UNIVERSITAT POLITÈCNICA DE CATALUNYA

DOCTORAL PROGRAMME IN CIVIL ENGINEERING

Doctoral Thesis

Decision analysis under uncertainty for sustainable development

by

FATINE EZBAKHE

Advisor: Dr. AGUSTÍ PÉREZ FOGUET

Barcelona, September 2019

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

ESaGD
EngineeringSciences&GlobalDev

Abstract

Policy-making for sustainable development becomes more efficient when it is reliably backed by evidence-based decision analysis. Concretely, this is crucial in the planning of public services delivery. By translating *raw* data into information, decision analysis illuminates our judgment, and ultimately the policies we adopt. In the context of public services provision, decision analysis can support the prioritization of policy options and the monitoring of progress. However, most models are deterministic – that is, they do not consider the uncertainty in their evidence. These *incomplete* models, through their impact in policy decisions, can ultimately lead to an inefficient use of resources. The main barriers to a wider incorporation of uncertainty are: (i) the complexity of the approaches currently available, and (ii) the need to develop methods tailored to the specific decision problems faced in public services delivery.

To overcome these limitations, this thesis intends to facilitate the incorporation of uncertainty in the evidence into decision analysis for sustainable development. We propose two methods. First, a non-compensatory multi-criteria prioritization under uncertainty model. Given multiple criteria and uncertain evidence, the model identifies the best policy option to improve service provision for sustainable development. The non-compensatory nature of our model makes it an attractive alternative to the widely used composite index approach. Second, a compositional trend analysis under uncertainty model to monitor service coverage. By considering the non-negativity and constant-sum constraints of the data, our model provides better estimates for measuring progress than standard statistical approaches.

These two methods are validated in real case studies in the energy, water and health sectors. We apply our prioritization model to the context of strategic renewable energy planning (Chapter 1) and the targeting of water, sanitation and hygiene services (Chapter 2). Furthermore, we use our trend analysis model to the global monitoring of water and sanitation (Chapter 3) and child mortality (Chapter 4).

Our results emphasize the importance of considering and incorporating uncertainty in the evidence into decision analysis, particularly into prioritization and monitoring processes, both central to sustainable development practice.

List of publications arising from this thesis

Papers published in journals indexed in the Web of Science:

- **Ezbakhe, F.** and Pérez-Foguet, A., 2019. Levels and trends in child mortality: a compositional approach. *Demographic Research* (Under Review)
- **Ezbakhe, F.** and Pérez-Foguet, A., 2019. Decision analysis for sustainable development: the case of renewable energy planning under uncertainty. *European Journal of Operational Research* (Under Review)
- **Ezbakhe, F.** and Pérez-Foguet, A., 2019. Estimating access to drinking water and sanitation: the need to account for uncertainty in trend analyses. *Science of the Total Environment*, 696. doi: 10.1016/j.scitotenv.2019.133830
- **Ezbakhe, F.** and Pérez-Foguet, A., 2018. Multi-Criteria decision analysis under uncertainty: two approaches to incorporating data uncertainty into water, sanitation and hygiene planning. *Water Resources Management*, 32(15), 5169—5182. doi: 10.1007/s11269-018-2152-9

Papers published in other journals:

- **Ezbakhe, F.** and Pérez-Foguet, A., 2018. Embracing data uncertainty in water decision-making: an application to evaluate water supply and sewerage in Spain. *Water Supply*, 19(3), 778—788. doi: 10.2166/ws.2018.122
- **Ezbakhe, F.** and Pérez-Foguet, A., 2017. Considering data uncertainty in the water and sanitation sector: application to large number of alternatives and criteria. *European Water*, 57, 215–222.

Papers presented in national and international conferences:

- **Ezbakhe, F.** and Pérez-Foguet, A., 2019. Monitoring the access to water, sanitation and hygiene in the MENA region: survey errors and compositional nature, in: *EWRA 2019 – 11th World Congress on Water Resources and*

Environment "Managing Water Resources for a Sustainable Future". 25-29 June 2019. Madrid, Spain.

- **Ezbakhe, F.** and Pérez-Foguet, A., 2019. WASH your data off: navigating statistical uncertainty in compositional data analysis, in: *CoDaWork 2019 – The 8th International Workshop on Compositional Data Analysis*. 03-08 June 2019. Terrasa, Spain.
- **Ezbakhe, F.**, Giné-Garriga, R. and Pérez-Foguet, A., 2018. Evaluating the human rights to water and sanitation: the use of participatory diagnosis tools at a local level, in: *X Congreso Ibérico de Gestión y Planificación del Agua*. 06-08 September 2018. Coimbra, Portugal.
- **Ezbakhe, F.** and Pérez-Foguet, A., 2018. Tackling the compositional nature of water, sanitation and hygiene (WASH) data in statistical analysis, in: *29th European Conference on Operational Research*. 08-10 July 2018. Valencia, Spain.
- **Ezbakhe, F.** and Pérez-Foguet, A., 2018. Aquifer contracts for groundwater resources planning in the MENA region: a means to support stakeholder participation?, in: *3rd International Conference on Integrated Environmental Management for Sustainable Development*. 02-05 May 2018. Sousse, Tunisia.
- **Ezbakhe, F.** and Pérez-Foguet, A., 2017. Embracing data uncertainty in ranking: an application to evaluate water supply and sewerage in Spain, in: *Young Water Professionals Spain Conference 2017*. 16-18 November 2017. Bilbao, Spain.
- **Ezbakhe, F.** and Pérez-Foguet, A., 2017. Improved prioritisation tool for local decision-making in the water, sanitation and hygiene sector, in: *40th WEDC International Conference*. 24-28 July 2017. Loughborough, United Kingdom.
- **Ezbakhe, F.** and Pérez-Foguet, A., 2017. Considering data uncertainty in the water and sanitation sector: Application to large number of alternatives and criteria, in: *EWRA 2017 – 10th World Congress on Water Resources and Environment "Panta Rhei"*. 05-09 July 2017. Athens, Greece.

For my parents, Hassan and Amina.

Acknowledgements

"If I have seen further it is by standing on the shoulders of Giants."

– Isaac Newton

D	A	K	E	A	L	B	A	P	S	O	M	A
B	E	M	F	T	A	M	I	N	A	M	J	Y
N	Q	S	P	Q	M	E	D	Y	W	A	T	Y
X	R	I	C	A	R	D	R	L	L	R	E	K
R	Z	S	W	G	R	B	U	L	J	Y	B	H
J	S	Z	K	Y	D	O	N	E	Z	P	V	O
G	O	W	H	A	S	S	A	N	P	A	X	L
O	J	R	D	A	N	N	A	U	V	Z	K	O
G	O	B	D	Q	Y	D	S	C	F	L	M	U
G	U	X	C	I	H	E	C	T	O	R	I	D
L	U	E	J	J	I	H	A	N	E	Y	A	K
Y	M	A	G	U	S	T	I	F	G	U	U	A
E	U	A	W	Y	A	S	S	I	N	M	I	S

Contents

List of Figures

List of Tables

Acronyms and abbreviations

CODA	Compositional Data
DHS	Demographic and Health Survey
EBRD	European Bank for Reconstruction and Development
ELECTRE	ELimination and Choice Translating REality
FBH	Full Birth History
GAM	Generalized Additive Model
GCPSE	UNDP's Global Centre for Public Service Excellence
IEA	International Energy Agency
IEAG	UN Secretary-General's Independent Expert Advisory Group
IGME	UN Inter-agency Group for Child Mortality
ILR	Isometric Log-Ratio
IWGHS	UN Intersecretariat Working Group on Household Surveys
JMP	WHO/UNICEF's Joint Monitoring Programme
LSMS	Living Standards Measurement Study
MAUT	Multi-Attribute Utility Theory
MCDA	Multi-Criteria Decision Analysis
MICS	Multiple Indicator Cluster Survey
NMR	Neonatal Mortality Ratio
NSE	Nash-Sutcliffe Efficiency
OHCHR	Office of the High Commissioner for Human Rights
OLS	Ordinary Least Squares
RSE	Relative Standard Error

RSME	Root-Mean-Square Error
SBH	Summary Birth History
SBP	Sequential Binary Partition
SDG	Sustainable Development Goal
SDSN	Sustainable Development Solutions Network
UN	United Nations
UNCTAD	United Nations Conference on Trade and Development
UNDP	United Nations Development Programme
UNESA	United Nations Department of Economic and Social Affairs
UNGA	United Nations General Assembly
UNICEF	United Nations International Children's Emergency Fund
U5MR	Under-five Mortality Ratio
WASH	Water, Sanitation and Hygiene
WHO	World Health Organization
WHS	World Health Survey

Introduction

"We know that we do not know, but that is almost all that we know"

– Michel Callon, *Acting in an uncertain world.*

This thesis explores uncertainty in decision analysis for sustainable development, in particular for public services delivery. We ask two questions. First: is it possible to include the uncertainty in the evidence base[1] when prioritizing public policy options? Second: is it possible to incorporate this uncertainty when monitoring and analyzing trends in service coverage? To that end, we propose two methods: one based on non-compensatory multi-criteria analysis, and a second one on compositional trend analysis. Both methods intend to characterize and incorporate uncertainty in the sources of evidence policy-makers use to inform their decisions. We show that uncertainty must be acknowledged and accounted for properly in decision analysis in order to achieve true evidence-based public policy and planning processes.

This introductory chapter is divided into four sections. First, to understand the theoretical framework of our research, we begin explaining the context of the 2030 Agenda for Sustainable Development Goals (SDGs) in which the thesis is embedded. We then shed light on the demand for data to support evidence-based decision-making. Next, we present decision analysis as an approach to help decision-makers translate evidence into recommendations for practice and policy, especially for public services. Finally, we highlight the challenge of dealing with imperfect or incomplete evidence.

[1] We define evidence base as the information and data used to support decision- and policy-making, and ultimately justify the *soundness* of the decisions.

Setting the scene: the Sustainable Development Goals

“Senseless, dreamy, garbled”, that is what Easerly (2015) claimed the UN's Sustainable Development Goals stood for. The Economist (2015) baptized them as the “Stupid Development Goals”, facetiously labelling the proposed targets as the “169 Commandments”. The article also argued that goals were so “sprawling and misconceived” and “unfeasibly expensive” that they would be “worse than useless”. Others echoed this sentiment by pithily stating SDGs were “too vague” and looked more “like an encyclopedia of development than a useful tool for action” and “no targets left behind” (Humanosphere 2015). Hickel (2015) also slated the “profoundly contradictory, to the point of being self-defeating nature of SDGs”, highlighting the conflicting relationship between concerns on environmental protection and the emphasis on economic development based on the “old model of industrial growth”, and concluded that the SDGs “aim to save the world without transforming it”. Furthermore, several human rights experts lamented the SDGs' “missed opportunity” to unequivocally put human rights principles at the heart of the UN development agenda (Ramcharan 2015; Feiring and Hassler 2016; Winkler and Williams 2017; Weber 2017; McInerney-Lankford 2017; Yap and Watene 2019).

Certainly from an implementation perspective, the criticism that SDGs are difficult and expensive to realize is not far off the mark. With an estimated USD 5-7 trillion annual price tag (UNCTAD 2014), there are reasonable grounds to doubt whether sufficient funding can be mobilized to achieve the SDGs. The financing gap is particularly acute in developing countries, where current annual investment levels lag behind at USD 1.4 trillion (UNCTAD 2014). Added to this is the difficulty to measure and monitor the 232 indicators proposed by the Inter-Agency and Expert Group on SDG indicators (UNGA 2017). Indeed, monitoring progress is a costly proposition. Major investments – valued at USD 1 billion per annum – are required to develop the statistical capacity of countries and facilitate their compilation of statistics needed for SDG monitoring (SDSN 2015). Therefore, the existing concern that such complex and expensive indicator framework will potentially divert already-scarce resources from SDG implementation (Hering 2017; Guppy et al. 2019) is surely justified.

Even though criticism of the SDGs has been notoriously outspoken and unyielding, it is also important to point out that they have emerged from an inclusive process where countries, including developing ones, have reached an agreement on a comprehensive vision for the development agenda up to 2030. As Bhattacharya and Kharas (2015) remark, although some SDG targets are "clearly not achievable", which in turn might "undercut the overall credibility of the package", this is the price of democracy. In their words, "like all democratic processes, the result may be messy and leave many dissatisfied, but it reflects compromise and a desire for consensus". Furthermore, amongst the several reasons identified by Kumar et al. (2016) as why the SDGs are better than the previous Millennium Development Goals, is the fact that "SDGs have evolved after a long and extensive consultative process", meaning that they are globally cooperative and applicable to all countries and actors.

Data for SDGs

Public pundits' opinions aside, the SDGs have been agreed upon by all 193 Member States of the United Nations, and now is the time to implement strategies, allocate resources, and continuously evaluate progress towards achieving the 2030 targets. It is thus indispensable to have adequate data to understand how efforts towards the SDGs are translating into better economic, environmental and social outcomes. As the UN Secretary-General's Independent Expert Advisory Group recognized in 2015, "data are the lifeblood of decision-making and the raw material for accountability; without high quality data providing the right information on the right things at the right time, designing, monitoring and evaluating effective policies becomes almost impossible" (IEAG 2015).

Development data are derived from various sources, such as censuses, household surveys, administrative records, civil registration and vital statistics systems, and geospatial technologies (SDSN 2015). Amid all these data, household survey data represent a cornerstone in addressing the data requirements for SDGs. Indeed, a preliminary analysis by the UN Intersecretariat Working Group on Household Surveys identified that 80 out of the 232 SDG indicators were sourced from household surveys, especially in goals related to public services delivery (e.g., health, education, hunger, poverty, inequality, cities and communities, water and sanitation, and energy) (IWGHS 2019).

In addition to available data, evidence can also be sourced from expert knowledge (Gold et al. 1990; Radaelli 1995; Renooij 2001; Fenton and Neil 2012; Head 2016; Constantinou et al. 2016). As Head (2016) put it: "within public policy discussions, it is axiomatic that [both] reliable information and expert knowledge are integral to sound processes for formulating and implementing policy". As a matter of fact, expert knowledge gains significance in data-scarce contexts such as developing countries, where it can help fill the gaps (Gibson and Le 2019).

Decision analysis for SDGs

Whilst it is true that data – and expert knowledge – are needed to measure progress towards delivering the 2030 Agenda for Sustainable Development, this does not imply gathering data just for the sake of it. Simply collecting more data is not enough: data must be analyzed and used to be valuable for monitoring and policy-making (Jütting and McDonnell 2017). Otherwise "data graveyards" are generated, and the cost, time and effort to collect them are wasted (Custer and Sethi 2017). It is hence crucial to transition from a data collection focus to one that promotes the use of those data to make evidence-based decisions for improved SDG implementation.

In an attempt to avoid spending meager resources on collecting data that will not be utilized, Shepherd et al. (2015) suggested to integrate decision analysis approaches with SDG monitoring. According to Shepherd and colleagues, the application of decision analysis concepts and tools is the most effective way to harness data for development decision- and policy-making. Along similar lines, Whitney et al. (2018) emphasized that decision analysis approaches could help decision-makers allocate resources more efficiently and consequently enhance the effectiveness of policy decisions. Rosenstock et al. (2017) also underlined the need to shift from a "more is better" monitoring framework to a "less is more" philosophy in which data are analyzed and interpreted for a cost-efficient decision-making.

It becomes evident that the markup from using decision analysis techniques in sustainable development is that it actually generates information that improves the way people make decisions, and ultimately the policies that are adopted. So the question that remains is: why is decision analysis not used more extensively in the context of international development? For instance, a survey conducted by Clapp et al. (2013) on the use of data analysis techniques by stakeholders in African agriculture revealed that more than half of respondents (54%) could not identify a policy decision that would be reinforced by data. If we know that policies for sustainable development affect millions of people and require vast amounts of money, then why are decisions not well supported by evidence?

A prime reason for why decision analysis is not embedded in policy development – both at the international level and down to sub-national levels – lies in that decision analysts are often missing from the decision-making process. As Kahneman explained, “decision analysts are not going to control the world, because the decision-makers, the people who are in charge, do not want to relinquish the intelligence function to somebody else” (Schrage 2003). Additionally, most decision analysts work for academic institutions that are seldom involved in real-world decision processes (Ferrier et al. 2016) and that decision and data analyses are viewed as “overtime activities” only “a few intrinsically motivated officials” use to base their policy decisions on (Development Gateway 2016). Consequently, for decision analysts to influence decision-making processes in a tangible, positive manner, it is crucial that they engage with decision- and policy-makers, and vice versa.

Certainly, as illustrated in Figure 1, decision analysis can aid decision-making for sustainable development by bridging science and policy. A decision analysis process can bring together the several actors playing a role: decision-makers (or policy-makers), stakeholders, potentially subject experts, and decision analysts. Decision-makers *own the problem*: they are responsible for making the decisions on which policies to implement. Decision-makers, and their decisions, are partially – and not necessarily fully – accountable to the stakeholders involved in the problem (e.g., the public, industry, interest groups, etc.), who share the impacts arising from a decision. Experts provide their professional advice and expertise on the content of the decision, whereas analysts provide their process skills to structure and conduct the analysis. Besides improving the decision analysis model, the participation of these different actors allows for the consideration of manifold perspectives on the decision and, more importantly, increases the chance that the model outputs are actually used in practice.

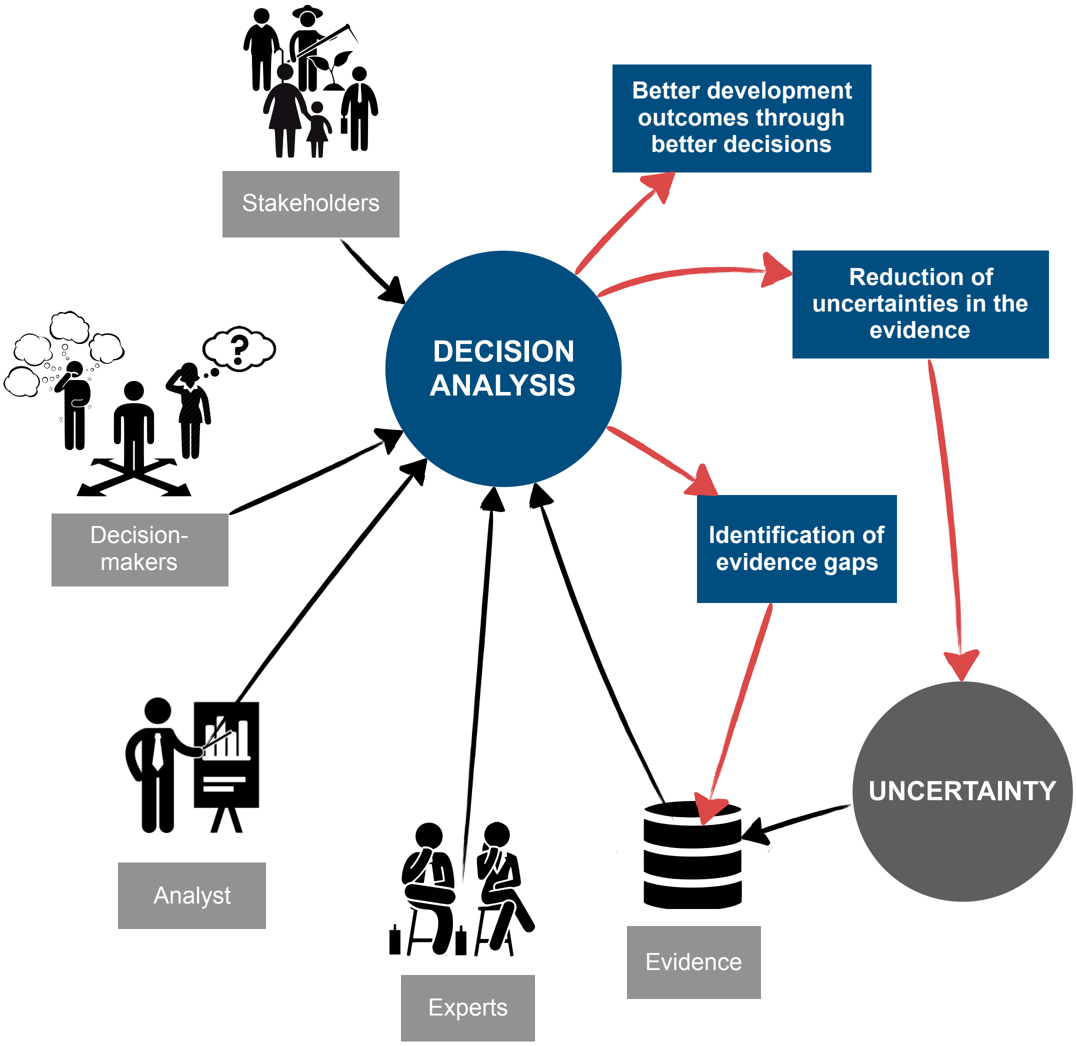

Figure 1: Decision analysis process for sustainable development (adopted from Luedeling and Shepherd (2016)). Black arrows connect the actors that deliver inputs to a certain decision analysis process, and the red arrows connect the output ports of such decision decision analysis process with key decision-making levers. These, in turn, may provide feedback into the decision analysis process through "uncertainty" and "evidence", i.e., by incorporating new or modified knowledge into it.

Therefore, decision analysis processes can become particularly relevant in decision-making for public services delivery, which is key to the 2030 Agenda. Indeed, all 17 SDGs and their 169 targets depend, directly or indirectly, on the provision of public goods or the implementation of public policies for their successful achievement (GCPSE 2016). By blending together different types of evidence from multiple sources and diverse perspectives from key actors, decision analysis can capture and simplify substantially the technical complexity of public services planning. Furthermore, decision analysis can tackle the increasing need to assess services delivery issues in a holistic manner, especially for the prioritization of public policy options and the monitoring of service coverage.

Multi-criteria prioritization approaches help tackle the siloed and often conflicted nature of public services planning. For instance, multi-criteria composite indexes are widely used for policy evaluation in health (Antony and Rao 2007; Peppard et al. 2008; Sartorius and Sartorius 2014; Lagravinese et al. 2019), water and sanitation (Cohen and Sullivan 2010; Giné Garriga and Pérez Foguet 2010; Jeyakumar and Ghugre 2017; Chaudhuri1 et al. 2018), and energy (Nussbaumer et al. 2012; Kılkış 2015; Iddrisu and Bhattacharyya 2015; Gouveia et al. 2019). However, improvements in the way these indexes are constructed remains a critical research issue at both theoretical and practical levels (Munda and Nardo 2009; Mazziotta and Pareto 2016).

Additionally, public services monitoring can be strengthened by retrospective and prospective analyses of service coverage over time. Analysis of patterns and trajectories is particularly important to gauge whether inequalities in access and service levels are being progressively reduced (McArthur and Rasmussen 2019). This means it helps identifying issues and collectives of people that are being left behind. For this, complete and consistent time-series data are needed, which are not always available. The challenges associated with data availability are further amplified by the unsound use of standard statistical approaches for the analysis of compositional data subject to non-negativity and constant-sum constraints (Lloyd et al. 2012).

Given the importance of both multi-criteria prioritization and trend analysis in public services provision, more efforts should be made to develop *improved* methods addressing the limitations of composite indexes and standard statistical analysis.

Coping with the uncertainty in the evidence

Apart from improving the decision analysis techniques for a better interpretability of the evidence base, it is paramount to remember that all evidence involves a degree of uncertainty. Acknowledging uncertainty in evidence might sound like an oxymoron, but uncertainty is inherent to evidence and its evaluation, and an important fraction of the literature fails to take it into account. Nonetheless, uncertainty should not be misinterpreted as a synonym for *absence of evidence*. Walker et al. (2003) described uncertainty as “any deviation from the unachievable ideal of complete deterministic knowledge of the relevant system” and classify it along three dimensions: location, level and nature. Location refers to where uncertainty manifests itself within the decision model, and can be specialized into: input data, model parameters, model structure and context. The level dimension focuses on the degree of uncertainty, which ranges from determinism (i.e., the absence of uncertainty) to statistical uncertainty, scenario uncertainty, recognized ignorance, and total ignorance. While determinism is the ideal situation where evidence is complete, exact and precise, it is rarely achieved in policy-making. Lastly, the nature dimension identifies its origin: is the uncertainty due to the lack of, or imperfection, of knowledge (i.e., epistemic uncertainty) or to the inherent variability of the system being modelled (i.e., ontological uncertainty)?

Understanding all expressions of uncertainty helps identifying those that are critical in the decision. This, in turn, is a key step for a better acknowledgement and management of uncertainty in decision analysis and support endeavors. However, accepting that uncertainty is ubiquitous to decision- and policy-making has substantial implications. First and foremost, it entails that ignoring uncertainties can undermine the policy decisions adopted and ultimately lead to an inefficient use of resources. Second, it requires characterizing and quantifying uncertainty, especially the one arising from the evidence base. For instance, data derived from household surveys – one of the main sources of information for policy-making – are intrinsically subject to a certain degree of statistical uncertainty. Knowledge is also prone to uncertainty, as it is often incomplete, imprecise, vague, not fully reliable or even contradictory (Janssen et al. 2010). Last but not least, it means that practical tools are needed to address and incorporate uncertainty into decision analysis processes.

Aims and methods

After explaining the theoretical underpinnings of our work in the Introduction, we now define the research problem we tackle, and present the objectives and research questions we focus on. We also provide a brief overview of the topics addressed in the thesis and the research method used.

The research problem

Despite the need to characterize and incorporate uncertainty of the evidence base when planning for sustainable development, uncertainty is seldom integrated in decision analysis models. The reason behind this is not the lack of conceptual approaches; quite the contrary, there is a multitude of mathematical methods for coping with uncertainty in decision problems that can be useful in policy-implementation decisions, in particular for public services provision. These include probability-based approaches, fuzzy numbers, Bayesian models, scenario analyses, or risk-based approaches (Broekhuizen et al. 2015). The main obstacle to using these approaches is their complexity and the high level of technical expertise required to apply them.

Matching the complexity of uncertainty approaches to the decision-making needs while keeping them as simple as possible is essential for widening the use of decision analysis under uncertainty. To date, however, only a few attempts have been made to facilitate the integration of uncertainty in sustainable development practice. Furthermore, only a handful of research studies have specifically focused on developing simple uncertainty approaches that could be applied to plan for public services provision.

Having identified such shortcoming in the literature, this thesis aims at enabling a systematic consideration of uncertainty in planning for the 2030 Development Agenda, in particular for the prioritization and monitoring of public services in health (SDG 3), water and sanitation (SDG 6), and energy (SDG 7).

The rationale underlying this work is that the incorporation of uncertainty into decision analysis methods is a necessary condition for better policy-making for services delivery and ultimately sustainable development. We argue that accounting for this uncertainty in the evidence has a twofold motivation. On one hand, policy and planning are more effective if uncertainty is characterized and addressed in decision analysis processes. On the other hand, highlighting the magnitude of uncertainty in critical evidence can trigger efforts towards improving it by reducing uncertainties. This would lead to evidence with well-bounded uncertainty, and thus better informed decision-making for sustainable development.

Objective and research questions

The overall objective of the research presented in this thesis is as follows:

To facilitate the incorporation of uncertainty in the evidence into decision analysis for sustainable development.

To address this objective, we develop and present simple methodological tools that shall improve the planning, monitoring, and evaluation of SDGs, especially in the frame of public services provision, by dealing with the inherent uncertainty of the evidence on which decisions are based.

In particular, this thesis examines two research questions:

- How can we include the uncertainty of the evidence in the prioritization of policy options for service provision?
- How can we incorporate this uncertainty in the trend analysis of service coverage for progress monitoring?

The first research question tackles the prioritization of policy options based on multiple – possibly conflicting – criteria or goals. It consists in the evaluation and ranking of different policy alternatives, and the selection of the best alternative with respect to the preferences of the actors involved in the decision. Alternatives can be policy measures, strategies, scenarios, or other actions that could solve the problem at hand, or the geographical areas (communities, regions, countries, etc.) that must be ranked from neediest to least needy for further services improvements. Our solution to this question is based on the ELimination and Choice Translating REality (ELECTRE) III model, a non-compensatory multi-criteria analysis tool that overcomes the limitation of composite indexes (i.e., their compensatory aggregation).

The second question focuses on the statistical analysis of temporal trends in service coverage for key development metrics (e.g., health-care and water and sanitation services) to evaluate performance and guide the design or targeting of policy interventions that could accelerate the attainment of SDGs. Trend analysis can be done at different levels (including sub-national, national, regional and global) to provide a comprehensive picture of SDG implementation. Our solution is based on compositional data analysis, a branch of statistics that deals specifically with relative data such as population and service coverage data.

Overview of the topics addressed

This work revolves around three axes, as illustrated in Figure 2. The first part focuses on multi-criteria decision analysis under uncertainty, where we deal with the prioritization of policy alternatives for energy (Chapter 1) and water and sanitation (Chapter 2) planning. The outcome is an improved prioritization model based on ELECTRE III. The second part targets trend analysis of service coverage, considering its underlying uncertainty, for water and sanitation (Chapter 3) and child mortality (Chapter 4). In this case, the outcome is an improved regression model based on compositional data theory. The third and final part lays out some guidance – and future perspectives – on the incorporation of uncertainty in decision analysis for sustainable development.

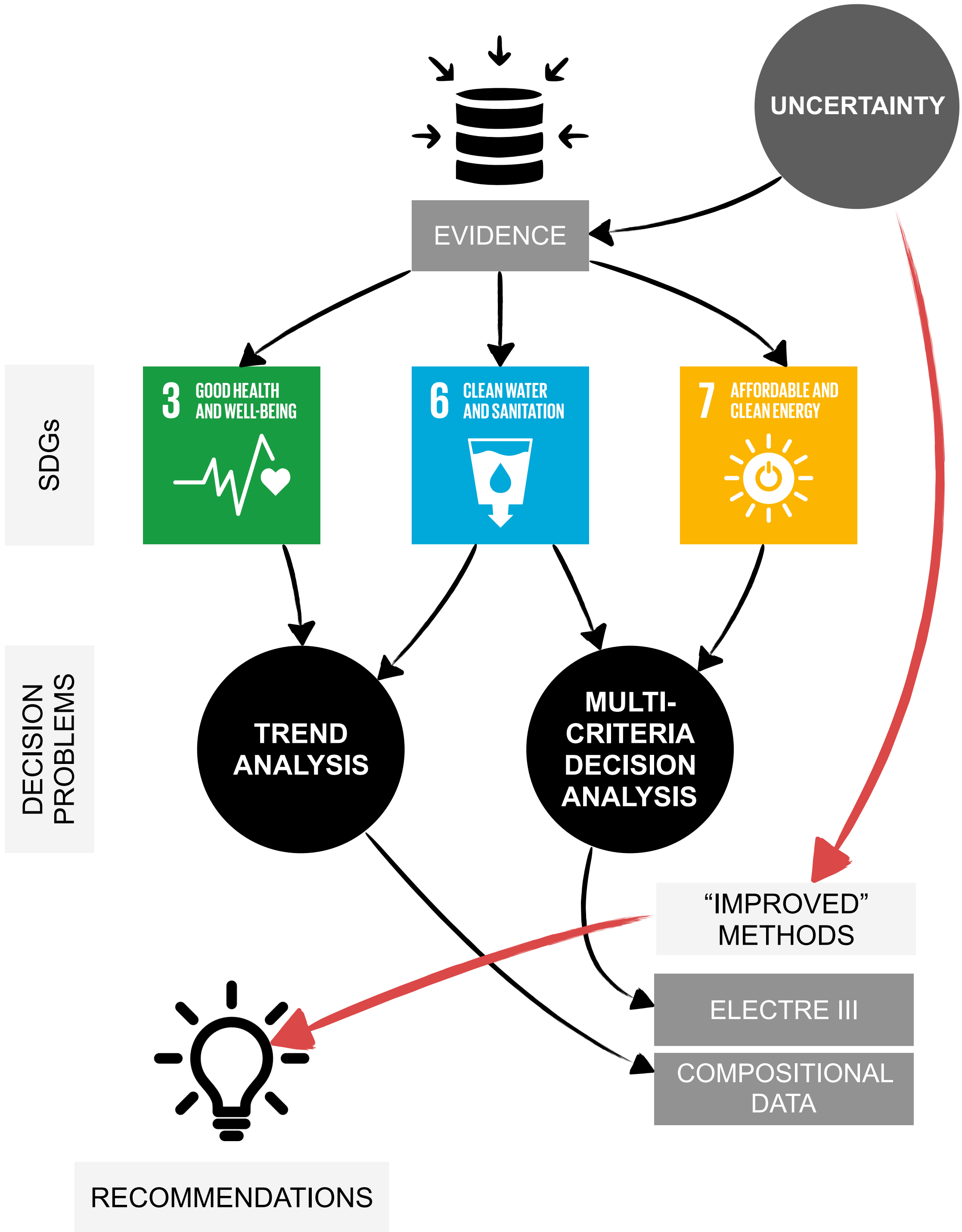

Figure 2: Overview of the topics addressed in this thesis. Our goal is to provide *improved* methods for the incorporation of uncertainty in the evidence into two decision problems: (i) multi-criteria analysis for the prioritization of public policy options, and (ii) trend analysis for the monitoring of service coverage. The two decision problems are tackled with ELECTRE III and compositional data theory, respectively. We focus on service provision for SDG 3 on "good health and well-being", SDG 6 on "clean water and sanitation" and SDG 7 on "affordable and clean energy". Finally, based on the application of these two methods, we provide some recommendations and guidance on the integration of uncertainty in decision analysis for sustainable development.

In more detail:

Chapter 1. Strategic renewable energy planning

We present Multi-Criteria Decision Analysis under uncertainty for the prioritization of policy options. The approach, based on ELECTRE III, incorporates uncertainty in the evidence base directly into the decision analysis, without the need of complex uncertainty approaches, such as fuzzy sets theory. It is tested in the case of renewable energy planning in Turkey, and compared with other uncertainty approaches from the literature.

Chapter 2. Targeting and prioritization in the WASH sector

We compare two data-driven Multi-Criteria Decision Analysis under uncertainty approaches for the prioritization of communities. The comparison looks at how these two approaches – based on compensatory and non-compensatory aggregation procedures – integrate data uncertainty into the decision analysis. It is applied to the case of targeting of rural communities for water, sanitation and hygiene (WASH) services in Kenya.

Chapter 3. Global monitoring of access to WASH

We introduce an approach for characterizing the uncertainty around temporal estimates of WASH. The approach provides a response to the issue of non-reporting of standard errors in household surveys, as well of producing "better" estimates that account for the compositional and non-linear nature of the data. It is illustrated in four countries (Bolivia, Gambia, Morocco and India) with data from WHO/UNICEF's Joint Monitoring Programme.

Chapter 4. Level and trends in child mortality

We improve the monitoring of child mortality by developing a trend analysis approach based on compositional analysis. The approach is applied to countries of sub-Saharan Africa, and compared with official estimates provided by the UN's Inter-agency Group for Child Mortality.

Research method

To tackle the formulated questions, we integrate three general research steps, which are briefly outlined below.

The first step consists in a literature review to identify the relevant theoretical and methodological debates in decision analysis. We consult a wide and diverse array of information sources, covering scientific papers, handbooks, published reports by UN agencies and other organizations, and grey literature such as working papers, conference proceedings and project reports. The broad themes addressed include, but are not limited to, decision analysis theory, monitoring strategies of SDG targets, uncertainty approaches and compositional data analysis. The insights from the literature review have informed the theoretical framework of this thesis.

The second step entails the development of *improved* methodological approaches, based on ELECTRE III and compositional data theory, for the integration of uncertainty in the decision analysis process. The underlying philosophy of these techniques is *keeping it simple*. Our main objective is to offer a set of uncertainty approaches with a reasonable balance between simplicity, robustness, and accuracy. We specifically target methods that could be applied to a wide range of decision problems in public services provision for SDG implementation. These methods are developed such that they could be replicated by other researchers, and decision analysis, in reasonable timescales. To this end, all technical details of the methods that this work puts forward are publicly available[1].

The third and last step is the implementation of the proposed approaches in real-life decision problems in order to examine their validity and usefulness. Four case studies are selected to test our methods and validate our research:

- Strategic planning for renewable energy development in Turkey.
- Geographical targeting of water and sanitation interventions in Kenya.
- Monitoring of water and sanitation in Bolivia, Gambia, Morocco and India.
- Monitoring of child mortality in countries of sub-Saharan Africa.

[1] All R scripts and datasets used to develop these methods are available in the Zenodo repository: https://zenodo.org/communities/escgd

We use two types of evidences in these case studies. To inform decision-making in the case of energy (Chapter 1) we refer to expert knowledge; in the cases of water and sanitation (Chapters 2 and 3) and child mortality (Chapter 4) we use available data obtained mainly from household surveys.

As seen, three different SDGs are addressed by this thesis, which play a central role in public services delivery for SDGs (Le Blanc 2015; Zhou and Mounuddin 2017):

- SDG 3 on “ensuring healthy lives and promoting well-being for all at all ages” (in Chapter 4).
- SDG 6 on “ensuring available and sustainable management of water and sanitation for all” (in Chapters 2 and 3).
- SDG 7 on “ensuring access to affordable, reliable, sustainable and modern energy for all” (in Chapter 1).

It is important to emphasize that our technical contribution goes beyond our application to these three SDGs. Indeed, our *improved* prioritization of public policies can be applied in other fields: from food (SDG 2) and education (SDG 4) to housing and transport (SDG 11). Furthermore, our *improved* monitoring of population groups – and their service levels – is particularly relevant to reducing inequalities (SDG 10) in service coverage.

1

Strategic renewable energy planning

Abstract:

Multi-Criteria Decision Analysis (MCDA) methods are increasingly used to aid decision making for renewable energy planning. However, although uncertainty is present in all decision environments and can be accounted for in MCDA in various ways, dealing with incomplete and vague information in decision analysis remains a challenge. The objective of this Chapter is to simplify the incorporation of uncertainty in the scoring of alternatives in MCDA processes. A modified ELimination and Choice Translating REality (ELECTRE) III model is presented, in which the uncertainty in the performance scores is expressed as lower/upper bounds and then it is added to the model's discrimination thresholds. Unlike other uncertainty approaches developed in the literature (such as those based on fuzzy set theory), our approach does not require additional knowledge apart from understanding the ELECTRE III model. The proposed approach is applied for the evaluation of renewable energy resources for Turkey – hydro, wind, geothermal, solar and biomass – under five main criteria: technological, technical, economic, environmental and socio-politic.

Keywords: Multi-Criteria Decision Analysis; Energy; SDG 7; Expert knowledge

This chapter is based on:

- Ezbakhe, F. and Pérez-Foguet, A., 2019. Decision analysis for sustainable development: the case of renewable energy planning under uncertainty. *European Journal of Operational Research* (Under Review)

1.1 Introduction

Sustainable energy development, and the transformation towards renewable energy sources, is a top priority on the international agenda. Evidence of this is that Sustainable Development Goal 7 (SDG 7) of the 2030 Agenda champions renewable energy adoption, and calls for governments to"ensure access to affordable, reliable, sustainable and modern energy for all". Although not explicitly stated, achieving this goal requires the integration and balance of all three dimensions of sustainable development: economic, social and environmental. As the former UN Secretary General Ki-Moon (2012) once declared, "energy is the golden thread that connects economic growth, social equity, and environmental sustainability". However, integrating multi-dimensional issues into decision-making and planning is far from an easy endeavor, as it requires achieving true policy coherence and linkages across sectors and actors (Stafford-Smith et al. 2017). If decision-making is to have a multi-dimensional perspective, then policy-makers require more mathematical tools to tackle these type of decision problems.

