Statistical Downscaling of Snow Cover Fraction in Svalbard

Jonny Lewis-Brown

October 16, 2023

Abstract

The Svalbard archipelago in the Arctic is one of the fastest warming areas on the planet, with important consequences for birds that migrate to the region. Resultant changes in the timing of the spring snow melt have an impact on the breeding success of populations. To study how these populations may be affected in the future, ecologists require high resolution climate data to perform ecological modelling. However, the output from General Circulation Models is too low resolution (roughly 100 km) to be applied directly to Species Distribution Models. As a result, downscaling routines are required to bridge this gap. In this project, a statistical downscaling method is introduced to produce high-resolution (1 km) snow cover fraction data directly from low-resolution model temperature. The resulting product is validated against satellite snow cover fraction and is found to successfully capture both the timing of the annual snow melt and the spatial patterns of snow cover throughout Svalbard. This method is then applied to projections from the EC-Earth climate model to provide estimates of the snow cover under different Shared Socioeconomic Pathways.

Contents

1 Introduction

As the climate warms, Arctic migratory birds face increasing difficulties in timing their migrations to optimise breeding success (Dickey et al., 2008; Lameris et al., 2018). Due to positive feedback mechanisms, the Arctic is warming faster than mid-latitude regions (Previdi et al., 2021). As a result, the signals that birds use as a reference for the arrival of spring change as they travel northwards, and this affects the migration process (Clausen and Clausen, 2013). Combined with these varying changes in temperature, birds also have to adapt to related issues such as habitat loss (Wauchope et al., 2017) and changes in snow cover.

In order to quantify some of these challenges, the Arctic Migrants Consortium was formed to assess the vulnerability of Arctic migratory birds to rapid climate change. This consortium combines climate and ecological expertise to examine how the fitness and distribution of the birds has been affected by existing climate change, and providing projections and options for mitigation that can be translated into policy. The High Arctic breeding locations studied within the consortium include Nova Zemlya and the Taimyr peninsula in Russia, and the Svalbard archipelago to the north of Scandinavia. The structure of the Work Packages within the consortium is shown in figure 1. The research presented here is being undertaken within the consortium as part of Work Package 1. It focuses on snow cover in Svalbard to provide climatic background data and improve upon future projections.

The distribution and duration of snow cover are amongst the most important features of the climate system in Svalbard, with repercussions for all types of flora and fauna in the area (Winther et al., 2003). For migratory birds, changes in snow cover can have a dramatic

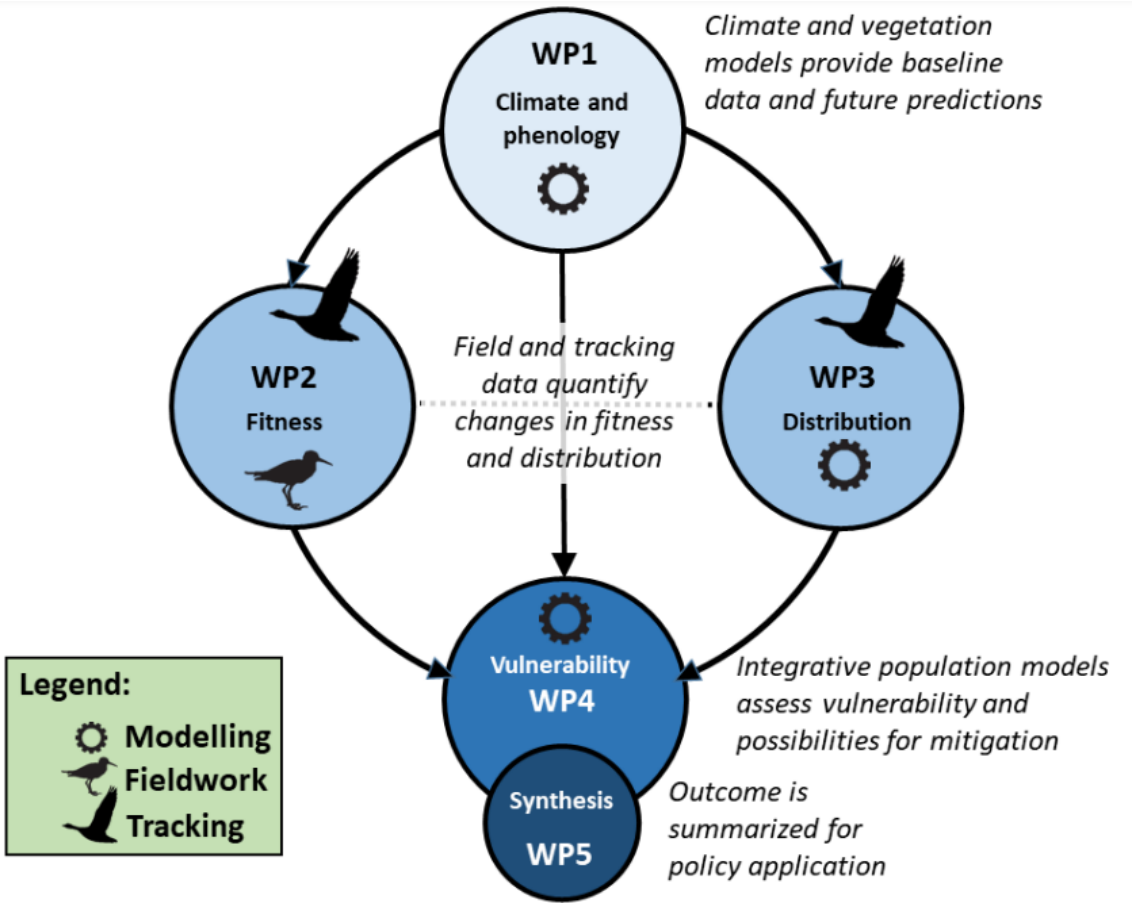


Figure 1: Structure of the Work Packages in the Arctic Migrants Consortium.

impact on breeding success in the region due to its effect on the timing of nesting and the number of birds that are able to nest (Madsen et al., 2007). Therefore bird ecologists desire accurate snow cover data when modelling the populations of migratory birds in Svalbard.

1.1 Modelling

Climatic impacts on bird populations can be studied using Species Distribution Models (SDMs) (Guisan and Zimmermann, 2000). These SDMs combine climate data with field observations to model specific ecological populations and to make future predictions, which can later be used for decisions on policy and conservation (Naimi and Araújo, 2016; Guisan et al., 2013). Temperature and precipitation are the most commonly used variables for these purposes, but in Arctic regions the impact of snow cover is particularly important. It is therefore vital that SDMs have access to the most appropriate and accurate climate projection data, including relating to snow cover, to produce the most reliable models. .

General Circulation Models (GCMs) are the foremost tools available to climate scientists to make predictions about future climate. However, due to data processing constraints even the best GCMs currently have a maximum spatial resolution of approximately 100km (Doblas-Reyes et al., 2021). This is insufficient for use in SDMs, as a much finer spatial resolution ($\leq$ 1km) is required to capture the variability present in most ecological systems (Seo et al., 2009). For example, SDMs can give vastly different outcomes for the inhabitable areas of certain species when using elevation data at high and low spatial resolutions. This is because a coarse grid scale cannot resolve important topographical features (Randin et

al., 2009). In order to bridge this discrepancy between GCMs and SDMs, downscaling procedures are required.

1.2 Downscaling

Downscaling is a method of increasing the spatial resolution of data. Simply interpolating data is not the same as downscaling, as no extra information has been put into the system and many variables (such as precipitation) do not follow the smooth pattern given by interpolation. Downscaling routines must instead add value to the original data (Luca et al., 2015). This can be achieved using one of the two primary methods of downscaling, dynamical or statistical.

1.2.1 Dynamical Downscaling

Dynamical downscaling involves the use of Regional Climate Models (RCMs) to simulate the climate over a specific area or region, using GCM output as boundary conditions. RCMs have the advantage that they can resolve climate behaviour at smaller scales than GCMs, while retaining physical consistency (Xu et al., 2019). However, this comes with a large computational burden and as a result RCMs typically have a maximum resolution of approximately 10-50km. Therefore, even with this increased resolution, RCMs are unable to resolve some of the scales that ecological models require (Doblas-Reyes et al., 2021). Due to both the computational requirements and the need for higher resolutions than can be provided by dynamical downscaling, the focus in this project has only been on statistical approaches.

1.2.2 Statistical Downscaling

In statistical downscaling, statistical relationships are identified between observations of the climate variable of interest at the local scale, called the predictand, and existing data describing the surrounding large scale climate, called the predictor(s) (Maraun and Widmann, 2018). These relationships can then be used to create transfer functions that can applied to climate model data in order to generate high resolution variables (Pielke and Wilby, 2012). This approach has the advantage compared to dynamical downscaling of being more computationally efficient, as well as easier to use (Xu et al., 2019). In addition, in many cases it is possible to produce much higher resolution products, as statistical downscaling methods are not limited by the complexity of a RCM.

However, the success of any statistical downscaling routine is dependent on the fulfilment of several assumptions (Hewitson and Crane, 1996; Hanssen-Bauer et al., 2005). These assumptions are well summarised in Schoof (2013), which is quoted below:

1. The predictor(s) must be adequately simulated by the GCM.
2. The predictor(s) must incorporate the climate change signal.
3. There must be a strong relationship between the predictor(s) and the predictand.
4. The relationship between the predictor(s) and predictand must be time invariant.

The first two requirements depend upon the choice of predictor, whereas the third can be established using existing data. The final assumption is the hardest to verify, as it is only possible to use historical data to create the transfer function. If the length of the record for the predictor(s) and/or predictand is too short, it may not be possible to directly establish whether the relationship is time invariant. In addition, any process that can only take place under climate change induced warmer conditions will not be taken into account by statistical downscaling.

Due to their frequent usage, the most commonly downscaled variables are precipitation and temperature. Temperature variations play an important role in the local distribution of snow cover and as a result some of the principles of temperature downscaling are presented later in section 3.2. In particular, the temperature downscaling method developed by the CHELSA team (Karger et al., 2017) is discussed in more detail. Downscaling of precipitation is not discussed here.

1.3 Snow Cover Fraction

1.3.1 Background

The snow cover fraction variable refers to the proportion of an area that is covered by snow, measured as a percentage. This is related to other snow variables, such as snow depth, but they are usually derived using different methods when making use of remote sensing (Dong and Menzel, 2016). Snow cover fraction maps are established from optical satellites that identify the proportion of snow cover in a pixel based on the reflectance of different visual wavelengths back to the satellite (Nolin, 2010).

Due to the difficulty involved, downscaling of snow cover fraction is less common than temperature or precipitation. One of the biggest problems when developing a downscaling regime is that snow cover has a large spatial heterogeneity (Anderson et al., 2014). Local elevation and orographic effects have an impact on the both the distribution of snowfall and the removal of snow due to wind-blown effects and ablation (Tennant et al., 2017). Local humidity is also important as it affects the proportion of precipitation that falls as snow (Jennings and Molotch, 2019). As a result, despite the desirability of a globally applicable procedure, the majority of studies tend to focus on a single region. Three case studies of previous approaches to snow cover downscaling are summarised below.

1.3.2 Previous Downscaling

First, Matiu and Hanzer (2022) focus on the Austrian Alps. Their method aggregates multiple pixels from a high-resolution satellite snow cover fraction dataset to create a low-resolution training dataset at the same resolution as several RCMs. This training dataset is then compared back against a long time series of the original high-resolution pixels. From this a relationship is formed that assesses the probability of a high-resolution pixel being snow covered given the low-resolution snow cover fraction. This provides an indication of how the fine scale spatial distribution of snow cover could be inferred from a broader dataset. The established relationship is then applied to low-resolution RCM snow cover fraction data to produce high-resolution projections of future snow.

Tryhorn and Degaetano (2013) focus on sets of observation stations in the North-Eastern US. Their downscaling routine uses a modelling software, the Statistical DownScaling Model (Wilby et al., 2002), with a set of different predictor variables such as temperature, wind velocity and surface humidity to directly produce downscaled snow cover fraction. The routine is adapted for different stations to use the most statistically relevant set of predictor variables for each area. This step attempts to ensure against overfitting.

Finally, Gauthier et al. (2022) use a different strategy, focused on the Western US. Instead of using local variables as predictors, they use principal component analysis (Hannachi et al., 2007) to find the large-scale climate modes that have the largest effect on the local snow cover. Leading principal components are selected for both the low-resolution predictor and high-resolution predictand fields. Canonical correlation analysis (Maraun and Widmann, 2018) is then used to yield patterns that maximise the shared correlation between the two fields. This pattern-based strategy fits well with the capabilities of GCMs to reproduce the

climate at a larger scale, allowing for ease of use in future simulations. This approach can also help to identify the primary causes of snow cover variation by comparing the identified patterns with observed climate data fields such as temperature and precipitation.

1.4 Aims

In this thesis, a different method will be presented that focuses on the Svalbard archipelago. A simple statistical relationship between reanalysis temperature and satellite snow cover will be established that directly produces high-resolution snow cover to be used in bird research. This method will then be evaluated against existing snow cover data and applied to produce some initial future projections.

2 Data

2.1 Study region

The study region is the archipelago of Svalbard, approximately spanning the area $74-81\,^{\circ}$N and $10-30\,^{\circ}$E. The largest islands of the archipelago are Spitsbergen, Nordaustlandet and Edgeøya, but there are many smaller islands lying further to the south and east (see figure 2). This region lies well inside the Arctic circle and glaciers cover approximately 60% of the available land mass (Schuler et al., 2020). This corresponds to approximately 10% of the glacier area of the Arctic outside of the Greenland ice sheet. The remainder of the Svalbard surface is mostly covered by snow throughout the winter. In an average year, the snow starts to accumulate towards the end of the Northern Hemispheric summer (September) and the spring melt occurs during the period from late May through June (Spolaor et al., 2016).

The red markers in figure 2 shows the areas where colonies of 3 types of migratory geese (barnacle geese, pink-footed geese and Brent geese) have been identified (data from ©Norwegian Polar Institute, `www.npolar.no`). The most important areas of study tend to be low-lying coastal areas (Mehlum, 1998) and most of the identified colonies in figure 2 are located on the coasts of the main islands or on smaller islands. The birds arrive from early to mid May as the snow is beginning to melt. The main breeding season occurs during June and July and they will have left by the end of September as the snow covering starts to return. The annual snow cycle is therefore of primary importance for migratory birds (Pedersen et al., 2020; Morrissette et al., 2010). Breeding success depends on the early availability of nesting areas and on food abundance during the breeding season. These factors rely heavily on the timing of the snow melt (Madsen et al., 2007; Jensen et al., 2014).

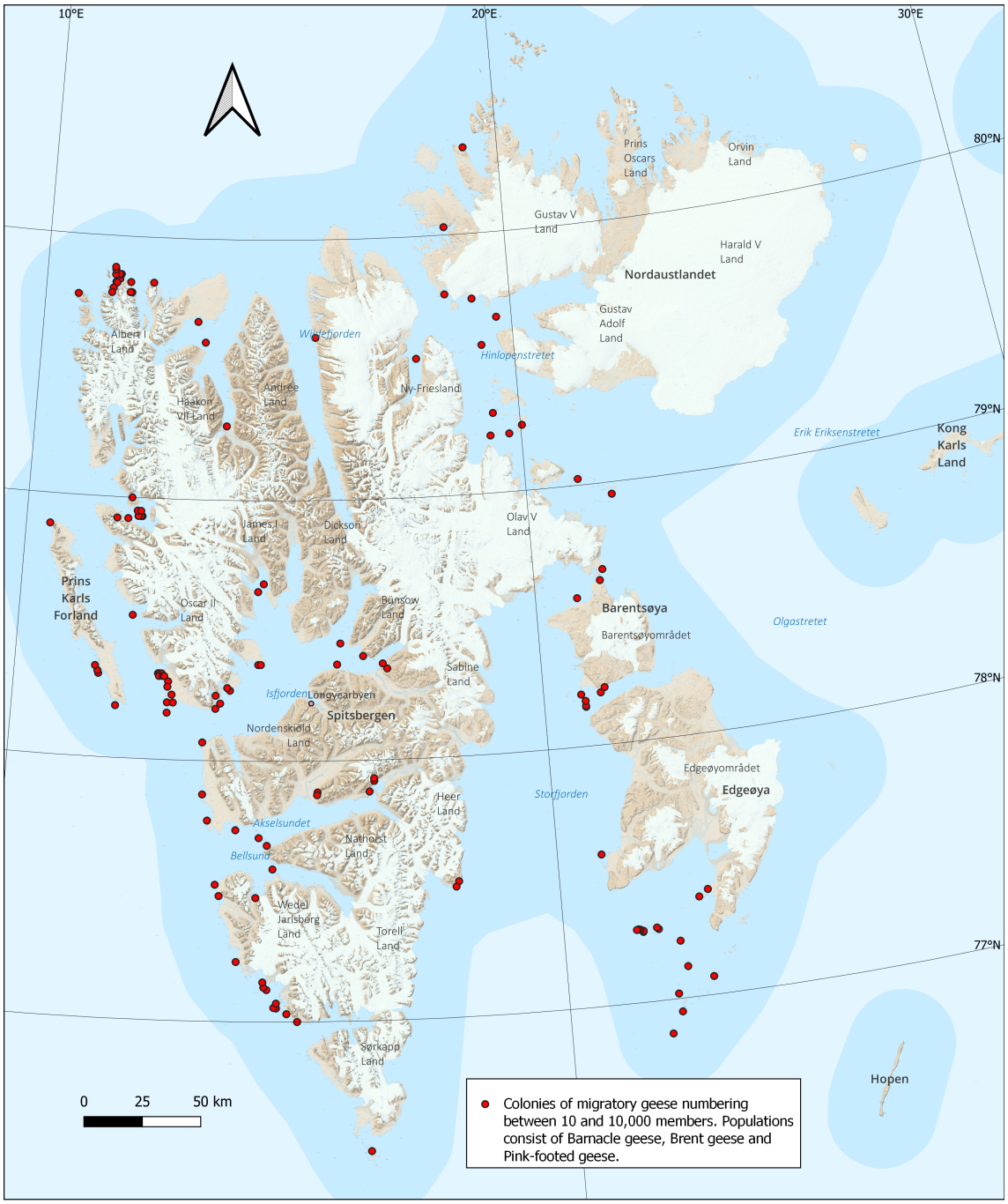


Figure 2: Topographic map of Svalbard. Red dots show the locations of colonies of migratory geese. Base map: Norwegian Polar Institute ©

2.2 Observed Snow Cover Fraction from Satellites

Remote sensing of observed snow cover fraction on the ground (SCFG) was obtained from the MODIS (Moderate Resolution Imaging Spectroradiometer) satellite via the Snow_cci project (ESA Snow climate change initiative). Daily data is available for Svalbard at a spatial resolution of $0.01° \times 0.01°$ for the majority of the period between February 2000 and December 2020 (Nagler et al., 2022). The SCFG product is derived from the viewable snow cover by applying a canopy correction based on the transmissivity of the forest canopy (Hansen et al., 2013). However, trees on Svalbard only reach a few centimetres in height so this effect has no impact on the observations in this region.

Due to significant missing data during the first two years, only the years from 2002 until 2020 have been used. Also, Svalbard is in polar night during the Northern Hemispheric winter, so satellite data is only available for the time period between March and September each year.

