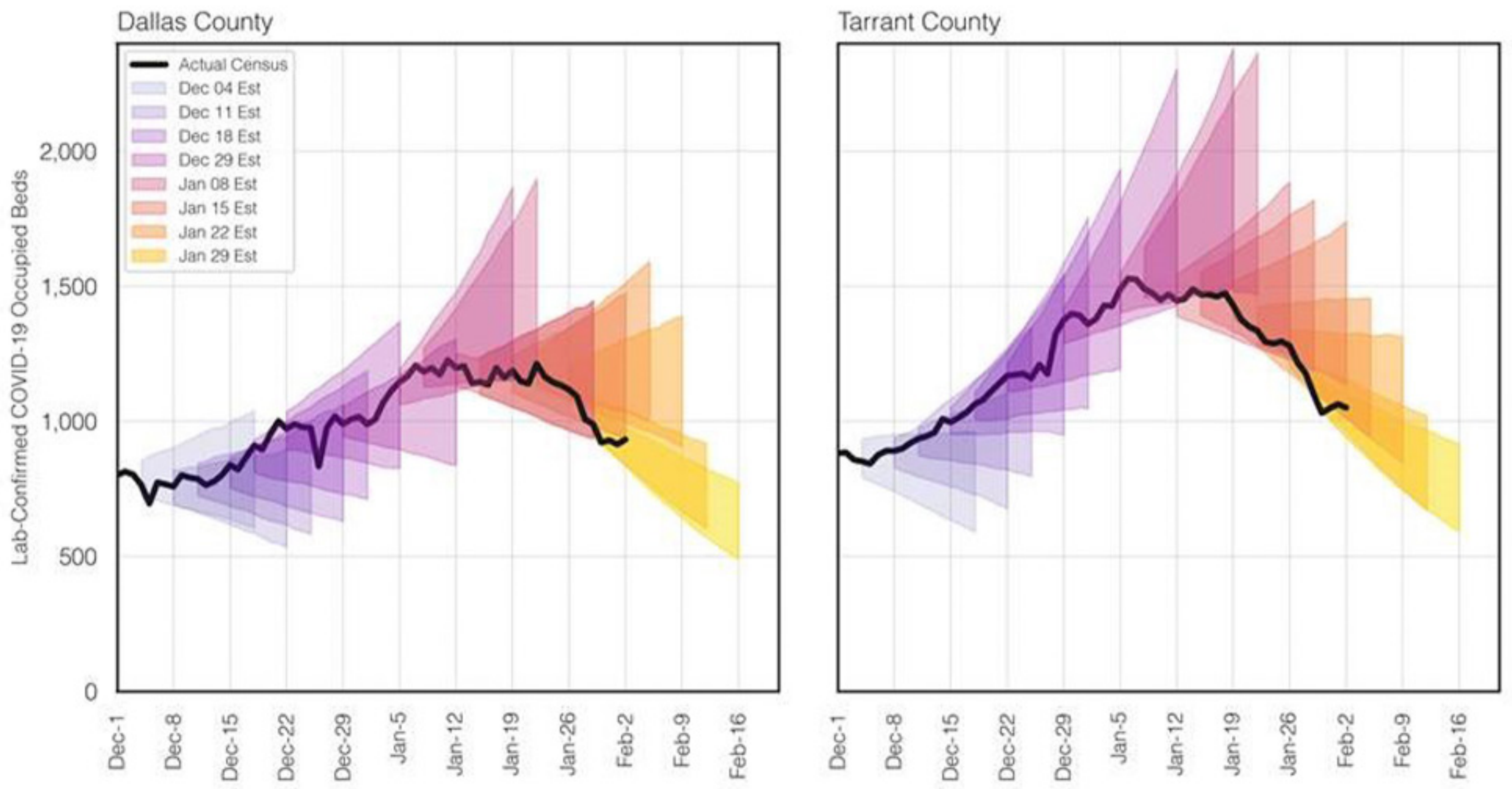
Copula Conformal Prediction for Multi-step Time Series Forecasting

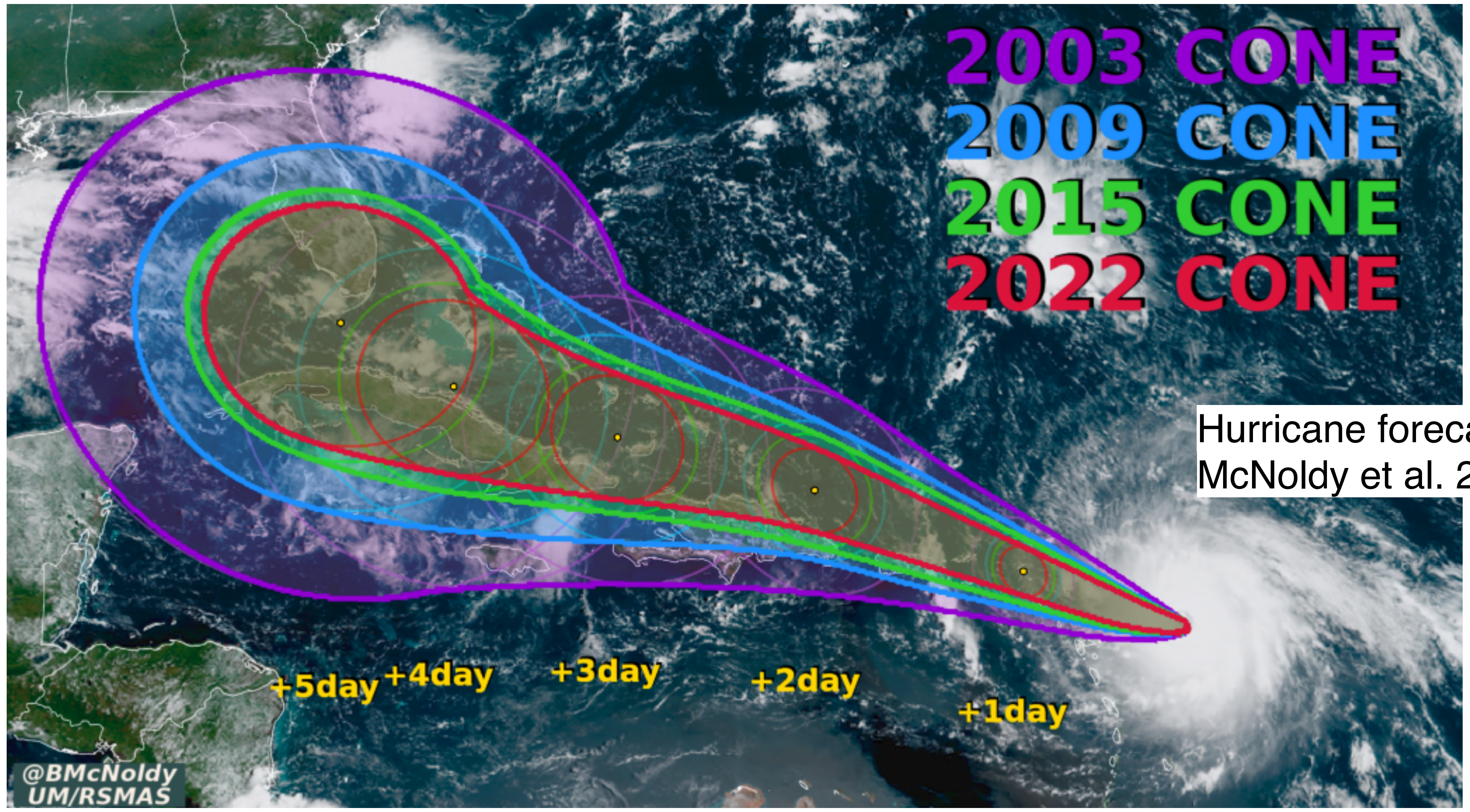
Sophia Sun, Rose Yu

ICLR 2024 Poster 17807

Uncertainty Quantification for Time Series



Covid Forecasts. Patrick McGee / UT Southwestern 2021



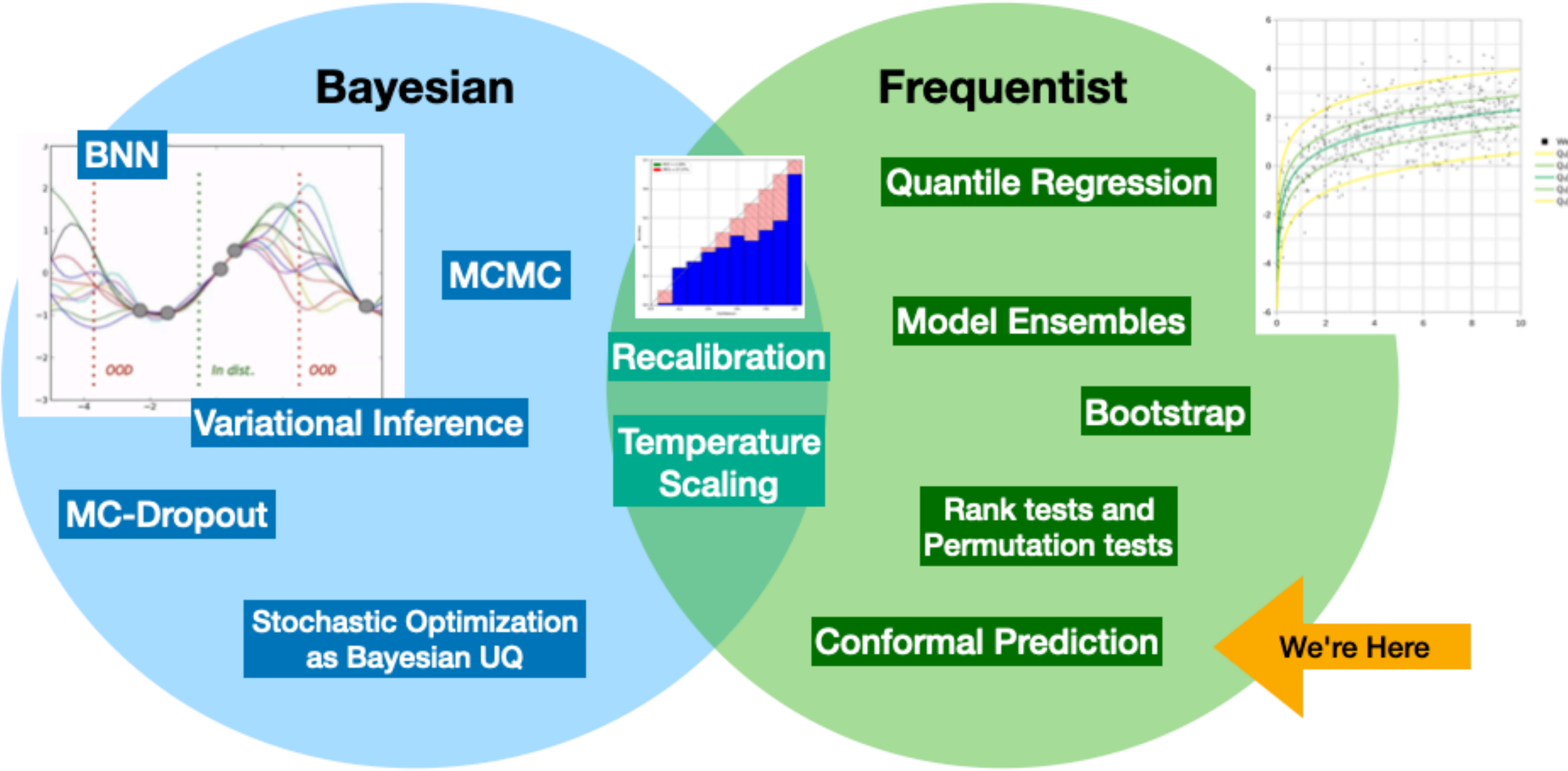
Hurricane forecasting
McNoldy et al. 2022

