Water!!!

TEAM #51350

February 1, 2016

Abstract

Key Words:

Contents

1 Introduction

2 Nomenclatures

Table 1: Nomenclatures System

P	Population
GDP	Gross Domestic Product
PCGDP	Per Capita Gross Domestic Product
GAP	Cross Agricultural Product
AWC	Agricultural Water Consumption per year
IWC	Industrial Water Consumption per year
DWC	Domestic Water Consumption per year
TWR	Total Water Resource
SWR	total Surface Water Resource
UWR	total Underground Water Resource
WWD	Waste Water Discharge
A	Annual water supplies per person

3 Model of water supply ability

When it come to the water supply ability of a region, a country or even the world. We often use the measurement called annual water supplies per person(A) for description[1]. We can set three levels to classify the ability of several regions:

level 1	$A > 1700$	Sufficient
level 2	$1700 > A > 1000$	stressful
level 3	$1000 > A$	scarce

To cover the internal dynamics of the water flow and the water storage change, we introduce following model.

3.1 Model Introduction

Water circulation is a rather complicated process, which make it almost impossible for us to design a purely fundamental model to include all the variables

and their relations. Nevertheless, if we just collect all the data and using fitting method as our predicting model, it will be too trivial and old-fashioned. To solve this paradox, we introduce a phenomenological model which is quite normal in particle physics and other related field. Our model includes two main parts:

- prominent external parameters: a several statistical parameters about a region like population, GDP and so on.

- internal dynamic process: a several dynamical relations between external parameters and internal parameters like water consumption ,water recycle and evolution equations of water storage and related variables.

The prominent external parameters comes from the statistic numbers and it's fitness, while the internal dynamics process is mainly based on the relation between state parameters and external parameters and the water cycle process.

3.2 Model Structure
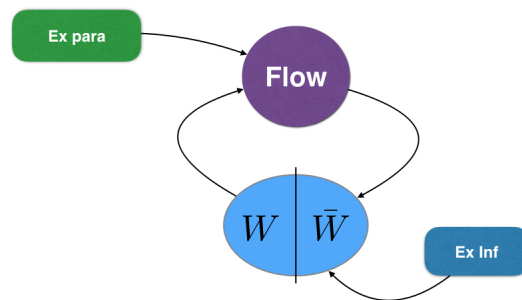
Our model's structure is pretty clear. We using some prominent external parameters as the input of our dynamic system. To determine the water supply ability of a region, the most important part is the water flow which means the total amount of industrial, agricultural and residential consumption. The season we can use the consumption to measure the supply it that they must be the same during a long period just like the electricity use equal to the electricity produce.

$$\langle \text{Consumption} \rangle = \langle \text{Supply} \rangle$$

The W and $\bar{W}$ is a crucial conception in our dynamic part, W stand for the total clean water resource over a period, while the $\bar{W}$ is the total wasted water resource which can be transfered into W after several processing steps.

3.3   Model Prominent Parameters

The prominent parameters are some thing the dynamics part heavily rely on, to find the prominent parameters, we need to investigate the relation between some alternative parameters and consumptions according to the past data. In other words, prominent parameters are not given before a specific research object is given, we just give some alternative parameters and using the past data to determine which is better related and which is less related to our dynamic part variables.

Generally, there are several important alternative parameters like: Population, GDP Per Capita are of course important statistic numbers. Moreover, Agricultural Irrigation Area play an important role in agricultural consumption, Iron and Steel Production is also important in industrial consumption, and Engel's coefficient is an important statistical number about residential living which make it important in residential consumption.

3.4   Model Dynamics

The dynamics part of our model is inspired by the water cycle(Figure 1)


Figure 1: water cycle[2]

3.5 How to apply our model

With good structure of our model and program, the application of our model can be done in a pretty clear way.

step 1 Using the data of alternative prominent parameters and consumption variables in the past to find the strength of correlation, Spearman or Pearson or other coefficients.

step 2 According to the analysis of step 1, we can determine the specific evolution equations of consumption variables which depend on the prominent parameters and water resource W and $\bar{W}$.

step 3 Using the past data of needed prominent parameters, we will be able to use fitness method to predict the future data of prominent parameters as the input of evolution equations.

step 4 Using the Iterative method, we can predict the variables in dynamics part thus we can get the Output – the water supply per person (A) to judge the water situation in a specific future time of a certain region.