The Snow_cci SCFG product uses masks to identify and remove areas such as oceans and glaciers for which the snow cover fraction is not relevant. Water areas are masked in places where 30% or more of a pixel is identified as water, and permanent snow and ice (PSI) areas are correspondingly masked when 50% or more of a pixel has previously been classified as such.

Clouds are also masked via a cloud detection algorithm (Metsämäki et al., 2015). Cloud cover is a frequent issue, consistently obscuring the view from the satellite for large areas of Svalbard. Although routines exist to reconstruct the snow cover underneath the clouds, such as Dong and Menzel (2016), it was decided that this would add extra uncertainty into the downscaling process without a significant gain in efficacy. As such, only the existing snow cover provided by the SCFG product is used.

2.3 Reanalysis Data

Whilst climate variables such as snow cover are available from remote sensing, other important variables such as 2-metre air temperature are not directly detectable by satellites, and so data collected from local weather stations are still the most accurate direct observations. However, these stations are spatially sporadic and can only produce data values at a specific point. This station data therefore cannot account for values over all areas of a grid cell and can be markedly different from the desired average value. To combat this, models and data assimilation techniques can be used to create analyses that incorporate the station observations into a physically consistent system. These analyses depend on the model and techniques used, which can introduce biases due to changes in models over time. Reanalysis products have therefore been introduced which use consistent specific models and data assimilation to create consistent gridded variables for historical climate (Slivinski, 2018). This also allows for longer historical time series from before the start of the satellite era (Madry, 2013).

2.3.1 Temperature

ECMWF (the European Centre for Medium-Range Weather Forecasts) provides ERA5 (ECMWF Reanalysis v5) as a reanalysis product, available from 1940 to the present at a resolution of $0.25° \times 0.25°$ (Hersbach et al., 2020). Hourly air temperature and height data from ERA5 are used on both single levels (Hersbach et al., 2018a) and pressure levels (Hersbach et al., 2018b), provided via the Copernicus Climate Change Service. ERA5 is one of the most widely used reanalyses and has been found to perform well for air

temperature in Arctic regions against other similar products (Isaksen et al., 2022; Graham et al., 2019).

2.3.2 Snow Cover Fraction

To give a comparison to the downscaling procedure for snow cover fraction, multiple reanalysis snow cover datasets were used as base data for a simple interpolation. Firstly, ERA5 provides the variables snow depth and snow density on single levels, from which snow cover fraction can then be derived via a simple formula (Hersbach et al., 2018a). Secondly, ECMWF also provides ERA5-Land, a related land-specific reanalysis at a finer resolution than ERA5 of $0.1° \times 0.1°$ (Muñoz-Sabater et al., 2021). The ERA5-Land dataset only has data on single levels, so cannot be used to create variables such as lapse rate that require multiple levels of temperature. However, snow cover fraction is included in the reanalysis as a variable in itself, rather than as a derived product (Muñoz-Sabater, 2021). Examples of ERA5 and ERA5-Land snow cover are shown in figures 3a and 3b respectively.

2.4 Orography

High resolution orography for the region is obtained from the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) (Danielson et al., 2011). This is initially provided at a resolution of 30 arc seconds, but has been interpolated to $0.01°$ in order to be compatible with the data provided by MODIS. Compared to the low resolution elevation data used by reanalyses or models, higher resolution data such as GMTED2010 captures far more of the local variation in mountainous regions such as Svalbard, as it takes the average elevation over a smaller area.

3 Methods

3.1 Unsuitability of Model/Reanalysis Snow Cover

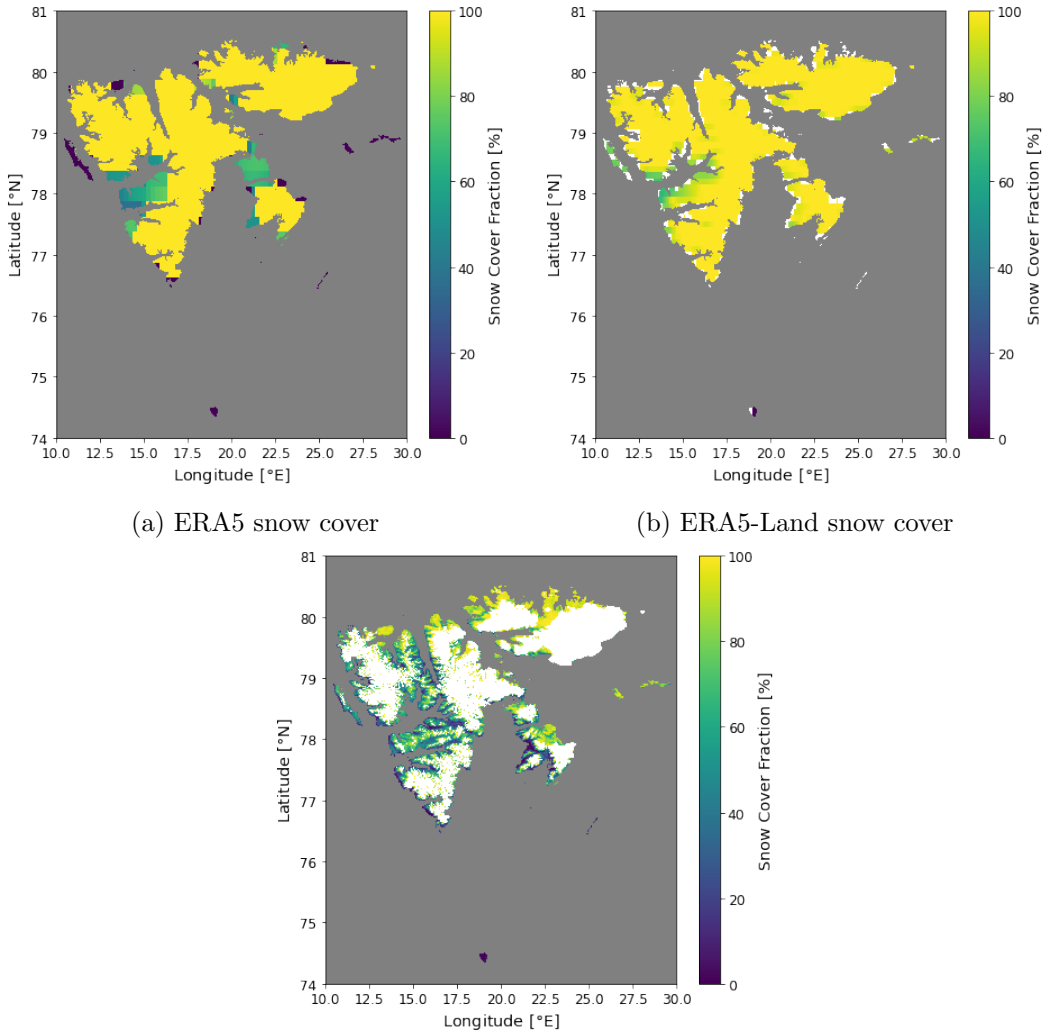
In order to downscale snow cover, it is first necessary to identify one or more variables provided by models that can be used in the process as predictors. Amongst the most obvious predictor variables is the low-resolution model snow cover itself. For this to be used, a relation must be found between the model (or reanalysis) snow cover and MODIS observations, such that future model snow cover can be used as an input variable to produce high resolution snow cover.

However, when examining snow cover from lower-resolution reanalyses there is an immediate problem that areas of permanent snow and ice (PSI), such as glaciers and ice caps, will occupy far more area on a low-resolution map than on an equivalent high-resolution map. As noted in section 2.2, MODIS implements a PSI mask, but the reanalysis models do not, and so in both ERA5 (Hersbach et al., 2018a) and ERA5-Land (Muñoz-Sabater, 2021) PSI areas have a year-round snow cover fraction of 100%. Due to this issue, large areas that are not PSI in MODIS, many of which are coastal and are therefore relevant for migratory birds, do not give reliable snow cover fractions in the reanalysis models. A similar problem occurs specifically in the ERA5 reanalysis near to coastal areas due to overlap with the ocean, and areas with significant snow cover are considered to be ocean pixels and are therefore snow-free.

A comparison between the two reanalyses and the MODIS satellite snow cover fraction is illustrated in figure 3. The figures show the monthly average of the snow cover fraction for June 2011. Because of the lower resolution, both ERA5 (figure 3a) and ERA5-Land (figure 3b) register extremely high snow cover values throughout June in areas where

MODIS (figure 3c) identifies a snow cover of 50% or less. Although ERA5-Land has a better resolution than ERA5, the snow cover persists too long even in areas without permanent snow and ice (see also figure 13 for a time series comparison with ERA5-Land). In fact, only very limited areas of Svalbard snow cover in either reanalysis product bear any resemblance to the MODIS observations.

This means that model snow cover is not a suitable predictor for the high-resolution snow cover in Svalbard. It was therefore decided to abandon the existing reanalysis and future model snow cover entirely for the downscaling procedure. Reanalysis snow cover will only be used as a comparison for the completed downscaled product.



(a) ERA5 snow cover

(b) ERA5-Land snow cover

(c) MODIS satellite snow cover

Figure 3: Illustrative comparison of the average snow cover from two reanalyses and from the MODIS satellite for June 2011. The ocean has been masked in grey. Permanent Snow and Ice is masked in white for MODIS.

3.2 Temperature Downscaling

As the model snow cover is not used in the downscaling procedure, other variables must be identified as predictors. To first order, temperature is a natural choice to use as a predictor, given the strong correlation between temperature and snow cover (see figure 6d). However, before reanalysis or model temperature can be used as a predictor, it must

first be downscaled to the same resolution as the MODIS observations so that a statistical relationship can be established.

There are multiple existing well-tested methods to produce statistically downscaled temperature. Two such methods have been developed by the WorldClim (Fick and Hijmans, 2017) and CHELSA (Karger et al., 2017) groups. WorldClim creates its temperature dataset by directly interpolating between thousands of collected station observations, taking into account various factors such as distance to the coast, elevation and cloud cover. In contrast, CHELSA's dataset takes ERA5 temperature and uses the statistical relationship between local temperature and elevation to re-project the temperature onto a higher resolution orography via the local lapse rate. Both of these datasets are regularly used to provide high-resolution input data for SDMs, with CHELSA-based models often considered to outperform WorldClim-based models for certain tasks (Bobrowski et al., 2021; Datta et al., 2020). For this study, the ease of use of the method provided by CHELSA made it the obvious choice to create downscaled temperature.

A summary of CHELSA's temperature algorithm is provided below, illustrated in figure 4. For more information see Karger et al. (2017) and Karger et al. (2021). The algorithm relies on the fact that elevation is the primary local driver of temperature variation (Rolland, 2003). First, the temperature lapse rate is derived for each pixel from the difference between the ERA5 temperature on the 850 hPa and 950 hPa levels (Hersbach et al., 2018b). This is then combined with the surface temperature to derive the sea level temperature (Hersbach et al., 2018a). Both the sea level temperature and the lapse rate are interpolated to 0.01°, and the resulting high-resolution temperature is projected onto the GMTED2010 orography to create the final surface temperature.

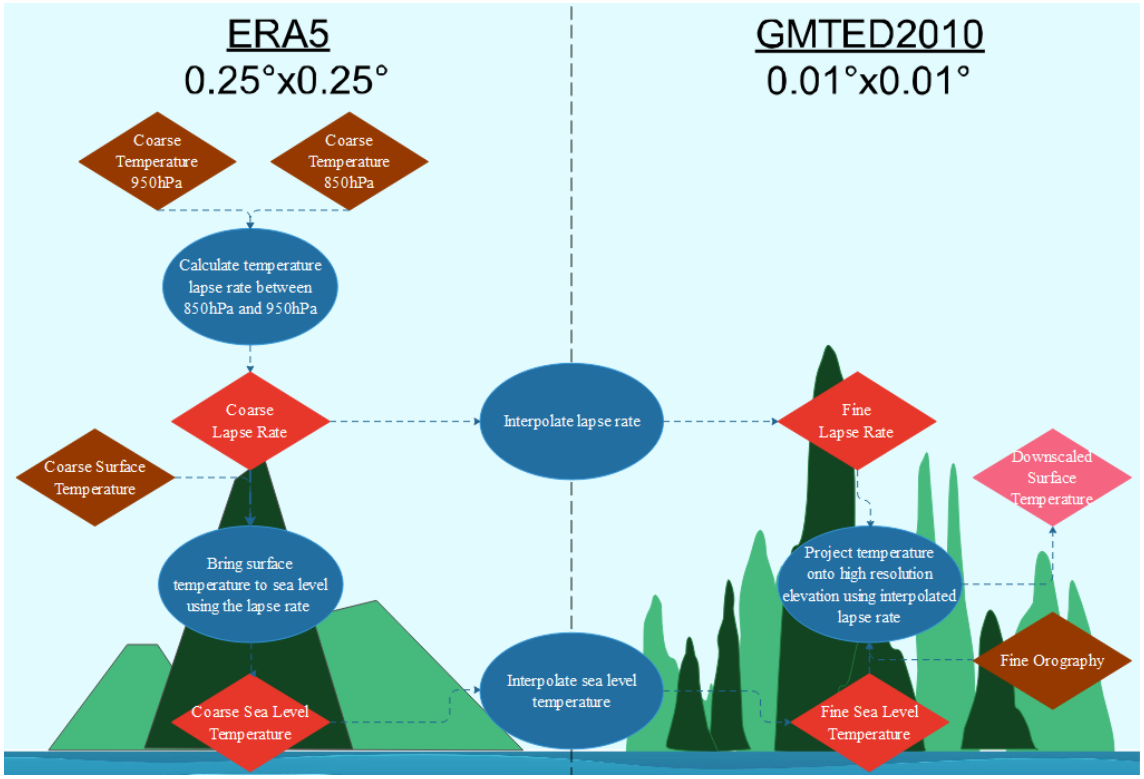
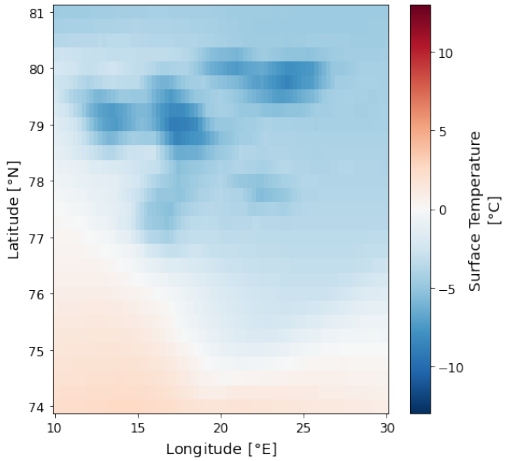


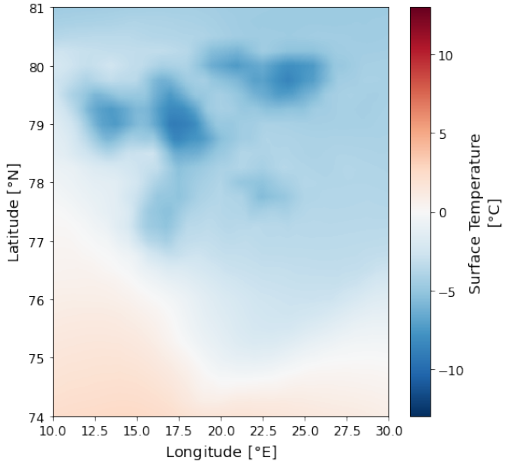
Figure 4: The temperature downscaling algorithm produced by CHELSA. Orange diamonds represent input data, red diamonds represent interim data products and the pink diamond represents the output data. Blue ovals show the processes that take place during the algorithm.

Note that this method can be applied directly to low-resolution model temperature data instead of reanalysis temperature data. However, every model contains biases, so this must be taken into account when interpreting any results derived from such downscaling.

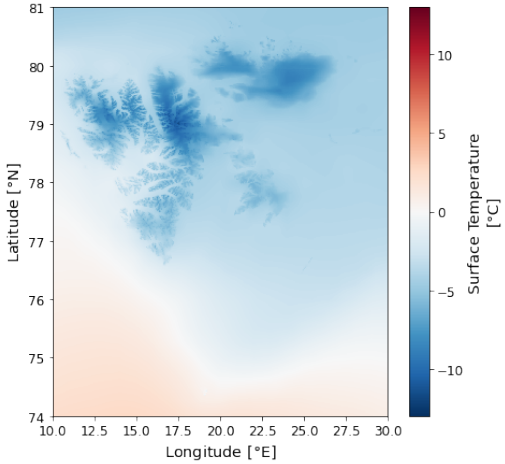
Figure 5 shows the difference between simply interpolating (figure 5b) and downscaling (figure 5c) the temperature on Svalbard in May 2011. It is clear that interpolation has no added value compared to the original ERA5 temperature (figure 5a), whereas the CHELSA downscaling procedure produces a much sharper contrast between the land and sea by including high-resolution orography. Ocean pixels are not affected, whereas areas with large topographic variation have large differences.



(a) Original ERA5 temperature



(b) Interpolated ERA5 temperature



(c) Downscaled ERA5 temperature

Figure 5: Difference between simple interpolation (b) and downscaling (c) of ERA5 temperature on Svalbard in May 2011.

