



Earth and Life Institute (ELI)

Participatory approaches to monitor water-related Sustainable Development Goal (SDG 6) in Tunisia

Raed Fehri

Thesis submitted in partial fulfillment
of the requirements for the degree of
Docteur en sciences agronomiques et ingénierie biologique

Dissertation committee:

Prof. Marnik Vanclooster (UCLouvain, advisor)

Dr. Slaheddine Khlifi (ESIM, advisor)

Prof. Kim Mens (UCLouvain)

Prof. Buytaert Wouter (Imperial College London, UK)

Prof. Leonardo Alfonso Segura (UNESCO-IHE, Delft)

Prof. Emmanuel Hanert (UCLouvain, president)

Version of March 2021.

Scientific and participatory approaches to monitor water-related Sustainable Development Goal (SDG 6) in Tunisia

by Raed Fehri

© Raed Fehri 2021

ELI

Université catholique de Louvain

Croix du sud, 2

1348 Louvain-la-Neuve

Belgium

This work was partially supported by the Islamic Development Bank (IDB), the Wallonie Bruxelles International (WBI), and the Université catholique de Louvain.

To my family...
I love you so much

ACKNOWLEDGMENTS

The success of this thesis was not possible without the support of the people who made it happen by closely collaborating, or just by being a helpful hand. My sincere thanks and gratitude go to:

Prof. Marnik Vanclooster and Dr. Slaheddine Khlifi, my advisors, for their wise advice, patience, and strong support throughout all the steps of this thesis; a special tribute goes to Marnik Vanclooster for giving me the opportunity to start this PhD with him and for helping me shape the different objectives of this work; similarly, my deep gratitude goes to Slaheddine Khlifi who was in my guidance since my engineering studies, I will never forget his unconditional support on the field during the citizen science campaigns and all the wonderful talks that we shared;

Prof. Kim Mens and Prof. Wouter Buytaert, the members of my PhD committee, for their guidance and support during this journey; I would like to thank Kim Mens for helping me bring the Together4Water platform to reality; also, a special thank goes to Wouter Buytaert for inspiring me to develop the Together4Water citizen science approach, his valuable work helped me discover in depth the concept of participatory science;

Prof. Patrick Bogaert, who was a sensational support for me during our collaboration on Data Fusion of citizen science discharge outcomes; Patrick Bogaert is not only an amazing professor with remarkable mathematical knowledge, but who is also a kind person with impressive dedication to work and attention to details;

ACKNOWLEDGMENTS

Both master students: Nicolas Deineko and Antoine Van Grootenbrulle (Rest in Peace) who contributed to the development of the Together4-water platform, their motivation and creativity were crucial to the achievement of the website;

All the citizen scientists who participated in the different citizen science activities of the Together4Water project and contributed to the success of the data collection campaigns, without their motivation and dedication this work couldn't be possible;

The Tunisian water institutes: DGRE, CRDA, SONEDE, ONAS, and the Ministry of Agriculture, for their help during our data collection campaign, without their collaboration this work was not possible;

My dear friend Mokrane Kadir, who stood with me during this journey, he was an endless help and support for me with his spontaneous spirit and very warm heart;

The members of GERU, who were a family for me that gave me the possibility to do more than research and be part of amazing activities such as sport, scientific animation, trips, and special lunches and dinners, I will never forget that warm atmosphere full of laughter and love everytime we get together;

My soulmate and my love, for her support "no matter what", her love, her courage when I lost mine, and for never losing faith in me;

And of course my Mom, my Dad, my two brothers, and my sister for being my biggest support in this journey and for believing in me even when I didn't believe in myself; no matter what I do, I can never give back to my Mom and Dad what they did for me, praying for me everyday without knowing much about what's going on with me, and more importantly raising me into this version of myself; their resilience and strength is the fuel that keeps me fighting everyday to fulfill my dreams and make them proud of me;

ABSTRACT

The Sustainable Development Goals (SDGs) are a global initiative lead by the United Nations to achieve “a better and more sustainable future for all”. Given the increasing impact of climate change on the available water resources over the past two decades, a particular attention is paid to water through SDG 6. The SDG-6 aims to “ensure availability and sustainable management of water and sanitation for all”. This goal comes with a set of highly interconnected targets and indicators that are essential for its global and local implementation. Tunisia, as a member of the United Nations, adopted the framework of the SDGs with a particular interest in SDG 6 since Tunisia is considered as one of the highly water stressed countries in North Africa. Despite the growing need for a rapid, yet consistent, implementation of the goal in the country, very few scientific initiatives are working on monitoring the different indicators at the national scale using nationally-produced data. However, lower scale implementation and monitoring of SDG 6 indicators are essential to provide a consistent overview of the current state of water resources in the country. The current data infrastructure in Tunisia represents a major constraint, since the water-related information is poorly monitored and not accessible at the regional scale. This is due to the complex water institutional system, which is mostly dominated by the lack of coordination and communication between the different data providers. Nevertheless, over the past decade, participatory approaches for water monitoring have been getting more attention in many parts of the world, especially with the emergence of novel cost-effective smart technologies. Such participatory approaches are often referred to as Citizen Science (CS). With CS, citizens are involved in scientific activities, such as environmental monitoring, and they also can be a valuable part of decision making and policy formulation. Together with

the generic available data, in particular from remote sensing technology, CS could be tested and implemented in Tunisia in order to improve the access to available and reliable data.

The main objective of this thesis was to strengthen the scientific basis for participatory approaches in conjunction with public available environmental data to support the monitoring of water resources in Tunisia. The strengthening of water observation is expected to improve the monitoring of the SDG 6 indicators in Tunisia. The study is focused on the Medjerda catchment, which is the most important river basin in Tunisia and considered as the largest national water resource. The focus was first put on the calculation of the water stress SDG 6 indicator (6.4.2). This indicator was selected as a specific case in this study since water stress is very relevant for the considered catchment. The indicator was evaluated at the scale of the Medjerda catchment and 4 regions within for the period 2000-2016. The spatial and temporal disaggregation was based on analysis of both governmental and remote sensing data, and provided a valuable understanding of the seasonal and regional behavior of water stress. The lack of access to reliable official data was the major constraint in measuring the water stress indicator. The second research section of the thesis was then dedicated to the improvement of the water-related database of the Medjerda catchment through data collection and the creation of a publicly available and standardized online platform "Together4Water", which provides access and dynamic visualizations of the collected data as well as interactive real-time communication tool with users. Consequently, we introduced in the third section of this work the concept of CS as a complementary source of data in Tunisia. The approach was based on the engagement of everyday citizens of a test region within the Medjerda catchment in monitoring rainfall, discharge, and water quality using cost-effective tools. Focus was mainly put on rainfall and discharge monitoring. The Together4Water website was the base for reaching a wide range of citizens from different generations in the test region allowing the initiative to engage a group of volunteers in data collection. For rainfall, traditional low-cost rain gauges were used at multiple locations, while the data transmission was ensured through the Together4Water website. Results yield a good agreement between citizens' observations and the official rainfall data. For discharge, a publicly available smartphone app "Discharge app" was used to estimate river flow of the Medjerda at two locations. A Data fusion technique was then used to validate and combine citizens' estimations to generate unique yet higher quality discharge datasets for both locations. Finally, an assessment of the

potential contribution of the CS initiative to the SDG 6 monitoring was performed. We conclude that around 54% of the goal's indicators could be supported by our CS initiative, which emphasizes the value of public engagement in SDG 6 monitoring.

Water...
is the driving force of all nature.
LEONARDO DA VINCI

TABLE OF CONTENTS

- 1 Introduction and Objectives 1**
 - 1.1 Introduction 2
 - 1.1.1 SDG 6: a global framework for Water 3
 - 1.1.2 SDG 6 related water challenges in Tunisia 5
 - 1.1.3 The current state of water-related data in Tunisia 7
 - 1.1.4 Citizen Science to support the monitoring of the implementation of SDG 6 8
 - 1.2 Objectives 11
- 2 Disaggregation of SDG-6 Water Stress indicator 17**
 - 2.1 Introduction 19
 - 2.2 Materials and methods 20
 - 2.2.1 Study area 20
 - 2.2.2 Data sources 22
 - 2.2.3 Estimation of irrigated lands’ water consumption 24
 - 2.2.4 SDG-6 water stress indicator 25
 - 2.2.5 First-order uncertainty analysis 28
 - 2.3 Results and discussion 29
 - 2.3.1 WS estimates at basin and administrative region scale using governmental data 29
 - 2.3.2 Governmental data vs. RS data-based irrigation water consumption and WS estimates 32
 - 2.3.3 Disaggregating yearly WS estimates 34
 - 2.3.4 Uncertainty analysis 35
 - 2.4 Conclusion 38
- 3 Creation of the Together4Water platform 41**
 - 3.1 Introduction 43
 - 3.2 Approach 43

TABLE OF CONTENTS

3.3	The available water-related database	44
3.4	Concept	46
3.5	Requirements of the web application	47
3.6	Existing Citizen Science platforms	47
3.7	Prototype	49
3.7.1	The prototype of Together4Water	49
3.8	User interface	51
3.8.1	Public website pages	51
3.8.2	Administration pages	53
3.9	Conclusion	55
4	Citizen Science for Water monitoring in Tunisia	57
4.1	Introduction	59
4.2	Materials and methods	61
4.2.1	Step-by-step citizens science approach	61
4.2.2	Data quality assessment and validation	63
4.3	Results and discussion	63
4.3.1	Data variability	63
4.3.2	Correlation analysis	64
4.3.3	Percent error (PE) assessment	65
4.3.4	Student's t-test and mean variability comparison	66
4.3.5	Importance of the step-by-step citizens' engagement	68
4.4	Challenges	69
4.5	Conclusion	70
5	Data Fusion of Citizen Science data	71
5.1	Introduction	73
5.2	Materials and methods	75
5.2.1	Study area	75
5.2.2	Citizens engagement methodology	77
5.2.3	Mobile phone discharge application	78
5.2.4	Data fusion and prediction of discharge	80
5.2.5	Sources of errors	82
5.3	Results and discussions	85
5.3.1	Data quality assessment	85
5.3.2	Distribution of the errors	87
5.3.3	Variance decomposition	88
5.3.4	Data fusion	90
5.4	Conclusions	92

6	The potential of Citizen Science in SDG 6 monitoring	95
6.1	Introduction	97
6.2	Materials and Methods	99
6.2.1	Study area	99
6.2.2	The Together4Water Citizen Science project	101
6.2.3	Review of SDG 6 indicators	102
6.3	Results	104
6.3.1	Citizen Science data collection	104
6.3.2	Potential contribution of CS to SDG 6 in Tunisia	105
6.4	Discussion	106
6.5	Conclusion	108
7	Conclusion and outlooks	111
7.1	Conclusion	112
7.1.1	General conclusion	112
7.1.2	Specific objectives	112
7.1.3	Specific research questions	114
7.2	Perspectives	117
7.2.1	More advanced and large scale implementation of Citizen Science	117
7.2.2	The added value of training in Citizen Science	118
7.2.3	More focus on water quality monitoring	119
7.2.4	Measurement of SDG 6 indicators	120
	References	121
	Appendices	131

LIST OF FIGURES

1.1	SDG 6 targets and indicators. (Source: unwater.org)	4
1.2	Specific objectives of the PhD thesis represented in their general context	13
1.3	Flowchart of the research questions investigated in this PhD	15
2.1	The study area. The right down figure is the situation of the study area, while the inset yields the location of the study area in North Africa.	21
2.2	Map of MODIS-ET (April 2016) and CCI LandCover 2015 in the Medjerda catchment.	23
2.3	Yearly WS at the scale of the Medjerda catchment.	30
2.4	Yearly WS at the scale of the administrative regions.	31
2.5	Comparison of yearly irrigation water consumption.	32
2.6	Level of WS at the scale of the Medjerda (black color refers to WS using only governmental data and the orange refers to WS that includes estimated irrigation).	33
2.7	Comparison of Level of WS at the scale of Medjerda subregions.	34
2.8	Monthly distribution of WS at the scale of Medjerda catchment.	34
2.9	Monthly distribution of WS at the scale of Medjerda subregions.	35
2.10	Comparison of irrigation uncertainty contribution (gov data vs estimated data).	36
2.11	Comparison of total uncertainty ($S_{WS(RS)}$ vs $S_{WS(GOV)}$).	36
3.1	Home page in the prototype. The chat window is opened.	49
3.2	Map page in the prototype. A data type is selected and a location popup is visible.	50

3.3	Data selection page in the prototype. By selecting a data type and a location and clicking “Show visualization”, the user will be redirected to the data visualization page where time series are displayed in line charts.	50
3.4	Screenshot of the website’s map page. A data type is selected, a location popup is shown and the background map selection menu is expanded.	52
3.5	Screenshot of the data visualization page for a specific dataset.	53
3.6	Screenshot of the website page in the administration site.	54
3.7	Specification of the CSV format of files containing datasets of multiple time series.	55
4.1	The study area. The right down inset is Medjez-El-Beb city, the CS test region.	61
4.2	CS step-by-step approach.	62
4.3	Rain gauge fixed by a citizen scientist during monitoring rainfall.	62
4.4	Variability of monthly cumulative CS and reference data (Ref) at ESIM site (a), Medjez-El-Beb city center (b) and Slouguia (c).	64
4.5	Correlation between CS measurements and reference data at all sites.	65
4.6	Variation of monthly rainfall PE at (a) Medjez-El-Beb city center, (b) ESIM, and (c) Slouguia.	66
4.7	Comparison of the mean and the variability (95% CI) of the CS measurement to the reference data.	67
4.8	Rainfall CS data transmission example with a short description of the event.	68
4.9	Photos of rain gauges sent by citizens during monitoring rainfall in the test region.	69
5.1	Location of Medjez-N5 and Slouguia study sites.	76
5.2	Picture of the reference sites with corresponding screenshots of the smartphone application for two events. Parts (a) and (b) refer to the Medjez-N5 site, while Parts (c) and (d) refer to the Slouguia site.	79
5.3	Screenshot of the app with the encoded parameters at the Medjez-N5 site.	80

5.4	Citizens-based versus reference-based discharge measurements at the Medjez-N5 site, along with the 1:1 line (black plain line), the estimated regression line (red plain line), and the corresponding 95% prediction interval (dashed lines). . .	86
5.5	Histograms (on diagonal) and scatterplots (off diagonal) of the discharge residuals for the 4 citizens at the Medjez-N5 site, with corresponding correlation coefficients between citizens' discharge measurements for the same events.	89
5.6	Predicted (fused) citizens-based discharges versus reference discharges, along with the corresponding prediction intervals (dashed lines) around the 1:1 line (plain line). Part (a) refers to the Medjez-N5 site, while Part (b) refers to the Slouguia site.	91
6.1	Study region.	100
6.2	Citizen science data collection tools.	102
6.3	Steps of SDG 6 indicators review for the potential Together4Water contribution.	104
6.4	Potential contributions of the Together4Water project to SDG 6 indicators.	106
B. 1	The study area.	135
B. 2	The monitoring sites (a) Noirhat, (b) Rosières, and (c) Suzeryl. The white spots in each site are the markers fixed on the river banks.	136
B. 3	Steps for site installation in the Discharge app.	136
B. 4	Discharge app estimates versus Aqualim data at (a) Noirhat, (b) Rosières, and (c) Suzeryl, along with the estimated regression line (red plain line), and the corresponding 95% prediction interval (dashed lines).	137

LIST OF TABLES

2.1	Summary of the requirements for the estimation of irrigation water consumption in the Medjerda	25
2.2	Parameter contribution to the uncertainty in WS assessment at the scale of the Medjerda catchment	37
3.1	Summary of the Together4Water database.	46
4.1	Summary of CS data validation methods.	63
4.2	Student’s t-test results.	67
5.1	Sample statistics and p -values (p_v) of the Shapiro–Wilk’s test for citizens-based (C_i) discharges at both sites, where left and right parts of the table refer to the Medjez-N5 and Slouguia site, respectively. Values $y_{(1)}$, $y_{(n)}$, $\bar{y}$ and σ are the minimum, maximum, mean and standard deviation, respectively. . . .	88
5.2	REML estimates of σ_m^2 and σ_u^2 at the Medjez-N5 and Slouguia sites, along with their respective contribution (in %) to the total variance $\sigma_m^2 + \sigma_u^2$	90
5.3	Estimated variances S_i^2 of the four citizens-based (C_i) discharge errors around the 1:1 line and the corresponding R_i^2 values, along with the results for the fused measurements (last line). Left and right parts of the table refer to the Medjez-N5 and Slouguia site, respectively.	92
6.1	Summary of SDG 6 indicators.	103
6.2	Comparison of the average, minimum, and maximum of Total Alkalinity, Nitrates and Chlorides among three citizens (numbers between parentheses indicate the number of collected observations by each citizen).	105

LIST OF ACRONYMS

BLUP Best Linear Unbiased Predictor. 15, 72, 92, 113, 116

CS Citizen Science. 8, 42, 58, 72, 96

EFR Environmental Flow Requirements. 19, 25

ERWR External Renewable Water Resources. 26

ET Evapotranspiration. 19

FOUA First Order Uncertainty Analysis. 18

GDWR General Direction of Water Resources. 7, 22, 24, 26, 32, 43, 45, 46

GEE Google Earth Engine. 19, 23

IRWR Internal Renewable Water Resources. 26

IWRM Integrated Water Resources Management. 107

JMP Joint Monitoring Program. 107

LC Land Cover. 19

MDGs Millennium Development Goals. 3

NIM National Institution of Meteorology. 8, 43

NSA National Sanitation Agency. 8, 43, 45, 46

LIST OF ACRONYMS

PCC Pearson Correlation Coefficient. 58, 64

RCAD Regional Commissary for Agricultural Development. 7, 22, 43

REML Restricted Maximum Likelihood. 85

RS Remote Sensing. 19, 38

SDGs Sustainable Development Goals. 3, 18

TDR Time Domain Reflectometry. 9

TFWW Total FreshWater Withdrawn. 25

TRWR Total Renewable Water Resources. 25

UN United Nations. 3, 96

VMF Variable Monthly Flow. 27

WEDNC Water Exploitation and Distribution National Company. 8, 22, 43, 44, 46

WS Water Stress. 18, 19, 38

1

Introduction and Objectives

1.1 Introduction

WATER resources management is pivotal for sustainable development at global and local levels. Access to safe drinking water and sanitation are pillars of human health and well-being. Water is also an essential natural resource supporting all economic activities and is a key factor for maintaining terrestrial ecosystems and associated ecosystem services. Water uses are highly interconnected and potentially conflicting, and intertwined with other sectors such food, energy, and industrial production. Yet, the water system is subjected to many pressures, affecting both of its quantitative and qualitative aspects in many regions of the world (United-Nations, 2015). In the past few decades, the impact of climate change has been reflected more often in shifts in water availability, for example, increasing water scarcity and stress in some regions and flooding in others. Consequently, water plays a key factor in managing global and local climate change-related risks such as famine, disease epidemics, political instability, and natural disasters.

Within these sectors, water is crucial in the implementation of integrated solutions and strategies. However, in many parts of the world, water resources are usually managed by different agencies, administrations, and institutions of governments and within different sectors, with poor coordination among them and without an overall picture and understanding of the state of the resource. This sectoral approach added a problem of coherence between the different water users, where policies and decision-making in one sector could be contradicted or duplicated in another sector. Furthermore, water resources are naturally confined to hydrological and hydrogeological entities such as basins and aquifers, which, from physical and ecological points of view, are the most appropriate for management. Nevertheless, water resources are often managed according to administrative units, institutes, and local agencies, which commonly don't consider water basins, resulting in further fragmentation and conflicts, especially in the case of transboundary catchments.

To ensure sustainable management of water resources at all levels, it is essential to look at the water cycle in its entirety, including all uses and users, and considering the potential impact of climate change. Countries need to consider moving from sectoral management of water resources, to an integrated, more innovative and efficient approach that can balance the different needs and pressures. In September 2015, heads of state from

all around the world adopted the UN 2030 agenda for Sustainable Development (UN-Water, 2018). The Agenda consists of 17 Sustainable Development Goals (SDGs) and 169 targets that should be reached by the global community by 2030 (UN-Water, 2018). The SDGs Agenda has been defined through a bottom-up approach and is intended to be applicable to all members of the UN, either belonging to the developed or developing world. In contrast to the previous Millennium Development Goals (MDGs), a specific goal on water (SDG 6), and 7 other goals indirectly related to this goal, has been included in the Agenda. These goals call on states to address the major issues related to water at the sub-national, national, regional, and global levels (UNESCO, 2016).

By expanding the MDGs, SDG 6 focuses on drinking water and basic sanitation. It includes water, wastewater, and ecosystem resources, which cover all the main aspects related to freshwater in the context of sustainable development. Bringing these aspects together under one goal is a first step towards addressing the different water-related issues and enabling coherent and sustainable management, thus making SDG 6 a major step forward towards a sustainable water future (United-Nations, 2015).

1.1.1 SDG 6: a global framework for Water

The United Nations has defined 8 targets and 11 indicators for SDG 6 to "ensure availability and sustainable management of water and sanitation for all" (UN-Water, 2018) (Figure 1.1). The Targets are set as specific aims of the goal, highly interconnected, sub-categorizing it into different major water-related issues that need to be addressed. On the other hand, the indicators represent the metrics by which the achievement of the targets can be tracked.

The targets 6.1 and 6.2 build on the equitable access to safe and affordable drinking water, and sanitation and hygiene. The two targets, monitored by indicators 6.1.1 and 6.2.1, contribute to a safely managed drinking water and sanitation chain in order to protect the health of individuals and communities, as well as the environment. Water "for all" represents an important share of total water use within a country, region or a basin (target 6.4). Access to safe drinking water is directly dependent on the quality of the water available in natural ecosystems including rainwater, groundwater, water from infiltration wells, and water from bodies such as lakes and rivers (targets 6.2, 6.3 and 6.6). Subsequently, target 6.3 sets out to improve ambient water quality, which is essential to protect both ecosystems and

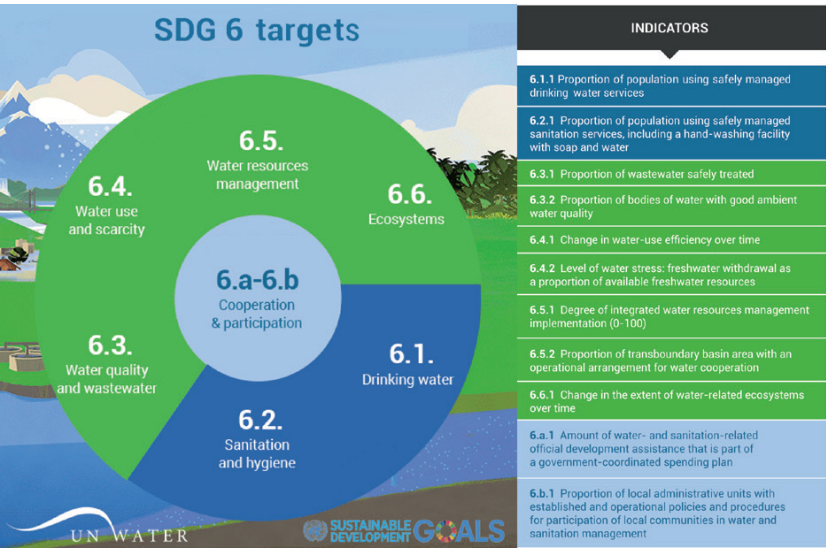


Figure 1.1: SDG 6 targets and indicators. (Source: unwater.org)

human health (Bhaduri et al., 2016). The target is monitored by the means of two indicators, 6.3.1 and 6.3.2, putting an attention on the proportion of wastewater safely treated and the water bodies with good ambient water quality. The focus on water recycling (e.g. recirculating water used in an industry) and reuse (e.g. using wastewater in irrigation) can contribute to the efforts for reducing freshwater withdrawals and increasing water use efficiency (target 6.4).

Target 6.4 addresses water scarcity (indicator 6.4.2), which aims at increasing water availability at both economic and environmental levels, by increasing water-use efficiency across all sectors within a country or a region (indicator 6.4.1) (Vanham et al., 2018). Addressing environmental water requirements, which ensures "fair" water flow in the environment to sustain its natural processes, is vital for maintaining ecosystem's health and resilience (Fehri et al., 2019). Several sectors are highly dependent on water, as a consequence, water-related conflicts arise where water resources are limited and not properly managed. Therefore, target 6.5 aims for the implementation of integrated water resources management (IWRM) at all levels (indicator 6.5.1) and the transboundary water cooperation (indicator 6.5.2) (UN-Water, 2018) by bringing together stakeholders from all sectors, regions, and neighbor countries, which provides a framework to

balance the drinking water consumption and sanitation services (linked with targets 6.1 and 6.2) and the sustainable management of water, wastewater and ecosystems (targets 6.3, 6.4 and 6.6).

Target 6.6 builds on the reduction of the degradation and destruction of the ecosystems (indicator 6.6.1) (United-Nations, 2015). This includes aquatic ecosystems such as rivers, lakes and wetlands, as well as mountains and forests, which are essential as freshwater reservoirs and for maintaining high water quality. Two additional targets were added to SDG 6, referred to as “tools for implementation”. Target 6.a seeks to enhance international cooperation and capacity building (indicator 6.a.1), and target 6.b aims at supporting and strengthening the participation of local communities (indicator 6.b.1) (UN-Water, 2018).

1.1.2 SDG 6 related water challenges in Tunisia

Tunisia lies between the hot desert in the south and the Mediterranean in the north. Similarly to the other North-African countries, Tunisia is dominated by arid and semi-arid climates, and highly affected by climate change and limited water supplies (Bouraoui et al., 2005; Sellami, 2014). Water resources in Tunisia are estimated at 4700 million cubic meters (Mm^3) including 650 Mm^3 of nonrenewable resources (13.8% of the total water resources). Groundwater resources represent 42.5% of the total potential (DGRE, 2016).

Under the variability of climate, influenced by the Mediterranean in the north and the Saharan domination in the south, rainfall is considered scarce and unequally distributed in space and time. Tunisia receives on average 230 $mm/year$ of rainfall (DGRE, 2016). The annual precipitation varies on average between 600 $mm/year$ in the north and 156 $mm/year$ in the south. The ratio between the highest and the lowest observed values of precipitation vary from 4.4 in the north to 15.8 in the south (NRC, 2007), which illustrates the temporal variability of rainfall in the country (DGRE, 2016).

Surface water resources are estimated at 2700 Mm^3 spatially distributed per year over three natural areas distinguished by their climatic and hydrological conditions (NRC, 2007). The northern region provides up to $2190 \text{ Mm}^3/\text{year}$, thus representing 82% of the total surface water potential while covering only 16% of the country's area (DGRE, 2016). Most of the northern water resources are located in the Medjerda Catchment, which represents the largest water resource in Tunisia. Around 60% of the population is provided with drinking water from the Sidi Salem reservoir (770 Mm^3) located in the catchment (Bouraoui et al., 2005; Fehri et al., 2019). The center part (22% of the country's area), is characterized by inconsistent water resources. It contributes to the total surface water potential by 12%. On the other hand, in the southern part of the country (approximately 62% of the total area), surface water is considered as a scarce resource, providing very irregular volumes evaluated at $190 \text{ Mm}^3/\text{year}$ which represents only 6% of Tunisia's total water potential (NRC, 2007).

Surface water quality, evaluated by its degree of salinity, is essential for water management in the country as it varies depending on the origin of the resource. In Tunisia, a salinity of less than 1.5g/l is considered as acceptable for freshwater, which, according to ANPE (2018), accounts for approximately 72% of the surface water. The spatial distribution of water quality is highly variable across the country, 82% of the water resources in the north is considered in good quality, compared to only 48% and 3% of the available water, respectively, in the center and the south (ANPE, 2018).

The groundwater resource in Tunisia is estimated at 2000 Mm^3 , situated within 212 shallow aquifers and 267 deep aquifers (ANPE, 2018). The spatial distribution of groundwater varies across the country, 55% and 18% of the shallow and deep aquifers, respectively, are located in the north. The center provides 30% of the shallow resources and 24% of the deep resources, while the southern part provides 15% and 58% of the shallow and deep resources, respectively (DGRE, 2016). Nevertheless, groundwater quality is considered relatively poor, as only 8% and 20% of shallow and deep aquifers, respectively, are considered in good quality. Furthermore, according to ANPE (2018), water with salinity up to 3g/l can be used for agricultural purposes and for drinking water production. Therefore, approximately 36% of groundwater resources are considered not suited for these two major sectors with increasing demand in the last two decades. The inequality in quantity and quality of water resources in Tunisia makes the management more complex and reflects the need to transfer surface water from the north to other parts of the country (NRC, 2007).

Various pressures are exerted on the available water resources in the country. Agriculture, services (mainly drinking water), and industry are considered as the main water consuming sectors in Tunisia. Irrigation constitutes the largest consumer of water, using 80% of the total water potential in the country. On the other hand, withdrawals for drinking water purposes represent 16% while the industrial sector consumes 4% of the total water potential (ANPE, 2018). The political changes in Tunisia that began in the last decade added more pressure on water resources, in particular with respect to the systems and legislation of water management in the country as the illegal water withdrawals has been increasing (Fehri et al., 2019) as well as the local conflicts between water users and sectors.

This water situation in Tunisia requires a consistent approach for better management strategies at local and national levels. Thus, as one of the members of the United Nations, Tunisia adopted the SDGs, in particular SDG 6. It is expected that with this adoption, Tunisia will reinforce its efforts for an integrated water management across the country. However, so far, no scientific based operational strategy exists to monitor the implementation of SDG 6.

The targets and indicators presented in Section 1.1.1 should be monitored at different scales in the country in order to improve the understanding of the current situation of water resources and to identify the needed actions that can enhance the resilience and integrity of ecosystems and biological diversity, and establish a strategic plan to move from increasing water production to demand management through new water monitoring systems, and legal and institutional cooperation mechanisms.

1.1.3 The current state of water-related data in Tunisia

The implementation of SDG 6 in Tunisia should consider the specificity of current and future water data structure, quality and availability at national and regional levels. Indeed, building reliable water-related databases ensures the best measurement of SDG 6 indicators. Therefore, establishing systems for efficient monitoring and transparent sharing and validation of data is critical for reinforcing the link between the national and regional measurements of the indicators. The Tunisian water-related information is monitored and provided by multiple institutions and agencies. The General Direction of Water Resources (GDWR) and its regional agencies (Regional Commissary for Agricultural Development (RCAD)) are the official governmental institutions that monitor water-related information at

both urban and rural regions. Data such as river flow, precipitation, land cover/land use, groundwater, and dams' statistics are collected locally and aggregated at national scale. However, the quality of the collected data is still relatively low. Further improvements are required with respect to data gaps, incomplete time series and uncertainty (Fehri et al., 2019). Drinking water availability and distribution are monitored by the Water Exploitation and Distribution National Company (WEDNC) at the regional scale. The National Sanitation Agency (NSA) provides data about sanitation networks, the proportion of population linked with sanitation facilities, and wastewater treatment. The quality of the WEDNC and NSA data is considered acceptable, yet, the complex administrative system within both organizations makes the access to data difficult and time consuming.

Other sources like the National Institution of Meteorology (NIM) monitor meteorology across the country at different spatial and temporal resolutions with high quality. However, the access to data is not free and restricted to specific governmental agencies.

In addition to governmental data, remote sensing has been evolving in the last two decades as a reliable source of water-related information, with high spatial and temporal resolutions. Remote sensing data can be used as a cost-effective surrogate for in-situ data. For instance, researches found in the literature used remote sensing evapotranspiration products to estimate water consumption over different ecosystems and irrigated lands (Fehri et al., 2019; Yang et al., 2018). Moreover, Van Dijk et al. (2016) and Tarpanelli et al. (2011) have been using satellite images to develop frameworks that estimate discharge at different rivers in many parts of the world, which may be a robust complement to the existing databases as well as a unique source of data where expert discharge data are not available.

1.1.4 Citizen Science to support the monitoring of the implementation of SDG 6

Smartphone technology is being widely used in science to collect water-related information such as precipitation, river discharge, water level, and land cover, in the framework of a new innovative participatory approach "Citizen Science (CS)" (Carrel et al., 2019; Fehri, Khlifi, & Vanclooster, 2020; See et al., 2015). Dickinson et al. (2012) define CS as the engagement of non-professionals in authentic scientific research, while Buytaert et al. (2014) described it as the participation of the general public (i.e., non-scientists) in the generation of new scientific knowledge. In our perspective, CS is

a complex process built on civic engagement into science, environmental monitoring, and capacity building. Nevertheless, the definition of CS depends mainly on the specific goals of the project, i.e. whether it aims at engaging volunteers for data collection (which is known as crowdsourcing) or empowering and including citizens in knowledge generation activities, training, and capacity building, which is considered as a more advanced form of CS (Pocock et al., 2019). In addition, Roy et al. (2012) provide a classification of CS based on the context of the projects. Three main approaches are recognized: contributory projects (entirely designed by scientists while citizens are mainly participating in data collection), collaborative projects (also designed by scientists while citizens are more involved in the scientific process such as collecting and analyzing data), and co-created projects (designed in collaboration between scientists and citizens as both work together in partnership). Whatever the exact definition one is using, participation and empowerment of the general public at any level within a region or a country suggests that everyday citizens have the potential to deliver timely and cost-effective alternatives to traditional data collection (Walker, Forsythe, Parkin, & Gowing, 2016).

