Andrey Martynov

Institute of Geography

University of Bern

Hallerstrasse 12

3012 Bern Switzerland

Sensitivity of hailstone size on choice of microphysical schemes and on climate change in simulations of hailstorms over Switzerland in June-August 2012 with the WRF model

23rd of September, 2018

ABSTRACT

The hailstone sizes, simulated by the WRF model in current climate conditions and in conditions, imitating the climate changes expected towards the end of the 21st century, have been examined. The statistical significance of applying climate change conditions and of using a different microphysical scheme has been estimated in comparison with the natural internal variability of the simulated climate system. It has been found that the deviations of results, obtained by using the Thompson microphysics instead of the default Morrison scheme, from Morrison simulations, can be explained by the internal variability, while the climate change produces changes in hailstone size distribution, significantly different from the present climate results.

INTRODUCTION

Simulating hailstorms using high-resolution Numerical Weather Predictions model is frequently used for studying the conditions of initiation, evolution and disintegration of hailstorms [e.g. Mahoney 2012]. Model data can provide essential information, not available from observations, on the 3D structure of hailstorms, paths of growing hailstones within the convective cell, etc. One important variable that can be obtained from simulations is the size of hailstones on the ground. From one side, this value can directly be compared with the ground observations (if available), and from another side, in simulations imitating the climate change conditions it can give a hint to changes of hailstone sizes, hailstorm frequency, spatial and temporal distributions in the future.

There are however some important features of numerical weather simulation approach that should be taken into consideration.

1. The equations, describing the atmospheric dynamics and convection, ansd solved by the weather models, have chaotic solutions (the so-called Lorenz chaos), and are thus instable to small perturbations of initial or boundary conditions. This means that two simulations starting at close but not identical initial conditions will deviate with time, and at certain point will be producing very different solutions. This is also true for similar initial and boundary conditions, but for numerically different solvers (i.e. for different compilers or different CPU types).

This factor is known to climate scientists. It is considered that different realizations of weather in such divergent simulations represent the natural, intrinsic variability of the climate system, which is denoted as “internal variability”. Ensemble simulations of up to hundreds of members are normally run for weather prediction or for climate studies, allowing thus for estimating the internal variability of the studied system. It is usually considered that the values, obtained from such an ensemble run, are normally distributed, because they represent different realizations of the same physical system; average values and standard deviation represent the most probable state of the system and its uncertainty.

2. The simulation results are in general strongly dependent on the numerical models, in particular on the choice of parameterizations for physical processes, not explicitly resolved at the spatial and temporal resolution used for simulation. There are many such processes, the most important of which are the microphysics of clouds, radiative processes in clouds, processes in soil and vegetation, convection in the planetary boundary layer. For each of these and other processes numerous parameterization solutions have been developed, and several options exist for each of them. For example, in the Weather Research and Forecasting model (WRF) model [Skamarock et al. 2008], the total number of combinations of different physical parameters is of the order of two millions [http://www2.mmm.ucar.edu/wrf/users/wrfv4.0/testing.html]

In many cases there is no evident choice of one or other parameterization; therefore an usual approach consists of running simulations with several most relevant options and performing sensitivity tests. This requires additional resources such as CPU hours, storage space and working hours for treating and analyzing data.

OUR STUDY AND RESEARCH QUESTIONS

The study performed in our research group is centered at studying the changes in hailstorm characteristics that can be expected in the future under the climate change conditions. To do it, in the pilot study we simulate hailstorms over the central part of Europe (Figure 1) in the summer period of 2012, using the WRF model in the current climate conditions and in codnitions, imitating the climate change environment, expected towards the end of the XXI century.

The pilot study has been designed in accordance to the features described above:

3 current climate simulations have been performed, using the Morrison microphysical scheme [Morrison et al. 2009] (simulations 1-3)

1 current climate simulation has been performed, using the Thompson microphysical sceme [Thompson et al. 2008] (simulation 4)

1 climate change simulation has been performed, using the Morrison microphysical scheme (simulation 5)

The research questions that the current study seeks to answer are:

1. Does the choice of the microphysical scheme affect the simulated hail sizes?

This is essential for the desing of the next phase of the project: if the choice of microphysics does not significantly affect the hail size characteristics, there is no need in performing simulations with different microphysical options.

