Neural Network Applications for NWP Models

V. Krasnopolsky

Acknowledgement: *A. Belochitski, L. Breaker, Y. Fan, M. Fox-Rabinovitz, W. Gemmill, Y-T. Hou, H-Ch. Kim, Y. Lin, S. Lord, C. Lozano, A. Mehra, S. Moorthi, S. Nadiga, D-B. Rao, P. Rusch, H. Tolman, F. Yang*

Some NN applications in Numerical Modeling

- **Model Initialization:**
  - Satellite Retrievals
  - Direct Assimilation
  - Assimilation of Surface Observations
- **Model Physics:**
  - Fast and Accurate NN Emulations of Model Physics
  - New NN parameterizations
  - NN Stochastic physics
- **Post-processing:**
  - NN nonlinear ensembles
  - NN nonlinear bias corrections
  - Upscaling and downscaling

NN Applications Developed (black) and Under Development (red)

- **NN for Model Initialization:**
  - Satellite Retrievals
    - SSMI retrieval algorithm (operational since 1998)
    - QuickScat retrieval algorithm
  - Direct Assimilation
    - Forward model for direct assimilation of SSMI BT
    - QuickScat forward model
  - Assimilation of Surface Observations
    - Observation operator for assimilation of SSH anomaly
    - Empirical biological model for ocean color
    - NN algorithm to fill gaps in ocean color fields and for creating long and consistent ocean color data sets

NN Applications Developed (black) and Under Development (red) – cont.

- **NN for Model Physics**:
  - Fast and accurate emulations of parameterizations of physics
    - Fast nonlinear wave-wave interaction for WaveWatch
    - Fast NN long and short wave radiation for NCAR CAM, CFS, and GFS models
    - NN emulation for CRM in MMF
    - Fast NN microphysics for NMMB and WARF
  - New parameterization
    - Convection parameterization for NCAR CAM learned by NN from CRM

NN Applications Developed (black) and Under Development (red) – cont.

- **NN for Post-processing:**
  - Nonlinear ensembles
    - Nonlinear multi-model NN ensemble for calculating precipitation rates over ConUS
    - Nonlinear NN averaging of wave models ensemble
  - Nonlinear bias corrections
    - Nonlinear bias corrections for GFS
    - Nonlinear bias corrections for wave model

Mapping

- Mapping: A ***rule of correspondence established between vectors*** in vector spaces ^{n} and ^{m} that associates each vector ***X*** of a vector space ^{n} with a vector ***Y*** in another vector space ^{m}.

$$Y = F(X)$$

$$X = \{x_1, x_2, \ldots, x_n\}, \quad ^{n}$$

$$Y = \{y_1, y_2, \ldots, y_m\}, \quad ^{m}$$

$$y_1 = f_1(x_1, x_2, \ldots, x_n)$$

$$y_2 = f_2(x_1, x_2, \ldots, x_n)$$

$$\vdots$$

$$y_m = f_m(x_1, x_2, \ldots, x_n)$$

Mapping $Y = F(X)$: examples

- **Time series prediction:**

  $X = \{x_t, x_{t-1}, x_{t-2}, ..., x_{t-n}\}$, - Lag vector

  $Y = \{x_{t+1}, x_{t+2}, ..., x_{t+m}\}$ - Prediction vector

  (Weigend & Gershenfeld, "Time series prediction", 1994)

- **Nonlinear ensemble average:**

  X = {Ensemble members, Metadata}

  Y = {Nonlinear ensemble average}

  (Krasnopolsky and Lin, 2012)

- **Retrieving surface wind speed over the ocean from satellite data (SSM/I):**

  X = {SSM/I brightness temperatures}

  Y = {W, V, L, SST}

  (Krasnopolsky, et al., 1999; operational since 1998)

- **Calculation of long wave atmospheric radiation:**

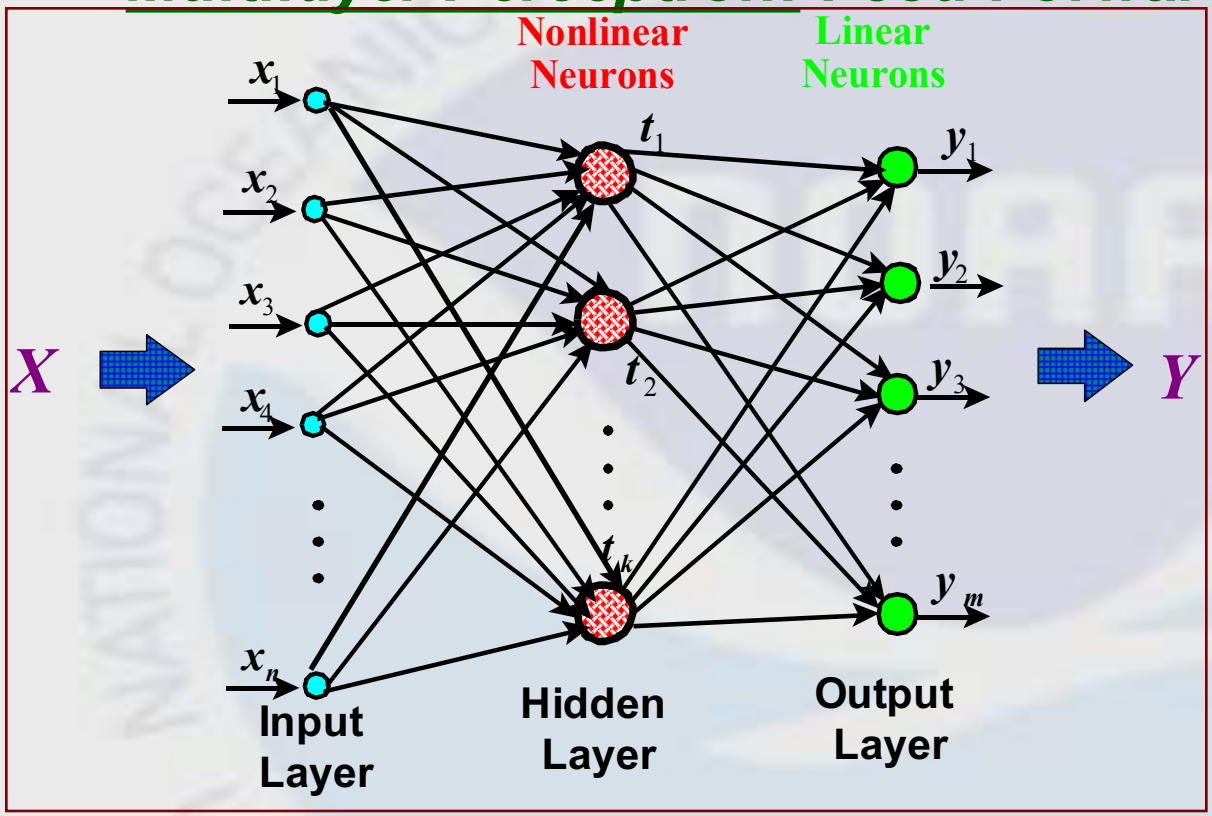
  X = {Temperature, moisture, O_3, CO_2, cloud parameters profiles, surface fluxes, Metadata}

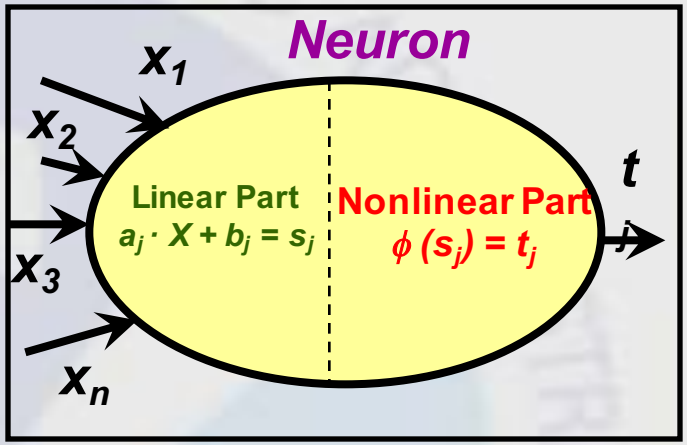
  Y = {Heating rates profile, radiation fluxes}

  (Krasnopolsky et al., 2005, 2010, 2012)

