



Bayesian networks modelling in support to cross-cutting analysis of water supply and sanitation in developing countries

C. Dondeynaz^{1,2}, J. López Puga³, and C. Carmona Moreno¹

¹Institute for Environment and Sustainability, Joint Research Centre, Ispra, Italy

²School of Engineering, Liverpool University, Liverpool, UK

³Departamento de Psicología, Facultad de Ciencias de la Salud, Universidad Católica San Antonio de Murcia, Murcia, Spain

Correspondence to: C. Dondeynaz (celine.dondeynaz@ext.jrc.ec.europa.eu)

Received: 1 February 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 28 February 2013

Revised: 5 July 2013 – Accepted: 17 July 2013 – Published: 5 September 2013

Abstract. Despite the efforts made towards the Millennium Development Goals targets during the last decade, improved access to water supply or basic sanitation still remains unavailable for millions of people across the world. This paper proposes a set of models that use 25 key variables and country profiles from the WatSan4Dev data set involving water supply and sanitation (Dondeynaz et al., 2012). This paper suggests the use of Bayesian network modelling methods because they are more easily adapted to deal with non-normal distributions, and integrate a qualitative approach for data analysis. They also offer the advantage of integrating preliminary knowledge into the probabilistic models. The statistical performance of the proposed models ranges between 20 and 5 % error rates, which are very satisfactory taking into account the strong heterogeneity of variables. Probabilistic scenarios run from the models allow an assessment of the relationships between human development, external support, governance aspects, economic activities and water supply and sanitation (WSS) access.

According to models proposed in this paper, gaining a strong poverty reduction will require the WSS access to reach 75–76 % through: (1) the management of ongoing urbanisation processes to avoid slums development; and (2) the improvement of health care, for instance for children.

Improving governance, such as institutional efficiency, capacities to make and apply rules, or control of corruption is positively associated with WSS sustainable development. The first condition for an increment of the HDP (human development and poverty) remains of course an improvement of the economic conditions with higher household incomes.

Moreover, a significant country commitment to the environment, associated with civil society freedom of expression

constitutes a favourable setting for sustainable WSS services delivery. Intensive agriculture using irrigation practises also appears as a mean for sustainable WSS thanks to multi-uses and complementarities. With a WSS sector organised at national level, irrigation practices can support the structuring and efficiency of the agriculture sector. It may then induce rural development in areas where WSS access often is set back compared to urban areas¹. External financial support, called Official Development Assistance (ODA CI), plays a role in WSS improvement but comes last in the sensitivity analyses of models.

An overall 47 % of the Official Development Assistance goes first to poor countries, and is associated to governance aspects: (1) political stability and (2) country commitment to the environment and civil society degree of freedom. These governance aspects constitute a good framework for aid implementation in recipient countries.

Modelling is run with the five groups of countries as defined in Dondeynaz et al. (2012). Models for profile 4 (essential external support) and profile 5 (primary material consumption) are specifically detailed and analysed in this paper. For countries in profile 4, fighting against water scarcity and progressing desertification should be the priority. However, for countries in profile 5, efforts should first concentrate on consolidation of political stability while supporting diversification of the economic activities. Nevertheless, for both profiles, reduction of poverty should remain the first priority as previously indicated.

¹JMP statistics, 2004 <http://www.wssinfo.org/data-estimates/table/>, last access: 22 July 2013.

1 Introduction

On 6 March 2012², the UN announced that the millennium target for safe access to drinking water was reached, while sanitation target was still out of reach. Only 63 % of the World has improved sanitation access, and this figure is projected to increase only to 67 % by 2015, far from the 75 % requested by the Millennium Development Goals (MDGs). At the same time, the UN estimates that by 2015, 92 % of the global population will have improved access to drinking water. Behind these global figures, disparities remain across regions and urban/rural areas: “Only 61 % of the people in sub-Saharan Africa have access to improved water supply sources compared with 90 % or more in Latin America and the Caribbean, Northern Africa, and large parts of Asia” (WHO, 2012). Access to basic sanitation still represents a challenge for the coming years with “1.1 billion people who still practice open defecation. The vast majority (949 million) lives in rural areas. This affects even regions with high levels of improved water access” (WHO, 2012).

Upstream to these targets towards water supply and sanitation (WSS), an integrated water resources management (IWRM) approach was adopted by the international community in 1992 at Dublin’s conference. This considers the whole water cycle, from water withdrawal to restitution to the environment, including sanitation, the involvement of users and the preservation of the natural resources. IWRM promotes an inclusive water resources management. This approach has shifted efforts from infrastructure development and operational maintenance to a wider management including, for example, all users depending on the same water source. Sanitation, often neglected, is to be handled jointly with water supply because of its negative impacts on water quality at the end of the cycle.

The IWRM echoes the complexity of the water sector. Therefore, tools are needed to better understand and make mechanisms clearer, in particular to support decision making. In line with these international concerns and approach, this work aims at analysing and modelling variables involved in the access to WSS at country level. A better understanding of which and how variables impact WSS will help to orient efforts and set priorities for intervention. To do so, the analyses performed previously on the WatSan4Dev database (Dondeynaz et al., 2012) are used as a basis for the modelling phase proposed in this paper. WatSan4Dev was developed by the Joint Research Centre (JRC) of the European Commission (EC) and contains 42 indicators (called variables) representing environmental, socio-economic, governance and financial aid flows data in developing countries. As explained in Dondeynaz et al. (2012), a subset of 25 variables is used for analysing WSS behaviour because of the correlations observed through multivariate analyses of the WatSan4Dev

database. The description of the data and multivariate analyses performed both on variables and observations (countries) are detailed in Dondeynaz et al. (2012). In line with this previous analytical work, this paper proposes various probabilistic models based on Bayesian network methods. Scenarios are tested to describe relationships and map processes behind WSS behaviours and country development. Five country profiles are modelled, and scenarios run.

In Sect. 2, this WatSan4Dev subset is briefly described. The modelling methods, namely Bayesian networks are described in Sect. 3. The description of models and simulations are presented in Sect. 4.

2 The data: WatSan4Dev database

The WatSan4Dev database is used for this modelling phase. The data set consists of indicators characterising socio-economic and environmental status at national scale. Further details and description of WatSan4Dev variables, applied data pre-processing and coherency verification are fully described in Dondeynaz et al. (2012). The next two sections (Sects. 2.1 and 2.2) briefly summarise the variables and the methodologies used to build the WatSan4Dev database and the 25 variables subset used in this work.

2.1 The variables

The WatSan4Dev indicators are collected from international providers like the World Bank, FAO, universities such as Harvard, Colombia and Yale or recognise non-governmental organisations such as Transparency International. The national scale is chosen, as most of the data are mainly supplied at this level. The data set for 2004 is used in this work because this was the last release of the Joint Monitoring Programme (JMP) report on WSS access levels at the beginning of this research. Building on the several country profiles defined during the analytical phase of the WatSan4Dev subset, 25 variables are thematically clustered under five thematic areas: water resources (WR), human activity pressure on water resources (AP), country environmental concern (CEC), human development and poverty (HDP) and Official Development Assistance (ODA CI). The reduction of the number of variables to 25 results from preliminary multivariate analyses: collinearity between variables or no significant correlations with any other variables constitute the exclusion criteria from the variable subset. Figure 1 outlines the organisation and main correlations between selected variables.

The Water Poverty Index (WPI) and the Human Development Index (HDI) among other well-known composite indicators are used as key references to check the coherency and validity of the data used in this paper. Because of their composite and synthetic nature, these composite indicators are neither adequate for establishing the origins, nor the

²http://www.who.int/mediacentre/news/releases/2012/drinking_water_20120306/en/, last access: 22 July 2013.

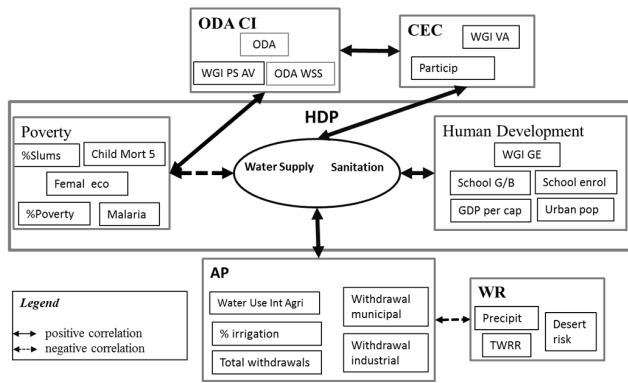


Fig. 1. Organisation and main correlations between selected variables.

consequences of the variations of simple variables related to WSS, which are the objective of this paper.

Throughout this paper, Worldwide Governance Indicators of governance effectiveness (WGI GE), which also represents control of corruption (WGI CofC), rule of law (WGI RofL), regulatory quality (WGI RQ), and environmental governance (Env gov), are called advanced governance. The last two governance indicators (WGI PS AV; WGI VA) are called basic governance because they ensure political stability; moreover, accountability towards citizens appears as an important basis for the development of other measures of governance.

The observations include 101 developing countries (our target group), therefore developed countries but also small states and islands are excluded.

2.2 Data pre-processing and coherency

Missing data are processed using multivariate imputation³ and normalised using standard normalization methods, namely square roots, logarithm and linear regression (Dondeynaz et al., 2012). The WatSan4Dev data set should be considered for qualitative analysis and not for quantitative interpretation purposes because of the strong heterogeneity of the sources and the data collection methods used by the official data providers.

Dondeynaz et al. (2012) uses the principal component analysis (PCA) and factor analysis (FA) for analysing relationships among and clustering of variables (Fig. 1). In this paper, the Bayesian models are built based on these previous multivariate analyses and conclusions.

³Software used: Amelia II by Honaker, King and Blackwell (Honaker et al., 2011).

3 Methodology

3.1 Description of Bayesian networks

Bayesian networks are statistical tools that originated in the field of artificial intelligence as models to manage uncertainty. A Bayesian network (also known as belief network, causal network, influence diagram or probabilistic expert system) is a statistical graphical model pertaining to *highly structured stochastic systems* (Cowell et al., 1999) conceived to represent probabilistic conditional relationships between variables. Following Aguilera et al. (2011), a Bayesian network could be formally specified through two different dimensions. Firstly, a Bayesian network is a graphical representation of a problem containing a set of related variables and, on the other hand, the network is defined by a set of probability distributions. The following section technically defines each of these dimensions and highlights the advantages of using such methods.

3.1.1 Graphical representation

A Bayesian network is a graph which represents a problem. In this context of modelling, the graph, called a directed acyclic graph (DAG), is defined as a pair $G = (V, E)$ where V refers to a finite set of vertexes, nodes or variables and E is a subset of ordered pairs in the Cartesian product $(V \times V)$ called links or edges.

The term *directed* refers to a directionality concept, implying that the edges or links between variables are directed. For instance, if $(A, B) \in E$ but $(B, A) \notin E$, that means there is a directed link between A and B and it is represented as $A \rightarrow B$.

The term *acyclic* refers to the fact that loops are forbidden in the network. A directed connection between two variables in a Bayesian network is interpreted as statistical or relevant dependency. Referring to the example above, it is said that B depends on A . Talking in causal terms, A causes B or, the other way around, B is the effect of A .

The *principle of conditional independence* is the concept which is used to spread evidence within the model. Let us consider three variables or a set of variables, x , y and z ; x and y are (conditionally) independent given z if $p(x|z) = p(x|yz)$. In other words, two variables x , y are independent given a third variable z if and only if $p(xy|z) = p(x|z) \times p(y|z)$. This principle is important because it makes probability updates possible in the three basic types of sub-structures (converging, diverging and serial) of a Bayesian network.

