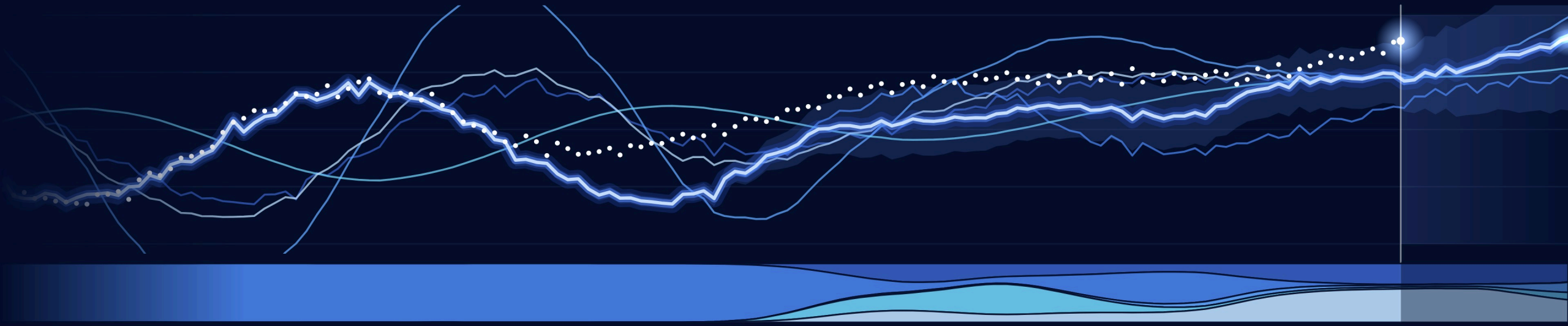
NEURIPS 2026

Ensemble Modeling for Time Series Forecasting:

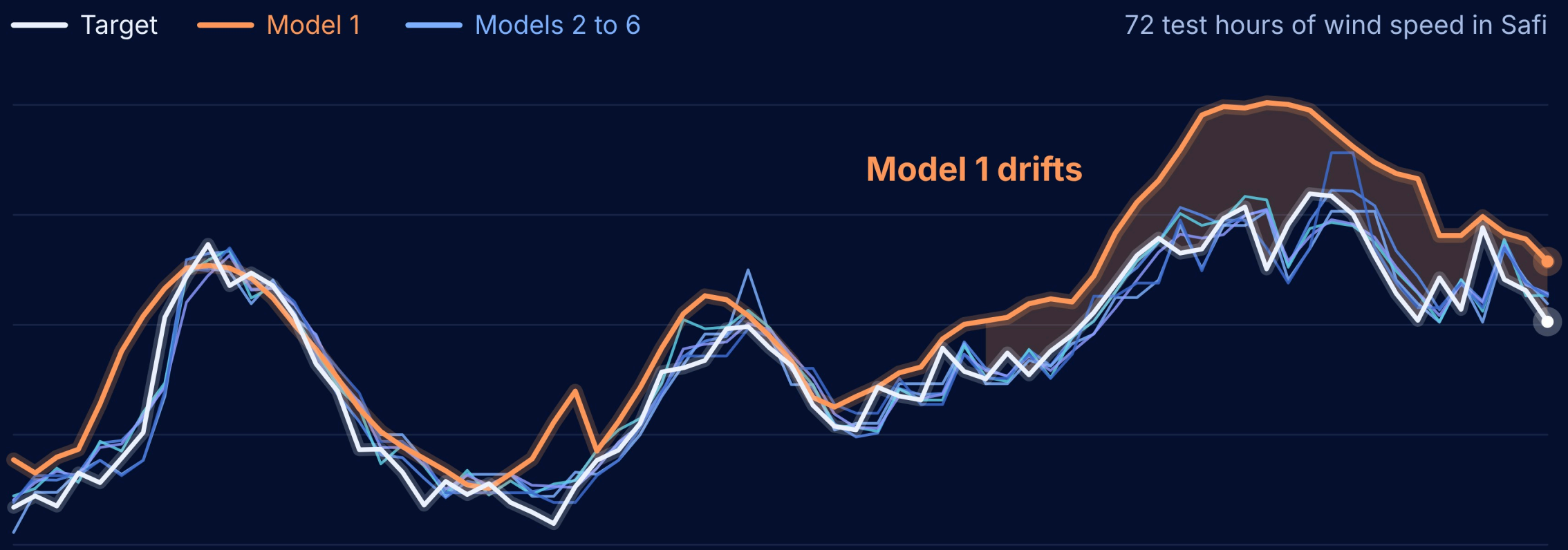
an Adaptive Robust Optimization Approach