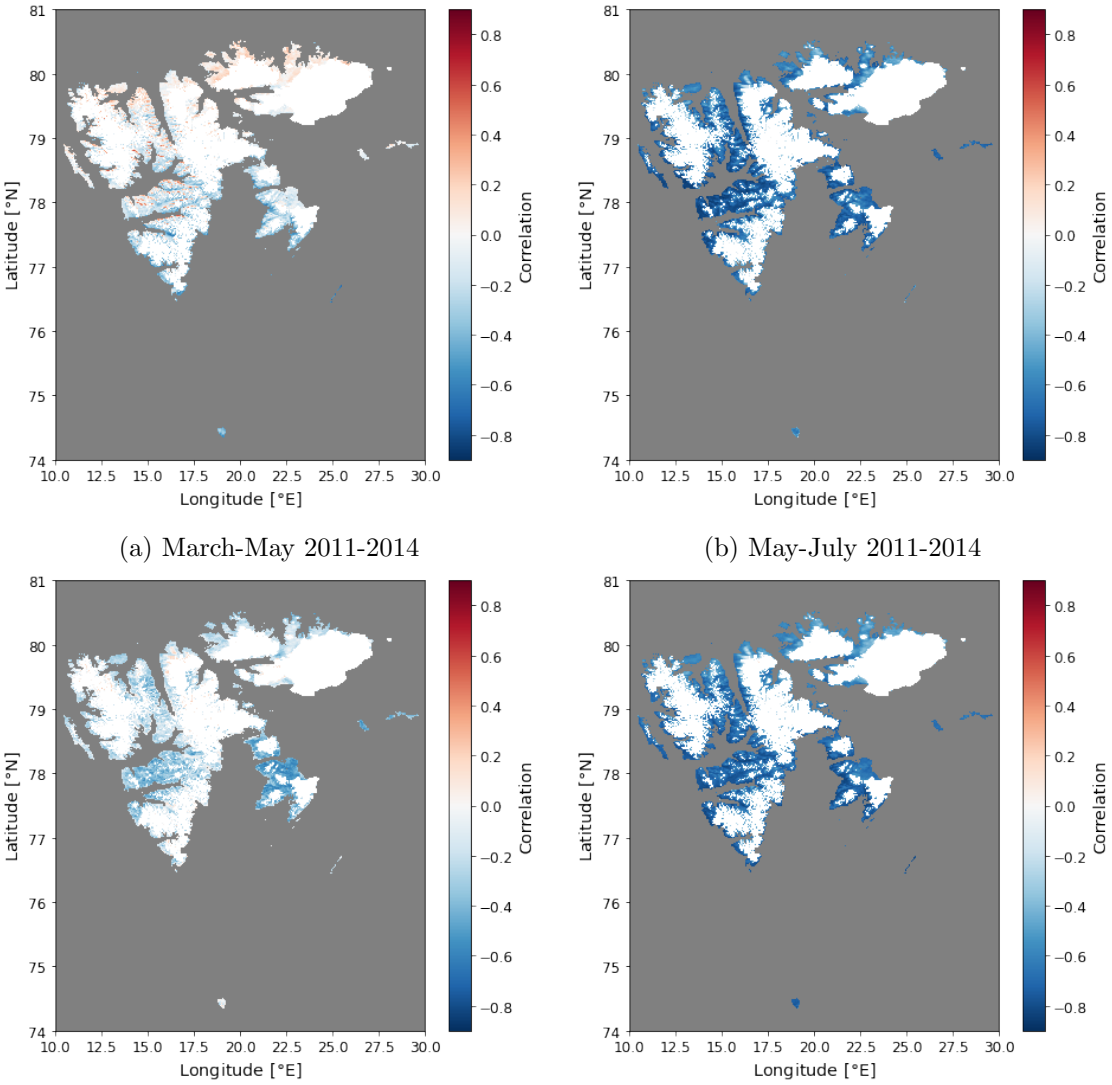
3.3 Snow Cover Downscaling

3.3.1 Temporal Coverage

To use temperature as a predictor in a downscaling regime for snow cover, the next task is to find a statistical relationship between historical temperature and snow cover observations. However, not all times of year in Svalbard are useful in providing data to establish such a relationship. As discussed in section 2.2, MODIS satellite observations of snow cover fraction are only available for Svalbard from March until September each year.

Also, note that during early spring (March and April) the surface is usually completely covered in snow, and in late summer (August and September) the surface is usually mostly bare of snow. Because of this, variations in temperature have less of an effect on the snow cover in the early and late season. Therefore, the relationship between snow cover and temperature is most effectively captured during the transition period of May, June and July (MJJ) from high snow to low snow. This is also the most important period for the arrival and nesting of migratory birds.

The seasonal differences in correlation are illustrated in figure 6, where the temporal correlation coefficient r between downscaled ERA5 temperature and MODIS snow cover has been calculated for each pixel over different time periods. Correlations from the years from 2011-2014 are shown for the period March, April and May (figure 6a, average $r = -0.093$), May, June and July (figure 6b, average $r = -0.631$), July, August and September (figure 6c, average $r = -0.240$) and the whole available year from March to September (figure 6d, average $r = -0.621$). It is clear that the strong negative correlation exhibited in figure 6d over the whole year is primarily due to the contribution from the middle of the MODIS year, whereas the correlation during early spring and late summer are not significant. The relationship between snow cover and temperature will therefore only be calculated based on data from the months May, June and July.



(a) March-May 2011-2014

(b) May-July 2011-2014

(c) July-September 2011-2014

(d) March-September 2011-2014

Figure 6: Maps of the correlation coefficient between snow cover and temperature over the years 2011-2014. Plots (a)-(c) show the correlation for the 3 month periods March-May, May-July and July-September respectively. Plot (d) shows the correlation for the year from March-September.

3.3.2 Algorithm

The first order relationship between snow cover and temperature is shown in the scatter plot in figure 7. This compares the daily spatial median temperature and snow cover fraction in May, June and July 2011-2014. Based on this figure, to first order a successful model would need to incorporate that there is consistently high snow cover at low temperatures ($T \lessapprox -1.5\,^{\circ}C$ and consistently low snow cover at higher temperatures ($T \gtrapprox 3\,^{\circ}C$, with a transition between the two states. It is visually apparent that attempting to simulate this relationship with a simple linear model would result in a significant underestimation of the snow cover at low temperatures and vice versa at high temperatures. Therefore a more complex modelling system is required.

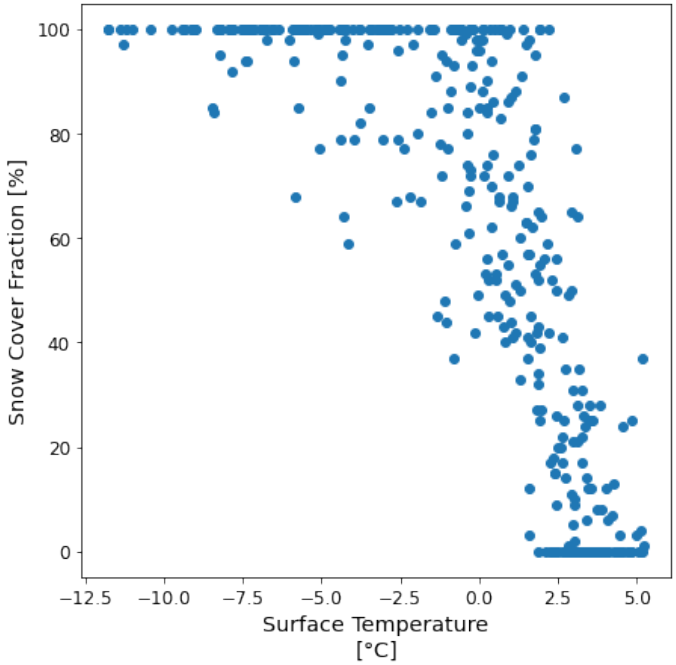


Figure 7: Scatter plot of downscaled ERA5 temperature against MODIS snow cover fraction. Each point represents the median value over Svalbard for a single day during May, June and July 2011-2014.

3.3.2.1 Modelling equation for snow cover

It is important to note here that there are two primary processes that contribute to changes in snow cover fraction - snowfall and snow melt. Dedicated snow models often use separate increase and decrease models to simulate both processes simultaneously. Previous studies (Legates and Willmott, 1990; Rawlins et al., 2006; Krasting et al., 2013) focusing specifically on snowfall have derived an equation that describes an empirical relationship between temperature and the proportion of precipitation that manifests as snow. Equation 1 gives a generalised version of this equation:

$$snc = \frac{100}{1 + a \times b^{T}} \tag{1}$$

Here $0 \leq snc \leq 100$ is the snow cover fraction (%) and T is the surface temperature (°C). In the original research, the parameters $a = 1.61$ and $b = 1.35$ were found empirically in order to fit data for snowfall. See figure 8 later for examples of the shape of the curve produced by this equation with different values of these parameters.

To account for the added effect of snow melt, in this research the assumption is made that the total snow cover fraction has a similar dependency on temperature as snowfall, such that equation 1 can also be used for snow cover. However, the values of the parameters a and b are to be determined by an optimisation process for each pixel individually so that they can vary throughout Svalbard. Changing these parameters across the region implicitly includes local orographic effects on the snow cover inside a grid cell. Examples of these local effects include shadowing of deep valleys by surrounding mountains and impacts due to the proximity of the grid cell to the coast or to nearby glaciers.

3.3.2.2 Temporal averaging

Another factor that affects the relationship between temperature and snow cover is that the snow cover (and in particular snow melt) on a given day is influenced by the temperature over a longer time period than just that given day. Snow melt is often modelled using degree-day methods (He et al., 2014) where the melt is calculated by applying a degree-day factor, measured in $\mathrm{mm\,day^{-1}{}^{\circ}C^{-1}}$, to air temperature data and so decreases in snow cover are caused by melt over a period of time.

As a result, it was decided to use a multi-day average of downscaled ERA5 temperature to compare with a single day of MODIS snow cover. This average acts over the days prior to the day for which the snow cover is to be estimated. Multiple options were tested for the multi-day average and these were evaluated based on the correlation between the resulting downscaled snow cover and the MODIS snow cover for several individual years. Using a 13-day mean of the temperature was found to increase this correlation by the largest amount when compared to only using one day, on average by 0.1 for each year.

3.3.2.3 Optimisation

The optimisation process uses the years between 2002 and 2016 as a training dataset. The years from 2017 to 2020 comprise the test dataset used for evaluation. A non-linear least squares fit is applied between the 13-day mean of the downscaled ERA5 temperature and the daily MODIS snow cover to find the optimal parameters a and b from equation 1 for each pixel. Using data from a large number of years is important to ensure that effects due to internal variability are minimised. A large sample size also reduces the effect of cloud cover as the fitting procedure can only be applied to cloud-free data, which could cause bias if certain months in a year are particularly cloudy.

3.3.3 Parameters

3.3.3.1 Parameter Effects

Once the parameters a and b have been calculated for each pixel, they can be applied directly to the 13-day mean of downscaled temperature to produce a new downscaled snow cover fraction. The parameters affect how the snow cover changes with temperature in different ways. Parameter b determines the range of temperatures over which the bulk of the variation in snow cover occurs. 10% and 90% snow cover are chosen to characterise the bounds of this variation. The range T_R can then be calculated using equation 2:

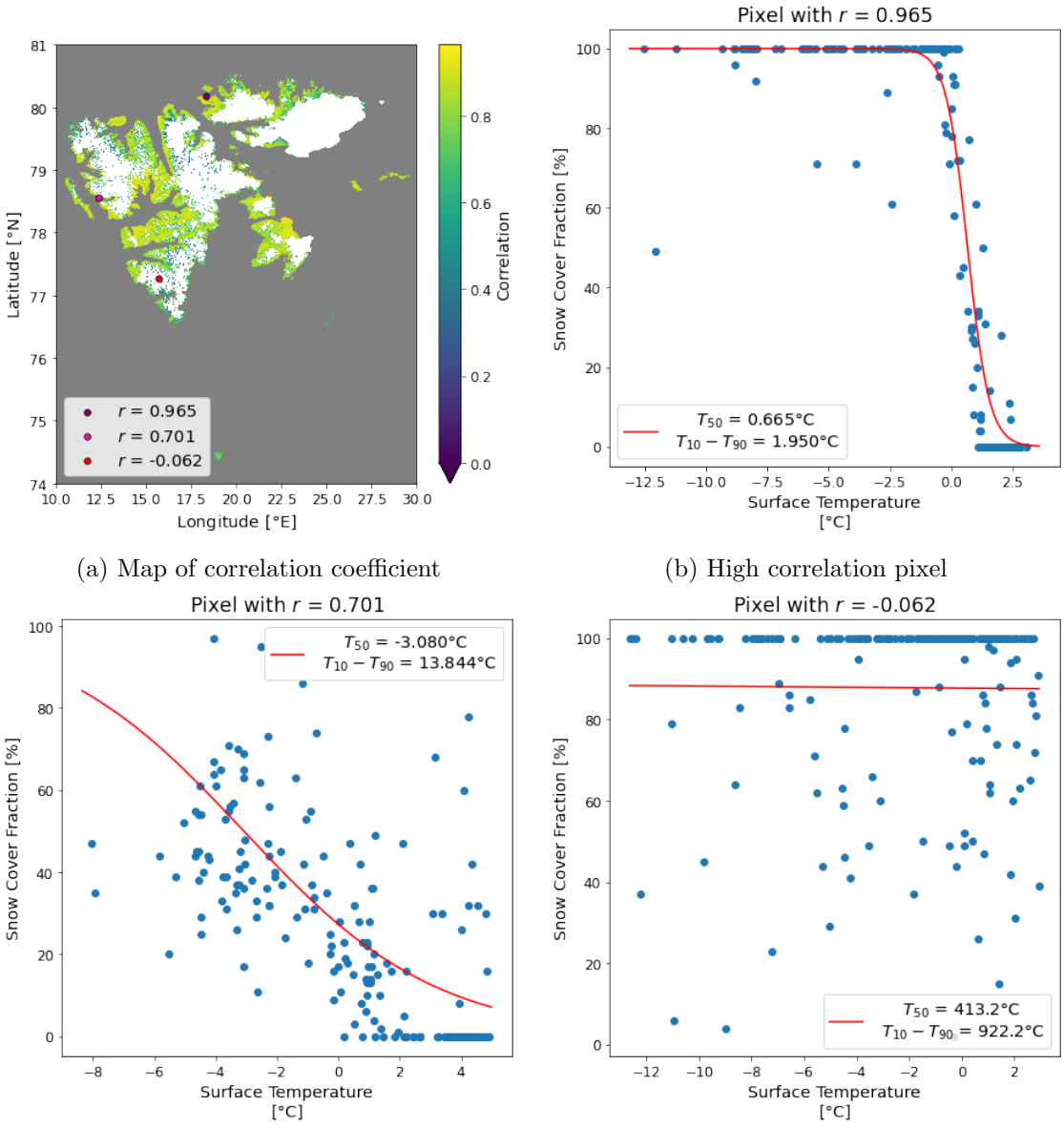
$$T_R = T_{10} - T_{90} = \frac{\log(81)}{\log(b)} \tag{2}$$

Parameter a determines the temperature at which the temperature range given by b is centred, i.e. the temperature for which the snow cover fraction reaches 50%. This median temperature T_{50} can be calculated using equation 3:

$$T_{50} = -\frac{\log(a)}{\log(b)} \tag{3}$$

The fit is anti-symmetric around the inflection point at T_{50}, so the parameters a and b fully describe the fit function.

To illustrate how these parameters change over the region, the new downscaled snow cover is compared with MODIS observations for May, June and July from 2011 to 2014. The temporal correlation coefficient between the two snow cover fraction datasets is then calculated for each pixel over this period and the results are shown in figure 8.



(a) Map of correlation coefficient

(b) High correlation pixel

(c) Mid-range correlation pixel

(d) Low correlation pixel

Figure 8: (a) shows a map of the correlation coefficient r per pixel over Svalbard for the years 2011-2014, with the locations of three specific pixels with different correlations marked. (b), (c) and (d) show scatter plots of MODIS snow cover against 13-day mean temperature and the optimised fit between them for the three pixels. The median temperature T_{50} and temperature range of snow cover changes T_R is noted for each one.

Figure 8a shows how the temporal correlation coefficient (r) changes over the region. The majority of pixels have high values of r, typically greater than 0.8. Low-lying areas near the coasts tend to have higher correlations, whereas pixels near to or within the confines of glaciers have far lower, often negligible correlations. The areas of high correlation correspond well with the aims of the project, given that migratory geese tend to nest near in low-lying coastal areas (Mehlum, 1998) as there is little food close to glaciers. Figure 8a also shows the locations of 3 specific pixels with r values that are representative of the different areas. In particular, the pixel with the highest correlation ($r = 0.965$) is near the coast at a height of 20 m and the pixel with the lowest correlation ($r = -0.062$) is in the middle of a glacier at a height of 765 m.

The remaining parts of figure 8 show scatter plots of MODIS snow cover against 13-day mean temperature over the period 2011-2014, along with the fitted function to demonstrate how the fit works against the data for the 3 representative pixels.

Figure 8b, showing the fit for a pixel with $r = 0.965$, follows an idealised profile, with few outliers. There is consistent 100% snow cover for temperatures $< -1\,^{\circ}\mathrm{C}$ and a sharp transition to 0% snow cover. The temperature range for snow cover changes, T_R, is very small ($< 2\,^{\circ}\mathrm{C}$) and so there is 90% snow at $-0.4\,^{\circ}\mathrm{C}$ and 10% snow at $1.6\,^{\circ}\mathrm{C}$.

Figure 8c, showing the fit for a pixel with $r = 0.701$, has a mid/lower-range correlation and the observations are scattered around the fitted profile. As a result of this the temperature range is much larger than in figure 8b and the median temperature is lower than would usually be seen when predicting snowfall data. However, there is still a clear transition that passes through the bulk of the observations.

Figure 8d, showing the fit for a pixel with $r = -0.062$, has negligible correlation. This is visible in the failure of the fit, which does not form a successful transition; indeed the model predicts that the snow cover would only reach below 50% at $413\,^{\circ}\mathrm{C}$. In the observations, there is mostly consistent 100% snow cover at all available temperatures and low snow cover occurs sporadically at all temperatures, rather than in a transition region. This is likely due to the presence of glaciers close to the pixel, and effects due to wind-blown snow on the higher elevation areas.

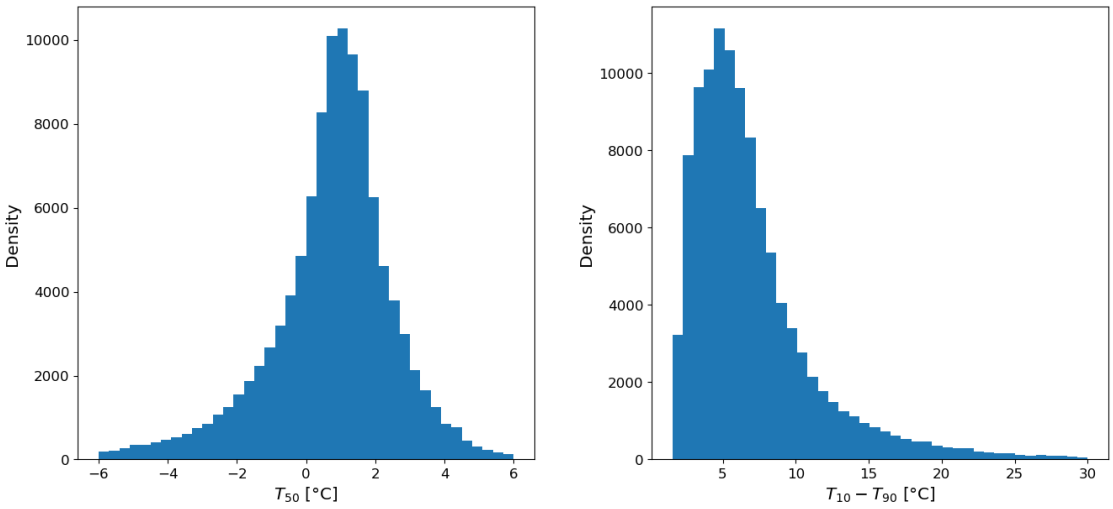
3.3.3.2 Parameter Maps

The following figures show different analyses of how the median temperature T_{50} and the temperature range T_R vary. Individual histograms showing the distributions of T_{50} and T_R are shown in figure 9. Maps showing how T_{50} and T_R change across the region are shown in figure 10.

In addition, in the appendix the added figure A1 shows extra 2-dimensional histograms that examine how the distribution of values of T_{50} and T_R affects the distribution of the correlation between MODIS and downscaled snow cover for each pixel. A1a compares T_{50} and the correlation. Figure A1b compares T_R and the correlation. Figure A1c compares T_{50} and T_R.

Figure 9a shows that within the range of values of T_{50} the majority are concentrated in a small region of temperature between approximately $0\,^{\circ}\mathrm{C}$ and $2.5\,^{\circ}\mathrm{C}$, with a fairly symmetric tail to either side. From figure A1a, it transpires that central values between approximately $-1.5\,^{\circ}\mathrm{C}$ and $2\,^{\circ}\mathrm{C}$ correspond with high correlations between the MODIS and downscaled data ($r \geq 0.8$), .