Multi-Criteria Decision Analysis (MCDA) methodologies have been increasingly recognized as a practical for decision problems in energy planning (Diakoulaki et al. 2005; Løken 2007; Kurka and Blackwood 2013; Wu et al. 2018; Marttunen et al. 2018; Bhardwaj et al. 2019). MCDA can help establish a coherent picture about complex decision problems by dividing them into three elements (Belton and Stewart 2002). First, the alternatives to be appraised, which can be policy options, strategies or action plans for the energy sector. Second, the objectives and criteria for assessing the consequences of each option (e.g., technical, economic, social or environmental characteristics). Third, the weights for each criterion to reflect their relative importance in the decision. These three elements are then combined to derive the rankings and inform decision-makers on the most suitable option. From this point of view, MCDA models can aid decision analysis by capturing the multi-faceted implications of energy choices across a wide range of evaluation criteria.

Adding to the inherent complexity of decision-making is the fact that available information is often incomplete and vague (Hokkanen et al. 2000). The uncertainties faced by decision-makers can arise in each and every step of the MCDA process, for instance, when identifying the alternatives and criteria, scoring the alternatives or weighting the criteria. To account for the associated uncertainties in decision analysis, different hybrid MCDA methodologies have been proposed, in which MCDA models are supplemented with uncertainty approaches such as deterministic sensitivity analysis, probability models, fuzzy set theory and grey systems theory (Broekhuizen et al. 2015). Albeit several options exist to address uncertainty, some require a relatively high technical expertise. More simple and easy-to-implement approaches for MCDA under uncertainty are hence needed for decision-makers – and the other actors involved– not familiarized with uncertainty analysis. By keeping the appraisal of uncertainty straightforward, the whole decision analysis process actually becomes more transparent and credible. This is particularly important in developing countries, where the applicability of some uncertainty approaches is oftentimes limited by the knowledge required to implement them. Indeed, as Inotai et al. (2018) underline in their guidance towards the implementation of MCDA frameworks in developing countries, "although good practices of MCDA development are widely published, approaches which were proven to be applicable in developed countries may not be feasible in developing countries with limited capacities for decision making".

In light of the need to achieve a balance between capturing the uncertainties in decision analysis and keeping the MCDA process comprehensible for decision-makers, we develop a simple MCDA approach for dealing with the uncertainty arising from the scoring of alternatives. The key features of our work are:

- We present modified version of ELECTRE III, where the uncertainty in performance scores is directly taken into account at the phase of determining the model's discrimination thresholds. Specifically, we express the discrimination thresholds of ELECTRE III as a function of the lower and upper score bounds elicited by decision-makers.

- We apply the proposed model to the case of sustainable energy planning in Turkey, in which we use evidence from expert knowledge to assess the suitability of 5 renewable energy technologies (hydro, wind, solar, geothermal and biomass) based on 31 different technological, technical, economic, environmental and socio-politic criteria.

- To demonstrate the effectiveness of the proposed approach, we compare the ranking of energy alternatives through our modified ELECTRE III model with those of three other hybrid MCDA methods proposed by Mousavi et al. (2017), Erdogan and Kaya (2015) and Kahraman and Kaya (2010). When comparing the methods we consider both the convergence of rankings and the ease of implementation of the uncertainty approaches used.
- Finally, we analyze the policy implications of the resulting rank orders on Turkey's energy sector.

We structure this Chapter as follows. In Section 1.2, we briefly explain the uncertainty approaches used in MCDA, particularly for renewable energy planning. Next, in Section 1.3 we introduce the case of energy planning in Turkey and describe the particular decision problem we analyze. In Section 1.4, we provide an overview of ELECTRE III methodology, together with an explanation of our approach for accounting for uncertainty in alternatives' scores. In Section 1.5, we present and discuss the results of applying our modified ELECTRE III model to the evaluation of renewable energy alternatives in Turkey, comparing it to other uncertainty approaches, and analyzing the policy implications for Turkey's energy sector. Finally, we summarize the main conclusions of our work in Section 1.6.

1.2 A background on MCDA under uncertainty

In this section we present a background on the approaches used to handle uncertainty in MCDA (1.2.1) as well as an overview of the application of MCDA under uncertainty for renewable energy planning (1.2.2).

1.2.1 Handling uncertainty in MCDA

Any MCDA involves eight main stages, which include: (i) establishing the decision context, (ii) identifying the options (i.e., the alternatives) to be appraised, (iii) identifying the objectives and criteria for assessing the consequences of each option, (iv) "scoring" the alternatives (i.e., assessing the expected performance of each option against the criteria), (v) "weighting" the criteria (i.e., assigning weights for each criterion to reflect their relative importance in the decision), (vi) applying the selected MCDA model to combine the weights and scores for each option and derive the ranking, (vii) examining the results, and (viii) conducting a sensitivity analysis to determine whether other preferences affect the ordering (Dodgson et al. 2009).

As seen, the outcome of any MCDA model depends on the assumptions that are made when building and populating it with criteria weights and performance scores (Goetghebeur and Wagner 2017; Groothuis-Oudshoorn et al. 2017). The incomplete knowledge about the structure of the decision problem and its input information is known as "uncertainty". There are many types of uncertainties, and they play different roles in the various stages of a MCDA. For instance, during the scoring stage, in which available data or expert judgment are used to evaluate the performance of the alternatives on the different criteria, four types of uncertainty can be distinguished: stochastic uncertainty, parameter uncertainty, heterogeneity, and structural uncertainty (Briggs et al. 2012). In the case of performance scores elicited from experts, stochastic uncertainty refers to the random variability performance values as assigned by the same person. Parameter uncertainty reflects the variability in the estimation of the performance values, for instance, the error in estimating the mean value given by a group of experts to an alternative. Heterogeneity is the between-person variability that can be explained by the person's characteristics, and structural uncertainty relates to the assumptions inherent to the decision model, such as those underlying the choice of the expert elicitation technique.

Different approaches have been proposed to handle uncertainty in MCDA (Durbach and Stewart 2011, 2012; Broekhuizen et al. 2015), namely deterministic sensitivity analysis, probabilistic models, fuzzy set theory and grey systems theory:

- **Deterministic sensitivity analysis**: uncertainty in performance scores is examined by varying one score at a time and studying the impact of this variation on the rank order of alternatives. If the ranking does not change with the induced variation, the decision is considered robust. Otherwise, one can assess the extent to which a performance score can be changed before a different rank order of alternatives is obtained. The range in which the particular performance score is likely to vary can be based on available data or experts' judgment. Deterministic sensitivity analysis is the most straightforward method for taking into account uncertainty in performance scores. Its drawback is the assumption of uni-variate uncertainty, which prevents the evaluation of the cumulative impact of uncertainty in multiple performance scores.

- **Probabilistic models**: uncertainty in performance scores is represented with probability distributions that can be either assigned from available data or elicited from the decision-makers (O'Hagan et al. 2006). After selecting the probability distribution, the uncertainty can be propagated through the MCDA model through simulation approaches (e.g., Monte Carlo simulations). This allows to assess the probability distribution of the alternatives' overall score, and make probabilistic statements to describe the chance of occurrence for a particular rank order of alternatives. Yet, the process of assigning probability distributions for each alternative on each criterion can be time consuming.

- **Fuzzy set theory**: this approach was first introduced by Zadeh (1965) to handle the vagueness and imprecision in human judgments. Elements have a degree of membership to a set, which is expressed as a value between zero (no membership) and one (full membership). Within the context of MCDA, decision-makers must define fuzzy sets and the membership functions to capture the uncertainty in performance scores. Similarly to the probabilistic models, fuzzy set theory can be time consuming, as it requires defining and agreeing on the fuzzy set membership functions.

- **Grey systems theory**: the approach, first developed by Deng et al. (1982), can also handle associate vagueness in human judgments. In this case, uncertainty in performance scores can be represented with three value ranges: black, white and grey. Black numbers indicate a complete lack of knowledge (i.e., the value can go from minus infinity to plus infinity), while white numbers denote complete knowledge (i.e., the performance score has a single value). Grey numbers are between these two extremes, and the performance score can be defined between a lower and upper bound. Contrary to fuzzy set theory, grey systems theory is more straightforward to use, since in case of disagreement on the performance scores, the lowest and highest values can be used as the lower and upper bounds for the grey numbers.

A review completed by Broekhuizen et al. (2015) revealed that fuzzy set theory has become the most common uncertainty approach in MCDA (306 of the 632 applications identified; i.e., 48%), followed by deterministic sensitivity analysis (32%), probabilistic models (16%) and grey systems theory (4%). A recent comprehensive literature analysis by Kaya et al. (2019) also demonstrates this approach's extensive application in the energy field. All three models (Kahraman and Kaya 2010; Erdogan and Kaya 2015; Mousavi et al. 2017) with which our model is compared are based on fuzzy set theory.

1.2.2 MCDA for renewable energy planning and policy

Adopting and selecting alternative energy sources is inherently a multidimensional decision making process: it involves looking at a broad spectrum of renewable energy resources or conversion technologies and analyzing their multiple characteristics at different levels (e.g., technical, economic, social and environmental). From this point of view, MCDA can provide an evidence-based decision-making support tool that allows to justify choices in the renewable energy sector. In this sense, four categories of MCDA application in renewable energy can be distinguished: renewable energy planning and policy, renewable energy evaluation and assessment, renewable energy technology and project selection, and environmental impact assessment (Abu Taha and Daim 2013). We focus on the first area, which refers to the assessment of a feasible energy plan and/or the diffusion of alternative renewable energy options in order to reach a certain national target.

There is a myriad of MCDA applications in renewable energy planning and policy making, as revealed by the reviews of Wu et al. (2018) and Marttunen et al. (2018). Some recent examples include the following. Seddiki and Bennadji (2019) applied a fuzzy Preference Ranking Organization METHod for Enrichment of Evaluations (PROMETHEE) model to assess renewable energy alternatives for electricity generation in residential buildings in Algeria. They concluded that photo-voltaic panels were the best alternative due to their good characteristics in the payback period and the energy production. Ighravwe and Babatunde (2018) used a fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) to address the problem of renewable energy planning for mini-grid energy distribution systems in Nigeria, and identified biomass as the best-ranked renewable energy. McKenna et al. (2018) combined Multi-Attribute Value Theory (MAVT) with multi-dimensional sensitivity analysis to evaluate alternative energy systems for 2030 in a small community of Germany. They determined that the best alternatives consisted in those maximizing both environmental sustainability and local energy autonomy. Lee and Chang (2018), on the other hand, presented a comparative analysis of ranking renewable energy sources for electricity generation in Taiwan using four MCDA methods (amongst them TOPSIS and ELECTRE), and determined that hydro power was the optimal alternative in all methods. Abdullah and Najib (2016) proposed a fuzzy Analytic Hierarchy Process (AHP) for sustainable energy planning in Malaysia, and after evaluating seven energy alternatives, they selected nuclear energy as the most suitable option.

In Turkey particularly, MCDA has been widely used for renewable energy planning. Erdin and Ozkaya (2019) have recently applied ELECTRE to decide on the most appropriate renewable energy alternative for different geographic regions in Turkey. Solar and biomass energy were found to be the most suitable for Central Anatolia region, while in the Aegean these were wind and geothermal, and hydropower for the Black Sea. Ervural et al. (2018) used fuzzy TOPSIS to prioritize between nine alternative energy strategies for Turkey's energy planning, revealing that an increase in the share of renewable energy resources within the energy supply was the third best policy (the first and second were turning the country into an energy hub and improving energy efficiency, respectively). Atilgan and Azapagic (2017) assessed future electricity scenarios for Turkey up to 2050 and suggested that renewable scenarios – particularly based on hydropower– outperformed those dominated by fossil fuels in environmental and social aspects. Other applications evaluating renewable energy alternatives in Turkey are shown in Table 1.1.

Table 1.1: Recent MCDA applications for renewable energy planning in Turkey.

STUDY	MCDA METHOD	RENEWABLE ENERGIES EVALUATED						BEST
		Hydro	Solar	Wind	Geothermal	Biomass	Other	ALTERNATIVE
Yilan et al. (2020)	MAVT	x	x	x	x		x	Hydro
Karakaş and Yıldıran (2019)	Fuzzy AHP	x	x	x	x	x		Solar
Çolak and Kaya (2017)	Hesistant fuzzy TOPSIS	x	x	x	x	x	x	Wind
Büyüközkan and Güleryüz (2017)	Fuzzy DEMATEL-ANP-TOPSIS	x	x	x	x	x		Geothermal
Balin and Baraçli (2017)	ANP-TOPSIS	x	x	x	x	x	x	Wind
Çelikbilek and Tüysüz (2016)	Grey-based DEMATEL-ANP-VIKOR	x	x	x	x	x		Solar
Büyüközkan and Güleryüz (2016)	DEMATEL-ANP	x	x	x	x	x		Wind
Erdogan and Kaya (2015)	Type-2 fuzzy TOPSIS	x	x	x	x	x		Wind
Kuleli Pak et al. (2015)	ANP-TOPSIS	x	x	x	x	x		Hydro
Şengül et al. (2015)	Fuzzy TOPSIS	x	x	x	x	x	x	Hydro
Büyüközkan and Güleryüz (2014)	Fuzzy TOPSIS	x	x	x	x	x		Wind
Kabak and Dağdeviren (2014)	ANP	x	x	x	x	x		Hydro
Ertay et al. (2013)	Fuzzy AHP	x	x	x	x	x		Wind
Demirtas (2013)	AHP	x	x	x	x	x		Wind
Baris and Kucukali (2012)	Composite index	x	x	x	x	x		Biomass
Erol and Kılkış (2012)	AHP	x	x	x	x		x	Solar
Boran et al. (2012)	Intuitionistic fuzzy TOPSIS	x	x	x	x			Hydro
Kaya and Kahraman (2011)	Fuzzy TOPSIS	x	x	x	x	x	x	Wind
Kaya and Kahraman (2010)	Fuzzy AHP	x	x	x	x	x	x	Wind
Kahraman and Kaya (2010)	Fuzzy AHP-VIKOR	x	x	x	x	x	x	Wind
Kahraman et al. (2010)	Choquet integral	x	x	x	x	x	x	Wind

1.3 The case of energy planning in Turkey

In this section, we first present a general overview of Turkey's energy planning (1.3.1), highlighting its national renewable energy action plan for 2023. We then describe in detail the decision problem we analyze (1.3.2), which consists in selecting the most appropriate renewable energy option in Turkey according to the preferences of three energy experts involved with energy policy and planning.

1.3.1 An overview of Turkey's energy planning

With a rapid population increase, economic development and urbanization, Turkey has one of the fastest growing energy markets in the world. According to the International Energy Agency (2019b), energy consumption in Turkey has risen from 40,000 ktoe in 1990 to 97,850 ktoe in 2016. However, whereas demand for energy has grown, production has remained low (Bulut and Muratoglu 2018). As a result, Turkey imports a substantial amount of the energy it consumes in the form of oil, coal and natural gas from neighbouring countries, mainly Russia, Iran and Azerbaijan (Berk and Ediger 2018). Recent statistics by the International Energy Agency (2019a) show that net energy imports rose from 51% in 1990 to 75% in 2015. This strong foreign dependency on energy, together with the substantial price fluctuations, make energy security one of Turkey's top liabilities (Fackrell 2013). In order to avoid the risks deriving from energy import dependence, Turkey has made significant reforms in its energy sector. The most prominent one if the adoption of the "National renewable energy action plan for Turkey" (EBRD 2014), which puts a special emphasis on the key role of renewable energies in meeting the increasing demand, as well as reducing greenhouse gas emissions.

Indeed, due to its geographic position and climatic conditions, Turkey has a great renewable energy potential that can assist in guiding its energy policy. It is estimated that the total electricity generation potential from renewable energy sources is 240,165 GW/year for an economic potential of 138,000 MW, which equals to 13% of EU-27's total potential (Ozcan 2018). Specifically, the potential is (in MW) 36,000 hydro, 48,000 wind, 50,000 solar, 2,000 geothermal and 2,000 biomass. However, this potential is far from being fully exploited. The utilization rates of renewable energy sources are: 71.2% for hydropower, 8.9% for wind, 0.45% for solar, 30.7% for geothermal and 17.3% for biomass (Ozcan 2018). Now the goal for 2023 is to develop 30% of Turkey's total electricity generation mix from renewable sources: 34,000 MW of hydro, 20,000 MW of wind, 5,000 MW of solar, 1,000 MW of geothermal and 1,000 MW of biomass (Table 1.2, source: MENR (2018)).

Table 1.2: Renewable energy targets for electricity generation in Turkey, in MW.

	2017	2019	2023
Hydro	27,700	32,000	34,000
Wind	9,500	10,000	20,000
Geothermal	420	700	1,000
Solar	1,800	3,000	5,000
Biomass	540	700	1,000

Nearly 68.5% of the new renewable capacity to be added in the next four years is planned to come from wind energy. Furthermore, the 2023 target implies a 200% increase in wind generation capacity between now and 2023. Notwithstanding this strong commitment towards wind power, it remains important to evaluate available renewable energy options to help decision- and policy-makers reaffirm the legislation and targets put in place, as well as guide them in the development of future strategies or initiatives for strengthening the energy sector.

1.3.2 The decision problem

The decision problem, adopted from the literature (Kahraman and Kaya 2010; Erdogan and Kaya 2015; Mousavi et al. 2017), consists in evaluating and selecting the most appropriate renewable energy alternative. The energy alternatives considered are those included in Turkey's national renewable energy action plan: Hydro, Wind, Solar, Biomass and Geothermal. These five energy alternatives are assessed based on five dimensions: technological, technical, economical, environmental and socio-politic. The full 31 sub-criteria are represented in Figure 1.1 and defined in Table 1.3.

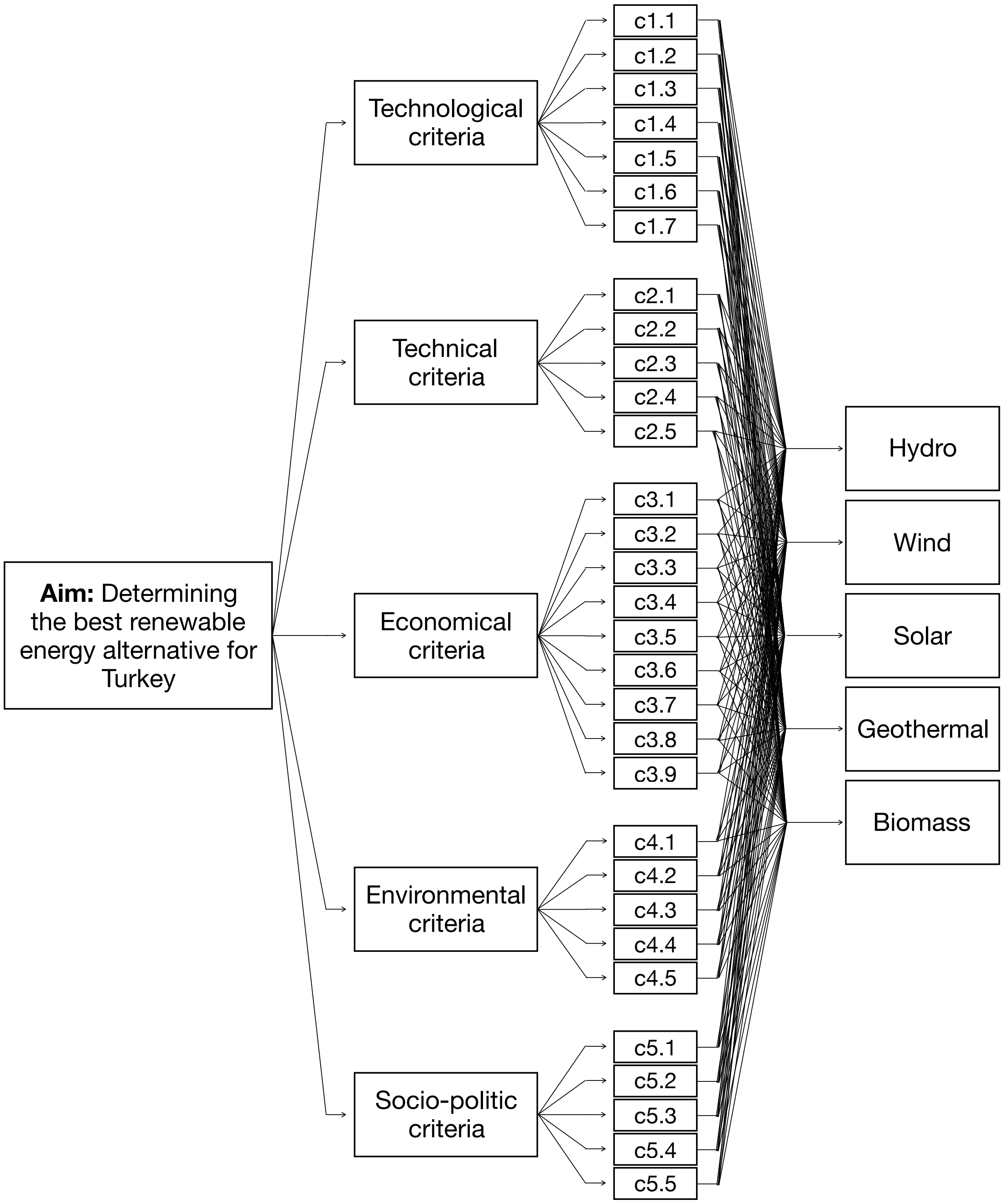

Figure 1.1: Structure of the energy decision problem analyzed. There are five criteria (technological, technical, economical, environmental and socio-politic). Each of these is split in several sub-criteria that we detail in Table 1.3. These sub-criteria are used to evaluate the performance of the five energy alternatives considered in Turkey's national renewable energy action plan: hydro, wind, solar, geothermal and biomass.

Table 1.3: Description of the criteria and sub-criteria of the energy decision problem. The corresponding direction of the sub-criteria means whether they need to be maximized ("bigger is better") or minimized ("smaller is better").

Criteria	Sub-criteria	Description	Direction
Technological	c1.1. Feasibility	Ability of the technology to be successfully implemented	Maximize
	c1.2. Risk	Extent to which the technology can fail	Minimize
	c1.3. Reliability	Ability of the technology to perform and operate without failure	Maximize
	c1.4. Duration of the preparation phase	Idle time between the technology decision and the kick-off of the implementation	Minimize
	c1.5. Duration of the implementation phase	Time until the technology is fully functioning	Minimize
	c1.6. Continuity and predictability of performance	Extent to which the technology can operate and perform continuously and confidently	Maximize
	c1.7. Local technical know-how	Assessment of the complexity of the considered technology and the capacity of local actors to ensure its operation and maintenance	Minimize
Technical	c2.1. Energy intensity	Quantity of energy produced by the technology	Maximize
	c2.2. Energy efficiency	Ratio between the output energy content and the total energy used	Maximize
	c2.3. Technical reliability	Ability of the whole system to perform and operate	Maximize
	c2.4. Technology readiness level	Maturity of the technology and its components	Maximize
	c2.5. Ease of access to the source	Accessibility to the energy source	Maximize
Economic	c3.1. Operation and maintenance costs	Costs for operating and maintaining the technology	Minimize
	c3.2. Investment costs	Total costs for the energy investment in order to be fully operational	Minimize
	c3.3. Economic value	Worth of the energy in economic terms (e.g., payback period, internal rate of return)	Maximize
	c3.4. Service life	Time period of intended use	Maximize
	c3.5. Local and regional economic development	Contribution to the local and regional economic growth	Maximize
	c3.6. Economic risks	Extent to which macroeconomic conditions (e.g., Government regulation or political stability) can affect the energy investment	Minimize
	c3.7. Security of energy supply	Energy supply at a price level that does not disrupt the operation	Maximize
	c3.8. Sustainability of energy resources	Degree of sustainability of the energy resources used	Maximize
	c3.9. Source durability	Extent to which the energy source remains serviceable	Maximize
Environmental	c4.1. Pollutant emission	Type and quantity of pollutant emissions (e.g., CO2) and their associated costs	Minimize
	c4.2. Land requirement	Total area of land used for implementing the technology	Minimize
	c4.3. Need of waste disposal	Extent of the damage on the quality of the environment	Minimize
	c4.4. Water pollution	Extent of the contamination of water resources	Minimize
	c4.5. Land disruption	Extent of the loss of environmental value due to the implementation of the technology	Minimize
Socio-politic	c5.1. Compatibility with the national energy policy objectives	Degree of convergence with the national energy policy	Maximize
	c5.2. Political acceptance	Degree of consensus among policy-makers on the energy candidate	Maximize
	c5.3. Social acceptance	Degree of consensus among civil society	Maximize
	c5.4. Labour impact	Direct and indirect positive repercussions on the local labour force	Maximize
	c5.5. Job creation potential	Direct and indirect employment generated from implementing the technology	Maximize

The decision process involved a group of three professional energy experts from the Ministry of Energy and Natural Resources (for more details refer to Kahraman and Kaya (2010)). These experts evaluated the energy alternatives under the selected criteria, and determined the criteria's weights. The performance values given by these experts were expressed in ranges between 0 and 1, with lower and upper bounds, as shown in Table 1.4.

Table 1.4: Decision matrix elicited from the group of energy experts.

Criteria	Sub-criteria	Weights	Alternatives *Hydro*	*Wind*	*Solar*	*Biomass*	*Geothermal*
	c1.1	3.728	[0.25 - 0.4]	[0.6 - 0.7]	[0.6 - 0.7]	[1.0 - 1.0]	[0.25 - 0.4]
	c1.2	4.307	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]
	c1.3	3.734	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[0.6 - 0.7]
Technological	c1.4	4.505	[1.0 - 1.0]	[1.0 - 1.0]	[0.6 - 0.7]	[0.6 - 0.7]	[0.25 - 0.4]
	c1.5	2.201	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]
	c1.6	2.844	[1.0 - 1.0]	[1.0 - 1.0]	[0.25 - 0.4]	[0.6 - 0.7]	[1.0 - 1.0]
	c1.7	0.811	[1.0 - 1.0]	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[0.6 - 0.7]
	c2.1	2.069	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[1.0 - 1.0]	[1.0 - 1.0]
	c2.2	3.856	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]
	c2.3	5.097	[1.0 - 1.0]	[1.0 - 1.0]	[1.0 - 1.0]	[1.0 - 1.0]	[0.25 - 0.4]
Technical	c2.4	2.237	[0.25 - 0.4]	[1.0 - 1.0]	[0.6 - 0.7]	[1.0 - 1.0]	[0.25 - 0.4]
	c2.5	3.248	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]
	c3.1	3.313	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[1.0 - 1.0]
	c3.2	4.779	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]
	c3.3	1.406	[0.25 - 0.4]	[0.6 - 0.7]	[1.0 - 1.0]	[1.0 - 1.0]	[1.0 - 1.0]
Economic	c3.4	1.215	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]	[1.0 - 1.0]
	c3.5	2.529	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]	[0.6 - 0.7]	[0.25 - 0.4]
	c3.6	0.625	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[1.0 - 1.0]	[0.25 - 0.4]
	c3.7	1.239	[0.25 - 0.4]	[1.0 - 1.0]	[0.6 - 0.7]	[1.0 - 1.0]	[1.0 - 1.0]
	c3.8	2.841	[1.0 - 1.0]	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]
	c3.9	5.824	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[0.6 - 0.7]	[1.0 - 1.0]
	c4.1	3.854	[1.0 - 1.0]	[1.0 - 1.0]	[0.6 - 0.7]	[0.25 - 0.4]	[1.0 - 1.0]
	c4.2	3.879	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]
Environmental	c4.3	5.096	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]
	c4.4	3.794	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[1.0 - 1.0]
	c4.5	3.970	[0.25 - 0.4]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]
	c5.1	4.459	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]
	c5.2	2.609	[1.0 - 1.0]	[0.25 - 0.4]	[0.6 - 0.7]	[1.0 - 1.0]	[0.25 - 0.4]
Socio-politic	c5.3	2.904	[1.0 - 1.0]	[1.0 - 1.0]	[0.25 - 0.4]	[1.0 - 1.0]	[0.6 - 0.7]
	c5.4	4.594	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]
	c5.5	2.433	[1.0 - 1.0]	[0.25 - 0.4]	[0.25 - 0.4]	[1.0 - 1.0]	[0.25 - 0.4]

1.4 Methodology

We first provide an overview of ELECTRE III method, in particular the concept of discrimination thresholds and the evaluation procedure (1.4.1). We then explain our modified version of the ELECTRE III model for incorporating uncertainty (1.4.2).

1.4.1 ELECTRE III

The problem of a multi-criteria decision analysis is usually composed by a set of m alternatives, $A = \{a_1, a_2, ..., a_i, ..., a_m\}$, and n criteria, $C = \{c_1, c_2, ..., c_j, ..., c_n\}$, with their relative importance coefficients (or criteria weights), $W = \{w_1, w_2, ..., w_j, ..., w_n\}$. This way, a $m \cdot n$ decision matrix can be constructed, M, where m_{ij} represents the performance value or score of the alternative a_i on criterion c_j, $g_j(a_i)$, for all $a_i \in A$ and $c_j \in C$. This decision matrix is Table 1.4 in our case study.

When using ELECTRE III, each alternative $a \in A$ is compared with every other alternative $b \in A - \{a\}$, with the aim of assessing the credibility of the assertion "alternative a is at least as good as alternative b" or, in other words, "alternative a outranks alternative b". The outranking relation between alternatives a and b is denoted as aSb (Roy 1991). Four situations may occur when comparing each pair of alternatives a and b: (i) aPb, a is strongly preferred to b; (ii) aQb, a is weakly preferred to b; (iii) aIb, a is indifferent to b; and (iv) aRb, a and b are incomparable.

To determine which of these four preference situations occurs and build the outranking relation, aSb, the method makes use of the concept of the pseudo-criterion and its three discrimination thresholds (Roy 1991). These thresholds account for the imperfect nature of the evaluations, and are as follows:

- Indifference threshold, q_j: difference between the alternatives' performances below which we are indifferent to the two alternatives for criterion c_j.
- Preference threshold, p_j: difference above which we show a clear strict preference of one alternative over the other for criterion c_j.
- Veto threshold, v_j: difference above which we negate any possible outranking relationship indicated by the other criteria.

These thresholds are not experimental values to be approximated to, but rather values defined by the decision-makers for assessing the appropriateness of the alternatives (Roy et al. 1986). Once chosen, these thresholds are used to test the outranking relation aSb based on the concordance/discordance principles (Figueira et al. 2010):

- Concordance: to validate an outranking relation aSb, a sufficient majority of criteria in favor of this assertion must occur.

- Discordance: the assertion aSb cannot be validated if a minority of criteria is strongly against this assertion.

The concordance concept represents the degree to which an alternative a outranks another alternative b, and is measured in two steps. First, a partial concordance is defined for each criterion: concordance is 1 when the j^{th} criterion fully supports the assertion aSb, and 0 when the criterion does not support aSb at all (Equation 1.1). A global concordance is then obtained by summing the weighted partial concordances for all criteria (Equation 1.2). On the other hand, discordance indicates the degree to which an alternative a cannot outrank alternative b: a value of 0 indicates that the j^{th} criterion does not oppose to the assertion aSb, while a value of 1 expresses a veto to aSb (Equation 1.3).

The evaluation procedure of the ELECTRE III method is illustrated in Figure 1.2, and follows the next steps (Roy 1991):

Step 1. Calculate the partial concordance index, $c_j(a,b)$, for each pair of alternatives and criterion.

$$c_j(a,b) = \begin{cases} 1, & \text{if } g_j(a) + q_j \geq g_j(b) \\ 0, & \text{if } g_j(a) + p_j \leq g_j(b) \\ \dfrac{g_j(a) - g_j(b) + p_j}{p_j - q_j}, & \text{otherwise} \end{cases} \tag{1.1}$$

Step 2. Compute the global concordance index, $C(a,b)$, for each pair of alternatives by summing the partial concordance indexes for all criteria according to their weights.

$$C(a,b) = \frac{1}{\sum_{j=1}^{n} w_j} \sum_{j=1}^{n} w_j \cdot c_j(a,b) \tag{1.2}$$

Step 3. Calculate the discordance index, $d_j(a,b)$, for each pair of alternatives and criterion.

$$d_j(a,b) = \begin{cases} 1, & \text{if } g_j(a) + v_j \leq g_j(b) \\ 0, & \text{if } g_j(a) + p_j \geq g_j(b) \\ \dfrac{g_j(b) - g_j(a) - p_j}{v_j - p_j} & \text{otherwise} \end{cases} \tag{1.3}$$

Step 4. Calculate the outranking credibility index, $S(a,b)$, for each pair of alternatives.

$$S(a,b) = \begin{cases} C(a,b), & \text{if } d_j(a,b) \leq C(a,b)\ \forall j \\ C(a,b) \cdot \prod_{j \in J} \dfrac{1 - d_j(a,b)}{1 - C(a,b)} & \text{where } J \text{ is the set of criteria} \\ & \text{such as } d_j(a,b) > C(a,b) \end{cases} \tag{1.4}$$

It is important to pinpoint that, for a single criterion analysis, the credibility degree is always equal to the concordance value, because $c_j(a,b)$ and $d_j(a,b)$ are complementary. For a multi-criteria analysis, the maximum value of the credibility degree, i.e., $S(a,b) = C(a,b)$, is reached when all criteria are in concordance with the assertion aSb. Otherwise, if one criterion is strongly against the assertion aSb, $d_j(a,b) > C(a,b)$, the credibility of the assertion is questioned and the concordance value is modified according to Equation 1.4. Furthermore, should the veto be imposed for one criterion, $d_j(a,b) = 1$, the credibility is zero whatever the concordance value.

Step 5. Exploit the outranking relations to obtain two partial pre-orders: $Z1$, which sorts and selects the alternatives from best to worst (i.e., descending distillation), and $Z2$, which does it from worst to best (i.e., ascending distillation). The steps to order the alternatives in the descending pre-order of $Z1$ are:

Step 5.0. Identify the set of alternatives included in the iteration step. For the first iteration, all alternatives are included (i.e. $\mathcal{D} = A$)

Step 5.1. Determine the maximum value of the credibility index.

$$\lambda_{max} = \max_{a,b \in \mathcal{D}, a \neq b} S(a,b) \tag{1.5}$$

Step 5.2. Calculate the cutoff level, λ, which represents the lower bound of the range of credibility indexes that will be taken into account in this iteration. It is based on the distillation coefficients α and β ($\alpha > \beta$). These coefficients are technical parameters usually assumed to be equal to 0.3 and 0.15, respectively.

$$\lambda = \lambda_{max} - (\alpha - \beta \cdot \lambda_{max}) \tag{1.6}$$

Step 5.3. At this cutoff level, define the relation $T(a, b)$ between each pair of alternatives $\forall a, b \in \mathcal{D}$.

$$T(a,b) = \begin{cases} 1, & \text{if } S(a,b) > \lambda \text{ and } S(a,b) - S(b,a) > \alpha - \beta \cdot S(a,b) \\ 0, & \text{otherwise} \end{cases} \tag{1.7}$$

Step 5.4. For each alternative a, determine its λ-strength and λ-weakness (i.e., the number of times alternative a is preferred over others and the number of times other alternatives are preferred over a, respectively).

$$\lambda\text{-strength}(a) = \mid \{b \in \mathcal{D} \ / \ T(a,b) = 1\} \mid \tag{1.8}$$

$$\lambda\text{-weakness}(a) = \mid \{b \in \mathcal{D} \ / \ T(b,a) = 1\} \mid \tag{1.9}$$

Step 5.5. Determine the qualification of each alternative, λ-qualification.

$$\lambda\text{-qualification}(a) = \lambda\text{-strength}(a) - \lambda\text{-weakness}(a) \tag{1.10}$$

Step 5.6. Select the subset of alternatives with the largest qualification, $\mathcal{D}^{max}$.

$$\mathcal{D}^{max} : \{a \in \mathcal{D} \ / \ \lambda\text{-qualification}(a) \geq \lambda\text{-qualification}(b) \ \ \forall b \in \mathcal{D}\} \tag{1.11}$$

Step 5.7. Remove the subset $\mathcal{D}^{max}$ from the process (i.e., $\mathcal{D} = \mathcal{D} - \mathcal{D}^{max}$).

Step 5.8. If $\mid \mathcal{D} - \mathcal{D}^{max} \mid \neq 0$ go to *Step 5.1.* Otherwise, the distillation is stopped. The ascending distillation follows the same steps, but at *Step 5.6*, the subset of alternatives with the lowest qualification are selected instead ($\mathcal{D}^{min}$).

Step 6. The two pre-orders are combined to form the final pre-order ($Z = Z_1 \cap Z_2$).

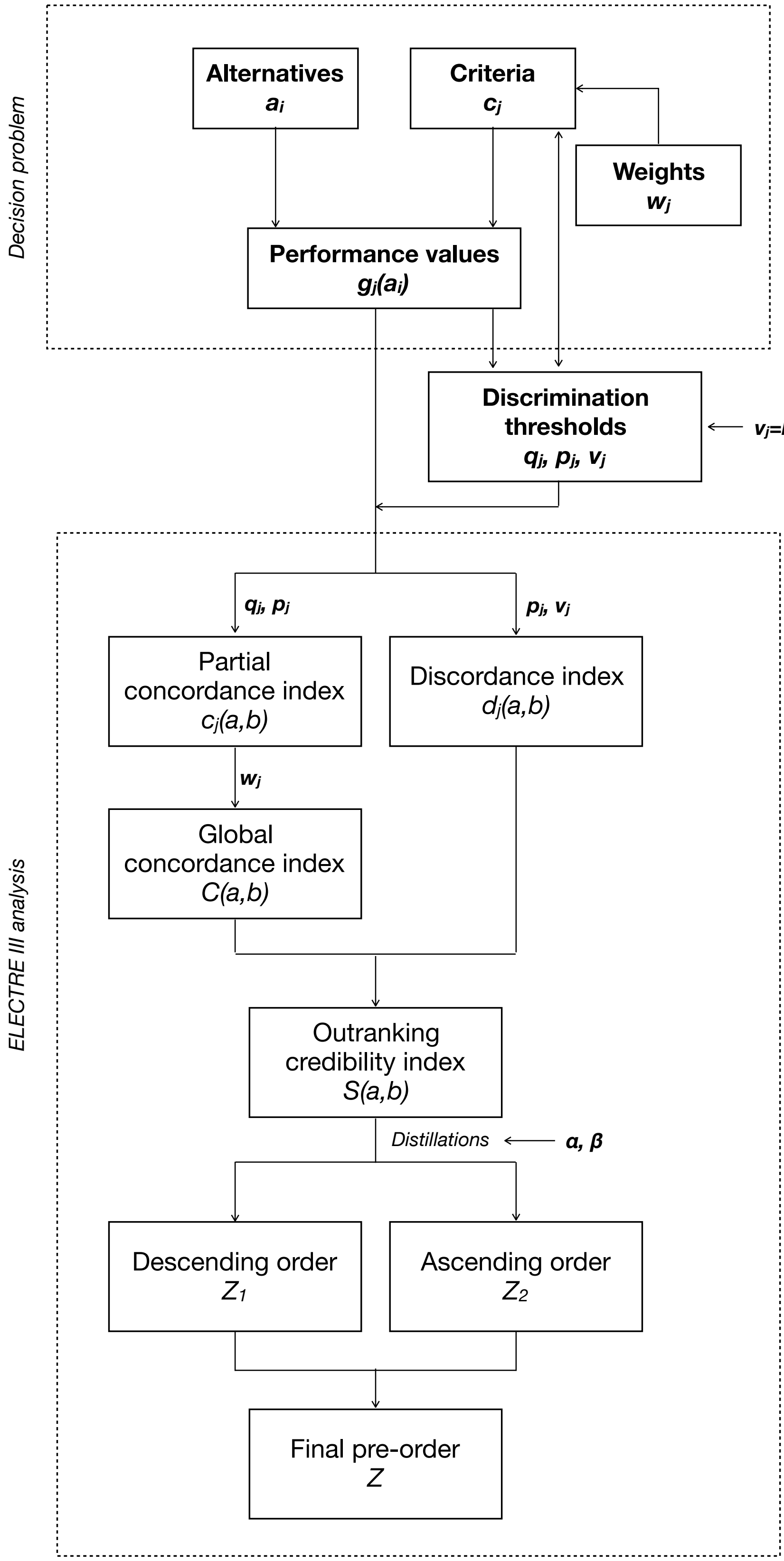

Figure 1.2: Overview of ELECTRE III method. The alternatives a_i, criteria c_j, weights w_j, and performance values $g_j(a_i)$ constitute the decision problem. The discrimination thresholds q_j, p_j, v_j are used to calculate credibility index for each pair of alternatives a, b. These credibility indexes are then exploited to obtain the final pre-order Z.