Goal: "Cone of uncertainty" valid for all time steps of $\mathbf{y}$

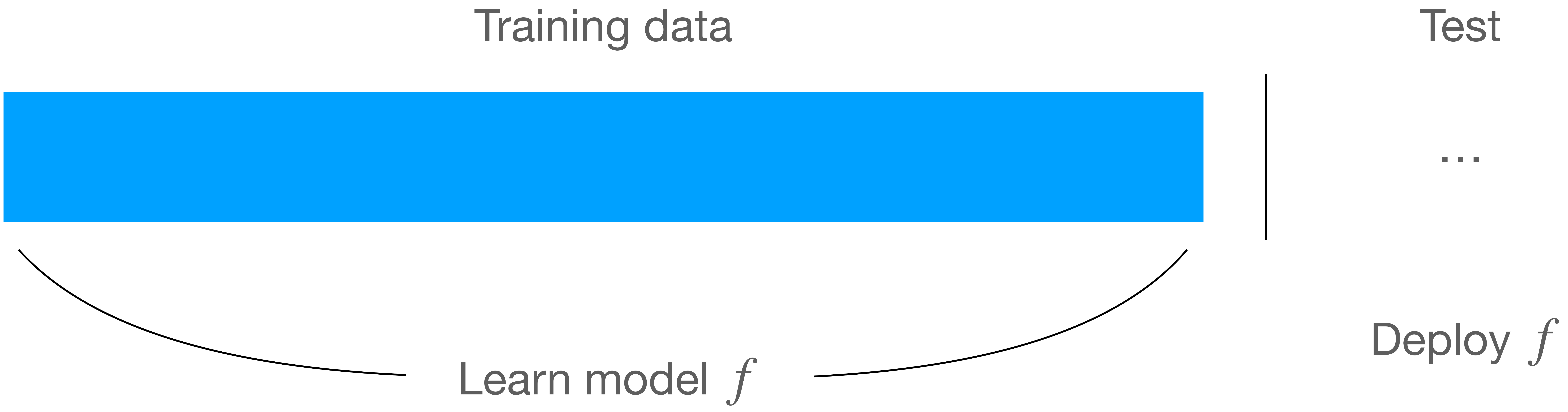
$$\mathbb{P}[\, \forall h \in \{1,\ldots,k\}, \mathbf{y}_{t+h} \in \Gamma_h^{1-\alpha} \,] \geq 1 - \alpha$$

Conformal prediction

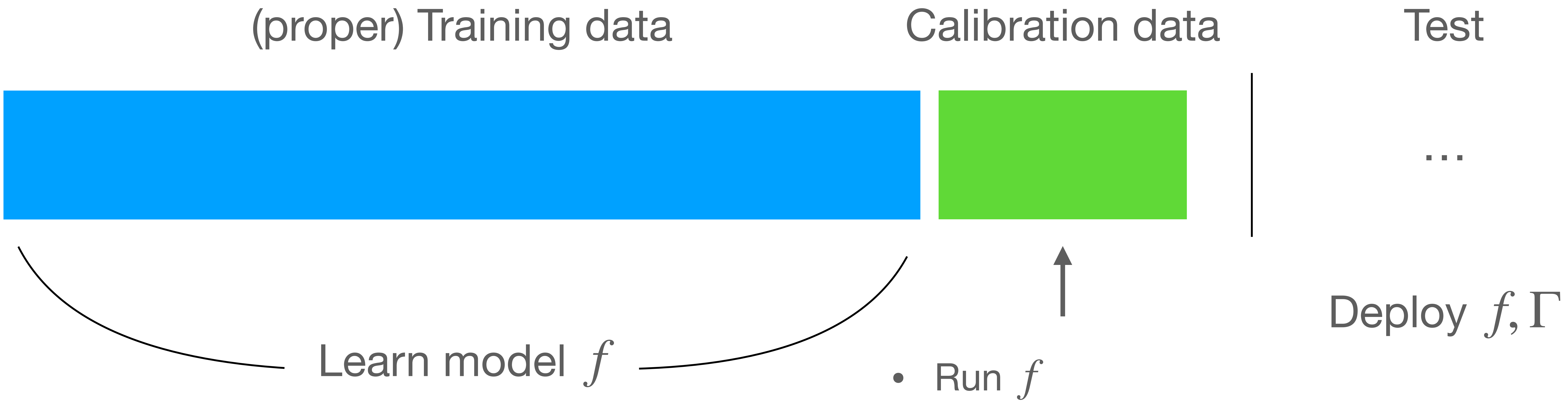
- Works for ***any*** underlying prediction model
- Works for ***any*** underlying data distribution
- Guaranteed validity in finite samples
- (Under mild assumptions)



Overview of Conformal Prediction

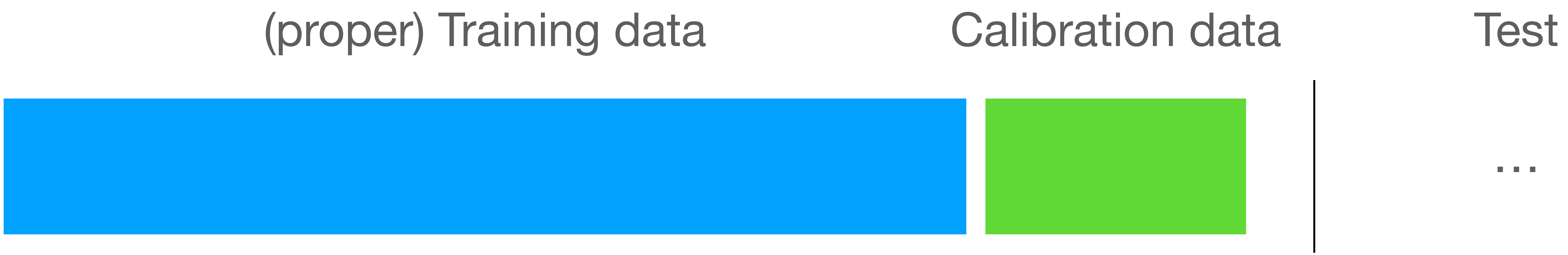


Overview of Conformal Prediction



- Run f
- Get nonconformity scores e.g. $||y - f(x)||$
- Calibrate confidence region Γ with scores

