Spring snow albedo feedback over Northern Eurasia: Comparing in-situ measurements with reanalysis products

Martin Wegmann\textsuperscript{1}, Emanuel Dutra\textsuperscript{2}, Hans-Werner Jacobi\textsuperscript{1} and Olga Zolina\textsuperscript{1,3}

\textsuperscript{1}Institute for Geosciences and Environmental Research (IGE), Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP*, Grenoble, France

* Institute of Engineering Univ. Grenoble Alpes

\textsuperscript{2}Instituto Dom Luiz, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal

\textsuperscript{3}P.P. Shirshov Institute of Oceanology, Moscow, Russia
This study uses daily observations and modern reanalyses in order to evaluate reanalysis products over Northern Eurasia regarding the spring snow albedo feedback (SAF) during the period from 2000 to 2013. We used the state of the art reanalyses ERA-Interim land and the Modern-Era Retrospective Analysis for Research and Applications Version 2 (MERRA2) as well as an experimental setup of ERA-Interim land with prescribed short grass as land cover to enhance the comparibility with the station data. While snow depth statistics derived from daily station data are well reproduced in all three reanalyses, the day-to-day variability of the albedo is notably higher at stations compared to any reanalysis product. The ERA-Interim grass setup shows an improved performance in representing albedo variability and generates comparable estimates for the snow albedo in spring. We find that modern reanalyses show a physically consistent representation of SAF, with realistic spatial patterns and area-averaged sensitivity estimates. However, station-based SAF values are significantly higher than in the reanalyses, which is mostly driven by the stronger contrast between snow and snow-free albedo. Switching to grass-only vegetation in ERA-Interim land increases the SAF values up to the level of station-based estimates. We found no significant trend in the examined 14-year timeseries of SAF, but inter-annual changes of about 0.5% K⁻¹ in both station-based and reanalysis estimates were derived. This inter-annual variability is primarily dominated by the variability in the snow melt sensitivity, which is correctly captured in reanalysis products. Although modern reanalyses perform well for snow variables, efforts should be made to improve the representation of dynamic albedo changes.
1. Introduction

Global warming is enhanced at high northern latitudes, where the Arctic near-surface air temperature has risen at twice the rate of the global average in recent decades—a feature called Arctic amplification (Serreze and Barry 2011). Climate model experiments for the 21st and 22nd centuries show that the Arctic warming will continue and intensify under all emission scenarios (Collins et al. 2013). Arctic amplification of the global warming signal results from several processes interacting with each other such as the albedo feedback due to a reduction in snow and ice cover, enhanced poleward atmospheric and oceanic heat transport, and changes in humidity (Serreze and Barry 2011).

Being one of the critical factors of the Arctic amplification, the surface albedo feedback implies that the additional amount of reflected shortwave radiation at the top of the atmosphere decreases with decreasing surface albedo whereas near-surface air temperature increases with decreasing surface albedo (Thackeray and Fletcher 2016). It is considered to be a positive feedback in the sense that an initial warming leads to a warming strengthening over time quantified through the change in surface albedo per unit change of temperature (Robock 1983, Cess et al. 1991, Qu and Hall 2007). Snow can cause such a feedback since a snow-free surface absorbs more shortwave radiation and converts the energy to longwave radiation and convection, which warm the lower layers of the troposphere (Curry et al. 1996). This snow albedo feedback (SAF) and its impact on climate have been studied for several decades (Wexler et al. 1953, Budyko 1969, Schneider and Dickinson 1974, Lian and Cess 1977). It got further attention in the wake of anthropogenic global warming accompanied by the reduction of snow and ice cover over the Northern Hemisphere (Bony et al. 2006, Qu and Hall 2007, Fernandes et al. 2009, Flanner et al. 2011, Qu & Hall 2014, Fletcher et al. 2015, Thackeray and Fletcher 2016).

During 1979–2011, the Arctic snow cover extent in June decreased at a rate of -21% per decade (Derkson and Brown 2012). Climate model projections for the end of the 21st century show an even more reduced Arctic cryosphere and, thus, the SAF will continue to modulate Arctic warming (Brutel-Vuille et al. 2013). The SAF is especially effective over the Northern Hemisphere (NH) since most of the NH is
covered by snow during boreal wintertime (Groisman et al. 1994). Hall (2004) found that 50% of the total NH SAF caused by global warming occurs during spring, while Qu and Hall (2014) estimated that the SAF variability accounts for 40-50% of the spread in the warming signal over the continents of the NH.

