BAYESIAN FUSION OF ENSEMBLE FORECASTS

A Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Master of Science

by

Zachary Armentrout

May

2013

APPROVAL SHEET

The thesis

is submitted in partial fulfillment of the requirements

for the degree of

Master of Science

_____
AUTHOR

The thesis has been read and approved by the examining committee:

Roman Krzysztofowicz
_____
Advisor

Donald Brown
_____

Guofen Yan
_____

_____

_____

_____

Accepted for the School of Engineering and Applied Science:

Dean, School of Engineering and Applied Science

May

2013

ABSTRACT

The probabilistic format of forecasts has been shown to add economic value to numerical weather prediction (NWP) models. To estimate the probability distributions of meteorological variates, a forecast center generates an ensemble forecast—a collection of two or more deterministic forecasts that verify at the same time. The spread of an ensemble characterizes the uncertainty about a meteorological variate due to uncertain initial conditions of the atmosphere.

This thesis builds on the analysis of Nah Youn Lee from 2010 by comparing the statistical properties of two ensemble forecasting systems for surface temperature—one from the Canadian Meteorological Centre (CMC), and one from the National Centers for Environment Prediction (NCEP)—and by fusing the two ensemble forecasts via a Bayesian processor. Each of these systems generates a 20-member ensemble in 12-hour steps for up to 16 days into the future.

A statistical analysis of the CMC ensemble supports these findings: (i) The ensemble mean is consistently among the most informative predictors of central tendency. (ii) An optimal combination of ensemble members is more informative than the ensemble mean; however, its composition varies with lead time and season. (iii) The best predictor of uncertainty varies across lead times and seasons with no apparent trend, but the width of the 90% central credible interval and an amalgam of ensemble range and coefficient of kurtosis are frequently among the best, while the ensemble standard deviation is not.

A statistical analysis of a combination of the CMC and NCEP ensembles supports these findings: (i) The most informative predictor of central tendency is the affine combination of ensemble means. (ii) An optimal combination of ensemble members (among all 40) is more informative than the affine combination of means; however, its composition varies with lead time and season. (iii) The best predictor of uncertainty is either the arithmetic mean or the geometric mean of the same predictors from each ensemble, although all predictors of uncertainty are generally poor.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Numerical Weather Prediction

Numerical weather prediction (NWP) is the use of mathematical models to forecast the weather. Modern NWP models incorporate a multitude of meteorological variates, both current and historical, to make predictions for a wide range of variates, over large areas and many lead times. Consequently, a NWP model has a substantial computational magnitude. At the National Centers for Environmental Prediction (NCEP) of the National Weather Service (NWS), for instance, current weather conditions around the globe are modeled every six hours, and forecasted for up to two weeks ahead, using observations taken at many ground stations and airborn sensors. In the northern hemisphere, these observations are used to define a meteorological state on a 1° x 1° grid, at several altitudes. A meteorological state in the continental United States is defined on a finer grid (5 km x 5 km) containing about 300,000 points. Current weather conditions cannot be measured exactly at every forecast point, and thus are estimated by applying a model to observations taken at relatively few stations. These estimates of current states (i.e., forecasts with zero lead time) are referred to as *analysis*, and serve as the initial conditions of a NWP model.

Like any forecast, a weather forecast contains some degree of uncertainty. The need to characterize uncertainty probabilistically has been recognized in several domains requiring risk analysis and decision-making, including NWP (Krzysztofowicz 1983, Krzysztofowicz 2001). One way to forecast future atmospheric states probabilistically is to generate a representative sample of trajectories of a stochastic process driven by a probability distribution of the initial state. This approach is central to a method employed by many meteorological centers today, known as *ensemble forecasting*.

1.2 Ensemble Forecasting

1.2.1 Definition

An *ensemble forecast* is a collection of two or more deterministic forecasts that verify at the same time (Sivillo et al. 1997). Together, the forecasts in an ensemble approximately represent the uncertainty in a modeled system state, as well as provide a means of characterizing the uncertainty in future states. The trajectory of a single deterministic forecast, conversely, does not take into account potential error in the initial predictand observation, and thus can quickly deviate from actual observations over time. Expanding on this logic empirically, Cloke and Pappenberger (2009) and Toth et al. (2001) suggest that the economic value of an ensemble forecast system is substantially greater than that of a single-forecast system.