4 China's water scarcity

According to the UN water scarcity map, China is a country with water stress,
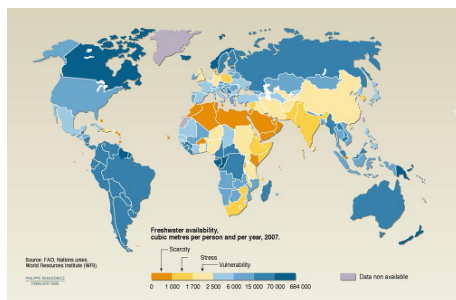
which make it a region where water is moderately overloaded. In our consideration, level 2, will provide more abundant behavior in a dynamical model(will it become water scarce or water sufficient in the future?). Thus we pick up China as our research object. To make our model more predictable and more reality connected.

Figure 2: water scarcity map[3]

4.1   Social drivers

4.2   Environmental drivers

5   China's water scarcity: Further Investigation and Prediction

For further investigation, we study with the data from National Bureau of Statistics of the People's Republic China[4].According to our models, we should predetermine the alternative prominent parameters and grab all the data and try to analyses them and they relation.

5.1   Prominent variables' tendency

Using proper assumption and the past data in databases. We can using fitness method to determine the time dependency of our alternative parameters. In our consideration, following factors especially eye-catching. For further con-

P	Population
PCGDP	Per Capita Gross Domestic Product
IA	Irrigation Area
ISP	Iron and Steel Production
ElP	Electricity Production.
EnC	Engel's Coefficient

venience, we will investigate their relation towards time at first.

Population   The population growth model is various, for best concern, we use Logistic Model as the Population evolution model thus we can use the solution–Standard logistic sigmoid function as our target function by adding three basic coefficients.

$$P(t) = \frac{K^P}{1 + \exp[-\lambda^P(t - t_0^P)]}$$

Per Capita Gross Domestic Product   As the largest developing country over the world, China has pretty high PCGDP growth rate. It seems like an exponential growth However, according to our data browse, some developed countries

have already reach some obstacles. Thus we think it will be suitable if we apply Logistic model in the fitness model of PCGDP.

$$\text{PCGDP}(t) = \frac{K^{\text{PCGDP}}}{1 + \exp[-\lambda^{\text{PCGDP}}(t - t_0^{\text{PCGDP}})]}$$

Irrigation Area According to our observation, there is a strong evidence about linear growth of Irrigation Area. Consider the complicity and the fact we are just required to predict about 15 years

$$\text{IA}(t) = A_0^{\text{IA}} + A_1^{\text{IA}}t + A_2^{\text{IA}}t^2$$

Iron and Steel Production

$$\text{ISP}(t) = A_0^{\text{ISP}} + A_1^{\text{ISP}}t + A_2^{\text{ISP}}t^2$$

Electricity Production

$$\text{ELP}(t) = A_0^{\text{ELP}} + A_1^{\text{ELP}}t + A_2^{\text{ELP}}t^2$$

Engel's Coefficient

$$\begin{cases} \text{Enc}_{\text{urban}}(t) = \frac{C^{\text{urban}}}{t - t_0^{\text{urban}}} \\ \text{Enc}_{\text{rural}}(t) = \frac{C^{\text{rural}}}{t - t_0^{\text{rural}}} \end{cases}$$

With the fitness method, we can get all the evolution of alternative prominent parameters in Figure 5.1, and the parameters in evolution equations in Table5.1:

