Bayesian Model Averaging을 이용한 풍속의 확률론적 예측

한근희·김찬식·김찬수

공주대학교 응용수학과

(2016년 8월 28일 접수, 2016년 9월 21일 수정, 2016년 9월 23일 게재 확정)

A Probabilistic Forecast of Wind Speed using Bayesian Model Averaging

Keunhee Han · Chansik Kim · Chansoo Kim

Department of Applied Mathematics, Kongju National University

(Received 28 August 2016, Revised 21 September 2016, Accepted 23 September 2016)

Abstract : This paper used the Bayesian model averaging (BMA) with gamma distribution that takes the form of probability density functions to calibrate probabilistic forecasts of wind speed. We considered the alternative implementation of BMA, which was BMA gamma exchangeable model. This method was applied for forecasting of wind speed over Pyeongchang area using 51 members of the Ensemble Prediction System for Global (EPSG). The performances were evaluated by rank histogram, means absolute error, root mean square error, continuous ranked probability score and skill score. The results showed that BMA gamma exchangeable models performed better in forecasting wind speed, compared to the raw ensemble and ensemble mean.

Key Words : bayesian model averaging, continuous ranked probability score, gamma distribution, rank histogram

주요어 : 베이지안 모델 평균, 연속순위확률점수, 감마분포, 순위히스토그램

1. 서론

날씨 예측의 가장 중요한 목적 중 하나는 미래 대기 상태에 대한 정확하고 신뢰할 수 있는 예측을 생성하는 것이다. 특히 2018년에 동계올림픽이 평창에서 개최됨에 따라 경기력 및 경기장 운영 등에 상당한 영향을 주는 온도, 풍속과 같은 예측 정보의 정확도는 동계 올림픽의 성공적인 운영에 매우 중요한 요소 중의 하나이다. 기후 변수들에 대

한 예측은 관측자료 및 대기의 물리법칙 등을 이용하여 대기 운동 상태를 모의할 수 있는 수치적 날씨 예측 시스템(Numerical Weather Prediction, NWP)을 기반으로 생성되지만, 다수의 편미분 방정식들로 구성되는 NWP 모델은 수치적인 해결 방안만을 제공할 수 있으며 초기조건에 매우 강한 의존성을 보이는 단점을 가지고 있다. 신뢰도가 낮은 가능성이 있는 초기조건들과 NWP 모델 자체의 문제점으로 인해 발생하는 불확실성을 줄

이기 위하여 다수의 초기조건들을 이용하여 NWP 모델을 적용한 후 앙상블을 생성하는 것이 가능하다(Leith, 1974). 앙상블 예측치들을 이용하면 앙상블 평균 혹은 중앙값과 같은 전형적인 점 추정치들을 생성하는 것이 가능할 뿐 아니라 미래 날씨 변수들의 분포에 대해 추정 하는 것이 가능해지며 이를 통하여 통계적 예측을 수행할 수 있다(Gneiting and Raftery, 2005). 앙상블 예측 시스템이 최초로 실현된 것은 1990년대로 거슬러 올라가며(Buizza *et al.*, 1993; Toth and Kalnay, 1997) 대기과학 분야에서는 지난 20 년간 매우 폭넓게 활용되고 있다.

최근에 모든 주요 국가들의 국립기상연구소들은 자체적인 앙상블 예측 시스템(Ensemble prediction system, EPS)을 운용하고 있으며 그러한 시스템의 예로서 Météo France 의 PEARP EPS(Descamps *et al.*, 2015) 또는 독일 대기서비스의 COSMO-DE EPS 등을 들 수 있지만 앙상블 예측에 관련하여 가장 잘 알려진 시스템은 European Centre for Medium-Range Weather Forecasts 이다(ECMWF Directorate, 2012). 그러나, 지금까지 운영되고 있는 다수의 EPS들의 분석에 의하면 예측앙상블들은 일반적으로 과소변동성(underdispersive)의 경향을 보이며 따라서 보정(calibration)이 되지 못하고 있음을 알 수 있다. 앙상블 예측의 정확성을 높일 수 있는 방법으로는 앙상블에 대한 통계적 후처리 기법을 적용하여 확률분포를 갖는 확률론적 예측을 제공하는 것이다. Murphy and Winkler (1987)와 Gneiting *et al.*(2007)에 따르면 확률론적 예측의 목적은 관측치들의 확률밀도함수와 그에 해당하는 앙상블 예측치들의 확률밀도함수들 간의 일치성을 제공하기 위해 예측분포의 변동성을 작게 하는 것이다.

