A Joint Model for Water Scarcity Evaluation

Jingyuan Wang and Li Li(✉)

School of Computer and Information Science,
Southwest University, Chongqing, China
wjykim@email.swu.edu.cn, lily@swu.edu.cn

Abstract. To make a real difference for our thirsty planet, we establish the water demand-supply model and the Advanced Water Poverty Index (AWPI). First, we develop a dynamic demand-supply model to measure the ability of a region to satisfy its water consumption. On the demand side, we fit agricultural and industrial water needs by the Grey Verhulst prediction model, then we consider domestic needs through the Logistic Growth model of total population and the Regression model of residential needs per capita. On the supply side, we estimate the impacts of multiple factors such as utilized internal river and rainfall, desalinated seawater and purified sewage. In the experiments, we use the sensor data from the World Bank. Also, the stability of our model has been proved by the evaluation. Second, we analyze the types of water scarcity by improving the Water Poverty Index to the Advanced Water Poverty Index, and we creatively add population as the sixth key component. The prediction can be used as an important indicator for the government to take some specific intervention measures to help alleviate the severe water shortage and achieve sustainable development of water resources.

Keywords: Water scarcity · Grey Verhulst prediction model · Logistic Growth model · AWPI

1 Introduction

Lack of sufficient available water resources to meet living and production needs, 1.6 billion people are suffering from water scarcity [8,9]. Therefore, all the countries are taking strategies to guard our thirsty planet.

Extracting more and more clean water for agricultural, industrial and residential purposes heavily threats the health of aquatic ecosystems and the life those ecosystems support. And the explosive population growth is exacerbating this matter. How to thoroughly analyze the degree and the causes of water scarcity has become a worldwide issue attracting more and more concern [5,7].

We make some necessary assumptions (for simplifying a realistic model) as follows.

- We only take the water demand of industry, agriculture and domestic into consideration.
- The consumption of water is determined and can be measured by statistic method.

© Springer International Publishing AG 2017
S. Song et al. (Eds.): APWeb-WAIM 2017 Workshops, LNCS 10612, pp. 11–20, 2017.
https://doi.org/10.1007/978-3-319-69781-9_2

– The country or region chosen is relatively stable.

Our contributions are as follows.

– **Modeling the water demand and supply to evaluate the scarcity without intervention.** For one thing, we develop the Grey Verhulst prediction model, the Logistic Growth model and the Regression model to analyze agricultural, industrial and residential water demand. For another, we estimate the water supply through utilized internal river and rainfall, desalinated seawater and purified sewage.
– **Analyzing the causes of water scarcity.** To distinguish the types of water scarcity, we establish the Advanced Water Poverty Index (AWPI), in which we consider resource, access, capacity, use, environment, and population as six key components.

2 Notations

In this section, we show the variables and parameters to be used in our work. They are fundamental and vital to the models in the following sections.

We list the variables and parameters used for constructing the models as follows. Notations and their definitions are shown in Table 1.

Table 1. Notations used in the paper

Symbol	Definition	Unit
$r(x)$	The population growth rate	Unitless
x_m	The maximum population which natural resources and environmental conditions can accommodate	Unitless
D_t, D_{at}, D_{it}	Dynamic total, agricultural and industrial water demand over time	Billion cubic meters
U_{dt}, P_{dt}	Dynamic residential water demand per capita and total population over time	Billion cubic meters, unitless
S_t	Dynamic water supply over time	Billion cubic meters
N	Natural water source capacity	Billion cubic meters
A_t, D_{kt}	Dynamic utilization rate of natural water source and sewage of k industry over time	Unitless
SE_t, p_t	Dynamic desalinated seawater and water pollution over time	Billion cubic meters
R_k	Purification rate of k industry	Unitless
ρ	Relative supply-demand ratio	Unitless
RI, AI, CI, UI, EI, PI	Resource, access, capacity, use, environment and population index	Unitless

3 The Demand-Supply Model

3.1 The Demand Model

To analyze the condition of water demand, we must consider the agricultural consumption, industrial consumption and domestic consumption, respectively. Because the monotonically increasing sequence corresponds to the GM (1, 1) model while S-shaped sequence corresponds to the Grey Verhulst prediction model [1,13], we randomly choose five countries to judge the shape of their agricultural and industrial usage to match a proper model. As for the domestic usage, we develop the Logistic Growth model [12] to predict the population growth, and we establish the Regression model to predict the domestic water consumption per capita.

