Oceanologia 64 (2022) 535—545

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.journals.elsevier.com/oceanologia

ORIGINAL RESEARCH ARTICLE

The decline of Svalbard land-fast sea ice extent as a result of climate change

Jacek A. Urbański[a,*], **Dagmara Litwicka**[b]

[a] *GIS Laboratory, Institute of Oceanography, University of Gdańsk, Gdynia, Poland*
[b] *Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland*

Received 14 July 2021; accepted 24 March 2022
Available online 8 April 2022

KEYWORDS
Svalbard;
Fast ice;
Warming;
Machine learning;
Random Forest

Abstract The Svalbard Archipelago has experienced some of the most severe temperature increases in the Arctic in the last three decades. This temperature rise has accelerated sea-ice melting along the coast of the archipelago, thus bringing changes to the local environment. In view of the importance of the near-future distribution of land-fast sea ice along the Svalbard coast, the available observation data on the ice extent between 1973 and 2018 are used herein to create a random forest (RF) model for predicting the daily ice extent and its spatial distribution according to the cumulative number of freezing and thawing degree days and the duration of the ice season. Two RF models are constructed by using either regression or classification algorithms. The regression model makes it possible to estimate the extent of land-fast ice with a root mean square error (RMSE) of 800 km^2, while the classification model creates a cluster of submodels in order to forecast the spatial distribution of land-fast ice with less than 10% error. The models also enable the reconstruction of the past ice extent, and the prediction of the near-future extent, from standard meteorological data, and can even analyze the real-time spatial variability of land-fast ice. On average, the minimum two-monthly extent of land-fast sea ice along the Svalbard coast was about 12,000 km^2 between 1973 and 2000. In 2005—2019, however, the ice extent declined to about 6,000 km^2. A further increase in mean winter air temperatures by two degrees, which is forecast in 10 to 20 years, will result in a minimum two-monthly land-fast ice extent of about 1,500 km^2, thus indicating a trend of declining land-fast ice extent in this area.

* Corresponding author at: GIS Laboratory, Institute of Oceanography, University of Gdańsk, Gdynia, 81-378, Poland.
E-mail address: jacek.urbanski@ug.edu.pl (J.A. Urbański).
Peer review under the responsibility of the Institute of Oceanology of the Polish Academy of Sciences.

Production and hosting by Elsevier

https://doi.org/10.1016/j.oceano.2022.03.008

1. Introduction

The Svalbard Archipelago is the largest land area in the European part of the Arctic. The West Spitsbergen Current and the semi-continuous weather front between the cold masses of Arctic air and the warmer air of the polar cell strongly influence the Svalbard climate. Because of the low air temperature and the highly-indented coastline, the coastal waters of Svalbard are covered every year by land-fast sea ice (referred to hereafter as fast ice), i.e., ice that holds fast to the coastline or the sea bottom. Fast ice usually accumulates in fjords, between islands, and in shallow inshore waters. In the Arctic, fast ice is biologically significant as a breeding and molting site for seals, mainly ringed seals (*Pusa hispida*), which are the principal prey of polar bears (*Ursus maritimus*) (Krafft et al., 2006; Smith and Lydersen, 1991). Moreover, fast ice protects coastal areas from erosion by wave action for as long as it persists. The effects of climate change are intensified in this region, with the Arctic experiencing one of the largest increases in air temperature in the present century (Førland et al., 2011; Isaksen et al., 2016). Since the beginning of the 20th century, meteorological observations have shown that the air temperature has always fluctuated in this region, and the temperature has risen by 4—5°C during the last 40—50 years (Hanssen-Bauer et al., 2019). Although the winter air temperatures between the 1960s and early 1990s were only slightly higher than at the beginning of the 20th century, the temperature increased by the beginning of the 21st century and continues to do so (Nordli et al., 2014, 2020).

The mean annual temperature changes at three meteorological stations in the Svalbard Archipelago between 1975 and 2018 are presented in Figure 1, along with the mean winter temperature changes, which is when land-fast ice predominantly occurs. Current forecasts envisage a mean annual temperature rise in this region of at least 1°C per decade until the mid-21st century (Hanssen-Bauer et al., 2019).

The large-scale permanent monitoring of local sea ice conditions in the fjords and coastal waters of Svalbard began just under 20 years ago using the new technology of high and medium resolution satellite imaging, mainly via C-band Synthetic Aperture Radar (SAR) sensors (Hanssen-Bauer et al., 2019; Johansson et al., 2020; Muckenhuber et al., 2016) and GIS-based automatic or semi-automatic systems for sea ice classification (Zakhvatkina et al., 2019). Since 2005, the Norwegian Ice Service has produced ice charts of the Svalbard area almost daily (Monday—Friday). Since the new methods of ice mapping were introduced, it has been possible gradually to improve the accuracy of the maps. Previously, fast ice conditions in Svalbard were assessed mainly from various observations made as part of several fast-ice related projects (Gerland et al., 2008; Hanssen-Bauer et al., 2019; Zhuravskiy et al., 2012). For example, analyses of the total number of fast-ice days before 1st April (the ring seal pupping date) were performed using satellite data from 1974 to 1988 to reveal a substantial interannual variability of 0—155 days in the fjords of northern Spitsbergen (the largest island of Svalbard), and 38—107 days on the western coast (Smith and Lydersen, 1991).

There have been several reports on local fast ice conditions in the last 20 years. For example, an analysis of the temporal changes in ice cover during 2000—2014 in Isfjorden and Hornsund using SAR and optical images revealed a significant decrease in the extent of fast ice in both fjords (Muckenhuber et al., 2016). Meanwhile, systematic observations in Kongsfjorden since 2003 initially detected substantial interannual variability in the fast ice extent with intervals of 2—3 years or more (Gerland and Renner, 2007), i.e., a similar pattern to that observed in the 1970s and 1980s. However, more recent observations have indicated that the

