Hydrol. Earth Syst. Sci. Discuss., 10, 2481–2548, 2013
www.hydrol-earth-syst-sci-discuss.net/10/2481/2013/
doi:10.5194/hessd-10-2481-2013

Hydrology and Earth System Sciences
Discussions
Open Access

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

C. Dondeynaz[1,2], J. López Puga[3], and C. Carmona Moreno[1]

[1]Institute for Environment and Sustainability, Joint Research Centre, Ispra, Italy
[2]Liverpool University, Liverpool, UK
[3]Catholic University San Antonio, Murcia, Spain

Received: 1 February 2013 – Accepted: 15 February 2013 – Published: 28 February 2013

Correspondence to: C. Dondeynaz (celinedondeynaz@yahoo.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Despite the efforts made towards the millennium goals targets during the last decade, access to improved water supply or basic sanitation remains still not accessible for millions of people across the world. This paper proposes a set of models that use 25 key variables from the WatSan4Dev dataset and country profiles involving Water Supply and Sanitation (Dondeynaz et al., 2012). This paper proposes the use of Bayesian Network modelling methods because adapted to the management of non-normal distribution, and integrate a qualitative approach for data analysis. They also offer the advantage to integrate preliminary knowledge into the probabilistic models. The statistical performance of the proposed models ranges between 80 and 95 % which is very satisfactory taking into account the strong heterogeneity of variables. Probabilistic scenarios run from the models allow a quantification 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, a strong poverty reduction will induce an increment of the WSS access equal to 75–76 % through: (1) the organisation of on-going urbanisation process to avoid slums development; and, (2) the improvement of health care for instance for children.

On one side, improving governance, such as institutional efficiency, capacities to make and apply rules or control of corruption will also have a positive impact on WSS sustainable development. The first condition for an increment of the WSS access remains of course an improvement of the economic development with an increment of household income.

Moreover, a significant country environmental commitment associated with civil society freedom of expression constitutes a favourable environment for sustainable WSS services delivery. Intensive agriculture through irrigation practises also appears as a mean for sustainable WSS thanks to multi-uses and complementarities. Strong and structured agriculture sector facilitates rural development in areas where WSS access

often steps behind compared to urban areas[1]. External financial support, named Official Development Aid (ODA), plays a role in WSS improvement but comes last in the sensitivity analyses of models.

This aid supports first poor countries at 47 %, and is associated to governance aspects: (1) political stability and (2) country environmental commitment 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, to fight against water scarcity and desertification pressure should be the priority. However, for countries in profile 5, efforts should first concentrate on political stability consolidation while supporting economic activity diversification. Nevertheless, for both profiles, reduction of poverty should remain the first priority as previously indicated.

1 Introduction

On 6 March 2012[2], the UN announced that Millennium target for safe drinking water is reached, while sanitation target is 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). On the same time, the UN estimates that by 2015, 92 % of the global population will have access to improved drinking water. Beyond these global figures, disparities appear 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).

[1]JMP statistics, 2004 http://www.wssinfo.org/data-estimates/table/

[2]http://www.who.int/mediacentre/news/releases/2012/drinking_water_20120306/en/

HESSD
10, 2481–2548, 2013

Access to basic Sanitation represents still a challenge for next 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), Integrated Water Resources Management (IWRM) approach has been adopted by the International Community in 1992 in Dublin's conference. This considers the whole water cycle from water withdrawal to restitution to the environment including also sanitation, the involvement of users and the preservation of the natural resources. IWRM promotes an inclusive way of water resources management. This approach has shifted efforts from infrastructure development and operational maintenance to wider management including, for example, all users depending on the same water source. Sanitation often neglected is to be jointly handled 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. Better understand which and how variables impact WSS will help to orient efforts and set priorities for intervention. To do so, the analyses performed previously on WatSan4Dev database (Dondeynaz et al., 2012) are used as a basis for the modelling phase proposed in this paper. WatSan4Dev has been developed by the Joint Research Centre (JRC) of the European Commission (EC) and contains 42 indicators (called variables) from environmental, socio-economic, governance and financial aid flows data in developing countries. As explained in Dondeynaz et al. (2012), a sub-set of 25 variables is used for analysing WSS behaviour because of correlations observed through multi-variate analyses of the WatSan4Dev database. Description of the data and multi-variate analyses preformed 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. Five country profiles are modelled, and scenarios run.

In Sect. 2, this WatSan4Dev sub-set 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 dataset consists of indicators characterizing socio-economic and environmental status at national scale. Further details and description of WatSan4Dev variables, data pre-processing applied and coherency verification are fully described in Dondeynaz (2012). The next two sections (Sects. 2.1 and 2.2) briefly summarize the variables and the methodologies used to build the WatSan4Dev database and the 25 variables sub-set 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 recognized Non-Governmental Organisations such as Transparency International. The national scale is chosen as most of the data are mostly supplied at this level. Data set for 2004 is used in this work because the last release of the Joint Monitoring Programme (JMP) report on WSS access level at the beginning of this research. Building on the several country profiles defined during the analytical phase of the WatSan4Dev sub set, 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 – Poverty (HDP) and Official Development Aid (ODA CI). The reduction of the variables to 25 results from preliminary multivariate analyses:

HESSD
10, 2481–2548, 2013

colinearity between variables or no significant correlations with any other variables constitute the exclusion criteria of the variable subset. Figure 1 outlines the organisation and main correlations between selected variables.

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 imputation[3] and normalized using standard normalization methods, namely square-roots, logarithm and linear regression (Dondeynaz et al., 2012). The WatSan4Dev dataset 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. Dondeynaz et al. (2012) uses the Principal Component Analysis (PCA) and Factor analysis (FA) for analysing relationships and variables clustering (Fig. 1). In this paper, the Bayesian models are built on theses previous multi-variate analyses and conclusions.

3 Methodology

3.1 Description of Bayesian networks

Bayesian networks are statistical tools generated 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

[3]Software used: Amelia II by Honaker, King and Blackwell.

HESSD
10, 2481–2548, 2013

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 *directed acyclic graph* or DAG, is defined a 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 directionality concept implying that the edges or links between variables are directed. For instance, if $(A,B) \in E$ but $(B,A) \notin E$, means that 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 networks 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 evidences within the model. If considered three variables or 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 possible probabilities update in the three basic types of sub-structures (converging, diverging and serial) in 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-known as *bayesian*) (Cox, 1946; Cowell et al., 1999; Dixon, 1964;

HESSD
10, 2481–2548, 2013

Heckerman, 1995). It considers that probability is just a *degree of belief* as 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 parametrically 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's node states. These probabilities or parameters are the materials on which Bayesian theorem works.

3. 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 evidences related to another linked event. In its simplest form, Bayesian theorem can be formulated as follow:

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

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

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

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

3.1.3 Advantages and limits of Bayesian networks

There are several reasons to use Bayesian network modelling in environmental sciences: (1) the most important advantage is that Bayesian networks are able to differentiate and efficiently manage the quantitative and qualitative dimensions of a problem (i.e. Edwards, 1998; Heckerman, 1995); (2) Bayesian networks can optimally manage the presence of missing data (Nadkarni and Shenoy, 2004; Jansen et al., 2003). This is especially important in datasets related to developing countries where the usual issue is the proportion of missing data. Additionally, Bayesian networks can combine prior knowledge with empirical data to develop models (Nadkarny and Shenoy, 2004) and interaction effects. Non-linear relations are modelled efficiently (Lee et al., 2005); (3) Bayesian networks carry out *local computations* in order to avoid computing the whole joint distribution in a model which makes the inference task a relative low demanding process (Pearl, 2001; Xiang, 2002).

However, Aguilera et al. (2011) have noted that there are several drawbacks associated with the use of Bayesian networks in the field of environmental sciences: (1) the building process might be challenged when the ratio between cases and variables is low; (2) although Bayesian networks can handle hybrid data bases (containing continuous and discrete variables) solutions are not yet available in commercial software. Additionally, (3) problems to model time series and fuzzy-based models have been also noted.

3.2 Bayesian Networks: processing steps

This section describes the processing four steps necessary to model our sub-set with Bayesian Networks method:

1. *Variables discretisation:* The variables are discretised using the *k*-mean method that is described in 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 dataset. The objective is to create three

HESSD
10, 2481–2548, 2013

categories or levels (low, medium, high) matching with the qualitative nature of the selected variables. The Appendix A indicates the limits of each class for each variable used in this paper. Water supply, Sanitation and WGI's variables only show two levels (high and low), because of the shape of their distribution which increases the error classification 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, that includes pillar's sub-variables, is categorised in three levels using k-means clustering method as for the variables.

3. *DAG design*: PCA/FA are applied to the 25 variables to understand the organisation of this sub dataset. PCA/FA results, correlations and country profiles analysis are input to design links and structure of the DAG.

4. *Beliefs computation*: Bayesian theorem is applied and probabilities are computed for each variable and each level (HIGH, MID, LOW) according to the input dataset (101 developing countries) or country observations of a specific profile.

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

4 Modelling WatSan4Dev sub-set

The dataset modelling is done in two mains steps: (1) integrate composite indicator for each pillar in a global model; (2) create a model for each pillar. Separated models are built for Water Supply and Sanitation. The global models propose a synthesis of the mechanisms behind a level of WSS while the modelling by pillars downscales the analyses into a specific dimension.

