Modeling clouds with probabilistic cellular automata

Jesse Dorrestijn[1], Daan Crommelin[1], Pier Siebesma[2,3], Harm Jonker[3], Joseph Biello[4], Steef Böing[3]

[1]CWI, Amsterdam, Netherlands. [2]KNMI, De Bilt, Netherlands. [3]Delft University of Technology, Delft, Netherlands. [4]Department of Mathematics, University of California, Davis, USA.

For climate and atmosphere models, the parameterization (or representation) of clouds and convection is of great importance. Climate and weather prediction models have too coarse resolutions to resolve these processes, therefore they must be represented in a simplified way. An emerging approach is the use of cellular automata (CA) to represent clouds, convection and other subgrid processes, thanks to their random character combined with the ability to self-organize into spatial structures. As clouds and convection interact with the atmospheric circulation, the CA are coupled to the partial differential equations (PDEs) that model the atmosperic flow. In a recent study, we used data of high-resolution convection simulations to construct probabilistic CA (PCA) emulating the development of convective clouds. Statistical inference for PCA is a key element in this approach.

1. Clouds & convection

An accurate representation (or "parameterization") of clouds in climate and atmosphere models is a major challenge that has to be tackled in order to improve weather prediction and climate models. The IPCC report of 2007 [1] states that "cloud feedbacks remain **the largest source of uncertainty** in climate sensitivity estimates". Convection, the vertical motion of air due to thermal effects, is a closely related process: clouds arise often in rising warm air. Therefore, there is a great need for new methods for representing clouds and convection. Because of the inherent uncertainty of these phenomena, stochastic methods for representing clouds and convection have become a hot research topic in recent years [2,3,4,5]. One promising approach is the use of PCA [6,7,8].

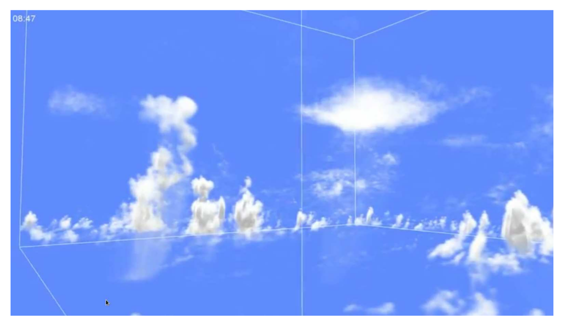

Figure 1 *Snapshot of a high-resolution model of clouds* *Credits: Jerome Schalkwijk*

2. A simplified representation of clouds

For a simplified representation of clouds in climate and global atmosphere models, one can classify clouds into several different **cloud types**. For global models, the local **area fractions** of these cloud types are needed. We write Ψ for the global atmospheric state (formed by variables as wind, temperature and humidity), and $\sigma_1, \sigma_2, \ldots, \sigma_M$ for the area fractions of the M cloud types under consideration. After discretizing Ψ on a 3D **macro-grid** spanned around (a part of) the world and the discretized PDE can be formulated as follows:

$$\frac{d\Psi}{dt} = F(\Psi, \sigma_1, \ldots, \sigma_M). \qquad (1)$$

We denote the **atmospheric state** of the ith vertical column of the macro-grid by Ψ_i, and the area fraction for cloud type m in the same column by σ_m^i.

