Member Survey of ML/AI Uses for Model Development

Fanglin Yang (NCEP)
Tim Graham (Met Office)

37th Session of the Working Group on Numerical Experimentation (WGNE)
8-10 November 2022, NCAR, Boulder, CO, USA

With contributions from

CMC -- Ron McTaggart-Cowan

DWD -- Günther Zangl

ECMWF -- Matthew Chantry

Hydrometcenter of Russia -- Michael Tsyrulnikov

INPE -- Ariane Frassoni

JMA -- Masashi Ujiie

Météo France -- Romain Roehrig

Met Office -- Tim Graham

NCEP -- Fanglin Yang

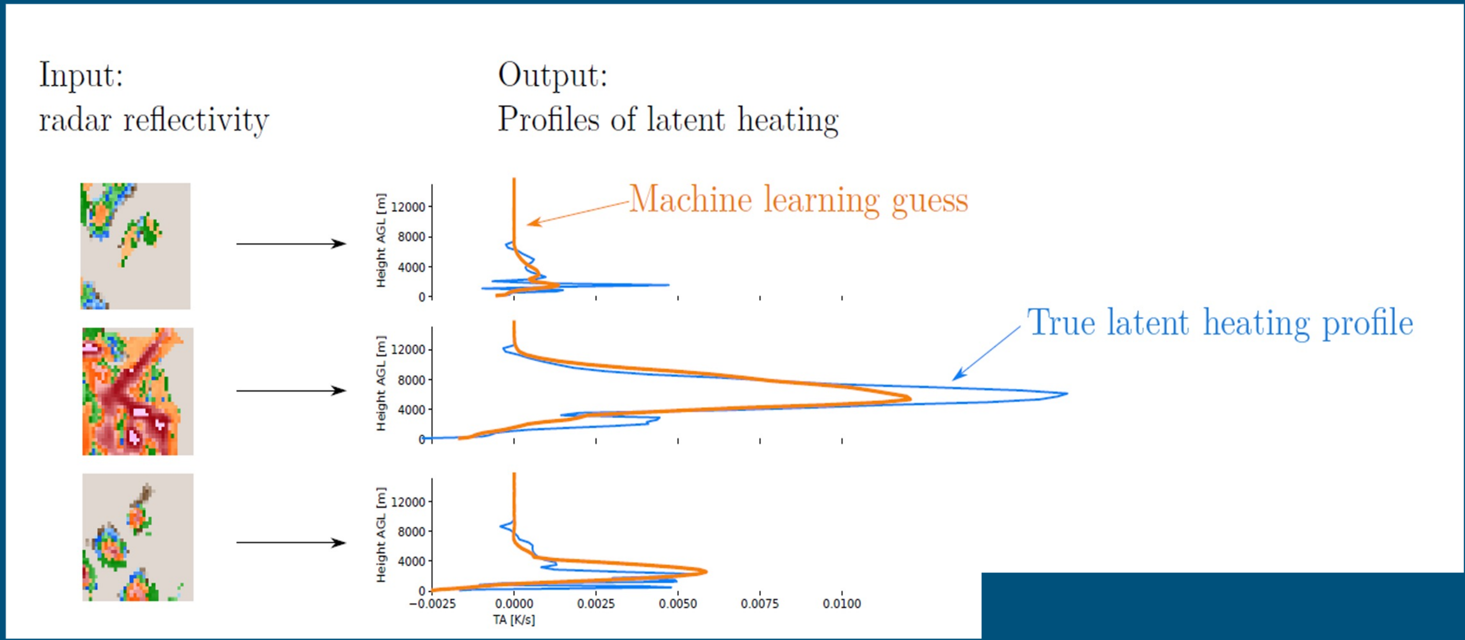
NRL -- Carolyn Reynolds

Canadian Meteorological Centre

Ron McTaggart-Cowan, Howard Barker,
Jason Cole, Stephane Gagnon and Dominik Jacques

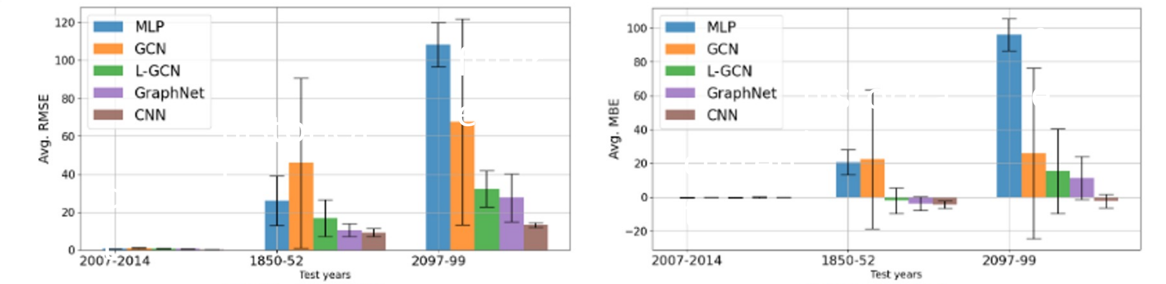
CMC: AI/ML in Data Assimilation

- Assimilating radar data in the 2.5km national deterministic prediction system yields positive results (Jacques et al. 2022)
- Latent heat nudging involves applying heating tendency profiles within the model in regions of observed precipitation



- We hope to replace the current heuristic-based heating profiles with ML-based profiles, trained from a large sample of modelled heating rates
- Promising results are obtained from a deep (50-layer) Residual Network (ResNet) model

CMC: AI/ML in Model Physics



(a) Test RMSE (b) Test MBE

Figure 3: Performance as a function of the main test set year (Fig. 2) and OOD test set (Fig. 3) for our different baseline models. Metrics are in W/m^2 and shown as the average over the vertical and over the up- and down-welling flux errors. The leftmost x-tick for the OOD plots corresponds to the average metric of the main test years 2007-2014, for which the metrics are shown on a yearly basis in Fig. 2. Generalization to pre-industrial and future pristine-sky conditions is a particularly challenging task because of the changes in gas concentrations More structured models like a CNN or GraphNet perform significantly better than an MLP.

- A project to replace the radiative transfer scheme with an ML-based emulator is under way in the context of the Canadian Earth System Model (CanESM).
- A benchmark dataset for training and evaluation of radiative transfer emulators (climART) has been developed and shared with the community (Cachay et al. 2022)

- The climART benchmark is used to test multiple ML techniques for applicability under current, historical and future climates under pristine and clear-sky conditions (to be extended to all-sky)
- “Structured” models perform better under changing backgrounds than multilayer perceptron (MLP)
- Speed-ups to >10x are observed, although fair comparison on GPUs is difficult

CMC: AI/ML in Post-Processing

- A new model output statistics (MOS)-type post-processing system is under development at CMC
- The “modularity” design requirement will allow the current regression-based prediction technique to be replaced with AI/ML in the future
- Planning to meet with Waves to Weather (T3) researchers (Rasp and Lerch 2018) to discuss deep neural network-based post-processing:
  - Applicability of wind gust algorithm developed with KNMI to larger Canadian region
  - Utility of AI for ensemble post-processing (higher order moments)
  - Use of explainable AI to develop forecaster confidence in AI-based guidance

DWD

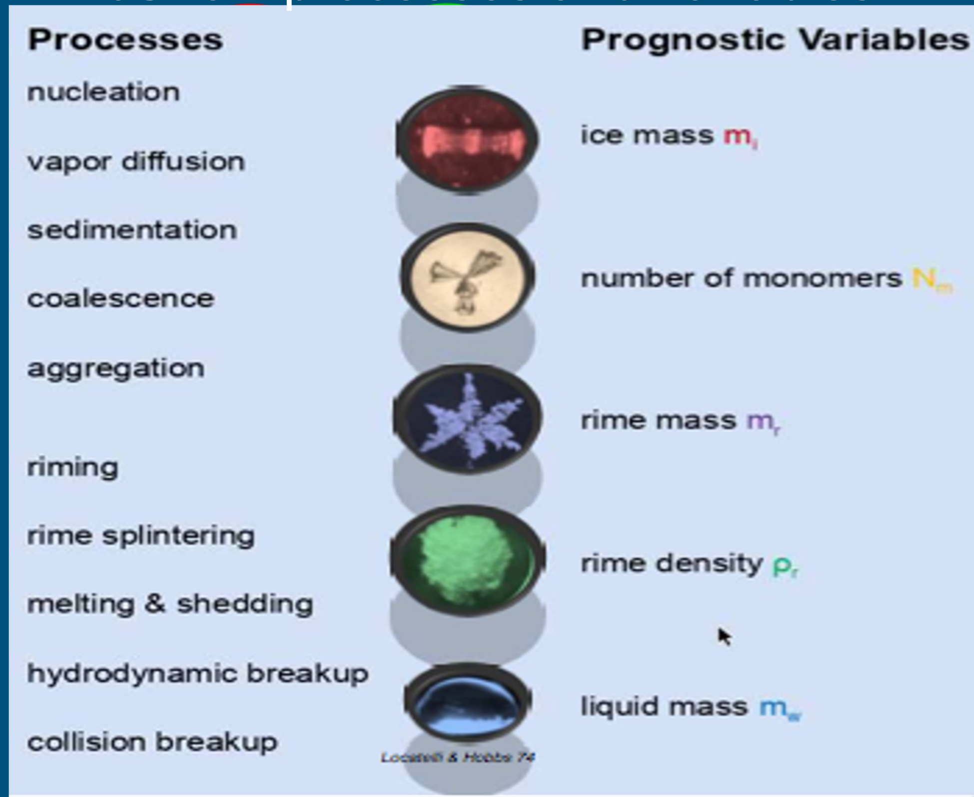
Günther Zangl

DWD: AI/ML in Model Physics

Lagrangian particle model

- To generate (microphysical states) training data using the Lagrangian particle model ***McSnow*** that explicitly resolves ice processes (Brdar and Seifert 2018)
- Each McSnow particle has several variables that describe its current microphysical state.
- Needs at least 1000 particles per grid point, better 10000 to reduce Monte-Carlo noise.
- These are expensive simulations that are even today hardly feasible in 3d.

McSnow processes and variables



ML for ice microphysics: Choice of bulk equations

- We try to built an ODE system with 6 particle classes:
  - *ice monomers, snow aggregates, rimed ice, rimed snow graupel and rain (and cloud droplets)*.
- All classes have mass and number densities, rimed classes (including graupel) have in addition rime mass, rime volume and liquid mass.
- Hence, for rimed particle classes with have rime fraction, rime density and melted fraction as bulk properties (P3-like scheme).
- This makes a total of 23 variables and more than 100 process rates.

