Support Vector Regression and Time Series Analysis for the Forecasting of Bayannur’ s Total Water Requirement

Xuchan Ju

June 5, 2014

contents

contents

- Introduction
- Support Vector Regression

contents

- Introduction
- Support Vector Regression
- Time Series Analysis

contents

contents

Introduction




Problem Description

- Select influencing factors from climate aspect, social aspect and economic aspect of Bayannur, which are related to the total water requirement;

Problem Description

- Select influencing factors from climate aspect, social aspect and economic aspect of Bayannur, which are related to the total water requirement;
- Built total water requirement model using these influencing factors;

Problem Description

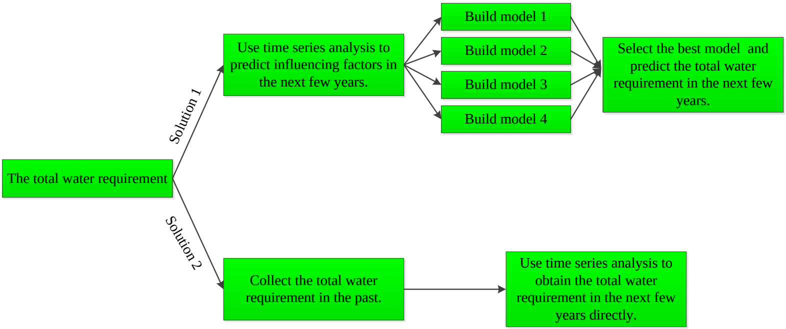
- Select influencing factors from climate aspect, social aspect and economic aspect of Bayannur, which are related to the total water requirement;
- Built total water requirement model using these influencing factors;
- Forecast total water requirement.

Influencing Factors

Table 1. the influencing factors which determine the total water requirement.

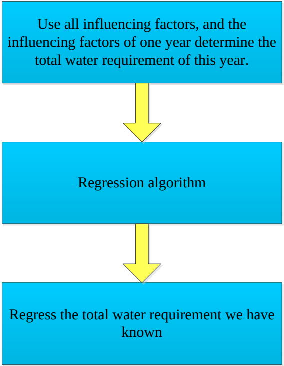
Influencing factors	Influencing factors
Precipitation	Gross value of production of the primary industry
Annual mean temperate	Gross value of production of the secondary industry
The highest mean temperate all the year round	Gross value of production of the tertiary industry
The lowest mean temperate all the year round	Per capita GDP
Annual average sunshine time	Comprehensive water consumption per capita
Annual frost-free period	Urban per capita disposable income
Annual average groundwater depth in the irrigation area	Per capita net income of farmers and herdsmen
Water diversion in the irrigation area	The planting area of crops in Hetao-irrigated region
The first irrigation water of the main canal	The gross irrigation quota in Hetao-irrigated region
Water diversion of main canal straight mouth	The livestock number
Loss of conveying water	The total grain output
Water use efficiency of main canal	Gross agricultural output value
Evaporation	Gross industrial output value
Population	Gross construction output value
GDP	

Two solutions

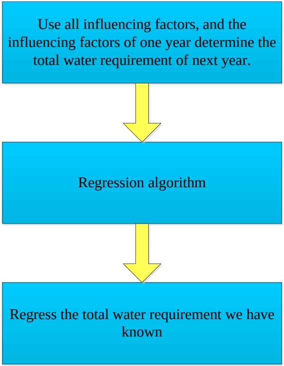


Four models

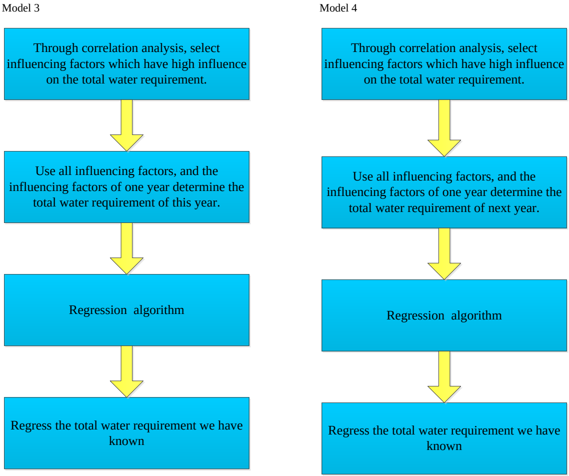
Model 1



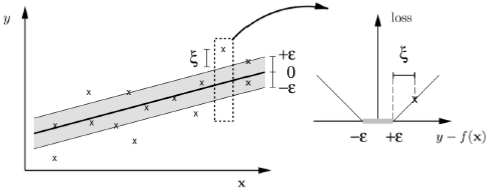
Model 2



Four models


Model 3
Through correlation analysis, select influencing factors which have high influence on the total water requirement.
Use all influencing factors, and the influencing factors of one year determine the total water requirement of this year.
Regression algorithm
Regress the total water requirement we have known
Model 4
Through correlation analysis, select influencing factors which have high influence on the total water requirement.
Use all influencing factors, and the influencing factors of one year determine the total water requirement of next year.
Regression algorithm
Regress the total water requirement we have known

