For office use only

T1 ________________

T2 ________________

T3 ________________

T4 ________________

Team Control Number

52831

Problem Chosen

E

For office use only

F1 ________________

F2 ________________

F3 ________________

F4 ________________

2016
MCM/ICM
Summary Sheet

(Your team's summary should be included as the first page of your electronic submission.)

Type a summary of your results on this page. Do not include the name of your school, advisor, or team members on this page.

Tackling Water Scarcity: Modeling Water Access in India Using Multiple Regression Analysis

"Water, water, every where," lamented the Ancient Mariner, "Nor any drop to drink" [8]. Samuel Taylor Coleridge might as well have been describing India's water crisis. It is not an absolute lack of water that threatens India's water security, but a challenge of managing this critical resource. Our task was to construct a mathematical model capable of measuring a country's ability to provide clean water for its population, to use this model to forecast India's water security fifteen years into the future, to develop an intervention policy to enhance water security, and to use our model to predict the policy's success over the same time frame.

We approached the problem by constructing a series of multiple linear regression models using publicly available data in order to account for the wide variety of factors that influence water access. We first used a simple two-predictor model that explained 43.7% of the variance in access to water. Using this basic model as a foundation, we considered a variety of predictors for water security and used a variable selection algorithm to develop a refined model that explained 71.6% of the variance in access to water.

After testing the model's applicability to India's water supply by comparing model-generated predictions of access to water against published statistics, we used an autoregressive multivariate model to forecast the change of each water access predictor variable over the next fifteen years and the resulting projections to predict changes to India's water security through the year 2031. Our model did not directly account for climate change, so we used published data on effects of climate change on water supply to create a climate-corrected forecast model.

Finally, we developed a multidisciplinary intervention policy designed to increase the availability of clean water in India. Using published data to determine appropriate alterations to the predictor values, we simulated the changes in access to water fifteen years after implementation of our proposed policy changes, resulting in a 3.78% increase in water access.

Tackling Water Scarcity: Modeling Water Access in India Using Multiple Regression Analysis

Contents

Introduction

Human life depends on access to clean water. While the world possesses an abundance of water, many populations lack access to the clean freshwater essential for life. The world's population is currently estimated to have exceeded 7.3 billion people and is projected to reach 8.5 billion by 2030 [20]. Ensuring all these people have access to clean water is a challenge that must be addressed on multiple fronts.

Though water is a renewable resource, groundwater and surface freshwater available for drinking or agriculture accounts for only 1% of the world's total water supply (the vast majority of the rest being ocean water) [14]. Human populations are generally established in regions near water, with only 10% of the world's population living more than 10 km from a surface freshwater body [12]. However, presence of a freshwater body alone does not guarantee access to safe and clean water. Despite improvements through technology and education, some regions still practice open defecation or lack sanitation facilities, resulting in contaminated water supplies [26].

Numerous factors, ranging from the physical and ecological to the political and socioeconomic, affect whether a country or region experiences water scarcity. Among these are population growth, agriculture and food production, climate, climate change, land use, water quality or water pollution, demand for water, economic status, professional capacity (including education and working conditions), political issues, social issues [3], and catastrophic events such as droughts or floods. As of 2007, an estimated 1.2 billion people experience physical scarcity and an additional 1.6 billion experience economic scarcity [21].

In India, attempts by the government and other organizations have had some success in developing municipal water supplies [11], but have failed to solve the problem for many Indians, in part due to the country's large and growing population, which is expected to exceed 1.5 billion people by 2030 [20].

In this paper, we address the challenge of modeling (and designing intervention policies against) water scarcity by developing a series of regression models of increasing complexity. We apply these models, combined with projections of future events such as population growth and climate change, to develop a forecast of future water scarcity in India absent any intervention plan. Finally, we use our model to propose a possible policy for addressing water scarcity in India and suggest the likely outcome of its implementation.

Definitions

Because water scarcity is a complicated interdisciplinary problem, it is important to exercise care in operationally defining one's terms. Toward that end, in the interest of clarity, we use the following definitions in this paper:

- **Access to electricity** is the percentage of a country's population with access to electricity according to data from "industry, national surveys, and international sources" [28].
- **Access to water** is the percentage of a country's population actively using an improved water source [38].
- **Air pollution**, expressed in micrograms per cubic meter, is "the average level of exposure of a nation's population to concentrations of suspended particles measuring less than 2.5 microns in aerodynamic diameter, which are capable of penetrating deep into the respiratory tract and causing severe health damage" [39].
- **Clean water** is freshwater that meets the criteria of an improved water source.
- **Freshwater** is defined as water that is not ocean water or other salt water.
- **GDP per capita** is a country's gross domestic product divided by that same country's midyear population and is reported in current US dollars [34].
- **Improved water sources** are defined as "piped water on premises" or "other improved drinking water sources" including "public taps or standpipes, tube wells or boreholes, protected dug wells, protected springs, [and] rainwater collection" [27].

- **Improved sanitation facilities** are defined as “likely to ensure hygienic separation of human excreta from human contact” and include systems that flush to a sewer, septic tank, or pit latrine, pit latrines with slab, and composting toilets [27].
- **Intervention**, in the context of this paper, refers to any measure or measures implemented by governments or organizations to directly or indirectly increase a population’s access to clean water.
- **Pollution** is any contaminant introduced into an otherwise safe water supply and which renders the water supply unfit (or less fit) for drinking or use in agriculture.
- **Population growth** is the annual increase in a country’s population counting all permanently settled residents, defined in our models as “the exponential rate of growth of midyear population from year *t* - 1 to *t*, expressed as a percentage” [41].
- **Sanitation** is the percentage of a country’s population using improved sanitation facilities [37].
- **Scarcity of water** is a relative measure of the supply of water available to a region against the region’s demand for water; specifically in this case, scarcity of water refers to scarcity of clean water, and is operationally defined as lack of access to water.
- **Water security** is a measure of the ability of a population to ensure supply of clean water meets or exceeds demand; in this paper, we operationally define water security as access to water.
- **Water supply**, expressed in cubic meters, is defined as “renewable internal freshwater resources per capita” [43]

Assumptions

Water scarcity is a complex interdisciplinary issue with international significance. Contained within the problem are variables pertaining to politics, economics, culture, human biology, ecology, climatology, geology, and many other disciplines. It is impossible to fully model every possible relevant factor. As such, we made several assumptions and simplifications.