Overview of Conformal Prediction

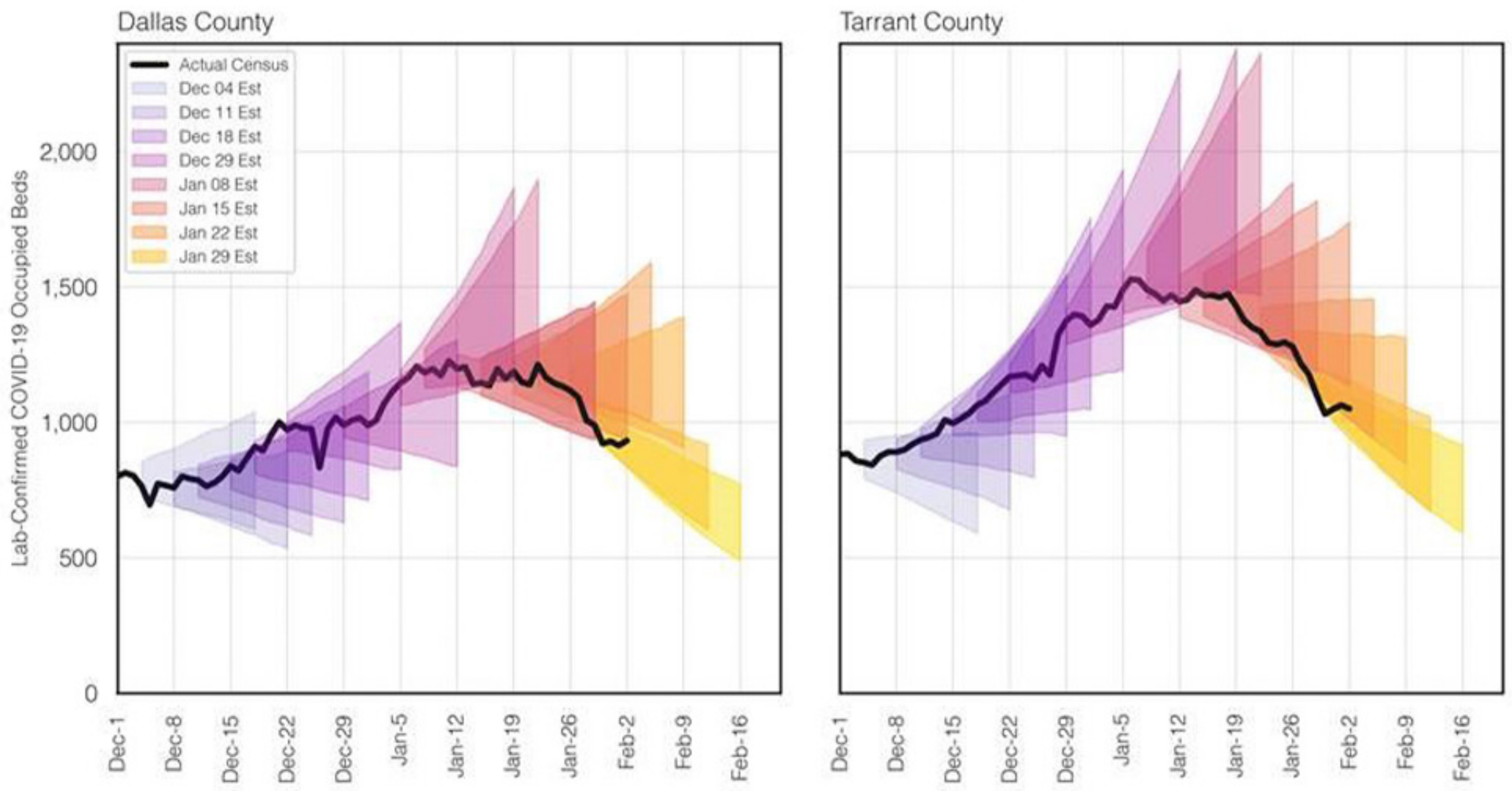


Definition 1: Validity Guarantee

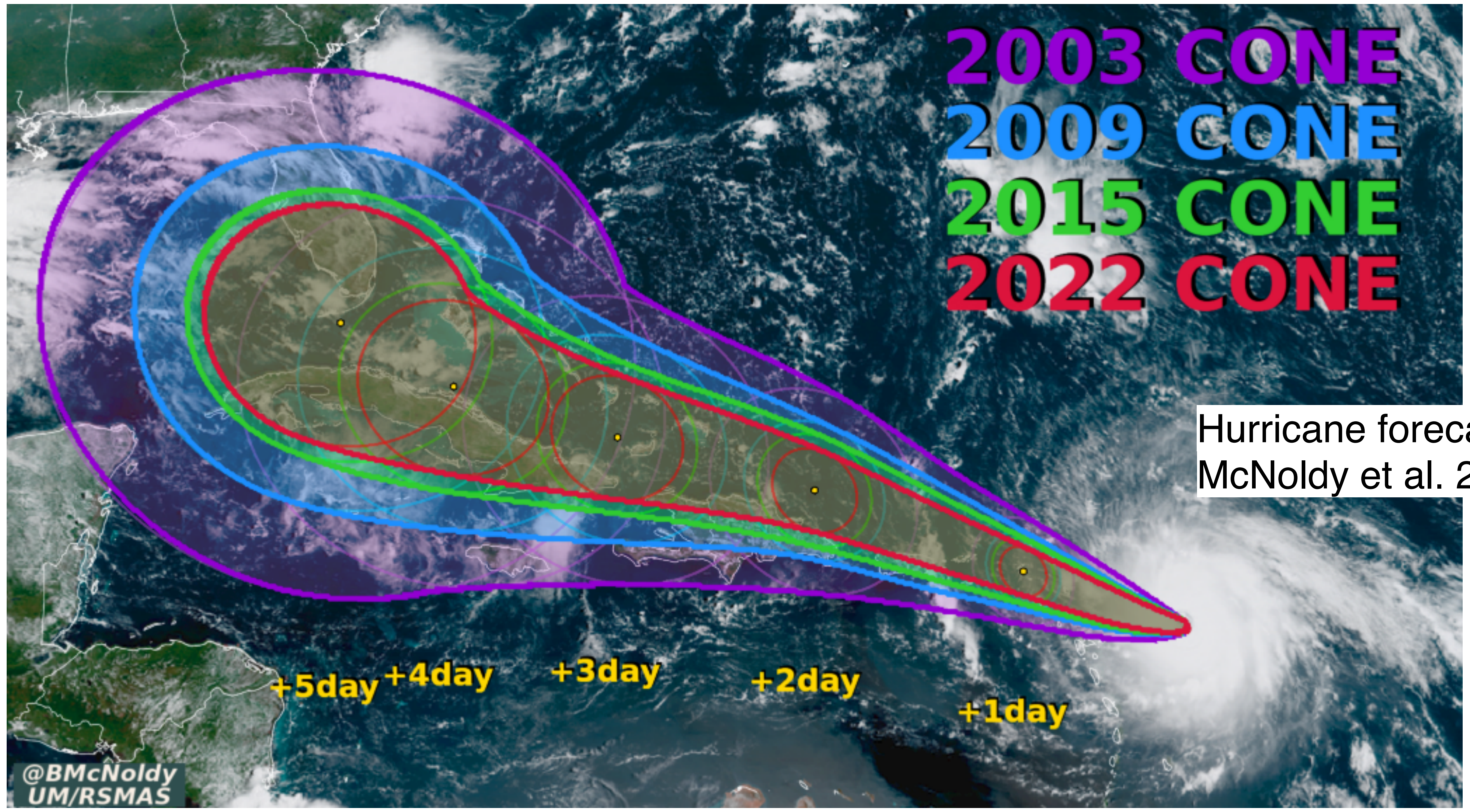
Given a test data pair (X, Y) and a desired coverage rate $1 - \alpha \in (0,1)$, a confidence region $\Gamma^{1-\alpha} : \mathbf{X} \to \{\text{subsets of } \mathbf{Y}\}$ is *valid.*

$$\mathbb{P}(Y \in \Gamma^{1-\alpha}(X)) \geq 1 - \alpha$$

UQ for Time Series Setting



Covid Forecasts. Patrick McGee / UT Southwestern 2021



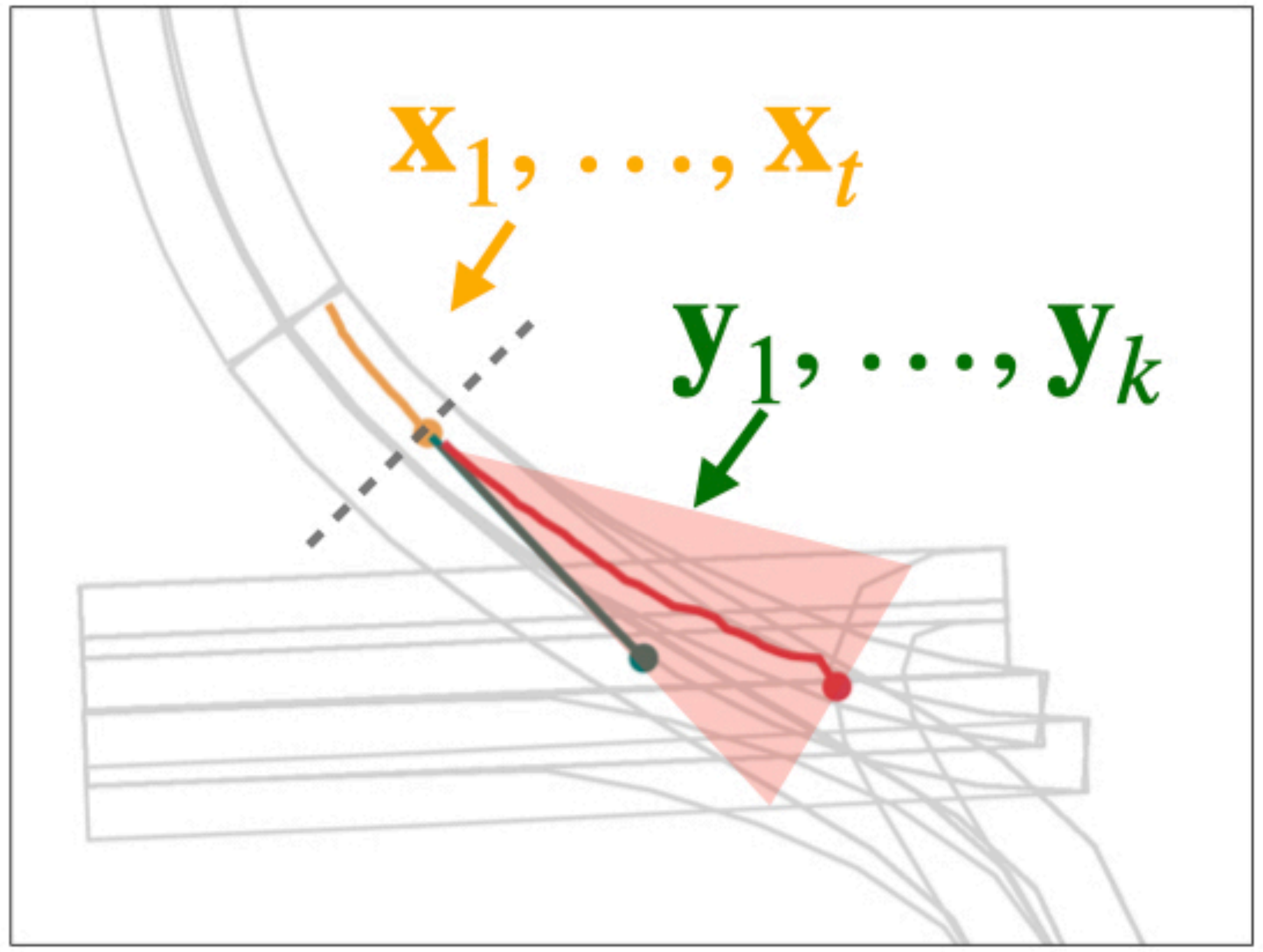
Hurricane forecasting
McNoldy et al. 2022

