Extraction of Sea Temperature in the Barents Sea by a Scale Space Multiresolution Method – Prospects for Atlantic Salmon

LEENA PASANEN1*, PÄIVI LAUKKANEN-NEVALA2, ILKKA LAUNONEN3, SERGEY PRUSOV4, LASSE HOLMSTRÖM5, EERO NIEMELÄ6, JAAKKO ERKINARO7

1 Department of Mathematical Sciences, University of Oulu, Pentti Kaiteran katu 1, PO Box 3000, FI-90014, University of Oulu, Finland. E-mail: Leena.Pasanen@oulu.fi
2 Natural Resources Institute Finland, PO Box 413, FI-90014 Oulu, Finland. E-mail: pivi@tilastoneuvonta.fi
3 Department of Mathematical Sciences, University of Oulu, Pentti Kaiteran katu 1, PO Box 3000, FI-90014, University of Oulu, Finland. E-mail: Ilkka.Launonen@oulu.fi
4 Knipovich Polar Research Institute of Marine Fisheries and Oceanography, 183763, Knipovich Street, 6, Murmansk, Russia. Phone: 815-7-25-3, E-mail: prusovs@gmail.com
5 Department of Mathematical Sciences, University of Oulu, Pentti Kaiteran katu 1, PO Box 3000, FI-90014, University of Oulu, Finland. E-mail: Lasse.Holmstrom@oulu.fi
6 Natural Resources Institute Finland, PO Box 413, FI-90014 Oulu, Finland. E-mail: eero.niemela@luke.fi
7 Natural Resources Institute Finland, PO Box 413, FI-90014 Oulu, Finland. E-mail: jaakko.erkinaro@luke.fi

*Corresponding author

The work of Leena Pasanen, Ilkka Launonen and Lasse Holmström was supported by the Academy of Finland under Grant 250862
Abstract

Variation of marine temperature at different time scales is a central environmental factor in the life cycle of marine organisms, and may have particular importance for various life stages of anadromous species, e.g. Atlantic salmon. To understand the salient features of temperature variation we employ scale space multiresolution analysis, that uses differences of smooths of a time series to decompose it as a sum of scale-dependent components. The number of resolved components can be determined either automatically or by exploring a map that visualizes the structure of the time series. The statistical credibility of the features of the components is established with Bayesian inference. The method was applied to analyze a marine temperature time series measured from the Barents Sea and its correlation with the abundance of Atlantic salmon in three Barents Sea rivers. Besides the annual seasonal variation and a linear trend, the method revealed mid time-scale (~10 years) and long time-scale (~30 years) variation. The 10-year quasi-cyclical component of the temperature time series appears to be connected with a similar feature in Atlantic salmon abundance. These findings can provide information about the environmental factors affecting seasonal and periodic variation in survival and migrations of Atlantic salmon and other migratory fish.

Keywords

Climate change, anadromous Atlantic salmon, Bayesian inference, time series decomposition, scale space

1. Introduction

The large-scale impacts of climate change on marine ecosystems and their fisheries
productivity have been widely recognized and clearly demonstrated (e.g. Lehodey et al., 2006; Barange et al., 2010; Hoegh-Gulberg and Bruno, 2010). Changes in ocean temperatures and related shifts in other physical conditions have altered marine ecosystems, e.g. plankton communities (Edwards and Richardson, 2004), and also resulted in distribution shifts among fish species (Perry et al., 2005). Furthermore, an overall change in oceanic fauna throughout the north-eastern Atlantic ecosystem has been attributed to climate change, subarctic-subtropical water-mass exchanges and a strong and persistent bottom-up influence of the North Atlantic subpolar gyre (Hátún et al., 2009).

The impacts of changing climate are particularly complicated on anadromous fish species, which have to cope with a variety of habitats and environmental conditions during their life cycle, across long geographic distances and wide temperature and salinity gradients, from fresh water to full sea water (Heino et al, 2016). For example, Atlantic salmon (*Salmo salar* L.) juveniles emigrate from fresh water after having achieved a certain size and physiological stage (called smolts at this life stage), and return to their natal rivers after one or more years of rapid growth and maturation at sea (Aas et al., 2011).

Over the past three decades, Atlantic salmon populations have declined over most of their range, despite marked reductions in fishing pressure (Chaput 2012). The decline has been associated to climate forcing and concurrently increasing sea temperature, pointing towards the marine environment as a key factor dictating the development in salmon survival and stock status (e.g. Todd et al., 2011; Friedland et al., 2014). This is true also with salmon populations in northernmost Europe where abundance indices and marine growth rates are correlated with sea surface temperatures in adjacent sea areas of the Barents Sea and Norwegian Sea, but fewer links have been
detected with the broad-scale climate indices, like the North Atlantic Oscillation (Niemelä et al., 2004). In addition to survival, temperature may influence the oceanic migration patterns and subsequent timing of maturation and return migration (Todd et al., 2011). Environmental factors affect the freshwater entry and run timing of salmon (e.g. Thorstad et al., 2008), and Vähä et al. (2011) demonstrated that different populations of salmon within a single large river system of the Barents Sea basin entered fresh water in a consistent, population-specific order, although the actual timing varied between years, likely reflecting variation in environmental conditions at sea and in the river. Moreover, large scale variation in climate in the Baltic Sea area has induced periodic cycles in abundance and growth of salmon (Huusko and Hyvärinen, 2012). Therefore, understanding the rapidly changing environment and predicting the consequences require appropriate measurements and long-term data sets from the environment and modeling tools for integrating population dynamics, life history stages, corresponding environments and their changes (Lassalle et al., 2008).

