POLITECNICO DI MILANO

Water Level Data in Finland Rivers and Lakes via Satellite Altimeters

Tesi di Laurea Magistrale in
Ingegneria per l'Ambiente e il Territorio/ Environmental and Land Planning Engineering

Author: **Anthony Mascia**

Student ID: 928531
Advisor: Prof. Carlo De Michele
Co-advisor: Cristina Deidda
Academic Year: 2021-22

Abstract

The water level monitoring of rivers and lakes is a fundamental issue in order to assess the health status of water bodies and to know the availability of freshwater resources. The presence of in-situ water level stations is quite rare (for example in mountains areas), and when available it may have discontinuous functioning, or lack of maintenance, making this kind of information difficult to be used in hydrological modelling. In the water level monitoring problem, recently, remote sensing data collected through satellite altimeters have showed great capability, not only in the case of rivers with wide sections (>100 m), but also in the case of rivers with narrow sections ($<$ 100 m). Here, it has been assessed the feasibility of satellite altimetric measurements (Jason, Sentinel 3A, Envisat) in rivers and lakes, in Finland, making a comparison to in-situ data. In particular it has been used three river virtual stations (only Sentinel 3A data), and four lake virtual stations (mix of satellites). It has been used four different methods of comparison between in-situ and remote water level measurements: (1) comparing the differences between successive times of satellite and in-situ measurements, (2) comparing the satellite and in-situ measurements assuming that the initial value of the joint measurements is equal (specifically the value of satellite water level is equal to the value of in-situ one), (3) comparing the satellite and in-situ measurements calibrating the initial water level of satellite measurements minimizing the MAE, (4) comparing the satellite and in-situ measurements calibrating the initial water level of satellite measurements minimizing the RMSE.

Keywords: water level; Finland; rivers; lakes; altimetry; satellite water level, in-situ water level; comparison; accuracy

Abstract in lingua italiana

Il monitoraggio dei livelli d'acqua dei fiumi e dei laghi è un tema fondamentale per valutare lo stato di salute dei corpi d'acqua e per conoscere la disponibilità delle risorse idriche. Spesso le stazioni di misure di livello in-situ o non sono presenti (per esempio nelle zone di montagna) o sono in disuso, o possono avere un funzionamento discontinuo dato dalla mancanza di manutenzione, rendendo questo tipo di informazioni difficili da usare nei modelli idrologici. I dati satellitari possono essere una valida alternativa agli strumenti tradizionali di monitoraggio del livello idrico. Recenti applicazioni hanno mostrato grandi potenzialità dei dati satellitari nel monitoraggio dei livelli idrici nel caso dei fiumi con larghe sezioni (>100m), ma anche nel caso di fiumi con sezioni strette(<100m). In questa tesi, è stata valutata l'affidabilità delle misure altimetriche satellitari (Jason, Sentinel 3A, Envisat) dei fiumi e dei laghi, in Finlandia, facendo un confronto con i dati in-situ. In particolare, sono state utilizzate tre stazioni virtuali fluviali (solamente i dati di Sentinel 3A), e quattro stazioni virtuali lacustri (un mix di satelliti). Sono stati utilizzati quattro differenti metodi di confronto tra le misure dei livelli idrici satellitari e quelle in-situ: (1) confrontando le differenze tra due successivi istanti temporali delle misure satellitari e in-situ, (2) confrontando le misure satellitari e in-situ assumendo che il valore iniziale delle due serie temporali sia uguale (in particolare il valore del livello d'acqua satellitare è uguale al primo valore della serie in-situ) (3) confrontando le misure satellitari e in-situ calibrando il valore iniziale del livello d'acqua dei dati satellitari minimizzando l'indice MAE, (4) confrontando le misure satellitari e in-situ calibrando il valore iniziale del livello d'acqua dei dati satellitari minimizzando l'RMSE.

Parole chiave: livello d'acqua; Finlandia; fiumi; laghi; altimetria; livello d'acqua satellitare; livello d'acqua in-situ; confronto; accuratezza

Contents

Introduction

Rivers and lakes freshwater are a fundamental resource for humans and for natural ecosystems [2]. They are the habitat of many aquatic and terrestrial species, but they are also used for energy production, civil, touristic, agricultural and industrial purposes [13, 35].

Due to the climate changes, there is an increasing trend over time in the number of extreme events like droughts and floods. They cause huge environmental and economic impacts, and loss of human life in all over the world [28]. Changes in the water levels of rivers and lakes, both increases and decreases, can have economic and ecological repercussions. This can particularly affects areas in whose economies are mainly based on tourism, agriculture, livestock on fishing or hydropower production [5, 7, 35, 51]. It is also important to underline the social impacts of these events in the less developed countries as Africa and Asia. Droughts lead to problems of food and water scarcity and floods generate high damages to houses and infrastructures [48].

Over time, in the world there has been an increasing difficulty to find freshwater because of the changing of its spatial and temporal distribution. Droughts are always more frequent, involving an increasing percentage of areas in the world and occasional water shortages are becoming intermittent or even permanent [13, 35].

Climate change has altered, and will continue to alter, the rhythms of the accumulation and melting of snow [4, 35, 48]. This leads to strong seasonal changes in water levels and the flow of water bodies. It is compounded by factors such as population growth and the economy impacting on the increase in per capita consumption. Consequently, the supply of fresh water appears to be a global problem of increasing importance [4, 35, 48].

The report "Climate Change Impacts on Water Resources and Lake Regulation in the Vuoksi Watershed in Finland" [66], published in 2010, showed that levels and flow rates in the south-eastern Finnish area of Vuoksi will change. In comparison to the past, they will decrease in late spring and summer and will increase during late autumn and winter. This means that modification of the current level dam adjustment calendars will be needed.[66].

Accurate monitoring of surface water level is increasingly required. Many studies have shown the urgent need to develop an efficient and effective water body monitoring system [35]. This is both to enable proper water resource management for the purposes mentioned above, but also for the proper functioning of risk reduction and mitigation systems such as flood alerts and extreme event forecasting systems. [13, 24]. This is due to the rapid changes in the water balance of the lakes, the change in flooding behaviors during the spring and summer periods, and ice cover in the winter [55]. For example the knowledge of the flow discharge of a river section useful to assess the the output of a catchment [25].

Many of the world's rivers and lakes are not properly monitored in-situ, although the need and importance of doing so is evident. Italy is an example [23]. In the past, in-situ manual station measurements have often been temporally and spatially scarce, discontinuous, not homogeneous and prone to human error. This has led in many cases to a lack of precise information about in the way and time of measurement, and the location of the place of execution. [23, 35, 55] Today's digital stations retain the punctual characteristics of manual stations and have the important advantage of being automatic. However, they have the problem of the continuous and necessary maintenance [23]. Over time, the monitoring of water levels has become more accurate and regular [35]. But, paradoxically in recent decades is decreasing the availability of measures [11, 34]. The problems that disadvantage in-situ monitoring are the high installation and maintenance costs of measuring stations [36], the difficulty of reaching remote areas and political issues in the interest areas [50, 58].

Satellite measurements may be a valid alternative to in-situ measurements where instruments are rare, missing, discontinuous, malfunctioning, abandoned or difficult to access,[23, 35]. They are also a source of information complementary to scarce shore-based measures [3, 15]. The punctual nature of in-situ measurements has been overcome by remote sensing which allows to obtain georeferenced images at a good temporal and spatial scale. Point measurements are more localized and are used as a tool for validating what satellite sensors detect [18].

Remote sensing satellite instruments have revolutionized the study of environmental issues such as meteorology, climatology, agriculture and irrigation as well as the monitoring of air pollutants, of surface water bodies, of snow cover, of landslide [54], of vegetation [33] and of hydrological and natural hazards. So they allow the study of issues that in the past were poorly studied because of the difficulty to obtain data and information.

Altimetry is a satellite technique of estimation of water level height referred to a surface of reference [23, 35]. It is calculated measuring the time taken by a radar pulse for the two-way travel from satellite to the water surface and ,after the back-scattering, from the water surface to the satellite. It is a technique useful for hydrological aims, for example to study water cycle and more in particular for monitoring the water level and the volume variation of lakes and reservoir [30, 39]. About rivers, it is possible to obtain the discharge from altimetric data [56, 60]. It is utilized an instrument named radar altimeter which is changed and evolved during the time. [12, 23, 58].

Altimetric sensors were initially designed for open ocean water level measurements [49]. Later they were modified to be used for lakes and inland sea giving high level of confidence after a very big process of validation using in-situ data [1, 10]. Contrariwise, the assessment of river satellite data is more complex than the previous case, one of the problems is the width of the channels [19, 58].

The challenge of the 21st century is to try to understand, with satellites, climate change and its impacts as well as possible. Mitigation of these impacts could be achieved through using satellite technology and creating models capable of describing future scenarios [6, 37, 43, 63, 64, 68]

The aim of this work it to evaluate the feasibility of using satellite altimetric data for water level monitoring focusing on Finland hydrological network, rivers and lake. It will be shown the comparison of these data with in-situ data. The analysis will be done over three different time period data: the whole year, winter and summer period.

1 | Satellite data for hydrological applications

1.1. Radar altimetry

1.1.1. History of altimetric missions

Altimetry became important to retrieve satellite data of water surface height when in-situ records are missing or they are not complete [23, 35]. First altimeters were used to carry out elevation of water level surface of open oceans [49]. At a second stage, altimetry was then applied to the study of rivers' water surface.

Seasat, Geosat and ERS1 launched in 1978, 1985 and 1991 respectively were the first satellite missions that showed the benefit of radar altimetry to the scientific community [32, 35]. The purpose of these first missions was to study sea level and ocean topography [13, 15, 23]; however, data were affected by many problems related to the ocean circulation [32, 35].

The launch of TOPEX/Poseidon satellite in 1992 developed by NASA and CNES was the turning-point of radar altimetry's technique: the improvement of the orbit's determination allowed for better quality of data as a result [59].

Jason 1, Envisat, Jason 2 and Jason 3 launched in 2001, 2002, 2008 and 2016 respectively, allowed further improvement compared to Topex/Poseidon [13, 15, 42, 65]. The satellites mentioned above were all developed by NASA and CNES, except from ENVISAT which was developed by ESA. Researches understood the potentiality of the use of satellite when in-situ data are scarce or missing [3, 15]. They offered, after validation with in-situ data, very good quality data over large bodies of water,such as lakes and inland seas [13, 15].

Jason 1 and Envisat had good performance for lakes and inland sea that were largely studied [29, 53].However, the agreement on the evaluation of rivers' altimetric data was not so good, except for the largest ones [17, 22].The results obtained were linked the channels' width [23] and better performances were found on exams taken overwider rivers [17, 22]. Birkett et all(2002) [8] studied the accuracy of TOPEX/Poseidon's data of surface water height of rivers of Amazon basin with a minimum theoretical width of 1.16 km, therefore not less than 1 km. Joecila Santos et al(2010) [20] studied ERS2 and ENVISAT water levels of rivers width till to 0.05 km and they found high quality data series. They discovered that the river's reach does not strongly affect the quality of results. Therefore, it was possible to include in the study less then 1 km-wide rivers.

Only through the technological improvement of the satellites was it possible to estimate the water levels of small water bodies with a width of less than 200m [52, 57]. In 2016, Sentinel 3 A was launched. This satellite is one of the missions of the Copernicus Earth Observation Programme of The European Union with the aim to have a proper Earth observation system [27]. It operates with a nadir Satellite Radar ALtimeter called SRAL [44]. This new altimeter has better characteristics in comparison to the previous one, such as it works with lower noise and higher spatial resolution than previous missions [45]. It was found good accuracy better accuracy for open sea [45] and lakes [17, 53] than previous missions [31]. Studies about the feasibility of Sentinel 3 data of rivers are recents [38, 61] and they showed a not clear link between data accuracy and river width [9, 40, 67]. But it's probable that terrain around virtual station can be influent, for example in a relief zone [40]. Likewise Silva et al (2010) [21] discovered that the quality of the water level series was not affected by river's width, but by the land cover near the river, by the river channel's morphology (width, direction and shape) and by the land's topography [46]. The report of "Copernicus Global Land Operations", published in 2008, studied rivers with a minimum width of 300 m [58].

Deidda et al. evaluated the quality of Jason, Sentinel 3A, Sentinel 3B data of narrow rivers in Italy that are less than 300 m wide [23]. They found the feasibility to use these data.

Therefore, radar altimeter satellite were born with the aim to measure altimetry in open oceans. Then, good results with particular conditions were found also for inland seas, lakes, rivers and wetlands [58].

1.1.2. How does the altimeter works?

The direct measure of the height of a water level referred to a generic surface can not be done using tools carried on satellites. It can be calculated indirectly using altimetry. [23, 35, 58]. Radar altimeter send a radar signal to the Earth surface. The electromagnetical wave it is reflected from the Earth and it comes back to the source that is also a receiver. It is measured the time of this signal travel. From that measure it is possible to calculate the distance between the satellite and the water surface, the satellite altitude. The distance just mentioned is shown in the Figure 1.1 with the name "range". It is proportional to the time of the wave travel [58].

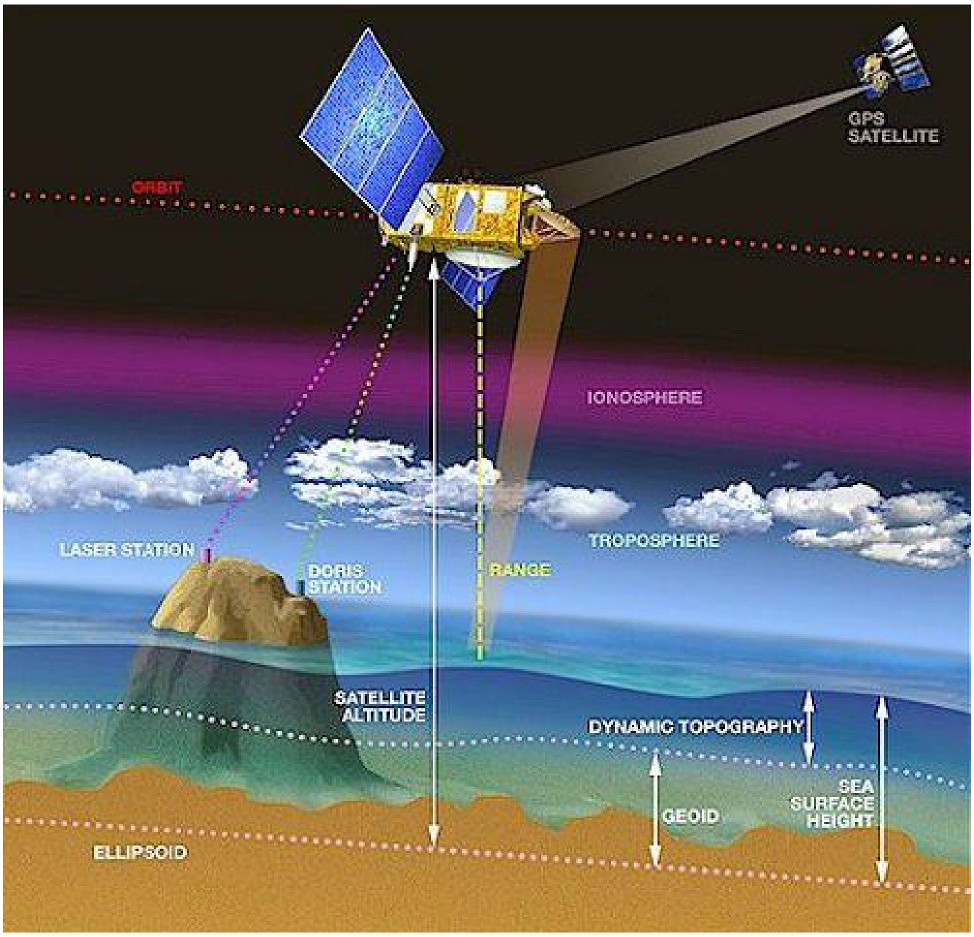


Figure 1.1: Satellite altimetry operating mechanism (Credit CNES)

It is necessary to know the distance between satellite and a surface reference that usually is the ellipsoid (as shown in the Figure 1.1) but could be also the geoid. Therefore, it is crucial the knowledge of the precise satellite orbit.

There are many typology of the reflecting surface, it can be small lakes and rivers or large lakes. This influence the shape and the magnitude of the wave echoes. Jason 3, Sentinel 3 A and Sentinel B use the Ku band, which has the frequency of 13.6 GHz. Ku band has a good trade-off characteristics between the potentiality of the technology that is linked with the power emitted, the available bandwith and the sensitivity to atmospheric and ionospheric perturbations that could influence the propagation of the radar wave [58].

On board of these satellites have been installed the following instruments:

- SRAL, a SAR radar altimeter. It is a dual-frequency SAR ALtimeter;
- MWR, a Microwave Radiometer;
- POD, a Precise Orbit Determination;
- OLCI, an Ocean and Land Colour Instrument;
- SLSTR, a Sea and Land Surface Temperature Radiometer

In addition to altimetric measurements it can also give visible and infrared measurements [26].

2 | Finland case study

2.1. Characteristic of the study area

Finland is a country in northern Europe, more precisely located in the Scandinavian Peninsula 2.1.
It borders with Norway to the north, with Russia to the east and with Sweden to the north-west. It has a coastline on the Baltic Sea, which includes the Gulf of Bothnia to the west and the Gulf of Finland to the south It spans over 1,000 km in length from north to south, between 60° and 70° of latitude and between 20° e 30° of longitude. This makes it one of the northernmost countries in the world: in fact, about a third of the territory is north of the Arctic Circle.

Figure 2.1: Location of Finland in Europe

The area of 338,145 km^2 makes it one of the 8 largest countries in Europe. The resident population is around 5.5 million and its population density is the lowest in the European

Union at 18 ab/km^2. The inhabitants of Finland are not equally distributed over the territory: 50% live in the southern zone and 3%, 180,000 inhabitants, live in the Lapland region, which has an area equal to 30% of that of the entire nation as shown in the Figure 2.2.

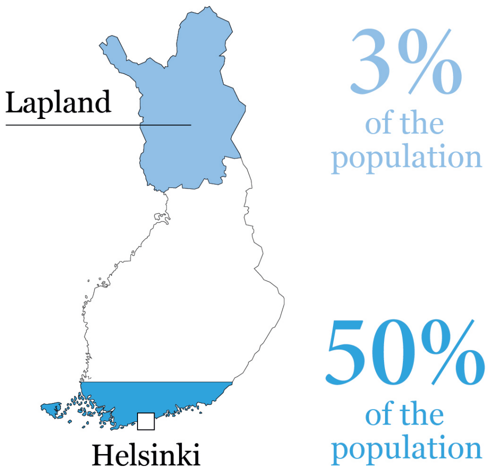


Figure 2.2: Population distribution in Finland

Finland is the country with the highest number of water bodies in the world, with almost a tenth of the surface of the country covered by water. The Figures 2.3 shows the finnish lakes and rivers.

Figure 2.3: Finnish lakes and rivers

There are about 188,000 lakes with an area of more than 500 m^2, which is a requirement to be defined as a lake by the Finnish Ministry of the Environment. Most of these lakes are of glacial origin [62]. In the south-eastern part of the country there is the Region of Lakes, whose name derives from the high number of this type of surface water bodies. The lakes with surface greater than 100 km^2 are 47 (an high number, considering that throughout the European Union there are 93), 309 lakes have surface greater than 10 km^2

and about 56,000 greater than 0.01 km^2. The combined volume of Finnish lakes is a total of 235 km^3 of water [14].

More than 80% of the lakes have good or excellent water quality, although many of them suffer from eutrophic phenomena due to their low depth and therefore require special attention. The elongated and jagged contours, in combination with the high number of islands, makes them unique in comparison to the lakes of other nations. In most cases, their particular shapes make it difficult to determine where a lake ends and where the next one begins. They are often connected to each other by numerous rivers and artificial canals also used for the transport of goods [14].

The most important lakes are:

- Lake Saimaa, with an area of 4,380 km2 [14], is the first in Finland and the fourth largest lake in Europe. It is located in the eastern part of the country, between the cities of Lappeenranta and Joensuu and hosts 13,710 islands. In fact, Lake Saimaa consists of several lakes connected to each other by narrow passages The main parts of Lake Saimaa are the lakes Suur-Saimaa, Pihlajavesi, Haukivesi, Puruvesi, Orivesi and Pyhäselkä .
- Lake Päijänne is the deepest lake and the second in size (1,100 km2)[cit A]. It is located in the southern part of the country and hosts 2,690 islands. It is important as it supplies drinking water to approximately one million people in the metropolitan area of Helsinki through a 120km long tunnel, due to its excellent chemical, physical and biological characteristics [14]. It's the major lake of the River Kymijoki watercourse.
- Lake Inarinjärvi is the third largest lake in Lapland (1,050 km^2). It is located in the northern part of Lapland and represents the largest lake in that area. I. The hydroelectric power plant constructed in the River Paatsjoki in Russia started operation in 1934 and it has regulated the water level of Lake Inari ever since. Because of the regulation, the water level of the lake is about 0.5 m higher than in its natural state.

The lakes analysed in this thesis are:

- Lake Inarinjärvi, described above;
- Lake Kemijärvi, the 19th largest lake (231 km^2) that is located in Lapland;
- The Lake Oulujärvi 5th lake by extension (930 km^2) and located in the northern central area of the country;

- Lake Vanajavesi, with an area of 150 km2, it located in the southern zone.

The river network is approximately 25,000 km long, as shown in the Figure 2.3, and consists of about 47 major rivers including

- The Kemijoki, which is the longest in the country (550 km). Its catchment area covers more than half of the territory of Lapland, which is the main river area of Finland, hosting some of the rivers with the greatest flows. This river is the main source of flooding in Finland and produces about 40% of the electricity generated throughout the state.
- The Iijoki, which is third in length (330 km). It is located in the central part of the country and flows into the Gulf of Bothnia.
- The Ounasjoki, which is the main tributary of the Kemijoki River, is therefore located in the northern area. It is 298 km long.
- The Kitinen, also a tributary of the Kemijoki and is 278 km long.

Rapids in Finland produce one sixth of the country's electricity [14].

In this paper it will be analyzed two rivers:

- The Kokemäenjoki which is the main waterway in the southwestern area and is around 121 km long;
- The Tankajoki, that is 34 km long and located in Lapland. Il Tankajoki, lungo 34 km, is located in Lapponia.

As stated above, the number of lakes and rivers is high and most of them are small. For this reason,although resorting to territorial information systems, it is often difficult to identify the flow directions, the trends of the rivers, as many of the smallest are not marked on the maps.

2.2. Satellite data

The initial aim of this thesis was to validate only Sentinel 3 altimetric data. The retrieval of the rivers data regarding this satellite was easy. It has been more difficult to find lakes data. After a long research it has been found lakes altimetric data that are a mix of satellite data. This lack could be seen also on its positive side, that is it has been possible to validate a multimission data record.

More informations about the two database mentioned just above will be illustrated in the next subsections after an explanation of what is a virtual station and its data.

2.2.1. Altimetry

Each satellite travels on a determined orbit making cycles around the Earth. Satellites are characterized by a proper revisit time which is the duration of each cycle. The projection on the Earth surface is called ground track. The intersection of the ground track with the river's center line is named virtual station.Each satellite have a specific number of virtual stations. The positions of the virtual stations are fixed for each mission. For each virtual station is possible to retrieve the referred altimetric time serie which contain one data for every cycle [58]. The duration of the orbital cycle of Sentinel 3A is 27 days[58].

The ground track mentioned above is theoretical. The real orbit change every cycles because of the dishomogeneity of the Earth's mass distribution. Therefore, it will not pass over the same intersection with the water body every cicle. Are needed small burns to correct the orbit to not change more than 1 km from the theoretical one [23].

But for Finland the number of virtual stations available is low, this because the raw data are under validation process. On this site there are not present Finland lakes data. After a meticulous research they are found on Theia website, but for some virtual stations the data are a mix of altimetric data of several satellites. Therefore the assessment agreement is not referred only to Sentinel 3. This can be also a step forward. Below are shown in more detail these things just mentioned.

2.2.2. Rivers altimetric data

Altimetric water level records of Finland rivers analyzed in this thesis were downloaded freely from the site https://land.copernicus.eu/global/products/wl (accessed date: 19 October 2022). At the moment of the download were available only 3 rivers virtual stations related to Sentinel 3A in Finland. They are shown in the Figure 2.4

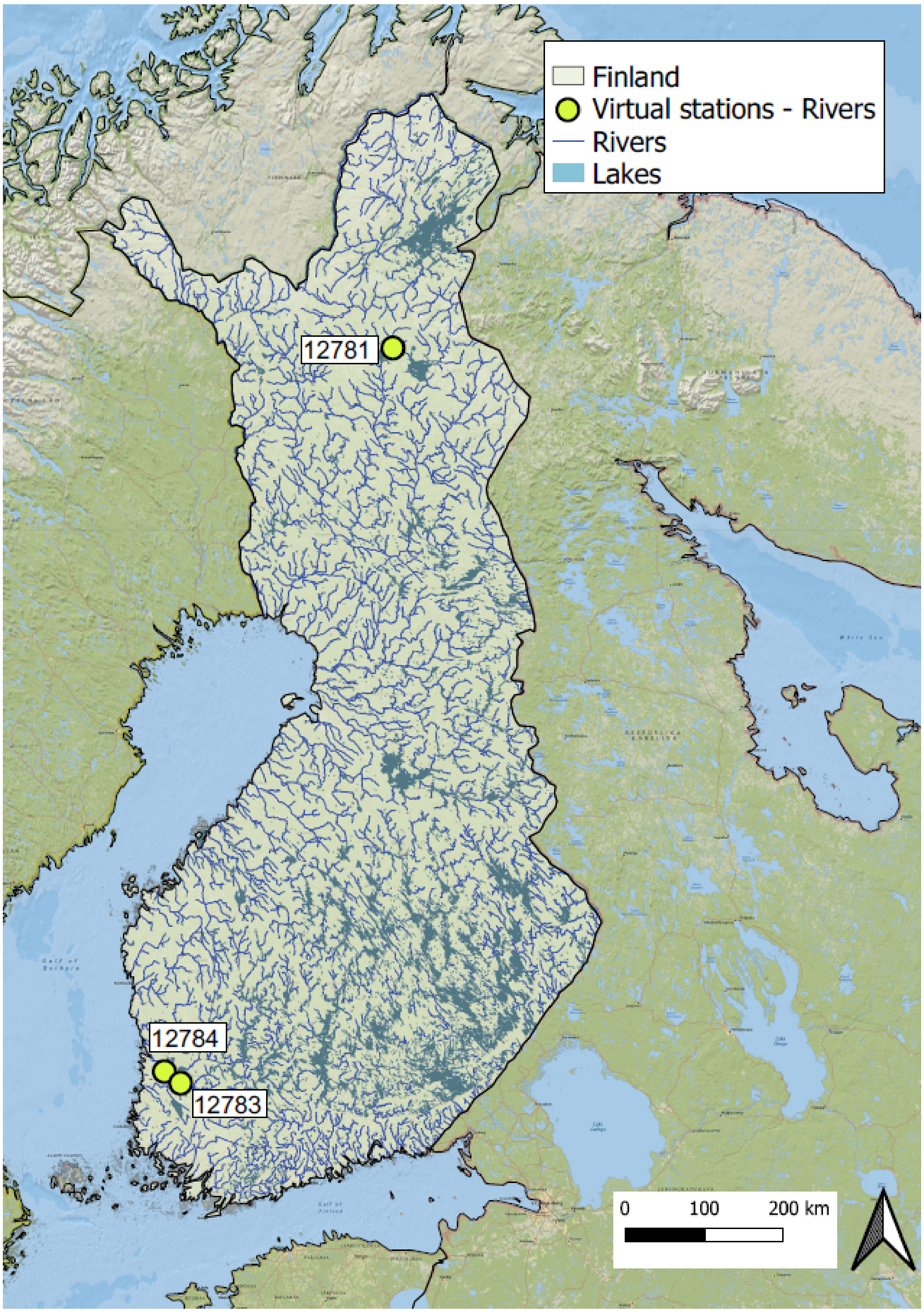


Figure 2.4: Rivers virtual stations analyzed

For each virtual station it is possible to download one file in ".json" format. It has been written an R code to open and save automatically this original files.

