Modeling Applied in Prediction and Allocation of Water Resources in Vast Area

Sunyechu Wang
School of Mathematics & Computational Science
Sun Yat-sen University
Guangzhou, China
e-mail: sophiawatson@163.com

Jieting Gong
School of Mathematics & Computational Science
Sun Yat-sen University
Guangzhou, China
e-mail: satanfever@163.com

Kaixi Huang
School of Mathematics & Computational Science
Sun Yat-sen University
Guangzhou, China
e-mail: hkxi@qq.com

***Abstract*—Thepaper mainly discusses prediction of water consumption and precipitation and allocation among areas. First of all, the vast area is divided into smaller parts according to its water resources, measured by precipitation in recent years. Precipitation and water consumption is predicted within each part before deciding water transportation plan among areas.**

Keywords: GM(1,1) model, Gray-Markov Chain model, linear programming, unbalanced transportation problem

I PREDICTION OF WATER RESOURCES

We first divided the vast area into smaller parts according to its water resources, measured by precipitation in recent years. Water consumption and precipitation in former years are analyzed within those smaller areas and prediction of their future volume is made. GM(1,1) model is adopted in prediction of water consumption. But it yields little fruit in predicting precipitation for precipitation fluctuates fiercely as time goes by. Deviation still exists when we ameliorate the model with unbiased GM(1,1) model. Finally the introduction of Markov chain is adopted to help gain an accurate result.

Allocation of water among areas is turn into an unbalanced transportation problem. We further turn the unbalanced problem into a balanced one by setting a visual consumer. Then linear programming aiming at minimum cost is adopted to solve the problem.

A. Prediction Of Water Consumption

Here we take mainland China as an example. Based on precipitation, we divided the area into ten parts.(see Fig. 1)

Figure 1.

Under such division, water consumption from 2003 to 2010 within each part is presented as below.(see Table I)

TABLE I

Zone\Year	2003	2004	2005	2006	2007	2008	2009	2010
1	640.3	649.73	645.81	646.65	658.79	659.57	664.85	671.46
2	927.1	1061.3	1058.9	1115.1	1121.6	1161.1	1164.1	1174.6
3	736.4	769.82	789.87	792.2	817.86	828.54	844.99	852.91
4	791.3	809.85	840.36	847.5	849.88	855.53	865.14	867.16
5	131.3	131.42	134.48	144.09	140.29	142.38	140.61	147.18
6	662.3	668.06	668.18	714.06	689.51	699.85	706.29	698.45
7	478.1	488.86	624.44	530.35	535.02	543.87	570.13	588.71
8	352.5	367.29	375.82	378.65	373.54	372.13	374.11	376.09
9	54.3	58.16	63.84	67.23	67.81	71.89	59.61	65.97
10	500.7	497.06	464.36	513.43	517.74	528.22	530.9	535.08

Let's present the data in columnar section(see Fig. 2).

Conclusion could be drawn that water consumption varies slightly as time goes by.

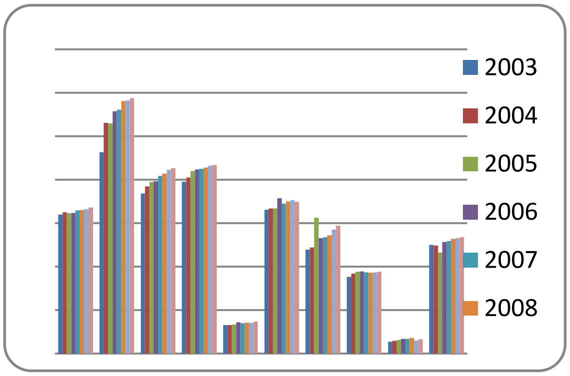

Figure 1.

For further prediction, variation of water consumption is needed. Here we adopt GM(1,1) model to achieve the goal.

