A 7-year dataset for driving and evaluating snow models at an Arctic site (Sodankylä, Finland)

Richard Essery¹, Anna Kontu², Juha Lemmetyinen³, Marie Dumont⁴, and Cécile B. Ménard⁵

¹School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FE, UK
²Arctic Research Unit, Finnish Meteorological Institute, 99600 Sodankylä, Finland
³Finnish Meteorological Institute, 00101 Helsinki, Finland
⁴Météo-France-CNRS, CNRM-GAME UMR3589, CEN, Grenoble 38000, France
⁵CORES Science and Engineering Ltd, Edinburgh, UK

Correspondence to: Richard Essery (richard.essery@ed.ac.uk)

Received: 7 December 2015 – Published in Geosci. Instrum. Method. Data Syst. Discuss.: 19 January 2016
Revised: 12 April 2016 – Accepted: 31 May 2016 – Published: 16 June 2016

Abstract. Datasets derived from measurements at Sodankylä, Finland, for driving and evaluating snow models are presented. This is the first time that such complete datasets have been made available for a site in the Arctic. The continuous October 2007–September 2014 driving data comprise all of the meteorological variables required as inputs for physically based snow models at hourly intervals: incoming solar and longwave radiation, snowfall and rainfall rates, air temperature, humidity, wind speed and atmospheric pressure. Two versions of the driving data are provided: one using radiation and wind speed measurements made above the height of the trees around the clearing where the evaluation data were measured and one with adjustments for the influence of the trees on conditions close to the ground. The available evaluation data include automatic and manual measurements of bulk snow depth and snow water equivalent, and profiles of snow temperature, snow density and soil temperature. A physically based snow model is driven and evaluated with the datasets to illustrate their utility. Shading by trees is found to extend the duration of both modelled and observed snow cover on the ground by several days a year.

1 Introduction

Many studies have used meteorological data to drive snow models and meteorological or hydrological data to evaluate model performance at instrumented sites. These studies have often only used limited periods of driving data (e.g. two winters for several sites in Essery et al. (2009)) or limited evaluation data (e.g. infrequent manual measurements of snow mass in Slater et al. (2001)). Recently, valuable datasets have been published with multiple years of driving data and multiple sources of evaluation data for several snow research sites: Reynolds Mountain East in the Owyhee Mountains of Idaho (Reba et al., 2011), Col de Porte in the Chartreuse Mountains of France (Morin et al., 2012), the Senator Beck Basin in the San Juan Mountains of Colorado (Landry et al., 2014), Snoqualmie Pass in the Cascade Range of Washington (Wayand et al., 2015) and Weissflujoch in the Plessur Alps of Switzerland (WSL, 2015). All of these are high-elevation, mid-latitude sites; there has been a lack of comparable datasets that could be used for evaluating snow models at high latitudes.

Snow models operating on energy balance principles form components of land surface models that are used to provide energy and moisture flux boundary conditions for the atmosphere in numerical weather prediction and climate models, but they can also be driven with measured meteorological data. The typical input data required are downwelling shortwave and longwave radiation fluxes, precipitation rate, air temperature, humidity, wind speed and atmospheric pressure. All of these variables can be measured with low-power instruments, but all are challenging to measure in cold and snowy environments where instruments can be covered by snow or ice and access for maintenance may be difficult. Model-driving data have to be continuous, so gap filling is required if instrument or power failures occur. Data time steps
Table 1. Instruments and missing data for meteorological driving variables between 1 October 2007 and 30 September 2014.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Instrument</th>
<th>Height</th>
<th>Missing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>Vaisala FD12P</td>
<td>2 m</td>
<td>0.67 %</td>
</tr>
<tr>
<td>Air pressure</td>
<td>Vaisala PTB201A</td>
<td>1 m</td>
<td>0.11 %</td>
</tr>
<tr>
<td>Air temperature</td>
<td>Pentronic PT100</td>
<td>2 m</td>
<td>0.11 %</td>
</tr>
<tr>
<td>Relative humidity</td>
<td>Vaisala HMP35D</td>
<td>2 m</td>
<td>0.19 %</td>
</tr>
<tr>
<td>Global solar radiation</td>
<td>Kipp &amp; Zonen CM11</td>
<td>14 m</td>
<td>0.61 %</td>
</tr>
<tr>
<td>Diffuse solar radiation</td>
<td>Kipp &amp; Zonen CM11 with shading ball</td>
<td>14 m</td>
<td>0.55 %</td>
</tr>
<tr>
<td>Longwave radiation</td>
<td>Kipp &amp; Zonen CG4</td>
<td>14 m</td>
<td>8.65 %</td>
</tr>
<tr>
<td>Wind speed</td>
<td>Vaisala WAA25</td>
<td>22 m</td>
<td>0.12 %</td>
</tr>
</tbody>
</table>