Although it is not a brand new concept, CS is becoming increasingly adopted in natural science research as a cost-effective way of data collection and knowledge generation through the engagement of population into science. Fritz et al. (2019) suggest that the required data for monitoring the different SDG 6 indicators can partially be delivered through a CS approach. Indeed, many hydrological and water resources studies that built on CS have shown promising monitoring techniques and results that can help in measuring some of the previous indicators. Canfield et al. (2002) showed promising outcomes of monitoring water quality through a CS approach in the framework of the Florida LAKEWATCH CS project. citizen scientists were able to provide strongly correlated water quality and clarity data to the mean values obtained by professionals. Soil moisture volunteered measurements studies using Time Domain Reflectometry (TDR) sensors, are used to detect moisture fluctuations in the ground after precipitation events (Buytaert et al., 2014) which can serve in irrigation and many ecosystem contexts. Vegetation characterisation studies using geo-tagged mobile photos are highly developed for field verification of vegetation types and change and identification of ecosystem stability (Buytaert et al., 2014).

In the same context, See et al. (2015) and Fritz et al. (2012) developed a volunteers-based study to produce land cover and agricultural field size

maps using the Geo-Wiki smartphone application. Precipitation measurements are also being largely used in a CS context despite its sensitivity to the local aerodynamic conditions and non-standard locations and calibration of the rain gauges. The Rainlog project for example is a promising cooperative initiative for monitoring rainfall in Arizona state by a large network of citizens using affordable rain gauges (Woodward, 2016). Stream flow velocity measurements are also under development in a CS context based on indirect measurements methods. However, recent researches showed that smartphone applications based on particle image velocimetry can be used to measure surface velocity and discharge of rivers (Carrel et al., 2019; Kim et al., 2008; Kwonkyu et al., 2015; Lüthi et al., 2014; Weijs & Chen, 2014).

The application of CS allows increasing the amount of collected data in a given region. The validation process is crucial in order to approve the validity of the data to be used in science and decision making. In some cases, multiple measurements can be made at the same location by different citizen scientists such as river discharge, precipitation, and water quality using various monitoring tools. Therefore, data assessment and integration frameworks are needed to improve the quality of CS-based data and to identify the errors and the level of coherence with the reference sources of information such as governmental databases. In this regard, data fusion techniques are being largely used as tools to combine various data with different qualities into a single optimal result (Bogaert & Fasbender, 2007; Fasbender et al., 2009; Mattern et al., 2012).

Hence, the application of CS concepts in the monitoring of water-related data in Tunisia represents an opportunity to empower communities to take actions themselves and change their attitude to the environment. Successful civic engagement including universities, schools, and NGO's is key to a consistent data collection which may support the implementation of SDG 6 in the country by reinforcing the existing databases as well as improving the water-related monitoring systems.

1.2 Objectives

THE main overall goal of this PhD thesis is to strengthen the scientific basis for an operational monitoring of the implementation of the SDG 6 in Tunisia. The monitoring system must be coherent with the indicators' framework that has been agreed upon at UN level, but should consider the specificity of current and future water data structure, quality and availability. Therefore, the thesis also aims to boost water data availability in the near future through exploring the potential of Citizen Science approaches.

To address this main goal, four specific objectives are defined. Their integration in a common framework is schematically represented in Figure 1.2.

Specific objectives

SO1: To develop a scientific-based method allowing to disaggregate an SDG 6 indicator at different spatial and temporal scales in Tunisia

As described in the previous section, many studies recommend the measurement of SDG 6 indicators at different levels within a country or a region (Vanham et al., 2018) in order to better understand its behavior as well as to support well placed actions for water management. However, so far, there are no scientific-based initiatives to monitor the implementation of SDG 6 in Tunisia at the local scale. The current state of water-related data, including its availability and structure, is still an issue and might be problematic for the monitoring of certain SDG 6 indicators, in particular at the local scale. The first specific objective of this thesis is therefore to develop a scientific-based method allowing to disaggregate the water stress indicator (6.4.2) at different temporal and spatial scales in Tunisia using the available data (official data and remote sensing products). This first objective will be developed for the case study of the Medjerda basin. Remote sensing products and governmental data will then be used to disaggregate the water stress indicator at different scales in the catchment.

Chapter 2

SO2: To design a Citizen Science-based approach for monitoring water-related data in Tunisia

Given the lack of coordination between the different Tunisian administrations and institutes that collect and document data, the access to informa-

tion is complex. Moreover, the quality of the available data is considered poor (data gaps, high uncertainty) and require further data assimilation procedure. In this thesis, the second specific objective is therefore to (i) perform a data collection campaign to create a consistent water-related database for the study region. A dedicated online platform will then be implemented for data visualisation and standardization, which will improve the quality of the information; (ii) and design a citizen science-based approach as a new way of water monitoring allowing to improve data quality and availability. A step-by-step citizen science method will be adopted to engage, train, and motivate citizens in the Medjerda catchment. Discharge and rainfall will be monitored by volunteers using cost-effective tools: smartphone application and rain gauges.

Chapter 3 | Chapter 4

SO3: To develop a data fusion framework allowing to combine various citizen science measurements

The implementation of citizen science will increase number of measurements of the target parameters, particularly discharge, in the test region. In some cases, different citizens will perform measurements at the same location, which will generate similar time series for the observed parameter. Hence, the third objective of this thesis is to develop a framework based on data fusion techniques in order to combine citizens measurements at the observation location and generate a single optimal result.

Chapter 5

SO4: To assess the potential contribution of Citizen Science to the SDG 6 indicators in Tunisia

The fourth and final objective of the thesis is to present an assessment of the potential contribution of citizen science data in measuring the SDG 6 indicators in Tunisia at the scale of the Medjerda catchment. Many studies reviewed the actual and potential inclusion of citizen science data in calculating the indicators at different scales as a part of the integrated monitoring of SDGs at global and national levels (Fraisl et al., 2020; Fritz et al., 2019).

Chapter 6

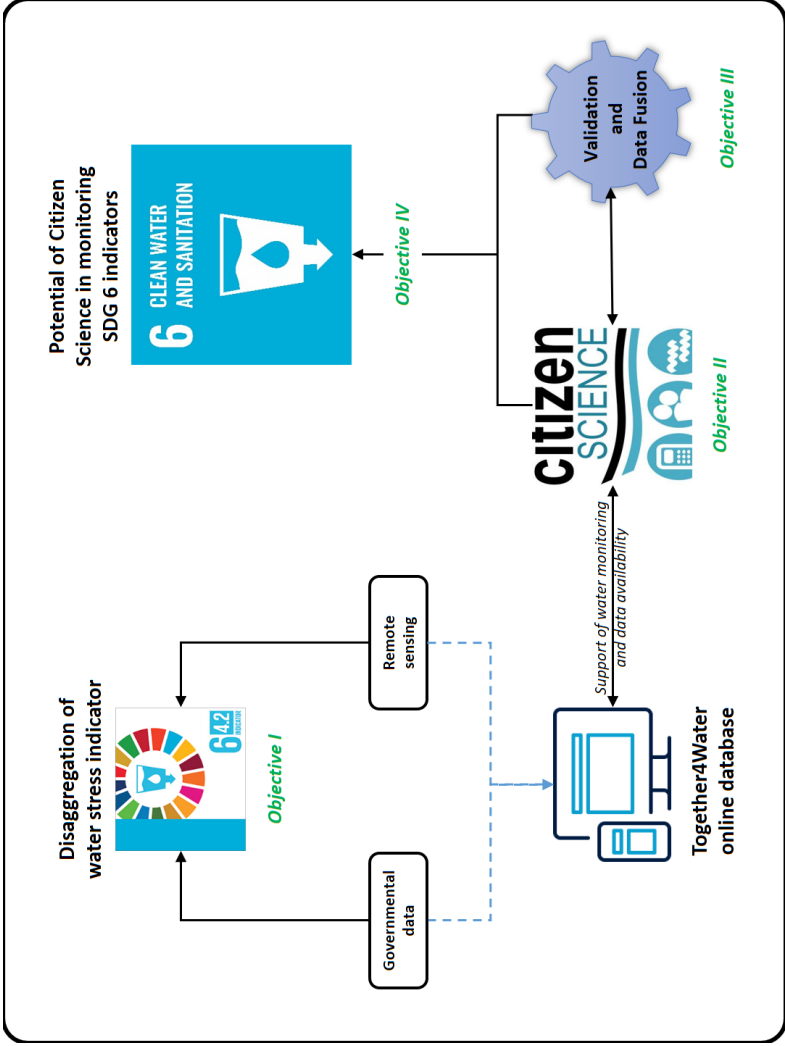


Figure 1.2: Specific objectives of the PhD thesis represented in their general context

Specific research questions

We also defined five research questions (RQ) that help us progressing in the way towards the implementation of SDG 6 in Tunisia. Their interrelations are explained in the flowchart of Figure 1.3.

RQ1: Does the disaggregation of the SDG 6 water stress indicator (6.4.2) at different temporal and spatial scales provides a better understanding of water stress issues?

The indicator 6.4.2 of SDG 6 measures the actual situation of water stress at the scale of a country or region. At a UN level, it is recommended to measure the indicator nationally at yearly scale. However, we show that the disaggregation of WS indicator at different temporal and spatial scales is crucial and may deepen the understanding about its regional and intra-annual variability.

RQ2: How should water-related data be collected and customized to improve its quality and availability?

Data availability and quality are crucial for monitoring SDG 6 indicators. The complex data access systems in Tunisia and the poor quality of the existing available databases question the monitoring of SDG 6 indicators in the country. Therefore, a citizen-based water monitoring project "Together4Water" was launched in the study region to support the existing database. A dedicated standardized online platform was created to facilitate citizen-generated data and to ensure dynamic visualisations of the information including historic environmental data collected from the official Tunisian water institutions.

RQ3: What is the potential intrinsic value of water information data collected in Tunisia using Citizen Science-based approaches?

Citizen science has been evolving as a novel approach of data collection and capacity building in environmental science. The major question is directly related to the value of the collected information and its pertinence in science and decision making at national and local levels. Therefore, citizen science has been tested in Tunisia in a test region in order to assess its social and scientific value.

RQ4: How can citizen science-based water information data optimally be combined with existing reference data from different water-related services?

In order to improve the quality of citizens' measurements, a data fusion approach can be used as a technique for combining various citizen science data. To do so, the Best Linear Unbiased Predictor (BLUP), a widely used statistical data fusion technique, has been evaluated in this context. The fused data has been compared to the reference observations to assess its quality.

RQ5: Whats the potential contribution of Citizen Science data to the monitoring of SDG 6 indicators in Tunisia ?

In the framework of the integrated monitoring of the SDGs, particularly SDG 6, at global, national, and local levels, many studies at a UN level showed the actual and potential use of various CS data to measure the indicators. Therefore, for the case of the Medjerda catchment in Tunisia, we will present an assessment of the potential contribution of citizen-based observations to the monitoring of SDG 6 indicators.

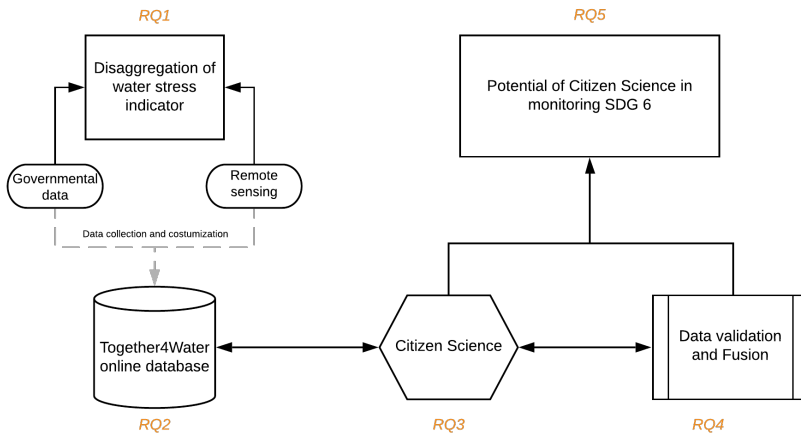


Figure 1.3: Flowchart of the research questions investigated in this PhD

2

Disaggregation of SDG-6 Water Stress indicator

This chapter is a modified version of a published manuscript in *Science of The Total Environment*

Fehri, R., Khelifi, S., and Vanclooster, M. (2019). Disaggregating SDG-6 waterstress indicator at different spatial and temporal scales in tunisia. *Science of The Total Environment*, 694:133766.

Abstract

THE recently adopted UN Sustainable Development Goals (SDGs) encompasses a specific goal for water (SDG-6). The target 6.4 deals with water scarcity and refers to two main indicators: water use efficiency and Water Stress (WS), monitored by the UN statistical services yearly at the country level. Yet, for more efficient development planning, indicators should also be provided with higher spatial and temporal resolutions. This study presents a data-driven method allowing to disaggregate the WS indicator at higher spatial and temporal resolution. We applied the method for the Medjerda catchment in Tunisia, known as being severely water-stressed. We disaggregated the WS indicator from the overall catchment to the administrative regional level at yearly and monthly scales. In order to overcome poorly documented irrigation water withdrawals, two approaches were adopted: 1) we used yearly governmental data at both catchment and regions scales; 2) we replaced governmental irrigation data by remote sensing-based irrigation estimation. First Order Uncertainty Analysis (FOUA) was performed to characterize the uncertainty associated with the assessment of WS. Results reveal that the WS at the scale of the catchment increases considerably in recent years, exceeding 50% from 2005 and surpassing the 100% threshold in 2015 and 2016 (102%, 108% respectively). The two adopted approaches result in similar WS trends. However, the second approach yields higher WS values compared to the first approach (144% versus 108% in 2016). The monthly-disaggregated WS at catchment scale exhibits a similar increasing trend. The highest WS values are at the end of the fall and during the summer season, which is mainly due to the increasing demand for irrigation and drinking water. Siliana region is the most affected by WS, while Beja is the least affected. The FOUA shows that the integration of remote sensing-based irrigation data reduces the WS uncertainty.

2.1 Introduction

IN the past 50 years, water has become a limited resource in many nations all over the world. Nowadays, about 4 billion people face severe water stress at least one month per year, in particular in arid and semi-arid regions like in North Africa (Mekonnen & Hoekstra, 2016). According to FAO (2018), North African countries face severe water scarcity, which exerts a significant burden on the overall development of these countries.

Given the global water crisis, a specific goal dedicated to water (SDG-6) has been integrated in the recently adopted UN 2030 Agenda for Sustainable Development. The different targets and indicators of SDG-6 are linked to all water functions and services and considered to be important for sustaining life on Earth. Indicator 6.4.2 of target 6.4 measures the level of Water Stress (WS). It is a blue water WS indicator. It is defined as the ratio of total yearly national freshwater withdrawn by all sectors to the total of fresh blue water resources minus the Environmental Flow Requirements (EFR) (Vanham et al., 2018). The quality of the assessment of this indicator depends primarily on the quality of available water data. Governments generally use traditional hydrometric data to assess the WS indicator. Yet, traditional hydrometric data archives suffer from data gaps, incomplete time series, poor spatial resolution, often leading to poor quality databases.

Alternatively, Remote Sensing (RS) offers hydrometric attributes with a high space-time resolution that may complete the poorly monitored national data such as precipitation, Land Cover (LC) and irrigation consumption. Yang et al. (2018) showed for instance that RS data could be used to estimate water consumption of ecosystems including arable lands using different Evapotranspiration (ET) products such as MODIS. RS data can, therefore, be used to reduce the uncertainty in assessing the WS indicator of SDG-6. In addition, with the availability of new RS products and RS processing platforms such as the Google Earth Engine (GEE), a capacity exists to downscale SDG-6 indicators at higher spatial and temporal resolutions. This offers an opportunity to address the recommendation of Vanham et al. (2018) stating that more advanced monitoring levels of the WS indicator are needed. It offers also perspectives to support more efficiently regional water management design and planning. Yet the processing streams of RS based methods for assessing SDG-6 indicators have not been tested, and the final quality of the disaggregated indicators is not yet known.

The objective of this study is to present a data-driven method allowing to evaluate and disaggregate the SDG-6 indicator using traditional and RS based data products. We evaluate the method for the case study of the Medjerda catchment in Tunisia. Tunisia is considered as one of the most arid countries in the Mediterranean region, suffering from extreme water scarcity in recent years (FAO, 2018). Tunisia's Medjerda catchment represents the most important river basin in the country, however, it is highly affected by climate change and the pressures on its water resources are high. We assess the WS indicator at the scale of the entire Tunisian catchment and at the scale of four administrative regions (the "governorates") within the catchment (Jendouba, Beja, El kef, and Siliana), both at the yearly and monthly time scales. Annual assessment of the WS indicator might not be sufficient enough to obtain insights into its intra-annual variations. Therefore, the disaggregation of the indicator at a monthly time scale is useful. The monthly assessment provides more knowledge about WS seasonality and its intra-annual variability which are not revealed by assessment at an annual temporal resolution (Degefu et al., 2018; Gain & Wada, 2014). We propose two data-driven approaches: (1) measurement using nationally produced data; and (2) measurement based on ET-derived irrigation water consumption as a substitute for governmental irrigation data. We include First Order Uncertainty Analysis (FUOA) to characterize in each approach the uncertainty in the measurement of WS resulting from uncertainties in the input data.

2.2 Materials and methods

2.2.1 Study area

Tunisia lies between the hot desert in the south and the Mediterranean in the north. It is dominated by arid and semi-arid climates and affected by limited water supplies. Most of the country's water resources are concentrated in the Medjerda river basin, which represents the most important catchment in the country. The upstream is located in the semi-arid Atlas Mountains of eastern Algeria and it runs over a distance of 312 km in the North of Tunisia. The study area lies in the sub-humid to the Mediterranean humid bio-climatic zone where the average annual rainfall is about 600 mm. It covers a total area of 23,700 km², of which 7,870 km² (33%) are located in Algeria. The Medjerda river meets with "Mellegue River" com-

ing from Algeria and its confluence is located in Jendouba Governorate in Tunisia, just before Sidi Salem Dam. This latter dam is one of the largest dams in the country (Fig 2.1). The population within the Medjerda river basin was estimated to be 2 million in 2014 (14% of country's population, (INS, 2014), while the basin occupies 9.8% of the land area of Tunisia. Agriculture represents the basin's major economic activity.

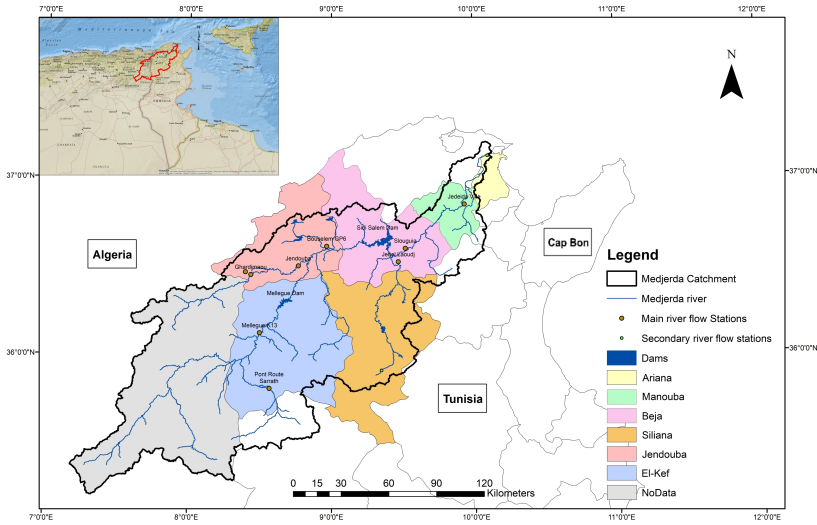


Figure 2.1: The study area. The right down figure is the situation of the study area, while the inset yields the location of the study area in North Africa.

In fact, most of the water resources are used for agricultural purposes (84%) and only about 10% is used for services and industries (Bouraoui et al., 2005). According to the Ministry of Agriculture (2005), the area of irrigated lands in the Medjerda is rapidly increasing over the years. It increased from 49,000 *ha* in 1987 to 260,000 *ha* in 2005. Irrigated lands are estimated to represent 9.4% of the total basin area putting considerable pressure on the available water resource very often leading to water scarcity.

In addition to water shortage, water quality issues exert an additional burden on the water resource. The salinity of both surface and ground water, reduce the water quality in particular in the downstream coastal part of the catchment. Soil erosion further affects the long-term storage, capac-

ity and water quality of dams. Additional pollution originates from urban areas, industries, and agriculture practices that intensively use fertilizers and pesticides (Bouraoui et al., 2005).

The Medjerda basin has a major economic importance in the country, it encompasses six governorates: Ariana, Mannouba, Beja, Siliana, Jendouba, and El Kef. The latter four Governorates are considered as the most “economically” and “hydrologically” sensitive regions of the Medjerda. The basin’s contribution to the Tunisian economy is significant. According to the INS (2014), 51% of total national wheat production is cultivated in the basin, and around 60% of the population is provided with drinking water from the Sidi Salem reservoir.

2.2.2 Data sources

Data on river flow, available groundwater, and withdrawals for the irrigation, domestic and industrial sector are needed to assess the WS indicator for the Medjerda catchment. Monthly and yearly data including averaged river flow of 9 hydrological stations (Fig 2.1), groundwater recharge, irrigation, and industrial water withdrawal for all 6 governorates of the Medjerda basin for the period of 2000-2016 were collected from the bulletins of the General Direction of Water Resources (GDWR) and the compilations made by the Regional Commissary for Agricultural Development (RCAD). Withdrawals for domestic and industrial consumption for the same period and at the same temporal and spatial scales were collected from the Water Exploitation and Distribution National Company (WEDNC).

Land cover (CCI-Landcover) and MODIS-ET data (Figure 2.2) where combined to estimate monthly and yearly water consumption in the irrigation sector. Yearly Land Cover (LC) data at a spatial resolution of 300m were provided by the ESA Climate Change Initiative (ESA-CCI, CCI-LC) (available at <http://maps.elie.ucl.ac.be/CCI/viewer/>) for the period of 1992 to 2015. The dataset is based on the MERIS FR and Reduced Resolution (RR) archive, and uses Advanced Very High-Resolution Radiometer (AVHRR), SPOT-Vegetation (SPOT-VGT) and Project for On-Board Autonomy (PROBA-V, with the V standing for Vegetation) for LC change detection (Bontemps et al., 2013) sensor data.

The MODIS-ET is a product of the National Aeronautics and Space Administration (NASA) under a project to estimate global terrestrial ET from remote sensing data. It is based on the improved ET algorithm by Mu, Zhao, and W.Running (2011) which uses the Penman-Monteith equa-

tion (Monteith, 1965). Data were extracted using the Google Earth Engine (GEE) at 500m spatial resolution and at 8-day time step for the entire Medjerda catchment. Data were aggregated at monthly and yearly scales in order to match the time scale of the WS indicator. All collected data were included in a well structured database within the Together4Water online platform, which allows dynamic visualisations of the different time series and maps (Chapter 2).

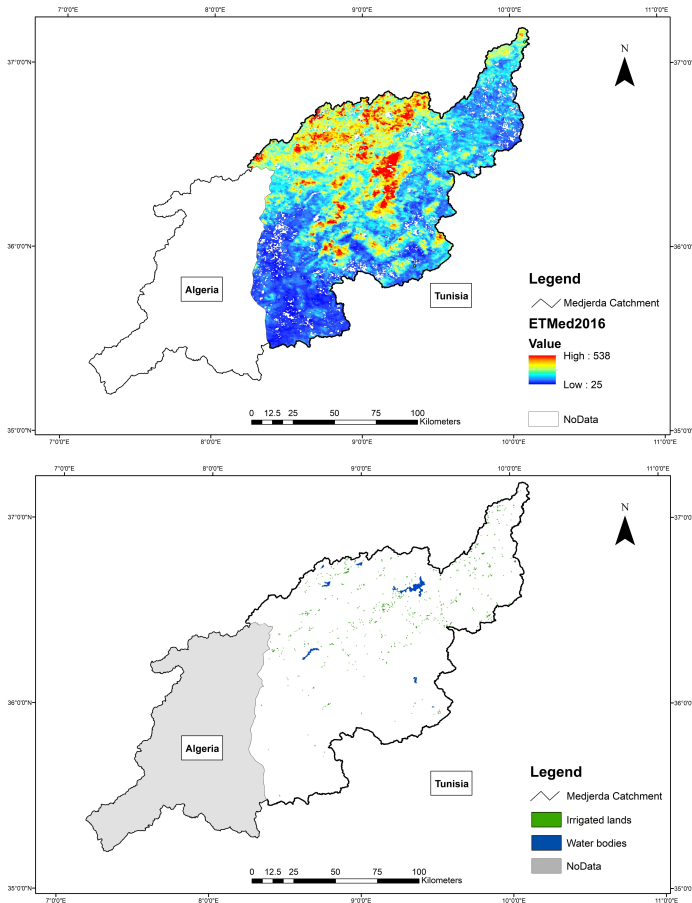


Figure 2.2: Map of MODIS-ET (April 2016) and CCI LandCover 2015 in the Medjerda catchment.

2.2.3 Estimation of irrigated lands' water consumption

At the scale of the Medjerda catchment, the information about the monthly irrigation water withdrawal is not consistent and limited as it doesn't cover all the Medjerda regions, only yearly data were delivered by the GDWR for all the Medjerda regions. Yang et al. (2018) used different ET products to estimate water consumption of different ecosystems at a basin scale, considering ET as a substitute for water consumption. We, therefore, adopted ET in irrigated land as a surrogate for irrigation water consumption. This will improve the calculation of WS indicator at different temporal and spatial scales, in particular at monthly scale for the whole catchment and its subregions.

Irrigated land was estimated from the land cover map (CCI-LC 300m). Land cover was considered to be constant within a year, and we adopted the land cover of 2016 as the same land cover for 2015 due to the unavailability of the 2016 CCI-LC product at the time of processing. ArcGIS software was used to assess yearly and monthly mean ET over irrigated lands. The LC-300m and ET-500m images (Table 2.1) were resampled in a 100m frame in advance in order to avoid processing errors. Yearly and monthly water consumption of irrigated lands in the entire basin and the regions within was determined from the product of the average water consumption per pixel and the number of pixels (Yang et al., 2018).

$$U_{Irrig,RS} = V_{pp} \times N_p \quad (2.1)$$

with $U_{Irrig,RS}$, estimated water consumption of irrigated lands; V_{pp} , average water consumption per pixel, and N_p , number of pixels.

Irrigation efficiency in the Medjerda region was taken into account in this assessment, irrigation water, which have been lost through drainage systems, infiltration, and through leakage in the irrigation networks should be added to the volume of evapotranspiration from irrigation lands in order to have a realistic estimation of irrigation water consumption in the Medjerda region. According to Dhahbi (2018), irrigation efficiency at the perimeter level in the Medjerda catchment ranged from 60% to 45% during the 2000-2016 period, which means 40% to 55% of irrigation water was lost in the system. The irrigation network in the basin performed well in the early 2000s with an efficiency of 60% as it was newly operational following the launch of new irrigated lands. However, in recent years, the efficiency decreased to reach 45% due to the lack of maintenance of the ir-

rigation network and the absence of operational systems of irrigation water management (Dhahbi, 2018). Hence, this irrigation efficiency range (60% to 45%) was added to the estimated volume of evapotranspiration at the scale of the irrigated lands of the Medjerda catchment and its subregions with a decrease of 1% each year of the 2000-2016 period. Although there is no available official assessment of monthly irrigation efficiency in the Medjerda, we adopted for the monthly scale the same efficiency as each year of the study period (2000-2016).

Table 2.1: Summary of the requirements for the estimation of irrigation water consumption in the Medjerda

Products & Parameters	Description	Spatial resolution	Temporal resolution
MODIS-ET	MODIS actual ET product	500m	2000-2016
CCI-LC	ESA CCI Landcover product	300m	1992-2015
Irrigation efficiency	Decreases from 60% to 45% (1% step)	-	2000-2016

2.2.4 SDG-6 water stress indicator

The WS indicator measures the level of water stress and is defined as the freshwater withdrawal as a proportion of available freshwater resources (WHO, UNISEF, & JMP, 2016). It is computed as follows:

$$WS(\%) = \frac{TFWW}{TRWR - EFR} \times 100 \quad (2.2)$$

with TFWW, the Total FreshWater Withdrawn ($m^3/year$); TRWR, the Total Renewable Water Resources ($m^3/year$); and EFR, the Environmental Flow Requirements ($m^3/year$).

The TFWW represents the sum of the water withdrawal of the three main sectors in the Medjerda catchment, i.e. the agricultural, the services (drinking water distribution) and industrial sector. In this research, we focus on the agricultural sector as the irrigation water consumption, it represents the major water consumer in the basin. The domestic and industrial sectors exert pressure on the water resources of the catchment as well. Domestic water demand increases significantly at the end of the spring and during the summer season as the available water resources in dams and

easily accessible groundwater decrease. Similarly, the industrial water use in the catchment is significant during the same period of the year as most of the existing industries are agri-food industries. They reach the peak of water consumption during the summer season. The TRWR refers to the maximum theoretical yearly amount of renewable freshwater available for the basin and the regions within. It is expressed as the sum of Internal (IRWR) and External Renewable Water Resources (ERWR) (FAO, 2017):

$$TRWR = IRWR + ERWR \quad (2.3)$$

IRWR is defined as the sum of the Surface Water Produced Internally, SW_p , and the groundwater recharge, GW_R . The exchange between surface water and groundwater ($Q_{overlap}$) is subtracted to avoid overestimation of TRWR (FAO, 2017). SW_p represents the sum of average annual of monthly exploited water of dams and the flow of rivers generated internally in the basin measured by official stations in each governorate located in the basin. GW_R is delivered by the General Direction of Water Resources (GDWR) at yearly and monthly scales. GW_R is estimated by the GDWR based on their piezometric archive for each aquifer in the basin. In the case of the Medjerda basin, $Q_{overlap}$ is set to 30% of the groundwater recharge according to the hydrogeological archives of the DGRE. IRWR is computed then as:

$$IRWR = SW_p + GW_R - (Q_{overlap}) \quad (2.4)$$

ERWR represent the part of average annual or monthly renewable water resources that are not generated in the basin (FAO 2017). The equation is then:

$$ERWR = SW_{IN} - SW_{OUT} \quad (2.5)$$

with SW_{IN} , the volume of surface water entering the basin or a region per time step ($m^3/year$); and SW_{OUT} , the volume of surface water leaving the basin or a region per time step ($m^3/year$).

At the scale of the entire basin, SW_{IN} encompasses the average annual or monthly flow entering the basin from Ghardimaou station right at the border with Algeria and from the Mellegue station (30 km away from the border), which measures the flow of Mellegue river entering the basin before connecting with the Medjerda. We did not perform a specific modeling to estimate the discharge of Mellegue river right at the border with Algeria given the significantly low flow ($0.10 m^3/s$ average annual) com-

ing from the tributaries before the Mellegue station (Sarrath river and two small streams).

The EFR is defined as the quantity, quality, and timing of water flows needed to maintain the components, functions, processes, and resilience of aquatic ecosystems that provide goods and services to people where flows are regulated (Hirji Richard, 2009). The explicit inclusion of EFR in the equation of SDG-6 WS indicator increases the reliability of the indicator and its usefulness as a tool to enhance informed water management decisions (FAO, 2017) and it is a key requirement for maintaining freshwater populations and habitats (Vanham et al., 2018). The quantification of EFR can be achieved through different methods grouped in literature into three categories: hydrological, hydraulic-habitat and holistic methods (European Commission, 2015; Pastor et al., 2014; Sood et al., 2017; Tharme, 2003). According to Vanham et al. (2018), hydrological methods are considered to be highly recommended at the planning level of water resource management. In our case, we adopted the hydrological method to quantify the EFR in the Medjerda basin. Pastor et al. (2014) proposed that, globally, EFR for “fair” ecological conditions ranges between 25% and 46% of the mean annual flow. For the annual assessment of the WS, we propose EFR as 45% of the mean annual flow according to the hydrological bulletins of the DGR. For the monthly assessments, we adopted the “Variable Monthly Flow (VMF)” method (Fasel et al., 2016; Pastor et al., 2014), which considers the classification of months into three groups: the low, intermediate and high-flow seasons. These groups are based on the monthly flow variability compared to the mean annual flow. During low-flow months (Jun, July, August, and September), we fixed EFR as 65% of the mean monthly flow. For the intermediate period (October, November, December, and January) we adopted 45% and for the high-flow season (February, March, April, and Mai) EFR was fixed as 30%. These definitions of EFR represent a default environmental protection policy implying the lowest level of recommended protection from a water management perspective (Smakhtin, Revenga, & Döll, 2004). Substituting Eq 2.3, Eq 2.4, and Eq 2.5 in Eq 2.2 yields:

$$WS(\%) = \frac{U_{Irrig} + U_{Indus} + U_{Serv}}{(SW_p + GW_R - (Q_{overlap})) + (SW_{IN} - SW_{OUT}) - EFR} \times 100 \quad (2.6)$$

with U_{Irrig} , the irrigation water use ($m^3/year$); U_{Indus} , the industries water

use ($m^3/year$); and U_{Serv} , the water use for other services ($m^3/year$).

