Exploring the interlinkages of water and sanitation across the 2030 Agenda: a Bayesian Network approach

D. Requejo-Castro[1], R. Giné-Garriga[1], A. Pérez-Foguet[1]

[1] *Engineering Science and Global Development (EScGD), Universitat Politècnica de Catalunya (UPC), Department of Civil and Environmental Engineering, Barcelona School of Civil Engineering (ETSECCPB), Barcelona, Spain, david.requejo@upc.edu, ricard.gine@upc.edu, agusti.perez@upc.edu*

Abstract

The 2030 Agenda for Sustainable Development recognizes the indivisible and integrated nature of its 17 Sustainable Development Goals (SDGs) and 169 targets, as well as the need to address these interlinkages to fully achieve its aims. In addition, the Agenda stresses the importance of "leaving no one behind", which can only be achieved by understanding the interlinkages between the Goals and by undertaking actions to bring them together for the benefit of all. Thus, the identification of these linkages will enable countries to implement the SDGs effectively by harnessing synergies between them while managing potential conflicts. Despite their significance in monitoring initiatives, indicators separately are not adequate to provide an insight into the complex cause and effect relations within global development issues. The suitability of Bayesian Networks (BNs) to integrate multiple and simultaneous relationships has been largely exploited in the literature. Taking a dedicated goal on water and sanitation (SDG 6) as starting point, this paper reviews the potential of a BNs approach to analyse the interdependency between the SDGs, the associated targets and the corresponding indicators. Available global data has been exploited to run the BNs model. Achieved results are compared with a recent research developed by UN-Water, where interlinkages between the targets under Goal 6 and other targets across the 2030 Agenda are conceptually described. The paper discusses the extent to which a BNs is a suitable system to identify and assess these linkages, relationships and synergies. The study concludes that a BNs approach is useful to accommodate the complexities and interdependencies of the SDGs targets and indicators.

Keywords: Sustainable Development Goal 6, Bayesian Networks, interlinkages, data-driven, decision-making

1. Introduction

On 25th September 2015, the 2030 Agenda resolution announced 17 Sustainable Development Goals (SDGs) with 169 associated targets, remarking their integrated and indivisible nature (United Nations General Assembly, 2015). This implies the need to balance the three dimensions of sustainable development (SD). Lack of integration across sectors in terms of strategies, policies and implementation has long been perceived as one of the main pitfall of previous approaches to SD (UN-DESA, 2015). Thus, integrated approaches are needed to elucidate the interdependencies among the SDGs and to facilitate their effective implementation by harnessing synergies among them while managing any potential conflicts (UN-Water, 2016b).

In this context, indicators represent the backbone of monitoring progress towards the SDGs at all levels (i.e. local, national, regional and global). Additionally, a sound indicator framework will turn the SDGs and their targets into a management tool to help countries develop implementation strategies, to allocate resources efficiently and to increase the accountability of all stakeholders (SDNS, 2015). To achieve this, the SDGs require annual reporting of high-quality data from all countries. This, in turn, will require much greater investments in building independent, impartial national statistical capacities and strengthening statistical quality and standards. To date every target has had at least one lead technical or specialist agency, responsible for coordinating data standards and collection, ensuring harmonization, and providing technical support where necessary (SDNS, 2015). Some of these custodian agencies are already elaborating data baselines and reports, and therefore, presenting a starting point for SDGs achievement. However, a baseline for several of the targets remains unavailable (United Nations General Assembly, 2015). It is remarkable in this context the development of an unofficial SDG Index (SDSN, 2016), as a tool to provide indicative country-level estimates, covering 63 indicators for 149 out of the 193 UN member countries. For each indicator, one ranking is produced (SDSN, 2016). Despite of the likely utility of this approach, the global nature of SDGs should permeate the monitoring framework, by exploiting the interlinkages among SDGs targets and indicators, as well as by promoting multi-sectorial approaches.

The literature shows several contributions that address, to a certain extent, the SDGs interlinkages. To take some examples, and from a more conceptual and qualitative perspective, SDGs have been read as a network of targets connecting the different goal areas (UN-DESA, 2015), they have been organized through a grading system of interactions (Nilsson et al., 2016) or they have been particularly analysed in relation to one SDG (e.g. UN-Water, 2016b). On the other hand, from a quantitative point of view, important efforts have undergone to establish the links among different SDGs and to support planning and decision-making for their implementation (Collste et al., 2017; Khalili et al., 2017).

In parallel, BNs have been extensively used to explore the interdependencies and cause-effect relationships, simulating complex problems that involve a large number of variables that are highly interlinked (Dondeynaz et al., 2013). Briefly, BNs are directed acyclic graphs that exploit the duality between an interaction graph and a probability model (Castelletti and Soncini-Sessa, 2007). While its graphical structure provides a visual representation of the logical relationship among variables, and thus providing an excellent language to communicate and discuss these relationships, conditional probabilities quantify these relationships. Basically, BNs are made up of three elements (Bromley, 2005): i) a series of nodes associated with the variables relevant to a particular study; ii) the links among these variables and their direction which represent, respectively, the existing dependency among them and the cause-effect relationships, and iii) the conditional probability tables (CPTs) that quantify the extent to which one node is likely to be affected by the others (Bromley, 2005; Cain, 2001).

BNs have been applied in fields such as medicine and artificial intelligence. They have been also applied as an aid to decision making in the field of water resources (Bromley et al., 2005; Castelletti and Soncini-Sessa, 2007; Henriksen and Barlebo, 2008; Molina et al., 2013, 2009), and to support monitoring WaSH services in an interdisciplinary, holistic way (Alok, 2002; Fisher et al., 2015). Among its strengths, BNs allow incorporating data and knowledge from different sources and domains (i.e. economic, social, physical or environmental), and they are especially helpful when there is scarcity or some degree of uncertainty in the data (Bromley et al., 2005; Castelletti and Soncini-Sessa, 2007; Henriksen and Barlebo, 2008).