Léonard Boussioux · Henry Mao · Dimitris Bertsimas

UNIVERSITY of WASHINGTON · FOSTER SCHOOL OF BUSINESS · PAUL G. ALLEN SCHOOL OF COMPUTER SCIENCE & ENGINEERING · MIT

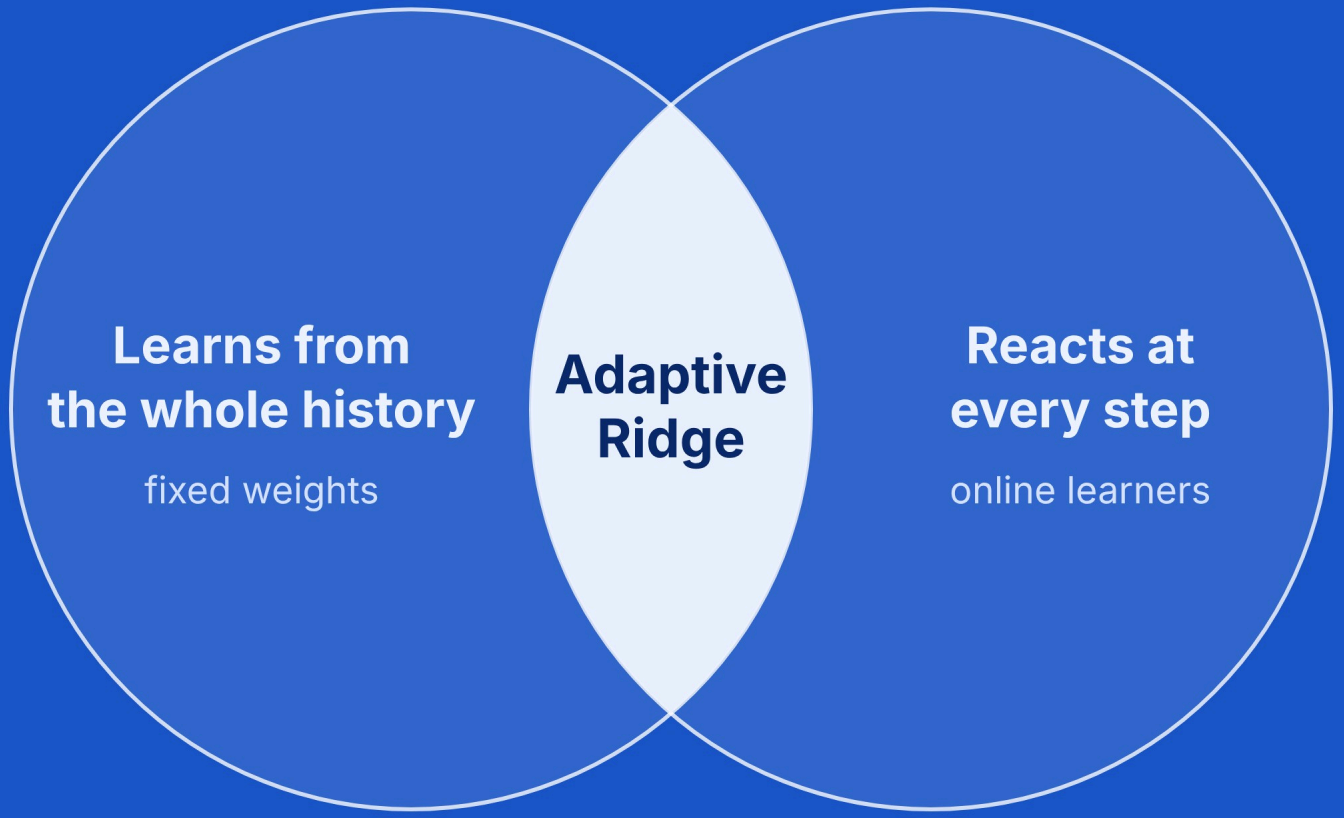


Which forecast should we trust?



