Introduction to Machine Learning Applications for Numerical Weather Prediction Systems.*

V. Krasnopolsky

NOAA/NWS/NCEP/EMC

Acknowledgments: A. Belochitski, M. Fox-Rabinovitz, H. Tolman, Y. Lin, Y. Fan, J-H. Alves, R. Campos, S. Penny

*** By no means this overview should be considered comprehensive.**

Abstract

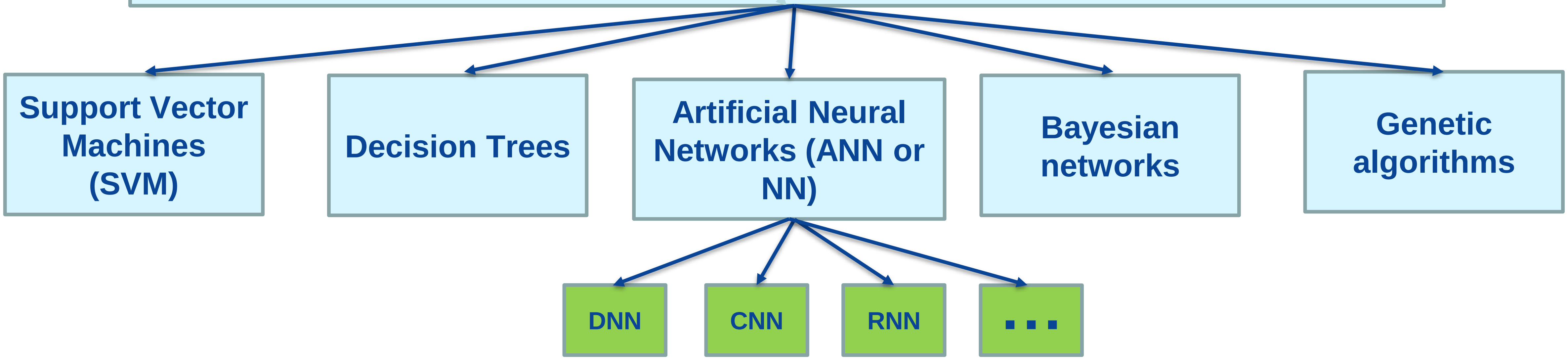
- Weather and Climate Numerical Modeling and related fields have been using ML for 25+ years
- Many successful ML applications have been developed in these fields
- **Our current plans for using ML are build on the solid basis of our community previous experience with ML in Weather and Climate Modeling and related fields**
- ML is a toolbox of versatile nonlinear statistical tools
- ML can solve or alleviate many problems but not any problem; **ML has a very broad but limited domain of application**

Outline

- I. Machin Learning
- II. Challenges
- III. A list of developed ML applications
- VI. Several examples

What is ML?

- ML is a subset of artificial intelligence (AI)
- ML algorithms build mathematical/statistical models based on training data - ML is Learning from Data Approach
- ML toolbox includes among other tools:



Mapping

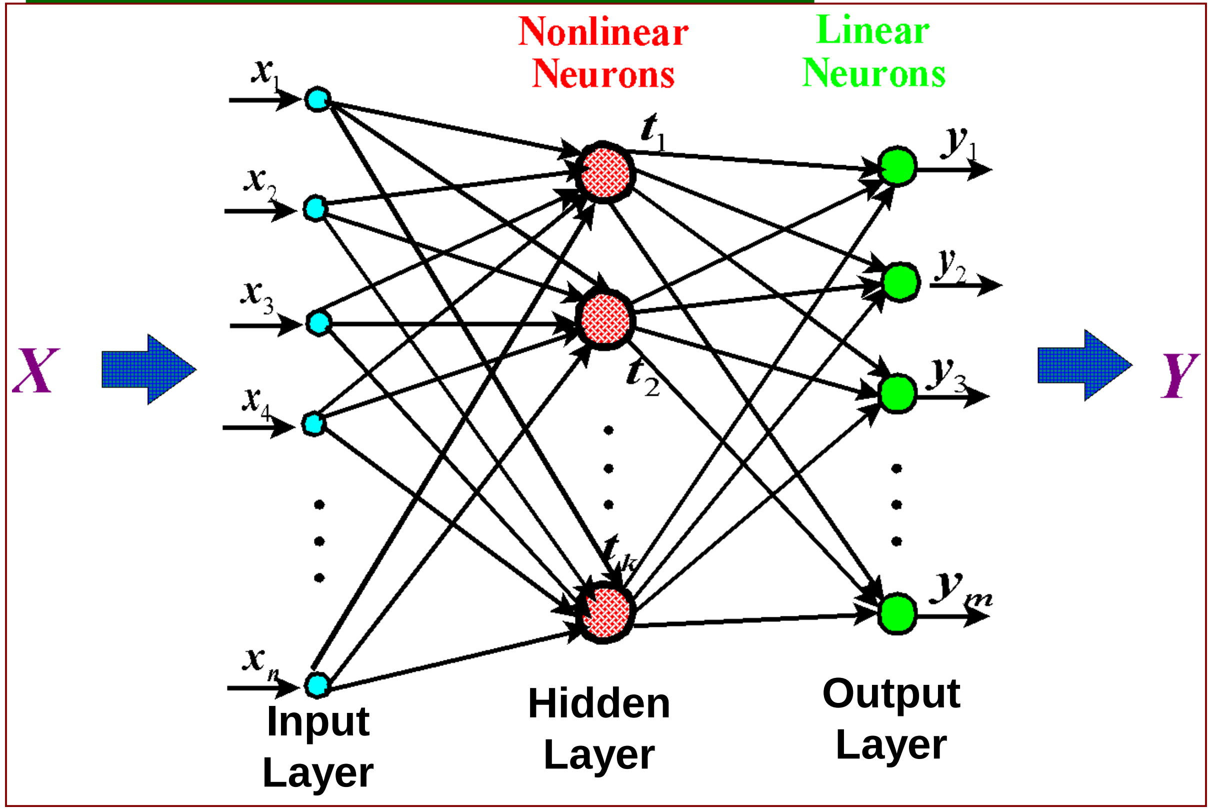
- Mapping: A *rule of correspondence established between two vectors* that associates each vector X of a vector space $\Re^n$ with a vector Y of another vector space $\Re^m$

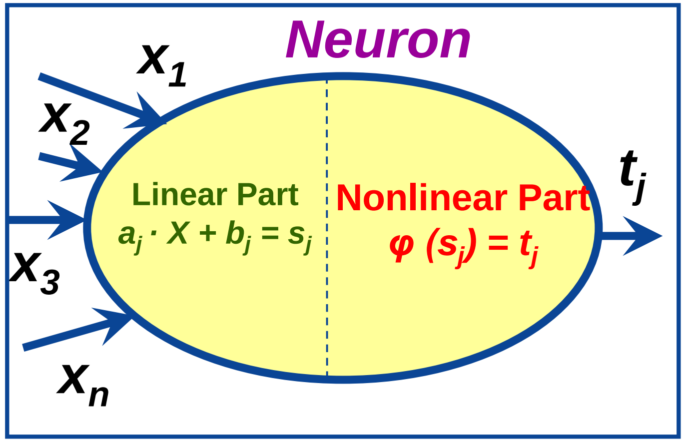
$$\left.\begin{array}{l} Y = F(X) \\ X = \{x_1, x_2, \ldots, x_n\}, \in \Re^n \\ Y = \{y_1, y_2, \ldots, y_m\}, \in \Re^m \end{array}\right\} \neq \begin{bmatrix} y_1 = f_1(x_1, x_2, \ldots, x_n) \\ y_2 = f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ y_m = f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}$$

<u>ML tools:</u> NNs, Support Vector Machines, Decision Trees, etc. are generic tools to approximate complex, nonlinear, multidimensional mappings.

NN - Continuous Input to Output Mapping

Multilayer Perceptron: Feed Forward, Fully Connected



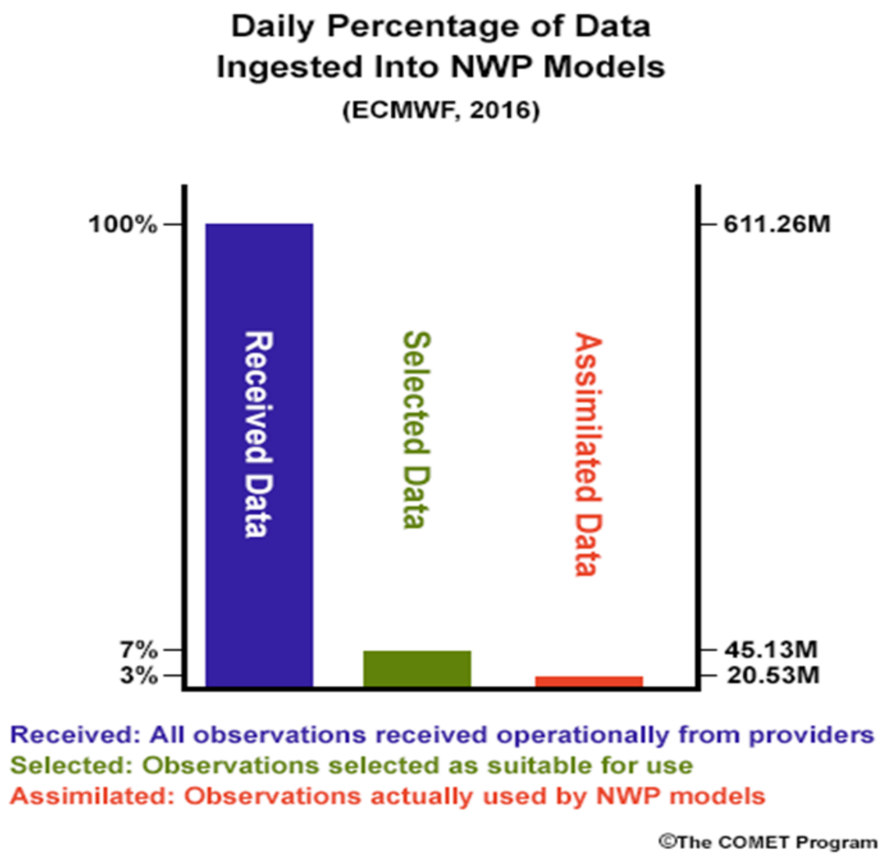


$$t_j = \phi\left(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i\right) =$$