Time series typically exhibit features that manifest themselves in different time scales. Fine scale, local details, such as yearly seasonal variation, show in short time scales, while large characteristics, such as changes in climate, appear only over longer time horizons, tending to be more smooth in character. For various life stages of anadromous Atlantic salmon, several of the different time scales in marine temperature variation may have particular importance. For example, the point of sea water entry of juvenile salmon (smolts) and the first months at sea are critical periods (Otero et al., 2014), but occur just once during the salmon life span, which typically varies between 3-6 years, but may cover even 10 or more years for salmon in northern latitude rivers (Niemelä et al., 2006). On the other hand, such salmon may spend up to four or five consecutive years at sea, and the environmental conditions, including their repeated
annual cycle, during the marine phase of their life cycle are crucial for feeding, growth, maturation, and ultimately survival of the animals. Another critical life stage is the return to sea by spent fish that need fast reconditioning after an exhaustive spawning migration and reproduction. For salmon in northern latitudes this may happen 1-3 times during the life time of salmon, often every second year, but sometimes also in consecutive years (Niemelä et al., 2006). Finally, cyclic variation of marine conditions over several years or decades affect several generations of salmon and their marine migrations, and thus the long term resilience and productivity of populations (Todd et al., 2011, Huusko and Hyvärinen, 2012).

In the present study, we analyze a marine temperature time series measured from the Barents Sea and its correlation with the abundance of Atlantic salmon in three Barents Sea rivers. We are interested in uncovering marine temperature features which may affect fish behavior at different time scales. These features are presented by additive components in a time series decomposition. The widely used classical time series decomposition (see e.g. Hyndman and Athanasopoulos, 2014) includes noise and seasonal components as well as a trend that may also contain large scale oscillatory or quasi-oscillatory features. The annual seasonal component of sea temperature and its effect are well known, but we are interested in extracting also other oscillatory or quasi-oscillatory large scale temperature features that may be embedded in the trend component when the classical approach is used. To this end, we propose to use a new statistical method, scale space multiresolution analysis (Pasanen et al. 2013), to resolve the salient features of a time series across multiple time scales. For a temperature time series, the analysis helps us establish the existence of a linear trend during past decades, explore the extent of long-scale temperature oscillations, and find the time intervals during which the warm and cool periods occurred.
The scale space multiresolution analysis uses differences of smooths of a time series to decompose it as a sum of scale-dependent components, the number of which is determined automatically or by visual analysis of the so-called scale-derivative map. As time series are often noisy, Bayesian statistical inference is used to establish the credibility of the features of the extracted structural components. In addition to sea temperature series, the method can be used to study long-term variation in for example salinity, oxygen conditions and chlorophyll content over time.

2. Materials and methods

2.1. Smoothing

A time series can be smoothed for example in order to remove its rough, small-scale features. Perhaps the most commonly used smoothing method is the moving average. Consider a time series \( y = [y_1, \ldots, y_n] \), a smoothing operator \( S_\lambda \) and a smoothed time series \( S_\lambda y \). Here \( \lambda \geq 0 \) is a “smoothing parameter” that controls the amount of smoothing in \( S_\lambda y \). For example, the window length of a moving average is such a smoothing parameter: the wider the window the smoother the result. Other popular smoothing methods include local linear regression and spline regression (Eubank, 1999). We consider here the discrete spline smoother (Green and Silverman, 1993) for which \( S_0 y = y \) and, as \( \lambda \) grows to infinity, the smooth \( S_\lambda y \) becomes the linear regression line of the time series.

Smoothing can reveal time series features that correspond to different time scales. When \( \lambda \) is small, little smoothing is applied and the smooth \( S_\lambda y \) resembles the observed time series \( y \) showing its small-scale features. On the other hand, a larger value of \( \lambda \) smooths out the smallest scale details and reveals only locally average behavior in \( y \). The scale-dependent features of \( y \) can be extracted by considering
differences $S_{\lambda_i}y - S_{\lambda_j}y$ of smooths, where $\lambda_j > \lambda_i$. The difference of smooths captures the features that are present at the smoothing level $\lambda_i$, but will be smoothed out at the higher smoothing level $\lambda_j$. Such smoothing based exploration of time series features is an example of statistical scale space analysis, a methodology that has gained considerable popularity in recent years (Holmström, 2010; Holmström and Pasanen, 2016).

