This is a repository copy of *Using emulators to understand the sensitivity of deep convective clouds and hail to environmental conditions*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/182657/

Version: Published Version

Article:
Wellmann, C., Barrett, A.I., Johnson, J.S. orcid.org/0000-0002-4587-6722 et al. (4 more authors) (2019) Using emulators to understand the sensitivity of deep convective clouds and hail to environmental conditions. Journal of Advances in Modeling Earth Systems, 10 (12). pp. 3103-3122. ISSN 1942-2466

https://doi.org/10.1029/2018MS001465

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Check for updates

AGU100 ADVANCING EARTH AND SPACE SCIENCE

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2018MS001465

Key Points:
- Comprehensive sensitivity analysis of convective clouds via statistical emulation
- Temperature profile and CCN concentration cause the largest variability
- Results depend on convection trigger (warm bubble, cold pool, and orography)

Correspondence to:
C. Wellmann,
constanze.wellmann@kit.edu

Citation:
Wellmann, C., Barrett, A. I., Johnson, J. S., Kunz, M., Vogel, B., Carslaw, K. S., & Hoose, C. (2018). Using emulators to understand the sensitivity of deep convective clouds and hail to environmental conditions. *Journal of Advances in Modeling Earth Systems*, *10*, 3103–3122. https://doi.org/10.1029/2018MS001465

Received 7 AUG 2018
Accepted 22 NOV 2018
Accepted article online 27 NOV 2018
Published online 15 DEC 2018

©2018. The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Using Emulators to Understand the Sensitivity of Deep Convective Clouds and Hail to Environmental Conditions

C. Wellmann[1], A. I. Barrett[1], J. S. Johnson[2], M. Kunz[1], B. Vogel[1], K. S. Carslaw[2], and C. Hoose[1]

[1]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany, [2]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK

Abstract This study aims to identify model parameters describing atmospheric conditions such as wind shear and cloud condensation nuclei (CCN) concentration, which lead to large uncertainties in the prediction of deep convective clouds. In an idealized setup of a cloud-resolving model including a two-moment microphysics scheme we use the approach of statistical emulation to allow for a Monte Carlo sampling of the parameter space, which enables a comprehensive sensitivity analysis. We analyze the impact of six uncertain input parameters on cloud properties (vertically integrated content of six hydrometeor classes), precipitation, and the size distribution of hail. Furthermore, we investigate whether the sensitivities are robust for different trigger mechanisms of convection. We find that the uncertainties of most cloud and precipitation outputs are dominated by the uncertainty in the temperature profile and the CCN concentration while the contributions of other input parameters to the uncertainties may vary. The temperature profile is also an important factor in determining the size distribution of surface hail. We also notice that the sensitivities of cloud water and hail to the CCN concentration depend on environmental conditions. Our results show that depending on the choice of the trigger mechanism, the contribution of the input parameters to the uncertainty varies, which means that studies with different trigger mechanisms might not be comparable. Overall, the emulator approach appears to be a powerful tool for the analysis of complex weather prediction models in an idealized setup.

Plain Language Summary Severe hailstorms have a large damage potential and cause harm to buildings and crops, for instance. However, important processes for the prediction of hailstorms are insufficiently represented in weather forecast models. This study examines the impact of environmental conditions on hailstorms and identifies sources of uncertainty in their prediction. We find that the temperature and the aerosol load of the atmosphere greatly influence the simulation of hailstorms. Especially, the prediction of the size of the hailstones is affected by changes of the temperature.

1. Introduction

Severe convective storms are frequently associated with heavy precipitation, strong wind gusts and hail often lead to serious damage to buildings, infrastructure, and crops. For example, on 27 and 28 July 2013, Germany was hit by two severe supercells (Kunz et al., 2018) causing insured losses of USD 3.8 billion (SwissRe, 2014). Thus, the modeling and prediction of severe storms and deep convective clouds are key areas of interest in atmospheric research. It is also prerequisite to better understand how changes in the environmental conditions feedback into the dynamics and microphysics of deep convective clouds.

Wind shear is an important parameter controlling the organizational form of convective systems; therefore, its impact on deep convection has been examined over several decades (Dennis & Kumjian, 2017; Schlesinger, 1978; Weisman & Klemp, 1984; Weisman & Rotunno, 2000). For instance, Dennis and Kumjian (2017) find that an increased deep-layer shear elongates the storm's updraft downshear, leading to an increased volume of the hail growth zone in combination with longer residence times, thus increasing the resulting hail mass. In recent years, many studies have been published regarding the sensitivity of deep convective clouds to the concentration of cloud condensation nuclei (CCN) and aerosol load (Cui et al., 2011; Fan et al., 2013; Khain et al., 2011; Morrison, 2012; Noppel et al., 2010; Rosenfeld et al., 2008; Tao et al., 2007, 2012; Yang et al., 2017).

For example, Tao et al. (2007) compared three storms observed during measurement campaigns for clean and polluted aerosol conditions. They found that compared to the low CCN concentration setups, rain is suppressed in the early stages of the storm for the high CCN cases. However, in the mature stage of the cloud the effect of the CCN varies between suppression and enhancement of rain depending on the modeled storm. In an idealized study, Morrison (2012) examined the effect of small perturbations of microphysical and thermodynamical processes in addition to the effect of CCN concentration on deep convective clouds. They showed that in general polluted conditions lead to a weakening of the convection, whereas only small changes are visible for the domain-mean accumulated precipitation. Moreover, they claimed that the effect of the CCN concentration clearly depends on additional perturbations: For example, turning off the hail initiation leads to an invigoration of convection for polluted conditions. Overall, the general outcome of these studies is not consistent, and research is still ongoing.

Several studies investigated the effect of not only one but two parameters on deep convection. Lee et al. (2008), for example, examined the effect of CCN concentrations in convective clouds for different environmental conditions in terms of changing wind shear and convective available potential energy (CAPE), and Fan et al. (2009) investigated the sensitivities to the wind shear and the CCN concentration. Moreover, large and well-defined perturbations were usually used in most of those studies to trigger deep convection at specific locations and without the need of long spin-up times in idealized model setups. For instance, Morrison et al. (2009) examined the impact of microphysics parameterizations on the modeling of squall lines by implementing an ellipsoidal positive temperature perturbation, also referred to as warm bubble (WB), to trigger the initial convective cell. Similarly, Adams-Selin et al. (2013), who investigated the impact of the graupel parameterization on the development of bow echos, triggered convection by a prescribed cold pool (CP). Both Chen and Lin (2005) and Storer et al. (2010) analyzed among other things the effect of CAPE on convection. Whereas in Chen and Lin (2005) the convection is generated by a mountain ridge, Storer et al. (2010) inserted a WB.

The above-mentioned publications indicate that several parameters affect the formation and dynamics of deep convective clouds and the corresponding in-cloud processes. In addition, as each study is confined to a single trigger mechanism and comparable studies have been performed for several triggers separately from each other, the results for different trigger mechanisms have not been directly compared. Here we conduct a comprehensive sensitivity study using high-resolution model simulations with a sophisticated microphysics scheme to determine the effects of six parameters and environmental conditions including the vertical temperature profile and the wind shear on deep convection. Our work aims to identify the environmental conditions that lead to large uncertainties in the prediction of deep convective clouds. We focus on the prediction of integrated hydrometeor mass, precipitation, and the size distribution of hail. Furthermore, we investigate whether the sensitivities are robust for different trigger mechanisms of deep convection by triggering convection either with a WB, a CP, or orography. In an idealized setup of a cloud resolving model we modify the selected input parameters to examine their impact on the model output.

The straightforward approach for analyzing the sensitivity of the model output to changes in the input parameters is to vary a chosen parameter in a given range while the other parameters are kept constant. This so-called *one-at-a-time* (OAT) analysis can give insight into problems where only the effect of a single model input is of interest. However, our goal is to identify not only the effect of each input parameter independently, but also to assess the amount of interactions that are happening between those parameters not captured by one-at-a-time analyses. Furthermore, an additional objective is to examine the relative contribution of the input parameters to the uncertainty of the output. Therefore, we make use of statistical emulation (O'Hagan, 2004; 2006) and variance-based sensitivity analysis (Saltelli, 2008), where the uncertainty of the output is decomposed into the contributions of the individual model input parameters but by also considering their interactions. The applicability of this approach for complex atmospheric models was demonstrated in L. A. Lee et al. (2011, 2013) where this method was successfully applied to determine the sensitivity of a global aerosol model to uncertain parameters. Of most relevance to the present study, Johnson et al. (2015) used statistical emulation to quantify the sensitivity of convective cloud properties to aerosol concentrations and microphysical processes. They found that overall the parameters corresponding to the concentration of the Aitken and accumulation aerosol modes and the collection efficiency of droplets by graupel had the largest impact on the output uncertainty of the cloud properties, whereas the parameters describing ice properties were only of minor importance. Furthermore, the main contributors to the uncertainty varied when only subdomains of the parameter space were examined, indicating that different cloud regimes can have different sensitivities to the parameters considered.

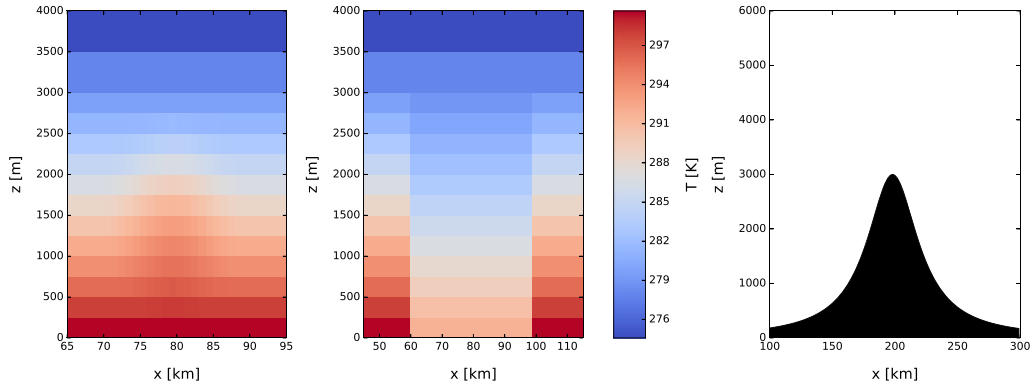

Figure 1. Triggers of deep convection used in this study. (a) Ellipsoidal warm bubble with a temperature perturbation of $\Delta T = 2$ K and a horizontal radius of $R_{hor} = 10$ km. (b) Cylindrical cold pool with a temperature perturbation of $\Delta T = -8$ K and a radius of $R_{hor} = 20$ km. (c) Bell-shaped mountain ridge with a height of 3, 000 m and a half width of 25 km.

The model setup and the input parameters are described in section 2 followed by an explanation of the methods of statistical emulation and variance-based sensitivity analysis in section 3. The results of the sensitivity analysis and further analyses are presented in section 4. In section 5 we summarize and conclude our findings.

2. Model Setup

For this study we use the COSMO (Consortium for Small-Scale Modeling) model (Baldauf et al., 2011; Schättler et al., 2016) developed by Deutscher Wetterdienst and the COSMO consortium. It is a limited-area numerical weather prediction model, which is nonhydrostatic and compressible. We run COSMO in an idealized setup with 700×500 grid points where the horizontal resolution is 1 km and consequently convection is explicitly resolved. The domain extends to a height of 23 km resolved by 64 vertical levels. Furthermore, we use open boundary conditions, the radiation scheme is switched off, and the Coriolis force is neglected in the simulations. The initial temperature and humidity profiles are taken from Weisman and Klemp (1982), hereinafter referred to as WK, to maintain atmospheric conditions that favor the development of deep convection. The maximum specific humidity q_{v0} according to WK is set to a value of 12 g/kg at the lowest level. The vertical profile of the wind speed is chosen to be similar to Figure 3b of Weisman and Rotunno (2000), where they presented the hodograph for quarter-circle shear. We also use the two-moment mixed-phase bulk microphysics scheme by Seifert and Beheng (Seifert & Beheng, 2006a), which predicts both the mass mixing ratios and the number densities of six hydrometeor classes, namely, cloud droplets, rain, cloud ice, snow, graupel, and hail.