다양한 통계적 후처리 기법들 중 가장 널리 사용되는 기법으로써 Bayesian model averaging(BMA;

Raftery *et al.*, 2005) 및 ensemble model output statistics(EMOS) 또는 비 동질성 회귀분석 기법(Gneiting *et al.*, 2005) 등을 들 수 있다. EMOS 기법의 예측확률밀도함수는 앙상블 멤버에 의존하는 하나의 모수적 분포를 사용한다. 온도 및 해수면 기압의 앙상블 예측을 보정하기 위해 EMOS 모델은 Gneiting *et al.*(2005)과 Han *et al.*(2016) 등에 의해 연구되었다. 또한, 풍속의 경우는 Thorarinsdottir and Gneiting(2010)과 Lerch and Thorarinsdottir(2013) 등에 의해 연구되었다. 또 다른 사후처리 방법 중 하나가 BMA이다. EMOS의 예측확률밀도함수는 하나의 모수적 확률분포를 이용한다. 반면에 BMA의 예측확률밀도함수는 앙상블 멤버들에 해당하는 개개의 확률밀도함수들의 가중합으로 표현된다. Raftery *et al.*(2005)은 온도 및 해수면 기압에 대하여, 그리고 Wilson *et al.*(2007)과 Han *et al.*(2016)은 기온에 대해 혼합 정규분포를 갖는 BMA 모델을 적용하였다. 풍속에 대해서는 감마분포(Sloughter *et al.*, 2010) 또는 절단정규분포(Baran, 2014)를 적용하여 분석하였다.

본 연구에서는 평창 지역 경기장 별 상세 풍속 예측치를 제공하기 위한 하나의 방법으로 전지구 앙상블 예측 시스템으로부터 생성된 51개 앙상블 멤버 예측치들과 그에 해당하는 관측치들에 대한 확률론적 예측치를 제공하고 그 성능을 분석하고자 한다. 풍속의 경우 강수량과 마찬가지로 한쪽으로 치우침을 갖는 분포(skewed distribution)의 특성을 가지고 있으며 또한, 주어진 앙상블 멤버 예측치들이 음이 아닌 값들을 가지므로 이러한 특성을 고려하여 감마분포를 적용한 BMA 모델을 고려하고자 한다. 앙상블 멤버 예측치들과 그에 해당하는 관측치들 간의 신뢰성을 평가하기 위해 순위 히스토그램(rank histogram)을 통해 분석하였으며, 확률론적 예측치들의 예측 성능을 평가하기 위해 MAE(Mean Absolute Error), RMSE(Root

Mean Square Error), CRPS(Continuous Ranked Probability Score)와 CRPSS(Continuous Ranked Probability Skill Score)를 이용하여 예측의 정확도를 평가하였다.

다음 장에서는 분석에 사용한 자료들에 대해 간략하게 언급하였고, 3장에서는 감마분포를 갖는 BMA 모델에 대해 간략하게 언급하였다. 관측치와 앙상블 예측치들 간의 신뢰성 및 BMA 모델들을 이용한 확률론적 예측치들의 예측 성능에 대한 결과는 4장에 주어져 있다. 마지막으로 5장에 요약이 주어져 있다.

2. 평창 풍속 자료

본 연구에서 사용된 자료는 평창 지역 26개 관측지점(Table 1)의 지상 풍속 관측 자료와 현재 운영 중인 앙상블 예측 시스템들 중 좋은 예측 성능을 보여주고 있는 ECMWF(European Centre for Medium-Range Weather Forecasts; ECMWF Directorate, 2012)를 통해 생성된 51개 앙상블 멤버 예측치들을 사용하였다. 전지구 앙상블 예측 시스템인 ECMWF는 51개의 exchangeable 앙상블 멤버들로 구성되어 있다. 앙상블 멤버들 중에서 50개의 앙상블 멤버들은 초기조건들에 대해 랜덤하게 변동성(perturbation)을 부여한 상태에서 예측치들이 생성되었고 하나의 앙상블 멤버 예측치는 초기조건에 변동성을 주지 않고 생성되었다. ECMWF는 00UTC와 12UTC 등 하루에 두 번 예측되며 수평해상도는 33km이다.

관측기간은 2013년 5월 1일부터 2016년 3월 18일까지 일 풍속 자료이고 예측시간은 00UTC에서 매 3시간 간격으로 240시간까지(10일)의 예측시간을 갖는다. 초기 자료는 관측치와 앙상블 예측치들이 일별 자료로 구성되었기 때문에 관측지점 별

분석을 위하여 일 자료를 관측지점 별 자료로 변환하였다. 분석에 앞서 51개의 앙상블 예측치들과 그에 해당하는 관측치들 중에서 결측이 존재하는지를 확인하였으며, 본 과정에서 앙상블 모델 예측치 또는 관측치들 중에서 어느 한 값이라도 결측이면 해당 날짜에 해당하는 모든 자료는 제거하였다.

평창 동계올림픽 경기가 2월에 개최되므로 결측을 제거한 관측지점 별로 1월과 2월의 앙상블 예측치들과 그에 해당하는 관측치들을 이용하여 분석하였다. 분석에 사용된 26개의 관측지점들에 대한 지리적 정보는 Table 1에 주어져 있다. 모든 관측지점들로부터 얻어진 결과들을 모두 기술하기에는 분석된 내용이 너무 많고 유사한 결과들이 존재하므로 관측지점 중에서 같은 경기가 열리면서 연관성이 있는 관측지점들을 분석에 사용하였다(Table 2). 분석에 사용된 관측지점들은 먼저, 봅슬레이와 루지 등과 같이 슬라이딩 경기가 개최되는 station 2553(슬라이딩 출발)과 2554(슬라이딩 도착)의 알펜시아 슬라이딩 센터와 설상 종목 중에서 스키점프 경기가 열리는 station 2573(스키점프 K98 출발), 2574(스키점프 K98 점프)와 2575(스키점프 착지) 등의 알펜시아 스키점프 센터이다.