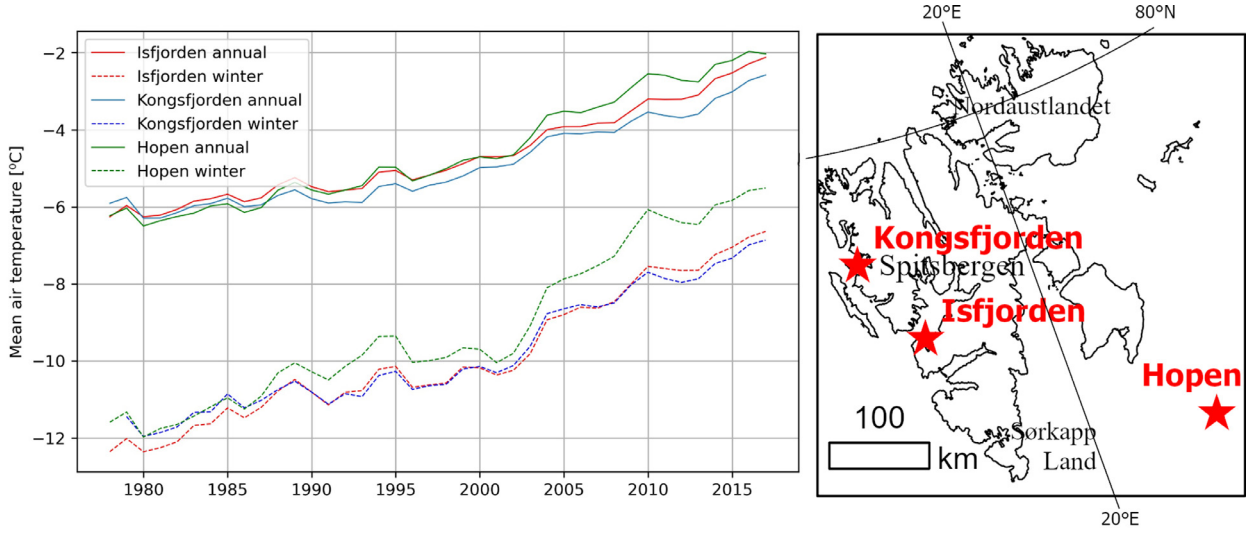

Figure 1 The air temperature changes at three meteorological stations in the Svalbard Archipelago between 1975 and 2018. The solid lines show the annual mean, and the dashed lines show the winter (December—May) means. Eight-year smoothing was used with a rolling yearly mean (left). The locations of the meteorolgical stations: Isfjorden—Barentsburg (78.1°N, 14.3°E), Kongsfjorden—Ny-Ålesund (78.923°N, 11.933°E), and Hopen (76.5°N, 25.067°E) (right).

ice extent is smaller in most years and that the ice season is becoming shorter (Pavlova et al., 2019). In particular, there has been a smaller extent of fast ice near the northern coasts of Svalbard, although occasional observations have shown that the fast ice cover can last from November until July (Wang et al., 2013). It is generally agreed that the duration of fast ice cover around Svalbard has become perceptibly shorter in the last ten years (Dahlke et al., 2020; Hanssen-Bauer et al., 2019; Pavlova et al., 2019). An understanding of this temporal distribution is also vital in studying environmental changes, coastal erosion, and the ecology of many species (Hanssen-Bauer et al., 2019; Krafft et al., 2006; Smith and Lydersen, 1991). Hence, the present project aims to determine the changes in the extent of fast ice and to predict its near-future extent and distribution.

2. Material and methods

2.1. Meteorological data

Two sets of data were used in the present work — one for modeling and the other for assessing the modeling method. In the first case, observational meteorological data from the Hopen, Barentsburg, and Ny-Ålesund stations between 1973 and 2019 were acquired as daily summaries, including minimum, maximum and average air temperatures, from the National Centres for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) (https://www.ncdc.noaa.gov/cdo-web/datatools/selectlocation) (Menne et al., 2012). In the second case, the Arctic Regional Reanalysis dataset of hourly short-term forecasts of surface meteorological variables at a 2.5-km resolution was used. This dataset was produced using the HARMONIE-AROME regional numerical weather prediction model (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-carra-single-levels?tab=overview), and the 2 m daily air temperatures at noon from 1998 to 2019 were downloaded as multivariate rasters for use in the present study.

The mean annual and winter air temperature changes between 1975 and 2018 at the above-mentioned meteorological stations reveal a significant increase in the winter mean temperature between 2000 and 2005 (Figure 1). This time period was therefore used herein to divide the analyzed period into two parts.

2.2. Ice data

The following three sets of sea-ice data were used in the present project: (i) the daily operational ice charts produced by the Norwegian Ice Service in 2005—2018 for the Svalbard Archipelago were downloaded from the archive dataset of the Norwegian Meteorological Institute (https://cryo.met.no/archive/ice-service/icecharts/quicklooks/); (ii) the ice charts for 1973—1998 were downloaded in vector.shp format from the Climate and Cryosphere Historical Ice Chart Archive (http://www.climate-cryosphere.org/resources/historical-ice-chart-archive/gis-data), and (iii) the 10 km x 10 km raster-gridded daily sea ice concentration data for the entire Arctic between 2003 and the present were obtained from the US. National Snow and Ice Data Center (https://nsidc.org/data/G10033/versions/1). The latter data set was created using various methods according to availability over time. Thus, the earliest charts were produced by combining sources such as aerial reconnaissance, surface observations, and airborne and ship reports. Subsequently, infrared and visible-band satellite imagery from the Advanced Very High-Resolution Radiometer (AVHRR) was used, and currently, data from the Synthetic Aperture Radar and Advanced Microwave Scanning Radiometer are used. According to Yu et al. (2014), the relative uncertainty in chart-derived fast ice extent can range from 5% to 25% on average. As this estimation was made for the 25-km resolution National Ice Center charts produced in 1975—2006, the relative uncertainty is assumed to be 5—20% herein. The GIS vector and raster layers with a georeferenced coastline and land-water mask of the Svalbard Archipelago were obtained from the GIS Centre, University of Gdańsk.

2.3. Data pre-processing

The observational meteorological data contained fields with incomplete values of the mean, maximum, and minimum daily air temperatures. The minimum and maximum values contained only occasional gaps, whereas gaps in the mean values were very frequent. The data pre-processing workflow involved the following two steps: (i) a linear regression model was built to estimate the average daily temperature from the minimum and maximum values, then (ii) a regression equation was used to fill gaps in the average daily temperature within each data set. These data are presented in Figure 1 above.

The next step was to calculate the cumulative number of freezing degree days (FDD) and thaw degree days (TDD) by employing Stefan's Law (Leppäranta, 1993):

$$FDD = \int_0^t [T_f - T_0(t)]dt \ \ for\ T_0 < T_f \quad (1)$$

$$TDD = \int_0^t [T_0(t) - T_f]dt \ \ for\ T_0 > T_f \quad (2)$$

where t is time, T_f is the freezing temperature, and $T_0(t)$ is the average daily temperature. The freezing temperature is −1.9°C (the freezing temperature of seawater). The formulas generally use the day as the unit of time. The FDD is also referred to as the sum of negative degree days, and is used to simplify the formula for estimating ice thickness (Leppäranta, 1993) by providing the cumulative sum of below-zero temperatures for each successive day. The TDD is similar, but with respect to temperatures above −1.9°C. Herein, it was assumed that the ice season starts on 1st September and lasts for 300 days until June. The definition of the ice season allows the number of ice seasons and the number of days within a particular ice season (ICSD) to be assigned to each day. Hence, a tidy text format file was created for 13,794 days (1/09/1973—27/06/2019). The following two spatial data models were used to organize the georeferenced data for analysis: (i) a set of 110 randomly distributed points in an area of fast ice localization, and (ii)

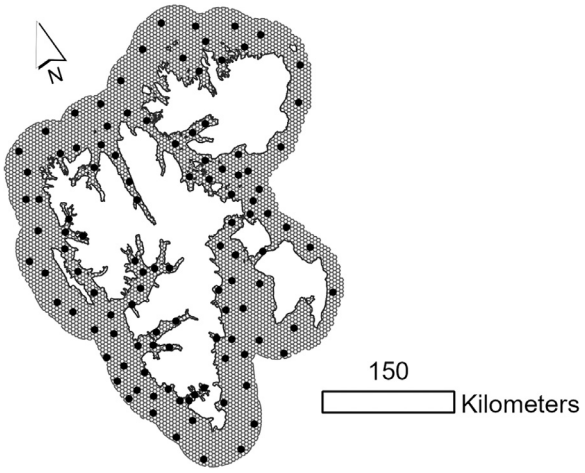

Figure 2 The spatial data models used in the present project: 110 points and a grid of 4782 hexagonal polygons.

a grid of 4782 polygonal cells with areas of 16 km^2. Each model was assigned a unique ID to points and polygons, as indicated in Figure 2. The grid of polygons defines the project area of interest, which is 76512 km^2.

