Study of Uncertainty Quantification through Machine Learning Techniques

Polynomial Chaos with Neural Networks

Rachel Cooper, Azam Moosavi, Vishwas Rao, and Adrian Sandu

Computational Science Laboratory
"Compute the Future!",
Department of Computer Science,
Virginia Polytechnic Institute and State University
Blacksburg, VA 24060
Argonne National Laboratory

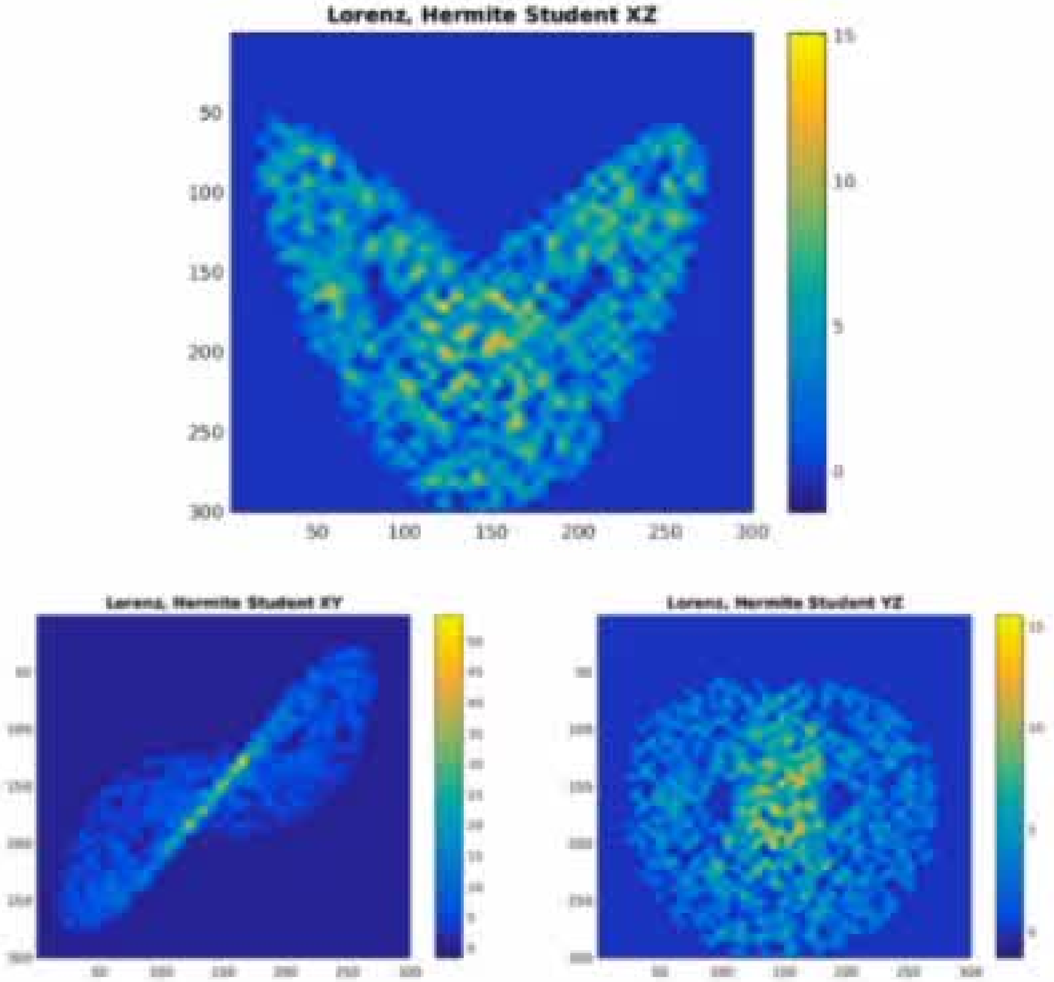
SIAM CSE, Feb. 25, 2019

Compute the Future!

VIRGINIA TECH.

The student learns the behaviors of the larger ReLU model in the smaller framework.

- Teacher student training
- Qualitatively closest result



Estimated polynomial for Lorenz System

$$
\begin{aligned}
f(x_1, x_2, x_3) = \max\{0, (&H_1(-0.65382x_1 + 0.0619x_2 - 0.0387x_3) + \\
&H_2(-0.6692x_1 + 0.8177x_2 - 0.5841x_3 - 0.0123) + \\
&H_3(0.4283x_1 + 0.5061x_2 + 0.1041x_3 - 0.0109) + 1.0125)\}
\end{aligned}
$$

$$
\begin{aligned}
H_1(x) &= x \\
H_2(x) &= x^2 - 1 \\
H_3(x) &= x^3 - 3x
\end{aligned}
$$

A learning-based approach for uncertainty analysis

Vishwas Rao
Mathematics and Computer Science Division
Argonne National Laboratory
Collaborators: A. Sandu, Azam Moosavi

SIAM CSE 19

02/25/2018

Data Assimilation

Data assimilation fuses information from prior, model, and observations to reduce uncertainty.