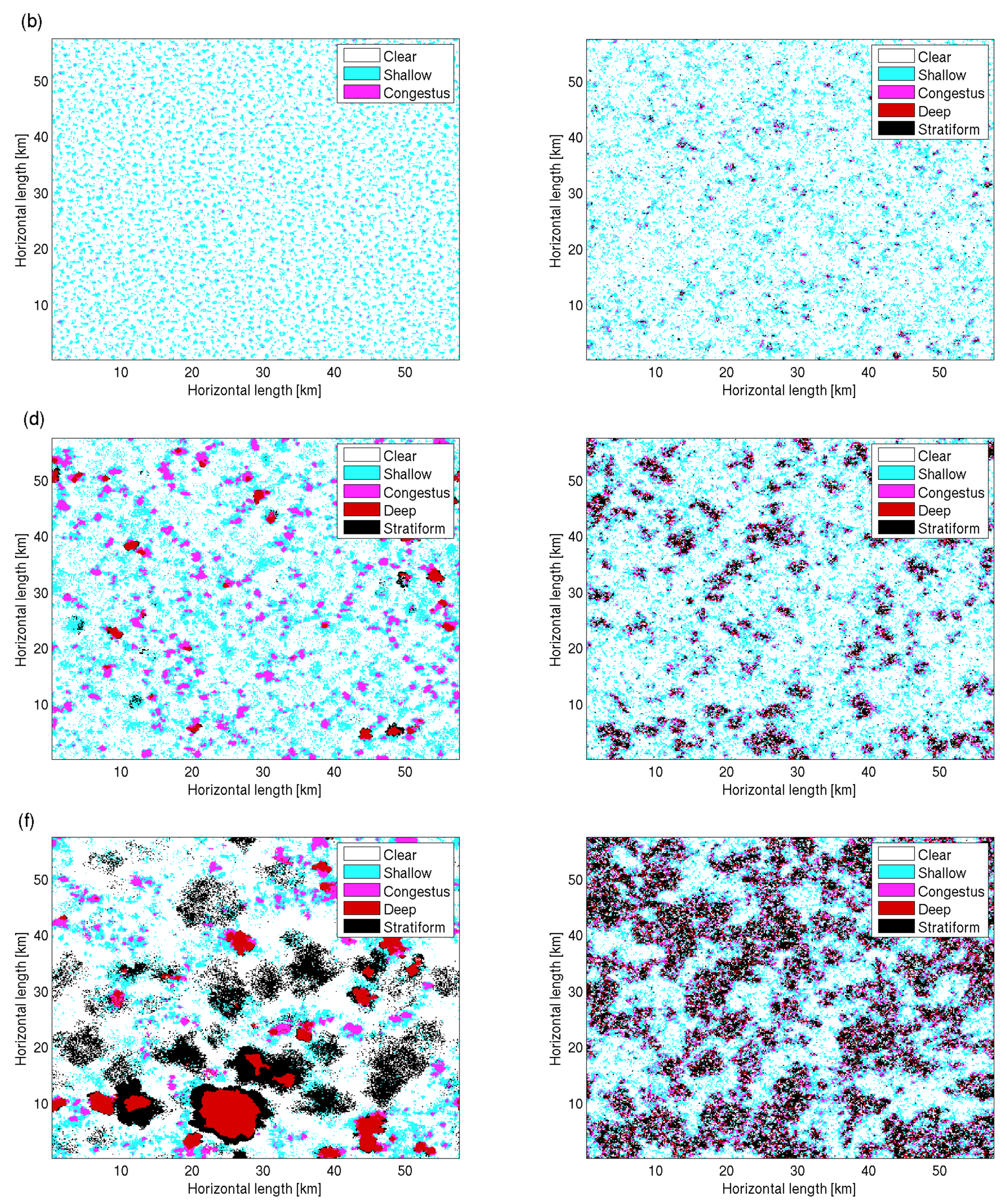

Figure 2 *(left) Snapshots of* ***high-resolution model*** *cloud type data and (right) reproduced PCA patterns*

3. Markov chains and PCA

We model transitions between cloud types by Markov chains. We associate a 2D **micro-grid** with the ith macro-grid column. The set of N Markov chains $Y_1^i(t), Y_2^i(t), \ldots, Y_N^i(t)$ living on this micro-grid determines the cloud type area fractions $\sigma_m^i(t)$:

$$\sigma_m^i(t) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}[Y_n^i(t) = m] \qquad (2)$$

at every time instance t. The area fractions calculated in (2) are used in (1). Conditioning the Markov chain transition probabilities on neighboring cell states results in a PCA, see Figure 3.

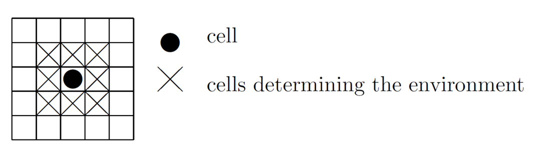

Figure 3 *Conditioning Markov chains on the neighboring cells gives PCA*

4. Interaction between PCA and PDE

The PCA on the 2D micro-grid determines the cloud type area fractions that are used in the discretized PDE (1). For a full coupling the time evolution of the σ_m^i should depend on Ψ_i. Thus, the rules of the corresponding PCA are made dependent on Ψ_i, and we consider the transition probabilities

$$\text{Prob}(Y_n^i(t+\Delta t) \,|\, Y_n^i(t), Y_{\{n\}}^i(t), \Psi_i(t)\,),$$

where $\{n\}$ denotes the neighborhood of the n-th cell of the micro-grid. In this manner, the PCA on the micro-grid interacts with the discretized PDE on the macro-grid.