Several studies investigated spring NH SAF based on satellite, reanalysis and model datasets (Fernandes et al. 2009, Fletcher et al. 2012, Qu and Hall 2014, Fletcher et al. 2015). Satellite-based estimates of SAF vary within ±10% depending on the analysed data set. Hall et al. (2008) used the International Satellite Cloud Climatology Project (ISCCP) data (Schiffer and Rossow 1983) to calculate an SAF strength of -1.13% K⁻¹, whereas Fernandes et al. (2009) using Advanced Very High Resolution Radiometer (AVHRR) data (Justice et al. 1985) found a slightly weaker SAF of -0.93% K⁻¹. Qu and Hall (2014) determined the SAF using Moderate Resolution Imaging Spectroradiometer (MODIS) data (Hall et al. 2002) and found a value of -0.87% K⁻¹ for springtime. Considering different spatial and temporal domains as well as the variety of methods applied, the SAF estimates around -1% K⁻¹ from satellite data can be considered as quantitatively consistent.

Model- and reanalysis-based estimates are somewhat higher compared to those derived from satellite data. Fletcher et al. (2015) investigated CMIP3 and CMIP5 ensembles to estimate the SAF for an assortment of Global Climate Models (GCMs). From a large set of SAF estimates for individual models, they found an ensemble mean of -1.2% K⁻¹ which is in fair agreement with MODIS values, but is higher compared to ISCCP- and AVHHR-based estimates. Within this comparison Fletcher et al. (2015) also investigated SAF computations based on ERA-Interim (Dee et al. 2011), Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al. 2011) and NCEP-2 (Kanamitsu et al. 2002) reanalyses, thus, providing the most up to date assessment of SAF in reanalysis datasets. While MERRA data resulted in a slightly weaker SAF of -1.17% K⁻¹ compared to ERA-Interim (-1.23% K⁻¹), both reanalyses show similar SAF values compared to MODIS.
Although satellite products cover large parts of the NH, they exhibit low temporal resolution and significant uncertainties for high solar zenith angles as well as complex terrains (e.g. Wang et al. 2014). Thackeray and Fletcher (2016) compared CMIP3/CMIP5 model families and found that the models represent the SAF process rather accurately. However, there are still inherent biases likely related to the use of outdated parameterizations. In this respect the use of in-situ observations would provide an opportunity for evaluating SAF estimates in different gridded datasets and especially among reanalyses. However, estimating SAF in the Arctic using in-situ data is challenging, mostly because of the lack of reliable, relevant observations, both in the temporal and spatial domain. Furthermore, the lack of in-situ SAF estimates hampers the understanding of SAF in high latitude climates (Graversen and Wang 2009, Gravesen et al. 2014).

In this study we use a unique dataset of daily observations and modern reanalyses over Northern Eurasia in order (1) to evaluate reanalysis products with respect to radiation and snow properties and (2) to determine the SAF in spring between 2000–2013 based on in-situ measurements. We compare different land-reanalysis products with modified vegetation settings. Specific questions to be addressed in this study are the following: How well do the modern reanalyses reproduce snow and radiation features on a daily resolution? What are realistic estimates of the SAF from the station data over Northern Eurasia and how well do they compare to the gridded reanalyses data? What are the major characteristics of space-time variability of the SAF in station and reanalysis data?

The paper is organized as follows. After describing the different datasets and the methods in sections 2 & 3, we evaluate the daily output for snow, radiation fluxes and temperature within these datasets in section 4.1. In section 4.2 we assess the results of the SAF computations and the differences between products including also an analysis of the spatial and temporal variabilities. Section 5 discusses the results and considers potential implications for future studies.

2. Data

2.1 Reanalysis Data
To investigate the SAF processes in reanalyses, we evaluated two products: the ERA-Interim-land (ERAI-L, Balsamo et al. 2015) and Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) (Gelaro et al. 2017). ERAI-L is a land-surface only simulation driven by the near-surface meteorology and fluxes from the ERA-Interim atmospheric reanalyses (Dee et al. 2011). The land-surface model in ERAI-L (HTESSEL) has several enhancements compared with the land-surface model used in ERA-Interim including the snowpack representation (Dutra et al. 2010). ERAI-L considers the prognostic evolution of snow mass and density, and for exposed areas there is also a prognostic evolution of snow albedo. For shaded snow, i.e. snow under high vegetation, the albedo is considered constant and dependent on vegetation type (see Dutra et al. 2010 for more details). Since the observations used in this study are local, and in the case of forest regions likely represent a clearcut in the forest, idealized simulations prescribing grassland everywhere were carried out with the ERAI-L configuration (hereafter ERA-Interim land grass only (ERAI-LG)). The main goal of this simulation is to evaluate the role of land cover when comparing point observations with gridded reanalysis and to evaluate pathways to improve reanalyses in representing albedo processes.

