UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Master Thesis

Elin Svensson

Reconstruction of Monthly Mean SST Fields for the Baltic Sea, 1883-2011

Master's thesis in physics

DMI
Vejr, klima og hav

Academic supervisor: Markus Jochum

Academic co-supervisors: Kristine S. Madsen, Jacob L. Høyer

Submitted: 14/10/2016

Preface

This thesis represents the completion of my master's degree in physics with qualification profile in geophysics, which was carried out in the period 2013-2016 at the Niels Bohr Institute, University of Copenhagen. The work of the thesis was conducted in a time frame of one year, corresponding to 60 ECTS points, in co-operation with the Danish Meteorological Institute under supervision of Markus Jochum, Kristine S. Madsen and Jacob L. Høyer.

University: University of Copenhagen
Institute: Niels Bohr Institute
ECTS: 60
Author: Elin Svensson
KU-username: fps532
Email: elinosvensson@hotmail.com
Title: Reconstruction of Monthly Mean SST Fields for the Baltic Sea, 1883-2011
Submitted: October 14th, 2016

Signature | Date

Abstract

Monthly mean sea surface temperature (SST) fields are reconstructed for the Baltic Sea by combining the Danish Meteorological Institute (DMI) SST Climate Data Record (CDR) with in situ records in a multivariate regression model. The resulting fields obtain the spatial resolution of the DMI SST CDR and the temporal coverage of the in situ records. To enable reconstructions from 1883 to 2011, long and continuous in situ records are constructed by combining in situ observations from adjacent stations. Correlation analyses are performed to ensure that the most optimal combination of in situ records are used. In addition, historical sea ice fields are produced and applied to the reconstructed monthly mean SST fields to avoid misrepresentation of ice covered regions.

The reconstructed SST fields are validated against independent in situ records and the result demonstrates high accuracy and temporal stability for large parts of the Baltic Sea. The most accurate reconstructions are found in the Transition Zone where yearly validation statistics reveal an accuracy of 0.38°C with insignificant temporal trends in the biases. The SST fields in the western and southern Baltic Sea have further proven to be of sufficient quality for long-term climatic studies. In this thesis, the linear SST trends between 1883 and 2011 are examined. The results display an average warming of 0.07 - 0.09°C/decade in the major basins of the Baltic Sea.

Resumé

Gennemsnitlige månedlige havoverflade-temperatur (SST) felter er her rekonstrueret for Østersøen ved at kombinere Danmarks Meteorologiske Instituts (DMI) SST Climate Data Record (CDR) med in situ målinger i en statistisk model. Denne model rekonstruerer SST felter med den rumlige opløsning af DMI SST CDR og med den tidslige dækning af in situ målingerne. For at opnå rekonstruktioner fra 1883 til 2011 måttes in situ målinger fra omkringliggende stationer kombineres til udvidede datasæt. Korrelationsanalyse er udført for at sikre, at den mest optimale kombination af in situ datasæt var opnået. Historisk baserede havis-felter var derudover produceret og benyttet på de rekonstruerede gennemsnitlige månedlige SST felter for at undgå fejlfortolkning af områder med havis.

De rekonstruerde SST felter er valideret med uafhængige in situ målinger, og resultaterne viser at SST felterne har en høj nøjagtighed og tidslig stabilitet for store dele af Østersøen. De mest nøjagtige rekonstruktioner er opnået for Transition Zonen hvor årligt validerings statistik påviser en nøjagtighed på 0.38°C med insignifikante tidslige bias trends. SST felterne er påvist til at være af tilstrækkelig høj kvalitet for langsigtede klimastudier i de vestlige og sydlige dele af Østersøen. I dette studie er de lineære SST trends mellem 1883 og 2011 undersøgt. Resultaterne viser en gennemsnitlig opvarmning på 0.07 - 0.09°C/årti i de største bassiner i Østersøen.

Acknowledgements

First of all, I would like to thank my supervisors Markus Jochum, Kristine S. Madsen and Jacob L. Høyer for their advice and encouragement throughout this project. They have always welcomed me into their offices and I have really appreciated all of the friendly conversations. I would especially like to thank Kristine for helping me collect the in situ data and for always pushing me in the right direction.

I would also like to thank Pekka Alenius at the Finnish Meteorological Institute (FMI) and Ivo Saaremäe at the Estonian Weather Service for providing data. The remaining data used in this thesis originate from the Danish Meteorological Institute (DMI), the Swedish Meteorological and Hydrological Institute (SMHI) and the Marine Environmental Monitoring Network in the North Sea and Baltic Sea (MARNET) of the Federal Maritime and Hydrographic Agency (BSH).

Finally, I am very grateful for all of the support that I have recieved from my friends and family. My most special thanks go to my boyfriend Christian, who has been my greatest support.

Contents

1 Introduction **1**

2 The Baltic Sea **3**
2.1 Area definition . . . 5
2.2 Overview of the physical oceanography . . . 7

3 Sea surface temperature **10**
3.1 Defining SST . . . 10
3.2 Processes affecting SST . . . 12
3.3 Diurnal warming . . . 14
3.4 Annual SST cycle . . . 15
3.5 Annual sea ice cycle . . . 17

4 Satellite observed SST **19**
4.1 Introduction to satellite derived SST . . . 19
4.1.1 Polar orbiting satellites . . . 19
4.1.2 Measuring SST using infrared sensors . . . 21
4.2 DMI SST CDR . . . 23
4.2.1 Data included in the CDR . . . 23
4.2.2 Construction of the CDR . . . 25
4.2.3 Validation of the CDR . . . 26

5 Sea ice fields **27**
5.1 Masking out sea ice from the CDR . . . 27
5.2 Historical sea ice fields . . . 28

6 In situ records **32**

7 Calculating monthly means **39**
7.1 Danish, Swedish and German records . . . 39
7.2 Finnish records . . . 41
7.2.1 Method 1 . . . 41
7.2.2 Method 2 . . . 41
7.3 Comparison against CDR . . . 44

8 Correlation analysis **48**
8.1 Theory . . . 48
8.2 Correlation between in situ records . . . 49
8.3 Correlation between in situ records and CDR . . . 54
8.4 Correlation between CDR time series . . . 55

9 Constructing long and continuous records **57**
9.1 Method . . . 57
9.1.1 Combining records with temporal overlaps . . . 58
9.1.2 Combining records without temporal overlaps . . . 61
9.2 Extended records . . . 62
9.2.1 Extended Record A . . . 66
9.2.2 Extended Record B . . . 68
9.2.3 Extended Record C . . . 69
9.2.4 Extended Record D . . . 70
9.2.5 Extended Record E . . . 71
9.2.6 Extended Record F . . . 72
9.3 Comparison against CDR . . . 73

10 Multivariate regression model **76**
10.1 Theory . . . 76
10.2 Model parameters . . . 78
10.3 Model quality . . . 83
10.4 Validation . . . 85

11 Analysis of the reconstructed SST fields **89**

12 Discussion **96**

13 Conclusion and outlook **101**

Bibliography **103**

A Appendix **106**

B Appendix **108**

C Appendix **111**

D Appendix **113**

1 Introduction

Sea surface temperature (SST) is an essential parameter for the study of the ocean and atmosphere. It governs important processes in the surface layer of the ocean, such as the exchange of heat, momentum and gas (Stigebrandt, 2001). The spatial and temporal development of SST can therefore affect both environmental conditions and dynamical processes in the surface layer. In return, a variety of different fields, such as coastal zone management, fisheries, the development of phytoplankton, weather forecasting and studies of climate change, are dependent on accurate SST information (Feistel et al., 2008; Karagali et al., 2012; Høyer and Karagali, 2016). There is therefore a demand for regional studies of climatic SST signals. Such studies should preferably be based on high resolution SST records covering long time periods.

In the Baltic Sea, previous analyses of climatic SST signals have been based on sparse in situ observations and short non validated satellite observations (Høyer and Karagali, 2016; The BACC Author Team, 2008). Høyer and Karagali (2016) therefore saw a need to develop a high resolution SST Climate Data Record (CDR) for the Baltic Sea. The product was developed with consideration of regional conditions and contains daily gap-free SST fields with a spatial resolution of 0.03^o x 0.03^o. The product is referred to as the DMI (Danish Meteorological Institute) SST CDR and is a great product to base Baltic Sea SST analyses on. However, as the product is satellite based, it is only able to resolve SST fields between 1982 and 2011. Climatic SST signals should preferably be analyzed using data sets with much greater temporal coverage. In the Baltic Sea, the residence time of water is 30 years (Stigebrandt, 2001). The climatic signals should therefore preferably be studied from data sets that at least exceeds this internal timescale (Feistel et al., 2008).

The aim of this thesis is therefore to reconstruct historical SST fields that can be used by anyone who wishes to study long-term climatic SST signals in the Baltic Sea. The fields are constructed by combining monthly mean SST fields from the DMI SST CDR with monthly mean in situ records in a multivariate regression model. The resulting fields obtain the spatial resolution of the DMI SST CDR and the temporal coverage of the in situ records. The temporal coverage of the reconstructed fields are therefore determined from the availability of in situ data. Due to the limited amount of early in situ observations, it was not possible to reconstruct SST fields prior to 1883. Similarly, more recent ice data were lacking, setting an upper limit for the reconstructions to 2011. Due to data gaps, individual months are also missing. Still, the resulting product is by far the longest and most complete observation based product containing 2D SST fields for the Baltic Sea.

The following thesis describes the construction and validity of the monthly mean SST fields. The thesis starts by giving background information. A brief description of the study area and the most important aspects of the Baltic Sea physical oceanography is given in Chapter 2. The concept of SST is defined and explained in Chapter 3. The chapter focuses on defining SST and explaining the most important processes that affect the SST development in the Baltic Sea. The thesis then reviews the data used for the monthly mean SST reconstructions. The DMI SST CDR is described in Chapter 4, the sea ice fields applied to the reconstructed SST fields are described in Chapter 5 and the in situ records used in the study are reviewed in Chapter 6.

Before the DMI SST CDR and in situ observations could be combined in the multivariate regression model, some preparing analyses had to be made. First of all, the in situ records had to be converted to correspond to monthly mean values. The methods and assumptions used to convert the in situ records into monthly mean records are described in Chapter 7. Thereafter, correlation analyses had to be performed on a variety of combinations of in situ records and DMI SST CDR grid cells in order to select the most optimal in situ records for the model. The correlation analyses are described in Chapter 8. As the individual in situ records are rather short, in situ observations from adjacent stations had to be combined in order to construct long enough records to be included in the model. The construction of the long and continuous in situ records is described in Chapter 9.

The multivariate regression model is described in Chapter 10. Here, the model quality and validity of the reconstructed SST fields are also discussed. Although it is beyond the scope of this report to analyze the reconstructed SST fields, a short analysis on the most obvious temperature trends is presented in Chapter 11. Discussions and conclusions on the validity of the reconstructed monthly mean SST fields are given in Chapter 12 and 13, respectively.

2 The Baltic Sea

The Baltic Sea is a large, intra-continental brackish sea situated between 10-30°E and 54-66°N (Feistel et al., 2008). The physical oceanography is characterized by a freshwater surplus, the narrow and shallow connection to the ocean and the topographic division of a series of basins (Kullenberg and Jacobsen, 1981; Stigebrandt, 2001; Madsen, 2009). The Baltic Sea is often divided into three major areas; the Baltic Proper, the Gulf of Finland and the Gulf of Bothnia. In this thesis, the Transition Zone, consisting of Kattegat, Öresund and the Belt Sea, is also considered to be a part of the Baltic Sea, despite having a different oceanographic regime (Kullenberg and Jacobsen, 1981). A detailed description of the area division is given in Section 2.1, while the most important aspects of the physical oceanography is reviewed in Section 2.2. Here, the heat balance is given the greatest attention as it to a large extent determines the development of SST, as shown in Chapter 3. A more thorough description of the Baltic Sea oceanography can be found in Kullenberg and Jacobsen (1981), Stigebrandt (2001) and Feistel et al. (2008).

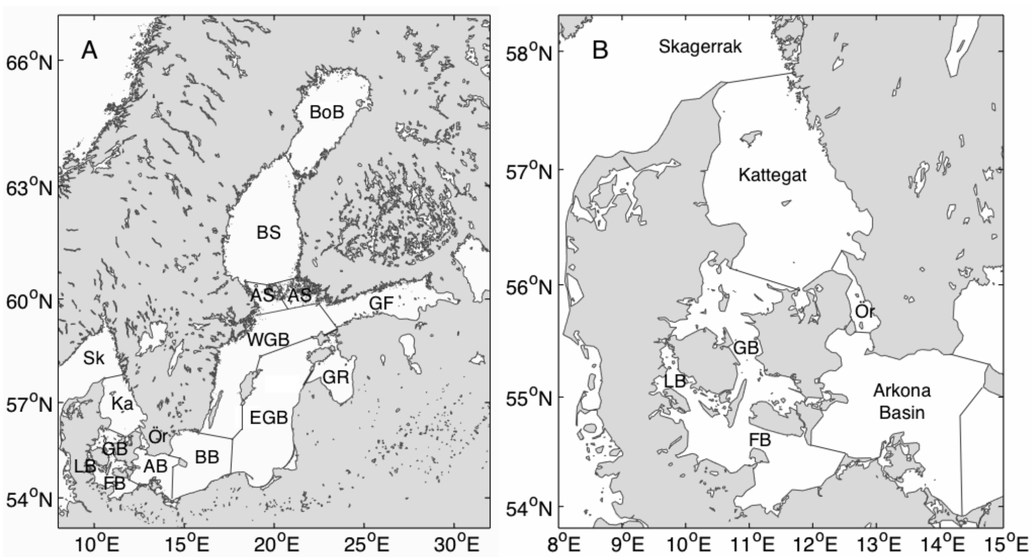

Figure 2.1: Major basins of the Baltic Sea (A) and zoom of the Transition Zone (B). In the figures, the following abbreviations are used: BoB = Bothnian Bay, BS = Bothnian Sea, ÅS = Åland Sea, AS = Archipelago Sea, GF = Gulf of Finland, GR = Gulf of Riga, WGB = Western Gotland Basin, EGB = Eastern Gotland Basin, BB = Bornholm Basin, AB = Arkona Basin, Ör = Öresund, FB = Fehmarn Belt, LB = Little Belt, GB = Great Belt, Ka = Kattegat, Sk = Skagerrak.

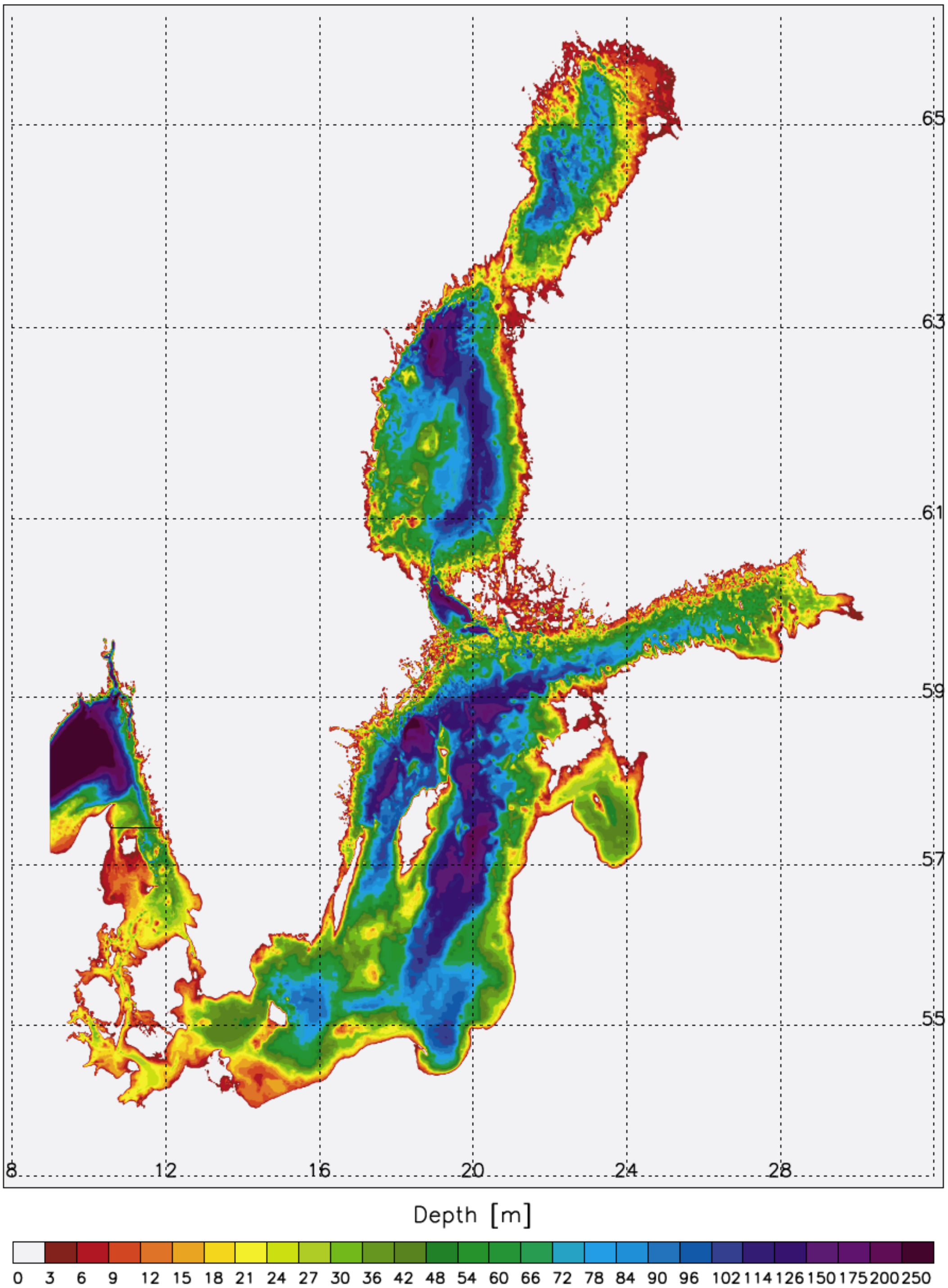

Figure 2.2: Bathymetry of the Baltic Sea, taken from Omstedt and Axell (2003).

2.1 Area definition

The Baltic Sea is often divided into three main areas; the Gulf of Bothnia, the Gulf of Finland and the Baltic Proper. The Bothnian Bay, Bothnian Sea, Åland Sea and Archipelago Sea collectively make up the Gulf of Bothnia. Similarly, the Baltic Proper consists of the Arkona, Bornholm, Eastern Gotland and Western Gotland Basins. The Gulf of Finland and the Gulf of Riga connects to the Baltic Proper from the east. In this thesis, the Transition Zone, consisting of Kattegat, Öresund and the Belt Sea, is also considered to be a part of the Baltic Sea. The division between the Transition Zone and Skagerrak is usually defined by the line Skaw-Marstrand (e.g. in Kullenberg and Jacobsen (1981)). However, in this thesis, a more simple division of 57.76°N is used. The different basins can be viewed in Figure 2.1, which is based on the area division made by Omstedt and Axell (2003). The bathymetry of the Baltic Sea is shown in Figure 2.2, which is taken directly from Omstedt and Axell (2003).

The Baltic Sea connects to the open ocean through the shallow Transition Zone. Kattegat has an average depth of 23 m, but deepens in the east to Skagerrak (Kullenberg and Jacobsen, 1981; Stigebrandt, 2001). The largest depths in Öresund reach 50 m while depths up to 80 m are found in narrow trenches in the Great Belt (Kullenberg and Jacobsen, 1981). Due to the shallowness, the water exchange between the Baltic Sea and North Sea becomes limited. The water exchange is further restricted by the shallow sills separating the Transition Zone from the Baltic Proper. A sill is defined as the deepest section of the most shallow barrier separating two basins (Madsen, 2006). It thereby sets a limit on the amount of water that can be exchanged between the basins. From Figure 2.2, it can be seen that water originating from Kattegat must pass either the Belt Sea or Öresund in order to enter the Arkona Basin. The shortest path is through the narrow and shallow Öresund. The sill in Öresund is located between Copenhagen and Malmö and has a depth of 7 - 8 m and a sill area of 0.1 km^2 (Kullenberg and Jacobsen, 1981; Stigebrandt, 2001). The Belt Sea constitutes a longer channel system, where water passing either the Great or Little Belt are connected at the Fehmarn Belt (Stigebrandt, 2001). The Gedser-Darss sill of the Fehmarn Belt has a depth of 17 - 18 m and a section area of 0.4 km^2 (Kullenberg and Jacobsen, 1981), and is thereby the deepest connection between Kattegat and the Arkona Basin (Winsor et al., 2001).

The Baltic Proper is by far the largest basin (Stigebrandt, 2001). The Arkona and Bornholm Basins are relatively shallow with average depths of 23 m and 46 m, respectively. The average depth of the Eastern and Western Gotland Basins are much greater, as shown in Figure 2.2. The maximum depth of the entire Baltic Sea is found at the Landsort Deep, located in the Western Gotland Basin. Here, the maximum depth reaches 459 m. In the Eastern Gotland Basin, the maximum depth reaches 249 m (Kullenberg and Jacobsen, 1981).

The Gulf of Riga connects to the Baltic Proper from the east through the Irbe Strait. The Irbe Strait is generally shallow but contains a narrow channel with a width of 27 km and sill depth of 21 m (Omstedt and Axell, 2003; Kullenberg and Jacobsen, 1981). To the north, the Gulf of Riga connects to the Gulf of Finland through the Suur Strait, which constitutes a complex system of straits. Due to the great number of straits surrounding the Gulf of Riga, it is almost completely closed off and is thereby characterized by its own set of physical characteristics (Omstedt and Axell, 2003). The average depth of the Gulf of Riga is 23 m (Kullenberg and Jacobsen, 1981). The Gulf of Finland is the easternmost extension of the Baltic Sea. There is no sill region connecting the Gulf of Finland to the Baltic Proper. The average depth of the Gulf of Finland is 37 m while the maximum depth is 115 m (Omstedt and Axell, 2003).

The Gulf of Bothnia is the northernmost extension of the Baltic Sea. With the exception of the shallow Archipelago Sea, the sub basins are separated by two sill areas (Omstedt and Axell, 2003). To the north, the Bothnian Sea and Bothnian Bay are separated through the Northern Quark Strait. The Northern Quark Strait is generally shallow, but contains two narrow channels with sill depths of approximately 25 m (Stigebrandt, 2001). To the south, the Bothnian Sea is separated from the Åland sea through the Southern Quark Strait. The sill is wider than at the Northern Quark Strait and has an approximate depth of 40 m (Omstedt and Axell, 2003). Of the four sub basins, the Åland Sea is the deepest with a mean depth of 75 m and a maximum depth of 300 m. In contrast, the neighboring Archipelago Sea is very shallow, with a maximum depth of 40 m (Kullenberg and Jacobsen, 1981).

2.2 Overview of the physical oceanography

The Baltic Sea hydrography is dominated by the division of a series of highly stratified basins, a net freshwater surplus and the limited connection to the open ocean (Winsor et al., 2001; Madsen, 2009). The only connection to the open ocean is through the shallow Transition Zone. Here, relatively fresh outflowing Baltic surface water meets high-saline Skagerrak surface water, forming a sharp west-east oriented front between Skagerrak and Kattegat (Stigebrandt, 2001). In Kattegat and the Belt Sea, the waters are mixed and form a surface layer of relatively low salinity. The thickness of the surface layer usually reaches 5 - 15 m. As this thickness approximately equal the sill depths of the entrance straits, the surface layer effectively block high-saline Kattegat deep water from entering the Baltic Proper (Stigebrandt, 2001).

The flow through the shallow straits of the Belt Sea and Öresund is mainly barotropic. That is, the flow is forced by horizontal pressure gradients along the straits. The pressure gradients are usually caused by weather related sea level differences between the Southern Kattegat and the western Arkona Basin. As the sea level difference varies a lot, so does the flow (Stigebrandt, 2001). The average circulation through the shallow straits can be viewed as a two layer-flow, where an outgoing fresh surface current removes freshwater from the Baltic, while a deep water current brings high salinity water into the Baltic (Kullenberg and Jacobsen, 1981). However, the instantaneous flow is often one-layered, where the entire water column of the Transition Zone either gives rise to an inflow or outflow. The salinity of the inflowing water therefore varies a lot depending on the recent history of inflow and outflow events (Stigebrandt, 2001). Water with high enough salinity to replace the bottom water of the Baltic Proper enter during so-called major inflows. These events are rare and are found during large and persistent inflows, usually caused by a high sea level in Kattegat being accompanied by intense westerly winds (Kullenberg and Jacobsen, 1981). In result, there is a strong wind driven vertical mixing, raising the surface salinity of the Transition Zone up to 28 psu (Stigebrandt, 2001).

While the Transition Zone controls the inflow of high salinity water, the majority of the freshwater is supplied to the large Gulfs (Omstedt and Axell, 2003). The largest freshwater supply originates from river inflow, while net precipitation contributes with freshwater of an order of magnitude smaller. In total, 80% of the river runoff and 85% of the net precipitation is supplied to the Gulf of Bothnia and Gulf of Finland (Omstedt and Axell, 2003). In the gulfs, the freshwater is mixed with sea water and enters neighboring seaward basins as low-saline surface water (Eilola and Stigebrandt, 1998). As all major basins have a net freshwater surplus, the surface salinity decreases away from the Transition Zone, and the lowest surface salinities are found in the Bay of Bothnia, the Bothnian Sea and the Gulf of Finland (Madsen, 2009; Stigebrandt, 2001). The surface salinity increases from approximately 3 psu in the Bay of Bothnia to 10 psu in the Arkona Basin (Winsor et al., 2001). In the Transition Zone, the salinity gradient is much larger, approaching 30 psu in Skagerrak (Stigebrandt, 2001).

In addition to the large horizontal salinity gradient, the water within each basin is highly stratified. Water of different densities are therefore applied at strata (Stigebrandt, 2001). This means that low salinity water enters from landward basins as surface water, while high salinity waters enter from seaward basins at greater depths. The dominating flow between the large gulfs are baroclinic (Omstedt and Axell, 2003). The baroclinic flow is caused by internal horizontal pressure gradients, which is caused by different vertical density distributions along the straits, separating the different basins (Stigebrandt, 2001).

The complete heat balance of the Baltic Sea has been characterized by Omstedt and Rutgersson (2000). They define the change in total heat content (H) with time to be a balance between the net heat flux of inflowing and outflowing waters (F_i - F_o) and the net heat loss through the surface (F_{loss}). In the latter, a positive heat flux indicates that heat is transmitted from the water to the atmosphere. As all heat fluxes are given in Wm^{-2}, the total change is obtained by multiplying the heat fluxes with the surface area of the Baltic Sea (A_s). The complete heat balance is thus given by:

$$\frac{dH}{dt} = (F_i - F_o - F_{loss})A_s \tag{2.1}$$

The heat flux through the surface can further be specified as:

$$F_{loss} = \underbrace{(1 - A_i)(F_h + F_e + F_l + F_{pr} + F_{sn} + F_s^o)}_{1} + \underbrace{A_i(F_w^i + F_s^i)}_{2} + \underbrace{F_{ic} + F_{ri}}_{3} \tag{2.2}$$

The first term of Equation 2.2 deals with heat fluxes in the open water. Here, A_i is the ice concentration, F_h the sensible heat flux, F_e the latent heat flux, F_l the net long-wave radiation, F_{pr} the heat flux associated with rain, F_{sn} the heat flux associated with snow and F_s^o the solar radiation received by the open water. The second term deals with heat fluxes in ice covered regions, where F_w^i is the heat flux between water and ice and F_s^i the solar radiation received through the ice. The last term deals with the net heat flux associated with advection, where F_{ice} is the heat flux of ice advected out of the Baltic Sea and F_{ri} is the heat flux of the river runoff entering the Baltic Sea.

Omstedt and Rutgersson (2000) show that the annual mean heat loss of the Baltic Sea is in close balance with the annual change of heat storage and the net exchange through the Transition Zone. The Baltic Sea is therefore found to act as a closed basin in a thermodynamical sense. Furthermore, it is also shown that the heat balance is dominated by two processes; the absorption of incoming solar radiation and the heat loss through the surface due to the latent heat flux, sensible heat flux and outgoing long-wave radiation. These are found to be an order of 10^2 larger than the other terms in Equation 2.1 and 2.2. These are also the main heat fluxes that affect the temperature of the surface waters (Stigebrandt and Gustafsson, 2003; Knauss, 1997), and will be described in greater detail in Chapter 3.

3 Sea surface temperature

This chapter gives a small introduction to SST and the typical SST cycles of the Baltic Sea. In order to do it properly, the temperatures that are referred to as SST have to be defined. There is often made a distinction between four types of SST; skin, sub-skin, bulk and foundation SST. The characteristic depth and temperature variations of each SST definition is described in Section 3.1. In Section 3.2, the main processes affecting the SST are described. The diurnal warming, annual SST cycle and annual sea ice cycle are reviewed in Section 3.3, 3.4 and 3.5, respectively.

3.1 Defining SST

In order to define SST, the surface layer has to be defined. The surface layer is found in the upper few meters of the surface and is usually divided into two main parts; the thermal skin layer and the diurnal warming layer (Robinson, 2010). The typical depths and vertical temperature distributions of the layers are illustrated in Figure 3.1. The figure is based on the schematics of Robinson (2010).

The thermal skin layer spans the uppermost 10 - 100 μm of the ocean and is where the processes that control heat loss takes place. As a consequence, there is a negative temperature gradient in the layer (Knauss, 1997). The temperature measured at the immediate surface is therefore typically some tenths of a degree colder than the temperature measured a millimeter further down. This is usually referred to as the cool skin effect (Karagali and Høyer, 2014). The temperature measured at the immediate surface is referred to as the skin SST while the temperature measured right below the thermal skin layer is referred to as the sub-skin SST. Satellites using infrared radiometers to observe SST measure skin temperatures, while satellites using microwave radiometers to observe SST measure sub-skin temperatures (Robinson, 2010).

The diurnal warming layer typically spans the upper meters of the surface. The temperature of the diurnal warming layer is affected by the amount of absorbed solar radiation and mixing (Knauss, 1997). The temperature difference between the thermal skin layer and diurnal warming layer can therefore vary significantly between day and night, under strong and calm winds, and cloudy and clear days. Most in situ observations are made in the diurnal warming layer, where the registered temperatures are referred to as bulk SST (Knauss, 1997).

The foundation SST is found beneath the diurnal warming layer and represents surface temperatures under well mixed conditions. From Figure 3.1, it can be seen that the foundation temperature is equivalent to the sub-skin temperature at night or during days with moderate to strong winds (Karagali and Høyer, 2014; Robinson, 2010). As the foundation SST remains constant throughout the day, it is usually used as a reference SST.

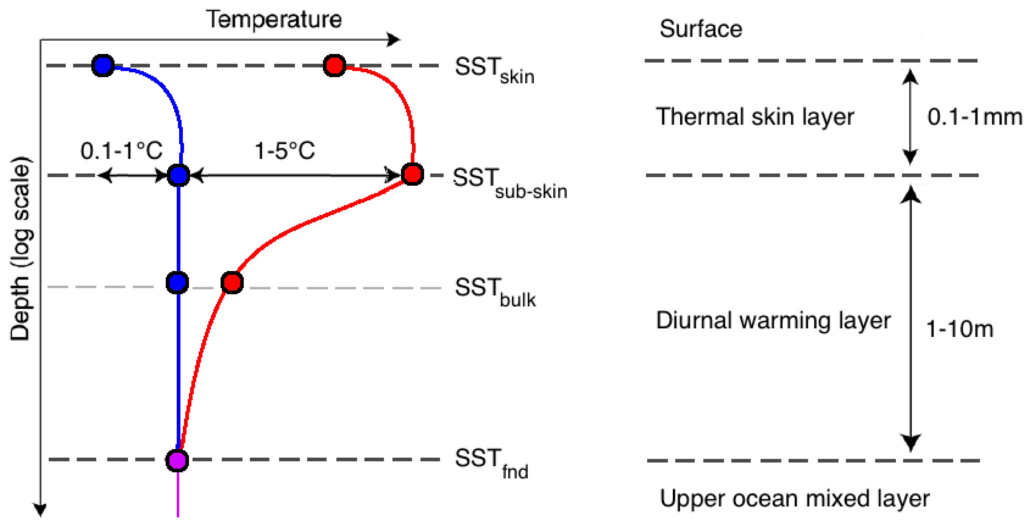

Figure 3.1: Illustration of the typical temperature profile of a calm and sunny day (red) and day with moderate to strong winds (blue). The temperature profile at night is the same as the temperature profile during moderate to strong winds (blue). The depths at which the skin, sub-skin and foundation SST is measured is marked. An example of the bulk SST is given. All temperatures measured in the diurnal warming layer are referred as bulk SST. The illustration is based on the schematics of Robinson (2010).

3.2 Processes affecting SST

In Section 2.2, it was shown that the dominating heat fluxes of the Baltic Sea include the absorption of incoming solar radiation and the heat loss through the surface due to the latent heat flux, sensible heat flux and outgoing long-wave radiation. These are also the main heat fluxes that affect the temperature of the surface waters (Stigebrandt and Gustafsson, 2003; Knauss, 1997). The rate at which heat enters a given surface area of the Baltic Sea can therefore be estimated from:

$$F_{loss} = F_h + F_e + F_l + F_s^o \tag{3.1}$$

Here, a positive value of F_{loss} indicates that more heat is leaving the surface than entering, resulting in a cooling of the surface layer.

The first term in Equation 3.1 (F_h) is the sensible heat flux and refers to the heat lost at the surface due to convection and conduction. When the ocean surface is warmer than the atmosphere, heat is transferred from the sea surface to the atmosphere. The amount of heat transferred increases with increasing temperature differences and increasing wind speed (Knauss, 1997). The sensible heat flux therefore varies with season, typically leading to a net heat flux entering the Baltic Sea between March and July, where atmospheric temperatures generally are warmer than the surface temperature (Omstedt and Rutgersson, 2000).

