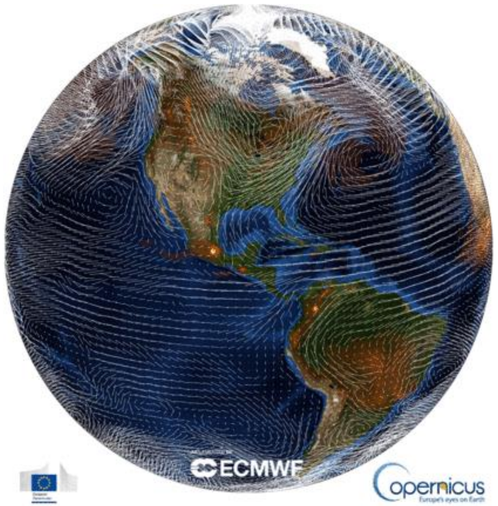
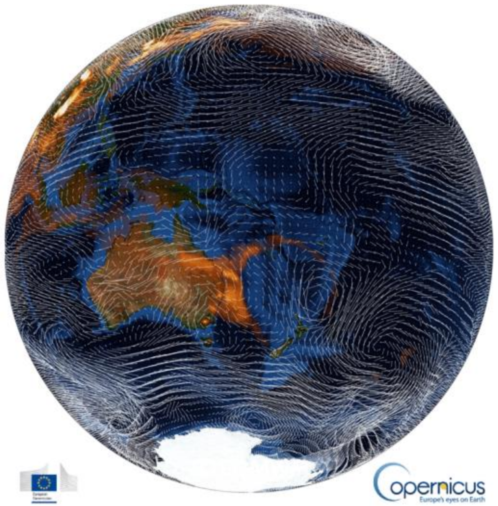




Courtesy: Antje Inness, ECMWF

Machine Learning in Earth System Science

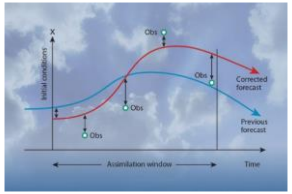
MARTIN G. SCHULTZ | JÜLICH SUPERCOMPUTING CENTRE, GERMANY

ESM User forum, JSC, Jülich, 4-6 Feb 2020

Mitglied der Helmholtz-Gemeinschaft

JÜLICH Forschungszentrum

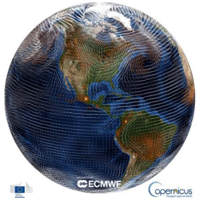
Workflow of an Earth system model

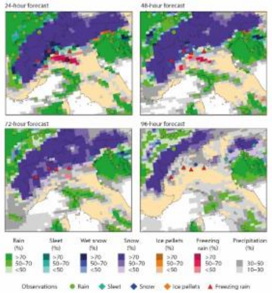


Data publication

↑

Data aquisition → Data processing → (Data assimilation) → Model run (forecast) → Data post-processing





Machine learning could play a role
in all steps of this workflow

After Peter Bauer, 2019

What is machine learning?

”Machine learning algorithms are described as learning a target function(f) that best maps input variables to an output variable (Y): $Y = f(X)$” (James Le)

Definition

- A network is a function $\hat{\phi}\left(\mathbf{x}^{(i)}\right) = \hat{\mathbf{y}}^{(i)}$ which tries to approximate the true function $\phi\left(\mathbf{x}^{(i)}\right) = \mathbf{y}^{(i)}$

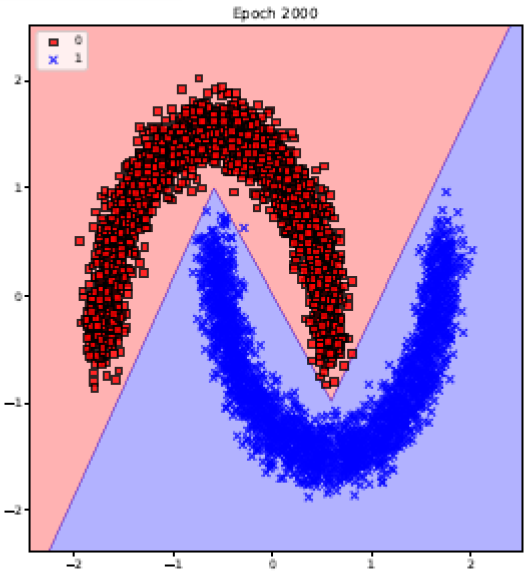
Neural networks can approximate non-linear functions

Activation of neurons in layer l :

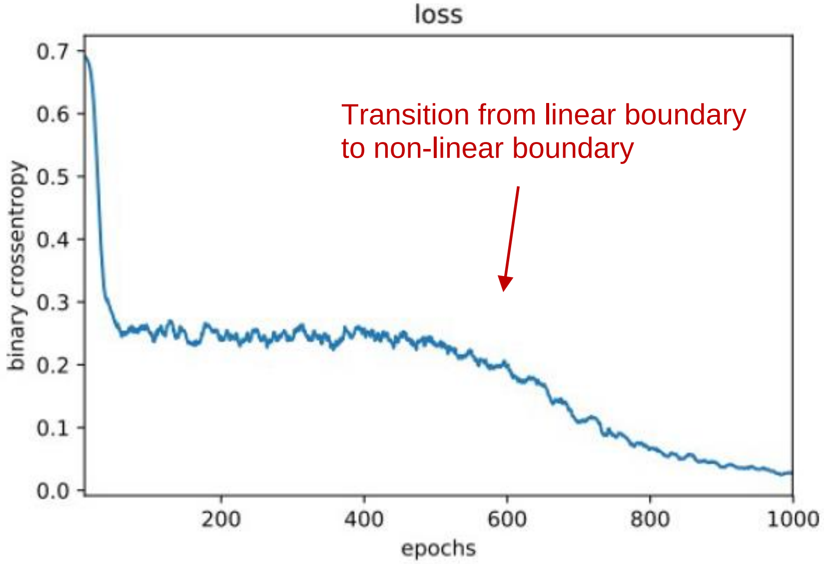
$$\mathbf{A}^{[l]} = g^{[l]}\left(\mathbf{Z}^{[l]}\right) \quad \text{with} \quad \mathbf{Z}^{[l]} = \mathbf{W}^{[l]}\mathbf{A}^{[l-1]} + \mathbf{b}^{[l]}$$

The capability to learn non-linear functions arises from the non-linearity of the activation function g