3. Bayesian Model Averaging

Bayesian model averaging는 통계적 모형들을 결합하기 위한 방법이고 동시에 training data를 이용하여 통계적 모형들을 평가하는 방법이다. BMA는 편이가 보정되어지는 방법은 아니지만, BMA가 적용되어지기 전에 임의의 편이된 오차에 대해 개별적으로 보정되어진 후 모형들을 결합하여 사용되어질 수 있다. 회귀분석과 같은 전형적인 통계적 분석에서는 하나의 통계적 모형이 가정되어

Table 1. Geographic characteristics of stations.

ID	Station	Lat.	Lon.	Elevation (m)
2553	Sliding Start	37°39'02"0	128°40'21"3	959
2554	Sliding Finish	37°39'21"2	128°40'32"2	784
2555	Ski Jumping K125 Jump	37°39'70"5	128°40'76"5	859
2557	Biathlon Start	37°39'89"2	128°41'23"5	756
2558	Biathlon Middle	37°39'57"5	128°41'06"3	776
2559	Cross-country Middle	37°39'35"9	128°41'19"3	782
2560	Alpine GS Start	37°36'76"2	128°40'46"0	1417
2561	Alpine GS Finish	37°37'18"0	128°39'53"0	976
2571	Alpine Base	37°28'24"2	128°36'51"5	373
2572	Ski Jumping K125 Start	37°39'40"6	128°40'43"9	884
2573	Ski Jumping K98 Start	37°39'40"5	128°40'45"4	866
2574	Ski Jumping K98 Jump	37°39'41"7	128°40'48"0	831
2575	Ski Jumping Landing	37°39'44"1	128°40'49"7	793
2576	Sliding Middle	37°39'09"1	128°40'33"7	875
2577	Cross-country Start	37°39'50"0	128°41'05"3	768
2578	Alpine GS Middle1	37°36'57"4	128°40'07"7	1302
2579	Alpine GS Middle2	37°36'06"5	128°40'08"5	1185
2580	Cross Start	37°34'39"7	128°18'44"35	903
2581	Cross Finish	37°35'03"8	128°19'19"4	653
2582	Mogul Start	37°34'41"4	128°19'23"5	791
2583	Slopestyle Finish	37°34'47"3	128°19'30"8	704
2584	Alpine DH Start	37°26'43.55"	128°35'56.17"	1370
2585	Alpine Super GS Start	37°27'1.06"	128°36'0.91"	1220
2586	Alpine DH Middle	37°27'26.81"	128°36'6.43"	927
2587	Alpine DH Finish	37°27'51.41"	128°36'10.39"	665
2588	Slopestyle Start	37°34'27.68"	128°19'23.61"	867

진다. 종종 이러한 모형은 자료에 근거하여 여러 가능한 모형들 중에서 하나가 선택되어지지만 분석자는 이 모형이 가장 좋은 모형인지에 대한 확신을 가질 수는 없다.

BMA(Leamer, 1978; Kass and Raftery, 1995; Hoeting *et al.*, 1999)는 이와 같은 문제를 해결하기 위한 방법으로 하나의 최적 모형을 가정하는 것이 아니라 가능한 통계적 모형들을 결합하기 위한 방법으로 고려되어진 모형으로써 Raftery *et al.*(2005)는 BMA를 통계적 모형으로부터 기상

분야로 확장하여 적용하였다. 기본적인 개념은 임의의 예측치가 주어져 있을 때, 즉, 최적의 모형은 존재하지만 어떤 모형이 최적인지 알 수가 없을 때, 최적 모형에 대한 불확실성을 BMA를 통해 평가하는 것이다(Wilson *et al.*, 2007; Kim and Suh, 2013). 각 앙상블 멤버의 예측치를 m_k 라 하고, $\tilde{m}_k$ 를 편이가 보정된 예측치라고 하자. 예측치 m_k 는 조건부 확률밀도함수 $g_k(y|\tilde{m}_k)$ 를 따른다고 하고, 이것은 m_k 가 앙상블에서 최적의 예측치로서 $\tilde{m}_k$ 가 주어져 있을 때의 조건부 확률밀도함수를 의미한

다. BMA 예측모형은 다음과 같다.

$$p(y|m_1, \dots, m_K) = \sum_{k=1}^K w_k g_k(y|\tilde{m}_k) \quad (1)$$

w_k 는 예측치 m_k 가 최적일 때의 사후확률이고 훈련기간 동안 예측치 m_k 의 성능에 근거하고 있다. w_k 는 가중치들이고, $\sum_{k=1}^K w_k = 1$ 이다.

