Development of a new spatially complete and daily continuous lake surface water temperature dataset for Lake Vänern, Sweden

Anna Schultze

Anna Schultze (2023).
Development of a new spatially complete and daily continuous lake surface water temperature dataset for Lake Vänern, Sweden

Master degree thesis, 30 credits in *Physical Geography and Ecosystem Science*
Department of Physical Geography and Ecosystem Science, Lund University

Level: Master of Science (MSc)

Course duration: *January* 2023 until *June* 2023

Development of a new spatially complete and daily continuous lake surface water temperature dataset for Lake Vänern, Sweden

---

Anna Schultze

Master thesis, 30 credits, in *Physical Geography and Ecosystem Science**

Supervisor
Zheng Duan
Department of Physical Geography and Ecosystem Science

Exam committee:
Martin Berggren
Hani Younes

Acknowledgment

Abstract

Lakes are an important part of the world's ecosystems. The ecological state of lakes is threatened by rising temperatures that affect the biological, physical and chemical cycles. Therefore, it is essential to monitor lake surface water temperature (LSWT) and its spatiotemporal variabilities. Currently monitoring LSWT employs three primary approaches: in-situ measurements, satellite remote sensing, and reanalysis products/modelling. Each has its advantages and limitations. In-situ measurements offer accuracy at the point scale but suffer from inconsistencies and infrequent data collection. Satellite remote sensing provides relatively high spatial resolution but is affected by cloud cover and data gaps. Reanalysis products offer all-weather data but often at a coarse spatial resolution, limiting their ability to capture fine spatial scale variations in LSWT. This study aims to develop a new spatially complete and daily continuous LSWT by fusing satellite LSWT product and reanalysis product for Lake Vänern, the largest lake in Sweden. The reanalysis product ERA5-Land providing hourly lake temperature at the spatial resolution of 0.1° was used. Five existing satellite LSWT products were evaluated against in-situ measurements. The MODIS LSWT product was identified as the most suitable satellite product to be fused with ERA5-Land data using the Enhanced Spatial and Temporal Adaptive Reflectance Model (ESTARFM). A bias correction was conducted to account for systematic bias resulting from the data fusion. The bias-corrected fused LSWT dataset was evaluated against in-situ measurements and showed higher accuracy than the MODIS and ERA5-Land data with a mean absolute error of 1.57 °C, root mean square error of 2.04 °C and R^2 of 0.87. The spatial and temporal variations of the bias-corrected fused LSWT were in good agreement with the ERA5-Land and MODIS-derived LSWT, as well as with in-situ measurements. Finally, the bias-corrected fused LSWT product was used to investigate the spatial and temporal dynamics of Lake Vänern, revealing the development of a thermal bar and seasonal LSWT changes. This study demonstrated the good performance of the data fusion approach in generating a spatially complete and temporally continuous LSWT dataset. This approach is valuable for LSWT monitoring and further investigation of ecological changes in lakes associated with shifting LSWT.

I Table of Contents

II   List of Figures

III   List of Tables

IV   List of Abbreviations

LSM	Land Surface Model
(A)ATSR	(Advanced) Along Track Scanning Radiometer
AVHRR	Advanced Very High-Resolution Radiometers
CGLOPS	Copernicus Global Land Operations
EM	Electromagnetic
(E)STARFM	(Enhanced) Spatial and Temporal Adaptive Reflectance Fusion Model
ENVISAT	Environmental Satellite
ERS-2	European Remote-Sensing Satellite 2
FLake	Freshwater Lake
IFS	Integrated Forecast System
IR	Infrared
LST	Land Surface Temperature
LSWT	Lake Surface Water Temperatures
MODIS	Moderate Resolution Imaging Spectroradiometer
MetOp-A	Meteorological Operation
NIR	Near-infrared
SLSTR	Sea and Land Surface Temperature Radiometer
TIR(S)	Thermal Infrared Sensor
TOA	Top of the Atmosphere

1   Introduction

There are about a hundred million lakes on Earth, storing 87 % of the available freshwater, providing important ecosystem services and are crucial for biodiversity (Woolway et al., 2020). Climate change poses a significant threat to the ecological function of lakes due to elevated air temperatures, which subsequently affect water temperatures (Dörnhöfer & Oppelt, 2016). These temperature changes, in turn, alter various biological, physical, and chemical processes within the lakes (Reinart & Reinhold, 2008). The increasing temperatures lead to earlier algae blooming which increases the negative effects of eutrophication. Further enhanced temperatures lead to a shifted timing of stratification leading to an overturn of the lake (Woolway et al., 2020). The difference between lake surface water temperature (LSWT) and the atmospheric temperatures plays a crucial role in governing the exchange of energy and water between these two spheres, thereby directly influencing the energy and hydrological cycles. LSWT serves as a valuable indicator of climate change and represents one of the essential climate variables to monitor the magnitude of change on a regional scale (Reinart & Reinhold, 2008). Lakes offer a robust proxy for assessing the impact of climate change, owing to their extensive research history, global distribution, and their immediate responsiveness to climate and environmental changes (Adrian et al., 2009).

Therefore, it is important to consistently monitor the spatial and temporal dynamics of LSWT. The primary techniques employed for monitoring LSWT include in-situ measurements, satellite remote sensing, and reanalysis products, the outcome of Land Surface Models (LSM). In-situ measurements are point-based measurements that are usually taken in combination with other lake ecology parameters. While this method can yield highly accurate measurements, such pointed-based measurements are limited in their ability to adequately capture the spatial dynamics of LSWT due to the typically sparse network of measurement points. Moreover, the poor temporal sampling frequency (bi-weekly or irregular intervals) of point-based measurements often fails to adequately capture temporal variations of LSWT. These disadvantages make the in-situ measurements almost unsuitable to monitor the spatial and temporal dynamics of large lakes (Reinart & Reinhold, 2008).

The second method, satellite remote sensing, has been widely used to estimate LSWT. Previous studies have highlighted the utility of remote sensing in complementing the limited data coverage by leveraging the advantages offered by thermal sensors, as well as the enhanced spatial and temporal coverage. Satellite remote sensing enables the monitoring of LSWT changes on a global scale, with the generation of daily data. Notable satellite sensors, such as the Moderate Resolution Imaging Spectroradiometer

(MODIS) (Chavula et al., 2009; Reinart & Reinhold, 2008), Thermal Infrared Sensor (TIRS) of Landsat 8 (Dyba et al., 2022) and the (A)ATSR-series (Llewellyn-Jones et al., 2001; MacCallum & Merchant, 2010; Schneider et al., 2009; Zhang et al., 2021), have been used and evaluated to investigate the temporal and spatial variations of LSWT in many lakes. Nevertheless, satellite remote sensing for LSWT estimation is susceptible to cloud influences, which frequently lead to data gaps or inaccurate estimates (Long et al., 2020).

The third method involves LSM which produces reanalysis products, consisting of many land variables, which are able to simulate lake water temperatures on an hourly basis under all weather conditions (Muñoz-Sabater et al., 2021; Stefanidis et al., 2022). The state-of-the-art reanalysis product, ERA5-Land, is generated by the European Centre for Medium-Range Weather Forecasts (ECMWF) using their Integrated Forecast System (IFS). The IFS is a comprehensive numerical weather prediction system that includes an LSM component called the Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (H-TESSEL). The FLake model is included in the IFS, and it is a one-dimensional model that simulates the vertical temperature and water content profiles of a lake. To assure the accuracy of the model the location, the bathymetry and the preliminary condition of the lake are required. The Freshwater Lake model (FLake) is able to produce seven predictive variables to describe the state of the lake (Muñoz-Sabater et al., 2021).

ERA5-Land provides data on lake temperature at hourly and 0.1 °(9 km) spatial resolution on the global scale (ECMWF, 2023). The spatial resolution is coarse compared to satellite products (e.g., MODIS 1 km) (Long et al., 2020). Previous studies have demonstrated the accuracy of ERA5-Land products in estimating LSWT. Stefanidis et al. (2022) used the skin temperature, total layer temperature, mix layer temperature and mix layer depth from the ERA5-Land product to investigate the water temperature change of 51 lakes across Europe. Another study evaluated and compared the ERA5-Land lake mix layer temperature dataset with other satellite LSWT products (GloboLakes and CGLOPS) to assess their performance for 11 lakes in North America. The lake mix layer temperature expresses the average temperature at the top water layer of a lake (Zhang et al., 2021). Both aforementioned studies demonstrated that it is suitable to use lake mix layer temperature to represent LSWT. In contrast, this study will not solely focus on the evaluation and comparison of different datasets, nor will it limit its LSWT products to LSM data alone. Instead, it will expand upon previous work by combining the reanalysis product with satellite LSWT products.

Previous studies have successfully addressed the issue of cloud gaps in satellite products by merging multiple data sources and employing gap-filling methods (Duan et al., 2017; Zeng et al., 2018). For example, Duan et al. (2017) used passive microwave data to account for data gaps in MODIS due to cloud contamination. Passive microwave measurements have the advantage of penetrating through clouds, enabling temperature estimation even on cloudy days. However, passive microwave data typically have a lower spatial resolution (25 km) and relatively lower accuracy (RMSE: 3.5-4.4 °C) (Long et al., 2020). Zeng et al. (2018) filled the gaps of the satellite-derived Land Surface Temperature (LST) data using NDVI data, but this NDVI-based approach is not suitable for water surfaces. Additionally, there have been attempts to merge satellite data with reanalysis products. Long et al. (2020) used the Enhanced Spatial and Temporal Adaptive Reflectance Model (ESTARFM), which is a heritage of the spatial and temporal adaptive reflectance model (STARFM) developed by Feng et al. (2006), to combine MODIS LST data with the reanalysis product generated by the China Land Data Assimilation Model to investigate the land surface temperature dynamics. These studies demonstrated that the data fusion approach is able to merge two different products; However, all these studies focused on land surfaces rather than water bodies (Long et al., 2020).

Lake Vänern, with a surface area of 5648 km^2, is the largest lake in Sweden and the European Union. Currently, the limited number (only 3) of existing measurement stations has made it challenging to capture the spatial and temporal variations of Lake Vänern. Therefore, there is a need to obtain these variations by other methods. Satellite LSWT data offer the advantage of high spatial resolution, but they tend to be spatially and temporally incomplete due to cloud influences. On the other hand, the reanalysis product ERA5-Land can provide hourly and spatially complete LSWT data but with a very coarse spatial resolution. By combining the strengths of satellite and reanalysis products through a data fusion model like ESTARFM, it is possible to develop a solution for generating a spatially complete and daily continuous LSWT dataset. Therefore, this study aims to fuse satellite and ERA5-Land reanalysis products to leverage their advantages and create a spatially complete and daily continuous LSWT dataset for Lake Vänern. To achieve this, the accuracy of various LSWT products for Lake Vänern will be first investigated to determine the most suitable products for data fusion. There are several LSWT products available, however, first, the accuracy of these products for Lake Vänern needs to be investigated to further decide which product will be used in the data fusion. The study will access the accuracy of the MODIS, TIRS, Copernicus Global Land Operations (CGLOPS), GloboLakes and ACR-Lake LSWT products as well as the lake mixed layer temperature of the ERA5-Land product. To achieve the overall aim, three specific objectives are defined in this study:

(1) Evaluating the accuracy of multiple satellite-derived LSWT products and ERA5-Land reanalysis product against in-situ measurements.

(2) Developing and evaluating a data-fusion approach for generating a merged LSWT dataset.

(3) Analyzing the spatial and temporal variability of LSWT in Lake Vänern using the generated LSWT dataset.

This thesis will answer the following research questions:

(1) Do satellite-derived LSWT products and the ERA5-Land product accurately represent the LSWT of Lake Vänern when compared to previous studies?

(2) Can the proposed data-fusion method fill data gaps in satellite-derived LSWT products and enhance the accuracy of LSWT estimates compared to the already existing LSWT products?

(3) Does the generated LSWT product show a similar temporal and spatial pattern compared to previous studies?

2   Theoretical background

This chapter provides a theoretical background for the thesis, introducing fundamental concepts related to remote sensing. Furthermore, it explains how lake surface temperatures can be estimated from satellite sensors and the reanalysis product ERA5-Land and finally, describes the concepts of the data fusion approach.

2.1   Satellite remote sensing of lake surface water temperature

Remote sensing is the process of observing the Earth's surface from a distance using instruments such as satellite sensors. This technique is based on physical theories that state that every object with a temperature above absolute zero (0 Kelvin) reflects and emits energy in every wavelength. Satellite sensors measure the reflectance of the Sun's light or the emission of electromagnetic (EM) radiation from objects on Earth, allowing for the identification and detection of properties of objects on a global scale (Read & Torrado, 2009).

2.1.1 Electromagnetic spectrum

EM radiation is energy transported in the form of photons with the properties of a wave traveling through space and time. It is swinging in all directions perpendicular to its travel direction (Emery & Camps, 2017). EM radiation can be characterized by two measures: wavelength and frequency. A wavelength describes the distance between the peaks of two waves and is expressed in units of length (meters, micrometres (μm) or nanometres (nm)). The frequency is defined as the number of cycles per one-second

period and has its units in hertz (Hz), megahertz (MHz) or gigahertz (GHz) (Martin, 2014). Wavelength and frequency are inversed to each other, meaning that a higher frequency results in shorter wavelengths (Emery & Camps, 2017). The transported energy (E) can be described in following equation:

$$E = f \times h$$

where f is the frequency and h represents the Planck constant (6.62607×10^{-34} J·s). EM radiation is commonly described as a flux of energy or power, in either joule per second or Watts (W), radiation incident on or emitted by an object (Martin, 2014).

The EM radiation spreads over a large spectrum of wavelengths and frequencies and can be split into different ranges along the EM spectrum (Figure 1). The most known spectrum is the visible range between 0.4 to 0.7 μm (10^{-7}). Since it is the range in which the human eye can detect EM radiation and therefore contains a lot of essential information for the human. The satellite sensor, however, captures a wider part of the EM spectrum, which makes remote sensing an important asset to humanity. It detects shorter wavelengths known as gamma, X-ray and ultraviolet but also longer wavelengths in the near-infrared (NIR) (10^{-7}- 10^{-5}), thermal infrared (TIR) (10^{-5} to 10^{-3} m) up to microwave (1 mm to 1 m) range. The radiation in the shorter wavelength up to the NIR is mainly reflected radiation from the sun, while the mid and longer infrared wavelengths contain wavelengths emitted by an object making them important for the determination of thermal properties (Emery & Camps, 2017).
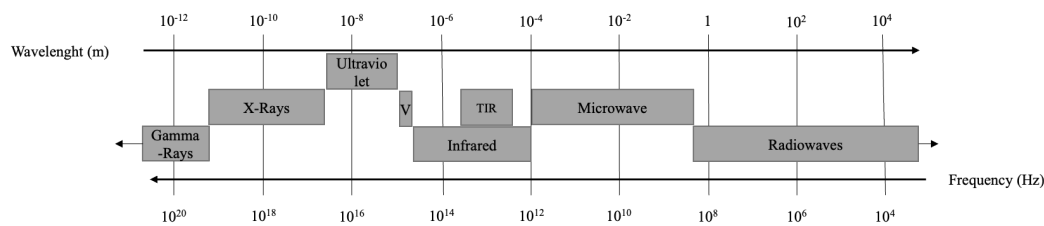
Figure 1. Electromagnetic spectrum. Modified from Emery & Camps, 2017; Martin et al. 2014.

The radiation received or emitted by a surface can be expressed as the radiant flux density ($\Delta\Phi$), which is measured in W/m^2. The received radiation at the surface coming from an external source is called irradiance, while the emitted radiance from a surface is defined as the exitance. The intensity of radiation is the radiant flux density for a solid angle, which can be described as the flux of radiation traveling from a point source through a cone in a certain angle onto a surface measured in W per steradians (Emery & Camps, 2017). The intensity can be further defined as radiance (L), which can be described in the following equation:

$$L = \frac{\Delta^2 \Phi}{\Delta \Omega \Delta A cos\theta} \qquad\qquad (2)$$

Radiance describes the radiance flux ($\Delta\Phi$) emitted by a surface (ΔA) for a solid angle ($\Delta\Omega$) and certain incline (θ) for a given direction. Radiance represents the measure that is detected by satellite sensors in W m^{-2} sr^{-1} (Martin, 2014).

The radiance from each of the spectra can be used to observe and identify different properties of objects. The variation of the reflected, absorbed or transmitted energy of an object by wavelength can be described as the spectral signature (Emery & Camps, 2017). However, the energy also varies with different angles of incidence and is influenced by atmospheric conditions. Therefore, objects can rather be distinguished based on their spectral response patterns than their spectral signature (Read & Torrado, 2009).