First, in all cases, we assume that the data we used in our models were both complete and accurate. Inherent in our use of linear regression for the construction of our model, we must assume that all variables included are linearly related to access to clean water, that none of our predictor variables are correlated, the data are normally distributed, and the errors from the regression are normally distributed with mean zero and constant variance. In all models, we assume that percent of population with access to clean water is a direct (inverse) proxy for water scarcity. Because our model incorporates a forecast, we assume that the relationships between variables will remain consistent for at least the next fifteen years. Finally, due to availability of data, we model 2012 as the “current” year.

In addition, each of the particular models we constructed have their own sets of assumptions.

Basic Model Assumptions

The basic model is our simplest model, so the assumptions we made are few and far-reaching. Specifically, we assume that population growth is a direct proxy for clean water demand and that GDP per capita is a direct proxy for a country’s ability to meet that demand.

Refined Model Assumptions

For the refined model, the same assumptions apply, with the following additions and modifications:

- GDP per capita is a direct proxy for economic ability to promote water security,
- A country’s continent and latitude together provide a direct proxy for geographic variance and climate,
- Access to electricity is a direct proxy for a country’s technological ability to promote water security,
- Air pollution is a direct proxy for industrial and agricultural water pollution.

Models of the Water Scarcity Crisis

To account for the numerous physical and socioeconomic factors that influence water scarcity, we developed a series of multiple linear regression models of increasing refinement. The Basic Model provides an extremely simplified view, but served as the basis on which we built the Refined Model, which accounts for more variables. We used the Refined Model, in turn, to forecast Indian water scarcity and to develop an intervention plan for improving water security in India.

For clarity, please note that in our models, all data given as a percentage are reported on a 0 to 100 scale rather than a 0 to 1 scale.

The Basic Model

For the Basic Model, our goal was simply to establish a baseline against which to measure the improvements of our Refined Model and to quickly locate the limitations of our technique.

Description of the Model

In this model, we examine two predictors of water security. First, we treated population growth (annual % increase), as a proxy for a population's increasing demand for clean water. Second, we treated GDP per capita (current US$) as a proxy for a population's ability to meet that demand. We measured water security as the percentage of the population with access to an improved water source.

Using data on 212 countries from the World Bank [34; 38; 41], we constructed a linear regression model in R, treating GDP per capita and population growth as predictors of access to water in 2012. Plots of the data are shown in Figure 1.

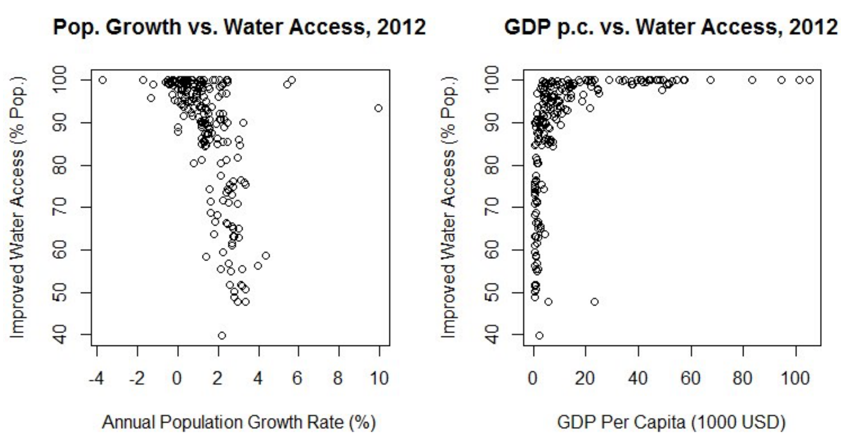

Figure 1. Plots of Basic Model data

Results

Our Basic Model yielded the linear regression equation:

$$W = 0.0002971G - 5.354P + 91.38$$

where:

W = Access to water
G = GDP per capita
P = Population growth,

with an adjusted R^2 value of 0.437 and statistical significance ($p_G = 6.37 \times 10^{-12}; p_P < 2 \times 10^{-16}$) on both variables.

Discussion

These preliminary results show that, under the assumptions of the Basic Model, an increase in the rate of a country's population growth by one percentage point is, on average, associated with a 5.35% *decrease* in access to water, holding all other variables constant; similarly, every $1000 increase in GDP per capita results in a 0.297% increase in access to water, with all other variables held constant. These results make intuitive sense if one considers GDP per capita a proxy for a country's economic ability to meet water demands and population growth a representation of increase in that demand.

Though it ignores numerous important factors, the Basic Model demonstrates the possibility of using multiple linear regression to predict water scarcity, so we used it as a starting point for our Refined Model.

Refined Model

Beginning from the skeleton of the Basic Model, we chose a set of twenty variables to consider as possible predictors of access to water. We then chose the subset of these variables resulting in the highest adjusted R^2 value. We used the resulting set of eight predictors to model access to water in 2012.

Description of refinements

In order to more accurately model access to water, we selected twenty variables from publicly-available datasets. An overview of these variables and our justification for their selection is presented in Table 1 below. Once we achieved this cross-section of the physical, social, and economic variables that can affect water scarcity, we culled the subset with the highest adjusted R^2 value using the "*regsubsets*" function in R. Our Refined Model contains this subset of eight variables (designated with an asterisk in Table 1).

We then used a multiple linear regression in R to model these selected variables over 147 countries as predictors of access to water in 2012. The resulting regression equation is our Refined Model.

Table 1. Overview of variables considered for Refined Model (* indicates variables included in Refined Model)

Variable	Units	Justification
*Population Growth [41]	Annual % increase	We treat population growth as a proxy for a country's demand for clean water; increased population growth is considered to be an increase in demand.
*GDP Per Capita [34]	Current US $	GDP per capita is our proxy for a country's *economic* ability to increase the supply of clean water.
*Water Supply Per Capita [43]	Cubic meters	Water supply per capita serves as a measure of a country's total supply of freshwater (clean or otherwise), controlled for population size.
*Continent	Categorical	Combined with latitude (see below), the continent on which a country resides serves as a partial metric of geographic variance in access to water between countries.
*Access to Electricity [28]	% of population	The proportion of a country's population has access to electricity is our proxy for that country's *technological* ability to increase the supply of clean water.
*Air Pollution [39]	Micrograms per cubic meter	Air pollution serves as our proxy for a country's level of industrial and agricultural water pollution in the absence of quality data on levels of water pollution.
*Latitude [9]	Degrees	Combined with the continent on which a country resides, we use latitude as a partial metric of geographic variance in access to water between countries.
*Sanitation [37]	% of population with access	The portion of a country's population with access to improved sanitation facilities is our (inverse) proxy for individual or domestic water contamination.
Total Population [42]	Number of individuals	The total population of a country serves as a metric of the current demand for clean water.
Population Density [40]	People per sq. km of land	We use population density as a measurement of a how efficiently a country can distribute clean water.
Total GDP [33]	Current US $	Total GDP is an alternative (to GDP per capita) proxy for a country's economic ability to increase the supply of clean water.
Annual Precipitation [32]	Millimeters per year	We used annual precipitation as a measure of the climatic variance between countries.

