Hydrographic responses to regional covariates across the Kara Sea

Jussi Mäkinen and Jarno Vanhatalo

1Department of Environmental Sciences, University of Helsinki, Helsinki, Finland, 2Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland, 3Department of Biosciences, University of Helsinki, Helsinki, Finland

Abstract The Kara Sea is a shelf sea in the Arctic Ocean which has a strong spatiotemporal hydrographic variation driven by river discharge, air pressure, and sea ice. There is a lack of information about the effects of environmental variables on surface hydrography in different regions of the Kara Sea. We use a hierarchical spatially varying coefficient model to study the variation of sea surface temperature (SST) and salinity (SSS) in the Kara Sea between years 1980 and 2000. The model allows us to study the effects of climatic (Arctic oscillation index (AO)) and seasonal (river discharge and ice concentration) environmental covariates on hydrography. The hydrographic responses to covariates vary considerably between different regions of the Kara Sea. River discharge decreases SSS in the shallow shelf area and has a neutral effect in the northern Kara Sea. The responses of SST and SSS to AO show the effects of different wind and air pressure conditions on water circulation and hence on hydrography. Ice concentration has a constant effect across the Kara Sea. We estimated the average SST and SSS in the Kara Sea in 1980–2000. The average August SST over the Kara Sea in 1995–2000 was higher than the respective average in 1980–1984 with 99.9% probability and August SSS decreased with 77% probability between these time periods. We found a support that the winter season AO has an impact on the summer season hydrography, and temporal trends may be related to the varying level of winter season AO index.

1. Introduction

The Arctic Ocean is subject to decreasing cover and volume of sea ice, increasing sea surface temperature (SST), and changes in sea surface salinity (SSS) [McPhee et al., 2009; Steele et al., 2008; Stroeve et al., 2007]. We need more accurate predictions of surface hydrography as it acts in an interplay with Arctic ice cover and affects the climate and weather conditions in the Arctic and in the midlatitudes [Bingyi and Jia, 2002; Petoukhov and Semenov, 2010]. In addition to climatology, also biological production and marine species distributions are affected by surface hydrography [Fetzer et al., 2002; Moore and Huntington, 2008]. Here we study the spatiotemporal variation of SST and SSS in the Kara Sea from 1980 until 2000. We use in situ observations and hierarchical statistical models to study the impact of climatic (Arctic oscillation index) and seasonal (river discharge and ice concentration) environmental covariates on the surface hydrography of the Kara Sea.

The Kara Sea is one of the shelf seas surrounding the Arctic Basin (Figure 1). Its hydrography is characterized by freshwater inflow from the continental rivers and saline water inflow from the Barents Sea [Pavlov and Pfirman, 1995], which together with seasonal ice conditions cause strong spatiotemporal hydrographic variation [Janout et al., 2015]. During the past few decades in the Kara Sea, ice concentration has declined and SST increased, which have resulted from strengthened heat flux through Atlantic currents and enhanced warm airflow from the midlatitudes [Gerdes, 2003; Polyakov et al., 2007]. Furthermore, increased river discharge has enhanced heat flux especially in the shallow shelf areas [Steele et al., 2008], but has also reduced the general level of SSS [Peterson et al., 2002; Steele and Erdold, 2004]. All these changes affect the biota in a manner that is difficult to predict [Doney et al., 2012]. There is also an increasing interest on the Kara Sea as a shipping route and oil and gas reservoir, which create an environmental threat on the local biota [Ho, 2010; Nevalainen et al., 2016]. Hence, there is a growing interest in the current and past environmental conditions of the Kara Sea and of the Arctic shelf seas in general.

Despite earlier efforts, we lack spatially and temporally accurate information about surface hydrography in the Kara Sea. The circulation studies of Harms and Karcher [1999, 2005] explained seasonal circulation
patterns, and Panteleev et al. [2007] explained interannual variation of the circulation patterns in relation to the Arctic oscillation (AO). Steele and Ermold [2004] and Steele et al. [2008] studied broader-scale trends of SSS and SST, respectively, and investigated temporal trends from the last 100 years in the Pan-Arctic region. However, the studies were carried out in a low spatial resolution and temporal trends were left without an assessment of the total uncertainty.

In this study we use hierarchical statistical models to study the spatiotemporal patterns of the surface hydrography. The novelty of the study is in the spatially varying responses of surface hydrography to environmental covariates. We also estimate the SST and SSS patterns from 1980 to 2000 and quantify the uncertainty related to these estimates.

2. Study Area

The Kara Sea (Figure 1) is a shallow shelf sea (the mean depth is 111 m and 40% of the sea area is shallower than 50 m), where warm Atlantic, cold Arctic, and fresh river inflow mix [Pavlov and Pfirman, 1995]. Sea currents transport relatively warm and salty Atlantic water to the western part of the Kara Sea through the Kara

Figure 1. The Kara Sea and its division into five subregions in relation to its physical characteristics. The $s_1$ and $s_2$ axis denote the axes of the coordinate system used in the analysis and in Figures 2 and 3. The intersection of these axes denotes the origo for the distances in Figures 2 and 3.
Strait and around the northern tip of Novaya Zemlya (NZ). The northern Kara Sea receives salty Atlantic water from a current heading north along the St. Anna Trough [McClimans et al., 2000; Schauer et al., 2002]. Freshwater is transported to the Kara Sea in larger extent through the rivers Ob and Yenisei, which make 40% of the total river runoff to the Arctic Ocean, and to a lesser extent through the rivers Nizhnanyaya and Pyasina, which are located eastward from the river Yenisei [Ruediger, 2003]. River discharge peaks in May and June and stays low outside of the summer season.

The Kara Sea is annually covered by sea ice. Ice formation starts in September and complete ice cover lasts from November to June. Ice formation creates brines and especially polynyas are prone to continuing ice formation and are a source of saline water [Martin and Cavalieri, 1989; Pavlov and Pfirman, 1995]. Wind and ice motion create horizontal stress on sea surface, which controls freshwater distribution and export from the Kara Sea [Panteleev et al., 2007].