2.1.2   Estimation of lake surface water temperature using remote sensing

As stated before, the longer infrared wavelengths can be used to determine the thermal properties of objects. This is done using the TIR channels of a sensor. Either way, the measured radiance at the sensor can only be converted to the top of the atmosphere temperature (TOA) and that is why it is necessary to determine the ground emissivity which can be achieved by the normalized difference vegetation index, also known as NDVI. The ground emissivity and the top of the atmosphere temperature can be used to calculate the land surface temperature (Avdan & Jovanovska, 2016). Furthermore, a water/land mask as well as land cover information is used to distinguish between land and water. The previous steps are conducted using algorithms. The most common algorithms to generate LST products, which include LSWT data, are the single-channel, the split-window/dual-window, the temperature and emissivity separation as well as the physics-based day/night algorithm. The single-channel and the split-window/dual-window require the land surface emissivity to estimate, but the processing can be considered simple (Li et al., 2023). Sensors such as TIRS and MODIS use the single-channel (TIRS) and the split-window/dual-window (MODIS) algorithms to generate LST products (Li et al., 2023; Sayler, 2023; Wan, 2006). The other two algorithms are able to generate the LST without knowing the emissivity, however, they also come with their disadvantages such as complexity and their dependency on the accuracy of the atmospheric correction and geometric registration.

The ASTER sensor onboard of Terra platform uses temperature and emissivity separation algorithm to generate its LST product. The physics-based day/night algorithm is used to generate the 8-days, 16-days and monthly 6 km spatial resolution level 3 LST products of MODIS (Li et al., 2023). Not only the LST can be estimated using the thermal bands of the sensors, further the LSWT can be determined by it. ARC-Lake is

a LSWT dataset, which was generated using the ATSR series. This was done by combining the three infrared channels of the ATSR-1, ATSR2 and AATSR. The authors used a simplified version of the inverse problem, which was originally used to determine sea surface temperature. The LSWT was estimated using the optimal estimation method. The method combines observations of the lake's surface water temperature with information about the predicted state of the atmosphere in terms of total column water vapor to account for the influence of clouds (MacCallum & Merchant, 2012). The microwave spectrum is used to distinguish between vegetation, snow as well as surface and ocean roughness (Emery & Camps, 2017). The radio wavelengths at about 1 m are mostly used for communication applications and are therefore not relevant for remote sensing (Martin, 2014).

2.1.3 Satellite sensors

Satellites are equipped with different sensors which detect the reflected and emitted EM radiance by objects for different wavelengths. This is called multispectral remote sensing (Emery & Camps, 2017), where the sensor measures simultaneously different ranges within the EM spectrum to generate bands for each wavelength (Read & Torrado, 2009). The most used wavelengths are in the visible and infrared spectrum. Nevertheless, atmospheric conditions such as clouds, fog and aerosols are disturbances in these wavelengths that affect the accuracy of these measurements (Martin, 2014). The sensor stores the received signal as digital numbers which can be used to analyse the radiance numerically or visualise the variation of the photon's intensity with different wavelengths (Emery & Camps, 2017).

Over the past decades, numerous satellite missions have been launched, each characterized by unique technical properties. The main differences among these missions lie in their temporal, spatial, spectral, and radiometric resolutions. The different resolutions determine various aspects, such as their revisit time, the size of the smallest detectable object, the number of bands as well as the number of brightness pixels (Read & Torrado, 2009). One notable satellite mission is the ERS-2 mission, which was launched in 1995 and ended in 2011. The mission featured a platform equipped with seven sensors, each serving various applications (ESA, n.d.-b). The primary goals of the mission were to study sea surface temperatures, winds, and atmospheric ozone. One of the onboard sensors was the ATRS-2 which had 7 bands ranging between the visible and the thermal spectrum. The spatial resolution was 30 m with a revisit time of 35 days (EoPortal, 2012). The thermal bands of the ATSR-2 sensors were utilized to contribute to the production of the GloboLakes LSWT product (MacCallum & Merchant, 2012).

The Envisat mission was launched in 2002 and extended the ERS mission with more instruments, however, the mission ended in 2012. The platform carried 10 sensors such as the AATSR. The objective of the mission was to investigate atmospheric chemistry and ocean studies (ESA, n.d.-a). The AATSR sensor was a heritage of the ATSR2 sensor. The sensor was designed to measure sea surface temperature using its seven bands ranging between the visible and TIR part of the spectrum with a spatial resolution of 1 km (Llewellyn-Jones et al., 2001). The AATSR sensor contributed, like ATSR2, to the GloboLakes LSWT product (MacCallum & Merchant, 2012).

The Terra/Aqua satellite is a mission that has operated since February 2000 up until today. The platform hosts several sensors such as MODIS, the Advanced Spaceborne Thermal Emission and Reflection Radiometer, Clouds and Earth Radiant Energy System, Multi-angle Imaging SpectroRadiometer and Measurement of Pollution in the Troposphere. The satellite operation is observing the atmosphere, land surface and oceans as well as snow, ice and energy budget (NASA, 2023). MODIS has a high temporal resolution with a revisit time of at least once a day, it also covers 36 spectral bands which range between 250 and 1000 m spatial resolution (Read & Torrado, 2009). Since MODIS has this high spatial-temporal and spectral resolution it is very useful for many applications, which vary between cloud detection, estimation of LST up to photosynthetic activity (NASA, 2023). The LST product of MODIS is covering water bodies as well, which enables the monitoring of LSWT.

Another satellite mission launched in early 2000 was the MetOpA satellite, which started in 2006 and ended in 2021 (Eumetsat, n.d.). The satellite carried, among others, the Advanced Very High-Resolution Radiometer (AVHRR) sensor which has a temporal resolution of a day, and measures radiation with five bands with a spatial resolution of 1.1 km. Originally, the sensor was built to globally monitor clouds, land surface temperature and vegetation, but, it was also used to observe fires, volcanic activity, radiation, snow and ice (Xiong et al., 2018). The AVHRR sensor was also utilized to contribute to the GloboLakes dataset (MacCallum & Merchant, 2012)

Relative new missions are the Landsat 8 and Landsat 9 satellite. The Landsat 8 has been operating since February 2013 and is continuing. The platform hosts two sensors the operational land image and TIRS with a temporal resolution of 16 days. The former covers nine spectral bands with eight of them with a spatial resolution of 30 m and one with 15 m. The latter has a temporal resolution of 16 days and two bands in the thermal infrared regime with a spatial resolution of 100 m. The thermal sensors of Landsat 8 can be used to estimate lake temperatures. The Band 10 of the collection 2 level 2 data represents the computed surface temperature, which also includes LSWT (Ihlen, 2019).

Other relatively new satellite operations are the Sentinel-3A satellite, which was launched in 2016, and its twin satellite Sentinel-3 B in 2018. On the Sentinel-3A four sensors are installed, namely Ocean and Land Colour Instrument, the Sea and Land Surface Temperature Radiometer (SLSTR), Synthetic Aperture Radar Altimeter and the Microwave Radiometer. The sensors serve different purposes such as monitoring sea level and sea surface temperature, land cover mapping, vegetation health and others. The sensors vary in their temporal, spatial and spectral resolution. The SLSTR sensor has a temporal resolution of 27 days and covers 9 spectral bands with a spatial resolution of 500 m for the visible range and 1 km for the thermal infrared channels. The Ocean and Land Colour Instrument has the same temporal resolution as the SLSTR, however, it has 21 spectral bands and a higher spatial resolution of 300 m (ESA, n.d.-c). The CGLOPS LSWT product is an outcome of the SLSTR sensor, where the SLSTR sensor is used to generate LSWT estimations since 2016 (Carrea & Merchant, 2020b).

2.2   ERA5-Land reanalysis product for lakes

The land surface is only one part of the earth's complex system but since its processes affect society and ecosystems it is important to model the changes of these processes due to climate change to make projections of how society will be affected. This can be done through LSM, which are models that offer the ability to solve complex interactions between the land surface and the atmosphere which includes water, energy, and carbon fluxes, through a numerical approach. They also consider direct and indirect anthropogenic impacts as well as ecological dynamics (Fisher & Koven, 2020). The original purpose of LSM was to set physical boundary conditions to model the impact of the land surface on the atmosphere. LSMs have since then further developed and cover nowadays a range of land surface processes such as surface hydrology, soil moisture dynamics and land surface heterogeneity (Pal & Sharma, 2021).

A commonly used product generated by the H-Tessel LSM is the ERA5-Land reanalysis product, which is able to generate various variables that represent the state of land components. ERA5-Land is based on the ERA5 and takes advantage of its atmospheric forcing. The spatial resolution of ERA5-Land (9 km) is higher compared to ERA5 (31 km). ERA5-Land is based on the H-Tessel model using version CY41R2 which is provided by the ECMWF IFS. The ERA5-Land data covers a 31 km horizontal grid and divides the atmosphere into 137 layers between the Earth's surface and an 80 km height which makes it a finer and more frequent product compared to previous products (Muñoz-Sabater et al., 2021). It is based on a single simulation, which improves the computing time of the model. The model's purpose is to simulate the surface exchanges including the land surface hydrology (ECMWF, 2023). The ERA5-Land product

includes the FLake. It is a one-dimensional model which works with the assumption that the temperature profile of a lake has a specific shape. The lake is separated into two layers, the uppermost (mixed) and the bottom (thermocline) level. The model provides several variables to represent the state of a lake: mix layer temperature which represents the average uppermost layer's temperature of lakes, the mixed-layer depth describing the thickness of the top layer, the bottom temperature, the mean temperature of the water column as well as the surface temperature of the upper ice as well as its ice thickness (Muñoz-Sabater et al., 2021; Zhang et al., 2021). ERA5-Land is able to generate spatially complete datasets of LSWT (Zhang et al., 2021) with high accuracy (Muñoz-Sabater et al., 2021; Stefanidis et al., 2022), as well as its higher spatial resolution compared to other reanalysis products, thus the ERA5-Land data will be further used in this study.

2.3   Data fusion concepts

Data fusion can be described as the approach of merging different data sources together to gain an improved dataset (Zhang, 2010), which either supplements the incomplete data of one sensor with data from other sensors or improves the estimation of a dataset by combining different data sources (Schmitt & Zhu, 2016). Schmitt and Zhu (2016) divided the methods of data fusion into four steps: data alignment, data correlation, attribute estimation and identify assessment. The first two steps validate that the data sources have a relationship with each other but also with the research object. The data alignment makes sure that the transformation, the units, and the spatial and temporal resolution of the object of interest correspond with each other between each input variable to gain a universal representation. The last two steps can be described as data fusion, which is the combination of aligned and correlated data. The aim of the steps is the optimization of the object of interest by fusing the information from different sources in an accurate framework (Schmitt & Zhu, 2016).

In the last decades, a range of various data fusion methods have been developed to advance the spatiotemporal resolution of datasets such as the Spatial and Temporal Adaptative Reflectance Model (STARFM) (Feng et al., 2006) or the flexible spatiotemporal data fusion method (Zhu et al., 2016). The approaches aimed to fuse satellite products varying in their spatial and temporal resolution to synthesize their advantages. This was achieved by blending sources with a high spatial but low temporal resolution (e.g., Landsat) with low spatial but high temporal resolution data (e.g., MODIS). In STARFM a set of fine and coarse-resolution images, acquired at the same time, as well as one coarse image at a prediction time is needed to forecast the pixel value of the fine-resolution image at the prediction time. The model searches for similar pixels within a moving window using fine-resolution data. Furthermore, the quality of the pixels needs

to be ensured to then weight the pixels based on their spectral, temporal and spatial distance using the fine and coarse data. The previous step is then used to predict the fine-resolution pixel (Feng et al., 2006).

In terms of data fusion for temperature retrieval purposes, previous studies have used data fusion models such as the ESTARFM (Long et al., 2020), a heritage of STARFM, multiresolution Kalman filter algorithm (Dong et al., 2022) to combine satellite and LSM data to fill gaps in the satellite data, conquer the scale difference between the sources, gain high spatial and temporal resolution and improve their accuracy. Dong et al. (2022) were able to generate a spatially complete and hourly LST available dataset that followed the spatial pattern and diurnal variation of LST well. The algorithm is based on the Kalman filter, but instead of treating the temporal dimension, it addresses the scale differences between different data sources. The multiresolution Kalman filters from the fine to the coarse resolution but also smoothens from the coarse to the fine resolution, which leads to spatial complete data (Dong et al., 2022). The validation of the proposed method reported an RMSE of 3.08-3.38 °C (Dong et al., 2022). Other methods such as the ESTARFM data fusion approach, used by Long et al. (2020), indicated a similar accuracy with an RMSE ranging between 2.77 and 3.96 °C. ESTARFM was developed to combine data sources with different spatial and temporal resolutions to make use of their advantages, to finally, gain a high spatial and temporal resolution dataset. This was achieved by integrating a linear spectral mixing model, consideration of distances between pixels and their similarity as well as their temporal correlation (Long et al., 2020). The different data fusion methods have shown that the combination of multisource data is able to fill gaps in satellite data, gain a high spatial and temporal resolution and finally, is able to keep a high accuracy.

3   Study area

The study focuses on Lake Vänern in Sweden, which is the largest lake in the European Union (58-60°N 12-14 °E) (Figure 2). The area of interest covers 5648 km^2 and is influenced by a maritime on the border to Central European climate (Reinart & Reinhold, 2008). Lake Vänern hosts a variety of fish communities and is an important freshwater supplier for the region. The lake's temperature varies spatially, due to the large size of Lake Vänern. During the winter months, Lake Vänern rarely freezes, only in areas close to the shore (Kvarnäs, 2001). The bathymetry of Lake Vänern is characterized by two basins which are located on the West and East of the lake. The maximum depth is about 100 m which is reached in the southwest as well as in the centre of the eastern part of the lake (Figure 2).
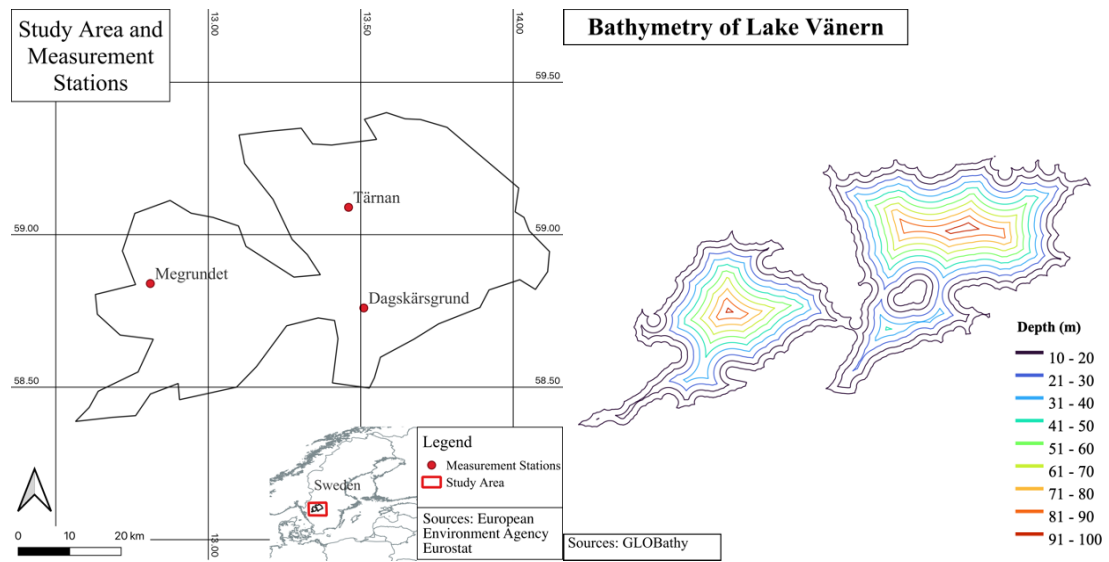
Figure 2. Location and three in-situ measuring stations and bathymetry of Lake Vänern based on data from European Environment Agency, Eurostat and GLOBathy.

4　　Data and methodology

This study utilized data from multiple satellite sensors, resulting in five satellite LSWT products, as well as in-situ measurements and ERA5-Land reanalysis output. The used datasets and their spatial resolution, temporal length and data source can be retrieved from Table 1.

Table 1. Summary of in-situ measurements, satellite LSWT products and ERA5-Land reanalysis product along with their data source, spatial and temporal properties.