NN - Continuous Input to Output Mapping

Multilayer Perceptron: Feed Forward, Fully Connected





$$t_j = (b_{j0} + \sum_{i=1}^{n} b_{ji} \; x_i) =$$

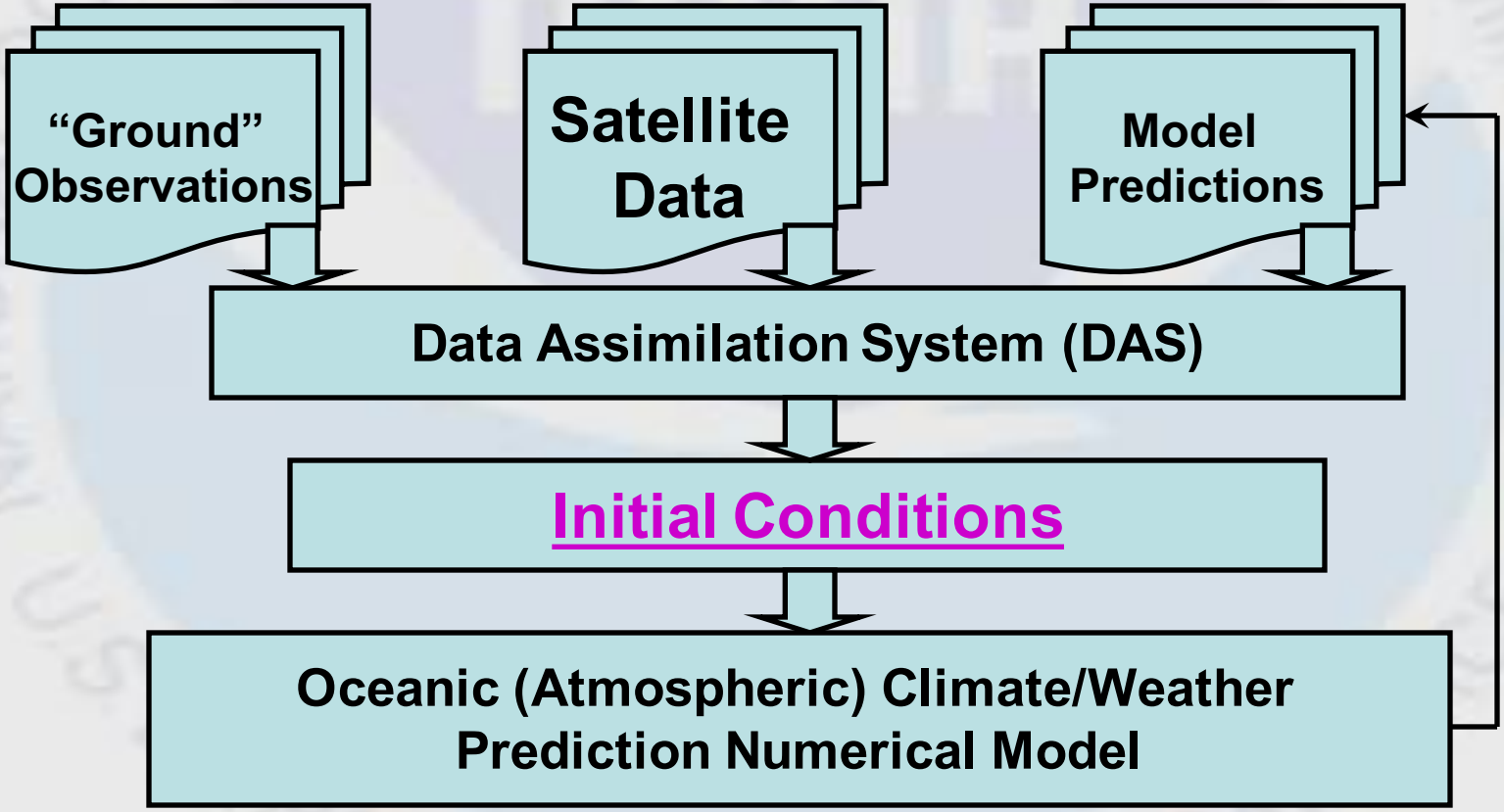
$$= \tanh(b_{j0} + \sum_{i=1}^{n} b_{ji} \; x_i)$$

$$Y = F_{NN}(X) \Rightarrow$$

$$y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \; t_j = a_{q0} + \sum_{j=1}^{k} a_{qj} \; (b_{j0} + \sum_{i=1}^{n} b_{ji} \; x_i) =$$

$$= a_{q0} + \sum_{j=1}^{k} a_{qj} \; \tanh(b_{j0} + \sum_{i=1}^{n} b_{ji} \; x_i); \quad q = 1,2,\ldots,m$$

I. NN in Model Initialization



Ingesting Satellite Data in DAS

- **Satellite Retrievals:**

$$G = f(S),$$

S *– vector of satellite measurements;*

G *– vector of geophysical parameters;*

f *– transfer function or retrieval algorithm*

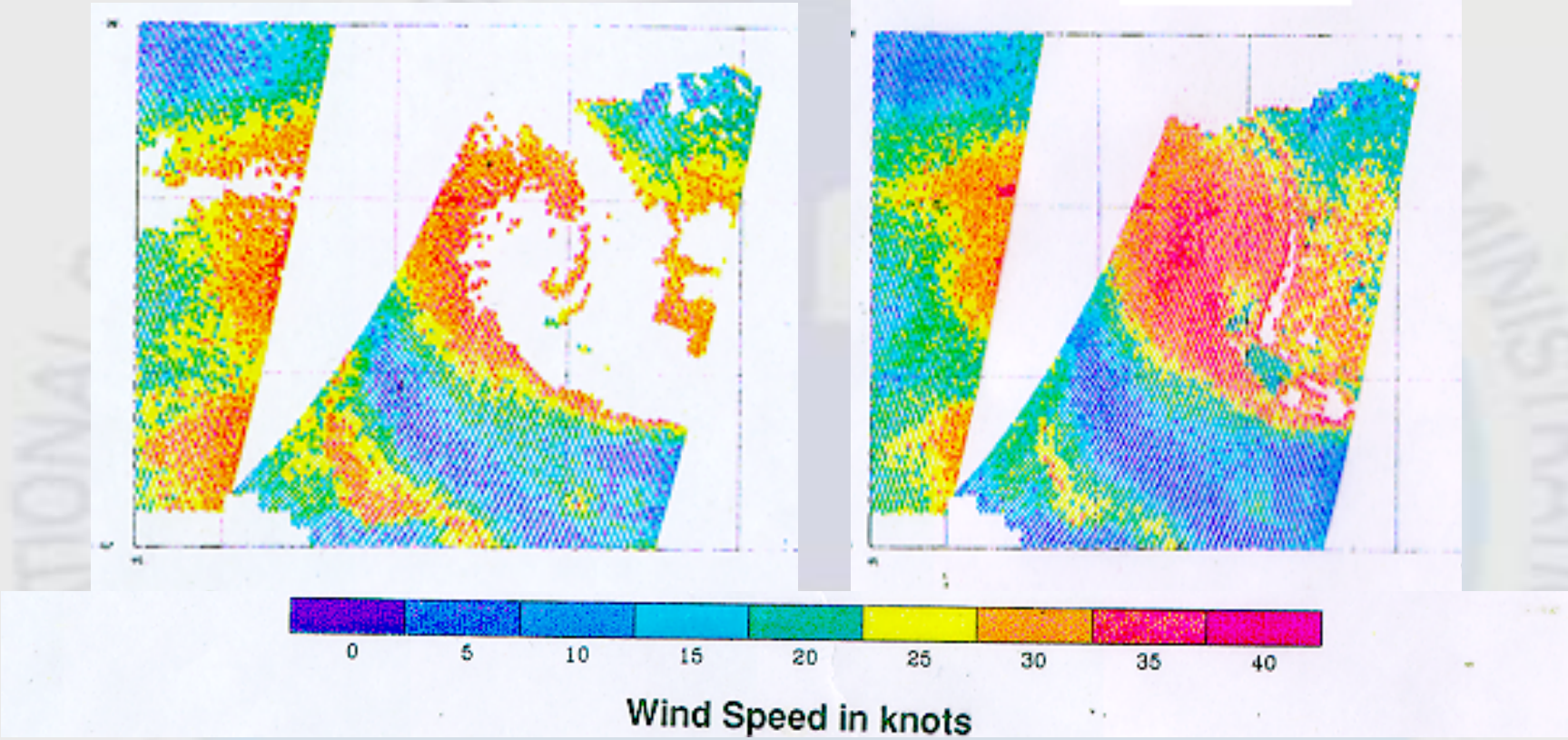
- **Direct Assimilation of Satellite Data:**