In this study the results are generated and compared for three different trigger mechanisms to initiate deep convection: a WB, a CP, and bell-shaped mountain ridge. Cross sections of the trigger mechanisms are shown in Figure 1.

First, the frequently used approach of a WB is employed (Figure 1a). As in WK the maximum temperature excess ΔT lies between 2 K and 5 K and is located in the center of the bubble. The perturbation decreases toward the edges following a $\cos^2$ function. The vertical extent of the WB is fixed at $R_z = 1,400$ m whereas the horizontal radius is varied between $R_{hor} = 5$ km and $R_{hor} = 15$ km. At model initialization the bubble is released at a distance of $\Delta x = 80$ km from the model boundary to the west and at a distance of $\Delta y = 200$ km from the boundary to the south where it ascends and triggers convection.

Second, we use a CP as the trigger mechanism, illustrated in Figure 1b. It is a negative temperature perturbation in the shape of a cylinder, which is placed on the ground and reaches a height of $z = 3,000$ m. The strongest temperature contrast is found at the bottom of the cylinder and the difference in temperature to the ambient air decreases linearly with increasing height. Within the CP the conditions are horizontally homogeneous. The simulations are initialized with the CP located at the same position as the WB.

The third trigger mechanism is a bell-shaped mountain ridge, which extends from the northern to the southern boundary of the domain (Figure 1c). Its height and half width range between 2,000 and 5,000 m and 5 and 50 km, respectively. The center of the ridge is situated at a distance of $\Delta x = 200$ km from the boundary to the west and a distance of $\Delta y = 200$ km from the boundary to the south. As the directional wind

Table 1
Overview Over the Uncertain Input Parameters and Their Ranges

Input	Min	Max	Units
CCN concentration	100	4,000	cm^{-3}
INP concentration	0.01	10	Factor
Wind shear (F_{shear})	0	1	Scaling factor
Potential temperature at the ground	295	302	K
θ_0 (WK profile)			
Temperature perturbation ΔT (WB)	2	5	K
Temperature perturbation ΔT (CP)	−10	−6	K
Height of ridge (Oro)	2,000	5,000	m
Radius of warm bubble $R_{hor, WB}$	5	15	km
Radius of cold pool $R_{hor, CP}$	10	30	km
Half width of ridge (Oro)	5	50	km

Note. CCN = cloud condensation nuclei; CP = cold pool; WB = warm bubble; INP = ice-nucleating particles; WK = Weisman and Klemp (1982).

shear in the lower atmosphere is altered during the simulations, the angle between the axis of the crest and a north-to-south axis is coupled to the directional shear. This means that the orientation of the crest varies between a north-to-south axis and a north-west-to-south-east axis depending on the wind direction in the lower troposphere in order to guarantee for a roughly perpendicular flow over the mountain ridge. In contrast to the thermal triggers, the ridge constantly produces small clouds right above the crest, which reduce the comparability to the WB setup and the CP setup. Therefore, a larger domain of 1,200 km × 500 km is used to give room to the orography. For the analysis, the western part of the domain containing the ridge is removed and only an area of 700 km × 500 km located in the east is considered. All simulations are run for 5 hr with a time step of $\Delta t = 6$ s. Before the evaluation period 1 hr of spin up is run for the WB and CP setups, whereas for the orography case a 5-hr spin-up is required to allow for more persistent convective development. Note that the excluded initial development of the cell differs for each trigger mechanism.

2.1. Input Parameters

A set of six input parameters is defined describing different regimes of atmospheric conditions. This set consists of the CCN concentration, the concentration of ice-nucleating particles (INP), the directional wind shear, and the surface potential temperature θ_0 determining the WK temperature profile. We also perturb the magnitude of the temperature, the radius of the WB and the CP, whereas in the orography case the height and the half width of the mountain ridge are perturbed. An overview over the chosen input parameters and their ranges is given in Table 1, along with a detailed description of each parameter in the following sections.

2.1.1. CCN Concentration

CCN play an essential role in the activation of cloud droplets. Thus, from the early stage of cumulus cloud formation to the dissipation, the CCN concentration influences the development and dynamics as well as the microphysics (Cui et al., 2006, 2011; Fan et al., 2013; Khain et al., 2011; Morrison, 2012; Noppel et al., 2010; Rosenfeld et al., 2008; Seifert & Beheng, 2006b; Tao et al., 2012). A cloud droplet nucleation scheme based on grid-scale supersaturation and empirical power law activation spectra is implemented in COSMO using lookup tables introduced by Segal and Khain (2006). Furthermore, the aerosol is assumed to have the largest concentration in the lowest 2 km above the ground, followed by an exponential decrease toward higher altitudes with a scale height of 1 km. In this study, the maximum CCN concentration is varied between 100 and 4,000 cm^{-3}, which corresponds to a change from maritime to polluted conditions.

2.1.2. Concentration of INP

Similar to CCN, the INP are particles that support the formation of cloud ice (Houze, 1993) and therefore primarily affect the number of ice particles in the cloud. Here a scaling factor is applied to the INP concentration for three microphysical processes: the deposition nucleation of cloud ice, the immersion freezing of cloud droplets, and the freezing of raindrops. In the case of the formation of cloud ice, the heterogeneous nucleation scheme of ice from Huffman and Vali (1973) is used in the simulations. For the freezing of cloud droplets and raindrops, a stochastical model is implemented in the two-moment scheme (Seifert & Beheng, 2006a) following the measurements of Bigg (1953). The values for the parameters were taken from Pruppacher and

Klett (1997). For the simulations in our study, we multiply the INP concentration in each of these three processes by the same scaling factor, which we vary from 0.01 to 10 on a logarithmic scale. This range is chosen according to DeMott et al. (2010) to represent measurements of different field campaigns.

2.1.3. Wind Shear

In general, wind shear describes the difference in wind velocity and wind direction between two heights and thus characterizes the wind field of the environment in the embedded layer. Here we focus on directional shear associated with streamwise vorticity as this is most important for the organization of convection. The vertical profile of velocity is not varied between the simulations. We introduce the factor F_{shear}, which determines the vertical profile of the wind direction WD according to

$$WD(z) = \begin{cases} 270° - F_{\text{shear}} \cdot 90° + \frac{F_{\text{shear}} \cdot 90°}{6000\,m} \cdot z & ,z \leq 6,000 \text{ m AMSL}, \\ 270° & ,z > 6,000 \text{ m AMSL}, \end{cases} \quad (1)$$

where $F_{\text{shear}} \in [0; 1]$. For example, $F_{\text{shear}} = 0$ represents westerly wind at all heights, and the maximum value of $F_{\text{shear}} = 1$ indicates southerly wind near the ground, which gradually turns into a westerly wind with increasing height.

2.1.4. Potential Temperature

The WK profile of the potential temperature used in our study provides atmospheric conditions favoring deep convection and is calculated as

$$\theta(z) = \begin{cases} \theta_0 + (\theta_{\text{tr}} - \theta_0) \left(\frac{z}{z_{\text{tr}}}\right)^{\frac{5}{4}} & ,z \leq z_{\text{tr}}, \\ \theta_{\text{tr}} \cdot \exp\left[\frac{g}{c_p T_{\text{tr}}} (z - z_{\text{tr}})\right] & ,z > z_{\text{tr}}, \end{cases} \quad (2)$$

with the potential temperature at the tropopause $\theta_{\text{tr}} = 343$ K, the height of the tropopause $z_{\text{tr}} = 12$ km, the gravity acceleration $g = 9.80665\,\text{m/s}^2$, the specific heat of dry air $c_p = 1,005\,\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$ and the temperature at the tropopause $T_{\text{tr}} = 213$ K. The near-surface potential temperature θ_0 is initially set to 300 K (Weisman & Klemp, 1982) but in this study θ_0 is varied between 295 K and 302 K in order to simulate different atmospheric conditions. This variation of the vertical temperature profile with constant surface humidity of 12 g/kg results in a change of the CAPE from 1, 048 to 1, 410 J/kg.

2.1.5. WB and CP Characteristics

The thermal triggers, WB and CP, are characterized by a temperature perturbation ΔT and a radius R_{hor}. The temperature perturbation describes the maximum initial deviation between the core temperature of the WB or the CP and the ambient air. In the WB setup, ΔT ranges from 2 K (Weisman & Klemp, 1982) to 5 K (Brooks, 1992). For the CP setup ΔT is set to approximately −8 K (Adams-Selin et al., 2013; Weisman et al., 1997) varying between −10 K and −6 K. The radius R_{hor} specifying the horizontal extent of the thermally perturbed air mass ranges from 5 to 15 km for the WB and from 10 to 30 km for the CP. Varying the temperature perturbation and the radius leads to different buoyancy gradients of the initial perturbation, which is a measure of the strength of the trigger.

2.1.6. Height and Half Width of Ridge

While the WB and the CP are defined by a temperature perturbation and a radius, the ridge is characterized by the height and the half width. The height ranges between 2,000 and 5,000 m; hence, it is similar to the height of the Alps and the Rocky Mountains. The half width specifies the horizontal extent from the center of the mountain ridge to its edge in x direction at an elevation equal to half of the maximum height. In this study, the half width is varied between 5 and 50 km representing both steep and extensive mountains.

3. Statistical Methods

We perform a global sensitivity analysis to identify contributions to the uncertainty in model outputs from both individual parameters and interactions between them. This is possible by variance-based sensitivity analysis (Saltelli, 2008) through which the output uncertainty is decomposed into the contributions of the model inputs and their interactions. To infer those measures, the variance-based approach requires a comprehensive sampling of the model output over the whole parameter uncertainty space defined by our parameter ranges given in Table 1. The necessary output is often produced by Monte Carlo simulations (Saltelli, 2008). As the COSMO model used here is a complex numerical weather prediction (NWP) model with high computational cost, the data for the sensitivity analysis cannot be generated using a Monte Carlo approach in a reasonable time for this model, and therefore, an alternative method is needed. Here we adopt the approach of statistical emulation to enable dense sampling of model output over the parameter uncertainty (Oakley & O'Hagan,

2004; O'Hagan, 2004). The idea behind emulation is to build a surrogate model using training data that represents the relationship between a particular model output and a set of uncertain inputs. This means that an emulator is able to estimate a specific model output for new input combinations without having to run full model simulations. Due to its low computational cost, the required data for the variance-based sensitivity analysis can easily be produced by the emulator. To construct emulators for different model outputs over the parameter uncertainty, a number of well-spaced input combinations are selected from within the parameter uncertainty space. Outputs from the COSMO simulations with these selected parameters are used to train the emulator, these are referred to as *training runs* (Johnson et al., 2015). After evaluation of the model output, we use Gaussian process emulation to construct the surrogate model for each model output of interest independently. Once they are validated (Bastos & O'Hagan, 2009), the emulators can be used to generate output across the whole parameter space, as needed for the variance-based sensitivity analysis. In the following sections we describe this procedure in more detail.

