



HAL
open science

Comments on “New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Longwave Radiation in a Climate Model”

Frederic Chevallier

► **To cite this version:**

Frederic Chevallier. Comments on “New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Longwave Radiation in a Climate Model”. *Monthly Weather Review*, 2005, 133 (12), pp.3721-3723. 10.1175/MWR3078.1 . hal-02957705

HAL Id: hal-02957705

<https://hal.science/hal-02957705>

Submitted on 13 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NOTES AND CORRESPONDENCE

Comments on “New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Longwave Radiation in a Climate Model”

FRÉDÉRIC CHEVALLIER

Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace, Gif-sur-Yvette, France

(Manuscript received and in final form 15 March 2005)

1. Introduction

The dramatic increase in computer power over the last decades has not fulfilled the ever-increasing needs of the climate and weather sciences. Improving the efficiency of numerical models is still topical. In their paper “New approach to calculation of atmospheric model physics: Accurate and fast neural network emulation of longwave radiation in a climate model,” Krasnopolsky et al. (2005, hereafter KFC05) develop the idea that artificial neural networks could accelerate model physics components in an atmospheric general circulation model (AGCM). As a first step, they apply it to the computation of longwave (LW) cooling rates and fluxes, which usually represent the main computational burden in an AGCM. KFC05 quote the studies previously performed at Laboratoire de Météorologie Dynamique and at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Chéruy et al. 1996; Chevallier et al. 1998, 2000b; Chevallier and Mahfouf 2001). While referring to these earlier works is fair, KFC05 hardly discuss how their “new” method differs from the previous one, nor how the conclusions disagree. I wish to make such a discussion in the present note, as a complement to the KFC05 paper. The following sections successively tackle the method and the prospects.

2. Method

An AGCM LW radiation model basically computes LW surface net fluxes F_s and vertical profiles of LW

cooling rates C_r . Fluxes and cooling rates are functions of surface variables \mathbf{S} , of the profiles of atmospheric temperature \mathbf{T} , and of atmospheric compounds \mathbf{A} ; \mathbf{T} and \mathbf{A} depend on the pressure discretization \mathbf{P} . Cooling rates in the vertical are proportional to the partial derivative of net fluxes as a function of atmospheric pressure:

$$C_r(P) \sim \frac{\partial F(P)}{\partial P}. \quad (1)$$

Among the variables representing the atmospheric compounds \mathbf{A} , one can make the distinction between cloud variables \mathbf{C} and the other ones \mathbf{V} , which mainly describe gas profiles, the former having usually more subgrid-scale variability than the latter. In the approach of Washington and Williamson (1977), used in several LW radiation codes (e.g., Morcrette 1991), a separation of variables is assumed between \mathbf{C} and \mathbf{V} in the expression of the fluxes:

$$F(\mathbf{S}, \mathbf{T}, \mathbf{V}, \mathbf{C}) = \sum_i a_i(\mathbf{C})F_i(\mathbf{S}, \mathbf{T}, \mathbf{V}), \quad (2)$$

where the sum is over the discretized atmospheric layers between the surface and the top of the atmosphere.

Chevallier et al. (1998, 2000b) introduced neural networks N_i to replace (or “emulate” in the words of KFC05) the F_i 's and call the result “NeuroFlux”:

$$F(\mathbf{S}, \mathbf{T}, \mathbf{V}, \mathbf{C}) = \sum_i a_i(\mathbf{C})N_i(\mathbf{S}, \mathbf{T}, \mathbf{V}). \quad (3)$$

KFC05 suggest to replace (or emulate) the whole function $C_r(\mathbf{S}, \mathbf{T}, \mathbf{V}, \mathbf{C})$ by a single neural network N :

$$C_r(\mathbf{S}, \mathbf{T}, \mathbf{V}, \mathbf{C}) = N(\mathbf{S}, \mathbf{T}, \mathbf{V}, \mathbf{C}). \quad (4)$$

Equation (3) allows one to speed up the radiation computation by sevenfold (Chevallier et al. 2000b). KFC05

Corresponding author address: F. Chevallier, Laboratoire des Sciences du Climat et de l'Environnement, Bat 701, L'Orme des Merisiers, 91191 Gif sur Yvette Cedex, France.
E-mail: frederic.chevallier@cea.fr

take advantage of the singleness of the neural network involved and demonstrate 50 to 80 times faster computing times.

