U.S. Department of Commerce
National Oceanic and Atmospheric Administration

National Weather Service

National Centers for Environmental Predictions

5830 University Research Court

College Park, MD 20740

Office Note 513

https://doi.org/10.25923/71tx-4809

Review: Using Machine Learning for Data Assimilation, Model Physics, and Post-Processing model outputs

Vladimir Krasnopolsky

Environmental Modeling Center

National Centers for Environmental Predictions

College Park, Maryland

April 2023

Corresponding author address: 5830 University Research Court, College Park, MD 20740. E-mail: Vladimir.Krasnopolsky@noaa.gov

Contents

Abstract

In this paper machine learning (ML) applications developed at EMC and at other leading weather centers for numerical weather and climate modeling systems (NWCMS) are briefly introduced. The most important papers published in this field during recent years are reviewed. Advantages and limitations of the ML approach in applications to NWCMS are briefly discussed.

List of acronyms

CFS – NOAA climate forecast system
ConUS – continental United States
CRM – cloud resolving model
DA – data assimilation
DAS – data assimilation system
DL – deep learning
DNN – deep neural network
ECMWF – European Centre for Medium-Range Weather Forecasts
EMC – Environmental Modeling Center of NCEP NOAA
EMLP – emulating ML parameterization
EMOS – ensemble MOS
FM – forward model
GCM – global circulation model
GWFS – global ocean wave ensemble forecast system
GFS – NOAA global forecast system
GPU – graphics processing unit
LES – large eddies system
ML – machine learning
MLP – ML parameterization
MMF – multi-scale modeling framework
MOS – model output statistics
MP – microphysics
NN – neural network
NWCMS – numerical weather/climate modeling system
NWP – numerical weather prediction
PP – post-processing
QC – quality control
SFS – NOAA seasonal forecast system
SMLP – stochastic ML parameterization
SP – super parameterization
UFS – NOAA unified forecast system
UKMO - United Kingdom's Meteorological Office

I. Introduction

During the last several decades a well pronounced trend emerged in numerical weather and climate prediction. It marks a transition from simple linear or weakly nonlinear single-disciplinary models like simplified atmospheric or oceanic models that include a limited description of the physical processes, to complex nonlinear multidisciplinary systems or Numerical Weather/Climate Modeling Systems (NWCMS) like NOAA Unified Forecast System (UFS) and Seasonal Forecast System (SFS) with fully coupled atmosphere, land, ocean, ice and wave components, and NCEP GFS and CFS. Data assimilation system (DAS), model physics, and post-processing are three important components of NWCMS. Currently NWCMSs face four major challenges:

1. The vast amounts of observational data available from satellites, in-situ scientific measurements, and in future, from internet-of-things devices, increase with tremendous speed. Even now only a small percentage of the available data is used in the modern DASs. The problems with assimilation of new data in DAS range from growing time consuming (with increasing amounts of data) vs. limited computational resources to the necessity of new approaches to assimilate new types of data.
2. The increasing requirements to improve the accuracy and the forecast horizon of NWCMSs cause their growing complexity due to increasing horizontal and vertical resolutions and related increasing the complexity of model physics. Thus, global and regional modeling activities consume a tremendous amount of computing resources, which presents a significant challenge despite growing computing capabilities. Model ensemble systems have already faced the computational resources problem that limits the resolution and/or the number of the ensemble members in these systems.
3. Currently model physics is the most computationally demanding part of NWCMSs. With the increase of model resolutions, many subgrid physical processes that are currently parameterized become resolved processes and should be treated correspondingly. However, not always the nature of these processes is sufficiently understood to develop a description of the processes based on the first principles.

Also, with the increase of model resolution the scales of the subgrid processes that should be parameterized become smaller and smaller. Parameterizations of such processes often become more and more time consuming and sometimes less accurate because underlying physical principles may be not fully understood.

4. Current numerical models produce improved weather forecasts and climate projections with better accuracy. A major part of these improvements is due to the increase in supercomputing power that has enabled higher model resolution, better physics, and more comprehensive data assimilation (Bauer et al. 2015). Yet, the "demise of the 'laws' of Dennard and Moore" (Bauer et al. 2021, Khan et al. 2018) indicates that this progress is unlikely to continue due to increase in the required computer power. Moore's law drove the economics of computing by stating that every 18 months, the number of transistors on a chip would double at approximately equal cost. However, the cost per transistor starts to grow with the latest chip generations, indicating an end of this law. Thus, due to aforementioned limitations, results produced by the NWP and climate projecting systems still contain errors of a different nature. Thus, the post-processing (PP) correction of model output errors becomes even more important (Haupt et al. 2021). Currently used operational post-processing systems like MOS (Carter et al. 1989) are based on linear techniques (linear regressions). However, because optimal corrections of model outputs are nonlinear, for correcting biases of even regional fields, many millions of linear regressions are introduced in MOS (Gneiting et al. 2005, Wilks and Hamill 2007), making such systems cumbersome and resource consuming.

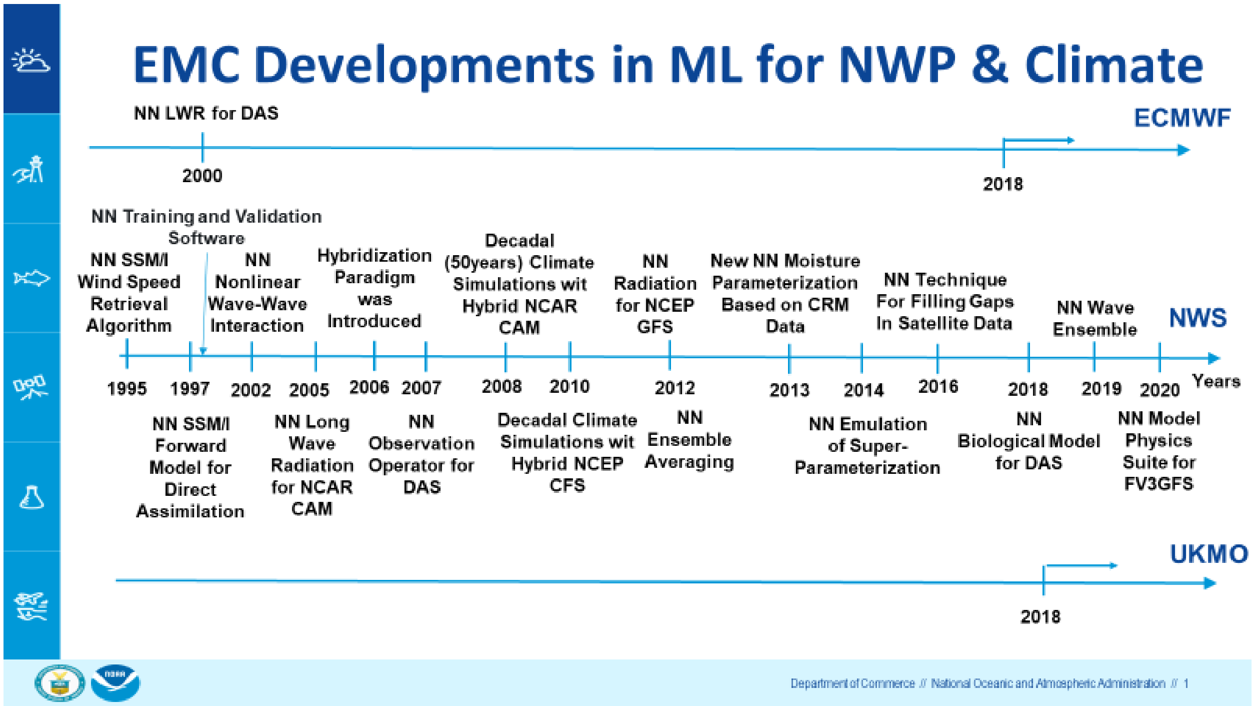

Figure 1. A timeline of ML developments at EMC.