In order to improve the calculation method of WS indicator at different temporal and spatial scales compared to the existing approaches (UN-Water, 2015), two methods were adopted based on two different data sources: a method using only governmental data and a method based on the estimated water consumption of irrigated lands ($U_{Irrig,RS}$) as a surrogate for governmental irrigation data. Therefore, Eq 2.6 yields:

$$WS(\%) = \frac{U_{Irrig,Gov} + U_{Indus} + U_{Serv}}{(SW_p + GW_R - (Q_{overlap})) + (SW_{IN} - SW_{OUT}) - EFR} \times 100 \quad (2.7)$$

with $U_{Irrig,Gov}$, governmental irrigation data.

and,

$$WS(\%) = \frac{U_{Irrig,RS} + U_{Indus} + U_{Serv}}{(SW_p + GW_R - (Q_{overlap})) + (SW_{IN} - SW_{OUT}) - EFR} \times 100 \quad (2.8)$$

Eq 2.7 and 2.8 of WS were used to compute the indicator at yearly scale for the whole catchment and its subregions. For the temporal disaggregation, we used Eq 2.8 to measure monthly WS at the same spatial scales.

2.2.5 First-order uncertainty analysis

FOUA is a widely used technique for estimating the uncertainty in a deterministic model generated by the uncertainty in the parameters (Loague et al., 1989). The theoretical basis and limitations of FOUA are detailed and described by Cornell (1972). In order to determine the total uncertainty in the WS indicator at yearly scale through FOUA, uncertainty measurement for each parameter is required. In this study, the uncertainties in U_{Irrig} , U_{Indus} , U_{Serv} , $SW_p + GW_R$, SW_{IN} and SW_{OUT} are included in the FOUA of WS indicator, hence, this application yields:

$$\overline{WS}(\%) \stackrel{1}{=} \frac{\overline{U_{Irrig}} + \overline{U_{Indus}} + \overline{U_{Serv}}}{(\overline{SW_p} + \overline{GW_R} - (\overline{Q_{overlap}})) + (\overline{SW_{IN}} - \overline{SW_{OUT}}) - \overline{EFR}} \times 100 \quad (2.9)$$

where the overbars designate mean values and $\stackrel{1}{=}$ signify equal in the first order sense. The uncertainty contributed by the i_{th} parameter is given

by:

$$C_i \stackrel{1}{=} \left| \frac{\partial WS}{\partial P_i} \right| S_{P_i} \quad (2.10)$$

where S_{P_i} represents the standard deviation of the parameter P_i . For this study, we used available monthly and yearly data of U_{Irrig} , U_{Indus} , U_{Serv} , $SW_R + GW_R$, SW_{IN} and SW_{OUT} from our reference database to compute the mean and variance. S_{P_i} was determined for each parameter at the yearly scale. Although S_{P_i} does not include all the uncertainty sources that might interfere directly in the measurement of the parameters such as errors related to gauging equipment and the observations, it represents the interannual variability of each dataset, which contains a part of uncertainty sources. The total uncertainty in WS indicator is:

$$S_{WS} \stackrel{1}{=} \left(\sum_{i=1}^n C_i^2 \right)^{1/2} \quad (2.11)$$

We adopted both of the governmental irrigation data ($U_{Irrig, Gov}$) and the estimated irrigation consumption ($U_{Irrig, RS}$) to yield the total uncertainty in the WS indicator for both approaches ($S_{WS(Gov)}$ and $S_{WS(RS)}$). Therefore, comparison can be made to assess the added value of the estimated irrigation data in the calculation of the WS indicator (Section 2.3.4).

2.3 Results and discussion

2.3.1 WS estimates at basin and administrative region scale using governmental data

Figure 2.3 shows the yearly WS at the scale of the Medjerda catchment for the period 2000-2016. It follows a significant increasing trend (Mann-Kendall trend test, $P_{value} < 0.05$). Lowest values of WS were in the early 2000s with only 30% of WS in 2000 and 36%, 41% and 42% of WS in 2001, 2002 and 2003 respectively. This low level of WS is a result of the low demand for irrigation (only 117.44 Mm^3 in 2000), the presence of important hydrological events (especially between 2002 and 2003), and the low population in the Medjerda in that period (around 950,000 thousand in 2004 according to the (INS, 2014)), which implies lower water demand for drinking water provision. For the period from 2004 to 2012, the WS increased and ranged between 50% and 64%. This increase is consistent with

a population increase during that period, as well as the increase of demand for irrigation and the rapid expansion of irrigated lands.

Since the year 2011, the WS increased further to reach 72% in 2013 and to exceed 100% in 2015 and 2016 (102% and 108% respectively). This fast increase of WS is a result of a high increase of irrigation demand (319 Mm^3 and 291 Mm^3 in 2016 and 2015 respectively versus 117 Mm^3 in 2000) in the basin and from neighboring regions such as Cap Bon connected to the Medjerda hydrosystem since 1982 through the Medjerda-Cap Bon channel.

In this study, the volume of water withdrawn by the Cap Bon region was not included due to the non-availability of public data. In addition, an increase of the domestic demand influences the level of WS as the population grew significantly in the basin (around 2 million in 2014). The industrial withdrawal in the basin wasn't significantly affected during the 2000-2016 period and remained stable between 4.5 Mm^3 and 6 Mm^3 per year. The increasing demand, mainly from the irrigation and domestic sectors, results in excessive exploitation of the available water resources. About 90% of the groundwater resource of the Medjerda aquifers were overexploited in 2015 and 2016 (DGRE, 2016). FAO (2017) and Vanham et al. (2018), consider "serious water scarcity" when the WS exceeds 70%. Hence, the Medjerda catchment suffers from serious water scarcity since 2013.

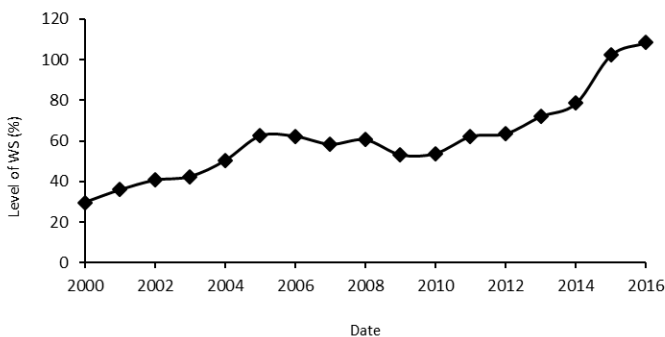


Figure 2.3: Yearly WS at the scale of the Medjerda catchment.

Different WS patterns are obtained when disaggregating the indicator at the level of four regions within the basin (Fig. 2.4). The yearly spatial distribution of WS followed a significant increasing trend ($P_{value} < 0.05$) in Jendouba, Beja and El Kef (Fig. 2.4a, 2.4b, 2.4c).

However, in Siliana (Fig. 2.4d) the time course of the WS was very irregular

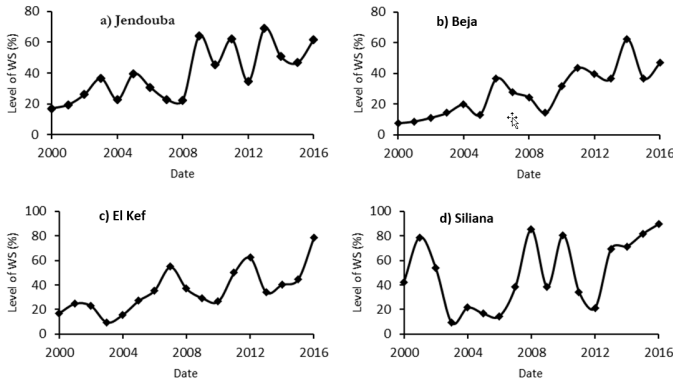


Figure 2.4: Yearly WS at the scale of the administrative regions.

marked with peaks, illustrating the dependence of the WS on the hydrologic regime of the Siliana river which flows into the Medjerda at Laouedj Station in Beja. The WS reached 70% (severe water scarcity) only in El Kef in 2016 and in Siliana in the majority of the period 2000-2016 with an increasing trend since 2013 (Fig. 2.4d). Jendouba and Beja regions didn't reach the severe water scarcity situation over the entire period of 2000-2016.

The critical situation in the Siliana region is related to the overexploitation of aquifers (100% of which are over exploited) as well as the decrease in the capacity of dams (the Lakhmes, Rmil and Siliana Dams). The dams in the Siliana region are exclusively dedicated to irrigation. They collect particle laden surface waters, drained from basins that are subjected to poor erosion management. Hence, these dams are prone to sedimentation and reduction of freshwater storage capacity.

Many factors may contribute to these different spatial patterns of WS in the Medjerda subregions. First, the water resources of each region vary greatly, in terms of the number of dams and the average annual discharge of rivers. For instance, Jendouba region contains 3 main dams (Bouhertma, Kesseb, and Bni Mtir) dedicated to domestic water supply and irrigation. The average annual discharge of the Medjerda river in Jendouba is $10 \text{ m}^3/\text{s}$ in the period 2000-2016. In contrast, El Kef region contains only one dam (Mellegue dam) and depends on Mellegue river and the Sarrath stream, with an average annual discharge of $6 \text{ m}^3/\text{s}$ and $0.8 \text{ m}^3/\text{s}$ respectively for the period 2000-2016. Second, the irrigation demand varies between the

regions depending on the available irrigated lands. The highest irrigation demand is in Beja region with 150 Mm^3 in 2016 compared to Jendouba, El Kef, and Siliana with 15 Mm^3 , 42 Mm^3 , and 43 Mm^3 respectively. Another important factor of the high variation of WS between regions is the different domestic water demand, which depends on the population of each region. The demand in Jendouba region is the highest of the 4 regions with 12 Mm^3 in 2016 compared to Beja, El Kef and Siliana with 11 Mm^3 , 8 Mm^3 and 7 Mm^3 respectively.

2.3.2 Governmental data vs. RS data-based irrigation water consumption and WS estimates

Figure 2.5 compares the yearly estimated irrigation water consumption to the governmental data provided at the scale of the entire Medjerda basin. We mark the presence of significant differences between the values coming from both data streams. The differences decrease with time. This may lead to the assumption that the remote sensing method tends to overestimate irrigation water consumption.

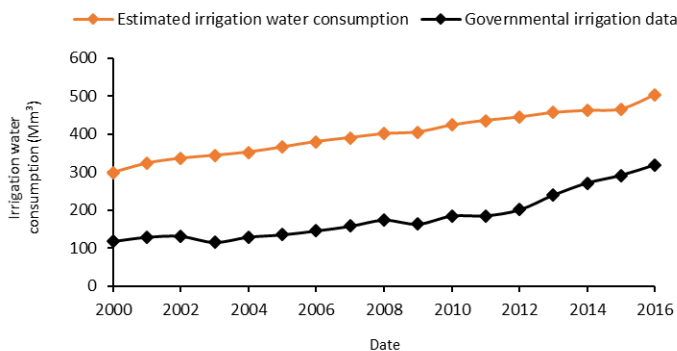


Figure 2.5: Comparison of yearly irrigation water consumption.

The differences in estimated water consumption data with both methods can be due to many reasons. First, governmental data don't provide highly accurate measurements of irrigation withdrawal in the basin. According to the GDWR hydrological bulletins, the provided data don't account the illegal water withdrawn from surface and groundwater. Such illegal withdrawals have been regularly reported in the basin, in particular since the early 2000s. Second, our estimation of irrigation water with-

drawal take into account the irrigation efficiency which ranges from 60% to 45% for the period 2000-2016 (Dhahbi, 2018) which adds 40% to 55% (water lost through the irrigation systems) to the volume of evapotranspiration from the irrigated lands in the basin (section 2.3.3).

Estimated irrigation withdrawals were integrated into the assessment of the WS indicator at both yearly and monthly temporal resolutions for the entire Medjerda catchment and Jendouba, Beja, El kef and Siliana regions. Figure 2.6, compares the estimated yearly WS values based on remote sensing and governmental data. At the scale of the whole catchment, both approaches lead to an increasing WS trend ($P_{value} < 0.05$), with higher WS values when remote sensing data are used. The WS level reached 71.8% (sever water scarcity) in 2004 and exceeded 100% in 2011 (102.31%).

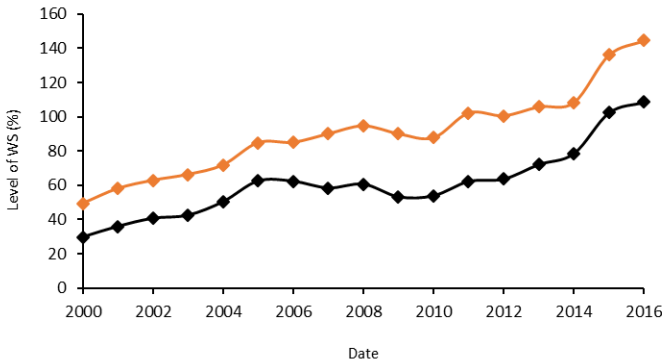


Figure 2.6: Level of WS at the scale of the Medjerda (black color refers to WS using only governmental data and the orange refers to WS that includes estimated irrigation).

At the regions' scale (Fig 2.7), similar observations can be made. Both methods follow the same pattern and trend ($P_{value} < 0.05$) for the entire period with higher WS values using remote sensing data.

This increasing trend of yearly WS is mainly due to the increasing demand for irrigation, domestic, and industrial water consumption in the period 2000-2016, which is directly related to the expansion of irrigated lands and the important population growth at the scale of the catchment and its sub-regions.

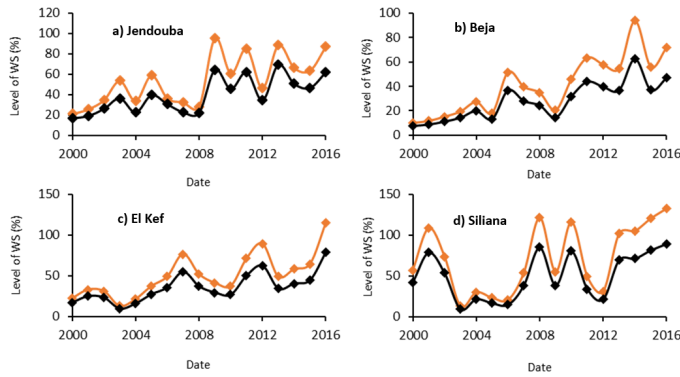


Figure 2.7: Comparison of Level of WS at the scale of Medjerda subregions.

2.3.3 Disaggregating yearly WS estimates

The monthly disaggregation of the WS indicator at the scale of the entire basin is shown in Figure 2.8.

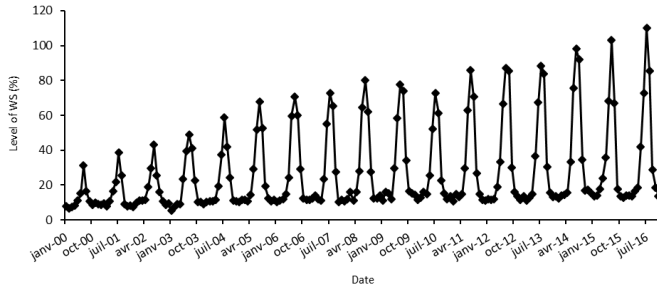


Figure 2.8: Monthly distribution of WS at the scale of Medjerda catchment.

There is a significant trend ($P_{value} < 0.05$) of increasing WS with a clear seasonal pattern for the whole period. We observed higher levels at the end of the spring season (May) and the entire summer season (Jun, July, and August). Severe water stress ($WS > 70\%$) is reached mainly in summer (dry season) due to the increase of irrigation withdrawals, domestic, and industrial water consumption and the decrease of precipitation.

Monthly-disaggregated WS data for the smaller Medjerda regions (Fig 2.9) showed a slightly less seasonal behavior with more significant irregularities in Siliana region (Fig 2.9d), which confirms the dependence of the WS

on the hydrologic regime of the Siliana river. Yet, the higher values of WS at the end of the spring season and the entire summer season are consistent with the assessment at the scale of the entire Medjerda basin.

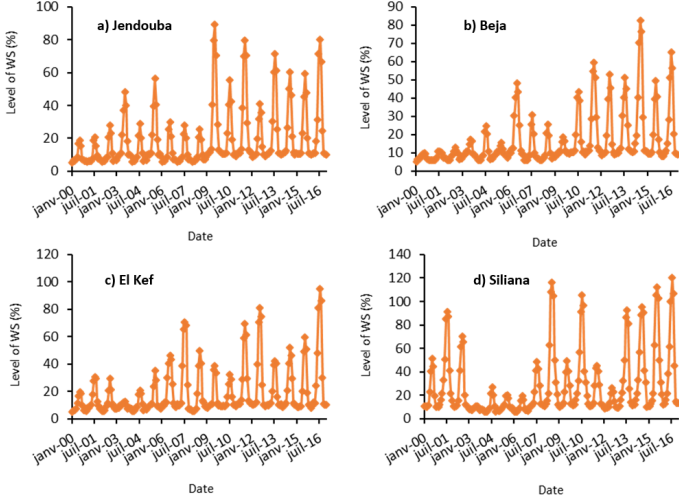


Figure 2.9: Monthly distribution of WS at the scale of Medjerda subregions.

The seasonal behavior of WS at the scale of the Medjerda catchment and the regions within is mainly due to the seasonality in the irrigation water consumption.

In addition, the domestic and industrial water withdrawals are marked by a similar pattern with a rise at the end of spring and during the summer season, which contributed to the specific dynamic behavior of WS.

2.3.4 Uncertainty analysis

The contribution of each input parameter to the uncertainty in yearly WS at the scale of the Medjerda is summarized in Table 2.2. The greatest uncertainty in WS assessment is contributed by the mean annual flow parameters (SW_{IN} , SW_{OUT} , and SW_P), which is due to the high seasonal variability of discharge measurements in the study area. GW_R , U_{Indus} , and U_{Serv} contributed less to the uncertainty in the WS assessment. Indeed, these parameters exhibit low seasonal variability. Irrigation water withdrawal provided by governmental data ($U_{Irrig,GOV}$) contributed also to less uncertainty. However, since 2012, the uncertainty in this parame-

ter increased significantly. This is mainly due to the higher seasonal variability of $U_{Irrig,GOV}$ in that period marked by high irrigation demands in Jendouba and Siliana regions. The estimated irrigation water consumption ($U_{Irrig,RS}$) was included in the UA as well, Table 2.2 and Figure 2.10, show that the remote sensing method contributed slightly more uncertainty than the governmental data until the year 2012, and significantly less uncertainty in the last four years.

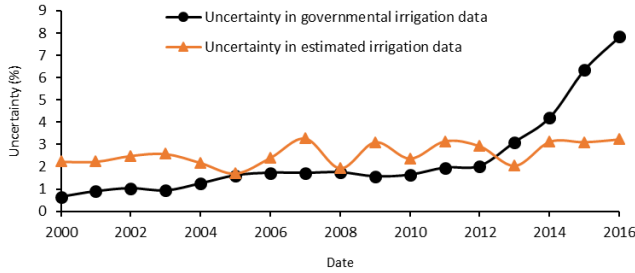


Figure 2.10: Comparison of irrigation uncertainty contribution (gov data vs estimated data).

The total calculated WS uncertainty (S_{WS}), both for assessments based on remote sensing ($S_{WS(RS)}$) and governmental data ($S_{WS(GOV)}$) is shown in Table 2.2 and Figure 2.11. The WS uncertainty for both approaches follows the same pattern and trend with slightly higher values of uncertainty when governmental irrigation withdrawal data are used. Therefore, the inclusion of remote sensing based irrigation water consumption data does not only allow to downscale WS assessments to the regional and monthly scale but also reduces the total uncertainty in the WS assessments.

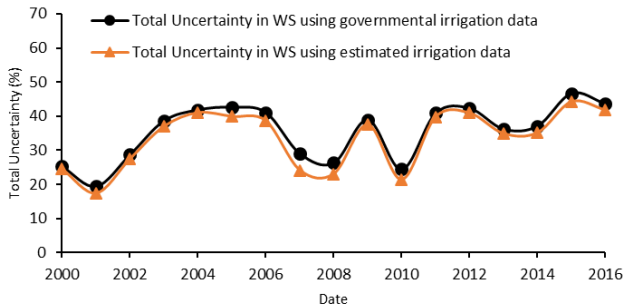


Figure 2.11: Comparison of total uncertainty ($S_{WS(RS)}$ vs $S_{WS(GOV)}$).

Table 2.2: Parameter contribution to the uncertainty in WS assessment at the scale of the Medjerda catchment

P_i	$C_1(SW_{IN})$	$C_2(SW_{OUT})$	$C_3(GW_R)$	$C_4(U_{Irrig, Gov})$	$C_5(U_{Irrig, RS})$	$C_6(U_{Indus})$	$C_7(U_{serv})$	$C_8(SW_P)$	$S_{WS}(GOV)$	$S_{WS}(RS)$
2016	30.31%	20.18%	3.14%	7.83%	3.23%	0.18%	0.48%	22.51%	43.64%	41.76%
2015	28.23%	20.02%	3.00%	6.35%	3.10%	0.17%	0.46%	30.57%	46.71%	44.22%
2014	21.89%	21.72%	1.79%	4.20%	3.12%	0.13%	0.42%	20.19%	37.14%	35.12%
2013	28.46%	14.44%	1.63%	3.11%	2.05%	0.14%	0.48%	16.81%	36.24%	34.85%
2012	29.92%	19.02%	1.43%	2.03%	2.92%	0.12%	0.51%	23.00%	42.34%	40.90%
2011	24.58%	25.84%	1.56%	1.96%	3.15%	0.13%	0.44%	19.91%	40.92%	39.80%
2010	18.32%	3.79%	1.22%	1.64%	2.37%	0.11%	0.35%	15.60%	24.45%	21.34%
2009	23.37%	23.61%	1.27%	1.56%	3.10%	0.11%	0.40%	20.29%	38.98%	37.51%
2008	19.90%	2.03%	1.64%	1.75%	1.94%	0.13%	0.44%	16.97%	26.34%	23.05%
2007	19.90%	2.03%	1.64%	1.75%	1.94%	0.13%	0.44%	16.97%	26.34%	23.05%
2006	21.33%	23.26%	1.91%	1.72%	2.40%	0.14%	0.44%	26.22%	41.12%	38.59%
2005	23.23%	21.54%	1.98%	1.61%	1.70%	0.14%	0.54%	28.44%	42.66%	39.89%
2004	26.60%	26.21%	1.35%	1.25%	2.17%	0.11%	0.41%	18.86%	41.88%	40.99%
2003	25.36%	20.25%	1.01%	0.93%	2.57%	0.10%	0.37%	20.97%	38.67%	36.97%
2002	18.69%	17.48%	0.86%	1.03%	2.49%	0.11%	0.25%	13.02%	28.75%	27.43%
2001	14.96%	5.50%	0.67%	0.89%	2.23%	0.09%	0.23%	11.14%	19.48%	17.35%
2000	23.11%	3.40%	0.48%	0.64%	2.23%	0.08%	0.19%	9.68%	25.30%	24.41%

2.4 Conclusion

We disaggregated in this chapter the SDG-6 WS indicator for the Tunisian case study of the Medjerda catchment at different spatial and temporal scales. The disaggregation was conducted by the means of two approaches to assess irrigation water withdrawals. The major conclusions are as follows:

1. Yearly WS in the Medjerda basin showed an increasing trend for the period 2000-2016. The level of WS remained acceptable for the first four years with 42% in 2003. However, from 2004 to 2012, WS ranged between 50% and 64%. In the last four years of the period, WS reached 70% (72% in 2013) and exceeded 100% in 2015 and 2016 (102% and 108% respectively), which considers the Medjerda catchment to be in a situation of severe water scarcity.
2. The pattern of yearly WS at a finer spatial scale of the regions Jendouba, Beja, EL Kef, and Siliana followed, in general, the same increasing trend as the entire Medjerda. The level of WS reached 70% only in El Kef (79% in 2016) and Siliana. The Silianan subregion showed a different response in WS than the other subregions.
3. Irrigation water consumption was estimated using MODIS-ET and CCI Land Cover (CCI-LC) products at the scale of the entire catchment and the subregions. The irrigation water consumption assessment method based on remote sensing provided slightly higher yearly values of irrigation water than the governmental data.
4. Yearly WS assessments based on governmental and remote sensing approaches were compared. As well at the catchment and at the sub-region scale, both assessments showed the same increasing pattern for the entire period. The monthly assessment of WS at the scale of the entire basin exhibited a seasonal pattern and an increasing trend over time. At the spatial scale of the subregions, the seasonal pattern of WS was less pronounced. The highest values of WS were at the end of the fall and during the summer.
5. FOUA was utilized to characterize the uncertainty contribution of each parameter in the WS indicator. Results showed that river flow data contributed the most to the total uncertainty (S_{WS}) due to its high seasonal variability in the period 2000-2016. The uncertainty on WS assessed using governmental $S_{WS(GOV)}$ and RS data ($S_{WS(RS)}$)

were compared. The comparison showed that the uncertainty of WS using RS data was smaller as compared to the uncertainty of WS using governmental data. Therefore, the inclusion of the estimated irrigation water consumption based on RS data reduced the total uncertainty in the indicator.

3

Creation of the Together4Water platform

This chapter is based on the outcomes of a defended master thesis in *Computer Science*

Deineko, Nicolas. Creation of a website to dynamically visualize and collect environmental data. Ecole polytechnique de Louvain, Université catholique de Louvain, 2020. Prom. : Mens, Kim ; Fehri, Raed.

The Together4Water website is fully hosted by the UCLouvain servers.

Abstract

The efficiency of the current water data infrastructure in Tunisia for assessing water related SDG indicators in a robust way is currently low. There is therefore a need to develop alternative water data generation and processing methodologies to foster the Tunisian water data infrastructure.

To meet this challenge, the Together4Water project was launched in 2018. The Together4Water is a scientific initiative, which aims at engaging citizens in Tunisia in water monitoring and water knowledge generation. The Together4Water project envisages, a.o., to reinforce the existing water-related databases that suffer from data gaps and low quality, to engage citizens in sustainable water resources management, to foster capacity building in water data collection and processing, and to encourage the use of cost-effective novel technologies in water monitoring. In addition, the project aims at supporting the implementation and monitoring of SDG 6 indicators at different scales in Tunisia.

One of the specific aims of the Together4Water project is to create a website to dynamically visualize and collect water-related data at the scale of the Medjerda catchment in Tunisia. On this website, generic data collected from public data providers and specific data collected by volunteers in the framework of Citizen Science (CS) should be compiled and published.

In this chapter, the website created for the initiative is presented along with the implementation choices and technologies used. The running website is available at <https://2gether4water.sipr.ucl.ac.be/>

3.1 Introduction

WATER is a key natural resource for sustaining the development of Tunisia, but unfortunately it is a resource under heavy pressure. Sustainable management of water resources should be based on a thorough understanding of the different stocks and fluxes in water system and should therefore be based on efficient monitoring of the water status. Yet the traditional methods for monitoring water resources in Tunisia suffer from many drawbacks such as reviewed in the introduction of this thesis. As an alternative to traditional monitoring of water resources by the local authorities, generic monitoring strategies based on remote sensing and participatory observation strategies based on the engagements of citizens can be proposed. Yet, for managing the data collected by local authorities in combination with data provided by generic data providers and citizens, specific data infrastructure systems need to be developed. This requires specific platforms for compiling, processing and publishing the data. In this chapter, we show how such water data infrastructure platform is developed for a test region in the Medjerda catchment area.

3.2 Approach

In the framework of supporting the data availability in Tunisia, particularly at the scale of the Medjerda catchment, a citizen science data collection campaign was organized in collaboration with multiple agencies and institutions in charge of water monitoring and management such as the General Direction of Water Resources (GDWR), National Sanitation Agency (NSA), Water Exploitation and Distribution National Company (WEDNC), National Institution of Meteorology (NIM), and the Regional Commissary for Agricultural Development (RCAD). These organisations are responsible for maintaining the official water data infrastructure of the region. The target parameters are primarily rainfall, discharge, and water quality (multiple indicators for both surface water and groundwater) data.

Additional historical data related to drinking water provision, sanitation, as well as generic hydroenvironmental data collected from public data providers were also compiled. The obtained data were pre-processed and converted in conventional formats such as Microsoft Excel sheets, Shapefiles, and Tagged Image File Format (Tiff). This part was crucial to ensure

acceptable initial quality of the data and to facilitate the creation of the database.

A specific platform was subsequently created to manage the collected data. The platform enables users/citizens to access dynamic visualisations of the data and to send measurements in the framework of CS using a built-in instant chat box. The chat feature of the website is essential to communicate with citizen scientists and to ensure easy data transmission throughout the CS field campaigns. The web application was built using modern tools and technologies. The code should be maintainable and documented. The software development methodology started with a rapid throwaway horizontal prototype illustrating the main functionalities of the application. Once the prototype was confirmed, the implementation of the different parts of the software was initiated. The development method of this website was close to the waterfall model, although, it's better described as a waterfall with overlapping sub-projects, since the project was made of multiple parts that can sometimes be developed in parallel, but also sometimes require other parts of the project. The code of the website was tracked using the git version-control system in order to maintain an history and traces of changes. The git repository was also hosted on a private GitHub repository available to whitelisted users at <https://github.com/ndeineko/EPL20-master-thesis> so that the different administrators of the project could have access to the source code.

3.3 The available water-related database

The collected historic data from the various Tunisian water institutions are the foundation of the Together4Water database. Given the fact that each provider has a specific format for data, a data format standardization is needed in order to facilitate the process of the creation of the Together4Water platform. All data are initially disaggregated at the level of 7 governorates: Tunis, Manouba, Bizerte, Beja, Jendouba, El Kef. In the following, a description of the collected historic data (summarized in Table 3.1).

Drinking water

Drinking water data were collected from the Water Exploitation and Distribution National Company (WEDNC) for the period 2000-2016, covering major water supply parameters such as the consumed volume by usage,

monthly distributed volume, number of created connections, number of subscribers per use, and the population served. The data are in good quality and without gaps for all 7 regions.

Sanitation

The National Sanitation Agency (NSA) provided good quality sanitation data for the period 2000-2016 that encompass 5 main parameters: the length of sewer network, number of registered users, level of connections to the network per region, billed consumed volume, and number of treatment plants. In addition, additional information about the treatment plants was provided. For example, the geographical location of each plant, daily and annual discharge, performance, and the quality of the treated water.

Rainfall

Rainfall data were provided by the General Direction of Water Resources (GDWR) for the period 1970-2016. A total of 147 monitoring stations are available with high quality monthly and daily data. The stations are also provided with precise geographic location and a conventional ID number.

Discharge

Daily and monthly high quality discharge data for the period 1995-2016 of 9 monitoring stations in the Medjerda catchment were provided by the General Direction of Water Resources (GDWR). Additional metadata were also delivered such as the geographic location, a conventional ID number, and the monitoring tool.

Groundwater

Good quality groundwater data were provided by the General Direction of Water Resources (GDWR) for the Medjerda's aquifers for the period 1995-2016. The data are divided into two categories: shallow aquifers and deep aquifers. The available parameters are mainly the groundwater recharge, exploitation, and the piezometry level. In addition, metadata were collected such as the number of monitoring wells, geographic location of wells, and conventional ID number of each well.

Water quality

Water quality data were collected from multiple sources. The WEDNC, NSA, and GDWR provided good quality data for all Medjerda regions for the period 2000-2016. The collected parameters are mainly, nitrate, phosphors, pH, salinity, and drinking water microbiology analysis. Additional metadata were also delivered such as the number and location of samples and the sampling method.

Table 3.1: Summary of the Together4Water database.

Data	Parameters	Metadata	Source	Temporal resolution
Drinking water	Consumed volume by usage; Monthly distributed volume; Number of created connections; Number of subscribers per use; Population served	Location of reservoirs	WEDNC	2000-2016 (Yearly and monthly)
Sanitation	Length of sewer network; Number of registered users; Level of connections to the network; Billed consumed volume; Number of treatment plants	Location of the treatment plants	NSA	2000-2016 (Yearly and monthly)
Rainfall	Rainfall	Location of stations; Stations' ID number	GDWR	1970-2016 (Monthly and daily)
Discharge	Discharge	Location of stations; Stations' ID number	GDWR	1995-2016 (Monthly and daily)
Groundwater	Groundwater recharge; Groundwater exploitation; Piezometry level	Number of monitoring wells; Location of wells; Wells' ID number	GDWR	1995-2016 (Yearly)
Water quality	Nitrate; Phosphors; pH; Salinity; Microbiology analysis	Location of samples; Sampling method	WEDNC; NSA; GDWR	2000-2016 (Yearly)

3.4 Concept

The creation of the Together4water platform is based on the following major steps. First, the website characteristics and features were specified and discussed with the project administrators. Then, multiple existing similar applications were analyzed to acquire knowledge on what a coherent website structure or page layouts could be used. A prototype of the application was then proposed based on the initial characteristics and features of the future website as well as the collected example of similar platforms. Afterwards, the details of the used technologies were analyzed and the implementation choices were made. The graphical user interface of the public

part of the site as well as the administrative pages were presented.