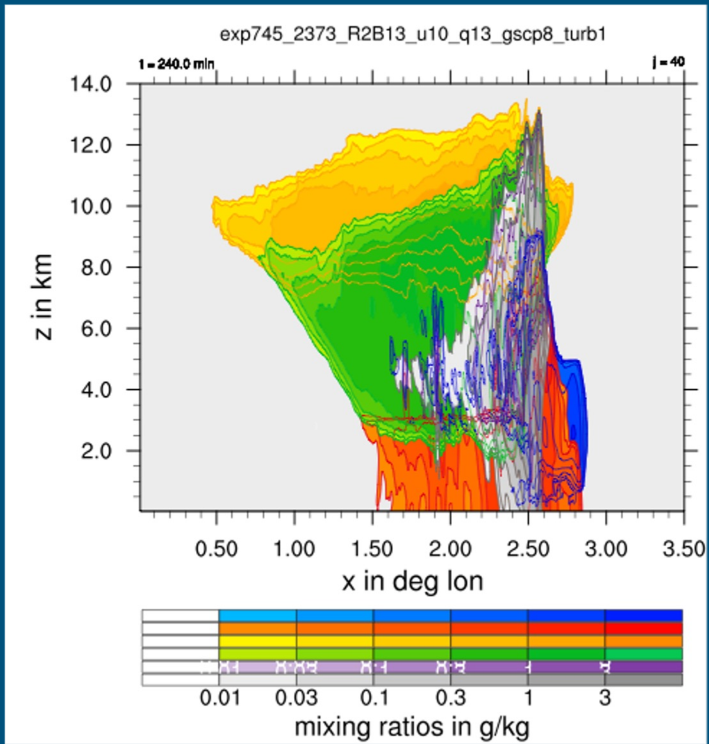
- **Can we „learn“ all those bulk process rates from McSnow output and come up with an ODE system that works reasonably well?**

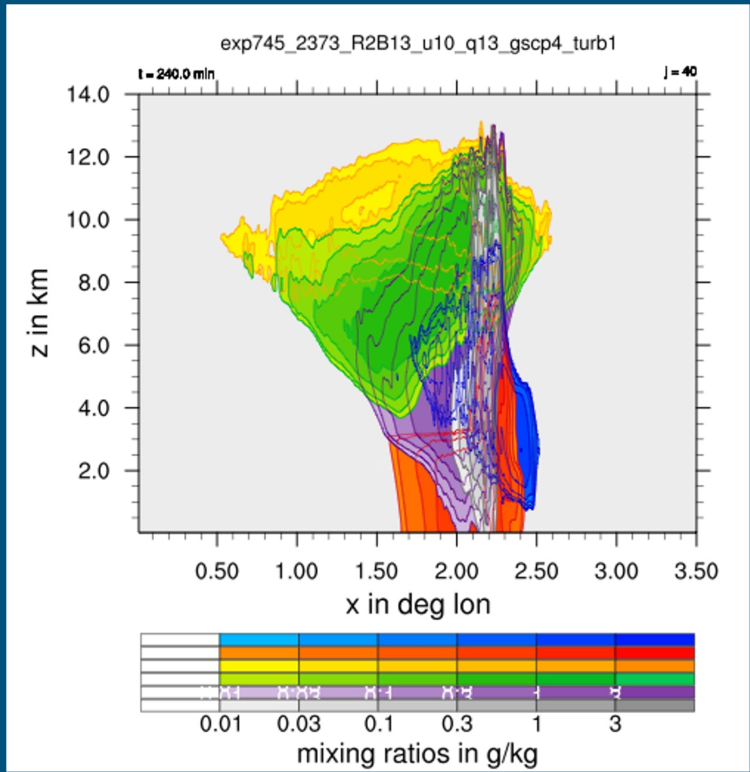
Training data for bulk ice microphysics:

- We use an idealized box model falling through a prescribed atmosphere.
- The idealized box model allows many simulations. This is preferred here over a few 3d simulations.
- Training data is generated by brute-force hypercube sampling of initial condition and atmospheric profile resulting more than 10.000 simulation.
- This can cover a large range of parameters.
- It proved to be necessary to include updraft parcels in the training data to better represent processes within the convective core.
- Maybe another advantage of the idealized box model approach is that it does not contain an imprint of the current climate, in contrast to more realistic simulations.
- Note: This is similar in spirit to the so-called „bin-emulating schemes“ used in some cloud-resolving models (e.g. RAMS).

Simulation of an idealized squall line with ICON

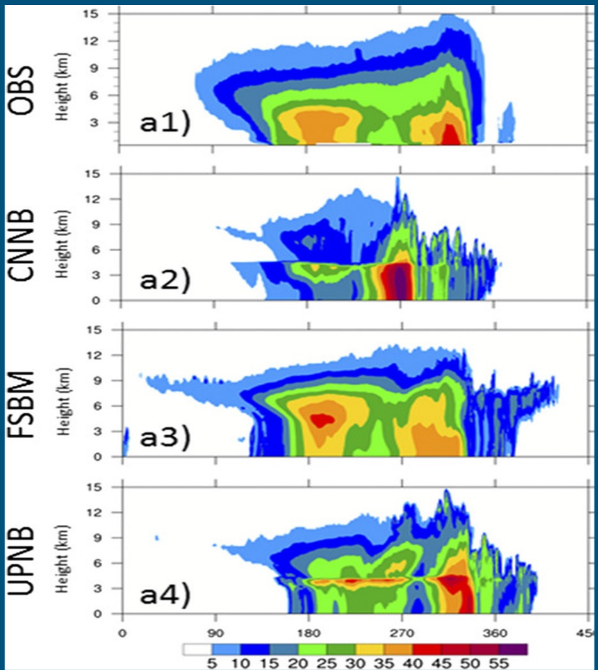
Vertical cross section of hydrometeors:
ML-based vs classic two-moment





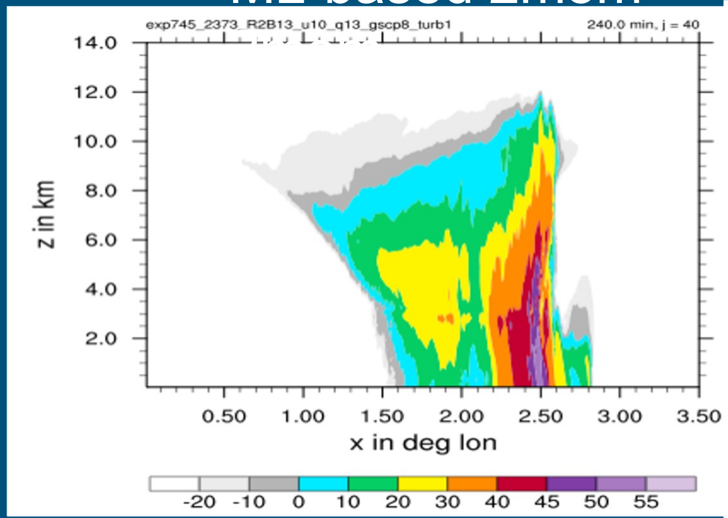
Simulation of an idealized squall line with ICON

- Radar reflectivity (Rayleigh approximation)



Xue et al. (2017): Observations and three bin microphysics schemes

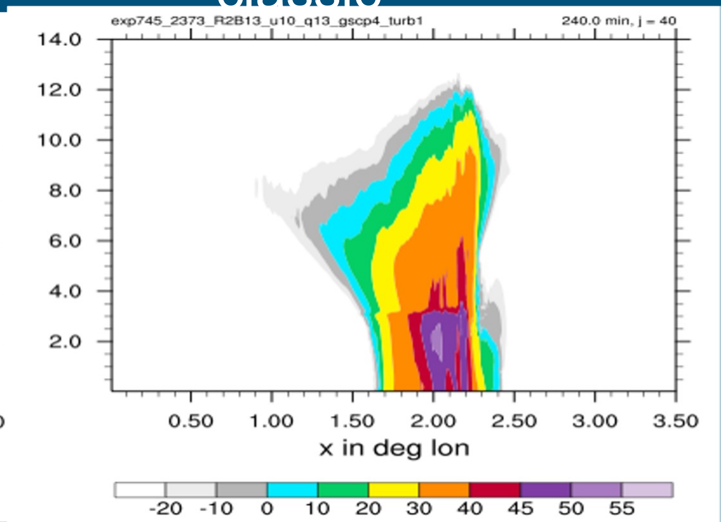
ML-based 2mom



dBZ

Extended stratiform region with secondary maximum

classic



dBZ

Lack of stratiform region (common for bulk schemes)

ECMWF

Peter Dueben, Matthew Chantry, Frederic Vitart, Zied Ben Bouallegue, Massimo Bonavita, Patrick Laloyaux, Antonino Bonanni

ECMWF: AI/ML in Data Assimilation

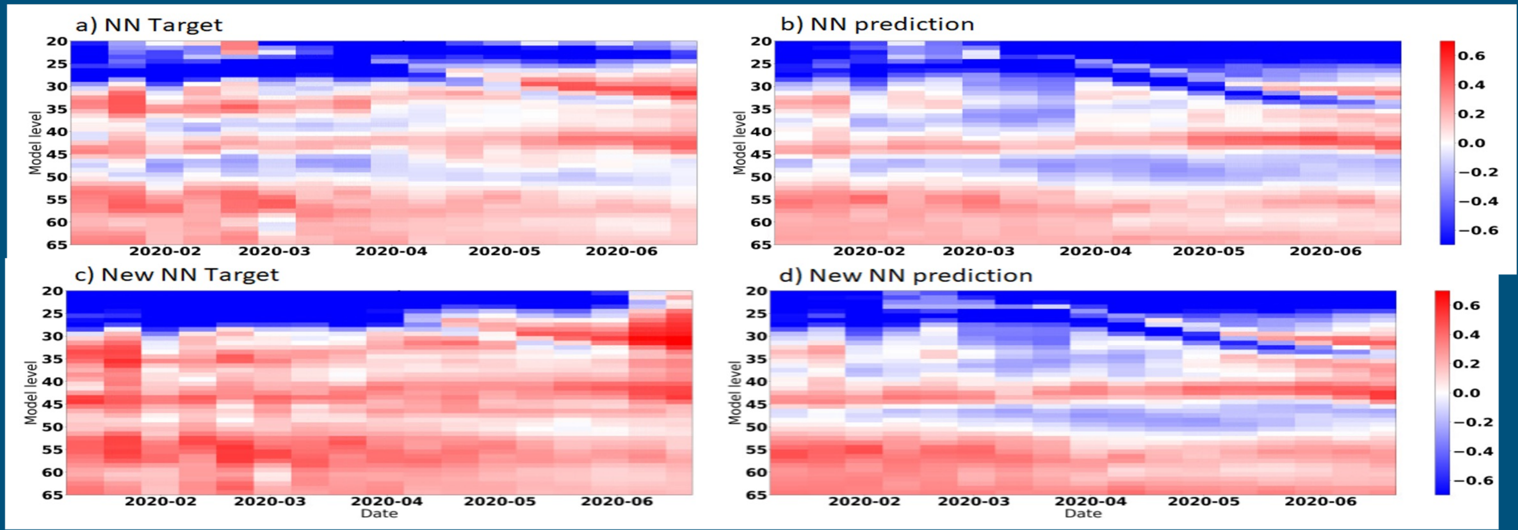
Learn how to combine operational models and machine learning