Satellite mission	Source/ Sensor	Dataset	Spatial/ temporal resolu- tions	Temporal length	Data Source
/	In-Situ	/	Point / ~ 10/year	1973 - 2022	miljo- data.slu.se
/	ERA5-Land	Lake Mix- Layer Temperature	0.1 ° (9 km) / hourly	1950 – today	cds.cli- mate.coperni- cus.eu
Terra/ Aqua	MODIS	MOD11A1	1 km / daily	2000 – today	earthdata.nasa .gov
Landsat 8	TIRS	Band 10	100 m /16 days	2013 – today	usgs.gov
ERS-2 / ENVISAT	ATSR2 / AATSR	ARC-Lake	0.05°/ daily average	1995 – 2012	laketemp.net
ERS-2 / ENVISAT / MetOpA	ATSR2 / AATSR / AVHRR	GloboLakes	0.025° / daily average	1995 – 2016	cata- logue.ceda.ac. uk
ENVISAT / Sentinel-3	AATSR / SLSTR	CGLOPS	0.0083 (1 km) / 10 days average	2002-2012 2016- today	land.coperni- cus.vgt.vito.b e

4.1　　In-situ measurements

The in-situ measurements utilized in this study were collected from three observation stations located in Lake Vänern: Tärnan, Megrundet and Dagskärsgrund (Figure 2). In-situ measurements from those three stations for Lake Vänern are available from 1973 to the present (Table 1), the measurements were taken at different depths between April

and October, but this study focuses on LSWT and thus we only used measurements within the first 0.5 m depth. The dataset shows variation in the number of measurements, for some years with a minimum of five measurements in 2011 and a maximum of ten measurements in 2004, 2006 and 2019. The complete dataset spanning the entire temporal length was utilized in this study to evaluate the performance of satellite and reanalysis products.

4.2   Satellite LSWT products

A total of five satellite LSWT products were evaluated in this study, each providing coverage over different spatial extents and temporal lengths. The specific properties of these products are summarized in Table 1.

4.2.1   MODIS

The land surface temperature/emissivity Level 3 product (MOD11A1) of Collection 6.1 was utilized in this study to evaluate the accuracy of the product compared to the in-situ measurement. The LST product is available daily with a spatial resolution of 1 km in a 1,200 by 1,200 km grid. MODIS passes Lake Vänern approximately at 9:30 am (Reinart & Reinhold, 2008) and therefore, the MODIS daytime LST product was chosen. The product is a Level 3 product and was obtained from the MOD11_L2 swath product (Wan et al., 2021). The evaluation of the product was conducted through the comparison with in-situ measurements on clear days following the common practices (Duan et al., 2019; Wan et al., 2004).

4.2.2   TIRS

The thermal infrared sensor (TIRS) installed on Landsat 8 obtains the surface temperature of the Earth. The used surface temperature data was derived from the Landsat 8 Collection 2 Level 2 data which was created with a single channel algorithm from band 10 of the Level 1 data using TOA reflectance, TOA brightness temperature, advanced spaceborne thermal emission and reflection radiometer global emissivity dataset data and normalized differenced vegetation index, atmospheric profiles of geopotential altitudes, humidity data, reanalysis data of and air temperature (Sayler, 2023). The validation of the Landsat LST product over water surfaces was conducted at nine sites at the border between Canada and the USA. The validation results reported RMSE values of 1.1 °C, 0.9 °C and 0.9-1.3 °C for Lake Eri, Lake Superior and Lake Michigan, respectively (Duan et al., 2021).

4.2.3   GloboLakes

The GloboLakes dataset contains daily estimates of LSWT with a 0.05 °C spatial resolution for 1000 lakes between 1995 and 2016. The estimates were derived from multiple sensors namely the AVHRR, AATSR and the ATSR2 using the optimal estimation

14

algorithm. The satellite data was harmonized with each other to gain a finer resolution than its predecessor ARC-Lake and to extend its temporal length as well as the number of processed lakes (Carrea & Merchant, 2019).

4.2.4 CGLOPS

CGLOPS is a 10-day averaged LSWT product that provides similar to GloboLakes 1000 lakes in a spatial resolution of 1 km. The data can be split into three types: historical, reprocessed and near-real-time. The historical data were obtained from the AATSR sensor, covering the period from 2002 and 2012. The reprocessed and near real-time data were obtained by the SLSTR-A and SLSTR-B sensors operating on Sentinel-3. Specifically, the LSWT data were obtained from SLSTR-A for the period between April 2018 and August 2020. However, the SLSTR-A data obtained between June 2016 and April 2018 represented reprocessed data due to technical issues. From August 2020 onwards, the SLSTR-A data was combined with the tandem sensor SLSTR-B to enhance the quality of the LSWT product. The LSWT product is based on Level-3 data, which have been divided into intervals corresponding to specific time periods within a month. These intervals cover the 1st to the 10th, 11th to the 20th, and the 21st until the end of the corresponding month (Carrea & Merchant, 2020b). The evaluation of the product was conducted through the comparison with in-situ measurements and quality control during and post-processing. Additionally, the reprocessed data was evaluated against the near real-time data. The combined LSWT obtained from SLSTR-A and B was compared against the measurements of SLSTR-A to show the improvement of the product by incorporating SLSTR-B (Carrea & Merchant, 2020a). The validation of CGLOPS reported a mean bias of < -0.469 °C for AATSR derived and < -0.375 °C for SLSTR-A LSWT with a quality level above three. The 10-day average product validation for Lake Vänern indicated a high accuracy with a mean bias of -0.139 °C for AATSR and -0.104 °C for SLSTR-A (Carrea & Merchant, 2020a).

4.2.5 ARC-Lake

The ARC product contained two LSWT datasets derived from the ATSR2 and the AATSR sensors, installed on the ERS-2 and ENVISAT satellites, respectively. From the ARC dataset, two LSWT sets were retrieved the datasets obtained by the ATSR2 and AATSR satellites (Table 1) (Merchant & Maccallum, 2018a). The temporal coverage of the product was from 1995 until 2012, covering approximately 250 lakes with a spatial resolution of 0.05 °. For each sensor, a day and a night dataset were provided. The used product contained unaveraged per lake estimations split into observations or reconstructions on a daily timestep. For this study, only the daytime observational data was used. The author created the dataset by first applying a Bayesian approach to detect clouds, next the optimal estimation algorithm was used to retrieve the LSWT (Merchant

& Maccallum, 2018a). The dataset was validated by comparing the LSWT estimation with 54 in-situ measurements across 18 lakes with a mean bias of -0.34 °C (MacCallum et al., 2013)

4.3   ERA5-Land reanalysis product

This study used the ERA5-Land product which is obtained from the ECMWF. The data is an hourly reanalysis product covering the time range from 1950 up to today. The reanalysis product was chosen because of its high temporal resolution and its relatively high spatial resolution. The ERA5-Land data has been validated against in-situ measurements of 51 lakes across Europe and North America (Stefanidis et al., 2022; Zhang et al., 2021). In this study, the lake mix layer temperature dataset was used as a representation of the LSWT. Previous studies have demonstrated the dataset's capability to accurately represent LSWT, reinforcing its suitability for use in this study (Zhang et al., 2021). Zhang et al. (2021) have pointed out that the in-situ measurements are taken at a different depth compared to satellite estimation which represents the skin temperature of the lake, while the mix layer temperature dataset of ERA5-Land represents the average temperature of the lake's top layer (Zhang et al., 2021). The data was retrieved daily for the time period between 1973 and 2022 to cover the temporal length of the in-situ measurements. The spatial resolution of the ERA5-Land product is 9 km (Table 1). To ensure consistency and comparability with the respective LSWT products, the ERA5-Land data underwent a resampling process. The spatial resolution was adjusted to match the size of each specific LSWT product, ensuring that the datasets are aligned and can be effectively compared and analysed in the study.

4.4   Data fusion and bias correction

4.4.1   Overall data fusion framework

Figure 3 shows the overall framework used in the study. The study followed the methodology outlined by Long et al. (2020), which was supplemented with an additional step to compare and evaluate the accuracy of the satellite and reanalysis products against the in-situ measurements. After the comparison, the satellite products were split into cloudy and cloud-free days. The cloud-free days were used to feed into the ES-TARFM model by Zhu et al. (2010), which fused the coarse spatial resolution ERA5-Land data with the fine resolution satellite data to generate a spatially complete and daily available LSWT dataset. The detailed steps of the model will be further explained in Chapter 4.4.2. Subsequently, the newly generated dataset underwent bias correction using the cloud-free pixels on cloudy days through linear and variance scaling. Finally, the bias-corrected dataset was evaluated using the in-situ measurements. The fusion model ESTARFM and the corresponding bias correction were selected based on the

16

findings of Long et al. (2020), which demonstrated their ability to generate a highly accurate LST dataset with a high spatial and temporal resolution. Furthermore, the proposed method eliminated the need for additional data reconstruction, unlike other approaches such as gap-filling methods.
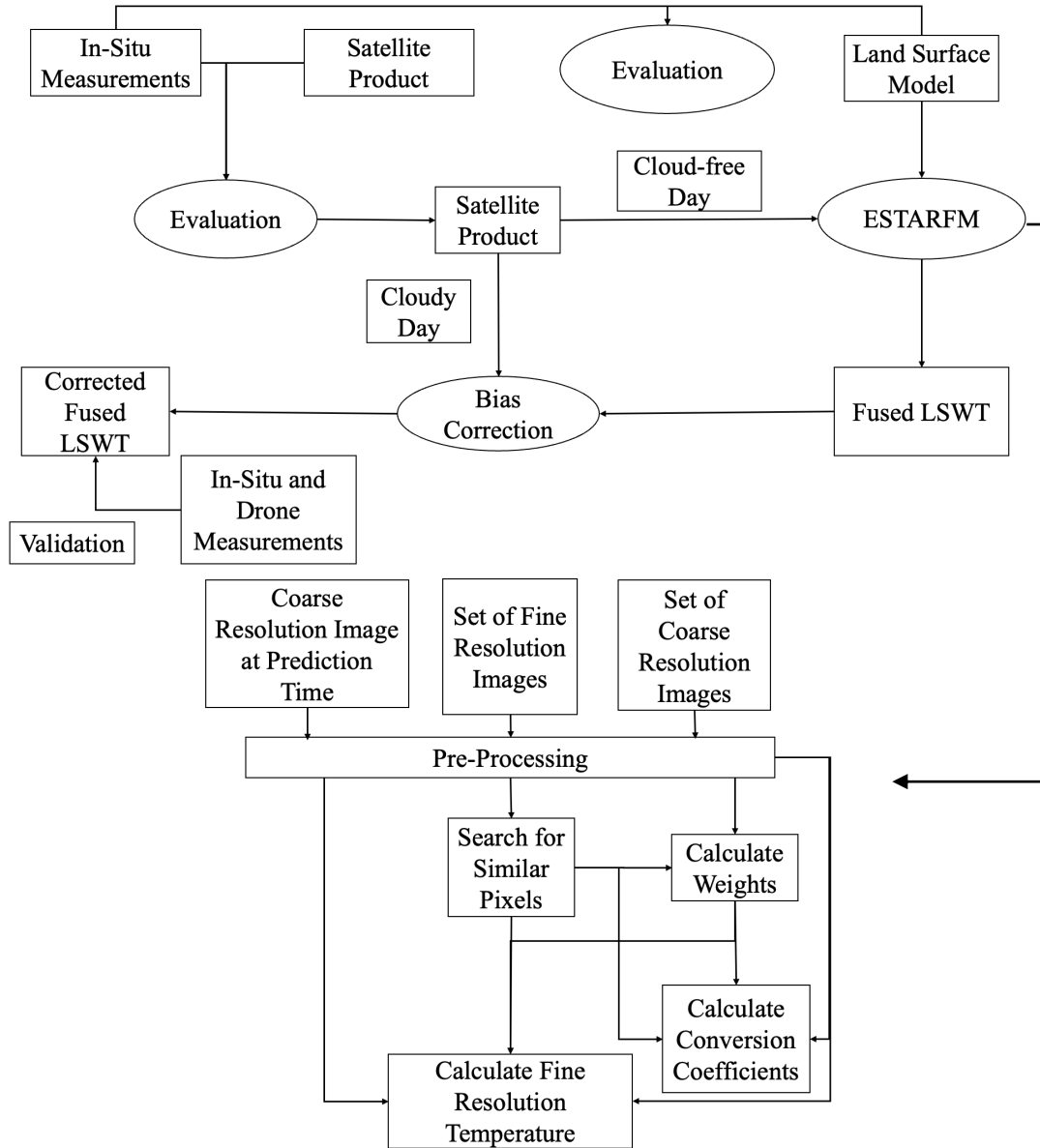
Figure 3. Overall data fusion framework. Modified from Long et al., 2020 & Zhu et al., 2010.

4.4.2 Pre-processing for data fusion and theoretical introduction of ESTARFM

The used data fusion model in this study is the ESTARFM by Zhu et al. (2010). For the implementation of the model, it was necessary to first resample the ERA5-Land data to the resolution of the satellite-derived LSWT product. This was done using the bilinear resampling method. Further, the satellite LSWT product was split into two subsets based on if the day corresponds to a clear or cloudy day. This chapter will discuss the data fusion method used, as well as the bias correction applied to the fused LSWT product. The bias correction was chosen to correct a systematic error in the fusion model.

17

Long et al. (2020) were able to significantly reduce the systematic error by following the steps outlined in the proposed bias correction (Chapter 4.4.2.3).

4.4.2.1   Division of datasets

The first step in implementing the model involved dividing the data from each satellite product into two groups: cloud-free days and cloudy days. This was done by examining specific layers within the product or quality flags associated with each satellite product. The approach by Long et al. (2020) was followed for MODIS, where a pixel with a quality flag of "cloud" and a "LST error > 3 K" was defined as a cloudy pixel. Additionally, days with pixels that did not have a corresponding LSWT value were also classified as cloudy days. For the ARC-Lake product, the division was based on the number of cloudy pixels present on each day. If the number of cloudy pixels exceeded zero, the day was considered as cloudy. According to the CGLOPS documentation, the effect of clouds had already been considered and a cloud detection algorithm was used to create the product. Therefore, days with pixels that contained no data or poor-quality levels were defined as cloudy days. For the GloboLakes product, the division was based on the quality of the data. Days with pixels lower than the quality flag of 4 were defined as cloudy days. For TIRS the pixel quality assessment band provided cloud confidence data, which was utilized to identify cloudy days.

4.4.2.2   Theoretical basis of ESTARFM

This study followed the data fusion approach ESTARFM proposed by Zhu et al. (2010). The ESTARFM is an advancement of the STARFM model developed by Feng et al. (2006). The STARFM has demonstrated its effectiveness in generating satellite-sensed data at a high spatial and temporal resolution. However, it does have limitations, particularly in heterogeneous landscapes. These limitations prompted the development of ESTARFM, which leverages the correlation between high and low spatial resolution data to obtain a high-resolution and temporally continuous dataset while also reducing systematic biases. Originally, the ESTARFM was developed to generate daily surface reflectance from MODIS and Landsat imagery. Long et al. (2020) was the first study that demonstrated the use of the ESTARFM for fusing LST.

The ESTARFM model requires a minimum of two pairs of coarse (e.g., ERA5-Land) and fine (e.g., MODIS) images acquired on the same day, as well as a coarse image for the desired prediction day. The underlying assumption of the model is that data obtained from different products for the same region are correlated and comparable, with minimal systematic bias, as long as they are acquired simultaneously and appropriately pre-processed. The pre-processing steps involve radiometric calibration, atmospheric correction, geometric rectification, and resampling of the coarse data to match the spatial

resolution of the fine data source. The model can be divided into four main steps (Figure 3). Firstly, the two fine-resolution images (F) are utilized to identify pixels within a moving window that exhibit similar characteristics. Subsequently, weights are calculated for these similar pixels, and conversion coefficients are computed through linear regression. Finally, the conversion coefficients and weights are applied to the coarse image (C) to generate temperature estimations for the fine-resolution image on the prediction date (Zhu et al., 2010). The model is based on a linear spectral mixing model, which can be expressed as followed:

$$F_k(x_{w/2}, y_{w/2}, t_p) = F(x_{w/2}, y_{w/2}, t_k) + \sum_{i=1}^{N} W_i V_i (C(x_i, y_i, t_p) - C(x_i, y_i, t_k))$$

where F and C correspond to the fine and coarse spatial resolution LSWT, respectively; the size of the moving window is expressed as w with the central pixel at the location $x_{w/2}, y_{w/2}$; t_k represents the time of the acquired high-resolution LSWT (k = m, n) while t_p is the prediction time; the number of pixels is expressed by N; the weight and the conversion coefficient of the similar pixel i can be described as W_i and V_i, respectively.