MERRA2 also includes a dedicated land module for surface variables. Furthermore, it applies an updated Goddard Earth Observing System (GEOS) model and analysis scheme and assimilates more observations than its predecessor MERRA (Rienecker et al. 2011). Finally, MERRA2 uses observation-based precipitation data to force its land-surface parameterizations, similar to what formerly was known as MERRA-land. Unlike ERAI-L, MERRA2 consists of a full land-atmosphere reanalysis. Its incremental analysis update (IAU) scheme improves upon 3D-Var by dampening the analysis increment. In IAU, a correction is applied to the forecast model gradually, limiting precipitation spinup in particular.

For near-surface temperature we use 2m air temperature for both the reanalyses and observations. Moreover, we do not use albedo diagnosed by the reanalysis, but calculate it from the radiative flux components consistent with the observed albedo. For this purpose we use upward and downward shortwave radiation at the surface as diagnosed by ERA-Interim and MERRA2 as well as surface net and surface incoming radiation from the station observations. Snow depth is used as diagnosed by reanalyses and, if needed, converted to cm.
More information about general characteristics of reanalysis products in the Arctic can be found in Lindsay et al. (2014), Dufour et al. (2016) and Wegmann et al. (2017).

2.2 Observational in-situ data

To evaluate reanalysis performance, we used newly assembled in-situ radiation observations from Russian meteorological stations. This dataset includes 4-hourly Solar Radiation and Radiation Balance Data from the WMO World Radiation Network of the World Radiation Data Center (WRDC) at the Voeikov Main Geophysical Observatory, Saint Petersburg, Russia. The original WRDC data contains time series (1964–2015) from 65 locations. Of these, we selected 47 stations for this study because they overlap with daily snow depth and 2m temperature observations (see Supplement Table 1). Of these 47 stations three were attributed by ERAI-L to ocean areas, so that the final dataset consists of 44 stations. Temperature and snow depth observations were taken from the All-Russian Research Institute of Hydrometeorological Information World Data Centre (RIHMI-WDC), Obninsk, Russia. A detailed description of this dataset is provided by Bulygina et al. (2010). This dataset includes snow depth as well as snow cover over an area around meteorological stations. Snow cover information in this data set is not stored in percentages, but rather in a scale of integers from 0 to 10 (for example, 50% is assigned a value of 5, but so is 53%). This makes these data hardly applicable for precise SAF calculations. Snow depth information is measured in centimeters with the precision of 1 cm. This might lead to an underestimation of snow depth in case of shallow snow (between 0 and 1 cm). All variables (temperature, snow depth and snow cover, surface LW radiation budget and surface SW radiation, the sum of the surface short-wave and long-wave radiation budgets) were represented as daily time series for the period 2000–2013.

Figure 1 shows the location of the stations together with the climatological 2000–2013 MAMJ snow depth as computed by ERAI-L. The distribution of stations is quite heterogeneous, with very few stations located in Eastern Siberia and in the Far East. Moreover, some stations have prolonged periods of missing values; six stations have more than 50% missing values in the daily timeseries for MAMJ. For monthly means, the total number of missing values generally decreases from 2000 to 2013 (see
Supplementary Figure 1), However, data for the year 2009 are missing at 44 out of 47 stations for the MAM period and for 3 stations also June values are missing. Nevertheless, spatial and temporal coverage of this data set is exceptional for the analysis of albedo in this region. It is also important to note that neither snow nor radiation from these stations were assimilated in the reanalysis datasets and, therefore, our inter-comparisons are completely independent.

Figure 1: Station location and snowdepth [cm] for the 2000–2013 MAMJ average taken from ERAI-L. Red colored stations are excluded by the land-sea mask of ERAI-L.

3. Methods

To evaluate the climatic variables needed for the SAF computation, we first compared daily values of snow depth, albedo and 2m temperature from the meteorological stations with those from the reanalyses. To co-locate observations with reanalyses, we extracted the information of the gridcell from the reanalysis, in which the station is located. We then derived long-term differences, performed a correlation analysis and also compared the variability among the datasets for the MAMJ period.
Since the SAF signals for the seasonal cycle and for the long-term climate change signal are highly correlated (Hall and Qu 2006), we focus here on the evaluation of the seasonal cycle. Snow cover is converted from snow depth following a logarithmic equation according to which 2.5 cm of snow depth was defined as equivalent to 100% snow cover (Fletcher et al. 2015). In most analyses, SAF is split into a snow cover component (SNC) and a temperature/metamorphosis component (TEM). SNC relates to the decrease of the albedo linked to the earlier melting of snow, which causes the exposition of the surface with a much reduced albedo. TEM concerns the reduction of snow albedo due to enhanced metamorphism and larger grain sizes at warmer temperatures. SAF is computed as sum of the two components, SNC and TEM, according to:

