A data-driven stochastic parameterization of deep convection

Jesse Dorrestijn[1], Pier Siebesma[2,3], Harm Jonker[3], Christian Jakob[4], Daan Crommelin[1], Frank Selten[2]

[1]CWI, Amsterdam, Netherlands, [2]KNMI, De Bilt, Netherlands, [3]Delft University of Technology, Delft, Netherlands, [4]ARC Centre of Excellence for Climate System Science, Monash University, Melbourne, Australia

J.Dorrestijn@cwi.nl

Introduction

- Variability in rainfall is missing in climate models;
- Parameterizations of deep convection and clouds have to be improved;
- We present a stochastic parameterization inferred from observational data [1];
- High-resolution ($\sim$ 2.5×2.5 km^2) observational rainfall data is combined with large-scale ($\sim 150 \times 150$ km^2) data to construct a multicloud model [4,5];
- The multicloud model is based on conditional Markov chains (CMCs) [1-4] that switch between several cloud types;
- Cross-correlation analysis shows that the mean (large-scale) vertical velocity $\langle\omega\rangle$ is an important indicator of deep convection;
- Therefore, we condition the Markov chains on $\langle\omega\rangle$;
- After the training process the Markov chains form cloud type area fractions that are comparable to the observations of the test data set;
- We are currently testing the stochastic multicloud model in a General Circulation Model (GCM).

Markov chains

The multicloud model we use in this study consists of Markov chains positioned on the N nodes of a 2-dimensional micro-grid. The state of each Markov chain at time t is denoted by $Y_n(t)$. Each Y_n can take on 5 different values corresponding to the cloud types defined in Table 1. The Markov chains make a transition between the cloud types every Δt =10 minutes. N Markov chains together form cloud type area fractions σ_m for the various cloud types:

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

in which $\mathbf{1}$ is the indicator function ($\mathbf{1}[A] = 1$ if A is true, 0 otherwise), N is the number of micro-grid nodes, and $m \in \{1, \dots, 5\}$ the cloud type. We use radar data to estimate the transition probabilities, needed in the Markov chain model.

The high-resolution observational data

The high-resolution observational rainfall data is obtained from a rain radar in Darwin in North-Australia. For two time periods of 5 and 3 months (the training and test data set) we have integer valued cloud top height (Figure 1a) and rain rate observations at 10-minute time steps. Using the criteria of Table 1, the radar data is classified into *clear sky* and the cloud types *moderate congestus*, *strong congestus*, *deep convective clouds* and *stratiform clouds* (Figure 1b). The 5×5 transition matrix for the Markov chains is estimated by counting transitions observed in the high-resolution data set.

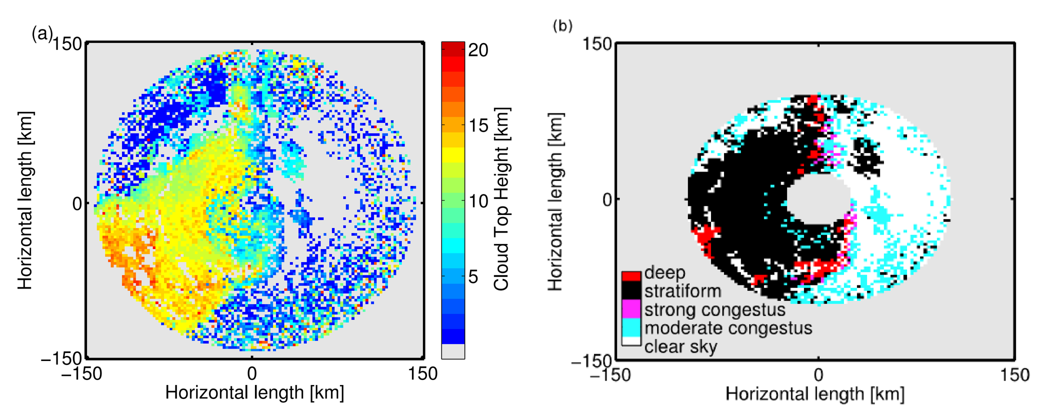

Figure 1: (a) A snapshot of the cloud top height derived from radar observations (b) the same data classified into several cloud types.

Cloud top height [km]	rain rate [mm h^{-1}]	
	≤ 12	> 12
≥ 6.5	stratiform ($m = 5$)	deep convective ($m = 4$)
	≤ 3	> 3
$\in [1.5, 6.5)$	moderate congestus ($m = 2$)	strong congestus ($m = 3$)
< 1.5	clear ($m = 1$)	

Table 1: Thresholds for the cloud top height and the rain rate are used to classify the radar data pixels into several cloud types.