The calculation of the high-resolution LSWT at the prediction time is based on the high-resolution LSWT acquired at a known time and the weights and conversion coefficients of similar pixels found in the coarse resolution LSWT. The weight of a similar pixel is determined by its similarity between the fine and coarse resolution image and its distance to the center of the moving window. A shorter distance and higher similarity lead to a higher weight of the similar pixel. The similarity is calculated by the correlation coefficient between the similar pixel and its corresponding coarse-resolution pixel. The similarity is based on the correlation coefficient of the similar pixel and the corresponding coarse-resolution pixel. To improve the accuracy of the calculations of the high-resolution temperature product a temporal weight was applied based on the distance to the prediction date (Zhu et al., 2010). The final high-resolution LSWT can be calculated using the equation below:

$$F(x_{w/2}, y_{w/2}, t_p) = T_m * F_m(x_{w/2}, y_{w/2}, t_p) + T_n * F_n(x_{w/2}, y_{w/2}, t_p)$$

where the variables used in the previous equations were supplemented with T_k as the temporal weight, which indicates the magnitude of change between the coarse resolution LSWT at time t_k and the prediction time t_p.

4.4.2. Bias correction

The ESTARFM model reduces the systematic biases of the used data, however, the bias might not be completely eliminated. Therefore, an additional bias correction approach was used by Long et al. (2020) to minimize the bias of the final fused results. We tested

this additional bias correction approach in our study. For this approach, the systematic biases were eliminated by using clear pixels of cloudy days from the fine-resolution LSWT product. The first step involved employing a linear scaling approach, which calculated the difference of the mean of the fused LSWT of clear pixels on clear days (T_D^a) and the LSWT of clear pixels on cloudy days (T_M^a) using this equation:

$$bias = \ \mu(T_D^a) - \mu(T_M^a)$$

The calculated bias was then subtracted from the fused LSWT values before the linear scaling to obtain a linear-corrected fused LSWT ($T_D{}^b$).

$$T_D^b = \ T_D^a - bias$$

In addition to the linear scaling approach, a variance scaling approach was implemented. This was done by normalizing $T_D{}^b$ by subtracting its mean ($\mu T_D{}^b$) to shift the mean to zero. The normalized LSWT was expressed as $T_D{}^c$. Furthermore, the standard deviation of the difference was scaled using the ratio of the standard deviation of $T_M{}^a$ and $T_D{}^c$ and was finally summed up with the mean of the scaled-corrected LSWT.

$$T_D^c = T_D^b - \mu(T_D^b)$$
$$T_D^d = \mu(T_D^b) + T_D^c \frac{\sigma(T_M^a)}{\sigma(T_D^c)}$$

$T_D{}^b$ was further adjusted by dividing it by its standard deviation *(std ($T_b{}^d$))* and finally, bias-corrected by multiplying it by the standard deviation of $T_b{}^d$ *(std ($T_b{}^d$)).*

$$Adjusted\ LSWT = \frac{T_D^b}{std(T_D^b)}$$
$$Corrected\ LSWT = \ \ Adjusted\ LSWT * std(T_D^d)$$

4.5   Evaluation of LSWT products and fused dataset against the in-situ measurements

The in-situ measurement represents a point measurement of the LSWT, while the satellite LSWT products have a spatial resolution ranging from 100 m up to 5 km. This spatial mismatch can introduce bias, especially when coastline pixels are included. Therefore, it is essential to consider this spatial mismatch when comparing the in-situ measurements with satellite data. Previous studies have mitigated the effects of spatial mismatch by employing a 3x3 grid centered around each measurement station (Reinart & Reinhold, 2008; Schneider et al., 2009). The grid accounts for the possible error due to shoreline pixels and simultaneously represents are large area of the lake (Schneider et al., 2009). The satellite pixels within the 3x3 grid were averaged and the standard deviation for each pixel within the grid was calculated to identify areas with homogenous temperatures. If the standard deviation of a pixel was above 0.75, then the

corresponding data was rejected (Reinart & Reinhold, 2008). This procedure was applied for all satellite-derived LSWT products, as well as for the fused and bias-corrected LSWT product.

Additionally, the correctness of each pixel was accessed for the products with a quality control layer. For the CGLOPS and GloboLakes products, only pixels with a quality flag higher than 3 were considered, ensuring that only pixels with acceptable or higher quality were included in the analysis. For the MODIS product, the data needed to be considered as good data and the LST error below 3. The TIRS data was only used when the quality layer indicated clear skies. For ARC-Lake no quality control layer was available, therefore, no quality control requirement was applied.

The ERA5-Land product was also compared to the in-situ measurements. Due to the high difference in spatial resolution between the LSM data and the in-situ measurements, it was decided to follow a station-based observation space approach which was previously used in Yilmaz (2023). The approach obtained the LSWT value of the ERA5-Land dataset for the cell in which the observation station was located (Yilmaz, 2023).

To evaluate the accuracy of each LSWT product and the fused dataset against the in-situ measurements, the mean absolute error (MAE), the root mean square error (RMSE) and the R^2-value. Additional to the calculation between fused and clear pixels on cloudy days, the bias *(bias$_{In-situ}$)* between the satellite-derived LSWT and in-situ measurement was calculated. This was done by using following equations:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |P_i - O_i|$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(P_i - O_i)^2}{N}}$$

$$R^2 = \frac{(\sum_{i=1}^{N}(O_i - \bar{O}_i)(P_i - \bar{P}_i))^2}{\sum_{i=1}^{N}(O_i - \bar{O}_i)^2 \sum_{i=1}^{N}(P_i - \bar{P}_i))^2}$$

$$bias_{In-situ} = Satellite_{LSWT} - InSitu_{LSWT}$$

where, N represents the number of observations, O_i and $\bar{O}_i$ the observation and mean observation data and P_i and $\bar{P}_i$ the fused and mean fused LSWT, respectively.

5 Results

<u>5.1 Comparisons of LSWT variation between the measurement stations</u>

After comparing the measured LSWT from the three measurement stations, it was found that there were 61 matching days. On average, there were 2.26 matching days per year since 1996. There were no matching dates for all three measurement stations before the year 1996. Spatial variations in LSWT were observed (as shown in Figure 4). No specific general pattern was discovered, but on average, Dagskärsgrund showed a higher temperature with an average LSWT of 8.79 °C compared to Tärnan and Megrundet with 7.82 °C and 8.15 °C, respectively. The maximum variation in LSWT on a single day was observed between Dagskärsgrund and Tärnan, with a difference of 7.1 °C. This highlights the spatial variability in LSWT among the measurement stations.

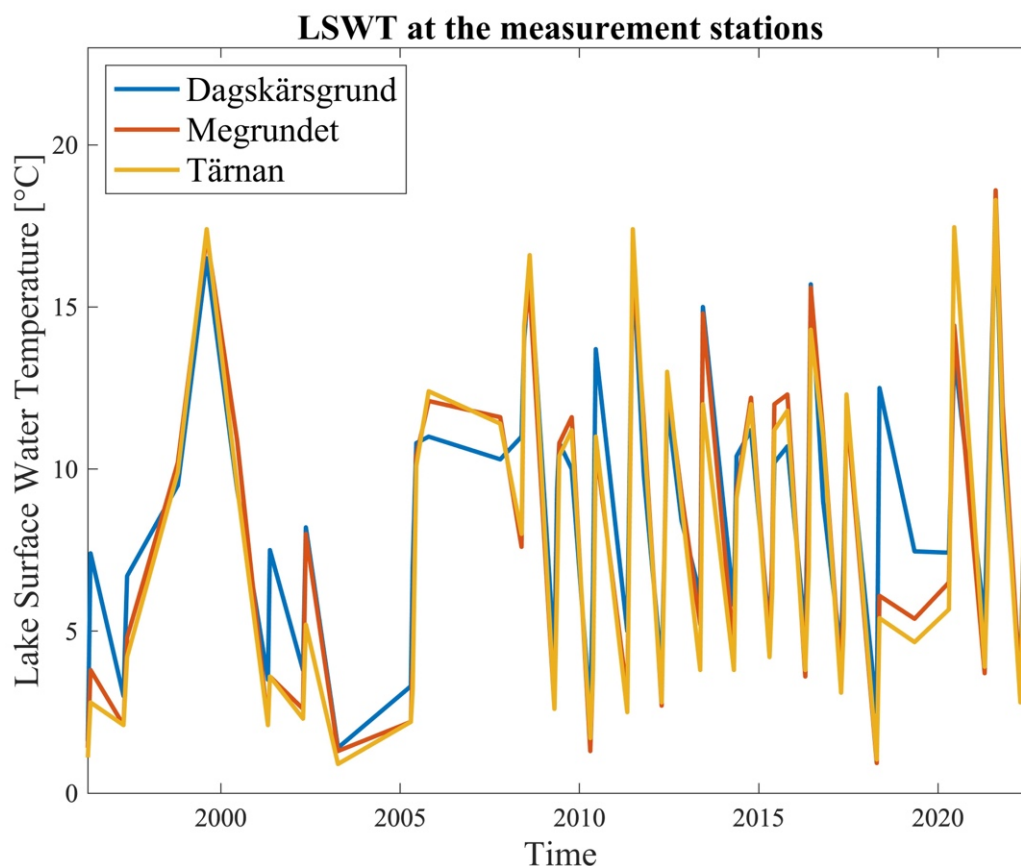***Figure 4.*** *Lake surface water temperature (LSWT) measurements for each station on days with available measurements for all three stations.*

<u>5.2   Evaluation of satellite LSWT products against in-situ measurements</u>

Figure 5-7 shows the comparison of all six satellite LSWT products against in-situ measurements with their evaluation metrics presented in Table 2. The TIRS product

22

only matched with the in-situ measurements on five days, which was the lowest number of matches among all the satellite-derived products. The MAE and RMSE were 0.91 °C and 1.11 °C, respectively, indicating a high accuracy in estimating LSWT in Lake Vänern, Additionally, a R^2-value of 0.96 was observed. It was found that TIRS under-estimated the LSWT compared to the in-situ measurement by -0.83 °C on average. In contrast, a total of 139 matching days were found for the MODIS product, which was the highest number of matches among all the satellite products. This suggests a better temporal overlap between MODIS and the in-situ measurements than other satellite products. The evaluation metrics show that the MODIS product had an MAE of 1.13 ° and RMSE of 1.64 °C, which were relatively higher compared to the other satellite products. The R^2 value was 0.93. Similar to TIRS, an underestimation of -0.57 °C was found (Table 2).
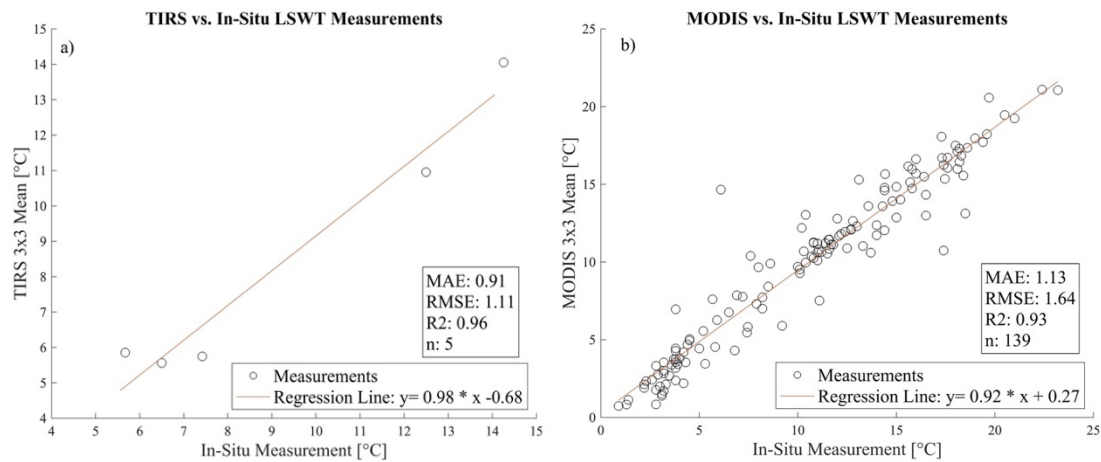

Figure 5. Comparisons of satellite lake surface water temperature (LSWT) products from TIRS (a) and MODIS (b) against in-situ measurements.

The GloboLakes dataset matches 68 times with the observation data, which was the second-highest value. Furthermore, the second highest accuracy of all the datasets with an MAE of 0.56 °C, RMSE of 0.79 and an R^2 value of 0.97. The dataset and the in-situ measurements had a strong positive correlation (Figure 6a). In contrast to the other LSWT products, it can be seen that GloboLakes overestimated the LSWT with an average bias of 0.04 °C, which is also the second lowest bias (Table 2). The CGLOPS product corresponded 21 times with the in-situ measurements. The evaluation metrics illustrate an MAE of 1.26 °C and RMSE of 1.82 °C and an R^2 of 0.93. The R^2 was higher compared to the other products, but at the same time, the RMSE was higher compared to the ARC-Lake, GloboLakes and TIRS datasets, while compared to MODIS it was lower. Like MODIS and GloboLakes a strong positive relationship with the in-situ measurement was seen (Figure 6b). The bias between CGLOPS and the in-

situ measurements indicated, similar to previous LSWT products, that CGLOPS was underestimating the LSWT, on average by -0.36 °C (Table 2).
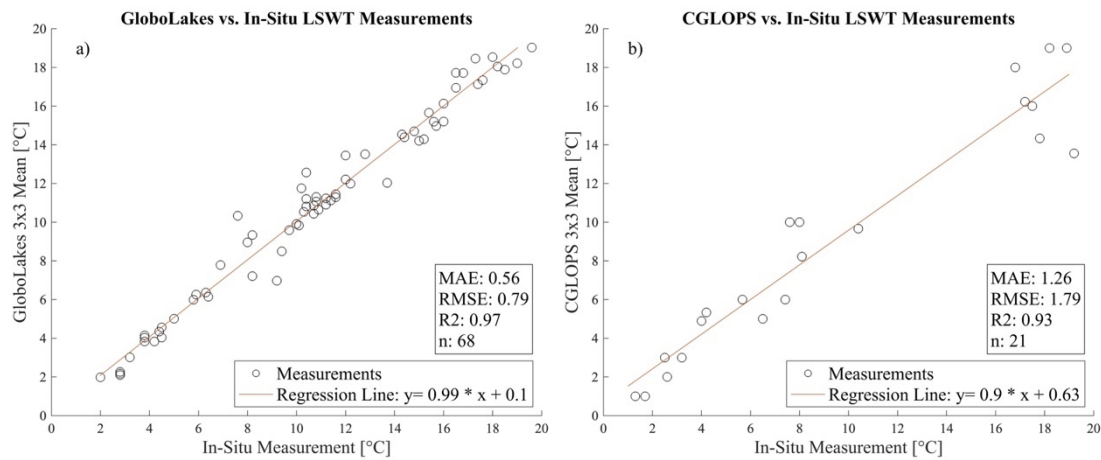

Figure 6. *Comparisons of satellite lake surface water temperature (LSWT) products from GloboLakes (a) and CGLOPS (b) against in-situ measurements.*

The AATSR dataset of ARC-Lake matched 13 times with the in-situ measurement. For these matches, it was found that the MAE and RMSE with 0.75 °C and 1.06 °C, respectively, were lower compared to other LSWT products. The R^2-value was 0.96 °C. The relationship between the AATSR and the field measurements showed a high correlation with 0.89. (Figure 7a). The AATSR LSWT product slightly overestimated the LSWT with 0.03, which is the lowest bias of all LSWT products (Table 2). ATSR2 corresponded six times with the in-situ measurements, which is less compared to other LSWT products. The MAE and the RMSE of the ATSR2 dataset were 0.32 °C and 0.46 respectively, which were the lowest compared to the other LSWT products. The R^2 was very high with a value of 1 (Figure 7b). In contrast to AATSR, it was calculated that the ATSR2 dataset estimated lower LSWT for the measurement stations and therefore underestimated the LSWT by -0.31 °C on average (Table 2).
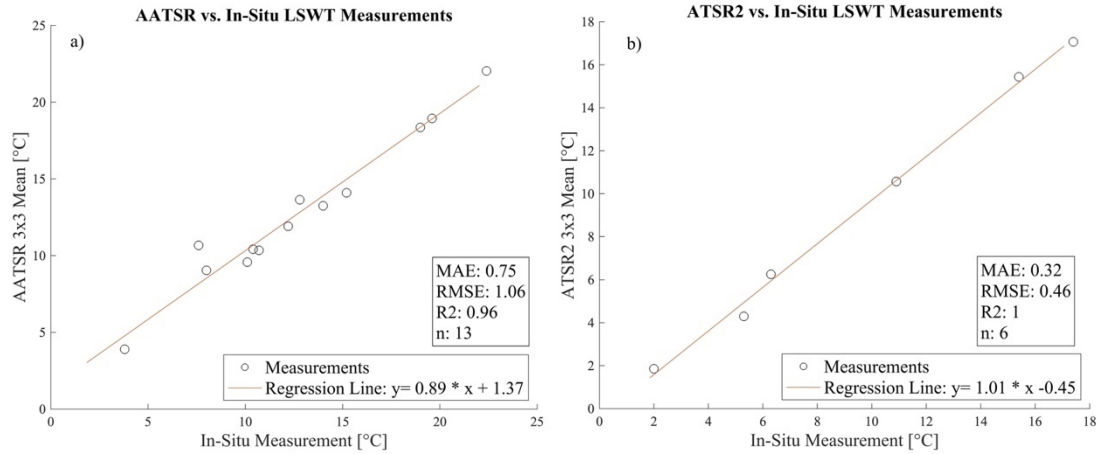
Figure 7. *Comparisons of lake surface water temperature (LSWT) products from AATSR (a) and ATSR2 (b) against in-situ measurements.*