Flexible and powerful numerical techniques are required to reduce growing demands for computer resources that outrun the actual growth of computer resources, enable new data types to be assimilated in DAS, meet aforementioned challenges of model physics, and develop flexible post-processing (PP) techniques to correct errors in model outputs. Developments in the various fields of artificial intelligence (AI), in particular in machine learning (ML), computer science, and statistics indicated the possibilities of using ML as one of such techniques. ML is increasingly being applied to solve and/or alleviate aforementioned problems of NWCMSs (e.g., Christensen and Zanna 2022, Dueben et al. 2021a, Haupt et. al 2008, Hsieh 2009, Krasnopolsky, 2013, Dueben and Bauer 2018, Boukabara et al 2019).

EMC scientists were among pioneers in the field of ML applications to NWCMS (see Fig. 1). They first developed many key approaches that are currently used in this field. EMC developments in this field during the period 1995 to 2012 are reviewed by Krasnopolsky (2013). The later developments are presented in Krasnopolsky (2020), Campos (2020) (and references there).

The majority of applications proposed in aforementioned works are based on two assumptions:

1. a large number of NWCMS applications, from a mathematical point of view, may be considered as *mapping, M,* that is a relationship between two vectors or two sets of parameters *X* and *Y*:

$$Y = M(X), \quad X \in R^n, Y \in R^m \tag{1}$$

   where *n* and *m* are the dimensionalities of vectors *X* and *Y* correspondingly.

2. ML provides an all-purpose non-linear fitting capability. Neural networks (NN), the major ML tool that is used in applications, are 'universal approximators' (Hornik, 1991) for complex multidimensional nonlinear mappings (Vapnik 1995, 2006, 2019). Such tools can be used and have been already used to develop a large variety of applications for NWCMSs.

Currently ML is considered as a powerful and prospective tool for further development and improvement of NWCMSs at ECMWF (e. g., Düben 2021b, see also Figs. 2 and 3), UKMO and other world centers. In accordance with NOAA AI Strategy, it is expected that the ML applications briefly described below and shown in Figures 4, 5, and 6 will be developed at NWS in close collaborations with Academy, NOAA Cooperative Institutes, NOAA Cooperative Science Centers, other NOAA divisions, private companies, and international communities.

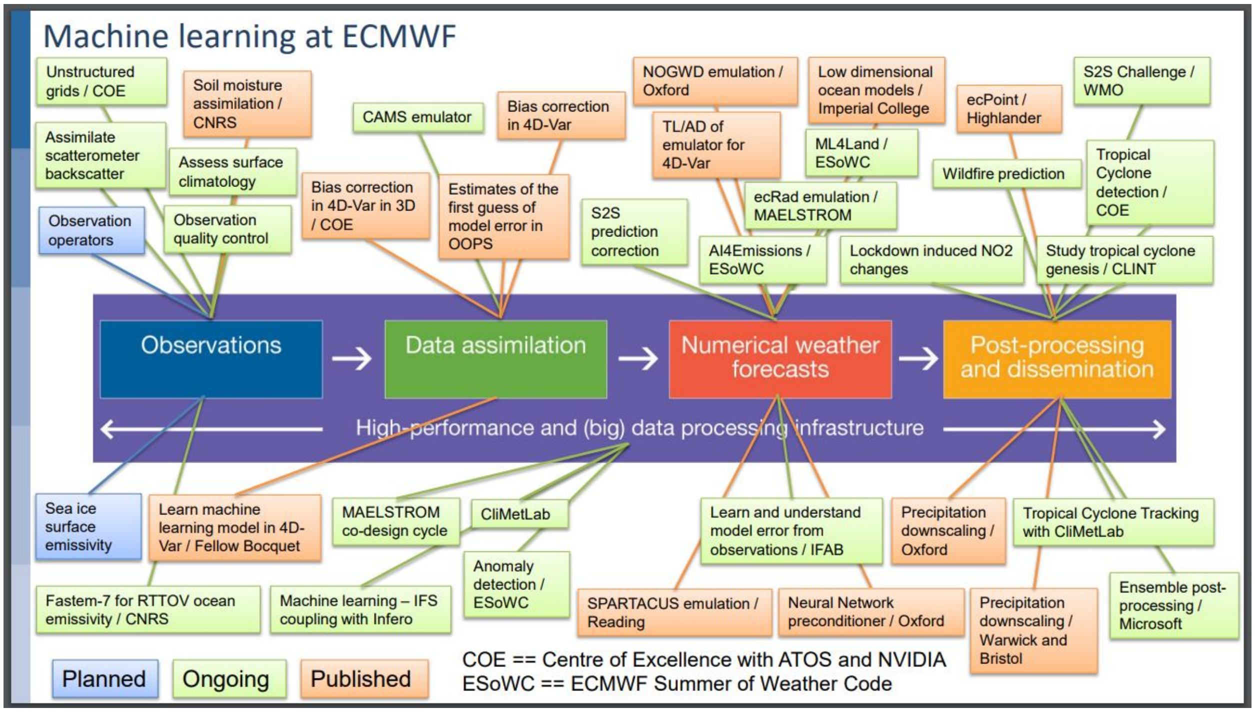

Figure 2: Machine learning applications at ECMWF that are already being explored or planned. The color-coding of the boxes corresponds to the respective component of the workflow for NWP (from Düben P., et al., 2021b)

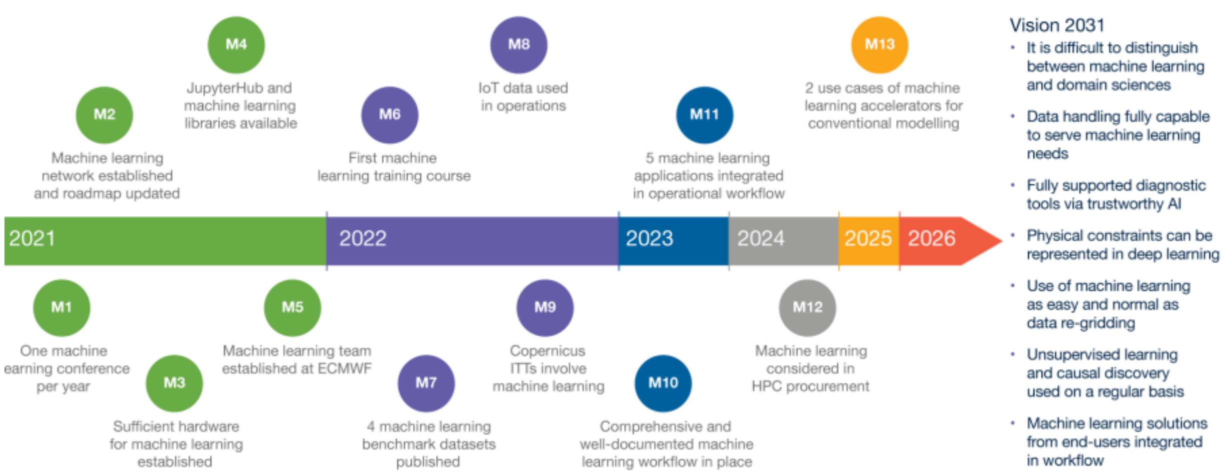

Figure 3: Timeline of machine learning developments at ECMWF with all milestones (from Düben P., et al., 2021b)’

Two major types of ML tools that have been applied to develop applications for NWCMS: (1) NNs (e.g., Krasnopolsky and Fox-Rabinovitz 2006; Brenowitz & Bretherton, 2018; Rasp et al., 2018, Geer 2020) and (2) tree algorithms (e.g., Belochitski et al., 2011; O'Gorman and Dwyer, 2018). There are many different types of NNs: shallow, deep, convolutional, recurrent, etc., as well as many types of tree algorithms. Advantages and limitations of different types of ML are discussed in detail in (Krasnopolsky and Belochitsky 2022).

Appendix 1 presents some estimates for time required for development at EMC of some ML applications reviewed in this note.

II. Machine Learning for Data Assimilation

Both DAS and ML, from a mathematical point of view, belong to the same class of optimization problems. Both methods apply a nonlinear optimization of an error function to determine optimal parameters of the system. Because DAS can be considered as a mapping between observations, first guess, and the final analysis, in principle, it may be possible to substitute the entire variational DAS with a ML DAS (Geer 2020, Boukabara et al. 2022, Dong et al 2022). However, while and if this approach is reaching maturity, it makes sense to focus on using ML for improvements of the existing EMC variational DAS. The following elements of the variational DAS are good candidates for applying ML.

