

# Water!!!

TEAM #51350

February 1, 2016

**Abstract**

**Key Words:**

## Contents

<b>1</b>	<b>Introduction . . . . .</b>	<b>3</b>
<b>2</b>	<b>Nomenclatures . . . . .</b>	<b>3</b>
<b>3</b>	<b>Model of water supply ability . . . . .</b>	<b>3</b>
3.1	Model Introduction . . . . .	4
3.2	Model Structure . . . . .	4
3.3	Model Prominent Parameters . . . . .	5
3.4	Model Dynamics . . . . .	5
3.5	How to apply our model . . . . .	6
<b>4</b>	<b>China's water scarcity . . . . .</b>	<b>6</b>
4.1	Social drivers . . . . .	8
4.2	Environmental drivers . . . . .	8
<b>5</b>	<b>China's water scarcity: Further Investigation and Prediction . . . . .</b>	<b>8</b>
5.1	Prominent variables' tendency . . . . .	8
5.2	Relation between prominent parameters and variables . . . . .	10
5.3	Fitness method in consumption variables . . . . .	13
<b>6</b>	<b>Intervention plan designing . . . . .</b>	<b>15</b>
<b>7</b>	<b>Prediction with Intervention plan . . . . .</b>	<b>15</b>
<b>8</b>	<b>Conclusion . . . . .</b>	<b>15</b>

# 1 Introduction

## 2 Nomenclatures

$P$	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
PCGDP	Per Capita Gross Domestic Product
IA	Irrigation Area
ISP	Iron and Steel Production
ELP	Electricity Production.
EnC	Engel's Coefficient
$A$	Annual water supplies per person

Table 1: Nomenclatures System

## 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

Table 2: water supply ability

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 variables: find out the relation between prominent external parameters and internal variables including Agriculture Water Consumption, Industrial Water Consumption and Domestic Water consumption.

The prominent external parameters comes from the statistic numbers and it's fitness, while the internal dynamics variables is mainly based on prominent parameters, which makes them indirectly depend on time in our model.

### 3.2 Model Structure

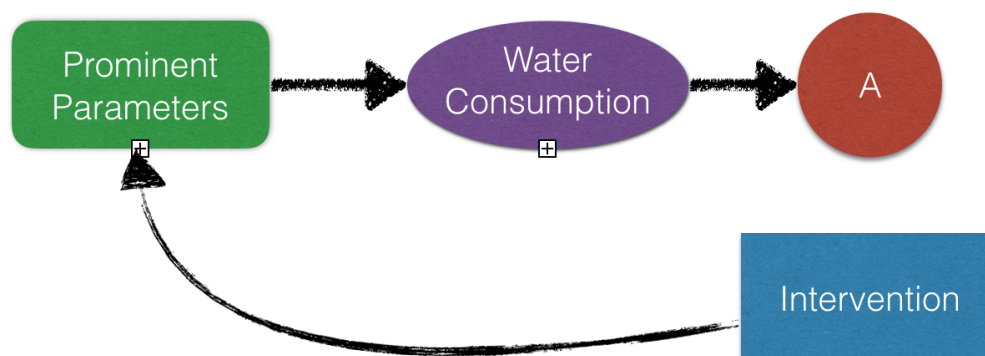


Figure 1: structure

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 transferred 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 2). Basically, our water is mainly refreshed through the precipitation. So it will be elegant if we using such water cycle process to build a mighty system about how the water flow and run with real data to determine the theoretical water consumption. However, after several data browse,

we find that using precipitation as part of the water storage's changing rate is ridiculous—The total water resource is steadily a certain ratio of precipitation. Therefore using prominent parameters as arguments of internal variables and find the relation according to the past data will be a proper choice.

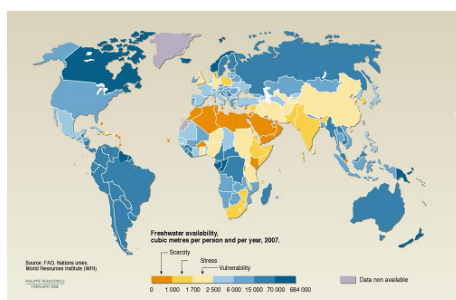
### 3.5 How to apply our model

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 past data of alternative prominent parameters, find proper relation between those parameters and time, therefore determine the time dependent equation of prominent parameters.
- step 2 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 3 According to the coefficients, structure the parameter dependent equations of consumption variables.
- step 4 With the equations construed above and the time dependent equation of prominent parameters, get the tendency of consumption variable therefore getting the prediction of water consumption and the water supply per person (A).