Sea surface temperature stays below 4°C throughout the year, except in estuaries where it may peak at 9°C. Summer temperatures decrease in the northeast direction but in winter surface water is constantly at the freezing point as the whole sea is covered by ice [Pavlov and Pfirman, 1995]. Salinity increases from east to west with the strengthening impact of the Atlantic water. Ice formation, ice export, and river discharge control local-scale variation of salinity [Harms and Karcher, 1999]. The location of freshwater dominance is controlled by wind conditions [Johnson et al., 1997].

The Kara Sea is divided into five regions with different physical characteristics (Figure 1). The southwestern corner (Region 1) has the highest annual temperature and is affected by Atlantic current through the Kara Strait [Pavlov and Pfirman, 1995]. Eastern coast of NZ (Region 2) is affected by strong ice formation and oceanic current flowing to southwest [Pfirman, 1995]. The region in front of the river estuaries (Region 3) is affected by annual pulses of freshwater [Harms and Karcher, 2005]. The northeastern coast (Region 4) is an outflow path for freshwater export to the Laptev Sea [Janout et al., 2015]. The northern Kara Sea (Region 5) is influenced by Atlantic currents and flow from the southern Kara Sea [Pavlov and Pfirman, 1995].

3. Materials and Methods

3.1. Data

The data comprises SST and SSS measurements, bathymetry, AO, discharge of the river Yenisei, and sea ice concentration from 1980 to 2000; the time period spanned by all the data sources. Bathymetry varies in space, AO, and discharge in time, and ice concentration in space and time. The results are presented and the hydrographic estimates generated in a lattice grid with 5 km cell size. This resolution was chosen as a compromise between computational effort and good visual mapping.

Hydrographic measurements of SST and SSS were compiled from World Ocean Data Base (WOD) 2013, which integrates quality controlled oceanographic data sets from several sources [Boyer et al., 2013; NOAA, Ocean Climate Laboratory, 2015] (see also Figure S1 in Supporting Information). We did not include remote sensed satellite data into the analysis. Most of the available high-resolution satellite data start from the 2000s [Stroh et al., 2015]. The high-resolution OISST data set is available from 1981 but the data set is an interpolation derived by analyzing satellite images and point observations from ships and buoys [Reynolds et al., 2007]. Hence, we would have data point duplicates between OISST data set and the point observations from the WOD. Moreover, satellite images do not detect SST through ice and hence would not have improved the input data in the season of sea ice cover when there is a shortage of point observations.

The environmental variables were selected based on their assumed impact on hydrography and their accessibility. Bathymetric data were derived from the International Bathymetric Chart of the Arctic Oceans (IBCAO v. 3.0) [NOAA, National Geophysical Data Center, 2015], which is generated from ship track data, contour maps, gridded sources, topography, and coastline information [Jakobsson et al., 2012]. The bathymetry data were transformed from the original 500 m to 5 km resolution. Bathymetry is used to describe the effect of shallow shelf area on hydrography. In general, the shallower the water column, the faster it heats up in the summer and affects SST. Furthermore, river discharge has the strongest impact in the shelf area and hence bathymetry is an index also for the effect of the freshwater.

AO describes the 1000 mb height anomaly north from 20°N latitude. AO indicates the polar vortex modulation and contributes to the Arctic climate and oceanic sea level pressure. AO is a surrogate for the
differences in air pressure systems on the northern hemisphere. It indicates the wind system, which regulates
the sea currents in the Arctic [Thompson and Wallace, 1998; Rogers and McHugh, 2002]. The AO data
set is archived by and accessible through Climate Prediction Center of NOAA [NOAA, Climate Prediction
Center, 2015]. We utilized the monthly mean values of AO. Water circulation patterns change in the Kara
Sea according to the prevailing AO regime [Panteleev et al., 2007]. The AO regime used in the literature
describes air pressure conditions which last from a couple of months to decades. We included the monthly
mean AO and the mean AO over the previous winter (from December to March) into the analyses. The for-
eroer explains the monthly scale water circulation patterns and its effects on hydrography and the latter
describes the mean air pressure conditions of the past winter, which affect hydrographic conditions of the
following spring, summer, and autumn through ice transport and formation. Especially in the Kara Sea, AO
is an important factor affecting the summer season sea ice conditions [Rigor et al., 2002; Rigor and Wallace,
2004].

Discharge values were measured from the river Yenisei in a gauge situated 697 km inland from the river
pour point. Of the two major rivers, Yenisei and Ob, we selected Yenisei to represent the continental fresh-
water inflow for its higher discharge and relatively short estuary zone. The rivers Ob and Yenisei vary sea-
sonally in a similar manner. Since our model is correlative and does not describe causal effects of different
discharge pathways from Yenisei or Ob, having two covariates with similar variation would not improve the
model performance. The river discharge data have been utilized by earlier studies [Shiklomanov and Lamners,
2009] and is accessible through R-ArcticNet [R-ArcticNet, 2015]. We counted a mean daily dis-
charge for each month based on the daily discharges. We took the logarithm of the discharge in order to
lower the discharge peak in May and June and to improve the fit of linear response compared to using dis-
charge directly. We tested the effect of time lag on the impact of discharge on hydrography. Based on the
model comparison (see section 3.2 and Appendix B), we selected a 2 months lag to describe the time need-
ed for discharge to affect hydrography.

The monthly mean sea ice concentration data were derived from the brightness temperature data from the
National Snow and Ice Data Center [Cavalieri et al., 1996; NSIDC, 2015]. The ice concentration data were reinter-
polated from 25 km cell size onto a 5 km cell size grid by using ordinary kriging interpolation. We used
spherical semivariogram for modeling the distance decay and infer lag, nugget and sill with ArcMap soft-
ware [ESRI, 2015]. The reinterpolation smoothed rasters but did not incorporate new data to improve the
accuracy.

3.2. Spatiotemporal Modeling With Spatially Varying Coefficient Processes

We built a hierarchical Bayesian spatiotemporal regression model for both end variables (SSS or SST). The
regression model was then used to examine the effects of environmental covariates (bathymetry, ice con-
centration, monthly mean AO, winter mean AO, and log of the river Yenisei discharge) on SST and SSS and
to estimate SST and SSS over the Kara Sea at the 5 km grid cells from 1980 to 2000.