In the Figure 2.5 it has been illustrated the general information of the virtual stations and of the records retrieved.

Identifier	Longitude	Latitude	Country	Basin	River	Time_start	Time_end	WaterSurfRef_Alt	Satellite
12781	270.037	680.914	Finland	Kemijoki	Tankajoki	20/04/2016 18:43	12/09/2021 18:44	21.82	SENTINEL3A
12784	22.009	613.788	Finland	Kokemaenjoki	Tattaranjoki	26/04/2016 09:36	18/09/2021 09:36	18.81	SENTINEL3A
12783	223.645	612.575	Finland	Kokemaenjoki	Kokemaenjoki	26/04/2016 19:26	18/09/2021 19:26	18.81	SENTINEL3A

Figure 2.5: General informations of the rivers virtual stations analyzed

It is possible to know the following information of the virtual stations:

- Identifier. It is the name of the virtual station;
- Longitude;
- Latitude;
- Country;
- Basin;
- River;
- Time_start. It is the date and the time of the first observation downloaded
- Time_end. It is the date and the time of the last observation downloaded
- WaterSurfRef_Alt. It is the height of the geoid, that is the level reference of the observations.
- Satellite. It is the satellite platform.

Each "json" file contains also the section of the altimetric data. A few parts of the are shown in the Figures 2.6, 2.7,2.8.

```
      Identifier Longitude Latitude Time              Time-h-d           Water level
 [1,] "12781"    "27.0037" "68.0914" "2016.3026772162" "20-04-2016-18:43" "264.65"
 [2,] NA         NA        NA        "2016.376447708"  "17-05-2016-18:43" "266.23"
 [3,] NA         NA        NA        "2016.5977591834" "06-08-2016-18:43" "266.8"
 [4,] NA         NA        NA        "2016.6715296752" "02-09-2016-18:43" "266.96"
 [5,] NA         NA        NA        "2016.745300167"  "29-09-2016-18:43" "267.08"
 [6,] NA         NA        NA        "2016.8190706588" "26-10-2016-18:43" "266.93"
 [7,] NA         NA        NA        "2016.8928411506" "22-11-2016-18:43" "266.54"
 [8,] NA         NA        NA        "2017.0404927702" "15-01-2017-18:43" "265.92"
 [9,] NA         NA        NA        "2017.1144653729" "11-02-2017-18:42" "265.48"
[10,] NA         NA        NA        "2017.1884379756" "10-03-2017-18:42" "265.08"
[11,] NA         NA        NA        "2017.2624105784" "06-04-2017-18:43" "264.45"
[12,] NA         NA        NA        "2017.3363831811" "03-05-2017-18:42" "264.21"
[13,] NA         NA        NA        "2017.4103557839" "30-05-2017-18:43" "265.06"
[14,] NA         NA        NA        "2017.4843283866" "26-06-2017-18:42" "265.92"
[15,] NA         NA        NA        "2017.5583009893" "23-07-2017-18:42" "266.32"
```

Figure 2.6: Altimetric data of the virtual station 12781 from the file ".json"

	Identifier	Longitude	Latitude	Time	Time-h-d	Water level
[1,]	"12783"	"22.3645"	"61.2575"	"2016.3191522465"	"26-04-2016-19:25"	"49.44"
[2,]	NA	NA	NA	"2016.3929227383"	"23-05-2016-19:25"	"49.16"
[3,]	NA	NA	NA	"2016.4666932301"	"19-06-2016-19:25"	"48.72"
[4,]	NA	NA	NA	"2016.5404637219"	"16-07-2016-19:25"	"48.64"
[5,]	NA	NA	NA	"2016.6142342137"	"12-08-2016-19:25"	"48.82"
[6,]	NA	NA	NA	"2016.6880047055"	"08-09-2016-19:25"	"48.69"
[7,]	NA	NA	NA	"2016.7617751973"	"05-10-2016-19:25"	"48.25"
[8,]	NA	NA	NA	"2016.8355456891"	"01-11-2016-19:25"	"48.45"
[9,]	NA	NA	NA	"2016.9093161809"	"28-11-2016-19:25"	"49.04"
[10,]	NA	NA	NA	"2016.9830866727"	"25-12-2016-19:25"	"48.58"
[11,]	NA	NA	NA	"2017.0570129376"	"21-01-2017-19:26"	"48.39"
[12,]	NA	NA	NA	"2017.2049581431"	"16-03-2017-19:26"	"48.83"
[13,]	NA	NA	NA	"2017.2789307458"	"12-04-2017-19:25"	"48.72"
[14,]	NA	NA	NA	"2017.3529033486"	"09-05-2017-19:26"	"48.94"
[15,]	NA	NA	NA	"2017.4268759513"	"05-06-2017-19:26"	"48.68"

Figure 2.7: Altimetric data of the virtual station 12783 from the file ".json"

	Identifier	Longitude	Latitude	Time	Time-h-d	Water level
[1,]	"12784"	"22.009"	"61.3788"	"2016.3180327869"	"26-04-2016-09:36"	"20.9"
[2,]	NA	NA	NA	"2016.3918032787"	"23-05-2016-09:36"	"20.79"
[3,]	NA	NA	NA	"2016.4655737705"	"19-06-2016-09:36"	"19.71"
[4,]	NA	NA	NA	"2016.5393442623"	"16-07-2016-09:36"	"19.51"
[5,]	NA	NA	NA	"2016.6131147541"	"12-08-2016-09:36"	"19.97"
[6,]	NA	NA	NA	"2016.6868852459"	"08-09-2016-09:35"	"20.01"
[7,]	NA	NA	NA	"2016.7606557377"	"05-10-2016-09:35"	"20.37"
[8,]	NA	NA	NA	"2016.8344262295"	"01-11-2016-09:35"	"19.54"
[9,]	NA	NA	NA	"2016.9081967213"	"28-11-2016-09:35"	"20.38"
[10,]	NA	NA	NA	"2016.9819672131"	"25-12-2016-09:35"	"19.5"
[11,]	NA	NA	NA	"2017.055890411"	"21-01-2017-09:36"	"19.52"
[12,]	NA	NA	NA	"2017.1298630137"	"17-02-2017-09:36"	"19.44"
[13,]	NA	NA	NA	"2017.2778082192"	"12-04-2017-09:36"	"20.71"
[14,]	NA	NA	NA	"2017.3517808219"	"09-05-2017-09:35"	"20.24"
[15,]	NA	NA	NA	"2017.4257534247"	"05-06-2017-09:36"	"19.59"

Figure 2.8: Altimetric data of the virtual station 12784 from the file ".json"

In the previous Figures there are the following information fields:

- Identifier. It is the name of the virtual station;
- Longitude
- Latitude;
- Time in decimal format;
- Time-h-d. It contains the data and the time of the altimetric measure;
- Water level. In this column there are the data that will be under the validation process. the water level height above surface reference (geoid).

For the comparison with in-situ measurements it is created a new variable, which is the height of water surface above ellipsoid. It is the sum of the height of geoid above ellipsoid with the height of the water surface over the geoid. It was not considered the uncertainty reported in the "json.file".

The Figure 2.9 shows the number of the observations of water levels downloaded from the year 2016 to the access day.

Identifier	Numer of observations downloaded
12781	53
12784	69
12783	65

Figure 2.9: Number of rivers virtual stations observations retrieved

As said in the 2.2.1, it is not possible to have Sentinel 3 A measurements for every day. There is to consider that the satellites, not only Sentinel 3 A, have a specific revisit time. In this case it is 27 days.

2.2.3. Lakes altimetric data

The retrieval of finnish lakes altimetric data has been not very easy as the case of the rivers. The site used for rivers data (https://land.copernicus.eu/global/products/wl) permitted to download (accessed date: 19 October 2022) in the first time at only the the altimetric data of one virtual station, the lake Vanajanselka. It has been checked many months later if new lakes virtual stations will be available, bur nothing changed. Probably the records are not uploaded on the site because the are under process of elaborations. It has been considered unsatisfactory to consider only one lakes virtual station because of the large number of this typology of water bodies present in Finland, as said in the section 2.1.

It has been found the site Theia, https://www.theia-land.fr/ that permitted to download the data of four lakes virtual stations located in Finland [16]. The only problem was that these records did not have monomission origin and they were not label with the proper satellite that made the measurement. In fact they are a mix of different satellites measurements. This different origin of the measurements could be judge as a discouraging problem. But watching the positive side, it is a possibility to analyze a multimission record. Considering a given time duration, a multimission record surely will have a higher number of observations than a monomission record.

The virtual stations retrieved are shown in the Figure 2.10.

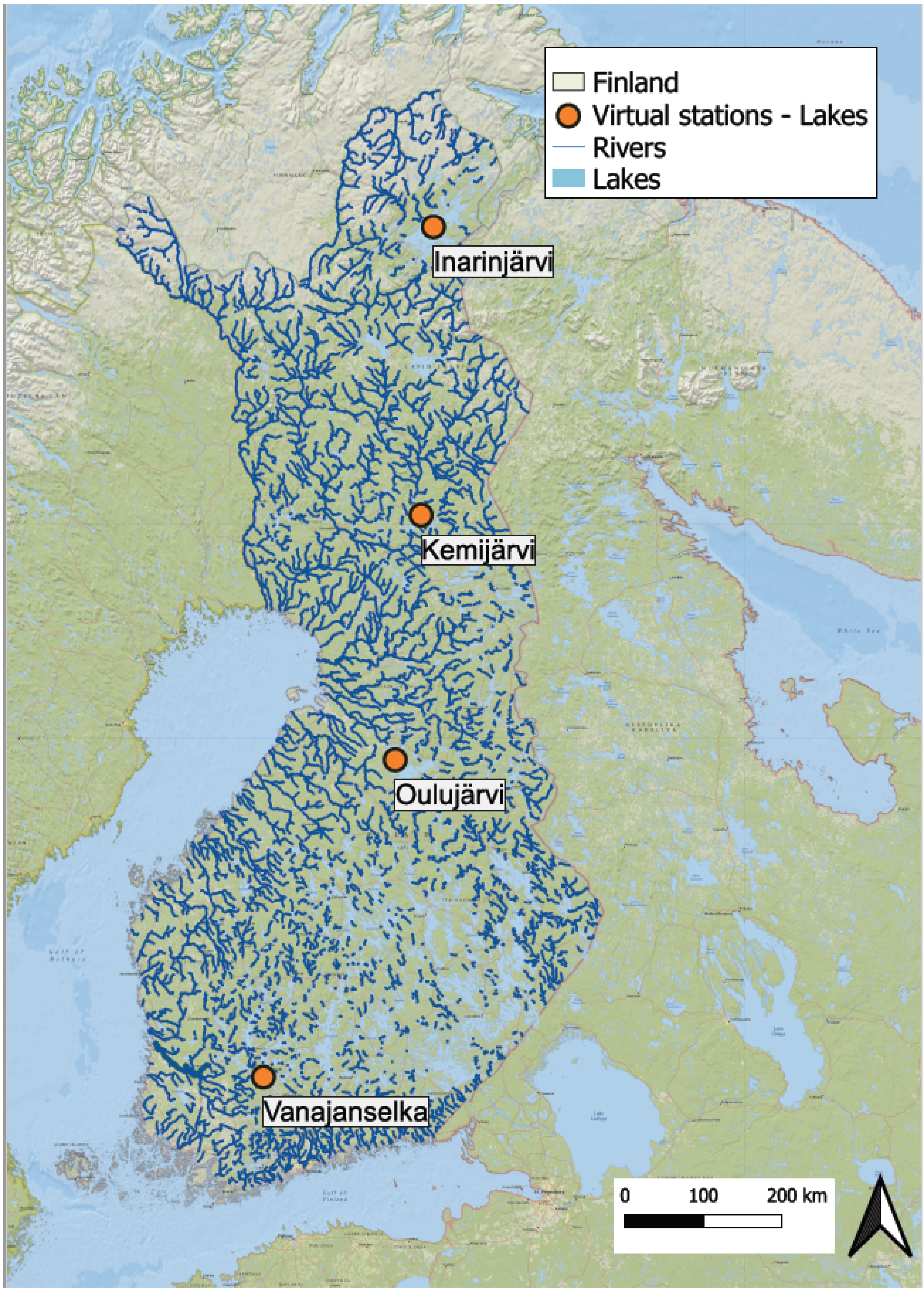


Figure 2.10: Lakes virtual stations analyzed

From the site mentioned above it has been downloaded one file ".txt" for each virtual station and one of these files is shown in the Figures 2.11 and 2.12. It is reported as example the virtual station Inarinjärvi.

*L_inarinjarvi - Blocco note di Windows

File Modifica Formato Visualizza ?

```
lake=inarinjarvi;country=Finland;basin=Inari;lat=69.04;lon=27.84;date=2022.32896591;
first_date=2002.87304033;last_date=2022.32266553;type=operational;diff=public
#
# Length: 58   km
# width:   19 km
# maximum of depth:   -- m
# Mean area:  1050  km2
# Catchment area:  13400  km2
# Mean volume:    15 km3
#
# saral            track number:   44 197 502 655 741
# sentinel3A       track number:  272 567 681
# envisat          track number:   44 197 502 655 741
# cryosat2         track number:  000
#
# corrections applied: Solid Earth tide, pole tide, ionospheric delay
# wet and dry tropospheric delay, altimeter biaises
#
# surface of reference: GGMO2C; high resolution global gravity model
# developped to degree and order 200 at CSR
# Center for Space research, university of Texas, austin, USA
# ref: Tapley B, Ries J, Bettatpur S, et al., (2005),
# GGM02 - an improved Earth Gravity field from GRACE, J.geod. 79: 467-478
```

Figure 2.11: Firts part of the ".txt" of the virtual station Inarinjärvi

```
#
# first date 2002 11 15 yr month day 15 hours 50 minutes
# last date 2022 04 28 yr month day 18 hours 33 minutes
#
# data file format
# (1): decimal year (2) date = aaaa/mm/dd (3): time = hh/mn
# (4): heigth above surface of ref (m), (5): standard deviation from heigth (m)
# (6): area (km2), (7): volume with respect to volume of first date (km3)
# (8): flag
#
# The water level, surface and volume algorithm developed at Legos, Toulouse, France
#
# citation: Cretaux J-F., Arsen A., Calmant S., et al., 2011.
# SOLS: A lake database to monitor in the Near Real Time water level
# and storage variations from remote sensing data, Advances in space Research, 47, 1497-1507
#
2002.87304033 ; 2002/11/15 ; 15:50 ;    119.7 ;     2.32 ; 9999.999 ; 9999.999 ;  ;
2002.97926370 ; 2002/12/24 ; 10:21 ;   119.56 ;     1.37 ; 9999.999 ; 9999.999 ;  ;
2003.07324581 ; 2003/01/27 ; 17:38 ;   119.32 ;     1.09 ; 9999.999 ; 9999.999 ;  ;
2003.17340944 ; 2003/03/05 ; 07:04 ;   118.65 ;     1.37 ; 9999.999 ; 9999.999 ;  ;
2003.26830479 ; 2003/04/08 ; 22:21 ;   118.05 ;     1.11 ; 9999.999 ; 9999.999 ;  ;
2003.45577816 ; 2003/06/16 ; 08:37 ;   119.19 ;     1.16 ; 9999.999 ; 9999.999 ;  ;
2003.55746195 ; 2003/07/23 ; 11:22 ;   119.33 ;     1.27 ; 9999.999 ; 9999.999 ;  ;
2003.64775685 ; 2003/08/25 ; 10:21 ;   120.43 ;     2.01 ; 9999.999 ; 9999.999 ;  ;
2003.74330479 ; 2003/09/29 ; 07:21 ;      120 ;     1.71 ; 9999.999 ; 9999.999 ;  ;
2003.94475647 ; 2003/12/11 ; 20:04 ;   119.15 ;     1.18 ; 9999.999 ; 9999.999 ;  ;
2004.03195203 ; 2004/01/12 ; 16:40 ;   119.05 ;     0.86 ; 9999.999 ; 9999.999 ;  ;
2004.12727497 ; 2004/02/16 ; 13:59 ;   119.14 ;     0.78 ; 9999.999 ; 9999.999 ;  ;
```

Figure 2.12: Second part of the ".txt" of the virtual station Inarinjärvi

This file mentioned above contain a lot of information about the virtual station. For example there are many general information about the virtual station like the name of the lake and of the country where is located, the basin, the geographical coordinates, the date of data download, the data of the first and the last altimetric measurement available. There are also information about the lake, for example the surface, the mean volume,the catchment area the mean area, the maximum depth and the width . There are also information about the satellite that gave the measurements and their track number and the corrections applied to rough altimetric data (Solid Earth tide, pole tide, ionospheric delay, wet and dry tropospheric delay, altimeter biaises), the surface of reference and the name of the Scientific center that developped the data.

About the record that will be validated are reported the day, the month, the year and the time when the measurement are taken, the altimetric measurements.

The virtual station retrieved are the followings:

- Inarinjärvi. 307 observations available of the satellites Saral, Sentinel 3A, Envisat and Cryosat 2;

- Kemijärvi. 103 observations available of the satellite Cryosat 2 ;
- Oulujärvi. 212 observations available of the satellites Jason 2, Jason 3, Saral and Sentinel 3A;
- Vanajavesi or also named Vanajanselka. 78 observations available of the satellites Sentinel 3A.

Therefore, only the virtual station Vanajanselka has Sentinel 3A monomission record.

2.3. In-situ data

In-situ data of rivers and lakes are downloaded by the site of the SYKE, the Finnish Environment Institute (https://www.syke.fi).

The first step has been to download the database. This operation has been done writing an R code. This online database is not easily to access and be handled.In the enormous database of the SYKE there are all the station and all the measures together in sequence.

It has been downloaded the informations of all the possible stations from the 1 January 2000 to October 2021. Then this database that is a mixture of stations has been re-organized with lists to be more user friendly. It has been downloaded also a database with in-situ stations information as the geographical coordinates, the river name, the station ID and more information. It has been retrieved in-situ water level records of 1690 stations.They are shown in the Figure 2.13.

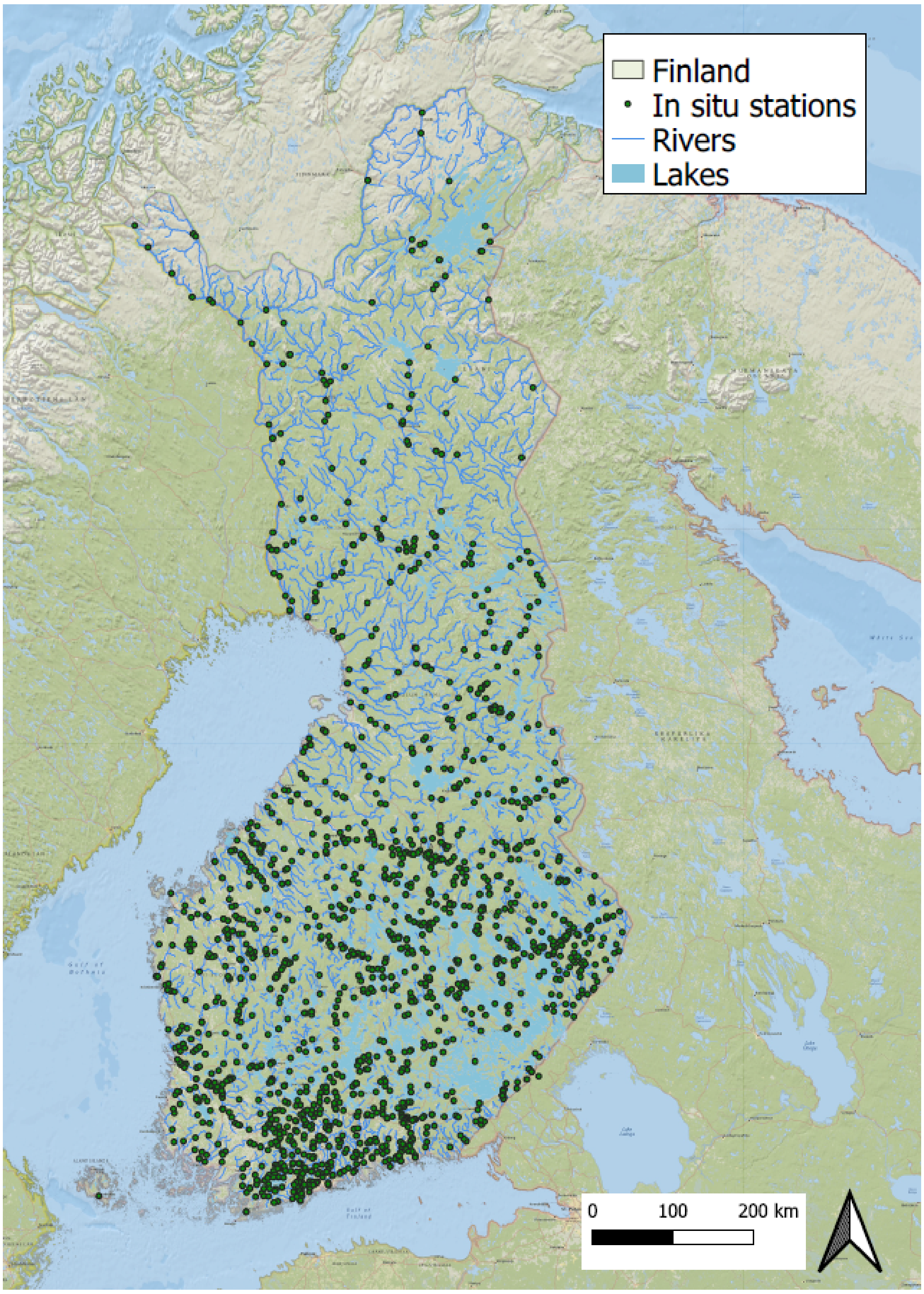


Figure 2.13: In-situ stations downloadable from the site of the SYKE

Each virtual station record has a different number of observations. Probably because many station are dismissed or are installed at different moment. There are many information: the ID identificators, the geographical coordinates (there are more than one typology of coordinates). Then it has been downloaded the records. On the site it was present only one entire database. It has been downloaded the entire database of the SYKE from 2000 to October 2022 through an R code. These data were organized into: id station, day/month/year, water level and others information often incomplete. It has been created list one list of measurements for each in-situ station. So finally, it has been disaggregated the big initial database.

3 | Methodology

3.1. Methods of comparison

The aim of this thesis is to understand the potentiality of using satellite altimetric data for water level retrieving.
In this section it will be illustrated the four methods applied to do the validation of satellite data to compare satellite and in-situ records. The same procedures have been applied both for rivers and lakes.

The validation will be made between a couple made up with one virtual station and one in-situ station. The probability that a virtual station may coincide with an in-situ station is very low [23]. Therefore to go on with the comparison it's necessary to define a criteria to built the couples and to define which ones have to be considered and which ones have to be discarded.

A data pre-processing has been applied to select just the stations for which there exists a continuous length of data series. Many stations from the SYKE database have does not have any observations after the year 2000. In this way we neglected the in-situ stations that are broken or dismissed and could influence the analysis.

Secondly, we needed a method to select couples in base of their relative distance, that is the distance between their geographical coordinates. To compare each virtual station with the in-situ one we selected different radius of distance. The starting point is the virtual station that become the center of a circumference with a given radius. In this work it will be considered the following radius: 5, 10, 15, 20, 25, 35, 50 km. The in-situ stations in the selected circle are used to build couples and the ones out the circle are discarded.

The third step is the elimination of the couples not hydraulically connected to be sure to consider just the couple belonging to the same hydraulic network that in Finland is very complicated and dense of channels. Therefore, for Finland case this operation could be done automatically only if it is available a good shapefile with all the correct name of channels and rivers or it can be done manually trying to be careful using the shapefile and the maps with visual check. Increasing the radius, it is more probable to consider an higher number of in-situ stations not hydraulically connected.

The fourth step is the joining of the in-situ record and the virtual station record. It's not a formal union. For each couple and for each station datum has been searched the corresponding temporal datum of the in-situ station that composes the couple.

The fifth step is to eliminate separately the outliers of the virtual station record and in-situ record of the couples. It has been considered an outliers a measurement outside the mean+/-3*standard deviation of the sample. Therefore, if a virtual station measure of a certain day is an outlier it will be discharged also the corresponding in-situ measure of the same day and vice versa.

The sixth step is to discard the couples with less than 20 observations.

It could be useful to the comprehension that it was very clear, from the beginning, the need to not consider the couples not hydraulically connected even if the have good indicators. By the elimination of the couples mentioned above, the agreement become higher and also became higher the percentage of the couples with good agreement for each radius.

Satellite and in-situ data are often referred to a different reference level, it is highly improbable that they may coincide [23]. Therefore it is not possible a direct comparison of the two records. There are in literature many options for the validation. A first approach could be to refers the two data series to the same reference. But it needs precise topographic and instrumental information that in many case are not possible to achieve. For example the details about river section of the water body in the exact location of the virtual station, the Digital elevation Model (DEM) and the reference surface related to the data collection [23].

Following Deidda et all. [23] four different methodologies have been applied to compare in-situ and satellite data without the knowledge the two reference levels:

- **Method 1:** It will be compared the differences between the differences between successive times of satellite and in-situ measurements;
- **Method 2:** It will be compared the satellite and in-situ measurements assuming that the initial value of the joint measurements is equal (specifically the value of satellite water level is equal to the value of in-situ one);
- **Method 3:** It will be compared the satellite and in-situ measurements calibrating the initial water level of satellite measurements minimizing the MAE;

- **Method 4:** It will be compared the satellite and in-situ measurements calibrating the initial water level of satellite measurements minimizing the RMSE.

These two records that will be compared must have the same unit of measure for example metres or centimeters. In this case study, for in-situ measurements has been downloaded data referred to centimeters. They has been converted to meters. Altimetric data of Sentinel 3 has been downloaded in meter directly.

3.2. Statistical indicators

For each methodology will be calculated the couples statistical indicators listed below:

- R, the Pearson correlation Coefficient;
- RMSE, the Root Mean Square Error;
- MAE,the Mean Absolute Error;
- MBE, the Mean Bias Error;
- NSE, the Nash-Sutcliffe efficiency.

These indicators are commonly used in satellite data validations analysis, see Deidda et all [23]. The NSE usage is suggested by Halicki et all [35].

The sample of in-situ water level measurements is identified as $\{y_1, y_2, \ldots, y_n\}$ and the sample of satellite water level as $\{x_1, x_2, \ldots, x_n\}$. The variable n is the number of sample elements.

The Pearson correlation Coefficient is described as

$$\mathrm{R} = \frac{s_{yx}}{s_y s_x} \tag{3.1}$$

where:

- $s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}}$ is the sample standard deviation of in-situ measurements;
- $s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ is the sample standard deviation of satellite measurements;
- $\bar{y}$ is the mean of the in-situ measurements sample;
- $\bar{x}$ is the mean of the satellite measurements sample;
- $s_{yx} = \frac{\sum_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x})}{n-1}$ is the sample covariance.

R evaluates the linear correlation between two variables and can take values between -1 and 1, extremes included (R $\in [-1, 1]$). The values -1 and $+1$ indicate the perfect linear relationship respectively negative and positive between the two data series. 0 states the absence of the linear relationship. The closer the index is to zero, the weaker the linear relationship. The closer it is to -1 or $+1$, the stronger the linear relationship is. More tents to zero and weaker the linear relationship is, more tents to -1 or $+1$ and stronger the linear relationship will be.

RMSE indicator is the Root Mean Square Error,

$$\text{RMSE} = \sqrt{\frac{\text{RSS}}{n}} \tag{3.2}$$

Where RSS $= \sum_{i=1}^{n} (y_i - x_i)^2$ is the Residual Sum of Squares . RMSE can be only equal or greater than 0, so non-negative (RMSE $\in [0, +Inf]$). If it is 0 it indicates that there is a perfect agreement between the two samples. It describes how much satellite water levels depart from in-situ water levels. More its value increase more the disagreement grows. To have a good agreement is necessary the RMSE value as low as possible. The advantage of using the RMSE index is that the terms present in the RSS sum are quadratic, so is not negative in order to avoid that positive and negative values of the deviations can cancel each other . Furthermore, more the error values increase more they are amplified with the squares.