Figure 1. (a) The location of FMI-ARC (dot), 90 km north of the Arctic Circle (dashed line) in Finland. (b) Orthophotograph of the FMI-ARC site, showing the locations of the automatic weather station (AWS), the radiometer tower (rad) and the intensive observation area (IOA). (c) The automatic weather station, with the radiometer tower in the background. (d) The radiometer tower. (e) The IOA, showing the locations of the ultrasonic depth gauge (UDG), the snow temperature profile and the snow pit for manual measurements.

2 Site

FMI-ARC (67.368° N, 26.633° E, 179 m above sea level, Fig. 1a) is collocated with the Sodankylä Geophysical Observatory beside the Kitinen River, 90 km north of the Arctic Circle and 7 km southeast of the town of Sodankylä in northern Finland. Snow typically lies from October until May; in daily records between 1951 and 2000, the annual maximum snow depth had a median of 83 cm, an interquartile range of 21 cm and a range from 62 cm (1954) to 119 cm (2000). Soil frost depths can reach over 2 m (Rautiainen et al., 2014), and air temperatures can fall below −30 °C in winter, but the sun only remains entirely below the horizon for a few days in December. Continuous meteorological measurements have been made at or near this site since 1908 (Kivi et al., 1999).
Figure 1b is an aerial orthophotograph of the site. The area around FMI-ARC is level and forested, predominantly with pine trees about 15 m tall, but many measurements are made in clearings or in a large wetland area to the east of the site. Driving data for this paper are taken from an automatic weather station (Fig. 1c) and a radiometer tower (Fig. 1d) 30 m apart, with instruments that are calibrated annually. Evaluation data are taken from an intensive observation area (IOA) (Fig. 1e) that was established 590 m to the south of the weather station for NoSREx (Lemmetyinen et al., 2016). A list of many other observations not discussed in this paper and contact information can be found at http://litdb.fmi.fi/index.php.

3 Driving data and gap filling

All of the meteorological variables necessary for model driving are measured by the AWS and the radiometer tower at FMI-ARC with the instruments and at the heights listed in Table 1; note that radiation and wind measurements are made at heights above the forest canopy. The radiometers are ventilated, the anemometer is heated to reduce problems with freezing or snow accumulation and instruments are cleaned after every snowfall or at least three times a week. Temperature and humidity sensors are naturally ventilated inside a Stevenson screen. Precipitation is measured using an optical sensor and two weighing gauges which give similar total amounts; data from the optical sensor are used here. There is no nearby wind speed measurement that could be used for gauge correction, but wind speeds are generally low, and measured snowfall has been adjusted to match snow accumulation on the ground as described below. FMI-ARC is staffed 5 days a week, and automatic error checking can identify instrument problems immediately. For the 7-year period collated here, fewer than 1 % of hourly data (visible as red points in Fig. 2) are missing for any variable with the exception of longwave radiation; the longest period of missing data is a 52-day gap in the longwave radiation measurements from 10 September to 31 October 2011 because of a faulty power supply. The archived driving data files include a flag that records which data were missing and had to be filled for each hour.

Measurements from the AWS and the radiometer tower are used for driving data whenever they are available, but gaps have to be filled to form a complete driving dataset. Gaps of 4 h or shorter are filled by linear interpolation. For shortwave radiation, air temperature, humidity and wind speed, longer gaps are filled with data from nearby instruments. No alternative longwave radiometer was operating at FMI-ARC for the full period, so longwave gaps are filled using ERA-Interim reanalyses (Dee et al., 2011). Longwave radiation fluxes in ERA-Interim are produced by short-range forecasts that can be expected to be accurate if the analysed vertical profiles of temperature and humidity in the atmosphere are accurate, although errors may be larger in cloudy conditions (Kangas et al., 2016). Data from both the surface synoptic station and the upper-air station at Sodankylä are available for assimilation in reanalyses, and ERA-Interim compares well with the...
in situ measurements; the longwave radiation measurements and forecasts have a correlation coefficient of 0.88 and a root mean square difference of 26.2 W m\(^{-2}\) after removal of a 5.1 W m\(^{-2}\) bias for periods when both are available (a scatter plot is included as the Supplement). Direct measurements of longwave radiation are rarely available for cold regions, and snow models are known to be sensitive to longwave driving data (Raleigh et al., 2016); having near-continuous longwave measurements is therefore a distinct advantage of the FMI-ARC site.