2. Does the climate change affect the simulated hail sizes?

This is the key question of the whole study.

METHODS

1. Reading and transforming arrays (1D to 2D)

2. Fitting source data to a distribution (scipy.stat)

3. Determining the p-values for a normal distribution (scipy.stat)

METADATA

The metadata for the simulation data are formed by:

- The model configuration files, technically allowing any qualified user of the WRF model to reproduce the entire simulations; these files are stored with the model output data.

- The metadata, contained in the model output, including the long and short names of variables, their units, and the history of manipulation with these files.

- The data treatment bash scripts with comments.

- The Jupyter notebook.

The complete data and metadata are currently only available to authorized uses of the server systems of the Institute of Geography. No journal publications based on these data has been made yet, thus no provisions have been made for making these data available to external users; eventually it will be done according to the applicable rules as part of the publication process.

DATA QUALITY

The source data are stored as 32-bit floating point values, which is sufficient for our analysis. From practical reasons, 2-3 significant digits for the hail size are sufficient – in particular because the hail size in ground reports rarely is given with more than 2 significant digits in cm or inches. The formal precision of simulated variables is largely exceeding these requirements.

DATA FLOW

The simulations in the present work have been performed, using the Weather Research and Forecasting model with the Advanced Research WRF dynamical core (WRF-ARW), version 3.6.1 at the 2.2 km horizontal resolution. The initial and boundary conditions for the WRF simulations were provided by the 6-hourly high-resolution operational model analysis of the European Centre for Medium-Range Weather Forecasts (ECMWF) at the 0.125° horizontal grid spacing. The simulation domain is presented on the Figure 1. The duration of all simulations (after a week-long spin-up period) was 92 days, of 2208 hours between 00Z, 1st of June and 00Z, 1st of September 2012.

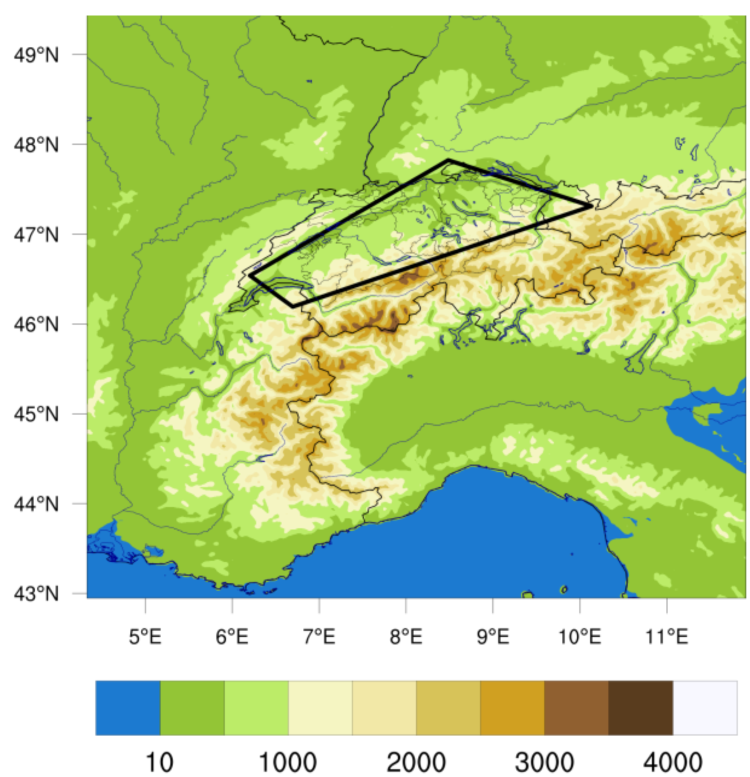


Figure 1. Simulation domain with orography (meters over sea level) , rivers, country borders (the outline area is not used in the current study)

All simulations were performed on the UBELIX cluster of the University of Bern.

The model produced hourly output files in the NetCDF format [https://en.wikipedia.org/wiki/NetCDF], which were transferred from UBELIX to the storage servers of the Institute of Geography. The total volume of data produced by one simulation is approx. 1.4 Tb.