Against this background, this study seeks to explore the interdependencies across the 2030 Agenda through a Bayesian Network (BN) approach. Specifically, the focus is on the goal dedicated on water and sanitation, thus the paper explores the interlinkages between the targets under Goal 6 and other targets across the 2030 Agenda. SDG 6 reads as follows: "to guarantee the availability of water and its sustainable management and sanitation for all", and includes eight specific targets: six of them are based on outcomes, and the remaining two are based on the means of implementation. In more detail, targets 6.1 and 6.2 relate to drinking-water and sanitation and hygiene, respectively. Targets 6.2 and 6.3 expand the framework beyond the use of sanitation facilities to cover the full sanitation chain. Targets 6.4 and 6.5 refer to water-use efficiency and the implementation of Integrated Water Resources Management (IWRM), respectively. Target 6.6 focuses on healthy water-related ecosystems. Finally, targets 6.a and 6.b suggest the importance of international cooperation and local stakeholder participation to achieve previous targets. This paper takes as a reference point a UN-Water Analytical Brief that describes the linkages across the 2030 Agenda with Goal 6, providing an overview of the target-level linkages and how they are interdependent (UN Water, 2016). We then apply a BN technique to identify and assess these linkages. While the literature shows similar analysis with specific focus on e.g. target 6.1 (Cronk and Bartram, 2017), 6.1 and 6.2 (Dondeynaz et al., 2013; Giné-Garriga et al., 2018) or 6.4 and 6.5 (Mohajerani et al. 2017; Molina et al. 2009), none of them consider Goal 6 and its targets in a holistic and interdependent way.

The article is structured as follows. Section 2 describes the methods for data selection and data processing, as well as for BNs construction. Main results are presented and discussed in Section 3. Major findings are highlighted in Section 4 to conclude the study.

2. Methods

This section provides a detailed description of the steps followed to conduct the analysis. First, we introduce the initial settings established in the study. Second, we present the criteria for data selection, including the assumptions made for missing data treatment. Third, we define the scenarios for the analysis. Fourth, data normalization process is described. Finally, we set up the conditions for the analysis.

Initial settings

To deal with BNs generation, it is possible to use a wide range of both free and commercial software available which, at the same time, employ one of the three types of structure learning methods, namely constraint-based, score-based and hybrid approaches (Liu et al., 2017; Madsen et al., 2016). This study uses "R" and its package "bnlearn" developed by Scutari (2010). In this case, constraint-based and score-based methods for network structure learning are available. Briefly, the former learns the network structure by analyzing the probabilistic relations with conditional independence (CI) tests; and the latter assigns a score to each candidate BN and try to maximize it with some heuristic search algorithm (Scutari, 2010). Further details of these algorithms are not provided, but have been extensively described recently by Liu et al. (2017). In this study, constraint-based methods will be used, as they also provide a wide range of options to combine network structure learning algorithms (SLA) and CI tests. To do this, a bootstrap technique is applied to estimate the strength of each link, taking into consideration the empirical frequency over a set of networks learned from bootstrap samples. Specifically, it computes the probability of each link and the probabilities of each link's directions conditional on the link being present in the network (Scutari and Ness, 2018). Through this technique, and comparing different SLA + CI tandems, we explore the interlinkages among the SDGs and we identify those strong ones from current data availability.

Indicator selection

A first screening of eligible indicators was conducted based on the study developed by UN-Water (2016b). Initially, a total of 74 targets were connected conceptually, but only 42 out of these 74 (57%) presented associated data (UNSD, 2018). To determine the final set of indicators for inclusion in this research, several criteria were defined as detailed below:

- *Outcome-related indicators.* As per the definition of the SDGs, indicators based on outcomes were selected and those related to the means of implementation were discarded (United Nations General Assembly, 2015). Under this rule, SDG 17 and a number of indicators were excluded from the analysis. Indicator of target SDG 6.a was however included because of its relevance for the study.
- *Data availability.* The list of 234 countries provided by the Joint Monitoring Programme for Water Supply, Sanitation and Hygiene was initially considered. However, this number was reduced to 200 by removing those countries with very low data availability. In terms of indicators' selection, it was limited to those indicators with data from at least 75% of countries (an exception was made in relation to SDG 1 indicators, where data were available for only 65% 200 countries).
- *Global relevance.* To improve data comparability, various decisions were made. First, from those indicators proposed to measure access to safely managed water and sanitation services (targets 6.1 and 6.2), the level of service "at least basic" (JMP, 2017) was taken into consideration. Second, five unofficial indicators were included when the Goals were misrepresented with just one indicator. Three of them were obtained from i the SDG Index baseline (SDSN, 2016). Third, other development-related indicators available in the UNDP database were employed, such as "mean year of schooling" and "Inequality-adjusted Human Development Index" (UNDP, 2018).

- *Timeliness.* For each indicator / country pair, the more recent data was selected. For each indicator, when data is obtained from different years, the time period is specified.

In total, a subset of 38 indicators was identified (see Table 1), covering the selected 200 countries. However, when data was merged, only 68 countries with complete information (regarding all indicators considered) were distilled due to the high rate of missing data (from now on, we refer it as NA data). Thus, special attention was paid to this aspect.

NA treatment

BNs can only be fully run when there is an absence of NA data. With the aim to increase the final sample of countries and address NA data problems, a set of assumptions were made.

- When SDG regional data was available, this was used to fill information gaps as a proxy. For example, there was information for 181 countries as regard SDG indicator 3.9.2 ("mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene") and regional data was applied to reach 200 countries with information. In total, this process was done to 65% of the indicators selected, varying the amount of countries using regional data for each indicator.
- Those countries with a value of Human Development Index (UNDP, 2016) defined as "very high" (VH HDI) were assigned to the best possible scenario, otherwise they would not have been included because of NA data. For example, SDG indicator 1.1.1 related to "proportion of population below the international poverty line of US$1.90 per day", we assumed 0% for VH HDI countries. Same assumption was made regarding indicator 6.a.1 ("total official flows for water supply and sanitation, by recipient"), considering the reception of 0 billions of USD for these countries.
- In contrast, a similar hypothesis was made in reference to those countries defined as "developing" (United Nations, 2017) by assigning a value corresponding to the worst scenario. It should be noted the existence of VH HDI countries but defined as "developing" as well (i.e. Argentina or Saudi Arabia). As examples, SDG indicators 6.5.1 ("integrated water resources management implementation") and 9.5.1 ("research and development expenditure as a proportion of GDP") were assumed as zero when there was NA data associated with these countries.
- Different values from the literature were considered. For example, 1.2% prevalence of undernourishment rate (SDG indicator 2.1.1) for each developed country with NA data (SDSN, 2016) or 6% of European urban dwellers living in extremely precarious conditions (UN-Habitat, 2015), which we computed it as proportion of urban population living in slums (SDG indicator 11.1.1).