Seven-year series of gap-filled hourly data are shown in Fig. 2 for all of the driving variables apart from precipitation. Measuring solid precipitation is particularly challenging, and uncertainties in snowfall inputs are a major source of uncertainty in snow model outputs (Raleigh et al., 2015). Total precipitation is usually measured but has to be partitioned into snow and rain for mass balance calculations, either in the driving data or by the model. This is usually done by selecting a threshold or function of air temperature or wet-bulb temperature discriminating between rain and snow (Auer, 1974; Sims and Liu, 2015). Figure 3a shows the annual average snowfall partitioned from total precipitation for Sodankylä with varying temperature or wet-bulb temperature thresholds; the snowfall is not very sensitive to the choice of temperature or wet-bulb temperature as a predictor because humidity is usually high during precipitation, but it is sensitive to the choice of threshold because a significant amount of precipitation falls at temperatures close to 0 °C. With precipitation classified as snow for temperatures lower than 2 °C, Fig. 3b shows that the cumulated amount of snowfall is less than the maximum observed snow water equivalent (SWE) on the ground in most winters but slightly greater in 2010–2011 and 2012–2013. Because the site is cold and little melting of snow occurs in autumn or winter, the cumulated snowfall should be close to the amount of snow on the ground at points that are not affected by canopy interception or wind redistribution. Snowfall data are therefore scaled by the factors required to match the maximum measured SWE each winter (Table 2); cumulated snowfall then also matches the rate of accumulation on the ground quite well, as shown in Fig. 3b.

The snow measurement points in the IOA (Fig. 2e) are not directly beneath trees, so snow accumulation there will not be greatly affected by canopy interception, but they are shaded from direct solar radiation by nearby trees. The presence of the trees will also increase the incoming longwave radiation and decrease the wind speed near the ground relative to more open locations. Measurements above the forest canopy height do not take these influences into account. To allow the use of snow models without representations of forest canopies, radiation fluxes and wind speed are adjusted in a modified driving dataset. From the hemispherical image of the canopy at the IOA in Fig. 4a, the sky view fraction is estimated as \( f_v = 0.8 \) and a transmissivity \( \tau \) for direct solar radiation is calculated by determining the fraction of each hour for which the sun would be blocked by the canopy. Modified solar radiation is given by

\[
SW' = f_v SW_{dif} + \tau (SW - SW_{dif}),
\]

where \( SW \) and \( SW_{dif} \) are the measured incoming global and diffuse solar radiation (Reid et al., 2014). Longwave radiation is modified by assuming that the canopy temperature can be approximated by the air temperature (Essery et al., 2008; Pomeroy et al., 2009) so that

\[
LW' = f_v LW + (1 - f_v)\sigma T^4,
\]

where \( LW \) is the measured incoming longwave radiation, \( \sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4} \) is the Stefan–Boltzmann constant and \( T \) is the air temperature in kelvin. The resulting decreases in solar radiation and increases in longwave radiation are shown in Fig. 4b. Solar and longwave radiation will both be underestimated by these modifications close to tree trunks at the sun-lit northern edge of the IOA clearing where the snow is observed to melt first.

An anemometer installed temporarily at 2 m height above the ground close to the IOA for 7 days in March 2012

### Table 2. Scaling factors required to match measured snowfall to measured snow accumulation

<table>
<thead>
<tr>
<th>Winter</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007–2008</td>
<td>1.373</td>
</tr>
<tr>
<td>2008–2009</td>
<td>1.165</td>
</tr>
<tr>
<td>2009–2010</td>
<td>1.131</td>
</tr>
<tr>
<td>2010–2011</td>
<td>0.922</td>
</tr>
<tr>
<td>2011–2012</td>
<td>1.093</td>
</tr>
<tr>
<td>2012–2013</td>
<td>0.971</td>
</tr>
<tr>
<td>2013–2014</td>
<td>1.092</td>
</tr>
</tbody>
</table>
recorded an average wind speed that was 35% of the wind speed at 22 m height (equal to the ratio given by a logarithmic wind profile with a roughness length of 0.55 m). This ratio is used to scale the wind speed in the modified driving dataset. There is no permanently installed anemometer below the canopy height at the IOA, so the wind adjustment is highly uncertain. Because the wind is rarely strong enough to move snow in the IOA and snowmelt is dominated by radiation in spring, however, it is expected that models will not be highly sensitive to the wind adjustment.