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

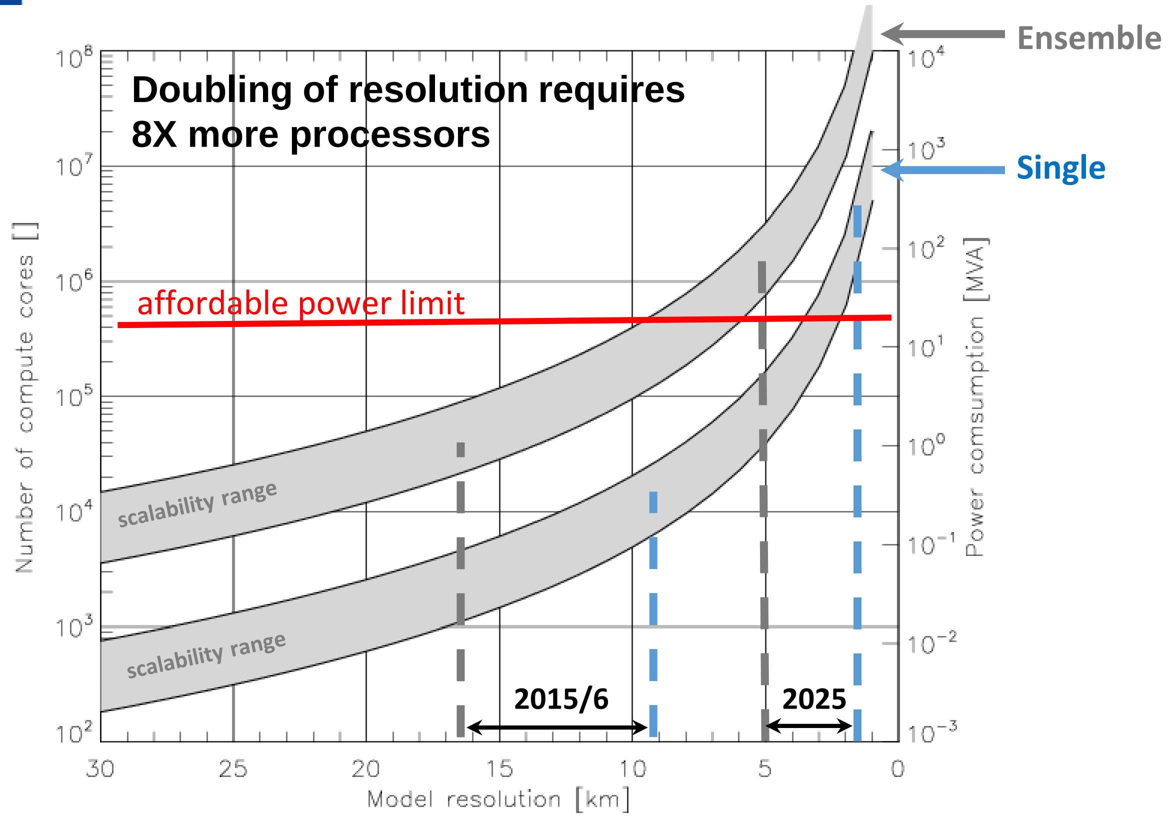
$$Y = F_{NN}(X) \Rightarrow \begin{cases} y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot t_j = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot \phi\left(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i\right) = \\ = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot \tanh\left(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i\right); \quad q = 1,2,\ldots,m \end{cases}$$

Why we need ML: Data challenge

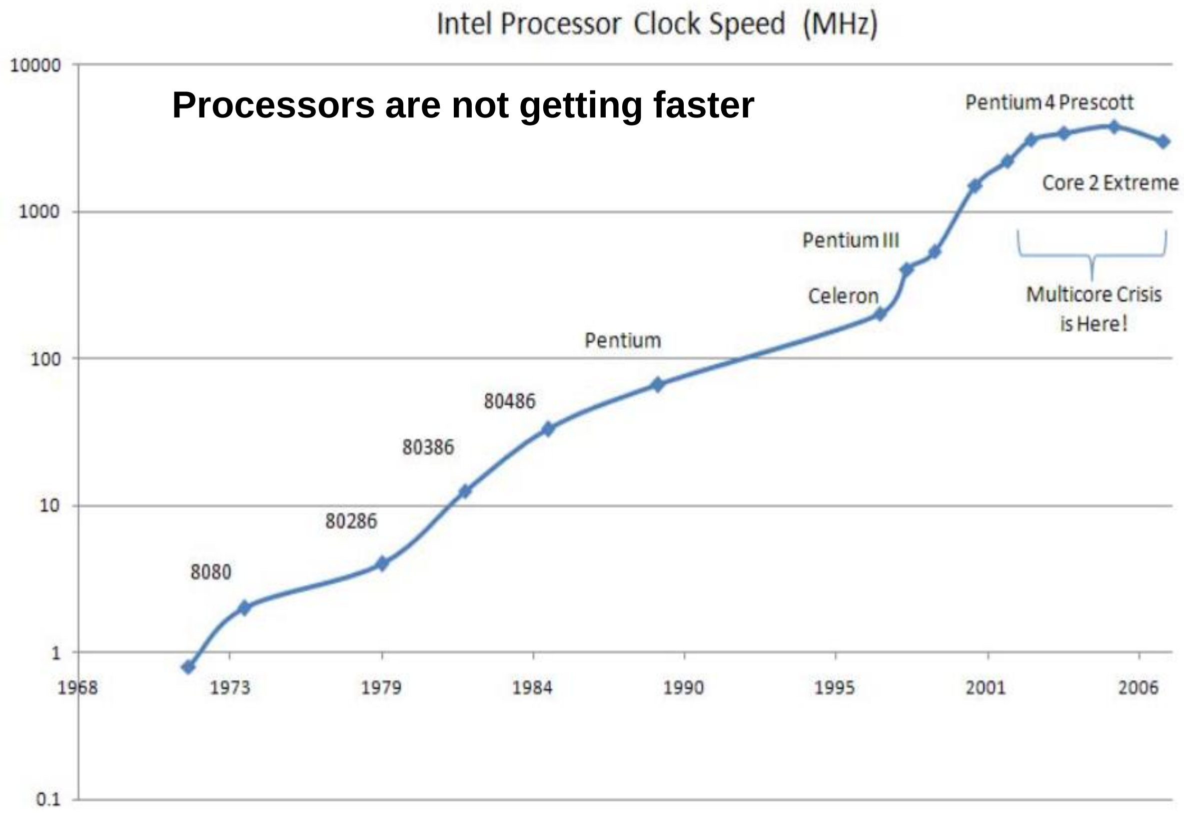


<u>ML response to the challenge:</u> Speed up data processing by orders of magnitude; improve extraction of information from the data; enhance assimilation of data in DASs

Why we need ML: Resolution Challenge

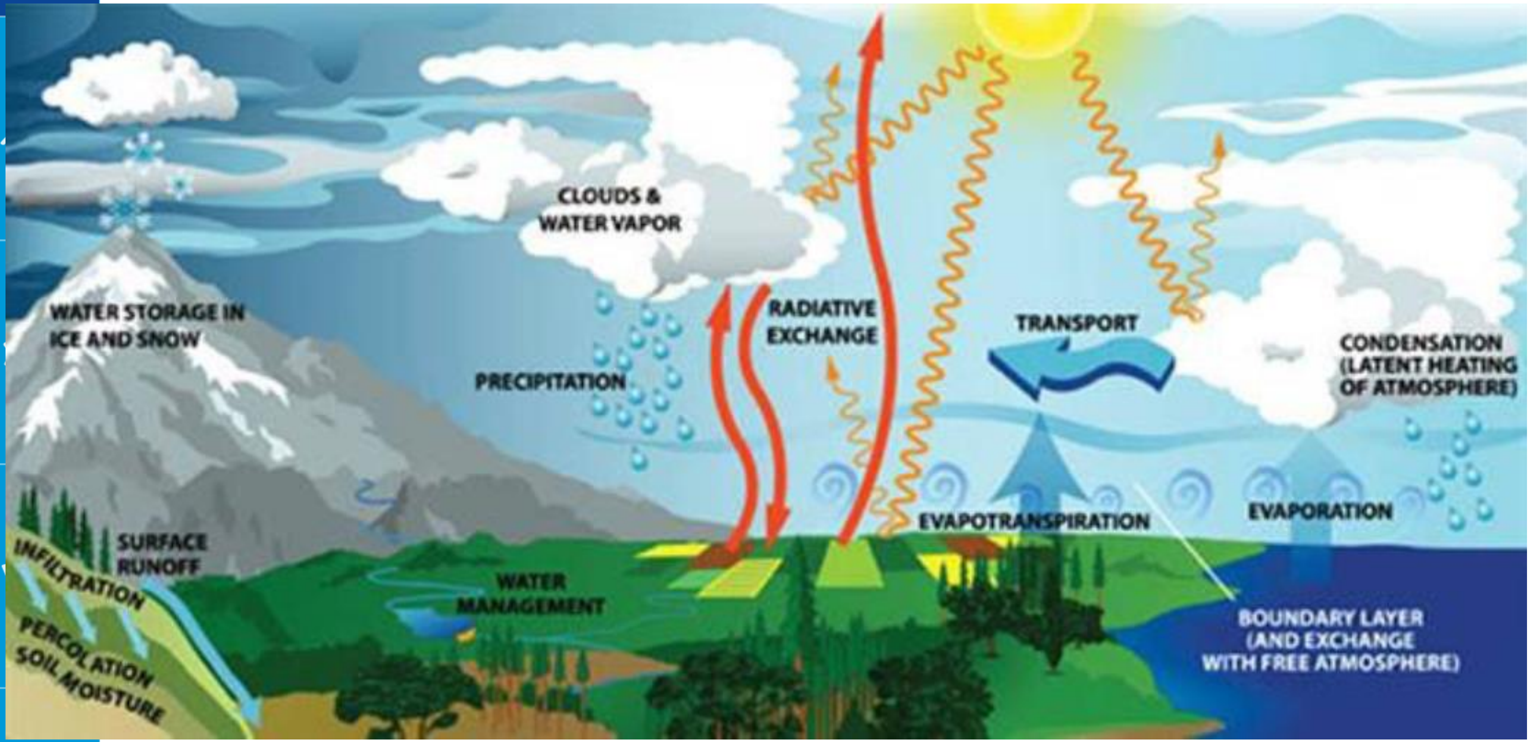


[ECMWF, Bauer et al. 2015]



