

A Changing Weights Spatial Forecast Combination Approach with an Application to Housing Price Prediction

Chuanhua Wei¹, Chenping Du² & Nana Zheng¹

¹ School of Science, Minzu University of China, Beijing, China

² School of Mathematics, Sun Yat-Sen University, Guangzhou, China

Correspondence: Chenping Du, School of Mathematics, Sun Yat-sen University, Guangzhou 510220, China. Tel: 86-178-2706-5504. E-mail: dcp9765@126.com

Received: January 31, 2020

Accepted: February 25, 2020

Online Published: March 5, 2020

doi:10.5539/ijef.v12n4p11

URL: <https://doi.org/10.5539/ijef.v12n4p11>

Abstract

Forecast combination has been widely applied in various fields since the seminal article of Bates and Granger (1969). However, these research were focused only on time series data. Few study focus on the spatial data, this paper proposes a novel adaptive spatial forecast combination method with varying weights based on the geographically weighted regression technique. Finally, the proposed method is applied to the Boston house prices prediction, and the results indicate that our procedure performs better than the other forecast combination methods.

Keywords: geographically weighted regression, combining forecast, local weighted least-squares approach, varying weights

1. Introduction

Forecast combination method has been theoretical studied in econometrics and applied in many fields for a long time since the seminal article of Bates and Granger (1969). The recent advances can be found in Granger (2006), Timmermann (2006), Claeskens et al. (2016) and Chan and Pauwels (2018).

As we all known, forecast combination method was focused only on time series forecasting problem. However, different to time series forecasting, spatial forecasting or spatial prediction is also very important in many areas, such as economics, geography, population studies, environment science, ecology and epidemiology. Spatial prediction methods can be divided into two classes, one is geostatistics (spatial statistics) methods, the other is spatial econometrics methods. In geostatistics, the prediction of unknown quantities at given locations is commonly made by the kriging technique. In addition to the kriging technique for modeling regular lattice spatial data, the spatial econometrics models including spatial autoregressive models and geographically weighted regression can also be used. More discussion and comparison on these two kinds of spatial prediction can be found in Tsutsumi and Seya (2009) and Mojiri et al. (2017). However, little efforts have been devoted to spatial forecast combination. This paper wants to fill up this gap, and proposes a novel spatial combining prediction method with varying weights based on the geographically weighted regression technique.

The rest of this paper is organized as follows. In Section 2, the spatial forecast combination method is proposed. In Section 3, the proposed approach is applied to the Boston house prices prediction. Conclusion is presented in Section 4.

2. Method

To proposed the new method, we briefly describe a simple combined forecast method based on the Ordinary Least Squares (OLS) regression. OLS-based forecast combination method was studied by Granger and Ramanathan (1984). They proposed to combine forecasts with unrestricted regression coefficients as weights.

Let f_i^j denote an out-of-sample prediction value of the variable y_i by the j th individual forecasting approach at spatial location (u_i, v_i) , for $i=1, 2, \dots, n$ and $j=1, 2, \dots, p$. Then, we can define the following linear regression model.

$$y_i = \beta_0 + \beta_1 f_i^1 + \dots + \beta_p f_i^p + \varepsilon_i, i = 1, 2, \dots, n \quad (1)$$

Apply the Ordinary Least Squares approach to linear regression model (1), we can obtain the OLS estimator for

regression coefficient $\beta = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)^T$ as

$$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)^T = (F^T F)^{-1} F^T Y, \quad (2)$$

with

$$F = \begin{bmatrix} 1 & f_1^1 & f_1^2 & \dots & f_1^p \\ 1 & f_2^1 & f_2^2 & \dots & f_2^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & f_n^1 & f_n^2 & \dots & f_n^p \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Then, we can construct prediction value for y_0 at a new location (u_0, v_0) by the OLS combined forecast method. Firstly, we obtain prediction $f_0^j, j = 1, 2, \dots, p$ by p individual forecasting methods. Then, combined forecast for y_0 can be defined as

$$\hat{y}_0^{\text{OLS}} = \hat{\beta}_0 + \hat{\beta}_1 f_0^1 + \dots + \hat{\beta}_p f_0^p. \quad (3)$$