The Arctic Regional Reanalysis dataset of daily air temperature was assigned to the randomly distributed points presented in Figure 2 by extracting its values from rasters. In the resulting tidy text format, the file rows represent days in 110 columns, with the air temperature at each point.

The scanned ice cover maps were manually or automatically digitized to time series of rasters with the following classes: 0 (water), 1 (land), 2 (open water), 3 (very open drift ice), 4 (open drift ice), 5 (close drift ice), 6 (very close drift ice), and 7 (fast ice). The spatial resolution of the rasters was 300 m, and the fast ice classes were used to convert the vector ice charts to the raster format with the same shape and spatial resolution. Combining these two ice sets gives the time series of 1960 rasters for the years 1973—2018, with a gap between 1999 and 2004. It contains, on average, one or two ice concentration maps per week. The tidy text format file of fast ice contains rows representing each map, with a separate column for each hexagon (4782 columns in total), with classes 1 and 0 indicating the presence and absence of fast ice.

The daily sea ice concentration raster gridded data for the entire Arctic was used along with the fast ice class to create an additional text file with two columns representing time and fast ice extent in the area of interest.

2.4. Machine learning modeling

The RF regression and classification models were used to predict the ice cover extent and to classify the hexagonal grid cells for the two classes (ice/no ice) on any day using the FDD, TDD, and ICSD as predictive features. The RF model, introduced by Breiman (2001), has been used in many geophysical and environmental applications (Lutz et al., 2018; Mutanga et al., 2012; Rodriguez-Galiano et al., 2014, 2015). In this method, random decision trees are created by bootstrapping data in which the sample data are drawn and replaced. Then the majority vote or average prediction across all trees is used to generate the result. The main advantages of the RF model are that overfitting can be reduced by averaging several trees and that no statistical assumptions are required regarding normal distribution and data linearity. The model also allows one to measure the relative importance of each feature for the prediction. Furthermore, it is easy to apply, and the few default hyperparameters usually give good results. However, one of the model's shortcomings is that it is impossible to extrapolate beyond the range of values in the training set. As a result, predictions can be made only in the range of values represented by the training data set. In the present study, the range of values in the training set is 0—2800 for the FDD, 1.1—300 for the TDD, and 0—38288 km^2 for the ice cover extent. The typical workflow uses a training set containing the dependent and independent variables to train the RF model, and a test data set to validate the results. The RF modeling was performed using the Python Scikit-learn package (Pedregosa et al., 2011). The following two models were created: (i) a regression model for predicting the total fast ice extent (in km^2) along the coast of Svalbard for every day with available FDD, TDD, and ICSN data, and (ii) a classification model for predicting the presence or absence of fast ice in each hexagonal cell of the grid (by assigning the values 1 or 0) and for building the cluster of independent models in order to predict the spatial distribution of fast ice.

The regression and classification modeling uses a training set to build a model and test sets to validate the model. Both sets are randomly created by splitting the basic set of daily data with known FDD, TDD, and ICSN, along with the total fast ice extents for the regression model and the presence or absence values of each cell in the cluster for the classification model. A standard method for evaluating the accuracy of the model on continuous data (total fast ice extent) is the root mean square error (RMS). However, the error rate is used for evaluating the machine learning classification model, where the error rate is defined as one minus the accuracy, and the accuracy is the ratio of the number of observations with correct classification to the total number of classified observations. The Scikit-learn functionality also allows the tuning of hyperparameters and the analysis of feature importance, the latter being a measure of how the random shuffle of a particular feature influences the result. From 1973 to the present time, the available data range makes the RF method applicable for the present project. In view of the spatial scale (a few hundred square kilometers) and the statistical nature of the models, it is assumed that the FDD and TDD are highly autocorrelated and that the values from one point in Isfjorden can be used. This assumption is discussed in detail later.

The random hyperparameters used in the present work are listed in Table 1. These were tuned by using a field cross-validation grid.

The accuracy of the model was evaluated via the following two steps: (i) the feature importance was evaluated using Scikit-learn, giving accuracies of 60%, 30%, and 10% for the FDD, TDD, and ICESD, respectively, and (ii) the RMS error was calculated as 802 km^2, which is approximately the area of 50 cells in Figure 2. The RF classification model using a cluster of 4782 independent submodels uses the same features as the regression model. The model maps the fast ice distribution in time. The array of fast ice cells can be

Table 1 The hyperparameters tested and used in sklearn.ensemble.RandomForestRegressor (scikit-learn 0.23.2).

Hyperparametres	Hyperparameter grid values tested	Best hyperparameters
Using bootstrap sample when building trees.	True, False	True
The maximum depth of tree.	10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None	80
The number of features considered when looking for the best split.	auto, sqrt	auto
The maximum number of samples required to be at a leaf node.	1,2,4	2
The minimum number of samples required to split an internal mode.	2,5,10	2
The number of trees in the forest.	200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000	1800

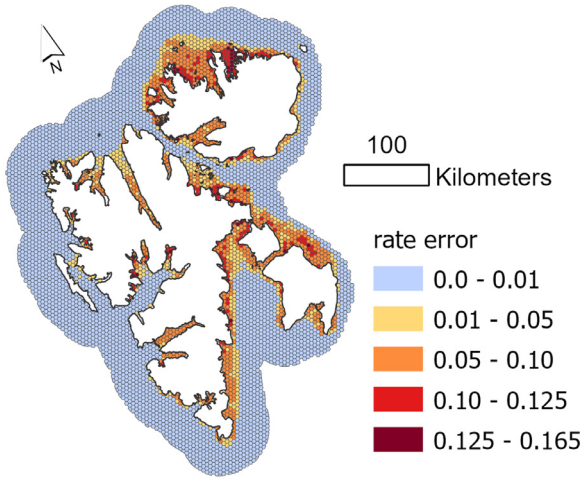

Figure 3 The rate errors in the predicted spatial distribution of fast ice by the classification model.

summed for any day to give the fast ice cover in km^2. The local ice cover depends on, among other things, the water depth, water heat flux, wave action, water circulation, water heat flux, and local microclimate. All these features are irrelevant in the machine learning process if the dependent variable (target) represents a single location. The model uses only basic features from one location and predicts new values only for that particular location. As a result, the robust autocorrelation of features in space is irrelevant. The results of error evaluation for the fast ice distribution model are presented as a rate error map in Figure 3. The rate errors were calculated independently for each cell using the corresponding rows of the validation set. Thus, the classification model exhibits a less than 10% error in predicting the spatial distribution of fast ice for most of the study area. Low error values cover the area, with occasional fast ice only.

3. Results

3.1. Cumulative freezing degree days

The cumulative freezing degree days were previously used by Yu et al. (2014) to explain the changes in fast ice extent in the Arctic because this approach provides a measure of both the changing surface air temperature and its cumulative effect throughout an entire ice growth season. The FDD is the simplest parameter that correlates well with fast ice duration due to its cumulative nature (Leppäranta, 2014). As an ice season straddles two consecutive years, starting in autumn and ending in early summer, all statistics should apply to an ice season rather than a year. The cumulative FDD in each ice season from 1973/74 to 2018/19 in Isfjorden is presented in Figure 4, where the values are assigned to the year when the ice season begins.