Let \( y(s, t) \) and \( x(s, t) \) denote, respectively, the end variable and the vector of 5 environmental covariates at
location \( s \) (coordinates in km) and time \( t \) (in years). Notice that AO and log river discharge covariates vary in
time but not in space, bathymetry is constant over time but varies in space, and ice concentration varies in
time and space. We followed Gelfand et al. [2003] and used a spatially varying coefficient process independ-
ently for both end variables

\[
y(s, t) = \alpha(s, t) + x(s, t)^T \beta(s) + \varepsilon(s, t),
\]

where \( \alpha(s, t) \) is a spatiotemporally varying intercept, \( \beta(s) = [\beta_1, \beta_2(s), \ldots, \beta_5(s)]^T \) is a \( 5 \times 1 \) vector of one con-
stant and four spatially varying coefficients, and \( \varepsilon(s, t) \) is the i.i.d. zero mean Gaussian observation error
with variance parameter \( \sigma^2 \). Different from the ordinary linear regression, where the linear coefficients are
assumed constant in space (\( \beta_d(s) = \beta_d \) for all \( s \)), the spatially varying coefficient process assumes that the
coefficients are (latent) functions of the spatial coordinates. Hence, the model is an extension to the tradi-
tional linear regression so that we allow the response of SST and SSS along covariates to change in space.
Moreover, the intercept is allowed to vary in space and time leading to spatiotemporally varying intercept
which accounts for temporally and spatially correlated variation that is not explained by the covariates.

We assumed that the coefficients related to ice concentration, monthly mean AO, winter mean AO, and riv-
er Yenisei discharge, respectively \( \beta_2(s), \ldots, \beta_5(s) \), vary in space and the coefficient related to bathymetry,
\( \beta_t \), is constant throughout the study region. We standardized all the covariates to have zero mean and standard deviation of one in order to help the assessment of their relative importance for explaining the data. We standardized also the end variables, y, to have zero mean and standard deviation of one for modeling and retransformed them to the original scale when presenting results.

We followed the Bayesian approach [Gelman et al., 2014] and gave a vague Gaussian prior for the constant coefficient related to the bathymetry, \( \beta_0 \sim N(0, 10) \), and independent Gaussian process (GP) priors [Gelfand et al., 2003] for the spatiotemporally varying intercept and spatially varying regression coefficients, \( \beta_d(s) \), where \( d \in \{2, 3, 4, 5\} \). A GP is a stochastic process that can be used to define distributions over functions. It is defined by a mean and a covariance function so that, e.g., \( \beta_d(s) \sim GP(0, k(s, s')) \) denotes that the coefficients corresponding to the second covariate have a zero mean GP prior with a covariance function \( k_2(s, s') = \text{Cov}(\beta_2(s), \beta_2(s')) \) that describes the correlation between coefficients at locations \( s \) and \( s' \). We assumed that the three coefficient processes are mutually independent and have zero mean GP priors with exponential covariance functions, so that the covariance function of the \( d \)th coefficient process is

\[
k_d(s, s') = \sigma_d^2 \exp\left(-\sqrt{\sum_{x=1}^{2} (x-x')^2/l_d^2}\right).
\]

The variance parameter, \( \sigma_d^2 \), governs the magnitude of the regression coefficients and the length-scale parameters, \( l_d \), govern the autocorrelation length of the GP along the \( x \) and \( y \) coordinates (Figure 1). The correlation between two locations drops below 5% of its maximum when these locations are approximately 3 times the length scale apart. Hence, the exponential covariance function implies continuous functions for the coefficients and that the coefficients at alternative locations are the more correlated the closer they are in space. The coordinate system for the spatial processes was chosen to reflect the main directions of the Kara Sea so that \( x \) axis corresponds roughly to the distance from the continent (see Figure 1).

The spatiotemporally varying intercept is given a GP prior with mean \( \beta_0 \sim N(0, 10) \) and an additive covariance function of three components. The first component is an exponential covariance function of spatial location \( k_{\text{exp}}(s, s') = \sigma_0^2 \exp\left(-\sqrt{\sum_{x=1}^{2} (x-x')^2/l_0^2}\right) \) (similar to the covariance functions of the regression coefficients) and it corresponds to a temporally constant process that describes spatially varying long-term averages over the study period. The second component is a product of exponential covariance function in space and a periodic covariance function in time, \( k_{\text{perST}}((s, t), (s', t')) = \sigma_{\text{perST}}^2 \exp\left(-\sqrt{\sum_{x=1}^{2} (x-x')^2/l_{\text{perST}}} - \frac{|t-t'|}{P_{\text{perST}}^2}\right) \). This component describes seasonally varying spatial patterns in the intercept that recur annually. The last component is a product of exponential covariance function in space and exponential covariance function in time, \( k_{\text{expST}}((s, t), (s', t')) = \sigma_{\text{expST}}^2 \exp\left(-\sqrt{\sum_{x=1}^{2} (x-x')^2/l_{\text{expST}}} + \rac{|t-t'|}{P_{\text{expST}}^2}\right) \), which captures spatiotemporal variation that does not have seasonal pattern. This covariance structure leads to an additive model where each covariance function corresponds to one spatiotemporal process [Rasmussen and Williams, 2006].

After formulating the model, we trained it with the data by applying the Bayes rule and calculating the posterior distribution over the model parameters (e.g., parameters of the covariance functions) and the latent coefficient functions. The trained model was then used to calculate the posterior probability distribution for SSS and SST at all grid cells for each month from 1980 to 2000 (see Supporting Information for maps summarizing these). These posterior distributions were used to calculate areal and temporal averages. All the calculations were conducted with GPy toolbox [Vanhatalo et al., 2013]. See Appendix A for more details on the model and its training as well as the posterior results for the hyperparameters. The models were validated with the posterior predictive checks [Gelman et al., 2014] and leave-one-out cross validation (LOO-CV) [Vehtari and Ojanen, 2012]. See Appendix B for details.