To improve the constant weights forecast combination methods, Deutsch et al. (1994) proposed a time-varying weights combined forecast method. They considered the combination of forecasts using changing weights derived from switching regression models or from smooth transition regression models. Discussions on time-varying weights forecast combination method can be found in Billio et al. (2000), Terui and van Dijk (2002), Guidolin and Timmermann (2009) and Raftery et al. (2010). We are motivated to propose a spatial combined forecast method with spatial-varying weights.

Based on model (1), we can get the following spatial varying coefficient model by allowing the regression coefficients $\beta = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)^T$ to change with spatial location,

$$y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i) f_i^1 + \dots + \beta_p(u_i, v_i) f_i^p + \varepsilon_i, i = 1, 2, \dots, n \quad (4)$$

Compared with linear model (1), $\beta(\cdot) = [\beta_0(\cdot), \beta_1(\cdot), \dots, \beta_p(\cdot)]^T$ is not a p -dimensional vector of constants but unknown functions of geographical locations,

To estimate spatial varying model (4), Brunson et al. (1996) proposed a geographically weighted regression (GWR) approach. The GWR method estimates the coefficients by the local weighted least-squares technique in the framework of kernel smoothing. More references on GWR can be found in the monograph Fotheringham et al. (2002). Now, for a spatial location (u_0, v_0) , we apply this technique to estimate the regression coefficients $\beta(u_0, v_0) = (\hat{\beta}_0(u_0, v_0), \hat{\beta}_1(u_0, v_0), \dots, \hat{\beta}_p(u_0, v_0))^T$. Let d_{0i} be the distance between (u_0, v_0) and the i th designed location (u_i, v_i) . The GWR estimation of $\beta(u_0, v_0)$ is to minimize the objective function

$$\sum_{i=1}^n \left[y_i - \beta_0(u_0, v_0) - \sum_{j=1}^p \beta_j(u_0, v_0) f_i^j \right]^2 K_h(d_{0i}) \quad (5)$$

and K is a kernel function, h is a bandwidth and $K_h(\cdot) = K(\cdot/h)/h$.

Let $W(u_0, v_0) = \text{diag}[K_h(d_{01}), K_h(d_{02}), \dots, K_h(d_{0n})]$, and the solution to the weighted least-squares problem (5) can be explicitly expressed by

$$\hat{\beta}(u_0, v_0) = (\hat{\beta}_0(u_0, v_0), \hat{\beta}_1(u_0, v_0), \dots, \hat{\beta}_p(u_0, v_0))^T = (F^T W(u_0, v_0) F)^{-1} F^T W(u_0, v_0) Y. \quad (6)$$

As usual, the weight function $K_h(\cdot)$ in equation (5) can be chosen to be the Gaussian weight

$$K_h(d_{0i}) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{d_{0i}}{h} \right)^2 \right], i = 1, 2, \dots, n, \quad (7)$$

For the choice of bandwidth h , various data-driven methods, such as cross-validation (CV), generalized cross-validation (GCV), and AIC, can be applied, more details can be found in Fotheringham et al. (2002). CV method will be employed in the next section.

Then, the GWR combined forecast for y_0 based on individual forecasting values $f_0^1, f_0^2, \dots, f_0^p$ can be defined as

$$\hat{y}_0^{\text{GWR}} = \hat{\beta}_0(u_0, v_0) + \hat{\beta}_1(u_0, v_0) f_0^1 + \dots + \hat{\beta}_p(u_0, v_0) f_0^p \quad (8)$$

3. Results

We now illustrate the proposed estimating procedure by analyzing the Boston Housing data set, which was used originally by Harrison and Rubinfeld (1978) and corrected for a few minor errors by Gilley and Pace (1996). The

data set consists of the median value of owner-occupied homes in 506 census tracts in the Boston Standard Metropolitan Statistical Area in 1970, along with 13 accompanying sociodemographic and related variables, and is now freely available through the **Sedep** package of the open source software R. Details of variables can be found in Table 1.

