Bias correcting weather models with machine learning

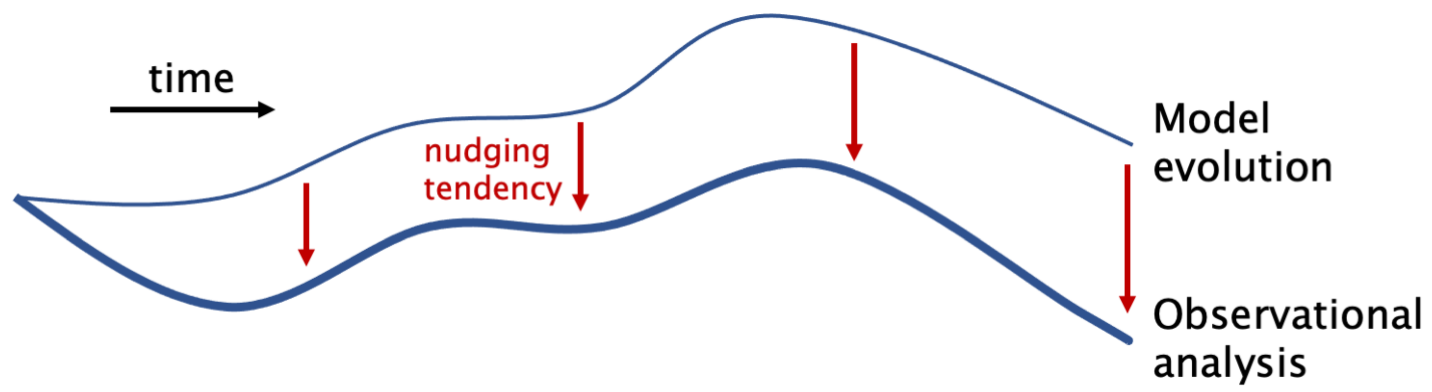
Oliver Watt-Meyer[1], Noah Brenowitz[1], Chris Bretherton[1], Spencer Clark[1,2], Lucas Harris[2], Brian Henn[1], Anna Kwa[1], Jeremy McGibbon[1], Andre Perkins[1]

[1]Vulcan Inc., [2]NOAA GFDL
contact: oliwm@vulcan.com

Methods

1. Perform a hindcast GCM run with a linear relaxation "nudging" term added to prognostic equations:

$$\frac{\partial x}{\partial t} = \dots - \frac{x - x_{obs}}{\tau}$$



2. Train an ML model to predict the tendencies $-\frac{x - x_{obs}}{\tau}$ from above run given only the model state.
3. Now make a forecast using same model as step 1 and at each timestep apply corrective tendencies predicted by the ML algorithm trained in step 2.

Atmospheric model details

- FV3GFS at C48 (~200km) resolution (https://github.com/NOAA-EMC/fv3atm)
- Nudging temperature, specific humidity, surface pressure and horizontal winds
- 6-hour nudging timescale τ
- Reference dataset for nudging is GFS analysis at approximately 1° resolution
- Two-year nudged simulation initialized on 1 January 2015
- Training data comes from 2015 year; 2016 is used for testing and verification

Machine learning details:

- Random forest (RF) used to predict nudging tendencies
- A single RF is trained to predict a column of tendencies given column profiles of the model state
- Inputs:
  temperature (T) , specific humidity (q), eastward wind (u), northward wind (v), land-sea mask, surface geopotential, cosine of zenith angle
- Outputs:
  temperature nudging tendency, specific humidity nudging tendency, eastward wind nudging tendency, northward wind nudging tendency
- Training data is 160 timesteps worth of column profiles; approx 2.2 million samples. 16 individual trees, with max depth of 13

Coupling of machine learning and GCM:

- We use a python wrapper of the Fortran FV3GFS code to allow calling python code during model simulation
- At end of each timestep, RF makes prediction of tendencies given current state of model and these tendencies are applied to the state
- The column-integrated drying induced by the nudging or machine learning is assumed to be converted to rainfall and is added to the surface precipitation rate felt by the land-surface model
- A limiter is applied to the specific humidity tendencies predicted by the RF so that the resulting specific humidity is non-negative

References

Huffman, G. J., R. F. Adler, M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, J. Susskind, 2001: Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations. J. Hydrometeor., 2(1), 36-50.

Key points

- A random forest (RF) can make skillful predictions of the nudging tendencies from a hindcast GCM simulation nudged toward observational analysis.
- Coupling a conventional GCM to this RF results in marked improvements in 1- to 10-day forecasts of Z500 and other variables.
- The root mean squared error of annual-mean precipitation is reduced by 24% in the RF-assisted GCM.