MBE index is the Mean Bias Error and it is described with the following formula,

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n} (x_i - y_i) \tag{3.3}$$

It is the average bias in the prediction. The range of possible values is unlimited(MBE$\in [-Inf, +Inf]$.). Positive values denote an average overestimation of altimetric measurements compared to in-situ measurements, while to negative values correspond an underestimation .

MAE index is the Mean Absolute Error,

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |x_i - y_i| \tag{3.4}$$

It is the index of the average of the absolute values of the errors made by the satellite compared with in-situ measurements. It can only assume values equal or greater than 0. Lower is the value lower is the average error, so satellite data is of good quality in comparison with in-situ data.

NSE index is the Nash-Sutcliffe Efficiency, defined as [47]:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y_i})^2} \tag{3.5}$$

It is a measure of satellite performance It is a good performance indicator also if the number of the sample is low [41]. It expresses the agreement between the satellite estimations and in-situ observations, but it does not shows errors, so it can be used also for different rivers comparison. The range values is from -Inf to 1,including extreemes(NSE$\in [-Inf, 1]$). If the NSE tends to 1, the altimetric measurements and in-situ water levels agree in amplitude, phase and mean. If NSE is near to 0 means that the satellite measurements have similar characteristics of the average in-situ data. So, closer the model efficiency is to 1 more satellite measurements are reliable.

4 | Results

In this chapter it will be shown the results of the validation process of Finland rivers and lakes altimetric data.

It could be useful to recall that the rivers and lakes satellite records, under study, have been illustrated in the subsubsections 2.2.2 and 2.2.3.

The preprocessing of data retrieved and the four methods of comparison have been illustrated in the subsection 3.1. Below there is a brief description of what compares each method:

- Method 1: Water level differences between two successive measurements, separately for satellite and in-situ records;
- Method 2: Satellite measurements are calibrated using the difference between the first value of the satellite and the corresponding in-situ value;
- Method 3: Satellite measurements that are calibrated using the value that optimize the MAE and the in-situ value;
- Method 4: Satellite measurements are calibrated using the value that optimize the RMSE and the in-situ value.

As said in the subchapter 3.2, it has been calculated the followings statistical indicators: R, MAE, MBE, RMSE and RMSE. The help to assess the performance of altimetric data to replace or be complementary to record measured in-situ. Therefore, it if the could be used for hydrological purposes. The most considered indicators during the analysis of validation of this work have been R and RMSE.

The Pearson Coefficient, R, is usually used to evaluate the linear correlation between two samples of observations. Or better, it is useful to understand if the two trends have the similar behaviour.

The RMSE is used to evaluate the differences between the two trends. It considers the sum or the squares of the errors. Therefore, more high is the error and more is its weight.

Firstly it will be shown the result of rivers validation and then the lakes one, respectively in the subsection 4.1 and 4.2. Rivers altimetric data has been analyzed in the following way:

- Annual period;
- Summer period;
- Winter period.

Lakes altimetric data are analyzed only considering the annual records.

4.1. Rivers

The satellite data used for the analysis have been downloaded from Copernicus website (//land.copernicus.eu/global/products/wl). Currently three virtual stations are available for rivers, named: 12781, 12783, 12784. All measures are referred to the satellite Sentinel 3.

Each virtual station has been compared with a set of in-situ stations considering six different radius of distance: 5, 10, 15, 20 ,25, 35 and 50 km. The virtual station is considered as the center of the circumference. Generally, the number of the couples selected for the comparison increase if the radius increase.

The couples with less than 20 observations as length of the sample of comparison are discarded because otherwise the statistical analysis could not be robust and significant.

The analysis has been done using the four different methodologies and calculating, for each one, the five metrics for each one (R, MAE, MBE, RMSE and NSE).

As first study have been considered all the couples considering all the in-situ stations that fall inside the various radius of circumference. That implies that we are considering both hydraulically and not hydraulically connected stations. Therefore it has been analyzed the general behaviour of every virtual station with all the in-situ stations located inside the circumference, considering different range of radius. This is a general starting analysis since it has been considered couples without a hydrological connection between the satellite station and the in-situ station.

In the Figure 4.1 are reported the number of the couples considered for the analysis for each radius, including the ones not hydraulically connected. Then, in the third column there are displayed the number of couples with positive correlation. For each methodology are displayed the percentage of couples that falls into two different category named "G" and

"VG". These percentages are referred just to the couples with R positive for all methods, that one with negative values of correlation have been discarded from the analysis.

The two categories "G" and "VG" reported in table 4.1 are defined in the following way:

- **G:** "Good" agreement. Couples with the R greater or equal to 0.7 and the RMSE minor or equal to 0.4 ($R >= 0.7$ and $RMSE <= 0.4$)
- **VG:** "Very Good" agreement. Couples with the R greater or equal to 0.8 and the RMSE minor or equal to 0.3 ($R >= 0.8$ and $MSE <= 0.3$)

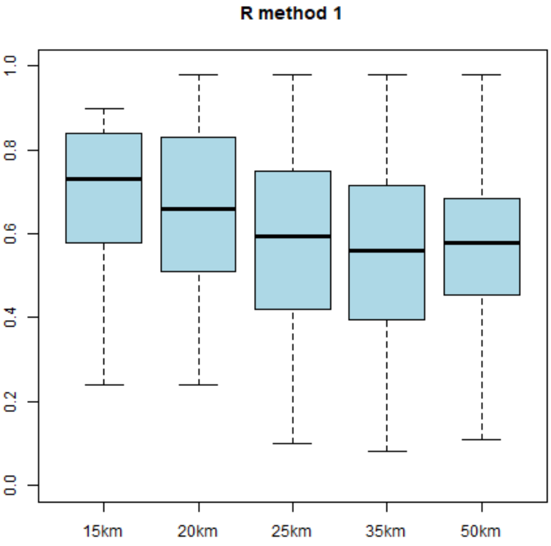
			Method 1		Method 2		Method 3		Method 4	
Radius	Couples	Couples with all R >0	G	VG	G	VG	G	VG	G	VG
[km]	[-]	[-]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
5	3	2	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0
10	7	4	50,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0
15	12	9	44,4	33,3	33,3	33,3	33,3	33,3	33,3	33,3
20	26	21	33,3	23,8	23,8	23,8	23,8	23,8	23,8	23,8
25	36	28	25,0	17,9	17,9	17,9	25,0	17,9	25,0	17,9
35	65	51	15,7	11,8	13,7	11,8	17,6	11,8	17,6	11,8
50	99	80	15,0	8,8	13,8	8,8	20,0	11,3	20,0	11,3

Figure 4.1: Number of couple, hydraulically connected or not, analyzed for each radius of distance, percentage of couples that perform positive values of Pearson correlation and percentage of couples with best results in terms of R and RMSE

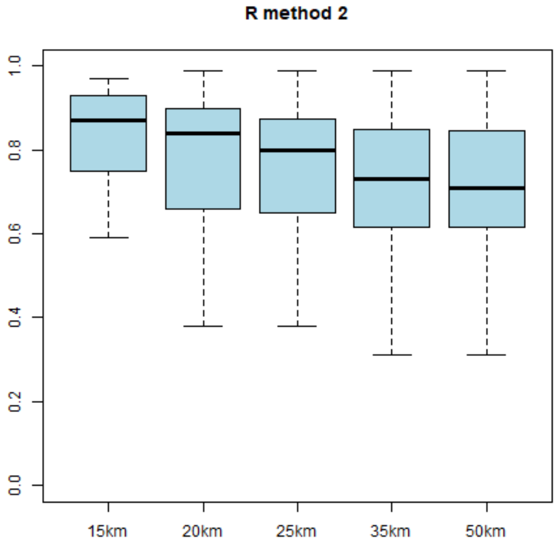
From the Figure 4.1 it can be seen that obviously considering smallest value of radius the percentage of couples with good and very good performance (**Group G** and **Group VG** as defined above) are the highest. It can be seen that also considering a radius of 20 km there are still the 25% of couples performing a good agreement. Still at 50 km of distance we have some couples for which the correlation is higher then 0.8. It has to be pointed out that these results include also the not hydraulically connected couples. This means that the low values of correlation can be done just for the fact that we are considering stations not laying in the same river section. A further development of the work, as it will be shown later, has been the selection of just hydraulically connected couples.

In the Figures 4.2 are shown the boxplots of the Pearson Coefficient related to different radius, for the method 1 (Figure 4.2a) , method 2 (Figure 4.2b), method 3 (Figure 4.2c) and method 4 (Figure 4.2d).

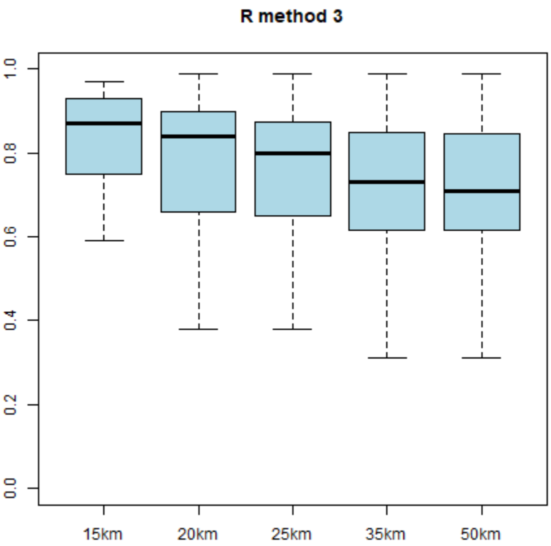
It is immediately important to emphasize that the boxplot of method 2,3 and 4 are identical because their virtual stations records are just a translation of a defined quantity that change from each method, as said in the section 3.1. Thus, the virtual stations trends



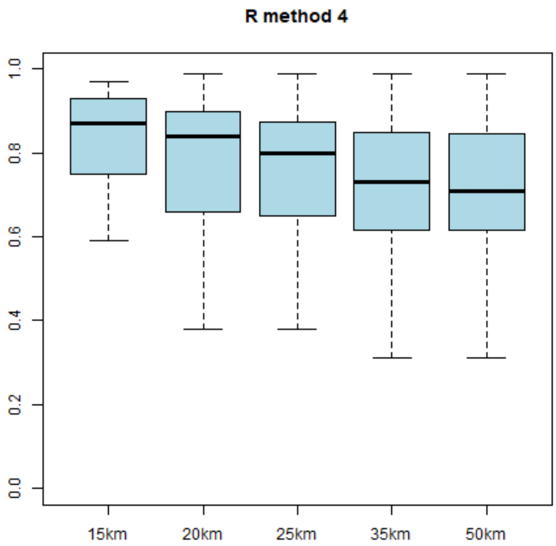
(a) R of method 1



(b) R of method 2



(c) R of method 3



(d) R of method 4

Figure 4.2: R boxplot of couples hydraulically connected and not related to the five radius

are not altered. But, surely the error changes, increasing or decreasing. As described in the subsection 3.1, the first method is a comparison of the temporal differences, and it is that one that gives the worst results respect the others three. So, setting the same reference system imposing the equality of first measure (method 2) or translating the altimetric record minimizing the MAE (method 3) or the RMSE (method 4), the results obtained have been better. Method 1 Pearson coefficient boxplot is different from the remaining methods because the sample considered is the temporal differences.

With lower distance the performance increases for all the methods. Increasing the radius, it is more probable that it has been considering an increasing number of couples that are not hydraulically connected and so that can not be hydrologically comparable. Despite that, for radius of 50 km the median of Pearson is around 0.6 considering the first method and around 0.7 considering the remaining methods.

For method 2,3 and 4, the median value of Pearson was on the range of 0.8 considering till 25 km of radius distance.

In the next boxplot, Figure 4.3 can be found the results in terms of Pearson correlation for the radius of 15 km for all the four methods. Considering a shorter radius it is more likely to consider just stations in the same river network, and as can be seen the results are quite good (correlation higher than 0.8 with the methods 2,3 and 4).

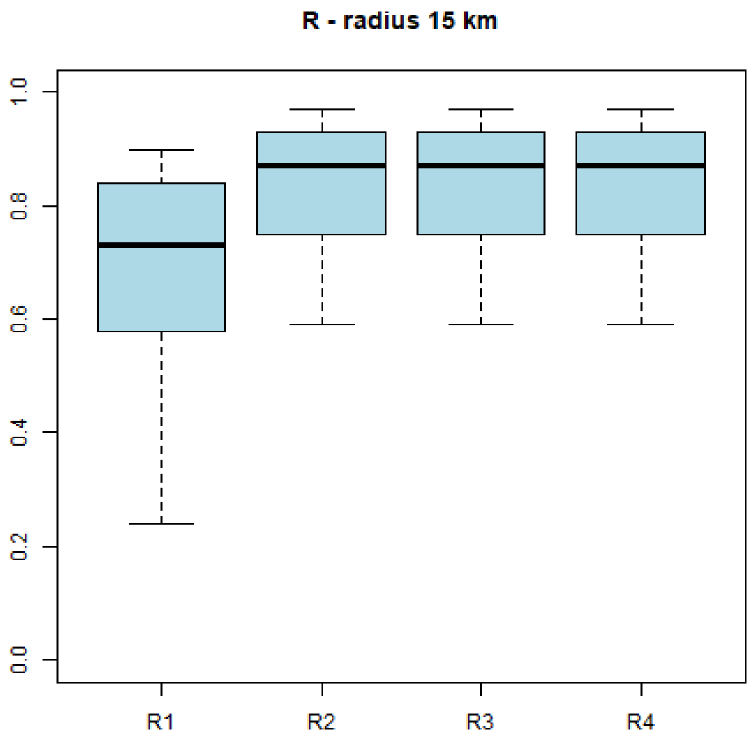


Figure 4.3: R boxplot of the four method related to the radius 15 km

As said before, the method 2, 3 and 4 have the same Pearson coefficient boxplots. This is reasonable because their altimetric records are shifted between them of a constant quantity specific quantity for all the temporal series. The specific quantity of virtual station record shifting depends from the method.

As can be seen from Figure 4.3 considering shorter radius the couples performs high values of correlation (in median higher than 0.8). Method 1 is the one with the worst performance.

The RMSE median tends to increase increasing the radius (Figure 4.4).

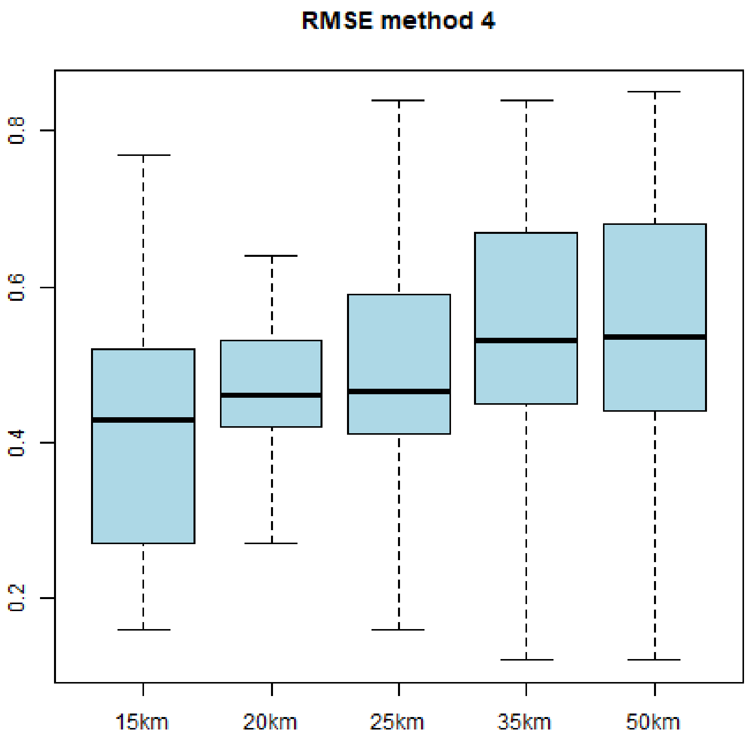


Figure 4.4: RMSE boxplot of the method four related to the five radius

To improve the analysis it has been excluded the couples not hydraulically connected. The results expected were the increasing of R indicator and the decreasing of the RMSE regarding the general behaviour of the circles.

The next step of the work has been to individuate the hydraulically connected stations with a manual selection, still considering annual records. In this way it has been possible to filter just the really comparable stations and to give more importance to the largest radius that will have higher agreement than in the previous example. The hypothesis was that the statistical indicators will express an increasing of the confidence, R will increase and RMSE will decrease, for largest diameter because of the decreasing of the probability to consider the couples not hydraulically connected in the radius lists. Thus, the number of comparable couples decreased respect to the initial example. But they are really comparable to drive to have more correct conclusions.

Many couples showed one or more than one R negatives indicators, considering the four methods. An example of that will be described later trying to give and explanation of this result. These couples has been deleted from the radius lists.

In the Figure 4.5 are reported the numbers of the couples for each radius of distance considering both the hydraulically connected and not, the numbers of the couples only hydraulically connected, the numbers of the couples with positive correlation for all methods and the percentages of couples that perform better in terms of R and RMSE 4.1. It has been considered the classification of the quality of the agreement used for the previous example and in addition the class named E that means "Excelent" agreement. It means that the R is greater or equal to 0.9 and the RMSE minor or equal to 0.2.

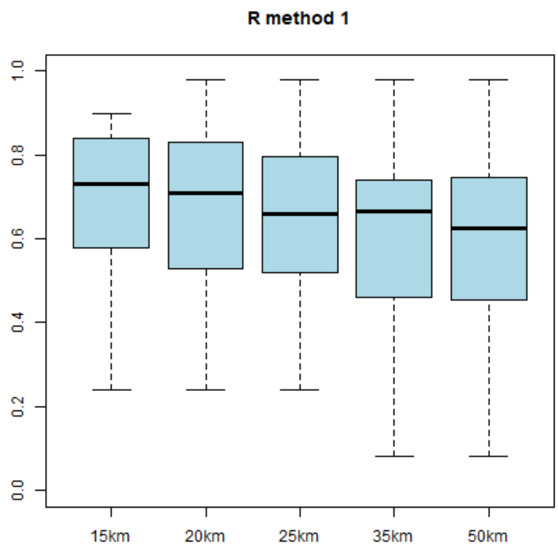
				Method 1			Method 2			Method 3			Method 4		
Radius	Couples	Couples hydr. Connected	Couples with all R >0	G	VG	E	G	VG	E	G	VG	E	G	VG	E
[km]	[-]	[-]	[-]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
5	3	3	2	50,0	50,0	0,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0	50,0
10	7	7	4	50,0	25,0	0,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0	25,0
15	12	12	9	44,4	33,3	0,0	33,3	33,3	11,1	33,3	33,3	22,2	33,3	33,3	22,2
20	26	21	17	41,2	29,4	5,9	29,4	29,4	11,8	29,4	29,4	23,5	29,4	29,4	23,5
25	36	29	23	30,4	21,7	4,3	21,7	21,7	8,7	30,4	21,7	17,4	30,4	21,7	17,4
35	65	45	34	23,5	20,6	2,9	20,6	17,6	8,8	26,5	17,6	14,7	26,5	17,6	14,7
50	99	63	48	22,9	22,9	2,1	22,9	14,6	8,3	31,3	18,8	12,5	31,3	18,8	12,5

Figure 4.5: Number of couple, hydraulically connected, analyzed for each radius of distance, percentage of couples that perform positive values of Pearson correlation and percentage of couples with best results in terms of R and RMSE

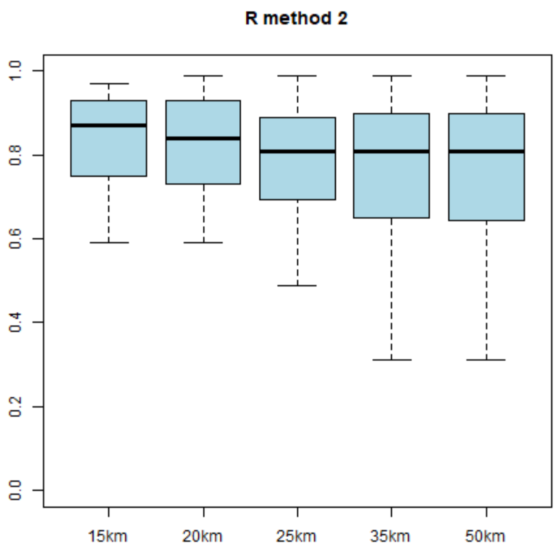
The Figure just above shows that the method 3 and 4 gave the best result in terms of percentage considering the index "Excelent".

The Pearson coefficient of the four method are reported in the Figure 4.6. The Figure 4.6b, 4.6c, 4.6d are identical for the same motivation expressed during the explanation of the couples hydraulically connected and not. Increasing the radius, the median of the Pearson coefficient tends to decrease considering all methods. But for 25, 30 and 50 km there are more or less a stability of the median. This because increasing the radius has not been considered only new couples with lower performance but also couple with good performance with the virtual stations. Thus, the median value didn't decrease having a sort of compensation adding good in-situ stations and not good ones. The variability increase as in the case which were considered all the couples without hydraulic discrimination and the first method is even less accurate than the others. But now the linear correlation is stronger, the median of the methods 2,3,4 is higher than 0.8 for every radius. The methods 2,3 and 4 have better performances than the method 1.

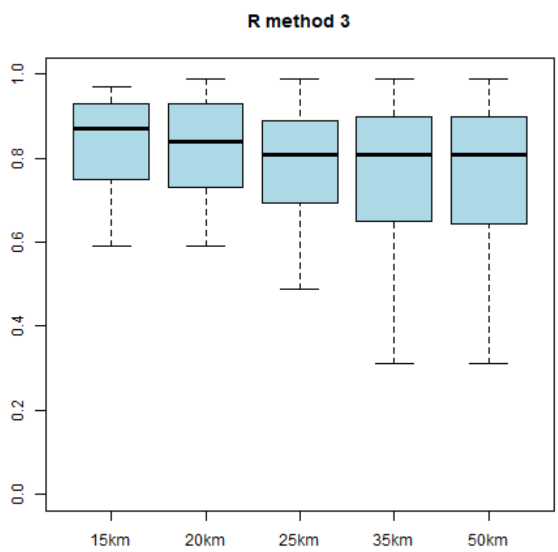
These results described above are reasonable and they as been more or less expected. Increasing the radius value means that are considered in-situ stations with higher distance from the virtual station. Therefore, the probability to have more perturbation between the virtual station and the in-situ station is higher. Can be possible the are non compared same river section regimes. Perturbations could be even natural even anthropic. Examples



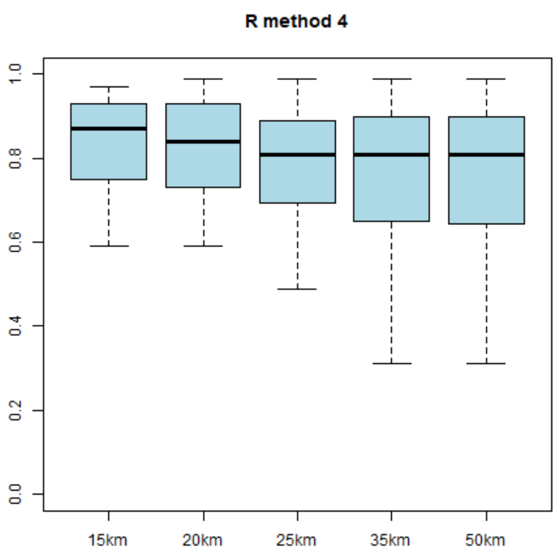
(a) R method 1



(b) R methood 2

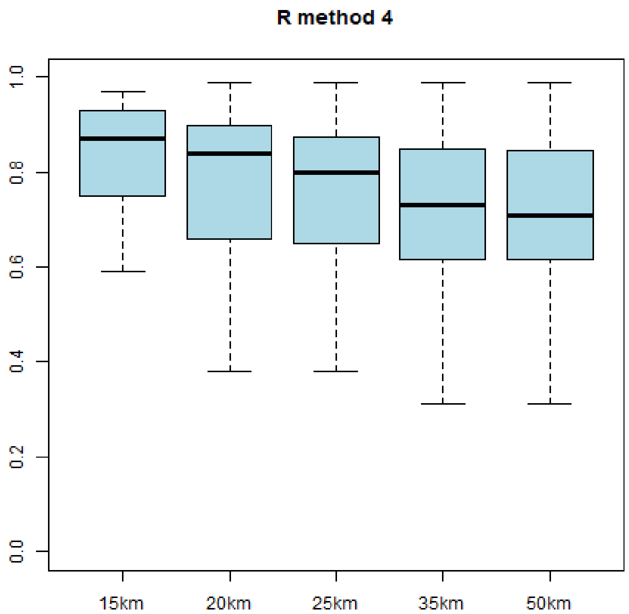


(c) R method 3

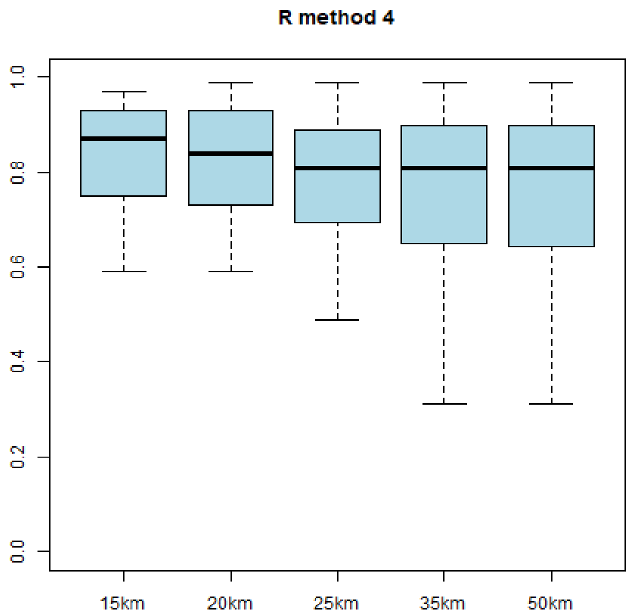


(d) R method 4

Figure 4.6: R boxplot of couples hydraulically connected



(a) R of the method 4 of the couples hydraulically connected and not



(b) R of the method 4 of the couples hydraulically connected

Figure 4.7: Comparison of the method 4 between the couples hydraulically connected and the couples hydraulically connected and not

of that could be the stacks of the bridges, the presence of river tributaries and effluents, the variations of the section (width and shape),the sharp variations in slope, the presence of island and bifurcation. It could be present also dam with the purpose of lamination the flow peaks or of hydroelectrical production. The consequence is on the error. The RMSE of the method 4, so minimizing this indicators, is higher increasing the radius, Figure 4.8.

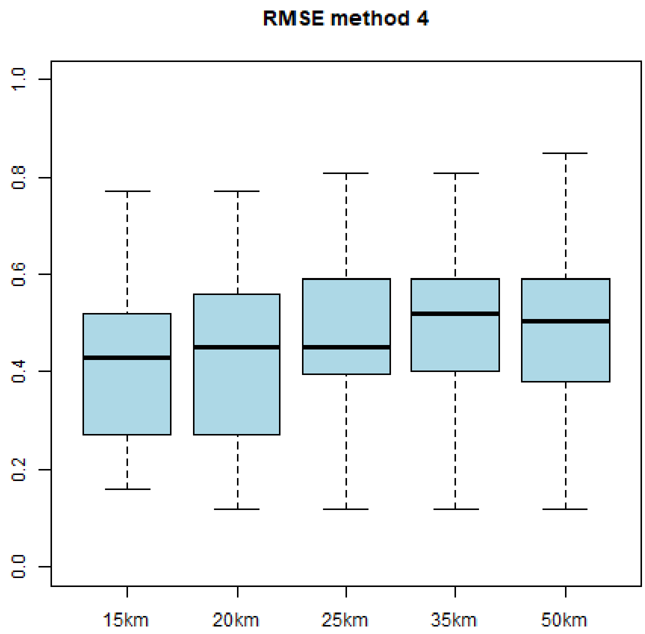


Figure 4.8: RMSE boxplot of method 4 related to the five radius

Now, it will be illustrated the rivers virtual stations and many examples of couples validated.

The virtual station 12781 and the in-situ stations related to the radius of 50 km are shown in the Figure 4.9. The virtual station 12781 is very important because the first two best couples among all the entire study are made up with that one. Therefore, they are also the best couples for this virtual station. These two couples are composed by the in-situ stations 2458 and 2468.