2.2. The decomposition

Using differences of smooths, a time series can be decomposed into additive scale-dependent components as follows. Let $0 = \lambda_1 < \lambda_2 < \cdots < \lambda_L \leq \infty$ be an increasing sequence of smoothing levels. Since $S_{\lambda_1}y = S_0y = y$, a decomposition of a time series $y$ is then given by

$$y = \sum_{j=1}^{L-1} (S_{\lambda_j} - S_{\lambda_{j+1}})y + S_{\lambda_L}y - \bar{y}1 + \bar{y}1 = \sum_{j=1}^{L+1} z_j,$$

where the $z_j$s are the scale-dependent components:

- $z_j = (S_{\lambda_j} - S_{\lambda_{j+1}})y$, $j = 1, \ldots, L - 1$,
- $z_L = S_{\lambda_L}y - \bar{y}1$,
- $z_{L+1} = \bar{y}1$,
- $\bar{y}$ is the mean value of $y$,
- $1$ is a vector of ones.

However, since we are not interested in the mean of the series, the last component $z_{L+1}$ will be omitted from our analyses. Notice that because of the particular smoothing method employed (discrete spline smoothing), $z_L$ is simply the linear trend if $\lambda_L = \infty$.

The choice of the smoothing parameter sequence $0 = \lambda_1 < \lambda_2 < \cdots < \infty$. 


\( \lambda_L \leq \infty \) is crucial for proper extraction of the salient scale-dependent features of a time series. A subjective, judicious choice of this sequence is one possibility. However, we have developed an objective approach where the smoothing parameter sequence is selected by an optimization algorithm as the local minima of the norm of the “scale-derivative” \( \frac{\partial S_{\lambda y}}{\partial \ln(\lambda)} \) (Pasanen et al., 2013). Time series analysis is visualized using a so-called scale-derivative map and, as an alternative to the optimization method, this map can also be used to support a subjective choice of the smoothing parameter sequence. As an example of scale-derivative analysis, the Kola Section sea temperature time series considered in this article and its associated map are displayed in Fig. 1 and Fig. 2, respectively. The color of a pixel in the map indicates the value of the scale-derivative at a given time and smoothing level, deep red corresponding to a large positive value, deep blue to a large negative value, while green indicates a value close to zero. The two uppermost black lines in Fig. 2 correspond to the values of \( \lambda \) produced by the optimization algorithm and the lowest one to the smoothing level determined by visual inspection.

[Figure 1 near here]

[Figure 2 near here]

Positive value of the scale-derivative for a time \( i \) and smoothing level \( \lambda \) indicates that the value of the smooth at time \( i \) increases when the smoothing level \( \lambda \) is increased. Therefore, the value of the smooth at time \( i \) is smaller than its average in a local neighborhood. An analogous interpretation holds for a negative value of the derivative. It follows that the scale-derivative map reveals the scale-dependent components as oscillating bands of colors. For example, there are three oscillating bands in Fig. 2. The bottom, high-frequency band of blue and red represents the annual seasonal component. Since the amplitude of this component is relatively high, the other
two components appear in the milder colors of yellow and cyan. The middle band of cyan and yellow represents a low-frequency, large-scale component with a period of approximately 10-15 years. The topmost band with light yellow in the middle and cyan at the edges represents a negative scale-derivative changing to positive and then back to negative again, indicating a u-shaped time series component. Note that these inferences can be made directly from the smoothed data without any reference to a particular statistical time series model.

If the data contain non-stationary features, straightforward minimization of the scale-derivative norm may not produce a satisfactory smoothing parameter sequence. This may happen, for example, if there is a transient oscillation which is not present for the whole duration of the series. Prior to choosing the sequence, it is therefore important to visually inspect the scale-derivate map because it can reveal such non-stationary features. It may then be necessary to choose the smoothing parameter sequence manually including values roughly at the boundaries between two oscillating bands (Pasanen et al., 2013).

2.3. Bayesian inference

If the observed time series $y$ is noisy, statistical inference may be necessary to establish the statistical credibility of the scale-dependent components $z_j$. In Bayesian statistics, this requires building a posterior model for the underlying time series. We assume an additive noise model for the observed time series,

$$y_i = \mu_i + \epsilon_i,$$

(2)

where $y_i$ is the observed value of the time series at time $i$, $\mu_i$ is the unknown true value, and $\epsilon_i$ is the residual value, which we formally assume to be uncorrelated Gaussian noise, $\epsilon_i \sim N(0, \sigma_i^2)$. 


According to the Bayes’ theorem, the posterior distribution of the unknown true time series $\mu$ conditional on the observed time series $y$ is
\[
p(\mu | y) = \frac{p(y | \mu)p(\mu)}{p(y)} \propto p(y | \mu)p(\mu)
\]
where $p(y | \mu)$ is the likelihood function, $p(\mu)$ is the prior distribution of the true time series underlying the observed data and $p(y)$ is the marginal distribution of $y$.

In Section 2.2, the scale-dependent components $z_j$ were defined by decomposing the observed noisy time series into a sum of differences of smooths. However, we are of course interested in the features of the noiseless, true time series $\mu$ and its scale-dependent components. Since $\mu$ is modeled as a random variable, its scale-dependent components $z_j$ are also random variables and the expansion (1) holds with the observed time series $y$ replaced by $\mu$. The smoothing parameter sequence $\lambda_j$ is obtained from the scale-derivative map of the posterior mean of the true time series.