Normalization

Prior BNs analysis, indicators were normalized between 0 and 100, with 0 denoting worst performance and 100 describing the optimum. In so doing, those indicators expressed in percentage were kept invariable. On the other hand, lasted indicators were normalized using "min-max" technique. This method normalizes indicators to have an identical range by subtracting the minimum value and dividing by the range of the indicator values (OECD, 2008). In this case, no limits were fixed and maximum and minimum values were defined according to the best and worst performance of the indicators, applying logarithms when valued differed in orders of magnitude. Finally, a last transformation was required for several indicators according to "more is better" scale. For example, SDG indicator 9.4.1 ("emissions of carbon dioxide"), which represents a "less is better" goal, was employed calculating the complementary value. This is, if after normalization this indicator is 30, then a final value of 70 was considered. In contrast, similar examples to SDG indicator 7.1.1 ("proportion of population with access to electricity") were kept constant according to the scale fixed.

__Table 1.__ SDG indicators employed in Bayesian Networks analysis

SDG	Indicator (units)	Year(s)	Source(s)	Countries with information	NA treatment				Notes
					Regional data	VH HDI	Developing	Literature	
1	1.1.1. Poverty line of US$1.90 per day (% population)	2010-2014	WB	129	✓	✓			a
	1. 1.2. Poverty line of US$1.90 per day, “working poor”, 15 years and over (% employees)	2016	ILO	137	✓	✓			a
2	2.1.1. Prevalence of undernourishment (% population)	2015	FAO	171	✓			✓	a
	2.2.1. Proportion of stunted children (% of children under the age of 5 years)	2016-2005	UNICEF/WHO/ WB	149	✓			✓	a
3	3.1.1. Maternal mortality ratio (number per 100,000 live births)	2015	UNICEF/WHO/ WB/UNFPA	183	✓				a
	3.2.1. Under-five mortality rate (number per 1,000 live births)	2015	UN IGME	195	✓				a
	3.9.2. Mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene (number per 100,000 people)	2012	GHO/WHO	183	✓				a
4	4.2.2. Participation rate in organized learning (% children)	2016-2006	UNESCO	161	✓				
	UNDP_4.1. Mean years of schooling (years)	2015	UNDP	188					
5	5.5.1. Proportion of seats held by women in national parliaments (% women)	2017	UN WOMEN/ IPU	191	✓				
	UNDP_5. Gender Inequality Index (0-1)	2015	UNDP	159					a, b
6	6.1.1. Proportion of population using AT LEAST BASIC drinking water services (% pop.)	2015	JMP	226	✓				b
	6.2.1. Proportion of population using AT LEAST BASIC sanitation services (% pop.)	2015	JMP	223	✓				b
	6.4.1. Water-Use Efficiency (USD/m3)	2014	FAO	166					a
	6.4.2. Freshwater withdrawal as a proportion of available freshwater resources (%)	2014	FAO	171	✓				a
	6.5.1. Integrated Water Resource Management implementation (%)	2012	UNDP	133			✓		
	6.a.1. Total official flows for water supply and sanitation, by recipient (billions USD)	2015	OECD	137	✓	✓			
7	7.1.1. Proportion of population with access to electricity (%)	2014	GTF	212	✓				
	7.1.2. Proportion of population with primary reliance on clean fuels and technology	2014	GHO/WHO	226	✓				

SDG	Indicator (units)	Year(s)	Source(s)	Countries with information	NA treatment				Notes
					Regional data	VH HDI	Developing	Literature	
8	8.1.1. Growth rate of real GDP per capita (%)	2015	UN	209	✓				
	8.5.2. Unemployment rate, 15 years old and over, both sexes (% population)	2016-2008	ILO	185	✓				a
9	9.1.2. Passenger volume, by air transport (number people)	2015	ICAO	154	✓				
	9.4.1. Emissions of carbon dioxide (millions metric tons)	2015	IEA / UNIDO	141	✓				a
	9.5.1. Research and development expenditure (% GDP)	2015-2002	UNESCO	134			✓		
10	*UNDP_10. Inequality-adjusted Human Development Index (0-1)*	2015	UNDP	151					b
	SDSN_10. GINI Index (0-100)	2012-2003	WB	134					a, b
11	11.1.1. Urban population living in slums (%)	2014-2005	UN HABITAT	94	✓			✓	a
	11.5.2. Direct disaster economic loss, average annual loss (USD)	2016	UNISDR	211					a
	11.6.2. Annual mean levels of fine particulate matter ($PM_{2.5}$) in cities ($\mu g/m^3$)	2012	WHO	184	✓				a
12	12.2.2. Domestic material consumption (millions metric tons)	2010	UNEP	192					a
	12.4.1. Compliance with the Stockholm Convention on hazardous waste and other chemicals (%)	2015	UNEP	185	✓				
13	13.1.2. Number of countries with legislative and/or regulatory provisions been made for managing disaster risk (Yes/No)	2015-2013	UNISDR	119			✓		
	SDSN_13. Climate Change Vulnerability Monitor Index (0-1)	2014	HCSS	146					a, b
14	14.5.1. Coverage of protected areas in relation to marine areas (%)	2016	UNEP/WCMC	175	✓				
15	15.1.1. Forest area as a proportion of total land area (%)	2015	FAO	225	✓				
	15.5.1. Red List Index (0-1)	2017	IUCN	226					b
16	16.3.2. Unsentenced detainees (% prison population)	2015	UNODC	142	✓				a
	16.10.1. Number of cases of killings of journalists and associated media personnel, both sexes (number people)	2015	OHCHR	223					

Notes

In italics: Indicators not included in model 2 / a: Indicators transformed from “less is better” to “more is better” scale / b: Indicators not included by the Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs)

BNs analysis and models generation

"bnlearn" package implements the following constraint-based SLA (Scutari, 2010): grow-shrink (gs), incremental association (iamb), fast incremental association (fast.iamb) and interleaved incremental association (inter.iamb). Additionally, CI tests must be chosen regarding data typology. In this case, in order to enable the analysis, we considered that all indicators follow a normal distribution. Thus, as all values were considered as continuous, we selected those CI tests available for this case. In addition, this selection was carried out according to a robustness criterion. This is, we chose those tests which provided the same results when running BNs several times. Thus, CI tests selection was reduced to linear correlation (cor), Fisher's Z (zf) and mutual information (mi-g) tests (Scutari, 2010). By applying the combination of both SLA and CI tests to the database, 12 networks were obtained. It is of standing out that the network structure is learnt from data (data-driven).

Bootstrap technique was applied to each network by generating 500 potential BNs. By establishing a strength threshold value (i.e. 0.8 over 1), which represents the frequency of links, we distilled those relationships common to each network generated by the tandem SLA + CI test. In addition, this technique also provides the direction of these links, which is understood as the frequency of link direction conditional on the link's presence (Scutari and Ness, 2018). A quantitative value between 0 and 1 denotes this direction. For example, if a link is established from SDG indicator 7.1.1 to 1.1.1 with a value 0.7, this also means that from 1.1.1 to 7.1.1 is 0.3. A value of 0.5 would be translated into an undirected link. On the other hand, it should be highlighted that, when applying this technique, the result is not visualized on the structure of the network obtained by combining the tandem SLA + CI test.