## 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

Figure 3: water scarcity map[3]

water sufficient in the future?). Thus we pick up China as our research object. To make our model more predictable and more reality connected.

[illegible][illegible]

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.

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
EIP	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, a polynomial fitting with power 2 is sufficient.

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

**Iron and Steel Production** The data of ISP perform in a similar way of IA, therefore we can use the same method to predict the ISP evolution.

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

**Electricity Production** The behavior of ELP is every bit as same as two parameters above, thus we just apply the same evolution equation form.

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

**Engel's Coefficient** Engel's Coefficient perform in a much different way. When it comes to the Engel, it make sense that with the development of society. It will keep decreasing as it used to be with a limit target: zero. Therefore we introduce a inverse proportion function to describe it's tendency.

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

With the fitness method, we can get all the evolution of alternative prominent parameters in Figure 4, and the parameters in evolution equations in Table 3:

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

Table 3: evolution parameters

## 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  $\{x_1, x_2, \dots, x_n\}$  and  $\{y_1, y_2, \dots, y_n\}$ , then the formula for  $r$  is[6]:

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

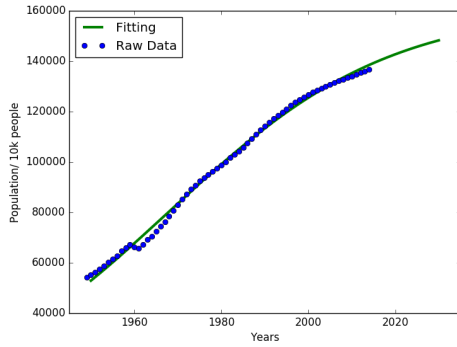
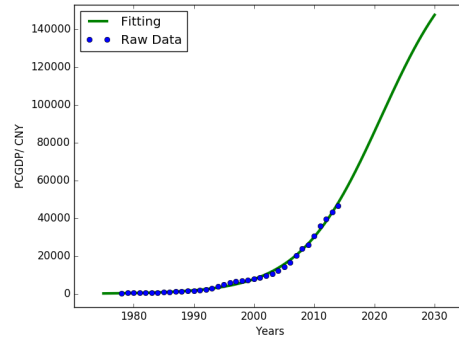
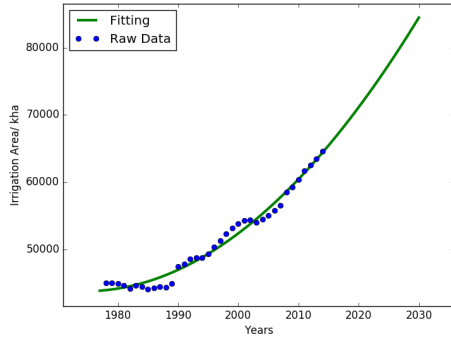
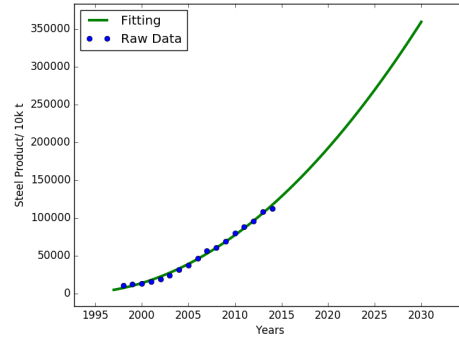
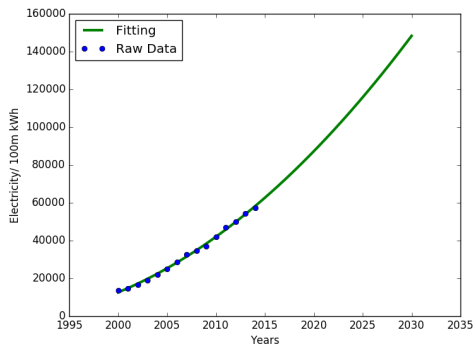
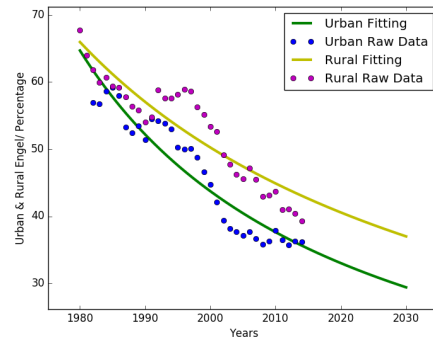
(a)  $p(t)$ (b)  $PCGDP(t)$ (c)  $IA(t)$ (d)  $ISP(t)$ (e)  $ELP(t)$ (f)  $Enc(t)$ 