The large-scale data

We have data available that defines the large-scale dynamical and thermodynamical state of the atmosphere around Darwin. The data has been prepared by [6] and are NWP analysis large-scale variable estimates improved by constraining moisture budgets with observational rain data from the Darwin rain radar. Given the timeseries of the deep convective area fraction $\sigma_4(t)$ and the timeseries of the large-scale variable $X(t)$, the normalized cross-correlation function (CCF) of $X(t)$ and $\sigma_4(t)$ is:

$$\mathrm{CCF}(\tau) = \int_{-\infty}^{\infty} \tilde{X}(t+\tau)\tilde{\sigma}_4(t)dt \tag{2}$$

with $\tilde{X}(t) = \frac{X(t)-\mu_X}{\sigma_X}$ (i.e. the large-scale variable normalized by subtracting its mean μ_X and dividing by its standard deviation σ_X), $\tilde{\sigma}_4$ defined analogously, and τ the time lag of X w.r.t. σ_4. We calculate the CCF for the convective available potential energy (CAPE), the mean vertical velocity $\langle\omega\rangle$ in hPa h^{-1} and for the relative humidity (RH) (Figure 2). We see that $\langle\omega\rangle$ gives the highest correlation and therefore we make the probabilities of the Markov chains dependent on $\langle\omega\rangle$, which results in a set of 5×5 transition matrices.

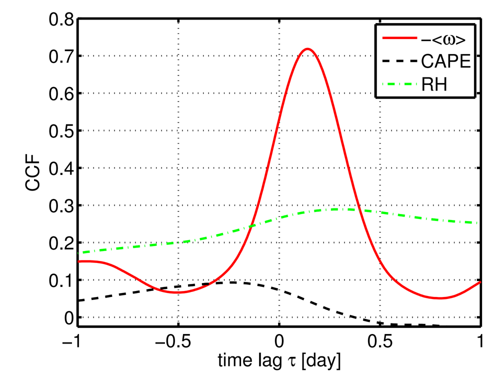

Figure 2: Cross-correlation between the deep convective area fraction σ_4 and the large-scale variables, $\langle\omega\rangle$, CAPE and relative humidity.

Results

After the training process we calculate cloud type area fractions with the multicloud model for the test data set and compare the fractions with the observations. In Figure 3-5 we see that the multicloud model is able to reproduce the observed fractions quite well. In Figure 3 adn 4 we choose $N = 100$ Markov chains to form the cloud type area fractions. The model can be adapted to the scale of the GCM grid box by choosing N larger for larger grid boxes. Further we chose 25 intervals for the large-scale variable $\langle\omega\rangle$.

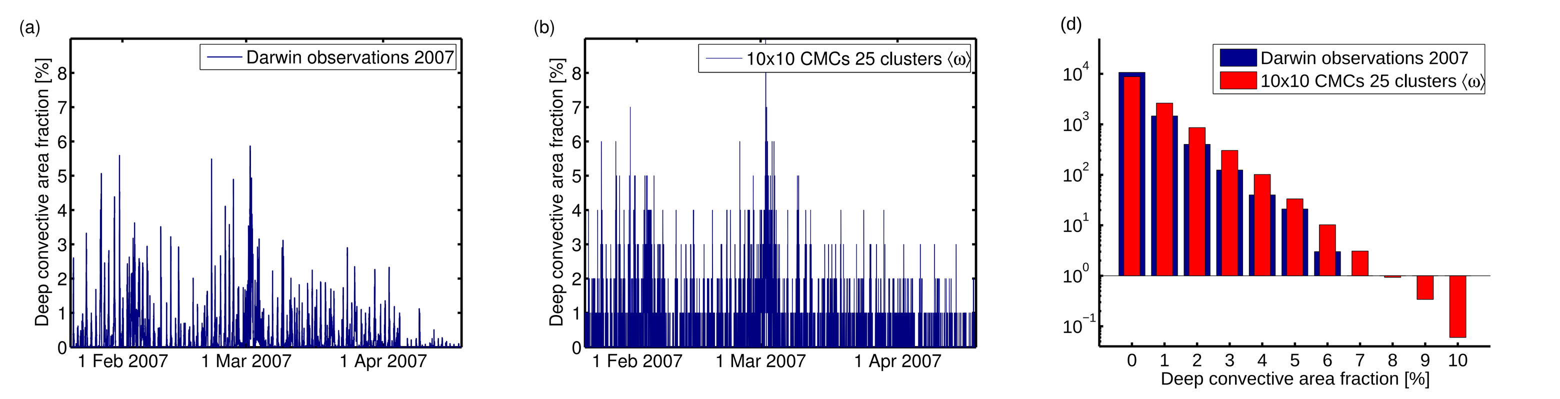

Figure 3: (a) Observed deep convective area fractions and (b) area fractions reproduced by the CMCs (d) histograms comparing the fractions.