- During data-assimilation the model trajectory is “synchronised” with observations
- It is possible to learn model error when comparing the model with (trustworthy) observations

Approach: **Learn model error from a direct comparison of the model trajectory and observations**

Benefit: Correct for model error and understand model deficiencies

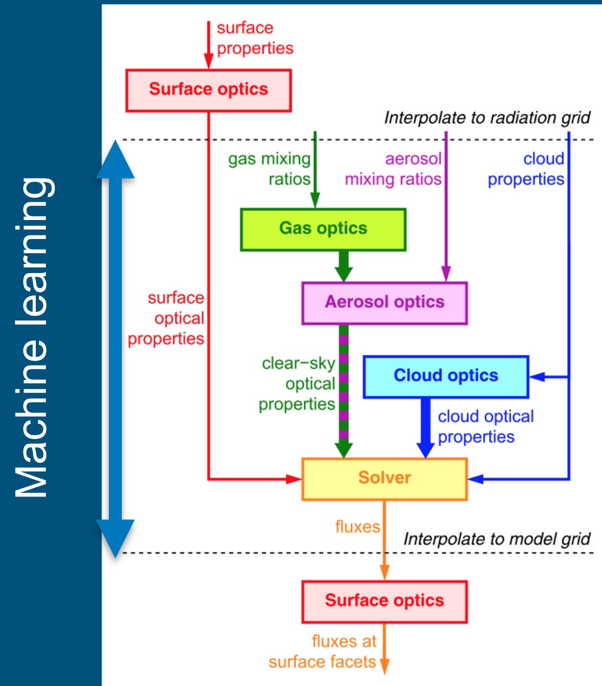
Question: **What happens when the model is upgraded and the error pattern change?**



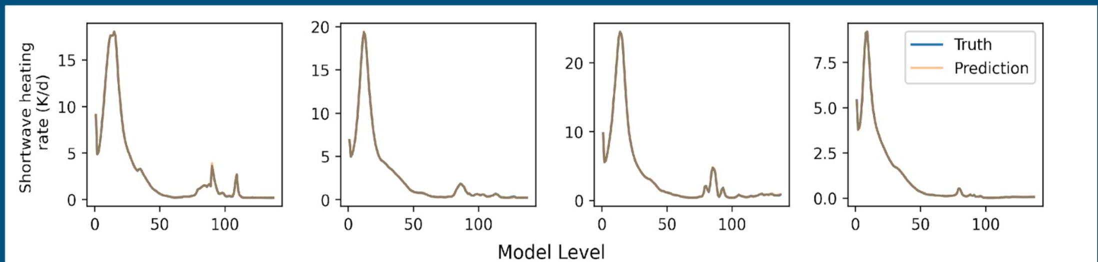
NVIDIA

Laloyaux, Kurth, Dueben, Hall JAMES 2022

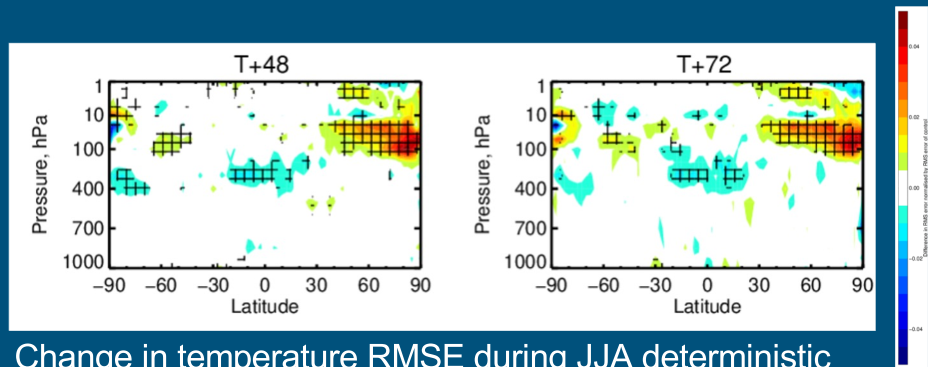
ECMWF: AI/ML in Model Physics



- Learning to **emulate radiative transfer scheme in the IFS.**
  - Reduce computational cost.
  - Leverage GPU nodes on HPC.
- Utilise bespoke physics-informed architecture.
- Offline errors small, O(10^{-2} K/d) & O(1 Wm^{-2})



- Coupled forecasts neutral below 100hPa.



Change in temperature RMSE during JJA deterministic forecasts at TCo399 versus existing radiation scheme.

- **Ongoing work to utilise GPUs in coupled forecast.**

ECWMF: AI/ML in Post-Processing

Ensemble post-processing using transformers

Collaboration with Microsoft

Utilise state-of-the-art machine learning architecture to coherently postprocess 2m-temperature IFS forecasts.

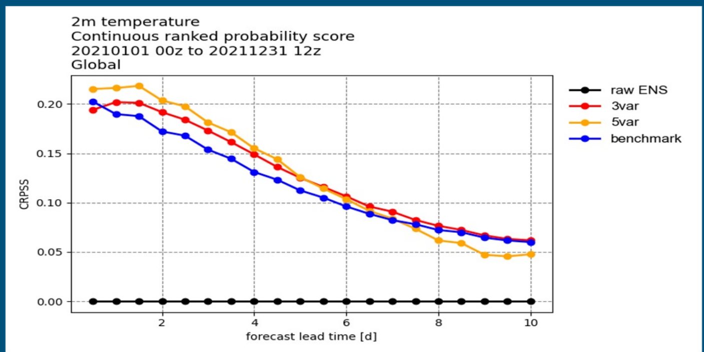
Solution ensures ensemble member fidelity while making state dependent adjustments to ensemble bias and spread.

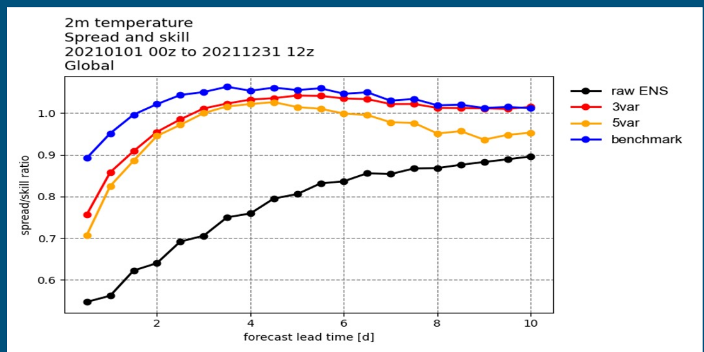
Ensemble-size agnostic – training from hindcast (11-members), predictions with ENS (51-members).

Trained to optimise the CRPS.

Improves CRPS compared with member-by-member benchmark (or raw ensemble).

Improves spread-skill parity.





ECWMF: AI/ML Application Competition

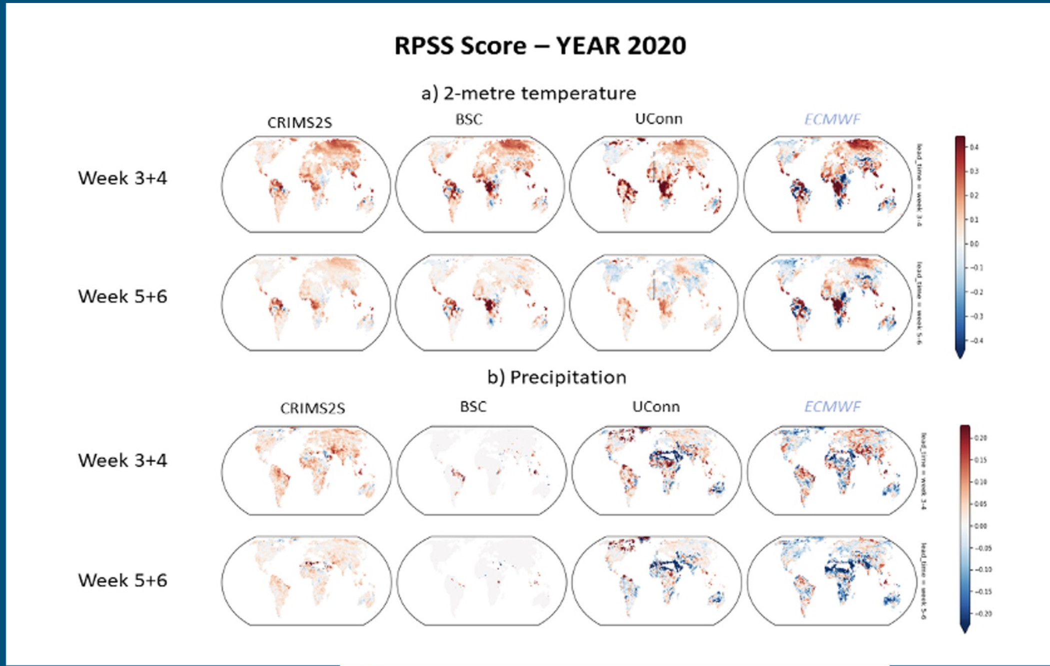
Challenge: Provide forecasts of near surface temperature and precipitation for weeks 3+4 and 5+6 more skilful than ECMWF operational forecasts for the year 2020.

- Hosted by Swiss Data Science Center at ETH Zürich, **with ECMWF support through the new European Weather Cloud for data access to S2S forecasts, the use the CliMetLab software and the provision of virtual machines to some participants from developing countries**.
- Timeline: June-November 2021
- **All codes and forecasts are open source** to foster community learning on AI/ML methods for S2S
- 30k Swiss Francs prize from WMO

ECMWF

Outcome of the competition:

- 49 registered teams
- 5 teams succeeded in providing better forecasts than the Benchmark (ECMWF S2S operational forecasts)
- Top 3 teams got rewarded a prize.



WORLD METEOROLOGICAL ORGANIZATION S2S WWRP WCRP SDSC ECMWF

Hydrometcenter of Russia

Michael Tsyrulnikov

Hydrometcenter: AI/ML in Data Assimilation

LSEF: A Locally Stationary Ensemble Filter

In the **LSEF analysis,** the background-error covariance matrix is represented as $$B = W \cdot W^T$$

The sparse **W matrix** is estimated directly from the ensemble in a four-stage procedure:

(1) A multi-scale band-pass filter is applied to ensemble perturbations.

