

Freie Universität



Berlin

**Trends in surface temperature from new long-term
homogenized thermal data by applying remote sensing
techniques and its validation using *in-situ* data of five
southern European lakes**

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)
submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Sajid Pareeth

Berlin

July 21, 2016

Reviewer 1:

Dr. Rita Adrian

Head of Ecosystem Research Department, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany

Reviewer 2:

Prof. Dr. Klement Tockner

Director, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany.
Professor of Aquatic Ecology, Freie Universität, Berlin.

Date of Disputation: July 12, 2016

Acknowledgements

First and foremost, thank you to my wife Meera, for all her love and endless support. Love to my little boy Rayan who cheered up my stressful days.

I would like to extend sincere gratitude to my supervisors at FEM-CRI, Markus Neteler and Nico Salmaso for their constant support and critical reviews, which is the foundation of this thesis. I am grateful to Markus Neteler for all the discussions, his tireless help in data processing and resolving cluster issues at critical stages. I thank Nico Salmaso for guiding me through the thesis with endless meetings and supporting me with statistical approaches and getting collaborators in place. I thank them for their patience in listening to me and providing excellent support to fulfill my research requirements. I also thank my supervisor at the university, Rita Adrian for her continuous support and critical reviews.

I want to thank my colleagues in exPGIS group – Luca Deluchi, Duccio Rocchini, Roberto Zorer, Markus Metz, Matteo Marcantonio, Veronica Andreo, Carol Garzon, Stepan Turek for creating a cheerful working environment, all the support and discussions. I am especially grateful to Luca Deluchi for all his help with programming. I want to thank my colleagues in Idrobiologia group in FEM-CRI – Adriano Boscaini, Leonardo Cerasino, Monica Tolotti, Andrea Zampedri for the support during sampling and other trips. I also thank Francesca Cagnacci for supporting me to obtain IRSAE funds. I thank Paolo Lenti for his help in solving cluster issues. I would like to express my gratitude to our collaborators – Alessandro Ludovisi, Barbara Leoni, Giuseppe Morabito, Fabio Buzzi, Fabio Lepori for sharing the data and critical reviews. I would like to thank all my friends at FEM-CRI for a wonderful atmosphere and support during the last three years in Trento as well as for their input on specific scientific questions.

I wish to thank Abhay Devasthale and Martin Raspaud at SMHI for their support with Pytroll development and discussions. I sincerely thank Juan-Carlos Jimenez-Muñoz of University of Valencia, for sharing with me the split-window coefficients. I express my sincere gratitude to the Open Source Software community, for their wonderful contributions and tireless support. My special thanks to the developers and user communities of GRASS GIS, R, Pytroll and Orfeo Toolbox. I want to thank them for their patience in understanding my problems and for helping me to resolve the issues. I wish to thank all the data providers specifically, NOAA for making AVHRR LAC data available through CLASS archive, NASA for MODIS swath data through LAADS archive, ESA for ATSR series of data through MERCI archive. I want to thank EUMETSAT for providing ERA-interim reanalysis data.

I would like to express my gratitude for the administrative and financial support provided by Fondazione Edmund Mach institute through my PhD scholarship under the FIRS> T (FEM International Research School e Trentino) program. I thank Alessandro Gretter, Elisabetta Perini and Flavia Zanon in FEM for their help and support. I wish to thank IRSAE network for granting financial support which allowed me to attend conferences and conduct research visits.

Finally, I wish to thank my parents for their unconditional support and love. I thank my sister for her constant motivation and faith in me. I want to thank my brother Hafeel, for his support, encouragement and paving the way in research. My love to dearest Abeni and Rumi. I also express sincere thanks to my extended family for their support and love.

List of Abbreviations

AATSR: Advanced Along-Track Scanning Radiometer
AOT: Aerosol Optical Thickness
ATSR: Along Track Scanning Radiometer
AVHRR: Advanced Very High Resolution Radiometer
BT: Brightness Temperature
C: Celsius
CDR: Climate Data Records
CLASS: Comprehensive Large Array-data Stewardship System
DN: Digital Number
DTC: Diurnal Temperature Cycle
EA: Eastern Atlantic
ECMWF: European Center for Medium-range Weather Forecasts
ECV: Essential Climate Variables
ERS: European Remote Sensing
GCP: Ground Control Point
GIS: Geographic Information System
GLTC: Global Lake Temperature Collaboration
GLEON: Global Lake Ecological Observatory Network
GPL: General Public License
GRASS: Geographic Resources Analysis Support System
HANTS: Harmonic ANalysis of Time Series
ICT: Information and Communication Technologies
K: Kelvin
LAC: Local Area Coverage
LST: Land Surface Temperature
LSWT: Lake Surface Water Temperature
MAE: Mean Absolute Error
MODIS: Moderate Resolution Imaging Spectroradiometer
NAO: North Atlantic Oscillation
NIR: Near Infrared
NOAA: National Oceanic and Atmospheric Administration
POD: Polar Orbiter Data
RMSE: Root Mean Square Error
SIFT: Scale Invariant Feature Transform
SPARC: Separation of Pixels Using Aggregated Rating over Canada
SST: Sea Surface Temperature
ST: Surface Temperature
SW: Split-Window
SZA Solar Zenith Angle
TIR: Thermal Infra-red Region
TOA: Top Of Atmosphere
UTC: Coordinated Universal Time
VZA: View Zenith Angle

Summary

Recent studies, based on a combination of long-term *in-situ* and satellite derived temperature data indicate that lakes are rapidly warming at the global scale. Since Lake Surface Water Temperature (LSWT) is highly responsive to long-term modifications in the thermal structure of lakes, it is a good indicator of changes in lake characteristics. There have not been done many studies at a regional scale to understand the lakes' response to climate change, mainly due to lack of high spatio-temporal data. Therefore, further studies are needed to understand variation in trends, impacts and consequences at a regional scale. It is essential to have highly frequent spatially explicit data to understand the spatio-temporal thermal variations of LSWT. Continuous *in-situ* water temperature data measured at high temporal resolution from permanently installed stations are becoming increasingly available through GLEON (Global Lake Ecological Observatory Network; <http://gleon.org/>) or NetLake (Networking Lake Observatories in Europe). But these data are often heterogeneous with different sources and time line, point based, and not available for many lakes around the globe. To establish permanent weather stations for all the large lakes in the world is also not economically viable. As an alternative to direct measurements, remote sensing is considered as a promising approach to reconstruct complete time series of LSWT where direct measurements are missing. Temperature of land/water surfaces is one of the direct and accurate measurements using satellite data acquired in the thermal infra-red spectral region. Furthermore, the availability of daily satellite data since the 1980s at a moderate resolution of 1 km from multiple polar orbiting satellites is an opportunity not to be missed. But owing to the complexities related to earlier satellite missions, and the need of high level of processing, the potential of the historical satellite data in deriving a homogenised LSWT is still not explored well.

There is a gap in the availability of long-term time series of LSWT from the satellite data which could be used in understanding the patterns and drivers of thermal variations in large lakes. This thesis aims to fill this gap by developing reproducible and extendable methods to derive homogenised daily LSWT for thirty years from 1986 to 2015. Hence, the main objectives of this thesis are i) to reconstruct thirty years (1986-2015) of daily satellite thermal data as a homogenised time series of LSWT for five large Italian lakes by combining thermal data from multiple satellites, ii) to assess the quality of the satellite derived LSWT using long-term *in-situ* data collected from the same lakes, iii) to report the seasonal and annual trends in LSWT using robust statistical tests.

The first part of the thesis deals with the accurate processing of historical Advanced Along-Track Scanning Radiometer (AVHRR) sensor data to derive time series of LSWT. A new method to resolve the complex geometrical issues with the earlier AVHRR data obtained from National Oceanic and Atmospheric Administration (NOAA) satellites has been developed. The new method can accurately process historical AVHRR data and develop time series of geometrically aligned thermal channels in the spectral range of 10.5-12.5 μm . The validation procedure to check the accuracy of image to image co-registration using 2000 random images (from a total of 22,507 images) reported sub-pixel accuracy with an overall Root Mean Square Error (RMSE) of 755.65 m. The usability of newly derived time series of thermal channels to derive LSWT for lakes were tested and validated. Furthermore, cross-platform and inter-platform validations were performed using corresponding same day observations which reported an overall RMSE of less than 1.5 °C.

In the second part of the thesis, a new method was developed to derive homogenised daily LSWT standardized at 12:00 UTC from thermal channels of thirteen different satellites. The new method is implemented for Lake Garda in Northern Italy developing time series of homogenised daily LSWT for last thirty years from 1986 to 2015. The sensors used in this study are the AVHRR from multiple NOAA satellites, Along Track Scanning Radiometer (ATSR) series from European Remote Sensing (ERS) satellites and Moderate Resolution Imaging Spectroradiometer (MODIS) from Aqua and Terra satellites. The LSWT time series are then validated using long-term *in-situ* data obtained from a deep and a shallow sampling location in the lake. Validation of LSWT from individual satellites against corresponding *in-situ* data reported an overall RMSE of 0.92 °C. The validation between final homogenised LSWT and the *in-situ* data reported a coefficient of determination (R^2) of 0.98 and a RMSE of 0.79 °C.

In the third part of the thesis, homogenised daily LSWT for the last thirty years (1986-2015) were developed for five large lakes in Italy using the newly developed methods. The LSWT time series was validated against the *in-situ* data collected from the respective lakes. Furthermore, long-term trend analysis to study the seasonal and annual variations in LSWT over thirty years was performed over the newly developed LSWT data. The validation procedure reported an average RMSE and Mean Absolute Error (MAE) of 1.2 °C and 0.98 °C, respectively, over all the lakes. The trend analysis reported an overall regional summer warming rate of 0.03 °C yr⁻¹ and an annual warming rate of 0.017 °C yr⁻¹. During summer, all studied sub-Alpine lakes showed high coherence in LSWT to each other. The summer mean LSWT of Lake Garda, located in the sub-Alpine region also exhibit high temporal coherence with that of central Italian Lake Trasimeno. Annually, mean LSWT of all sub-Alpine lakes were found to be highly coherent to each other, while mean LSWT of Lake Trasimeno resulted less coherent to the other lakes.

Overall, the thesis aims at contributing to the accurate processing of the various historical satellite data and the development of a new method which allows to merge them into a unified, longest possible time series of LSWT. The newly developed methods used open source geospatial software tools, which ensure the reproducibility and also extensibility to any other geographic location given the availability of satellite data. Although this study is using LSWT as the primary physical variable, the developed methods can be used to derive any other time series of land and water based regional products from satellite data.

Zusammenfassung

Jüngste Studien, die auf einer Kombination von langfristigen *in-situ*- und satellitenabgeleiteten Temperaturdaten basieren, zeigen, dass Seen auf globaler Skala sich schnell erwärmen. Da die Oberflächentemperatur von Seen (Lake Surface Water Temperature, LSWT) stark auf langfristige Veränderungen in der thermischen Struktur der Seen reagieren, sind sie ein geeigneter Indikator für Veränderungen in den Eigenschaften der Seen. Auf regionaler Skala wurden bisher nicht viele Studien durchgeführt, um den Einfluss des Klimawandels auf Seen zu verstehen, vor allem wegen des Mangels an hochauflösenden raum-zeitlichen Daten. Daher sind weitere Studien erforderlich, um Trendvariationen, Auswirkungen und Folgen auf regionaler Ebene zu verstehen. Es ist unumgänglich, über zeitlich hochauflösende und räumlich explizite Daten zu verfügen, um die räumlich-zeitlichen Schwankungen der Oberflächentemperatur von Seen (LSWT) zu verstehen. Kontinuierliche *in-situ*-Daten von Wassertemperatur mit hoher zeitlicher Auflösung, die von fest installierten Stationen gemessen werden, stehen zunehmend in größerer Anzahl durch GLEON (Global Lake Ecological Observatory Network, <http://gleon.org>) oder NetLake (Networking Lake Observatories in Europe) zur Verfügung. Aber diese Daten sind oft heterogen, da sie aus verschiedenen Quellen stammen und unterschiedliche Zeithorizonte umfassen, auch sind sie nur punktuell und stehen nicht für viele Seen rund um den Globus zur Verfügung. Allerdings ist es ökonomisch betrachtet nicht realistisch, permanente Wetterstationen für alle großen Seen der Welt aufzustellen. Ein vielversprechender Ansatz alternativ dazu ist die Fernerkundung, die dort erlaubt, wo direkte Messungen fehlen, vollständige Zeitreihen der Oberflächentemperatur von Seen zu rekonstruieren. Die Oberflächentemperatur von Land- und Wasserflächen ist eine der direkten und genauesten Messungen, die mit Hilfe von Satellitendaten im thermischen Infrarot-Spektralbereich durchgeführt werden kann. Darüber hinaus bieten die täglichen Satellitendaten einer Reihe polar-umlaufender Satelliten mit einer moderaten Auflösung von 1 km seit den 1980er Jahren eine Möglichkeit, die genutzt werden sollte. Aufgrund der komplexen Probleme bei der Datenverarbeitung früherer Satellitenmissionen wurde jedoch das Potential im Zusammenhang mit diesen historischen Satellitendaten zur Erstellung einer homogenisierten LSWT-Zeitreihe bislang nicht gut genutzt.

Es gibt eine Lücke in der Verfügbarkeit von langfristigen LSWT-Zeitreihen aus Satellitendaten, die für das Verständnis von Mustern und Auslösern von thermischen Veränderungen in großen Seen dienen könnten. Diese Arbeit zielt darauf ab, diese Lücke zu füllen, indem sie reproduzierbare und erweiterbare Methoden zur Ableitung einer dreißigjährigen Zeitreihe (1986 bis 2015) von homogenisierten täglichen LSWT-Werten vorstellt. Die wichtigsten Ziele dieser Arbeit sind i) 30 Jahre (1986-2015) täglicher, thermischer Satellitendaten von mehreren Satelliten als homogenisierte LSWT-Zeitreihe für fünf große italienische Seen zu rekonstruieren, ii) die Qualität der aus Satellitendaten abgeleiteten LSWT-Daten unter Verwendung von langfristigen *in-situ*-Daten aus den gleichen Seen zu beurteilen, und iii) die saisonalen und jährlichen LSWT-Trends mit statistischen Tests zu ermitteln.

Der erste Teil der Arbeit beschäftigt sich mit der genauen Verarbeitung von historischen Sensordaten des Advanced Along-Track Scanning Radiometers (AVHRR), um LSWT-Zeitreihen abzuleiten. Es wurde eine neue Methode entwickelt, die die komplexen geometrischen Probleme, die in den früheren AVHRR Sensordaten (von der National Oceanic and Atmospheric Administration (NOAA) bereitgestellt) enthalten sind, lösen kann. Die neue Methode kann historische AVHRR Daten exakt verarbeiten und Zeitreihen von geometrisch ausgerichteten thermischen Kanälen im Spektralbereich von 10,5-12,5 μm prozessieren.

Das Validierungsverfahren zur Prüfung der Genauigkeit der Bild-zu-Bild-Ausrichtung wurde mit 2000 (von insgesamt 22.507 Bildern) zufällig ausgewählten Bildern durchgeführt und ergab eine Subpixel-Genauigkeit mit einer mittleren quadratischen Abweichung (Root Mean Square Error, RMSE) von 755,65 m. Die Nutzbarkeit der von thermischen Kanälen neu abgeleiteten LSWT-Zeitreihen wurde getestet und validiert. Des Weiteren wurden plattformübergreifende Validierungen unter Verwendung von am gleichen Tag gemessenen Beobachtungen durchgeführt, die einen Gesamt-RMSE von weniger als 1,5 °C ergaben.

Im zweiten Teil der Arbeit wurde eine neue Methode entwickelt, um homogenisierte, tägliche, auf 12:00 Uhr UTC bezogene LSWT-Daten aus thermischen Kanälen von dreizehn verschiedenen Satelliten standardisiert abzuleiten. Die neue Methode wurde für den Gardasee in Norditalien umgesetzt, wobei die entwickelte homogenisierte LSWT-Zeitreihe mit täglicher Auflösung die letzten dreißig Jahre von 1986 bis 2015 umfasst. Als Sensoren wurden in dieser Studie sowohl das AVHRR (auf mehreren NOAA-Satelliten) sowie das Along Track Scanning Radiometer (ATSR, auf European Remote Sensing (ERS) Satelliten) und das Moderate Resolution Imaging Spectroradiometer (MODIS, auf Aqua und Terra-Satelliten) verwendet. Die LSWT-Zeitreihen wurden an einer tiefen und an einer flachen Probenahmestelle des Garda-Sees mit langfristigen *in-situ*-Daten validiert. Die Validierung von LSWT-Werten einzelner Satelliten in Bezug auf die *in-situ*-Daten ergab einen Gesamt-RMSE von 0,92 °C. Die Validierung zwischen den endgültigen homogenisierten LSWT und den *in-situ*-Daten ergab ein Bestimmtheitsmaß (R^2) von 0,98 und einen RMSE von 0,79 °C.

Im dritten Teil der Arbeit wurden tägliche LSWT für die letzten dreißig Jahren homogenisiert (1986-2015) und diese Zeitreihen für fünf große Seen in Italien mit Hilfe der neu entwickelten Methoden erstellt. Die LSWT-Zeitreihe wurde mit *in-situ*-Daten der jeweiligen Seen validiert. Darüber hinaus wurde eine langfristige Trendanalyse durchgeführt, die die saisonalen und jährlichen Schwankungen in LSWT von mehr als dreißig Jahren in den neu entwickelten LSWT Daten analysiert. Das Validierungsverfahren ergab einen mittleren RMSE von 1,2 °C bzw. einen mittleren absoluten Fehler (MAE) von 0,98 °C für alle Seen. Die Trendanalyse ergab eine regionale Sommer-Erwärmungsrate von 0,03 °C yr⁻¹ und eine jährliche Erwärmungsrate von 0,017 °C yr⁻¹. Alle subalpinen Seen der Studie zeigten für die Sommermonate eine hohe Kohärenz in Bezug auf LSWT. Der Sommer-LSWT des Gardasees wies in der subalpinen Region auch mit dem mittellitalienischen Trasimenischen See eine hohe zeitliche Kohärenz auf. Der jährliche-LSWT aller subalpinen Seen wies ebenfalls eine große Kohärenz auf, während der mittlere LSWT des Trasimenischen Sees weniger kohärent im Vergleich zu den anderen Seen war.

Insgesamt zielt die Arbeit auf eine genaue Bearbeitung verschiedener historischer Satellitendaten ab und will zur Entwicklung eines neuen Verfahrens beitragen, die sie zu einer einheitlichen, möglichst langen Zeitreihe von LSWT vereinen. Die neu entwickelten Methoden verwenden Open Source GIS-Software-Tools, die die Reproduzierbarkeit und auch Erweiterbarkeit auf andere geographische Regionen ermöglichen, sofern Satellitendaten verfügbar sind. Obwohl diese Studie LSWT als primäre physikalische Größe verwendet, können diese Methoden verwendet werden, um auch andere Land- und Wasser-basierte regionale Zeitreihen-Produkte aus Satellitendaten abzuleiten.

Thesis outline

This thesis is structured as a cumulative work with three manuscripts presented as Chapters 2, 3 and 4. Out of three, one has already been published as a peer reviewed article (Chapter 2), second one has been accepted and in press (Chapter 3) while the third one is submitted to a journal (Chapter 4). Each of these chapters has its own abstract, introduction, methods, results and conclusions. Additionally, two further chapters are included: the Chapter 1 giving a general introduction to the entire work and Chapter 5 with a general discussion summarizing the overall work pursued. The references are listed at the end of the thesis. The layout of manuscripts have been modified with respect to the publishers' article styles in order to maintain a homogeneous thesis design.

Details concerning the article manuscripts:

Chapter 2

Pareeth, S., Delucchi, L., Metz, M., Rocchini, D., Devasthale, A., Raspaud, M., Adrian, R., Salmaso, N., Neteler, M., 2016. *New automated method to develop geometrically corrected time series of brightness temperatures from historical AVHRR LAC data*. Remote Sensing 8, 169. doi:10.3390/rs8030169. Reprinted by permission from MDPI AG (Basel, Switzerland)

Author contributions

SP designed the study, processed and analyzed the data and wrote the manuscript. LD, MM, AD and MR co-developed the tools required for the study and contributed to the final draft. DR contributed to the data analysis and the manuscript. RA advised on the preparation of manuscript and contributed to the text. MN and NS co-designed the study, advised on the preparation of manuscript and contributed to the text.

Chapter 3

Pareeth, S., Salmaso, N., Adrian, R., Neteler, M. (in press), *Homogenised daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-Alpine lake*. Sci. Rep. 6, 31251; doi:10.1038/srep31251 (2016). Reprinted by permission from Macmillan Publishers Ltd: SCIENTIFIC REPORTS, copyright (2016)

Author contributions

SP designed the study, processed and analyzed the data and wrote the manuscript. RA advised on the preparation of the manuscript and contributed to the text. MN and NS co-designed the study, advised on the preparation of manuscript and contributed to the text.

Chapter 4

Pareeth, S., Bresciani, M., Buzzi, F., Leoni, B., Lepori, F., Ludovisi, A., Morabito, G., Adrian, R., Neteler M., Salmaso N. (Submitted), *Warming trends of perialpine lakes from homogenised time series of historical satellite and in-situ data*.

Author contributions

SP designed the study, processed and analyzed the data and wrote the manuscript. RA advised on the preparation of manuscript and contributed to the text. MB, FB, BL, FL, AL and GM provided the *in-situ* data required for the validation and contributed to the text. NS and MN co-designed the study, advised on the preparation of manuscript and contributed to the text.

Contents

Acknowledgements	v
List of abbreviations	vii
Summary	ix
Zusammenfassung	xi
Thesis outline	xiii
1 General Introduction	1
1.1 Background	1
1.1.1 Thermal remote sensing	2
1.1.2 Global warming of lakes in the last decades	8
1.2 Thesis rationale	9
1.3 Research objectives	11
2 New automated method to develop geometrically corrected time series of brightness temperatures from historical AVHRR LAC data	13
2.1 Abstract	13
2.2 Introduction	14
2.3 Methods	17
2.3.1 Data and software	17
2.3.2 Pre-processing	19
2.3.3 Precise geo-rectification using SIFT	23
2.3.4 Validation of thermal calibration by estimating LSWT	24
2.4 Results	26
2.4.1 Time series of calibrated AVHRR LAC data	26
2.4.2 Geo-rectification of calibrated AVHRR LAC data using SIFT	27

2.4.3	Quality assessment of the time series using estimated LSWT	29
2.4.4	Software code for deploying the method elsewhere	31
2.5	Discussion	32
2.6	Conclusion	35
3	Homogenised daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-Alpine lake	37
3.1	Abstract	37
3.2	Introduction	38
3.3	Study area	40
3.4	Methods	42
3.4.1	Data	42
3.4.2	Pre-processing	42
3.4.3	Deriving LSWT using split-window algorithm	44
3.4.4	Homogenising LSWT from multiple satellites	45
3.4.5	Gap filling using harmonic analysis	48
3.4.6	Temporal database and trend analysis	48
3.5	Results	49
3.5.1	Validation	50
3.5.2	Long term warming trends	51
3.6	Discussion	53
4	Warming trends of perialpine lakes from homogenised time series of historical satellite and <i>in-situ</i> data	59
4.1	Abstract	59
4.2	Introduction	60
4.3	Study area	61
4.4	Methods	64
4.4.1	Data	64
4.4.2	Homogenized LSWT from satellite data	64
4.4.3	Validation, trends and regional coherence	65
4.5	Results	67
4.6	Discussion	73
4.7	Conclusion	75

5 General Discussion	77
5.1 Summary	77
5.1.1 Resolving important geometrical issues of AVHRR data	78
5.1.2 Homogenisation of LSWT	80
5.1.3 Warming of large lakes in Italy	81
5.2 Conclusion and future perspectives	83
List of Figures	85
List of Tables	89
Appendix	91
A.1 Python script to read and calibrate level 1B AVHRR LAC data using Pytroll libraries	91
A.2 Shell script to process the level 1B AVHRR LAC data - entire workflow	94
Bibliography	101

Chapter 1

General Introduction

1.1 Background

In the last few decades, innovation in the Information and Communication Technologies (ICT) has revolutionized the research paradigm with access to an “unlimited” amount of data. Satellite remote sensing is one of the domains which is growing rapidly with an increasing number of satellites on Earth’s orbit which are monitoring land and water (*Campbell and Wynne, 2011*). This enabled researchers to formulate their research making use of these data to monitor Earth’s surface understanding the patterns and processes on a particular snapshot of time or over a period of time (*Kerr and Ostrovsky, 2003; Lillesand et al., 2014*). The basic principle behind remote sensing is based on the measurement of reflected and emitted radiation from the Earth surface at different wavelengths in the electromagnetic spectrum (Fig. 1.1).

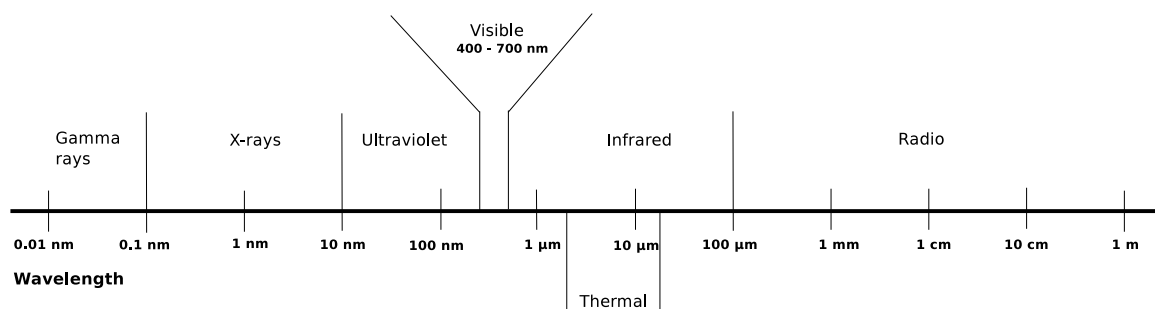


Figure 1.1: Electromagnetic spectrum. The wavelengths of visible and thermal region are shown separately.

The reflected energy from a single land cover/land use type has similar properties which form the basic principle behind identifying and studying different land use types using remote sensing. Satellite sensors record the reflected energy in the optical range of the electromagnetic spectrum while they record the emitted energy in the thermal region (*Lillesand et al., 2014*).

Moreover, the availability of satellite data over more than the last three decades has enabled researchers to use it for time series analyses, to understand changes over time (*Verbesselt et al.*, 2010). This has especially been used in studying land use, land cover changes over years, mapping physical and geographical changes in water bodies, and climate change studies. The time series of satellite data obtained from multiple polar orbiting and geo-stationary satellites are also used to develop Climate Data Records (CDR) with time scales of several hours, days, weeks, months and years, which are used extensively in climate change studies (*NOAA*, 2004). The CDR's are often grouped into atmospheric, oceanic and terrestrial records. In the atmospheric genre, the major Essential Climate Variables (ECV) developed using satellite data are Aerosol Optical Thickness (AOT), cloud properties, solar irradiance, precipitation etc. (*Ashouri et al.*, 2014). In the oceanic genre, the most common ECV's are ocean colour, sea salinity, sea surface temperature, sea ice etc. In the terrestrial genre there is surface reflectance, Leaf Area Index (LAI), Normalized Difference Vegetation Index (NDVI), Land Surface Temperature (LST) etc. An exhaustive list of various ECV's produced from combining satellite data and *in-situ* observation from global networks are given by NOAA (<https://www.ncdc.noaa.gov/cdr>).

1.1.1 Thermal remote sensing

Thermal remote sensing deals with acquisition, processing and interpretation of data recorded in the Thermal Infrared Region (TIR) (Fig. 1.1). The basic theory behind thermal remote sensing is that all objects with a temperature above absolute zero (0 K or -273.15 °C) emit electromagnetic radiation. The satellite sensors with thermal detectors can record this emitted TIR radiation at the long wavelengths ranging from 3 - 14 μm . Due to possible contamination by reflected sunlight in the 3 - 5 μm , the common approach for mapping temperature of the Earth's surface is by using the data acquired between 9 - 14 μm (*Kuenzler and Dech*, 2013; *Li et al.*, 2013).

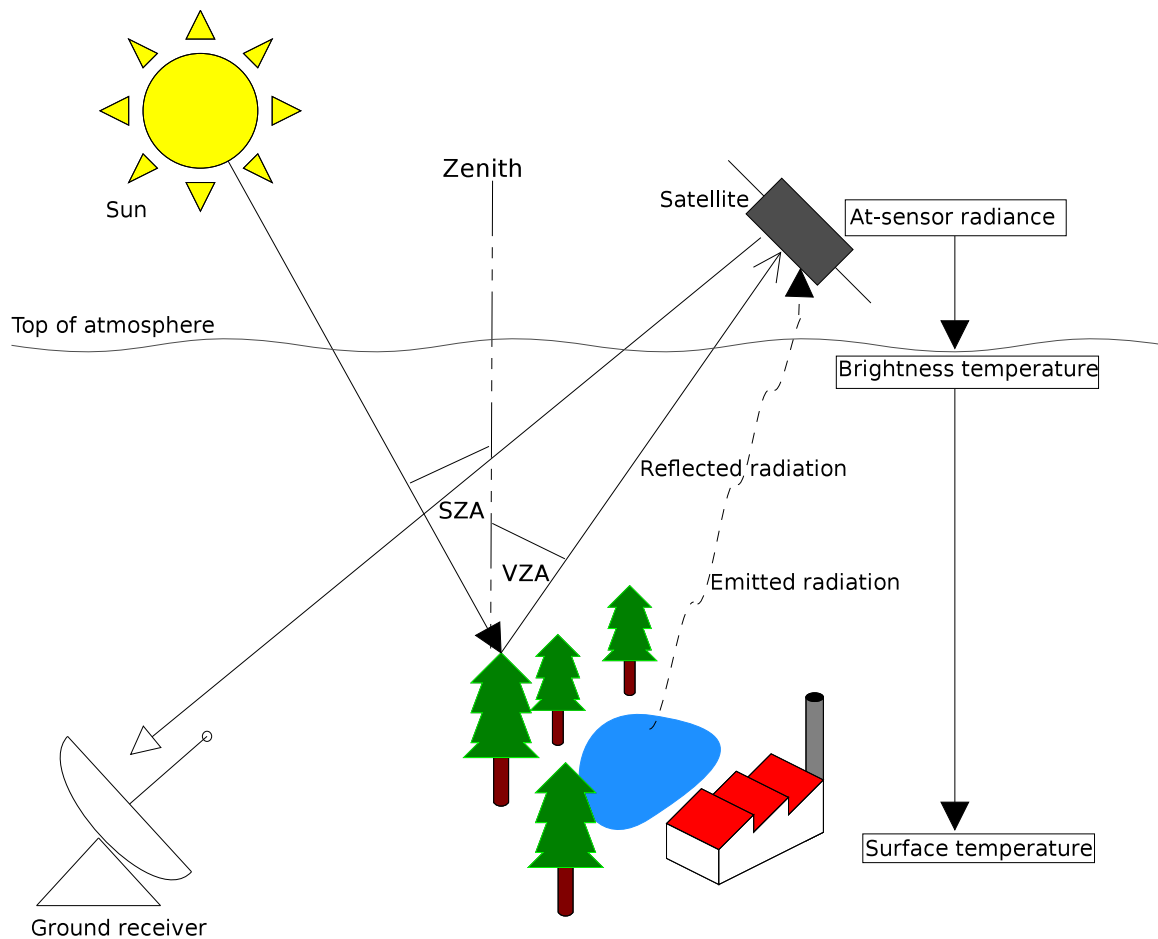


Figure 1.2: Principle of remote sensing. The emitted and reflected radiation from the Earth's surface is recorded by the satellite and sent to a ground receiver as radiances. VZA - Viewing Zenith Angle; SZA - Solar Zenith Angle

Many of the polar orbiting and geo-stationary satellites are equipped with sensors which can record TIR radiations in the 10 - 12 μm region. Thus the emitted radiation recorded by the thermal detectors in a sensor is scaled to integer values and stored as Digital Numbers (DN) in order to avoid precision issues in the pixel values. These data are then distributed to the user (Fig. 1.2).

From at-sensor radiance to TOA brightness temperature

The at-sensor radiance is a measure of reflected or emitted energy received by the radiometers in the satellite sensors. The DN is converted to at-sensor radiance using the calibration coefficients provided with the data. The relation between radiance emitted from a *Black Body*¹ and its absolute temperature is explained using Planck's law (Eq. 1.1).

¹*Black Body* is an ideal hypothetical object which emits equal amount of energy it absorbed, with emissivity equals one

$$I_{\lambda} = \frac{2hc^2}{\lambda^5(e^{hc/k\lambda T} - 1)} \quad (1.1)$$

where I_{λ} [$\text{W m}^{-2} \text{steradian}^{-1} \mu\text{m}^{-1}$] is the radiance measured in the wavelength λ [μm] from a body of absolute temperature T in Kelvin.

The constants h, c and k in the Eq. 1.1 are:

$h = 6.626 \times 10^{-34}$ J s; Planck's constant

$c = 2.9979246 \times 10^8$ m s⁻¹; speed of light

$k = 1.3806 \times 10^{-23}$ J K⁻¹; Boltzmann constant

Furthermore, by inverting the Planck's equation the absolute temperature of an object can be derived from the spectral radiance (Eq. 1.2).

$$T = \left(\frac{hc}{k\lambda}\right) \left(\frac{1}{\ln((2hc^2\lambda^{-5})/I_{\lambda} + 1)}\right) \quad (1.2)$$

In thermal remote sensing, I_{λ} represents the at-sensor spectral radiance measured for objects in Earth surface at wavelength λ of electro magnetic spectrum. Absolute temperature T represents the Top Of Atmosphere (TOA) Brightness Temperature (BT_{λ}). The final step is to derive the Surface Temperature (ST) from BT_{λ} after considering atmospheric effects and emissivity of different objects measured. The accuracy of the ST is based on how accurately we can quantify the atmospheric effects (in this case, mainly the water vapor column) at the time of image acquisition. Most objects emit less energy than the predicted radiance using Planck's law due to the emissivity of the object. Moreover, emissivity (range: 0 - 1) varies with wavelength and type of the object. The objects with high emissivity absorb and emit large amount of energy while those with low emissivity absorb and emit lower amount of energy. The emissivity related error depends upon the type of the object. According to *Kuenzer and Dech* (2013), for an object with a temperature of 288 K and an emissivity of 0.98, an emissivity change of 0.01 could result in a temperature difference of 0.73 K.

The Split-Window technique to obtain surface temperature from space

One of the most common method used in estimating ST from BT_{λ} is based on differential absorption method (*Quattrochi and Luvall*, 2004). The difference between radiances from two adjacent channels can be used to estimate the atmospheric contribution to the signal. This

purely empirically based methodology is well established in thermal remote sensing to derive surface temperature of different land use types. From the TIR region, dual thermal channels recorded in the 9 - 14 μm are used for estimating surface temperature using differential absorption method. This method is generally known as Split-Window (SW) technique. The dual thermal channels used in this method are termed (for this thesis) as T_i [10.5 - 11.5 μm] and T_j [11.5 - 12.5 μm]. In this study, we focus on satellite sensors which provide dual thermal channels T_i and T_j to estimate lake surface temperature. Table 1.1 lists the satellites with sensors having the dual thermal channels which can be used to estimate surface temperature using split window algorithm (used in this thesis).

Table 1.1: Important specifications of sensors with dual thermal channels in the 10.5-12.5 μm wavelength range. These are the satellites used in this study.

Sensor	Satellites	Spatial resolution	Time line	Revisit time (days)	Orbit height (km)	Wavelength(μm) of thermal channels (T_i , T_j)
AVHRR/2	NOAA-9/11/12/14	1.1 km	1986-2001	0.5	833	(10.3-11.3), (11.4-12.4)
AVHRR/3	NOAA-16/17/18/19	1.1 km	2001-2015	0.5	870	(10.3-11.3), (11.5-12.5)
ATSR1	ERS-1	1 km	1991-1997	3	785	(10.35-11.35), (11.5-12.5)
ATSR2	ERS-2	1 km	1995-2003	3	785	(10.35-11.35), (11.5-12.5)
AATSR	Envisat	1 km	2002-2012	3	783	(10.35-11.35), (11.5-12.5)
MODIS	Aqua, Terra	1 km	2000-2015	2	705	(10.78-11.28), (11.77-12.27)

The accurate estimate of surface temperatures using SW technique depends primarily on the split-window coefficients. These coefficients are often derived by regressing simulated BT's against atmospheric profiles obtained from radiative transfer models such as MODTRAN or by comparing satellite data with *in-situ* measurements (*Quattrochi and Luvall, 2004; Li et al., 2013*). Split-window coefficients represent linear or nonlinear functions of the emissivity, atmospheric water vapour and column and View Zenith Angle (VZA) of the satellite. These three parameters are considered critical in determining the accuracy of the surface temperature. The basic linear split-window equation (*McMillin, 1975; Li et al., 2013*) to derive ST from dual thermal channels is given in Eq. 1.3.

$$ST = c_0 + c_1 T_i + c_2 (T_i - T_j) \quad (1.3)$$

where c_0 , c_1 , c_2 are split window coefficients, T_i and T_j are TOA brightness temperatures derived from the dual thermal channels. The accuracy of the derived ST depends on the coefficients which in turn are estimated using atmospheric profiles derived from a radiative transfer model (*Li et al., 2013*). However, the linear approach produced large errors during hot and wet atmospheric conditions. To overcome this problem a nonlinear approach (Eq. 1.4) is introduced (*Walton, 1988*).

$$ST = c_0 + c_1T_i + c_2(T_i - T_j) + c_3(T_i - T_j)^2 \quad (1.4)$$

A combined approach is used in operational ST retrievals from MODIS and AVHRR data, where output from linear equation is used as a first estimate of ST in the actual nonlinear equation. There exist many variants of linear and nonlinear approaches in the literature which include additional parameters like emissivity, atmospheric water vapor and VZA explicitly in the SW equation. *Sobrino et al. (1993)* and *Coll and Caselles (1997)* introduced a nonlinear SW algorithm considering the emissivity and demonstrated less error compared to the ST estimated using basic nonlinear approach (Eq. 1.4). *Becker and Li (1995)* introduced cosine of VZA into the SW equation to model the angular dependencies of accurate ST retrieval. Furthermore, *Sobrino et al. (1991)*; *Becker and Li (1995)*; *Jimenez-Munoz and Sobrino (2008)* extended the split-window equation by including both emissivity and atmospheric water vapor factor in the ST retrieval. Eq. 1.5 is one such approach developed by *Jimenez-Munoz and Sobrino (2008)* which includes the corrections due to emissivity and atmospheric water vapour.

$$ST = T_i + c_1(T_i - T_j) + c_2(T_i - T_j)^2 + c_0 + (c_3 + c_4W)(1 - \epsilon) + (c_5 + c_6W)\Delta\epsilon \quad (1.5)$$

where $c_0 - c_6$ are split window coefficients, T_i and T_j are at-sensor brightness temperatures derived from the dual thermal channels, ϵ is mean emissivity, $\Delta\epsilon$ is emissivity difference and W is the total atmospheric water vapour column (g/cm^2). The SW equation with atmospheric water vapor and emissivity factors though proved to improve the accuracy, they are often difficult to implement due to lack of required data. An exhaustive review on existing algorithms for retrieving ST from space data is given in *Li et al. (2013)*. The two most important products derived from satellite thermal data using SW technique are Land Surface Temperature (LST) and Sea Surface Temperature (SST). Though the basic SW approach is similar in both cases the coefficients has to be parametrised depending on land and water.