3.1.2 Qualitative dimension

There are three quantitative basic elements characterising a Bayesian network:

1. The Bayesian networks are based on a vision considering probability as a subjective issue (so-called *bayesian*) (Cox, 1946; Cowell et al., 1999; Dixon, 1964; Heckerman, 1995). It considers that probability is just a *degree of belief* with regards to the occurrence of events. The probability is understood as a subjective evaluation based on our state of knowledge about nature or reality (Dixon, 1964). The probability is a measure to quantify uncertainty and Bayesian networks use it in subjective terms.
2. A Bayesian network is perimetrically determined by a set of conditional probability functions. Generally, those parameters are specified in a conditional probability table (CPT). The CPT contains a set of probability values corresponding to all possible combinations of node states and parent node states. These probabilities or parameters are the bases on which the Bayesian theorem is founded.
3. The Bayesian theorem is a useful rule derived from the concept of conditional probability applied to the intersection of events. Derived from Thomas Bayes' work (1763), it is a valuable tool when willing to update the knowledge about an event based on evidence related to another linked event. In its simplest form, Bayesian theorem can be formulated as follows:

$$p(B|A) = \frac{p(A|B) \times p(B)}{p(A)}. \quad (1)$$

The *principle of conditional independence* allows factorising probabilistically the whole model, meaning, in practical terms, the assessment of probabilities given the evidence. Given a Bayesian network, the factorisation processes are carried out according to the following Eq. (1):

$$p(x) = \prod_{v \in V} p(x_v | x_{pa(v)}), \quad (2)$$

where x_v refers to the variables in the model and $x_{pa(v)}$ stands for x_v 's parents or ancestors (i.e., Aguilera et al., 2011; Hoeting et al., 1999; Nadkarny and Shenoy, 2001).

3.1.3 Advantages and limits of Bayesian networks

The most important advantage of the use of Bayesian network (BN) modelling in environmental science is that they are able to differentiate and efficiently manage the quantitative and qualitative dimensions of a problem (i.e., Edwards, 1998; Heckerman, 1995). Bayesian networks can also optimally manage missing data (Nadkarny and Shenoy, 2004; Jansen et al., 2003) and this is particularly relevant when analysing data from developing countries because of the high proportion of missing data. Additionally, Bayesian networks allow: (1) the combination of prior knowledge with empirical data to develop models (Nadkarny and Shenoy, 2004);

and (2) taking into account the interaction effects among variables. It is also possible to carry out *local computations* to avoid the re-estimation of the whole joint distribution of a model. This makes the inference task a relatively low computing-demanding process (Pearl, 2001; Xiang, 2002). It is also to be noted that non-linear relationships are modelled efficiently (Lee et al., 2005).

However, Aguilera et al. (2011) noted that there are several drawbacks associated to the use of Bayesian networks in the field of environmental sciences: (1) the building process might be challenging when the ratio between cases and variables is low; (2) problems to model time series and fuzzy-based models have also been noted.

Concerning the specific application of Bayesian networks (BN) to the water sector, these have extensively been used from ecological applications to their integration in decision support systems at local/river basin level. Bayesian networks are used to model strains from which the water ecosystems suffer, such as climate change (Varis and Kuikka, 1997) and farming/irrigation practices (Batchelor and Cain, 1999; Borsuk et al., 2004; Martín de Santa Olalla et al., 2005), for instance. On another side, Baran and Jantunen (2004), Bromley et al. (2005), Castelletti and Soncini-Sessa (2007b), Zorilla and al. (2009) and Carmona et al. (2011) highlight the use of Bayesian networks as tools to facilitate stakeholders' participation and planning. BN modelling has been used alone but also combined with other statistical techniques to support the management and decision-making processes at river-basin level. This was the case, for instance, in Castelletti and Soncini-Sessa (2007a), Molina et al. (2010), and Sušnik et al. (2012). These works combined BN to model socio-economic components within a hydrological model in a decision-making framework at local level. Giné Garriga et al. (2009) uses BN at river basin level to model several dimensions (physical availability of water resources, extend access to water supply (WS), people's ability and capacity for sustaining access, various uses of water resources and environmental factors that impact water resources) following the framework of the Water Poverty Index (Sullivan et al., 2003).

The work proposed in this paper follows this applied multi-dimensional modelling approach and takes advantage of the flexibility of BN methods mainly at the level of missing data management and non-linear behaviour modelling. It aims at scaling it up to national level to support regional or worldwide cross-country analysis as JMP monitors WSS at this scale.

3.2 Bayesian networks: processing steps

This section describes the necessary four processing steps to model our subset with the Bayesian networks method:

1. Variables discretisation: the variables are discretised using the k mean method (McQueen, 1967). Although the discretisation process produces a loss of information (Aguilera et al., 2011), it is performed to make

the problem tractable and to homogenise the data set. The objective is to create three categories or levels (low, medium, high) matching with the qualitative nature of the selected variables. Table A1 indicates the limits of each class for each variable used in this paper. Water supply, sanitation and WGI (Worldwide Governance Indicators) variables show only two levels (high and low), because of the shape of their distribution: this increases the classification error rate.

2. Building thematic composite indicators: composite indicators are built for each thematic pillar namely, HDP, CEC, AP, WR and ODA CI (Fig. 1). The first component of the PCA, which includes the pillar subvariables, is categorised in three levels, using the k means clustering method as it was done for the variables.
3. DAG design: two different modelling strategies are used to develop the general and specific models. Firstly, global or general models are built using the PCA/FA results and using expert knowledge (literature and field experience) to specify link directions. This expert and statistically guided method to develop graphs allows testing different structures in order to assess the goodness of fit for each of them. The DAGs for specific models are created using the generic structure of the naïve Bayes classifier or simple Bayesian classifier. Those structures are considered appropriate to model relationships between the composite indicator (CI) and its subvariables. This method is used as a non-parametric alternative to regression models trying to generate more parsimonious (in terms of the number of parameters to estimate) and usable models from the decision-making point of view.
4. Beliefs computation: the Bayesian theorem is applied and probabilities are computed for each variable and each level (HIGH, MID, LOW) according to the input data set (101 developing countries) or country observations from a specific profile.

Different probabilistic scenarios are then run to describe the interrelations between variables or in the frame of a specific country profile.

4 Modelling WatSan4Dev subset

The data set modelling is done in two main steps: (1) the integration of thematic composite indicators in a global model; and (2) the creation of a specific model for each pillar. Separated models are built for water supply and sanitation. The global models offer a synthesis of the mechanisms involved in the WSS level and its articulation within a national context. The modelling by pillars scales down the analyses.

4.1 General models: structure, statistical validation and scenarios

Statistical validation is carried out⁴ and measured by the error rate, logarithmic loss (LL), quadratic loss (QL) and spherical payoff (SP) coefficients (Pearl, 1978). The error rate represents the percentage of cases incorrectly classified; the lower the value, the better the predictive validity of the model. The logarithmic loss (LL) ranges between zero and infinity where zero is the best fit. The quadratic loss (QL) ranges between zero and two (zero is the best fit). The spherical payoff (SP) varies between zero and one (one is the best fit).

A sensitivity analysis is carried out to evaluate the relative impact of each variable in the model. Entropy reduction (ER) (or mutual information) refers to the expected reduction in the query variable due to a finding in any other variable of the model (Pearl, 1991). Its value varies between zero (meaning complete independence between the query and the instantiated variable) and the entropy value of the query without any evidence about the model. The variance of node belief (BV) and the root mean square (RMS) change of belief are also computed (Neapolitan, 1990). Both statistics range from zero to one (the closer the value to zero, the stronger the independence between the query and instantiation variable).

The global model is based on composite indicators compiled from the previously defined pillars of variables: human development and poverty (HDP), water resources (WR), human activity pressure on WR (AP), Official Development Assistance (ODA CI) and country environmental concern (CEC). It is to be noted that WR variables as well as ODA CI are poorly correlated with the WSS variables. Therefore, WR variables are set with an indirect link via the AP pillar, with WS and S (sanitation) as dependent variables (Fig. 2). The reason resides in the existing correlation between the amount of water resources available (TWRR) and water use intensity in agriculture (Water Use Int Agri).

Financial aid flow (ODA CI integrating variables ODA, ODA WSS and WGI PS AV) is included in the models following clues given by previous PCA analysis: (i) ODA shows higher positive correlations with poverty variables than with development variables and shows a negative correlation with WS (−0.450) and S (−0.412); (ii) ODA WSS is correlated with the political stability and absence of violence index (WGI PSAV, correlation 0.363). The non-significant correlation with WSS may be due to a scale bias; ODA WSS disbursements were 8% of the total ODA between 2004–2005 (OECD, 2007); and (iii) WGI PSAV is positively correlated with other governance variables such as the indicator measuring voice and accountability within a country (WGI VA – correlation 0.524). Therefore, ODA CI is linked with HDP, CEC and WSS (Fig. 2).

⁴To build and test the models, the Netica Application (Norsys Software Corp.) version 4.16 is used (www.norsys.com).

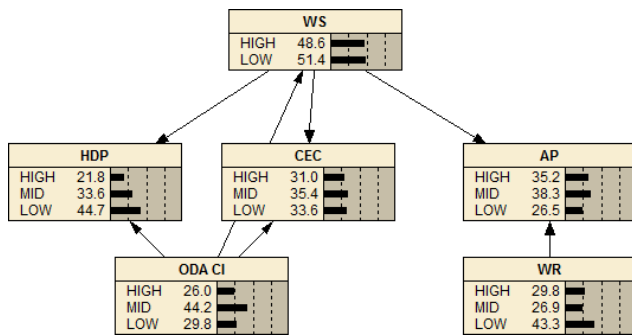


Fig. 2. WS general model, DAG shows the relationships between the WS and pillars and initial probabilities computed.

After developing the Bayesian networks, several simulations are run to measure the probabilistic variations of key variables. The interpretation of these results constitutes a hypothesis on the potential underlying mechanisms which are introduced in the following paragraphs. Hypotheses are mainly formulated from the simulation but also according to relevant scientific literature and international field related experiences.

4.1.1 Water supply (WS) general model

The general model correctly classifies 86 % of the Wat-San4Dev subset (Table 1a). Figure 2 presents the direct acyclic graph (DAG) of this model.

The sensitivity analysis provides the influence (probability change) of the pillars on the WS variable (Table B1a). The variable influencing the most WS access level is the development of the country (HDP). Poverty reduction, as a component of the HDP pillar, is a key result of the increased access to an improved water supply. WR is excluded from the sensitivity analysis because of its indirect link with WS. Table 2 presents the initial probabilities (IP) of each variable of the model for a specific level. For each simulation, a specific variable is set at 100 % probability for one category. The model then computes the new probability distributions of the other variables according to this change. The results of simulations are reported in Table 3. The positive increments of variables (recomputed probability minus initial probability) are represented in bold and the negative increments in italic. Simulations are ordered according to WS variation (HDP implies the highest change, ODA-CI the lowest).

Simulation 1 (HDP set to 100 %) directly regards the estimation of improvement on WS needed to support poverty reduction efforts. Indeed, in line with the Millennium Development Goals initiative, countries should halve the poverty worldwide through the improvement of WS access (halving the population without access to improved water supply and basic sanitation – target 7C). Simulation 2 (CEC set to 100 %) estimates the influence of the WS access and WS management on the environmental concern indicator.

Table 1. Performance and error analysis for WS, S, HDP, CEC and AP models.

Model	Error rate	LL*	QL*	SP*
(a) WS general model	11.88 %	0.2814	0.1673	0.9083
(b) S general model	9.901 %	0.2921	0.1622	0.9143
(c) HDP model	4.95 %	0.168	0.09021	0.9517
(d) CEC model	19.80 %	0.4931	0.2967	0.8353
(e) AP model	11.88 %	0.2688	0.157	0.9129

* LL = logarithmic loss, and QL = quadratic loss where best value is 0, SP = spherical payoff where best value is 1.

Simulation 3 examines the type and level of pressure of human activity (water demand according to sectors) on resources in the light of WSS development. Simulation 4 (ODA CI set to 100 %) observes Official Development Assistance flow (ODA CI) to measure the impact of such financial mobilization.

Simulation WS1 (HDP set to 100 %): pushing poverty reduction to foster development

In this simulation, increasing HDP calls for major efforts on poverty reduction. In that case, high access to WS needs to reach almost 85 %, thus a very important improvement of WS access when compared to IP. Various mechanisms can be suggested to explain it:

- Safe access to a source of drinking water close to customers is a primary basic service that enables health, social and ultimately economic development of a country. Indeed, the impact on health is well known (i.e. on waterborne disease, or child mortality), but its positive impact goes far beyond by participating in the creation of favourable context for economic development (Sect. 4.2.1).
- Governments engage in WSS programs as an important component of their poverty reduction strategy (to increase their HDP). Associated to the state efforts, the international donors finance programs fighting against poverty, and mobilise resources to WSS services. In 2004, DAC donors committed to 4.4 billion USD (bilateral and multilateral ODA) and disbursed around 2.4 billion USD in Official Development Assistance for WSS improvement (OECD-DAC, 2007).

The current model estimates the contribution of WS to HDP but the relationships between HDP and WS can be multiple. Indirectly, in a context of economic development, households tend to dedicate financial resources to improve their WS access when this is possible.