3.2. Fast Ice extent in Svalbard

An RF regression model was used to model the changes in fast ice extent in the area of interest, with a total surface of about 76,000 km^2 (Figure 5). The extent of fast ice fluctuates from year to year with 3—4 year cycles of high and low values. Two periods of ice extents larger than 25,000 km^2 appeared before 1990, separated by periods with less extensive fast ice. The winter season fast ice extent is larger than 15,000 km^2 on only a few occasions after 2000, whereas such ice seasons are common before 2000. The modeled extents are compared with the results of Yu et al. (2014) for the period 1975—2007. However, their analysis was performed using data with a spatial resolution of 25 km^2, which may be too coarse for the Svalbard fjords. Nevertheless, the two-time series show several similarities. They both have ice extents of less than 15,000 km^2 for nearly all ice seasons after 2000, and only a few such low extents before. The winters with large extents, such as 1977—1978 and 1981—1982 are also visible in both. However, the maximum fast ice extents reported by Yu et al. (2014) are larger than 40,000 km^2. This is because they occasionally recorded fast ice beyond the area of interest of the present work.

3.3. Distribution of fast ice extent

The cell net classification model makes it possible to estimate the spatial distribution of fast ice for any day. Furthermore, as the created time series is complete, it is easy to calculate statistics for fast ice extent within any time range. For example, the average duration of fast ice in the periods 1973—2000 and 2005—2019 are seen to differ significantly (Figure 6a and b). According to Dahlke et al. (2020), such time ranges emphasize the stationary ice situation be-

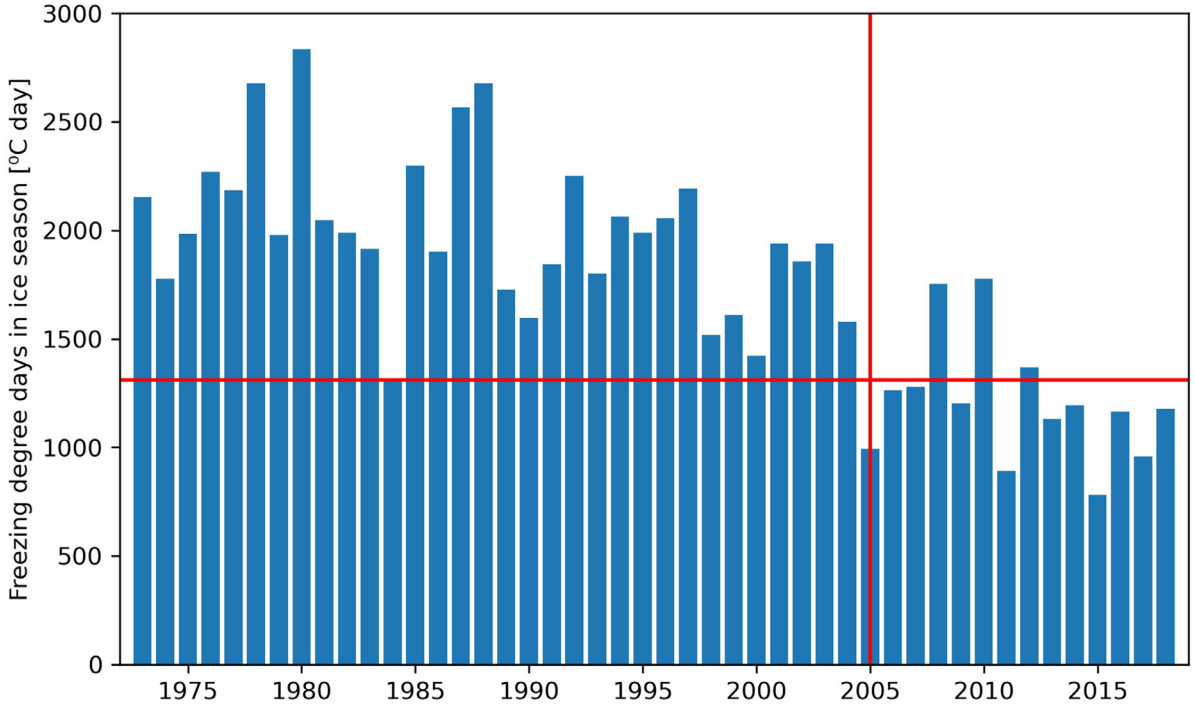

Figure 4 The cumulative freezing degree days (FDD) in Isfjorden during the ice seasons 1973/74 to 2018/19 (assigned to the year when the ice season starts). The vertical red line indicates the year 2005 as the start of warmer winters, and the horizontal line shows the minimum FDD before 2005 (the season 1984/85 was the one that determined this level).

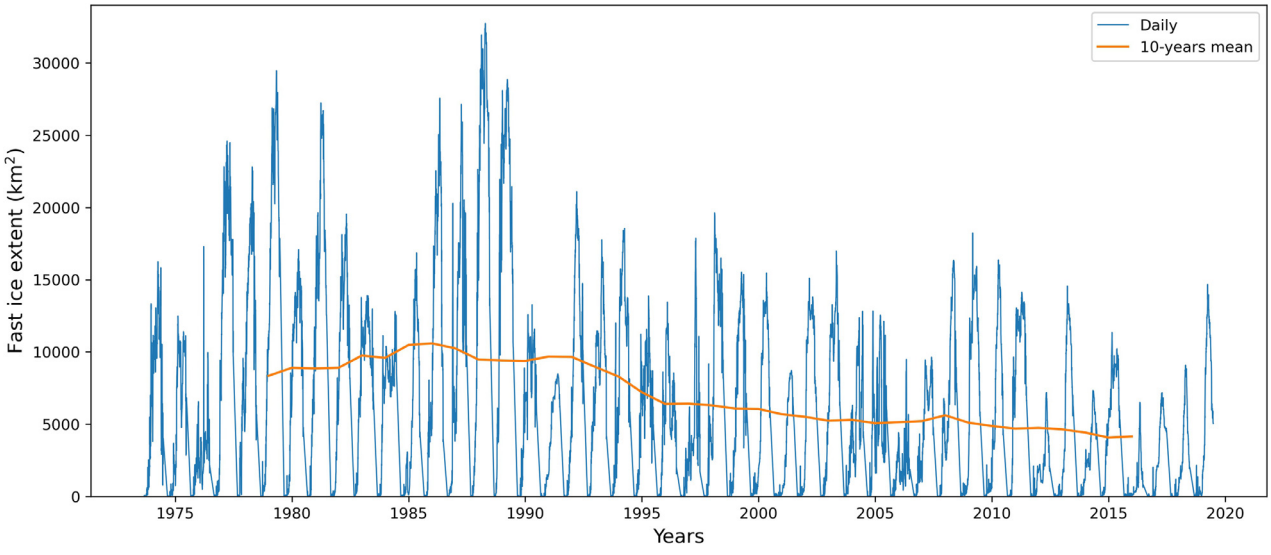

Figure 5 The time series of fast ice extent obtained using the regression model, where the orange line shows 10-year rolling mean.

fore 2000 and the significantly reduced ice cover in recent times.