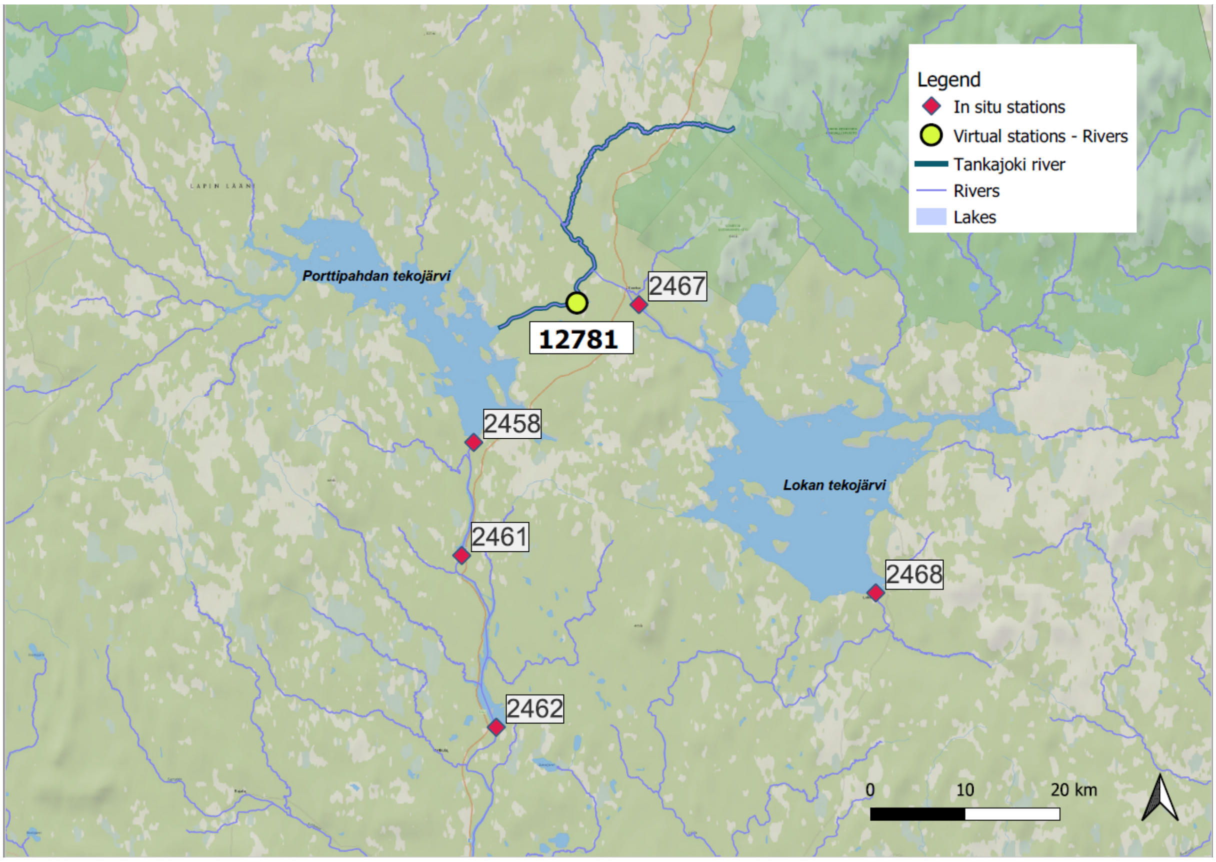


Figure 4.9: Virtual station 12781 and the in-situ station related to the radius of 50 km

The couple composed by the in-situ station 2458 has also the best performance of all the rivers couples of the entire annual study. The in-situ station 2458 is shown in the Figure 4.10.

Figure 4.10: In-situ station 2458

This couple has the highest R for the methods 2,3 and 4 (they are equal) and the lowest RMSE for the method 4. It has also the highest NSE indicators for method 4. Methods 2, 3 and 4 have higher R value and lower RMSE value than the method 1. But the differences from the two groups of methods are very minimal. These results are shown in the Figure 4.11

Virtual station	In situ station	Number of observations	R1	RMSE1	R4	RMSE4	NSE4
12781	2458	53	0.98	0.16	0.99	0.12	0.98

Figure 4.11: Number of the observations and statistical indicators of the couple 12781-2458

In the Figure 4.12 are illustrated the in-situ temporal record and the altimetric record modified using method 2. Their firsts observations values coincide, they are imposed equal.

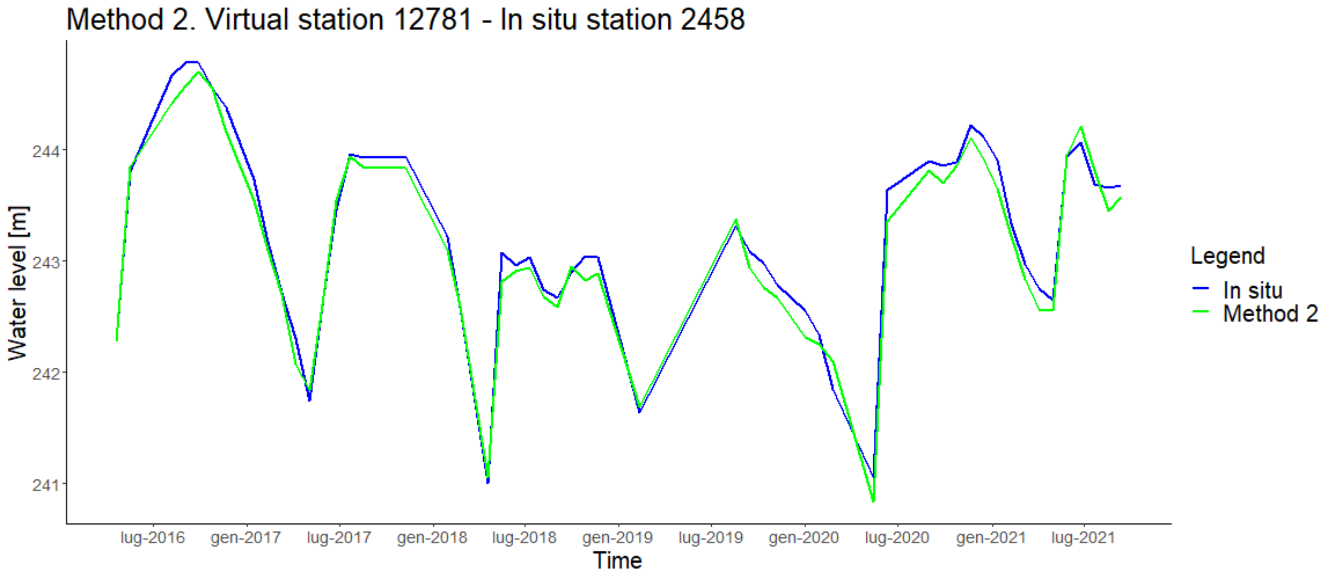


Figure 4.12: Temporal trends of virtual station 12781 (method 2) and of in-situ station 2458

In the Figure 4.13 are shown the in-situ temporal record and the altimetric record related to the minimization of the MAE (method 3) and the RSME (method 4).

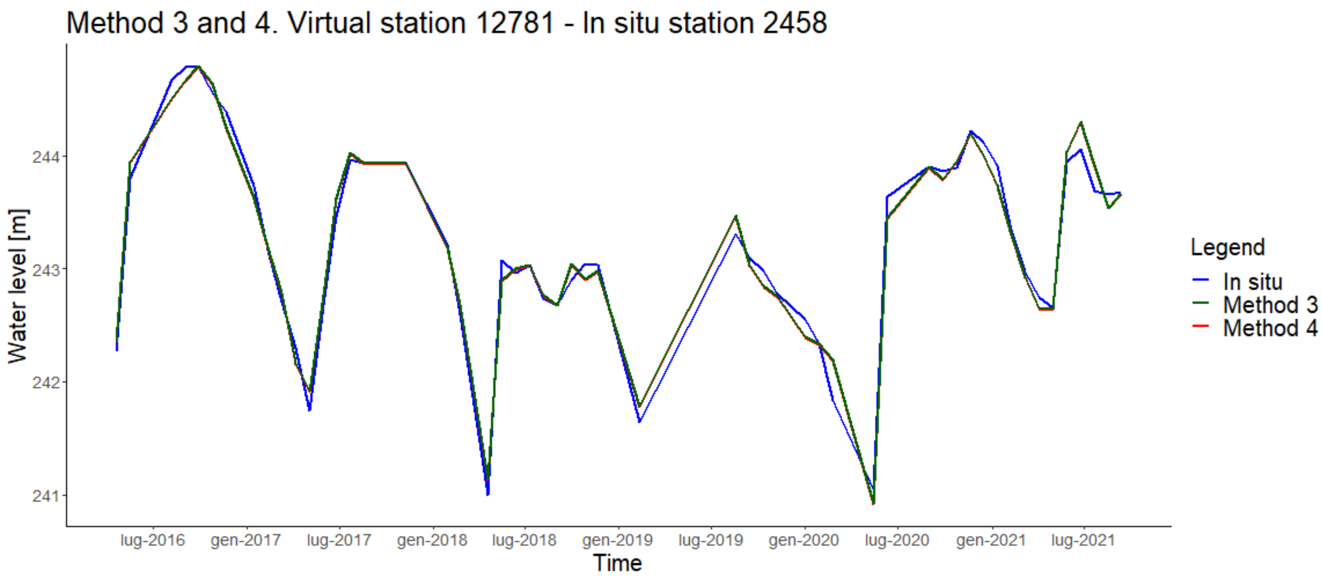
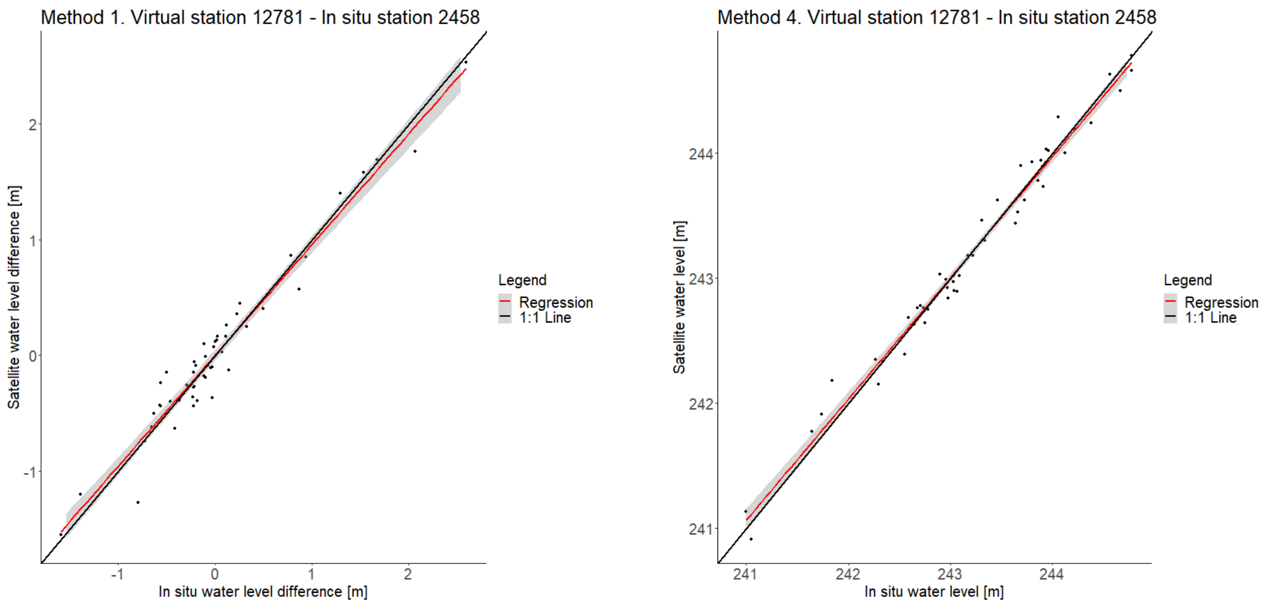


Figure 4.13: Temporal trends of virtual station 12781 (method 3 and 4) and of in-situ station 2458

The union of the last two Figures (the 4.12 and the 4.13) is presented in the Figure 4.14. Even a visual analysis makes clear the good agreement of the comparison in this case. In



(a) Method 1 (b) Method 4

Figure 4.15: Scatter plot of the couple 12781-2458 (Method 1 and 4)

fact, the lines overlap and it is difficult to distinguish them.

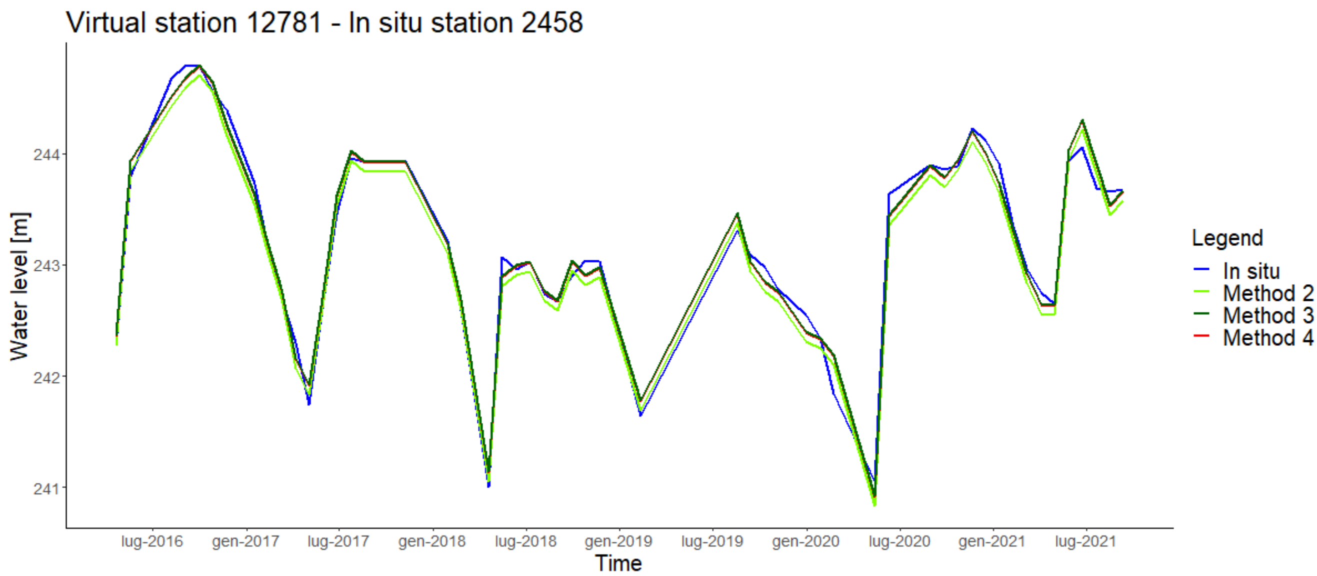


Figure 4.14: Temporal trends of virtual station 12781 (method 2, 3 and 4) and of in-situ station 2458

The scatter plot related to method 1 and 4 are shown in the Figures 4.15a and 4.27b. It is possible to understand what means to have a very good correlations. Therefore, this satellite virtual station is able to retrieve water level measures with good accuracy.

The second best couples with a high agreement is the one made up with the in-situ station 2468. This in-situ station is showed in the Figure 4.16. It is located upstream the dam.

Figure 4.16: In-situ station 2468

In the Figure 4.17 are reported the most many indicators related of this couple.

Virtual station	In situ station	number of observation	R1	RMSE1	R2,R3,R4	RMSEcalRMSE	NSE cal rmse
12781	2468	53	0,95	0,16	0,98	0,19	0,94

Figure 4.17: Number of observations and statistical indicators of the couple 12781-2468

The Figures show the same type of graphs (Figures 4.18, 4.19 and 4.20 and boxplot (Figures 4.21a and 4.21b) of the couple above. The results are similar. It has been found high accuracy of altimetric measurements compared with the in-situ station and also high R and low RMSE for the methods 1 and 4.

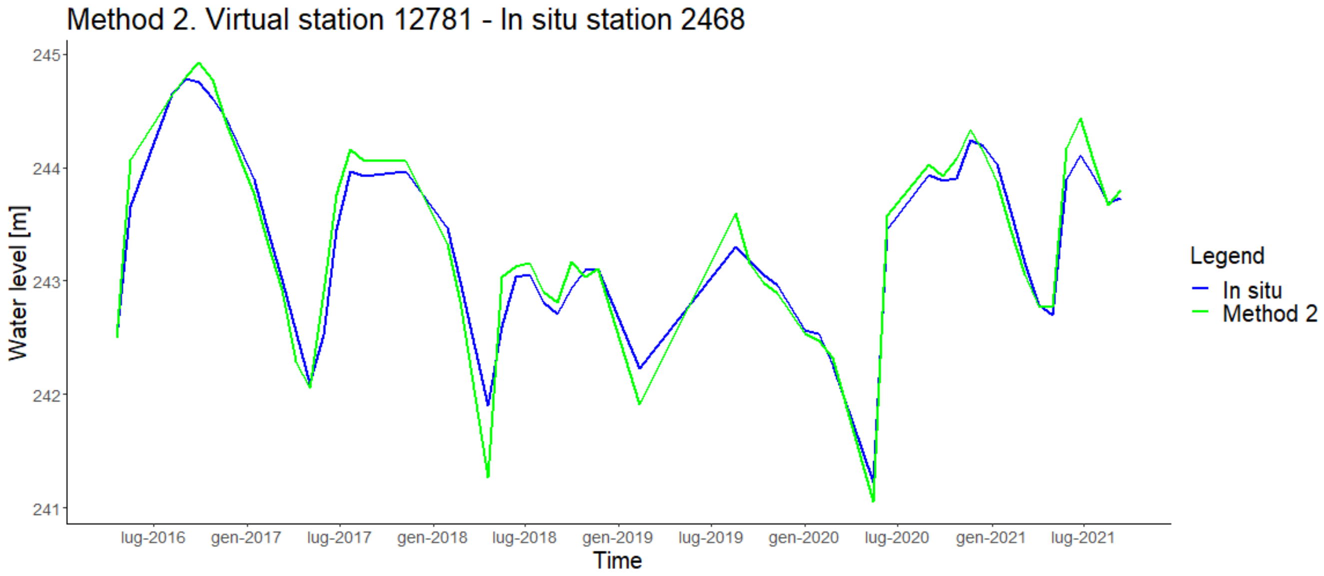


Figure 4.18: Temporal trends of virtual station 12781 (method 2) and of in-situ station 2468

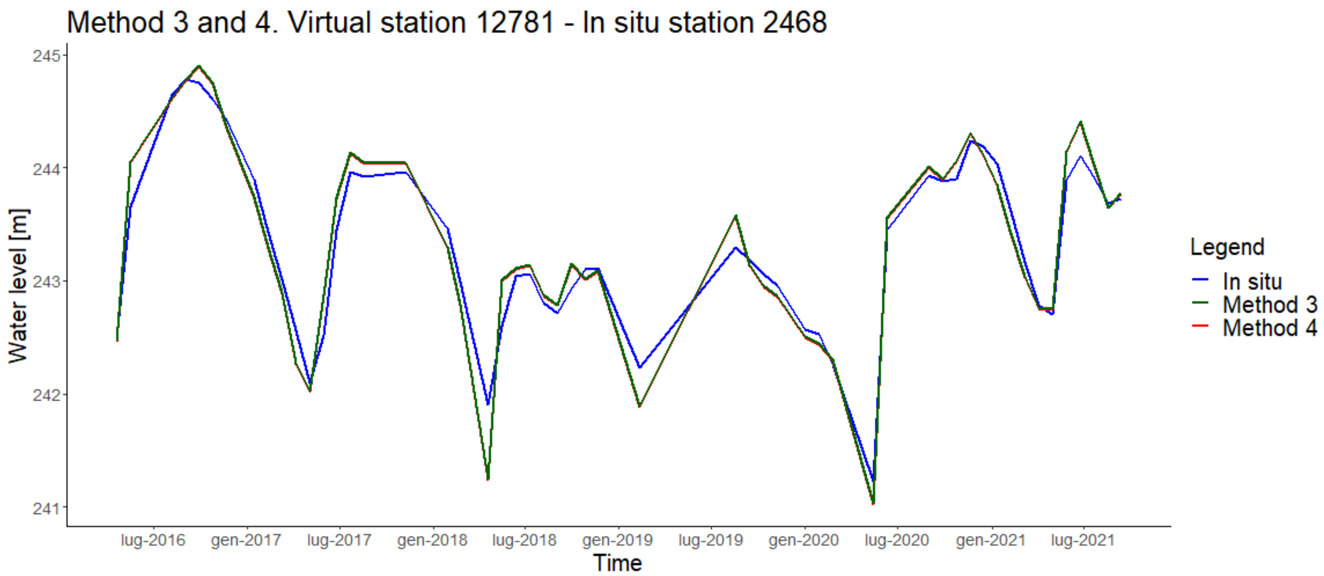
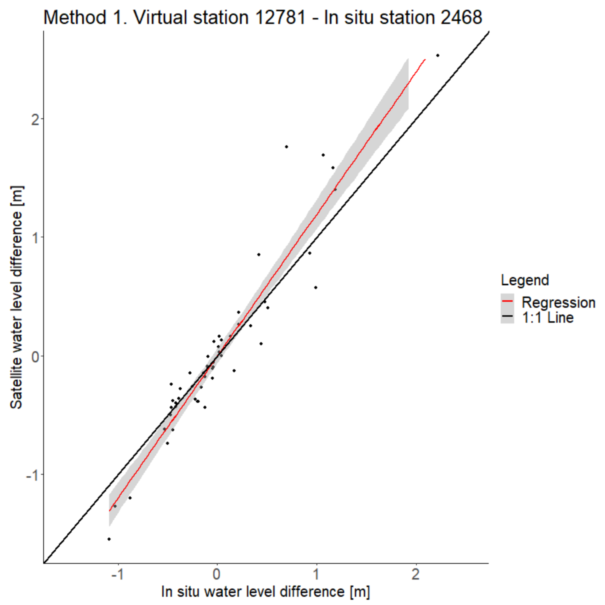
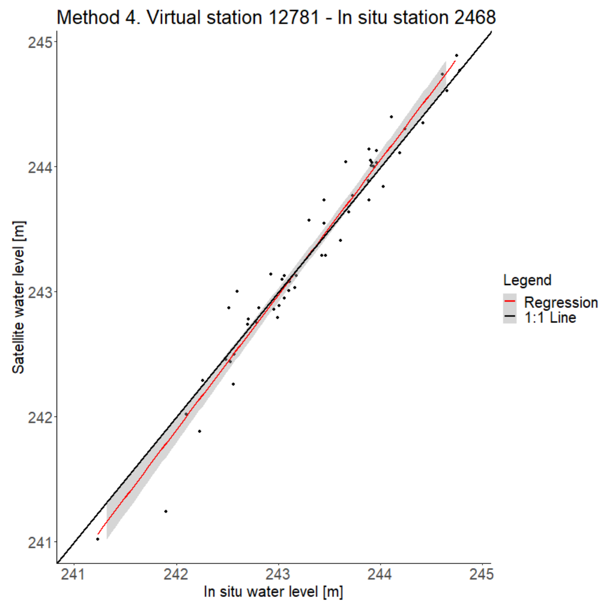


Figure 4.19: Temporal trends of virtual station 12781 (method 3 and 4) and of in-situ station 2468



(a) Method 1



(b) Method 4

Figure 4.21: Scatter plot of the couple 12781-2468 (Method 1 and 4)

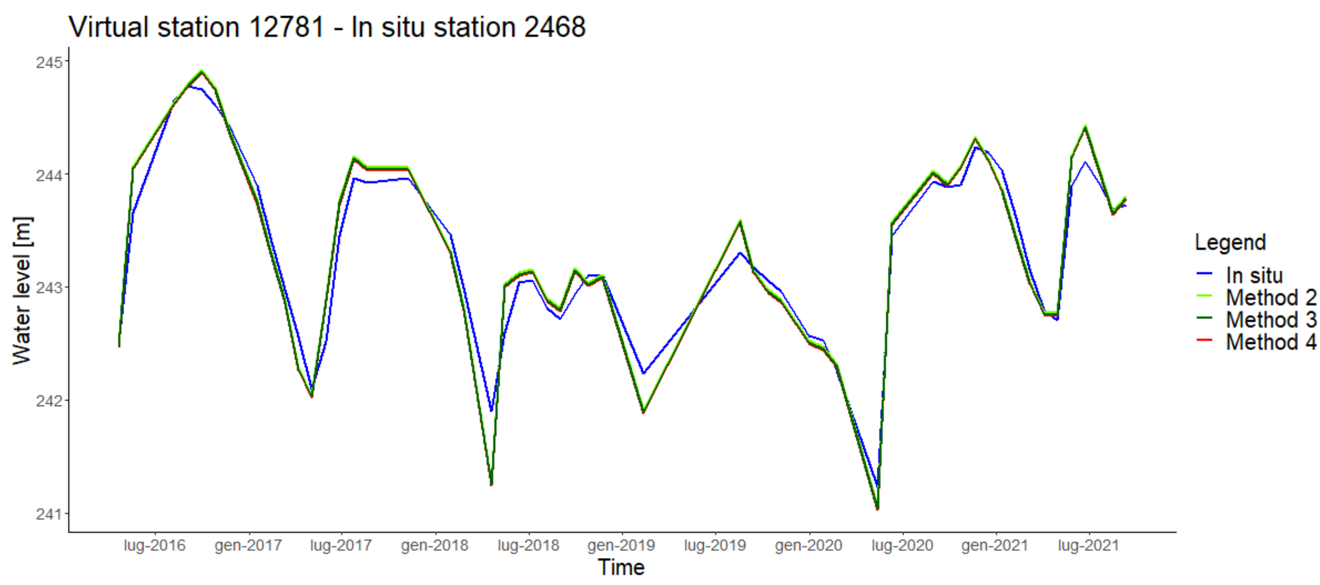


Figure 4.20: Temporal trends of virtual station 12781 (method 2,3 and 4) and of in-situ station 2468

The virtual station 12783 has an higher number of in-situ stations than the 12781 to make data comparison with. There are two particular in-situ stations: 2277 and 2278. The 2277 is downstream to a dam and the 2278 is upstream. Both two are upstream respect the virtual station. Near the dam there is a power plant.

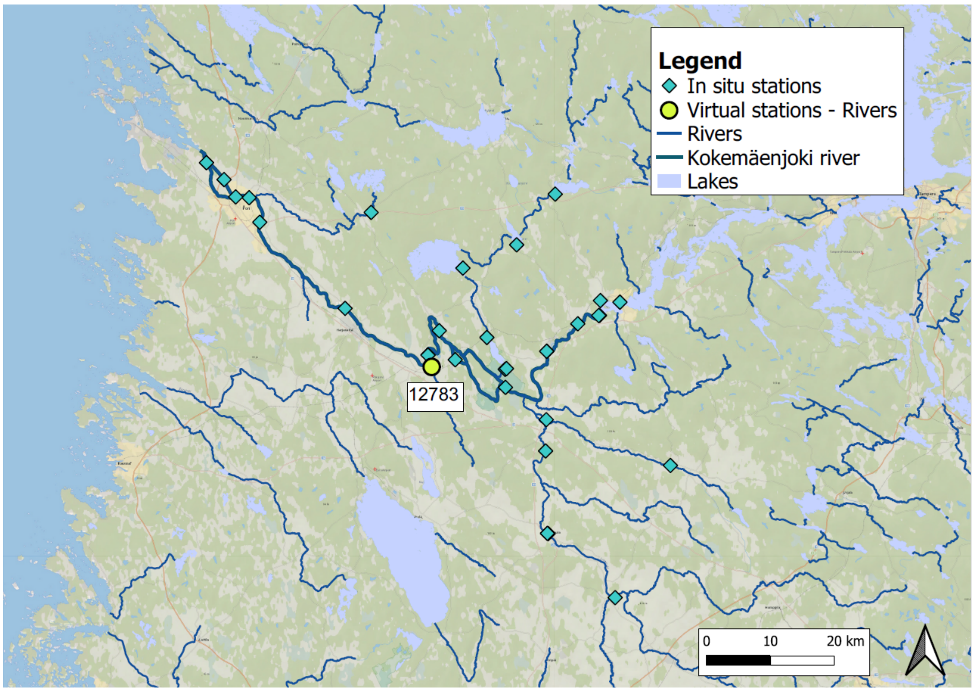


Figure 4.22: Virtual station 12783 and in-situ stations related to the radius of 50 km

The Figure 4.23 shows the locations of the virtual station 12783 and the in-situ stations 2277 and 2278.