(2) In each pass band, sample variances (in the univariate 2D case) or sample covariance matrices (in the multivariate 3D case) are computed.

(3) From sample band variances (which can be viewed as aggregated spatial spectra), so-called **local spectra** are restored at each grid point. **It is here where ML comes into play.** Without ML, the new technique did outperform EnKF, 3D-Var, and EnVar with 2D synthetic truth and observations, but the improvement was not large. With ML, the improvement became substantial.

(4) From the local spectra, a *spatially varying convolution kernel* is computed, from which W is built.

ML details: vanilla neural network (NN), 2 hidden layers (with 36 and 72 neurons), Adam optimizer. The learning sample was of the same size for **LSEF, 3D-Var** (to **learn the mean B matrix**), and for **EnKF** (to learn **the localization length**).
Importantly, in our experiments, the NN needs to be *trained only once*, and then it can be applied to different ensemble sizes, different spatial background-error covariances etc.

Ref.: *M Tsyrulnikov and A Sotskiy. The ensemble Kalman filter regularized with non-parametric non-stationary spatial convolutions (in preparation).*

Experimental setup: synthetic truth on the sphere, synthetic randomly located observations with observation noise st.dev. equal to FG error st.dev.

Static analyses. Verification against truth. 90% bootstrap tolerance intervals (stripes). RMSE are normalized (see the y-axis label)

6 schemes are compared:

Mean-B: 3D-Var with time-mean B matrix

EnKF-B: stochastic EnKF with tuned localization

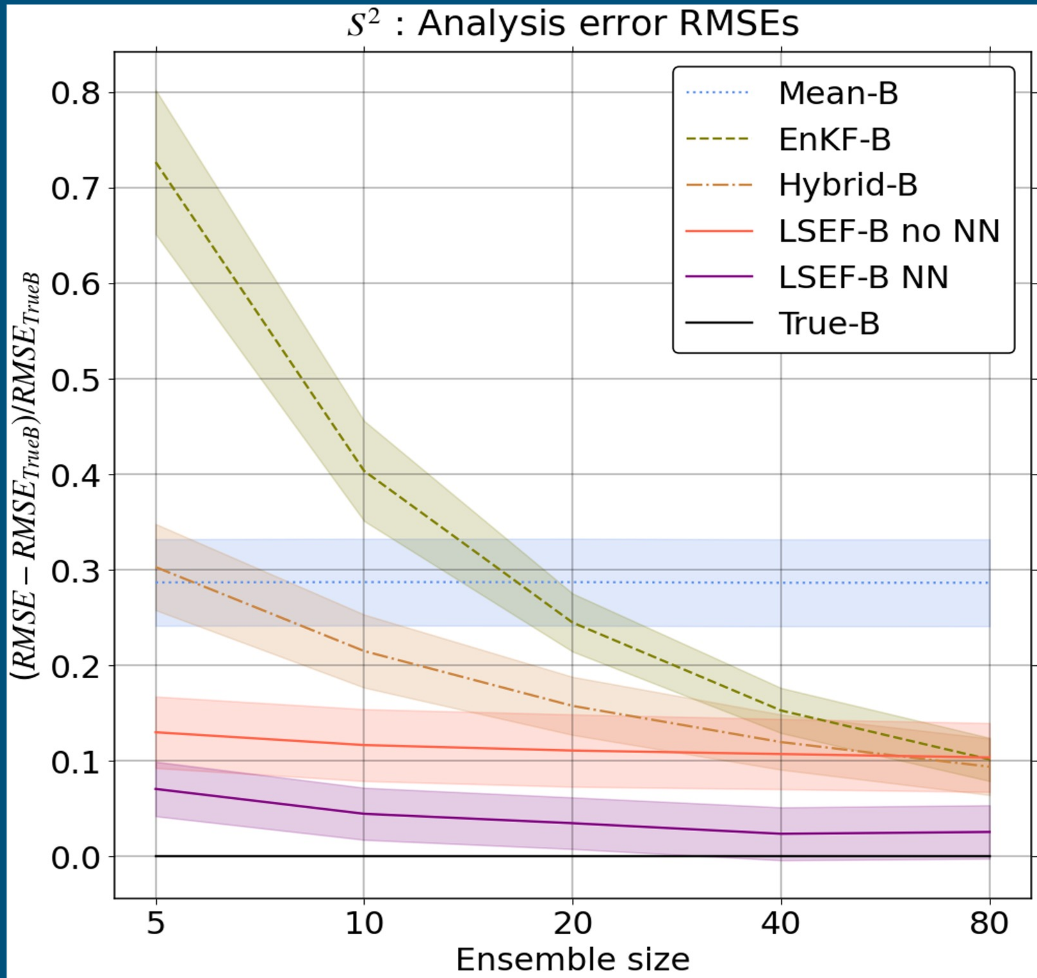
Hybrid-B: EnVar with 50% B-mean and 50% EnKF

LSEF-B no NN: the new analysis without ML

LSEF-B NN: the new analysis with ML (2-layer NN)

True-B: the optimal analysis (which has access to the true B matrix)

Conclusion: **ML (vanilla 2-layer NN) leads to a major improvement in the analysis accuracy**



INPE

Brazilian National Institute for Space Research

Ariane Frassoni

INPE: AI/ML in Data Assimilation

https://doi.org/10.5194/gmd-2022-50
Preprint. Discussion started: 9 September 2022
© Author(s) 2022. CC BY 4.0 License.

Geoscientific Model Development
Open Access
Discussions

Neural networks for data assimilation of surface and upper-air data in Rio de Janeiro

Vinícius Albuquerque de Almeida[1], Haroldo Fraga de Campos Velho[2], Gutemberg Borges França[1], and Nelson Francisco Favilla Ebecken[3]

[1]Laboratory for Applied Meteorology - Federal University of Rio de Janeiro
[2]National Institute for Space Research
[3]Civil Engineering/COPPE - Federal University of Rio de Janeiro

Correspondence: Vinícius Albuquerque de Almeida (vinicius@lma.ufrj.br)

Method	Mean Error (K)	RMSE (K)	SD (K)
3D-Var	0.61	1.06	0.50
NN-TensorFlow	0.48	0.99	0.59
NN-Weka	0.79	1.12	0.44

Library	Single cycle	Total
3D-Var	00:00:08.00	00:02:27.00
TensorFlow	00:00:00.04	00:00:01.21
Weka	00:00:00.21	00:00:05.88

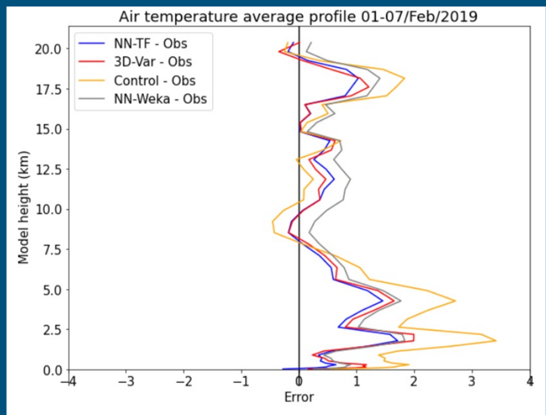
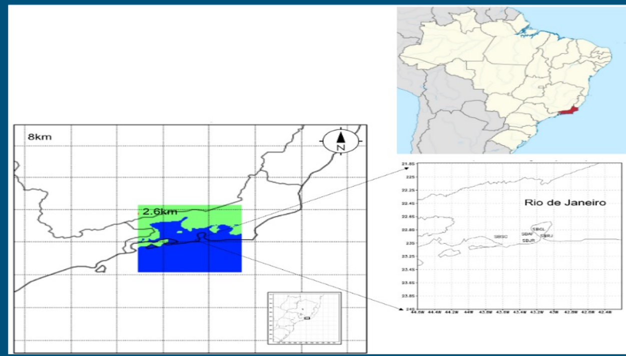


Figure 5. Mean temperature profile difference of 6-hour forecast (from a field without data assimilation), 6-hour forecast with 3D-Var and neural networks - trained in TensorFlow and Weka - in relation to the observed profiles between Feb 1st and 8th, 2019 at SBGL.



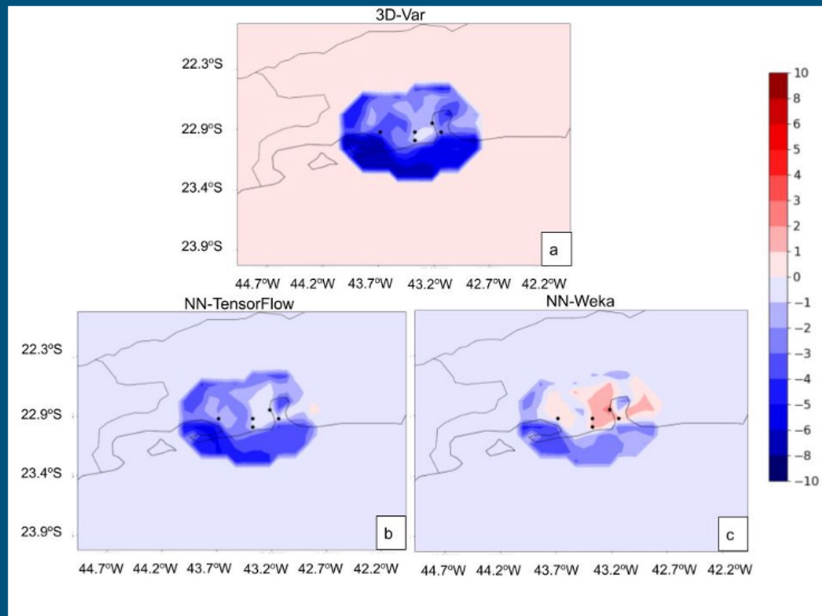


Figure 4. Temperature error map of (a) 3D-Var and neural networks - trained in (b) TensorFlow and (c) Weka - applied to 6-hour forecas fields in relation to the observed map on Feb 1st, 2019 12 UTC.