The second term in Equation 3.1 (F_e) is the latent heat flux and refers to the heat lost at the surface due to evaporation. The amount of heat lost due to evaporation is dependent on the difference between the specific humidity of the atmosphere and the sea surface, and increases with increasing wind speed (Stigebrandt and Gustafsson, 2003; Omstedt, 1990). The latent heat flux is cooling the Baltic Sea during all seasons but has the greatest cooling effect during fall (Omstedt and Rutgersson, 2000).

The third term in Equation 3.1 (F_l) refers to the net long-wave radiation emitted from the surface. According to Stefan-Boltzmann's law, all bodies radiate heat by an amount that is proportional to the fourth power of the absolute temperature of the body. However, much of the radiated energy is absorbed by clouds and greenhouse gases, and is therefore partly re-emitted back to the surface (Knauss, 1997). The net long-wave radiation is therefore dependent on the SST, cloud cover and the amount of greenhouse gases in the atmosphere (Stigebrandt and Gustafsson, 2003; Knauss, 1997). The net long-wave radiation is positive for all seasons, thereby cooling the surface layer (Omstedt and Rutgersson, 2000).

The last term in Equation 3.1 (F_s^o)refers to the short-wave solar radiation absorbed by the surface layer. The amount of received solar radiation is dependent on the latitude, earth-sun distance and cloud-cover (Stigebrandt and Gustafsson, 2003). While the incoming solar radiation is warming the surface layer during all seasons, the largest heat fluxes are found between May and July, with a peak in June (Omstedt and Rutgersson, 2000).

On an annual scale, the net heat loss to the atmosphere is in close balance with the received solar radiation. The Baltic Sea is therefore found to be in almost perfect thermal balance with the atmosphere above (Stigebrandt and Gustafsson, 2003; Omstedt and Rutgersson, 2000). However, on shorter time scales this is not the case, and the associated energy fluxes result in important temperature changes within the sea (Stigebrandt, 2001). In general, there is an input of heat (negative F_{loss}) between late March and August. Maximum temperatures are therefore usually found in August while minimum temperatures are found in February or March (Feistel et al., 2008). The maximum annual sea ice extent is also found in February or March (Omstedt and Axell, 2003). Sea ice forms when there is a negative heat flux through the surface at the same time as the surface water reaches its freezing temperature (Stigebrandt, 2001). The freezing temperature is salinity dependent and therefore shows great variation within the Baltic Sea. The freezing temperature can be estimated from (Stigebrandt, 2001):

$$T_f = -k_f S \tag{3.2}$$

Here, T_f is the freezing temperature (°C), S is the salinity (psu) and k_f is a constant of 0.054°C/psu. This means that the freezing temperature varies from approximately -0.16°C in the Bay of Bothnia (using a value of 3 psu) to -1.51°C in Kattegat (using a value of 28 psu).

3.3 Diurnal warming

A significant amount of the incoming solar radiation is absorbed by the upper meters of the surface (Karagali et al., 2012). The heating of the surface water due to the absorption of solar radiation is referred to as diurnal warming. The warming is especially pronounced in the upper millimeters of the surface, where the warming typically becomes 1°C larger than in the underlying water (Karagali and Høyer, 2014). Diurnal warming thereby has an especially large influence on skin and sub-skin temperatures.

The magnitude of the diurnal warming is found by subtracting the foundation SST from the SST of the heated water. During particularly sunny days with weak winds, the warming can reach values over 5°C (Robinson, 2010; Karagali et al., 2012). The typical temperature profile of such warming is illustrated by the red curve in Figure 3.1. Under moderate to strong winds, the diurnal warming may never establish (Robinson, 2010). The temperature profile is then equivalent to the nighttime temperature profile, which is shown by the blue curve in Figure 3.1.

Karagali et al. (2012) have quantified the diurnal warming of the Baltic Sea using observations from the Spinning Enhanced Visible Infrared Imager (SEVIRI) on board the Meteosat Second Generation satellite. As the SST is observed using an infrared radiometer, the observed temperatures correspond to skin SST. The analysis covers 10°W to 30°E and 48°N to 60°N, and thereby includes the Transition Zone, Baltic Proper and most of the Gulf of Finland. The Gulf of Bothnia is not included in the study. For each day, a foundation SST is calculated based on observations made between 00:00 and 03:00 LT (local time) of the given day, three previous days and three coming days. The diurnal warming is then quantified by subtracting the foundation SST from the corresponding daytime skin SST. Only temperatures exceeding the foundation SST with 2°C are characterized as diurnal warming events.

Karagali et al. (2012) find that the majority of the diurnal warming events occur between April and August. The local time of the maximum warming is usually found between 14:00 and 17:00, where the majority is observed at 15:00. While 75% of all diurnal warming events are within 3°C of the foundation SST, warmings exceeding the foundation SST with 5°C are observed. In fact, diurnal warmings of 6.4°C are encountered. The majority of the diurnal warming events are observed in semi-closed shallow areas of the Baltic Sea such as the Gulf of Riga, Gulf of Finland and the Danish Straits. Furthermore, most of the events are observed within 10 km of the coast.

As this thesis deals with monthly mean SST fields, it is also important to quantify the impact that the diurnal warming can have on the monthly mean SSTs. This has been done by Karagali and Høyer (2014), who also bases their study on the SEVIRI SST fields. The selected domain of the study covers 73°W to 45°E and 60°S to 60°N. Parts of the Gulf of Finland and the Gulf of Bothnia are therefore not resolved. Furthermore, it should also be mentioned that the SST fields have been corrected for the cool skin bias by adding 0.2°C to the temperature fields, and thus represent sub-skin temperatures. The foundation SSTs used in this study are calculated from 00:00-04:00 LT observations.

The impact of the diurnal warming was found by subtracting the mean daytime SST from the foundation SST of the same day. The monthly mean diurnal estimates were then obtained by averaging the daily mean diurnal warming estimates of the same month. In result, 12 monthly mean diurnal estimates were obtained. Karagali and Høyer (2014) found that the monthly mean diurnal warming signal can reach 0.4°C in large parts of the Baltic Sea. Combining the findings of Karagali et al. (2012) and Karagali and Høyer (2014), it can therefore be expected that the diurnal warming of monthly means based on local afternoon temperatures can become even greater, especially in shallow and coastal areas during the late spring and summer.

3.4 Annual SST cycle

The DMI SST CDR created by Høyer and Karagali (2016) is used to illustrate the annual SST cycle of the Baltic Sea. The DMI SST CDR is one of the main data sets used in this thesis and is described in detail in Chapter 4. Here, it is enough to know that the analysis is based on monthly mean SST fields. These are based on nighttime values and should therefore be free from diurnal warming and represent foundation SST. In addition, it should be mentioned that ice covered regions were assigned a temperature of -1°C in order to avoid misrepresentation of the monthly mean values.

The annual SST cycle of the Baltic Sea is illustrated in Figure 3.3. For each month, the average temperature was calculated by averaging over all monthly mean SST fields corresponding to that particular month. From the figure, it can be seen that the SST varies from approximately 15°C - 20°C in the summer to freezing temperatures during winter. The lowest temperatures are typically found in February and March while the highest temperatures are found in August. It can also be seen that the development of the temperature cycle is characterized by regional differences, such as longer winters and shorter summers in the northern regions of the Baltic Sea. The regional differences are particularly pronounced in the warming period of May and June and the cooling period of September and October.

As the Baltic Sea is located at high latitudes, the temperature cycle is greatly influenced by the seasonal variation in solar radiation (Stigebrandt, 2001). Furthermore, the amplitude of the seasonal cycle is much greater in the Baltic Sea than in open seas located at the same latitudes. This is mainly an effect of the strong stratification, as this limits the amount of convection and wind induced mixing (Stigebrandt, 2001). In addition, the effect is strengthened by the shallowness and enclosed nature of the Baltic Sea, as this causes the temperature variations to approach that of land (Høyer and Karagali, 2016).

The average amplitude of the seasonal cycle is illustrated in Figure 3.2. For each year, the annual amplitude has been calculated by subtracting the coldest monthly mean SST from the warmest monthly mean SST of each grid cell. The average annual amplitude was then found by averaging over all annual amplitudes. From Figure 3.2, it can be seen that the average amplitude of the seasonal cycle exceeds 13°C in the entire Baltic Sea. The largest values are found in the Gulf of Finland, the Gulf of Riga, along the southeastern coast of the Baltic Proper and along the eastern coast of the Gulf of Bothnia, where the amplitude of the seasonal cycle approaches 20°C.

Due to the limited surface area of the Baltic Sea, air advected over the Baltic Sea is not able to equilibrate with the underlying water (Stigebrandt and Gustafsson, 2003). Instead, the surface temperatures of the Baltic Sea are to a large extent determined by the atmosphere. Large variations in the seasonal cycle can therefore be found from year to year. This is often illustrated by the winter temperatures, where mild westerly winds causes relatively high temperatures, while cold easterly winds are accompanied by low temperatures and extensive ice covers.

3.5 Annual sea ice cycle

Every winter, sea ice forms in the Baltic Sea. The sea ice extent varies from year to year as the frequency and spatial distribution not only depends on the latitude but also on meteorological and hydrographic conditions (Feistel et al., 2008). Typical ice conditions can therefore vary on small spatial scales. As a minimum the Bothnian Bay, the eastern part of the Gulf of Finland and the Archipelago sea freezes over (Seinä and Palosuo, 1996). In an extremely severe winter, the entire Baltic Sea (including the Transition Zone) freezes. On average the sea ice reaches a latitude of approximately 59°N. The typical ice season starts in early November and the annual maximum sea ice extent is usually reached in late February or early March (Omstedt and Axell, 2003). The ice starts to melt in April and the Baltic Sea is at latest ice free in the beginning of June. The typical sea ice extent of mild, average and severe winters is described in Chapter 5. The typical sea ice extent between October and May can be viewed in Figure 5.1 and 5.2.

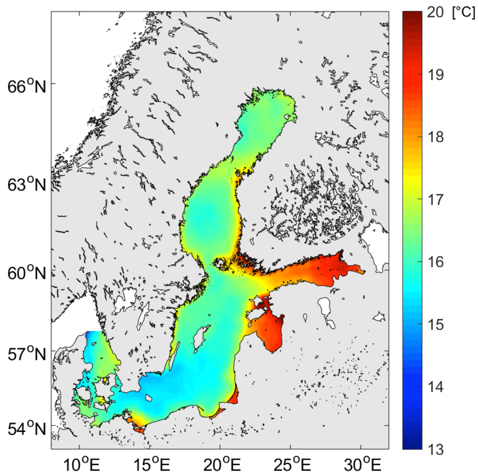

Figure 3.2: Average amplitude of the seasonal cycle between 1982 and 2011.

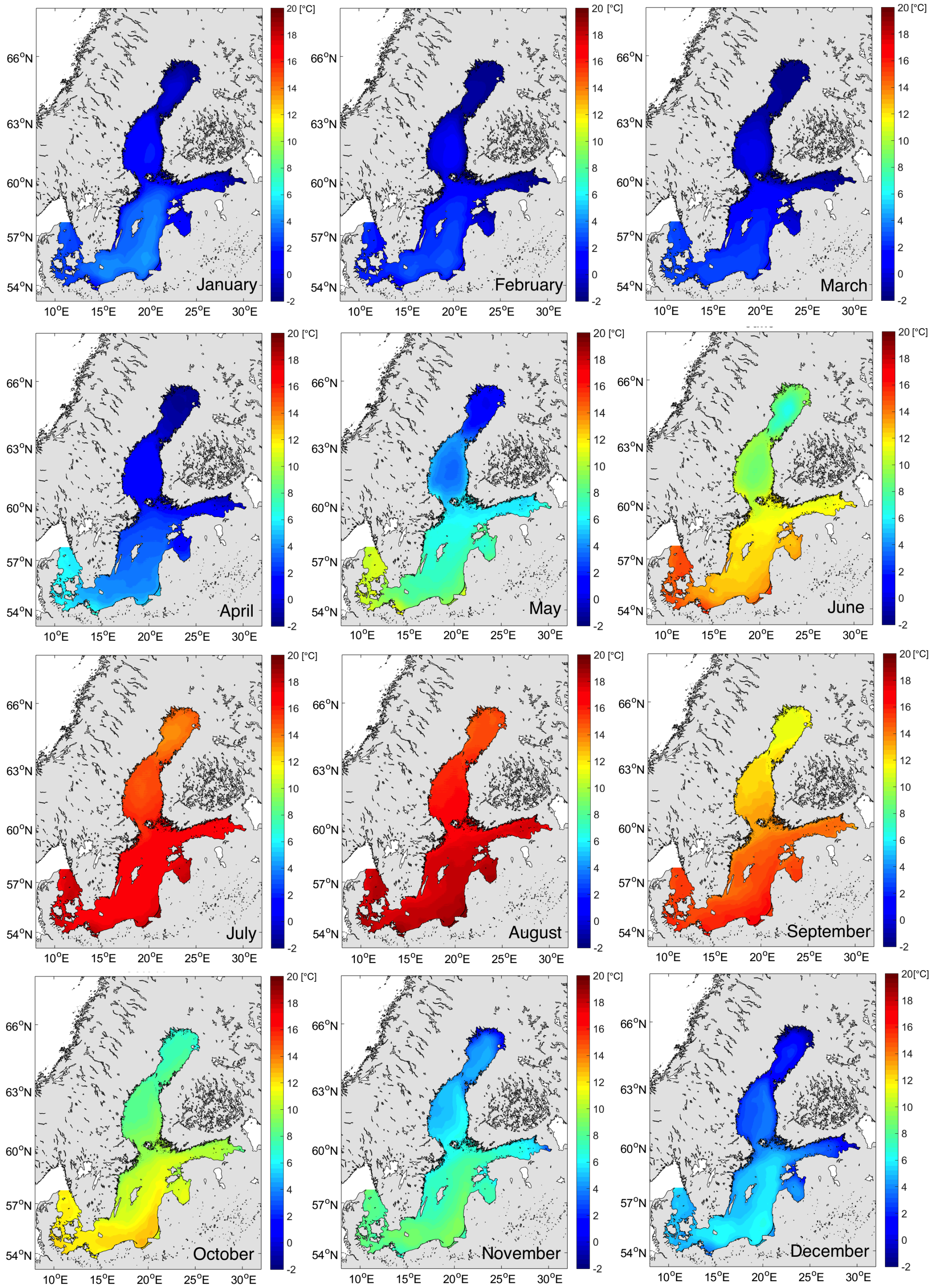

Figure 3.3: Average monthly mean SST between 1982 and 2011.

4 Satellite observed SST

One of the main data sets used in this thesis is the DMI SST CDR created by Høyer and Karagali (2016). The DMI SST CDR is part of the Copernicus Marine Environmental Monitoring Services and can be accessed at `http://marine.copernicus.eu/`. The DMI SST CDR is satellite based and consists of daily gap-free SST fields with a spatial resolution of 0.03° x 0.03°. The record covers the entire Baltic Sea and spans from January 1982 to December 2011. In this thesis, the corresponding monthly mean SST fields were accessed. As the DMI SST CDR is an important part of this thesis, a more detailed description on the construction and validity of the record is presented in Section 4.2. To help clarify the advantages and uncertainties of the record, an introduction to satellites and their ability to measure SST is given in Section 4.1. In the remainder of the report, the term CDR will refer to the DMI SST CDR.

4.1 Introduction to satellite derived SST

In this section, a brief introduction to satellites and their ability to monitor SST is given. The section is based on Robinson (2010) and deals with theory relevant for understanding the satellite data included in the CDR. The CDR is based on two satellite data sets; the Advanced Very High Resolution Radiometer (AVHRR) Pathfinder data set and the Along Track Scanning Radiometer (ATSR) Reprocessing for Climate (ARC) data set. As both data sets originate from polar orbiting satellites, the section will start of by describing the properties and sampling capabilities of such satellites. An overview of the sensors used to measure the SST is given in Section 4.1.2.

4.1.1 Polar orbiting satellites

Polar orbiting satellites are satellites that orbit the earth in an almost north-south direction, passing close to both poles in each orbit. The altitude of the orbit is rather low and typically ranges from 700 to 1350 km. The corresponding orbital period is approximately 100 min, resulting in 14 - 15 orbits a day. Each orbit has a so called ascending and descending track. The ascending track constitutes the part of the orbit where the satellite scans the earth from the southeast to northwest, while the descending track refers to the part of the orbit where the satellite scans the earth from the northeast to southwest.

The number of orbits the satellite has to complete before it has returned to its starting position varies. The time it takes for the satellite to return to the starting position is called the orbit repeat period. If the orbit repeat period is one day, the satellite has returned to its starting position after 14 - 15 tracks. If the orbit repeat period is longer than a day, the satellite needs to orbit the earth a greater amount of times before it has returned to its starting position, leading to a greater amount of tracks around the Earth. The tracks are often defined such that each orbit passes the equator at the same local time. The satellite is then said to be sun-synchronous.

Detectors or sensors used to monitor ocean properties are attached to the satellite. A single detector can monitor an area corresponding to its instantaneous field of view (IFOV). The resulting measurement is therefore given by the averaged ocean property registered by the intersection of the IFOV and the surface. However, as a measurement can't be made instantaneously, the sensor moves while making the measurement, resulting in each measurement corresponding to a slightly larger area than the IFOV. This effective area is called the footprint and determines the spatial resolution of the sensor. Often, the sensor moves such that it scans perpendicular to the satellite track. The scan lines are usually designed such that the time it takes to make a complete scan corresponds to the time it takes for the satellite to move a distance equal to the footprint size in the along-track direction. The spatial resolution will thereby equal the footprint size also in the along-track direction.

The distance the sensor scans perpendicular to the satellite track is referred to as the swath width. If the swath width is a minimum of 2700 km, the satellite is able to scan the entire earth in one day. Every place on earth will then be viewed from a descending and ascending track, resulting in a daytime and nighttime field. Due to geometry, the tracks will overlap at high latitudes, resulting in an even greater coverage in these regions. If the swath width is narrow, a greater amount of orbits are needed in order to scan the entire earth, leading to a greater orbit repeat period. It takes approximately 15 days to obtain a global coverage from a satellite with a swath width of 200 km. There is therefore a trade-off between high spatial and temporal resolution.

4.1.2 Measuring SST using infrared sensors

Satellites monitor the ocean through sensors that can detect electromagnetic radiation. Infrared (IR) radiometers are used to measure the thermal IR radiation emitted from the sea surface. The magnitude of the radiance can then be used to calculate the SST. It is therefore important that the IR radiometers measure the radiance in so-called window regions, as these correspond to wavebands in which the radiation only is weakly attenuated (absorbed or scattered) by atmospheric gases and particles. Nearly all of the radiation is thereby transmitted directly from the sea surface to the top of the atmosphere and can therefore be used to measure the SST. In the thermal IR region, window regions are found between 3.5 - 4.1 μm and 10.0 - 12.5 μm. The latter is often divided into two separate wavebands, 10.3 - 11.3 μm and 11.5 - 12.5 μm. All main sensors used to monitor SST (including the AVHRR and ATSR) contain these three channels. The channel measuring between 3.5 - 4.1 μm is usually called the 3.7 μm channel, while the latter two are referred to as the split window channels.

If it is assumed that the sea surface is a black body (has an emissitivity of 100%) and that the transmission through the atmosphere is 100%, the radiance (L) measured at the top of the atmosphere obeys Planck's Radiation Law:

$$L(\lambda, T) = \frac{C_1}{\pi\lambda^5[\exp(C_2/\lambda T) - 1]} \tag{4.1}$$

Here, L is the spectral radiance (Wm^{-2}m^{-1}str^{-1}) emitted from a surface of temperature T (K) with a wavelength of λ (m), C_1 is a constant of 3.74 x 10$^-$16 Wm^2 and C_2 is a constant of 1.44 x 10^{-2} mK.

According to Equation 4.1, it is possible to calculate the SST if the radiance of a specific wavelength can be measured. As each channel measures the radiance over a waveband, the equation would have to be integrated with respect to wavelength and convolved with the spectral sensitivity in order to obtain the radiance of the channel. However, as the sea surface isn't a black body and as the atmosphere attenuates some of the emitted radiation, Equation 4.1 is used to calculate the so-called equivalent black-body brightness temperature T_{bn}. This temperature is specific to channel n and represent the temperature that a black body would have if it emitted the same radiation as detected by the satellite.

Knowing that the emissitivity of seawater is greater than 0.98 in the thermal infrared region, that a small contribution of the satellite detected radiance is reflected sky radiance, and that T_{bn} is cooler than the actual SST due to the absorption of gases and particles, it is possible to calculate the SST from T_{bn} using an atmospheric correction procedure. The main idea behind such procedure is that the thermal IR radiation is attenuated differently in each of the three channels. It is therefore possible to estimate the amount of atmospheric gases and particles (attenuation) by calculating the difference in T_{bn} between two channels. During the daytime, this difference is calculated from the split window channels as channel 3.7 μm is contaminated with reflected solar radiation. At night, the difference between any two channels can be used.

If channel i records a brightness temperature of T_{bi} and channel j a brightness temperature of T_{bj}, the difference is given by $\Delta_{i,j,}(T_b)$. The larger the value of $\Delta_{i,j,}(T_b)$ is, the larger is the atmospheric attenuation and the greater correction is needed. This idea is utilized in many algorithms, which uses the top of the atmosphere brightness temperature to calculate the SST. An example of such algorithm is the Non-Linear SST (NLSST) algorithm developed by Walton et al. (1998). The NLSST is used to estimate the SST from the brightness temperature of the window channels according to:

$$\mathrm{SST} = AT_4 + B\Gamma(T_4 - T_5) + C[\mathrm{Sec}(\theta) - 1](T_4 - T_5) + D \tag{4.2}$$

Here, SST is the satellite derived SST estimate, T_4 and T_5 are the brightness temperatures of channel 4 (10.3 - 11.3 μm) and 5 (11.5 - 12.5 μm), Γ is a first guess of the SST and θ is the satellite zenith angle. A, B, C and D are constant coefficients which are estimated from linear regression to in situ observations.

The SST can only be calculated in cloud free pixels as clouds completely changes the transmission properties of the atmosphere. If clouds are not detected, the resulting SST will be inaccurate and underestimated. It is therefore important to be able to detect and remove cloud contaminated pixels. Cloud detection is usually better during the day as visible and near-infrared channels can be used. During nighttime, the temperature difference measured by any two infrared channels (as described above) is used to detect clouds. It is often the most difficult to detect clouds which only partly cover a pixel.

4.2 DMI SST CDR

This section describes the DMI SST CDR created by Høyer and Karagali (2016). As the CDR is one of the main data sets in this thesis, this section hopes to clarify the advantages and uncertainties of the CDR. The section is constructed such that it starts by describing the data included in the CDR. This is followed by a short description of the construction of the CDR. Finally, the validity of the CDR is reviewed.

4.2.1 Data included in the CDR

AVHRR Pathfinder data set

One of the main satellite data sets used to construct the CDR was the Advanced Very High Resolution Radiometer (AVHRR) Pathfinder data set, version 5.2 (Casey et al., 2010). The AVHRR Pathfinder started to observe SST on board polar orbiting NOAA platforms in 1981, with the primary purpose of imaging clouds and measuring SST (Walton et al., 1998). The period of the orbit is 102 min and the swath width of the satellite is approximately 2500 km. The orbits are defined such that there is an overlap between adjacent orbits, ensuring two global coverages each day. The Pathfinder data set hereby offers separate day and night products, both with a resolution of 4 km (Feistel et al., 2008).

In order to accurately estimate the SST, the AVHRR satellite sensor system contains five channels. Channel 1 measures in the visible spectral range, channel 2 in the near-IR reflective region and channel 3 - 5 in the thermal IR spectral region (Walton et al., 1998). As previously described, the latter channels are used to measure the SST. The SSTs are obtained from the brightness temperatures using the NLSST algorithm developed by Walton et al. (1998). The regression coefficients are estimated from linear regression to in situ observations. The Pathfinder data set thus represent bulk temperatures. Unfortunately, the AVHRR SST observations have been shown to have significant seasonal bias variations in middle and high latitudes (Høyer et al., 2012).

When constructing the CDR, Høyer and Karagali (2016) included data from January 1982 to December 2011 and decided to use the night product to ensure minimum diurnal warming effects.

ATSR Reprocessing for Climate

The other main satellite data set used to construct the CDR was the Along-Track Scanning Radiometer (ATSR) Reprocessing for Climate (ARC) data set (Merchant et al., 2012; O. Embury, 2012), version 1.1.1. The data set is based on observations made by three different ATSRs, which have monitored the SST from the European Space Agency's (ESA) Earth Observation satellites (Smith et al., 2012). Each ATSR have monitored the SST during a different time period. The resulting data set spans from August 1991 to April 2012 and has a spatial resolution of 0.1° (Merchant et al., 2012; Høyer and Karagali, 2016). Unfortunately, the data return of the ARC is rather low. In fact, the data return of the ARC is only about 10% of the Pathfinder data return for the Baltic Sea region (Høyer and Karagali, 2016). The limited data return is mainly caused by the narrow swath width of 500 km, resulting in an orbit repeat period of 3 days (Merchant et al., 2012).

The ATSR measures the SST using three infrared sensors with dual view capability. This means that every sensor observes the earth from two different angles; one vertically down through the atmosphere and the other at about 55° from the vertical in the direction of travel (Smith et al., 2012; Merchant et al., 2012). All places are therefore observed twice, which result in very accurate SST retrievals (Merchant et al., 2012). In addition, the ATSR is very well calibrated (Smith et al., 2012), which also increases the accuracy of the SST retrievals. In return, the accuracy of the ARC is much greater than the accuracy of the Pathfinder data set (Høyer and Karagali, 2016).

The SST retrieval coefficients are based on the radiative transfer equation (Merchant et al., 2012). The SST retrieval is therefore based on physics and is independent of in situ observations. The resulting SST estimates have been created to represent a variety of different depths. Høyer and Karagali (2016) chose to use SST estimates of 1 m, where skin and diurnal warming effects had been accounted for. Due to the low data return, both day and night data was used.

In situ observations

The in situ observations used to construct the CDR originated from the International Comprehensive Ocean Atmosphere Data Set (ICODAS) (Woodruff et al., 1987, 2011). These consist of observations made from buoys and ships. To gain a greater coverage, buoy observations from the Marine Environmental Monitoring Network (MARNET) was also included. Only the observations made by the most shallow sensor of each buoy was used, with a maximum depth of 4 m.

Sea ice

The sea ice information used to construct the CDR was provided by the Swedish Meteorological and Hydrological Institute (SMHI) and originated from digitized ice charts. During the beginning of the period, ice charts were provided twice a week, while daily ice charts were available by the end of the period. To obtain daily fields for the entire period, the ice field closest in time was used to fill a missing day. The spatial resolution of the ice charts were 5 km.

4.2.2 Construction of the CDR

From the description of the Pathfinder data set and the ARC data set, it can be seen that the ARC data set is the most accurate, while the Pathfinder data set has the greatest data return. In addition, both data sets contain gaps due to clouds. Høyer and Karagali (2016) therefore found the complimentary of the two data sets to be a great reason to construct a level 4 product. A level 4 product is a product that contains gap-free analyzed fields (Robinson, 2010). Such fields can be constructed through interpolation and/or by supplementing with data from additional sources. In this case, the two satellite data sets were combined with in situ observations and ice fields through an optimal interpolation method.

Before the level 4 product was created, the accuracy of the Pathfinder data set was improved through a dynamical high latitude bias correction method, as described by Høyer et al. (2014). The aim of the bias correction was to dynamically adjust the less accurate Pathfinder data set against more accurate (reference) data sets. The reference data set was chosen to be the ARC data set from the 14th of August 1991, and in situ observations prior to the availability of the ARC data set. It is shown that the biases found prior to and after the switch of reference record are consistent. It is thereby demonstrated that the bias correction is independent on the reference record, which should yield a consistent data set.

The bias corrected Pathfinder data set is combined with the ARC data set, in situ observation and sea ice fields according to the optimal interpolation method described by Høyer and She (2007). The method is designed to account for regional characteristics of mid and high latitude coastal and shelf seas, and should therefore yield accurate fields. The resulting product contains daily gap-free fields between 1982 and 2011, which have a spatial resolution of 0.03°.

4.2.3 Validation of the CDR

Høyer and Karagali (2016) used in situ observations that were not included in the analysis to ensure an independent validation of the CDR. These included observations from drifting buoys, moored buoys and ship observations from the ICOADS, as well as moored buoy from MARNET. The stability of the record was examined by calculating yearly average validation statistics for the entire region (North Sea and Baltic Sea). Disregarding the validation against ship observations, the yearly validation of the DMI SST CDR was found to meet the updated Global Climate Observing System (GCOS) requirements (GCOS, 2006, 2011). This means that the DMI SST CDR has an accuracy within 0.1°C and a stability within 0.03°C/decade. 15-day average validation statistics were also calculated to examine the performance throughout the year. The seasonal bias was found to be insignificant for the drifting buoys while the moored buoys and ship comparisons showed negative biases during winter and positive biases during summer. All data types showed the largest standard deviations during summer and the smallest during winter.

Overall, the DMI SST CDR has been demonstrated to be stable and accurate as it meets GCOS requirements for satellite based climate data records. The DMI SST CDR is therefore an excellent data set to use in this thesis. As the purpose of the thesis is to reconstruct monthly mean SST fields, the corresponding monthly mean SST fields were accessed. These were constructed by J. L. Høyer by averaging over all daily SST values for a given month and grid cell. The spatial resolution of 0.03° was thus kept.

The largest disadvantage of the data set is that regions covered with sea ice have been assigned a value of -1°C. This is a very rough estimate considering that the freezing temperature of the Baltic Sea varies from approximately -0.16°C to -1.51°C (as shown in Chapter 3). In addition, it is difficult to deduce the number of ice covered days that are included in the monthly mean estimate.

5 Sea ice fields

One of the main difficulties when working with SST is to deal with sea ice in a meaningful manner. Neglecting ice covered periods when calculating monthly or annual SST averages generally lead to overestimated values. Regions covered by sea ice are therefore set to have a temperature of -1°C in the CDR (Høyer and Karagali, 2016), as this is the approximate freezing temperature of the brackish water. At the same time, most in situ records lack data from ice covered periods as most in situ stations were withdrawn at the presence of ice. This makes it difficult to combine the CDR with the in situ records in a meaningful manner. Ice contaminated monthly mean values are therefore eliminated from the CDR such that the coming analyses (Chapter 7 - Chapter 10) only are based on actual SST. The method used to eliminate the ice contaminated data is described in Section 5.1.

Furthermore, it is important that the final reconstructed SST fields include sea ice to accurately represent past temperature conditions. Historical sea ice fields are therefore constructed. As the overall ice coverage, maximum extent and date of maximum extent varies from year to year (Seinä and Palosuo, 1996), it would require a study on its own to reconstruct accurate sea ice fields. Instead, characteristic fields representing mild, average and severe winters are created. The construction is described in Section 5.2.

5.1 Masking out sea ice from the CDR

In order to eliminate ice contaminated data from the CDR, ice fields with the same spatial and temporal resolution as the CDR are created. The ice fields are designed such that every grid cell contains information on the number of days that were covered with sea ice for each month. If at least one day has been covered by sea ice, the corresponding CDR grid cell is set to `NaN` (Not a Number). In this way, only monthly means that are based on actual SST remain in the CDR, making it consistent with the in situ records.

The ice fields are based on sea ice concentrations provided by the Swedish Meteorological and Hydrological Institute (SMHI). The data set spans from the 1979/80 winter season to the 2010/11 winter season, thereby providing information for all winters in the CDR. The sea ice concentration is gridded with a spatial resolution of 0.1° in the west-east direction and 0.05° in the north-south direction. The sea ice concentration of every grid cell is expressed as a value between 0 and 1, where 0 corresponds to a completely ice free cell and 1 corresponds to a completely ice covered cell. In the beginning of the period, a sea ice concentration is given approximately twice a week, which increases to daily values at the end of the period.

As Høyer and Karagali (2016) used the same ice concentrations to detect ice covered grid cells in the CDR, it is believed that the same information can be used to accurately filter out the ice contaminated data. However, before the sea ice concentrations can be used to detect ice contaminated grid cells, the data must be regridded to match the spatial and temporal resolution of the CDR. This is done in accordance with Høyer and Karagali (2016)[1]. The spatial resolution of the CDR is obtained by assigning every 0.03° x 0.03° grid cell with the ice concentration of the closest grid cell in the SMHI data. Similarly, daily values are obtained by assigning each day with the ice concentration of the day closest in time. A monthly value is then obtained by counting the number of days with sea ice in each month. In doing so, an ice concentration of 0.3 was used to separate ice covered cells from ice free cells. In result, the ice fields have a spatial resolution of 0.03° x 0.03° where every grid cell contains information on the number of days that were covered with sea ice for each month. If at least one day contains ice, the corresponding CDR grid cell is set to `NaN`.

5.2 Historical sea ice fields

The final reconstructed SST fields must represent sea ice in order to accurately capture past temperature conditions. However, it is not possible to construct corresponding monthly mean ice fields due to the large amount of work and data gathering required. Instead, historical ice fields representing the typical sea ice extent of mild, average and severe winters are created. These fields are based on the sea ice concentrations provided by SMHI and the winter severity definitions of Seinä and Palosuo (1996).

Seinä and Palosuo (1996) defines the severity of a winter season from the maximum annual sea ice extent. The classification is based on the ice extent of 276 ice seasons (1719-1995) and is defined such that there is an approximate equal amount of mild, average and severe ice seasons (approximately 33% each). The

[1]Method was described in person by Jacob L. Høyer

winter season is defined to be mild if the sea ice extent is smaller than 139 000 km^2, average if the maximum sea ice extent falls between 139 001 and 279 000 km^2, and severe if the maximum sea ice extent exceeds 279 001 km^2. A completely frozen Baltic Sea corresponds to an area of 420 000 km^2 (Seinä and Palosuo, 1996).