1.4.2 Our modified ELECTRE III

Some assumptions of our proposed approach are defined as follows:

- Performance scores of alternatives on the different criteria, $g_j(a_i)$, are not unique values, but a range of values defined by lower and upper bounds, $g_j(a_i)_L$ and $g_j(a_i)_U$, that represent the set of possible values for that alternative. These bounds can be either obtained from decision-makers or available data.

- The definition of the indifference and preference thresholds, q_j and p_j, can be linked to the uncertainty associated with the criterion in question. In this case, the indifference threshold is interpreted as the minimum margin of uncertainty in the performance values, while the preference threshold is taken as the maximum margin of uncertainty.

- The definition of the veto threshold, v_j, cannot be associated to the uncertainty, but to the conditions under which a discordant criterion can exert a veto on an outranking relationship. Nonetheless, the size of the veto threshold is generally fixed based on the preference threshold (Roy et al. 1986). The further v_j is from p_j (i.e., the bigger k is), the less the veto threshold will affect the outranking of one alternative over another. Therefore, this veto to preference ratio, $k = v_j/p_j$, needs to be specified. In our case, since we do not have access to the experts involved in the decision, we test different values of k to evaluate the impact of the ratio of veto and preference thresholds on the final rank orders.

Therefore, a set of discrimination thresholds q-p-v can be defined for each pair of alternatives a and b, and criterion j, as follows:

$$q_j(a,b) = \max\left(|g_j(a) - g_j(a)_U|, |g_j(b) - g_j(b)_L|\right) \tag{1.12}$$

$$p_j(a,b) = |g_j(a) - g_j(a)_U| + |g_j(b) - g_j(b)_L| \tag{1.13}$$

$$v_j(a,b) = k \cdot p_j(a,b) \tag{1.14}$$

Note that, if $g_j(a)$ is greater than $g_j(b)$, then the bounds considered when calculating the discrimination thresholds are switched.

The proposed approach is illustrated in Figure 1.3. As seen, we consider two alternatives a and b to be indifferent for the j^{th} criterion when their performance scores are included in each others' intervals. If the two performance intervals overlap, but this intersection does not span the average performance score of the other alternative, we weakly prefer b over a. If, on the other hand, the two performance intervals do not overlap, we strictly prefer b. When the difference of the evaluation between the two alternatives is greater than our veto threshold, we also reject that alternative a could outrank b for the rest of criteria.

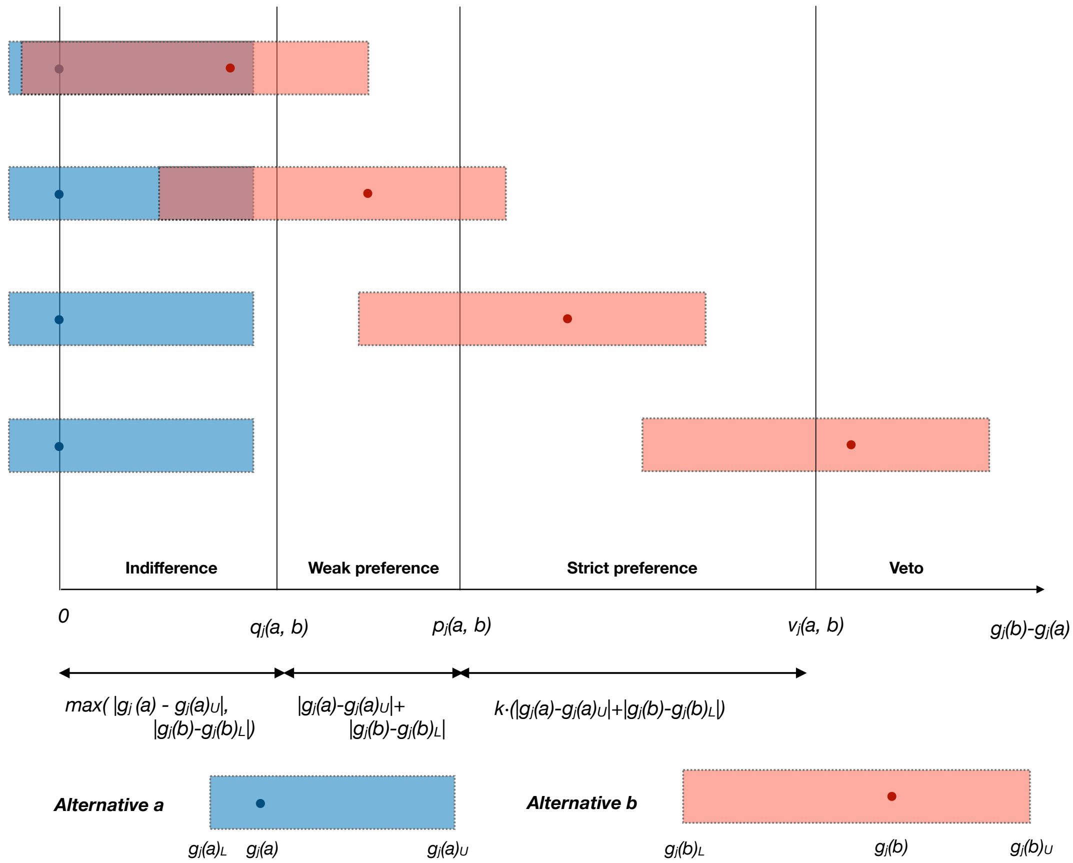

Figure 1.3: Proposed approach for incorporating uncertainty of performance scores. a and b are the pair of alternatives being compared for the j^{th} criterion, $g_j(*)$ is the performance score of the alternative, $g_j(*)_L$ and $g_j(*)_U$ are its lower and upper bounds, respectively, and k is the veto to preference ratio selected.

1.5 Results and Discussion

In this section, we first present and discuss the results of applying ELECTRE III with our proposed approach, highlighting the impact of the veto to preference ratio (k) on the rank order of renewable energies (1.5.1). Then we compare the results to those obtained by Mousavi et al. (2017), Erdogan and Kaya (2015) and Kahraman and Kaya (2010) (1.5.2). Finally, we analyze and discuss the implications of these rankings for energy planning and policy in Turkey (1.5.3).

1.5.1 Ranking of renewable energy alternatives

As explained in Section 1.4.2, in our proposed approach indifference and preference thresholds are directly linked to the uncertainty in the performance scores of the different alternatives. However, the veto threshold cannot be associated to the uncertainty but rather to the limit of tolerance that the actors involved in the decision are willing to accept for any compensation between criteria (i.e., a decrease of one criterion can be compensated by an equivalent gain on any other criterion). Therefore, veto thresholds must be elicited from them (in this case the three energy experts involved in the decision analysis process). Without access to the energy experts, we had to test different veto to preference ratios to study how the evaluation of renewable energy alternatives changed.

As we can see in Figure 1.4, for veto to preference ratios bellow 5, all renewable energy alternatives occupy the same rank order. This is because low veto thresholds, in conjunction with the overlapping scores of energy alternatives, lead to more possibility of incomparability between pairs of energy alternatives. For bigger k, hydropower is always considered the least appropriate option, preceded by biomass. The case of geothermal power is noteworthy: with a k of 6, it is the best energy alternative, but takes the second and third place for higher k. On the other hand, both wind and solar energies rank the best: for k between 7 and 10, wind is preferred over solar power, while for k values higher than 10, both options occupy the first rank. Interestingly, in the case of $k \geq 10$, the ranking of renewable energy alternatives stabilizes. These results highlight the impact of the veto concept: increasing the veto thresholds leads to a reduction in the concordance part as a veto effect becomes more difficult to observe.

However, it is important to point out that this does not imply a low robustness of our approach. Indeed, Roy et al. (1986) emphasized that veto thresholds are merely numbers that represent the decision-makers' deliberate policy decisions. There is hence no "correct" value for the veto threshold (nor the veto to preference ratio, k), and all rankings are equally valid depending on the decision-makers' preferences and judgments. Therefore, for our approach to be accurately reflect decision-makers' actual values and opinions, the veto to preference ratio should also be elicited from them, by means of trade-off questions for example.

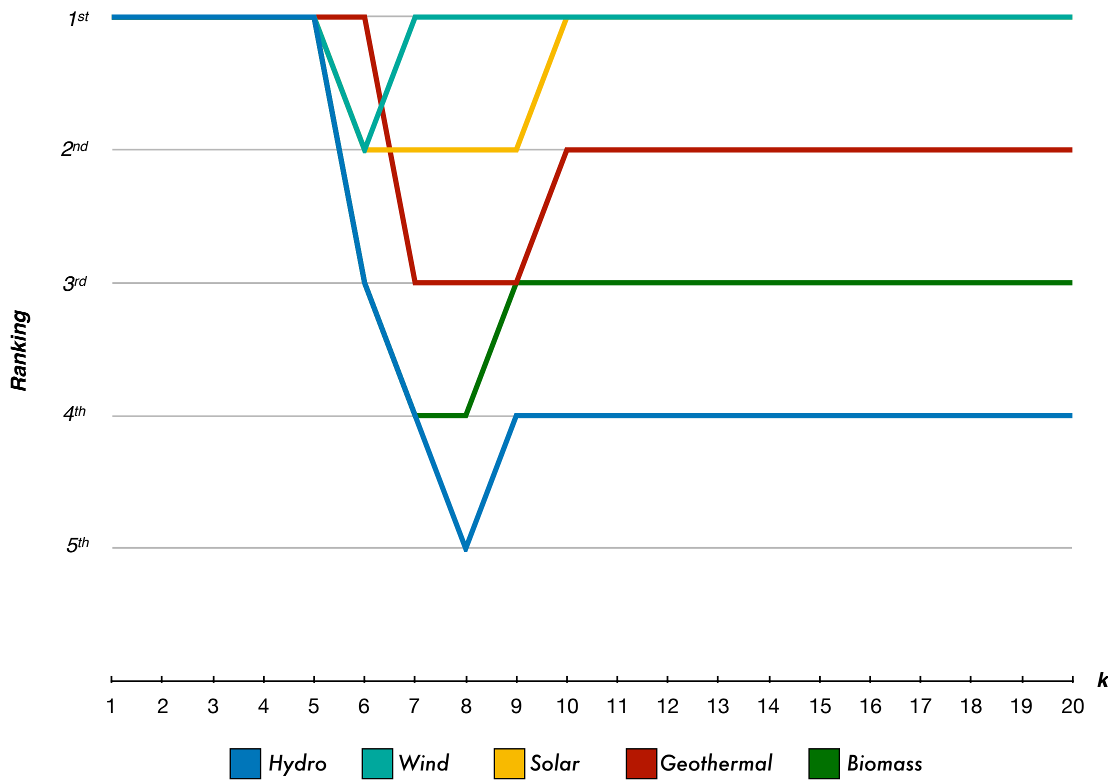

Figure 1.4: Influence of the veto to preference ratio, k, on the rankings of the five renewable energy alternatives evaluated (hydro, wind, solar, geothermal and biomass).

1.5.2 Comparison with other uncertainty approaches

The ranking of energy alternatives obtained by Mousavi et al. (2017), Erdogan and Kaya (2015) and Kahraman and Kaya (2010) is: *Wind* > *Solar* > *Geothermal* > *Biomass* > *Hydro*; which is the same result that we obtain with a veto to preference ratio of 8 ($k = 8$).

However, although the comparative analysis show that the ranking results obtained from our proposed methods and those by Mousavi et al. (2017), Erdogan and Kaya (2015) and Kahraman and Kaya (2010) are the same, the methods differ significantly in the way in which uncertainty is considered. As mentioned previously in Section 1.2.1, Mousavi et al. (2017), Erdogan and Kaya (2015) and Kahraman and Kaya (2010) use fuzzy set theory to cope with uncertainty in performance scores. Kahraman and Kaya (2010) apply an AHP model with ordinary fuzzy set theory. They use a trapezoidal membership function for each criterion to convert the experts' crisp judgments into a fuzzy value between 0 and 1. Mousavi et al. (2017) employed ELECTRE under a hesitant fuzzy set approach, in which the performance values of the decision matrix are represented as hesitant fuzzy elements. They use hesitant fuzzy sets, which are an extension of ordinary fuzzy sets, to allow for the selection of multiple membership functions under each criterion. Erdogan and Kaya (2015) also propose a similar approach, now based on interval type-2 fuzzy set, for the TOPSIS model. In this case, the membership functions also have an uncertainty associated with it, described by its two bounding functions (i.e., lower and upper membership functions).

The addition of fuzzy logic in decision analysis can significantly increase the difficulty of implementation and the complexity of the MCDA model (Hanratty and Joseph 1992). First and foremost because the actors involved in the decision – particularly decision analysts – must invest time learning about fuzzy set theory and the mathematical definitions behind it in order to apply it in MCDA under uncertainty. Second, because the arithmetic of fuzzy sets require much more time – especially during the definition of the membership functions – and calculation steps, which consequently limits their application. Our approach, on the other hand, does not require additional knowledge apart from understanding the ELECTRE III model.

By considering performance scores as a range of values defined by lower and upper bounds, and using these bounds to determine the model's discrimination thresholds, the process of incorporating uncertainty to the MCDA problem becomes less complex. Nonetheless, three key limitations must be acknowledged:

- It can only incorporate the uncertainty in scoring the alternatives in MCDA. For other uncertainties (e.g., in weighting) other approaches must be used.
- It is only suitable for decision problems where the performance scores can be expressed as a range of values (with lower and upper bounds).
- It requires decision-makers (and other actors involved in the decision) to define the veto to preference ratio (i.e., parameter k). If it cannot be defined, it is necessary to assess the extent of the influence of parameter k (i.e. the ratio between the veto and the preference thresholds) on the final rankings.

1.5.3 Policy implications for the Turkish energy sector

According to the results of our analysis, wind energy is found out to be the best energy alternative for Turkey, followed by solar, geothermal, biomass and hydro. This is largely in line with the national renewable energy action plan: most of the new renewable capacity to be added in the next four years is planned to come from wind energy (68.5%), followed by solar and hydro (13.7% each) and geothermal and biomass (2.05% each). Given the primary role of wind energy in Turkey's 2023 energy strategy, it is important that Turkish policy-makers provide measures to promote wind energy development.

The Turkish government has already put various renewable support schemes to increase the attractiveness of investments in renewable energy projects, notably in wind power. Among these is the creation of the Renewable Energy Resources Support Mechanisms (YEKDEM in Turkish) in 2011 (Kaplan 2015). Under the YEKDEM scheme, wind power producers can apply for favourable feed-in-tariffs which guarantees a sale price of 7.3 US dollar cents per kWh for the first 10 years of operation (plus a local-content bonus ranging from 0.6 and 1.3 US dollar cents per kWh if the plant components are produced in Turkey) (International Energy Agency

2011). In addition, the new Electricity Market Law also introduces an exemption from licensing for wind generation facilities with a capacity below 1 MW (or up to 5 MW if authorized by the Council of Ministers) (Kaydul et al. 2017). Another important means of support is the Renewable Energy Resources Area (YEKA) program, under which the government offers an advantageous auction policy model for investors committed to develop local manufacturing and research capacity, and employ a large share of domestic workers (Tagliapietra et al. 2019).

Notwithstanding these wind power support mechanisms, which have indeed boosted investment in wind energy, there remains a number of bottlenecks that must be addressed in order to achieve the 2023 target (International Energy Agency 2016; Ozcan 2014, 2018), including:

- Constraints in the transmission and distribution networks that hinder the connection of wind power capacity. It is important that policy-makers start devising grid integration strategies to ensure the interconnection and reliable operation of wind power.

- Administrative hurdles that lead to lengthy permitting and spatial planning processes. More efforts are thus needed to improve bureaucratic efficiency in licensing and granting permits.

- Lack of knowledge and training on wind power, in particular on strategic planning of the whole wind energy system. Training and capacity-building programs are necessary to keep energy stakeholders updated with the respect to the legal, commercial and legal changes in the road map of wind energy.

1.6 Key messages

Prioritizing and assessing policy alternatives under uncertainty is regarded as a complex decision making process, in particular when decision- and policy-makers are not familiarized with sophisticated uncertainty approaches such as fuzzy set theory. To cope with this issue, we have presented a modified ELECTRE III model for dealing with the uncertainty arising from the scoring of alternatives.

In our proposed approach, the uncertainty in performance scores is directly incorporated when the model's discrimination thresholds are established, by expressing them as a function of scores bounds. We have illustrated the validity and suitability of our approach by applying it to a real case study on renewable energy policy selection in Turkey, and by comparing it with other approaches developed in the literature.

Some key messages can be highlighted from this Chapter:

- The approach requires that actors involved in the decision elicit their preferences by specifying a range of performance values for each alternative (i.e., lower and upper bounds of performance scores). It also requires them to specify the veto to preference ratio (parameter k in our model).

- Unlike fuzzy set-based approaches, where actors need to be familiar with fuzzy set theory, the proposed approach does not require any additional knowledge apart from ELECTRE III. This makes it easier to understand and be applied in decision analysis.

- In the specific case study of Turkey, the ranking of energy alternatives are determined as: *Wind, Solar, Geothermal, Biomass and Hydro.* These results agree with the rankings obtained with other methods available in the literature. In addition, the ranking that our method produces is aligned with Turkey's national renewable energy action plan.

- Given the primary role of wind energy in Turkey's 2023 energy strategy, it is crucial that decision-makers address some current constraints for wind development, mainly ensuring grid integration, bureaucratic efficiency and capacity-building of energy stakeholders.

2

Targeting and prioritization in the WASH sector

Abstract:

In the era of *leaving no one behind*, where one of the most pressing issues is to provide universal access to safe Water, Sanitation and Hygiene (WASH) services, it is imperative to target and prioritize the most disadvantaged. Multi-Criteria Decision Analysis (MCDA) models can play a key role in informing resource allocation for effective WASH planning. However, data uncertainty – intrinsic to the available data collection tools used in the sector – must be accounted for in the decision analysis process in order to avoid misleading conclusions. This Chapter presents two data-driven MCDA models, based on compensatory and non-compensatory aggregation techniques, to incorporate data uncertainty. We use WASH planning in rural Kenya as a case study to illustrate and compare the two models. The comparison shows that, while both approaches integrate data uncertainty in a considerably different manner, they lead to similar prioritization of districts in Kenya.

Keywords:
Multi-Criteria Decision Analysis; Water, Sanitation and Hygiene; SDG 6; Household Surveys

This Chapter is based on:

- Ezbakhe, F. and Pérez-Foguet, A., 2018. Multi-Criteria decision analysis under uncertainty: two approaches to incorporating data uncertainty into water, sanitation and hygiene planning. *Water Resources Management*, 32(15), 5169—5182. doi: 10.1007/s11269-018-2152-9
- Ezbakhe, F. and Pérez-Foguet, A., 2018. Embracing data uncertainty in water decision-making: an application to evaluate water supply and sewerage in Spain. *Water Supply*, 19(3), 778—788. doi: 10.2166/ws.2018.122
- Ezbakhe, F. and Pérez-Foguet, A., 2017. Considering data uncertainty in the water and sanitation sector: Application to large number of alternatives and criteria. *European Water*, 57, 215–222.

2.1 Introduction

Achieving universal access to safe water, sanitation and hygiene (WASH) services by 2030 is a vast endeavour for the global community (UN-Water 2018). Targets 6.1 and 6.2 of the Sustainable Development Goals challenge governments to tackle the *unfinished business* of extending WASH services to those who remain unserved, as well as to progressively improve the level of services provided. The progressive realization of universal access to WASH and the reduction of inequalities in service levels is also consistent with the UN's resolution on the human rights to water and sanitation (UNGA 2010). However, the commitment to *leave no one behind* will demand increased attention on those most in need of better WASH services, and deliberate efforts to reduce and eliminate inequalities. As the former Special Rapporteur on the human rights to safe drinking water and sanitation declared, governments must give "priority to realizing a basic level of service for everyone before improving service levels for those already served" (OHCHR 2011).

Targeting and prioritization are thus central to WASH planning. By identifying under-served and disadvantages geographical areas and social groups, governments can allocate resources more equitably, and altogether implement more effective policies for poverty alleviation (Giné-Garriga et al. 2015). However, improved priority-setting and targeting remains a difficult task for WASH policy-makers. Indeed, inadequate targeting of resources towards those without access to WASH services is one of the three main challenges in the WASH sector, next to the lack of finance for strengthening service delivery and the enabling environment, and the slow implementation of integrated water resources management (UN-Water 2018). Other studies reported inefficient resource allocation as one of the major weaknesses constraining the overall performance of the sector (Gutierrez 2007; Jiménez and Pérez-Foguet 2010; Cairncross et al. 2010; De Palencia and Pérez-Foguet 2011; Wayland 2019; Giné-Garriga and Pérez-Foguet 2019). For this reason, a step forward in the WASH sector would be to establish appropriate decision analysis tools that assist policy-makers in targeting and prioritizing the most vulnerable communities.

Multi-Criteria Decision Analysis (MCDA) models can play an important role in informing WASH planning. By evaluating and ranking population groups against multiple criteria – be it service coverage, quality, availability or capacity –, MCDA models can provide insight on priority-setting of WASH interventions for various reasons. First, they steer interventions that tackle multiple dimensions affecting access to WASH services. Second, contrary to ad-hoc priority-setting, in which policy-makers tend to use rather intuitive processing of complex data, MCDA enables a more transparent and systematic assessment of evidence. Furthermore, MCDA can enhance stakeholder participation, which contributes not only to a better framing of policy issues but also to a higher chance to adopt the solutions.

A wide variety of MCDA models exist today, and can be grouped in two main approaches (Ishizaka and Nemery 2013): (i) value measurement models (*American school*), based on the construction of a numerical score for each alternative, and (ii) outranking models (*European school*), built on the pairwise comparison between alternatives. The differences between these two MCDA families are substantial. For instance, there is no underlying utility function in outranking models: the output is a ranking of alternatives without any scores to indicate the extent to which one alternative is preferred to another. In addition, the set of decision rules describing the aggregation procedure in outranking models are only partially compensatory, which limits the trade-offs between the different criteria (Stewart and Losa 2003).

Despite these considerable disparities, only few studies have discussed the relative merits on similar decision problems between these two main MCDA models. As several authors dealing with this issue have emphasized (Ozernoy 1992; Olson et al. 1995; Zanakis et al. 1998; Gavurova et al. 2017; Motahari Farimani 2018), it is intricate to compare MCDA models when it is not even possible to define a clear, objectively "best" decision in a multiple attribute environment. It is indeed not straightforward to answer questions such as: which model is more appropriate for the type of decision problem in hand? Does the decision change when using a different model? And, most importantly, how differently can they incorporate uncertainty?

Against this background, we present and compare two MCDA models, based on Multi-Attribute Utility Theory (MAUT, compensatory) and ELimination and Choice Translating REality (ELECTRE III, non-compensatory), to examine how they would integrate data uncertainty into the decision analysis process.

In particular:

- We extend the MAUT and ELECTRE III models to tackle the uncertainty in the input data. Specifically, we deal with the statistical uncertainty present in the data used in the decision analysis process.

- We apply both extended models to the case of WASH planning in Kenya. We use data from household surveys to assess the WASH status of rural communities in 21 counties and prioritize them based on 9 different criteria, which are related to the availability, quality and accessibility of services.

- We assess the convergence between the counties rankings from each model.

- Finally, we analyze the implications of the rankings for Kenya's WASH sector.

We structure the remainder of the Chapter as follows. In Section 2.2, we provide an overview of compensatory and non-compensatory MCDA models. Then, we present the case of WASH planning in rural Kenya and describe our decision problem in Section 2.3. In Section 2.4, we briefly introduce our extensions of MAUT and ELECTRE III models, drawing special attention to the way each incorporates data uncertainty into the analysis process. Our results are presented and discussed in Section 2.4, in particular the comparison between the two models and the policy implications. Finally, in Section 2.5 we highlight the key messages of our work.

2.2 A background on compensation in MCDA

In MCDA, weights are assigned to the decision criteria to represent their relative importance. This recognizes that not all criteria have the same importance to the actors involved in the decision. Weights are also used to combine the alternatives' performances for each individual criterion and obtain their overall performance. A wide variety of techniques are available to undertake this aggregration, and can be split into two categories: compensatory and non-compensatory (Jeffreys 2004).

The distinction between compensatory and non-compensatory techniques is based on the trade-offs between evaluation criteria. In compensatory approaches, poor performance of an alternative in some criteria can be "compensated" for by high performance in other criteria. In other words, a good aggregated performance may be the result of a very good value in some criteria which masks potentially critical values for other criteria. Some examples of compensatory techniques include: simple additive weighting (SAW), technique for order of preference by similarity to ideal solution (TOPSIS), analytic hierarchical processes (AHP), and multi-attribute utility or value functions (MAUT or MAVT). In contrast, in non-compensatory approaches, poor performances of an alternative in some criteria cannot be compensated for even with high performances in other criteria, and thus it will be reflected its aggregated performance. ELimination and Choice Translating REality (ELECTRE) is perhaps the best example of this approach.

Choosing which aggregation mechanism is a non-trivial decision, as each makes a number of assumptions regarding the decision-makers' preferences that are not always easily identifiable (Ouerdane 2011). One way to select a particular aggregation procedures is the Conjoint Measurement Theory, which "examines the conditions under which a relation on a set of alternatives – described by a vector of evaluations – is determined by a sort of synthetic measurement that takes relevant attributes of the objects into account in an appropriate manner" (Krantz et al. 1971). In other words, this theory seeks to identify a system of axioms to represent the conditions (or preferences) under which a procedure can be used. During the decision analysis process, these axioms can help the analyst and decision-makers' to choose an aggregation mechanism that fits the preferences of the latter. In practice, it is not straightforward to rely on such axioms to make the choice. As Ouerdane (2011) explains, axioms can be overly abstract, hard to test and, very often, sufficient but not necessary.

That is why the strategy generally followed to select between compensatory and non-compensatory models consists in applying both and compare the rankings. Table 2.1 shows an example of studies from the water sector following such strategy. In general, compensatory approaches are considered easier to understand and implement, but their outcomes depend highly on the weights of dominant criteria; a fact that becomes their main drawback.

Table 2.1: Summary of previous studies from the water sector comparing different MCDA models.

Study	MCDA models compared	Decision problem	Results of the comparison
Duckstein et al. (1982)	ELECTRE III; MAUT; CP	River basin planning: selection of the best alternative for flood control in Tucson Basin (USA)	There is no significant difference in ranking across models, but MAUT is more time consuming than the others
Goicoechea et al. (1992)	MATS-PC; ARIADNE; EXPERT CHOICE; ELECTRE III	Water resources planning: ranking of water supply project plans in Washington Metropolitan Area (USA)	There is no significant difference in ranking across models
Mahmoud and Garcia (2000)	MAUT; PROMETHEE II; AHP; ELECTRE III	Water resources management: evaluation of operation alternatives of the Red Bluff Diversion Dam (USA)	There is no significant difference in ranking across models, but some models (MAUT) are easier to understand and require little interaction from the decision-makers
Kangas et al. (2001)	ELECTRE III; PROMETHEE-II	Strategic natural resources planning: ranking of forest strategies in Kainuu (Finland)	The models give somewhat different rankings
Chitsaz and Banihabib (2015)	ELECTRE I; ELECTRE III; MAUT; AHP; TOPSIS	Flood management: prioritization of flood management alternatives in Golestan (Iran)	There is no significant difference in ranking across models, but they present different sensitivity to changes in criteria weights
Sikder and Salehin (2015)	MAUT; AHP	Rural water supply planning: selection of the best technology for water supply in Bangladesh	There is no significant difference in ranking across models
Banihabib et al. (2017)	MAUT; AHP; ELECTRE III	Water resources strategic management: ranking of flood management alternatives in Shahrood (Iran)	The models give somewhat different rankings and ELECTRE III's results are considered the most reliable due to their low dependency to criteria weights

2.3 The case of WASH planning in Kenya

In this section, we first present a general overview of Kenya's WASH sector (2.3.1). We then describe in detail the decision problem we analyze (2.3.2), which consists in selecting the rural counties most in need for better WASH services.

2.3.1 An overview of Kenya's WASH sector

The architecture of the Kenya's WASH sector has undergone significant changes in the previous decades. Since the introduction of the new Water Act in 2002, significant policy revision and restructuring has been undertaken to enhance access to safe and improved WASH services[1][2]. One of the major outcomes of this process culminated in the constitutional recognition of access to WASH as a basic human right. According to Article 31 of the Constitution, every Kenyan is entitled to "clean and safe water in adequate quantities" as well as "reasonable standard of sanitation" (Republic of Kenya 2010). This legal commitment has been further affirmed in recent policies such as the Kenya Vision 2030, the National Environmental Sanitation and Hygiene Policy (2015-2030), the National Open Defecation Free Kenya (2030) and the National Water Master Plan (2030) (Mansour et al. 2017).

However, in spite of the constitutional right to WASH and the political push, access to WASH services remains a challenge. With a population of almost 50 million, nearly 42% of Kenyans still rely on unimproved water sources and 70% use unimproved sanitation facilities (JMP 2017a). These figures are even more alarming for rural areas, where improved coverage lags behind the national average (i.e., only 37% and 36% of the population has access to improved water and sanitation, respectively). The inadequate WASH has been linked to negative health outcomes, especially amongst children (Garrett et al. 2008; du Preez et al. 2011; Darvesh et al. 2017).

[1] Improved drinking water sources are those that have the potential to deliver safe water by nature of their design and construction, and include: piped water, boreholes, protected dug wells, protected springs, rainwater, and bottled water.

[2] Improved sanitation facilities are those designed to hygienically separate excreta from human contact, and include: flush/pour flush to piped sewer system, septic tanks or pit latrines, ventilated improved pit latrines, composting toilets, and pit latrines with slabs.

Amid the many causes of this low coverage, ineffective resource allocation has been identified as a priority concern in Kenya (Kenya National Commission on Human Rights 2014; Mansour et al. 2017; Development Indicatives 2018). This ineffectiveness stems from two main issues. First, financing of the WASH sector is insufficient. For instance, 15% of the 2019-2020 budget (i.e., 3,650 out of total 23,645 million Kenyan shillings) will be allocated to the water and sanitation sector, of which no more than 27% will be dedicated to water and sewerage infrastructure development (Republic of Kenya 2019). A closer look at the budget allocations at the local level reveals that most 13 out of the 47 counties do not have any budget for water or sanitation (Kenya National Commission on Human Rights 2014). Added to this insufficient financing is the significant institutional fragmentation and overlap of competences. For instance, at the national level, policy development for the sector is shared between three ministries (the Ministry of Water and Irrigation, the Ministry of Environment and Natural Resources, and the Ministry of Health); at the local level, responsibility for water and sanitation service provision is in the hand of each of the 47 counties, and their sub-counties are responsible for managing implementation. The lack of coordination between the different institutions leads to duplication of roles and inefficiencies that undermine efforts for better WASH services.

Furthermore, the institutional capacity of the sector falls short of its requirements. The capacity of county governments, for example, is insufficient for an effective planning and allocation of resources, both in terms of quantity and *skillfulness* of public employees, and there is no explicit initiative from the national government to fill this gap (Mansour et al. 2017). Another important problem lies in the poor data collection and analysis at both national and county levels (Development Indicatives 2018), which hampers the understanding of the coverage levels and the cost of prioritizing services in under-served areas. Consequently, establishing appropriate data-driven planning tools will surely result in a positive impact on the WASH sector.

2.3.2 The decision problem

The decision problem, inspired by Giné Garriga and Pérez Foguet (2013b), consists in the evaluation and prioritization of areas most in need for better WASH services.

The 21 counties considered in this analysis (Figure 2.1) have been previously identified by the Kenyan government as those with the most vulnerable rural populations.

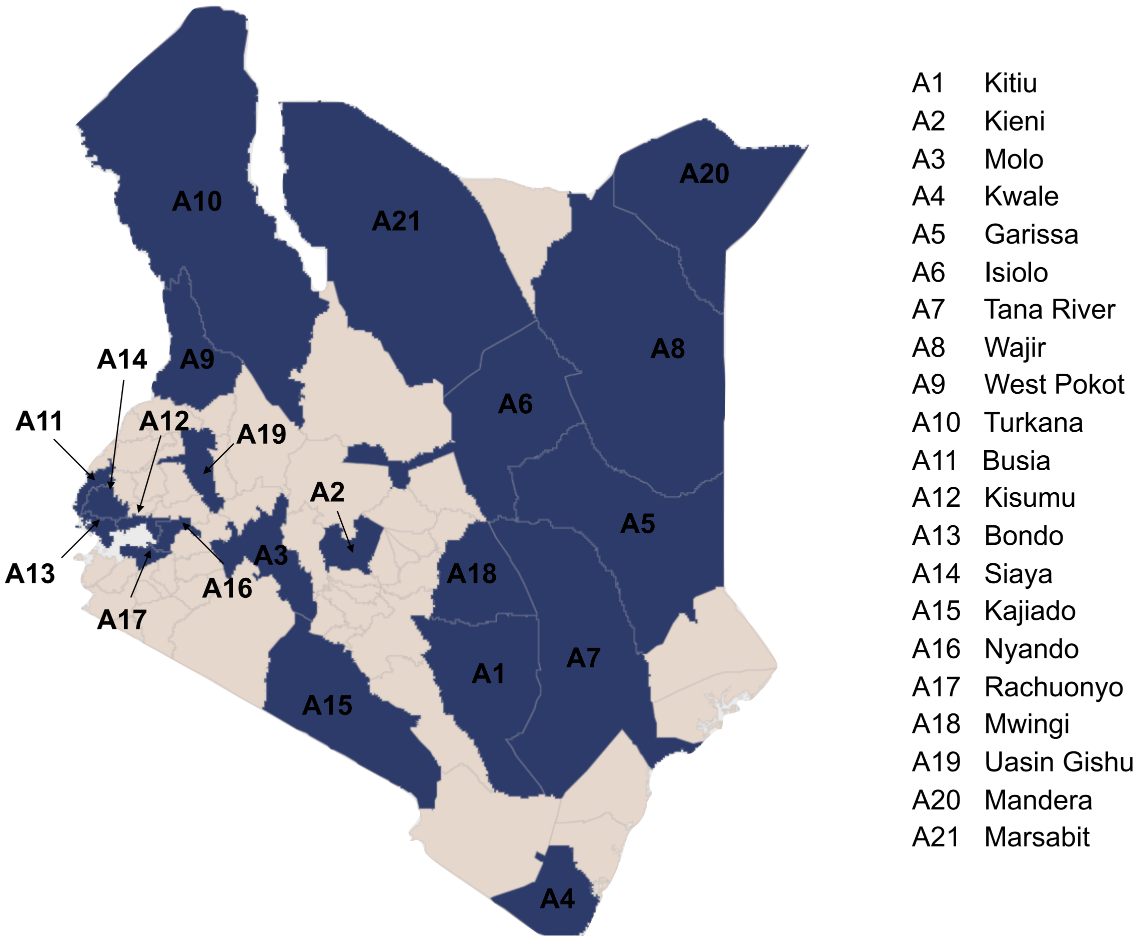

Figure 2.1: Map of Kenya with the 21 counties considered in the decision problem.

Data on the WASH coverage levels in these 21 counties was collected through household surveys. A total of 4,925 households were selected, and multiple WASH-related issues were covered in each one of them, including: (i) quality of the water source, (ii) type of main drinking water source, (iii) distance from dwelling to the water source, (iv) functionality of water supply in the household, (v) person responsible for dwelling water, (vi) domestic water consumption, (vii) type of sanitation facilities, (viii) sanitary inspection of water supplies, and (iv) point-of-use water treatment. The standards (or minimum levels) for these issues are shown in Table 2.2.

Table 2.2: WASH criteria considered in the decision problem, with the standard required.

c_j	Criteria	Standard (or minimum level)
c_1	Quality of the water	Water with good analysis results
c_2	Type of main drinking water source	Access to improved water sources
c_3	Distance from dwellings to water source	Less than 30 minutes spent in water fetching
c_4	Functioning water supply	Source is in an operational condition
c_5	Person responsible for dwelling water	Person responsible is not a child
c_6	Domestic water consumption	More than 20 liters of per capita per day
c_7	Type of sanitation facilities	Access to improved sanitation facilities
c_8	Sanitary inspection of water supplies	No identified risk of contamination
c_9	Household water treatment	Adequate treatment technology used

Each household was given a value of 0 or 1 depending on whether it met the standard (1) or not (0). This provided the proportion of households that met the standards. The survey data is shown in Table 2.4. These data constituted the performance scores in our MCDA models.

A data-driven approach, based on Principal Component Analysis (PCA)[1] (Nardo et al. 2005), was used to derive the weights of criteria (shown in Table 2.3), as done in other WASH-related studies (Giné Garriga and Pérez Foguet 2010, 2013a; Pérez-Foguet and Giné-Garriga 2011). The idea behind it is to account for the highest possible variation in the data set using the smallest possible number of criteria. As a result, weights are no longer a measure of importance of the associated criterion but a way to correct for the overlapping information of two or more criteria.

Table 2.3: Criteria weights used in both MCDA models (obtained from a Principal Component Analysis).

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
Weights	0.152	0.160	0.101	0.054	0.148	0.052	0.073	0.112	0.147

[1]Principal Components Analysis groups together individual variables which are collinear to form linearly uncorrelated variables – the principal components – that account for as much of the variability in the data as possible.

Table 2.4: Results of the household surveys in the 21 counties of Kenya. n is the sample size, and x_i and p_i the number and proportion of households meeting the standard for each i^{th} criterion.