HESSD
10, 2481–2548, 2013

4.1 General models: structure, statistical validation and scenarios

Statistical validation is carried out[4] and measured by the error rate, logarithmic loss (LL), quadratic loss (QL) and spherical payoff (SP) coefficients (Pearl, 1991). 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 (WSS) due to a finding in any other variable of the model (Pearl, 1988). Its value varies between zero (meaning complete independence between the query and the instantiation 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 strongest the independence between the query and instantiation variable).

The global model is based on composite indicators deduced from the previously defined pillars of variables: Human Development-Poverty (HDP), Water Resources (WR), Human activity pressure on WR (AP), Official Development Aid (ODA CI) and Country Environmental Concern (CEC). It is to be noted that WR variables as well as ODA CI are few correlated with the WSS variables. Therefore, WR variables are set with an indirect linked via AP pillar with WS and S dependent variables (Fig. 2). The reason stands in the existing correlation between the amount of water resources available (TWRR) and water use intensity in agriculture (Water Use Int Agri).

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

As for, financial aid flows (ODA CI integrating variables ODA, ODA WSS and WGI PS AV) are included in the models following clues given by previous PCA analysis: (i) ODA shows higher positive correlations with poverty variables than with development variables. It also 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); (iii) WGI PSAV is correlated positively with other governance variables 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).

4.1.1 Water Supply (WS) general model

The general model classifies correctly 86 % of the WatSan4Dev sub-set (Table 1a). Figure 2 presents the Direct Acyclic Graph (DAG) of this model.

The sensitivity analysis provides the influence of the pillars on WS variable (Table 2). The main variable influencing the most the WS access level is the development of the country (HDP). Poverty reduction, as a component of the HDP pillar, is a key element to increase access to an improved water supply. This is in line with the current efforts made towards the Millennium Goals that pushes forward countries as well as the international community to halve the poverty worldwide. WR is not included in the sensitivity analysis because of indirect link with WS.

The Table 3 presents the initial probabilities of each variables of the model for a specific level (column Initial probabilities). For each simulation, a specific variable is set at 100 % probability for simulation an important increment, and then the model computes the impact on the other variables. The results of simulations variables are reported in Table 3. In green are represented the positive increments of variables (recomputed probability minus initial probability) and in red the negative increment. The estimated WS level is computed and reported for each simulation. Simulations are ordered by decreasing range of variable impact (HDP is the highest, ODA-CI the lowest) on WS.

HESSD
10, 2481–2548, 2013

Simulation 1 (HDP set to 100 %) directly concerns the impact of poverty reduction efforts. Simulation 2 (CEC set to 100 %) estimates the influence of an environmental concern context. Simulation 3 examines the impact of the type and level of pressure of human activity (water demand according to sectors) on development of WSS. Simulation 4 (ODA CI set to 100 %) observes Official Development Aid 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 reaches almost 85 %, thus a very positive impact. Various mechanisms can be suggested to explain it:

- In a context of economic development, households tend to dedicate financial resources to improve WS as soon as this is possible.
- Governments engage WSS program as part of the strategy of poverty reduction, MDG's target, thus look for reducing diseases, children mortality, ... Associated to the state's efforts, the international donors finance programs fighting against poverty, and mobilize resources for economic development and health purposes. In 2004, all DAC Donors committed 3 billion and disbursed around 2.1 billion $ in Official Development Aid dedicated to WSS improvement (OECD, 2007).

Note that in case of unequal distribution of wealth, the impact of development on WS access is a priori undermined. Such indicator could not be integrated into models. Indeed, the Gini index that measures the inequality of income at country scale showed too many missing values when considering developing countries.

HESSD
10, 2481–2548, 2013

Simulation WS2 (CEC set to 100 %): the impact of environmental concern on WS

In this simulation, the increment of environmental concern (CEC) implies the WS increment up to 70.3 %. CEC expresses the country commitment toward major environmental issues (Particip). The degree of civil society freedom of expression and the level of accountability of government (WGI VA) is also included. The bigger the increment of CEC, the more the country is committed toward the environment and citizens. In that case, CEC constitutes a favourable framework: "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" (Buntaine and Parks, 2010). In details, the provision of safe water and sanitation facilities implies a concern but also capacities for environmental management (soils-land management, wet lands preservation, infrastructure norms in particular for sanitation and waste water treatment ...). In addition to government strong commitment and capacities, civil society may play a role of alert and control (see Sect. 4.2.2).

Simulation WS3 (AP set to 100 %): the role of human activities

According to the model, the increment of activity pressure on WR (AP) represented an increment of WS access up to 66 %. AP composite index should be interpreted as follow: the bigger the AP factor increment, the more the agriculture monopolises water resources available often through irrigation practises. The performance of agriculture is an essential element in defining its impact on WS access. Agriculture is the main consumer of 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 (network supply). In case of irrigated agriculture, sustaining water supply for agriculture is a priority and the existing infrastructure is maintained by the population. Thanks to multi-purpose infrastructure, the water supply facility is likely to be used for domestic use. Irrigation

HESSD
10, 2481–2548, 2013

practises (above 20 % 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, CEC and WS

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

In 2004, investments in social infrastructure and services (government and civil society support, education, and health issues) represented 37 % of the Official Development Aid provided by all DAC donors and EU (OECD statistics of 2004[5]). Humanitarian, emergency aid is excluded from ODA. The model shows that poor countries (with Low HDP) received high Official development Aid at 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 still 21 % of the countries that benefit from relatively high ODA CI show either high HDP. This may raise questions about the motivations behind ODA delivery. Over the last decades, it has been quite polemic stemming from donor's visions and strategies of cooperation (Carbonnier, 2010). In addition, excessive dependence to 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); elements included in HDP variable. The ways in which external aid is delivered may complement the explanation of the ODA CI, HDP, WS observed behaviour. For instance, donor fragmentation combined with the low capacities of a country to fund implementation can favour corruption practises.

[5] OECD: http://stats.oecd.org/qwids/ (last access: 15 February 2013).

Concerning other governance aspects, the model estimates an increment of high CEC from 31 to 34.3 % with an increment of ODA CI; the majority of countries (43.9 %) show a medium CEC level.

Several hypotheses can be formulated to understand this link:

- Programs/projects supported by external aid include more 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) ensure financial efficiency of the funds and infrastructure sustainable implementation. In fact, in the last decades, the experience in cooperation has shown that ODA, and by extension ODA WSS, hardly provides results with unstable public authorities or corruption spread at all scales and generally, poor governance (Knack, 1999). 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 than WS (Fig. 3). Model performance reaches up to 90 % of correct classification (Table 1b).

Table 4 shows the sensitivity analysis of the variables involved in Sanitation. Variables/pillars are ordered by level of impact on S. Table 5 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 first sensitive to human development (HDP). Therefore the improvement of living conditions through the reduction of informal urbanisation, the improvement of health and education, balanced economic development ... Commitment toward the environment and governance are favourable framework for sanitation services development. Financial external aid (ODA CI) is mainly (44.5 %) delivered in countries with low HDP. The main difference stands in the relative more influence of human activity pressure variable (AP), the role of agriculture is playing in this variable.

Agriculture and sanitation can develop synergies like, for instance, the reuse of human excreta for soils fertilisation (i.e eco-sanitation detailed by Jenssen, 2004; SIDA, 2004) Si. With little expenses, crops benefits from fertilizer at individual or small scale. The sanitation products can be collected and transformed in fertilizer at affordable price with composting toilets or by solar drying process. This opportunity of valorisation of sanitation/wastewater products in agriculture is possible both in rural and urban areas. This is easier in rural areas where a simple dry latrine allows composting human excreta. Regarding urban areas, wastewater sludge, if treated correctly for instance by solar drying, can be reused as well for agriculture. However, it is difficult to estimate the extensions of such practices because statistics regarding sanitation/wastewater products reuse in developing countries are barely available. A linked issue is the cultural and psychological barrier related to the reuse of human excreta or wastewater products that counterbalances the potential positive impact of agriculture in WSS. Indicators that measure such aspects are currently missing at country scale and should be include in the models.

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 downscaled and sub-models for each pillar are described.

HESSD
10, 2481–2548, 2013

4.2.1 Human Development and Poverty model (HDP)

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

Table 6 shows the sensitivity analysis computed for the HDP model. The three main variables influencing HDP are the income per capita, fighting efficiently against mortality (mainly children mortality under 5 yr) and an organised urbanisation (fighting against slums). These are the variables where authorities and/or donors should focus their efforts to foster development. Other important aspects are school enrolment and malaria.

Table 7 sums up the probabilities computed according to several scenarios presented in this paper. Simulation 1 estimates what impact the reduction of children mortality under 5 yr would have on HDP. Simulation 2 estimates the improvement of HDP in case of significant reduction of informal urbanisation (% slums). Simulation 3 observes the control of malaria and how it affects human development. Simulation 4 observes the benefit of good governance conditions on HDP. Simulation 5 examines the impact of an increase of education (from primary level to university). Finally, simulation 6 examines the role of gender issue through girl's primary education (School G/B) and female economic participation (Femal eco). The order of variable impact is different between sensitivity analysis and the simulations table. The sensitivity analyses reports the cumulated impact of each variable on HDP across all categories (high, mid, low). The simulations table reports the probability variations for a specific category of the variable.

