Assessing the uncertainty of glacier mass-balance simulations in the European Arctic based on variance decomposition

T. Sauter\(^1\)\(^2\) and F. Obleitner\(^1\)

\(^1\)Institute of Meteorology and Geophysics, University of Innsbruck, Innsbruck, Austria
\(^2\)Institute of Geography, University of Erlangen-Nürnberg, Erlangen, Germany

Correspondence to: T. Sauter (tobias.sauter@uibk.ac.at)

Received: 8 January 2015 – Published in Geosci. Model Dev. Discuss.: 13 March 2015
Revised: 3 August 2015 – Accepted: 23 November 2015 – Published: 10 December 2015

Abstract. State-of-the-art numerical snowpack models essentially rely on observational data for initialization, forcing, parametrization, and validation. Such data are available in increasing amounts, but the propagation of related uncertainties in simulation results has received rather limited attention so far. Depending on their complexity, even small errors can have a profound effect on simulations, which dilutes our confidence in the results. This paper aims at quantification of the overall and fractional contributions of some archetypical measurement uncertainties on snowpack simulations in arctic environments. The sensitivity pattern is studied at two sites representing the accumulation and ablation area of the Kongsvegen glacier (Svalbard), using the snowpack scheme Crocus. The contribution of measurement errors on model output variance, either alone or by interaction, is decomposed using global sensitivity analysis. This allows one to investigate the temporal evolution of the fractional contribution of different factors on key model output metrics, which provides a more detailed understanding of the model’s sensitivity pattern. The analysis demonstrates that the specified uncertainties in precipitation and long-wave radiation forcings had a strong influence on the calculated surface-height changes and surface-energy balance components. The model output sensitivity patterns also revealed some characteristic seasonal imprints. For example, uncertainties in long-wave radiation affect the calculated surface-energy balance throughout the year at both study sites, while precipitation exerted the most influence during the winter and at the upper site. Such findings are valuable for identifying critical parameters and improving their measurement; correspondingly, updated simulations may shed new light on the confidence of results from snow or glacier mass- and energy-balance models. This is relevant for many applications, for example in the fields of avalanche and hydrological forecasting.

1 Introduction

Snow is a key component of the earth system, and it has a vital importance for the structure and dynamics of the atmospheric boundary layer by modifying, for example, the exchange processes between the atmosphere and the underlying ground. Bridging the gap between the inherent microphysical snow processes and the exchange processes at the snow surface, still constitutes major challenges to scientists.

Sophisticated snowpack models summarize the present knowledge and prove themselves to be a useful tool in simulating the spatial and temporal evolution of snowpacks. Thus, snow models have been successfully adopted for avalanche forecasting (e.g. Bellaire et al., 2013; Durand et al., 1999; Lehning et al., 1999), glacier modelling (e.g. Obleitner and Lehning, 2004; Gallée et al., 2001), hydrological research (e.g. Magnusson et al., 2014; Lehning et al., 2006; Liston and Elder, 2006; Bernhardt et al., 2010), and climate impact studies (e.g. Durand et al., 2009). The currently used snow models can be roughly classified by their degree of complexity, ranging from simplified bulk or single-layer models to detailed physical snowpack models (Etchevers et al., 2004; Feng et al., 2008; Rutter et al., 2009). In general, the development and use of higher-order models also induces a need for more and better data to constrain the initialization, forcing, parametrizations, and validation of the simulations. However, the quality of relevant data (model and observations) is still difficult to assess. In that sense, the true value
of a measured quantity remains a rather theoretical concept and can often not be determined. One therefore usually estimates a range of values within which the true value is likely to fall. Probability density functions are widely recognized as appropriate measures for describing the uncertainty of input data and model parameters, and they are used in this study (see Sect. 3.2). In practise, however, these can often not be specified in a straightforward manner due to the complex nature of the measurement errors. It is nevertheless a major methodical issue to account for best estimated measurement errors, which allows scientists to objectively quantify their impact on the model outcome, and to provide information on the robustness of the results. A corresponding approach is based on Monte Carlo methods considering randomly drawn samples for each input factor from previously derived distribution functions. First- and higher-moment statistics can be computed to quantify the integrated model uncertainty. In this context, integrated is understood as the total effect of all measurement or parameter uncertainties on the model’s variability. At this point, there is still no information on how uncertainty in the model output can be assigned to different sources of uncertainty in the input data set or parameter setting. For example, interaction effects make it difficult to unambiguously allocate the uncertainty of model parameters and forcing data on the model’s variance. To achieve a full understanding of the sensitivity pattern of highly interconnected and non-linear models, such as sophisticated snow models, it is necessary to decompose the complete variance into its specific influences due to the arctic environmental conditions. However, it may be helpful to demonstrate the benefits of the applied method, to identify critical issues concerning model input and parametrization and to establish future priorities in corresponding research. An attractive approach for estimating sensitivity measures independently of the degree of linearity (model free) is based on global sensitivity analysis (GSA), which is introduced in Sect. 2.4. We then developed reference runs, which are validated by key observations at the two glacier sites. Based on these reference runs and the specification of the uncertainties of key variables and parameters, we performed Monte Carlo simulations. The results are presented in Sect. 3.1 and are mainly discussed regarding the impact of key drivers in terms of first- and total-order indices and inherent limitations, as well as regarding differences concerning the two sites at the glacier.

2 Data and methods

2.1 Crocus model set-up

Crocus is a physically based finite-element and one-dimensional multilayer snow scheme implemented in the land-surface model ISBA of the modelling platform SURFEX. The model is extensively described elsewhere (Vionnet et al., 2012; Brun et al., 1992), and we therefore simply provide a basic description and note modifications important for this study.

Snow is considered a porous material, whose properties are described through bulk physical properties (thickness, density, temperature, and liquid water content) and basic microstructure characteristics – dendricity, sphericity, grain size, and snow grain history. The parameter of snow grain history indicates whether there once was liquid water or faceted crystals in the layer (Brun et al., 1992; Vionnet et al., 2012). The changes in the morphological shape of snow crystals depend on snow metamorphism in response to atmospheric forcing and internal processes. To adequately treat the internal processes, the model employs a number of parametrizations derived from specific field and laboratory experiments (Brun et al., 1989). The governing equations are numerically solved in a vertical domain with space- and time- varying grid distances (which are necessary in order to
cope with accumulation or settling processes). The model is forced by the basic meteorological parameters (air temperature, humidity, wind speed, and precipitation rate as well as incoming solar and long-wave radiation) and is initialized by the vertical profiles of the key physical properties of snow and its underlying substrate. Model outputs comprise the vertical profiles of the snow temperature, density, liquid water content and structure parameters, and prognostic time series of surface temperature, snow depth, and energy- and mass-balance components, the latter two being coupled. Following, for example, Armstrong and Brun (2008), the model treats layerwise changes of internal energy according to

\[
\text{SEB} = \text{NR} + \text{SHF} + \text{LHF} + R + G,
\]

(1)

\[
\text{SEB} = L_{li}(R_t - R_M) - \int_{z=0}^{HS} \left[ \frac{d}{dt} \left( \rho_c c_p T_z \right) \right] \, dz.
\]

(2)

Therein, the surface-energy balance (SEB) denotes the surface-energy budget, that is the sum of net radiation (NR), the turbulent fluxes of sensible (SHF) and latent heat (LHF), the heat transferred by precipitation and blowing snow (R), and by conduction from the underlying material (G; glacier ice in our case). The associated changes in energy can be used for changes in the cold content of the snowpack throughout its total depth (HS) (second term in Eq. 2) or phase changes (melt or freeze; first term in Eq. 2). \( R_t \) and \( R_M \) are the freezing and melting rate, \( L_{li} \) is the latent heat of the fusion of ice (3.34 \times 10^3 \text{J} \text{kg}^{-1}), \( c_p \) is the specific heat capacity of ice (2.1 \times 10^3 \text{J} \text{kg}^{-1} \text{K}^{-1}), and \( \rho_c \) and \( T_z \) denote the density and snow temperature at depth \( z \), respectively.