County	n	c_1		c_2		c_3		c_4		c_5		c_6		c_7		c_8		c_9	
		x_1	p_1	x_2	p_2	x_3	p_3	x_4	p_4	x_5	p_5	x_6	p_6	x_7	p_7	x_8	p_8	x_9	p_9
A1	247	193	0.781	83	0.336	51	0.206	232	0.939	237	0.960	154	0.623	159	0.644	175	0.709	160	0.648
A2	195	167	0.856	9	0.046	159	0.815	176	0.903	194	0.995	169	0.867	169	0.867	105	0.538	134	0.687
A3	186	131	0.704	93	0.500	110	0.591	184	0.989	178	0.957	118	0.634	156	0.839	132	0.710	92	0.495
A4	238	177	0.744	133	0.559	143	0.601	210	0.882	228	0.958	140	0.588	72	0.303	80	0.336	53	0.223
A5	209	107	0.512	61	0.292	106	0.507	173	0.828	203	0.971	83	0.397	27	0.129	75	0.359	60	0.287
A6	230	159	0.691	119	0.517	85	0.370	162	0.704	227	0.987	90	0.391	50	0.217	45	0.196	47	0.204
A7	224	146	0.652	104	0.464	44	0.196	191	0.853	218	0.973	63	0.281	106	0.473	26	0.116	58	0.259
A8	230	188	0.817	25	0.109	20	0.087	230	1.000	223	0.970	120	0.522	149	0.648	125	0.543	108	0.470
A9	429	338	0.788	123	0.287	63	0.147	378	0.881	387	0.902	210	0.490	211	0.492	280	0.653	201	0.469
A10	240	173	0.721	160	0.667	116	0.483	238	0.992	214	0.892	97	0.404	77	0.321	124	0.517	162	0.675
A11	236	191	0.809	101	0.428	4	0.017	157	0.665	207	0.877	114	0.483	57	0.242	157	0.665	22	0.093
A12	218	157	0.720	169	0.775	113	0.518	198	0.908	205	0.940	155	0.711	100	0.459	159	0.729	138	0.633
A13	246	132	0.537	52	0.211	109	0.443	242	0.984	233	0.947	129	0.524	118	0.480	218	0.886	153	0.622
A14	244	205	0.840	130	0.533	74	0.303	226	0.926	219	0.898	150	0.615	161	0.660	176	0.721	140	0.574
A15	242	134	0.554	61	0.252	88	0.364	223	0.921	220	0.909	122	0.504	107	0.442	182	0.752	167	0.690
A16	249	203	0.815	167	0.671	134	0.538	237	0.952	241	0.968	166	0.667	135	0.542	207	0.831	190	0.763
A17	230	190	0.826	199	0.865	97	0.422	124	0.539	227	0.987	114	0.496	108	0.470	95	0.413	87	0.378
A18	128	71	0.555	51	0.398	54	0.422	102	0.797	125	0.977	27	0.211	5	0.039	14	0.109	22	0.172
A19	229	177	0.773	72	0.314	176	0.769	218	0.952	227	0.991	116	0.507	103	0.450	182	0.795	160	0.699
A20	240	195	0.812	41	0.171	60	0.250	224	0.933	233	0.971	63	0.263	70	0.292	89	0.371	16	0.067
A21	235	168	0.715	159	0.677	114	0.485	210	0.894	230	0.979	116	0.494	151	0.643	61	0.260	151	0.643

2.4 Methodology

We first explain how we characterize the uncertainty in the input data (2.4.1). Then, we provide an overview of MAUT and ELECTRE III models, and describe the extensions we have developed to incorporate the uncertainty of the data (2.4.2).

2.4.1 Characterization of data uncertainty

We use confidence intervals to characterize uncertainty associated with the input data. Since these data are proportions of populations estimated from the household surveys, we consider that they follow a binomial probability distribution[1], $B(n, p)$. We parametrize the the number of households surveyed in each county (n) and the proportion of households verifying the minimum levels required for each criterion (p).

The lower and upper limits of the confidence interval are calculated according to Clopper and Pearson (1934) *exact* method:

$$p_{i_L} = \left[1 + \frac{n - x_i + 1}{x_i \cdot F_{1-\alpha/2,\ 2x_i,\ 2(n-x_i+1)}}\right]^{-1} \tag{2.1}$$

$$p_{i_U} = \left[1 + \frac{n - x_i}{(x_i + 1) \cdot F_{\alpha/2,\ 2(x_i+1),\ 2(n-x_i)}}\right]^{-1} \tag{2.2}$$

where p_{i_L} and p_{i_U} are the lower and upper bounds of the $100(1-\alpha)\%$ confidence interval for the i^{th} criterion, n is the number of households surveyed, x_i is the number of households that meet the required standard, and $F(c, df_1, df_2)$ is the $1-c$ quantile from the F distribution with degrees of freedom df_1 and df_2.

While there are other formulas to calculate the confidence intervals of binomial proportions, we choose the Clopper-Pearson interval because it is based on the cumulative probabilities of the binomial distribution rather than an approximation to the normal distribution[2]. The confidence intervals are shown in Figure 2.2.

[1] We assume that sample sizes are much smaller than the population size.

[2] Clopper-Pearson intervals are more exact than those obtained with any approximation to the binomial distribution (Agresti and Coull 1998).

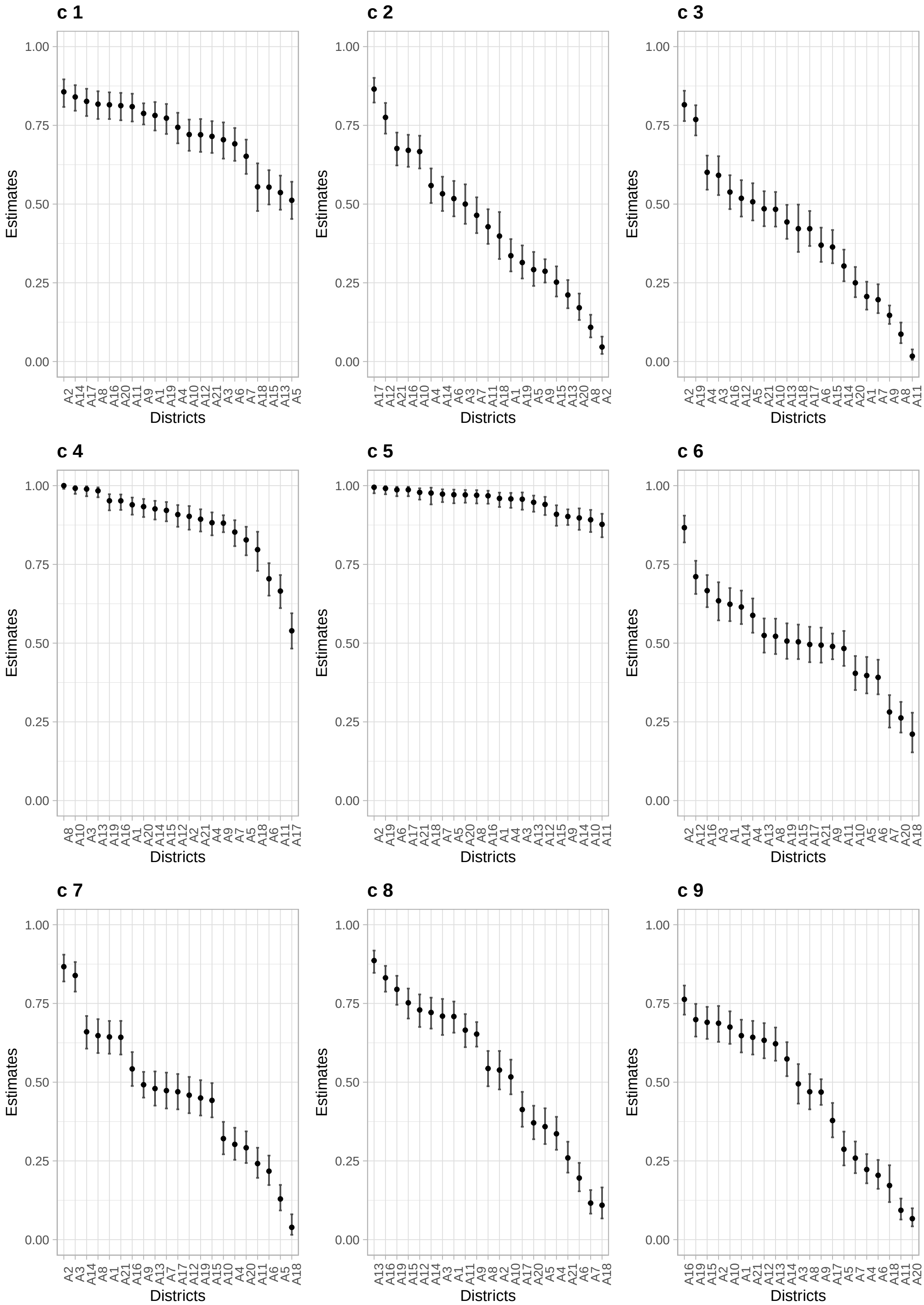

Figure 2.2: Confidence intervals of population estimates. A_i are the 21 counties (detailed in Figure 2.1) and c_j are the 9 criteria (described in 2.2) considered. Counties are ordered in descending order for each criterion.

2.4.2 MCDA models

The Multi-Attribute Utility Theory (MAUT) methodological framework can be divided into two steps. First, a utility function is defined to construct the global value of each county. Among the several functions that can be used (e.g., additive, multiplicative and multi-linear) we use the additive form for simplicity reasons[1]. Thus, the utility value for each county is calculated as the weighted sum of performance scores. Then, these utility values are ordered to obtain the county ranking.

In our extended version of MAUT, denoted as Model U, the statistical uncertainty in the input data is incorporated as follows:

- Uncertainty propagation. This lets us calculate the effect of the individual uncertainties of the data on the global uncertainty of the utility values. To do so, we use Monte Carlo simulations, in which 10,000 performance scores are randomly generated for each county and the probability distribution of the county's utility value is derived.

- Hypothesis testing. This lets us determine the statistical significance between the utility distribution of a pair of counties. We use the Welch's t-test (Welch 1947). If the null hypothesis of no differences in the utility value means is accepted, the counties are considered to occupy the same ranking position; otherwise, one county ranks higher than the other.

The ELECTRE III model also involves two steps (explained in detail in Chapter 1). First, an outranking relation is constructed for each pair of counties so as to assess the preference relation between the two. Then, the outranking relations between all pairs are used to build two pre-orders through descending and ascending distillations, and a final pre-order is obtained from their intersection.

Our extended version of ELECTRE III, denoted as Model S, incorporates data uncertainty directly into the discrimination thresholds, as explained in Chapter 1. Figure 2.3 illustrates the different ways models U and S integrate data uncertainty. Model S is more straightforward, as data uncertainty is directly included through

[1]The final results will depend on the particular utility function selected.

the discrimination thresholds. In contrast, model U requires more steps to propagate uncertainty and conduct hypothesis testing before obtaining the final ranking.

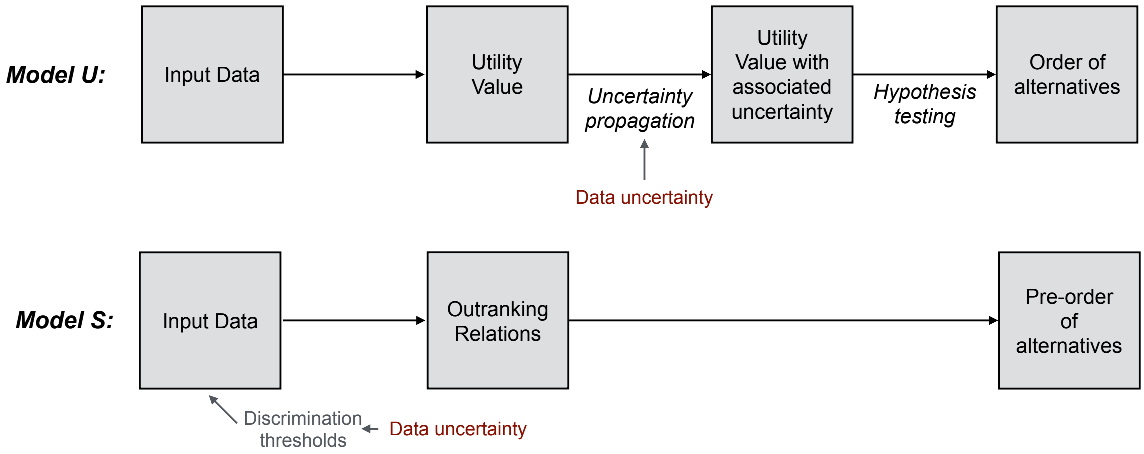

Figure 2.3: Incorporation of data uncertainty in the extended MCDA models. Model U, based on compensatory MAUT, requires propagating uncertainty of the utility values, and hypothesis testing. Model S, based on non-compensatory ELECTRE, incorporates uncertainty directly into the discrimination thresholds.

2.5 Results and Discussion

We first compare the rankings of models U and S (2.5.1). We then discuss the implications of this county prioritization for WASH planning in Kenya (2.5.2).

2.5.1 Comparison of rankings

The two MCDA models result in similar county orders (Figure 2.4). In both cases, counties of Molo (A3), Kisumu (A12), Nyando (A16) and Uasin Gishu (A19) occupy the leading positions. A closer look at the survey data (Table 2.4) reveals why these four counties have better WASH services than the rest. For instance, in terms of water supply (c_4), more than 95% of their populations have access to functioning water points, 8% above the national average. The same happens in respect to the distance from dwelling to water (c_3): while on average only 40% of the population has access to a water source in less than 30 minutes, in these four counties the proportion is more than 12% higher. In addition, more than 71% of households own latrines in good hygienic conditions (c_8), far from the average of 53%.

On the other hand, both models place counties of Garissa (A5), Isiolo (A6), Mwingi (A18) and Mandera (A20) in the lowest ranks. These four counties severely lack adequate quantities of water for domestic purposes (c_6): only 21-39% of their populations have access to more than 20 liters of water per capita per day, 30% below the national estimate. Furthermore, whereas the access to improved sanitation services is 46% on average (c_7), it does not reach 29% in these counties.

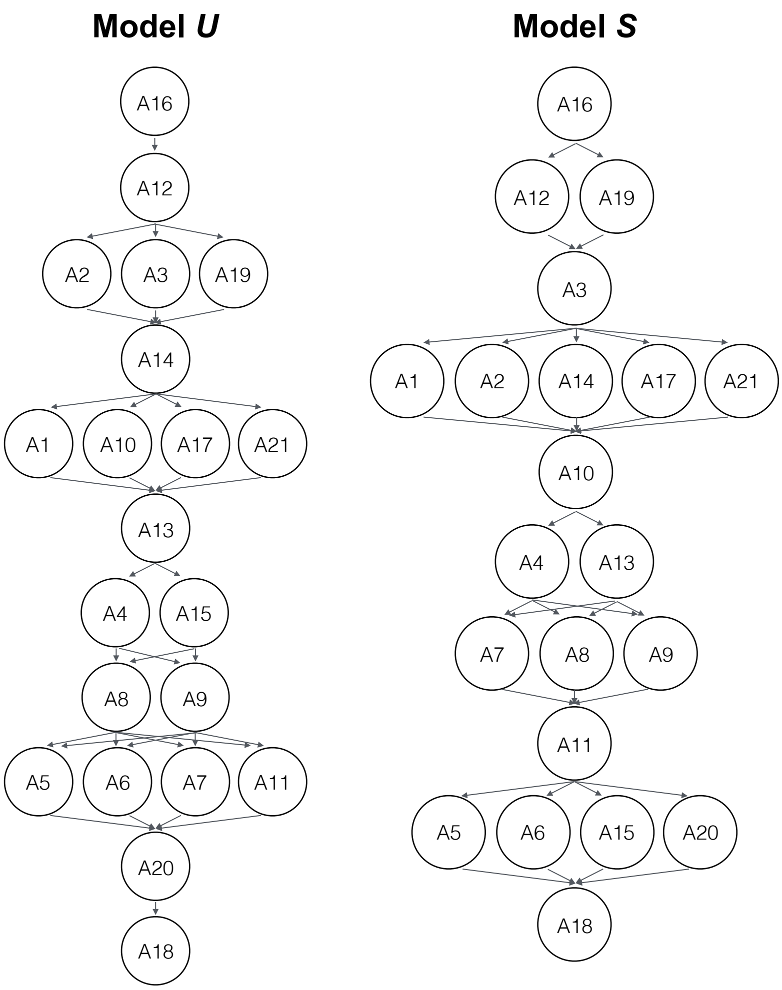

Figure 2.4: Rankings of the 21 counties obtained with models U (compensatory) and S (non-compensatory). A_i are the 21 counties (detailed in Figure 2.1).

The only major divergence between the two models is the position of counties Tana River (A7) and Kajiado (A15). Model U ranks Kajiado higher than Tana River, while model S results in the opposite. This reflects the different principles underlying the two models, especially concerning the compensatory nature of their aggregation procedures. The Kajiado county has better services in terms of distance to the source, functionality of water supplies, domestic water consumption, household water quality and water treatment (i.e., c_3, c_4, c_6, c_8 and c_9), but performs poorly in criteria related to improved water sources and person responsible for collecting water (c_2 and c_5). Model U, being fully compensatory, places Kajiado in a higher position because the bad performances on these two criteria are compensated by the rest. On the contrary, model S, which is only partially compensatory, results in a lower position for Kajiado county.

The overall convergence between the rankings coincides with results of other studies (Duckstein et al. 1982; Roy et al. 1986; Goicoechea et al. 1992; Mahmoud and Garcia 2000), where rankings obtained by MAUT and ELECTRE III were largely the same.

2.5.2 Policy implications for WASH planning in Kenya

Prioritization maps are powerful instruments for displaying information and easily identifying where to target future WASH investments. They are widely used in the sector to identify geographical inequalities and areas for improvement (Pullan et al. 2014; Yu et al. 2014; Giné-Garriga et al. 2013; Giné-Garriga et al. 2015; Patunru 2015; Jia et al. 2016; Chaudhuri and Roy 2017; Cole et al. 2018; He et al. 2018).

In our case, both models lead to similar targeting and prioritization maps (Figure 2.5). An important gap in WASH services can be identified in the North and North-Eastern regions of the country, which have been historically under-served (The World Bank 2018). Such regions are mainly arid and semi-arid, with water scarcity leaving a majority of the inhabitants dependent on unimproved water sources (Kurui et al. 2019). Furthermore, counties in these areas have the lowest population densities (less than 10 people per squared kilometer, compared to the average 250 people per square kilometer in Western counties) (Jayne and Muyanga 2012).

This has an immediate policy implication: although these regions are more difficult to target because of their dispersed population, more efforts – and allocated budgets – must be made to improve WASH service provision.

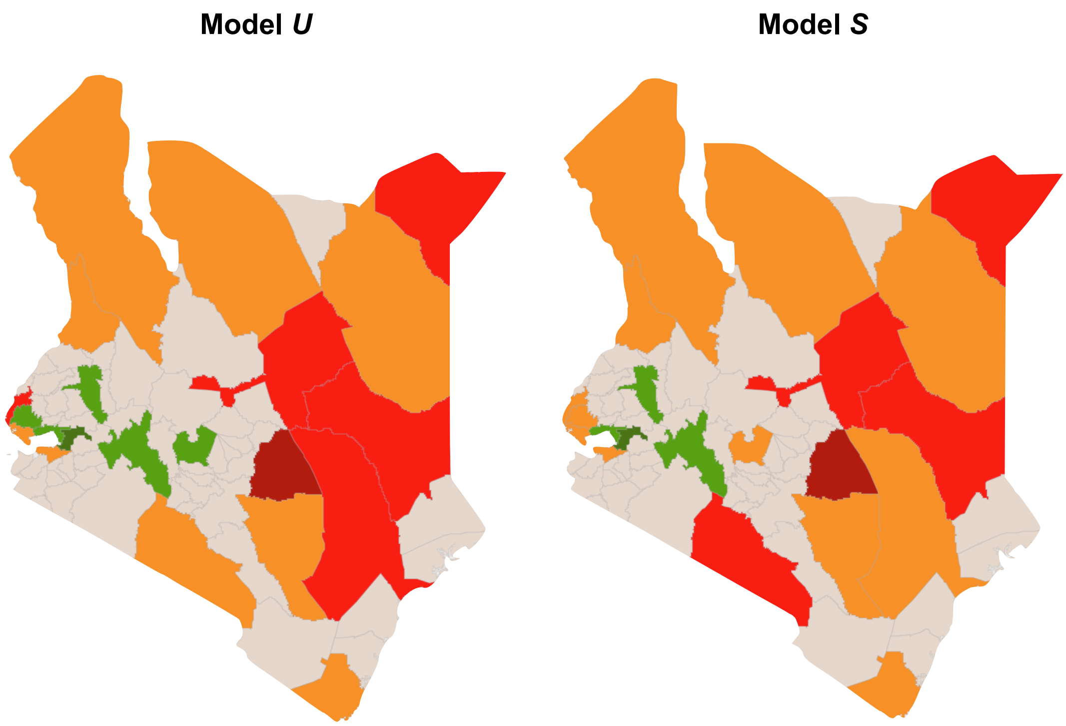

Figure 2.5: Prioritization of counties based on their ranking, from the most (red) to the least (green) disadvantaged.

2.6 Key messages

Multi-Criteria Decision Analysis (MCDA) models can help decision- and policy-makers target and prioritize those populations (and regions) most in need for better WASH services. However, selecting the most appropriate MCDA model can be challenging, even more when dealing with data with a certain degree of uncertainty, as there is a lack of MCDA models integrating this uncertainty. We have presented and compared two extended, data-driven MCDA models, using different criteria aggregation and uncertainty incorporation techniques.

The key messages to highlight from this Chapter are:

- Data used in the WASH sector are collected from household surveys, and are thus subject to statistical uncertainty that needs to be characterized. A simple way to define the uncertainty in household estimates, and measure its effect on the MCDA output, is through confidence intervals.

- The two models incorporate data uncertainty in a considerably different manner. Model U, based on MAUT, requires a step of "uncertainty propagation" in order to calculate the uncertainty in global utility values, and another of "hypothesis testing" to determine the final ranking. Model S, based on ELECTRE III, presents a more straight-forward ranking procedure, as data uncertainty is incorporated when defining the discrimination thresholds.

- Both models can be useful decision-aid instruments for targeting and prioritization in the WASH sector. In our case study, the two models yield similar rankings of counties. However, we must remember that MCDA models should not be used to reveal a *right* prioritization, but rather to guide the decision analysis process.

3

Global monitoring of access to WASH

Abstract:

Nationally representative household surveys are the main source of data for tracking drinking water, sanitation and hygiene (WASH) coverage. However, all survey point estimates have a certain degree of error that must be considered when interpreting survey results for decision- and policy-making. In this Chapter, we develop an approach to characterize and quantify uncertainty around WASH estimates. We apply it to four countries – Bolivia, Gambia, Morocco and India – representing different regions, number of data points available and types of trajectories, in order to illustrate the importance of communicating uncertainty for temporal estimates, as well as taking into account both the compositional nature and non-linearity of the data. While it only considers the uncertainty arising from sampling, our approach is particularly useful in the WASH sector, where the dissemination and analysis of uncertainty lags behind.

Keywords: Trend analysis; Water, Sanitation and Hygiene (WASH); SDG 6; Household Surveys

This chapter is based on:

- Ezbakhe, F. and Pérez-Foguet, A., 2019. Estimating access to drinking water and sanitation: the need to account for uncertainty in trend analyses. *Science of the Total Environment*, 696. doi: 10.1016/j.scitotenv.2019.133830

3.1 Introduction

Substantial progress has been made worldwide in increasing people's access to water, sanitation and hygiene (WASH) services. Between 1990 and 2015, almost 2.6 billion people gained access to improved drinking water sources, and 2.1 billion gained access to improved sanitation facilities (JMP 2015). Notwithstanding this laudable achievement, there remains a tremendous effort to reach the millions of unserved people: nearly 845 million people still lack access to basic drinking water, and 2,300 million to adequate sanitation (JMP 2017b). This has severe health implications: in 2016 alone, inadequate WASH was estimated to cause 829,000 diarrhoeal deaths, constituting 60% of the total diarrhoeal deaths, that would have been preventable through access to improved WASH services (Prüss-Ustün et al. 2019). The economic burden of poor WASH services is also considerable. For instance, in 2015, the lack of access to sanitation was estimated to cost the global economy 223 billion USD, corresponding to 0.9% of the global GDP (LIXIL 2016). To address this, the sixth Sustainable Development Goal (SDG 6) of the 2030 Agenda specifically calls for countries to "achieve universal and equitable access to safe and affordable drinking water for all" (target 6.1) as well as to "achieve access to adequate and equitable sanitation and hygiene for all and end open defecation" (target 6.2) (UNGA 2015).

Realizing these bold targets will require both greater investments in WASH, and understanding the levels and trends in service coverage in order to evaluate progress, and identify and prioritize successful policies (Cronk et al. 2015). The responsibility of monitoring progress of the SDG 6 targets related to WASH lies within the WHO/UNICEF's Joint Monitoring Programme (JMP). Since 1990, the JMP has been producing estimates of national, regional and global progress on WASH access. The JMP currently produces estimates for a total of 26 indicators related to WASH, all of which refer to the proportion of the population using a specific level of WASH services (JMP 2018). The JMP estimation method begins with the identification and compilation of all nationally-representative data relevant to the use of WASH services. A linear least-squares regression is then used to model the proportions of the population over time.

However, the JMP estimation method presents some limitations. First, the use of linear regression introduces substantial bias in estimates, particularly when coverage rates show non-linear patterns (Bartram et al. 2014; Fuller et al. 2016). Furthermore, as Luh et al. (2013) highlight, simple linear regressions fail to capture progressive realization of the human rights to water and sanitation. Several alternative regression approaches have been proposed to address this shortcoming, including quadratic, logit, piecewise linear and generalized additive models (Wolf et al. 2013; JMP 2014).

Second, the JMP estimation method fails to account for the compositional nature of the data. Proportions of the population are subject to a unit-sum constraint and thus cannot vary independently, which invalidates most standard statistical approaches. Pérez-Foguet et al. (2017) have recently addressed this issue by modelling JMP data with compositional data analysis. They concluded that the log-ratio transformation[1] of data did not only avoid misleading results from when proportions were analyzed separately, but also helped improve the performance of regression models, especially when coverage rates are near 0% or 100%.

Third, the characterization and representation of uncertainty around estimates remains an untackled issue by the JMP (JMP 2014). This is crucial, as estimates are largely based on data from nationally representative household surveys, subject to both sampling and non-sampling errors. Thus, in addition to further minimize these errors, uncertainty assessment of the estimates – in the form of confidence intervals, for example – is indispensable for an evidence-based analysis of levels and trends in WASH coverage. Failure to conduct and report such confidence intervals may lead to misinterpretation of coverage rates and trends, and ultimately undermine effective policy-making for WASH.

But reporting confidence intervals in WASH estimates is far from an easy endeavour. On one hand, information on sampling errors is seldom included in household survey reports and, even when published, it is often unclear whether they have been computed accurately (Betti et al. 2018). On the other hand, the general assumption that estimates from household surveys are approximately normally distributed can be problematic when coverage rates are near 0% or 100% (Janicki 2019).

[1] Log-ratio analysis uses log-ratio transformations of the data to take them from the "simplex" to real space, hence avoiding many of the problems associated with constrained data.

As the interest in estimates of WASH coverage will continue to grow in the SDG era, we can learn much more by showing the uncertainty around WASH estimates. In this Chapter, we present an approach to characterize and communicate uncertainty around water and sanitation estimates, and, at the same time, take into account both the compositional and non-linear nature of the data. In particular:

- We develop an approach to characterize uncertainty around water and sanitation estimates, based on compositional data analyses, non-linear regressions and Monte Carlo simulations.
- We apply it to four countries (Bolivia, Gambia, Morocco and India), representing different SDG regions, number of data points available and trajectory.
- We assess the magnitude of confidence intervals for WASH estimates, together with the effect of compositional and non-linear patterns in the available data.

We organize the rest of the Chapter as follows. We first provide a background on compositional data analysis in Section 3.2. In Section 3.3, we present the case of global monitoring of WASH access. We then present our method for characterizing uncertainty in WASH estimates, and the four countries selected in Section 3.4. We present and discuss the results of applying our approach in Section 3.5. Finally, in Section 3.6 we highlight the main conclusions of the Chapter.

3.2 A background on compositional data analysis

Compositional data are arrays of non-negative multivariate components that are some part of a whole. They are usually recorded in closed form, summing to a constant (e.g., proportions summing to 1 or percentages summing to 100%). Such data are widespread in many disciplines, such as geosciences, biology, economics, and population studies (Lloyd et al. 2012; Ferrer-Rosell et al. 2016; Bergeron-Boucher et al. 2017; Wei et al. 2018; Linares-Mustaros et al. 2018; Marcillo-Delgado et al. 2019). By definition, data on WASH access are compositional: the individual proportions of the population using each WASH service level are not independent of each other, but related by being expressed as percentages of the total population.

Compositional data have particular and essential properties that arise from the fact that they represent parts of some whole (Pawlowsky-Glahn and Egozcue 2006). They are a vector of strictly positive real numbers with a constant sum constraint:

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_D);\ x_i > 0 \qquad i = 1, 2, \ldots, D;\ \sum_{i=1}^{D} x_i = \kappa \tag{3.1}$$

The elements of a composition, x_i, are called components or parts, and the only relevant information is contained in the ratios between them (Pawlowsky-Glahn et al. 2015). This conditions the relationships that variables have to one another. For instance, if the values of one component are decreasing over time, values of at least one other component will have to increase to preserve the constant sum. As a result, compositional data are enclosed in a subspace where they can only vary between 0 and the radix value (κ). Such subspace, known as the simplex, does not follow the rules of Euclidean geometry, which makes standard statistical techniques inappropriate for the analysis of compositional data (Aitchison 1999).

Because of this particular geometry, working in the simplex can be counterintuitive. As an alternative, compositional data may be transformed to the real scale where classic statistics can be applied (Pawlowsky-Glahn et al. 2015). These transformations are based on log-ratios between components, and lead to "open" data, called coordinates, that can take any real value between $-\infty$ and ∞. Several log-transformations have been proposed, including the additive log-ratio (ALR), the centred log-ratio (CLR) and the isometric log-ratio (ILR) (Aitchison 1982; Egozcue et al. 2003).

In the following, the ILR transformation is applied to perform the statistical analysis of WASH data. It represents the composition given a particular orthonormal basis in the simplex (Egozcue et al. 2003), given by:

$$\boldsymbol{y} = ilr(\boldsymbol{x}) = \log(\boldsymbol{x}) \cdot \boldsymbol{V} \tag{3.2}$$

where $\boldsymbol{x}$ is the vector with the D parts of the composition, $\boldsymbol{V}$ a $D \cdot (D-1)$ matrix representing the orthonormal basis in the simplex, and $\boldsymbol{y}$ the resulting vector with the $D-1$ coordinates of the composition in that basis $\boldsymbol{V}$.

There are several ways to define orthonormal bases in the simplex, one of which consists in a sequential binary partition (SBP) of the composition (Pawlowsky-Glahn et al. 2015). A SBP represents a hierarchy of the parts of a composition, and contains successive splits of the parts into two groups, coded by the signs + and –, respectively (Pawlowsky-Glahn and Egozcue 2011). The orthonormal basis $\boldsymbol{V}$ can be obtained from the SBP as:

$$y_i = \sqrt{\frac{r_i s_i}{r_i + s_i}} \log\Big(\frac{(\prod_+ x_{ij})^{1/r_i}}{(\prod_+ x_{ik})^{1/s_i}}\Big) \qquad i = 1, 2, \ldots, D-1 \tag{3.3}$$

where y_i is the i^{th} coordinate (or balance) of the composition, x_{ij} and x_{ik} are the components coded as + and – in the i^{th} partition, and r_i and s_i are the number of parts with positive and negative signs in that partition, respectively.

Once the data are transformed into ILR balances, standard statistical approaches can be applied. Finally, regression points can be back-transformed to the original space using the inverse ILR:

$$\boldsymbol{x} = ilr^{-1}(\boldsymbol{y}) = C\ [exp(\boldsymbol{V} \cdot \boldsymbol{y})] \tag{3.4}$$

where $\boldsymbol{y}$ contains the ILR coordinates of $\boldsymbol{x}$ with respect to the basis $\boldsymbol{V}$, and C is the closure operator:

$$C[\boldsymbol{x}] = \Big(\frac{x_1}{\sum_{i=1}^{D} x_i}, \frac{x_2}{\sum_{i=1}^{D} x_i}, ..., \frac{x_D}{\sum_{i=1}^{D} x_i}\Big) \tag{3.5}$$

For a 4-part composition, $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$, an example SBP can be:

Order	x_1	x_2	x_3	x_4	r	s
1	+1	+1	-1	-1	2	2
2	+1	-1	0	0	1	1
3	0	0	+1	-1	1	1

and therefore, the orthonormal basis is:

$$\boldsymbol{V} = \begin{bmatrix} 1/2 & 1/\sqrt{2} & 0 \\ 1/2 & -1/\sqrt{2} & 0 \\ -1/2 & 0 & 1/\sqrt{2} \\ -1/2 & 0 & -1/\sqrt{2} \end{bmatrix}$$

The ILR coordinates can be computed, following Equation 3.3, as:

$$y_1 = \frac{1}{2}\log\frac{x_1 x_2}{x_3 x_4} \ ; \ y_2 = \frac{1}{\sqrt{2}}\log\frac{x_1}{x_2} \ ; \ y_3 = \frac{1}{\sqrt{2}}\log\frac{x_3}{x_4} \tag{3.6}$$

3.3 The case of global WASH monitoring

The JMP currently monitors coverage of WASH services in 230 countries and territories. Six primary indicators are used to monitor water and sanitation access, each reported separately for urban and rural populations (Table 3.1).

Table 3.1: Primary water and sanitation indicators used by the JMP.

Water	*The proportion of the population that uses...*
W_1	All improved drinking water sources
W_2	Piped drinking water sources
W_3	No drinking water sources (i.e., surface water)
Sanitation	*The proportion of the population that uses...*
S_1	All improved sanitation facilities
S_2	Improved sanitation facilities connected to sewers
S_3	No sanitation facilities (i.e., open defecation)

For a given indicator, the data points available vary depending on the country (Table 3.2). More than one third of all countries lack any data on surface water and open defecation (34.9% and 40.6%, respectively), and nearly four fifths have less than 10 data points for these two indicators (77.3% and 77.7%). Furthermore, most countries do not provide complete data for all three indicators: only 21.4% and 17.5% of all countries present more than 10 complete data points for water and sanitation, respectively.

Table 3.2: Data availability for indicators in the JMP database (1990-2015) (JMP database was extracted on May 28th, 2019).

Service	Setting	Indicator	Number of countries with the following data points						
			0	*1*	*2*	*3-5*	*6-10*	*11-15*	*>16*
Water	Urban	W_1. Improved	29	13	17	46	48	28	48
		W_2. Piped	31	15	18	43	47	28	47
		W_3. Surface	93	16	10	25	34	19	32
		W_{123}. All indicators	93	16	10	25	36	18	31
Water	Rural	W_1. Improved	34	16	15	42	46	32	44
		W_2. Piped	36	19	14	40	46	31	43
		W_3. Surface	93	16	10	25	33	19	33
		W_{123}. All indicators	93	16	10	25	36	17	32
Sanitation	Urban	S_1. Improved	23	23	10	35	57	26	55
		S_2. Sewer	32	25	19	49	48	20	36
		S_3. Open defecation	80	17	12	34	35	20	31
		S_{123}. All indicators	87	20	17	36	29	19	21
Sanitation	Rural	S_1. Improved	27	23	9	33	56	27	54
		S_2. Sewer	42	24	14	46	47	23	33
		S_3. Open defecation	82	17	12	32	34	21	31
		S_{123}. All indicators	88	21	16	35	29	18	22

Data are obtained from nationally representative household surveys – including Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), World Health Surveys (WHS), and Living Standards Measurement Studies (LSMS) – and national censuses conducted by governments. Household survey data are subject to both sampling and non-sampling errors. Measuring the WASH services of one or another sample of households taken from the same population would give different estimates; this is the origin of sampling errors. Non-sampling errors, on the other hand, arise from biases in data collection, such as omission of households, inappropriate interview methods, and errors in data processing (Banda 2003).

Sampling errors are generally measured with the standard error statistic, which reflects the variability between estimates we would obtain from different samples of the population. However, the JMP does not report the standard errors in its household database. This is partly due to “concerns that non-sampling errors are likely to dominate sampling errors, especially since the underlying household survey data used to assess basic services often have large sample sizes” (Bain et al. 2018). While it is true that sampling errors represent only one component of the total survey error and may underestimate non-sampling errors, they still need to be accounted for. For instance, in Burkina Faso, the relative standard errors of water and sanitation estimates range between 0.7% and 19.2%, with an average of 7.1% (WHS 2003; MICS 2006). The challenge, however, lies in obtaining all standard errors of water and sanitation data, as they are not available in the majority of survey reports. For instance, in DHS reports, errors are only provided for a small selection of variables that does not include those related to water and sanitation (Verma and Lê 1996; Vaessen et al. 2005).

3.4 Methodology

In this section, we first describe our method for the characterization of uncertainty around water and sanitation estimates (3.4.1). We then present the four case studies – urban water in Bolivia and Gambia, and rural sanitation in Morocco and India – we selected for the analysis (3.4.2).

3.4.1 Proposed approach

To characterize uncertainty around WASH estimates and, simultaneously, consider the compositional and non-linear nature of data, our method encompasses the following steps: (i) pre-process the JMP data to express them as 4-part compositions, (ii) treat the zero values in the compositional data by imputation techniques, (iii) estimate the standard errors of the proportions with a generalized relative standard error function, (iv) generate simulations of the compositions for each year following an extended Beta distribution, (v) fit the non-linear regression model to the data, and (vi) calculate the 95% confidence intervals from the 2.5th and 97.5th regression percentiles. The proposed approach is illustrated in Figure 3.1.