3.5 Requirements of the web application

As stated in the previous sections, an objective of this work is to create a website to dynamically visualize water-related data. Another technical objective of this work is to include an easy communication tool that can facilitate the interaction with citizens and to collect new data sent by volunteers.

After analyzing the different requirements and potential features to be integrated in the web application, the design criteria were fixed. The platform should:

- contain a database with data imported from the existing files (historical data);
- yield a dynamic map of the location of each dataset;
- allow multiple dynamic visualizations of the time series;
- contain a brief descriptive statistics of each time series (Min, Max, Mean, and Median);
- allow to link the map with the different time series visualizations;
- be validated by the client and other users;
- give the possibility for the administrators to add, export, and delete data; and
- contain an interactive instant chat box to facilitate CS data collection either by sharing text messages, images, or files.

3.6 Existing Citizen Science platforms

Several active citizen science projects implemented interactive online platforms in order to improve their communication with volunteers and to facilitate data transfer and visualization. The benefit of having well structured platform in the context of CS activities is also related to the motivation of volunteers since it could be a way of keeping their motivation at highest levels through constant feedbacks, presentation of the collected

data, easy access to the different visualizations, and direct communication and support.

The Maasai Mara citizen observatory website (<http://mara.info.ke>) is part of the Ground Truth consortium (Ground Truth, 2020), which supports several citizens observatories around the world and guide them to have consistent implementation within the societies as providers of capacity building opportunities as well as reliable data. The website provides visualizations of a wide range of citizen-generated data using a smartphone application such as precipitation and biodiversity parameters, and also various data sent by automated weather stations.

The website presents dynamic maps of the data with different types of layers as well as multiple charts, mostly line and bar charts for hydro-climatic data, with a drop-down menu to select the stations. However, it doesn't provide a real-time communication tool as well as the ability to register within the platform to add comments and share data. Another existing CS project that provides an interactive online platform is the iFADO project (iFADO, 2020). It aims at improving the monitoring of water quality of various water bodies using advanced tools as well as citizens' participation. The iFADO Platform (<http://www.ifado.eu/>) provides dynamic visualizations of data with interactive visual effects, which increases the quality of the presented information. Also, the platform gives rich information about the progress of the project as well as the different tools used in the data collection. Nevertheless, the platform doesn't provide communication tools with citizens. EyeOnWater (<https://www.eyeonwater.org/content/map.php>) is an online platform directly linked to the EyeOnWater smartphone application, which is used to collect water quality data and produce interactive maps of the outcomes from users around the world. The map shows the location of the measurements allows visitors, by clicking on individual measurements on the map, to see the picture that was taken along with brief metadata such as the date and time, and a button to report incorrect pictures.

The common aspects between the presented platforms are the high level visualization of the data and the detailed information about the latest news and activities of citizen science campaigns and the progress of the project in terms of achieving the defined objectives. These details are essential for any CS platform, however, other features were missing such as a real-time communication space where citizens can register and communicate with each other and report data, pictures, and questions to the project coordinators. These particular additions could increase the citizens' mo-

tivation and willingness to contribute to the projects. The Together4Water project considers these aspects as basis for the implementation of the online platform.

3.7 Prototype

Before starting the development the web application, it was important to make sure that the objectives of this work are interpreted in the same way by the client and the implementer. To do so, a prototype of website was created, which provides a representation of the different characteristics and features of the website.

There are multiple available prototyping tools for web applications. A selection of 3 tools was proposed: InVision, Fluid UI, and Proto.io.

3.7.1 The prototype of Together4Water

Before starting the development of the Together4Water platform, a prototype was created in order to check in advance the different features as well as the potential design of the pages. This step is crucial as it allows testing a "theoretical" version of the website that has most of the fixed requirements, which gives the ability to modify them if necessary.

The home page is shown in Figure 3.1, the map page in Figure 3.2 and the data selection page in Figure 3.3.

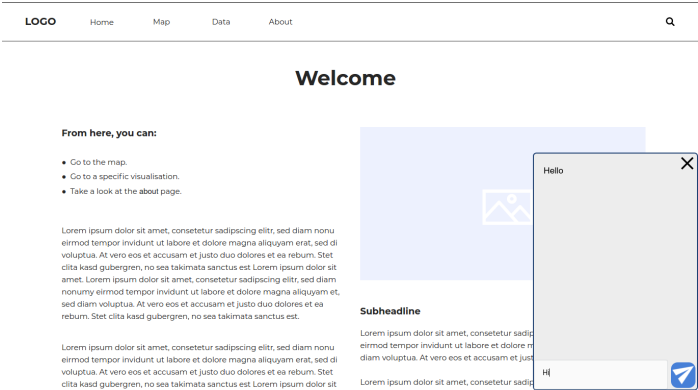


Figure 3.1: Home page in the prototype. The chat window is opened.

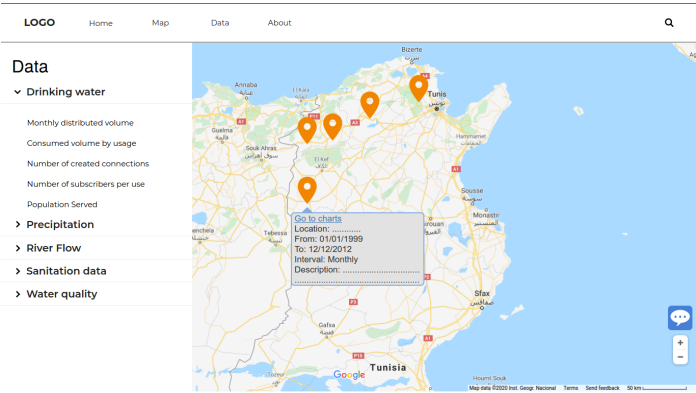


Figure 3.2: Map page in the prototype. A data type is selected and a location popup is visible.

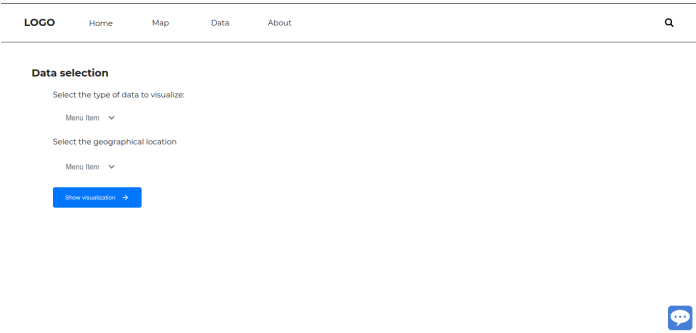


Figure 3.3: Data selection page in the prototype. By selecting a data type and a location and clicking “Show visualization”, the user will be redirected to the data visualization page where time series are displayed in line charts.

3.8 User interface

3.8.1 Public website pages

The public pages are designed based on a similar visual style with low complexity in order to reduce visual noise, which should facilitate the user's understanding of the different features of the Together4Water platform. The color scheme of the website is composed of white, gray and 3 shades of blue, with the text in very dark gray.

The design of the web pages is kept standard, each page is composed of a header, content, and a footer. The header stays at the top of the screen even while scrolling the page, which increases the simplicity of the website. It contains direct links to the main pages and a search bar.

The footer contains the copyrights as well as links to secondary pages such as the terms and conditions page.

Home

The home page represents a welcome page to the website. It contains a brief description of the Together4Water project and its different objectives as well as some details about the citizen science monitoring tools.

Map

The map page is divided in 2 parts; the available data types on the left hand side and the interactive map on the right with markers for all locations associated with the selected data type (Fig. 3.4). Data types are grouped into expandable categories, which makes the search and selection of data much easier.

When a data type is selected, markers appears at various locations on the map. Clicking on a marker shows a popup menu with information about the dataset (description, number of charts, etc.) and a link to the data visualization page where charts can be visualized.

Data selection and visualization

The data selection page allows the user to search and choose a data type then select the associated locations. Subsequently, the user can go to the associated data visualisation page where several charts and visualizations of the data can be found (Fig. 3.5).

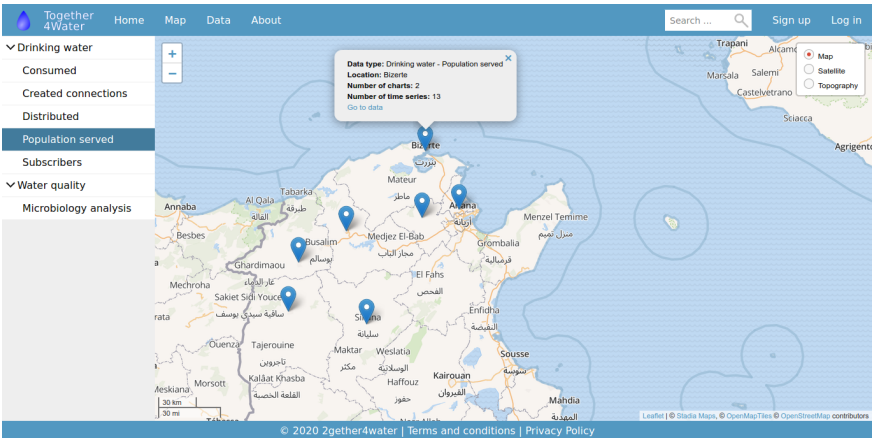


Figure 3.4: Screenshot of the website’s map page. A data type is selected, a location popup is shown and the background map selection menu is expanded.

Following the initial website requirements, a button that shows and hides descriptive statistics about the time series was implemented (the number of missing values, the minimum value, the maximum value, the mean, and the median).

With respect to data download, only users with approved login and a permission from the administrator have the right to download the dataset. A button to download the current dataset in JSON format is therefore shown at the bottom of the content. On the other hand, for the users who are not connected with a login or don’t have the right to download data, the button is replaced with a message to contact the administrator to apply for the permission to download.

Login and Sign up

The login page provides the user with a form to login with a user name and a password while the sign up page proposes a simple form to create a new profile by entering a user name, a password, and the affiliation of the user. All the fields in both pages are associated with a help text to guide the user throughout the process.

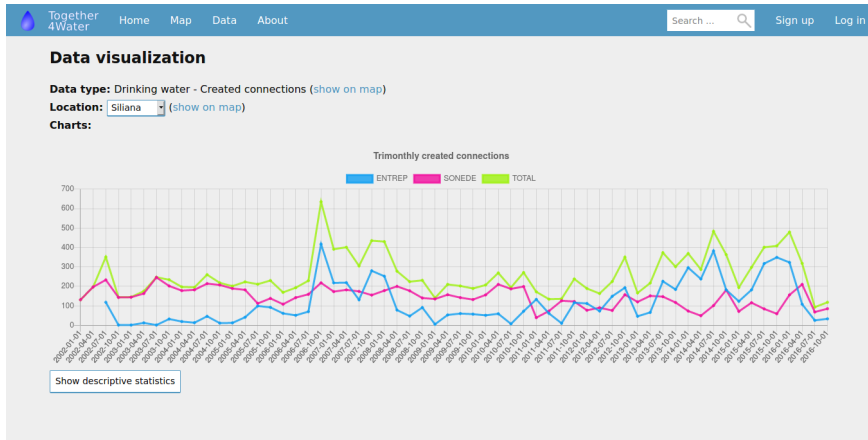


Figure 3.5: Screenshot of the data visualization page for a specific dataset.

About

The about page contains information about the funding partners of the Together4Water project and more importantly a full description of the data sharing policy. A direct link to the official protocol of data sharing is included in order to ensure that the users have a clear understanding of the process of data access.

Instant chat box

When the user is connected, the instant chat box appears at the bottom left of the page. The user can click on it to show the chat popup where text messages, images, and files can be sent.

3.8.2 Administration pages

The administrative pages are based on the default Django administration style. Within the site, Django models' objects can be listed, modified, and deleted (Fig. 3.6).

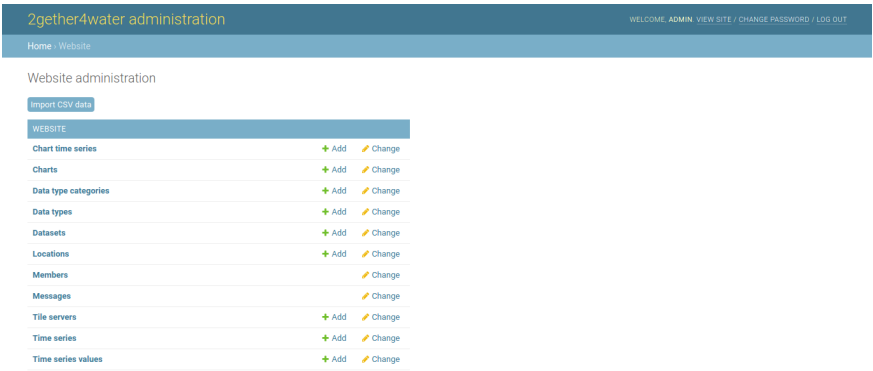


Figure 3.6: Screenshot of the `website` page in the administration site.

Changes have been made to the default administration settings for some models in order to facilitate the management of the platform:

In the objects' models page, a search bar was included and objects can be sorted according to several criteria. In the website models page, a button was added to import datasets in CSV format that should follow the specifications defined in Figure 3.7. Data import is simplified to guarantee easy transformation of raw citizens' data to standardized information.

More technical details about the development of the Together4Water platform can be found in Deineko (2020).

The CSV file should begin with zero or more metadata lines.

A metadata is defined as a line of comma-separated elements where:

- The first element is a string starting with META:x where x must be replaced by the type of metadata.

The different possible metadata types are defined as follows:

- Location: represents the location of the dataset.
- Data type: accepts 2 arguments: the category and the name of the dataset.
- Chart time series: accepts at least 2 arguments:
 - the name of a chart and the names of time series as they appear in the headers of data columns.
- Chart display names: accepts at least 2 arguments:
 - the name of a chart and the labels that will be associated with the time series.
- Time series units: accepts at least 2 arguments:
 - the name of a chart and the units associated with the time series.

The following metadata are mandatory: Location and Data type

Figure 3.7: Specification of the CSV format of files containing datasets of multiple time series.

3.9 Conclusion

In this chapter we described how a fully operational website for the Together4Water project based was designed and implemented based on the specific requirements fixed before the start of the development. The platform is implemented using advanced tools and can easily be extended and improved. The major conclusions of this chapter are:

1. The Together4Water website provides dynamic visualizations (Charts and Maps) of the historical data collected for the Medjerda catchment in the framework of SDG 6 monitoring. The collected raw data are implemented in a consistent database within the website, which ensures an easy and proactive management of the different datasets as

well as the ability to add new data or complete the existing ones.

2. The website include a direct link between the various visualizations. For instance, the user is able to visualize locations of discharge monitoring stations on the map and still can directly access the chart representation of a chosen station via a permanent link implemented in the description of each station.
3. Users are provided with the ability to create a personal profile in the Together4Water website, which give them a more complete experience with the different implemented features such as the possibility to fill a request form to download data, the use of the instant chat box, and to receive feedback and updates about the project. The creation of the user profile is made simple and requires very few personal information, which gives users more privacy while using the platform.
4. The instant chat box is implemented as a tool for data transmission and a space for interactions with and between citizens who are registered with an approved profile in the website. This tool enables participants of the different citizen science monitoring campaigns to send their measurements using text messages as well as image files. The administrator can access and collect the chat box inputs via the administration site.
5. A full description of the data sharing policy of the Together4Water is added in the "About" page, which complies with the policies of the different partners Tunisian institutions as well as the UCLouvain. This is essential in order to ensure transparent and regulated data management and sharing with all parties (users and partners).

4

Citizen Science for Water monitoring in Tunisia

This chapter is a modified version of a published manuscript in *Environmental Science & Policy*, part of special issue on *Water Innovation in Africa*

Fehri, R., Khlifi, S., and Vanclooster, M. (2020). Testing a citizen science water monitoring approach in tunisia. *Environmental Science and Policy*,104:67–72.

Abstract

CITIZEN (CS) has been emerging in the last decade as a new field of environmental monitoring involving a direct collaboration between everyday citizens and scientists. In Tunisia, several recent governmental efforts aimed at reinforcing and renovating the existing official water monitoring systems. However, the lack of reliable hydrological data is still an issue, which could be better addressed through integrating a CS approach. The latter approach is tested for rainfall monitoring in the Medjerda catchment in Tunisia using cost-effective and publicly available manual rain gauges. We used a step-by-step approach to target, engage, and train citizens on using the monitoring tools and transmitting the data to a user-friendly online platform. The ongoing approach involved 7 citizens from different generations and different educational backgrounds. The collected daily CS data are compared with data from reference stations. Results yield a significant correlation between CS data collected at 3 different sites and the reference stations with r (Pearson Correlation Coefficient (PCC)) ranges between 0.91 and 0.98 for all citizens. Student's t -test was applied to evaluate the significance of the agreement between the CS and reference data. In addition, the variability of the CS data is compared with the variability associated with the official governmental data. The CS approach delivered consistent outcomes to complement existing Tunisian monitoring systems, and also to enhance innovation, adaptation, and local capacity building in the Tunisian water sector.

4.1 Introduction

UNDER the current global water crisis and the future climate change scenarios, several water-related issues are expected to exacerbate many regions around the world and especially in southern Mediterranean, identified as a water-scarce region in recent years (FAO, 2018). Within this region, Tunisia is considered as one of the most water-stressed countries.

The Medjerda catchment, the most important river basin in the country, reached severe water scarcity with the SDG-6 water stress indicator exceeding 100% in recent years (Fehri et al., 2019). Yet, the robust evaluation of water stress remains complicated due to the lack of sufficient, meaningful, and consistent data. Despite the recent governmental efforts to reinforce the existing hydrometrological monitoring systems, and availability and accuracy of the data, several issues persist for assessing water scarcity.

Current official hydrometric and meteorological water monitoring networks (especially rainfall and discharge) are not spatially dense and suffer from data gaps, incomplete timeseries, and often leading to poor quality databases. Novel water-related data acquisition and handling strategies are therefore required to improve the characterization of catchments' behavior and to support national and local decision-making. This major issue can be partially addressed through CS.

The engagement of citizens in science have been evolving over the past decades as it provides opportunities for scientists to apply cost-effective approaches in order to generate useful information about the environment, which could be used in science as valuable additions to the existing models. Despite its relevance at the scientific level, CS is highly appreciated from a social aspect. The impact of citizen science on volunteers within projects is considerable and gaining more acknowledgment within interdisciplinary initiatives (Albert, Balázs, Butkevičienė, Mayer, & Perelló, 2021). Capacity building through training, social events, workshops, and collaboration with multiple volunteers play a major role in strengthening the added value of CS in society. In addition, being part of decision making and policy formulation in several CS indicatives around the world, citizens have more power and influence in their local communities regarding environmental issues such as pollution, water scarcity, and water availability (See et al., 2015). Nevertheless, it's worth mentioning that despite the major focus of CS on natural science with 80% of the current practices in Europe

(Tauginienè et al., 2020), social scientists have a crucial role in bridging participatory traditions from the natural/technical sciences, the social sciences, and the humanities. This role of social scientists help increasing the engagement of citizens in all forms of CS and most importantly strengthening the impact of the activities from a social aspect (Albert et al., 2021; Heiss & Matthes, 2017).

Although CS and the activity of community-based monitoring, known in the literature as crowdsourcing, citizen observatory, and participatory monitoring, is rapidly expanding around the world (Buytaert et al., 2014; See et al., 2016), considerable attention needs to be paid to the validation of CS approaches. Indeed, the quality of the CS-collected data may strongly depend on the CS data monitoring strategy, the variables that are observed and the socio-economic profile of the region in which the CS program is implemented. Hence, scope exists to evaluate the validation status of CS-based hydrometrological monitoring programs. Since there are currently no reported CS initiatives in Tunisia to monitor water-related data, we launched in 2018 the Together4Water project in collaboration with local Tunisian institutes and schools in a test region of the Medjerda catchment “Medjez-El-Beb city” (Fig. 4.1).

The study area is mainly based on agricultural activities and lacks operational water resources management (Fehri et al., 2019). One aspect of the program aims at engaging citizens from different generations in monitoring precipitation using cost-effective tools. The precipitation is highly variable in space and time in the study area, however, existing official observing networks are too far apart to see much of this variability and suffer from inconsistent and low-quality data.

In this chapter, we present results from the ongoing validation of the Together4Water CS field campaign in the test region. We used a step-by-step approach to engage and train volunteered citizens in the region. We assess the quality of daily precipitation monitored in 3 different sites in the test region (Medjez-El-Beb city center, ESIM, and Slouguia) based on data of 3 reference stations. We evaluate the agreement and variability of measurements between citizens for each site and we compare this with the mean and variability of the reference stations.

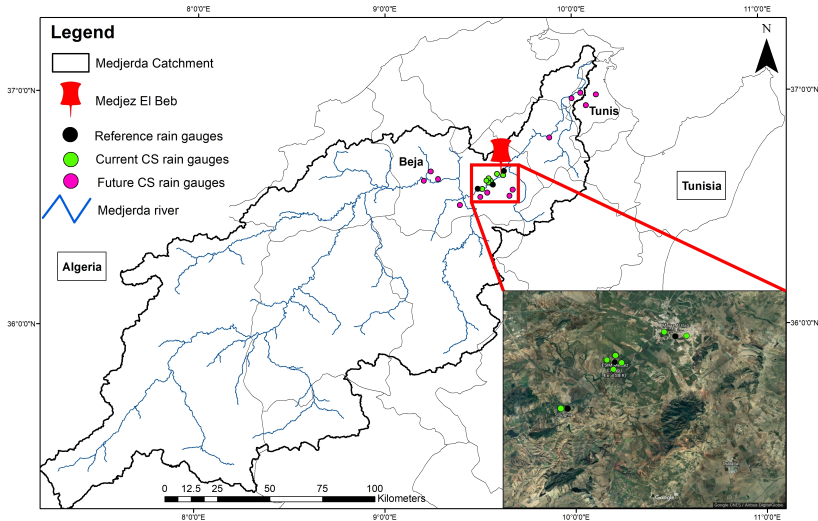


Figure 4.1: The study area. The right down inset is Medjez-El-Beb city, the CS test region.

4.2 Materials and methods

4.2.1 Step-by-step citizens science approach

In this research, we adopted a step-by-step citizens' engagement method (Fig. 4.2) in the test region at the start of the Together4Water field campaigns in October 2018.

First, the Together4Water online platform (<https://2gether4water.sipr.ucl.ac.be/>) was created to facilitate reaching a larger number of citizens via social media regardless of their educational backgrounds. The online platform provides all the needed information about the main objectives of the project, the different field activities, and the tools to be used. In addition, an instant message tool was setup (Section 3.1.5), which is designed for real-time data transmission. Then, 3 potential sites (Medjez-El-Beb city center, ESIM, and Slouguia) were selected in the region based on the spatial distribution of the volunteers and the availability of reference monitoring stations. Cost-effective traditional rain gauges (Fig. 4.3) were tested in

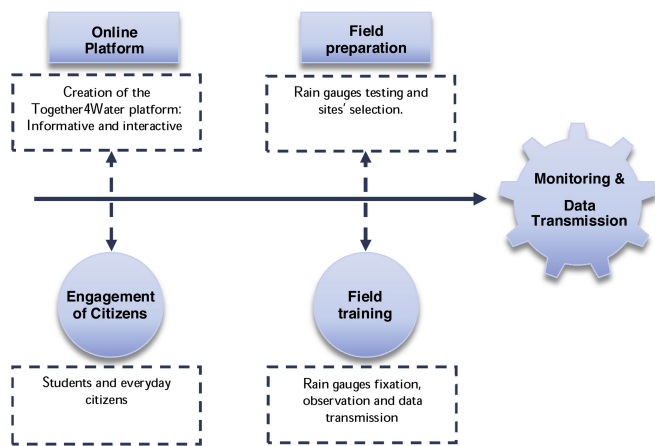


Figure 4.2: CS step-by-step approach.

the field and distributed to citizens to be fixed at their households around the chosen sites. Citizens were trained in advance on how to use the rain gauges so that they were confident to participate and collect good quality observations in all sites. This field preparation phase is crucial to any CS field campaign to avoid technical issues and to guarantee continuous monitoring. After successfully finalizing the previous steps, we launched the start of monitoring and data transmission.



Figure 4.3: Rain gauge fixed by a citizen scientist during monitoring rainfall.

4.2.2 Data quality assessment and validation

The quality assurance of data is crucial within a CS program. Hence, we focus in this chapter on the evaluation and validation process of CS rainfall data. The validation methods reported in the literature typically compare CS with reference data using statistical techniques based on the correlation coefficient, the bias, and comparative tests such as Student's t-test (Wolff et al., 2005). For this study, the CS and reference data were compared using the Pearson Correlation Coefficient (PCC) r , the percent error, and the Student's t-test (Table 4.1). The latter assesses the significance of the difference between the means of two datasets. The variability associated with the means of the two data sources is further compared. The theoretical basis and limitations of Student's t-test are detailed and described by Zabell (2008).

Table 4.1: Summary of CS data validation methods.

Method	Formula
PCC	$r = \frac{\sum_{i=1}^n C_S \cdot R - (\sum_{i=1}^n C_S) \times (\sum_{i=1}^n R)}{\sqrt{(N \sum_{i=1}^n C_S^2 - (\sum_{i=1}^n C_S)^2) \times (N \sum_{i=1}^n R^2 - (\sum_{i=1}^n R)^2)}} \quad (4.1)$
Percent error (PE)*	$PE_j(\%) = \frac{C_{S,j} - R_j}{R_j} \times 100 \quad (4.2)$
Student's t-test	$t = \frac{ \bar{X}_1 - \bar{X}_2 }{\sqrt{\frac{2(SD_1^2 + SD_2^2)}{2n-2}}} \quad (4.3)$

* C_S , citizen science data; R , reference data; N , number of data pairs; j , the j_{th} measurement/observation

** $\bar{X}_1$ and $\bar{X}_2$, means of the two data sets, SD_1 and SD_2 are their standard deviations

4.3 Results and discussion

4.3.1 Data variability

The collected data so far cover the period from October 2018 till July 2019. We removed the dry months (June and July) from the analysis since there was no precipitation measured during these months. The monthly variability of cumulative CS and reference rainfall for the 3 sites is shown in

Figure 4.4. Daily rainfall data were converted at a monthly scale for better visualization of the time series. The variability increases slightly with higher precipitation events. This variability could be due to the spatial distribution of citizens' rain gauges in the study sites and experimental bias.

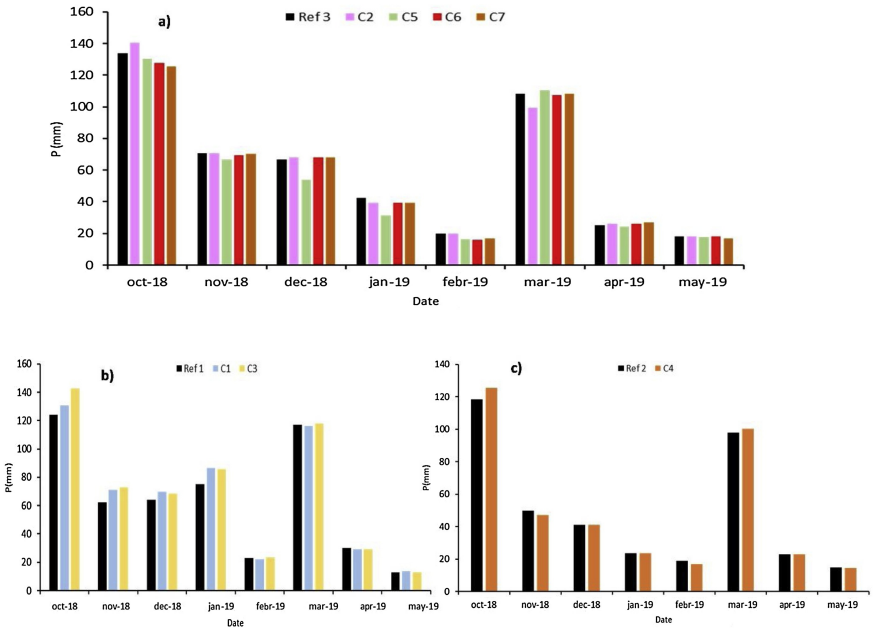


Figure 4.4: Variability of monthly cumulative CS and reference data (Ref) at ESIM site (a), Medjez-El-Beb city center (b) and Slouguia (c).

4.3.2 Correlation analysis

The PCC of each citizen's daily measurements and reference data yielded good agreements at all sites with r ranges between 0.91 and 0.98 (Fig. 4.5). Many reasons for this high correlation are proposed. First, the spatial variability of rainfall is not affecting the CS measurements as all rain gauges are mostly 900m apart from the reference stations. In addition, citizens were trained to take measurements at the same hour daily as the reference measurements (at 07:00 am) to guarantee the consistency in sampling time between CS data and the reference ones.

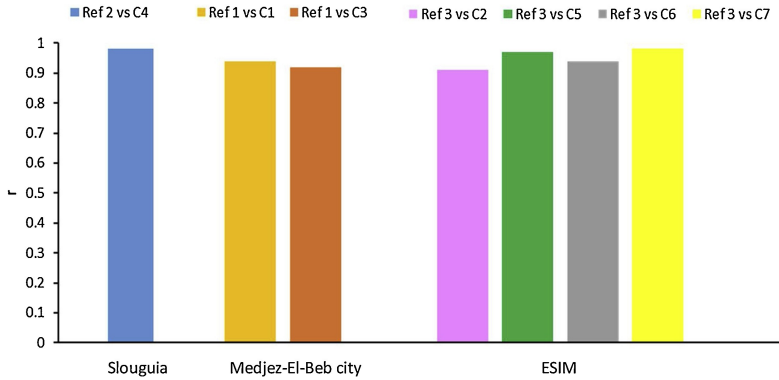


Figure 4.5: Correlation between CS measurements and reference data at all sites.

4.3.3 Percent error (PE) assessment

We performed a PE assessment at the study sites for all citizens' daily measurements. The average daily PE at Medjez-El-Beb city center site was 9% for citizen-1 and 13% for citizen-3. At ESIM site, the daily PE averaged 8%, 20.5%, 9%, and 11% for citizen-2, citizen-5, citizen-6, and citizen-7 respectively. At Slouguia, the average daily PE of citizen-4 measurement was 13.5%. In order to have more understanding of the error in citizens' measurements, the monthly PE was computed.

Figure 4.6 shows the variation of monthly PE at all sites. Citizens at Medjez-El-Beb city center site tend to slightly overestimate rainfall with relatively more consistent measurements delivered by citizen-1. At ESIM site, volunteers slightly underestimated rainfall. Citizen-2 measurements showed more consistency than the other volunteers. For Slouguia, citizen-4 delivered consistent measurement with a slight overestimation in February 2019. Hence, CS data showed acceptable error and, importantly, good consistency with the official data at all sites, suggesting that the measurement errors are not strongly contaminated by systematic bias. The variable measurement error between citizens at all sites could be due to short-range spatial variability, under catch as the citizens' rain gauge could be close to a small tree, wall or exposed to wind; and other observational measurements.

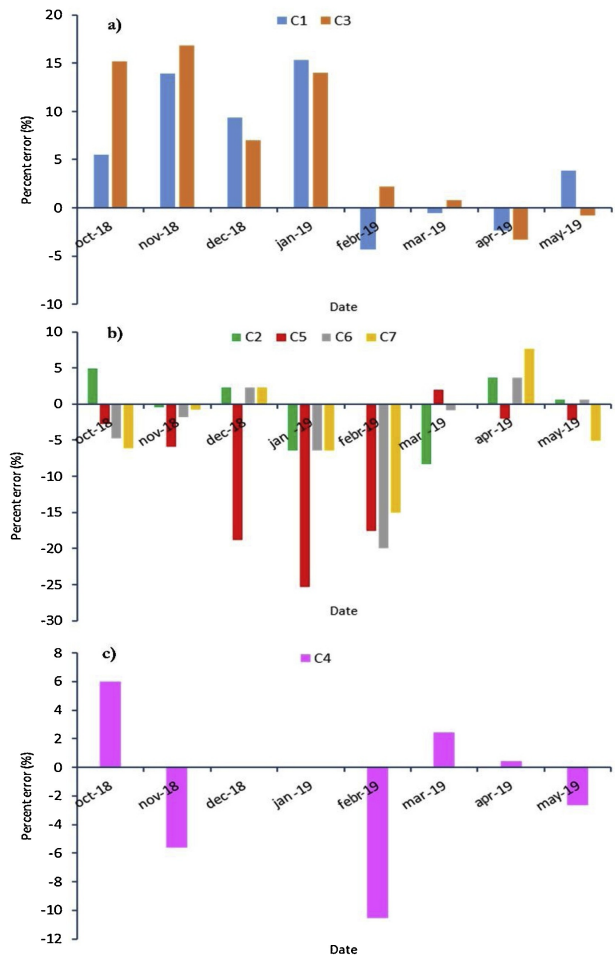


Figure 4.6: Variation of monthly rainfall PE at (a) Medjez-El-Beb city center, (b) ESIM, and (c) Slouguia.