Between 1973 and 2000, at least half of the surface area of the fjords in west Spitsbergen was covered by fast ice for 4 to 5 months, while the fjords in northern Spitsbergen were entirely covered by fast ice for the same length of time. The northwestern Spitsbergen coast also had fast ice cover for 4–5 months (Figure 6a). The east coast, where there are no fjords, was covered by fast ice, mainly in the bays. Between 2005 and 2019, however, the distribution and duration of fast ice cover changed dramatically. In the fjords of West Spitsbergen, fast ice persisted for four months only at their heads, whereas those in northern Spitsbergen were covered by fast ice for less than half the fjord lengths, usually for two months and only locally for three months (Figure 6b). The most significant differences in duration are in the fjords of Spitsbergen, with the largest changes being in the northern ones. Throughout this area, the duration has been reduced by 3–4 months. Such dramatic changes have occurred only locally; elsewhere, the reduction has

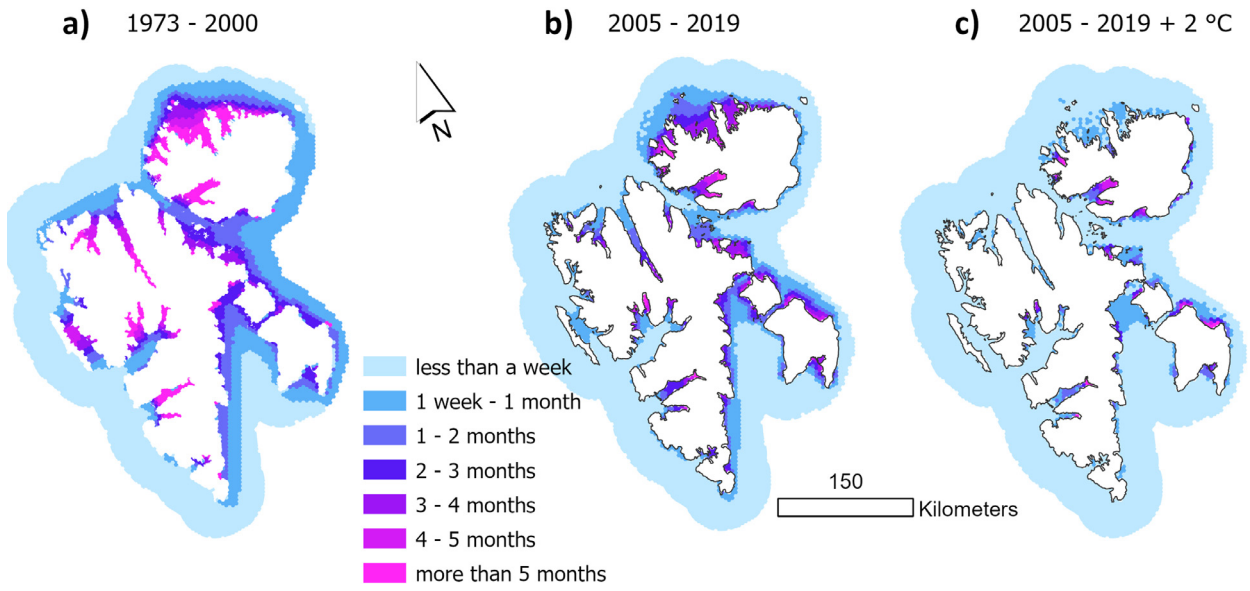

Figure 6 The mean distribution of the fast ice duration in Svalbard (a) during the ice seasons of 1973—2000, (b) during the ice seasons of 2005—2019, and (c) in the near future, assuming a 2°C increase in winter air temperature.

Table 2 The extent of fast ice for various durations of ice cover, with durations of more than 2 months indicated in bold.

Duration of fast ice cover	1973—2000 extent (km^2)	2005—2019 extent (km^2)	2005—2019 + 2°C; near-future extent (km^2)
1 week—1 month	18 059	11 255	6 085
1—2 months	7 464	4 119	1 341
2—3 months	4 467	2 564	712
3—4 months	2 363 **12 320**	2 391 **6 177**	301 **1 522**
4—5 months	1 830	930	356
more than 5 months	3 660	292	153

generally been from 1 to 3 months. In northern Svalbard, the duration of fast ice has decreased by two months; elsewhere, the changes are no greater than one month, and locally no more than two months. The significant changes relate mainly to the fjord's branches (e.g. in Isfjorden), where the fast ice cover before 2000 was long-lasting.

The map in Figure 6c shows the model scenario used in the present study, in which the average fast ice duration for the years 2005—2019 is projected into the future 10—20 years, based on the two-degree increase in the winter air temperature predicted by Hanssen-Bauer et al. (2019). Here, a dramatic decrease in the ice extent is predicted across the entire area. The areas of fast ice extent according to duration are presented in Table 2. However, these total values will differ from the modeled maximum areas because the ice cover may occur at different times. According to the validation procedure described above, the values in Table 2 have an error rate no greater than 10%.

4. Discussion

Although many papers describe the decrease of fast ice extent in Svalbard (e.g., Gerland et al., 2008; Hanssen-Bauer et al., 2019; Zhuravskiy et al., 2012), no prediction was made therein regarding the near future. Nevertheless, given the firm basis of air temperature forecasts, these can be used to model the future fast ice extent. Indeed, the prediction of fast ice extent is vital for the analysis of many physical and biological processes. Among these, the most important physical processes are the terrestrial discharge of freshwater (McClelland et al., 2012) and coastal erosion (Frederick et al., 2016). In addition, changes in the extent of fast ice have significant ecological consequences (Krafft et al., 2006; Smith and Lydersen, 1991).

The important assumption in the present project is that the air temperature changes have strong spatial autocorrelation, thus enabling spatial modeling of the entire area based on the temperature at one point only. The cumulative FDD was previously used by Yu et al. (2014) to understand the long-term changes in fast ice in the Arctic. Hence, the temporal and spatial distributions and their correlation with the FDD were analyzed in the present study using the Arctic Regional Reanalysis dataset of hourly short-term forecasts of surface meteorological variables. As the aim was to use the FDD series from Isfjorden, all series were correlated with this point. Due to the massive amount of data, the FDD was calculated using the air temperature at noon instead of the mean daily air temperature. This is justified because the elevation of the sun doesn't change significantly during the day and, hence, the daily temperature range is small. The analysis was performed using a 110-point