4 Evaluation data

Physically based snow models may include snow temperature, mass, density, liquid water content and grain size in layers as state variables. Predicted fluxes will include reflected shortwave radiation, emitted longwave radiation, sensible and latent heat exchanges with the atmosphere, and conducted heat flux and drainage of water at the base of the snowpack. Snow properties that have to be predicted include albedo and thermal conductivity. Measurements of any state variable, flux or property may be used as evaluation data for models, and the measurements need not be continuous; measured and modelled variables can be compared at whatever times for which measurements are available.

FMI-ARC data that will be used in the model evaluation below are listed in Table 3. Again, many more measurements are made in the IOA in addition to those discussed here, including snow grain size, hardness, wetness, microwave brightness temperatures and soil moisture. The microstructure of snow samples taken during special experiments has been measured in great detail by X-ray-computed tomography (Proksch et al., 2015). Outgoing radiative and turbulent flux measurements are made above the canopy height at FMI-ARC, so they would be most useful for evaluating models that include vegetation canopies.

Snow depth and SWE are measured in the IOA both manually about once a week and many times daily with automatic instruments. These measurements are compared in Fig. 5. The output of the experimental SWE sensor, which works by measuring the attenuation of gamma radiation from a source beneath the snow, is noisy but tracks the manual measurements well after calibration and averaging. Snow accumulation varies spatially. Figure 6 compares the snow depth in the IOA for the winter of 2012–2013 with snow depths measured in the forest beside the IOA and 900 m to the northeast on the wetland. The snow was deepest throughout the winter and melted latest in the IOA. Some snow is intercepted by the forest canopy as it falls and can sublimate, reducing the depth of snow on the forest floor. Wind can remove and compact snow in the open wetland area, again reducing the snow depth. Differences in snow accumulation and melt rates lead to differences in the persistence of snow cover at different sites; the measured snow depth fell to zero on 3 May 2013 on the wetland, 6 May in the forest and 13 May in the IOA. Photographs of the IOA in Fig. 7 show small-scale variations in cover as the snow melts. Bare patches first appear around the bases of trees, and the snow lies longest at the shady side of the clearing.

Snow temperatures are measured continuously by an array of thermistors supported on a stick that becomes buried in the snow (http://litdb.fmi.fi/ioa0007_data.php) and inter-
Table 3. Evaluation data from the IOA.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow depth</td>
<td>Campbell Scientific SR50</td>
</tr>
<tr>
<td></td>
<td>Manual sampling</td>
</tr>
<tr>
<td>Snow water equivalent</td>
<td>Astrock Gamma Water Instrument</td>
</tr>
<tr>
<td></td>
<td>Manual sampling</td>
</tr>
<tr>
<td>Snow density profiles</td>
<td>Toikka Snow Fork sampling at 10 cm height increments from 09/10/2009</td>
</tr>
<tr>
<td></td>
<td>Manual sampling at 5 cm height increments from 07/12/2009</td>
</tr>
<tr>
<td>Snow temperature profiles</td>
<td>Campbell Scientific 107-L at 10 cm height increments from 06/09/2011</td>
</tr>
<tr>
<td></td>
<td>Manual sampling at 10 cm height increments</td>
</tr>
<tr>
<td>Soil temperature profiles</td>
<td>Decagon Devices 5TE at 5, 10, 20, 40 and 80 cm depths from 06/09/2011</td>
</tr>
</tbody>
</table>

Figure 7. Snow melting in the IOA at noon on (a) 1 May and (b) 13 May 2013.

mittently by inserting a stem thermometer into the snow face when pits are dug. Both methods are subject to biases; it has been observed that the thermistor stick interferes with the accumulation of snow and can form a depression up to 30 cm deep in the snow surface, and digging a snow pit brings air into contact with snow beneath the surface. Density is measured by weighing 250 or 500 cm$^3$ snow samples cut from snow pits (Leppänen et al., 2016) and also by a dielectric method that relates density and wetness to the measured permittivity of snow (Sihvola and Tiuri, 1986). The dielectric method generally gives lower densities than gravimetric sampling of snow at Sodankylä.

5 Model results

Preliminary versions of the driving and evaluation datasets were used in a study with the Joint UK Land Environment Simulator (JULES) land surface model by Ménard et al. (2015). The above-canopy and modified driving datasets are used here to drive Crocus (Vionnet et al., 2012), which is a detailed multi-layer snowpack model originally developed for avalanche forecasting in the French mountains. Although physically based, some of the processes in Crocus have been parametrized using experimental results from the mid-latitude site at Col de Porte (45.3° N, 5.8° E, 1325 m a.s.l.), which is much warmer than Sodankylä in winter and has heavier snowfall.