A better education is a long-term element that contributes to the increase of income (Houthakker, 1959) and, therefore, living conditions including WS become a priority. Note that in the case of unequal distribution of wealth, these positive

Table 2. General model – water supply simulations. In bold, positive variations of probabilities.

Variables	Levels	Initial Probabilities	Simulation 1 HDP	Simulation 2 CEC	Simulation 3 AP	Simulation 4 ODA CI
HDP	HIGH	21.8 %	100	6.9	5.5	0.7
ODA CI	LOW	29.8 %	1	2.1	2.3	100
CEC	HIGH	31.0 %	10.5	100	5.1	3.1
AP	HIGH	34.4 %	8.8	5.4	100	2.6
WR	HIGH	29.8 %	(=)	(=)	(=)	(=)
WS	HIGH	48.6 %	84.5 %	70.3 %	66.3 %	59.4 %

Table 3. General model – sanitation simulations. In bold, the positive variations of probabilities.

Variables	Levels	Initial Probabilities	Simulation 1 HDP	Simulation 2 AP	Simulation 3 CEC	Simulation 4 ODA CI
HDP	HIGH	21.6 %	100	6.3	1.4	0.7
ODA CI	LOW	29.8 %	1	1.8	4.9	100
CEC	HIGH	31.1 %	1.9	2.1	100	5.1
AP	HIGH	34.2 %	9.9	100	2.3	2.1
WR	HIGH	29.8 %	(=)	(=)	(=)	(=)
S	HIGH	45.8 %	80.3 %	66.7 %	53.9 %	53.1 %

impacts are a priori undermined. The indicator of inequality could not be integrated into these models. Indeed, the Gini index that measures the inequality of income at country level showed too many missing values when considering developing countries. The advanced governance included in HDP (control of corruption, government efficiency, policy making capacities) is definitely favourable to WSS sustainability: “The strength and quality of government institutions and a strong policy environment are just as important to achieving success in the environment sector as they are to other sectors” (Burtaine and Parks, 2012).

Simulation WS2 (CEC set to 100 %): WS management fostering environmental concern and civil society role

In this simulation, the increment of environmental concern (CEC) implies an increment of WS up to 70.3 %. CEC expresses the country commitment towards major environmental issues (Particip). The degree of civil society freedom of expression and the level of accountability of government (WGI VA) are also included. The bigger the increment of CEC, the more the country is committed towards its environment and citizens. According to the model, the improvement of safe water facilities reinforces the commitment associated with a strong civil society. A management of the water services made in a participatory way, with for instance water user associations and/or consultation processes⁵, directly reinforces CEC. Indeed, water services management opens room for discussion and participation of citizens, at least at

⁵Principle no. 2 of IWRM, formalised in DUBLIN Statement 1992: <http://www.wmo.int/pages/prog/hwrp/documents/english/icwedece.html#principles> (last access: 28 August 2013).

local scale. The water sector can be a starting point, calling citizens to organise themselves to participate in decision-making and management processes beyond its warning role, i.e. pollution or WSS services interruption (Sect. 4.2.2).

Simulation WS3 (AP set to 100 %): WS related to the development of human activities

According to the model, the increment of activity pressure on WR (AP) entailed an increment of WS access up to 66 %. AP composite index should be interpreted as follow: the bigger the increment of the AP factor, the more the agriculture monopolises available water resources, often through irrigation practices. It is important to remember that the AP behaviour is related to countries where agriculture often represents an important economic activity. The municipal water supply network (called by FAO municipal demand) is still used to draw a small part the water in the areas considered. Agriculture is the main consumer of drawn water resources (consuming above 50 % of water resources) in 80 % of the selected countries. Low values indicate that a significant part, still not a main part, of water consumption goes for industrial activities, including raw material extraction (implying a specific water supply) and/or domestic/municipal uses (served by network supply).

This behaviour indicates that agriculture is positively associated with the water supply coverage creating a positive synergy between these variables: (1) in the case of irrigation agriculture, water supply infrastructures can be implemented following a multi-purpose logic, providing water to cover population, agricultural and often energy needs; but also, (2) a strong water supply sector ensures the service to

the population, the priority, but the sector (capacities, market organisation...) can be beneficial to agriculture and in particular irrigation plants. However, irrigation practices (above 20 % of irrigated areas) concerned exclusively South East Asia, Middle East and Egypt. Latin America and Africa are little concerned.

Simulation WS4 (ODA CI set to 100 %): the multiple relations between ODA CI, and HDP, WS CEC

The first group: ODA CI, HDP and WS

In this simulation, high donors' investments (ODA CI) to support partner countries indicate a low-medium HDP situation (78 % probability) and therefore, there is a 61 % probability that access to WS is poor. This is compliant to the purpose of Official Development Assistance: providing financial support to less advanced countries (low HDP).

In 2004, investments in social infrastructures and services (government and civil society support, education, and health issues) represented 37 % of the Official Development Assistance provided by all DAC donors and the EU (OECD statistics in 2004⁶). Humanitarian, emergency aid is excluded from ODA. The model shows that poor countries (with low HDP) received high Official Development Assistance with almost 45 % of probability. In a general way, the lower the HDP of the country is, the higher the ODA CI is. However, the model estimates that 21 % of the countries that benefit from relatively high ODA CI show high HDP. This may raise questions about the motivations behind ODA delivery. Over the last decades, there has been a rather polemical stemming from donor's visions and strategies of cooperation (Carbonnier, 2010). In addition, excessive dependence on external financial support can imply negative effects on the quality of governance like "corruption in government, bureaucratic quality, and the rule of law" (Knack, 1999): all elements included in the HDP variable. The ways in which external aid is delivered may complement the explanation of the ODA CI, HDP, and WS observed behaviour. For instance, donor fragmentation combined with the low capacities of a country to fund implementation can favour corruption practices.

The second group: ODA CI, WS and CEC

The model estimates that the increment of ODA CI (also including political stability, WGI PS AV indicator) induces an increment of high CEC (also including civil society voice, WGI VA indicator) from 31 to 34.3 %, but with the majority of countries (43.9 %) showing a medium CEC level.

Several hypotheses can be formulated to understand this link:

- Programs/projects supported by external aid include activities for citizen involvement in management of WSS services because the IWRM participatory approach pushes in that direction.
- Minimum political stability (WGI PS AV) and civil society voice (WGI VA) facilitate financial efficiency of the funds and sustainable infrastructure implementation. In fact, the experience in cooperation has shown that external aid has been conditioned to political stability, democratisation and more generally to governance (Santiso, 2001). Good context facilitates the mobilisation of funds while poor governance represents a risk.

These two hypotheses may coexist depending of the governance context and donors' strategy in a specific country. Donors may deliver aid despite risks in some countries because of their long-term or geo-political strategy. Necessity of supporting less advanced countries can also prevail over high aid effectiveness expectations.

4.1.2 Sanitation (S) general model

The S model follows the same structure as WS (Fig. 3). Model performance reaches up to 90 % of correct classification (Table 1b).

Table 3 includes computed probabilities of the different simulations made as for WS. Simulations are ordered by their level of impact.

As for water supply, sanitation is mainly sensitive to human development and poverty (HDP) (Table B1b). The development of sanitation services participates in the improvement of living conditions as a key element of the slums' formalisation, health improvement and education (see Sect. 4.2.1). Commitment towards the environment and good governance is improved in a context of high sanitation access. Financial external aid (ODA CI) is mainly (44.5 %) delivered in countries with low HDP and low S. The main difference stands in the relatively higher sensitivity to the human activity pressure variable (AP) and therefore to agriculture pressure.

Agriculture and sanitation can develop synergies like, for instance, the reuse of human excreta for soils fertilisation. This opportunity of valorisation of sanitation/wastewater products in agriculture is available both in rural and urban areas. This is easier in rural areas where sanitation products can be collected and transformed into fertiliser at affordable prices with composting toilets⁷ or by a solar drying process. Regarding urban areas, wastewater sludge, if treated correctly, can be reused as well for agriculture. However, it is difficult to estimate the extension of such practices because statistics regarding sanitation or wastewater products reuse

⁶OECD: <http://stats.oecd.org/qwids/>, last access: 15 February 2013.

⁷See composting toilet fiche: <http://www.sswm.info/category/implementation-tools/wastewater-treatment/hardware/site-storage-and-treatments/composting-t>, last access: 4 July 2013.

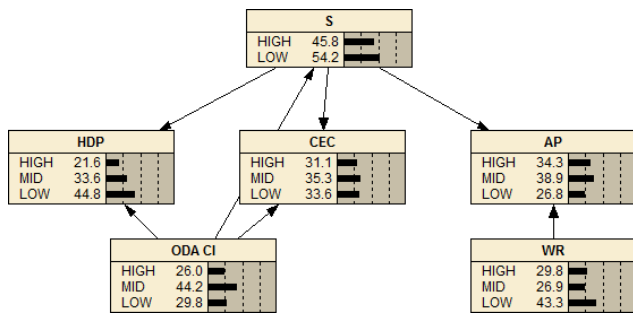


Fig. 3. S general model, DAG shows the relationships between the S and pillars.

in developing countries are seldom available. Raschid-Sally and Jayakody (2008) studied wastewater reuse in agriculture in 53 cities across the world.

4.2 Specific models for thematic pillars

General models provide the core idea about the relationships between the thematic pillars and WSS while specific models refine this analysis. In this section, the analysis of relationships is scaled down and submodels for each pillar are described.

The specific models’ DAGs are set with the composite indicator (HDP, CEC or ODA CI) as a parent node to its sub-variables nodes. As a consequence, the variation change on the subvariables is estimated according to the CI behaviour within the global model. This allows measuring the association between subvariables and the corresponding CI. Therefore, the link between CI and its subvariables is rather a matter of strength of relationship than a conditional link between two variables. An indirect link is created from WSS variables to subvariables with the CI as intermediate variable. As for global models, each following section summarises the sensitivity analysis results (described in Table B1), and details several simulations run.

Note that the order of variation can be different between the sensitivity analyses and the simulations tables. The sensitivity analyses report the cumulated variation across all categories (high, mid, low). The simulation table reports the probability variations for a specific category.

4.2.1 Human development and poverty model (HDP)

The statistical analyses establish a model with direct links between HDP, the variables in the HDP pillar, and the ODA CI pillar. The suggested model explains the variability up to 95% (Table 1c). Figure 4 shows the structure of the HDP model and initial probabilities calculated for all variables.

The three main variables influencing HDP are the income per capita, child mortality under 5 yr (fighting efficiently against mortality) and the proportion of slums (pushing the organisation of urbanisation) (Table B1c). These are

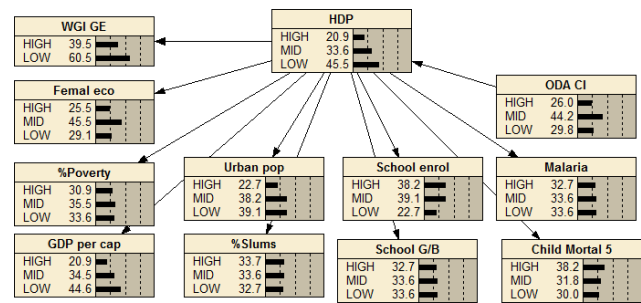


Fig. 4. Probabilities in % for the variables involved in the HDP pillar, discretised in high (HIGH), medium (MID) and low (LOW) levels.

the variables where authorities and/or donors should focus their efforts to foster development. Two other important aspects are the support of school enrolment and the control of malaria.

Table 4 sums up the probabilities computed according to several scenarios presented in this paper. Simulation 1 estimates the HDP level when significantly reducing child mortality under 5 yr. Simulation 2 estimates the improvement of HDP necessary to support the reduction of informal urbanisation (% slums). Simulation 3 observes how the improvement in human development is translated in terms of malaria control. Simulation 4 observes the benefits of good development on governance conditions. Simulation 5 examines the association between human development and education (from primary level to university). Finally, simulation 6 examines the role of gender through girls’ primary education (School G/B) and female economic participation (Femal eco) variables.