4.3.4 Student's t-test and mean variability comparison

Student's t-tests were performed to assess the significance (at the 5% level, $\alpha = 0.05$) of the difference between the means of citizens' measurements and the reference data at all sites. The null hypothesis (H_0) considers the difference between the means equals zero while the alternative hypothesis (H_a) considers the difference is different from zero. Results of Student's t-test are shown in Table 4.2. At all sites, we can't reject the null hypothesis

since p-values are greater than the level of significance which suggests that the mean of CS measurements for all citizens is significantly close to the mean of the reference data. This confirms the consistency and good quality of the CS data.

Table 4.2: Student's t-test results.

Statistic	Medjez El Beb city			ESIM					Slougua	
	Ref_1	C1	C3	Ref_3	C2	C5	C6	C7	Ref_2	C4
min	0	0	0	0	0	0	0	0	0	0
max	87	95	100	90	92	92	92	91	85	90
SD*	8.659	9.097	9.402	8.705	8.675	8.559	8.743	8.726	7.565	7.834
$\bar{X}_i$ or $\bar{X}_{Ref,i}^*$	2.256	2.463	2.525	2.122	2.200	2.061	2.155	2.159	1.769	1.789
$\bar{X}_{Ref,i} - \bar{X}_i$	-	-0.208	-0.269	-	-0.078	0.061	-0.033	-0.037	-	-0.020
Student's t-test parameters										
t	-	-0.245	-0.311	-	-0.094	0.074	0.040	-0.044	-	-0.027
t critical	-	1.965	1.965	-	1.965	1.965	1.965	1.965	-	1.965
DOF*	-	436	436	-	436	436	436	436	-	436
P_{value}	-	0.807	0.756	-	0.925	0.941	0.968	0.965	-	0.978

*SD, standard deviation; $\bar{X}_i$, mean of CS data; $\bar{X}_{Ref,i}$, mean of reference data; DOF, degree of freedom

We assessed as well the variability (95% confidence interval (CI)) associated with the mean of each data source. Figure 4.7 compares the mean and variability of citizens' measurements to the mean and variability of the reference data. At all sites and for all citizens, similar uncertainty of the mean to the reference ones can be observed, which confirms the goodness of CS-collected data.

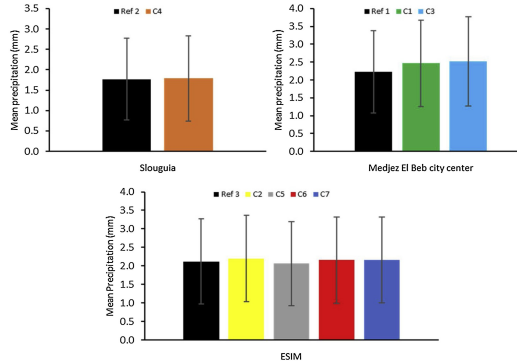


Figure 4.7: Comparison of the mean and the variability (95% CI) of the CS measurement to the reference data.

4.3.5 Importance of the step-by-step citizens' engagement

The step-by-step citizens' engagement approach delivered promising outcomes in terms of effectiveness and rapidity with respect to message spreading and communication with motivated and interested citizens. At the start of the field campaigns in October 2018, 7 citizens from different generations and educational backgrounds (students and everyday citizens) were selected to measure rainfall in the 3 sites: two citizens (C1 and C3) at Medjez-El-Beb city center, four (C2, C5, C6, and C7) at ESIM site, and one citizen (C4) at Slouguia. At the time of writing, 17 more citizens were engaged and equipped with rain gauges in newly selected sites (Fig. 4.1). The progressive approach of CS engagement helped volunteers better understand the importance of environmental monitoring in the region, which created a motivation to take actions and participate in the Together4Water field campaigns. Younger citizens (students) were more acquainted to utilize the online platform for data transmission via the instant messaging tool. Older citizens did need more training, which helped them better learn how to deal with data in terms of management and description. Citizens send daily rainfall measurements via the platform with a description of the circumstances of the measurements. Such an approach allowed to collect actual data as well as metadata of the target variable (Fig. 4.8).

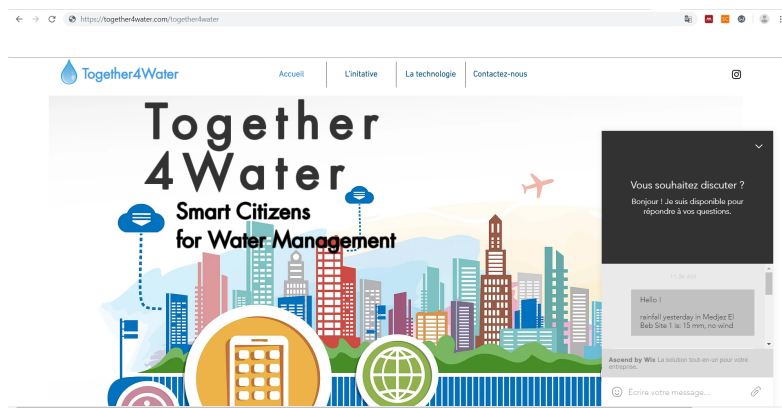


Figure 4.8: Rainfall CS data transmission example with a short description of the event.

This is highly important since metadata such as observation time, environment and weather at the time of measurement, or a photo of the rain gauge at time of measurement (Fig. 4.9), allows improving the quality of the data and understanding precipitation variability.



Figure 4.9: Photos of rain gauges sent by citizens during monitoring rainfall in the test region.

4.4 Challenges

With the Together4Water initiative, new challenges emerge at regional and national scales. The main ones are described as follows.

The Together4Water program involved everyday citizens from different generations in a small region. Moving to a larger scale is a challenge at different levels such as training process and data management. It is important then to improve our training and communications programs.

One of the main goals of the initiative is involving stakeholders in the monitoring program. It is a challenge to establish an operational link between citizen scientists and stakeholders to move forward towards an integrated water monitoring system in the Medjerda catchment. A more advanced step-by-step engagement approach is therefore needed.

4.5 Conclusion

We presented in this chapter a test of the Together4Water CS initiative in Tunisia. The step-by-step citizens engagement approach showed promising results in motivating citizens in the test region. The Together4Water platform helped to reach a range of locals through social media and regional websites. In addition, the variability of CS data between citizens, at all sites, was slightly higher for high rainfall events. This could be due to short-range spatial variability of rainfall, the presence of other disturbing factors like the wind, which affects the quality of measurement, and measurement error. In contrast, lower events measurements showed more consistency for all citizens. Statistical comparisons (correlation analysis, PE, Student's t-test, and means' variability) of CS measurements against the reference data at all sites validate their quality for use in further studies.

In addition to the validated quality of data, we consider the following lessons learned from the field. The field training is a crucial part of the field campaign. We provided consistent guidelines to citizens to better understand the technical aspects of rainfall measurement, data transmission and the added value of metadata. Younger citizen scientists (students) are more familiar with using the rain gauge and the Together4Water platform for data transmission than older citizens. Therefore, the latter category of citizens requires further assistance throughout the start of monitoring. Furthermore, we learned that the local community had a greater understanding of their local hydrology, a sense of responsibility of their water resources, and a sense of being a research partner in a community project.

5

Data Fusion of Citizen Science data

This chapter is a modified version of a published manuscript in *Journal of Hydrology*

Fehri, R., Bogaert, P., Khelifi, S., and Vanclooster, M. (2020). Data fusion of citizen-generated smartphone discharge measurements in Tunisia.

Abstract

WATER resources management techniques have been evolving over the years, introducing new ways of monitoring and collecting data that improve both the quality and quantity of water-related information. Among them, Citizen Science (CS) has been introduced in the field of environmental monitoring as a novel approach that involves a direct collaboration between citizens, scientists and local authorities.

In Tunisia, the lack of reliable hydrological data about rivers' discharge remains a major issue, despite the recent governmental efforts to reinforce the existing official monitoring systems. In this study, we show how this problem can be efficiently addressed using a CS methodology. Citizens at two locations in Tunisia (Slouguia and Medjez-N5) have monitored discharge of the Medjerda river (the most important water resource in the country) using a mobile phone application for a series of hydrological events. Using a Best Linear Unbiased Predictor (BLUP) as a data fusion procedure for combining the various CS measurements, the results show that predicted discharge at both locations is in very good agreement with the reference data collected for the same hydrological events. Based on a variance decomposition, this approach allows us as well to properly assess the respective part of the random errors that are related to the citizens and measuring devices.

It is concluded that CS-based discharge data collection is a promising cost-effective way for obtaining reliable and numerous measurements. The monitoring approach based on the use of a mobile phone application is thus quite valuable for complementing the existing Tunisian monitoring system, as well as for empowering local communities. Furthermore, this approach can be applied at larger scales in the country by involving more citizens and adding other sites, which should support the national efforts for better and smarter water resources management.

5.1 Introduction

EFFICIENT strategies of water resources management require the availability of reliable hydrometric quantities such as rivers' discharge and rainfall. These data are needed to assess and model catchments' hydrological behavior over time. Nevertheless, in many regions of the world, the availability of water-related data (especially discharge) is still a major concern, especially in Africa (Hannah et al., 2011). One of the main reasons are the high acquisition and maintenance costs of the traditional discharge measurement methods (which are relying on advanced sensors like pressure transducers); costly, dangerous, and time-consuming site management (e.g., improvement of cross-sections); and expert knowledge (Davids et al., 2019). Within the African regions, Tunisia is no exception with respect to this issue (Fehri et al., 2019; Sellami et al., 2013). Despite the recent governmental efforts to boost the availability and accuracy of relevant hydrological data, several issues persist. The access to data is still not publically available and is limited to specific governmental institutions, which represents a major disadvantage for local groups and stakeholders. In addition, current official discharge monitoring networks are not spatially dense and suffer from data gaps, leading to scarce and incomplete time series and consequently poor quality databases (Fehri, Khlifi, & Vanclooster, 2020).

By the light of these challenges and accounting for the growing demand of water-related data, alternative cost-effective discharge monitoring solutions have been developed thanks to the availability of new smart sensing technologies, offering novel approaches that can be used for improving the availability and reliability of hydrometric information (Paul et al., 2018). As an example, discharge estimation using remote sensing techniques have been successfully used in several studies (Brakenridge et al., 2007; Tarpanelli et al., 2011; Van Dijk et al., 2016). However, applications of remote sensing-based technology in the context of small streams remain problematic (Tauro et al., 2018) and motivate the use of ground-based indirect discharge measurements. In this context, image and video analysis techniques such as Particle Tracking Velocimetry (PTV), Particle Image Velocimetry (PIV), Surface Structure Imaging Velocimetry (SSIV), and Large-Scale Particle Image Velocimetry (LSPIV) are being widely used for extracting water level and surface flow velocity using simple and inexpensive equipment such as fixed cameras and smartphones ((Kantoush et

al., 2011; Lüthi et al., 2018; Nezu & Sanjou, 2011)). Indeed, camera-based discharge measurements become more available and can be used in floods assessment and rating curve establishment (Dramais et al., 2011; Le Bour-sicaud et al., 2016). In parallel, over the last decade, the involvement of citizen scientists through Citizen Science (CS) initiatives has been emerging. It can potentially help filling the existing hydrological data gap by generating additional hydrologic information, including discharge (Davids et al., 2019; Fehri, Khelifi, & Vanclooster, 2020; Fujita et al., 2013; Le Coz et al., 2016; See, 2019).

Despite the rapid expansion of CS and the activity of community-based monitoring, more efforts need to be put to improving the quality of CS-based data, as they are collected by non-specialists using non-reference methods. The quality of the citizen-collected data might vary depending on the design of the CS project. For instance, projects that engage volunteers to collect data regardless of the quality of the outcomes are prone to high error rates associated with citizens' measurements. On the other hand, projects that include a training phase as well as precise monitoring guidelines before the start of the data collection should expect less errors, which helps the validation process of the data. Yet, there is always a need for methods that aim at improving the final quality of CS-collected data if the goal is to combine them at some point with existing databases. Among these methods, data fusion techniques are widely found in the literature. They are based on the idea that combining at best information coming from different sources with unusual formats, different resolutions, units, and locations contributes to improve their final accuracy (see, e.g., Fasbender et al. (2009) and Bogaert and Fasbender (2007)).

In order to improve water monitoring networks in Tunisia and since no existing CS initiatives were already reported to monitor water-related data in the country, the Together4Water project was started in October 2018 in collaboration with local Tunisian institutes and schools (see Together4Water homepage) in a test region "Medjez-El-Beb city" (Fehri, Khelifi, & Vanclooster, 2020). The project aims at engaging the population in environmental monitoring and knowledge generation to support the existing water-related databases. This should also foster the local communities' awareness about their environment and empower them to take actions and contribute to the sustainable management of water resources. Consequently, one goal of the project is to involve citizens from different generations in monitoring discharge of the Medjerda river at two different locations (denoted Medjez-N5 and Slouguia) using the publicly available

smartphone application "Discharge app" that estimates river discharge based on the processing of a five-seconds video recording (Photrack, 2018a). The application was first successfully used by the "Global iMoMo Initiative" to monitor discharge of the Themis River in Tanzania in collaboration with local stakeholders (Photrack, 2018b). In our study case, more attention will be paid to the validation of the citizen-based data as well as to the fusion of the various measurements.

In the first part of this chapter, we will shortly describe the CS approach that is adopted to engage citizens in the region and the way they were trained to properly use the smartphone application. We subsequently assess the quality of the CS-based discharge measurements by comparing them to two official reference stations located at the same places and maintained by the local water authority. In the second part, a fusion of these citizens-based measurements is proposed using a Best Linear Unbiased Predictor (BLUP) approach, that aims at combining them for obtaining a single and improved estimation. Results of this data fusion will be compared to the reference discharges in order to assess the quality of the prediction. Two main objectives will thus be achieved here, i.e. (i) to show that citizens-based discharge measurement using a smartphone application can provide sound estimation of the real discharge and (ii) to emphasize that data fusion is relevant for combining these measurements, thus delivering higher quality data at the end. As this quality is depending both on the measurement errors caused by the citizens and by the smartphone application, a special attention will be paid to the assessment of the respective contribution of these two sources of errors.

5.2 Materials and methods

5.2.1 Study area

Known for its strategic location in North Tunisia, the test region "Medjez-El-Beb city" is located in the center of the Medjerda basin (see Figure 5.1), about 60 km from Tunis (the capital of Tunisia) and 50 km from Beja (the capital of the governorate). The study area is characterized by a semi-arid climate with an average annual rainfall of 420 mm (DGRE, 2016). The Medjerda catchment is the most important freshwater resource in the country. It covers a total area of 23,700 km^2 of which 33% are located in Algeria. About 84% of the basin water resources are used for agricultural purposes, while

about 10% are used for services and industries (Fehri et al., 2019). Similarly, the main economic activities of the Medjez-El-Beb region are related to agriculture and agri-food industries that exert significant pressure on the available water resources. The population within the city of Medjez-El-Beb has been growing significantly over the last two decades, from 38,964 in 2004 to 41,749 in 2014 (INS, 2014). Multiple schools, institutes and NGO's are located in the city, which represents a suitable environment for citizen science activities.

The city of Medjez El Beb provides easy access to the existing discharge monitoring sites. The official discharge monitoring in the study area consists of two hydrometric stations (namely Medjez-N5 and Slouguia). The monitoring method includes automated stations based on water level radar sensors for measuring open channel flow. These stations send daily measurements to a centralized database. The hydrological regime at both sites is similar. The mean annual discharge is around $15 \text{ m}^3/\text{s}$ and $17 \text{ m}^3/\text{s}$ for Medjez-N5 and Slouguia, respectively, while water level is typically ranging from 0.7 to 1.8 m for Medjez-N5 and from 0.8 to 2 m for Slouguia. The physical characteristics of both sites are quite similar too, with cross-section lengths of 29 m and 30 m at Medjez-N5 and Slouguia, respectively.

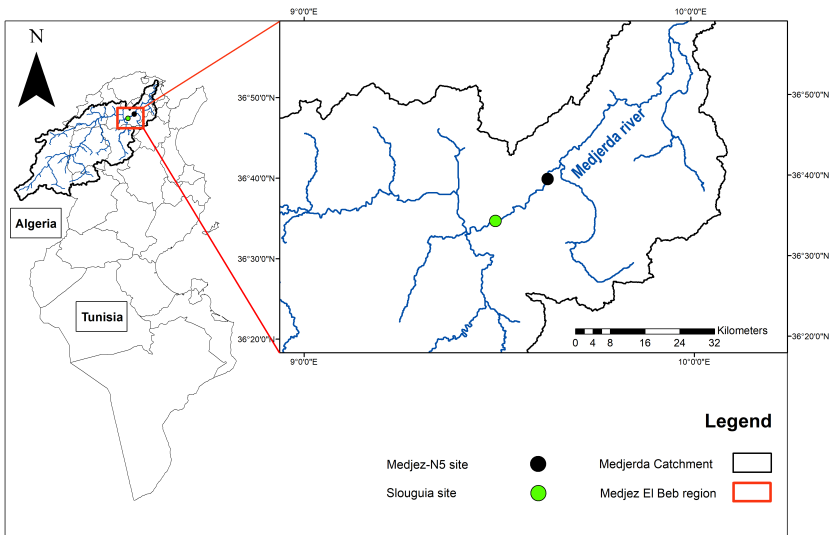


Figure 5.1: Location of Medjez-N5 and Slouguia study sites.

5.2.2 Citizens engagement methodology

Thanks to its high demography and diversified activities in both rural and non-rural environments, the Medjez-El-Beb city fulfills appropriate conditions for a successful CS program. In the framework of the Together4Water project, the citizens involvement has been realised using a step-by-step CS approach (Fehri, Khlifi, & Vanclooster, 2020) in order to ensure consistent and reliable discharge data collection using the publicly available smartphone application "Discharge app". Citizens were first reached in the region via the project website, that includes a detailed presentation of the project's goals and the available monitoring tools. The website was advertised on the local social media and at the schools and universities of the region, with the aim of reaching a wide and diversified panel of people. A total of 20 candidate citizens from different educational backgrounds and generations were afterwards selected to perform the discharge measurements based on their motivation, their location (close to the measurement sites), the fact that they own a smartphone, and their access to an internet connection. The final step of the engagement process consisted in providing them a complete field training for proper use of the measurement tool as well as for proper data transmission. The training program was organized in the framework of group meetings with citizens at the monitoring sites. Assistance was provided to explain the different steps of the measurement such as locating the sites created on the app, choosing the right standing position located on one of the riverbanks, and filming a 5-seconds video with steady hands to avoid low quality videos. A specific attention was paid to the data transmission process, which can be made immediately after the processing of the video. In addition, a list of factors that can affect the quality of the measurement was delivered and explained to citizens such as the presence of wind, light reflection on the water surface, and the color of the water. The knowledge of the impact of these factors is crucial to help citizens evaluate the quality of the measurements on site and to make multiple measurements if needed. The entire training program was video-recorded and documented in a public handbook distributed to citizens. This preparation and training step was crucial in order to ensure consistent data collection and to reduce knowledge-related issues at the start of the measurement campaign. In this study, we present the outcomes of eight citizens (four citizens per measuring site) who collected the most data from the start of the monitoring campaign.

5.2.3 Mobile phone discharge application

In this study, a publicly available smartphone application named "Discharge app" was used. The application was developed by Photrack (2018a) as a cost-effective measurement tool. The Discharge app (available for Android devices) uses the smartphone's built-in camera and accelerometer to optically measure open channels' water level and surface velocity and therefore to derive an estimation of the discharge (Carrel et al. (2019); Lüthi et al. (2014)). The approach is based on the Surface Structure Imaging Velocimetry (SSIV) technique, which is a correlation-based technique typically applied to large channels or rivers, similar to Large Scale Particle Imaging Velocimetry (LSPIV) approach (Kantoush et al. (2011); Fujita et al. (1998)).

For a proper use of the mobile phone app, a technical field preparation and setup was first needed, as described in the Discharge app developers' instructions guide (Carrel et al., 2019). Four markers were positioned on both riverbanks (Fig. 5.2) and a geometrical survey was performed at both sites in order to gather information about the exact position of the markers, the cross-section of the river, and the shoreline. For this last task, a Disto S910 telemeter (*LeicaTM*) was used. The cross-sections obtained at both sites were encoded into the Discharge app, along with the Manning-Strickler coefficient. The video resolution of the mobile phone was set at its highest value in order to maximize the accuracy of the results (for more technical details, see Appendix B).

As an illustration, Figure 5.3 presents a screenshot of the application with the window where this information was encoded. The main steps of Discharge estimation can be described as follows: (i) the selection of the implemented site (Fig. 5.3); (ii) a five-seconds video of the river flow is to be recorded. Instructions for the recording are user-specific and related to the user's standing position and the selection of the water column; (iii) after locating the four markers on both riverbanks, the intersection between the water surface and the shoreline is manually defined to characterize the actual water level (the blue line in Figure 5.2b and 5.2d); (iv) finally, the recorded video is processed locally on the smartphone using the SSIV algorithm.

The estimated discharge as well as the surface velocity are then displayed (Fig. 5.2). The theoretical aspects of the Discharge app are detailed and described by Carrel et al. (2019), Lüthi et al. (2014), and Leitão, Peña-Haro, Lüthi, Scheidegger, and de Vitry (2018).

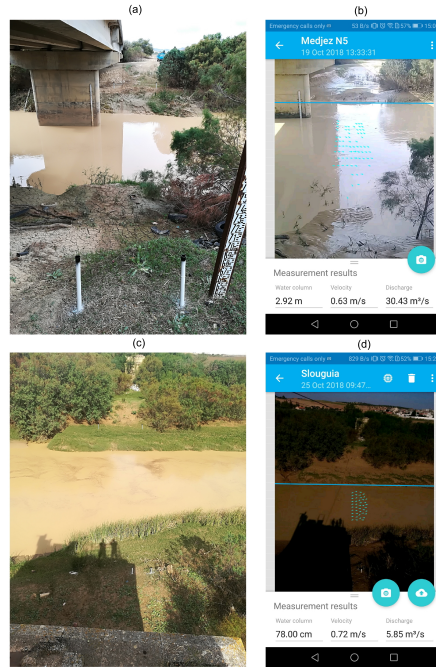


Figure 5.2: Picture of the reference sites with corresponding screenshots of the smartphone application for two events. Parts (a) and (b) refer to the Medjez-N5 site, while Parts (c) and (d) refer to the Slouguia site.

Using this mobile phone app, the data collection campaign started in October 2018 at both sites, where citizens were in charge of measuring the river discharge for a same series of hydrological events. Measurements were uploaded via the app itself to a cloud-based online platform where all measurements can be accessed for all citizens in a structured database, including the date and hour of the measurement along with the name of the volunteer. A total of 100 events were monitored by eight citizens (four citizens per site), but only a part of them was considered in this study, covering the period from October 2018 to October 2019.

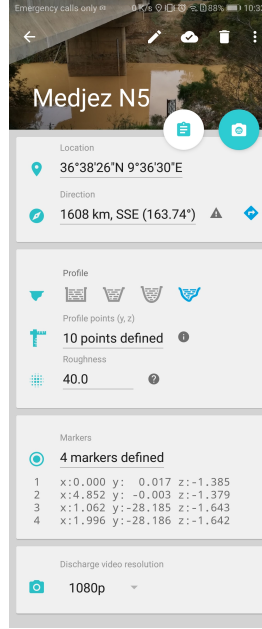


Figure 5.3: Screenshot of the app with the encoded parameters at the Medjez-N5 site.

5.2.4 Data fusion and prediction of discharge

Let us consider a given hydrological event for which the true discharge is denoted by z , while the various discharge measurements as collected from n distinct users at the same location are given by $Y_1, \dots, Y_n$. Let us consider that we can assume that $E[Y_i|z] = z$, i.e. on the average these users measurements correspond to the true discharge z . Stated in other words, the regression $E[Y|z]$ is the 1:1 line and users are all providing distinct but unbiased measurements of the true discharge z . Generalizing this for any event such that Z is a random variable too, we can thus consider the random model

$$Y_i = Z + \varepsilon_i \quad \forall i = 1, \dots, n \quad \Longleftrightarrow \quad \mathbf{Y} = Z\mathbf{1} + \boldsymbol{\varepsilon} \quad (5.1)$$

(with $\mathbf{1}$ the unit column vector of size n) where Z is the single true but unknown discharge and where $\mathbf{Y} = (Y_1, \dots, Y_n)'$ are the various corresponding users measurements. From the fact that $E[Y_i|z] = z$, the errors $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)'$ are such that $E[\boldsymbol{\varepsilon}] = \mathbf{0}$ (with $\mathbf{0}$ the null column vector of

size n), with a covariance matrix Σ_ε . We will additionally consider that $Z \perp \varepsilon$, i.e. the measurement errors ε are uncorrelated with the discharge Z .

In a data fusion framework, let us consider that what is sought for is a way to combine the various Y_i 's in order to obtain a single predictor Z^p of the unknown discharge Z . One can rely on the Best Linear Unbiased Predictor (BLUP) of Z , i.e. the best linear combination $Z^p = \sum_i \lambda_i Y_i = \lambda' \mathbf{Y}$ such that the prediction variance $\text{Var}[Z - \lambda' \mathbf{Y}]$ is minimum subject to the unbiasedness constraint $E[Z - \lambda' \mathbf{Y}] = 0$, where $\lambda = (\lambda_1, \dots, \lambda_n)'$ is a vector of weights that need to be estimated. This can be done using the Lagrangian formalism, where the objective function

$$L(\lambda, \eta) = \text{Var}[Z - \lambda' \mathbf{Y}] + 2\eta E[Z - \lambda' \mathbf{Y}] \quad (5.2)$$

needs to be minimized with respect to the weights λ and to the Lagrangian multiplier η . Eq. (5.2) can however be strongly simplified. Using eq. (5.1), the unbiasedness constraint can be rewritten as

$$\begin{aligned} E[Z - \lambda' \mathbf{Y}] &= E[Z] - E[Z] \lambda' \mathbf{1} - \lambda' \overbrace{E[\varepsilon]}^{=0} \\ &= E[Z](1 - \lambda' \mathbf{1}) = 0 \end{aligned}$$

so that the unbiasedness condition to be fulfilled is $\lambda' \mathbf{1} = 1$, i.e. the weights must sum up to one. Using again eq. (5.1), the predictor variance can now be rewritten as

$$\begin{aligned} \text{Var}[Z - \lambda' \mathbf{Y}] &= \text{Var}[Z \overbrace{(1 - \lambda' \mathbf{1})}^{=0} - \lambda' \varepsilon] \\ &= \text{Var}[\lambda' \varepsilon] = \lambda' \Sigma_\varepsilon \lambda \end{aligned} \quad (5.3)$$

Plugging these two results in eq. (5.2) now gives the classical problem of minimizing

$$L(\lambda, \eta) = \lambda' \Sigma_\varepsilon \lambda + 2\eta(1 - \lambda' \mathbf{1})$$

where $\lambda' \Sigma_\epsilon \lambda$ is a quadratic form, so that the minimum is unique and is reached when

$$\begin{cases} \frac{\partial L(\lambda, \eta)}{\partial \lambda} = 0 \\ \frac{\partial L(\lambda, \eta)}{\partial \eta} = 0 \end{cases} \iff \begin{cases} 2\Sigma_\epsilon \lambda + 2\eta \mathbf{1} = 0 \\ 1 - \lambda' \mathbf{1} = 0 \end{cases} \iff \begin{pmatrix} \Sigma_\epsilon & \mathbf{1} \\ \mathbf{1}' & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Solving this linear system of equations leads to the solution

$$\lambda = \frac{1}{\mathbf{1}' \Sigma_\epsilon^{-1} \mathbf{1}} \Sigma_\epsilon^{-1} \mathbf{1} \quad (5.4)$$

(where $\mathbf{1}' \Sigma_\epsilon^{-1} \mathbf{1}$ is a scalar), with an associated prediction variance as given by eq. (5.3) and thus equal to

$$\text{Var}[Z - \lambda' \mathbf{Y}] = \frac{1}{\mathbf{1}' \Sigma_\epsilon^{-1} \mathbf{1}} \quad (5.5)$$

Eqs. (5.4) and (5.5) thus provide a way to fuse the various discharge measurements $\mathbf{Y}$ in order to predict at best the unknown discharge Z , by providing a set of weights λ and an associated prediction variance $\text{Var}[Z - \lambda' \mathbf{Y}]$ which is measuring the performance of this predictor Z^p .

5.2.5 Sources of errors

In the previous section, we focused on the BLUP without any specific hypothesis on the covariance matrix Σ_ϵ , so that these results hold true in general. We will now focus on the structure of Σ_ϵ by considering mixed sources of errors of various origins. The first possible origin is a random error ϵ_m linked to the event, that is assumed to be random (i.e., this error will vary from one event to another one) but that affects identically all users for the same event. For our case study, this error could be linked to the smartphone application that underestimate or overestimate in a similar way for all users the discharge value of a given event, while this error differs between events. A second possible origin is a random error $\epsilon_{u,i}$ linked to the user itself, that can be assumed as independent between users and independent from ϵ_m , with a variance that only depends on the user.

These sources of errors (the smartphone application and the users) were considered based on the fact that they can easily be interpreted, which led

to this simple and classical decomposition of errors. On the other hand, a possible site effect could also be considered, but due to the fact that only two sites were at hand, no clear conclusions could be drawn, and it would be impossible to reliably estimate this kind of effect in the context of a random factors model. Furthermore, it is quite possible (and even likely) that pure random errors exist, but they could not be assessed from the data at hand. Indeed, this would require repeated measurements for the same citizen and the same event. If such pure random errors exist in our data, they are thus included into the user's variance in our model. It is worth mentioning that crossed effect between the user and the event might be possible and is not theoretically impossible, although, it would be unlikely at the same time, at least if the measurement conditions of the same event are comparable for all users. In such conditions, it is hard to imagine how there could be an interaction between the event and the user. Nevertheless, it remains true that our model could possibly include pure random errors and/or interaction effects, as long as these effects can be assessed from the data at hand and that there is some evidence to consider them in the model. It is unfortunately not the case for our data.

Stated in other words, considering that $E[\varepsilon_m] = 0$ and $E[\varepsilon_{u,i}] = 0$ and assuming first that all users have a similar variance of errors, we now have

$$\varepsilon = \varepsilon_m \mathbf{1} + \varepsilon_u$$

where ε_m is a scalar and $\varepsilon_u = (\varepsilon_{u,i}, \dots, \varepsilon_{u,n})'$, so that

$$\Sigma_\varepsilon = \sigma_m^2 \mathbf{1}\mathbf{1}' + \sigma_u^2 \mathbf{I} \quad (5.6)$$

(with $\mathbf{I}$ the identity matrix), where σ_m^2 and σ_u^2 are the measuring device's and user's variance, respectively. The corresponding correlation matrix is given by

$$\mathbf{R}_\varepsilon = \mathbf{D}_\varepsilon^{-1/2} \Sigma_\varepsilon \mathbf{D}_\varepsilon^{-1/2}$$

where $\mathbf{D}_\varepsilon = \text{diag}(\Sigma_\varepsilon)$. Using eq. (5.6), this gives $\mathbf{D}_\varepsilon = (\sigma_m^2 + \sigma_u^2) \mathbf{I}$ and so the (i, j) element of $\mathbf{R}_\varepsilon$ is then given by

$$r_{ij,\varepsilon} = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_u^2} \quad \forall i \neq j \quad (5.7)$$

where $r_{ij,\varepsilon}$ (also known as the intraclass correlation coefficient in the random model terminology) is the correlation between the measurements of

the i th and j th users when they are measuring the discharge for the same event. As seen from eq. (5.7) and under the condition of eq. (5.6), this correlation coefficient corresponds to the part of the variance that comes from the measuring device, where $\sigma_m^2 + \sigma_u^2$ is the total variance. Under the condition of eq. (5.6) again, it can also be proved that all weights must then be equal to $1/n$, so that $\lambda = (1/n)\mathbf{1}$, and an explicit formulation for the prediction variance $\text{Var}[Z - \lambda'\mathbf{Y}]$ as a function of n , σ_m^2 and σ_u^2 can also be obtained (see Appendix A). In general, depending on training and experience, it could however be expected that the same variance σ_u^2 does not necessarily hold for all users, so that eq. (5.6) is more generally written as

$$\Sigma_\epsilon = \sigma_m^2 \mathbf{1}\mathbf{1}' + \mathbf{D}_u \quad (5.8)$$

where $\mathbf{D}_u$ is a diagonal matrix having $\sigma_{u,i}^2$ as i th element (i.e., the variance for the i th user) and where the weights λ are then no more identical in that case.