3.1.1 The Grey Verhulst Prediction Model

First, we use the annual freshwater withdrawals, industry/agriculture from World Bank database[1] (Romania, Cyprus, Iraq, India and China) (Fig. 1), which apparently justifies a non-monotonic increase; thus we establish the Grey Verhulst prediction model instead of the GM (1, 1) Model.

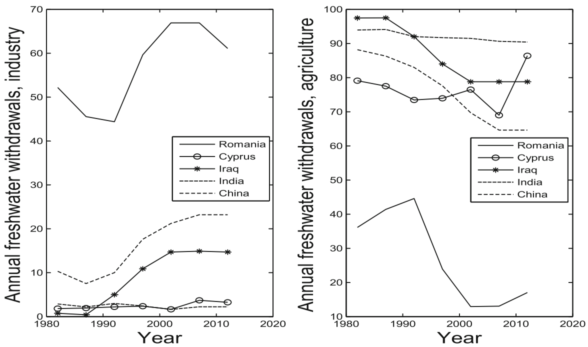

Fig. 1. The annual freshwater withdrawals, industry/agriculture

We select a sequence

$$x^{(0)} = x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n) \tag{1}$$

and we can get another sequence within the following relationship

$$\begin{aligned} x^{(1)} &= x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n) \\ &= x^{(0)}(1), x^{(0)}(1) + x^{(0)}(2), ..., x^{(0)}(1) + ... + x^{(0)}(n) \end{aligned} \tag{2}$$

[1] http://data.worldbank.org.

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i),\ k = 1, 2, ..., n$

Also,

$$z^{(1)} = z^{(1)}(2), z^{(1)}(3), ..., z^{(1)}(n) \tag{3}$$

where the correction among these variables is

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1),\ k = 2, 3, ..., n \tag{4}$$

We then obtain $x^{(0)} + az^{(1)} = b(z^{(1)})^2$ as well as the differential equation $\frac{dx^{(1)}}{dt} + ax^{(1)} = b(x^{(1)})^2$.

To estimate the parameters, we make

$$u = [a, b]^T, B = \begin{bmatrix} -z^{(1)}(2) & (z^{(1)}(2))^2 \\ -z^{(1)}(3) & (z^{(1)}(3))^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & (z^{(1)}(n))^2 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \tag{5}$$

then we get that $\hat{u} = [\hat{a}, \hat{b}]^T = (B^T B)^{-1} B^T Y$ though least squares method.

Therefore the solution to the differential equation is

$$\hat{x}^{(1)}(k+1) = \frac{\hat{a}x^{(0)}(1)}{\hat{b}x^{(0)}(1) + \left[\hat{a} - \hat{b}x^{(0)}(1)\right] e^{\hat{a}k}} \tag{6}$$

so the accumulated reduction series is

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{7}$$

3.1.2 The Regression Model

Then we consider the domestic water consumption in which the population growth plays a key role. We develop the Logistic model to estimate the population size and establish the Regression model to fit the personal water consumption.

The population growth rate varies as $r(x) = r(1 - \frac{x}{x_m})$ where r is a constant, x is the population, and x_m is the maximum population which natural resources and environmental conditions can accommodate.

Then, we get the ordinary differential equations

$$\begin{cases} \frac{dx}{dt} = r(1 - \frac{x}{x_m})x \\ x(t_0) = x_0 \end{cases} \tag{8}$$

and its solution

$$x(t) = \frac{x_m}{1 + (\frac{x_m}{x_0} - 1)e^{-r(t-t_0)}} \tag{9}$$

Third, we use the Regression model to fit the domestic water consumption per capita. As a benchmark model, we simply assume the time series data meets a linear relationship. It's rational because the difference between this fundamental model and the improved one just exists in the complexity of a function.

The dependent variable y represents domestic water consumption per capita (cubic meter/person), and the independent variable x represents time (year).