Goal: "Cone of uncertainty" valid for all time steps of $\mathbf{y}$

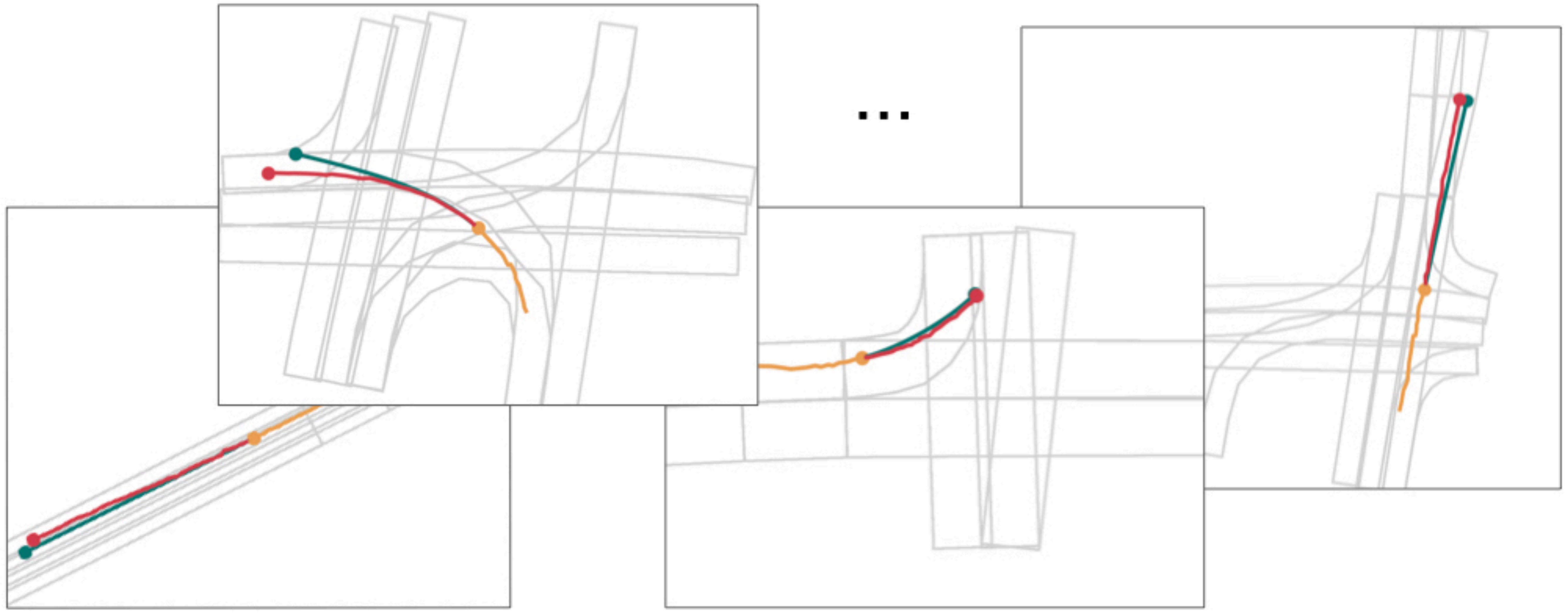
$$\mathbb{P}[\ \forall h \in \{1,\ldots,k\}, \mathbf{y}_{t+h} \in \Gamma_h^{1-\alpha}\ ] \geq 1-\alpha$$

UQ for Time Series Setting

Time steps: temporal dependence (non-i.i.d.)



Dataset $\mathscr{D}$: independent observations (i.i.d.)



Dataset $\mathscr{D} = \{(\mathbf{x}^{(i)}_{1:t}, \mathbf{y}^{(i)}_{t+1:t+k}\}^{n}_{i=1}$

Goal: "Cone of uncertainty" valid for all time steps of $\mathbf{y}$

$$\mathbb{P}[\ \forall h \in \{1,\ldots,k\}, \mathbf{y}_{t+h} \in \Gamma_h^{1-\alpha}\ ] \geq 1 - \alpha$$

Producing k-step-valid confidence regions

Naive solution: Bonferroni correction (aka. Union bounding) [2]

Find the $(1 - \alpha/k)$ confidence region for each of the k time steps.

Then by Boole's inequality:

$$\mathbf{P}[\ \exists h \in \{1, \ldots, k\}, \mathbf{y}_{t+h} \notin \Gamma_h^{1-\alpha}\ ] \leq \alpha$$

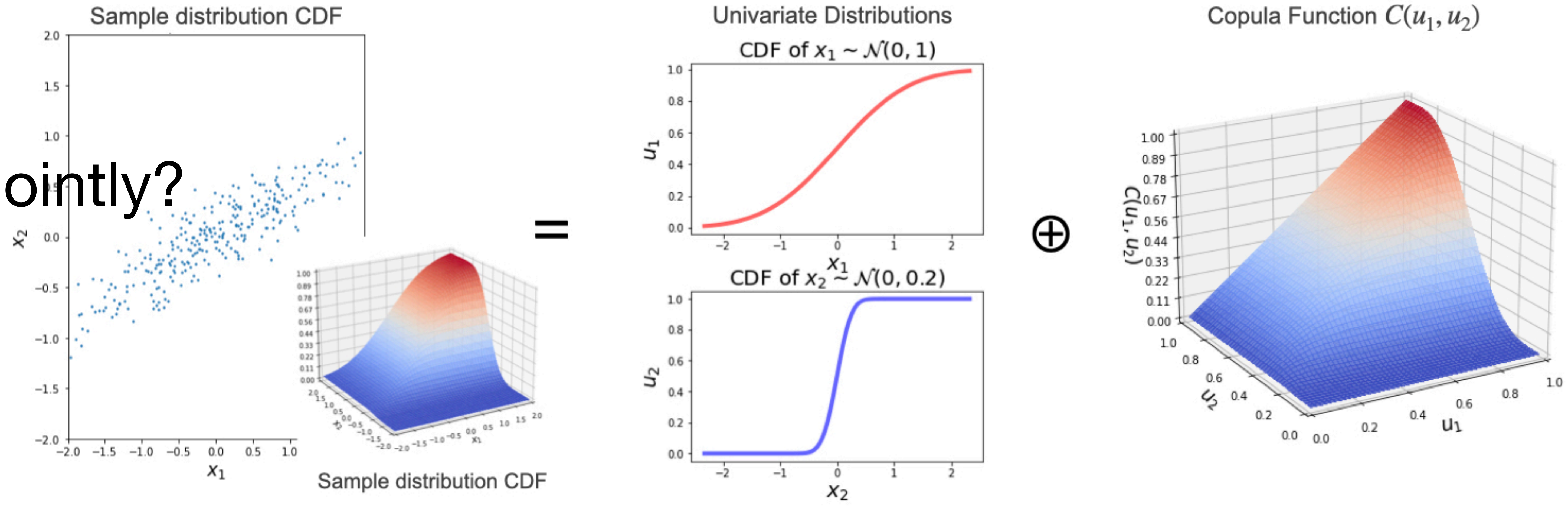
Problem: when k is large, this is very inefficient.

[2] Stankeviciute, Kamile, Ahmed M Alaa, and Mihaela van der Schaar. "Conformal time-series forecasting." *NeurIPS,* 2021

Copula Conformal Prediction for Time Series

How can we model the distributions jointly?

Idea: **Copulas**



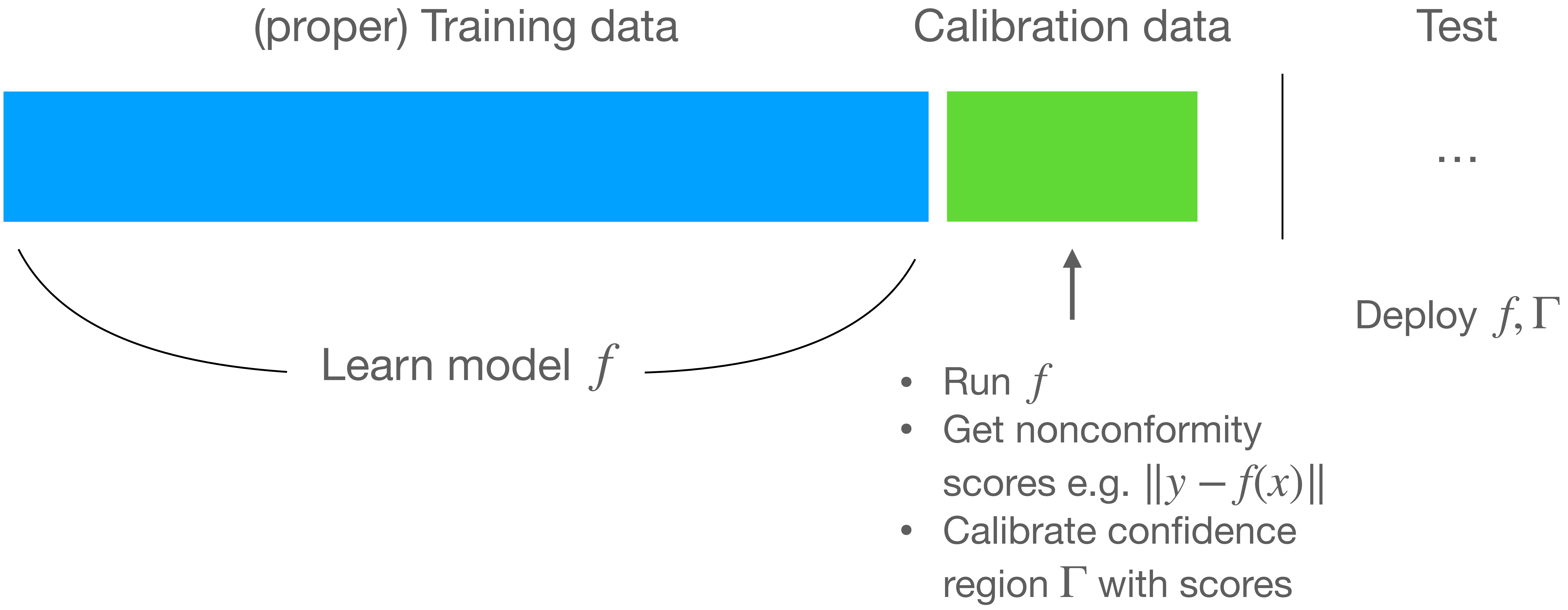
A copula is a function that synthesizes multiple CDFs to a joint CDF

$$C(u_1, \cdots, u_k) = \mathbb{P}(U_1 \leq u_1, \cdots, U_k \leq u_k)$$

$$F(x_1, \cdots, x_k) = C(F_1(x_1), \cdots, F_k(x_k))$$ (Sklar's theorem)

For joint coverage guarantees, we only have to bound the Copula.