Simulation HDP1 (Child Mortal-5 % set to 100 %): WSS is essential for reducing child mortality

According to the model, if child mortality (Child Mortal-5) is reduced by increasing efforts on basic health care, HDP is expected to be high with 53.6% probability. Fighting child mortality requires better access to safe drinking water, together with basic health care and quality of nutrition. Improved sanitation considerably reduces the human/water contamination by dissemination of pathogenic material. With WSS access, the incidence of waterborne diseases such as diarrhoea is considerably reduced, and consequently the risk of mortality, in particular for young children. Pneumonia and diarrhoea are the two first causes of young child mortality in developing countries: diarrhoeal disease kills 1.5 million children every year (UNICEF WHO, 2009). Africa and Southeast Asia remain most affected by diarrhoea and malaria.

Table 4. Thematic model-HDP Simulations. In bold, the positive variations of probabilities.

Variables	Category	Initial Probabilities	Simulation 1 Child Mortal-5	Simulation 2 %slums	Simulation 3 Malaria	Simulation 4 WGI GE	Simulation 5 School enrol	Simulation 6 School G/B
School enrol	HIGH	37.6 %	30.7	30.6	29.7	7.4	100	17.5
School G/B	HIGH	32.4 %	16.4	16.1	15.2	12.6	15.1	100
Malaria	LOW	33.1 %	27.5	27.2	100	21.3	26.1	15.5
%slums	LOW	32.2 %	28.6	100	26.4	21.9	26.1	16
%Poverty	LOW	33 %	30.1	29.5	27.7	23	27.2	16.8
Child Mortal-5	LOW	29.5 %	100	26.2	24.5	20.4	23.8	14.9
Femal eco	LOW	28.8 %	15.1	15	14.4	11.7	14.8	8.6
Urban pop	HIGH	22.3 %	22.8	22	20	17.1	18.7	12.4
GDP per cap	HIGH	20.4 %	23.8	22.7	20.1	17.7	18.2	12.8
WGI GE	HIGH	39 %	27	26.5	25	100	24.7	12.4
ODA CI	LOW	29.8 %	1.1	1.1	1.2	0.9	1.4	0.7
HDP	HIGH	20.2 %	53.6 %	51.9 %	47.7 %	44.7 %	44.3 %	37.9 %

Simulation HDP2 (low% slums set to 100 %): the role of urbanisation development

The type of urbanisation plays a central role, as already mentioned (Dondeynaz et al., 2012). If slums (%slums) are reduced, that suggests major efforts to control and organise urbanisation processes; in that case, high HDP probability is estimated to rise to 52 %. Organising and implementing services in such suburbs/districts is an obvious priority. The availability of WSS services represents a good leverage for improving living conditions. The consequences of lack of sanitation are even more amplified in terms of health and water contamination because of the density of the population. The limitation of slums' extension from high to medium level is translated in the model by the reduction of probability that countries have a low level of HDP from 91.2 % to 38.3 %. This increment positively impacts on essential variables: school enrolment, household income (GDP per cap), child mortality rate and malaria prevalence.

The level of urbanisation (Urban pop) presents a negative correlation (-0.650) with urban slums (% slums) accordingly to the dynamics of the rural-urban population transition. World urbanisation prospects (UNPD, 2001) analyse urban and rural population growths, providing detailed information on this transition at regional level. Before the beginning of this transition, the majority of the population lived in rural areas with agriculture as the principal activity and child mortality rates were relatively high. For economic and demographic reasons, the population started moving from rural areas to urban centres. Therefore, at this time, the urban population represents proportionally a small percentage but its growth is significant. The population migrating to cities often resides in informal settlements around urban areas, therefore incrementing the slums. Progressively, as the proportion of urban population increases, urban development should be more and more organised and as a consequence, slum proportion is supposed to be reduced. This process ends up with the majority of the population living in formal urban areas

with an urban population growth mainly due to demographic and economic dynamics. The model proposed in this paper estimates that in 2004 almost 40 % of developing countries present a low urban population rate, while the level of population living in slums is high in 34 % of cases. Generally speaking, urbanisation represents an opportunity for development of WSS services as the concentration of the population provides the necessary critical mass for collective infrastructures and improves cost effectiveness. The organisation and development/rehabilitation planning of urban areas is crucial to limit the extension and the population living in slums.

Simulation HDP3 (low Malaria set to 100 %): WSS improvement limits malaria prevalence

The reduction of malaria prevalence (Malaria) means an increase of HDP up to 47.7 %. According to the CDC⁸, the majority of the countries included in the WatSan4Dev database is affected by malaria either on the whole or on part of their territory. Among other socio-economic aspects, WSS services availability also contributes to reduce malaria prevalence (HDP as connecting variable). The lack of water supply infrastructure leads to outdoors water storage that facilitates the development of mosquitoes and therefore dissemination of malaria (Lewis, 2011). Limiting inappropriate storage is essential, and avoided in the case of piped water and rainwater harvesting infrastructure.

Malaria prevalence can also be considered as a proxy of good and efficient sanitation. The existence of sanitation facilities reduces the development of mosquitoes by reducing favourable conditions. Improving sanitation, and by extension wastewater collection, reduces stagnant and potentially pathogenic water areas. Indeed, in the case of pipe infrastructure "when population growth outpaces the existing infrastructure, wastewater treatment systems are unable to cope with the influx, garbage; sanitation facilities cannot contain

⁸CDC malaria map: <http://cdc-malaria.ncsa.uiuc.edu/>, last access: 4 July 2013.

the increased refuse, and access to clean, treated drinking water may not be available. All of these conditions contribute to create the perfect environment for disease causing mosquitoes to breed” (Lewis, 2011).

Simulation HDP4 (WGI GE set to 100 %): advanced governance and HDP relationships

In the case of simulation of high advanced governance (WGI-GE), HDP increases to 44.5 %, more than doubling the initial probability. In fact, the capability of a nation to organise public services, to be committed towards implementation of suitable policies, and to fight corruption (as represented by the WGI-GE) is associated with good development of the country. This positive association can appear as a chicken-egg situation. Corruption practices in a context of weak institutional authorities are favoured in a poverty background. As the same time, country development is also boosted by good governance.

Applied to the WSS sector, the development and maintenance of WSS services can build capacities and create participatory management rules, at least at local level. Good governance supports a sustainable WSS delivery (Transparency International, 2008), while the interest and participation of users pushes for better governance.

Simulation HDP 5 and HDP 6: WSS role in education and gender issues (school enrolment, ratio girls to boys at school and female economic activity analysis)

In the HDP 5 simulation, school enrolment (School enrol) is fostered, and HDP is enhanced to reach up to 44.3 % probability. In the HDP 6 simulation, the improvement of the girls’ schooling rate (School G/B) shows a positive correlation with an increment of HDP which reaches up to 37.9 %. A rather good HDP (High HDP with 36.6 % probability) indicates a limited female economic participation (Femal eco) (Boserup, 1989). Beyond the HDP improvement, several hypotheses on the relationships with WSS can be depicted:

- Water availability in households is an important factor in the enrolment, attendance and drop-out rates of children. When water access is improved in communities, school enrolment rates increased in Tanzania, India and Bangladesh by up to 15 per cent (UNDP and UNICEF, 2006). Back-to-school campaigns in countries in transition are especially dependent on water supplies: “Teachers are hesitant to relocate to communities without a reliable and safe source of water” (UNDP and UNICEF, 2006).
- Girls in the developing world are generally in charge of the daily water fetching, setting time barriers to school attendance. Separate sanitation facilitates school attendance of girls even more, particularly after reaching puberty thanks to privacy, and security advantages.

Despite of these advantages, “in some countries in Africa and Asia as few as 10 per cent of schools have adequate and separate sanitation facilities, while student-to-latrine ratios can be as high as 150 : 1” (UNDP and UNICEF, 2006).

- In a context of poverty, women are massively active in the economy to sustain their families, which mean, translated in statistics, high female economic activity rate. With the increment of household income, this rate drops to a certain threshold, before rising up in a second phase because of a completely different dynamic (Boserup, 1989). Women in most developing countries handle both domestic work, including water fetching, and an economic activity: “Providing water sources frees up time for women and leads to both direct and indirect opportunities to engage in activities which provide an income” (UNDP and UNICEF, 2006). Therefore, the level of women’s activity remains more or less the same, but the benefit resides in the qualitative improvement of women’s daily lives.

4.2.2 Country Environmental Concern model (CEC)

The performance of the CEC model is higher than 80 % (Table 1d). Figure 5 shows the initial probabilities of variables in the CEC model following the same methodology as for HDP. The World Governance Voice and Accountability index (WGI VA) only presents two levels (high and low), following the structure of the variable: its range is $[-2.5, 2.5]$, where 0 can be considered as the limit between poor and good governance.

Table B1d provides the CEC sensitivity to subvariables in a decreasing order: participation in international environmental agreement (Particip), government accountability and citizen freedom of expression (WGI VA) and finally the ODA CI (including ODA, ODAWSS and WGI PS AV).

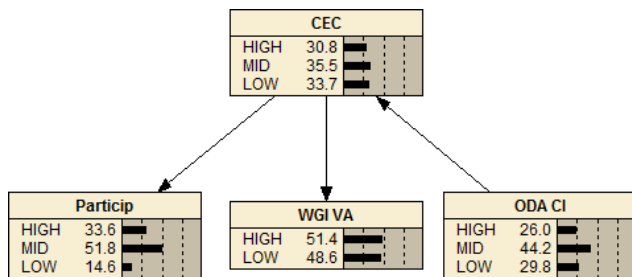
Simulation 1 analyses the country commitment towards international environmental concerns. Simulation 2 estimates the impact of citizens’ freedom of action and expression on environment issues and, more widely, democratic conditions and government accountability. Simulation 3 considers ODA CI relations with environmental and democratic commitment (Table 5).

Simulation CEC1 (Particip set to 100 %): commitment towards environment includes WSS management

In this simulation, high international environmental commitment of countries (Particip) induces an increment of CEC up to 70.1 %. Signing environmental agreements (Particip) formally engages countries towards global environment protection through the United Nations Framework Convention on Climate Change (UNFCCC), Vienna Convention on the Protection of the Ozone Layer, Convention on the Trade in Endangered Species (CITES), and the Basel Convention on

Table 5. Thematic model, CEC simulations. In bold, positive variations and in italic, negative variations of probabilities.

Variables	Level	Initial probabilities	Simulation 1 Particip	Simulation 2 WGI VA	Simulation 3 ODA CI
Particip	HIGH	33.6 %	100	18.47	1.37
WGI VA	HIGH	51.4 %	28.3	100	<i>-1.1</i>
ODA CI	LOW	29.8 %	1.3	<i>-0.6</i>	100
CEC	HIGH	30.8 %	70.1 %	54.4 %	36.2 %

**Fig. 5.** Probabilities in % for the variables involved in the CEC pillar discretised in high (HIGH), medium (MID) and low (LOW) levels.

the Transboundary Movement of Hazardous Waste. This indicator may reveal the national commitment toward environment and environmental services such as water sanitation, energy, and waste. Water and sanitation delivery starts with infrastructure implementation, but rapidly calls to integrated and environmental management to ensure its sustainability.

However, this positive association between international concern and WSS may be different according to the priority and different perception given to WS and S. Safe drinking water supply directly supports human survival, hence it is a priority, while sanitation is often perceived as very much secondary. Sanitation is generally neglected by governments/public institutions even when it is a good leverage for the improvement of livelihood and health. With the launch of the “International Year for Sanitation 2008”, the UN tried to push countries to fight the lack of sanitation and to make progress towards the MDGs. In 2012, the sanitation target was still not reached. This mechanism is supported by global models: CEC is more important in the sensitivity analysis for WS than for S (See Table B1a and b).

Simulation CEC2 (WGI-VA set to 100 %): the role of the freedom of expression and accountability to citizens in the country commitment towards environment

The model estimates that an improvement of the degree of civil society freedom (WGI VA) means an increment of CEC from 30.8 to 54.4 %. A country in which citizens are able to express their criticisms and put pressure on their governments according to their interest is key in the construction of an efficient environmental management.

As mentioned in Dondeynaz et al. (2012), the voice of citizens both directly or indirectly pushes the government and other economic actors to sustainably manage the environment. The population is primarily concerned by lack of water resources, poor waste management, ecosystems degradation and impact of pollution. Being first in line, civil society plays a major role in control, alerting and pushing for sustainable management of the environment (Barrett and Graddy, 2000) if allowed by the institutional and political context.