3.1. Experiment Design

We require the emulator to predict the model output with acceptable accuracy; thus, sufficient input information via the training runs is needed. On the one hand, features of the output might get lost and thus cannot be reproduced by the emulator if there are too few training runs. On the other hand, if there are too many training runs, the emulator might reproduce the output almost perfectly, but a lot of computational time is needed, which is impractical. Loeppky et al. (2009) suggested a data set of $n = 10d$ training runs to obtain reliable results, where d is the number of uncertain input parameters (here $d = 6$). However, the number can be increased if the accuracy of the first result is not as high as expected. Here we raise the number to $n = 15d$ in order to add information while keeping the computational effort feasible. We require the emulator to predict the model output across the whole multidimensional parameter uncertainty space, and so our training data simulations must have good coverage of this space and be well-spaced throughout it. To achieve this, the "maximin Latin hypercube sampling" (Morris & Mitchell, 1995) is applied to place the input combinations within the parameter space for the training runs. This algorithm maximizes the minimum distance between the input combinations for the training runs and thus ensures good coverage of the parameter uncertainty space.

3.2. Gaussian Process Emulation

A Gaussian process is a generalization of the Gaussian distribution and the multivariate Gaussian distribution to an infinite number of variables (Rasmussen, 2004). Similar to the mean and the variance of a normal distribution such a process is defined by a mean function $m(\mathbf{x})$ and a covariance structure $V(\mathbf{x}, \mathbf{x}')$, where $\mathbf{x} = (x_1, \ldots, x_6)$ is an input combination within the defined parameter uncertainty. The mean function is specified by

$$m(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \boldsymbol{\beta}, \tag{3}$$

where $\mathbf{h}(\mathbf{x})$ contains known regression functions of $\mathbf{x}$ and $\boldsymbol{\beta}$ consists of unknown coefficients. The regression functions are arbitrary; however, they should be chosen such that they reflect prior beliefs about the form of the emulator (Johnson et al., 2015; Oakley & O'Hagan, 2004). Here we assume a linear trend. The covariance structure is given by

$$V(\mathbf{x}, \mathbf{x}') = \sigma^2 c(\mathbf{x}, \mathbf{x}'), \tag{4}$$

where σ^2 is an unknown scale parameter and the function $c(\mathbf{x}, \mathbf{x}')$ a correlation function. The correlation function is designed to decrease as $|\mathbf{x} - \mathbf{x}'|$ increases and furthermore $c(\mathbf{x}, \mathbf{x}) = 1$ has to be valid. Here we chose the Matérn correlation structure

$$c_{\text{Matern}}(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} r}{l} \right)^{\nu} K_{\nu} \left(\frac{\sqrt{2\nu} r}{l} \right), \tag{5}$$

with positive parameters ν and l, Gamma function $\Gamma(\nu)$, and a modified Bessel function K_{ν}. The Matérn choice leads to a stationary and isotropic covariance function since it only depends on $r = |\mathbf{x} - \mathbf{x}'|$ and thus is invariant to rigid motions (Rasmussen & Williams, 2006). This specification of the mean function and covariance structure leads to additional parameters such as ν and l being introduced. We have no prior information about these so-called *hyperparameters*; hence, they are estimated from the training data. This is done by optimizing the marginal likelihood, which is the probability of the data given the hyperparameters (Rasmussen, 2004; Rasmussen & Williams, 2006). To construct the emulator, these prior specifications of the Gaussian process are updated through a Bayesian statistical framework using the information in the training runs to produce a

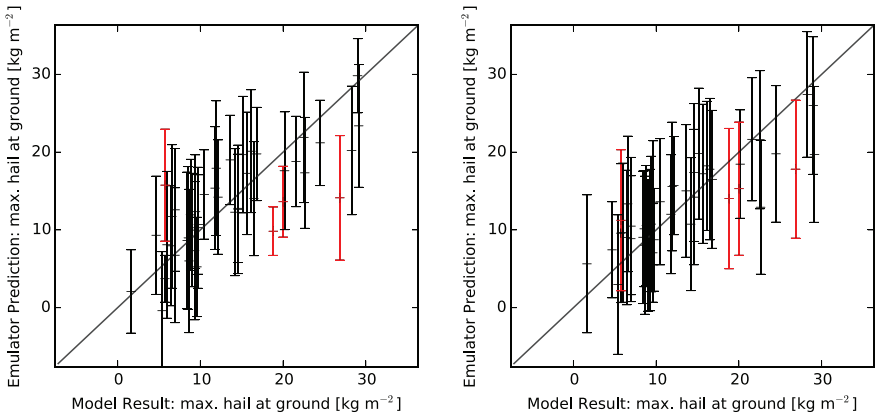

Figure 2. (a) Emulator predictions of the maximum amount of hail at the ground with 95% confidence intervals versus the model output (cold pool setup). Outliers are depicted in red. (b) Same as (a) but with expanded training data. For comparison the same points as in (a) are depicted in red.

posterior Gaussian process specification that has an updated mean and covariance specification conditioned on the training data. The updated posterior mean and covariance take the following form (Johnson et al., 2015; O'Hagan, 2004)

$$m^*(x) = \mathbf{h}(x)^T\hat{\beta} + \mathbf{c}(x)^T\mathbf{C}^{-1}\left(\mathbf{y} - \mathbf{H}^T\hat{\beta}\right) \tag{6}$$

and

$$\begin{aligned} V^*(x,x') &= \hat{\sigma}^2 c^*(x,x') \\ &= \hat{\sigma}^2\left[c(x,x') - \mathbf{c}(x)^T\mathbf{C}^{-1}\mathbf{c}(x')\right. \\ &\quad \left. + \left[\mathbf{h}(x) - \mathbf{H}\mathbf{C}^{-1}\mathbf{c}(x)\right]^T\left[\mathbf{H}\mathbf{C}^{-1}\mathbf{H}^T\right]^{-1}\left[\mathbf{h}(x') - \mathbf{H}\mathbf{C}^{-1}\mathbf{c}(x')\right]\right], \end{aligned} \tag{7}$$

where

$$\begin{aligned} \mathbf{C} &= \begin{pmatrix} c(x_1,x_1) & c(x_1,x_2) & \dots & c(x_1,x_n) \\ c(x_2,x_1) & c(x_2,x_2) & \dots & c(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ c(x_n,x_1) & c(x_n,x_2) & \dots & c(x_n,x_n) \end{pmatrix}, \\ \mathbf{H} &= \left(\mathbf{h}(x_1), \mathbf{h}(x_2), \dots, \mathbf{h}(x_n)\right), \\ \hat{\beta} &= \left[\mathbf{H}\mathbf{C}^{-1}\mathbf{H}^T\right]^{-1}\mathbf{H}\mathbf{C}^{-1}\mathbf{y}, \\ \hat{\sigma}^2 &= \frac{\mathbf{y}^T\left(\mathbf{C}^{-1}\mathbf{H}\left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}^{-1}\right)\mathbf{y}}{n-q-2}. \end{aligned}$$

A detailed derivation and additional information regarding statistical emulation can be found in O'Hagan (2004, 2006) and Johnson et al. (2015).

3.3. Validation

Once an emulator is built, it is necessary to evaluate it to determine whether it produces an accurate estimation of the model at points in the parameter space that were not originally included in the training data. Therefore, we perform 45 additional model runs whose input combinations are also generated by the maximin Latin hypercube algorithm to cover the parameter space (*validation data*). Emulator predictions at these points can be plotted against the COSMO model output, and if the emulator is a reasonable representation of the model these points should follow the 45° line (line of equality). Moreover, 95% confidence bounds on the emulator predictions can be obtained from the Gaussian process posterior specification of the emulator. We consider an emulator to be valid if the errors in prediction are reasonably small (points lie close to the line of equality) and if the 95% confidence bound on the emulator prediction crosses the line of equality for at least 95% of the validation points. Figure 2a shows this validation plot for the maximum amount of accumulated hail at the ground observed at a grid point in the CP setup.

In general the points follow the line of equality, but as 95% of the validation points should contain the COSMO output within the 95% confidence interval on the emulator prediction, only two to three outliers are acceptable in a set of 45 validation runs. For this example, there are four outliers marked in red, which exceeds the

allowed amount of two and thus the emulator is not validated here. To address this, 10 new points are added to the initial Latin hypercube sampling and the gained information from these extra simulations is used to complement the information contained in the training data. New emulators are fitted for this extended training data (Figure 2b). For clarity the same points including their confidence bounds are depicted in red. The additional information in the training data improved the emulator fit as the number of outliers is clearly reduced in Figure 2b compared to Figure 2a. A wide range of diagnostics is available to validate Gaussian process emulators such as quantile-quantile plots and individual prediction errors, which can be plotted against the index of the simulation, the model output or input parameters to detect possible correlations (Bastos & O'Hagan, 2009). For our emulators all of the mentioned measures (not shown) indicate that, similar to Figure 2b, there are a small number of outliers with larger errors than the threshold mentioned in Bastos & O'Hagan, 2009 (2009; not shown). However, based on the increase of the prediction accuracy through including the additional training runs and the lack of systematic errors in the results of the above-mentioned validation methods we consider all emulators as validated.

Finally, we assess the robustness of the emulators in order to rule out potential dependencies of the result on the choice of the training data. To do this, we take a random sample of 45 new validation runs from the original training data and swap out this sample from the training set, replacing it with the original validation data. Although this new training data does not guarantee for a well-spaced coverage of the parameter space as two different Latin hypercube samplings are mixed, it is still used to build another emulator. This is done for a total sum of 10 random samples leading to a set of independent emulators describing the same conditions. The results of these emulators and the original one are compared and as no structural changes in the results emerge. Hence, we find that the emulators are robust to changes in the training data.

3.4. Variance-Based Sensitivity Analysis

When running a model or emulator with different combinations of d different input parameters, it is common to induce an uncertainty in the output's value, which can be quantified as a variance. The goal of the variance-based sensitivity analysis is to decompose this output variance into different contribution sources related to the input parameters that were varied. This decomposition of the variance includes not only contributions from each parameter individually but also contributions from interactions between two or more input parameters. Assuming independence between the input parameters, this decomposition can be written as (Oakley & O'Hagan, 2004)

$$V = \sum_{i=1}^{d} V_i + \sum_{i<j} V_{ij} + \ \ldots \ + V_{1\ldots p}, \tag{8}$$

where V_i are contributions from each parameter. Terms with more than one index indicate contributions from parameter interactions, where the number of indices corresponds with the number of interacting parameters. V_{ij} are contributions from two interacting parameters, while $V_{1\ldots p}$ describes the contribution from all considered parameters interacting with each other. Here we use the extended Fourier amplitude sensitivity test (FAST) introduced by Saltelli et al. (1999) to generate this variance decomposition. The concept of FAST is a transformation to 1-D Fourier space such that the d-dimensional input space can be explored by a monodimensional curve. This has the advantage that along the path all input parameters change simultaneously, which means that FAST offers a global sensitivity analysis. In order to gain a space-filling curve and to avoid possible overlap with upper frequencies, a sample size of several thousands of runs is desirable (Saltelli et al., 1999). As this amount of simulations is not feasible with the conservative approach of running the full NWP model, the emulators are necessary to simulate the required model output. We compute the sensitivity measures of the FAST approach using the statistical software R (R Core Team, 2017) and the R package "sensitivity" (Pujol et al., 2017). We obtain the so-called *main effect* S_i of each input parameter i by normalizing the calculated individual variance contribution of the parameter V_i with the overall variance V in the output

$$S_i = \frac{V_i}{V}, \tag{9}$$

which is a measure of the contribution of input i to the output variance. Hence, the main effect shows the percentage by which the output variance could be decreased if there was no uncertainty in the input i. Furthermore, the *total effect index* S_{T_i} generated by this method includes all terms that are linked to input i, that is, both individual contribution and interaction terms. The total effect is given by $S_{T_i} = 1 - \frac{V_{\sim i}}{V}$, where $V_{\sim i}$ denotes all terms of equation (8) that do not include contributions from input parameter i. As the total effect