ϵ-Support Vector Regression



$$
\begin{aligned}
\min_{w,b,\xi^{(*)}} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*), \\
\text{s.t.} \quad & (w \cdot x_i) + b - y_i \leqslant \varepsilon + \xi_i, i = 1, \cdots, l, \\
& y_i - (w \cdot x_i) - b \leqslant \varepsilon + \xi_i^*, i = 1, \cdots, l, \\
& \xi_i^{(*)} \geq 0, i = 1, \cdots, l,
\end{aligned}
$$

- Dual problem

$$\min_{\alpha^{(*)}\in R^{2l}} \quad \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) + \varepsilon\sum_{i=1}^{l}(\alpha_i^* + \alpha_i) - \sum_{i=1}^{l} y_i(\alpha_i^* - \alpha_i),$$

$$\text{s.t.} \quad \sum_{i=1}^{l}(\alpha_i^* - \alpha_i) = 0,$$

$$0 \le \alpha_i^{(*)} \le C, i = 1, \cdots, l,$$

Decision function

$$y = (w \cdot x) + b = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)(x_i \cdot x) + b.$$

ϵ-Support Vector Regression

Algorithm 1 ε-Support Vector Regression

(1) Input the training set $T = \{(x_1, y_1), \cdots, (x_l, y_l)\} \in (R^n \times \mathcal{Y})^l$,where $x_i \in R^n, y_i \in \mathcal{Y} = R, i = 1, \cdots, l$;
(2) Choose an appropriate kernel $K(x, x')$, an appropriate accuracy $\varepsilon > 0$ and the penalty parameter $C > 0$;
(3) Construct and solve the convex quadratic programming problem (9), obtaining a solution $\alpha^{(*)} = (\alpha_1, \alpha_1^*, \cdots, \alpha_l, \alpha_l^*)^\top$;
(4) Compute b: choose a component of $\alpha^{(*)}$, $\alpha_j \in (0, C)$ or $\alpha_k^* \in (0, C)$. If α_j is chosen, compute

$$b = y_j - \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) K(x_i, x_j) + \varepsilon; \tag{10}$$

if α_k^* is chosen, compute

$$b = y_k - \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) K(x_i, x_k) - \varepsilon; \tag{11}$$

(5) Construct the decision function

$$y = g(x) = \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) K(x_i, x) + b. \tag{12}$$

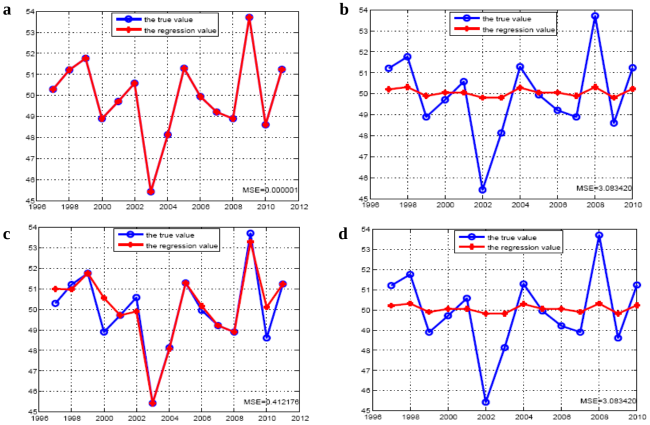
Time Series Analysis

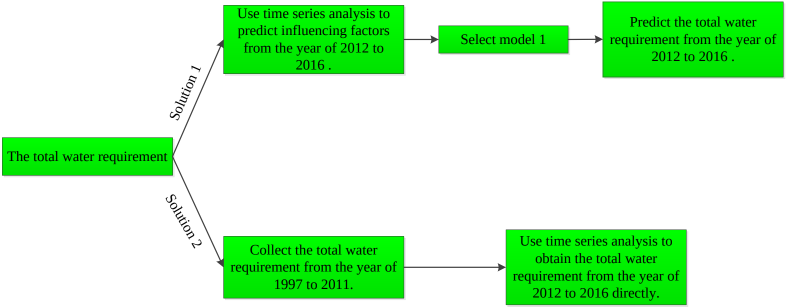
- A time series is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals;

Time Series Analysis

- A time series is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals;
- Autoregressive moving average(ARMA) model is the most-used model in time series analysis.