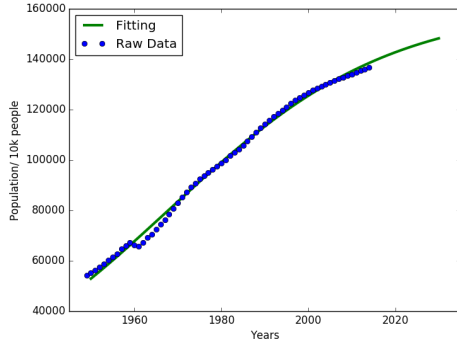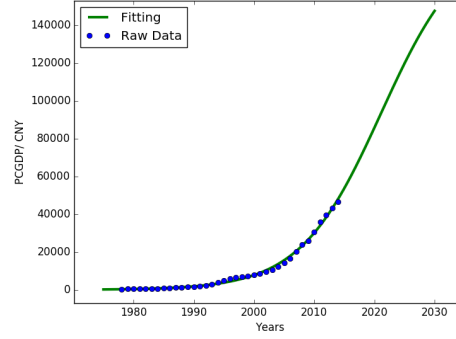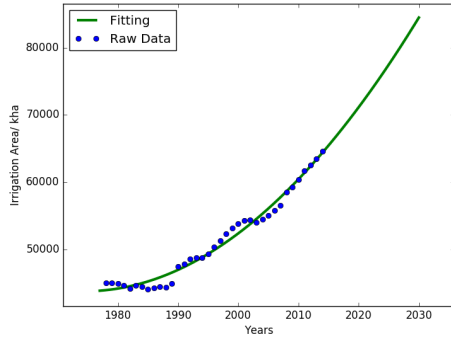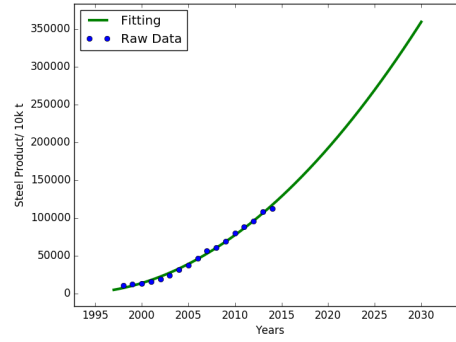| $\left[t_0^P, \lambda^P, K^P\right]$	[1968.14, 3.94%, 161162.404]
$\left[t_0^{\text{PCGDP}}, \lambda^{\text{PCGDP}}, K^{\text{PCGDP}}\right]$	[2021.12, 14.9%, 187060.984]
$\left[A_0^{\text{IA}}, A_1^{\text{IA}}, A_2^{\text{IA}}\right]$	[5.155e+07, -5.217e+04, 13.21]
$\left[A_0^{\text{ISP}}, A_1^{\text{ISP}}, A_2^{\text{ISP}}\right]$	[1.027e+09, -1.031e+06, 258.6]
$\left[A_0^{\text{ELP}}, A_1^{\text{ELP}}, A_2^{\text{ELP}}\right]$	[3.125e+08, -3.146e+05, 79.19]
$\left[C^{\text{urban}}, t_0^{\text{urban}}, C^{\text{rural}}, t_0^{\text{rural}}\right]$	[1938.24, 2700.15, 1916.02, 4218.31]

Table 2: evolution parameters

(a) $p(t)$

(b) $\text{PCGDP}(t)$

(c) $\text{IA}(t)$

(d) $\text{ISP}(t)$
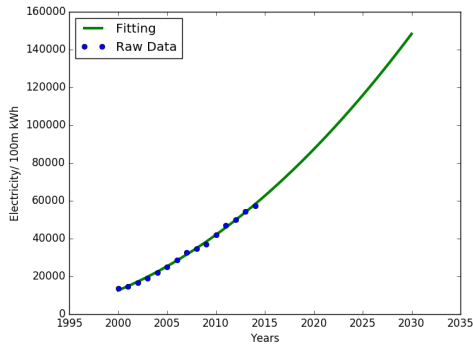
(e) $\text{ELP}(t)$
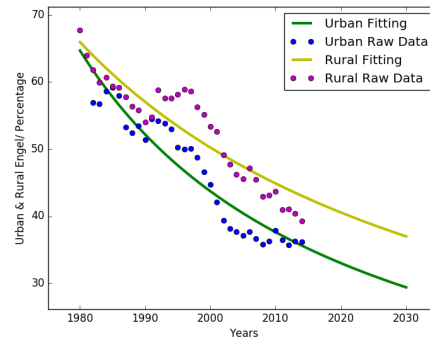
(f) $\text{Enc}(t)$

Figure 3: Parameters fit: blue dots means the existed data which we refer from the national database, green lines stand for the optimized curve and the prediction value in the future

5.2   Relation between prominent parameters and variables

To investigate the relation between prominent parameters and variables, we firstly introduce the conception of correlation. In statistics, dependence is any statistical relationship between two random variables or two sets of data, and correlation refers to any of a broad class of statistical relationships involving dependence.