Net radiation itself is composed of the sum of incoming and outgoing solar- and long-wave radiation. Crocus treats solar radiation in three spectral bands ([0.3–0.8], [0.8–1.5], and [1.5–2.8] \mu m). For each band, the spectral albedo is computed as a function of the snow properties (microstructure), and the incoming radiation in each band is then depleted as a function of the spectral albedo. The remaining energy penetrates into the snowpack and is assumed to decay exponentially with snow depth. Turbulent fluxes are parametrized following the standard micrometeorological framework based on the Monin–Obukhov similarity theory, which employs a bulk-transfer approach and a stability correction for stable stratification of the atmospheric surface layer.

Layerwise changes in internal energy induce either-varying cold content (warming/cooling; second term of Eq. 2) or phase changes of individual snow layers (Eq. 2, first term). Precipitation \((P)\), meltwater refreezing and/or sublimation/evaporation rates \((E)\) as well as runoff \((R_{\text{runoff}})\) couple the energy and the mass budget of a snowpack according to

\[
\frac{dM}{dt} = P \pm E - R_{\text{runoff}}.
\]

(3)

The key parameters of this coupled system will also be addressed in this study. Crocus has not yet been applied to Kongsvegen. The following paragraphs summarize the main modifications and set-up used in this study in order to develop reference runs properly reproducing the seasonal evolution of snow and ice at the two glacier sites.

**Water flow and refreezing.** The formation of superimposed ice is an important factor for the mass balance of arctic glaciers, and an appropriate treatment is also important in this study. We refer to superimposed ice as all water that percolates through the snowpack and refreezes on the glacier surface (Wright et al., 2007; Brandt et al., 2008; König et al., 2002). Obleitner and Lehning (2004) and Karner et al. (2013) showed that on the Kongsvegen glacier, the superimposed ice layer can reach a thickness of several decimetres in some years. The water percolation and refreezing routine in the current Crocus version basically simulates the gravitational water flow through the snowpack (Gascon et al., 2014). The energy available for refreezing is calculated at the beginning of each iteration step. If the snow layer temperature is below the melting point, water refreezes and the residual liquid water is retained up to a maximum holding capacity, which is difficult to determine. Default Crocus assumes a value of 5% of the pore volume and reproduces an increase of the average density of layers affected by the refreezing of water (Vionnet et al., 2012). This implementation is appropriate for the simulation of snow evolution in Alpine terrain but it fails to reproduce the formation of superimposed ice because water can percolate through glacier ice as well. To overcome the issue, all water exceeding the maximum liquid water holding capacity at an impermeable snow–ice interface is assumed to contribute to the runoff, and the water flow to the next layer is simply set to zero. This prevents percolation of water into the glacier ice and increases the refreezing potential at the snow–ice interface. This approach has been successfully applied in a similar setting using a different snow model (Obleitner and Lehning, 2004).

**Model input/output.** The Crocus model is forced by air temperature \((T)\), relative humidity \((\text{RH})\), wind speed \((U)\), incoming shortwave radiation \((\text{SW})\), incoming long-wave radiation \((\text{LW})\), \( P \), and atmospheric pressure (see Sect. 2.2). These time-dependent parameters were measured at both sites and are provided to the model by NetCDF file for hourly time steps. The input file was modified to include the roughness length for momentum as well as the fraction of total pore volume used to calculate the maximum holding capacity. Constant model parameters are provided by a static option file describing the initial and boundary conditions and the basic model parameters.

**2.2 Input data**

To run Crocus, we use meteorological and glaciological observations from two sites at the Kongsvegen glacier, located
in north-western Svalbard. The Kongsvegen glacier currently covers a total surface area of \(\sim 100 \text{ km}^2\) and extends over a total length of 26 km. From the highest point (750 m a.s.l.) in the east, the glacier flows towards the north-west coast of the archipelago. Several automatic weather stations were operated along the flow line of the glacier. The study makes use of two of them: KNG8 (78.75\(^\circ\) N, 13.33\(^\circ\) E; 668 m a.s.l.), located in the accumulation zone, and KNG1 (78.84\(^\circ\) N, 12.66\(^\circ\) E; 162 m a.s.l.), located in the ablation zone (see Fig. 1). Due to computational limitations, we had to restrict our simulations and error analysis to a 1-year period. The stations are equipped with state-of-the-art sensors for air temperature (unventilated), relative humidity, wind speed, and direction as well as shortwave and long-wave radiation components (see Table 2). Surface-height changes were measured by an ultrasonic ranger. Comprehensive quality control of the recorded data was performed following the method of Karner et al. (2013). The data have been further corrected.

**Filling data gaps.** For shorter gaps, the missing values were estimated by linear regression from surrounding stations, where it was possible. When the surrounding stations had missing values, the values were estimated by a stochastic nearest-neighbour resampling conditioned on the remaining variables (Beersma and Buishand, 2003). This was achieved by first calculating the Euclidean distance between the present-day and all other days without gaps. Then the missing value was replaced by randomly drawing out one of the 20 most analogous days. This approach is convenient for small gaps and guarantees physically consistent fields.

**Conversion of snow depth changes to water equivalent.** Snow precipitation rates were calculated offline from surface-height changes measured by the ultrasonic ranger, and converted to snow water equivalent (SWE) for input to the model. The density of freshly fallen snow \(\rho_{\text{new}}\) was calculated according to the default parametrization used by Crocus, which is a function of wind speed \(U\), and air temperature \(T_{\text{air}}\), given as

\[
\rho_{\text{new}} = a_\rho + b_\rho \cdot (T_{\text{air}} - 273.16) + c_\rho \cdot \sqrt{U},
\]

where \(a_\rho = 300 \text{ kg m}^{-3}\), \(b_\rho = 6 \text{ kg m}^{-3} \text{ K}^{-1}\), and \(c_\rho = 26 \text{ kg m}^{-3/2} \text{ s}^{-1/2}\) (Brun et al., 1992; Vionnet et al., 2012). Note, that the default value for \(a_\rho\) is set to 109 kg m\(^{-3}\) (Pahaut, 1976). The modification of this parameter accounts for the systematic underestimation of simulated settling and compaction of the near-surface snow layers compared to repeated snow pit observations. The latter reveal that the mean density of the near-surface snow layers is usually in the range of 100–200 kg m\(^{-3}\). It was further necessary to reduce the amount of noise in the original snow records in order to avoid erratic precipitation events, which lead to unrealistically high accumulation. The main factors that affect the sensor signal are blowing snow, intense snowfall, uneven snow surfaces, extreme temperatures, and snow crystal type. Blowing and drifting snow are frequent processes in the European Arctic and often result in the formation of sastrugi, which introduce additional surface variability not associated with precipitation events (Sauter et al., 2013). In principle, the associated small-scale variability can be usually reduced by moving average filter, but the very different event durations sometimes make it difficult to determine an appropriate subset size. We therefore decided to take the mean saltation trajectory height as a criterion of the uncertainty, which is assumed to be proportional to the surface shear stress \(u_\text{s}^2\) [m\(^2\) s\(^{-2}\)] (Pomeroy and Gray, 1990):

\[
h_{\text{salt}} = \frac{1.6 \cdot u_\text{s}^2}{2 \cdot g},
\]

where \(g\) (m s\(^{-2}\)) is the gravitational acceleration. The surface-shear stress has been estimated by assuming a logarithmic wind profile and an arbitrarily chosen constant roughness length of \(z_0 = 0.02\) m. Finally, snow depth variations smaller than \(0.8 \cdot h_{\text{salt}}\) were considered as noise. The factors \(z_0\) and 0.8 were used for calibration and to determine how much signal was removed from the original time series. At KNG8, for example, this procedure yields a simulated end-winter snow accumulation, which is well validated by independent stake observations.