Table 1. Overview of variables considered for Refined Model (* indicates variables included in Refined Model)

Variable	**Units**	**Justification**
Rural Population [44]	% of total population	Because water supply strategies may vary between urban and rural populations, the proportion of a population living in rural areas could affect a country's ability to supply clean water.
Water Used: Agriculture [29]	% of total freshwater withdrawal	Because the amount of water used for various purposes (agricultural, industrial, domestic) differs from country to country, the relative proportions of water used in each context could affect the efficiency of a country's water program.
Water Used: Industry [31]	% of total freshwater withdrawal	See above.
Water Used: Domestic [30]	% of total freshwater withdrawal	See above.
Public Health Expenditures [35]	% of total health expenditures	Because water-borne communicable diseases are a possible result of water scarcity, a country with large public health expenditures may be well-invested in water distribution programs in the public interest; alternatively, countries with large public health expenditures may be facing the consequences of water scarcity.
Educational Attainment [5]	Average years per person	Educated individuals may be able better situated to contribute individually to water conservation; educated societies may be more likely to develop technological advances to improve access to water.
Distance to Navigable River [9]	Kilometers	Especially in countries without a municipal water supply, the distance to a freshwater body such as a river may affect the proportion of the population with access to water.

Plots of population growth and GDP per capita are identical to those from the Basic Model and are presented in Figure 1. Plots of the remaining six variables we used are presented in Figure 2.

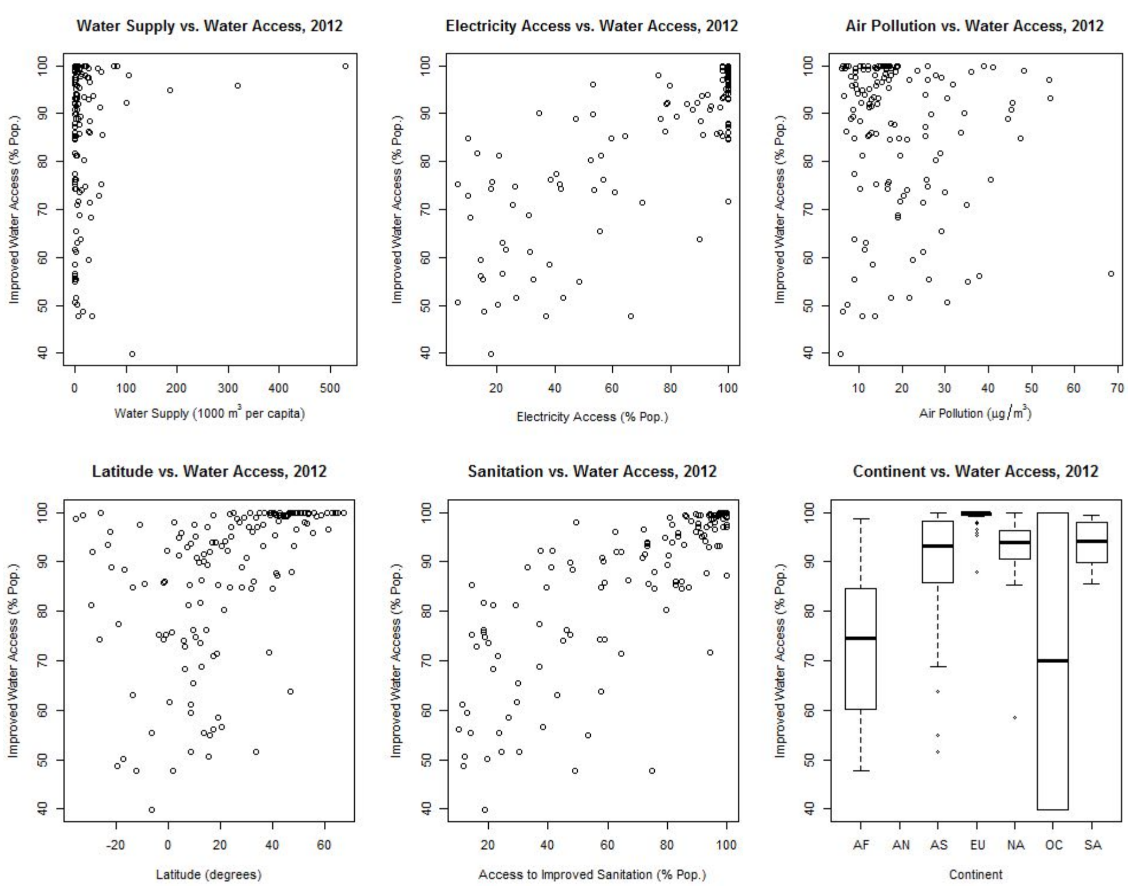

Figure 2. Plots of Refined Model data

Results

Our multiple linear regression resulted in the coefficients and p-values reported in Table 2 below, with an adjusted R^2 value of 0.716. Note that because the continent on which a country resides is categorical rather than quantitative data, the continent variable has been split into sub-variables with the following continent codes: AS (Asia), EU (Europe), NA (North America), OC (Oceania), SA (South America).

Table 2. Coefficients and *p*-values for Refined Model variables (variables in **bold** are significant at the $p<0.05$ level)

Variable	Coefficient	*p*-value
Population Growth	**-1.789**	**0.01523**
GDP Per Capita	**1.009×10^{-4}**	**0.02860**
Water Supply Per Capita	1.772×10^{-5}	0.19477
Continent: AS	-2.448	0.37877
Continent: EU	2.790	0.48747
Continent: NA	-0.1394	0.96595
Continent: OC	**-16.99**	**0.00853**
Continent: SA	-6.617	0.08423
Access to Electricity	**0.2771**	**7.16×10^{-7}**
Air Pollution	**0.1864**	**0.02161**
Sanitation	**0.1490**	**0.00808**
Latitude	**-0.1664**	**0.00511**

Discussion

All of the data used in the multiple regression to construct this model was from the year 2012. While we were able to locate some more recent data, 2012 was the most recent year in which we located annual data for a majority of our variables. Our data on air pollution, however, was the only variable used which did not have any values for 2012. In order to repair this deficiency, we took the arithmetic mean of 2011 and 2013 values for each country to approximate the 2012 values. Because this gap is only one year long and we believe catastrophic leaps in air quality are unlikely in such a time frame, we are confident that this estimation introduces a smaller error into our model than would using data from different years.