To find out which features of a scale-dependent component $z_j$ are real and not just artifacts of noise, we infer the credibly positive or negative time intervals for each such component. For this, we need to estimate the posterior distributions $p(z_j | y)$. This can be done by first drawing a large sample from the posterior distribution $p(\mu | y)$ of the true time series $\mu$. For the sea temperature series displayed in Fig. 1 and described below, we used sample size $10^4$. Each sampled time series $\mu$ is then transformed into a difference of smooths, i.e. $z_j = (S_{\lambda_j} - S_{\lambda_{j+1}})\mu$ and the transformed sample is used to approximate the posterior distribution $p(z_j | y)$.

The credibly positive or negative time intervals for each component $z_j$ could be found by finding the time points where the proportion of positive or negative values in the transformed sample exceeds a threshold value, for example 0.95. Such point-wise inference is, however, prone to false-positives, and we therefore use
simultaneous inference over all time points by applying the method of highest point-
wise probabilities (HPW), first described by Erästö and Holmström (2005).

The statistical inference is summarized by plotting the posterior means
\[ E(z|y) \]
of the scale-dependent components and coloring the time interval background
white, black or gray when the component is credibly positive, negative or neither,
respectively. Such a summary is shown in Fig. 3 for the Kola Section sea temperature
time series.

[Figure 3 near here]

As explained in Section 2.2, a positive value of the scale-derivative for a
time \(i\) and smoothing level \(\lambda\) indicates that the value of the smooth at time \(i\) increases
when the smoothing level \(\lambda\) is increased and analogously for a negative scale-
derivative. A positive scale-derivative therefore indicates a possible local minimum in
the time series. Note that since a local minimum corresponds to a negative value in an
extracted component, the scale-derivative and the extracted components then in fact
have opposite signs. This is evident in Fig. 2 and Fig. 3, especially in the case of the u-
shaped component (panel (e) in Fig. 3).

2.4. Analyzing correlation between two time series

When investigating the dependence between two time series, the problem often is that
the type of correlation they exhibit depends on the time scale considered. Thus, a small
scale correlation structure in the data may hamper the detection of large scale
correlation features, and vice versa. In such situations, scale space multiresolution
analysis can help discover correlation structures that manifest themselves in different
time scales.
We assume that the two time series contain interesting features in similar time scales and that our goal is to discover the correlations between such features. To this end, both time series are decomposed into additive scale-dependent components using the same smoothing parameter sequence and the correlation between the corresponding component pairs is analyzed. Scale-derivatives can be used to choose this smoothing parameter sequence (cf. Section 2.2), but a compromise is needed between the smoothing levels suggested on the basis of the two scale-derivative maps. Our approach here is simply to select a common smoothing parameter sequence for the two time series so that the levels $\lambda_j$ are located as nearly as possible between the oscillating bands of blue and red in the scale-derivative maps of each time series. Dependence between each multiresolution component pair is summarized by reporting their Pearson correlation coefficient and displaying the components in the same plot.

2.5. Sea temperature data

Knipovich Polar Research Institute of Marine Fisheries and Oceanography (PINRO) has over the years accumulated a large array of oceanographic data on the Barents Sea and adjacent waters. The longest datasets are available for the Kola Section running along the 33°30'E meridian from 69°30' to 77°00'N (Fig. 4). Observations along the Kola Section started in 1900 and are more frequent (6–15 times a year) than at other sections in the Barents Sea (Boitsov et al., 2012). The Kola Section runs across the Coastal Water having warm temperatures and low salinity (stations 1-3; 69°30’-70°30’N, 33°30E). The northern stations (8 through 10; 73°00’-74°00’N, 33°30E) with highest salinities measured are considered as Atlantic Water. Stations 3-7 (70°30’-72°30’N, 33°30E) are usually associated with the Murman Current, which is the continuation of the North Cape Current entering the Barents Sea through its western boundary.
To study seasonal and interannual variability of temperature, the data are averaged vertically over 4 depth intervals (0-50, 0-200, 50-200 and 150-200 m) and horizontally over 3 parts of the Section. Temperature averaged over the upper 200 m and over stations 3 through 7 (70°30’-72°30’N, 33°30E) produces a well-known and widely used time-series (Yndestad et. al, 2004). The Section has been occupied most frequently since the mid-1950s with often more than 12 surveys each year. In 1970-1980 the annual number of sampling occasions exceeded 15. A considerable decrease in sampling occasions of the section occurred in the early 1990s. Since the 1990s approximately 10 surveys per year along the transect have been undertaken by PINRO (Karsakov, 2009).

The general opinion is that the Kola Section time series can serve as a reliable indicator of climate variability in the whole southern Barents Sea (coastal and Atlantic water domain) and the temperature time series (stations 3-7, 0-50 m, 0-200 m) can be used to assess seasonal and interannual variability of hydrographic conditions in the southern Barents Sea. In this study long-term variations in the Barents Sea were analyzed using monthly mean temperature data of the 0–50 m layer of the Kola Section, 1951–2013 (stations 3-7). The time series is displayed in Fig. 1.