<u>ML response to the challenge:</u> Speed up model calculations

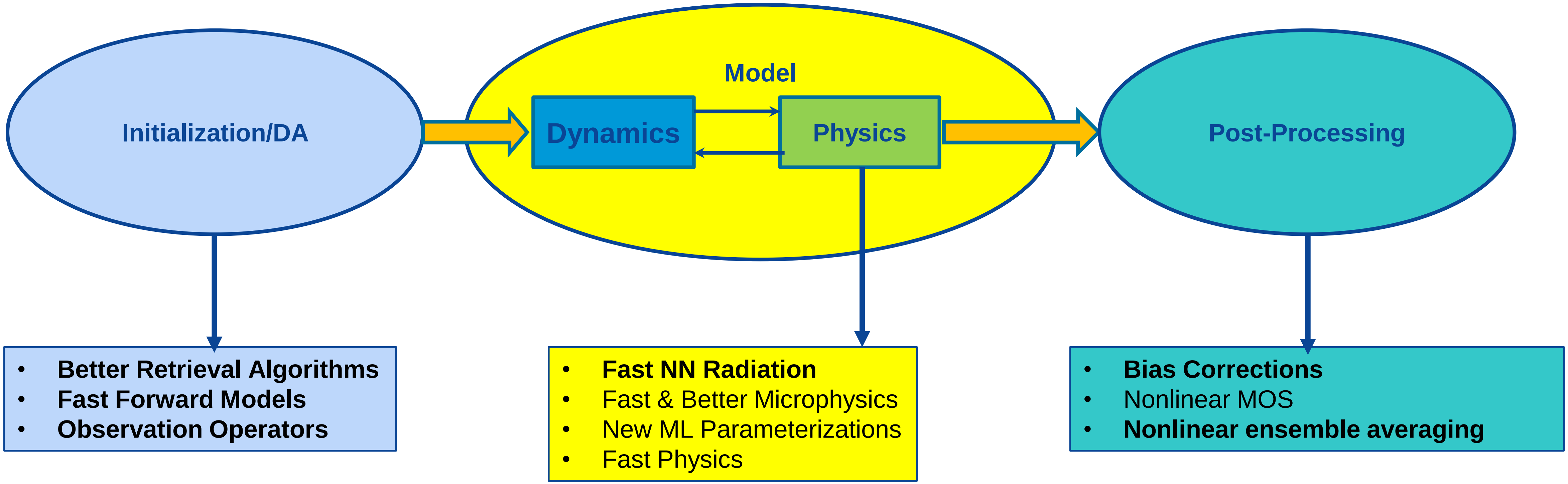
Why we need ML: Model Physics Challenge



- With increased resolution, scales of subgrid processes become smaller and smaller
- Subgrig processes have to be parameterized
- Physics of these processes is usually more complex
- The parametrizations are complex and slow

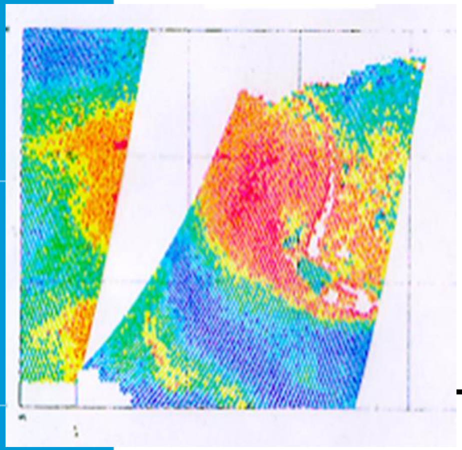
<u>ML response to the challenge:</u> Speed up calculations via developing fast ML emulations of existing parameterizations and developing fast new ML parameterizations

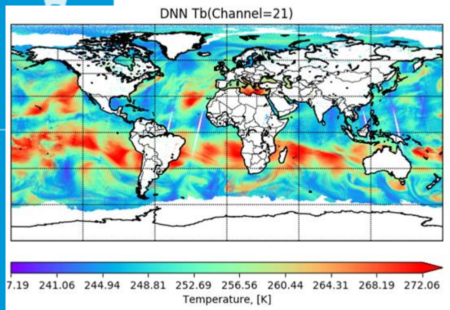
Using ML to Improve Numerical Weather/Climate Prediction Systems



I. ML for Model Initialization

- **Developed NN Applications (examples)**
  - *Satellite Retrievals*
    - Fast ML retrieval algorithms based on inversion of fast ML emulations of RT models
      - Clement Atzberger, 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models, *Remote Sensing of Environment*, Volume 93, Issues 1–2, 53-67. https://doi.org/10.1016/j.rse.2004.06.016
    - ML empirical (based on data) retrieval algorithms
      - Krasnopolsky, V.M., et al., 1998. "A multi-parameter empirical ocean algorithm for SSM/I retrievals", *Canadian Journal of Remote Sensing*, Vol. 25, No. 5, pp. 486-503 (operational since 1998)
  - *Direct Assimilation*
    - ML fast forward models
      - H. Takenaka, et al., 2011. Estimation of solar radiation using a neural network based on radiative transfer. *Journal Of Geophysical Research*, Vol. 116, D08215, https://doi.org/10.1029/2009jd013337
  - *Assimilation of surface observations and chemical and biological observations*
    - ML empirical biological model for ocean color
      - Krasnopolsky, V., S. Nadiga, A. Mehra, and E. Bayler, 2018: Adjusting neural network to a particular problem: Neural network-based empirical biological model for chlorophyll concentration in the upper ocean. *Applied Computational Intelligence and Soft Computing*, 7057363, 10 pp. doi:10.1155/2018/7057363.
    - ML algorithm to fill gaps in ocean color fields
      - V. Krasnopolsky, S. Nadiga, A. Mehra, E. Bayler, and D. Behringer, 2016, “Neural Networks Technique for Filling Gaps in Satellite Measurements: Application to Ocean Color Observations”, *Computational Intelligence and Neuroscience*, Volume 2016 (2016), Article ID 6156513, 9 pages, doi:10.1155/2016/6156513

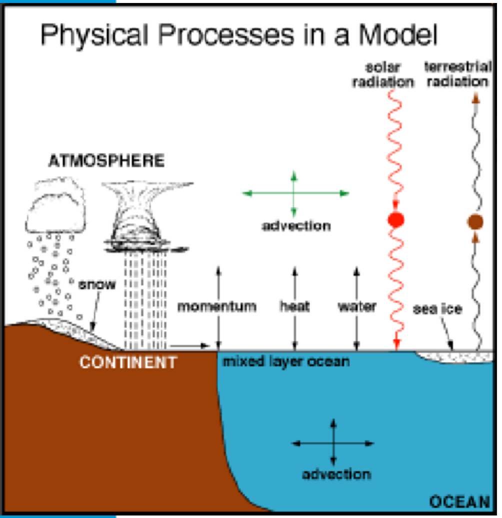




II. ML for Numerical Model

ML Applications developed & under development

- *Fast and accurate ML emulations of model physics*
    - Fast NN nonlinear wave-wave interaction for WaveWatch model
        - Tolman, et al.(2005). Neural network approximations for nonlinear interactions in wind wave spectra: direct mapping for wind seas in deep water. *Ocean Modelling*, 8, 253-278
    - Fast NN long and short wave radiation for NCEP CFS, GFS, and FV3GFS models
        - V. M. Krasnopolsky, M. S. Fox-Rabinovitz, Y. T. Hou, S. J. Lord, and A. A. Belochitski, 2010: "Accurate and Fast Neural Network Emulations of Model Radiation for the NCEP Coupled Climate Forecast System: Climate Simulations and Seasonal Predictions", *Monthly Weather Review*, 138, 1822-1842, doi: 10.1175/2009MWR3149.1
    - Fast NN emulation of super-parameterization (CRM in MMF)
        - Rasp, S., M. S. Pritchard, and P. Gentine, 2018: Deep learning to represent subgrid processes in climate models. *Proceed. National Academy Sci*., 115 (39), 9684–9689, doi:10.1073/pnas.1810286115
    - Fast NN PBL
        - J. Wang, P. Balaprakash, and R. Kotamarthi, 2019: Fast domain-aware neural network emulation of a planetary boundary layer parameterization in a numerical weather forecast model; in press, https://doi.org/10.5194/gmd-2019-79

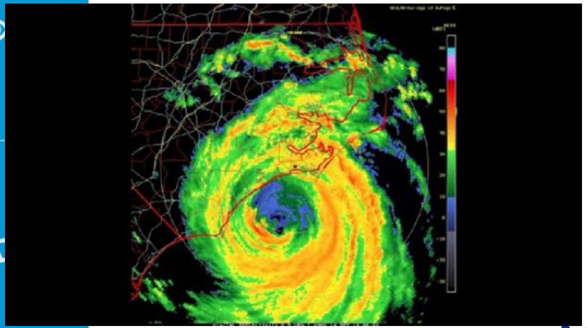


- *New ML parameterizations*
    - NN convection parameterization for GCM learned by NN from CRM simulated data
        - Brenowitz, N. D., and C. S. Bretherton, 2018: Prognostic validation of a neural network unified physics parameterization. *Geophys. Res. Lett.*, 35 (12), 6289–6298, doi:10.1029/2018GL078510.
- *ML emulation of simplified GCM*
    - Scher, S., 2018: Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. *Geophys. Res. Lett*., 45 (22), 12,616–12,622, doi:10.1029/2018GL080704.