Nudging tendencies, offline predictions

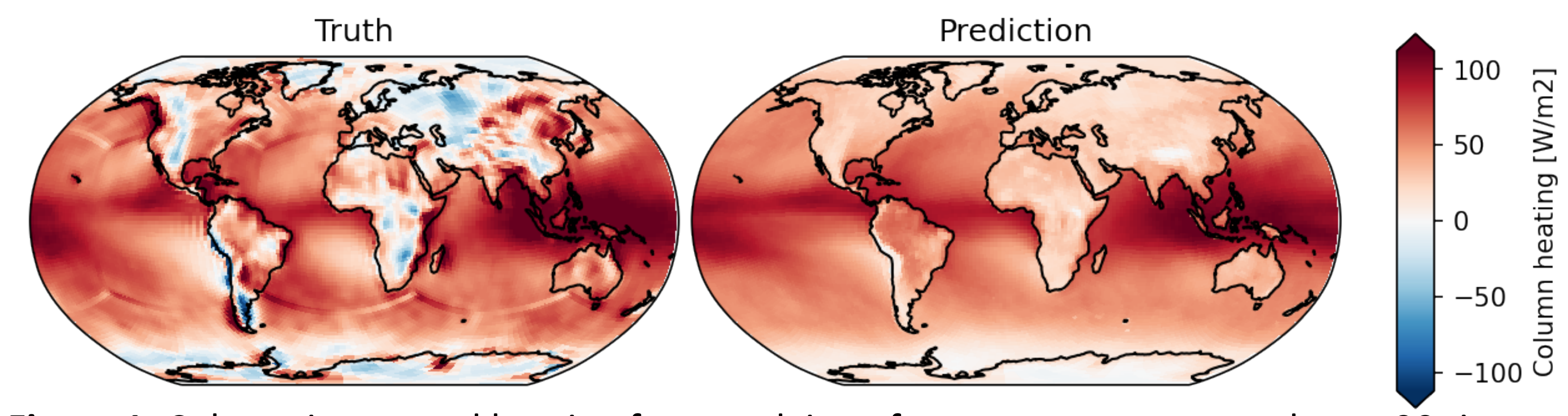


Figure 1: Column integrated heating from nudging of temperature, averaged over 90 timesteps of test data. Left: truth from nudged simulation, right: random forest prediction.

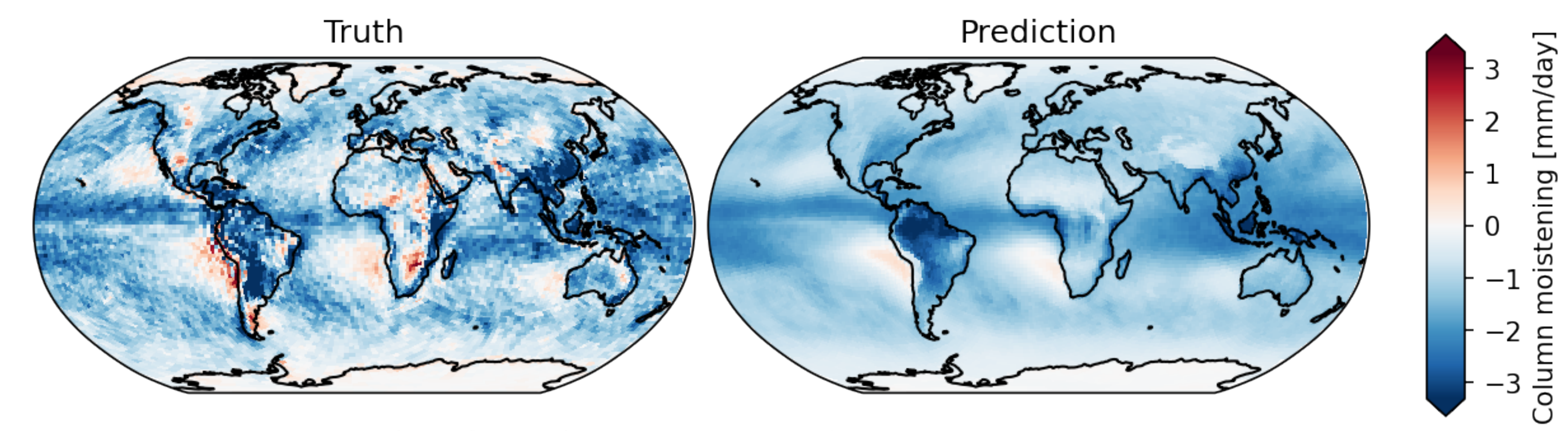


Figure 2: As in Figure 1 but for column integrated moistening.