\[
SNC = (\overline{\alpha_{\text{snow}}} - \overline{\alpha_{\text{land}}}) \frac{\Delta S_c}{\Delta T_{2m}}
\]  

\[
TEM = \overline{S_c} \frac{\Delta \alpha_{\text{snow}}}{\Delta T_{2m}}
\]

where \(\alpha_{\text{snow}}\) is the snow-covered surface albedo, \(\alpha_{\text{land}}\) is the snow-free surface albedo, \(S_c\) is the snow cover fraction and \(T_{2m}\) is the 2 m temperature. The first term of (1) is also known as albedo contrast, whereas the second term will be referred to as snow melt sensitivity. In (1) and (2) deltas indicate month-to-month changes and the overbars indicate means over the two adjacent months. Note that \(\Delta T_{2m}\) does not represent a hemispheric mean but rather the difference at an individual location. It was found that the contribution of SNC and TEM to the overall SAF is between 60 to 70% and 30 to 40% for the NH (Fletcher et al. 2015).

Since daily data are available, we define \(\alpha_{\text{snow}}\) as the monthly mean over all daily estimates during the specific month when \(S_c = 100\%\). Moreover, we define \(\alpha_{\text{land}}\) as the mean over all daily estimates during MAMJ when \(S_c = 0\%\). This allows for a more realistic estimation of \(\alpha_{\text{land}}\) than conventionally using summer (e.g. August) albedo.

4 Results

4.1 Daily data evaluation
Since 2m air temperature in reanalyses has been comprehensively evaluated in previous studies (e.g. Schubert et al. 2014, Lindsay et al. 2014), we only perform a general comparative assessment of the daily values of albedo and snow depth involved in the SAF computations.

Figure 2 shows an overall comparison between station data and reanalyses in terms of correlations, differences and magnitude of variability quantified by the standard deviation for the albedo and snow depths. On a day-to-day basis MERRA2 and ERAI-L are underestimating average albedo values compared to observations by about 0.1 during MAMJ (Figure 2a). On the other hand, ERAI-LG shows a much smaller average deviation from the station data with differences close to zero. However, the overall range of the boxplot for ERAI-LG is similar to the other two reanalyses resulting in only slightly less absolute deviations from the observations.

For snow depth (Figure 2b), all three reanalysis datasets show an overestimation of daily values for MAMJ. Interestingly, ERAI-LG shows the largest deviations from observed values, although the grass represents better the conditions at the observational sites. This can be caused by biases in the observations due to surrounding higher vegetation creating a snowfall shadow or negative instrumental biases (Rasmussen et al. 2012). Moreover, positive biases in particular for precipitation can occur in reanalysis products (Brun et al. 2013).

The analysis of daily correlations (Figure 2c and d) demonstrates that the correlations for the albedo are generally low among all three experiments, whereas for some stations they can reach correlation coefficients higher than 0.8. Surprisingly, the correlations between MERRA2 and station data are the highest for albedo and the lowest for snow depth. The observed difference between MERRA2 and the ECMWF experiments regarding the correlation for albedo can likely be explained by the introduction of aerosols (and their respective deposition) in MERRA2. These findings suggest that further studies are needed to investigate the impact of aerosols on snow albedo representation. For snow depth, the correlation values are dominated by snowfall and melting events. Also in this case, the grass-only experiment shows no increased performance compared to the classic ERAI setup.

Considering the representation of day-to-day variability (Figure 2e and f), all reanalyses severely underestimate the day-to-day variability of the albedo. MERRA2...
and ERAI-L show similar means, but reach the overall station level only in specific grid cells. A clear improvement is observed in ERAI-LG, which shows the smallest deviation from station estimates. Nevertheless, all modern reanalyses fail to adequately reproduce daily variability in the observed albedo. On the other hand, for snow depth the agreement is very good. The means of all four products are around the values of 8 to 10 cm, with the grass-only experiment being the closest to the average station variability.

In summary, the boxplot analysis (Figures 2) reveals that there is a general improvement in agreement between stations and ERAI-L if vegetation is set to grass only. However, none of the reanalysis products can accurately reproduce day-to-day albedo variability. This is likely explained by the comparison of grid versus point observations, where small-scale variations are averaged out. Moreover, observed snow-free albedo depends on short-term changes linked to the vegetation and meteorology for example causing frost or modifying soil moisture.
Figure 2: Boxplot analysis for daily albedo (a, c, e) and snow depth (b, d, f) estimates using data from 44 locations over 2000–2013 MAMJ period. (a) and (b) Difference between station and reanalysis, (c) and (d) linear correlation between station and reanalysis, (e) and (f) standard deviation. Triangle indicates the mean value.
4.2 Analysis of feedback components

To assess regional patterns of key SAF components, we show their spatial distribution over Russia as revealed by the observations in Figure 3 (See Supplement Figures 2-4 for the respective distribution from the reanalyses data).