Deriving Lake Surface Water Temperatures (LSWT)

For inland water bodies like large lakes, SST algorithms were commonly used in past studies. Studies by *Oesch et al. (2005)*; *Politi et al. (2012)* proved the usability of linear multichannel SST and nonlinear SST estimations from ATSR2 and AVHRR sensor data over large lakes in Europe. The multichannel linear approach reported higher accuracy than the nonlinear approach. *Crosman and Horel (2009)* used MODIS LST product (MOD11L2) to study the

spatial diurnal, seasonal and annual variations of water surface temperature of hypersaline Great Salt Lake in Utah. *Reinart and Reinhold* (2008) used MODIS SST products to study the temperature development over three years for the Lakes Vänern and Vättern in Sweden. On the other hand, *Hook et al.* (2003) used ATSR2 brightness temperatures from 11 and 12 μm to retrieve skin and bulk temperatures of Lake Tahoe by regressing them with *in-situ* data. In all the multiple approaches the RMSE obtained against *in-situ* data ranged from 0.5 - 1.5 $^{\circ}\text{C}$ reiterating the usability of satellite data in studying lake surface temperatures. There is a multitude of surface temperature products available in public which are often sensor specific and limited to a specific time frame. More recent study by *Hulley et al.* (2011) developed a list of lake and sensor specific coefficients to derive Lake Surface Water Temperature (LSWT) for 169 lakes globally. By applying these coefficients to SW equation with a VZA component, the derived LSWT exhibited lower RMSE compared to MODIS operational product. While this approach demands further data processing to derive LSWT using the published coefficients, *Riffler et al.* (2015) published long-term LSWT (1989-2013) for 25 European lakes using their own lake specific split-window coefficients and the SW equation used by *Hulley et al.* (2011). ArcLakes is another LSWT product derived exclusively from ATSR series data for a set of selected lakes (*MacCallum and Merchant, 2012*).

In this thesis, we used satellite specific split-window coefficients obtained from *Jimenez-Munoz and Sobrino* (2008) to derive LSWT from different sensors on board of 13 satellites. Given that the emissivity of water surface is stable spatially and close to unity (*Quattrochi and Luvall, 2004*), the mean emissivity is taken as unity in the methodology developed in this thesis. Hence the final SW equation to retrieve LSWT used in this thesis is modified from Eq. 1.5 and is given in Eq. 1.6.

$$LSWT = T_i + c_1(T_i - T_j) + c_2(T_i - T_j)^2 + c_0 \quad (1.6)$$

The *in-situ* data used in this study for validation represents bulk temperature from the epilimnion of the lake at 0.0 - 0.5 m depth. The satellites, however, measure the temperature above the sub-micron layer between water surface and the air which is highly variable according to the meteorological conditions (*Merchant, 2013*). This is due to the skin effect existing over water surfaces due to the presence of a thermal skin. The skin temperature is often cooler than the bulk temperature in the order of 0.2 - 0.5 $^{\circ}\text{C}$ (*Hook et al., 2003; Wilson et al., 2013*). The effect is normally low during daytime due to the presence of strong solar heating which heats up the upper layer of lakes (*Merchant, 2013*).

1.1.2 Global warming of lakes in the last decades

A recent study on 235 globally distributed lakes using a combination of *in-situ* and satellite data reported an average warming of summer surface water temperatures of $0.034\text{ }^{\circ}\text{C yr}^{-1}$ between the years 1989 and 2009 (O'Reilly *et al.*, 2015; Sharma *et al.*, 2015). Schneider and Hook (2010) reported $0.045\text{ }^{\circ}\text{C yr}^{-1}$ warming of night time mean LSWT over the months of January through March and July through September from 1985-2009. The lakes in the mid and high latitudes of the northern hemisphere showed higher warming rates than those in the tropics and southern hemisphere. A recent study by Layden *et al.* (2015) demonstrated the use of satellite derived LSWT to study the detailed climatology of 246 globally distributed lakes. They further explained the role of lake depth, lake altitude and distance from coast in determining the inter-lake variations of LSWT. Schneider *et al.* (2009) reported rapid night time summer warming of lakes in California and Nevada at the rate of $0.11\text{ }^{\circ}\text{C yr}^{-1}$ since 1992. The Great Lakes in the US except for Lake Erie were reported to be warming at high rates of $0.5\text{ }^{\circ}\text{C}$ per decade (O'Reilly *et al.*, 2015). Antarctic lakes are reported to be warming at $0.06\text{ }^{\circ}\text{C yr}^{-1}$ between 1980 and 1995 (Quayle *et al.*, 2002). Lake Tanganyika, one of the African Great Lakes, is reported to be warmed by $0.9 - 1.3\text{ }^{\circ}\text{C}$ since 1913 (Verburg *et al.*, 2003). The surface temperature of Lake Kivu, another deep lake among the African Great Lakes were reported to have been increased by $0.5\text{ }^{\circ}\text{C}$ in last thirty years (Katsev *et al.*, 2014). In the recent global study by O'Reilly *et al.* (2015), the rates of warming followed a regional variation, highlighting the need of regional understanding of lakes' response to climate change.

In Europe, there are multitude of studies on different lakes reporting rapid warming over recent decades. Adrian *et al.* (2009) reported July warming since 1970 at the rate of $0.02 - 0.05\text{ }^{\circ}\text{C yr}^{-1}$ computed from *in-situ* data of multiple northern hemisphere lakes. In western Europe, George *et al.* (2005); Dokulil *et al.* (2006) reported an increase of surface temperature by $1.4\text{ }^{\circ}\text{C}$ between 1960 and 2000. In central Europe, Dokulil *et al.* (2006); Livingstone (2003) reported warming of lakes at rates between $0.1 - 0.2\text{ }^{\circ}\text{C}$ per decade. Investigations carried out on the large lakes south of the Alps showed a rapid warming ($0.1 - 0.3\text{ }^{\circ}\text{C}$ per decade), at a rate comparable with that of other European lakes (Salmaso and Mosello, 2010).

Most of the studies use long-term *in-situ* sampling data to understand the thermal variations over time. There are multiple approaches in determining trends. Trends are estimated either by using slope of linear regressions or with non-parametric tests like Mann-Kendall to determine significant Sen slope (Sen (1968)). With a different approach, Salmaso and Mosello (2010), Coats *et al.* (2006) and Dokulil *et al.* (2006) reported warming at similar rates using volume weighted mean of *in-situ* profile data at different depths of large lakes in Europe

and North America. In all the cases, availability of high-frequent *in-situ* data was crucial to study the long-term variations. In case of missing data or sparse *in-situ* sampling, LSWT estimated from satellite data using thermal remote sensing techniques is considered as a good alternative.

1.2 Thesis rationale

The advancements in remote sensing and wide acceptance of open data policy have led to a plethora of datasets in the public. To study the long-term thermal dynamics of lakes, it is important to have spatio-temporal datasets at high temporal resolution. Remote sensing is a promising alternative to sparse *in-situ* data in providing the required data to study temporal dynamics. The availability of more than thirty years of data from multiple polar orbiting satellites at 1 km spatial resolution and daily global coverage is an opportunity not to be missed in deriving time series of necessary variables. The products derived from satellite data are often termed as level-2 data, which offers comparatively easy processing and are accompanied with quality control layers. These products, such as LST and SST, are different because they are often sensor specific, follow different algorithms, time scales and quality. Most of the products are often spatially aggregated in order to minimize the geometrical issues, but then become unsuitable to study smaller features like lakes. Hence it is not possible to merge in order to study long-term changes. Moreover, the older polar orbiting satellites (like AVHRR data from NOAA-9/11/12/14 satellites) affected largely with position errors due to the orbital drifts, and navigational errors due to discrepancies with the satellite orientation (generally termed as “satellite attitude”) and satellite clock (Brunel and Marsouin, 2000; Baldwin and Emery, 1995). Though there are numerous studies dealing with an accurate geocorrection of AVHRR data, there is a lack of a complete solution to accurately process them (Rosborough *et al.*, 1994). They are rarely used in research due to lack of robust methods and tools which can perform the geometric corrections to accurately process the data. The available solutions are either outdated, they offer limited success, or are not available in public.

Hence, the first part of my thesis (Chapter 2) deals with the issue of software availability by having developed a new methodology and the necessary software tools to accurately process thirty years of AVHRR LAC data (1986-2015) at 1 km spatial resolution for Northern Italy. Chapter 2 explains in detail the developed methodology to accurately process AVHRR LAC level-1B data and does inter-sensor comparisons to demonstrate the robustness of the method. Furthermore, the usability of processed AVHRR data to estimate LSWT is demonstrated.

In addition to the issues with multiple sources of data as mentioned above, another major obstacle to combine data from multiple satellites is different acquisition times. Though the data is spectrally similar, the varying acquisition times is a major obstacle to merge data from different satellites. Moreover, the orbital drift experienced by the older satellites shifted the acquisition time by hours. This affects temperature products significantly as it follows a cyclic daily pattern at least over lakes in Northern Italy. Thus, in order to develop a unified LSWT time series which combines data from multiple satellites, a homogenisation procedure is necessary, which corrects for the varying acquisition times. Chapter 3 deals with the homogenisation procedure by developing a new method to derive thirty years (1986-2015) of gap-filled daily LSWT with a 1 km spatial resolution and a standardized time at 12:00 UTC using a diurnal pattern based homogenisation procedure is described. In the process, dual thermal channels at level-1B from thirteen satellites were processed to obtain brightness temperatures which in turn were used to estimate LSWT. As a case study, satellite derived time series of LSWT were developed for Lake Garda validated using long-term *in-situ* measurements.

The main objective behind developing new homogenised daily LSWT data covering thirty years is to study long-term trends of water temperature of large lakes in Italy. The newly developed satellite derived LSWT time series with daily temporal resolution offers the opportunity to study the seasonal and annual trends of surface water temperature. The selection of study lakes was driven by the availability of long-term *in-situ* data for the validation of satellite derived data. Much of these *in-situ* data were monthly sampled. In order to establish the usability of the developed methodology satellite derived LSWT were validated against *in-situ* measurements of water temperatures in five large lakes in Italy Lake Garda, Lake Iseo, Lake Como, Lake Maggiore and Lake Trasimeno. The results are described in Chapter 4 along with the estimates of long-term seasonal and annual trends in water temperature based on non-parametric tests and empirical models.

To ensure future re-usability of the developed methods by other researchers, it is important to make them openly available to the research community. This is a significant part of science where further openness in sharing tools and data are promoted rigorously. In the last decades, the growth of open geospatial tools like GRASS GIS and the addition of newer specialized geospatial tools are promoting open science in this domain. New developments are deposited in public repositories and appended to the following chapters. From a data perspective, all the level-1B raw data obtained from thirteen satellites are freely available from the respective agencies, making it further extensible to other parts of the globe where data scarcity is main hindrance in doing science.

1.3 Research objectives

This thesis aims to fill the gap in availability of methods and tools to leverage the access to the plethora of satellite data for the benefit of understanding land/water properties over time. Despite the availability of long-term historical satellite data, lack of methods and tools to accurately process them has been a main hindrance to study the long-term changes in biological and physical variables derivable from space, at high spatial and temporal scale. Here, LSWT as the key physical variable under study, new methods will be implemented to develop thirty years of homogenised time series of LSWT from thirteen polar orbiting satellites for five large lakes in Italy. On a data perspective, it is a significant contribution, as the new methods will be reproducible and extensible, making it feasible for other scientists to develop their own datasets. In the context of climate change, the developed time series of LSWT for large lakes in Italy south of the Alps, give new insights into the seasonal and annual warming trends of lakes within that region.

The **first objective** of the thesis is to develop a new method to accurately process historical level-1B data from NOAA series of satellites. To develop the new method, we made use of multiple open source tools which can read level-1B radiances, apply solar and thermal calibration to the channels, remove bow-tie effects on wider zenith angles, correct for clock drifts on earlier images and perform precise geo-rectification by automated generation and filtering of ground control points using a feature matching technique. Thus, data from nine NOAA satellites (NOAA-9/11/12/14/16/17/18/19) were processed using the new method to develop a time series of brightness temperatures from the dual thermal channels. The usability of the time series of brightness temperatures in estimating LSWT using the split-window algorithm (Eq. 1.6) was tested further. As part of this objective new readers written in Python language were developed for the AVHRR level-1B data and are made publicly available. The entire method and results are described in Chapter 2 of this thesis.

The **second objective** of the thesis was to develop a homogenisation method to merge data from thirteen satellites (including the NOAA satellites) to develop longest possible time series of satellite derived LSWT. The derived LSWT using Eq. 1.6 and satellite specific split-window coefficients were then homogenised to a standardized time of 12:00 UTC using a modified typical pattern technique based on diurnal cycles. All the satellite data used were acquired during day between 8:00 to 17:00 UTC. As a case study, this method was applied to develop homogenised LSWT for Lake Garda and validated against long-term *in-situ* data. With this study, a new method was established to homogenise LSWT from multiple satellites which is both reproducible and applicable for other lakes.

The **third objective** was to extend the LSWT time series to five large lakes in northern Italy. Long-term *in-situ* water temperature data of the lakes were collected from the respective partners to validate the developed LSWT for studied lakes. Furthermore, long-term seasonal and annual trends of water temperatures over the last thirty years (1986-2015) were estimated based on the new dataset. Temporal coherence in water temperatures between lakes were estimated on annual and seasonal time scales based on the satellite derived homogenised LSWT.

Chapter 2

New automated method to develop geometrically corrected time series of brightness temperatures from historical AVHRR LAC data

2.1 Abstract

Analyzing temporal series of satellite data for regional scale studies demand high accuracy in calibration and precise geo-rectification at higher spatial resolution. The Advanced Very High Resolution Radiometer (AVHRR) sensor aboard the National Oceanic and Atmospheric Administration (NOAA) series of satellites provide daily observations for the last 30 years at a nominal resolution of 1.1 km at nadir. However, complexities due to on board malfunctions and orbital drifts with the earlier missions hinder the usage of these images at their original resolution. In this study, we developed a new method using multiple open source tools which can read level-1B radiances, apply solar and thermal calibration to the channels, remove bow-tie effects on wider zenith angles, correct for clock drifts on earlier images and perform precise geo-rectification by automated generation and filtering of ground control points using a feature matching technique. The entire workflow is reproducible and extendable to any other geographical location. We developed a time series of brightness temperature maps from AVHRR local area coverage images covering the sub alpine lakes of Northern Italy at 1 km resolution (1986-2014; 29 years). For the validation of derived brightness temperatures, we extracted Lake Surface Water Temperature (LSWT) for Lake Garda in Northern Italy and performed inter-platform (NOAA-x vs NOAA-y) and cross-platform (NOAA-x vs MODIS/AT-

SR/AATSR) comparisons. The MAE calculated over available same day observations between the pairs—NOAA-12/14, NOAA-17/18 and NOAA-18/19 are 1.18 °C, 0.67 °C, 0.35 °C, respectively. Similarly, for cross-platform pairs, the MAE varied between 0.5 to 1.5 °C. The validation of LSWT from various NOAA instruments with *in-situ* data shows high accuracy with mean R^2 and RMSE of 0.97 and 0.91 °C respectively.

2.2 Introduction

The AVHRR Local Area Coverage (LAC) data which are available at nominal 1.1 km spatial resolution (hereafter AVHRR LAC) with an extensive coverage of the Earth is an unique source of the longest medium resolution time series satellite data for regional studies. However, the number of studies which use older AVHRR LAC data are limited, owing to the complexity in achieving acceptable accuracy due to well documented orbital drifts and clock errors with the instruments (*Brunel and Marsouin, 2000; Privette et al., 1995; Baldwin and Emery, 1995; Krasnopolsky and Breaker, 1994; Emery et al., 1989*). The archived AVHRR LAC data is stored in 10 bit precision and consist of sensor data in raw digital counts, calibration coefficients, time codes, quality indicators, angles, telemetry etc, which are appended at different bit locations (*Cracknell, 1997*). The data are provided in NOAA level-1B raw format where Earth location and calibration parameters are appended with the sensor data (digital numbers) but not applied (*Robel, 2009*). The data are freely available on the Comprehensive Large Array-data Stewardship System (CLASS)—<http://www.class.ncdc.noaa.gov/>. CLASS is NOAA's online facility to distribute data from multiple satellites. Indeed, it requires high level processing taking care of inter-satellite calibrations, multiple versions of AVHRR sensor and correcting for the known issues with the older data (*Baldwin and Emery, 1995; Trishchenko et al., 2002; Robel, 2009; Kidwell, 1998*). Developing a long-term time series combining AVHRR LAC data from multiple NOAA instruments demand two different approaches, one for the POD (Polar Orbiter Data) series of instruments (NOAA-9/11/12/14) and another for the KLM (KLM represents the instrument codes) series (NOAA-16/17/18/19) due to: (i) significant differences in the headers of the POD and KLM series data (*Robel, 2009; Kidwell, 1998*), (ii) varying spectral characteristics of the different versions of the AVHRR sensor—AVHRR/1 and AVHRR/2 for POD and AVHRR/3 for KLM (Table 2.1), and (iii) pre-processing steps depending upon the corrections to be applied for accurate navigation (*Baldwin and Emery, 1995*).

The POD instruments suffered largely with position errors due to the orbital drifts, and navigational errors due to discrepancies with the satellite orientation (generally termed as satellite attitude) and satellite clock (*Brunel and Marsouin, 2000; Baldwin and Emery, 1995*). The

satellite position error that occurred from the orbital drift of the instruments, is due to the lack of orbit adjustment systems (*Privette et al.*, 1995). This caused a change in the equatorial crossing time from the original (by hours) and the shift often increased with the age of the instrument (*Price*, 1991). In Fig. 2.1, the observation time is plotted for all the NOAA instruments used in this study, showing a rather larger shift with the earlier missions. The effects of orbital drift on the derived vegetation indices and sea surface temperature is well documented by *Privette et al.* (1995) and *Kaufmann et al.* (2000). On the other hand, the clock error is due to drift in the satellite clock over time. To overcome this, a periodical reset of the clock was performed in the earlier NOAA instruments (*Baldwin and Emery*, 1995). This also contributed to the navigational error due to clock mis-synchronization. To correct the navigational errors, historical ephemeris (position of satellite in space at a given day and time) data stored in two line element format are widely used (*Baldwin and Emery*, 1995). The attitude control system comprises of three rotation angles - roll, pitch and yaw which control the actual orientation by maintaining the spacecraft fixed axes along the desired geodetic directions (*Rosborough et al.*, 1994). In the earlier AVHRR LAC data attitude errors are difficult to correct due to missing data of roll, pitch and yaw angles. To correct the navigational errors due to attitudinal shifts, the angles are modeled from a set of known landmarks identified from the images, but this approach has limited success (*Rosborough et al.*, 1994; *Baldwin and Emery*, 1995). The KLM instruments on the other hand were launched with much improved AVHRR/3 six channel sensor and had advanced systems on board to minimize the navigational errors. Numerous studies in the past deal with the accurate geolocation of the AVHRR LAC data to develop continuous time series irrespective of the NOAA instrument (*Bordes et al.*, 1992; *Fontana et al.*, 2012; *Moreno and Melia*, 1993; *Rosborough et al.*, 1994; *Brunel and Marsouin*, 2000; *Latifovic et al.*, 2005; *Khlopenkov et al.*, 2010; *Hüsler et al.*, 2011; *Riffler et al.*, 2015). One of the earlier methods is the automatic adjustment of the AVHRR images by applying a physical deformation model of the actual image using the ephemeral elements followed by landmark adjustment using reference Ground Control Points (GCP) along coastal lines (*Bordes et al.*, 1992; *Brunel and Marsouin*, 2000). With this method, an accuracy within one pixel is achieved on images acquired from NOAA-11/12/14. *Hüsler et al.* (2011) used this method as part of their AVHRR archiving process for images over Europe. *Moreno and Melia* (1993) used a small sample of GCP's to refine the orbital elements of the satellite thereby improving the registration accuracy of the images. More recently, advanced image matching techniques have been used to extract GCP's to achieve sub-pixel accuracy in geo-rectification of AVHRR images (*Latifovic et al.*, 2005; *Khlopenkov et al.*, 2010). *Latifovic et al.* (2005) developed historical AVHRR 1.1 km baseline data over entire Canada and used all the AVHRR LAC data acquired by NOAA POD and KLM instruments. They used an image matching procedure to automatically extract GCP's.



Figure 2.1: Plot of variation in observation times of National Oceanic and Atmospheric Administration (NOAA) instruments, note the large orbital drifts of the earlier NOAA-9/11/12/14 instruments.

Table 2.1: Spectral resolution of Advanced Very High Resolution Radiometer (AVHRR) sensors in μm .

Channels	AVHRR/1	AVHRR/2	AVHRR/3
1	0.58-0.68	0.58-0.68	0.58-0.68
2	0.725-1.10	0.725-1.10	0.725-1.10
3a	N.A	N.A	1.58-1.68
3b	3.55-3.93	3.55-3.93	3.55-3.93
4	10.50-11.50	10.3-11.3	10.3-11.3
5	Ch.4 repeated	11.5-12.5	11.5-12.5

However, the potential use of earlier AVHRR LAC images is often hindered by lack of a complete solution to accurately process them (*Rosborough et al.*, 1994). They are seldom used in research due to lack of open software tools which can perform the aforementioned corrections and accurately process the data. The available solutions are either outdated, they offer limited success, or are published with proprietary licenses. Newer methods and geospatial libraries have become available in the last decade and it is now worthwhile to revisit the AVHRR navigation (*Brovelli et al.*, 2012; *Neteler et al.*, 2012). At present, the ability to accurately process all the AVHRR LAC images are limited to very few research centers with satellite receiving stations. There is also lack of publicly available processed AVHRR LAC

dataset covering a particular area of interest; on the other hand, the level-1B data are freely available for any part of the world. Therefore, we see a strong demand for an open source solution to process the AVHRR LAC data in level-1B format for any area of interest which will support the scientists in regional level studies.

Our aim is to develop an automated workflow using multiple open source geospatial libraries and software packages to process all existing AVHRR LAC data (since 1986) for our study area. As a proof of concept, geo-rectified time series (1986-2014) of brightness temperature from the two thermal bands (channels 4 and 5) of AVHRR are developed, which is then intended to be used for estimating the historical surface temperature of sub-alpine lakes (*Riffler et al.*, 2015). The quality of the developed time series is validated for the automatic geo-rectification and thermal calibration. We estimated long-term LSWT of Lake Garda from the time series of brightness temperatures and performed inter-platform and cross-platform comparisons along with *in-situ* data to validate the product. We used Pytroll, a set of open source Python libraries originally meant for processing weather satellite data to read, correct and calibrate the AVHRR data [urlhttp://www.pytroll.org/](http://www.pytroll.org/). We extended the Pytroll to enhance its support for POD and KLM series of AVHRR LAC data. We used the feature matching technique in Orfeo toolbox for precise geo-rectification (*Inglada and Christophe*, 2009). Finally all the post processing including cloud masking, workflow automation, and time series development were accomplished in the GRASS GIS 7.0 software (*GRASS Development Team*, 2015; *Neteler et al.*, 2012; *Neteler and Mitasova*, 2008).

2.3 Methods

2.3.1 Data and software

We obtained all the available AVHRR LAC data from the NOAA public archive (CLASS) in the period 1986-2014 covering the study area of Northern Italy (Longitude: 42.9762°E–46.2371°E; Latitude: 8.2127°N–12.3456°N; Area: 360 km × 340 km). Only day time images acquired between 8:00 to 17:00 UTC are considered for this study. The data are stored in binary format and consist of two major parts—data header and the scan line data. The data header contains generic data on the instrument, acquisition time, data quality, orbital parameters, and conversion factors for calibration and telemetry data, while the scan line data contain data specific to a single scan or time segment including calibration, navigation, telemetry and sensor data, among many others. A complete list of headers is given here: POD–*Kidwell* (1998); KLM–*Robel* (2009). There are 2,048 points in a LAC scan line on which the solar zenith angle and Earth location data (*latitude* and *longitude*) are sampled at every

40th sample, while the sensor data are recorded for each of the 2,048 samples. Fig. 2.2a,b shows the year-wise distribution of the obtained level-1B data and the time line of all the NOAA instruments, respectively. As listed in the Table 2.1, AVHRR/1 lacks dual thermal bands required for the split window technique of surface temperature calculation, which is our main goal in processing this data. Hence the AVHRR LAC images acquired by the earliest NOAA instruments—TIROS-N, NOAA-6/8/10 were avoided. Due to non-availability of day images over the study area, data from NOAA-15 were also discarded. In line with the main objective of estimating day time Lake Surface Water Temperature (LSWT) for the sub-alpine lakes in the study area, we processed only the day images.

We combined multiple open source geospatial software packages in a single workflow to meet the specific requirements of AVHRR processing. The software used, its source code links and the licenses are listed in Table 2.2. Pytroll is an open source project, contributed by several developers primarily from national meteorological services. Among the Pytroll libraries, *mpop*, *pygac* and *pyresample* are used respectively for reading, calibrating, and re-sampling the AVHRR LAC images. Fig. 2.3 shows their specific roles in the workflow. We used these libraries to read all the AVHRR LAC level-1B data, apply corrections, calibrate and write all the channels to separate TIFF images. Orfeo toolbox is a remote sensing image processing toolbox developed in C++ primarily by the Centre National d'Études Spatiales, France (*Inglada and Christophe, 2009*). We used the Scale Invariant Feature Transform (SIFT) feature matching algorithm in Orfeo toolbox to extract homologous points (hereafter GCP's; *homologous* meaning *same relative position*, a term used in Orfeo toolbox module – *otbcli_HomologousPointsExtraction* – for extracting similar GCP's from two images using feature matching algorithms like SIFT) in order to apply final matching of each image against a reference image (*Lowe, 2004*). The GCP's created are then filtered using the newly developed *m.gcp.filter* tool in GRASS GIS (*GRASS Development Team, 2015*). GRASS GIS is one of the most stable and well established open source geographical information systems and remote sensing software which can handle both vector and raster processing in two and three dimensions (*Neteler et al., 2012*). The calibrated AVHRR LAC images are rectified using the filtered GCP's with *i.rectify* module in GRASS GIS. We developed all the raster processing and the cloud removal procedure in GRASS GIS (*Neteler and Mitsova, 2008*). Fig. 2.3 shows the workflow diagram with the respective libraries listed horizontally against the specific requirement they meet.

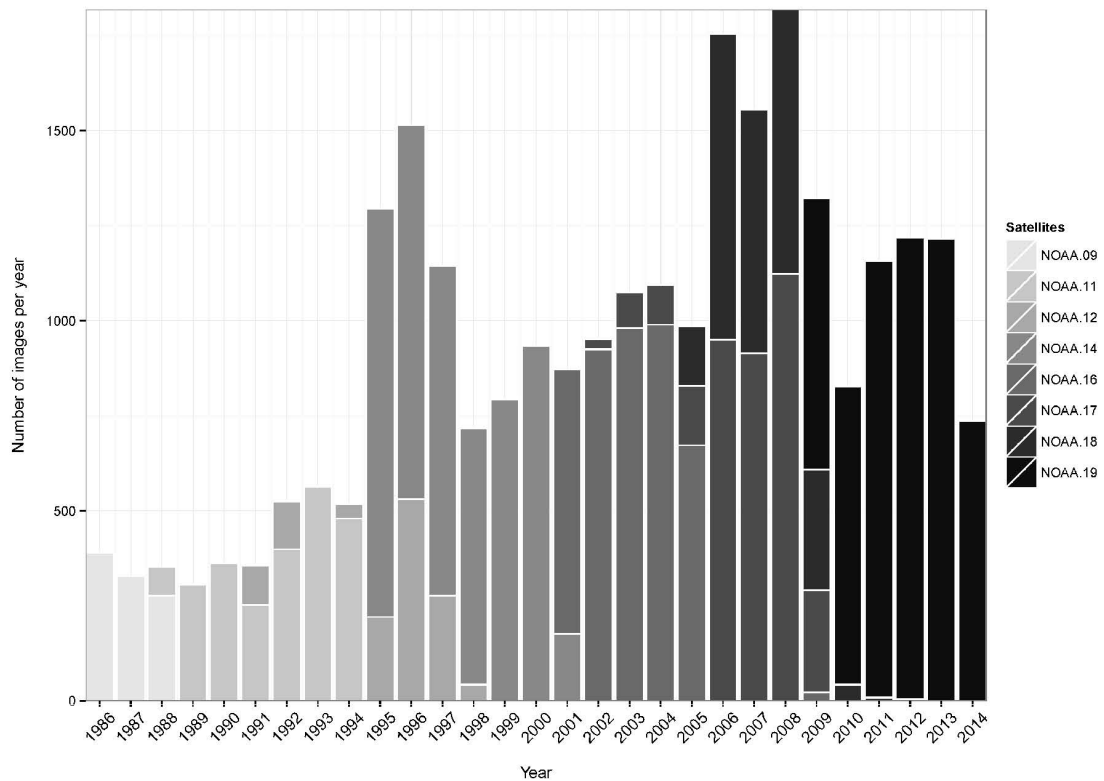
2.3.2 Pre-processing

The pre-processing of AVHRR LAC data in level-1B format include reading the data, performing solar and thermal calibration to the respective channels, applying clock drift correction and cloud removal. The foremost requirement of this methodology was to develop an AVHRR LAC data reader which can read all the versions of AVHRR LAC level-1B data irrespective of the NOAA instrument. The Python library *pygac* in the Pytroll framework is developed to read and calibrate AVHRR global area coverage level-1B data of 4 km spatial resolution which is derived from LAC data by resampling at every 5th sample along a scanline. We enhanced the *pygac* library to add support to AVHRR LAC data (Fig. 2.3).

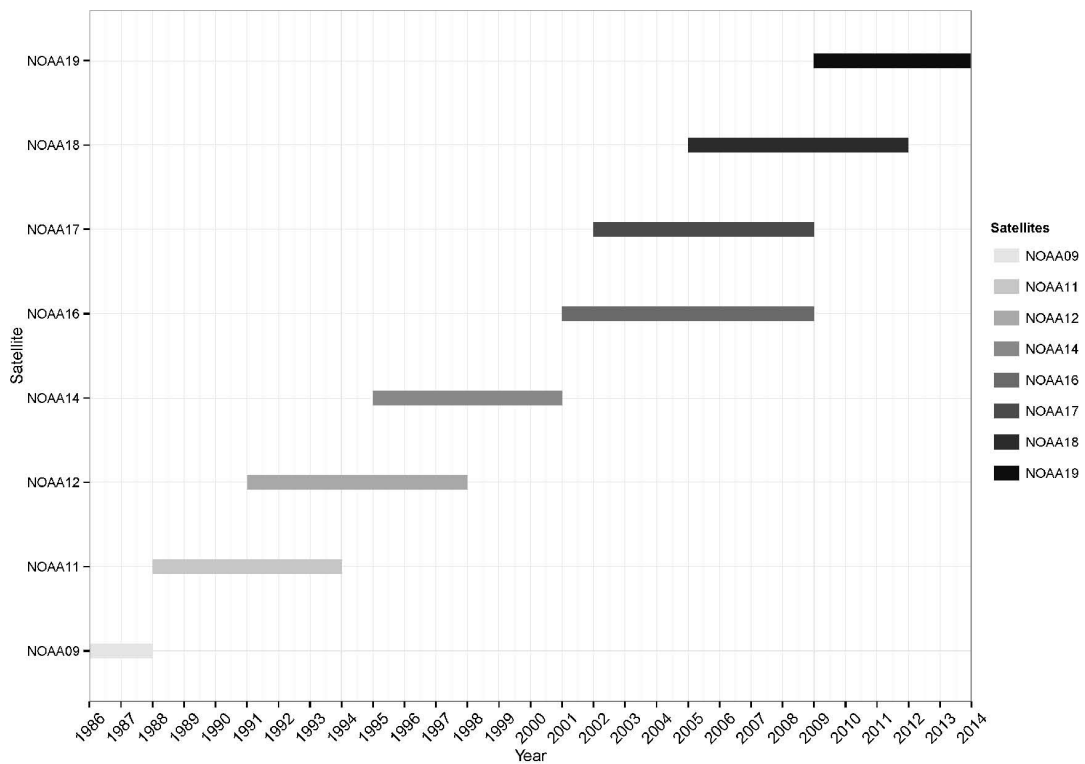
Table 2.2: List of software packages used in this study (License type urls: <http://www.gnu.org/copyleft/gpl.html>, http://www.cecill.info/licences/Licence_CeCILL_V2-en.html).

Software	Source	Version	License
Pytroll/ <i>mpop</i>	https://github.com/pytroll/mpop	-	GNU GPL v3
Pytroll/ <i>pygac</i>	https://github.com/adybbroe/pygac	-	GNU GPL v3
Pytroll/ <i>pyresample</i>	https://github.com/pytroll/pyresample/	-	GNU GPL v3
Orfeo Toolbox	https://www.orfeo-toolbox.org/	4.4.0	CeCILL v2
GRASS GIS	http://grass.osgeo.org/	7.0.0	GNU GPL v3

The fundamental structure of the AVHRR LAC level-1B header and sensor data records is consistent across data types, however there are significant differences between records originating from various instruments (POD vs KLM) which have to be dealt separately. Hence we developed and added two readers (*lac_pod*—POD data and *lac_klm*—KLM data) in *pygac* to successfully read and calibrate the AVHRR LAC data irrespective of the NOAA instrument. The newly developed readers are released to the public in the source code repository here—<https://github.com/adybbroe/pygac>. The main difference between the POD and KLM drivers is the mapping of the header structure. All the different releases of level-1B data types by NOAA over POD are mapped in the new readers following the NOAA documentation. The Python library *mpop* in the Pytroll framework can read multiple satellite data in raw format and use the coefficients given in the header to apply calibration and corrections. *mpop* has an extendable objective design which enables us to write plugins to support additional data formats. We developed a plugin (*lac_l1b*) which is a wrapper to the new LAC readers in *pygac* to read and calibrate the AVHRR LAC data.



(a)



(b)

Figure 2.2: Data distribution. (a) Distribution of AVHRR Local Area Coverage (LAC) images over study area from 1986 to 2014 obtained from NOAA Comprehensive Large Array-data Stewardship System (CLASS) archive, separated by different instruments. (b) Time line of NOAA instruments used in this study.

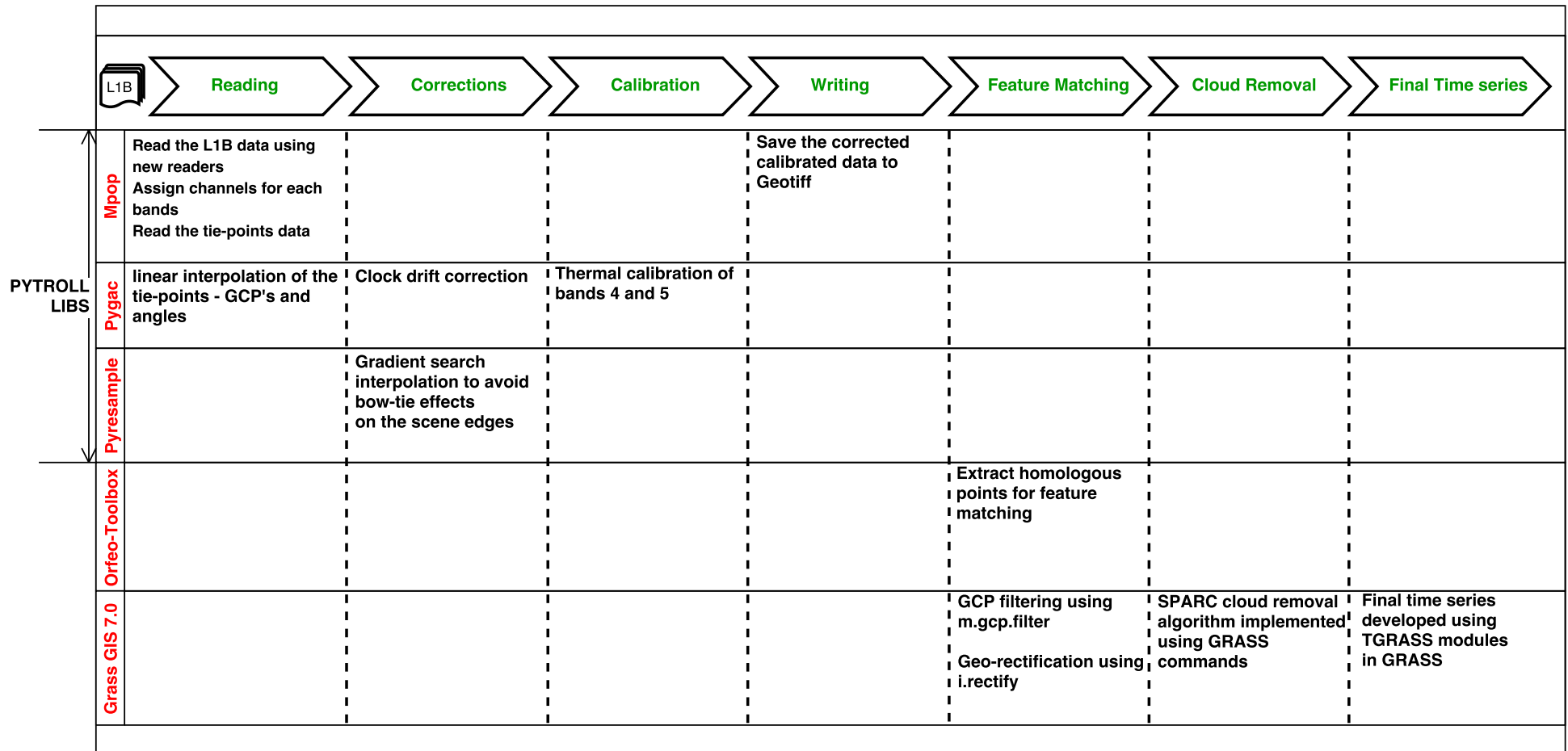


Figure 2.3: Workflow diagram. On top, the horizontal arrows show the main steps involved in the workflow, the software are listed vertically on the left. Role of each software in the workflow is explained in the corresponding cells.

For the calibration of the solar channels ($0.63\ \mu\text{m}$ and $0.86\ \mu\text{m}$) we used the inter-sensor consistent calibration coefficients developed by *Heidinger et al.* (2010) which ensures the continuity of data even with the transition of satellites. Unlike the solar channels, an on board calibration is performed on the thermal channels (*Robel, 2009; Kidwell, 1998*). In spite of on board calibration, the data are prone to contamination (leading to bias of more than $1\ ^\circ\text{C}$) with unwanted fluctuations due to multiple factors including; (i) atmospheric attenuation of the signal, (ii) decaying of the instruments, and (iii) solar contamination (*Trishchenko et al., 2002; Cao et al., 2001*). To remove these perturbations in the true signal *Trishchenko* (2002) introduced a robust median based approach to remove short time fluctuations from a given sample of calibration elements followed by Fourier transform filtering to remove the persisting errors over long time. For more details on thermal calibration, refer to *Trishchenko et al.* (2002) and *Trishchenko* (2002). The aforementioned calibration techniques are successfully implemented in operation for generating historical AVHRR baseline data records over Canada by *Latifovic et al.* (2005) and more recently to develop LSWT time series for Alpine lakes in Europe by *Riffler et al.* (2015). In Pytroll, *pygac* library has implemented thermal calibration procedure as per *Robel* (2009); *Kidwell* (1998) followed by a Fourier transform filtering to remove high frequency fluctuations (*Trishchenko, 2002*); which we used to successfully calibrate the thermal channels (Fig. 2.3). The ancillary data like GCP's and the viewing zenith angle are provided as tie points (51 per scanline). We used *pygac* library to apply linear interpolation on *latitude*, *longitude* and viewing zenith angle tie points to all the points in the geographical grid covered by the data (Fig. 2.3).