Ignoring extreme outliers, T_{50} has a mean value of $0.72\,^{\circ}\mathrm{C}$, and a median of $0.93\,^{\circ}\mathrm{C}$. Outliers for T_{50} are considered to be anything above $10\,^{\circ}\mathrm{C}$ or below $-10\,^{\circ}\mathrm{C}$, excluding



(a) Median temperature T_{50}

(b) Range of temperature T_R

Figure 9: Histograms showing the frequency and range of values of the median temperature T_{50} and the range T_R across Svalbard from the downscaling algorithm.

0.5% of the pixels. These average values for T_{50} correspond well with commonly used threshold temperatures for snow melt under degree-day methods (Senese et al., 2014).

Note that a value of T_{50} above 0 °C implies that for the pixel in question the snow cover stays high (above 50%) at temperatures above freezing. This is important from an ecological perspective as these areas will therefore have to be reach warmer temperatures before the snow cover is low enough for nesting to be a possibility for migratory birds.

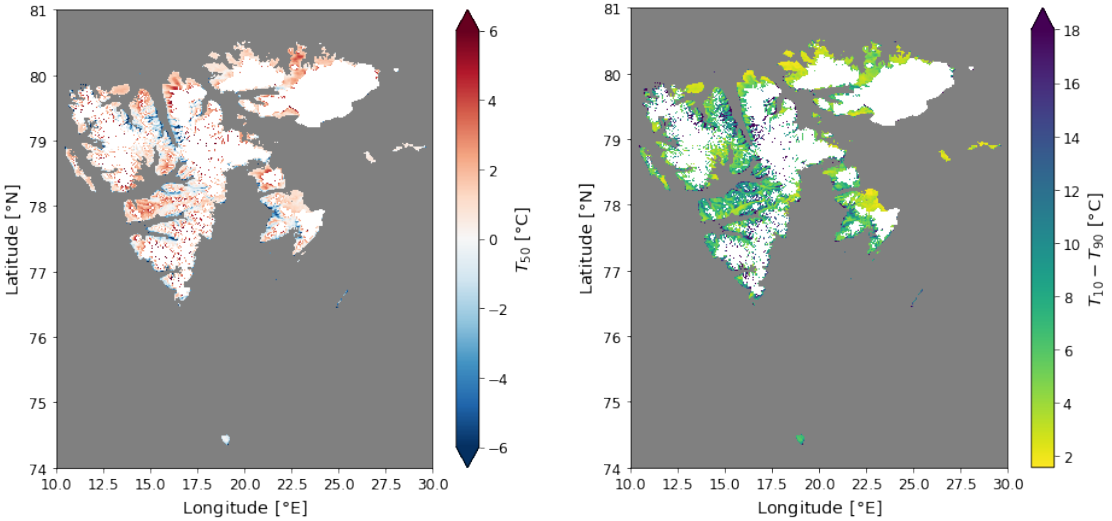
Figure 10a analyses the spatial distribution of T_{50} over Svalbard. Coastal areas either have a slightly positive T_{50} close to the average value, particularly in the north, or a moderately negative T_{50}, particularly in valleys carved out by glaciers. This is likely because runoff and other processes that remove snow mean that snow cover in such areas reduces below 50% at lower temperatures compared to inland areas. Indeed, moving inland the values of T_{50} tend to increase significantly, particularly in areas close to glaciers where there is often snow cover all year round.

Figure 9b shows that the values of the temperature transition range T_R are concentrated in a wide region approximately between 2 °C and 9 °C, with an elongated tail through higher ranges. From A1b it transpires that the sharper transitions in the given region between 2 °C and 9 °C usually indicate good correlations ($r \geq 0.8$), with an increase of T_R causing a gradual decrease in the effectiveness of the fit.

Again ignoring extreme outliers, T_R has a mean value of 7.13 °C, and a median of 5.94 °C. Outliers for T_R are considered to be anything above 40 °C, excluding 0.5% of the pixels.

In the spatial distribution shown in figure 10b T_R varies widely across Svalbard. Low values, indicating sharp transitions, are particularly prevalent on the north coast. High values, indicating slow transitions, are most prevalent inland in glacier dominated areas and also in some of the glacial valleys identified with negative values of T_{50} in figure 10a. Other coastal and inland regions tend to have mid-range values.

Sharp transitions indicate that the local snow cover has a particularly strong dependence on temperature. This suggests that in areas with slow transitions other factors such as wind-blown snow or runoff have a more important role to play than in areas with a sharp transition. This is backed up by the distribution in figure 10b as the flatter areas distant



(a) Median temperature T_{50}

(b) Range of temperature T_R

Figure 10: Maps showing how the median temperature T_{50} and range T_R vary over Svalbard.

from glaciers in Nordaustlandet in the north of Svalbard have sharp transitions whereas areas close to glaciers and in the valleys have slow transitions. The majority of the coastal areas used as breeding grounds for migratory birds have mid-range values of $4 \leq T_R \leq 9$, indicating that temperature is a dominant factor in the snow cover but that other processes have some effect.

In comparison to the original empirical research on snowfall (Legates and Willmott, 1990; Rawlins et al., 2006; Krasting et al., 2013), the values of the parameters $a = 1.61$ and $b = 1.35$ established there yield a value for T_{50} of $-1.59\,^{\circ}\mathrm{C}$ and for T_R of $14.6\,^{\circ}\mathrm{C}$. These are quite different from the mean values of T_{50} and T_R over Svalbard for the algorithm presented here. This suggests that the addition of other processes aside from snowfall has a significant impact on the temperature dependence of the snow cover.

Finally, using both a comparison of figures 10a and 10b and figure A1c shows that low values of T_R correspond to values of T_{50} just above $0\,^{\circ}\mathrm{C}$ and high values of T_R correspond to values of T_{50} well above or below $0\,^{\circ}\mathrm{C}$. In other words, a sharp transition is always centred close to $0\,^{\circ}\mathrm{C}$, but slow transitions vary significantly.

4 Evaluation

In order to evaluate the algorithm, the results must be compared to satellite data from a period that was not used for the training process. In this section, the algorithm is applied to the test temperature dataset from 2017 to 2020, and the resulting downscaled snow cover is compared to the equivalent data from MODIS using monthly averages and time series.

Accompanying statistical data for these four years is shown in table 1. This table contains comparisons of the monthly and 3-monthly means and standard deviations of both the downscaled and MODIS snow cover. It also contains the Root Mean Square Error (RMSE) and correlation coefficient between the two products over each month and over the full three months.

In addition, a metric estimating the earliest nesting opportunity for the migrating birds is

Table 1: Statistical differences between MODIS snow cover and downscaled snow cover for the years 2017-2020. The mean and standard deviation are shown for both snow cover products, as are the RMSE and correlation between the two products. These are available for each month individually and for the total across all 3 months.

Snow Cover Fraction		**Mean (%)**		**Standard Deviation (%)**		**RMSE (%)**	**Correlation**
		MODIS	**DS**	**MODIS**	**DS**		
2017	**May**	86.5	96.8	21.6	8.54	22.4	0.395
	June	64.0	64.7	37.2	29.4	23.6	0.786
	July	19.1	27.8	30.0	26.2	20.9	0.756
	MJJ	54.8	63.1	42.0	36.6	22.2	0.862
2018	**May**	74.7	80.0	27.8	23.5	28.2	0.439
	June	53.1	58.0	37.6	28.9	25.3	0.745
	July	18.3	17.9	29.2	18.9	19.9	0.727
	MJJ	46.9	51.9	39.1	35.3	24.5	0.808
2019	**May**	81.0	93.6	25.0	12.6	26.7	0.36
	June	61.7	61.2	37.6	30.3	24.3	0.766
	July	19.1	25.7	28.5	21.6	20.6	0.719
	MJJ	52.0	60.2	40.3	35.9	23.8	0.821
2020	**May**	71.7	88.7	30.8	18.9	31.7	0.536
	June	31.8	53.0	35.0	26.3	31.8	0.695
	July	10.2	22.9	22.1	21.3	24.7	0.449
	MJJ	38.0	54.9	39.8	35.1	29.5	0.777

also shown for each product - more information on this metric is available in section 4.3.

Similar data restricted to pixels less than 100 m above sea level is shown in table 2. This is intended to represent the areas that migratory birds are most likely to use as nesting grounds.

To provide an alternative evaluation, the algorithm was trained instead on data from 2006-2020, and the resulting parameters were applied to a secondary test dataset of temperature data from May, June and July for the years from 2002-2005. Extra statistical data based on these years are shown in the Appendix in tables A1 and A3. It is possible to use non-consecutive years to train the data but this has not been done here.

4.1 Monthly Maps

Figure 11 contains maps showing monthly averages of both the downscaled and MODIS snow cover fraction, along with the difference between the two products, for May, June and July 2017.

In May the downscaled product predicts that Svalbard essentially has a uniform layer of full snow cover. Differences between the two products are smallest on the north and west coasts of Svalbard because there is maximum snow cover in both, whereas inland and in the south-east the satellite observations record lower snow covers. This discrepancy is verified in the data in table 1, with a mean value of 97% for the downscaled compared to 87% for MODIS. There is also a large decrease in the standard deviation for the downscaled

Table 2: Statistical differences between MODIS snow cover and downscaled snow cover for the years 2017-2020, restricted to data points with an elevation below 100 m. The mean and standard deviation are shown for both snow cover products, as are the RMSE and correlation between the two products. These are available for each month individually and for the total across all 3 months.

Snow Cover Fraction		**Mean (%)**		**Standard Deviation (%)**		**RMSE (%)**	**Correlation**
		MODIS	**DS**	**MODIS**	**DS**		
2017	**May**	84.6	95.1	22.9	11.7	22.8	0.464
	June	52.8	57.3	40.0	33.2	24.1	0.82
	July	11.0	15.5	23.6	18.8	18.0	0.679
	MJJ	47.6	56.0	42.8	39.9	21.4	0.882
2018	**May**	69.3	72.0	30.0	28.1	28.7	0.525
	June	31.7	44.2	34.8	28.8	25.9	0.728
	July	6.75	7.03	17.6	9.3	16.5	0.393
	MJJ	35.2	41.0	38.7	35.8	23.8	0.804
2019	**May**	77.3	91.0	26.5	16.3	26.9	0.479
	June	45.7	50.1	40.1	32.5	25.0	0.781
	July	11.9	17.9	22.9	17.8	19.5	0.595
	MJJ	42.9	53.0	40.4	38.0	23.8	0.834
2020	**May**	71.1	84.5	31.8	23.2	29.9	0.596
	June	24.1	41.3	31.5	25.5	29.8	0.626
	July	7.05	16.2	18.2	18.1	23.1	0.242
	MJJ	34.4	47.4	39.2	36.2	27.6	0.793

product which indicates that the product doesn't capture the spatial variation present in MODIS.

In June the spatial pattern is visually much better than for May. It slightly underestimates the snow cover in the west and overestimates it in the east. This balances so that there is only a 1% difference in the mean over Svalbard. In July the algorithm generally provides an over estimate, particularly in the east, and has a mean of 28% compared to 19% for MODIS.

Overestimation of the snow cover fraction in May is present to some extent in every year of the evaluation (see tables 1 and A1 and the time series in figures 13 and A2). This suggests that snow cover fraction has a smaller temperature dependence earlier in the season. Therefore when May is particularly cold the algorithm assumes that most areas of Svalbard have 100% snow cover. In reality effects that redistribute snow through the local area such as wind-blown snow and leeward shading from mountains can cause a small but significant reduction in the local snow cover percentage. Further discussion is given in section 6.3 of this issue and similar problems with overestimation in July.

Figure 12 looks at the equivalent data for 2018. The snow cover is once again overestimated in May, although not by as much as in 2017, with a mean of 80% for the downscaled product and 75% for MODIS. The spatial pattern is stronger for June with only a slight overestimation in the west and underestimation inland. July 2018 provides the best RMSE for any month with consistently low snow cover everywhere and only minor differences

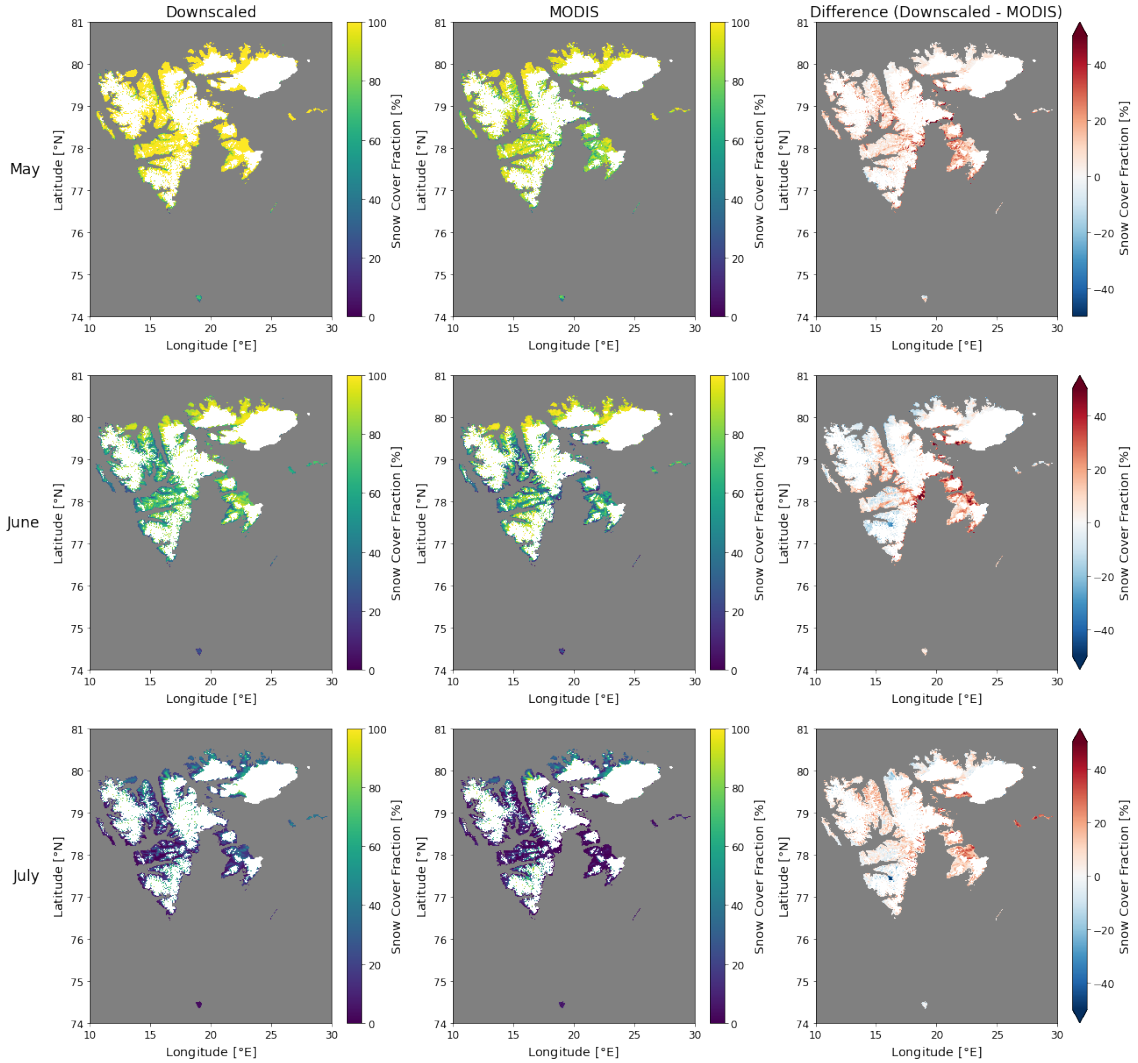


Figure 11: Comparison of monthly averaged downscaled snow cover and MODIS snow cover for May, June and July 2017.

between the two datasets. Indeed the monthly differences are rarely more than 10% for any pixel throughout Svalbard. The mean is 18% for both data sets.

In both 2017 and 2018 the downscaled snow cover manages to capture a sizeable portion of the spatial variation present in MODIS. The standard deviation is always below that given by MODIS. However, that is to be expected when using a smooth function to represent a process with discontinuities such as snow cover fraction's temperature dependence. This shows that the algorithm can perform, albeit with limitations, both in warmer and in colder years.

To simulate the areas in which birds nest, the statistical data was restricted to values below 100 m (see table 2). This gives in general a small improvement to the RMSE and closes the discrepancy between the standard deviations of the two products. The improvement is only slight however so it is not possible to say with confidence that the algorithm performs significantly better under 100 m.

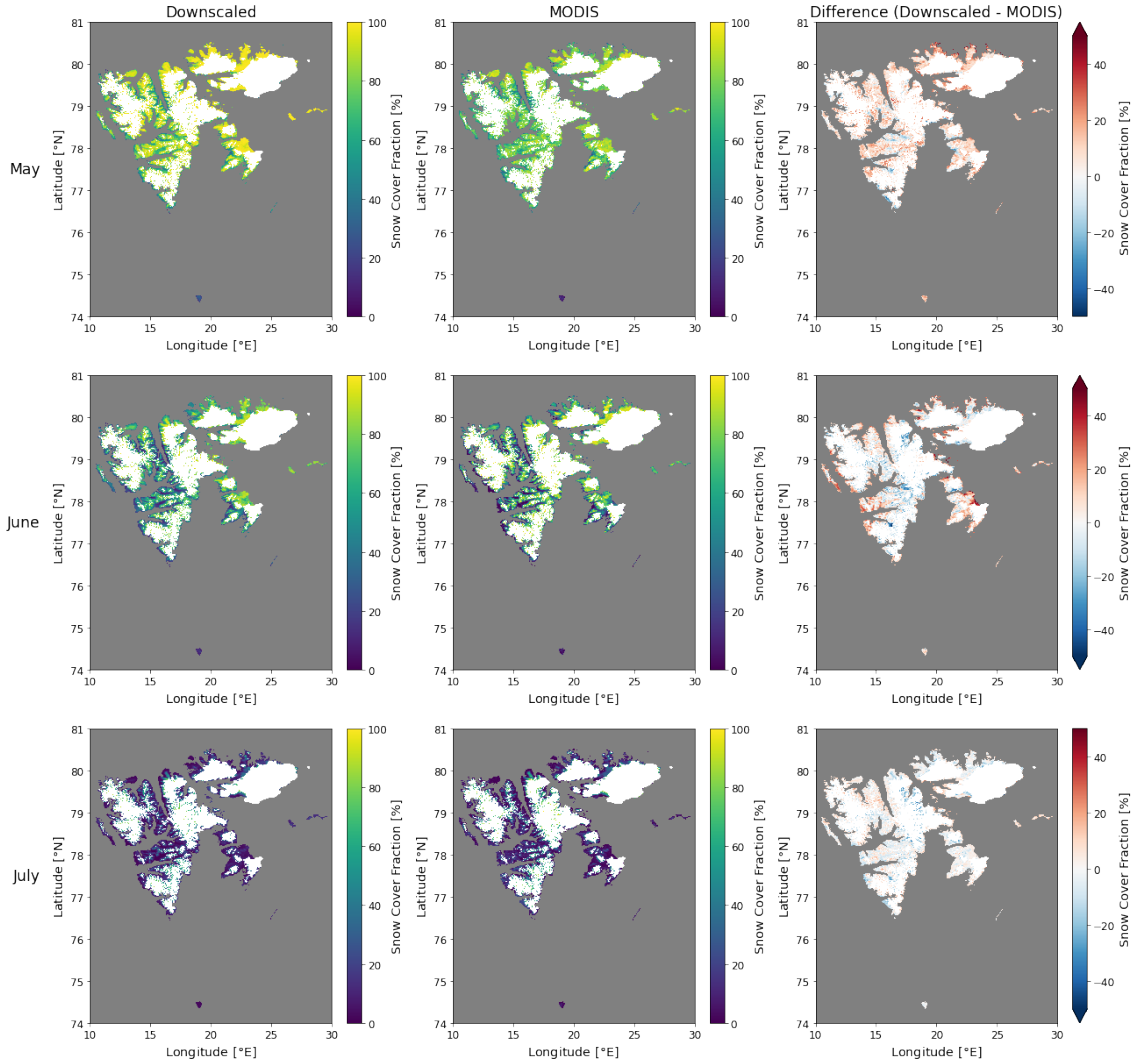


Figure 12: Comparison of monthly averaged downscaled snow cover and MODIS snow cover monthly averages for May, June and July 2018.