Strong SNC (Figure 3a) responses in the station data are observed in Southern European Russia and Western Siberia as well as over the Far East. The weaker responses are observed in Southern Eastern Siberia. TEM (Figure 3b) follows a similar distribution but is more homogeneously distributed with most negative values in Central Siberia and towards the Arctic coastline. Snow melt sensitivity (Figure 3c) is strongest in the mid-latitudinal and subpolar regions north of 50° N, such as Finland to the southeast, west and north of Lake Baikal and along the Pacific Coast. Here the temperatures react most strongly to seasonal snow melt. While there is a broad agreement between the stations and ERAI-LG in this region, stations show a somewhat stronger snow melt sensitivity (not shown). Snow melt sensitivity is a key factor for the SNC calculations and, thus, shapes the spatial variability of SNC.

The other key factor in the SNC calculations is the contrast in albedo between snow-covered and snow-free periods (Figure 3d). The observed albedo contrast is characterized by a relatively homogeneous pattern with somewhat smaller values in the southern regions, especially over Southern Eastern Siberia east of the Lake Baikal. In general, a north-south gradient is visible with similar patterns as in SNC. Mean albedo for the spring season (Figure 3e) shows that highest values are found closer to the Arctic coastline, in Central Siberia and towards the western border. Lower mean albedo values are mostly located east of Lake Baikal. This distribution is in general agreement with the reanalyses datasets, especially for the lower values in the south east.

Finally, since TEM follows closely the general MAMJ snow distribution, we show average snow depth in Figure 3f. A clear north-south gradient is visible with hotspots at the Pacific coast and towards the Barents-Kara sea. Moreover, snow depths from stations follow closely the ERA-L snowdepth distribution shown in Figure 1.
To analyse the differences between the datasets and to highlight the context of the station data, Figure 4a shows the response for SAF computed for the entire period 2000-2013 and all 44 locations. Stations show much stronger SAF (-2.5% K⁻¹) compared to MERRA (-1.6% K⁻¹) and ERAI-L (-1.8% K⁻¹). At the same time ERAI-LG shows SAF estimate close to that derived from the station data (-2.8% K⁻¹). Thus, changing the vegetation to short grass adds about 1 K to the responses revealed by classic reanalyses making the results close to observations.
The further analysis of the two components of SAF (SNC and TEM, Figure 4b and c) shows that ERAI-LG reproduces well the SNC signal derived from the station data (-1.6% K\(^{-1}\) mean for stations and -1.7% K\(^{-1}\) mean for ERAI-LG), whereas the other two reanalyses show much weaker SNC values. The lowest value of -0.56% K\(^{-1}\) was obtained from the MERRA2 data. In general, SNC responses largely explain differences in SAF (Figure 4a).

For TEM values (Figure 4c), all three reanalyses are in a good agreement with the observations with MERRA2 showing the best agreement. Changing the vegetation to grass in ERA-Interim results in a TEM component, which is 0.4-0.5% K\(^{-1}\) stronger compared to the standard version of ERA-Interim. Given that TEM represents the response to snow metamorphosis, good performance of MERRA2 is in agreement with findings implied by Figure 2. However it is worth noting that for the station network as well as for the ECMWF experiments, locations with positive TEM are calculated. This is due to snow albedo changes being positive in some instances (Figure 4c).

To further investigate the nature of the SNC and TEM responses we show in Figure 4d the results for snow melt sensitivity, which is one of the two key components in the SNC response (1). This component is barely influenced by the underlying vegetation. All three reanalysis datasets agree very well with the station network, with ERAI-LG showing the closest agreement for both mean and median. This indicates an accurate representation of this relationship in both NASA and ECMWF land surface modules.

Figure 4d implies that the changes in the SNC should stem from the albedo contrast, the second key component expressed as the average difference between albedo values for a complete snowcover and snow-free conditions (Figure 4e). Indeed, MERRA2 shows the lowest albedo contrast among all datasets, resulting in very low SNC values. Albedo contrast in ERAI-L is higher than MERRA2, but is on average still lower compared to the observations, which show average values around 0.35. ERAI-LG shows the strongest albedo contrast, which is twice as large compared to the experiment with classic vegetation cover. These striking differences among the datasets mainly drive the SNC results.
Figure 4: Boxplot analysis for MAMJ 2000–2013 a) SAF, b) SNC, c) TEM, d) snow melt sensitivity, e) albedo contrast and f) snow albedo. Triangle indicates the mean value.