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

According to the model, if the children mortality (Child Mortal-5) is reduced by increasing efforts on basic health care, HDP could grow up to 53.6 %. The main link with

WSS stands in the fact that fighting children mortality supposes better access to safe drinking water together with basic health care and quality of nutrition. The incidence of waterborne diseases such as diarrhoea is considerably reduced and thus the risk of mortality in particular young children. Pneumonia and Diaorrhea are the two first causes of young children mortality in developing countries: diarrhoeal disease kills 1.5 million children every year (UNICEF, 2009). South East Asian and Africa remain most affected by diarrhoea and malaria.

Simulation HDP2 (low % slums set to 100 %): the role of urbanization 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 could rise up to 52 %. The development of informal urbanisation worsens living conditions including WSS provision and thus their limitation/reduction represents a good leverage. The consequences of lack of sanitation take even more amplitude in term of health and water contamination because of the density of the population. The limitation of slums extension just from high to medium level implies an important improvement of living conditions. In such case, countries with low level of HDP would decrease from 91.2 % to 38.3 % probability. This also impacts positively essential variables: school enrolment, the household income (GDP per cap), children mortality rate and malaria prevalence.

The level of urbanization (Urban pop) presents a negative correlation (−0.650) with urban slums (% slums) accordingly to the dynamic of the rural-urban population transition. World urbanisation prospects (UNDP, 2001) analyse urban and rural population growths providing detailed information on this transition at regional level. Before the beginning of this transition, the population lives in majority in rural areas with agriculture as principal activity and children mortality rate relatively high. For economic and demographic reasons, the population starts moving from rural areas to urban centres. Therefore at this moment, the urban population represents proportionally a small

HESSD
10, 2481–2548, 2013

percentage but its growth is important. The population migrating to cities often feeds informal settlements around urban areas therefore the increment of slums. Progressively, the proportion of urban population increases, urban development should be more and more organised and thus slums proportion is supposed to be reduced. This process ends up with the majority of 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 counties present low urban population rate while the level of population living in slums is high in 34 % of cases. Despite potential slums development, urbanisation represents an opportunity for development of WSS services as the concentration of the population offers the necessary critical mass for collective infrastructure and improves cost effectiveness.

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

The reduction of malaria prevalence (Malaria) contributes to the increment HDP up to 47.7 %. According to the CDC[6], a majority of the countries included in the WatSan4Dev database is concerned by the malaria either on the whole or part of their territory. Being a disease spreading through mosquitos, lack of water supply infrastructure leading to outdoors water storage facilitates mosquito's development and thus malaria's dissemination (Lewis, 2011). Limiting inappropriate storages are essential and avoided in 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 mosquitos by reducing favourable conditions. Improving sanitation, and by extension wastewater collection, reduces stagnant and pathogenic water areas. Indeed, in case of pipe infrastructure "when population growth outpaces the existing infrastructure, wastewater treatment systems are unable to cope with the influx, garbage and sanitation facilities

[6]CDC malaria map: http://cdc-malaria.ncsa.uiuc.edu/

cannot contain 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" (Sergo, 2007).

Simulation HDP4 (WGI GE set to 100 %): governance as a favourable framework

In case of simulation of high government effectiveness (WGI-GE), HDP increases to 44.5 %, multiplying by more than two the initial probability. In fact, the capability of a nation to organize public services, to be committed towards implementation of suitable policies, and to fight corruption (as represented by the WGI-GE) bodes well for the WSS delivery in a country. Butaine and Parks (2012) demonstrated that the governance effectiveness has a strong and positive relationship with the success of environmental projects, including WSS. It should be mentioned that description of corruption mechanisms, for example by Transparency International or UNDP, are not breakdown by WSS sub-sectors; some further details or case studies are available for water supply while sanitation is often enclosed in WSS or neglected. Therefore, it is hardly possible to highlight reasons on different potential impacts of governance on Sanitation versus Water supply.

Simulation HDP5 (School enrol set to 100 %): the influence of School enrolment on WSS

In this simulation, if school enrolment (School enrol) is fostered, HDP is enhanced and reaches 44.3 % probability. Beyond the HDP improvement, several hypotheses on the relationship with WSS can be depicted:

1. The access to improved water supply reduces the time need to fetch water and improves health conditions that allow a better assiduity to school. Water availability in households is an important factor in the enrolment, attendance and drop-out rates of children. This is most often seen in African countries with low water coverage rates, but the correlation has also been documented in countries in Asia, the

Middle East and Latin America. When water access is improved in communities, school enrolment rates increased in Tanzania, India, Bangladesh by up to 15 %. 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). The equipment of schools with improved water and Sanitation facilities is a more direct and simple expression of this link between education and WSS.

2. The increase of school enrolment indicates a better education of the population and thus better individual opportunities on the labour market. This access to a better position implies a better income that may also be accompanied by a displacement toward equipped urban areas, leading to an increased access to WS. In fact, it is easily assumed that with an increase of income, households will dedicate part of it to ensure its provision in safe water. The model proposed in this paper shows a positive impact of school enrolment on the income per capita (GDP per cap) and vice et versa confirming this process.

3. The level of school enrolment (from primary to university degree) is an expression of the level of competences available on the labour market. WSS sector needs technical, high educated professionals that the education system should provide. Adequacy between skills required and qualified people available on the labour market in developing countries is often an important constraint to functioning WSS services.

Simulation HDP6 (School G/B set to 100 %): the role of gender in WSS (ratio girls to boys at school and female economic activity analysis)

According to the model, the improvement of the girls'schooling rate (School G/B) increments HDP up to 37.9 %. Limited female economic participation (Femal eco) indicates a rather good HDP (high HDP at 36.6 % probability). Several hypotheses can support these observed relationships:

1. Girls in the developing world are generally in charge of the daily water fetching setting first time barriers to school attendance. In this view, girl's better attendance is rather a consequence of the improvement of the WSS pushing for human development of the country. However improving girl attendance at school is an important incentive for schools to get WSS facilities. In the same way, teachers are easier to recruit and retain, if good water and sanitation is provided (WSSCC, 2006). Separated sanitation facilities for girls facilitate school attendance of girls 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).

2. In a context of poverty, women are massively active in the economy to sustain their family, translated in statistics by high economic activity rate. With the increment of household's income, this rate drops until 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" (WSSCC, 2006). Therefore, the level of women activity remains rather the same, but the benefit improves women daily lives.

4.2.2 Country Environmental Concern model (CEC)

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

HESSD
10, 2481–2548, 2013

Table 8 shows the variables ordered by level of importance in the model: 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 9).

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

In this simulation, high international commitment (Particip) of countries increments CEC that reaches 70.1 %. Signing environmental agreements (Particip) formally engage countries towards global environment protection: 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), Basel Convention on the Transboundary Movement of Hazardous Waste. A country worried by environmental issues is assumed to increase its commitment towards water resources management in order to ensure their sustainability. This approach includes the extension and reinforcement of WSS infrastructure and above all of its management.

However, this potential commitment toward WSS may be handled differently because of difference in priority and perception between WS and S. Safe drinking water supply is directly touching human survival hence, is a priority when sanitation is often perceived as much secondary. Sanitation is generally neglected by governments/public institutions even being a good leverage for the improvement of livelihood, and health. With the launch of "International Year for Sanitation 2008", the UN tried to push countries to fight the lack of sanitation and to make progress towards the MDG's. In 2012,

HESSD
10, 2481–2548, 2013

the Sanitation target is still not reached. This mechanism is supported by the global models: CEC is more important in WS than in S model (see Tables 2 and 4).

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

The model estimates that the improvement of degree of civil society freedom (WGI VA) pushes CEC from 30.8 to 54.4 %. A country in which citizen are able to express their criticisms and pressure their government according to their interest is a favourable framework for an efficient environmental management.

As mentioned in Dondeynaz et al. (2012), the voice of citizens both directly or indirectly pressures the government and also economic actors to manage environment sustainably. Indeed, the population is first concerned by lack of good management of water resources, wastes, ecosystems ... and by degradation or pollution impacts. Being in first line, civil society plays a major role in control, alert and pushes for sustainable management of the environment if allowed by the institutional and political context. Civil society is also less sensitive to industrial-economic interest and is expected to push more for environmental policy making than other stakeholders.

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 still invested at 21 % in poor environment concern context (low CEC). However, a high ODA CI may also be found in cases of good CEC (34 % of countries). Indeed, minimum governance conditions, political stability or civil society involvement opportunity are appealing conditions for international donors. This behaviour also indicates the attention paid by donors to

HESSD
10, 2481–2548, 2013

support local population involvement and more generally governance aspects (refer to Sect. 4.1.1 Simulation 4).

4.2.3 Activity Pressure on water resources (AP)

The structure of the AP model is set differently according to correlations found out by Principal Component Analysis (Dondeynaz et al., 2012). Municipal and industrial withdrawals variables are not significantly correlated to the WSS variables but are correlated with total withdrawal (Fig. 6). The model shows a performance better than 88 % (Table 1e).