As regard network visualization, from all candidates, a unique BN is presented to exemplify the graphical results obtained. For its selection, a criterion of robustness, regarding the SLA, was applied. First, each SLA was combined with the 3 CI tests. Second, the resulting networks were compared. Third, those sets of 3 options presenting more similarities allowed selecting the SLA. Finally, the tandem SLA + CI test identifying a higher number of interlinkages was selected.

We generated two models:

- First model is made up of 98 countries with full information regarding the final list of 38 indicators. This represents 49% of the countries considered. From this sample, 19% correspond to "developed" countries.
- In second model, with the aim increase country sample, we reduced the number of indicators to 34. Thus, 130 countries are considered finally for the analysis (15% of them "developed"). However, this causes a representation absence of SDG 10.

Finally, we carried out a detailed analysis for each model separately, identifying similarities and differences in relation to the interlinkages between water and sanitation indicators pointed out in the UN-Water Analytical Brief (2016). In addition to this, we focused on the strongest links as a relevant output from the models. As a last step, we provided an assessment of both models by identifying common interlinkages.

3. Results and Discussion

This section presents and discusses achieved results. For both models, the scheme of the network is visualized in Figures 1 and 2; and key causal relationships are highlighted (Table 2). Main interlinkages between the targets under Goal 6 and other targets across the 2030 Agenda are presented in Table 3. They are compared with conceptual linkages described in the UN-Water Analytical Brief (2016). The discussion seeks to highlight main synergies and potential conflicts between data-driven results (Model 1 and Model 2) and theoretical interdependencies (UN-Water Brief).

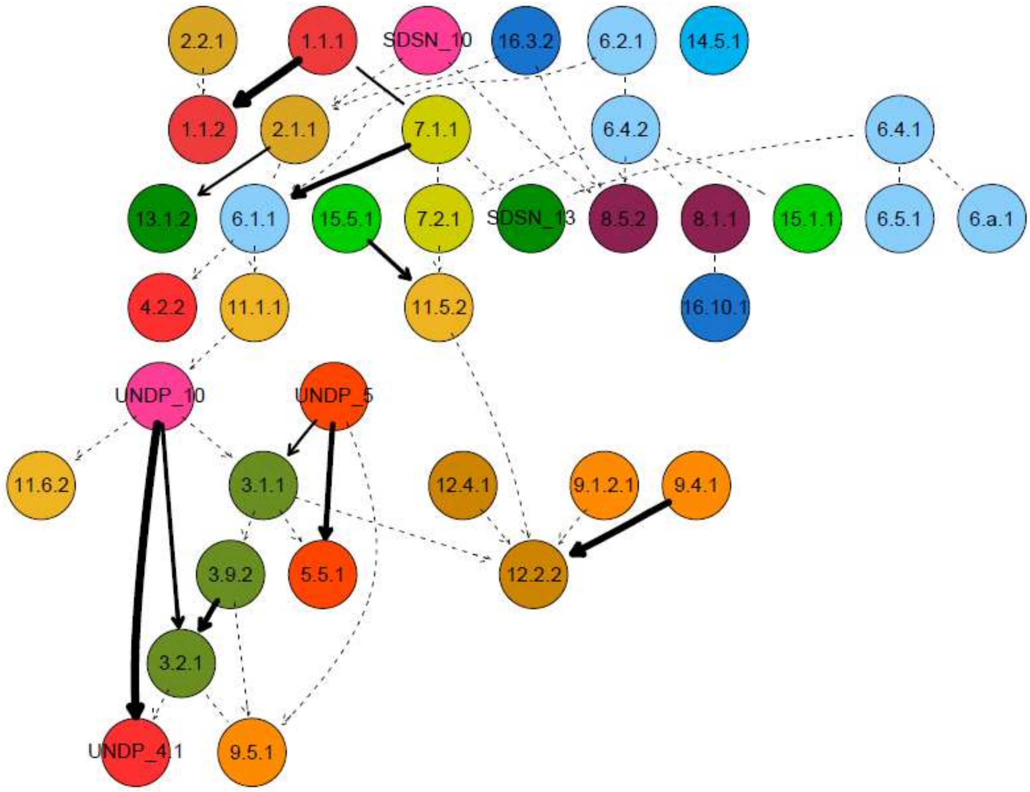

Figure 1. *BN model 1 defined by SLA "inter.iamb" and CI test "mi-g". Bold links show relationships with strength values between 0.8 and 1 (see Table 2)*

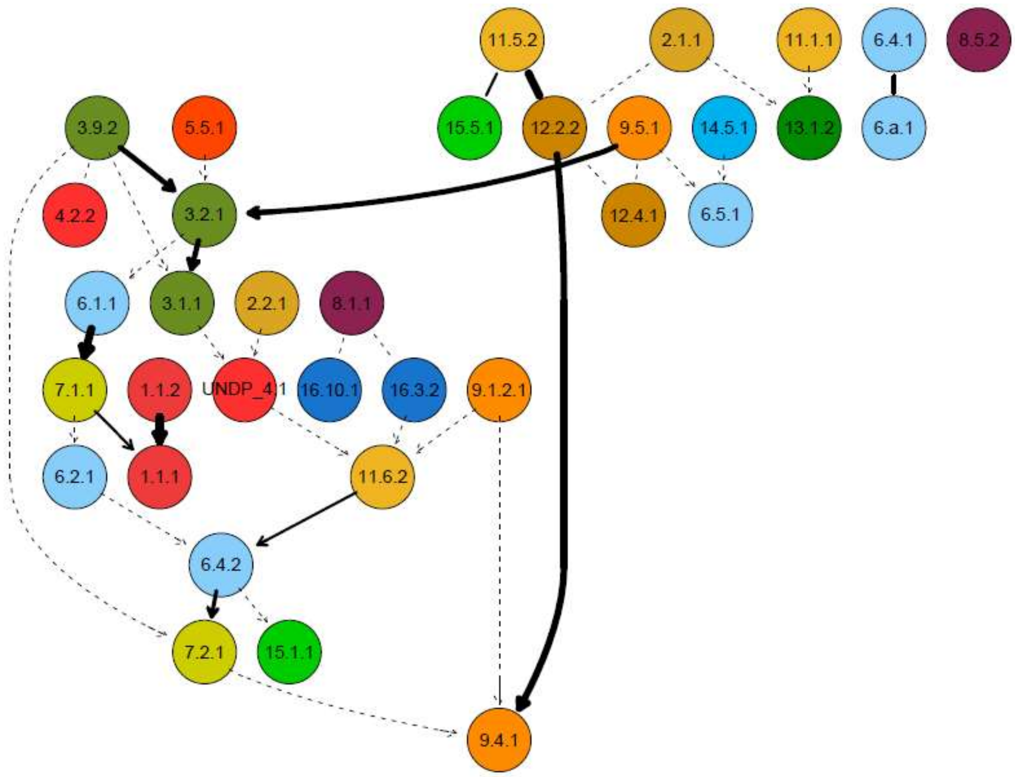

Figure 2. *BN Model 2 defined by SLA "gs" and CI test "mi-g". Bold links show relationships with strength values between 0.8 and 1 (see Table 2).*