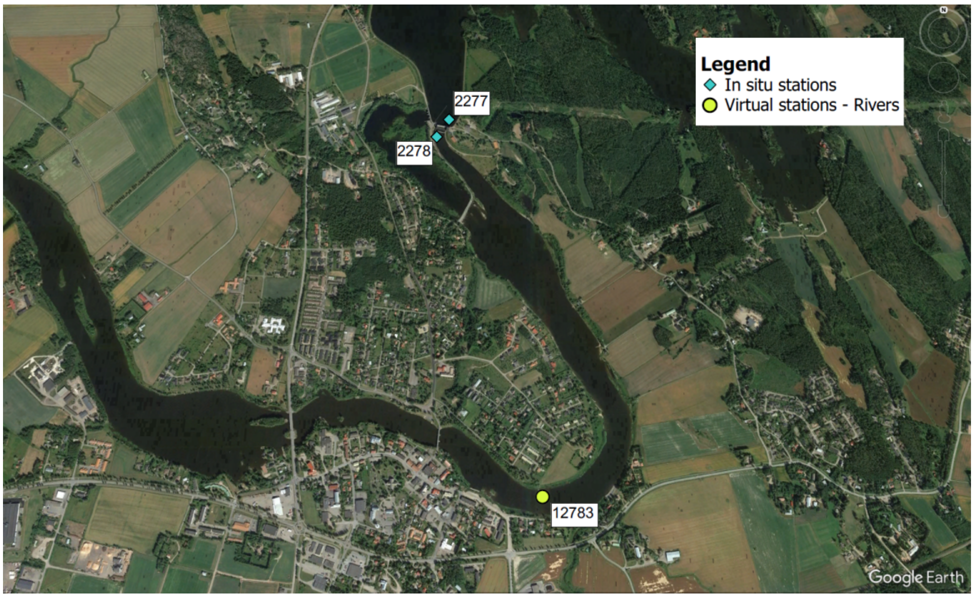


Figure 4.23: Virtual station 12783 and in-situ stations 2278 and 2277

The Figure 4.24 is a zoom of the Figure 4.23. It is possible to see the of the dam and the power plant.

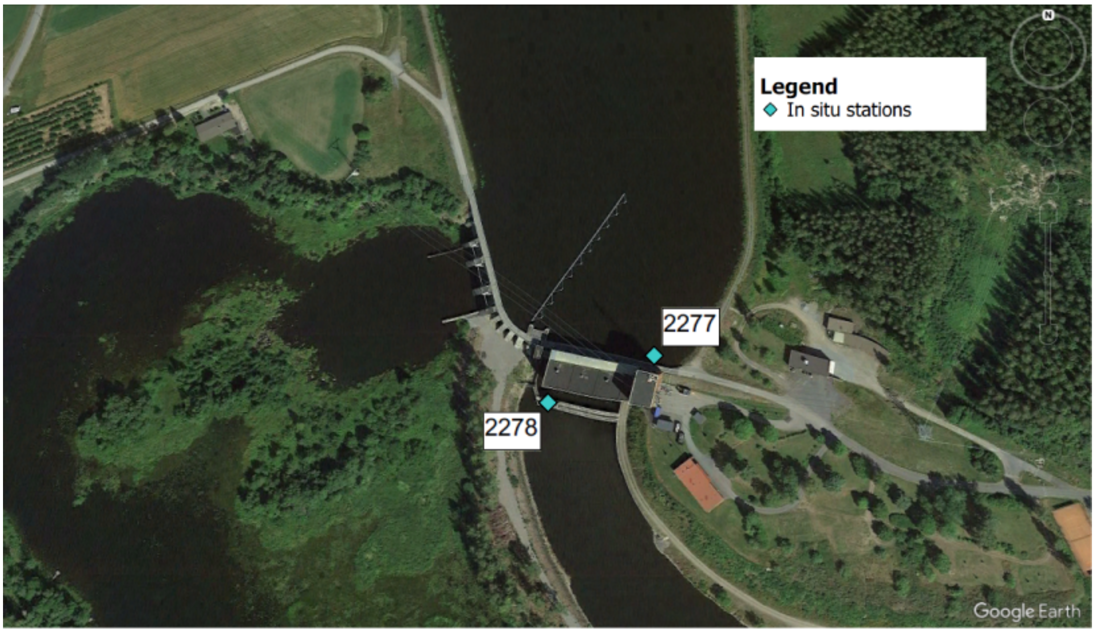
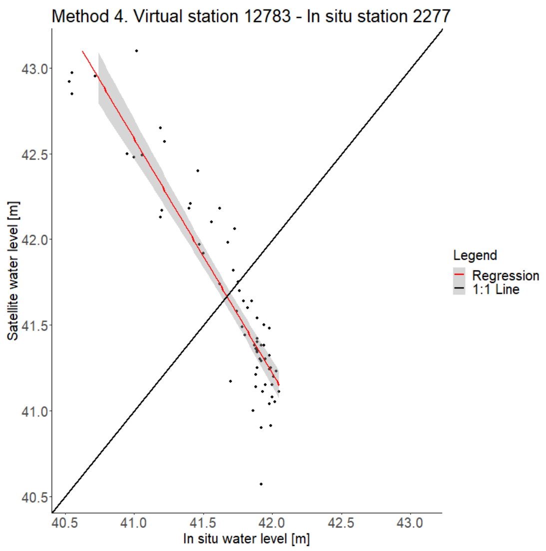


Figure 4.24: In-situ stations 2277 and 2278 (Google Earth)

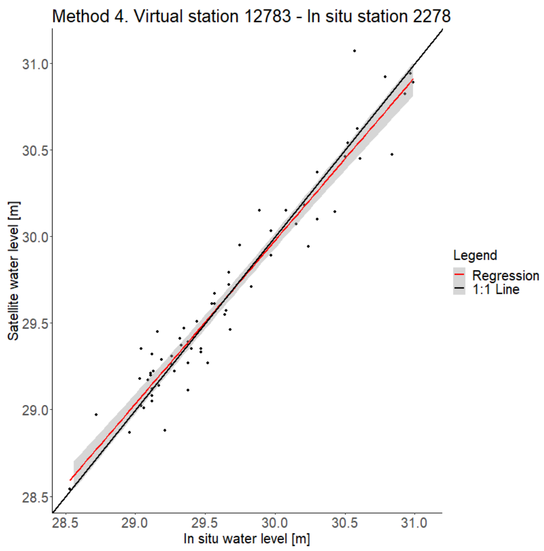
The Figure 4.25 is a picture of the dam and the power plant. It has reported here to have more comprehension.

Figure 4.25: Location of the in-situ stations 2277 and 2278

The two in-situ stations, 2277 and 2278, have different behaviour respect one each other and respect the virtual station 12783.



(a) Couple 12783-2277



(b) Couple 12783-2278

Figure 4.27: Scatter plot of the couple 12783-2277 and 12783-2278 (Method 4)

The Figure 4.26 show the statistical indicators of the two couples.

Virtual station	In situ station	Number of observations	R4	RMSE4	NSE4
12783	2278	64	0.97	0.16	0.93
12783	2277	65	-0.92	0.97	-4.96

Figure 4.26: Number of observations and statistical indicators of the couples 12783-2277 and 12783-2278

The linear correlation of the virtual station 12783 with the 2278 is positive and very strong. The Rmse of method 4 is very low. With the 2277, R is resulted negative and near -1. This means that there is a strong linear negative correlation. The growth of one of the two water leves corresponds to the decrease of the other one. In this case if water level upstream the dam increase consequently the water level downstream decrease. The cause is the hydroelectric power plant. The in-situ station 2277 is upstream the dam. But the RMSE is high, it is near to 1.

The scatter plot showed in the Figures **??** is the one of the couples 12783 and 2277. It is possible to understand what means to have a negative correlation. The 1:1 line has negative inclination instead the previous examples where the inclination were positive.

The virtual station 12783 have good agreement also with the in-situ station 2270. That one is showed in the Figure 4.28

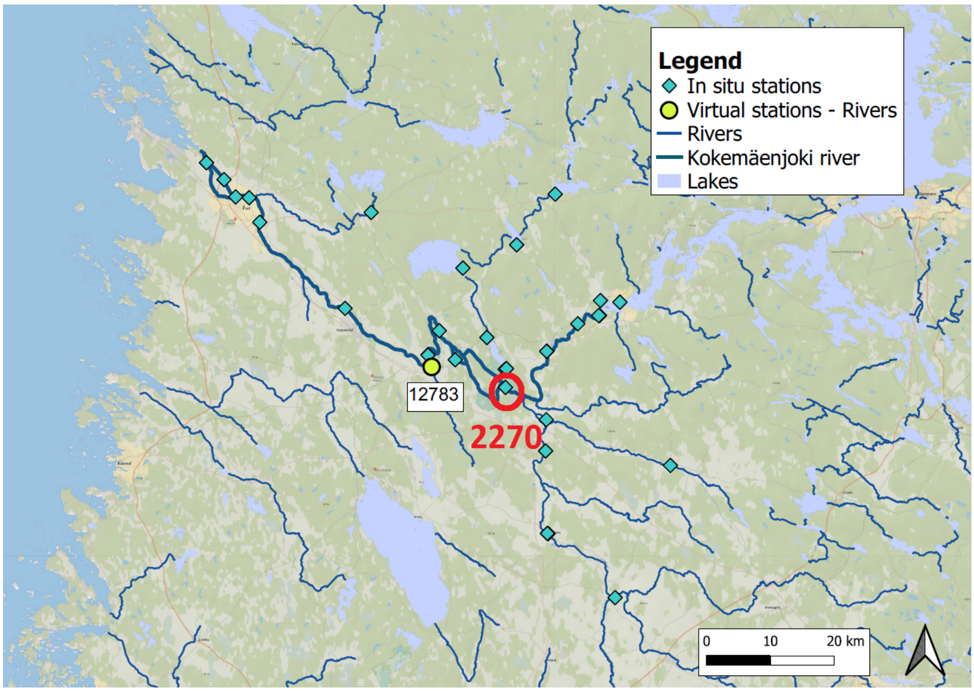


Figure 4.28: Location of the in-situ station 2270

The indicators of the method 4, referred to this couples, are Pearson Coefficient equal to 0.95 and RMSE equal to 0.27. Even if between them there is the dam mentioned in the previous example and also some curves of the channel, the RMSE is very high. Probably the value of RMSE not vert low is caused by these disturbs present between the two stations.

The virtual station have good performance also with the in-situ station 3744 (R equal 0.94 and RMSE equal to 0.19 , both for the method 4). This in-situ station is shown in the Figure 4.29 4.30

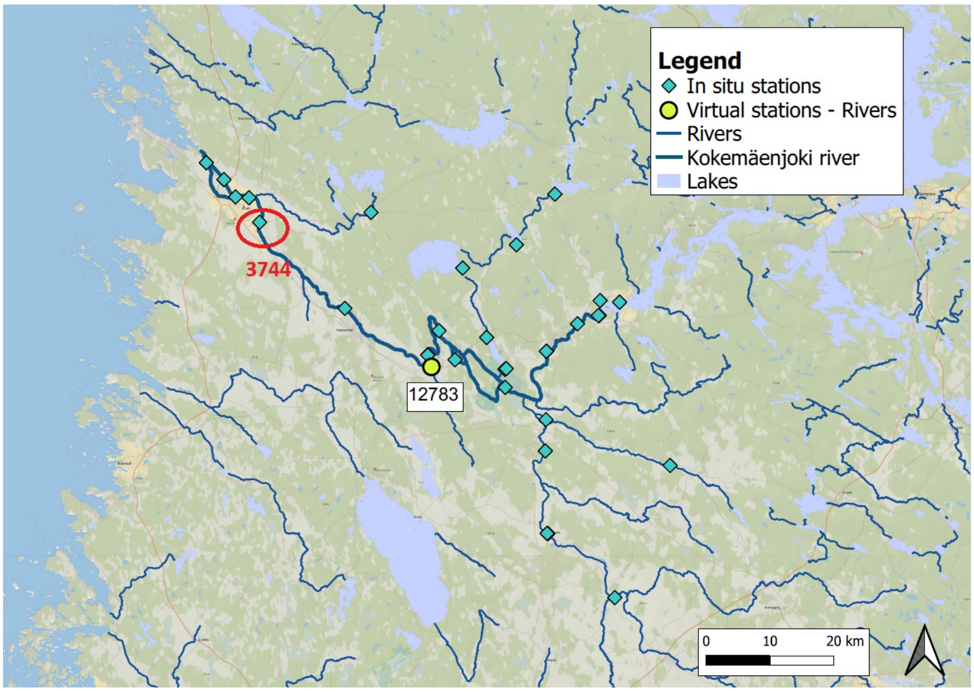


Figure 4.29: Location of the in-situ station 3744

Figure 4.30: In-situ station 3744

The last couple that will be shown for regarding the station 12783 is the one made up with the in-situ station 2253 (Figure 4.31). The R and the RMSE, regarding the method 4, are 0.94 and 0.2. The agreement is high. This in-situ station is near and downstream a dam.

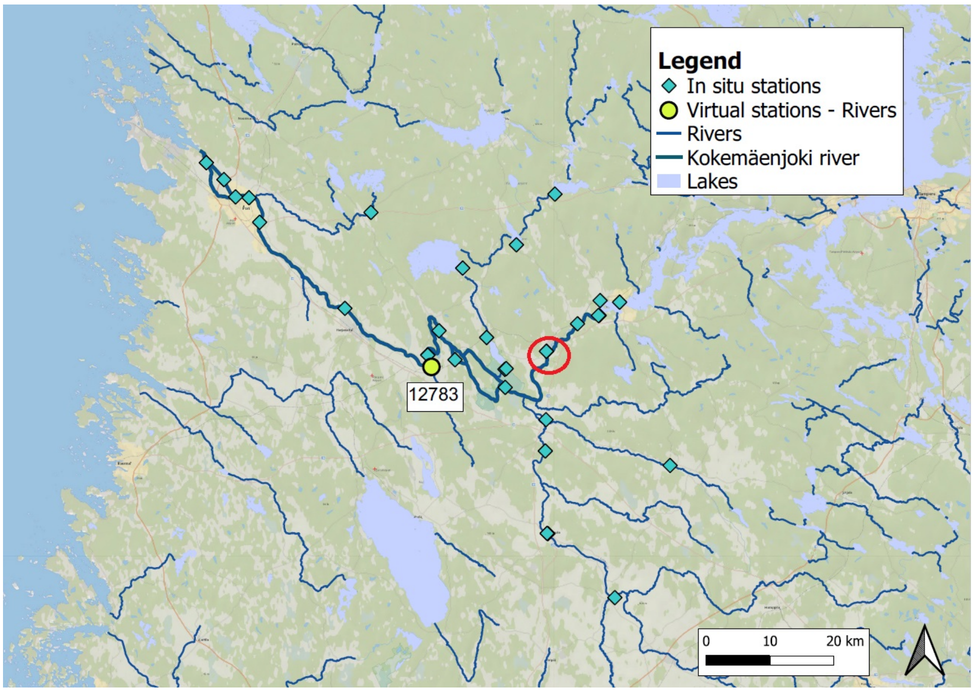


Figure 4.31: Location of the in-situ station 2253

It is necessary to underline that these dam and power plant are upstream respect the dam and power plant cited during the explanation of the couples made up with the virtual station 12783 and the in-situ station 2277 and 2278 (the dam and the power plant of Kolsin), Figure 4.32. Therefore the virtual station 12783 is downstream to both the dam. The good agreement found for the couple 12783-2277 and the couple 12783-2253 could means that the regulation of the flow mad up from the power plants are the same. Thus, the water levels have the same trends. This explanation could be reinforced also considering the couple with the in-situ station 2250 (R4 equal to 0.88 and RMSE4 equal to 0.38). Also this in-situ station in near and downstream to a power plant.

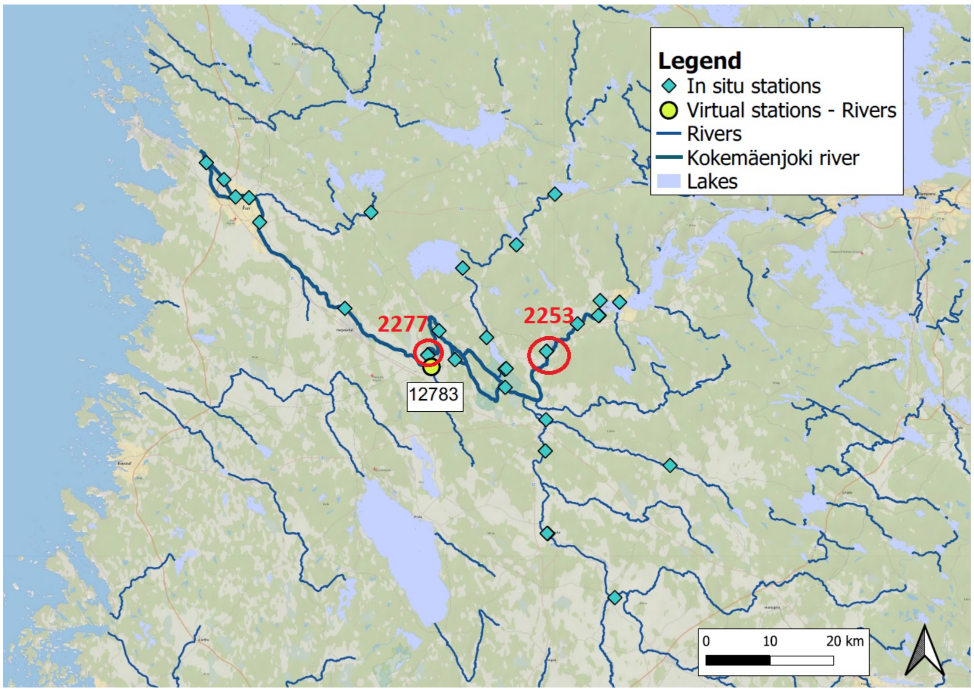


Figure 4.32: Location of the in-situ station 2277 and 2253

Downstream the virtual station 1783 there is the last one examinated in this thesis, the 12784. This virtual station and the in-situ station hydraulically connected are shown in the Figure 4.33

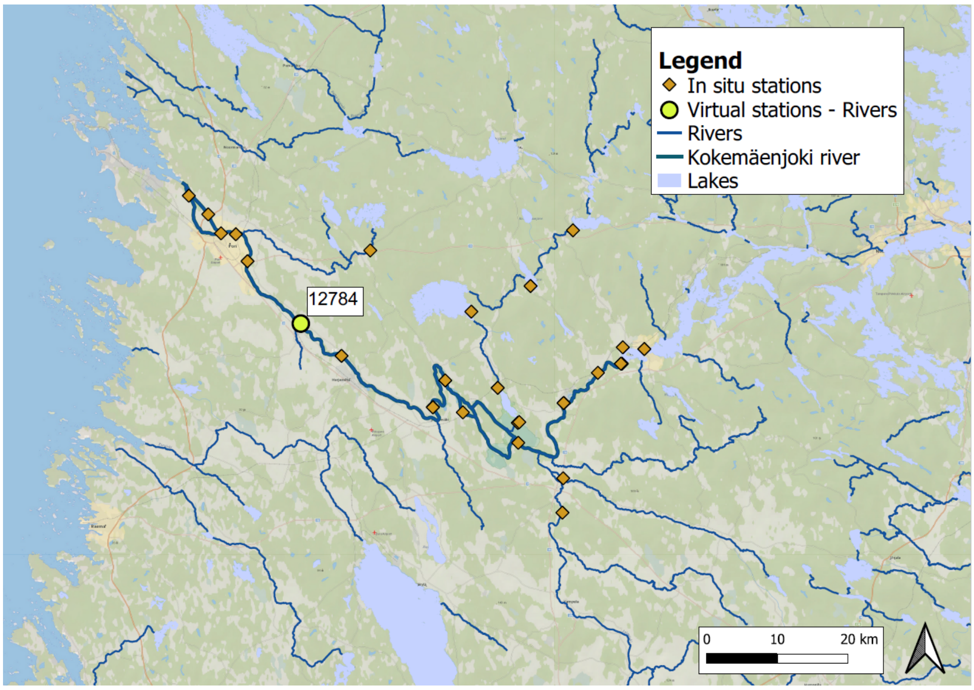


Figure 4.33: Virtual station 12784 and in-situ stations related to the radius of 50 km

The couples with the higher agreement are the two with the in-situ station 2250, 2253 and 2270. Both this two in-situ station are near and downstream to three different power plants. Also the virtual station 12784 is downstream respect the power plant where is located the in-situ station 2279. To have more comprehension see the Figure 4.34.The statistical indicators are shown in the Figure 4.35. This situation could be explained as the case of the virtual station and the in-situ stations downstream the power plants.

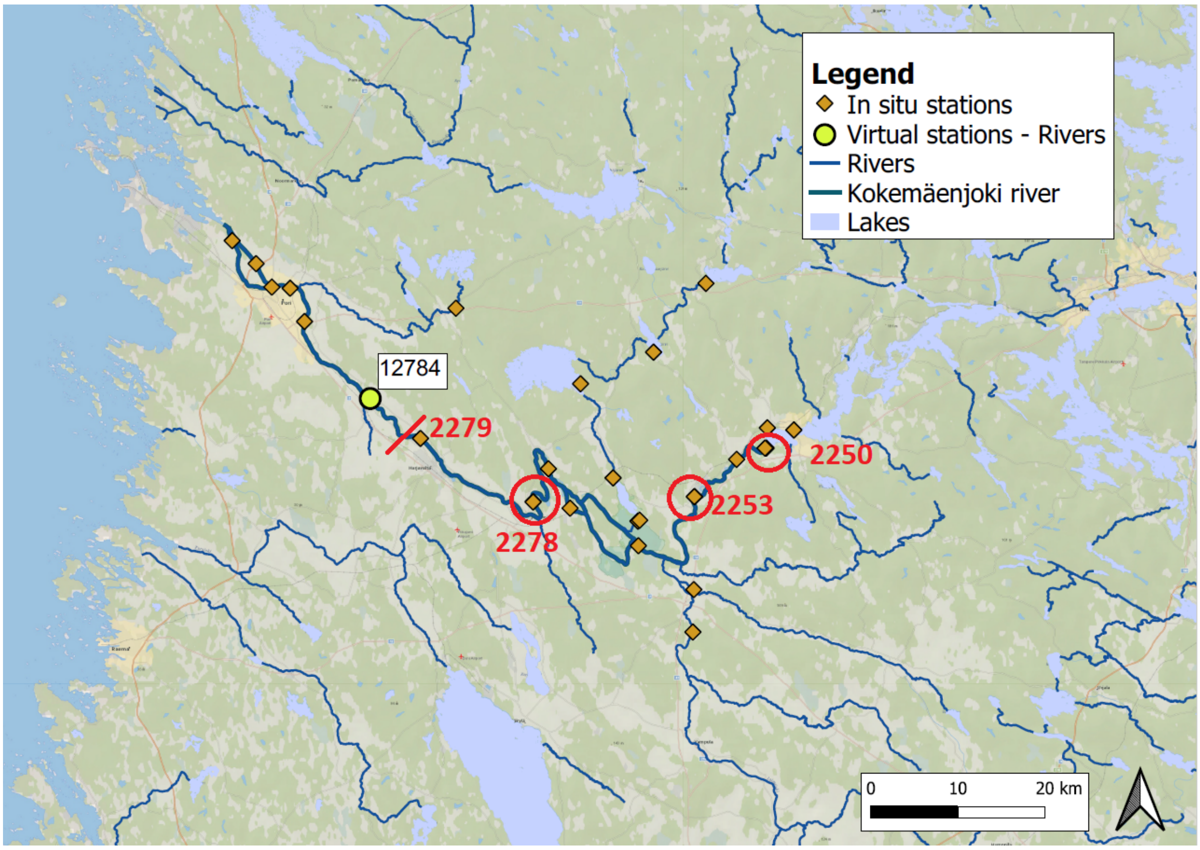
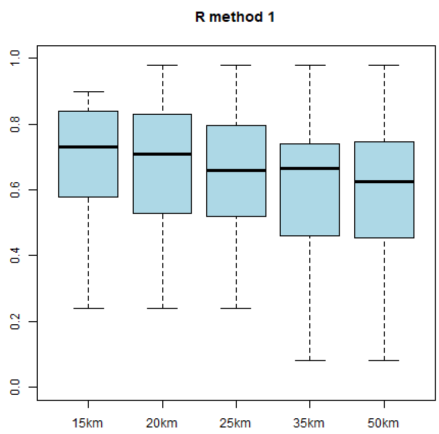


Figure 4.34: Location of the in-situ stations 2279, 2278, 2253 and 2250

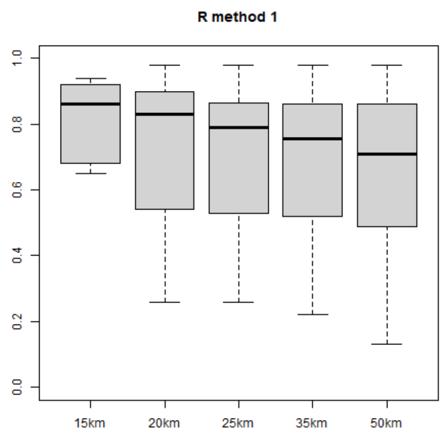
Identifier_VS	In situ station	Number of observations	R1	RMSE1	R2	RMSE2	R4	RMSE4	NSE4
12784	2250	69	0.84	0.38	0.95	0.37	0.95	0.27	0.89
12784	2253	69	0.82	0.42	0.94	0.37	0.94	0.37	0.56
12784	2278	69	0.74	0.47	0.88	0.45	0.88	0.4	0.71

Figure 4.35: Statistical indicators of the couple 12784-2250, 12784-2253 and 12784-2278

It has been assessed for rivers virtual stations a further analysis based on seasonality. For this work it has been decided to start with only two seasons. Considering that, it is possible to divide the year in two periods that in this thesis are called: summer and winter.



(a) Annual period



(b) Winter period

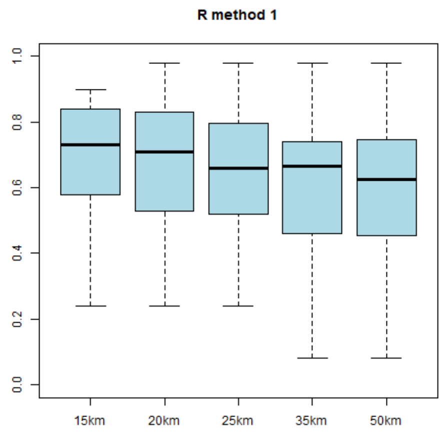
Figure 4.36: Box plot of method 1 related to the annual and the winter periods

For the north part of Finland the summer could be the period from June to September and the winter from October to May. For south part of Finland the summer could be the period from May to October and the winter from November to April. This rough subdivision as been suggested from the SYKE to make a first seasonal analysis. The virtual station 12783 belongs to the North and the virtual stations 12783 and 12784 belong to the South part of the country.

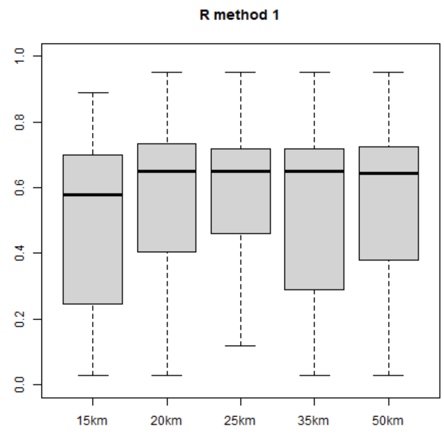
After the subdivision of the dataset in the two groups, it has been calculated the statistical indicators. It is necessary to mark that the individuals groups have a lower number of observation than the original dataset because this last one has been subdivided.