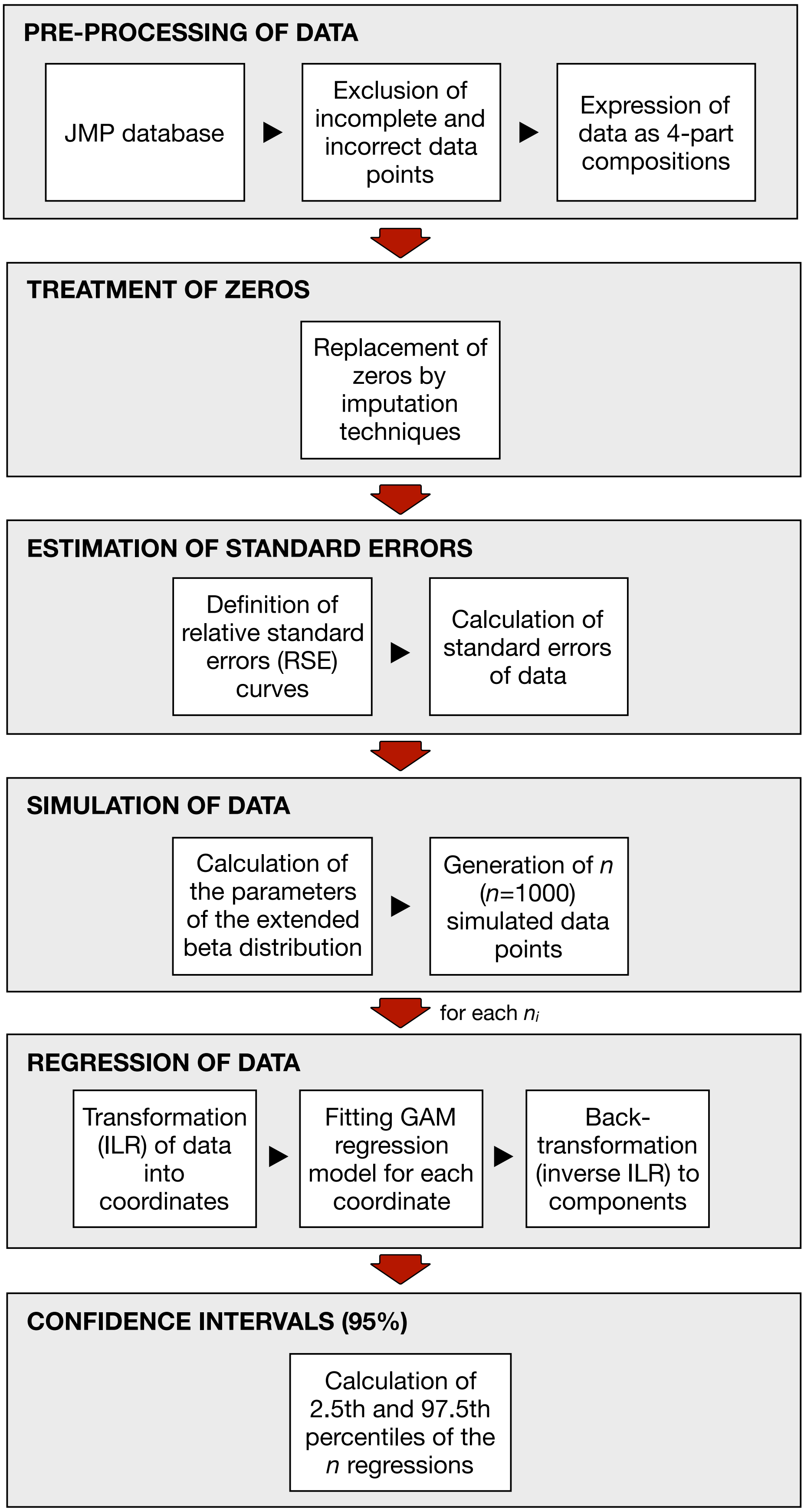

Figure 3.1: Proposed approach for characterizing uncertainty in WASH estimates.

Step 1. Pre-processing of data

The 3 primary indicators for water and sanitation used by the JMP (from Table 3.1) are expressed as 4-part compositions, $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$, as shown in Table 3.3.

Table 3.3: Pre-processing of JMP indicators for water and sanitation.

Water	Means of estimation	*The proportion of the population that uses...*
x_1	W_2	Piped drinking water sources
x_2	$W_1 - W_2$	Other improved drinking water sources
x_3	W_3	No drinking water facility (i.e., surface water)
x_4	$1 - W_1 - W_3$	Other unimproved drinking water sources
Sanitation	**Means of estimation**	***The proportion of the population that uses...***
x_1	S_2	Improved sanitation facilities connected to sewers
x_2	$S_1 - S_2$	Other improved sanitation facilities
x_3	S_3	No sanitation facilities (i.e., open defecation)
x_4	$1 - S_1 - S_3$	Other unimproved sanitation facilities

Therefore, only years with complete data for all 3 indicators (i.e., all parts of the composition) are included in the analysis. Furthermore, years with out-of-range data (i.e., $W_1 + W_3 > 1$ and $W_1 < W_2$ for water, and $S_1 + S_3 > 1$ and $S_1 < S_2$ for sanitation) are excluded. For instance, according to the JMP database, the percentages of people using the different types of drinking water sources in urban Botswana in 2007 were 98.9% (W_1), 99.0% (W_2) and 1.3% (W_3), which is clearly erroneous: the sum of the people using improved and unimproved drinking water sources cannot exceed 100%, and the percentage of people using piped supplies cannot be greater than those using all forms of improved sources.

Step 2. Treatment of zeros

The compositional analysis of JMP data is based on log-ratios of parts. For this, zeros must be treated in the first place. In this case, since data are mainly sourced from household surveys, we consider zero values as non-structural zeros. In other words, since we cannot be completely sure that there is not a single household using a particular WASH service, zeros present in the data can be seen as rounded zeros.

As such, we replace them with the following imputation technique (Martín-Fernández et al. 2003):

$$r_j = \begin{cases} \delta, & \text{if } x_j = 0 \\ \left(1 - \frac{\sum_{k|x_j=0} \delta}{c}\right) x_j & \text{otherwise} \end{cases} \tag{3.7}$$

where r_j is the non-zero composition, δ is the imputed value on the part x_j, and c is the constant sum-constraint (in this case $c = 1$). As Martín-Fernández et al. (2003) explains, the imputed value δ can be associated to the rounding-off error (i.e., the precision of the data). In this case, since data included in the JMP database are given in percentages of the population with a precision of 1 digit after the decimal point, $\delta = 0.5 \cdot 10^{-3}$.

Step 3. Estimation of standard errors

To overcome the problem of non-reporting of standard errors of survey data, we use a generalized relative standard error (RSE) function, which defines a relationship between the relative standard errors (i.e., the standard errors expressed as a percentage of the estimated proportion) and the corresponding proportion. We use a modified version of Gabrel (2000) formula:

$$RSE(p) = \sqrt{a + b\frac{1-p}{p}} \tag{3.8}$$

where p is the estimated proportion (i.e., x_i from our compositional data), and a and b are coefficients derived from the RSE curves. These RSE curves indicate the magnitude of the relative standard error for estimated proportions of various sizes and should be interpreted as an approximation rather than exact values for any specific proportion. They have the following meaning: for small values of p, relative standard errors are relatively high (when p approaches zero, the relative standard error approaches infinite), and decrease with a square-root dependence as p increases, reaching its minimum value for $p = 1$.

Therefore, coefficients a and b can be obtained by fixing minimum and maximum values for the relative standard errors:

$$a = {RSE_{min}}^2 \tag{3.9}$$

$$b = \frac{\delta}{1-\delta}({RSE_{max}}^2 - {RSE_{min}}^2) \tag{3.10}$$

where RSE_{min} and RSE_{max} are the selected minimum and maximum relative standard errors, and δ is the precision of the data.

To fix RSE_{min} and RSE_{max}, we differentiate between three types of sources. In household surveys from MICS, DHS, LSMS and WHS, which are generally considered of higher quality, they are set as 4% and 40%, respectively. In other household surveys, 8% and 80%. In censuses, where all households are counted, they are considered zero. The standard errors of the data are finally obtained by multiplying the RSE by the estimated proportion.

Step 4. Simulation of data

Confidence intervals of estimates are constructed via simulation techniques. This involves generating n simulations ($n = 1,000$) of the data for each year assuming a generalized Beta distribution, also known as Pearson Type I (Bowman and Shenton 2007). The use of a generalized Beta distribution, instead of the Normal distribution, is motivated by its ability to model proportions near the boundaries (i.e., 0 or 1), where the normal approximation of the sampling distribution is no longer valid (Bowman and Shenton 2007). Essentially, Pearson Type I distributions are location-scale transformations of Beta distributions. The probability density function, supported in the interval (a_1, a_2) and shape parameters $\alpha, \beta > 0$, is a power function of the variable x and its complement, as follows:

$$B(x|a_1, a_2, \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(x-a_1)^{\alpha-1}(a_2-x)^{\beta-1}}{(a_2-a_1)^{\alpha-\beta-1}} \tag{3.11}$$

where $\Gamma(k)$ is the complete gamma function.

We use the method of moment estimators to obtain the parameters α and β of the generalized Beta distributions (Bain and Engelhardt 1987). This involves equating the moments of the generalized Beta distribution with the sample proportion p and variance σ:

$$p = a_1 + (a_2 - a_1)\frac{\alpha}{\alpha + \beta} \tag{3.12}$$

$$\sigma = (a_2 - a_1)^2 \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{3.13}$$

We express the limits a_1 and a_2 in terms of the precision of the data (i.e., $a_1 = 0.5 \cdot 10^{-3}$ and $a_2 = 9.5 \cdot 10^{-3}$). With these a_1 and a_2, and the resulting α and β parameters from equations 3.12 and 3.13, we simulate 1,000 proportions using the generalized Beta probability distribution.

Step 5. Regressions of data

For each simulation, we fit a regression model to the data following our compositional approach: we first log-transform the 4 components into 3 coordinates, which we model separately and then back-transform the regression results, as explained in Section 3.2. We also follow the standard approach, where the 4 components are modelled separately (as in the JMP estimation method) to compare it with ours.

We apply a non-linear regression based on the generalized additive model (GAM), in which the linear form is replaced by a sum of smooth functions[1]. We also apply ordinary least squares (OLS) linear regression to compare results. For this comparison, we use the Root-Mean-Square Error (RMSE)[2] and Nash-Sutcliffe Efficiency (NSE)[3].

Step 6. Confidence intervals

Finally, we calculate the 95% confidence intervals from the 2.5th and 97.5th percentiles of the regressions results.

[1] GAM is applied with thin-plate regression splines with four degrees of freedom (Wood 2003).

[2] RMSE represents the quadratic mean of the differences between the observed and the modelled proportions. A coefficient of 0 indicates a perfect fit to the data.

[3] NSE computes the square differences between the observed values and their mean. The efficiency is 1 if the model fits all data, 0 if it performs equally to the mean of the observed data, and below 0 if the observed mean is a better predictor than the model.

3.4.2 WASH data

We illustrate the application of the proposed approach in four case studies: (i) urban water in Bolivia, (ii) urban water in Gambia, (iii) rural sanitation in Morocco, and (iv) rural sanitation in India. These four case studies represent different SDG regions, number of data points and types of trajectories, as seen in Table 3.4.

Table 3.4: Case studies considered in the analysis. x_1, x_2, x_3 and x_4 refer to the proportions of the population using piped water/sewer connections, other improved water/sanitation, surface water/open defecation and other unimproved water/sanitation, respectively.

Country	Service	Setting	Region	Data points	Trajectories x_1	x_2	x_3	x_4
Bolivia	Water	Urban	Latin America and the Caribbean	25	S	A	NC	LD
Gambia	Water	Urban	Sub-Saharan Africa	7	LG	LD	D	LD
Morocco	Sanitation	Rural	Northern Africa and Western Asia	8	LD	LG	LD	NC
India	Sanitation	Rural	Central and Southern Asia	12	LG	LG	D	NC

Linear trajectories indicate that a country is making a constant change – either growth (LG), decline (LD) or none (NC) – and are the most common. Acceleration (A) occurs when a country starts making significant progress after a period of no or little progress. Deceleration (D) happens when the progress abruptly stalls, or even declines, before reaching high coverage levels. Saturation (S) takes place when progress slows down as it approaches full coverage, at which the rate flattens out.

3.5 Results and Discussion

We present and analyze the results from applying our method to the case studies. We first examine the importance of considering the compositional nature of data in trend analyses by comparing the standard and compositional approaches (3.5.1). Second, we evaluate the effect of non-linear patterns in the data by comparing OLS and GAM regressions estimates (3.5.2). Then, we analyze the impact of the standard errors on the confidence intervals of WASH estimates (3.5.3). Finally we discuss some policy implications for the global monitoring of WASH (3.5.4)

3.5.1 Compositional nature

Figure 3.2 shows the coverage estimates for the case of rural sanitation in Morocco, obtained with the standard statistical approach (i.e., OLS regression model fitted to each indicator separately) and our compositional approach. It also shows the official estimates provided by the JMP.

According to our OLS estimates, the percentages of the population using each service level in 2015 are: -0.1% for sewer connections, 77.1% for other improved facilities, 20.2% for open defecation and 2.8% for other unimproved facilities. These figures differ slightly from those estimated by the JMP (3%, 76.2%, 18% and 1.9%, respectively) for two main reasons:

- Our estimates are constructed from all data points available, whereas JMP only uses data from 2000 onwards. This has an important effect on the trend of service coverage: for sewer connections, for example, excluding the data points before 2000 leads to a growth trajectory instead of a decline.
- Although by definition OLS regression estimates would add up to 100%, there is no way to ensure that all four values lie between 0 and 1. The current JMP method avoids out-of-boundary values by adjusting the extrapolation results: if estimates are below 0% or above 100%, they are capped at 0% and 100%, respectively (JMP 2018).

However, this ad hoc post-process does not address the underlying problem: WASH data are compositional, and ignoring their compositional nature would lead to spurious results. Consequently, a compositional approach must be applied to obtain more theoretically sound estimates of the different proportions of the population having access to water or hygiene facilities, especially when coverage rates are near 0% and 100%. For instance, for access to sewer connections (x_1), the standard approach provides a negative percentage (-0.1%), while in the compositional approach it is positive (0.05%).

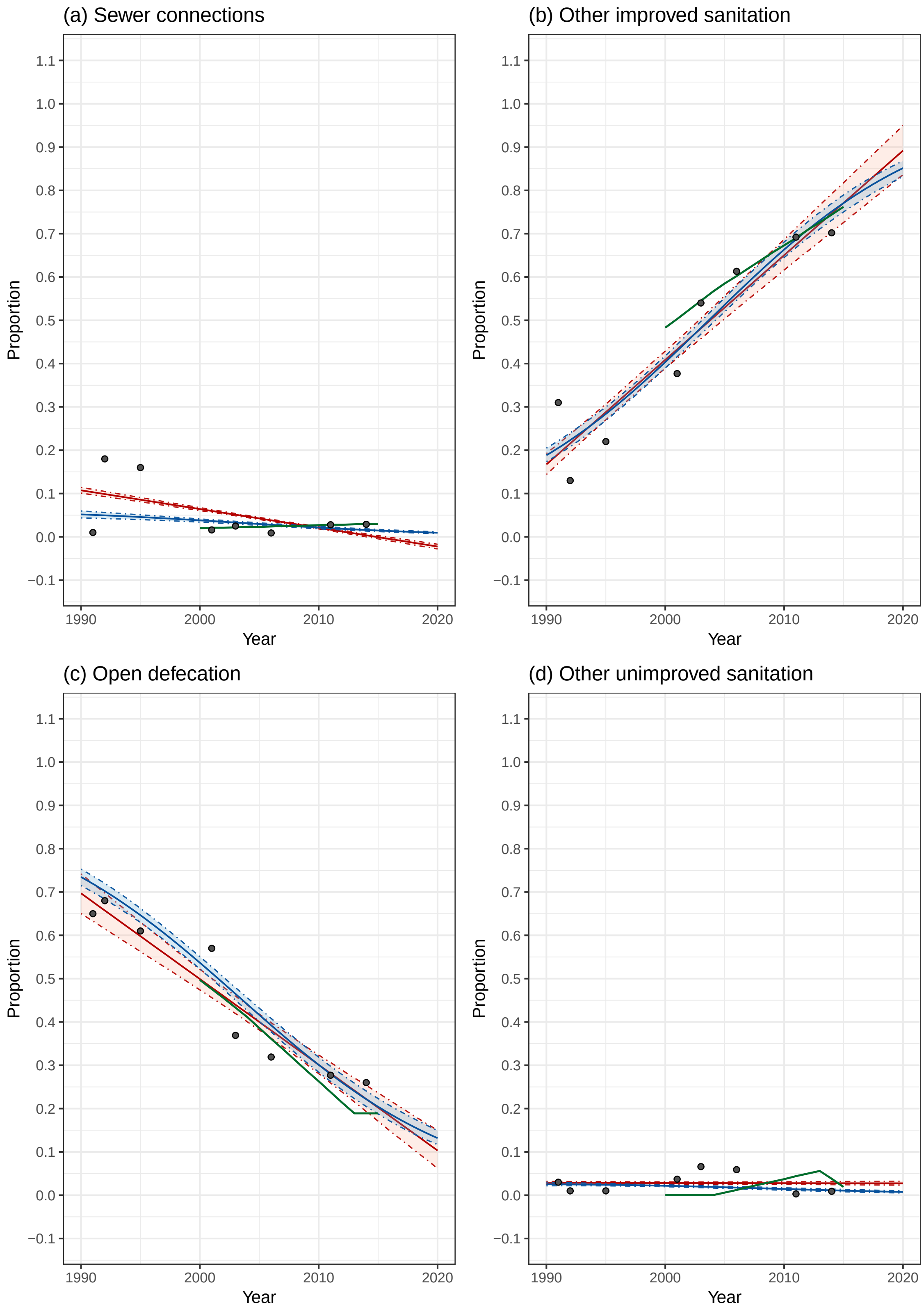

Figure 3.2: Coverage estimates for rural sanitation in Morocco, with OLS regression. In red, our estimates with the standard approach; in blue, our estimates with the compositional approach; in green, JMP estimates; in black, the coverage data points used. The shaded areas represent the 95% confidence intervals.

3.5.2 Non-linear patterns

Figure 3.3 illustrates the coverage estimates for all four case studies, obtained with linear (OLS) and non-linear (GAM) regression analyses. For countries with linear trajectories in their water and sanitation coverage, the differences between OLS and GAM regression models are not substantial. In India (Figure 3.d) for instance, where most components present linear patterns, OLS and GAM estimates never differ by more than 2 percentage points. However, when trajectories are non-linear, OLS regression is a bad estimator for coverage levels. This is evident in the case of Bolivia (Figure 3.a). For the first component (i.e., access to piped water sources), where data points follow a saturation trajectory, OLS overvalues coverage in the 1990–1995 and 2011–2015 periods, but undervalues the coverage levels between 1995 and 2011. The reverse occurs in the second component (i.e., access to other improved water sources), in which progress shows an accelerated pattern.

In order to further compare OLS and GAM regression estimates, we use the RMSE and NSE coefficients to quantitatively describe the accuracy of the models to the JMP data (Table 3.5). In all case studies, GAM provides better RMSE and NSE values, which translates in an improved accuracy of the regression models. This is particularly noticeable in non-linear trends. In Bolivia, the access to other improved water sources, with an acceleration pattern, presents a 0.0056 reduction in the root-mean-square errors; compared to the 0.0009 decrease in the case of access to surface water, where the trajectory is linear.

Despite the “superior” statistical power of non-linear regression models such as GAM, the JMP still chooses OLS for the estimation of coverage. There are three main reasons for this. First, the majority of countries show linear trajectories (62%-85.3%, depending on the service and setting (Fuller et al. 2016)), which makes the use of OLS appropriate. Second, OLS is easier to understand and implement by the JMP’s non-technical audience, which includes WASH sector stakeholders and policy-makers. In addition, OLS can be implemented without the need of specialized statistical software.

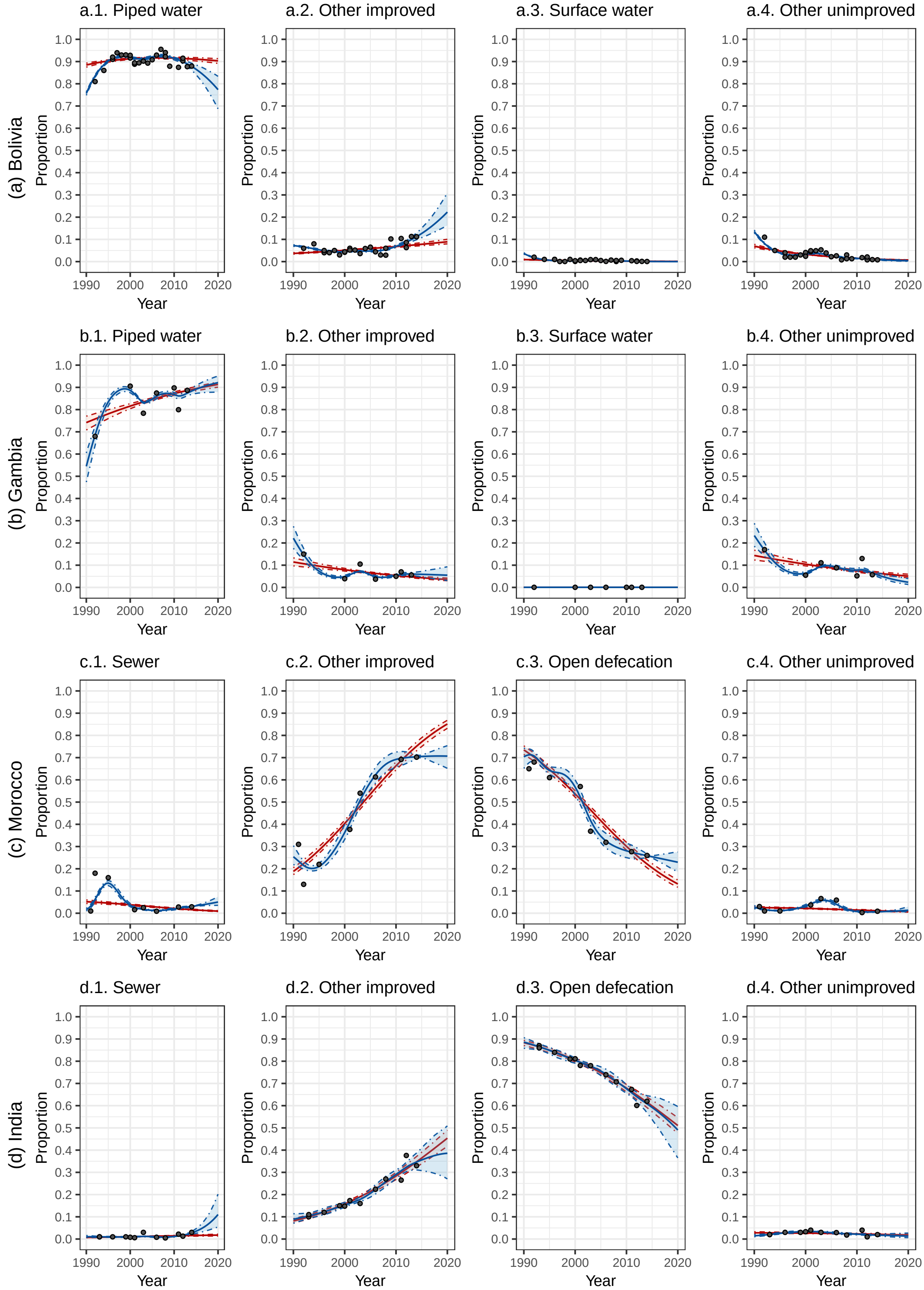

Figure 3.3: Coverage estimates for the countries, with the compositional approach. Countries include: (a) urban water in Bolivia, (b) urban water in Gambia, (c) rural sanitation in Morocco, and (d) rural sanitation in India. In red, estimates with OLS regression; in blue, estimates with GAM regression; in black, the coverage data points used. The shaded areas represent the 95% confidence intervals.

Table 3.5: Values of the root-mean-square error (RMSE) and the Nash-Sutcliffe Efficiency coefficient (NSE) for coverage estimates with OLS and GAM regression models. Estimates correspond to those from the compositional approach. x_1, x_2, x_3 and x_4 refer to the proportion of the population using piped water/sewer connections, other improved water/sanitation, surface water/open defecation and other unimproved water/sanitation, respectively.

Case study	Component	RSME		NSE	
		OLS	GAM	OLS	GAM
Urban water in Bolivia	x_1	0.0288	0.0183	0.0930	0.6315
	x_2	0.0211	0.0155	0.2894	0.6154
	x_3	0.0040	0.0031	0.2046	0.5431
	x_4	0.0161	0.0102	0.4437	0.7779
Urban water in Gambia	x_1	0.0584	0.0337	0.4183	0.8062
	x_2	0.0287	0.0147	0.4344	0.8527
	x_3	0.0000	0.0000	NA[1]	NA[1]
	x_4	0.0337	0.0228	0.3390	0.6970
Rural sanitation in Morocco	x_1	0.0640	0.0437	0.0514	0.5576
	x_2	0.0660	0.0435	0.8967	0.9551
	x_3	0.0568	0.0360	0.8827	0.9528
	x_4	0.0242	0.0086	-0.1327	0.8568
Rural sanitation in India	x_1	0.0078	0.0061	0.1483	0.4701
	x_2	0.0235	0.0215	0.9262	0.9378
	x_3	0.0150	0.0136	0.9702	0.9753
	x_4	0.0086	0.0068	0.0298	0.4010

[1]NA values are obtained because all observed data are equal.

3.5.3 Magnitude of uncertainty in estimates

In addition to generating better WASH estimates by considering both the compositional and non-linear nature of the data, one of the main contributions of our approach is the characterization of uncertainty around estimates. This is done by constructing the 95% confidence intervals (Table 3.6).

Confidence intervals are generally wider for the GAM model. In India, for example, the 2020 projection of the percentage of people practicing open defecation (i.e., x_3) is 47.2%-54.8% and 36.0%-61.0% with OLS and GAM, respectively. Furthermore, with few data points, the GAM model results in even wider confidence bounds. For example, in Morocco, the widths of the confidence intervals for 2020 are 3.5, 10.5, 9.6 and 2.2 percentage points (which represent 68.6%, 14.9%, 41.9% and 146.7% of the mean estimated values, respectively).

Constructing these confidence intervals around estimates can be extremely useful for two main purposes. First, it allows us to describe the precision of the estimate and represent its sampling distribution. Second, it provides context for policy-making. In the case of rural sanitation in India, 48.7% of population is expected to be practicing open defecation in 2020, with a 95% confidence interval of 36%-61%. This means that the true coverage of all the population is likely to be between 36% and 61%, but it might not be: the “95%” indicates that, if we repeated the same household survey many times, 95% of the them would include the true percentage, but 5% would not. Therefore, if the target is to decrease the prevalence of open defecation to below 50% in 2020, we could not be 100% certain that it is achieved, even with a mean point estimate at 48.7%. That is why policy-makers must consider the errors in the WASH coverage estimates when assessing progress against coverage targets. In this sense, our approach can help improve the use of JMP data to evaluate trends in coverage and inform decision-making.

However, it is important to recall that our approach estimates standard errors. As explained in Section 3.3, the JMP global database does not provide information on standard errors, because it is rarely included in household survey reports. That is why we approximate standard errors by defining a generalized curve for the relative standard errors, with fixed minimum and maximum RSE. These values are merely a first approximation, since the real RSE curve must be extracted from each household survey’s microdata.

Table 3.6: Coverage estimates with OLS and GAM regression models – with the compositional approach – for years 1990, 2000, 2010 and 2010. In parenthesis the 95% confidence intervals. x_1, x_2, x_3 and x_4 refer to the proportion of the population using piped water/sewer connections, other improved water/sanitation, surface water/open defecation and other unimproved water/sanitation, respectively.

		OLS				GAM			
Case study	x_i	1990	2000	2010	2020	1990	2000	2010	2020
Urban	x_1	88.4 (87.5-89.3)	91.3 (90.9-91.6)	91.5 (91.0-92.0)	90.3 (89.2-91.3)	75.6 (74.8-76.5)	91.9 (91.4-92.3)	91.5 (90.8-92.1)	77 (67.6-83.4)
water	x_2	3.7 (3.4-4.0)	5.1 (4.8-5.3)	6.8 (6.4-7.2)	9.0 (8.0-10.0)	7.2 (6.9-7.5)	4.5 (4.2-4.8)	6.9 (6.3-7.5)	22.6 (16.2-31.8)
in	x_3	0.9 (0.8-1.1)	0.4 (0.4-0.5)	0.2 (0.2-0.2)	0.1 (0.1-0.1)	3.7 (3.4-4.0)	0.3 (0.3-0.4)	0.2 (0.2-0.3)	0.0 (0.0-0.0)
Bolivia	x_4	7.0 (6.2-7.8)	3.3 (3.1-3.4)	1.5 (1.4-1.6)	0.7 (0.6-0.8)	13.5 (12.9-14.2)	3.3 (3.1-3.6)	1.4 (1.3-1.6)	0.4 (0.2-0.7)
Urban	x_1	74.1 (70.8-77.1)	81.6 (80.4-82.7)	87.2 (86.5-87.9)	91.3 (90.1-92.4)	54.5 (47.5-60.8)	88.7 (87.7-89.6)	86.8 (85.5-87.8)	91.9 (87.5-94.8)
water	x_2	11.4 (9.7-13.2)	7.9 (7.3-8.5)	5.4 (5.0-5.8)	3.6 (3-4.2)	22.1 (17.6-27.1)	4.9 (4.4-5.5)	5.4 (4.8-6.0)	5.7 (3.2-9.3)
in	x_3	0.1 (0.0-0.1)	0.1 (0.0-0.1)	0.1 (0.0-0.1)	0.0 (0.0-0.1)	0.0 (0.0-0.1)	0.1 (0.0-0.1)	0.1 (0.0-0.1)	0.0(0.0-0.1)
Gambia	x_4	14.4 (12.3-16.7)	10.4 (9.6-11.2)	7.3 (6.8-7.9)	5.1 (4.3-5.9)	23.4 (18.4-28.7)	6.4 (5.7-7.1)	7.8 (7.1-8.6)	2.4 (1.4-4.0)
Rural	x_1	5.2 (4.4-6.0)	3.8 (3.5-4.2)	2.1 (1.9-2.4)	0.9 (0.8-1.2)	1.6 (1.1-2.2)	3.6 (2.9-4.4)	2 (1.5-2.4)	5.1 (3.7-7.2)
sanitation	x_2	18.8 (17.2-20.4)	40.3 (38.9-41.8)	66.3 (64.5-68.2)	85.1 (83.3-86.8)	25.4 (21.4-29.7)	36.3 (33.1-39.6)	69.1 (65.7-72.6)	70.5 65.0-75.5)
in	x_3	73.5 (71.5-75.3)	53.7 (52.2-55.3)	30.1 (28.3-31.9)	13.2 (11.6-14.9)	70.4 (65.7-74.4)	56.7 (53.3-60)	28.1 (24.7-31.5)	22.9 (18.6-28.2)
Morocco	x_4	2.5 (2.2-3.0)	2.2 (2.0-2.4)	1.4 (1.2-1.7)	0.7 (0.5-1.0)	2.7 (2.0-3.5)	3.3 (2.8-3.9)	0.8 (0.5-1.1)	1.5 (0.7-2.9)
Rural	x_1	0.7 (0.6-0.9)	1 (0.9-1.2)	1.4 (1.3-1.6)	1.8 (1.4-2.2)	1.1 (0.7-1.6)	0.9 (0.8-1.1)	1.1 (0.9-1.4)	11.5 (5.6-19.9)
sanitation	x_2	8.2 (7.3-9.2)	15.8 (15-16.6)	28.4 (26.8-30)	45.4 (41.7-49.3)	9 (7.2-11.4)	15.3 (14.4-16.2)	29.1 (27.2-31.1)	38.4 (27.1-50.6)
in	x_3	88.3 (87-89.4)	80.5 (79.7-81.4)	67.9 (66.1-69.6)	51.1 (47.2-54.8)	88.5 (85.7-90.5)	80.4 (79.3-81.4)	67.6 (65.6-69.6)	48.7 (36.0-61.0)
India	x_4	2.8 (2.4-3.3)	2.6 (2.5-2.8)	2.3 (2.0-2.6)	1.8 (1.4-2.2)	1.5 (1.1-2.0)	3.4 (3.1-3.7)	2.2 (1.9-2.5)	1.4 (0.6-2.6)

3.5.4 Implications for global monitoring of WASH access

The use of confidence intervals to characterize uncertainty around regression estimates is not a novelty, and neither is the application of simulation techniques to construct these confidence bounds. In the health sector, for instance, confidence intervals are widely used to communicate uncertainty around child mortality indicators (Hodge et al. 2014; Bermejo et al. 2015; Minnery et al. 2015). By contrast, in the WASH sector, confidence bounds of estimates are rarely reported. The JMP (2014) justifies this by asserting that "it may be more important to be transparent about the level of uncertainty than being able to calculate a quantitative measure that could be misleading". However, how can we be transparent when the margins of error of survey data are not even included in the JMP's public database? And even if errors were available, how can we model their distributions when the assumption of normally distributed errors is no longer acceptable? These are precisely the two questions our approach tackles.

On one hand, our compositional and non-linear approach generates more theoretically sound coverage estimates, which could potentially better serve the needs not only of global monitoring agencies such as the JMP but also of country decision-makers. On the other hand, our uncertainty approach shows that approximating the standard errors with generalized RSE curves solves the issue of not reporting them. Indeed, this approach can be applied by the international community when dealing with trend analysis of WASH access (Jeuland et al. 2013; Cumming et al. 2014; Beyene et al. 2015; Pérez-Foguet et al. 2017; Armah et al. 2018; Chitsaz and Banihabib 2015; Nhamo et al. 2019). However, our approach provides a mere approximation of these errors. In order to obtain results more coherent with reality, more efforts should be made to include standard errors of WASH variables in household survey reports (and ultimately in the JMP database).

3.6 Key messages

Characterizing uncertainty around WASH estimates is crucial for a correct assessment of coverage trends over time. However, reporting confidence intervals around WASH estimates is not an easy endeavour, as survey data compiled by the JMP does not include publicly available margins of error. In this Chapter, we have presented a simple approach to characterize and communicate uncertainty in WASH estimates, and, simultaneously, produce better estimates by considering the compositional nature and non-linearity of the data.

Three key messages can be summarized, as follows:

- WASH data are compositional, and thus they should not be modelled with standard statistical analysis. Log-ratio transformations designed for compositional data lead to more conceptually sound estimates, especially in the occurrence of coverage rates near 0% or 100%.

- OLS regression may underestimate or overestimate coverage of WASH services when coverage data show non-linear patterns such as acceleration, deceleration and saturation. Non-linear methods such as GAM are alternative to account for the non-linear trajectories in WASH access.

- Standard errors of survey data can be approximated with our approach, but to obtain a more accurate measure of the magnitude of uncertainty around WASH estimates, more efforts should be made to include errors in household survey reports and the JMP global database.

4

Levels and trends in child mortality

Abstract:

Child mortality is a matter of great concern to the global community. In the context of decision-making, trend analysis of two child mortality indicators – neonatal (NMR) and under-five (U5MR) mortality rates – is key to evaluate countries' progress and identify what works for effective public-health policy-making. The estimation of these child mortality indicators is, however, challenging for the great majority of developing countries, where vital registration systems are often incomplete and/or unreliable. Therefore, models are required to construct NMR and U5MR estimates. In this Chapter, we present and compare two approaches – based on logit and isometric log-ratio transformations of the data – to monitor progress under uncertainty. We apply them to the case of child mortality in sub-Saharan Africa. Our analysis show that, albeit both approaches lead to similar NMR and U5MR estimates, only the isometric log-ratio transformation designed for compositional data produces conceptually sound results, where child mortality components fulfill the constant-sum constraint.

Keywords: Trend analysis; Child mortality; SDG 3; Household Surveys

This chapter is based on:

- Ezbakhe, F. and Pérez-Foguet, A., 2019. Levels and trends in child mortality: a compositional approach. *Demographic Research* (Under Review)

4.1 Introduction

The ongoing decline in child mortality is considered one of the most important successes in public and population health in the past three decades. Death of children under five years old has fallen from 12.6 million per year in 1990 to 5.4 million per year in 2017, even as the world's population under age five grew by nearly 30 million (UNICEF 2018; UNESA 2017). Notwithstanding this substantial progress, there is still a heavy burden of child deaths due to preventable or treatable causes. This has both social and economic consequences: in the WHO African region alone, the cost of child mortality amounted to 150.3 billion USD in 2013 (i.e., approximately 6% of the combined GDP in the region) (Kirigia et al. 2015). In recognition of the crucial need to further combat child mortality, the third Sustainable Development Goal (SDG 3) of the 2030 Agenda specifically calls for countries to "ensure healthy lives and promote well-being for all at all ages" (UNGA 2015). In particular, Target 3.2 specifies the end of preventable deaths of newborns and children under five by lowering the neonatal and under-five mortality rates to at least 12 and 25 deaths per 1,000 live births by 2030, respectively.

Achieving this ambitious child survival target goes beyond ensuring universal access to effective, good-quality and affordable health care for women and children. It also requires understanding the levels and trends in child mortality in order to evaluate countries' performances and identify effective policies (UNICEF 2018). That is why measuring and monitoring child mortality is a global priority. However, tracking progress towards reducing child mortality can be challenging, particularly in developing countries with dysfunctional vital registration systems. According to Mahapatra et al. (2007), vital statistics are unavailable or of poor quality in 111 countries, mainly in Sub-Saharan Africa, South-East Asia and Western Pacific, which represent 72% of the world's population. This lack of reliable data inevitably takes a toll on the effectiveness of public-health policy-making.

To overcome the absence of reliable vital registration data in many countries, the United Nations Inter-agency Group for Child Mortality Estimation (UN IGME) produces and publishes estimates of child and young adolescent mortality rates every year (IGME 2018). Child mortality indicators are provided for three different age intervals: neonatal mortality ratio (NMR), i.e., the number of deaths within the first 28 days of life per 1000 live births; the infant mortality ratio (IMR), i.e., the number of deaths among children under age of 1 year per 1000 live births; and under-five mortality ratio (U5MR), i.e., the number of deaths of children up to age of 5. With these input data, the UN IGME generates child mortality estimates for years of interest using a Bayesian B-splines Bias-adjusted (B3) regression model (Alkema and New 2014; Alkema et al. 2014; Alexander and Alkema 2018). In addition, the UN IGME's B3 model also adjusts the errors and biases in the data.

However, besides accounting for the inherent uncertainty in child mortality data, there is also a compositional nature that should be considered when analyzing the data. Strictly speaking, child mortality indicators are not rates but probabilities calculated according to the conventional life-table approach (Rutstein 1984) and are thus naturally constrained. Indeed, the sum of probabilities of dying in the neonatal (0–28 days), post-neonatal (29–364 days) and childhood (1–4 years) age intervals and the probability of surviving beyond 5 years must equal 1. This constant sum constraint makes it not possible to follow the usual Euclidean geometry, since data belong to a subspace of the Euclidean space, as explained in Chapter 3.2. Therefore, child mortality variables must not be analyzed separately, as this may produce spurious correlations and misleading results. In the B3 model, the compositional nature of data is accounted for to some extent by considering logarithms and ratios. For instance, the U5MR is modelled in the log-scale (i.e., log(U5MR)). For the IMR, the model is fitted to the log-odds transform of the ratio r between the IMR and the median B3 model estimates of U5MR (i.e., log(r/(1-r))). And for the NMR, it models the ratio between NMR to the difference between U5MR and NMR (i.e., NMR/(U5MR-NMR)). However, in order to be theoretically sound with respect to the principles of compositional data analysis, log-ratio transformations between the compositional parts are needed.

The literature on the application of compositional analysis to mortality data is extensive. For instance, Oeppen (2008) explored the use of centered log-ratio transformation for forecasting mortality by cause of death. Similarly, Salomon and Murray (2001) developed a compositional model based on additive log-ratios to predict cause-of-death patterns by age and sex; and Bergeron-Boucher et al. (2017,

2018) also applied centered log-ratios to coherently forecast the distribution of deaths of sub-populations. Other researchers have focused on applying isometric log-ratios to model trends in health-related outcomes (Carson et al. 2016; Fairclough et al. 2017). However, the modelling of composition trends in child mortality have yet to be fully explored.

In this context, our aim is to assess the application of compositional data analysis for estimating trends in child mortality, with their associated uncertainty. Specifically:

- We estimate child mortality with two data transformations – logit and isometric log-ratio – and compare them with the official estimates provided by the IGME.
- We use all publicly available household survey data on the two child mortality indicators used in SDG 3 monitoring: neonatal (NMR) and under-5 (U5MR) mortality indicators.
- We apply the trend analysis to the countries of sub-Saharan Africa. We select this region because it accounted for nearly 54% of global under-5 deaths in 2015 (Wang et al. 2017).