Table 2. *Bias$_{In\text{-}situ}$, MAE, RMSE, R^2 and number of matches between lake surface water temperature product and in-situ measurements.*

	Bias$_{In\text{-}situ}$ (°C)	MAE (°C)	RMSE (°C)	R2	Number of Matches
MODIS	-0.56	1.13	1.64	0.93	139
TIRS	-0.83	0.91	1.11	0.96	5
Glo-boLakes	0.04	0.56	0.79	0.97	68
CGLOPS	-0.36	1.26	1.79	0.93	21
ATSR2	-0.31	0.32	0.46	1	6
AATSR	0.03	0.75	1.06	0.96	13
ERA5-Land	0.50	1.32	1.62	0.93	780

5.3 Evaluation of ERA5-Land reanalysis product against in-situ measurements

Figure 8 shows the comparison of the ERA5-Land product with the in-situ measurement. The accuracy of the ERA5-Land product is presented in Table 2. In total 780 matches were found during the temporal length of the observation data from 1973 to 2022. The ERA5-Land product agreed well with the in-situ measurements with an R^2 value of 0.93. The error was low with an MAE of 1.32 °C and RMSE of 1.62 °C, which was similar to the error of the CGLOPS and MODIS products.
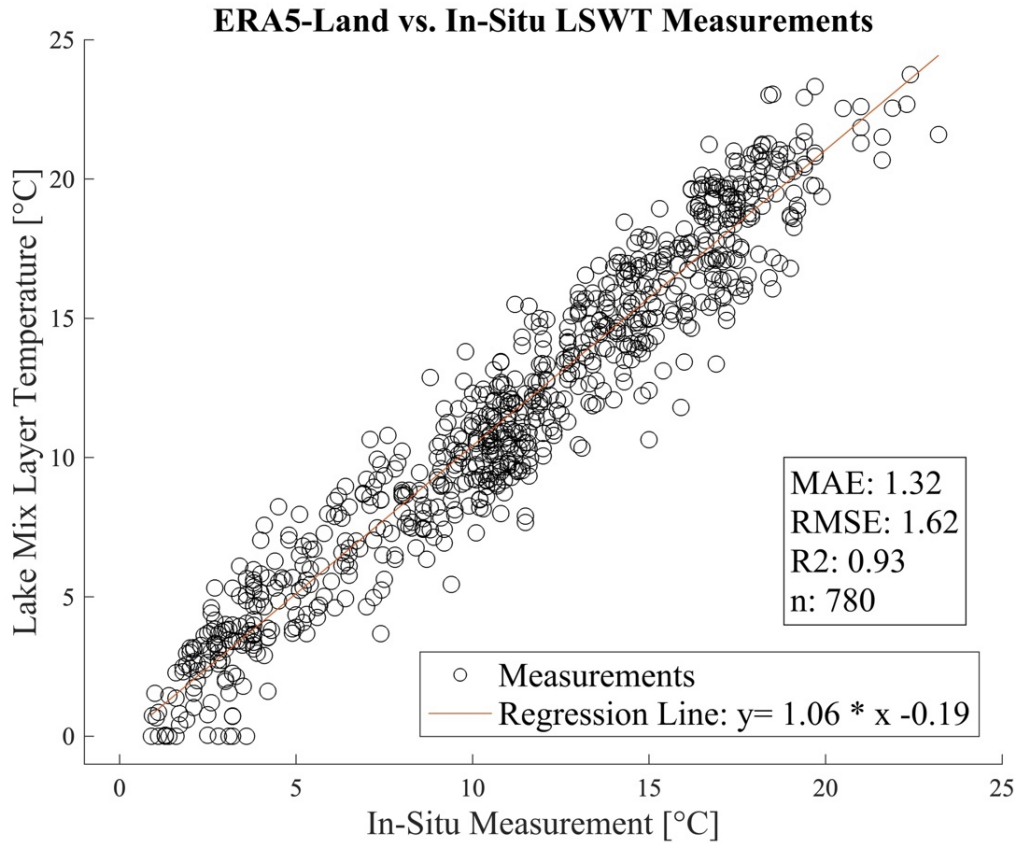
Figure 8. *Evaluation of ERA5-Land against in-situ measurements.*

5.4 Generation of fused LSWT dataset and evaluation

5.4.1 Division of satellite LSWT products

Table 3 shows the clear and cloudy days for each satellite-derived LSWT product. The performance of the model is depending on the accuracy of the input data. Since it can be assumed that for clear days the accuracy is high the dataset was split into clear and cloud days to use the clear days as an input for the model. During the time period 2000-2023, the MODIS product had a total of 273 clear days and 8,051 cloudy days. The ARC-Lake product was split into 592 clear days for ATSR2 and 731 clear days for AATSR for the time period 1995-2012. The GloboLakes product had the third lowest number of clear days, with only 19 days identified as clear during its 21 yearlong lengths For the CGLOPS product, only 8 clear days were found throughout its operational time so far. TIRS has been operating since 2016, however, due to its poor temporal resolution, only every 16 days data is available. This caused a generally low number of matches in which none of those days had a complete dataset or were completely free of clouds.

The highest number of clear days were found in ARC-Lake; however, these products ended in 2012, due to the discontinuance of the ATSR2 and AATSR sensors. Therefore,

finally based on the availability of clear days, good spatial and temporal resolution, and continuous dataset, it was decided to conduct the data fusion using the MODIS product in this study.

Table 3. Number of clear and cloudy days for each lake surface water temperature product.

	MODIS	TIRS	GloboLakes	CGLOPS	ATSR2	AATSR
Clear Days	273	/	19	8	592	731
Cloudy Days	8051	166	3829	552	1043	1505

5.4.2 Evaluation of the fused LSWT dataset without bias correction

The data fusion was conducted for the time period between 2018 and 2019 using the MODIS dataset. The time period was chosen because 37 clear days (23 in 2018 and 14 in 2019) as well as 18 in-situ measurements were available during this time period. It was the second highest number of in-situ measurements for two years, only 2003/2004 and 2006/2007 had one extra in-situ measurement. However, 2018/2019 had with 37 clear days a higher number of clear days compared to 2003/2004 with 18 and 2006/2007 with 24 clear days. The uncorrected fused LSWT product will be further named as the fused product, while the bias-corrected LSWT product will be referred to as the corrected LSWT product.

Figure 9 illustrates the spatial pattern of the LSWT on a clear day next to the fused LSWT pattern on the 3rd of June 2018. As can be seen, the fused LSWT not only captured the spatial pattern well, but it also reflected the magnitude of the LSWT on the clear day in a similar range.
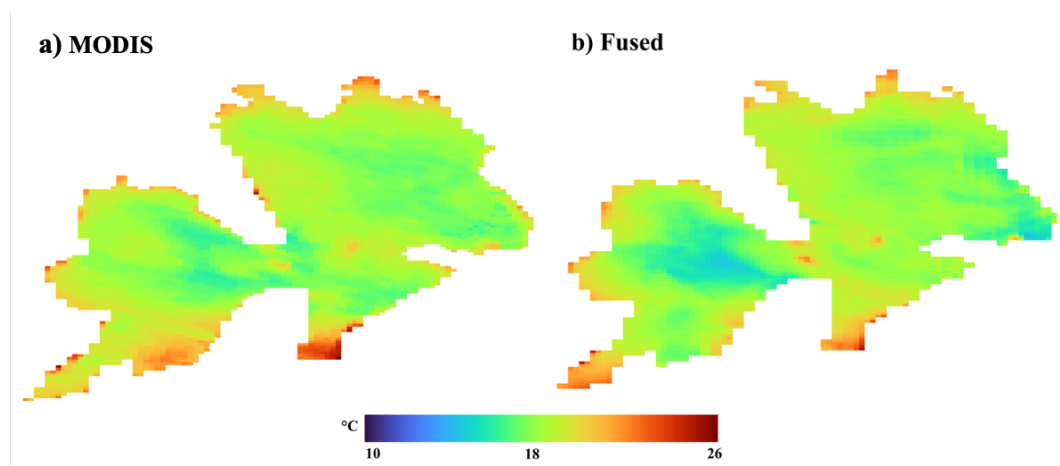

Figure 9. Spatial LSWT pattern of the MODIS lake surface water temperature (LSWT) (a) and the fused (b) LSWT on a clear day in June 2018.

Figures 10 and 11 show the LSWT spatial variation of the MODIS, ERA5-Land, fused and bias-corrected LSWT for the 7th of July 2018 and the 20th of January 2019. The dates were chosen to illustrate the difference between the LSWT during summer and winter. As it can be seen the MODIS LSWT indicated data gaps due to clouds. Higher temperatures were seen in the center, the minimum varied much for the rest of the region. In contrast, the ERA5-Land LSWT had no data gaps, although the spatial resolution was spare. The maximum temperature for this day was reached in the North of Lake Vänern and the minimum temperature was found in the center, indicating a different spatial variation of LSWT than MODIS. The fused LSWT was able to illustrate the LSWT in a high spatial resolution and a similar pattern as the ERA5-Land but with a lower magnitude. Like the ERA5-Land, the highest LSWT was reached in the North of the Eastern basin as well as in the South of the Western basin. The minimum was slightly more south-located than the ERA5-Land (Figure 10).
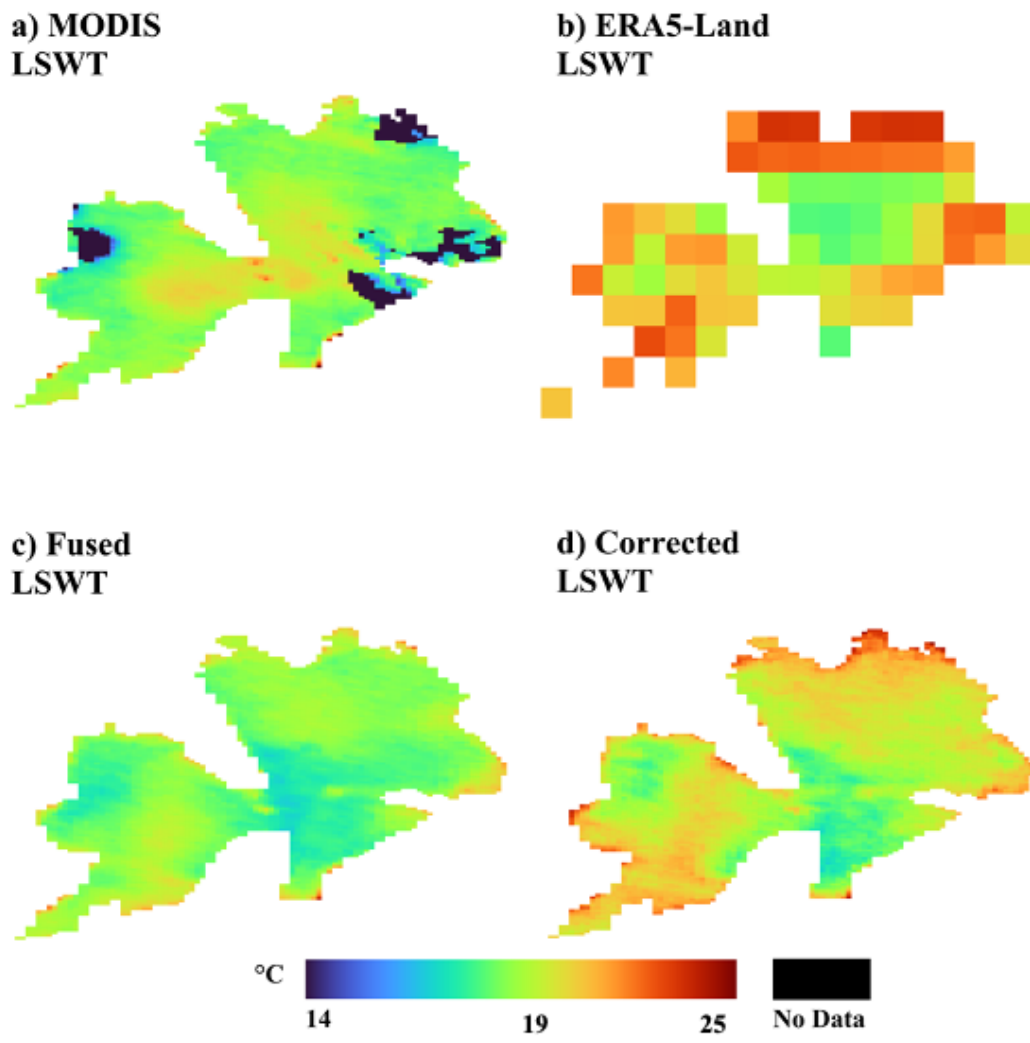

Figure 10. Lake surface water temperature (LSWT) for MODIS (a), ERA5-Land (b), fused (c) and bias-corrected (d) LSWT in July 2018.

Similar to Figure 10, the MODIS LSWT indicated a spatially incomplete dataset, while the ERA5-Land data represented the LSWT with a low spatial resolution. The fused LSWT product was able to fill the gaps of MODIS as well as maintain a high spatial resolution. In contrast to Figure 10, it can be seen in Figure 11 that the spatial dynamics between the LSWT products did not vary as much. All LSWT products reach their maximum temperatures in the center of Lake Vänern. MODIS showed a slightly lower temperature. The minimum temperatures were found close to the shore for MODIS and ERA5-Land. The fused LSWT product showed its lowest temperature in the south of the western basin of Lake Vänern (Figure 10).
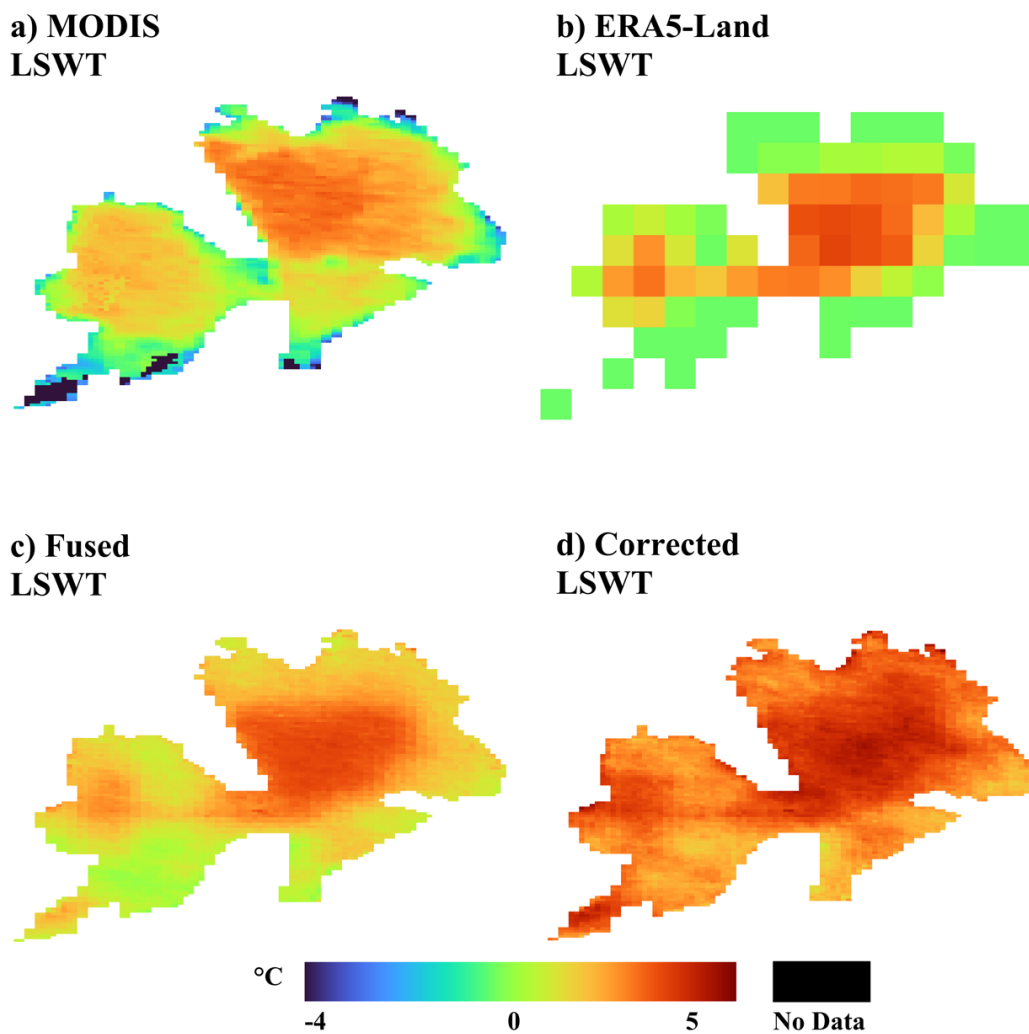
Figure 11. Lake surface water temperature (LSWT) variation for MODIS (a), ERA5-Land (b), fused (c) and bias-corrected (d) LSWT product in January 2019.