5. Inference from high-resolution cloud data

We **infer PCA rules from data**. This can be high-resolution cloud model data (Figure 1) as well as high-resolution observational cloud data, e.g. radar data. We estimate the probability of a transition from state α to state β given the state of the neighboring cells $Y_{\{n\}} = \delta$ and the atmospheric state $\Psi_i = \gamma$

$$p_{\delta,\gamma}(\alpha,\beta) = \text{Prob}(Y_n^i(t+\Delta t) = \beta | Y_n^i(t) = \alpha, Y_{\{n\}}^i(t) = \delta, \Psi_i(t) = \gamma), \qquad (3)$$

using the estimator

$$\hat{p}_{\delta,\gamma}(\alpha,\beta) = \frac{T_{\delta,\gamma}(\alpha,\beta)}{\sum_\beta T_{\delta,\gamma}(\alpha,\beta)}$$

where

$$T_{\delta,\gamma}(\alpha,\beta) = \sum_{t,n} \mathbf{1}[Y_n^i(t+\Delta t) = \beta]\mathbf{1}[Y_n^i(t) = \alpha]\mathbf{1}[Y_{\{n\}}^i(t) = \delta]\mathbf{1}[\Psi_i(t) = \gamma].$$

We choose neigborhoods of 8 cells, giving M^8 possible neighboring cell configurations. Clearly, the number of configurations grows very rapidly as M increases. Therefore, we reduce the number of possibilities by using a reduction function g, and instead of (3) we use

$$p_{\delta,\gamma}(\alpha,\beta) = \text{Prob}\big(Y_n^i(t+\Delta t) = \beta | Y_n^i(t) = \alpha, g\big(Y_{\{n\}}^i(t)\big) = \delta, \Psi_i(t) = \gamma\big).$$

Choosing a good reduction function is difficult, and a systematic procedure to do so is currently lacking.

6. Two examples

In [7] we inferred a PCA from high-resolution **model data**. For classification, we used 5 different cloud types inspired by [4], distinguished by their cloud top height and rain water content. Thus, the PCA on the 2D mico-grid had $M = 5$ states, giving 5^8 neighborhood configurations. Using a suitable reduction function, we reduced the number of possibilities to 33 (i.e., $1 \leq \delta \leq 33$ in (4)). Furthermore, we considered 20 states for Ψ_i. In Figure 2 we depict cloud type patterns of the high-resolution model data (left) and PCA patterns (right). The corresponding cloud type area fractions can be found in Figure 5.

Currently we are working on the construction of a cloud type PCA using **observational data** from radar measurements. This PCA has $M = 3$ states. See Figure 4 for preliminary results.

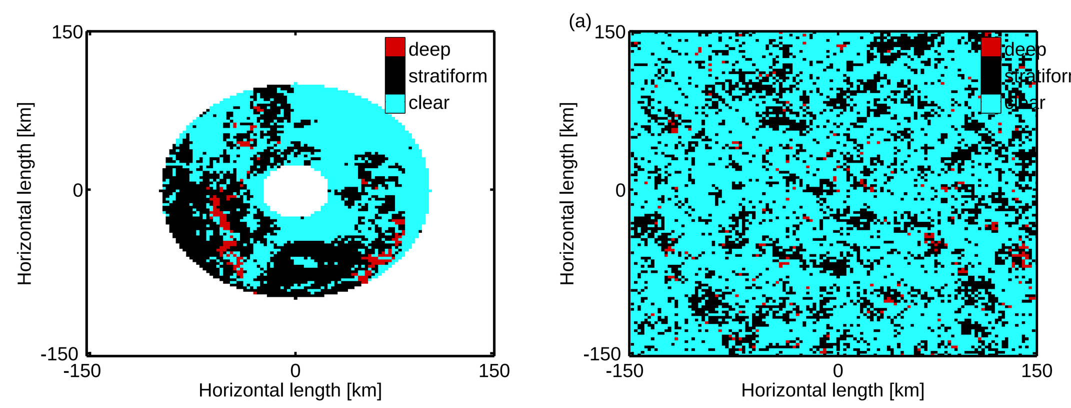

Figure 4 *(left) Snapshots of* ***observational*** *cloud type data and (right) reproduced PCA patterns*