The network has learned a non-linear shape

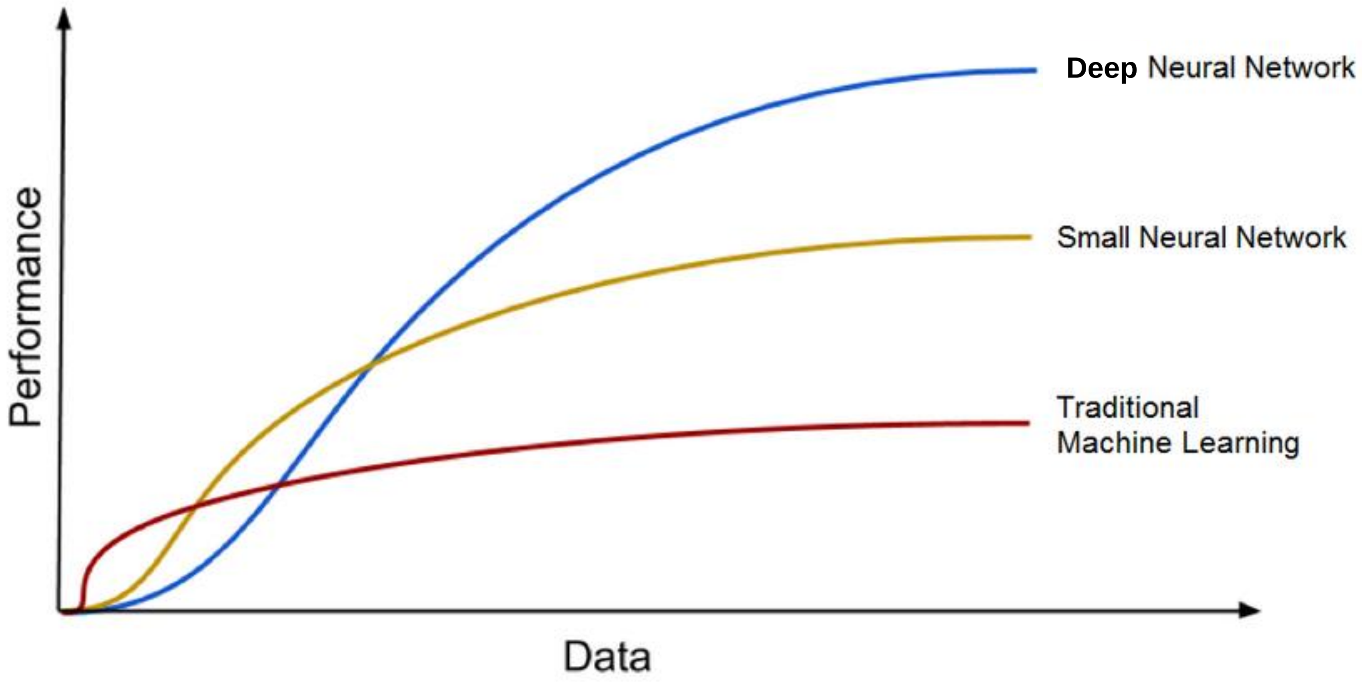


The network abstracted from the actual shape



The learning process minimizes a cost function

The power of deep learning



source: https://www.deeplearning-academy.com/p/ai-wiki-machine-learning-vs-deep-learning

Machine learning in weather forecasting

AMS
American Meteorological Society

Accurate and Fast Neural Network Emulations of Model Radiation for the NCEP Coupled Climate Forecast System: Climate Simulations and Seasonal Predictions*

V. M. Krasnopolsky
NOAA/NCEP, Science Applications International Corporation, Camp Springs, and Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

See all authors & affiliations

https://doi.org/10.1175/2009MWR3149.1
Received: 14 July 2009
Final Form: 2 December 2009
Published Online: 1 May 2010

Krasnopolsky et al., 2010

NCEP model:

- Two neural networks (LWR and SWR)
- Ca 600 inputs and 70 outputs each
- Time saving: ~factor 12 (LWR), factor 60 (SWR)
- Accuracy similar to control run

→ Possible to run radiation at every time step

https://doi.org/10.1175/2009MWR3149.1

Krasnopolsky et al., 2010

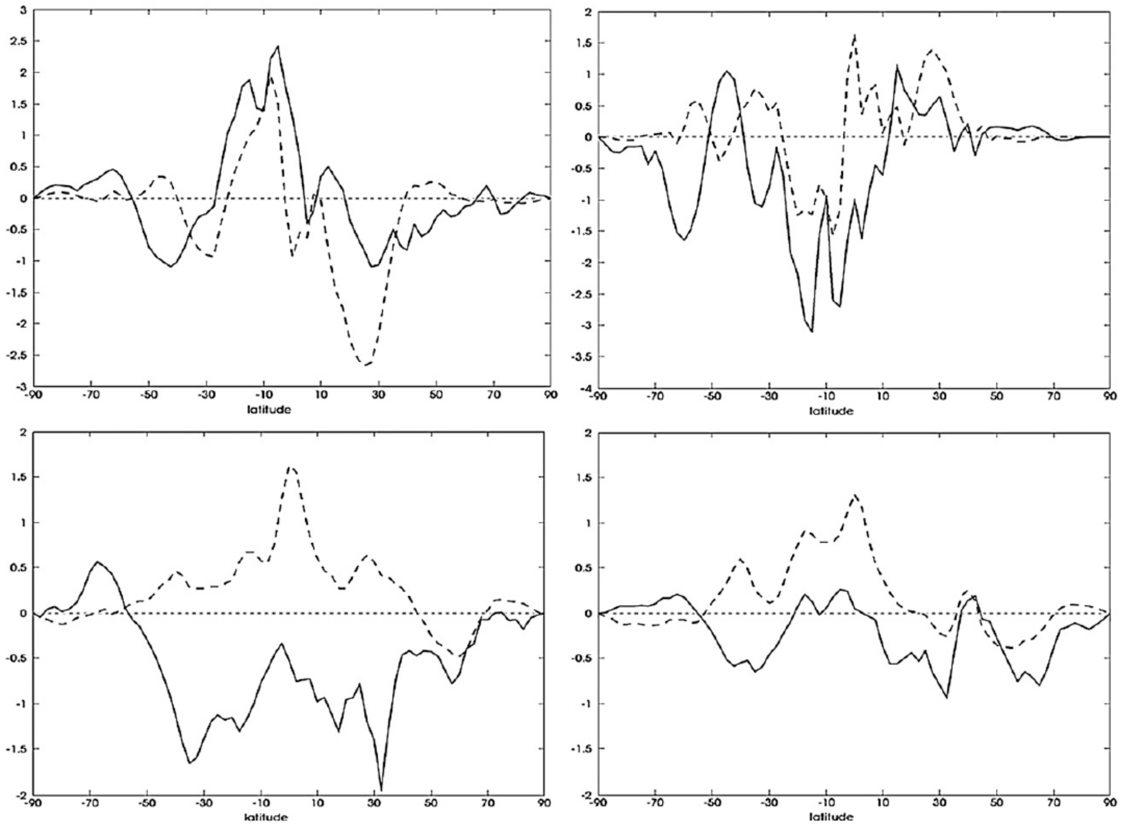


FIG. 2. (top) Zonal and time-mean top-of-atmosphere upward (left) LWR and (right) SWR flux (W m^{-2}) for DJF. The solid line is NN − CTL (the difference between the NN and a control run), and the dashed line is CTL1 − CTL (the difference between two control runs) presented for comparison. (bottom) Zonal and time-annual mean (left) downward and (right) upward surface LWR flux (W m^{-2}). The fluxes' differences are multiplied by cos(lat) to equalize the areas.

https://doi.org/10.1175/2009MWR3149.1

Recent studies

frontiers in Physics

REVIEW
published: 10 October 2019
doi: 10.3389/fphy.2019.00153

Check for updates

The Application of Machine Learning Techniques to Improve El Niño Prediction Skill

Henk A. Dijkstra[1,2*], *Paul Petersik*[1], *Emilio Hernández-García*[3] *and Cristóbal López*[3]

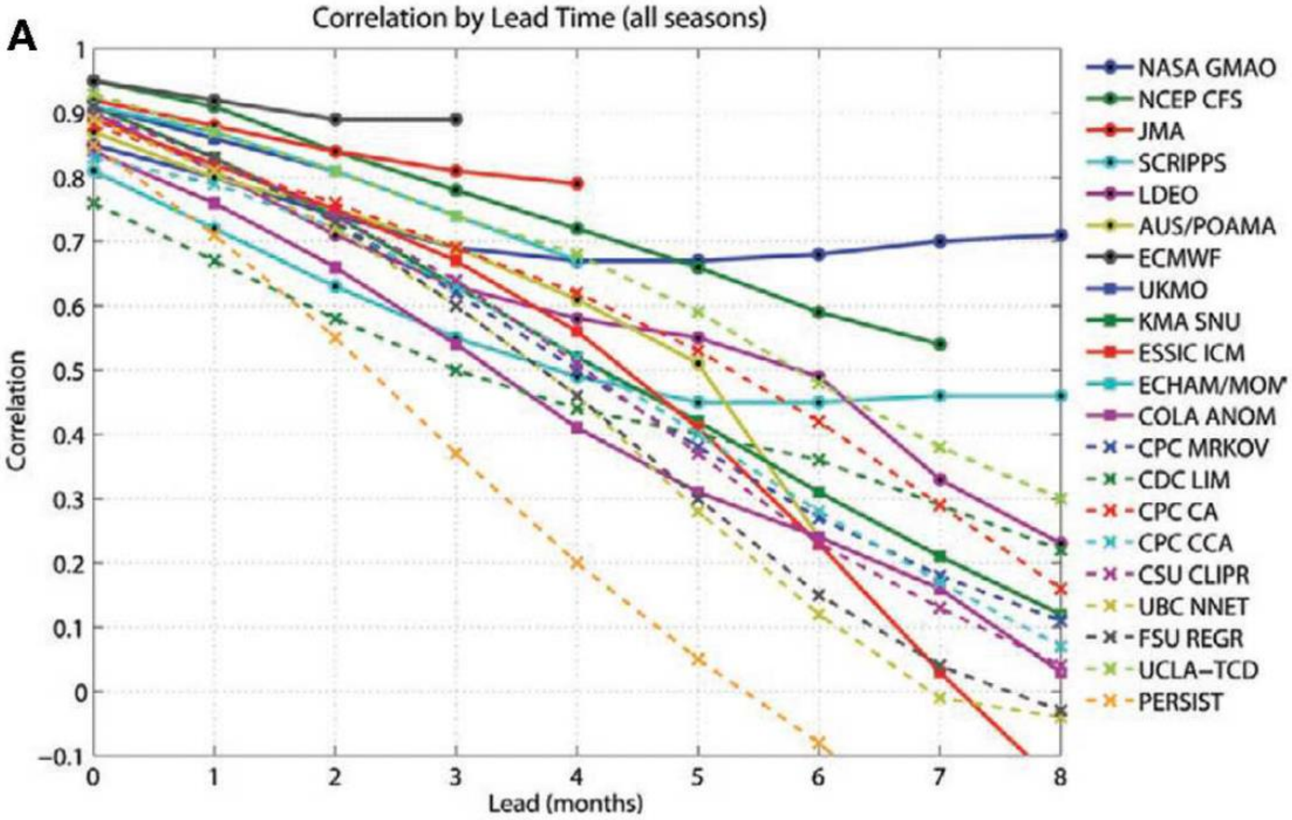
[1] Department of Physics, Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands, [2] Department of Physics, Center for Complex Systems Studies, Utrecht University, Utrecht, Netherlands, [3] IFISC (Spanish National Research Council - University of the Balearic Islands), Instituto de Física Interdisciplinar y Sistemas Complejos, Palma de Mallorca, Spain