Six models forecast the same wind speed. Model 1 drifts in the second half.

The goal



One simple rule, learned from past errors, tells the weights how to react.

Adaptive robust optimization

1 Here and now

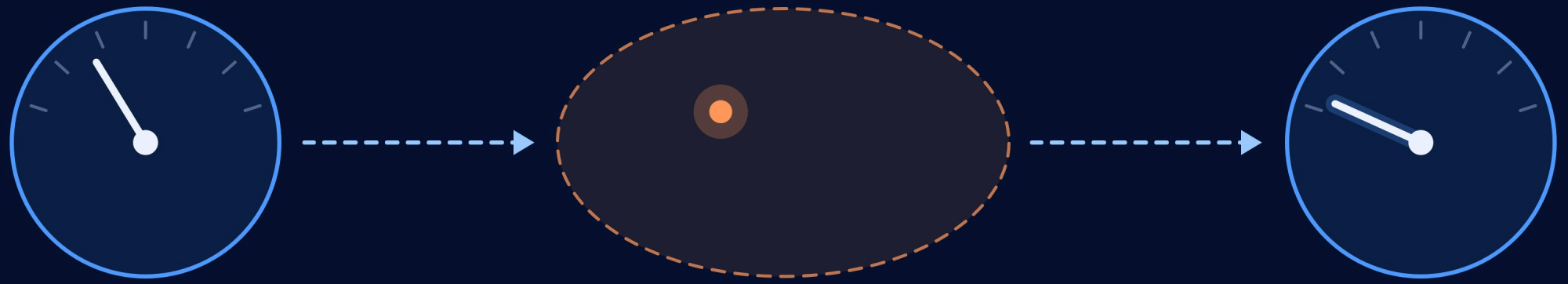
decide before the uncertainty is revealed

2 Uncertainty is revealed

anywhere inside an uncertainty set

3 Wait and see

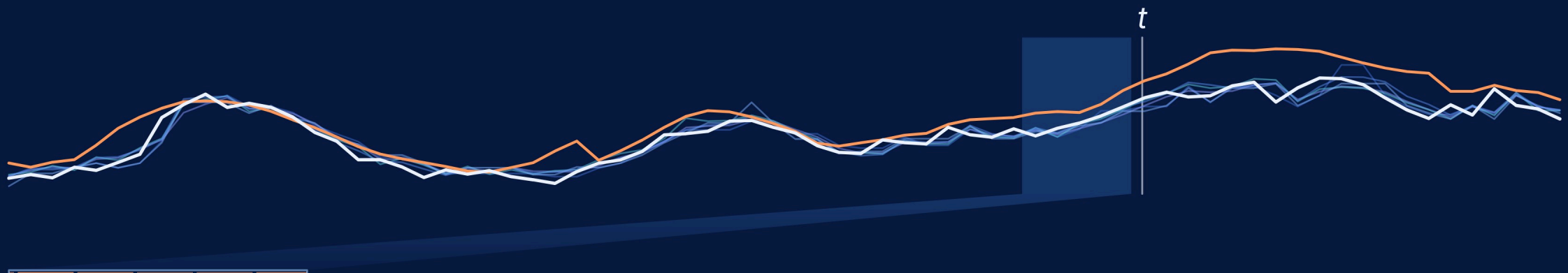
decide after, with a rule

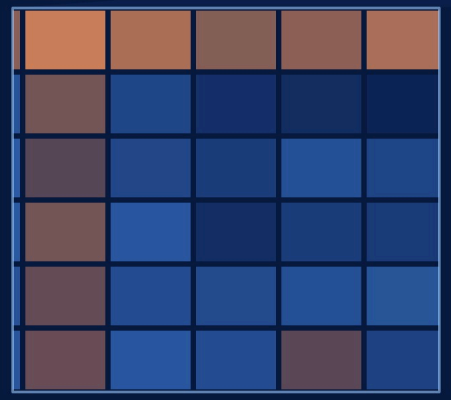


x z $y(z) = x_0 + V z$

Plan for the worst case, and let some decisions wait for what is revealed.