**Large amplitude spikes.** Large amplitude data spikes in recorded snow depth can occur during intense snowfall events when snow particles obstruct the propagation of the sensor signal (ultrasonic pulses). Sudden snow depth changes in excess of 50 mm h\(^{-1}\) are assumed to belong to this class of events, and were simply ignored. Transition from rain to snow was assumed to take place in the range from 0 to 1 °C, with half the precipitation falling as snow, and the other half...
as rain. There was no direct information available to better constrain this threshold. Input of calculated changes in precipitation water equivalent are considered as part of the calibration procedure of the reference runs and yield overall satisfactory reproduction of the independently observed end of winter snow height, that is accumulation at both sites.

2.3 Reference run set-ups

Table 2. Specification of basic model input uncertainties and assigned probability density functions. The Sobol sequence has been generated from the distributions given in the last column, where $N(\mu, \sigma)$ is a normal distribution with mean $\mu$ and standard deviation $\sigma$, and $U([b, u])$ is a uniform distribution in the interval $[b, u]$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Uncertainty</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{air}$</td>
<td>Air temperature</td>
<td>±0.3 K</td>
<td>$N(0.00, 0.30)$</td>
</tr>
<tr>
<td>RH</td>
<td>Relative humidity</td>
<td>±3.0 %</td>
<td>$N(0.00, 3.00)$</td>
</tr>
<tr>
<td>SW</td>
<td>Shortwave incoming radiation</td>
<td>±10.0 %</td>
<td>$N(0.00, 0.10)$</td>
</tr>
<tr>
<td>LW</td>
<td>Long-wave incoming radiation</td>
<td>±10.0 %</td>
<td>$N(0.00, 0.10)$</td>
</tr>
<tr>
<td>$U$</td>
<td>Wind speed</td>
<td>±0.3 m s$^{-1}$</td>
<td>$N(0.00, 0.30)$</td>
</tr>
<tr>
<td>$P$</td>
<td>Precipitation</td>
<td>±25.0 %</td>
<td>$N(0.00, 0.25)$</td>
</tr>
<tr>
<td>$z_0$</td>
<td>Aerodynamic roughness length</td>
<td>0.001–0.10 m</td>
<td>$U(0.001, 0.10)$</td>
</tr>
<tr>
<td>PVOL</td>
<td>Pore volume fraction for maximum liquidwater holding capacity</td>
<td>0.03–0.05</td>
<td>$U(0.03, 0.05)$</td>
</tr>
</tbody>
</table>

The reference runs serve as basis for the subsequent assessment of the uncertainty of the simulation results (see Sect. 3.2), and the corresponding decomposition of the model variance (see Sect. 3.4). The modified Crocus model (see Sect. 2.1) is forced with the pre-processed and adjusted input data introduced in Sect. 2.2. The most relevant model parameters are given in Table 1. Following Björnsson et al. (1996) and Brandt et al. (2008), the model at KNG8 is initialized with an isothermal firm layer (273.16 K) with a mean density of 600 kg m$^{-3}$ and a total thickness of 20 m. At the bottom of the model domain a constant base temperature of 271 K has been applied. The maximum number of snow layers is set to 50 in order to obtain a detailed snowpack stratigraphy. The initial grid spacing is increased from 0.01 m at the surface to 10 m at the bottom. The number of grid cells and their spacing is updated during the simulation according to the accumulation, temperature, density, and melt. KNG8 is located in the accumulation zone of the glacier, where the near-surface layers consist of perennial snow rather than bare ice (Björnsson et al., 1996; Brandt et al., 2008). KNG1 is located in the ablation area of the glacier, where surface conditions are characterized by less snow accumulation during winter, stronger melt during summer, and a corresponding prevalence of bare ice at the surface. At both sites, the simulations start at the end of the ablation season, with the lowest recorded snow depth (defined by the minimum recorded surface height), and they are forced by hourly measurements. Simulation results are stored every 6 h for analysis. At the lower station, KNG1, the glacier ice reappears at the surface in the course of the ablation season. To represent the site-specific condition, the initial density is set to 830 kg m$^{-3}$, which is corroborated by observations from ice cores. Measurements of surface temperature and albedo are used for validation only and are considered as key indicators to judge the model’s ability to calculate the energy balance. Validation of mass-balance calculations is performed by comparing simulated and observed snow temperature and density profiles. Note, however, that the reference simulations were not optimized to fully reproduce the available observations. This would have required further tunings, which are not necessary for the purpose of this methodical study. There is no doubt, however, that the two reference runs truly reflect the basic
characteristics of the seasonal evolution of snow and ice at the two considered sites.

2.4 Global sensitivity analysis

In general, sensitivity analysis (SA) permits inferences on the different sources of uncertainty in model inputs (Sauter and Venema, 2011). This section gives an overview of how model-free sensitivity measures can be derived from variance-based methods. For the purpose of illustration, let us assume a generic model

\[ Y = f(X_1, X_2, \cdots, X_k), \]  

where the model output is \( Y \), the input quantity \( X_i \), and the corresponding first-order unconditional variance \( V(Y) \). Most common SA measures are based on local derivatives \( \partial Y / \partial X_i \) to estimate the relative importance of individual components. It is convenient to normalize the derivatives by the standard deviation so that the measures are weighted and sum up to 1. It is also interesting to note in this context that in linear models, the normalized derivatives coincide with the well-known standardized (linear) regression coefficients (Saltelli et al., 2006). Obviously, both measures rely on the assumption of linearity, which makes them unsuitable for complex models. This is particularly true when interaction effects become important, which is a characteristic property of non-linear and non-additive models. However, such effects may be addressed by so-called model-free measures, which can be effectively estimated by the GSA method described here. In contrary to the commonly used SA, GSA calculates the sensitivity measures in broader regions of parameter space by selecting appropriate distributions instead of a specific value of each parameter.

If one forcing input \( X_i \) is fixed at a particular value \( x_i^{*} \), the resulting conditional variance of \( Y \) is given by \( V_{X \sim i}(Y|X_i = x_i^{*}) \). This measure characterizes the relative importance of the factor \( X_i \), since the conditional variance will be less than the unconditional variance. The fact that this sensitivity measure depends on the value of \( x_i^{*} \) makes it rather impractical. Taking instead the average over the uncertainty distribution of \( x_i \), the undesired dependence will disappear (Saltelli et al., 1999, 2006).

The expression

\[ V(Y) = E_{X_i}(V_{X \sim i}(Y|X_i = x_i^{*})) + V_{X_i}(E_{X \sim i}(Y|X_i = x_i^{*})) \]  

(7)

decomposes the total variance \( V(Y) \) into the first-order (second right-hand-side term) and higher-order (first right-hand-side term) contributions. The corresponding first-order sensitivity index of \( X_i \) is given by

\[ S_i = \frac{V_{X_i}(E_{X \sim i}(Y|X_i = x_i^{*}))}{V(Y)}. \]  

(8)

This sensitivity index indicates the importance of individual factors without considering interactions effects. When the model belongs to the class of additive models, the first-order terms add up to 1, for example \( \sum_{i=1}^{k} S_i = 1 \). If this is not the case, the remaining variance must be explained by the higher-order effects induced by the interaction of input factor uncertainties. Interactions represent an important feature of non-linear non-additive models. The total sensitivity \( S_{T_i} \) of a factor \( X_i \) is made up of the first- and all higher-order terms where a given factor \( X_i \) is participating, consequently giving information on the non-additive character of the model. The \( S_{T_i} \) can be computed using

\[ S_{T_i} = \frac{E(V(Y|X_{\sim i}))}{V(Y)}. \]  

(9)

where \( X_{\sim i} \) indicates that all factors have been fixed and only \( X_i \) varies over its uncertainty range. This approach permits, even for non-additive models, the recovery of the complete variance of \( Y \). The sum of \( S_{T_i} \) is equal to 1 for perfectly additive models, otherwise it is always greater than 1. The difference between \( S_i \) and \( S_{T_i} \) is a useful measure of how much each factor is involved in interactions with any other factor (Saltelli et al., 2010).