Figure 8 compares Crocus simulations driven by the above-canopy and below-canopy datasets with measurements of snow depth, SWE and soil temperature. Simulated snow depths are generally close to the measurements but are sometimes overestimated after snowfall because of Crocus predicting densities for fresh snow that are lower than observed at Sodankylä. Simulated SWE follows the measurements during the accumulation periods, which is to be expected because of the lack of mid-winter melt and the scaling of the snowfall in the driving data to the SWE measurements. Snowmelt starts at about the right time each spring in the simulations but then proceeds faster than observed. The modified driving data reduce melt rates; simulations with the above-canopy driving data remove the snow on average 13 days earlier than the snow disappearance dates inferred from the ultrasonic depth gauge at the IOA, but simulations with the modified below-canopy driving data remove the snow on average only 6 days earlier than observed. As shown by Fig. 7, the dates of snow disappearance can differ by 2 weeks even over short distances in reality; this spatial variability is not represented by a one-dimensional model such as Crocus. Simulated soil temperatures have cold spikes that are greater than observed at the start of some winters but then remain
close to 0°C once the snowpack has become established. Measured soil temperatures also show a strong influence of insulation by snow but can fall a couple of degrees lower than the simulations in late winter.

The frequent snow pit measurements in the IOA and the multi-layer outputs of Crocus give a large amount of data for comparison. Profiles of temperature and density for 140 snow pits dug between 7 December 2009 and 14 May 2014 are plotted in the Supplement, but the evolution of the snowpack over the winter of 2012–2013 alone is shown in Fig. 9. Snow pits were dug once a week, usually on Tuesday but sometimes on Wednesday or Thursday, for the 28 weeks between 31 October 2012 and 7 May 2013. Simulations and measurements both show temperatures remaining close to 0°C at the base of the snowpack with periods of much colder temperatures in snow layers close to the surface. The snow then rapidly warms and becomes wet and isothermal at 0°C when melt begins in spring. Density generally increases with depth in the snowpack and with time, again increasing rapidly once the snow becomes wet.

Quantitative comparisons between simulated and measured profiles of snow properties are complicated by differences in simulated and measured snow depths. Simply making scatter plots (Fig. 10) of variables at the same heights above the ground and at the same times shows strong correlations of 0.80 between simulated and measured snow temperatures and 0.74 for densities. The simulated temperatures tend to be higher than observed for the warmer temperatures found near the base of the snowpack.
6 Conclusions

Data from the FMI Arctic Research Centre at Sodankylä have been used to construct datasets that will allow driving of snow models for multiple years and evaluation of model outputs against multiple types of observations. There are some gaps in the data, but the availability of additional instruments and high-quality atmospheric reanalyses gives confidence in the filling of gaps to provide continuous driving data. The utility of the datasets has been demonstrated by driving the Crocus snow model and evaluating its outputs against snow depth, SWE, snow density, snow temperature and soil temperature measurements. The physical basis of the model allows it to perform well in an Arctic environment very different to the mid-latitude mountain environments for which it was first developed.

It is intended that Sodankylä will be used as a reference site in an upcoming evaluation of snow simulations in Earth System models (http://www.climate-cryosphere.org/activities/targeted/esm-snowmip/). Under the open-data policy of the Finnish Ministry of Transport and Communications, FMI is committed to the long-term upkeep and public distribution of its data; the datasets used in this paper can be downloaded from the FMI litdb archive at http://litdb.fmi.fi/ESMSnowMIP.php.

The Supplement related to this article is available online at doi:10.5194/gi-5-219-2016-supplement.

Acknowledgements. The staff at FMI-ARC are thanked for data collection and maintenance of instruments. Collection of evaluation data in the IOA was supported by ESA ESTEC contract 22671/09/NL/IA/ef. Visits to FMI-ARC by the first author were supported by NERC grant NE/H008187/1 and ESA ESTEC contract 23103/09/NL/JC. Samuel Morin and Matthieu Lafaysses assisted with the Crocus simulations. The orthophotograph in Fig. 1b was supplied by the Finnish Geospatial Research Institute. We are grateful to Charles Fierz and Mark Raleigh for their helpful comments in reviewing this paper.

Edited by: J. Pulliainen

References


WSL Institute for Snow and Avalanche Research SLF: Meteorological and snowpack measurements from Weissfluhjoch, Davos, Switzerland, Dataset, doi:10.16904/1, 2015.