Weights as wait-and-see decisions





$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_0 + \mathbf{V}_0\,\mathbf{Z}_t$$

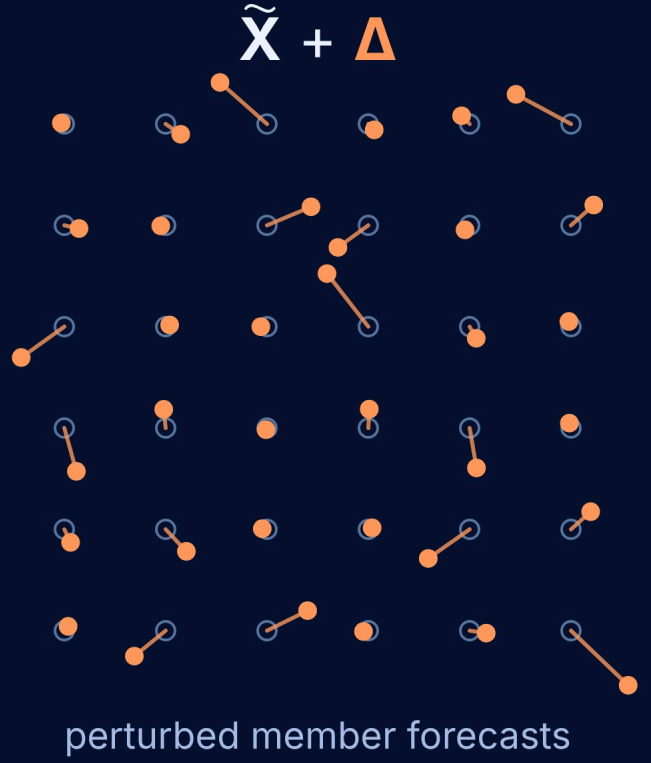


$$\hat{y}_t = \mathbf{X}_t^{\top}\boldsymbol{\beta}_t$$

errors of the last τ hours — fitted once — weights, every hour — forecast

The rule is fitted once. The weights move at every step.

Robust training is a regularized regression



perturbed member forecasts

$$\min_{\boldsymbol{\beta} \in \Omega_{\text{adapt}}} \max_{\|\boldsymbol{\Delta}\|_F \leq \lambda} \|\mathbf{y} - (\tilde{\mathbf{X}} + \boldsymbol{\Delta})\,\boldsymbol{\beta}\|_2$$

Theorem 1 $=$

$$\min_{\boldsymbol{\beta} \in \Omega_{\text{adapt}}} \|\mathbf{y} - \tilde{\mathbf{X}}\,\boldsymbol{\beta}\|_2 + \lambda\,\|\boldsymbol{\beta}\|_2$$

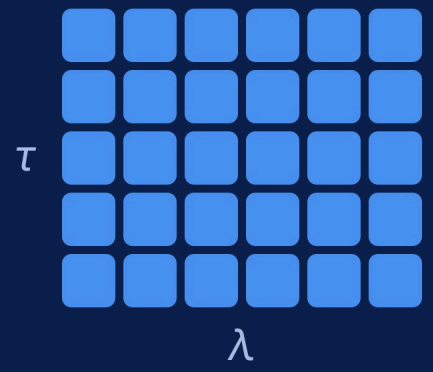
Robustness to perturbed forecasts becomes a ridge penalty on the weights.