The polar orbiting NOAA satellites are launched in sun synchronous orbits at an altitude of around 870 km above Earth (*Cracknell, 1997*). They are designed to acquire two daily snapshots of Earth, one in ascending and other in the descending mode. This is possible due to the wider scan angles (also called zenith angles) of $\pm 55.4^\circ$ covering approximately a 2,300 km wide scene. Subsequently, the spatial resolution also varies significantly off-nadir from 1.1 km at nadir to 4 km at the edges both in along-scan and along-track directions (*Cracknell, 1997*). Due to this wide off-nadir field of view, panoramic bow-tie effects are seen towards the edges of the images (Fig. 2.4a) (*Khlopenkov and Trishchenko, 2008*). This creates artifacts when applying re-projection with nearest-neighbour resampling on the level-1B data. In this methodology, we used an algorithm based on 2D gradient search in the latitude and longitude geolocation fields using their local gradients to project level-1B data to Lambert Azimuth equal-area projection, which corrects for the bow-tie effects. Complete details of the algorithm and its implementation is explained in *Khlopenkov and Trishchenko* (2008). In Pytroll, this method is implemented in *pyresample* library which is used to resample all the AVHRR level-1B data (Fig. 2.3).

The POD satellites (NOAA-9/11/14) were affected by drifts in the satellite clock which in turn contributed to the position errors (*Baldwin and Emery, 1995*). The difference between satellite clock time and the actual coordinated universal time (offset Δt) was provided by NOAA periodically (see Section 2 at *Kidwell (1998)*). The clock drift leads to location error upto 4 km, as the GCP's are estimated based on the satellite clock (*Kidwell, 1998*). The known offsets Δt for the satellites NOAA-9/11/14 are used in the *pygac* library to correct for clock drifts.

Unlike for level-2 MODIS products, there are no existing cloud mask available for the level-1B AVHRR LAC data. For this, we adapted the algorithm developed by *Khlopenkov and Trishchenko (2007)*, Separation of Pixels Using Aggregated Rating over Canada (SPARC) originally implemented for creating a cloud mask over Canada. As our main interest is to remove thick clouds and thin cirrus above the lakes, we used two relevant tests from the original SPARC algorithm, (i) Brightness temperature test (T-test) using channel 4 and (ii) thin cirrus test (C-test) which uses the difference between channels 4 and 5. T-test uses the channel 4 brightness temperature and compares it with a dynamic threshold which is the surface skin temperature data of the corresponding day and time derived from climatic models. In the original study, North American regional reanalysis is used. Here we replaced it with the European regional analysis interim dataset developed by the European Center for Medium-range Weather Forecasts (ECMWF) (*Dee et al., 2011*) following the successful implementation over Europe by *Riffler et al. (2015)*. The SPARC algorithm is implemented using raster processing tools in GRASS GIS 7.0 (Fig. 2.3).

2.3.3 Precise geo-rectification using SIFT

After calibration and corrections, the AVHRR LAC data still suffer from position errors due to erroneous satellite attitude angles (*Baldwin and Emery, 1995; Rosborough et al., 1994; Brunel and Marsouin, 2000*). For any long-term analysis of landscape properties using historical AVHRR data, it is important to obtain precise geometric position for all the images to avoid spurious trends and artifacts. Hence we used an automatic feature matching technique called Scale Invariant Feature Transform (SIFT) by *Lowe (2004)* to extract matching GCP's between an image pair to achieve sub-pixel accuracy in image to image rectification (*Khlopenkov et al., 2010; Latifovic et al., 2005*). We selected a cloud-free geometrically corrected image from NOAA-19 ascending mode (dated 01 August 2012) as reference image to all the other data to which image to image matching is applied. The homologous points extraction module in the Orfeo Toolbox 4.4.0 is used to automatically extract the matching GCP's using the SIFT algorithm (Fig. 2.3). The SIFT algorithm is widely used in object recognition from

images (Lowe, 1999). Yu *et al.* (2008) and Fan *et al.* (2013) successfully implemented SIFT features in registering multiple source remote sensing imageries. The SIFT workflow starts with identifying and populating the SIFT features from the input and reference images. The SIFT features are scale invariant and highly distinctive which makes them suitable for image matching techniques. The matching SIFT features are then identified based on euclidean distance of their feature vectors. The SIFT algorithm was developed and explained in detail by (Lowe, 2004).

We applied SIFT individually to all the bands of a single AVHRR LAC image and combined all the extracted GCP's. If the total number of matching GCP pairs are less than 20, often due to wide coverage of cloud, then the input image is discarded from further processing. Before applying image rectification using the extracted GCP's, it is important to discard the GCP pairs with larger Root Mean Square Error (RMSE). We used the newly developed GRASS GIS add-on *m.gcp.filter* to apply this filter. GCP filtering is applied in iterative mode till all the GCP's with individual RMSE larger than 500 m (half a pixel) are discarded. Finally, we applied polynomial second order image rectification using *i.rectify* tool in GRASS GIS. The entire process of GCP extraction using SIFT, GCP filtering and geo-rectification is automated (see Appendix A.1 for the code). For validating the accuracy of automated geo-rectification process, we extracted independent GCP's (12,340 GCP's) from 2,000 randomly selected geo-rectified AVHRR LAC images (around 10% of total number of images). The x-deviation and y-deviation along with RMSE are then calculated for validating the quality of geo-rectification (Rocchini and Di Rita, 2005; Khlopenkov *et al.*, 2010).

2.3.4 Validation of thermal calibration by estimating LSWT

The thermal calibration computes top of atmosphere brightness temperatures from channels 4 and 5 emitted radiances. In order to validate the performance of thermal calibration on thermal channels from multiple NOAA instruments, we estimated Lake Surface Water Temperature (LSWT) for the Lake Garda in Northern Italy using the split-window technique from all valid brightness temperatures. LSWT is estimated using the modified non-linear split-window Eq. (2.1) (Jimenez-Munoz and Sobrino, 2008).

$$T = T_4 + c_1(T_4 - T_5) + c_2(T_4 - T_5)^2 + c_0 \quad (2.1)$$

where T_4 and T_5 are brightness temperatures derived from NOAA AVHRR channels 4 and 5 respectively; c_0, c_1, c_2 are split-window coefficients. We used the instrument/sensor specific split-window coefficients derived and published by Jimenez-Munoz and Sobrino (2008). We

further estimated LSWT using the same method from the MODIS, ATSR1, ATSR2 and AATSR sensors which have similar dual thermal channels in the same spectral range. We used the MODIS channels 31 and 32 while for ATSR1, ATSR2 and AATSR, we used the channels 11 and 12 for LSWT estimation. To correct for the different acquisition time (between 8:00 to 17:00 UTC) of the satellites and the orbital drifts experienced with earlier NOAA instruments (Fig. 2.1) we applied a correction to the derived LSWT based on Diurnal Temperature Cycle (DTC) model. This procedure is intended to homogenise the derived LSWT's to the noon time at 12:00 UTC irrespective of the acquisition time. The model used is explained by the Eq. 2.2.

$$T_s(t) = T_0 + T_a \cos(\pi/\omega(t - t_m)) + T_b \sin(\pi/\omega(t - t_m)) \quad (2.2)$$

where,

$$\omega = 4/3(t_m - t_{sr}) \quad (2.3)$$

where T_0 is the residual temperature around sun rise; T_a and T_b are temperature amplitudes; $T_s(t)$ is surface temperature at time t ; t_m is the time at which temperature is maximum; t_{sr} is the time of sun rise; ω is calculated using Eq. 2.3. The correction factor is calculated for each LSWT from the corresponding diurnal cycle as the absolute difference between LSWT at $T_s(t)$ and $T_s(12)$. Finally the LSWT's are homogenised to 12:00 UTC noon fitting to the diurnal cycle by applying the correction factor.

A global filter is then applied to remove the outliers due to undetected clouds discarding all the LSWT's beyond the range 7.0 °C–30 °C which is the long-term minimum and maximum temperature respectively measured over Lake Garda. Further, we limited the LSWT estimation to those pixels acquired at viewing zenith angle less than 45° (Hulley *et al.*, 2011). We estimated the RMSE and bias between inter-platform (AVHRR between different NOAA instruments) and cross-platform (AVHRR/NOAA versus other sensors) LSWT's averaged over lake. We took all the days where a pair of observations is available from different sensor/instruments. For inter-platform comparisons, same day observations for the pairs NOAA-12/14, NOAA-17/18 and NOAA-18/19 are compared. Similarly for cross-platform comparison, the pairs NOAA-11/ATSR1, NOAA-14/ATSR2, NOAA-16/MODIS-Aqua, NOAA-17/MODIS-Terra, NOAA-17/AATSR, NOAA-18/MODIS-Aqua and NOAA-19/MODIS-Aqua are compared. We compared mean LSWT computed over lake per image between these sensors.

Further, we validated the satellite derived LSWT for Lake Garda using historical *in-situ* data. Monthly *in-situ* data measured using a multi-parameter probe from the deepest point of the lake covering the period 1991-2013 are used for validation. The *in-situ* data are measured

around 12:00 UTC noon. The corresponding AVHRR images matching the dates of *in-situ* data and acquired between 8:00 to 17:00 UTC are used for validation. We estimated the correlation statistics, R^2 , RMSE and Mean Absolute Error (MAE) between the *in-situ* data and the corresponding LSWT extracted from the pixel representing the field location (Table 2.4).

2.4 Results

2.4.1 Time series of calibrated AVHRR LAC data

The historical AVHRR LAC images at 1.1 km spatial resolution available for more than last 30 years is the longest time series of satellite data and a potential alternative to ground observations. Indeed, to develop the longest satellite derived time series of AVHRR LAC data, it is inevitable to combine data from multiple instruments which in turn require superior inter-sensor consistent calibration (*Heidinger et al.*, 2010; *Trishchenko*, 2002). A total of 22,507 input images were acquired and processed using the new method (Fig. 2.2a,b). The newly developed AVHRR LAC data readers in *pygac* and the new *mpop* plugin works with all the level-1B data across the POD and KLM series of instruments from the CLASS archive. The code segment to read the POD and KLM data using the new plugin in *mpop* is given in Appendix A.1. The new *mpop* plugin has been made available in a public repository with GNU general public license <https://github.com/pytroll/mpop>.

The thermal calibration applied to the data as conceived by *Trishchenko* (2002) is robust and takes care of the inter-calibration between different NOAA instruments. A conclusive validation of the quality of calibration between NOAA instruments is beyond the scope of this study. Instead we did inter-platform comparison of estimated LSWT which is explained in subsection 2.4.3. AVHRR is a wide angle $\pm 55.4^\circ$ sensor using the whisk broom technique to acquire large swath in one acquisition. Nevertheless, it is important to resolve the artifacts in the wider angles due to the bow-tie effects. Visual comparison of multiple images before and after the resampling shows that the gradient-search method eliminates the bow-tie effects restoring the original geometry over the edges (*Khlopenkov and Trishchenko*, 2008). Fig. 2.4 compares the Nearest Neighbour resampling technique with our approach. Gradient search approach (Fig. 2.4b) provides a clear geometry of the lakes in Northern Italy, while the Nearest Neighbour (Fig. 2.4a) interpolation distorts the image making it impossible to capture the geometry of small objects like lakes.

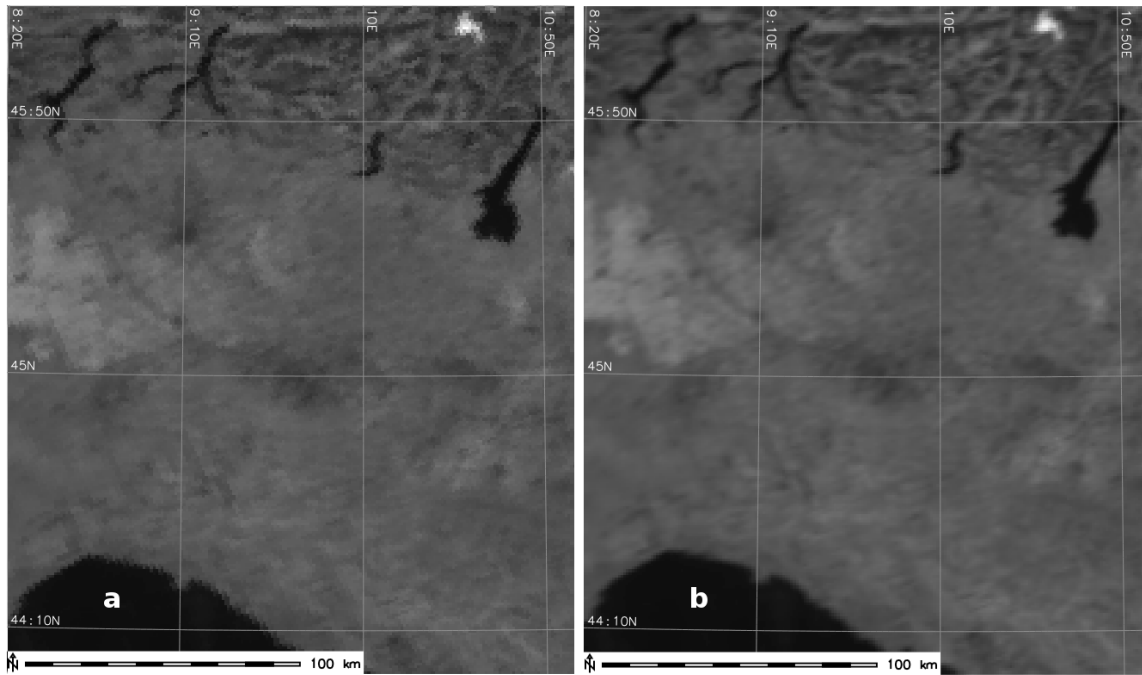


Figure 2.4: NIR (Near Infrared) channel 2 of AVHRR LAC data resampled using (a) Nearest Neighbour interpolation; (b) interpolated using the gradient search algorithm. The image is NOAA-18 AVHRR LAC taken on 1 August 2011.

2.4.2 Geo-rectification of calibrated AVHRR LAC data using SIFT

We found that the number of extracted GCP's using SIFT depends upon two major factors; (i) the overlapping area between the input and the reference image (ii) the cloud coverage. SIFT tends to fail on input images with cloud coverage more than 50% of the total area. On a clear sky image, combining all the corresponding bands, large number of GCP's are extracted (> 300) for an area of $122,400 \text{ km}^2$, which is then filtered iteratively to remove outliers ($\text{RMSE} > 500 \text{ m}$). We observed that with large number of GCP's, linear model rectification often failed due to islands of localized point clouds in the image. Hence we used the polynomial second order rectification for all the images irrespective of the number of GCP's obtained from SIFT. Fig. 2.5 shows different steps of the feature matching based geometric correction applied on a NOAA-14 AVHRR LAC image (POD) acquired on 9 August 1997. The input image is shown in Fig. 2.5a (band 2-NIR) overlaid with the extracted GCP's in which the position error is clearly visible along the national boundary and the lakes. Fig. 2.5c shows the final output which is precisely aligned to the country and lake boundaries.

The validation output shows that the SIFT based geo-rectification procedure is able to achieve sub-pixel accuracy. Fig. 2.6a shows the scatter plot of x deviation and y deviation of the GCP's and Fig. 2.6b shows the kernel density plots. The overall RMSE of 12,340 independent GCP's from 2,000 images is 755.63 m, which is below 1 pixel. The mean deviation in x and y

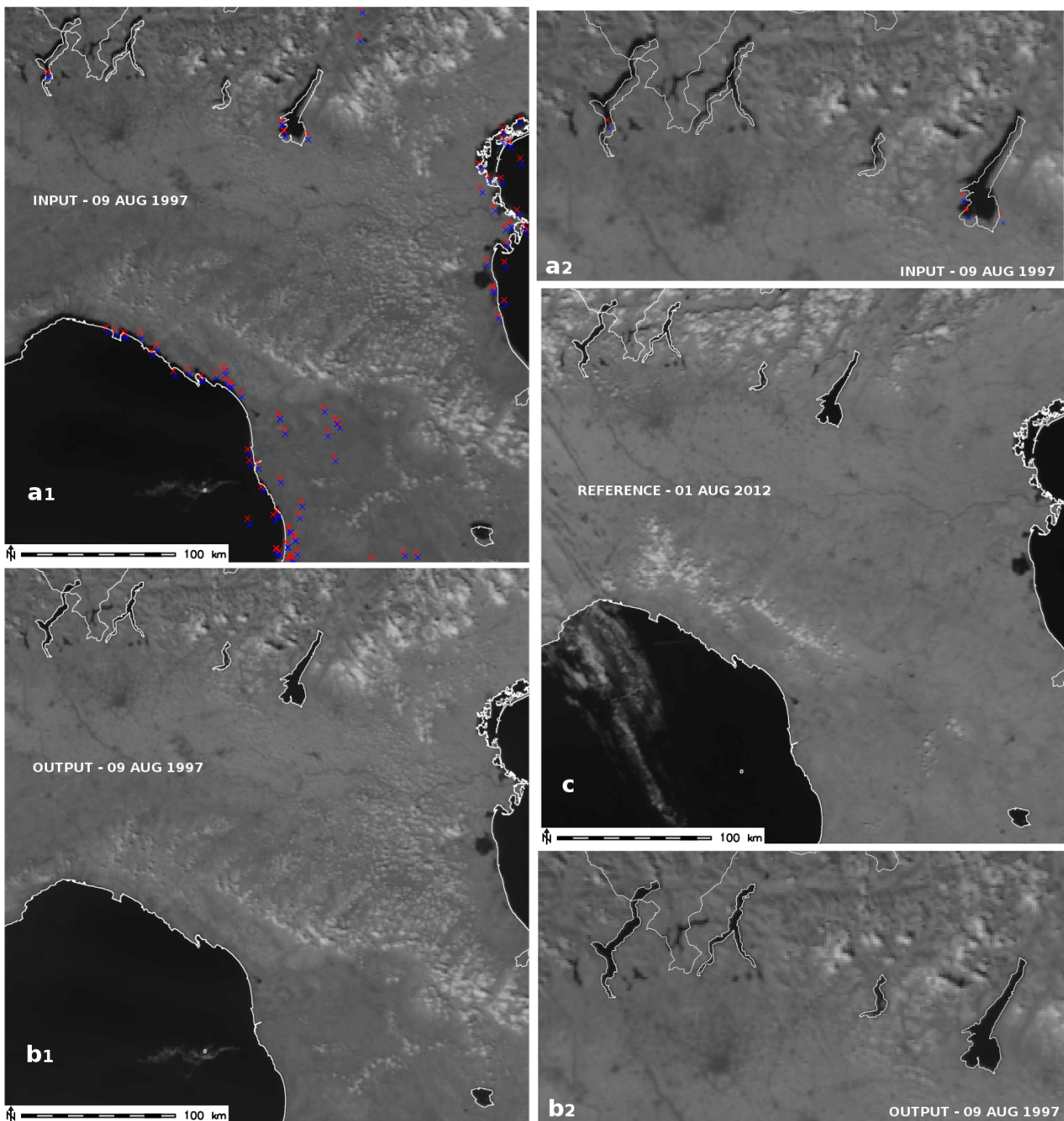


Figure 2.5: (a–c) explains geometric correction done on NOAA-14 AVHRR LAC data (NIR channel 2 shown here) acquired on 09 August 1997. (a₁) The input image with lake and country boundaries overlaid. Also shown is blue (input) and red (reference) crosses, the Ground Control Points (GCPs) extracted using the SIFT algorithm; (b₁) The reference image used to extract the GCP points, NOAA-19 AVHRR LAC acquired on 01 August 2012; (c) The NOAA-14 AVHRR LAC image after polynomial second level rectification using the GCP's; (a₂) Fig. a₁ zoomed to the sub-alpine lakes in Northern Italy; & (b₂) Fig. b₁ zoomed to the sub-alpine lakes in Northern Italy.

direction is 607.82 m and 624.21 m respectively. The mean deviations are found to be on the same range suggesting that there is no geometric errors in either direction.

2.4.3 Quality assessment of the time series using estimated LSWT

For the validation of the calibrated AVHRR LAC time series data, we compared the LSWT of the Lake Garda estimated from all the valid thermal channels, irrespective of the NOAA instrument, after pre-processing, geo-rectification and orbital drift correction. Fig. 2.7 shows boxplots of absolute lake mean LSWT differences between pair of NOAA instruments. The boxplots (Fig. 2.7) are based on the same day observations by multiple NOAA instruments. We could only get three NOAA instrument pairs with same day observations for the inter-sensor comparison—NOAA-12/14, NOAA-17/18 and NOAA-18/19. The MAE for NOAA instrument pairs are 1.18 °C, 0.67 °C, 0.35 °C for NOAA-12/14, NOAA-17/18 and NOAA-18/19 respectively. Moreover the RMSE is 1.36 °C, 0.88 °C and 0.44 °C respectively.

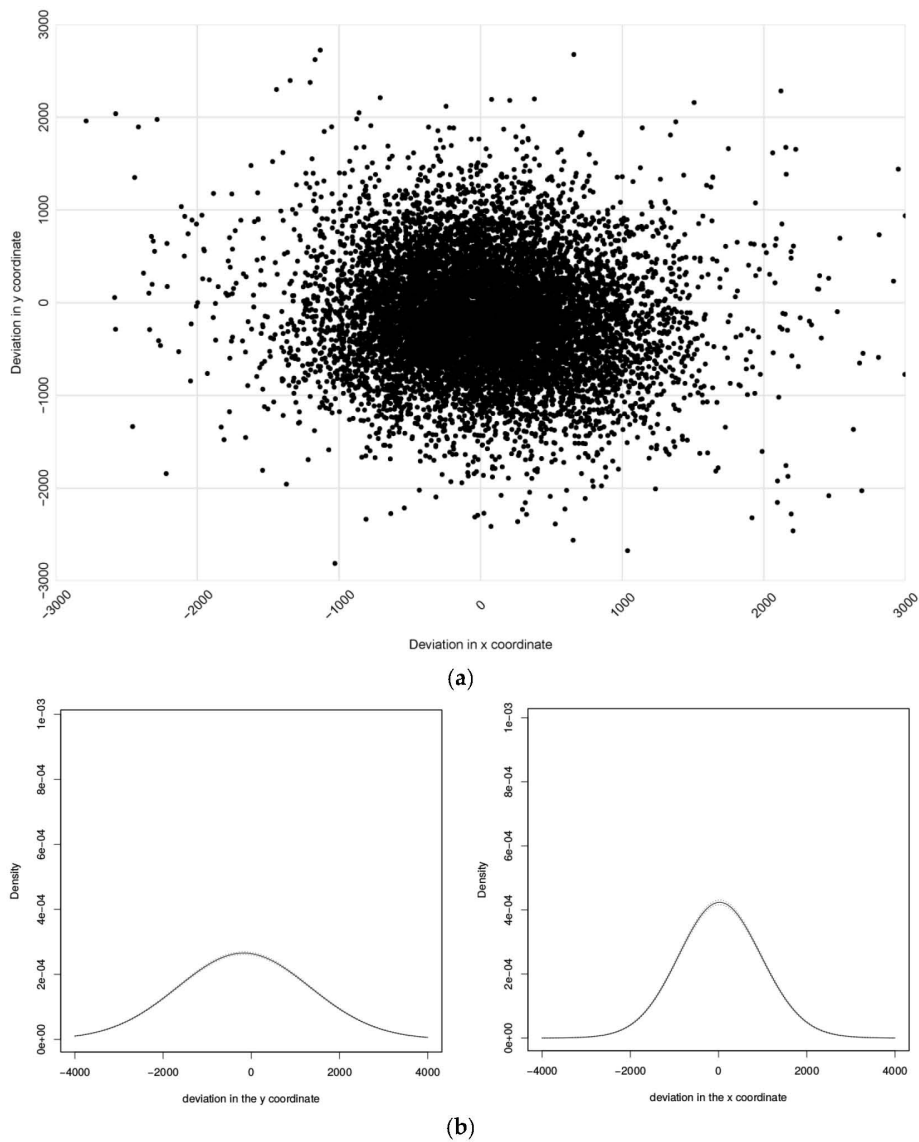


Figure 2.6: Validation of Scale Invariant Feature Transform (SIFT) algorithm. (a) Independent GCP point-cloud plotted in x-deviation, y-deviation; (b) Kernel density plot of x and y deviations.

Boxplots of the absolute lake mean LSWT differences between the cross-platform pairs are shown in Fig. 2.8. The boxplots (Fig. 2.8) are based on the same day observations by multiple instruments. The MAE and RMSE in all the cases are found to be between 0.5 °C to 1.5 °C (Table 2.3). In all the cases except for NOAA-17/MODIS-Terra, the median of lake mean LSWT difference lies below 1.5 °C (Fig. 2.8), which indicates that the majority of the differences are in the lower range of 0-1.5 °C. The lowest MAE, RMSE of 0.6 °C, 0.85 °C respectively are reported with NOAA-17/AATSR observations, while the highest errors of 1.2 °C, 1.35 °C are with NOAA-17/MODIS-Terra (Table 2.3).

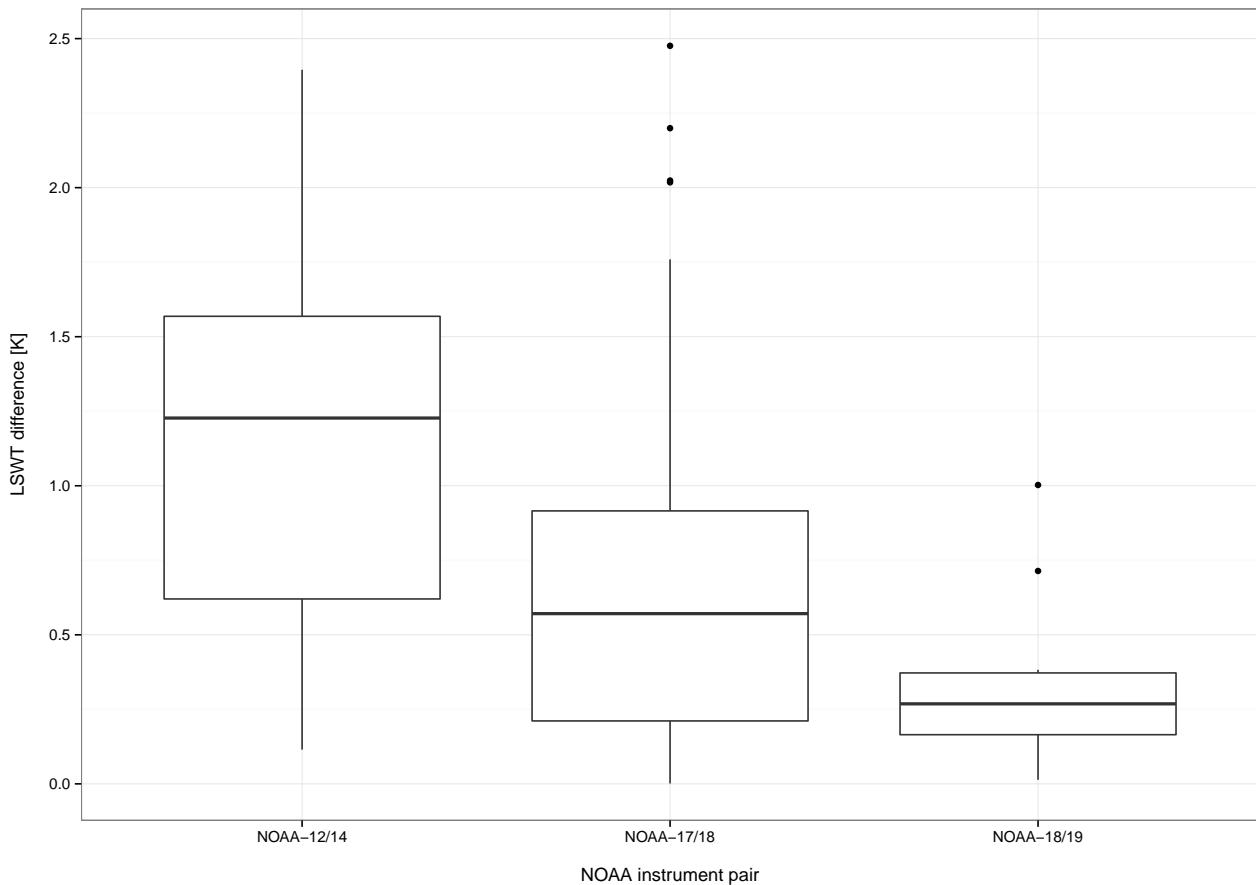


Figure 2.7: Boxplots of absolute difference between lake mean Lake Surface Water Temperature (LSWT) estimated from multiple NOAA same day observations; horizontal thick line inside the box represents median; lower and upper end of the box represents first and third quartiles respectively; the bottom whisker ranges from first quartile to the smallest non-outlier and the top whisker ranges from third quartile to the largest non-outlier; the dots outside whiskers are outliers.

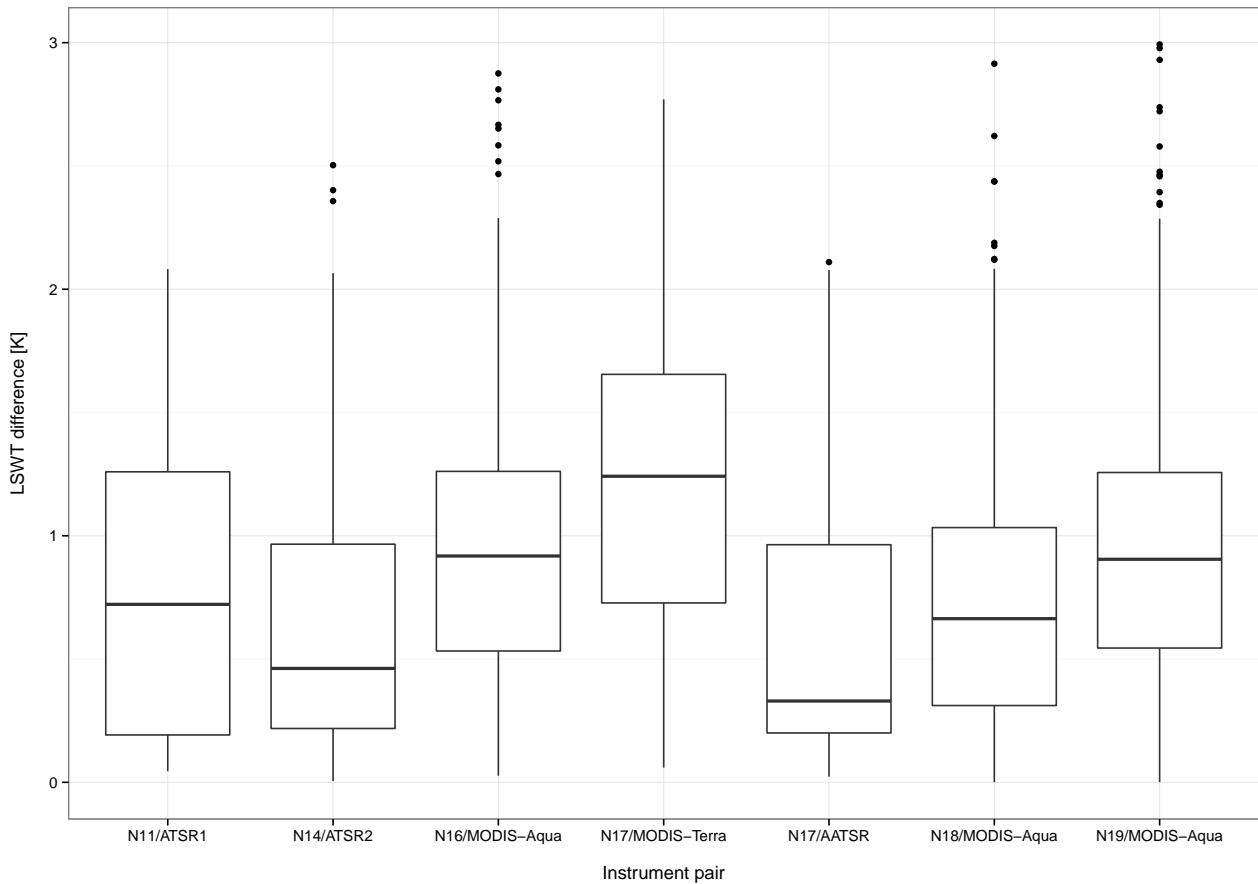


Figure 2.8: Boxplots of absolute difference between lake mean LSWT estimated from same day observations from NOAA and other instruments; horizontal thick line inside the box represents median; lower and upper end of the box represents first and third quartiles respectively; the bottom whisker ranges from first quartile to the smallest non-outlier and the top whisker ranges from third quartile to the largest non-outlier; the dots outside whiskers are outliers.

Table 2.4 lists important correlation statistics derived between LSWT and the *in-situ* data. Slopes are estimated from the linear models between LSWT from each NOAA instrument and the corresponding *in-situ* data. All the slope values are close to each other which shows the similar relationship between LSWT and *in-situ* data irrespective of the NOAA instrument. In all the cases, coefficient of determination R^2 is higher than 0.93. An average RMSE and MAE of 0.92 °C and 0.71 °C are reported from the models.

2.4.4 Software code for deploying the method elsewhere

To enable reproducibility of our method following *Rocchini and Neteler (2012)* and *Reichman et al. (2011)* we publish the code snippets used to process the entire AVHRR LAC data. Appendix A.1 is the Python script to process the AVHRR LAC data using the Pytroll libraries—*mpop*, *pyresample*, *pygac*—which will read level-1B format, apply solar and thermal calibra-

Table 2.3: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) (in °C) between lake mean LSWT estimated between pair of sensors.

Sensor pair	N12/N14	N17/N18	N18/N19	N11/ATS1	N14/ATS2	N17/AATSR	N17/MODT	N16/MODA	N18/MODA	N19/MODA
MAE	1.18	0.67	0.35	0.81	0.72	0.6	1.2	0.99	0.75	0.95
RMSE	1.36	0.88	0.44	1.0	0.98	0.85	1.35	1.16	0.92	1.12
N	26	75	10	38	98	28	75	149	262	385

N = Number of observations; N11 = NOAA-11 (likewise for other NOAA instruments - N11,N14,N16,N17,N18,N19); ATS = AATSR; MODT = MODIS-Terra; MODA = MODIS-Aqua.

Table 2.4: Correlation statistics between *in-situ* data and corresponding LSWT derived from various NOAA instruments; N = Number of samples used in the linear model.

Satellites	Slope	R ²	RMSE	MAE	N
NOAA-11	0.97	0.98	0.81	0.65	12
NOAA-14	0.90	0.93	1.38	1.05	22
NOAA-16	0.98	0.99	0.36	0.28	10
NOAA-17	0.89	0.99	0.56	0.46	7
NOAA-18	0.98	0.97	1.04	0.77	12
NOAA-19	0.92	0.94	1.40	1.08	21

tion, resample using gradient search method, write all the channels and viewing zenith angle to image format. Appendix A.2 is a bash script which is the second part of processing where the output images from Python script are processed inside GRASS GIS session. The bash script executes the SIFT matching using Orfeo toolbox, GCP filtering, geo-rectification and cloud removal.

2.5 Discussion

In this study we introduced a new automated method to accurately develop time series of AVHRR LAC data at 1 km spatial resolution covering 29 years from 1986-2014. The method is developed by chaining several open source geospatial packages (Table 2.2 and Fig. 2.3) which is reproducible and extendable. To the best of our knowledge there are no other published studies which demonstrate successful processing of this data using tools which are available open to all researchers. The new additions to the Pytroll libraries—*mpop* and *pygac*—support all the AVHRR LAC data in level-1B format irrespective of the NOAA instrument, version of the AVHRR sensor and the various data header types released by NOAA. Validation of automated geo-rectification of all the AVHRR LAC data using independent GCP's extracted from 2,000 randomly selected images reported sub-pixel accuracy with an overall RMSE of 755.63 m. The thermal calibration takes care of inter-platform variations (*Trishchenko*, 2002) and our validation procedure shows that the same day LSWT estimated from AVHRR on board different NOAA instruments are close to each other with overall mean absolute differences (Fig. 2.7) below 1.5 °C.

One of the important aspects of this methodology is its ability to deal with the earlier AVHRR LAC data from POD satellites (Baldwin and Emery, 1995; Kaufmann et al., 2000; Rosborough et al., 1994). As proved by earlier studies (Brunel and Marsouin, 2000; Privette et al., 1995; Emery et al., 1989), it is difficult to achieve accurate geo-rectification between temporal AVHRR images acquired from POD satellites. In order to achieve precise navigation, we applied feature matching technique—SIFT to extract homologous GCP's (Lowe, 2004; Fan et al., 2013). We found that with this approach, high quality GCP's can be extracted after applying an iterative filter based on RMSE. The independent GCP's produced to validate the geo-rectification showed that the majority (> 80%) of the GCP displacements either in x and y directions are in the sub-pixel range (Fig. 2.6a). The reason for larger GCP displacements (> 1000 m) could be that SIFT failed due to the cloud coverage or due to isolated cases of spurious calibration. To derive any meaningful long-term bio-physical variables from the AVHRR LAC data it is crucial to take care of the inter-platform calibration which enables continuity of data over the transition of instruments (Trishchenko et al., 2002; Trishchenko, 2002; Heidinger et al., 2010). Furthermore, the orbital drift (Fig. 2.1) experienced with the earlier NOAA satellites poses a serious challenge to develop high quality time series data (Schneider and Hook, 2010; Riffler et al., 2015). However, the acquisition time correction procedure based on the DTC model shows promising results towards homogenising data from multiple satellites. The inter-platform validation of LSWT shows that the data are indeed comparable and that it is possible to combine the data over time to create longer time series (see Fig. 2.7 and Table 2.3). The median of the inter-platform boxplots (Fig. 2.7) are all below 1.5 °C depicting the superior thermal calibration (Fig. 2.7). Moreover, with the cross-platform validation we showed that the developed LSWT from AVHRR LAC data are close to those from similar sensors on board instruments other than NOAA. We found exceptionally good agreement between NOAA-17 and AATSR, both morning overpasses, with lowest MAE and RMSE of 0.6 °C and 0.85 °C. On the other hand, the highest MAE and RMSE of 1.2 °C and 1.35 °C is reported between NOAA-17 and MODIS-Terra in spite of both being morning overpass data, which needs to be further investigated (Table 2.4). The outliers in both inter-platform and cross-platform boxplots could be due to undetected clouds, unresolved navigational errors or isolated cases of spurious thermal calibration. Though the validation with *in-situ* data show good agreement, it is important to note that the satellite sensor measures temperature of a sub-micron layer between water surface and air which is highly variable according to the meteorological conditions, whereas *in-situ* data represents bulk temperatures in the lake (Wilson et al., 2013; Hook et al., 2003). Due to the skin effect over the lake surface there is a considerable difference between the skin and bulk temperature which may also reflect the actual differences between satellite derived LSWT and the *in-situ* data (Hook et al., 2003; Schneider and Hook, 2010). The accuracy of the derived LSWT also depends upon the viewing zenith angle at which the observation is taken (nadir or wide angle) (Hulley et al., 2011).

The developed method is fully automated which will enable the users to process bulk data in a single workflow. *Rocchini and Neteler (2012)* emphasizes the need for adopting Stallman's four freedom paradigms in ecological research. With the growing interest in open research and publishing the data publicly it is also important to have the software packages used in ecological research in a public domain (*Reichman et al., 2011*). Thus by taking advantage of the development of open source geospatial libraries over the last decade, we successfully revisited the accurate navigation of AVHRR LAC data and implemented a robust methodology to process them. Though the method is robust, manually checking the precision of each of the thousands of images is not practical, there could be unresolved navigational errors in the final time series. The automated cloud masking may also leave undetected cloud pixels as clear sky ones. Though the overall accuracy of SIFT based geo-rectification procedure is found to perform well within the extent of a pixel, it does not always ensure accurate positioning of inland pixels especially in mountainous terrain like the surroundings of Lake Garda. Our procedure does not correct for the surface elevation which may lead to pixel misplacements and could have adversary effects especially on small features like lakes. *Khlopenkov et al. (2010)* explains accurate geolocation of AVHRR historical time series which includes a orthocorrection scheme taking care of the accurate positioning of the pixels in complicated terrains. Hence it is very important to design and perform robust statistical outlier detection before any trend analysis is performed with the developed time series data. Moreover, while we tested this method for a small study area covering Northern Italy, it can be extended to any other area provided the data are in level-1B format and obtained from the NOAA CLASS archive. It is recommended to use high performance computing solutions if you are processing 29 years of AVHRR data over a larger area. An ideal way to compare the performance of different sensors is the simultaneous nadir observations approach, but it is unlikely to obtain such a pair of images over our study area (*Karlsson and Johansson, 2014*).