Copula Conformal Prediction

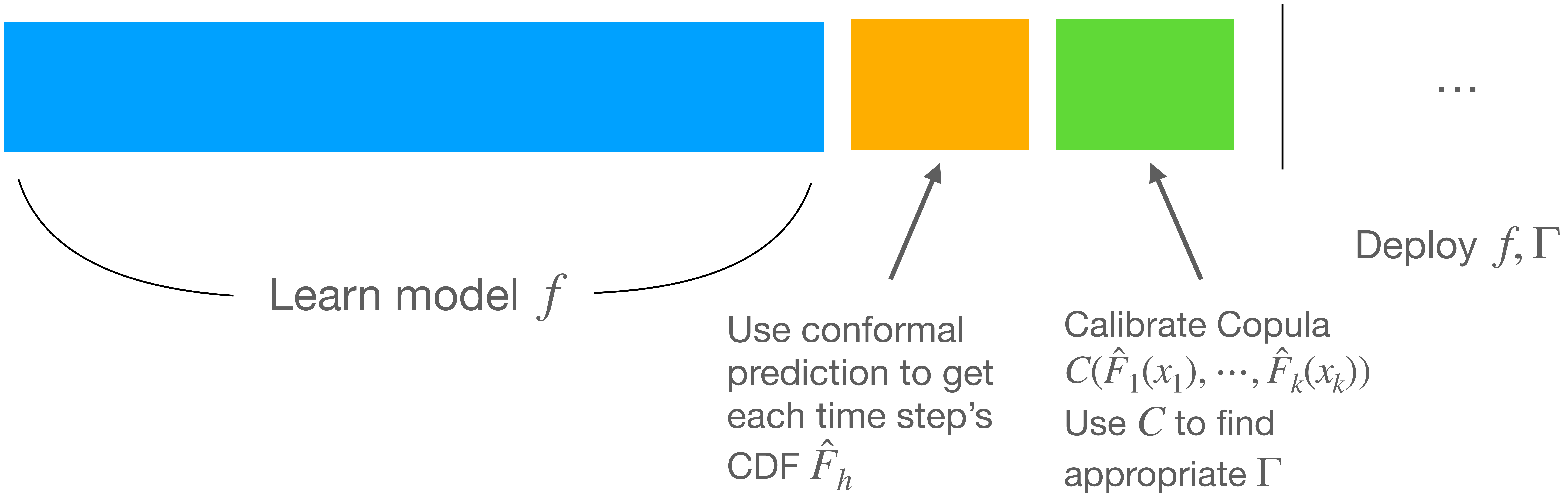


Copula Conformal Prediction

We prove that it also has finite-sample
tee

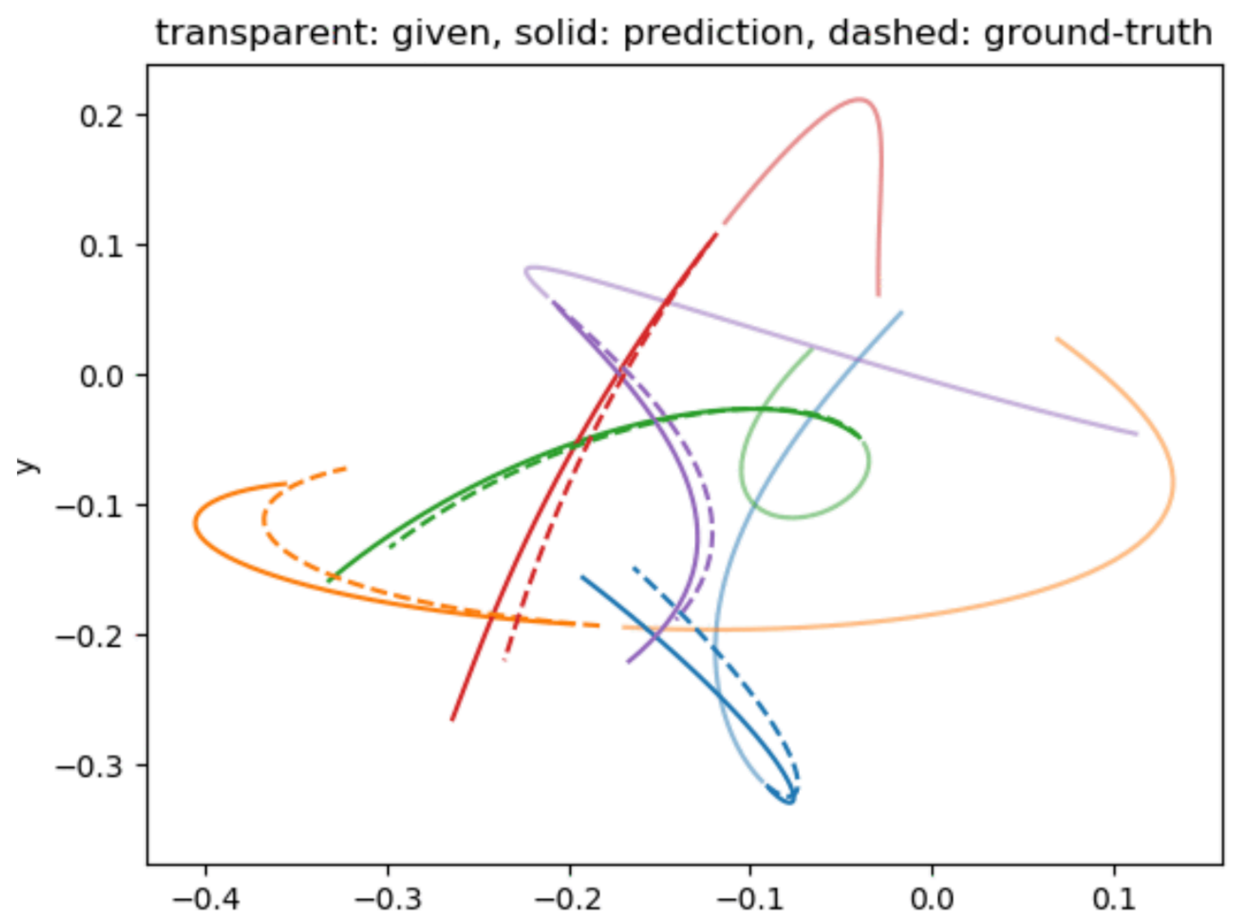
Theorem 4.1 (Validity of CopulaCPTS). *CopulaCPTS (algorithm 1) produces valid confidence regions for the entire forecast horizon. i.e.*

$$\mathbb{P}[\,\forall j \in \{1,\ldots,k\},\ y_j \in \Gamma_j^{1-\alpha}\,] \geq 1-\alpha.$$