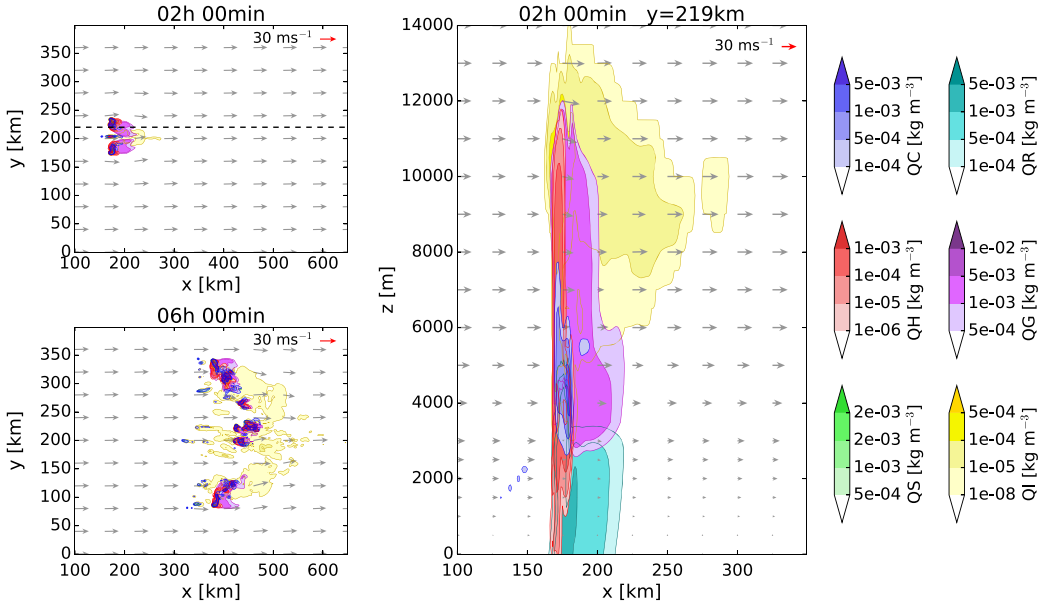

Figure 3. Cross sections of the cloud at different time steps. (a, b) Horizontal cross section at a height of $z = 5,000$ m for (a) 2 hr and (b) 6 hr after the start of the simulation. (c) Vertical cross section at $y = 219$ km for 2 hr after the start of the simulation. The colors denote the different hydrometeors (dark blue = QC: cloud water; light blue = QR: rain; red = QH: hail; purple = QG: graupel; green = QS: snow; yellow = QI: ice), and the dashed line in (a) denotes the location of the vertical cross section of (c). The arrows illustrate the wind field.

includes both individual and interaction contributions and the individual contributions are accessible through the main effect, the *interaction effect* S_{I_i} can be determined by $S_{I_i} = S_{T_i} - S_i$ giving an indication of how much input parameter i interacts with other parameters.

3.5. Output Variables for the Construction of Emulators

Many aspects are of interest when analyzing the sensitivities of clouds and hail formation to different conditions. We focus on variables that describe the properties of the clouds, precipitation, and the size distribution of hail. The hydrometeors are analyzed as the vertically integrated content of each particle class (cloud water, graupel, hail, ice, rain, and snow). The spatial and temporal mean is taken over the grid points showing hydrometeor concentrations larger than 0 kg/m^3 and all time steps, which results in a single value for each particle class for every simulation. Regarding the precipitation, we consider area-mean accumulated precipitation and precipitation rate, both for total precipitation and hail. Instead of the spatial and temporal means, the maximum values of these output variables are used for the analysis here, except for the precipitation of hail where both the mean and the maximum value are evaluated, in order to concentrate on the extremes of precipitation, which have a large impact on the ground. Moreover, the size distribution of hailstones reaching the ground is of interest. In the two-moment scheme of Seifert and Beheng (2006a) a generalized Γ distribution is implemented for the distributions of hydrometeors:

$$\frac{dN}{dx} = N_0 x^{\nu} \exp\left(-\lambda x^{\mu}\right), \tag{10}$$

where N is the number density, x is the particle mass, and ν and μ are the shape parameters of the assumed distribution. Following the standard definition of the hail class in COSMO, we use $\mu = 0.333333$ and $\nu = 1.0$ in this study. The coefficients N_0 and λ are given by gamma distributions of the number and mass density, respectively (Seifert & Beheng, 2006a). Via a conversion from mass x to particle diameter D, the term $\frac{dN}{dx}$ can be transformed to $\frac{dN}{dD}$, and thus, a measure for the number of particles per diameter is available. Emulators are constructed for the number density $N(D)$ at 13 fixed diameters for the spatiotemporal mean of the size distribution of hail at the ground.

4. Results

4.1. Cloud Structure

A short overview of the typical development and the structure of the simulated clouds is given in this section using the WB setup as an example.

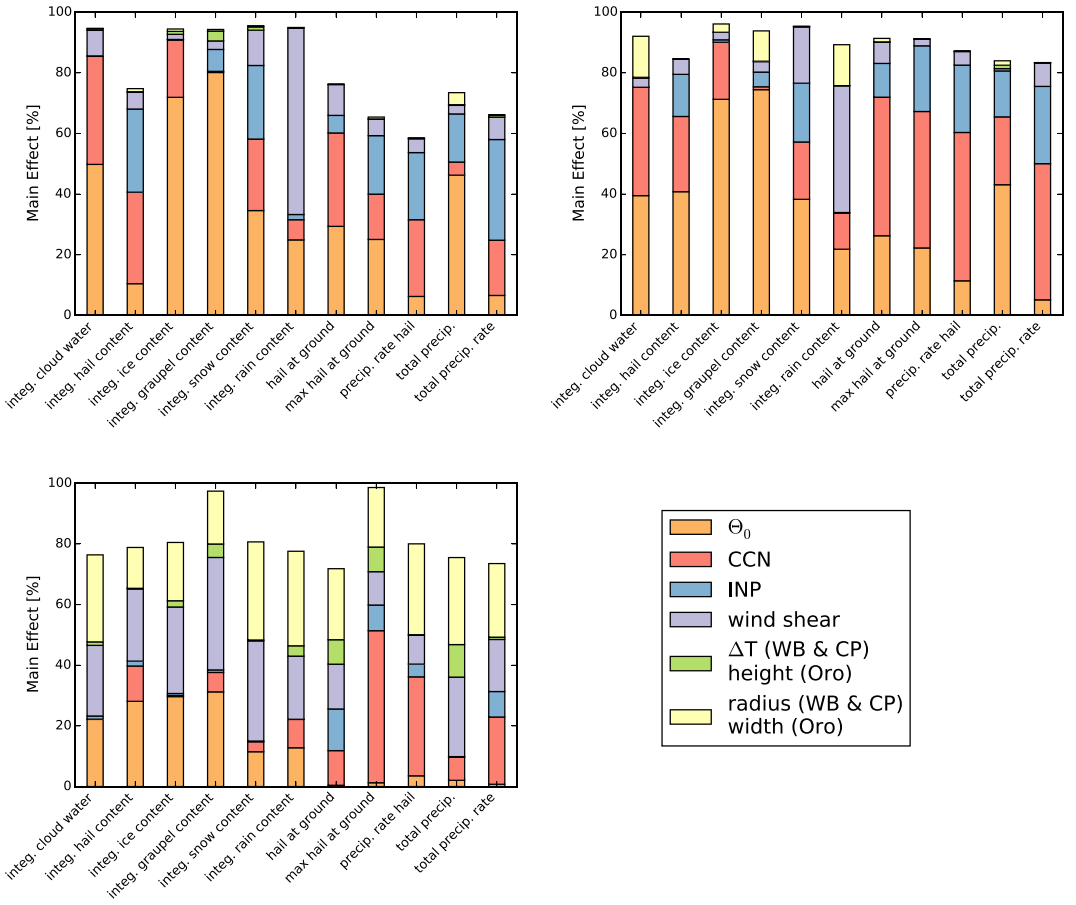

Figure 4. Bar plot of the main effect for vertically integrated hydrometeor masses and precipitation of cloudy grid points when (a) a warm bubble, (b) a cold pool, and (c) a mountain ridge is used as trigger mechanism. CCN = cloud condensation nuclei; CP = cold pool; INP = ice-nucleating particles; WB = warm bubble.

The WB initiates a single cell that develops further. After 2 hr the storm has split into two cells (Figure 3a). The updraft regions contain graupel, hail, and cloud water; graupel is also present in larger areas outside of the updrafts. After 6 hr, the convective cells have moved to the east but have developed into a curved structure extending in the *y* direction similar to squall lines (Figure 3b). In the vertical cross section after 2 hr (Figure 3c) graupel and hail appear at all heights of the cloud, whereas ice is found in the anvil near the tropopause and cloud water is found mainly between $z = 2,000$ m and $z = 6,000$ m. Rain, graupel, and hail reach the ground as precipitation.

4.2. Sensitivity Analysis Using the Predefined Parameter Space

To distinguish the contributions of the input parameters to the uncertainty of the model outputs, we have performed a variance-based sensitivity analysis where the input parameters are varied according to Table 1. Figure 4 shows stacked bar charts of the calculated main effects—the contributions from the individual parameters to the overall output uncertainty—for the mean integrated hydrometeors and the outputs characterizing the precipitation.

If the main effect percentages in a bar do not add up to 100%, this means that further contributions of interacting parameters can be found. Comparing the overall appearance of Figures 4a–4c, we find that the sum of the main effects reaches 80% for most output variables in the WB and the CP setups, whereas in the orography setup only a few bars reach this percentage. Hence, large parts of the uncertainty in the WB and CP output can be explained by the individual parameters, whereas interaction effects contribute substantially to the output uncertainty in the orography setup. This implies that the input parameters are interacting more to jointly affect the uncertainty compared to the thermal trigger setups. A possible explanation for this is that in the simulations the bell-shaped mountain may trigger several cells along the ridge simultaneously forming

a complex system where the parameters are interacting more compared to the initial single cell triggered in the WB and CP setups.

Looking at the mean of the vertically integrated hydrometeor masses in Figures 4a–4c, we find that for all three trigger mechanisms the potential temperature θ_0 contributes significantly to the uncertainty of these variables. In the WB and CP setups, θ_0 is the largest source of uncertainty for the integrated contents of ice and graupel. In the orography setup θ_0 adds 31% to the output uncertainty of the integrated graupel content and between 12% and 29% to the other variables. Via equation (2) θ_0 directly determines CAPE. CAPE is transformed into kinetic energy of ascending air parcels, such that it affects the updraft velocity and thus the storm organization. In addition, θ_0 is a controlling factor of the moisture availability. Consequently, it acts as source of uncertainty for the integrated masses of the hydrometeors. The CCN concentration has a considerable main effect of 30–35% for the integrated cloud water and the integrated hail content in the WB and CP setups. The effect on the remaining hydrometeor species is less than 20%. In contrast, the main effect of the CCN concentration is less than 1% for the integrated cloud water and less than 11% for the other hydrometeors in the orography setup. In the WB setup and in the CP setup, the uncertainty of the INP concentration takes an effect only on the integrated snow, graupel, and hail content, where the contribution reaches up to 27% for the integrated hail content in the WB setup. In the orography setup the INP concentration adds less than 2% for all integrated hydrometeors.