https://doi.org/10.3389/fphy.2019.00153

Dijkstra et al., 2019

Physical models

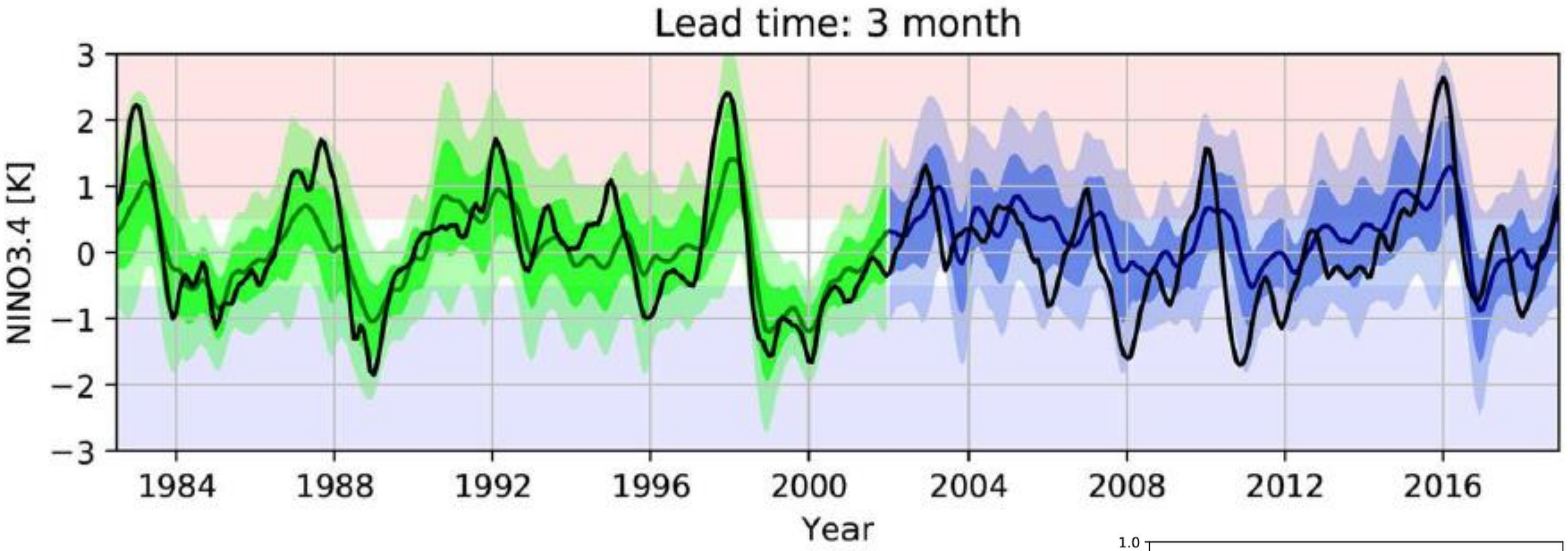
Statistical models

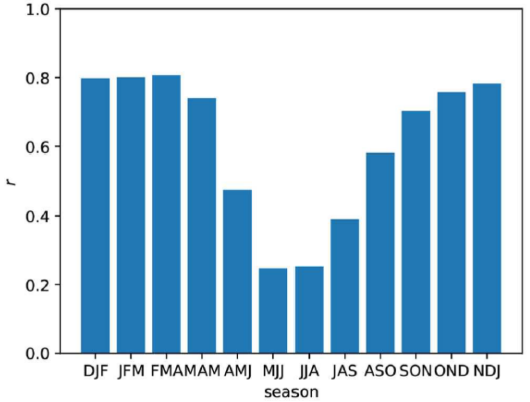


https://doi.org/10.3389/fphy.2019.00153

Original figure:
Barnston et al. (2012)

Dijkstra et al., 2019





https://doi.org/10.3389/fphy.2019.00153

Has the future arrived already?

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER

10.1029/2018GL080704

Key Points:
- A neural network can emulate the dynamics of a simple general circulation model
- The trained network can successfully forecast the model weather
- The network can produce a realistic representation of the model climate

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Video S1
- Video S2
- Video S3

Toward Data-Driven Weather and Climate Forecasting: Approximating a Simple General Circulation Model With Deep Learning

S. Scher[1]

[1]Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

Abstract It is shown that it is possible to emulate the dynamics of a simple general circulation model with a deep neural network. After being trained on the model, the network can predict the complete model state several time steps ahead—which conceptually is making weather forecasts in the model world. Additionally, after being initialized with an arbitrary model state, the network can through repeatedly feeding back its predictions into its inputs create a climate run, which has similar climate statistics to the climate of the general circulation model. This network climate run shows no long-term drift, even though no conservation properties were explicitly designed into the network.