First and total-order indices can be computed by Monte Carlo-based numerical procedures (Saltelli et al., 2010; Sobol et al., 2007). Estimating the conditional variances, such as \( V_X(E_{X \sim i}(Y|X_i = x_i^{*})) \), is computationally expensive, but Saltelli et al. (2010) provided an efficient algorithm for the simultaneous computation of \( S_i \) and \( S_{T_i} \). The calculation requires two independent sampling matrices \( A \) and \( B \), with the elements \( a_{ji} \) and \( b_{ji} \). The subscript \( i \) runs from one to the number of factors \( k \), while \( j \) runs from one to the number of samples \( N \). A third matrix, \( A_B^{(i)} \), is introduced, where all columns are taken from \( A \), except for the \( i \)th column, which is from \( B \). The first-order effect can then be computed as

\[ V_{X_i}(E_{X \sim i}(Y|X_i = x_i^{*})) = \frac{1}{N} \sum_{j=1}^{N} f(B)_{j}(f(A_B^{(i)})_{j} - f(A)_{j}). \]  

(10)

where \( f(\cdot)_{j} \) denotes the model output of the \( j \)th row. Similarly, total effect can be estimated by

\[ E(V(Y|X_{\sim i})) = \frac{1}{2N} \sum_{j=1}^{N} (f(A)_{j} - f(A_B^{(i)})_{j})^2. \]  

(11)

The indices are estimated at a total cost of \( N \cdot (k+2) \) model runs with \( N \), a sufficiently large number of base samples. In this study, we performed 20 000 model runs with \( k = 8 \) factors and \( N = 2000 \) base samples, which proved to be a reasonable compromise between computational feasibility and robustness of the results. The elements of the matrices \( A \) and \( B \) are generated from quasi-random Sobol sequences (see Sect. 2.6). The Sobol sequence generates quasi-random numbers in a range between [0,1]. The random numbers are then
mapped to match the uncertainty distributions given in Table 2 (see also Sect. 2.5). The roughness length for heat $z_{0h}$ is derived from the roughness length for momentum using the relation $z_{0h} = z_0/10$ as its a default setting for Crocus. The snowpack model is forced with each of the 20 000 parameter combinations.

Sensitivity indices are computed from the 6-hourly model output of these Monte Carlo runs and are analysed with regard to snow depth, surface-energy balance, turbulent heat flux, and latent heat flux. The calculations are based on the reference runs performed at the two glacier sites. Therefore, this strategy allows for the study of the detailed temporal evolution and distinction of patterns during summer and winter and in different mass-balance regimes of the glacier (accumulation and ablation area), respectively. The accuracy of the sensitivity indices was assessed from 1000 empirical bootstrapping samples being randomly drawn with replacement from the original data set. The indices $S_i$ are calculated for each of the bootstrap data sets, and the 95% confidence regions are estimated.

### 2.5 Measurement error characteristics

The model uncertainty is estimated from a set of quasi-Monte Carlo sequences (see Sect. 2.6), based on the calibrated reference runs and specified uncertainty measures of key input factors and model parameters (Table 2). The probability density distributions of the measurement errors are either derived from simultaneous measurements with two sensors (as for air temperature) or by the accuracy of the sensor given by the manufacturer specifications (this set-up is similar to the NB_lab scenario in Raleigh et al., 2015). When dealing with measurement errors, there is usually insufficient information on how the given uncertainties were determined and how the underlying distribution functions look. Regarding field applications, additional factors come into play that are usually not considered in calibration procedures. For example, temperature measurements may be affected by aging or insufficient shielding from solar radiation, both also being crucial in glacier environments. To characterize the uncertainty of the measured meteorological parameters used to force the model, we follow a common approach and assign normally distributed errors considering the standard deviation derived from the manufacturer specifications. Roughness length and pore volume fraction are assumed to vary uniformly in a predefined range, which appears justified by observational evidence indicating high local and temporal variability of snow surface conditions due to the formation of sastrugi or meltwater conduits (Sauter et al., 2013). It therefore also seems reasonable to use a uniform range of pore volume fractions rather than assuming a truncated normal distribution.

### 2.6 Sobol’s sampling

Randomly drawn samples from a hypercube space tend to have clusters and gaps. Such sequences are said to have a high discrepancy. Low-discrepancy sequences, also known as quasi-random sequences, are designed to have well-distributed numbers in a multidimensional space, even for small quantities. Quasi-random algorithms bias the selection of points to maintain an even spread across the hypercube. These sequences are commonly used in sensitivity analysis and provide better estimates of the model-free sensitivity measures (see Sect. 2.4). Sobol sequences, which are used in this study, belong to this class of sequences (Sobol, 1998; Sobol et al., 2007).

### 3 Results

#### 3.1 Reference run

Here we mainly examine the accuracy of the reference run at KNG8, which is representative of the accumulation area of the glacier and prevailing snow conditions. Validation of the reference run for KNG1 (representative of the ablation area of the glacier) reveals similar skills, and so we more or less forego a detailed description of those results. Comparison of the simulation at KNG8 with the snow pit profile from 6 April 2011 shows a difference in snow depth at the end of the winter period of less than 0.1 m (see Fig. 2).

In terms of water equivalent, the accumulated mass during the winter amounts to +0.76 m, compared to +0.82 m having been observed. Figure 2 also shows the comparison of simulated snow surface temperature with observational data. The simulated snow surface temperature is derived from upwelling long-wave radiation assuming a snow emissivity of 0.99. Surface temperature is a key variable for flux parametrizations and also links the calculated mass and energy balance. Its temporal variability is well captured ($R^2 = 0.93$), and a root mean squared error of 2.3 K conforms to the general skill of most sophisticated snowpack models (Obleitner and De Wolde, 1999; Rutter et al., 2009; Etchevers et al., 2004). The spread increases in the winter, which, for example, could be associated with undetected rimming of the sensor or structural model uncertainties. The vertical temperature gradient in the snowpack is an important driver of snow metamorphism and is depicted in Fig. 2. In the upper 0.6 m, the observed temperature is slightly higher than modelled and the RMSE $= 1.2 \degree C$ is in part also attributed to measurement shortcomings (Obleitner and De Wolde, 1999). The corresponding density profile confirms that the model is able to simulate the gross snowpack layering (see Fig. 2). The relatively large difference within the upper 0.1 m is due to the fact that the constant $d_p$ in Eq. (4) is set to 300 kg m$^{-3}$. The RMSE between the measured and modelled albedo over the entire simulation period is 0.06. Note that the measured
albedo ranges between 0.65 in the ablation period and 0.92 in the accumulation period, which is characteristic for a site in the accumulation region (Armstrong and Brun, 2008; Greuell et al., 2007). Albedo is significantly depleted at the lower site during summer, as is typical for a site in the ablation area due to exposure of darker glacier ice, which has also been confirmed by Greuell et al. (2007).