Figure 12 shows the statistical measures for the fused LSWT pixels against the clear pixels of cloudy days of MODIS. Each pixel represents the mean value of the statistical measure. The MAE showed a generally low value with 1 °C for the whole lake. The maximum value of 3.09 °C was reached close to the shore (Figure 12b). The bias had slightly higher values with about 1.5 °C for the center of the lake. Lower values were

found in the North of the Lake where the values were about 0.5 °C (Figure 12c). The RMSE followed a similar pattern to the MAE with higher values closer to the shore. The maximum RMSE is 4.16 °C and the lowest RMSE value was about 1.43 °C (Figure 12d). The R^2 pattern did not follow an overall trend. However, the R^2 generally showed a high value for the whole lake with a maximum R^2 of 0.99 and a minimum of 0.93 (Figure 12e).

Figure 12. *Spatial dynamics of mean statistical measures of pixel number (a), MAE (b), Bias (c), RMSE (d) and R2(e) between fused lake surface water temperature pixels and clear pixels on a cloudy day of MODIS pixel.*
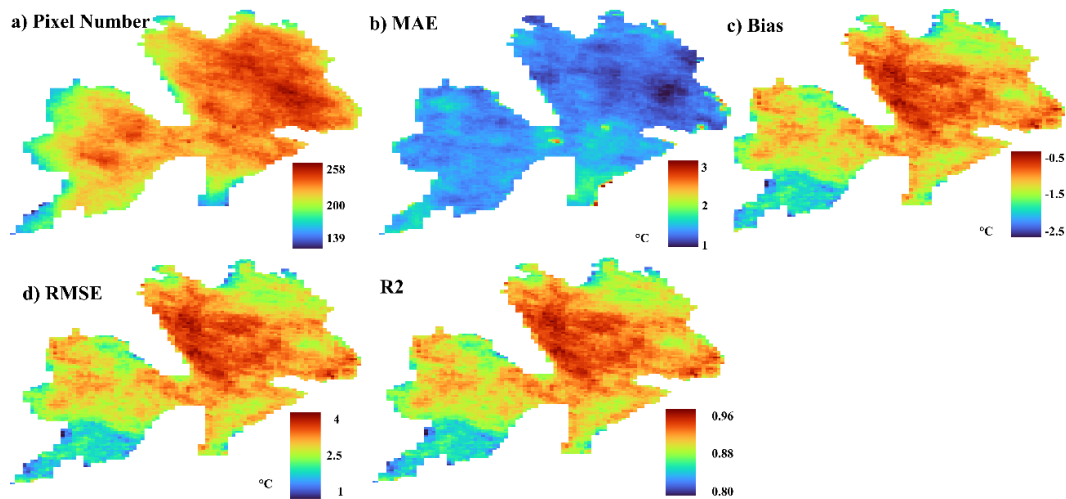
Figure 13 shows the evaluation of the MODIS, ERA5-Land and the fused LSWT against the in-situ measurements for the time period 2018/2019. For the MODIS product, only six days were found where MODIS had sufficient data. The low number of matches was caused by clouds which result in data gaps. The evaluation metrics show that the MODIS LSWT product had a low accuracy in terms of RMSE, MAE and R^2 with 2.21 °C, 3.64 °C and 0.41, respectively. Additionally. The ERA5-Land product also had much more days where it matched with the in-situ measurements. The ERA5-Land performed better with an MAE of 1.65 °C, RMSE of 2.13 °C and an R^2 of 0.88, the fused LSWT had the same number of matches with the in-situ measurements as the ERA5-Land. The fused LSWT indicated a higher accuracy compared to the MODIS and ERA5-Land LSWT in terms of the MAE with 1.49 °C, however, the RMSE of 2.15 °C and R^2 of 0.87 was slightly lower compared to the ERA5-Land LSWT. Indicating it was able to reach the accuracy of ERA5-Land (Figure 13).
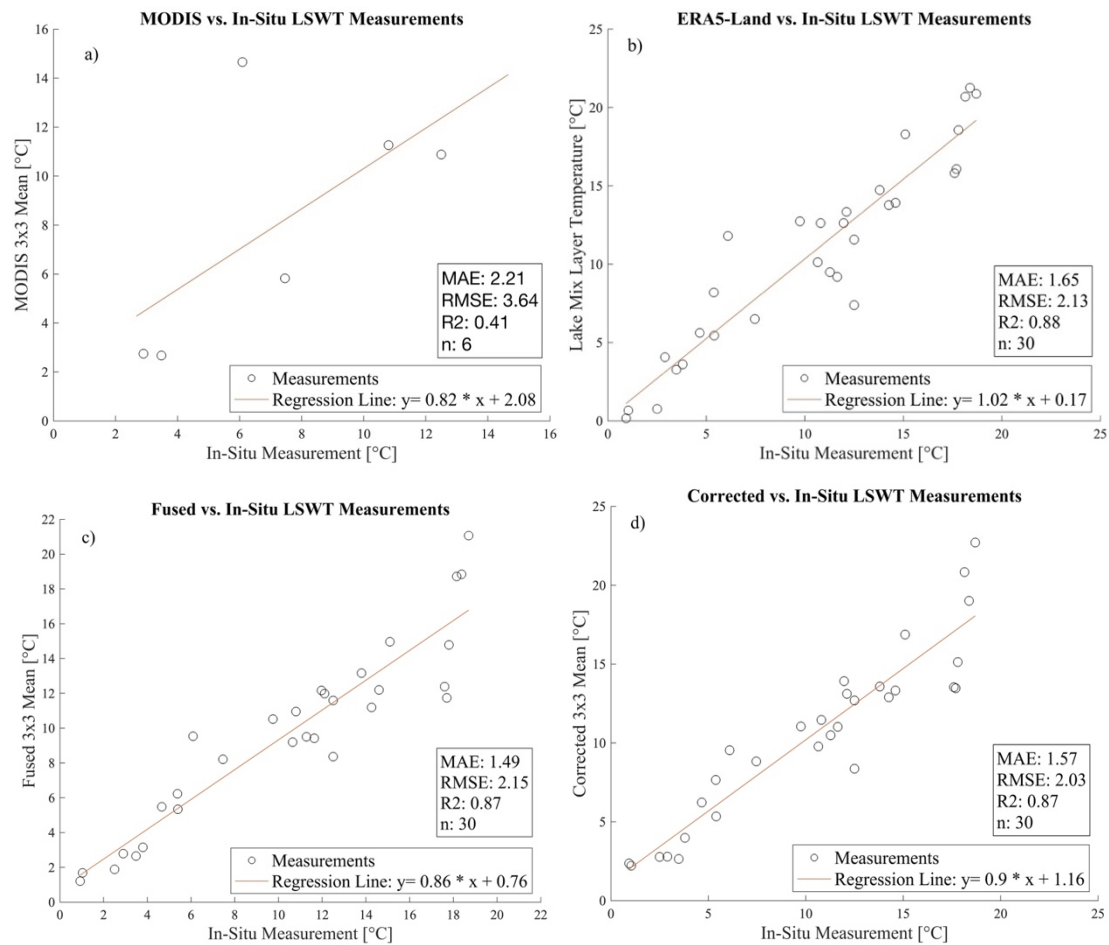
Figure 13. *Evaluation of MODIS LSWT (a), ERA5-Land (b), fused (c) and bias-corrected lake surface water temperature (LSWT) (d) against in-situ measurements for the period 2018/2019.*

5.4.3 Evaluation of the bias-corrected LSWT product

The fused LSWT was corrected based on a linear and variance scaling approach to account for the systematic bias in the results. In Figure 10 and 11 the spatial dynamics of LSWT products were displayed. The fused LSWT product was able to illustrate the LSWT in a high spatial resolution, this was also seen for the bias-corrected LSWT. Similar to the fused LSWT product, the bias-corrected LSWT found higher temperatures in the North of the Eastern basin and in the South of the Western basin in July 2018, although, with a higher magnitude (Figure 10). Figure 11 presents the spatial dynamics of the fused and bias-corrected LSWT in January 2019, displaying that the fused LSWT indicated higher temperature in the center of Lake Vänern, and the lowest temperatures were found in the South of the Western basin. This trend was also found in the bias-corrected LSWT product, however, similar to Figure 10, it was seen that the magnitude of the LSWT is higher (Figure 11).

In Figure 13 the MODIS LSWT, the ERA5-Land, the fused and the bias-corrected LSWT products were compared against the in-situ measurements. After the bias correction, it can be seen that the bias-corrected LSWT was able to maintain the high number of matches with the in-situ measurements, nonetheless, the MAE was somewhat higher with 1.57 °C compared to the fused LSWT. In contrast, the RMSE (2.03 °C)) indicated that the bias-corrected LSWT performed slightly better than the fused and the MODIS LSWT. The R^2 did not change (Figure 13).

Further, the temporal variation of the bias-corrected LSWT was compared to the ERA5-Land, clear pixel of MODIS LSWT as well as the in-situ measurement to assess if the LSWT products followed a similar temporal pattern. The temporal variation of the LSWT products for each station is displayed in Figure 14-16. Gaps in the MODIS data are due to clouds influence and therefore no data was available for that day/period. Moreover, it has to be noted that between the 21st of February 2018 and the 7th of April 2018, no ERA5-Land data was available for the measurement stations, therefore, the corresponding values in the bias-corrected LSWT were set to not a number. In the Figures, it can be seen that all three products followed a similar trend in the range of the in-situ measurement for all three measurement stations. Only for Dagskärsgrund a slight difference between the in-situ measurements and the LSWT products was seen (Figure 16). An interesting finding to point out is that the bias-corrected aligned with the available clear pixels almost perfectly. When no clear pixel was available the bias-corrected LSWT showed a higher magnitude compared to ERA5-Land LSWT, this was mainly the case during the winter months.

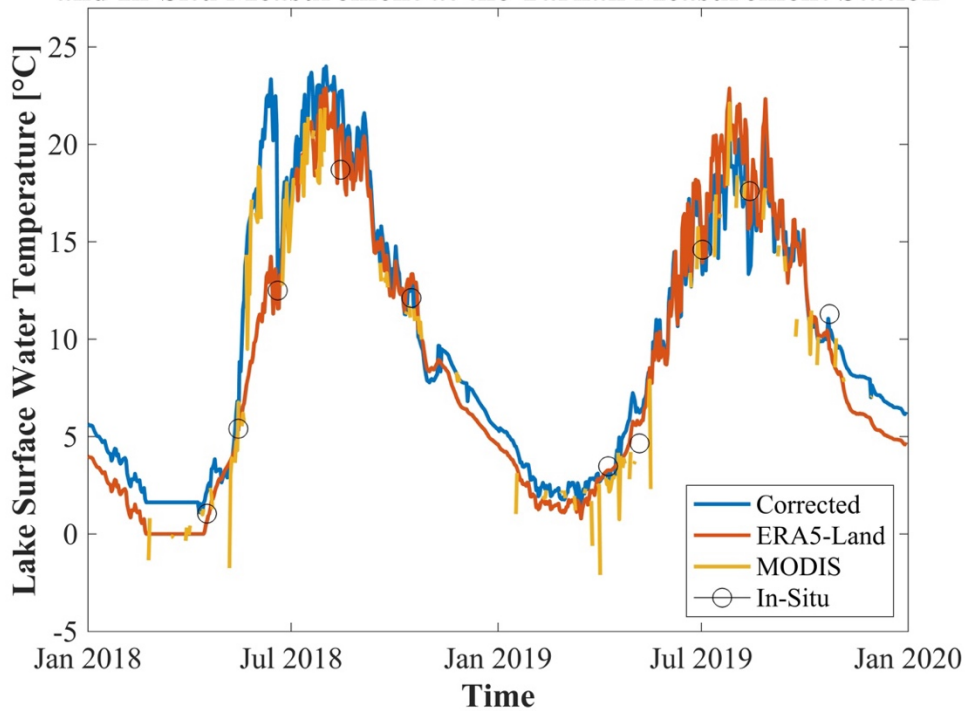***Figure 14.*** *Daily lake surface water temperature (LSWT) of the bias-corrected, clear pixels of MODIS and ERA5-Land products at the Tärnan measurement station from January 2018 to December 2019.*
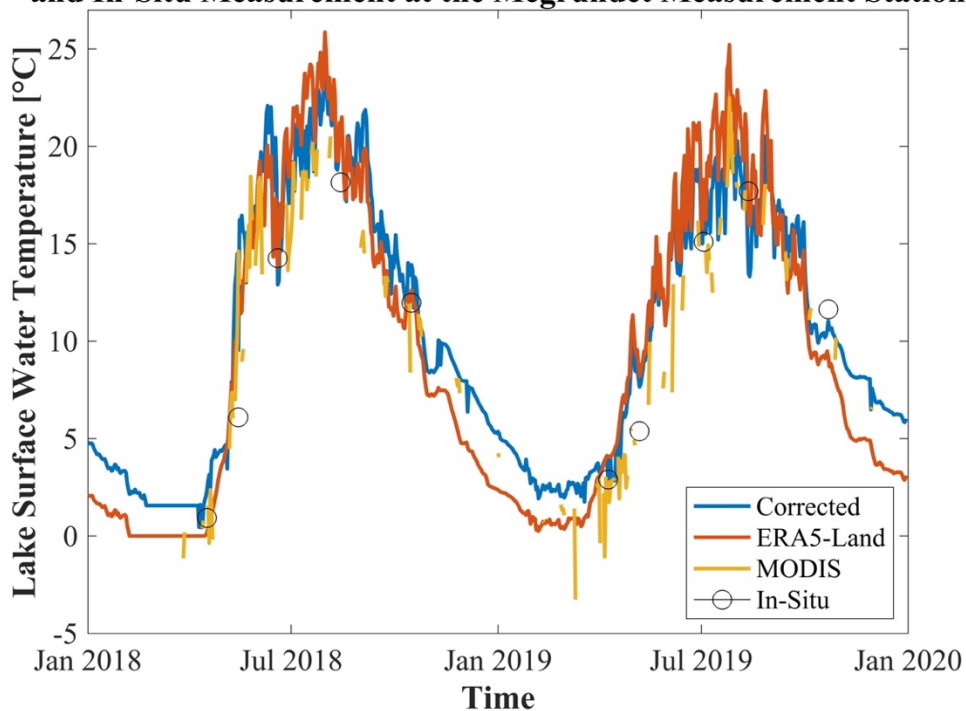

*F**igure 15.*** *Daily lake surface water temperature (LSWT) of the bias-corrected, clear pixels of MODIS and ERA5-Land products at the Megrundet measurement station from January 2018 to December 2019.*

Figure 16. *Daily lake surface water temperature (LSWT) of the bias-corrected, clear pixels of MODIS and ERA5-Land products at the Dagskärsgrund measurement station from January 2018 to December 2019.*

5.5   Spatial and temporal dynamics of lake surface water temperature

5.5.1 Temporal variability of lake surface water temperature of Lake Vänern

As described above the temporal variation of the bias-corrected product followed the ERA5-Land data good, therefore, the monthly mean temporal LSWT variation of the study period as well as the spatial variation of the monthly means of each year will be further investigated. Figure 17 illustrates the monthly mean LSWT of the bias-corrected LSWT product. The monthly mean LSWT was calculated by lake-wise averaging the LSWT and further, the mean of all daily bias-corrected LSWT was taken to obtain the monthly average LSWT.