For this study the output data were treated in the following way, using the Climate Data Operators tool [https://code.mpimet.mpg.de/projects/cdo/] and bash scripts:

- For each simulation hour, maximum values of all 2D variables were collected;

- Hourly maximum values were combined into a singe column, spanning all 145 variables and 2208 hours.

From this step on, all data treatment has been performed, using a Jupyter notebook.

- All five datasets were read and converted into 2D arrays.

For the presented study, the variable #139, AFWA_HAIL_NEWMEAN, representing the simulated hailstone size, has been selected (other variables have been used in different studies, in particular for the search of hail size proxies, presented in the CAS-M2 poster).

Each of the five hailstone size datasets has been fit, using the Generalized Extreme Value Distribution [https://en.wikipedia.org/wiki/Generalized_extreme_value_distribution]. This distribution can be applied to these data because they represent maximum values of spatial and time series of rare events.

For each dataset, the fitting procedure produced three parameters (shape, location and scale).

The fitting parameters of simulations 1-3 ("internal variability" runs) were used for determining parameters of the normal distribution, to which they best fit.

For simulations 3 (different microphysics) and 4 (climate change) a statistical test has been applied, allowing to determine whether the results of these simulations do also match the Poisson distribution of the internal variability runs. The confirmation of the null hypotheses (that is, simulations 4 and 5 are matching the distributions of simulations 1-3) would mean negative answers to the corresponding research questions.

DATASETS

The source data (simulation output sets) are stored on secure servers of the Institute of Geography, University of Bern. Servers are located in a physically protected area, online access to these systems is restricted to authorized users.

The source data are not confidential and are available to all users of these servers. They are frequently used by different user, for example, bachelor and master students.

The data treatments scripts and intermediate files are also located on these servers.

A copy of source data has been stored as a backup on a off-site server, belonging to another department of the University of Bern. Numerous copies of simulation configuration files exist, have been shared with colleagues and can be easily retrieved.

DATA MODELS

The logical data model is presented on Figure 2.

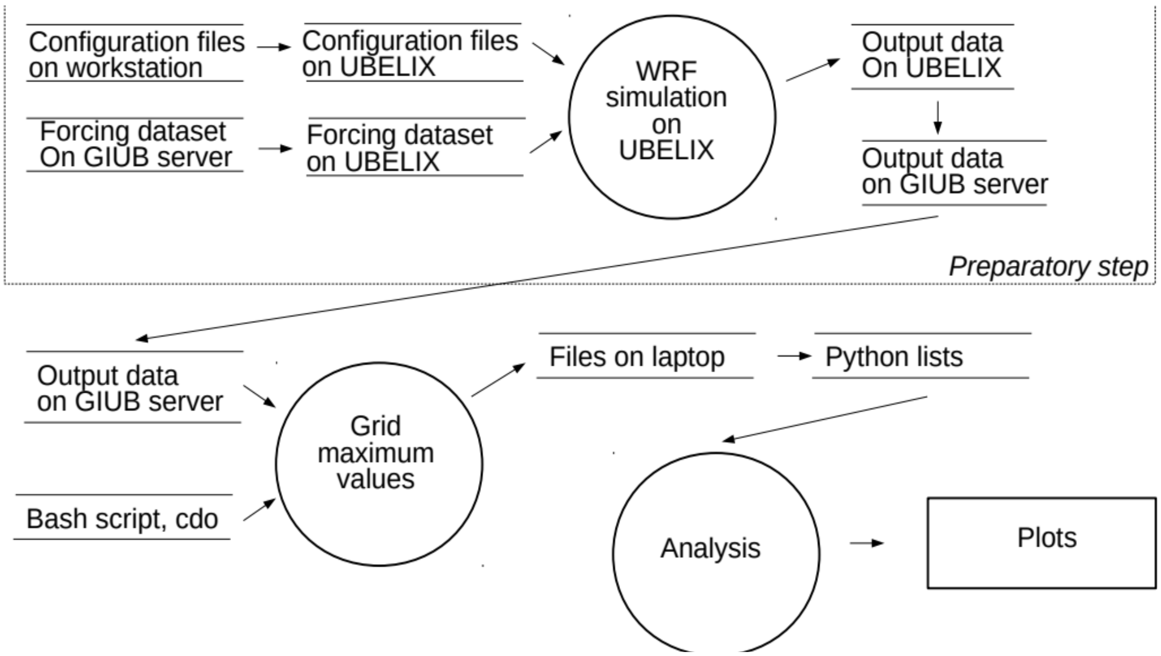


Figure 2. The logical data model of the presented analysis. The "preparatory step" has been carried out erlier, not as part of this project.