INPE: AI/ML in Post-Processing

Springer Link

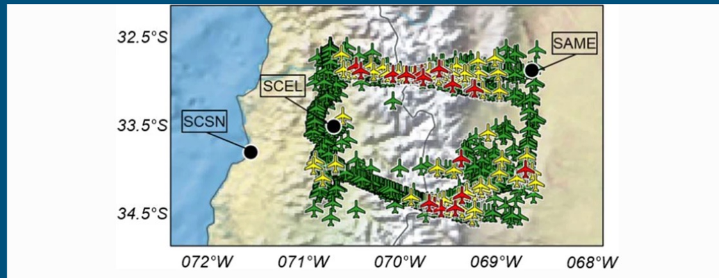
Published: 27 May 2022

In-Flight Turbulence Forecast Model Based on Machine Learning for the Santiago (Chile)–Mendoza (Argentina) Air Route

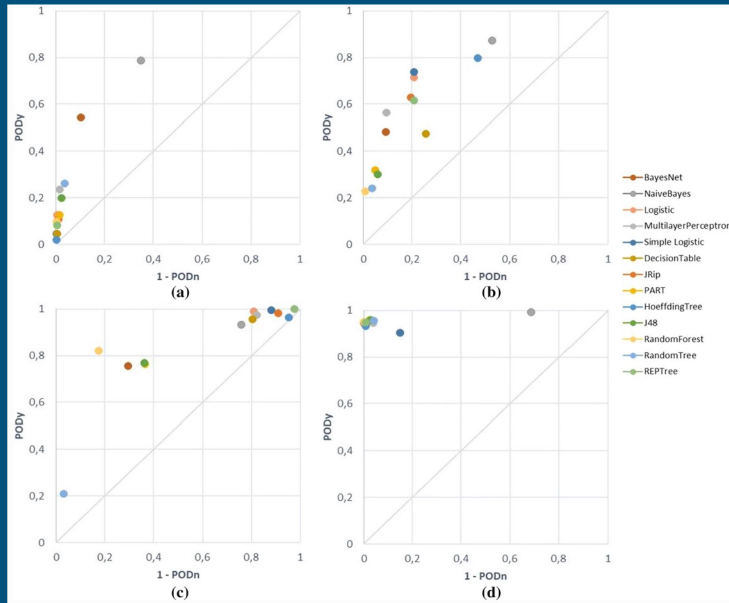
Filipe Menegardo-Souza, Gutemberg Borges França, Wallace Figueiredo Menezes & Vinícius Albuquerque de Almeida

Pure and Applied Geophysics **179**, 2591–2608 (2022) | Cite this article

75 Accesses | Metrics



Spatial distribution of the 1300 VRTG records above the FL100, from March 2018 to December 2019 in the study area, indicated according to their intensities. VRTG Class 1 (green), Class 2 (yellow), Class 3 (red). In black circles the airports of Santo Domingo (SCSN), Santiago (Arturo Merino Benítez International Airport—SCEL) and Mendoza (Governor Francisco Gabrielli International Airport—SAME)



ROC curve with PODy versus 1-PODn values of the training and testing results of the ML algorithms: **a** unbalanced original data, **b** ClassBalancer 50/50, **c** CostMatrix 380FN, **d** Random Replication "YES"—mode 1

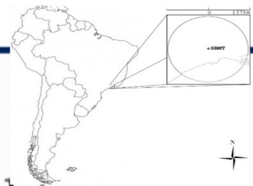
	0Z	3Z	6Z	9Z	12Z	15Z	18Z	21Z
OBSERVED								
Index 5								
Unbalance Original Data (BayesNet)								
ClassBalancer 96/04 (RandomForest)								
CostMatrix 380FN (RandomForest)								
Random Removal "NO" (RandomForest)								
Random Replication "YES" - mode 1 (RandomForest)								
Random Replication "YES" - mode 2 (RandomForest)								
Random Replication "YES" - mode 3 (BayesNet)								
Random Replication "YES" - mode 3 (RandomForest)								

Results of optimal forecast MOG turbulence models for September 28th, 2018 compared to observed data. In gray absence of turbulence.

https://link.springer.com/article/10.1007/s00024-022-03053-5/tables/10

INPE: AI/ML in Post-Processing

SpringerLink



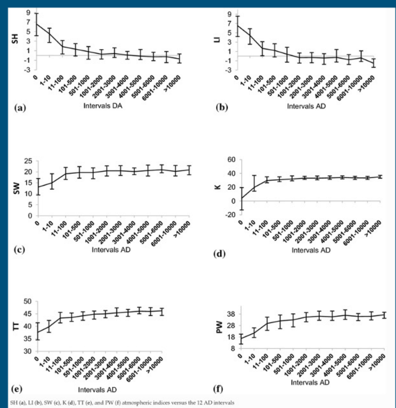
Published: 30 June 2022

Severe Convective Weather Forecast Using Machine Learning Models

Jimmy Nogueira de Castro, Gutemberg Borges França, Vinícius Albuquerque de Almeida & Valdonel Manoel de Almeida

Pure and Applied Geophysics 179, 2945–2955 (2022) | Cite this article

Result	AD interval per event	Model	POD	1-FAR	BIAS	Kappa	F-measure
1	0	RandomForest	0.91	0.95	0.92	0.74	0.88
2	> 10	RandomForest	0.94	0.88	0.91	0.83	0.84
3	> 100	RandomForest	0.83	0.85	0.97	0.74	0.84
4	> 200	BayesNet	0.74	0.91	0.82	0.71	0.82
5	> 300	BayesNet	0.76	0.80	0.96	0.68	0.78
6	> 400	BayesNet	0.68	0.92	0.74	0.67	0.78
7	> 500	RandomForest	0.75	0.67	1.12	0.62	0.71
8	> 600	BayesNet	0.67	0.77	0.88	0.61	0.72
9	> 700	LMT	0.60	0.91	0.66	0.62	0.73
10	> 800	AdaBoostM1	0.63	0.77	0.82	0.60	0.69
11	> 900	AdaBoostM1	0.58	0.83	0.70	0.58	0.69
12	> 1000	Multilayer Perceptron	0.70	0.56	1.25	0.55	0.63
13	> 2000	Multilayer Perceptron	0.66	0.71	0.94	0.60	0.68
14	> 3000	JRip	0.56	0.49	1.15	0.45	0.52
15	> 4000	BayesNet	0.26	0.84	0.31	0.31	0.40
16	> 5000	BayesNet	0.21	0.83	0.26	0.27	0.34



In this study, prediction models based on machine learning were created with the goal of detecting severe convective events in the TMA-SP. The main findings are

Atmospheric discharges occur more frequently in the summer and are associated with UHI

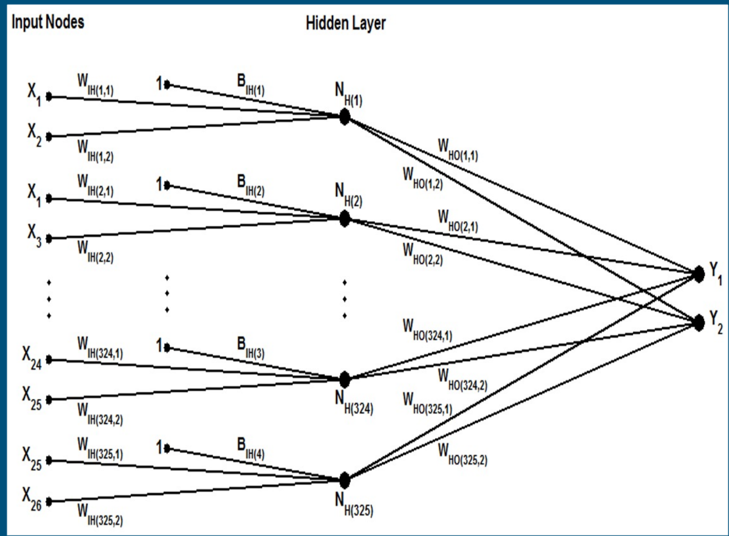
ML models predict and classify the severity (up to 5 h in advance),

Hindcast show models do not currently have a satisfactory performance in predicting convective events with frontal origin

INPE: AI/ML in Post-Processing

Convective systems forecast

Use of trained neural networks with data from analysis and model prediction to predict the occurrence of strong convective activity. Tests with eta model 20 km, 2 classes (strong, non-strong), 26 attributes



(2013-2015) use of neural networks trained with data from analysis and model prediction to predict the occurrence of convective activity tests with eta model 20 km, nsca/sca classes, 48 hs prediction, averages and maximum/minimum of 20 runs… best result below!

		Predicted NSCA	Predicted SCA	Predicted NSCA	Predicted SCA
Actual	NSCA training	506	44	498/520	30/52
	SCA training	25	250	20/39	236/255
	NSCA test	44	2	42/45	1/4
	SCA test	2	21	0/4	19/23

Training **POD/FAR**
0.909/0.149
Tests **POD/FAR**
0.913/0.087

Courtesy: Stephan Stephany

JMA

Masashi Ujiie

JMA: Overview of AI/ML Activities

- Familiarization with machine learning for model developers (extra slides)
  - Use a multiscale Lorenz96 model as a toy system to emulate subgrid parameterization with neural networks (NN)
- Emulate a non-orographic gravity wave (NGW) scheme by NN
  - Recognize importance of domain scientist's knowledge
    - **Momentum fluxes as predictors rather than tendencies of U and V winds for momentum conservation in the emulated scheme.**
    - Choice of loss functions by considering the targeted region
  - A "Learning by Python and Prediction by Fortran" approach makes implementation of ML into NWP models easy (extra slides)
    - Python: learn by ML package (e.g. PyTorch) and write weight coefficients to external files.
    - Fortran: load the files and calculate NN as matrix-matrix multiplication

JMA: AI/ML in Model Physics

U wind tendency from NGW and momentum conservation

$$\left(\frac{\partial U}{\partial t}\right)_{\mathrm{NGW}} = -\frac{1}{\rho}\underline{\frac{\partial F_U}{\partial z}}$$

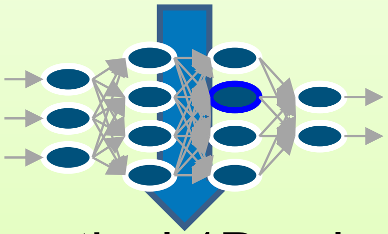
Convergence of momentum flux

$$\int \rho \left(\frac{\partial U}{\partial t}\right)_{\mathrm{NGW}} dz = 0$$

$$(\because F_U(z=0) = F_U(z=z_{\mathrm{top}}) = 0)$$

Input (vertical 1D column)

$$U,\ V,\ T,\ P,\ F_{\mathrm{launch\_level}}$$



2 hidden layers with 1000～2000 units/layer

Output (vertical 1D column)

$$\left(\frac{\partial U}{\partial t}\right)_{\mathrm{NGW}},\ \left(\frac{\partial V}{\partial t}\right)_{\mathrm{NGW}}$$

or

$$F_U,\ F_V$$

Loss functions

(a) MSE (Mean Square Error)

$$\mathrm{LOSS}^{[s]} = \frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M}\left(p_{nm}^{[s]} - y_{nm}\right)^2$$

predictor reference

(b) Δp (*pressure thickness*) weighted MSE

$$\mathrm{LOSS}^{[s]} = \frac{1}{N}\sum_{n=1}^{N}\frac{\sum_{m=1}^{M}\left(p_{nm}^{[s]} - y_{nm}\right)^2 \Delta p_m}{\sum_{m=1}^{M}\Delta p_m}$$

