Short-Range Skill



At H12 *(training boundary)*

0.929 LSTM-NGF

0.898 Learned-noise UKF

0.887 Hand-tuned UKF

At H24 *(beyond training)*

0.861 LSTM-NGF

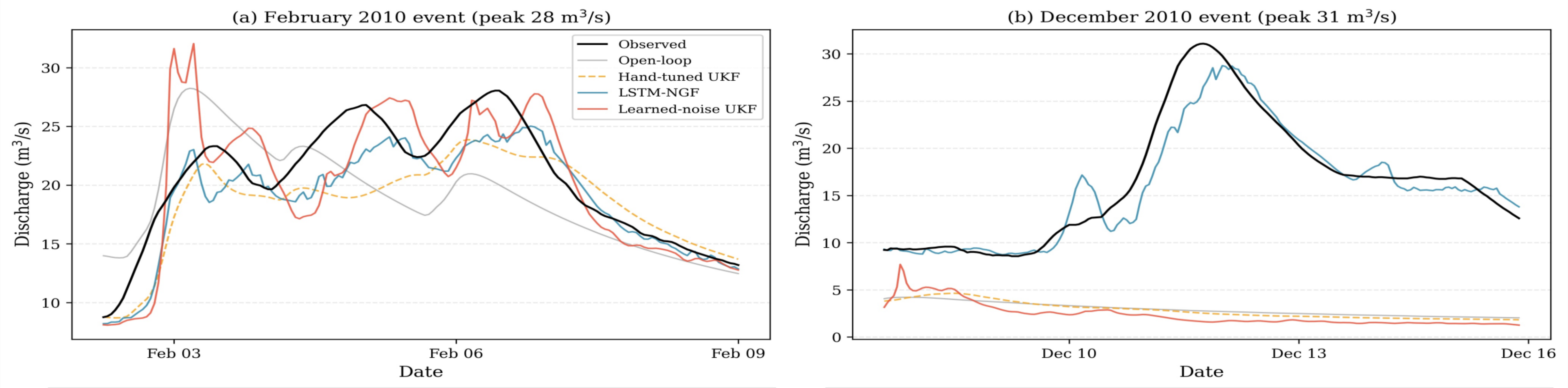
0.870 Hand-tuned UKF

0.788 Learned-noise UKF

LSTM-NGF gives the strongest short-range skill and shows the strongest H12 performance across all three flow regimes, while the hand-tuned UKF stays more conservative beyond the training horizon.

Why the Methods Differ: Flood-Event Behavior

12-hour-ahead (H12) forecasts on independent test set (2009–2011)



(a) Multi-peak event — Feb 2010

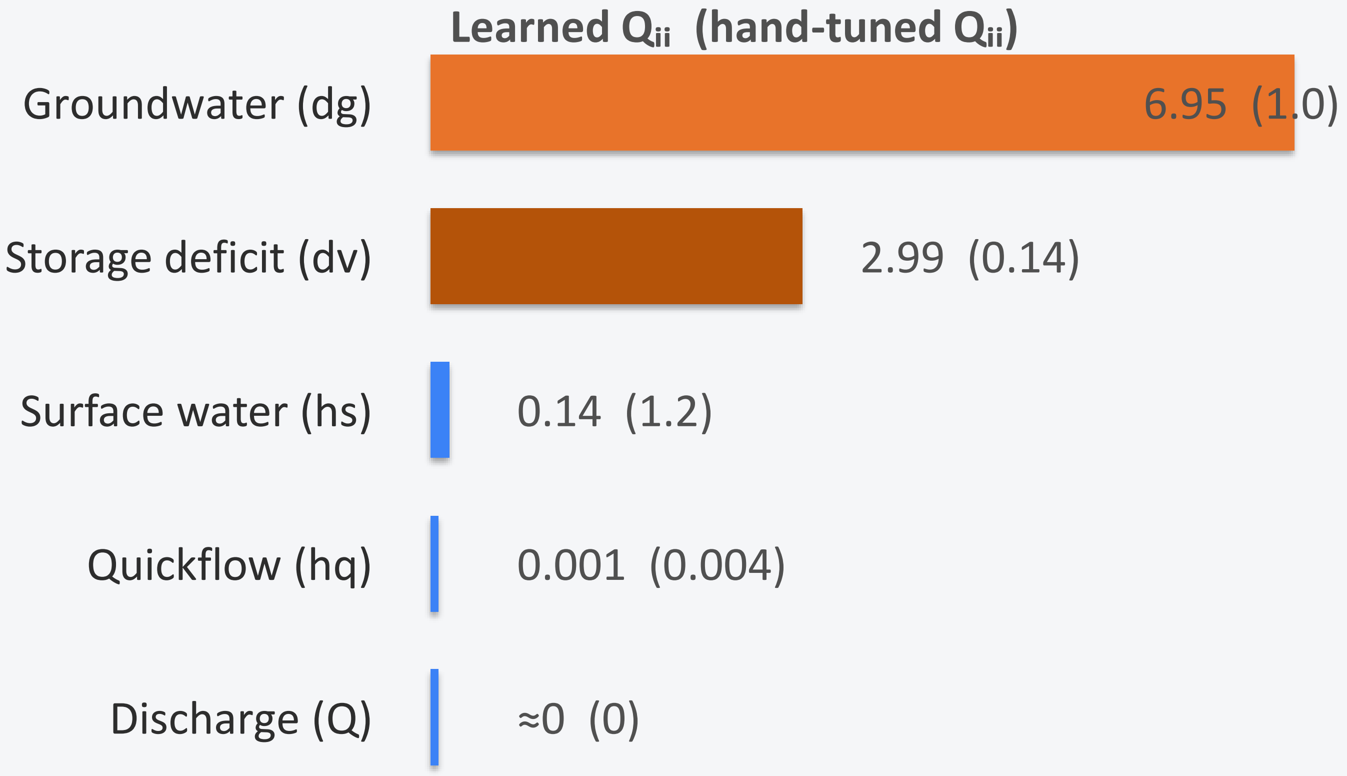
- LSTM-NGF tracks the observed hydrograph most closely
- Learned-noise UKF oscillates (near-zero R)
- Hand-tuned UKF underestimates peaks

(b) Sensitivity-collapse-like case — Dec 2010

- Both UKF variants give weak corrections in this sensitivity-collapse-like regime
- LSTM-NGF still captures ~92.5% of the flood peak, consistent with its gain not being directly tied to local covariance sensitivity

Learned Q/R Provide Forecast-Oriented Diagnostics

What indirect learning adds: transparent Q/R diagnostics



Key Findings

- Keeps covariance-based Kalman structure
- Learns only 6 parameters: diag(Q) + R
- Largest learned Q: groundwater + storage deficit
- Learned R ≈ 0: observations dominate
- Shows where forecast loss allows stronger corrections

Forecast-oriented effective parameters — not physical noise estimates.

Take-Home Messages

1. **A unified differentiable Kalman framework makes end-to-end update learning possible**

   Self-supervised with discharge only; retains the sequential predict-update loop.

2. **Direct gain learning gives the strongest short-range forecast skill**

   Strongest H12 performance across all three flow regimes; shows more robust flood-event behavior in this case study.

3. **Learning Q/R remains useful**

   Preserves Kalman structure and exposes forecast-oriented Q/R diagnostics.

Outlook: Multi-catchment validation · Hybrid short-range / long-range strategy

Thank You!

I'd be happy to discuss both the methodological side and the hydrological implications.

Manuscript under review

Contact

yiqun.sun@hotmail.com