Snow albedo is well captured by the grass-only experiment showing the same average value around 0.6 as determined from the observations (Figure 4f). The standard vegetation schemes used in MERRA2 and ERAI-L reduce the snow albedo in the analyzed grid cells to 0.33 and 0.37. The differences in snow albedo between the products is the main driver for the differences in the albedo contrast since the snow-free albedo values are remarkably similar for all reanalysis products (Figure 5a). Nevertheless, they strongly deviate from the snow-free albedo determined from the observations, which is roughly twice as large compared to the reanalyses with a mean
value of about 0.21 and which is very close to albedo values for grass (see e.g. Betts and Ball 1997, Wei et al. 2001).

To explore the impact of different factors on the TEM estimates, we show in Figure 5 mean values of temperature, snow cover and albedo, as well as the average change of snow albedo during spring. Also, to underline the crucial role of in-situ snow depth information, mean snow depth is shown. Mean station snow depth lies within the range of reanalyses values, with higher values reported by ERAI-LG. Moreover, stations have the lowest snow cover among all datasets (Figure 5b and c). This difference is likely due to the conversion of snow depth to snow cover as well as the precision (in centimeters) of the Russian snow depth measurement. Precision of snow depth diagnosed by reanalysis is much finer and the logarithmic conversion here can be performed more accurately. As a result, TEM values diagnosed by stations are probably too low. If we consider instead in-situ snow cover information from stations, the average snow cover is quite similar to reanalyses (ca. 55%), and the average TEM value gets stronger. However, replacing converted snow cover with observed snow cover in Eq. (2) is a questionable procedure, as the remaining terms were computed using snow depth conversion. Thus, for consistency we show lower values of TEM in Figure 4.

Temperature is well represented by all datasets with MERRA2 being about 1 K colder compared to stations, which is quite notable for such a robust variable. However, absolute values of temperature do not have a strong impact on the computation of TEM, since month-to-month changes in temperature affect both TEM and SNC computations. For ERAI-LG, the effect of the underestimated snow-free albedo and overestimated complete snow cover albedo cancel each other out. Finally, the snow albedo change during spring season (Figure 5f) is very similar in station data and in MERRA2 (-0.09 average in both datasets), which points towards an adequate representation of snow metamorphosis and aerosol deposition in MERRA2. The ERAI-LG experiment shows a stronger change of snow albedo during spring than the standard version. ERAI-L potentially keeps the temperature and therefore snow metamorphosis more constant throughout spring season due to a more stable local temperature climate induced by the vegetation. Note also, that some stations show an increase of snow albedo during spring. This can be caused by fresh snow accumulation in late spring in some locations.
Figure 5: Boxplot analysis for MAMJ 2000–2013 a) snow free albedo, b) snow cover fraction, where the light grey boxplot is the originally observed snow cover from stations, c) snow depth, d) 2m temperature, e) mean albedo and f) snow albedo change within the season. Triangle indicates the mean value.

Figure 6 shows timeseries (2000–2013) for the mean values for SAF-related variables. Timeseries for SNC (Figure 6a) and TEM (Figure 6b) show that inter-annual
variations of up to 0.5% K\textsuperscript{-1} are possible for both stations and reanalyses. Moreover, for both SNC and TEM, ERAI-LG seems to reproduce well the overall baseline and the magnitude of variability.

For snow melt sensitivity (Figure 6c) the agreement among the datasets is very good for both magnitude and interannual variability, with MERRA2 showing an amplified inter-annual variability (up to 1.5% K\textsuperscript{-1}), which is beyond the magnitudes observed at stations. As already noted above, snow melt sensitivity seems to be a rather well reproduced process in modern reanalyses. Since snow-free albedo is quite constant over time in the reanalyses, the albedo contrast is dominated by the snow albedo (Figure 6d). ERAI-LG and the station network agree very well on the magnitude of snow albedo, whereas ERAI-L and MERRA2 fail to reproduce such high values. Magnitudes of inter-annual variability can reach up to ±0.05 in stations, with slightly weaker response in reanalyses. Correlation between stations and reanalyses is rather low, only individual years are captured correctly by ERAI-LG (see Supplement for correlation values).

Snow albedo change within spring season (Figure 6e) is well captured by MERRA2 and ERAI-LG. Furthermore, ERAI-LG captures well the inter-annual variability for this metric. Specifically, variability during 2001–2004 and 2005–2008 periods is quite well represented. On the other hand, ERAI-L seems to lack the consistency with observations. Finally, as it was mentioned in section 4.1, snow depth variability (Figure 6f) is very well captured by all reanalyses. Again, ERAI-LG overestimates snow depth by up to 5 cm, with the other two reanalyses being on average 1-2 cm above the station values.
Figure 6: Yearly timeseries of selected MAMJ SAF components averaged over all 44 locations. 

a) SNC, b) TEM, c) snow melt sensitivity, d) snow albedo, e) snow albedo change within the season, f) snow depth.