$$S = F(G),$$

F *– forward model*

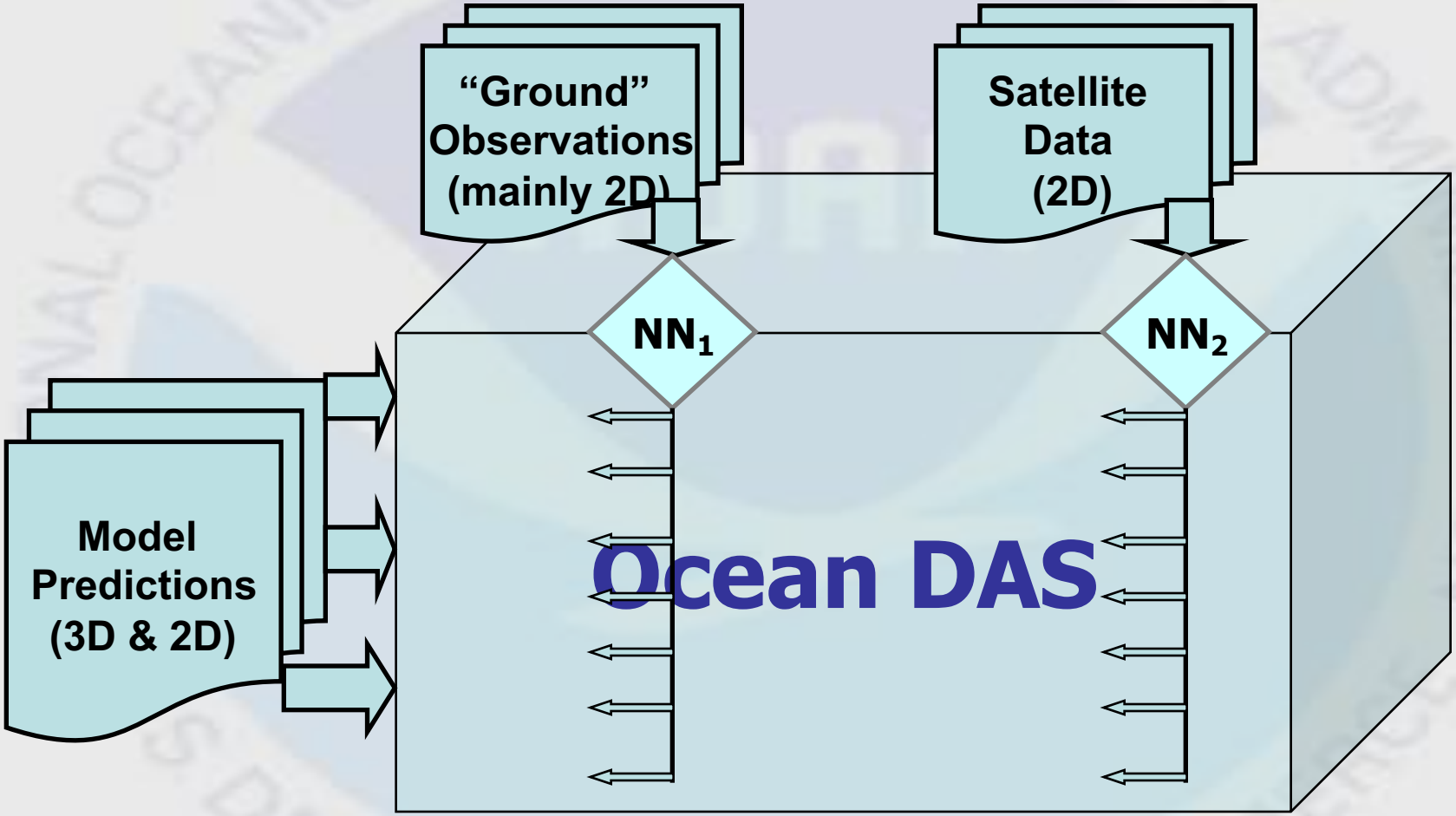
- Both ***F*** & ***f*** are mappings and NN are used
  - *Fast and accurate NN retrieval algorithms* $\mathbf{f_{NN}}$
  - *Fast NN forward models* F_{NN} *for direct assimilation*

Satellite Retrievals



Wind speed fields retrieved from the SSM/I measurements for a mid-latitude storm. Two passes (one ascending and one descending) are shown in each panel. Each panel shows the wind speeds retrieved by (left to right) GSW (linear regression) and NN algorithms. The GSW algorithm does not produce reliable retrievals in the areas with high level of moisture (white areas). NN algorithm produces reliable and accurate high winds under the high level of moisture. 1 knot ≈ 0.514 m/s

DAS: Propagation of Information Vertically Using NNs



NN1 and NN2 – observation operators

Observation operator for SSH

$$\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X}) \approx \boldsymbol{NN}(\boldsymbol{X}),$$

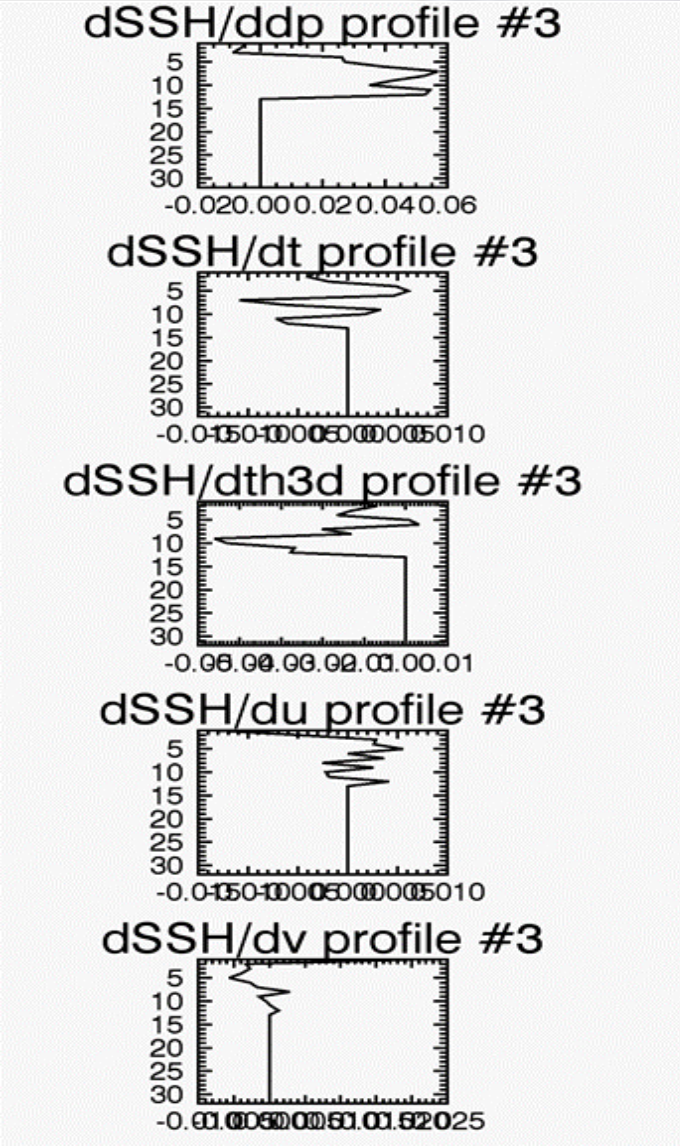
Y *– vector of SSH satellite measurements;*

X *– vector of ocean model variables;*

F *– observation operator –* ***mapping***

Input #	Names	Units	Input	Size
1	Day of the year		$\sin(\frac{2 \cdot day \cdot \pi}{366})$	1
2	---- " -----		$\cos(\frac{2 \cdot day \cdot \pi}{366})$	1
3	Lon		$Sin(lon)$	1
4	---- " ---		$Cos(lon)$	1
5	Lat		$Sin(lat)$	1
6	bottom pressure	?	pbot	1
7	vertically average x-veolcity	cm/s	ubavg	1
8	vertically average y-veolcity	cm/s	vbavg	1
9	temperature	°C	temp	32
10	layer thickness at p-points	m	dp	32
11	potential density	kg/m^3	th3d	32
12	internal x-velocity	cm/s	u	32
13	internal y-velocity	cm/s	v	32
Total				168
Output #			Output	Size
1	SSH anomaly	m	srfhgt	1
2	Montgomery potential	m^2/s^2	montg1	1
Total				2

Jcobian of Observation Operator



Profiles of derivatives of SSH. Derivatives over *dp* (upper row), *temp* (second row from the top), *th3d* (third row from the top), *u* (fourth row from the top) and *v* (bottom row).

II. NN for Model Physics

- Deterministic First Principles Models, 3-D Partial Differential Equations on the Sphere + the set of conservation laws (mass, energy, momentum, water vapor, ozone, etc.)

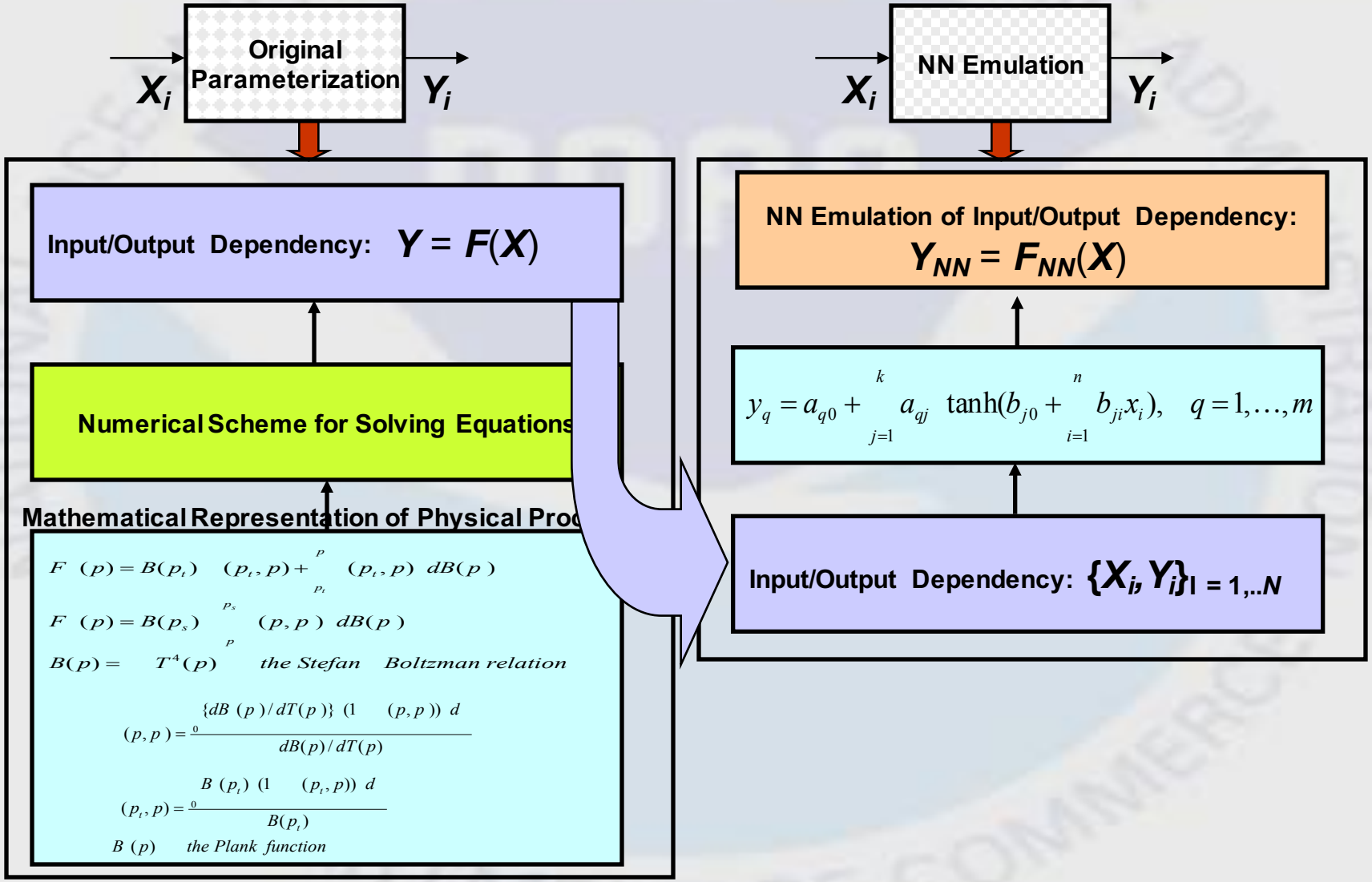
$$\frac{\quad}{t} + D(\quad, x) = P(\quad, x)$$