본 연구에서 예측하고자 하는 풍속의 경우, 음이 아닌 실수 값을 가지므로 이러한 특성을 고려한 감마분포를 이용하여 분석하고자 한다. 따라서, 식(1)에서 각 앙상블 멤버들의 조건부 확률밀도함수는 $g_k(y|\tilde{m}_k) \sim G(\alpha_k, \beta_k)$ 인 감마분포를 갖는다고 가정하였다. 각 앙상블 모델 예측치들을 감마분포와 결합하기 위해 각 앙상블 멤버 예측치 m_k 에 대해, $\mu_k = b_0 + b_1 m_k$ 이고 $\sigma_k = c_0 + c_1 m_k$ 로 앙상블 예측치들의 선형함수로 표현하였다. 이들로부터 감마분포의 모수들은 다음과 같이 정의된다.

$$\alpha_k = \frac{\mu_k^2}{\sigma_k^2}, \quad \beta_k = \frac{\mu_k}{\sigma_k^2}$$

여기서 $\mu_k = b_0 + b_1 m_k$ 이고 $\sigma_k = c_0 + c_1 m_k$ 이고 평균은 원 앙상블 예측치들의 선형함수로 표현되었다.

분석에 사용된 풍속 자료는 앞에서 언급한 것처럼, 앙상블들이 exchangeable 멤버들로 구성되어 있으므로 이러한 특성을 반영하는 BMA 감마 exchangeable 모델을 적용하고자 한다(Slougher *et al.*, 2010; Fraley *et al.*, 2010). 먼저, 초기조건에 변동성(perturbation)을 주지 않고 생성된 앙상블 예측치를 m_1 이라고 하고 초기조건에 서로 다른 변동성을 주고 생성된 50개의 앙상블 멤버들을 $m_{l,1}, \dots, m_{l,50}$ 라 할 때, BMA exchangeable 모델은 식(1)로부터 다음과 같이 정의된다.

$$p(y|m_1, m_{l,1}, \dots, m_{l,K}) = w_1 g_1(y|\tilde{m}_1, b_0, b_1, c_0, c_1) + \frac{w_2}{50} \sum_{j=1}^{50} g_j(y|\tilde{m}_{lj}, b_0, b_1, c_0, c_1) \quad (2)$$

BMA 예측치들을 생성하기 위해 훈련기간의 자

료를 이용하여 BMA 모형의 모수들인 w_1, w_2, c_0 와 c_1 을 추정하기 위하여 Slougher *et al.* (2010)이 제시한 방법을 적용하여 분석하고자 한다. 먼저, 식(1)로부터 $\theta = (w_k, c_0, c_1)$, $k=1, \dots, K$ 이고 위치 s 와 시간 t 에 대해 앙상블의 k 번째 예측치를 m_{kst} 라 하고, 그에 해당하는 관측치를 y_{st} 라고 하자. 평균 모수들 b_0 와 b_1 는 선형 회귀식에 의해 추정되어지며, BMA 모수 θ 는 최우추정법에 의해 추정하고자 한다. 최우추정량(maximum likelihood estimator)은 우도함수를 최대화하는 θ 의 값을 구하는 것보다는 로그 우도함수를 이용하여 구하는 것이 편리하다. 로그우도함수는 다음과 같다.

$$l(\theta) = \sum_{s,t} \sum_{k=1}^K w_k g_1(y_{st}|\tilde{m}_{kst})$$

수리적으로 최대화하는 값을 유도할 수 없기 때문에 뉴턴-랩슨과 같은 비선형 최대화 방법을 이용하여 해를 구해야 하며 따라서 EM(expectation-maximization) 알고리즘(Dempster *et al.*, 1977)을 확장한 ECME 알고리즘(Liu and Rubin, 1994)을 이용하고자 한다.

BMA 모델로부터 생성된 예측치들과 관측치들 간의 분포적 특성을 반영하는 확률론적 예측 성능을 평가하기 위해 CRPS(Continuous Ranked Probability Score)를 사용하였다(Unger, 1985; Hersbach, 2000; Gneiting and Raftery, 2007). CRPS는 예측치의 누적확률분포와 관측치의 누적 분포함수 간의 차이를 나타내고 다음과 같이 정의된다.

$$crps(F, y) = \int_{-\infty}^{\infty} [F(t) - I(t \geq y)]^2 dt \quad (3)$$

여기서, y 는 관측치, F 는 예측치의 누적확률분포함수이고 $I(\cdot)$ 는 지시함수이다. 또한, 모든 자료에 대한 crps는 다음과 같이 표현된다.

$$CRPS = \frac{1}{n} \sum_{i=1}^n crps(F_i, y_i)$$

Skill Score는 모델로부터 생성된 예측치들 간 상호 비교를 위해 사용되어지는 방법 중 하나이다. 그 중에서 CRPS(Continuous Ranked Probability Skill Score)는 위에서 언급한 CRPS를 이용하여 다음과 같이 정의된다.

$$CRPS = 1 - \frac{\overline{CRPS}_{forecast}}{\overline{CRPS}_{reference}} \quad (4)$$

여기서 $\overline{CRPS}$ 는 주어진 기간에서 CRPS의 평균을 나타내고 참조 예측치(reference forecast)는 일반적으로 기후 예측치들을 사용하지만 비교하고자 하는 다른 방법에 의해 생성된 예측치들을 사용할 수도 있다. CRPS는 모델로부터 예측된 예측치가 참조된 예측치에 비해 평균적으로 얼마나 개선되었는지를 보여주는 척도이고 최적의 CRPS의 값은 1이다.