4. Results and Discussion

4.1. The Effects of Covariates on SSS

Both AO indices, discharge of Yenisei, and ice concentration have a significant and spatially varying effect on SSS (Figures 2 and 3) whereas bathymetry does not have any effect (its posterior mean was \(-0.01\) and its 95% credible interval is between \(-0.06\) and \(0.03\)). River discharge has a negative impact on SSS over
most of the study area so that the magnitude of the response decreases to west and north. This is reasonable since river discharge increases the freshwater content in the central Kara Sea. Only the southwestern corner has a positive response to river discharge which may be due to an increased transport of saline water from the Barents Sea to the Kara Sea during the positive monthly AO, which is coupled with high river discharge [Panteleev et al., 2007]. The effect of ice concentration on SSS is mostly positive in the Kara Sea and strongest in the central part. Ice formation increases SSS and correlates negatively with discharge of freshwater. During winter the discharge of freshwater does not compensate the year-round operating saline Atlantic water inflow which leads to an increase of SSS [Harms and Karcher, 1999]. Ice concentration has a negative effect on SSS only in the southwestern corner, which may be related to some local hydrographic process that we could not take into account with the model parameters.

The coefficients of both AO indices vary spatially (Figures 2 and 3). Monthly mean AO has negative coefficients on the shelf area and positive coefficients in the northwestern Kara Sea. During negative phase of monthly mean AO the easterly winds prevail, which creates a weaker circulation pattern: saline currents
from the Barents Sea and from the Arctic Ocean entering the northern Kara Sea are depressed and the saline current along the eastern coast of NZ is reduced. Also, the freshwater current out from the Kara Sea along the northeastern coast declines [Janout et al., 2015; Panteleev et al., 2007]. These are called blocking conditions because of the reduced levels of saline water inflow and freshwater export. The positive coefficients of monthly mean AO in the northwestern Kara Sea indicate an intensified inflow of saline Atlantic water during positive monthly mean AO [Panteleev et al., 2007]. In the northeastern coast negative coefficients in turn reflect the strengthened freshwater export to the Laptev Sea [Harms and Karcher, 2005]. The southeastern Kara Sea has a negative response to monthly mean AO. During negative monthly mean AO northeasterly wind pattern pushes water from the southcentral Kara Sea to the southern corner. According to the circulation study of Panteleev et al. [2007], such a mechanism would cause relatively saline water to accumulate there.

As the monthly mean AO index is a surrogate for wind pattern, the mean winter AO indicates the impact of the past winter meteorological conditions on the hydrography of the following spring, summer, and
autumn. AO index varies more strongly in winter than in summer and affects the ice transport and the hydrography of the following year [Rigor et al., 2002]. The two AO indices impact in different temporal scales. SSS has a negative response to winter mean AO in most parts of the Kara Sea, as only the southern part has a neutral response. During a negative winter AO the ice export decreases and the following spring and summer are characterized by higher sea ice concentration and volume, and lower SST and air surface temperatures. Furthermore, there is an intensified sea ice formation in the following autumn leading to higher ice volume [Rigor et al., 2002]. Stronger ice formation would increase SSS and negative winter mean AO would affect positively SSS. In the southern Kara Sea ice conditions may be less affected by ice transport to the Arctic Ocean and the response to the winter mean AO is neutral. The winter mean AO describes, thus, the severity of ice conditions in the Kara Sea which cannot be fully captured by ice concentration that does not directly account for ice volume. In addition to ice conditions, the winter mean AO indicates also the water circulation patterns in the Arctic Ocean, which affects the northern Kara Sea [Morison et al., 2000, 2006]. The negative response of SSS in the northern Kara Sea results from a decreased freshwater content in the Eurasian basin under an anticyclonic circulation pattern (negative AO) [Morison et al., 2012]. The change in the response from neutral to strongly negative is along the shelf break where the Arctic circulation pattern starts to affect hydrography [Morison et al., 2012].

It should be noted that the model does not uniquely indicate causality and hence the coefficients may represent mere correlation with causal drivers. An assumption of causality between river discharge and SSS in the central Kara Sea seems justified. The coefficients are weaker further away from the estuaries and outside of the shelf region. The negative responses continue along the northeastern coast, which is a pathway of freshwater export from the Kara Sea [Harms and Karcher, 2005; Janout et al., 2015]. River discharge is a seasonal process which follows an annual pattern peaking in May and June (Figure S4 in Supporting Information). Still, some part of the negative response may be related to ice melt driven decrease of SSS in summer after the discharge peak. The southern corner of NZ has a controversial response to discharge and sea ice which may be caused by a local factor not included in the model. This could be, for example, a local-scale saline current from south which is intensified during the years of positive AO regime [Panteleev et al., 2007].

The spatiotemporal variation of SSS is explained by variables operating in different spatial scales and having causal and correlative effects. As river discharge and ice concentration can be interpreted to have a causal relationship on the hydrography, both AO indices correlate with climatological variation and control circulation patterns which have a causal relationship with the hydrography [Panteleev et al., 2007; Rigor et al., 2002]. The impact of discharge and ice concentration is mostly unidirectional and only the magnitude of the impact varies in the Kara Sea. AO indices on the contrary correlate with climatic factors that drive circulation patterns in multiple scales. These circulation patterns may have different effects on the surface hydrography in different regions in the Kara Sea. Positive AO regime drives an eastward wind pattern and saline water penetration from the Atlantic, which increases the average SSS in the Kara Sea [Harms and Karcher, 1999; Panteleev et al., 2007]. On the other hand, the positive AO regime is characterized by higher precipitation on the continent and higher discharge in continental rivers [Peterson et al., 2002]. Hence, positive mean monthly AO correlates with lower SSS and higher freshwater export along the northeastern coast. The responses of SST and SSS are linked to the effects of AO on the water circulation, which is highlighted with the strongly negative responses of SSS to mean winter AO in the northern Kara Sea [Morison et al., 2000, 2012].

The mean winter AO correlates with the summer ice concentration [Rigor et al., 2002]. As it indicates the mean conditions of the winter, it sums up a larger-scale atmospheric pressure conditions than the monthly mean AO. The different coefficients in the northern and southern Kara Sea indicate regionally different effects on SSS. Accordingly winter mean AO explains a part of variation of SSS which is not explained by ice concentration only but is also related to ice volume. The high volume of sea ice transport from the northern Kara Sea, higher surface air temperature and less intensive ice formation decrease SSS. As the coefficients of monthly mean AO varied in a west-east direction, the coefficients of winter mean AO varied in a north-south direction. There are different drivers of SSS correlating with different AO indices.