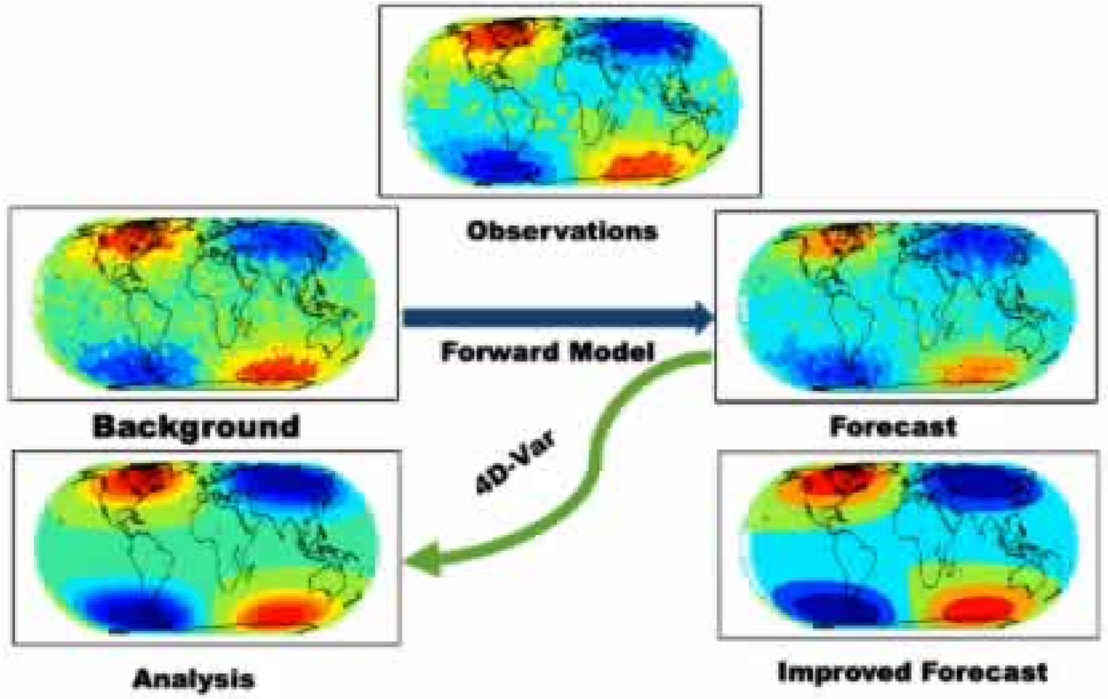
- We need to estimate the true state of the system($\mathbf{x}^{\text{true}}$)
- DA combines the following sources of information
  - **Prior** $\mathcal{P}^{\text{b}}(\mathbf{x})$: captures our current knowledge of $\mathbf{x}^{\text{true}}$
  - **The model**: captures our knowledge of the physical laws that govern the reality

$$\mathbf{x}_{k+1} = \mathcal{M}\left(\mathbf{x}_k\right) + \varepsilon^{\text{model}}, \quad k = 0, \ldots, N-1\,.$$

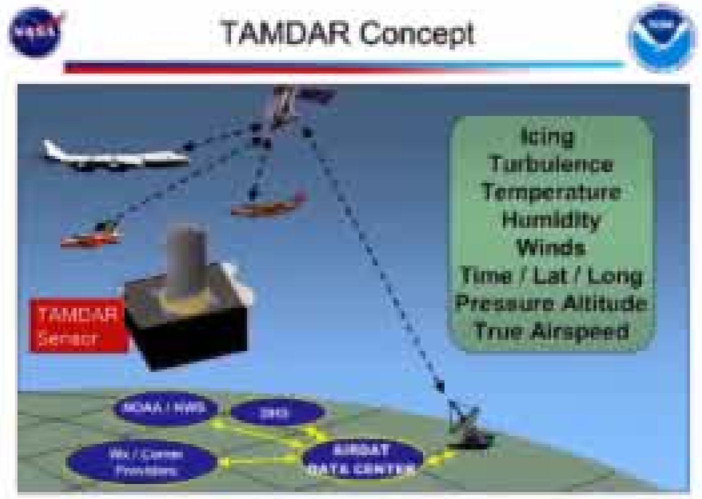
  - **Observations**: Noisy and sparse measurements of reality at different times

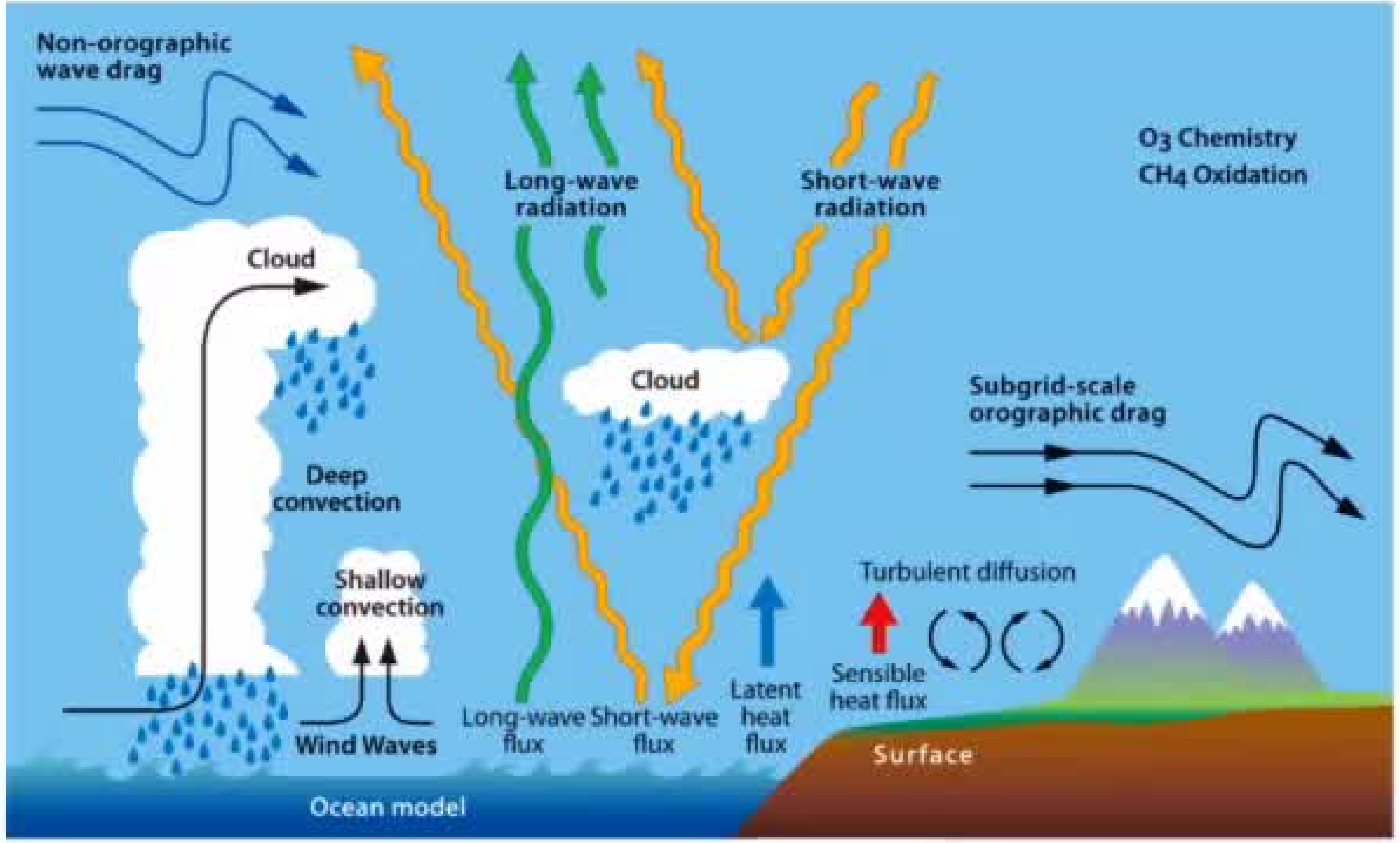
$$\mathbf{y}_i = \mathcal{H}_i\left(\mathbf{x}_i\right) + \varepsilon_i^{\text{obs}}, \quad i = 1, \ldots, N\,.$$