III. ML for Post-processing

ML Applications Developed

- *Nonlinear ensembles*
    - Nonlinear multi-model NN ensemble for predicting precipitation rates over ConUS
        – Krasnopolsky, V. M., and Y. Lin, 2012: A neural network nonlinear multimodel ensemble to improve precipitation forecasts over Continental US. *Advances in Meteorology*, 649450, 11 pp. doi:10.1155/2012/649450.
    - Nonlinear NN averaging of wave models ensemble
        – Campos, R. M., V. Krasnopolsky, J.-H. G. M. Alves, and S. G. Penny, 2018: Nonlinear wave ensemble averaging in the Gulf of Mexico using neural network. *J. Atmos. Oceanic Technol*., 36 (1), 113–127, doi:10.1175/JTECH-D-18-0099.1.
    - Nonlinear NN ensemble for hurricanes: improving track and intensity
        – Shahroudi N., E. Maddy, S. Boukabara, V. Krasnopolsky, 2019: Improvement to Hurricane Track and Intensity Forecast by Exploiting Satellite Data and Machine Learning. The 1st NOAA Workshop on Leveraging AI in the Exploitation of Satellite Earth Observations & Numerical Weather Prediction, https://www.star.nesdis.noaa.gov/star/documents/meetings/2019AI/Wednesday/S3-2_NOAAai2019_Shahroudi.pptx



– *Nonlinear bias corrections*
    - Nonlinear NN bias corrections
        – Rasp, S., and S. Lerch, 2018: Neural networks for postprocessing ensemble weather forecasts. *Mon. Wea. Rev*., 146 (10), 3885–3900, doi:10.1175/MWR-D-18-0187.1.
    - Nonlinear NN approach to improve CFS week 3 an 4 forecast
        – Fan Y., C-Y. Wu, J. Gottschalck, V. Krasnopolsky, 2019: Using Artificial Neural Networks to Improve CFS Week 3-4 Precipitation & 2m Temperature Forecasts, The 1st NOAA Workshop on Leveraging AI in the Exploitation of Satellite Earth Observations & Numerical Weather Prediction, https://www.star.nesdis.noaa.gov/star/documents/meetings/2019AI/Thursday/S5-6_NOAAai2019_Fan.pptx

Several Examples of ML Applications

Ingesting Satellite Data in DAS

- **Satellite Retrievals:**

$$\boldsymbol{G = f(S),}$$

S *– vector of satellite measurements;*

G *– vector of geophysical parameters;*

f *– transfer function or retrieval algorithm*

- **Direct Assimilation of Satellite Data:**

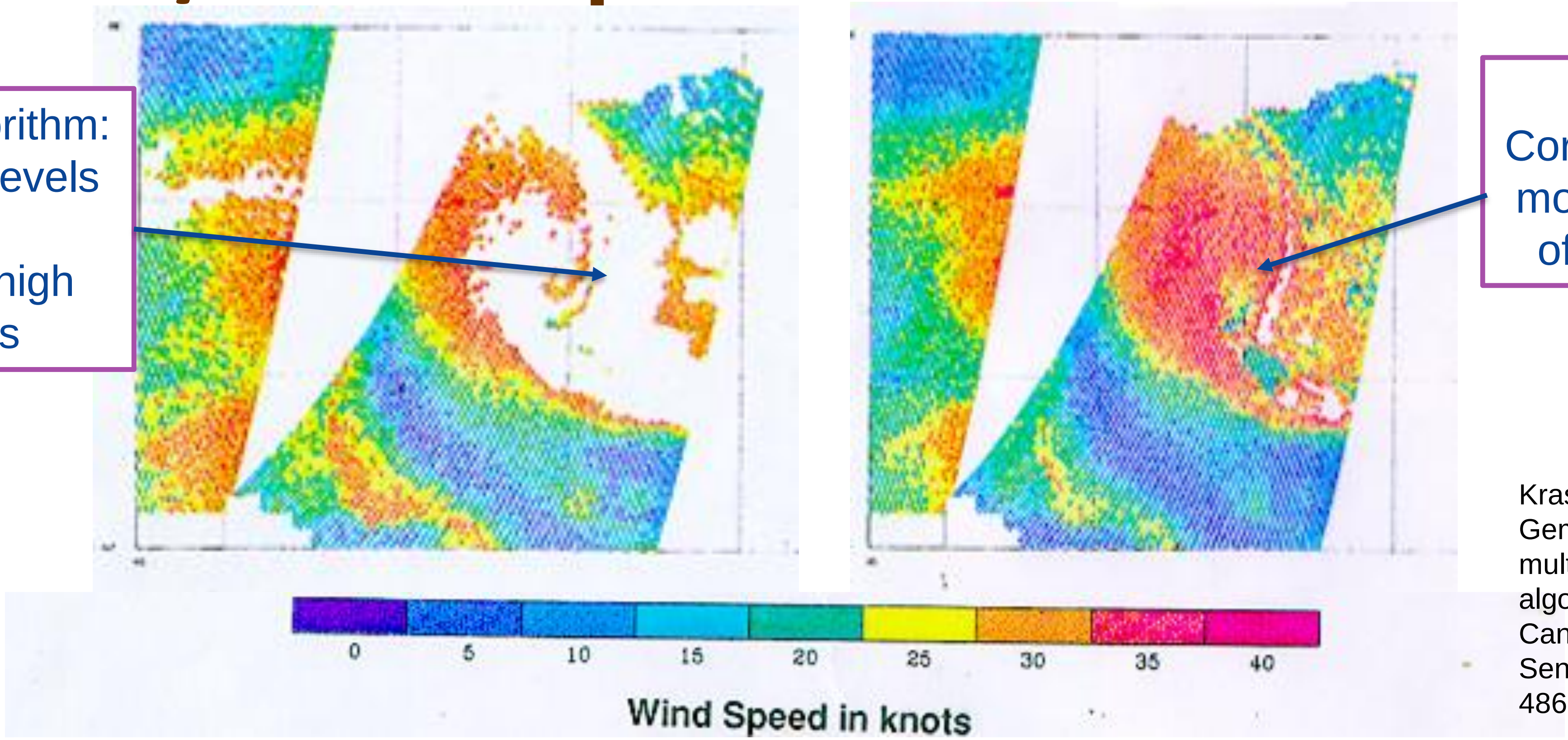
$$\boldsymbol{S = F(G),}$$

F *– forward model*

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

SSM/I Wind Speed Satellite Retrievals

Regression algorithm: Confuses high levels of moisture with high wind speeds



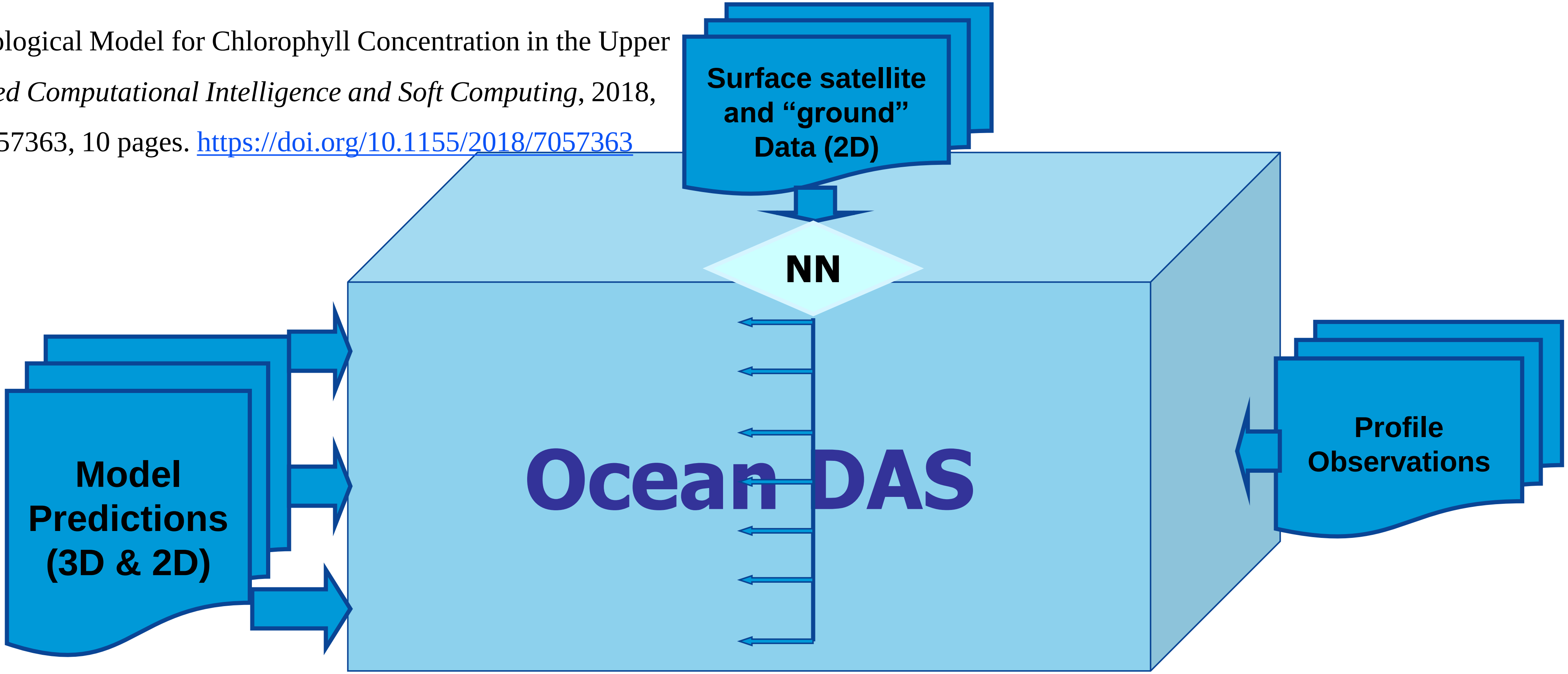
NN algorithm: Correctly retrieves the mostly energetic part of wind speed field

Krasnopolsky, V.M., W.H. Gemmill, and L.C. Breaker, "A multi-parameter empirical ocean algorithm for SSM/I retrievals", Canadian Journal of Remote Sensing, Vol. 25, No. 5, pp. 486-503, 1999

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: Propagating Information Vertically Using NNs, Assimilating Chemical and Bio data

Krasnopolsky, V., Nadiga, S., Mehra A., Bayler, E. (2018), Adjusting Neural Network to a Particular Problem: Neural Network-based Empirical Biological Model for Chlorophyll Concentration in the Upper Ocean, *Applied Computational Intelligence and Soft Computing*, 2018, Article ID 7057363, 10 pages. https://doi.org/10.1155/2018/7057363