The sensitivity analysis is computed for the four directed-linked variables: Water Use Intensity in Agriculture (Water Use Int Agri), proportion of irrigated areas (% irrigation), the total amount of water withdrawal (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, account for little in WSS (Table 10).

Simulation 1 concerns the impact of agricultural intensive water use, combining the Water use intensity in Agriculture (Water Use Int Agri) and the proportion of irrigated areas (% irrigation). Simulation 2 aims at examining the competition between different uses of water through their respective withdrawal amounts (total withdrawal, withdrawal industrial, withdrawal municipal). Simulation 3 observes the impact of potential reduction of WR available, under pressure of climate change, on economic activity.

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

Agriculture intensity 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 agriculture intensive practices produces an increment AP up to 98.5 %.

The positive impact on WSS can be explained by several hypotheses: (1) the amount of water resources mobilized by such practises underlies the existence of facilities, capacities and financial resources that benefit at the end the WSS and the population; (2) irrigation and intensive agricultural practices allow multi-uses of water supply infrastructure developed both for crops and for municipal supply; and, (3) irrigation schemes, even small, structure rural areas and facilitate organisation of WSS services.

However, only 23 % of the countries considered present an irrigation rate higher than 10 % mainly in South East Asia, Middle East and some Maghreb countries. 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 yield. In this context, the population is poor and live in rural areas where WSS access is generally low (Low WSS is estimated around 33 % of probability despite high AP). Dissemination or isolation of population can reinforce the effect of barrier to WSS delivery.

Simulation AP2 (total withdrawal set to 100 %): the analysis of the water withdrawals according uses

The AP model estimates that an increment of total water withdrawal induces an increment of AP up to 48.3 %; the municipal withdrawal variable decreases to 24.6 % and industrial withdrawal slightly goes up to 26.5 %. This means that industrial variable has 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 behavior of the total withdrawal 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 withdrawal (water amount per capita) are negatively correlated (respectively −0.553 and −0.251). Looking at WatSan4Dev database, the majority of the developing countries (80 %) dedicate more than 50 % of water available to agricultural purpose with three main profiles: (i) 35 % of the countries that dedicated water almost exclusively

HESSD
10, 2481–2548, 2013

to agriculture (above 85 %); (ii) 40 % of the countries dedicate the majority of water resources to agriculture but nibbled step by step by the household consumption (corresponding to classification at AP medium level); (iii) 25 % of countries present substantial industrial/mining/forestry activities higher than 10 % often combined with high municipal consumption, reducing significantly the share dedicated to agriculture. This is the case, for instance, of Lesotho with 40 % or Papua New Guinea with 42 % water withdrawals due to mining activities. Therefore, the total withdrawal is mainly lead by water agriculture demand, that competes with municipal and industrial demands altogether.

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

According to the AP model, the reduction of water resources due to climate change, for instance, slightly decreases high activity pressure (AP) probability from 33.3 to 31.9 % but, at the same time, low human activity pressure probability decreases from 26 to 18.8 % (−27 %). This suggests that independently of the amount of WR, countries can find adapted solutions to mitigate the consequences of declining resources. This mechanism may be also confirmed by input observation data: arid countries such as 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 adapted infrastructure to water scarcity such as desalination plant.

Irrigation can be an adapted solution to increase the water use efficiency in agriculture. In that case, the limiting factor in such development is rather political commitment, policies, investments, agronomic and management skills than the amount of water resources. In that sense, FAO published recommendations encouraging the reinforcement of policies for planning and infrastructure development including social aspects, capacities building towards staff and farmers (FAO, 1987). Kweku Baah Inkoo and Zomanaa Nangu (2011) analysed the reasons of under-utilization of irrigation infrastructure in the upper west region of Ghana. They highlighted the importance of

HESSD
10, 2481–2548, 2013

the non-environmental factors such as lack of ownership, management and skills. Irrigation development and the intensity use of water resources result of rather complex processes and factors, even if water scarcity or desertification may be a starting point for such intensive practise in agriculture.

Note that indicators on the quality of resources are not included on the database and associated models because of being not available at country scale (Dondeynaz et al., 2012). Such water quality indicator will be affected by human activities variables; the level of change could highlight the quality of WR management.

Summarising the analyses of the AP model, strong agriculture sector supports WSS development particularly in case of intensive practises. In that case, agriculture participates to rural development and a structuration of these areas. However, this positive effect can be undermined when agriculture monopolises water resources and/or express poor economic development (agriculture generates not enough income above survival). The existence municipal water demand indicates on-going urbanisation processes that can also contribute to WSS development as described in Sect. 4.2.1.

4.3 Modelling country profiles

Dondeynaz et al. (2012) proposed five country profiles that were built around 5 axes corresponding to first five PCA components: (HDP – Human Development against Poverty, AP – human Activity Pressure, WR – Water Resources, ODA CI – Official development Aid, 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 term 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 when considered human development (HDP). However, profile 4 benefits from higher level of freedom (CEC), political stability and high external Aid support. Profile

HESSD
10, 2481–2548, 2013

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. It finally allows measuring mechanisms specific to a group of countries and running scenarios. The statistical performance of country profile models are greater than 79 % (Table 12).

The model offers a precise picture of the characteristic of each group of countries confirms outlines of each profile and provides quantitative and qualitative measures.

4.3.1 Analysing WSS behaviours across profiles

In a general way, water supply (WS) and sanitation (S) in the WatSan4Dev show a linear behaviour that can be expressed by (Fig. 7):

$$S = 1.0285WS - 21.82 \quad (3)$$

with $R^2 = 0.5265$.

The linear model highlights that S value is below the WS value (−21.82 %) as often observed in data and on the field. However, this linear model only explains 52.65 % (R^2) of data variability. Some examples of an increased gap between WS and S are: Botswana (53 % between WS and S), Nepal (55 %) and Namibia (62 %). On the contrary, Libya in 2004 experienced a negative gap between WS and S (72 % versus 97 %, respectively). These examples cannot be explained by the linear model.

In 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, countries from profiles 4 (62 % for WS and 35 % for S) and profile 5 (65 % for WS and 38 % for S) show on the contrary low average WSS levels.

Countries from the profile 3 follow a different WS and S statistical profile (purple dots on Fig. 7). WSS values in this case are less homogeneous showing higher standard deviations: WS average value is 73 % with 15.2 % standard deviation while S reaches 55 % with 22 % of standard deviation. Profile 3 is built around the AP pillar with the predominant role played by agriculture. WSS services are more or less developed according to the agriculture sector and its organisation/dynamics. In this group, irrigation practises 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 either good access to WS without irrigation use. This can be explained by an on-going urbanisation that favours WS access still in a context of agriculture dominance. 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.

4.3.2 Country profiles: analysis of Profile 4 (essential external a 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 have the less favourable status regarding WSS (Fig. 7).

Model running Profile 4 (essential external support)

For these countries, the sensitivity analyses of the WS model ranges 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 less advantageous situation regarding WSS access than countries from profile 5 (see Sect. 4.3.1). However, they benefit from stable political context that attracts external financial flows: 56 % of countries show a high ODA CI. The environmental concern constitutes also a favourable basis to support efforts in WSS sector: 73 % of these countries show a medium-high level for CEC.

HESSD
10, 2481–2548, 2013

Countries in this profile are first concerned by poverty described by 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 general governance (WGI-GE). Simulation 1 estimates the increment of WS 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 level. In fact, 23 out of 24 countries are in Africa and have part of their territories under desertification treat. All countries overpass 40 % of national territories being potentially affected by desertification in 2004 (FAO statistics). Sustaining adaption or mitigation measures appears essential to fight against this constraint and ensure long term WSS services.

In simulation 3, CEC behaves differently from general model (Sects. 4.1.1 and 4.1.2) regarding S. With the increment of CEC, S drops up to 11.3 % while with the increment of ODA CI, S drops up to 12.5 %. The maximum impact on S is estimated when CEC is at a medium level. In that case, high S becomes 21.7 % and high ODA CI 68 % (note that Table 14 is computed for high CEC category). Therefore, from these computations, no additional increments of S are expected after a certain increment of CEC and ODA-CI. This illustrates the influence of high external aid dependence on the governance aspects as described by Knack (1999). For instance, “aid weakens governmental accountability (WGI –VA) by retarding the development of a healthy ‘civil society’ underpinning democracy and the rule of law” (Knack, 1999). Indeed, external support may short-circuit processes for such accountability development in developing countries by reducing government’s dependence on its citizenry for tax revenues (Moore, 1998; Karl, 1997). The civil society strengthening and role is therefore undermined. More investigation in these specific countries may shed light on this behaviour and why in particular S is concerned.

HESSD
10, 2481–2548, 2013

Model running Profile 5 (Primary material consumption)

According to both sensitivity analyses, WSS variables are influenced in order by HDP, AP, CEC and finally ODA CI. Tables 15 and 16 indicate that major efforts should be put on the three main dimensions (AP, HDP, and CEC) to improve substantially the level of WSS. Simulations are ordered in the same way for WS and S. Indeed, countries from this profile present low access to WS in 75 % of cases while low S concerns 88 %.