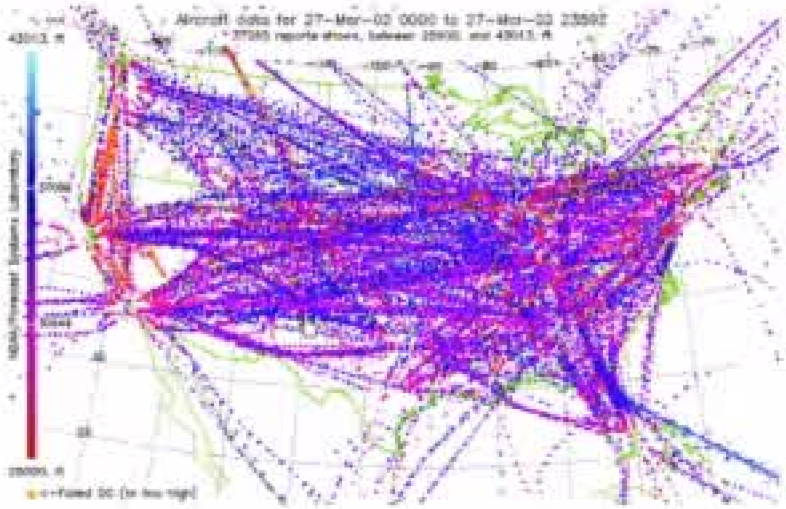
- DA computes the analysis (posterior, $\mathcal{P}^{\text{a}}(\mathbf{x})$) that represents our improved understanding of $\mathbf{x}^{\text{true}}$

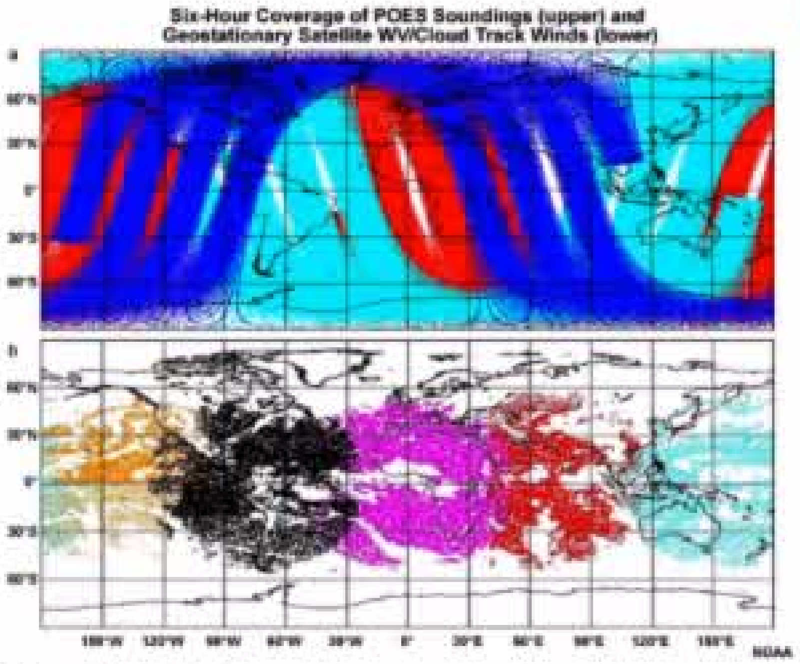


Challenges: Complex models and big data









- Complex Models (**10s of processes** and **10^8 Variables** in the state space).
- Lots and lots of data: **10^7** Observations from different sources in a period of 24 Hours.