We assume that

$$y = \beta_0 + \beta_1 x \tag{10}$$

Then we calculate the estimates of the parameters by minimizing the differences between data points and the line:

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \\ \hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \end{cases} \tag{11}$$

where $\bar{x} = \sum_{i=1}^{n} x_i$, $\bar{y} = \sum_{i=1}^{n} y_i$. Thus we obtain the regression equation

$$y = \hat{\beta}_0 + \hat{\beta}_1 x \tag{12}$$

With these three dynamic consumption processes, we get water demand sequence which changes over time and is influenced by population growth and economic development as follows:

$$D_t = D_{at} + D_{it} + U_{dt} \cdot P_{dt} \tag{13}$$

3.2 The Supply Model

We divide the water supply into two parts: clean water from natural source which inherently depends on the geological, topographical and ecological conditions, and fresh water from regeneration (different industries have different renewable rate). Plus, increasingly serious climate growth and pollution reduce the water provision.

The key equation is

$$S_t = (N \cdot A_t + SE_t + \sum_{kt} R_k \cdot D_{kt})(1 - p_t) \tag{14}$$

where S represents the total supply quantitation, N represents natural water source, A represents access to clean and fresh water (rely on techniques, management and infrastructures), SE represents the water from desalination plants, R represents renewable rate of water, D represents water needs, p represents the degree of environmental pollution, and $k = a, i, d$ represents agriculture, industry and domestic, respectively.

3.3 The Demand-Supply Model

In conclusion, we define ρ as an indicator which shows the relative situation of water supply and water demand. In other words, it is a measure of the ability of a region to provide clean water to meet the needs of its population, and it contains all the dynamic factors as mentioned above.

$$\rho = \frac{S_t}{D_t} = \frac{(N \cdot A_t + SE_t + \sum_{kt} R_k \cdot D_{kt})(1 - p_t)}{D_{at} + D_{it} + U_{dt} \cdot P_{dt}} \tag{15}$$

When the value of ρ is in the following interval, the ability to provide clean water to meet the needs and corresponding degree of water scarcity are shown in Table 2.

Table 2. The results of traffic flow prediction with 15-min time interval

The degree of water scarcity	The value of ρ
Very poor ability and severe water scarcity	$\leq$0.7
Low ability and serious water scarcity	0.7–1
Medium ability and little water scarcity	1–1.2
High ability and no water scarcity	$\geq$1.2

3.4 Evaluation

In this section, we evaluate the two pivotal models both in theory and in practice to prove its stability and feasibility.

3.4.1 Evaluating the Grey Verhulst Prediction Model

Making $\varepsilon(k)$ represents relative error, we then calculate

$$\varepsilon(k) = \frac{x^0(k) - \hat{x}^0(k)}{x^0(k)}, k = 1, 2, ..., n \tag{16}$$

where $\hat{x}^0(1) = x^0(1)$. If $\varepsilon(k) < 0.5$, we consider that it meets the general requirements; and if $\varepsilon(k) < 0.2$, we consider that it meets a high requirement.

We choose Romania and Cyprus as the test objects. Their agricultural and industrial water consumptions from the year 1982 to 2012 are shown in Table 3. We can see that only two relative errors are beyond a reasonable interval, that is, our model possesses a high accuracy.

Table 3. Real/predicted water needs of agriculture/industry in Romania and Cyprus

Year	Real needs	Predicted needs	Residuals	Relative error	Year	Real needs	Predicted needs	Residuals	Relative error
Romania_Agricutrue					*Romania_Industry*				
1982	36.12	36.12	0	0	1982	52.18	52.18	0	0
1987	41.41	45.156	−3.746	0.090	1987	45.56	47.702	−2.142	0.047
1992	44.59	34.337	10.252	0.229	1992	44.39	51.234	−6.844	0.154
1997	23.92	26.110	−2.190	0.091	1997	59.63	55.027	4.602	0.077
2002	12.93	19.854	−6.924	**0.535**	2002	66.91	59.100	7.809	0.116
2007	13.04	15.097	−2.057	0.157	2007	66.91	63.475	3.434	0.051
2012	17.03	11.480	5.549	0.325	2012	61.08	68.174	−7.094	0.116
Cyprus_Agricutrue					*Cyprus_Industry*				
1982	79.1	79.1	0	0	1982	1.83	1.83	0	0
1987	77.5	73.644	3.855	0.049	1987	1.94	1.805	0.134	0.069
1992	73.45	74.618	−1.168	0.015	1992	2.212	2.043	0.168	0.076
1997	73.93	75.605	−1.675	0.022	1997	2.37	2.311	0.058	0.024
2002	76.48	76.604	−0.124	0.001	2002	1.663	2.615	−0.952	**0.572**
2007	68.98	77.617	−8.637	0.125	2007	3.704	2.958	0.745	0.201
2012	86.41	78.643	7.766	0.089	2012	3.261	3.347	−0.086	0.026