Simulation 1 indicates that the increment of AP leads to an increase of WS up to 40.3 % and S up to 29.6 %, respectively. These countries are characterised by water resources abundance (81 % of the countries benefit from medium (48 %) to high (33 %). However, an important part of these WR are mainly used for industrial and raw material exploitation (low values in AP) and less dedicated to agriculture (below 50 % water withdrawals). The natural resource exports have counteracting effects on growth, by weakening the manufacturing and agricultural sectors (Knack, 1999). This type of activities undermines at the end, the development of WSS infrastructure for households (see Sects. 4.1 and 4.2.3). Sustaining agricultural sector in these countries, or at least, diversifying activities can be a good leverage for WSS services development.

Simulation 2 shows that an increment of HDP produces an increment of WS up to 40.2 % and S up to 30.9 %, respectively. Poverty in these countries is a crucial issue that includes fighting against, high demography, informal urban settle development, and poor general health in particular for children, as well as poor average education. Governance is globally weak (WGI-GE included in HDP) 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) is often 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.

HESSD
10, 2481–2548, 2013

The improvement of CEC will have a positive impact on WSS as shown in simulation 3; high WS reaches 40.2 % and S 30.9 %. As already mentioned, country environmental concern is a favourable context fostering WSS management. Environmental management appears even more essential in an economy focused on natural resources exploitation. In addition, the increment of CEC also supposes efforts to improve civil society freedom and government accountability to citizens.

5 Discussions

The models proposed appear as flexible tools to measure interactions between the 25 variables that influence the WSS access in developing countries (101 countries). These models highlight key variables and mechanisms pushing for WSS either for all developing countries but also specific profiles of countries. In addition, the Bayesian Network method used for modelling allows running probabilistic scenarios for better understanding the impact on WSS.

Concerning the issue related to the development of Worldwide Governance indicators (WGI), Thomas (2008) argues that the governance data are based on expert assessments often subjective, indicators are spotty, with standards error issues and thus, non-extendable and comparable between countries. Langbein and Knack (2010) questions the number of indices created to measure governance, and justifies the number of distinct governance variables. They call for caution in particular in making country evaluation using these indicators. Thomas (2008) also noted that "evidence must be provided to show that a purported measure of a theoretical construct is valid both in its conceptualization and its operationalization, by exploring predicted relationships with other observable variables". Kaufman et al. (2009) responds that "construct validity is not a useful tool to assess the merits of the WGI, and even if it were, Thomas provides no evidence of any practical consequences of failure to meet the criteria of construct validity".

HESSD
10, 2481–2548, 2013

Despite these limits and criticisms, the authors' point of view is that measuring governance is relatively new and subject to caution since 90's. This paper proposes the use of the WGIs but verifies their coherency with other indicators provided by others institutions, namely environmental governance (sub-indicator 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). Even if 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 relevant indicators are selected according to their correlations: (1) government effectiveness (WGI-GE) that represents also corruption control (WGI-CoC), rule of law (WGI-RofL), regulatory quality (WGI-RQ) and environmental governance (env gov) (2) political stability and absence of violence (WGI PS-AV) and (3) voice and accountability (WGI-VA) (Dondeynaz et al., 2012).

Moreover, it is to be noted that, for instance, the corruption is measured at scale of the water sector as a whole often including water supply and sanitation, agriculture and hydropower. Corruption analysis is done rather at different scales than at sub-sector level. The disaggregation of data by sub-sectors, specific analyses and/or case studies, in particular related to sanitation, are rare at national country scale. Description of corruption mechanisms (i.e. those proposed by Transparency International, 2008) are declined for the whole water sector, with specific information in case of water supply, sanitation being enclosed in WSS or neglected. Such breakdown or at minimum, compared case studies on corruption mechanisms in both sub-sectors could help in making more accurate interpretations.

Other 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. This missing aspect in characterizing water resources

HESSD
10, 2481–2548, 2013

could explain the negligible weight of water resources in the WSS models proposed in this paper. Such indicators at national scale could allow modelling the interactions between environment and development through the analysis of the qualitative impacts on water resources. This would be particularly useful in case of sanitation. Lack of adequate sanitation services is known to impact water quality and thus health (i.e. high rates of children mortality in sub-Saharan Africa).

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 of this gap stands in cultural and psychological aspects linked to Sanitation. Indeed, the "Sanitation for all" year (2008) highlights its "unpopularity" leading to a "sanitation crisis"[7]. Improved sanitation is often not considered as necessary/vital improving the quality of life. However, understanding these cultural and psychological factors is essential to allow its sustainable development. Up to now, there is no quantitative or qualitative means to appreciate and measure these effects. In many countries, awareness campaigns have been proven as a crucial leverage to convince population to implement, and keep functioning sanitation facilities. Community-led total sanitation approach[8] developed and applied in Bangladesh since 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 cross-interaction between multiple factors and issues. Bayesian Network method is chosen to model the WatSan4Dev sub-dataset that includes 25 indicators and five dimensions in WSS (Dondeynaz et al., 2012). Indeed, Bayesian networks present the advantage of providing probabilistic causal dependence

[7]International Year of Sanitation 2008: http://esa.un.org/iys/docs/IYS_flagship_web_small.pdf
[8]http://www.communityledtotalsanitation.org/

HESSD
10, 2481–2548, 2013

relationships between variables. The proposed models in this paper are efficient tools to identify and measure probabilistic impact of key elements on WSS levels (statistical performances range between 80–95 %).

In a first step, general models were built separately for Water Supply and Sanitation to observe and identify potential differences. Therefore, composite indicators are computed for the following dimensions/pillars defined in Dondeynaz et al. (2012): Human Development-Poverty (HDP), Country Environmental Concern (CEC), Activity Pressure on water resources (AP), Official Development Aid flows (ODA CI) and Water Resources availability (WR). These models synthesise large mechanisms involved in WSS access existing in the common literature in the domain. The statistical classification error rates are less than 12 % for Water Supply (WS) and 10 % for Sanitation (S).

The priorities set by the MDG's initiative in reducing poverty are first comforted by the key findings from this modelling exercise (Table 17). Indeed, general models confirm that alleviation of poverty is the principal leverage to improve WSS access. In addition, external aid (ODA CI) is provided preferentially to the poorest countries, but not exclusively, and has an impact on WSS sector. Funds are also invested in relatively less poverty-concerned countries (still not considered developed) probably because of the geo-strategy policy of donors.

The development of irrigated agriculture and the emergence of urban activities are also favourable to WSS development. Models put in advance that Sanitation benefits more from the development of human activities in particular from an efficient agriculture (i.e. irrigation) than Water Supply. This is also suggested by field's experience: fertilisers based on the reuse of human excreta or/and wastewater treatment sludge can be an integrated solution for rural development once the cultural/psychological barrier is overcome. On another side, urbanisation processes bring incentives and the critical mass necessary for WSS infrastructure development and costs recovery.

Finally, the models highlight that water availability (WR) counts for little in the population access to WSS. This link would be certainly improved if indicators related to water

HESSD
10, 2481–2548, 2013

quality exist. This work suggests that water quality is an essential element that international institutions in charge of MDGs monitoring should take into account for linking sanitation to WR but also, human economic activities and the management of WR.

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

1. *Human Development-Poverty model (HDP):* key variables are the household income (GDP per cap), urbanisation processes where the reduction of informal settlement is crucial (urban pop, %slums, Poverty) and finally health improvement (children mortality and malaria). The urbanization process facilitates the WSS access except in case of informal urban settlement development. It also emphasizes the importance of the control of disease (malaria) and mortality of children. Governance provides essential framework for sustainable WSS services delivery. Education of population facilitates WSS development through the creation of specific demand but also through the provision of necessary skills to maintain services. Gender issue would also be such upstream element enhancing WSS.

2. *Country Environmental Concern model (CEC)*: in both cases, the environmental country commitment in front of the international community (Particip) rather indicates a base strategy towards WSS improvement. This environmental concern is supported by the civil society if an enabling context exists. This framework is attractive to external donors (ODA CI).

3. *Activity Pressure on Water Resources model (AP)* includes agriculture intensity, water demand and the water resources availability. This pillar highlights synergies with the development of the economic activities: organised agriculture sector allows a rural development and facilitates access to WSS services. The availability of water resources (WR) has a little influence on the pattern of activity development as well as the level of water demand.

Table 17 summarises the potential probability changes of WS, S and composites indicators according to the increments of 25 selected variables.

HESSD
10, 2481–2548, 2013

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 to run scenarios for the different profiles. In a general way, S access mainly steps behind WS (−21 % in average) (Fig. 7). Profile 1 (towards well-being) and 2 (freedom/democracy black spot) are the more advanced countries while profile 3 (agricultural economy) shows spread 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 lowest levels in developing countries. In the limit of the extension of the paper, groups of countries from profile 4 and 5 are specifically analysed because being the less advanced countries among the subset considered.

Countries belonging to profile 4 experiment desertification effects in a context of limited water resources availability (66 % with low WR). Therefore, the development of adapted or mitigation measures is crucial. Benefiting from minimum governance (political stability and civil society freedom) and concern about environment, the main effort should be oriented to a development of HDP pillar meaning: control of urbanisation, organisation-reinforcement of health care, education, and improve "advanced governance" related to state effectiveness (corruption control, make and rule laws). The model identifies a different behaviour for environmental commitment (CEC) impact on S. The model clearly proves that after a certain CEC level no additional improvement on S is expected. More investigation in these specific countries may shed light on this behaviour and why only Sanitation is concerned.