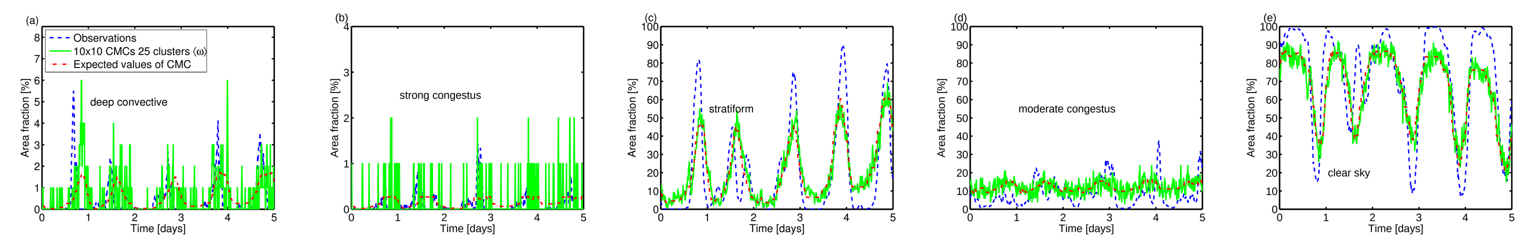

Figure 4: (a) Deep convective (b) strong congestus (c) stratiform (d) moderate congestus and (e) clear sky area fractions for a period of 5 days.

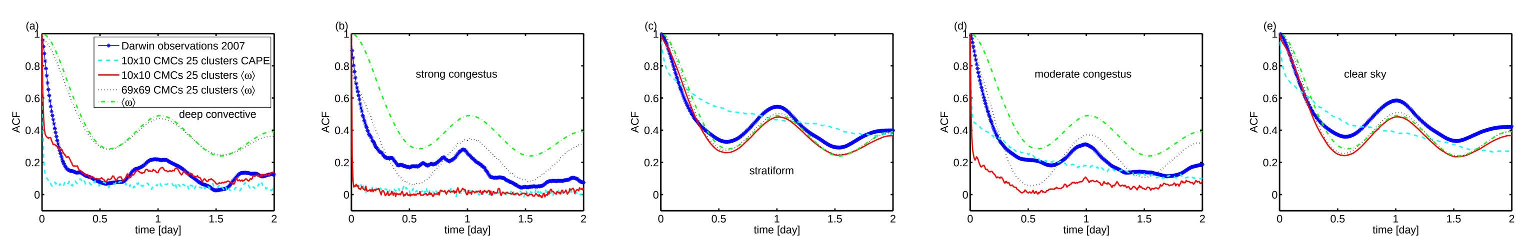

Figure 5: Autocorrelation functions (ACFs) for the (a) deep convective (b) strong congestus (c) stratiform (d) moderate congestus and (e) clear sky area fractions.

Current work: Implementation in a GCM

The multicloud model can easily be coupled to a GCM, for example by using the deep convective area fractions to serve as a closure for the mass flux at cloud base:

$$M_b = \rho\, \sigma_4\, w_{cb}$$

with ρ the density and w_{cb} the vertical velocity at cloud base in a deep convective updraft (which both can be chosen to be constants). We are currently testing the multicloud model by implementing it in "SPEEDY", a model solving the primitive equations with a set of simplified parameterizations for the most relevant physical processes only. We analyze the results by looking at Hovmöller diagrams of equatorial precipitation and zonal-wavenumber frequency diagrams.

Conclusions

- We have been able to construct a parameterization of cloud type area fractions entirely from observational data;
- The mean vertical velocity $\langle\omega\rangle$ was shown to be the best variable to condition on;
- The cloud type area fractions produced by the Markov chains are comparable to the observational fractions;
- The multicloud model can be used in the Grey zone and is scale-adaptive;

References

[1] J. Dorrestijn, D.T. Crommelin, A.P. Siebesma, H.J.J. Jonker, C. Jakob: Stochastic parameterization of convective area fractions with a multicloud model inferred from observational data, *Submitted to J. Atmos. Sci.*

[2] D. Crommelin and E. Vanden-Eijnden: Subgrid-Scale Parameterization with Conditional Markov Chains, *J. Atmos. Sci.*,2008,**65**,2661–2675

[3] J. Dorrestijn, D.T. Crommelin, A.P. Siebesma, H.J.J. Jonker: Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data, *Theor. Comput. Fluid Dyn.*,2013,**27**,133–148

[4] J. Dorrestijn, D.T. Crommelin, J.A. Biello, S.J. Böing: A data-driven multi-cloud model for stochastic parametrization of deep convection, *Phil. Trans. R. Soc. A.*, 2013,**371**,20120374

[5] B. Khouider, J. Biello, A.J.Majda: A Stochastic Multicloud Model for Tropical Convection, *Comm. Math. Sci.*, 2010,**8**,187–216

[6] L. Davies, C. Jakob, P. May, V.V. Kumar, S. Xie: Relationships between the large-scale atmosphere and the small-scale convective state for Darwin, Australia, *J. Geophys. Res. Atmos.*, 2013,**118**, 534,11–545

Acknowledgement

The project is funded by the NWO-program "Feedbacks in the Climate System".