The Baltic Sea maximum annual sea ice extent between the 1719/20 winter season and the 2012/13 winter season is available at `http://www.eea.europa.eu/data-and- maps/daviz/maximum-extent-of-ice-cover#tab-chart`[2]. The reconstructions from 1719/20 to 1939/40 were created by Professor Jurva, wherefrom the Ice Service of the Finnish Institute of Marine Research has continued the time series (Seinä and Palosuo, 1996). The data is considered to be reliable since the late 1800's (Seinä and Palosuo, 1996).

Using the maximum annual sea ice extent, all winter seasons included in the SMHI data set are classified as either mild, average or severe. This corresponds to 14 mild, 14 average and 4 severe winters. For every month, classification and grid cell, an average sea ice concentration is calculated. If the resulting sea ice concentration is 0.3 or higher, the grid cell is set to contain ice. The ice fields are regridded to obtain the spatial and temporal resolution of the CDR using the method described in Section 5.1. The resulting monthly mean sea ice extent for mild, average and severe winters are shown in Figure 5.1 - 5.2. As the severity of the past winter seasons can be deduced from the maximum annual sea ice extent chart, the corresponding monthly ice fields can easily be applied to the reconstructed monthly mean SST fields (see Chapter 10).

[2]Data were downloaded in June 2016

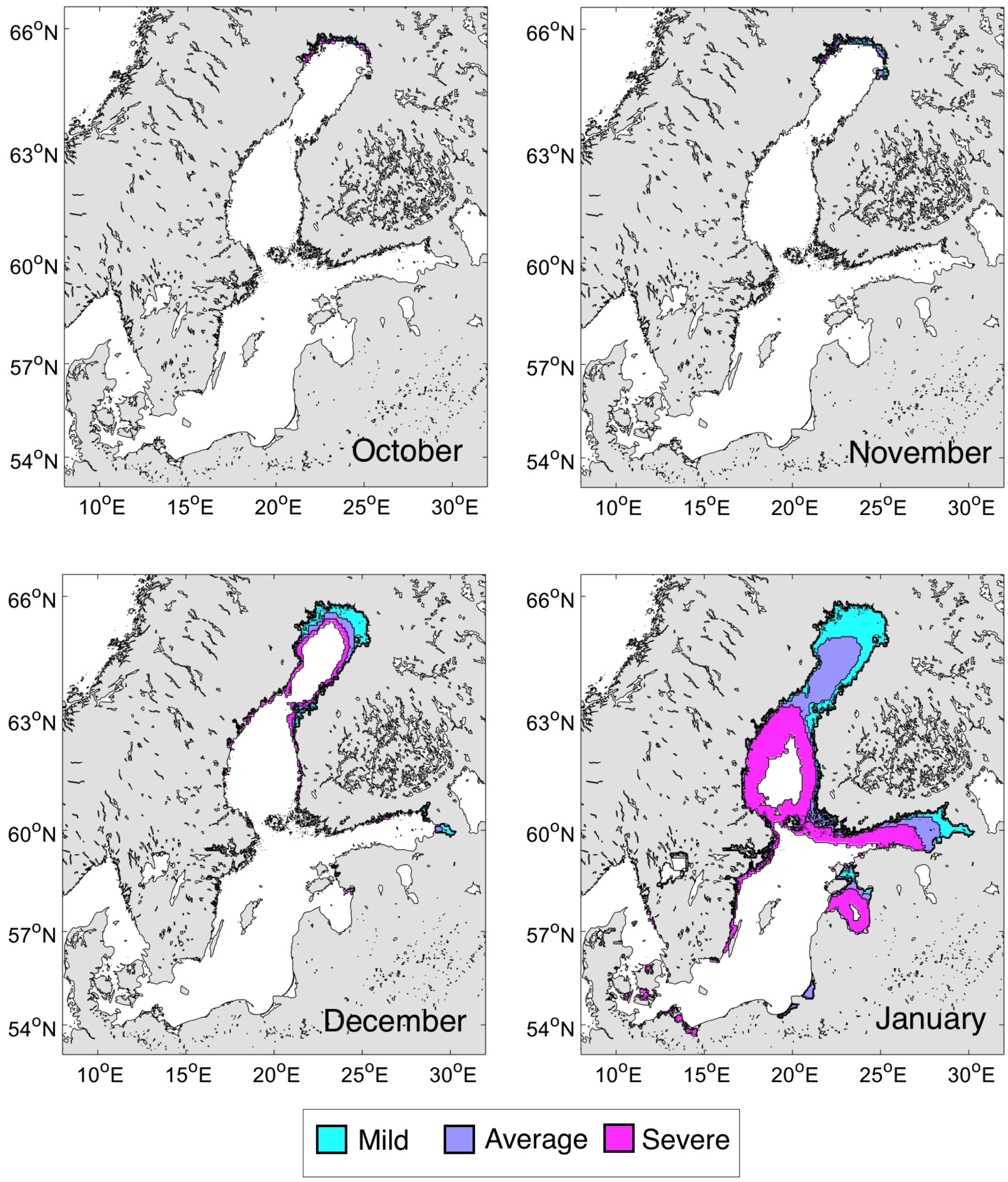

Figure 5.1: Average sea ice extent in October, November, December and January for mild, average and severe winters.

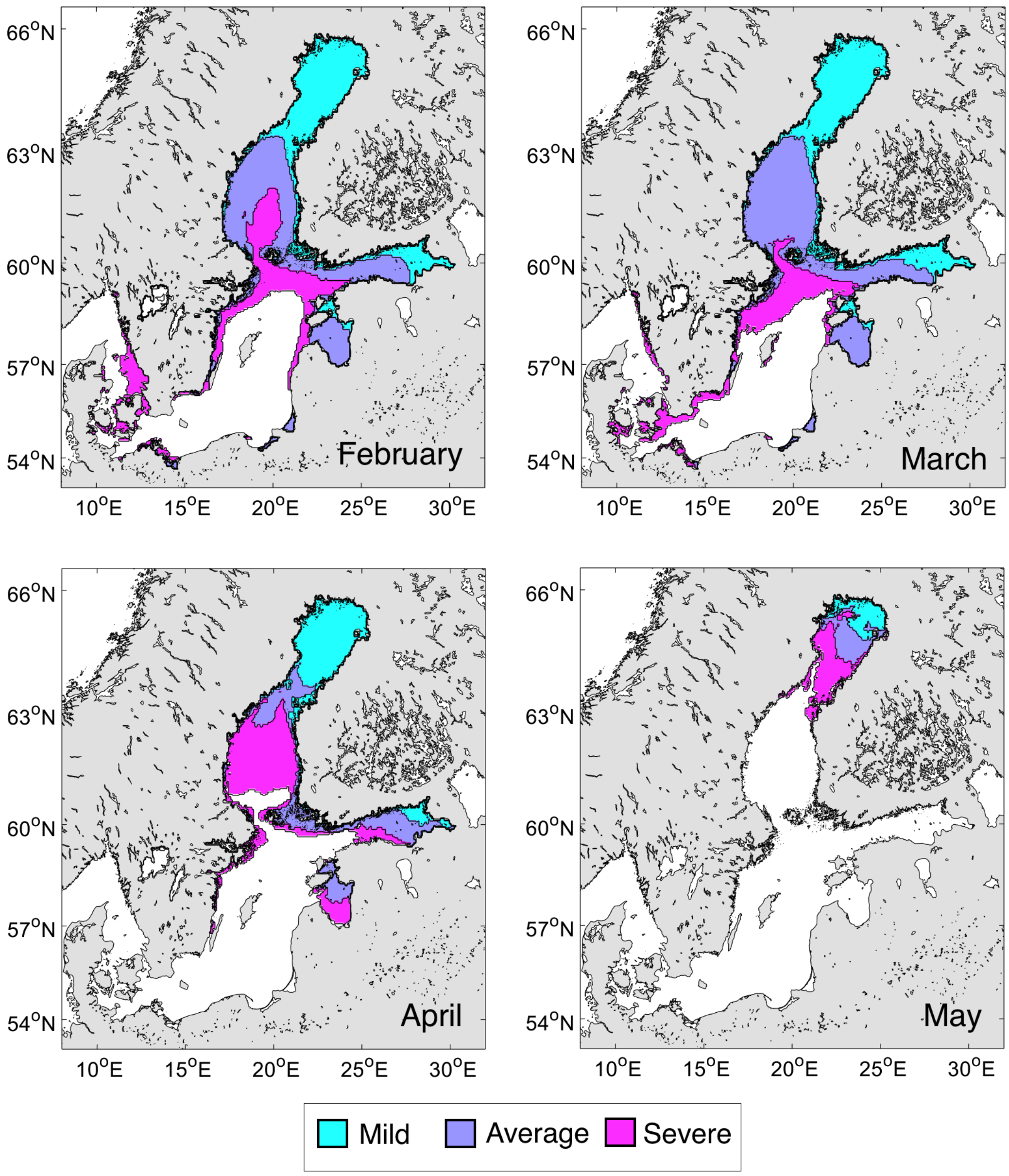

Figure 5.2: Average sea ice extent in February, March, April and May for mild, average and severe winters.

6 In situ records

This chapter describes the in situ records used in the study. The data originate from Denmark, Sweden, Germany and Finland and include observations made at coastal stations (maintained and automatic), lightships and buoys. In total, 78 in situ records are used. Of these, 32 are used for the reconstructions while the rest are used to validate the result. The spatial distribution of the in situ stations is shown in Figure 6.1 and detailed information on the stations is given in Table 6.1. In the table, the listed years correspond to the data available for this thesis. It is possible that observations from additional time periods exist. For example, earlier (not yet digitized) observations from Danish lightships are available in the Nautical-Meteorological annuals of the Danish Meteorological Institute (Madsen, 2009). Similarly, the depths and times noted in the table refer to the times and depths of the data used in this thesis. Many of the in situ stations have recorded temperatures at additional times and depths.

The Danish in situ records were provided by the Danish Meteorological Institute (DMI) and includes observations from lightships, maintained coastal stations and automatic coastal stations. The majority of the lightship records span from the 1930's to the 1970's. Drogden constitutes by far the longest record, spanning from 1900 to 1998. Besides recording daily surface temperatures at 07:00 or 08:00 LT, the lightships have also recorded vertical temperature profiles. While Læsø Trindel, Anholt Knob and Gedser Rev/Kadetrenden have measured daily 5 m temperatures in 1976, data from 0 m is lacking in all records. In some data sets, observations from the Second World War (1939-1945) are also lacking.

From Table 6.1, it can be seen that the earliest Danish coastal measurements begin in 1931 and that most records continue until the 1990's. Unfortunately, data from 1976 are lacking from all records. In the late 1990's to the early 2000's, the maintained coastal stations were replaced by automatic coastal stations (Madsen, 2009). In contrast to the daily measurements obtained at the maintained coastal stations, data are available for every 10 - 15 min at the automatic stations.

The Swedish in situ records were obtained from the Swedish Meteorological and Hydrological Institute (SMHI). The data were downloaded from `http://opendata-download-ocobs.smhi.se/explore/?parameter=3`[1], which contain SMHI's open

[1]Data were downloaded in September 2015

data. The data constitute observations made by buoys, lightships and coastal hydrographic stations. In the data sets, every observation is assigned a quality flag of either 'G' (green) or 'Y' (yellow). 'G' marks quality controlled and approved observations while 'Y' is used to label suspicious or uncontrolled values. Only quality controlled observations are included in this study.

The Swedish lightships and maintained coastal stations included in this study have observed ocean temperatures since the 1800's. While the longest coastal records end in the 1920's, most lightship records continue until the 1960's. From Table 6.1, it can be seen that Falsterborev and Finngrundet contain measurements from as early as the 1860's. Unfortunately, there is a large data gap from the mid 1860's to the mid 1880's in both records. All records also contain gaps during the First World War (1914-1918) and the Second World War (1939-1945). From the northernmost stations, data from many (or all) winters are missing. This is most likely caused by the withdrawal of the ships due to sea ice. More recent data (from 2010 until present) can be accessed from six automatic coastal stations. Unfortunately, all of the data were marked with a quality of 'Y' and were therefore not used. The temperature at the coastal stations and lightships have been observed on a daily basis. Most observations were registered at 08:00 LT, but measurements from other times also exist. Besides measuring surface temperatures (0 m), temperature profiles were also obtained.

The Swedish buoys have also been used to measure temperature profiles, where the most shallow depth varies between the buoys. The surface temperatures obtained by the buoys therefore correspond to depths between 0 m and 5 m. The temperatures are recorded on an hourly to sub-hourly basis. However, the data sets are found to be rather uncontinuous. Data are often missing for several months or years. It should also be noted that Almagrundet only has data for January 1979, data are then missing until 1986.

In addition to the open data records, SMHI has access to the in situ records listed at `http://www.smhi.se/klimatdata/oceanografi/havstemperatur/stationslista-havstemperatur-1.2289`. Due to financial limitations, additional records could not be purchased. If possible, it would have been especially interesting to obtain the data recorded at Hanö (station number 37216, 56.00°N 14.83°E), Landsort (station number 37205, 58.73°N 17.87°E) and Bjuröklubb (station number 37197, 64.48°N 21.57°E) as these constitute long records that are evenly distributed along the Swedish east coast.

The German buoys included in the study are part of the Marine Environmental Monitoring Network in the North Sea and Baltic Sea (MARNET) of the Federal Maritime and Hydrographic Agency (BSH). Information on the buoys can be found at `http://www.bsh.de/de/Meeresdaten/Beobachtungen/MARNET-Messnetz/MARNET.jsp`. In this thesis, only observations from the most shallow sensors were used, which varies from 0 m to 3 m. All buoys have recorded the temperature on an hourly basis. Kiel constitutes the longest record, which spans from 1987 to 2015.

The Finnish in situ records were obtained from personal contact with Pekka Alenius at the Finnish Meteorological Institute (FMI). Of 48 available coastal and lightship records, 12 records were accessed. The records were selected such that they would cover the Gulf of Finland and the Gulf of Bothnia in the most effective way (both spatially and temporally). As none of the Danish, German or Swedish in situ records cover the Gulf of Finland, 4 of the records were chosen to represent the Gulf of Finland. The remaining records were selected to complement the Swedish records of the Bothnian Sea and to cover the Bay of Bothnia. All measurements originate from the upper half meter of the ocean and were registered at 14:00 LT. Approximately 3 measurements are available for each month.

Finally, 6 in situ records were obtained from personal contact with Ivo Saaremäe at the Estonian Weather Service. The records cover the southern coast of the Gulf of Finland, the northeastern Baltic Proper and the Gulf of Riga, with a temporal coverage from the 1940's until present. The records are therefore a great complement to the already accessed records. Unfortunately, when the records were compared to the CDR (as the other records will be in Chapter 7), the records were found to deviate significantly from the corresponding CDR time series. The standard deviations between the in situ records and corresponding CDR time series were found to vary between 1.26°C and 3.09°C. In addition, individual months were found to differ with as much as 9.01°C from the corresponding monthly mean SST in the CDR. The reason for the large differences are not known, and the Estonian in situ records were therefore not included in the study.

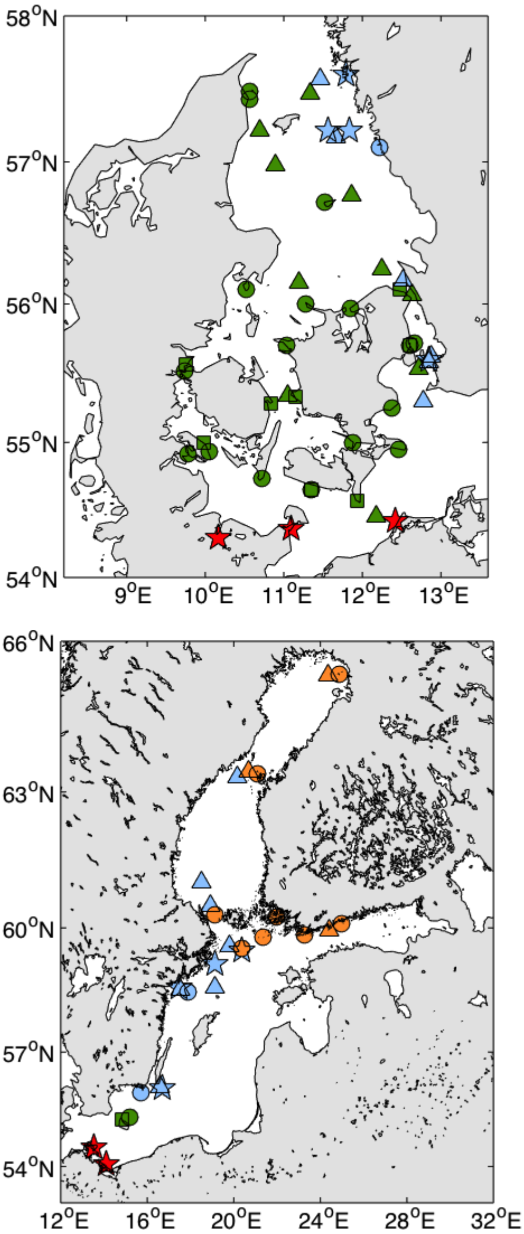

Figure 6.1: Map showing the spatial distribution of in situ stations. The colors correspond to stations belonging to different countries where green represents Denmark, blue Sweden, red Germany and orange Finland. The shapes correspond to different station types where circles represent maintained coastal stations, squares automatic coastal stations, triangles lightships and stars buoys.

St. Num.	Station Name	Years	Lat.	Lon.	Time	Depth
Danish maintained coastal stations						
20098	Frederikshavn	1939-1990	57.43	10.57	08.00	0m
20108	Hirsholm	1990-1995	57.48	10.57	08.00	0m
22160	Sletterhage	1931-1985	56.10	10.52	08.00	0m
26458	Mommark	1931-1967	54.93	10.05	08.00	0m
26478	Sønderborg	1931-1999	54.92	9.78	08.00	0m
27022	Anholt	1990-1999	56.72	11.52	08.00	0m
28118	Middelfart	1931-1999	55.52	9.73	08.00	0m
28548	Keldsnor	1931-1952	54.73	10.72	08.00	0m
"	Bagenkop	1952-1999	54.75	10.67	08.00	0m
29007	Gniben	1991-1999	56.00	11.28	08.00	0m
29008	Hundested	1931-1942	55.97	11.85	08.00	0m
"	Rørvig	1942-1999	55.95	11.77	08.00	0m
29118	Refsnæs	1931-1983	55.70	11.03	08.00	0m
30332	Langelinie	1936-1983	55.70	12.60	08.00	0m
30342	Middelgrundsfortet	1931-1999	55.72	12.67	08.00	0m
31062	Rødvig	1931-1999	55.25	12.38	08.00	0m
31248	Klintholm	1931-1990	54.95	12.47	08.00	0m
31308	Masnedø	1931-1937	55.00	11.88	08.00	0m
"	Storstrøm	1939-1988	54.97	11.88	08.00	0m
"	Farø	1988-1998	54.95	11.98	08.00	0m
31572	Rødbyhavn	1931-1999	54.65	11.35	08.00	0m
32002	Christiansø	1931-1998	55.32	15.18	08.00	0m
Danish automatic coastal stations						
23293	Fredericia	2002-2014	55.57	9.75	08.00	0m
26457	Fynshav	2002-2014	55.00	9.98	08.00	0m
28234	Slipshavn	1999-2015	55.28	10.83	08.00	0m
29393	Korsør	2000-2014	55.33	11.15	08.00	0m
30017	Hornbæk	2000-2015	56.10	12.47	08.00	0m
30336	København	1999-2015	55.70	12.60	08.00	0m
31573	Rødby	1999-2014	54.65	11.35	08.00	0m
31616	Gedser	2001-2014	54.57	11.93	08.00	0m
32048	Tejn	2001-2015	55.25	14.83	08.00	0m
Danish lightships						
6047	Læsø Trindel	1931-1943	57.47	11.34	07-08.00	0m
"	"	1943-1945	57.52	11.26	07-08.00	0m
"	Læsø Nord	1945-1975	57.53	11.34	07-08.00	0m
"	Læsø Trindel	1975-1977	57.47	11.42	07-08.00	0m
6057	Læsø Rende	1931-1943	57.21	10.69	07-08.00	0m
"	"	1943-1962	57.21	10.73	07-08.00	0m
"	"	1962-1965	57.20	10.73	07-08.00	0m
6067	Østre Flak	1931-1942	56.97	10.90	07-08.00	0m
"	Ålborg Bugt	1943-1973	56.85	10.80	07-08.00	0m

St. Num.	Station Name	Years	Lat.	Lon.	Time	Depth
Continuation - Danish lightships						
6077	Schultz's Grund	1931-1944	56.15	11.19	07-08.00	0m
"	Kattegat SW	1945-1971	56.10	11.15	07-08.00	0m
6087	Anholt Knob	1931-1945	56.77	11.86	07-08.00	0m
"	"	1945-1948	56.75	11.99	07-08.00	0m
"	Anholt Nord	1948-1975	56.85	11.80	07-08.00	0m
"	Anholt Knob	1975-1985	56.75	11.88	07-08.00	0m
6127	Halsskov Rev	1931-1973	55.34	11.05	07-08.00	0m
6147	Gedser Rev	1931-1939	54.45	12.18	07-08.00	0m
"	"	1945-1955	54.42	12.15	07-08.00	0m
"	"	1955-1976	54.45	12.18	07-08.00	0m
"	Kadetrenden	1976-1979	54.78	12.75	07-08.00	0m
"	Møn SE	1979-1988	54.80	12.78	07-08.00	0m
6157	Kattegat S	1966-1975	56.25	12.25	07-08.00	0m
6167	Lappegrund	1931-1969	56.07	12.63	07-08.00	0m
6183	Drogden Fyrskib	1900-1937	55.53	12.72	07-08.00	0m
"	Drogden Fyr	1937-1998	55.53	12.72	07-08.00	0m
Swedish maintained coastal stations						
35014	Varberg	1879-1887	57.10	12.22	08-09.00	0m
35009	Utklippan	1880-1922	55.95	15.70	08-09.00	0m
35007	Landsort	1880-1922	58.48	17.87	08-09.00	0m
Swedish lightships						
35016	Vinga	1930-1965	57.57	11.47	08.00	0m
35015	Fladen	1893-1969	57.17	11.67	08.00	0m
35013	Svinbådan	1884-1960	56.17	12.52	08.00	0m
35012	Kalkgrundet	1883-1922	55.62	12.88	08.00	0m
35011	Oskarsgrundet	1883-1961	55.58	12.85	08.00	0m
35010	Falsterborev	1860-1972	55.30	12.78	08.00	0m
35008	Ölands rev	1923-1951	56.12	16.57	08.00	0m
35006	Hävringe	1951-1967	58.55	17.52	08.00	0m
35005	Kopparstenarna	1883-1915	58.58	19.15	08.00	0m
35004	Svenska Björn	1883-1968	59.58	19.78	08.00	0m
35003	Grundkallen	1883-1960	60.50	18.92	08.00	0m
35002	Finngrundet	1860-1969	61.03	18.52	08.00	0m
35001	Sydostbrotten	1883-1963	63.32	20.17	08.00	0m
Swedish buoys						
35070	Trubaduren	1978-2004	57.60	11.79	08.00	2m
33001	Läså Ost	2001-2009	57.22	11.57	08.00	2m
35068	Fladen	1988-1999	57.22	11.83	08.00	2m
35067	Oskarsgrundet	1983-1999	55.60	12.85	08.00	2m
35063	Ölands Södra Grund	1979-1991	56.07	16.68	08.00	2m
35057	Gustav Dahlén	1982-1987	58.60	17.47	08.00	5m
35056	Almagrundet	1979-1992	59.15	19.13	08.00	2m
35054	Svenska Björn	1984-1992	59.47	20.35	08.00	2m

St. Num.	Station Name	Years	Lat.	Lon.	Time	Depth
German buoys						
-	Kiel	1987-2015	54.30	10.16	08.00	0m
-	Fehmarn Belt	1985-2012	54.36	11.09	08.00	1m
-	Darßer Schwelle	1998-2015	54.42	12.42	08.00	2m
-	Oder Bank	1997-2015	54.05	14.10	08.00	3m
-	Arkona Becken	2002-2015	54.53	13.52	08.00	2m
Finnish maintained coastal stations						
-	Bogskär	1899-1968	59.52	20.38	14.00	0m
-	Utö	1900-2005	59.78	21.34	14.00	0m
-	Märket	1906-1970	60.31	19.15	14.00	0m
-	Seili	1967-2012	60.25	21.96	14.00	0m
-	Valassaaret	1919-2009	63.39	21.08	14.00	0m
-	Krunnit	1968-2007	65.37	24.89	14.00	0m
-	Harmaja	1900-1993	60.10	24.96	14.00	0m
-	Tvärminne-A	1926-2012	59.85	23.25	14.00	0m
-	Tvärminne-B	1967-2012	59.85	23.25	14.00	0m
Finnish lightships						
-	Snipan	1900-1960	63.43	20.67	14.00	0m
-	Kemi	1900-1974	65.35	24.35	14.00	0m
-	Porkkala	1899-1968	59.93	24.42	14.00	0m

Table 6.1: Information on the in situ stations. The depths and times correspond to the depths and times of the data used in this thesis. All times are given as local times.

7 Calculating monthly means

As the aim of this thesis is to reconstruct monthly mean SST fields, the in situ records must be converted to correspond to monthly mean SST records. The methods, assumptions and criteria used to calculate the monthly mean values from the Danish, Swedish and German records are described in Section 7.1. As the sampling rates at the Finnish stations are much lower than at the Danish, Swedish and German stations, the Finnish monthly means have to be calculated based on other criteria. Two different methods are tested, which are described in Section 7.2. Finally, all records that have a temporal overlap with the CDR are compared to the CDR. The result of the comparison is shown in Section 7.3.

7.1 Danish, Swedish and German records

The first step in calculating monthly means is to extract the data on which the monthly means are going to be based. This data should be chosen such that the in situ records are consistent with each other as well as the CDR in regard to depth and time. As nearly all stations have registered a 0 m temperature at 08:00 LT, and as these values are fairly consistent with the CDR, all data corresponding to 08:00 LT and 0 m were extracted from the records. Furthermore, the most shallow sensor of some buoys was located at deeper depths. In result, the data extracted from buoys correspond to depths between 0 m and 5 m. Similarly, some Danish lightships have registered the temperature at 07:00 LT instead of 08:00 LT and some Swedish maintained coastal stations have observed the temperature at 09:00 LT instead of 08:00 LT. The temperatures registered at these times were therefore used equivalent to the 08:00 LT temperatures. Furthermore, the monthly means of Læsø Trindel, Anholt Knob and Gedser Rev/Kadetrenden are *not* based on the 5 m temperatures when the 0 m temperatures are lacking in 1976 as this would cause an inconsistency within the records. The times and depths extracted from each record are shown in Table 6.1.

The second step is to carefully view all data sets such that outliers can be removed. These values are found by calculating a climatology for each in situ record. The climatologies are generated by averaging over all SSTs corresponding to the same month. In result, the climatologies consist of 12 values, which represent the average SST of the record for each discrete month. All values differing with more than 3 standard deviations from the climatology are marked so that they can be examined. Of these, only values differing significantly from the overall trend and neighboring values are discarded. As an example, Figure 7.1 shows the SST recorded at Frederikshavn between the 14th of February 1978 and the 16th of February 1981. Here, the temperatures recorded at the 12th of October 1979 and the 31st of October 1979 are found to differ with more than 3 standard deviations from the monthly climatology. These observations are therefore marked (shown in blue) and the overall trend and neighboring values are examined to determine if they should be discarded or not. As the SST registered on the 12th of October 1979 was found to be 1.0°C while the two neighboring values showed 12.2°C (11th of October 1979) and 11.2°C (15th of October 1979), this observation was classified as an outlier. In contrast, the value of 5.2°C registered on the 31st of October 1979 followed the overall trend and was not found to differ significantly from the neighboring values (5.8°C on the 30th of October 1981 and 5.0°C on the 2nd of November 1981) and was thus kept. The evaluation procedure was performed manually for all in situ records.

After all outliers have been removed, monthly means are calculated from the remaining observations. A monthly mean is only calculated for months that contain at least 15 days with data. The monthly mean of months containing less than 15 observations are set to `NaN`.

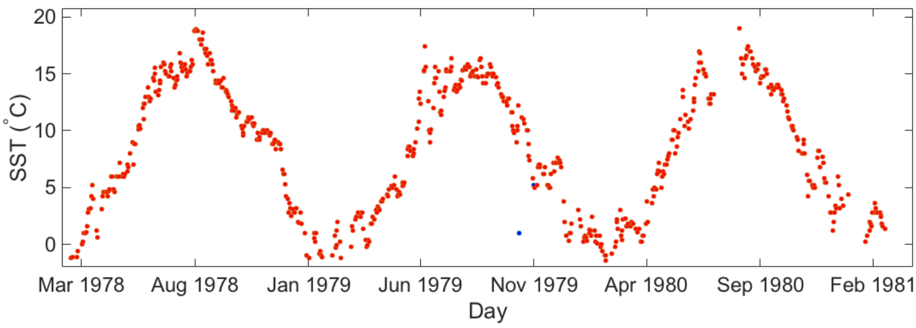

Figure 7.1: SST observed at Frederikshavn between the 14th of February 1978 and the 16th of February 1981. All observations were made at 08:00 LT and at a 0 m depth. The blue dots represent the values that differ with more than 3 standard deviations from the climatology.

7.2 Finnish records

At the Finnish in situ stations, the surface temperatures are observed at 14:00 LT and all measurements correspond to the upper half meter of the surface. In contrast to the other stations, measurements are only made approximately 3 times a month. It is therefore not possible to calculate monthly means based on the criteria described in Section 7.1. Due to the sparse data, two different methods are tested to try to yield as accurate monthly means as possible. The methods are described in Section 7.2.1 and Section 7.2.2.

7.2.1 Method 1

In the first method, monthly means are calculated in the same way as in Section 7.1, but with the constraint that the months needs to contain at least 3 measurements in order for a monthly mean to be calculated. This is assumed to give a somewhat accurate representation of the monthly mean as the measurements generally are evenly distributed throughout the month (samples are usually taken around the 1st, 11th and 21st of each month). As the data are very sparse, it is difficult to detect outliers and all observations are therefore kept.

7.2.2 Method 2

The second set of monthly means are calculated by fitting a sinusoidal function to the time series using harmonic analysis. The main advantage of this method is that the harmonic analysis takes the timing of the measurements into account. In result, it is possible to estimate daily SSTs for the entire time series and thereby also for months containing less than 3 observations. The monthly means can then be calculated by extracting the daily values from the fitted sinusoidal function. The theory behind the method is described in the subsection 'Harmonic analysis' while the application on the Finnish records is described in the subsection 'Harmonic analysis of the Finnish in situ records'.

Harmonic analysis

Harmonic analyses are used to examine well-defined cyclic oscillations (such as annual or semi-annual oscillations) in overdetermined problems (Emery and Thomson, 1997). In other words, harmonic analyses are used to determine the amplitudes and phase lags of oscillations with well-known frequencies in problems in which the data series contain more observations than the number of prescribed frequencies. After the amplitudes and phase lags have been found, it becomes possible to reconstruct the original signal. The derivation of the theory is based on Emery and Thomson (1997) and thereby follows their notations.

Let's say that M specified frequencies are examined in a time series. It is then the easiest to examine the time series when it is expanded in its harmonic constituents, such that $q = 0, 1, ..., M$ specifies the frequency components:

$$x(t_i) = \bar{x} + \sum_{q=1}^{M} C_q \cos(2\pi f_q t_i - \phi_q) \tag{7.1}$$

Here, $x(t_i)$ represents a time series with N values where $i = 1, 2, ..., N$, and in which the mean value is given by $\bar{x}$. The amplitude, frequency and phase lag of component q is given by C_q, f_q and ϕ_q, respectively. As the frequency components already are prescribed, the amplitudes, phase lags and mean value must be estimated in order to reconstruct the original signal. These are usually estimated from another form of Equation 7.1, namely from the Fourier series:

$$x(t_i) = \bar{x} + \sum_{q=1}^{M} \left[A_q \cos\left(t_i \frac{2\pi}{T_q}\right) + B_q \sin\left(t_i \frac{2\pi}{T_q}\right) \right] \tag{7.2}$$

Here, T_q is the period of component q. The period relates to the frequency according to $T_q = 1/f_q$. A_q and B_q are the Fourier coefficients of component q. These are related to the amplitude and phase lag of component q according to $C_q = (A_q^2 + B_q^2)^{1/2}$ and $\phi_q = \tan^{-1}(B_q/A_q)$. The Fourier coefficients and the mean value should be estimated such that they apply for the entire time series:

$$\begin{bmatrix} x(t_1) \\ x(t_2) \\ \vdots \\ x(t_N) \end{bmatrix} = \begin{bmatrix} 1 & \cos(t_1 \frac{2\pi}{T_1}) & \sin(t_1 \frac{2\pi}{T_1}) & \dots & \cos(t_1 \frac{2\pi}{T_M}) & \sin(t_1 \frac{2\pi}{T_M}) \\ 1 & \cos(t_2 \frac{2\pi}{T_1}) & \sin(t_2 \frac{2\pi}{T_1}) & \dots & \cos(t_2 \frac{2\pi}{T_M}) & \sin(t_2 \frac{2\pi}{T_M}) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \cos(t_N \frac{2\pi}{T_1}) & \sin(t_N \frac{2\pi}{T_1}) & \dots & \cos(t_N \frac{2\pi}{T_M}) & \sin(t_N \frac{2\pi}{T_M}) \end{bmatrix} \begin{bmatrix} \bar{x} \\ A_1 \\ B_1 \\ \vdots \\ A_M \\ B_M \end{bmatrix} \tag{7.3}$$

However, as actual time series rarely correspond to perfect sinusoidal functions, it is not possible to find coefficients that will satisfy Equation 7.3 completely.