First, E0 (the origin sequence) is as below.

$$E_0=(E(1),E(2),\dots,E(n))$$

E1 is gained after accumulated generating operation

$$E1=(E1(1),E1(2),\dots,E1(n))$$

in which

$$E1(k)=\sum_{i=1}^{k} E0(k),\ k=1,2,\dots,n$$

Then back ground value array Z is constructed according to E0

$$Z=(z(2),z(3),\dots,z(n))$$

in which

$$z(k)=\alpha\ (E1(k)+E1(k-1))\text{(in most cases } \alpha=0.5)$$

As a result, we have

$$Y=\begin{bmatrix} E(2) \\ E(3) \\ \dots \\ E(n) \end{bmatrix},\ B=\begin{bmatrix} -z(2) & 1 \\ -z(3) & 1 \\ \dots & \dots \\ -z(n) & 1 \end{bmatrix}$$

After least square estimation of parameter sequence, we have

$$\hat{a}=[a,b]^T=\left(B^T B\right)^{-1} B^T Y.$$

Then model can be decided and estimation of E0 computed.

$$E_1\left(\hat{k}+1\right)=\left(T(1)-\frac{b}{a}\right)e^{-ak}+\frac{b}{a}, k=1,2,\cdots n$$

$$\hat{E}(k)=\hat{E}_1(k)-\hat{E}_1(k-1)$$

Error checking is needed at last. The residual is presented as

$$\frac{1}{n}\sum_{i=1}^{n}\left(\hat{E}(i)-E(i)\right)$$

According to the model and the data, the outcome is presented as below(see Fig. 3).

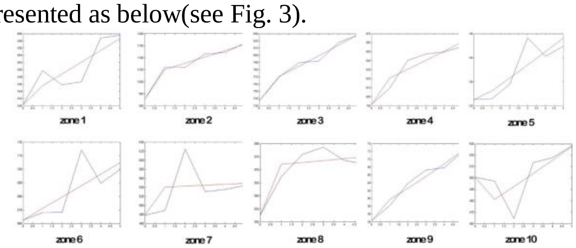

Figure 2.

It is learned that variation of water consumption could be highly stimulated by the model except abnormal data from a few years. So we can reasonably predict perspective water consumption in 2015,2020, 2025 within each part.(see Table.II)

TABLE II

Zone\Year	2015	2020	2025
1	690. 78	713. 02	735. 97
2	1304. 5	1433. 71	1575. 72
3	934. 35	1019. 04	1111. 41
4	915. 55	960. 65	1007. 97
5	157. 4	169. 35	182. 22
6	736. 74	766. 18	796. 8
7	615. 96	657. 94	702. 79
8	378. 66	382. 64	384. 65
9	70. 55	74. 32	78. 28
10	592. 32	648. 77	710. 59

We notice certain degree of growth in water consumption, but with no noticeable fluctuation. Therefore, the result is reliable.

B. Prediction Of Precipitation

Since large amount of data is needed, here we take Nanjing as an example in case the data amount is guaranteed. Precipitation in each part could be predicted similarly.

1)GM(1,1)

Based on precipitation per year in Nanjing during 1971 to 2008, prediction of precipitation using GM(1,1) is presented as below.

TABLE III

Year	precipitation	Year	precipitation	Year	precipitation	Year	precipitation
1971	1009	1981	1049	1991	1828	2001	737
1972	1279	1982	1106	1992	884	2002	1076
1973	930	1983	1125	1993	1241	2003	1659
1974	1289	1984	961	1994	647	2004	895
1975	1341	1985	1014	1995	771	2005	994
1976	893	1986	722	1996	1232	2006	1106
1977	1148	1987	1379	1997	903	2007	1073
1978	532	1988	925	1998	1241	2008	974
1979	959	1989	1258	1999	1215		
1980	1025	1990	952	2000	1031		

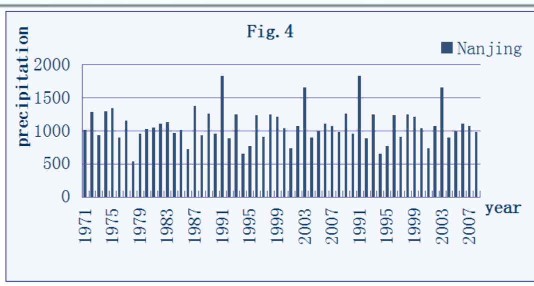

Figure 3.