We note also that the coefficient for air pollution is a positive number, where one would intuitively expect a negative number, but we consider this a byproduct of the multiple regression, rather than an error in our model.

Application of the Refined Model to India

Our goal in this section is to apply our Refined Model to the country of India, use it to forecast India's water security for the next fifteen years, and then develop and model the success of an intervention designed to increase the Indian population's access to water.

Background on the Indian Water Crisis

Though India has the world's seventh largest economy by nominal GDP [10], and efforts to improve access to water in India have had some success [18] a confluence of factors have placed the country on the brink of water crisis. Though the drivers of water scarcity are too numerous to address in this paper, we have identified several key causes of the problem.

We first note that, though there are environmental factors related to water scarcity (including climate change [36], rainfall patterns, and geographic features such as rivers and groundwater sources), the impending water crisis in India is "predominantly a manmade problem" [7] resulting from failures of socioeconomic systems.

India's population growth is easily the leading threat to water security. Though India is not a particularly dry country [7], its large and growing population places increasing demands on the water supply, with the result that, though much of the country has access to some water, the current water supply in India is less than 1,000 cubic meters per person, already less than the 1,700 cubic meter per person threshold for considering a country water-stressed [13] and the population is still increasing [20].

Additional problems include an insufficient sanitation system (in 2012, only 38% of the Indian population had access to improved sanitation facilities [37], and many individuals are reluctant to use such facilities for cultural reasons [6]) and over-extraction of groundwater resources for agricultural purposes [13].

Though our Refined Model does not account for all of these variables, it was drawn from a pool of variables that included measures of water used for agriculture [29], sanitation [37], water supply [43], and population growth [41] (see Table 1). Additionally, a later extension to the model (see the "Forecast to 2031" section below) accounted for climate change.

Testing the Model's Applicability to India

In order get a sense of our forecast's accuracy, we first tested the model for its applicability to India. Because our model is based on values from 2012, we looked up a literature value and determined that 92.2% of the Indian population had access to water that year [38].

Because India was included in the 147 countries in our model, we re-ran the regression with India's values removed in order to perform and out-of-sample test. The coefficients and *p*-values for each variable in this modified model are presented in Table 3.

Table 3. Coefficients and *p*-values for Refined Model with India removed (variables in **bold** are significant at the p<0.05 level)

Variable	Coefficient	*p*-value
Population Growth	**-1.723**	**0.01980**
GDP Per Capita	**9.826×10^{-5}**	**0.03320**
Water Supply Per Capita	1.741×10^{-5}	0.20258
Continent: AS	-2.838	0.31205
Continent: EU	2.537	0.55940
Continent: NA	-0.4585	0.88885
Continent: OC	**-17.11**	**0.00808**
Continent: SA	-6.712	0.07994
Access to Electricity	**0.2698**	**1.59×10^{-6}**
Air Pollution	**0.1713**	**0.03735**
Sanitation	**0.1615**	**0.00509**
Latitude	**-0.1608**	**0.00701**

With India's values removed, our model suggests that 83.2% of India's population should have had access to water in 2012, while the version with India's values included suggests that 84.2% of the population should have had access to water in the same year.

Though our model's values for 2012 do not match the actual values, they are close enough to each other that we don't feel the model is significantly influenced by the presence of India's own data. Indeed, we suspect these data support the model's validity. The difference between our model's values and the actual value will be considered in the discussion below.

Forecast to 2031

To turn our model from a simple multiple linear regression into a forecast model, we generated forecast data for each variable in the model independently and then re-ran the model as applied to India with the projected data for each of the next fifteen years. Specifically, we used four techniques to generate the forecast data.

For population growth, GDP per capita, and sanitation, we used autoregression in R on each predictor. That is, we ran a regression for each predictor using its most recent value as the dependent variable and the previous 10 years of values as independent variables. The resulting coefficients were used to forecast the data through the year 2031. Continent and latitude, as non-variable predictors, were held constant.

Because we had limited data available for the remaining three variables, we accounted for their forecasts through various means. In the case of water supply, we assumed that (absent changes from climate change accounted for later) the total water supply of a country should remain relatively constant. As such, we averaged the 2012 and 2013 values for water supply and held that value constant through 2031. For access to electricity, we lacked sufficient data for an autoregression, so we assumed the percentage of people without access to electricity would decrease at the same rate as between 1990 and 2012 (approximately 3.7% annually in India). We then used this rate to project values from 2012 to 2031. Finally, we similarly lacked sufficient data for an autoregression on our air pollution statistic, so we assumed its rate of increase to be constant from 2000 to 2013 (approximately 2.5% annually) and used this rate to project values from 2013 to 2031.

Our forecast suggests that by the year 2031, 94.4% of the Indian population should have access to water, up from 84.2% in 2012. The remainder of the forecast data are shown in Table 4.

Table 4. Access to water (% of population) in India from 2012 to 2031

Year	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**
Access to Water	84.2	84.8	85.5	85.9	86.5	87.2	87.5	88.2	88.6	89.0
Year	**2022**	**2023**	**2024**	**2025**	**2026**	**2027**	**2028**	**2029**	**2030**	**2031**
Access to Water	89.7	90.1	90.5	91.2	91.6	91.9	92.8	92.9	93.3	94.4

Forecast accounting for climate change

In order to account for the effects of climate change, we made a simple alteration to our projections. According to a report by the World Bank, a 3° C increase in temperature by the 2080s could result in a 10% decrease in water supply in South Asia (including India) [45]. We interpret "water supply" in the context of this report to mean a supply of usable clean water, so we treat it as directly applicable to our projections of access to water. Assuming, then, that the report's projections hold and that both increase in temperature and decrease in availability of water follow a linear trend, we determined that India might experience a 0.044° C increase in temperature per year and thus a 0.149% decrease in access to water per year, beginning in 2013.

Projecting these values forward, we predict a 0.792° C increase in temperature in India by 2031, accompanied by a reduction in access to water of 2.68 percentage points, resulting in a 2031 value for access to water of 91.7% of the population.