The remainder of the Chapter is structured as follows. In Section 4.2 we present the case of child mortality monitoring, in particular the use household survey data. We then present an overview of the method and data used in our analyses in Section 4.3. In Section 4.4, we present the results from applying the two data transformations to child mortality, altogether highlighting the differences between our estimates and those provided by the IGME. In Section 4.5, we conclude the Chapter summarizing the main outcomes of our analyses.

4.2 The case of child mortality monitoring

The responsibility for monitoring and assessing child mortality at the global, regional and country level lies within the United Nations Children's Fund (UNICEF). Together with other members of the UN Inter-agengy Group for Child Mortality Estimation (IGME), UNICEF estimates child mortality every year for monitoring progress and shares them in their public database (http://www.childmortality.org).

To do so, the UN IGME first reviews and compiles all available nationally representative data relevant to the estimation of child mortality – including data from civil registration systems, population censuses and household surveys – and assesses their quality to exclude those with substantial errors.

The most reliable data sources for child mortality monitoring are civil registration and vital statistics (CRVS) systems, in which all births and deaths are routinely registered and certified. Unfortunately, there are not comprehensive CRVS systems in most developing countries. In the absence of continuous recording systems, measures of child mortality are derived from alternative data sources, most notably periodic, nationally representative household surveys (Hill et al. 2015), using both direct or indirect methods. Direct estimation approaches collect child mortality from full birth histories (FBHs) of women in reproductive age (i.e., 15 to 49 years old). In a FBH, women are asked to report on the date of birth, sex, survival status, age (if alive) and age at death (if dead), for each of their births. Probabilities of dying in childhood are then computed based on synthetic cohort life tables (Rutstein and Rojas 2006). However, this approach is time-consuming and expensive due to the extensive questionnaires and training of interviewers. Indirect estimation methods, on the other hand, use summary birth histories (SBHs), whereby women only report the total number of children they have given birth to and the number who have died – or equivalently the number still alive – at the time of the survey. Thus, instead of a full distribution of births and deaths over time as in FBHs, SBHs only provide the proportions of children dead at the time of the survey. In SBHs, probabilities of dying in childhood are derived from modeled relationships between the proportions of children dead and the age of the women (Brass 1975; Zlotnik and Hill 1981).

As seen in Table 4.1, most countries turn to household surveys to collect data on child mortality. Indeed, household surveys account for 91.1% and 79.4% of the total number of data series compiled for NMR and U5MR, respectively; which represent 40.2% and 55.7% of all mortality observations in the database. Amongst the most common household surveys are Demographic and Health Surveys (DHS), providing 59.8% and 52.4% of household data in NMR and U5MR, respectively. Two other important features can be highlighted when examining the child mortality database. First, the UN IGME excludes a significant number of data points from their estimation process (an average of 10.5% and 43.9% for NMR and U5MR, respectively), because of their substantial degree of non-sampling errors or omissions. Second, information on sampling errors in NMR observations is missing in more than 15% of the household surveys, and nearly 62% in U5MR observations.

Table 4.1: Data availability for neonatal (NMR) and under five (U5MR) mortality ratios from the UN IGME public database (extracted on May 28th, 2019). Other MICS includes National MICS; and other DHS includes Interim DHS, Special DHS, National DHS, World Fertility Survey, Malaria Indicator Survey and AIDS Indicator Survey.

Indicator	Type	Source	Number of countries	Number of series	Number of observations	% excluded by IGME	% with unreported standard errors
NMR	Vital Registration	VR	115	20	3320	29.2	-
		SVR	4	4	91	1.1	-
	Censuses	CEN	0	0	0	0	-
	Household surveys	MICS	37	22	222	27.0	16.7
		Other MICS	0	0	0	0.0	0.0
		DHS	90	73	1372	8.1	4.6
		Oher DHS	72	49	362	14.1	21.5
		LSMS	0	0	0	0	0
		Other surveys	59	103	337	15.4	47.2
U5MR	Vital Registration	VR	136	78	5619	42.5	-
		SVR	4	7	140	21.4	-
	Censuses	CEN	137	128	2332	48.1	-
	Household surveys	MICS	79	58	1008	44.5	21.4
		Other MICS	4	3	24	16.7	100.0
		DHS	90	176	5333	40.1	22.5
		Oher DHS	96	123	1465	39.9	37.8
		LSMS	1	1	1	100.0	100.0
		Other surveys	130	462	2349	41.6	88.7

Non-sampling errors may arise due to many different factors, including non-coverage, non-response, poor quality questionnaire, or defective survey implementation and data processing (Lesser and Kalsbeek 1999). Typical errors in data collection, such as under-reporting of deceased children (specially of neonatal deaths) or misreporting of ages at death (in particular age heaping around age 1) (Guillot et al. 2012).

While non-sampling errors can be minimized in many ways (e.g., proper design of survey questionnaire and data collection), sampling errors will always exist as the sample size is always smaller than the population size. Sampling errors for child mortality estimates can be quite large. A review by Korenromp et al. (2004) of sampling errors from Demographic and Health Surveys in various sub-Saharan African countries revealed median relative errors of 5.6% and 4.4% for infant and under-five mortality, respectively. This considerable amount of uncertainty is mainly due to the fact that "most household surveys are not designed to produce highly accurate estimates of child mortality, but rather aim for high accuracy of a number of other indicators" (UNESA 2011). For instance, in Multiple Indicator Cluster Surveys, child mortality rate is not selected as a key indicator on which to base the calculation of the sample size. This is because the sample sizes that would be required to measure child mortality indicators – with the same precision as recommended for other indicators – are too large and would be impractical. Other indicators such as immunization coverage are recommended instead (UNICEF 2006).

That is why, in addition to further minimize these errors, uncertainty assessment of the estimates – in the form of uncertainty intervals, for example – is indispensable for an evidence-based analysis of child mortality levels and trends. Failure to conduct and report such uncertainty intervals may lead to misinterpretation of rates and trends, and ultimately undermine effective policy-making for child mortality reduction.

4.3 Methodology

In this section, we first describe our method for the estimation of child mortality and its associated uncertainty (4.4.1). We then present the data used to estimate child mortality in sub-Saharan countries (4.4.2).

4.3.1 Proposed approach

We estimate the NMR and U5MR for each country during 1990-2018 – or earlier if data were available – by the use of a Generalized Additive Model (GAM) (as in Chapter 3.2). Data are considered 3-part compositions, $\boldsymbol{x} = (x_1, x_2, x_3)$, where:

$$x_1 = \mathrm{NMR}/1000; \quad x_2 = (\mathrm{U5MR} - \mathrm{NMR})/1000; \quad x_3 = 1 - \mathrm{U5MR}/1000 \tag{4.1}$$

We perform two data transformations: (i) logit transformation, where we fit the model to the logarithm of the odds (as in Equation 4.2), and (ii) isometric log-ratio (ILR) transformation, in which we fit it to the $D-1$ balances (as in Chapter 3.2).

$$z_i = log(\frac{x_i}{1-x_i}) \quad i = 1, 2, 3 \tag{4.2}$$

For ILR, we create a SBP that mimics the UN IGME's NMR/(U5MR-NMR) ratio:

Order	x_1	x_2	x_3	r	s
1	+1	+1	-1	2	1
2	+1	-1	0	1	1

With this established SBP, the orthonormal basis is:

$$\boldsymbol{V} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \\ -\sqrt{2/3} & 0 \end{bmatrix}$$

The ILR coordinates can be computed, following Equation 3.3 from Chapter 3.2, as:

$$z_1 = \sqrt{\frac{2}{3}}\log\left(\frac{\sqrt{x_1x_2}}{x_3}\right) = \sqrt{\frac{2}{3}}\log\left(\frac{\sqrt{\mathrm{NMR}\cdot(\mathrm{U5MR}-\mathrm{NMR})}}{1000-\mathrm{U5MR}}\right) \tag{4.3}$$

$$z_2 = \frac{1}{\sqrt{2}}\log\left(\frac{x_1}{x_2}\right) = \sqrt{\frac{2}{3}}\log\left(\frac{\mathrm{NMR}}{\mathrm{U5MR}-\mathrm{NMR}}\right) \tag{4.4}$$

The regression estimates are back-transformed to the original space using the inverse logit function (i.e., $e^{z_i}/(1+e^{z_i})$) and the inverse ILR shown in Equation 3.4 from Chapter 3.2. The NMR and U5MR estimates are finally derived from x_1 and x_2.

Confidence intervals of the mortality rates are constructed via simulation techniques. This involves generating 1,000 simulations of the survival probability for each age-group – neonatal and under five – assuming a Binomial distribution, $B(p, n)$[1]. The survival probability p is computed as $1 - \mathrm{MR}$ and the sample size n is derived from the standard errors of the ratios (i.e., $se = \sqrt{\mathrm{MR}(1-\mathrm{MR})/n}$), where MR is the mortality ratio (i.e., NMR/1000 and U5MR/1000, respectively). Finally, the 90% confidence intervals are obtained from the 5th and 95th percentiles of the simulations.

4.3.2 Child mortality data

In our analysis, we only consider data from censuses and household surveys series, and those series deemed of good quality by the IGME. In addition, for data with unreported sampling standard errors, we impute an error of 2.5% for census observations and 10% for those from household surveys, as done by Alkema and New (2014). On the other hand, since we consider data as 3-part compositions, only data-series with complete information on both ratios, NMR and U5RM, are included in the analysis. This has a major impact on the amount of data incorporated into the modelling procedure, since there is notably less data available for NMR than U5MR (i.e., 247 year series with 2,293 observations for U5MR versus 823 series with 10,180 observations for NMR).

The number of compositional data points considered for each of the 48 countries of sub-Saharan Africa are shown in Table 4.2. On average, only 18% of the observations include both NMR and U5MR. A clear example of this is Senegal: from the 44 and 73 time series for NMR and U5MR, respectively, only 34 included information for both ratios (resulting in 41 observations instead of the 186 available for U5MR). Furthermore, there are 9 countries with less than 4 observations for both NMR and U5MR – Central African Republic, Comoros, Djibouti, Equatorial Guinea, Gabon, Gambia, Seychelles, Sierra Leone and South Sudan – that are excluded from the analysis because of their lack of sufficient data.

[1] We use a Binomial distribution to model child mortality because it is the preferred distribution for dealing with counts (in this case, the number of deaths of children). In Chapter 3, we assumed a generalized Beta distribution to model WASH coverage because the two location parameters (a_1 and a_2) allow us to specify the support of the distribution and consider the rounding-off error.

Table 4.2: Data availability for neonatal (NMR) and under five (U5MR) mortality ratios from household surveys and censuses in countries of sub-Saharan Africa. "Years" represents the number of year series and "Points" the number of data observations.

COUNTRY		NMR Years	NMR Points	U5MR Years	U5MR Points	BOTH Years	BOTH Points
AGO	Angola	5	5	23	24	5	5
BEN	Benin	15	25	65	105	14	21
BWA	Botswana	9	9	28	28	4	4
BFA	Burkina Faso	17	23	67	110	15	19
BDI	Burundi	15	15	49	50	5	5
CPV	Cabo Verde	8	8	16	16	5	5
CMR	Cameroon	25	30	68	92	19	20
CAF	Central African Republic	5	5	32	32	3	3
TCD	Chad	12	15	49	57	8	8
COM	Comoros	5	5	10	10	0	0
COG	Congo	13	13	28	28	6	6
CIV	Cote d'Ivoire	25	25	68	83	14	14
COD	Democratic Republic of the Congo	10	10	28	28	10	10
DJI	Djibouti	6	6	16	16	1	1
GNQ	Equatorial Guinea	3	3	10	10	3	3
ERI	Eritrea	13	13	30	34	4	4
ETH	Ethiopia	12	20	47	77	6	10
GAB	Gabon	6	6	8	8	2	2
GMB	Gambia	0	0	27	28	0	0
GHA	Ghana	26	43	88	115	11	12
GIN	Guinea	21	21	58	78	10	10
GNB	Guinea-Bissau	10	10	18	18	10	10
KEN	Kenya	19	35	87	127	13	21
LSO	Lesotho	12	20	55	58	5	5
LBR	Liberia	15	20	57	63	4	4
MDG	Madagascar	12	20	49	68	7	10
MWI	Malawi	28	35	82	164	24	25
MLI	Mali	16	20	41	94	16	20
MRT	Mauritania	17	20	52	71	12	12
MOZ	Mozambique	11	15	44	76	10	14
NAM	Namibia	20	20	26	29	5	5
NER	Niger	23	25	45	88	20	20
NGA	Nigeria	16	25	45	75	7	15
RWA	Rwanda	21	29	60	108	17	23
STP	Sao Tome and Principe	10	10	24	24	10	10
SEN	Senegal	44	60	73	186	34	41
SYC	Seychelles	0	0	10	10	0	0
SLE	Sierra Leone	5	5	68	69	3	3
SOM	Somalia	5	5	10	10	5	5
ZAF	South Africa	10	10	24	24	7	7
SSD	South Sudan	1	1	13	13	1	1
SDN	Sudan	17	17	88	93	6	6
SWZ	Swaziland	15	15	34	36	8	8
TGO	Togo	10	15	58	70	5	7
UGA	Uganda	23	35	68	126	13	19

4.4 Results and Discussion

In this section we present and analyze the child mortality obtained with the two data transformations (4.4.1), comparing them with those provided by UN IGME (4.4.2). We also discuss some policy implications for the global monitoring of child mortality.

4.4.1 LOGIT and ILR data transformations

Figure 4.1 shows the comparison of the child mortality estimates obtained with the two data transformations, logit and isometric log-ratio, for Malawi.

The LOGIT data transformation provides three univariate coordinates, each representing the log odds of the mortalities and survival probabilities. In the case of Malawi, the log odds for mortality under 1 month of age (z_1) and mortality from ages 1 to 5 years (z_2) display decreasing values (between 1970 and 2018, z_1 and z_2 declined by 1.3 and 1.9 points, respectively); whereas log odds for survival beyond age 5 (z_3) increase by 1.9 points. This indicates both a reduction in child mortality and improvement of life expectancy over time.

On the other hand, the ILR data transformation results into two multivariate coordinates: z_1 captures all relative information about the survival ratio, while the z_2 captures the relationship between the mortalities under 1 month and from ages 1 month to 5 years. In Malawi, z_1 declines almost constantly from -1.2 in 1970 to -2.7 in 2018, while in z_2 there is no change in the first 23 years, but starts increasing from 1993 onwards. These values indicate not only a decline in child mortality over the years, but specifically in the ratio of mortalities under 1 month and between 1 month and 5 years from 1993.

For both data transformations, the resulting mortality components (i.e., x_1, x_2 and x_3) are essentially the same. For instance, in Malawi, the estimates for neonatal mortality (x_1) obtained with the LOGIT approach are 46.5, 37.4, 28 and 23.3 per mil for 1990, 2000, 2010 and 2018, respectively; with ILR, these estimates are 46.5, 37.5, 27.9 and 22.9.

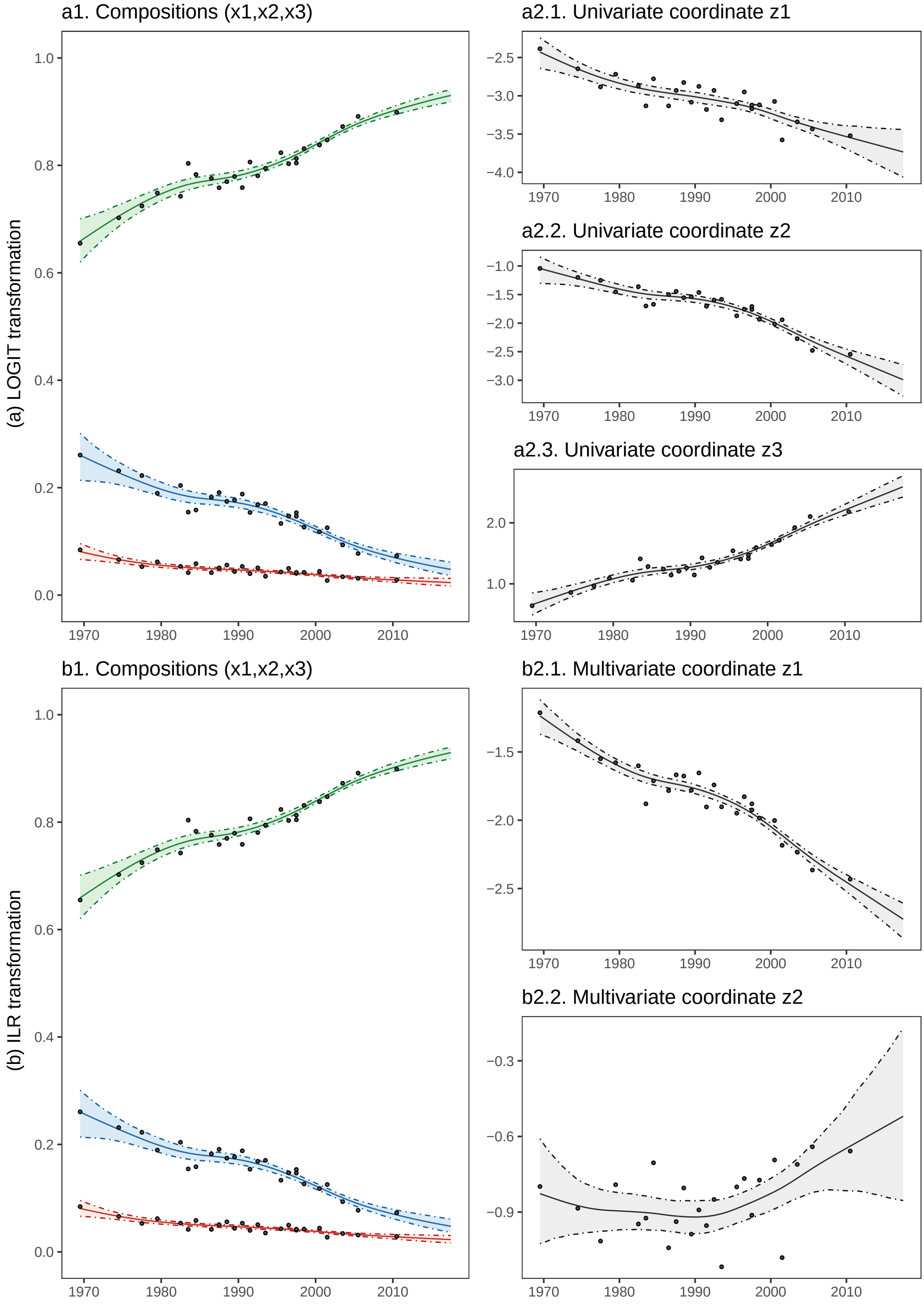

Figure 4.1: Mortality estimates for Malawi, with logit (a) and isometric-log ratio (b) transformations. In black, the regression results in the coordinates: the three univariate coordinates for logit transformation (in a2.1, a2.2 and a2.3) and the two multivariate coordinates for the ilr transformation (in b2.1 and b2.2). In color (a1 and b1), the regression results in the original scale: in red, mortality under 1 month (x_1); in blue, mortality between 1 month and 5 years (x_2); in green x_3 survival beyond 5 years.

A closer analysis of the differences between LOGIT and ILR estimates for all countries (Figure 4.2) reveals that the average difference is of 0.9 points per mil, which is negligible for all practical purposes. The maximum differences are found in Liberia, reaching 5.28 and 6.03 per mil for the NMR and U5MR, respectively. Furthermore, although difficult to appreciate, the differences in U5MR are slightly higher than in NMR. This is because U5MR is obtained from the sum of x_1 and x_2, while NMR is directly x_1. In addition, the differences between ILR and LOGIT are nearly zero for years with data available, since estimates are close to the observed mortality ratios.

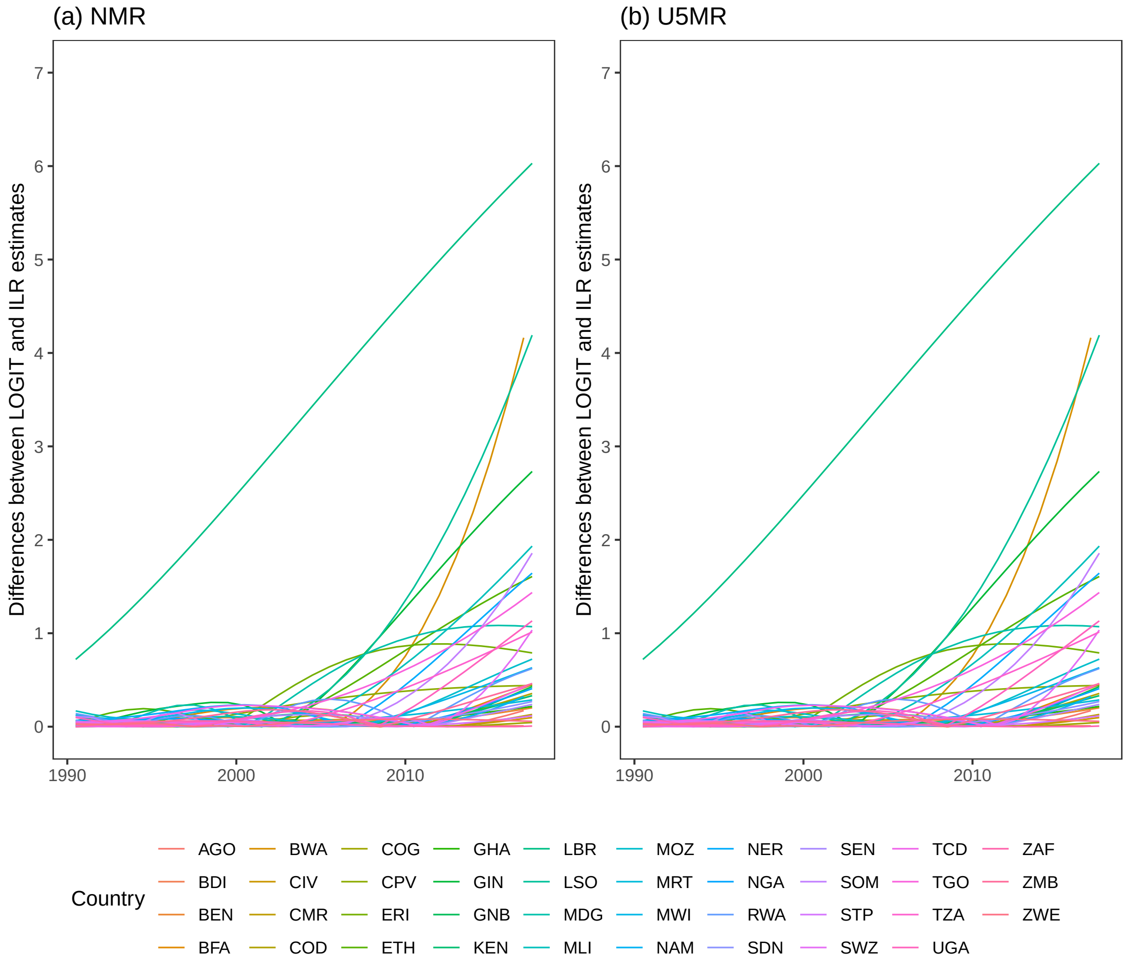

Figure 4.2: Differences, in absolute value, between estimates obtained with logit and isometric-log ratio transformations for (a) neonatal (NMR) and (b) under-5 (U5MR) mortality ratios. The countries with the highest differences (i.e., more than 2 points per mil in 2018) are Liberia (LBR), Botswana (BWA), Lesotho (LSO) and Guinea (GIN).

However, although the resulting mortality estimates are essentially identical, only the ILR approach strictly fulfills the unit-sum constraint (Figure 4.3). In the majority of countries, the sum of mortality estimates with LOGIT is higher than one. In Lesotho, Botswana, Somalia and Liberia, for instance, it reaches 1.0951, 1.0584, 1.0305 and 1.0113, respectively (i.e., 95.1, 58.4, 30.5 and 11.3 deaths per 1,000 live births).

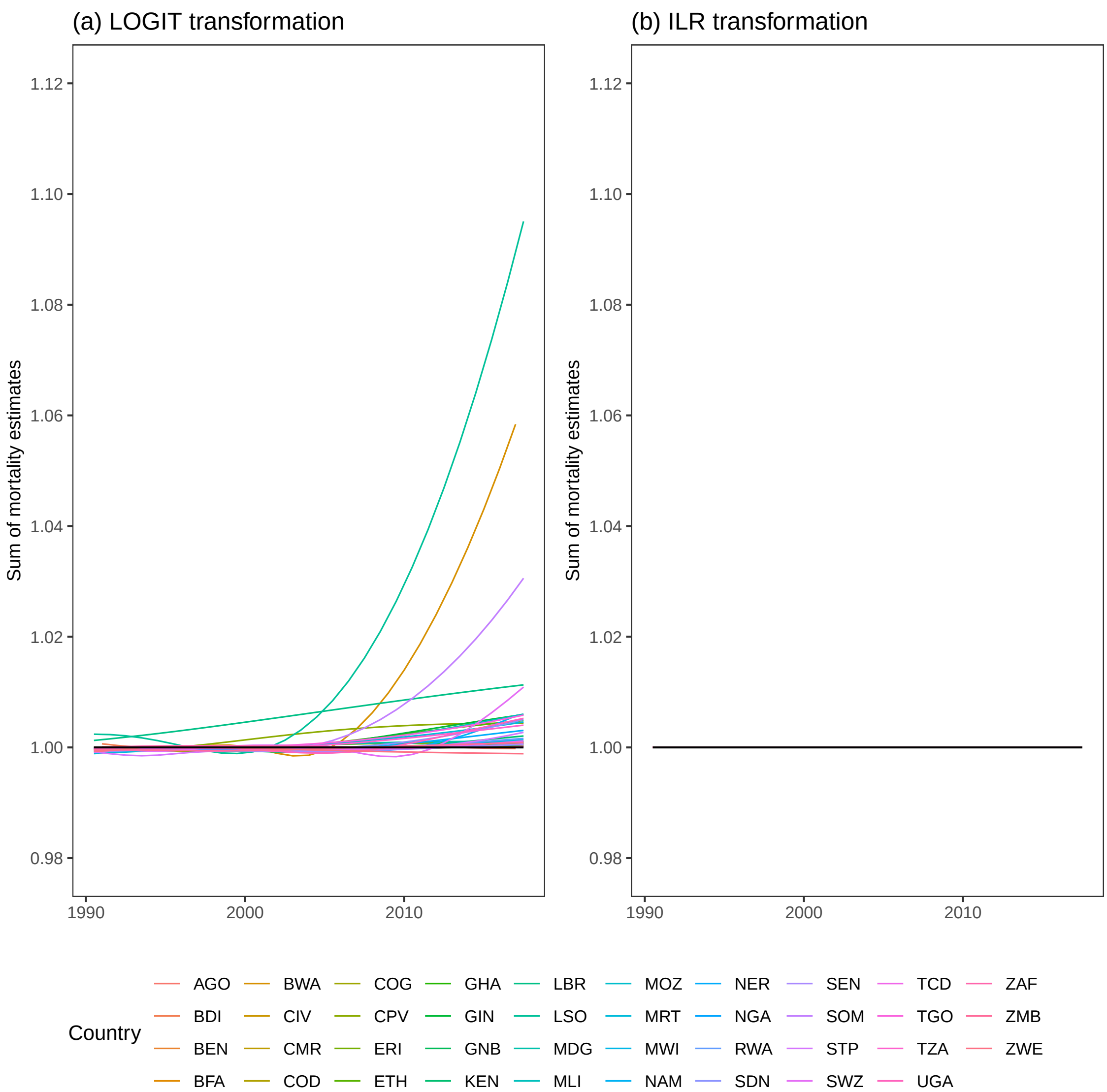

Figure 4.3: Sum of mortality estimates with (a) logit and (b) isometric-log ratio transformations. Notice that only the ILR transformation fulfills the unit-sum constraint, whereas the LOGIT transformation results in sums higher than one, especially in the countries of Lesotho (LSO), Botswana (BWA), Somalia (SOM) and Liberia (LBR).

This non-unit sum is even greater for the 90% confidence intervals (Figure 4.4). In 2018, for example, the 90% confidence intervals of the sum of mortality estimates in these four countries are: (1.0015-1.3463) in Lesotho, (1.0060-1.2003) in Botswana, (0.9745-1.5971) in Somalia and (1.0015 -1.1987) in Liberia. In contrast, the sum of the estimates based on ILR-transformed data adds up to one in all cases.

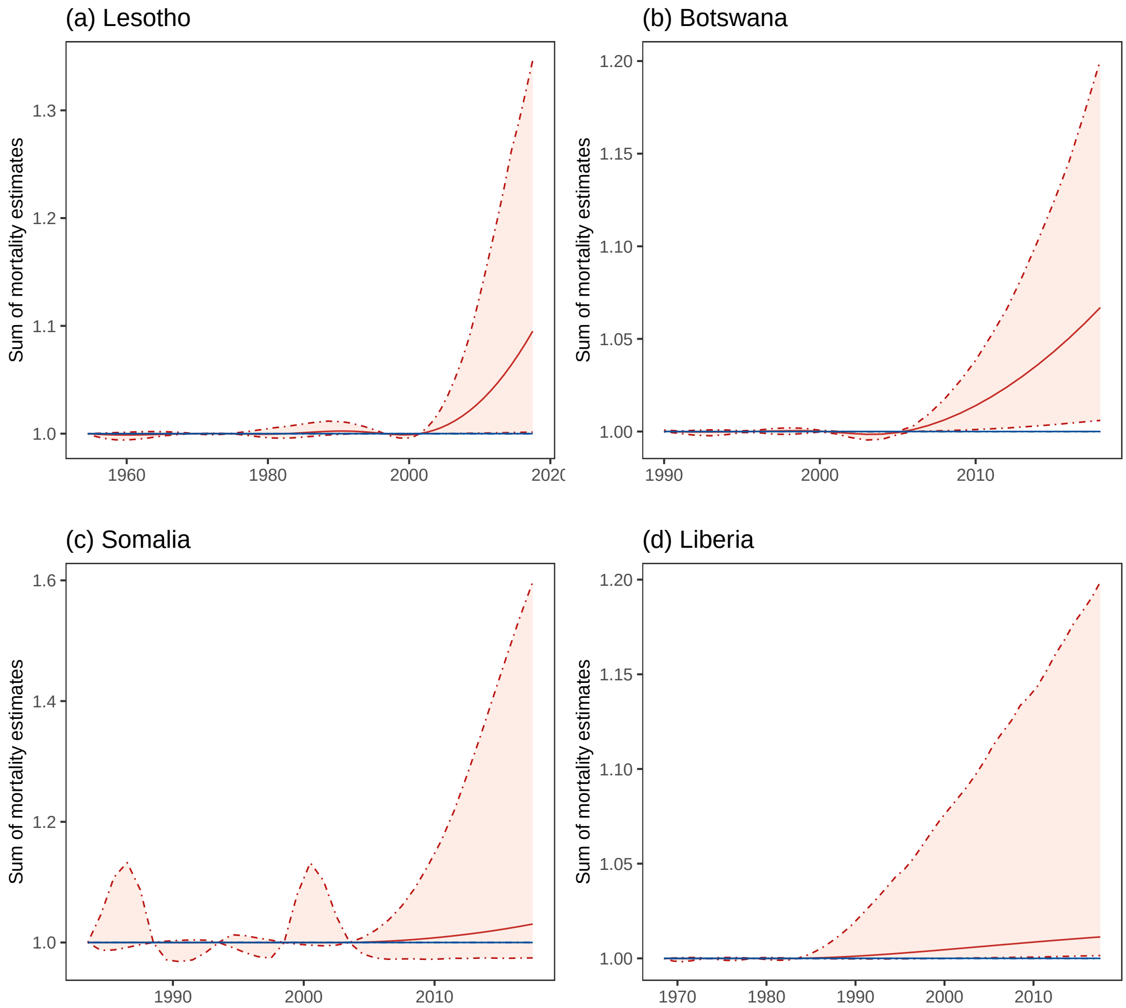

Figure 4.4: Sum of mortality estimates with logit and isometric-log ratio transformations in: (a) Lesotho, (b) Botswana, (c) Somalia and (d) Liberia. In red, the LOGIT estimates; in blue, the ILR estimates. The shaded areas represent the 90% confidence intervals. Notice that only the ILR transformation (in blue) fulfills the unit-sum constraint for both observed and simulated data. In years with data, the LOGIT transformation (in red) provides estimates closer to the unit sum (thus the narrower confidence intervals).

4.4.2 Official child mortality estimates

Figure 4.5 shows the comparison of the NMR and U5MR obtained with ILR and those provided by UN IGME for Malawi, Mauritania, Lesotho and Liberia. As expected, estimates differ greatly between the two approaches. This happens for two main reasons. First, in our model, we fit a simple GAM regression model instead of a Bayesian B-splines approach, and we do not adjust for bias due to AIDS[1]. Second, since we follow a compositional approach, our regressions are only based on year series with observation for both NMR and U5MR. In the case of Malawi, for example, our estimates are obtained with 25 data points, whereas the IGME regression model uses all 164 and 35 data points available for the estimation of NMR and U5MR, respectively. This becomes more evident in countries like Lesotho and Liberia, where the ILR estimates are based on merely 5 and 4 data points. These few observations may lead to completely different NMR and U5MR trends. For instance, in Lesotho, ILR estimates show a substantial increase in U5MR from 2000 onwards, mainly because it excludes later observations where U5MR falls below 125 deaths.

This lack of data leads to wider confidence intervals. In the case of Liberia, our estimates in 2018 are 40.4 for NMR and 84.8 for U5MR, with 90% confidence intervals of (5.5-192.4) and (199.7-534.2). This translates into interval widths of 187.0 and 514.5 deaths per 1,000 live births for NMR and U5MR, respectively. On the contrary, the widths of IGME's confidence intervals are only 25.3 and 51.3 deaths per 1,000 live births for NMR and U5MR, respectively. However, as seen in Section 4.4.1, the LOGIT transformation used in IGME' B3 model do not guarantee that the three child mortality probabilities will sum up to 1. As Lloyd et al. (2012) explain, "the LOGIT transform is a valid to analyze a two-part composition [but] has serious problems when there are more than two parts". Only the ILR transformation leads to estimates that are theoretically sound with respect to the most important properties of compositional data.

[1]In populations severely affected by HIV/AIDS (i.e., those where the prevalence reaches 5% of the adult population), there is a correlation between the mortality risks of mothers and their children: HIV-positive children will be more likely to die than other children, and will be less likely to be reported since their mothers will have been more likely to die also. Therefore, child mortality estimates will be biased downwards. That is why the IGME adjusts for bias due to AIDS in child mortality estimation.

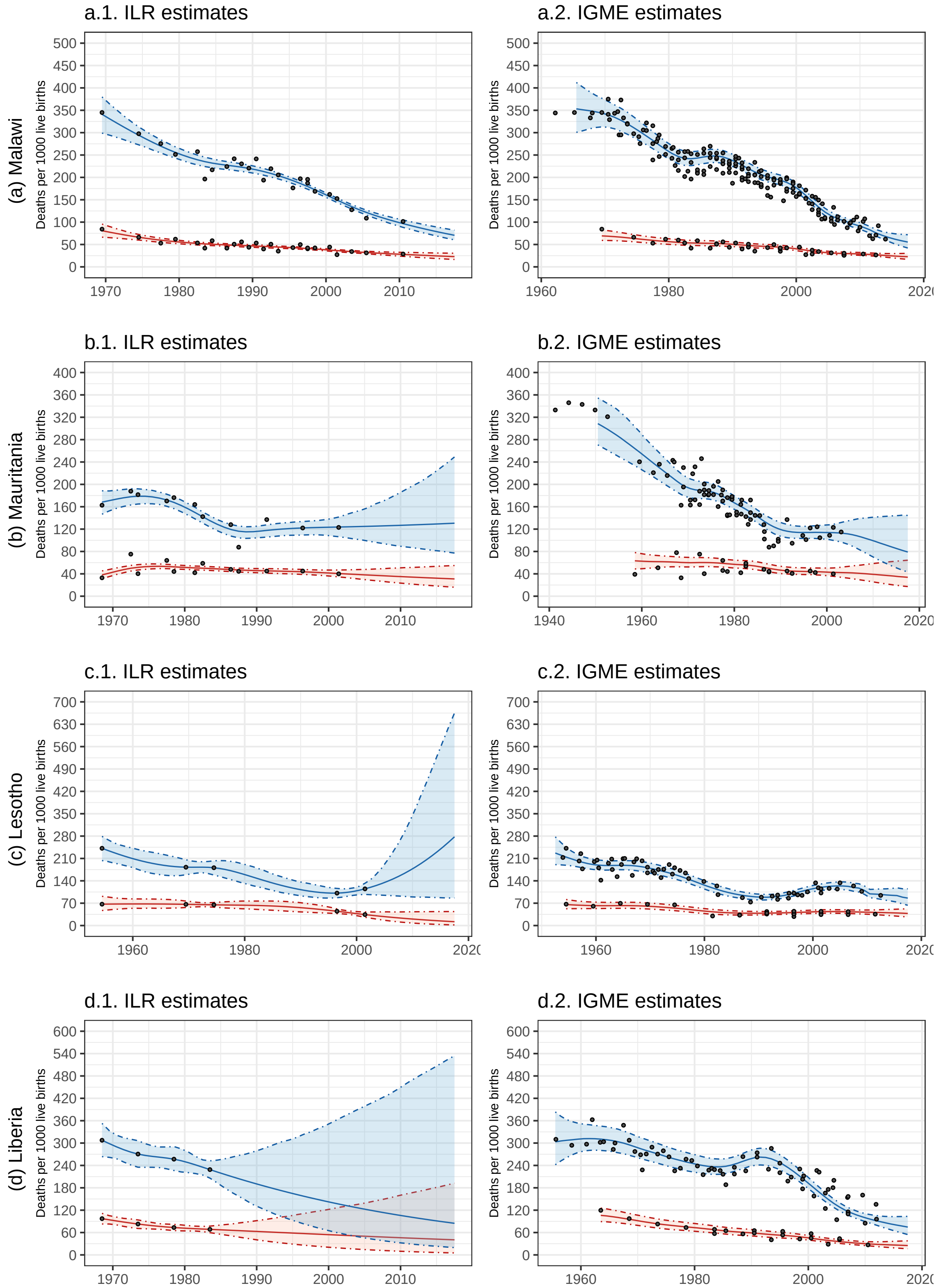

Figure 4.5: Neonatal (NMR) and under-five (U5MR) mortality estimates for: (a) Malawi, (b) Mauritania, (c) Lesotho and (d) Liberia. In red, neonatal mortality estimates; in blue, under-five mortality estimates; in black, the mortality data points. The shaded areas represent the 90% confidence intervals. In Figures (x.1), estimates obtained with the isometric-log ratio transformation, and in (x.2) those provided by UN IGME's model. Notice that, in ILR, only year series with observation for both NMR and U5MR are considered, which leads to less data points for the regression analysis and consequently wider confidence intervals.

But our approach has one main drawback: time series with incomplete data must be excluded. We have seen that this a substantial impact on the number of observations available for analysis, since merely 18% of the time series data from the UN IGME database included observations for both NMR and U5MR in sub-Saharan Africa. The reason why there are less data available for NMR than for U5MR lies in the impossibility to indirectly estimate neonatal mortality from summary birth histories, as it is done for under-five child mortality (Burstein et al. 2018). In extreme cases such as Gambia, where there are no complete time series, the country is omitted from the analysis. In other cases, the limited data available for NMR implies that the regression is done with few observations and hence the resulting trends in child mortality should be taken with caution – especially when comparing it to trends provided by the IGME.