Figure 17 shows that the year 2018 started with a mean LSWT. In January, which was followed by a decrease in February and further reached its minimum in March. Afterward, the LSWT started to rise and reached its maximum in August and then decreased in December. The year 2019 followed a similar trend, however, it started with the lowest temperatures in January with and increased up until July when it reached its maximum. Except for January, it can be seen that 2019 indicated a higher mean LSWT up until August compared to 2018. The fall LSWT was in the same magnitude for both years. Comparing the two years with each other no clear trend was seen for the first three months. However, for both years a strong positive trend in temperature was

34

detected and was peaking in the summer months. The following month the trend for the years a similar indicating a strong decrease in temperature (Figure 17).
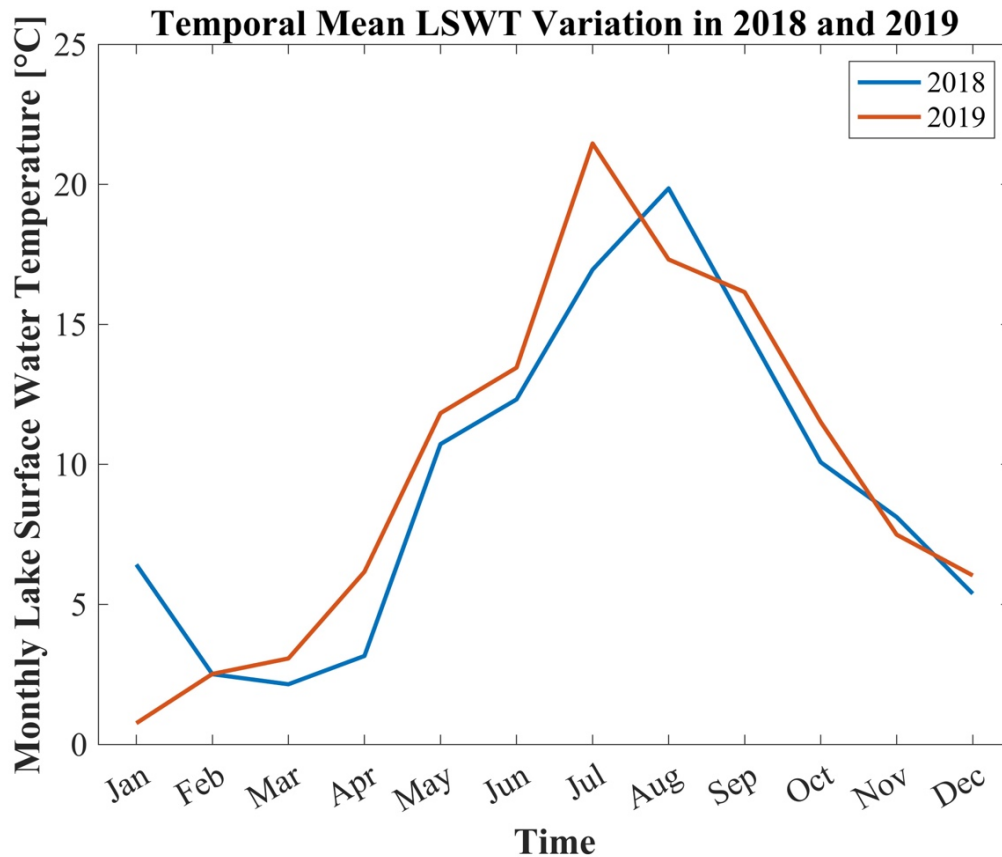
Figure 17. *Monthly mean lake surface water temperature (LSWT) of Lake Vänern in 2018 and 2019.*

5.5.2   *Spatial variability of lake surface water temperature of Lake Vänern*

To further investigate the spatial dynamics of Lake Vänern the monthly mean temperature for each pixel was calculated for 2018 and 2019. Figure 18 illustrates the spatial variation of the monthly averages for Lake Vänern. The results show that the bias-corrected LSWT not only filled the data gaps of MODIS but was also able to provide detailed spatial dynamics found for each month. Low LSWT were found in the winter months for both years. The year 2018 indicated slightly higher temperatures in the Eastern basin in January. The winter of 2019 showed lower temperatures compared to the previous year. During May and June, the spatial variation and magnitude of LSWT was comparable between the years, showing lower temperatures in the basins and higher temperatures closer to the coast. 2019 reached its maximum already in July, while the highest temperature in 2018 was in August with a lower magnitude compared to the following year. The LSWT decreased for both years from September on, following a similar pattern with decreasing LSWT at the coast faster (Figure 18).
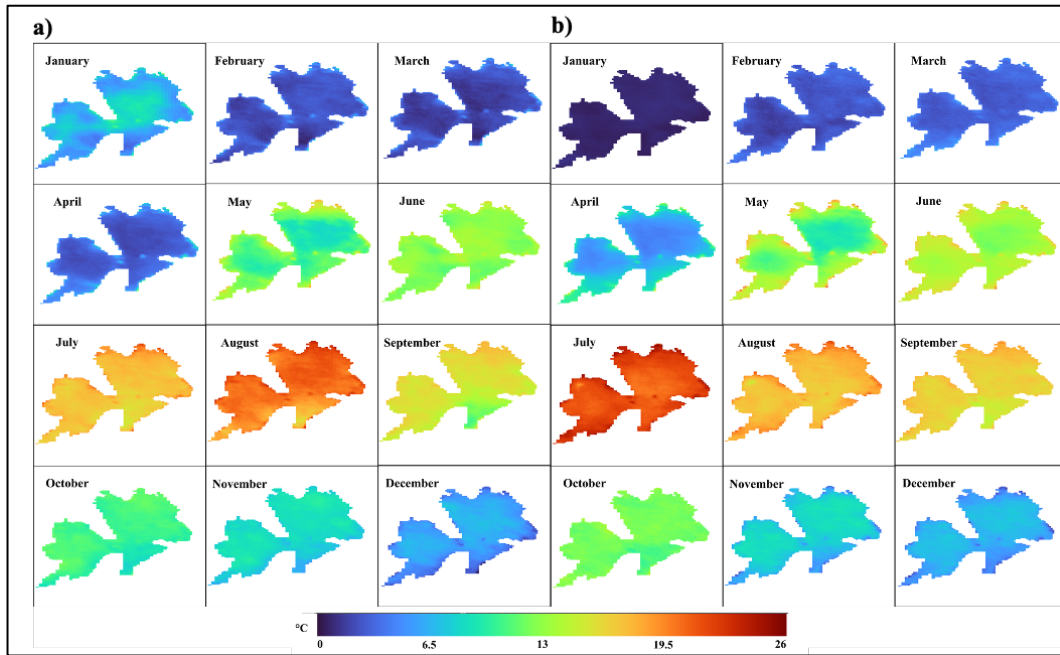
Figure 18. *Spatial dynamics of monthly mean lake surface water temperature of Lake Vänern in 2018 (a) and 2019 (b).*

6 Discussion

6.1 Evaluation of the satellite and reanalysis LSWT products

The evaluation of the LSWT products indicates that all of them were able to capture the LSWT with high accuracy, however, variations in their accuracy were found. Overall, the MAE varied between 0.32 and 1.32 °C, the RMSE ranged between 0.46 and 1.73 °C. The results indicate that the ATSR2 and GloboLakes LSWT products performed with higher accuracy. Comparing the results of this study with previous studies it can be found that the LSWT products fall into the same range or even outperform previous studies. The MAE and RMSE of MODIS found in this study of 1.13 °C were coherent with the MAE range found in Sima et al. (2013) (0.8-1.9 °C) and the RMSE < 2 °C in Lazhu et al. (2022). A higher accuracy compared to previous studies was found for GloboLakes, CGLOPS, ARC-Lake and TIRS. Zhang et al. (2021) indicated that the RMSE for GloboLakes and CGLOPS was approximately 2.91 and 2.33 °C, respectively, which was significantly higher compared to the RMSE found in this study for GloboLakes with 0.79 °C. The RMSE of CGLOPS was in a comparable range with 1.79 °C (Table 2). During the validation of ARC-Lake, the mean difference of both subsets to in-situ measurement was found to be 0.3 ± 0.9 °C (MacCallum & Merchant, 2010), covering similar ranges seen in this study with a bias$_{\text{In-situ}}$ ranging from -0.31 to 0.03 for ATSR2 and AATSR, respectively. Table 2 shows that TIRS had a similar accuracy compared to the results seen in Sharaf et al. (2019), which displayed an RMSE of 0.73 °C, an MAE of 0.71 °C and an R^2-value of 0.97. Last but not least, the ERA5-Land lake mix layer temperature dataset was evaluated against the in-situ measurement.

36

The results show that the values discovered in this study were significantly higher than those found in Zhang et al. (2021). The previous study indicated that the RMSE was 3.41 °C and the bias$_{\text{In-situ}}$ 1.64 °C, which did not align with the values presented in Table 2. Since no LSWT product showed a significantly lower accuracy compared to previous studies it can be said that all evaluated products represented the LSWT accurately. The presented MAE and RMSE of ERA5-Land are lower compared to other studies (Muñoz-Sabater et al., 2021; Zhang et al., 2021), indicating a better performance for Lake Vänern.

It has to be noted that the satellite-based LSWT estimations represent the skin temperature of the lake, while the in-situ measurements indicate the bulk temperature of the lake. The cool skin of the lake causes the LSWT products to underestimate the LSWT compared to the in-situ measurements (Carrea & Merchant, 2020a; Dörnhöfer & Oppelt, 2016; Zhang et al., 2021). The literature reports that the skin effect can cause bias ranging between 0.1 and 0.6 °C (Sima et al., 2013). The relationship between the bulk and skin temperature is influenced by the heat flux between the atmosphere and the lake as well as the water column's thermal stratifications. The relationship is also determined by wind speed, the time of the day as well as the depth of the measurement (Reinart & Reinhold, 2008). Additional to the skin temperature effect other factors such as satellite biases, errors in the in-situ measurements and near-surface stratification influence the accuracy of the LSWT product (Carrea & Merchant, 2020a). The negative values of the bias$_{\text{In-situ}}$ can be seen for almost all satellite-derived LSWT products in this study. According to the quality assessment report of CGLOPS, the bias ranges between -0.24 °C ± 0.88 °C, supporting the findings of this study for CGLOPS. Only GloboLakes and AATSR indicate an overestimation, however, the bias is comparably low (Table 2). This study is not addressing the bias caused by the cooling effect, since the wind speed and heat exchange parameters are unknown. In contrast to the satellite-obtained LSWT products, the ERA5-Land overestimates the LSWT showing positive values. The overestimation of ERA5-Land is caused by the dependency on the accuracy of the lake depth in the input data (Muñoz-Sabater et al., 2021).

It is also important to note that all satellite-derived LSWT products show a very low number of matches compared to available in-situ measurements. For instance, the GloboLakes dataset was available daily for 21 years, during that time 180 in-situ measurements were taken, however, the evaluation shows that only 68 matches with the in-situ measurements were found (Table 2). The low number of matches was due to missing data because of pixels with low quality and high variation of LSWT within the 3x3 mismatch matrix. That underlines the need for a spatially complete and temporal continuous LSWT dataset.

6.2   Evaluation of the generated LSWT dataset

The data fusion was conducted by combining the ERA5-Land data with the MODIS data. As described before MODIS did not perform as well as GloboLakes or the ARC-Lake subsets. However, the accuracy was still considered as good, as well due to the fact that MODIS is still continuing it was decided to conduct with it. Additionally, it could be argued that the ERA5-Land data represents the LSWT accurately enough, however, the poor spatial resolution of the product cannot account for the spatial dynamics found in lakes and therefore it is necessary to obtain an LSWT with a high spatial resolution.

The comparison of the fused LSWT with the MODIS LSWT on a clear day shows that the approach was able to represent the LSWT with similar spatial patterns and magnitude (Figure 9). The results of the conducted data fusion supported that the method was able to fill the gaps of MODIS and therefore generated a spatially complete LSWT dataset for Lake Vänern . Additionally, it was able to create a dataset that has a higher spatial resolution, compared to ERA5-Land, providing more detailed spatial information. The evaluation of the fused LSWT against the clear pixels on cloudy days indicates that there was a negative bias but that the overall accuracy as well as the correlation was high (Figure 12). Compared to Long et al. (2020) the range of the MAE and RMSE was slightly lower in the fused LSWT, while the range of R^2 shows a similar range (Figure 12). Long et al. (2020) argue that the bias is caused by different land cover classes for LST studies, however, this is not the case for this application. Therefore, other factors influence the systematic error in the model, which needs to be further investigated. Feng et al. (2006) reported that the systematic error in the original STARFM method is due to the difference in the data processing of the sources, the discrepancy of the acquisition time and the bandwidth as well as geolocation errors. Even though the ERA5-Land data is not considered a satellite product and therefore the difference in bandwidth cannot be attributed to the error found in the generated dataset. Nevertheless, due to the different nature of the input data, it can be assumed that this causes an inaccuracy in the output.

To account for the systematic bias found in the fused LSWT a bias correction was conducted. Even though the bias-corrected LSWT shows a higher agreement with the in-situ measurement in terms of the RMSE compared to the fused LSWT product, the error did not decrease enough to account for the systematic error. Comparing the difference of the RMSE before and after the bias correction with the results reported by Long et al. (2020) it shows a rather small decrease in the error. Nevertheless, the results indicate that the bias correction is suitable to correct for biases in the LSWT application.

It can be said that the spatial pattern of the fused and bias-corrected LSWT product is in good agreement with the ERA5-Land data, however, does not follow as much the spatial dynamics found in the MODIS LSWT data (Figure 10 & Figure 11) Finally, comparing the accuracy of the generated LSTW product with the evaluated LSWT products, it can be seen that the accuracies fall in a similar range. Previous studies such as Schwab et al. (1999) developed an approach where the data gaps in AVHRR were filled with temperature estimation based on earlier temperatures. The results shown in Figure 13d show are comparable to the RMSE of 1.10-1.76 °C found by Schwab et al. (1999), however, the spatial resolution in our study is higher with 1 km compared to 2.6 km. More recent studies made use of the MODIS cloud product by merging it with the MODIS LST product to obtain spatially complete data. The results indicated an RMSE ranging between 1.17 and 2.05 °C. It is important to note, that the cloud layer comes with an uncertainty of ±5 °C, which affects the accuracy of the results (Moukomla & Blanken, 2016). The range of the RMSE in this study is similar to the reported from previous studies, indicating that the method is suitable to generate spatially complete LSWT maps although with a lower accuracy than the ERA5-Land LSWT.

The evaluation of the spatial dynamics of the generated LSWT product showed that the fused and the bias-corrected LSWT products follow the spatial pattern of the ERA5-Land LSWT product. However, this was not expected since it can be seen that in equation 3 the fine-resolution data provides detailed information on spatial variation, which should determine the accuracy of the LSWT at the prediction time more than the coarse-resolution image. The relatively low number of clear days could have caused, that the fused and bias-corrected LSWT shadowed rather the ERA5-Land than the MODIS LSWT product.

The results of the temporal pattern indicated a good alignment of the bias-corrected with both the ERA5-Land LSWT and the clear pixels of the MODIS LSWT at all three measurement stations (Figure 14-16). Long et al. (2020) found a similar relationship with the MODIS and bias-corrected data, although, the paper mainly compared the in-situ measurement with the bias-corrected data. Since all three products implied a slight shift discrepancy to the in-situ measurements for the Dagskärsgrund measurement station, indicating that the accuracy of the bias-corrected might be influenced by this difference. Long et al. (2020), suggested that the temporal distance between the MODIS images does not influence the results of the data fusion as much, rather the accuracy of the MODIS LSWT may be the main influence of the accuracy. Our study had a similar result, the highest discrepancy between the in-situ measurement was found on days

where no clear pixel was found (Figure 14-16). Overall, it can be said that if clear pixels were available the bias-corrected LSWT represented the trend and magnitude of the MODSI LSWT perfectly, while for the ERA5-Land a positive bias was visible. This supports the effectiveness of the bias correction approach to partly account for the systematic error.

6.3   Spatial and temporal analysis of LSWT variation of Lake Vänern

The data fusion method was able to generate a spatially complete and temporal continuous dataset for Lake Vänern with good accuracy in terms of temporal and spatial variation, thus, the newly generated dataset was used to investigate the spatial and temporal dynamics of Lake Vänern.

The results in this study did not show a matching trend in LSWT for the beginning of the years. In 2018 a decrease in LSWT from January to February was found, while 2019 showed a slow steady increase in temperature for the first three months. Similar results were found in Reinart and Reinhold (2008) where no clear trend was seen at the beginning of the years. The LSWT in January 2018 showed higher temperatures , this might be due to the great heat capacity of large lakes (Kvarnäs, 2001), which causes the LSWT cools down slower resulting in higher temperatures in winter. The decrease in the LSWT in February and March might be caused due to the previously mentioned missing data at the beginning of 2018, which influences the accuracy of the spatial pattern. Previous studies link that the strong increase in LSWT in the following month was due to increasing air temperatures, which positively influences the LSWT (Kvarnäs, 2001; Reinart & Reinhold, 2008). The year 2018 shows a faster decrease in LSWT for the fall months compared to 2019. This trend is not supported in other studies, since a later maximum of LSWT usually results in relatively higher temperatures for the following month. The strong decline in temperature in fall can be an indicator of strong winds which usually increase in fall (Kvarnäs, 2001).