As the predictor $Z^p = \lambda'\mathbf{Y}$ is a linear combination of users measurements $\mathbf{Y}$, the prediction error $Z - \lambda'\mathbf{Y}$ will tend to be Gaussian distributed as long as the correlation between these measurements (as measured by eq. (5.7)) is not too high. It is then possible to provide a simple prediction interval to the discharge Z . Assuming that

$$Z - \lambda'\mathbf{Y} \sim N(\mu_p, \sigma_p^2)$$

with $\mu_p = E[Z - \lambda'\mathbf{Y}] = 0$ from the unbiasedness condition and with $\sigma_p^2 = \text{Var}[Z - \lambda'\mathbf{Y}]$ as computed from eq. (5.5), a two-sided prediction interval at the $1 - \alpha$ confidence level is then

$$-z_{1-\alpha/2} \leq \frac{Z - \lambda'\mathbf{Y}}{\sqrt{\sigma_p^2}} \leq z_{1-\alpha/2}$$

or equivalently

$$\lambda'\mathbf{Y} - z_{1-\alpha/2} \sqrt{\sigma_p^2} \leq Z \leq \lambda'\mathbf{Y} + z_{1-\alpha/2} \sqrt{\sigma_p^2} \quad (5.9)$$

where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the zero-mean unit variance Gaussian distribution. This confidence interval will narrow when n is increased, but it is incorrect to state that $\lim_{n \rightarrow \infty} \sigma_p^2 = 0$ in general, except when $\sigma_m^2 = 0$. Stated in other words, increasing the number n of users mea-

measurements will help to reduce the uncertainty about Z , but there is a lower bound equal to σ_m^2 for the prediction variance, that relates to the errors ε_m which cannot be reduced by increasing n . The proper estimation of σ_m^2 and σ_u^2 is thus important, as these variances are conditioning the sampling strategy, i.e. by assessing the benefit of including more users in the process or the benefit of improving the measuring device for reducing σ_m^2 . In practice, the estimation of σ_m^2 and σ_u^2 can be done from a set of observed hydrological events and user measurements using any standard statistical computer package. Indeed, eq. (5.1) corresponds to a one-way ANOVA model with random effect, so that a Restricted Maximum Likelihood (REML) procedure can be used for estimating at best these variance components. See, e.g., Galecki and Burzykowski (2013) and Fox (2008) for a presentation of the details related to random models and the corresponding estimation procedures.

5.3 Results and discussions

A careful quality assessment of the CS-based discharge measurements is crucial, both for evaluating the performance of the smartphone application and for assessing the consistency of these measurements over time. With these goals in mind, extreme discharge events (that represent four events out of the total) were excluded from our analysis at both sites, since measurements were inaccurate due to the difficulty of locating the markers on both sides of the river, as caused by the high water level during these events. Therefore, only discharge measurements corresponding to values lower than $35 \text{ m}^3/\text{s}$ were considered in this analysis. This represents 90% of the collected data. More precisely, 90 events were considered in this study, where each event was monitored by the same four citizens at each study site.

5.3.1 Data quality assessment

Figure 5.4 compares at the Medjez-N5 site the discharge measurements obtained from the reference station to those obtained from the four citizens using the application. In order to check for the possible presence of bias during the measurements, a regression line $E[Y_i|z] = \beta_{0,i} + \beta_{1,i}z$ was fitted for all citizens, leading to slope estimates $\hat{\beta}_{1,i}$ that range from 0.992 to 0.996. However, separate F -test based on the null hypothesis $H_0 : \beta_{0,i} =$

$0, \beta_{1,i} = 1$ lead to the conclusion that these regression lines do not significantly differ from the 1:1 line ($p_v > 0.05$ for all citizens) and therefore citizens-based discharge measurements can be considered as statistically unbiased estimates of the true discharge (as evaluated from the reference station).

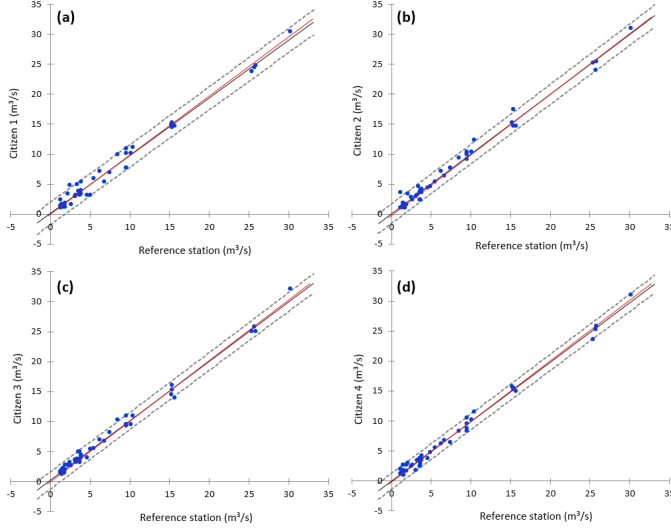


Figure 5.4: Citizens-based versus reference-based discharge measurements at the Medjez-N5 site, along with the 1:1 line (black plain line), the estimated regression line (red plain line), and the corresponding 95% prediction interval (dashed lines).

Although, small biases are always possible, but it appears to us that they could be neglected in our study. For practical applications, it seems to us that what is relevant here is the effect of possible biases by comparison with other random effects. For our data, it is clear that biases can be neglected, as long as we refer to the official gauging stations as our reference data. If the conclusion of the test would have been that biases exist, they must ideally be accounted for in the model, as long as their impact is not minimal for the final results. If these biases are coming from the users, this would lead to considering that measurements from each user should be corrected for each user's bias prior to the data fusion. If these biases come from systematic errors, they would be however much more complicated to assess, as doing so would require other reference data (ideally the real discharge values) for comparison purposes. It is still possible to conduct a sensitivity analysis in order to assess the impact that a misspecification

(e.g., in the conversion of surface velocities to depth-average velocities) might have, but still this would not allow us to assess the real impact of these misspecifications without additional reference data or alternate measurement methodologies, again for comparison purposes.

Besides the absence of bias, the R^2 values are ranging from 0.984 to 0.990 and thus indicate a good agreement between citizen-based and reference measurements. Similar observations and conclusions can be made for all citizens at the Slouguia site (not shown here). One can also see that the pattern of deviations from the 1:1 line is comparable between citizens, thus suggesting that the training provided at the start of the campaign was beneficial, as these citizens were from different generations and educational backgrounds.

5.3.2 Distribution of the errors

Focusing on the measurements errors for each citizen, it is possible to test the equality of variance between citizens, as well as to test the hypothesis that these errors are Gaussian distributed for each citizen. It's worth noting that the official discharge measurements are prone to errors. However, the data provider does not communicate the information about these errors. Accordingly, the additional variance that might be induced by such errors is not accounted for here. Variance equality was tested using the Levene's test (Levene (1960), Iachine et al. (2010)), that leads to the conclusion they can be assumed as equal between citizens for both sites (with $p_v = 0.157$ and 0.319 at the Medjez-N5 and Slouguia sites, respectively). Normality was tested separately for each citizen using the Shapiro–Wilk test (Shapiro & Wilk, 1965) and leads to the conclusion that measurement errors can be assumed as normally distributed for all citizens at both sites too (with $p_v > 0.05$ for all citizens; see Table 5.1).

Errors can be expected as being proportional to discharge in general, and this was partly observed on the raw data, that included 10% of discharge values that were above $35 \text{ m}^3/\text{s}$. From the examination of measurements above that threshold (not shown in the Thesis), they were considered as quite unreliable and they are also associated with extreme events, for which the measurement app is not expected to deliver sound results. These values were thus discarded. For the 90% of values below that threshold, no proportional effect can be observed, as illustrated in Figure 5.4. Hence, the assumption of a constant variance of errors for citizen-based discharge below the $35 \text{ m}^3/\text{s}$ threshold seems valid. Although Figure 5.4

might show slightly wider and narrower uncertainty interval between users (especially for citizen 1 and 3), the Levene statistical test confirms that all variances can indeed be considered as identical. Though the bivariate scatterplots might appear as different, testing the equality of variance only relies on the data used to build the histograms that appear on the diagonal in Figure 5.5 (that are not so different). It is worth noting too that the number of measurements is limited, so (i) the statistical power of the Levene test (i.e. its ability to detect a difference between variances when it exists) is not very high, and (ii) few values might appear as more isolated on the histograms, thus giving the impression that the distributions are different as well.

Table 5.1: Sample statistics and p -values (p_v) of the Shapiro–Wilk’s test for citizens-based (C_i) discharges at both sites, where left and right parts of the table refer to the Medjez-N5 and Slouguia site, respectively. Values $y_{(1)}$, $y_{(n)}$, $\bar{y}$ and σ are the minimum, maximum, mean and standard deviation, respectively.

	C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4
$y_{(1)}$	-1.908	-1.580	-1.800	-1.770	-1.097	-1.699	-1.699	-1.398
$y_{(n)}$	2.239	2.330	1.930	1.190	1.303	1.919	1.944	1.908
$\bar{y}$	-0.163	0.063	0.144	-0.102	-0.083	-0.184	-0.086	-0.178
σ	0.945	0.793	0.695	0.640	0.670	0.854	0.847	0.697
W	0.970	0.941	0.963	0.974	0.922	0.985	0.952	0.957
p_v	0.273	0.219	0.141	0.371	0.145	0.959	0.245	0.308

5.3.3 Variance decomposition

In order to assess the amount of intraclass correlation as defined by eq. (5.7), one can estimate the Pearson correlation coefficient $\hat{r}_{ij}$ between CS-based discharge that are measured for the same set of events, as presented in Figure 5.5 for the Medjez-N5 site. All correlations are statistically significant ($p_v > 0.05$) and range from 0.33 to 0.48, except between citizens 3 and 4 at the Medjez-N5 site. Comparable results (not shown here) are obtained at the Slouguia site.

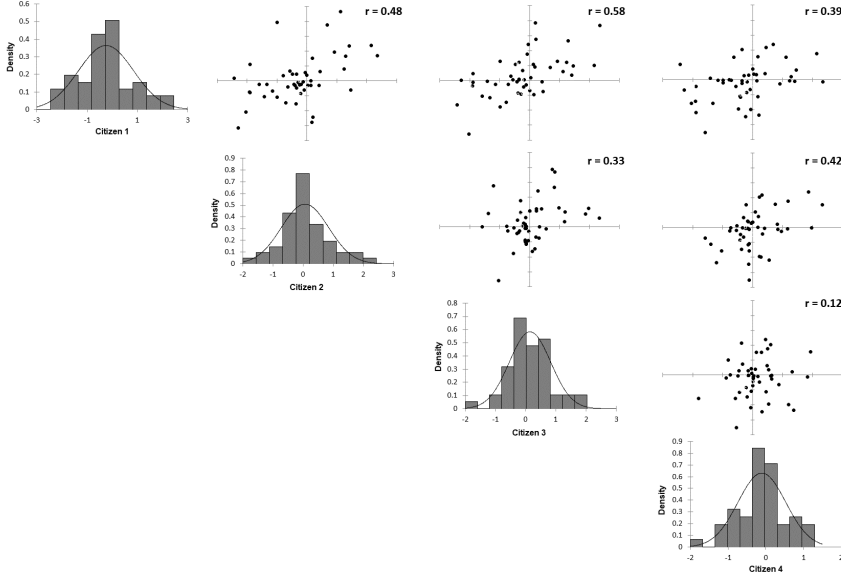


Figure 5.5: Histograms (on diagonal) and scatterplots (off diagonal) of the discharge residuals for the 4 citizens at the Medjez-N5 site, with corresponding correlation coefficients between citizens' discharge measurements for the same events.

The fact that these correlations are significant and comparable between citizens is in agreement with the hypothesis that the total variance of errors can indeed be decomposed in two parts, as suggested in Section 5.2.5, where σ_m^2 is the part of the variance linked to the smartphone application while $\sigma_{u,i}^2$ is the part linked to each citizen. As the hypothesis of variance equality between citizens has already been accepted in Section 5.3.2, we can consider that $\sigma_{u,i}^2 = \sigma_u^2$ for all citizens, so that a single σ_u^2 value need to be estimated for each site.

Being able to assess the respective amount of each source of errors is important in our context, as it eases the identification of the part of the measurement chain that need to be improved in priority. It also allows us to assess the expected benefit of increasing the number of citizens measuring the same event in order to reduce the error, knowing that σ_m^2 is a lower bound for the total variance, as previously explained.

Based on a one-way ANOVA model with random effect and using a REML procedure, the corresponding results are given in Table 5.2. It can be seen that the estimated part of the variance σ_m^2 related to the smartphone application is equal to 37% and 39% of the total variance at the Medjez-N5

and Slouguia sites, respectively, so that the variance σ_u^2 attributed to the citizens account for more than 60 % of the total at both sites.

Table 5.2: REML estimates of σ_m^2 and σ_u^2 at the Medjez-N5 and Slouguia sites, along with their respective contribution (in %) to the total variance $\sigma_m^2 + \sigma_u^2$.

	Medjez-N5	Slouguia
σ_m^2	0.259 (36.8%)	0.398 (38.7%)
σ_u^2	0.444 (63.2%)	0.630 (61.3%)
total	0.703 (100%)	1.028 (100%)

As stated before, ε_m represents the random error related to the event, which varies between events but which identically affects all citizens for the same event. In our study, the main reason for this effect is attributed to the smartphone application. Its importance can be reduced by improving the sites setup by using a better geometrical survey of the cross-section, with more visible markers at both river banks (especially for high discharge events). In parallel, a better estimation of the Manning-Strickler coefficient that was empirically selected and encoded in the application, is expected to reduce this random error and lead to an improved discharge estimation. On the other hand, the $\varepsilon_{u,i}$ error is related to the user itself and depends on various factors such as the citizen's position on the river-bank while recording the video sequence, the absence of hand movement during the recording, and the quality of the camera of the citizen's mobile phone (mainly in terms of video resolution). With respect to the reference data, it is worth mentioning that the official measurements at both sites are also affected by errors. However, the quantitative information about these errors is not available and is limited to specific governmental institutions.

5.3.4 Data fusion

Following the reasoning in Sections 5.2.4 and 5.2.5, one can use eqs. (5.4) and (5.5) for fusing the various CS-based discharge measurements of the same event and for providing a two-sided prediction interval for the true discharge using eq. (5.9). Corresponding results are shown in Figure 5.6 for both sites. For the sake of comparison with the initial results obtained separately for each citizen, Table 5.3 provides the estimated variances of the errors around the 1:1 line, along with the corresponding R^2 values.

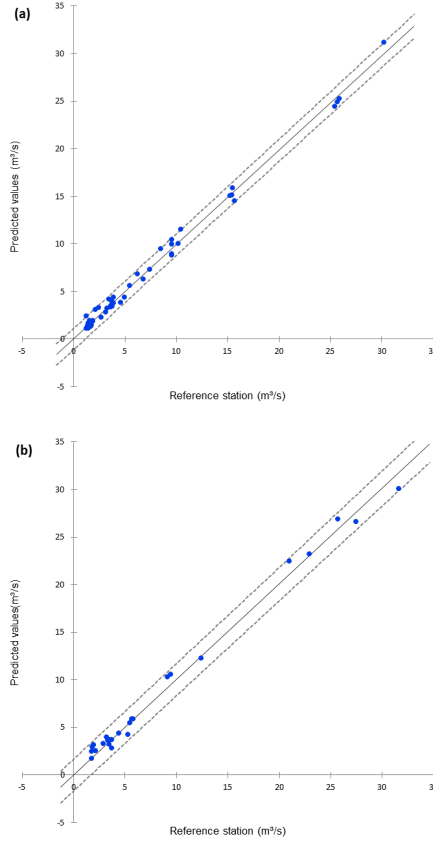


Figure 5.6: Predicted (fused) citizens-based discharges versus reference discharges, along with the corresponding prediction intervals (dashed lines) around the 1:1 line (plain line). Part (a) refers to the Medjez-N5 site, while Part (b) refers to the Slougua site.

One can see from the reduced S^2 and increased R^2 values that fusing CS-based discharge measurements leads to improved results, with a good agreement between the fused predicted discharges and the reference discharge at both sites. The 95% prediction intervals that are obtained under a Gaussian distribution hypothesis are thus reduced and remain in good agreement with the fused data as well. Though the benefit of the fusion procedure might appear as modest here, it should also be reminded that (i) σ_m^2 (the part of the variance that cannot be reduced by fusing measurements) represents here about 40% of the total variance at both sites, and (ii) the fusion only involved eight citizens' measurements in our study. As a

result, the fusion procedure leads here to a moderate reduction of the total variance, but this variance reduction can however greatly vary in general, depending on the study site and the number of involved citizens.

Table 5.3: Estimated variances S_i^2 of the four citizens-based (C_i) discharge errors around the 1:1 line and the corresponding R_i^2 values, along with the results for the fused measurements (last line). Left and right parts of the table refer to the Medjez-N5 and Slouguia site, respectively.

C_i	S_i^2	R_i^2	S_i^2	R_i^2
1	0.8737	0.9838	1.1557	0.9854
2	0.6157	0.9883	2.1590	0.9720
3	0.4722	0.9897	1.2191	0.9859
4	0.4002	0.9894	0.9815	0.9875
fused	0.2682	0.9970	0.4141	0.9960

5.4 Conclusions

In the framework of the Together4Water project, discharge data were collected for a series of 90 hydrological events by a group of 8 citizens (4 citizens per site) using a publicly available smartphone application. In this chapter, a Best Linear Unbiased Predictor (BLUP) method was used as a data fusion procedure in order to combine various CS-based discharge measurements at two locations along the Medjerda river, our test river in Tunisia. The major conclusions are as follows:

1. The step-by-step CS approach used in this study was successful, as we were able to engage and motivate a group of citizens from different generations and educational backgrounds. The training program was crucial to streamline the data collection process before the start of the monitoring campaign, which helped to guarantee a consistent measurement of the discharge as well as appropriate data transmission.
2. For low to moderate discharge values ($< 35 \text{ m}^3/\text{s}$), the use of the mobile phone application provides promising results at both sites and for all users. The good quality (absence of bias and high correlation with true discharges) of these results was confirmed from the various analyses. For extreme discharge events ($> 35 \text{ m}^3/\text{s}$), citizens

were however unable to provide reliable measurements (difficulty to access the sites, to locate the markers on both sides of the river, etc.) and more work still needs to be done for handling these situations.

3. A random model and a REML procedure were successfully used to model and to estimate the respective contribution of two sources of errors. For our study sites, these results have shown that the errors linked to the event/measuring device (ε_m) and to the users (ε_u) represent about 40% and 60% of the total variance, respectively. Reducing ε_m would require improving the mobile phone application and measuring conditions (more precise geometrical survey, better positioning of the markers, better estimation of the Manning-Strickler coefficient). Reducing the final impact of ε_u was however successfully done by simply fusing the various citizens' discharge measurements.
4. Although no firm conclusions can be obtained so far for other sites that might present other challenges for collecting and fusing CS-based discharge measurements, this study still shows that the use of a low-cost measuring device (the smartphone application) by non-specialist volunteers (our trained citizens) is able to provide reliable results for discharge measurements and is thus a very promising approach for improving water discharge monitoring in countries that have limited monitoring networks.

6

The potential of Citizen Science in SDG 6 monitoring

Abstract

THE UN Sustainable Development Goal, SDG 6, of the 2030 Agenda is a dedicated goal for water that aims at ensuring the availability and sustainable management of water and sanitation for all. To achieve this goal at national and regional scales, the availability of consistent and reliable data is crucial. In Tunisia, data availability and quality is still a point of concern. Nevertheless, Citizen Science (CS) represents a promising source of data that could be utilized for monitoring the implementation of SDG 6. In this regard, the Together4Water project was launched in Tunisia to foster the engagement of citizens in environmental monitoring. One of the main objectives of the project is to support the monitoring of SDG 6 at the scale of the Medjerda basin. This study presents an assessment of the potential contribution of the Together4Water initiative to the different SDG 6 indicators. Our review shows that the CS data could potentially contribute directly to 6 of the 11 indicators, which represents around 54% of the total SDG 6 indicators. We conclude that the Together4Water initiative represents a new source of data that could be used for supporting the monitoring of the SDG 6 indicators at the scale of the Medjerda basin in Tunisia.

6.1 Introduction

Sustainable management of water resources require more efficient approaches that include all parts of the water cycle integrated strategies that can balance the different needs and pressures. The the United Nations (UN) 2030 Agenda for sustainable development (UN-Water, 2017a) was introduced as new sets of goals (17 goals and 169 targets) that aim at improving the global well being through a bottom-up approach. A specific goal for water, SDG 6, has been included in the agenda that calls on states to “Ensure availability and sustainable management of water and sanitation for all” (UNESCO, 2016).

Several studies analyzed and examined SDG 6 at different levels in many parts of the world to understand the interlinkages between the indicators and propose different strategies for achieving the goal, which could pave the way for achieving other SDGs. For example, Hall, Ranganathan, and C. (2017) showed that SDG 6 has existing interconnections with other goals and targets such as SDG 1 on *ending poverty in all its forms everywhere*, SDG 2 on *hunger and achieving food security*, and SDG 5 on *gender equality*. In addition, Nhemaehena et al. (2018) developed a composite baseline index of agriculture related SDGs (SDG 1, 2, 6, 7 and 15) to evaluate the performance of Southern African countries in meeting the identified goals. Their results showed that the SDG 6 water stress indicator recorded the lowest performance among most countries, which highlights the impact of water scarcity in the region. Also, several factors are essential for meeting SDG 6 at national level. For instance Martínez-Córdoba, Raimo, Vitolla, and Benito (2020) suggested that the taxes associated with water supply and sanitation services, local government budget revenues, the income of the inhabitants of the municipality, and population density are favorable factors for reaching the SDG 6 by Spanish local governments. The fragmentation of local governments however is an unfavorable factor for reaching SDG 6.

The United Nations has defined 8 targets and 11 indicators for SDG 6. The computation methods of the SDG 6 indicators were developed based on the existence of a global methodology and data availability, therefore taking into account a potential global and local implementation and monitoring.

In the context of this SDG 6 indicators framework, data availability and accessibility are essential requirements in the computation and monitoring

process. However, in many parts of the world, the current state of data with respect to availability and access is still a major point of concern that needs to be addressed. Nevertheless, alternative data sources have been introduced in the computation of SDG 6 indicators. For example, international space agencies, such as NASA, JAXA, and ESA, provide a wide range of data acquired by advanced sensors and at different spatial and temporal resolutions such as rainfall, evapotranspiration, land cover, and land use. Fehri et al. (2019) showed that irrigation water consumption, estimated from remote sensing products, can be integrated in the calculation of SDG 6 indicators, particularly for indicator 6.4.2 related to water stress. Also, Dickens, Rebelo, and Nhamo (2017) used satellite images to describe the location and spatial extent of water-related ecosystems in order to assess the 6.6.1 indicator.

Another strategy that could be followed is the use of citizen-generated data. Over the last decade, Citizen Science (CS) and community-based data collection approaches have been evolving with the emergence of smart cost-effective sensors, providing new source of information that could be used by local authorities as surrogate or alternative to the nationally produced in-situ environmental data (Buytaert et al., 2016; Fehri, Bogaert, et al., 2020; Fehri, Khelifi, & Vanclooster, 2020). This represents a valuable support for the implementation process of SDG 6 given the fact that most of the indicators are data driven that take into account a variety of parameters (UN-Water, 2017a, 2017b; Vanham et al., 2018). Fraisl et al. (2020) showed that CS is already contributing or could contribute to around 33% of the entire SDGs indicators (246 indicators), which thus highlights the potential of the community contribution to the implementation of the 2030 Agenda.

Tunisia, as one of the members of the United Nations, adopted the SDGs, particularly SDG 6 in order to improve the national water resources management. However, so far, only few initiatives were taken for shaping the monitoring of the SDG 6 in Tunisia, considering the specificity of the Tunisian water system context (UNOSD, 2019). Indeed, data availability represents a major constraint towards implementing SDG 6 in the country (Fehri et al., 2019). Also access to data through the official administration remains a bottleneck. As an example, Fehri et al. (2019) disaggregated the water stress indicator (6.4.2) at different spatial and temporal scales in the Medjerda catchment in Northern Tunisia enriching governmental data with remote sensing products.

Considering the growing potential of CS as a valuable alternative contributor to the monitoring of progress of the SDGs (Fritz et al., 2019; Sarah

& Rachel, 2017), particularly to SDG 6, the Together4Water project (<https://2gether4water.sipr.ucl.ac.be/>) was launched in Tunisia in 2018. The project aims to support the environmental monitoring using cost-effective smart technologies, to reinforce the existing official databases, and to contribute to the implementation of SDG 6 in the country (Fehri, Bogaert, et al., 2020; Fehri, Khlifi, & Vanclooster, 2020). The project includes the monitoring of rainfall, discharge, and water quality in a pilot area of the Medjerda catchment.

To our knowledge, there are no existing studies that relate CS to the monitoring of SDG 6 in Tunisia, which represents a major challenge for Together4water in order to provide the first steps towards implementing an operational CS-based SDG 6 monitoring system in the country. In this study, we assess the potential contribution of the Together4Water citizen science initiative in the calculation of SDG 6 indicators in Tunisia based on the outcomes of the citizen-based data collection. In the first part of this chapter, we will shortly describe the environmental monitoring activities of the Together4Water project. We then summarize the SDG 6 indicators' review process. In the second part, we assess the possible contributions of the Together4Water project to the monitoring of the SDG 6 indicators. Two main objectives will thus be achieved here, i.e. (i) to show that CS could represent a reliable source of environmental data in Tunisia and (ii) to highlight the potential contribution of such a novel approach in the monitoring of SDG 6 indicators in the country.

6.2 Materials and Methods

6.2.1 Study area

Tunisia, as most of the southern Mediterranean regions, is characterized by arid and semi-arid climates and suffers from limited water supplies. The biggest portion of the country's water resources is originated in the North, in particular in the Medjerda catchment (Fig. 6.1), which represents the most important river basin in Tunisia. The Medjerda covers a total area of $23,700 \text{ km}^2$, of which $7,870 \text{ km}^2$ (33%) are located upstream in Algeria. The population within the Medjerda catchment has been growing over the last decades and it is estimated to be around 3 million in 2022 (15% of country's population, (INS, 2014)). The Medjerda represents 9.8% of the land area of Tunisia, which makes it the largest river basin in the country. Agricul-

ture dominates the economic activities in the basin, as it contributes with 51% to the total national wheat production. Agricultural activities exert the most important pressure on water resources in the basin as 84% of the total available freshwater is used for agriculture and 9.4% of the total basin area is considered as irrigated land (Fehri et al., 2019). On the other hand, industry is also an important economic activity in the catchment, mostly within the agri-food sector. Similarly, services such as drinking water supply, represent a pivotal economic activity in the study region. According to (INS, 2014), around 60% of the Tunisian population is provided with drinking water from the Sidi Salem reservoir, which is located in the northern part of the basin.

Not only water shortage is considered as a main concern in the study region, water quality issues are also exerting an additional burden on the available water resources. The high salinity of both surface and ground water affects the water quality and therefore the crop production as well as industry and drinking water supply. This is a particular point of concern in the downstream coastal part of the catchment (Bouraoui et al., 2005).

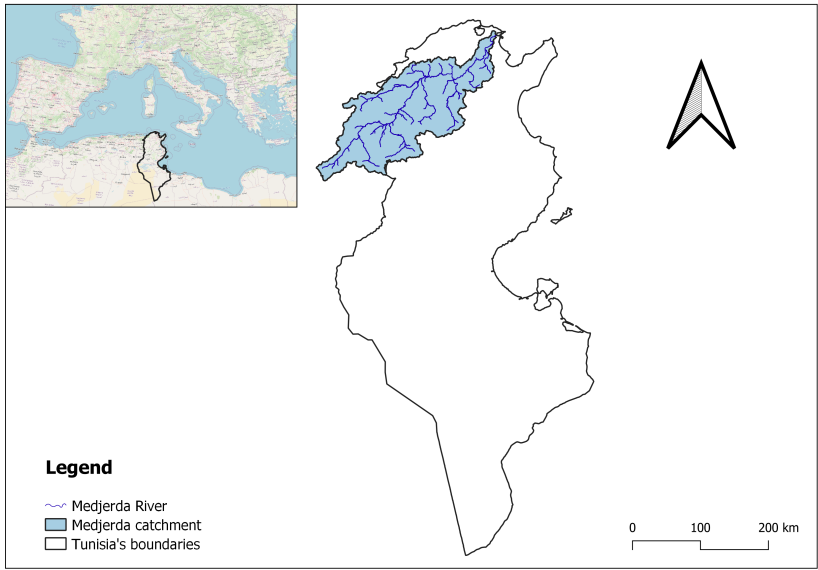


Figure 6.1: Study region.

6.2.2 The Together4Water Citizen Science project

The Together4Water project is, upon our knowledge, the first CS initiative on water in Tunisia. It was launched in October 2018 in the framework of a partnership between Tunisian institutions and UCLouvain in Belgium and was implemented in a test region within the Medjerda catchment "Medjez El Beb region" (Fehri, Bogaert, et al., 2020). The CS approach of Together4Water was designed using a Step-by-Step methodology (Fehri, Khelifi, & Vanclooster, 2020), which builds on 3 main tasks (1) the engagement of citizens in the Medjerda catchment through social media and the Together4Water online platform; (2) the selection of the monitoring sites and tools; and (3) the training and capacity building of the engaged citizens, i.e. training on the use of the monitoring tools and the data transmission protocols. The main goal of the project is to design a new environmental monitoring strategy that involves citizens in the Medjerda catchment. The project could potentially reinforce the existing official water monitoring capacity with respect to rainfall, discharge, and water quality (Total Alkalinity, Nitrate (NO₃), and Chloride) (Fig. 6.2). For rainfall, traditional rain gauges were used at multiple locations for daily measurements. The data were transmitted via the Together4Water online platform using the instant chat box. Discharge was measured using the "Discharge app" smartphone application at two sites in the Medjerda catchment. The application estimates discharge based on the processing of a five-second video recording of the river flow (Carrel et al., 2019; Fehri, Bogaert, et al., 2020). The transmission of the observations was done using the built-in data transfer tool within the application itself, which sends measurements to the "Discharge app" online platform where data can be visualized and downloaded by the administrator (Fehri, Bogaert, et al., 2020; Photrack, 2018a). With respect to water quality, total alkalinity, nitrate, and chloride content were monitored using cost-effective water quality test strips. In addition, domestic water quality was monitored as well. The Together4Water project aims also the strengthening of the local communities capacity to take actions themselves and contribute to their local water resources management strategies through data collection, collective knowledge generation and collective decision making. In addition, and given the importance of SDG 6 framework in Tunisia, one of the specific objectives of Together4Water is to support the monitoring and implementation of SDG 6 indicators in Tunisia.



Figure 6.2: Citizen science data collection tools.

6.2.3 Review of SDG 6 indicators

To assess the potential contribution of the Together4Water CS project to the monitoring of the SDG 6 indicators in Tunisia, we used a three-step review methodology (Fig. 6.3). The first step was to consult the UN Statistics Division online resources and compile a list of the SDG 6 indicators (Table 6.1). Then, we performed a thorough review of the metadata documents of the indicators that describe the methodology and data sources associated with each indicator (<https://unstats.un.org/sdgs/metadata/>). This step provided us with an understanding of the designed methodologies for each indicator and the needed data for the calculation process. The last step was to identify the potential contribution of the Together4Water CS project to the calculation of SDG 6 indicators. This process is based on the 3 monitored parameters by citizens: rainfall, discharge, and water quality. The identification of a possible contribution of CS data to one or multiple SDG 6 indicators requires that at least one or all 3 parameters can be used as input data in the calculation method of the indicators. Reviewing the outcomes is required to validate the identified contribution or to adjust the result depending on the indicator.

Table 6.1: Summary of SDG 6 indicators.