So the question is: should we prioritize more conceptually sound mortality estimates, even if it entails less data points in the regression analysis and therefore wider confidence intervals? The answer to this question is not that simple. On one hand, poor and theoretically inconsistent estimates lead to a misleading overview of the health situation and child mortality trends, specially when data are closer to the limits of their range (i.e., 0 or 1,000 deaths). On the other hand, a trend analysis based on few observations should be interpreted with care, as the risks of getting nonsensical results are relatively high.

In this sense, more efforts should be made to improve the data availability for neonatal mortality, as it represents the indicator with fewer observations. In particular, it is important to develop and validate new methods for the indirect estimation of neonatal mortality from summary birth histories, as it is currently done for under-5 mortality. As highlighted by Burstein et al. (2018), "[the] use of such methods allows research to utilize a massive amount of SBH data for estimation of trends in neonatal mortality" and consequently "further improve the evidence base for monitoring of trends and inequalities".

4.4.3 Implications for child mortality in sub-Saharan Africa

The 1990-2018 evolution of neonatal (NMR) and under 5 (U5MR) mortality ratios of sub-Saharan African countries are shown in Figure 4.6. Child mortality rates show substantial decline in the last 30 years in all countries throughout sub-Saharan Africa except for Botswana, Somalia and Sudan for the case of NMR, and Lesotho and Togo for U5MR. However, much of this rise in child mortality is due to extrapolation on the basis of very few data points. In Botswana, for instance, only 4 data points are included in the regression analysis, the most recent being in 2005. Furthermore, in nearly half of the countries the analysis is done with less than 10 observations (as seen in Table 4.2), which hinders the reliability of mortality estimates. Consequently, in countries with limited data, trends in child mortality should be taken with caution.

On the other hand, only three countries – Eritrea, Cabo Verde and Madagascar – (7.9% of all countries) already meet in 2018 the SDG 3 of reducing NMR and U5MR to 12 and 25 deaths per 1000 lives births by 2030, respectively. The distribution of NMR for the rest of countries is as follows: 44.7% of them are between the target value and tow-times the target value (i.e., between 12 and 24 deaths), 34.2% between two- and three-times the target value (24 and 36 deaths), and 13.2% with rates over triple the target. For U5MR, these figures are 23.7%, 26.3% and 42.1%, respectively, which presents far less hopeful figures of under-five mortality in sub-Saharan Africa.

Furthermore, a geographical analysis of child mortality estimates shows great disparities amongst regions in sub-Saharan Africa. The maps of the NMR and U5MR rates both show a concentration of high mortality regions mostly in Western and Central Africa. For instance, in 2018, the average neonatal mortality ratio in countries of Western and Central Africa is higher than 25 per mil, while the regional average in Southern and Eastern Africa – excluding Botswana and Somalia – is less than 18. One immediate policy implication can be drawn from this analysis: in order to achieve the SDG 3 targets for child mortality, policy-makers should dedicate more means to increasing access to and use of maternal and child care services, especially in Western and Central African countries. This requires a stronger political commitment from the national governments along with international support.

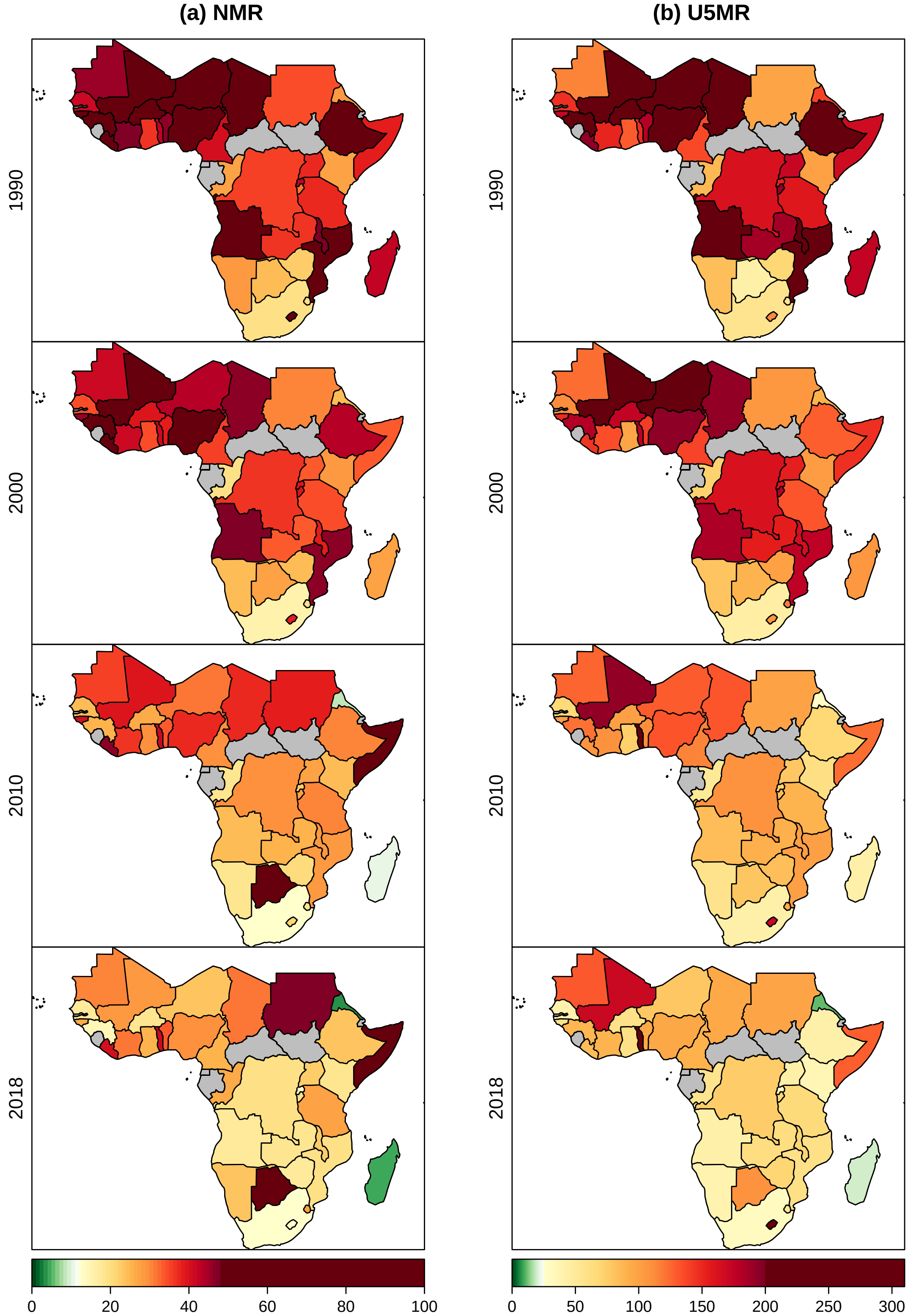

Figure 4.6: Evolution of mortality in sub-Saharan African for years 1990, 2000, 2010 and 2018: (a) neonatal (NMR) and (b) under-five (U5MR) mortality ratios. In green, the counties with mortality ratios lower than the SDG 3 target values of 12 and 25 deaths per 1,000 live births for NMR and U5MR, respectively. In red, the countries with mortality ratios greater than the target values. In grey, the countries with insufficient data that are excluded from the analysis.

4.5 Key messages

The estimation of child mortality is challenging for the great majority of developing countries, where vital registration systems are often incomplete and/or unreliable, and thus models are required to construct NMR and U5MR estimates for years of interest. In this Chapter, we consider and compare two data transformations, logit and isometric log-ratio, to produce these estimates during 1990-2018 – or earlier if data were available – for sub-Saharan African countries.

Three key messages can be summarized for this Chapter, as follows:

- While logit transformations are widely used in child mortality estimation (for instance in the UN IGME's B3 model), isometric log-ratio transformation designed for compositional data would lead to more conceptually sound results.
- One of the downfalls of the isometric log-ratio transformation is the need to exclude the time series with incomplete data (because there are less data for NMR than for U5MR), which leads to more uncertainty in the child mortality estimates.
- More efforts should be made to utilize the vast amount of summary birth history data available from household surveys and censuses to estimate NMR and help improve the data availability for this indicator and, ultimately, the compositional trend analysis of child mortality.

Conclusions

We laid out the content of this thesis along a pair of entwined dimensions, both aimed at providing quantitative answers to the integration of uncertainty in decision analysis for sustainable development. The first dimension – *"how can we include the uncertainty of the evidence in the prioritization of policy options for service provision?"* – deals with the development of a simplified non-compensatory multi-criteria decision analysis under uncertainty for the prioritization of alternatives. The second dimension – *"how can we incorporate this uncertainty in the trend analysis of service coverage for progress monitoring"* – focuses on a straightforward characterization of uncertainty in compositional data analysis to track progress.

Both uncertainty approaches are applied to real decision problems for sustainable development, in particular child mortality reduction (SDG 3), water, sanitation and hygiene targeting (SDG 6) and renewable energy planning (SDG 7).

Overall, this thesis offers a framework – together with a full set of methods and case studies – that can support the incorporation of uncertainty in decision analysis for the provision of public services for SDG implementation. We also provide some recommendations and future perspectives for policy-makers who intend to integrate uncertainty in their decision-making processes.

Main conclusions

The main takeaway of this thesis is that **uncertainty matters**. We have confirmed that uncertainty is inherent to the evidence used to support decision-making for the SDGs, in particular for public services delivery. In Chapter 1, we have seen that the performance scores of renewable energy alternatives provided by the decision-makers where not exact values, but rather wide intervals. In Chapters 2, 3 and 4, we have also seen that household survey data, which remain the main source of information for the planning of services provision, has a great deal of uncertainty due to the sampling process. However, this uncertainty is often ignored when utilizing these data. The case of water and sanitation monitoring (Chapter 3) is a clear example of this, as uncertainties relating to data are seldom measured nor included in the analysis by the official monitoring program.

Besides, **decisions made disregarding uncertainty are likely to be unreliable and misleading**. Indeed, our results show that prioritization and trend analysis of SDGs and related targets are inevitably inaccurate, due to both low availability and bad quality of data. In the case of prioritization (Chapters 1 and 2), it becomes evident that the numerical values assigned to the alternatives are not deterministic. Instead, they are drawn from performance intervals that influence the subsequent rankings. In trend analysis (Chapters 3 and 4), it is also apparent that the resulting estimates are not single points: they are a range of likely values that depend on the level of data uncertainty. Therefore, the interpretation of prioritization rankings and/or forecasted estimates must consider this uncertainty.

Furthermore, **uncertainty characterization is not a novelty, but simple and reliable uncertainty characterization methods are**. We have seen that uncertainty can be incorporated in decision analysis in a number of ways, from probabilistic frameworks to fuzzy set theory in prioritization problems (Chapters 1 and 2), and the construction of confidence intervals in trend analysis (Chapters 3 and 4). However, these methods are often too complex to be understood and used by decision-makers – and even analysts – unfamiliar with uncertainty assessment. Simplicity is especially important to ensure that uncertainty is constructively incorporated into sustainable development policy-making.

At last, **our approaches allow for a simple and practical integration of uncertainty into decision analysis**. In the case of prioritization (Chapters 1 and 2), our version of the ELECTRE III model incorporates uncertainty in a direct manner, by expressing the discrimination thresholds as a function of the performance scores bounds. At the same time, our ELECTRE III model overcomes the issue of compensation between the criteria that composite indexes are subject to. In trend analysis (Chapters 3 and 4), our regression model, based on compositional data theory, provides more theoretically sound estimates – with their confidence intervals – than standard approaches used by the official SDG monitoring mechanisms.

In practice, the simplicity of our approaches can potentially grant a twofold improvement. On one hand, our methods are versatile: decision-makers would be able to use a mix of evidence to make their decision for various SDGs. For instance, our prioritization model can be used with both expert knowledge (Chapter 1) and household survey data (Chapter 2). On the other hand, our methods do not undermine the transparency of the decision analysis process, which in turn would make decision-makers comfortable with applying them to real life problems.

However, although this thesis presents a comprehensive illustration of the incorporation of uncertainty in prioritization and trend analysis problems for sustainable development, it falls short of showing how our methods may be integrated into existing decision support systems for services provision. Indeed, despite the implementation of various case studies dealing with different settings (i.e., energy in Chapter 1, water and sanitation in Chapters 2 and 3, and health in Chapter 4), they are not sufficient to give conclusive information on whether our methods are useful to improve decision-making for sustainable development. The main reason for this relates to our limited resources, which restricted severely our ability to engage with the different actors involved in the decision analysis process.

Recommendations

Based on our results, we offer some recommendations and guidance for the incorporation of uncertainty in decision analysis, which we summarize as follows:

- **Put more emphasis on uncertainty**. Uncertainty needs to be mainstreamed in sustainable development in order to make it an integral part of decision analysis. Only then can uncertainties be identified, acknowledged and accounted for. Mainstreaming uncertainty not only involves reporting it, but also reflecting on questions such as: what are the main sources of uncertainty in sustainable development? How can these sources of uncertainty be reduced? What implications do they have in a given policy or decision context? And how can they be dealt with in the decision-support process?

- **Some uncertainty is better than *no* uncertainty**. Although a particular type or level of uncertainty seldom manifests itself in isolation, it is useful to begin by considering a single uncertainty source. This analysis will lead to the clear understanding of why this particular source of uncertainty matters in decision analysis and, consequently, to the identification and integration of other sources of uncertainties. In our approaches, even a basic consideration of uncertainty in the evidence provides insightful perspectives on the role (and impact) of uncertainty in decision analysis.

- **Fully document the methods**. One main constraint to decision analysis under uncertainty lies in the difficulty to apply existing approaches for uncertainty characterization. This is due to the fact that a comprehensive documentation of these methods is rarely reported. A meticulous documentation of methods can facilitate their replicability by other practitioners and researchers, which will feedback constructively into the methods themselves.

- **Simplicity is not a flaw, but a prerequisite**. While real world decision problems – especially under uncertainty – are rarely straightforward, it is important to keep the decision analysis model as simple as possible to guarantee transparency. Otherwise, decision- and policy-makers might perceive the model as a *black-box*, and consequently they may feel reluctant towards using its outcomes. For this reason, keeping the process simple to ensure a close interaction between decision analysts, decision-makers, and other stakeholders, is a more productive, dynamic, and efficient alternative.

Way forward

This thesis deals with the development of simple tools to enable the integration of uncertainty present in the evidence into decision analysis processes for sustainable development. It covers two important decision problems in public services delivery – the prioritization of alternatives and the monitoring of progress – and provides practical approaches for incorporating uncertainty in planning for SDG implementation. However, having simple uncertainty approaches does not necessarily imply their proper application for sound decision-making. To effectively encourage policy- and decision-makers in considering uncertainty in their decisions, other specific challenges need to be addressed.

First, we must **continue improving the methods for a better integration of uncertainty in decision analysis for sustainable development**. On one hand, given the importance of prioritization in policy domains, not only for services provision but also for academic performance, quality of life assessment or industrial competitiveness, future work will need to envision the improvement of multi-criteria analysis tools. In particular, research needs to focus on the computational efficiency of non-compensatory aggregation procedures, such as ELECTRE III, so they can be applied to more complex prioritization problems (i.e., more alternatives and/or criteria). On the other hand, future efforts should be directed towards extending compositional data approaches beyond the monitoring of public services provision. Indeed, compositional data analysis is necessary for accurate statistical modelling, and should be extended to other areas of global sustainable development. Furthermore, more research must be direct towards modelling compositional time-series with other trend analysis approaches (e.g., autoregressive or moving average models).

Second, we must **focus less on data acquisition, and instead focus more on the decision problems that need to be solved**. In a day and age when the "data revolution" is at the heart of the international community's approach to policy engagement and capacity development (IEAG 2015), there has been an explosion in the volume and the types of data available, stemming from new technologies such as the Internet Of Things or real-time collaborative platforms. However, with enormous amounts of data being produced constantly, it is crucial to take a step back and reflect on which data are required to support better decision-making. This calls for shifting the strategic question from *"How can we get more data for sustainable development?"* to *"What are the data needed to make the decisions for sustainable development?"*. Decision-makers who are implementing and planning for SDGs must thus structure their data gathering efforts around the practical decisions on the ground. This will enable them to identify critical uncertainties and evidence gaps and create incentives to close them.

Last but not least, **decision analysis must be embedded in decision-making**. More than often, decision analysts are viewed as researchers making sophisticated models in their academic work that are unsuitable for real-world practice. This must change: analysts need to work together with decision-makers to structure and evaluate the problem and use evidence to inform the policy-making. Although this will require governments and the international community to invest in bridging the gap between science and policy, a stronger collaboration between decision-makers and analysts will help increase the policies effectiveness and, ultimately, the sustainable deployment of key resources. As Shepherd et al. (2015) put it: "training a generation of decision analysts to work with policy-makers could do more for development than any other single intervention".

Addressing these challenges might not be an easy job, but we believe are the way forward for a systematic improvement in our decision-making processes that could dramatically help make a better use of evidence for sustainable development.

Bibliography

Abdullah, L. and Najib, L. (2016). Sustainable energy planning decision using the intuitionistic fuzzy analytic hierarchy process: Choosing energy technology in Malaysia. *International Journal of Sustainable Energy*, 35(4):360–377. doi: 10.1080/14786451.2014.907292.

Abu Taha, R. and Daim, T. (2013). Multi-Criteria Applications in Renewable Energy Analysis, a Literature Review. In *Research and Technology Management in the Electricity Industry*, pages 17–30. Springer. doi: 10.1007/978-1-4471-5097-8_2.

Agresti, A. and Coull, B. A. (1998). Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician*, 52(2):119–126. doi: 10.2307/2685469.

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):139–160. Retrieved July 19, 2019, from: https://www.jstor.org/stable/2345821.

Aitchison, J. (1999). Logratios and natural laws in compositional data analysis. *Mathematical Geology*, 31(5):563–580. doi: 10.1023/A:1007568008032.

Alexander, M. and Alkema, L. (2018). Global estimation of neonatal mortality using a Bayesian hierarchical splines regression model. *Demographic Research*, 38:335–372. doi: 10.4054/DemRes.2018.38.15.

Alkema, L. and New, J. R. (2014). Global estimation of child mortality using a Bayesian B-spline bias-reduction model. *The Annals of Applied Statistics*, 8(4):2122–2149. doi: 10.1214/14-AOAS768.

Alkema, L., New, J. R., Pedersen, J., and You, D. (2014). Child mortality estimation 2013: an overview of updates in estimation methods by the United Nations Inter-agency Group for Child Mortality Estimation. *PloS One*, 9(7):e101112. doi: 10.1371/journal.pone.0101112.

Antony, G. and Rao, K. V. (2007). A composite index to explain variations in poverty, health, nutritional status and standard of living: Use of multivariate statistical methods. *Public Health*, 121(8):578–587. doi: 10.1016/j.puhe.2006.10.018.

Armah, F. A., Ekumah, B., Yawson, D. O., Odoi, J. O., Afitiri, A.-R., and Nyieku, F. E. (2018). Access to improved water and sanitation in sub-Saharan Africa in a quarter century. *Heliyon*, 4(11):e00931. doi: 10.1016/j.heliyon.2018.e00931.

Atilgan, B. and Azapagic, A. (2017). Energy challenges for Turkey: Identifying sustainable options for future electricity generation up to 2050. *Sustainable Production and Consumption*, 12:234–254. doi: 10.1016/j.spc.2017.02.001.

Bain, L. J. and Engelhardt, M. (1987). *Introduction to probability and mathematical statistics*. Brooks/Cole.

Bain, R., Johnston, R., Mitis, F., Chatterley, C., and Slaymaker, T. (2018). Establishing sustainable development goal baselines for household drinking water, sanitation and hygiene services. *Water*, 10(12):1711. doi: 10.3390/w10121711.

Balin, A. and Baraçli, H. (2017). A fuzzy multi-criteria decision making methodology based upon the interval type-2 fuzzy sets for evaluating renewable energy alternatives in Turkey. *Technological and Economic Development of Economy*, 23(5):742–763. doi: 10.3846/20294913.2015.1056276.

Banda, J. P. (2003). Nonsampling errors in surveys. In *Expert Group Meeting to Review the Draft Handbook on Designing of Household Sample Surveys*. United Nations Secretariat. Retrieved July 19, 2019, from: https://unstats.un.org/unsd/demographic/meetings/egm/Sampling_1203/docs/no_7.pdf.

Banihabib, M. E., Hashemi-Madani, F.-S., and Forghani, A. (2017). Comparison of Compensatory and non-Compensatory Multi Criteria Decision Making Models in Water Resources Strategic Management. *Water Resources Management*, 31(12):3745–3759. doi: 10.1007/s11269-017-1702-x.

Baris, K. and Kucukali, S. (2012). Availability of renewable energy sources in Turkey: Current situation, potential, government policies and the EU perspective. *Energy Policy*, 42:377–391. doi: 10.1016/j.enpol.2011.12.002.

Bartram, J., Brocklehurst, C., Fisher, M., Luyendijk, R., Hossain, R., Wardlaw, T., and Gordon, B. (2014). Global monitoring of water supply and sanitation: history, methods and future challenges. *International Journal of Environmental Research and Public Health*, 11(8):8137–8165. doi: 10.3390/ijerph110808137.

Belton, V. and Stewart, T. (2002). *Multiple Criteria Decision Analysis: an integrated approach*. Springer US.

Bergeron-Boucher, M.-P., Canudas-Romo, V., Oeppen, J., and Vaupel, J. W. (2017). Coherent forecasts of mortality with compositional data analysis. *Demographic Research*, 37:527–566. doi: 10.4054/DemRes.2017.37.17.

Bergeron-Boucher, M.-P., Simonacci, V., Oeppen, J., and Gallo, M. (2018). Coherent modeling and forecasting of mortality patterns for subpopulations using multiway analysis of compositions: an application to Canadian provinces and territories. *North American Actuarial Journal*, 22(1):92–118. doi: 10.1080/10920277.2017.1377620.

Berk, I. and Ediger, V. Ş. (2018). A historical assessment of Turkey‘s natural gas import vulnerability. *Energy*, 145:540–547. doi: 10.1016/j.energy.2018.01.022.

Bermejo, R., Firth, S., Hodge, A., Jimenez-Soto, E., and Zeck, W. (2015). Overcoming stagnation in the levels and distribution of child mortality: The case of the Philippines. *PloS One*, 10(10):e0139458. doi: 10.1371/journal.pone.0139458.

Betti, G., Gagliardi, F., and Verma, V. (2018). Simplified Jackknife variance estimates for fuzzy measures of multidimensional poverty. *International Statistical Review*, 86(1):68–86. doi: 10.1111/insr.12219.

Beyene, A., Hailu, T., Faris, K., and Kloos, H. (2015). Current state and trends of access to sanitation in Ethiopia and the need to revise indicators to monitor progress in the Post-2015 era. *BMC Public Health*, 15(1):451. doi: 10.1186/s12889-015-1804-4.

Bhardwaj, A., Joshi, M., Khosla, R., and Dubash, N. K. (2019). More priorities, more problems? Decision-making with multiple energy, development and climate objectives. *Energy Research & Social Science*, 49:143–157. doi: 10.1016/j.erss.2018.11.003.

Bhattacharya, A. and Kharas, H. (2015). Worthy of support: Our piece on the "Sustainable Development Goals". The Economist. Retrieved July 19, 2019, from: https://www.economist.com/free-exchange/2015/04/08/worthy-of-support.

Boran, F., Boran, K., and Menlik, T. (2012). The evaluation of renewable energy technologies for electricity generation in Turkey using intuitionistic fuzzy TOPSIS. *Energy Sources, Part B: Economics, Planning, and Policy*, 7(1):81–90. doi: 10.1080/15567240903047483.

Bowman, K. O. and Shenton, L. (2007). The beta distribution, moment method, Karl Pearson and RA Fisher. *Far East Journal of Theoretical Statistics*, 23(2):133. Retrieved July 19, 2019, from: https://www.csm.ornl.gov/~bowman/fjts232.pdf.

Brass, W. (1975). *Methods for estimating fertility and mortality from limited and defective data*. Laboratories for Population Statistics Occasional Publication.

Briggs, A. H., Weinstein, M. C., Fenwick, E. A., Karnon, J., Sculpher, M. J., and Paltiel, A. D. (2012). Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. *Medical Decision Making*, 32(5):722–732. doi: 10.1177/0272989X12458348.

Broekhuizen, H., Groothuis-Oudshoorn, C. G., van Til, J. A., Hummel, J. M., and IJzerman, M. J. (2015). A review and classification of approaches for dealing with uncertainty in multi-criteria decision analysis for healthcare decisions. *Pharmacoeconomics*, 33(5):445–455. doi: 10.1007/s40273-014-0251-x.

Bulut, U. and Muratoglu, G. (2018). Renewable energy in Turkey: Great potential, low but increasing utilization, and an empirical analysis on renewable energy-growth nexus. *Energy Policy*, 123:240–250. doi: 10.1016/j.enpol.2018.08.057.

Burstein, R., Wang, H., Reiner Jr, R. C., and Hay, S. I. (2018). Development and validation of a new method for indirect estimation of neonatal, infant, and child mortality trends using summary birth histories. *PLoS Medicine*, 15(10):e1002687. doi: 10.1371/journal.pmed.1002687.

Büyüközkan, G. and Güleryüz, S. (2014). A new GDM based AHP framework with linguistic interval fuzzy preference relations for renewable energy planning. *Journal of Intelligent & Fuzzy Systems*, 27(6):3181–3195. doi: 10.3233/IFS-141275.

Büyüközkan, G. and Güleryüz, S. (2016). An integrated DEMATEL-ANP approach for renewable energy resources selection in Turkey. *International Journal of Production Economics*, 182:435–448. doi: 10.1016/j.ijpe.2016.09.015.

Büyüközkan, G. and Güleryüz, S. (2017). Evaluation of Renewable Energy Resources in Turkey using an integrated MCDM approach with linguistic interval fuzzy preference relations. *Energy*, 123:149–163. doi: 10.1016/j.energy.2017.01.137.

Cairncross, S., Bartram, J., Cumming, O., and Brocklehurst, C. (2010). Hygiene, sanitation, and water: what needs to be done? *PLoS Medicine*, 7(11):e1000365. doi: 10.1371/journal.pmed.1000365.

Carson, V., Tremblay, M. S., Chaput, J.-P., and Chastin, S. F. (2016). Associations between sleep duration, sedentary time, physical activity, and health indicators among Canadian children and youth using compositional analyses. *Applied Physiology, Nutrition, and Metabolism*, 41(6):S294–S302. doi: 10.1139/apnm-2016-0026.

Çelikbilek, Y. and Tüysüz, F. (2016). An integrated grey based multi-criteria decision making approach for the evaluation of renewable energy sources. *Energy*, 115:1246–1258. doi: 10.1016/j.energy.2016.09.091.

Chaudhuri, S. and Roy, M. (2017). Rural-urban spatial inequality in water and sanitation facilities in India: A cross-sectional study from household to national level. *Applied Geography*, 85:27–38. doi: 10.1016/j.apgeog.2017.05.003.

Chaudhuri1, S., Roy1, M., and Jain, A. (2018). Appraisal of WaSH (Water-Sanitation-

Hygiene) Infrastructure using a Composite Index, Spatial Algorithms and Sociodemographic Correlates in Rural India. *Journal of Environmental Informatics*. doi: 10.3808/jei.201800398.

Chitsaz, N. and Banihabib, M. E. (2015). Comparison of Different Multi Criteria Decision-Making Models in Prioritizing Flood Management Alternatives. *Water Resources Management*, 29(8):2503–2525. doi: 10.1007/s11269-015-0954-6.

Clapp, A., DauSchmidt, N., Millar, M., Hubbard, D., and Shepherd, K. D. (2013). A survey and analysis of the data requirements for stakeholders in African agriculture. World Agroforestry Centre (ICRAF). Retrieved July 19, 2019, from: https://hdl.handle.net/10568/34532.

Clopper, C. and Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413. doi: 10.2307/2331986.

Cohen, A. and Sullivan, C. A. (2010). Water and poverty in rural China: developing an instrument to assess the multiple dimensions of water and poverty. *Ecological Economics*, 69(5):999–1009. doi: 10.1016/j.ecolecon.2010.01.004.

Çolak, M. and Kaya, İ. (2017). Prioritization of renewable energy alternatives by using an integrated fuzzy MCDM model: A real case application for Turkey. *Renewable and Sustainable Energy Reviews*, 80:840–853. doi: 10.1016/j.rser.2017.05.194.

Cole, M. J., Bailey, R. M., Cullis, J. D., and New, M. G. (2018). Spatial inequality in water access and water use in South Africa. *Water Policy*, 20(1):37–52. doi: 10.2166/wp.2017.111.

Constantinou, A. C., Fenton, N., and Neil, M. (2016). Integrating expert knowledge with data in Bayesian networks: Preserving data-driven expectations when the expert variables remain unobserved. *Expert Systems with Applications*, 56:197–208. doi: 10.1016/j.eswa.2016.02.050.

Cronk, R., Slaymaker, T., and Bartram, J. (2015). Monitoring drinking water, sanitation, and hygiene in non-household settings: Priorities for policy and practice.

International Journal of Hygiene and Environmental Health, 218(8):694–703. doi: 10.1016/j.ijheh.2015.03.003.

Cumming, O., Elliott, M., Overbo, A., and Bartram, J. (2014). Does global progress on sanitation really lag behind water? An analysis of global progress on community- and household-level access to safe water and sanitation. *PloS One*, 9(12):e114699. doi: 10.1371/journal.pone.0114699.

Custer, S. and Sethi, T. (2017). Avoiding Data Graveyards: Insights from Data Producers Users in Three Countries. USAID and AidData Center for Development Policy. Retrieved July 19, 2019, from: https://www.developmentgateway.org/sites/default/files/2018-10/Avoiding%20Data%20Graveyards%20Final%20Report.pdf.

Darvesh, N., Das, J. K., Vaivada, T., Gaffey, M. F., Rasanathan, K., and Bhutta, Z. A. (2017). Water, sanitation and hygiene interventions for acute childhood diarrhea: a systematic review to provide estimates for the Lives Saved Tool. *BMC Public Health*, 17(4):776. doi: 10.1186/s12889-017-4746-1.

De Palencia, A. J. F. and Pérez-Foguet, A. (2011). Implementing pro-poor policies in a decentralized context: the case of the Rural Water Supply and Sanitation Program in Tanzania. *Sustainability Science*, 6(1):37–49. doi: 10.1007/s11625-010-0121-1.

Demirtas, O. (2013). Evaluating the best renewable energy technology for sustainable energy planning. *International Journal of Energy Economics and Policy*, 3(4S):23–33. Retrieved July 19, 2019, from: https://dergipark.org.tr/ijeeep/issue/31907/350760.

Deng, J.-L. et al. (1982). Control problems of grey systems. *Systems & Control. Lett.*, 1(5):288–294. doi: 10.1016/S0167-6911(82)80025-X.

Development Gateway (2016). Policy Brief: Increasing the Impact of Results Data. Retrieved July 19, 2019, from: https://www.developmentgateway.org/sites/default/files/2017-02/RDI-PolicyBrief.pdf.

Development Indicatives (2018). Enhancing access to safe water and improved sanitation services in Kenya: are we on track? Retrieved July 19, 2019, from: http://devinit.org/wp-content/uploads/2018/12/Enhancing-access-to-safe-water-and-improved-sanitation-services-in-Kenya.pdf.

Diakoulaki, D., Antunes, C. H., and Martins, A. G. (2005). MCDA and energy planning. In *Multiple criteria decision analysis: state of the art surveys*, pages 859–890. Springer. doi: 10.1007/0-387-23081-5$_2$1.

Dodgson, J., Spackman, M., Pearman, A., and Phillips, L. (2009). Multi-criteria analysis: a manual. London School of Economics and Political Science, Department of Economic History. Retrieved July 19, 2019, from: http://eprints.lse.ac.uk/12761/1/Multi-criteria_Analysis.pdf.

du Preez, M., Conroy, R. M., Ligondo, S., Hennessy, J., Elmore-Meegan, M., Soita, A., and McGuigan, K. G. (2011). Randomized intervention study of solar disinfection of drinking water in the prevention of dysentery in Kenyan children aged under 5 years. *Environmental Science & Technology*, 45(21):9315–9323. doi: 10.1021/es2018835.

Duckstein, L., Gershon, M., and McAniff, R. (1982). Model selection in multiobjective decision making for river basin planning. *Advances in Water Resources*, 5:178–184. doi: 10.1016/0309-1708(82)90040-9.

Durbach, I. N. and Stewart, T. J. (2011). An experimental study of the effect of uncertainty representation on decision making. *European Journal of Operational Research*, 214(2):380–392. doi: 10.1016/j.ejor.2011.04.021.

Durbach, I. N. and Stewart, T. J. (2012). Modeling uncertainty in multi-criteria decision analysis. *European Journal of Operational Research*, 223(1):1–14. doi: 10.1016/j.ejor.2012.04.038.

Easerly, W. (2015). The SDGs Should Stand for Senseless, Dreamy, Garbled. Foreign Policy. Retrieved July 19, 2019, from: https://foreignpolicy.com/2015/09/28/the-sdgs-are-utopian-and-worthless-mdgs-development-rise-of-the-rest.

EBRD (2014). National Renewable Energy Action Plan for Turkey. European Bank for Reconstruction and Development and Turkish Ministry of Energy. Retrieved July 19, 2019, from: https://www.ebrd.com/documents/comms-and-bis/turkey-national-renewable-energy-action-plan.pdf.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barcelo-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3):279–300. doi: 10.1023/A:1023818214614.

Erdin, C. and Ozkaya, G. (2019). Turkey's 2023 Energy Strategies and Investment Opportunities for Renewable Energy Sources: Site Selection Based on ELECTRE. *Sustainability*, 11(7):2136. doi: 10.3390/su11072136.

Erdogan, M. and Kaya, I. (2015). An integrated multi-criteria decision-making methodology based on type-2 fuzzy sets for selection among energy alternatives in Turkey. *Iranian Journal of Fuzzy Systems*, 12(1):1–25. doi: 10.22111/IJFS.2015.1839.

Erol, Ö. and Kılkış, B. (2012). An energy source policy assessment using analytical hierarchy process. *Energy Conversion and Management*, 63:245–252. doi: 10.1016/j.enconman.2012.01.040.

Ertay, T., Kahraman, C., and Kaya, İ. (2013). Evaluation of renewable energy alternatives using MACBETH and fuzzy AHP multicriteria methods: the case of Turkey. *Technological and Economic Development of Economy*, 19(1):38–62. doi: 10.3846/20294913.2012.762950.

Ervural, B. C., Zaim, S., Demirel, O. F., Aydin, Z., and Delen, D. (2018). An ANP and fuzzy TOPSIS-based SWOT analysis for Turkey's energy planning. *Renewable and Sustainable Energy Reviews*, 82:1538–1550. doi: 10.1016/j.rser.2017.06.095.

Fackrell, B. (2013). Turkey and Regional Energy Security on the Road to 2023. pages 83–89. Turkish Policy Quarterly. Retrieved July 19, 2019, from: http://turkishpolicy.com/article/635/turkey-and-regional-energy-security-on-the-road-to-2023-summer-2013.

Fairclough, S. J., Dumuid, D., Taylor, S., Curry, W., McGrane, B., Stratton, G., Maher, C., and Olds, T. (2017). Fitness, fatness and the reallocation of time between children's daily movement behaviours: an analysis of compositional data. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1):64. doi: 10.1186/s12966-017-0521-z.

Feiring, B. and Hassler, A. (2016). Human rights in follow-up and review of the 2030 Agenda for Sustainable Development. Danish Institute for Human Rights. Retrieved July 19, 2019, from: https://www.humanrights.dk/sites/humanrights.dk/files/media/dokumenter/udgivelser/sdg/may_17_follow-up_and_review_sdg_docx.pdf.

Fenton, N. and Neil, M. (2012). *Risk assessment and decision analysis with Bayesian networks*. CRC Press.

Ferrer-Rosell, B., Coenders, G., and Martínez-Garcia, E. (2016). Segmentation by tourist expenditure composition: an approach with compositional data analysis and latent classes. *Tourism Analysis*, 21(6):589–602. doi: 10.3727/108354216X14713487283075.

Ferrier, S., Ninan, K., Leadley, P., Alkemade, R., Acosta, L., Akçakaya, H., Brotons, L., Cheung, W., Christensen, V., Harhash, K., Kabubo-Mariara, J., Lundquist, C., Obersteiner, M., Pereira, H. M., Peterson, G., Pichs-Madruga, R., Ravindranath, N., Rondinini, C., and Wintle, B. A. (2016). The methodological assessment report on scenarios and models of biodiversity and ecosystem services. Secretariat of the Intergovernmental Platform for Biodiversity and Ecosystem Services. Retrieved July 19, 2019, from: https://www.ipbes.net/sites/default/files/downloads/pdf/2016.methodological_assessment_report_scenarios_models.pdf.

Figueira, J. R., Greco, S., Roy, B., and Słowiński, R. (2010). ELECTRE methods: main features and recent developments. In *Handbook of multicriteria analysis*, pages 51–89. Springer. doi: 10.1007/978-3-540-92828-7_3.

Fuller, J. A., Goldstick, J., Bartram, J., and Eisenberg, J. N. (2016). Tracking progress towards global drinking water and sanitation targets: A within and among country analysis. *Science of the Total Environment*, 541:857–864. doi: 10.1016/j.scitotenv.2015.09.130.

Gabrel, C. (2000). The National Nursing Home Survey: 1995 summary. *Vital and Health Statistics*, 13(146):1–83. Retrieved July 19, 2019, from: https://www.cdc.gov/nchs/data/series/sr_13/sr13_146.pdf.

Garrett, V., Ogutu, P., Mabonga, P., Ombeki, S., Mwaki, A., Aluoch, G., Phelan, M., and Quick, R. (2008). Diarrhoea prevention in a high-risk rural Kenyan population through point-of-use chlorination, safe water storage, sanitation, and rainwater harvesting. *Epidemiology & Infection*, 136(11):1463–1471. doi: 10.1017/S095026880700026X.

Gavurova, B., Belas, J., Kocisova, K., and Kliestik, T. (2017). Comparison of selected methods for performance evaluation of Czech and Slovak commercial banks. *Journal of Business Economics and Management*, 18(5):852–876. doi: 10.3846/16111699.2017.1371637.

GCPSE (2016). SDG Implementation Framework: Effective public service for SDG implementation. UNDP Global Centre for Public Service Excellence. Retrieved July 19, 2019, from: https://www.localizingthesdgs.org/library/136/Effective-public-service-for-SDG-implementation-SDG-Implementation-Framework-Note-1.pdf.

Gibson, J. and Le, T. (2019). Using local expert knowledge to measure prices: Evidence from a survey experiment in Vietnam. Munich Personal RePEc Archive (MPRA) Paper No. 92533. Retrieved July 19, 2019, from: https://mpra.ub.uni-muenchen.de/92533/1/MPRA_paper_92533.pdf.

Giné-Garriga, R., Fernández de Palencia, A., and Pérez-Foguet, A. (2015). Improved monitoring framework for local planning in the water, sanitation and hygiene sector: From data to decision-making. *Science of The Total Environment*, 526:204–214. doi: 10.1016/j.scitotenv.2015.04.078.

Giné-Garriga, R., Jiménez-Fernéndez de Palencia, A., and Pérez-Foguet, A. (2013). Water-sanitation-hygiene mapping: An improved approach for data collection at local level. *Science of the Total Environment*, 463-464:700–711. doi: 10.1016/j.scitotenv.2013.06.005.