Simulation CEC3 (ODA CI set to 100 %): external aid impacts on country environmental concern

The relationship between ODA CI and the CEC is not linear. The main trend is that high external aid flow (ODA CI) mainly indicates a context of medium (45 %) or good (34 %) CEC. The model estimates that high financial external flows are less invested in contexts where environmental concern is poor (only in 21 % of cases). It reveals the attention paid by international donors to intervene in relatively stable political context, support local population involvement and more general governance aspects (Sect. 4.1.1-simulation WS4).

4.2.3 Activity pressure on water resources (AP)

The structure of the AP model is set differently according to the correlations found out by a principal component analysis (Dondeynaz et al., 2012). Municipal and industrial withdrawals variables are correlated with total withdrawal (Fig. 6). The model shows a classification performance higher than 88 % (Table 1e).

The sensitivity analysis is computed for the four directly linked variables: water use intensity in agriculture (Water Use Int Agri), proportion of irrigated areas (% irrigation), the total amount of water withdrawn (total withdrawal) and water resources composite index (WR). In agreement with Dondeynaz et al. (2012), the weight of the water resources (WR) is the lowest in the model and therefore, it counts little in AP (Table B1e).

Simulation 1 is concerned with the weight of intensive agriculture water usage in AP, that combines the intensity of water usage in agriculture (Water Use Int Agri) and the proportion of irrigated areas (% irrigation). Simulation 2 aims at examining the competition between different uses

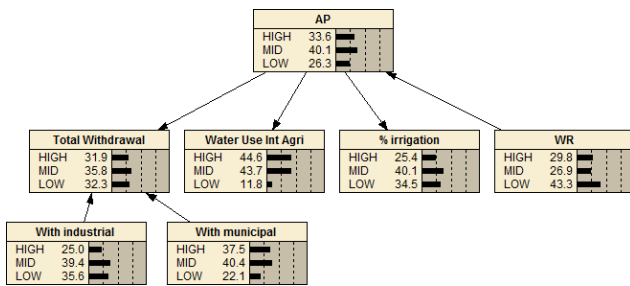


Fig. 6. DAG and initial probabilities for the variables involved in the AP pillar discretised in high (HIGH), medium (MID) and low (LOW) levels.

of water through their respective withdrawal amounts (total withdrawal, industrial withdrawal, municipal withdrawal). Simulation 3 observes the impact of potential reduction of available WR, under the pressure of climate change, on economic activity (Table 6).

Simulation AP1 (% irrigation and Water Use Int Agri set to 100 %): WSS sector and the water intensity in agriculture

Intensive agriculture practices are represented in this model by the percentage of irrigated areas (% irrigation) and the amount of water provided per hectare (Water Use Int Agri). In the simulation, an increment of intensive agriculture practices means an increment of AP up to 98.5 %.

The positive association between WSS and the intensive agriculture activities can be explained by several hypotheses: (1) the amount of water resources mobilised by such practices underlies the existence of facilities, and capacities. An organised sector at national level (providers with appropriated technical skills, institutional organisations. . .) can be the common denominator between WSS and Irrigation practices. (2) intensive agriculture practices allow multiple uses of water supply infrastructures developed for both crops and population in rural areas.

Looking at the WatSan4Dev subset, only 23 % of the countries considered present an irrigation rate higher than 10 %, mainly in Southeast Asia, the Middle East and some Maghreb countries. These countries show high WS access (but not necessarily high S) associated with irrigation schemes. These practices even small irrigation plants, structure rural areas and thus facilitate organisation of WSS services. For the rest of developing countries, the exclusive use of water resources for agriculture (high AP) can reveal an economy based on rainfall agriculture with low and weather-driven harvests. In this context, the population is poor and lives in rural areas where WSS access is generally low (Low WSS is estimated with around 33 % of probability despite high AP). Dissemination of population over rural areas can reinforce the effect of barriers to WSS delivery.

Simulation AP2 (Total withdrawal set to 100 %): the analysis of the water withdrawals by usage

The AP model estimates that an increase of the total water withdrawal increments AP up to 48.3 %; the municipal withdrawal variable decreases to 24.6 % and industrial withdrawal goes slightly up, to 26.5 %. This means that the industrial variable has a slightly different behaviour than municipal withdrawal with respect to Total Withdrawal: a certain industrial water demand can co-exist with dominant agricultural pressure.

Indeed, the behaviour of the total amount of withdrawals mainly depends on the agriculture water needs as confirmed by the AP model. Therefore, the relation between municipal and industrial withdrawal (in percentage of the total withdrawal) with the total amount of water withdrawals (amount of water drawn per capita) are negatively correlated (respectively -0.553 and -0.251). Looking at WatSan4Dev subset, the majority of developing countries (80 %) dedicate more than 50 % of available water to agricultural purposes with three main profiles: (i) 35 % of the countries dedicate water almost exclusively to agriculture (above 85 %); (ii) 40 % of the countries dedicate the majority of water resources to agriculture but this proportion is slowly being taken over by household consumption (corresponding to classification at AP medium level); (iii) 25 % of countries present substantial industrial/mining/forestry activities, i.e. higher than 10 %, often combined with high municipal consumption, significantly reducing the share dedicated to agriculture. This is the case, for instance, of Lesotho with 40 %, or Papua New Guinea with 42 % of water withdrawals dedicated to mining activities. Therefore, the total withdrawals variable is mainly driven by water agriculture demand, which competes with municipal and industrial demands.

Simulation AP3 (WR set to 100 %): the availability of water resources impacts on the activity development

According to the AP model, the reduction of water resources due to, for instance, climate change, slightly decreases the probability of high activity pressure (AP), from 33.6 % to 31.9 % (-5 %). At the same time, the probability of low human activity pressure still decreases from 26 % to 19.1 % (-26.5 %) to aliment also medium AP category ($+20$ %). This suggests that countries can find solutions adapted to mitigate the consequences of declining resources. This mechanism may be also confirmed by inputting observation data: arid countries such as the ones in the Middle East show high WSS access thanks to economic activities less conditioned by water resources. Therefore, these countries have funds available to implement infrastructures adapted to water scarcity such as desalination plants.

Irrigation can be an adapted solution to increase the efficiency of water use in agriculture. In that case, the limiting factors in such development are political commitment,

Table 6. Thematic model-AP simulations. In italic, negative variations and in bold, the positive variations of probabilities.

Variable	LEVEL	Initial probabilities	Simulation 1	Simulation 2	Simulation3
%irrigation	HIGH	25.5 %	100	9.8	<i>-1.3</i>
Water Use Int Agri	HIGH	44.7 %	100	11.1	0.9
Total withdrawal	HIGH	31.9 %	13.7	100	0.2
Withdrawal industrial	HIGH	25 %	(=)	1.5	(=)
Withdrawal municipal	HIGH	37.5 %	(=)	<i>-12.7</i>	(=)
WR	LOW	29.8 %	(=)	(=)	100
AP	HIGH	33.7 %	98.5 %	48.3 %	31.9 %

policies, investments, agronomic and management skills rather than the amount of water resources. In that sense, the FAO published recommendations encouraging the reinforcement of policies for planning and infrastructure development, including social aspects and capacity building for staff and farmers (FAO, 1987). As a concrete case, Kweku Baah Inkoo and Zomanaa Nangu (2011), analysed the reasons of under-utilisation of irrigation infrastructures in the upper west region of Ghana. They highlighted the importance of non-environmental factors such as lack of ownership, management and skills. The development of irrigation and the intensive use of water resources are the results of rather complex processes and factors, even if water scarcity or desertification may be a starting point for such intensive practice in agriculture.

Summarising the analyses of the AP model, the WSS sector supports agriculture, particularly in the case of intensive practices. In that case, agriculture participates in rural development and the structuring of these areas can then favour the development of WSS. However, this positive effect can be undermined when agriculture monopolises water resources and/or expresses poor economic development (e.g. agriculture does not generate enough income above survival).

4.3 Modelling country profiles

Dondeynaz et al. (2012) proposed five country profiles that were built around 5 axes corresponding to the first five PCA components: (HDP – human development and poverty, AP – human activity pressure, WR – water resources, ODA CI – Official development Assistance, CEC – country environmental concern). Profile 1 (towards well-being) is considered as the most favourable profile with high values on the WR, HDP and CEC pillars. Profile 2 (freedom/democracy black spot) shows weaknesses in terms of accountability and civil society's freedom, associated with a low commitment towards environment (CEC). Profile 3 (agricultural economy) presents an economy mainly driven by agricultural activities in a context of abundant natural resources. Profile 4 (essential external support) and profile 5 (primary material consumption) are the less favourable profiles with regards to human development and poverty (HDP). However, profile 4

benefits from higher levels of freedom, environmental concern (CEC), political stability and high external aid support (ODA CI). Profile 5 shows an economy mainly based on natural resources exploitation, often in a context of political instability.

The general models, described in Sect. 4.1, are used to model each country profile defined in Dondeynaz et al. (2012). The advantage of the models is to refine the analysis of each country profile, which improves the interpretation related to water supply and sanitation status. They also complete the characterisation of each group of countries by providing the probabilistic distribution of each axis. Finally, they allow to measure mechanisms specific to a profile and to run different scenarios. The statistical performances of country profile models are greater than 79 % (Table 7).

4.3.1 Analysing WSS behaviours across profiles

In a general way, as shown in Fig. 7, WS and S in Wat-San4Dev database show a coarse linear behaviour that can be expressed by the following equation:

$$S = 1.0285 \text{ WS} - 21.82, \quad (3)$$

with adjusted $R^2 = 0.5265$.

This linear model only explains 52.65 % (R^2) of data variability, but it highlights the general gap existing between S and WS (-21.82 %) as often observed in data and on the field (UNICEF and WHO, 2008). Some examples of an increased gap between WS and S are: Botswana (53 % between WS and S), Nepal (55 %) and Namibia (62 %). Conversely, Libya in 2004 experienced a negative gap between WS and S (72 % versus 97 %, respectively). These latter examples are part of those that cannot be explained by the linear model introduced in this paper.

In the top right-hand corner of Fig. 7, countries belonging to profile 1 (91 % for WS and 75 % for S) and profile 2 (93 % for WS and 86 % for S) show high-average WSS levels. On the bottom left corner, instead, countries from profiles 4 (62 % for WS and 35 % for S) and profile 5 (65 % for WS and 38 % for S) show low average WSS levels.

Countries from profile 3 follow a different WS and S statistical profile (purple dots on Fig. 7). WSS values in this

Table 7. Statistical performance of country profile models for WS and S.

Model name	Nb. of observations	error rate	LL*	QL*	SP*
WS-profile1	23	4.35 %	0.084	0.033	0.98
S-profile1	23	8.7 %	0.31	0.18	0.90
WS-profile2	17	5.88 %	0.16	0.079	0.96
S-profile2	17	0 %	0.054	0.0071	0.998
WS-profile3	19	21.05 %	0.42	0.27	0.85
S-profile3	19	10.53 %	0.32	0.19	0.91
WS-profile4	24	0 %	0.096	0.036	0.96
S-profile4	24	8.33 %	0.24	0.14	0.93
WS-profile5	18	16.67 %	0.37	0.22	0.88
S-profile5	18	0 %	0.034	0.0052	0.998

* LL = logarithmic loss, QL = quadratic loss where best value is 0, SP = spherical payoff where best value is 1.

case are less homogeneous, showing higher standard deviations: the WS average value is 73 % with 15.2 % standard deviation while S reaches 55 % with 22 % standard deviation. Profile 3 is built around the AP pillar with the predominant role being played by agriculture. WSS services show higher or lower levels of development according to the agriculture sector and its organisation/dynamics. In this group, irrigation practices are linked to relatively high WS: all countries with a high irrigation level show a high access to WS (above 75 %). However, countries like Guatemala, Honduras, Nicaragua, and Suriname show good access to WS as well, without irrigation use. This can be explained by an ongoing urbanisation that favours WS access in a context where agriculture is still dominant. In fact, these countries show an urban population rate around 45–50 % and informal slums are relatively limited (below 45 %). This relationship is less clear regarding sanitation. Country profiles: analysis of Profile 4 (essential external support) and 5 (primary material consumption).

The modelling is run for all five profiles, but because of the limited space in this paper, only profile 4 and 5 will be analysed in this section, since they show a less favourable status regarding WSS (Fig. 7).

4.3.2 Model running Profile 4 (essential external support)

For these countries, the sensitivity analyses of the WS model classify the HDP variables as the most important, followed by AP CEC and finally ODA CI. In the case of the S model, AP ranks fourth for S model switching with ODA CI. Countries belonging to this profile are in a less advantageous situation regarding WSS access than countries from profile 5 (see Sect. 4.3.1). However, they benefit from a stable political context which attracts external financial flows: 56 % of countries show a high ODA CI. The environmental concern is also good: 73 % of these countries show a medium-high level for CEC.

