Combining forecasts: Performance and coherence

Mary E. Thomson a, Andrew C. Pollock, Dilek Önkal, M. Sinan Gönül

2019.10.24

Content

- Introduction
- Methodology: Performance and coherence measures
- Demonstration
- Summary

Introduction

- There is general agreement in many forecasting contexts that **combining individual predictions** leads to **better final forecasts**
- However, the relative error reduction in a combined forecast depends upon the extent to which the component forecasts contain **unique/independent information**
- Unfortunately, obtaining independent predictions is difficult in many situations, as these forecasts may be based on similar statistical models and/or overlapping information

Introduction

- The current study addresses this problem by incorporating a measure of **coherence** into an analytic evaluation framework so that the degree of independence between sets of forecasts can be identified easily
- This analytical framework makes an important contribution to the extant work by decomposing the performance and coherence using an integrated method in order to **illustrate the underlying aspects that are responsible for the error reduction** displayed in composite predictions

Methodology

- mean squared error for performance(MSEP)
- mean squared error for coherence (MSEC)

-Decompose MSEP and MSEC into **bias, resolution** and **error variation**, where bias is the difference between the mean of the forecasts and the mean of the actual values **(measuring under-/overestimation)**; resolution is effectively the slope coefficient of a linear relationship between the forecasts and the actual values **(measuring discrimination ability)**; and error variation is the **variation in the actual values that is not explained by variation in the forecasts**.

Methodology (notation)

- Forecasts f_{ij}: i denotes the forecaster (i = 1,2,. . . , n) and j denotes the specific forecast period (j = 1,2,. . .k)

- Composite forecasts f_{mj}

$$f_{mj} = \frac{1}{n}\sum_{i=1}^{n} f_{ij}$$

- The mean of the composite forecasts for all periods $M(f_m)$

$$M(f_m) = \frac{1}{k}\sum_{j=1}^{k} f_{mj}$$

Methodology (notation)

- The perfect forecaster (PF), who would make forecast changes that were precisely in line with the actual value, such that $f_{ij} = a_j$ for all j and a_j is the actual value
- The constant value forecaster (CVF), who would make all predictions with a constant value,c_i, such that$f_{ij} = c_i$ for all j.

Methodology: Performance measures

- The mean squared error for performance

$$MSEPI_i = \frac{1}{k}\sum_{j=1}^{k}(f_{ij} - a_j)^2$$

$$MSEPM = \frac{1}{k}\sum_{j=1}^{k}(f_{mj} - a_j)^2$$

Methodology: Performance measures

- The mean squared error for performance

$$MSEPI_i = BPI_i + RVPI_i + EVPI_i$$

$$MSEPM = \mathrm{BSPM} + \mathrm{RVPM} + \mathrm{EVPM}$$

Bias squared for performance

- The bias (B) is measured by the difference between the mean of forecast values, M(f), and the mean of the actual values, M(a)

$$BSPI_i = {B_i}^2$$

$$BSPM = {B_m}^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} B_i\right)^2$$

where

$$B_i = M(f_i) - M(a) \text{ and } B_m = M(f_m) - M(a)$$

Resolution variation for performance

- The resolution variation component of performance (RVP) for individual and composite forecasters is related to the slope, which is a measure of discrimination that reflects the ability to detect a one point change in the actual value

$$RVPI_i = (1 - SL_i)\ {}^2V(a)$$

$$\mathrm{RVPM} = (1 - SL_m)\ {}^2V(a)$$

Error variation for performance

$$EVPI_i = SC_i = V(u_i)$$

$$EVPM = SC_m = \frac{1}{n^2} V(\sum_{i=1}^{n} u_i)$$

where

$$V(u_i) = V(f_i) - {SL_i}^2 V(a)$$

Methodology: Coherence Measures

- Bias squared for coherence

$$BSC_{hi} = [M(f_h) - M(f_i)]^2 = (B_h - B_i)^2$$

- Resolution variation for coherence

$$RVC_{hi} = (SL_h - SL_i)^2 V(a)$$

- Error variation for coherence

$$EVC_{hi} = SC_h + SC_i - 2SC_{hi}$$

where

$$SC_{hi} = \mathrm{C}(u_h, u_i) = (\frac{1}{k}\sum_{j=1}^{k} u_{hj}u_{ij})$$

Linking performance and coherence measures

- A link exists between the performance measures for individual composite forecasts and the coherence measures between the individual pairs of forecasts