Instead, the coefficients are estimated such that the difference between the reconstructed sinusoidal function and the actual data becomes as small as possible. The difference is usually expressed by the residual vector $\mathbf{r}$ according to:

$$\mathbf{r} = \mathbf{d} - \mathbf{G}\mathbf{m} \tag{7.4}$$

Here, $\mathbf{d}$ is the vector containing the measured data, $\mathbf{G}$ is the matrix containing the sine and cosine terms, and $\mathbf{m}$ is the vector containing the model parameters ($\bar{x}$, A_q and B_q). The number of model parameters are given by $2M + 1$. As Equation 7.3, corresponds to a discrete (finite amount of model parameters and observations) and overdetermined ($N > 2M + 1$) inverse problem, the least squares solution is the statistically most likely solution (Aster et al., 2013). Such solution is found by minimizing the 2-norm of the residual vector and is given by:

$$\mathbf{m}_{L_2} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d} \tag{7.5}$$

The model coefficients obtained from Equation 7.5 are inserted into Equation 7.2, which yields a fitted sinusoidal function for the time series.

Harmonic analysis of the Finnish in situ records

In the Finnish time series, the prescribed frequency corresponds to the annual oscillation with a period of $T_1 = 365.25$ days. To allow for asymmetries between the different halves of the year, the fit is also based on a semi-annual oscillation with a period of $T_2 = 185.625$. As the temperature development varies between years, each year is fitted separately.

Monthly means are calculated from the sinusoidal fits by extracting the daily SSTs obtained for each month. However, as some of the sinusoidal fits are based on very few observations, there still needs to be some limitations. First of all, a monthly mean is only calculated when there is at least 10 measurements in the five closest months. Secondly, a monthly mean is not calculated for the warmest months (July and August) or coldest months (February and March) if data from these months are missing, as this affects the amplitude of the signal. Finally, as the fits are made on a yearly basis, there is often a discontinuity between December and January. To avoid the discontinuity from affecting the signal, the monthly mean for January is calculated by averaging the monthly means of December and February. An example of the sinusoidal fit is shown in Figure 7.2.

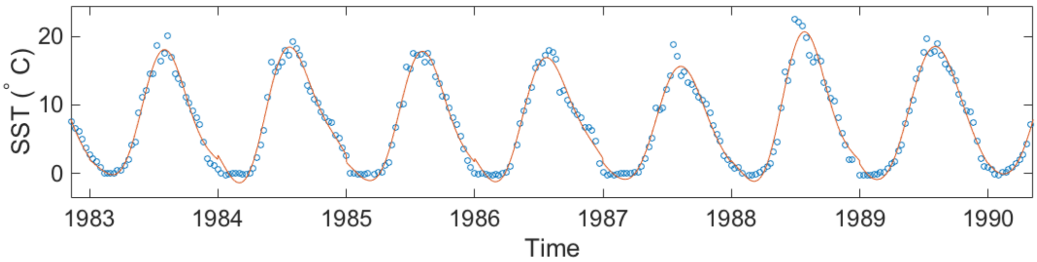

Figure 7.2: Sinusoidal fit (red curve) of the SST (blue dots) measured at Seili between 1983 and 1990.

7.3 Comparison against CDR

The monthly mean in situ records are compared against the CDR in order to check the consistency between the time series. As the CDR spans from 1982 to 2011, it is only possible to make the comparison for records covering the same period (partly or fully). For each of these in situ records, the time series in the closest CDR grid cell is extracted. The consistency between the time series is then checked by calculating the root mean square error (RMSE), mean bias (Bias), standard deviation (Std), minimum difference (Min) and maximum difference (Max) between the series. All calculations are made by subtracting the monthly means of the in situ records from the CDR. A positive bias therefore indicates that the CDR is warmer than the in situ record. The result of the comparison is shown in Table A.1, where N represents the number of match-ups on which the comparisons are based.

Of the statistics listed in Table A.1, the RMSE is especially useful to check the overall agreement between the records. Denoting two time series as x and y, the RMSE is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x(t_i) - y(t_i))^2} \tag{7.6}$$

From Equation 7.6, it can be seen that the RMSE approaches zero as the difference between the two time series approaches zero. A small RMSE therefore represents a good agreement between the two time series. From Table A.1, it can be seen that the RMSE generally is lower than 1.0°C for the in situ records used in this study. While only 9 out of the 42 Danish, German and Swedish in situ records have a RMSE larger than 1.0°C, 5 out of the 7 Finnish in situ records have a RMSE greater than 1.0°C (regardless of the method). Tvärminne-B even has a RMSE of 2.0°C for both methods. Table A.1 thereby confirms that the monthly means obtained from the Finnish in situ records are less accurate than the ones obtained from the Danish, German and Swedish records. Besides the low sampling rate, the large RMSE could be the result of the CDR and Finnish in situ records representing different times (the CDR is a night product while the Finnish in situ records are based on SSTs measured at 14:00 LT).

In order to check the temporal stability between the in situ records and the CDR, monthly and yearly mean differences are calculated. The result is shown in Figure 7.4, where the left and right column corresponds to the monthly mean bias and yearly mean bias, respectively. To distinguish the more uncertain Finnish records from the Danish, German and Swedish records, the comparison is made separately for the Finnish records. In Figure 7.4, green represents the Finnish records obtained from Method 1, red represents the Finnish records obtained from Method 2 and blue represents the Danish, German and Swedish records.

Figure 7.4 demonstrates that the monthly means obtained from the Finnish records are less stable with time than those obtained from the Danish, German and Swedish records. The standard deviations are also larger, indicating larger uncertainties in the monthly means. In contrast, it can be seen that the Danish, German and Swedish in situ records are very stable with time. The yearly mean bias is generally smaller than 0.1°C while the standard deviation generally is lower than 1.0°C. The monthly mean bias demonstrates that the records also are stable with respect to the season. The first six months of the year show a slight negative bias while the latter six months show a slight positive bias. The standard deviation is the greatest in May and June, indicating the largest uncertainties in these months.

To examine if the small seasonal bias seen in Figure 7.4 could be a result of the sampling technique, the same comparison as in Figure 7.4 is made separately for each station type (only including Danish, German and Swedish records). The result is shown in Figure 7.3, where lightships are represented with black, maintained coastal stations with blue, buoys with red and automatic coastal stations with green. Here, it can be seen that the seasonal bias changes with station type. The buoys and lightships show similar seasonal biases, with a negative bias during winter and a positive bias during summer. The coastal stations also show similar biases, where the largest negative bias is found during spring while the maximum positive bias is found during fall.

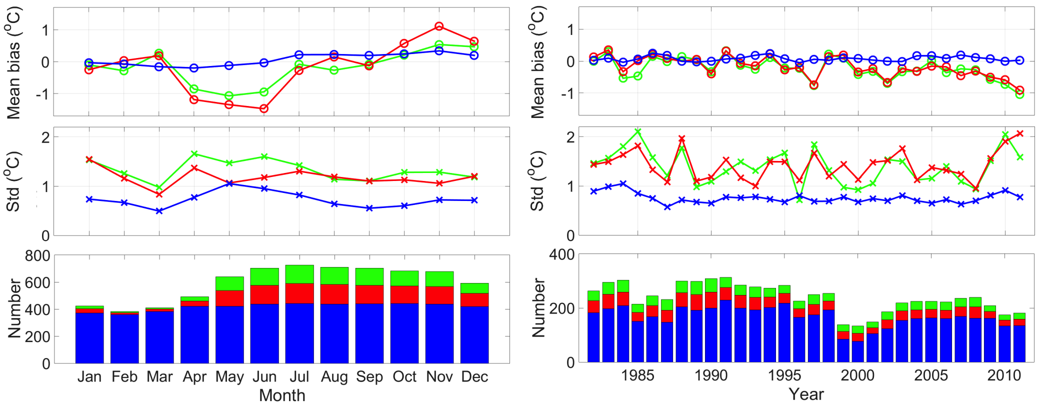

Figure 7.3: Monthly mean difference (left column) and yearly mean difference (right column) between the CDR and in situ records. The top graphs show the mean difference for the specified months or years, the middle graphs show the corresponding standard deviations and the bottom graphs show the number of match-ups on which the calculations were based. In all graphs, green represents the Finnish records obtained from Method 1, red represents the Finnish records obtained from Method 2 and blue represents the Danish, German and Swedish records.

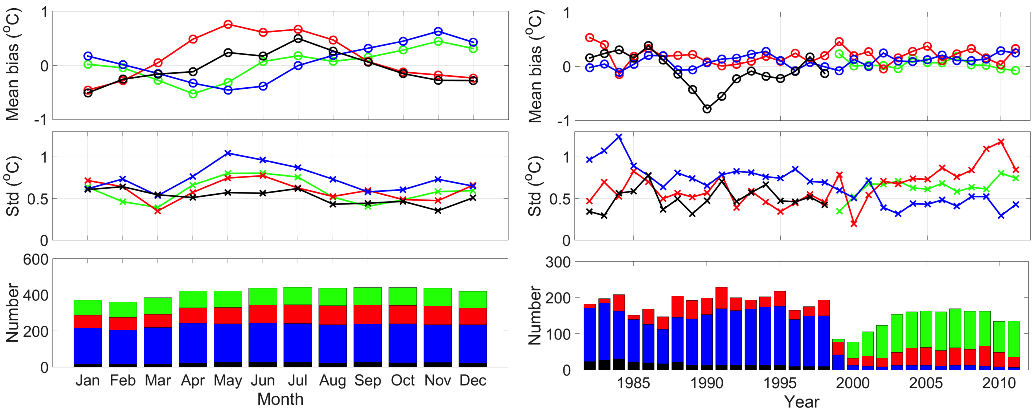

Figure 7.4: As in Figure 7.3, but where green represents automatic coastal stations, red represents buoys, blue represents maintained coastal stations and black represents lightships. Only Danish, German and Swedish records are included in the analysis.

The accuracy of the two sets of Finnish records are further examined by calculating the correlation coefficient (ρ) between the records and the CDR. The correlation analysis is described in detail in Chapter 8. Here, it is enough to know that the correlation coefficient describes the covariability between two time series, where a coefficient of 1 indicates a completely positive linear relationship between the series. The record in which the correlation coefficient is the closest to 1 is thereby the record which follows the temporal development of the CDR the best. The correlation coefficient of each record can be viewed in Table 7.1, together with the RMSE and the number of monthly means (Num) generated by each method.

With the exception of Krunnit, the records obtained from Method 1 are better correlated with the CDR. Similarly, the RMSE is lower for all records obtained from Method 1 except for Krunnit. The monthly means obtained from Method 1 are therefore assumed to be the most accurate. Furthermore, it can be seen that Method 2 often generates a larger number of monthly means. As seen in Chapter 9 and Chapter 10, Utö and Seili will be used for the monthly mean SST reconstructions. It is therefore important that they contain as many observations as possible. From Table 7.1, it can be seen that Seili gains 90 extra values if Method 2 is used instead of Method 1. Similarly, Utö gains 195 extra values. From Table 7.1, it can also be seen that the RMSEs and correlation coefficients only become slightly worse when Method 2 is used instead of Method 1. In the remaining of the report, the Seili and Utö records obtained from Method 2 are therefore used, while the other records are obtained form Method 1.

	Method 1			Method 2		
Station Name	ρ	RMSE (°C)	Num	ρ	RMSE (°C)	Num
Tvärminne-A	0.73	1.45	939	0.67	1.61	990
Tvärminne-B	0.79	2.00	530	0.71	2.00	527
Harmaja	0.68	1.20	852	0.46	1.33	971
Seili	0.83	1.22	478	0.77	1.28	568
Utö	0.77	0.67	476	0.73	0.83	671
Valassaaret	0.84	0.84	696	0.83	0.87	822
Krunnit	0.65	1.76	61	0.78	1.33	66

Table 7.1: Comparison of the accuracy of the two methods used to calculate monthly means from the Finnish in situ data.

8 Correlation analysis

The goal of this thesis is to combine the CDR and in situ records in a multivariate regression model to yield historical SST fields. To obtain good model performance, it is a requirement that the CDR is highly correlated within some distance of the in situ stations (Høyer, 2002). The correlation between the CDR and in situ records is therefore examined to ensure that the most optimal records are included in the model. Furthermore, none of the in situ records are long or continuous enough to be used directly. Instead, long and continuous records have to be constructed by combining observations of adjacent stations (see Chapter 9). To achieve accurate and consistent records, it is important that the constructions are based on highly correlated records. In addition, correlation analyses can be used to reveal 'odd behaving' in situ records and to examine the expected spatial patterns of the SST variations. Correlation analyses therefore serve multiple important purposes in this thesis.

In total, three different types of correlation analyses are made; the in situ records are correlated against each other, the in situ records are correlated against the CDR, and selected CDR time series are correlated against the CDR. The analyses are reviewed in Section 8.2 - 8.4, respectively. First, the theory behind the correlation analysis is described.

8.1 Theory

Correlation analyses are used to examine the covariability between time series with respect to time. To follow the notations used by Emery and Thomson (1997), the examined time series are denoted x and y. The correlation coefficient of the time series is calculated after the mean values, $\bar{x}$ and $\bar{y}$, have been subtracted from the time series. If time series y consists of N values, the mean value is estimated as the sample mean according to:

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{8.1}$$

However, as this thesis deals with time series of monthly mean SSTs, there is a strong seasonal cycle inherited in the time series. This means that any two records will have a high correlation after the mean value (which corresponds to the yearly mean SST) has been subtracted from the records. Instead, $\bar{x}$ and $\bar{y}$

must be calculated for each month individually in order to remove the seasonal cycle from the measurements. The covariance function can then be calculated from the original time series, but where the mean value depends on the month m in which value i was observed. The covariance function, C_{xy}, can then be calculated as:

$$C_{xy} = \frac{1}{N} \sum_{i=1}^{N} [y_i(m) - \bar{y}(m)] [x_i(m) - \bar{x}(m)] \tag{8.2}$$

The correlation coefficient, ρ_{xy} is obtained after the covariance function has been normalized using the standard deviations of the time series:

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \tag{8.3}$$

The correlation coefficient can obtain values between -1 and 1. If there is a complete positive linear relationship between two time series, a correlation coefficient of 1 is found. Similarly, a complete negative linear relationship results in a correlation coefficient of -1. If there is absolutely no linear relationship between the two time series, the correlation coefficient becomes 0.

In practice, Equations 8.1 - 8.3 are only applied to the part of time series x and y where there is a temporal overlap. As all anomalies would equal zero if only data from one year is used, the temporal overlap needs to cover at least two years. The correlation coefficient of any two data sets can then easily be found using MATLAB's function `corr`. As both the in situ records and the CDR contain `NaN`s (due to the presence of sea ice or data gaps), the function is run with the addition `pairwise` to ensure that match-ups containing `NaN` are excluded from the calculation.

8.2 Correlation between in situ records

All in situ stations are assigned an index according to Table B.1. The indexes span from 1 to 78, where stations assigned 1-16 are situated in Kattegat, 17-36 in the Belt Sea, 37-46 in Öresund, 47-51 in the Arkona Basin, 52-54 in the Bornholm Basin, 55-65 in the Western Gotland Basin, 66-69 in the Gulf of Finland and 70-78 in the Gulf of Bothnia. The exact position of the stations can be viewed in Figure 8.1.

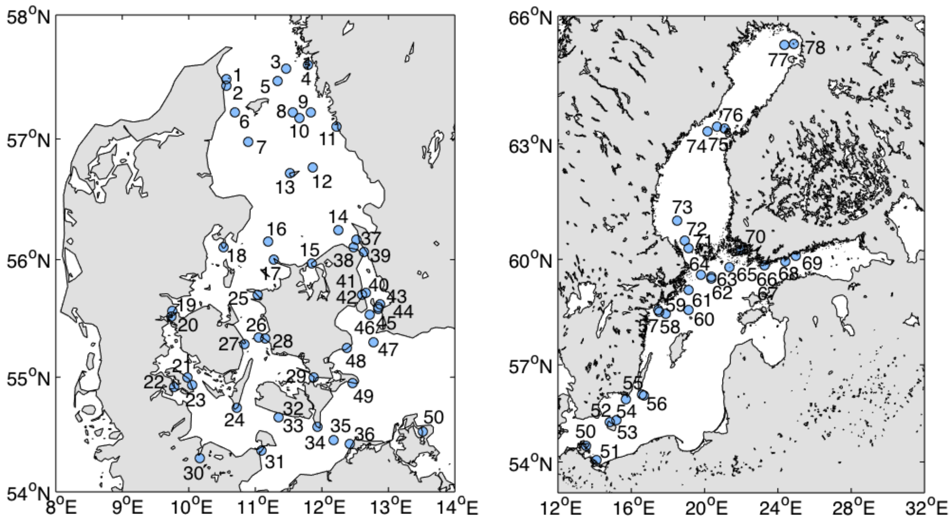

Figure 8.1: Positioning of the stations used in this study. Indexes are assigned according to Table B.1. Note that station 50 is shown on both maps, and that station 32 and 33, 41 and 42, and 66 and 67 are located at the same position.

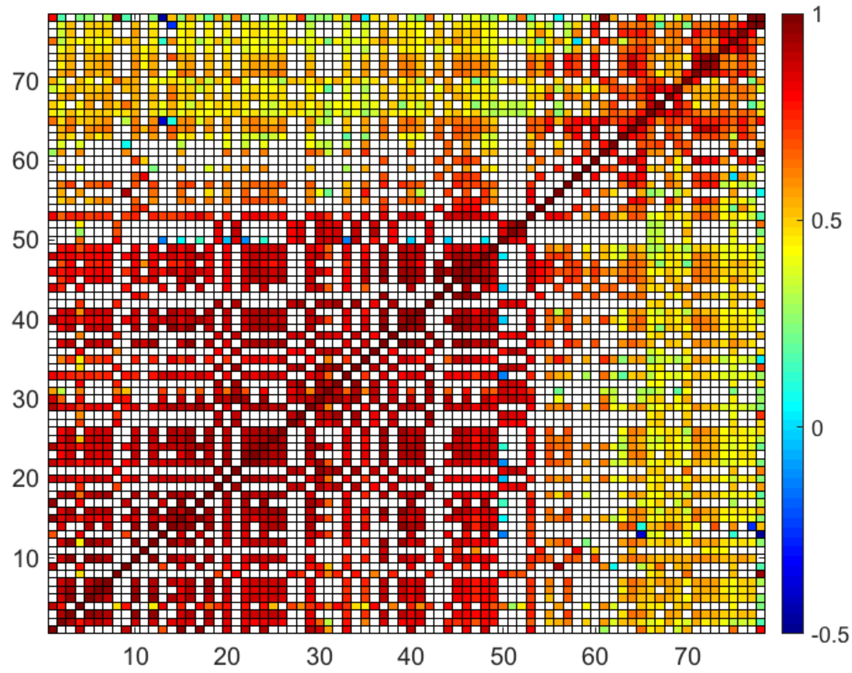

Figure 8.2: Correlation between each set of in situ stations. Indexes are assigned according to Table B.1. White represent a temporal overlap of less than 24 months.

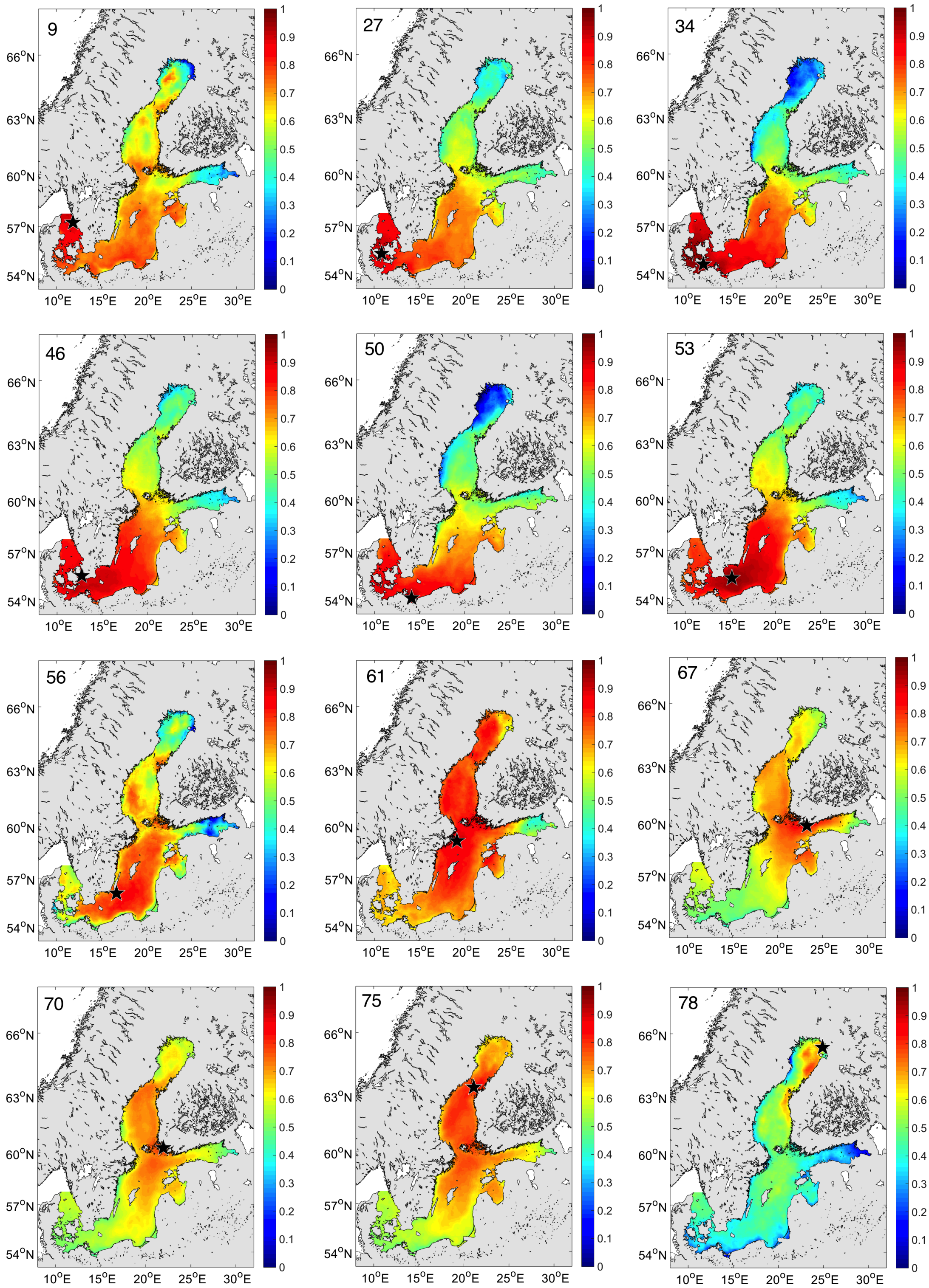

Figure 8.3: Correlation between 12 selected in situ records and the CDR. Indexes are assigned according to Table B.1. The position of the stations are shown by the black stars.

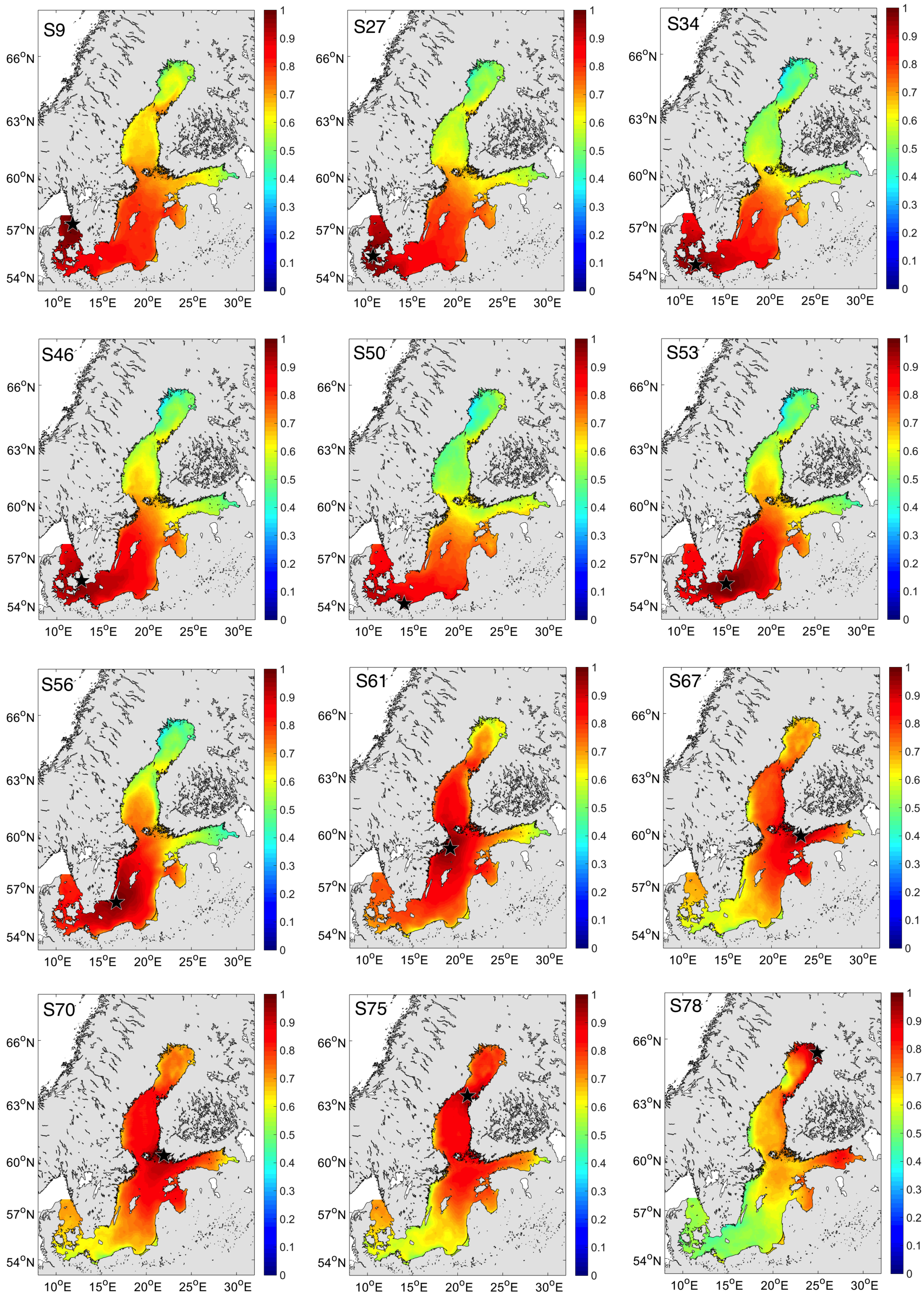

Figure 8.4: Correlation between the CDR time series corresponding to the 12 in situ stations shown in Figure 8.3 and the CDR. The position of the time series are shown by the black stars.

Each in situ record is correlated against all other in situ records. In result, each match-up is assigned a correlation coefficient between -1 and 1. The correlation coefficients are shown in Figure 8.2, where the color of row i and column j corresponds to the correlation between station i and j. It follows from Equation 8.2 and 8.3 that the matrix is symmetric. Thus, the same coefficient is found at row j and column i. Correlation coefficients based on less than 24 match-ups are set to `NaN`, which is shown as white in Figure 8.2. Note that the axis of Figure 8.2 spans from -0.5 to 1. The lower boundary is set to -0.5 as the lowest correlation was found to be -0.47, corresponding to the correlation between Anholt (13) and Krunnit (78).

From Figure 8.2, it can be seen that adjacent stations generally have a high correlation and that the correlation decreases with distance. Furthermore, it can be seen that Figure 8.2 can be divided into smaller areas with similar characteristics. For example, it can be seen that all stations in the Transition Zone (indexes 1-46) are highly correlated. In the Transition Zone, Trubaduren (4) and Fehmarn Belt (31) have the lowest correlations with the other stations in the area. While Anholt (13) generally have high correlations with the other stations in the Transition Zone, it has the lowest correlations with the northern Baltic stations.

Stations located in the Arkona and Bornholm Basin (indexes 47-54) correlate well with stations in the Transition Zone. However, Oder Bank (50) stands out. From Figure 8.2, it can be seen that Oder Bank varies between having high (0.65 - 0.90) and low (-0.14 - 0.14) correlations with stations in the Transition Zone. The correlations are therefore examined in greater detail, where it is found that all low values correspond to correlations against maintained coastal stations (and one buoy), while all high values correspond to correlations against automatic stations and buoys. As similar trends are not shown by any other record, it is not believed to be caused by a discrepancy between station types. Instead, it can be seen that all low correlations correspond to records that end in 1999. As Oder Bank started recording temperatures in 1997, the low correlation coefficients are assumed to be caused by the analysis being based on too few observations (too short temporal overlaps). This could therefore indicate that a temporal overlap of at least 36 months is needed in order to get valuable information from the correlation coefficients.

Moving northward through the Western Gotland Basin (indexes 55-65), the correlation against stations in the Transition Zone slowly decreases while the correlation against stations in the Gulf of Bothnia and the Gulf of Finland increases. Especially Svenska Björn (64) and Utö (65) seem to correlate well with all stations north of the Bornholm Basin. In the Gulfs, the correlation between adjacent stations is still high, while the overall correlations are lower than in the Transition Zone.

8.3 Correlation between in situ records and CDR

All in situ records that have a temporal overlap of at least 24 months with the CDR are correlated against the CDR. Correlations between the in situ records and CDR are useful as it reveals how spatially representative the in situ records are. A selection of the correlations are shown in Figure 8.3. These were selected to represent different regions of the Baltic Sea.

In Figure 8.3, it can be seen that the correlation coefficient is the highest in the proximity of the in situ stations and decreases with distance from the stations. The correlation patterns obtained by Fladen (9), Slipshavn (27), Gedser (34), Drogden (46), Oder Bank (50) and Christiansø (53) are quite similar. All show high correlations in the Transition Zone and in the majority of the Baltic Proper, with decreasing correlations in the Gulf of Bothnia and the Gulf of Finland. In general, Ölands Södra Grund (56), Almagrundet (61), Tvärminne-B (67), Seili (70), Valassaaret (75) and Krunnit (78) show much lower correlations in the proximity of the stations. This was expected for the Finnish stations due to the large uncertainties in the records. The reason for the relatively low correlations of Ölands Södra Grund and Almagrundet are not known.

Although the correlation in the proximity of Almagrundet was lower than expected, Almagrundet is still found to have relatively high correlations (over 0.7) in the majority of the Baltic Proper and the Gulf of Bothnia. This is in accordance to the results in Section 8.2, where Svenska Björn and Utö were found to show similar characteristics. The northwestern Baltic Proper is thereby found to represent the SST variations of a large extent of the Baltic Sea rather well. It is therefore desirable to include in situ records from the northwestern Baltic Proper in the multivariate regression model. Furthermore, it should be mentioned that Almagrundet showed better spatial representativeness than Svenska Björn and

Utö in the correlations with the CDR.

The correlation between the Finnish in situ records and the Gulf of Bothnia and the Gulf of Finland was expected to be higher. It was especially disappointing to see that all records were poorly correlated with the eastern part of the Gulf of Finland. Krunnit is found to have especially poor correlations with the CDR. This is believed to be caused by the great discontinuity (extremely many gaps) in the Krunnit record, leading to the correlation between 1982 and 2007 only being based on 35 match-ups.

8.4 Correlation between CDR time series

The CDR time series corresponding to the stations investigated in Section 8.3 are also correlated against the CDR. This makes it possible to determine whether the spatial structures shown in Figure 8.3 were expected. The result is shown in Figure 8.4. Here, an S prior to the station index represents the CDR time series corresponding to the in situ record with the same index. When comparing Figure 8.3 and 8.4, it can be seen that the correlation patterns between the in situ records and the CDR generally are more patchy, and that the overall correlations are lower. The differences are most likely explained by the inclusion of different time periods, sampling rates and sampling depths.

From Figure 8.4, it can be seen that the correlation patterns obtained from the time series corresponding to Fladen (S9), Slipshavn (S27), Gedser (S34), Drogden (S46), Oder Bank (S50), Christiansø (S53) and Ölands Södra Grund (S56) are quite similar to those obtained by the corresponding in situ records. The largest differences are found in the Gulf of Bothnia and the Gulf of Finland, where the correlations with the CDR time series generally are higher. Furthermore, the relatively high correlation between Fladen (9) and the Gulf of Bothnia are also found between the CDR time series corresponding to Fladen (S9) and the Gulf of Bothnia.

The correlation patterns obtained from the time series corresponding to Ölands Södra Grund (S56), Almagrundet (S61) and Valassaaret (S75) are also found to be quite similar to those obtained by the corresponding in situ records, although the correlations between the CDR time series and CDR are much higher. As expected, the correlation patterns of Tvärminne-B (67), Seili (70) and Krunnit (78) are greatly underestimated compared to the corresponding correlation patterns of the CDR time series. Figure 8.4 demonstrates that the correlation patterns of Tvärminne-B and Seili should have been similar to that of Valassaaret (75). It can also be seen that Krunnit is the station which should be able to represent the northernmost extension of the Gulf of Bothnia and the easternmost extension of the Gulf of Finland the best.

From Figure 8.3 and 8.4, it can be concluded that evenly distributed in situ records have to be included in the multivariate regression model to obtain good model performance for the entire Baltic Sea. As the highest correlations are found in a localized area around the in situ stations, it would be an advantage to include a dense network of in situ records. Furthermore, due to the large spatial extent of relatively high correlations, it is expected that the inclusion of just a handful of records will be sufficient.

9 Constructing long and continuous records

The multivariate regression model is only able to reconstruct monthly mean SST fields for months in which all the included in situ records contain data (Chapter 10). It is therefore important that the model is based on long and continuous in situ records. Furthermore, the description of the in situ records in Chapter 6 reveals that none of the in situ records are long or continuous enough to be used directly in the model. Instead, the desired records have to be constructed by combining shorter records obtained by adjacent stations. In total, six data sets are constructed. The method used to construct the long and continuous records is described in Section 9.1, while detailed information about the specific constructions is found in Section 9.2. The accuracy and spatial representativeness of the long and continuous records are examined in Section 9.3.