The flagship NOAA series of instruments ended with NOAA-19, though the AVHRR radiometer continued aboard the MetOp-A and MetOp-B satellites as part of a collaborative meteorological program by the European Space Agency and the European organization for the exploitation of meteorological satellites. However, data from MetOp series are not considered in this study to focus exclusively on NOAA instruments. The methodology also works with solar channels of AVHRR (*Heidinger et al., 2010*), though it is not the scope of this paper to validate the performance of solar calibration. For the SIFT method, bringing a seasonal perspective by using different reference images for multiple seasons (or months) may improve the quality of the extracted GCP's (*Khlopenkov et al., 2010*). It has to be noted that for inter-platform and cross-platform validation, we took the same day observations between 08:00 to 17:00 UTC which is a wide window and explains some of the outliers in spite of performing the acquisition time correction procedure. The gaps in the final time series after

cloud masking and the outlier removal could be statistically filled using methods like harmonic analysis (Roerink *et al.*, 2000) and spatio-temporal modelling (Metz *et al.*, 2014). We avoided night time acquisitions as the specific project objectives demanded processing of day time data and due to non-availability of field data during night time to validate the product. In general, checking the consistency with night time data is important to make complete use of the AVHRR dual acquisitions. It is recommended to use a long-term high frequent *in-situ* data for validating the satellite derived products to match the time of observations. Comparing the long-term trends from both *in-situ* data and satellite derived LSWT would be a recommended approach in testing the temporal stability of AVHRR LAC data over transition of instruments (Schneider and Hook, 2010). But due to lack of high frequent *in-situ* data we avoided comparing the long-term trends, instead we used linear models to check the integrity of derived LSWT. For LSWT estimation, we used a generic non-linear split window equation (Eq. (2.1)), but using satellite specific coefficients provided by Jimenez-Munoz and Sobrino (2008). However, previous studies (Hulley *et al.*, 2011; Riffler *et al.*, 2015) have shown that lake specific coefficients will lead to accurate LSWT estimation. Due to non-availability of lake specific coefficients for the POD satellites and other non AVHRR sensors (Hulley *et al.*, 2011), it is difficult to follow the same method to derive LSWT from all the satellites used in this study. Nevertheless, we obtained high accuracy (see Tables 2.3 and 2.4) by using satellite specific coefficients in a non-linear split window equation.

2.6 Conclusion

To conclude, in this paper we present a new automated method for processing historical AVHRR LAC data using open source geospatial packages (Fig. 2.3). The method is extendable and reproducible to any other geographic location provided the data is taken from NOAA CLASS archive. All the software development related to this work is been published in public archives making it accessible to maximum potential users (Table 2.2). We found that it is possible to resolve the much documented navigational errors with the AVHRR LAC data from POD satellites by implementing a robust geo-rectification procedure using SIFT feature extraction procedure. This automated geo-rectification procedure produced promising results and the validation using independent GCP's reported an overall RMSE of 755.63 m. Further we did a three step validation procedure for the derived time series of brightness temperatures, (i) inter-platform validation by comparing estimated LSWT's between different NOAA instruments; (ii) cross-platform validation by comparing estimated LSWT's between NOAA instruments and other polar orbiting sensors like MODIS and AATSR; (iii) validation using *in-situ* data (Tables 2.3 and 2.4). The overall RMSE reported for inter-platform, cross-platform

and *in-situ* validation are 0.89 °C, 1.05 °C and 0.92 °C respectively. The high accuracy obtained with the validation procedure after performing an acquisition time correction procedure based on DTC model shows that it is indeed possible to develop a homogenised dataset from AVHRR LAC data over multiple NOAA instruments. Our new method, along with the code snippets (Appendices A and B) to process AVHRR LAC data will enable researchers to leverage the availability of historical medium resolution AVHRR data in their research, otherwise impossible to use due to well documented issues (*Brunel and Marsouin, 2000; Privette et al., 1995; Krasnopolsky and Breaker, 1994*). It takes care of multiple sources of errors to a large extent giving an opportunity to develop a homogenised time series of brightness temperatures from AVHRR LAC data. The new methodology opens a new paradigm for researchers to create long-term (29 years) daily temporal datasets of land/water surface temperature derived from satellite imageries. It should be noted that the current study is validated only for lake surface temperature. For land surface, corresponding changes must be made on the split window algorithm and the coefficients. The effects of orbital drift will also vary depending on the land use demanding case specific correction procedures. It is recommended to establish robust statistical filtering methods to remove undetected outliers before performing any trend analysis using the derived data. The data thus derived can play a significant role in research as an alternative to long-term field measurements. Moreover, we could not find any public dataset derived from AVHRR data available at its original spatial resolution. Much of the earlier AVHRR data are underused due to lack of software tools to accurately process them. This study will fill this gap and give scientists the opportunity to use the AVHRR LAC data at its original spatial resolution for their research.

Acknowledgements

The PhD Scholarship of Sajid Pareeth is supported by FIRS>T (FEM International Research School e Trentino) of Fondazione Edmund Mach, Italy. We acknowledge NOAA for providing historical AVHRR LAC data in level-1B format through the CLASS archive. We also acknowledge EUMETSAT for providing European reanalysis data. We acknowledge NASA and ESA for providing MODIS and ATSR series data in public through LAADS and MERCI respectively. We sincerely thank Juan-Carlos Jimenez-Muñoz of University of Valencia, for publishing the satellite specific split-window coefficients and providing us with the missing ones upon request. We acknowledge the communities behind the development and fine-tuning of the open source geospatial packages, especially the ones used here—Pytroll, Orfeo toolbox and GRASS GIS for their valuable support through mailing lists and personal communications. We also thank the anonymous reviewers for their critical reading and suggestions which substantially helped to improve the manuscript.

Chapter 3

Homogenised daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-Alpine lake

3.1 Abstract

Availability of remotely sensed multi-spectral images since the 1980's, which cover three decades of voluminous data could help researchers to study the changing dynamics of biophysical characteristics of land and water. In this study, we introduce a new methodology to develop homogenised Lake Surface Water Temperature (LSWT) from multiple polar orbiting satellites. Precisely, we developed homogenised 1 km daily LSWT maps covering the last 30 years (1986 to 2015) combining data from 13 satellites. We used a split-window technique to derive LSWT from brightness temperatures and a modified diurnal temperature cycle model to homogenise data which were acquired between 8:00 to 17:00 UTC. Gaps in the temporal LSWT data due to the presence of clouds were filled by applying Harmonic ANalysis of Time Series (HANTS). The satellite derived LSWT maps were validated based on long-term monthly *in-situ* bulk temperature measurements in Lake Garda, the largest lake in Italy. We found the satellite derived homogenised LSWT being significantly correlated to *in-situ* data. The new LSWT time series showed a significant annual rate of increase of $0.020\text{ }^{\circ}\text{C yr}^{-1}$ (* $P < 0.05$), and of $0.036\text{ }^{\circ}\text{C yr}^{-1}$ (***) $P < 0.001$) during summer.

3.2 Introduction

Lakes are considered worldwide as sentinels of climate change (Adrian *et al.*, 2009; Williamson *et al.*, 2009). Any change in the surrounding catchment due to climate forcings will reflect on the physical, chemical and biological processes in lakes (Williamson *et al.*, 2009). The surface temperature as being direct and sensitive to long-term changes in thermal structure of the lakes is a good indicator to understand the changes in the lake characteristics (Gerten and Adrian, 2000; Adrian *et al.*, 2009). Thermal variations on the lake surface and epilimnion are crucial for key responses like vertical mixing and stratification, nutrient and oxygen dynamics, as well as spread and geographical expansion of biota (Verburg *et al.*, 2003; Williamson *et al.*, 2009; Adrian *et al.*, 2009). Recent studies on a large number of lakes worldwide indicated a global trend of rapid warming related to climate change (O'Reilly *et al.*, 2015; Schneider *et al.*, 2009; Coats *et al.*, 2006; Kraemer *et al.*, 2015). Global warming has a direct impact on thermal characteristics of lakes, influencing variations in the physical as well as biological characteristics (Fink *et al.*, 2014; Kirill Ya and Filatov, 1999). Geographical expansion of toxic cyanobacteria is reported owing to the warming of lakes (Winder *et al.*, 2009; Paerl and Huisman, 2008; Paerl *et al.*, 2011; Jöhnk *et al.*, 2008). The expansion of tropical bloom-forming cyanobacterium *Cylindrospermopsis raciborskii* (Woloszynska) to the temperate lakes in the mid latitudes is triggered by global warming (Briand *et al.*, 2004). The study "Blooms like it hot" (Paerl and Huisman, 2008) explain climate change as a potent catalyst for the expansion of toxic blooms. In Sweden, annual phytoplankton bloom in larger lakes was found to be anticipated by a month due to the early warming in spring (Weyhenmeyer, 2001).

Lake Surface Water Temperature (LSWT) exhibits a rapid and direct response to climate forcing prominently induced by changes in air temperature, cloud cover, short wave radiation, latitude, and the lakes morphometry (O'Reilly *et al.*, 2015). Detailed synthesis of *in-situ* and satellite derived LSWT over lakes globally, indicated rapid warming during summer. Moreover, LSWT and air temperature trends during summer are found to diverge over many lakes globally which signifies the need of regional studies in understanding the change dynamics (O'Reilly *et al.*, 2015). Investigations carried out on the large lakes south of the Alps using non-parametric tests on volume weighted spring mean temperature derived from long-term *in-situ* data series from 1970 to 2009 showed a significant warming rate of 0.012 - 0.028 °C yr⁻¹, at a rate comparable with that of other European and North American lakes (Salmaso and Mosello, 2010; Schneider and Hook, 2010). The year-to-year fluctuations in the thermal structure were shown to be strictly controlled by large scale atmospheric dynamics between the N-Atlantic and the Mediterranean regions (Salmaso and Cerasino, 2012; Salmaso, 2012). Results of limnological studies carried out in the largest and deepest lakes south of the Alps (Garda, Iseo, Como, Lugano and Maggiore) (Salmaso and Mosello, 2010), pointed out the

need of integrating interdisciplinary approaches to the scientific based long-term monitoring. Lack of long-term *in-situ* data at a high temporal frequency is the main obstacle in identifying long-term trends. In this context, data from remote sensing as a substitute to *in-situ* data could play a key role in limnological studies.

Surface temperature is one of the accurate and reliable measurement using remote sensing (Kirill Ya and Filatov, 1999). Remote sensing of the temperature is based on recording the emitted radiation from earth surface in the spectral domain of 8 - 14 μm (Kuenzer and Dech, 2013). Inverse Planck's law is applied to convert the emitted radiance recorded in the thermal infrared region to Top Of Atmosphere (TOA) Brightness Temperatures (BT). There are multiple approaches to estimate land/water surface temperature from the brightness temperatures (Li *et al.*, 2013). In the case of inland water bodies like large lakes, the most common approach is the split-window technique where the difference between the two adjacent thermal channels (10.5 - 11.5 μm , 11.5 - 12.5 μm) is taken as a measure of atmospheric attenuation to derive the Surface Temperature (ST) (Maul and Sidran, 1971). The accuracy of the derived ST depends on the split-window coefficients, which in turn is based upon multiple parameters like spectral response function and emissivity of the channels, column water vapour in the atmosphere and View Zenith Angle (VZA) of the sensor. These coefficients are derived by regressing simulated BT's from radiative transfer models like MODTRAN against the atmospheric profiles. The coefficients are generally derived on a regional or global level for the estimation of Land Surface Temperature (LST) and Sea Surface Temperature (SST) (Czajkowski *et al.*, 1998; Li *et al.*, 2001; Jimenez-Munoz and Sobrino, 2008). For lakes, both LST and SST based algorithms are used interchangeably and with higher accuracy of RMSE between 0.5 - 1.5 $^{\circ}\text{C}$ (Hook *et al.*, 2003; Oesch *et al.*, 2005, 2008; Schneider *et al.*, 2009; Schneider and Hook, 2010; Riffler *et al.*, 2015). Furthermore, lake and sensor specific constants are published for an exhaustive list of global lakes by Hulley *et al.* (2011).

There are a multitude of LST and SST datasets derived from multiple sensors publicly available such as Pathfinder from AVHRR data (Kilpatrick *et al.*, 2001), EuroLST (Metz *et al.*, 2014), GlobTemperature from ATSR series (Ghent, 2012) and the MODIS temperature products. ArcLakes (MacCallum and Merchant, 2012) and the LSWT data by (Riffler *et al.*, 2015) are two lake specific products derived exclusively from satellite data for a set of selected lakes. A global database of lake summer surface water temperatures for 291 lakes combining both *in-situ* and satellite data was recently published by Sharma *et al.* (2015). Nevertheless, the products are often based on selected sensors aboard multiple instruments with different specifications and acquisition times. The algorithm behind surface temperature retrieval also differ among the products. In addition, many temperature products from moderate resolution (~ 1 km) sensors are re-sampled to coarser resolution (~ 5 km) which results in

missing spatial coverage or even complete data loss at lake locations. Since multiple datasets cover different time frames, we need to combine them in order to generate a time series appropriate for trend analysis in the context of climate impact studies. But combining these data from multiple sources without applying acquisition time correction and homogenisation would result in spurious trends.

Our aim was to develop a new methodology to derive daily homogenised LSWT data from multiple moderate resolution sensors. As a result, we developed a new homogenised LSWT time series for Lake Garda using the dual thermal channels from the sensors - AVHRR/2 (NOAA-9/11/12/14), AVHRR/3 (NOAA-16/17/18/19), ATSR1 (ERS-1), ATSR2 (ERS-2), AATSR (Envisat) and MODIS (Aqua/Terra). The new time series offers 30 years of daily LSWT from 1986 to 2015 developed using split-window algorithm with satellite specific coefficients published by *Jimenez-Munoz and Sobrino (2008)*, acquisition time/orbital drift corrected, spatially sampled at 1 km resolution and gap filled using harmonic analysis. To our knowledge, this is the first time a homogenised LSWT was developed merging data from 13 satellites. The new method is reproducible and extensible to other lakes globally, provided that satellite data is available. The entire methodology is implemented using open source software packages. Hereafter, in the following sections "AVHRR" represents both AVHRR/2 and AVHRR/3 sensors and "ATSR series" represents ATSR1, ATSR2 and AATSR unless explicitly specified.

3.3 Study area

Lake Garda is located east of the southern sub-Alpine region between the coordinates - Longitude: $45.44^{\circ}E$ - $45.92^{\circ}E$; Latitude: $10.44^{\circ}N$ - $10.91^{\circ}N$. With a surface of 368 km^2 and a volume of 49 km^3 , it is the largest lake in Italy. Along with the lakes Iseo, Como, Maggiore and Lugano it forms part of the group of deep lakes located south of the Alpine chain. These lakes are an important economic resource in Europe. Besides representing a major tourist destination, their waters are also used for drinking water purposes. Lake Garda is divided into two basins, northwestern and southeastern, which are divided by an underwater ridge connecting the Sirmione peninsula with Punta S. Vigilio (Fig. 3.1). The northwestern basin is large and deep, the shores descend at sharp slopes and the bottom spreads over 20 km at depths from 300 to 350 m (maximum depth). The maximum depth in the southeastern basin is around 80 m, whereas the shape is nearly conical. From a practical point of view, the lake can, however, be divided into two major zones separating the deeper northern area from the shallower southern area (mostly $< 150 \text{ m}$). This distinction is consistent with the observations in this study *Bresciani et al. (2011)* where, using temporal multi-spectral

satellite images, found significant spatial and temporal variability in the concentrations of chlorophyll-*a* in the northern and southern areas of Lake Garda. The low ratio between the surface of the catchment area and the surface of the lake (6.1) explains the long theoretical water renewal time of about 27 years in comparison with other deep lakes in the area.

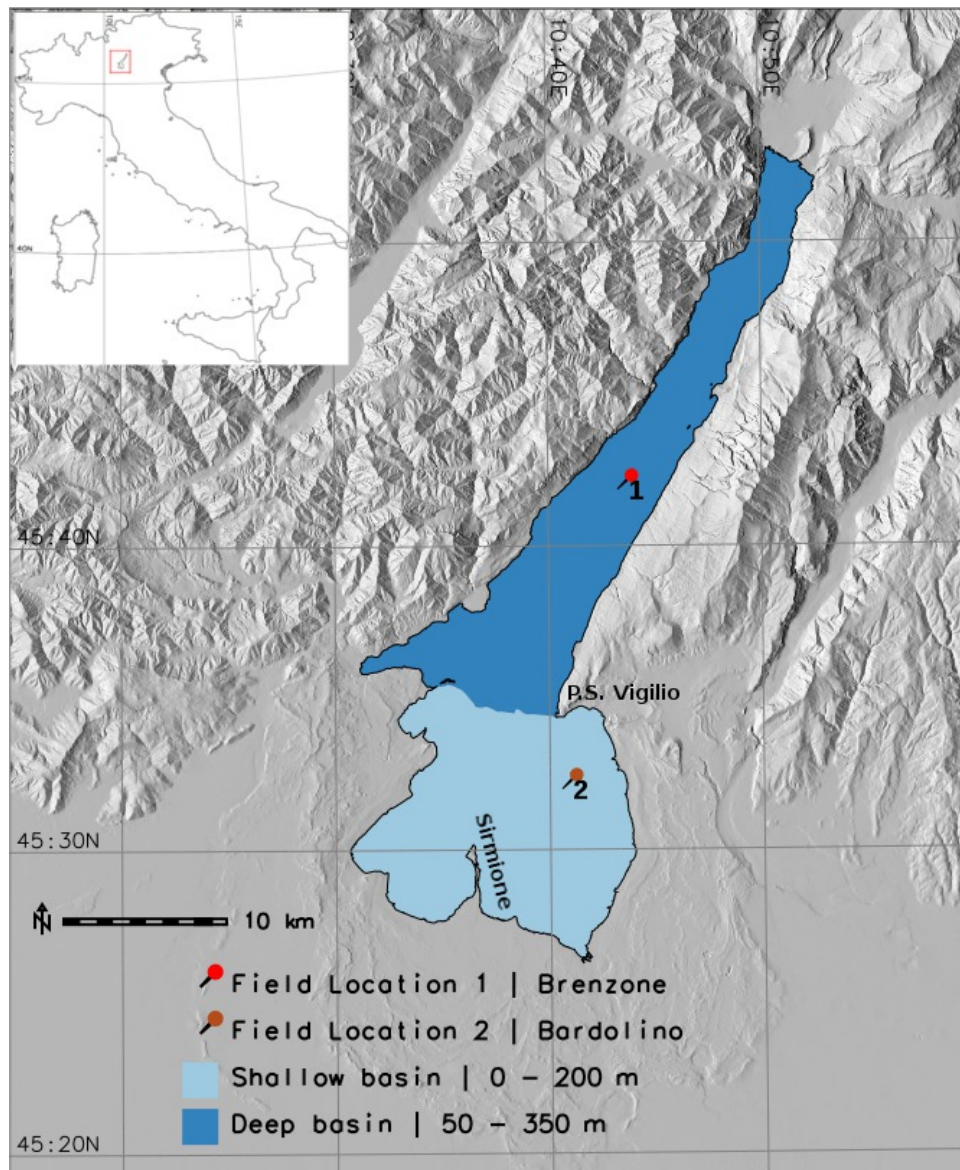


Figure 3.1: Study area map of Lake Garda; dark blue area depicts the deep basin and light blue area depicts the shallow basin. Push pins represent locations of *in-situ* water temperature monitoring. The maps were generated using the software GRASS GIS 7.0 Neteler and Mitasova (2008) (URL - <https://grass.osgeo.org/grass7/>)

3.4 Methods

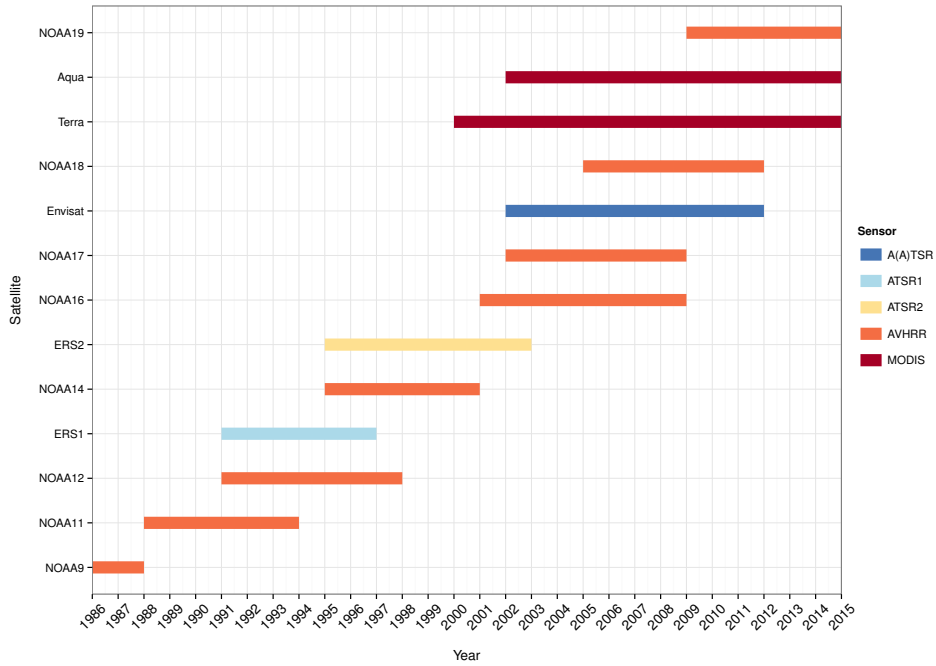
3.4.1 Data

In this study we used multiple moderate resolution sensor data to estimate long-term LSWT's over the past 30 years. We used data acquired by six sensors - AVHRR/2, AVHRR/3, ATSR1, ATSR2, A(A)TSR and MODIS which were on board 13 satellites (Fig. 3.2a). These sensors offer dual thermal channels at 10.5 - 11.5 μm and 11.5 - 12.5 μm with a spatial resolution of approximately 1 km. This was ideal for deriving surface water temperature using a split-window algorithm. The sensors AVHRR and MODIS provide daily images, while ATSR series provide data every three days. We downloaded and processed level-1B data of AVHRR, ATSR series and MODIS from Comprehensive Large Array-data Stewardship System (CLASS - <http://www.class.ncdc.noaa.gov/>), Mercia data archive (<http://ats-merci-ds.eo.esa.int/merci/welcome.do>) and Level 1 and Atmosphere Archive and Distribution System (LAADS - <https://ladsweb.nascom.nasa.gov/>) respectively for the study area. For AVHRR, we used Local Area Coverage (LAC) data available at its original spatial resolution of 1.1 km. For ATSR series, we downloaded and processed nadir viewing TOA data. The level-1B swath products MYD021KM and MOD021KM for Aqua and Terra satellites respectively are used to extract dual thermal channels from MODIS sensor. Fig. 3.2b shows the year wise distribution of data from multiple satellites. We processed the AVHRR and MODIS data using Pytroll libraries (<http://www.pytroll.org/>). Furthermore, we used BEAM software provided by European Space Agency to process the ATSR series of data (<http://www.brockmann-consult.de/cms/web/beam/>). For validation of the satellite derived water temperatures, long-term monthly *in-situ* data between 1991 - 2015 collected from two locations in Lake Garda were used. We used bulk temperature (0 - 0.5 m) measured at two locations - Brenzone in the deep basin and Bardolino in the shallow basin (see Fig. 3.1).

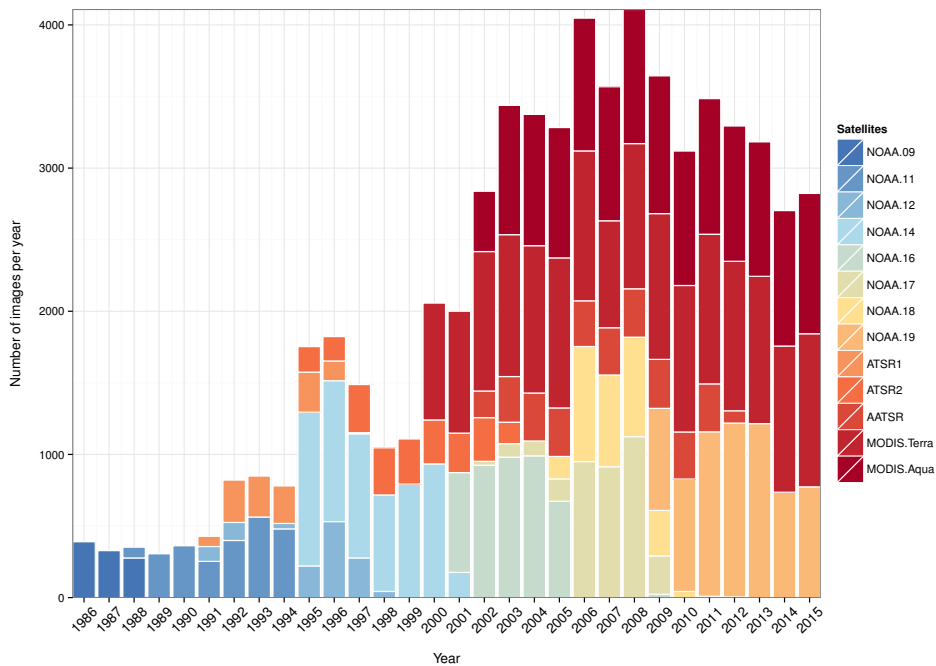
3.4.2 Pre-processing

The pre-processing of thermal channels at level-1B include thermal calibration to convert from radiances to brightness temperatures using inverse Planck's law followed by removal of cloud pixels. The data from MODIS and ATSR series were calibrated using the coefficients provided in the header. For AVHRR, the calibration procedure was much more complicated due to (well documented) navigational errors with the older NOAA satellites (*Privette et al.*, 1995). Despite on board thermal calibration in NOAA instruments, the calibration data were prone to solar contamination and atmospheric attenuation of the signal and satellite decays

Figure 3.2: Availability of moderate resolution satellite data between 1986 and 2015



(a) Time line of the satellites in orbit



(b) Distribution of level-1B data downloaded for the study area from 1986 to 2015 obtained from respective archives - CLASS (AVHRR), LAADS (MODIS) and Mercr (ATSR series) separated by different instruments

(Trishchenko et al., 2002). In this study, we used the geometrically corrected time series of brightness temperatures derived from dual thermal channels of AVHRR LAC data from 1986 to 2015 developed by Pareeth et al. (2016). Furthermore, we considered only those pixels acquired at a zenith angle less than 45 degrees. All the further data processing was

restricted to lake pixels masked using a Lake Garda boundary layer. For the cloud masks of level-1B AVHRR LAC and MODIS data, we used the Separation of Pixels Using Aggregated Rating over Canada (SPARC) algorithm developed by *Khlopenkov and Trishchenko (2007)*, originally implemented for creating cloud mask over Canada. We used two relevant tests to remove thick clouds and thin cirrus above the lakes, i) brightness temperature test (T-test) using channel T_i and ii) thin cirrus test (C-test) which uses the difference between channel T_i and T_j . The T-test uses channel T_i brightness temperature and compares it with a dynamic threshold which is the surface skin temperature data of the corresponding day and time derived from climatic models. In this study, we replaced the North American regional reanalysis data with European centre for medium-range weather forecasts (*Dee et al., 2011*) following the successful implementation over Europe by *Riffler et al. (2015)*. The SPARC algorithm is implemented using raster processing tools in GRASS GIS 7.0 (*Neteler et al., 2012; Neteler and Mitasova, 2008*). For ATSR series, we used the quality layer provided with the data to remove the cloud pixels.

3.4.3 Deriving LSWT using split-window algorithm

From the brightness temperatures derived by thermal calibration, the next step was to estimate LSWT's using split-window algorithm. We used a non-linear split-window Eq. 3.1 to estimate the surface temperature as proposed by *Jimenez-Munoz and Sobrino (2008)*:

$$T = T_i + c_1(T_i - T_j) + c_2(T_i - T_j)^2 + c_0 + (c_3 + c_4W)(1 - \epsilon) + (c_5 + c_6W)\Delta\epsilon \quad (3.1)$$

where $c_0 - c_6$ are split-window coefficients, T_i and T_j are at-sensor brightness temperatures derived from the dual thermal channels, ϵ is mean emissivity, $\Delta\epsilon$ is emissivity difference and W is the total atmospheric water vapour column (g/cm^2). For water surface, the mean emissivity is close to unity. As our study is dealing with lakes, the Eq. 3.1 was modified by assuming ϵ as one and $\Delta\epsilon$ to be zero. Hence the modified Eq. 3.2 was used in this study to retrieve LSWT from dual thermal channels:

$$T = T_i + c_1(T_i - T_j) + c_2(T_i - T_j)^2 + c_0 \quad (3.2)$$

We used satellite specific split-window coefficients published by *Jimenez-Munoz and Sobrino (2008)*. These coefficients were derived by statistical minimization from a simulated database of brightness temperatures and atmospheric profiles obtained from MODTRAN radiative

transfer model, emissivity spectra extracted from spectral libraries and the spectral response functions (*Jimenez-Munoz and Sobrino, 2008*).

After estimating the LSWT, we applied two levels of filters to remove outliers due to undetected clouds and spurious calibrations. First, we applied a global filter based on a minimum and maximum thresholds derived from the long-term *in-situ* data. Thus we discarded any LSWT values beyond the range 6.5 - 29 °C. We further applied an advanced filter based on inter-quartile range of LSWT's derived every 16 days (*Neteler, 2010; Metz et al., 2014*). The lower and upper thresholds for this filter was computed using the Eq. 3.3 and Eq. 3.4 respectively:

$$Lower_threshold = 1^{st}quartile - 1.5(3^{rd}quartile - 1^{st}quartile) \quad (3.3)$$

$$Upper_threshold = 1^{st}quartile + 1.5(3^{rd}quartile - 1^{st}quartile) \quad (3.4)$$

where the 1stquartile and 3rdquartile were derived climatologically from the LSWT data every 16 days over the years 1986 to 2015. This filter was then applied using the derived thresholds to those LSWT data of corresponding time frames.

3.4.4 Homogenising LSWT from multiple satellites

We considered all the satellite observations for a single day over a wide window of time; from 08:00 to 17:00 UTC (Fig. 3.3). Fig. 3.3 clearly shows the orbital drift of earlier satellites - NOAA-9/11/12/14/16.

We applied a physical model based on typical pattern technique proposed by *Jin and Treadon (2003)* to standardize the acquisition time to 12:00 UTC. In this approach, monthly diurnal cycles were used as a reference pattern to correct for the varying observation times to produce a standard observation time. The monthly diurnal cycles computed from hourly averages were derived from the existing satellite observations of LSWT and were limited to day time (Fig. 3.4). We followed the Diurnal Temperature Cycle (DTC) model to derive the diurnal variations of LSWT based on *Duan et al. (2014)*. The model used was described by the Eq. 3.5:

$$T_s(t) = T_0 + T_a \cos(\pi/\omega(t - t_m)) + T_b \sin(\pi/\omega(t - t_m)) \quad (3.5)$$

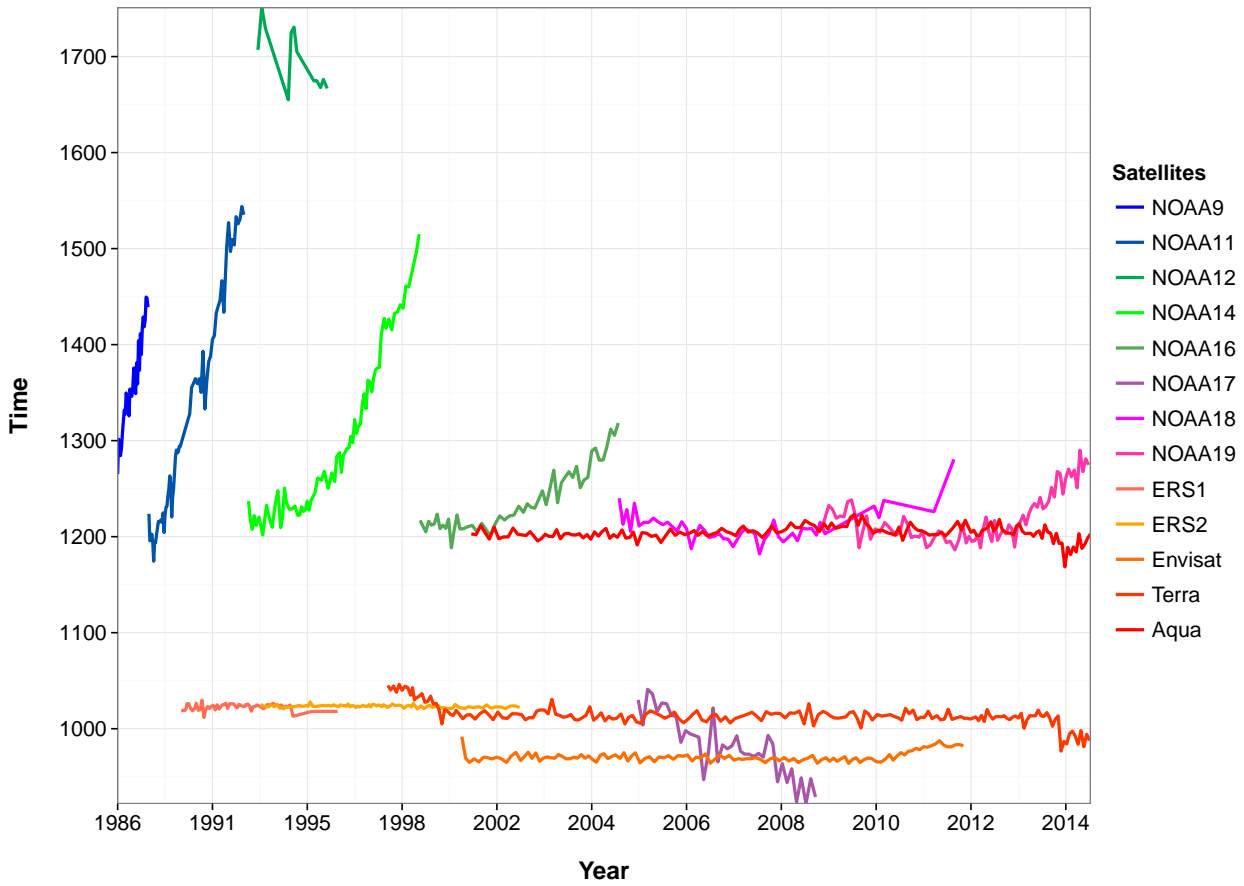


Figure 3.3: Plot of variation in observation times of all the satellites used in this study; Note the large orbital drifts of the earlier NOAA-9/11/12/14 instruments.

with

$$\omega = 4/3(t_m - t_{sr}) \quad (3.6)$$

where T_0 is the residual temperature at sun rise; T_a and T_b are temperature amplitudes; $T_s(t)$ is surface temperature at time t ; t_m is the time at which temperature is maximum; t_{sr} is the time of sun rise; ω is calculated using the Eq. 3.6. T_0 , T_a and T_b are obtained by statistically fitting the model given in the Eq. 3.5. Finally, to apply the time correction to all the LSWT data, for each image with an observation time t , $T_s(t)$ and $T_s(12)$ were calculated using the model as given in Eq. 3.5. The correction factor was then computed using Eq. 3.7 to correct the LSWT to a standard time at 12:00 UTC. A condition based approach is used for the correction as given in Eq. 3.8 and Eq. 3.9:

$$cf = abs(T_s(t) - T_s(12)) \quad (3.7)$$

$$T_c = T_{\text{ori}} - cf; \quad T_s(t) > T_s(12) \quad (3.8)$$

$$T_c = T_{\text{ori}} + cf; \quad T_s(t) < T_s(12) \quad (3.9)$$

where cf is the correction factor, T_c is the corrected LSWT to a standard time 12:00 UTC and T_{ori} is the actual LSWT.

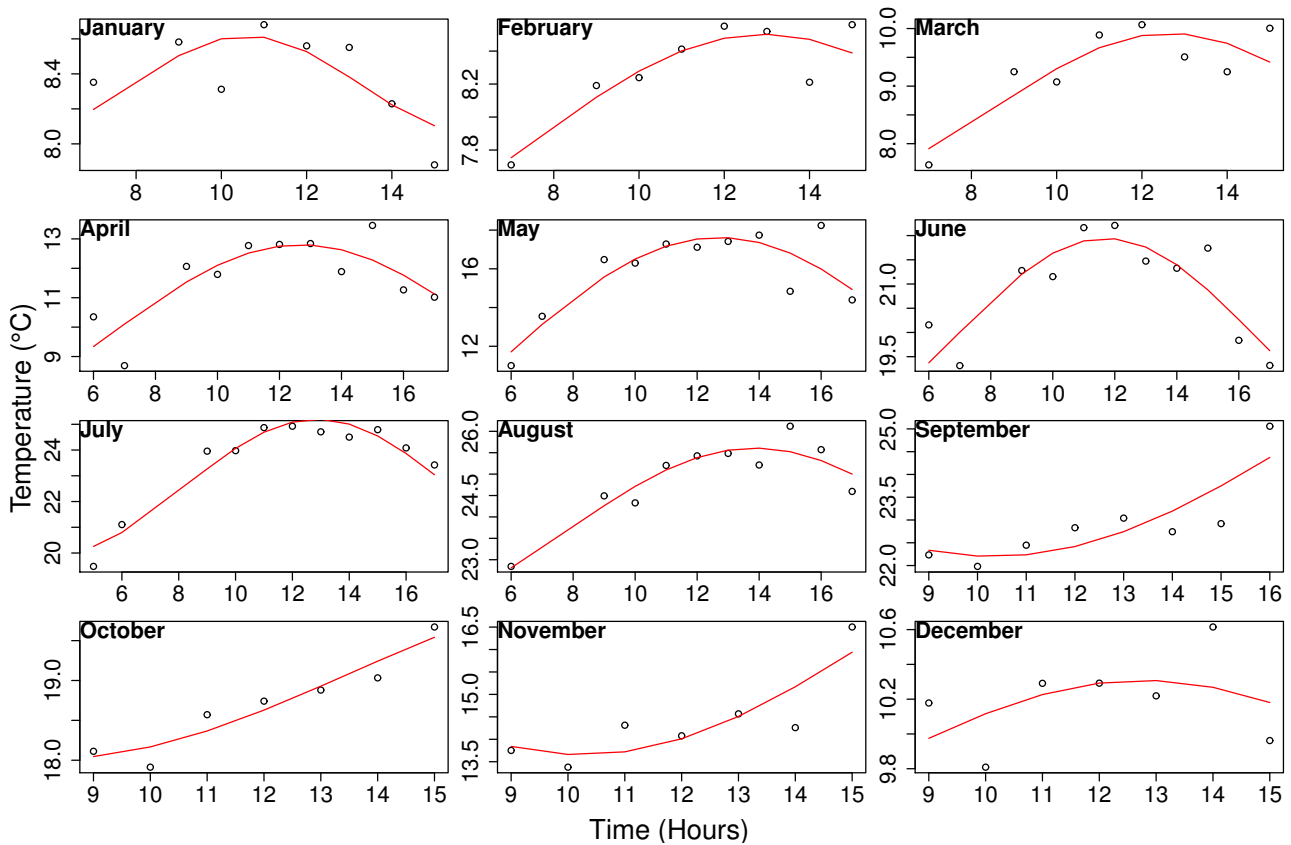


Figure 3.4: Monthly diurnal cycles during day time estimated using long-term hourly climatologies from satellite observations. The points in plot are the long-term hourly averages for a particular month, and the red line is DTC model fit.

In order to derive one homogenised LSWT per day, the subsequent step was to merge by averaging the single day observations from multiple satellites. For this, we first performed a satellite specific linear regression analysis between LSWT and *in-situ* data to compare the slope of regression lines between different satellites. We performed the analysis on the deep and the shallow basins separately. The slopes varied between 0.85 and 1.1, depicting that the data used for the models were similar to each other and exhibited similar residual patterns. We used the *r.mapcalc* module in GRASS GIS 7.0 software to apply the time correction model to the LSWT data.