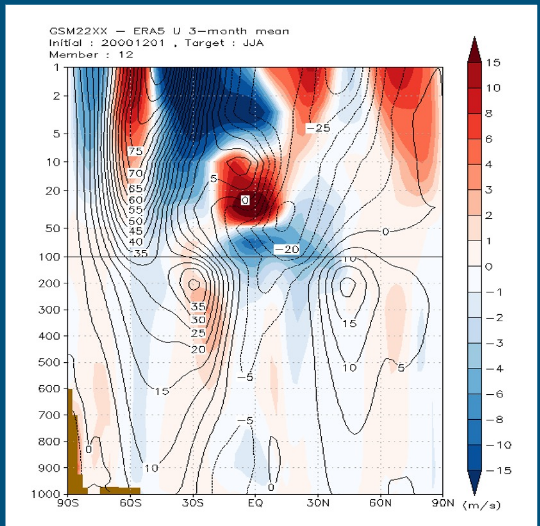
Larger weight over upper troposphere and stratosphere

MATSUKAWA Chihiro

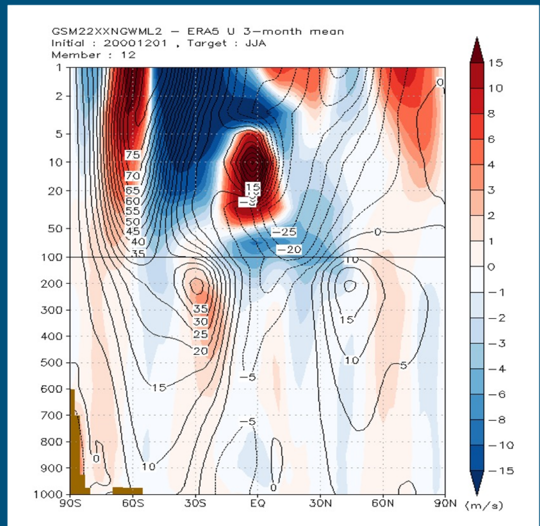
JMA: AI/ML in Model Physics

JJA zonal mean U-wind (contours) and its bias (colors) [m/s] against ERA-5 from TL159L128 1-year simulation

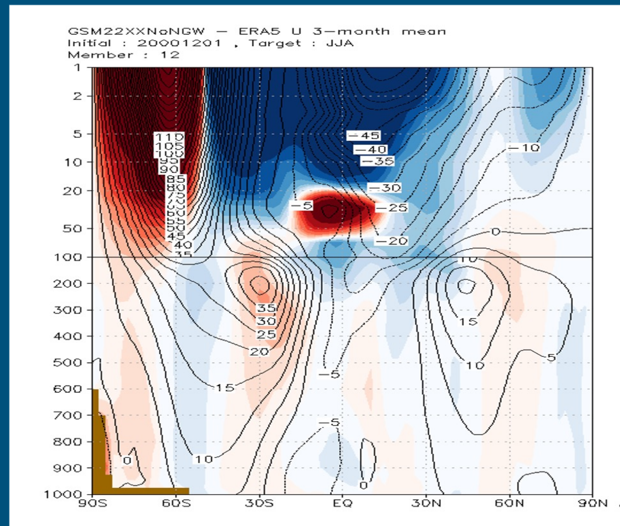
Scinocca(2003) NGW
(reference)

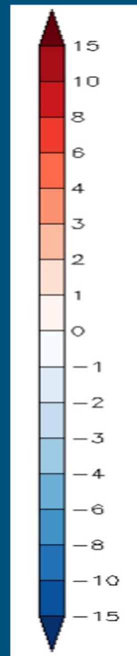


Emulated
(tendencies)



No NGW





The emulated NGW scheme represents the characteristics of the true NGW scheme qualitatively.

Met France

Romain Roehrig

Météo-France: **Overview of AI/ML Activities**

- ***Calibration of ensemble prediction forecasts:*** AI-based post-processing of meteorological parameters (e.g., wind gusts)

- ***Innovative products*** based on kilometer-scale NWP system
  i. Detection of weather objects (e.g., bow echo, TCs)
  ii. Synthesis of EPS forecasts by coherent meteorological scenario
  iii. Detection of mesoscale convective system in nowcasting system.

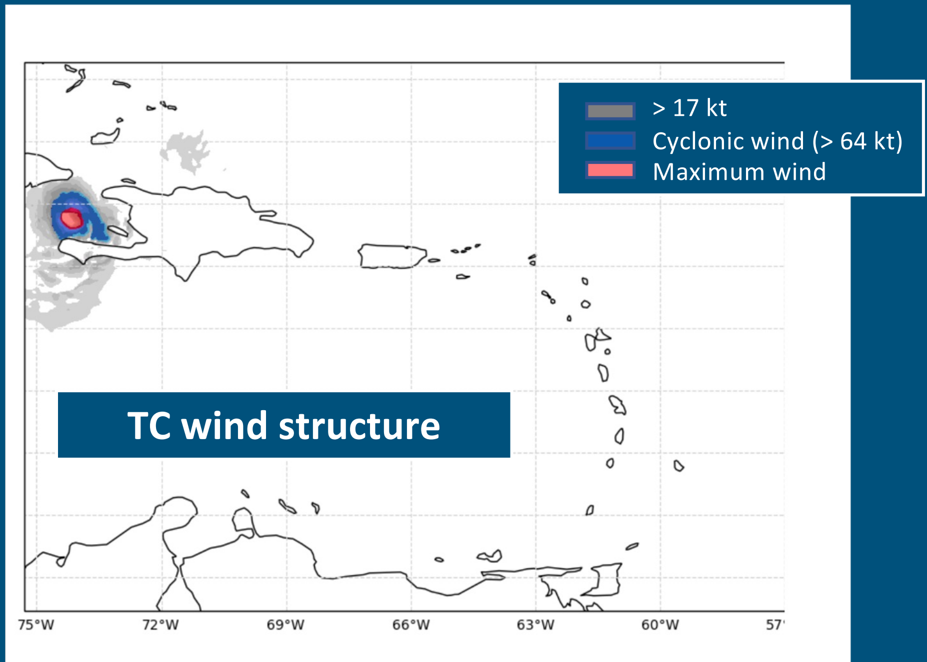
- ***AI-based generation of new EPS members*** for the kilometer-scale NWP system

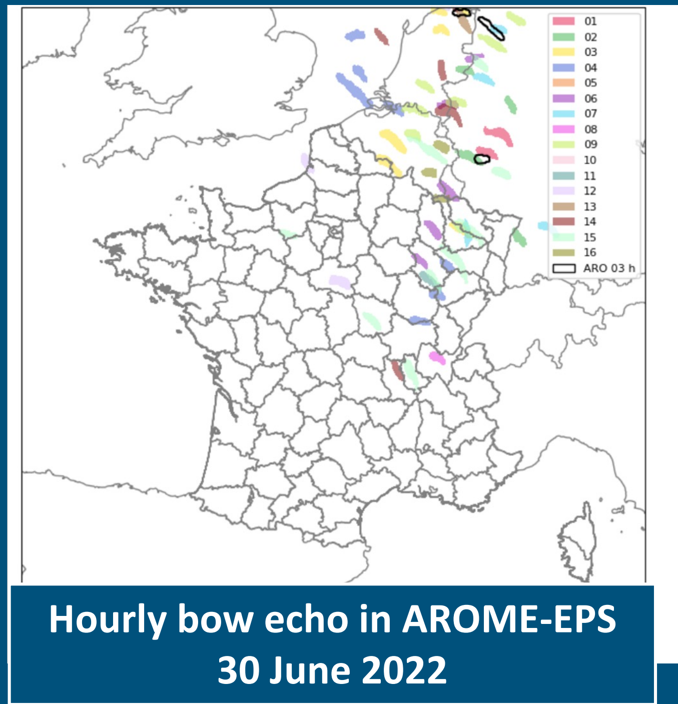
- + **continuing to explore the use of ML techniques for climate model calibration** (CNRM and IPSL).

Météo-France: **AI/ML in Post-Processing**

AI for weather object detection

Use of convolutional neural network to detect objects in AROME/AROME-EPS forecasts





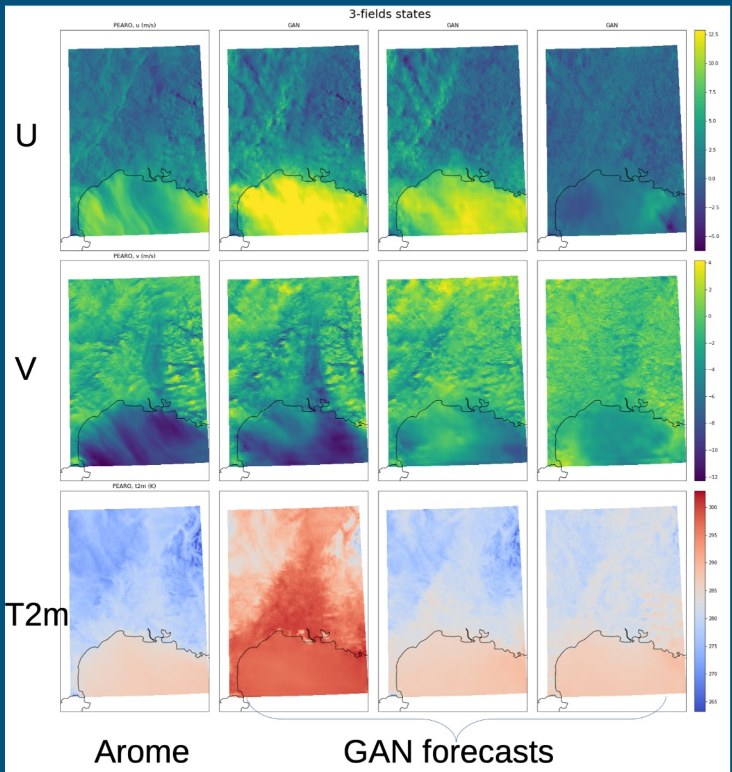
+ other promising tests for the dectection of supercells and weather fronts

Mounier et al. (2022, AIES)