Discussion

We observe that, even though our model's initial (2012) values for access to water in India are below published values as mentioned above, the positive trend in our forecast is at odds with other forecasts of an impending water crisis in India. We will discuss possible extensions of this work to further explain this discrepancy in the "Future Work" section. For now, we interpret this difference in projections as a result of differing metrics of water scarcity. Even though our model accounts for population growth and sanitation, it is likely that our model is predicting the total freshwater available to the Indian population, and not necessarily the amount the Indian infrastructure is capable of delivering to individual communities or residents. Therefore, it is possible that India may not have a water shortage in absolute terms, but may still experience a devastating water crisis due to failures in distribution or treatment.

Intervention Plan

Drinking water is the obvious concern most people would consider when they think of water scarcity. Indeed, according to the World Health Organization, 1.1 billion people lack access to any improved source of clean drinking water [24]. In India, despite past and current efforts to improve drinking water, 21% of communicable diseases remain water-related [18]. Clearly, for this reason alone, any intervention capable of successfully improving access to clean water could improve the lives of millions or billions of people.

Despite this obvious consideration, improvement of access to clean water has other, far-reaching consequences. About 70% of the world's freshwater supply is used in agriculture [23]. Water scarcity, then, plays a significant role in feeding the world's population. In particular, India has become the world's sixth largest agricultural exporter [22].

Water scarcity in India threatens not only the availability of drinking water, but the agricultural supply both for India itself and for importers of Indian crops. All of this can have complex and far-reaching consequences, but a partial list of effects Indians could experience from water scarcity include: drinking water shortage, poorer health, diminishing food crops, and diminishing international trade. On the other hand, improved water security can provide exactly the opposite benefits: plentiful clean drinking water, improved health, a strong agricultural industry, and robust international trade. As such, it is essential to develop interventions designed to improve access to water.

In this section, we suggest several different policies designed to improve India's water security. While these interventions were considered specifically with India in mind, many of them could easily be applied to other countries or regions.

Overview and justification of interventions

Because our model accounts for several different variables, we developed intervention techniques focused on different aspects of the water scarcity problem. Specifically, we looked at:

- Water supply,
- Sanitation,
- Population growth,
- GDP per capita,
- Pollution,
- Technology.

Though our model accounts for water pollution through the proxy of air pollution and technological advancement through the proxy of access to electricity, we consider the variables of pollution and technology more broadly in our intervention plan. In the end, we selected two of these interventions to test in our model, but present our entire collection of proposals for consideration.

Interventions for water supply

Though water is a renewable resource, it's part of a dynamic system and regularly changes location. Practically speaking, it is possible to consider the total supply of water on Earth as constant. It is also generally inefficient to transport large quantities of water long distances. Furthermore, India is not a particularly dry region [7]. Therefore, we do not consider any methods for increasing the absolute quantity of water available to India. Instead, we focus on methods of increasing the availability or utility of water that already exists.

First, we consider that removal of contaminants could open up vast new reserves of usable freshwater in India. For instance, the Ganges River is so polluted that the total of toxins, hazardous chemicals, and bacteria in the river are almost 3000 times the WHO's "safe" limit [1]. Increased regulation to prevent industrial pollution and the introduction of an improved treatment infrastructure, while perhaps costly in the short term, could significantly reduce pollution in existing bodies of water, making available vast quantities of otherwise-unusable freshwater.

Furthermore, much of the freshwater in India is transitory. As much as 45% of the water from precipitation in India is never used as rivers deposit the water into either the Arabian Sea or the Bay of Bengal [2]. Our suggested intervention plan thus includes the collection and storage of rainwater for later use. This could also reduce the current over-exploitation of groundwater sources for agriculture [13].

Interventions for sanitation

One of the biggest contributors to water pollution in India is poor individual sanitation practices. Part of the problem is institutional, in that as of 2012, only 38% of the Indian population had access to improved sanitation facilities [37]. Compounding the problem, certain subsets of the Indian population are reluctant to give up the practice of open defecation for cultural reasons [6]. In order to increase access to improved sanitation facilities, we suggest a large-scale government program dedicated to providing every Indian with access to improved sanitation. In order to overcome cultural resistance, this program must be accompanied by an educational campaign, and perhaps a program of incentives, to maximize use of the new facilities.

Interventions for population growth

Three general approaches, best implemented in concert with one another, suggest themselves as interventions to decrease growth in population. The first and simplest is a program to increase distribution of condoms and increase their use. Primarily, this must be an education campaign, as the Indian government already distributes condoms free of charge, but the rate at which they are used has recently declined from 26 million in 2007 to 16 million in 2011 [15]. Fortunately, this change is hypothesized to be due to a preference for other contraceptive measures. However, maintaining or even increasing condom distribution should help to reduce India's population growth as well as contribute to public health.

The second approach, which we have considered in our model, is to increase women's education. It is well-established that increasing women's liberation correlates with declining rates of population growth [17]. We propose a government program to incentivize (and perhaps to subsidize) enrollment in secondary education for women with a goal of increasing women's secondary education enrollment rates by 1% annually.

We also considered a program to educate citizens about family planning alternatives and, perhaps, to incentivize certain family-planning practices. While the former remains a good idea to help people voluntarily reduce population growth, we suspect the latter could breed contempt among the people for cultural reasons [4]. Additionally, a more aggressive plan such as China's one-child policy would likely be doomed to failure because as a democracy, India lacks the centralized control to make such policies feasible.

Interventions for GDP

Because India already has the world's seventh-largest nominal GDP [10], we don't feel any intervention is necessary to increase GDP in absolute terms (though any healthy economy should strive toward increasing GDP in general). However, India also has one of the world's lowest values of GDP per capita (1262.64 US dollars in 2014), due to its overwhelming population size [19]. Because of this, our recommendation to increase GDP per capita is to focus on reducing population growth, perhaps using the methods suggested above, while maintaining an increase in nominal GDP.

Interventions for pollution and technology

Our model used proxy statistics for pollution and technology, so we were unable to model any direct interventions for these predictors of water scarcity. However, we recommend increased investment in renewable clean energy and research for innovative water treatment facilities as an essential part of any plan to maximize water security. By developing and deploying new water treatment technologies, India can begin to clean up existing pollution, and by investing in renewable clean energy, they can minimize future pollution. Because India's shortage is of usable clean water, not water in general, this would go a long way toward solving their problem. Furthermore, though the effects would not likely be felt within the fifteen years our model forecasts, investment in alternative energy sources can, in the long term, minimize the effects of climate change on water availability.