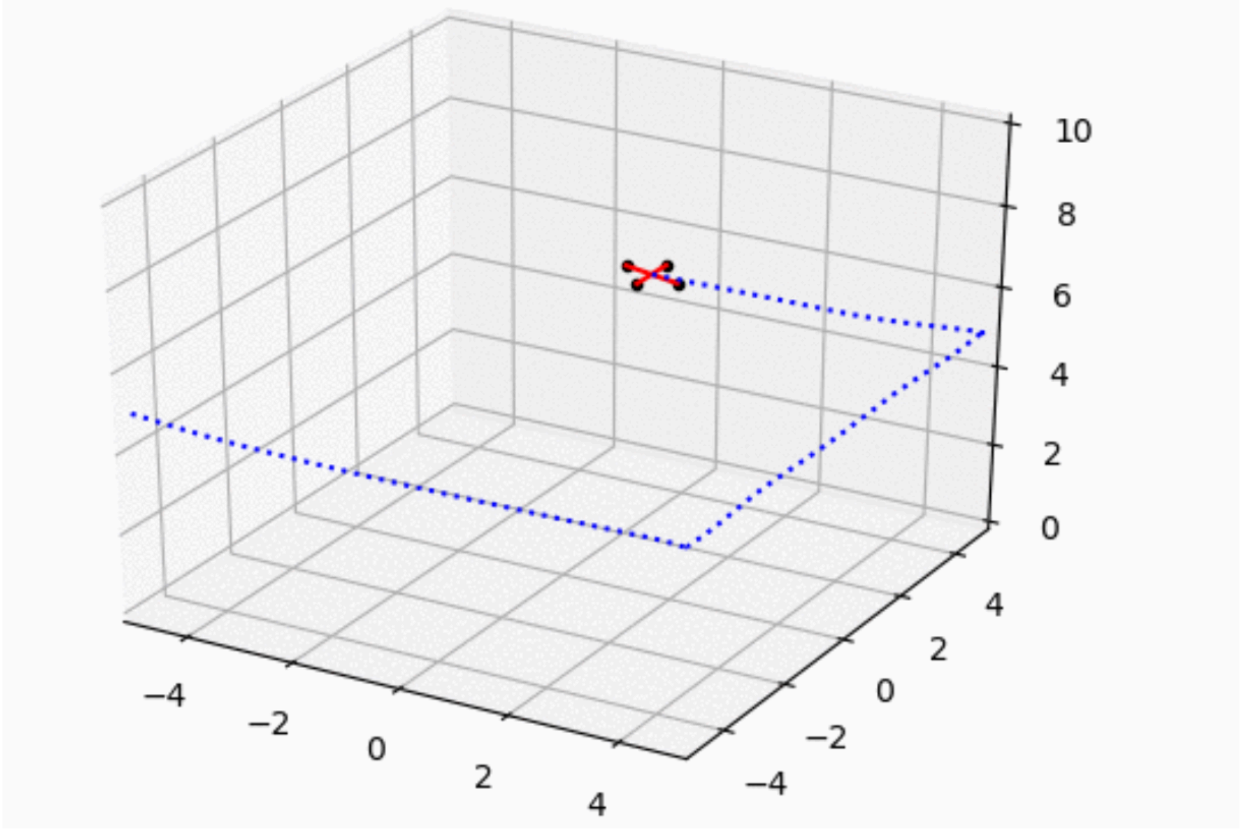


Experiment datasets

Synthetic

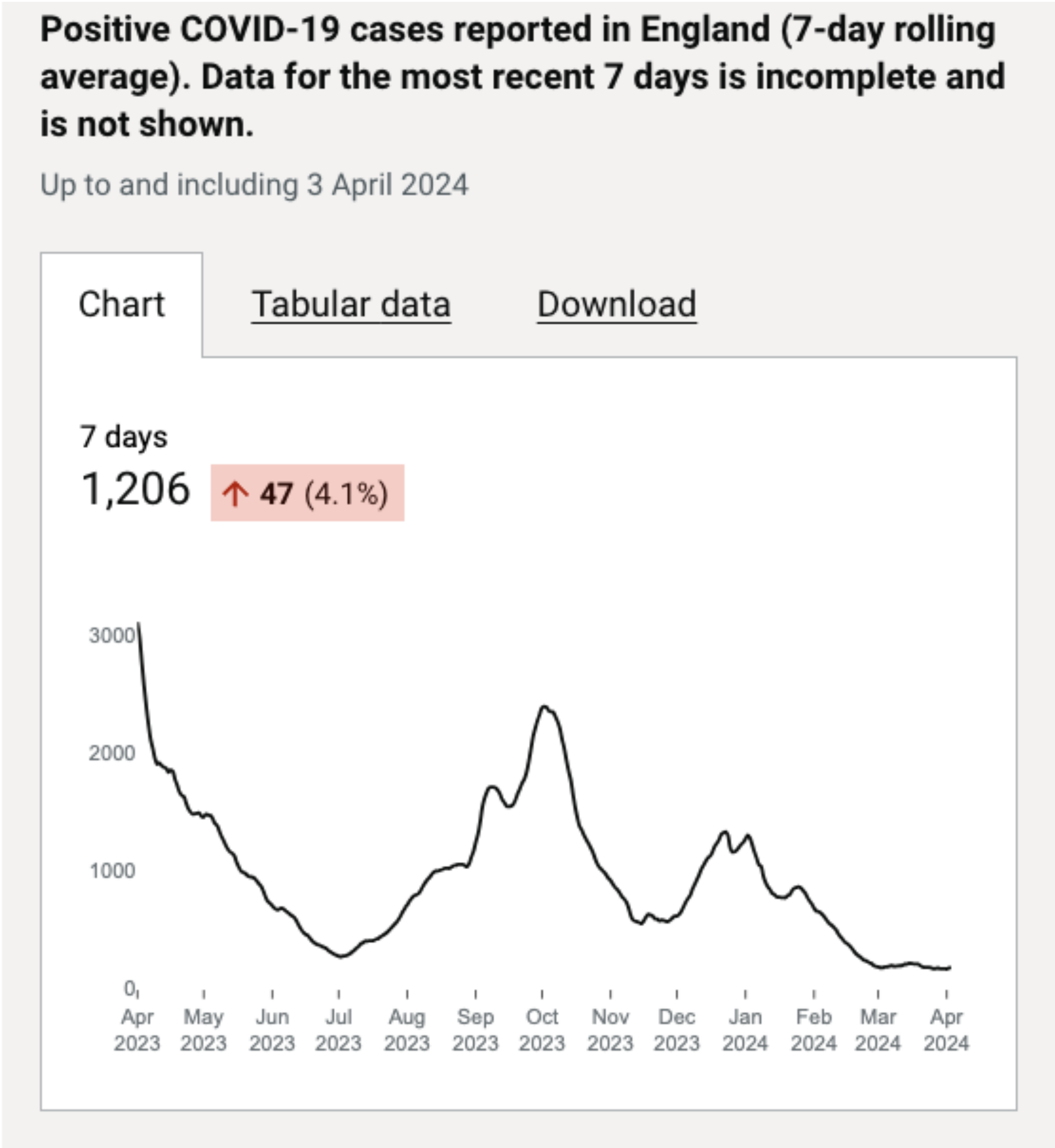


Interacting particles + Gaussian noise

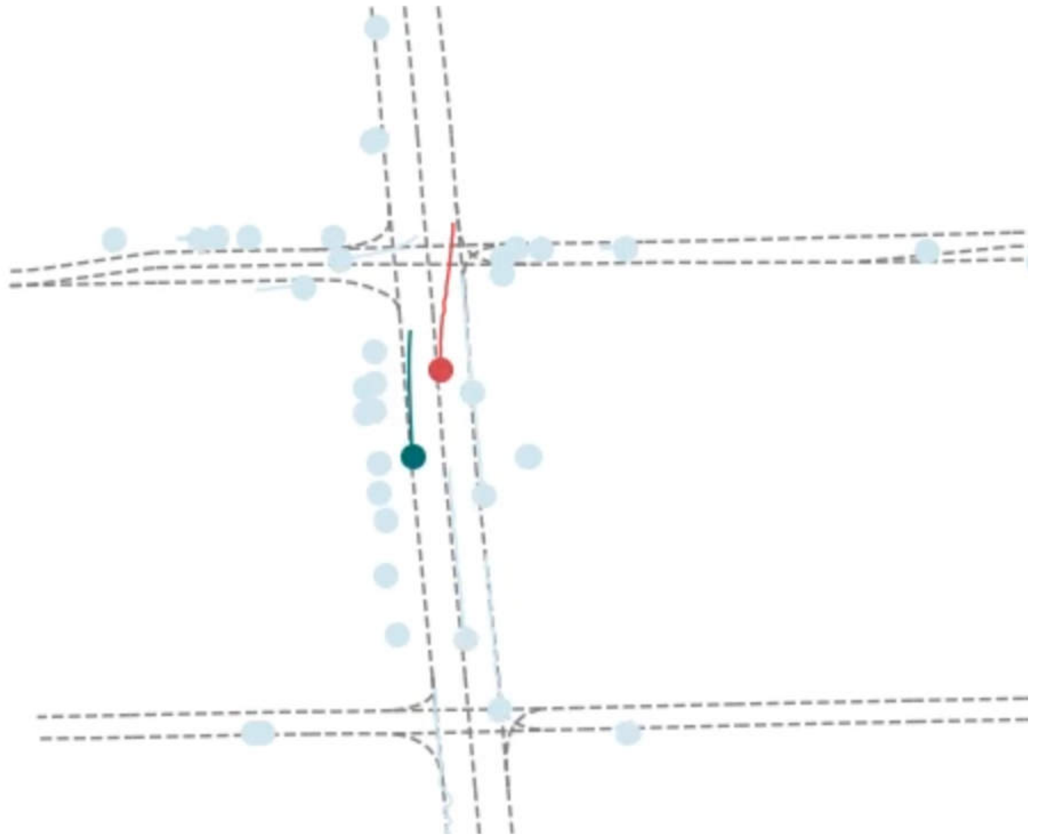


Drone trajectory following + Gaussian noise

Real-world

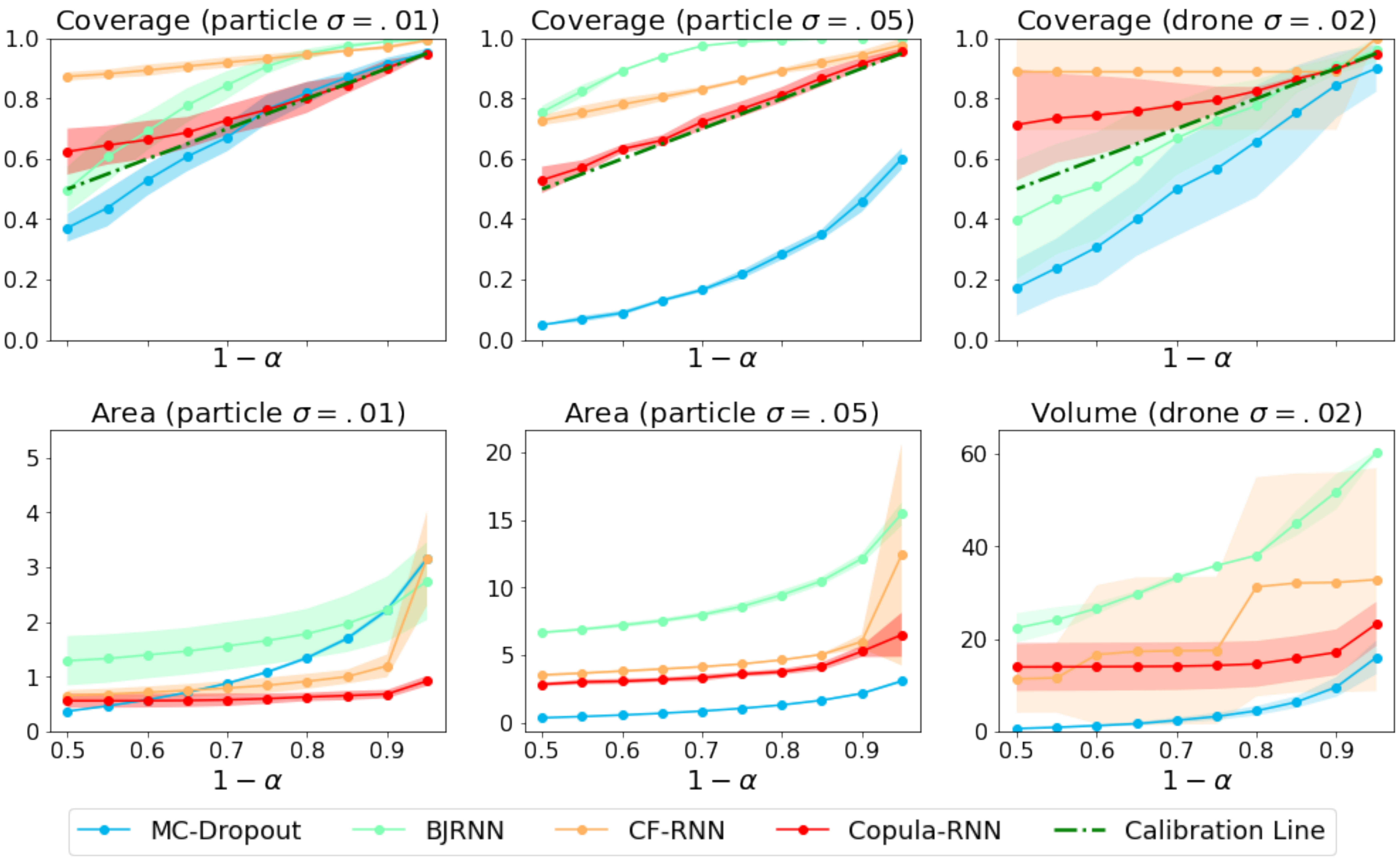


New Covid cases
(380 regions, 500 days)