In the WB and CP setups the percentage of the main effect of the wind shear is rather small for the integrated hydrometeor masses, except for the integrated snow and rain contents where there are larger contributions reaching as high as 61% for the rain content in the WB setup. This relates to the impact of the wind shear on the convectively generated CP and the connected evaporation rate of rain, as discussed, for example, by Weisman and Klemp (1982) and Weisman et al. (1997). Furthermore, snow and rain are advected in COSMO for high-resolution simulations (Doms & Baldauf, 2015), and thus, they are directly affected by variations of the wind field. In the orography setup, the uncertainty contribution of the wind shear is comparable for all hydrometeor species varying between 37% for the integrated graupel content and 21% for the rain content.

The contribution of the trigger characteristics to the uncertainties varies for each trigger mechanism. In the WB setup neither the temperature perturbation ΔT nor the radius of the WB R_{hor} have a noteworthy contribution to the overall uncertainty for all condensates. In the CP setup the initial size of the CP affects the results more clearly by adding about 14% to the uncertainty of the integrated cloud water, graupel, and rain contents. However, the contribution of the temperature perturbation is hardly visible. In the orography setup the properties of the mountain ridge explain in combination approximately the same amount of uncertainty as the wind shear, whereas the half width is more important than the height. Horizontal cross sections of the cells reveal that their initial size barely depends on the radius of the WB, whereas a large radius of the CP leads to an extensive initial cell (not shown). Thus, uncertainty in the CP radius has an increased impact on the output uncertainty compared to the WB setup. Furthermore, several cells are triggered along the ridge in the orography setup compared to single cells in the WB and CP setups, which further adds to the differences in the results found for each trigger mechanisms.

For the precipitation variables the main effects in the WB setup and in the CP setup show a similar behavior where the uncertainty contributions from θ_0, the CCN concentration and the INP concentration dominate the overall uncertainties. The main effect of θ_0 is below 30% except for the total precipitation where it is above 45% in both setups. The contribution of the CCN concentration varies between 4% for the total precipitation and 31% for the mean amount of hail at the ground in the WB setup, whereas it varies between 18% and 48% in the CP setup. Further, the INP concentration accounts for about 6–33% of the overall uncertainty. Compared to θ_0, CCN concentration, and INP concentration, the remaining input parameters have only minor contributions. In the orography setup θ_0 does not contribute; however, there are additional contributions from the wind shear and the half width of the mountain ridge. Thus, the output uncertainties are composed of 11–50% CCN concentration, up to 14% INP concentration, 10–26% wind shear, and 20–30% half width of the ridge. These results lead to the conclusion that the properties of the orography lead to greater uncertainty in the output of the precipitation variables than the properties of the thermal perturbations.

Our findings are in good agreement with the work of Fan et al. (2013), Yang et al. (2017), and Lee et al. (2008). Fan et al. (2013) observed changes of 25% of the anvil expansion due to changes in the CCN concentration and Yang et al. (2017) found clear differences in the vertically integrated condensate mixing ratio such as an increase of ice from 6 to 18 g/kg for increasing CCN. Furthermore, the CAPE, thus the potential temperature

Table 2
Changes of the Parameter Ranges of the Uncertain Input Parameters when Related to Forecast Errors

input	min	max	units
wind shear (F_{shear})	0.3333	0.6666	scaling factor
θ_0 (WK profile)	299	301	K
height of ridge (Oro)	3000	4000	m
half width of ridge (Oro)	26500	28500	m

profile, and the wind shear were identified by Lee et al. (2008) to affect the amount of cloud condensate and precipitation.

4.3. Sensitivity Analysis Related to Typical Forecast Errors

So far we considered the whole predefined parameter space (cf. Table 1) for our analysis, which was chosen to cover a wide range of atmospheric conditions. For example, the potential temperature was varied over a range of 7 K. By definition the main effect indicates how much the variance in the output could be reduced if the input parameter was known exactly (Johnson et al., 2015). However, the forecast of the temperature is typically more accurate. Therefore, we now relate the parameter space used for the sensitivity analysis to the typical forecast errors of the NWP model COSMO in an operational setup for Central Europe. Table 2 gives an overview of the adapted input parameter ranges.

The ranges of wind shear and θ_0 are derived from their root mean square errors of the COSMO prediction for different lead times. The ranges for the height and the half width of the mountain ridge are estimated as the maximum change of terrain that is not resolved by the model grid. These new ranges are then centered around the central value of the original parameter ranges. The remaining input parameters such as CCN and INP concentrations are not part of the operational model forecast, and thus, their ranges cannot be restrained by model errors. Nevertheless, the trigger characteristics of the WB and the CP represent values used in other studies (Adams-Selin et al., 2013; Brooks, 1992; Brooks & Wilhelmson, 1992; James et al., 2006). Also, the chosen ranges of CCN and INP concentrations are found throughout Europe (Boose et al., 2016; Bougiatioti et al., 2009). The sensitivity analysis is run again using the same emulators with the parameter ranges given in Table 2. The results are shown in Figure 5. As the possible parameter range is now reduced to average forecast or resolution errors the overall composition of the bar plots changes. Particularly, the main effect of θ_0, which strongly affected most outputs in Figure 4 is considerably reduced.

In the WB setup (Figure 5a) the highest contribution of the potential temperature is found for the integrated graupel content where it reaches up to 14% compared to values of about 80% when the full parameter range was used. For all other variables the contribution is less than that. Thus, other input parameters emerge to explain the output variance. For the integrated condensates the CCN concentration is responsible for more than 40% of the uncertainty except for the graupel content. Furthermore, the INP concentration adds about 35% to the output uncertainty of the integrated hail, graupel, and snow contents being the largest contributor for the snow content. Noticeable uncertainty contributions of more than 30% from the initial temperature perturbation of the bubble and the wind shear can be found for the integrated graupel and rain contents, respectively. Regarding the precipitation variables, the WB characteristics and the wind shear are only of minor importance, instead the overall uncertainty is dominated by the CCN and INP concentrations. For the variables describing precipitation of hail, the CCN concentration contributes at least 49% of the uncertainty, whereas for the total precipitation the INP concentration is the major contributor (up to 48%). Additionally, the sum of the main effect is larger, which means that less interaction occurs.

In the CP setup (Figure 5b) the effect of θ_0 is clearly reduced compared to when the original parameter range was used. Instead, the variance of all of the integrated hydrometeors is dominated by the main effects of the CCN and INP concentrations and the radius of the CP. In contrast, the variance in the precipitation variables hardly depends on the uncertainty in the radius of the initial CP. For these variables it is the CCN and INP concentrations that contribute most to the overall output uncertainty.

In the orography setup the impact of the potential temperature is not reduced as much as in the WB and CP setups. Nevertheless, the CCN concentration is the largest uncertainty source varying between 19% and 80% except for the integrated cloud water and ice content where it is less than 3%. This negligible contribu-

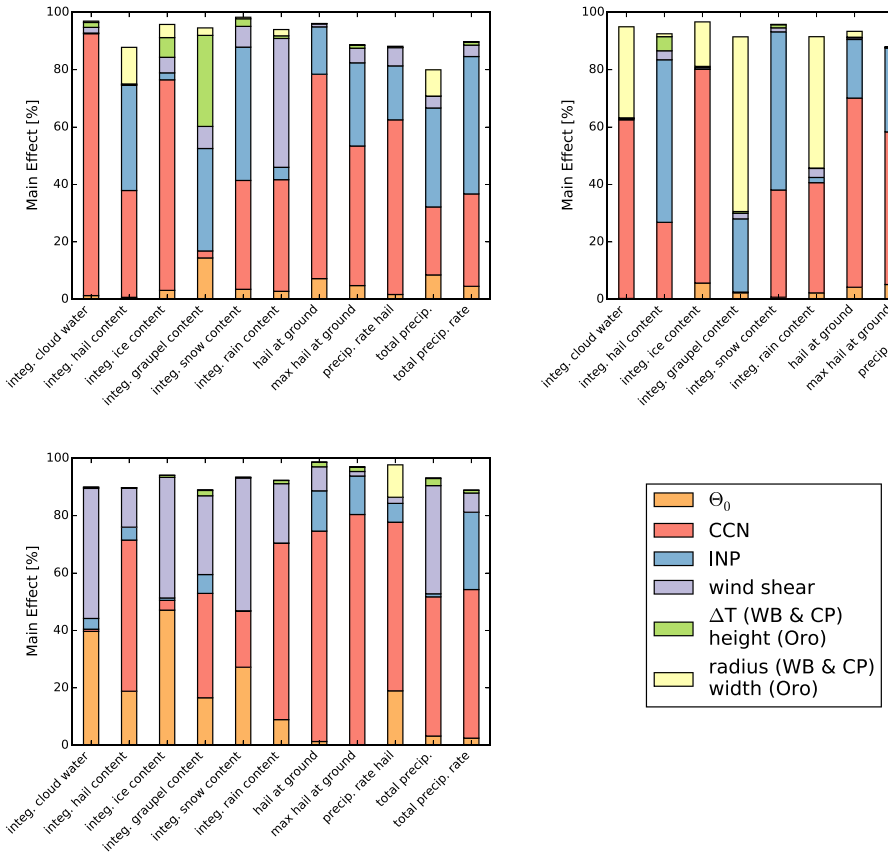

Figure 5. Bar plot of the main effect for vertically integrated hydrometeor masses and precipitation of cloudy grid points using (a) a warm bubble, (b) a cold pool, and (c) orography as trigger mechanism when the parameter space is restrained by mean forecast errors. (d) Interaction effect of the hydrometeors in the orography setup using the original parameter ranges. CCN = cloud condensation nuclei; CP = cold pool; INP = ice-nucleating particles; WB = warm bubble.

tion from the CCN concentration to the output uncertainty of cloud water stands out as both WB setup and CP setup show opposite results. An additional analysis reveals that the contribution from the CCN concentration is increased if the clouds over the ridge are included (not shown). Thus, the uncertainty in the CCN concentration mainly affects the cloud water during the formation of the cloud over the ridge, while its impact is reduced further downwind. For the precipitation output, a similar behavior arises: The individual parameters account for more of the output uncertainty here. In particular, the CCN and INP concentrations have larger effects.

4.4. Size Distribution of Hail

Regarding the damage potential of hail events, the size distribution of hailstones is of great importance. Thus, we additionally investigate the response of the surface size distribution of the hydrometeor class classified as hail by the COSMO model to different environmental conditions. Each uncertain input parameter is assigned two discrete values representing regimes within the parameter ranges with low and high values, respectively. The two regimes are denoted by "-" and "+." The values were chosen such that a reasonable difference occurs without the necessity of using the full parameter ranges and such that there is always a reasonable amount of wind shear present. The chosen values are given in Table 3.

For each setup and for all 64 possible combinations of these parameters, emulators are used to simulate the size distribution of hail. The emulators predict the number concentration at given particle diameters. Figure 6a shows the mean size distributions of surface hail for the WB setup where the shadings denote the areas in which the surface hail size distributions are found for the environmental conditions specified by the legends.

The maximum of the size distribution of surface hail is located at a diameter of 5 mm and an amount in the order of $\mathcal{O}(1)$ m^{-4}. The size distributions can be divided into three groups. By analyzing the input combinations of each group, three governing input parameters can be identified: the CCN concentration, the

Table 3
Input Values Representing Both Lower and Higher Regimes of the Parameter Ranges, Which Are Used to Analyze the Size Distribution of Hail

Input	Lower value (−)	Higher value (+)	Units
CCN concentration	500	3,000	cm^{-3}
INP concentration	0.1	10	scal. factor
Wind shear (F_{shear})	0.5	1.0	scal. factor
Potential temperature θ_0	298	302	K
Temperature perturbation ΔT (WB)	2	5	K
Temperature perturbation ΔT (CP)	−10	−6	K
Height of ridge (Oro)	2,000	5,000	m
Radius of warm bubble R_{hor}	7	13	km
Radius of cold pool R_{hor}	14	26	km
Half width of ridge (Oro)	14	41	km

Note. CCN = cloud condensation nuclei; CP = cold pool; INP = ice-nucleating particles; WB = warm bubble.