4. 실험 결과 및 성능 평가

먼저, 풍속 자료에 대해 앙상블 예측치들과 그에 해당하는 관측치들이 얼마나 유사한지 살펴보기 위해 순위 히스토그램(Rank Histogram; Hamil, 2001)을 이용하여 분석하였다. 순위 히스토그램은 앙상블 멤버들의 예측치들을 크기 순으로 정렬하였을 때, 그에 해당하는 관측치들의 상대적인 순위를 결정하여 이를 히스토그램으로 나타낸 그림이다. 순위 히스토그램이 균등한 형태를 갖는 경우 앙상블 예측치와 그에 해당하는 관측치들은 같은 분포를 따른다고 할 수 있다.

Figure 1은 2014년부터 2016년 1월과 2월 풍속 자료에 대해 선택된 다섯 곳의 관측지점에 대한 앙상블 예측치들과 그에 해당하는 관측치들의 순위 히스토그램을 보여주고 있다. 먼저, 순위 히스토그램에서 보여주듯이 U-형태를 가지고 앙상블 예측

치들이 관측치들에 비해 과소변동성(underdispersion)을 가지고 있음을 알 수 있다. 또한, station 2553과 2573은 음의 편이(negative bias)를 가지고 있고 이것은 앙상블 예측치들이 관측치들에 비해 작게 예측되었음을 의미한다. station 2554, 2574와 2575는 양의 편이(positive bias)를 가지고 있어 앙상블 예측치들이 그에 해당하는 관측치들에 비해 평균적으로 크게 예측되었음을 알 수 있다.

순위 히스토그램에서 균등성으로부터 순위 분포가 얼마나 벗어났는지를 나타내는 척도가 신뢰성 지수(reliability index)이다. 신뢰성 지수는 $\Delta = \sum_{i=1}^c |p_i - 1/c|$ 으로 정의되며, 여기서 c 는 순위 히스토그램에서 분류되는 순위의 수이고 $1/c$ 은 각 순위에서 얻어지는 기대 상대도수이고, p_i 는 i 번째 순위에서 관측되는 기대 상대도수이다. 다섯 곳의 관측지점들에 대한 신뢰성 지수는 Table 2에 주어졌다. Table 2로부터, 신뢰성 지수들이 일정한 값들을 가지고 있어 균등성으로부터 많이 벗어나고 있음을 알 수 있다. 따라서, 앙상블들은 과소변동성을 가지고 있으며 관측지점에 따라 일관되게 음 혹은 양의 편이를 가지고 있어 보정(calibration)이 필요하다는 것을 알 수 있다.

BMA 감마 exchangeable 모델에 적용된 모수들을 추정하기 위해 앙상블 예측치들과 그에 해당하는 관측치들을 훈련기간(training period)과 검증기간(validation period)으로 나누어 분석하였다. 훈련기간은 2013년 1월 1일부터 2015년 2월 28일 기간의 1월과 2월에 해당하는 예측치와 그에 해당하는 관측치들을 사용하였고, 검증기간에 사용된 자료는 2016년 1월 1일부터 2016년 2월 29일에 해당하는 자료를 사용하였다. 훈련기간의 자료들을 이용하여 BMA 모수들을 추정하였고 이를 이용하여 예측확률밀도함수로부터 추정된 예측치와 80% 신뢰영역을 제시하였다.

BMA 감마 exchangeable 모델이 어떻게 적용되

Figure 1. Rank histogram for the selected stations.

Table 2. Reliability indexes for the selected stations.

Station	Station	Elevation (m)	Reliability index
2553	Sliding Start	959	0.7697006
2554	Sliding Finish	784	1.023035
2573	Ski Jumping K98 Start	866	1.038154
2574	Ski Jumping K98 Jump	831	1.440378
2575	Ski Jumping Landing	793	1.415077

Table 3. Summary of ensemble mean, BMA weights, BMA forecast and 80% prediction intervals for the station 2573.

Station	Date	Projection	Obs (m/s)	Ensemble mean (m/s)	BMA forecast (m/s)	80% prediction intervals(m/s)
2573	2016.02.18	3h	6.8	5.15	6.68	(3.95, 10.48)

BMA weight

w_1	0.3004
w_2	0.6996

Figure 2. BMA gamma predictive probability density function (thick curve) and its ensemble members (thin curves) for station 2573 (Feb. 18, 2016, 3h), respectively. The 80% prediction intervals (dotted lines), BMA forecast (thin line) and the observation (thick line) are displayed.