Pearson product-moment correlation coefficient [5] is a measure of the linear correlation between two variables X, Y. If there is two datasets $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_n\}$, then the formula for r is[6]:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

Spearman's rank correlation coefficient is a nonparametric measure of statistical dependence between two variables. It assesses how well the relationship between two variables can be described using a monotonic function. If there are no repeated data values, a perfect Spearman correlation of $+1$ or -1 occurs when each of the variables is a perfect monotone function of the other[7][8]. The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranked variables[?]. For a sample of size n, the n raw scores X_i, Y_i are converted ranks x_i, y_i, and ρ is computed from:

$$\rho = 1 - \frac{6 \sum (x_i - y_i)^2}{n(n^2 - 1)}$$

Kendall rank correlation coefficient is a statistic used to measure the association between two measured quantities. Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be a set of observations of the joint random variables X and Y respectively, such that all the values of (x_i) and (y_i) are unique. Any pair of observations (x_i, y_i) and (x_j, y_j), where $i \neq j$, are said to be concordant if the ranks for both elements agree: that is, if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. They are said to be discordant, if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant. Thus the τ coefficient is defined as[10][11]:

$$\tau = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{\frac{1}{2}n(n - 1)}$$

Using coefficients above, we get a map about the relation between prominent parameters and consumption variables.

Data	Pearson (r_p, p)	Spearman (r_s, p)	Kendall (τ, p)
IWC $\sim P$	(0.669,0.0245)	(0.627,0.0388)	(0.455,0.0516)
IWC $\sim$ PCGDP	(0.614,0.0444)	(0.627,0.0388)	(0.455,0.0516)
IWC $\sim$ ELP	(0.648,0.0311)	(0.627,0.0388)	(0.455,0.0516)
IWC $\sim$ ISP	(0.653,0.0294)	(0.627,0.0388)	(0.455,0.0516)
AWC $\sim P$	(0.924,4.95e-5)	(0.927,3.97e-5)	(0.782,0.000815)
AWC $\sim$ PCGDP	(0.936,2.28e-5)	(0.927,3.97e-5)	(0.782,0.000815)
AWC $\sim$ IA	(0.927,4.10e-5)	(0.927,3.97e-5)	(0.782,0.000815)
DWC $\sim P$	(0.844,0.00108)	(0.836,0.00133)	(0.745,0.00141)
DWC $\sim$ PCGDP	(0.806,0.00276)	(0.836,0.00133)	(0.745,0.00141)
DWC $\sim$ Enc$_{\text{urban}}$	(-0.785,0.00417)	(-0.809,0.00256)	(-0.600,0.0102)
DWC $\sim$ Enc$_{\text{rural}}$	(-0.419,0.199)	(-0.315,0.345)	(-0.204,0.383)

Table 3: Coefficients: Coefficients of correlation in different standard

Coefficients Table

5.3   Fitness of past data using dynamic model

6   Intervention plan designing

7   Prediction with Intervention plan

8   Conclusion

References

[1] Falkenmark and Lindh 1976, quoted in UNEP/WMO.Climate Change 2001: Working Group II: Impacts,Adaptation and Vulnerability. UNEP. Retrieved 3 February 2009.

[2] trinityrivertexas.org, Living with the Trinity Lesson Plan 1: The Natural Water Cycle and the Urban Water Cycle

[3] http://www.unep.org/dewa/vitalwater

[4] http://www.stats.gov.cn

[5] https://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient#Definition

[6] http://www.statsoft.com/textbook/glosp.html#Pearson%20Correlation

[7] https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient

[8] Zwillinger, D. and Kokoska, S. (2000). CRC Standard Probability and Statistics Tables and Formulae. Chapman & Hall: New York. 2000. Section 14.7

[9] Myers, Jerome L.; Well, Arnold D. (2003). Research Design and Statistical Analysis (2nd ed.). Lawrence Erlbaum. p. 508

[10] https://en.wikipedia.org/wiki/Kendall_rank_correlation_coefficient

[11] W.R. Knight, A Computer Method for Calculating Kendalls Tau with Ungrouped Data, Journal of the American Statistical Association, Vol. 61, No. 314, Part 1, pp. 436-439, 1966