Mathematical description of the intervention plan

Because of time constraints and due to the large scope of our intervention plan, we chose to use our model to test two of the interventions: specifically, the sanitation intervention and the women's education intervention for population growth.

To model the effects of the sanitation program, we examined the results of a similar sanitation education program in Bangladesh, the result of which was a drop in open defecation rates from 34% in 1990 to 3% in 2012 [25]. We assumed the results of a comparable program would be similar in India, an estimate we believe to be conservative because our proposed sanitation intervention also includes a large program of distribution of toilets or other sanitation facilities. Thus, we estimated that the sanitation program would increase our access to sanitation statistic by 1.41% per year, so we modified the sanitation statistic in our Refined Model to reflect this annual increase.

It was more difficult to determine a value for the effect of women's education on population growth. We first chose a 1% annual increase in women's secondary education as an arbitrary but, we think, realistic value for the results of a directed campaign to increase women's social standing through education. We then used a linear regression analysis comparing published women's secondary education enrollment data [16] to population growth in order to determine the effect this change would have on the rate of population growth. The results of this regression are shown in Figure 3.

The regression yielded the equation:

$$G = -0.023589E + 3.212242,$$

where:

G = Population growth (%)
E = Women's gross enrollment in secondary school (%).

The result was significant ($p = 7.38\text{x}10^{-8}$; $R^2 = 0.2105$), so we used the result of this regression analysis to determine the effect on population growth of our hypothetical 1% increase in women's secondary education. Specifically, the 1% increase in women's education resulted in a 0.24% decrease in population growth.

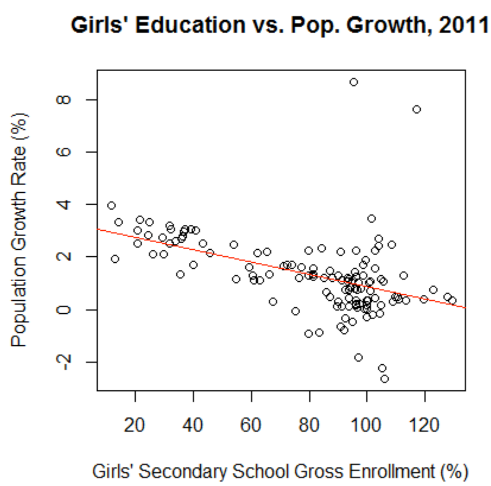

Figure 3. Effect of women's secondary education on population growth

To account for the results of this intervention, therefore, we subtracted 0.24% from the population growth statistic each year in our Refined Model.

Analysis of intervention plan

We used our Refined Model to forecast access to water in India fifteen years into the future with and without these interventions in various combinations. First, we projected water access with and without the interventions without accounting for climate change. These results are presented in Table 5 and shown graphically in Figure 4.

Table 5. Projected access to water in India (2031) with and without interventions and climate change

Conditions	Access to Water (% of Population)
No intervention; not controlled for climate change	94.4
With interventions; not controlled for climate change	98.2
No interventions; controlled for climate change	92.1
With interventions; controlled for climate change	95.9

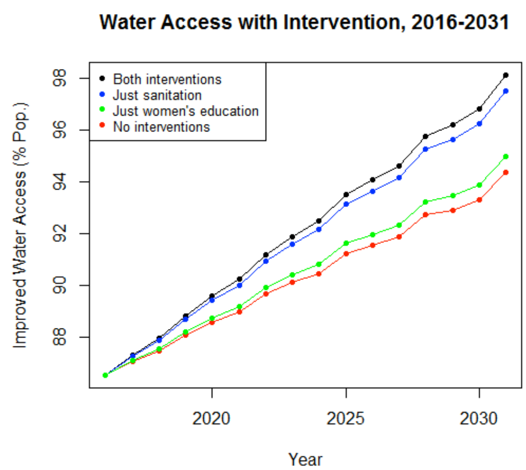

Figure 4. Effects of interventions on access to water in India

Finally, we re-ran this forecast to project Indian access to water with and without the interventions and with our model of climate change introduced. Again, the results are presented in Table 5, and these results are shown graphically in Figure 5.

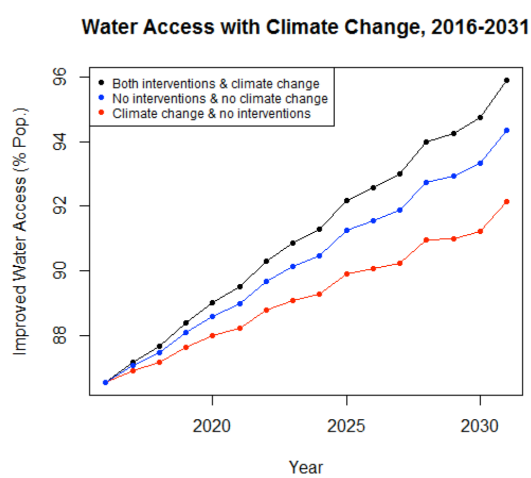

Figure 5. Effects of interventions on access to water in India, controlled for climate change

Discussion

Though we are not convinced that our model accurately measures the precise amount of clean water available in India, we are convinced that it accurately demonstrates the relative impacts of the various predictors we considered. We note that both of the interventions we modeled each improved access to water in India over the timeframe considered, and that the conjunction of both interventions resulted in a greater improvement than either did alone.

When we introduced climate change into the model, the prospects of access to water worsened. However, the introduction of both interventions completely mitigated the predicted effects of climate change, resulting in greater access to water than without the interventions, even with climate change excluded from the forecast.

This gives us confidence that our intervention plan would significantly improve access to clean water in India. However, we must also consider the possible impacts of each intervention on neighboring regions or on the ecosystem as a whole. Fortunately, we anticipate that most interventions will either have no measurable impact on other regions or will have a net positive impact:

- Because the bulk of India's river water is deposited into oceans rather than traversing other countries, collection of precipitation will not deprive downstream regions of an essential water supply.
- Removal of pollution from India's rivers will result in a decrease in pollution of ocean waters.
- Reduction of population growth will place less strain on the regional ecology.
- Increased women's education will open up new business opportunities, growing GDP within India and fostering international trade.

We do offer this word of caution, however: though we designed our interventions to include alternative energy sources that could combat climate change, it is possible that if only some of the interventions are pursued, or if they are pursued incompletely, an increase in India's economic productivity could increase carbon dioxide emissions, driving increased climate change.