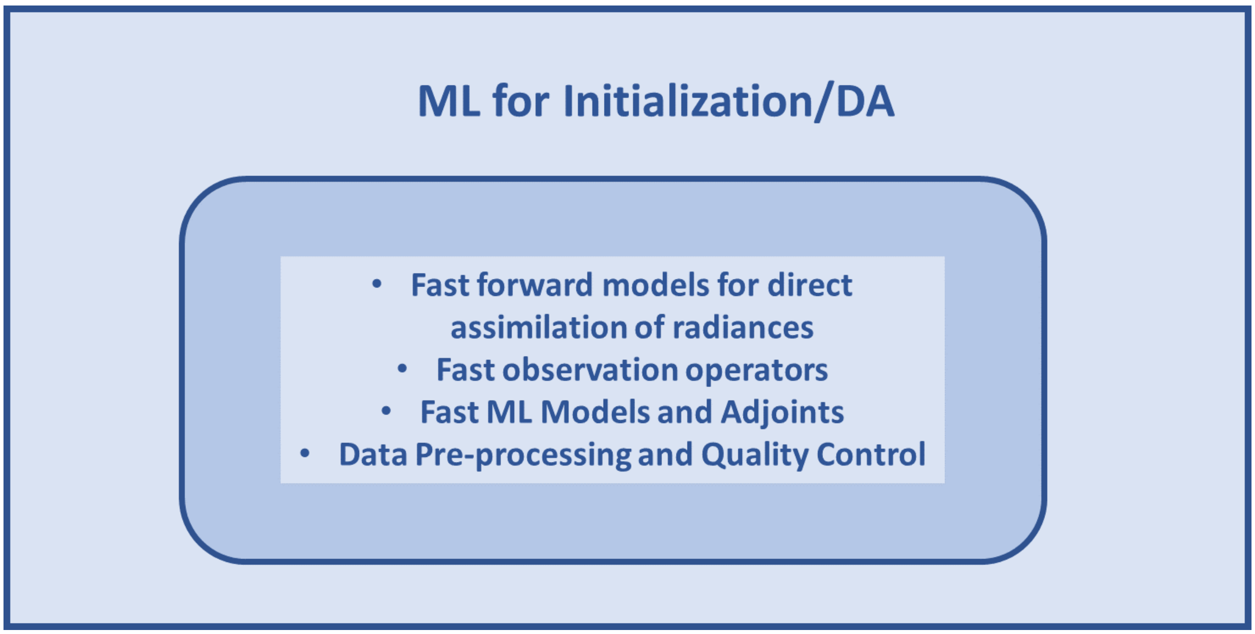

Figure 4. The summary of ML applications for EMC DAS that have been developed, are under development, and are considered.

II.1 Fast ML forward models for direct assimilation of satellite measurements

Forward models (FM) are used for direct assimilation of satellite radiances in DAS (Bauer et al. 2010). FMs are usually complex due to the complexity of the physical processes which they describe and the complexity of the first principle formalism on which they are based (e.g., radiative transfer theory). Thus, the dependence of satellite radiances on the geophysical parameters, which FMs describe, is a complicated and nonlinear mapping. These mappings may exhibit different types of nonlinear behavior. Direct assimilation is an iterative process where FMs and their Jacobians are calculated many times for each satellite measurement. As a result, this process becomes very time consuming and sometimes even prohibitively expensive for operational (real-time) applications.

For such applications, it is essential to have fast and accurate versions of FMs. Usually despite the functional complexity of FM mappings, ML techniques like NNs can provide fast and accurate emulations of FMs (Krasnopolsky 2013, Chapter 3.2). Moreover, an NN can also provide an entire Jacobian matrix with only a small additional computational effort.

II.2 Fast ML observation operators

When 2-D observations like surface winds, surface currents, SST, or sea surface elevation are assimilated into an atmospheric or oceanic DAS, the impact of these data in the DAS is mostly localized at the vertical level where they are assimilated. There is usually no explicit mechanism in the DAS to propagate the impact of these data to other vertical levels and to other variables except for error covariances and cross correlations in the variational solver that can to some extent spread the influence of 2-D observations to other vertical layers and other fields. Usually, this propagation occurs later, with a delay, during the integration of the model, in accordance with dependencies determined by the model physics and dynamics.

Recently, several attempts have been made to extract these dependencies from model simulations (Mellor and Ezer 1991) or observed data (Guinehut et al. 2004) in a simplified linear form for use in an ocean DAS to allow for 3-D assimilation of the 2-D surface data. However, these simplified and generalized linear dependencies that are often derived from local data sets do not properly represent the complicated nonlinear relationships (mappings) between the model variables. If we were able to extract or emulate these mappings in a simple, but not overly simplified and yet in an adequately nonlinear analytical form, they could be used in the DAS to facilitate a more effective 3-D assimilation of the 2-D surface data. ML observation operators have been developed for some surface observations (e.g., a ML observation operator for ocean surface elevation is described in Krasnopolsky 2013, Chapter 5.1.1). Also, assimilating chemical and biological observations in physical models that do not have corresponding prognostic variables require fast chemical and biological models to describe complex relationships between chemical/biological and physical prognostic variables. ML chemical and biological models can be built to play this role in DAS. For example, an ocean color NN empirical model has been developed (Krasnopolsky et al. 2018).

II.3 Fast ML Models and Adjoints

Fast hybrid and ML models for fast calculation of the first guess in DAS can be developed (Cheng et al. 2023, see also Sections III.4 and III.5 of this paper). Also, because some ML tools (e.g., NNs) are analytically differentiable, using such hybrid and ML models alleviates the problem of calculating adjoints, simplifying and speeding up calculations in 4Dvar DAS (Hatfield et al. 2021, Geer 2021, Maulik 2022). Although the differentiation of statistical models is an ill-posed problem, an NN ensemble technique has been developed to regularize the problem (Krasnopolsky 2007).

II.4 Data Pre-processing and Quality Control

ML promises to enhanced assimilation of satellite measurements, including radiances affected by clouds, precipitation, and surface properties (requiring more complete radiative transfer models accounting for these effects), and using improved or more efficient thinning, quality control (QC), observation bias correction, and cloud clearing procedures (e.g., Geer et al. 2018). There is the potential for ML techniques to help with QC decisions, either of the categorical (accept or reject) kind, or the more flexible "nonlinear" or "variational" kind where possibly dubious measurements are down weighted. For example, Sha et al. (2021) developed an automated DNN-based QC of precipitation for a sparse station observation network within the complex terrain area.

III. Machine Learning for Model Physics:

Any parameterization of model physics, even the entire model physics, and entire model, is a mapping (1) between a vector of input parameters (e.g., profiles of atmospheric state variables) and a vector of output parameters (e.g., a profile of heating rates in radiation parameterization). In terms of Y vs. X dependencies, parameterization mappings may be continuous or almost continuous, that is, they contain only a finite number of step-function-like discontinuities. Usually, parameterizations of physics do not contain singularities. ML can be used; (1) to develop emulating ML parameterizations (EMLP) that accurately emulate the original physically based parameterization schemes, speeding up the calculation by orders of magnitude; (2) when underlying physics of processes is not well understood, ML can be used to develop new ML parameterizations (MLP) by learning from data (simulated by high resolution models or/and observed); (3) ML as statistical tools can be used to develop stochastic ML parameterizations (SMLP).