3.4.5 Gap filling using harmonic analysis

The daily LSWT data after time correction and averaging over the same day still suffered from gaps due to cloud cover. We used Harmonic ANalysis of Time Series (HANTS) to reconstruct the gaps in the LSWT data. The harmonic analysis decomposes a time series into several sinusoidal functions with unique amplitude and phase values (*Jakubauskas et al.*, 2001). It applies a least square fitting procedure based on the derived harmonics (*Xu et al.*, 2013). The fitting procedure is repeated until the LSWT values with large deviations are removed from the fitted curve (*Roerink et al.*, 2000). Hence, the role of HANTS is twofold in this process; i) to remove the outliers, and ii) fill the gaps with fitted values. The fitting procedure based on HANTS is explained well by *Xu et al.* (2013) and *Roerink et al.* (2000). We implemented HANTS based reconstruction on LSWT images using the *r.hants* addon in GRASS GIS 7.0 software (*Metz*, 2015).

3.4.6 Temporal database and trend analysis

The gap filled daily LSWT maps from 1986 to 2015 were then imported into a spatio-temporal database using TGRASS modules in GRASS GIS 7.0. TGRASS is an advanced set of modules which perform time series analysis on spatio-temporal data (*Gebbert and Pebesma*, 2014). We developed annual and seasonal climatologies from the daily LSWT data to analyse the long-term trends. The seasonal climatologies were developed by aggregating the LSWT over four seasons; winter (December/January/February), spring (March/April/May/), summer (June/July/August) and fall (September/October/November). We applied non-parametric tests on annual and seasonal mean LSWT separately for deep and shallow basins to detect the long-term trends. In particular, we used the Mann-Kendall test to identify the presence of monotonic upward or downward trends. We computed the Theil-Sen slope for the quantitative estimation of significant trends (*Sen*, 1968). We used the Durbin-Watson test to detect potential serial correlation issues in the time series and checked for its statistical significance. To avoid mixed pixels with land along the shore, we used an inner lake buffer considering only water pixels while extracting the annual and seasonal climatologies (Fig. 3.5). Furthermore, for comparative analysis, we computed the same statistics on long-term *in-situ* data and satellite derived LSWT from AVHRR sensor by *Riffler et al.* (2015) for Lake Garda.

3.5 Results

We processed a total of 62,799 level-1B input images in order to derive daily LSWT's. A single LSWT map has 363 pixels of 1 km resolution representing the entire Lake Garda. After applying the inner mask to remove the edge pixels, a total of 223 pixels were considered for deriving long-term climatologies (Fig. 3.5). The aggregated seasonal and annual maps were developed from daily data. To demonstrate the spatial variability of the new LSWT maps, Fig. 3.5 shows the aggregated summer mean temperature maps for the years 1992 and 2003, along with a temperature difference map. The year 1992 had a relatively cool summer, while the year 2003 had a record warm summer due to a heatwave all across Europe. Harmonic analysis filled data gaps by repetitively applying a least square fit and removing the outliers until the remaining data was within the valid range. The amplitude of the first harmonic in all the years was larger than the other harmonics depicting a strong unimodal pattern of LSWT. To demonstrate the data reconstruction using HANTS, Fig. 3.6 shows the smoothed LSWT plotted over the daily averaged LSWT after homogenisation for the year 2003.

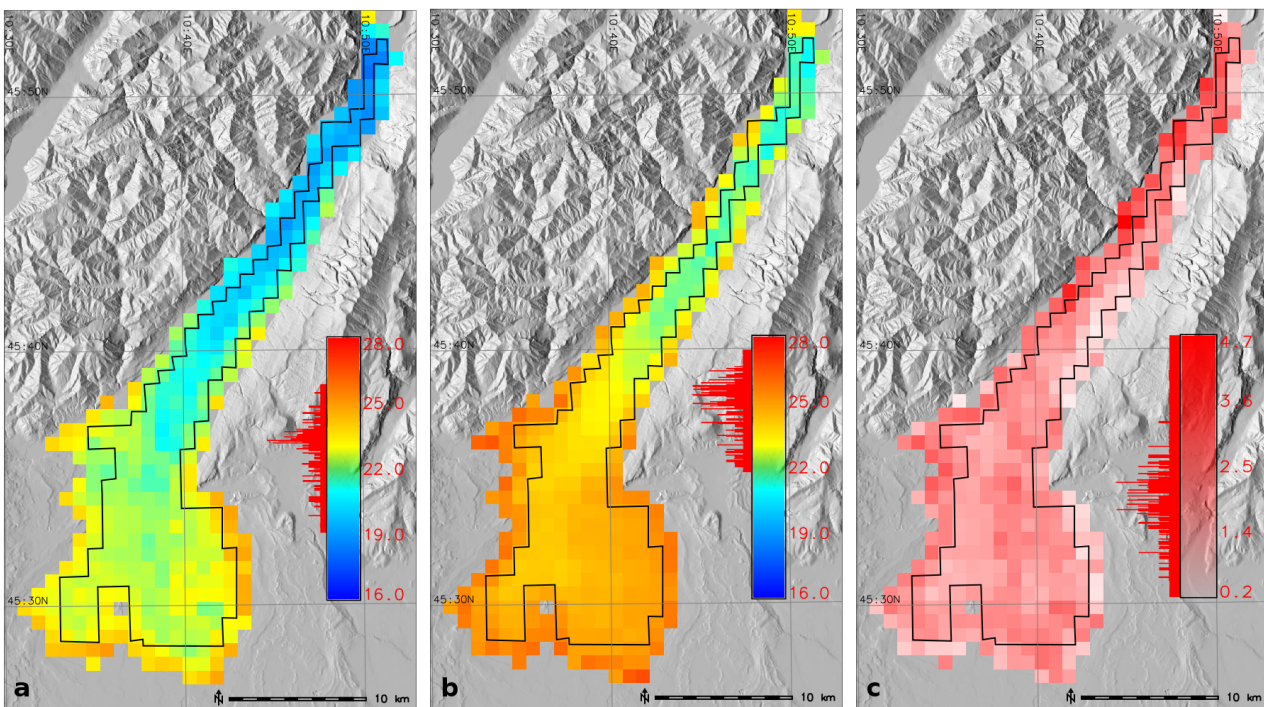


Figure 3.5: a) Summer mean LSWT of year 1992, b) summer mean LSWT of year 2003, c) difference map between summer means of 2003 and 1992. The boundary layer shown in black over Lake Garda is the inner buffer used to mask out the edge pixels. The maps were generated using the software GRASS GIS 7.0 *Neteler and Mitasova (2008)* (URL - <https://grass.osgeo.org/grass7/>)

Table 3.1: RMSE in °C reported between same day observations of different pair of satellites for the deep and shallow lake basin. Number inside brackets represents N - Number of observations.

Satellite pair	RMSE (Deep basin)	RMSE (Shallow basin)
NOAA17/NOAA18	0.74(71)	0.71(85)
NOAA18/NOAA19	0.45(10)	0.38(9)
NOAA11/ERS-1	0.95(37)	0.91(37)
NOAA14/ERS-1	0.78(19)	0.82(16)
NOAA14/ERS-2	0.78(90)	0.89(90)
NOAA16/ERS-2	0.76(38)	0.93(37)
NOAA16/Envisat	0.85(58)	0.87(50)
NOAA17/Envisat	0.67(26)	0.57(29)
NOAA18/Envisat	0.73(84)	0.76(95)
NOAA19/Envisat	0.66(75)	0.75(78)
ERS-2/Envisat	0.79(37)	0.51(33)
NOAA14/Terra	1.19(23)	1.28(22)
NOAA12/NOAA14	1.11(22)	1.05(20)
NOAA16/Terra	1.11(147)	1.17(139)
NOAA17/Terra	1.20(68)	1.16(76)
NOAA18/Terra	0.98(205)	1.02(220)
NOAA19/Terra	1.07(323)	1.12(310)
ERS-2/Terra	0.94(55)	0.87(56)
Envisat/Terra	1.18(212)	1.05(243)
NOAA16/Aqua	0.97(138)	0.97(130)
NOAA17/Aqua	1.06(63)	0.99(70)
NOAA18/Aqua	0.82(252)	0.73(256)
NOAA19/Aqua	0.97(397)	0.86(380)
ERS-2/Aqua	0.99(20)	0.78(20)
Envisat/Aqua	1.10(215)	0.98(211)
Terra/Aqua	0.78(858)	0.86(809)

3.5.1 Validation

To validate the new LSWT data, we performed a cross platform comparison of same day observations between the satellite pairs, followed by regression analysis between final homogenised LSWT and the *in-situ* data. The RMSE of the various cross-platform pairs for both deep and shallow basins varied between 0.38 °C and 1.28 °C (Table 3.1). The minimum RMSE of 0.38 °C was reported for the satellite pair NOAA18/NOAA19 for both the deep and shallow basins. The highest RMSE of 1.2 °C was reported for the pair NOAA17/Terra in the deep basin and of 1.28 °C for NOAA14/Terra in the shallow basin. An average RMSE of 0.88 °C was reported for all the satellite pairs taking into consideration both the basins. Furthermore, we validated LSWT derived from individual satellites with the corresponding *in-situ* data. We were only able to use data from 11 satellites for the validation due to the non-availability of matching *in-situ* data against NOAA9 and NOAA12 (Table 3.2). An average RMSE of 0.86 °C and 0.94 °C was obtained for the deep and shallow basins respectively.

Table 3.2: RMSE in °C reported at deep basin and shallow basin from absolute difference of LSWT between *in-situ* data and satellite derived LSWT. Number inside brackets refer to N - the number of same day observations.

Satellites	RMSE (Deep basin)	RMSE (Shallow basin)
NOAA11	0.33(9)	0.93(9)
NOAA14	0.54(14)	0.51(17)
NOAA16	0.33(8)	0.91(8)
NOAA17	0.37(4)	0.86(3)
NOAA18	1.25(11)	0.90(13)
NOAA19	0.67(16)	1.14(14)
ERS-1	0.41(5)	0.74(4)
ERS-2	0.47(14)	0.70(12)
Envisat	0.50(10)	0.50(11)
Terra	0.92(41)	1.04(36)
Aqua	0.89(46)	1.17(38)

For the deep basin, LSWT derived from NOAA16 had the lowest RMSE against *in-situ* data (0.42 °C), while NOAA18 had the highest RMSE (1.29 °C). In the shallow basin, the lowest RMSE was reported for ERS-2 (0.67 °C), whereas the highest RMSE was reported for Aqua against *in-situ* data (1.18 °C). An average RMSE of 0.92 °C was estimated between satellite derived LSWT and *in-situ* observations. The regression analysis between the final homogenised LSWT and the *in-situ* data revealed a coefficient of determination (R^2) of 0.98 for both basins. The mean RMSE estimations were 0.83 °C and 0.75 °C for deep and shallow basins respectively (Fig. 3.7).

3.5.2 Long term warming trends

We found a significant (* $P < 0.05$) warming trend for lake mean LSWT at an annual rate of 0.020 °C yr⁻¹ (Fig. 3.8). For the deep basin, we found a similar warming trend (0.014 °C yr⁻¹; * $P < 0.05$) for annual mean LSWT, but for shallow basin there was no significant annual trend. For summer months, we report a significant warming trend at the rate of 0.036 °C yr⁻¹ (** $P < 0.001$) for the lake mean LSWT (Fig. 3.8). We report a similar increasing rate for the mean LSWT over the deep basin (0.039 °C yr⁻¹; ** $P < 0.001$), while no significant warming was found for the shallow basin. We did not observed any significant trends during the other seasons. Furthermore, standard deviation of 0.32 °C and 0.53 °C were estimated from the time series of annual and summer mean LSWT respectively. For comparison, we performed summer and annual trend analysis on the lake mean temperature computed from *in-situ* data and recently published satellite derived LSWT (Riffler *et al.*, 2015). We obtained a significant warming trend at the rate of 0.039 °C yr⁻¹ (* $P < 0.05$) and 0.027 °C yr⁻¹ (* $P < 0.05$) for both

summer and annual time series from the data by *Riffler et al.* (2015), while no significant warming was found from the *in-situ* data (Fig. 3.9).

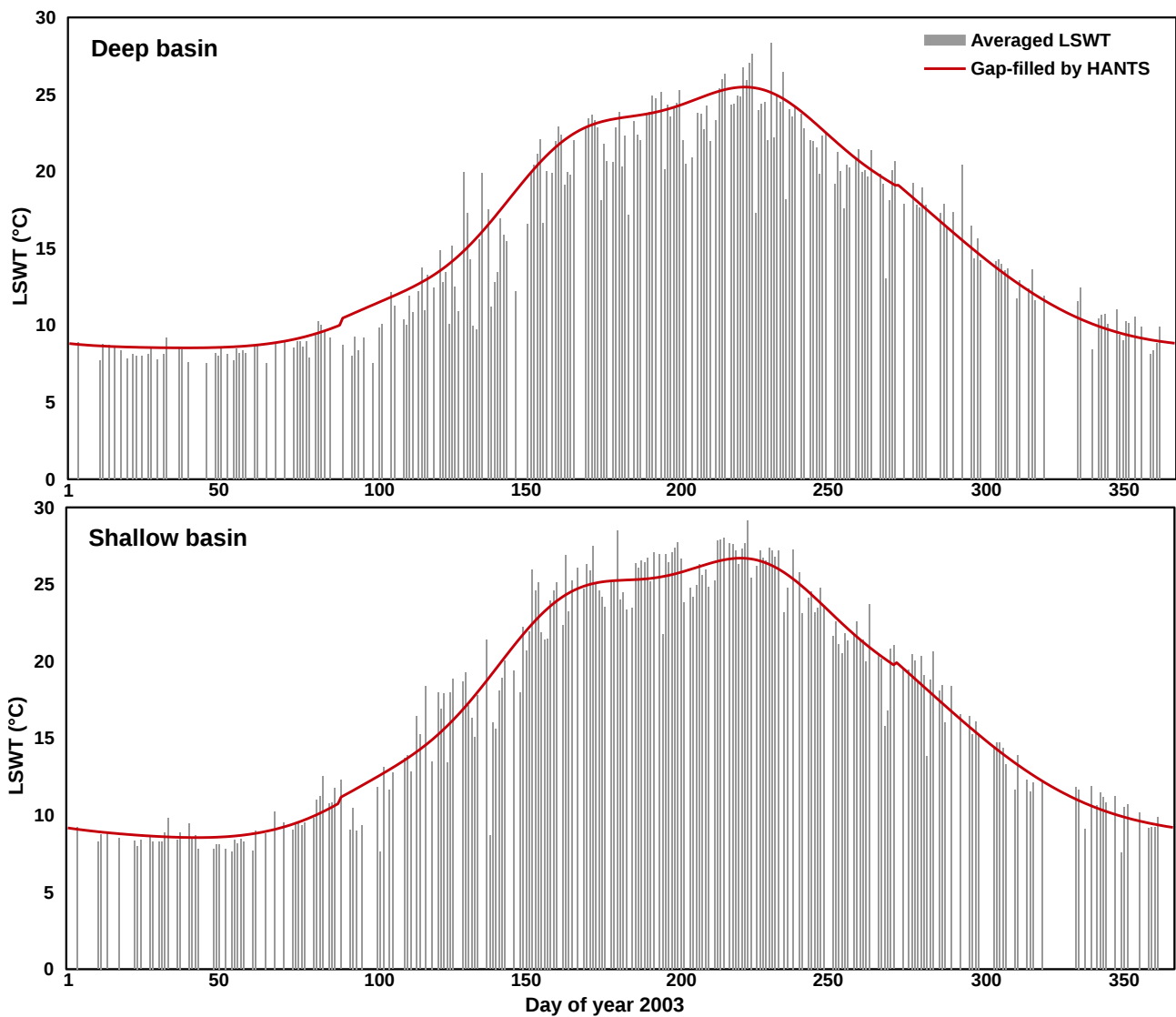


Figure 3.6: Plots showing gap-filled reconstructed LSWT using HANTS (red line) over the averaged homogenised LSWT (grey bars) from multiple observations over a particular day of the year. The daily LSWT time series from the deep and shallow basins for the year 2003 is shown as an example.

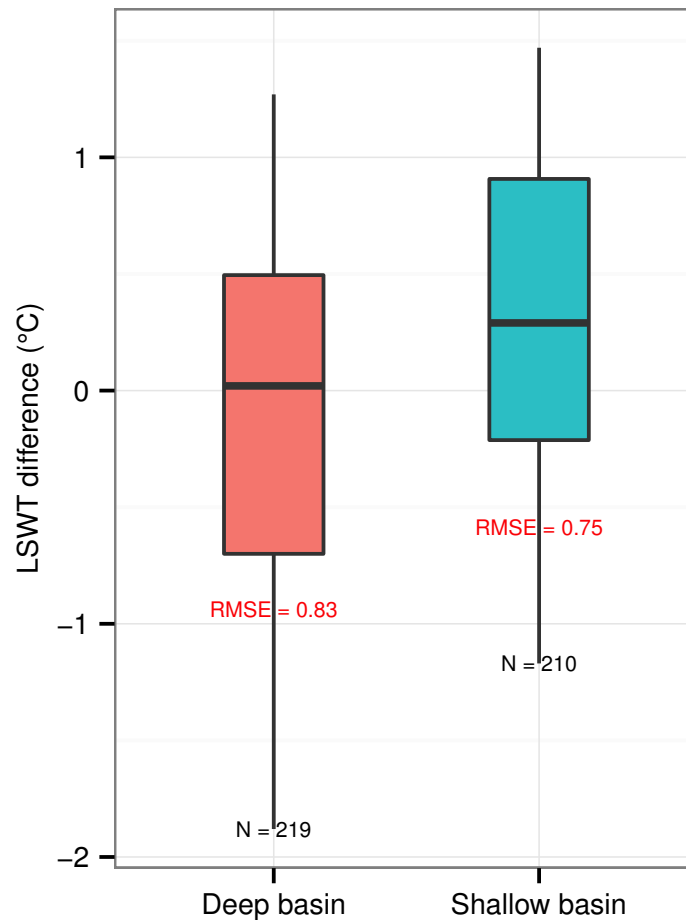


Figure 3.7: Boxplots of lake mean LSWT difference between final homogenised LSWT and corresponding in-situ data. The values shown in black over the lower whisker represents N and the values in red is the reported RMSE in °C.

3.6 Discussion

Remotely sensed satellite data offer a great alternative to *in-situ* data. We demonstrated the usability of satellite derived LSWT in estimating long-term annual and seasonal trends of lake mean LSWT. The inevitable breaks in historical satellite data due to different life periods and the overpass time of multiple satellites is the main challenge in developing a continuous time series (Fig. 3.3). In order to make a long-term time series of any bio-geophysical parameter from satellite data, it is necessary to proceed with a combination of different time series, which demand homogenisation to correct for the different acquisition times. Hence, we developed a workflow to derive daily LSWT maps of Lake Garda for the last 30 years (1986 - 2015) at 1 km spatial resolution recorded by six moderate resolution sensors on board of 13 different satellites. For AVHRR data acquired from multiple NOAA instruments, we used

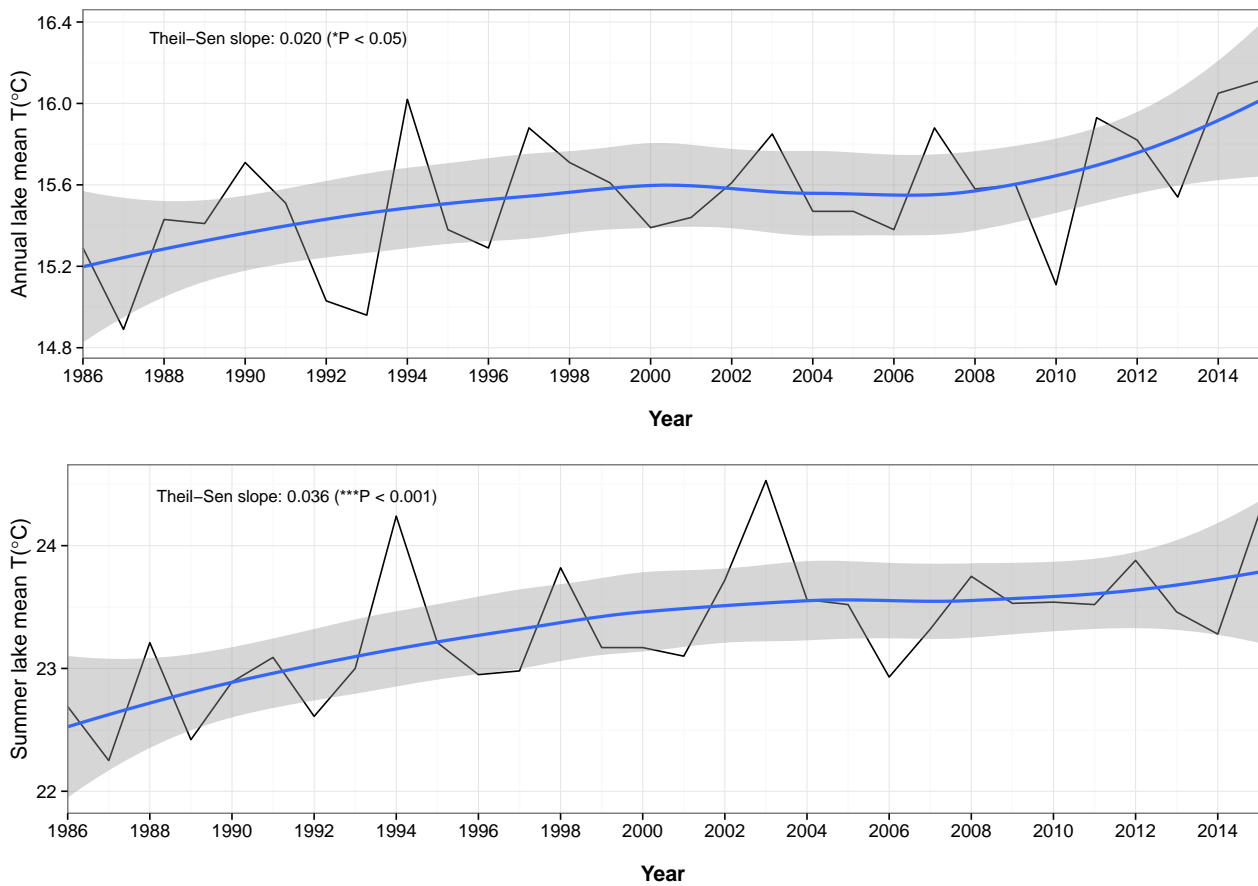


Figure 3.8: Annual (top) and summer (bottom) trends of lake mean LSWT derived from the new homogenised daily LSWT between 1986 to 2015. Data were smoothed using the loess interpolation (blue line). The gray area is the 95% confidence interval.

the calibration techniques which take into account the inter-satellite calibration (*Trishchenko et al., 2002; Trishchenko, 2002; Heidinger et al., 2010*). For the ATSR series, the calibration coefficients provided with the data were of superior quality ensuring continuity of data with transition of the instrument (*Schneider et al., 2009*). Moreover, the split-window technique with satellite specific coefficients used in this study could derive LSWT's at an average RMSE of 0.88 °C on a single day.

We found similar RMSE (< 1 °C) before and after applying time correction procedures between satellite derived LSWT and corresponding *in-situ* data. This could be due to the fact that lake surface temperature of a deep sub-Alpine lake is a slow changing parameter during the day. Precisely, the temperature measured at different times during the day was close enough to exhibit any noticeable change in accuracy indicators before and after the time correction procedure. Nevertheless, for the older satellites, like earlier NOAA instruments which underwent considerable orbital shifts during their course in orbit, this step was crucial. We argue that the acquisition time correction is necessary while combining data from

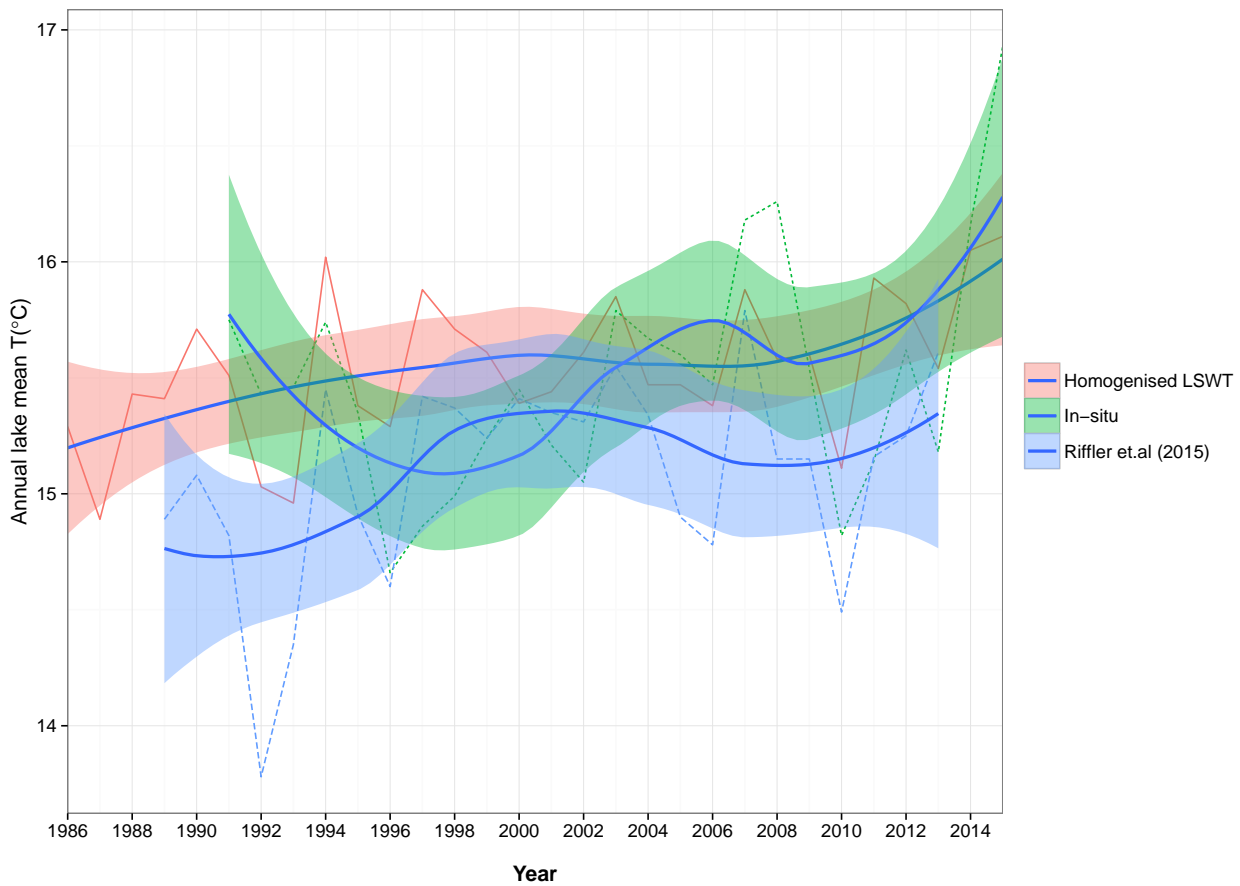


Figure 3.9: Annual trends of lake mean LSWT computed from homogenised LSWT (deep basin), *in-situ* data and satellite derived LSWT *Riffler et al.* (2015). Data is smoothed using the loess interpolation (blue line). The area shown around the smoothed line is the 95% confidence interval.

different satellites due to at least two reasons: i) to merge the data obtained from multiple satellites which have different quality, acquisition times and orbital decays; ii) for more dynamic surfaces where temperature varies at short temporal scales, this step may prove to be crucial. Moreover, identical slopes from different linear models between individual satellites and *in-situ* data confirmed the comparability of different satellite derived LSWT to each other.

Harmonic analysis was used to remove the gaps in the daily data due to undetected clouds and unresolved spurious calibration. We found that the gap filled daily data was over smoothed which could inadvertently remove some of the actual temporal dynamic change. In spite of this smoothing, the final homogenised LSWT data was able to pick up the long-term thermal dynamics over the lake surface. From the long-term summer trend of LSWT (Fig. 3.8), it was evident that the new LSWT data was able to pick up the low and high peaks in the last 30 years. Moreover, the spatial coverage of 1 km is another advantage of this dataset in comparing the thermal dynamics of different locations in the lake. The temperature difference map in Fig. 3.5 clearly illustrates how LSWT is distributed over the surface

of Lake Garda, whereby the shallow basin is more warm than the deep basin. Fig. 3.5 also shows that in the deep basin, the western shores are warmer than at the eastern side. This is explained by the high mountain chain of Mt. Baldo (2218 m) in the east, which blocks the sun for a long time during the day. The amplitude and phase maps obtained for each harmonic per year provide additional information on lake's thermal variation over time. The amplitude of the first harmonic is the highest for all the years, showing the influence of annual cycle of LSWT on the lake characteristics.

We performed validation at different stages of the work flow before and after homogenisation. We found that the RMSE obtained between individual satellite derived LSWT and *in-situ* data varied significantly between the deep and the shallow basins (Table 3.2). The higher RMSE in the shallow basin may be due to the varying skin effects in Lake Garda. The difference between skin and bulk temperature often termed as *skin effect* could be a deciding factor in varying accuracy over the deep and the shallow basins (Wilson *et al.*, 2013; Schneider *et al.*, 2009; Hook *et al.*, 2003). It must be noted that, the satellite measured LSWT represents temperature of a sub-micron layer between the lake surface and the air, and is highly variable according to the meteorological conditions. The *in-situ* data on the other hand represents bulk temperature, often measured between 0 to 0.5 m. The RMSE values were closer in the cross-platform comparison analysis in both the deep and shallow basins (Table 3.1), which demonstrates the superior quality of split-window coefficients. With respect to the sensor performance in deriving LSWT, we found that the ATSR series exhibited lower RMSE of 0.75 °C and 0.74 °C, followed by AVHRR with RMSE of 0.88 °C and 0.96 °C, MODIS with RMSE of 0.92 °C and 1.18 °C for deep and shallow basins respectively. A similar study by Riffler *et al.* (2015) also reported RMSE between 1 to 1.6 °C for AVHRR derived LSWT against *in-situ* data from multiple European lakes. For a better comparison it is important to have high frequency *in-situ* data matching the exact acquisition time of the satellites (Oesch *et al.*, 2008; Schneider *et al.*, 2009). The majority of the LSWT difference in both basins were below 1.0 °C. This suggests that the homogenisation procedure followed by HANTS retained the same pattern of accuracy and the RMSE in a similar range. Moreover, the comparison between summer lake mean trends obtained from our new LSWT data series and a similar satellite derived LSWT by Riffler *et al.* (2015) showed that both the data series are comparable and exhibits similar trends (Fig. 3.9).

To understand the long-term dynamics of lake surface temperature, it is crucial to have high frequency *in-situ* observations. In this study, we demonstrate that the satellite derived data can complement the missing *in-situ* observations of surface water temperature, with high spatial and temporal resolution. Global coverage of the satellite data is an added advantage which enables us to use the same data set for analyzing multiple lakes subsequently for

comparative analyses. The correct evaluation of the long-term trends of water temperature should take into account the advantages and limits of measurements made by both remote sensing techniques and *in-situ* recordings. Remote sensing technologies are limited to the detection of long-term trends circumscribed to the surface of water bodies. While this limit can be compensated by measurements performed at high temporal (daily) and spatial (whole lake) scales, the direct measurement of surface data in the field at lower temporal frequency can seriously impair the evaluation of long-term trends. Conversely, the lower temporal and spatial coverage of *in-situ* data can be partly balanced by a complete availability of data collected in the water column. The thermal structure of the mixolimnetic layer is less vulnerable to transient (daily to weekly scales) fluctuations of surface water temperatures originating from changes in local meteorological conditions. Therefore, most of the studies which have focused on the long-term trends of lake water temperatures have been based on measurements recorded in the water column. Lake Garda being a deep sub-Alpine lake, undergo complex mixing processes during winter and spring which have a strong influence on surface water temperatures (Salmaso, 2012). The inter-annual change of LSWT could be helpful in identifying impacts on the associated physical and biological processes in the lake.

Recent studies by Schneider and Hook (2010) and O'Reilly *et al.* (2015) have confirmed global warming of lakes due to climate change between 1985 and 2009 at the rate of 0.03 - 0.04 °C yr⁻¹. The main difference between the two studies is that, Schneider and Hook (2010) used exclusively the satellite derived LSWT, while O'Reilly *et al.* (2015) used a combination of *in-situ* and satellite data to derive the long-term trends. There are regional studies showing rapid warming of North American large lakes and European lakes which uses satellite derived LSWT (Schneider *et al.*, 2009; Riffler *et al.*, 2015). With a different approach, Salmaso and Mosello (2010), Coats *et al.* (2006) and Dokulil *et al.* (2006) reported warming at similar rates using *in-situ* volume weighted mean bulk temperature for the large lakes in Europe and North America. In contrast to our approach, many earlier studies used LST and SST data for lake studies which reported similar results. More recently, studies by Riffler *et al.* (2015) and Hulley *et al.* (2011) both developed lake specific surface temperature data using optimized split-window coefficients. We used the data by Riffler *et al.* (2015) for comparing the trends and obtained similar increasing temperature trends during summer. Study by Adrian *et al.* (2009) reported July warming since 1970 at the rate of 0.02 - 0.05 °C yr⁻¹ computed from *in-situ* data for multiple northern hemisphere lakes. Although we did our analysis separately for the deep and shallow basins of Lake Garda, we could not find any significant trend in the shallower, southern basin. Moreover, the thermal variation over the deep basin was similar to the entire lake basin, depicting that the lake is more influenced by the characteristics of the deep basin.

One of the main advantages of the developed method is that it gives the opportunity to study the long-term thermal dynamics of other lakes given the availability of the satellite data over study area. The spatial and temporal coverage obtained from this method is unique and cannot be achieved by any other data source. The outliers due to undetected clouds and other unsuitable data, are often difficult to remove completely. Though the method is robust, manually checking the accuracy of cloud masks for thousands of images is not practical. The automated cloud masking may leave undetected cloud pixels as clear sky ones. A future enhancement for refining cloud masks would be to use seasonal thresholds in SPARC algorithm instead of global ones. Nevertheless in this study, HANTS were able to remove the outliers as shown by the validation results at the expense of losing some of the short term temporal variability. The spatial resolution of 1 km is the highest possible when considering moderate resolution satellite data, but it is not high enough to study the smaller lakes. Moreover at this resolution, the issue with mixed pixels along the shore were solved by using an inner lake buffer while computing long-term climatologies.

Acknowledgements

The PhD Scholarship of Sajid Pareeth is supported by FIRS> T (FEM International Research School e Trentino) of Fondazione Edmund Mach. We acknowledge NOAA for providing historical AVHRR LAC data in level-1B format through CLASS archive. We acknowledge NASA and ESA for providing MODIS and ATSR series data in public through LAADS and MERCI respectively. We also acknowledge Eumetsat for providing European reanalysis data. We sincerely thank Juan-Carlos Jimenez-Muñoz of University of Valencia, for publishing the satellite specific split-window coefficients and providing us with the missing ones upon request. We acknowledge the communities behind developing and fine tuning the open source geospatial packages, especially the ones used here - Pytroll, Orfeo toolbox and GRASS GIS for their valuable support through mailing lists and personal communications. Investigations were carried out in the framework of the LTER (Long Term Ecological Research) Italian network, site "Southern Alpine lakes", IT08-000-A (<http://www.lteritalia.it/>). We are grateful to our colleagues in FEM and to the ARPA Veneto (G. Franzini) for logistic support in the field.

Chapter 4

Warming trends of perialpine lakes from homogenised time series of historical satellite and *in-situ* data

4.1 Abstract

Availability of more than thirty years of historical satellite data is a valuable source which could be used as an alternative to the sparse *in-situ* data. We developed a new homogenised time series of daily day time Lake Surface Water Temperature (LSWT) over last thirty years (1986-2015) at a spatial resolution of 1 km from thirteen multiple polar orbiting satellites. The new homogenisation procedure implemented in this study corrects for the different acquisition times of the satellites standardizing the derived LSWT to 12:00 UTC. In this study, we developed new time series of LSWT for five large lakes in Italy and validated the product with *in-situ* data from the respective lakes. Furthermore, we estimated the long-term annual and summer warming rates, the temporal coherence of mean LSWT between the lakes, and studied the intra-annual variations and long-term trends from the newly developed LSWT time series. We found a regional warming trend at the rate of $0.017\text{ }^{\circ}\text{C yr}^{-1}$ annually and $0.03\text{ }^{\circ}\text{C yr}^{-1}$ during summer. Mean annual and summer LSWT temporal patterns in these lakes were found to be highly coherent. Amidst the reported rapid warming of lakes globally, it is important to understand the long-term variations of surface temperature at a regional scale. This study contributes a new method to derive long-term accurate LSWT for lakes with sparse *in-situ* data thereby facilitating understanding of regional level changes in lake's surface temperature.

4.2 Introduction

To understand a lake's response to climate change, it is important to have long-term data with high temporal resolution. Data acquired from satellites and processed using remote sensing techniques is considered a good alternative to sparse *in-situ* data. Among many variables derived from satellite data, surface temperature is considered as reliable source due to its high level of accuracy (Reinart and Reinhold, 2008; Crosman and Horel, 2009). Many studies reported an average Root Mean Square Error (RMSE) less than 1 °C between satellite derived LSWT and *in-situ* data (Oesch *et al.*, 2008; Schneider *et al.*, 2009). Furthermore, the satellite derived LSWT is successfully used for estimating warming trends on a set of global lakes by Schneider and Hook (2010). Split-window algorithm is used to convert the top of atmosphere brightness temperatures measured from thermal channels (10 - 12 μm) radiances to the surface temperature (Li *et al.*, 2013). The split window coefficients are derived using radiative transfer models, by regressing the atmospheric profiles against simulated brightness temperatures. The accuracy of the estimated surface temperature from satellite data depends upon these coefficients, which in turn accounts for the correction of atmospheric attenuation (Maul and Sidran, 1971). Lake specific split-window coefficients and LSWT for a set of global lakes were derived and published by Hulley *et al.* (2011); MacCallum and Merchant (2012); Riffler *et al.* (2015). Moreover, Jimenez-Munoz and Sobrino (2008) published satellite specific split window coefficients to derive surface water temperature using a nonlinear split-window equation. There are multiple satellite derived surface temperature datasets that have been made available by national agencies to the public, for example, Pathfinder (Kilpatrick *et al.*, 2001), EuroLST (Metz *et al.*, 2014), GlobTemperature from ATSR series (Ghent, 2012) and the MODIS temperature products. More recently Sharma *et al.* (2015) prepared a global database of LSWT by acquiring both satellite based and *in-situ* data. However, these datasets are often of limited quality and satellite/sensor specific. Given that the availability of thermal data now span more than thirty years, it is important to combine data from multiple satellites to derive longest possible time series.

Many lakes around the globe are reported to be warming rapidly and the warming rates are found to be highly variable (O'Reilly *et al.*, 2015). Due to the close coupling of air and surface temperature, lakes are often directly affected by changes in air temperature due to climate change (Livingstone *et al.*, 2005; Salmaso *et al.*, 2014). It is well known that basically all system levels of lakes are responding to enhanced water temperature and subsequent changes in thermal regime (Adrian *et al.*, 2009). The changes in lake characteristics due to climate change are found to be highly coherent in temporal and spatial scales (Salmaso *et al.*, 2014; Livingstone and Dokulil, 2001; Dokulil *et al.*, 2006). For example, Lake Surface Water Temperature (LSWT) is highly coherent owing to the large scale of the climatic forcing

controlling this factor. In Europe, changes in surface temperature over lakes are largely controlled by the North Atlantic Oscillation (NAO) (Blenckner *et al.*, 2007) and, mostly in the Mediterranean area, by the East-Atlantic Pattern (Salmaso, 2012). Nevertheless, the actual rate of change in temperature is highly variable between lakes and depends upon micro climatic conditions like local air temperature, wind speed and landscape properties like altitude and latitude (O'Reilly *et al.*, 2015; Livingstone *et al.*, 2005). Although at very large spatial scales (global, Northern hemisphere) patterns are found to be consistent, variation among regional lakes can be important to understand. In this study, we look into a regional scale assessment of variation in LSWT among large lakes in Italy. In Italy, an increase of surface temperature has been reported for the large lakes south of Alps in Northern Italy (Salmaso and Mosello, 2010; Lepori and Roberts, 2015) and the largest lake of Peninsular Italy - Lake Trasimeno (Ludovisi and Gaino, 2010).