$$MSEPM = \frac{1}{n}\sum_{i=1}^{n} MSEPI_i - \frac{1}{n^2}\sum_{h=1}^{n-1}\sum_{i=2,i>h}^{n} MSEC_{hi}$$

Linking performance and coherence measures

$$BSPM = \frac{1}{n}\sum_{i=1}^{n} BPI_i - \frac{1}{n^2}\sum_{h=1}^{n-1}\sum_{i=2,i>h}^{n} BSC_{hi}$$

$$RVPM = \frac{1}{n}\sum_{i=1}^{n} RVPI_i - \frac{1}{n^2}\sum_{h=1}^{n-1}\sum_{i=2,i>h}^{n} RVC_{hi}$$

$$EVPM = \frac{1}{n}\sum_{i=1}^{n} EVPI_i - \frac{1}{n^2}\sum_{h=1}^{n-1}\sum_{i=2,i>h}^{n} EVC_{hi}$$

- The composite MSEPM (BSPM,RVPM,EVPM) measure will be less than the sum of the individual measures, provided that a degree of diversity exists, as measured by the MSEC for any pair of forecasters

The relative percentage improvement of composite forecasts

$$RMSEPM = 100 * (\frac{1}{n^2}\sum_{h=1}^{n-1}\sum_{i=2,i>h}^{n} MSEC_{hi})/(\frac{1}{n}\sum_{i=1}^{n} MSEPI_i)$$

Demonstration

Table 1
Actual and forecast yearly Q4 RPI percent inflation values and forecast errors.

Q4 Year	Actual	Forecasts				Forecast errors			
		F1	F2	F3	F4	F1	F2	F3	F4
1998	3.0	2.6	3.1	2.9	3.6	−0.4	0.1	−0.1	0.6
1999	1.5	1.5	1.5	1.3	1.6	0.0	0.0	−0.2	0.1
2000	3.1	2.9	2.5	3.1	3.4	−0.2	−0.6	0.0	0.3
2001	1.0	2.2	2.3	2.0	2.4	1.2	1.3	1.0	1.4
2002	2.5	2.4	2.5	2.6	2.1	−0.1	0.0	0.1	−0.4
2003	2.6	2.8	2.7	2.6	3.0	0.2	0.1	0.0	0.4
2004	3.4	3.4	2.7	2.5	3.0	0.0	−0.7	−0.9	−0.4
2005	2.4	2.9	2.6	2.8	2.0	0.5	0.2	0.4	−0.4
2006	4.0	2.9	1.9	1.5	2.5	−1.1	−2.1	−2.5	−1.5
2007	4.2	3.0	3.5	2.1	2.3	−1.2	−0.7	−2.1	−1.9
2008	2.7	2.1	2.1	1.7	2.2	−0.6	−0.6	−1.0	−0.5
2009	0.6	−0.4	−1.5	−1.5	−0.2	−1.0	−2.1	−2.1	−0.8
2010	4.7	3.4	2.9	4.7	2.8	−1.3	−1.8	0.0	−1.9
2011	5.1	3.9	3.5	2.8	3.5	−1.2	−1.6	−2.3	−1.6
2012	3.1	3.3	3.6	3.2	3.3	0.2	0.5	0.1	0.2
2013	2.6	3.0	3.3	2.5	2.3	0.4	0.7	−0.1	−0.3
2014	1.9	2.7	3.5	2.8	2.3	0.8	1.6	0.9	0.4
Mean	2.8	2.6	2.5	2.3	2.5	−0.2	−0.3	−0.5	−0.4
SD	1.2	1.0	1.2	1.3	0.9	0.8	1.1	1.1	0.9

Demonstration

Table 2
Correlations between forecast errors.