4.2 Time Series

Figure 13 contains time series showing how snow cover changes during May-July for the years from 2017-2020. The figures compare the spatial mean of three different snow cover fractions: MODIS (blue), the downscaled product (orange) and ERA5-Land reanalysis snow cover (green). For the ERA5-Land snow cover, pixels that do not change throughout the year (indicating areas of permanent snow and ice) are removed from the time series, to emulate the permanent snow and ice mask present in MODIS and, by extension, the downscaled product. Individual days in the MODIS time series are often affected by cloud cover, so a 7-day average is taken for all datasets to smooth out this variation.

During May 2017 the downscaled snow cover is consistent with the results shown in figure 11, with an over estimation of the snow cover compared to the satellite data during the first month. However, the downscaled snow cover accurately follows the transition from high snow to low snow that the satellite snow cover takes during June, running only a day or two behind. It then over estimates the snow cover again during the early part of July.

In 2018, apart from a high point at the beginning of May the downscaled snow cover follows a similar path to the satellite snow cover through the year.

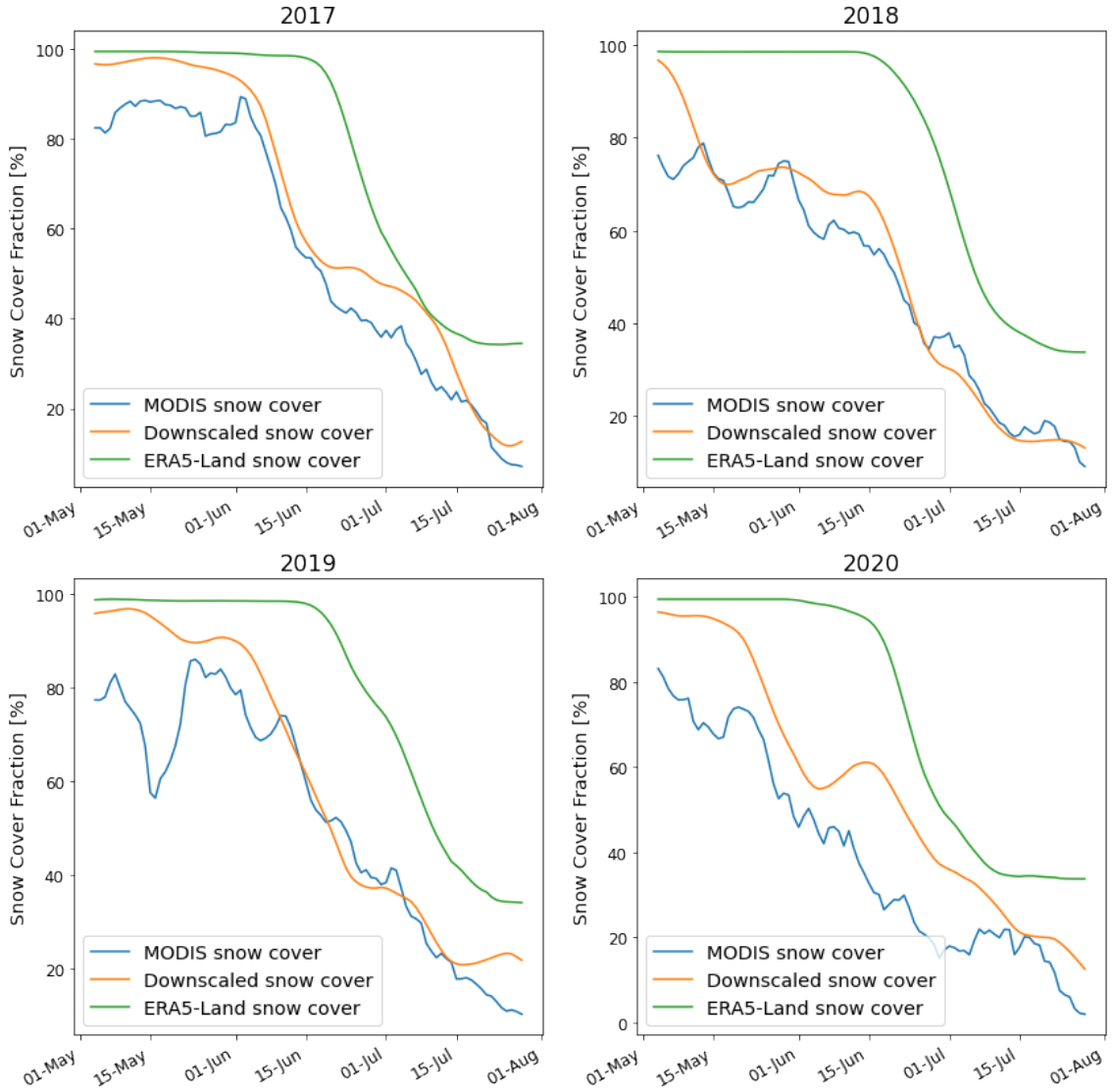


Figure 13: Time series showing how the spatial mean of 3 different snow cover fraction datasets changes over the summer period for the years from 2017-2020. The data is taken from MODIS, the downscaling routine and ERA5-Land. A 7-day mean of each time series is taken to smooth out variability in MODIS due to cloud cover.

At the beginning of the time series for 2019 there is a sudden reduction in the satellite snow cover centred in the middle of May. This is not due to a genuine decrease of snow but is instead caused by several days of persistent cloud cover over the whole of mainland Svalbard, with only the most southerly islands remaining unaffected. These islands tend to have much less snow than the mainland, so the spatial average of the snow cover is significantly reduced when only these areas are visible by satellite. Aside from this blip 2019 follows a similar pattern to 2017. Snow cover is once again overestimated in May but accurately follows the observed transition to low snow cover from June to mid July.

2020 is the only consistently poor year where snow cover is overestimated throughout the season. The reasons for this are unclear. However, it may be that errors are present in the MODIS product, as some individual days of snow cover data during 2020 do not follow the patterns observed in all other years. For further discussion of this issue see section 6.

Aside from 2020, the other three years shown in figure 13 have established that the algo-

rithm is capable of capturing the observed transition from high snow cover to low snow cover. Ecologically this is the most important time period for the algorithm to accurately represent because migratory geese require the snow cover to have dropped to a certain percentage before they can set up their nests. This idea is explored further in section 4.3.

The equivalent time series for the years 2002-2005 are shown in in the appendix in figure A2. The general pattern of results is similar, with the downscaled snow cover overestimating the satellite snow cover early in May and then following the transition relatively well over the next two months. 2004 tracks the transition particularly well.

4.3 First Nesting Day

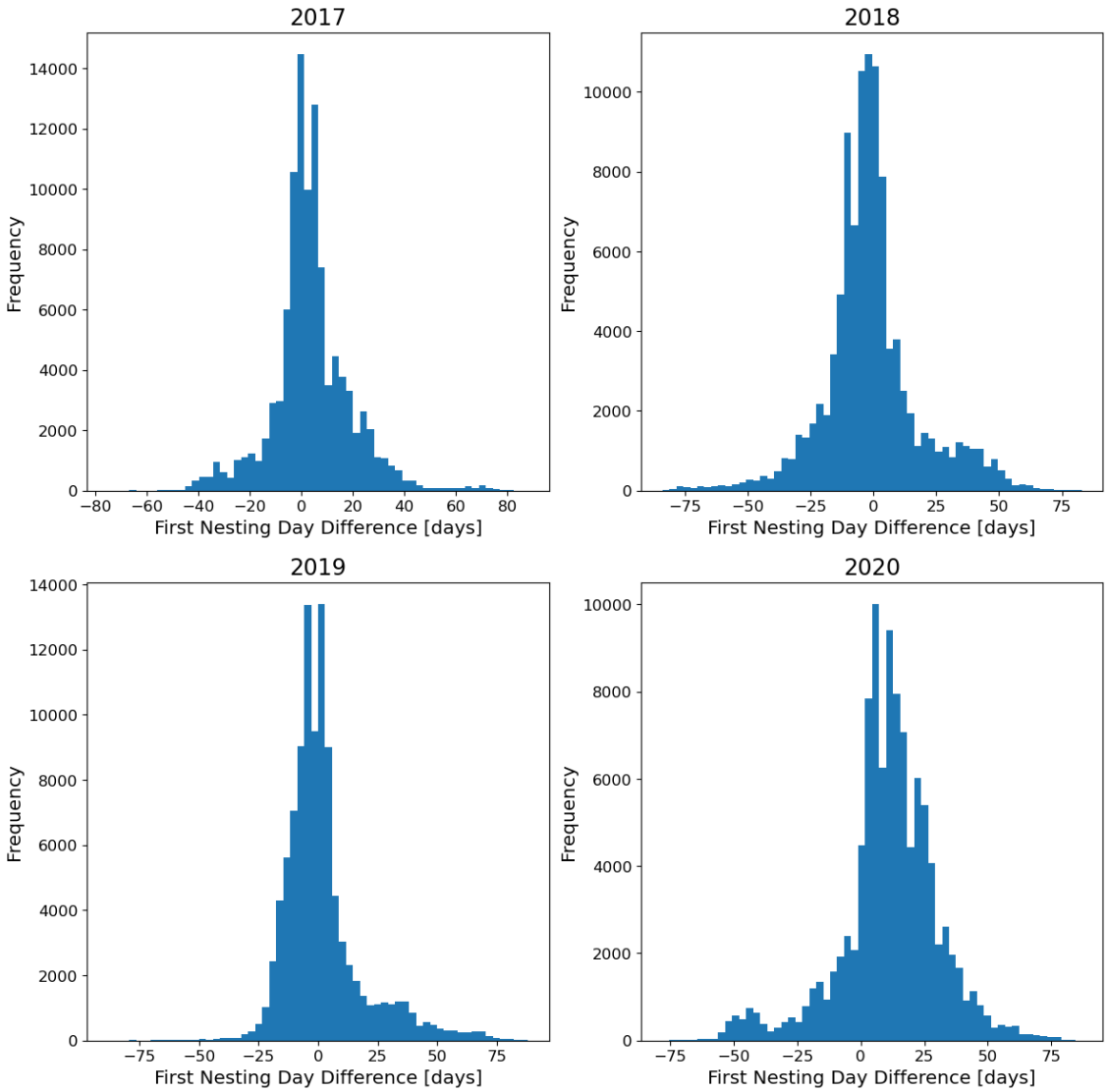


Figure 14: Histogram showing the differences per pixel between the first nesting day metric of the downscaled snow cover and the MODIS snow cover for each year between 2017 and 2020. A positive value means that MODIS predicts an earlier first nesting day than the downscaled product.

An important use of high resolution snow cover is to allow ecologists to make accurate estimations of the first opportunity for migratory birds to nest. One metric used to measure this is to find the first day of the year on which the 7-day mean of snow cover fraction for a pixel falls below 50% (Doiron et al., 2015; Lameris et al., 2018; Lameris

et al., 2019). It is measured by the day of the year, so, given the data restrictions, the earliest possible value is 120 or 121, on May 1st, and the latest is 212 or 213, on July 31st. Statistical data for this metric is shown in table 3.

Table 3: Statistical differences between the predicted first available nesting day across Svalbard as given by both the MODIS snow cover and the downscaled snow cover for the years 2017-2020. The mean of the predicted first nesting day, measured by the date of the year, and the standard deviation across Svalbard, measured in days, are shown for both products. The RMSE between the two products is also shown.

First Nesting Day	**Mean (date of the year)**		**RMSE (days)**	**Standard Deviation (days)**	
	MODIS	**DS**		**MODIS**	**DS**
2017	June 23rd	June 26th	15.6	19.6	16.3
2018	June 16th	June 15th	19.2	23.4	20.3
2019	June 20th	June 21st	16.8	21.8	15.1
2020	June 9th	June 19th	22.4	21.4	19.7

Two figures are shown to illustrate this metric. Firstly, figure 14 shows histograms of the differences between the first nesting day as predicted by MODIS and by the downscaled snow cover. This is done per pixel for each year from 2017-2020. The difference is calculated such that a positive difference means that the downscaled snow cover predicts a later first nesting day than MODIS, indicating that the snow cover stays above 50% for a longer period in the downscaled data.

Secondly, spatial maps comparing this metric across Svalbard for each snow cover product are shown in figure 15. Purple colours show an early first nesting day, whereas yellow show a late first nesting day. On the difference plots, red (positive) colours mean that the downscaled product predicts a later first nesting day than MODIS.

2017 and 2019 were years with similar properties in that they both experienced colder conditions in May compared to other years. In figure 15 for these two years the downscaled product in general predicts a first nesting day that is slightly earlier in the year compared to MODIS in the west of Svalbard but predicts a later first nesting day on the islands in the south east.

In both of these years the majority of pixels had differences of less than 10 days (see figure 14), with a positive skew in the less accurate values. This leads to a low RMSE of under 17 days for each year. For 2017 the average first nesting day is June 23rd for MODIS and June 26th for the downscaled product. For 2019 the average first nesting day is June 20th for MODIS and June 21st for the downscaled product.

2018 was generally a warmer year with an earlier spring. Throughout most of Svalbard the downscaled product predicts a slightly earlier first nesting day than MODIS. However, many individual spots with a much later first nesting day means that the average is within 1 day for each product. For this year there is no clear spatial pattern regarding where the algorithm overestimates the first nesting day.

In 2020, as noted in previous sections, the algorithm consistently overestimates the snow cover as produced by the satellite product. As such the first nesting day is significantly later in the downscaled product over the entirety of Svalbard, with the majority of individual pixels overestimating by between 0 and 25 days. On average MODIS predicts June 9th as the first nesting day and the downscaled product predicts June 19th.

Extra statistical data restricted to points below 100 m is shown in table 4. As expected the lower elevation areas have an earlier decrease in snow cover and as such an earlier first nesting day. There is generally an improvement in the RMSE by 1 or 2 days for most years.

Table 4: Statistical differences between the predicted first available nesting day across Svalbard as given by both the MODIS snow cover and the downscaled snow cover for the years 2017-2020, restricted to data points with an elevation below 100 m. The mean of the predicted first nesting day, measured by the date of the year, and the standard deviation across Svalbard, measured in days, are shown for both products. The RMSE between the two products is also shown.

First Nesting Day	**Mean (date of the year)**		**RMSE (days)**	**Standard Deviation (days)**	
	MODIS	**DS**		**MODIS**	**DS**
2017	June 17th	June 20th	13.9	18.5	14.0
2018	June 6th	June 5th	20.0	20.7	20.7
2019	June 15th	June 15th	15.7	21.5	14.6
2020	June 6th	June 11th	20.4	20.3	19.4

Equivalent data for the years 2002-2005 is shown in the appendix in figures A3 and A4 and tables A2 and A4. The results are similar to those for the primary test dataset, with similar spatial patterns and the majority of pixels having an agreement between MODIS and the downscaled product within 10 days. The mean first nesting days between MODIS and the downscaled product are very close and the RMSE takes values of about 17 days.

Overall these results indicate that the downscaling routine has the capability to reproduce the timing of the snow melt to within 10 days for the majority of pixels over most years. This is likely to be within the range of internal variability of climate models and as such the algorithm can be applied to future projections. Further discussion of the bias and usability of the algorithm can be found in section 6.

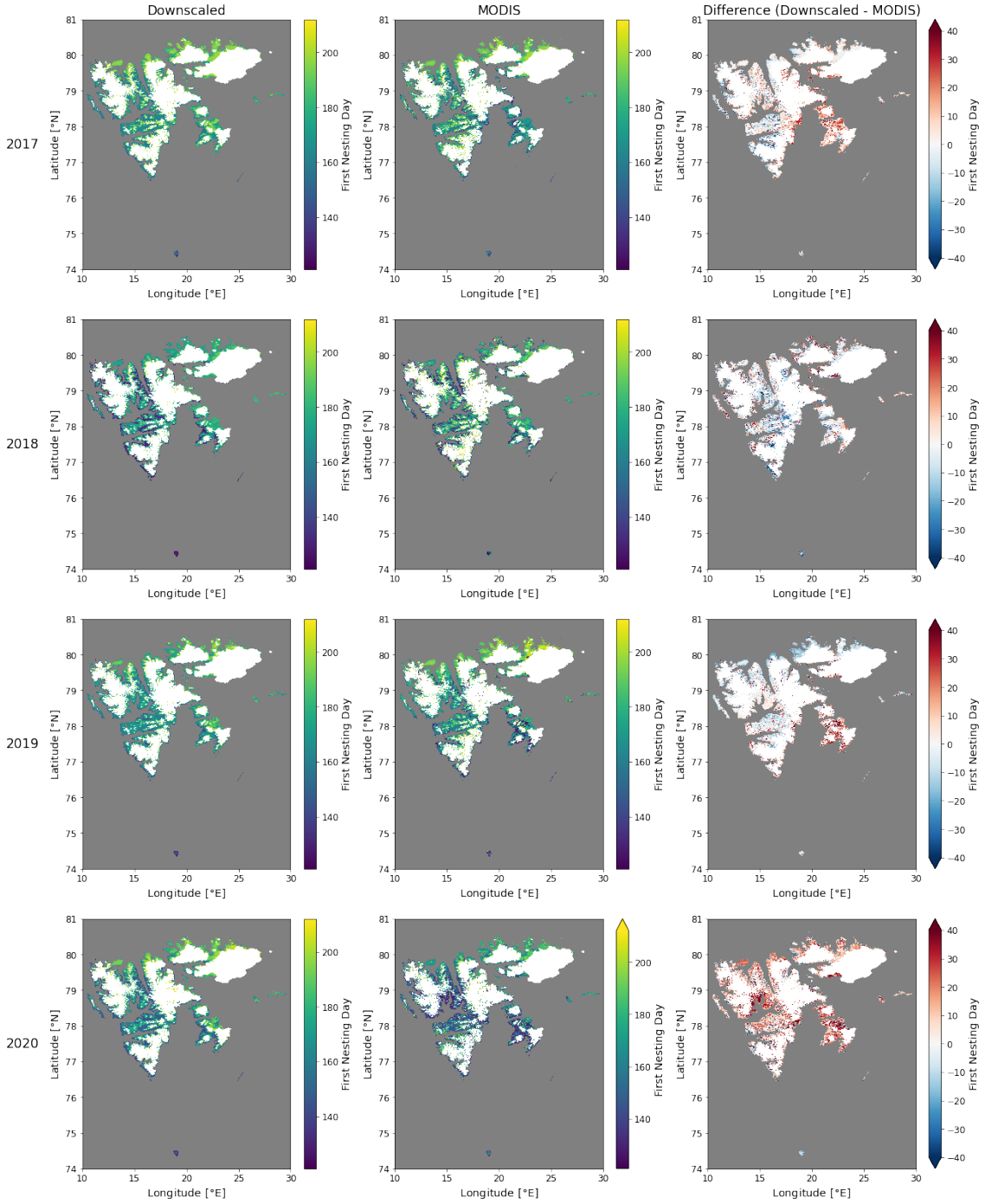


Figure 15: Metric comparing the first available nesting day for each pixel for the downscaled product and the MODIS product, shown for each year from 2017 to 2020. In the comparison column on the right, red colours mean that the downscaled product predicts a later first nesting day than MODIS, and blue means that the downscaled product predicts an earlier first nesting day than MODIS.