Conclusions

India's water crisis is not caused by any single factor. Instead, it is an interdisciplinary problem resulting from a perfect storm of otherwise unrelated variables: political, economic, social, geophysical, meteorological, and so on. While our model does not (indeed, probably can not) account for all of these factors, it used a representative cross-section of the available data to show inexact values for India's water security, but clearly-defined trends in India's water security.

Our model shows that it is possible to improve India's water situation and that the most important factors in need of intervention are socioeconomic. Even though climate change will worsen the problem for India, as our model shows, the greatest challenge for India is to develop the cultural attitudes and to build the critical infrastructure necessary to guarantee access to clean water to all of its people.

Future Work

Due to time constraints, we were unable to include every variable in our model that we considered. Furthermore, published data on access to water in India belie the severity of India's water crisis. As such, though we feel our model was successful in our goal of demonstrating the effects of our variables of interest upon the water supply, we also feel that there is more work to be done to expand this into a more robust model.

We would like to address several key areas in future work. First, some of the necessary data in our model was missing from our publicly-available sources, so any future expansion of this work should strive to expand the data pool for the regression. Even more importantly, though the initial twenty predictors of water access we used were carefully considered to be a representative cross-section of the factors that influence water security, we would like to reconstruct the model using a much larger pool of variables upon which our variable selection function could act. In particular, some of our variables (including air pollution and access to electricity) were used as proxy variables that we assume correlated strongly with variables of interest (such as water pollution and technological advancement) for which we could not find quality data. In the next model, these proxy variables should be replaced with more direct measures. We would also like to include more quantifiable information about Indian culture in the expanded model.

Because the model was not constructed in a country-specific manner, it would be useful to apply the same model to other countries or regions to develop a more complete portrait of the world's water supply. Results from such an analysis could lead to the development of a model accounting for changes not only to the national water supply of a particular country but to the international water supply. It could also lead to the development of a concerted international intervention against water scarcity.

Finally, moving beyond expansions to the current model, we would like to construct an entirely new model to determine if and how the predictors of water scarcity are interrelated. Additionally, such a model should account for the possibility of transient or catastrophic events such as droughts. This could lead to a more refined intervention strategy by targeting specific variables with the goal of creating a cascade effect.

No single model will likely ever be able to fully account for the complexity of water scarcity, but through efforts to develop increasingly accurate and refined models, the world can inform its countries and its peoples about continuously improved management strategies so that, in time, no human being should ever have to worry about access to clean water.

References

[1] All About India. "Ganges River Pollution." *The Complete Guide to India.* <http://www.all-about-india.com/Ganges-River-Pollution.html> (accessed 31 January, 2016)

[2] Almeida, J. & Dias, R (January, 2016). "Impact of Climate Change on Indian Water Resources." *Journal of Global Resources*. <http://isdesr.org/sites/vol/jan_2016/6.pdf> (accessed 31 January, 2016)

[3] Appelgren, B.G (November 1997). "Proceedings of the Second Expert Consultation on National Water Policy Reform in the Near East; Appendix 3: Keynote Paper - Management of Water Scarcity: National Water Policy Reform in Relation to Regional Development Cooperation." *FAO Corporate Document Repository*. <http://www.fao.org/docrep/006/ad456e/ad456e0a.htm> (accessed 29 January, 2016)

[4] Balaiah, D., et al. (April, 2001). "Fertility Attitudes and Family Planning Practices of Men in a Rural Community of Maharashtra." *The Journal of Family Welfare*. <http://medind.nic.in/jah/t01/i1/jaht01i1p56g.pdf> (accessed 1 February, 2016)

[5] Barro, R., & Lee, J. (2015). "Projections of Educational Attainment by Country." *Barro-Lee Educational Attainment Dataset*. <http://barrolee.com/data/oup_download_c.htm> (accessed 30 January, 2016)

[6] Biswas, S. (October, 2014). "Why India's sanitation crisis needs more than toilets." *BBC News*. <http://www.bbc.com/news/world-asia-india-29502603> (accessed 31 January, 2016)

[7] Brooks, N (August, 2007). "Imminent Water Crisis in India." *Arlington Institute*. <http://www.arlingtoninstitute.org/wbp/global-water-crisis/606> (accessed 31 January, 2016)

[8] Coleridge, S.T. (1798). "The Rime of the Ancient Mariner." Republished by *The Poetry Foundation.* <http://www.poetryfoundation.org/poem/173253> (accessed 31 January, 2016)

[9] Gallup, J.L., Mellinger, A.D., & Sachs, J.D (November 2007). "Geography Datasets." *Center for International Development at Harvard University*. <http://www.cid.harvard.edu/ciddata/geographydata.htm> (accessed 30 January, 2016)

[10] Knoema (2015). "World GDP Ranking 2015." *Knoema*. <http://knoema.com/nwnfkne/world-gdp-ranking-2015-data-and-charts> (accessed 31 January, 2016)

[11] Kulkarni, V. "Overcoming Water Scarcity." *UNICEF India*. <http://unicef.in/Story/516/Overcoming-Water-Scarcity> (accessed 29 January, 2016)