Forecaster	F1	F2	F3	F4
F1	1			
F2	0.904	1		
F3	0.816	0.812	1	
F4	0.850	0.777	0.739	1

Demonstration

Table 3

Performance measures: individual and composite for sets of forecasters.

Measure	MSEP	BiasSqP	Bias	Mean	ResVarP	Slope	ErrVarP	Variance
F1	0.584	0.050	−0.224	2.624	0.202	0.620	0.331	0.868
F2	1.257	0.112	−0.335	2.512	0.253	0.575	0.892	1.354
F3	1.425	0.268	−0.518	2.329	0.207	0.615	0.950	1.479
F4	0.957	0.137	−0.371	2.476	0.378	0.480	0.441	0.763
C1234	0.899	0.131	−0.362	2.485	0.256	0.572	0.512	0.970
C123	0.951	0.129	−0.359	2.488	0.220	0.603	0.602	1.110
C124	0.823	0.096	−0.310	2.537	0.273	0.558	0.454	0.889
C134	0.843	0.137	−0.371	2.476	0.257	0.571	0.449	0.906
C234	1.048	0.166	−0.408	2.439	0.275	0.557	0.607	1.040
C12	0.851	0.078	−0.279	2.568	0.227	0.597	0.546	1.045
C13	0.880	0.137	−0.371	2.476	0.205	0.617	0.538	1.071
C14	0.708	0.088	−0.297	2.550	0.284	0.550	0.336	0.758
C23	1.224	0.182	−0.426	2.421	0.229	0.595	0.813	1.308
C24	0.992	0.125	−0.353	2.494	0.312	0.527	0.555	0.944
C34	1.051	0.197	−0.444	2.403	0.286	0.547	0.567	0.986
PF	0.000	0.000	0.000	2.847	0.000	1.000	0.000	1.398
RWF	2.622	0.011	0.106	2.953	1.239	0.058	1.372	1.377
CVF(2)	2.115	0.718	−0.847	2.000	1.398	0.000	0.000	0.000
CVF(3)	1.421	0.023	0.153	3.000	1.398	0.000	0.000	0.000
CVF(4)	2.727	1.329	1.153	4.000	1.398	0.000	0.000	0.000

Demonstration

Table 4
Coherence measures between pairs of forecasts.

Measure	F1,F2	F1,F3	F1,F4	F2,F3	F2,F4	F3,F4
MSEC	0.277	0.495	0.251	0.466	0.460	0.559
BiasSqC	0.012	0.087	0.022	0.033	0.001	0.022
ResVarC	0.003	0.000	0.027	0.002	0.013	0.026
ErrVarC	0.262	0.409	0.202	0.431	0.446	0.512

Demonstration

Table 5

Relative percentage improvement of composite forecasts on performance.

Measure	RMSEP	RBiasSqP	RResVarP	RErrVarP
C1234	$\boxed{15}$	8	2	22
C123	13	10	0	17
C124	12	4	6	18
C134	$\boxed{15}$	10	2	22
C234	$\boxed{14}$	4	2	20
C12	8	4	0	11
C13	12	$\boxed{14}$	0	16
C14	8	6	2	13
C23	9	4	0	12
C24	10	0	1	17
C34	12	3	2	18
Mean	11	6	2	17

Summary

- The passage shows that composite forecasts should be obtained by pooling heterogeneous forecasters that show effective performances in specialized aspects that target a customized use of the pooled forecasts
- The framework can be used to aid in the identification of good/superior forecasters who can be included in the formation of composite forecasts, while filtering out other (sub-standard/not-so-good) forecasters, as guided by coherence comparisons among the forecasters