A Deep Hybrid Model for Weather Forecasting

Aditya Grover*
IIT Delhi
aditya.grover1@gmail.com

Ashish Kapoor
Microsoft Research
akapoor@microsoft.com

Eric Horvitz
Microsoft Research
horvitz@microsoft.com

2015

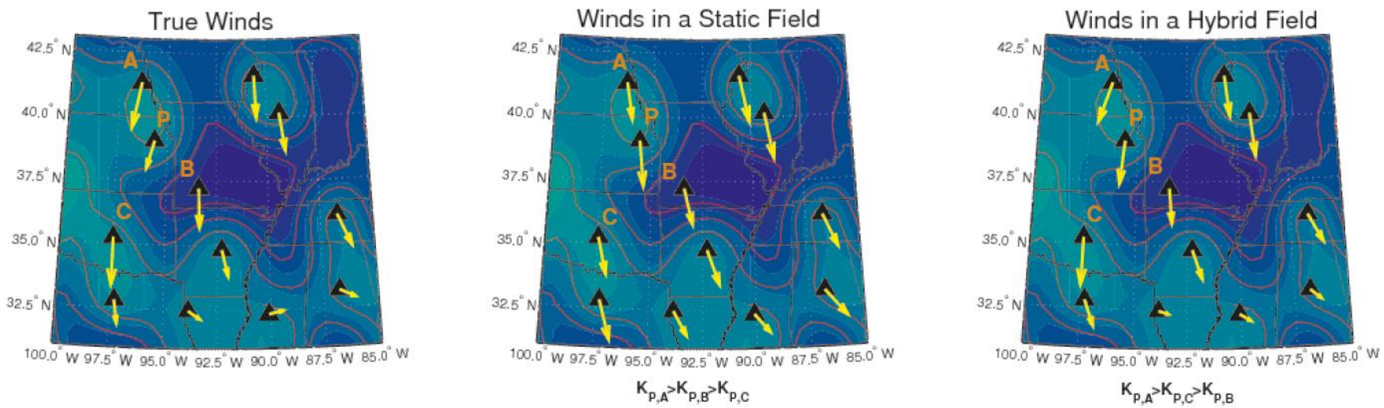
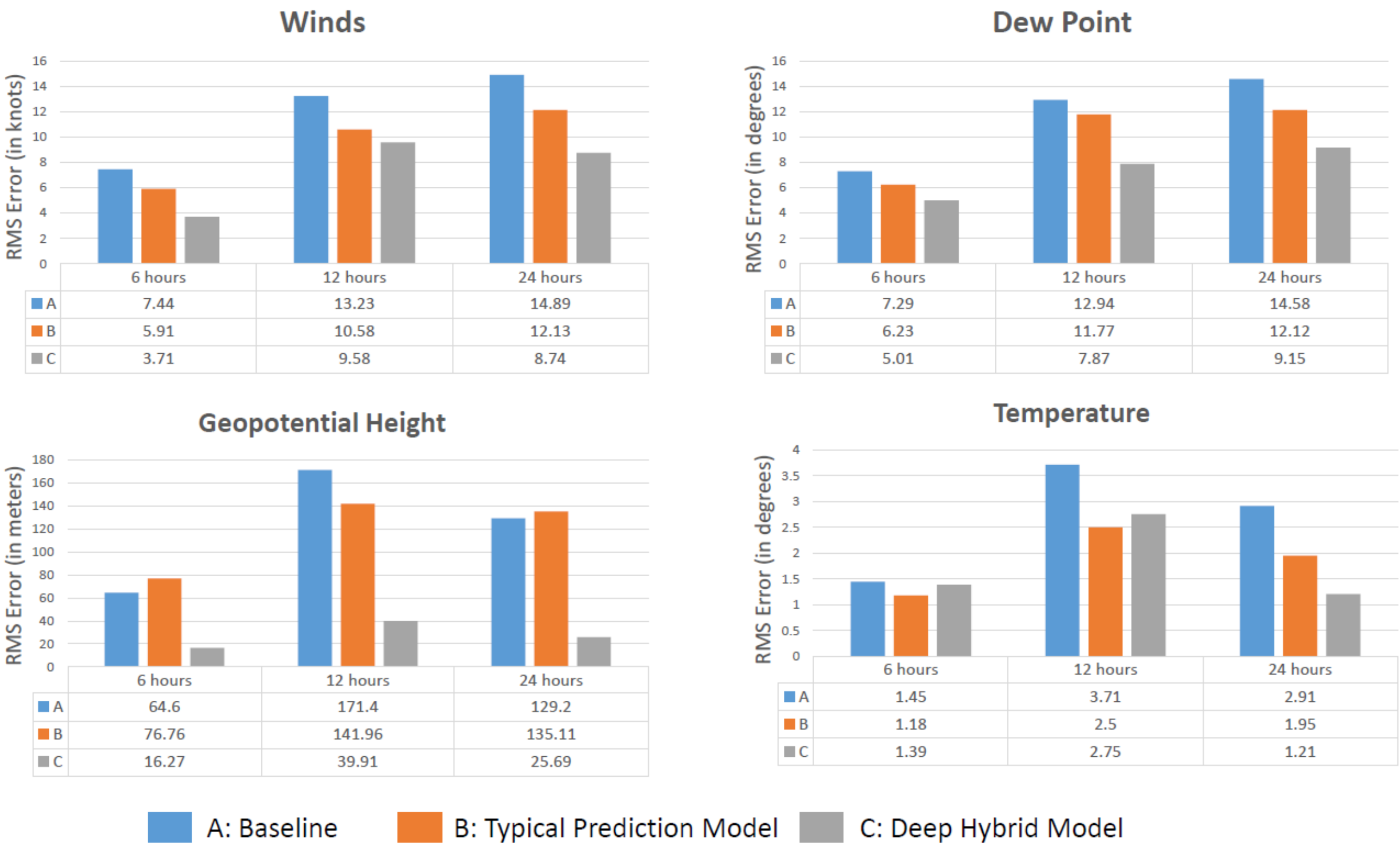


Figure 1: Spatial Interpolation of winds in a static and hybrid field. Filled contours represent temperature and isobar lines are marked in red. In the hybrid field, the interpolated wind vectors are closely aligned with the true values. However, the static field fails to account for the long-range dependencies.

https://doi.org/10.1145/2783258.2783275

Grover et al., 2015



Winds

	6 hours	12 hours	24 hours
A	7.44	13.23	14.89
B	5.91	10.58	12.13
C	3.71	9.58	8.74

Dew Point

	6 hours	12 hours	24 hours
A	7.29	12.94	14.58
B	6.23	11.77	12.12
C	5.01	7.87	9.15

Geopotential Height

	6 hours	12 hours	24 hours
A	64.6	171.4	129.2
B	76.76	141.96	135.11
C	16.27	39.91	25.69

Temperature

	6 hours	12 hours	24 hours
A	1.45	3.71	2.91
B	1.18	2.5	1.95
C	1.39	2.75	1.21

https://doi.org/10.1145/2783258.2783275

Satellite data processing

Near-real-time forecast of satellite-based soil moisture using long short-term memory with an adaptive data integration kernel

Kuai Fang and Chaopeng Shen*

Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA.

https://doi.org/10.1175/JHM-D-19-0169.1

Fang and Shen, 2019

One year training, two years testing, 500 epochs
Includes a „step-wise data intgeration kernel“

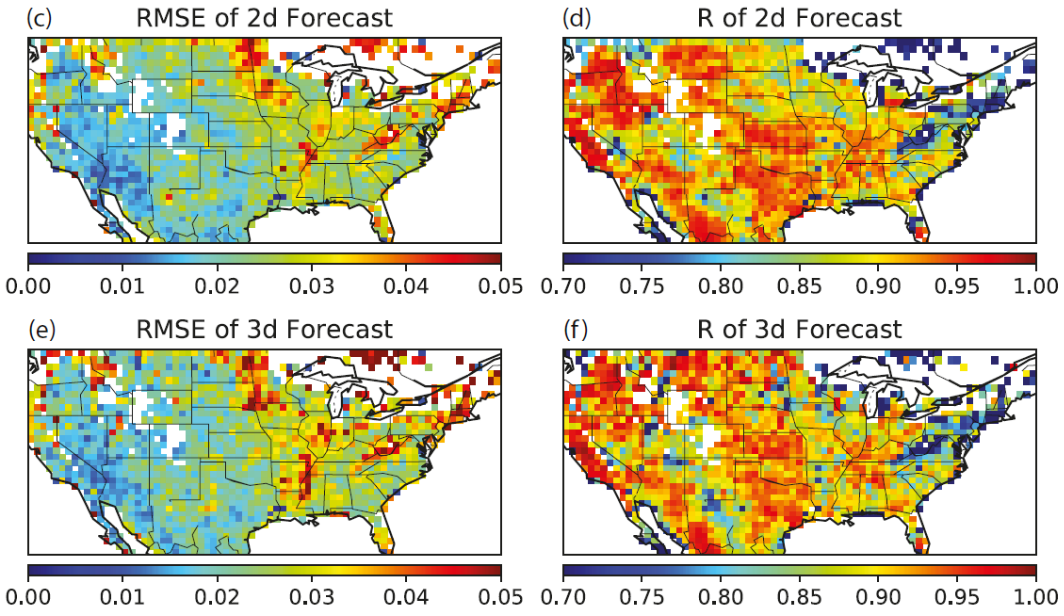


FIG. 3. Performance of soil moisture forecasts evaluated against SMAP L3 product. Panels (a), (c) and (e) show the RMSE for 1-, 2-, and 3-day forecast, respectively, while (b), (d) and (f) show R.