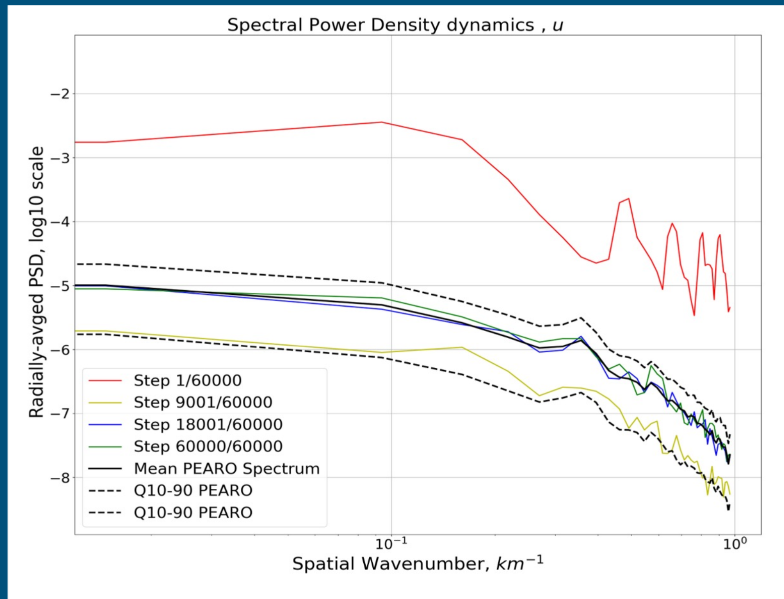
Météo-France: AI/ML for weather forecast generation

Design a hybrid EPS combining 'true' forecasts and AI-generated forecasts

- Use of Generative Adversarial Networks (GAN) to generate AROME-like forecasts



Zonal wind mean power spectrum



AROME-EPS MEAN
AROME-EPS Q10-90
1 iteration
9001 iterations
18001 iterations
60000 iterations

- Promising realistic GAN-generated atmospheric states

PhD C. Brochet, with ideas from <u>Besombes et al. 2021, NPG)</u>

Met Office

Tim Graham, Philip Brohan, Tom Dunstan

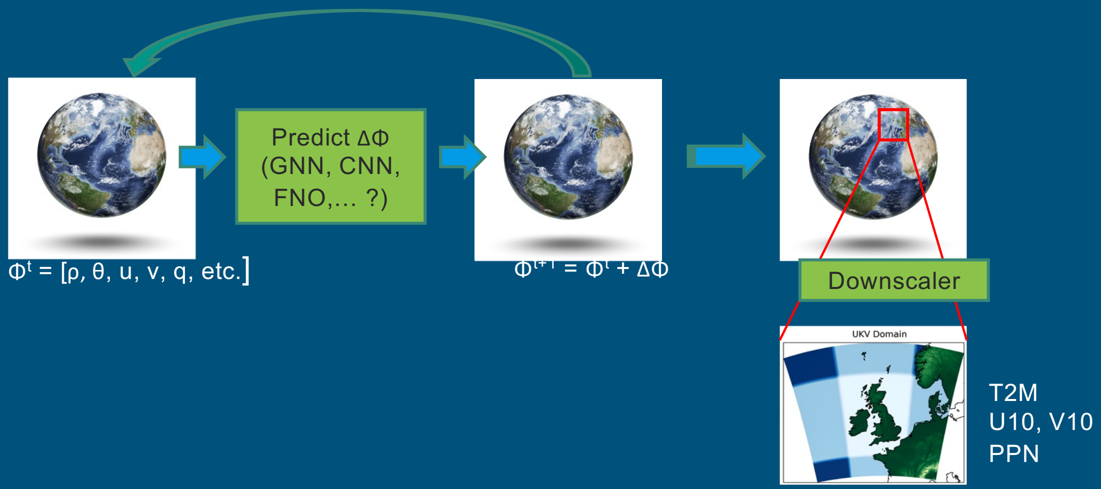
MetOffice: Overview of AI/ML Activities

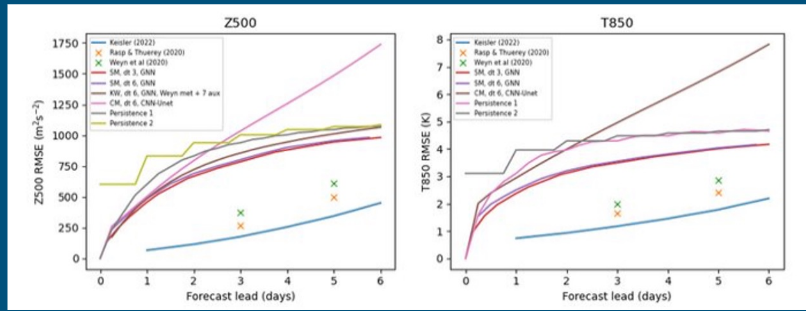
Data Science at the Met Office

- Key theme of the Met Office research and innovation strategy
- Training to develop skills of more staff
  - Provided by informatics lab.
  - At a team level developing examples relevant to current work.
- Working with partners (e.g. Microsoft)
- Four Capability areas:
  - Discovery and Attribution
  - Fusing Simulation with data science
  - Uncertainty and decisions
  - Data to decisions

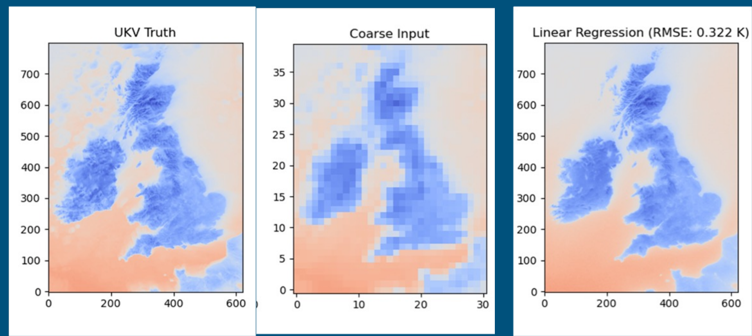
MetOffice: AI/ML in Downscaling

Moonshot Project: Data-Driven NWP with UK focus





Initial, low-resolution tests show skill against persistence



Simple, localised models work well for orographic corrections. Generative models needed for convective processes.

- Global forecast model + UK regional downscaler
- Forecast model trained on ERA5 reanalysis data
- Ongoing trials using Graph Neural Nets (GNN), Convolutional Neural Nets (CNN), and novel methods.

MetOffice: AI/ML in Reanalysis

ML based reanalysis over China

ML model of monthly climate state: MSLP, Precip, T2M, & SST anomalies at 0.25 degrees resolution.

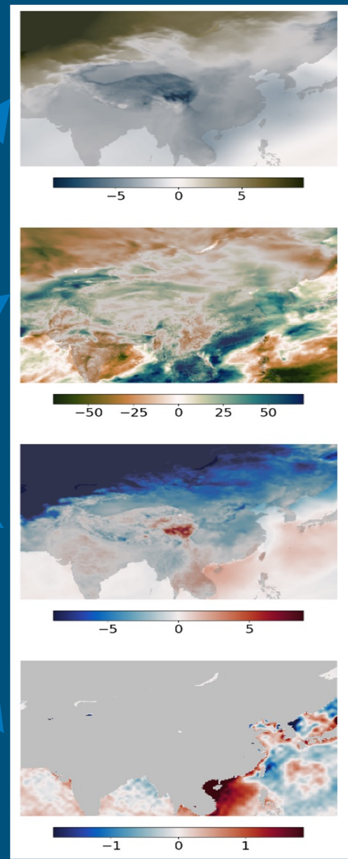
Model trained to represent multivariate climate state as a 100-dimensional latent-space (LS) vector. ML model is a Deep Convolutional Variational AutoEncoder, trained on ERA5.

ML model is bidirectional – can estimate LS vector for a month from an arbitrary subset of real climate state, and then recover full climate state from LS vector. => Data Assimilation: recover full state from sparse observations.

Application to reanalysis, observations sensitivity experiments, climate sensitivity experiments and extension to impacts variables.

VAE trained on ERA5 1959-2021

Latent-space vector $(x_1, x_2, \ldots, x_{100})$ ⟷ ML Model



Philip Brohan

MetOffice: AI/ML in Model Physics

- ML radiation parameterisation radiation trained on line by line code.
  - Currently a bit more expensive than existing radiation scheme
  - But based on line by line code so potentially more accurate
  - Will it be faster when running on GPU based systems?

- Non-Orographic gravity wave drag scheme
  - Produces similar results to the model parameterisation
  - On current CPU systems is slightly slower
  - Will it be faster or easier to port to GPUs?

MetOffice: AI/ML in Post-Processing

- Precipitation from model columns
  - Using NWP output from global model ensemble columns
  - Trained on radar data data.
  - Aim to provide better predictions of precipitation (probability of rainfall within different classes) than the raw model output.

- Downscaling climate data (wind) to energy security decisions.
  - Testing Autoencoder and Long Short-Term Memory (LSTM) methods
  - LSTM performs best and shows improvement over linear interpolation.

NCEP

Fanglin Yang

NCEP: Overview of AI/ML Activities

Observations	Data Assimilation	Forecast	Post/Product
Radiosonde processing	Physics emulation	Air Quality Forecast: Accelerated Transport	Wave feature identification (Wave-Watch-III)
Satellite Thinning	Improved Background	Atmospheric Chemistry Emulator	Rip Currents
	Background Error Covariances	Physics Suite Emulation	Air Quality Bias Correction
	Use analysis increments to diagnose and correct model biases	Radiation Parameterizations	Sub-Seasonal/ Seasonal forecast products
		GL Wave Emulation	

NCEP: AI/ML in Model Dynamics

AI Emulator for Accelerated Transport of Chemical Species

Evaluate the use of AI-based emulators for accelerated transport of atmospheric tracers: chemical species and aerosols.