있는지를 설명하기 위해, 관측지점 2573(스키점프 K98 출발)에서 2016년 2월 10일 예측시간이 3h에서 계산된 BMA 결과들이 Table 3과 Figure 2에 주어졌다. 먼저, Table 3으로부터 관측지점 2573에서 2016년 2월 18일 관측된 풍속은 6.8(m/s)이다. 관측지점 2573은 Figure 1로부터 음의 편이를 가지고 있어 양상블 예측치들이 관측치에 비해 평균적으로 작게 예측되고 있음을 알 수 있다. 따라서, 양상블 평균은 관측치보다 작게 예측되고 있다. 반면에 BMA 예측치는 관측된 풍속과 거의 유사하게 추정되었음을 알 수 있다. 또한, 관측된 풍속은 BMA 감마 예측분포로부터 예측된 80% 예측구간에 포함되어 있어 예측의 정도가 좋다는 것을 알 수 있다. 식(2)에서 양상블 예측치 m_1 에 대한 BMA 가중치는 0.3004이고 나머지 m_2 부터 m_{51} 의 50개 양상블 예측치들에 해당하는 BMA 가중치는 0.6996으로 50개 양상블 멤버들의 개개의 가중치는 $0.699/50=0.01398$ 으로 동일한 가중치를 갖는다.

Figure 2는 Table 3에서 얻어진 BMA 결과로부터 생성된 BMA 감마 예측확률밀도함수가 주어졌다. 진한 실선은 BMA 감마 예측확률밀도함수이고 점선은 80% 예측구간의 하한과 상한이다. 가는 수직선과 짙은 수직선은 각각 BMA 예측치와 그에 해당하는 관측치를 표현하고 있다. 예측확률밀도함수에서 일련의 점들은 2016년 2월 18일에 예측된 51개의 양상블 멤버 예측치들이며 관측치에 비해 상당히 작게 예측되고 있음을 알 수 있지만, BMA 감마 모델을 적용했을 때 얻어진 예측치는 관측치와 상당히 유사하다는 것을 볼 수 있고 이것은 Table 2에서 언급한 것과 동일한 결과를 보여주고 있다.

모든 예측시간(3h부터 240h)에서 양상블과 BMA 감마 모델 예측치들의 예측 성능을 비교하기 위해 다섯 곳의 관측지점 2553, 2554, 2573, 2574와 2575에서 계산된 MAE, RMSE와 CRPS를 비교한 그림이 Figure 3과 4에 주어졌다. 원형(o) 점은 예측시간이 3h에서 63h, 녹색 플러스(+)

Figure 3. Comparison of the projection-specified mean MAE and RMSE values (m/s) for the selected stations.

점은 예측시간이 66h부터 129h까지이고 빨간 곱하기($\times$) 점은 예측시간이 132h에서 240h까지의 예측 시간에 따른 MAE와 RMSE의 변화를 나타낸다. Figure 3에서는 각 관측지점에서 모든 예측 시간에 대해서 BMA 감마 모델 하에서 얻어진 예측치들과 앙상블 평균에 대한 MAE와 RMSE 산점도가 주어져 있다. 그림에서 보듯이, 대부분 45도 직선 위쪽에 모든 값들이 분포되어 있음을 알 수 있고 이것은 모든 예측시간에 대해 BMA 감마 모델의 예측치들이 앙상블 평균보다 좋은 예측 성능을 보여주고 있다는 것을 의미한다. 그러나, 관측지점 2553의 경우, 예측시간에 따라 앙상블 평균과 BMA 감마 모델의 예측 성능은 다르게 나

타하고 있지만 전반적으로 BMA 감마 모델의 예측 성능이 다소 좋다는 것을 볼 수 있다(Table 4와 Figure 5). 또한, 관측지점 2553과 2573의 경우, 다른 관측지점들의 예측치들에 비해 45도 직선을 따라 넓게 분포되어 있어 예측의 변동성이 크다는 것을 알 수 있다. 예측시간의 증가에 따른 비교에서 관측지점 2553과 2573의 경우, 대체적으로 예측시간이 길어질수록 예측시간이 짧을 때보다 상대적으로 MAE(RMSE)가 커지는 경향을 가지고 있어 예측시간이 길어질수록 예측성능은 감소하고 있음을 알 수 있다. 그러나, 다른 관측지점에서는 예측시간이 길어지더라도 MAE와 RMSE에는 변화가 거의 없음을 알 수 있다. Figure 5의 CRPS 비교에

Figure 4. Same as Figure 3, but for the mean CRPS values (m/s).

Table 4. Averaged MAE, RMSE and CRPS for the selected stations.

(a) MAE (m/s) and RMSE (m/s)

	Station 2553		Station 2554		Station 2573		Station 2574		Station 2575	
Measure	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Emsemble mean	1,511	2,040	2,103	2,477	2,495	3,206	2,598	2,961	2,264	2,583
BMA Gamma	1,478	1,971	0,928	1,206	2,096	2,713	0,563	0,909	0,704	0,966

(b) CRPS (m/s)

	Station 2553	Station 2554	Station 2573	Station 2574	Station 2575
Measure	CRPS	CRPS	CRPS	CRPS	CRPS
Raw ensemble	1,226	1,696	2,129	2,156	1,843
BMA Gamma	1,049	0,684	1,455	0,425	0,496

Figure 5. Comparison of prediction skills of the selected stations in terms of MAE (m/s), CRPS (m/s), MAESS and CRPSS.

Figure 6. Comparison of CRPS of raw ensemble and BMA gamma exchangeable model forecasts for all stations.