5 Application to Future Climate

5.1 Background

A primary reason for the development of this method was to apply it to climate models and examine how migratory birds will be affected by snow cover in the future. To illustrate an example usage, the algorithm is applied to data taken from runs of the EC-Earth3 model. EC-Earth3 is an Earth System Model which is part of the CMIP6 generation of climate models, developed by the EC-Earth consortium. The model has a spatial resolution of 100 km (Döscher et al., 2022).

EC-Earth provides both historical runs and future runs under different Shared Socioeconomic Pathways (SSPs). Historical runs are available for the years from 1850 to 2014 and are forced based on observational climate data as outlined in the CMIP6 protocol (Eyring et al., 2016). Only historical years that overlap with the operation of the MODIS satellite are used in this project.

The SSPs are different pathways of projected socioeconomic developments up to 2100 (Riahi et al., 2017). Each scenario provides plausible narratives of future progress and international politics, accompanied by data such as population growth and GDP. These SSP scenarios are combined with Representative Concentration Pathways (RCPs), which proscribe greenhouse gas concentrations until 2100 (van Vuuren et al., 2011), to provide climate projections through the coming century (Moss et al., 2010). In this project, SSP2-4.5 and SSP5-8.5 are used to examine middle of the road and high emission scenarios respectively.

5.2 Application

To apply the snow cover downscaling algorithm to model data, EC-Earth temperature was first downscaled from 1° to 0.01° resolution. This process was almost identical to that used for ERA5 temperature (see section 3.2). The difference in resolution between ERA5 and EC-Earth has no effect on the downscaling procedure. The only differences in this process compared to that of ERA5 was due to the fact that EC-Earth runs do not contain temperature data on the 950 hPa pressure level, so the difference between the 850 hPa levels and 1000 hPa levels were used instead to estimate the lapse rate.

The values of the parameters a and b used for the snow cover downscaling are those established from the fit between ERA5 temperature and MODIS snow cover (see section 3.3). Applying a new fit between historical EC-Earth temperature and the equivalent MODIS snow cover was considered, as this would help to incorporate any temperature biases in EC-Earth that are different in ERA5. However, given that EC-Earth weather is randomised in historical simulations, rather than accurately following daily weather patterns, it was decided that this would not represent an accurate fit between real snow cover and temperature data.

Three different time periods in the 21st century are compared - historical, mid-century and late-century. For each period an average over six years is used. Because the transition from high snow to low happens at different times each year, the averaged transition is less steep than any individual year. However, when using model projections, it is necessary to average in order to account for internal variability. From the historical modelling the years 2005-2010 are used. For future runs both SSP2-4.5 and SSP5-8.5 are compared for both 2045-2050 and 2095-2100.

5.3 Time Series

The time series in figure 16 show how the spatial mean of the downscaled snow cover fraction for different EC-Earth scenarios changes over the period from May to July. MODIS snow cover over the years from 2005-2010 is shown as a comparison. For each time series, data is averaged over the 6 years by day of the year, and a 7-day rolling mean is taken as in section 4.2.

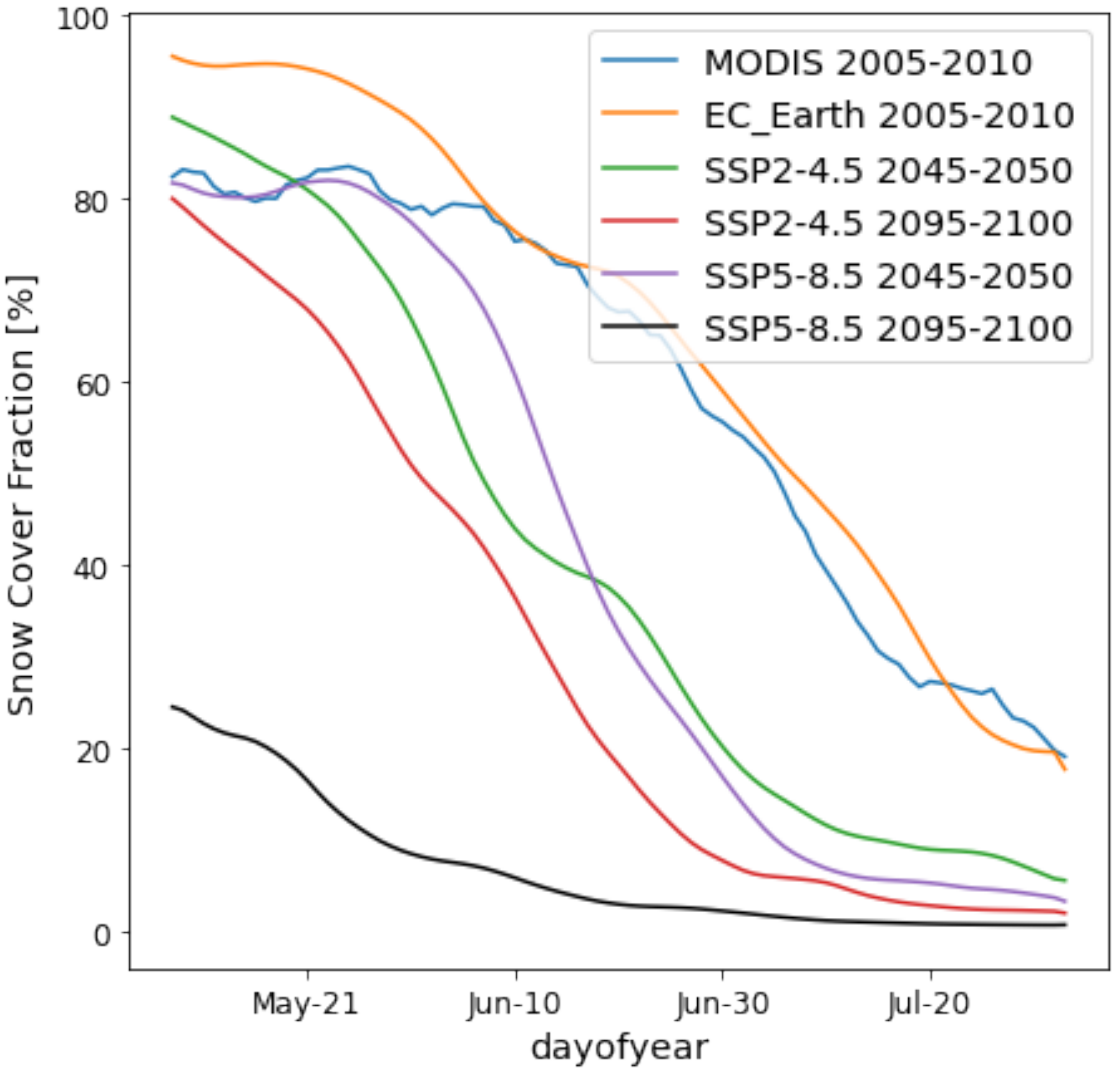


Figure 16: Time series showing how the spatial mean of the downscaled snow cover fraction changes over the summer period for different EC-Earth pathways. The time series included are the MODIS satellite data and downscaled historical EC-Earth runs from 2005-2010, plus the downscaled future EC-Earth projections from both SSP2-4.5 and SSP5-8.5 for the mid-century and the end of the century. A 6-year mean organised by day of the year, followed by a 7-day rolling mean, is taken for each time series.

The EC-Earth historical average (orange) follows the MODIS average (blue) well from the start of June. Similarly to the downscaled ERA5 snow cover, the EC-Earth downscaled snow cover fraction suffers from an overestimation of 10-15% during May. As noted earlier, the transition from high snow cover to low snow cover is steeper in each individual MODIS year compared to the aggregate, shown for example in figure A2.

By the middle of the 21st century from 2045-2050 both of the projections, the 'middle-of-

the-road' SSP2-4.5 (green) and the 'business as usual' SSP5-8.5 (purple), are showing a significant decrease in snow cover compared to the historical modelling. Snow cover has decreased only slightly in May compared to the orange line, but the transition to low snow occurs much earlier in both SSP2-4.5 and SSP5-8.5, starting on average around the end of May, whereas the annual drop in MODIS snow cover occurs on average around the middle of June. The two projections do not differ much between themselves at this stage, with internal variability likely to be the largest factor differentiating the two for any individual year.

By the end of the 21st century from 2095-2100 the two different projections have completely diverged. The SSP2-4.5 projection (red) has decreased by approximately 10% throughout the season compared to the mid-century SSP2-4.5 projection. In contrast, snow cover fraction in the SSP5-8.5 projection (black) is drastically reduced, with less than 30% average snow cover even at the start of May, falling to minimal snow cover as early as mid-June.

5.4 Ecological Impact

The mid-century changes in snow cover under both these scenarios will have large scale effects on birds migrating to Svalbard. Some of these changes can be extrapolated from the effects of climate change that have already been observed. For example, the earlier decrease in snow cover is likely to have some positive impacts as previous years with an early spring melting season have caused a significant increase in the number of birds attempting to nest (Madsen et al., 2007; Anderson et al., 2015). However, an increase in the availability of nesting grounds is unlikely to be the only effect of climate change in the region. Arctic amplification of the climate change signal may result in migrating birds being unable to re-tune their migration timing in order to arrive in Svalbard at the optimal time to capitalise on the high availability of food in the region at the beginning of the spring (Clausen and Clausen, 2013; Lameris et al., 2018). In addition, the warming climate will have separate impacts particularly in regards to marine life, which may have knock-on effects on other parts of the local ecosystem (Descamps et al., 2017). There may also be increased predation due to higher survival rates of regional predators (Layton-Matthews et al., 2020). Species distribution modelling methods are necessary to paint a larger holistic picture of the combined impacts.

By the end of the century it is clear that under the SSP5-8.5 projection temperatures have risen to such an extent that the climate state of Svalbard will have changed completely. This will undoubtedly have drastic impacts on all local and migratory wildlife. Such extreme change makes it difficult to predict exactly what the effect on migratory birds will be. Multiple other species will continue to expand polewards as conditions change in their current regions, resulting in increased competition which will be exacerbated by (Hastings et al., 2020; Lenoir and Svenning, 2015; Virkkala and Lehikoinen, 2014). The birds will also have to adapt to the fact that snow will now have disappeared from their current breeding grounds long before their present day arrival time. However, it may be that other global climate effects will have much larger impacts.

6 Discussion

6.1 Algorithm Success

Using interpolated reanalysis snow cover data has been shown to be extremely ineffective for the small scales required by ecologists. RCM dynamical downscaling is also too

high complexity and is therefore unable to produce data at the desired resolution. Under this reasoning a statistical downscaling approach was constructed without excessive complexity. Despite the simplicity of using only temperature as a predictor, the downscaling method described in this report offers a significant improvement on the original model and reanalysis snow cover data. The statistical analysis in section 4 shows that the downscaling has succeeded in several of the regimes of interest.

Most importantly, it successfully models the transition between high snow and low snow, as shown in figures 13 and A2. For the majority of the individual years, both in the range from 2002-2005 and from 2017-2020, the average snow cover from the downscaling tracks the satellite snow cover to within a few percent during the transition phase. This is reflected in the accuracy of the metric to estimate the first nesting date of the migratory birds. When comparing the satellite and downscaled data, on average the 50% snow cover threshold has been reached by the two datasets within the space of 2 days (see tables 1 and A1).

6.2 Downscaling Assumptions

It is important to establish whether this method conforms to the assumptions required for a successful statistical downscaling (see section 1.2.2).

The first requirement is that the predictor(s) must be well-simulated by the GCM used to make projections. Temperature is the only predictor and so any GCM that can successfully model temperature in the Arctic will satisfy this condition. The EC-Earth model used in this report has been shown to have a cold bias in some areas (Hazeleger et al., 2012) but this can be accounted for when making predictions.

The second requirement is that the predictor must incorporate the climate change signal. Temperature clearly contains this signal and so satisfies this condition.

The third requirement states that there must be a strong relationship between the predictor and predictand. In section 3.3.1, it is established that there is a strongly negative correlation between temperature and snow cover, in particular over the period of snow melt from May to July. Indeed the success of the algorithm in reproducing the satellite snow cover, using only temperature, shows that this requirement is also satisfied.

The final requirement is that the relationship between the predictor and the predictand must be time-invariant. This is the hardest condition to verify and is not an explicit part of the algorithm's design. Indeed, there are some effects of climate change that cannot be taken into account via this method. For example, under glacial retreat in warmer climates new areas would be revealed which are currently masked as PSI in MODIS. These areas would have no data associated with them under this downscaling method. The retreat of glaciers will have further effects on the lapse rate and the pattern of snow accumulation, which are likely to change in areas in the immediate vicinity of the glaciers (Gardner et al., 2009). Lapse rate changes would adversely affect the temperature downscaling methodology, whereas snow accumulation changes would affect the parameters used in comparing local snow cover to temperature. However, areas of Svalbard affected by these issues are likely to be closer to glaciers and therefore unlikely to be the most important areas from an ecological perspective for migratory birds.

6.3 Potential Improvements

Using temperature as the only predictor has resulted in a simple and effective algorithm for producing high-resolution snow cover fraction data. However, there are areas where

the downscaling procedure is inaccurate. Some of the problems and potential solutions are discussed below.

Firstly, in every year assessed in the evaluations the snow cover fraction in May is overestimated. The transition period from high snow cover to low cover centred on June is the most important time period for the nesting of migratory birds, but their initial arrival currently occurs from the beginning of May so it is still important to accurately capture this period of higher snow cover and lower temperatures. One reason for this is that snow cover has a decreased dependence on temperature at the lower end of the temperature range. This means that other factors such as effects of wind-blown snow are more likely to be more important for determining smaller changes in the snow cover. Also, early in the season the snow still has a thick layer so changes in temperature are unlikely to have a major impact on the percentage snow cover, even if the snow depth decreases. This will affect the accuracy of the fitting process. Another minor effect could be that there are fewer data points in the training dataset during May compared to other months because of increased cloud cover.

Although the bias is not as prevalent as it is in May, snow cover in July is also overestimated by the algorithm in several years. This once again suggests that factors other than temperature were more important for reducing the snow cover in some areas. It may also be that the inherent bias in the model equation itself becomes more problematic at the temperature extremes. The fitting process is designed to provide a smooth fit between temperature and snow cover according to equation 1. However, whilst this works well as an approximation it cannot account for some discontinuities in the system. For example, it may be that in certain areas snow melt acts more rapidly above a certain temperature than can be accounted for by equation 1, resulting in a slow decrease in snow cover in the algorithm but a rapid one in reality. This issue is linked to the assumption that snow melt follows the same temperature dependence as snow fall (see section 3.3.2.1).

One option for improving the modelling of snow cover throughout the year is to develop a bias correction routine to systematically reduce the snow cover in areas where the overestimation is most prevalent. Another option would be to adapt the use of equation 1 such that the function follows a separate path at temperature extremes or so that it can be adapted for specific pixels where the fit is poor.

Finally, the lack of inclusion of other relevant climate predictors in the downscaling process could result in drawbacks when making future projections. Variables such as wind speed, humidity and precipitation are also important for snow cover, particularly if large scale patterns change under global warming.

Including other such variables would involve making extensive changes to the downscaling algorithm as it currently stands, which could be difficult to achieve without significantly increasing the complexity. Also, as the current process involves downscaling temperature as the only predictor variable, before being able to use other predictors in the algorithm it may be necessary to downscale them. Precipitation is considerably more difficult to downscale than temperature and any further downscaling introduces extra uncertainties into the system. Despite these challenges, optimising using climatic factors other than temperature would clearly be the next step in the development of this algorithm.

7 Conclusion

The primary aim of this research has been to improve the high-resolution snow cover fraction data in Svalbard available for use by bird ecologists in their SDMs. This has been achieved by means of a downscaling process that establishes a local relationship between

temperature and snow cover. This relationship can then be applied to temperature from climate models to create high-resolution snow cover projections.

The transition between high snow cover and low snow cover is well simulated by the algorithm. As a result the time when local nesting areas become available to migratory birds can be predicted to within a few days. The spatial pattern of the snow cover is usually well simulated. In general the algorithm is most accurate in coastal regions and at lower elevations. These are also the most important areas for use as nesting grounds.

Snow cover is overestimated by the algorithm by approximately 10% on average in May. However, this is early in the season before birds are looking to nest so is of less ecological importance.

The downscaling method has therefore succeeded in its primary task and can be used to offer an improved high-resolution Svalbard dataset for use in future projections.

Acknowledgements

Many thanks to my supervisors Willem Jan van de Berg, Richard Bintanja and in particular Nomikos Skyllas for all of their advice and patience. Thanks also to my family for all of their support.

References

Anderson, Brian T. et al. (2014). "Insights into the physical processes controlling correlations between snow distribution and terrain properties". In: *Water Resources Research* 50 (6), pp. 4545–4563. ISSN: 19447973. DOI: 10.1002/2013WR013714.

Anderson, Helen B. et al. (Feb. 2015). "The dilemma of where to nest: influence of spring snow cover, food proximity and predator abundance on reproductive success of an arctic-breeding migratory herbivore is dependent on nesting habitat choice". In: *Polar Biology* 38 (2), pp. 153–162. ISSN: 14322056. DOI: 10.1007/s00300-014-1574-y.

Bobrowski, Maria, Johannes Weidinger, and Udo Schickhoff (2021). "Is new always better? Frontiers in global climate datasets for modeling treeline species in the himalayas". In: *Atmosphere* 12 (5). ISSN: 20734433. DOI: 10.3390/atmos12050543.

Clausen, Kevin Kuhlmann and Preben Clausen (Nov. 2013). "Earlier Arctic springs cause phenological mismatch in long-distance migrants". In: *Oecologia* 173 (3), pp. 1101–1112. ISSN: 00298549. DOI: 10.1007/s00442-013-2681-0.

Danielson, Jeffrey J, Dean B Gesch, and U S Geological Survey (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. DOI: 10.3133/ofr20111073. URL: http://pubs.er.usgs.gov/publication/ofr20111073.

Datta, Arunava, Oliver Schweiger, and Ingolf Kühn (2020). "Origin of climatic data can determine the transferability of species distribution models". In: *NeoBiota* 59, pp. 61–76. ISSN: 13142488. DOI: 10.3897/NEOBIOTA.59.36299.