We build the GM(1,1) model according to raw data in Table I. Then we have X(1) after accumulated generation operation of the sequence.

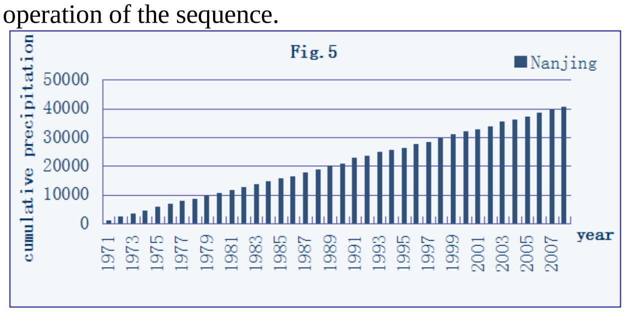

Figure 4.

Similar to GM(1,1) model adopted in prediction of water consumption, expression of the model could be computed in matlab as below:

$$\hat{E}^{(0)}_{(k+1)} = 1060.9e^{0.000186}$$

Prediction is presented as below.

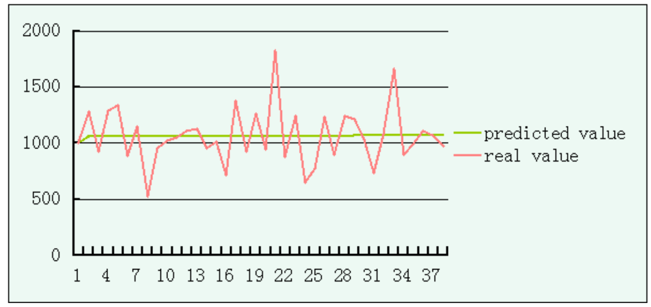

Figure 5.

TABLE IV

Year	predicted value of GM	relative error	Year	predicted value of GM	relative error	Year	predicted value of GM	relative error	Year	predicted value of GM	relative error
1971	1009	0	1981	1062.92	0.01	1991	1064.90	0.42	2001	1066.89	0.45
1972	1061.15	0.17	1982	1063.12	0.04	1992	1065.10	0.20	2002	1067.09	0.01
1973	1061.34	0.14	1983	1063.32	0.05	1993	1065.30	0.14	2003	1067.29	0.36
1974	1061.54	0.18	1984	1063.52	0.11	1994	1065.50	0.65	2004	1067.48	0.19
1975	1061.74	0.21	1985	1063.71	0.05	1995	1065.71	0.38	2005	1067.68	0.07
1976	1061.93	0.19	1986	1063.91	0.47	1996	1065.90	0.13	2006	1067.88	0.03
1977	1062.13	0.07	1987	1064.11	0.23	1997	1066.09	0.18	2007	1068.08	0.00
1978	1062.33	1.00	1988	1064.31	0.15	1998	1066.29	0.14	2008	1068.28	0.10
1979	1062.53	0.11	1989	1064.51	0.15	1999	1066.49	0.12	2009	1068.48	
1980	1062.72	0.04	1990	1064.71	0.12	2000	1066.69	0.03	2010		

Quite different from prediction of water consumption, precipitation is not effectively stimulated by the model for drought and flood could lead to abnormal data. To alleviate the impact of abnormal data, optimization with smoothing treatment is introduced into the model.