Train once, adapt at every step

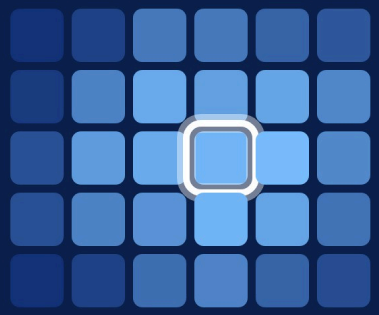


split in time order, no shuffling

1 Fit

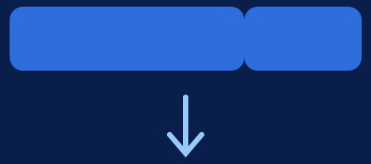


2 Select



lowest validation error

3 Refit



the rule

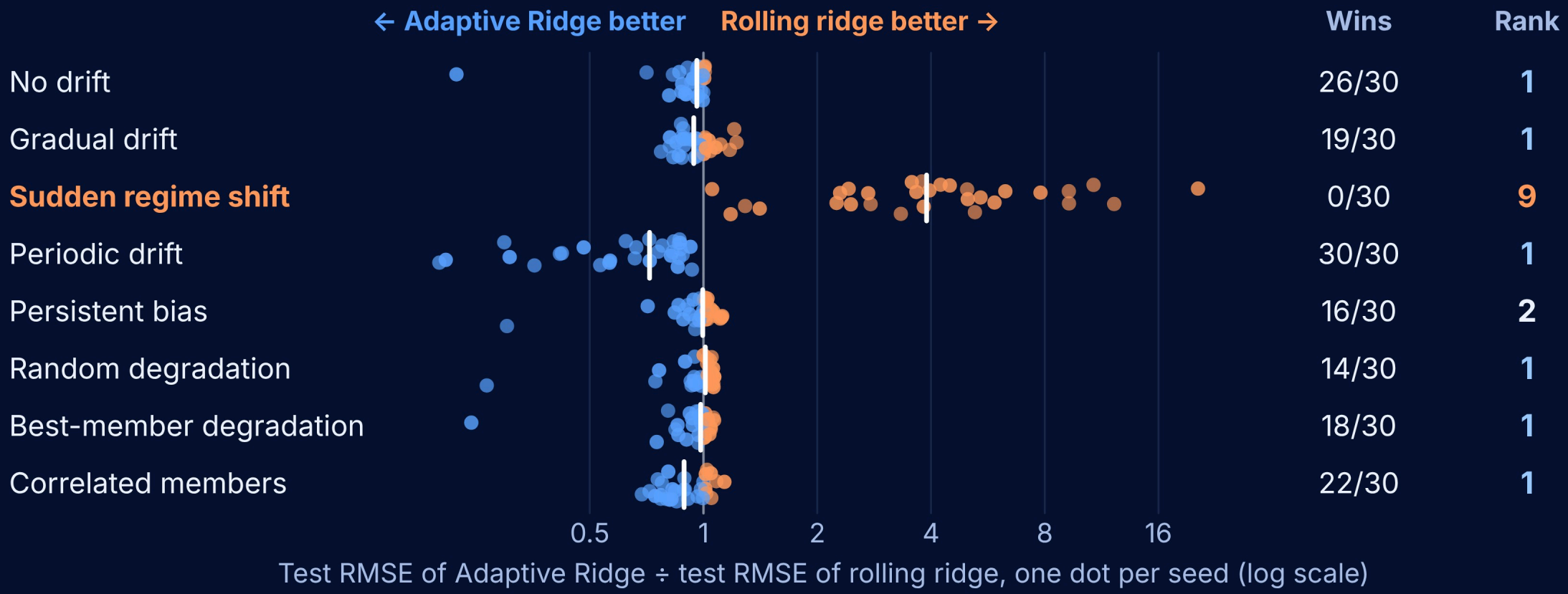
4 Deploy



weights at each step

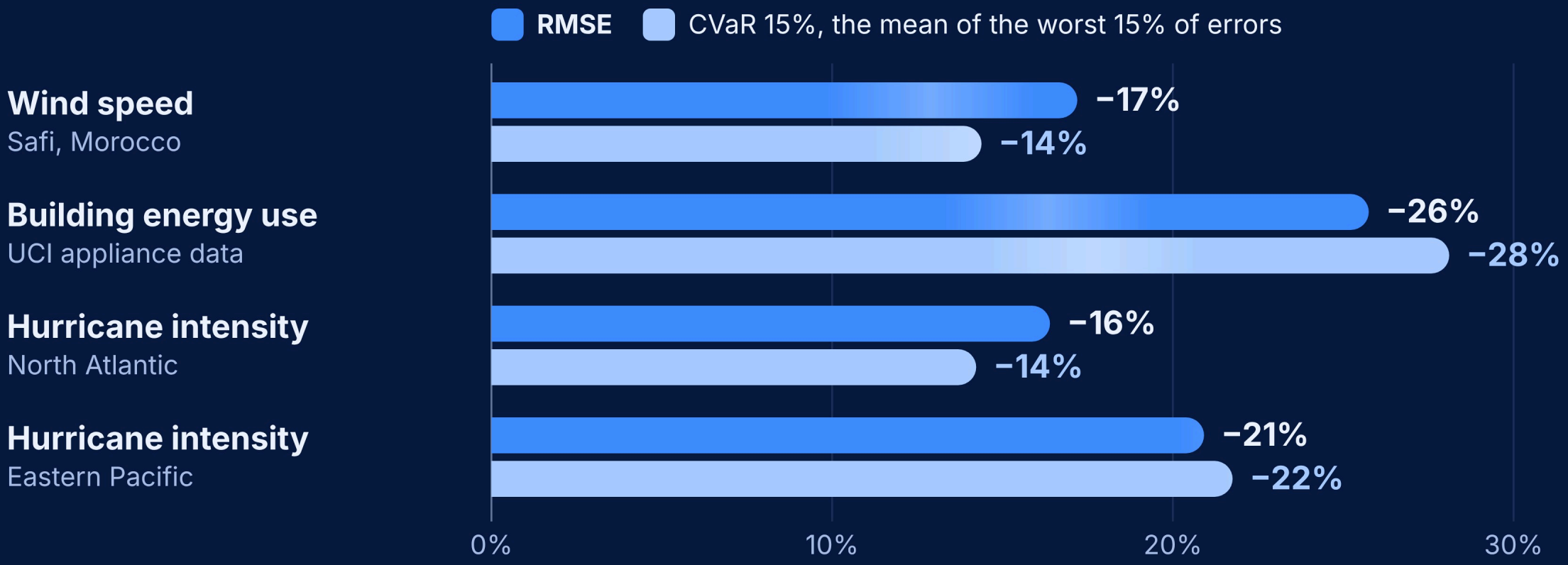
The whole pipeline takes about 0.1 seconds for 10 members and 3,000 observations.

Eight disturbances, 15 methods, 30 seeds



Best of 15 methods in six of eight scenarios. The sudden shift is the failure case.

Three real-world forecasting tasks



16–26% lower RMSE and 14–28% lower CVaR 15% than the best member in hindsight.