Target	Indicator	Monitoring methodology
Goal 6. Ensure availability and sustainable management of water and sanitation for all		
6.1 By 2030, achieve universal and equitable access to safe and affordable drinking water for all	6.1.1 Proportion of population using safely managed drinking water services	Household surveys and censuses at county, regional, and global levels.
6.2 By 2030, achieve access to adequate and equitable sanitation and hygiene for all, paying special attention to the needs of women and girls and those in vulnerable situations	6.2.1 Proportion of population using (a) safely managed sanitation services and (b) a hand-washing facility with soap and water	Household surveys and censuses at county, regional, and global levels.
6.3 By 2030, improve water quality by reducing pollution, halving the proportion of untreated wastewater and substantially increasing recycling and safe reuse globally	6.3.1 Proportion of domestic and industrial wastewater flows safely treated	The proportion of wastewater flows which are safely treated is calculated as a ratio of the amount of wastewater safely treated to the amount of wastewater generated.
	6.3.2 Proportion of bodies of water with good ambient water quality	The indicator computation method consists in two main steps: first classifying all assessed water bodies based on the compliance of the monitoring data collected for selected parameters. Then the indicator is computed using the classification results as the proportion of the number of water bodies classified with good quality to the total number of classified water bodies.
6.4 By 2030, substantially increase water-use efficiency across all sectors and ensure sustainable withdrawals and supply of freshwater and substantially reduce water scarcity	6.4.1 Change in water-use efficiency over time	The sum of the major economic sectors, weighted according to the proportion of water used by each sector over the total use.
	6.4.2 Level of water stress: freshwater withdrawal as a proportion of available freshwater resources	The freshwater withdrawal as a proportion of available freshwater resources minus the environmental flow requirements.
6.5 By 2030, implement integrated water resources management at all levels, including through transboundary cooperation as appropriate	6.5.1 Degree of integrated water resources management	Questionnaires through consultations between relevant stakeholders: ministries and institutions involved in water resources management as well as NGOs, academia, and business.
	6.5.2 Proportion of transboundary basin area with an operational arrangement for water cooperation	The total transboundary surface area in the country is the sum of the surface areas in the country of each of the transboundary basins and aquifers.
6.6 By 2020, protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes	6.6.1 Change in the extent of water-related ecosystems over time	The indicator calculation methodology includes five sub-indicators: 1) vegetated wetlands, 2) rivers and estuaries, 3) lakes, 4) aquifers, and 5) artificial water bodies.
6.a By 2030, expand international cooperation and capacity-building support to developing countries in water- and sanitation-related activities and programmes	6.a.1 Amount of water- and sanitation-related official development assistance that is part of a government-coordinated spending plan	The indicator is computed as the proportion of total water and sanitation-related Official Development Assistance (ODA) that is included in the government budget, i.e. the amount of water and sanitation-related ODA in the government budget divided by the total amount of water and sanitation-related ODA.
6.b Support and strengthen the participation of local communities in improving water and sanitation management	6.b.1 Proportion of local administrative units with established and operational policies and procedures for participation of local communities in water and sanitation management.	The indicator assesses through national, regional, and global questionnaires, the percentage of local administrative units that have an established and operational mechanism by which individuals and communities can meaningfully contribute to decisions and directions about water and sanitation management.

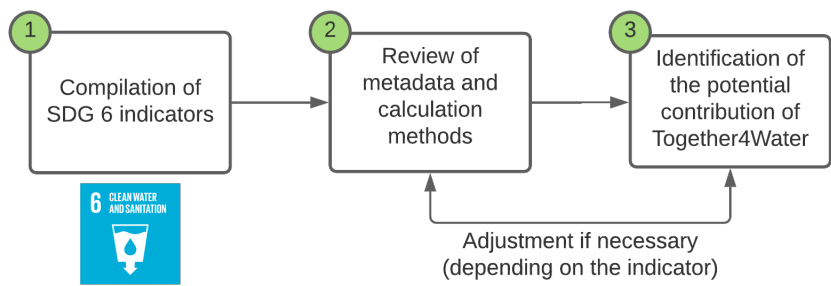


Figure 6.3: Steps of SDG 6 indicators review for the potential Together4Water contribution.

6.3 Results

6.3.1 Citizen Science data collection

The ongoing Together4Water project started the first field data collection campaign with citizens in the Medjerda catchment in October 2018. A total of 50 volunteers were involved in the monitoring process of rainfall (24 citizens), discharge (20 citizens), and water quality (6 citizens) using cost-effective tools provided to volunteers prior to the field campaign (Fig. 6.2). For rainfall, traditional rain gauges were used at multiple locations for daily observations and the data was transmitted via the Together4Water online platform using the instant chat box. The citizens' measurements delivered consistent outcomes in good agreement with the reference official rainfall data provided by the local authorities at all locations (Fig. 4.4) (Fehri, Khelifi, & Vanclooster, 2020). Regarding discharge measurement, a group of volunteers used the "Discharge app" at two sites in the test region within the Medjerda catchment. The estimated discharge showed good agreement with the reference discharge at both locations (Fig. 5.4), which reflects the high performance of the "Discharge app" in estimating river flow when used by an everyday citizen. Also, the outcomes suggest that the estimated discharge could be used as a complement for the official measurements. In addition, a data fusion procedure was applied to combine citizens measurements for both locations, which generated unique yet more robust discharge data (Fehri, Bogaert, et al., 2020). For water quality, total alkalinity, nitrate, and chloride were monitored by citizens using the test strips in several locations of the Medjerda catchment. The citizens'

measurements provided reliable estimations of all parameters and showed consistent agreement with expert laboratory analysis. The different outcomes of the citizens’ observations are summarized in (Table 6.2).

Table 6.2: Comparison of the average, minimum, and maximum of Total Alkalinity, Nitrates and Chlorides among three citizens (numbers between parentheses indicate the number of collected observations by each citizen).

Parameter	Statistic	Citizen 1 (33)		Citizen 2 (31)		Citizen 3 (32)	
		Test Strips	Lab	Test Strips	Lab	Test Strips	Lab
Total alkalinity (°f)	Average	10.90	11.20	12.88	11.52	12.53	12.13
	Min	3	3.03	2.7921	1.9374	3.3564	3.2187
	Max	20	32.57	19.14	16.3760	19.89	19.56
Nitrates (ppm)	Average	7.34	10.40	7.57	8.17	4.49	5.28
	Min	0	0.07	0	0	0	0
	Max	25	25.73	25.19	26.1615	25	25.0654
Chloride (ppm)	Average	1088.67	1263.56	2754.86	3475.07	3767.29	2897.70
	Min	370	163.80	620	702	620	732.70
	Max	3400	3726.45	6100	8383.05	7510	9695.30

6.3.2 Potential contribution of CS to SDG 6 in Tunisia

The assessment of the potential contributions of the Together4Water CS initiative in monitoring SDG 6 indicators was based on the project’s monitored parameters: water quality, rainfall, and discharge. The review of the indicators’ calculations methodologies provided a consistent understanding of the requirements of each indicator in terms of data input. A potential contribution of the Together4Water action is considered if the observed parameter could be used as input in the calculation of one or multiple SDG 6 indicators. Figure 6.4 shows the results of the SDG 6 indicator review process. Data collected in the framework of the ongoing Together4Water project could potentially contribute to 6 of the 11 indicators, which represents around 54% of the total SDG 6 indicators. In addition, the individual potential contribution of each parameter was assessed in order to determine their unique value in monitoring SDG 6 in Tunisia. Water quality data can be used in the calculation of 5 indicators: indicator 6.1.1 on *drinking water quality and management*, 6.3.2 on *the quality of natural water bodies*, 6.4.2 related to *water stress*, 6.5.1 on the *integrated water resources management*, and 6.6.1 on the *change in the extent of water-related ecosystems*. Regarding rainfall, CS observations can be used as input data for 4 indicators: indicator 6.4.1 on *water use efficiency*, 6.4.2, 6.5.1, and 6.6.1. Similarly, discharge data can also potentially be included in the calculation of the latter 4 indicators.

Moreover, based on the identified indicators, the assessment showed that 5 out of the 8 SDG 6 targets (Table 6.1) could potentially be supported by the Together4Water CS initiative, which represents around 62% of all the goals targets. This gives more value of the project’s CS monitoring program as a future reliable contributor to monitoring the implementation of SDG 6 in Tunisia.

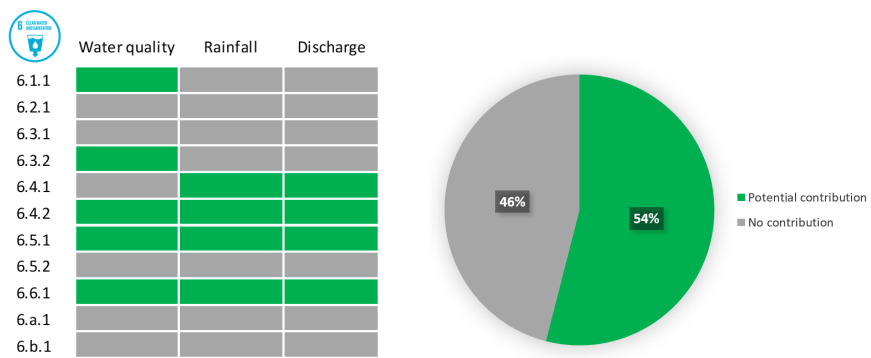


Figure 6.4: Potential contributions of the Together4Water project to SDG 6 indicators.

6.4 Discussion

Our study is the very first systematic assessment of the potential contribution of CS to SDG 6 monitoring at the indicator level in Tunisia. Other studies found in the literature introduced the analysis of the actual contribution and the potential of CS globally at the levels of all goals, targets, and indicators (Fraisl et al., 2020; Fritz et al., 2019; Sarah & Rachel, 2017). The main outcome of our assessment is that the Together4Water CS project has a real potential to be a major contributor in monitoring the implementation of SDG 6 in Tunisia through the citizen-based data collection of 3 essential parameters: water quality, rainfall, and discharge. Based on the review of all the indicators’ calculation methodologies (Table 6.1), the citizen-based data can be used as input in 6 of the 11 SDG 6 indicators, representing around 54% of the total goal’s framework. In compliance with this outcome, similar observations could be made regarding the SDG 6 targets, as 5 of the 8 targets could potentially be supported by the Together4Water CS initiative, which constitutes 62% of the goal’s targets.

Figure 6.4 shows that water quality data is an essential asset to the monitoring of SDG 6 with the most potential contributions to the indicators' framework, particularly to indicators 6.1.1, 6.3.2, 6.4.2, 6.5.1, and 6.6.1. Regarding the indicator 6.1.1 on the *proportion of population using safely managed drinking water services* (monitored globally by the Joint Monitoring Program (JMP) (JMP, 2020)), the designed calculation methodology, which is based on household surveys and censuses, includes the aspects of the quality of drinking water as a required information for the monitoring of the indicator (UNStats, 2020). Also, Moreno et al. (2020) highlighted the importance of water-quality analysis in national household surveys in Ecuador in order to support the monitoring of SDG 6.1. Similarly, the indicator 6.3.2 on the *proportion of bodies of water with good ambient water quality* relies directly on water quality input data and CS monitoring approaches for national and regional calculations (UNStats, 2020). In contrast, although the indicator 6.4.2 focuses mainly on the *level of water stress*, the water quality component is highly crucial and should be included in the monitoring process. In this regard, Vanham et al. (2018) addressed water quality as an integral part of the indicator 6.4.2 in the sense of water quality degradation which represents a form of water stress in a given region. Concerning the indicator 6.5.1 on the *degree of Integrated Water Resources Management (IWRM)*, the inclusion of water quality information is highly recommended, although the monitoring process of the indicator is based on questionnaires through consultations between relevant stakeholders such as ministries and institutions involved in water resources management as well as NGOs and academia. The information related to water quality is essential in the evaluation and scoring of the indicator since IWRM, at any given region, strongly relates to the provision of safe water in services such as drinking water and also in water-related ecosystems through recycling and safe reuse (UN-Environment, 2018). Our assessment showed also a potential contribution of water quality CS data to the indicator 6.6.1 on the *change in the extent of water-related ecosystems over time*. Indeed, the progressive monitoring approach of the indicator proposes 5 sub-indicators, 3 of them are directly related to the surface and ground water quality (Sub-indicators 2, 4, and 5) (UNStats, 2020), which emphasises its importance in the monitoring process of the indicator.

The potential contribution of rainfall and discharge CS data to the SDG 6 indicators is sound. Four indicators are directly related to both parameters. The indicators 6.4.1 and 6.4.2 (Target 6.4) on *water use efficiency and water scarcity*, include information on rainfall and discharge as main inputs

in the calculation process (Fehri et al., 2019; UNStats, 2020; Vanham et al., 2018). In addition, the indicator 6.5.1 requires a clear knowledge of the quantitative information about surface water and groundwater in order to provide appropriate evaluation of the degree of implementation of IWRM by local stakeholders in a given region (UN-Environment, 2018). With respect to the indicator 6.6.1, the progressive monitoring approach proposes 2 sub-indicators related to surface water extent and quantity (sub-indicators 1 and 3, respectively) (UNStats, 2020). Hence, rainfall and discharge are essential and represent main input data in the calculation of the indicator 6.6.1. Also, it's worth noting that CS-based methodologies were particularly encouraged in the monitoring of the sub-indicator 3 on the *quantity of water (discharge) in rivers and estuaries* in order to increase data availability especially in data-scarce regions (UNStats, 2020).

Finally, we acknowledge that CS only represents a complementary source of data for the calculation of SDG 6 indicators in Tunisia. Indeed, nationally-produced data by the official institutions represent the basis for a consistent monitoring of the indicators together with remote sensing products. Consequently, efforts are needed in order to improve the joint water monitoring and data collection activities by citizens and official institutions.

6.5 Conclusion

In the framework of the Together4Water project, citizen scientists were engaged to generate reliable environmental data that can reinforce the existing databases of the Medjerda catchment. The potential contribution of the project to the SDG 6 indicators in Tunisia was assessed based on their calculation methodologies. It is concluded that the CS data collection campaign provided reliable outcomes with respect to rainfall, discharge, and water quality. The data were validated using official expert measurements in all CS monitoring sites. In addition, the review of the calculation methodologies of the SDG 6 indicators showed that the Together4Water project could potentially contribute to 6 of the 11 indicators, which represents around 54% of the total indicator framework of the goal. Water quality has the most potential contribution to SDG 6 with 5 indicators (6.1.1, 6.3.2, 6.4.2, 6.5.1, and 6.6.1) compared to rainfall and discharge with 4 indicators (6.4.1, 6.4.2, 6.5.1, and 6.6.1).

Notwithstanding the demonstrated potential of CS to contribute to the monitoring of the implementation of the SDG 6 Agenda, some limitations

with CS should be considered. The potential of CS to contribute to the monitoring of the implementation of SDG 6 in Tunisia has only been demonstrated for a pilot project area in Tunisia. The upscaling of the approach to larger areas could be limited by several factors. For instance, the governmental strategy to report on the implementation of SDG 6 is currently based only on nationally-produced data, rather than on the combination of several sources of information (e.g remote sensing and CS). While appropriate fusion methodologies of CS and generic data with reference data could be proposed for the pilot study area, such fusion methodologies are expected to vary between different regions in terms of the quality of CS and reference data. It is therefore not clear how the data collection and fusion methodologies developed in the Together4Water project could be generalised over the whole Tunisian territory. Also, the sustainability of the CS initiative is a point of concern, since the current Together4Water was supported only by a funding mechanism limited in time. It is believed that the successful contribution of CS in monitoring the implementation of the SDG Agenda would need sustained basic funding, which goes beyond the time scale of a short term research program.

7

Conclusion and outlooks

7.1 Conclusion

7.1.1 General conclusion

THE fundamental goal of this PhD thesis was to strengthen the scientific basis for an operational monitoring system for the implementation of SDG 6 in Tunisia. A particular attention was paid to the coherence between the monitoring system and the indicators' framework with respect to the methodologies and the specificity of current and future water data structure, quality and availability in Tunisia. The Medjerda basin was used as a case study. In this last chapter of this PhD, we provide a summary of our work by going through the specific research questions and specific objectives in order to demonstrate our way of achieving them. In the second part, we give a research outlook and suggest how future research can contribute to an operational monitoring system of SDG 6 indicators in Tunisia.

7.1.2 Specific objectives

SO1: To develop a scientific-based method allowing to disaggregate an SDG 6 indicator at different spatial and temporal scales in Tunisia

In this first objective, focus was put in Chapter 2 on the disaggregation of the water stress indicator (6.4.2) at different spatial and temporal scales in the Medjerda catchment. To do so, we used the available official database together with remote sensing products. The major outcome from this analysis is that water stress has been significantly increasing in the period 2000-2016 and exceeded the threshold of 100% at the scale of the Medjerda catchment and also in 4 regions within.

Chapter 2

SO2: To evaluate the potential of Citizen Science in monitoring water-related data in Tunisia

For reaching this specific objective, it was needed to design a specific data infrastructure to collect, validate, and publish CS data for the selected case study. In Chapter 3, we set the ground for the implementation of a citizen-based water monitoring project in the region. The Together4Water online platform was built to share and dynamically visualise citizen-generated

observations as well as historic official water-related data collected from the official Tunisian water institutions. In addition, the platform provides an easy data update system, which allows to add new data in the database. In chapter 4, a citizen science approach was presented for monitoring rainfall in the selected case study. Rainfall was monitored by a group of citizens using traditional low-cost rain gauges in different locations in a pilot region in the Medjerda catchment. Chapter 4 shows that the collected data are in good quality based on validation using official rainfall observations.

Chapter 3 | Chapter 4

SO3: To develop a data fusion framework allowing to combine various citizen science measurements

In Chapter 5, we presented a CS-based approach for monitoring the Medjerda river discharge using a publicly available smartphone application "Discharge app". Our analysis showed that the citizens' data are in good agreement with official discharge data collected on 2 sites in the case study area. The Best Linear Unbiased Predictor (BLUP) was used as a data fusion procedure to combine the citizen-based discharge and generate unique robust time series of river flow which could be useful for the final users of the data.

Chapter 5

SO4: To assess the potential contribution of Citizen Science to the SDG 6 indicators in Tunisia

For the fourth objective of the thesis, we assessed the potential contribution of the Together4Water citizen science project to the SDG 6 indicators in Tunisia. We performed a review based on the outcomes of the project and the calculation methodologies of the indicators. Results showed that the Together4Water project could potentially contribute to 54% of the total goal's indicator framework through rainfall, discharge, and water quality data.

Chapter 6

7.1.3 Specific research questions

RQ1: Does the disaggregation of the SDG 6 water stress indicator (6.4.2) at different temporal and spatial scales provides a better understanding of water stress issues?

The United Nations as well as the UN Statistics Division, provide several recommendations for a consistent monitoring of the SDGs at all levels such as including remote sensing data, disaggregating the indicators at different scales, and considering an uncertainty analysis for each outcome. For SDG 6, we were able to calculate the indicator 6.4.2 on Water Stress (WS) using the gathered official data at different spatial (the whole catchment and 4 governorates) and temporal (yearly and monthly) scales in the Medjerda catchment. Remote sensing products were also used to estimate irrigation water consumption in order to overcome the lack of this particular information since its poorly monitored at national level. Chapter 2 shows that using both data sources (governmental data and remote sensing products), the level of WS has been increasing in the Medjerda in the period 2000-2016. The catchment exceeded 100% of WS, which indicates a critical water situation. The spatial disaggregation of the indicator at the scale of 4 gouvernorates within the catchment showed similar behavior with increasing yearly levels of WS. At monthly scale, the behavior of WS showed a significant seasonality as it increases in the end of the fall and during the summer, while it slightly decreases in the other seasons. In addition, we performed an uncertainty analysis of the indicator at yearly scale. Results showed that the highest uncertainty contribution is related to discharge data while remote sensing-based estimation of the irrigation water consumption reduced the uncertainty related to the official irrigation data.

RQ2: How should water-related data be collected and customized to improve its quality and availability?

The SDG 6 comes with 11 data-driven indicators (Table 6.1) that are linked to several water-related issues such as drinking water, sanitation, water use efficiency, and water scarcity. To be able to monitor one or some of the indicators, we performed a data collection campaign that allowed us to gather a variety of water-related information that could be included as input in the calculation of the SDG 6 indicators. Several Tunisian institutions such as the DGRE, SONEDE, ONAS, and the Ministry of Agriculture, were consulted to access the data and to acquire them in a standard format (numer-

ical format). The collected data were upgraded in a conventional format in order to facilitate their management, analysis, and transfer. Nevertheless, the aim of this specific question is not only to prepare a local database for the study but also set the ground for a citizen-based water monitoring project "Together4Water" through the creation of a dedicated online platform. The platform ensures the accessibility of the collected historic data and the citizen-generated ones, which represents a massive improvement of the value of the water-related database. As shown in Chapter 3, an interactive online platform (<https://2gether4water.sipr.ucl.ac.be/>) was created for that purpose. The website provides dynamic visualisations of the collected data as well as a standardized format according to Open Geospatial Consortium (OGC) WaterML2.0 standard. Among the different characteristics of the website, an instant chat box was setup to support the Citizen Science data collection campaigns (Chapter 4).

RQ3: What is the potential intrinsic value of water information data collected in Tunisia using Citizen Science-based approaches?

In Chapter 4, we presented Citizen Science (CS) as a novel approach of environmental monitoring and data collection using cost-effective tools and based on the engagement of citizens from the Medjerda catchment. We launched the Together4Water project in Tunisia as the first CS initiative in the country with an overall objective to reinforce the existing water-related database and to support the monitoring of SDG 6 indicators. We focused on 3 main parameters: rainfall, discharge, and water quality. As illustrated in Chapter 4, we followed a Step-by-Step methodology to reach and engage citizen in a pilot region of the Medjerda catchment (Medjez El Beb) to monitor rainfall using traditional low-cost rain gauges. The Together4Water online platform was used as a key tool in reaching and interacting with citizens in the early stages of the project. The engaged volunteers monitored rainfall at multiple locations in Medjez El Beb (Medjez El Beb City Center, ESIM, and Slouguia). The data transmission was performed using the instant chat box of the Together4Water website. Results showed that the citizens observations are in good agreement with the official rainfall data at all locations, which indicates the reliability of the outcomes considering the used low-cost tools in the monitoring process.

RQ4: How can citizen science-based water information data optimally be combined with existing reference data from different water-related services?

Chapter 5 highlights the importance of novel technology in water monitoring, particularly with respect to discharge estimation. Discharge app is a smartphone application developed by (Lüthi et al., 2014) that estimates river discharge by processing a 5-second video of the flow. This app was first tested in Belgium at 3 different locations by an expert user in order to evaluate its performance (Appendix B). Results of the test showed that the app's estimates are in good agreement with the official discharge measurements provided by Aqualim project in all locations. This highlights the potential of the app in delivering consistent estimations of discharge. In the context of CS, the smartphone application was used by citizens in the Medjerda catchment to measure discharge at two locations (Medjez-N5 and SLouguia). The volunteers were trained first on how to use the app and how to perform the data transmission to the dedicated Discharge app cloud. Chapter 5 shows the outcomes of the CS discharge monitoring. Results yield good agreement between the citizen-based estimations and the official discharge data provided by the DGRE at both locations. Although, in order to increase the quality of the estimated discharge, a data fusion procedure was used. We proposed the Best Linear Unbiased Predictor (BLUP) technique to combine the citizen-based data at both locations to generate unique more robust time series of discharge, which could be more useful for the final user of the data (e.g. the Ministry of Agriculture). The fused discharge showed good agreement with the official data at both locations, which suggests that similar approaches can be applied to different CS data to improve their quality and to be easily used by the official institutions.

RQ5: What's the potential contribution of Citizen Science data to the monitoring of SDG 6 indicators in Tunisia ?

In this last chapter of the thesis, the potential contribution of CS to the monitoring of SDG 6 indicators was assessed. A thorough review of the goal's indicators was performed to identify possible uses of the Together4Water outcomes as input data in the calculation methodologies of the indicators. Chapter 6 shows that the Together4Water project could potentially contribute to 6 of the 11 SDG 6 indicators at the scale of the Medjerda catchment, which represents around 54% of the total goal's framework. Our review showed also that Water quality data has the most potential contri-

bution to SDG 6 with 5 indicators (6.1.1, 6.3.1, 6.4.2, 6.5.1, and 6.6.1). On the other hand, rainfall and discharge data could contribute to 4 indicators (6.4.1, 6.4.2, 6.5.1, and 6.6.1). Hence, our assessment emphasizes the importance of CS as complementary source of data towards achieving the SDG 6 in Tunisia.

7.2 Perspectives

In the previous chapters, we presented our approach for strengthening the scientific basis for monitoring the implementation of SDG 6 in Tunisia, in particular at the scale of the Medjerda catchment. Nevertheless, several other open questions remain. In the following, we suggest some potential ideas that are not yet fully explored in the current literature with respect to the Tunisian case or at larger scale (globally).

7.2.1 More advanced and large scale implementation of Citizen Science

Our work proved that CS, in the framework of the Together4Water project, has the potential to deliver reliable environmental data based on low-cost tools and a proper engagement and training of citizens. Although, the scale of the implementation of the project is still limited to a small region within the Medjerda catchment. The limited scale represents a constraint towards a more consistent monitoring of water resources at the basin scale and the national scale as well. Therefore, an upscaling of the Together4Water initiative is crucial and should be considered to increase the monitoring sites, the number of volunteers, and the target environmental parameters. In addition, the inclusion of local stakeholders as well as governmental representatives should add more value to the CS data at both regional and national levels. Another essential element of the Together4Water project is the online platform, which represents a key tool for reaching and interacting with citizens as well as ensuring data transmission. The website could be further improved to provide more interactive information about the project's objectives, the ongoing data collection campaigns, and the collected data. The visualization page could be upgraded to include more types of visualizations such as pie charts, boxplots, and scatter plots. The existing charts could also include a way to visualize only a specified period of time instead of the whole period of the time series. This should give the

user more insights about the seasonal behavior the each dataset as well as the specific extreme events (e.g. for rainfall and discharge). The image-sharing feature in the chat box could be improved to include more type of files. This should be implemented carefully since currently, the server checks the content of the files to verify that they are images of the accepted types. Supporting new file formats could represent a security issue for the website, as they could include viruses. The admin must add checks to make sure that the files are safe. Nevertheless, the data-sharing part of the application could also be visualized outside of the chat box, for example in a new dashboard page, where the users could have an overview of the data they shared and how it was integrated. It could be interesting to add private rooms in the chat box that divide citizens into multiple groups according to the parameter they are monitoring. For instance, citizens who measure rainfall can be grouped into one room so they can interact and exchange experience. This feature should improve the data transmission and collection via chat box as each category of inputs can be grouped separately by the admin. User profiles could include more attributes such as an avatar/photo, age, gender, social media links, and a brief description of the interests, hobbies, and professional activities. This should provide more details about the participants, which could improve the assessment of the citizen-generated data in a more objective way. The profile features could be updated in a new settings page, where users could also have the ability to delete their account without asking the administrator.

7.2.2 The added value of training in Citizen Science

We showed that training is a crucial part of any citizen science project to guarantee a consistent data collection procedure with reliable outcomes. However, the actual value of training has not been assessed in a concrete manner yet. It's worth looking at this particular point in more detail by dividing the volunteers in two categories: (i) a trained group of citizens and (ii) a non-trained group. This should deliver two different outcomes of data collection, particularly for Discharge since it requires more IT knowledge from the citizens than in the case of rainfall and water quality where the monitoring process is relatively straightforward. Both data sets should be assessed and compared to the official measurements in order to quantify the associated errors. Although we assume that the error in both cases could be high (depending on other external factors such as wind, light reflection, obstacles etc.), it is conceivable that the error associated with

the non-trained citizens would be higher than the trained group. Consequently, a thorough analysis of variance is required in order to quantify the part of the expertise of citizen in the total variance of errors associated with each data set. Hence, by this assessment, it could be possible to identify the value of the training provided to citizens, which can be evaluated through the amount of variance contribution related to the expertise of the volunteer in both groups. Nevertheless, other factors related to the citizen could also be assessed such as the age group and the educational background, which should bring additional selection criteria in the beginning of every citizen science project depending on the monitoring tools and the final objective of the project (whether data quality matters or not).

7.2.3 More focus on water quality monitoring

In this research, we focused mainly on rainfall and discharge measurement in the context of citizen science. Although results were reliable and could be used in monitoring the SDG 6 indicators (Chapter 6), we also showed that water quality could potentially be a major part of the goal's framework. Water quality monitoring through CS has been evolving recently with the emergence of several novel low-cost technologies (Farnham et al., 2017). Water quality data availability and accessibility are major issues in Tunisia given the lack of coordination between the multiple related agencies and institutions that monitor the different water resources in the country. Therefore, citizen-based approaches could provide remarkable support to the existing databases and should foster the data accessibility. Within the Together4Water project, the main focus with respect to water quality monitoring is on nitrate, total alkalinity, and salinity, that represent crucial indicators for assessing the quality of water and also some of the points of concerns related to the Tunisian waters, particularly within the Medjerda catchment. Nevertheless, other water quality indicators should be considered such as phosphate, turbidity, total hardness, and biochemical or chemical oxygen demand. These parameters are also essential for a consistent evaluation of the quality of water resources (surface water and groundwater). The engagement of more volunteers in the project should facilitate the monitoring of all the water quality indicators consistently not only within the Medjerda catchment but also in other regions where water quality issues are numerous and affecting agriculture productivity as well as ecosystems. In addition, as stressed in the previous section (7.2.1), the participation of local stakeholders should increase the impact of water

quality monitoring and could bring more support from the governmental agencies and institutions, which would be a major support for the Together4Water project, particularly from a financial aspect.

7.2.4 Measurement of SDG 6 indicators

The monitoring of the SDGs, particularly SDG 6, is an international task supported by the United Nations in order to achieve the goal by 2030 at global, national, and regional scales (UN-Water, 2018). Given the importance of SDG 6 for human and ecosystems' health, national efforts are put together in several countries in order to monitor the implementation of the goals indicators and make sure they are well met through the adaptation of the water resources management strategies or the increase of citizens' awareness with respect to the major water-related issues. In Tunisia, to our knowledge, there are limited initiatives that strive towards the monitoring of SDG 6 indicators in the country. Our work is considered as one of the very few first steps for implementing an operational system for SDG 6 monitoring. The disaggregation of the water stress (WS) indicator (6.4.2) at different spatial and temporal scales provided a deep understanding about the level of WS in the Medjerda catchment. It also allowed alerting about current and the future situation of the available water resources compared to the multiple demands. Future work should consider monitoring all SDG 6 indicators in order to provide a consistent overview on the national and regional status of the available water resources. We believe that, based on our work, the use of governmental data only is not sufficient for a reliable calculation of the goals' indicators since the official databases suffer from data gaps and low quality. Therefore, it is recommended to include all possible sources of data in the monitoring process, particularly remote sensing products and citizen science. These sources of information are able to provide timely and cost-effective reliable data that could help improve the calculation of the indicators and reduce the associated uncertainties.

References

- Albert, A., Balázs, B., Butkevičienė, E., Mayer, K., & Perelló, J. (2021). Citizen social science: New and established approaches to participation in social research. In K. Vohland et al. (Eds.), *The science of citizen science* (pp. 119–138). Cham: Springer International Publishing.
- ANPE. (2018). *Rapport National sur l'état de l'environnement* (Tech. Rep.). Tunis: Agence Nationale de Protection de l'Environnement (ANPE).
- Aqualim. (2020). *Réseau de mesure limnimétrique de la Direction des Cours d'Eau non navigables*. Retrieved 2020-10-23, from <http://aqualim.environnement.wallonie.be/login.do?time=1604763138644>
- Bhaduri, A., Bogardi, J., Siddiqi, A., Voigt, H., Vörösmarty, C., Pahl-Wostl, C., ... Osuna, V. R. (2016). Achieving sustainable development goals from a water perspective. *Frontiers in Environmental Science*, 4, 64. doi: 10.3389/fenvs.2016.00064
- Bogaert, P., & Fasbender, D. (2007). Bayesian data fusion in a spatial prediction context: a general formulation. *Stochastic Environmental Research and Risk Assessment*, 21(6), 695–709.
- Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., ... Arino, O. (2013). Consistent global land cover maps for climate modelling communities: current achievements of the ESA' and cover CCI. In *Esa living planet symposium 2013* (pp. 9–13). Edinburgh, UK.
- Bouraoui, F., Benabdallah, S., Jrad, A., & Bidoglio, G. (2005, jan). Application of the SWAT model on the Medjerda river basin (Tunisia). *Physics and Chemistry of the Earth, Parts A/B/C*, 30(8-10), 497–507. doi: 10.1016/J.PCE.2005.07.004
- Brakenridge, G. R., Nghiem, S. V., Anderson, E., & Mic, R. (2007). Orbital microwave measurement of river discharge and

- ice status. *Water Resources Research*, 43(4), W04405. doi: <https://doi.org/10.1029/2006WR005238>
- Buytaert, W., Dewulf, A., Bièvre, B. D., Clark, J., & Hannah, D. M. (2016). Citizen science for water resources management: Toward polycentric monitoring and governance? *Journal of Water Resources Planning and Management*, 142(4), 01816002. doi: [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000641](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000641)
- Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T. C., Bastiaensen, J., ... Zhumanova, M. (2014). Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. *Frontiers in Earth Science*, 2, 26. doi: <https://doi.org/10.3389/feart.2014.00026>
- Canfield, D. E., Brown, C. D., Bachmann, R. W., & Hoyer, M. V. (2002, mar). Volunteer Lake Monitoring: Testing the Reliability of Data Collected by the Florida LAKEWATCH Program. *Lake and Reservoir Management*, 18(1), 1–9. doi: 10.1080/07438140209353924
- Carrel, M., Detret, M., Peña-Haro, S., & Lüthi, B. (2019). Evaluation of the DischargeApp: a smartphone application for discharge measurements. In *Hydrosensoft, international symposium and exhibition on hydro-environment sensors and software*. (p. 8). Madrid, Spain.
- Cornell. (1972). First order analysis of model and parameter uncertainty. In Cornell C.A. (Ed.), *Int. symp. on uncertainties in hydrologic and water resource systems* (pp. 1245–1274). Tucson, AZ.
- Davids, J., Rutten, M., Pandey, A., Devkota, N., van Oyen, W., Prajapati, R., & van de Giesen, N. (2019). Citizen science flow – an assessment of simple streamflow measurement methods. *Hydrology and Earth System Sciences*, 23(2), 1045–1065. doi: <https://doi.org/10.5194/hess-23-1045-2019>
- Degefu, D. M., Weijun, H., Zaiyi, L., Liang, Y., Zhengwei, H., & Min, A. (2018). Mapping Monthly Water Scarcity in Global Transboundary Basins at Country-Basin Mesh Based Spatial Resolution. *Scientific Reports*, 8(1), 2144. doi: 10.1038/s41598-018-20032-w
- Deineko, N. (2020). *Creation of a website to dynamically visualize and collect environmental data* (Doctoral dissertation, UCL - Ecole polytechnique de Louvain). Retrieved from <http://hdl.handle.net/2078.1/thesis:25268>
- DGRE. (2016). *Rapport hydrologique 2016* (Tech. Rep.). Tunis: Ministry of Agriculture, Water Resources and Fisheries.