서도 Figure 4와 동일한 결과를 보여주고 있다.

선택된 관측지점들에서, BMA 감마 exchangeable 모델과 앙상블 간 예측 성능을 보다 상세하게 비교하기 위해 MAE, CRPS, Mean Absolute Error Skill Score(MAESS)와 CRPSS를 이용하여 비교하였다. Figure 5a와 5c에서는 관측지점에서 얻어진 MAE와 CRPS에 대한 그래프가 주어지고 있다. 모든 관측지점에서 BMA 감마 모델이 앙상블보다는 좋은 예측성능을 보여주고 있다. 두 예측치들 간에 보다 상세한 차이를 비교하기 위해 MAESS와 CRPSS를 사용하였다. MAESS는 식 (4)에서 CRPS 대신 MAE를 사용하였고 식 (4)에서 분모의 참조 예측치(reference forecast)는 앙상블을 사용하였다. Figure 5b와 5d로부터 두 척도 값들이 모두 양의 값을 가지고 있어 BMA 감마 모델의 MAE(혹은 CRPS)가 앙상블들의 MAE(혹은 CRPS)보다 평균적으로 작다는 것을 알 수 있고 이것은 BMA 감마 모델의 예측성능이 더 좋다는 것을 의미한다.

Table 4는 모든 예측시간에서 앙상블, 앙상블 평균과 BMA 감마 모델 예측치들의 MAE, RMSE와 CRPS를 앞에서 선택된 다섯 곳의 관측지점들 각각에 대해 평균한 값이 주어지고 있다. Figure 5에서 언급한 것처럼 BMA 감마 모델 예측치들이 MAE, RMSE와 CRPS 관점에서 좋은 예측 성능을 보여주고 있음을 볼 수 있다.

모든 관측지점들에 대해서 앙상블 예측치와 BMA 감마 모델 예측치들의 CRPS를 비교한 그림이 Figure 6에 주어지고 있다. 삼각점선과 점선은 각각 앙상블과 BMA 감마 모델 예측치들의 CRPS이다. 그림에서 보듯이, 모든 관측지점들에서 BMA 감마 모델의 예측치들이 앙상블들에 비해 좋은 예측성능을 보여주고 있다. 따라서, 앙상블 모델에서 생성된 예측치들보다는 더 개선된 예측성능을 보여주고 있다.

5. 결론

정확한 기상 정보를 제공하기 위해, 앙상블 수치예보 시스템을 적용하고 있지만, 모델이 갖는 한계성과 불확실성으로 인해 편이(bias)와 변동성(dispersion)을 가지게 된다. 이러한 문제점을 해결하기 위한 하나의 방법은 미래 날씨 사건들에 대해 확률분포를 갖는 확률론적 예측치를 제공하는 것이다. 본 연구에서 풍속의 경우, exchangeable 멤버들로 구성된 앙상블 예측치들로서 두 개의 exchangeable 그룹을 가지므로 확률론적 예측치를 생성하기 위해 Bayesian model averaging에 감마분포를 결합한 BMA 감마 exchangeable 모델을 적용하였다.

먼저, 주어진 앙상블 멤버 예측치들과 그에 해당하는 관측치들의 연관성을 살펴보기 위해 순위 히스토그램과 신뢰성 지수를 계산하여 분석하였다. 그 결과, 앙상블 예측치들이 관측된 풍속 자료보다 평균적으로 과소 혹은 과대 추정하고 있으며 변동성은 과소하게 모의되고 있음을 알 수 있었다. 이와 같은 편이와 변동성을 보정하기 위해 통계적 사후처리 방법인 BMA 감마 exchangeable 모델을 적용한 확률론적 예측치를 생성하였다. BMA 감마 모델의 모수들은 주어진 훈련기간에서 최대우도추정 방법을 통해 추정되었고 추정된 예측확률밀도함수를 이용하여 예측치와 80% 예측구간을 제시하였다.

확률론적 예측치들의 예측 성능을 비교하기 위해 MAE, RMSE와 CRPS를 이용하여 예측의 정도를 비교하였다. 먼저, 예측시간에 따른 예측성능 평가에서, 선택된 다섯 곳의 관측지점에 대해 BMA 감마 모델이 앙상블 예측치들보다 좋은 예측 성능을 보여주었다. 그러나, 관측지점에 따라 예측시간에 대한 예측의 변동성과 성능은 다르게 나타나고 있음을 알 수 있었다. 또한, 모든 관측지

점에서의 비교에서도 BMA 감마 모델의 예측치들이 앙상블들에 비해 평균적으로 예측 오차를 개선하고 있음을 알 수 있고 모든 관측지점에서 좋은 예측 성능을 보여주었다. 그러나, 관측지점의 지역적 특성에 따라 다소 다른 예측 성능을 보여주고 있어 이런 관측지점에 따른 특성을 반영할 수 있는 사후처리기법을 개발하고 적용하여 그 예측 성능을 비교해 보는 것이 필요하다.

사사

본 연구는 기상청 기상국립과학원 주요사업 '2018 평창동계올림픽 기상지원'의 일환으로 수행되었습니다.