INP concentration, and the potential temperature θ_0. If the values of all CCN concentration, INP concentration, and θ_0 represent the higher values, the least amount of hail is produced. Conversely, the largest amount of hail is produced when CCN, INP, and θ_0 take the lowest values. Medium amounts of hail are produced for other combinations of the governing parameters. As in COSMO hail is either produced by freezing of raindrops or growth of graupel, a decrease of the CCN concentration leads to larger raindrops and faster growth of graupel and thus larger hail. However, between the cloud base and the ground large amounts of hail are lost due to melting. Therefore, large hailstone are necessary for a detection of surface hail such that low CCN and INP concentrations provide a higher number of hailstones at the ground. Presumably, high temperatures in the lower troposphere further enhance the melting process and accordingly, more hailstones are found for a low value of θ_0. To summarize, the least amount of hail at the ground is found for conditions with a high number of CCN and INP and rather warm temperatures with high CAPE, whereas few CCN and INP and colder conditions produce more hail when a WB is used as trigger mechanism.

The parameter main effects for the size distribution shown in the right plot of Figure 6b confirms the result found above. Mostly, θ_0 contributes to the output uncertainty of the number concentrations at the given diameters. In the medium size range with diameters between $2.5 \cdot 10^{-4}$ and $2.5 \cdot 10^{-3}$ m the potential temperature θ_0 is responsible for more than 80% of the output uncertainty. On the one hand the main effect of θ_0 decreases to roughly 20% toward larger diameters, and on the other hand the sum of the uncertainty contributions from the CCN and INP concentrations increases to a total of 46%. Contributions to the overall main effect of other input parameters are only of minor importance, and thus, the three parameters causing the largest uncertainty in the size distribution are the potential temperature θ_0 and the CCN concentration and the INP concentration.

When using a CP as trigger mechanism, the analysis of the size distribution yields similar results with some differences. As it can be seen from Figure 6c the maximum of the distribution is also found at a diameter of 5 mm. However, four groups of size distributions can be distinguished. Identical to the WB setup, a group of simulations identified by high CCN and INP concentrations and high temperatures can be separated toward lower number concentrations and a group characterized by low CCN and INP concentrations, and low temperatures can be separated toward higher number concentrations. In addition, the remaining size distributions can be divided into one group with medium high number concentrations related to low potential temperatures and one group with medium low number concentrations related to high potential temperatures. Figure 6d shows the main effects for the size distribution of surface hail for the CP setup. From a diameter of 0.25 mm to a diameter of 1 mm over 80% of the uncertainty is dominated by θ_0. From a diameter of 2.5 mm onward the influence of θ_0 gets smaller and the contributions of the CCN and INP concentrations start to grow reaching values of 44% and 12%, respectively. Compared to the WB setup these contributions of the CCN and INP concentrations are larger.

Results deviating from those of the WB and CP setups are found when a mountain ridge is used to trigger convection. Figure 6e shows the size distributions for each of the 64 combinations of the input parameters

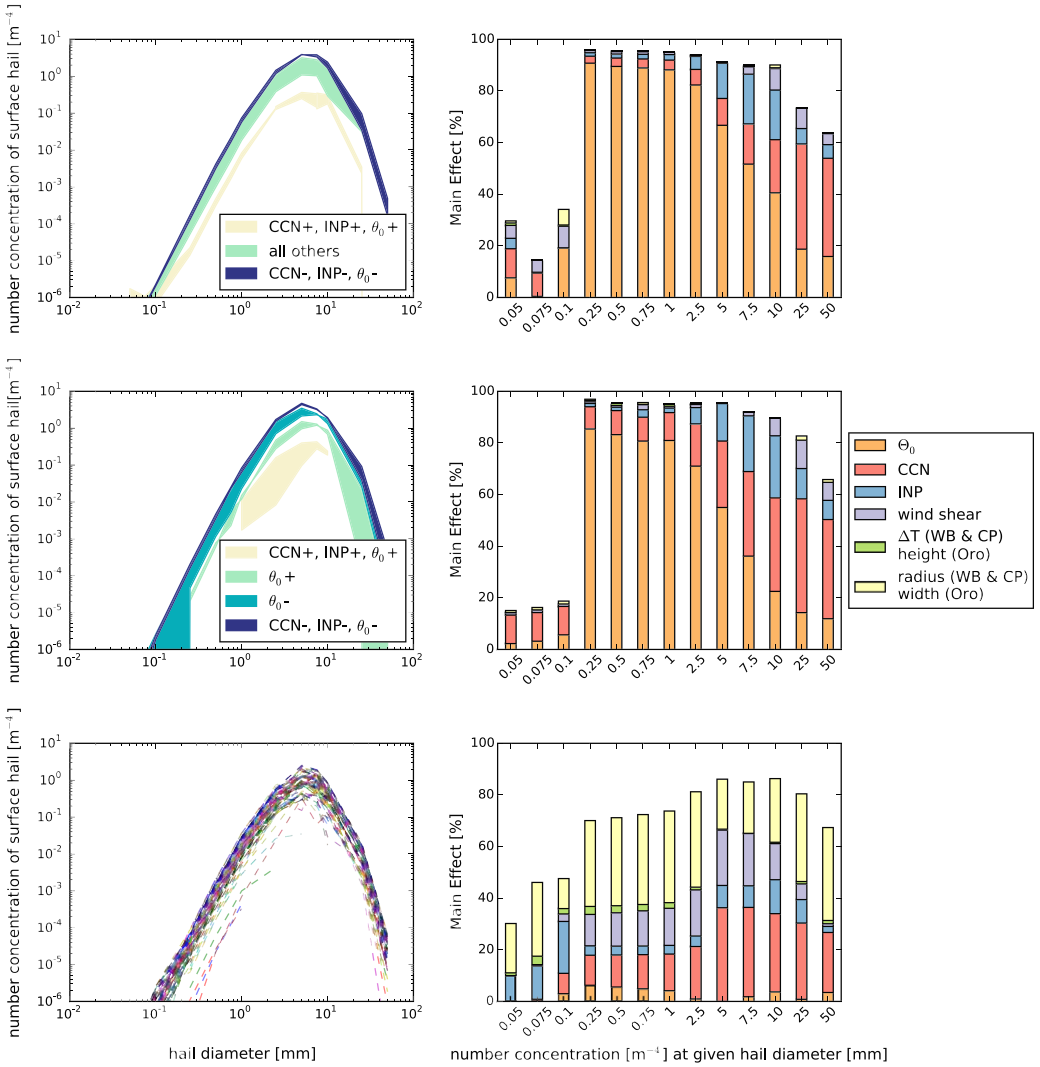

Figure 6. (left) Size distributions of hail at $z = 0$ m for all trigger mechanisms. Each line represents the distribution in one of the 64 environmental conditions. (right) Bar plots of the main effect for the mean size distribution of surface hail for all trigger mechanisms using the whole parameter range. The shading in (a) and (c) illustrates regimes of the size distributions controlled by CCN and INP concentrations and θ_0. CCN = cloud condensation nuclei; CP = cold pool; INP = ice-nucleating particles; WB = warm bubble.

stated in Table 3 and Figure 6f depicts the main effects. Compared to the WB and CP setups, the shape of the distributions is more diverse. Some of them peak at a diameter of 5 mm, whereas others peak at a diameter of 7.5 mm or have similar values at both diameters. A small number of simulations shows size distributions with lower number concentrations. However, these differences are only distinct in the range of the maximum, and thus, a clear separation into groups is not possible. The sum of the main effects in each case varies around 70%, which means that more parameter interactions occur in the orography setup as already seen in section 4.2. Several input parameters contribute almost equally to the output uncertainty: the CCN concentration ($\sim$ 20%), the INP concentration ($\sim$ 10%), the wind shear ($\sim$ 20%), and the half width of the mountain ridge ($\sim$ 20%). Furthermore, there are also minor contributions from θ_0 and the height of the ridge. The fact that various input parameters have similar large contributions to the uncertainty in the size distribution here prevents us from naming a specific key parameter in this case.

4.5. Sensitivity to CCN Concentration

There are many studies analyzing the cloud response and consequences of variations in CCN, for example, Khain et al. (2011), Morrison (2012), Noppel et al. (2010), and Yang et al. (2017). However, the results of these studies are not uniform, and thus, the relation between changes in the CCN concentration and the cloud

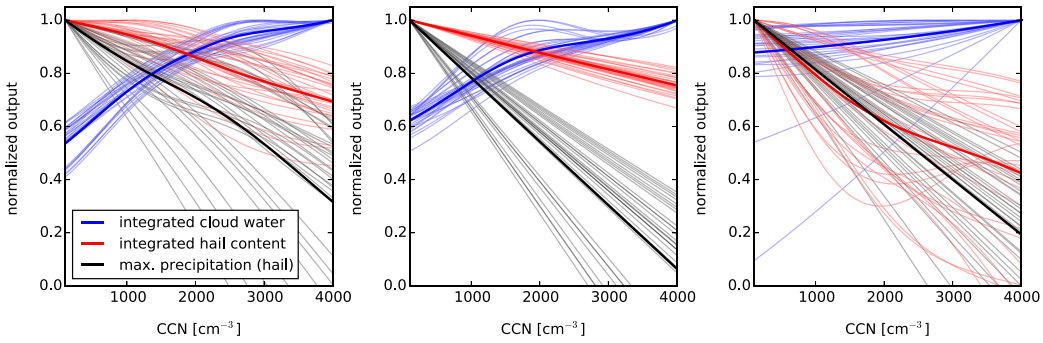

Figure 7. Sensitivity of mean integrated cloud water, mean integrated hail, and precipitation by hail to variations in the cloud condensation nuclei concentration for (a) the warm bubble setup, (b) the cold pool setup, and (c) the orography setup. Thin lines represent the results for 32 different environmental conditions (regimes), and the bold line denotes the mean. CCN = cloud condensation nuclei.

responses is not clearly determined. For example, on the one hand Khain et al. (2011) simulated a severe supercell that caused large hail up to 10 cm in diameter over southwest Germany for different CCN concentrations and found that an increase of the CCN concentration leads to an increase of hail fall. On the other hand Noppel et al. (2010) obtain in the simulations of the same storm an opposite result: An increase of the CCN concentration is mainly connected to a decrease of the precipitation of hail.

As the emulators require hardly any computational power, they are a convenient tool to run many simulations and investigate the cloud response to variations of the CCN concentration in different environmental conditions in our idealized setup. Similar to section 4.4 each input parameter except for the CCN concentration is assigned two values representing the lower and higher values of the parameter space. The chosen values are the same as in Table 3, while the parameter range of the CCN concentration is systematically sampled with a set of 10,000 points. All possible combinations of the five remaining input parameters generate 32 different environmental regimes for comparison. During the analysis we focus on the mean integrated cloud water, the mean integrated hail mass, and the maximum amount of precipitation by hail. Figure 7 shows the considered output versus the CCN concentration for all 32 regimes and for all trigger mechanisms. The output values are normalized with their respective maximum reached in each regime. For most of the considered regimes this maximum is reached at the lowest CCN concentration for the integrated hail content and the amount of hail at the ground and at the highest CCN concentration for the integrated amount of cloud water.

In general, the mean sensitivity of the three output variables to the CCN concentrations behaves similarly for all trigger mechanisms. The cloud water increases whereas both the in-cloud hail and the precipitation of hail decrease with increasing CCN concentration. When a WB is used to trigger convection, the integrated cloud water shows a strong increase for low CCN concentrations, which flattens toward higher concentrations. The integrated hail content decreases to approximately 70% of its maximum value when a polluted atmosphere is prescribed. However, the differences in hail content between the regimes are larger than for the cloud water as the minimal values of the hail content vary between 40% and 90% of the maximum value. Both the decrease and the spread are more pronounced for the precipitation of hail than for the integrated hail content. On average, the precipitation of hail is reduced to less than 40% when the CCN concentration is increased, while in some regimes the reduction is even stronger such that no hail is found at the ground.