- Dhahbi, G. (2018). *Suivi des périmètres irrigués en Tunisie* (Tech. Rep.). Tunis: CRDA, Ministère de l'Agriculture.
- Dickens, C., Rebelo, L.-M., & Nhamo, L. (2017). Guidelines and indicators for target 6.6 of the sdgs: "change in the extent of water-related ecosystems over time". *CGIAR Research Program on Water, Land and Ecosystems (WLE)*, 56. Retrieved from <https://hdl.handle.net/10568/96083>
- Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., ... Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, 10(6), 291-297. doi: <https://doi.org/10.1890/110236>
- Dramais, G., Coz, J. L., Camenen, B., & Hauet, A. (2011). Advantages of a mobile lispiv method for measuring flood discharges and improving stage-discharge curves. *Journal of Hydro-environment Research*, 5(4), 301 - 312. (IMAGE VELOCIMETRY) doi: <https://doi.org/10.1016/j.jher.2010.12.005>
- European Commission, D.-G. f. t. E. (2015). *Ecological flows in the implementation of the Water Framework Directive : guidance document n31*. European Commission. Retrieved from <http://ec.europa.eu> doi: 10.2779/775712
- FAO. (2017). *Step-by-Step Monitoring Methodology for Indicator 6.4.2* (Tech. Rep.). Author. Retrieved from <http://www.fao.org/sustainable-development-goals/indicators/642/en/>
- FAO. (2018). *SDG Indicator 6.4.2 - Water stress*. Retrieved 2018-08-21, from <http://www.fao.org/sustainable-development-goals/indicators/642/en/>
- Farnham, D. J., Gibson, R. A., Hsueh, D. Y., McGillis, W. R., Culligan, P. J., Zain, N., & Buchanan, R. (2017). Citizen science-based water quality monitoring: Constructing a large database to characterize the impacts of combined sewer overflow in new york city. *Science of The Total Environment*, 580, 168 - 177. doi: <https://doi.org/10.1016/j.scitotenv.2016.11.116>
- Fasbender, D., Brasseur, O., & Bogaert, P. (2009). Bayesian data fusion for space-time prediction of air pollutants: The case of no2 in belgium. *Atmospheric Environment*, 43(30), 4632 - 4645. doi: <https://doi.org/10.1016/j.atmosenv.2009.05.036>
- Fasel, M., Bréthaut, C., Rouholahnejad, E., Lacayo-Emery, M. A., & Lehmann, A. (2016, dec). Blue water scarcity in the Black Sea

- catchment: Identifying key actors in the water-ecosystem-energy-food nexus. *Environmental Science and Policy*, 66, 140–150. doi: 10.1016/j.envsci.2016.09.004
- Fehri, R., Bogaert, P., Khlifi, S., & Vanclooster, M. (2020). Data fusion of citizen-generated smartphone discharge measurements in tunisia. *Journal of Hydrology*, 590, 125518. doi: <https://doi.org/10.1016/j.jhydrol.2020.125518>
- Fehri, R., Khlifi, S., & Vanclooster, M. (2019). Disaggregating sdg-6 water stress indicator at different spatial and temporal scales in tunisia. *Science of The Total Environment*, 694, 133766. doi: <https://doi.org/10.1016/j.scitotenv.2019.133766>
- Fehri, R., Khlifi, S., & Vanclooster, M. (2020). Testing a citizen science water monitoring approach in tunisia. *Environmental Science and Policy*, 104, 67–72. doi: <https://doi.org/10.1016/j.envsci.2019.11.009>
- Fox, J. (2008). *Applied regression analysis and generalized linear models*, 2nd ed. Sage Publications.
- Fraisl, D., Campbell, J., See, L., Wehn, U., Wardlaw, J., Gold, M., ... Fritz, S. (2020). Mapping citizen science contributions to the un sustainable development goals. *Sustainability Science*. doi: 10.1007/s11625-020-00833-7
- Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., ... Obersteiner, M. (2012). Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software*, 31, 110–123. doi: <http://dx.doi.org/10.1016/j.envsoft.2011.11.015>
- Fritz, S., See, L., Carlson, T., Haklay, M. M., Oliver, J. L., Fraisl, D., ... West, S. (2019). Citizen science and the united nations sustainable development goals. *Nature Sustainability*, 2(11), 1063. doi: <https://doi.org/10.1038/s41893-019-0426-8>
- Fujita, I., Kunita, Y., & Tsubaki, R. (2013). Image analysis and reconstruction of the 2008 toga river flash flood in an urbanised area. *Australasian Journal of Water Resources*, 16(2), 151–162. doi: 10.7158/13241583.2013.11465411
- Fujita, I., Muste, M., & Kruger, A. (1998). Large-scale particle image velocimetry for flow analysis in hydraulic engineering applications. *Journal of Hydraulic Research*, 36(3), 397–414. doi: 10.1080/00221689809498626
- Gain, A. K., & Wada, Y. (2014, mar). Assessment of Future Water Scarcity at Different Spatial and Temporal Scales of the Brahmaputra River Basin. *Water Resources Management*, 28(4), 999–1012. doi:

- doi.org/10.1007/s11269-014-0530-5
- Galecki, A., & Burzykowski, T. (2013). *Linear mixed-effects models using R*. Springer.
- Ground Truth, . (2020). *Ground Truth 2.0 – A bottom-up, citizen-focused approach to environmental monitoring*. Retrieved 2020-12-12, from <https://cordis.europa.eu/article/id/421807-ground-truth-2-0-a-bottom-up-citizen-focused-approach-to-environmental-monitoring>
- Hall, R., Ranganathan, S., & C., R. G. (2017, 2). A general micro-level modeling approach to analyzing interconnected sdgs: Achieving sdg 6 and more through multiple-use water services (mus). *Sustainability*, 9, 314. Retrieved from <http://www.mdpi.com/2071-1050/9/2/314> doi: 10.3390/su9020314
- Hannah, D. M., Demuth, S., van Lanen, H. A. J., Looser, U., Prudhomme, C., Rees, G., ... Tallaksen, L. M. (2011). Large-scale river flow archives: importance, current status and future needs. *Hydrological Processes*, 25(7), 1191-1200. doi: <https://doi.org/10.1002/hyp.7794>
- Heiss, R., & Matthes, J. (2017). Citizen science in the social sciences: A call for more evidence. *GAIA - Ecological Perspectives for Science and Society*, 26(1), 22-26. doi: doi:10.14512/gaia.26.1.7
- Hirji Richard, R. D. (2009). *Environmental Flows in Water Resources Policies, Plans, and Projects*. The World Bank. doi: 10.1596/978-0-8213-7940-0
- Iachine, I., Petersen, H. C., & Kyvik, K. O. (2010). Robust tests for the equality of variances for clustered data. *Journal of Statistical Computation and Simulation*, 80(4), 365-377. doi: <https://doi.org/10.1080/00949650802641841>
- iFADO, . (2020). *iFADO Project*. Retrieved 2020-12-20, from <http://www.ifado.eu/>
- INS. (2014). *Institut National de la Statistique, Tunisie*. Retrieved 2019-12-06, from <http://dataportal.ins.tn/>
- JMP. (2020). *The WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP)*. Retrieved 2020-10-21, from <https://washdata.org/how-we-work/about-jmp>
- Kantoush, S. A., Schleiss, A. J., Sumi, T., & Murasaki, M. (2011). Lspiv implementation for environmental flow in various laboratory and field cases. *Journal of Hydro-environment Research*, 5(4), 263 - 276. Retrieved from <http://www.sciencedirect.com/science/>

- article/pii/S1570644311000736 (IMAGE VELOCIMETRY) doi: <https://doi.org/10.1016/j.jher.2011.07.002>
- Kim, Y., Muste, M., Hauet, A., Krajewski, W. F., Kruger, A., & Bradley, A. (2008, sep). Stream discharge using mobile large-scale particle image velocimetry: A proof of concept. *Water Resources Research*, 44(9), n/a–n/a. doi: 10.1029/2006WR005441
- Kwonkyu, Y., Byungman, Y., & Seokmin, L. (2015). Surface image velocity measurement system for wide rivers using smartphones. In *The 36th iahr world congress* (pp. 1–3). The Hague.
- Le Boursicaud, R., Pénard, L., Hauet, A., Thollet, F., & Le Coz, J. (2016). Gauging extreme floods on youtube: application of lspiv to home movies for the post-event determination of stream discharges. *Hydrological Processes*, 30(1), 90-105. doi: <https://doi.org/10.1002/hyp.10532>
- Le Coz, J., Patalano, A., Collins, D., Guillen, N. F., Garcia, C. M., Smart, G. M., ... Braud, I. (2016). Crowdsourced data for flood hydrology: feedback from recent citizen science projects in argentina, france and new zealand. *Journal of Hydrology*, 16(2), 766-777. doi: 10.1016/j.jhydro1.2016.07.036
- Leitão, J. P., Peña-Haro, S., Lüthi, B., Scheidegger, A., & de Vitry, M. M. (2018). Urban overland runoff velocity measurement with consumer-grade surveillance cameras and surface structure image velocimetry. *Journal of Hydrology*, 565, 791 - 804. doi: <https://doi.org/10.1016/j.jhydro1.2018.09.001>
- Levene, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, 52(3-4), 278-292.
- Loague, K. M., Yost, R. S., Green, R. E., & Liang, T. C. (1989, apr). Uncertainty in a pesticide leaching assessment for Hawaii. *Journal of Contaminant Hydrology*, 4(2), 139–161. doi: 10.1016/0169-7722(89)90018-1
- Lüthi, B., Philippe, T., & Peña-Haro, S. (2014). Mobile device app for small open-channel flow measurement. In *Smart and mobile devices used for environmental applications* (pp. 283–287). San Diego, California, USA: International Environmental Modelling and Software Society (iEMSs). Retrieved from <http://www.iemss.org/society/index.php/iemss-2014-proceedings>
- Lüthi, B., Philippe, T., & Peña-Haro, S. (2018, May). *Method and system for determining the velocity of a moving fluid surface* (Tech. Rep.). Zurich:

- Google Patents.
- Martínez-Córdoba, P.-J., Raimo, N., Vitolla, F., & Benito, B. (2020, 4). Achieving sustainable development goals. efficiency in the spanish clean water and sanitation sector. *Sustainability*, 12, 3015. Retrieved from <https://www.mdpi.com/2071-1050/12/7/3015> doi: 10.3390/su12073015
- Mattern, S., Raouafi, W., Bogaert, P., Fasbender, D., & Vanclooster, M. (2012). Bayesian Data Fusion (BDF) of monitoring data with a statistical groundwater contamination model to map groundwater quality at the regional scale. *Journal of Water Resource and Protection*, 4, 929–943. doi: 10.4236/jwarp.2012
- Mekonnen, M. M., & Hoekstra, A. Y. (2016, feb). Four billion people facing severe water scarcity. *Science Advances*, 2(2). doi: 10.1126/sciadv.1500323
- Miller, K. S. (1981). On the inverse of the sum of matrices. *Mathematics-Magazine*, 54(2), 67–72.
- Ministry of Agriculture. (2005). *Irrigation, rapport technique* (Tech. Rep.). Ministry of Agriculture. Retrieved from <http://www.agriculture.tn>
- Monteith, J. L. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology*, 19, 205–234.
- Moreno, L., Pozo, M., Vancraeynest, K., Bain, R., Palacios, J. C., & Jácome, F. (2020). Integrating water-quality analysis in national household surveys: water and sanitation sector learnings of ecuador. *npj Clean Water*, 3, 23. Retrieved from <https://doi.org/10.1038/s41545-020-0070-x> doi: 10.1038/s41545-020-0070-x
- Mu, Q., Zhao, M., & W. Running, S. (2011, aug). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115(8), 1781–1800. doi: 10.1016/J.RSE.2011.02.019
- Nezu, I., & Sanjou, M. (2011). Piv and ptv measurements in hydrosciences with focus on turbulent open-channel flows. *Journal of Hydro-environment Research*, 5(4), 215 - 230. (IMAGE VELOCIMETRY) doi: <https://doi.org/10.1016/j.jher.2011.05.004>
- Nhemachena, C., Matchaya, G., Nhemaehena, C., Karuaihe, S., Muchara, B., & Nhlengethwa, S. (2018, 3). Measuring baseline agriculture-related sustainable development goals index for southern africa. *Sustainability*, 10, 849. Retrieved from <http://www.mdpi.com/2071-1050/10/3/849> doi: 10.3390/su10030849
- NRC. (2007). Strengthening science-based decision making in develop-

- ing countries. In L. Holliday (Ed.), *Agricultural water management: Proceedings of a workshop in tunisia*. Washington, DC: The National Academies Press. doi: doi.org/10.17226/11880.
- Pastor, A. V., Ludwig, F., Biemans, H., Hoff, H., & Kabat, P. (2014). Accounting for environmental flow requirements in global water assessments. *Hydrology and Earth System Sciences*, 18(12), 5041–5059. doi: [10.5194/hess-18-5041-2014](https://doi.org/10.5194/hess-18-5041-2014)
- Paul, J. D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J. A., Bhusal, J., Cieslik, K., ... Supper, R. (2018). Citizen science for hydrological risk reduction and resilience building. *WIREs Water*, 5(1), e1262. doi: <https://doi.org/10.1002/wat2.1262>
- Photrack, H. (2018a). *Discharge app*. Retrieved 2019-11-22, from <https://discharge.ch>
- Photrack, H. (2018b). *Global iMoMo Initiative*. Retrieved 2020-06-26, from <https://www.imomohub.com/>
- Pocock, M. J. O., Roy, H. E., August, T., Kuria, A., Barasa, F., Bett, J., ... Trevelyan, R. (2019). Developing the global potential of citizen science: Assessing opportunities that benefit people, society and the environment in east africa. *Journal of Applied Ecology*, 56(2), 274–281. doi: [10.1111/1365-2664.13279](https://doi.org/10.1111/1365-2664.13279)
- Roy, H., Pocock, M., Preston, C., Roy, D., Savage, J., Tweddle, J., & Robinson, L. (2012). *Understanding citizen science and environmental monitoring: final report on behalf of uk environmental observation framework* (Tech. Rep.). Wallingford: Natural History Museum. Retrieved from <http://nora.nerc.ac.uk/id/eprint/20679/> (Freely available online - click on Official URL link for full text)
- Sarah, W., & Rachel, P. (2017). *How could citizen science support the Sustainable Development Goals?* (Tech. Rep.). Stockholm Environment Institute (SEI). Retrieved from <https://mediamanager.sei.org/documents/Publications/SEI-2017-PB-citizen-science-sdgs.pdf>
- See, L. (2019). A review of citizen science and crowdsourcing in applications of pluvial flooding. *Frontiers in Earth Science*, 7, 44. doi: <https://doi.org/10.3389/feart.2019.00044>
- See, L., Fritz, S., Perger, C., Schill, C., McCallum, I., Schepaschenko, D., ... Obersteiner, M. (2015). Harnessing the power of volunteers, the internet and Google Earth to collect and validate global spatial information using Geo-Wiki. *Technological Forecasting and Social Change*, 98. doi: [http://dx.doi.org/10.1016/j.techfore.2015.03.002](https://doi.org/10.1016/j.techfore.2015.03.002)

- See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., ... Rutzinger, M. (2016). Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. *ISPRS International Journal of Geo-Information*, 5(5), 55. Retrieved from <http://www.mdpi.com/2220-9964/5/5/55>
- Sellami, H. (2014). *Uncertainty assessment of hydrological impacts of climatic change in small mediterranean catchments* (Doctoral dissertation, UCL UCL - SST/ELI/ELIE - Environmental Sciences Ingénierie biologique, agronomique et environnementale). Retrieved from <http://hdl.handle.net/2078.1/146526>
- Sellami, H., Vanclooster, M., Benabdallah, S., & La Jeunesse, I. (2013). Assessment of the swat model prediction uncertainty using the glue approach a case study of the chiba catchment (tunisia). In *2013 5th international conference on modeling, simulation and applied optimization (icmsao)* (p. 1-6). doi: <https://doi.org/10.1109/ICMSAO.2013.6552605>
- Shapiro, S. S., & Wilk, M. B. (1965, 12). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. doi: <https://doi.org/10.1093/biomet/52.3-4.591>
- Smakhtin, V., Revenga, C., & Döll, P. (2004, sep). A Pilot Global Assessment of Environmental Water Requirements and Scarcity. *Water International*, 29(3), 307-317. doi: 10.1080/02508060408691785
- Sood, A., Smakhtin, V., Eriyagama, N., Villholth, K. G., Liyanage, N., Wada, Y., ... Dickens, C. (2017). *Global environmental flow information for the sustainable development goals* (Tech. Rep.). Colombo, Sri Lanka: International Water Management Institute. Retrieved from <http://www.iwmi.cgiar.org/publications/iwmi-research-reports/iwmi-research-report-168/> doi: 10.5337/2017.201
- Tarpanelli, A., Brocca, L., Lacava, T., Faruolo, M., Melone, F., Moramarco, T., ... Tramutoli, V. (2011). River discharge estimation through MODIS data. In C. M. U. Neale & A. Maltese (Eds.), *Remote sensing for agriculture, ecosystems, and hydrology xiii* (Vol. 8174, pp. 77 - 84). SPIE. doi: <https://10.1117/12.898201>
- Tauginienė, L., Butkevičienė, E., Vohland, K., Heinisch, B., Daskolia, M., Suškevičs, M., ... Prūse, B. (2020). Citizen science in the social sciences and humanities: the power of interdisciplinarity. *Palgrave Communications*, 6(1), 89. Retrieved from <https://doi.org/10.1057/s41599-020-0471-y> doi: 10.1057/s41599-020-0471-y
- Tauro, F., Selker, J., van de Giesen, N., Abrate, T., Uijlenhoet, R., Porfiri,

- M., ... Grimaldi, S. (2018). Measurements and observations in the xxi century (moxxi): innovation and multi-disciplinarity to sense the hydrological cycle. *Hydrological Sciences Journal*, 63(2), 169-196. doi: <https://doi.org/10.1080/02626667.2017.1420191>
- Tharme, R. E. (2003). A global perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers. *River Research and Applications*, 19(5-6), 397-441. doi: 10.1002/rra.736
- UN-Environment. (2018). *Progress on Integrated Water Resources Management. Global Baseline for SDG 6 Indicator 6.5.1: Degree of IWRM Implementation* (Tech. Rep.). UN Environment, UN Environment-DHI Centre, Cap-Net, the Global Water Partnership, and UN-Water. Retrieved from <https://www.unepdhi.org/publications-iwrm/>
- UNESCO. (2016). *Integrated Monitoring for SDG 6 (GEMI) Work in Progress Workshop* (Tech. Rep.). Delft: UNESCO-IHE.
- United-Nations. (2015). *Transforming our world: the 2030 agenda for sustainable development*. United Nations.
- UNOSD. (2019). *Workshop on Using SDG 6 Policy Support System (SDG-PSS) to facilitate countries in Africa and Middle East for water-related sustainable development*. United Nations Office for Sustainable Development (UNOSD). Retrieved 2020-10-21, from <https://unosd.un.org/events/2019-workshop-using-sdg-6-policy-support-system-sdg-pss-facilitate-countries-africa-and>
- UNStats. (2020). *SDG Indicators Metadata repository*. Retrieved 2020-10-21, from <https://unstats.un.org/sdgs/metadata/>
- UN-Water. (2015). *UN-Water: Indicators and Monitoring*. Retrieved 2016-08-11, from <http://www.unwater.org/sdgs/indicators-and-monitoring/en/>
- UN-Water. (2017a). *A/RES/71/313 Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. Seventy-first session of the General Assembly on 6 July 2015* (Tech. Rep.). United Nations. Retrieved from http://ggim.un.org/documents/A_RES_71_313.pdf
- UN-Water. (2017b). *Integrated Monitoring Guide for Sustainable Development Goal 6 – Good practices for country monitoring systems* (Tech. Rep.). United Nations. Retrieved from https://www.unwater.org/app/uploads/2017/09/G1_Good-practices-for-country-monitoring-systems_Version-2017-07-12a.pdf
- UN-Water. (2018). *Sustainable Development Goal 6: Synthesis Report 2018 on Water and Sanitation* (Tech. Rep.). United Nations. Retrieved from

- <https://www.unwater.org>
- Van Dijk, A. I. J. M., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groeve, T., & Schellekens, J. (2016). River gauging at global scale using optical and passive microwave remote sensing. *Water Resources Research*, 52(8), 6404-6418. doi: <https://doi.org/10.1002/2015WR018545>
- Vanham, D., Hoekstra, A., Wada, Y., Bouraoui, F., de Roo, A., Mekonnen, M., ... Bidoglio, G. (2018, feb). Physical water scarcity metrics for monitoring progress towards SDG target 6.4: An evaluation of indicator 6.4.2 "Level of water stress". *Science of The Total Environment*, 613-614, 218-232. doi: 10.1016/J.SCITOTENV.2017.09.056
- Walker, D., Forsythe, N., Parkin, G., & Gowing, J. (2016). Filling the observational void: Scientific value and quantitative validation of hydrometeorological data from a community-based monitoring programme. *Journal of Hydrology*, 538, 713 - 725. doi: <https://doi.org/10.1016/j.jhydrol.2016.04.062>
- Weijis, S. V., & Chen, Z. (2014). Measuring surface flow velocity with smartphones: potential for citizen observatories. In *école polytechnique fédérale de lausanne epfl* (pp. n/a-n/a). Vienna: EGU general assembly.
- WHO, UNISEF, & JMP. (2016). *WHO / UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation*. Retrieved 2016-08-09, from <http://www.wssinfo.org/>
- Wolff, D. B., Marks, D. A., Amitai, E., Silberstein, D. S., Fisher, B. L., Tokay, A., ... Pippitt, J. L. (2005, apr). Ground Validation for the Tropical Rainfall Measuring Mission (TRMM). *Journal of Atmospheric and Oceanic Technology*, 22(4), 365-380. doi: 10.1175/JTECH1700.1
- Woodward, G. (2016). *Rainlog.org*. Retrieved 2018-08-16, from <http://www.rainlog.org>
- Yang, P., Xia, J., Zhan, C., Mo, X., Chen, X., Hu, S., & Chen, J. (2018, feb). Estimation of water consumption for ecosystems based on Vegetation Interfaces Processes Model: A case study of the Aksu River Basin, Northwest China. *Science of The Total Environment*, 613-614, 186-195. doi: 10.1016/J.SCITOTENV.2017.09.045
- Zabell, S. L. (2008). On student's 1908 article "the probable error of a mean". *Journal of the American Statistical Association*, 103(481), 1-7. doi: <https://doi.org/10.1198/016214508000000030>

Appendices

Appendix A. Additional equations for the BLUP framework

If we assume that Σ_ϵ is given by eq. (5.6), it is possible to provide an explicit expression for Σ_ϵ^{-1} based on Miller (1981). In general, if $\mathbf{A}$ is a positive definite matrix and $\mathbf{B}$ is a matrix of rank 1, then

$$(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \frac{1}{1+g} \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1} \quad (\text{A.1})$$

where $g = \text{trace}(\mathbf{A}^{-1} \mathbf{B})$. By identifying $\mathbf{A} = \sigma_u^2 \mathbf{I}$ and $\mathbf{B} = \sigma_m^2 \mathbf{1} \mathbf{1}'$ so that $\Sigma_\epsilon^{-1} = (\mathbf{A} + \mathbf{B})^{-1}$ and knowing that $\mathbf{A}^{-1} = (1/\sigma_u^2) \mathbf{I}$, this gives after elementary manipulations

$$g = \frac{\sigma_m^2}{\sigma_u^2} n \quad ; \quad \frac{1}{1+g} = \frac{\sigma_u^2}{\sigma_u^2 + n\sigma_m^2} \quad ; \quad \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1} = \frac{\sigma_m^2}{\sigma_u^4} \mathbf{1} \mathbf{1}' \quad (\text{A.2})$$

and so we obtain

$$\Sigma_\epsilon^{-1} = \frac{1}{\sigma_u^2} \left(\mathbf{I} - \frac{\sigma_m^2}{\sigma_u^2 + n\sigma_m^2} \mathbf{1} \mathbf{1}' \right) \quad (\text{A.3})$$

which does not require the computation of any inverse matrix. Plugging this result in eq. (5.4) now gives

$$\lambda \propto \Sigma_\epsilon^{-1} \mathbf{1} = \frac{1}{\sigma_u^2} \left(1 - \frac{n\sigma_m^2}{\sigma_u^2 + n\sigma_m^2} \right) \mathbf{1} \quad (\text{A.4})$$

where all factors in front of $\mathbf{1}$ are scalars, thus showing that all weights

are the same and must be equal to $1/n$, as these weights sum up to one. Plugging this result in eq. (5.5) finally gives

$$\text{Var}[Z - \lambda' \mathbf{Y}] = \frac{1}{\frac{n}{\sigma_u^2} \left(1 - \frac{n\sigma_m^2}{\sigma_u^2 + n\sigma_m^2} \right)} \quad (\text{A.5})$$

Appendix B. Test of the Discharge smartphone app in Belgium

In this appendix we present details about the test of the "Discharge app" in Belgium. This test was carried out before the start of the CS field campaign in Tunisia to evaluate the performance of the app and check its different features and settings. An expert user of the app performed measurements of a series of hydrological events at three sites in the Walloon region: Noirhat, Suzryl, and Rosières (Fig.B. 1). The collected daily data were compared with data from reference stations at every location provided by the Aqualim project (Aqualim, 2020).

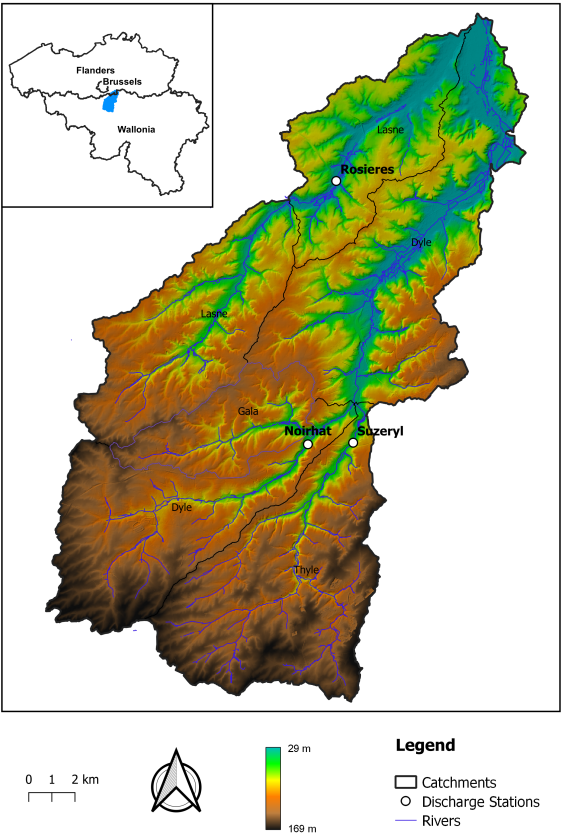


Figure B. 1: The study area.

B. 1 Steps for site installation

Before conducting the test of the app in the 3 locations, we performed sites' installation following the guidelines of the Discharge app development team. Figure B. 2 illustrates the different on-field parts of the installation.

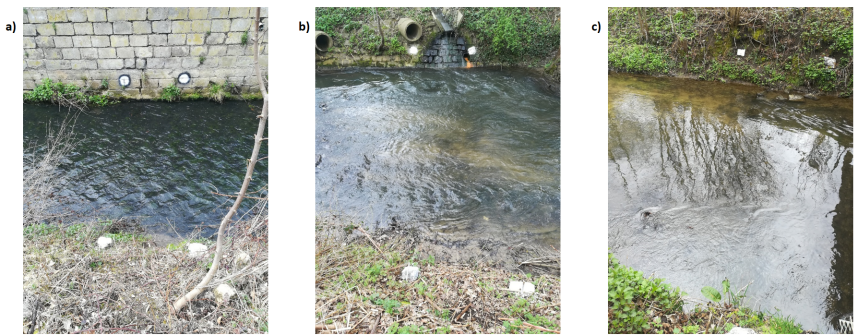


Figure B. 2: The monitoring sites (a) Noirhat, (b) Rosières, and (c) Suzeryl. The white spots in each site are the markers fixed on the river banks.

The first step is to gather precise measurements of each site's cross-section through conducting a geometrical survey (bathymetry) using the Disto S910 device. This part is essential to ensure reliable representations of each stream's geometry, which is a major factor for a reliable estimation of discharge. The main elements of the cross-section geometrical survey are, (i) to locate 4 markers fixed on both river banks (Fig. B. 2); (ii) to locate the shore line, which represents the extent of water in the river from the standing spot of the user; and (iii) to measure the river bed profile with as many points as possible.

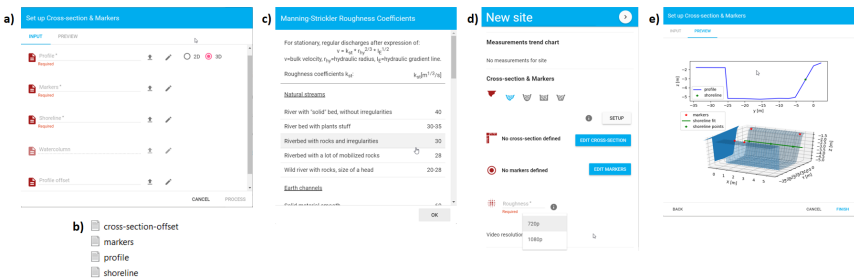


Figure B. 3: Steps for site installation in the Discharge app.

After the extraction of the river's cross-section, the second step is to configure the sites in the application's online platform discharge.ch. Figure B. 3 shows the different parts of a typical site set up. The bathymetry outputs (in text format) are loaded in the "Set up Cross-section & Markers" window (Fig. B. 3a & B. 3b). Then, we chose an appropriate Manning strikler coefficient based on the app's guide and the knowledge of the site (Fig. B. 3c). On the same window, we define the video resolution (Fig. B. 3d). It is recommended to set it at the highest resolution possible (1080p) to guarantee more reliable estimation of the discharge (Carrel et al., 2019). The processing of the different input information generates a representation of the river's cross-section as well as the markers and the shore line (Fig. B. 3e). After approving the generated cross-section, the site becomes ready to be used on-site to estimate the discharge.

B. 2 Outcomes of the discharge measurements

Daily discharge was measured by an expert user of the app for 3 months at all locations: Noirhat, Suzeryl, and Rosières.

Results of the expert user measurements compared with the Aqualim data are shown in Figure B. 4.

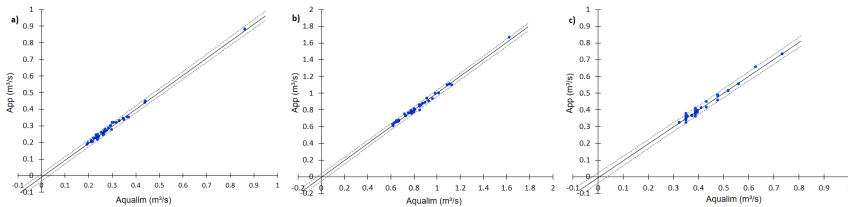


Figure B. 4: Discharge app estimates versus Aqualim data at (a) Noirhat, (b) Rosières, and (c) Suzeryl, along with the estimated regression line (red plain line), and the corresponding 95% prediction interval (dashed lines).

The Discharge app estimates are in good agreement with the reference data at all sites. The corresponding R^2 values are as follows: 0.993 for Noirhat, 0.986 for Suzeryl, and 0.994 for Rosières. Therefore, it's concluded that the discharge estimated by the smartphone application is in a good quality, which confirms the reliability of the sites' installation performed described in the previous section.