  - ψ - a 3-D prognostic/dependent variable, e.g., temperature
  - *x* - a 3-D independent variable: *x, y, z* & *t*
  - *D* - dynamics (spectral or gridpoint)
  - *P* - physics or parameterization of physical processes (1-D vertical r.h.s. forcing) – mostly time consuming part > **<u>50% of total time</u>**

Accurate and Fast NN Emulations of Model Physics

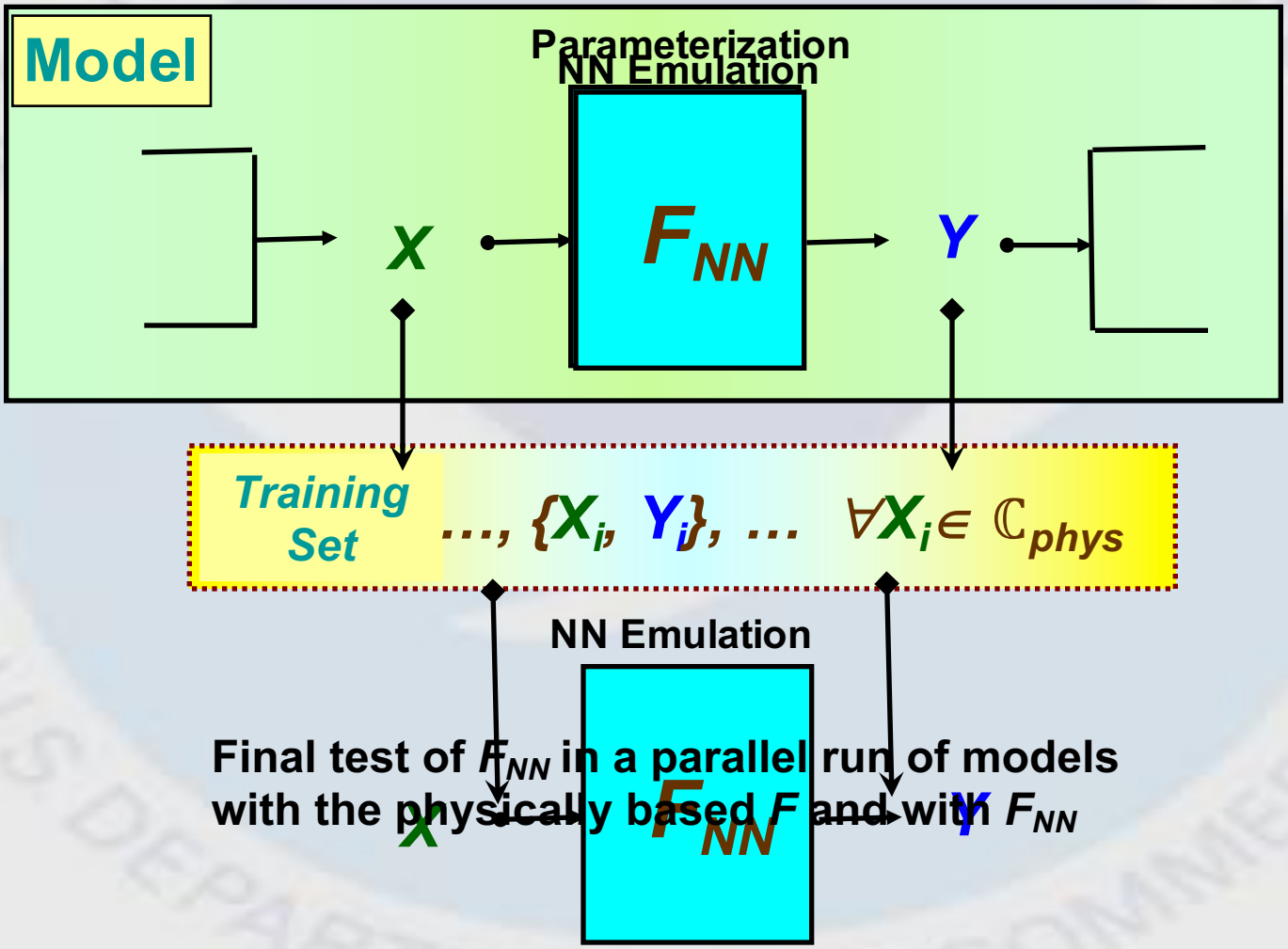
- Any **parameterization** of model physics is a mapping and can be emulated by NN
- The **entire model physics** is a mapping and can be emulated by NN
- NN emulation is usually **1 to 2 orders of magnitude faster** than the original parameterization: ~20 times for SWR and ~100 times for LWR

The Magic of NN Performance



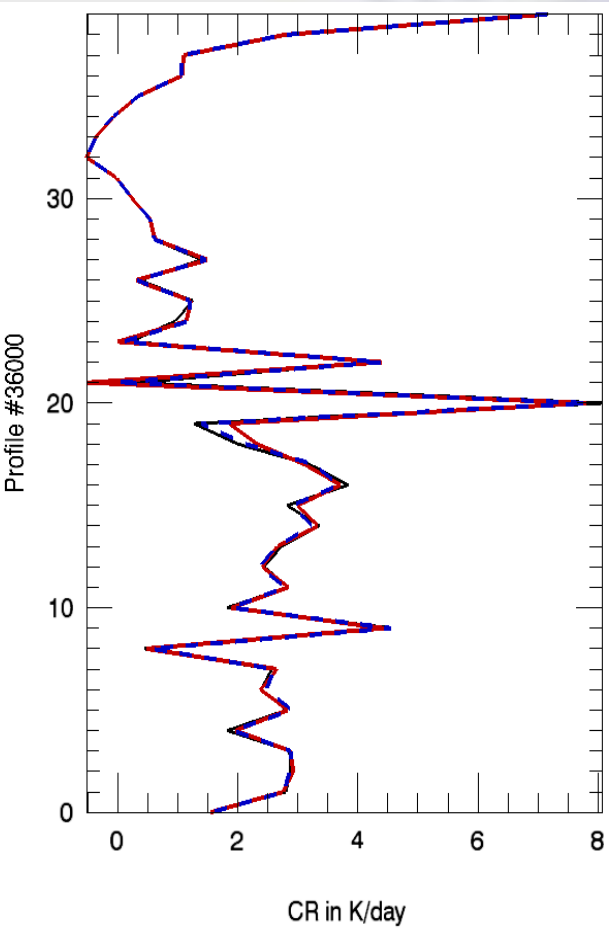
NN Emulations of Model Physics Parameterizations