9.1 Method

The first step in creating long and continuous data sets is to choose the records on which the data sets are going to be based. When choosing the records, there are several important aspects that needs to be taken into account:

1. When possible, the records should be chosen such that they have temporal overlaps. This makes it possible to check and correct for biases, and to make sure that the chosen records are highly correlated.

2. The constructions should be based on the smallest number of records as possible as this leaves a greater amount of records for validation.

3. At least one of the chosen records must be relatively long and continuous such that it can be used as a reference record.

4. The reference records should be evenly distributed throughout the Baltic Sea.

It is emphasized that a great amount of work has been put in to analyzing all of the records such that the most optimal combinations were chosen.

After the records were chosen, the long and continuous data sets were constructed by extending and filling in gaps in the reference records. In order to fill in gaps or extend the reference record using data obtained at another location (and possibly at another depth), the relation between the two records must be known. Such relation is easy to find when there is a temporal overlap between the records, whereas the relation becomes more uncertain without a temporal overlap. As a result, a slightly different method has to be used when combining records without a temporal overlap than when combining records with a temporal overlap. The method used to combine records with a temporal overlap is described in section 9.1.1 and the technique used to combine records without a temporal overlap is described in section 9.1.2.

9.1.1 Combining records with temporal overlaps

When two different stations have measured SST over the same period, it is possible to find a relationship between the recorded temperatures. In the construction of the long and continuous records, the stations are located close to each other and the recorded temperatures are highly correlated, which makes it safe to assume that the relationship is linear. This means that the temperatures measured at the reference station can be determined from the temperatures measured at the extending station through the relationship:

$$d(t_n) = m_1 + g(t_n)m_2 \tag{9.1}$$

Here, $d(t_n)$ represents the SST measured at the reference station at time t_n while $g(t_n)$ represents the SST measured at the extending station at time t_n, where $n = 1, 2, ..., N$ and N is the number of match-ups. The two linear regression parameters m_1 and m_2 determine the relationship between the recorded temperatures. The two linear parameters are assumed to be constant in time and should therefore apply at all times according to:

$$\begin{bmatrix} d(t_1) \\ d(t_2) \\ \vdots \\ d(t_N) \end{bmatrix} = \begin{bmatrix} 1 & g(t_1) \\ 1 & g(t_2) \\ \vdots & \vdots \\ 1 & g(t_N) \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} \tag{9.2}$$

However, as there exist noise in the data and as it is very unlikely that the true relationship is completely linear, it is not possible to find two model parameters that will satisfy Equation 9.2 exactly. Instead, the parameters are found according to the inverse method described in Chapter 7. That is, the model parameters are found such that the residual between the predicted ($\mathbf{Gm}$) and actually measured ($\mathbf{d}$) temperatures becomes minimized. In Equation 9.2, $\mathbf{d}$ is the vector containing the temperatures measured at the reference station, $\mathbf{G}$ is the matrix containing the temperatures measured at the extending station and $\mathbf{m}$ is the vector containing the two linear regression parameters. The solution that minimizes the 2-norm of the residual is called the least squares solution, and is given by:

$$\mathbf{m}_{L_2} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d} \tag{9.3}$$

As the linear regression parameters are constants, the reconstructed temperatures are only dependent on the temperatures of the extending records. However, when examining the data sets, the relationship is found to vary with season. As an example, the monthly mean SSTs registered by Drogden and Falsterborev between January 1930 and January 1932 are shown in Figure 9.1. In the figure, it can be seen that Drogden (blue) usually records warmer SST than Falsterborev (red) during spring (April-June) and colder SST during autumn (October-December). As the same span of temperatures are found during spring and autumn, a seasonal dependency must be added. In order to include such seasonal dependency, Equation (9.3) is applied to data from every month separately such that a specific set of linear regression parameters are found for each month.

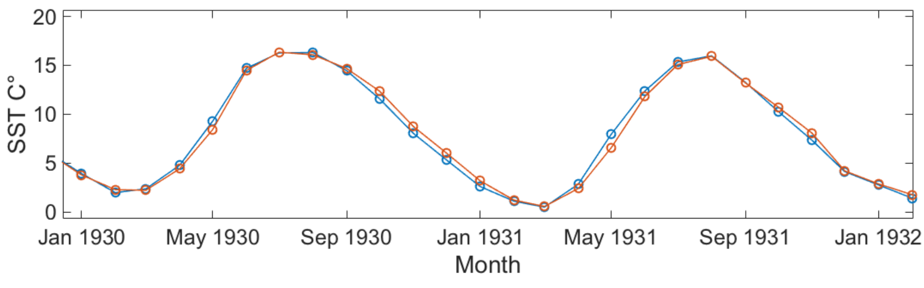

Figure 9.1: Monthly mean SST at Drogden (blue) and Falsterborev (red) between January 1930 and January 1932.

Similarly, if there is found to be an obvious drift in the relationship between two records, different model parameters are calculated for different time periods. As an example, the comparison of Falsterborev and Drogden is considered. The Drogden record is long and nearly gap free, and is therefore used as a reference record. Falsterborev is used to fill in gaps and extend the Drogden record back in time. As Falsterborev was inactive during the First and Second World War, the relationship between the two stations can be examined for three different periods; January 1900 - October 1914, December 1919 - December 1939 and January 1950 - November 1972.

In Figure 9.2 it is possible to view the difference in the monthly mean SST recorded by Drogden and Falsterborev. Obviously, the relationship between the two data sets has changed between the different periods. In order to increase the accuracy of the reconstructed temperatures, model parameters calculated from the first period are used to reconstruct Drogden temperatures up to the First World War, model parameters calculated from the second period are used to reconstruct Drogden temperatures between the First and Second World War, and model parameters calculated from the third overlapping period are used to reconstruct Drogden temperatures after the Second World War. Similarly, different parameters are calculated for different periods when stations have been moved. In the absence of moves and obvious drifts, the monthly model parameters are assumed to stay constant with time.

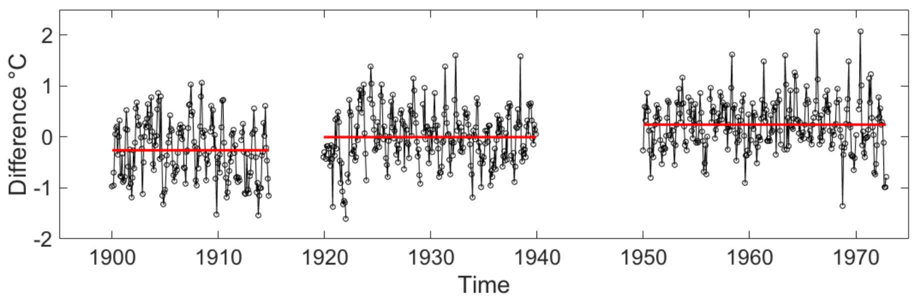

Figure 9.2: Comparison of the monthly mean SST obtained by Drogden and Falsterborev. The black line represents the monthly mean difference, which is calculated by subtracting the monthly mean SST recorded by Falsterborev from the monthly mean SST measured by Drogden. The red lines represent the average difference of the three periods. The average difference of the first, second and third period is -0.26°C, 0.00°C and 0.24°C, respectively.

9.1.2 Combining records without temporal overlaps

Unfortunately, it is sometimes necessary to extend the reference data sets with records that have no temporal overlap with the reference data. There are different ways of dealing with such problem, and the chosen technique is mainly based on whether the data sets have a temporal overlap with the CDR and/or other in situ records. In total, the problem is dealt with in four ways:

1. If neither data set has a temporal overlap with the CDR or any other in situ record it becomes difficult to accurately relate the two series. Furthermore, the CDR covers the entire Baltic Sea, and it is therefore possible to relate the data series in the grid cell corresponding to the location of the reference station with the data series in the grid cell corresponding to the position of the extending station. The model parameters found from the CDR can then be applied to the in situ data. However, it is important to remember that this does not remove any biases that may exist between the in situ records.

2. If both data sets have a temporal overlap with the CDR, it becomes possible to relate the two in situ records by finding two sets of model parameters. The first set of parameters (m_{1_s} and m_{2_s}) are found from relating the in situ data of the extending station to the data series of the CDR grid cell corresponding to the position of the reference station. The second set of parameters ($m_{1_{ref}}$ and $m_{2_{ref}}$) are found from relating the data series of the CDR grid cell corresponding to the reference station to the reference record. It is then possible to reconstruct reference temperatures from the extending record according to:

$$d(t_n) = m_{1_{ref}}(t_n) + [m_{1_s}(t_n) + g(t_n)m_{2_s}(t_n)]m_{2_{ref}}(t_n) \tag{9.4}$$

3. If both records have a temporal overlap with a third in situ record, it is possible to relate them in the same way as in case 2, but where the CDR is replaced by the third in situ record.

4. It is possible to relate two records using a greater amount of intermediate records by combining or adding more steps to case 2 and 3.

9.2 Extended records

In total, six long and continuous data sets are constructed. The goal was to create the data sets such that they would be uniformly distributed throughout the Baltic Sea. However, due to the rather uneven allocation and coverage of the in situ stations, this was not possible. Instead, four of the six records are positioned in or very close to the Transition Zone and are created to represent Kattegat, the Great Belt, Öresund and the Bornholm Basin. This is considered to be a good coverage of the Transition Zone and southern Baltic Proper. It was much more difficult to create long and continuous data sets in the central and northern part of the Baltic Sea due to sparse data. Considering the available data, it was considered a success that it was possible to construct a long and continuous data set in both the Bothnian Sea and the Gulf of Finland. Without financial limitations, it would have been possible to access a long record for the Western Gotland Basin (Landsort) and the Bay of Bothnia (Bjuröklubb) as well. Depending on the quality of these data sets, it is likely that the coverage of the central and northern Baltic Sea would have been better.

As described in Section 9.1, there are several important aspects that need to be taken into account when selecting a reference record and the records that are going to be used to extend it. Information about the selected records are found in Table 9.1. The horizontal lines in the table separate the records used for the different data sets. In each of the sections, the reference record is marked with a star. In the table and the remainder of the report, stations used to construct the continuous record in Kattegat is marked with the letter A. Similarly, stations used to construct the continuous record in the Great Belt is marked with a B, Öresund is marked with a C, the Bornholm Basin is marked with a D, the Bothnian Sea is marked with an E and the Gulf of Finland is marked with the letter F. The positioning of the stations are illustrated in Figure 9.3. The reference stations are again marked by stars. Finally, it is possible to view the correlation between the selected in situ records in Figure 9.4 (which continues in Figure 9.5). In the figure, the correlation between the in situ records and corresponding CDR time series, as well as the correlation between the CDR time series, are also shown. In all cases, an S prior to the station index represent the CDR time series corresponding to the in situ record with the same index. Information about the specific extended records are found in Section 9.2.1 - 9.2.6.

Index	Station Number	Station Name	Type	Start Month	End Month	Lat.	Lon.
A1	35013	Svinbådan	S L	1884-01	1960-09	56.17	12.52
*A2	35015	Fladen	S L	1893-06	1969-10	57.17	11.67
A3	35016	Vinga	S L	1930-01	1965-12	57.57	11.47
A4	6097	Anholt Knob	D L	1931-01	1985-07	56.77	11.86
A5	20098	Frederikshavn	D M	1939-01	1990-03	57.43	10.57
A6	27022	Anholt	D M	1990-10	1999-02	56.72	11.52
A7	20108	Hirsholm	D M	1990-04	1995-08	57.48	10.57
A8	30017	Hornbæk	D A	2000-07	2015-09	56.10	12.47
*B1	6127	Halsskov Rev	D L	1931-01	1973-08	55.34	11.05
B2	28548	Keldsnor	D M	1931-01	1999-03	54.73	10.72
B3	-	Kiel	G B	1987-01	2015-03	54.30	10.16
B4	28234	Slipshavn	D A	1999-12	2015-09	55.28	10.83
C1	35010	Falsterborev	S L	1860-07	1972-11	55.30	12.78
*C2	6183	Drogden Fyrskip	D L	1900-01	1998-09	55.53	12.72
C3	30342	Middelgrundsfortet	D M	1931-01	1999-03	55.72	12.67
C4	30332	Langelinie	D M	1936-01	1983-03	55.70	12.60
C5	35067	Oskarsgrundet	S B	1983-09	1999-06	55.60	12.85
C6	30336	København	D A	1999-07	2015-09	55.70	12.60
D1	35009	Utklippan	S M	1880-04	1922-06	55.95	15.70
D2	35008	Ölands Rev	S L	1923-09	1951-08	56.12	16.57
*D3	32002	Christiansø	D M	1931-01	1998-05	55.32	15.18
D4	-	Oder Bank	G B	1997-06	2015-11	54.05	14.10
D5	32048	Tejn	D A	2001-07	2015-09	55.25	14.83
*E1	35002	Finngrundet	S L	1860-06	1969-06	61.03	18.52
E2	35003	Grundkallen	S L	1883-06	1960-01	60.50	18.92
E3	35004	Svenska Björn	S L	1883-05	1968-09	59.58	19.78
E4	-	Utö	F M	1900-01	1996-02	59.78	21.34
E5	-	Seili	F M	1967-01	2014-12	60.25	21.96
F1	-	Porkkala	F L	1900-01	1968-09	59.93	24.42
F2	-	Harmaja	F M	1900-02	1993-09	60.10	24.96
*F3	-	Tvärminne-A	F M	1926-09	2012-12	59.85	23.25
F4	-	Tvärminne-B	F M	1967-01	2013-02	59.85	23.25

Table 9.1: Information about the stations used to create the continuous records. The first letter under 'Type' represents the nationality of the station (D = Denmark, F = Finland, G = Germany, S = Sweden) and the second letter represents the type of station (A = Automatic coastal, B = Buoy, L = Lightship, M = Maintained coastal). Reference stations are marked with stars.

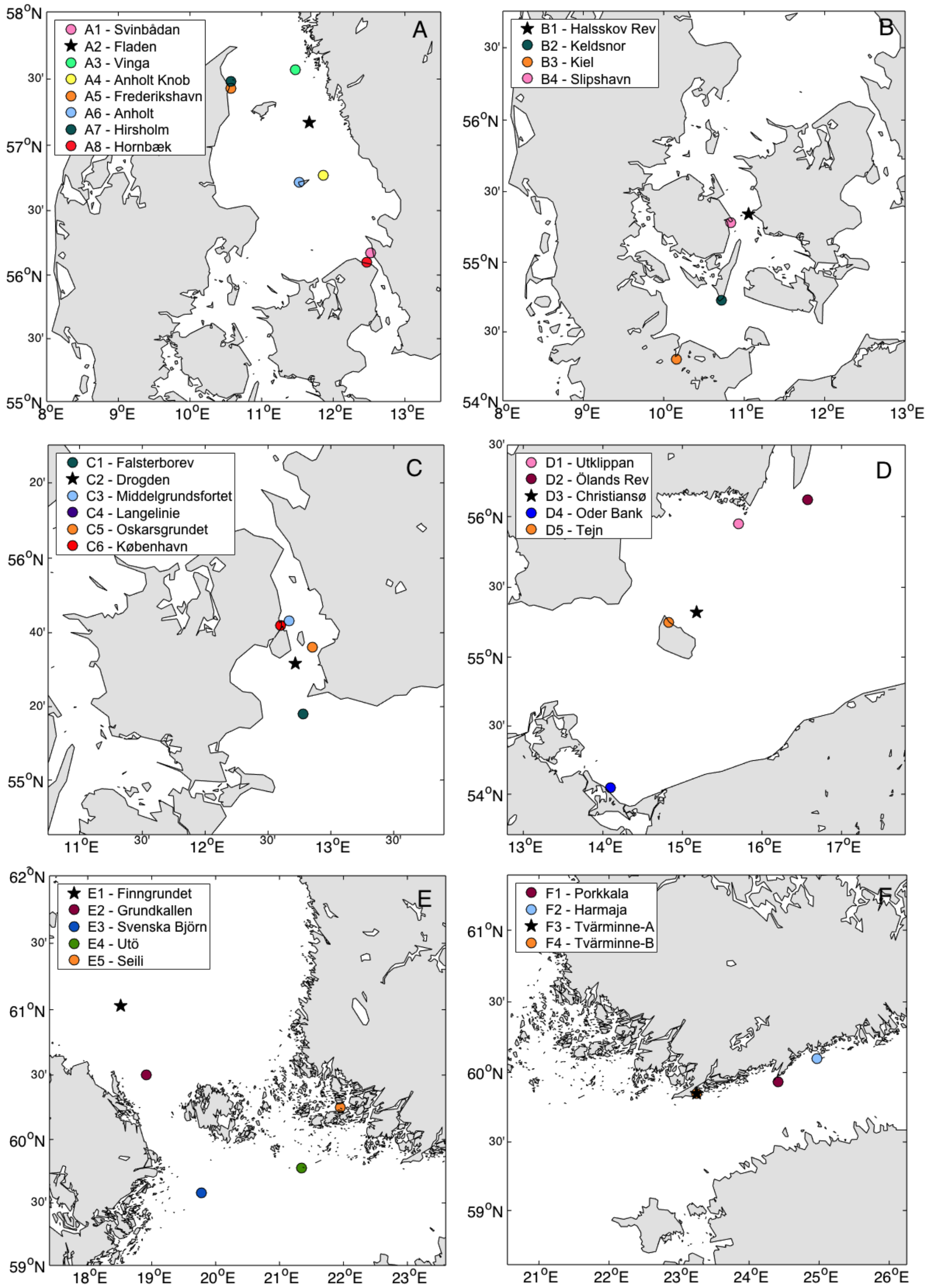

Figure 9.3: Positioning of the stations constituting the long and continuous records. In all six areas, the reference station is marked with a star. It is not possible to view the position of Langelinie as it coincides with København (subfigure C) or Tvärminne-B as it coincides with Tvärminne-A (subfigure F).

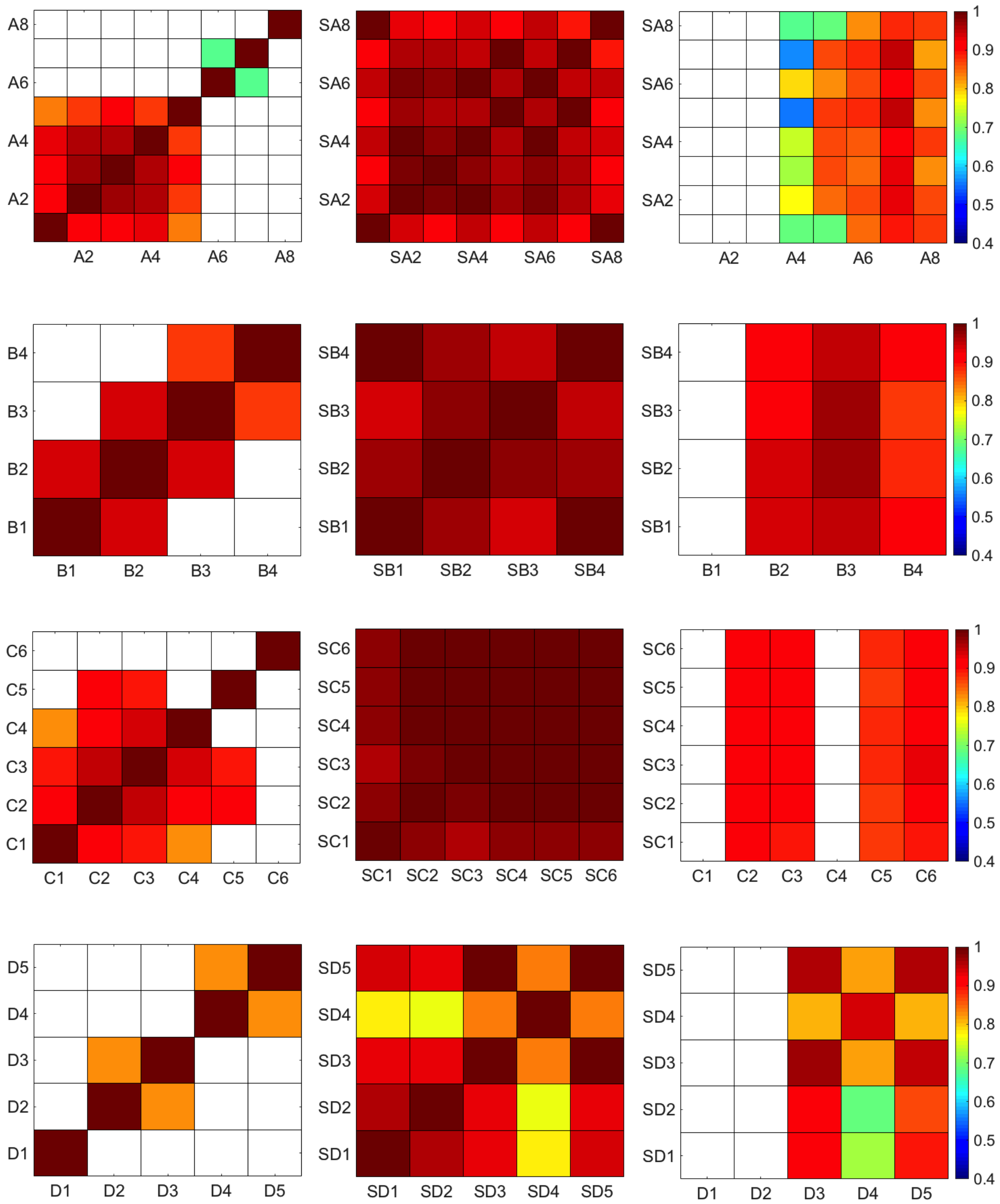

Figure 9.4: The left column shows the correlation between each set of in situ records. The middle column shows the correlation between the corresponding CDR time series. The right column shows the correlation between the in situ records and the CDR time series. White represents temporal overlaps shorter than 24 months. The labels refer to the indexing in Table 9.1, where an S prior to the index represent the corresponding CDR time series.

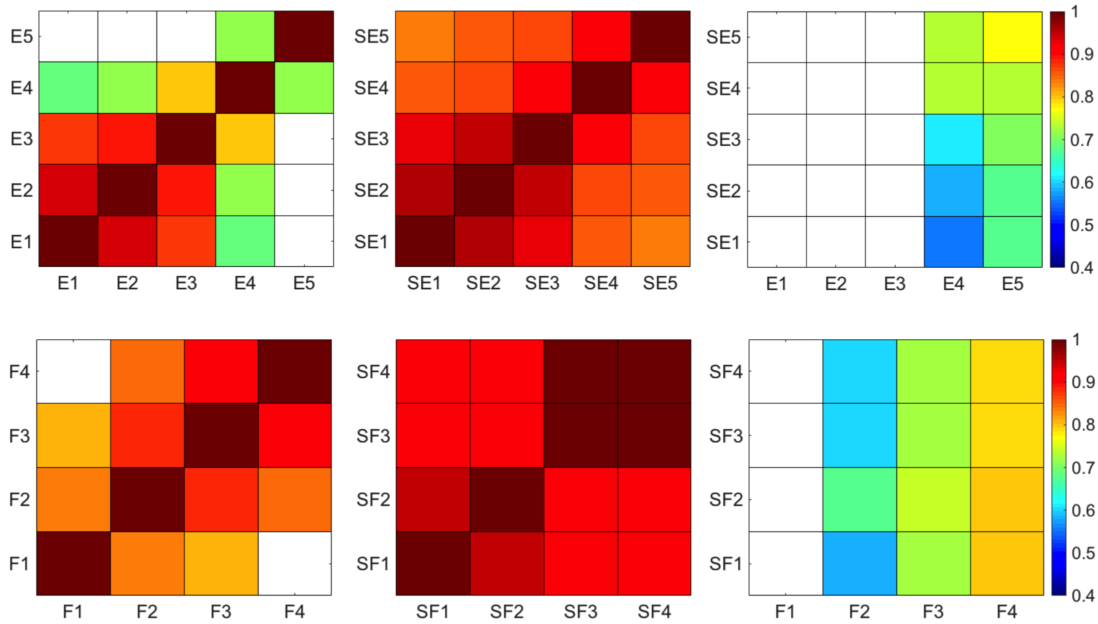

Figure 9.5: Continuation of Figure 9.4.

9.2.1 Extended Record A

The first continuous data set is created to represent Kattegat and uses Fladen as reference station. This record will be called Extended Record (ER) A in the remaining of the report. In total, 7 additional data sets were needed to extend and fill in gaps in the Fladen data set in order for ER A to span from January 1884 to September 2015. It should be noted that there still exist three larger data gaps; November 1914 - October 1918, January 1976 - December 1976 and March 1999 - June 2000. Data for individual months are also missing.

From Table 9.1, it can be seen that data from Svinbådan, Vinga, Anholt Knob and Frederikshavn overlap with data from Fladen. These can therefore easily be used to extend and fill in gaps in the Fladen data sets through the least square method described in Section 9.1.1. Furthermore, it should be noted that Anholt Knob has experienced several small moves. Two sets of model parameters are therefore calculated. The first set of parameters is calculated from data measured at 56.77°N 11.86°E, 56.75°N 11.99°E and 56.75°N 11.88°E and the second set of parameters is calculated from data measured at 56.85°N 11.80°E.

Neither Anholt, Hirsholm or Hornbæk have recorded temperatures during the same period as Fladen. As neither the CDR or any other in situ record overlaps with both data sets, the method described by case 4 in Section 9.1.2 has to be used. In other words, a station that overlaps both with the CDR and Fladen record has to be used in order to reconstruct Fladen data from Anholt, Hirsholm and Hornbæk. There are two possible candidates; Anholt Knob and Frederikshavn. From Figure 9.4, it can be seen that the correlation between Anholt Knob (A4) and Fladen (A2) is higher than the correlation between Frederikshavn (A5) and Fladen (A2). Anholt Knob is therefore initially assumed to be the better choice. Furthermore, Anholt Knob has a rather short temporal overlap with the CDR, which is reflected in the low correlation between Anholt Knob and the CDR. As the reconstructions partly will be based on model parameters found from relating the in situ record to the CDR, Frederikshavn is chosen. In this way, the model parameters relating the in situ record to the CDR will be based on eight years rather than just three, which is assumed to give more accurate reconstructions. Another advantage of choosing Frederikshavn over Anholt Knob is that Frederikshavn has stayed stationary throughout its measurements while Anholt Knob has experienced several moves. The model parameters found from Frederikshavn are therefore believed to be the most reliable. After all, the correlation between Anholt Knob (A4) and Fladen (A2) is still high.

It is possible to view the RMSEs and correlations between the reconstructed data and reference observations in Table C.1. In the table, it is also possible to view the maximum and minimum number of months that the model parameters are based on[1]. Statistics for Anholt (A6), Hirsholm (A7) and Hornbæk (A8) are displayed in gray as it is not possible to compare the reconstructed temperatures to actual Fladen data. Instead, statistics from the intermediate steps are shown.

From Table C.1, it can be seen that Vinga (A3) reconstructs Fladen data the most accurately as it has the smallest RSME and highest correlation. Anholt Knob (A4) was found to be the second best as it has the second smallest RMSE and second highest correlation, followed by Svinbådan (A1) and lastly Frederikshavn (A5). If it is possible to use more than one station to extend the Fladen data, the stations are prioritized in this order. It is not possible to prioritize Anholt, Hirsholm and Hornbæk based on the RMSE and correlation of the reconstructed temperatures, as they are not compared to actual Fladen data. Instead, the prioritizing is based on the correlation between the corresponding CDR time

[1]The model parameters are calculated specifically for each month. The number of months with data varies depending on the start date, end date and the number of data gaps.

series. Here, it is found that the time series corresponding to Anholt (SA6) has the highest correlation with the time series corresponding to Fladen (SA2), followed by Hirsholm (SA7) and Hornbæk (SA8). If more than one of these stations can be used to extend the Fladen record, they are prioritized in this order.

9.2.2 Extended Record B

The second data set is created to represent the Great Belt and will be called Extended Record B (ER B) in the remainder of the report. Halsskov Rev was used as reference station although it could be argued that Keldsnor would have been a better choice as it constitutes the longest and most continuous record in the area. Keldnor is not selected as reference station as it moved in October 1952 (where it also switched name to Bagenknop). Even though the move is rather small, it is reflected in a change in the average bias against the other in situ records (not shown). Keldsnor is therefore not an ideal reference station as the long and continuous record should represent stationary data. Instead, Halsskov Rev is chosen as the reference station as it has stayed stationary throughout its measurements, at the same time as it is has a more central position in the Great Belt. Data from Keldsnor (Bagenknop), Kiel and Slipshavn are used to extend and fill in gaps in the record. Furthermore, due to the lack of in situ data prior to 1931, it was only possible to construct ER B from January 1931 to September 2015. It should also be mentioned that data from 1976 and individual months are missing.

From Table 9.1, it can be seen that only Keldsnor has recorded SST in the same period as Halsskov Rev. Temperatures measured at Keldsnor can therefore be used to extend and fill in gaps in the Halsskov Rev record according to the method described in Section 9.1.1. However, it should be noted that two sets of model parameters are used; one set for data prior to the move in October 1952, and the other set for data from November 1952. Neither Kiel nor Slipshavn have a temporal overlap with Halsskov Rev. As Kiel has a temporal overlap with Keldsnor, it is possible to reconstruct Halsskov Rev temperatures according to case 3 in Section 9.1.2. By adding an extra step to case 3, it is also possible to reconstruct Halsskov Rev temperatures from the Slipshavn data. In this case, model parameters are found for the overlap between Slipshavn and Kiel. The model parameters which reconstructs Keldsnor data from Kiel and Halsskov Rev data from Keldsnor are then applied. The correlation and RMSE of each of these steps can be viewed in Table C.1.

As only one station overlaps with the reference station, it is not possible to compare the RMSE and correlations with the reference record in order to determine the most accurate reconstructions. Instead, the records are prioritized after the CDR time series that has the highest correlation with the CDR time series corresponding to Halsskov Rev. It is found that the CDR time series corresponding to Slipshavn (SB4) has the highest correlation with the CDR time series corresponding to Halsskov Rev (SB1), followed by Keldsnor (SB2) and Kiel (SB3). If more than one data set can be used to extend the record or fill in a certain gap, the records are prioritized in this order.

9.2.3 Extended Record C

The third data set is created to represent the Öresund region and will be called Extended Record C (ER C) in the remainder of the report. Drogden is chosen as reference station as it covers a long time period (1900 - 1998) at the same time as it has been maintained and therefore constitutes a highly continuous record. The record is extended to span between January 1880 and September 2015 by complementing with data from Falsterborev, Middelgrundsfortet, Oskarsgrundet and København. The Danish coastal station Langelinie is used to fill in gaps within the data set (such as during World War II). However, it should be noted that data from 1976 and some months (especially winter months) are still missing.

From Table 9.1, it can be seen that all stations except for København have a temporal overlap with Drogden. Data from these stations can therefore be used to extend the Drogden record according to the least square method described in Section 9.1.1. However, a drift is found in the Falsterborev record. Three sets of model parameters are therefore calculated, which apply prior to World War I, between World War I and II, and after World War II, respectively. Similarly, a drift is found in Middelgrundsfortet. One set of model parameters is therefore calculated based on data prior to the gap in 1976, while the other set is calculated based on data after the gap. As both Drogden and København overlaps with the CDR, the København record is used to extend the Drogden record according to the method described by case 2 in Section 9.1.2.

It is possible to view the RMSE and correlations of the reconstructed data in Table

C.1. It can be seen that the records with the smallest RMSE doesn't necessarily correspond to the records with the highest correlations. As all correlations are high (over 0.9), the accuracy of the records is ranked after the RMSE. Based on this criteria, it can be seen that Falsterborev (C1) reconstructs the Drogden data the most accurately, followed by Middelgrundsfortet (C3), Langelinie (C4) and Oskarsgrundet (C5). If it is possible to use more than one station to extend the Drogden data set or fill in a certain gap, the stations are prioritized in this order. København is used to extend the Drogden data set from July 1999.

9.2.4 Extended Record D

The fourth long data set is created to represent the Bornholm Basin and is denoted Extended Record D (ER D). Christiansø is chosen as the reference station due to the continuity and length of the record. Data from Utklippan and Ölands Rev are used to extend the record back in time, while Oder Bank and Tejn are used to add more recent data. The resulting record spans between April 1880 and November 2015. It should be noted that there exist three large gaps; July 1922 - August 1923, January 1975 - December 1976 and August 1998 - March 1999.

In Table 9.1, it can be seen that Ölands Rev and Oder Bank have a temporal overlap with Christiansø. However, as the temporal overlap between Oder Bank and Christiansø is shorter than a year, only Ölands Rev can be used to extend the Christiansø record according to the method described in Section 9.1.1. As both Oder Bank and Tejn have a temporal overlap with the CDR, these are used to reconstruct Christiansø temperatures according to case 2 in Section 9.1.2. Utklippan doesn't have a temporal overlap with any in situ record or the CDR. Utklippan was therefore used to extend the record according to the method described by case 1 in Section 9.1.2.

From Table 9.1, it can be seen that only Oder Bank and Tejn have overlapping time periods. If both records can be used to fill in a certain gap in the Christiansø record, Tejn is chosen as Figure 9.4 shows that the correlation between Tejn (D5) and the CDR time series corresponding to Christiansø (SD3) is much higher than the correlation between Oder Bank (D4) and the CDR time series corresponding to Christiansø (SD3). The correlations between the corresponding CDR time series support the result (the correlation between SD5 and SD3 is higher than the correlation between SD4 and SD5). The reconstructions obtained from the Tejn record are therefore believed to be more accurate than the reconstruction obtained from the Oder Bank record.

9.2.5 Extended Record E

The fifth long data set is created to represent the Bothnian Sea and is denoted Extended Record E (ER E) in the remainder of the report. From Figure 9.5, it can be seen that Grundkallen and Svenska Björn has the highest correlations with the in situ records in the area. However, as the resulting record will represent the northernmost record in the multivariate regression model, Finngrundet was chosen as reference station as it is located the farthest north. Instead, Grundkallen and Svenska Björn was used to fill in gaps in the Finngrundet record. Utö was also used to fill in gaps in the record, while Seili was used to add more recent data. The resulting record spans from June 1883 to September 2015. While there are no large data gaps, individual months are missing throughout the record. These are usually found during the winter months and are believed to be caused by the withdrawal of the stations due to sea ice. The data gaps are especially common prior to the addition of Utö in 1900, where data between December and April are missing for nearly all of the years.