https://doi.org/10.1175/JHM-D-19-0169.1

Fang and Shen, 2019

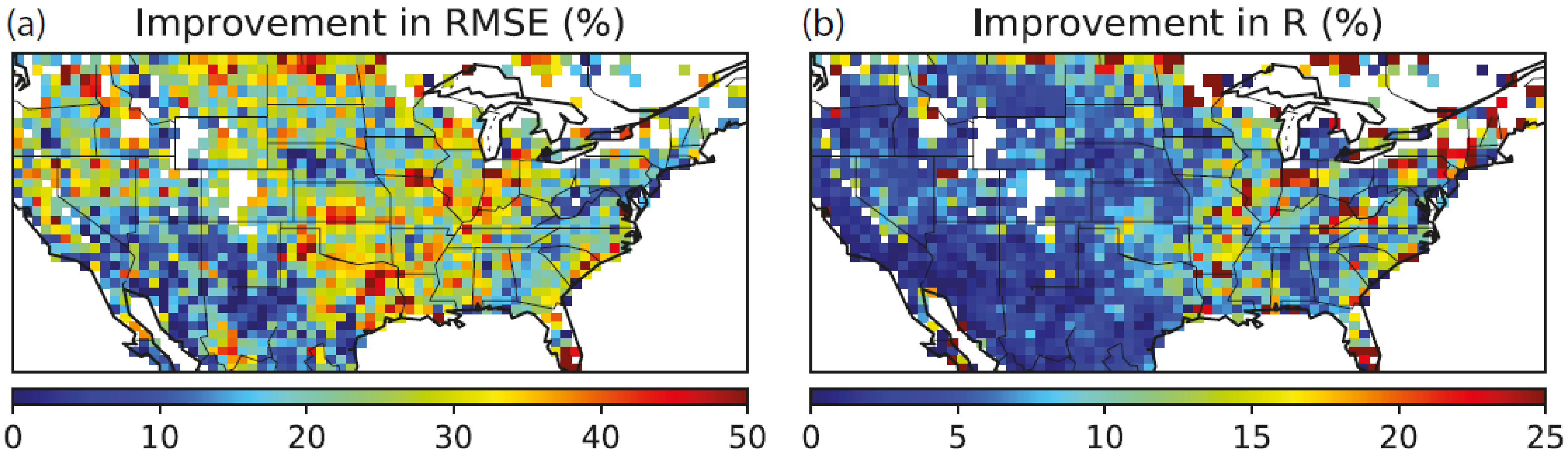
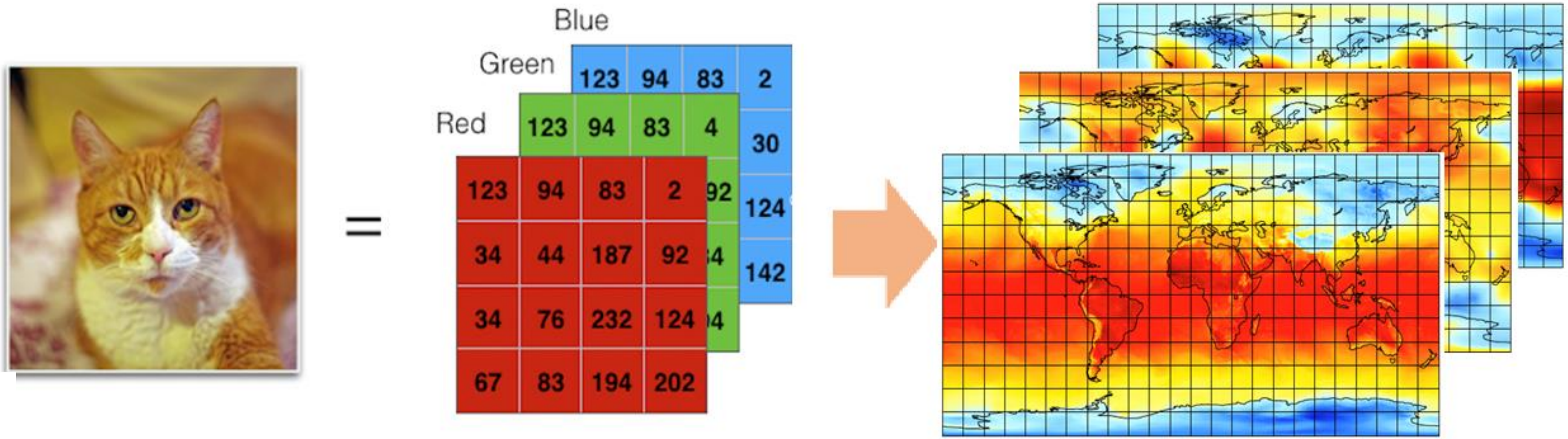


FIG. 4. Improvements from projection to forecast models. RMSE and R improvements were calculated as $\frac{RMSE(Project)-RMSE(Forecast)}{RMSE(Project)} \times 100\%$ and $\frac{R(Forecast)-R(Project)}{R(Project)} \times 100\%$.

https://doi.org/10.1175/JHM-D-19-0169.1

B. Gong (in preparation)

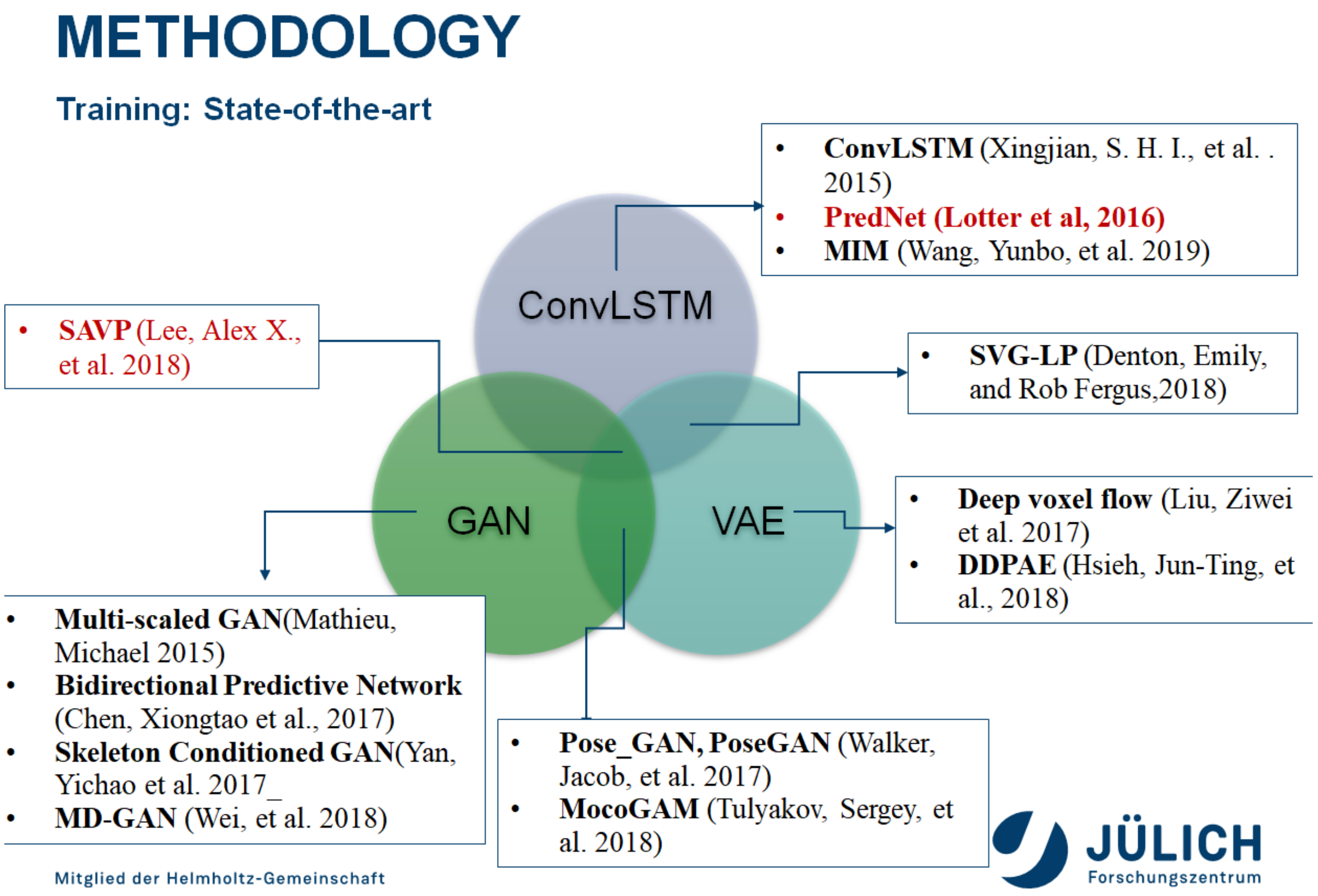
Relation between video prediction and weather forecasting