NN – observation operator and/or empirical ecological model

ML Fast Model Physics

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

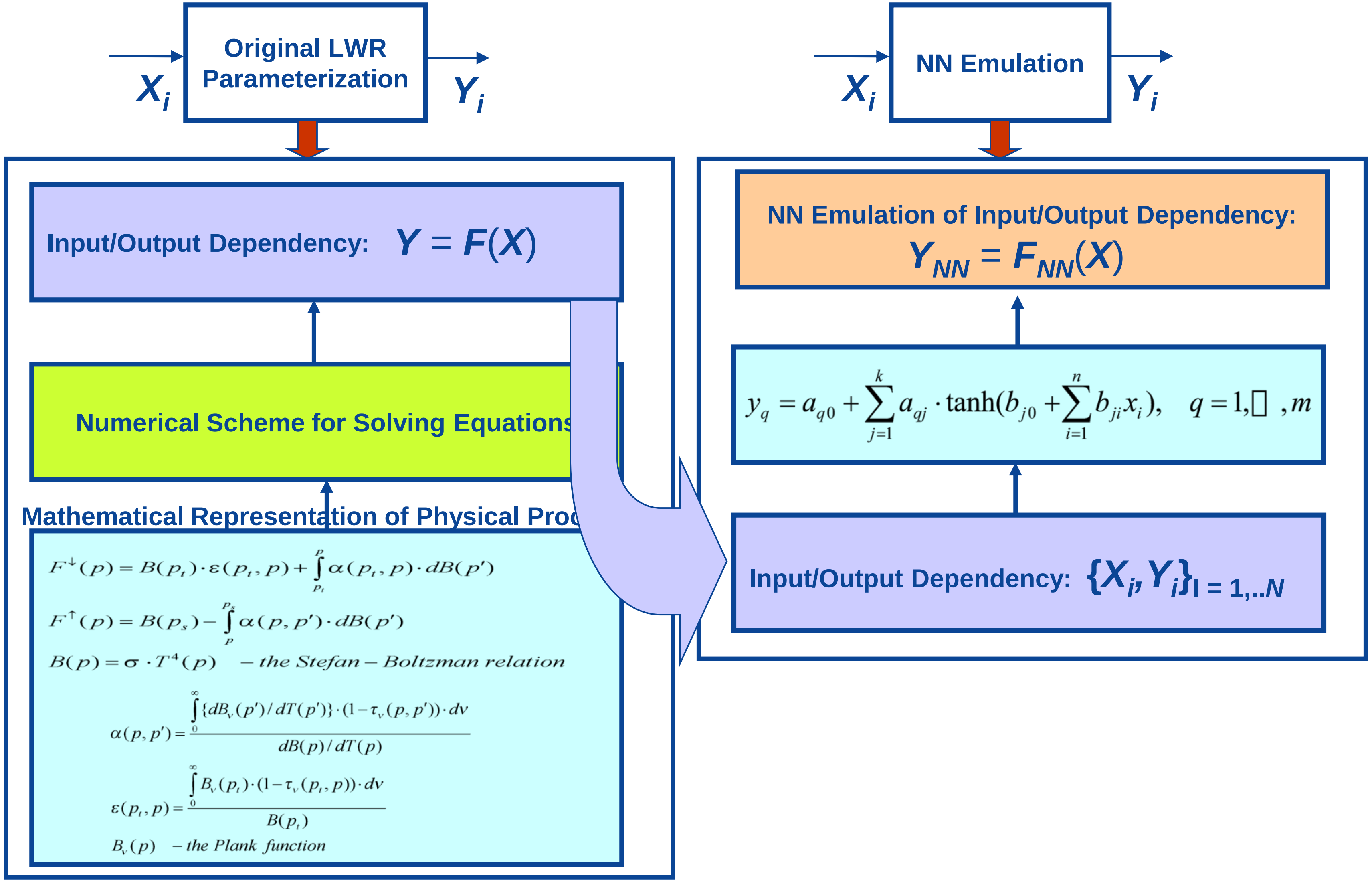
Model Dynamics

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

Model Physics

  - 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) or resolved physics
  - *P* - physics or parameterization of subgrid physical processes (1-D vertical r.h.s. forcing) **– mostly time consuming part > <u>50% of total time</u>**

The Magic of NN Performance (LWR)



Accurate and fast neural network (NN) emulations of long- and short-wave radiation parameterizations in NCEP GFS/CFS

- **Neural Networks perform radiative transfer calculations *much faster* than the RRTMG LWR and SWR parameterizations they emulate:**

	RRTMG LWR	RRTMG SWR
Average Speed Up by NN, *times*	16	60
Cloudy Column Speed Up by NN, *times*	20	88

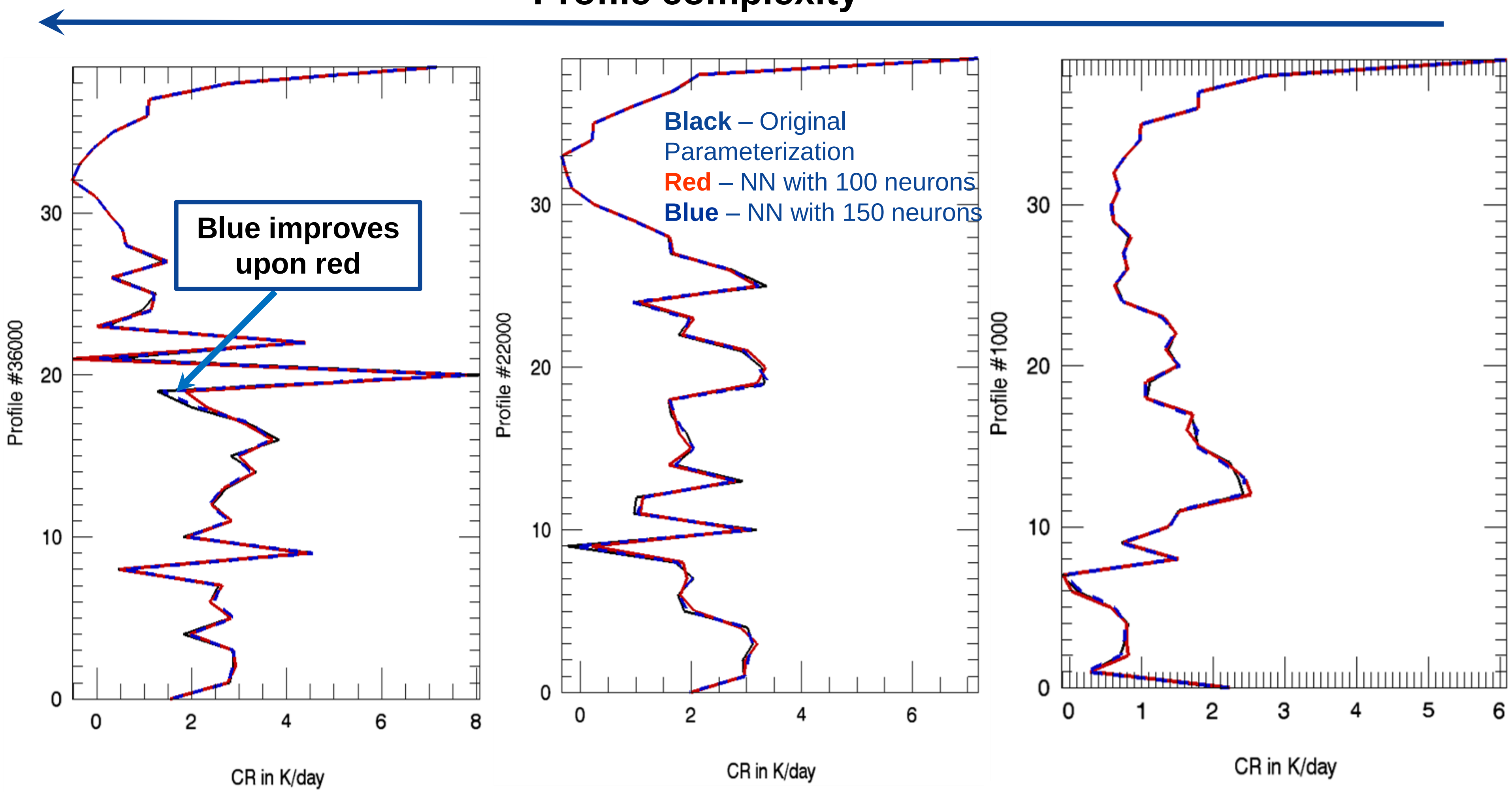
- **As a result of the speed up, GFS with NN radiation calculated with the same frequency as the rest of the model physics, or 12 times per model hour, takes up as much time as GFS with RRTMG radiation calculated only once per model hour.**

- **Neural network emulations are *unbiased*, and affect model evolution only as much as round off errors (see next slide).**

V. M. Krasnopolsky, M. S. Fox-Rabinovitz, Y. T. Hou, S. J. Lord, and A. A. Belochitski, 2010: "Accurate and Fast Neural Network Emulations of Model Radiation for the NCEP Coupled Climate Forecast System: Climate Simulations and Seasonal Predictions", *Monthly Weather Review*, 138, 1822-1842, doi: 10.1175/2009MWR3149.1