From Table 9.1, it can be seen that all records except for Seili has a long temporal overlap with Finngrundet. Data from these stations can therefore be used to extend Finngrundet according to the least square method described in Section 9.1.1. As Finngrundet doesn't have a temporal overlap with the CDR, an in situ record that overlaps both with Seili and Finngrundet has to be used in order to reconstruct Finngrundet temperatures from the Seili record. As Utö is the only in situ record that overlaps both with Seili and Finngrundet, Utö was used to reconstruct Finngrundet temperatures from the Seili record according to case 3 in Section 9.1.2.

It is possible to view the RMSEs and correlations of the reconstructed data in Table C.1. It can be seen that Grundkallen (E2) reconstructs Finngrundet temperatures the best as it has the lowest RMSE and highest correlation, followed by Svenska Björn (E3) and Utö (E4). If more than one data set can be used to fill in a certain gap, the records are prioritized in this order. From Figure 9.5, it can be seen than the correlation between the CDR time series corresponding to

Utö (SE4) and the CDR time series corresponding to Finngrundet (SE1) is higher than the correlation between the CDR time series corresponding to Seili (SE5) and the CDR time series corresponding to Finngrundet (SE1). Utö is therefore chosen over Seili if both can be used to fill in a certain gap in the Finngrundet record.

9.2.6 Extended Record F

The last long data set is created to represent the Gulf of Finland and is denoted Extended Record F (ER F) in the remainder of the report. The record constitutes of data originating from Porkkala, Harmaja, Tvärminne-A and Tvärminne-B. Of the four records, Harmaja and Tvärminne-A are best suited as reference stations due to the length and continuity of the records. From Figure 9.5, it can be seen that the correlation between Tvärminne-A (F3) and Tvärminne-B (F4) is higher than the correlation between Harmaja (F2) and Tvärminne-B (F4). Furthermore, the correlation between Harmaja (F2) and Porkkala (F1) exceeds the correlation between Tvärminne-A (F3) and Porkkala (F1). The reference station can therefore not be determined from the correlation between the in situ records. Instead, Tvärminne-A is chosen as reference station as it has the highest correlation with the CDR time series (SF1, SF2, SF3 and SF4) and as it is situated very close to Tvärminne-B.

As all data sets have a temporal overlap with Tvärminne-A, all temperatures are reconstructed according to the method described in Section 9.1.1. This means that it is possible to determine how well the reconstructed temperatures correspond to Tvärminne-A temperatures as the RMSE and correlation can be calculated. The RMSEs and correlations can be viewed in Table C.1. From the table, it can be seen that Tvärminne-B (F4) reconstructs Tvärminne-A temperatures the best as it has the highest correlation and smallest RMSE. Harmaja (F2) makes the second best reconstructions as it has the second highest correlation and second smallest RMSE, followed by Porkkala (F1). If more than one station can be used to extend the record or fill in a certain gap, the stations are prioritized in this order. The resulting data set spans between January 1900 and February 2013. The only large data gap is found between June 1918 and November 1918. However, it could be discussed whether or not it is realistic that nearly all winter months contain data.

9.3 Comparison against CDR

In order to determine the validity and representativeness of the extended records, the extended records are compared against the CDR. As the CDR only contains data between January 1982 and December 2011, it is only possible to check the corresponding parts of the extended records. Furthermore, as different time periods of the extended records are based on different in situ data sets, it is possible that the quality of the records vary with time. Due to the relatively short temporal coverage of the CDR, quality variations cannot be revealed here. Instead, the results are assumed to be valid for the entire records.

The extended records are compared against the corresponding CDR time series by calculating the correlation (ρ), RMSE, bias, minimum difference (Min) and maximum difference (Max) between the time series. The result is shown in Table 9.2. Here, N shows the number of match-ups on which the comparisons are based. All differences are calculated by subtracting the temperatures of the extended records from the CDR. A positive bias therefore indicates that the temperatures in the CDR are warmer than in the extended record.

From Table 9.2, it can be seen that the extended records of the Transition Zone and the southern Baltic Proper agrees well with the corresponding CDR time series. All correlations are 0.9 or higher, while all RMSEs are lower than 0.65°C. The results obtained from ER A and ER B especially proves the success of the constructions as these are based on reference records without temporal overlaps with the CDR. However, the results obtained for ER E and ER F reveals much lower consistency between the northern records and the CDR. This was expected as they are based on Finnish in situ records, which have been shown to be much less consistent with the CDR in Chapter 7 and 8.

Each extended record is also correlated against the CDR in order to examine the spatial representativeness of the records. The resulting correlation maps are shown in Figure 9.6, where the location of the extended records are marked with black stars. The correlation maps for ER A, ER B, ER C and ER D are similar and show the expected patterns based on the correlation analysis in Chapter 8. All records show the highest correlations in the proximity of the stations, while the correlations in the Transition Zone and the majority of the Baltic Proper are high. It can also be seen that the correlation is above 0.6 in most of the Gulf of

Riga and in the entrance of the Gulf of Bothnia and the Gulf of Finland. The correlation decreases to approximately 0.4 in the northernmost extension of the Gulf of Bothnia and easternmost extension of the Gulf of Finland.

The overall correlation between the northern records and the Baltic Sea is much lower. The correlation pattern obtained for ER F agrees well with the correlation pattern found for Tvärminne-B in Figure 8.3. The spatial representativeness is therefore as expected. However, the representativeness of ER E is much lower than expected. In Figure 8.3, it can be seen that both Almagrundet (61) and Valassaaret (75) (located south and north of ER E, respectively), show much higher correlations in the Gulf of Bothnia, Gulf of Finland and the northern Baltic Proper. The correlation pattern was therefore expected to be much more similar to these. The generally low correlations could therefore indicate that the extending in situ records were incompatible with the reference record. This is especially believed to be the case for Utö and Seili, as Figure 9.5 proves that Finngrundet, Grundkallen and Svenska Björn are highly correlated.

Record	ρ	RMSE (°C)	Bias (°C)	Min (°C)	Max (°C)	N
ER A	0.90	0.62	-0.01	-1.57	2.22	321
ER B	0.93	0.45	0.16	-1.13	1.53	345
ER C	0.93	0.58	-0.14	-1.83	2.55	336
ER D	0.96	0.48	0.32	-0.99	1.42	344
ER E	0.64	1.20	0.54	-1.92	5.87	315
ER F	0.72	1.21	0.04	-4.53	4.05	246

Table 9.2: Comparison of the extended records with the time series of the corresponding CDR grid cells.

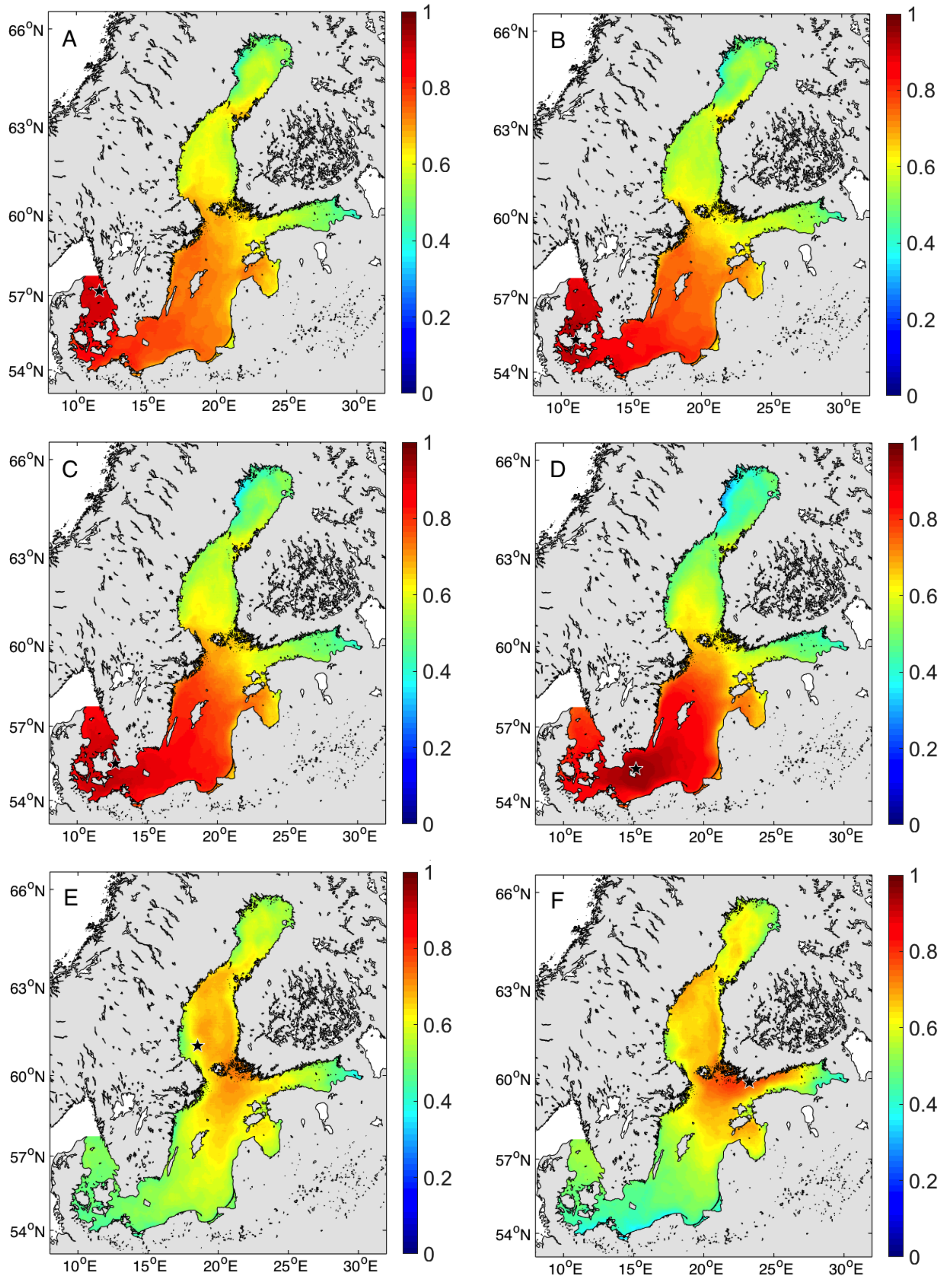

Figure 9.6: Correlation between the CDR and Extended Record A, B, C, D, E and F. The location of the extended records are shown by the small black stars.

10 Multivariate regression model

As described in previous chapters, the DMI SST CDR has a high spatial resolution (0.03° x 0.03°) but a limited time coverage (1982 - 2011). In contrast, the extended records (created in Chapter 9) sample localized regions but have a much longer time coverage. By finding a relationship between the extended records and CDR, it is possible to reconstruct historical SST fields with the spatial resolution of the CDR and temporal coverage of the extended records. Such relationship can be found by constructing a multivariate regression model where the SST of every position in the Baltic Sea (CDR grid cell) is expressed as a weighted sum of the observations from the extended records. Such model can thereby reconstruct the SST at any location in the Baltic Sea as long as the extended records contain observations for the same time. Similar models have been constructed by Madsen (2006) and Høyer (2002), where the aim was to reconstruct the sea surface height of the Baltic Sea and North Sea, respectively. Even though the examined parameter differs, the basic concept is the same and the model is based on their studies.

The theory behind the model is presented in Section 10.1. The resulting model coefficients are reviewed in Section 10.2 and the model quality is discussed in Section 10.3. Finally, the model is validated by comparing the reconstructed monthly mean SST fields to in situ records not included in the model. The result of the validation is presented in Section 10.4.

10.1 Theory

The purpose of the multivariate regression model is to reconstruct historical SST fields for the Baltic Sea using in situ observations. The model is based on the idea that the SST of each position in the Baltic Sea (CDR grid cell) can be expressed as a weighted sum of in situ observations according to:

$$d(i, j, t_n) = \sum_{k=1}^{K} m_k(i, j) g_k(t_n) \tag{10.1}$$

Here, $d(i, j, t_n)$ is the CDR SST of grid cell i, j at time t_n, where $n = 1, 2, ..., N$ and N is the number of months where all extended records contain data. The corresponding observations in the extended records are given by $g_k(t_n)$. Here, $k = 1, 2, ..., K$ where K is the number of extended records included in the model.

The weight assigned to each record is given by $m_k(i,j)$. Each extended record is therefore assigned a specific weight in every position (CDR grid cell) of the Baltic Sea. If the weights are known, Equation 10.1 can be used to reconstruct the SST in every position of the Baltic Sea for all times where the extended records contain data.

The weights relating the extended records to the CDR are found by setting up a system of equations that are valid for all times where both the CDR and the extended records contain data. The complete set of equations is given by:

$$\begin{bmatrix} d(i,j,t_1) \\ d(i,j,t_2) \\ \vdots \\ d(i,j,t_N) \end{bmatrix} = \begin{bmatrix} g_1(t_1) & g_2(t_1) & \dots & g_K(t_1) \\ g_1(t_2) & g_2(t_2) & \dots & g_K(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(t_N) & g_2(t_N) & \dots & g_K(t_N) \end{bmatrix} \begin{bmatrix} m_1(i,j) \\ m_2(i,j) \\ \vdots \\ m_K(i,j) \end{bmatrix} \tag{10.2}$$

As the data contains noise and as it is very unlikely that the relationship between the CDR and extended records is completely linear, it is not possible to find a set of model coefficients (weights) that will satisfy Equation 10.2 exactly. Instead, the model coefficients are found from the least square solution described in Chapter 7. That is, the model coefficients are found such that the residual between the model predicted SST ($\mathbf{Gm}$) and actual SST ($\mathbf{d}$) becomes minimized. In this case, $\mathbf{d}$ is the vector containing the CDR time series, $\mathbf{G}$ is the matrix containing the extended records and $\mathbf{m}$ is the vector containing the weights. The model coefficients are then estimated from the least square solution according to:

$$\mathbf{m} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d} \tag{10.3}$$

In order to keep the climatology and seasonal cycle of each position in the Baltic Sea, the model is based on SST anomalies. As the model coefficients obtained from Equation 10.3 will be based on data spanning from 1982 to 2011, the anomalies are obtained by subtracting the corresponding monthly climatologies from all time series included in the model. The model parameters obtained from Equation 10.3 are then used to reconstruct SST anomalies from the extended (anomaly) records according to Equation 10.1. The reconstructed SST fields are obtained by adding the grid cell specific climatologies to the reconstructed anomalies. Finally, the reconstructed ice fields (created in Chapter 5) are applied to the reconstructed

SST fields such that ice covered regions are assigned a temperature of -1°C. All reconstructed SSTs below -1°C are also assigned a temperature of -1°C.

As there exist data gaps in the extended records, the analysis is made for different combinations of extended records such that the model can be used to reconstruct SST fields despite there being a data gap in one or more of the records. The number of extended records included in the model is altered in Equation 10.1 - 10.3 by changing the value of K. As there are six extended records, it is possible to base the model on 63 different combinations. As this is very extensive, the records are examined to find the most necessary combinations. These are found based on two criteria. First of all, at least three of the six extended records must contain data. Two records are considered too few to describe the entire Baltic Sea. Secondly, the specific combination must be able to reconstruct at least five months. The different combinations and the number of reconstructions (months) that are based on each combination is shown in Table 10.1. As only two extended records (ER C and ER D) contain data prior to 1883, the earliest reconstructions correspond to 1883. Similarly, as the historical ice fields only spans until 2011, it is only possible to make reconstructions until 2011. The reconstructed monthly mean SST fields thus span from 1883 to 2011.

10.2 Model parameters

The weight assigned to the extended records in every position (CDR grid cell) of the Baltic Sea is shown in Figure 10.2. The location of the extended records are shown by black stars. With some exceptions, the weights are largest in the vicinity of the stations and approaches zero with distance. The spatial scale on which the extended records have a great influence varies, but seems to mostly depend on the distance to the adjacent stations. The overall result was expected, though some of the weights in the Gulf of Finland and the Gulf of Bothnia were found surprising.

The weights found for the Transition Zone agrees well with the correlation patterns shown in Figure 9.6. ER A is found to have the greatest influence in Kattegat, ER B in the Belt Sea and ER C in the Öresund region. It can also be seen that ER B has a rather large influence in Kattegat. This was anticipated as the correlation between ER B and the CDR is higher than the correlation between ER A and the CDR for large parts of Kattegat. ER D, ER E and ER

F show weights close to zero and are therefore found to have nearly no influence on the reconstructed temperatures in the Transition Zone.

ER D is found to dominate the signal in the Baltic Proper, which agrees well with the correlation pattern shown in Figure 9.6. Furthermore, ER C has a rather large influence on the western part of the Arkona Basin. This is expected as ER C is located at the boundary to the Arkona Basin. ER B is found to have some influence along the southern and eastern coasts along the Baltic Proper. This is assumed to be a coastal phenomena caused by similar depths in the Belt Sea and coastal regions of the Baltic Proper (see Figure 2.2). What is more surprising is that ER E has a relatively large influence on the northeastern Baltic Proper. This was not expected as the correlation between the CDR and ER D is higher than the correlation between the CDR and ER E in most of the region.

One of the more surprising results was the rather low weight of ER F in the Gulf of Finland. Even along the northern coast, where the correlation with the CDR is high, the weight barely exceeds 0.5. Instead, ER E and ER B are found to dominate the signal in the region. Ones again, ER B is believed to have been assigned a large influence due to similar bathymetries.

Finally, ER E is found to dominate the signal in the Gulf of Bothnia, but there are two startling features. First of all, ER F was expected to be assigned large weights in the Archipelago Sea due to the high correlation with the CDR in the area. Instead, ER F was given a much greater weight. Secondly, the relatively high weights of particularly ER A, but also ER B, were unexpected. The high weights could be an artifact of similar bathymetry or simply represent noise. Generally, ER A, ER B and ER C show coastal characteristics, which can be seen by the slightly higher weights in the coastal regions along the Baltic Sea.

In order to determine the reliability of the weights, the model coefficients found for the corresponding CDR time series are shown in Figure 10.3. Although the magnitude of the weights in the proximity of the stations are strengthened, the overall patterns agree well with Figure 10.2. However, there are some differences. As expected, ER B loses influence in Kattegat while ER A gains influence. Something less expected are the negative weights assigned to ER A along the eastern coast of the Baltic Sea. At the same time, the weights assigned to ER B are greatly strengthened for the same area. This is assumed to show a deficiency in the model. The largest differences are found for ER F, which now is assigned large weights in the Gulf of Finland, the eastern part of the Baltic Proper, the Gulf of Riga, along the eastern border of the Gulf of Bothnia and along the Northern Quark Strait. This could indicate that the small weights assigned to the actual record are due to the undersampling (and time difference) of the Finnish in situ records.

As a final test, the absolute sum of all model parameters are calculated. These are shown in Figure 10.1, where the left figure corresponds to the sum of the weights found for the actual records while the right figure shows the sum of the weights found for the corresponding CDR time series. It is expected that the sum is close to 1 in areas which are well represented by the model. According to this theory, the model seems to be more accurate when it is based on the extended records than the corresponding CDR time series. The rather high and patchy total weights in the Gulf of Bothnia, Gulf of Finland and northern Baltic Proper could indicate that the model is having trouble resolving these areas. This is also most likely the explanation for the unexpected model coefficients in the area.

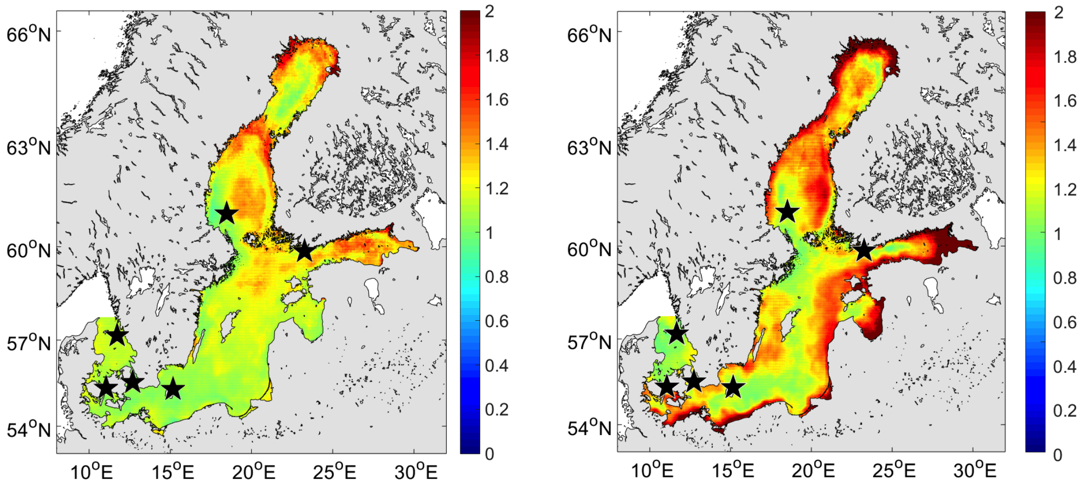

Figure 10.1: Absolute sum of the weights of the extended records (left) and corresponding CDR time series (right). The position of the stations are shown by black stars.

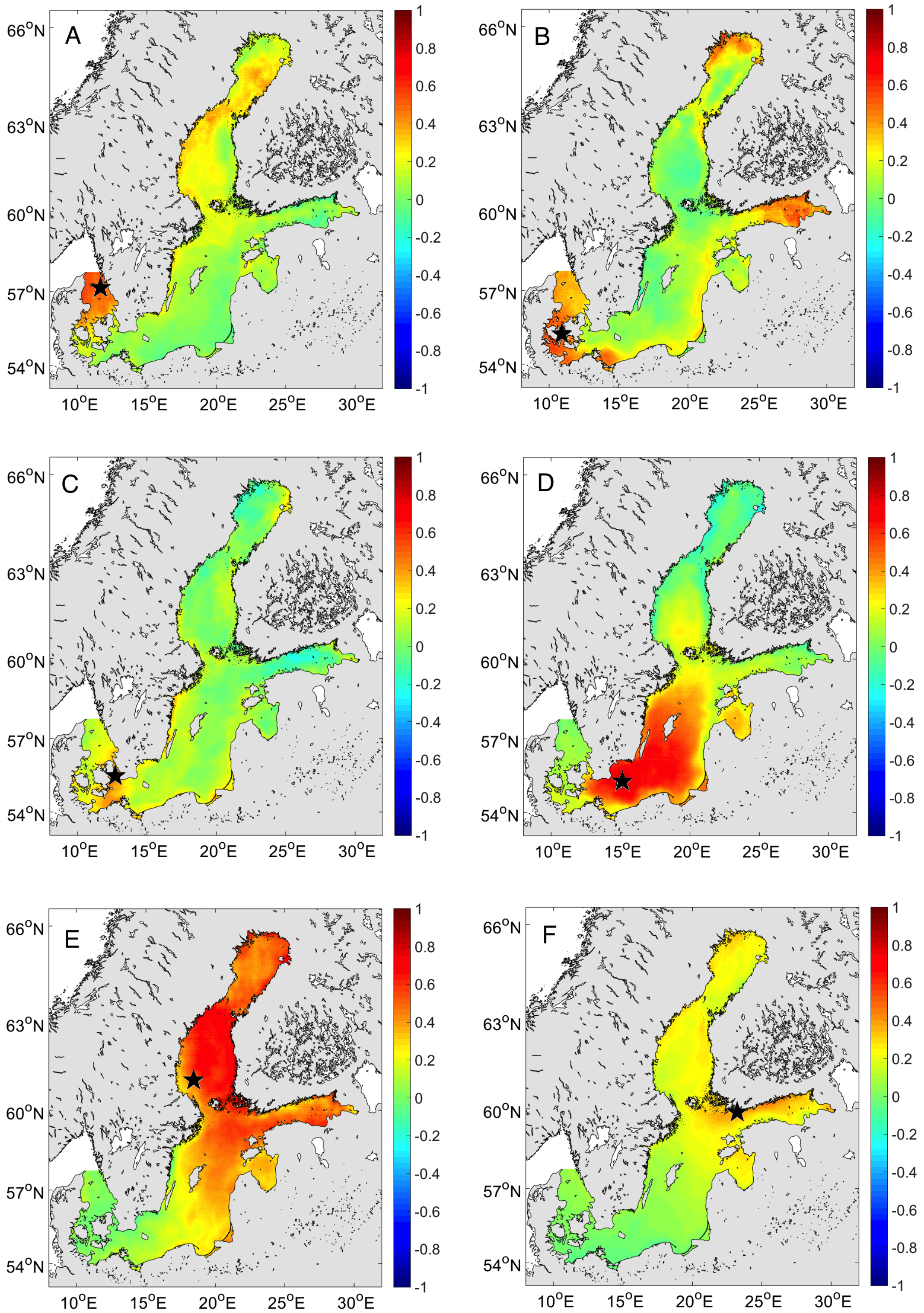

Figure 10.2: Weights assigned to Extended Record A, B, C, D, E and F. The position of the records are shown by black stars.

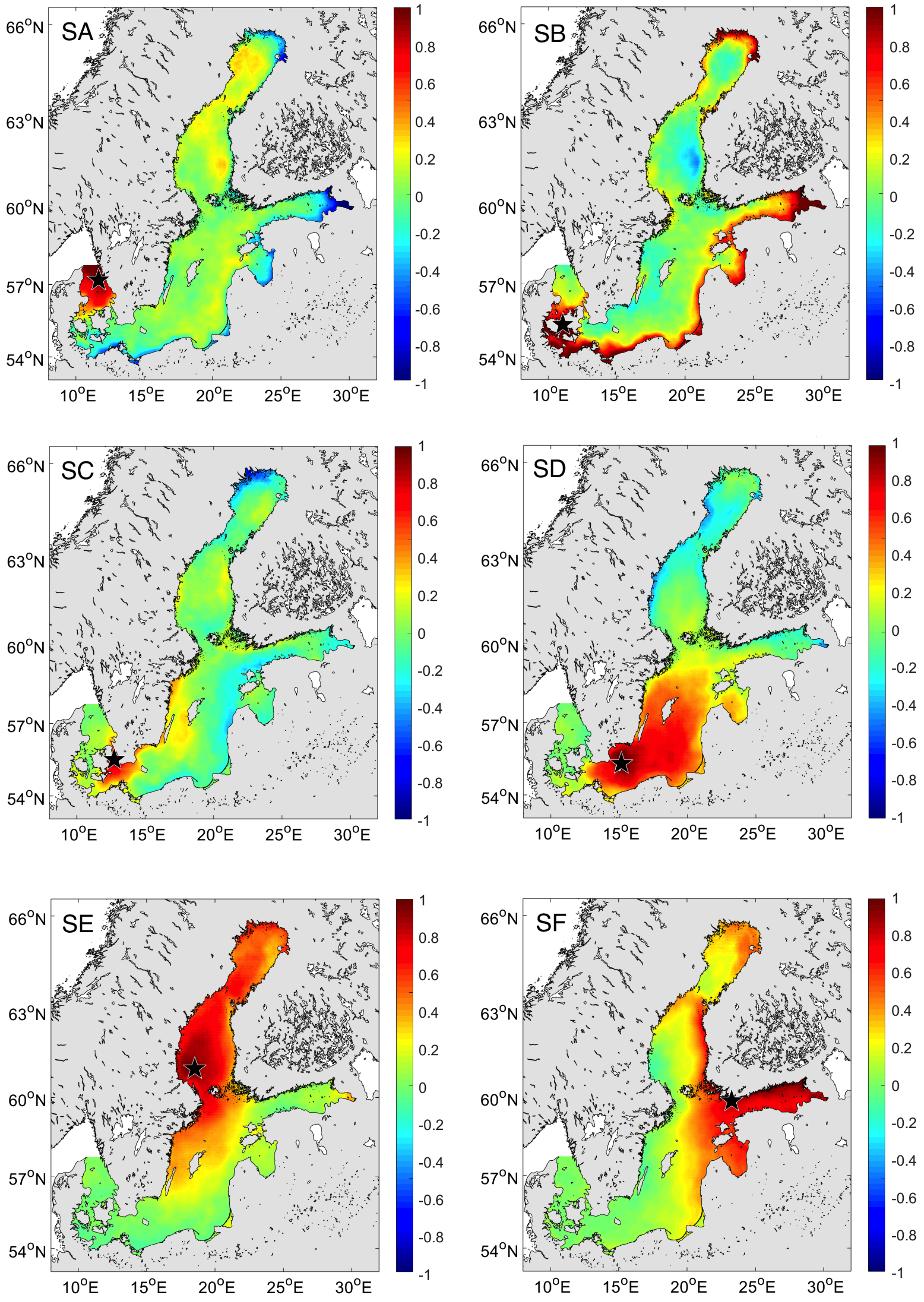

Figure 10.3: Weights assigned to the CDR time series corresponding to Extended Record A, B, C, D, E and F. The position of the time series are shown by black stars.

10.3 Model quality

The RMSEs between the reconstructed SST fields and the DMI SST CDR are shown in Figure 10.4. It is important to remember that this isn't a validation of the reconstructed fields, but an indication on the reliability of the model. The lowest RMSEs are found in the Transition Zone and the Baltic Proper, with increasing values toward the Gulf of Bothnia and the Gulf of Finland. The highest RMSEs are found in the eastern part of the Gulf of Finland where the values approach 2.5°C. It can also be seen that the RMSEs generally are a tad higher along the coast than in the open water.

For comparison, the RMSEs obtained from the reconstructions based on the CDR time series are also shown. Although the RMSEs generally are lower, the same patterns are found. It can therefore be concluded that the reduced model quality in the Gulf of Bothnia and the Gulf of Finland is caused by an underrepresentation of these areas rather than being caused by the poor quality of ER E and ER F. The model would therefore benefit from including additional records, preferably from the Bay of Bothnia and the eastern extension of the Gulf of Finland.

So far, the discussion has only dealt with the multivariate regression model which includes all 6 records. Besides this model, 12 other models based on the record combinations shown in Table 10.1 are also constructed. For each of the combina-

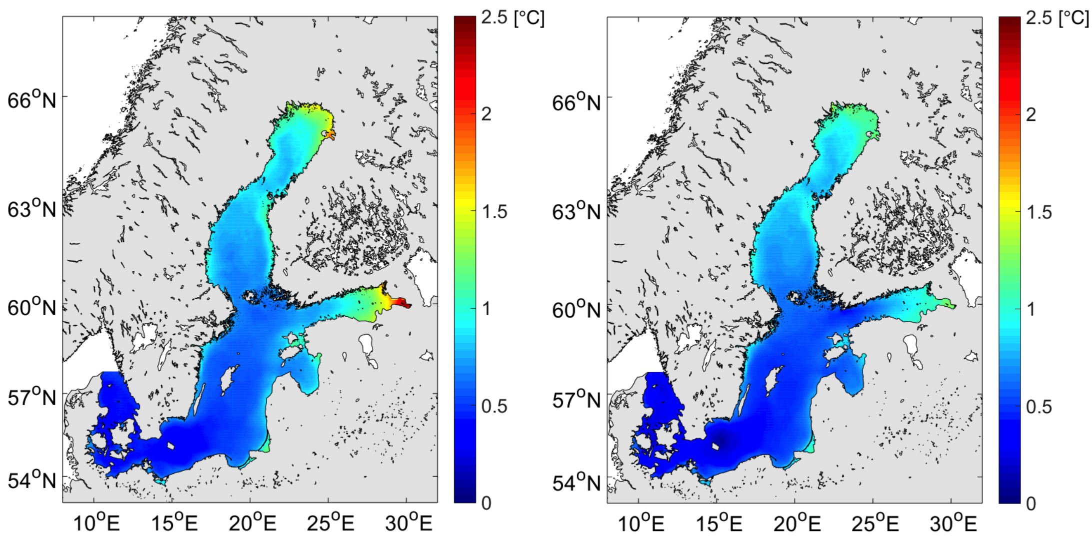

Figure 10.4: RMSEs between the reconstructed SST fields and CDR (left) and the corresponding CDR based fields and the CDR (right).

tions, new model parameters are calculated. To be able to compare the model quality of the 13 different models, an average RMSE (<RMSE>) is calculated for the entire Baltic Sea. These can be viewed in Table 10.1, together with the number of monthly mean SST fields that the model has reconstructed. It can be seen that the model quality generally decreases (increasing <RMSE>) with a decreasing amount of records. Furthermore, when comparing the result of models which are based on the same amount of records, it can be seen that models missing either ER E or ER F are the worst. This highlights the importance of using evenly distributed stations. It is especially interesting to see that the model quality is worse when ER F is missing than ER E, despite ER F generally having lower weights. The worst model combination is found when both ER E and ER F are missing. Furthermore, it can be seen that the RMSE only varies between 0.75°C and 0.89°C, indicating that all combinations are valid representations of the Baltic Sea. Furthermore, there exist 46 months where it wasn't possible to reconstruct monthly mean SST fields. Of these, 12 months are from 1976. Of the remaining months, most gaps are found during winter (December-March) or in the 1800's. Only 4 of the remaining months fall outside these categories.

Combination	Number of Months	<RMSE> (°C)
ABCDEF	841	0.75
BCDEF	20	0.76
ACDEF	288	0.76
ABCEF	22	0.78
ABCDF	38	0.79
ABCDE	9	0.80
ACEF	21	0.78
ACDF	7	0.79
ACDE	128	0.80
CDEF	32	0.78
ACD	65	0.89
CDF	16	0.81
CDE	15	0.83
Missing	46	-
Total	1548	-

Table 10.1: Models based on different combinations of Extended Record A, B, C, D, E and F. The number of monthly mean SST fields that each model was used to reconstruct are listed under Number of Months. The average RMSEs of the reconstructions are listed under <RMSE>.