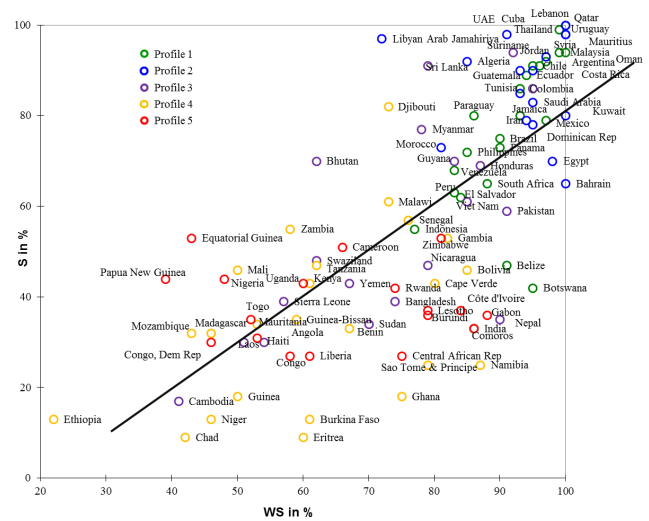


Fig. 7. WS versus S according to 5 country profiles with linear equation as $S = 1.0285WS - 21.82$ and $R^2 = 0.5265$.

Countries in this profile are first affected by poverty described in the model by the organisation of urbanisation (% slums), health access (Child Mortal-5, Malaria), education from basic to college (school enrol, school G/B) and advanced governance (WGI-GE). Simulation 1 estimates the increment of WS going up to 35.8 % and S up to 37 % with an increment of the HDP pillar (high HDP set to 100 %).

Water resources availability is also limited with around 96 % of countries having low (66 %) or medium (30 %) WR. In fact, 23 out of 24 countries are in Africa and have part of their territories under desertification threat. All countries had part of their national territory at risk of desertification in 2004; except Djibouti which is already a desertified area (FAO statistics). Sustaining adaptation or mitigation measures appears essential to fight against this constraint and ensure long-term WSS services.

In simulations 3 and 4, ODA CI highlights some differences from the general model (Sects. 4.1.1 and 4.1.2) regarding S and CEC. With the increment of ODA CI, high S drops down to 12.5 % and CEC increases up to 46.9 % (note that Table 8 is computed for High ODA CI category). In contrast with WS model, high S still increases up to 20 %, while CEC is concentrated on the medium-low category (34 % probability) when incrementing ODA CI at a medium level (100 % probability). In addition, an increment of ODA CI pushes the CEC at a higher level (46.9 % high probability) with respect to the S global model (34.9 % probability). This suggests that ODA CI has a better efficiency either on S and/or CEC than in a general behaviour.

More investigation in these specific countries may shed light on this specific behaviour for this profile and explain the reason in particular S is concerned.

Table 8. Profile 4: water supply and sanitation simulations.

Variables* – – WS model	Initial Probabilities	Simulation 1 HDP	Simulation 2 AP	Simulation 3 CEC	Simulation 4 ODA CI
HDP	11.5 %	100	1.1	0.1	<i>–2.33</i>
ODA CI	55.6 %	<i>–11.4</i>	0.5	11.6	100
CEC	36.2 %	0.3	1.1	100	7.6
AP	16.4 %	2	100	0.7	0.2
WR	3.7 %	(=)	(=)	(=)	(=)
WS	16.5 %	35.8 %	23 %	22.9 %	18.8 %
Variables* – S model	Initial Probabilities	Simulation 1 HDP	Simulation 2 AP	Simulation 3 CEC	Simulation 4 ODA CI
HDP	11.7 %	100	1.7	<i>–1.75</i>	<i>–3.11</i>
ODA CI	55.6 %	<i>–14.8</i>	<i>–1.5</i>	12.4	100
CEC	38.3 %	<i>–5.7</i>	<i>–1.3</i>	100	8.6
AP	15.6 %	2.2	100	<i>–0.4</i>	<i>–0.4</i>
WR	3.7 %	(=)	(=)	(=)	(=)
S	16.4 %	37 %	25.7 %	11.3 %	12.5 %

* Values concern HIGH category of each variable. In italic, negative variations and in bold, positive variations of probabilities.

4.3.3 Model running Profile 5 (Primary material consumption)

According to both sensitivity analyses, the WSS variables are influenced, in order, by HDP, AP, CEC and lastly ODA CI. Table 9 indicates that major efforts should be put on the three main dimensions (AP, HDP, and CEC). Simulations are ordered in the same way for WS and S. Countries from this profile present low access to WS in 75 % of cases while low S affects 88 % of cases.

Simulation 1 indicates that the increment of AP means an increase of WS up to 40.3 % and S up to 29.6 %. These countries are characterised by water resource abundance: 81 % of the countries benefit from medium (48 %) to high (33 %) WR. All countries (except Zimbabwe) withdraw a limited amount of water, all uses considered, relative to the full set.

Overall, 63.5 % of them show a low AP value, indicating that a significant part of the drawn WR is mainly used for industrial and raw material exploitation. The municipal use (urban activities and population) absorbs another important share of the remaining WR. As the total amount withdrawn is limited, the necessary water dedicated to domestic use appears bigger in relative terms. The agriculture share is therefore low, below 50 % of water withdrawals, and the share for irrigation practices near 0. Good WSS also appears dissociated with the industrial water demand, highlighting no significant synergies between industrial and population supply despite intensive agriculture/urban activities (see Sects. 4.1 and 4.2.3).

In this case, the natural resource exports have “counteracting effects on growth, by weakening the manufacturing and agricultural sectors” (Knack, 1999). Diversifying activities such as the agricultural sector in these countries may

evolve with WSS services improvement thanks to synergies (Sect. 4.2.3). Equatorial Guinea, Central Africa, Lesotho, Papua New Guinea, Congo, and the Democratic Congo are examples of these economies strongly oriented to oil extraction and mining, where agriculture is weak and characterised by subsistence farming. In these cases, the redistribution and investment of this revenue is crucial.

The countries that are showing better AP value (16 %), because of a lower industrial share, can focus on the direct reinforcement of water supply-sanitation capacities. This will benefit the population and also the agricultural sector (main employer of labour force⁹). Comoros, Cameroon, Burundi, Togo and Rwanda are representative of this category.

Simulation 2 shows that an increment of HDP induces an increment of WS and S up to 40.2 % and 0.9 %, respectively. Poverty in these countries is a crucial issue encouraging high fecundity, informal urban settle development, and generally poor health, particularly children, as well as poor education. Advanced governance is globally weak (WGI-GE) and therefore the capacities of the states are small.

These states should be supported by international donors; however their political instability (measured by WGI PS AV) can be a barrier. In fact, only 15 % of countries in profile 5 benefit from high external aid (ODA CI) while the majority (57 %) show medium ODA CI, and almost 30 % a low ODA CI. Simulation 4 shows that an increment of ODA CI will produce an increase of HDP up to 23.3 % for both WS and S models.

⁹Consult country profile from the CIA Factbook: <https://www.cia.gov/library/publications/the-world-factbook/>, last access: 4 July 2013.

Table 9. Profile 5: water supply and sanitation simulations.

Variables*	Initial	Simulation 1	Simulation 2	Simulation 3	Simulation 4
WS model	Probabilities	AP	HDP	CEC	ODA CI
HDP	14.6 %	1.8	100	3.1	8.7
ODA CI	14.3 %	(=)	8.5	8.5	100
CEC	14.6 %	1.8	3.1	100	8.7
AP	15.4 %	100	2.4	2	(=)
WR	33.3 %	(=)	(=)	(=)	(=)
WS	24.9 %	40.3 %	40.2 %	40.2 %	25 %

Variables	Initial	Simulation 1	Simulation 2	Simulation 3	Simulation 4
S model	Probabilities	AP	HDP	CEC	ODA CI
HDP	13 %	4.1	100	5.6	10.3
ODA CI	14.3 %	3.1	11.4	11.4	100
CEC	13 %	4.1	5.6	100	10.3
AP	13.6 %	100	4.2	4.2	2.9
WR	33.3 %	(=)	(=)	(=)	(=)
S	12 %	29.6 %	30.9 %	30.9 %	25 %

* Values are intended as HIGH category for each variable. In italic, negative variations and in bold, positive variations.

The improvement of CEC is associated with increments of WSS as shown in simulation 3: high WS should reach 40.2 %, and S 30.9 %. Environmental management appears even more essential in an economy focused on industrial/natural resources exploitation. The CEC is mainly low in these countries (48.5 %) also meaning that efforts are made to improve civil society freedom and government accountability to citizens. Basic governance should be strengthened.

5 Discussions

The models proposed appear to be flexible tools to measure associations between the 25 variables that interact with the WSS access in developing countries (101 observation countries). These models highlight key variables and mechanisms on how WSS basic services contribute to the national development. The analyses are performed both for all developing countries and also for specific country profiles. In addition, the Bayesian networks method used for modelling allows for the running of probabilistic scenarios.

Concerning the issue related to the development of World-wide Governance Indicators (WGI), Thomas (2009) argues that the governance data are based on expert assessments, which are often subjective, that indicators are spotty, with standard error issues and as a consequence, non-extendable and non-comparable between countries. Langbein and Knack (2010) question the number of indices created to measure governance, and justify the number of distinct governance variables. They call for caution, in particular in making country evaluations using these indicators. Thomas (2009) also noted that “evidence must be provided to show that a purported measure of a theoretical construct is valid both in its conceptualisation and its operationalisation,

by exploring predicted relationships with other observable variables”. Kaufmann et al. (2009) respond that “construct validity is not a useful tool to assess the merits of the WGI, and even if it were, Thomas (2009) provides no evidence of any practical consequences of failure to meet the criteria of construct validity”.

Despite these limits and criticisms, the authors’ point of view is that measuring governance is relatively new and has been subject to caution since the 1990s. This work uses the WGIs but verifies their coherency with other indicators provided by others institutions, namely environmental governance (subindicator included in the Environmental Sustainability Index 2005 provided by the World Economic Forum) and corruption perception index (CPI provided by Transparency International). The coherency is confirmed throughout the preliminary analysis performed (FA and PCA) in Dondeynaz et al. (2012). Although the models proposed in this paper show the limits and constraints linked to the accuracy of the WGIs values, the paper, together with literature and field experience, demonstrates the crucial role of governance in sustainable WSS (Plummer and Cross, 2006). Therefore, the number of WGIs is limited and the relevant indicators are selected according to their correlations: (1) government effectiveness (WGI GE), which also represents corruption control (WGI CofC), rule of law (WGI RofL), regulatory quality (WGI RQ), Corruption Perception Index (CPI) and environmental governance (Env Gov) (called advanced governance throughout the paper); (2) political stability and absence of violence (WGI PS AV); and (3) voice and accountability (WGI-VA) (Dondeynaz et al., 2012). The last two are called basic governance.

Moreover, it is to be noted that governance aspects, when measured at the scale of the water sector as a whole, often

include water supply and sanitation, agriculture and hydropower. The disaggregation of data by subsectors, specific analyses and/or case studies, are rare at national country level, particularly in relation to sanitation. For instance, the description of corruption mechanisms (i.e. by Transparency International, 2008) are described for the whole water sector, with specific information in the case of water supply, sanitation being enclosed in WSS or omitted. Having such breakdowns or at least, compared case studies on corruption mechanisms in both subsectors could help in providing more accurate interpretations.

Another aspect is, as mentioned in Dondeynaz et al. (2012), the absence of indicators on water quality that should be further developed by the international institutions in charge of monitoring MDGs. For this purpose, the JMP set a task force on “monitoring drinking water quality” (UNICEF WHO, 2011). This missing aspect in characterising water resources could explain the negligible weight of water resources in the models suggested in this paper. Such indicators at national level could allow for the modelling of the interactions between environment and development through the analysis of the qualitative impacts on water resources. This would be particularly useful in the case of sanitation. Lack of adequate sanitation services is known to impact water quality and hence health (i.e. high rates of child mortality in sub-Saharan Africa are partly due to waterborne diseases).