Individual LWR Heating Rates Profiles

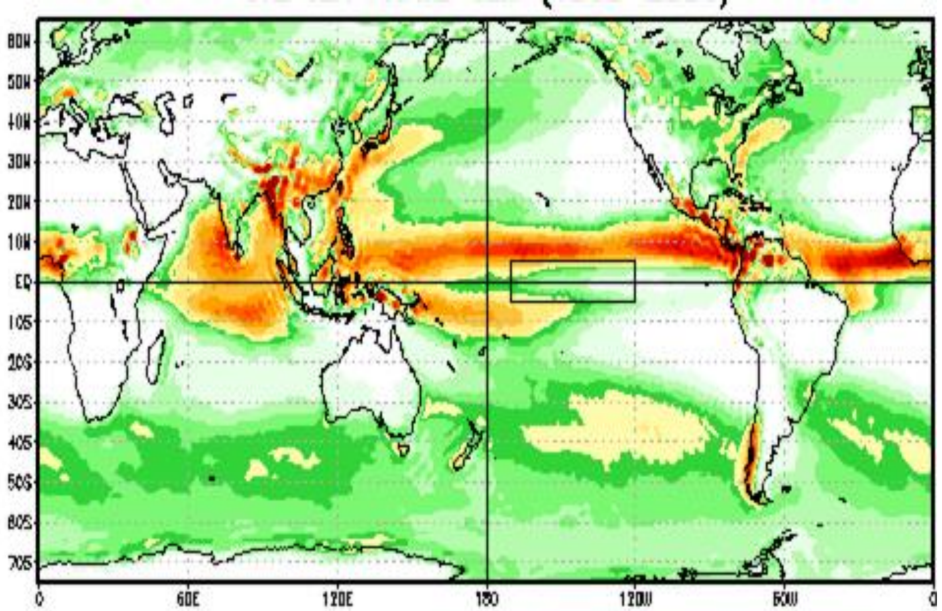
Profile complexity



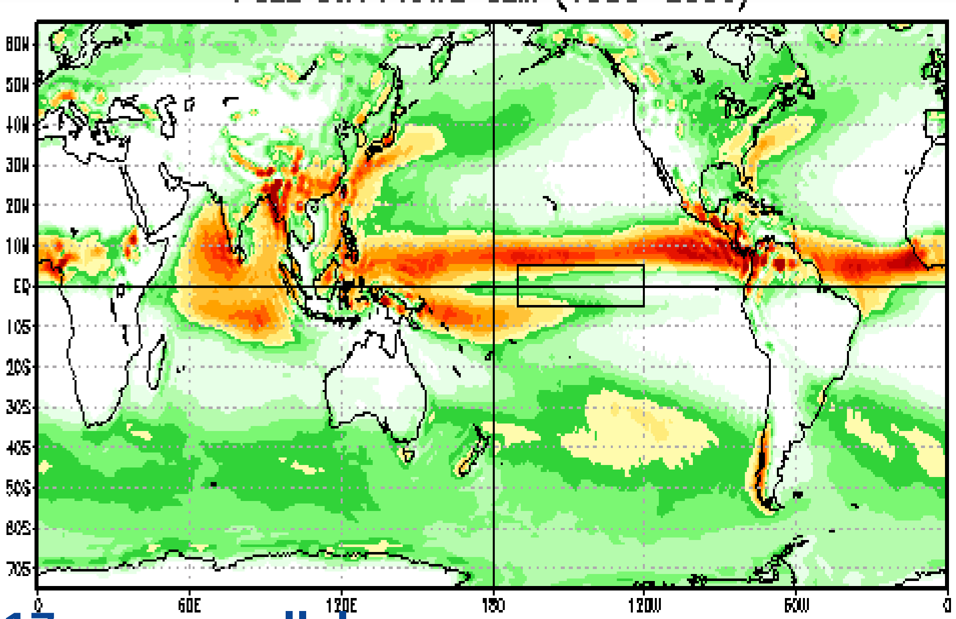
PRMSE = 0.18 & 0.10 K/day **PRMSE = 0.11 & 0.06 K/day** **PRMSE = 0.05 & 0.04 K/day**

NCEP CFS *PRATE* – 17 year parallel runs

CTL run with RRTMG LW and SW radiations

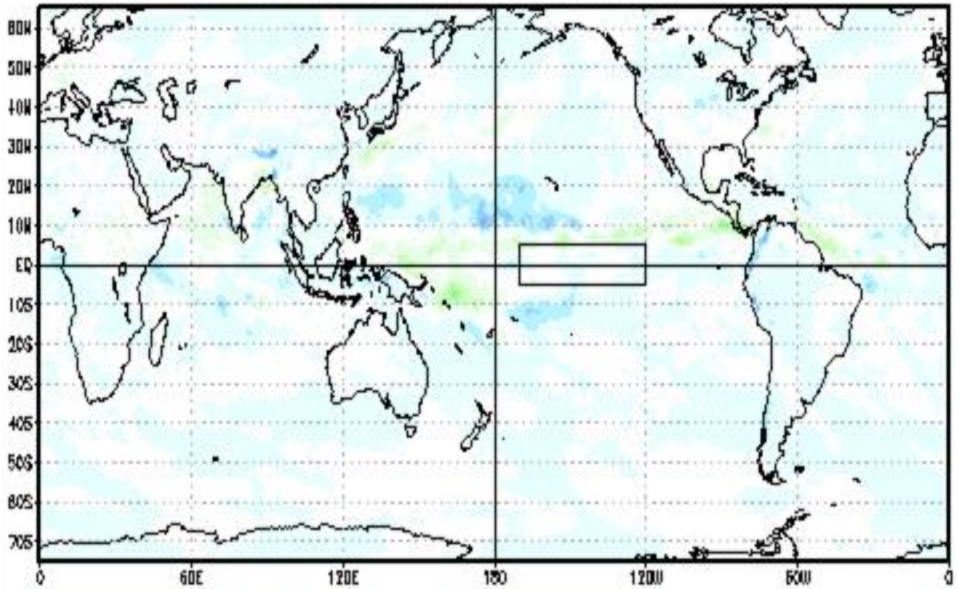


NN run with NN LW and SW radiations

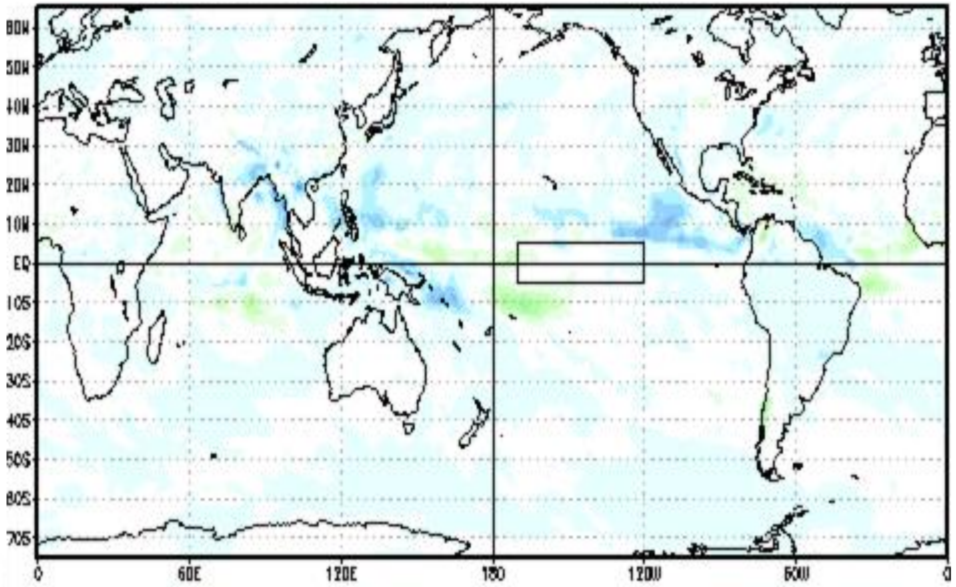


JJA

NN – CTL run differences



Differences between two control runs with different versions of FORTRAN compiler



Calculating Ensemble Mean

- **Conservative ensemble (standard)**:

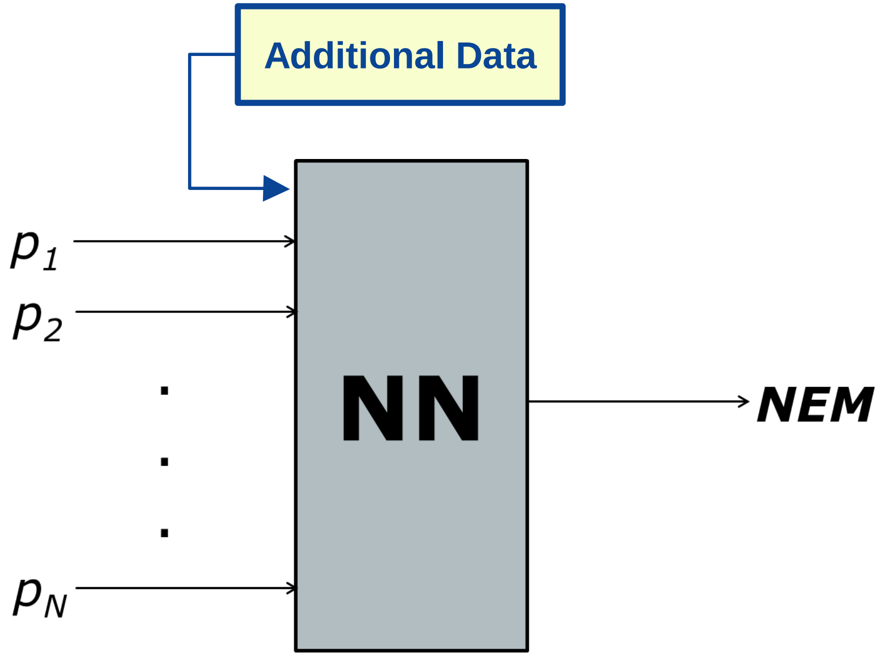
$$EM = \frac{1}{N}\sum_{i=1}^{N} p_i, \quad p_i \ \textit{is an ensemble member}$$

- **If past data are available, a nonlinear ensemble mean can be introduced:**

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

$$\boldsymbol{P} = \{p_1, p_2, \ldots, p_N\}$$

- **NN is trained on past data**

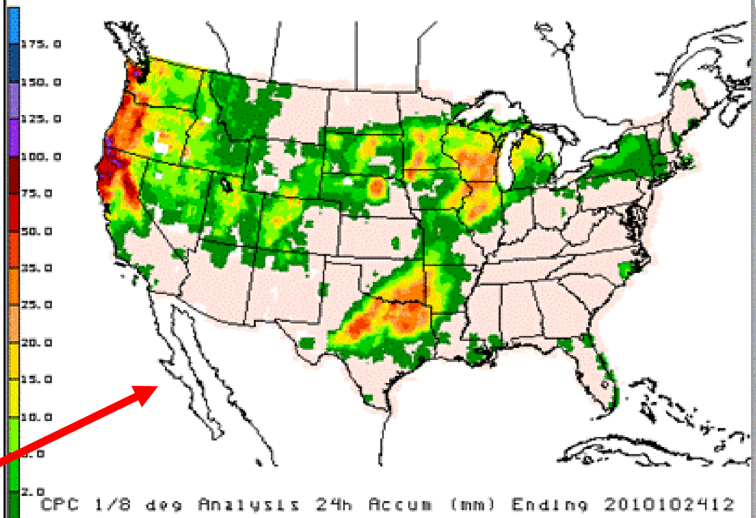


Example of ML (NN)-based Ensemble: Nonlinear Multi-model Ensemble Mean

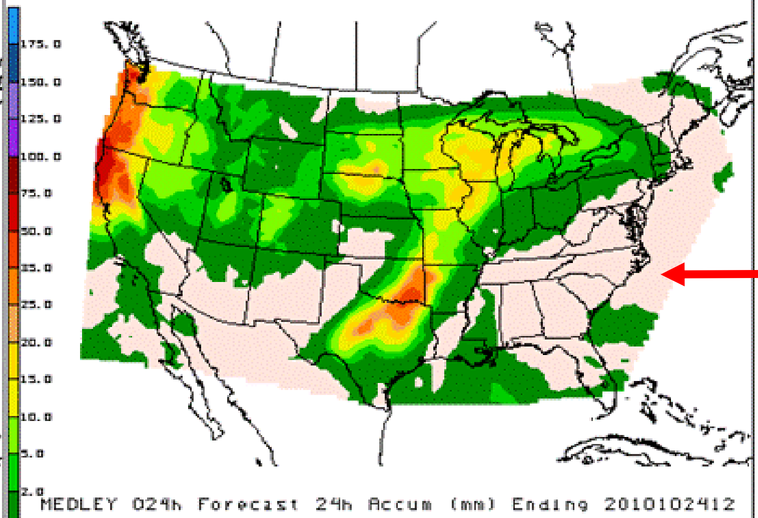
24 hour precipitation forecast over ConUS

Ensemble members:
NCEP (global and regional), UKMO, ECMWF, JMA, Canada (global and regional), German.