1.2.2 Ensemble Generation

There are several ways to generate ensemble members. The NCEP uses the breeding method, a procedure that simulates the growth of errors in the forecast cycle by carrying forward initial condition errors into subsequent forecast cycles. First, a small, random perturbation is added to the predictand observation to initialize each of 20 ensemble forecast trajectories, resulting in 20 different arbitrary perturbations. Next, the NWP model is run to produce the control forecast and perturbed forecasts for the next six hours. The vector of differences between the perturbed forecasts and control forecast, a "bred vector," is then rescaled and added to the control to form a new perturbed initial condition for the following forecast cycle. After a transient period of forecast cycles (3–4 days), the bred vectors become dominated by the fastest-growing instabilities of the dynamical process. Both experimentally and operationally, the breeding method has been shown to offer superior forecast skill to that of smoothed control and randomly generated (Monte Carlo) ensemble forecasts (Toth and Kalnay 1993).

The Canadian Meteorological Centre (CMC) instead generates each ensemble member with a separate perturbed Global Environmental Multiscale (GEM) model. Each GEM model has a

different combination of parameters and/or physical parameterizations, and is initialized with a perturbed predictand realization generated by an ensemble Kalman filter. Every six hours, an ensemble of random perturbation fields is added to the state estimates to maintain sufficient ensemble spread (Charron et al. 2010).

1.2.3 Notation

Realizations of all variates in this analysis are indexed by k, denoting the day of the year: k = 1, 2, … , 366. (An index for 29 February is included because the data contain a leap year.) The predictand (forecasted variate) is $W(k)$, and its realization is $w(k)$. The analysis on day k, not the actual predictand observation, is used as a proxy for $w(k)$ to verify forecasts. The predictors include high resolution forecast $Y(k)$ with realization $y(k)$, low resolution control forecast $Y_0(k)$ with realization $y_0(k)$, and ensemble forecast—a vector $Y(k) = (Y_1(k),\ldots,Y_J(k))$ with realization $y(k) = (y_1(k), \ldots , y_J(k))$, where J is the number of ensemble members, equal to 20 in this study. An arbitrary predictor is denoted by X, with corresponding realization x.

1.3 Overview

Chapter two discusses the benefits and challenges of ensemble forecasts, introduces the Bayesian processor of ensemble, and the need for approximately sufficient statistics of an ensemble.

Chapter three discusses pre-processing of CMC ensemble data, standardization of ensemble members, normality of ensemble members and predictand, seasonality, and correlation and conditional correlation between ensemble members.

Chapter four discusses the theory behind the Bayesian processor of ensemble and the Gaussian-Gamma model. Also, the definition of a sufficient statistic is given along with the motivation behind the search for approximately sufficient statistics.

Chapter five details the main results of hypothesizing and evaluating predictors of central tendency. In addition, methodologies are outlined for single predictors and the combination of predictors, evaluated using the Informativeness Score.

Chapter six summarizes secondary results for hypothesized predictors of central tendency, which include the mean of a subset of ensemble members, and predictors constructed from the upper/lower mean and majority/minority mean.

Chapter seven details the methodology for evaluating predictors of uncertainty. Predictors of uncertainty are hypothesized and evaluated, and the assumptions of the Gaussian-Gamma model are validated.

Chapter eight makes comparisons between the statistical properties of the NCEP and CMC ensembles. The comparisons motivate the combination of the two ensembles in order to produce more informative predictors of central tendency and uncertainty.

Chapter nine is an exploratory analysis of the fusion of NCEP and CMC ensembles. Various predictors of central tendency and uncertainty from the joint sample of ensemble forecasts are hypothesized and evaluated. More hypothesized predictors are to be evaluated. The hypothesized Gaussian-Gamma model has yet to be evaluated.

Chapter 10 presents conclusions, recommendations for operational forecasting centers, and suggestions for future research.