European Research Council
Established by the European Commission

Advanced Grant
ERC-2017-ADG
#787576

B. Gong (in preparation)

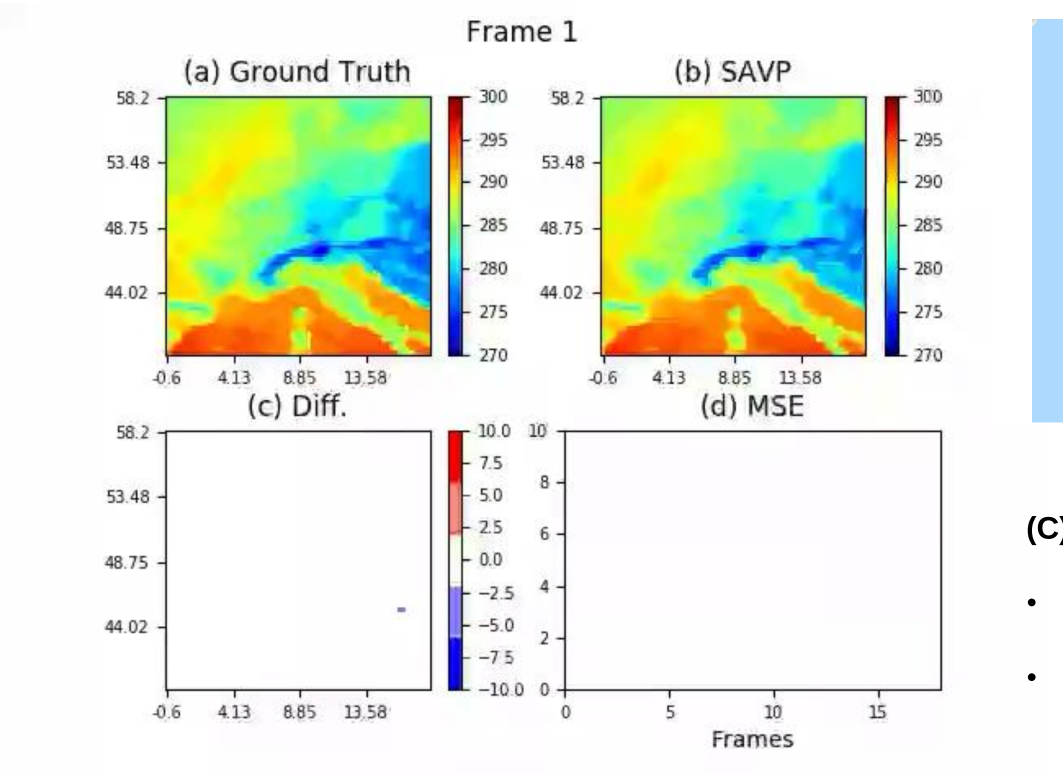


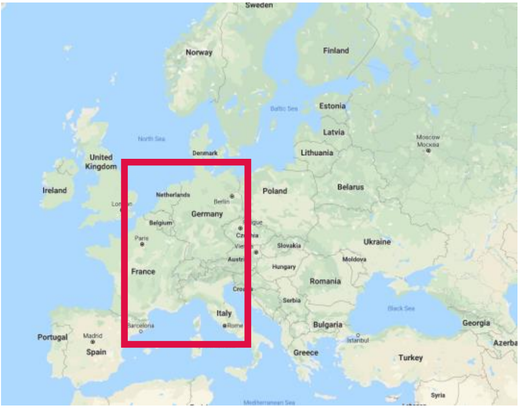
European Research Council
Established by the European Commission

Advanced Grant
ERC-2017-ADG
#787576

B. Gong (in preparation)

Post-processing: Visualization for Temperature Prediction (SAVP)





(C) Diff (= GT - SAVP):

- The value towards red color means the temperature were underestimated;
- The blue color corresponds to the overestimated values.

European Research Council
Established by the European Commission

Advanced Grant
ERC-2017-ADG
#787576

JÜLICH
Forschungszentrum

Kleinert et al., 2020 (in prep.)

Station wise data (Germany)

- Chemical data are taken from **Tropospheric Ozone Assessment Report** (TOAR)
  Statistics according to the EU definition of ozone concentrations
  Station type: Background
- Meteorological data are taken from **COSMO-REA6**

$\rightarrow$ **330 stations are used**

Variable	Description	Type	Unit
O_3	Daily maximum 8h average ozone	measur.	nmol mol^{-1}
NO	Daily maximum 8h average nitrogen oxide	measur.	nmol mol^{-1}
NO_2	Daily maximum 8h average nitrogen dioxide	measur.	nmol mol^{-1}
u,v	Daily mean of wind's components	model rea	m/s
T-2M	Daily maximum of 2m temperature	model rea	$^\circ$ C
relhum	Daily mean of relative humidity	model rea	%
pblheight	Daily maximum height of planetary boundary layer	model rea	m
cloudcover	Daily mean cloud cover	model rea	%

European Research Council
Established by the European Commission

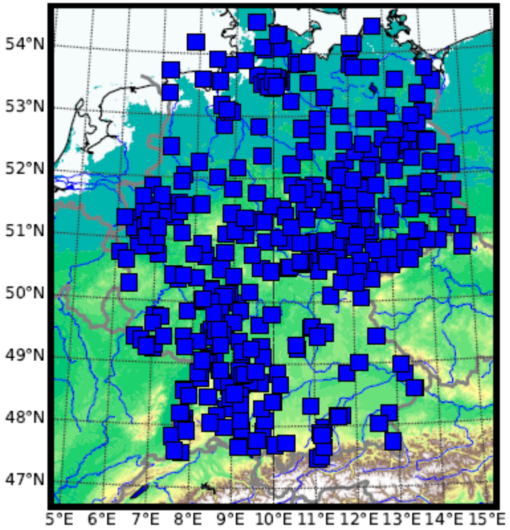
Advanced Grant
ERC-2017-ADG
#787576

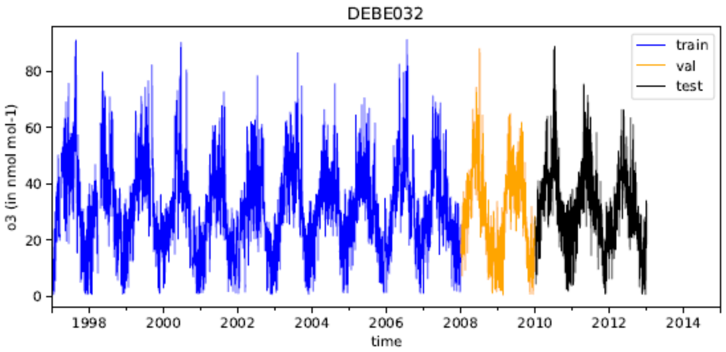
Kleinert et al., 2020 (in prep.)

Experimental Setup

Data preparation

- Input (X):
  - Create 2D data for each time step (daily data of **previous 7 days** $\times$ **variables** (as channels))
  - Split time series into (**no** overlap)
    - training data (Jan 1997-Dec 2007) $\sim 642,000$ samples
    - validation data (Jan 2008- Dec 2009) $\sim 145,000$ samples
    - test data (Jan 2010 - Dec 2015) $\sim 214,000$ samples
  - Standardised per station
- Output (y):
  - Create target vector of daily **maximum 8-hour average** ozone concentrations (**lead time of up to 4 days**)





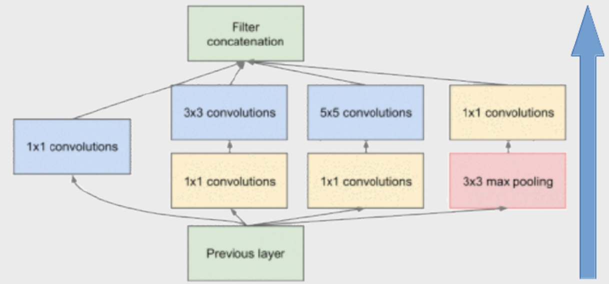
European Research Council
Established by the European Commission

Advanced Grant
ERC-2017-ADG
#787576

Kleinert et al., 2020 (in prep.)