Image sources: NASA, ECMWF, Weather.com

Model Errors

- We use measurements at sparse locations to obtain information about the global physical state

$$\mathbf{y}_t = h(v_t) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \mathbf{R}_t), \quad t = 1, \cdots, T.$$

- The model predicted values at observed locations is given by

$$\mathbf{o}_t = \mathcal{H}(\mathbf{x}_t), \quad t = 1, \cdots, T.$$

- Hence the model error in observation space is:

$$\boldsymbol{\Delta}_t = \mathbf{o}_t - \mathbf{y}_t \in \mathbb{R}^m, \quad t = 1, \cdots, T.$$

- Hence we can write the evolution equations for the physical system:

$$\begin{aligned} v_t &= \mathcal{M}\big(v_{t-1}, \Theta\big) + \boldsymbol{\delta}_t(v_t), \quad t = 1, \cdots, T, \\ \mathbf{y}_t &= h(v_t) + \epsilon_t. \end{aligned}$$

Model Errors

- Consider the following NWP model that describes the time evolution of the atmosphere

$$\mathbf{x}_t = \mathcal{M}(\mathbf{x}_{t-1}, \Theta), \quad t = 1, \cdots, T\,.$$

- The atmosphere as an abstract process, evolves in time as follows:

$$v_t = \mathcal{P}\left(v_{t-1}\right), \quad t = 1, \cdots, T.$$

- The model state seeks to approximate the physical state:

$$\mathbf{x}_t \approx \psi(v_t), \quad t = 1, \cdots, T.$$

- Assuming that the model state at t – 1 has the ”ideal” value, the model prediction will differ from reality:

$$\psi(v_t) = \mathcal{M}\left(\psi(v_{t-1}), \Theta\right) + \boldsymbol{\delta}_t\left(v_t\right), \quad t = 1, \cdots, T.$$

Model Errors

- We use measurements at sparse locations to obtain information about the global physical state

$$\mathbf{y}_t = h(v_t) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \mathbf{R}_t), \quad t = 1, \cdots, T.$$

- The model predicted values at observed locations is given by

$$\mathbf{o}_t = \mathcal{H}(\mathbf{x}_t), \quad t = 1, \cdots, T.$$

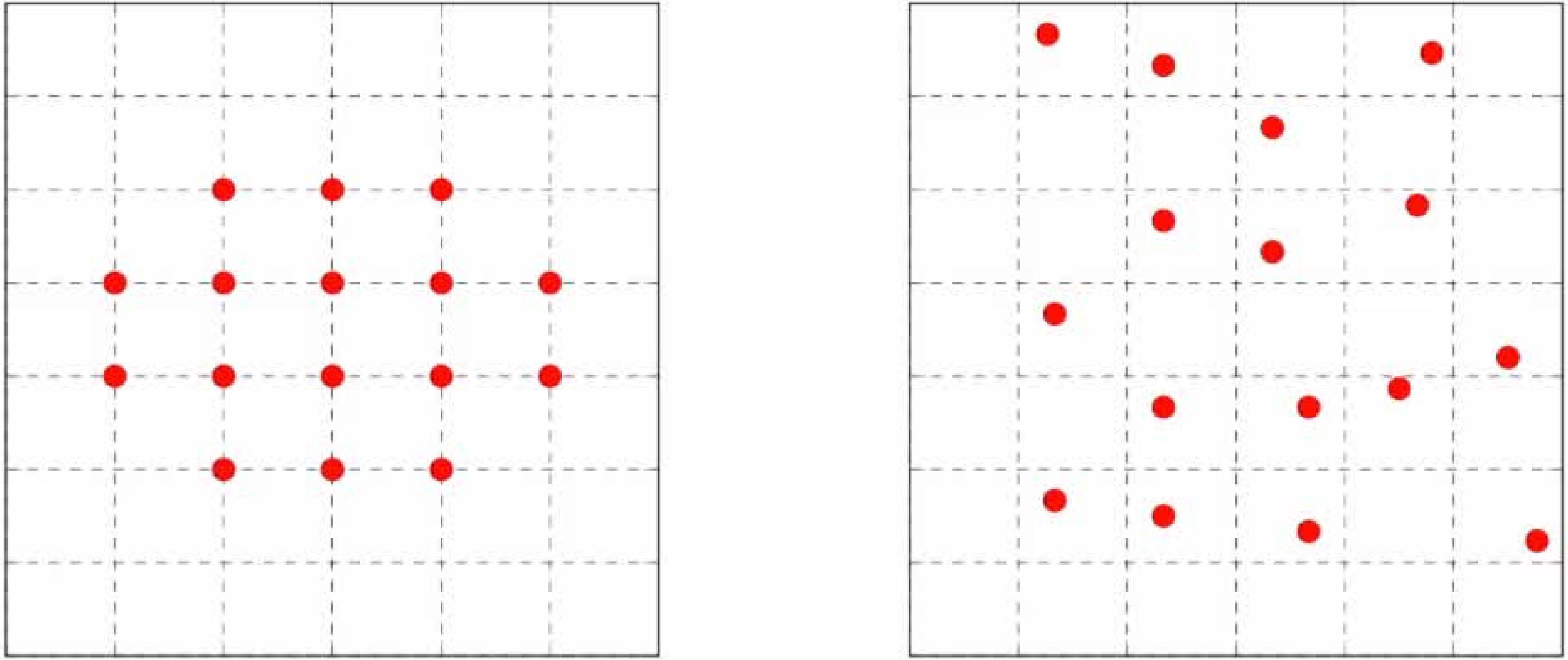
- Hence the model error in observation space is:

$$\boldsymbol{\Delta}_t = \mathbf{o}_t - \mathbf{y}_t \in \mathbb{R}^m, \quad t = 1, \cdots, T.$$

- Hence we can write the evolution equations for the physical system:

$$\begin{aligned} v_t &= \mathcal{M}\big(v_{t-1}, \Theta\big) + \boldsymbol{\delta}_t(v_t), \quad t = 1, \cdots, T, \\ \mathbf{y}_t &= h(v_t) + \epsilon_t. \end{aligned}$$

Observation operator ($\mathcal{H}$)