2. BACKGROUND

2.1 Ensemble Forecasting Systems

As per the methodology of Lee (2010), this research focuses on forecasts of 2-meter temperature at the Savannah, GA, weather station. Both the NCEP and CMC generate 20-member ensemble forecasts of this variate. Figure 2.1 shows examples of the CMC and NCEP ensemble forecasts, respectively, made on a specific date. The CMC ensemble forecasts exhibit much greater spread than do the NCEP ensemble forecasts; further, the two sets of ensemble forecasts vary in central tendency and uncertainty with increasing lead time. Ensemble forecast systems have been shown to exhibit higher skill than systems producing a single deterministic forecast. An ensemble forecast system makes extended-range forecasts feasible, as the ensemble members collectively provide information about forecast uncertainty. Further, a best estimate of a predictand (the primary feature of a single-forecast system) can be obtained from ensemble members, for example, as their mean, median, or some other statistic.
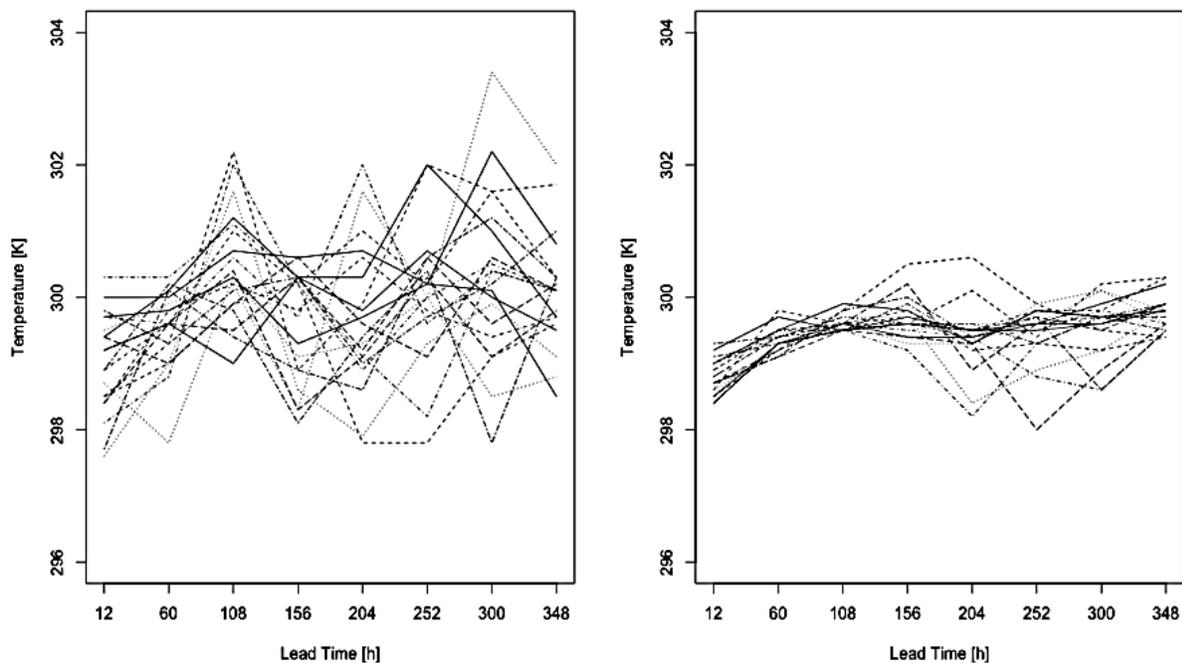
Figure 2.1. Twenty-member ensemble forecasts of surface temperature in Savannah, GA, made by CMC (left) and NCEP (right) on 24 July 2008 with lead times of 12, 60, 108, 156, 204, 252, 300, 348 h.

Meteorological centers that employ ensemble forecasting systems face a number of challenges. First, an estimate of a probability distribution directly from the ensemble members is often poorly calibrated, especially at longer lead times (Gneiting et al. 2004). In addition, an ensemble often underestimates total forecast uncertainty, as the ensemble spread directly reflects only the uncertainty due to initial atmospheric conditions. The ensemble members also may be stochastically dependent or have non-identical distributions, thereby invalidating the interpretation of the ensemble as a random sample of predictand trajectories. These challenges motivate the operational use of a statistical processor, which can extract more useful information from an ensemble. One such processor, the Bayesian processor of ensemble, is outlined in the following section.

2.2 Bayesian Processor of Ensemble

To address the challenges identified above as they apply to the NCEP ensemble, Lee (2010) chooses to process the ensemble statistically using the Bayesian processor of ensemble (BPE), which takes the form (Krzysztofowicz 1983):

$$\phi(w|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|w)\,g(w)}{\kappa(\boldsymbol{y})}, \tag{2.1}$$

$$\kappa(\boldsymbol{y}) = \int f(\boldsymbol{y}|w)g(w)dw, \tag{2.2}$$

where $\boldsymbol{y}$ is the vector of 20 ensemble members, treated as a realization of random vector $\boldsymbol{Y}$, g is the prior density function of the temperature predictand W, f is the family of likelihood functions, κ is the marginal density function of $\boldsymbol{Y}$, and ϕ is the family of posterior density functions of W.

While the BPE defined in (2.1) and (2.2) requires the processing of all 20 ensemble members, there may exist a summary statistic X that provides as much information about ensemble central tendency as $\boldsymbol{y}$ provides. In addition, there may exist a predictor T—perhaps a measure of ensemble spread, a higher-order moment, or some combination thereof—that contains information

about total uncertainty. Replacing vector $\boldsymbol{y}$ with statistic (X, T) in the BPE may allow for more efficient processing, in that the likelihood function of (X, T) may be more easily estimated than is the likelihood function of all 20 ensemble members. Depending on the joint distribution of ensemble members, such a statistic will not necessarily be sufficient, and must therefore be hypothesized based on theoretical models and empirical analyses. For the NCEP ensemble forecast system, Lee (2010) found that the combination of ensemble mean (a predictor of central tendency) and ensemble range (a predictor of ensemble spread) exhibited comparable informativeness to that of the full ensemble. This study, likewise, hypothesizes and tests several predictors of central tendency and uncertainty within the CMC ensemble, with the goal of finding a similar replacement statistic for statistical processing.