Verification Data



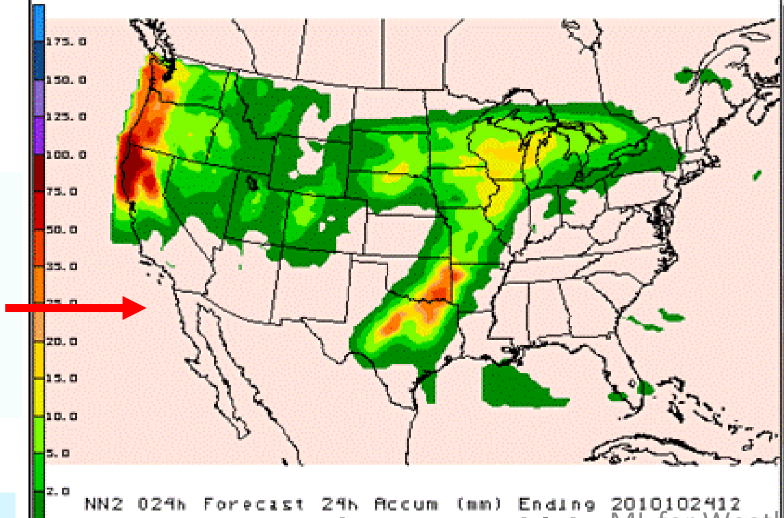
CPC analysis (ground truth)



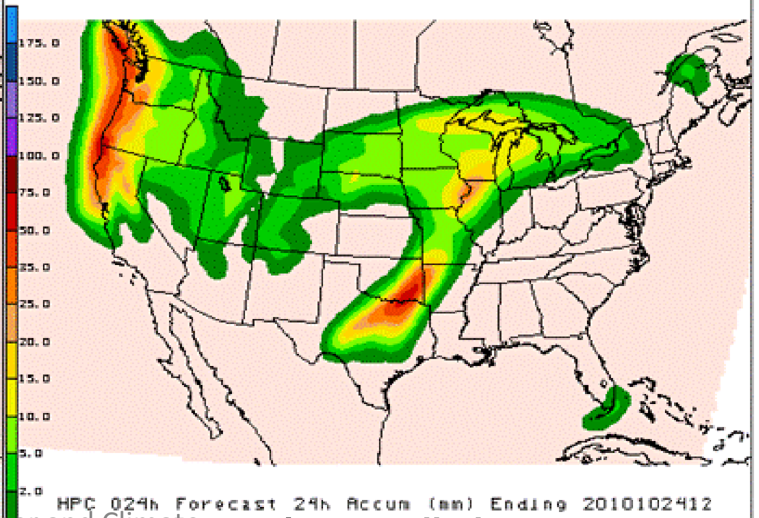
EM (arithmetic ensemble mean)

Reduced maximum and diffused sharpness and fronts. A lot of false alarms.
Due to slightly shifted maps from ensemble members.

ML-based Ensemble. Closer to CPC with maintained sharpness and minimal false alarm.



NEM (NN ensemble)



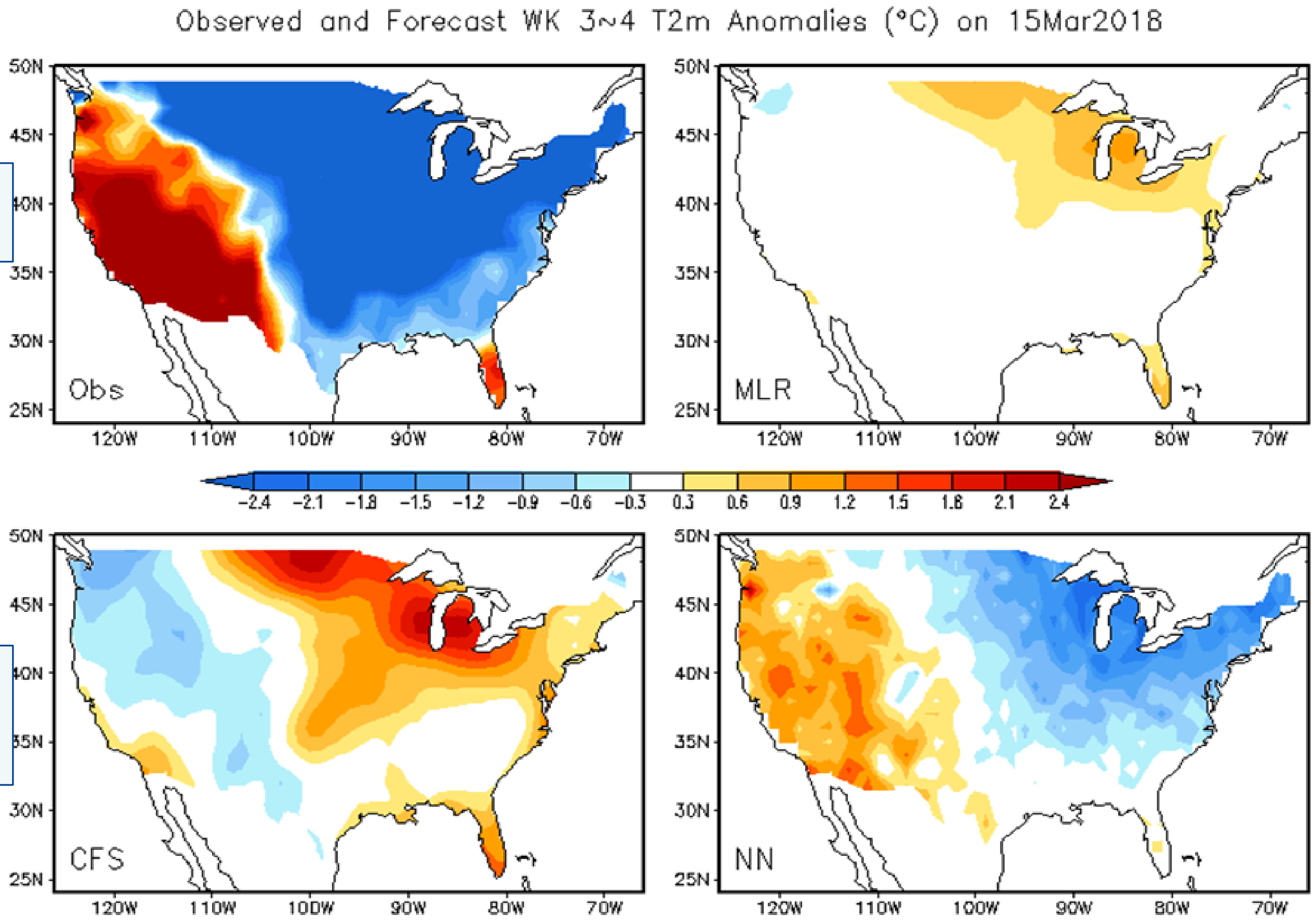
Human analyst prediction

Krasnopolsky, V. M., and Y. Lin, 2012: A neural network nonlinear multi-model ensemble to improve precipitation forecasts over Continental US. *Advances in Meteorology*, 649450, 11 pp. doi:10.1155/2012/649450

Neural Network Improves CFS Week 3-4 2 Meter Air Temperature Forecasts

Observed anomaly

Multiple Linear Regression



CFS ensemble (EM)