The claim of novelty made by KFC05 is mainly based on the inclusion of cloud variables \mathbf{C} in the neural network. The reader is free to approve it or not. One of the questions *I* wish to answer here is why my former co-authors and myself did choose the more complicated formulation (3) rather than the straightforward Eq. (4). The simulation of all-sky fluxes from the Morcrette (1991) radiation model with a single neural network actually showed poor accuracy (Chevallier 1994). My analysis of this preliminary result is twofold. First the performance of a neural network is tightly linked with the statistical properties of its training dataset. Ideally a neural network in an AGCM should process seldom input configurations and frequent ones with the same quality. Strategies have to be implemented to even out the probability distribution functions of the variables in the training dataset (Chevallier et al. 2000a), even though KFC05 do not mention this issue. In practice, since atmospheric variables are coupled to each other in a nonlinear way, a training dataset is a compromise, which is all the more difficult to find when the number of variables is increased. In this respect, the variable separation (3) is advantageous. A more fundamental problem may also explain the failure of that preliminary test. There always exists a neural network that can simulate any continuous function over a given compact interval at any desired accuracy (Cybenko 1989). Now, subgrid-scale cloudiness may not be properly treated by continuous functions. In the case of the maximum-random overlap scheme Morcrette (1991), a series of condition tests is used, the simulation of which may be poor with neural networks. Consequently, the success of KFC05 in dealing with the problem may be due to a better neural network technique (type of network, choice of architecture, quality of the training dataset, implementation skill) than the previous studies, and/or a more continuous radiation scheme to simulate. KFC05 do not give any clue about this matter. Note that the latter reason would considerably limit the prospects of the method.

3. Prospects

In addition to the discontinuities in the modeling of subgrid-scale processes, important issues have raised skepticism about the use of neural networks for the modeling of the atmosphere. It is important to recall them in order to define the range of possible neural network applications.

The validity domain of a trained neural network is

usually the first topic to be raised. Faster parameterizations allow easier simulations of the atmosphere in particular and of the earth system in general over long time scales. Over such scales the climate may differ from the conditions of the present day. The corresponding evolution of the neural network errors is a matter of concern. Indeed the interpretation of a climate simulation requires the distinction of the real climate signals from the numerical artifacts.

The flexibility of a trained neural network is another major problem for further development of the approach. The development of observation campaigns and the inclusion of more processes in the models ever increase the list of inputs to an AGCM parameterization block. For LW radiation, one tends to make aerosols and minor gases interactive with the rest of the AGCM. The inclusion of new atmospheric compounds is all the more complicated now that the code is less physically (i.e., more statistically) based. As an example, to introduce a new aerosol type in the LW radiation computation, a simple extension of existing arrays is needed in the code of Morcrette (1991), whereas a redefinition of the training datasets and of the neural network architecture is required in the NeuroFlux approach. This argument also applies to any change in the vertical resolution. For instance the ECMWF vertical grid has been changed twice since 1999 and another refinement is being prepared.

Further to several studies that highlighted undesirable features in the neural network Jacobians (Aires et al. 1999; Chevallier and Mahfouf 2001), the accuracy of the derivatives usually comes third in the critics. Indeed sensitivity studies like those with ensemble simulations (e.g., Stainforth et al. 2005) form another major application for faster parameterizations. In the case of small perturbations, such studies critically depend on the accuracy of the AGCM Jacobians. Satisfactory accuracy for the neural network Jacobians can actually be obtained with larger neural networks (i.e., more neurons on the hidden layers), which complicates the training and penalizes the computing time. Demonstrating a useful trade-off is a subject for future research.

These issues are mentioned in KFC05, but they are not addressed. They show research directions where new developments are needed. In comparison, the question of which variables have to be inputs to the neural networks is of lesser importance.