Table 1. Description of the Boston dataset

Variable	Description	Defined variable
MEDV	median value of owner-occupied homes in 1,000USD	$Y = \ln(MEDV)$
CRIM	the per capita crime rate by town	$X_1 = CRIM$
ZN	proportion of area zoned with large lots	$X_2 = ZN$
INDUS	the proportion of non-retail business acres per town	$X_3 = INDUS$
CHAS	the Charles River dummy variable	$X_4 = CHAS$
NOX	the nitrogen oxide concentration in parts per 10 million	$X_5 = NOX^2$
RM	the average number of rooms in owner-occupied homes per dwelling	$X_6 = RM^2$
AGE	the proportion of owner-occupied homes built prior to 1940	$X_7 = AGE$
DIS	the weighted distances to five employment centers in the Boston region	$X_8 = \ln(DIS)$
RAD	the index of accessibility to radial highways	$X_9 = \ln(RAD)$
TAX	the full value property tax rate per 10,000USD	$X_{10} = TAX$
PTRATIO	the pupil-teacher ratio by town school district	$X_{11} = PTRATIO$
B	the black proportion of population by town	$X_{12} = B$
LSTAT	the proportion of population that is in the lower status	$X_{13} = \ln(LSTAT)$

We apply the following five models as individual forecasting procedures. M_1 is a traditional linear regression model. To capture the “large-scale” locational effects between response variable and associated covariates, Pace and Gilley (1997) considered M_2 with (u_i, v_i) is the latitude (LAT) and longitude (LON) of the i th observation. M_3 is a spatial error model, M_3 is a spatial autoregressive model, and M_4 is a spatial Durbin model. We use the Euclidean distance in terms of longitude and latitude to set up the spatial weight matrix \mathbf{W} . We choose the threshold distance to be 1 which gives a \mathbf{W}_n matrix with 1.18% nonzero elements.

$$M_1: Y_i = \beta_0 + \sum_{j=1}^{13} \beta_j X_{ij} + \varepsilon_i,$$

$$M_2: Y_i = \beta_0 + \sum_{j=1}^{13} \beta_j X_{ij} + \beta_{14} u_i + \beta_{15} v_i + \beta_{16} u_i v_i + \beta_{17} u_i^2 + \beta_{18} v_i^2 + \varepsilon_i,$$

$$M_3: Y_i = \beta_0 + \sum_{j=1}^{13} \beta_j X_{ij} + \varepsilon_i, \varepsilon_i = \sum_{j=1}^n W_{ij} \varepsilon_j + e_i,$$

$$M_4: Y_i = \beta_0 + \sum_{j=1}^n W_{ij} Y_j + \sum_{j=1}^{13} \beta_j X_{ij} + \varepsilon_i,$$

$$M_5: Y_i = \beta_0 + \sum_{j=1}^n W_{ij} Y_j + \sum_{j=1}^{13} \beta_j X_{ij} + \sum_{j=1}^{13} \theta_j W_{ij} X_{ij} + \varepsilon_i.$$

For 506 observations, we divided the data into a training set (486 observations) and a test set (20 observations). Based on the training set and the above five individual models M_1 - M_5 , we can obtain out-of-sample forecasting value for response variable Y at each spatial location. Then, based on the true values of Y and the according four forecasting values, we can apply the above five combined forecast approaches to get weights of each individual forecasting method. Next, based on the training set and the above five individual models M_1 - M_4 , we can obtain individual forecasting values for response variable Y of test set. Finally, we get combined forecasting values by their individual forecasting values and weights.