Lower error than the official hurricane forecast



North Atlantic

10.2 kt

against 11.4 kt for the official forecast

Eastern Pacific

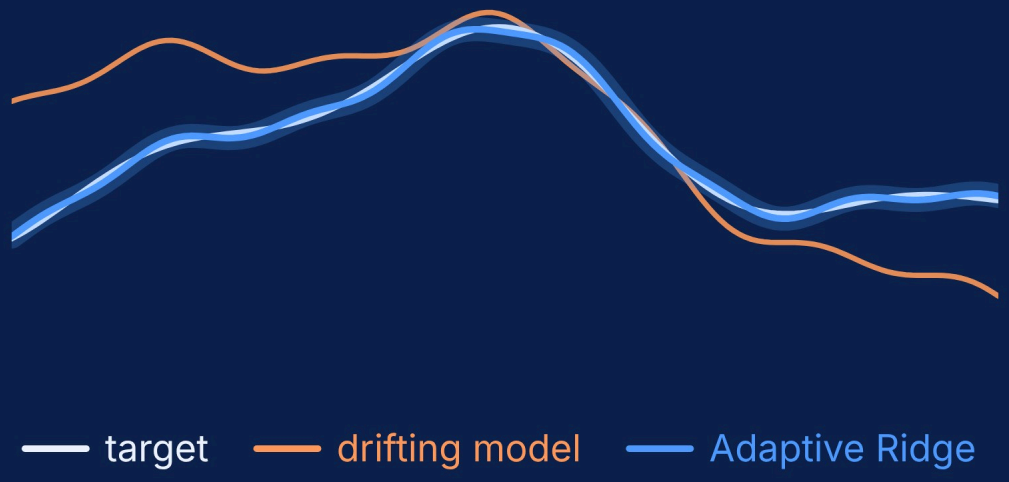
12.1 kt

against 15.6 kt for the official forecast

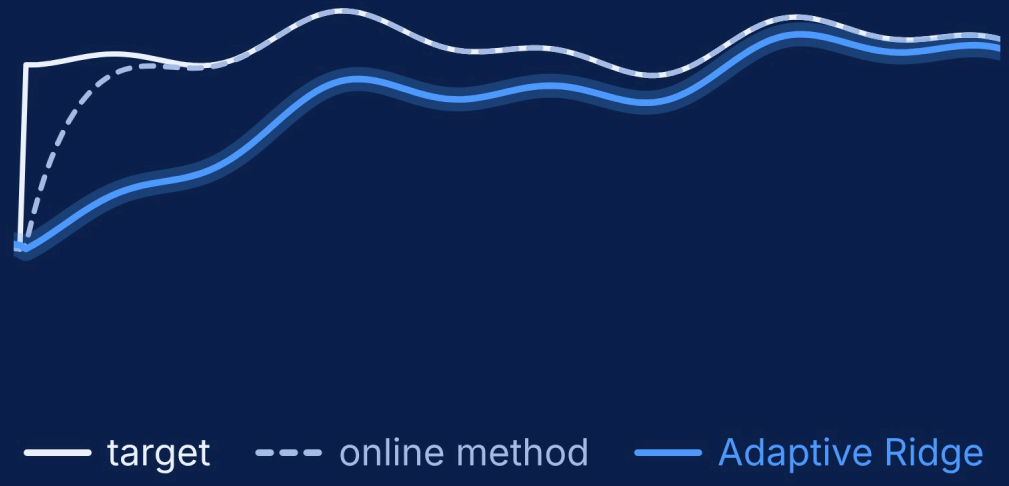
Adaptive Ridge RMSE for 24-hour intensity, against the official NHC forecast.

When to use it

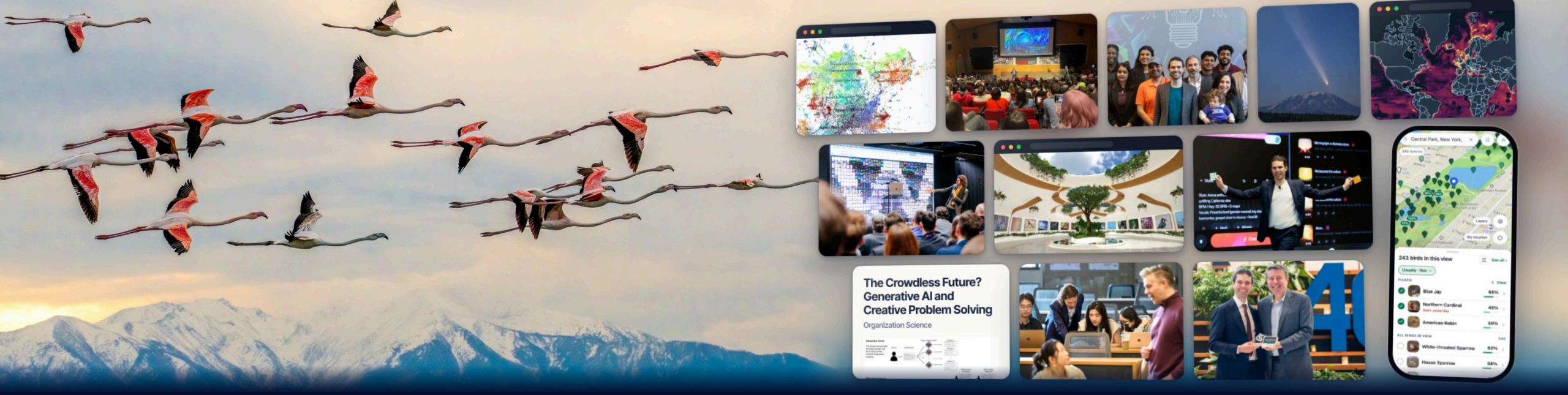
Gradual or periodic drift



Sudden, never-seen regime



With 10 members, it needs about 1,000 observations to match the best online methods.



Looking forward to connecting!

leobix@uw.edu · boussioux.com

Paper

Code

Paper on arXiv. Code at github.com/leobix/adaptive_ensemble.