The Figures 4.36, 4.37, 4.38 illustrate the comparison, for the method 1, between each possible couple of the period considered. The median of the annual and the winter period decrease with the increasing of the radius and the winter period performs better than the annual period. The boxplot of the summer period are more or less stable to the same median for all the radius. The boxplot of the winter have the median higher than the one of the summer period. Therefore the winter period performs better than the summer and the annual.

In the Figures 4.39, 4.40 and 4.41 are shown the boxplot of the period considered regarding the method 4. As previously mentioned, the method 4 for annual period gives higher agreement than the method 1. For the winter period has been found better performances than the summer period and respect the annual period more or less the behaviour is the same.

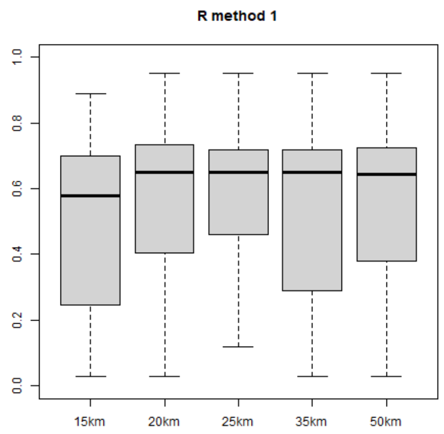


(a) Annual period

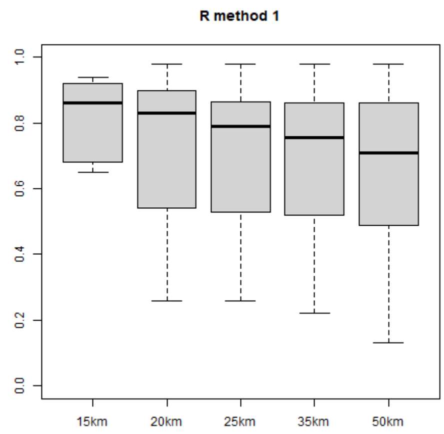


(b) Summer period

Figure 4.37: Box plot of method 1 related to annual and summer periods

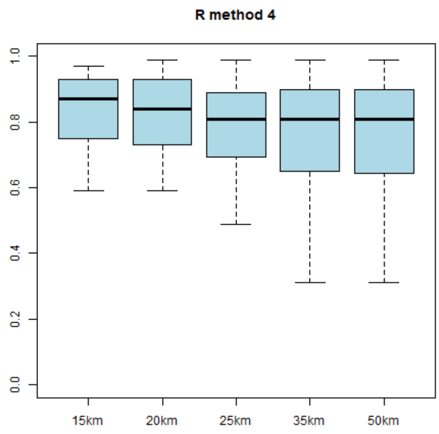


(a) Summer period

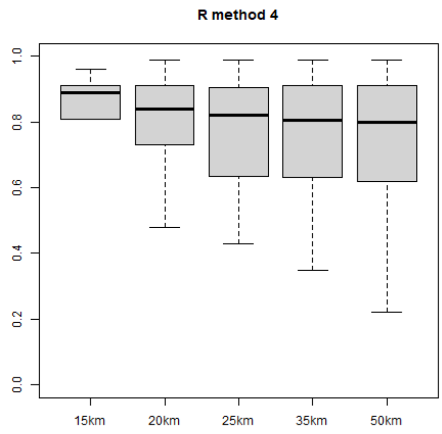


(b) Winter period

Figure 4.38: Box plot of method 1 related to summer and winter periods

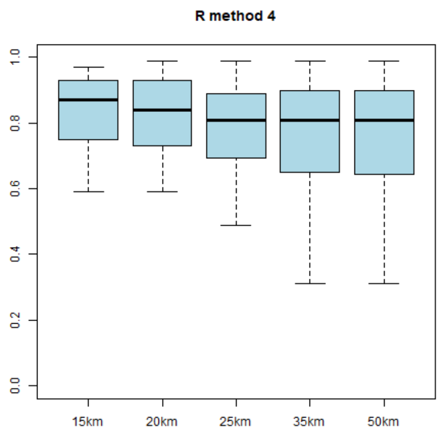


(a) Annual period

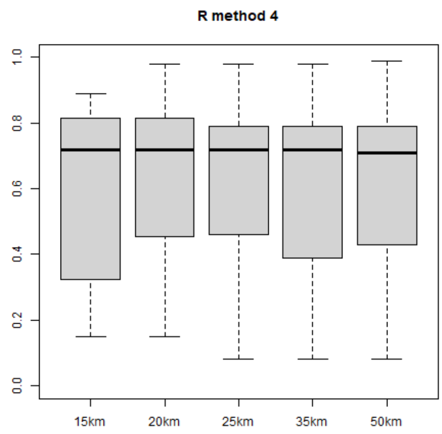


(b) Winter period

Figure 4.39: Box plot of method 4 related to the annual and the winter periods

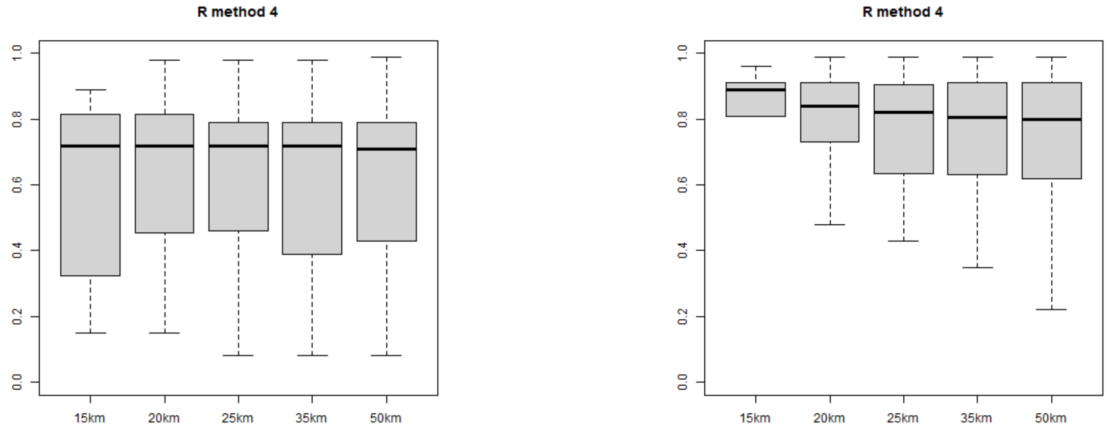


(a) Annual period



(b) Summer period

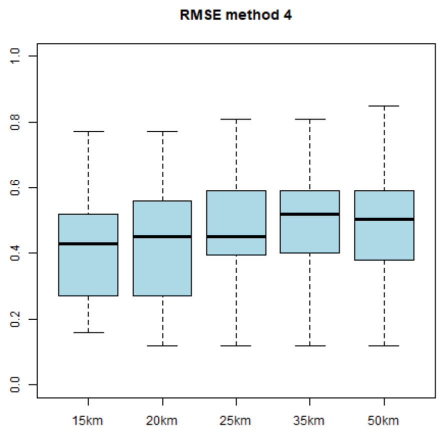
Figure 4.40: Box plot of method 4 related to annual and summer periods



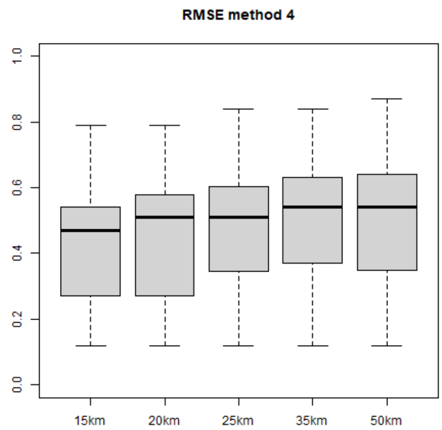
(a) Summer period (b) Winter period

Figure 4.41: Box plot of method 4 related to summer and winter periods

In the Figures 4.42, 4.43 and 4.44 are shown the boxplot of the RMSE of the method 4 referred to the period considered regarding the method 4. The RMSE4 for annual and winter periods increase if the radius increase. RMSE4 of summer period is less than the winter and annual period.

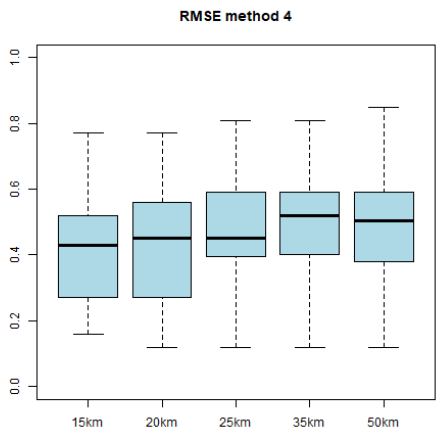


(a) Annual period

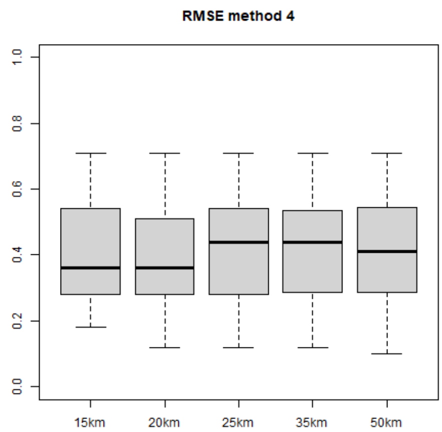


(b) Winter period

Figure 4.42: Box plot of method 4 related to the annual and the winter periods

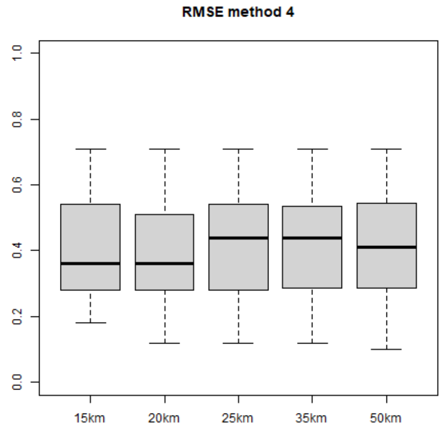


(a) Annual period

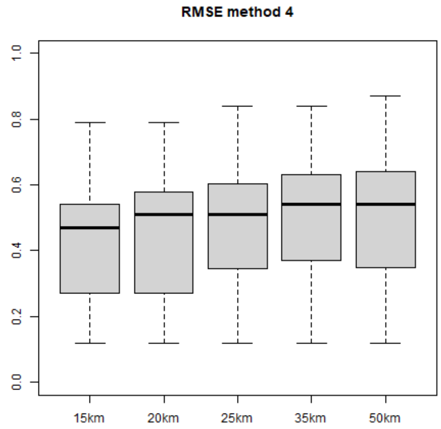


(b) Summer period

Figure 4.43: Box plot of method 4 related to annual and summer periods



(a) Summer period



(b) Winter period

Figure 4.44: Box plot of method 4 related to summer and winter periods

4.2. Lakes

The original dataset of lakes altimetric data is described in the section 2.2.3. It has been retrieved from the site Theia. The virtual stations analyzed are four and they are located in different part of Finland territory. Only the virtual station Vanajanselka has a monomission record that is referred to Sentinel 3A. The others lakes altimetric data analyzed are multimission or monomission but without Sentinel 3A data.

Finnish lakes have a particular conformation and shape. They are not single water bodies. They are fragmented in many parts as described in the section 2.1. For this reason, the research of the couples has not been done as the rivers case. Then, for every virtual station it has been founded the couples referred only to the radius of 50 km because lower radius did not have so many couples. From these groups it has been discarded manually the couples not hydraulically connected and also the couples hydraulically connected that fall outside the principal part of the lake. It has been calculated the statistical indicators and it has been discarded the couples with low agreement (R low and RMSE high). Therefore, it has been considered only the couples with the in-situ station located in the lake. Many of these in-situ stations are located near dams and power plants. The records analyzed were filtered by outliers as the rivers virtual stations validation case.

The virtual stations, under the process of validation, are located in the followings lakes:

- Inarinjärvi (1). 307 observations available of the satellites Saral, Sentinel 3A, Envisat and Cryosat 2;

- Kemijärvi (2). 103 observations available of the satellite Cryosat 2 ;
- Oulujärvi (3). 212 observations available of the satellites Jason 2, Jason 3, Saral and Sentinel 3A;
- Vanajavesi (4) or also named Vanajanselka. 78 observations available of the satellites Sentinel 3A.

The numbers reported in the list are sometimes used at the place of the finnish names. Their locations are shown in the Figure 2.10.

In the Figures 4.45 and 4.46, referred to the virtual station Inarinjärvi, it is possible to see the in-situ stations respectively in the radius of 50 km (hydraulically connected and not) and the ones that are simultaneously hydrologically connected and have good performances.

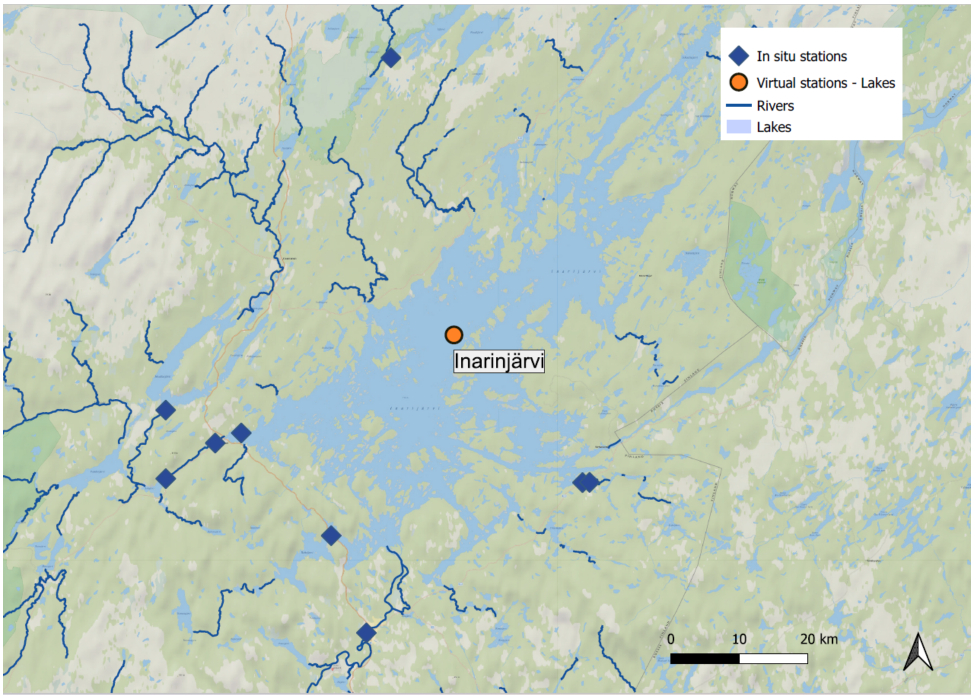


Figure 4.45: Virtual station Inarinjärvi and in-situ stations in the radius of 50 km

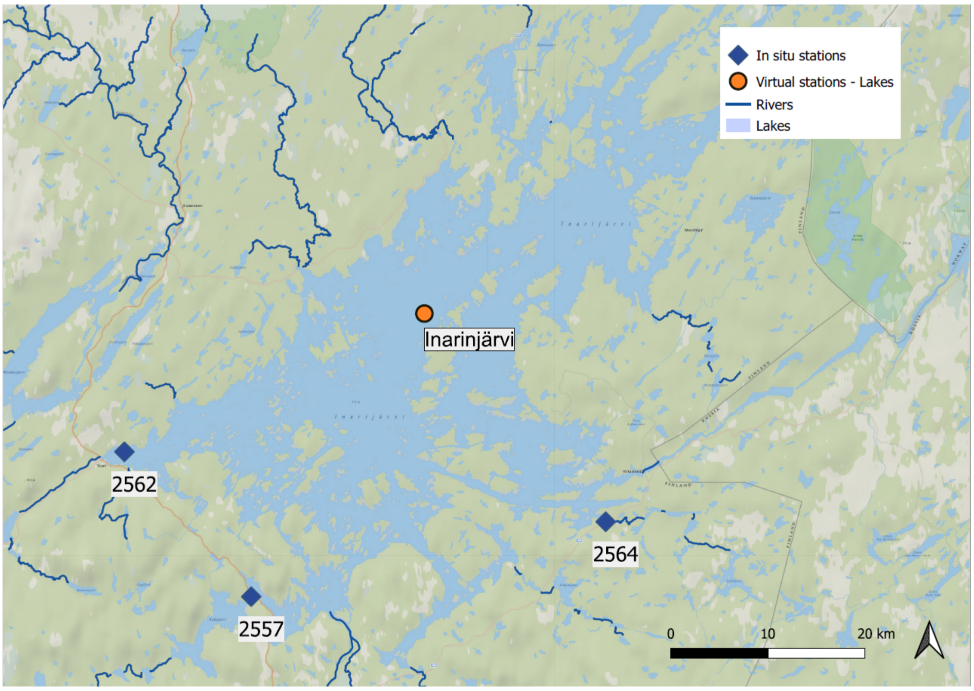


Figure 4.46: Virtual station Inarinjärvi and in-situ stations selected

In the next Figures are shown the same of 4.45 and 4.46 but referring to the other virtual stations: Kemijärvi (Figures 4.47 and 4.48), Oulujärvi (Figures 4.49 and 4.50) and Vanajavesi (Figures 4.51 and 4.52).

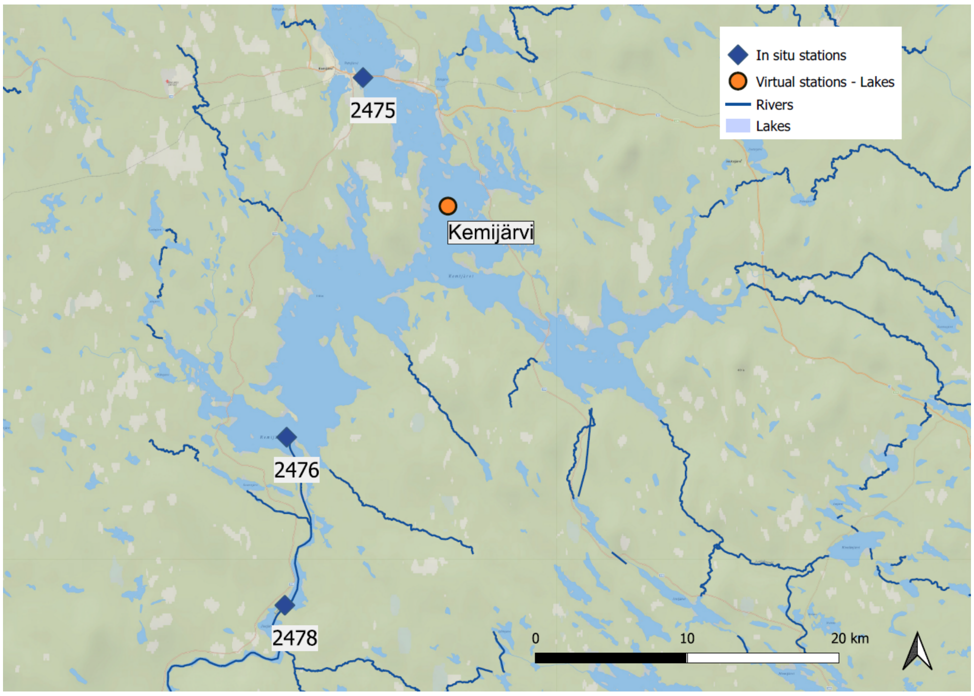


Figure 4.47: Virtual station Kemijärvi and in-situ stations in the radius of 50 km

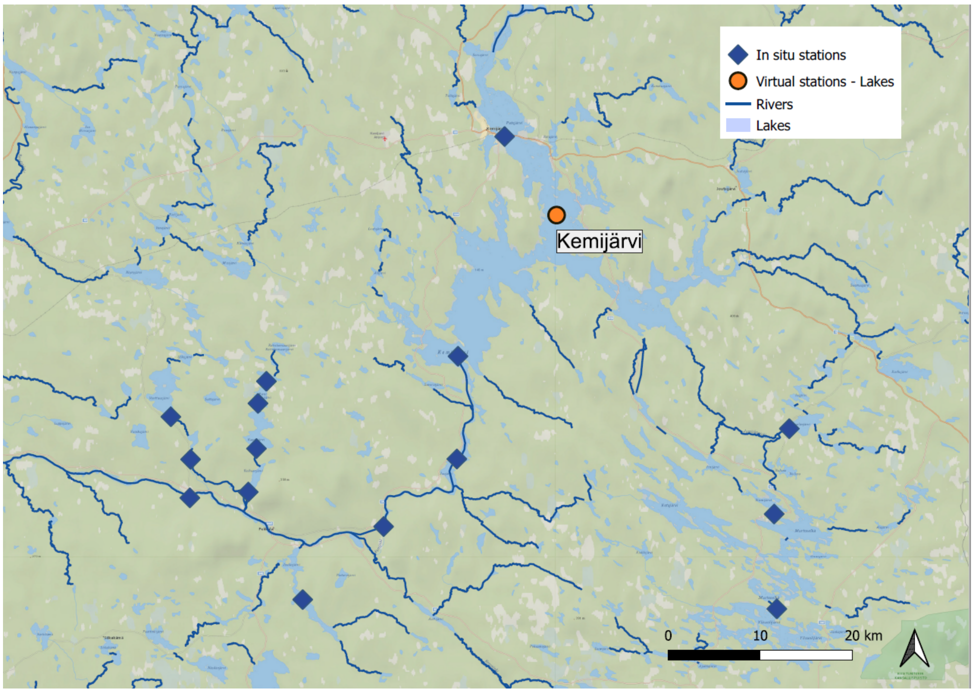


Figure 4.48: Virtual station Kemijärvi and in-situ stations selected

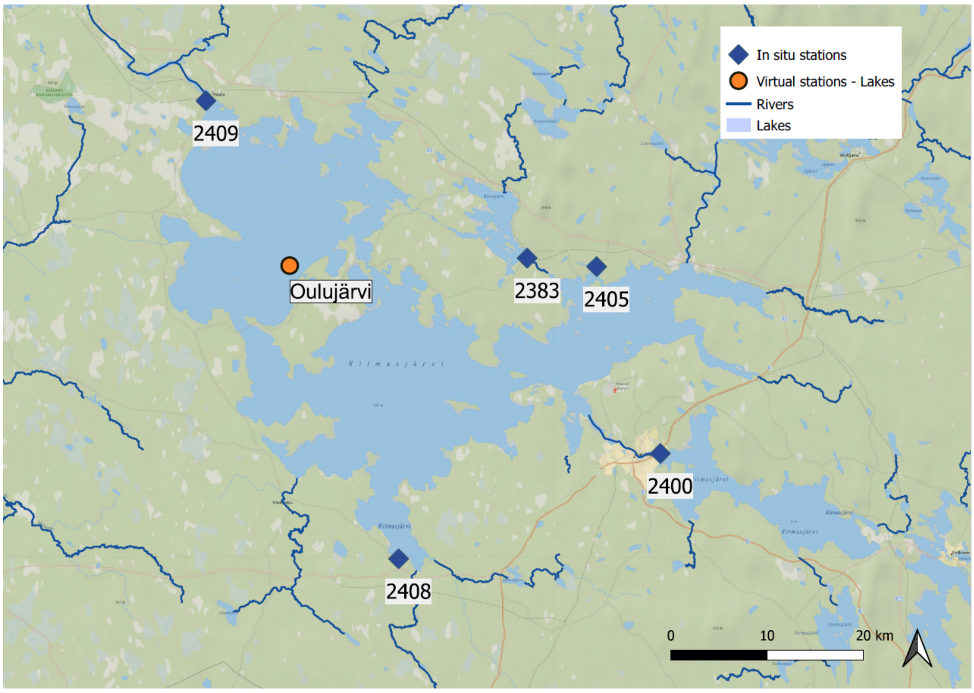


Figure 4.49: Virtual station Oulujärvi and in-situ stations in the radius of 50 km

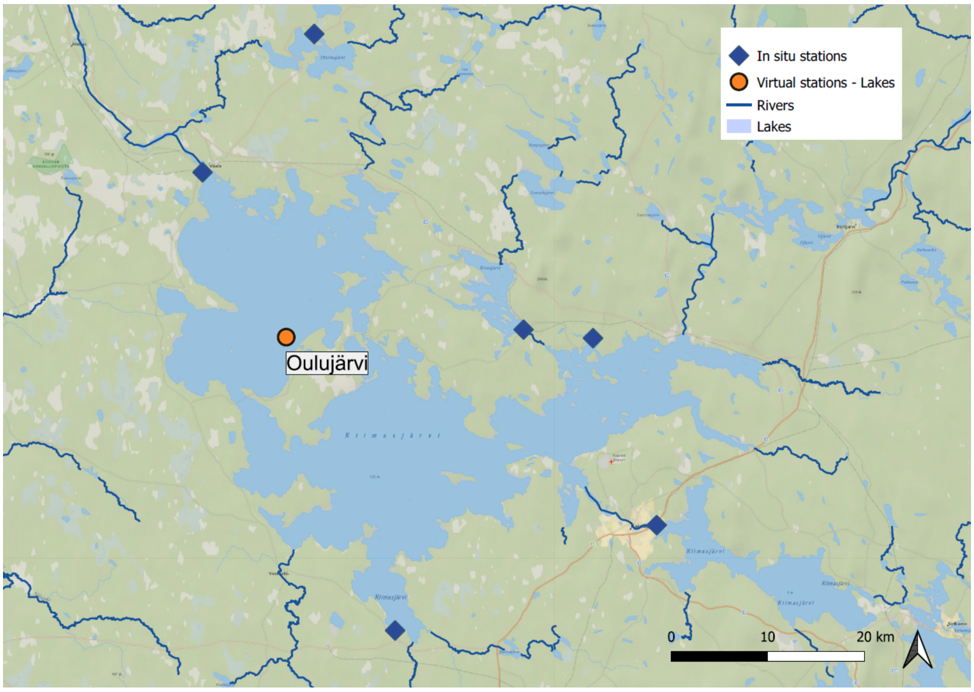


Figure 4.50: Virtual station Oulujärvi and in-situ stations with the best performances

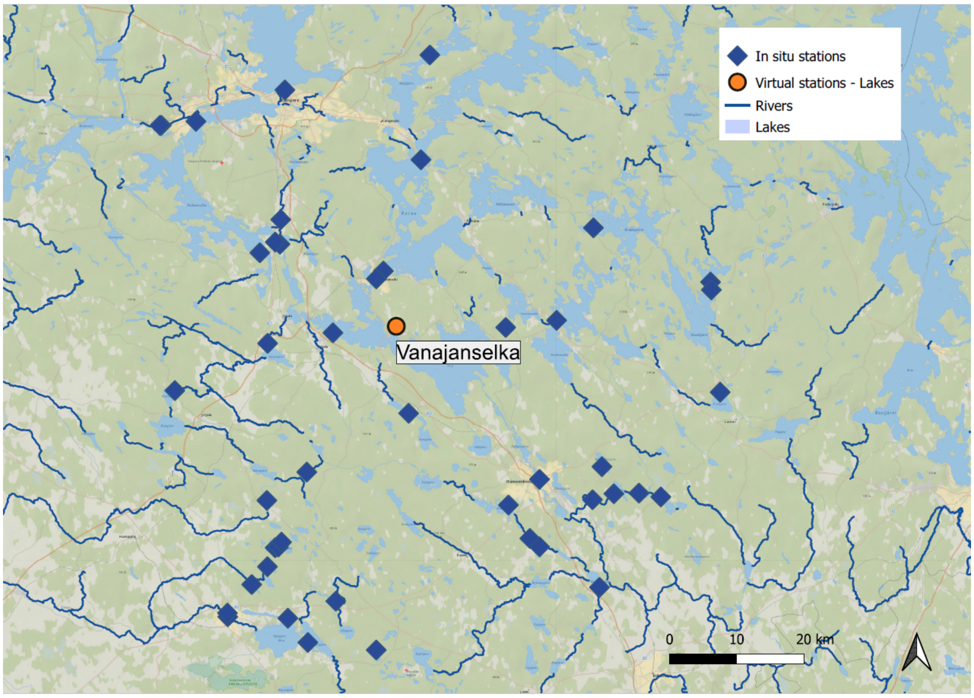


Figure 4.51: Virtual station Vanajavesi and in-situ stations in the radius of 50 km

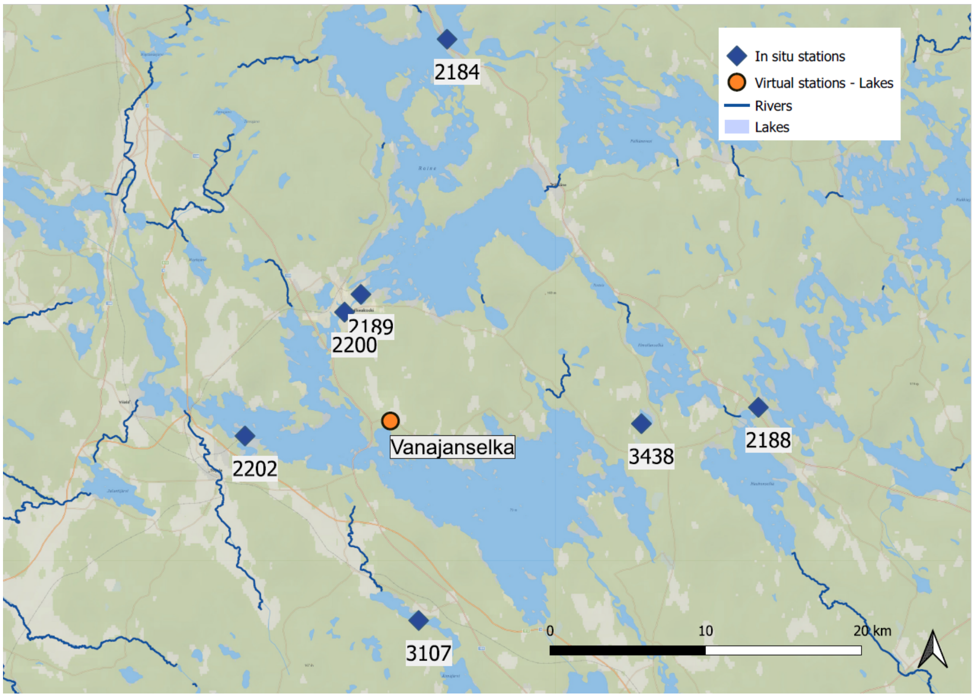


Figure 4.52: Virtual station Vanajavesi and in-situ stations selected

In the Figure 4.53 are reported the couples with the best indicators in terms of R of method 4. In particular they are ordered in descending way, from the highest value of R4 to the lowest. It has been reported the number of the observations in common between the virtual station and the corresponding in-situ station. For each couples the R of the methods 2 and 3, that are equal, are higher than the R of the method 1. Thus, the agreement improved using the original trends retrieved from the site THEIA respect using temporal differences. The RMSE of the method of the temporal differences (method 1) sometimes is higher and sometimes is lower than the RMSE of the method 2. But considering the indicator related to method 4 (R4, RMSE4 and NSE4) is possible to understand that it is the method that gave the best results in terms of high R and low RMSE.