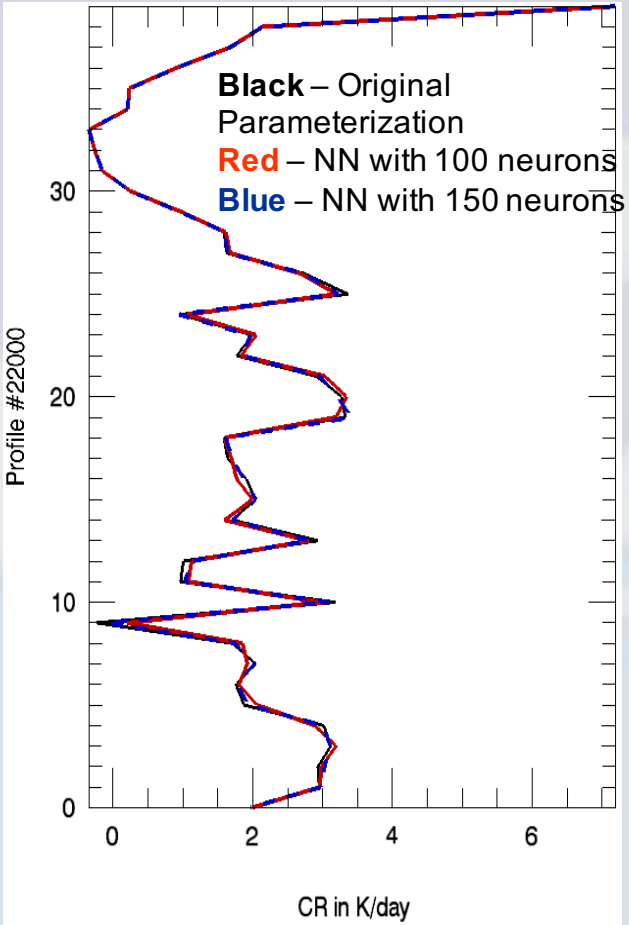
Learning from Data

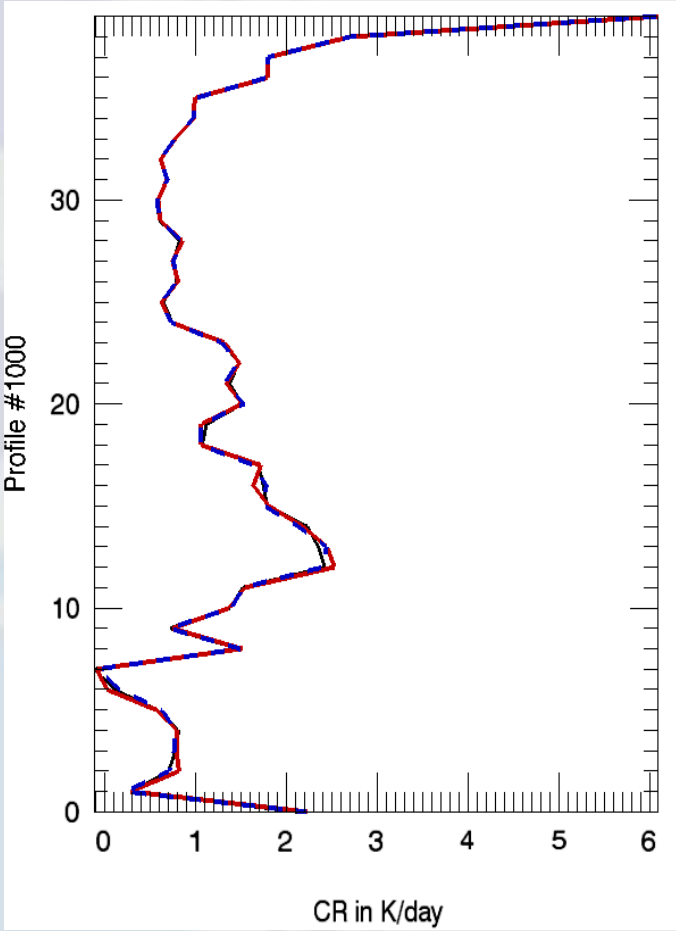


Final test of F_{NN} in a parallel run of models with the physically based F and with F_{NN}