Profile 5 shows weaknesses on almost all dimensions, having low HDP (68–70 %), low CEC (48–50 %), low values in AP (64–67 %) indicating pressure from industrial activities and limited development of agriculture. Despite this situation, these countries benefit from a medium level of investment by donors (mainly medium level ODA CI involves 57 % of countries) that can be often explained by the persistence of political instability. Therefore, ensuring basic governance conditions, stability and absence of

HESSD
10, 2481–2548, 2013

violence and accountability to citizen environment is essential: (1) to foster economic development and reinforce rural development; (2) to sustain efforts made for poverty reduction; and (3) to appeal for additional external support if necessary.

Acknowledgement. The present work benefited from the input from the various international institutions mentioned in this paper that provide the necessary datasets 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 AP 2009/220-992, respectively) component: stakeholder mapping and participation. The comments of the reviewers are greatly appreciated.

References

Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., and Salmerón, A.: Bayesian networks in environmental modelling, Environ. Model. Softw., 26, 1376–1388, doi:10.1016/j.envsoft.2011.06.004, 2011.

Bayes, T.: An essay towards solving a problem in the doctrine of chances, Philos. Transact., 53, 370–418, doi:10.1098/rstl.1763.0053, 1763.

Boserup, E.: Population, the status of women and rural development, Populat. Develop. Rev., 15, 45–60, 1989.

Burnside, C. and Dollar, D.: Aid, policies, and growth, Am. Econ. Rev., 90, 847–868, doi:10.1257/aer.90.4.847, 2000.

Butaine, M. T. and Parks, B. C.: When Do Environmentally-Focused Aid Projects Achieve their Objectives? Evidence from World Bank Post-Project Evaluations, http://irtheoryandpractice.wm.edu/bcparks/publications/achieve.pdf, last access: 27 February 2013, 40 pp., 2012.

Carbonnier, G.: Official development assistance once more under fire from critics, Rev. Int. Pol. Develop., 1, 137–142, 2010.

Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J.: Probabilistic Networks and Expert Systems, Springer, Harrisonburg, VA, 1999.

Cox, R. T.: Probability, frequency and reasonable expectation, Am. J. Phys., 14, 1–13, 1946.

Dixon, J. R.: A Programmed Introduction to Probability, John Wiley & Sons, New York, 1964.

Dondeynaz, C., Carmona Moreno, C., and Céspedes Lorente, J. J.: Analysing inter-relationships among water, governance, human development variables in developing countries: WatSan4Dev database coherency analysis, Hydrol. Earth Syst. Sci. Discuss., 9, 485–539, doi:10.5194/hessd-9-485-2012, 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, in: Proceeding of Consultation on Irrigation in Africa, Lome, Togo, 1987.
Heckerman, D.: A tutorial on learning with bayesian networks (Tech. Rep. MS-TR-95-06), Microsoft Research, Redmon, WA, 1995.
Hoeting, J. A., Madigan, D., and Raftery, A. E.: Bayesian model averaging: a tutorial, Statist. Sci., 14, 382–401, 1999.
Honaker, J., King, G., and Blackwell, M., Amelia, I. I.: a program for missing Data, Version 1.2-17, available at: http://r.iq.harvard.edu/docs/amelia/amelia.pdf (last access: 27 February 2013), 2011.
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-protein interactions from genomic data, Science, 302, 449–453, 2003.
Jenssen, P. D., Heeb, J., Huba-Mang, E., Gnanakan, K., Warner, W. S., Refsgaard, K., Stenström, T.-A., Guterstam, B., and Werner Alsén, K.: Ecological Sanitation and Reuse of Wastewater ecosan, A Thinkpiece on Ecological Sanitation, Norway, available at: http://www.iees.ch/pdf_files/en-thinkpiece-2004.pdf (last access: 27 February 2013), 2004.
Karl, T. L.: The Paradox of Plenty: Oil Booms and Petro-States, University of California Press, Berkeley, USA, 380 pp., 1997.
Knack, S.: Aid Dependence and the Quality of Governance: A Cross-Country Empirical Analysis, World Bank Publications, Washington, DC, USA, 38 pp., doi:10.1596/1813-9450-2396, 1999
Kweku Baah Inkoom, D. and Zomanaa Nanguo, C.: Utilisation of irrigation facilities towards poverty reduction, available at: http://www.jsd-africa.com/Jsda/V13No2_Spring2011_B/PDF/Utilisation%20of%20irrigation%20facilities.pdf, last access: 27 February 2013, J. Sust. Develop., 13, 335–351, 2011.
Langbein, L. and Knack, S.: The worldwide governance indicators: six, one, or none?, J. Develop. Stud., 46, 350–370, 2010.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

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: 27 February 2013), 2011.

McQueen, J. B.: Some methods for classification and analysis of multivariate observations, in: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, available at: http://projecteuclid.org/euclid.bsmsp/1200512992, last access: 27 February 2013, University of California Press, 281–297, 1967.

Moore, M.: Death without taxes: democracy, state capacity, and aid dependence in the Fourth World, in: Towards a Democratic Developmental State, edited by: White, G. and Robinson, M., Oxford University Press, 19 pp., 1998.

Nadkarni, S. and Shenoy, P. P.: A bayesian network approach to making inferences in causal maps, Eur. J. Operat. Res., 128, 479–498, 2001.

Nadkarni, S. and Shenoy, P. P.: A causal mapping approach to constructing bayesian networks, Decision Supp. Sys., 38, 259–281, doi:10.1016/S0167-9236(03)00095-2, 2004.

Neapolitan, R. E.: Probabilistic Reasoning in Expert Systems: Theory and Algorithms, Wiley & Sons, New York, 1990.

OECD-DAC: Measuring Aid to the Water sector,http://www.oecd.org/dac/stats/36365514.pdf, last access: 27 February 2013, 4 pp., 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., Morgan Kaufmann, San Mateo, CA, 1991.

Pearl, J.: Bayesian networks, causal inference and knowledge discovery, Tech. Rep. 281, University of California, Los Angeles, CA, 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, available at: http://www.wsp.org/sites/wsp.org/files/publications/712200782528_Tackling_Corruption_in_the_Water_and_Sanitation_Sector_in_Africa.pdf (last access: 28 February 2013), 2006.

SIDA: Ecological Sanitation, Department for Natural Resources and the Environment, available at: https://wiki.umn.edu/pub/EWB/Uganda/SIDAGuidebook.pdf, last access: 28 February 2013, Sida, Stockholm, Sweden, 90 pp., 2004.

HESSD

10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Transparency International: Global Corruption report 2008, available at: http://archive.transparency.org/publications/gcr/gcr_2008 (last access: 28 February 2013), 2008.
UNDP: World urbanisation prospects 2001, available at: http://www.un.org/esa/population/publications/wup2001/WUP2001_CH3.pdf, last access: 27 February 2013, 35–49, 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 February 2013), 2006.
UNICEF, WHO: Diarhhoea: why children are still dying and what can be done?, available at: http://whqlibdoc.who.int/publications/2009/9789241598415_eng.pdf, last access: 28 February 2013, WHO publications, Geneva, 57 pp.,2009.
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: 6 March 2012.
Xiang, Y.: Probabilistic Reasoning in Multi-Agent Systems, A Graphical Models Approach, Cambridge University Press, Cambridge, doi:10.1017/CBO9780511546938, 2002.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

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

Mode	Error rate	LL[a]	QL[a]	SP[a]
(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

[a] LL = Logarithmic Loss, and QL = Quadratic Loss where best value is 0, SP = Spherical Payoff where best value is 1.

HESSD
10, 2481–2548, 2013

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 2. WS sensitivity analysis.

Variables	ER[a]	BV[a]	RMS[a]	rank
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

[a] ER = Entropy reduction, BV= Belief Variance, RMS = Root Mean Square change.

Table 3. General model – water supply WS simulations. Marked with *, 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 %

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 4. S sensitivity analysis.

Variable	Sanitation			
	ER[a]	BV[a]	RMS[a]	rank
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

[a] ER = entropy reduction, BV = belief variance, RMS = root mean square change.

HESSD
10, 2481–2548, 2013

Table 5. General model – sanitation S simulations. Marked with *, 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 %

HESSD
10, 2481–2548, 2013

Table 6. HDP sensitivity analysis.

Variables	ER[a]	BV[a]	Rank
GDP_per_cap	0.75459	0.195912	1
Child Mortal-5	0.51866	0.109089	2
% Slums	0.50799	0.097389	3
School enrol	0.44736	0.059545	4
Malaria	0.43611	0.076922	5
%Poverty	0.38475	0.049855	6
Urban Pop	0.34193	0.066113	7
School G/B	0.28867	0.067878	8
WGI GE	0.26531	0.0345821	9
Femal eco	0.18287	0.038452	10
ODA CI	0.00196	0.000424	11

[a] ER = entropy reduction, BV = belief variance.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 7. Thematic model – HDP simulations. Marked with *, 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	30 %	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 %

Table 8. CEC sensitivity analysis.