Identifier_VS	In situ station	number of observations	R 1	RMSE 1	R 2	RMSE 2	R 4	RMSE 4	NSE 4
3	2405	212	0.86	0.14	0.94	0.18	0.94	0.15	0.88
3	2408	212	0.85	0.14	0.94	0.17	0.94	0.15	0.88
3	2409	212	0.85	0.14	0.94	0.18	0.94	0.15	0.88
4	2184	75	0.64	0.14	0.9	0.17	0.9	0.13	0.75
4	2188	76	0.61	0.14	0.9	0.17	0.9	0.13	0.8
4	2189	75	0.65	0.13	0.9	0.17	0.9	0.12	0.77
2	2475	103	0.69	1.09	0.89	0.84	0.89	0.81	0.43
1	2562	296	0.69	0.27	0.87	0.34	0.87	0.24	0.7
1	2564	296	0.7	0.27	0.87	0.34	0.87	0.24	0.72
2	2476	103	0.65	1.59	0.87	1.25	0.87	1.13	0.74
2	2484	103	0.64	1.24	0.87	1.16	0.87	1.1	-2.11
2	2478	98	0.66	1.69	0.86	1.37	0.86	1.24	0.7
2	2483	100	0.68	1.17	0.86	1	0.86	0.97	-0.24
4	3438	51	0.65	0.13	0.85	0.27	0.85	0.16	0.4

Figure 4.53: Number of observations and statistical indicators of the couples selected related to lakes

As can be seen from the Figure 4.53, it has been found many in-situ station that perform very well in terms of R4 and RMSE4 with the virtual stations 3 and 4. In particular, it has been found the R4 greater or equal to 0.9 and the RMSE less or equal to 0.15. These two station are located in the south part of the country. The virtual station 4 is the one that have only observations related to Sentinel 3A.

For the virtual stations number 1 and 2, Inarinjärvi and Kemijärvi, the best performances for R4 are between 0.85 and 0.89 and the corresponding RMSE4 are between 0.24 and 1.24. Therefore, the R4 indicators are still high but lower than the virtual station 1 and 2,but there is an increasing not negligible of the RMSE4.

The next Figures will show the most important temporal series and the most relevant scatter plot regarding one of the best lakes couple made up by the virtual station 3 (Oulujärvi) and the in-situ station 2405.

In the Figure 4.54 are shown the temporal record of the in-situ station 2405 and the temporal record of the method 2 of the virtual station.

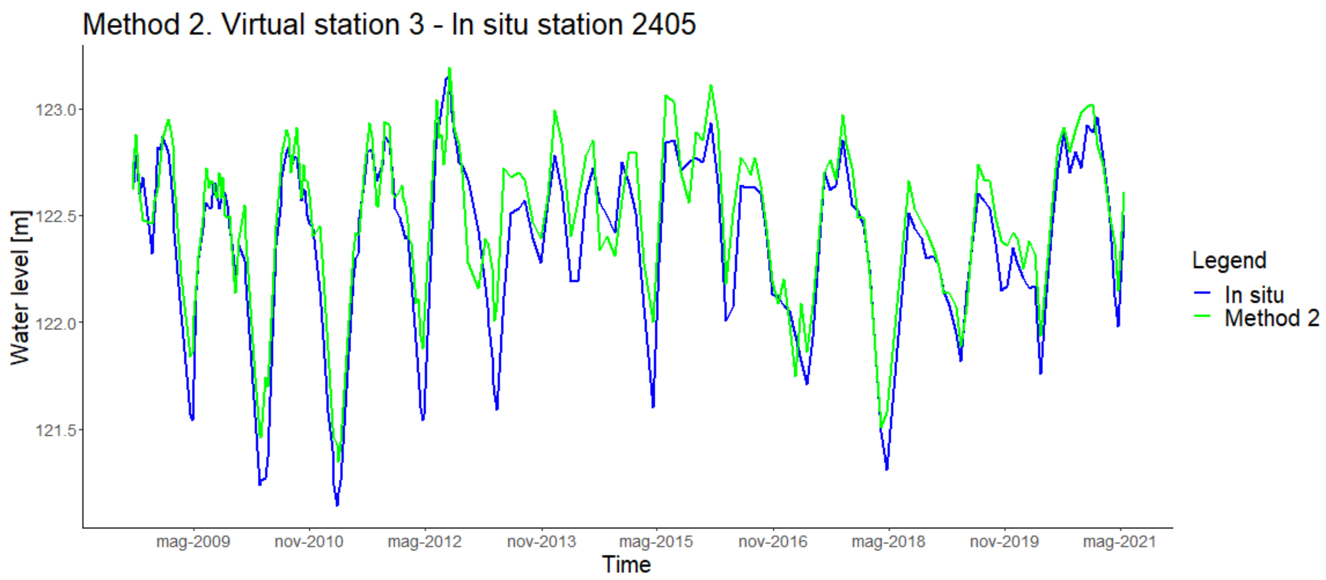


Figure 4.54: Temporal trends of virtual station 3 (method 2) and of in-situ station 2405

In the next Figure, 4.55 are shown the in-situ record and the records related to the calibration with MAE and RMSE.

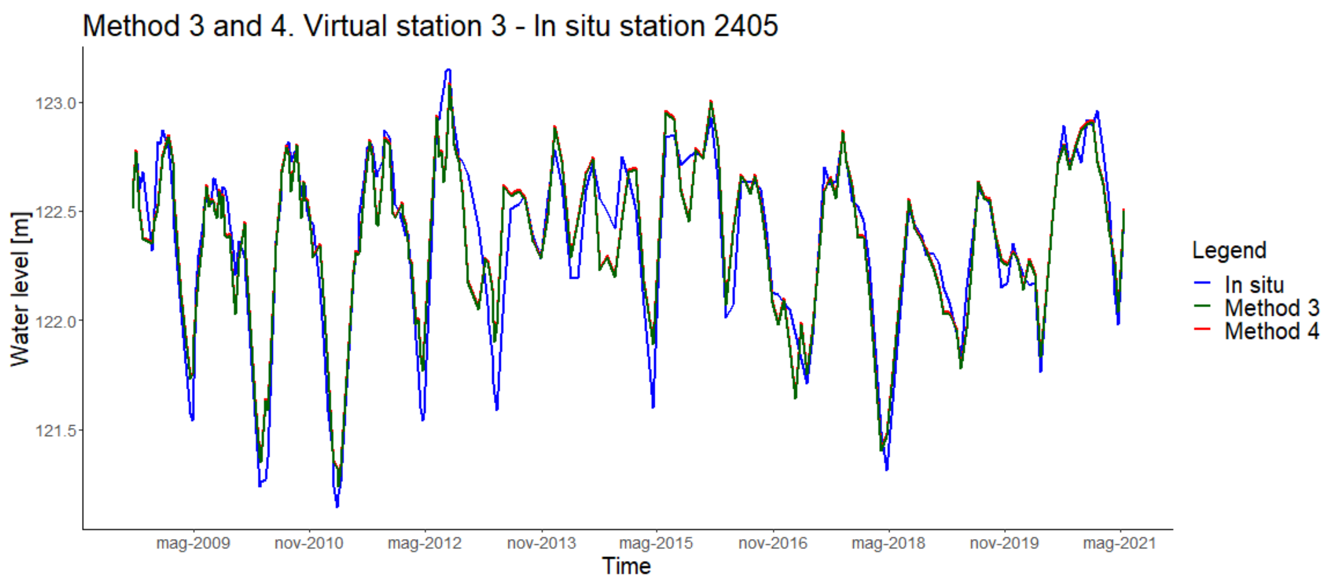


Figure 4.55: Temporal trends of virtual station 3 (method 3 and 4) and of in-situ station 2405

The scatter of the Figure 4.56 plot shows good agreement between of the record calibrated with RMSE (method 4) and the record of the in-situ station.

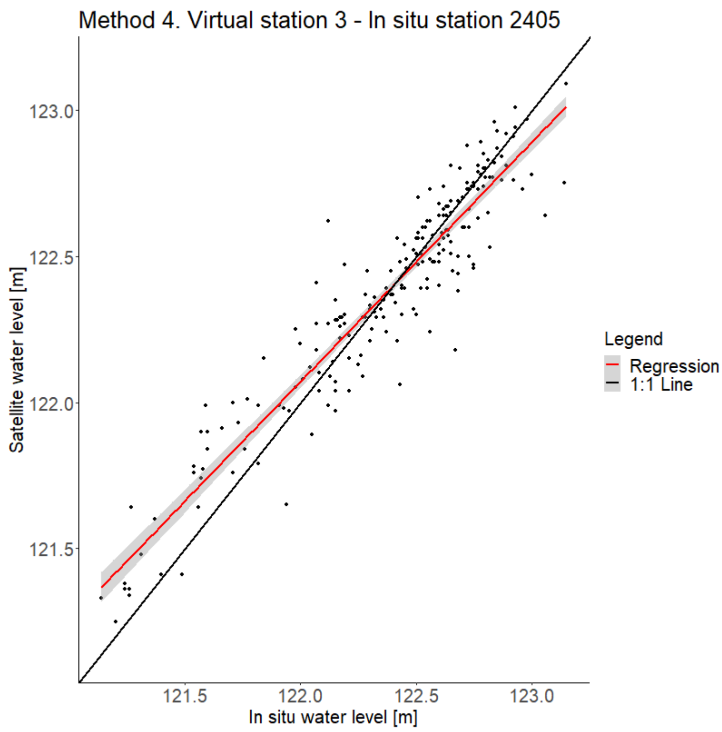


Figure 4.56: Scatter plot of the couple 3-2405 (Method 4)

For the lakes 4, Vanajavesi, it will be shown the temporal trends and the scatter plot referred to the couple made up with the in-situ station 2189 (Figures 4.57 and 4.58

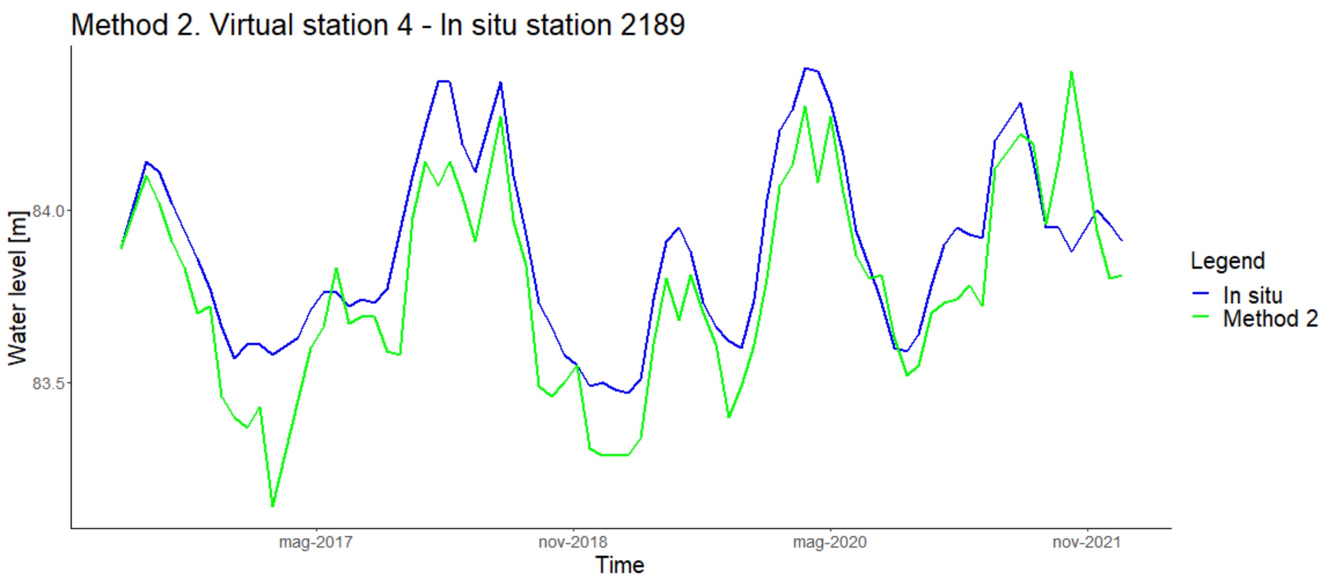


Figure 4.57: Temporal trends of virtual station 4 (method 2) and of in-situ station 2189

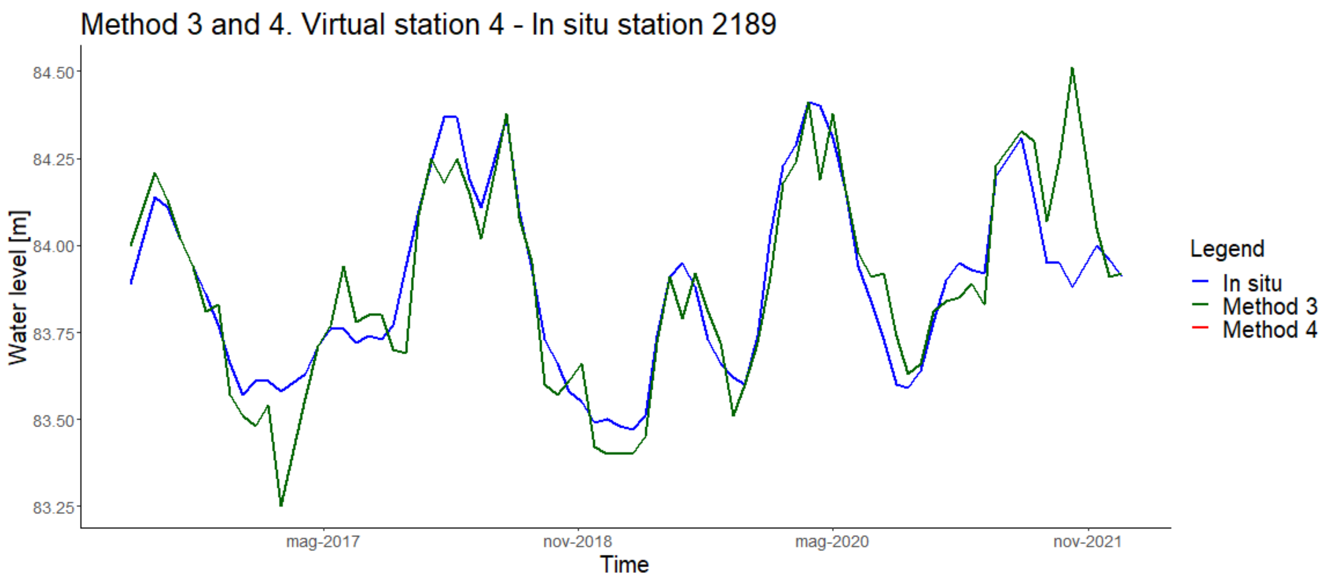


Figure 4.58: Temporal trends of virtual station 4 (method 3 and 4) and of in-situ station 2189

The scatter plot of the Figure 4.59 illustrates the good agreement of the couple 4-2189 related the method 4. The RMSE of the method 4 is 0.12 and the R is 0.9.

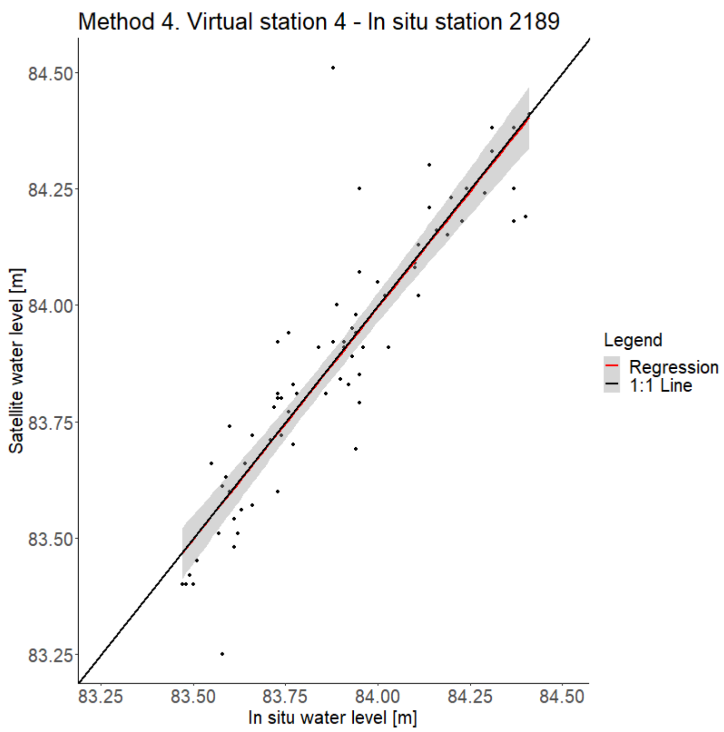


Figure 4.59: Scatter plot of the couple 4-2189 (Method 4)

For the virtual station 2 are shown the graphical results related to the couple made up with the in-situ station 2475 (Figures 4.60)

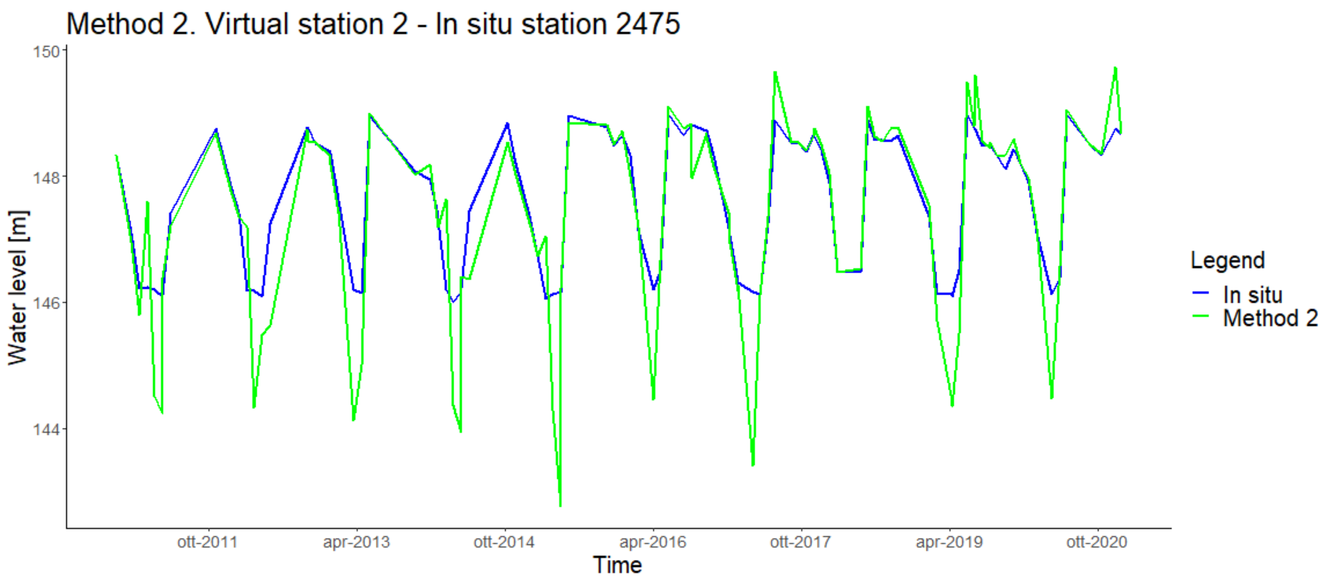


Figure 4.60: Temporal trends of virtual station 2 (method 2) and of in-situ station 2475

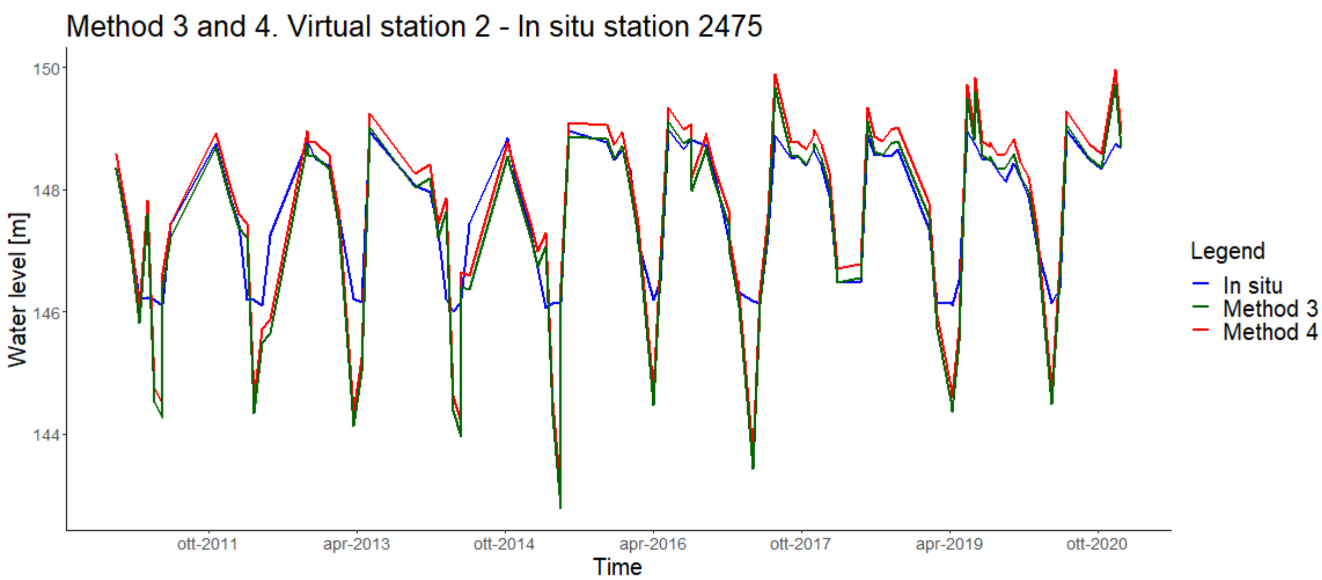


Figure 4.61: Temporal trends of virtual station 2 (method 3) and of in-situ station 2475

These two last Figures show that the satellite periodically underestimate the real water level with a big peack. R4 is high (0.89) but also the RMSE is high (0.81). In fact this error could be linked to the difference at the peack.