For the sea temperature data analyzed, no reasonable prior information was available for the true values \( \mu_i \). One possible choice could be to impose a smoothing prior for the temperatures (cf. Erästö and Holmström, 2005; Erästö and Holmström, 2007). However, we did not want the smoothing prior to reduce the seasonal effects and therefore used a flat non-informative prior distribution for the time series \( \mu \) which leads to the posterior distribution.
\[ \mu_i | y_i \sim N(y_i, \sigma_i^2) \quad (4) \]

when the noise variance \( \sigma_i^2 \) is considered fixed. As the posterior mean of the true time series now equals the observed time series, the posterior means of the scale-dependent components \( z_j \) are actually obtained by decomposing the observed time series \( y \) into differences of smooths. For the variance parameter \( \sigma_i^2 \), we used a value of 1 for the summer months of June, July and August and a value 0.52 otherwise. These values were based on the estimation of variance in similar sea temperature time series obtained from Laksefjord and Varangerfjord. The sea temperatures were received from the Institute of Marine Research, Norway. We could have considered the two variances as unknown parameters, but given the reliable information about their likely values, chose to treat them as known quantities.

2.6 Atlantic salmon abundance data

Abundances of Atlantic salmon were available from three rivers running to the Barents Sea, the rivers Kola and Tuloma in the Russian Kola Peninsula, and the River Teno, which forms the border between northern Finland and Norway (Fig. 4) and runs into the Barents Sea through the Tana fiord on the Norwegian coast. In the River Kola, all ascending salmon are caught at a fish counting fence, except for those that have passed the trapping site during the spring flood before the fence is installed (Jensen et al. 1997). In the River Tuloma, all ascending salmon are captured and counted at the fish ladder next to a hydroelectric power plant (Karppinen et al., 2002). All sea age groups of salmon are included in the counts at both Kola and Tuloma. In the River Teno, annual estimated catches of one-sea-winter salmon (in numbers) has been used as a proxy of abundance (see Niemelä et al. 2004).

Although the oceanic migration of these three salmon populations is not
known in detail, young salmon from these rivers spend at least their first months of sea
migration in the Barents Sea and thus the temperatures measured at the Kola section are
likely reflecting the environmental conditions the salmon are facing at sea (Niemelä et
al. 2004; Aas et al. 2011). Moreover, the temperatures at the Kola section are highly
correlated with those measured at other gauging stations along the Norwegian coast of
the Barents Sea (Niemelä, pers.comm.; http://www.imr.no/sjomil/index.html).

3. Results

3.1 Multiresolution decomposition of the monthly temperature data

The observed temperature series shows an obvious seasonal component (Fig. 1). In order to reveal longer scale and quasi-oscillatory patterns, the time series is
smoothed with a discrete spline smoother where the smoothing level is chosen based on
the scale-derivative map so that the seasonal component is smoothed out but the other
features remain intact. This corresponds to the middle black line in the scale-derivative
map (Fig. 2). The smooth is shown with a blue line in Fig. 1 and it does seem to include
quasi-oscillatory components and there might also be an increase in the temporal trend
after 2000.

For a closer examination of the potential quasi-oscillatory components and
the possible increase in trend in the new millennium the time series is resolved into
scale-dependent components using the decomposition (1). The decomposition was
based on an automatic selection of two smoothing levels, shown by the two upper black
lines in the scale-derivative map in Fig. 2. A visual inspection of the map confirms this
as a reasonably good representation of the data at different time scales, since the aforementioned components are shown by three alternating bands of red and blue or yellow and cyan, and no other bands are seen.

The extracted components are shown in Fig. 3 and they include a noise component (b), the annual seasonal component (c), mid time-scale (~10 years) variation (d), long time-scale (~30 years) variation (e), and a linear trend (f). The smallest smoothing parameter, which produced the noise component, was not obtained from the scale-derivative map, but was manually added to the smoothing level sequence in order to clarify the credibility of the annual seasonal component. The centered linear trend corresponds to $\lambda_5 = \infty$. The credibility of the features was quantified using the HPW method with the credibility threshold 0.95. In Fig. 3, the credibly positive or negative features are indicated with white or black background whereas the background of non-credible patterns is gray.

The smallest scale component is noise (Fig. 3, panel (b)). The strongest credible pattern of variation with the greatest amplitude is the annual seasonal variation with an approximately 4 °C difference between the warmest and coolest months (Fig. 3, panel (c)). The effect of the other scale-dependent components on the temperature variation was smaller but still mostly statistically credible.

For each decade, warm and cool periods are suggested by the mid-scale component (Fig. 3 panel (d)). Five cool periods are repeated every 8-13 years and six warm periods every 7-15 years. The extreme temperature difference between the credible warm and cool periods was on average 0.9 °C. The u-shaped long-scale component shows an approximately 30 years long cool period from 1967 to 1998 and a warmer period in the new millennium (Fig. 3, panel (e)). Relative to the zero level, the temperature dropped approximately 0.34 °C during the cool period. A slightly
increasing linear trend (0.94 °C over the whole study period) is shown over the whole 60-year time period (Fig. 3, panel (f)). The additional increase in temperature since 2000 contributed by the u-shaped component has been 0.55 °C.