- Solution process involves some form of **comparison** between what the model predicts and the observations
- On the left: **Restrict** the solution to compare
- On the right: **Restrict + Interpolation** to compare

If we had access to the model errors

- Good estimates of the discrepancy when available, could improve model predictions by applying a correction:

$$\mathbf{v}_t \approx \mathbf{x}_t + \boldsymbol{\delta}_t.$$

- We are interested in the following problems

  - Estimating the QoI of the model error in advance

$$\phi^{\text{error}}\left(\Theta, \mathbf{o}_\tau, \boldsymbol{\Delta}_\tau, \mathbf{o}_t\right) \approx \boldsymbol{\Delta}_t \quad \tau < t.$$

  - Identifying the physical packages that contribute most to the forecast uncertainty

$$\phi^{\text{physics}}\left(\boldsymbol{\Delta}_t\right) \approx \Theta\,.$$

Numerical experiments

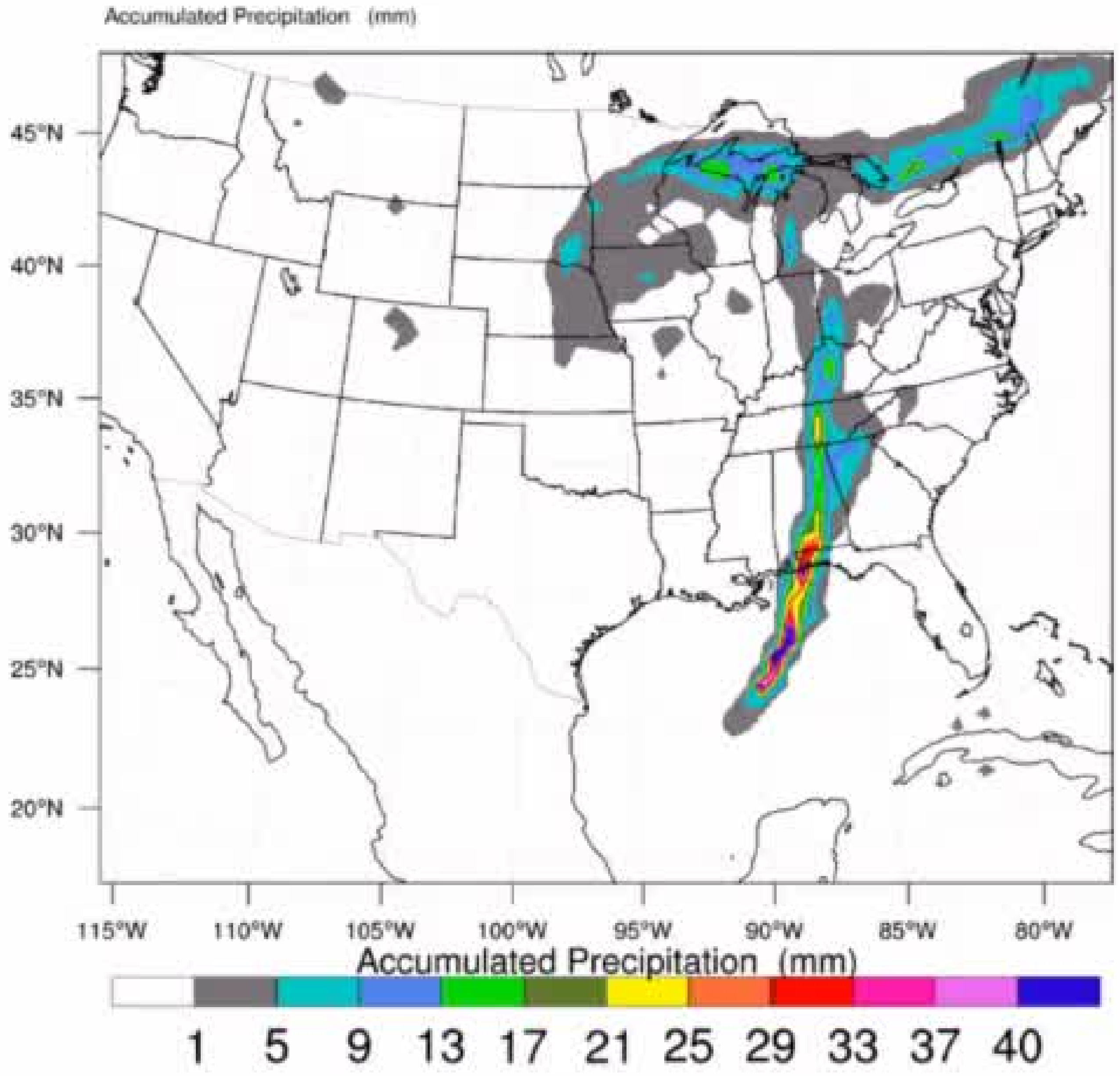
- All our experiments are conducted with WRF model and we estimate the uncertainties and model error for the prediction of precipitation.
- Training data: 7 AM to 12 AM (May 1st 2017).
- Testing period: 1 PM to 6PM (May 1st 2017).
- We use the "analysis provided" by NCEP as a proxy for truth.
- Input features which are used to train the machine

$$\phi^{\text{error}}\left(\Theta, \mathbf{o}_\tau, \boldsymbol{\Delta}_\tau, \mathbf{o}_{t=12\text{PM}}\right) \approx \boldsymbol{\Delta}_{t=12\text{PM}}, \quad 7\text{AM} \le \tau < 12\text{PM}.$$