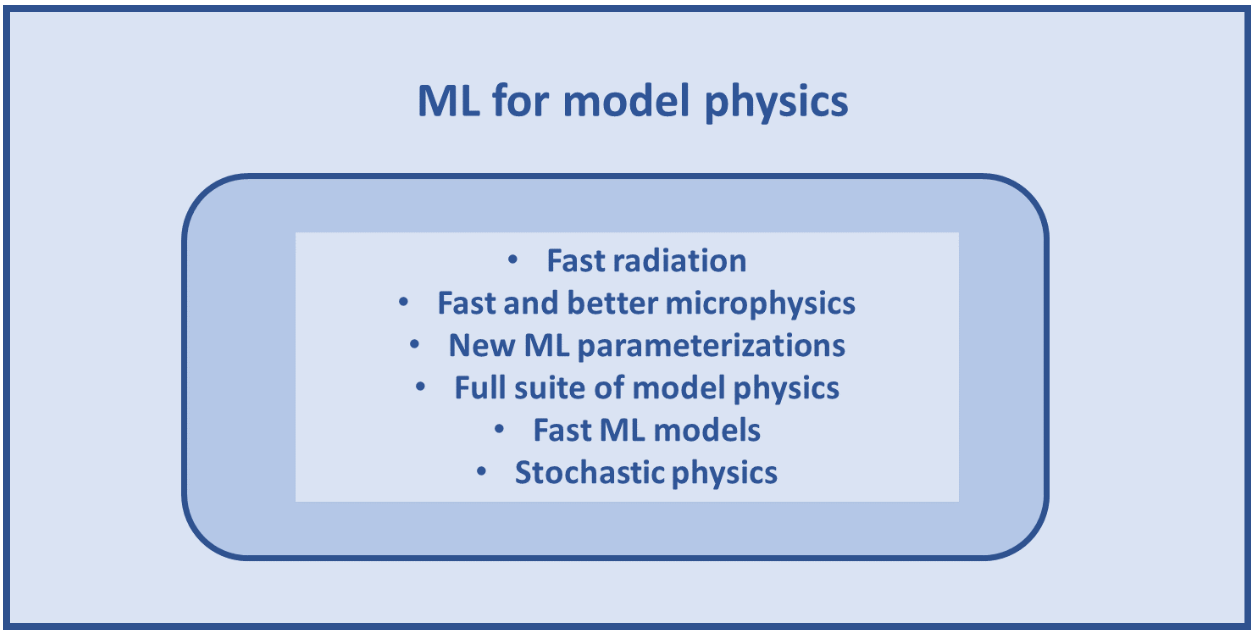

Figure 5. The summary of ML applications for EMC numerical models that have been developed, are under development, and are considered.

III.1 Fast ML radiation

Radiation parameterizations are among the most time-consuming components of model physics. Because of high computational cost they are never calculated at each time step and in each grid point of NWP models. At EMC and UKMO radiation is calculated every model hour and prorated in between. At ECMWF and Canadian Meteorological Center it is calculated at reduced horizontal or vertical resolution and then interpolated. Both these approaches are detrimental for the accuracy of the model forecast. Multiple NN emulators have been developed for radiation parameterizations (Chevallier et al.1998, 2000, Krasnopolsky et al. 2002, 2005, 2008b, 2010, 2012a, Pal et al. 2019, Roh and Song 2020, Ukkonen et al, 2020; Veerman et al, 2021, Lagerquist et al, 2021); however, to our knowledge, most of them have not yet been tested in an online setting to demonstrate their accuracy and stability in interactive coupling to an atmospheric model. EMC scientists demonstrated that accurate and fast radiation EMLPs can be developed for the CFS and GFS (Krasnopolsky et al. 2010, 2012a) that do not deteriorate the accuracy and stability of the model predictions and provide a speedup that allows calculating radiation at each time step in each grid point. They demonstrate high robustness and stability of EMLPs in model integrations even after significant changes in model physics and dynamics (Belochitski and Krasnopolsky 2022).

III.2 Fast and better ML Microphysics

State-of-the-art microphysical cloud modeling (e.g., Khain et al., 2000) is tremendously time consuming and cannot be introduced in atmospheric models without parameterization. Parameterizations significantly simplify the original microphysics and limit the number of atmospheric scenarios represented. However, even in a parameterized form microphysics calculations are computer resources and time consuming. Also, introducing parameterizations limits the number of atmospheric scenarios represented by each particular parameterization of MP. Often it is found that MP schemes perform well in certain atmospheric situations and perform not so well in others. When and why one scheme outperforms others is often not well understood. It appears that none of the existing MP parameterizations may offer comprehensive treatment of the natural processes involved.

In this case, ML tools can perform two different but related tasks when applied to MP parameterizations. First, ML can be used to create fast EMLPs by emulating various MP parameterizations; for example, the Thompson MP scheme (Thompson 2008) was emulated with an ensemble of shallow NNs (Krasnopolsky et al. 2017), Zhao-Carr microphysics was emulated by a two-layer vanilla recurrent neural network (Berkowitz et al. 2022), or Jensen et al. (2023) trained a random forest ML model, which is then used to predict supercooled large drops from several variables derived from High-Resolution Rapid Refresh model output. Second, ML tools can be applied to integrate existing MP parameterizations in a more comprehensive scheme that is able to offer better treatment of sub grid processes involved, cover a greater variety of sub grid scenarios, and stochastically represent uncertainty in MP schemes.

III.3 New ML parameterizations

The ML techniques can also be used to improve model physics. Because of the simplified parameterized physics that GCMs use, they cannot accurately simulate many important fine scale processes like cloudiness and convective precipitations (e.g., Brenowitz & Bretherton, 2018; Rasp et al., 2018). CRMs resolve many of the phenomena that lower resolution global and regional models do not resolve (e.g., higher resolution fluid dynamic motions supporting updrafts and downdrafts, convective organization, meso-scale circulations, and stratiform and convective components that interact with each other, etc).

An ML approach has been developed at EMC (Krasnopolsky et al., 2011) that uses ML/NN to develop a ML moisture parameterization trained using CRM simulated data. This MLP can be used as a moisture parameterization in GCMs and can effectively account for major sub-grid scale effects taken into account by other approaches (e.g., MMF approach). MLP emulates the behavior of a CRM or LES and can be run at larger scales (closer to GCM scales) in a variety of regimes and initial conditions. It can be used as a novel, and computationally viable parameterization of moisture processes in a GCM. Currently this approach is extensively applied and developed in many places for building MLPs for moisture physics (e.g., Schneider et al., 2017; Brenowitz & Bretherton, 2018; Gentine et al., 2018; O'Gorman & Dwyer, 2018; Pal, 2020), planetary boundary layer processes (Wang et al. 2019, Wang and Tan 2023), and other processes. This approach produces ML parameterizations of similar or better quality compared to the super parameterization, effectively taking into account subgrid scale effects at a fraction of the computational cost. Also, a combination of simulated and observed data can be used for developments of MLP when observed data are available.

III.4 ML Full Physics

Developing ML emulation of the entire model physics (or diabatic forcing) is a very attractive task. If successful, it could speed up model calculation significantly (especially for high resolution models). On one hand, a lot of challenges are faced when approaching this problem, on the other hand, the full model physics may be better balanced than each particular parameterization separately. It means that the full physics mapping may be smoother and easier for approximation than separate parameterization mappings. Krasnopolsky et al. (2009) discussed problems arising when emulating full physics using NNs. A NN emulation of the entire model physics is analytically differentiable, which will greatly simplify the calculation of an adjoint. Another approach to emulating the columnar physics was proposed in (Krasnopolsky et al. 2014) by emulating super-parameterization (SP) or columnar CRM embedded into the GCM. This approach was successfully applied by other authors (e. g. Wang et al. 2022).

III.5 Fast ML models

Recently, it was shown that it is possible to emulate the dynamics of a simple GCM with a deep neural network (Scher 2018). After being trained on the model, the network can predict the complete model state several time steps ahead. Scher and Messori (2019) assessed how the complexity of the climate model affects the emulating NN's forecast skill, and how dependent the skill was on the length of the provided training period. They showed that using the neural networks to reproduce the climate of general circulation models including a seasonal cycle remained challenging - in contrast to earlier promising results on a model without seasonal cycle. However, further attempts (e.g., Yik et al. 2022) to develop cheap machine learning models for the task of climate model emulation show some progress. Dueben and Bauer (2018) used a toy model for global weather predictions to identify challenges and fundamental design choices for a forecast system based on neural networks. Also, simplified atmospheric and ocean ML models can be developed for use in data assimilation systems for fast first guess calculations (Cheng et al. 2023) and to speed up integration of coupled models (Pawar and Sun 2022). Schultz et al. (2021) considered some evidence that better weather forecasts can be produced by introducing big data mining and deep NNs into the weather prediction workflow. They discuss the question of whether it is possible to completely replace the current numerical weather models and data assimilation systems with deep learning approaches using state-of-the-art machine learning concepts and their applicability to weather data with its pertinent statistical properties. They conclude that it is not inconceivable that numerical weather models may one day become obsolete, but a number of fundamental breakthroughs are needed before this goal comes into reach. In some sense the approach discussed by the authors is a reviving, at the new more sophisticated level, of the statistical weather prediction that existed before the NWP era.