The monthly average ice concentration of the Kara Sea is significantly correlated with mean monthly AO (corr. = -0.5301). There was no significant correlation between monthly mean AO and monthly mean daily river discharge (corr. = 0.0303). We found a stronger correlation between the annual median AO and the annual sum of river discharge (corr. = 0.3506). This was expected as the multiannual variation of AO affects
the precipitation and river discharge in high latitudes [Peterson et al., 2002]. The mean winter AO and annual sum of discharge were even more strongly correlated (corr. = 0.3961). The winter mean AO catches more of the long-term variation of AO than the monthly mean index. There was also a correlation between annual mean ice concentration and winter mean AO (corr. = −0.4169).

4.2. The Effects of Covariates on SST

The 95% posterior credible interval of the coefficient of bathymetry is between 0.07 and −0.01 and its posterior mean was 0.03. Hence, bathymetry has a positive effect on SST as assumed. SST increases in the shelf area compared to the deep sea. The response of SST to ice concentration is constantly negative, since ice concentration is highly correlated with winter season and cold surface air temperatures (Figures 2 and 3). The most negative coefficients are concentrated in the central Kara Sea, where the uprising of warm Atlantic water prevents the formation of constant ice cover and creates polynyas [Anselme, 1988]. Under these conditions low ice cover is an index for warm water uprising and an increase of SST.

The response of SST to river discharge varies spatially. Discharge has a positive effect on SST in the southern Kara Sea and a negative effect in the northern Kara Sea and along the northeastern coast line. Discharged water induces water column stratification which accelerates surface cooling and ice formation in autumn. The northeastern coast is affected by freshwater more than the southern Kara Sea, where stratification is weaker [Harms and Karcher, 1999]. The positive responses in the southern Kara Sea may be related to seasonality.

In general, the effects of wind patterns and water circulation on SST are less studied than on SSS. SST in the south corner of the Kara Sea has a positive response to monthly mean AO. Positive monthly mean AO is characterized by inflow of relatively warm Atlantic water through the Kara Strait, which increases SST in the southern Kara Sea [Zhang et al., 1998]. The cyclonic circulation pattern of the Kara Sea during positive monthly mean AO may enhance SST in the estuary region. The negative responses to the monthly mean AO in the northern parts and close to the NZ in the southern Kara Sea are related to warm water accumulation due to blocking conditions during negative AO, which increases SST [Panteleev et al., 2007].

The spatial pattern of the coefficients of the winter mean AO differs from that of the monthly mean AO. They are distributed in a decreasing pattern from south to north. The southern Kara Sea has a positive response to winter mean AO, which is due to the lower ice concentration, increased surface air temperature, and increased Atlantic water inflow from the Barents Sea [Rigor et al., 2002; Morison et al., 2012]. As discussed earlier, the positive winter mean AO has an effect on the cyclonic water circulation pattern, which increases warm water inflow from the Barents Sea to the Kara Sea [Morison et al., 2012]. The northern Kara Sea on the other hand is dominated by negative response which is in conflict with the study of Rigor et al. [2002]. The higher ice transport and lower ice concentration should be coupled with increased SST throughout the Kara Sea. The negative responses may be related to Ekman pumping, which forces the warm surface water to diverge from the central and northern Kara Sea [Janout et al., 2015].

4.3. Temporal Changes in SSS and SST

The Kara Sea consists of hydrographically different regions. There is more inter annual hydrographic variability in the Regions 1–3 located on the shelf area than in the Regions 4 and 5 in the deep-sea area, which agrees with the finding of Simstich et al. [2005] (Figure 5). The shelf area and the western Kara Sea are affected by inflows of fresh and saline water, respectively, as the northern part has more stable hydrographic conditions.

The uncertainties related to hydrographic estimates make it difficult to assess temporal trends with high certainty (Figures 4 and 5). This has been acknowledged in earlier studies, which have struggled with sparse data sets [Harms and Karcher, 1999; Panteleev et al., 2007; Simstich et al., 2005]. The uncertainty in our model estimates depends on the spatial and temporal availability of observations so that uncertainty is the lowest at locations and times that are near the observation points and hence, summer seasons have lower uncertainties than winter seasons (Figure 4). During winter and early summer the uncertainty is the highest and during this period of time the Gaussian distribution for the SST is slightly suboptimal since the posterior distribution gives nonnegligible probability mass to temperatures below the physical limit of approximately −1.85°C (seen as 95% interval dropping below this limit in Figure 4).
We divided the study period into subperiods of 5–6 years and compared them pairwise. There is a probable increase of SST from the first study period (1980–1984) to the last one (1995–2000) in the month of August. The expected increase in August is 0.7°C with 95% probability interval between 0.2 and 1.2°C. This supports the results of Steele et al. [2008], who showed a significant increase of SST in the whole Arctic and in the Kara Sea in summer (July–September) in years 1965–1995. The year 1995 stands out with a summer season SST that was almost 2°C higher than the study period average in many subregions (Figure 5). The anomaly of the whole Kara Sea is over 1.5°C with a probability of 0.95. This extreme SST may result from high air surface temperatures in the Kara Sea region in and around the year 1995 [Lawrimore et al., 2011].

The probability that salinity has decreased between the first and the last subperiod in August is 0.76. Large-scale study of Steele and Ermold [2004] has shown a statistically significant freshening of the Kara Sea between 1965 and 1995. Our study supports the results of Simstich et al. [2005] that suggested a gradual freshening of the whole water column between years 1996 and 2002 in the central Kara Sea. Our results show a decreasing trend of SSS after year 1995, which assumingly correlates with the freshening of the whole water column (Figure 4).

The temporal changes of the average hydrographic conditions of the Kara Sea are probably related to climatic conditions. Both AO variables are straight indices of the climatic variation as ice concentration and discharge are affected by the climate. The changes of AO explain the variation of SSS and SST in multidecadal scales [Steele and Ermold, 2004; Steele et al., 2008] and in shorter study periods [Panteleev et al., 2007; Rigor et al., 2002; Simstich et al., 2005]. Our study period consists of negative (1980–1988) and positive (1989–1995) AO regimes, when we compared the annual mean winter AO. The last period 1996–2000 shows strong annual fluctuations. With a probability of 0.8 there is an increase of the mean SST in month.
August between negative and positive AO regimes. During the same time SSS has decreased with a probability of 0.9. These results support the study of Rigor et al. [2002] about the effects of winter season AO on hydrography. In a multiyear scale the hydrographic variation follows the changes of AO with a time lag of a few years [Morison et al., 2006]. The high air surface temperatures and peaks of SST and SSS in 1995 are possibly caused by extremely high AO around 1989 and 1990, which increased the warm and saline water content in the Arctic Basin 4–5 years later [Morison et al., 2006; Steele and Boyd, 1998].