- Nudging tends to heat and dry the atmosphere, especially in regions of convection (e.g. ITCZ, warm pool, midlatitude fronts)
- Random forest predictions do a good job of capturing the mean spatial pattern of nudging tendencies

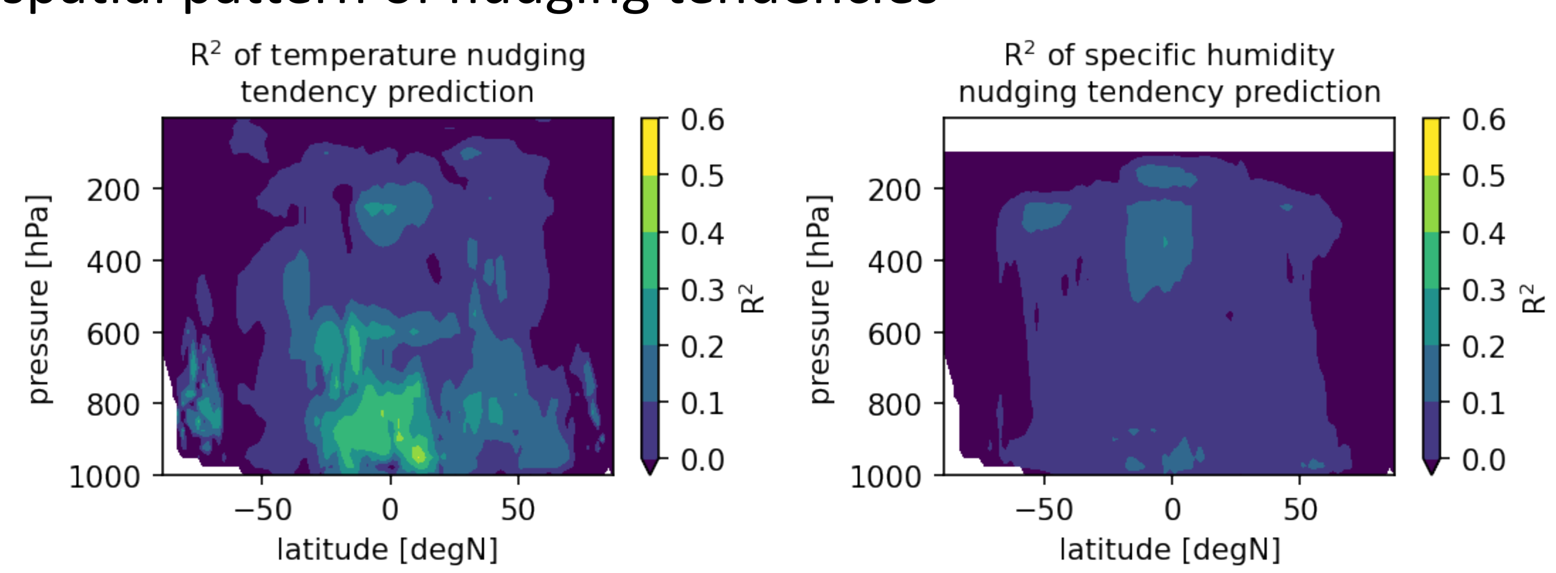


Figure 3: Offline R^2 skill for prediction of nudging tendencies. Up to 50% for temperature, less for moisture. But since we are learning a corrective tendency, we don't necessarily need high offline skill to get a benefit online.

Improved weather and climate forecasts

baseline: standard C48 FV3GFS model, **rf-control**: C48 FV3GFS model with random forest predictions of q, T, u and v tendencies at each timestep