III.6 ML Stochastic Physics

In some cases, the parameterization mapping contains an internal source of stochasticity. It may be due to several reasons: a stochastic process that the mapping describes, a stochastic method (e.g., Monte Carlo methods) implemented in mathematical formulation of the mapping, contribution of subgrid processes, or uncertainties in the data that are used to define the mapping. Such stochastic parameterizations can be emulated using an ensemble of ML/NNs (Krasnopolsky 2013).

ML can be used to create fast stochastic physics. Usually perturbed physics (or parameterization) *P* is created by adding a small random value to a deterministic physics. Using ML, the j^{th} perturbed version of the deterministic model physics, *P*, can be written as,

$$P_j = P_j^{ML} = P + \varepsilon_j \qquad (2)$$

where P_j^{ML} is a ML emulation number *j* of the original model physics, *P*, and ε_j is an emulation error for the ML emulation number *j.* As we have shown in our previous investigations (Krasnopolsky & Fox-Rabinovitz, 2006), ε_j can be controlled and changed significantly by varying the number of hidden neurons in NN so that not only the value but also the statistical properties of ε_j can be controlled. For example, the systematic components of the emulation errors (biases) can be made negligible (therefore, ε_j are purely random in this case). Thus, ε_j can be made the same order of magnitude as the natural uncertainty of the model physics (or of a particular parameterization) due to the unaccounted variability of sub-grid processes. Actually, a single ML emulation (each member of the aforementioned ensemble) can be considered as a stochastic version of the original deterministic parameterization and can be used for creating different ensembles with stochastic physics (Krasnopolsky et al. 2008a).

IV. Machine Learning for Post-Processing

Currently numerical models produce improved weather forecasts and climate projections with better accuracy. However, results produced by the NWP and climate projecting systems still contain errors of different nature. Errors from multiple sources have a detrimental effect on the skill of weather forecasts. One of the sources of errors is associated with the construction of an initial condition for numerical weather forecasting systems. The sensitivity to initial conditions makes errors grow rapidly during the course of the forecasts until they reach a level beyond which the forecasts do not display any useful skill. The boundary-condition errors and the model structural errors are two other important categories of errors that reduce forecast skill. Model structural errors include a missing or poor representation of subgrid dynamical and physical processes and inaccuracies associated with the numerical scheme.

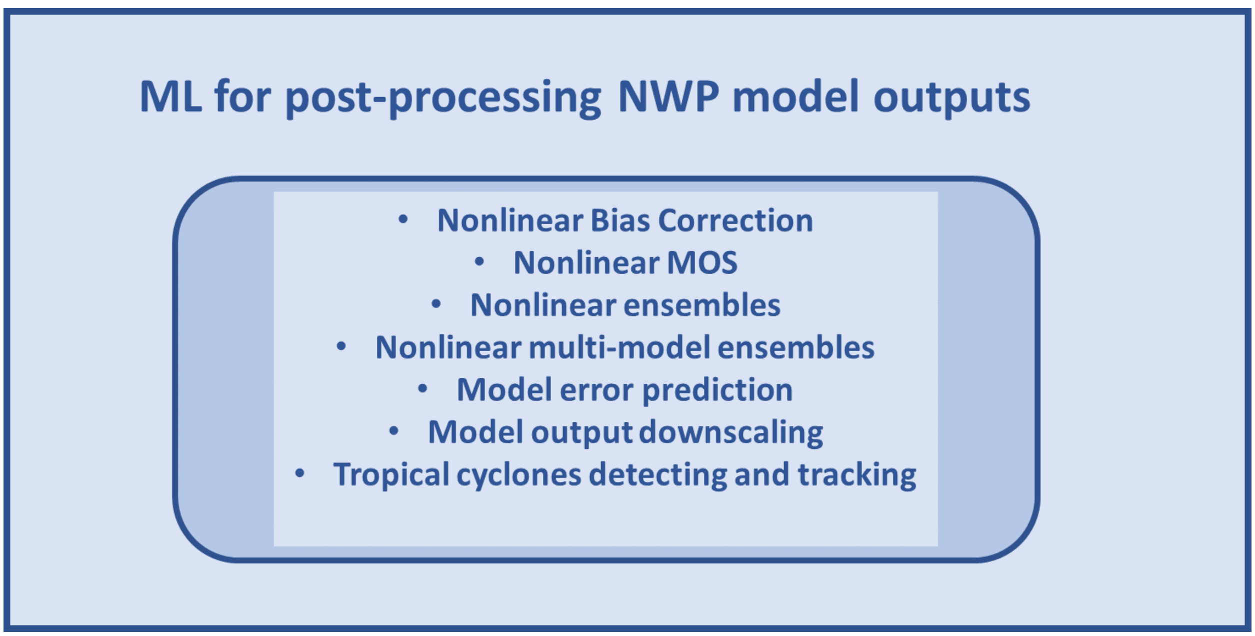

Figure 6. The summary of ML applications for EMC post-processing numerical model outputs that have been developed, are under development, and are considered.

All these NWP model deficiencies induce errors that are rapidly amplified in time due to the chaotic nature of the model dynamics, and in turn affect the forecasts by inducing errors (systematic and random). Thus, the post-processing (PP), correction of errors in model outputs/forecasts, becomes even more important (Vannitsem et al. 2021, Haupt et al, 2021). Figure 6 shows some important ML PP applications.

Statistical PP approaches correct errors in model output by comparing hindcasts to observations. Since the beginning of the era of the NWP forecast, attempts have been made to statistically correct model outputs, given observational data (e.g., Glahn and Lowry 1972). Most current weather forecasting centers rely on statistical methods that have been proven successful. The first approach that was used for statistical PP known as Model Output Statistics (Klein et al. 1959, Glahn and Lowry 1972) was based on multiple multilinear regressions. The U.S. National Weather Service has used these statistical methods to improve systematic model error since 1968 (Carter et al. 1989, Wilks and Hamill 2007). This approach has been also applied to correct errors in ensembles becoming Ensemble Model Output Statistics (EMOS) (Gneiting et al. 2005). These methods demonstrate significant reduction of errors in numerical forecasts (Hemri et al. 2014}. However, these approaches have several significant limitations: (1) they are essentially linear techniques; to account for the nonlinear character of errors (e.g., due to different atmospheric regimes, terrain types, etc.), multiple multilinear regressions are introduced to correct errors in different variables, at different locations, and under different weather conditions, thus, increasing tremendously the number of linear regressions used by the system; (2) they require significant amount of additional information about statistical properties of parameters (Glahn and Lowry 1972).

At the same time, these linear statistical approaches can be viewed as a supervised machine learning task, that is as a direct linear predecessor of nonlinear machine learning (ML) approaches. ML/AI methods, like neural networks (NN) and deep NNs (DNN), which usually are nonlinear and nonparametric, have capabilities to describe complex, multiscale, and nonlinear character of model errors significantly better and more compact and provide more effective corrections.

Initial efforts using machine learning in the context of PP NWP model output have shown promising results (see Haupt et al, 2021 and references there) in both probabilistic and deterministic settings. At ECMWF work (Grönquist et al. 2021) is mainly focused on post-processing ensemble predictions using deep neural networks (DNNs) on precipitation downscaling and tropical detection cyclone and tracing (see also Fig. 2). Also, Bouallègue et al. (2023) used ML technique to correct global 2m temperature and 10m wind speed forecast errors. Rojas-Campos et al. (2023) analyzed the potential of deep learning using probabilistic NN for post-processing ensemble precipitation forecasts at four observation locations. NNs show a higher performance at three of the four stations for estimating the probability of precipitation and at all stations for predicting the hourly precipitation. Benáček et al. (2023) used the tree-based ML techniques, namely, natural gradient boosting, quantile random forests, and distributional regression forests to adjust hourly 2-m temperature ensemble prediction at lead times of 1–10 days. They showed that key components to improving short-term forecasting are additional atmospheric/surface state predictors and the 4-yr training sample size.