LWR Individual Profiles



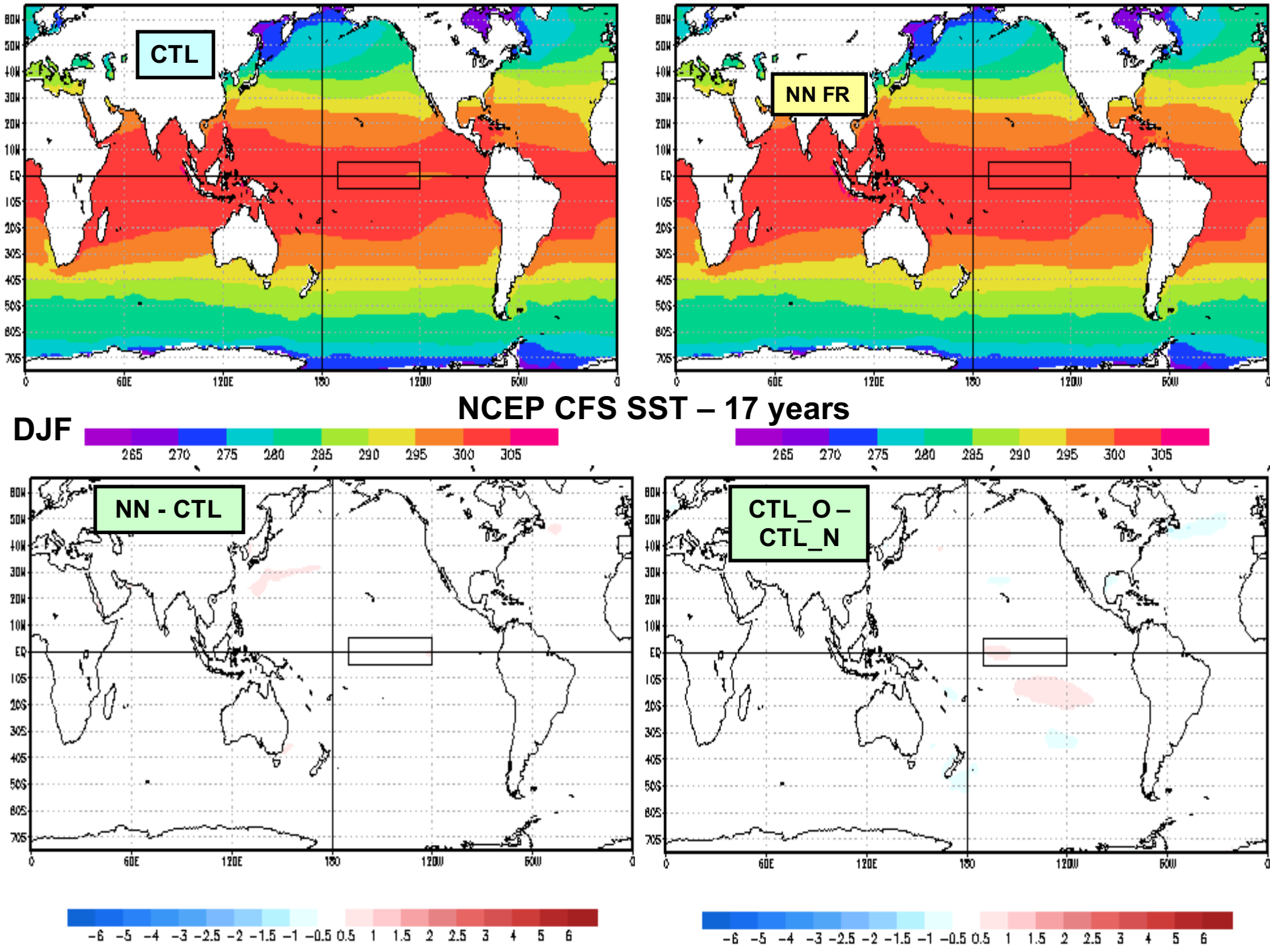


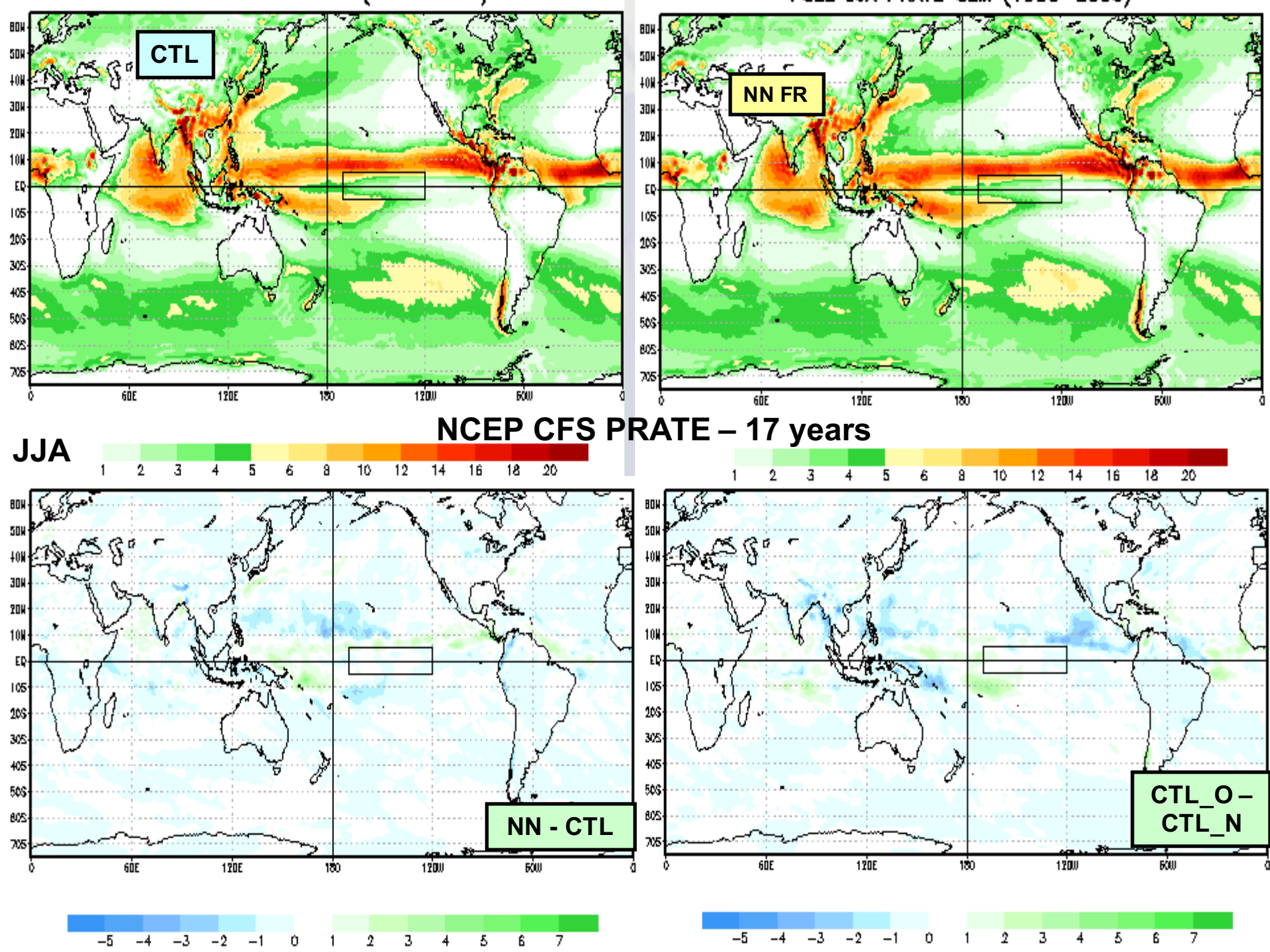


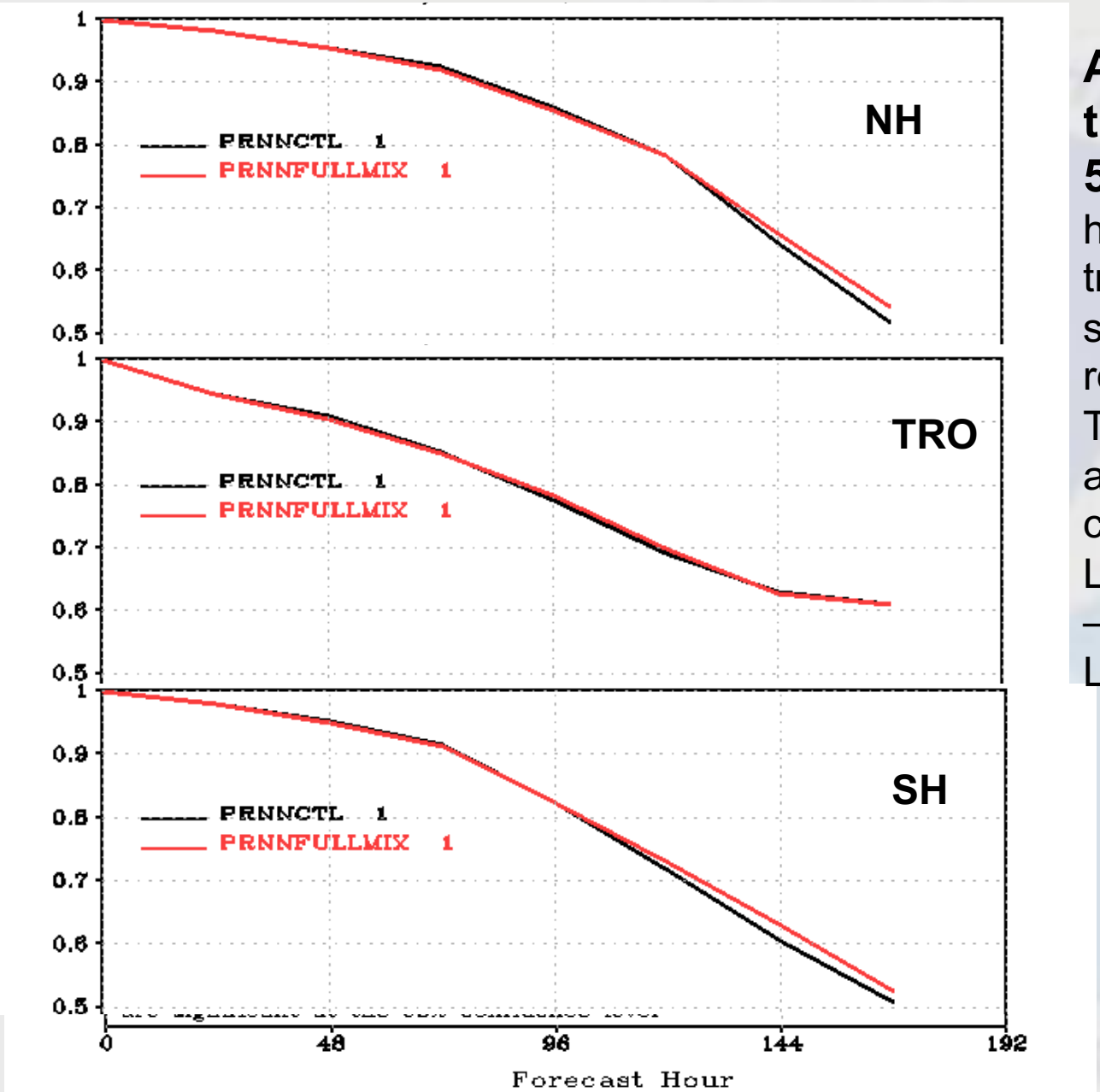
PRMSE = 0.18 & 0.10 K/day

PRMSE = 0.11 & 0.06 K/day

PRMSE = 0.05 & 0.04 K/day


CTL
NN FR
NCEP CFS SST – 17 years
DJF
265 270 275 280 285 290 295 300 305
NN - CTL
CTL_O – CTL_N
-6 -5 -4 -3 -2.5 -2 -1.5 -1 -0.5 0.5 1 1.5 2 2.5 3 4 5 6


CTL
NN FR
NCEP CFS PRATE – 17 years
JJA
1 2 3 4 5 6 8 10 12 14 16 18 20
NN - CTL
CTL_O – CTL_N
-5 -4 -3 -2 -1 0 1 2 3 4 5 6 7



Anomaly correlation for temperature fields at 500 mb for the northern hemisphere (upper row), tropics (medium row), and southern hemisphere (lower row).

Two GFS (T574L64) runs are shown: black line – control run with the original LWR and SWR and red line – run with NN SWR and LWR .

III. Post-processing Model Output

- **<u>Nonlinear multi-model ensembles for precipitation forecast</u>.**
- **Precipitation forecasts available from 8 operational models:**
  - **NCEP's mesoscale & global models (NAM & GFS)**
  - **the Canadian Meteorological Center regional & global models (CMC & CMCGLB)**
  - **global models from the Deutscher Wetterdienst (DWD)**
  - **the European Centre for Medium-Range Weather Forecasts (ECMWF) global model**
  - **the Japan Meteorological Agency (JMA) global model**
  - **the UK Met Office (UKMO) global model**
- **Also NCEP Climate Prediction Center (CPC) precipitation analysis is available over ConUS.**

Calculating Ensemble Mean

- **Conservative ensemble**:

$$EM = \frac{1}{N}\sum_{i=1}^{N} p_i, \quad p_i \ \ is\ an\ ensemble\ member$$

- **Weighted ensemble**:

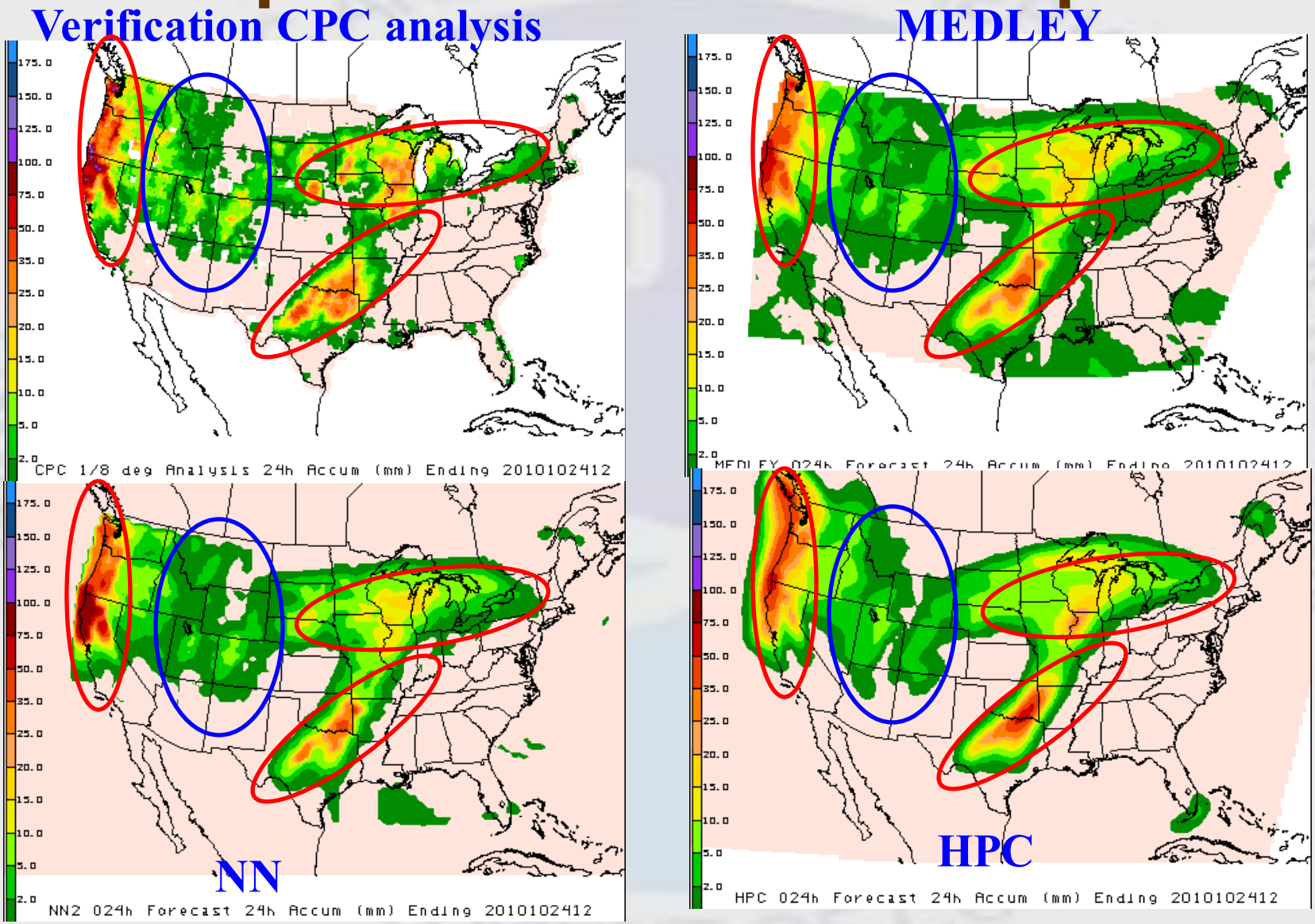
$$WEM = \frac{\sum_{i=1}^{N} W_i p_i}{\sum_{i=1}^{N} W_i}$$