3.2 Connection between Atlantic salmon abundance and temperature

The Atlantic salmon abundance time series from the rivers Kola and Tuloma follows a seemingly similar pattern, whereas the Teno time series shows somewhat different characteristics, not only in magnitude but in other respects, too (Fig. 5). However, closer inspection of these three data sets in fact shows similarity in their patterns of variation.

To analyze the connection between salmon abundance and the Kola Section sea water temperature, we extracted the multiresolution components separately for each of the three temperature-abundance pairs using a separate smoothing parameter sequence for each pair. These sequences were chosen by eye as a compromise between the corresponding two scale-derivative maps associated with each pair. As the salmon abundance datasets have been collected yearly, annual means are used also for the Kola Section sea temperatures, instead of the finer resolution, monthly time series considered in Section 3.1.

The Kola river salmon abundance and annual sea temperature time series for the Kola Section are shown in panel (a) of Fig. 6. The corresponding scale-derivative maps are shown panels (e) and (f) of Fig. 6 with the selected smoothing
parameter values highlighted with black lines. The extracted component pairs are shown in the panels (b)-(d). The time series are decomposed into four components: noise, small scale features, large scale features and linear trends. Both time series contain positive linear trends (not shown). Note that the small and large scale temperature components shown in the panels (c) and (d) of Fig. 6 are similar to those in Fig. 3; the differences are due to different sampling resolutions and different smoothing levels used for extracting the time series components. This is true also for the two other temperature-abundance comparisons presented below.

[Figure 6 about here]

The first component pair in Fig. 6 (panel (b)) contains the annual variation and can here be considered as noise. As expected, these components do not appear to exhibit any connection. The second component pair (panel (c)) contains variation in roughly 10-year cycles and, for most of the time period considered, the two components are positively correlated. The Pearson correlation coefficient is 0.65 but the connection between the components seems to vanish in the new millennium. The third component pair contains the deviations from the positive linear trends. As previously discussed, the temperature data show a cool period (relative to the linear trend) during 1967 - 1998. On the other hand, salmon abundance during 1968-1998 has increased more than the linear trend would predict and therefore the correlation between these long-term components is negative.

The results for the River Tuloma are shown in Fig. 7. The salmon abundance has a negative linear trend and therefore a negative correlation with the linear trend in the temperature (not shown).

[Figure 7 about here]
Again, the first component pair appears to be uncorrelated and the second components show roughly 10-year cycles and strong positive correlation (the Pearson correlation coefficient is 0.80). For the Tuloma river, the positive correlation between the quasi-cyclical components persists even in the new millennium. The third components contain the deviation from the linear trends and exhibit a positive correlation.

[Figure 8 about here]

The results for the River Teno are presented in Fig. 8. The salmon abundance time series for the Teno river is shorter than for the two other rivers, starting in 1972. This shorter time interval is now used also for the Kola Section sea temperature. The first component pair contains the yearly variation and no apparent correlation, while the second pair shows roughly 10-year cycles and a positive correlation (Pearson correlation coefficient 0.46). The sea temperature contains positive linear trend also in this interval and the salmon abundance has a slightly negative linear trend (result not shown). The components that represent the deviations from a linear trend do not seem to have a clear connection.

All salmon abundance data sets contain a roughly 10-year cycle that seems to be correlated with the 10-year cycle in the temperature data. The 10-year abundance cycle of the Tuloma seems to track the 10-year cycle in temperature rather closely for the whole time period considered. The corresponding abundance pattern of the Kola river follows the temperature even more closely during 1990-2000 but after 2000 the connection seems to be lost. The disappearance of the connection may, in part, be due to the fact that the 2006 observation for the River Kola salmon abundance time series is missing whereas both the Tuloma and the Teno series then have a local maximum (see
Fig. 5). The 10-year abundance cycle of Teno does not follow the temperature as closely, but there still seems to be at least some correlation.

4. Discussion

We used a novel time series method developed by Pasanen et al. (2013) to extract marine temperature features at different time scales in the Barents Sea and assessed their statistical credibility with Bayesian inference. The features in the temperature time series were then compared with Atlantic salmon abundance data to explore their time-scale dependent connections. Traditional time series analysis would decompose the data into seasonal variation and a single large scale component that would include both the linear trend and quasi-oscillatory patterns. The novel method is able to extract more useful information from the observed data. Another advantage of the new method is that it can also assess the credibility of the discovered components.

Other time series methods for finding oscillatory and quasi-oscillatory features include Fourier-based spectral analysis, wavelets and singular spectrum analysis (SSA). The classical estimate of time series spectrum is the periodogram, originally defined by Schuster (1898). The REDFIT program by Schulz and Mudelsee (2002) combines the calculation of the Lomb-Scargle periodogram with false-alarm confidence bounds for the spectrum peaks, based on an assumption of autocorrelated red noise. For a review of spectral methods for climatic time series, see Ghil et al. (2002).