In this study, we developed a homogenised time series of LSWT from multiple satellites for five large lakes in Italy using a new method developed by (Pareeth *et al.*, Submitted). The satellite derived homogenised LSWT cover thirty years (1986-2015) of daily data at a spatial resolution of 1 km, obtained from thirteen polar orbiting satellites. In this study, we (i) evaluated the possibility to use satellite data to reconstruct frequent (daily) LSWT, and (ii) used the reconstructed data to assess long-term and seasonal trends of LSWT in the study lakes. For validation of the satellite derived data, *in-situ* measurements of lake surface water temperatures at mostly monthly temporal resolution were available for the study lakes over years to decades during the 1986-2015 time window. We tested for long-term trends in LSWT, intra-annual variability and the coherence in temperature development between the lakes using the new LSWT data. Our study showed that the lakes are undergoing significantly rapid annual and summer warming with a high temporal coherence.

4.3 Study area

We selected five sub-alpine lakes and one lake in the Central Italy for this study. Table 4.1 lists the main features of the investigated lakes and Fig. 4.1 their location. For the satellite data processing we selected a bounding box (red box in the Fig. 4.1) Longitude: 42.9762°E - 46.2371°E; Latitude: 8.2127°N - 12.3456°N; Area: 360 km × 340 km, which covers all the five lakes. The four large lakes south of the Alps (Lake Garda, Lake Iseo, Lake Como and Lake Maggiore) have maximum depths and volumes ranging between 251 and 410 m, and 4.7 and 49 billions of m³, respectively (Table 4.1). These lakes contribute water for about 40% of the discharge of the River Po, the largest Italian river. Theoretical renewal times are around 4-5 years (lakes Maggiore, Como and Iseo), 12 years (Lake Lugano) and 27

years (Lake Garda). All these lakes are characterized by steep shores and, excluding Lake Garda, the absence of wide shallow areas. One of the most relevant characteristics of lake Garda is the presence of a wide shallower (< 40-80 m) in the southernmost part, i.e. in the region located in correspondence of the Po Plain, outside the sub-Alpine chain. Owing to their climatic location, these lakes should be warm monomictic, i.e. characterized by one complete circulation in the winter months at or above 4 °C and stable thermal stratification from spring to early autumn (Wetzel, 2003). Actually, owing to their large depth, this group of lakes is mero-oligomictic, showing long periods of incomplete spring mixing, interrupted by occasional and irregular complete overturns following harsh and windy winters (Lakes Garda, Como and Maggiore), or basically meromictic, i.e. characterized by a continuous separation of the deep hypolimnion from the upper layers (Mosello *et al.*, 1997; Salmaso *et al.*, 2003). Lake Trasimeno is the largest lake of the peninsular Italy. Despite its large extension, the lake is shallow, with a maximum nominal depth of 6.3 m, reached when the waters exceeds the artificial outlet threshold located at 257.50 m a.s.l. Because of the small extent of its watershed, meteorological conditions have caused flood and drought events over the centuries and the problems related to water level still remain unsolved (Ludovisi *et al.*, 2013).

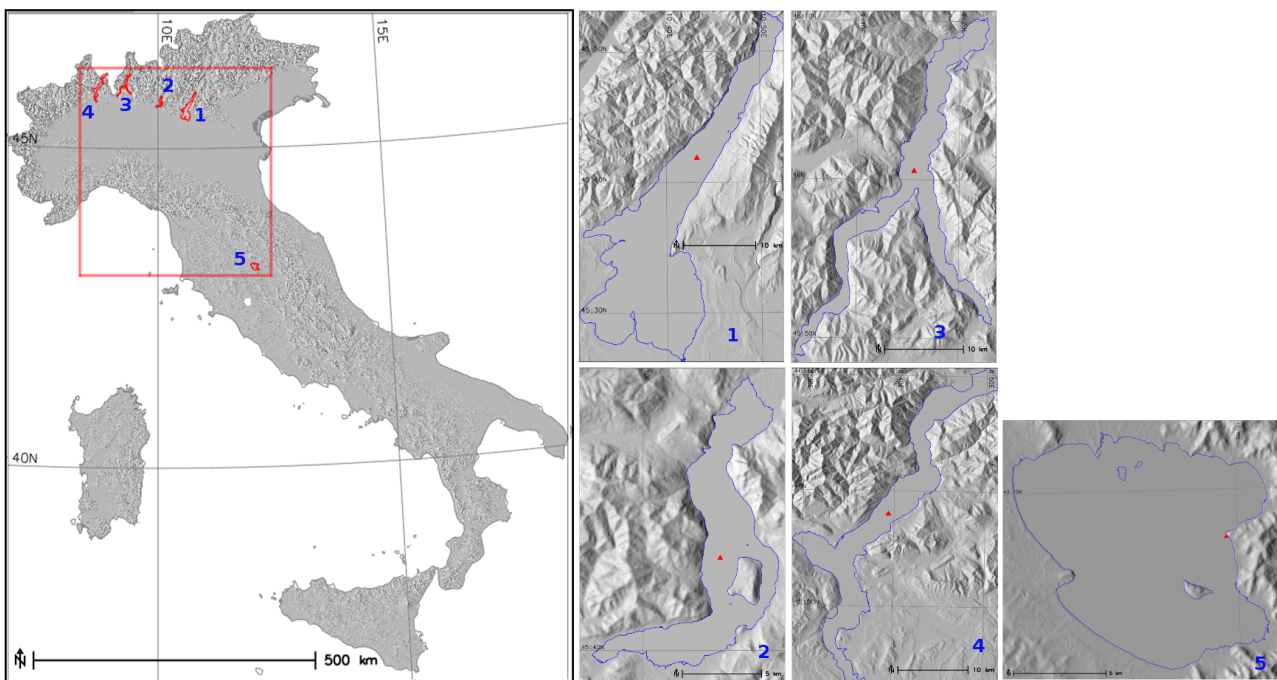


Figure 4.1: Study area in Italy depicting location of the study lakes and the bounding box used for image analysis (red box). Each lake and the location for the *in-situ* water temperature measurements (marked as red triangles) are shown in sub figures 1–5. 1) Lake Garda; 2) Lake Iseo; 3) Lake Como; 4) Lake Maggiore; 5) Lake Trasimeno

Table 4.1: Main characteristics of the investigated lakes and type of *in-situ* data used in this study

Name	Trophic status	Max depth (m)	Altitude a.s.l. (m)	Surface area (km ²)	<i>in-situ</i> time period	Temporal frequency	Time (UTC)	Method	References
Lake Garda	oligomesotrophy	350	65	366.77	1991-2015	Monthly	12:00-13:00	Probe	<i>Salmaso (2012)</i> <i>Salmaso and Cerasino (2012)</i>
Lake Iseo	mesoeutrophy	251	186	61.12	1993-2012	Monthly	10:30-11:30	Probe	<i>Leoni et al. (2014a,b)</i> <i>Marti et al. (2016)</i>
Lake Como	oligomesotrophy	410	198	145.24	2000-2013	Monthly	10:30-12:00	Probe	<i>Buzzi (2002)</i>
Lake Maggiore	oligotrophy	370	193	212.5	1986-2012	Monthly	10:00-12:00	mercury thermometer	<i>Morabito et al. (2012)</i>
Lake Trasimeno	mesoeutrophic	6.3	257.5	124	1986-2006	Daily	09:00-12:00	mercury thermometer	<i>Ludovisi et al. (2013)</i> <i>Ludovisi and Gaino (2010)</i>

4.4 Methods

4.4.1 Data

We used data (total of 62,799 images) from 13 polar orbiting satellites with moderate resolution sensors which provides dual thermal channels at 10.5-11.5 μm and 11.5-12.5 μm at 1 km spatial resolution to derive daily LSWT (Fig. 4.2). The satellites collectively cover a time span of 30 years (1986-2015) with daily multiple (day and night) snapshots of the entire globe. We used satellite observations acquired during the day to match the acquisition time of the *in-situ* data used for validation. We downloaded and processed level-1B data of AVHRR, ATSR series and MODIS from Comprehensive Large Array-data Stewardship System (CLASS - <http://www.class.ncdc.noaa.gov/>), Mercator data archive (<http://ats-mercator-ds.eo.esa.int/mercator/welcome.do>) and Level 1 and Atmosphere Archive and Distribution System (LAADS - <https://ladsweb.nascom.nasa.gov/>) respectively for the study area. *In-situ* data consisted of the temperatures measured at 0.0 - 0.5 m (bulk temperature) from surface of each lake. We collected the *in-situ* data from the research collaborators working on the respective lakes. Besides my institute (Fondazione Edmund Mach, Italy) which collects monthly samples from Lake Garda, data for other lakes were obtained from University of Milan-Bicocca (Lake Iseo; Contact person–Dr.Barbara Leoni), CNR–Institute for the Study of Ecosystems (Lake Maggiore; Contact person–Dr.Giuseppe Morabito), ARPA Lombardia (Lake Como; Contact person–Dr.Fabio Buzzi), Università degli Studi di Perugia (Lake Trasimeno; Contact person–Dr.Alessandro Ludovisi). The time span and temporal frequency of this data vary with lakes (Table 4.1). The methods and the locations of *in-situ* data measurement are shown in Table 4.1 and Fig. 4.1 respectively.

4.4.2 Homogenized LSWT from satellite data

We developed a homogenised daily LSWT time series corrected at 12:00 UTC for 30 years from 1986 to 2015. The entire workflow of the new method is shown in Fig. 4.3. The thermal channels obtained from the downloaded satellite data were calibrated to derive top of atmosphere brightness temperatures from at-sensor radiances. We used a modified method based on Separation of Pixels Using Aggregated Rating over Canada (SPARC) remove clouds for the months from April to October (*Khlopenkov and Trishchenko, 2007*). For the remaining colder months, we used an inter-quartile filter to remove the outliers due to clouds. We used a nonlinear split-window algorithm with satellite specific coefficients to derive surface temperature from brightness temperatures as proposed by *Jimenez-Munoz and Sobrino (2008)*. The acquisition time correction or the homogenisation procedure to correct for the varying

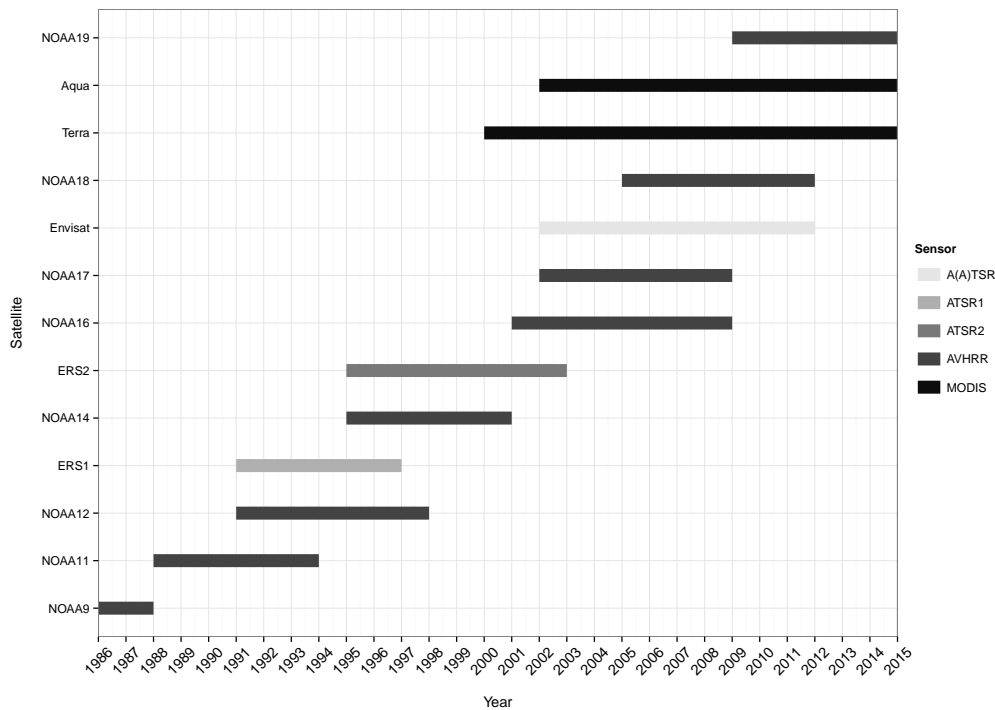


Figure 4.2: Graph showing time line of the thirteen satellites used in this study, color coded according to sensors. Y-axis represents the name of satellites used in this study. The legend represents the sensor on board each satellite.

acquisition time of the multiple satellites was based on a physical model using typical pattern technique (Jin and Treadon, 2003). The gaps in the time series due to cloud coverage were statistically filled using the harmonic analysis method. The entire methodology to derive the homogenised LSWT from multiple satellites is explained in detail by Pareeth *et al.* (Submitted), where time series of satellite derived homogenised LSWT for Lake Garda were developed and validated. Here, we extend this new LSWT time series to four other large lakes in the region. All the image processing was performed using the open source software GRASS GIS 7.0 (Neteler and Mitasova, 2008; Neteler *et al.*, 2012).

4.4.3 Validation, trends and regional coherence

To validate the newly developed homogenised LSWT, we used *in-situ* data collected from study lakes over the last three decades. We used lake mean for validation against *in-situ* data to avoid mixed pixels from the sampling locations. We used an inner lake mask to avoid land/water mixed edge pixels in the analysis and used this masked pixels to calculate the average LSWT. We developed linear regression models for each lake to assess the relation between *in-situ* and the satellite derived LSWT. Moreover, to quantify the accuracy of the satellite derived LSWT, we estimated Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for all the lakes. The daily LSWT derived from satellite images were then

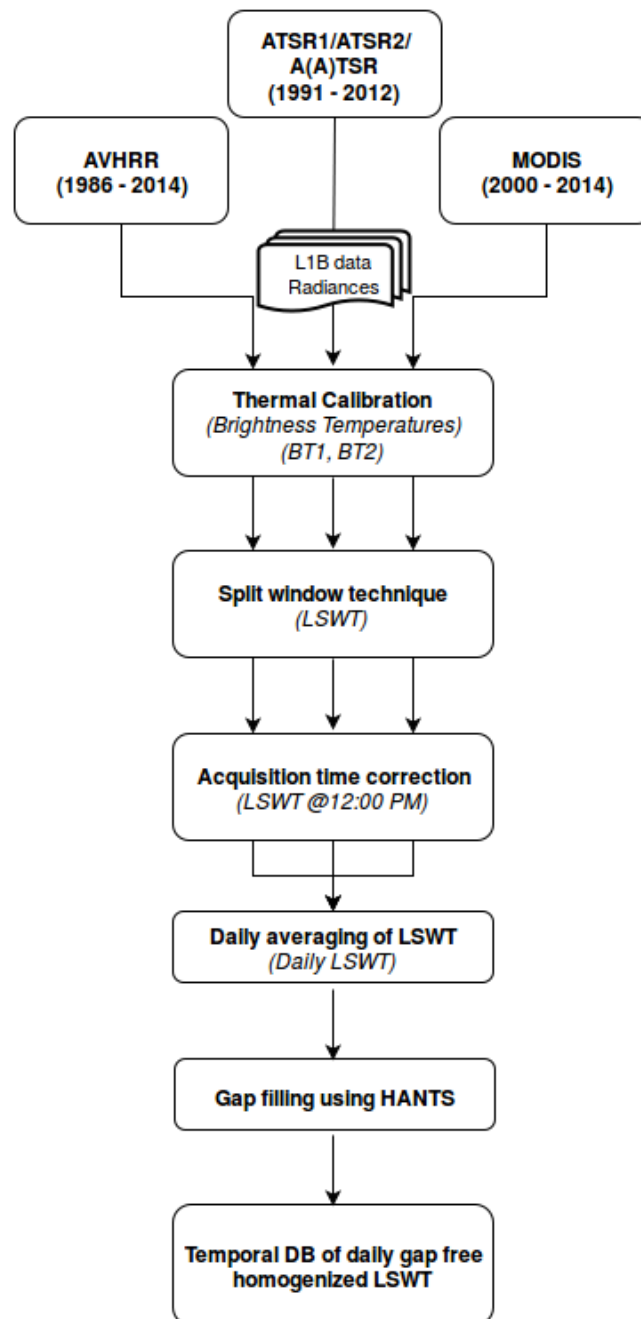


Figure 4.3: Block diagram of entire workflow implemented to derive time series of homogenised LSWT with level-1B thermal radiances obtained from thirteen polar orbiting satellites in the time window from 1986 to 2015.

aggregated to produce annual and summer time series. We avoided other seasons due to insignificant warming reported with our earlier study on Lake Garda. To produce summer values, daily values were averaged within the months June, July and August. We performed non-parametric trend analysis on the summer and annual averages using Mann-Kendal tests. The Durbin-Watson test was applied to check for potential serial correlation in the time series data. For comparison, trend analysis was performed on similar satellite based LSWT data from *Riffler et al.* (2015) for all the lakes except Lake Trasimeno, but for a shorter time scale

Table 4.2: Results of validation of the homogenised LSWT derived from satellite data against the *in-situ* data of the lakes. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are reported in °C. N is the number observations used in the linear models. Coefficient of determination (R^2) and slope are obtained from the linear models

Name	RMSE (°C)	MAE (°C)	R^2	Slope	N
Lake Garda	1.06	0.83	0.98	0.95	217
Lake Iseo	1.08	0.95	0.97	0.99	129
Lake Como	1.14	0.96	0.96	0.99	83
Lake Maggiore	1.13	0.97	0.97	0.96	207
Lake Trasimeno	1.38	1.13	0.98	0.97	4392

from the year 1989 to 2013. The temporal coherence of seasonal and annual LSWT between lakes was evaluated by estimating correlation coefficient between time series of seasonal data. The lakes were then grouped together using cluster analysis (complete method) with dissimilarity measure as $1 - \text{correlation}$ (*R Core Team*, 2013). We further fitted a Generalized Additive Mixed Model (GAMM) with an Auto-Regressive (AR1) model structure to take into account the effect of autocorrelation, to study the intra-annual variability and long-term trend of LSWT during the past three decades. We used R statistical software (Version 3.3.0) to run the statistical tests, with package *wq* for Mann-Kendall tests and package *mgcv* for GAMM model (*Jassby and Cloern*, 2016; *Wood*, 2006, 2004).

4.5 Results

The linear models relating *in-situ* and satellite data exhibited high coefficient of determination (R^2) in the analysed lakes, with an average of 0.97. Moreover, the slopes of the linear models were in the same range, implying that the relations between *in-situ* and satellite derived LSWT were similar in all lakes (Table 4.2). An average RMSE and MAE of 1.2 °C and 0.97 °C, respectively, were showed over all the lakes (Table 4.2, Fig. 4.4). The minimum respectively maximum RMSE were reported for Lake Garda and Lake Trasimeno (Table 4.2).

The results of the non-parametric trend analysis are given in Table 4.3. Four of five study lakes showed significant warming trends with an average rate of 0.03 °C yr⁻¹. The highest rate of warming during summer was found for Lake Garda and Lake Trasimeno at a rate of 0.04 °C yr⁻¹ (***P < 0.001). Significant warming during summer were also reported for Lake Como and Maggiore at a rate of 0.03 °C yr⁻¹ (***P < 0.05). For Lake Iseo, a non-significant warming trend of 0.03 °C yr⁻¹ was found. Fig. 4.5 shows the summer warming trend of all the lakes over the last thirty years. We found similar warming trends from the satellite derived LSWT developed by *Riffler et al.* (2015) (Table 4.3). To allow direct comparison of the trends in LSWT we adjusted the time period of our data to those used by *Riffler et al.* (2015).

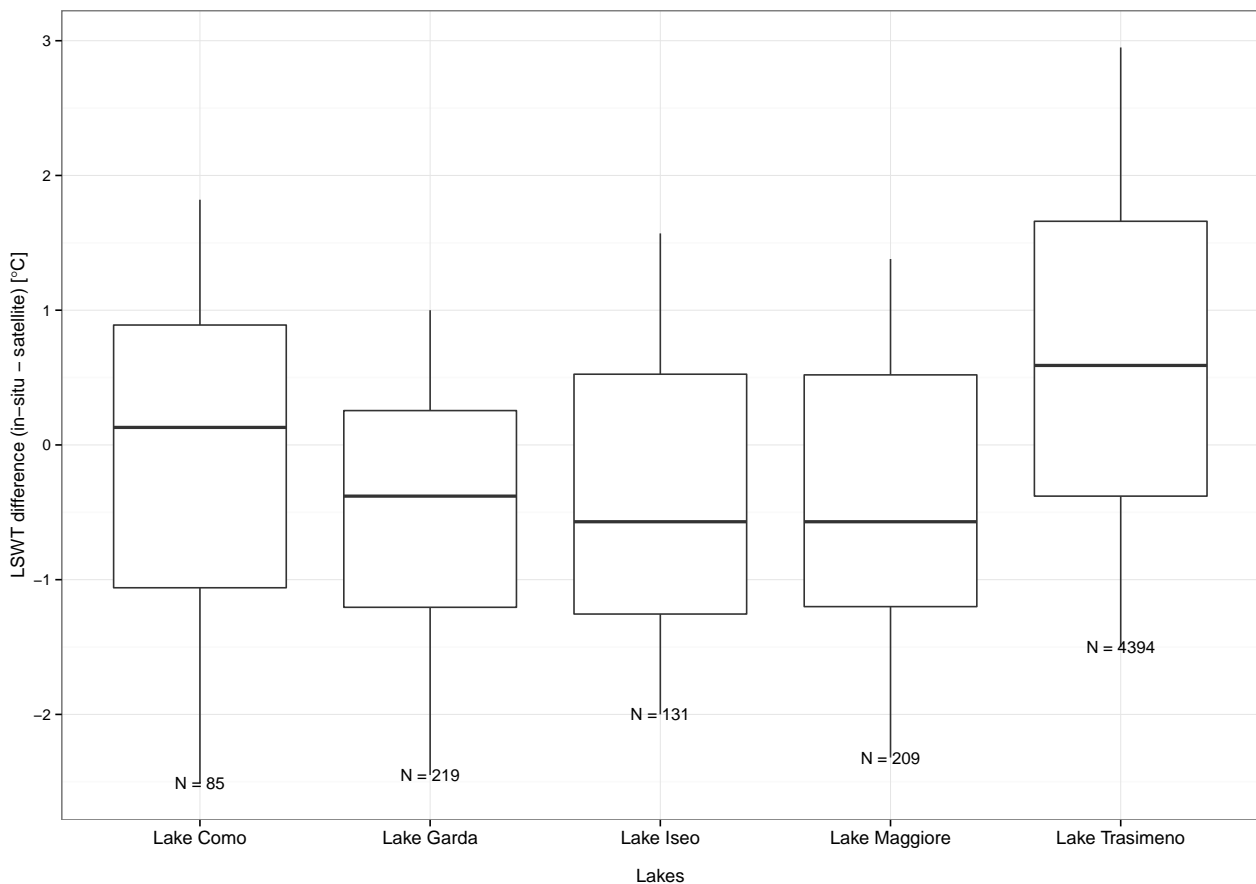


Figure 4.4: Boxplots of the difference between *in-situ* and lake mean LSWT estimated from satellites. The outliers are removed using an inter-quartile filter with 10 and 90 percentile cutoffs. Number of observations (N) in each lake used in the assessment is shown over the boxplot.

Table 4.3: Summer (June/July/August) trends obtained from two different sources of LSWT data. The first column lists the trends obtained from new homogenised LSWT over thirty years (1986-2015). Second column lists the summer trends from the same data but for a reduced time scale (1989-2013) to match the data from *Riffler et al. (2015)*.

Name	Homogenized LSWT (°C) (1986-2015)	Homogenized LSWT (°C) (1989-2013)	<i>Riffler et al. (2015)</i> (°C) (1989-2013)
Lake Garda	0.036 (**P < 0.001)	0.033 (*P < 0.05)	0.049 (**P = 0.01)
Lake Iseo	0.017 (P > 0.1)	0.016 (P > 0.1)	0.051 (**P = 0.01)
Lake Como	0.032 (*P < 0.05)	0.024 (P < 0.1)	0.041 (*P = 0.05)
Lake Maggiore	0.033 (*P < 0.05)	0.025 (P = 0.1)	0.003 (P > 0.1)
Lake Trasimeno	0.044 (**P < 0.001)	0.05 (**P < 0.001)	N.A

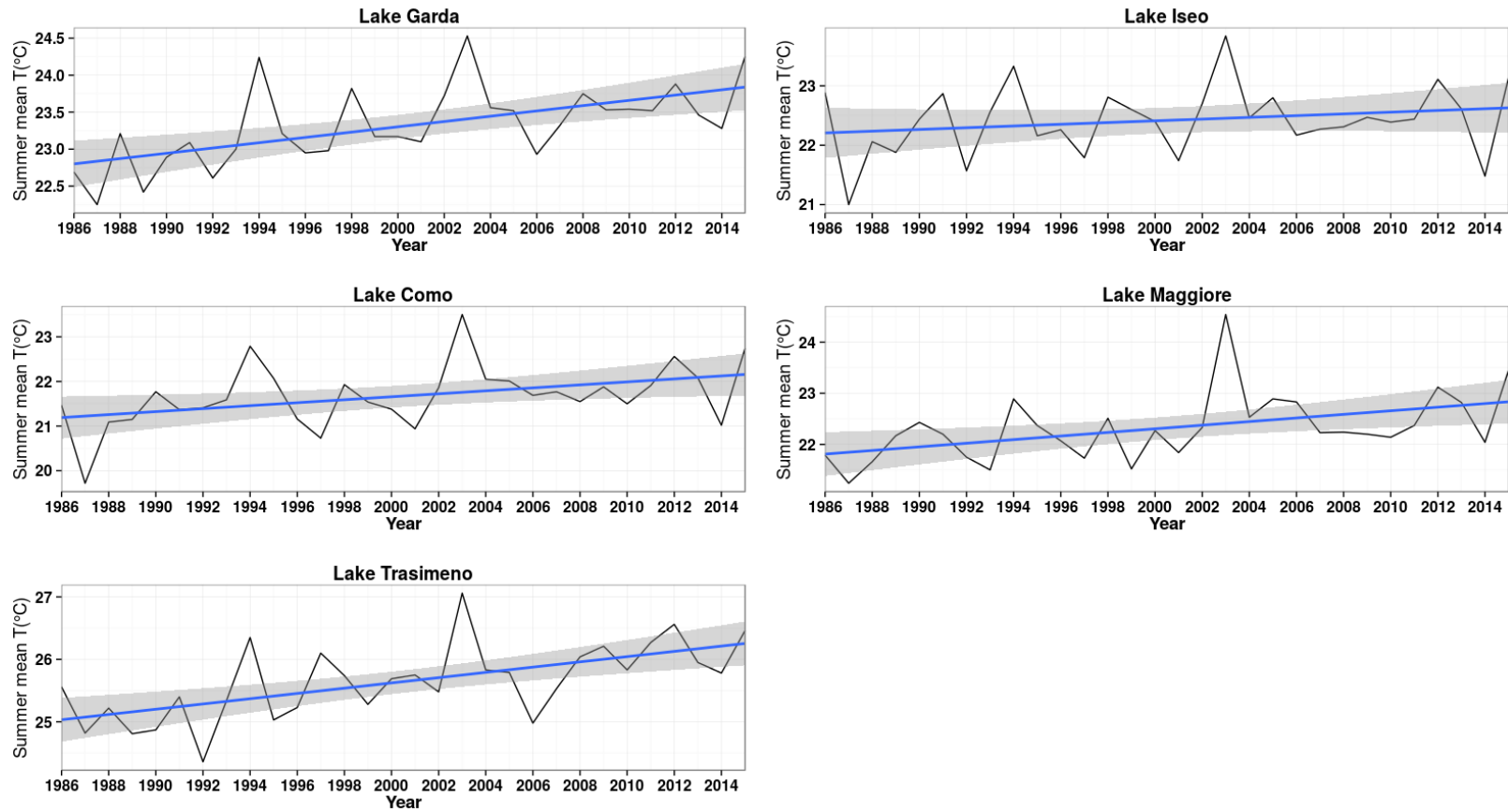


Figure 4.5: Summer mean LSWT derived from the new homogenised daily LSWT between 1986 to 2015. Data were smoothed using the linear model (blue line). The gray area represents 95% confidence interval

Table 4.4: Annual trend obtained from Generalized Additive Mixed Model (GAMM) and Mann-Kendall test for all the lakes. The long-term daily trend from GAMM is divided by 30 (number of years) to obtain the annual trend listed below.

Name	Mann-Kendall (°C) (1986-2015)	GAMM (°C) (1986-2015)
Lake Garda	0.020 (*P < 0.05)	0.018 (*P = 0.05)
Lake Iseo	0.019 (*P < 0.05)	0.019 (P = 0.1)
Lake Como	0.012 (P > 0.1)	0.011 (P > 0.1)
Lake Maggiore	0.017 (P < 0.1)	0.014 (P > 0.1)
Lake Trasimeno	0.017 (P > 0.1)	0.006 (P > 0.1)

Annually, Mann-Kendall tests computed on all the five lakes revealed an average warming of $0.017\text{ }^{\circ}\text{C yr}^{-1}$. All the analysed lakes showed an increase in the water temperatures with trends significant in lakes Garda, Iseo and Maggiore (Table 4.4). The highest rate of annual warming was reported for Lake Garda at $0.02\text{ }^{\circ}\text{C yr}^{-1}$ (*P < 0.05).

Pairwise correlation of LSWT between different lakes are depicted in Fig. 4.6 and Fig. 4.7. During summer, temporal coherence was highest between Lake Iseo and Como (correlation coefficient of 0.87) and weakest between Lake Trasimeno and Lake Como (correlation coefficient of 0.62). In spite of their differences in physical and biological properties, coherence in summer LSWT was high between Lake Garda and Lake Trasimeno, while the other three lakes are grouped into a single branch (Fig. 4.6). At the annual scale, the four sub-Alpine lakes showed a high temporal coherence, whereas Lake Trasimeno clearly stood out with lower correlation coefficients to other lakes.

The intra-annual variance in the LSWT over last thirty years obtained from GAMM were found to be similar for all the sub-alpine lakes, where as Lake Trasimeno had a long and hotter summers and cooler winters compared to other lakes (Fig. 4.8a). The long-term trends of daily LSWT increased by $0.34\text{ }^{\circ}\text{C}$ for Lake Como, $0.41\text{ }^{\circ}\text{C}$ for Lake Maggiore, $0.17\text{ }^{\circ}\text{C}$ for Lake Trasimeno, $0.55\text{ }^{\circ}\text{C}$ for Lake Garda and $0.56\text{ }^{\circ}\text{C}$ for Lake Iseo, over thirty years from 1986 to 2015 (Fig. 4.8b).

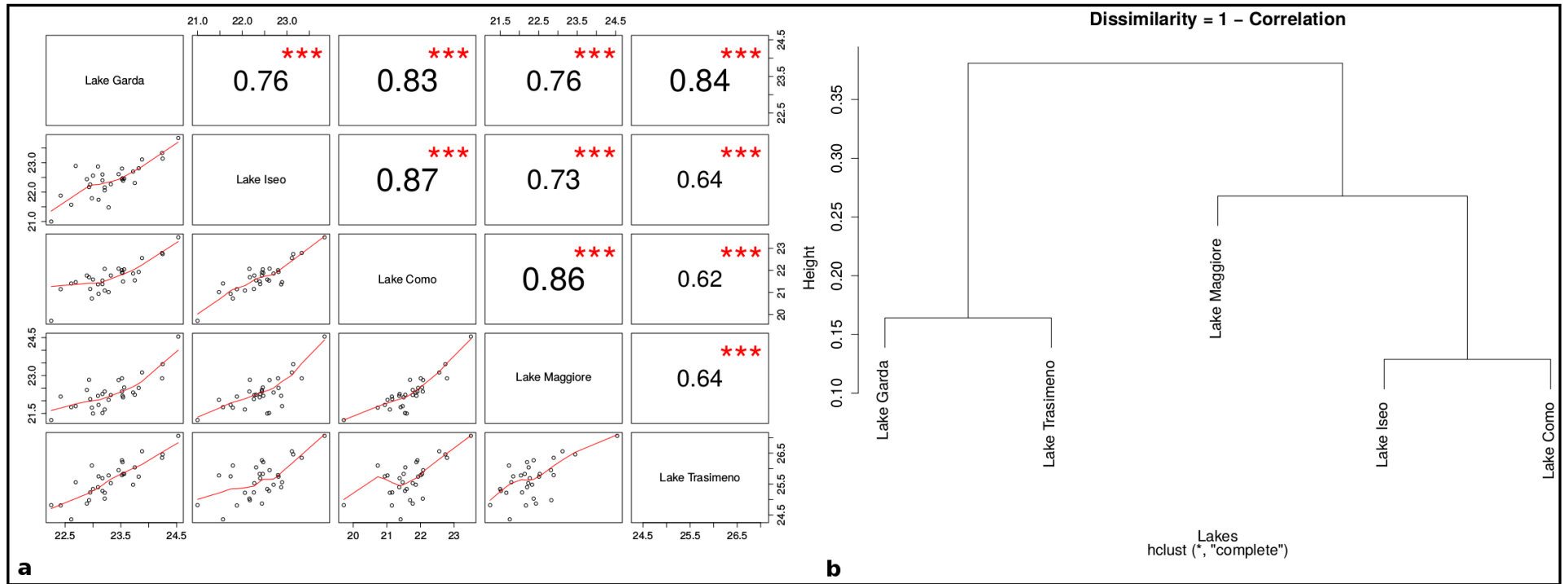


Figure 4.6: a) Scatter plot matrix showing temporal coherence between summer mean LSWT of all the lakes. Pearson's correlation coefficient is given as measure of coherence. b) Dendrogram showing the clustering of lakes based on summer mean LSWT with 1 - Pearson's correlation coefficient the dissimilarity measure.

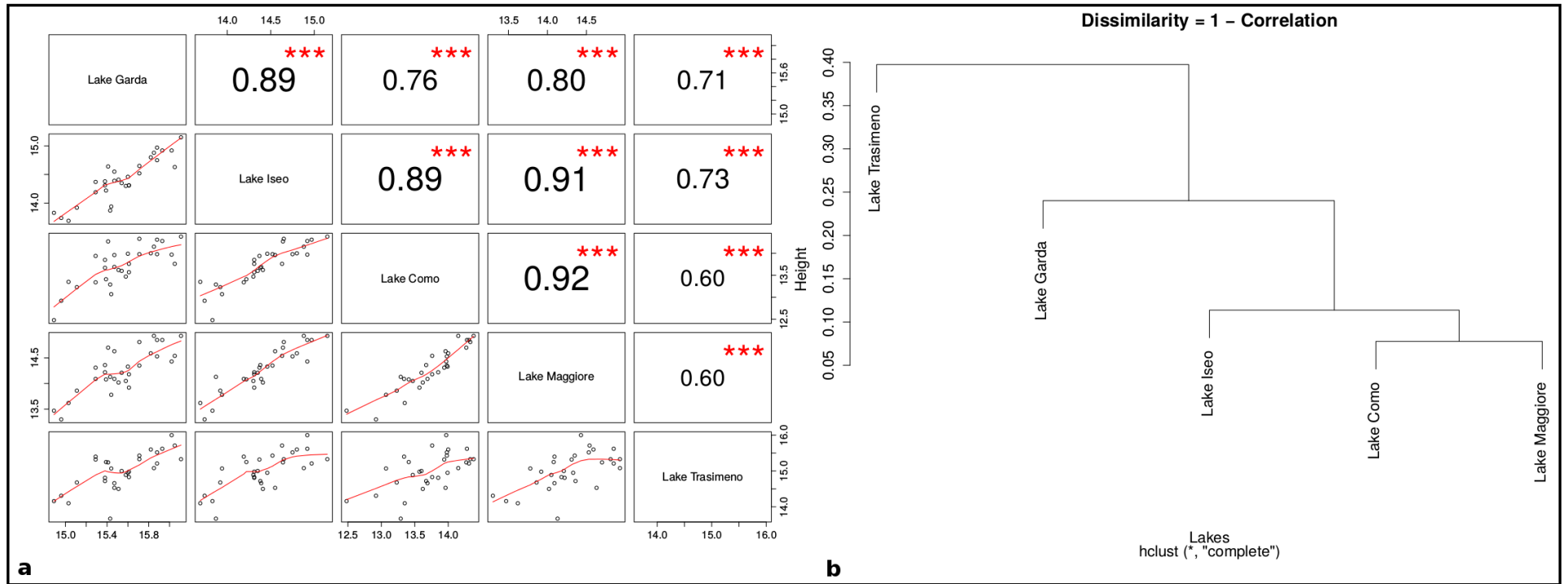


Figure 4.7: a) Scatter plot matrix showing temporal coherence between annual mean LSWT of all the lakes. Pearson's correlation coefficient is given as measure of coherence. b) Dendrogram showing the clustering of lakes based on annual mean LSWT with 1 - Pearson's correlation coefficient the dissimilarity measure.

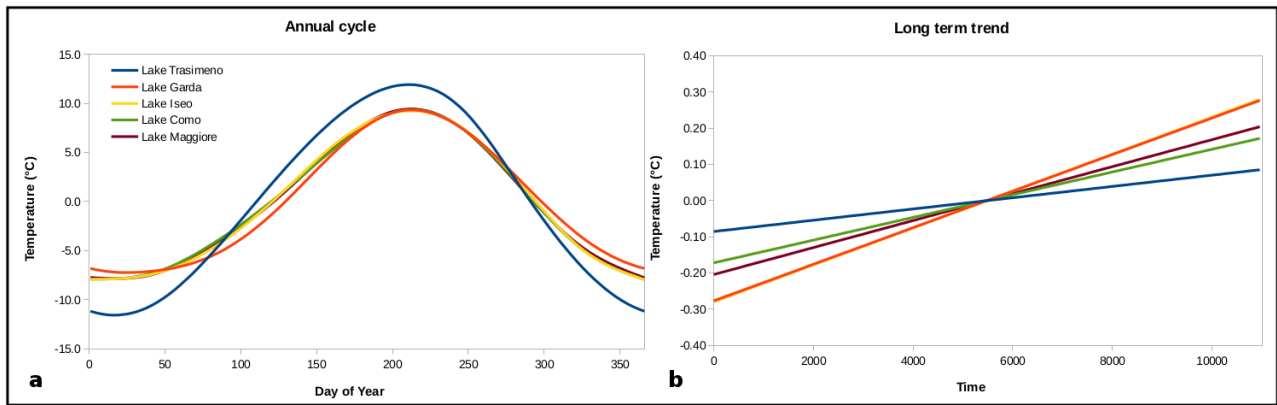


Figure 4.8: a) Annual cycles and b) long-term trends from the GAMM analysis. For the annual cycles, the x-axis represents day of the year and for long-term trend, the x-axis represents count of days from 1986 to 2015 representing the time factor in the model. In (b) the trend lines of the lakes Garda and Iseo practically coincide and are not distinguishable.