Descamps, Sébastien et al. (Feb. 2017). "Climate change impacts on wildlife in a High Arctic archipelago - Svalbard, Norway". In: *Global Change Biology* 23 (2), pp. 490–502. ISSN: 13541013. DOI: 10.1111/gcb.13381.

Dickey, Marie Hélène, Gilles Gauthier, and Marie Christine Cadieux (2008). "Climatic effects on the breeding phenology and reproductive success of an arctic-nesting goose species". In: *Global Change Biology* 14 (9), pp. 1973–1985. ISSN: 13541013. DOI: 10.1111/j.1365-2486.2008.01622.x.

Doblas-Reyes, F J et al. (2021). "Linking Global to Regional Climate Change". In: ed. by V Masson-Delmotte et al. Cambridge University Press, pp. 1363–1512. DOI: 10.1017/9781009157896.012.

Doiron, Madeleine, Gilles Gauthier, and Esther Lévesque (Dec. 2015). "Trophic mismatch and its effects on the growth of young in an Arctic herbivore". In: *Global Change Biology* 21 (12), pp. 4364–4376. ISSN: 13652486. DOI: 10.1111/gcb.13057.

Dong, Chunyu and Lucas Menzel (Dec. 2016). "Producing cloud-free MODIS snow cover products with conditional probability interpolation and meteorological data". In: *Remote Sensing of Environment* 186, pp. 439–451. ISSN: 00344257. DOI: 10.1016/j.rse.2016.09.019.

Döscher, R et al. (2022). "The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6". In: *Geoscientific Model Development* 15 (7), pp. 2973–3020. DOI: 10.5194/gmd-15-2973-2022. URL: https://gmd.copernicus.org/articles/15/2973/2022/.

Eyring, V et al. (2016). "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". In: *Geoscientific Model Development* 9 (5), pp. 1937–1958. DOI: 10.5194/gmd-9-1937-2016. URL: https://gmd.copernicus.org/articles/9/1937/2016/.

Fick, Stephen E. and Robert J. Hijmans (Oct. 2017). "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas". In: *International Journal of Climatology* 37 (12), pp. 4302–4315. ISSN: 10970088. DOI: 10.1002/joc.5086.

Gardner, Alex S. et al. (Aug. 2009). "Near-surface temperature lapse rates over arctic glaciers and their implications for temperature downscaling". In: *Journal of Climate* 22 (16), pp. 4281–4298. ISSN: 08948755. DOI: 10.1175/2009JCLI2845.1.

Gauthier, Nicolas, Kevin J. Anchukaitis, and Bethany Coulthard (June 2022). "Pattern-based downscaling of snowpack variability in the western United States". In: *Climate Dynamics* 58 (11-12), pp. 3225–3241. ISSN: 14320894. DOI: 10.1007/s00382-021-06094-z.

Graham, Robert M., Stephen R. Hudson, and Marion Maturilli (June 2019). "Improved Performance of ERA5 in Arctic Gateway Relative to Four Global Atmospheric Reanalyses". In: *Geophysical Research Letters* 46 (11), pp. 6138–6147. ISSN: 19448007. DOI: 10.1029/2019GL082781.

Guisan, Antoine and Niklaus Zimmermann (2000). "Predictive habitat distribution models in ecology". In: *Ecological Modelling* 135 (2), pp. 147–186. ISSN: 0304-3800. DOI: https://doi.org/10.1016/S0304-3800(00)00354-9. URL: https://www.sciencedirect.com/science/article/pii/S0304380000003549.

Guisan, Antoine et al. (Dec. 2013). "Predicting species distributions for conservation decisions". In: *Ecology Letters* 16 (12), pp. 1424–1435. ISSN: 1461023X. DOI: 10.1111/ele.12189.

Hannachi, Abdel, Ian T Jolliffe, and David B Stephenson (2007). "Empirical orthogonal functions and related techniques in atmospheric science: A review". In: *International Journal of Climatology: A Journal of the Royal Meteorological Society* 27 (9), pp. 1119–1152.

Hansen, M C et al. (2013). "High-Resolution Global Maps of 21st-Century Forest Cover Change". In: *Science* 342 (6160), pp. 850–853. DOI: 10.1126/science.1244693. URL: https://www.science.org/doi/abs/10.1126/science.1244693.

Hanssen-Bauer, I et al. (2005). "Statistical downscaling of climate scenarios over Scandinavia". In: *Climate Research* 29 (3). 10.3354/cr029255, pp. 255–268. URL: https://www.int-res.com/abstracts/cr/v29/n3/p255-268.

Hastings, Reuben A et al. (2020). "Climate Change Drives Poleward Increases and Equatorward Declines in Marine Species". In: *Current Biology* 30 (8), 1572–1577.e2. ISSN:

0960-9822. DOI: https://doi.org/10.1016/j.cub.2020.02.043. URL: https://www.sciencedirect.com/science/article/pii/S0960982220302505.

Hazeleger, W. et al. (Dec. 2012). "EC-Earth V2.2: Description and validation of a new seamless earth system prediction model". In: *Climate Dynamics* 39 (11), pp. 2611–2629. ISSN: 14320894. DOI: 10.1007/s00382-011-1228-5.

He, Z. H. et al. (Dec. 2014). "Estimating degree-day factors from MODIS for snowmelt runoff modeling". In: *Hydrology and Earth System Sciences* 18 (12), pp. 4773–4789. ISSN: 16077938. DOI: 10.5194/hess-18-4773-2014.

Hersbach, Hans et al. (2018a). "ERA5 hourly data on pressure levels from 1959 to present". In: *Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 14-APR-2022)*. DOI: 10.24381/cds.bd0915c6.

– (2018b). "ERA5 hourly data on single levels from 1959 to present". In: *Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 14-APR-2022)*. DOI: 10.24381/cds.adbb2d47.

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis". In: *Quarterly Journal of the Royal Meteorological Society* 146 (730), pp. 1999–2049. DOI: https://doi.org/10.1002/qj.3803. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803.

Hewitson, Bruce C and Robert George Crane (1996). "Climate downscaling: techniques and application". In: *Climate Research* 7 (2), pp. 85–95.

Isaksen, Ketil et al. (Dec. 2022). "Exceptional warming over the Barents area". In: *Scientific Reports* 12 (1). ISSN: 20452322. DOI: 10.1038/s41598-022-13568-5.

Jennings, Keith S. and Noah P. Molotch (Sept. 2019). "The sensitivity of modeled snow accumulation and melt to precipitation phase methods across a climatic gradient". In: *Hydrology and Earth System Sciences* 23 (9), pp. 3765–3786. ISSN: 16077938. DOI: 10.5194/hess-23-3765-2019.

Jensen, Gitte Høj et al. (Jan. 2014). "Snow conditions as an estimator of the breeding output in high-Arctic pink-footed geese Anser brachyrhynchus". In: *Polar Biology* 37 (1), pp. 1–14. ISSN: 0722-4060. DOI: 10.1007/s00300-013-1404-7.

Karger, Dirk N. et al. (Sept. 2017). "Climatologies at high resolution for the earth's land surface areas". In: *Scientific Data* 4. ISSN: 20524463. DOI: 10.1038/sdata.2017.122.

– (2021). *Climatologies at high resolution for the earth's land surface areas*. EnviDat. DOI: http://dx.doi.org/10.16904/envidat.228. URL: https://www.envidat.ch/dataset/chelsa-climatologies.

Krasting, John P. et al. (2013). "Future changes in northern hemisphere snowfall". In: *Journal of Climate* 26 (20), pp. 7813–7828. ISSN: 08948755. DOI: 10.1175/JCLI-D-12-00832.1.

Lameris, Thomas K. et al. (Aug. 2018). "Arctic Geese Tune Migration to a Warming Climate but Still Suffer from a Phenological Mismatch". In: *Current Biology* 28 (15), 2467–2473.e4. ISSN: 09609822. DOI: 10.1016/j.cub.2018.05.077.

Lameris, Thomas K. et al. (Dec. 2019). "Climate warming may affect the optimal timing of reproduction for migratory geese differently in the low and high Arctic". In: *Oecologia* 191 (4), pp. 1003–1014. ISSN: 14321939. DOI: 10.1007/s00442-019-04533-7.

Layton-Matthews, Kate et al. (Feb. 2020). "Contrasting consequences of climate change for migratory geese: Predation, density dependence and carryover effects offset benefits of high-arctic warming". In: *Global Change Biology* 26 (2), pp. 642–657. ISSN: 13652486. DOI: 10.1111/gcb.14773.

Legates, David R and Cort J Willmott (1990). "Mean seasonal and spatial variability in gauge-corrected, global precipitation". In: *International Journal of Climatology* 10 (2), pp. 111–127. DOI: https://doi.org/10.1002/joc.3370100202. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.3370100202.

Lenoir, J and J.-C. Svenning (2015). "Climate-related range shifts – a global multidimensional synthesis and new research directions". In: *Ecography* 38 (1), pp. 15–28. DOI: https://doi.org/10.1111/ecog.00967. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.00967.

Luca, Alejandro Di, Ramón de Elía, and René Laprise (Mar. 2015). "Challenges in the Quest for Added Value of Regional Climate Dynamical Downscaling". In: *Current Climate Change Reports* 1 (1), pp. 10–21. ISSN: 21986061. DOI: 10.1007/s40641-015-0003-9.

Madry, Scott (2013). "Introduction and History of Space Remote Sensing". In: ed. by Scott, Camacho-Lara Sergio Pelton Joseph N., and Madry. Springer New York, pp. 657–666. ISBN: 978-1-4419-7671-0. DOI: 10.1007/978-1-4419-7671-0_37. URL: https://doi.org/10.1007/978-1-4419-7671-0_37.

Madsen, Jesper et al. (Oct. 2007). "Effects of snow cover on the timing and success of reproduction in high-Arctic pink-footed geese Anser brachyrhynchus". In: *Polar Biology* 30 (11), pp. 1363–1372. ISSN: 07224060. DOI: 10.1007/s00300-007-0296-9.

Maraun, Douglas and Martin Widmann (Jan. 2018). *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press. ISBN: 9781107066052. DOI: 10.1017/9781107588783.

Matiu, M and F Hanzer (2022). "Bias adjustment and downscaling of snow cover fraction projections from regional climate models using remote sensing for the European Alps". In: *Hydrology and Earth System Sciences* 26 (12), pp. 3037–3054. DOI: 10.5194/hess-26-3037-2022. URL: https://hess.copernicus.org/articles/26/3037/2022/.

Mehlum, Fridtjof (1998). "Areas in Svalbard important for geese during the pre-breeding breeding and post-breeding periods". In: *Skrifter-Norsk Polarinstitutt*, pp. 41–56.

Metsämäki, Sari et al. (2015). "Introduction to GlobSnow Snow Extent products with considerations for accuracy assessment". In: *Remote Sensing of Environment* 156, pp. 96–108. ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2014.09.018. URL: https://www.sciencedirect.com/science/article/pii/S0034425714003630.

Morrissette, Manon et al. (July 2010). "Climate, trophic interactions, density dependence and carry-over effects on the population productivity of a migratory Arctic herbivorous bird". In: *Oikos* 119 (7), pp. 1181–1191. ISSN: 00301299. DOI: 10.1111/j.1600-0706.2009.18079.x.

Moss, Richard H et al. (2010). "The next generation of scenarios for climate change research and assessment". In: *Nature* 463 (7282), pp. 747–756. ISSN: 1476-4687. DOI: 10.1038/nature08823. URL: https://doi.org/10.1038/nature08823.

Muñoz-Sabater, J. (2021). "ERA5-Land hourly data from 1950 to present". In: *Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 16-JUN-2022)*. DOI: 10.24381/cds.e2161bac.

Muñoz-Sabater, J. et al. (2021). "ERA5-Land: a state-of-the-art global reanalysis dataset for land applications". In: *Earth System Science Data* 13 (9), pp. 4349–4383. DOI: 10.5194/essd-13-4349-2021. URL: https://essd.copernicus.org/articles/13/4349/2021/.

Nagler, T. et al. (Mar. 2022). "ESA Snow Climate Change Initiative (Snow_cci): Daily global Snow Cover Fraction - snow on ground (SCFG) from MODIS (2000-2020), version 2.0." In: *NERC EDS Centre for Environmental Data Analysis*. DOI: doi:10.5285/8847a05eeda646a29da58b42bdf2a87c.

Naimi, Babak and Miguel B. Araújo (Apr. 2016). "Sdm: A reproducible and extensible R platform for species distribution modelling". In: *Ecography* 39 (4), pp. 368–375. ISSN: 16000587. DOI: 10.1111/ecog.01881.

Nolin, Anne W (2010). "Recent advances in remote sensing of seasonal snow". In: *Journal of Glaciology* 56 (200), pp. 1141–1150. DOI: 10.3189/002214311796406077.

Pedersen, Åshild Ønvik et al. (Jan. 2020). "Climate-Ecological Observatory for Arctic Tundra". In: DOI: 10.5281/zenodo.4704475. URL: https://doi.org/10.5281/zenodo.4704475.

Pielke, Roger A. and Robert L. Wilby (2012). "Regional climate downscaling: What's the point?" In: *Eos* 93 (5), pp. 52–53. ISSN: 00963941. DOI: 10.1029/2012EO050008.

Previdi, Michael, Karen L Smith, and Lorenzo M Polvani (Sept. 2021). "Arctic amplification of climate change: a review of underlying mechanisms". In: *Environmental Research Letters* 16 (9), p. 93003. DOI: 10.1088/1748-9326/ac1c29. URL: https://dx.doi.org/10.1088/1748-9326/ac1c29.

Randin, Christophe F. et al. (2009). "Climate change and plant distribution: Local models predict high-elevation persistence". In: *Global Change Biology* 15 (6), pp. 1557–1569. ISSN: 13541013. DOI: 10.1111/j.1365-2486.2008.01766.x.

Rawlins, Michael A. et al. (Apr. 2006). "Evaluation of trends in derived snowfall and rainfall across Eurasia and linkages with discharge to the Arctic Ocean". In: *Geophysical Research Letters* 33 (7). ISSN: 00948276. DOI: 10.1029/2005GL025231.

Riahi, Keywan et al. (2017). "The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview". In: *Global Environmental Change* 42, pp. 153–168. ISSN: 0959-3780. DOI: https://doi.org/10.1016/j.gloenvcha.2016.05.009. URL: https://www.sciencedirect.com/science/article/pii/S0959378016300681.

Rolland, Christian (2003). "Spatial and Seasonal Variations of Air Temperature Lapse Rates in Alpine Regions". In: *Journal of Climate* 16 (7), pp. 1032–1046. DOI: https://doi.org/10.1175/1520-0442(2003)016<1032:SASVOA>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/clim/16/7/1520-0442_2003_016_1032_sasvoa_2.0.co_2.xml.

Schoof, Justin T. (Apr. 2013). "Statistical downscaling in climatology". In: *Geography Compass* 7 (4), pp. 249–265. ISSN: 17498198. DOI: 10.1111/gec3.12036.

Schuler, Thomas V et al. (Jan. 2020). "New data, new techniques and new challenges for updating the state of Svalbard glaciers". In: DOI: 10.5281/zenodo.4704575. URL: https://doi.org/10.5281/zenodo.4704575.

Senese, A. et al. (Oct. 2014). "Using daily air temperature thresholds to evaluate snow melting occurrence and amount on Alpine glaciers by T-index models: The case study of the Forni Glacier (Italy)". In: *Cryosphere* 8 (5), pp. 1921–1933. ISSN: 19940424. DOI: 10.5194/tc-8-1921-2014.

Seo, Changwan et al. (Feb. 2009). "Scale effects in species distribution models: implications for conservation planning under climate change". In: *Biology Letters* 5 (1), pp. 39–43. ISSN: 1744-9561. DOI: 10.1098/rsbl.2008.0476.

Slivinski, Laura C (2018). "Historical Reanalysis: What, How, and Why?" In: *Journal of Advances in Modeling Earth Systems* 10 (8), pp. 1736–1739. DOI: https://doi.org/10.1029/2018MS001434. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001434.

Spolaor, Andrea et al. (Sept. 2016). "Evolution of the Svalbard annual snow layer during the melting phase". In: *Rendiconti Lincei* 27, pp. 147–154. ISSN: 17200776. DOI: 10.1007/s12210-015-0500-8.

Tennant, Christopher J. et al. (Aug. 2017). "Regional sensitivities of seasonal snowpack to elevation, aspect, and vegetation cover in western North America". In: *Water Resources Research* 53 (8), pp. 6908–6926. ISSN: 19447973. DOI: 10.1002/2016WR019374.

Tryhorn, Lee and Art Degaetano (Oct. 2013). "A methodology for statistically downscaling seasonal snow cover characteristics over the Northeastern United States". In: *International Journal of Climatology* 33 (12), pp. 2728–2743. ISSN: 08998418. DOI: 10.1002/joc.3626.

van Vuuren, Detlef P et al. (2011). "The representative concentration pathways: an overview". In: *Climatic Change* 109 (1), p. 5. ISSN: 1573-1480. DOI: 10.1007/s10584-011-0148-z. URL: https://doi.org/10.1007/s10584-011-0148-z.

Virkkala, Raimo and Aleksi Lehikoinen (2014). "Patterns of climate-induced density shifts of species: poleward shifts faster in northern boreal birds than in southern birds". In: *Global Change Biology* 20 (10), pp. 2995–3003. DOI: https://doi.org/10.1111/gcb.12573. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.12573.

Wauchope, Hannah S. et al. (Mar. 2017). "Rapid climate-driven loss of breeding habitat for Arctic migratory birds". In: *Global Change Biology* 23 (3), pp. 1085–1094. ISSN: 13652486. DOI: 10.1111/gcb.13404.

Wilby, Robert L., C W Dawson, and E M Barrow (2002). "sdsm — a decision support tool for the assessment of regional climate change impacts". In: *Environmental Modelling & Software* 17 (2), pp. 145–157. ISSN: 1364-8152. DOI: https://doi.org/10.1016/S1364-8152(01)00060-3. URL: https://www.sciencedirect.com/science/article/pii/S1364815201000603.

Winther, Jan-Gunnar et al. (Jan. 2003). "Snow research in Svalbard–an overview". In: *Polar Research* 22 (2), pp. 125–144. DOI: 10.3402/polar.v22i2.6451. URL: https://polarresearch.net/index.php/polar/article/view/2104.

Xu, Zhongfeng, Ying Han, and Zongliang Yang (Feb. 2019). "Dynamical downscaling of regional climate: A review of methods and limitations". In: *Science China Earth Sciences* 62 (2), pp. 365–375. ISSN: 18691897. DOI: 10.1007/s11430-018-9261-5.

A Appendix

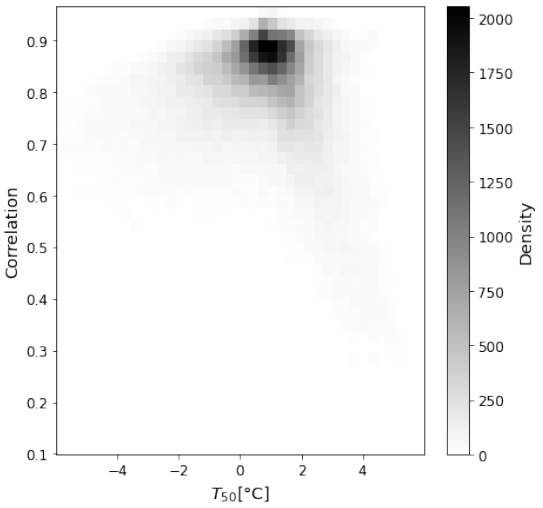
A.1 Tables

Table A1: Statistical differences between MODIS snow cover and downscaled snow cover for the years 2002-2005. The mean and standard deviation are shown for both snow cover products, as are the RMSE and correlation between the two products. These are available for each month individually and for the total across all 3 months.