Table 3 gives a summary of the observed meteorological variables and the calculated energy-balance components at KNG8 and KNG1. We thereby distinguish values for consistent summer and winter periods covering the months JJA and DJF, respectively. These must not be interpreted as ablation or accumulation periods, which are, in fact, of different durations at both sites. Air temperatures decrease with elevation and remain negative all over the glacier during the considered winter period, while they are positive during the central summer months. This is basically reflected in the observed surface temperatures, which indicate that the glacier (snow) surface melts during JJA while remaining frozen during the DJF period. Bulk vertical temperature gradients between the 2 m level (nominal) and the surface indicate inversion conditions prevailing throughout the year. Humidity increases with elevation along the glacier as expected. The otherwise observed decrease of vapour pressure with elevation during the summer may be related to low-lying clouds (as suggested by long-wave incoming radiation data). The local vertical gradients in vapour pressure are calculated by assuming saturation at the surface, and they reveal higher values in the air. This leads to positive latent heat fluxes providing mass and energy to the surface. Wind speeds are generally higher at KNG1, which is more pronounced during the winter months, when the air is more stably stratified. Katabatic winds play a role in this context, as is obvious from analysis of wind directions (shown in Karner et al., 2013). With regard to the radiation components, there is virtually no input from solar radiation during the winter months. During summer, global radiation, that is the sum of direct and diffuse solar radiation, increases by about 5 W m$^{-2}$ per 100 m elevation. About 80 % of incoming solar radiation is reflected at the higher site (KNG8) during the summer and reflects the persistence of snow. An albedo of about 48 % is calculated for the lower site (KNG1), where snow disappears early in the spring and exposes darker glacier ice at the surface. Long-wave incoming radiation is an important source of energy during both seasons. Its increase with elevation during the winter and the decrease during the summer, which reflects corresponding changes in cloud characteristics (low level fog in winter). These characteristics of the radiation components induce a decrease of net radiation with elevation, with overall negative values during the winter and positive ones during the summer. Sensible heat fluxes are generally directed towards the surface, which is more pronounced during the winter and
at the lower site KNG1. This is also true with regard to latent heat fluxes, which by magnitude equally contribute to the calculated surface-energy budget. The latter is characterized by negative values during the winter, when small gradients along the glacier occur. During the considered summer months, the energy budget is strongly positive and fosters melt at both sites. Naturally, this is more effective at the lower site, which mainly can be traced back to stronger input from solar radiation (lower albedo) and turbulent fluxes. There were corresponding developments of the mass balance at both sites. Note that further energy and mass-balance components were calculated by the simulations, which on average were small in magnitude. Therefore, they are not considered in the overall context of this study, which does not aim at a detailed investigation of the individual fluxes and associated processes themselves.

3.2 Integrated model uncertainty

Figure 3 shows the time series of snow depth for the reference run as well as of the quantiles estimated from the ensemble simulations for KNG8 and KNG1. At KNG8, the 95 % quantile range can be clearly divided into two regimes: (i) the build up of the snowpack when the 95 % interquartile range increases towards ±1.2 m until the end of June, and (ii) the melt period when the interquartile range experiences an additional increase. At the end of the 1-year simulation period, the uncertainty (95 % quantile range) in snow depth caused by the systematic measurement errors reaches more than 3 m. Note that the interquartile range shows a clear asymmetry, which is more pronounced after June 2011. This marks the onset of effective melt, which induces a higher liquid water content of the near-surface snow layers. The associated wet snow metamorphism drives a decrease in albedo. The development is enhanced upon exposure of snow from the previous year, which is characterized by even lower albedo and higher density. Sporadic snowfall events in August 2011 led to an increase of the upper 99 % quantile bound. The simulations are also quite sensitive to disturbances during the first 2 months, when the amounts of snowfall are small. The overall evolution and the final characteristics of the ensemble variability at KNG1 are similar to that at KNG8. Note, however, that the accumulation season is significantly shorter (beginning in November compared to August at KNG8) and is characterized by a smaller ensemble spread compared to KNG8. The latter reflects that precipitation at this elevation mostly comes in the form of rain. Throughout the accumulation season, the ensemble spread is low and is related to small snowfall amounts and widens significantly in the ablation season, when the glacier reappears at the surface. The point at which the glacier ice reappears depends on the maximum snow depth and can occur between May and the beginning of July. The final uncertainty (95 % quantile range) in snow depth due to measurement errors is almost 5.5 m at the end of the ablation season.

3.3 Mean total-order sensitivity indices

Figure 4 shows the mean $S_{Ti}$ of individual factors on the variance of calculated surface-height changes (SHC), SEB, the SHF, LHF flux at KNG8 and KNG1. Recall that total-order indices, $S_{Ti}$, measure the contribution of each factor to the ensemble variance, including all interaction effects.

At KNG8, SHC is mainly affected by uncertainties in precipitation $P$ (0.58) and incoming long-wave radiation LW (0.29). The remaining factors are very likely to have little impact. SEB, SHF, and LHF are most sensitive to LW, with $S_{Ti}$ values ranging from 0.53 to 0.77. Of note is the sensitivity of SEB to $P$ (0.25) and $z_0$ (0.4). Hence, $z_0$ is the second-most important parameter for SEB and SHF, and it even explains 

<table>
<thead>
<tr>
<th>Variable</th>
<th>KNG8</th>
<th>KNG1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air temperature (2 m) °C</td>
<td>253.5 (8.0)</td>
<td>271.5 (2.6)</td>
</tr>
<tr>
<td>Surface temperature °C</td>
<td>252.3 (9.1)</td>
<td>270.1 (3.9)</td>
</tr>
<tr>
<td>Rel. humidity (2 m) [%]</td>
<td>97 (6)</td>
<td>91 (7)</td>
</tr>
<tr>
<td>Water vapour pressure (2 m) [hPa]</td>
<td>3.3 (0.8)</td>
<td>5.3 (0.5)</td>
</tr>
<tr>
<td>Wind speed (2 m) [m s⁻¹]</td>
<td>1.2 (1.9)</td>
<td>0.8 (1.9)</td>
</tr>
<tr>
<td>SW incoming radiation [W m⁻²]</td>
<td>0.1 (0.9)</td>
<td>239.5 (167.9)</td>
</tr>
<tr>
<td>SW outgoing radiation [W m⁻²]</td>
<td>0.1 (0.8)</td>
<td>193.6 (137.4)</td>
</tr>
<tr>
<td>LW incoming radiation [W m⁻²]</td>
<td>223.4 (43.9)</td>
<td>268.5 (39.6)</td>
</tr>
<tr>
<td>LW outgoing radiation [W m⁻²]</td>
<td>231.5 (32.4)</td>
<td>301.8 (169.9)</td>
</tr>
<tr>
<td>Sensible heat flux [W m⁻²]</td>
<td>3.1 (36.0)</td>
<td>3.5 (17.1)</td>
</tr>
<tr>
<td>Latent heat flux [W m⁻²]</td>
<td>2.9 (4.5)</td>
<td>-1.1 (6.1)</td>
</tr>
<tr>
<td>Surface-energy balance [W m⁻²]</td>
<td>-1.9 (14.5)</td>
<td>14.8 (34.3)</td>
</tr>
</tbody>
</table>
most of the LHF variance (0.27). A smaller share of SHF and LHF uncertainty is explained by \( U \) and RH. In particular, RH is important for LHF. In order to make an important contribution to the ensemble spread, the total-order indices should exceed the 0.05 limit (Saltelli et al., 2006). Following this criteria, some factors (\( T \), \( SW \), and pore volume, \( PVOL \)) can be designated as insensitive and with little influence on SHC, SEB, SHF, and LHF. The averaged first-order indices vary between 0.66 and 0.82, depending on the considered model output (see Fig. 4). The sensitivity pattern at KNG1 differs from that at KNG8. SHC is sensitive to LW, \( P \), and \( z_0 \). In contrast to KNG8, \( P \) has less influence on SHC variability than LW at KNG1. In total, the model is less affected by the uncertainty of \( z_0 \). However, RH (\( S_{Ti} = 0.1 \)) explains slightly more of the LHF variability at KNG1 than at KNG8.