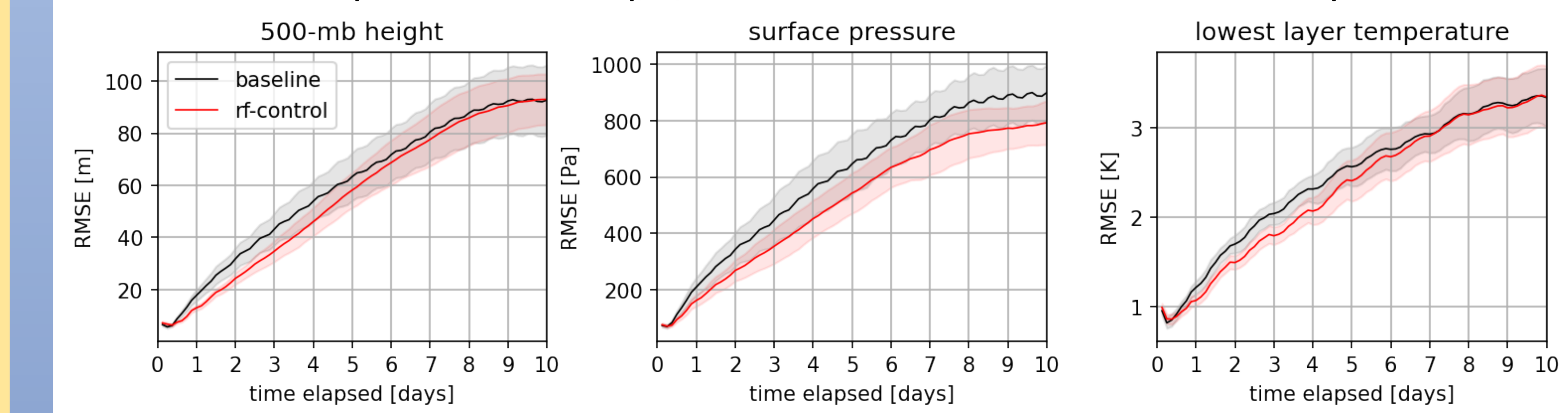


Figure 4: RMS error averaged across twelve forecasts initialized on the 1st of each month of 2016. Shading shows ± one s.d. Forecasts are verified against the simulation that was nudged towards GFS analysis.

- Model with online RF correction improves lead-time of Z500 and surface pressure forecasts by about one day and lowest model layer temperature by about half a day

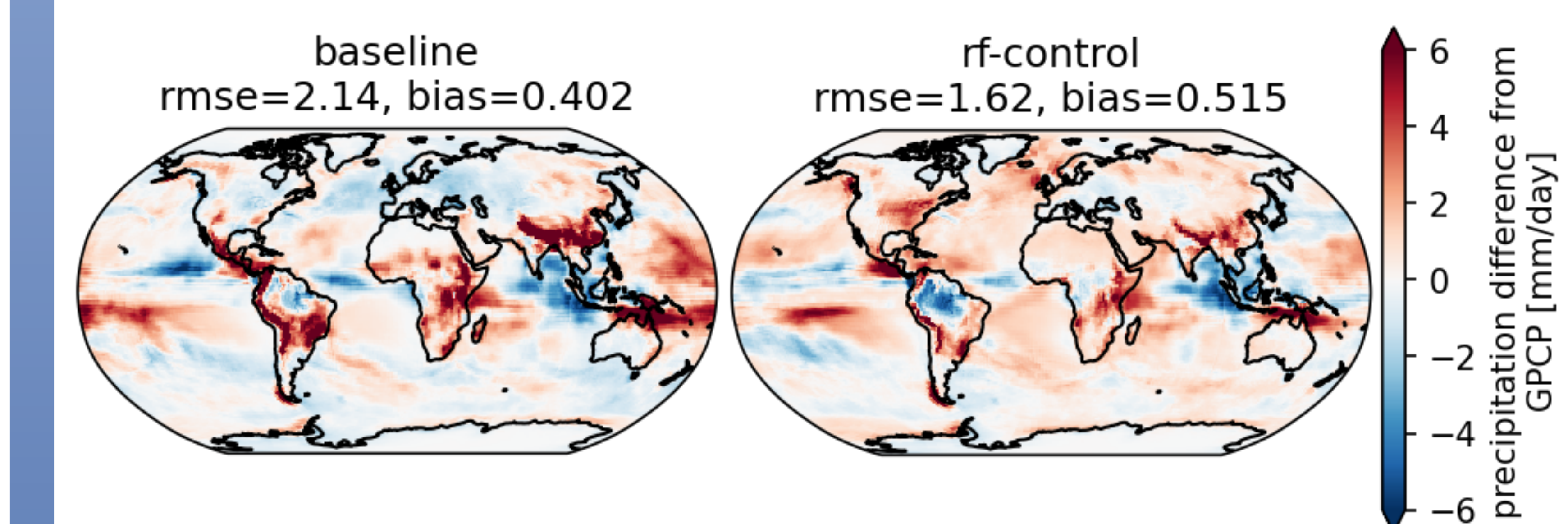


Figure 5: Precipitation bias relative to GPCPv1.3 product [Huffman et al., 2001] averaged over 2016. Titles show global mean RMSE and bias in mm/day.

- RMSE of time-mean precipitation is reduced by 24% in **rf-control** run compared to **baseline**
- The large positive bias of precipitation over Himalaya and Andes in **baseline** run is significantly smaller in **rf-control** run

Future work

- How much of an improvement does this method provide if the baseline model is higher resolution than 200km?
- Some temperature/moisture biases emerge in coupled GCM-RF simulations on monthly timescales, particularly in polar regions. How to avoid this?
- Use a neural network? RFs require significant memory which is a drawback. Have successfully trained a NN which produces stable simulations if it predicts heating/moistening. Instabilities if it predicts wind accelerations.
- How best to account for ML moistening tendencies in terms of soil moisture changes? Small land surface biases can emerge with current setup.