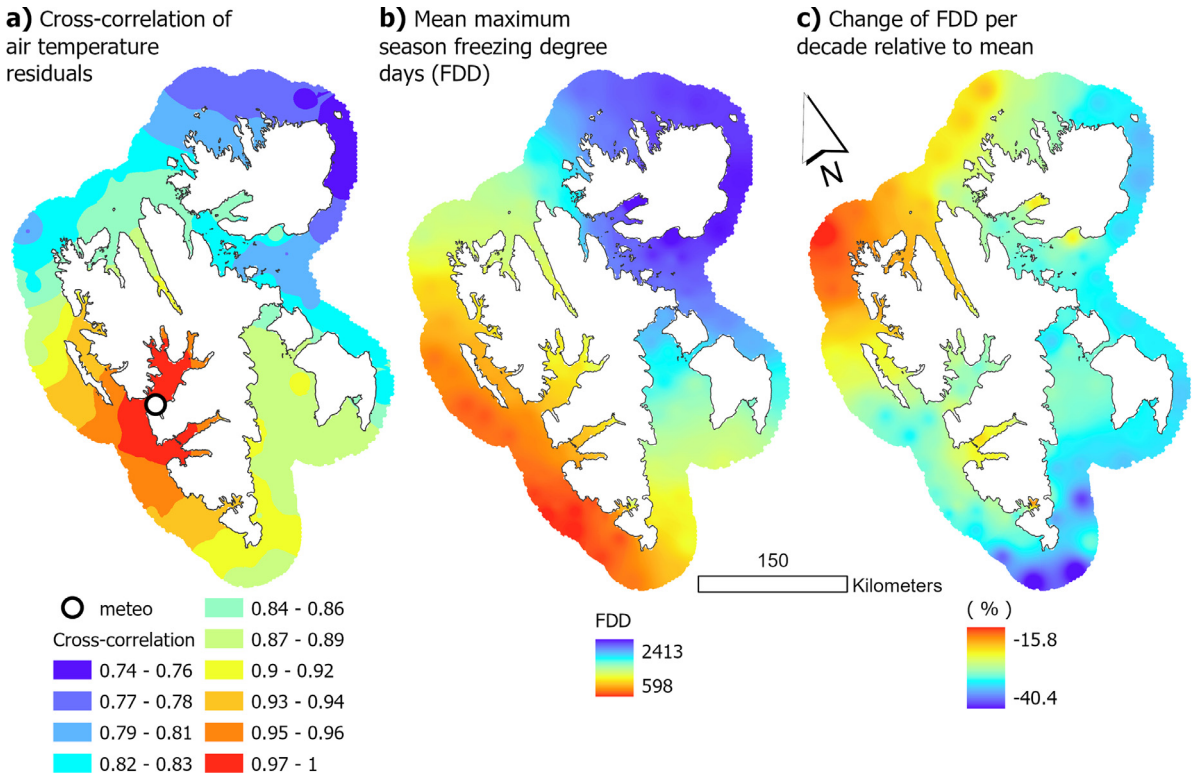

Figure 7 The cross-correlation of air temperature residuals (a). The mean maximum seasonal cumulative freezing degree days (FDD) (b). The decadal change in FDD relative to the mean (c).

time series of daily mid-day air temperatures during 1998–2019 (Figure 2). Because the cross-correlation is impacted by time series dependence, the correlation analyses were performed on temperature residues.

The results of the correlation analyses are presented in Figure 7. For most of the analyzed areas, the correlation coefficients of temperature residues with Isfjorden are higher than 0.8 (Figure 7a). Only in the northern-western parts of Nordaustlandet do the coefficients decrease to 0.74. Despite the influence of various water masses, the west and east coast air temperatures correlate well, with values of about 0.9. This strong correlation is essential when building a total regression model and distributed classification models of fast ice extent based on one-point FDD and TDD values. Meanwhile, the mean maximum ice season FDD is seen to change gradually from the southwest to north-east in the range 600 to 2400 days (Figure 7b), in agreement with the results of Yu et al. (2014) for the years 1977–2007. Further, the results in Figure 7c indicate that the estimated changes in FDD relative to the mean value per decade are higher than 15% for all areas. The fastest changes take place near Sørkapp, with a 40% decrease per decade, while the smallest changes occur at northern-western Spitsbergen.

As detailed in the Materials and methods section, three different data sets were used to create and evaluate a continuous picture of fast ice changes near the coasts of Svalbard (Figure 8). Here, all three data sets contain classified land-fast ice. The reliability of this classification is reflected in the metadata for each dataset, and is not within the scope of the present project.

The time series of FDD was calculated for Isfjorden. As the air temperature in Kongsfjorden is nearly identical, the FDD will be very similar in both fjords. The results in Figure 4 confirm the observed shortening of the fast ice season from 2005/2006 onwards (Pavlova et al., 2019). Prior to that season, the FDD oscillated with a frequency of a few years, varying between 1800 and 2300 days, with an overall decreasing trend. After 2005/2006, the FDD was significantly *higher* than the 1973–2005 (32-season) minimum in only two out of 13 subsequent seasons, but was lower than the minimum (between 900 and 1200 days) in 9 seasons. Interannual variability still occurs, but the decreasing trend in FDD now applies to both the mean and minimum values. The general trends reflected in the above analysis thus confirm the earlier results (Gerland and Hall, 2006; Gerland and Renner, 2007; Pavlova et al., 2019).

Machine learning models have two advantages over the observational mapping of sea ice extents. As noted above, such maps were created using various methods according to availability, thus improving with time. By contrast, the accuracy of estimation using a machine learning model is constant, as determined by the evaluation process. Furthermore, when accurate data are unavailable, better results can be obtained by teaching the model using more accurate contemporary data. In addition, such models can be used for forecasting the future fast ice extent, and the performance of the model can be evaluated by comparing its predictions with observations in the area of interest (the grid of hexagonal polygons in Figure 2). The results of the RF regression and classification models are compared with the Norwegian Ice Service ice-chart data set and the grid data from the US

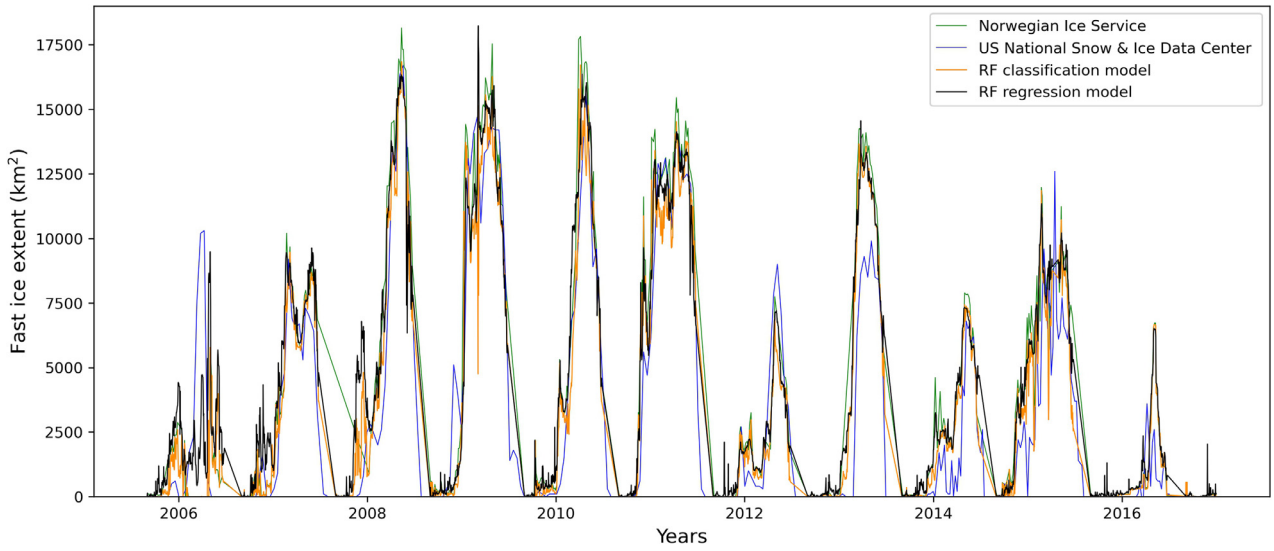

Figure 8 The time series of fast ice extent for project AOI (Figure 2) according to the data from the ice charts produced by the Norwegian Ice Service, the grid data from the US National Snow and Ice Data Center, the random forest classification model results, and the random forest regression model results. The ticks mark link each year to the 1st January of the next year.