[12] Kummu, M., de Moel, H., Ward, P.J., & Varis, O. (June, 2011) "How Close Do We Live to Water? A Global Analysis of Population Distance to Freshwater Bodies." *PLoS ONE*. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3110782/> (accessed 29 January, 2016)
[13] Luthra, S., & Kundu, A. (August, 2013). "India's Water Crisis: Causes and Cures." *The National Bureau of Asian Research*. <http://www.nbr.org/research/activity.aspx?id=356> (accessed 31 January, 2016)
[14] McMichael, C.E. (September, 2014). "Freshwater." *The Encyclopedia of Earth*. <http://www.eoearth.org/view/article/152861/> (accessed 29 January, 2016)
[15] Mohan, K. (October, 2013). "Condom Use in India Drops By Nearly 40 Percent In 5 Years As Use Of Women-Centric Contraceptives Rises." *International Business Times.* <http://www.ibtimes.com/condom-use-india-drops-nearly-40-percent-5-years-use-women-centric-contraceptives-rises-1429642> (accessed 31 January, 2016)
[16] Nation Master. "Education > School enrollment, secondary, female > % gross: Countries Compared." *Nation Master.* <http://www.nationmaster.com/country-info/stats/Education/School-enrollment%2C-secondary%2C-female/%25-gross#country> (accessed 31 January, 2016)
[17] Palmer, L. (2014). "Faminie Is a Feminist Issue: To feed the world, start by teaching girls to read." *Slate*. <http://www.slate.com/articles/health_and_science/feed_the_world/2014/04/educate_women_and_save_babies_how_to_control_population_and_end_hunger.html> (accessed 1 February, 2016)
[18] Snyder, S. "Water in Crisis - India." *The Water Project.* <https://thewaterproject.org/water-in-crisis-india> (accessed 31 January, 2016)
[19] Trading Economics (2016). "India GDP per capita 1960-2016." *Trading Economics* <http://www.tradingeconomics.com/india/gdp-per-capita> (accessed 1 February, 2016)
[20] United Nations Department of Economic and Social Affairs (2015). "World Population Prospects: Key findings & advance tables: 2015 revision." *United Nations Department of Economic and Social Affairs*. <http://esa.un.org/unpd/wpp/Publications/Files/Key_Findings_WPP_2015.pdf> (accessed 29 January, 2016)
[21] United Nations Water (March 2007). "Coping With Water Scarcity: Challenge of the Twenty-First Century." *Food and Agriculture Organization of the United Nations*. <http://www.fao.org/nr/water/docs/escarcity.pdf> (accessed 29 January, 2016)
[22] USDA (August, 2014). "India's Agricultural Exports Climb to Record High." *United States Department of Agriculture Foreign Agricultural Service.* <http://www.fas.usda.gov/data/india-s-agricultural-exports-climb-record-high> (accessed 31 January, 2016)
[23] USGS. "Irrigation water use." *The USGS Water Science School*. <http://water.usgs.gov/edu/wuir.html> (accessed 31 January, 2016)
[24] WHO. "Health through safe drinking water and basic sanitation." *World Health Organization*. <http://www.who.int/water_sanitation_health/mdg1/en/> (accessed 31 January, 2016)
[25] WHO (2014). "Progress on Drinking Water and Sanitation: 2014 Update." *World Health Organization*. <http://www.wssinfo.org/fileadmin/user_upload/resources/JMP_report_2014_webEng.pdf> (accessed 31 January, 2016)
[26] WHO, UNICEF (2014). "Progress on sanitation and drinking-water - 2014 update." *World Health Organization*. <http://apps.who.int/iris/bitstream/10665/112727/1/9789241507240_eng.pdf?ua=1> (accessed 29 January, 2016)
[27] WHO, UNICEF (2015). "Progress on sanitation and drinking-water: 2015 update and MDG Assessment" *World Health Organization*. <http://www.wssinfo.org/fileadmin/user_upload/resources/JMP-Update-report-2015_English.pdf> (accessed 29 January, 2016)
[28] World Bank (2015). "Access to electricity (% of population)." *World Bank Data*. <http://data.worldbank.org/indicator/EG.ELC.ACCS.ZS> (accessed 30 January, 2016)
[29] World Bank (2015). "Annual freshwater withdrawals, agriculture (% of total freshwater withdrawal)." *World Bank Data.* <http://data.worldbank.org/indicator/ER.H2O.FWAG.ZS> (accessed 30 January, 2016)

[30] World Bank (2015). "Annual freshwater withdrawals, domestic (% of total freshwater withdrawal)." *World Bank Data.* <http://data.worldbank.org/indicator/ER.H2O.FWDM.ZS> (accessed 30 January, 2016)
[31] World Bank (2015). "Annual freshwater withdrawals, industry (% of total freshwater withdrawal)." *World Bank Data.* <http://data.worldbank.org/indicator/ER.H2O.FWIN.ZS> (accessed 30 January, 2016)
[32] World Bank (2015). "Average precipitation in depth (mm per year)." *World Bank Data*. <http://data.worldbank.org/indicator/AG.LND.PRCP.MM> (accessed 30 January, 2016)
[33] World Bank (2015). "GDP at market prices (current US$)." *World Bank Data*. <http://data.worldbank.org/indicator/NY.GDP.MKTP.CD> (accessed 30 January, 2016)
[34] World Bank (2015). "GDP per capita (current US$)." *World Bank Data*. <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD> (accessed 30 January, 2016)
[35] World Bank (2015). "Health expenditure, public (% of total health expenditure)." *World Bank Data.* <http://data.worldbank.org/indicator/SH.XPD.PUBL> (accessed 30 January, 2016)
[36] World Bank (June, 2013). "India: Climate Change Impacts." *The World Bank*. <http://www.worldbank.org/en/news/feature/2013/06/19/india-climate-change-impacts> (accessed 31 January, 2016)
[37] World Bank (2015). "Improved sanitation facilities (% of population with access)." *World Bank Data*. <http://data.worldbank.org/indicator/SH.STA.ACSN> (accessed 30 January, 2016)
[38] World Bank (2015). "Improved water source (% of population with access)." *World Bank Data*. <http://data.worldbank.org/indicator/SH.H2O.SAFE.ZS> (accessed 30 January, 2016)
[39] World Bank (2015). "PM2.5 air pollution, mean annual exposure (micrograms per cubic meter)." *World Bank Data* <http://data.worldbank.org/indicator/EN.ATM.PM25.MC.M3> (accessed 30 January, 2016)
[40] World Bank (2015). "Population density (people per sq. km of land area)." *World Bank Data*. <http://data.worldbank.org/indicator/EN.POP.DNST> (accessed 30 January, 2016)
[41] World Bank (2015) "Population growth (annual %)." *World Bank Data*. <http://data.worldbank.org/indicator/SP.POP.GROW> (accessed 30 January, 2016)
[42] World Bank (2015). "Population, total." *World Bank Data.* <http://data.worldbank.org/indicator/SP.POP.TOTL> (accessed 30 January, 2016)
[43] World Bank (2015). "Renewable internal freshwater resources per capita (cubic meters)." *World Bank Data*. <http://data.worldbank.org/indicator/ER.H2O.INTR.PC> (accessed 30 January, 2016)
[44] World Bank (2015). "Rural population (% of total population)." *World Bank Data.* <http://data.worldbank.org/indicator/SP.RUR.TOTL.ZS> (accessed 30 January, 2016)
[45] World Bank (June, 2013). "Turn Down the Heat: Climate Extremes, Regional Impacts, and the Case for Resilience." *The World Bank*. <http://www.worldbank.org/content/dam/Worldbank/document/Full_Report_Vol_2_Turn_Down_The_Heat_%20Climate_Extremes_Regional_Impacts_Case_for_Resilience_Print%20version_FINAL.pdf> (accessed 31 January, 2016)