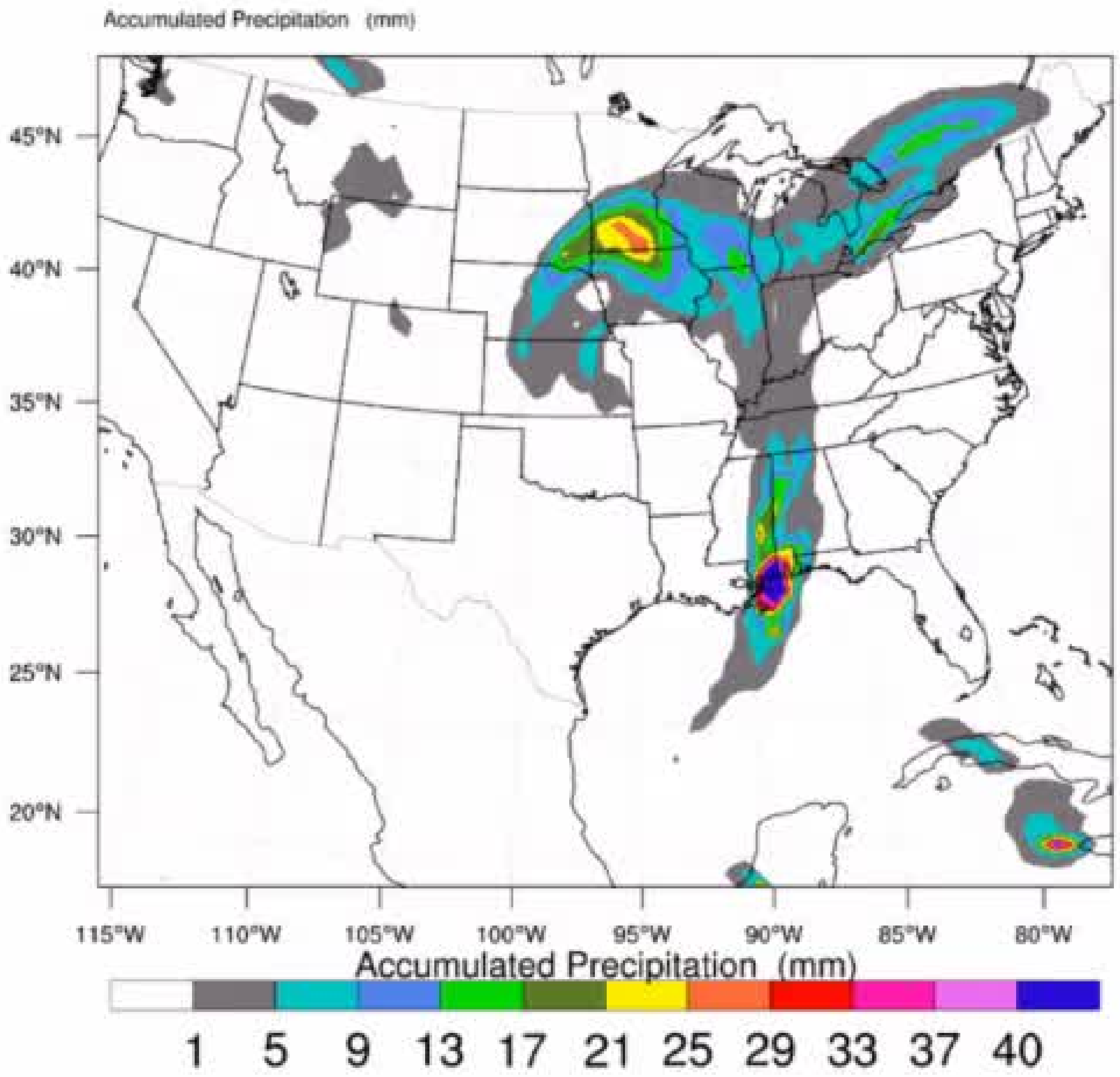
- Predict the forecast error at 6PM

$$\widehat{\boldsymbol{\Delta}}_{t=6\text{PM}} \approx \widehat{\phi}^{\text{error}}\left(\Theta, \mathbf{o}_\tau, \boldsymbol{\Delta}_\tau, \mathbf{o}_{t=6\text{PM}}\right), \quad 1\text{PM} \le \tau < 6\text{PM}.$$