Based on the above five individual forecasting models, besides the proposed GWR-based forecast combinations method, we will apply the following four frequently used forecast combinations schemes. It is obviously that the four methods are all constant weights procedures. The proposed GWR method is a varying weights procedure. In our real data analysis, the weights for the individual forecasting models M_1 to M_5 according to these four forecast combinations methods are computed by the R package **ForecastCombinations**. Denote f_0^k , $k=1,2, \dots, 5$

are out-of-sample individual forecasting values for housing price at location (u_0, v_0) according to individual models M_1 to M_5 .

1) Simple average (SA). The most natural approach to combine forecasts is using the mean of all those forecasts. The combined forecast is defined as

$$\hat{y}_0^{SA} = \frac{1}{5} \sum_{k=1}^p f_0^k$$

2) OLS. The OLS-based combine forecasts method was introduced in Section 2, and the combined forecast is given as

$$\hat{y}_0^{OLS} = -0.0907 - 0.7492 f_0^1 + 0.2953 f_0^2 - 0.0716 f_0^3 + 1.3072 f_0^4 + 0.2487 f_0^5$$

3) Constrained Least Squares (CLS) regression. For linear model (1), OLS estimator was obtained without any restriction. To deal with a problem sometimes dubbed as “bouncing betas”, we allow

$$\frac{1}{p} \sum_{k=1}^p \beta_k = 1, \beta_k \geq 0, k = 1, 2, \dots, p$$

Then, apply the constrained Least Squares for model (1) with constrained conditions, we can get the CLS estimator for β , and this combined forecast is given

$$\hat{y}_0^{CLS} = -5.6888 * 10^{-18} f_0^1 - 8.7542 * 10^{-18} f_0^2 - 2.8844 * 10^{-16} f_0^3 + 0.9686 f_0^4 + 0.0031 f_0^5$$

4) Variance-based, or Inverse Mean Squared Error (IMSE). This method was applied by Stock and Watson (2004) for combining forecasts of output growth. More accurate forecasting methods (lower MSE) are weighted more heavily in this method. MSE here is computed based on out-of-sample forecasts, sometimes referred to as Mean Squared Prediction Error. The IMSE combined forecast is defined as

$$\hat{y}_0^{IMSE} = 0.1755 f_0^1 + 0.1782 f_0^2 + 0.1771 f_0^3 + 0.2806 f_0^4 + 0.1866 f_0^5$$

To measure the performance of the individual and combined forecasting approaches, we use the following four indexes, they are computed by the true vales and forecasting values of test set. Results are shown in Table 2.

$$MSE = \frac{1}{20} \sum_{k=1}^{20} (Y_i - \hat{Y}_i)^2, MAE = \frac{1}{20} \sum_{k=1}^{20} |Y_i - \hat{Y}_i|,$$

$$MSPE = \frac{1}{20} \sum_{k=1}^{20} \left(\frac{Y_i - \hat{Y}_i}{Y_i} \right)^2, MAPE = \frac{1}{20} \sum_{k=1}^{20} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|.$$

From the results, it is obviously that M_5 perform the worst compared with other individual forecasting model in all tables. Simple-Mean combined forecasting approache perform the worst compared with other four forecast combination methods. The proposed GWR-based combined method perform best except that its MSE is larger than that of individual model M_3 .

Table 2. Results of the individual and combined forecasting approaches

	Single Model Forecasting					Combination Forecasting Method				
	M_1	M_2	M_3	M_4	M_5	SA	OLS	CLS	IMSE	GWR
MSE	0.0601	0.0610	0.0526	0.0619	0.0808	0.0610	0.0591	0.0530	0.0599	0.0565
MAE	0.1696	0.1705	0.1376	0.1698	0.1790	0.1607	0.1318	0.1382	0.1579	0.1295
MSPE	0.0062	0.0063	0.0051	0.0063	0.0073	0.0060	0.0052	0.0052	0.0059	0.0050
MAPE	0.0548	0.0551	0.0444	0.0547	0.0570	0.0518	0.0417	0.0446	0.0509	0.0410

4. Discussion

To combining different individual forecasting values for one event with spatial information, this paper proposes a novel forecast combination method with varying weights. By the geographically weighted regression technique, we can obtain spatial varying weights according to the individual forecasting methods. The results on Boston house prices indicates that the procedure performs better than the other forecast combination methods.