W_i - from a priori considerations or from past data (linear regression)

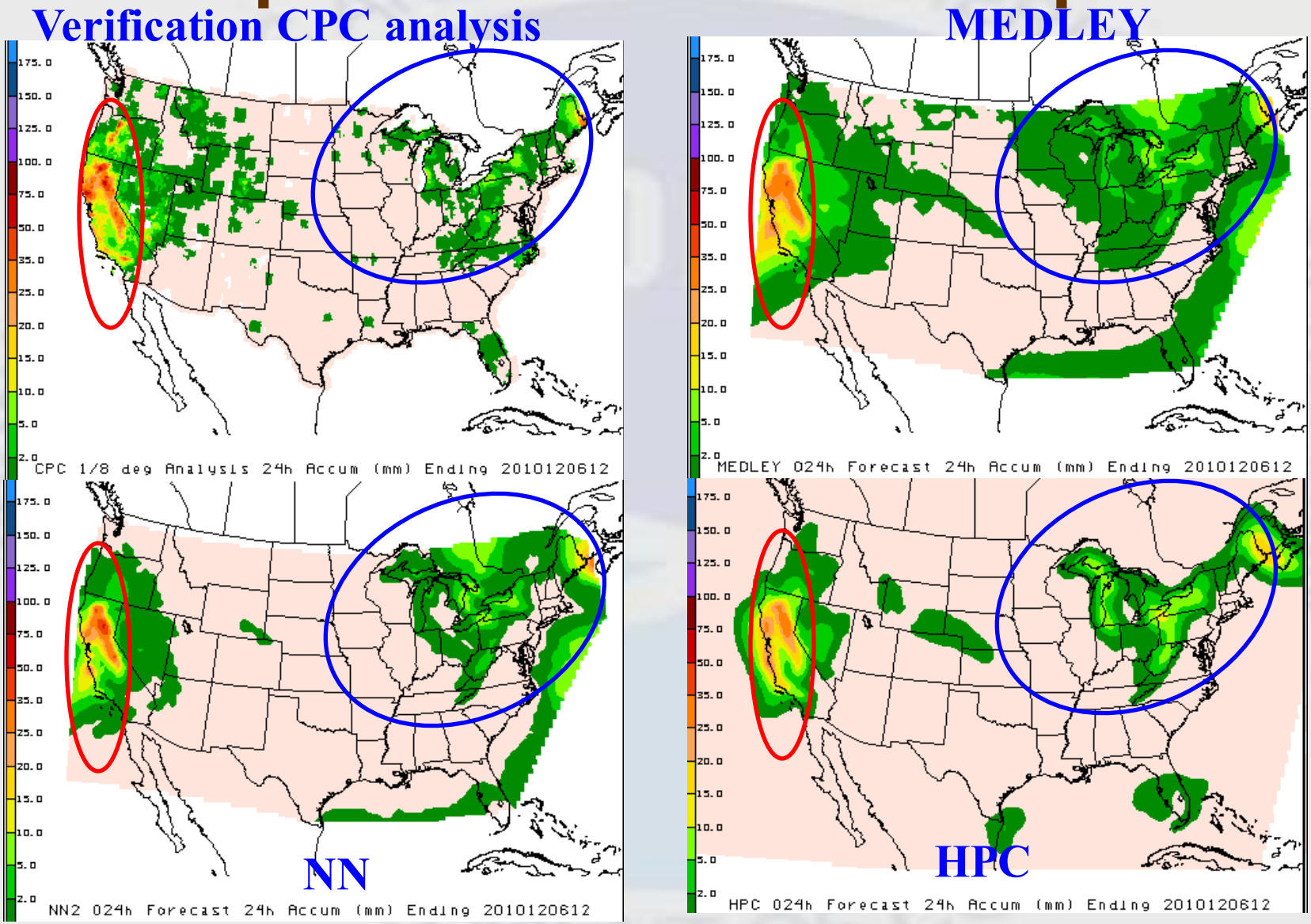
- **If past data are available the assumption of linear dependency can be relaxed:**