To further demonstrate the effect of the vegetation changes in the ERA-Interim land reanalysis, Figure 7 shows anomalies between ERAI-L and ERAI-LG. The structure
follows Figure 6, with SNC and TEM shown in Figure 7a&b. As is clearly visible both variables are generally less negative in ERAI-L, a fact already known from timeseries and boxplot analysis. The largest impact of the vegetation changes is found for Northern Russia, the Pacific coast and the western region between Black and Caspian Sea. Interestingly, but as expected, snow melt sensitivity (Figure 6c) is not the key driver behind this distribution. Since snow melt sensitivity is not directly linked to vegetation changes, the anomaly distribution is very heterogenous, with positive and negative anomalies over the whole domain. As known from the timeseries plot, snow sensitivity in ERAI-LG is overall slightly weaker than in ERAI-L, probably due to positive feedbacks such as reduction of nighttime cooling over higher vegetation types. The main driver behind the distribution of SNC is albedo contrast (Figure 7d). Albedo contrast is overall higher in ERAI-LG, especially along the borders of the domain, highlighted already for SNC.
5. Discussion

We compared spring SAF and its components determined from in-situ measurements over Russia for the period 2000–2013 with data derived from three modern reanalysis products restricted to the grid cells including the observational sites. This was achieved by using a unique collection of station measurements of radiation and snow characteristics investigating for the first time observed SAF over this broad spatial and temporal domain. Besides ERAI-L we also used a customized version of ERAI-L (ERAI-LG), in which vegetation was set to grass in all concerned grid cells. All three reanalysis datasets are completely independent from the analyzed station data. While
a direct comparison of point measurements with grid cell output always introduces uncertainties properties due to the spatial variability of the surface, this is for now the only way to evaluate reanalyses data using in-situ observations. An alternative option would be the satellite data, which come with their own uncertainties (e.g. Romanov et al. 2002, Foster et al. 2005, Wang et al. 2014).

Snow depth statistics derived from daily station data are reasonably well reproduced in all three modern reanalyses, which is in agreement with Wegmann et al. (2017) who investigated April snow depth in ERAI-L. While snow depth differences between ERAI-L and ERAI-LG are small, ERAI-LG shows slightly higher deviations from the station data than ERAI-L that might be caused by the higher vegetation in station surroundings and by underestimation of snowfall due to instrumentation used at the Russian station network (Rasmussen et al. 2012).

Day-to-day variability of albedo is notably higher in station data compared to any reanalysis product. Besides spatial averaging over the reanalyses grid cells, this is potentially caused by land surface changes due to weather (e.g. vegetation changes, flooding, frost, aerosol deposition), which are not represented in the reanalyses. However, ERAI-LG demonstrates increasing albedo variability, nearly doubling the standard deviations diagnosed by ERAI-L with the standard vegetation scheme.

The limitations of the station data imply some constraints for comparisons with reanalysed data. As near-surface temperature is unavailable in station data, we used for both stations and reanalyses 2m air temperature, which reduces the strength of the SAF feedback. Secondly, snow cover is underestimated in station data due to the measurement precision of 1cm, which reduces the strength of the TEM component. The snow albedo and the snow-free albedo are substantially higher in station data than in the reanalyses with classic vegetation boundary conditions (MERRA2 and ERAI-L). Compared to other observation-based studies, spring snow albedo and grass albedo derived from our station network is quite realistic (Roesch et al. 2009, Stroeve et al. 2006). Thus, the difference revealed by reanalyses is likely due to averaging over grid cells.

Results from ERAI-LG clearly demonstrate that SAF and its components are very close to those in the station data. The largest improvement was found for albedo contrast and for snow albedo, which both are more realistic in ERAI-LG. At the same
time snow-free albedo in all three reanalyses (including ERAI-LG) was found to be lower than in the station database because snow-free albedo in all reanalysis data sets is prescribed as a monthly climatology from MODIS data.

MERRA2 shows the lowest SAF values resulting from a very low albedo contrast, which is probably a consequence of the vegetation scheme in the MERRA2 land module. On the other hand, MERRA2 represents TEM reasonably well most likely due to the accurate representation of the intra-seasonal snow albedo changes. Thus, relative snowpack changes appear to be well represented in MERRA2, probably also due to a more accurate representation of aerosols.