### 3.4 Temporal evolution of the total-order sensitivity indices

Figures 5 and 6 show the temporal evolution of the \( S_{Ti} \) values with respect to SHC, SEB, SHF, and LHF. The \( S_{Ti} \) values are...
Figure 5. Evolution of the 6-hourly total-effect indices affecting modelled surface-height change (SHC), surface-energy balance (SEB), sensible heat flux (SHF) and latent heat flux (LHF) at KNG8. Refer to Table 2 for the explanation of the indicated uncertainty factors.

calculated for each time step using the 20,000 Monte Carlo runs.

The variability of SHC at KNG8 is mainly caused by the uncertainty of $P$ and LW. From November to May, almost all uncertainty is attributed to $P$, with $S_T$ ranging between 0.7 and 0.9. During the ablation season, LW becomes a dominant factor. Other factors, such as $U$, $z_0$, and SW, make less of a contribution (<0.2) to overall SHC variability, even though they can have an intermittently strong impact (≥0.3) on the variance of SEB, SHF, and LHF. Errors in LW and $z_0$ have a strong impact on SEB all year, while $P$ is only relevant in the winter season. During the spring, SW has an increased influence on SEB and drops to zero during the arctic winter. RH and $U$ contribute most to SEB variability in the period from August to March. Along with LW, $U$ and $z_0$ have a significant effect on both SHF and LHF variance. The uncertainty in $T$ and PVOL do not have an influence on either SHC or SEB.

At KNG1, the contribution of $P$ and LW is lower and not as consistent as at KNG8. In August and September, $z_0$ temporarily contributes (up to 0.8) to the SHC variability. In contrast to the sensitivity pattern at KNG8, other factors ($z_0$, $U$, RH, SW) contribute substantially to SHC. SEB is by far most sensitive to errors in LW. SW gains importance for a short period in May, with $S_T$ up to 0.4, although most of the time the contribution is very low, which is also true for RH, $T$, and PVOL. In general, the sensitivity pattern of SHF and LHF is similar to the pattern observed at KNG8. Here again, $z_0$ and $U$ temporarily explain a large share of the variability in turbulent fluxes (≥0.75) in the summer. Errors in RH mainly impact LHF variability.

4 Discussion

We investigated the seasonal pattern of the sensitivity of snow model output to uncertainties in input data and some key model parameters. Eight metrics characterizing forcing uncertainties (LW, $P$, PVOL, RH, SW, $T$, $U$, and $z_0$) and four metrics characterizing the model response (SHC, SEB, SHF, and LHF) have been considered. The introduced uncertainties represent the typical measurement errors of data used
Figure 6. Evolution of the 6-hourly total effect indices affecting modelled surface-height change (SHC), surface-energy balance (SEB), sensible heat flux (SHF) and latent heat flux (LHF) at KNG1. Refer to Table 2 for the explanation of the indicated uncertainty factors.

to force the model. The presented results are based on Monte Carlo simulations and subsequent application of Sobol’s sensitivity analysis to decompose first- and higher-order effects on the resultant variance. Simulations and analysis were applied to two sites at an arctic glacier to address characteristics in different mass-balance regimes.

The results from the reference simulations at the two sites allow for an in-depth discussion of the typical meteorological conditions at the two study sites and the related surface exchange processes, which are reflected in the constellation of the energy and mass-balance components. Such a study has not been performed at the Kongsvegen glacier thus far, but related aspects will nevertheless not be discussed in more detail here due to the rather methodical outline of this study. Note, however, that Obleitner and Lehning (2004) and Karner et al. (2013) have already studied this issue at another site close to the average equilibrium line of the glacier (ca. 537 m a.s.l.). This location is only about 137 m below KNG8, but the results concerning energy and mass balance are not directly comparable because the sites are located in different glaciological regimes (equilibrium line altitude vs. accumulation area). Some common features may be inferred from Table 3 though, which in part is addressed in Sect. 3.1. Consideration of sites other than KNG6 was mainly motivated by the availability of correspondingly suitable data. It is to be noted in this context that for the purpose of this study, the reference runs were not fully calibrated towards the observations, which would have been necessary for, for example, quantitative mass-balance studies. The overall results of this work show that on average about 80 % of the total variance of SHC and SEB can be explained by first-order effects (Fig. 4). This means that the remaining 20 % of the variance is due to non-linear interaction effects. There is no significant difference between the two sites at the glacier. This is in partial contrast to the findings of Raleigh et al. (2015), who performed similar investigations for different snow regimes and found that first- and total-order indices are of comparable magnitude. However, the results cannot be directly compared because they analysed different model output variables and used simpler parametrizations (i.e. bulk model), which possibly enhances the interaction effects. The performed sensitivity
Long-wave incoming radiation depends on column integrated air temperature humidity and cloudiness and is the dominant source of energy for the glacier, independent of site and season. This is typical for glacier environments (Greuell and Smeets, 2001) and is enhanced in the Arctic, where input from shortwave insolation is missing during the polar night conditions (e.g. Obleitner and Lehning, 2004; Van den Broeke et al., 2011; Karner et al., 2013). Variability in LW therefore directly impacts NR and hence SEB. This also holds true for corresponding measurement uncertainties, which are comparatively large. The sensitivity analysis shows, that about 50% of SEB variance can be explained by total-order effects due to LW (Fig. 4). The effect is slightly reduced at the higher site (KNG8), which may be related to the general decrease of long-wave incoming radiation with elevation (Table 3). Neither study site shows a pronounced seasonal variability in the corresponding SEB sensitivity pattern, which may be related to the rather continuous nature of long-wave incoming radiation and its dominance for NR (Figs. 5 and 6). LW uncertainty also strongly impacts on the variance of the calculated turbulent fluxes. Yearly averaged total-order indices are somewhat lower than for SEB, ranging at about 0.3 (KNG8) and ca. 0.5 (KNG1), respectively. The sensitivity analysis further reveals a stronger impact on SHF and an outstanding seasonal dependency of the sensitivity of the turbulent fluxes (Figs. 5 and 6). Feedback related to surface temperature provides a key for understanding these features, which couples the (long-wave) radiation budget and the turbulent fluxes. The stronger input by long-wave radiation, the more positive NR is, which in part is absorbed at the surface and increases surface temperature. Hence, surface temperature fluctuations are larger than those of air temperature and respond very sensitively to changes (uncertainties) in LW. This in turn effectively changes the stability of the near-surface air, which drives turbulent exchange therein. This feedback is most effective during dry snow, that is winter conditions, with large total-order sensitivity indices from autumn until spring (Fig. 5). In the ablation season, when the surface temperature is more or less at the melting point, SHF and LHF are no longer sensitive to uncertainties in LW. LHF is also affected (though to a lesser extent) because of the associated changes in vapour pressure at the surface. LW also strongly impacts SHC variability, which is more pronounced at KNG1 (Fig. 4) and during the summer (Figs. 5 and 6) when LW uncertainty explains more than 80% of SHC variability. This can be related to the fact that in our approach the input uncertainty range (±10%) proportionally increases with the magnitude of LW. The latter is essentially true during summer when air temperature and humidity are high. LW is further enhanced due to cloudiness and during precipitation events. Note that in the Kongsvegen area the percentage of low clouds rises over 60% from April to October (Kupfer et al., 2006). Stronger long-wave radiation input leads to higher surface temperatures, which induce steeper temperature gradients within the near-surface snow layers and enhance their metamorphism (settling or even melt). To put these findings in a broader context, Karner et al. (2013) applied another snow model to data from KNG6 (Fig. 1) and also identified LW uncertainty as the most influential factor on calculated mass balance and SEB. However, their study is based on consideration of single-order effects only. Another hint regarding the outstanding influence of uncertainties in LW is provided by Raleigh et al. (2015), who systematically explored the propagation of forcing uncertainties to snow model output based on Sobol’s sensitivity analysis. Their results confirm the importance of LW uncertainty, but a straightforward comparison to our results is hampered due to the different metrics used for input uncertainties and model output.