RISKS

The worst possible scenario which would occur with these data is a complete lost. However, this is a pilot project that has been used for preparing the next phase of simulations, which is about to start in the nearest weeks. If needed, these simulations can easily be repeated (possibly not numerically exactly, but within the internal variability framework) within 2-3 months.

If these data are lost, this will be of the least concern compared to losing other datasets stored on the same servers, including the backup server. Other risks are of the negligible importance.

ANALYSIS

The maximum grid hail size values are presented on the Figure 3 for all five simulations. Please note that a non-zero value signifies that a hailstorm has been simulated at some point over the large simulation domain (see Figure 1). The red line denotes the minimum hailstone size which is considered as severe (20 mm).

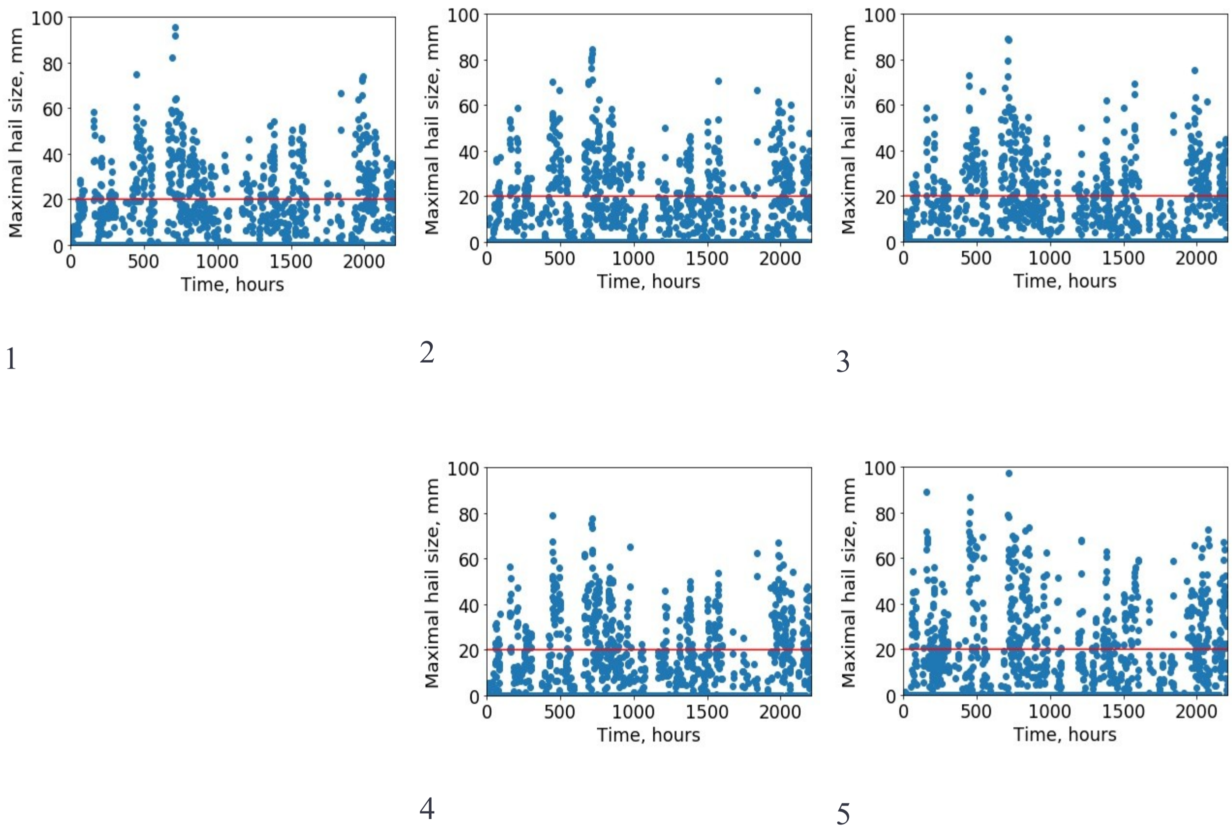


Figure 3. Maximal hail size over the simulation grid as function of time. Simulations 1-3 - "internal variability" runs, 4 – different microphysics, 5 – climate change conditions.