In general, we found higher SAF values in ERAI-L than in the recent CMIP3/5 analyses of NH SAF by Fletcher et al. (2015). This disagreement results from a variety of factors. First, our domain is limited to Russia only, thus excluding considerable parts of Eurasia as well as North America. In this respect our domain is set within a high SAF region, which may explain higher SAF values compared to the NH average by Fletcher et al. (2015). On the other hand, MERRA2 shows good agreements with the NH CMIP4/5 SAF results, however mostly because the albedo contrast is very low. Furthermore, as we pointed out above, in-situ observations used here tend to slightly overestimate SAF, mainly due to higher snow albedo values. This is because in-situ snow albedo is typically measured by a sensor installed over a vegetation-free snow pack. The vegetation scheme used in reanalyses gives lower snow albedo values implying realistic vegetation cover such as taiga or tundra. However, our MERRA2 results agree fairly well with the findings of Fletcher et al. (2015). Moreover, mean values of the albedo independent variable snow melt sensitivity are very close to the "observational" snow melt sensitivity computed by Fletcher et al. (2015).

We also found agreements with Fletcher et al. (2015) in the representation of the spatial pattern of the SAF components. Fletcher et al. (2015) as well as Fernandes et al. (2009) have shown maxima in SAF over northern Canada, northern Siberia and southwestern Eurasia. The relation of 60:40 found in satellites and reanalysis for SNC to TEM was replicated by our station network. We found similar spatial patterns for SAF and its components in both stations and gridded data specifically for Southern Russia, while the pattern of station responses is less homogenous compared to the
gridded data. Also consistent with Fletcher et al. (2015), we found higher snow melt
sensitivity north of 50° N. Finally, albedo contrast distribution, which closely follows
the snow albedo pattern, is in very good agreement with the gridded analysis of snow
albedo by Fletcher et al. (2015).

6. Conclusions

Reanalyses including land surface modules show a physically consistent
representation of SAF with realistic spatial patterns and area-averaged sensitivity
estimates. ERAI-LG shows a better performance in representing station-based
estimates considering the uncertainty associated with "point to grid cell" comparisons.
Accounting for aerosol-related processes would likely improve this performance in
future reanalysis releases. Thus, for the analysis and validation of large-scale temporal
and spatial averages of SAF modern reanalyses seem to be an appropriate tool.

However, for analysing processes on smaller scales and high temporal resolution
studies, a healthy dense station network is required. The idealized ERAI-LG
simulation also highlights the caveats of comparing in-situ observations with gridded
model data. In this study, we show these discrepancies in terms of albedo and snow
depth. Other variables, in particular 2m temperature, can be expected to have a similar
signal arising from the differences between the model’s gridcell land cover and the
actual station conditions. Our findings show that the experimental approach in ERAI-
LG allows for an enhanced use of in-situ observations to diagnose the SAF in not-
forested areas.

Considering future studies, the extension to other regions and use of other regional in-
situ data might give further insights into regional hotspots of SAF. Cross-validation
efforts employing model, reanalysis, satellite and station data may help to generate
blended products to investigate radiation and albedo feedbacks in the changing Arctic,
a region where SAF is especially strong. Regional modelling, including a variety of
multi-layer land surface models over areas with a relatively dense observation
network can provide a quantitative estimation of uncertainties among complex
variables such as snow depth, albedo or SAF.
Acknowledgements. This study was supported by the ARCTIC-ERA project funded by the Belmont Forum Fund through the ANR. OZ also benefited from the support by the Russian Ministry of Education and Science (project no. 14.B25.31.0026). ED was supported by FCT (IF/00817/2015).
References


Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P.,
Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C.,
Weaver, A.J. & Wehner M. 2013: Long-term Climate Change: Projections,
Commitments and Irreversibility. In: Climate Change 2013: The Physical Science
Basis. Contribution of Working Group I to the Fifth Assessment Report of the
Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner,
M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley
(eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York,
NY, USA.

Curry, J. A., Schramm, J. L., Rossow, W. B., & Randall, D. 1996: Overview of Arctic
cloud and radiation characteristics. Journal of Climate, 9, 1731-1764.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S.,
Andrae, U., Balsamada, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.
C. M., van de Berg, L., Bid- lot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes,
M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hölm, E. V., Isaksen,
L., Källberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M.,
Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-
N., & Vitart, F. 2011: The ERA–interim reanalysis: Configuration and
performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–
597

2012 period exceeding climate model projections. Geophysical Research Letters,
39.

Dufour, A., Zolina, O. and Gulev, S.K., 2016: Atmospheric moisture transport to the
Arctic: Assessment of reanalyses and analysis of transport components. Journal of
Climate, 29, 5061-5081.

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P. M., Beljaars, A., Schär, C., & Elder,
K. 2010: An improved snow scheme for the ECMWF land surface model:

Fernandes, R., H. Zhao, X. Wang, J. Key, X. Qu, & A. Hall 2009: Controls on
Northern Hemisphere snow albedo feedback quantified using satellite earth observations, Geophys. Res. Lett., 36, L21702


<table>
<thead>
<tr>
<th>Page</th>
<th>Reference</th>
</tr>
</thead>
</table>