Our results agree with previous studies about the importance of AO on SST and SSS, but we want to highlight the spatially varying impact of AO on hydrography. The different regions of the Kara Sea respond in different ways to winter mean AO and monthly mean AO. Moreover, the effects come through changes in ice concentration and movements, water circulation patterns, and continental river discharge [Pavlov and Pfriman, 1995].

5. Conclusions

On the level of the whole Arctic there have been major changes in environmental conditions during the past decades. These affect also SST and SSS which are important factors for species communities and biological production. Hence, more detailed information on SST and SSS from the Arctic shelf sea is of practical information when, for example, predicting effects of climate change on marine biota. Our study concentrated on the Kara Sea. We analyzed the effect of key environmental variables on SST and SSS with

![Figure 5. Areal mean SST and SSS of the month August of the Kara Sea subregions. Black line shows the posterior mean and gray lines the central 95% credible interval.](image-url)
hydrographic in situ measurements and model-based statistical inference. Our spatially varying regression model allows the regression coefficients related to environmental covariates to vary spatially, which created an analytic way to study areal responses to environmental variables. However, the modeling approach used here is based on correlative relationships between hydrographic observations and spatially or spatiotemporally varying covariates but does not describe physical dynamics. This data-driven methodology benefits from extensive use of available data but lacks conclusive causal and physical explanatory skills. Hence, the causal reasons behind the relationships between surface hydrography and environmental covariates need to be interpreted with care. However, according to model validation, the model explained well the spatiotemporal variation of surface hydrography.

AO variables and river discharge had spatially varying impacts on hydrography whereas sea ice concentration had spatially constant impact. Our results agree in general with previous studies about the spatial and temporal variation of the Kara Sea surface hydrography. The effects of AO on hydrography come through wind patterns and sea currents which control the spread of freshwater and the volume of Atlantic water entering the Kara Sea. AO index is related to climatic conditions in the Arctic and affects ice concentration and river discharge [Peterson et al., 2002; Rigor et al., 2002]. The winter mean AO affects surface hydrography through ice transport and water circulation, as suggested by the comparison of mean SST and SSS levels between negative and positive AO regimes. In this study both AO indices correlated with ice concentration and the mean winter AO correlated with the annual sum of river discharge. Despite that, we found different responses of hydrography to the covariates. Ice concentration has the least spatially varying impact on SST and SSS as monthly mean AO has a west to east varying impact and winter mean AO a south to north varying impact on SSS and SST. Discharge has the biggest effect on SSS close to the estuaries. We observed a positive trend in SST when calculating the probability for an increase in average SST from 1980–1984 to 1996–2000. Earlier studies have proved an increase of SST between years 1965 and 1995. SSS had a decreasing temporal trend that would speak for freshening of the Kara Sea. As shown above, long and short-term changes of surface hydrography are shadowed by seasonality and annual fluctuations. In order to deepen our understanding of the environmental changes in the Arctic shelf areas, we need to take better into account the regional and local-scale patterns. The southern Kara Sea varies seasonally and is mostly affected by Atlantic currents. The northern Kara Sea is more tightly related to changes in ice formation and transport and in patterns of Arctic circulation. The continental river discharge has a local impact close to the estuaries. These regional differences affect how marine ecosystem is affected by the warming climate and by the changes in Arctic ice extent and volume.

This is the first study to present spatially and temporally detailed information about annual and seasonal hydrographic variation in the Kara Sea and to display the spatially varying effects of environmental covariates. Our methodology is statistical and does not consider water circulation patterns as physical hydrographic models do. The benefits of the methodology are the extensive use of in situ measurements and the quantification of uncertainties. The study provides new, data-driven, knowledge about the spatial behavior of environmental drivers of surface hydrography in an Arctic shelf sea. Furthermore, we can make spatially and temporally high-resolution estimates on SST and SSS, which benefit other geophysical and ecological studies. With estimated map layers we can more accurately study the spatial dynamics of water circulation, ecosystems, and species distributions.

Appendix A

Mutually independent GP priors for the spatially varying coefficients and an independent Gaussian prior for the constant coefficient (corresponding to the bathymetry) imply that the linear term of the spatially varying coefficient model (1) has a GP prior with additive covariance function

\[ x(s, t) \sim GP(0, k_s((x, s, t), (x', s', t'))) \]

(A1)

where \( k_s((x, s, t), (x', s', t')) = 10 + k_{x_1}(s, t)x_1(s', t') \) and \( \sum_{d=2}^{4} k_d(s, s')x_d(s, t)x_d(s', t') \). Similarly, the spatiotemporally varying intercept has a GP prior

\[ x(s, t) \sim GP(0, k_i((s, t), (s', t'))) \]

(A2)

where \( k_i((s, t), (s', t')) = 10 + k_{exp}(s, s') + k_{perST}((s, t), (s', t')) + k_{expST}((s, t), (s', t')) \). Hence, the spatially varying coefficient model (1) can be rewritten as a hierarchical Bayesian model.
\[ y(s, t)|f(s, t) \sim N(f(s, t), \sigma_f^2), \quad (A3) \]
\[ f(s, t)|\theta \sim GP(0, k_c(s, t), (s', t')) + k_p((x, s, t), (x', s', t')), \quad (A4) \]
\[ \theta \sim p(\theta), \quad (A5) \]

where \( f(s, t) \) denotes a latent process corresponding to the true SSS or SST at location \( s \) and time \( t \) and \( p(\theta) \) denotes the prior distribution for the vector of hyperparameters, \( \theta \), that collects all the parameters of the covariance functions \( k_c(\cdot, \cdot) \) and \( k_p(\cdot, \cdot) \) (that is, all the length-scale and variance parameters) and the noise variance \( \sigma_n^2 \). We gave weakly informative half Student-t priors [Gelman, 2006] for all length-scale parameters and priors for the variance parameters of covariance functions. The observation error variance was given a log-uniform prior.