The spatial dynamics of Lake Vänern follow a distinct pattern for most of the years. Only in January, a significant difference was found between the years. As can be seen in January 2018, the basins of the LSWT still indicate higher temperatures compared to the other regions of the lake (Figure 18). This is due to the previously mentioned heat capacity of deep waters (Kvarnäs, 2001; Reinart & Reinhold, 2008). This was not seen in January 2019, but looking at the spatial pattern of December 2018 the LSWT did not vary as strong which might cause that January shows a rather similar pattern in LSWT for the whole lake. In the following month, it can be seen that a so-called thermal bar was developed. The thermal bar represents a border of 4 °C between coastal and central areas and moves with increasing water further into the center of the lake. In the area

40

between the thermal bar and the coasts, the LSWT increases and develops a thermocline. This was not seen as fast in the central area of the lake due to the depth of the basin which requires more heat per surface area is needed resulting in delayed warming. This trend was described in previous studies and is a common process in larger lakes (Kvarnäs, 2001; Reinart & Reinhold, 2008). However, the results indicate that this phenomenon does not happen during the fall months, which was also reported by Reinart and Reinhold (2008).

6.4   Limitations of this study and recommendations for future studies

Although the results of the study indicate a high accuracy of the applied method several limitations were found. The evaluation of the LSWT products could have been improved by a higher number of measurements per station which could improve the number of matches with the LSWT products, especially products with a low temporal resolution such as the TIRs and the CGLOPS dataset. Additionally, the low number of measurement stations limited the evaluation of the LSWT products to the three measurement stations, which makes it difficult to extrapolate the accuracy for the whole lake. To overcome this limitation, it was planned to conduct drone measurements to obtain LSWT measurements for a greater area and assess the spatial variation of LSWT in Lake Vänern, however, due to the latency of ERA5-Land with 2-3 months it was not possible to conduct this step at this point, however, will be conducted at a later time.

Another limitation that came across was the difference in the spatial resolution between MODIS and ERA5-Land. Due to the coarse spatial resolution of ERA5-Land the generated LSWT product included pixels of the shore. This caused that the results needed to be cropped to a smaller extent excluding parts of Lake Vänern. The coarse resolution of ERA5-Land also makes this approach unsuitable for smaller lakes. The evaluation of the generated dataset could have been expanded by investigating the performance of the product over a longer period to include more in-situ measurements, however, due to long computing times the study period was limited to two years. Additionally, as previously mentioned the missing data for ERA5-Land for spring 2018 causes uncertainty in the accuracy of the results during that period.

For future studies, it is recommended to improve the spatial and temporal frequency of in-situ measurements for a better evaluation. Furthermore, the presented LSWT products should be tested to investigate if other LSWT products are able to improve the accuracy of the fusion method. Since the proposed bias correction method was not able to remove the systematic error completely, it is recommended to apply the bias correction method for each season or explore further bias correction methods. Last but least, the data fusion should be conducted for a longer period to investigate the impact of

external factors, such as the air temperature, on the spatial and temporal variation of LSWT.

7 Conclusion

This study aimed to develop a new spatially complete and daily continuous lake surface water temperature (LSWT) for Lake Vänern, Sweden by fusing satellite and reanalysis products. Firstly, 6 different LSWT products were evaluated against the in-situ measurements, MODIS, TIRS, GloboLakes, CGLOPS ARC-Lake and ERA5-Land. Results showed that the GloboLakes and the ATSR2 product performed the best in terms of MAE (0.56 °C and 0.32 °C), RMSE (0.79 °C and 0.46 °C) and R^2 (0.97 and 1). The accuracy of all other LSWT products was within an acceptable range (MAE: 0.75-1.32 °C, RMSE: 1.06-1.79 °C, R^2: 0.92–0.96). The MODIS product was identified as the most suitable satellite LSWT to be fused with the reanalysis product ERA5-Land because of its high temporal and spatial resolution as well as the ongoing data collection of the sensor. The results prior to bias-correction revealed that the applied data fusion method effectively filled data gaps and enhanced the accuracy of the MODIS LSWT product. Specifically, the MODIS dataset exhibited an MAE of 2.21 °C, RMSE of 3.64 °C, and R^2 of 0.41 °C, whereas the fused dataset demonstrated improved performance with an MAE of 1.49 °C, RMSE of 2.15 °C, and R^2 of 0.87. Notably, the spatial resolution of 1 km was successfully maintained throughout the data fusion process.

Furthermore, the fused LSWT underwent bias correction to address systematic bias, resulting in improved RMSE (2.03 °C)). However, the MAE increased from 1.49 °C to 1.57 °C. Overall, the applied data fusion and bias correction method successfully generates reasonable LSWT estimations and improved the accuracy of the LSWT data compared to MODIS and ERA5-Land. The evaluation of the corrected LSWT showed that it captured the temporal dynamics, in good agreement with the MODIS LSWT. When comparing the temporal variation to in-situ measurements, it was observed that the discrepancy was highest when clear pixels were unavailable. This suggests that the bias correction contributed to the improvement of the fused LSWT.

Finally, the spatial and temporal dynamics of LSWT in Lake Vänern were investigated. The results revealed that the LSWT was lower at the beginning of the year but experienced a significant increase in temperature from April onwards due to rising air temperatures. The LSWT reached its maximum in July/August followed by a linear decrease of the LSWT. The temporal dynamics were reflected in the spatial pattern of the LSWT. The pattern exhibits the development of a thermal bar, which can only be seen in the spring. The bar can be connected to the gradient in temperature, which is mainly

42

caused by the bathymetry. As the LSWT rises, the thermal bar moved towards the center of the lake and finally disappeared once the LSWT exceeds 4 °C.

Overall, this study confirms the suitability of the existing LSWT products and demonstrates the good performance of the data fusion method in generating a spatially complete and temporal continuous LSWT dataset for Lake Vänern. This method and generated new dataset are valuable for LSWT monitoring and further investigation of ecological changes in lakes associated with shifting LSWT.

References

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingstone, D. M., Sommaruga, R., Straile, D., Van Donk, E., Weyhenmeyer, G. A., & Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, *54*, 2283-2297.

Avdan, U., & Jovanovska, G. (2016). Algorithm for Automated Mapping of Land Surface Temperature Using LANDSAT 8 Satellite Data. *Journal of Sensors*, *2016*, 1480307.

Carrea, L., & Merchant, C. (2020a). *Copernicus Global Land Operations Cyrosphere and Water. CGLOPS2. Quality Assessment Report. Lake Surface Water Temperature 1km Products.* I1.09).

Carrea, L., & Merchant, C. (2020b). *Copernicus Global Land Operations Cyrosphere and Water. CGLOPS2. Product User Manual. Lake Surface Water Temperature. 1km Products* I1.09).

Carrea, L., & Merchant, C. J. (2019). *GloboLakes: Lake Surface Water Temperature (LSWT)* Version 4.0).

Chavula, G., Brezonik, P., Thenkabail, P., Johnson, T., & Bauer, M. (2009). Estimating the surface temperature of Lake Malawi using AVHRR and MODIS satellite imagery. *Physics and Chemistry of the Earth, Parts A/B/C*, *34*(13), 749-754.

Dong, S., Cheng, J., Shi, J., Shi, C., Sun, S., & Liu, W. (2022). A Data Fusion Method for Generating Hourly Seamless Land Surface Temperature from Himawari-8 AHI Data. *Remote Sensing*, *14*(20), 5170.

Dörnhöfer, K., & Oppelt, N. (2016). Remote sensing for lake research and monitoring – Recent advances. *Ecological Indicators*, *64*, 105-122.

Duan, S.-B., Li, Z.-L., & Leng, P. (2017). A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data. *Remote Sensing of Environment*, *195*, 107-117.

Duan, S.-B., Li, Z.-L., Li, H., Göttsche, F.-M., Wu, H., Zhao, W., Leng, P., Zhang, X., & Coll, C. (2019). Validation of Collection 6 MODIS land surface temperature product using in situ measurements. *Remote Sensing of Environment*, *225*.

Duan, S.-B., Li, Z.-L., Zhao, W., Wu, P., Huang, C., Han, X.-J., Gao, M., Leng, P., & Shang, G. (2021). Validation of Landsat land surface temperature product in the conterminous United States using in situ measurements from SURFRAD, ARM, and NDBC sites. *International Journal of Digital Earth*, *14*(5), 640-660.

Dyba, K., Ermida, S., Ptak, M., Piekarczyk, J., & Sojka, M. (2022). Evaluation of Methods for Estimating Lake Surface Water Temperature Using Landsat 8. *Remote Sensing*, *14*(15), 3839.

ECMWF, E. C. f. M. R. W. F. (2023). *ERA5-Land: data documentation*. Retrieved 15.03.2023 from https://confluence.ecmwf.int/display/CKB/ERA5-Land%3A+data+documentation

Emery, W., & Camps, A. (2017). *Introduction to satellite remote sensing : atmosphere, ocean, land and cryosphere applications*.

EoPortal. (2012). *ERS-2 (European Remote-Sensing Satellite-2)*. European Space Agency. Retrieved 5th of May 2023 from https://www.eoportal.org/satellite-missions/ers-2#ers-2-european-remote-sensing-satellite-2

ESA. (n.d.-a). *Envisat instruments*. Retrieved 5th of May 2023 from
https://earth.esa.int/eogateway/missions/envisat#instruments-section

ESA. (n.d.-b). *ERS-2*. Retrieved 5th of May 2023 from
https://www.esa.int/Enabling_Support/Operations/ERS-2

ESA. (n.d.-c). *Sentinel-3 - Facts and Figures*. Retrieved 5th of May 2023 from
https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-3/Facts_and_figures

Eumetsat. (n.d.). *MetOP Series*. Retrieved 15.03.2023 from
https://www.eumetsat.int/our-satellites/metop-series

Feng, G., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat
and MODIS surface reflectance: predicting daily Landsat surface reflectance.
IEEE Transactions on Geoscience and Remote Sensing, *44*(8), 2207-2218.

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the Future of Land Surface
Models and the Challenges of Representing Complex Terrestrial Systems.
Journal of Advances in Modeling Earth Systems, *12*(4), e2018MS001453.

Ihlen, V. (2019). *Landsat 8 (L8) Data Users Handbook*.
https://www.usgs.gov/landsat-missions/landsat-8-data-users-handbook

Kvarnäs, H. (2001). Morphometry and Hydrology of the Four Large Lakes of
Sweden. *AMBIO: A Journal of the Human Environment*, *30*(8), 467-474, 468.

Lazhu, Yang, K., Qin, J., Hou, J., Lei, Y., Wang, J., Huang, A., Chen, Y., Ding, B., &
Li, X. (2022). A Strict Validation of MODIS Lake Surface Water Temperature
on the Tibetan Plateau. *Remote Sensing*, *14*(21), 5454.

Li, Z.-L., Wu, H., Duan, S.-B., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X.,
Zhu, J., Sun, Y., Si, M., Liu, M., Li, J., Zhang, X., Shang, G., Tang, B.-H.,
Yan, G., & Zhou, C. (2023). Satellite Remote Sensing of Global Land Surface
Temperature: Definition, Methods, Products, and Applications. *Reviews of
Geophysics*, *61*(1), e2022RG000777.

Llewellyn-Jones, D., Edwards, M. C., Mutlow, C. T., Birks, A. R., Barton, I. J., &
Tait, H. (2001). AATSR: Global-change and surface-temperature
measurements from Envisat [Article]. *ESA Bulletin*, *105*, 11-21.

Long, D., Yan, L., Bai, L., Zhang, C., Li, X., Lei, H., Yang, H., Tian, F., Zeng, C.,
Meng, X., & Shi, C. (2020). Generation of MODIS-like land surface
temperatures under all-weather conditions based on a data fusion approach.
Remote Sensing of Environment, *246*, 111863.

MacCallum, S., & Merchant, C. (2010). *ATSR Reprocessing for Climate Lake Surface
Temperature: ARC-Lake Validation Report - v.1.0* 1).

MacCallum, S. N., & Merchant, C. J. (2012). Surface water temperature observations
of large lakes by optimal estimation. *Canadian Journal of Remote Sensing*,
38(1), 25-45.

MacCallum, S. N., Merchant, C. J., & Layden, A. (2013). Global lake surface water
temperatures from ATSR. EGU General Assembly Conference Abstracts,

Martin, S. (2014). *An Introduction to Ocean Remote Sensing* (2 ed.).

Merchant, C., & Maccallum, S. (2018a). *Lake Surface Water Temperature ARC-Lake
v3 (1995-2012). Data Product Description*.

Moukomla, S., & Blanken, P. D. (2016). Remote Sensing of the North American
Laurentian Great Lakes' Surface Temperature. *Remote Sensing*, *8*(4), 286.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G.,
Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens,
B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E.,
Buontempo, C., & Thépaut, J. N. (2021). ERA5-Land: a state-of-the-art global

reanalysis dataset for land applications. *Earth Syst. Sci. Data*, *13*(9), 4349-4383.

NASA. (2023). *Terra - Terra Instruments*. Retrieved 5th of May 2023 from https://terra.nasa.gov/about/terra-instruments

Pal, S., & Sharma, P. (2021). A Review of Machine Learning Applications in Land Surface Modeling. *Earth*, *2*(1), 174-190.

Read, J. M., & Torrado, M. (2009). Remote Sensing. In R. Kitchin & N. Thrift (Eds.), *International Encyclopedia of Human Geography* (pp. 335-346). Elsevier.

Reinart, A., & Reinhold, M. (2008). Mapping surface temperature in large lakes with MODIS data. *Remote Sensing of Environment*, *112*(2), 603-611.

Sayler, K. (2023). *Landsat 8-9 Collection 2 (C2) Level 2 Science Product (L2SP). Guide*. https://www.usgs.gov/media/files/landsat-8-9-collection-2-level-2-science-product-guide

Schmitt, M., & Zhu, X. X. (2016). Data Fusion and Remote Sensing: An ever-growing relationship. *IEEE Geoscience and Remote Sensing Magazine*, *4*(4), 6-23.

Schneider, P., Hook, S. J., Radocinski, R. G., Corlett, G. K., Hulley, G. C., Schladow, S. G., & Steissberg, T. E. (2009). Satellite observations indicate rapid warming trend for lakes in California and Nevada. *Geophysical Research Letters*, *36*(22).

Schwab, D. J., Leshkevich, G. A., & Muhr, G. C. (1999). Automated Mapping of Surface Water Temperature in the Great Lakes. *Journal of Great Lakes Research*, *25*(3), 468-481.

Sharaf, N., Fadel, A., Bresciani, M., Giardino, C., Lemaire, B., Slim, K., Faour, G., & Vinçon-Leite, B. (2019). Lake surface temperature retrieval from Landsat-8 and retrospective analysis in Karaoun Reservoir, Lebanon. *Journal of Applied Remote Sensing*, *13*(4), 044505.

Sima, S., Ahmadalipour, A., & Tajrishy, M. (2013). Mapping surface temperature in a hyper-saline lake and investigating the effect of temperature distribution on the lake evaporation. *Remote Sensing of Environment*, *136*, 374-385.

Stefanidis, K., Varlas, G., Papaioannou, G., Papadopoulos, A., & Dimitriou, E. (2022). Trends of lake temperature, mixing depth and ice cover thickness of European lakes during the last four decades [Article]. *Science of the Total Environment*, *830*, Article 154709.

Wan, Z. (2006). *MODIS Land Surface Temperature Users' Guide*. https://lpdaac.usgs.gov/documents/447/MOD11_User_Guide_V4.pdf

Wan, Z., Hook, S., & Hulley, G. (2021). *MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1 km SIN Grid* Version 061).

Wan, Z., Zhang, Y., Zhang, Q., & Li, Z.-L. (2004). Quality Assessment and Validation of the MODIS Global Land Surface Temperature. *International Journal of Remote Sensing - INT J REMOTE SENS*, *25*.

Woolway, R. I., Kraemer, B. M., Lenters, J. D., Merchant, C. J., O'Reilly, C. M., & Sharma, S. (2020). Global lake responses to climate change. *Nature Reviews Earth & Environment*, *1*(8), 388-403.

Xiong, X., Butler, J., Cao, C., & Wu, X. (2018). 1.13 - Optical Sensors—VIS/NIR/SWIR. In S. Liang (Ed.), *Comprehensive Remote Sensing* (pp. 353-375). Elsevier.

Yilmaz, M. (2023). Accuracy assessment of temperature trends from ERA5 and ERA5-Land. *Science of the Total Environment*, *856*, 159182.

Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., & Hong, Y. (2018). A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 30-45.

Zhang, J. (2010). Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, *1*(1), 5-24.

Zhang, R., Chan, S., Bindlish, R., & Lakshmi, V. (2021). Evaluation of global surface water temperature data sets for use in passive remote sensing of soil moisture [Article]. *Remote Sensing*, *13*(10), Article 1872.

Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, *114*(11), 2610-2623.

Zhu, X., Helmer, E. H., Gao, F., Liu, D., Chen, J., & Lefsky, M. A. (2016). A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sensing of Environment*, *172*, 165-177.