Table 2. *Common interlinkages, regarding model 1, identified among 12 BNs as the combination of detailed SLA + CI tests. For "strength" and "direction" columns, maximum and minimum values from all networks are presented.*

Model 1 - Synergies and Conflicts			
Cause	**Effect**	**Strength**	**Direction** [a]
Extreme poverty (1.1.1)	Multidimensional poverty (1.1.2)	0.99 - 0.97	0.64 - 0.57
Maternal mortality (3.1.1)	Gender Inequality (UNDP_5)	0.90 - 0.83	0.85 - 0.80
Mortality rate attributed to unsafe WaSH services (3.9.2)	Under-five mortality rate (3.2.1)	0.88 - 0.80	0.57 - 0.47
Mean years of schooling (UNDP_4.1)	Inequality Human Development Index (UNDP_10)	1 - 0.99	0.57 - 0.49
Seats held by women in national parliaments (5.5.1)	Gender Inequality (UNDP_5)	0.89 - 0.83	0.84 - 0.77
Access to at least basic drinking water services (6.1.1)	Access to electricity (7.1.1)	0.95 - 0.86	0.55 - 0.51
Domestic material consumption (12.2.2)	Emissions of carbon dioxide (9.4.1)	0.98 - 0.95	0.53 - 0.47
Model 2 - Synergies and Conflicts			
From	**To**	**Strength**	**Direction** [a]
Extreme poverty (1.1.1)	Multidimensional poverty (1.1.2)	0.98 - 0.96	0.56 - 0.53
Under-five mortality rate 3.2.1	Maternal mortality (3.1.1)	0.97 - 0.92	0.57 - 0.51
Mortality rate attributed to unsafe WaSH services (3.9.2)	Under-five mortality rate (3.2.1)	0.97 - 0.94	0.81 - 0.78
Access to at least basic drinking water services (6.1.1)	Access to electricity (7.1.1)	1 - 0.99	0.58 - 0.50
Freshwater withdrawal (6.4.2)	Levels of fine particulate matter in cities (11.6.2)	0.87 - 0.82	0.62 - 0.49[1]
Cooperation and capacity-building (6.a.1)	**Water-Use Efficiency (6.4.1)**	0.89 - 0.85	0.61 - 0.55
Access to electricity (7.1.1)	Extreme poverty (1.1.1)	0.82 - 0.80	0.74 - 0.69
Primary reliance on clean fuels and technology (7.2.1)	**Freshwater withdrawal (6.4.2)**	0.90 - 0.83	0.59 - 0.52
Research and development expenditure (9.5.1)	Under-five mortality rate (3.2.1)	0.96 - 0.62	0.77 - 0.69
Direct disaster economic loss (11.5.2)	Forest area as a proportion of total land area (15.5.1)	0.91 - 0.84	0.66 - 0.56
Domestic material consumption (12.2.2)	Emissions of carbon dioxide (9.4.1)	0.98 - 0.96	0.57 - 0.48
Domestic material consumption (12.2.2)	Direct disaster economic loss (11.5.2)	0.99 - 0.98	0.61 - 0.50

Notes:

a) A value < 0.50 means the inverse cause-effect relationship

b) In bold, indicators related to SDG 6

c) Underlined, those interlinkages commonly identified by both models

Table 3. *Summary of interlinkages between targets related to Sustainable Development Goal 6.*

Indicator	**Main synergies / Potential Conflicts**		**UN-Water Analytical Brief (2016)**
	Model 1	**Model 2**	
Access to at least basic drinking water services (SDG 6.1.1)	Access to at least basic sanitation services (1.4.1 & 6.1.2); Undernourishment (2.1.1); Improve access to quality pre-primary education (4.2.2); Access to electricity (7.1.1); Upgrade slums and access to safe and affordable housing (11.1.1)	Under-five mortality (3.2.1); Access to electricity (7.1.1)	Eradicate poverty (Target 1.1 & 1.2); Access to basic services (T1.4); End hunger and malnutrition (T2.1 & T2.2); Reduce maternal mortality (T3.1) and end preventable deaths of children under 5 years of age (T3.2); Combat water-borne diseases (T3.3 & T3.9); Improve quality education (T4.1, T4.2, T4.3 & T4.5); Achieve gender equality (T5.1, T5.2, T5.4 & T5.5); Economic growth (T8.1); Create and maintain decent jobs (T8.5, T8.6 & T8.8); Reduce inequalities (T10.1 to 10.3)
Access to at least basic sanitation services (SDG 6.2.1)	Access to at least basic drinking water services (1.4.1 & 6.1.1); Freshwater withdrawal (6.4.2)	Freshwater withdrawal (6.4.2); Access to electricity (7.1.1)	
Water-Use Efficiency (SDG 6.4.1)	Integrated Water Resource Management (6.5.1); Cooperation and capacity-building (6.a.1); Climate Change Vulnerability (SDSN_13)	Cooperation and capacity-building (6.a.1)	Build the resilience of the poor (T1.5); End hunger and malnutrition (T2.1 & T2.2); Increase agricultural productivity (T2.3); Promote resilient agricultural practices (T2.4); Increase energy efficiency (T7.3); Promote inclusive and sustainable industrialization (T9.2); Adopt environmentally sound technology (T9.4); Reduce the impacts of water-related disasters (T11.5); Increase sustainable consumption and production (T12.1 to 12.8)
Freshwater withdrawal (SDG 6.4.2)	Primary reliance on clean fuels and technology (7.2.1); Growth rate of real GDP per capita (8.1.1); Unemployment rate (8.5.2); Forest area as a proportion of total land area (15.1.1)	Access to at least basic sanitation services (1.4.1 & 6.1.2); Primary reliance on clean fuels and technology (7.2.1); Levels of fine particulate matter in cities (11.6.2); Forest area as a proportion of total land area (15.1.1)	
Integrated Water Resource Management (SDG 6.5.1)	Water use efficiency (6.4.1)	Research and development expenditure (9.5.1); Coverage of protected areas in relation to marine areas (14.5.1)	Build the resilience of the poor (T1.5); End hunger and malnutrition (T2.1 & T2.2); Increase agricultural productivity (T2.3); Promote resilient agricultural practices (T2.4); Achieve gender equality (T5.5); Create and maintain decent jobs (T8.5 & T8.6); Adopt environmentally sound technology (T9.4); Reduce the impacts of water-related disasters (T11.5); Increase sustainable consumption and production (T12.1 – 12.8); Combat Climate Change (T13.2 & T13.3); Protect and conserve marine and coastal ecosystems (T14.1, T14.2 & T14.5); Protect and conserve terrestrial ecosystems (T15.1, T15.3 & T15.5); Promote coherent policies and the rule of law, and build effective, accountable and inclusive institutions (T16.3, T16.6, T16.7 & T16.10)
Note: Main synergies: Links that are likely to be mainly positive in that they may be mutually reinforcing or have positive interdependencies; Potential conflict: Links that usually still have positive aspects, but there exists a potential conflict in one or both directions unless policies, plans and implementation address the constraints and trade-offs.			