10.4 Validation

The reconstructed monthly mean SST fields are validated against all available independent in situ records, i.e. all in situ records not included in the model. The correlation, RMSE, average bias, standard deviation, minimum difference and maximum difference between the in situ records and the reconstructed monthly mean SST fields can be viewed in Table D.1. As Figure 10.1 and 10.4 have indicated that the model quality decreases with distance from the Transition Zone, the table is constructed such that the in situ records are listed according to the indexes in Table B.1. All comparisons are made by subtracting the in situ observations from the values of the corresponding grid cells in the reconstructed monthly mean SST fields. Positive biases therefore indicate that the reconstructed fields are warmer than the in situ records.

The correlations and RMSEs show great agreement between the Transition Zone (stations with indexes 1-46) and the reconstructed monthly mean SST fields. With the exception of Trubaduren and Varberg, all correlations are higher than 0.81. Furthermore, 21 out of the 29 records even have correlations of 0.90 or higher. With the exception of Trubaduren and Darβer Schwelle, all RMSEs are lower than 1.00°C. The results obtained from Trubaduren (correlation of 0.77, RMSE of 1.00°C and a maximum difference of 6.91°C) was not unexpected as similar statistics were found for the comparison with the CDR (see Table A.1). In Chapter 8, Trubdauren was also found to be the in situ station with the lowest correlations against the other in situ stations in the Transition Zone. The results obtained from Trubaduren is therefore assumed to reflect the quality of Trubaduren rather than the reconstructed SST fields.

The remainder of the Baltic Sea is validated using the remaining 17 in situ records. Of these, 3 represent the Arkona Basin, 8 the Western Gotland Basin and 6 the Gulf of Bothnia. Unfortunately, there are no in situ records from the Gulf of Finland. The correlation with the Arkona Basin varies between 0.89 and 0.93 with corresponding RMSEs of 0.86°C to 1.13°C. In the Western Gotland Basin, Hävringe (Lightship 1951-1967), Landsort (Maintained Coastal station 1880-1922) and Kopparstenarna (Lightship 1883-1915) show promising results with correlations of 0.91, 0.84 and 0.84 and RMSEs of 0.78°C, 1.13°C and 1.12°C, respectively. However, it can be seen that all stations in the Western Gotland Basin show warm biases. The reconstructed monthly mean SST fields are therefore believed to be less accurate in the Baltic Proper than in the Transition Zone.

The lowest correlations and highest RMSEs are found in the Gulf of Bothnia, where the correlations vary between 0.53 and 0.78 and the RMSEs span from 1.13°C to 1.77°C. Moving northward through the Gulf of Bothnia, the biases go from positive to negative. Furthermore, negative biases were expected for all Finnish in situ records as these represent day time values while the reconstructed SST fields correspond to night values. The lack of negative biases for the Finnish records in the northern Baltic Proper and southern Gulf of Finland therefore supports the assumption that these regions may be slightly too warm in the reconstructed SST fields.

To check the temporal stability of the reconstructed monthly mean SST fields, yearly mean biases and standard deviations are calculated. The result is shown in Figure 10.5, where black represents the Transition Zone (stations with index 1-46), red the Baltic Proper (stations with index 47-65) and green the Gulf of Finland (stations with index 70-78). It would have been interesting to make the analysis even more region specific, but this was not possible due to the limited amount of in situ records. Still, the yearly mean differences and standard deviations indicate that the accuracy of the reconstructed monthly mean SST fields decreases away from the Transition Zone.

In the Transition Zone, all yearly differences are within 0.38°C, whereof the majority (105 out of 124 years) are within 0.25°C. The corresponding standard deviations are lower than 1.0°C for all years. The yearly validation statistics therefore demonstrate the great accuracy of the reconstructed SSTs in the Transition Zone. Furthermore, it can also be seen that the reconstructed temperatures have a high temporal stability. Fitting a straight line to the yearly differences yields a trend of -0.0004°C/year (not shown), proving the great temporal stability of the reconstructed monthly mean SST fields in the Transition Zone.

In the Baltic Proper, all yearly differences are within 1.33°C. The majority of the years (110 out of 117) have an accuracy within 1.0°C. The accuracy of the reconstructed SST in the Baltic Proper is therefore found to be lower than in the Transition Zone. Still, the least accurate reconstructions are found in the Gulf of Bothnia. Here, all yearly validations are within 1.83°C, whereof the majority (114 out of 124) are within 1.0°C. Fitting a straight line through both time series (not shown) yields a trend of -0.0077°C/year for the Baltic Proper and -0.0052°C/year in the Gulf of Bothnia. The reconstructions of these areas are therefore also shown to be less stable.

However, the bar chart of Figure 10.5 reveals that the validation of the Baltic Proper and the Gulf of Finland are based on very few match-ups. While the average comparison in the Transition Zone is based on 83 match-ups, the average comparison in the Baltic Proper is only based on 21 match-ups while the comparisons of the Gulf of Finland on average are based on 16 match-ups. The validation of these areas are therefore not as robust as the validation of the Transition Zone. Lower accuracies were also expected as 1 of the records used to validate the Baltic Proper and 5 of the records used to validate the Gulf of Bothnia were Finnish. It is therefore possible that the lower accuracies partly represents the lower quality of the Finnish records. The 'true' accuracy of the fields may therefore be higher than shown in Figure 10.5.

In Chapter 7, it was shown that there exist seasonal biases that are specific for the station types. As the reconstructed monthly mean SST fields should be consistent with the CDR, the same biases should be found for the reconstructed SST fields. The result is shown in Figure 10.6, where black represents lightships, blue represents maintained coastal stations, red represents buoys and green represents automatic coastal stations. For comparison, the biases found between the CDR and in situ records are also shown. To be consistent with the analysis in Chapter 7, the analysis is only based on Danish, German and Swedish records.

Figure 10.6 shows that the station specific seasonal biases are consistent with those found for the CDR. This proves that the reconstructed monthly mean SST fields are consistent with the CDR on a seasonal basis, and therefore supports the validity of the reconstructed monthly mean SST fields.

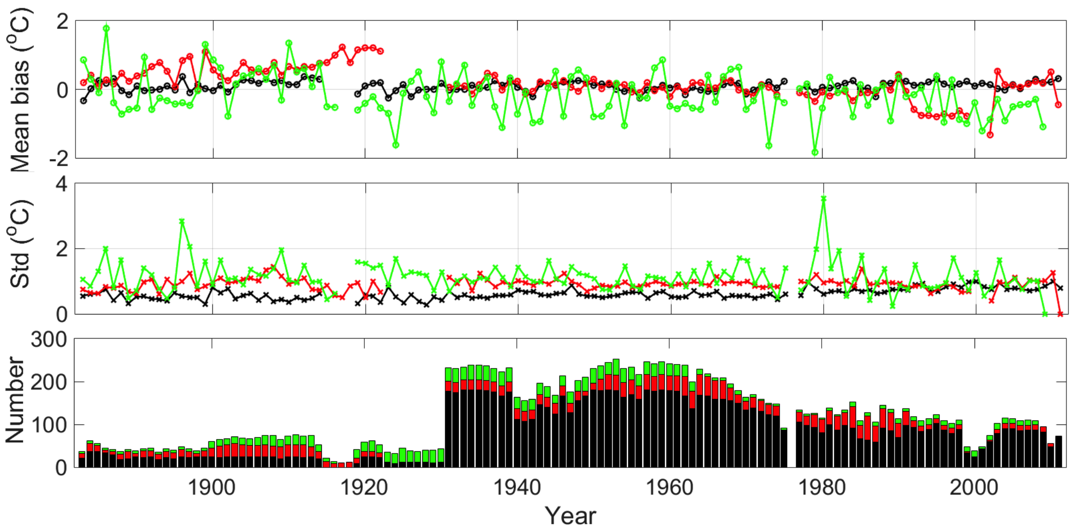

Figure 10.5: Yearly mean difference between the reconstructed monthly mean SST fields and the independent in situ records. The top graph show the mean difference for the specified years, the middle graph shows the corresponding standard deviations and the bottom graph shows the number of match-ups on which the calculations were based. The analysis differs between in situ records from the Transition Zone (black), the Baltic Proper (red) and the Gulf of Bothnia (green).

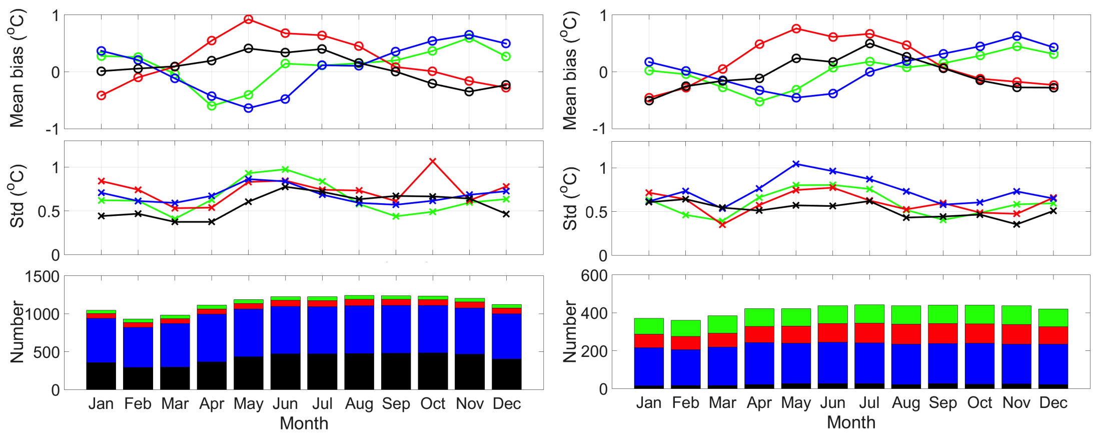

Figure 10.6: Monthly mean difference between the reconstructed monthly mean SST fields and in situ records (left) and DMI SST CDR and in situ records (right). The top graphs show the mean difference for the specified months, the middle graphs show the corresponding standard deviations and the bottom graphs show the number of match-ups on which the calculations were based. Green represents automatic coastal stations, red represents buoys, blue represents maintained coastal stations and black represents lightships. Only Danish, German and Swedish records are included in the analyses.

11 Analysis of the reconstructed SST fields

To demonstrate the usefulness of the product, this chapter analyses the interannual variability and linear trends of the reconstructed fields. A more in-depth analysis of the fields will most likely reveal interesting details of the Baltic Sea, but is out of the scope of this thesis and is left for future work. To reduce the seasonal impact, the analysis is based on monthly anomalies rather than monthly mean SST. The anomaly time series are obtained by subtracting the grid cell specific climatologies from the monthly mean SST time series of each grid cell.

Representative monthly anomaly time series are calculated for the Transition Zone, Baltic Proper, Gulf of Finland and Gulf of Bothnia. The regions are separated into four 'boxes' using rough and simplified estimates of the boundaries; the Baltic Proper is set to border the Transition Zone at 13.0°E, the Gulf of Bothnia at 60.3°N and the Gulf of Finland at 22.5°E. Using these boundaries, the Åland and Archipelago Sea become a part of the Baltic Proper, approximately half of the Arkona Basin is included in the Transition Zone, and the Gulf of Riga becomes a part of the Gulf of Finland. Representative monthly anomaly time series are then obtained by averaging over all grid cells in the regions. The resulting time series can be viewed by the thin gray lines in Figure 11.1.

The temperature variations are found to be the largest in the Transition Zone. The lower variations in the Gulf of Riga and the Gulf of Finland could be attributed to the high prevalence of sea ice as this results in a constant temperature of -1°C. If sea ice forms every winter, a temperature anomaly of 0°C is obtained for all winter months. Furthermore, the large temperature variations of the Transition Zone are most likely strengthened by the variations in the frontal position between the relatively fresh water of the Baltic Sea and the high salinity water of Skagerrak (Høyer and Karagali, 2016).

To focus on interannual variability, 3-year running means are calculated for each time series. To do this, yearly mean temperature anomalies are calculated for all years containing at least 11 months with data. The 3-year running mean is then estimated by averaging over the year prior to and after the given year. Single missing years are allowed in the calculations. The result is shown by the thick red line in Figure 11.1. To easily be able to compare the running means of the different regions, the running means are plotted together in Figure 11.2.

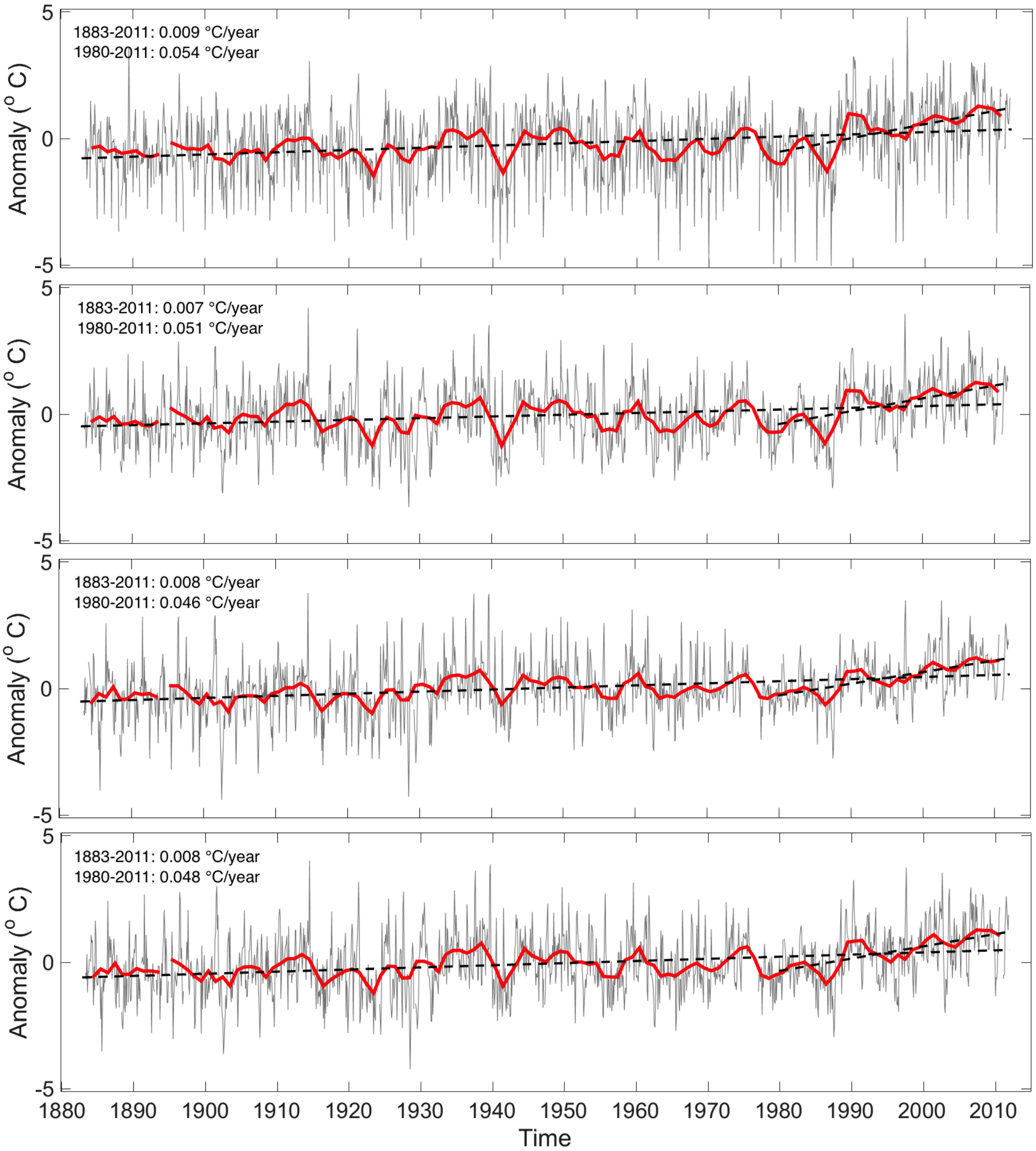

Figure 11.1: Average monthly SST anomalies (thin gray lines), 3-year running mean (thick red lines) and linear temperature trends (dashed black lines) in the Transition Zone (top), the Baltic Proper (second from the top), the Gulf of Bothnia (second from the bottom) and Gulf of Finland (bottom). The slope of the temperature trends are provided in the figures.

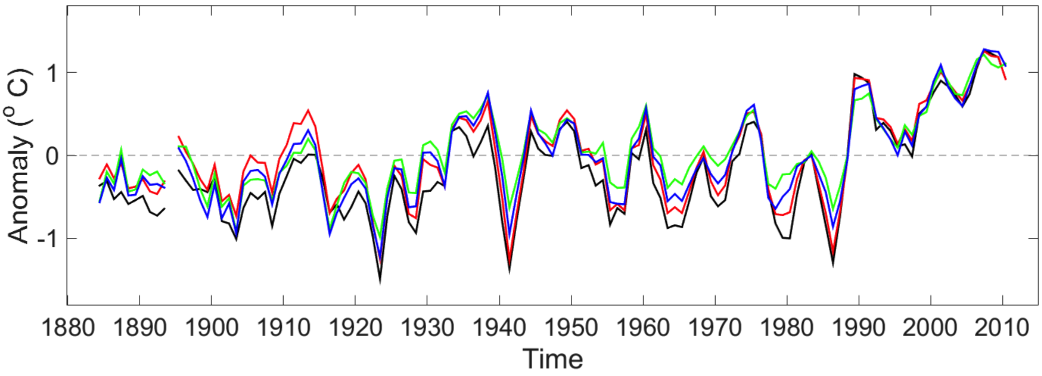

Figure 11.2: 3-year running mean of the yearly mean anomalies in the Transition Zone (black), Baltic Proper (red) Gulf of Bothnia (green) and the Gulf of Finland (blue). The dashed gray line marks 0°C.

Figure 11.2 shows highly correlated SST variations in the four regions. In the beginning of the time series, the running means are mostly negative, while they become mainly positive by the end of the period. Three main minima are found, representing especially cold periods around 1926, 1941 and 1986. All three years have been classified to have severe winters according to Seinä and Palosuo (1996). Negative temperature anomalies were thus expected. The minima in 1941 is believed to be extra pronounced as the winters of 1940 and 1942 were classified as extremely severe. In fact, the maximum annual sea ice extent of both winters reached 420 000 km^2, indicating that the entire Baltic Sea froze. The high prevalance of sea ice in the Gulf of Bothnia and the Gulf of Finland is therefore assumed to cause the more dampened signal in these areas, while the minima is much more pronounced in the Transition Zone and the Baltic Proper. In all four regions, the maximum temperature anomaly is found in 2007, where the anomalies reached 1.28°C, 1.26°C, 1.21°C and 1.28°C in the Transition Zone, Baltic Proper, Gulf of Bothnia and Gulf of Finland, respectively. In general, an unprecedented temperature increase is seen from the 1980's. Besides calculating temperature trends for the entire time series, there is therefore an incentive to also examine temperature trends between 1980 and 2011.

For each of the regions, temperature trends are calculated based on linear regression of the monthly anomaly time series. The resulting fits are shown by the black dashed lines in Figure 11.1 and represent an average warming of 0.009°C/year, 0.007°C/year, 0.008°C/year and 0.008°C/year between 1883 and 2011 in the Transition Zone, Baltic Proper, Gulf of Bothnia and Gulf of Finland, respectively. The average warming between 1980 and 2011 is found to be approximately 6 times higher, with an average warming of 0.054°C/year, 0.051°C/year, 0.046°C/year and 0.048°C/year in the Transition Zone, Baltic Proper, Gulf of Bothnia and the Gulf of Finland, respectively.

The main advantage of the reconstructed monthly mean SST fields is that it is possible to examine the spatial structure of the temperature trends in much greater detail. The linear temperature trends of every position in the Baltic Sea is shown in Figure 11.3, where the left figure shows the trends between 1883 and 2011, while the right figure shows the trends between 1980 and 2011. Figure 11.3 demonstrates that the temperature trends between 1980 and 2011 exceeds the temperature trends between 1883 and 2011 for the entire domain.

Between 1883 and 2011, the greatest temperature trends are found in Kattegat, the Archipelago Sea, along the Northern Quark and in localized regions of the Gulf of Finland, the Gulf of Riga and the Bothnian Sea. Due to the change in axis, it is difficult to infer the position of the largest temperature trends between 1980 and 2011. However, the largest temperature trends are to a large extent found in the same positions as between 1883 and 2011. The main difference in the spatial structure is that the smallest temperature trends are found in the Baltic Proper between 1883 and 2011, while the smallest temperature trends are found in the Bay of Bothnia between 1980 and 2011.

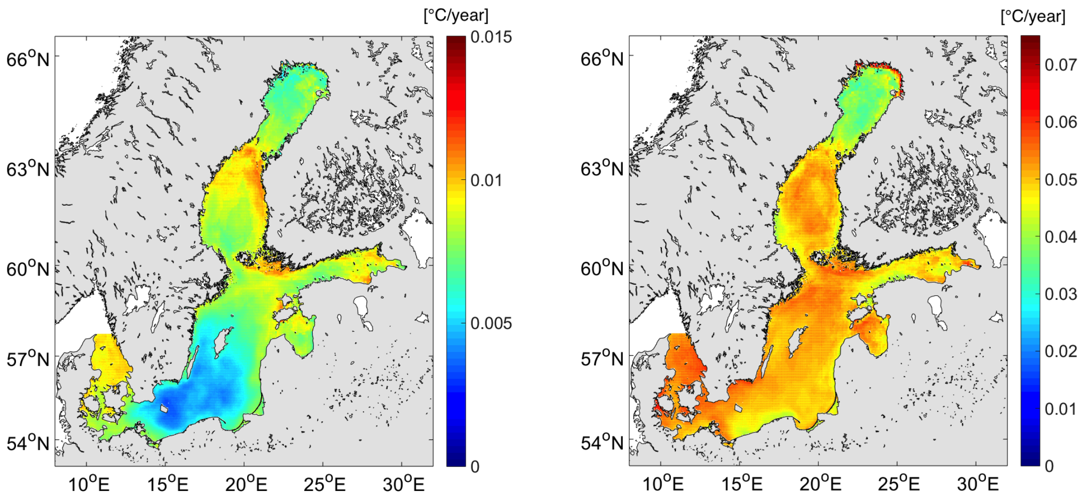

Figure 11.3: Linear temperature trends between 1883 and 2011 (left) and 1980 and 2011 (right). Note the different axis.

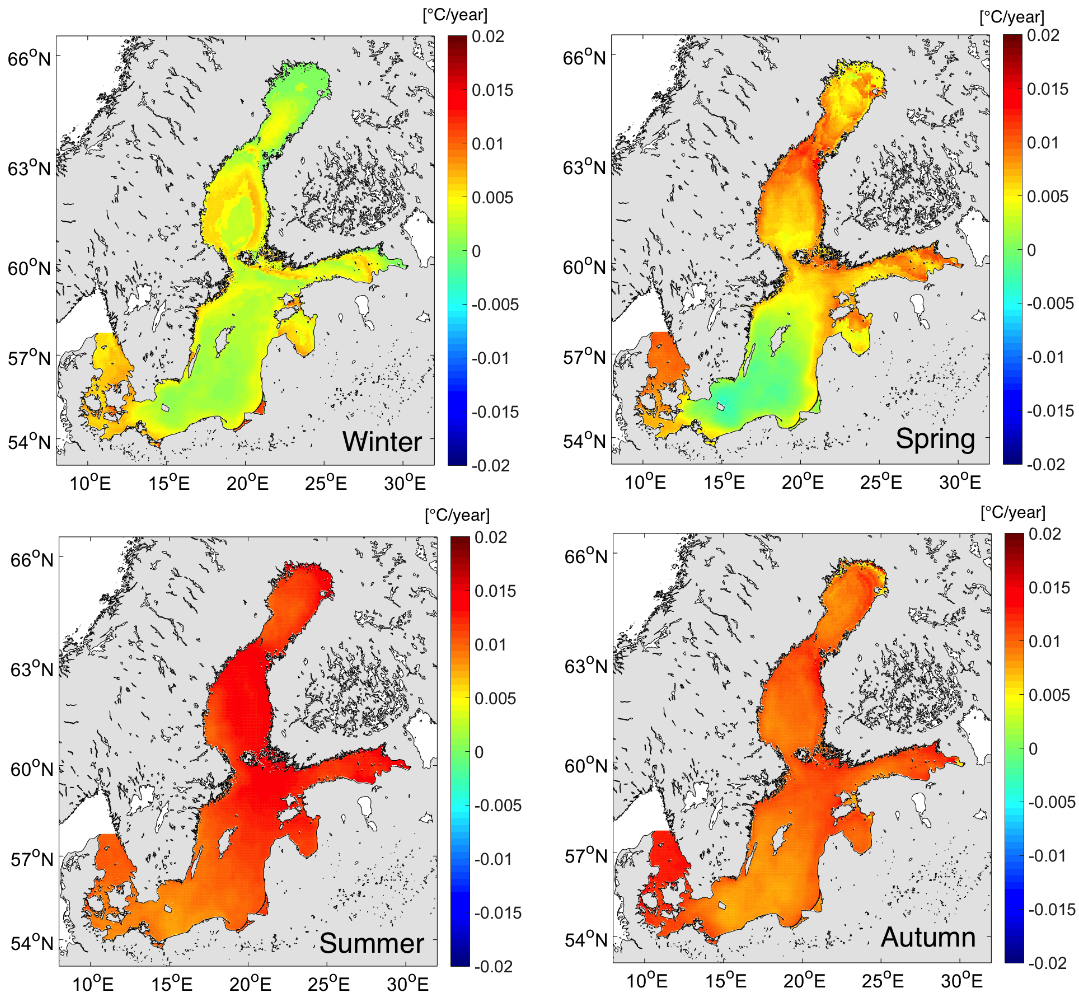

Figure 11.4: Linear temperature trends between 1883 and 2011 during winter (Jan-Mar), spring (Apr-Jun), summer (Jul-Sep) and autumn (Oct-Dec).

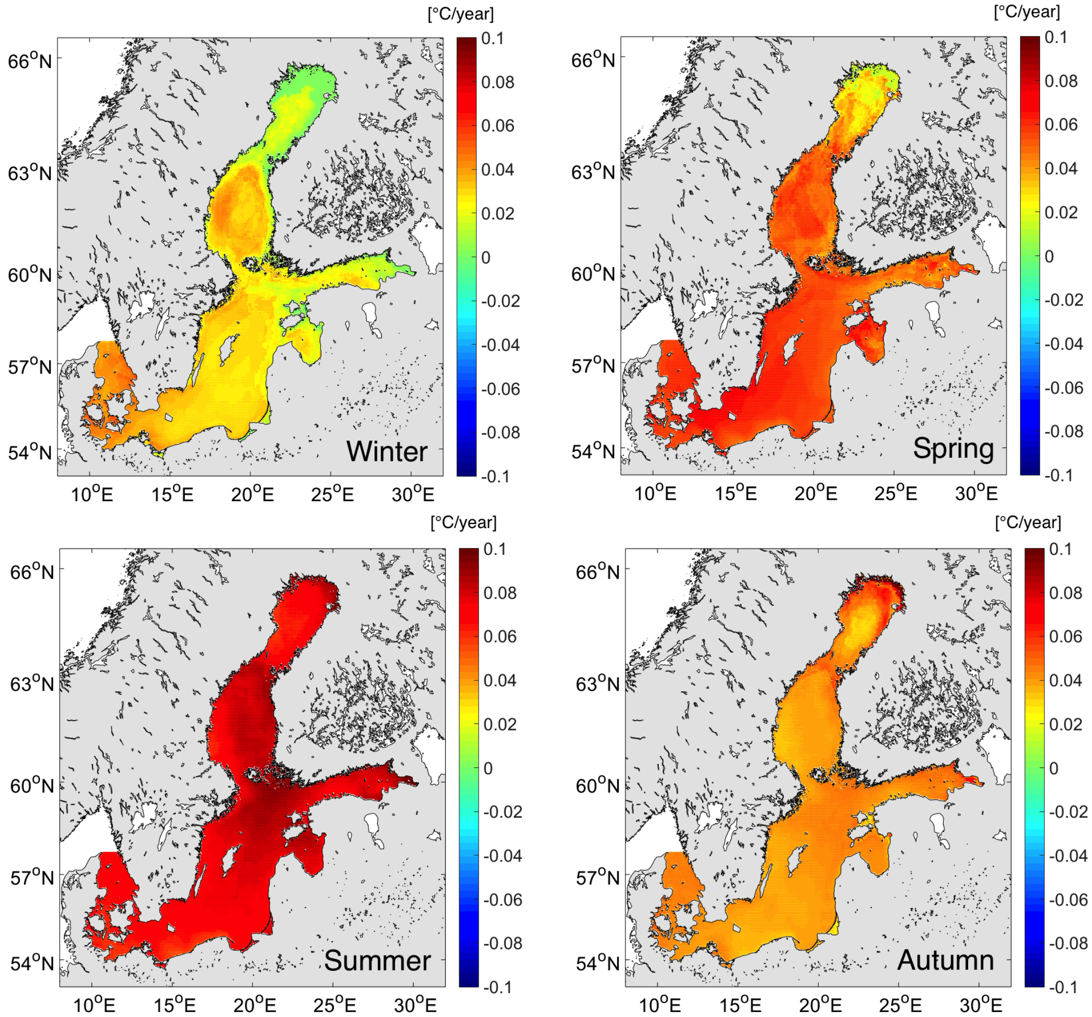

Figure 11.5: Linear temperature trends between 1980 and 2011 during winter (Jan-Mar), spring (Apr-Jun), summer (Jul-Sep) and autumn (Oct-Dec).

Høyer and Karagali (2016) have calculated linear temperature trends between 1982 and 2011 for the DMI SST CDR. The results agree to a large extent with the temperature trends obtained between 1980 and 2011 in this study. However, there is one major difference. Høyer and Karagali (2016) report temperature trends of 0.1°C/year in the eastern extension of the Gulf of Finland. In this study, temperature trends of 0.05°C/year are found in the same region. The difference is believed to be too large to be caused by the difference in the record lengths. Instead, the difference is interpreted as additional proof for a lower accuracy in the reconstructed fields in this region.

To examine the possibility for seasonally varying trends, season specific temperature trends are calculated. The seasonal temperature trends between 1883 and 2011 are shown in Figure 11.4 while the seasonal temperature trends between 1980 and 2011 are shown in Figure 11.5. For both time periods, the smallest temperature trends are generally found during winter while the largest temperature trends are found during summer. Furthermore, it can be seen that relatively large temperature trends are found during autumn between 1883 and 2011, while greater temperature trends are found during spring between 1980 and 2011. The difference is especially large in the Baltic Proper, where the temperature trends during spring are slightly negative between 1883 and 2011, while trends of 0.06°C/year are found between 1980 and 2011. The reason for the varying trends and observed spatial structures will not be speculated on here. Instead, in-depth analysis of the fields are left for future work.

12 Discussion

This thesis has shown that it is possible to reconstruct monthly mean SST fields by combining in situ observations with a satellite based climate data record in a multivariate regression model. The validation in Chapter 10 further demonstrates that it is possible to make reconstructions of high accuracy and temporal stability. This is especially evident in the validation of the Transition Zone, where the yearly validation statistics show an accuracy within 0.38°C and a stability of 0.004°C/decade. The yearly validation statistics for the reconstructed SST in the Baltic Proper and Gulf of Bothnia demonstrate lower accuracies and temporal stabilities in these regions. The accuracy of the Baltic Proper was found to be within 1.33°C while the stability of the record was found to be 0.077°C/decade. In the Gulf of Bothnia, the accuracy and temporal stability was found to be 1.83°C and 0.052°C/decade, respectively.

The Global Climate Observing System (GCOS) have formulated target requirements for the horizontal resolution, accuracy and stability of SST fields used to resolve decadal changes on regional (and global) scales (GCOS, 2006). According to the updated requirements, the SST fields should have a horizontal resolution of 10 km, an accuracy of 0.1°C and a stability of 0.03°C/decade (GCOS, 2011). Even though the requirements are designed for satellite based products of daily resolution, they can still be used to indicate the ability of the reconstructed SST fields to resolve accurate decadal SST signals. After all, the purpose of this thesis is to create historical SST fields that can be used to study long-term SST signals.

While the reconstructed fields meet the requirements for the horizontal resolution, only the reconstructed SST of the Transition Zone fulfills the stability requirements. None of the areas have the accuracy required by GCOS. Furthermore, the reconstructed SST of the Transition Zone is close to meeting all requirements. According to the yearly validation shown in Figure 10.5, 49 out of the 124 years included in the validation have an accuracy within 0.1°C, while 105 out of the 124 years have an accuracy within 0.25°C. As most years are very close to meeting the requirements, the reconstructed SST of the Transition Zone is considered suitable to base long-term SST studies on. Furthermore, the reconstructed monthly mean SST fields were not expected to achieve much higher accuracies than this. After all, it should be remembered that the reconstructions are based on a statistical model and that uncertainties have been introduced at several stages of the reconstructions.

The first set of uncertainties originate from the calculations of the monthly means (Chapter 7). In order to calculate a monthly mean from the Danish, German and Swedish records, the specific month was required to contain at least 15 days with data. It is therefore possible that monthly means based on a low number of observations deviate from the true monthly mean of the station. This is especially a risk when the majority of the observations originate from one half of the month.

The inconsistency between the Finnish records and the CDR proves that the largest uncertainties are found in the Finnish records. This is most likely an artifact of the sparse sampling rate. In addition, the SST measurements were performed during the afternoon and may therefore contain diurnal warming signals. It is therefore possible that the true SST development at the Finnish in situ stations have been aliased into wrong frequencies. The relatively low correlations between the CDR and Finnish in situ records in Table 7.1 supports the theory.

A second set of uncertainties originate from the construction of the long and continuous records, usually referred to as the extended records (Chapter 9). In the constructions, the reference records were extended with in situ records from adjacent stations. In order to keep the seasonal cycle of the reference station, the linear regression parameters relating the two records were calculated specifically for each month. In hindsight, it would have been better to base the reconstructions on SST anomalies. One set of linear regression parameters would then be able to relate the entire records. In result, the linear regression parameters would have been based on a greater number of observations and the extended records would most likely be more accurate.