By definition, a GP prior implies that any finite number of latent variables has a multivariate Gaussian prior distribution. Hence, a vector of latent variables \( \mathbf{f} \) corresponding to the vector of observations of SSS or SST, \( \mathbf{y} \), has a Gaussian prior \( f(\mathbf{y}|\mathbf{f}, \sigma_f^2)N(\mathbf{f}|0, \mathbf{K}) \), where the entries of the covariance matrix are given by the covariance function \( \mathbf{K}_{ij} = k_c((s_i, t_i), (s_j, t_j)) + k_p((x_i, s_i, t_i), (x_j, s_j, t_j)) \).

We optimize the hyperparameters to their Maximum a posterior (MAP) estimate

\[ \hat{\theta} = \arg\max_{\theta} p(\mathbf{y}|\theta)p(\theta), \quad (A6) \]

where \( p(\mathbf{y}|\theta) = {\prod}_j N(\mathbf{y}_j|\mathbf{f}_j, \sigma_f^2)N(\mathbf{f}|0, \mathbf{K})df = N(\mathbf{y}|0, \mathbf{K} + \sigma_f^2\mathbf{I}) \) is the marginal likelihood of the hyperparameters. The MAP estimate was located with scaled-conjugate gradient optimization in GPstuff [Vanhatalo et al., 2013]. The MAP estimates for SST and SSS models are summarized in Tables A1 and A2.

Let us denote by \( \mathbf{f} \) a vector of latent variables corresponding to every grid cell and every month from 1980 to 2000. After finding the MAP estimate for the hyperparameters, we calculated the (conditional) posterior predictive distribution of \( \mathbf{f} \) which is again Gaussian \( \mathbf{f}|\hat{\theta}, \mathbf{y} \sim N(\hat{\mathbf{m}}, \hat{\mathbf{K}}) \). See, e.g., Rasmussen and Williams [2006] for details on how to calculate the posterior mean, \( \hat{\mathbf{m}} \), and covariance, \( \hat{\mathbf{K}} \). We can similarly calculate the (conditional) posterior predictive distribution for each of the components of the latent process. For example, we can calculate the posterior predictive distribution of the second coefficient \( \beta_2|\hat{\theta}, \mathbf{y} \sim N(\hat{\mu}_{\beta_2}, \hat{\sigma}_{\beta_2}^2) \). See Rasmussen and Williams [2006] and Gelfand et al. [2003] for details.

### Table A1. The Maximum a Posterior (MAP) Estimate of the Hyperparameters in the SST Models

<table>
<thead>
<tr>
<th>Model Component</th>
<th>Hyperparameter</th>
<th>MAP Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatially varying coefficient of ice concentration</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 2.2 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td>Spatially varying coefficient of monthly mean AO</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 1.7 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 5.3 \times 10^{2} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td>Spatially varying coefficient of winter mean AO</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 1.7 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td>Spatially varying coefficient of log of the river Jenisei discharge</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 2.3 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 1.0 \times 10^{0} ) km</td>
</tr>
<tr>
<td>Temporally constant intercept term</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 0.1 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 2.2 \times 10^{2} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 4.8 \times 10^{2} ) km</td>
</tr>
<tr>
<td>Unseasonal spatiotemporal term</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 0.2 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 1.2 \times 10^{2} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 1.5 \times 10^{2} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( t ): ( l_{\text{per}} )</td>
<td>( 7.2 \times 10^{-2} )</td>
</tr>
<tr>
<td>Periodic, seasonal intercept term</td>
<td>Variance: ( \sigma_n^2 )</td>
<td>( 0.2 \times 10^{-1} )</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_1 ): ( l_{\text{per}} )</td>
<td>( 4.1 \times 10^{2} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( s_2 ): ( l_{\text{per}} )</td>
<td>( 6.9 \times 10^{1} ) km</td>
</tr>
<tr>
<td></td>
<td>Length-scale along ( t ): ( l_{\text{per}} )</td>
<td>( 0.7 \times 10^{-1} )</td>
</tr>
<tr>
<td>Observation error (likelihood)</td>
<td>Noise variance, ( \sigma_n^2 )</td>
<td>( 2.7 \times 10^{-2} )</td>
</tr>
</tbody>
</table>
we are interested on models predictive performance at unseen locations and times, we conducted the posterior predictive check with replicate data. As the original data, we did not find significant differences between the two. We checked this first by simulating replicate measurements from the posterior predictive distribution at the same time and spatial locations as the training data and comparing the samples of $y$ to the measured data. We compared the replicate data with original data by sample histogram and a scatterplot between the posterior mean of replicate data and the original data. We did not find significant differences between the two.

### Appendix B

The purpose of the model assessment is to check the reliability of a model in order to recognize whether possible model’s deficiencies have noticeable effect on the substantive inferences [Gelman et al., 2013]. Model comparison is used to assess the relative performance of alternative models and to choose the best possible model’s deficiencies have noticeable effect on the substantive inferences [Gelman et al., 2013] and posterior predictive comparison [Vehtari and Ojanen, 2012] for model assessment and for choosing the best time lag for the River Yenisei discharge. Notice, here we use the term predict to refer to building probability distribution for an unseen observation, $y$, conditional to the training data, $\{y, x\}$, that is $p(y|x; y, x)$, whether that unseen observation was historical or in the future.

The posterior predictive check builds upon assumption that in order for a model to work well, predictive data generated by a trained model should be similar to the measured data. We checked this first by simulating replicate measurements from the posterior predictive distribution at the same time and spatial locations as the training data and comparing the samples of $y$ to the measured data. We compared the replicate data with original data by sample histogram and a scatterplot between the posterior mean of replicate data and the original data. We did not find significant differences between the two.