The scale space multiresolution method differs from spectral methods in that, as a time-
domain method, it provides time-localized information of the scale-dependent components, an advantage which it shares with the wavelet methods and SSA. The SSA, which is related to principal component analysis, represents the time series in phase space, where it is decomposed by singular value decomposition into additive components, some of which may relate to trend and oscillations and others to noise. However, the SSA is a purely algebraic method which does not contain a statistical model of noise. The posterior SSA by Holmström and Launonen (2013) combines the SSA with Bayesian inference to separate signal components from noise, where the inference is performed in a similar fashion as in the scale space multiresolution method. Compared to the relatively large number of components produced by SSA, the scale-dependent components of scale space multiresolution analysis form a compact representation of the time series with a smaller risk of noise artifacts identified as signal. The monographs by Golyandina et. al (2001) and Golyandina and Zhigljavsky (2013) provide extensive reviews of various aspects of SSA.

In the scale space multiresolution method, the scale is a continuous parameter and hence can be flexibly chosen to produce the best possible representation of the data, whereas in most wavelet methods (e.g. Vidakovic, 1999) a dyadic set of scales is used. Holmström et al. (2011) compared the results of scale space multiresolution decomposition to the results of wavelet analysis in the case of digital images and concluded that wavelets can be more susceptible to artifacts that do not correspond to clear image features.

In time series analysis, the data are often assumed to be equispaced in time without any missing observations. Sometimes the time intervals with missing values are simply excluded from the analysis (see e.g. Friedland et al., 2003a and Friedland et al., 2003b). Our method does not require temporally equispaced data but the results would
be different if a time period containing a local minimum or maximum were missing.

In this paper, we used scale space multiresolution analysis also for discovering dependence structures between two time series that appear in different time scales. Although we only used the simple Pearson correlation coefficient to assess the dependence between the multiresolution components of the two time series, a deeper scale space oriented analysis would consider correlation also for a range of time scales. Such an approach is described in Pasanen and Holmström, 2016. We leave such analyses for future work.

The scale space multiresolution method revealed cool and warm periods in the Barents Sea, separated by 7-15 years. Since the marine life cycle of northern populations of Atlantic salmon varies typically between five and eight years (Niemelä et al., 2006), some cohorts of fish will fall inline well with temperature trends while others will fall between the trends. This fluctuation of environmental factors may therefore induce variable effects among salmon year classes. Understanding the periodic variation in environmental conditions may facilitate interpretation of trends in situations where both anthropogenic and environmental factors are acting behind changes in population abundance and/or structure (Niemelä et al., 2006; Chaput, 2012).

The three salmon populations seem to show an overall, roughly 10-year cyclic pattern in their abundance. Similarly, many other anadromous salmonid populations undergo marked variations in abundance, typically attributed to climate variation (Huusko & Hyvärinen 2012 and references therein). Moreover, corresponding roughly 10-year cycles have been detected in several populations of Atlantic salmon in various parts of the species’ range (e.g. Niemelä et al. 2004; Aas et al. 2011, Huusko & Hyvärinen 2012). Interestingly, all three populations, especially the Kola, show a less cyclic fluctuation and increasing deviation from the sea water temperature trend in the
last part of the time series, roughly since the turn of the millennium. The overall increase in the Barents Sea temperature may have improved survival of certain life stages of salmon, especially the previously spawned fish (Niemelä et al. 2006), but the changing environment may affect other life stages, like outmigrating smolts in a different manner, causing deviations from the earlier cycles through more variable, unpredictable mechanism (cf. Otero et al. 2014; Heino et al. 2016). In addition, the abundance trend in the Teno seems to deviate from the two Russian rivers especially after the turn of the millennium. Similar deviation has been detected in the River Teno time series compared with some other rivers in the close vicinity (Anon 2012), and several reasons for this phenomenon can be considered. Given its large population size and relative contribution to the coastal salmon fishery in the Barents Sea, the Teno salmon has long been exploited rather heavily both in the sea and in the river (Niemelä et al. 2006; Anon. 2012; Vähä et al. 2014), and this may have affected its population cycle in a way that differs from several other, smaller rivers in the Barents Sea area. It should be remembered, however, that the abundance data from the Teno are estimated catches whereas those from the Kola and Tuloma rivers are actual fish counts at counting facilities. Caution should therefore be practiced in direct comparisons of the three rivers.

Population dynamics of anadromous salmonid fish are profoundly impacted by survival rates of marine growth, and timing of subsequent maturation, i.e. sea age at maturity, which in turn are strongly affected by environmental conditions at sea (Jonsson and Jonsson, 2007; Wells et al., 2007). For both Pacific salmon (Chinook salmon, *Oncorhynchus tshawytcha*; Wells et al., 2007), and Atlantic salmon (Friedland and Haas, 1996), it has been postulated that fast marine growth during the year before return is associated with earlier maturation, whereas other studies suggest that Atlantic
salmon tend to mature and return to freshwater earlier if growth during the first months has been slow (Jonsson and Jonsson, 2007). In any case, critical parameters in population dynamics of anadromous salmonids that are also crucial in life cycle modeling (Michielsens et al., 2008; Massiot-Granier et al., 2014) are typically related to sea temperatures at certain specific periods of time.