National Snow and Ice Data Center in Figure 8. The comparison is for the time window when both data sources were available. A significant similarity between these four time series is observed, although some differences between the two observational datasets exist.

The two main features controlling the model results are the FDD and TDD, although the influence of the FDD is twice that of the TDD. The machine learning models are designed to make the most accurate predictions possible. For example, in the distributed classification model, all local factors such as water temperature or circulation are indirectly included when designing the rules for a particular model cell. The fast ice extent time series obtained using the regression model is presented in Figure 5, and the classification model gives nearly identical results (Figure 8). The characteristic feature is one of fluctuations in the ice extent that are superimposed onto the long-term trend represented by the 10-year rolling mean (orange line on Figure 5). In the latter, the fast ice extent is seen to decrease gradually over time, being halved within the last forty years (a rate of about 100 km^2 per year). This result in in agreement with that of Dahlke et al. (2020), who indicated a decrease of 10—20% per decade in the fjords. In our study, the analysis of spatial change was made using a cluster of local random forest models; this is a new method, but similar solutions have been used by others (e.g., Georganos et al., 2019). The present results show a significant decline in fast ice (Figure 6 and Table 2). The coverage of fast ice lasting more than two months was 12,320 km^2 in 1973—2000, but was halved in 2005—2019. Further, according to the forecast, an increase in temperature by two degrees will decrease the two-month ice coverage four-fold compared to that of 2005—2019.

The amount of fast ice is not a function of air temperature only. The surface air temperature can explain around 25—26% of the fast ice variability (Dahlke, et al., 2020). Many factors have smaller or larger impacts locally. The increased freshwater input from glaciers may alter the fjord fast ice persistence. The temperature and salinity of the water column also play important roles. For example, Kongsfjorden is affected by heat transport in the upper water column, which may impact the ice loss (Cottier et al., 2007; Sundfjord et al., 2017). Also, surface wind stress can cause advecting drift ice, and the mechanical destruction of the existing fast ice may change its extent (King et al., 2017). The snow cover has also important effects on development of fast ice (Wang et al., 2015). As a result, the presence of ice is determined by the unique relationship between air temperature and other factors at each locality. The classification model creates a separate RF model for each locality with an accuracy of greater than 80% (Figure 3). Additionally, the use of air temperature predictions enables these models to predict the future fast ice extent.

5. Conclusions

(1) The quality of ice extent data depends on the time of its acquisition, with older data being less accurate than newer data. However, it is possible to use current observations to model the past via machine learning with errors that are mostly independent of time. Machine learning is also an effective method for modeling the near-future extent and spatial distribution of fast ice. The modeling uses cumulative freezing degree days, thawing degree days, and the total number of days of a particular ice season. Due to the statistical nature of the modeling process, it is possible to use data from a single station if a strong spatial correlation exists.

(2) The most crucial parameter is the cumulative freezing degree days (FDD), with a 60% influence upon the fast ice extent and duration. Since 2005, a significant de-

crease in the FDD has been observed near Svalbard. The changes in air temperature and, hence, the FDD, are strongly correlated in space. The FDD has a significant increasing trend from the southwest to the northeast, with about 400 days across 100 km. As a result of warming, the highest (40%) decrease in FDD per decade is observed near Sørkapp. The lowest (10%) decrease is observed at northwest Spitsbergen.

(3) The maximum winter fast ice extent in the fjords and coastal waters of Svalbard has oscillated from 5,000 to 30,000 km^2 in the last 50 years, with a few-yearly sequence of higher and lower values. The greatest fast ice extent occurred during certain years between 1975 and 1990.
During the last forty years, the mean ice extent has decreased by half at a rate of about 100 km^2 per year.

(4) The most crucial changes in ice extent duration, before and after 2000, are in the western and northern parts of Spitsbergen. The reduction in ice presence is from 1 to 3 months. An increase in mean air temperature by two degrees will reduce the duration of fast ice extent by approximately four times compared to the present.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was carried out as part of the project ACCES "Deicing of Arctic Coasts: Critical or new opportunities for marine biodiversity and Ecosystem Services?" funded by the Belmont Forum, Call Title: Biodiversity2017 – Scenarios of Biodiversity and Ecosystem Services.

References

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Cottier, F., Nilsen, F., Inall, M., Gerland, S., Tverberg, V., Svendsen, H., 2007. Wintertime warming of an Arctic shelf in response to large-scale atmospheric circulation. Geophys. Res. Lett. 34 (10), 1–5. https://doi.org/10.1029/2007GL029948

Dahlke, S., Hughes, N., Wagner, P, Gerland, S., Wawrzyniak, T., Ivanov, B., Maturilli, M., 2020. The observed recent surface air temperature development across Svalbard and concurring footprints in local sea ice cover. Int. J. Climatol. 40, 5246–5265. https://doi.org/10.1002/joc.6517

Førland, E., Benestad, R., Hanssen-Bauer, I., Haugen, J., Skaugen, T., 2011. Temperature and Precipitation Development at Svalbard 1900 –2100. Adv. Meteorol. 2011. https://doi.org/10.1155/2011/893790

Frederick, J.M., Thomas, M.A., Bull, D.L., Jones, C.A., Roberts, J.D., 2016, The Arctic Coastal Erosion Problem, Sandia Report, SAND2016-9762, 122 pp.

Georganos, S., Grippa, T., Niang Gadiaga, A., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., Wolff, E., Kalogirou, S., 2019. Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. Geocarto Int. 36 (2), 1–16. https://doi.org/10.1080/10106049.2019.1595177

Gerland, S., Hall, R., 2006. Variability of fast-ice thickness in Spitsbergen fjords. Ann. Glaciol. 44 (9296), 231–239. https://doi.org/10.3189/172756406781811367

Gerland, S., Renner, A.H.H., 2007. Sea-ice mass-balance monitoring in an Arctic fjord. Ann. Glaciol. 46 (9296), 435–442. https://doi.org/10.3189/172756407782871215

Gerland, S., Renner, A.H.H., Godtliebsen, F., Divine, D., Løyning, T.B., 2008. Decrease of sea ice thickness at Hopen, Barents Sea, during 1966-2007. Geophys. Res. Lett. 35 (6), 1–5. https://doi.org/10.1029/2007GL032716