For the majority of developing countries, the access to water supply is higher than the access to basic sanitation (21.82 % average – see Sect. 4.3.1). In fact, one reason for this gap lies in cultural and psychological aspects linked to sanitation. Indeed, the “sanitation for all” year (2008) highlights its “unpopularity” leading to a “sanitation crisis”¹⁰. Improved sanitation is often not considered as necessary/vital for improving the quality of life. However, understanding these cultural and psychological factors is essential to its sustainable development. Up to now, there has been no quantitative or qualitative means to appreciate and measure these effects. In many countries, awareness campaigns have been proven as a crucial leverage to convincing populations to implement, and keep functioning sanitation facilities. A community-led total sanitation approach¹¹ developed and applied in Bangladesh starting in 1999 demonstrates this necessary cultural/educational work to be included in sanitation programs.

6 Conclusions

The identification of mechanisms influencing water supply and sanitation (WSS) is complex because of the

¹⁰International Year of Sanitation 2008: http://esa.un.org/iys/docs/IYS_flagship_web_small.pdf, last access: 4 July 2013.

¹¹<http://www.communityledtotalsanitation.org/>, last access: 4 July 2013.

Table 10. Recapitulative table on probability variations (difference between recomputed and initial probability).

Variables*	Suitable Level	High WS probability change	High S probability change
		Global models	
HDP	HIGH	35.9	34.5
AP	HIGH	18.1	21.3
CEC	HIGH	24.1	8.1
ODA CI	LOW	10.8	7.3
	Suitable Level	High HDP probability change	
GDP per cap	HIGH	53	
Urban pop	HIGH	43.1	
Child Mortal -5	LOW	33.7	
%Poverty	LOW	33.2	
%Slums	LOW	31.9	
Malaria	LOW	27.8	
WGI GE	HIGH	24.8	
School enrol	HIGH	24.4	
School G/B	HIGH	18	
Femal eco	LOW	16.6	
ODA CI	LOW	0.3	
	Suitable level	High CEC probability change	
Particip	HIGH	39.4	
WGI VA	HIGH	23.8	
ODA CI	LOW	5.6	
	Suitable level	High AP probability change	
%irrigation	HIGH	59.3	
Water Use in Agriculture	HIGH	35.7	
Total Withdrawal	HIGH	14.7	

* Water resource variables are subject to very long-term pattern changes, therefore they are only used as descriptive indicators.

cross-interaction between multiple factors and issues. The Bayesian networks method is chosen to model the Wat-San4Dev sub-data set that includes 25 indicators and five dimensions in WSS (Dondeynaz et al., 2012). Indeed, Bayesian networks present the advantage of providing probabilistic conditional dependence relationships between variables. The proposed models in this paper are efficient tools to identify and measure probabilistic changes of the selected elements and WSS levels (statistical performances range between 80–95 %). For thematic models, the link rather expresses the relationship between the composite indicator and its subvariables than a causal link.

In a first step, general models were built separately for water supply and sanitation to observe and identify potential differences. Accordingly, composite indicators are computed for the following dimensions/pillars defined in Dondeynaz et al. (2012): human development and poverty (HDP), country environmental concern (CEC), activity pressure on water resources (AP), Official Development Assistance flows (ODA CI) and water resources availability (WR). These models synthesise and estimate large mechanisms involved in WSS access, existing in the common field literature. The statistical classification error rates are less than 12 % for water supply (WS) and 10 % for sanitation (S).

Table A1. Limits of variable categories, high, mid, low.

Variables	Categories			Maximum	Unit or intervals*
	LOW Minimum	MID Upper limit	HIGH Upper limit		
Water supply (WS)	22	82	na	100	%
Sanitation (S)	9	58	na	100	%
%Poverty	8	29	52	71	%
Urban Pop	10	45	78	98	%
School enrol	23	39	67	94	%
Femal eco	18	39	73	92	%
% Slums	0	32	64	97	%
Withdrawal municipal	1	19	53	84	%
Withdrawal industrial	0	26	55	73	%
TWRR	8	40771	203825	326116	m ³ yr ⁻¹ per cap
Child Mortal-5	7	85	166	265	⁰ / ₀₀
School G/B	0.63	0.72	0.87	1.02	[0.63,1.02]
Particip	0.30	0.38	0.70	0.90	[0, 1]
GDP per cap	0.58	11.11	19.17	24.03	thousand \$per cap
%irrigation	0	34	68	100	%
ODA WSS	0	5.55	11.26	16.51	\$ per cap
ODA	0.60	103	195	289	\$ per cap
Malaria	0	144	288	428	⁰ / ₀₀
Water Use Int Agri	0.50	7565	17 708	28 333	m ³ ha ⁻¹ per cap
Precipit	51	1062	2103	3142	mm yr ⁻¹
Desert risk	0	33.70	66.40	100	%
Total withdrawal	6	748	1476	2161	m ³ yr ⁻¹ per cap
WGI VA.	-2.14	-0.53	na	1.17	[-2.5, +2.5]
WGI PS AV	-2.46	-1.11	na	1.10	[-2.5, +2.5]
WGI GE	-1.69	0.10	na	1.25	[-2.5, +2.5]
CEC	-2.70	-0.86	1.08	2.78	na
HDP	-4.56	-0.27	2.96	5.04	na
ODA CI	-2.35	-0.75	1.62	3.01	na
AP	-4.39	-2.30	0.61	3.23	na
WR	-2.43	-1.62	-1.39	2.87	na

* na: composite indicator with dimensionless unit.

First, the priorities set by the MDGs initiative in reducing poverty are supported by the key findings from this modelling exercise (Table 10). Indeed, general models confirm that the improvement of WSS access is a good leverage for the alleviation of poverty. In particular, the reduction of child mortality and slums areas are the first two variables strongly dependent on the increase of WSS coverage (Sect. 4.2.1). The improvement of living conditions in slums strongly depends on the availability and organisation of WSS basic services. Moreover, it should be mentioned that the urbanisation processes bring incentives and the critical mass necessary for WSS infrastructure development and costs recovery.

In addition, external aid (ODA CI) is provided preferentially to the poorest countries, but not exclusively. Funds are also invested in relatively less poverty-affected countries (still not considered as developed) probably because of the geo-strategy policy of donors.

The development of irrigated agriculture and the appearance of urban activities are also associated with WSS development. Models point out that sanitation benefits more from an efficient agriculture (i.e. irrigation) respect to water supply. This may also be suggested by field experience: fertilisers based on the reuse of human excreta or/and wastewater sludge treatment can be solutions integrated into rural development once the cultural/psychological barrier is overcome.

Finally, the models highlight that water availability (WR) counts for little in the population access to WSS, as well as for the economic activities. This link between water resources, usages and WSS would be certainly observed if indicators related to water quality existed. This work suggests that water quality is an essential element; international institutions are working on including this dimension in the MDGs monitoring.

Thematic models allow for the identification of key elements within the five dimensions/pillars:

Table B1. Sensitivity analysis results by model.

Models	Variables	ER	BV	RMS	rank
(a) WS general model	HDP	0.36475	0.1136945	0.3372	1
	CEC	0.06549	0.0221944	0.149	2
	AP	0.05969	0.0202661	0.1424	3
	ODA CI	0.01701	0.005855	0.07652	4
	WR	na	na	na	5
(b) S general model	HDP	0.3201	0.099637	0.3157	1
	AP	0.12310	0.0394685	0.1987	2
	CEC	0.01982	0.006738	0.08208	3
	ODA CI	0.01258	0.004285	0.06546	4
	WR	na	na	na	na
Models	Variables	ER	BV	rank	
(c) HDP model	GDP_per_cap	0.75752	0.197325	1	
	Child Mortal-5	0.52090	0.110159	2	
	% Slums	0.51033	0.098481	3	
	School enrol	0.44786	0.059585	4	
	Malaria	0.43765	0.077565	5	
	% Poverty	0.38850	0.050855	6	
	Urban Pop	0.34604	0.067429	7	
	School G/B	0.28987	0.068078	8	
	WGI GE	0.26849	0.035107	9	
	Femal eco	0.18374	0.038455	10	
	ODA CI	0.00195	0.000416	11	
(d) CEC model	Particip	0.41790	0.083801	1	
	WGI VA	0.37631	0.048051	2	
	ODA CI	0.03400	0.004356	3	
(e) AP model	% Irrigation	0.68640	0.161407	1	
	Water Use Int Agri	0.52215	0.090917	2	
	Total withdrawal	0.07318	0.011538	3	
	WR	0.03537	0.005556	4	

* ER = entropy reduction, BV = belief variance, RMS = root mean square change.

1. Human development and poverty model (HDP): the key variables are, in order, household income (GDP per cap), health improvement (child mortality and malaria), urbanisation processes, where the reduction of informal settlement is crucial (Urban Pop, % slums, Poverty) and lastly, education (school enrol). The WSS access facilitates formal urbanisation processes and is low in the case of development of informal urban settlements. It also emphasises the importance of disease control (e.g. malaria) and child mortality, where WSS services play an essential role. Sustainable management of WSS services can support governance, but the interaction between these two variables probably follows a virtuous cycle. Education of the population is strongly facilitated in the case of availability of WSS services. The first beneficiaries are women that see their daily burden alleviated, while girls' attendance at school is fostered.
2. Country environmental concern model (CEC): in both cases, the country environmental commitment in front of the international community (Particip) rather indicates a basic strategy toward WSS improvement. This environmental concern is supported by the civil society if the context makes it possible. This framework is attractive to external donors (ODA CI).
3. Activity pressure on water resources model (AP): includes agriculture intensity, water demand and water resources' availability. This pillar highlights synergies with the development of the economic activities: an organised agriculture sector benefits from a strong WSS sector. It may allow for a rural development that can then facilitate delivery of WSS services. The availability of water resources (WR) has little influence on the pattern of activity, development or the level of water demand.

Table 10 summarises the potential probability changes of WS, S and composite indicators according to the increments of the 25 selected variables.

In a second step, this general model is run for country profiles presented in Dondeynaz et al. (2012). Five profiles were built around the five pillars: HDP, CEC, AP, ODA CI and WR. The models and their behaviours highlight divergences in key variables or mechanisms and enable running scenarios for the different profiles. Generally speaking, S access mainly falls behind WS (−21 % on average); however, the relationship between the two variables also shows a non-linear behaviour (Fig. 7). Profile 1 (towards well-being) and 2 (freedom/democracy black spot) include more advanced countries, while profile 3 (agricultural economy) shows spread-out WSS distributions. Countries from profile 4 (essential external support) and 5 (primary material consumption), mainly sub-Saharan countries, are struggling with both water supply and sanitation access and record the lowest levels in developing countries. Within the limits of this paper, groups of countries from profile 4 and 5 are specifically analysed because they are the least advanced countries among the considered subset.

Countries belonging to profile 4 experience desertification effects in a context of limited availability of water resources (66 % with low WR). Therefore, the development of adapted or mitigation measures is crucial. Taking advantage of conditions such as minimum governance (political stability and civil society freedom) and concern about the environment, the main effort should be oriented towards the development of the HDP pillar, meaning: control of urbanisation, organisation-reinforcement of health care, education, and improved advanced governance related to state effectiveness. The model identifies a slightly different behaviour of ODA CI, S and environmental commitment (CEC). ODA CI appears to be more effective on S and CEC when considered in the general model. More investigation in these specific countries may shed light on this behaviour and why only sanitation is concerned.

Profile 5 shows weaknesses in almost all dimensions, having low HDP (68–70 %), low CEC (48–50 %), and low values in AP (64–67 %). The majority of these countries have an economy oriented towards a combination of industries and activities that exploit natural resources, in a context of population poverty and low access to WSS. Despite this situation, these countries benefit from a medium level of investment by donors (mostly medium level ODA CI involves 57 % of countries) that can be often explained by the persistence of political instability. Therefore, ensuring an environment of basic governance conditions, stability, absence of violence and accountability to citizen is essential: (1) to foster economic diversification and reinforce rural development; (2) to sustain efforts made for poverty reduction; and (3) to appeal for additional external support if necessary.

Overall, 16 % of countries that show better AP values are countries where agriculture, combined with some industries,

is the main source of employment, but globally this also reflects poverty. Increasing the revenue produced by this activity appears crucial. Investing in the WSS sector may also support this process (building capacities, creating complementary market opportunities, etc.).