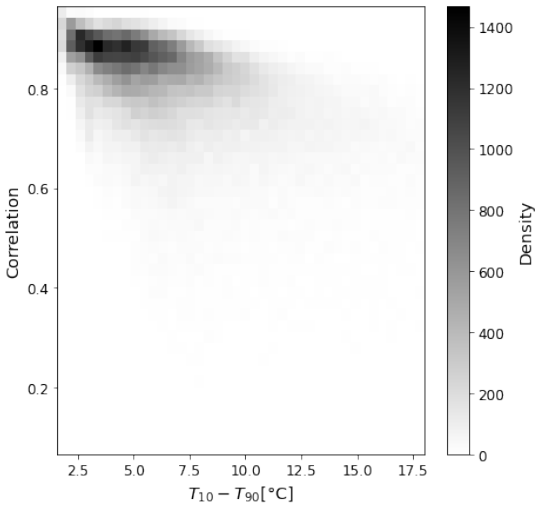
Snow Cover Fraction		**Mean (%)**		**Standard Deviation (%)**		**RMSE (%)**	**Correlation**
		MODIS	**DS**	**MODIS**	**DS**		
2002	**May**	89.5	95.8	18.7	10.3	17.5	0.505
	June	44.8	53.1	35.6	25.9	24.8	0.742
	July	11.4	21.4	21.6	20.3	21.2	0.627
	MJJ	50.9	56.8	41.4	36.5	21.1	0.889
2003	**May**	81.1	92.3	23.8	14.8	26.3	0.35
	June	55.5	65.3	37.0	28.0	26.8	0.707
	July	13.6	23.2	24.0	20.0	23.7	0.521
	MJJ	46.5	60.2	40.2	35.8	25.4	0.824
2004	**May**	81.2	89.2	23.4	17.1	23.4	0.448
	June	66.4	66.8	34.0	27.3	24.7	0.697
	July	12.6	19.8	22.5	19.2	20.7	0.596
	MJJ	46.9	58.5	40.4	36.2	22.6	0.854
2005	**May**	83.3	93.2	22.9	13.4	24.1	0.356
	June	63.8	56.9	35.4	26.6	25.6	0.72
	July	18.4	23.1	28.2	21.9	26.3	0.486
	MJJ	53.2	57.7	39.8	35.8	25.4	0.797

Table A2: Statistical differences between the predicted first available nesting day across Svalbard as given by both the MODIS snow cover and the downscaled snow cover for the years 2002-2005. The mean of the predicted first nesting day, measured by the date of the year, and the standard deviation across Svalbard, measured in days, are shown for both products. The RMSE between the two products is also shown.

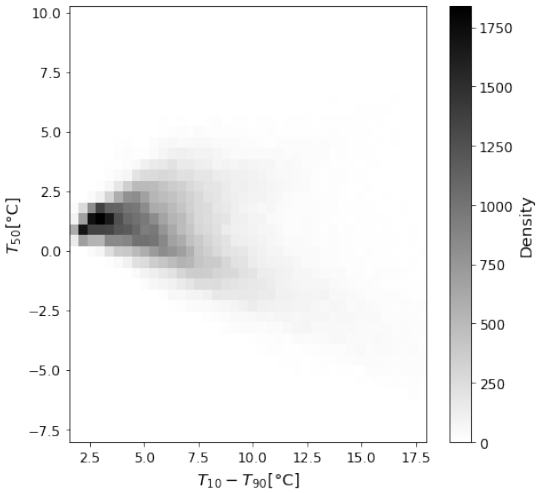
First Nesting Day	**Mean (date of the year)**		**RMSE (days)**	**Standard Deviation (days)**	
	MODIS	**DS**		**MODIS**	**DS**
2002	June 17th	June 18th	15.4	17.2	15.6
2003	June 20th	June 21st	17.6	20.6	14.8
2004	June 22nd	June 22nd	16.6	19.3	16.3
2005	June 23rd	June 20th	18.7	19.3	16.3



(a) Correlation against median temperature



(b) Correlation against temperature range



(c) Median temperature against temperature range

Figure A1: 2-dimensional histograms showing how the median temperature T_{50}, the temperature range T_R and the correlation between the downscaled snow cover and the MODIS snow cover r affect each other. Figure A1a shows the joint distribution of r and T_{50}, figure A1b shows the joint distribution of r and T_R and figure A1c shows the joint distribution of T_{50} and T_R.

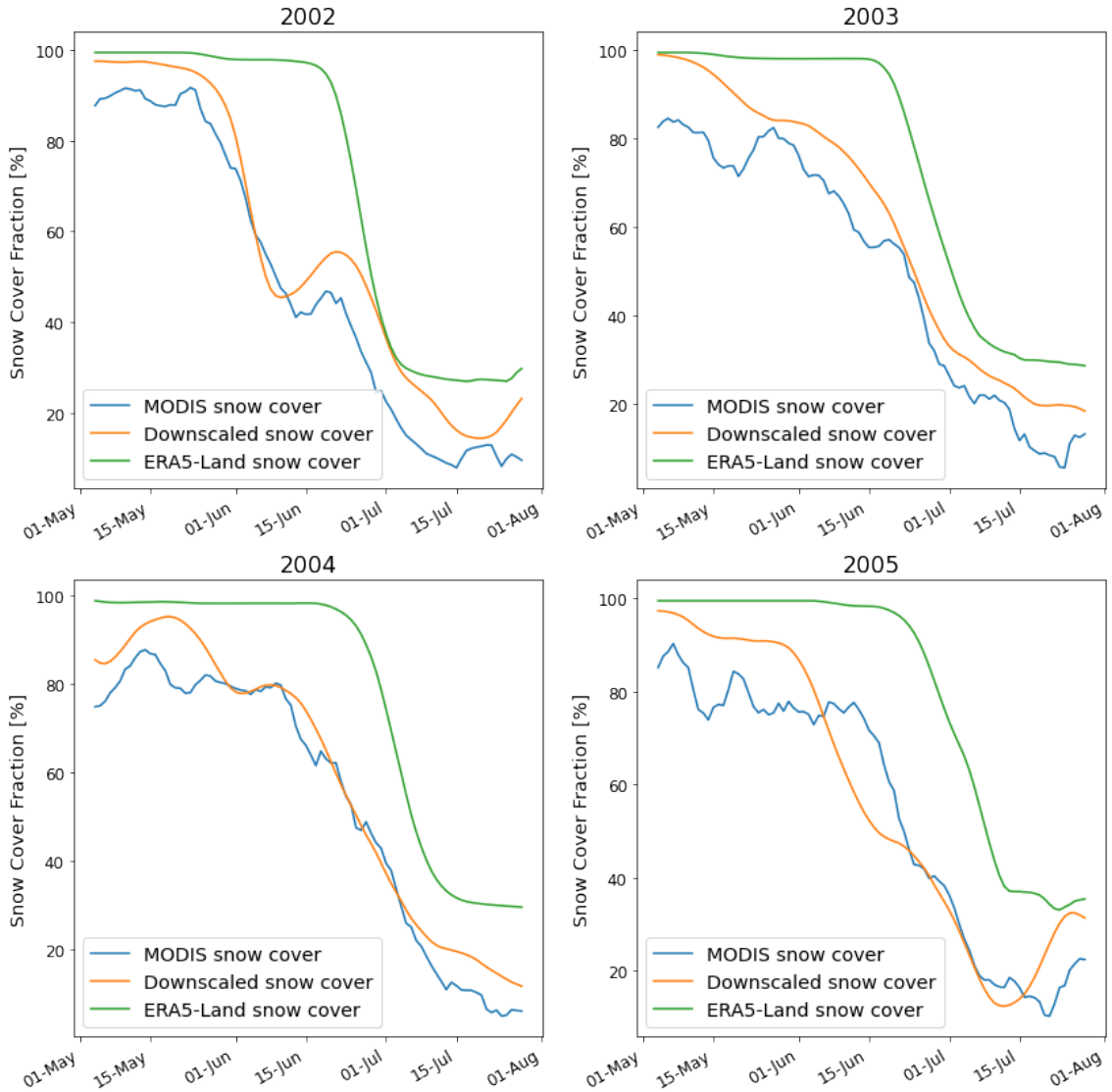


Figure A2: Time series showing how the spatial mean of 3 different snow cover fraction datasets changes over the summer period for the years from 2002-2005. The data is taken from MODIS, the downscaling routine and ERA5-Land. A 7-day mean of each time series is taken to smooth out variability in MODIS due to cloud cover.

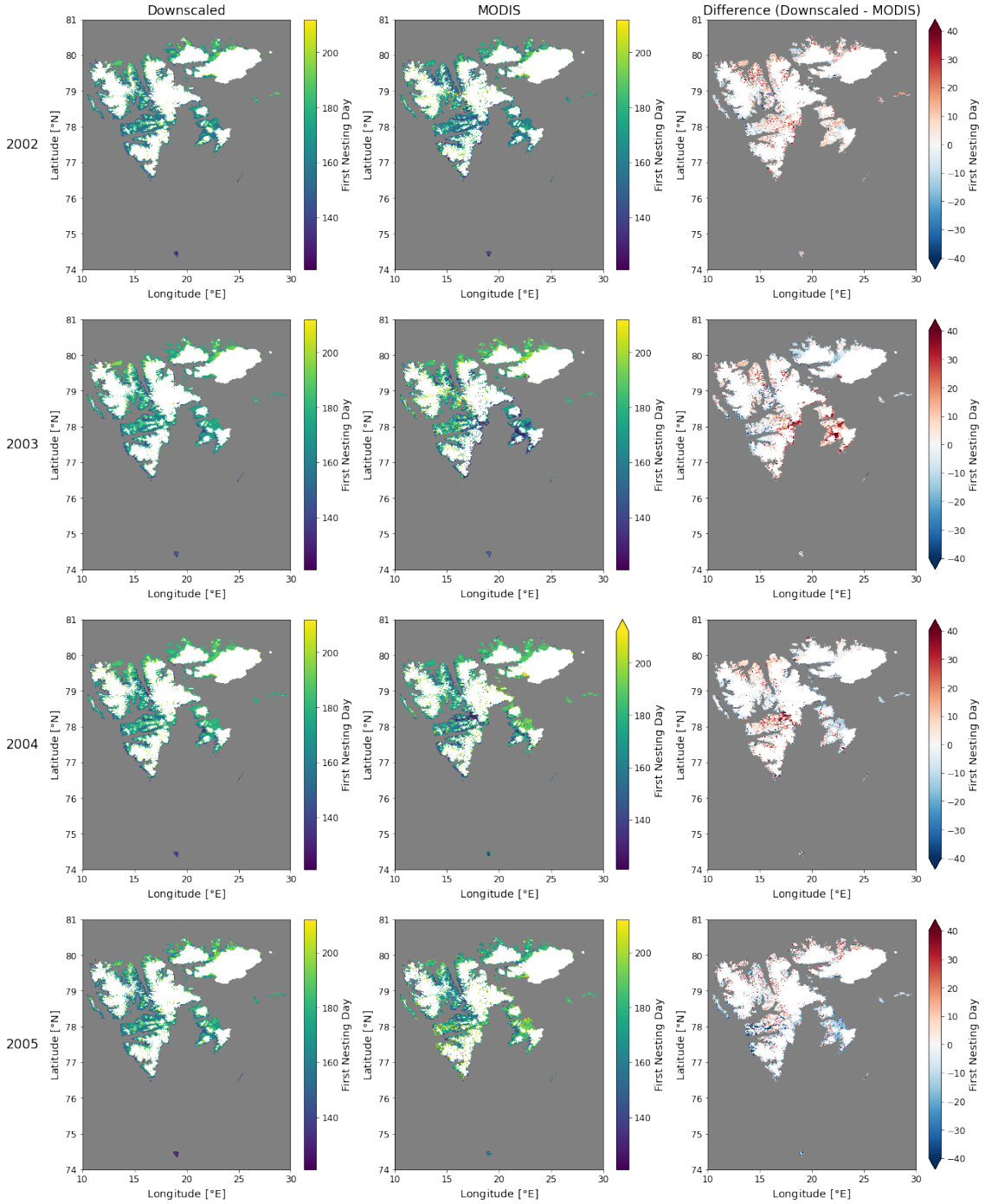


Figure A3: Metric comparing the first available nesting day, measured by day of the year. This is done for each pixel for the downscaled product, the MODIS product and the difference between the two, shown for each year from 2002 to 2005.

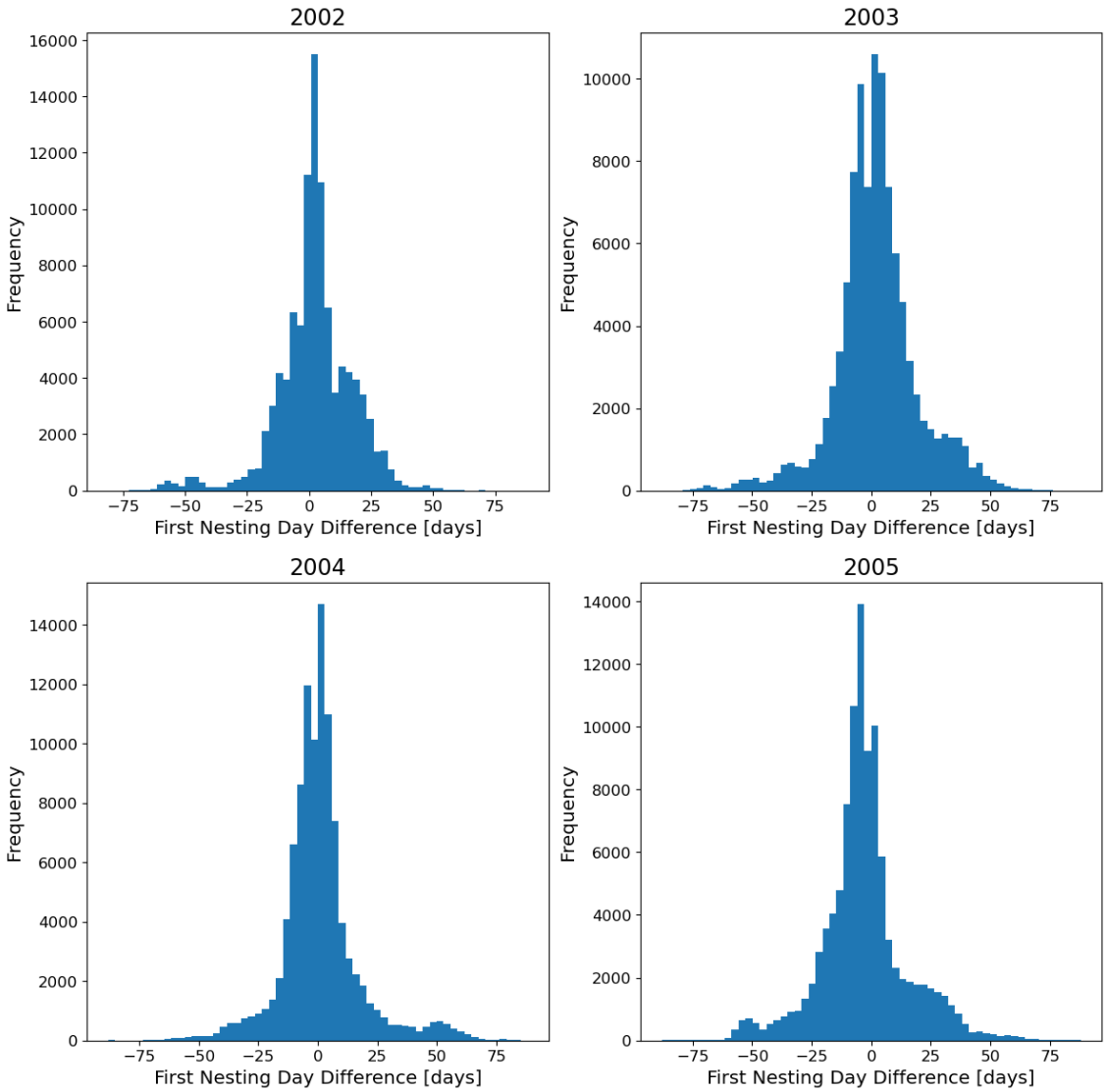


Figure A4: Histogram showing the differences per pixel between the first nesting day metric of the downscaled snow cover and the MODIS snow cover for each year between 2002 and 2005. A positive value means that MODIS predicts an earlier first nesting day than the downscaled product.

Table A3: Statistical differences between MODIS snow cover and downscaled snow cover for the years 2002-2005, restricted to data points with an elevation below 100 m. The mean and standard deviation are shown for both snow cover products, as are the RMSE and correlation between the two products are also shown. These are available for each month individually and for the total across all 3 months.

Snow Cover Fraction		**Mean (%)**		**Standard Deviation (%)**		**RMSE (%)**	**Correlation**
		MODIS	**DS**	**MODIS**	**DS**		
2002	**May**	88.4	93.8	19.8	13.5	17.7	0.558
	June	34.2	41.8	34.2	26.2	24.6	0.715
	July	5.02	11.1	13.5	11.8	16.5	0.243
	MJJ	45.1	49.0	42.7	38.8	19.8	0.894
2003	**May**	80.7	88.9	24.5	19.0	25.8	0.409
	June	48.7	52.6	38.0	29.9	26.4	0.72
	July	8.66	13.5	18.8	13.2	20.0	0.289
	MJJ	43.0	51.7	40.7	37.9	23.8	0.835
2004	**May**	79.6	84.3	24.5	21.6	23.8	0.498
	June	58.9	55.0	37.1	29.3	26.6	0.712
	July	8.55	10.9	18.9	11.2	17.9	0.38
	MJJ	43.6	50.0	40.8	37.4	22.3	0.847
2005	**May**	82.9	89.6	22.9	17.5	22.8	0.443
	June	55.3	46.6	37.4	26.3	27.5	0.715
	July	10.4	13.7	21.8	16.3	24.4	0.213
	MJJ	47.7	50.0	41.5	37.3	24.8	0.805

Table A4: Statistical differences between the predicted first available nesting day across Svalbard as given by both the MODIS snow cover and the downscaled snow cover for the years 2002-2005, restricted to data points with an elevation below 100 m. The mean of the predicted first nesting day, measured by the date of the year, and the standard deviation across Svalbard, measured in days, are shown for both products. The RMSE between the two products is also shown.

First Nesting Day	**Mean (date of the year)**		**RMSE (days)**	**Standard Deviation (days)**	
	MODIS	**DS**		**MODIS**	**DS**
2002	June 13th	June 11th	15.2	16.2	13.0
2003	June 17th	June 14th	17.6	20.4	15.2
2004	June 18th	June 14th	16.8	19.1	17.3
2005	June 21st	June 13th	20.6	19.3	15.6