3.4.2 Evaluating the Regression Model

This is a significant test where we take $H_0 : \beta = 0$ as the null hypothesis and $H_1 : \beta \neq 0$ as the alternative hypothesis. We construct a test statistic as follows:

$$T = \frac{\hat{\beta}_1}{\hat{\sigma}} \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{17}$$

where $\hat{\sigma} = \frac{1}{n-2} \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2$

When H_0 is met, T belongs to the distribution $t(n-2)$. Thus at the α confidence level, the rejection region is $\left\{ T \geq t_{\alpha/2}(n-2) \right\}$.

4 The Advanced Water Poverty Index (AWPI)

In this section, we improve the water poverty index by adding population as the sixth key factor, for population growth rate has already continually exacerbated the severe water scarcity problem [2–4,6,10,11].

We adopt the Advanced Water Poverty Index (AWPI) to measure the water supply of a country or a region, where we take the six components (Table 4) into account.

For every index, we decompose it into several sub-indexes and calculate it as below:

Table 4. The key components of the AWPI and their definitions

Components	Definitions
Resource (R)	Physical amount and distribution of internal river flows and groundwater from rainfall
Access (A)	Availability to natural water source
Capacity (C)	Ability to get clean fresh water
Use (U)	The comprehensive efficiency of water uses for agricultural, industrial and residential purposes
Environment (E)	Stability of the whole ecosystem
Population (P)	Population growth rate and population density

– Resource (The World Bank, 2015):

$$Resource\,Index\,(RI) = renewable\,internal\,freshwater\,resources\,index \tag{18}$$

– Access:

$$Access\,Index\,(AI) = \frac{S+W}{2} \tag{19}$$

where S is the sanitation coverage index (The World Bank, 2014), W is the percentage of the population with access to improved water sources, urban (The World Bank, 2015).

– Capacity:

$$Capacity\,Index\,(CI) = \frac{M+T_c+R+E+(1-A)}{4} \tag{20}$$

where M is the evaluation of management, T_c is the evaluation of technology advances, R is the recycle rate of used water, A is the average death age index.

$$\begin{aligned} E &= 1, if\,average\,schooling\,year\ \geq 12 \\ &= 0.8, if\,9 < averag\,schooling\,year < 12 \\ &= 0.6, if\,6 < average\,schooling\,year < 9 \\ &= 0.4, if\,3 < average\,schooling\,year < 6 \\ &= 0.2, if\,1 < average\,schooling\,year < 3 \\ &= 0, if\,average\,schooling\,year < 1 \end{aligned} \tag{21}$$

– Use:

$$Use\,Index\,(UI) = \frac{U-U_a-U_i-U_d}{3} \tag{22}$$

where $U_k = \frac{Actual\,Water\,Consumptions\,in\,k\,Industry}{Water\,Consumption\,in\,k\,Industry}$ and (k = a,i,d) represents agriculture, industry and domestic, respectively.

– Environment:

$$Environmental\,Index\,(EI) = \frac{(1-C)+G+T+(1-P_m)+(1-P_s)+F+H}{7} \quad (23)$$

where C is the evaluation of influence caused by climate change, G is geological condition, T is topographical feature, P_m is average exceedance of particulate matter, P_s is average exceedance of sewage, F is the evaluation of forest health, H is the evaluation of habitat and biodiversity health.

– Population:

$$Population\,Index\,(PI) = 1 - \frac{P_g + P_d}{2} \quad (24)$$

where P_g is the population growth rate, P_d is the population density.
In conclusion:

$$AWPI = RI + AI + CI + UI + EI + PI \quad (25)$$

5 Sensitivity

When we develop the demand model to predict agricultural and industrial demand for water, we use the data from the World Bank as the initial data to simulate the trend. However, all into the long time span of the data, it's difficult for the World Bank to investigate all developing countries one by one. Thus there may exist some inaccurate initial data, which will impact the accuracy of the prediction model. We call these inaccurate initial data "Outlier".