2)Optimized GM(1,1)

According to our optimization, we suggest that a sequence of n data is acquired.

$$X^{(0)} = X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)$$

Newly processed data could be expressed as:

$$X^{(0)}(t) = \frac{\{X^{(0)}(t-1) + 2X^{(0)}(t) + X^{(0)}(t+1)\}}{4} \qquad t \in [2, n-1]$$

$$X^{(0)}(1) = \frac{(3X^{(0)}(1) + X^{(0)}(2))}{4}$$

$$X^{(0)}(n) = \frac{X^{(0)}(n-1) + 3X^{(0)}(n)}{4}$$

And the expression of the ameliorated model is presented as below:

$$\hat{E}^{(0)}_{(k+1)} = 1053.3e^{0.00048}$$

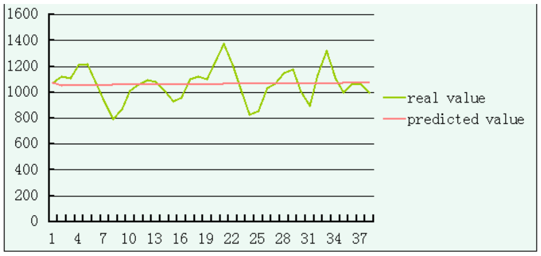

Figure 6.

TABLE V

Year	predicted value of optimized GM	relative error	Year	predicted value of optimized GM	relative error	Year	predicted value of optimized GM	relative error	Year	predicted value of optimized GM	relative error
1971	1076.5	0	1981	1058.31	0.00	1991	1063.86	0.23	2001	1068.93	0.19
1972	1053.80	0.06	1982	1058.82	0.03	1992	1064.36	0.12	2002	1069.43	0.06
1973	1054.30	0.05	1983	1059.32	0.02	1993	1064.87	0.06	2003	1069.94	0.19
1974	1054.80	0.13	1984	1059.82	0.04	1994	1065.38	0.29	2004	1070.45	0.04
1975	1055.30	0.13	1985	1060.33	0.14	1995	1065.88	0.25	2005	1070.96	0.07
1976	1055.80	0.05	1986	1060.83	0.11	1996	1066.39	0.03	2006	1071.47	0.00
1977	1056.30	0.14	1987	1061.33	0.04	1997	1066.90	0.00	2007	1071.98	0.01
1978	1056.81	0.33	1988	1061.84	0.05	1998	1067.40	0.07	2008	1072.49	0.07
1979	1057.31	0.22	1989	1062.34	0.03	1999	1067.91	0.09	2009		
1980	1057.81	0.04	1990	1062.85	0.15	2000	1068.42	0.06	2010		

After comparison of the two models, we learn that relative error of the ameliorated model is smaller than the original one. However, deviation still exists. For a more accurate prediction, we build a new model combining advanced gray model stated above and Markov chain.

3)Optimized Grey- Markov Chain Model

Relative error of the advanced GM is analyzed and divided into 6 states as below

Error burst%	16~7	7~4	4~0	0~-4	-4~-7	-7~-16
condition	1	2	3	4	5	6

According to the precipitation data, we have state matrix and single-step transition matrix.

$$\begin{pmatrix} 2 & 0 & 1 & 0 & 2 & 0 \\ 2 & 1 & 2 & 0 & 0 & 0 \\ 1 & 2 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 1 & 2 & 0 \\ 0 & 1 & 0 & 3 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2/5 & 0 & 1/5 & 0 & 2/5 & 0 \\ 2/5 & 1/5 & 2/5 & 0 & 0 & 0 \\ 1/7 & 2/7 & 1/7 & 1/7 & 2/7 & 0 \\ 0 & 0 & 2/5 & 1/5 & 2/5 & 0 \\ 0 & 1/5 & 0 & 3/5 & 0 & 1/5 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

We decide the calculation formula based on the state of relative error of former year. For example, we predict the precipitation in 2008 to be 1071.98 whereas actually it was 974, then we indicate that precipitation in 2009 to be:
1071.98 + 1/2*(0.11*(0.2)-0.04*(0.6)-0.23*(0.2))* 974= 1048.6

We witness only small deviation from the actual value after the precipitation is predicted. As a result, the prediction is well-rounded.

However, this model leaves something to be desired:
State division of relative error is decided arbitrary.
Influence of abnormal data is not perfectly eliminated.