From all trigger setups, the CP turns out to be most robust to changes of the environmental conditions. The cloud water increases from 60% to 100% and the hail content decreases linearly to 70% for CCN concentrations between 100 and 4, 000 cm^{-3}. Both variables show a maximum spread of about 20%. The precipitation of hail is reduced to values between 40% of the original value and no hail at all. Although the spread is higher for the hail reaching the ground, all atmospheric regimes exhibit a linear decrease.

In the orography setup large deviations from the mean trend are found. Beginning at values around 80% of the maximum the mean of the integrated cloud water increases slowly for low CCN concentrations and slightly faster for high concentrations. However, there are also regimes where a sensitivity to CCN is hardly visible. Two variations are possible for the trend of the integrated hail content. First, most of the simulations and thus the mean of all results are rapidly decreasing with increasing CCN concentration. Second, some regimes show a moderate increase of the hail content or are insensitive to changes of the CCN concentration between

approximately $2,000-3,000\ cm^{-3}$. In contrast, the results of the precipitation of hail are more uniform with linear decreases reaching the minimum value for the highest CCN concentrations.

Summarizing, we see that the sensitivity of hail on the ground, integrated cloud water, and hail content to the CCN concentration varies for each regime of environmental conditions, but its sign is never reversed. Especially in the orography setup large variations are possible, which reflects the complexity of convection above mountainous terrain (Kirshbaum et al., 2018). Still, the overall trend is the same for all considered trigger mechanisms, and in all cases the sensitivity to the CCN concentration is linked to an increase in the integrated cloud water and to a decrease in the integrated hail content and the maximum precipitation of hail. Thus, the results are robust for different environmental conditions and trigger mechanisms.

However, several studies note that there are problems concerning bulk microphysics schemes and hydrometeor sizes. For example, Dennis and Kumjian (2017) state that bin schemes are necessary to track changes of the hail size as bulk schemes do not treat the processes satisfactorily that affect the size distributions. Further, Lee et al. (2008) mention that the calculation of precipitation, and thus, also the hail at the ground is not as accurate in bulk models as in bin models. To approach this issue, Loftus and Cotton (2014) introduce a modified microphysics setup where a three-moment scheme is implemented for an improved prediction of hail. They find that an increase of the CCN concentration induces an increase of the number of hailstones, which corresponds to our findings where the CCN concentration is identified as controlling parameter of the size distribution as well. Since Loftus and Cotton (2014) investigate the effect of the CCN concentration only, the classification of controlling parameters of the size distribution of hail is assumably appropriate although a bulk microphysics scheme is implemented in the COSMO model. Further studies similar to Loftus and Cotton (2014) incorporating modifications of the microphysics scheme and the variation of not only one but several parameters are necessary to confirm these findings. Further, this means that the discrepancies between the findings of Khain et al. (2011) and Noppel et al. (2010) cannot be explained by different environmental conditions but rather by the differences of the microphysics schemes used in these studies.

5. Summary and Conclusions

In this study we have explored how variations in six different environmental conditions affect deep convective clouds. In particular we have explored integrated cloud water contents, surface precipitation, and hail size spectrum. Furthermore, we have investigated whether the results are robust by considering different trigger mechanisms of convection.

Using a variance-based sensitivity analysis technique, made possible through an implementation using statistical emulators, the output variance in model outputs has been decomposed into different components, which relate to changes of the environmental conditions (main effect) and their interactions (interaction effect). An analysis of the main effect sensitivity indices for each output has shown that for most output variables, the uncertainty in the potential temperature, which controls the CAPE in the initial profile, and the CCN concentration are the largest contributors to the overall output uncertainty. This means that the impact of processes depending on these parameters, such as the nucleation of cloud droplets, on the output uncertainty is larger compared to dynamical responses driven by wind shear, for example.

With the systematic covariation of different environmental conditions over a wide range, we also cover conditions that have been previously studied by smaller sets of sensitivity experiments. Storer et al. (2010), for example, analyzed the impact of CAPE and CCN concentration on deep convective clouds and also found that overall the clouds are quite sensitive to changes of CAPE. Further, they stated that some parameters such as the cloud water path do not strongly depend on CAPE similar to the results of our analysis. Another result of Storer et al. (2010) was that the effect of the CCN concentration is comparable to the effect of CAPE, which is consistent with our findings. Other studies such as those from Tao et al. (2007) and Fan et al. (2013) also showed a clear impact of the CCN concentration on the cloud properties and precipitation. Furthermore, the role of interaction effects changes and the composition of the parameter contributions to the output uncertainty for the different outputs differs for each of the trigger mechanisms.

However, the results of the sensitivity analysis depend on the ranges of the input parameters that were chosen to represent various atmospheric conditions. For a more applied approach, we subsequently restrained the ranges of the environmental conditions to mean prediction errors of the forecast of the COSMO model, and thereby, the focus has changed from a general analysis of environmental conditions to the composition of the

uncertainty of the actual model forecast. This analysis indicates that the CCN concentration is still important for the output uncertainty. Further controlling environmental conditions are the INP concentration in the WB and CP setups and the wind shear in the orography setup. Hence, a more precise representation of the vertical wind profile and an inclusion of observations of the aerosol concentration may improve the forecast error of the COSMO model.

To understand how the mean size distribution of surface hail is affected by environmental conditions and the choice of the convection trigger, we have defined 64 setups of distinct environmental conditions and have compared the corresponding size distributions for each trigger. It appears that the sensitivity of the surface hail size distribution to environmental conditions depends on the trigger mechanism as different input parameters impact the results. When a WB or CP is used, the uncertainty in the number concentration of surface hail is controlled by the CCN concentration, the INP concentration and the potential temperature θ_0. In contrast, no controlling parameters are identifiable in the orography setup as several parameters interact and determine the size distribution of hail. Overall, a shift of the size distributions toward smaller or larger hailstone diameters caused by the changes in the environment is not found.

Similarly, the sensitivity of integrated cloud water, integrated hail, and precipitation by hail to the CCN concentration has been examined in different environmental conditions. For this purpose the parameter range of the CCN concentration has been sampled with 10,000 points for which the output of cloud water, in-cloud hail and precipitation has been predicted by the emulators. The comparison of the results of the different environmental conditions and the three trigger mechanisms reveals that on average the sensitivities are robust for all considered setups. An increase of the CCN concentration is related to an increase of the integrated cloud water, whereas it is related to a decrease of the integrated hail content and the precipitation of hail. However, the exact curves vary depending on the environmental conditions. Especially in the orography setup, large deviations from the mean trend are possible for the different atmospheric regimes.

Summarizing, we see that for all of our analyses the results differ depending on the trigger mechanism with mainly the orography setup being different from the WB and CP setups. So far, sensitivity studies have usually been conducted using a single method to trigger convection (Adams-Selin et al., 2013; Chen & Lin, 2005; Storer et al., 2010) and thus the same sensitivities have not been compared for different trigger mechanisms. Our results indicate that sensitivities found for one trigger mechanism cannot be transferred to other triggers one to one. Part of the differences can be attributed to the structure of the initial convection in the orography setup where several cells are triggered along the mountain ridge compared to the WB and CP setups where single cells are triggered. However, there are also minor differences between the WB setup and the CP setup. This limits the transferability of identified sensitivities to real-case simulations, in which all three trigger mechanisms can happen in modified forms.

Concluding, statistical emulation enables a dense sampling of the whole parameter space in a short amount of time and thus a comprehensive sensitivity analysis. The quantification and comparison of contributions from input parameters to the output uncertainty would have not been possible without these methods. However, in the process information is lost because mean and maximum values are used in this analysis. In future studies this problem could be reduced by multivariate emulation (Hankin, 2012; Overstall & Woods, 2016), where multiple variables are fitted with a single emulator. However, such emulators can be more difficult to generate as further specifications must be made as to how the multiple output variables depend on each other, which are not always known. Furthermore, an emulator is not a replacement for a numerical model based on physical equations, yet the validation ensures that the specific relation between input and output described by each emulator is well represented. So far, we have considered deep convective clouds in idealized setups only, and the results depend on assumptions in the two-moment scheme such as the use of saturation adjustment. Thus, the sensitivities might vary for real-case simulations. Nevertheless, we are optimistic that our findings are also partly applicable to real events. Overall the emulator approach appears to be a powerful tool for the analysis of complex weather prediction models, which we recommend for further use.

Acknowledgments
The research leading to these results has been done within the project "Microphysical uncertainties in deep convective clouds and their implications for data assimilation" of the Transregional Collaborative Research Center SFB/TRR 165 "Waves to Weather" funded by the German Science Foundation (DFG). This work was performed on the computational resource ForHLR I funded by the Ministry of Science, Research and the Arts Baden-Württemberg and DFG. Furthermore, we thank Felix Fundel from Deutscher Wetterdienst for providing data on the prediction errors of the COSMO model. The processed training data set and the emulators are published (doi:10.5445/IR/1000085083). The full model simulations are available upon request.

References

Adams-Selin, R. D., van den Heever, S. C., & Johnson, R. H. (2013). Impact of graupel parameterization schemes on idealized bow echo simulations. *Monthly Weather Review*, *141*(4), 1241–1262. https://doi.org/10.1175/MWR-D-12-00064.1

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the cosmo model: Description and sensitivities. *Monthly Weather Review*, *139*(12), 3887–3905. https://doi.org/10.1175/MWR-D-10-05013.1

Bastos, L. S., & O'Hagan, A. (2009). Diagnostics for Gaussian process emulators. *Technometrics*, *51*(4), 425–438. https://doi.org/10.1198/TECH.2009.08019

Bigg, E. K. (1953). The formation of atmospheric ice crystals by the freezing of droplets. *Quarterly Journal of the Royal Meteorological Society*, *79*(342), 510–519. https://doi.org/10.1002/qj.49707934207

Boose, Y., Kanji, Z. A., Kohn, M., Sierau, B., Zipori, A., Crawford, I., et al. (2016). Ice nucleating particle measurements at 241 K during winter months at 3580 m msl in the Swiss Alps. *Journal of the Atmospheric Sciences*, *73*(5), 2203–2228. https://doi.org/10.1175/JAS-D-15-0236.1

Bougiatioti, A., Fountoukis, C., Kalivitis, N., Pandis, S. N., Nenes, A., & Mihalopoulos, N. (2009). Cloud condensation nuclei measurements in the marine boundary layer of the Eastern Mediterranean: CCN closure and droplet growth kinetics. *Atmospheric Chemistry and Physics*, *9*(18), 7053–7066. https://doi.org/10.5194/acp-9-7053-2009

Brooks, H. (1992). Operational implications of the sensitivity of modeled thunderstorms to thermal perturbations. *Fourth AES/CMOS Workshop on Operational Meteorology*, *110*, 398–407.