Acknowledgements. The present work benefited from the input from the various international institutions mentioned in this paper that provided the necessary data sets to handle this research. We also thank Alan Belward, head of the Global Environmental Monitoring Unit (JRC) and Giovanni Bidoglio, head of Water Resources Unit (JRC) for their constant support to this research project. This study benefited from the support of the RALCEA and ACE-Water projects (EC contracts AA 2010/241-167 and AA 2009/220-992, respectively) component: stakeholder mapping and participation. The comments of the reviewers are greatly appreciated.

Edited by: A. Ghadouani

References

- Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., and Salmerón, A.: Bayesian networks in environmental modelling, *Environ. Modell. Softw.*, 26, 1376–1388, doi:10.1016/j.envsoft.2011.06.004, 2011.
- Baran, E. and Jantunen, T.: Stakeholder consultation for bayesian decision support systems in environmental management, *Proceedings of ECOMOD*, 15–16 September, Penang, Malaysia 2004.
- Barrett, S. and Graddy, K.: Freedom, Growth, and the Environment, *Environ. Dev. Econ.*, 5, 433–456, 2000.
- Batchelor, C. and Cain, J.: Application of belief networks to water management studies, *Agr. Water Manage.*, 4, 51–57, 1999.
- Bayes, T.: An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions*, 53, 370–418, doi:10.1098/rstl.1763.0053, 1763.
- Boserup, E.: Population, the status of women and rural development, *Population and development review*, 15, Supplement: rural development and population: institutions and policy, 45–60, 1989.
- Bromley, J., Jackson, N., Clymer, O., Giacomello, A., and Jensen, F.: The use of Hugin to develop Bayesian Network as an aid to integrated water resource planning, *Environ. Modell. Softw.*, 20, 231–242, 2005.
- Borsuk, M., Stow, C., and Reckhow, K.: A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis, *Ecol. Modell.*, 173, 219–239, 2004.
- Burtaine, M. T. and Parks, B. C.: When Do Environmentally-Focused Aid Projects Achieve their Objectives? Evidence from World Bank Post-Project Evaluations, 40 pp., available at: <http://irtheoryandpractice.wm.edu/bcparks/publications/achieve.pdf>, 2012.
- Carbonnier, G.: Official development assistance once more under fire from critics, *Revue internationale de politique de développement*, 1, 137–142, doi:10.4000/poldev.141, 2010.
- Carmona, G., Varela-Ortega, C., and Bromley, J.: The use of participatory object-oriented Bayesian networks and agro-economic

- models for groundwater management in Spain, *Water Resour. Manage.*, 25, 1509–1524, doi:10.1007/s11269-010-9757-y, 2011.
- Castelletti, A. and Soncini-Sessa, R.: Coupling real-time control and socio-economic issues in participatory river basin planning, *Environ. Modell. Softw.*, 22, 1114–1128, doi:10.1016/j.envsoft.2006.05.018, 2007a.
- Castelletti, A. and Soncini-Sessa, R.: Bayesian Networks and participatory modelling in water resource management, *Environ. Modell. Softw.*, 22, 1075–1088, doi:10.1016/j.envsoft.2006.06.003, 2007b.
- Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J.: *Probabilistic networks and expert systems*, Harrisonburg, VA: Springer, 1999.
- Cox, R. T.: Probability, frequency and reasonable expectation, *Am. J. Phys.*, 14, 1–13, 1946.
- Dixon, J. R.: *A programmed introduction to probability*, New York: John Wiley & Sons, 1964.
- Dondeynaz, C., Carmona Moreno, C., and Céspedes Lorente, J. J.: Analysing inter-relationships among water, governance, human development variables in developing countries, *Hydrol. Earth Syst. Sci.*, 16, 3791–3816, doi:10.5194/hess-16-3791-2012, 2012.
- Edwards, W.: Hailfinder, Tools for and experiences with bayesian normative modeling, *Am. Psychol.*, 53, 416–428, doi:10.1037/0003-066X.53.4.416, 1998.
- FAO: Consultation on irrigation in Africa, FAO irrigation and drainage paper, 42, Proceeding of consultation on irrigation in Africa, Lome, Togo, ISBN:92-5-102547-9, 1987.
- Giné Garriga, R., Pérez Foguet, A., Molina, J. L., Bromley, J., and Sullivan, C.: Application of Bayesian Networks to assess water poverty, II International Conference on Sustainability Measurement and Modelling, Barcelona, <http://upcommons.upc.edu/e-prints/bitstream/2117/7804/1/bayesiannetworks.pdf> (last access: 28 August 2013), 2009.
- Heckerman, D.: *A tutorial on learning with bayesian networks* (Tech. Rep. MS-TR-95-06), Redmon, WA: Microsoft Research, 1995.
- Hoeting, J. A., Madigan, D., and Raftery, A. E.: Bayesian model averaging: a tutorial, *Stat. Sci.*, 14, 382–401, 1999.
- Honaker, J., King, G., and Blackwell, M.: *Amelia II: A Program for Missing Data*, Version 1.2-17, <http://r.iq.harvard.edu/docs/amelia/amelia.pdf> (last access: 28 August 2013), 2011.
- Houthakker, H. S.: Education and Income, *Rev. Econ. Stat.*, 41, 24–28, 1959.
- Jansen, R., Yu, H., Greenbaum, D., Kluger, N., Krogan, N. J., Chung, S., Emili, A., Snyder, M., Greenblatt, J. F., and Gerstein, M.: A bayesian network approach for predicting protein interactions from genomic data, *Science*, 302, 449–453, 2003.
- Kaufmann, D., Kraay, A., and Mastruzzi, M.: Response to “What Do the Worldwide Governance Indicators Measure?”, 4 pp., <http://info.worldbank.org/governance/wgi/pdf/KKMResponseEJDR2Final.pdf>, (last access: 28 August 2013), 2009.
- Knack, S.: Aid Dependence and the Quality of Governance: A Cross-Country Empirical Analysis, World Bank publications, 38 pp., doi:10.1596/1813-9450-2396, 1999.
- Kweku Baah Inkoom, D. and Zomanaa Nanguo, C.: Utilisation of irrigation facilities towards poverty reduction, *Journal of Sustainable Development*, 13, 335–351, 2011.
- Langbein, L. and Knack, S.: The Worldwide Governance Indicators: Six, One, or None?, *J. Dev. Stud.*, 46, 350–370, 2010.
- Lee, S. M., Abbott, P., and Johantgen, M.: Logistic regression and Bayesian networks to study outcomes using large data sets, *Nurs. Res.*, 2, 133–138, 2005.
- Lewis, L.: Cholera, Dengue Fever, and Malaria: The Unquestionable Link to Water, the water project, available at: <http://thewaterproject.org/cholera-dengue-fever-malaria-water.php> (last access: 29 August 2013), 2011.
- Martín de Santa Olalla, F., Dominguez, A., Ortega, F., Artigao, A., and Fabeiro, C.: Bayesian networks in planning a large aquifer in Eastern Mancha, Spain, *Environ. Modell. Softw.*, 22, 1089–1100, doi:10.1016/j.envsoft.2006.05.020, 2007.
- McQueen, J. B.: Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 281–297, available at: <http://projecteuclid.org/euclid.bsm/1200512992> (last access: 29 August 2013), 1967.
- Molina, J. L., Bromley, J., Garcia-Arostegui, J. L., Sullivan, C. A., and Benavente, J.: Integrated water resources management of overexploited hydrogeological systems using Object-Oriented Bayesian Networks, *Environ. Modell. Softw.*, 25, 383–397, 2010.
- Nadkarni, S. and Shenoy, P. P.: A bayesian network approach to making inferences in causal maps, *Eur. J. Oper. Res.*, 128, 479–498, 2001.
- Nadkarni, S. and Shenoy, P. P.: A causal mapping approach to constructing bayesian networks, *Decis. Support Syst.*, 38, 259–281, doi:10.1016/S0167-9236(03)00095-2, 2004.
- Neapolitan, R. E.: *Probabilistic reasoning in expert systems: Theory and algorithms*, New York: Wiley & Sons, 1990.
- OECD-DAC: Measuring Aid to the Water sector, 4 pp., available at: <http://www.oecd.org/dac/stats/36365514.pdf>, 2007
- Pearl, J.: An economic basis for certain methods of evaluating probabilistic forecasts, *Int. J. Man-Mach. Stud.*, 10, 175–183, doi:10.1016/S0020-7373(78)80010-8, 1978.
- Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 2nd Edn., San Mateo, CA: Morgan Kaufmann, 1991.
- Pearl, J.: *Bayesian networks, causal inference and knowledge discovery* (Tech. Rep. 281), Los Angeles, CA: University of California, 2001.
- Plummer, J. and Cross, P.: Tackling corruption in the water and sanitation sector in Africa: starting the dialogue, Water and Sanitation Program (WSP), Working Paper, 2006.
- Raschid-Sally, L. and Jayakody, P.: Drivers and characteristics of wastewater agriculture in developing countries: Results from a global assessment, International Water Management Institute, Colombo, Sri Lanka, IWMI Research Report 127, 35 pp., 2008.
- Santiso, C.: International co-operation for democracy and good governance: moving toward a second generation?, *Eur. J. Dev. Res.*, 13, 154–180, 2001.
- Sullivan, C. A., Meigh, J. R., Giacomello, A. M., Fediw, T., Lawrence, P., Samad, M., Mlote, S., Hutton, C., Allan, J. A., Schulze, R. E., Dlamini, D. J. M., Cosgrove, W., Delli Priscoli, J., Gleick, P., Smout, I., Cobbing, J., Calow, R., Hunt, C., Hussain, A., Acreman, M. C., King, J., Malomo, S., Tate, E. L., O’Regan, D. O., Milner, S., and Steyl, I.: The water poverty index: De-

- velopment and application at the community scale, *Nat. Resour. Forum*, 27, 189–199, doi:10.1111/1477-8947.00054, 2003.
- Sušnik, J., Vamvakeridou-Lyroudia, L. S., Savić, D. A., and Kapelan, Z.: Integrated System Dynamics Modelling for water scarcity assessment: Case study of the Kairouan region, *Sci. Total Environ.*, 440, 290–306, doi:10.1016/j.scitotenv.2012.05.085, 2012.
- Thomas, M.: What do the worldwide governance indicators measure? *European Journal of Development Research*, 16 July 2009, available at SSRN: <http://ssrn.com/abstract=1007527> (last access: 28 August 2013), 2009.
- Transparency International: Global Corruption report 2008, available at: http://www.transparency.org/whatwedo/pub/global_corruption_report_2008_corruption_in_the_water_sector (last access: 28 August 2013), 2008.
- UNPD: World urbanisation prospects 2001, 35–49, available at: http://www.un.org/esa/population/publications/wup2001/WUP2001_CH3.pdf, 2001.
- UNDP and UNICEF: Human Development Report 2006: Children and Water, Sanitation and Hygiene: The Evidence, occasional paper, available at: <http://hdr.undp.org/en/reports/global/hdr2006/papers/unicef.pdf> (last access: 28 August 2013), 2006.
- UNICEF WHO: Progress on Drinking water and Sanitation: special focus on Sanitation, UNICEF, New York and WHO, Geneva, 58 pp., available at: http://www.wssinfo.org/fileadmin/user_upload/resources/1251794333-JMP_08_en.pdf, 2008.
- UNICEF WHO: Diarrhoea: Why children are still dying and what can be done?, WHO publications, Geneva, 57 pp., available at: http://whqlibdoc.who.int/publications/2009/9789241598415_eng.pdf, 2009.
- UNICEF WHO: Drinking water, Equity, Safety and sustainability, WHO publication, Geneva, Switzerland, 66 pp., available at: http://www.wssinfo.org/fileadmin/user_upload/resources/report_wash_low.pdf, 2011.
- Varis, O. and Kuikka, S.: Bayesian approach to expert judgment elicitation with case studies on climatic change impact assessment on surface waters, *Climatic Change*, 37, 539–563, 1997.
- WHO press release: Millennium Development Goal drinking water target met, available at: http://www.who.int/mediacentre/news/releases/2012/drinking_water_20120306/en/index.html (last access: 28 August 2013), 6 March 2012.
- Xiang, Y.: Probabilistic reasoning in multi-agent systems. A graphical models approach, Cambridge, Cambridge University Press, doi:10.1017/CBO9780511546938, 2002.
- Zorrilla, P., Carmona, G., De la Hera, A., Varela-Ortega, C., Martínez-Santos, P., Bromley, J., and Henriksen, H. J.: Evaluation of Bayesian Networks in Participatory Water Resources Management, Upper Guadiana Basin, Spain, *Ecol. Soc.*, 15, 17 pp., 2009.