For the virtual station 1 are shown the graphical results related to the couple made up with the in-situ station 2562 (Figures 4.62, 4.63, **??**, 4.64)

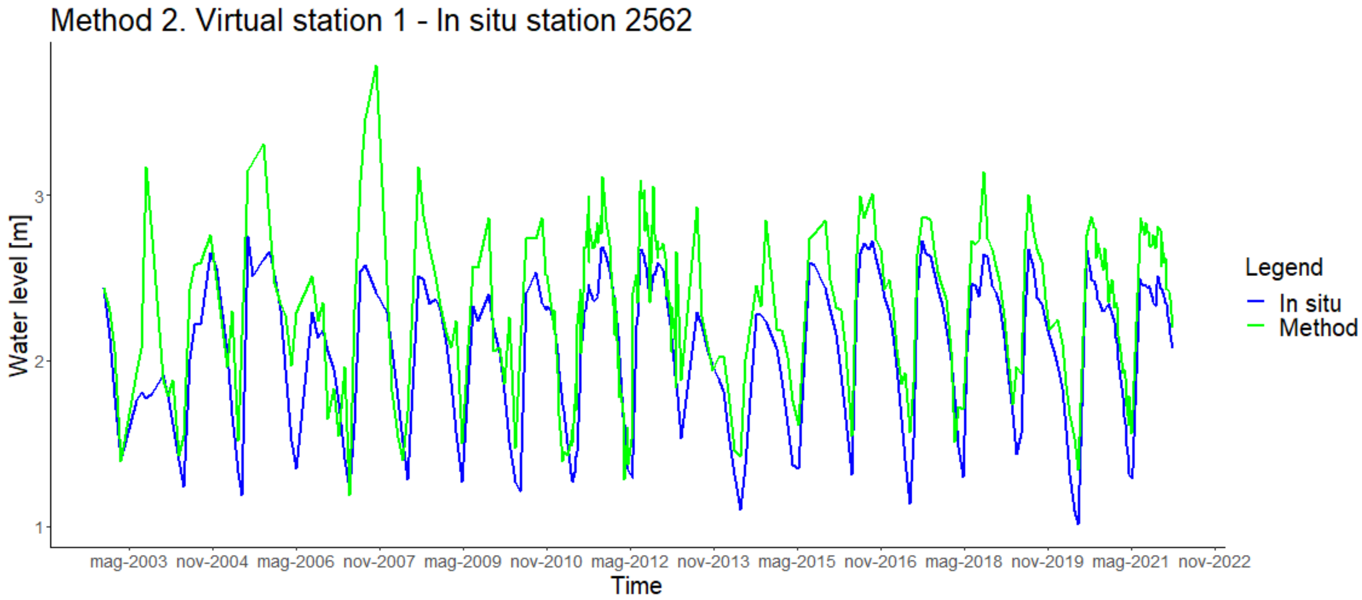


Figure 4.62: Temporal trends of virtual station 1 (method 2) and of in-situ station 2562

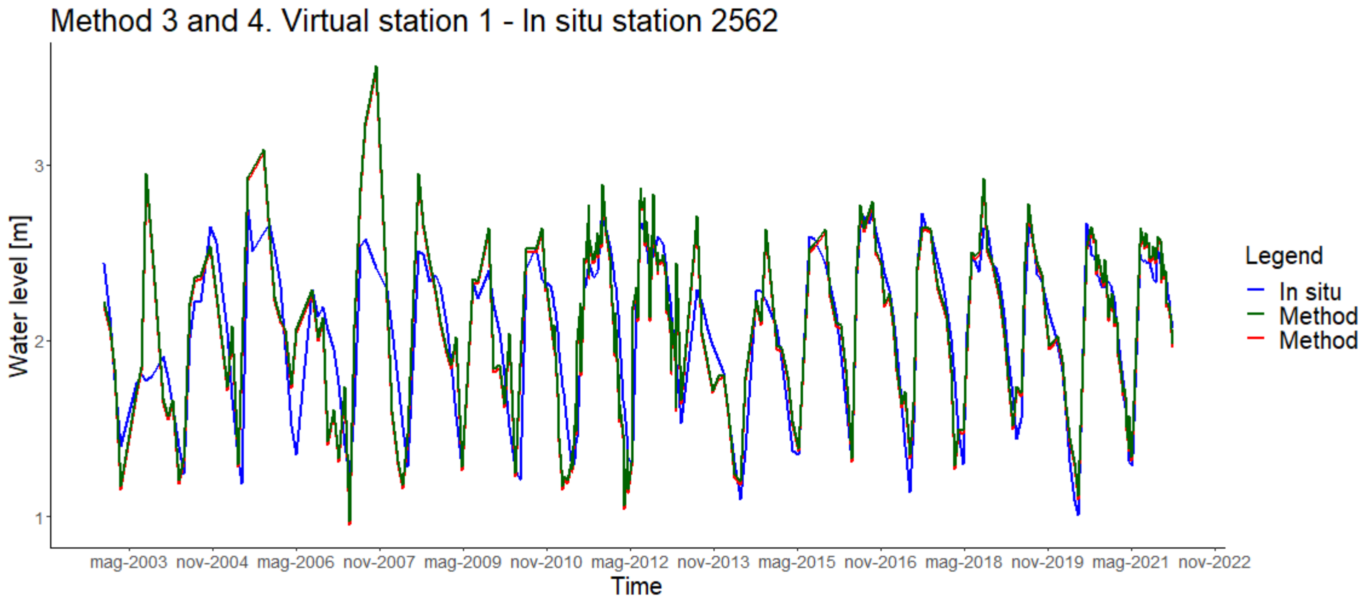


Figure 4.63: Temporal trends of virtual station 1 (method 3 and 4) and of in-situ station 2562

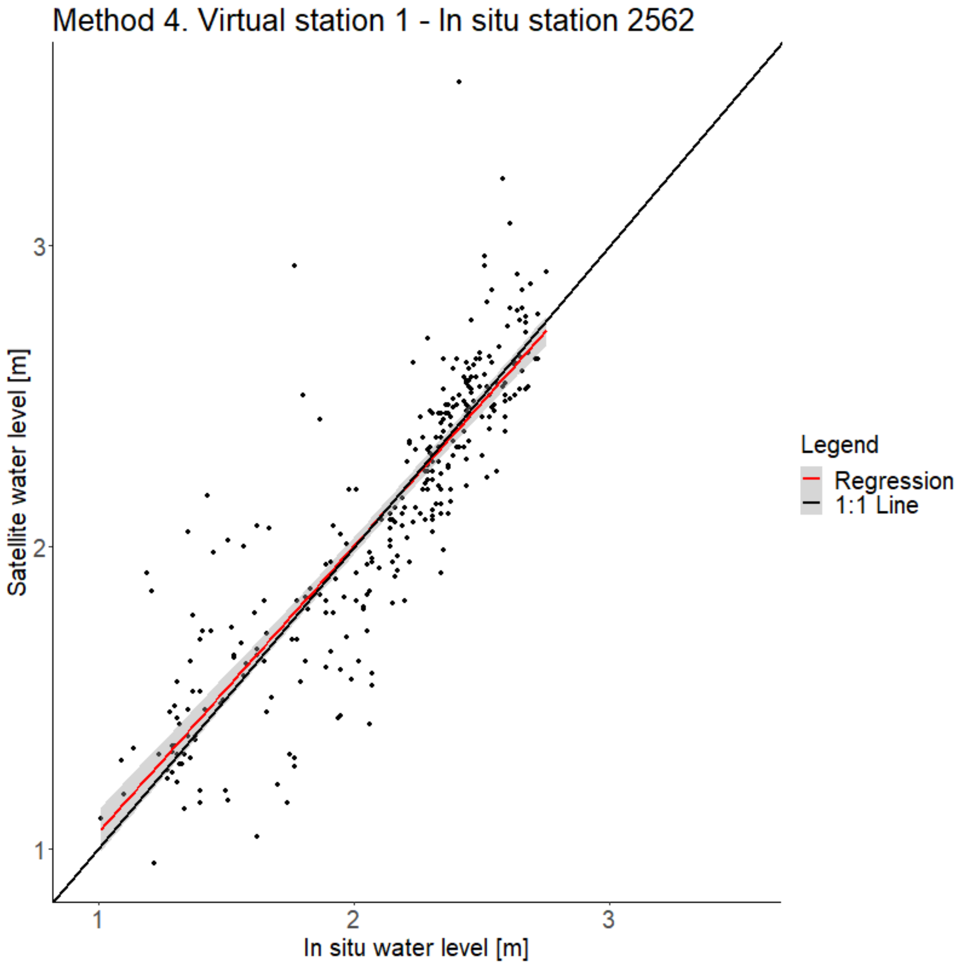


Figure 4.64: Scatter plot of the couple 2-2562 (Method 4)

This scatter plot shows the good agreement of this couple. R is 0.87 and RMSE 0.24.

5 | Conclusion

The aim of this work is a feasibility study of using satellite altimetric data (in particular Sentinel 3 data) for water level retrieving. This work of Thesis is part of a European project called GAUSS of Copernicus. The project has been done in collaboration with the Finnish Environment Institute (SYKE). A results of the thesis is also the creation of a code that can be implement and automatically do the comparison among satellite and in-situ stations with the future satellite altimetric data that will be available from Copernicus database (https://land.copernicus.eu/global/products/wl).

The altimetric data validated are related to Finland. This country has about 188,000 lakes and approximately 25,000 km of rivers. The water is an important resource of the finnish territory. It is very well distributed in the country. It is linked, for example, with tourism, economy and energy production. Its monitoring is also important trying to limit the damages coming from extreme events as droughts and floods. Satellite data can be a new source of water level data that can be used as integration of in-situ data or replace it in case of lack of in-situ measures. Their advantage is the free use and the disadvantage is the long duration of the orbit cycle that cause a low frequency of data available in the time. Satellite data have been validated comparing them with in-situ measurements to evaluate their performance in water level estimate. First altimeter were able to retrieve water level of large water bodies as oceans and during the time the improving of the technology permitted to retrieve also water levels of inland waters, then rivers with large width and now they are trying to go toward rivers wit narrow sections.

In this work we applied four methods for comparing in-situ and satellite data: (1) comparing the differences between successive times of satellite and in-situ measurements, (2) comparing the satellite and in-situ measurements assuming that the initial value of the joint measurements is equal (specifically the value of satellite water level is equal to the value of in-situ one), (3) comparing the satellite and in-situ measurements calibrating the initial water level of satellite measurements minimizing the MAE, (4) comparing the satellite and in-situ measurements calibrating the initial water level of satellite measurements minimizing the RMSE. The best performances have been found for the method 3 and 4 in terms of Pearson Coefficient and RMSE. These methods permit to compare directly the two data series without considering the surface of reference of the two records. The strength of the proposed methodology is that they can be applied also in the cases when the reference systems for the various in-situ sensors is unknown.

A first analysis has been done on the annual measurements finding good results despite the complexity of the hydrological network of Finland. The rivers sections are small and are present an high number of dams and power plants that could also regulate the flow, meanders and changing of the section in shape and dimensions. In some cases if there are the presence of some disturb between the two stations analyzed the agreement among satellite and in-situ measures could collapse. Increasing the distance between the virtual station and the in-situ stations the median value of the Pearson Coefficient decreases and the one of the RMSE increases.

Considering the radius 5, 10, 15, 20, 25, 35 it has been found respectively 2, 4, 9, 17, 23, 34, and 48 couples. The mean of the R are 0.86, 0.74, 0.80, 0.77, 0.76, 0.75 and 0.74 and the mean of the RMSE are 0.34, 0.42, 0.42, 0.43, 0.47, 0.49, 0.49 . These two indicators are related to method 4. It can be appreciate the trends described above. The numbers of the respectively couples increase if the radius increase and are 2, 4, 9, 17, 23, 34, 48. For rivers couples it has been found a median value greater than 0.8 also considering a distance of 20 km among the stations. It has been found also some specific couples that performs very well, also with R greater than 0.9 and RMSE lower than 0.2. These particular couples with the best performance could be used to continue the validation operation. A good agreement has been found for some rivers virtual stations that are downstream of a dam and the in-situ station just downstream of the previous dam. It has been assessed a further analysis based on seasonality trying to understand when during the year satellite altimetry has the best performances and when it fails. The reason of a different behaviour of the altimeter considering the same section could be the differences of the temperatures of the season considered that influence the physical status of the water. It is also possible that in the winter period the turbolence are less than the summer period when the water flux increase For winter were found better performances in terms of R1 and R4 than in the summer. But at the same time winter season has higher value of RMSE considering the method 4 than the summer. So the altimeter is more able to retrieve the trends but less to minimize the error. Winter respect to the annual period has better performances in terms of R for the method of the temporal differences but considering the method 4 they are equal. Therefore, for the couple that does not performs well using annual data it is possible to go more on detail and try to understand which one is the period when the altimeter fails. A consequence could be that during this period it not possible to use these measurements or maybe it could possible to use but carefully.

The lakes records, retrieved from the site https://www.theia-land.fr/, have been validated with excellent results finding many in-situ station that are located on the banks of these lakes. Wherease for rivers different radius have been considered to compare in-situ and satellite data, for lakes have been selected directly the stations that fall inside it and hydraulically connected.

From the results, we can say that altimetric satellite data (Sentinel 3) can be exploit to retrieve water level measures, also in a dense and intricate river network as the Finland one. Good results have been found also for the lakes. Regarding all the couples selected have been found the median of R equal to 0.88 and the median of the RMSE equal to 0.2. Both indicator are related to method 4. Have been found couples with R very high (more than 0.9) and RMSE very high (less than 0.15). This open the possibility of a new data source of water level that can be useful also to monitor hydro-power plants, drought assessment or helping decision making of water resources.

From more operational point of view, the analysis done in the work of thesis has been also the set-up of an algorithm that collect all the in-situ water level measures in a database and directly permit the comparison with satellite altimeter data. The algorithm developed are operational tools that can be exploited by the partners of the GAUSS (Generating Advanced Usage of Earth Observation for Smart Statistics) project.

6 | Future developments

The actual available satellite stations for Finland are 3 rivers virtual stations and 4 lakes virtual stations. The analysis could be continued integrating the new satellite data that will be available in future in Copernicus database. Finland case study has a good spatial distribution of in-situ stations (around 1690) that allow to validate the new virtual stations whereas they will be located. The high spatial distribution of in-situ stations permit to compare each virtual station with a series of in-situ ones, having a high number of couples for the validation of each virtual station. Moreover, it could be interesting to have lakes altimetric records only related to Sentinel 3 to evaluate this satellite individually. At the same time could be interesting to evaluate the feasibility of a mixed record of rivers virtual stations, as for lakes ones. As said in the Section 4 it has been found high agreement between the lakes mixed altimetric record and the in-situ stations. Moreover it is possible to have information with higher frequency.

Here it has been assessed an analysis considering seasonality. It could be done using different months for winter and summer or also increasing the number of seasons to consider all the climatic shades of the year. Some couples could improve or get worse their agreement going from annual record to a specific season. A little focus could be done for example for the lake couple 2-2475 where there are satellite measure peaks very pronounced respect the in-situ measurements. For this couple the R4 is high (0.89) and the RMSE4 is high too (0.81) and this last indicators it is not so good.

The code developed for the Finland case could be used also for the validation of virtual station of other countries. This work could be extended to the Europe and to all the countries of the world if the virtual station record are o will available. For each country the type of the in-situ dataset could be different so it will be necessary to make some little changes to the code to manage that records.

Bibliography

[1] D. E. Alsdorf and D. P. Lettenmaier. Tracking fresh water from space. *Science*, 301 (5639):1491–1494, 2003.

[2] D. E. Alsdorf and D. P. Lettenmaier. Tracking fresh water from space. *Science*, 301 (5639):1491–1494, 2003. doi: 10.1126/science.1089802. URL `https://www.science.org/doi/abs/10.1126/science.1089802`.

[3] D. E. Alsdorf, E. Rodríguez, and D. P. Lettenmaier. Measuring surface water from space. *Reviews of Geophysics*, 45(2), 2007.

[4] N. W. Arnell and S. N. Gosling. The impacts of climate change on river flow regimes at the global scale. *Journal of Hydrology*, 486:351–364, 2013. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2013.02.010. URL `https://www.sciencedirect.com/science/article/pii/S0022169413001224`.

[5] T. P. BARNETT, J. C. ADAM, and D. P. LETTENMAIER. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature (London)*, 438(7066):303 – 309, 2005. ISSN 0028-0836.

[6] M. Bavay, M. Lehning, T. Jonas, and H. Löwe. Simulations of future snow cover and discharge in alpine headwater catchments. *Hydrological Processes: An International Journal*, 23(1):95–108, 2009.

[7] M. Beniston. Climatic change in mountain regions: a review of possible impacts. *Climate variability and change in high elevation regions: Past, present & future*, pages 5–31, 2003.

[8] C. M. Birkett, L. Mertes, T. Dunne, M. Costa, and M. Jasinski. Surface water dynamics in the amazon basin: Application of satellite radar altimetry. *Journal of Geophysical Research: Atmospheres*, 107(D20):LBA–26, 2002.

[9] S. Bogning, F. Frappart, F. Blarel, F. Nino, G. Mahé, J.-P. Bricquet, F. Seyler, R. Onguéné, J. Etame, M.-C. Paiz, et al. Monitoring water levels and discharges

using radar altimetry in an ungauged river basin: The case of the ogooué. *Remote Sensing*, 10(2):350, 2018.

[10] S. Calmant and F. Seyler. Continental surface waters from satellite altimetry. *Comptes Rendus Geoscience*, 338(14-15):1113–1122, 2006.

[11] S. Calmant and F. Seyler. Continental surface waters from satellite altimetry. *Comptes Rendus Geoscience*, 338(14-15):1113–1122, 2006.

[12] D. B. Chelton, J. C. Ries, B. J. Haines, L.-L. Fu, and P. S. Callahan. Satellite altimetry. In *International geophysics*, volume 69, pages 1–ii. Elsevier, 2001.

[13] M. T. Coe and C. M. Birkett. Calculation of river discharge and prediction of lake height from satellite radar altimetry: Example for the lake chad basin. *Water Resources Research*, 40(10), 2004.

[14] T. I. Committee. Land of Islands and Waters.

[15] J.-F. Crétaux and C. Birkett. Lake studies from satellite radar altimetry. *Comptes Rendus Geoscience*, 338(14-15):1098–1112, 2006.

[16] J.-F. Crétaux, A. Arsen, S. Calmant, A. Kouraev, V. Vuglinski, M. Bergé-Nguyen, M.-C. Gennero, F. Nino, R. A. Del Rio, A. Cazenave, et al. Sols: A lake database to monitor in the near real time water level and storage variations from remote sensing data. *Advances in space research*, 47(9):1497–1507, 2011.

[17] J.-F. Cretaux, M. Berge-Nguyen, S. Calmant, N. Jamangulova, R. Satylkanov, F. Lyard, F. Perosanz, J. Verron, A. Samine Montazem, G. Le Guilcher, et al. Absolute calibration or validation of the altimeters on the sentinel-3a and the jason-3 over lake issykkul (kyrgyzstan). *Remote Sensing*, 10(11):1679, 2018.

[18] P. Da Ronco and C. De Michele. Cloud obstruction and snow cover in alpine areas from modis products. *Hydrology and earth system sciences*, 18(11):4579–4600, 2014.

[19] J. S. Da Silva, S. Calmant, F. Seyler, O. C. Rotunno Filho, G. Cochonneau, and W. J. Mansur. Water levels in the amazon basin derived from the ers 2 and envisat radar altimetry missions. *Remote sensing of environment*, 114(10):2160–2181, 2010.

[20] J. S. Da Silva, S. Calmant, F. Seyler, O. C. Rotunno Filho, G. Cochonneau, and W. J. Mansur. Water levels in the amazon basin derived from the ers 2 and envisat radar altimetry missions. *Remote sensing of environment*, 114(10):2160–2181, 2010.

[21] J. S. Da Silva, S. Calmant, F. Seyler, O. C. Rotunno Filho, G. Cochonneau, and

W. J. Mansur. Water levels in the amazon basin derived from the ers 2 and envisat radar altimetry missions. *Remote sensing of environment*, 114(10):2160–2181, 2010.

[22] I. de Oliveira Campos, F. Mercier, C. Maheu, G. Cochonneau, P. Kosuth, D. Blitzkow, and A. Cazenave. Temporal variations of river basin waters from topex/poseidon satellite altimetry. application to the amazon basin. *Comptes Rendus de l'Académie des Sciences-Series IIA-Earth and Planetary Science*, 333(10): 633–643, 2001.

[23] C. Deidda, C. De Michele, A. N. Arslan, S. Pecora, and N. Taburet. Accuracy of copernicus altimeter water level data in italian rivers accounting for narrow river sections. *Remote Sensing*, 13(21), 2021. ISSN 2072-4292. doi: 10.3390/rs13214456. URL `https://www.mdpi.com/2072-4292/13/21/4456`.

[24] E. W. Directors. Best practices on flood prevention, protection and mitigation, meetings, budapest, 30 november and 1 december 2002, bonn, 5 and 6 february 2003, 2003.

[25] H. Dixon, J. Hannaford, and M. J. Fry. The effective management of national hydrometric data: experiences from the united kingdom. *Hydrological Sciences Journal*, 58(7):1383–1399, 2013.

[26] C. Donlon, B. Berruti, A. Buongiorno, M.-H. Ferreira, P. Féménias, J. Frerick, P. Goryl, U. Klein, H. Laur, C. Mavrocordatos, et al. The global monitoring for environment and security (gmes) sentinel-3 mission. *Remote sensing of Environment*, 120:37–57, 2012.

[27] J. Fernández, C. Fernández, P. Féménias, and H. Peter. The copernicus sentinel-3 mission. In *ILRS workshop*, pages 1–4, 2016.

[28] P. Francesco, C. Pandolfo, M. Stelluti, N. Berni, L. Brocca, and T. Moramarco. Assessment of rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in the umbria region (central italy). *Landslides*, 9:229–237, 06 2012. doi: 10.1007/s10346-011-0287-3.

[29] F. Frappart, S. Calmant, M. Cauhopé, F. Seyler, and A. Cazenave. Preliminary results of envisat ra-2-derived water levels validation over the amazon basin. *Remote sensing of Environment*, 100(2):252–264, 2006.

[30] F. Frappart, F. Papa, J. S. Famiglietti, C. Prigent, W. B. Rossow, and F. Seyler. Interannual variations of river water storage from a multiple satellite approach: A case

study for the rio negro river basin. *Journal of Geophysical Research: Atmospheres*, 113(D21), 2008.

[31] F. Frappart, F. Blarel, I. Fayad, M. Bergé-Nguyen, J.-F. Crétaux, S. Shu, J. Schregenberger, and N. Baghdadi. Evaluation of the performances of radar and lidar altimetry missions for water level retrievals in mountainous environment: The case of the swiss lakes. *Remote Sensing*, 13(11):2196, 2021.

[32] L.-L. Fu, E. J. Christensen, C. A. Yamarone Jr, M. Lefebvre, Y. Menard, M. Dorrer, and P. Escudier. Topex/poseidon mission overview, 1994.

[33] W. Gould. Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecological applications*, 10(6):1861–1870, 2000.

[34] A. H. Group, C. Vörösmarty, A. Askew, W. Grabs, R. Barry, C. Birkett, P. Döll, B. Goodison, A. Hall, R. Jenne, et al. Global water data: A newly endangered species. *Eos, Transactions American Geophysical Union*, 82(5):54–58, 2001.

[35] M. Halicki and T. Niedzielski. The accuracy of the sentinel-3a altimetry over polish rivers. *Journal of Hydrology*, 606:127355, 2022. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2021.127355. URL `https://www.sciencedirect.com/science/article/pii/S0022169421014050`.

[36] D. M. Hannah, S. Demuth, H. A. van Lanen, U. Looser, C. Prudhomme, G. Rees, K. Stahl, L. M. Tallaksen, et al. Large-scale river flow archives: importance, current status and future needs. *Hydrological Processes*, 25(7):1191–1200, 2011.

[37] P. Horton, B. Schaefli, B. Hingray, A. Mezghani, and A. Musy. Prediction of climate change impacts on alpine discharge regimes under a2 and b2 sres emission scenarios for two future time periods (2020-2049, 2070-2099). Technical report, 2005.

[38] Q. Huang, X. Li, P. Han, D. Long, F. Zhao, and A. Hou. Validation and application of water levels derived from sentinel-3a for the brahmaputra river. *Science China Technological Sciences*, 62(10):1760–1772, 2019.

[39] L. Jiang, K. Nielsen, O. B. Andersen, and P. Bauer-Gottwein. Cryosat-2 radar altimetry for monitoring freshwater resources of china. *Remote Sensing of Environment*, 200:125–139, 2017.

[40] L. Jiang, K. Nielsen, S. Dinardo, O. B. Andersen, and P. Bauer-Gottwein. Evaluation of sentinel-3 sral sar altimetry over chinese rivers. *Remote Sensing of Environment*, 237:111546, 2020.

[41] P. Krause, D. Boyle, and F. Bäse. Comparison of different efficiency criteria for hydrological model assessment. *Advances in geosciences*, 5:89–97, 2005.

[42] T. Lafon. Jason 1: lessons learned from the development and 1 year in orbit. *Acta Astronautica*, 56(1-2):45–49, 2005.

[43] A. Laghari, D. Vanham, and W. Rauch. To what extent does climate change result in a shift in alpine hydrology? a case study in the austrian alps. *Hydrological Sciences Journal*, 57(1):103–117, 2012.

[44] Y. Le Roy, M. Deschaux-Beaume, C. Mavrocordatos, M. Aguirre, and F. Heliere. Sral sar radar altimeter for sentinel-3 mission. In *2007 IEEE International Geoscience and Remote Sensing Symposium*, pages 219–222. IEEE, 2007.

[45] A. Liibusk, T. Kall, S. Rikka, R. Uiboupin, Ü. Suursaar, and K.-H. Tseng. Validation of copernicus sea level altimetry products in the baltic sea and estonian lakes. *Remote Sensing*, 12(24):4062, 2020.

[46] P. Maillard, N. Bercher, and S. Calmant. New processing approaches on the retrieval of water levels in envisat and saral radar altimetry over rivers: A case study of the são francisco river, brazil. *Remote Sensing of Environment*, 156:226–241, 2015.

[47] J. Nash. River flow forecasting through conceptual models, i: A discussion of principles. *Journal of Hydrology*, 10:398–409, 1970.

[48] W. W. A. P. U. Nations) and UN-Water. Water in a changing world. 2009.

[49] E. Obligis, C. Desportes, L. Eymard, M. Fernandes, C. Lázaro, and A. Nunes. Tropospheric corrections for coastal altimetry. In *Coastal altimetry*, pages 147–176. Springer, 2011.

[50] F. Pan and J. Nichols. Remote sensing of river stage using the cross-sectional inundation area-river stage relationship (iarsr) constructed from digital elevation model data. *Hydrological Processes*, 27(25):3596–3606, 2013.

[51] S. L. Postel and C. Gretchen. Daily, and paul r. *Ehrlich."Human appropriation of renewable fresh water." Science-AAAS-Weekly Paper Edition*, 271:785–787, 1996.

[52] R. Schneider, A. Tarpanelli, K. Nielsen, H. Madsen, and P. Bauer-Gottwein. Evaluation of multi-mode cryosat-2 altimetry data over the po river against in situ data and a hydrodynamic model. *Advances in water resources*, 112:17–26, 2018.

[53] C. Schwatke, D. Dettmering, W. Bosch, and F. Seitz. Dahiti–an innovative approach

for estimating water level time series over inland waters using multi-mission satellite altimetry. *Hydrology and Earth System Sciences*, 19(10):4345–4364, 2015.

[54] H. Shahabi and M. Hashim. Landslide susceptibility mapping using gis-based statistical models and remote sensing data in tropical environment. *Scientific Reports*, 5, 2015.

[55] S. Shu, H. Liu, R. A. Beck, F. Frappart, J. Korhonen, M. Lan, M. Xu, B. Yang, and Y. Huang. Evaluation of historic and operational satellite radar altimetry missions for constructing consistent long-term lake water level records. *Hydrology and Earth System Sciences*, 25(3):1643–1670, 2021. doi: 10.5194/hess-25-1643-2021. URL `https://hess.copernicus.org/articles/25/1643/2021/`.

[56] A. W. Sichangi, L. Wang, K. Yang, D. Chen, Z. Wang, X. Li, J. Zhou, W. Liu, and D. Kuria. Estimating continental river basin discharges using multiple remote sensing data sets. *Remote Sensing of Environment*, 179:36–53, 2016.

[57] Y. Sulistioadi, K.-H. Tseng, C. Shum, H. Hidayat, M. Sumaryono, A. Suhardiman, F. Setiawan, and S. Sunarso. Satellite radar altimetry for monitoring small rivers and lakes in indonesia. *Hydrology and Earth System Sciences*, 19(1):341–359, 2015.

[58] N. Taburet, L. Zawadzki, S. Calmant, J.-F. Crétaux, F. Mercier, and M. Vayre. Copernicus Global Land Operations "Cryosphere and Water": Product user manual - lake and river water level. pages 1–41, 2020.

[59] B. D. Tapley, J. Ries, G. Davis, R. Eanes, B. Schutz, C. Shum, M. Watkins, J. Marshall, R. Nerem, B. Putney, et al. Precision orbit determination for topex/poseidon. *Journal of Geophysical Research: Oceans*, 99(C12):24383–24404, 1994.

[60] A. Tarpanelli, S. Barbetta, L. Brocca, and T. Moramarco. River discharge estimation by using altimetry data and simplified flood routing modeling. *Remote Sensing*, 5 (9):4145–4162, 2013.

[61] A. Tarpanelli, S. Camici, K. Nielsen, L. Brocca, T. Moramarco, and J. Benveniste. Potentials and limitations of sentinel-3 for river discharge assessment. *Advances in Space Research*, 68(2):593–606, 2021.

[62] M. Tikkanen. Long-term changes in lake and river systems in finland. *Fennia*, 180: 31–42, 01 2002.

[63] D. Vanham. The alps under climate change: implications for water management in europe. *Journal of Water and Climate Change*, 3(3):197–206, 2012.

[64] D. Vanham, E. Fleischhacker, and W. Rauch. Impact of snowmaking on alpine water resources management under present and climate change conditions. *Water Science and Technology*, 59(9):1793–1801, 2009.

[65] P. Vaze, S. Neeck, W. Bannoura, J. Green, A. Wade, M. Mignogno, G. Zaouche, V. Couderc, E. Thouvenot, and F. Parisot. The jason-3 mission: Completing the transition of ocean altimetry from research to operations. In *Sensors, Systems, and Next-Generation Satellites XIV*, volume 7826, pages 264–268. SPIE, 2010.

[66] N. Veijalainen, T. Dubrovin, M. Marttunen, and B. Vehviläinen. Climate change impacts on water resources and lake regulation in the vuoksi watershed in finland. *Water resources management*, 24(13):3437–3459, 2010.

[67] E. Zakharova, K. Nielsen, G. Kamenev, and A. Kouraev. River discharge estimation from radar altimetry: Assessment of satellite performance, river scales and methods. *Journal of Hydrology*, 583:124561, 2020.

[68] B. Zierl and H. Bugmann. Global change impacts on hydrological processes in alpine catchments. *Water Resources Research*, 41(2), 2005.

List of Figures