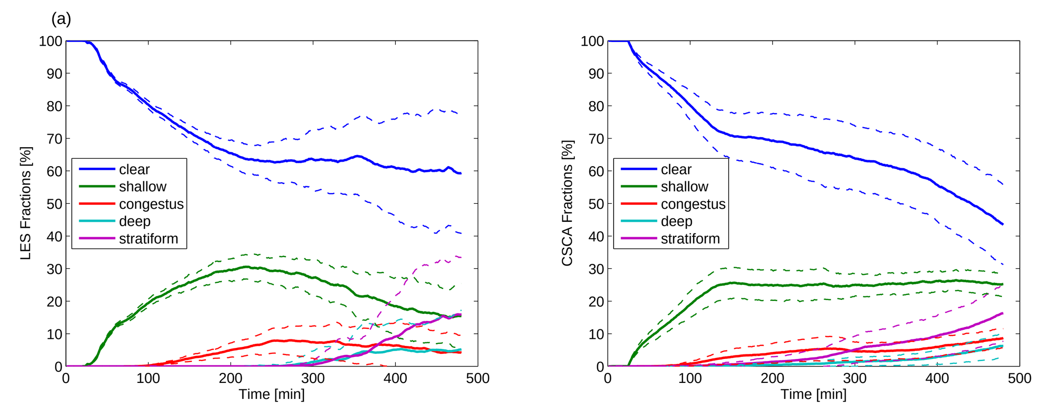

Figure 5 *Cloud type area fractions of high-resolution model data (left) and PCA (right)*

Conclusions

Modeling clouds and convection is a new and challenging application for PCA. A first step has been made by using statistically inferred PCA. Specific challenges lie in controlling the large number of rules for the PCA, the interaction between PCA and discretized PDE and the inference of rules from data.

Contact

J.Dorrestijn@cwi.nl Daan.Crommelin@cwi.nl siebesma@knmi.nl

H.J.J.Jonker@tudelft.nl joseph.biello@gmail.com s.j.boing@tudelft.nl

References

[1] Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report: Climate Change 2007, 8.6.3.2

[2] Palmer TN. 2001 A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Q.J.R. Meteorol. Soc.* **127**, 279-304

[3] Palmer TN. 2012 Towards the probabilistic Earth-system simulator: a vision for the future of climate and weather prediction. *Q.J.R. Meteorol. Soc.*,**138**: 841-861

[4] Khouider B, Biello J, Majda AJ. 2010 A stochastic multicloud model for tropical convection. *Commun. Math. Sci***8**, 187-216

[5] Dorrestijn J., Crommelin DT, Siebesma AP, Jonker HJJ. 2013 Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data. *Theor. Comput. Fluid Dyn.* **27**, 133-148

[6] Berner J, Doblas-Reyes FJ, Palmer TN, Shutts G, Weisheimer A. 2008 Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model. *Phil. Trans. Roy. Soc. A* **366**: 2559-2577

[7] Dorrestijn J, Crommelin DT, Biello, JA, Böing, SJ. 2013 A data-driven multicloud model for stochastic parameterization of deep convection. *Phil. Trans. Roy. Soc. A* **371**(1991):20120374

[8] Bengtsson L, körnich H, Källén E, Svensson E. 2011 Large-scale dynamical response to subgrid-scale organization provided by cellular automata, *J. Atmos. Sci.***68** 3132-3144

Acknowledgement

This research was supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organisation for Scientific Research (NWO). The visit of J.A.B. to CWI was financially supported through an NWO visitor grant. In addition, we acknowledge sponsoring by the National Computing Facilities Foundation (NCF) for the use of supercomputer facilities, with financial support of NWO. J.A.B. is supported by a grant from the National Science Foundation, DMS-1009959. The authors are grateful to Jessica Loriaux.