4. Conclusions

There exist specific applications for which neural networks can usefully simulate environmental processes. For instance, the neural-network-based longwave ra-

diation scheme NeuroFlux is operationally used in the ECMWF data assimilation system, where Janisková et al. (2002) could demonstrate that its advantages exceed its drawbacks. Krasnopolsky et al. (2002) present another successful application, in the field of ocean modeling. However, the limitations that my former coauthors and myself found for AGCM modeling are still topical and indicate directions for future research. Like any other parameterization technique, the relevance of neural networks needs to be regularly reevaluated with respect to the particular computational and scientific contexts where they are developed and used.

Acknowledgments. I wish to thank all the people who exercised their neurons to develop the “NeuroFlux” approach with me, in particular A. Chédin, F. Chérury, L. Li, N. A. Scott at LMD, and M. Janisková, J.-F. Mahfouf, and J.-J. Morcrette at ECMWF. This note owes them a lot. I am also grateful to V. Krasnopolsky (NOAA) for interesting discussions about the topic. M. Janisková, J.-J. Morcrette, and V. Krasnopolsky helped to improve the initial version of this paper.

REFERENCES

- Aires, F., M. Schmitt, A. Chédin, and N. A. Scott, 1999: The “weight smoothing” regularization of MLP for Jacobian stabilization. *IEEE Trans. Neural Networks*, **10**, 1502–1510.
- Chérury, F., F. Chevallier, J.-J. Morcrette, N. A. Scott, and A. Chédin, 1996: Une méthode utilisant les techniques neuronales pour le calcul rapide de la distribution verticale du bilan radiatif thermique terrestre. *C. R. Acad. Sci. Paris*, **322**, 665–672.
- Chevallier, F., 1994: Une nouvelle approche, par réseau de neurones, de la modélisation du transfert radiatif à des fins climatiques. Rep. from DEA “Méthodes Physiques en Télédétection,” University of Paris, 45 pp.
- , and J.-F. Mahfouf, 2001: Evaluation of the Jacobians of infrared radiation models for variational data assimilation. *J. Appl. Meteor.*, **40**, 1445–1461.
- , F. Chérury, N. A. Scott, and A. Chédin, 1998: A neural network approach for a fast and accurate computation of longwave radiative budget. *J. Appl. Meteor.*, **37**, 1385–1397.
- , A. Chédin, F. Chérury, and J.-J. Morcrette, 2000a: TIGR-like atmospheric situation databases for accurate radiative flux computation. *Quart. J. Roy. Meteor. Soc.*, **126**, 777–785.
- , J.-J. Morcrette, F. Chérury, and N. A. Scott, 2000b: Use of a neural network-based longwave radiative transfer scheme in the ECMWF atmospheric model. *Quart. J. Roy. Meteor. Soc.*, **126**, 761–776.
- Cybenko, G., 1989: Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.*, **2**, 303–314.
- Janisková, M., J.-F. Mahfouf, J.-J. Morcrette, and F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation. *Quart. J. Roy. Meteor. Soc.*, **128**, 1505–1528.
- Krasnopolsky, V. M., D. V. Chalikov, and H. L. Tolman, 2002: A neural network technique to improve computational efficiency of numerical oceanic models. *Ocean Modell.*, **4**, 363–383.
- , M. S. Fox-Rabinovitz, and D. V. Chalikov, 2005: New approach to calculation of atmospheric model physics: Accurate and fast neural network emulation of long wave radiation in a climate model. *Mon. Wea. Rev.*, **133**, 1370–1383.
- Morcrette, J. J., 1991: Radiation and cloud radiative properties in the European Center for Medium Range Weather Forecasts forecasting system. *J. Geophys. Res.*, **96** (D5), 9121–9132.
- Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406.
- Washington, W. M., and D. L. Williamson, 1977: A description of the NCAR GCM’s in general circulation models of the atmosphere. *General Circulation Models of the Atmosphere*, Vol. 17, *Methods in Computational Physics*, J. Chang, Ed., Academic Press, 111–172.