Experimental Results

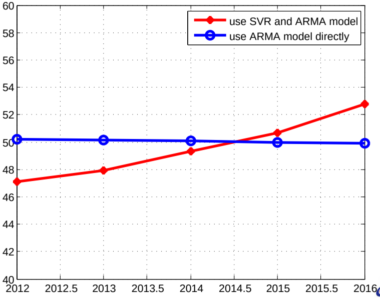



The total water requirement
Solution 1
Use time series analysis to predict influencing factors from the year of 2012 to 2016 .
Select model 1
Predict the total water requirement from the year of 2012 to 2016 .
Solution 2
Collect the total water requirement from the year of 1997 to 2011.
Use time series analysis to obtain the total water requirement from the year of 2012 to 2016 directly.

Experimental Results

Table 5. the total water requirement from the year of 2012 to 2016 via SVR and ARMA model and ARMA model directly.

Year	Using SVR and ARMA model	Using ARMA model directly
2012	47.120	50.200
2013	47.950	50.132
2014	49.341	50.064
2015	50.676	49.997
2016	52.807	49.929



Conclusion

- propose two solutions to the forecasting of Bayannur's total water requirement;

Conclusion

- propose two solutions to the forecasting of Bayannur's total water requirement;
- Two methods both provide scientific basis for the forecasting of Bayanuur' s total water requirement;

Conclusion

- propose two solutions to the forecasting of Bayannur's total water requirement;
- Two methods both provide scientific basis for the forecasting of Bayanuur’ s total water requirement;
- Forecasting the total water requirement accurately is the precondition of allocating the water resources reasonably;

Conclusion

- propose two solutions to the forecasting of Bayannur's total water requirement;
- Two methods both provide scientific basis for the forecasting of Bayanuur' s total water requirement;
- Forecasting the total water requirement accurately is the precondition of allocating the water resources reasonably;
- Our work is important and meaningful to promote the development of Bayannur.

Further work

- Use more methods to forecast Bayanuur's total water requirement.

Further work

- Use more methods to forecast Bayanuur's total water requirement.
- allocate the water resources reasonably

References

1. Xia RL. Analysis about policies and institutional arrangements of Efficient allocation of water resources in Bayanuur. *Thesis of master degree of Capital University of Economics and Business*; 2008.
2. Fu Q, Xing ZX, Ma YS. Applying multivariate auto-regression model to forecast the water requirement of well irrigation rice in Sanjiang plain. *Nature and Science*; 2004. p. 8-14.
3. Draper NR, Smith H. *Applied regression analysis*. John Wiley and Sons; 1981.
4. Anthony M, Bartlett PL. *Neural Network Learning: Theoretical Foundations*. Cambridge University Press; 1999.
5. Bowden GJ, Dandy GC, Maier HR. Input determination for neural network models in water resources applications. part 1-background and methodology. *Journal of Hydrology*; 2005. p. 75-92.
6. Maier HR, Dandy GC. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Modelling and Software*; 2000. p. 101-124.
7. Burges CJ. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*; 1998. p. 121-167.
8. Basak D, Pal S, Patranabis DC. Support vector regression. *Neural Information Processing-Letters and Reviews*; 2007. p. 203-224.
9. Deng NY, Tian YJ, Zhang YC. *Support Vector Machines: Optimization Based Theory, Algorithms, and Extensions*. Chapman and Hall/CRC; 2012.
10. Hamilton JD. *Time series analysis*. Princeton Unversity Press; 1994.
11. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis: Forecasting and Control*. John Wiley and Sons; 2008.
12. Montgomery DC, Johnson LA. *Forecasting and time series analysis*. McGraw-Hill; 1976.
13. Rao TS, Rao SS, Rao CR. *Handbook of Statistics, Volume 30: Time Series Analysis: Methods and Applications*. Elsevier; 2012.
14. Otnes RK, Enochson L. *Applied Time Series Analysis, Vol. 1*. New York: Wiley; 1978.
15. Cryer JD, Chan KS. *Time Series Analysis: With Applications in R*. Springer Science and Business Media; 2008.
16. Ezekiel M. *Methods of correlation analysis*. Oxford, England: Wiley; 1930.

Thank you for your listening!