One of the most critical periods within the life cycle of Atlantic salmon are the first months at sea, and mortality during this period has increased over recent decades across most of the species’ geographic range (e.g. Chaput, 2012; Mäntyniemi et al., 2012). Factors and processes affecting the declining post-smolt survival are poorly understood but elements related to the marine environment, e.g. sea temperature and related feeding opportunities, in addition to direct predation, are likely candidates behind this development (Mäntyniemi et al., 2012 and references therein). Moreover, marine migration of salmon at other phases of their life cycle include elements in their behavior in swimming, feeding, predator avoidance etc. which are typically directly affected by environmental conditions, especially the sea water temperature (Todd et al., 2011). In this study, the sea temperature measurements are from a sea area which is representing the critical environment the salmon in the Barents Sea area are facing upon their ocean entry. The scale space multiresolution method presented here provides useful information by extracting specific cool and warm periods and trends from the basic time series data. It is able to reveal the environmental elements behind critical mortality factors connected to certain temporal periods of salmon life history, and trends and periodicity in environmental conditions. Deeper insights of the environmental factors help us understand, and eventually counteract the development in marine mortality rates.

Although an anadromous fish species has been used as an example in this
study, the novel statistical approach described in the article would certainly be useful in any analysis where temperature-related critical events and parameters of a species’ life cycle are examined. For instance, studies of long-lived marine species of fish or mammals would benefit from such a multiresolution technique able to extract specific periods and trends from a time series.
References


Figure 1. The observed sea temperatures in the Kola Section of the Barents Sea for 1951–2013, averaged over the depth interval 0-50 m (black line) together with a smooth (blue line). The level of smoothing applied is indicated by the middle black line in Fig. 2 and it smooths out the annual seasonal trend while leaving the other salient features intact.
Figure 2. The scale-derivative map of the observed sea temperatures in the Kola Section of the Barents Sea for 1951–2013, averaged over the depth interval 0-50 m. Time is on the horizontal axis and the vertical axis represents the logarithm of the smoothing parameter. The color of a pixel indicates the value of the scale-derivative for a given time and smoothing level, with deep red indicating a large positive value, deep blue a large negative value, and green corresponds to a value close to zero. The black lines indicate the smoothing levels used in the decomposition shown in Fig. 3. The two uppermost black lines indicate the local minima of the norm of the scale-derivative and the bottom line was chosen by visual inspection.
Figure 3. Panel (a): the observed sea temperature data of the Kola section of the Barents Sea for 1951-2013, averaged over the depth interval 0-50m. Panels (b)-(f): the additive components in the multiresolution decomposition (1), together with their statistical credibility. The smoothing levels used for in the decomposition are indicated in Fig. 2. The posterior means of the components are shown in blue. The horizontal yellow line is the zero level. The background color indicates the time intervals where the component is credibly positive (white), negative (black) or neither (gray).
Figure 4. The map showing the Kola Section of the Barents Sea and the three rivers from which Atlantic salmon abundance data were obtained.
Figure 5. Atlantic salmon abundance in the rivers Kola, Teno and Tuloma.
Figure 6. Analysis of the dependence between Kola Section sea temperature and Atlantic salmon abundance in the Kola river. Panel (a): the observed temperature (blue) and salmon abundance (black). Panels (b)-(d): the additive multiresolution components in the decomposition (1) for temperature (blue), and salmon abundance (black). Panels (e) and (f): The scale-derivative maps of the temperature and salmon abundance, respectively. Time is on the horizontal axis and the vertical axis represents the logarithm of the smoothing parameter. The color of a pixel indicates the value of the scale-derivative for a given time and smoothing level. See caption of Fig. 2 for more information. The black lines indicate the smoothing levels used to define the
multiresolution components displayed in panels (b)-(d). The sequence of smoothing parameters was chosen by eye as a compromise between the values suggested by the two scale-derivative maps. (High resolution image is available at http://cc.oulu.fi/~lpasanen/BarentsSea)

Figure 7. Analysis of the dependence between Kola Section sea temperature and Atlantic salmon abundance in the Tuloma river. Panel (a): the observed temperature (blue) and salmon abundance (black). Panels (b)-(d): the additive multiresolution components in the decomposition (1) for temperature (blue), and salmon abundance (black). Panels (e) and (f): The scale-derivative maps of the temperature and salmon
abundance, respectively. See captions of Fig. 2 and Fig. 6 for more information. (High resolution image is available at http://cc.oulu.fi/~lpasanen/BarentsSea)

Figure 8. Analysis of the dependence between Kola Section sea temperature and Atlantic salmon abundance in the Teno river. Panel (a): the observed temperature (blue) and salmon abundance (black). Panels (b)-(d): the additive multiresolution components in the decomposition (1) for temperature (blue), and salmon abundance (black). Panels (e) and (f): The scale-derivative maps of the temperature and salmon abundance, respectively. See caption of Fig. 2 and Fig. 6 for more information. (High resolution
image is available at http://cc.oulu.fi/~lpasanen/BarentsSea)