Precipitation is another influential factor on the variance of snow model output. This mainly concerns the simulated surface height changes (which is considered as a metric of calculated mass balance) and surface energy balance. Total-order sensitivity indices are particularly high during the winter and at KNG8 (Fig. 4). The lower values in summer are bound up with the fact that no liquid precipitation is measured at this site, and hence has no impact on the variability. In these higher regions of the Kongsvegen glacier, recurrent snowfall events may occur year round, which results in a deeper snowpack (2.2 m) and a longer accumulation period (October through April). This is evident from Fig. 3 and the corresponding SHC sensitivity patterns. Snowfall occurs comparatively infrequently and is overall inefficient at the glacier tongue and during the summer months. This is mainly an effect of temperature lapse-rate determining the rain-snow transition and the tendency of cloud formation at the crest of mountains. Similar to Raleigh et al. (2015), we find that P uncertainty is a critical factor for the snow disappearance in the ablation zone (see Fig. 3). Depending on the winter conditions, the reappearance of glacier ice typically occurs between May and July. However, we find little evidence that ablation rates are significantly controlled by P. Note that in our study, precipitation was derived from ultrasonic sensors and corresponding uncertainty was specified from the manufacturer specifications. Frequently, however, snow precipitation is derived from standard gauges. As previously mentioned, even small
 Errors due to wind-induced under-catch or conversion of snow depth changes to precipitation rates in terms of SWE (see also Sect. 2.2) might thus have a significant impact on the simulations. According to Eq. (4), the conversion is sensitive to air temperature ($\partial \rho/\partial T_{\text{air}} = b_\rho$) and wind velocity ($\partial \rho/\partial U = c_\rho/(2 \cdot \sqrt{U})$). This demonstrates that the conversions are particularly sensitive to measurement errors at low wind speed. However, precipitation measurements at higher wind velocities usually show a systematic under-catch. Schmucki et al. (2014) showed that for standard precipitation measurements, a correction of under-catch may reduce the mean absolute percentage error by 14% for snow depth at high alpine stations. Forlånd and Hanssen-Bauer (2000) demonstrated the importance of this issue for Svalbard environments. During winter and spring the calculated SEB is strongly affected by uncertainties in precipitation input, which explains about 25% of the total variance. There is no indication of important interaction with shortwave radiation (missing during winter) or turbulent fluxes. Hence, precipitation-induced perturbation of LW is considered as the most important factor linking the variability of $P$ and SEB. The effect is more pronounced at the upper site. At the lower part of the glacier, fresh snow events are comparatively infrequent and inefficient. During the summer in particular, fresh snow usually melts within a short period without leaving a significant impact on SEB.

Surface roughness length has a strong impact on the turbulent fluxes and hence on SEB variances. Overall, this is due to the associated processes and their parametrizations (Vionnet et al., 2012). The sensitivity is particularly pronounced regarding SHF and at the upper study site (KNG8), where total-order sensitivity indices reach 0.3 on average throughout the year (Fig. 4) and are highest during the period from April until June. Interestingly enough, KNG1 experiences the most pronounced impact of $z_0$ uncertainty on SEB during the period from July to September, which constitutes the main ablation period at this site (Fig. 3). This feature is attributed to the concurrent appearance of bare ice and the accordingly parametrized increase of surface roughness, which represent the formation of, for example, meltwater channels. Uncertainty in $z_0$ also impacts the simulations of SHC, which is most pronounced at the lower site and during the summer. This again is related to the overall increased roughness (factor 10) and accordingly enhanced turbulent fluxes contributing the surface melt. These findings basically conform to the first-order sensitivity studies by Karner et al. (2013), changing $z_0$ by an order of magnitude. However, as a straightforward comparison is difficult due to the choice of error range, which can have a strong influence on the results (Raleigh et al., 2015).

Errors in wind speed significantly impact the calculated turbulent fluxes during the period from April to June (KNG8) and July to September (KNG1). It is notable that the latter periods correspond to those when $z_0$ uncertainties exert the most influence, indicating combined effects. We therefore attribute their impact on SEB and SHC mainly to their direct involvement in the calculation of the turbulent fluxes. The largest sensitivity of $U$ is associated with lower wind speeds (Table 3). This is in line with the findings of Dadić et al. (2013), who found higher sensitivity of the turbulent fluxes with respect to wind speed in the range of 3–5 m s$^{-1}$. The effect of local wind velocity variations on turbulent fluxes and ablation rates has also been addressed by other studies (Mott et al., 2013; Marks et al., 1998).

Air temperature may be expected to strongly influence the calculation of the turbulent fluxes and, therefore, also on the SEB. However, this is not seen in the results of the sensitivity analysis, which at both stations does not show significant impacts on any of the considered model output metrics (SHF, LHF, SEB, and SHC). Further, this result must be considered in light of the variances rather than absolute values. The driving temperature gradients between the surface and the air are of the order of 2–5 K (Table 3), which reduces the sensitivity of the calculated fluxes due to the comparatively small measurement errors that have been assumed ($\pm 0.3$ K). Further, Raleigh et al. (2015) found that $T$-forcing biases had a stronger impact on ablation rates (which may be considered as measure of summer SEB) compared to random errors, while peak snow water equivalent (comparable to winter SHC) was hardly affected. Similarly, Karner et al. (2013) found a strong impact of $T$ biases on the calculated SEB. The seasonal $T$-sensitivity patterns on SEB and its components are characterized by relatively strong impacts in the spring and autumn. During this period, temperature is crucial whether precipitation is considered as snow or rain. Feedback related to albedo or LW may play an additional role there. The sensitivity study was performed based on standard laboratory specifications given by the manufacturers. However, the actual uncertainties of air temperature measurements can be much larger depending on the efficiency of the used radiation shields or ventilation devices. Relevant to this study, Karner et al. (2013) did not find significant biases between ventilated or unventilated air temperature measurements. However, this result may not be generalized.

The impact of humidity forcing errors on the simulation metrics was analysed concerning the directly measured variable (relative humidity). By definition, however, the latter combines humidity and temperature information and is therefore not an ideal metric, which may be considered in forthcoming studies. Irrespective of that, our results reveal that on average RH uncertainty has an overall small but somewhat stronger impact on calculated SEB compared to $U$ (Fig. 4). The impact is less pronounced at the lower site and during the summer (Figs. 5 and 6). The overall variability
of the seasonal SEB sensitivity pattern is small, however, and is difficult to interpret due to the inherent temperature effects. There are indications of stronger impacts in the spring when conditions are favourable for sublimation due to high saturation deficits occurring simultaneously with strong winds and moderate temperatures (Sauter et al., 2013; Obleitner and Lehning, 2004; Karner et al., 2013). Calculated RH sensitivity is generally stronger regarding LHF compared to SHF, which is reasonable.