The posterior predictive check with replicate data can be seen as a first-order quality check for a model. As we are interested on models predictive performance at unseen locations and times, we conducted the

### Table A2. The Maximum a Posterior (MAP) Estimate of the Hyperparameters in the SSS Models

<table>
<thead>
<tr>
<th>Model Component</th>
<th>Hyperparameter</th>
<th>MAP Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatially varying coefficient of ice concentration</td>
<td>Variance, $\sigma^2_{\text{ice}}$</td>
<td>$9.6 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>$1.0 \times 10^2$ km</td>
</tr>
<tr>
<td>Spatially varying coefficient of monthly mean AO</td>
<td>Variance, $\sigma^2_{\text{perST}}$</td>
<td>$1.5 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>$1.0 \times 10^1$ km</td>
</tr>
<tr>
<td>Spatially varying coefficient of winter mean AO</td>
<td>Variance, $\sigma^2_{\text{perST}}$</td>
<td>$3.9 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>$1.0 \times 10^1$ km</td>
</tr>
<tr>
<td>Spatially varying coefficient of log of the river Jenisei discharge</td>
<td>Variance, $\sigma^2_{\text{perST}}$</td>
<td>$8.0 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>$1.0 \times 10^1$ km</td>
</tr>
<tr>
<td>Temporally constant intercept term</td>
<td>Variance, $\sigma^2_{\text{tempST}}$</td>
<td>$5.5 \times 10^{-1}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>$4.5 \times 10^0$ km</td>
</tr>
<tr>
<td>Unseasonal spatiotemporal term</td>
<td>Variance, $\sigma^2_{\text{unST}}$</td>
<td>$2.0 \times 10^{-1}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>$1.8 \times 10^0$ km</td>
</tr>
<tr>
<td>Periodic, seasonal intercept term</td>
<td>Variance, $\sigma^2_{\text{perST}}$</td>
<td>$9.7 \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td>Length-scale along $s_{1:45000}$</td>
<td>16 km</td>
</tr>
<tr>
<td>Observation error (likelihood)</td>
<td>Noise variance, $\sigma_{\text{noise}}^2$</td>
<td>$7.7 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

After solving the posterior predictive distribution for the latent variables, we can calculate areal and temporal averages. For example, let $\mathbf{f}$ denote the latent variables corresponding to SST and let $\mathbf{w}$ be a vector of the same length as $\mathbf{f}$ so that those elements of $\mathbf{w}$ are $1/N$ that correspond to the $N$ July 2000 grid cells over the Kara Sea in $\mathbf{f}$ and all other elements in $\mathbf{w}$ are zero. Then the July 2000 average SST over the Kara Sea has a Gaussian distribution, $\mathbf{w} \sim N(\mathbf{w}', \mathbf{w}'K\mathbf{w}')$. Similarly, if we want to calculate the difference between July 2000 and 1980 average SSTs, we can form a vector $\mathbf{w}_2$ where all the elements corresponding to July 2000 have value $1/N$ and all the elements corresponding to July 1980 have a value $-1/N$; all other elements of $\mathbf{w}_2$ are zero. Then the difference between July average of 2000 and 1980 has a Gaussian distribution $\mathbf{w}_2 \mathbf{f} \sim N(\mathbf{w}_2 \mathbf{w}', \mathbf{w}_2K\mathbf{w}_2')$. In practice, $\mathbf{K}$ is too large (it has $(N\times21\times12)^2$ elements where $N \approx 45000$) to be handled as a full matrix. Hence, we never formed full $\mathbf{K}$ but conducted all the calculations in pieces.
model validation also using leave-one-out cross validation (LOO-CV) [Vehtari and Ojanen, 2012]. The basic idea of LOO-CV is to leave each data point at time out of the training data, train the model with the rest of the data and use the trained model to form a LOO-CV predictive distribution at the left out data point; that is, $p(y_n|x_{-n}, y_{-n}, x_{-n})$, where $y_n$ and $x_n$ denote all the end variables and covariates excluding the $n$th data point. We compared the LOO-CV predictive mean to the left out data points with scatterplot and calculated the root mean squared error (RMSE) between the LOO-CV predictive mean and the left out data points. The RMSE of SST model was 0.2288 and that of the SSS model was 0.3314. Whereas the sample standard deviation of data was one (since we standardized $y$ to have standard deviation of one before modeling, see section 3.2), hence, the model explains 77% of the variation in SST and 67% in SSS. We used the LOO-CV RMSE also to choose the best time lag for the River Yenisei discharge from alternatives of 0, 1, 2, and 3 months.

Acknowledgments

The work was funded by the Academy of Finland (grants 292985, 266349, and 304531), the Lloyd’s Register Foundation and Research Funds of the University of Helsinki (decision No. 465/51-2014). The Lloyd’s Register Foundation supports the advancement of engineering-related education, and funds research and development that enhances the safety of life at sea, on land, and in the air. We thank also the two anonymous reviewers of this manuscript who helped to significantly improve our manuscript. The ice concentration data were generated by National Snow and Ice Data Center (NSIDC, http://nsidc.org/data/). We used the data set “Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1” (Data Set ID: NSIDC-0051), which is accessible through Polaris web service (http://nsidc.org/data/polaris/). The Arctic Oscillation data were generated by NOAA Climate Prediction Center (http://www.ncdc.noaa.gov/). We utilized the data set “Monthly mean AO index since January 1950” (http://www.cpc.ncep.noaa.gov/products/prep/C0/C1/CWlink/daily_ao_index/ao.shtml). The bathymetry data set, International Bathymetric Chart of the Arctic Ocean (IBCAO v. 3.0), is generated by several data sets. The discharge data for large Siberian rivers, which is accessible through R-ArcticNET v. 4.0 (http://www.r-arcticnet.sr.unh.edu/ObservedAndNaturalizedDischarge-Website/). We used the observed daily discharge values. Supporting Information holds SST and SSS map layers, which were generated with the studied methodology and with the utilized data sets.

References

Ho, J. (2010), The implications of Arctic sea ice decline on shipping, Mar. Policy, 34, 713–715.
NOAA, Climate Prediction Center (2015), Daily AO Index, College Park, Md. [Available at http://www.cpc.ncep.noaa.gov/products/predic t/ CWlink/daily_ao_index/ao.shtml].
NOAA, National Geophysical Data Center (2015), International Bathymetric Chart of the Arctic Ocean, Boulder, Colo. [Available at http:// www.ngdc.noaa.gov/mgg/bathymetry/arctic/]


Polyakov, I., et al. (2007), Observational program tracks Arctic Ocean transition to a warmer state, Eos Trans. AGU, 88(40), 398–399.

R-ArcticNET (2015), Observed and Naturalized Discharge Data for Large Siberian Rivers, Durham, N. H. [Available at http://www.r-arcticnet.sr.unh.edu/ObservedAndNaturalizedDischarge-Website/]