Hanssen-Bauer, I., Førland, E. J., Hisdal, H., Mayer, S., Sandø, A. B., Sorteberg, A., Adakudlu, M., Andresen, J., Bakke, J., Beldring, S., Benestad, R., Bilt, W., Bogen, J., Borstad, C., Breili, K., Breivik, Ø., Børsheim, K. Y., Christiansen, H. H., Dobler, A., Engeset, R., Frauenfelder, R., Gerland, S., Gjelten, H. M., Gundersen, J., Isaksen, K., Jaedicke, C., Kierulf, H., Kohler, J., Li, H., Lutz, J., Melvold, K., Mezghani, A., Nilsen, F., Nilsen, I. B., Nilsen, J. E. Ø., Pavlova, O., Ravndal, O., Risebrobakken, B., Saloranta, T., Sandven, S., Schuler, T. V, Simpson, M. J. R., Skogen, M., Smedsrud, L. H., Sund, M., Vikhamar-Schuler, D., Westermann, S., Wong, W. K. 2019. Climate in Svalbard 2100, 1/2019, [online]. Available from: https://www.miljodirektoratet.no/globalassets/publikasjoner/M1242/M1242.pdf

Isaksen, K., Nordli, Ø., Førland, E.J., Łupikasza, E., Eastwood, S., 2016. Recent warming on Spitsbergen—Influenceof atmospheric circulation and sea ice cover. J. Geophys. Res. Atmos. 121, 11,913–11,931. https://doi.org/10.1002/2016JD025606

Johansson, A.M., Malnes, E., Gerland, S., Cristea, A., Doulgeris, A.P., Divine, D.V., Pavlova, O., Lauknes, T.R., 2020. Consistent ice and open water classification combining historical synthetic aperture radar satellite images from ERS-1/2, Envisat ASAR, RADARSAT-2 and Sentinel-1A/B. Ann. Glaciol. 61 (82), 1–11. https://doi.org/10.1017/aog.2019.52

King, J., Spreen, G., Gerland, S., Haas, C., Hendricks, S., Kaleschke, L., Wang, C., 2017. Sea-ice thickness from field measurements in the northwestern Barents sea. J. Geophys. Res.-Oceans 122 (2), 1497–1512. https://doi.org/10.1002/2016JC012199

Krafft, B.A., Kovacs, K.M., Andersen, M., Aars, J., Lydersen, C., Ergon, T., Haug, T., 2006. Abundance of ringed seals (Pusa hispida) in the fjords of Spitsbergen, Svalbard, during the peak molting period. Mar. Mammal Sci. 22 (2), 394–412. https://doi.org/10.1111/j.1748-7692.2006.00035.x

Leppäranta, M., 1993. A review of analytical models of sea-ice growth. Atmos.- Oceans 31 (1), 123–138. https://doi.org/10.1080/07055900.1993.9649465

Leppäranta, M., 2014. Freezing of Lakes and the Evolution of their Ice Cover. Springer-Praxis, Heidelberg, 301 pp. https://doi.org/10.1007/978-3-642-29081-7

Lutz, S.R., Krieg, R., Müller, C., Zink, M., Knöller, K., Samaniego, L., Merz, R., 2018. Spatial Patterns of Water Age: Using Young Water Fractions to Improve the Characterization of Transit Times in Contrasting Catchments. Water Resour. Res. 54 (7), 4767–4784. https://doi.org/10.1029/2017WR022216

McClelland, J.W., Holmes, R.M., Dunton, K.H., Macdonald, R.W., 2012. The Arctic Ocean Estuary. Estuar. Coast. 35, 353–368. https://doi.org/10.1007/s12237-010-9357-3

Menne, M.J., I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, X. Yin, S. A. and R. Ray, R.S. Vose, B.E. Gleason, Houston, T. G. 2012. Global Historical Climatology Network – Daily (GHCN-Daily), Version 3.

Muckenhuber, S., Nilsen, F., Korosov, A., Sandven, S., 2016. Sea ice cover in Kongsfjorden and Hornsund, Svalbard (2000-2014) from remote sensing data. Cryosphere 10 (1), 149–158. https://doi.org/10.5194/tc-10-149-2016

Mutanga, O., Adam, E., Cho, M.A., 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. Int. J. Appl. Earth Obs. Geoinf. 18, 399—406. https://doi.org/10.1016/j.jag.2012.03.012

Nordli, Ø., Przybylak, R., Ogilvie, A.E.J., Isaksen, K., 2014. Long-term temperature trends and variability on spitsbergen: The extended svalbard airport temperature series, 1898—2012. Polar Res. 33 (1 SUPPL). https://doi.org/10.3402/polar.v33.21349

Nordli, Ø., Wyszyński, P., Cjelten, H.M.J., Isaksen, K., Łupikasza, E., Niedźwiedź, T., 2020. Revisiting the extended Svalbard Airport monthly temperature series, and the compiled corresponding daily series 1898—2018. Polar Res. 39. https://doi.org/10.33265/polar.v39.3614

Pavlova, O., Gerland, S., Hop, H., 2019. Changes in Sea-Ice Extent and Thickness in Kongsfjorden, Svalbard (2003—2016). In: Hop, H., Wiencke, C. (Eds.), The Ecosystem of Kongsfjorden, Svalbard. Advances in Polar Ecology, vol 2. Springer, Cham. https://doi.org/10.1007/978-3-319-46425-1_4

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Doubourg, V., Vaderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825—2830.

Rodriguez-Galiano, V., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., Ribeiro, L., 2014. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Sci. Total Environ. 476—477, 189—206. https://doi.org/10.1016/j.scitotenv.2014.01.001

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804—818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Smith, T.G., Lydersen, C., 1991. Availability of suitable land-fast ice and predation as factors limiting ringed seal populations, Phoca hispida, in Svalbard. Polar Res. 10 (2), 585—594. https://doi.org/10.1111/j.1751-8369.1991.tb00676.x

Sundfjord, A., Albretsen, J., Kasajima, Y., Skogseth, R., Kohler, J., Nuth, C., Skarðhamar, J., Cottier, F., Nilsen, F., Asplin, L., Gerland, S., Torsvik, T., 2017. Effects of glacier runoff and wind on surface layer dynamics and Atlanticwater exchange in Kongsfjorden, Svalbard; a model study. Estuar. Coast. Shelf Sci. 187, 260—272.

Wang, C., Shi, L., Gerland, S., Granskog, M.A., Renner, A.H.H., Li, Z., Hansen, E., Martma, T., 2013. Spring sea-ice evolution in Rijpfjorden (80°N), Svalbard, from in situ measurements and ice mass-balance buoy (IMB) data. Ann. Glaciol. 54 (62), 253—260. https://doi.org/10.3189/2013AoG62A135

Wang, C., Cheng, B., Wang, K., Gerland, S., Pavlova, O., 2015. Modelling snow ice and suoerimposed ice on landfastsea ice in Kongsfjorden, Svalbard. Polar. Res. 34, 20828. https://doi.org/10.3402/polar.v34.20828

Yu, Y., Fowler, C., Fetterer, F., Maslanik, J., 2014. Interannual Variability of Arctic Landfice Ice between 1976 and 2007. J. Clim. 27 (1), 227—243. https://doi.org/10.1175/JCLI-D-13-00178.1

Zakhvatkina, N., Smirnov, V., Bychkova, I., 2019. Satellite SAR data-based sea ice classification. An overview. Geosci. 9 (4), 3—5. https://doi.org/10.3390/geosciences9040152

Zhuravskiy, D., Ivanov, B., Pavlov, A., 2012. Ice conditions at Gronfjorden Bay, Svalbard, from 1974 to 2008. Polar Geogr. 35 (2), 169—176. https://doi.org/10.1080/1088937X.2012.662535