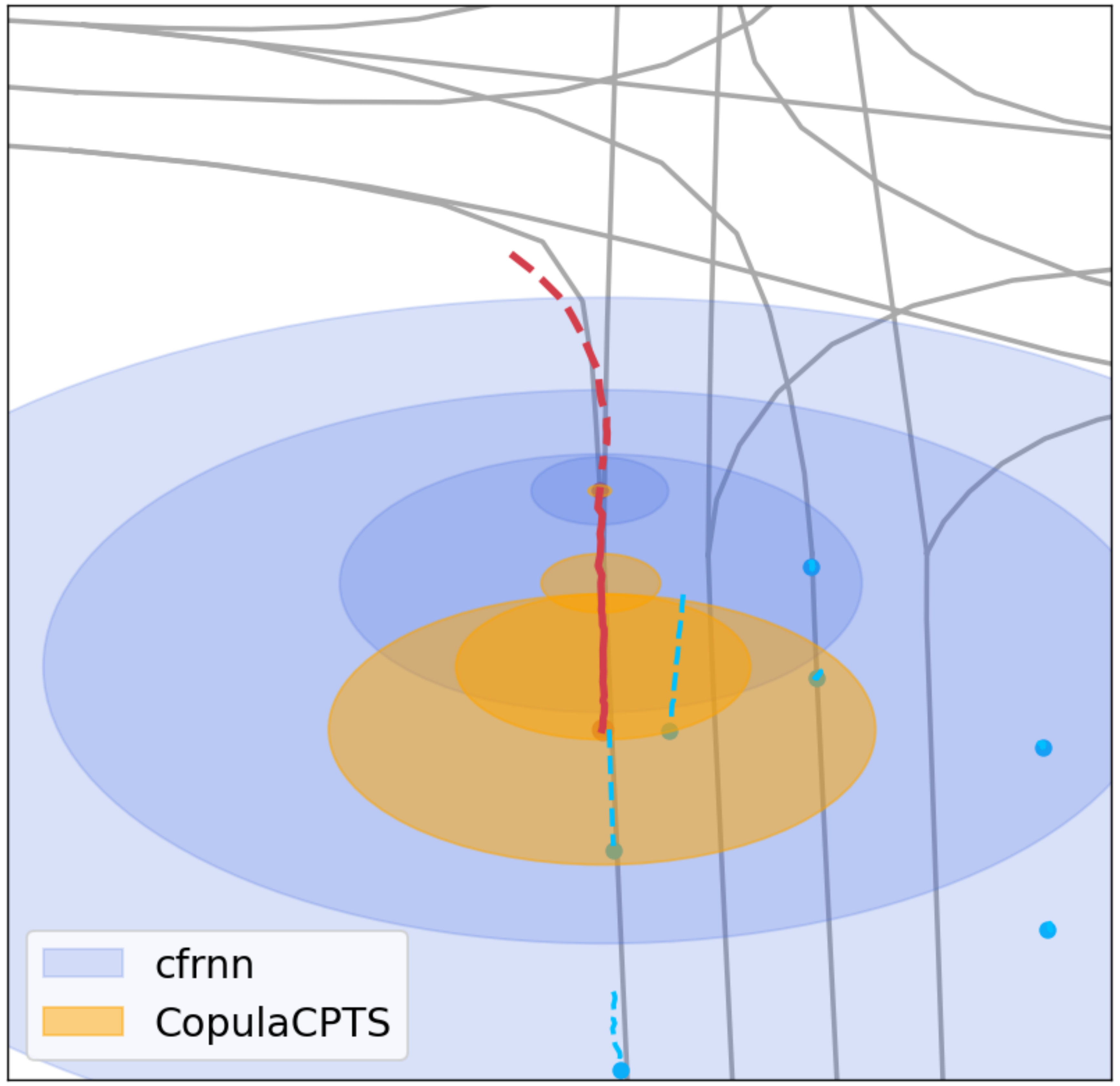


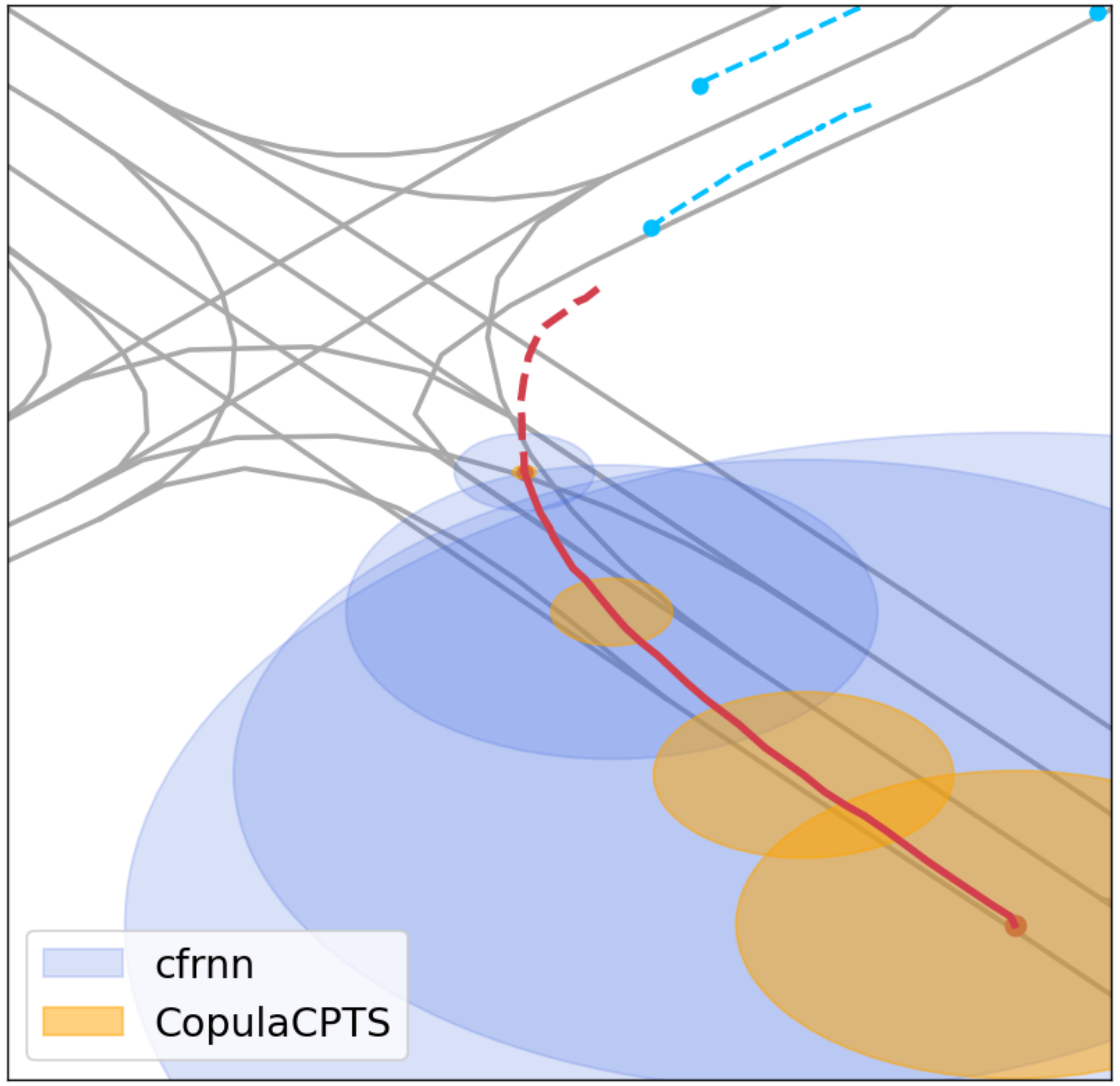
Argoverse Motion
Forecast Dataset

Results: calibration and sharpness



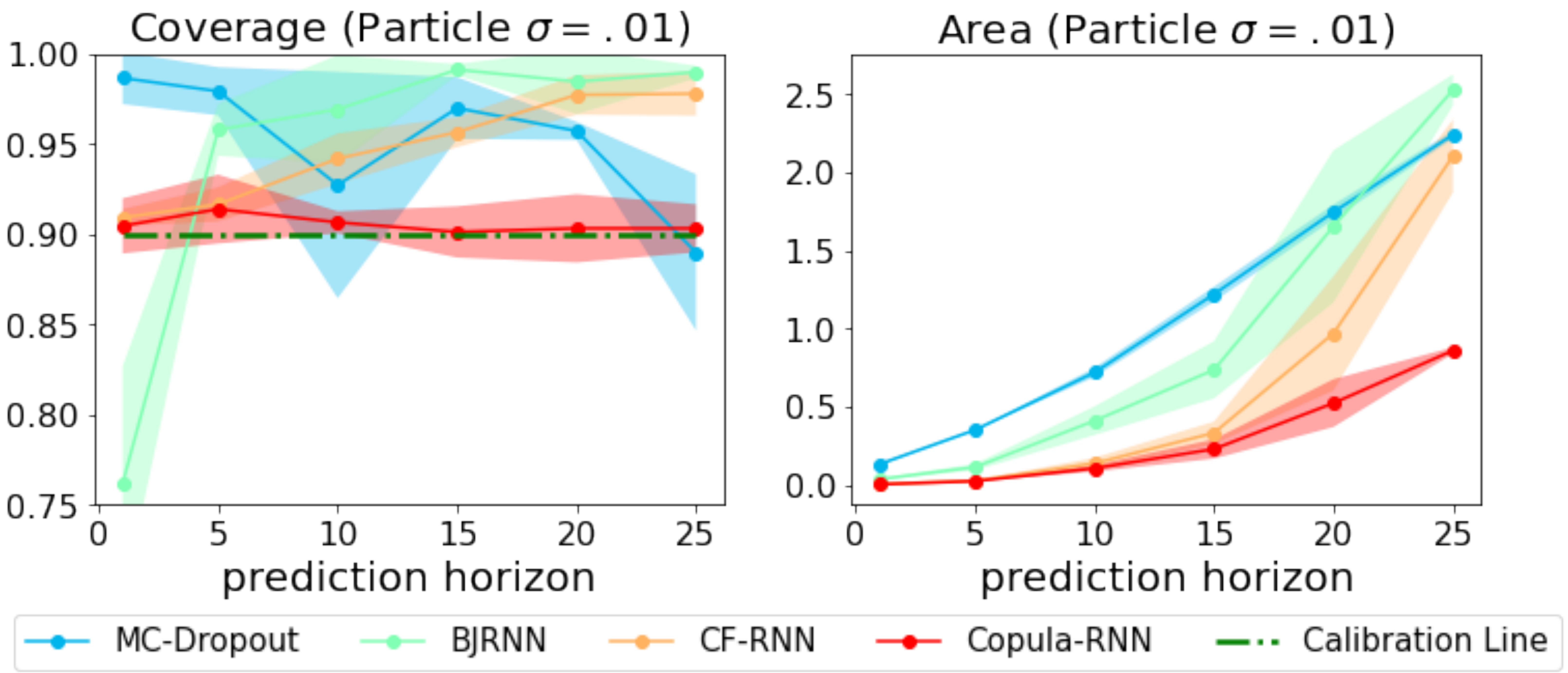
Results: examples



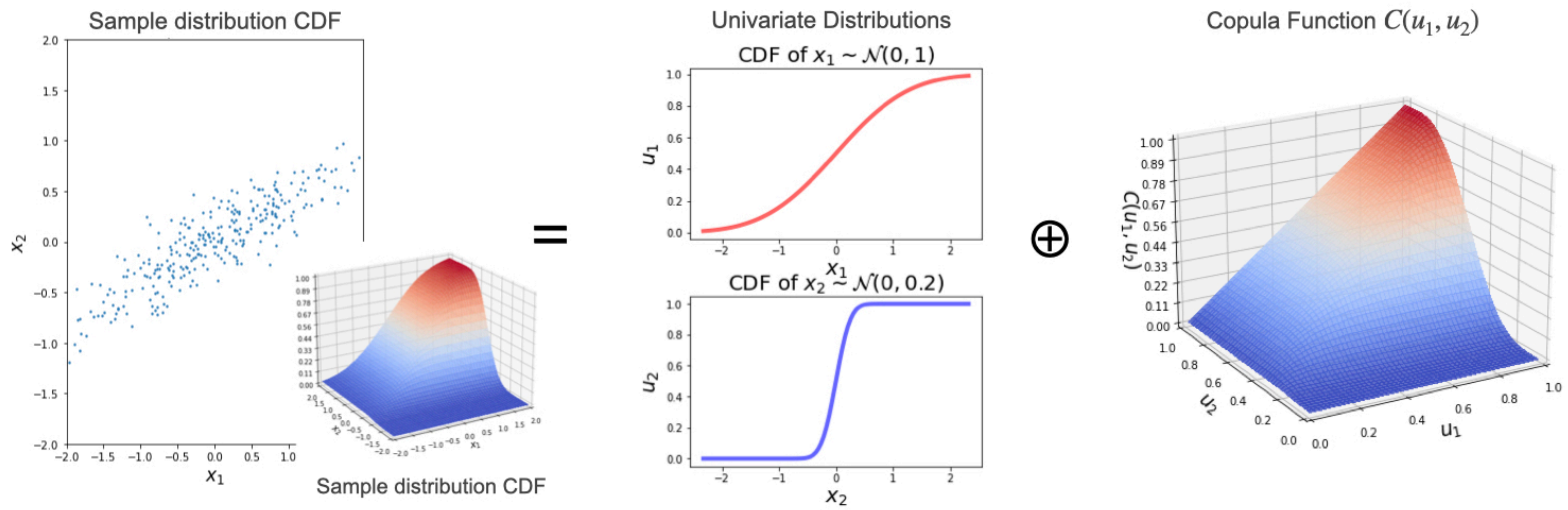


Argoverse autonomous vehicles dataset

Results: Over different horizon length



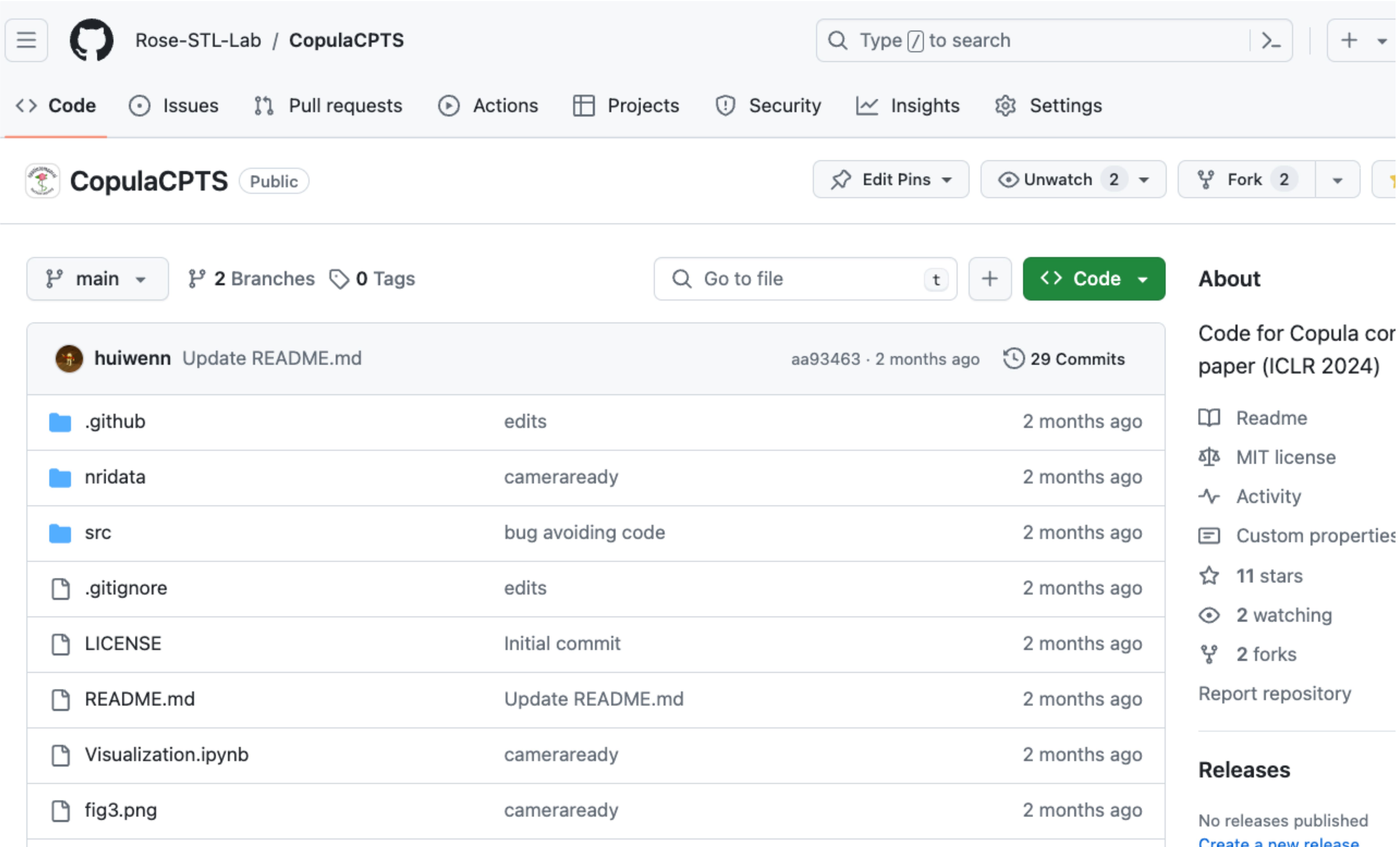
Takeaway



- We can use **copula functions** to better **capture uncertainty propagation through time.**
- At the cost of having an additional copula calibration dataset, we achieve **much sharper confidence intervals**.
- Improvement is more significant for **higher dimension** data and l**onger prediction horizons**.

Code is open sourced at

github.com/Rose-STL-Lab/CopulaCPTS



Thank you!

If you'd like to chat or collaborate,
reach me at shs066@ucsd.edu or on X @huiwensun_
:)