Figure 4: 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

**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  $\rho$  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  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  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  $\tau$  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 ( $\tau, p$ )	Average
IWC $\sim P$	(0.669,0.0245)	(0.627,0.0388)	(0.455,0.0516)	0.584
IWC $\sim$ PCGDP	(0.614,0.0444)	(0.627,0.0388)	(0.455,0.0516)	0.565
IWC $\sim$ ELP	(0.648,0.0311)	(0.627,0.0388)	(0.455,0.0516)	0.577
IWC $\sim$ ISP	(0.653,0.0294)	(0.627,0.0388)	(0.455,0.0516)	0.577
AWC $\sim P$	(0.924,4.95e-5)	(0.927,3.97e-5)	(0.782,0.000815)	0.878
AWC $\sim$ PCGDP	(0.936,2.28e-5)	(0.927,3.97e-5)	(0.782,0.000815)	0.882
AWC $\sim$ IA	(0.927,4.10e-5)	(0.927,3.97e-5)	(0.782,0.000815)	0.879
DWC $\sim P$	(0.844,0.00108)	(0.836,0.00133)	(0.745,0.00141)	0.808
DWC $\sim$ PCGDP	(0.806,0.00276)	(0.836,0.00133)	(0.745,0.00141)	0.796
DWC $\sim$ Enc <sub>urban</sub>	(-0.785,0.00417)	(-0.809,0.00256)	(-0.600,0.0102)	-0.731
DWC $\sim$ Enc <sub>rural</sub>	(-0.419,0.199)	(-0.315,0.345)	(-0.204,0.383)	-0.313

Table 4: Coefficients: Coefficients of correlation in different standard

### 5.3 Fitness method in consumption variables

To investigate the evolution of consumption variables, We first investigate the correlation between one specific consumption with one specific prominent parameters.

Parameters	$p_0$	$p_1$	$p_2$
$P$	-1.31e3	5.07e-2	3.42e4
PCGDP	-1.50e3	9.54e-3	-5.18e5
IA	-1.09e3	3.20e-2	-9.09e4

Table 5: AWC - parameters:  $p_0 + p_1(\text{para} - p_2)$

Parameters	$p_0$	$p_1$	$p_2$
$P$	1.43e3	7.62e-9	1.34e5
PCGDP	1.45e3	3.35e-10	3.31e4
ELP	1.44e3	3.10e-10	4.45e4
ISP	1.44e3	5.50e-11	8.40e4

Table 6: IWC - parameters:  $p_0 \exp[-p_1(\text{para} - p_2)^2]$

Parameters	$p_0$	$p_1$	$p_2$
$P$	3.95e2	-2.16e-5	1.05e5
PCGDP	3.98e2	-3.82e-6	-1.30e5