Acknowledgments

Author Contributions: Investigation and Methodology, Chuanhua Wei; Software: Chenping Du; Writing review editing, Chuanhua Wei, Zheng Nana. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4), 451-468. <https://doi.org/10.1057/jors.1969.103>
- Billio, M., Sartore, D., & Toano, C. (2000). Combining forecasts: Some results on exchange and interest rates. *The European Journal of Finance*, 6(2), 126-145. <https://doi.org/10.1080/13518470050020806>
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28, 281-298. <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>
- Chan, F., & Pauwels, L. L. (2018). Some theoretical results on forecast combinations. *International Journal of Forecasting*, 34, 64-74. <https://doi.org/10.1016/j.ijforecast.2017.08.005>
- Claeskens, G., Magnus, J. R., Vasnev, A. L., & Wang, W. (2016). A simple theoretical explanation of the forecast combination puzzle. *International Journal of Forecasting*, 32(3), 754-762. <https://doi.org/10.1016/j.ijforecast.2015.12.005>
- Deutsch, M., Granger, C. W. J., & Terasvirta, T. (1994). The combination of forecasts using changing weights. *International Journal of Forecasting*, 10(1), 47-57. [https://doi.org/10.1016/0169-2070\(94\)90049-3](https://doi.org/10.1016/0169-2070(94)90049-3)
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression the Analysis of Spatial Varying Relationships*. Wiley, West Sussex.
- Granger, C. W. J. (2006). Invited review combining forecast twenty years later. *Journal of Forecasting*, 8, 167-173. <https://doi.org/10.1002/for.3980080303>
- Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3, 197-204. <http://dx.doi.org/10.1002/for.3980030207>
- Guidolin, M., & Timmermann, A. (2009). Forecasts of US short-term interest rates: A flexible forecast combination approach. *Journal of Econometrics*, 150(2), 297-311. <https://doi.org/10.1016/j.jeconom.2008.12.004>
- Mojiri, A., Waghei, Y., Sani, H. R. N., & Borzadaran, G. R. M. (2017). Comparison of Predictions by Kriging and Spatial Autoregressive Models. *Communications in Statistics-Simulation and Computation*, 47(6), 1785-1795. <https://doi.org/10.1080/03610918.2017.1324980>
- Pace, R. K., & Gilley, O. W. (1997). Using the spatial configuration of the data to improve estimation. *Journal of Real Estate Finance and Economics*, 14, 333-340. <https://doi.org/10.1023/A:1007762613901>
- Raftery, A. E., Karny, M., & Ettler, P. (2010). Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics*, 52, 52-66. <https://doi.org/10.1198/TECH.2009.08104>
- Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23, 405-430. <https://doi.org/10.1002/for.928>
- Terui, N., & van Dijk, H. K. (2002). Combined forecasts from linear and nonlinear time series models. *International Journal of Forecasting*, 18, 421-438. [https://doi.org/10.1016/S0169-2070\(01\)00120-0](https://doi.org/10.1016/S0169-2070(01)00120-0)
- Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (pp. 135-196). North-Holland, Amsterdam. [https://doi.org/10.1016/S1574-0706\(05\)01004-9](https://doi.org/10.1016/S1574-0706(05)01004-9)
- Tsutsumi, M., & Seya, H. (2009). Hedonic approaches based on spatial econometrics and spatial statistics: Application to evaluation of project benefits. *Journal of Geographical Systems*, 11, 357-380. <https://doi.org/10.1007/s10109-009-0099-3>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).