Giné Garriga, R. and Pérez Foguet, A. (2010). Improved Method to Calculate a Water Poverty Index at Local Scale. *Journal of Environmental Engineering*, 136(11):1287–1298. doi: 10.1061/(ASCE)EE.1943-7870.0000255.

Giné Garriga, R. and Pérez Foguet, A. (2013a). Unravelling the Linkages Between Water, Sanitation, Hygiene and Rural Poverty: The WASH Poverty Index. *Water Resources Management*, 27(5):1501–1515. doi: 10.1007/s11269-012-0251-6.

Giné Garriga, R. and Pérez Foguet, A. (2013b). Water, sanitation, hygiene and rural poverty: issues of sector monitoring and the role of aggregated indicators. *Water Policy*, 15(6):1018–1045. doi: 10.2166/wp.2013.037.

Giné-Garriga, R. and Pérez-Foguet, A. (2019). Monitoring and targeting the sanitation poor: A multidimensional approach. *Natural Resources Forum*, 43(2):82–94. doi: 10.1111/1477-8947.12171.

Goetghebeur, M. and Wagner, M. (2017). Identifying value(s): a reflection on the ethical aspects of MCDA in healthcare decision making. In *Multi-Criteria Decision Analysis to support healthcare decisions*, pages 29–46. Springer. doi: 10.1007/978-3-319-47540-0_3.

Goicoechea, A., Stakhiv, E. Z., and Li, F. (1992). Experimental Evaluation of Multiple Criteria Decision Models for Application to Water Resources Planning. *Journal of the American Water Resources Association*, 28(1):89–102. doi: 10.1111/j.1752-1688.1992.tb03156.x.

Gold, H. J., Wilkerson, G. G., Yu, Y., and Stinner, R. E. (1990). Decision analysis as a tool for integrating simulation with expert systems when risk and uncertainty are important. *Computers and Electronics in Agriculture*, 4(4):343–360. doi: 10.1016/0168-1699(90)90039-R.

Gouveia, J. P., Palma, P., and Simoes, S. G. (2019). Energy poverty vulnerability index: A multidimensional tool to identify hotspots for local action. *Energy Reports*, 5:187–201. doi: 10.1016/j.egyr.2018.12.004.

Groothuis-Oudshoorn, C. G., Broekhuizen, H., and van Til, J. (2017). Dealing with uncertainty in the analysis and reporting of MCDA. In *Multi-Criteria Decision Analysis to Support Healthcare Decisions*, pages 67–85. Springer. doi: 10.1007/978-3-319-47540-0_5.

Guillot, M., Gerland, P., Pelletier, F., and Saabneh, A. (2012). Child mortality estimation: a global overview of infant and child mortality age patterns in light of new empirical data. *PLoS Medicine*, 9(8):e1001299. doi: 10.1371/journal.pmed.1001299.

Guppy, L., Mehta, P., and Qadir, M. (2019). Sustainable Development Goal 6:

Two gaps in the race for indicators. *Sustainability Science*, 14(2):501–513. doi: 10.1007/s11625-018-0649-z.

Gutierrez, E. (2007). Delivering pro-poor water and sanitation services: The technical and political challenges in Malawi and Zambia. *Geoforum*, 38(5):886–900. doi: 10.1016/j.geoforum.2005.09.010.

Hanratty, P. and Joseph, B. (1992). Decision-making in chemical engineering and expert systems: application of the analytic hierarchy process to reactor selection. *Computers & Chemical Engineering*, 16(9):849–860. doi: 10.1016/0098-1354(92)80037-A.

He, W.-J., Lai, Y.-S., Karmacharya, B. M., Dai, B.-F., Hao, Y.-T., and Xu, D. R. (2018). Geographical heterogeneity and inequality of access to improved drinking water supply and sanitation in Nepal. *International Journal for Equity in Health*, 17(1):40. doi: 10.1186/s12939-018-0754-8.

Head, B. W. (2016). Toward more "evidence-informed" policy making? *Public Administration Review*, 76(3):472–484. doi: 10.1111/puar.12475.

Hering, J. (2017). Managing the 'Monitoring Imperative' in the Context of SDG Target 6.3 on Water Quality and Wastewater. *Sustainability*, 9(9):1572. doi: 10.3390/su9091572.

Hickel, J. (2015). The problem with saving the world. The Jacobin Magazine. Retrieved July 19, 2019, from: https://www.jacobinmag.com/2015/08/global-poverty-climate-change-sdgs.

Hill, K., Brady, E., Zimmerman, L., Montana, L., Silva, R., and Amouzou, A. (2015). Monitoring change in child mortality through household surveys. *PloS One*, 10(11):e0137713. doi: 10.1371/journal.pone.0137713.

Hodge, A., Firth, S., Marthias, T., and Jimenez-Soto, E. (2014). Location matters: trends in inequalities in child mortality in Indonesia. Evidence from repeated cross-sectional surveys. *PLoS One*, 9(7):e103597. doi: 10.1371/journal.pone.0103597.

Hokkanen, J., Lahdelma, R., and Salminen, P. (2000). Multicriteria decision support in a technology competition for cleaning polluted soil in Helsinki. *Journal of Environmental Management*, 60(4):339–348. doi: 10.1006/jema.2000.0389.

Humanosphere (2015). Gates Foundation rallies the troops to attack UN development goals. Retrieved July 19, 2019, from: http://www.humanosphere.org/world-politics/2015/05/gates-foundation-rallies-the-troops-to-attack-un-development-goals.

Iddrisu, I. and Bhattacharyya, S. C. (2015). Sustainable Energy Development Index: A multi-dimensional indicator for measuring sustainable energy development. *Renewable and Sustainable Energy Reviews*, 50:513–530. doi: 10.1016/j.rser.2015.05.032.

IEAG (2015). A World That Counts: Mobilizing the data revolution for Sustainable Development. Independent Expert Advisory Group Secretary. Retrieved July 19, 2019, from: http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf.

Ighravwe, D. E. and Babatunde, M. O. (2018). Determination of a suitable renewable energy source for mini-grid business: A risk-based multicriteria approach. *Journal of Renewable Energy*, 2018. doi: 10.1155/2018/2163262.

IGME (2018). Explanatory Notes: Child mortality trend series to 2017. UN Inter-agency Group for Child Mortality Estimation. Retrieved July 19, 2019, from: https://childmortality.org/wp-content/uploads/2018/12/English_IGME_country_consultation_note.pdf.

Inotai, A., Nguyen, H. T., Hidayat, B., Nurgozhin, T., Kiet, P. H. T., Campbell, J. D., Németh, B., Maniadakis, N., Brixner, D., Wijaya, K., and Kaló, Z. (2018). Guidance toward the implementation of multicriteria decision analysis framework in developing countries. *Expert Review of Pharmacoeconomics & Outcomes Research*, 18(6):585–592. doi: 10.1080/14737167.2018.1508345.

International Energy Agency (2011). Renewable Energy Law 2010 - Turkey. Retrieved July 19, 2019, from: https://www.iea.org/policiesandmeasures/pams/turkey/name-24961-en.php.

International Energy Agency (2016). Energy policies of IEA countries: Turkey 2016 review. OECD/IEA. Retrieved July 19, 2019, from: https://www.iea.org/publications/freepublications/publication/EnergyPoliciesofIEACountriesTurkey.pdf.

International Energy Agency (2019a). Statistics and Data: Energy imports, net (% of energy use) in Turkey. Retrieved July 19, 2019, from: https://data.worldbank.org/indicator/EG.IMP.CONS.ZS.

International Energy Agency (2019b). Statistics and Data: Total Primary Energy Supply (TPES) by source in Turkey. Retrieved July 19, 2019, from: https://www.iea.org/statistics/?country=WORLD&year=2016&category=Energy%20supply&indicator=TPESbySource&mode=chart&dataTable=BALANCES.

Ishizaka, A. and Nemery, P. (2013). *Multi-criteria decision analysis: methods and software*. John Wiley & Sons, Ltd. doi: 10.1002/9781118644898.

IWGHS (2019). Achieving the Full Potential of Household Surveys in the SDG Era. Inter-Secretariat Working Group on Household Surveys. Retrieved July 19, 2019, from: https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item4c-ISWGHS-E.pdf.

Janicki, R. (2019). Properties of the beta regression model for small area estimation of proportions and application to estimation of poverty rates. *Communications in Statistics-Theory and Methods*, pages 1–21. doi: 10.1080/03610926.2019.1570266.

Janssen, J., Krol, M. S., Schielen, R. M. J., Hoekstra, A. Y., and de Kok, J.-L. (2010). Assessment of uncertainties in expert knowledge, illustrated in fuzzy rule-based models. *Ecological Modelling*, 221(9):1245–1251. doi: 10.1016/j.ecolmodel.2010.01.011.

Jayne, T. S. and Muyanga, M. (2012). Land constraints in Kenya's densely populated rural areas: implications for food policy and institutional reform. *Food Security*, 4(3):399–421. doi: 10.1007/s12571-012-0174-3.

Jeffreys, I. (2004). The use of compensatory and non-compensatory multi-criteria analysis for small-scale forestry. *Small-scale Forest Economics, Management and Policy*, 3(1):99–117. doi: 10.1007/s11842-004-0007-0.

Jeuland, M. A., Fuente, D. E., Ozdemir, S., Allaire, M. C., and Whittington, D. (2013). The long-term dynamics of mortality benefits from improved water and sanitation in less developed countries. *PloS One*, 8(10):e74804. doi: 10.1371/journal.pone.0074804.

Jeyakumar, A. and Ghugre, P. (2017). Assessment of personal hygiene and sanitation using a composite index among adolescent girls and their households in urban slums of Pune, Maharashtra. *Journal of Water, Sanitation and Hygiene for Development*, 7(2):323–330. doi: 10.2166/washdev.2017.127.

Jia, P., Anderson, J. D., Leitner, M., and Rheingans, R. (2016). High-resolution spatial distribution and estimation of access to improved sanitation in Kenya. *PloS One*, 11(7):e0158490. doi: 10.1371/journal.pone.0158490.

Jiménez, A. and Pérez-Foguet, A. (2010). Challenges for water governance in rural water supply: lessons learned from Tanzania. *International Journal of Water Resources Development*, 26(2):235–248. doi: 10.1080/07900621003775763.

JMP (2014). Report of Task Force on Methods. WHO/UNICEF's Joint Monitoring Programme. Retrieved July 19, 2019, from: https://washdata.org/file/478.

JMP (2015). Progress on sanitation and drinking water: 2015 update and MDG assessment. WHO/UNICEF's Joint Monitoring Programme. Retrieved July 19, 2019, from: https://data.unicef.org/wp-content/uploads/2015/12/Progress-on-Sanitation-and-Drinking-Water_234.pdf.

JMP (2017a). JMP data on access to Water, Sanitation and Hygiene on Households. Retrieved July 19, 2019, from: https://washdata.org/data/household.

JMP (2017b). Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baselines. WHO/UNICEF's Joint Monitoring Programme. Retrieved July 19, 2019, from: https://www.unicef.org/publications/files/Progress_on_Drinking_Water_Sanitation_and_Hygiene_2017.pdf.

JMP (2018). JMP Methodology: 2017 Update and SDG Baseline.

WHO/UNICEF's Joint Monitoring Programme. Retrieved July 19, 2019, from: https://washdata.org/sites/default/files/documents/reports/2018-04/JMP-2017-update-methodology.pdf.

Jütting, J. and McDonnell, I. (2017). Overview: What will it take for data to enable development? In *Development Co-operation Report 2017*. OECD. Retrieved July 19, 2019, from: https://read.oecd-ilibrary.org/development/development-co-operation-report-2017_dcr-2017-en.

Kabak, M. and Dağdeviren, M. (2014). Prioritization of renewable energy sources for Turkey by using a hybrid MCDM methodology. *Energy Conversion and Management*, 79:25–33. doi: 10.1016/j.enconman.2013.11.036.

Kahraman, C. and Kaya, İ. (2010). A fuzzy multicriteria methodology for selection among energy alternatives. *Expert Systems with Applications*, 37(9):6270–6281. doi: 10.1016/j.eswa.2010.02.095.

Kahraman, C., Kaya, I., and Çebi, S. (2010). Renewable energy system selection based on computing with words. *International Journal of Computational Intelligence Systems*, 3(4):461–473. doi: 10.1080/18756891.2010.9727714.

Kangas, A., Kangas, J., and Pykäläinen, J. (2001). Outranking methods as tools in strategic natural resources planning. *Silva Fennica*, 35(2):215–227. doi: 10.14214/sf.597.

Kaplan, Y. A. (2015). Overview of wind energy in the world and assessment of current wind energy policies in Turkey. *Renewable and Sustainable Energy Reviews*, 43:562–568. doi: 10.1016/j.rser.2014.11.027.

Karakaş, E. and Yıldıran, O. V. (2019). Evaluation of Renewable Energy Alternatives for Turkey via Modified Fuzzy AHP. *International Journal of Energy Economics and Policy*, 9(2):31–39. doi: 10.32479/ijeep.7349.

Kaya, İ., Colak, M., and Terzi, F. (2019). A comprehensive review of fuzzy multi criteria decision making methodologies for energy policy making. *Energy Strategy Reviews*, 24:207–228. doi: 10.1016/j.esr.2019.03.003.

Kaya, T. and Kahraman, C. (2010). Multicriteria renewable energy planning using an integrated fuzzy VIKOR & AHP methodology: The case of Istanbul. *Energy*, 35(6):2517–2527. doi: 10.1016/j.energy.2010.02.051.

Kaya, T. and Kahraman, C. (2011). Multicriteria decision making in energy planning using a modified fuzzy TOPSIS methodology. *Expert Systems with Applications*, 38(6):6577–6585. doi: 10.1016/j.eswa.2010.11.081.

Kaydul, E., Demir, Z., and Çolak, N. (2017). Turkey's Renewable Energy Sources and Governmental Incentives. In *International Sustainable Buildings Symposium*, pages 71–81. doi: 10.1007/978-3-319-64349-6_7.

Kenya National Commission on Human Rights (2014). Status on the right to safe drinking water and sanitation in Kenya. Retrieved July 19, 2019, from: https://www.ohchr.org/Documents/Issues/Water/HRViolations/KenyaNCHR.doc.

Ki-Moon, B. (2012). Secretary-General to Global Development Center: Energy is the golden thread connecting economic growth, social equity, environmental sustainability. Secretary-General: Statements and Messages, SG/SM/15839-EN/288. Retrieved July 19, 2019, from: https://www.un.org/press/en/2014/sgsm15839.doc.htm.

Kılkış, Ş. (2015). Composite index for benchmarking local energy systems of Mediterranean port cities. *Energy*, 92:622–638. doi: 10.1016/j.energy.2015.06.093.

Kirigia, J. M., Muthuri, R. D. K., Nabyonga-Orem, J., and Kirigia, D. G. (2015). Counting the cost of child mortality in the World Health Organization African region. *BMC Public Health*, 15(1):1103. doi: 10.1186/s12889-015-2465-z.

Korenromp, E., Arnold, F., Williams, B., Nahlen, B., and Snow, R. (2004). Monitoring trends in under-5 mortality rates through national birth history surveys. *International Journal of Epidemiology*, 33(6):1293–1301. doi: 10.1093/ije/dyh182.

Krantz, D., Luce, D., Suppes, P., and Tversky, A. (1971). *Foundations of measurement, Vol. I: Additive and polynomial representations*. Dover Publications.

Kuleli Pak, B., Albayrak, Y. E., and Erensal, Y. C. (2015). Renewable energy perspective for Turkey using sustainability indicators. *International Journal of Computational Intelligence Systems*, 8(1):187–197. doi: 10.1080/18756891.2014.963987.

Kumar, S., Kumar, N., and Vivekadhish, S. (2016). Millennium development goals (MDGS) to sustainable development goals (SDGS): Addressing unfinished agenda and strengthening sustainable development and partnership. *Indian Journal of Community Medicine*, 41(1):1. doi: 10.4103/0970-0218.170955.

Kurka, T. and Blackwood, D. (2013). Selection of MCA methods to support decision making for renewable energy developments. *Renewable and Sustainable Energy Reviews*, 27:225–233. doi: 10.1016/j.rser.2013.07.001.

Kurui, E. J., Ogendi, G. M., Moturi, W. N., and Nyawanga, D. O. (2019). Household Water Handling Practices in the Arid and Semi-Arid Lands in Kenya. In *The Relevance of Hygiene to Health in Developing Countries*. IntechOpen. doi: 10.5772/intechopen.8039.

Lagravinese, R., Liberati, P., and Resce, G. (2019). Measuring Health Inequality in US: A Composite Index Approach. *Social Indicators Research*, pages 1–26. doi:10.1007/s11205-019-02177-x.

Le Blanc, D. (2015). Towards integration at last? The sustainable development goals as a network of targets. *Sustainable Development*, 23(3):176–187. doi: 10.1002/sd.1582.

Lee, H.-C. and Chang, C.-T. (2018). Comparative analysis of MCDM methods for ranking renewable energy sources in Taiwan. *Renewable and Sustainable Energy Reviews*, 92:883–896. doi: 10.1016/j.rser.2018.05.007.

Lesser, V. M. and Kalsbeek, W. D. (1999). Nonsampling errors in environmental surveys. *Journal of Agricultural, Biological, and Environmental Statistics*, pages 473–488. doi: 10.2307/1400503.

Linares-Mustaros, S., Coenders, G., and Vives-Mestres, M. (2018). Financial performance and distress profiles. From classification according to financial ratios to compositional classification. *Advances in Accounting*, 40:1–10. doi: 10.1016/j.adiac.2017.10.003.

LIXIL (2016). The true cost of poor sanitation. LIXIL and Oxford Economics, with the collaboration of WaterAid. Retrieved July 19, 2019, from: https://www.lixil.com/en/sustainability/pdf/the_true_cost_of_poor_sanitation_e.pdf.

Lloyd, C. D., Pawlowsky-Glahn, V., and Egozcue, J. J. (2012). Compositional data analysis in population studies. *Annals of the Association of American Geographers*, 102(6):1251–1266. Retrieved July 19, 2019, from: https://www.jstor.org/stable/41805896.

Løken, E. (2007). Use of multicriteria decision analysis methods for energy planning problems. *Renewable and Sustainable Energy Reviews*, 11(7):1584–1595. doi: 10.1016/j.rser.2005.11.005.

Luedeling, E. and Shepherd, K. (2016). Decision-focused agricultural research. *The Solutions Journal*, 7(5):46–54. Retrieved July 19, 2019, from: http://www.thesolutionsjournal.com/article/decision-focused-agricultural-research.

Luh, J., Baum, R., and Bartram, J. (2013). Equity in water and sanitation: Developing an index to measure progressive realization of the human right. *International Journal of Hygiene and Environmental Health*, 216(6):662–671. doi: 10.1016/j.ijheh.2012.12.007.

Mahapatra, P., Shibuya, K., Lopez, A. D., Coullare, F., Notzon, F. C., Rao, C., Szreter, S., et al. (2007). Civil registration systems and vital statistics: successes and missed opportunities. *The Lancet*, 370(9599):1653–1663. doi: 10.1016/S0140-6736(07)61308-7.

Mahmoud, M. R. and Garcia, L. A. (2000). Comparison of different multicriteria evaluation methods for the Red Bluff diversion dam. *Environmental Modelling & Software*, 2815(5):471–478. doi: 10.1016/S1364-8152(00)00025-6.

Mansour, G., Oyaya, C., and Owor, M. (2017). Situation analysis of the urban sanitation sector in Kenya. Water and & Sanitation for the Urban Poor (WSUP). Retrieved July 19, 2019, from: https://www.wsup.com/content/uploads/2017/09/Situation-analysis-of-the-urban-sanitation-sector-in-Kenya.pdf.

Marcillo-Delgado, J., Ortego, M., and Pérez-Foguet, A. (2019). A compositional approach for modelling SDG7 indicators: Case study applied to electricity access. *Renewable and Sustainable Energy Reviews*, 107:388–398. doi: 10.1016/j.rser.2019.03.028.

Martín-Fernández, J. A., Barceló-Vidal, C., and Pawlowsky-Glahn, V. (2003). Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology*, 35(3):253–278. doi: 10.1023/A:1023866030544.

Marttunen, M., Belton, V., and Lienert, J. (2018). Are objectives hierarchy related biases observed in practice? A meta-analysis of environmental and energy applications of Multi-Criteria Decision Analysis. *European Journal of Operational Research*, 265(1):178–194. doi: 10.1016/j.ejor.2017.02.038.

Mazziotta, M. and Pareto, A. (2016). On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Social Indicators Research*, 127(3):983–1003. doi: 10.1007/s11205-015-0998-2.

McArthur, J. W. and Rasmussen, K. (2019). Classifying Sustainable Development Goal trajectories: A country-level methodology for identifying which issues and people are getting left behind. *World Development*, 123:104608. doi: 10.1016/j.worlddev.2019.06.031.

McInerney-Lankford, S. (2017). Human Rights and the SDGs: progress or a missed opportunity? Retrieved July 19, 2019, from: https://ohrh.law.ox.ac.uk/human-rights-and-the-sdgs-progress-or-a-missed-opportunity.

McKenna, R., Bertsch, V., Mainzer, K., and Fichtner, W. (2018). Combining local preferences with multi-criteria decision analysis and linear optimization to develop feasible energy concepts in small communities. *European Journal of Operational Research*, 268(3):1092–1110. doi: 10.1016/j.ejor.2018.01.036.

MENR (2018). Strategic plan 2015-2019. Ministry of Energy and Natural Resources. Retrieved July 19, 2019, from: https://www.etkb.gov.tr/File/?path=ROOT%2f1%2fDocuments%2fStrategic+Plan%2fStrategicPlan2015-2019.pdf.

MICS (2006). Enquête par Grappes à Indicateurs Multiples (MICS) et Indicateurs des Objectifs du Millénaire pour le Développement (OMD), Burkina Faso. UNICEF. Retrieved July 19, 2019, from: https://mics-surveys-prod.s3.amazonaws.

com/MICS3/West%20and%20Central%20Africa/Burkina%20Faso/2006/Final/Burkina%20Faso%202006%20MICS_French.pdf.

Minnery, M., Firth, S., Hodge, A., and Jimenez-Soto, E. (2015). Neonatal mortality and inequalities in Bangladesh: differential progress and sub-national developments. *Maternal and Child Health Journal*, 19(9):2038–2047. doi: 10.1007/s10995-015-1716-z.

Motahari Farimani, N. (2018). Decision Context Structuring: An Introductory Guide for Selecting MCDM. *SSRN*. doi: 10.2139/ssrn.3183184.

Mousavi, M., Gitinavard, H., and Mousavi, S. (2017). A soft computing based-modified ELECTRE model for renewable energy policy selection with unknown information. *Renewable and Sustainable Energy Reviews*, 68:774–787. doi: 10.1016/j.rser.2016.09.125.

Munda, G. and Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics*, 41(12):1513–1523. doi: 10.1080/00036840601019364.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005). Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD. Retrieved July 19, 2019, from: https://www.oecd.org/sdd/42495745.pdf.

Nhamo, G., Nhemachena, C., and Nhamo, S. (2019). Is 2030 too soon for Africa to achieve the water and sanitation sustainable development goal? *Science of The Total Environment*, 669:129–139. doi: 10.1016/j.scitotenv.2019.03.109.

Nussbaumer, P., Bazilian, M., and Modi, V. (2012). Measuring energy poverty: Focusing on what matters. *Renewable and Sustainable Energy Reviews*, 16(1):231–243. doi: 10.1016/j.rser.2011.07.150.

Oeppen, J. (2008). Coherent forecasting of multiple-decrement life tables: a test using Japanese cause of death data. Universitat de Girona, Departament d'Informàtica i Matemàtica Aplicada. Retrieved July 19, 2019, from: http://hdl.handle.net/10256/742.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006). *Uncertain judgements: eliciting experts' probabilities*. John Wiley & Sons. doi: 10.1002/0470033312.

OHCHR (2011). Report of the Special Rapporteur on the human right to safe drinking water and sanitation, Catarina de Albuquerque. A/HRC/18/33. Office of the High Commissioner for Human Rights. Retrieved July 19, 2019, from: https://www2.ohchr.org/english/bodies/hrcouncil/docs/18session/A-HRC-18-33_en.pdf.

Olson, D. L., Moshkovich, H. M., Schellenberger, R., and Mechitov, A. I. (1995). Consistency and accuracy in decision aids: Experiments with four multiattribute systems. *Decision Sciences*, 26(6):723–747. doi: 10.1111/j.1540-5915.1995.tb01573.x.

Ouerdane, W. (2011). Multiple criteria decision aiding: a dialectical perspective. *4OR: A Quarterly Journal of Operations Research*, 9(4). doi: 10.1007/s10288-010-0147-3f.

Ozcan, M. (2014). Assessment of renewable energy incentive system from investors' perspective. *Renewable energy*, 71:425–432. doi: 10.1016/j.renene.2014.05.053.

Ozcan, M. (2018). The role of renewables in increasing Turkey's self-sufficiency in electrical energy. *Renewable and Sustainable Energy Reviews*, 82:2629–2639. doi: 10.1016/j.rser.2017.09.111.

Ozernoy, V. M. (1992). Choosing the "Best" multiple criteria decision-making method. *Information Systems and Operational Research*, 30(2):159–171. doi: 10.1080/03155986.1992.11732192.

Patunru, A. A. (2015). Access to Safe Drinking Water and Sanitation in Indonesia. *Asia & the Pacific Policy Studies*, 2(2):234–244. doi: 10.1002/app5.81.

Pawlowsky-Glahn, V. and Egozcue, J. J. (2006). Compositional data and their analysis: an introduction. *Geological Society*, 264(1):1–10. doi: 10.1144/GSL.SP.2006.264.01.01.

Pawlowsky-Glahn, V. and Egozcue, J. J. (2011). Exploring compositional data with the CoDa-dendrogram. *Austrian Journal of Statistics*, 40(1–2):103–113. doi: 10.17713/ajs.v40i12.202.

Pawlowsky-Glahn, V., Egozcue, J. J., and Tolosana-Delgado, R. (2015). *Modeling and analysis of compositional data*. John Wiley & Sons. doi: 10.1002/9781119003144.

Peppard, P. E., Kindig, D. A., Dranger, E., Jovaag, A., and Remington, P. L. (2008). Ranking community health status to stimulate discussion of local public health issues: the Wisconsin County Health Rankings. *American Journal of Public Health*, 98(2):209–212. doi: 10.2105/AJPH.2006.092981.

Pérez-Foguet, A. and Giné-Garriga, R. (2011). Analyzing Water Poverty in Basins. *Water Resources Management*, 25(14):3595–3612. doi: 10.1007/s11269-011-9872-4.

Pérez-Foguet, A., Giné-Garriga, R., and Ortego, M. I. (2017). Compositional data for global monitoring: The case of drinking water and sanitation. *Science of the Total Environment*, 590:554–565. doi: 10.1016/j.scitotenv.2017.02.220.

Prüss-Ustün, A., Wolf, J., Bartram, J., Clasen, T., Cumming, O., Freeman, M. C., Gordon, B., Hunter, P. R., Medlicott, K., and Johnston, R. (2019). Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: An updated analysis with a focus on low-and middle-income countries. *International Journal of Hygiene and Environmental Health*, 222(5):765–777. doi: 10.1016/j.ijheh.2019.05.004.

Pullan, R. L., Freeman, M. C., Gething, P. W., and Brooker, S. J. (2014). Geographical inequalities in use of improved drinking water supply and sanitation across sub-Saharan Africa: mapping and spatial analysis of cross-sectional survey data. *PLoS Medicine*, 11(4):e1001626. doi: 10.1371/journal.pmed.1001626.

Radaelli, C. M. (1995). The role of knowledge in the policy process. *Journal of European Public Policy*, 2(2):159–183. doi: 10.1080/13501769508406981.

Ramcharan, B. G. (2015). Human Rights and the SDGs: A Side-lined Priority? Future United Nations Development System (FUNDS), Briefing 31. Retrieved July 19, 2019, from: https://futureun.org/media/archive1/briefings/FUNDS_Brief31_Human_Rights_SDGs_July2015.pdf.

Renooij, S. (2001). Probability elicitation for belief networks: issues to consider. *The Knowledge Engineering Review*, 16(3):255–269. doi: 10.1017/S0269888901000145.

Republic of Kenya (2010). The Constitution of Kenya: 2010. Retrieved July 19, 2019, from: https://www.wipo.int/edocs/lexdocs/laws/en/ke/ke019en.pdf.

Republic of Kenya (2019). Report of the Budget and Appropriations Committee on the proposed budget estimates for the financial year 2019/2020. Retrieved July 19, 2019, from: http://www.parliament.go.ke/sites/default/files/2019-07/REPORT%20OF%20THE%20BAC%20ON%20BUDGET%20ESTIMATES%20YEAR%202019-2020.pdf.

Rosenstock, T. S., Lamanna, C., Chesterman, S., Hammond, J., Kadiyala, S., Luedeling, E., Shepherd, K., DeRenzi, B., and van Wijk, M. T. (2017). When less is more: Innovations for tracking progress toward global targets. *Current Opinion in Environmental Sustainability*, 26:54–61. doi: 10.1016/j.cosust.2017.02.010.

Roy, B. (1991). The outranking approach and the foundations of ELECTRE methods. *Theory and Decision*, 31(1):49–73. doi: 10.1007/BF00134132.

Roy, B., Présent, d. M., and Silhol, D. (1986). A programming method for determining which Paris metro stations should be renovated. *European Journal of Operational Research*, 24(2):318–334. doi: 10.1016/0377-2217(86)90054-8.

Rutstein, S. O. (1984). *Infant and child mortality: levels, trends and demographic differentials*. Voorburg Netherlands International Statistical Institute.

Rutstein, S. O. and Rojas, G. (2006). Guide to DHS statistics. USAID. Retrieved July 19, 2019, from: https://dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_DHS-7.pdf.

Salomon, J. A. and Murray, C. J. (2001). Compositional models for mortality by age, sex and cause. Retrieved July 19, 2019, from: https://www.who.int/healthinfo/paper11.pdf.

Sartorius, B. K. and Sartorius, K. (2014). A new multidimensional population health indicator for policy makers: absolute level, inequality and spatial clustering-an empirical application using global sub-national infant mortality data. *Geospatial health*, 9(1):7–26. doi: 10.4081/gh.2014.2.

Schrage, M. (2003). Daniel Kahneman: The thought leader interview. Strategy+

Business. Retrieved July 19, 2019, from: https://www.strategy-business.com/article/03409?gko=d1233.

SDSN (2015). Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development. Sustainable Development Solutions Network. Retrieved July 19, 2019, from: http://unsdsn.org/wp-content/uploads/2015/04/Data-for-Development-Full-Report.pdf.

Seddiki, M. and Bennadji, A. (2019). Multi-criteria evaluation of renewable energy alternatives for electricity generation in a residential building. *Renewable and Sustainable Energy Reviews*, 110:101–117. doi: 10.1016/j.rser.2019.04.046.

Şengül, Ü., Eren, M., Shiraz, S. E., Gezder, V., and Şengül, A. B. (2015). Fuzzy TOPSIS method for ranking renewable energy supply systems in Turkey. *Renewable Energy*, 75:617–625. doi: 10.1016/j.renene.2014.10.045.

Shepherd, K., Hubbard, D., Fenton, N., Claxton, K., Luedeling, E., and de Leeuw, J. (2015). Policy: Development goals should enable decision-making. *Nature News*, 523(7559):152. doi: 10.1038/523152a.

Sikder, A. and Salehin, M. (2015). Multi-criteria decision making methods for rural water supply: a case study from Bangladesh. *Water Policy*, 17(6):1209–1223. doi: 10.2166/wp.2015.111.

Stafford-Smith, M., Griggs, D., Gaffney, O., Ullah, F., Reyers, B., Kanie, N., Stigson, B., Shrivastava, P., Leach, M., and O'Connell, D. (2017). Integration: the key to implementing the Sustainable Development Goals. *Sustainability Science*, 12(6):911–919. doi: 10.1007/s11625-016-0383-3.

Stewart, T. J. and Losa, F. B. (2003). Towards reconciling outranking and value measurement practice. *European Journal of Operational Research*, 145(3):645–659. doi: 10.1016/S0377-2217(02)00221-7.

Tagliapietra, S., Zachmann, G., and Fredriksson, G. (2019). Estimating the cost of capital for wind energy investments in Turkey. *Energy Policy*, 131:295–301. doi: 10.1016/j.enpol.2019.05.011.

The Economist (2015). The 169 commandments. Retrieved July 19, 2019, from: https://www.economist.com/leaders/2015/03/26/the-169-commandments.

The World Bank (2018). Boosting Shared Prosperity for the North and North Eastern Counties of Kenya. Retrieved July 19, 2019, from: https://www.worldbank.org/en/news/feature/2018/05/08/boosting-prosperity-improving-equity-in-north-and-north-eastern-kenya.

UN-Water (2018). Sustainable Development Goal 6: Synthesis Report on Water and Sanitation 2018. Retrieved July 19, 2019, from: https://sustainabledevelopment.un.org/content/documents/19901SDG6_SR2018_web_3.pdf.

UNCTAD (2014). World investment report 2014 – Investing in the SDGs: An action plan. United Nations Conference on Trade and Development. Retrieved July 19, 2019, from: https://unctad.org/en/PublicationsLibrary/wir2014_en.pdf.

UNESA (2011). Mortality estimates from major sample surveys: towards the design of a database for the monitoring of mortality levels and trends. Technical Paper. No 2011/2. United Nations Department of Economic and Social Affairs. Retrieved July 19, 2019, from: https://www.un.org/en/development/desa/population/publications/pdf/technical/TP2011-2_MortEstMajorSampSurv.pdf.

UNESA (2017). Population Division: World population prospects 2017. United Nations Department of Economic and Social Affairs. Retrieved July 19, 2019, from: https://population.un.org/wpp/.

UNGA (2010). Resolution on the Human rights and access to safe drinking water and sanitation. A/HRC/RES/15/9. United Nations General Assembly. Retrieved July 19, 2019, from: https://undocs.org/A/HRC/RES/15/9.

UNGA (2015). Transforming our world: the 2030 Agenda for Sustainable Development. A/RES/70/1. United Nations General Assembly. Retrieved July 19, 2019, from: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

UNGA (2017). Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. A/RES/71/313. United Nations General Assembly. Retrieved July 19, 2019, from: http://ggim.un.org/documents/A_RES_71_313.pdf.

UNICEF (2006). Multiple Indicator Cluster Survey Manual 2005: Monitoring the Situation of children and women. UNICEF's Division of Policy and Planning. Retrieved July 19, 2019, from: http://www.childinfo.org/files/Multiple_Indicator_Cluster_Survey_Manual_2005.pdf.

UNICEF (2018). Levels and Trends in Child Mortality Report 2018. Retrieved July 19, 2019, from: https://www.unicef.org/publications/index_103264.html.

Vaessen, M., Thiam, M., and Lê, T. (2005). Chapter XXII: The Demographic and Health Surveys. In *Household Sample Surveys in Developing and Transition Countries*. United Nations Department of Economic and Social Affairs Statistics Division. Retrieved July 19, 2019, from: https://unstats.un.org/unsd/hhsurveys/pdf/Household_surveys.pdf.

Verma, V. and Lê, T. (1996). An analysis of sampling errors for the demographic and health surveys. *International Statistical Review*, 64(3):265–294. doi: 10.2307/1403786.

Walker, W. E., Harremoës, P., Rotmans, J., Van Der Sluijs, J. P., Van Asselt, M. B., Janssen, P., and Krayer von Krauss, M. P. (2003). Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated assessment*, 4(1):5–17. doi: 10.1076/iaij.4.1.5.16466.

Wang, H., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N. M., et al. (2017). Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet*, 390(10100):1084–1150.

Wayland, J. (2019). Constraints on aid effectiveness in the water, sanitation, and hygiene (WASH) sector: evidence from Malawi. *African Geographical Review*, 38(2):140–156. doi: 10.1080/19376812.2017.1340169.

Weber, H. (2017). Politics of 'leaving no one behind': contesting the 2030

Sustainable Development Goals agenda. *Globalizations*, 14(3):399–414. doi: 10.1080/14747731.2016.1275404.

Wei, Y., Wang, Z., Wang, H., Yao, T., and Li, Y. (2018). Promoting inclusive water governance and forecasting the structure of water consumption based on compositional data: A case study of Beijing. *Science of the Total Environment*, 634:407–416. doi: 10.1016/j.scitotenv.2018.03.325.

Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. *Biometrika*, 34(1/2):28–35.

Whitney, C., Shepherd, K., and Eike, L. (2018). Decision analysis methods guide: agricultural policy for nutrition. World Agroforestry Centre (ICRAF). Retrieved July 19, 2019, from: https://hdl.handle.net/10568/98329.

WHS (2003). Report of the World Health Survey (WHS) for Burkina Faso. WHO. Retrieved July 19, 2019, from: https://www.who.int/healthinfo/survey/whsbfa-burkinafaso.pdf.

Winkler, I. T. and Williams, C. (2017). The Sustainable Development Goals and human rights: a critical early review. *The International Journal of Human Rights*, 21(8):1023–1028. doi: 10.1080/13642987.2017.1348695.

Wolf, J., Bonjour, S., and Prüss-Ustün, A. (2013). An exploration of multilevel modeling for estimating access to drinking-water and sanitation. *Journal of Water and Health*, 11(1):64–77. doi: 10.2166/wh.2012.107.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1):95–114. doi: 10.1111/1467-9868.00374.

Wu, T., Xu, D.-L., and Yang, J.-B. (2018). Multiple Criteria Performance Modelling and Impact Assessment of Renewable Energy Systems – A Literature Review. In *Renewable Energies*, pages 1–15. Springer. doi: 10.1007/978-3-319-45364-4_1.

Yap, M. L.-M. and Watene, K. (2019). The Sustainable Development Goals (SDGs) and Indigenous Peoples: Another Missed Opportunity? *Journal of Human Development and Capabilities*, pages 1–17. , doi: 10.1080/19452829.2019.1574725.

Yilan, G., Kadirgan, M. N., and Çiftçioğlu, G. A. (2020). Analysis of Electricity Generation Options for Sustainable Energy Decision Making: the Case of Turkey. *Renewable Energy*, 146:519–529. doi: 10.1016/j.renene.2019.06.164.

Yu, W., Bain, R. E., Mansour, S., and Wright, J. A. (2014). A cross-sectional ecological study of spatial scale and geographic inequality in access to drinking-water and sanitation. *International Journal for Equity in Health*, 13(1):113. doi: 10.1186/s12939-014-0113-3.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353. doi: 10.1016/S0019-9958(65)90241-X.

Zanakis, S. H., Solomon, A., Wishart, N., and Dublish, S. (1998). Multi-attribute decision making: A simulation comparison of select methods. *European Journal of Operational Research*, 107(3):507–529. doi: 10.1016/S0377-2217(97)00147-1.

Zhou, C. and Mounuddin, M. (2017). Sustainable Development Goals Interlinkages and Network Analysis: A practical tool for SDG integration and policy coherence. Institute for Global Environmental Strategies (IGES). Retrieved July 19, 2019, from: https://pub.iges.or.jp/pub/sustainable-development-goals-interlinkages.

Zlotnik, H. and Hill, K. (1981). The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality. *Demography*, 18(1):103–122. doi: 10.2307/2061052.