Shortwave incoming radiation is a strongly influential factor on the SEB of snow and ice, and corresponding uncertainties are expected to have a strong impact on respective simulations. Contrasting this general anticipation, our sensitivity analysis reveals that on an annual basis, only a small amount of the total SEB variance can be explained by the assumed uncertainty of SW input data (Fig. 4). This concerns both sites and basically reflects that in the Arctic, the anticipated effect is generally reduced due to the lack of solar insolation during winter. Recall that in the Kongsvegen environment, the polar night conditions last from late October to early February. This is also reflected in the seasonal sensitivity pattern, which do not show any signal during the winter. There is, however, a significant influence on SEB in the spring and autumn (Figs. 5 and 6). This might be related to the previously mentioned influence of intermittent fresh snow on older surfaces with lower albedo, whose effectiveness also depends on SW and its variability (and temperature as addressed above). Another reasoning is based on the consideration of energy supplied by uncertainties in SW compared to those in LW. Hence, the sensitivity of net shortwave radiation (\(\partial E \) SW) to measurement errors (\(\partial G \) ) to measurement errors (\(\partial E \) SW) is given by \(\partial G/\partial E \) SW = \(1 - \alpha \), with \(\alpha \) denoting albedo. Therefore, the ratio (\(R \)) of the sensitivities of incoming long-wave radiation and available net shortwave radiation at the ground is \(R = 1/(1 - \alpha) \). By multiplying \(R \) with the error ratio, we obtain a properly scaled ratio \(\hat{\rho} = (E_{LW}/E_{SW}) \cdot (1/(1 - \alpha)) \). Assuming a 10% error of typical daytime values in the summer (\(E_{SW} = 40 \text{ W m}^{-2} \) and \(E_{LW} = 26 \text{ W m}^{-2} \)) and a \(\alpha = 0.75 \), we obtain \(\hat{\rho} = 2.6 \). This means that the changes in energy due to the measurement uncertainty of LW are about 2.6 times greater than that for SW. In spring and autumn, the ratio becomes larger due to an increasing albedo and decreasing incoming shortwave radiation. This also leads to the conclusion that increasing the accuracy of SW measurements by a few percent would not increase our confidence in simulations of snow depth or SEB components. The sensitivity of SHC on uncertain specification of shortwave radiation SW is negligible overall, except in summer the latter being more pronounced at the lower site. This again reflects a coupling to albedo, which is lower at KNG1. The results conform to Karner et al. (2013) showing that the overall influence of SW is strikingly smaller compared to that of LW. As was pointed out by Raleigh et al. (2015), overall, this is attributed to the high albedo of snow (reducing absorbed energy and the associated impact of uncertainties) and the non-linear (amplifying) interactions of LW, which through surface temperature is coupled to the calculation of the turbulent fluxes.

The liquid water holding capacity of snow, PVOL, strongly depends on snow microstructure and related surface/subsurface developments throughout the winter season, and it is difficult to measure (Armstrong and Brun, 2008). However, investigation of the propagation of corresponding uncertainties in the snow model results was hardly addressed and therefore was considered in this study. According to our results, the uncertainty in specifying liquid water holding capacity of snow makes the least contribution to the total model variance of virtually all considered output metrics, mainly by interactions. The seasonal PVOL sensitivity pattern reveals some enhanced impact on SEB variability in the spring and autumn, which is more pronounced at the upper site (KNG8). Tentatively, this feature could be attributed to the percolation of rain or meltwater and subsequent refreezing. However, it remains to be investigated whether the associated release of energy can explain the observed variance pattern. Gascon et al. (2014) remarked that the Crocus percolation scheme tends to favour near-surface freezing and insufficient refreezing at depth, which could be another factor in this context. Overall, the assumption of default values (as in this study) does not have a significant impact on the calculated mass-balance (SHC).

5 Conclusions

We investigated the seasonal pattern of the sensitivity of snow model output to uncertainties in input data and some key model parameters. A set of eight metrics characterizing forcing uncertainties and four metrics characterizing the model response have been considered. The introduced uncertainties characterize typical measurement errors of data used to force a state-of-the-art snow model, and the presented results are based on Monte Carlo simulations and subsequent application of Sobol’s GSA. Simulations and analysis are applied to two sites at an arctic glacier to address characteristics in different mass-balance regimes. The results clearly demonstrate that even conservatively estimated input uncertainties can lead to a significant loss of confidence in key simulation results concerning the surface energy and mass budget. The overall impact of individual error sources on the sensitivity pattern varies between the two stations considered. In the accumulation zone (higher elevation station), precipitation and long-wave radiation are key factors for the evolution of the snowpack and contribute most to the model uncertainty. The precipitation variability is of less significance at the lower elevation station, while other factors, such as wind velocity or surface roughness, gain importance. Uncertainties in the measurement of incoming shortwave radiation
and air temperature have little influence on the model outcome, the former being biased by the specific arctic conditions. The calculated seasonal sensitivity patterns are similar overall at both study sites. The most temporally continuous influence on model output is exerted by variance of long-wave radiation and surface roughness. Precipitation tends to have the strongest impact during the winter, while wind velocity, air temperature, humidity, and liquid water holding capacity mainly impact the simulations in the summer or transitional seasons. The results thus allow for the identification of the most critical parameters and environmental conditions, which together with the consideration of relevant model parametrizations, provide directions for future improvements. The analysis is based on rather conservative though commonly used uncertainty estimations. These are mostly based on manufacturer specifications and hence on laboratory settings. In field applications, however, effective uncertainty is likely enhanced but is difficult to quantify. Moreover, we did not systematically consider effects of different uncertainty types (bias vs. random), different probability distributions or their combined propagation effects. Correspondingly, set-up ensemble simulations fed by sampling from quasi-random sequences are therefore recommended for future investigations. Overall, the performed decomposition of the snow model output sensitivity by GSA proved valuable for enhancing our understanding of key snow model output sensitivity patterns in response to uncertainties in forcing data. The key findings either confirm or complement those derived from a few other studies employing GSA. The revealed importance of long-wave radiation input may be considered as a trend-setting example. No doubt, however, more common efforts are necessary to further test and improve the method. This concerns enhanced consideration of the effects of different combinations of error types and probability distributions, including the propagation of parametrization uncertainties, which are mostly even less constrained than measurement errors. Detailed consideration of the parametrization of albedo in Crocus is suggested for the future, which was not addressed in this study. The presented approach is universal and can be applied to earth system models in general and may be applied to snow and glacier mass- and energy-balance modelling in all climate regions. From a practical and methodical point of view, the main limitations of this study are the high computational effort and proper specification of the probability density functions of the parameter uncertainties. Finally, we would like to note that measurement uncertainties are independently sampled and do not possess any correlation structures. Consequently, the approach cannot be used to investigate the response of snow or ice depending on systematic changes in the environmental (climate) conditions. This requires appropriate sampling strategy to obtain the same correlation structure as those observed in nature.

Acknowledgements. This work was supported by the Austrian Science Fund (FWF, grant I 369-B17). Field work at Kongsvegen was performed in cooperation with the Norwegian Polar Institute (Tromsø, J. Kohler) and University of Oslo (Ch. Nuth). The measurements of snow data were supported by Österreichische Polarforschungsgesellschaft, Julius-Payer Stipendium 2010. F. Karner, F. Bilgeri, and W. Steinkogler are thanked for performing the field work and pre-evaluation of the data used within this study. We thank M. S. Raleigh and both anonymous reviewers for their detailed comments and constructive criticism of the original manuscript. We would also like to thank Samuel Morin and Matthieu Lafaysse for their support in applying the Crocus snowpack model.

Edited by: P. Huybrechts

References


Braithwaite, R. J. and Zhang, Y.: Sensitivity of mass balance of five Swiss glaciers to temperature changes assessed by tuning a degree-day model, J. Glaciol., 46, 7–14, 2000.


T. Sauter and F. Obleitner: Assessment of the uncertainty of snowpack simulations


Norwegian Polar Institute. Kartdata Svalbard 1:100 000 (S100 Kart-data), Tromsø, Norway, Norwegian Polar Institute, https://data.npolar.no/dataset/645336c7-adfe-4dd5-97d6-9426fe7868e3 (last access: August 2015).


Pahaut, E.: La métamorphose des cristaux de neige (Snow crystal metamorphosis), 96, Monographies de la Météorologie Nationale, Météo France, 1976.