It can be seen that in approximately 1/3 of all simulation days a hailstorm has been simulated at least in one location over the domain.

For each simulation the distribution of hailstone size has been approximated using the Generalized Extreme Values Distribution (GEV).

The hail size histograms and corresponding cumulative distribution functions are presented on the Figure 4.

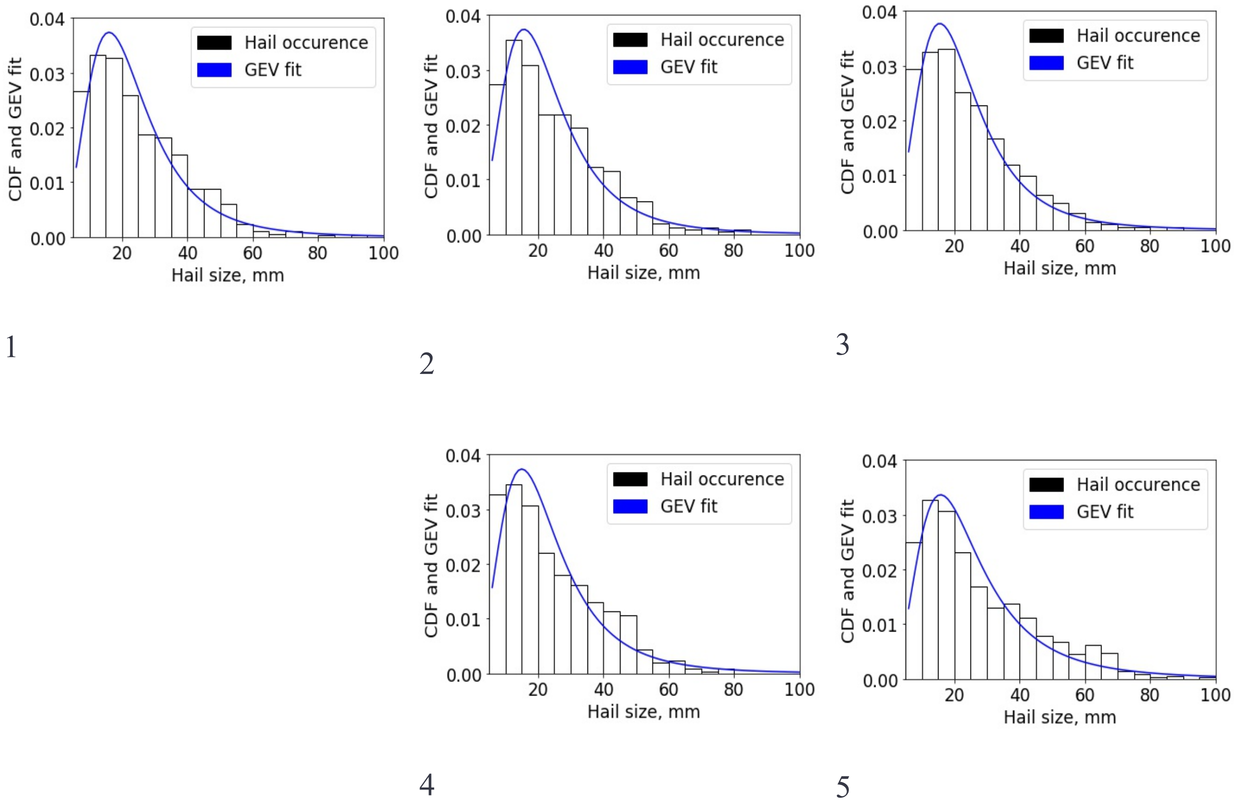
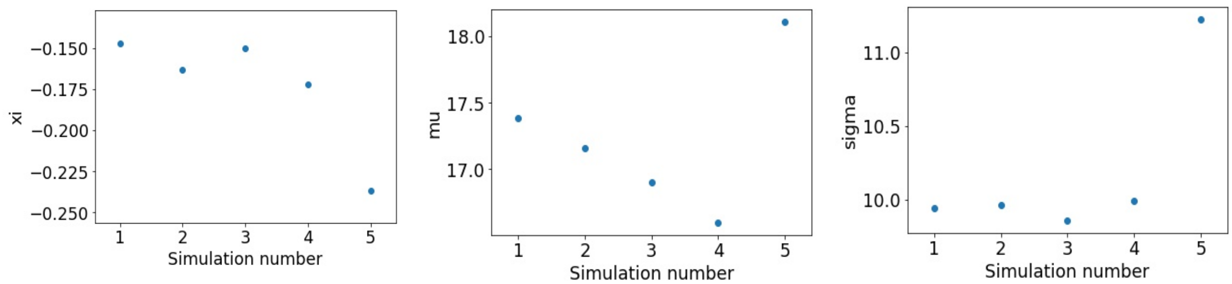