As regards the access to safe and affordable drinking water (SDG 6.1), it is observed that Model 1 does not predict the linkages between access to basic water services (6.1.1) and good health. Such linkages are to a certain extent represented in Model 2, particularly by relating access to improved water sources to under-five mortality rate (3.2.1). There is an extensive literature dealing with the effects of improved water supply, sanitation and hygiene on health (e.g. Bartram et al, 2005; Cairncross et al, 2010; Esrey et al, 1991; Feachem, 1984). Interestingly, both models underline the links between under-five mortality rate (3.2.1) and mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene (3.9.2). On the other hand, Model 1 shows the interlinkages between access to basic water services and prevalence of undernourishment (2.1.1), access to at least basic sanitation services (6.2.1), access to electricity (7.1.1) and population living in slums (11.1.1). This is coherent with several statements provided by UN-Water (2016b) which read "The Goal on poverty calls for universal access to basic services [T1.4], which include food [T2.1], water and sanitation [T6.1, T6.2], energy [T7.1] and housing [T11.1]… An adequate and reliable water resource is a prerequisite for the production of food [T2.1] and energy [T7.1], as well as industry [T9.1, T9.2], which in turn provide jobs [T8.5] and sustain cities and communities [T11.1]… Water supply, sanitation and hygiene services are also key factors in improving student health and thus educational outcomes [T4.1 – 4.3]". However, Model 2 only identifies the link with access to energy (7.1.1). However, none of the models show clear links with Goals related to sustainable economic growth (8), sustainable industrialization (9) and reduction of inequalities (10).

In relation to sanitation services (SDG 6.2), both models do not predict relevant interlinkages between access to basic sanitation services and the indicators of Goals 3 (health), 4 (education) and 5 (gender equality), which are identified in UN-Water Brief (2016). Here, it is pointed out that increasing the access to these services in the public sphere, including facilities for menstrual hygiene management, supports effective participation of women and girls at all levels in educational, political, economic and public life. However, Model 1 shows a direct link to access to basic water services (6.1.1) which, in turn, is linked to the indicators detailed above. On the other hand, both models underline the link to freshwater withdrawal (6.4.2), which reinforces the relationship between access to water and sanitation. Finally, Model 2 presents a relationship between access to sanitation and access to energy (7.1.1). This might be interpreted as a positive correlation (i.e. countries access or lack to both services), since wastewater requires energy for its treatment and management.

The analysis of water use and scarcity (SDG 6.4) shows clear interlinkages between water-use efficiency (6.4.1) and integrated water resource management (6.5.1), international cooperation in the water sector (6.a.1) and climate change. As pointed out in the UN-Water Brief (2016b), IWRM balances the water demands from various sectors. This is achieved by considering all levels of management, including transboundary cooperation and upstream–downstream uses as appropriate. This statement is coherent with the results obtained. On the other hand, CCVMI is constructed by three indicators; increase in weather-related disasters, sea level rise and loss of agricultural productivity. It is in relation to the latter where a relationship could be established due to the key role that water plays in agriculture. On the other hand, Model 2 only shows a direct relationship with official flows for water supply and sanitation (6.a.1). Although these results are coherent, several interlinkages related to Goals 2 (hunger and malnutrition), 7 (energy), 9 (sustainable industrialization) or 12 (sustainable consumption and production patterns) are missed. Nonetheless, when paying attention to freshwater withdrawal associated indicator (6.4.2), several of previously links are identified by both models. Thus, Model 1 predicts the relationships between 6.4.2 and indicators 6.2.1 (sanitation), 7.2.1 (renewable energy share in the total final energy consumption), 8.1.1 (growth rate of real GDP per capita), 8.5.2 (unemployment rate) and 15.1.1 (forest area as a proportion of total land area). As highlighted in UN-Water (2016b): "investments in water and sanitation provide significant economic and social returns, as well as generate employment [1, 8]. For example, there is a positive linkage between using more efficient water technologies [6.3, 6.4] and support to poor farmers [2.3, 8.5, 10.1]". This is coherent with the results provided by the model. In addition to this, how land is used and managed affects water availability. This fact supports the link between 6.4.2 and 15.1.1. Similarly, Model 2 identifies the interlinkages with indicators 6.2.1 (sanitation), 7.2.1 (clean energy) and 15.1.1 (forest area). While missing those relationships with Goal 8 (sustainable economic growth), a new connection is identified in the indicator 11.6.2 (mean levels of fine particulate matter in cities). There are different sources in which these particles are ejected to the atmosphere. One of the most common is the combustion of carbon for energy production. This, in turn, requires a high amount of water, increasing the level freshwater withdrawal.

As far as the implementation of integrated water resources management (SDG 6.5), both models do not predict most of the interlinkages defined by the UN-Water Brief (2016b), which states that implementing IWRM provides a framework for addressing many of the linkages by balancing the needs of different sectors and stakeholders. On the one hand, Model 1 shows a relationship with water-use efficiency (6.4.1), which was mentioned previously. On the other hand, Model 2 identifies new connections. In this case, relationships with indicators 9.5.1 (research and development expenditure) and 14.5.1 (coverage of protected areas in relation to marine areas) are observed. Implementing IWRM requires important investments, including research and development. Thus, it is coherent to understand this connection. On the other hand, and as affirmed in UN-Water (2016b), implementing target 6.5 (in combination with targets 6.3and 6.6) mutually reinforce targets on protecting and conserving marine and coastal ecosystems (targets 14.1, 14.2 and 14.5). This is also coherent with the results obtained.