Variables	ER[a]	BV[a]	Rank
Particip	0.41742	0.083617	1
WGI VA	0.37562	0.047798	2
ODA CI	0.03427	0.004406	3

[a] ER = entropy reduction, BV = belief variance.

HESSD
10, 2481–2548, 2013

Table 9. Thematic model, CEC simulations. Marked with [a], positive variations and marked with [b], 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[a]	1.37[a]
WGI VA	HIGH	51.4 %	28.3[a]	100	−1.1[b]
ODA CI	LOW	30 %	1.3[a]	−0.6[b]	100
CEC	HIGH	30.8 %	70.1 %	54.4 %	36.2 %

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 10. AP sensitivity analysis.

Variables	ER[a]	BV[a]	Rank
% Irrigation	0.68479	0.160761	1
Water Use Int Agri	0.52047	0.090606	2
Total withdrawal	0.07282	0.011413	3
WR	0.03535	0.0055417	4

[a] ER = entropy reduction and BV = belief variance.

HESSD
10, 2481–2548, 2013

Table 11. Thematic model – AP simulations. Marked with [a], negative variations and marked with [b], the positive variations of probabilities.

Variable	LEVEL	Initial probabilities	Simulation 1	Simulation 2	Simulation 3
% irrigation	HIGH	25.5 %	100	9.8[b]	−1.3[a]
Water Use Int Agri	HIGH	44.7 %	100	11.1[b]	0.9[b]
Total withdrawal	HIGH	31.9 %	13.7[b]	100	0.2[b]
Withdrawal industrial	HIGH	25 %	(=)	1.5[b]	(=)
Withdrawal municipal	HIGH	37.5 %	(=)	−12.7[a]	(=)
WR	LOW	30 %	(=)	(=)	100
AP	HIGH	33.7 %	98.5 %	48.3 %	31.9 %[a]

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 12. Statistical performance of country profile models for WS and S.

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

HESSD
10, 2481–2548, 2013

Table 13. Profile 4 – Water Supply WS simulations.

Variables[a]	Initial probabilities	Simulation 1 HDP	Simulation 2 AP	Simulation 3 CEC	Simulation 4 ODA CI
HDP	11.5 %	100	1.1[c]	0.1[c]	−2.33[b]
ODA CI	55.6 %	−11.4[b]	0.5[c]	11.6[c]	100
CEC	36.2 %	0.3[c]	1.1[c]	100	7.6[c]
AP	16.4 %	2[c]	100	0.7[c]	0.2[c]
WR	3.7 %	(=)	(=)	(=)	(=)
WS	16.5 %	35.8 %	23 %	22.9 %	18.8 %

[a] Values concern HIGH category of each variable. Marked with [b], negative variations and marked with [c], positive variations of probabilities.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 14. Profile 4 – sanitation S simulations.

Variables[a]	Initial probabilities	Simulation 1 HDP	Simulation 2 AP	Simulation 3 CEC	Simulation 4 ODA CI
HDP	11.7 %	100	1.7[c]	−1.75[b]	−3.11[b]
ODA CI	55.6 %	−14.8[b]	−1.5[b]	12.4[c]	100
CEC	38.3 %	−5.7[b]	−1.3[b]	100	8.6[c]
AP	15.6 %	2.2[c]	100	−0.4[b]	−0.4[b]
WR	3.7 %	(=)	(=)	(=)	(=)
S	16.4 %	37 %	25.7 %	11.3 %[b]	12.5 %[b]

[a] Values concern HIGH category of each variable.
[b] Negative variations.
[c] Positive variations of probabilities.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 15. Profile 5 – Water Supply WS simulations.

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

[a] Values concern HIGH category for each variable.
[b] Positive variations.

HESSD
10, 2481–2548, 2013

Table 16. Profile 5 – sanitation S simulations.

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

[a] Values concern HIGH category for each variable.
[b] Positive variations.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Table 17. Recapitulative table on probability variations (difference between recomputed and initial probability).

Variables[a]		High WS probability change	High S probability change
	Suitable Level	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	52.9	
Urban pop	HIGH	43	
Child Mortal-5	LOW	33.4	
%Poverty	LOW	33	
% Slums	LOW	31.7	
Malaria	LOW	27.5	
ODA CI	MID	25.6	
WGI GE	HIGH	24.6	
School enrol	HIGH	24.1	
School G/B	HIGH	17.7	
Femal eco	LOW	16.4	
	Suitable level	High CEC probability change	
Particip	HIGH	39.3	
WGI VA	HIGH	23.6	
ODA CI	LOW	5.4	
	Suitable level	High AP probability change	
% irrigation	HIGH	59.2	
Water use in agriculture	HIGH	35.5	
Total withdrawal	HIGH	14.6	

[a] Water resources variables are submit to very long term pattern changes therefore they are only use as descriptive indicator.

HESSD
10, 2481–2548, 2013

Table A1. Clustering limits for 25 variables from WatSan4Dev sub-set. Limits of variable categories, high, mid, low.

Variables	Classes				Unit or intervals[a]
	Low		Mid	High	
	Minimum	Upper limit	Upper limit	Maximum	
Water Supply (WS)	22.00	62.17	na	100.00	%
Sanitation (S)	9.00	58.50	na	100.00	%
%Poverty	7.60	29.00	51.62	71.30	%
Urban Pop	10.00	44.58	78.29	98.00	%
School enrol	22.70	38.69	66.75	94.10	%
Femal eco	17.60	39.12	73.10	91.80	%
% Slums	0.00	31.72	64.31	97.00	%
Withdrawal Municipal	1.00	19.34	52.88	84.00	%
Withdrawal industrial e	0.00	25.99	54.75	73.00	%
TWRR	7.64	40 771.23	203 825.60	326 116.37	$m^3 y^{-1} cap^{-1}$
Child Mortal-5	7.00	84.14	166.19	265.00	‰
School G/B	0.63	0.72	0.87	1.02	[0.63,1.02]
Particip to IEA	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	0.00	33.87	67.93	99.90	%
ODA WSS	0.00	5.55	11.26	16.51	$ per cap
ODA	0.60	102.79	194.59	289.27	$ per cap
Malaria	0.00	143.86	287.71	428.14	‰
Water Use int Agri	0.50	7565.36	17 708.50	28 333.30	$m^3 ha^{-1} yr^{-1}$
Precipit	51.20	1061.79	2103.29	3141.70	$mm yr^{-1}$
Desert risk	0.00	33.70	66.40	100.00	%
Total withdrawal	6.43	747.60	1475.85	2161.00	$m^3 cap^{-1} yr^{-1}$
WGI.V.A.	−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

[a] na: composite indicator with dimensionless unit.

HESSD
10, 2481–2548, 2013

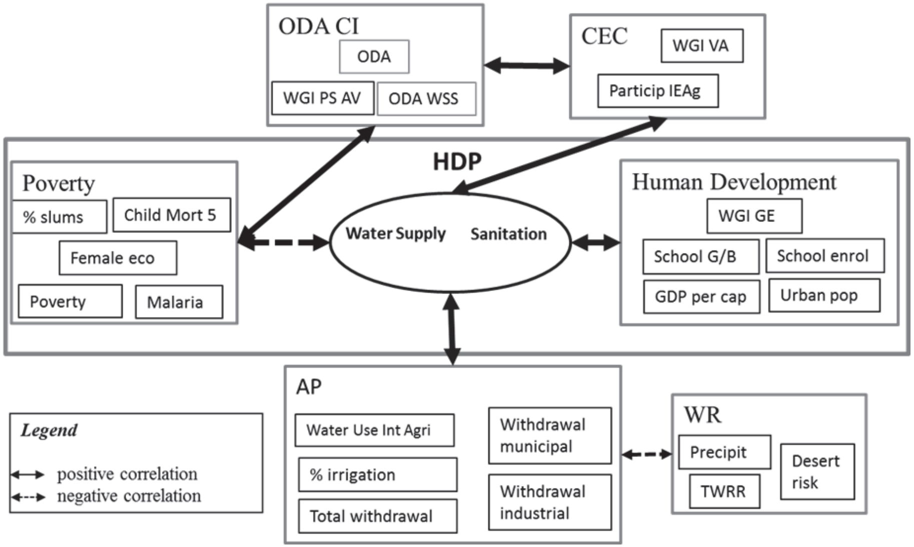

Fig. 1. Organisation and main correlations between selected variables.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

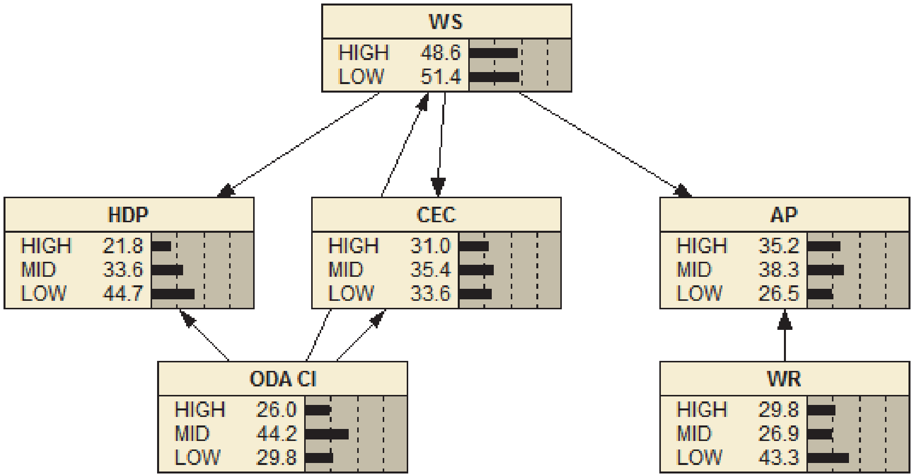

Fig. 2. WS general model, DAG shows the relationships between the WS and pillars.