4.6 Discussion

We introduced daily time series of satellite derived LSWT for five large lakes in Italy using a newly established methodology, which can be used to study the long-term thermal characteristics of lakes. We found that the new satellite derived LSWT is highly accurate as validated by *in-situ* long-term water temperature measurements. Importantly, for all five lakes similar accuracy of satellite derived LSWT was found, indicating the robustness of the new LSWT time series. This is further reiterated by the similar slopes of the fitted line in the linear models between *in-situ* and satellite derived LSWT. All studied lakes showed significantly warming trends during summer except for Lake Iseo. The temporal patterns during summer were similar among lakes with the year 2003 being the hottest reported in last thirty years. Moreover, mean summer LSWT were found to be temporally coherent. The underlying pattern of coherence matched with the spatial closeness of the lakes. Lakes in close spatial connection such as Lake Trasimeno and Lake Garda showed the highest coherence in LSWT. Coherence in LSWT diminished towards the west with Lake Como situated at the most westerly boarder of the study area showing the least coherent pattern in LSWT. Interestingly, Lake Garda was highly coherent to other sub-Alpine lakes in the study and also to Lake Trasimeno, while other sub-Alpine lakes showed less correlation to Lake Trasimeno. This pattern could reflect the influence of Mediterranean weather on Lake Garda during summer.

Our results on validation of satellite derived LSWT were in line with other similar studies. In a previous study establishing the homogenisation methodology we found higher performance by ATSR series of sensors with RMSE of 0.75 °C, followed by AVHRR and MODIS with

reported RMSE of 0.88 °C and 0.92 °C respectively (Pareeth *et al.*, Submitted). Oesch *et al.* (2005) used AVHRR and MODIS data to estimate LSWT for Swiss lakes and reported a bias ranging from 0.9 to 1.6 °C. Riffler *et al.* (2015) reported a RMSE of less than 1.5 °C using AVHRR data from multiple NOAA satellites. Multiple studies dealt with long-term warming of lakes around the globe and reported similar rates as obtained by our study. Schneider and Hook (2010) used satellite observations to report rapid warming of inland water bodies. Most recently O'Reilly *et al.* (2015) used a combination of satellite and *in-situ* observations to estimate long-term trends of LSWT and found an increasing summer average at the rate of 0.03 °C yr⁻¹. In Europe, multiple studies confirmed late spring/summer warming of North European lakes (Adrian *et al.*, 2009; George *et al.*, 2005), Central Europe (Dokulil *et al.*, 2006) and for the lakes south of Alps (Salmaso and Mosello, 2010).

The high temporal coherence between Lake Garda and Trasimeno during summer, sheds light to how differences in regional climate can influence the epilimnion of lakes located at the border of the Alpine chain. The shallower area of Lake Garda is located for a significant part in the Po Plain, i.e. outside the Alpine chain where the dominant eco-system is Mediterranean. Conversely the northern part of Lake Garda and the other large subalpine lakes are more influenced by Alpine climate. Moreover, the presence of a wider shallower area characterizing the southern portion of Lake Garda could expose the lake to a greater influence of the Mediterranean climate, and to a greater impact of short and medium frequency climatic fluctuations acting especially on the southern basin. These factors could contribute to explain the higher coherence in the temporal development of water temperature in lakes Garda and Trasimeno during summer. On the other hand, deep sub-Alpine lakes considered in this study undergo complex mixing processes during winter and spring, which have a strong influence on surface water temperatures (Salmaso *et al.*, 2014). This explains the weaker coherence of sub-Alpine lakes with Lake Trasimeno on an annual scale. Moreover, the GAMM output clearly showed a distinctive annual cycle of Lake Trasimeno compared to other lakes. Lake Trasimeno is shallow with a maximum depth of 6.3 m at an higher altitude of 257.5 m above a.s.l., which would explain the warm summer and coolest winter among the five lakes under study. The long-term trend from GAMM shows a clear upward trend in LSWT. The annual rate computed from the long-term trends is similar to what we obtained from Mann-Kendall tests. For Mann-Kendall tests, we averaged the daily observations to annual and seasonal time series thereby losing the variance to a greater extent. Moreover, the GAMM model takes care of autocorrelation by including a AR1 model.

To the best of our knowledge, this is the first time homogenised satellite derived LSWT time series were developed for large lakes to study long-term warming and temporal coherence between lakes. The advantages of using such a dataset are multitude, though certain caveats

should be considered before using it for analysis. First, the satellite sensor measures the emitted radiation in the thermal infra-red region, and the energy is emitted from a sub-micron layer above the lake surface, known as skin layer. Therefore, the LSWT from satellites represent skin temperature (Wilson *et al.*, 2013), while the *in-situ* data represent bulk data (0 - 5 m). This difference explains the higher RMSE we obtained in our study. The satellite data from Riffler *et al.* (2015) were converted to bulk temperature following the method developed by Minnett *et al.* (2011), whereas our data represents skin temperature. This could be one reason for the slightly different trends obtained from either dataset. Second, the higher RMSE reported in this study might also reflect the lack of high temporally frequent *in-situ* data (monthly time scale) which did not correspond to the actual acquisition time of the satellite derived LSWT (daily time scale). In this study, we considered only day time satellite data, as for night time, corresponding *in-situ* data were not available to validate the observations. Our dataset is at a resolution of 1 km, which could be useful for lakes as big as Lake Garda, but not an ideal choice to study the spatial variability for smaller and narrower lakes. There are high resolution thermal satellite data available from Landsat missions available at a spatial resolution of 30 m which could be used to study spatial and temporal variability of LSWT for smaller lakes. Future studies will look into the influence of larger climatic forcings like North Atlantic Oscillations (NAO) and Eastern Atlantic (EA) on the long-term warming changes of LSWT followed by ecological consequences in the large lakes south of the Alps.

4.7 Conclusion

In this study, we used newly developed daily LSWT time series of thirty years (1986-2015) of five large lakes in the Italy to study their temporal dynamics. The new LSWT data are purely satellite derived combining data from thirteen different satellites. The data are unique with its new methodology adopted to develop homogenised time series correcting for the different acquisition time of the satellites. We demonstrated that the satellite data could become a promising alternative to *in-situ* measurements of LSWT. The satellite derived LSWT exhibited high accuracy when compared to the available *in-situ* data of the lakes. With addition of new satellites like Sentinel-3 with similar spectral bands in the thermal region, this study gives the opportunity to expand further the dataset by adding future observations. We found a regional warming trend at the rate of $0.03\text{ }^{\circ}\text{C yr}^{-1}$ during summer and at an annual rate of $0.017\text{ }^{\circ}\text{C yr}^{-1}$. Annual and summer means of LSWT of the studied lakes were highly coherent. We further estimated the long-term trend and the intra-annual cycle using generalized additive mixed models. The method is extendable to any other geographical region of the globe given

the availability of satellite coverage. This gives higher flexibility to the researchers to study the long-term thermal dynamics of the lakes with sparse *in-situ* data.

Acknowledgements

The PhD Scholarship of Sajid Pareeth is supported by FIRS > T (FEM International Research School e Trentino) of Fondazione Edmund Mach. We acknowledge NOAA for providing historical AVHRR LAC data in level-1B format through CLASS archive. We acknowledge NASA and ESA for providing MODIS and ATSR series data in public through LAADS and MERCI respectively. We also acknowledge Eumetsat for providing European reanalysis data. We sincerely thank Juan-Carlos Jimenez-Muñoz of University of Valencia, for publishing the satellite specific split-window coefficients and providing us with the missing ones upon request. Investigations were carried out in the framework of the LTER (Long Term Ecological Research) Italian network, site "Southern Alpine lakes", IT08-000-A (<http://www.lteritalia.it/>). We are grateful to our colleagues in FEM and to the ARPA Veneto (G. Franzini) for logistic support in the field. We also thank all our collaborators who kindly permitted us to use their data for validation.

Chapter 5

General Discussion

5.1 Summary

Globally, lakes are subject to rapid warming due to climate change over the last century (O'Reilly *et al.*, 2015; Schneider and Hook, 2010). Long-term trends in the thermal variation of surface temperature of lakes point towards changing climatic regime in the surrounding landscape (Blenckner *et al.*, 2007). The global trend is often regionalised by specific meteorological and environmental conditions (O'Reilly *et al.*, 2015). Lakes in different geographical region respond differently to the environmental changes. Therefore, a deep understanding of the trends and patterns of thermal variations at a regional scale is necessary to quantify its effects and to develop adaptation options. For this reason, the availability of long-term data at daily temporal resolution is a prerequisite to study lake dynamics in a changing environment. There are estimates of about 304 million lakes in the world, with over 17,000 lakes having a surface area of greater than 10 km² (Downing *et al.*, 2006). Out of this huge number of lakes, long-term *in-situ* data are available only for a small fraction. Moreover, the *in-situ* data vary substantially with lakes with respect to the method of collection, quality, time line of availability, etc. In order to address these issues, there are strong efforts within the lakes related research community (Global Lake Temperature Collaboration (GLTC), GLEON, NetLake, Globolakes (<http://www.globolakes.ac.uk/>) etc.) to gather, homogenise and disseminate *in-situ* data through common data platforms in order to benefit the larger scientific community (Sharma *et al.*, 2015).

In this context, the availability of satellite data at high temporal resolution and high spatial scale for the more than last thirty years is an opportunity not to be missed in developing a homogenised uniform time series of water temperature data for multiple lakes (Pareeth

et al., 2016). Such a time series will facilitate the study of long-term trends of lakes using a single source of uniformly processed data of similar quality covering the same time period. In this thesis, we established a new method which allows the development of homogenised time series of daily LSWT between 1986 and 2015 for alpine lakes from multiple moderate resolution polar orbiting satellites (see Chapter 3). The method was applied to five large lakes in Italy where long-term *in-situ* data is available to validate the satellite derived time series (Chapter 4). The new method is reproducible and extendable to any other larger lakes in the world where satellite coverage is available. On a larger scale, this work has the potential to fill the gap in lack of *in-situ* data availability in studying long-term thermal dynamics of lakes.

5.1.1 Resolving important geometrical issues of AVHRR data

The first part of the thesis (Chapter 2) was aimed to resolve the well documented issues of geometrical discrepancies with the historical satellite data from AVHRR sensor. From the possible satellite data sources for the time line 1986-2015, AVHRR sensor data aboard multiple NOAA satellites had the highest share. The earlier NOAA satellites (1986-2001) undergone clock drifts and attitude errors, resulting in discrepancies with geometrical accuracy of the acquired data (*Baldwin and Emery, 1995; Brunel and Marsouin, 2000*). The orbital decay over time further deteriorated the quality of data Fig. 2.1. This prohibited the wide usage of AVHRR data at its original resolution of 1.1 km. Hence we developed a new workflow (see Chapter 2) to correct for these errors and to derive geometrically aligning time series of brightness temperatures from thermal channels of AVHRR sensors on board eight NOAA satellites (*Pareeth et al., 2016*).

The new method was developed by chaining existing multiple open source software tools to achieve different objectives (Fig. 2.3). We developed two new data drivers (one for AVHRR LAC data from POD satellites and another for AVHRR LAC data from KLM satellites) in Python programming language as part of the Pytroll libraries (<http://pytroll.org/>). These new drivers which are now openly available as part of Pytroll, can read and calibrate all the AVHRR LAC level-1B radiances obtained from the CLASS data archive irrespective of the NOAA satellite. Nonetheless this study was focused on lakes, the newly developed data drivers can be used for any land or water based applications. Here we want to point out the larger significance of this study, as this method can be used beyond lakes to also study land/water surface using bio-geophysical variables which can be derived from AVHRR data. A two step validation procedure was performed to test the robustness of thermal calibration. The inter-platform validation between LSWT of Lake Garda derived from multiple NOAA satellites were compared and reported an average RMSE less than 1 °C (Fig. 2.7). This shows

the superior quality of thermal calibration (*Trishchenko et al.*, 2002; *Trishchenko*, 2002) and allow to combine data from these satellites to extend the time series of LSWT. Moreover, cross-platform validation against other similar sensors like ATSR1, ATSR2, AATSR and MODIS employed in this study also revealed high accuracy confirming robustness of the proposed method (Fig. 2.8). The NOAA satellite series ended with NOAA-19 which is still operational, though the AVHRR sensor continued to be flown on the MetOp series of satellites. This study does not include data from MetOp which could be of interest in the future. Moreover, the night time observations were not taken into consider in this study due to non-availability of *in-situ* data to validate the product.

For precise geo-rectification of AVHRR data, a feature matching technique based on computer vision algorithm SIFT was used to extract homologous GCP's from a pair of images. This approach was critical to the overall workflow as there was a need of automate the GCP extraction (in order to process 22,507 AVHRR images). This technique can retrieve matching points and is followed by the application of a robust filter to remove any outliers. The SIFT algorithm was able to detect suitable pairs of homologous points as input to the subsequent image to image rectification process. As a limitation, it fails in case of cloud coverage greater than 50% affecting an AVHRR image scene. The resulting overall RMSE was always in sub-pixel range (Fig. 2.6), which is critical in time series analysis. For this entire process of geo-rectification, we used the SIFT algorithm implemented in the Orfeo-Toolbox software along with the newly developed GRASS GIS addon module *m.gcp.filter* and the enhanced *i.rectify* module (*Inglada and Christophe*, 2009; *GRASS Development Team*, 2015). Importantly, our procedure does not correct for the surface elevation which may lead to pixel misplacements and could have adversary effects especially on small features like lakes. *Khlopenkov et al.* (2010) explains accurate geolocation of AVHRR historical time series which includes an ortho-correction scheme taking care of the accurate positioning of the pixels in complicated terrains. Hence it is very important to design and perform robust statistical outlier detection before any trend analysis is performed with the developed time series data. The workflow can be enhanced further by including an ortho-correction procedure in the workflow which would correct for the surface undulation.

In this study, we explored the availability of open source geospatial libraries to find a solution to one of the most pressing challenges in AVHRR satellite data processing. With this approach, new possibilities are now open to researchers who may develop their own custom based regional datasets of more than thirty years from the freely available AVHRR data, which would otherwise be difficult to achieve due to the lack of proper tools to accurately process them. More details on the methodology, software tools and their availability are explained in Chapter 2. Some of the possible geo-physical variables derivable from AVHRR data at 1 km

spatial resolution and at daily scale for the past thirty years are e.g. LST, NDVI, 2-band EVI, and land cover map time series.

5.1.2 Homogenisation of LSWT

The chapter 3 of the thesis deals with a new approach to develop a homogenised LSWT from multiple satellites. *Schneider and Hook* (2010) point out the need of a unified dataset with compensated inter-satellite biases. For addressing this need, we used the thermal data obtained from six sensors, on board of thirteen different satellites with an acquisition time varying between 08:00 to 17:00 UTC to derive daily LSWT (Fig. 3.2). We implemented a modified pattern technique based on diurnal cycles from sun rise to sun set which was used to implement the time correction procedure to standardize LSWT to 12:00 UTC. With this approach, the maximum possible observations originating from multiple satellites could be included in the daily 30 year time series. The Lake Garda case study gave promising results with an overall RMSE of 0.79 °C against *in-situ* data (Fig. 3.8). Further analysis was done separately on deep and shallow basins of Lake Garda to study the spatial variability of change in surface temperature, which became possible due to the high spatial resolution of the newly created LSWT product. The application of the same split-window algorithm with satellite specific coefficients to derive the LSWT from the TOA brightness temperatures ensured data compatibility. The cross-platform validation confirmed that the derived LSWT are comparable. This shows the superior quality of the split-window coefficients provided by *Jimenez-Munoz and Sobrino* (2008) in correcting atmospheric attenuation. Interestingly, the time correction procedure did not improved the overall RMSE. A reason might be that the variations in lake surface temperature during day time are too low, except for unusual drops due to sudden weather variations, to exhibit any large difference in the overall accuracy estimators before and after time correction (Fig. 3.4). Nevertheless, the time correction procedure is a crucial step due to orbital drifts of the earlier satellites and given the wide window of acquisition times during the day.

Due to the differences in physical properties of shallow and deep basins of Lake Garda, different biases were observed in deep and shallow basins. In shallow basin, there was an overall positive bias when compared to *in-situ* data which means the satellite data were underestimating the surface temperature, probably due to the heating of epilimnion layer due to low depths. While in the deep basin, a negative bias was observed, which means most of the satellite derived LSWT were overestimating LSWT compared to the *in-situ* bulk temperature (Fig. 3.8). The gap filling procedure using harmonic analysis was able to retain the seasonal and annual variations. But the gap filled LSWT was found to be over-smoothed, hence losing

some of the local variation. The trend analysis reported significant warming of $0.036\text{ }^{\circ}\text{C yr}^{-1}$ during summer and an annual warming of $0.020\text{ }^{\circ}\text{C yr}^{-1}$ (Fig. 3.9). This results are in line with the other studies which reported warming of lakes at global and European level (O'Reilly *et al.*, 2015; Adrian *et al.*, 2009).

This study developed a new method to homogenise LSWT derived from multiple satellites. We show that satellite derive LSWT match well with *in-situ* records of one of our case study sites – Lake Garda. To the best of our knowledge, based on a thorough literature search, we found that it is the first time a homogenised daily LSWT data series covering 30 years was developed from historical satellite observations. The split-window algorithm with satellite specific coefficients rather than lake specific values makes it possible to extend the dataset to other lakes. The algorithm by Jimenez-Munoz and Sobrino (2008) has been implemented as a new add-on for GRASS GIS called *i.lswt* and is available with the current stable version of GRASS GIS (<https://grass.osgeo.org/grass70/manuals/addons/i.lswt.html>). Moreover, with the addition of more satellites in orbit like MetOp-A/B, Sentinel-3A with similar sensors, more thermal data become available to the public which could be easily processed by the new methodology. This method is introduced as a possible economical and feasible alternative to long-term *in-situ* data. Though the accuracy obtained from the validation procedure is high, comparing it to daily *in-situ* data obtained at the exact time of satellite data acquisition would be an ideal approach to test the stability of new datasets. With optical data, the issue with undetected clouds could result in spurious trends. Here, by applying harmonic analysis we were able to remove outliers at the expense of losing local variability. But for any surface other than water, this may not be suitable and require different spatio-temporal interpolation techniques to fill the gaps due to clouds (Metz *et al.*, 2014; Neteler, 2010). Moreover, the spatial resolution of 1 km is the highest which we can obtain from moderate resolution sensors, but may not be suitable for small lakes.

5.1.3 Warming of large lakes in Italy

In the last part of the thesis (Chapter 4), the newly developed methods (covered in Chapters 2 and 3) were used to develop a time series of homogenised daily LSWT covering thirty years (1986-2015) at a spatial resolution of 1 km. Specifically, it comprises five large lakes in Italy: Lake Garda, Lake Iseo, Lake Como, Lake Maggiore and Lake Trasimeno (Fig. 4.1). The selection of lakes was based on availability of long-term *in-situ* data to validate the satellite derived product. The validation results showed similar results in all the lakes with an overall RMSE of $1.2\text{ }^{\circ}\text{C}$ and R^2 of 0.98 (Table 4.2). The linear models developed with *in-situ* data as dependent variable and LSWT as independent variable reported similar slopes for all the

lakes, documenting the stability of the new LSWT datasets for the case study lakes. LSWT of Lake Trasimeno being the only shallow lake among the five lakes under study showed a positive bias against *in-situ* data. While LSWT of all the deep sub-Alpine lakes reported a negative bias, due to a cooler epilimnion compared to the surface layer (Fig. 4.4). The observed biases were in line with the results obtained between deep and shallow basins of Lake Garda.

The trend analysis reported significant warming during summer at an average rate of $0.03\text{ }^{\circ}\text{C yr}^{-1}$ over all the lakes (Table 4.3). Higher warming rates were reported for Lake Trasimeno and Lake Garda. Annually, an average rate of $0.017\text{ }^{\circ}\text{C yr}^{-1}$ was reported for all lakes. The warming rates obtained from the new homogenised LSWT were found to be in line with similar studies using different sources of data (Table 4.4). The temporal patterns during summer were found to be similar among lakes with the year 2003 being the hottest reported in the last thirty years (Fig. 4.5). Summer mean LSWT were found to be temporally coherent among the lakes. The highest coherence was reported between Lake Garda and Trasimeno (Fig. 4.5). This may be due to the influence of Mediterranean weather on Lake Garda during summer time. However, mean summer LSWT of other sub-Alpine lakes showed less coherence with that of Lake Trasimeno and were found to be in decreasing order with respect to the spatial distance between the lakes. The annual mean LSWT was found to be highly coherent between sub-Alpine lakes while exhibiting lower coherence with LSWT of Lake Trasimeno (Fig. 4.6). The local meteorological conditions above these lakes play a major role in determining the coherence between mean LSWT's. The significant part of the shallower area of Lake Garda is part of the Po river plain, i.e. outside the Alpine mountain chain where the dominant Mediterranean climate exert its major influence. Conversely, the northern part of Lake Garda and the other large subalpine lakes are more influenced by Alpine climate. Moreover, the presence of a wider shallower area characterizing the southern portion of Lake Garda exposes the lake to a greater influence of the Mediterranean climate, and to a greater impact of short and medium frequency climatic fluctuations acting especially on the southern basin. These factors could contribute to explain the higher coherence in the temporal development of water temperature in Lake Garda and Lake Trasimeno. On the other hand, deep sub-Alpine lakes considered in this study undergo complex mixing processes during winter and spring, which have a strong influence on surface water temperatures (*Salmaso et al.*, 2014). This explains the weaker coherence of sub-Alpine lakes with Lake Trasimeno on an annual scale. Moreover, the long-term annual trend using GAMM showed a clear upward slope with a similar rate of warming as obtained from Mann-Kendall tests. The intra-annual variations obtained from GAMM further showed the clear distinctive summer and winter patterns of Lake Trasimeno different from sub-Alpine lakes (Fig. 4.8).

In this study the usability of new satellite derived LSWT in detecting the long-term variations over large lakes in Italy was demonstrated. The results showed rapid warming during summer and upward annual trends. With GAMM, we could reconstruct the intra-annual variations showing distinctive pattern for Lake Trasimeno, also estimated the long-term trend as the smoothing was linear. For such analysis, it is crucial to have long-term data with high temporal frequency which could be obtained by using satellite data as an alternative to sparse *in-situ* data.

5.2 Conclusion and future perspectives

In this thesis, new methods to develop long-term homogenised daily LSWT of thirty years (1986-2015) using freely available satellite data and remote sensing techniques are presented (Chapters 2 and 3). Open source libraries were used throughout the study to attain reproducibility and extensibility. The newly developed code snippets are publicly available wherever applicable (Appendices A.1 and A.2). The LSWT developed using new method for the large five lakes in Italy were used to estimate the summer warming trends from 1986 to 2015. The study demonstrated the usability of satellite data in studying long-term variations of surface water temperature (Chapters 4). With the newly developed LSWT dataset, direct effects of climate change like warming of surface temperature has been studied in detail. The new LSWT data can be used to understand deeply the climatological behaviour and its dependencies (*Layden et al.*, 2015). It can also be used to find the cooling and warming periods, peak LSWT day in a year, ice melting days, their inter-annual variations etc. As a continuation of this thesis, next step would be to understand the indirect ecological consequences of warming due to climate change on these lakes. Another aspect will be to look into the influence of larger climatic forcings like North Atlantic Oscillation (NAO) and Eastern Atlantic (EA) climatic oscillations on the long-term changes in the LSWT of large lakes south of Alps and Central Italy. This new methods gives an opportunity to develop thirty years of unified LSWT from single source for all the large lakes at a global scale. Though, such an extension demands a lot of processing power, the results will be of high value to the scientific community to understand the long-term LSWT dynamics. Given recent extreme events like disappearance of large lakes around the world (*Micklin*, 1988, 2007; *Birkett*, 2000; *AghaKouchak et al.*, 2015), satellite based homogenised LSWT could play a key role in understanding the patterns and drivers of climate change related variations.

List of Figures

1.1	Electromagnetic spectrum. The wavelengths of visible and thermal region are shown separately.	1
1.2	Principle of remote sensing. The emitted and reflected radiation from the Earth's surface is recorded by the satellite and sent to a ground receiver as radiances. VZA - Viewing Zenith Angle; SZA - Solar Zenith Angle	3
2.1	Plot of variation in observation times of National Oceanic and Atmospheric Administration (NOAA) instruments, note the large orbital drifts of the earlier NOAA-9/11/12/14 instruments.	16
2.2	Data distribution. (a) Distribution of AVHRR Local Area Coverage (LAC) images over study area from 1986 to 2014 obtained from NOAA Comprehensive Large Array-data Stewardship System (CLASS) archive, separated by different instruments. (b) Time line of NOAA instruments used in this study.	20
2.3	Workflow diagram. On top, the horizontal arrows show the main steps involved in the workflow, the software are listed vertically on the left. Role of each software in the workflow is explained in the corresponding cells.	21
2.4	NIR (Near Infrared) channel 2 of AVHRR LAC data resampled using (a) Nearest Neighbour interpolation; (b) interpolated using the gradient search algorithm. The image is NOAA-18 AVHRR LAC taken on 1 August 2011.	27

2.5 (a–c) explains geometric correction done on NOAA-14 AVHRR LAC data (NIR channel 2 shown here) acquired on 09 August 1997. (a₁) The input image with lake and country boundaries overlaid. Also shown is blue (input) and red (reference) crosses, the Ground Control Points (GCPs) extracted using the SIFT algorithm; (b₁) The reference image used to extract the GCP points, NOAA-19 AVHRR LAC acquired on 01 August 2012; (c) The NOAA-14 AVHRR LAC image after polynomial second level rectification using the GCP's; (a₂) Fig. a₁ zoomed to the sub-alpine lakes in Northern Italy; & (b₂) Fig. b₁ zoomed to the sub-alpine lakes in Northern Italy. 28

2.6 Validation of Scale Invariant Feature Transform (SIFT) algorithm. (a) Independent GCP point-cloud plotted in x-deviation, y-deviation; (b) Kernel density plot of x and y deviations. 29

2.7 Boxplots of absolute difference between lake mean Lake Surface Water Temperature (LSWT) estimated from multiple NOAA same day observations; horizontal thick line inside the box represents median; lower and upper end of the box represents first and third quartiles respectively; the bottom whisker ranges from first quartile to the smallest non-outlier and the top whisker ranges from third quartile to the largest non-outlier; the dots outside whiskers are outliers. 30

2.8 Boxplots of absolute difference between lake mean LSWT estimated from same day observations from NOAA and other instruments; horizontal thick line inside the box represents median; lower and upper end of the box represents first and third quartiles respectively; the bottom whisker ranges from first quartile to the smallest non-outlier and the top whisker ranges from third quartile to the largest non-outlier; the dots outside whiskers are outliers. 31

3.1 Study area map of Lake Garda; dark blue area depicts the deep basin and light blue area depicts the shallow basin. Push pins represent locations of *in-situ* water temperature monitoring. The maps were generated using the software GRASS GIS 7.0 *Neteler and Mitasova (2008)* (URL - <https://grass.osgeo.org/grass7/>) 41

3.2 Availability of moderate resolution satellite data between 1986 and 2015 . . . 43

3.3 Plot of variation in observation times of all the satellites used in this study; Note the large orbital drifts of the earlier NOAA-9/11/12/14 instruments. . . 46

3.4	Monthly diurnal cycles during day time estimated using long-term hourly climatologies from satellite observations. The points in plot are the long-term hourly averages for a particular month, and the red line is DTC model fit. . . .	47
3.5	a) Summer mean LSWT of year 1992, b) summer mean LSWT of year 2003, c) difference map between summer means of 2003 and 1992. The boundary layer shown in black over Lake Garda is the inner buffer used to mask out the edge pixels. The maps were generated using the software GRASS GIS 7.0 <i>Neteler and Mitasova</i> (2008) (URL - https://grass.osgeo.org/grass7/) . .	49
3.6	Plots showing gap-filled reconstructed LSWT using HANTS (red line) over the averaged homogenised LSWT (grey bars) from multiple observations over a particular day of the year. The daily LSWT time series from the deep and shallow basins for the year 2003 is shown as an example.	52
3.7	Boxplots of lake mean LSWT difference between final homogenised LSWT and corresponding in-situ data. The values shown in black over the lower whisker represents N and the values in red is the reported RMSE in °C.	53
3.8	Annual (top) and summer (bottom) trends of lake mean LSWT derived from the new homogenised daily LSWT between 1986 to 2015. Data were smoothed using the loess interpolation (blue line). The gray area is the 95% confidence interval.	54
3.9	Annual trends of lake mean LSWT computed from homogenised LSWT (deep basin), <i>in-situ</i> data and satellite derived LSWT <i>Riffler et al.</i> (2015). Data is smoothed using the loess interpolation (blue line). The area shown around the smoothed line is the 95% confidence interval.	55
4.1	Study area in Italy depicting location of the study lakes and the bounding box used for image analysis (red box). Each lake and the location for the <i>in-situ</i> water temperature measurements (marked as red triangles) are shown in sub figures 1–5. 1) Lake Garda; 2) Lake Iseo; 3) Lake Como; 4) Lake Maggiore; 5) Lake Trasimeno	62
4.2	Graph showing time line of the thirteen satellites used in this study, color coded according to sensors. Y-axis represents the name of satellites used in this study. The legend represents the sensor on board each satellite.	65

4.3 Block diagram of entire workflow implemented to derive time series of homogenised LSWT with level-1B thermal radiances obtained from thirteen polar orbiting satellites in the time window from 1986 to 2015. 66

4.4 Boxplots of the difference between *in-situ* and lake mean LSWT estimated from satellites. The outliers are removed using an inter-quartile filter with 10 and 90 percentile cutoffs. Number of observations (N) in each lake used in the assessment is shown over the boxplot. 68

4.5 Summer mean LSWT derived from the new homogenised daily LSWT between 1986 to 2015. Data were smoothed using the linear model (blue line). The gray area represents 95% confidence interval 69

4.6 a) Scatter plot matrix showing temporal coherence between summer mean LSWT of all the lakes. Pearson’s correlation coefficient is given as measure of coherence. b) Dendrogram showing the clustering of lakes based on summer mean LSWT with 1 - Pearson’s correlation coefficient the dissimilarity measure. 71

4.7 a) Scatter plot matrix showing temporal coherence between annual mean LSWT of all the lakes. Pearson’s correlation coefficient is given as measure of coherence. b) Dendrogram showing the clustering of lakes based on annual mean LSWT with 1 - Pearson’s correlation coefficient the dissimilarity measure. 72

4.8 a) Annual cycles and b) long-term trends from the GAMM analysis. For the annual cycles, the x-axis represents day of the year and for long-term trend, the x-axis represents count of days from 1986 to 2015 representing the time factor in the model. In (b) the trend lines of the lakes Garda and Iseo practically coincide and are not distinguishable. 73

List of Tables

1.1	Important specifications of sensors with dual thermal channels in the 10.5-12.5 μm wavelength range. These are the satellites used in this study.	5
2.1	Spectral resolution of Advanced Very High Resolution Radiometer (AVHRR) sensors in μm	16
2.2	List of software packages used in this study (License type urls: http://www.gnu.org/copyleft/gpl.html , http://www.cecill.info/licences/Licence_CeCILL_V2-en.html).	19
2.3	Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) (in $^{\circ}\text{C}$) between lake mean LSWT estimated between pair of sensors.	32
2.4	Correlation statistics between <i>in-situ</i> data and corresponding LSWT derived from various NOAA instruments; N = Number of samples used in the linear model.	32
3.1	RMSE in $^{\circ}\text{C}$ reported between same day observations of different pair of satellites for the deep and shallow lake basin. Number inside brackets represents N - Number of observations.	50
3.2	RMSE in $^{\circ}\text{C}$ reported at deep basin and shallow basin from absolute difference of LSWT between <i>in-situ</i> data and satellite derived LSWT. Number inside brackets refer to N - the number of same day observations.	51
4.1	Main characteristics of the investigated lakes and type of <i>in-situ</i> data used in this study	63

4.2 Results of validation of the homogenised LSWT derived from satellite data against the *in-situ* data of the lakes. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are reported in °C. N is the number observations used in the linear models. Coefficient of determination (R^2) and slope are obtained from the linear models 67

4.3 Summer (June/July/August) trends obtained from two different sources of LSWT data. The first column lists the trends obtained from new homogenised LSWT over thirty years (1986-2015). Second column lists the summer trends from the same data but for a reduced time scale (1989-2013) to match the data from *Riffler et al. (2015)*. 68

4.4 Annual trend obtained from Generalized Additive Mixed Model (GAMM) and Mann-Kendall test for all the lakes. The long-term daily trend from GAMM is divided by 30 (number of years) to obtain the annual trend listed below. . . . 70

Appendix

A.1 Python script to read and calibrate level 1B AVHRR LAC data using Pytroll libraries

```
# -*- coding: utf-8 -*-
"""
Created on Fri Feb 25 2015
filename = pytroll.py
This script is executed by the workflow shell script (see Appendix B).
This Python script is generic, will not work as such, will need local tweaking.
@author: Sajid Pareeth
@email: spareeth@gmail.com
"""
import os
#set following variables first
os.environ['PPP_CONFIG_DIR'] = '/usr/local/src/mpop/etc'
os.environ['PYGAC_CONFIG_FILE'] = '/usr/local/src/pygac/etc/pygac.cfg'
# conversion of avhrr l1b bands to tiff starts here...
import numpy as np
import pyresample
import pyproj
import geotiepoints
import argparse
import datetime
from pyresample import import *
from pyresample.gradient_search import gradient_search
from pyresample.geometry import SwathDefinition
```

```
from pyresample.geometry import AreaDefinition
from mpop.satellites import PolarFactory
from mpop.projector import get_area_def
from mpop.satin.lac_l1b import KLMReader

#Pytroll starts here
reader = KLMReader()
ts = datetime.datetime.strptime('{year} {doy} {hour} {minute}', '%y %j %H %M')
reader.read("{ipdir}/{name}".format(ipdir=mipdir,name=mname))
reader.get_lonlat()
area_swath = SwathDefinition(reader.lons, reader.lats)
channels = reader.get_calibrated_channels()
sat_azimuth, sat_zenith, sun_azimuth, sun_zenith, rel_azimuth = reader.get_angles()
ID=reader.instrument_id
global_data = PolarFactory.create_scene("noaa", "{ID}".format(ID=ID), "avhrr", ts)
area_def = get_area_def("euro_laea_AVHRR")
scene = global_data
global_data.area = area_swath
# To save zenith angle as tiff, use the following line...
channels[:, :, 2] = sat_zenith
scene[0.63] = channels[:, :, 0]
scene[0.9125] = channels[:, :, 1]
scene[3.74] = channels[:, :, 2]
scene[10.8] = channels[:, :, 3]
scene[12.0] = channels[:, :, 4]
scene.area = area_swath
```

```

l = scene.project("euro_laea_AVHRR")
l1 = gradient_search(scene[0.63].data.astype(np.float64), scene[0.63].area.lons, scene[0.63].area.lats, area_def)
l[0.63] = np.ma.masked_values(l1, 0)
b1 = l.image(0.63, mode="L")
b1.save("{opdir}/NSS.LHRR.{inst}.D{year}{doy}.S{hour}{minute}_b1.tif", floating_point=True)
l2 = gradient_search(scene[0.9125].data.astype(np.float64), scene[0.9125].area.lons, scene[0.9125].area.lats, \
area_def)
l[0.9125] = np.ma.masked_values(l2, 0)
b2 = l.image(0.9125, mode="L")
b2.save("{opdir}/NSS.LHRR.{inst}.D{year}{doy}.S{hour}{minute}_b2.tif", floating_point=True)
lz = gradient_search(scene[3.74].data.astype(np.float64), scene[3.74].area.lons, scene[3.74].area.lats, area_def)
l[3.74] = np.ma.masked_values(lz, 0)
bz = l.image(3.74, mode="L")
bz.save("{opdir}/NSS.LHRR.{inst}.D{year}{doy}.S{hour}{minute}_bz.tif", floating_point=True)
l4 = gradient_search(scene[10.8].data.astype(np.float64), scene[10.8].area.lons, scene[10.8].area.lats, area_def)
l[10.8] = np.ma.masked_values(l4, 0)
b4 = l.image(10.8, mode="L")
b4.save("{opdir}/NSS.LHRR.{inst}.D{year}{doy}.S{hour}{minute}_b4.tif" floating_point=True)
l5 = gradient_search(scene[12.0].data.astype(np.float64), scene[12.0].area.lons, scene[12.0].area.lats, area_def)
l[12.0] = np.ma.masked_values(l5, 0)
b5 = l.image(12.0, mode="L")
b5.save("{opdir}/NSS.LHRR.{inst}.D{year}{doy}.S{hour}{minute}_b5.tif", floating_point=True)
#####ENDS HERE#####
if __name__ == "__main__":
    main()

```

A.2 Shell script to process the level 1B AVHRR LAC data - entire workflow

Note about script formatting: Long lines are broken with `\\` which means that the command continues on the next line. This symbol is usually not necessary when typing but it is used here for formatting reasons.

```
#!/bin/sh
## Processing NOAA level 1B AVHRR LAC images###
## Author: Sajid Pareeth, 2015
##The code below will not run as such, you have to adapt it to the local environment
##The below code should run inside GRASS GIS session

##Exit strategy if you are not inside GRASS GIS
if [ -z "$GISBASE" ] ; then
    echo "You must be in GRASS GIS to run this program." >&2
    exit 1
fi
##Setting the GRASS environments
export GRASS_OVERWRITE=1
export GRASS_MESSAGE_FORMAT=plain # percent output as 0..1..2..
# setting environment, so that awk works properly in all languages
unset LC_ALL
LC_NUMERIC=C
export LC_NUMERIC
##Setting the study area region
g.region n=2531000 s=2480000 w=4360000 e=4390000 res=$RES -a
#setting the Python environment:
PYVERSION='python --version 2>&1 | cut -d' ' -f2 | cut -d',' -f1-2'
```

```
MYPYSITES=/usr/local/lib64/python$PYVERSION
# for runtime, attach to existing PYTHONPATH:
export PYTHONPATH=$PYTHONPATH:$MYPYSITES/site-packages
#setting paths required for PYTROLL
PPP_CONFIG_DIR=/home/sajid/mpop_etc
export PPP_CONFIG_DIR=$PPP_CONFIG_DIR
PYGAC_CONFIG_FILE=/home/sajid/pygac_etc/pygac.cfg
export PYGAC_CONFIG_FILE=$PYGAC_CONFIG_FILE
for yyyy in `seq 1986 2014`; do #The loop over all the days from 1986 to 2014 starts here
    cd ${MYDATA}
    yy=`echo ${yyyy}|cut -c3-4`
    leap=`is_leap_year.sh $yyyy` #small script to check leap year or not
    if [ $leap -eq 1 ]; then
nd=366
    else
nd=365
    fi
    for d in `seq 1 $nd`; do
        cd ${MYDATA}
        dy=`echo $d | awk '{ printf("%03d\n", $1) }'`
        NUM=`ls NSS.LHRR.NP.D${yy}${dy}*|wc -l`
    if [ ${NUM} -eq 0 ]; then
        echo "No images on ${yyyy}${dy}"
        continue
    else
        for y in `ls NSS.LHRR.NP.D${yy}${dy}*`; do
```

```
echo "#####Processing ${y} starts here#####"
cd ${MYDATA}/"LACdata"${yyyy}"_NOheader"
# Set the region
g.region n=$N s=$S w=$W e=$E res=$RES -a
YEAR='echo ${y}|cut -c14-15'
DOY='echo ${y}|cut -c16-18'
HOURL='echo ${y}|cut -c21-22'
MIN='echo ${y}|cut -c23-24'
TIME='echo ${y}|cut -c21-24'
i='echo ${y}|cut -c1-24'
j='echo ${y}|cut -c1-18'
if [ ${TIME} -ge ${MINTIME} -a ${TIME} -le ${MAXTIME} ]; then
    echo "##PYTROLL starts here##"
    #below pytroll.py should be modified from Appendix A with parsing.
    python pytroll.py -i ${y} -o ${OUTDIR}
    echo "##PYTROLL ends here##"
    #Geo-correction using OTBcli_homologous points
    otbcli_HomologousPointsExtraction -in1 input_b1.tif -band1 1 -in2 ref_b1.tif -band2 1 -algorithm sift \\
        -mode full -out $OUTB1
    otbcli_HomologousPointsExtraction -in1 input_b2.tif -band1 1 -in2 ref_b2.tif -band2 1 -algorithm sift \\
        -mode full -out $OUTB2
    otbcli_HomologousPointsExtraction -in1 input_b4.tif -band1 1 -in2 ref_b4.tif -band2 1 -algorithm sift \\
        -mode full -out $OUTB3
    otbcli_HomologousPointsExtraction -in1 input_b5.tif -band1 1 -in2 ref_b5.tif -band2 1 -algorithm sift \\
        -mode full -out $OUTB4
    echo "##Feature matchin (SIFT) using OTB ends here##"
```

```

##To make the tie-points from SIFT Grass compatible
cat $OUTB1 $OUTB2 $OUTB3 $OUTB4 > ${MYTMPDIR}/${i}_all.txt
## adding a column with enable/disable
awk '{ $5="1\t"$5 }' ${MYTMPDIR}/${i}_all.txt > ${MYTMPDIR}/${i}_all_GRASS.txt
# proceed in the GRASS GIS database; importing the input TIFF files
r.in.gdal input=${y}_b1.tif output=${i}_b1 memory=${MEMORY}
r.in.gdal input=${y}_b2.tif output=${i}_b2 memory=${MEMORY}
r.in.gdal input=${y}_bz.tif output=${i}_zenith memory=${MEMORY}
r.in.gdal input=${y}_b4.tif output=${i}_b4 memory=${MEMORY}
r.in.gdal input=${y}_b5.tif output=${i}_b5 memory=${MEMORY}
i.group group=${y} input=${y}_b1,${y}_b2,${y}_b4,${y}_b5,${y}_zenith
mv ${MYTMPDIR}/${i}_all_GRASS.txt ${GRASSLOC}/group/${i}/POINTS #Moving the POINTS file to the GRASS group folder
PTCNT='wc -l ${GRASSLOC}/group/${i}/POINTS|cut -d' ' -f1' ##GCP filtering and geo-rectification
unset use
if [ ${PTCNT} -le 3 ]; then
    g.remove group name=${i} -f
    echo "Not enough gcps to filter (< 3), hence ${y} is not usable"
    continue
else
    eval 'm.gcp.filter group=${i} order=1 threshold=500 -b'
    USE=${use}
    i.target group=${i} -c
fi

if [ ${USE} -lt 20 ]; then
    g.remove group name=${i} -f