Deep CNN with inception blocks

Original structure proposed by Szegedy et al. (2015)

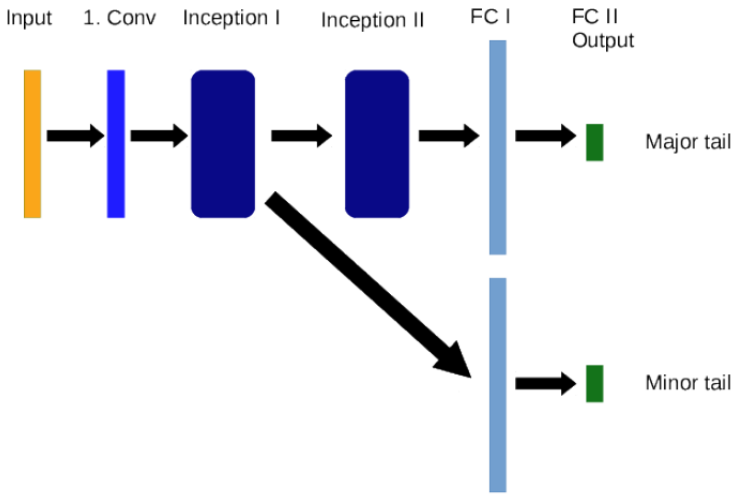


Modified Settings:

Conv. kernel size	pooling layers
1×1	Max pooling 3×1
3×1	mean pooling 3×1
5×1	

$\Rightarrow$ five pathways per block

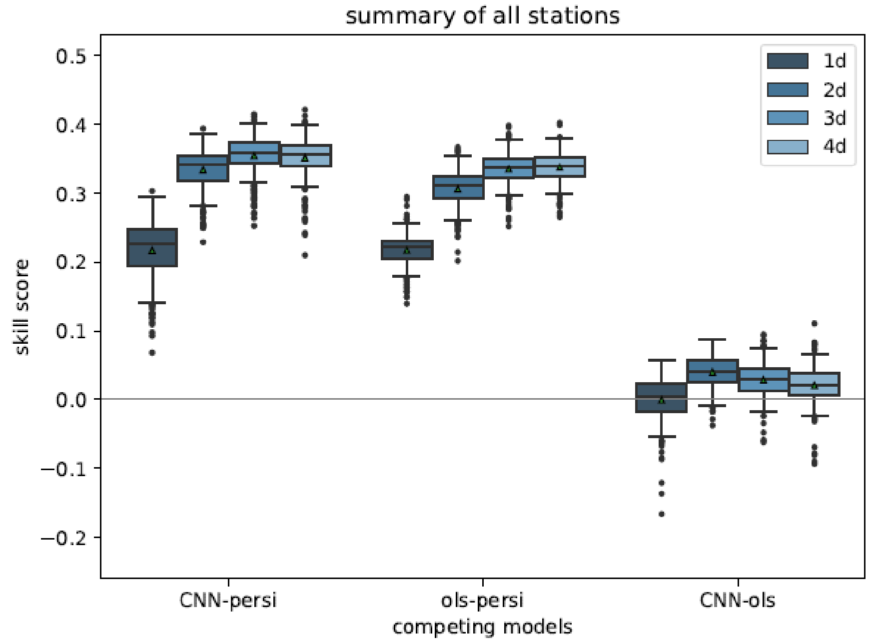
- Activations: PReLU
- Dropout rate: 0.1



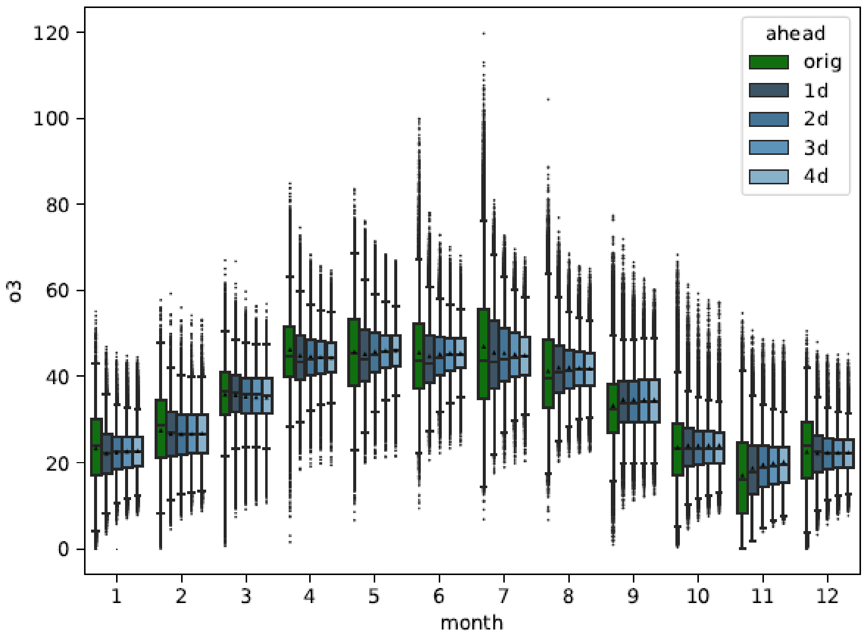
European Research Council
Established by the European Commission

Advanced Grant
ERC-2017-ADG
#787576

Kleinert et al., 2020 (in prep.)



(a) [<model of interest> - < reference model>]



(b) CNN: monthly comparison for up to 4 days lead

European Research Council
Established by the European Commission

Advanced Grant
ERC-2017-ADG
#787576

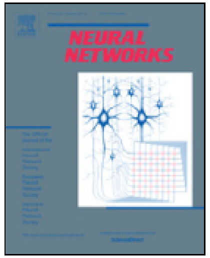
Ozone forecasting also in the US …

Neural Networks 121 (2020) 396–408

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

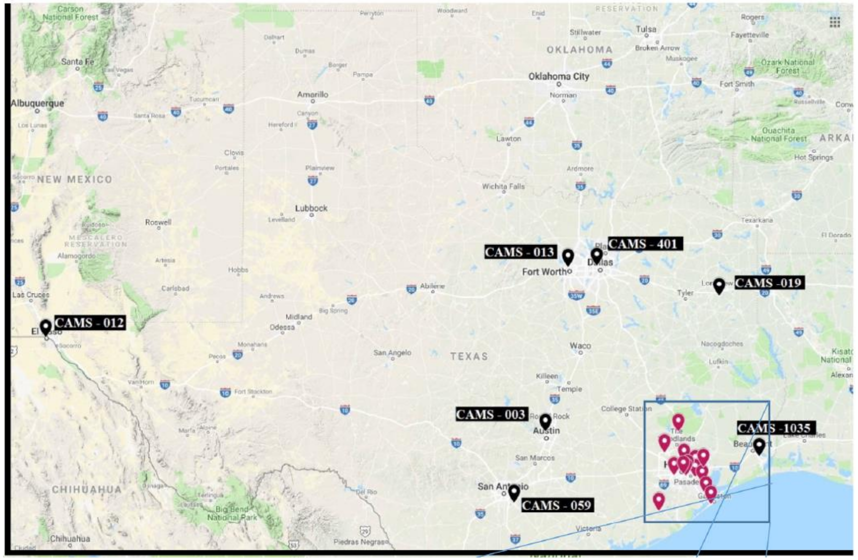


Using a deep convolutional neural network to predict 2017 ozone concentrations, 24 hours in advance

Alqamah Sayeed, Yunsoo Choi *, Ebrahim Eslami, Yannic Lops, Anirban Roy, Jia Jung

Department of Earth and Atmospheric Sciences, University of Houston, TX 77004, United States of America