BNs and sustainable development

This paper shows the likely utility of a BNs approach to describe cause-effect relationships among global development concepts. As stated previously, understanding these linkages enables the full exploitation of synergies, conflict resolution and trade-offs balances. On the basis of this linkages, integrated planning and management can support decision-making, reduce investment costs and facilitate implementation of a number of strategies that are geared towards sustainable development.

This study exploits official national data to run the BNs. It then describes the particular nexus between i) water and health, and ii) water and energy. This reinforces the idea of undertaking actions from a multi-sectorial perspective if the goal of “leaving no one behind” is to be achieved (UN-Water, 2016b).

Having said this, current data gaps for a number of indicators and countries clearly limit the scope of the study and the lessons learnt. Credible data are the lifeblood of decision-making (IEAG, 2014), as are needed to underpin sector advocacy, stimulate political commitment and trigger well-placed investment towards optimum health, environment and economic gains (UN-Water, 2016a). Moreover, data are the raw material for BNs models, since the validity of their outcomes are directly dependant on both data quality and quantity. At present, there are several global initiatives that are monitoring different aspects of the development Agenda, but a coherent framework is missing. To the extent that existing efforts expands to ensure harmonised and integrated monitoring of SDGs, increasing high-quality data availability, it is expected that a wide range of potential uses may emerge in relation to BNs applications. For instance, the study might focus on the sub-national level where decisions are taken by decentralized administrations. In this sense, it exits the possibility to simulate different scenarios and to infer the impact of potential interventions. To illustrate this idea, and taking as an example Figure 2, different levels of improvement (investments) regarding node 3.9.2 (mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene) might be simulated (through CPTs modification). Thus, the impacts on immediate nodes 3.2.1 (under-five mortality rate), 4.2.2 (percentage of children participating rate in organized learning) and 7.2.1 (proportion of population with primary reliance on clean fuels and technology) could be evaluated. Similarly, further connected nodes to the latest could be analyzed as well. This requires a shift on the way that potential interventions are assessed, encouraging decision-makers to integrate interdisciplinary perspectives which, in turn, it is essential for sustainable development.

In this scenario-based analysis, and considering the difficulty of countries to collect all the information related to the 17 SDGs, BNs might be also especially helpful when there is scarcity or some degree of uncertainty in the data. For example, and following with the same example, if there is no information related to the mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene, it is possible to set up different hypothesis and assess the potential impacts on associated nodes.

A last potential application of BNs falls on its inverse use (Carriger et al., 2016; Pérez-Foguet et al., 2017). Contrary to the methodology in scenario simulation, where one or more nodes can be modified to assess the impact in subsequent nodes, this use permits to establish a desirable value in an objective node and obtain the new values of connected nodes in order to reach that value. For example, and regarding Figure 2, it is possible to define a goal for node 1.1.1 (population below the international poverty). Then, when running the model, the values of associated nodes (in terms of conditional probabilities) are modified consequently. Finally, the analysis should be focused on input nodes. In this example, it would be possible to observe to which extent nodes 3.9.2 (water, sanitation and hygiene services), 5.5.1 (proportion of seats held by women in national parliaments) and 9.5.1 (research and development expenditure) change. Thus, specific interventions can be design, assess and implement. This possibility might be useful for planning purposes, allowing decision-makers to optimize these interventions in order to achieve the expected results.

4. Conclusions

In this study, a BN analysis have been carried out to elucidate the interlinkages of SDG 6, related to water and sanitation, across the 2030 Agenda. In so doing, a data-driven approach has been applied.

Taking as a starting point the Analytical Brief of UN-Water, we have compared the results provided by BNs models with the main water and sanitation interlinkages on the SDGs identified by this publication. We have demonstrated the coherence of the results obtained and we conclude that BNs approach is a potential and useful tool to accommodate the complexities and interdependency of the targets formulated within the 2030 Agenda.

We are concerned of the limitation of this study due to current data availability. We highlight the correlation between strong interlinkages and number of countries considered. This fact, in combination with the differences between both models applied, requires global consolidated data for further analysis. In addition to this, we have pointed out potential uses which could support the required multi-sectorial approach if sustainable development is to be achieved. These aspects suggest the way forward.

References

Bartram, J., Lewis, K., Lenton, R., and Wright, A., 2005. Focusing on improved water and sanitation for health. Lancet 365, 810–812.

Bromley, J. 2005. Guidelines for the Use of Bayesian Networks as a Participatory Tool for Water Resource. Based on the Results of the MERIT Project. Wallingford (United Kingdom).

Bromley, J., Jackson, N. A., Clymer, O. J., Giacomello, A. M., and Jensen, F.V. 2005. The Use of Hugin® to Develop Bayesian Networks as an Aid to Integrated Water Resource Planning. Environmental Modelling & Software 20 (2), 231–42. doi: 10.1016/j.envsoft.2003.12.021.

Cain, J. 2001. Planning Improvements in Natural Resources Management: Guidelines for Using Bayesian Networks to Support the Planning and Management of Development Programmes in the Water Sector and beyond. Centre for Ecology and Hydrology. Oxford.

Cairncross, S., Hunt, C., Boisson, S., Bostoen, K., Curtis, V., Fung, I.C., and Schmidt, W.P. 2010. Water, sanitation and hygiene for the prevention of diarrhoea. International Journal of Epidemiology 39, i193–i205. doi: 10.1093/ije/dyq035.

Carriger, J. F., Barron, M. G., and Newman, M. C. 2016. Bayesian Networks Improve Causal Environmental Assessments for Evidence-Based Policy. Environmental Science & Technology 50, 13195−13205. doi: 10.1021/acs.est.6b03220.

Castelletti, A., and Soncini-Sessa, R.. 2007. Bayesian Networks and Participatory Modelling in Water Resource Management. Environmental Modelling & Software 22 (8), 1075 – 1088. doi:10.1016/j.envsoft.2006.06.003.

Collste, D., Pedercini, M., and Cornell, S.E. 2017. Policy Coherence to Achieve the SDGs: Using Integrated Simulation Models to Assess Effective Policies. Sustainable Science 12, 921–931. doi: 10.1007/s11625-017-0457-x.

Cronk, R., and Bartam, R. 2017. Factors Influencing Water System Functionality in Nigeria and Tanzania: A Regression and Bayesian Network Analysis. Environment Science & Technology 51, 11336−11345. doi: 10.1021/acs.est.7b03287.