HESSD
10, 2481–2548, 2013

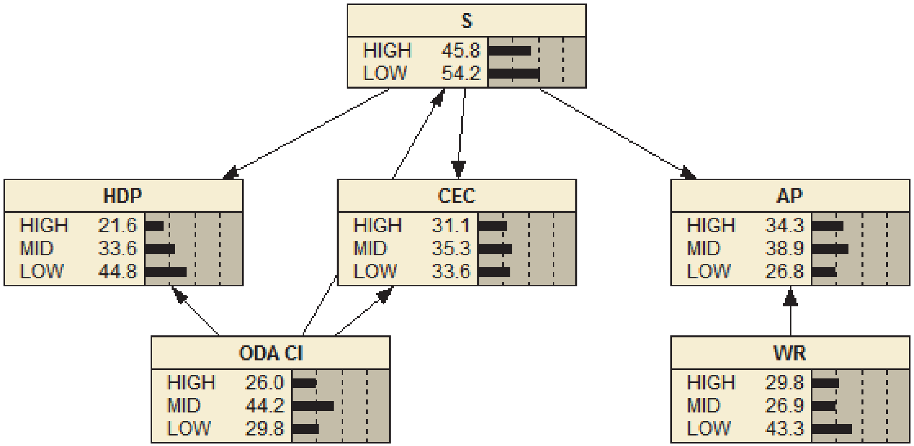

Fig. 3. S general model, DAG shows the relationships between the S and pillars.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

HESSD
10, 2481–2548, 2013

Bayesian networks modelling to cross cutting analysis of water supply

C. Dondeynaz et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

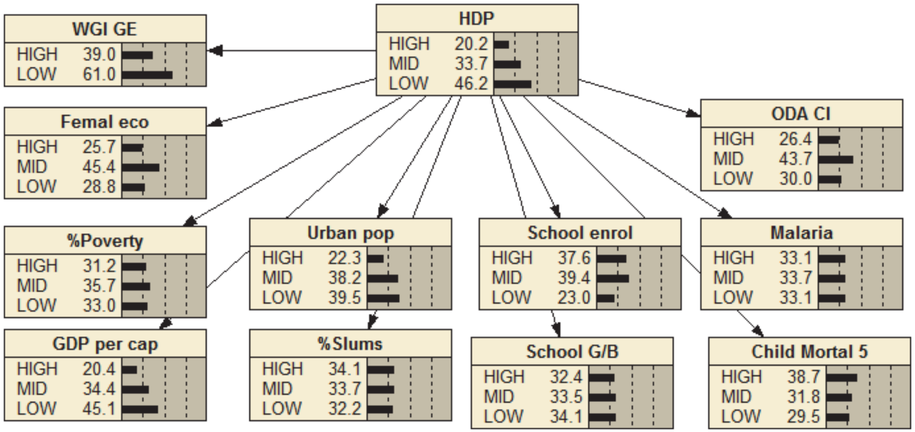

Fig. 4. DAG and probabilities in % for the variables involved in the HDP pillar discretised in high (HIGH), medium (MID) and low (LOW) levels.

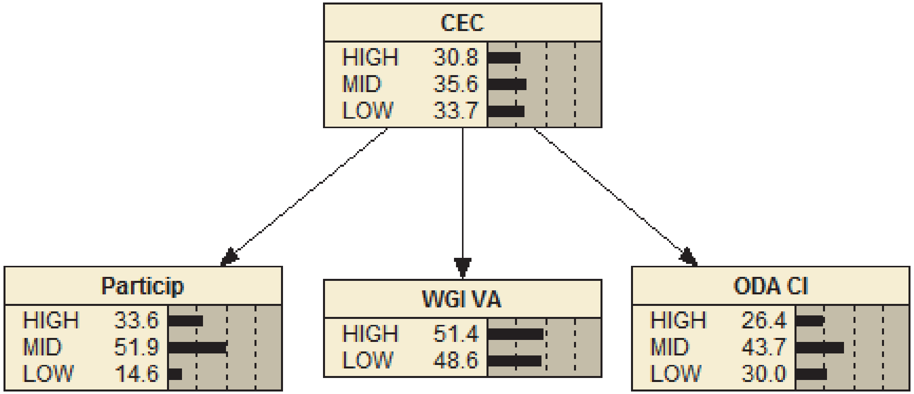

Fig. 5. Probabilities in % for the variables involved in the CEC pillar discretised in high (HIGH), medium (MID) and low (LOW) levels.

HESSD
10, 2481–2548, 2013
Bayesian networks modelling to cross cutting analysis of water supply
C. Dondeynaz et al.

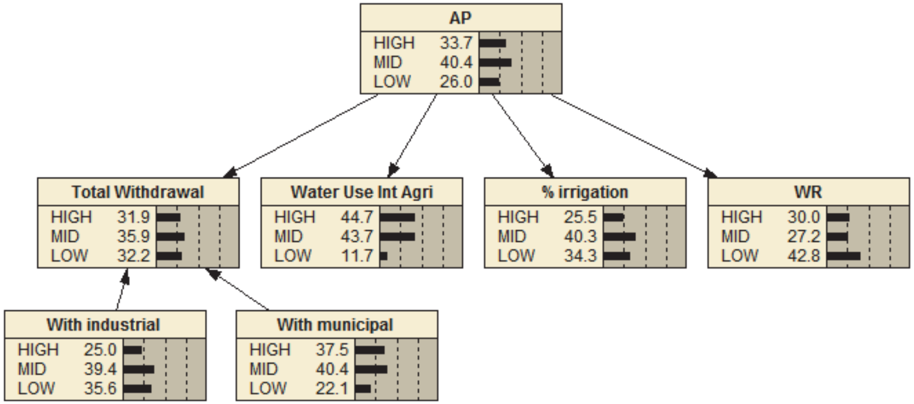

Fig. 6. DAG and probabilities for the variables involved in the AP pillar discretised in high (HIGH), medium (MID) and low (LOW) levels.

HESSD
10, 2481–2548, 2013

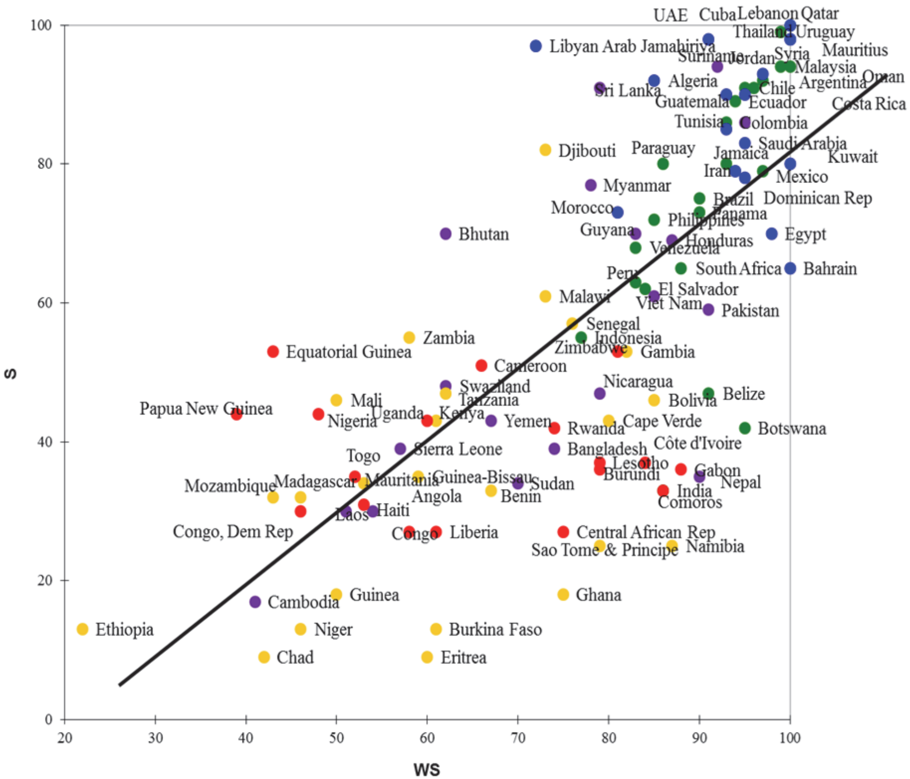

Fig. 7. WS versus S according to 5 country profiles with linear equation as S = 1.0285WS – 21.82 and R^2 = 0.5265.

HESSD

10, 2481–2548, 2013