Sayeed et al., 2020



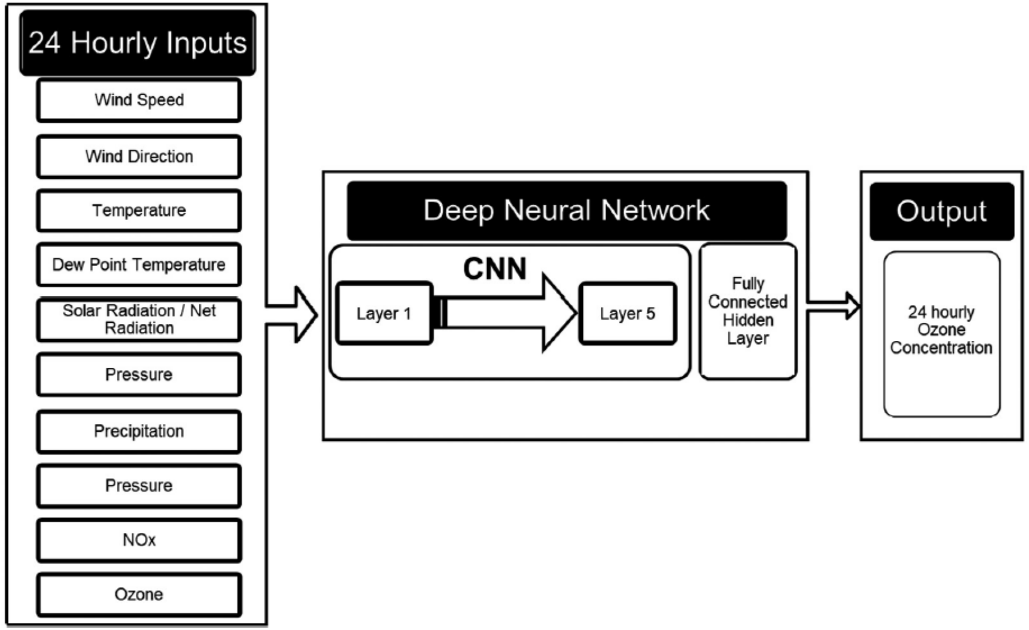
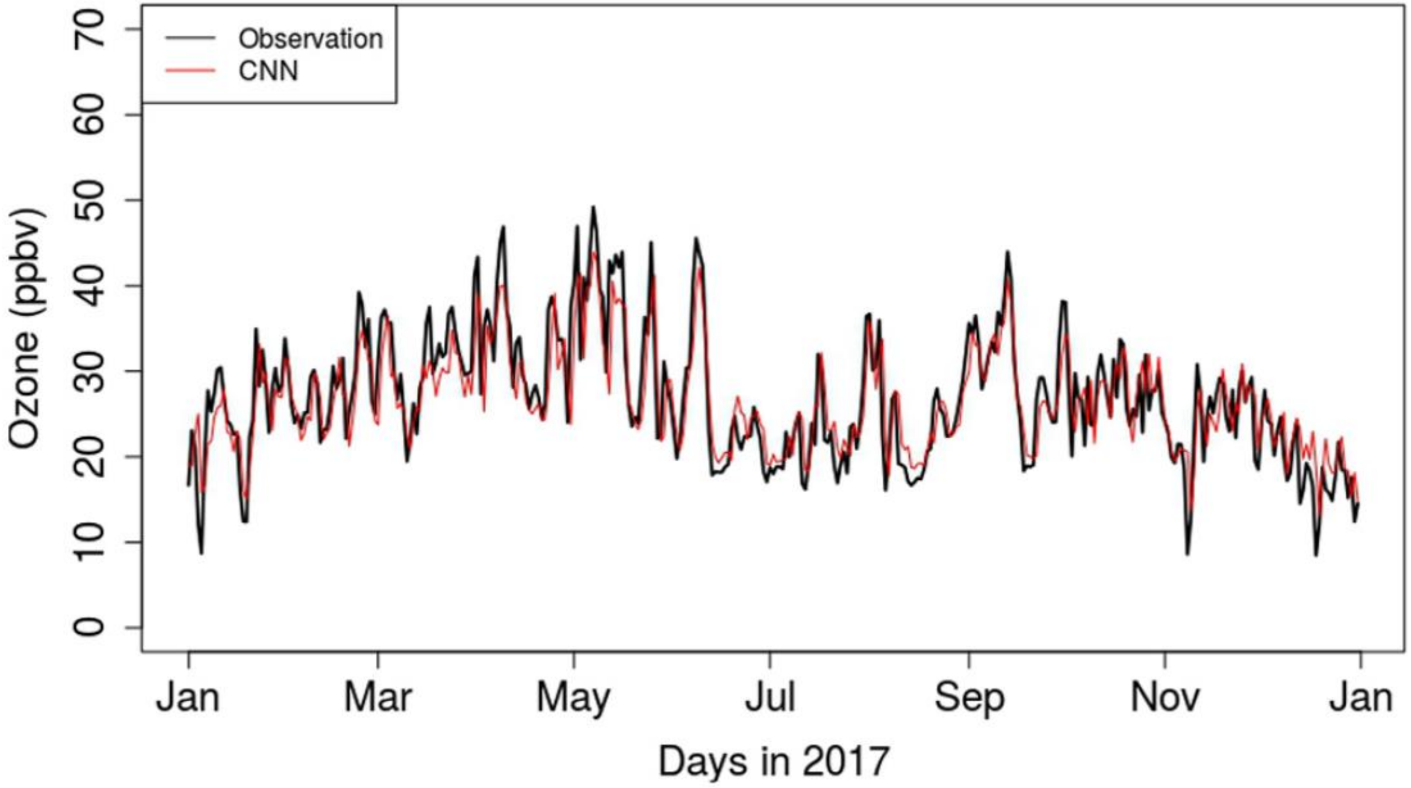


Fig. 2. Model Architecture: Detailed process flow of the deep CNN model.

https://doi.org/10.1016/j.neunet.2019.09.033

Sayeed et al., 2020



https://doi.org/10.1016/j.neunet.2019.09.033

Sayeed et al., 2020

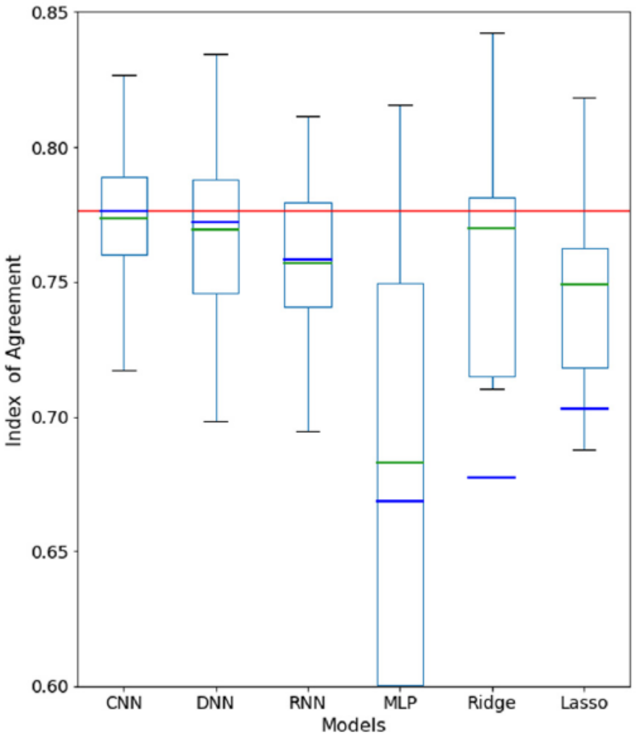


Fig. 12. Box and whisker plot for the index of agreement based on daily maximum ozone concentration of all stations. *Y*-axis represents IOA, and *X*-axis represents the models in the study.

https://doi.org/10.1016/j.neunet.2019.09.033

Conclusions

- The field of ML in weather and climate science is exploding
- NN approaches can be applied at all stages of ESM workflow
- Some recent studies show results between „promising“ and „astonishing“
- NN solutions are faster by factors of 10 or more
- Knowledge of Earth system science and statistics helps
- Be aware of (many) poor studies with improper set-ups or poor validation
- So far, vast majority of ESS studies employed supervised learning