II WATER ALLOCATION AMONG SMALL AREAS

Precipitation, x and water demand, y could be predicted by Gray-Markov model built above. We simply assume precipitation to be the only resource of water and it can be fully exploited. Water abundance is measured by D=x-y. Water allocation among small areas is in fact an unbalanced transition problem. Those areas with D>0 are decided as supplier while those with D<0 are decided as consumer.

We suggest that water is transported via present river or pipeline so that cost for newly-built transportation will be zero. Transportation cost per distance is the same so overall transportation cost is proportional to distance between two areas. When the areas are big enough, we simply decide the distance to be linear distance.

Taking mainland China as example, we have cost matrix as below (based on division in Table.VI) :

TABLE VI

	1	2	3	4	5	6	7	8	9	10
1	0	1109	748.4	1608	1857	1802	3299	2522	3354	4380
2	1109	0	831.6	1913	1220	803.8	2218	1774	3326	4075
3	748.4	831.6	0	1137	1137	1220	2883	1802	2744	3687
4	1608	1913	1137	0	1441	1968	3659	1940	1774	2938
5	1857	1220	1137	1441	0	693	2245	665.2	2273	3493
6	1802	803.8	1220	1968	693	0	1691	1053	2994	3493
7	3299	2218	2883	3659	2245	1691	0	2051	4296	4380
8	2522	1774	1802	1940	665.2	1053	2051	0	2245	2522
9	3354	3326	2744	1774	2273	2994	4296	2245	0	1331
10	4380	4075	3687	2938	3493	3493	4380	2522	1331	0

Since overall precipitation overwhelms overall water consumption, a visual consumer is decided as fm. Fm(X)=0, fm(y)=$\sum$x-$\sum$y. Then unbalanced transportation is now turn into balanced transportation problem. In this model, water transported to fm represents water stored with in each area.

The transportation model we build is a black box. Input precipitation and water consumption and the model will automatically identify supplier and consumer as long as overall precipitation overwhelms overall water consumption. The optimal plan will computed aiming at minimum cost.

Input precipitation x_i and y_i in each area.D_i=x_i-y_i. $D_i$is supply for suppler while D_i is demand for consumer. Then the problem could be describe by linear equations as below

$$sum = \sum cost_{ij} * tr_{ij}$$
$$0 < tr_{ij} \le D_i tr_{ij} \le y_j$$
$$Min(sum)$$

In which $const_{ij}$ presents transportation cost per distance from area i to area j; tr_{ij} represents transportation volume from area i to area j.

The linear equations could be solved by linear programming.

We input precipitation and water consumption in 2008 to exam the model and the result is as below.

TABLE VII

Area	1	2	5	7	8
3	133.89	142.39	266.77	404.99	104.71
4	239.50	246.16	383.34	598.72	219.93
6	435.13	246.16	508.43	118.51	381.66
9	214.06	294.66	360.62	10.06	189.01
10	61.81	41.19	233.10	80.95	10.73
store	40.42	57.37	177.61	76.51	10.96

REFERENCE

[1] Feng X, China Academic Journal Electronic Publishing House, vol.18, No.6, Dec 2008, ‘Supply Water Capacity Prediction Based on Mathematic Model’, number: 1009-3907(2008)06-0049-02;

[2] Hongbin L,Yong Y, Suiningmeteorologic bureau, ‘Rainfall Prediction in SuiningRegion Based on Grey-Markov Model’;

[3] Jieming ZH, China Academic Journal Electronic Publishing House,vol.18, No.4, ‘Urban Supply Water Capacity Prediction Based on Multiple Linear Regressions’;

[4] Zheng Z, Jiayan T, Zhiming L, China Academic Journal Electronic Publishing House vol.2 No.2 Apr.2008, ‘The mathematic model of city water supply capacity’, number: 1673-9353(2008)02-0011-05;

[5] Yuanjian ZH, Jing L, Long WChina Academic Journal Electronic Publishing House vol.23, No.4, Oct 2009, ‘Precipitation Predicting Model Based on Improved Markov Chain’, number: 1671-3559(2009)04-0402-04