Input: 3D tracer concentrations for transported species, pressure, temperature, wind, humidity, grid information (lat/lon/height,date,time)

Output: 3D tracer concentrations for transported species, grid information(lat/lon/height, date,time)

Format: netcdf

Geographic coverage: CONUS

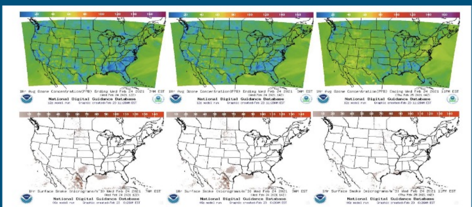
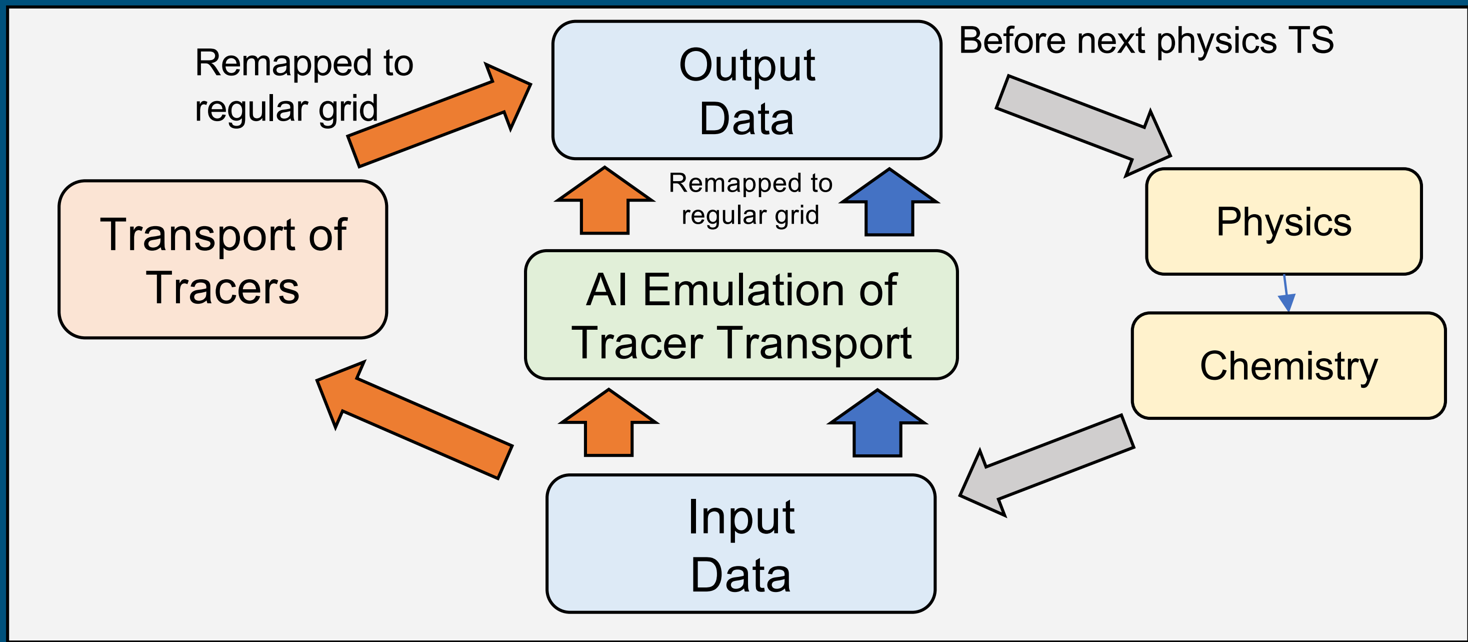


Figure 1: Concentrations of ozone (top) and smoke (bottom) over 4 hours in the lowest model layer in the operational NWS AQ predictions for 7 am (left), 9 am (middle) and 11 am EST (right) on 2/24/2021. Ozone increases, but smoke is dispersed as sun rises (source: https://airquality.weather.gov/).



Learning Mode

Prediction Mode

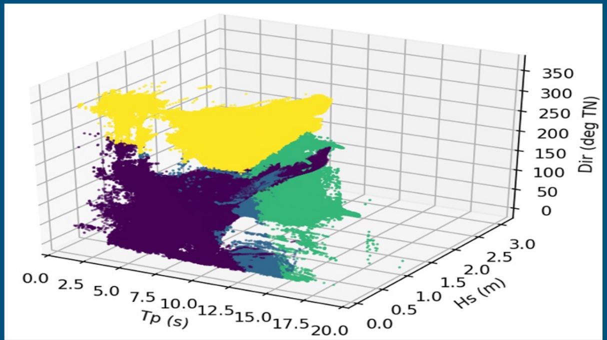
Sleeman et al

NCEP: AI/ML in Post-Processing

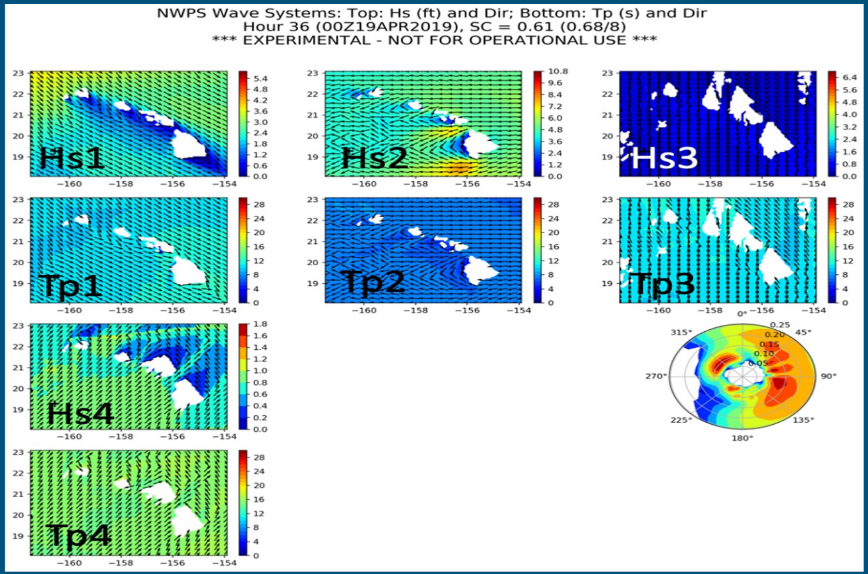
Wave System Identification Using Clustering

Using k-means clustering to identify spatially and temporally consistent wave systems from the output of Nearshore Wave Prediction System v1.3

k-means clustering in wave parameter space



Clustered wave systems in geo space with heights (Hs1-Hs4) and periods (Tp1-Tp4).



(Van der Westhuysen, 2020)

NRL

Carolyn Reynolds

Machine Learning of Atmospheric Boundary Layer Turbulence

David Flagg[1], Jeffrey Byers[1], Katarina Doctor[1], James Doyle[1], Hao Jin[1], Saša Gaberšek[1]
[1]U.S. Naval Research Laboratory

Q: Can we use machine learning of large-eddy simulation (LES) of the boundary layer to improve our ability to parameterize turbulence in numerical weather prediction (NWP)?

Concept

- Create a series of feature vectors ($\boldsymbol{\pi}$) through a Buckingham-Pi analysis of fields contributing to turbulence . . .common framework for comparison of NWP (COAMPS®) and LES
- Build regression function (f_H) for turbulence metric π_0 from LES fields
- Apply function (f_H) to NWP fields to predict $\pi_0{}'$
- Compare π_0 and $\pi_0{}'$ to learn about deficiencies in NWP boundary layer parameterization

π_0 = Reference turbulence metric

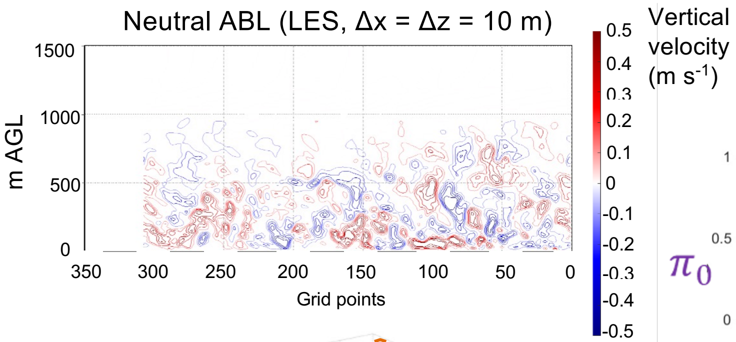
LES fields

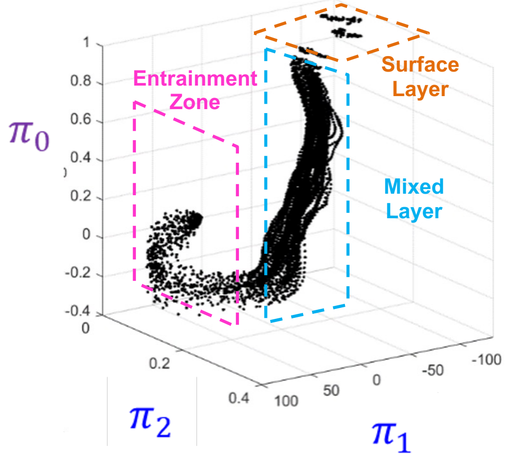
$$\pi_0 = f_H\,(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \boldsymbol{\pi}_3, \ldots)$$

$$\Downarrow$$

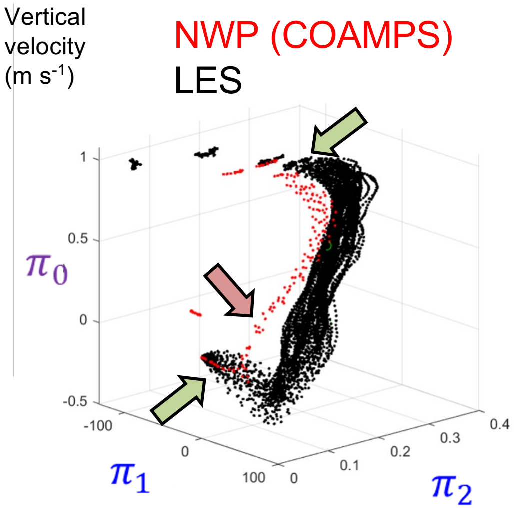
NWP fields

$$\pi_0{}' = f_H\,(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \boldsymbol{\pi}_3, \ldots)$$





LES fields over 1 hr; dry, convective boundary layer



NWP & LES matching well in the surface layer & above the boundary layer

Large discrepancy in the entrainment zone

Contact:
david.flagg@nrlmry.navy.mil

COAMPS® is a registered trademark of the U.S. Naval Research Laboratory
Distribution Statement A. Approved for public release. Distribution is unlimited

Summary

Data assimilation

- CMC, DWD, ECMWF, INPE, NCEP, HydroMet of Russia

Model Physics

- Radiative Transfer (CMC, ECMWF, Met Office, NCEP)
- Ice Microphysics (DWD)
- Non-Orographic gravity Wave Drag (JMA, Met Office)
- Boundary Layer Turbulence (NRL)
- Model calibration (Météo-France)

Model Dynamics

- Accelerated tracer transport (NCEP)

Emulation

- Generation of reanalysis without model (Met Office)
- ML based forecast systems (Met Office, NCEP)
- S2S forecast competition (ECMWF)

Post-processing

- Temperature postprocessing (ECMWF, Met Office)
- Flight turbulence (INPE)
- Convective Weather (INPE)
- Precipitation prediction (Met Office)
- Downscaling climate model data (Met Office)
- Wave Systems (NCEP)
- RIP currents (NCEP)
- Air quality bias correction (NCEP)
- Ensemble products (Météo-France)
- Weather object detection (Météo-France)

Thank You