At EMC shallow NNs were used to calculate nonlinear multi-model (eight global and regional models) ensembles and to correct 24-hour precipitation forecasts over the ConUS (Krasnopolsky and Lin 2012b). It was shown that, compared with the conservative ensemble (arithmetic mean of ensemble members) and linear regression approach, the ML approach provides slight improvements in gross statistical scores; however, it significantly reduces the number of false alarms and improves forecast of maxima and fronts shape and position. Recently, papers on using ML for multi-model ensemble forecasts of surface air temperatures (Wang et al. 2023) and for probabilistic multi-model ensemble predictions of Indian summer monsoon rainfall (Acharya and Hall 2023) have been published.

A nonlinear ensemble averaging technique using neural networks was applied to NCEP Global Ocean Wave Ensemble forecast System (GWES) data (Campos et al. 2020). Post-processing algorithms are developed based on shallow NNs trained with altimeter data to improve the global forecast skill from nowcast to forecast ranges up to 10 days, including significant wave height (Hs) and wind speed (U10). It is shown that a simple NN model with few neurons can reduce the systematic errors for short-range GWES forecasts, while a NN with more neurons is required to minimize the scatter error at longer forecast ranges. The RMSE of day-10 forecasts from the NN simulations indicated a gain of two days in predictability when compared to the conservative ensemble, using a reasonably simple post-processing model with low computational cost.

In a recent publication (Sebbar et al. 2023) several ML techniques have been compared and used for spatial downscaling of hourly model air temperature over mountainous regions.

A collaborative Google-NOAA study (Agrawal et al. 2023) is focused on investigating the benefits and challenges of using non-linear neural network (NN) based methods to post-process multiple weather features – temperature, moisture, wind, geopotential height, precipitable water – at 30 vertical levels of the NOAA GFS.

V. Summary

Advantages and limitations of ML technique (and NN technique in particular) applications in earth system sciences are discussed in detail in Chapters 2 and 4 of (Krasnopolsky, 2013). Here we will mention only major advantages and limitations that are relevant for development of ML applications for EMC DAS, model physics, and PP.

For DAS, ML can provide fast forward models for direct assimilation of satellite radiances, fast observation operators for instantaneous 3D assimilation of surface observations, fast environmental models for assimilating chemical and biological observations, fast adjoints for 4Dvar DAS, and fast hybrid and ML models for calculating first guess. For model physics ML can provide fast emulating ML parameterizations, fast and improved ML parameterization of physics, fast ML emulations of entire atmospheric physics, and fast ML stochastic physics. For PP ML can enable developments of nonlinear bias corrections, nonlinear ensemble averaging, etc. (see Fig. 6).

Some limitations of ML techniques should be mentioned. ML tools are not very good at far extrapolation. Nonlinear extrapolation is an ill-posed problem that requires regularization to provide meaningful results. The development of ML applications depends significantly on our ability to generate/collect a representative training set to avoid using NNs for extrapolation far beyond the domain covered by the training set. Because of the high dimensionality, *n*, of the input domain that is often several hundred or more, it is rather difficult to cover the entire domain. At least 2^n points are required to cover the entire domain. Especially difficult is to cover the "far corners" associated with rare events, even when we use simulated data for MLP training. A significant help here can be the ML ensemble approach. Using an ensemble of ML tools can help to regularize the extrapolation and deliver MLPs that are more stable when the inputs approach "far corners" or cross the boundary of the training domain.

Another related problem arises when ML emulations are developed for a non-stationary environment or climate system that changes with time. This means that the domain configuration for a climate simulation may evolve when using, for example, a future climate change scenario. In such situations, the ML emulation may be forced to extrapolate beyond its generalization ability leading to errors in MLP outputs and resulting in simulation errors in the corresponding model. Here compound parameterization (Krasnopolsky et al. 2008c) and dynamical adjustment as well as using the ML ensemble approach could be helpful.

It is noteworthy that ML still requires human expertise to succeed. The development of ML applications for NWP models and DAS is not a standard ML problem. While ML applications can, in principle, be used as a black box, the development of ML physics will require domain knowledge about the Earth system physics. Close collaborations between computer scientists, atmospheric physicists, and modelers will be essential even if petabytes of training data and GPU supercomputers are available. A deep understanding of how to use physical knowledge of the Earth system to improve the development of ML architectures and ML training and how to preserve conservation properties and take into account other physical constraints will be required. There are a lot of decisions that must be made in the process of developing ML applications that cannot be made automatically.

Like any other statistical models (e. g. MOS) ML applications have to be maintained and periodically updated.

VI. Appendix 1

The table below contains my very rough estimates of the time needed at EMC to develop ML applications mentioned in this article. In my evaluations, I assumed that at least one full-time ML and full-time field experts are available for the development.

ML Application	Methodology developed at EMC	Prototype developed at EMC	Estimated time for development
Forward model and Jacobian	yes	yes	<1 year
Observation operator	yes	yes	1 year
Fast models and adjoints for DAS	no	no	a few years
Data QC and pre-processing	no	no	1 year
Fast radiation	yes	yes	1 year
Fast microphysics	yes	no	1 to several years
New ML parameterizations	yes	no	several years
Full ML physics	yes	no	1 to several years
Fast ML models	no	no	several years
Stochastic physics	yes	no	1 - 3 years
Bias corrections	yes	yes	1 year
Nonlinear MOS	no	no	several years
Nonlinear ensembles	yes	yes	a few years
Nonlinear multi-model ensembles	yes	yes	a few years
Prediction of model errors	no	no	1 to several years
Model output downscaling	no	no	several years
Tropical cyclones detecting and tracking	no	no	several years

References:

Acharya N. and Hall K. (2023). A machine learning approach for probabilistic multi-model ensemble predictions of Indian summer monsoon rainfall. *MAUSAM*, 74, 421-428; https://mausamjournal.imd.gov.in/index.php/MAUSAM/article/view/5997

Agrawal S., R. Carver, C. Gazen, E. Maddy, V. Krasnopolsky, C. Bromberg, Z. Ontiveros, T. Russell, J. Hickey, and S. Boukabara (2023). A Machine Learning Outlook: Post-processing of Global Medium-range Forecasts. *arXiv:2303.16301,* https://doi.org/10.48550/arXiv.2303.16301

Bauer P., A. J. Geer, Ph. Lopez, D. Salmond (2010); Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Quat. J. Royal Met. Soc.*, Vol. 136, 1868-1885; https://doi.org/10.1002/qj.659

Bauer, P., Thorpe, A. & Brunet, G (2015). The quiet revolution of numerical weather prediction. *Nature* 525, 47–55. https://doi.org/10.1038/nature14956

Bauer, P., Dueben, P.D., Hoefler, T. et al. (2021). The digital revolution of Earth-system science. *Nat Comput Sc*i 1, 104–113. https://doi.org/10.1038/s43588-021-00023-0

Belochitski, A., Binev, P., DeVore, R., Fox-Rabinovitz, M., Krasnopolsky, V., & Lamby, P. (2011). Tree Approximation of the Long Wave Radiation Parameterization in the NCAR CAM Global Climate Model, *Journal of Computational and Applied Mathematics*, 236, 447-460. doi: 10.1016/j.cam.2011.07.013

Belochitski, A., and Krasnopolsky, V. (2021). Robustness of neural network emulations of radiative transfer parameterizations in a state-of-the-art general circulation model. *Geosci. Model Dev*., 14, 7425–7437. https://doi.org/10.5194/gmd-14-7425-2021

Bouallègue, Z. B., Cooper, F., Chantry, M., Düben P., Bechtold P., and Sandu, I. (2023). Statistical Modelling of 2m Temperature and 10m Wind Speed Forecast Errors. *Monthly Weather Review*, DOI: https://doi.org/10.1175/MWR-D-22-0107.1

Boukabara, S.-A., V. Krasnopolsky, J. Q. Stewart, E. S. Maddy, N. Shahroudi, and R. N. Hoffman, 2019: Leveraging modern artificial intelligence for remote sensing and NWP: Benefits and challenges. *Bull. Am. Meteorol. Soc.,* 100(12), ES473–ES491.