Numerical experiments

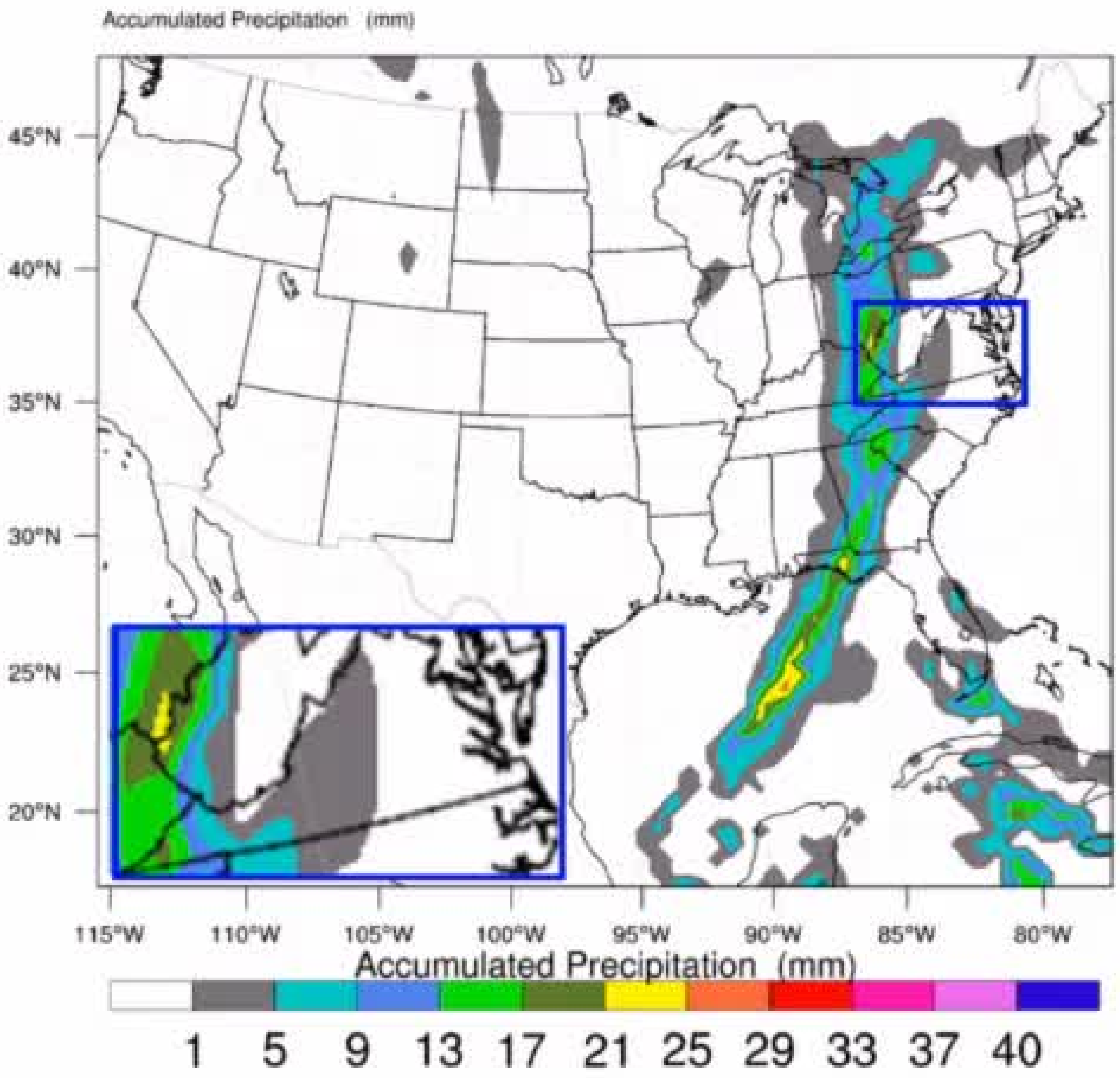


- Analysis at 6 AM
- Used as a proxy for truth for the training period

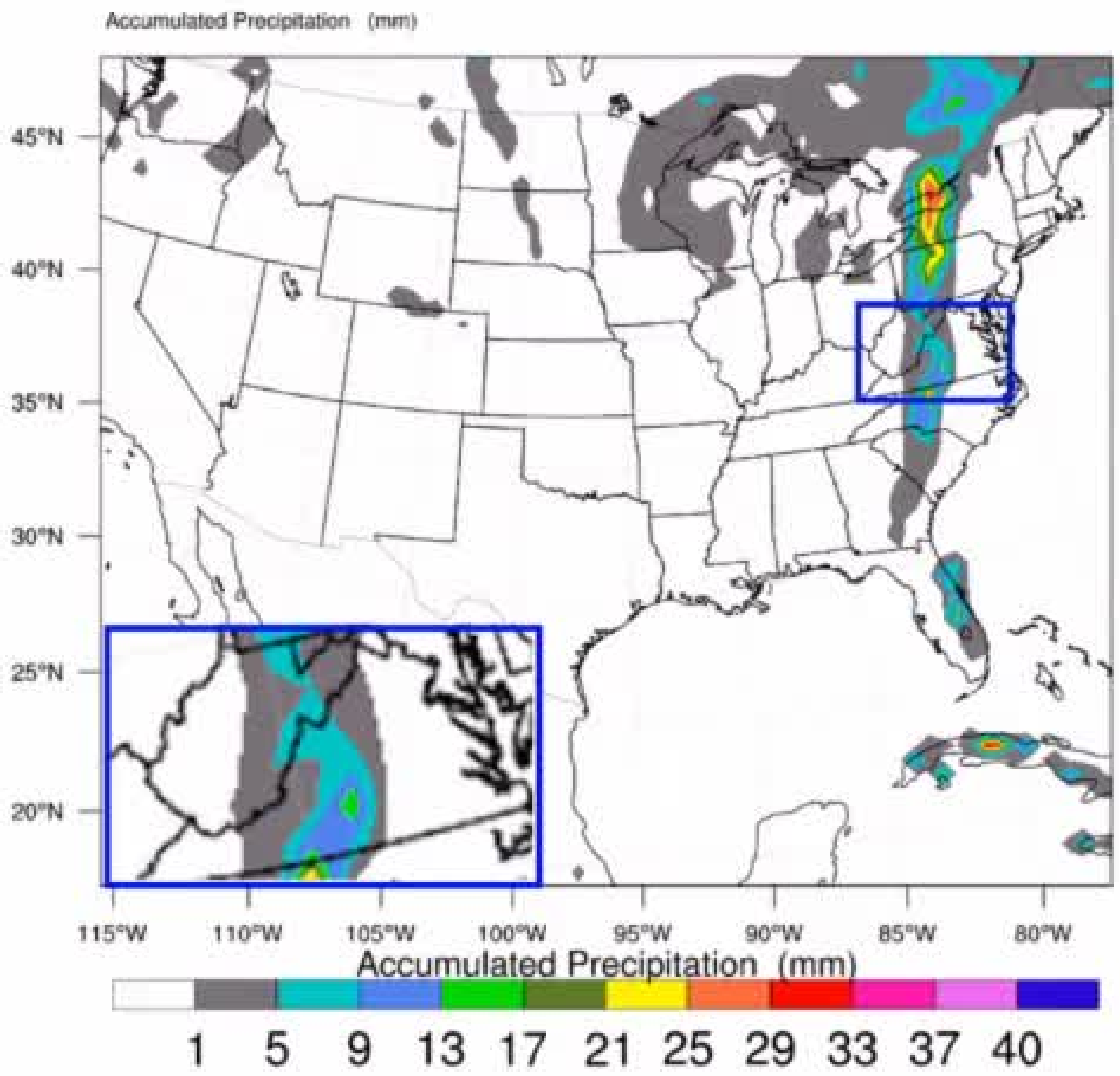


- Analysis at 12 PM
- Used as a proxy for truth for the testing period

Difference between WRF forecasts and Analysis

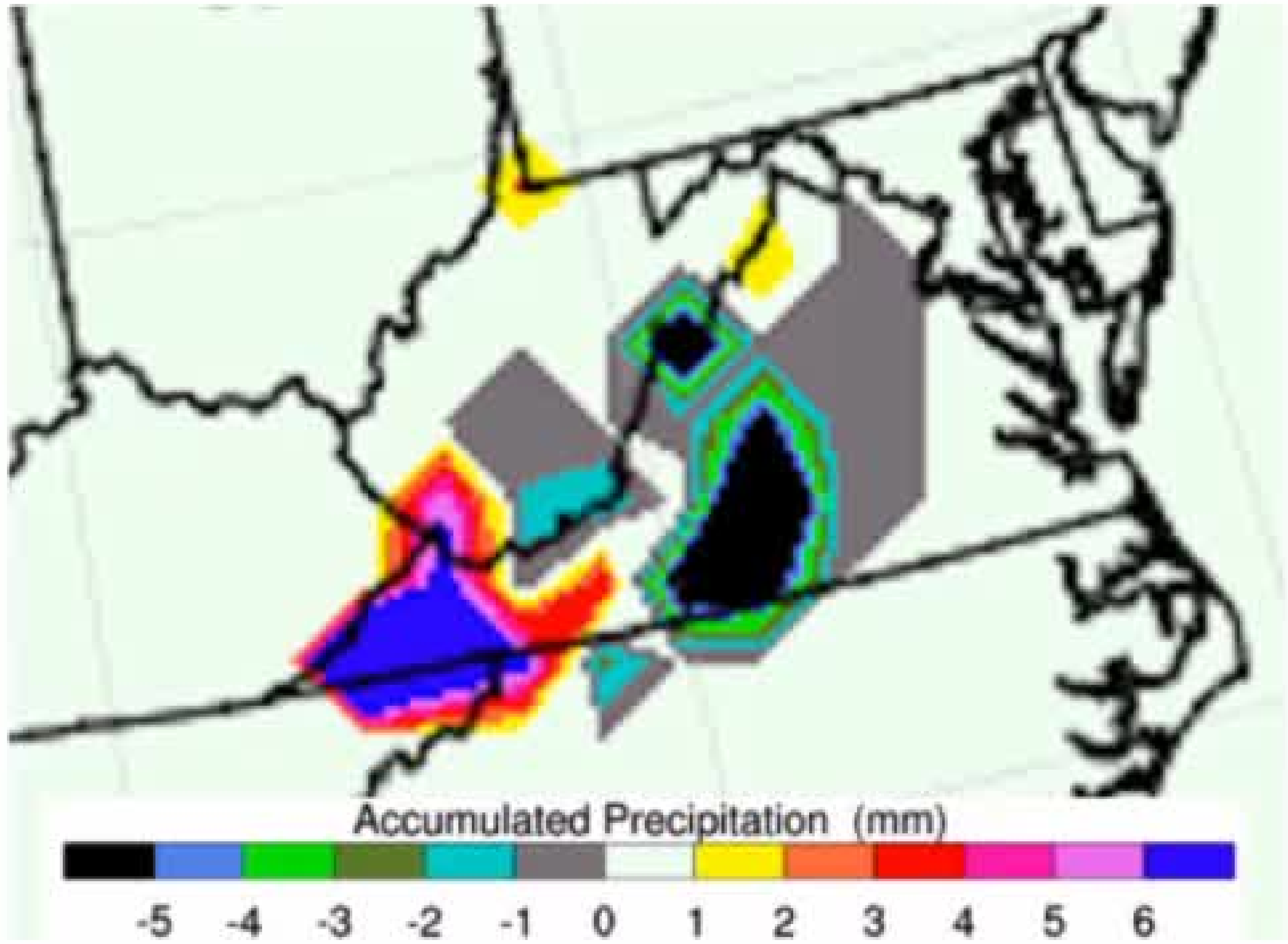


- WRF forecast at 6 PM



- NCEP Analysis at 6PM

Forecasts with and without corrections



- WRF Forecast
- WRF Forecast + Corrections

	RMSE
ANN	4.3 e-3
RF	5.8e-3

Packages that contribute to uncertainty

- The model configuration parameters represent various combinations of different physics such as microphysics schemes, cumulus parametrizations, short-wave, and long-wave radiation schemes
- The interaction of different physics schemes affect the accuracy of precipitation forecast.
- We construct a physics mapping using the norm and other statistical characterizations of the model data discrepancy as input features

$$\phi^{\text{physics}}\left(\bar{\mathbf{\Delta}}_{t=12\text{PM}}, \|\mathbf{\Delta}_{t=12\text{PM}}\|_2\right) \approx \Theta.$$

- WRF model is simulated for each of the possible physical combinations for the current forecast window and obtain the model errors for the current forecast window.
- From all the collected data, 80% is used for training the machine and on the remaining 20% we evaluate the model configuration, $\widehat{\Theta}_1$ for the given model errors.

Packages that contribute to uncertainty

- We repeat the test phase for each of the 50 samples with the scaled values of observable discrepancies ($\mathbf{\Delta}_{t=12\text{PM}}^{\text{test}}/2$) as inputs and obtain predicted physical combinations $\widehat{\Theta}_2$.
- The large variability in the predicted physical settings indicate that the WRF forecast error is sensitive to the corresponding physical packages.