참고문헌

- Baran, S., 2014, Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components, *Computational Statistics and Data Analysis*, 75, 227-238.
- Buizza, R., Tribbia, F., Molteni, F., and Palmer, T. N., 1993, Computation of optimal unstable structures for a numerical weather prediction model, *Tellus, A*, 45(5), 388-407.
- Dempster, A. P., Laird, N. M., and Rubin, D. B., 1977, Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society Series B*, 39, 1-38.
- Descamps, L., Labadie, C., Joly, A., Bazile, E., Arbogast, P., and Cébron, P., 2015, PEARP, the Météo-France short-range ensemble prediction system, *Quarterly Journal of the Royal Meteorological Society*, 141(690), doi:10.1002/qj.2469.
- ECMWF Directorate, 2012, Describing ECMWF's

- forecasts and forecasting system, *ECMWF Newsletter*, 133, 11-13.
- Fraley, C. A., Raftery, A. E., and Gneiting, T., 2010, Calibrating multimodel forecast ensembles with exchangeable and missing members using Bayesian model averaging, *Monthly Weather Review*, 138(1), 190-202.
- Gneiting, T., Balabdaoui, F., and Raftery, A. E., 2007, Probabilistic forecasts, calibration and sharpness, *Journal of the Royal Statistical Society Series B*, 69(2), 243-268.
- Gneiting, T. and Raftery, A. E., 2005, Weather forecasting with ensemble methods, *Science*, 310, 248-249.
- Gneiting, T. and Raftery, A. E., 2007, Strictly proper scoring rules, prediction, and estimation, *Journal of the American Statistical Association*, 102(477), 359-378.
- Gneiting, T., Raftery, A. E., Westveld, A. H., and Goldman, T., 2005, Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation, *Monthly Weather Review*, 113(5), 1098-1118.
- Hamill, T. M., 2001, Interpretation of rank histograms for verifying ensemble forecasts, *Monthly Weather Review*, 129(3), 550-560.
- Han, K., Kim, C., and Kim, C., 2016, A probabilistic forecast of temperature in Pyeongchang using homogeneous and nonhomogeneous regression models, *Journal of Climate Research*, 11(1), 87-103 (in Korean).
- Han, K., Choi, J., and Kim, C., 2016, Calibration of probabilistic forecast of temperature in PyeongChang area using Bayesian model averaging, *Journal of Climate Research*, 11(1), 49-67 (in Korean).
- Hersbach, H., 2000, Decomposition of the continuous ranked probability score for ensemble prediction systems, *Weather Forecasting*, 15(5), 559-570.
- Hoeting, J. A., Madigan, D. A., Raftery, A. E., and Volinsky, C. T., 1999, Bayesian model averaging: A tutorial, *Statistical Science*, 14(4), 382-401.
- Kass, R. E. and Raftery, A. E., 1995, Bayes Factors, *Journal of the American Statistical Association*, 90(430), 773-795.
- Kim, C. and Suh, M. S., 2013, Prospects of using Bayesian model averaging for the calibration of one-month forecasts of surface air temperature over South Korea, *Asia-Pacific Journal of Atmospheric Sciences*, 49(3), 301-311.
- Leamer, E. E., 1978, *Specification Search -Ad Hoc inference with nonexperimental data*, Wiley (New York), 370pp.
- Leith, C. E., 1974, Theoretical skill of Monte-Carlo forecasts, *Monthly Weather Review*, 102(6), 409-418.
- Lerch, S. and Thorarinsdottir, T. L., 2013, Comparison of non-homogeneous regression models for probabilistic wind speed forecasting, *Tellus A*, 65:21206, doi:10.3402/tellusa.v65i0.21206.
- Liu, C. and Rubin, D. B., 1994, The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence, *Biometrika*, 81(4), 633-648.
- Murphy, A. H. and Winkler, R. L., 1987, A general framework for forecast verification, *Monthly Weather Review*, 115(7), 1330-1338.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M., 2005, Using Bayesian model averaging to calibrate forecast ensembles, *Monthly Weather Review*, 133, 1155-1174.
- Sloughter, J. M., Gneiting, T., and Raftery, A. E., 2010, Probabilistic wind speed forecasting using ensembles and Bayesian model averaging, *Journal of the American Statistical Association*, 105(489), 25-35.
- Thorarinsdottir, T. L. and Gneiting, T., 2010, Probabilistic forecasts of wind speed: Ensemble model

- output statistics by using heteroscedastic censored regression, *Journal of the Royal Statistical Society Series A*, 173, 371-388.
- Toth, Z. and Kalnay, E., 1997, Ensemble forecasting at NCEP: the breeding method, *Monthly Weather Review*, 125(12), 3297-3318.
- Unger, D. A., 1985, A method to estimate the continuous ranked probability score in *Preprints, Ninth Conference on Probability and Statistics in Atmospheric Sciences, Virginia Beach, Virginia, Boston: American Meteorological Society*, 206-213.
- Wilson, L. J., Beauregard, S., Raftery, A. E., and Verret, R., 2007, Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging, *Monthly Weather Review*, 135(4), 1364-1385.

책임편집: 이승호
영문교열: 이은결