Brenowitz, N. D., & Bretherton, C. S. (2018). Prognostic Validation of a Neural Network Unified Physics Parameterization, *Geophysical Research Letters*, 45, 6289-6298. https://doi.org/10.1029/2018GL078510

Brenowitz, N. D. et al., 2022. Emulating Fast Processes in Climate Models. arXiv:2211.10774 [physics.ao-ph], https://doi.org/10.48550/arXiv.2211.10774

Campos R.M., V. Krasnopolsky, J-H. Alves, S. G. Penny (2020), Improving NCEP's global-scale wave ensemble averages using neural networks, *Ocean Modelling,* 149, https://doi.org/10.1016/j.ocemod.2020.101617

Carter GM, Dallavalle JP, Glahn HR. (1989). Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Weather Forecast*. 4, 401–412. (doi:10.1175/1520-0434(1989)004<0401:SFBOTN>2.0.CO;2)

Cheng, S., Chen, J., Anastasiou, C. et al. (2023). Generalized Latent Assimilation in Heterogeneous Reduced Spaces with Machine Learning Surrogate Models. *J Sci Comput,* 94, 11 https://doi.org/10.1007/s10915-022-02059-4

Chevallier, F., Chéruy, F., Scott, N. A., Chedin, A. (1998). A Neural Network Approach for a Fast and Accurate Computation of Longwave Radiative Budget. *Journal of Applied Meteorology*, 37, 1385-1397

Chevallier, F., Morcrette, J.-J., Chéruy, F., Scott, N. A. (2000). Use of a neural-network-based longwave radiative transfer scheme in the ECMWF atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, 126, 761-776

Christensen, H. and Zanna, L. (2022). Parametrization in Weather and Climate Models. https://doi.org/10.1093/acrefore/9780190228620.013.826

Dong R., H. Leng, J. Zhao, J. Song and S. Liang (2022). A Framework for Four-Dimensional Variational Data Assimilation Based on Machine Learning. *Entropy*, 24, 264; https://doi.org/10.3390/e24020264

Düeben, P., P. Bauer, S. Adams, (2021a). Deep Learning to Improve Weather Predictions. In "Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences" Eds:G. Camps-Valls, D. Tuia, X. X. Zhu, M. Reichstein, John Wiley & Sons Ltd., DOI:10.1002/9781119646181

Düben P., et al., (2021b). Machine learning at ECMWF: A roadmap for the next 10 years. *ECMWF Technical Memorandum* 878, https://www.ecmwf.int/sites/default/files/elibrary/2021/19877-machine-learning-ecmwf-roadmap-next-10-years.pdf

Dueben, P. D. and P. Bauer, 2018: Challenges and design choices for global weather and climate models based on machine learning. *Geosci. Mod. Dev*., 11(10), 3999–4009.

Geer A.J., P. Bauer, Ph. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Quat. J. Royal Met. Soc.*, Vol. 136, 1886-1905; https://doi.org/10.1002/qj.681

Geer, A. J., and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Quart. J. Roy. Meteor. Soc.*, 144, 1191–1217, https://doi.org/10.1002/qj.3202

Geer A.J., 2021: Learning Earth System Models from Observations: Machine Learning or Data Assimilation? *Phil. Trans. R. Soc.* A 379: 20200089; https://doi.org/10.1098/rsta.2020.0089

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., Yacalis G. (2018). Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, 45, 5742–5751. https://doi.org/10.1029/ 2018GL078202

Glahn HR, Lowry DA. 1972. The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.* **11**, 1203–1211. (doi:10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2)

Gneiting T, Raftery AE, Westveld AH, Goldman T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Wea. Rev*. 133, 1098–1118. (doi:10.1175/MWR2904.1)

Grönquist P., et al. (2021). Deep learning for post-processing ensemble weather forecasts. *Phil. Trans. R. Soc*. A 379: 20200092. https://doi.org/10.1098/rsta.2020.0092

Hatfield, S., M. Chantry, P. Dueben, P. Lopez, A. Geer, T. Palmer (2021). Building Tangent-Linear and Adjoint Models for Data Assimilation With Neural Networks. *Journal of Advances in Modeling Earth Systems*, v. 13, https://doi.org/10.1029/2021MS002521

Haupt S.E., A. Pasini, and C. Marzban (Eds.), (2008): *Artificial Intelligence in Environmental Sciences* (pp. 191-205), Springer; ISBN 978-1-4020-9119-3

Haupt S. E., et al. (2021). Towards implementing artificial intelligence post-processing in weather and climate: proposed actions from the Oxford 2019 workshop. *Phil. Trans. R. Soc*. A 379: 20200091.https://doi.org/10.1098/rsta.2020.0091

Hemri S, Scheuerer M, Pappenberger F, Bogner K, Haiden T. (2014). Trends in the predictive performance of raw ensemble weather forecasts. *Geophys. Res. Lett.* **41**, 9197–9205. (doi:10.1002/2014GL062472)

Hsieh W.W. (2009): *Machine learning methods in the environmental sciences.* Cambridge University Press, Cambridge; ISBN: 9780521791922

Jensen A., Weeks C., Xu M., Landolt S., Korolev A., Wolde M., and DiVito S. (2023). The prediction of supercooled large drops by a microphysics and a machine-learning model for the ICICLE field campaign. *Weather and Forecasting*, DOI 10.1175/WAF-D-22-0105.1

Klein, W. H., B. M. Lewis, and I. Enger (1959): Objective prediction of five-day mean temperatures during winter. *J. Meteor.,* 16, 672–682, https://doi.org/10.1175/1520-0469(1959)016<0672:OPOFDM>2.0.CO;2

Khain, A., Rosenfeld, D., Pokrovsky, A. (2000). Aerosol impact on the dynamics and microphysics of deep convective clouds, *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, 131, 2639-2663

Khan, H.N., Hounshell, D.A. & Fuchs, E.R.H. (2018). Science and research policy at the end of Moore's law. *Nat Electron* **1**, 14–21. https://doi.org/10.1038/s41928-017-0005-9

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Chalikov, D. V. (2005). New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Long Wave Radiation in a Climate Model. *Monthly Weather Review*, 133, 1370-1383; https://doi.org/10.1175/MWR2923.1

Krasnopolsky, V. M. & Fox-Rabinovitz, M. S. (2006). Complex Hybrid Models Combining Deterministic and Machine Learning Components for Numerical Climate Modeling and Weather Prediction. *Neural Networks*, 19, 122-134; DOI: 10.1016/j.neunet.2006.01.002

Krasnopolsky V.M., 2007: "Reducing Uncertainties in Neural Network Jacobians and Improving Accuracy of Neural Network Emulations with NN Ensemble Approaches", *Neural Networks*, 20, pp. 454–461; DOI: 10.1016/j.neunet.2007.04.008

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Belochitski, A. (2008a). *Using neural network emulations of model physics in numerical model ensembles.* Paper presented at 2008 IEEE International Joint Conference on Neural Networks, Hong Kong, China (2008 IEEE World Congress on Computational Intelligence). https://doi.org/10.1109/IJCNN.2008.4633998.

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Belochitski, A. (2008b). Decadal climate simulations using accurate and fast neural network emulation of full, long- and short wave, radiation. *Monthly Weather Review*, 136, 3683-3695. doi:10.1175/2008MWR2385.1

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Rasch, P. J., and Wang M. (2014). Fast NN Emulation of The Super-Parameterization In The Multi-Scale Modeling Framework. Office note (NOAA, National Centers for Environmental Prediction (U.S.)); 476. https://repository.library.noaa.gov/view/noaa/6961

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Tolman, H. L., Belochitski, A. (2008c). Neural network approach for robust and fast calculation of physical processes in numerical environmental models: Compound parameterization with a quality control of larger errors. *Neural Networks*, 21, 535-543

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Hou, Y. T., Lord, S. J., & Belochitski, A. A. (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

Krasnopolsky, V., Belochitski, A., Hou, Y.-T., Lord, S., & Yang, F. (2012a). Accurate and Fast Neural Network Emulations of Long and Short Wave Radiation for the NCEP Global Forecast System Model, NCEP Office Note, 471. http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on471.pdf.