When testing our model, we find an obvious bias in the prediction of demand in 2002. Thus, we delete the data in 2002 and use the remaining data to re-predict the demand by our model. Take Romania as an example, from Table 5, it is clear that our model is not sensitive to outliers.

Table 5. Agricultural and industrial demand with/without outliers in Romania

Year	Agricultural demand			Industrial demand		
	With outliers	Without outliers	Change rate	With outliers	Without outliers	Change rate
1982	36.12	36.12	0.00%	52.18	52.18	0.00%
1987	45.157	46.019	1.90%	47.703	44.596	−6.50%
1992	34.337	34.595	0.70%	51.234	49.94	−2.50%
1997	26.11	26.006	−0.40%	55.027	55.923	1.60%
2002	19.854	19.55	−1.50%	59.101	62.624	6.00%

6 Conclusion

The key point to measure the ability of a region to satisfy its water consumption is whether the supply can meet the demand. To measure water supply capacity of a country, it necessitates consideration of the supply and demand situation respectively. In our paper, we separate the measure model into two parts. First, we incorporate the Grey Verhulst model, the Logistic Growth and the Regression model in the demand model. Then, we develop the supply model and take various factors into account.

Moreover, we adopt the Advanced Water Poverty Index (AWPI) to measure the water supply and analyze the reasons of water scarcity of a country or a region from another angle, where we consider six key components and define each one clearly. The experimental results on World Bank dataset have proved the stability and feasibility of our models. According to the evaluation results, all the countries can take some specific intervention measures to help alleviate the severe water shortage and achieve sustainable development of water resources.

References

1. Cui, J., Liu, S.F., Zeng, B., Xie, N.M.: Parameters characteristics of grey Verhulst prediction model under multiple transformation. Kongzhi Yu Juece/Control Decis. **28**(4), 605–608 (2013)
2. Feitelson, E., Chenoweth, J.: Water poverty: towards a meaningful indicator. Water Policy **4**(3), 263–281 (2002)
3. Garriga, R.G., Foguet, A.P.: Improved method to calculate a water poverty index at local scale. J. Environ. Eng. **136**(11), 1287–1298 (2010)
4. Gong, L., Jin, C.: Urban water security evaluation system based on water poverty index. Shuili Fadian Xuebao/J. Hydroelectr. Eng. **33**(6), 84–90 (2014)
5. Kahil, M.T., Diner, A., Albiac, J.: Modeling water scarcity and droughts for policy adaptation to climate change in arid and semiarid regions. J. Hydrol. **522**, 95–109 (2015)
6. López Álvarez, B., Ramos Leal, J.A.: Water poverty index in subtropical zones: the case of Huasteca Potosina, Mexico. Rev. Int. Contam. Ambient. **31**(2), 173–184 (2015)
7. Mekonnen, M.M., Hoekstra, A.Y.: Four billion people facing severe water scarcity. Sci. Adv. **2**(2), e1500323 (2016)
8. Postel, S.L.: Entering an era of water scarcity: the challenges ahead. Ecol. Appl. **10**(10), 941–948 (2008)
9. Santini, G.: Coping with water scarcity. Unesco Tech. Doc. Hydrol. **58**(4), 77–98 (2015)
10. Sullivan, C.: Calculating a water poverty index. World Dev. **30**(7), 1195–1210 (2002)
11. Sullivan, C., Meigh, J., Lawrence, P.: Application of the water poverty index at different scales a cautionary tale. Water Int. **31**(3), 412–426 (2006)
12. Tsoularis, A., Wallace, J.: Analysis of logistic growth models. Math. Biosci. **179**(1), 21–55 (2002)
13. Wang, Z.X., Dang, Y.G., Liu, S.F.: Unbiased grey Verhulst model and its application. Syst. Eng.-Theory. **29**(10), 138–144 (2009)

http://www.springer.com/978-3-319-69780-2

Web and Big Data
APWeb-WAIM 2017 International Workshops: MWDA, HotSpatial, GDMA, DDC, SDMA, MASS, Beijing, China, July 7-9, 2017, Revised Selected Papers
Song, S.; Renz, M.; Moon, Y.-S. (Eds.)
2017, XV, 268 p. 83 illus., Softcover
ISBN: 978-3-319-69780-2