```

```

    echo "${y} is not usable due to lack of enough homologous points - ${USE}, hence avoiding"
    continue
elif [ ${USE} -gt 300 ]; then
    m.gcp.filter group=${i} order=2 threshold=500 -b -u
    i.rectify -a group=${i} extension=_rectified order=2 method=nearest --o
fi
unset use
r.colors map=${i}_b1_rectified rules=${CLRDIR}/avhrrb1.clr
r.colors map=${i}_b2_rectified rules=${CLRDIR}/avhrrb2.clr
r.colors map=${i}_b4_rectified color=kelvin
r.colors map=${i}_b5_rectified color=kelvin
echo "SPARC Cloud detection starts here" #Cloud mask based on SPARC
if [ ${MIN} -lt 30 ]; then
    ECMWFHR=${HOUR}
    elif [ ${MIN} -gt 30 ] && [ ${HOUR} -eq 23 ]; then
        ECMWFHR=00
else
    ECMWFHR='echo $(( ${HOUR} + 1 ))'
fi
D='echo $(( ${d} - 1 ))'
DATE='date -d "${D} days ${yyyy}-01-01" +"%d%m%Y"'
M=00
echo "${yyyy}${DOY}_ecmwf@sp_ecmwf is used for cloud removal" #offset and scale factors from Trischenko et.al 2006
r.mapcalc "${i}_Tindex = 1.0 * ( ${i}_b4_rectified - ${yyyy}${DOY}_ecmwf_hants@sp_ecmwf + 6.0 ) * -0.42"
r.mapcalc "${i}_Bindex_land = 1.0 * ( ${i}_b1_rectified - 0.30 ) * 67"
r.mapcalc "${i}_Bindex_water = 1.0 * ( ${i}_b2_rectified - 0.24 ) * 75"

```

```

    r.mapcalc "${i}_Cindex = 1.0 * (${i}_b4_rectified - ${i}_b5_rectified - 1.5) * 4.0"
r.mapcalc << EOF
${i}_b4_masked = eval( \
T_test      = if(${i}_Tindex > 8, null(), ${i}_b4_rectified), \
C_test      = if(${i}_Cindex > 15, null(), T_test), \
viewangle_test = if(${i}_zenith_rectified > 45, null(), C_test), \
viewangle_test)
EOF

r.mapcalc << EOF
${i}_b5_masked = eval( \
T_test      = if(${i}_Tindex > 8, null(), ${i}_b5_rectified), \
C_test      = if(${i}_Cindex > 15, null(), T_test), \
viewangle_test = if(${i}_zenith_rectified > 45, null(), C_test), \
viewangle_test)
EOF

r.mapcalc "${i}_zenith_masked = if(isnull(${i}_b4_masked), null(), ${i}_zenith_rectified)"
echo "SPARC Cloud detection ends here"
else
    echo "${y} is outside the time frame, hence avoiding";
    continue
fi
done
fi
done
echo "#####Processing ${yyyy} finishes here#####"
done

```

Bibliography

- Adrian, R., C. M. O'Reilly, H. Zagarese, S. B. Baines, D. O. Hessen, W. Keller, D. M. Livingstone, R. Sommaruga, D. Straile, E. Van Donk, G. A. Weyhenmeyer, and M. Winder (2009), Lakes as sentinels of climate change, *Limnology and oceanography*, 54(6), 2283–2297.
- AghaKouchak, A., H. Norouzi, K. Madani, A. Mirchi, M. Azarderakhsh, A. Nazemi, N. Nasrollahi, A. Farahmand, A. Mehran, and E. Hasanzadeh (2015), Aral Sea syndrome desiccates Lake Urmia: Call for action, *Journal of Great Lakes Research*, 41(1), 307–311, doi:10.1016/j.jglr.2014.12.007.
- Ashouri, H., K.-L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat (2014), PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies, *Bulletin of the American Meteorological Society*, 96(1), 69–83, doi:10.1175/BAMS-D-13-00068.1.
- Baldwin, D., and W. J. Emery (1995), Spacecraft attitude variations of NOAA-11 inferred from AVHRR imagery, *International Journal of Remote Sensing*, 16(3), 531–548, doi:10.1080/01431169508954417.
- Becker, F., and Z.-L. Li (1995), Surface temperature and emissivity at various scales: Definition, measurement and related problems, *Remote Sensing Reviews*, 12(3-4), 225–253, doi:10.1080/02757259509532286.
- Birkett, C. M. (2000), Synergistic Remote Sensing of Lake Chad: Variability of Basin Inundation, *Remote Sensing of Environment*, 72(2), 218–236, doi:10.1016/S0034-4257(99)00105-4.
- Blenckner, T., R. Adrian, D. M. Livingstone, E. Jennings, G. A. Weyhenmeyer, D. G. George, T. Jankowski, M. Järvinen, C. N. Aonghusa, T. Nöges, D. Straile, and K. Teubner (2007), Large-scale climatic signatures in lakes across Europe: a meta-analysis, *Global Change Biology*, 13(7), 1314–1326, doi:10.1111/j.1365-2486.2007.01364.x.
- Bordes, P., P. Brunel, and A. Marsouin (1992), Automatic Adjustment of AVHRR Navi-

gation, *Journal of Atmospheric and Oceanic Technology*, 9(1), 15–27, doi:10.1175/1520-0426(1992)009<0015:AAOAN>2.0.CO;2.

Bresciani, M., D. Stroppiana, D. Odermatt, G. Morabito, and C. Giardino (2011), Assessing remotely sensed chlorophyll-a for the implementation of the Water Framework Directive in European perialpine lakes, *Science of The Total Environment*, 409(17), 3083–3091, doi:10.1016/j.scitotenv.2011.05.001.

Briand, J.-F., C. Leboulanger, J.-F. Humbert, C. Bernard, and P. Dufour (2004), *Cylindrospermopsis Raciborskii* (cyanobacteria) Invasion at Mid-Latitudes: Selection, Wide Physiological Tolerance, Orglobalwarming?1, *Journal of Phycology*, 40(2), 231–238, doi:10.1111/j.1529-8817.2004.03118.x.

Brovelli, M. A., H. Mitasova, M. Neteler, and V. Raghavan (2012), Free and open source desktop and Web GIS solutions, *Applied Geomatics*, 4(2), 65–66, doi:10.1007/s12518-012-0082-4.

Brunel, P., and A. Marsouin (2000), Operational AVHRR navigation results, *International Journal of Remote Sensing*, 21(5), 951–972, doi:10.1080/014311600210371.

Buzzi, F. (2002), Phytoplankton assemblages in two sub-basins of Lake Como, *Journal of Limnology*, 61(1), 117–128.

Campbell, J. B., and R. H. Wynne (2011), *Introduction to Remote Sensing, Fifth Edition*, Guilford Press.

Cao, C., M. Weinreb, and J. Sullivan (2001), Solar contamination effects on the infrared channels of the advanced very high resolution radiometer (AVHRR), *Journal of Geophysical Research*, 106(D24), 33,463, doi:10.1029/2001JD001051.

Coats, R., J. Perez-Losada, G. Schladow, R. Richards, and C. Goldman (2006), The Warming of Lake Tahoe, *Climatic Change*, 76(1-2), 121–148, doi:10.1007/s10584-005-9006-1.

Coll, C., and V. Caselles (1997), A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison, *Journal of Geophysical Research: Atmospheres*, 102(D14), 16,697–16,713, doi:10.1029/97JD00929.

Cracknell, A. P. (1997), *advanced very high resolution radiometer AVHRR*, CRC Press.

Crosman, E. T., and J. D. Horel (2009), MODIS-derived surface temperature of the Great Salt Lake, *Remote Sensing of Environment*, 113(1), 73–81, doi:10.1016/j.rse.2008.08.013.

- Czajkowski, K. P., S. N. Goward, and H. Ouaidrari (1998), Impact of AVHRR filter functions on surface temperature estimation from the split window approach, *International Journal of Remote Sensing*, 19(10), 2007–2012, doi:10.1080/014311698215126.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597, doi:10.1002/qj.828.
- Dokulil, M. T., A. Jagsch, G. D. George, O. Anneville, T. Jankowski, B. Wahl, B. Lenhart, T. Blenckner, and K. Teubner (2006), Twenty years of spatially coherent deepwater warming in lakes across Europe related to the North Atlantic Oscillation, *Limnology and Oceanography*, 51(6), 2787–2793, doi:10.4319/lo.2006.51.6.2787.
- Downing, J. A., Y. T. Prairie, J. J. Cole, C. M. Duarte, L. J. Tranvik, R. G. Striegl, W. H. McDowell, P. Kortelainen, N. F. Caraco, J. M. Melack, and J. J. Middelburg (2006), The global abundance and size distribution of lakes, ponds, and impoundments, *Limnology and Oceanography*, 51(5), 2388–2397, doi:10.4319/lo.2006.51.5.2388.
- Duan, S.-B., Z.-L. Li, B.-H. Tang, H. Wu, R. Tang, Y. Bi, and G. Zhou (2014), Estimation of Diurnal Cycle of Land Surface Temperature at High Temporal and Spatial Resolution from Clear-Sky MODIS Data, *Remote Sensing*, 6(4), 3247–3262, doi:10.3390/rs6043247.
- Emery, W. J., J. Brown, and Z. P. Nowak (1989), AVHRR image navigation - Summary and review, *Photogrammetric Engineering and Remote Sensing*, 4, 1175–1183.
- Fan, B., C. Huo, C. Pan, and Q. Kong (2013), Registration of Optical and SAR Satellite Images by Exploring the Spatial Relationship of the Improved SIFT, *IEEE Geoscience and Remote Sensing Letters*, 10(4), 657–661, doi:10.1109/LGRS.2012.2216500.
- Fink, G., M. Schmid, B. Wahl, T. Wolf, and A. Wüest (2014), Heat flux modifications related to climate-induced warming of large European lakes, *Water Resources Research*, pp. n/a–n/a, doi:10.1002/2013WR014448.
- Fontana, F. M., N. C. Coops, K. V. Khlopenkov, A. P. Trishchenko, M. Riffler, and M. A. Wulder (2012), Generation of a novel 1km NDVI data set over Canada, the northern United States, and Greenland based on historical AVHRR data, *Remote Sensing of Environment*, 121, 171–185, doi:10.1016/j.rse.2012.01.007.

- Gebbert, S., and E. Pebesma (2014), A temporal GIS for field based environmental modeling, *Environmental Modelling & Software*, 53, 1–12, doi:10.1016/j.envsoft.2013.11.001.
- George, D., D. Hewitt, E. Jennings, N. Allott, and P. McGinnity (2005), The impact of changes in the weather on the surface temperatures of windermere (uk) and lough feeagh (ireland), in *Proceedings of the Fourth Inter-Celtic Colloquium on Hydrology and Management of Water Resources, Guimaraes, Portugal*, vol. 310, pp. 86–93.
- Gerten, D., and R. Adrian (2000), Climate-driven changes in spring plankton dynamics and the sensitivity of shallow polymictic lakes to the North Atlantic Oscillation, *Limnology and Oceanography*, 45(5), 1058–1066, doi:10.4319/lo.2000.45.5.1058.
- Ghent, D. (2012), Land surface temperature validation and algorithm verification, *Report to European Space Agency*, pp. 1–17.
- GRASS Development Team (2015), *Geographic Resources Analysis Support System (GRASS GIS) Software*, Open Source Geospatial Foundation, USA.
- Heidinger, A. K., W. C. Straka, C. C. Molling, J. T. Sullivan, and X. Wu (2010), Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record, *International Journal of Remote Sensing*, 31(24), 6493–6517, doi:10.1080/01431161.2010.496472.
- Hook, S. J., F. J. Prata, R. E. Alley, A. Abtahi, R. C. Richards, S. G. Schladow, and S. Pálmarsson (2003), Retrieval of Lake Bulk and Skin Temperatures Using Along-Track Scanning Radiometer (ATSR-2) Data: A Case Study Using Lake Tahoe, California, *Journal of Atmospheric and Oceanic Technology*, 20(4), 534–548, doi:10.1175/1520-0426(2003)20<534:ROLBAS>2.0.CO;2.
- Hulley, G. C., S. J. Hook, and P. Schneider (2011), Optimized split-window coefficients for deriving surface temperatures from inland water bodies, *Remote Sensing of Environment*, 115(12), 3758–3769, doi:10.1016/j.rse.2011.09.014.
- Hüsler, F., F. Fontana, C. Neuhaus, M. Riffler, J. Musial, and S. Wunderle (2011), Avhrr archive and processing Facility at the university of Bern: A comprehensive 1-km satellite data set for climate change studies, *EARSeL eProceedings*, 10(2), 83–101.
- Inglada, J., and E. Christophe (2009), The orfeo toolbox remote sensing image processing software., in *IGARSS (4)*, pp. 733–736.
- Jakubauskas, M., D. Legates, and J. Kastens (2001), Harmonic analysis of time-series AVHRR NDVI data, *Photogrammetric Engineering and Remote Sensing*, 4(67), 461–470.

- Jassby, A. D., and J. E. Cloern (2016), *wq: Exploring water quality monitoring data*, r package version 0.4.7.
- Jimenez-Munoz, J.-C., and J. Sobrino (2008), Split-Window Coefficients for Land Surface Temperature Retrieval From Low-Resolution Thermal Infrared Sensors, *IEEE Geoscience and Remote Sensing Letters*, 5(4), 806–809, doi:10.1109/LGRS.2008.2001636.
- Jin, M., and R. E. Treadon (2003), Correcting the orbit drift effect on AVHRR land surface skin temperature measurements, *International Journal of Remote Sensing*, 24(22), 4543–4558, doi:10.1080/0143116031000095943.
- Jöhnk, K. D., J. Huisman, J. Sharples, B. Sommeijer, P. M. Visser, and J. M. Stroom (2008), Summer heatwaves promote blooms of harmful cyanobacteria, *Global Change Biology*, 14(3), 495–512, doi:10.1111/j.1365-2486.2007.01510.x.
- Karlsson, K.-G., and E. Johansson (2014), Multi-Sensor Calibration Studies of AVHRR-Heritage Channel Radiances Using the Simultaneous Nadir Observation Approach, *Remote Sensing*, 6(3), 1845–1862, doi:10.3390/rs6031845.
- Katsev, S., A. A. Aaberg, S. A. Crowe, and R. E. Hecky (2014), Recent Warming of Lake Kivu, *PLOS ONE*, 9(10), e109,084, doi:10.1371/journal.pone.0109084.
- Kaufmann, R., L. Zhou, Y. Knyazikhin, V. Shabanov, R. Myneni, and C. Tucker (2000), Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data, *IEEE Transactions on Geoscience and Remote Sensing*, 38(6), 2584–2597, doi:10.1109/36.885205.
- Kerr, J. T., and M. Ostrovsky (2003), From space to species: ecological applications for remote sensing, *Trends in Ecology & Evolution*, 18(6), 299–305, doi:10.1016/S0169-5347(03)00071-5.
- Khlopenkov, K., and A. Trishchenko (2008), Implementation and Evaluation of Concurrent Gradient Search Method for Reprojection of MODIS Level 1b Imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 46(7), 2016–2027, doi:10.1109/TGRS.2008.916633.
- Khlopenkov, K., A. Trishchenko, and Y. Luo (2010), Achieving Subpixel Georeferencing Accuracy in the Canadian AVHRR Processing System, *IEEE Transactions on Geoscience and Remote Sensing*, 48(4), 2150–2161, doi:10.1109/TGRS.2009.2034974.
- Khlopenkov, K. V., and A. P. Trishchenko (2007), SPARC: New Cloud, Snow, and Cloud Shadow Detection Scheme for Historical 1-km AVHRR Data over Canada, *Journal of Atmospheric and Oceanic Technology*, 24(3), 322–343, doi:10.1175/JTECH1987.1.

- Kidwell, K. B. (1998), NOAA POD user's guide, <http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/podug/index.htm>, accessed: 2016-01-22.
- Kilpatrick, K. A., G. P. Podestá, and R. Evans (2001), Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database, *Journal of Geophysical Research: Oceans*, 106(C5), 9179–9197, doi:10.1029/1999JC000065.
- Kirill Ya, K., and n. Filatov (1999), *Limnology and Remote Sensing - A Contemporary Approach*, Springer-Praxis Series in Remote Sensing.
- Kraemer, B. M., S. Hook, T. Huttula, P. Kotilainen, C. M. O'Reilly, A. Peltonen, P.-D. Plisnier, J. Sarvala, R. Tamatamah, Y. Vadeboncoeur, B. Wehrli, and P. B. McIntyre (2015), Century-Long Warming Trends in the Upper Water Column of Lake Tanganyika, *PLOS ONE*, 10(7), e0132490, doi:10.1371/journal.pone.0132490.
- Krasnopolsky, V. M., and L. C. Breaker (1994), The problem of AVHRR image navigation revisited, *International Journal of Remote Sensing*, 15(5), 979–1008, doi:10.1080/01431169408954129.
- Kuenzer, C., and S. Dech (Eds.) (2013), *Theoretical Background of Thermal Infrared Remote Sensing*, no. 17 in *Remote Sensing and Digital Image Processing*, 1–26 pp., Springer Netherlands.
- Latifovic, R., A. P. Trishchenko, J. Chen, W. B. Park, K. V. Khlopenkov, R. Fernandes, D. Pouliot, C. Ungureanu, Y. Luo, S. Wang, A. Davidson, and J. Cihlar (2005), Generating historical AVHRR 1 km baseline satellite data records over Canada suitable for climate change studies, *Canadian Journal of Remote Sensing*, 31(5), 324–346, doi:10.5589/m05-024.
- Layden, A., C. Merchant, and S. MacCallum (2015), Global climatology of surface water temperatures of large lakes by remote sensing, *International Journal of Climatology*, 35(15), 4464–4479, doi:10.1002/joc.4299.
- Leoni, B., C. L. Marti, J. Imberger, and L. Garibaldi (2014a), Summer spatial variations in phytoplankton composition and biomass in surface waters of a warm-temperate, deep, oligo-holomictic lake: Lake Iseo, Italy, *Inland Waters*, 4(3), 303–310.
- Leoni, B., L. Garibaldi, and R. Gulati (2014b), How does interannual trophic variability caused by vertical water mixing affect reproduction and population density of the *Daphnia longispina* group in Lake Iseo, a deep stratified lake in Italy?, *Inland Waters*, 4(2), 193–203.

- Lepori, F., and J. J. Roberts (2015), Past and future warming of a deep European lake (Lake Lugano): What are the climatic drivers?, *Journal of Great Lakes Research*, 41(4), 973–981, doi:10.1016/j.jglr.2015.08.004.
- Li, X., W. Pichel, P. Clemente-Colón, V. Krasnopolsky, and J. Sapper (2001), Validation of coastal sea and lake surface temperature measurements derived from NOAA/AVHRR data, *International Journal of Remote Sensing*, 22(7), 1285–1303, doi:10.1080/01431160151144350.
- Li, Z.-L., B.-H. Tang, H. Wu, H. Ren, G. Yan, Z. Wan, I. F. Trigo, and J. A. Sobrino (2013), Satellite-derived land surface temperature: Current status and perspectives, *Remote Sensing of Environment*, 131, 14–37, doi:10.1016/j.rse.2012.12.008.
- Lillesand, T., R. W. Kiefer, and J. Chipman (2014), *Remote Sensing and Image Interpretation*, John Wiley & Sons.
- Livingstone, D. M. (2003), Impact of Secular Climate Change on the Thermal Structure of a Large Temperate Central European Lake, *Climatic Change*, 57(1-2), 205–225, doi:10.1023/A:1022119503144.
- Livingstone, D. M., and M. T. Dokulil (2001), Eighty years of spatially coherent Austrian lake surface temperatures and their relationship to regional air temperature and the North Atlantic Oscillation, *Limnology and Oceanography*, 46(5), 1220–1227, doi:10.4319/lo.2001.46.5.1220.
- Livingstone, D. M., A. F. Lotter, and H. Kettle (2005), Altitude-dependent differences in the primary physical response of mountain lakes to climatic forcing, *Limnology and Oceanography*, 50(4), 1313–1325, doi:10.4319/lo.2005.50.4.1313.
- Lowe, D. (1999), Object recognition from local scale-invariant features, in *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999*, vol. 2, pp. 1150–1157 vol.2, doi:10.1109/ICCV.1999.790410.
- Lowe, D. G. (2004), Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 60(2), 91–110, doi:10.1023/B:VISI.0000029664.99615.94.
- Ludovisi, A., and E. Gaino (2010), Meteorological and water quality changes in Lake Trasimeno (Umbria, Italy) during the last fifty years, *Journal of Limnology*, 69(1), 174–188.
- Ludovisi, A., E. Gaino, M. Bellezza, and S. Casadei (2013), Impact of climate change on the hydrology of shallow Lake Trasimeno (Umbria, Italy): History, forecasting and management, *Aquatic Ecosystem Health & Management*, 16(2), 190–197, doi:10.1080/14634988.2013.789776.

- MacCallum, S. N., and C. J. Merchant (2012), Surface water temperature observations of large lakes by optimal estimation, *Canadian Journal of Remote Sensing*, 38(01), 25–45, doi:10.5589/m12-010.
- Marti, C. L., J. Imberger, L. Garibaldi, and B. Leoni (2016), Using time scales to characterize phytoplankton assemblages in a deep subalpine lake during the thermal stratification period: Lake Iseo, Italy, *Water Resources Research*, 52(3), 1762–1780, doi:10.1002/2015WR017555.
- Maul, G. A., and M. Sidran (1971), Estimation of sea surface temperature from space, *Remote Sensing of Environment*, 2, 165–169, doi:10.1016/0034-4257(71)90089-7.
- McMillin, L. M. (1975), Estimation of sea surface temperatures from two infrared window measurements with different absorption, *Journal of Geophysical Research*, 80(36), 5113–5117, doi:10.1029/JC080i036p05113.
- Merchant, C. J. (2013), Thermal Remote Sensing of Sea Surface Temperature, (17), 287–313.
- Metz, M. (2015), r.hants - Addon to perform HANTS algorithm on raster time series data, <https://grass.osgeo.org/grass70/manuals/addons/r.hants.html>, accessed: 2016-01-22.
- Metz, M., D. Rocchini, and M. Neteler (2014), Surface Temperatures at the Continental Scale: Tracking Changes with Remote Sensing at Unprecedented Detail, *Remote Sensing*, 6(5), 3822–3840, doi:10.3390/rs6053822.
- Micklin, P. (2007), The Aral Sea Disaster, *Annual Review of Earth and Planetary Sciences*, 35(1), 47–72, doi:10.1146/annurev.earth.35.031306.140120.
- Micklin, P. P. (1988), Desiccation of the Aral Sea: A Water Management Disaster in the Soviet Union, *Science*, 241(4870), 1170–1176, doi:10.1126/science.241.4870.1170.
- Minnett, P. J., M. Smith, and B. Ward (2011), Measurements of the oceanic thermal skin effect, *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(6), 861–868, doi:10.1016/j.dsr2.2010.10.024.
- Morabito, G., A. Oggioni, and M. Austoni (2012), Resource ratio and human impact: how diatom assemblages in Lake Maggiore responded to oligotrophication and climatic variability, *Hydrobiologia*, 698(1), 47–60, doi:10.1007/s10750-012-1094-0.
- Moreno, J., and J. Melia (1993), A method for accurate geometric correction of NOAA AVHRR HRPT data, *IEEE Transactions on Geoscience and Remote Sensing*, 31(1), 204–226, doi:10.1109/36.210461.

- Mosello, R., A. Calderoni, and R. De Bernardi (1997), Le indagini sulla evoluzione dei laghi profondi sudalpini svolte dal c.n.r. istituto italiano di idrobiologia, *Idrobiol.*, 61, 19–32.
- Neteler, M. (2010), Estimating Daily Land Surface Temperatures in Mountainous Environments by Reconstructed MODIS LST Data, *Remote Sensing*, 2(1), 333–351, doi:10.3390/rs1020333.
- Neteler, M., and H. Mitasova (2008), *Open Source GIS: A GRASS GIS Approach*, third ed., Springer, New York, doi:10.1007/978-0-387-68574-8.
- Neteler, M., M. H. Bowman, M. Landa, and M. Metz (2012), GRASS GIS: A multi-purpose open source GIS, *Environmental Modelling & Software*, 31, 124–130, doi:10.1016/j.envsoft.2011.11.014.
- NOAA (2004), *Climate Data Records from Environmental Satellites: Interim Report*, National Academies Press, Washington, D.C.
- Oesch, D., J.-M. Jaquet, R. Klaus, and P. Schenker (2008), Multi scale thermal pattern monitoring of a large lake (Lake Geneva) using a multi sensor approach, *International Journal of Remote Sensing*, 29(20), 5785–5808, doi:10.1080/01431160802132786.
- Oesch, D. C., J.-M. Jaquet, A. Hauser, and S. Wunderle (2005), Lake surface water temperature retrieval using advanced very high resolution radiometer and Moderate Resolution Imaging Spectroradiometer data: Validation and feasibility study, *Journal of Geophysical Research: Oceans*, 110(C12), doi:10.1029/2004JC002857.
- O'Reilly, C. M., S. Sharma, D. K. Gray, S. E. Hampton, J. S. Read, R. J. Rowley, P. Schneider, J. D. Lenters, P. B. McIntyre, B. M. Kraemer, G. A. Weyhenmeyer, D. Straile, B. Dong, R. Adrian, M. G. Allan, O. Anneville, L. Arvola, J. Austin, J. L. Bailey, J. S. Baron, J. D. Brookes, E. de Eyto, M. T. Dokulil, D. P. Hamilton, K. Havens, A. L. Hetherington, S. N. Higgins, S. Hook, L. R. Izmet'seva, K. D. Joehnk, K. Kangur, P. Kasprzak, M. Kumagai, E. Kuusisto, G. Leshkevich, D. M. Livingstone, S. MacIntyre, L. May, J. M. Melack, D. C. Mueller-Navarra, M. Naumenko, P. Noges, T. Noges, R. P. North, P.-D. Plisnier, A. Rigosi, A. Rimmer, M. Rogora, L. G. Rudstam, J. A. Rusak, N. Salmaso, N. R. Samal, D. E. Schindler, S. G. Schladow, M. Schmid, S. R. Schmidt, E. Silow, M. E. Soyly, K. Teubner, P. Verburg, A. Voutilainen, A. Watkinson, C. E. Williamson, and G. Zhang (2015), Rapid and highly variable warming of lake surface waters around the globe, *Geophysical Research Letters*, p. 2015GL066235, doi:10.1002/2015GL066235.
- Paerl, H. W., and J. Huisman (2008), Blooms Like It Hot, *Science*, 320(5872), 57–58, doi:10.1126/science.1155398.

- Paerl, H. W., N. S. Hall, and E. S. Calandrino (2011), Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change, *Science of The Total Environment*, 409(10), 1739–1745, doi:10.1016/j.scitotenv.2011.02.001.
- Pareeth, S., L. Delucchi, M. Metz, D. Rocchini, A. Devasthale, M. Raspaud, R. Adrian, N. Salmaso, and M. Neteler (2016), New Automated Method to Develop Geometrically Corrected Time Series of Brightness Temperatures from Historical AVHRR LAC Data, *Remote Sensing*, 8(3), 169, doi:10.3390/rs8030169.
- Pareeth, S., N. Salmaso, R. Adrian, and M. Neteler (Submitted), Homogenized daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-alpine lake, *Scientific Reports*.
- Politi, E., M. E. J. Cutler, and J. S. Rowan (2012), Using the NOAA Advanced Very High Resolution Radiometer to characterise temporal and spatial trends in water temperature of large European lakes, *Remote Sensing of Environment*, 126, 1–11, doi:10.1016/j.rse.2012.08.004.
- Price, J. C. (1991), Timing of NOAA afternoon passes, *International Journal of Remote Sensing*, 12(1), 193–198, doi:10.1080/01431169108929644.
- Privette, J., C. Fowler, G. Wick, D. Baldwin, and W. Emery (1995), Effects of orbital drift on advanced very high resolution radiometer products: Normalized difference vegetation index and sea surface temperature, *Remote Sensing of Environment*, 53(3), 164–171, doi:10.1016/0034-4257(95)00083-D.
- Quattrochi, D. A., and J. C. Luvall (2004), *Thermal Remote Sensing in Land Surface Processing*, CRC Press.
- Quayle, W. G., L. S. Peck, H. Peat, J. C. Ellis-Evans, and P. R. Harrigan (2002), Extreme Responses to Climate Change in Antarctic Lakes, *Science*, 295(5555), 645–645, doi:10.1126/science.1064074.
- R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
- Reichman, O. J., M. B. Jones, and M. P. Schildhauer (2011), Challenges and Opportunities of Open Data in Ecology, *Science*, 331(6018), 703–705, doi:10.1126/science.1197962.
- Reinart, A., and M. Reinhold (2008), Mapping surface temperature in large lakes with MODIS data, *Remote Sensing of Environment*, 112(2), 603–611, doi:10.1016/j.rse.2007.05.015.
- Riffler, M., G. Lieberherr, and S. Wunderle (2015), Lake surface water temperatures of European Alpine lakes (1989–2013) based on the Advanced Very High Resolution Radiometer (AVHRR) 1 km data set, *Earth System Science Data*, 7(1), 1–17, doi:10.5194/essd-7-1-2015.

- Robel, J. (2009), NOAA KLM user's guide, <http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/klm/index.htm>, accessed: 2016-01-22.
- Rocchini, D., and A. Di Rita (2005), Relief effects on aerial photos geometric correction, *Applied Geography*, 25(2), 159–168, doi:10.1016/j.apgeog.2005.03.002.
- Rocchini, D., and M. Neteler (2012), Let the four freedoms paradigm apply to ecology, *Trends in Ecology & Evolution*, 27(6), 310–311, doi:10.1016/j.tree.2012.03.009.
- Roerink, G. J., M. Menenti, and W. Verhoef (2000), Reconstructing cloudfree NDVI composites using Fourier analysis of time series, *International Journal of Remote Sensing*, 21(9), 1911–1917, doi:10.1080/014311600209814.
- Rosborough, G., D. Baldwin, and W. Emery (1994), Precise AVHRR image navigation, *IEEE Transactions on Geoscience and Remote Sensing*, 32(3), 644–657, doi:10.1109/36.297982.
- Salmaso, N. (2012), Influence of atmospheric modes of variability on a deep lake south of the Alps, *Climate Research*, 51(2), 125–133, doi:10.3354/cr01063.
- Salmaso, N., and L. Cerasino (2012), Long-term trends and fine year-to-year tuning of phytoplankton in large lakes are ruled by eutrophication and atmospheric modes of variability, *Hydrobiologia*, 698(1), 17–28, doi:10.1007/s10750-012-1068-2.
- Salmaso, N., and R. Mosello (2010), Limnological research in the deep southern subalpine lakes: synthesis, directions and perspectives, *Advances in Oceanography and Limnology*, 1(1), 29–66, doi:10.1080/19475721003735773.
- Salmaso, N., R. Mosello, L. Garibaldi, F. Decet, M. C. Brizzio, and P. Cordella (2003), Vertical mixing as a determinant of trophic status in deep lakes: a case study from two lakes south of the Alps (Lake Garda and Lake Iseo), *Journal of Limnology*, 62(1s), 33–41.
- Salmaso, N., F. Buzzi, L. Cerasino, L. Garibaldi, B. Leoni, G. Morabito, M. Rogora, and M. Simona (2014), Influence of atmospheric modes of variability on the limnological characteristics of large lakes south of the Alps: a new emerging paradigm, *Hydrobiologia*, pp. 1–18, doi:10.1007/s10750-013-1659-6.
- Schneider, P., and S. J. Hook (2010), Space observations of inland water bodies show rapid surface warming since 1985, *Geophysical Research Letters*, 37(22), doi:10.1029/2010GL045059.
- Schneider, P., S. J. Hook, R. G. Radocinski, G. K. Corlett, G. C. Hulley, S. G. Schladow, and T. E. Steissberg (2009), Satellite observations indicate rapid warming trend for lakes in California and Nevada, *Geophysical Research Letters*, 36(22), doi:10.1029/2009GL040846.

- Sen, P. K. (1968), Estimates of the Regression Coefficient Based on Kendall's Tau, *Journal of the American Statistical Association*, 63(324), 1379–1389, doi:10.1080/01621459.1968.10480934.
- Sharma, S., D. K. Gray, J. S. Read, C. M. O'Reilly, P. Schneider, A. Quadrat, C. Gries, S. Stefanoff, S. E. Hampton, S. Hook, J. D. Lenters, D. M. Livingstone, P. B. McIntyre, R. Adrian, M. G. Allan, O. Anneville, L. Arvola, J. Austin, J. Bailey, J. S. Baron, J. Brookes, Y. Chen, R. Daly, M. Dokulil, B. Dong, K. Ewing, E. de Eyto, D. Hamilton, K. Havens, S. Haydon, H. Hetzenauer, J. Heneberry, A. L. Hetherington, S. N. Higgins, E. Hixson, L. R. Izmest'eva, B. M. Jones, K. Kangur, P. Kasprzak, O. Köster, B. M. Kraemer, M. Kumagai, E. Kuusisto, G. Leshkevich, L. May, S. MacIntyre, D. Müller-Navarra, M. Naumenko, P. Noges, T. Noges, P. Niederhauser, R. P. North, A. M. Paterson, P.-D. Plisnier, A. Rigosi, A. Rimmer, M. Rogora, L. Rudstam, J. A. Rusak, N. Salmaso, N. R. Samal, D. E. Schindler, G. Schladow, S. R. Schmidt, T. Schultz, E. A. Silow, D. Straile, K. Teubner, P. Verburg, A. Voutilainen, A. Watkinson, G. A. Weyhenmeyer, C. E. Williamson, and K. H. Woo (2015), A global database of lake surface temperatures collected by in situ and satellite methods from 1985–2009, *Scientific Data*, 2, 150,008, doi:10.1038/sdata.2015.8.
- Sobrino, J., C. Coll, and V. Caselles (1991), Atmospheric correction for land surface temperature using NOAA-11 AVHRR channels 4 and 5, *Remote Sensing of Environment*, 38(1), 19–34, doi:10.1016/0034-4257(91)90069-I.
- Sobrino, J. A., V. Caselles, and C. Coll (1993), Theoretical split-window algorithms for determining the actual surface temperature, *Il Nuovo Cimento C*, 16(3), 219–236, doi:10.1007/BF02524225.
- Trishchenko, A. P. (2002), Removing Unwanted Fluctuations in the AVHRR Thermal Calibration Data Using Robust Techniques, *Journal of Atmospheric and Oceanic Technology*, 19(12), 1939–1954, doi:10.1175/1520-0426(2002)019<1939:RUFITA>2.0.CO;2.
- Trishchenko, A. P., G. Fedosejevs, Z. Li, and J. Cihlar (2002), Trends and uncertainties in thermal calibration of AVHRR radiometers onboard NOAA-9 to NOAA-16, *Journal of Geophysical Research: Atmospheres*, 107(D24), 4778, doi:10.1029/2002JD002353.
- Verbesselt, J., R. Hyndman, G. Newnham, and D. Culvenor (2010), Detecting trend and seasonal changes in satellite image time series, *Remote Sensing of Environment*, 114(1), 106–115, doi:10.1016/j.rse.2009.08.014.
- Verburg, P., R. E. Hecky, and H. Kling (2003), Ecological Consequences of a Century of Warming in Lake Tanganyika, *Science*, 301(5632), 505–507, doi:10.1126/science.1084846.

- Walton, C. C. (1988), Nonlinear Multichannel Algorithms for Estimating Sea Surface Temperature with AVHRR Satellite Data, *Journal of Applied Meteorology*, 27(2), 115–124, doi:10.1175/1520-0450(1988)027<0115:NMAFES>2.0.CO;2.
- Wetzel, R. G. (2003), *Limnology: Lake and River Ecosystems. Third Edition.* By Robert G Wetzel., vol. 78, 368-369 pp., doi:10.1086/380040.
- Weyhenmeyer, G. A. (2001), Warmer Winters: Are Planktonic Algal Populations in Sweden's Largest Lakes Affected?, *AMBIO: A Journal of the Human Environment*, 30(8), 565–571, doi:10.1579/0044-7447-30.8.565.
- Williamson, C. E., J. E. Saros, and D. W. Schindler (2009), Sentinels of Change, *Science*, 323(5916), 887–888, doi:10.1126/science.1169443.
- Wilson, R. C., S. J. Hook, P. Schneider, and S. G. Schladow (2013), Skin and bulk temperature difference at Lake Tahoe: A case study on lake skin effect, *Journal of Geophysical Research: Atmospheres*, 118(18), 10,332–10,346, doi:10.1002/jgrd.50786.
- Winder, M., J. E. Reuter, and S. G. Schladow (2009), Lake warming favours small-sized planktonic diatom species, *Proceedings of the Royal Society B: Biological Sciences*, 276(1656), 427–435, doi:10.1098/rspb.2008.1200.
- Wood, S. (2006), *Generalized Additive Models: An Introduction with R*, Chapman and Hall/CRC.
- Wood, S. N. (2004), Stable and efficient multiple smoothing parameter estimation for generalized additive models, *Journal of the American Statistical Association*, 99(467), 673–686.
- Xu, Y., Y. Shen, and Z. Wu (2013), Spatial and Temporal Variations of Land Surface Temperature Over the Tibetan Plateau Based on Harmonic Analysis, *Mountain Research and Development*, 33(1), 85–94, doi:10.1659/MRD-JOURNAL-D-12-00090.1.
- Yu, L., D. Zhang, and E.-J. Holden (2008), A fast and fully automatic registration approach based on point features for multi-source remote-sensing images, *Computers & Geosciences*, 34(7), 838–848, doi:10.1016/j.cageo.2007.10.005.