Figure 4. The hailstone size distributions in 5 simulations and corresponding GEV fits.

The GEV distribution has three parameters: shape, location and scale, commonly denoted as ξ, μ, and σ, respectively. The fitting values of these parameters are shown for 5 simulations on the Figure 5.



a *b* *c*

Figure 5. Fitting parameters of the GEV distribution at five simulations: a) ξ, b) μ, c)σ

It can be qualitatively seen on the Figure 5 that the fitting parameters of the simulation 5 are notably different from those of other simulations; at the same time the values for the simulation 4 are rather close to those of simulations 1 to 3. As the next step the statistical significance of these observations will be estimated.

The three simulations 1-3, which were initiated under slightly different initial conditions, are representing the internal variability of the climate system. Using the fitted GEV parameters for these simulations, and considering, that the probability distribution of the internal variability runs is normal, (the Shapiro normality test has been passed) we can estimate the mean values and the variance of the fitted parameters of these simulations. Table 1 presents the mean values and the standard deviations of the ξ, μ, and σ parameters for simulations 1-3.

Table 1. ξ, μ, and σ for internal variability simulations 1-3: mean values and standard deviation

Parameter	Mean	Standard deviation
ξ	-0.153	0.007
μ	17.1	0.2
σ	9.92	0.04

Now it is possible to test the values of ξ, μ, and σ of the simulations 4 and 5 and to check whether the zero hypothesis (that these simulations are following the same probability distribution) can be accepted. The significance threshold of 0.05 (0.025 on the each side of the distribution for the two-sided test) will be used.

Table 2 represents the p-values of ξ, μ, and σ of the simulations 4 and 5 in relation to their normal distributions for simulations 1-3.

Table 2. p-values for fit parameters of simulations 4 and 5.

Parameter	p-value, simulation 4	p-value, simulation 5
ξ	0.997 (1-0.03) passed	1.0 (1.0-0.0) failed
μ	0.997 (1-0.03) passed	0.7E-6 failed
σ	0.36 passed	1.1E-11 failed

For the simulation 4, the null hypothesis can not be rejected. For the simulation 5, the null hypothesis has been rejected.

It can thus be concluded that the hail size distribution in the simulation 4 (Thompson microphysical scheme) agrees with the intrinsic internal variability of the climate system, determined by using three simulations with the Morrison microphysical scheme. The answer to the first research question is negative.

The hailstone size distribution in climate change conditions is significantly different from the present-climate values, and this difference can not be explained by the internal variability of simulation results. The answer to the second research question is thus positive.

REFERENCES

Lorenz, E.N (1963). "Deterministic nonperiodic flow". Journal of the Atmospheric Sciences. **20** (2): 130–141. doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Mahoney, K, Alexander, MA, Thompson G, Barsugli JJ, Scott JD (2012) Changes in hail and flood risk in high-resolution simulations over Colorado's mountains. Nature Climate Change *2*(2):125–131. *https://doi.org/10.1038/nclimate1344*

Morrison H, Thompson G, Tatarskii V (2009) Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One– and Two–Moment Schemes. Mon Wea Rev 137:991–1007. *https://doi.org/10.1175/2008MWR2556.1*

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X-Y, Wang W, Powers JG (2008) A Description of the Advanced Research WRF Version 3. NCAR Technical Note NCAR/TN-475+STR, 113 pp. http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf

Thompson G, Field PR, Rasmussen RM, Hall WD (2008) Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. Mon Wea Rev 136:5095–5115. https://doi.org/10.1175/2008MWR2387.1