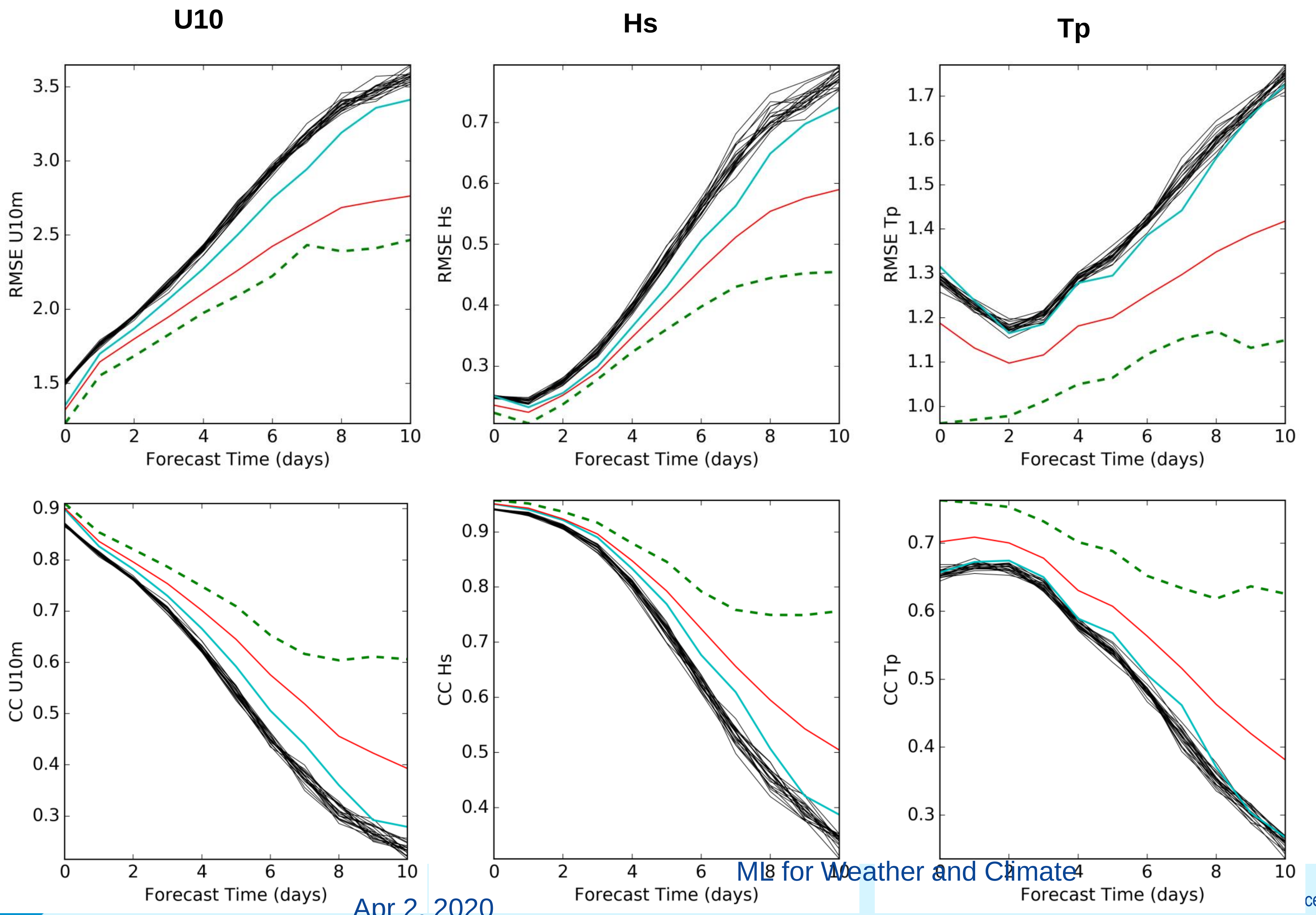
Neural Network (NEM)

Y. Fan, et al., 2019: Using Artificial Neural Networks to Improve CFS Week 3-4 Precipitation and 2 Meter Air Temperature Forecasts, submitted

NN wind-wave model ensemble (buoy data)

RMSE

CC

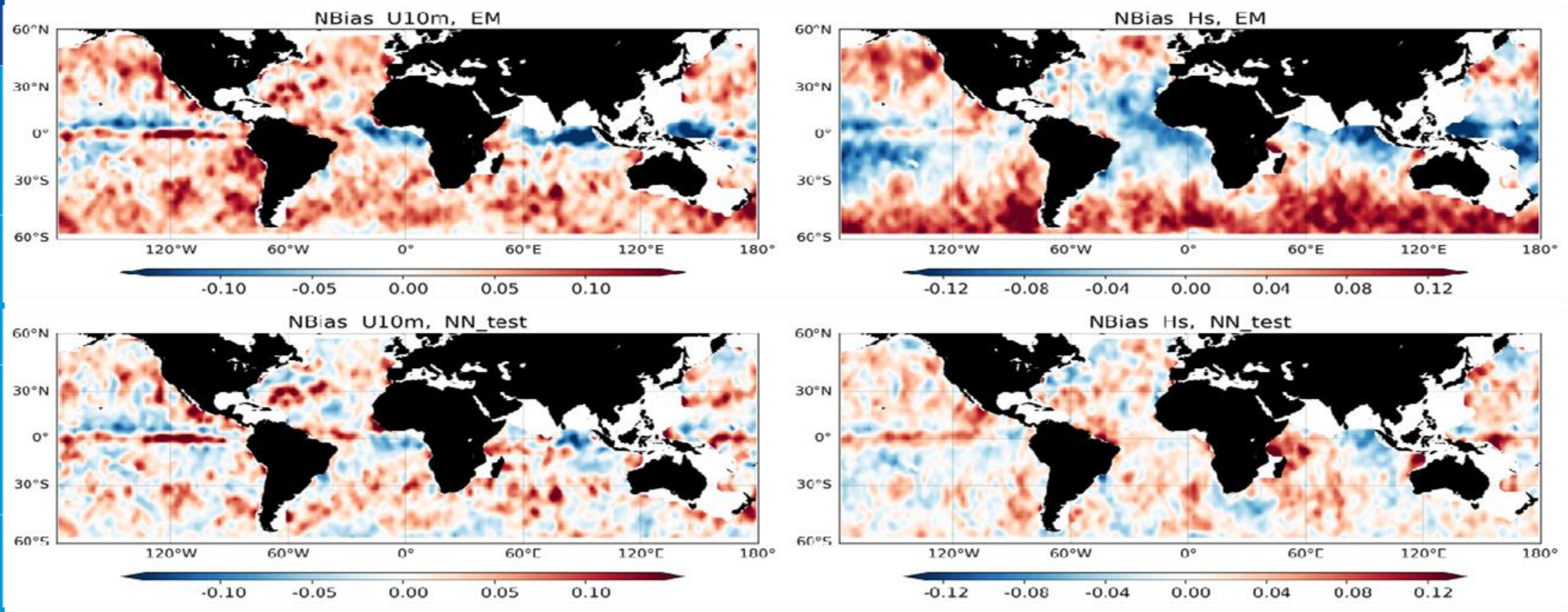


NCEP Global Wave Ensemble System
21 ensemble members

- Black: ensemble members
- Red: conservative ensemble mean (EM)
- Cyan: control run
-- Green: NN ensemble (NEM)

Campos, R. M., V. Krasnopolsky, J.-H. G. M. Alves, and S. G. Penny, 2018: Nonlinear wave ensemble averaging in the Gulf of Mexico using neural network. J. Atmos. Oceanic Technol., 36 (1), 113–127, doi:10.1175/JTECH-D-18-0099.1.

Global NN wind-wave model ensemble (altimeter data)



Normalized bias (NBias) for GWES ensemble mean (EM, top), and for NN ensemble mean (bottom) on an independent test set. The columns represent U10 (left) and Hs (right). Red indicates overestimation of the model compared to altimeter observations while blue indicates underestimation. Great part of large-scale biases in the mid- to high-latitudes has been eliminated by the NN ensemble mean simulation. *Campos, R. M., V. Krasnopolsky, J.-H. G. M. Alves, and S. G. Penny, 2020: Improving NCEP's Global-Scale Wave Ensemble Averages Using Neural Networks*, *Ocean Modeling*, 149, 101617

Summary of ML in NWP

MACHINE LEARNING:

- Is generic and versatile AI technique
- a lot of ML successful applications has been developed in NWP and related fields:
  - Model Initialization/data assimilation
  - Model improvements
  - Post-processing model outputs

SPEED UP MODEL CALCULATIONS

Fast ML Physics:
- -Radiation
- -Convection
- -Microphysics
- -PBL

Fast ML Chemistry and Biology
Fast Simplified ML GCMs

IMPROVE DATA UTILIZATION IN DAS

- Fast forward models for direct assimilation of radiances
- Improved retrieval algorithms
- Observation operators to better utilize surface observations
- Ecological models for assimilating chemical and biological data

SPEED UP CAN BE USED TO

- Improve parameterizations of physics
- Develop fast interactive chemistry and biology
- Increase the number of ensemble members in ensembles
- Increase model resolution

IMPROVE POST-PROCESSING

- -Bias corrections
- -Uncertainty prediction
- -Storm track and intensity
- -Ensemble averaging
- -Multi-model ensembles

BETTER PARAMETRIZATIONS

New parametrizations:
- From data simulated by higher resolution models
- From observed data

There is no free lunch

- ML has its domain of application; do not go beyond
- ML, as any statistical modeling, requires data for training; it is Learning from Data approach
- ML, as any nonlinear statistical modeling, requires more data, than linear models/regressions
- As any numerical models, ML applications should be periodically updated; however, ML can be updated on-line
- Interpretation of ML models, as any nonlinear statistical models, is not obvious

Some Additional References:

Krasnopolsky, V. (2013). The Application of Neural Networks in the Earth System Sciences. Neural Network Emulations for Complex Multidimensional Mappings. *Atmospheric and Oceanic Science Library. (Vol. 46),* 200pp., Springer: Dordrecht, Heidelberg, New York, London. DOI 10.1007/978-94-007-6073-8

Schneider T., Lan, S., Stuart, A., Teixeira, J. (2017). Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted High-Resolution Simulations, *Geophysical Research Letters*, 44, 12,396-12,417. https://doi.org/10.1002/2017GL076101

Dueben P.D. and Bauer P. (2018) . Challenges and design choices for global weather and climate models based on machine learning, *Geosci. Model Dev*., 11, 3999–4009, https://doi.org/10.5194/gmd-11-3999-2018

Cintra, R.S. and H. F. de Campos Velho (2018). Data Assimilation by Artificial Neural Networks for an Atmospheric General Circulation Model, DOI: 10.5772/intechopen.70791, https://www.intechopen.com/books/advanced-applications-for-artificial-neural-networks/data-assimilation-by-artificial-neural-networks-for-an-atmospheric-general-circulation-model

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., & Belochitski, A. A. (2013). Using Ensemble of Neural Networks to Learn Stochastic Convection Parameterization for Climate and Numerical Weather Prediction Models from Data Simulated by Cloud Resolving Model, *Advances in Artificial Neural Systems*, 2013, Article ID 485913, 13 pages. doi:10.1155/2013/485913

O'Gorman, P. A., and J. G. Dwyer, 2018: Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *Journal of Advances in Modeling Earth Systems*, 10 (10), 2548–2563, doi:10.1029/2018MS001351.

Krasnopolsky, V. M., M. S. Fox-Rabinovitz, and A. A. Belochitski, 2008: Decadal climate simulations using accurate and fast neural network emulation of full, longwave and shortwave, radiation. *Mon. Wea. Rev.,* 136 (10), 3683–3695, doi:10.1175/2008MWR2385.1.

Gentine, P., M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis, 2018: Could machine learning break the convection parameterization deadlock? *J. Geophys. Res*., 45 (11), 5742–5751, doi:10.1029/2018GL078202.

Questions?