Development of stochastic/ensemble neural network convection parameterizations for climate models using CRM simulated data

Vladimir Krasnopolsky, NOAA/NCEP and University of Maryland (UMD), Michael Fox-Rabinovitz, UMD, Philip Rasch DOE PNNL, Yefim Kogan, University of Oklahoma/UCSD-Scripps, Alexei Belochitski, UMD, and Peter Blossey, UW
Acknowledgements: Marat Khairoutdinov (SUNY – Stony Brook) for providing SAM

A novel approach based on the neural network (NN) technique has been formulated and used for development of NN stochastic/ensemble convection parameterizations for climate and NWP models. This fast NN convection parameterization is built based on direct learning cloud physics from CRM (Cloud Resolving Model)/SAM (System for Atmospheric Modeling, Khairoutdinov and Randall, 2003) simulated data. SAM/CRM simulations have been initialized with and driven/forced by 120-day long TOGA-COARE data. SAM/CRM simulated data have been averaged to produce hourly and horizontally, 256 km x 256 km, means. The data was projected onto a GCM space of atmospheric states to implicitly define a stochastic convection parameterization. Here NN serves as an interface transferring information about sub-grid scale processes from fine-scale data or models (SAM/CRM) into GCM (up-scaling). Actually, the averaged SAM output is “projected” on the CAM space or in other words, only a subset of relevant SAM variables available in a climate model (NCAR CAM) is selected, resulting in creating an NN training data set. The developed NNs have been trained and tested (i.e., their accuracy is estimated vs. SAM simulated data). The first 96 days of SAM simulated data are used for NN training, and the last 24 days for validation/testing, i.e., testing is done on an independent data set.

Different developed NN architectures use temperature and moisture vertical profiles as inputs and Q1C – the “apparent heat source”, Q2 – the “apparent moist sink” and cloudiness (CLD) vertical profiles as well as total precipitation as outputs. Some of the NN architectures also use vertical velocity and radiative heating rates profiles as additional inputs. The stochastic NN convection parameterization is defined by an ensemble of NNs with different architectures. The inherent uncertainty of the stochastic convection parameterization is indicated and estimated.

The accuracy of NN convection parameterization or similarity of time series of cloudiness and precipitation to SAM simulated data, is illustrated in Figs. 1 and 2.

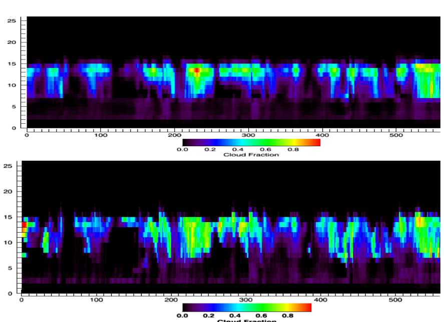

Fig. 1 Hovmöller diagrams for CLD profile (in fractions) time series; top - independent test data, bottom – NN convection.

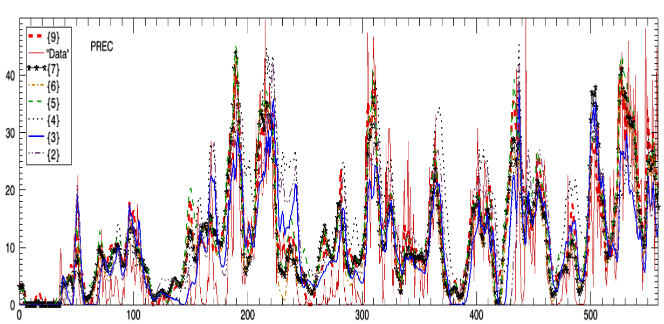

Fig. 2 Precipitation (in mm/day) time series for different NN architectures ({2}-{9}) vs. independent test data (“Data”).

An initial validation of NN convection in NCAR CAM has been done in an off-line/diagnostic mode. Actually, CAM inputs have been used, at every time step of the control/original CAM integration, for parallel (off-line) calculations of the NN convection parameterization (CAM-NN) to produce its outputs as a diagnostic byproduct.

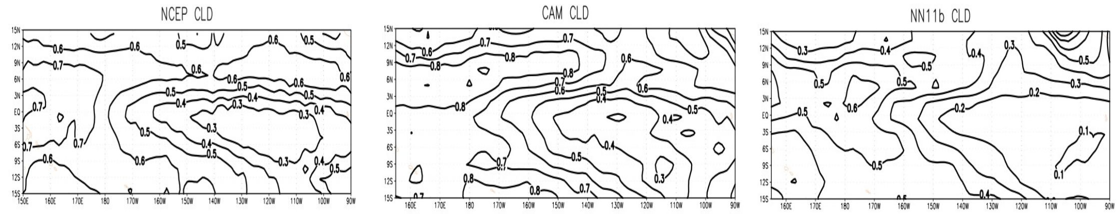

Fig. 3 Total CLD (in fractions, the contour interval is 0.1) for the Eastern Tropical Pacific Ocean (15 S to 15 N, 150 E to 90 W) for the 4-month TOGA-COARE period (Nov.-92 – Feb.-93) for the: NCEP reanalysis (left), control CAM (middle), and CAM-NN (right).

The CLD patterns shown in Fig. 3 are similar; their further in-depth analysis is needed. Note that the time series of precipitation and CLD for the CAM-NN- and control CAM runs (not shown) are similar in terms of both magnitude and frequency.

Conclusions

The presented initial results show the possibility of development of NN convection parameterizations based on learning cloud physics from CRM/SAM simulated data.

1. A novel approach based on using NNs is formulated and used for development of NN ensemble/stochastic convection parameterizations for climate models.
2. SAM/CRM simulations initialized with and driven/forced by TOGA-COARE data have been temporally and horizontally averaged and projected onto the GCM space of atmospheric states and used to derive very fast NN convection parameterizations with different architectures, and their accuracy is estimated.
3. Developed NN convection parameterizations have been validated in an off-line/diagnostic CAM (CAM-NN) run vs. the control CAM run. The initial results are encouraging: Total precipitation and cloudiness time series and tropical distributions for CAM-NN and CAM are realistic and consistent.
4. The NN convection parameterization can be applied to NWP models as well.

Future plans

1. Using long (3-year) ARM SGP data for driving SAM/CRM simulations for further developing and testing/validating NN convection parameterizations for mid-latitudes.
2. Using SAM/CRM simulations driven by CAM forcing for longer times, more geographic locations, and more diverse weather conditions so that NNs can be used globally and for all seasons, and validation of NN convection in CAM.

References

Khairoutdinov, M., and D. Randall, 2003: Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *J. Atmos. Sci.,* **60,** 607–625.

Krasnopolsky, V., M. Fox-Rabinovitz, Y. Kogan, A. Belochitski, and P. Blossey, 2010: Development of neural network convection parameterizations for climate models using Cloud Resolving Model simulations, NCEP Tech. Note, in preparation.

Krasnopolsky, V., M. Fox-Rabinovitz, P. Rasch, Y. Kogan, A. Belochitski, and P. Blossey, 2010: Neural network convection parameterizations for climate models based on statistical learning from Cloud Resolving Model simulations, paper in preparation.