The accuracy of the extended records are assumed to vary with time as the extended records are based on different in situ records for different time periods. The accuracy is assumed to be the lowest when the reference record is extended with an in situ record that has no temporal overlap with the reference record. When there is no temporal overlap, it is not possible to check the correlation between the records, check if there exist any biases, or calculate the RMSE. These can only be checked at the intermediate stages, such as when the CDR is used to relate the two records. Furthermore, it is assumed that any biases are removed from the intermediate reconstructions. It is also assumed that the correlation between the records is high if the correlation at all intermediate stages are high. After all, the data sets are obtained from closely located stations. However, it was sometimes necessary to extend the reference record with an in situ record that had no temporal overlap with the reference record, CDR or any other in situ record (case 1 in Section 9.1.2). The SST obtained from such record is considered to be the least accurate as the linear regression parameters obtained from the corresponding CDR time series are assumed to be valid for the in situ records despite representing a completely different time period and measuring technique.

The varying accuracies of the extended records affect the accuracy of the reconstructed SST fields as the weights assigned to each record in the multivariate regression model is constant in time. It is therefore expected that the accuracy of the reconstructed monthly mean SST fields reflects the accuracies of the extended records. For example, this could be the cause for the relatively low accuracies in the Baltic Proper prior to 1923, as seen in Figure 10.5. In Figure 10.2, it is shown that ER D has the largest influence on the reconstructions in the Baltic Proper. Utklippan was used to construct the record between 1880 and 1922 despite not having a temporal overlap with the reference record, the CDR or any other in situ records. This could have caused an inconsistency within ER D and be the reason for the shift in mean bias in the Baltic Proper prior to 1923. Yet, an even more plausible explanation is that only in situ stations located in the northwestern Baltic Proper (index 58, 60 and 63) were used to validate the Baltic Proper in this period. It could therefore also indicate that the reconstructions of the Arkona and Bornholm Basins are more accurate than the reconstructions of the Western Gotland Basin.

Due to the uneven coverage and quality of the extended records, it is not surprising that the accuracy and stability of the reconstructed SST fields vary within the Baltic Sea. The validation in Chapter 10 have demonstrated that the accuracy of the reconstructed SST fields decreases away from the Transition Zone. This is believed to be caused by the increasing influence of ER E and ER F in the multivariate regression model away from the Transition Zone. These records are based on Finnish in situ records and have therefore been shown to be of lower quality than the other extended records (Table 9.2 and Figure 9.6). It is therefore believed that it will be possible to increase the accuracy of the reconstructions outside the Transition Zone by increasing the quality of ER E and ER F. After all, the correlation patterns in Figure 8.4 demonstrates that ER E and ER F should be able to describe the SST development in the majority of the Baltic Proper, Gulf of Bothnia and the Gulf of Finland.

Furthermore, additional records have to be included in the model in order to resolve the SST development in the northernmost extension of the Gulf of Bothnia and the easternmost extension of the Gulf of Finland. Figure 10.4 shows that the model quality is low in these areas despite being based on CDR time series instead of the extended records. This demonstrates that the northernmost extension of the Gulf of Bothnia and the easternmost extension of the Gulf of Finland are underrepresented in the model. These areas are believed to need extra treatment as they have their own set of oceanographic characteristics due to the strong influence of river runoff and sea ice (Omstedt and Axell, 2003). Similarly, Figure 10.4 and 8.4 also demonstrates a need to include extra coastal records in the model. If enough records are included in the model, it would also be possible to set a constraint for the spatial influence of each record. In return, it would be possible to limit the amount of noise included in the model, such as the relatively large weights assigned to ER A and ER B in the Gulf of Bothnia.

Overall, the validation shows that the accuracy of the reconstructed SST fields decreases away from the Transition Zone. However, it is rather difficult to limit the area in which accurate studies of climatic SST signals can be based. By combining the information obtained from the validation statistics in Table D.1, the weights found in Figure 10.2 and the correlation patterns shown in Figure 9.6 it is assumed that the accuracy and temporal stability of the reconstructed fields in the Arkona Basin, Bornholm Basin and the southern regions of the Gotland Basins is similar to that found in the Transition Zone.

A more robust validation is needed in order to determine the accuracy in the remainder of the Baltic Sea. This is especially the case in the Gulf of Finland, which was not possible to validate at all. In the Gulf of Bothnia, 5 out of the 6 records used for the validation were Finnish. These are expected to differ from the reconstructed monthly mean SST fields as they represent day values. The validation results obtained for the Gulf of Bothnia should therefore be used as an indication of the accuracy, rather than the absolute accuracy.

From the above discussion, it can be concluded that both the model and the validation of the reconstructed monthly mean SST fields are limited by the quality and amount of in situ records. It is therefore emphasized that a large part of this study has been dedicated to collecting and analyzing the in situ records included in the thesis. Open source data was available from SMHI, while all other in situ records were accessed internally from DMI, FMI and BSH. It is possible that additional records are available at these institutes. For example, it is known that additional data can be purchased from SMHI, and that FMI has access to additional lightship and coastal records. It is also possible that Poland, Lithuania, Latvia and Russia have access to SST records as these countries also border the Baltic Sea. Due to the time limitation of this thesis, it was not possible to spend more time on data collection. However, to increase the accuracy of the product, an even more thorough data collection is needed. Additional data representing the Gulf of Bothnia, Gulf of Finland, Gulf of Riga and the eastern Baltic Proper is especially needed.

In addition to collecting more data, the data coverage could also increase by making further analyses on the already accessed data. For example, more effort could be put into analyzing the errors obtained by basing the monthly means on observations from different times and depths. For example, it was not possible to reconstruct any monthly mean SST fields in 1976 due to the lack of data in all Danish and Swedish records. Furthermore, Læsø Trindel, Anholt Knob and Gedser Rev/Kadetrenden have recorded SST at a 5 m depth during this period. These values were not used to keep the monthly mean SST records consistent, as all other values originated from 0 m. However, as the records are based on 07:00 - 08:00 LT measurements, the inconsistency would probably be very small. After all, the SST in the surface layer should be almost constant at this time, as illustrated in Figure 3.1. The data coverage of the in situ records could thus increase by basing the monthly means on observations made at different times and depths.

13 Conclusion and outlook

This thesis demonstrates that it is possible to reconstruct monthly mean SST fields by combining in situ records with the DMI SST CDR in a multivariate regression model. The resulting fields have a spatial resolution of 0.03° x 0.03° and covers the Baltic Sea from 1883 to 2011. This is the first time that 2D SST fields representing the Baltic Sea have been reconstructed so far back in time only using observations. The reconstructed fields are thus independent of atmospheric and hydrodynamical models and can therefore be used to validate model reanalysis.

Yearly validation statistics have demonstrated that the reconstructed SST fields are the most accurate in the Transition Zone. Here, an accuracy of 0.38°C with insignificant temporal trends are found. It is argued that similar accuracies apply in the Arkona Basin, Bornholm Basin and the southern region of the Gotland Basins. The reconstructed SST fields of these areas are close to meeting GCOS requirements and are therefore considered well-suited for long-term climatic studies. A more robust validation is needed in order to characterize the accuracy and stability of the reconstructed fields in the central and northern Baltic Sea. However, the analyses in this thesis have indicated that the accuracies of these regions are lower than in the Transition Zone.

The accuracy and temporal stability of the reconstructed SST fields are shown to be dependent on the quality and spatial representativeness of the extended records. It is therefore possible to increase the accuracy of the fields by basing the model on a greater number of high-quality records. Future efforts should therefore be put into constructing an even greater data base of in situ records.

Representative monthly anomaly time series were calculated based on the reconstructed SST fields for the Transition Zone, Baltic Proper, Gulf of Bothnia and Gulf of Finland. Analyses of the time series demonstrated large variability on monthly time scales, while the regions were shown to be highly correlated on interannual and decadal time scales. For each region, linear temperature trends were calculated between 1883 and 2011, and 1980 and 2011. The linear temperature trends between 1883 and 2011 were found to vary between 0.07-0.09°C/decade, while an average warming of 0.46-0.54°C/decade was found between 1980 and 2011. Using the complete fields, it was also demonstrated that spatial and seasonal variations in the linear SST trends exist within each region.

More in-depth analyses of the fields are left for future work. Such analyses could include investigations on the models ability to reconstruct sea ice. The multivariate regression model was found to reconstruct negative temperatures and it would therefore be very interesting to see how well these temperatures agree with the histoircal ice fields. It is possible that the historical ice fields weren't needed after all.

Bibliography

Aster, R. C., B. Borchers, and C. H. Thurber (2013). *Parameter Estimation and Inverse Problems* (2nd ed.). Elsevier.

Casey, K. S., T. B. Brandon, P. Cornillon, and R. Evans (2010). *Oceanography from Space: Revisited.* Springer, pp. 273-287.

Eilola, K. and A. Stigebrandt (1998). Spreading of juvenile freshwater in the Baltic proper. *Journal of Geophysical Research 103*(C12), pp. 27795–27807.

Emery, W. J. and R. E. Thomson (1997). *Data Analysis Methods in Physical Oceanography* (2nd and Revised ed.). Elsevier.

Feistel, R., G. Nausch, and N. Wasmund (2008). *State and Evolution of the Baltic Sea, 1952-2005: A detailed 50-Yea Survey of Meteorology and Climate, Physics, Chemisry, Biology and Marine Environment.* John Wiley & Sons, Inc.

GCOS (2006). Systematic observation requirements for satellite-based products for climate: Supplemental details to the satellite-based component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC".

GCOS (2011). Systematic observation requirements for satellite-based products for climate 2011 update: Supplemental details to the satellite-based component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update)".

Høyer, J. L., P. L. Borgne, and S. Eastwood (2014). A bias correction method for Arctic satellite sea surface temperature observations. *Remote Sensing of Environment 146*, pp. 201–213.

Høyer, J. L. and I. Karagali (2016). Sea Surface Temperature Climate Data Record for the North Sea and Baltic Sea. *Journal of Climate 29*(7), pp. 2529–2541.

Høyer, J. L., J. I. Karagali, G. Dybkjaer, and R. Tonboe (2012). Multi sensor validation and error characteristics of Arctic satellite sea surface temperature observations. *Remote Sensing of Environment 121*, pp. 335–346.

Høyer, J. L. and J. She (2007). Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea. *Journal of Marine Systems 65*, pp. 176–189.

Høyer, J. L. (2002). *On the Combination of Satellite and In Situ Observations to Detect Oceanic Processes.* Ph. D. thesis, University of Copenhagen.

Karagali, I. and J. L. Høyer (2014). Characterisation and quantification of regional diurnal SST cycles from SEVIRI. *Ocean Science 10*, pp. 745–758.

Karagali, I., J. Høyer, and C. Hasager (2012). SST diurnal variability in the North Sea and the Baltic Sea. *Remote Sensing of Environment 121*, 159–170.

Knauss, J. A. (1997). *Introduction to Physical Oceanography* (2nd ed.). Waveland Press, Inc.

Kullenberg, G. and T. S. Jacobsen (1981). The Baltic Sea: an Outline of its Physical Ocenaography. *Marine Pollution Bulletin 12*(6), 183–186.

Madsen, K. S. (2006). Coastal satellite altimetry: Sea surface topography in the North Sea - Baltic Sea system. Master's thesis, University of Copenhagen.

Madsen, K. S. (2009). *Recent and future climatic changes in temperature, salinity, and sea level of the North Sea and the Baltic Sea* . Ph. D. thesis, University of Copenhagen.

Merchant, C. J., O. Embury, N. A. Rayner, D. I. Berry, G. K. Corlett, K. Lean, K. L. Veal, E. C. Kent, D. T. Llewellyn-Jones, J. J. Remedios, and R. Saunders (2012). A 20 year independent record of sea surface temperature for climate from Along-Track Scanning Radiometers. *Journal of Geophysical Research 117*, 299–310.

O. Embury, C. J. M. (2012). A reprocessing for climate of sea surface temperature from the along-track scanning radiometers: A new retrieval scheme. *Remote Sensing of Environment 116*, pp. 47–61.

Omstedt, A. (1990). A coupled one-dimensional sea ice-ocean model applied to a semi-enclosed basin. *Tellus 42A*, 568–582.

Omstedt, A. and L. B. Axell (2003). Modeling the variations of salinity and temperature in the large Gulfs of the Baltic Sea. *Continental Shelf Research 23*, 265–294.

Omstedt, A. and A. Rutgersson (2000). Closing the water and heat cycles of the Baltic Sea. *Meteorologische Zeitschrift, 9*(1), 59–66.

Robinson, I. S. (2010). *Discovering the ocean from space: the unique applications of satellite oceanography.* Springer/Praxis Publishing.

Seinä, A. and E. Palosuo (1996). The Classification of the Maximum Annual Extent of Ice Cover in the Baltic Sea 1720-1995. *MERI-Report Series of the Finnish Institute of Marine Research 27*, 79–91.

Smith, D., C. Mutlow, J. Delderfield, B. Watkins, and G. Mason (2012). ATSR infrared radiometric calibration and in-orbit performance. *Remote Sensing of Environment 116*, 4–16.

Stigebrandt, A. (2001). Physical Oceanography of the Baltic Sea. *A System Analysis of the Baltic Sea 148*, 19–73.

Stigebrandt, A. and B. G. Gustafsson (2003). Response of the Baltic Sea to climate change - theory and observations. *Journal of Sea Research 49*, 243–256.

The BACC Author Team (2008). Assessment of climate change for the Baltic Sea Basin.

Walton, C., W. G. Pichel, and J. F. Sapper (1998). The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. *Journal of Geophysical Research 103*(C12), pp. 27999–28012.

Winsor, P., J. Rodhe, and A. Omstedt (2001). Baltic Sea ocean climate: an analysis of 100 yr of hydrographic data with focus on the freshwater budget. *Climate Research 18*, pp. 5–15.

Woodruff, S. D., R. J. Slutz, R. L. Jenne, and P. M. Steurer (1987). A comprehensive ocean–atmosphere data set. *Bulletin American Meteorological Society 68*, pp. 1239–1250.

Woodruff, S. D., S. J. Worley, S. J. Lubker, Z. Ji, J. E. Freeman, D. I. Berry, P. Brohan, E. C. Kent, R. W. Reynolds, S. R. Smith, and C. Wilkinsond (2011). ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *International Journal of Climatology 31*, pp. 951–967.

A Appendix

St. Num.	Station Name	RMSE	Bias	Std	Min	Max	N
Danish maintained coastal stations							
20098	Frederikshavn	0.58	0.14	0.56	-2.21	1.42	91
20108	Hirsholm	1.06	0.80	0.70	-0.51	2.29	64
22160	Sletterhage	0.64	0.33	0.55	-1.33	1.65	155
26478	Sønderborg	0.68	-0.05	0.68	-3.93	1.33	189
27022	Anholt	0.88	0.32	0.83	-1.98	1.89	65
28118	Middelfart	0.82	0.06	0.82	-3.50	2.25	203
28548	Bagenkop	0.54	0.21	0.50	-1.76	1.26	198
29007	Gniben	0.58	0.36	0.45	-0.49	2.21	90
29008	Rørvig	1.00	0.29	0.96	-2.27	2.68	189
29118	Refsnæs	0.92	-0.05	0.94	-1.67	1.22	23
30332	Langelinie	1.16	-0.01	1.22	-2.51	1.65	11
30342	Middelgrundsfortet	0.49	0.19	0.45	-0.81	1.81	149
31062	Rødvig	1.17	-0.76	0.89	-3.07	1.55	180
31248	Klintholm	1.15	-0.17	1.14	-2.77	2.26	91
31308	Storstrøm/Farø	1.23	-0.21	1.21	-3.04	2.29	183
31572	Rødbyhavn	1.02	-0.11	1.01	-3.54	2.39	186
32002	Christiansø	0.48	0.33	0.34	-0.59	1.32	192
Danish automatic coastal stations							
23293	Fredericia	0.88	0.38	0.80	-1.43	2.31	112
26457	Fynshav	0.43	0.19	0.39	-0.72	1.53	86
28234	Slipshavn	0.60	0.15	0.58	-1.37	1.97	140
29393	Korsør	0.73	0.27	0.68	-1.63	1.66	132
30017	Hornbæk	0.71	-0.37	0.61	-2.18	1.43	129
30336	København	0.61	-0.33	0.52	-2.30	0.82	139
31573	Rødby	0.53	0.06	0.53	-1.56	1.48	142
31616	Gedser	0.87	-0.09	0.87	-1.82	1.71	119
32048	Tejn	0.43	0.17	0.39	-0.82	1.35	122
Danish lightships							
6087	Anholt Knob	0.41	-0.00	0.41	-0.93	0.61	26
6147	Møn SE	0.58	0.40	0.43	-0.49	1.30	63
6183	Drogden Fyr	0.60	-0.15	0.58	-1.73	2.55	181

St. Num.	Station Name	RMSE	Bias	Std	Min	Max	N
Swedish buoys							
35070	Trubaduren	0.92	-0.09	0.92	-4.57	3.29	168
33001	Läså Ost	0.98	0.18	0.97	-2.14	2.85	59
35068	Fladen	0.51	-0.13	0.50	-1.64	0.77	53
35067	Oskarsgrundet	0.64	0.05	0.64	-2.97	3.05	155
35063	Ölands Södra Grund	0.55	0.35	0.43	-0.32	1.39	45
35057	Gustav Dahlén	0.66	0.09	0.66	-1.15	1.53	36
35056	Almagrundet	0.53	0.19	0.50	-1.32	1.23	44
35054	Svenska Björn	0.69	0.50	0.47	-0.48	2.23	63
German buoys							
-	Kiel	0.45	0.12	0.43	-1.26	1.30	260
-	Fehmarn Belt	0.53	0.18	0.50	-1.23	1.61	153
-	Darβer Schwelle	1.16	0.38	1.10	-3.66	2.62	111
-	Oder Bank	0.44	0.13	0.42	-1.50	1.05	118
-	Arkona Becken	1.01	0.16	1.00	-2.40	1.98	99
Finnish maintained coastal stations - Method 1							
-	Utö	0.67	0.23	0.64	-1.02	1.91	55
-	Seili	1.22	-0.81	0.91	-3.47	1.62	223
-	Valassaaret	0.84	-0.02	0.84	-3.89	2.61	153
-	Krunnit	1.76	-0.93	1.51	-5.05	1.41	35
-	Harmaja	1.20	0.31	1.17	-3.56	3.30	63
-	Tvärminne-A	1.45	0.11	1.45	-4.53	4.05	234
-	Tvärminne-B	2.00	-0.16	2.00	-5.61	3.93	234
Finnish maintained coastal stations - Method 2							
-	Utö	0.83	0.09	0.83	-2.09	2.72	116
-	Seili	1.28	-0.75	1.03	-4.17	1.72	252
-	Valassaaret	0.87	-0.02	0.87	-2.20	2.65	191
-	Krunnit	1.33	-0.73	1.13	-4.15	1.15	34
-	Harmaja	1.33	0.29	1.31	-2.60	2.76	69
-	Tvärminne-A	1.61	0.17	1.60	-4.60	4.05	245
-	Tvärminne-B	2.00	-0.08	2.00	-4.97	4.31	236

Table A.1: Comparison between in situ records and the DMI SST CDR. The root mean square error (RMSE), bias, standard deviation (Std), minimum difference (Min) and maximum difference (Max) are given in °C. N represents the number of match-ups on which the comparisons were based.

B Appendix

Index	St. Num.	Station Name	Type	Active Years	Lat.	Lon.
1	20108	Hirsholm	D M	1990-1995	57.48	10.57
2	20098	Frederikshavn	D M	1939-1990	57.43	10.57
3	35016	Vinga	S L	1930-1965	57.57	11.47
4	35070	Trubaduren	S B	1978-2004	57.60	11.79
5	6047	Læsø Trindel	D L	1931-1943	57.47	11.34
”	”	”	”	1943-1945	57.52	11.26
”	”	Læsø Nord	”	1945-1975	57.53	11.34
”	”	Læsø Trindel	”	1975-1977	57.47	11.42
6	6057	Læsø Rende	D L	1931-1943	57.21	10.69
”	”	”	”	1943-1962	57.21	10.73
”	”	”	”	1962-1965	57.20	10.73
7	6067	Østre Flak	D L	1931-1942	56.97	10.90
”	”	Ålborg Bugt	”	1943-1973	56.85	10.80
8	33001	Läså Ost	S B	2001-2009	57.22	11.57
9	35068	Fladen	S B	1988-1999	57.22	11.83
10	35015	Fladen	S L	1893-1969	57.17	11.67
11	35014	Varberg	S M	1879-1887	57.10	12.22
12	6087	Anholt Knob	D L	1931-1945	56.77	11.86
”	”	”	”	1945-1948	56.75	11.99
”	”	Anholt Nord	”	1948-1975	56.85	11.80
”	”	Anholt Knob	”	1975-1985	56.75	11.88
13	27022	Anholt	D M	1990-1999	56.72	11.52
14	6157	Kattegat S	D L	1966-1975	56.25	12.25
15	29008	Hundested	D M	1931-1942	55.97	11.85
”	”	Rørvig	”	1942-1999	55.95	11.77
16	6077	Schultz's Grund	D L	1931-1944	56.15	11.19
”	”	Kattegat SW	”	1945-1971	56.10	11.15
17	29007	Gniben	D M	1991-1999	56.00	11.28
18	22160	Sletterhage	D M	1931-1985	56.10	10.52
19	23293	Fredericia	D A	2002-2014	55.57	9.75
20	28118	Middelfart	D M	1931-1999	55.52	9.73
21	26457	Fynshav	D A	2002-2014	55.00	9.98
22	26478	Sønderborg	D M	1931-1999	54.92	9.78
23	26458	Mommark	D M	1931-1967	54.93	10.05
24	28548	Keldsnor	D M	1931-1952	54.73	10.72
”	”	Bagenkop	”	1952-1999	54.75	10.67
25	29118	Refsnæs	D M	1931-1983	55.70	11.03

Index	Number	Station Name	Type	Active Years	Lat	Lon
26	6127	Halsskov Rev	D L	1931-1973	55.34	11.05
27	28234	Slipshavn	D A	1999-2015	55.28	10.83
28	29393	Korsør	D A	2000-2014	55.33	11.15
29	31308	Masnedø	D M	1931-1937	55.00	11.88
”	”	Storstrøm	”	1939-1988	54.97	11.88
”	”	Farø	”	1988-1998	54.95	11.98
30	-	Kiel	G B	1987-2015	54.30	10.16
31	-	Fehmarn Belt	G B	1985-2012	54.36	11.09
32	31573	Rødby	D A	1999-2014	54.65	11.35
33	31572	Rødbyhavn	D M	1931-1999	54.65	11.35
34	31616	Gedser	D A	2001-2014	54.57	11.93
35	6147	Gedser Rev	D L	1931-1939	54.45	12.18
”	”	”	”	1945-1955	54.42	12.15
”	”	”	”	1955-1976	54.45	12.18
”	”	Kadetrenden	”	1976-1979	54.78	12.75
”	”	Møn SE	”	1979-1988	54.80	12.78
36	-	Daβler Schwelle	G B	1998-2015	54.42	12.42
37	35013	Svinbådan	S L	1884-1960	56.17	12.52
38	30017	Hornbæk	D A	2000-2015	56.10	12.47
39	6167	Lappegrund	D L	1931-1969	56.07	12.63
40	30342	Middelgrundsfortet	D M	1931-1999	55.72	12.67
41	30332	Langelinie	D M	1936-1983	55.70	12.60
42	30336	København	D A	1999-2015	55.70	12.60
43	35012	Kalkgrundet	S L	1883-1922	55.62	12.88
44	35067	Oskarsgrundet	S B	1983-1999	55.60	12.85
45	35011	Oskarsgrundet	S L	1883-1961	55.58	12.85
46	6183	Drogden Fyrskib	D L	1900-1937	55.53	12.72
”	”	Drogden Fyr	D M	1937-1998	55.53	12.72
47	35010	Falsterborev	S L	1860-1972	55.30	12.78
48	31062	Rødvig	D M	1931-1999	55.25	12.38
49	31248	Klintholm	D M	1931-1990	54.95	12.47
50	-	Oder Bank	G B	1997-2015	54.05	14.10
51	-	Arkona Becken	G B	2002-2015	54.53	13.52
52	32048	Tejn	D A	2001-2015	55.25	14.83
53	32002	Christiansø	D M	1931-1998	55.32	15.18
54	35009	Utklippan	S M	1880-1922	55.95	15.70
55	35008	Ölands rev	S L	1923-1951	56.12	16.57
56	35063	Ölands Södra Grund	S B	1979-1991	56.07	16.68
57	35006	Hävringe	S L	1951-1967	58.55	17.52
58	35007	Landsort	S M	1880-1922	58.48	17.87
59	35057	Gustav Dahlén	S B	1982-1987	58.60	17.47
60	35005	Kopparstenarna	S L	1883-1915	58.58	19.15

Index	Number	Station Name	Type	Active Years	Lat	Lon
61	35056	Almagrundet	S B	1979-1992	59.15	19.13
62	35054	Svenska Björn	S B	1984-1992	59.47	20.35
63	-	Bogskär	F M	1899-1968	59.52	20.38
64	35004	Svenska Björn	S L	1883-1968	59.58	19.78
65	-	Utö	F M	1900-2005	59.78	21.34
66	-	Tvärminne-A	F M	1926-2012	59.85	23.25
67	-	Tvärminne-B	F M	1967-2012	59.85	23.25
68	-	Porkkala	F L	1899-1968	59.93	24.42
69	-	Harmaja	F M	1900-1993	60.10	24.96
70	-	Seili	F M	1967-2012	60.25	21.96
71	-	Märket	F M	1906-1970	60.31	19.15
72	35003	Grundkallen	S L	1883-1960	60.50	18.92
73	35002	Finngrundet	S L	1860-1969	61.03	18.52
74	35001	Sydostbrotten	S L	1883-1963	63.32	20.17
75	-	Valassaaret	F M	1919-2009	63.39	21.08
76	-	Snipan	F L	1900-1960	63.43	20.67
77	-	Kemi	F L	1900-1974	65.35	24.35
78	-	Krunnit	F M	1968-2007	65.37	24.89

Table B.1: Indexes assigned to the in situ records. The first letter under 'Type' represents the nationality of the station (D = Denmark, F = Finland, G = Germany, S = Sweden) and the second letter represents the type of station (A = Automatic coastal, B = Buoy, L = Lightship, M = Maintained coastal).

C Appendix

Reconstruction	Reference	Extending	RMSE (°C)	ρ	Min	Max
Extended Record A						
A2 - A1	A2	A1	0.41	0.92	45	55
A2 - A3	A2	A3	0.26	0.97	23	29
A2 - A4	A2	A4	0.26	0.96	9	22
A2 - A5	A2	A5	0.44	0.90	17	24
A2 - A6	SA5	A6	0.36	0.90	4	7
”	A5	SA5	0.47	0.88	6	8
”	A2	A5	0.44	0.90	17	24
A2 - A7	SA5	A7	0.28	0.96	4	6
”	A5	SA5	0.47	0.88	6	8
”	A2	A5	0.44	0.90	17	24
A2 - A8	SA5	A8	0.52	0.84	8	12
”	A5	SA5	0.47	0.88	6	8
”	A2	A5	0.44	0.90	17	24
Extended Record B						
B1 - B2	B1	B2	0.26	0.96	15	21
B1 - B3	B2	B3	0.25	0.95	9	11
”	B1	B2	0.26	0.96	15	21
B1 - B4	B3	B4	0.50	0.89	10	16
”	B2	B3	0.25	0.95	9	11
”	B1	B2	0.26	0.96	15	21
Extended Record C						
C2 - C1	C2	C1	0.33	0.95	14	23
C2 - C3	C2	C3	0.34	0.96	14	44
C2 - C4	C2	C4	0.42	0.93	39	46
C2 - C5	C2	C5	0.46	0.94	11	14
C2 - C6	SC2	C6	0.40	0.93	9	13
”	C2	SC2	0.45	0.91	11	17
Extended Record D						
D3 - D1	SD3	SD1	0.38	0.94	23	30
D3 - D2	D3	D2	0.51	0.85	13	16
D3 - D4	SD3	D4	0.49	0.87	6	13
”	D3	SD3	0.30	0.97	15	17
D3 - D5	SD3	D5	0.29	0.96	9	11
”	D3	SD3	0.30	0.97	15	17

Reconstruction	Reference	Extending	RMSE (°C)	ρ	Min	Max
Extended Record E						
E1 - E2	E1	E2	0.44	0.94	8	69
E1 - E3	E1	E3	0.60	0.88	9	76
E1 - E4	E1	E4	0.89	0.73	8	49
E1 - E5	E4	E5	0.64	0.73	10	14
"	E1	E4	0.89	0.74	8	49
Extended Record F						
F3 - F1	F3	F1	0.85	0.82	12	20
F3 - F2	F3	F2	0.78	0.88	22	57
F3 - F4	F3	F4	0.67	0.91	40	46

Table C.1: Statistics for the construction of the long and continuous records. The accuracy of the reconstructed temperatures are checked by calculating the RMSE and correlation with the reference record. 'Min' and 'Max' represent the minimum and maximum number of months that the model parameters were based on. When there is no time overlap between the reference station and extending station, the temperatures are reconstructed in several steps. The statistics for all intermediate steps are shown in gray text.

D Appendix

Index	Station Name	ρ	RMSE	Bias	Std	Min	Max	N
4	Trubaduren	0.77	1.00	0.07	1.00	-2.85	6.91	196
5	Læsø Trindel	0.92	0.43	0.05	0.42	-1.38	1.66	505
6	Læsø Rende	0.92	0.44	0.04	0.44	-1.57	1.56	358
7	Østre Flak	0.92	0.46	-0.01	0.46	-1.65	1.65	440
8	Läså Ost	0.82	0.91	0.15	0.90	-1.62	2.64	61
9	Fladen	0.86	0.56	0.02	0.57	-1.34	1.06	53
11	Varberg	0.74	0.96	0.55	0.80	-1.35	2.21	47
14	Kattegat S	0.96	0.31	-0.01	0.32	-0.65	0.87	105
15	Hundested	0.90	0.96	0.29	0.91	-2.78	2.53	746
16	Schultz's Grund	0.96	0.34	0.00	0.34	-1.16	1.18	457
17	Gniben	0.95	0.62	0.45	0.43	-0.46	1.59	92
18	Sletterhage	0.93	0.53	0.00	0.53	-2.23	1.50	621
19	Fredericia	0.91	0.86	0.30	0.81	-1.16	2.90	107
20	Middelfart	0.92	0.66	-0.02	0.66	-1.79	2.41	789
21	Fynshav	0.93	0.48	0.10	0.47	-1.07	1.52	84
22	Sønderborg	0.92	0.54	0.05	0.54	-1.93	1.60	797
23	Mommark	0.94	0.42	0.11	0.41	-1.51	1.70	404
25	Refsnæs	0.94	0.70	0.05	0.69	-1.87	1.98	610
28	Korsør	0.93	0.78	0.28	0.73	-1.52	1.87	129
29	Masnedø	0.85	0.76	0.11	0.75	-3.04	2.39	728
31	Fehmarn Belt	0.86	0.60	0.26	0.55	-1.19	1.85	150
32	Rødby	0.94	0.58	0.06	0.58	-1.96	1.69	140
33	Rødbyhavn	0.87	0.94	0.13	0.93	-3.48	5.29	770
34	Gedser	0.95	0.90	-0.21	0.88	-2.04	1.39	114
35	Gedser Rev	0.92	0.58	0.24	0.53	-1.48	1.94	571
36	Darβer Schwelle	0.87	1.04	0.40	0.96	-2.20	2.40	108
39	Lappegrund	0.96	0.32	-0.15	0.28	-1.34	0.50	436
43	Kalkgrundet	0.92	0.55	0.12	0.54	-2.11	1.67	398
45	Oskarsgrundet	0.94	0.46	0.06	0.46	-2.49	1.55	737
48	Rødvig	0.91	0.86	-0.14	0.85	-3.11	2.41	755
49	Klintholm	0.89	1.13	-0.06	1.13	-2.91	2.98	646
51	Arkona Becken	0.93	1.01	0.16	1.00	-2.08	1.99	94
56	Ölands Södra Grund	0.83	0.65	0.18	0.63	-1.08	1.91	74
57	Hävringe	0.91	0.78	0.28	0.73	-1.33	3.84	176
58	Landsort	0.84	1.13	0.83	0.77	-1.39	3.45	300
59	Gustav Dahlén	0.62	1.00	0.08	1.01	-1.77	3.56	43
60	Kopparstenarna	0.84	1.12	0.44	1.03	-1.93	3.19	268

Index	Station Name	ρ	RMSE	Bias	Std	Min	Max	N
61	Almagrundet	0.75	0.79	0.31	0.73	-1.47	1.85	47
62	Svenska Björn	0.71	0.94	0.63	0.71	-0.77	2.46	64
63	Bogskär	0.78	1.20	0.37	1.15	-2.63	3.51	222
71	Märket	0.78	1.30	0.71	1.10	-2.42	3.89	181
74	Sydostbrotten	0.69	1.27	-0.01	1.27	-4.38	4.11	532
75	Valassaaret	0.67	1.13	-0.30	1.09	-4.81	5.04	688
76	Snipan	0.74	1.38	-0.03	1.38	-4.22	3.84	300
77	Kemi	0.65	1.48	-0.48	1.40	-4.49	2.65	271
78	Krunnit	0.53	1.77	-0.78	1.60	-4.60	3.06	59

Table D.1: Comparison between the reconstructed monthly mean SST fields and the independent in situ records. Only the first name is displayed for stations that have switched names. The root mean square error (RMSE), bias, standard deviation (Std), minimum difference (Min) and maximum difference (Max) are given in °C. ρ represents the correlation coefficient and N represents the number of match-ups on which the comparisons were based.