Krasnopolsky V. M. & Lin, Y. (2012b). A Neural Network Nonlinear Multimodel Ensemble to Improve Precipitation Forecasts over Continental US, *Advances in Meteorology*, 2012, Article ID 649450, 11 pages. http://dx.doi.org/10.1155/2012/649450

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). Dordrecht, Heidelberg, New York, London: Springer. DOI 10.1007/978-94-007-6073-8

Krasnopolsky, V., Middlecoff, J., Beck, J, Geresdi, I., & Toth, Z. (2017). A Neural Network Emulator for Microphysics Schemes. Paper presented at AMS annual meeting. https://ams.confex.com/ams/97Annual/webprogram/Paper310969.html

Krasnopolsky V., S. Nadiga, A. Mehra, 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,* Volume 2018, Article ID 7057363, 10 pp., https://doi.org/10.1155/2018/7057363

Krasnopolsky, V. and Belochitski, A. (2022). Using Machine Learning for Model Physics: An Overview. https://arxiv.org/ftp/arxiv/papers/2002/2002.00416.pdf

Lang M., P. J. Van Leeuwen, Ph. Browne, (2016): A systematic method of parameterisation estimation using data assimilation. *Tellus A: Dynamic Meteorology and Oceanography,* Volume 68. https://doi.org/10.3402/tellusa.v68.29012

Lagerquist, R., Turner, D., Ebert-Uphoff, I., Stewart, J., & Hagerty, V. (2021). Using deep learning to emulate and accelerate a radiative-transfer model. *Journal of Atmospheric and Oceanic Technology*, https://doi.org/10.1175/JTECH-D-21-0007.1

Maulik R, V. Rao, J. Wang, G. Mengaldo, E. Constantinescu, B. Lusch, P. Balaprakash, I. Foster, and R. Kotamarthi (2022). Efficient high-dimensional variational data assimilation with machine-learned reduced-order models. *Geosci. Mod. Dev*., 15, 3433–3445; https://doi.org/10.5194/gmd-15-3433-2022

O'Gorman, P. A., & Dwyer, J. G. (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.* https://doi.org/10.1029/2018MS001351

Pal, A., Mahajan, S., & Norman, M. R. (2019). Using deep neural networks as cost-effective surrogate models for Super-Parameterized E3SM radiative transfer. *Geophysical Research Letters*, 46, 6069–6079. https://doi.org/10. 1029/2018GL081646

Pal, A. (2020). Deep Learning Emulation of Subgrid-Scale Processes in Turbulent Shear Flows, *Geophysical Research Letters*, 47, e2020GL087005. https://doi.org/10.1029/2020GL087005

Pawar, S. and O. San (2022). Equation-Free Surrogate Modeling of Geophysical Flows at the Intersection of Machine Learning and Data Assimilation. *Journal of Advances in Modeling Earth Systems*, 14, Issue 11, https://doi.org/10.1029/2022MS003170

Rasp, S., Pritchard, M. S., Gentine, P. (2018). Deep learning to represent subgrid processes in climate models, *PNAS Latest Articles*. https://doi.org/10.1073/pnas.1810286115

Roh, S., & Song, H.-J. (2020). Evaluation of neural network emulations for radiation parameterization in cloud resolving model. *Geophysical Research Letters,* 47, e2020GL089444, https://doi.org/10.1029/2020GL089444

Rojas-Campos R., M. Wittenbrink, P. Nieters, E. J. Schaffernicht, J. D. Keller, and G. Pipa (2023). Postprocessing of NWP Precipitation Forecasts Using Deep Learning. *Weather and Forecasting*, 38, 487-497; DOI: https://doi.org/10.1175/WAF-D-21-0207.1

Scher, S. (2018). Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. *Geophysical Research Letters*, 45. https://doi.org/10.1029/2018GL080704

Scher, S., Messori G. (2019). Weather and climate forecasting with neural networks: using GCMs with different complexity as study-ground. *Geosci. Model Dev. Discuss*., https://doi.org/10.5194/gmd-2019-53

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

Schultz MG, Betancourt C, Gong B, Kleinert F, Langguth M, Leufen LH, Mozaffari A, Stadtler S. (2021). Can deep learning beat numerical weather prediction? *Phil. Trans. R. Soc. A* 379: 20200097. https://doi.org/10.1098/rsta.2020.0097

Sebbar, B.-e., Khabba, S., Merlin, O., Simonneaux, V., Hachimi, C.E,; Kharrou M.H.; Chehbouni, A. (2023). Machine-Learning-Based Downscaling of Hourly ERA5-Land Air Temperature over Mountainous Regions. *Atmosphere*, 14, 610. https://doi.org/10.3390/atmos14040610

Thompson, G., Field, P. R., Rasmussen, R. M., Hall, W. D. (2008). Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Monthly Weather Review*, 138, 5095-5115

Ukkonen, P., Pincus, R., Hogan, R. J., Nielsen, K. P., & Kaas, E (2020).: Accelerating radiation computations for dynamical models with targeted machine learning and code optimization. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002226, https://doi.org/10.1029/2020MS002226

Vannitsem et al. (2021). Statistical Postprocessing for Weather Forecasts: Review, Challenges, and Avenues in a Big Data World. *BAMS*, E681; https://doi.org/10.1175/BAMS-D-19-0308.1

Vapnik, V. N. (1995). The nature of statistical learning theory (p. 189). New York: Springer

Vapnik, V. N., & Kotz, S. (2006). Estimation of Dependences Based on Empirical Data (Information Science and Statistics). New York: Springer

Vapnik, V. N., Complete Statistical Theory of Learning, *Automation and Remote Control*, 2019, Vol. 80, No. 11, pp. 1949–1975.

Veerman, M. A., Pincus, R., Stoffer, R., van Leeuwen, C. M., Podareanu, D. & van Heerwaarden, C. C. (2021). Predicting atmospheric optical properties for radiative transfer computations using neural networks. *Phil. Trans. R. Soc*. A 379: 20200095, https://doi.org/10.1098/rsta.2020.0095

Wang, J., P. Balaprakash, and R. Kotamarthi, (2019). Fast domain-aware neural network emulation of a planetary boundary layer parameterization in a numerical weather forecast model. *Geosci. Model Dev.*, 12, 4261–4274, https://doi.org/10.5194/gmd-12-4261-2019

Wang, Le-Yi, Zhe-Min Tan (2023). Deep Learning Parameterization of the Tropical Cyclone Boundary Layer. *Journal of Advances in Modeling Earth Systems*, 15, Issue 1. https://doi.org/10.1029/2022MS003034

Wang, T., Zhang, Y., Zhi, X., Ji, Y. (2023). Multi-Model Ensemble Forecasts of Surface Air Temperatures in Henan Province Based on Machine Learning. *Atmosphere*, 14, 520. https://doi.org/10.3390/atmos14030520

Wang, X., Han, Y., Xue, W., Yang, G., and Zhang, G. J. (2022). Stable climate simulations using a realistic general circulation model with neural network parameterizations for atmospheric moist physics and radiation processes. *Geosci. Model Dev*., 15, 3923–3940, 2022 https://doi.org/10.5194/gmd-15-3923-2022

Wilks DS, Hamill TM. (2007). Comparison of ensemble-MOS methods using GFS reforecasts. *Mon. Wea. Rev*. 135, 2379–2390. doi:10.1175/MWR3402.1

Yik, W., Sam J. Silva, S.J., Geiss, A., Watson-Parris, D. (2022). Exploring Randomly Wired Neural Networks for Climate Model Emulation. arXiv:2212.03369v2 [physics.ao-ph]; https://arxiv.org/pdf/2212.03369.pdf