$$\boldsymbol{NEM = f(P) \approx NN(P)}$$

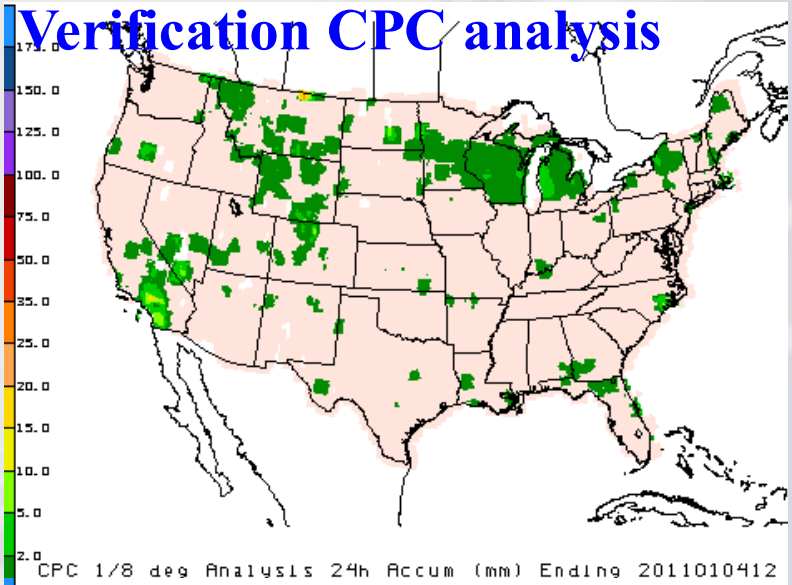
Sample NN forecast: example 1

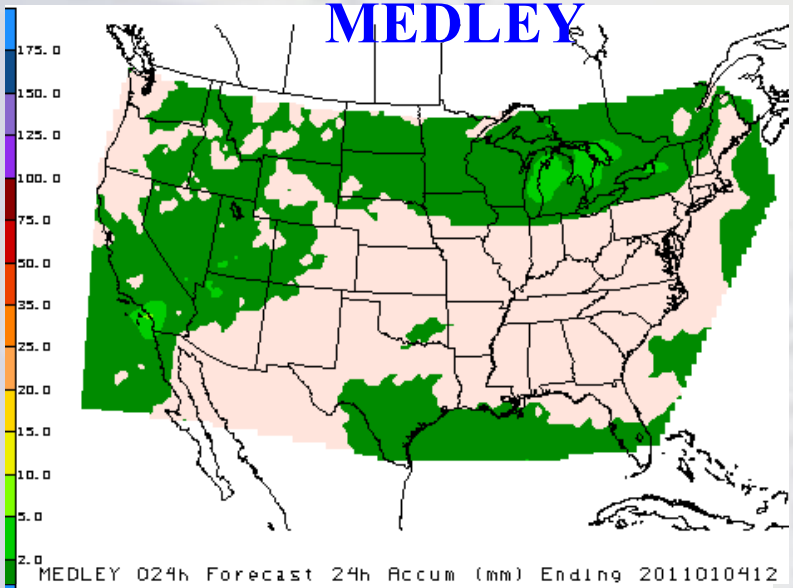


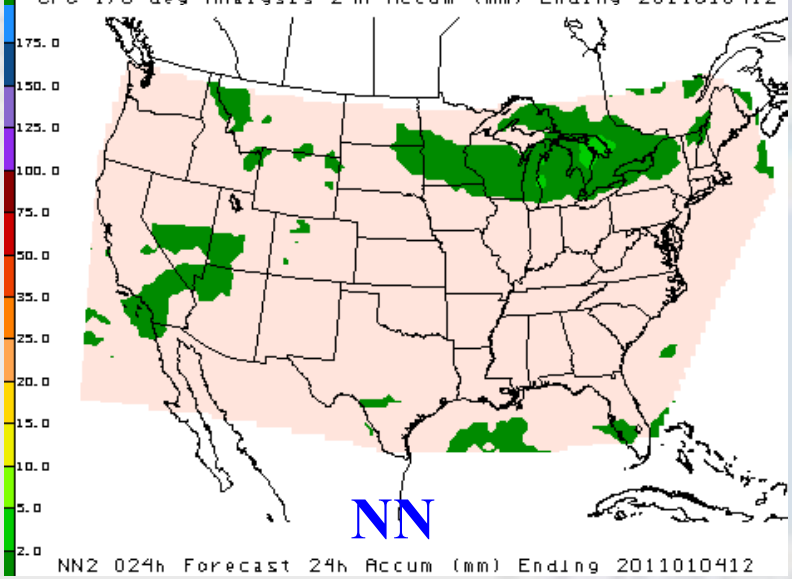
Sample NN forecast: example 2

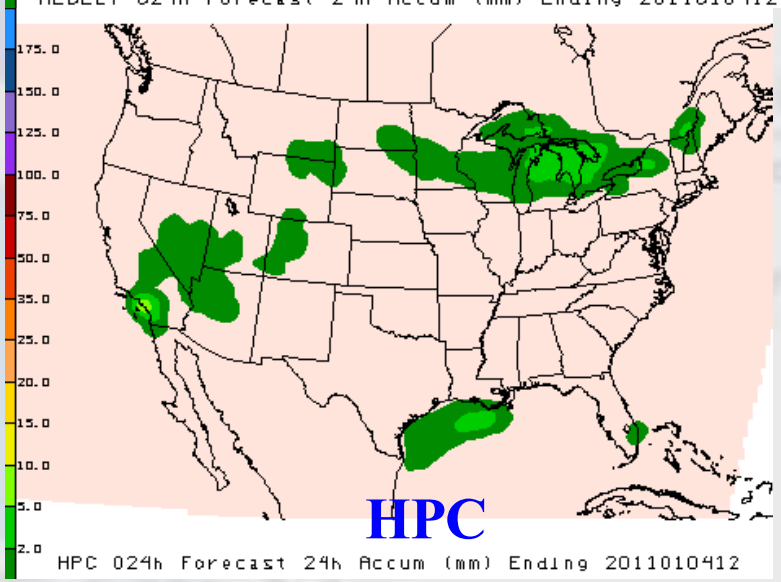


Sample NN forecast: example 3









Nonlinear Bias correction

- **Current approaches:**
  - Relate observed weather elements (**PREDICTANDS**) to appropriate model variables (**PREDICTORS**) via a statistical approach.
- **Predictors are obtained from:**
  1. Numerical Weather Prediction (NWP) Model Forecasts
  2. Prior Surface Weather Observations
  3. Geoclimatic Information
- **Predictands are obtained from:**
  - Historical record of observations at forecast points

Methodology

- **Statistical Method:**
  - **<u>MULTIPLE LINEAR REGRESSION</u>**

$$y_i = a_0 + a_1 x_1 + a_2 x_2 + \ \ldots + a_n x_n$$

y_i - predictand; i = 1,2,…,m

$X = \{x_1, x_2, \ldots, x_n\}$ – vector of predictors

FOR EACH SITE and FOR EACH PREDICTAND and FOR EACH WEATHER REGIME!

- **Motivations:**
  - **Actual dependence $y_i = f(X)$ may be nonlinear!**
  - **NNs are a flexible nonlinear tool that has a potential to represent these nonlinear dependencies more efficiently and compact**

Data used in this study:

- observations for **two** p**redictands** variables T_{2m} and T_d
- collected at **3,000 stations over the ConUS**
- during the period of **669 days** starting June 1, 2013 and ending March 31, 2015.
- GFS 24h forecast was saved for **five predictors**, GFS variables T_{2m}, T_d, rh, T_h, and wd during the same period of time. These variables plus some metadata (see below) were used as predictors.
- All days and locations with missing data are removed from the data sets.
- First 365 days were used for NN training, and last 304 days for independent validation
- Only 24 h projection time was considered

Some results

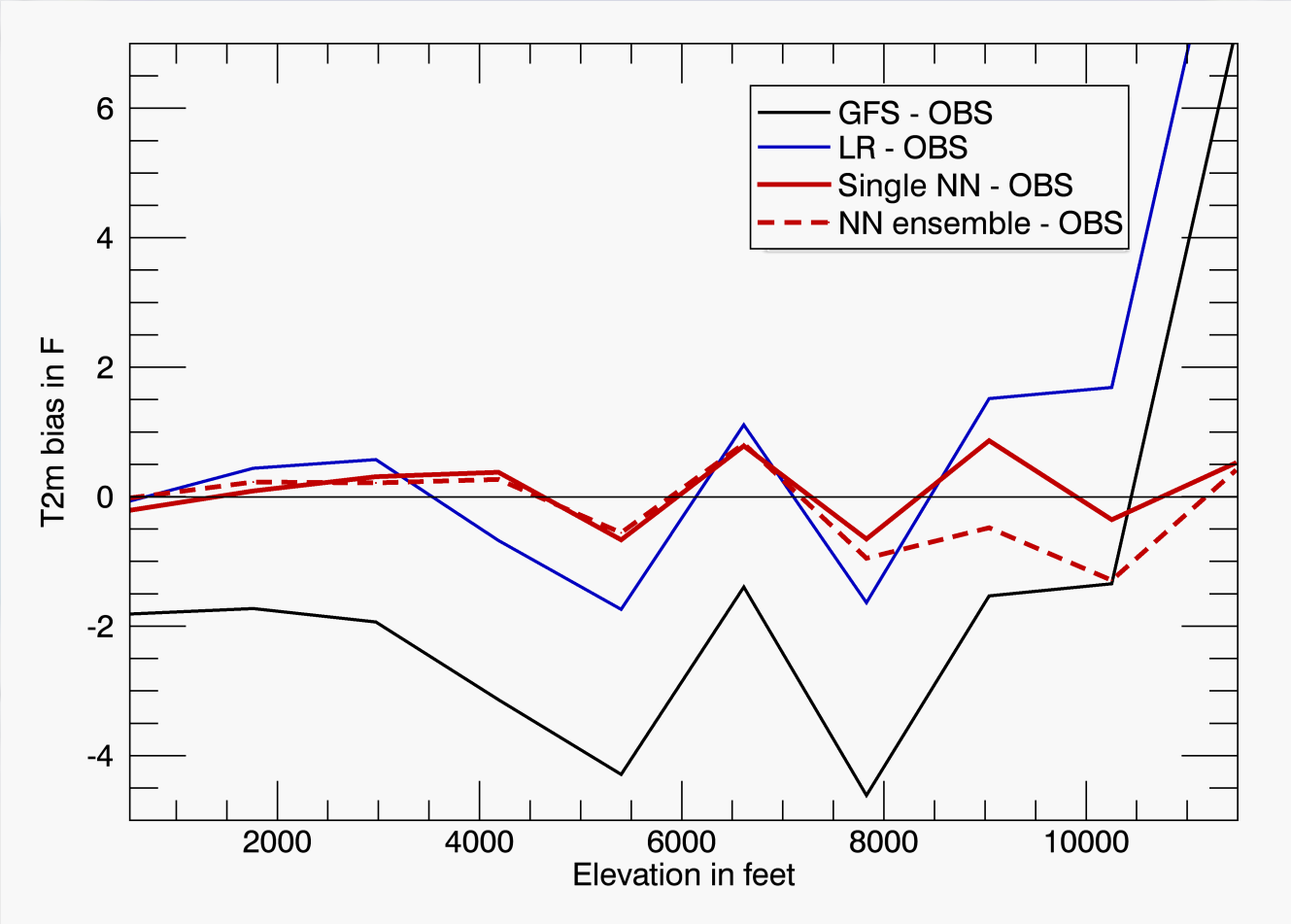
NN for all (3,000) stations, two variables (NN – 10:3:2)

in#	in	out
0	Sind(jd)	T2mo[0](jd+1)
1	Cosd(jd)	Tdo[0](jd+1)
2	Lat	
3	Lon	
4	GFS - T2m(jd)	
5	GFS – Td(jd)	
6	GFS – rh(jd)	
7	GFS – Th(jd)	
8	GFS – wd(jd)	
9	elevation	

NN performance, comparison with Linear regression bias correction

	T_{2m} in °F			T_d in °F		
	Bias	RMSE	CC	Bias	RMSE	CC
OBS - GFS	1.38	5.16	0.972	0.67	4.84	0.966
Linear Regression	-0.081	4.84	0.973	0.238	4.56	0.969
Best single NN	0.068	4.11	0.98	0.321	4.16	0.974
Ensemble of NN	-0.078	4.09	0.981	0.306	4.13	0.975

Some results (cont.)



Conclusions

- NN is a generic and versatile AI technique
- There exist numerous applications in Numerical Weather Prediction that can be successfully approached using NNs
- In NWP models NNs can be used in model initialization, as parts of the model physics, and for post-processing model outputs
- A significant experience in developing NN applications for NWP models has been accumulated at EMC

There is no free lunch

- NN, as any statistical model, requires data for training
- NN, as any nonlinear statistical model, requires more data, than linear model/regression
- As any nonlinear statistical model, NN may be over fitted
- As any statistical model, NN should be periodically updated to changes in environment; NN can be updated on-line