Dondeynaz, C., López Puga, J., and Carmona Moreno, C. 2013. Bayesian Networks Modelling in Support to Cross-Cutting Analysis of Water Supply and Sanitation in Developing Countries Water Supply and Sanitation in Developing Countries. Hydrology and Earth System Sciences 17 (9), 3397–3419. doi: 10.5194/hess-17-3397-2013.

Esrey, S.A., Potash, J.B., Roberts, L., and Shiff, C. 1991. Effects of Improved Water Supply and Sanitation on Ascariasis, Diarrhoea, Dracunculiasis, Hookworm Infection, Schistosomiasis, and Trachoma. Bulletin of the World Health Organization 69, 609–621.

Feachem, R.G. 1984. Interventions for the Control of Diarrhoeal Diseases among Young Children: Promotion of Personal and Domestic Hygiene. Bulletin of the World Health Organization 62, 467–476.

Giné-Garriga, R., Pérez-Foguet, A., Molina, J. L., Bromley, J., and Sullivan, C. A. 2009. Application of Bayesian Networks to Assess Water Poverty Index. In proceedings: International Conference on Sustainability Measurement and Modelling. Barcelona.

Giné-Garriga, R., Requejo, D., Molina, J. L., and Pérez-Foguet, A. 2018. A Novel Planning Approach for the Water, Sanitation and Hygiene (WaSH) Sector: The Use of Object-oriented Bayesian Networks. Environmental Modelling & Software 103, 1−15. doi: 10.1016/j.envsoft.2018.01.021.

Henriksen, H. J., and Barlebo, H.C. 2008. Reflections on the Use of Bayesian Belief Networks for Adaptive Management. Journal of Environmental Management 88 (4), 1025–36. doi:10.1016/j.jenvman.2007.05.009.

IEAG (Independent Expert Advisory Group on a Data Revolution for Sustainable Development). 2014. A World That Counts: Mobilising The Data Revolution for Sustainable Development. New York and Geneva.

JMP (Joint Monitoring Programme for Water Supply, Sanitation and Hygiene). 2017. WaSH in the Agenda 2030: New global indicators for drinking water, sanitation and hygiene. New York and Geneva.

JMP (Joint Monitoring Programme for Water Supply, Sanitation and Hygiene). 2015. Progress on Sanitation and Drinking Water: 2015 Update and MDG Assessment. New York and Geneva.

Liu, H., Zhou, S., Lam, W. and Guan, J. 2017. A New Hybrid Method for Learning Bayesian Networks: Separation and Reunion. Knowledge-Based Systems 121, 185–97. doi.org/10.1016/j.knosys.2017.01.029.

Madsen, A. L., Jensen, F., Salmerón, A., Langseth, H., and Nielsen, T. D. 2016. A Parallel Algorithm for Bayesian Network Structure Learning from Large Data Sets. Knowledge-Based Systems 117, 46–55. doi.org/10.1016/j.knosys.2016.07.031).

Mohajerani, H., Kholghi, M., Mosaedi, A., Farmani, R., Sadoddin, S., and Casper, M. 2017. Application of Bayesian Decision Networks for Groundwater Resources Management Under the Conditions of High Uncertainty and Data Scarcity. Water Resources Management 31 (6), 1859–79. doi: 10.1007/s11269-017-1616-7.

Molina, J. L., Pulido-Velázquez, D., García Aróstegui, J. L., and Pulido-Velázquez, M. 2013. Dynamic Bayesian Networks as a Decision Support Tool for Assessing Climate Change Impacts on Highly Stressed Groundwater Systems. Journal of Hydrology 479, 113–129. doi: 10.1016/j.jhydrol.2012.11.038.

Molina, J. L., García Aróstegui, J. L., Benavente, J., Varela, C., de la Hera, A., and López Geta, J. A. 2009. Aquifers Overexploitation in SE Spain: A Proposal for the Integrated Analysis of Water Management. Water Resources Management 23 (13), 2737–60. doi: 10.1007/s11269-009-9406-5.

Nilsson, M., Griggs, D., Visbeck, M. 2016. Policy: Map the Interactions Between Sustainable Development Goals. Nature 534, 320-322. doi: 10.1038/534320a.

OECD (Organisation for Economic Co-operation and Development). 2008. Handbook on constructing composite indicators. Methodology and user guide. Paris.

Pérez-Foguet, A., Requejo-Castro, D., Giné-Garriga, R., Martínez, G., Rodríguez, A. 2017. Construction and Application of Bayesian Networks to Support Decision-making in the Water, Sanitation and Hygiene Sector: A Case Study of SIASAR

Initiative in Central America. In Proceedings: International Sustainable Development Research Society (ISDRS) Conference. Bogotá.

Scutari, M. 2010. Learning Bayesian Networks with the Bnlearn R Package. Journal of Statistical Software 35, 1–22. doi: 10.18637/jss.v035.i03.

Scutari, M., and Ness, R. 2018. Package: Bayesian Network Structure Learning, Parameter Learning and Inference. http://www.bnlearn.com (accessed 20.02.2018).

SDSN (Sustainable Development Solutions Network). 2016. SDG Index & Dashboards: A Global Report. http://unsdsn.org (accessed 04.12.2017).

SDSN (Sustainable Development Solutions Network). 2015. Indicators and a Monitoring Framework for Sustainable Development Goals: Launching a Data Revolution for the SDGs. (accessed 29.11.2017).

UN-DESA (United Nations Department of Economic & Social Affairs). 2015. Towards Integration at Last? The Sustainable Development Goals as a Network of Targets. ST/ESA/2015/DWP/141. New York.

UNDP (United Nations Development Programme). 2018. Human Development Data (1990-2015). http://hdr.undp.org/en/data (accessed 15.01.2018).

UNDP (United Nations Development Programme). 2016. Human Development Report 2016: Human Development for Everyone. New York.

UN-Habitat (United Nations Human Settlements Programme). 2015. Slum Almanac 2015-2016: Tracking Improvement in the Lives of Slum Dwellers. Nairobi.

United Nations. 2017. World Economic Situation and Prospects 2017. New York.

United Nations General Assembly. 2015. Transforming Our World: the 2030 Agenda for Sustainable Development. Resolution A/RES/70/1. New York.

UNSD (United Nations Statistics Division). 2018. Sustainable Development Goal indicators website. https://unstats.un.org/sdgs (accessed 15.12.2017).

UN-Water. 2016a. Monitoring Water and Sanitation in the 2030 Agenda for Sustainable Development: An introduction. Geneva.

UN-Water. 2016b. Water and Sanitation Interlinkages Across the 2030 Agenda for Sustainable Development. Geneva.