Brooks, H. E., & Wilhelmson, R. B. (1992). Numerical simulation of a low-precipitation supercell thunderstorm. *Meteorology and Atmospheric Physics*, *49*(1), 3–17. https://doi.org/10.1007/BF01025398

Chen, S.-H., & Lin, Y.-L. (2005). Effects of moist Froude number and cape on a conditionally unstable flow over a mesoscale mountain ridge. *Journal of the Atmospheric Sciences*, *62*(2), 331–350. https://doi.org/10.1175/JAS-3380.1

Cui, Z., Carslaw, K. S., Yin, Y., & Davies, S. (2006). A numerical study of aerosol effects on the dynamics and microphysics of a deep convective cloud in a continental environment. *Journal of Geophysical Research*, *111*, D05201. https://doi.org/10.1029/2005JD005981

Cui, Z., Davies, S., Carslaw, K. S., & Blyth, A. M. (2011). The response of precipitation to aerosol through riming and melting in deep convective clouds. *Atmospheric Chemistry and Physics*, *11*(7), 3495–3510. https://doi.org/10.5194/acp-11-3495-2011

DeMott, P. J., Prenni, A. J., Liu, X., Kreidenweis, S. M., Petters, M. D., Twohy, C. H., et al. (2010). Predicting global atmospheric ice nuclei distributions and their impacts on climate. *Proceedings of the National Academy of Sciences*, *107*(25), 11,217–11,222. https://doi.org/10.1073/pnas.0910818107

Dennis, E. J., & Kumjian, M. R. (2017). The impact of vertical wind shear on hail growth in simulated supercells. *Journal of the Atmospheric Sciences*, *74*(3), 641–663. https://doi.org/10.1175/JAS-D-16-0066.1

Doms, G., & Baldauf, M. (2015). *A description of the nonhydrostatic regional COSMO-model, Part I: Dynamics and numerics*. Offenbach, Germany: Deutscher Wetterdienst (DWD).

Fan, J., Leung, L. R., Rosenfeld, D., Chen, Q., Li, Z., Zhang, J., & Yan, H. (2013). Microphysical effects determine macrophysical response for aerosol impacts on deep convective clouds. *Proceedings of the National Academy of Sciences*, *110*(48), E4581–E4590. https://doi.org/10.1073/pnas.1316830110

Fan, J., Tianle, Y., Jennifer, M. C., Steven, G., Alexander, K., Ruby, L. L., et al. (2009). Dominant role by vertical wind shear in regulating aerosol effects on deep convective clouds. *Journal of Geophysical Research*, *114*, D22206. https://doi.org/10.1029/2009JD012352

Hankin, R. (2012). Introducing multivator: A multivariate emulator. *Journal of Statistical Software, Articles*, *46*(8), 1–20. https://doi.org/10.18637/jss.v046.i08

Houze, R. A. (1993). *Cloud dynamics, International Geophysics Series* (pp. 53). San Diego, CA: Academic Press.

Huffman, P. J., & Vali, G. (1973). The effect of vapor depletion on ice nucleus measurements with membrane filters. *Journal of Applied Meteorology*, *12*(6), 1018–1024. https://doi.org/10.1175/1520-0450(1973)012<1018:TEOVDO>2.0.CO;2

James, R. P., Markowski, P. M., & Fritsch, J. M. (2006). Bow echo sensitivity to ambient moisture and cold pool strength. *Monthly Weather Review*, *134*(3), 950–964. https://doi.org/10.1175/MWR3109.1

Johnson, J. S., Cui, Z., Lee, L. A., Gosling, J. P., Blyth, A. M., & Carslaw, K. S. (2015). Evaluating uncertainty in convective cloud microphysics using statistical emulation. *Journal of Advances in Modeling Earth Systems*, *7*, 162–187. https://doi.org/10.1002/2014MS000383

Khain, A., Rosenfeld, D., Pokrovsky, A., Blahak, U., & Ryzhkov, A. (2011). The role of CCN in precipitation and hail in a mid-latitude storm as seen in simulations using a spectral (bin) microphysics model in a 2D dynamic frame. *Atmospheric Research*, *99*(1), 129–146. https://doi.org/10.1016/j.atmosres.2010.09.015

Kirshbaum, D. J., Adler, B., Kalthoff, N., Barthlott, C., & Serafin, S. (2018). Moist orographic convection: Physical mechanisms and links to surface-exchange processes. *Atmosphere*, *9*(3), 80. https://doi.org/10.3390/atmos9030080

Kunz, M., Blahak, U., Handwerker, J., Schmidberger, M., Punge, H. J., Mohr, S., et al. (2018). The severe hailstorm in SW Germany on 28 July 2013: Characteristics, impacts, and meteorological conditions. *Quarterly journal of the Royal Meteorological Society*, *144*(710), 231–250. https://doi.org/10.1002/qj.3197

Lee, L. A., Carslaw, K. S., Pringle, K. J., Mann, G. W., & Spracklen, D. V. (2011). Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmospheric Chemistry and Physics*, *11*(23), 12,253–12,273. https://doi.org/10.5194/acp-11-12253-2011

Lee, S. S., Donner, L. J., Phillips, V. T. J., & Ming, Y. (2008). The dependence of aerosol effects on clouds and precipitation on cloud system organization, shear and stability. *Journal of Geophysical Research*, *113*, D16202. https://doi.org/10.1029/2007JD009224

Lee, L. A., Pringle, K. J., Reddington, C. L., Mann, G. W., Stier, P., Spracklen, D. V., et al. (2013). The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei. *Atmospheric Chemistry and Physics*, *13*(17), 8879–8914. https://doi.org/10.5194/acp-13-8879-2013

Loeppky, J. L., Sacks, J., & Welch, W. J. (2009). Choosing the sample size of a computer experiment: A practical guide. *Technometrics*, *51*(4), 366–376. https://doi.org/10.1198/TECH.2009.08040

Loftus, A. M., & Cotton, W. R. (2014). Examination of CCN impacts on hail in a simulated supercell storm with triple-moment hail bulk microphysics. *Atmospheric Research*, *147*, 183–204. https://doi.org/10.1016/j.atmosres.2014.04.017

Morris, M. D., & Mitchell, T. J. (1995). Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference*, *43*(3), 381–402. https://doi.org/10.1016/0378-3758(94)00035-T

Morrison, H. (2012). On the robustness of aerosol effects on an idealized supercell storm simulated with a cloud system-resolving model. *Atmospheric Chemistry and Physics*, *12*(16), 7689–7705. https://doi.org/10.5194/acp-12-7689-2012

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. *Monthly Weather Review*, *137*(3), 991–1007. https://doi.org/10.1175/2008MWR2556.1

Noppel, H., Blahak, U., Seifert, A., & Beheng, K. D. (2010). Simulations of a hailstorm and the impact of CCN using an advanced two-moment cloud microphysical scheme. *Atmospheric Research*, *96*(2), 286–301. https://doi.org/10.1016/j.atmosres.2009.09.008

O'Hagan, A. (2004). *Kendalls advanced theory of statistics, Bayesian Inference* (Vol. 2B, 2nd ed.). Arnold, London: Halsted Press.

O'Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety*, *91*(10), 1290–1300. https://doi.org/10.1016/j.ress.2005.11.025

Oakley, J. E., & O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: A Bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *66*(3), 751–769. https://doi.org/10.1111/j.1467-9868.2004.05304.x

Overstall, A. M., & Woods, D. C. (2016). Multivariate emulation of computer simulators: Model selection and diagnostics with application to a humanitarian relief model. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *65*(4), 483–505. https://doi.org/10.1111/rssc.12141

Pruppacher, H. R., & Klett, J. D. (1997). *Microphysics of clouds and precipitation, Atmospheric and oceanographic sciences library* (Vol. 18). Dordrecht, Netherlands: Kluwer Academic Publishers.

Pujol, G., Iooss, B., Khalid Boumhaout, A. J., Veiga, S. D., Delage, T., Fruth, J., et al. (2017). Sensitivity: Global sensitivity analysis of model outputs. R package version 1.15.0.

R Core Team (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rasmussen, C. E. (2004). Gaussian processes in machine learning. In O. Bousquet, U. von Luxburg, & G. Rätsch (Eds.), *Advanced lectures on machine learning: ML summer schools 2003, Canberra, Australia, February 2–14, 2003, Tübingen, Germany, August 4–16, 2003, Revised Lectures* (pp. 63–71). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-28650-9_4

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. London: MIT Press.

Rosenfeld, D., Lohmann, U., Raga, G. B., O'Dowd, C. D., Kulmala, M., Fuzzi, S., et al. (2008). Flood or drought: How do aerosols affect precipitation? *Science*, *321*(5894), 1309–1313. https://doi.org/10.1126/science.1160606

Saltelli, A. H. (2008). *Sensitivity analysis Edited by* paperback, Wiley paperback series. Chichester: Wiley.

Saltelli, A., Tarantola, S., & Chan, K. P.-S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, *41*(1), 39–56. https://doi.org/10.1080/00401706.1999.10485594

Schättler, U., Doms, G., & Schraff, C. (2016). A description of the nonhydrostatic regional COSMO-model. Part VII: User's guide.

Schlesinger, R. E. (1978). A three-dimensional numerical model of an isolated thunderstorm: Part I. comparative experiments for variable ambient wind shear. *Journal of the Atmospheric Sciences*, *35*(4), 690–713. https://doi.org/10.1175/1520-0469(1978)035<0690:ATDNMO>2.0.CO;2

Segal, Y., & Khain, A. (2006). Dependence of droplet concentration on aerosol conditions in different cloud types: Application to droplet concentration parameterization of aerosol conditions. *Journal of Geophysical Research*, *111*, D15204. https://doi.org/10.1029/2005JD006561

Seifert, A., & Beheng, K. D. (2006a). A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteorology and Atmospheric Physics*, *92*(1), 45–66. https://doi.org/10.1007/s00703-005-0112-4

Seifert, A., & Beheng, K. D. (2006b). A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 2: Maritime vs. continental deep convective storms. *Meteorology and Atmospheric Physics*, *92*(1), 67–82. https://doi.org/10.1007/s00703-005-0113-3

Storer, R. L., van den Heever, S. C., & Stephens, G. L. (2010). Modeling aerosol impacts on convective storms in different environments. *Journal of the Atmospheric Sciences*, *67*(12), 3904–3915. https://doi.org/10.1175/2010JAS3363.1

SwissRe (2014). Sigma: Natural catastrophes and man-made disasters in 2013. Zurich: Swiss Re Economic Research & Consulting.

Tao, W., Chen, J., Li, Z., Wang, C., & Zhang, C. (2012). Impact of aerosols on convective clouds and precipitation. *Reviews of Geophysics*, *50*, RG2001. https://doi.org/10.1029/2011RG000369

Tao, W., Li, X., Khain, A., Matsui, T., Lang, S., & Simpson, J. (2007). Role of atmospheric aerosol concentration on deep convective precipitation: Cloud resolving model simulations. *Journal of Geophysical Research*, *112*, D24S18. https://doi.org/10.1029/2007JD008728

Weisman, M., & Klemp, J. (1982). The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Monthly Weather Review*, *110*, 504–520.

Weisman, M. L., & Klemp, J. B. (1984). The structure and classification of numerically simulated convective stormsin directionally varying wind shears. *Monthly Weather Review*, *112*(12), 2479–2498. https://doi.org/10.1175/1520-0493(1984)112<2479:TSACON>2.0.CO;2

Weisman, M. L., & Rotunno, R. (2000). The use of vertical wind shear versus helicity in interpreting supercell dynamics. *Journal of the Atmospheric Sciences*, *57*(9), 1452–1472. https://doi.org/10.1175/1520-0469(2000)057<1452:TUOVWS>2.0.CO;2

Weisman, M. L., Skamarock, W. C., & Klemp, J. B. (1997). The resolution dependence of explicitly modeled convective systems. *Monthly Weather Review*, *125*(4), 527–548. https://doi.org/10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2

Yang, H., Xiao, H., Guo, C., Wen, G., Tang, Q., & Sun, Y. (2017). Comparison of aerosol effects on simulated spring and summer hailstorm clouds. *Advances in Atmospheric Sciences*, *34*(7), 877–893. https://doi.org/10.1007/s00376-017-6138-y