Table 7: DWC - parameters:  $p_0 \exp[-p_1(\text{para} - p_2)]$

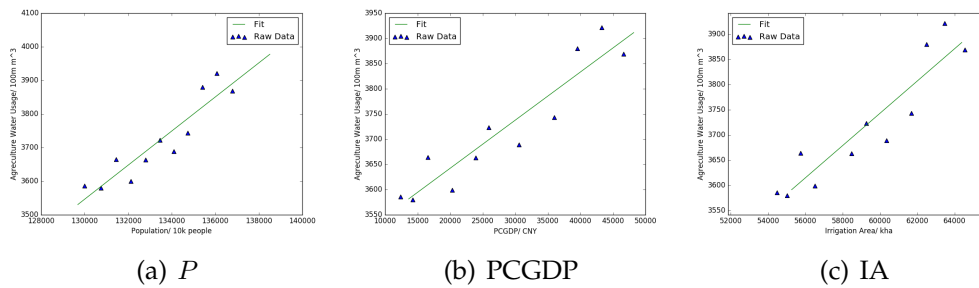
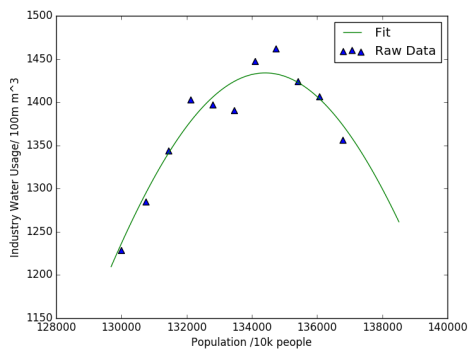
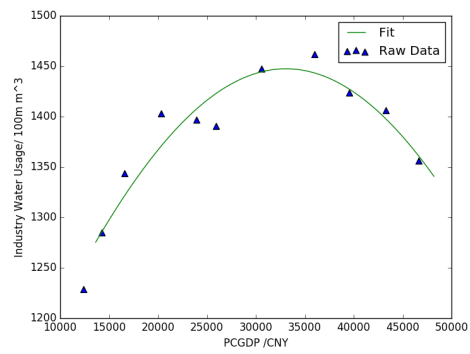
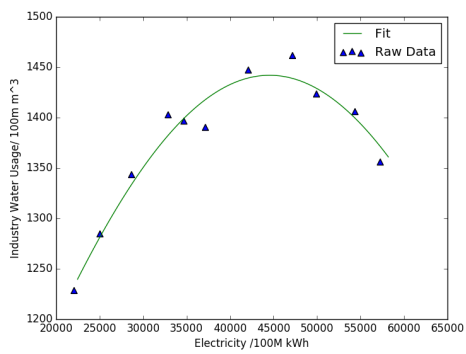


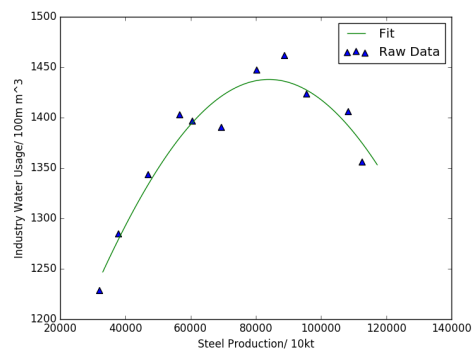
Figure 5: AWC - parameters

(a)  $P$ 

(b) PCGDP

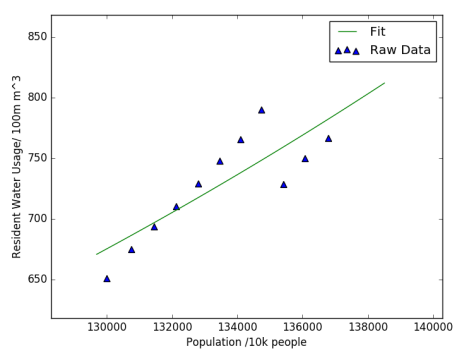
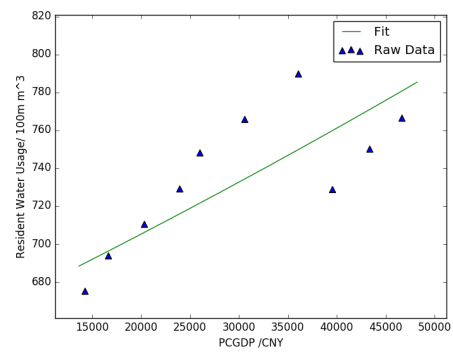


(c) ELP



(d) ISP

Figure 6: IWC - parameters

(a)  $P$ 

(b) PCGDP

Figure 7: DWC - parameters

## 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](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](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](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