2.3 Ensemble Spread

Among the common ensemble summary statistics, ensemble spread (i.e. standard deviation) has been especially popular for its presumed value as a predictor of uncertainty. While ensemble spread is a function of forecast error, it is also a function of season, lead time, and geographical location—that is, it not only reflects error due to initial condition, but also error due to atmospheric instability. Therefore, ensemble spread more accurately constitutes an upper bound on forecast error rather than the expected forecast error. Toth et al. (2001) instead suggest taking a probabilistic approach to uncertainty estimation. In this approach, for a single forecast, predictability is expected to vary with the number of ensemble members that fall within a climatically equiprobably interval centered at the ensemble mean. This measure of predictability is normalized by season, lead time, and geographical location. Additionally, the measure takes into account the forecast value with respect to the climatological distribution, with wider climatically equiprobable intervals corresponding to more extreme forecasts.

Ensemble spread is but one hypothesized predictor of uncertainty. The informativeness of ensemble spread is evaluated and compared to the informativeness of other hypothesized predictors in Chapter 7.

3. EMPIRICAL PROPERTIES OF THE CMC ENSEMBLE

3.1 Pre-Processing Ensemble Data

3.1.1 The Ensemble Data

The CMC data contain complete forecasts and predictand realizations for 599 days, made on dates from 07/11/2007 to 02/28/2009. The CMC ensemble forecast system employs a control forecast, but does not employ a high-resolution forecast. Forecasts are made every 12 hours (00 UTC and 12 UTC). In this study, focus is limited to forecasts made at 00 UTC with one of the following eight lead times: 12 h, 60 h, 108 h, 156 h, 204 h, 252 h, 300 h, 348 h.

3.1.2 Prior Mean and Variance

Prior mean m_k and prior variance s_k^2 of the predictand are defined as follows for each day k of the year:

$$m_k = E(W(k)), \tag{3.1}$$

$$s_k^2 = Var(W(k)). \tag{3.2}$$

Lee (2010) estimates the prior mean and variance using a 40-year (1959-1998) climatic sample. For each day k, temperatures from a 15-day sampling window $\{k - 7, ..., k, ..., k + 7\}$ over 40 years are incorporated to obtain a substantially larger sample size. (40 years $\times$ 15 days $= 600$ observations.) In both the prior and joint samples, a high resolution forecast at 12 UTC with lead time 0 h is used as analysis (the best available proxy for a true realization of W). A missing analysis datum on 11/12/2008 was filled in by linear interpolation using the analysis from the previous day (11/11/2008) and from the following day (11/13/2008). Similarly, the climatic mean and standard deviation for 29 February (day 60 of the year) were each filled in by linear interpolation using climatic estimates from the previous day (m_{59}, s_{59}^2) and from the following day (m_{61}, s_{61}^2).

3.1.3 Standardization

Temperature is non-stationary around the year. Therefore, the predictand W, control forecast Y_0, and ensemble members Y_j ($j = 1, \ldots, 20$) will be standardized using climatic estimates of mean m_k and standard deviation s_k of temperature on day k of the year at 12 UTC, which are obtained from 40 years (1959–1998) of observations. This method of standardization has been shown to bring temperature to near margin-stationarity (Krzysztofowicz and Evans 2007), and is performed using the following formulae:

$$w'(k) = \frac{w(k) - m_k}{s_k}, \quad k = 1, 2, \ldots, 366, \tag{3.3.1}$$

$$y_j'(k) = \frac{y_j(k) - m_k}{s_k}, \quad k = 1, 2, \ldots, 366, \quad j = 0, 1, \ldots, 20, \tag{3.3.2}$$

where $w(k)$ and $y_j(k)$ are, respectively, realizations of W and Y_j. When it is unnecessary to distinguish between days of the year, the standardized realizations are notated as follows:

$$w = w'(k), \tag{3.4.1}$$

$$y_j = y_j'(k), \quad j = 0, 1, \ldots, 20. \tag{3.4.2}$$

See Figure 3.1 for plots of all 20 original and standardized CMC ensemble forecasts with lead time 108 h. The standardized forecasts show a seasonal trend: Forecasts for roughly the middle six months (April through September) of each year have a greater central tendency and wider spread than do forecasts for the remaining months. This seasonality is addressed in the next section.
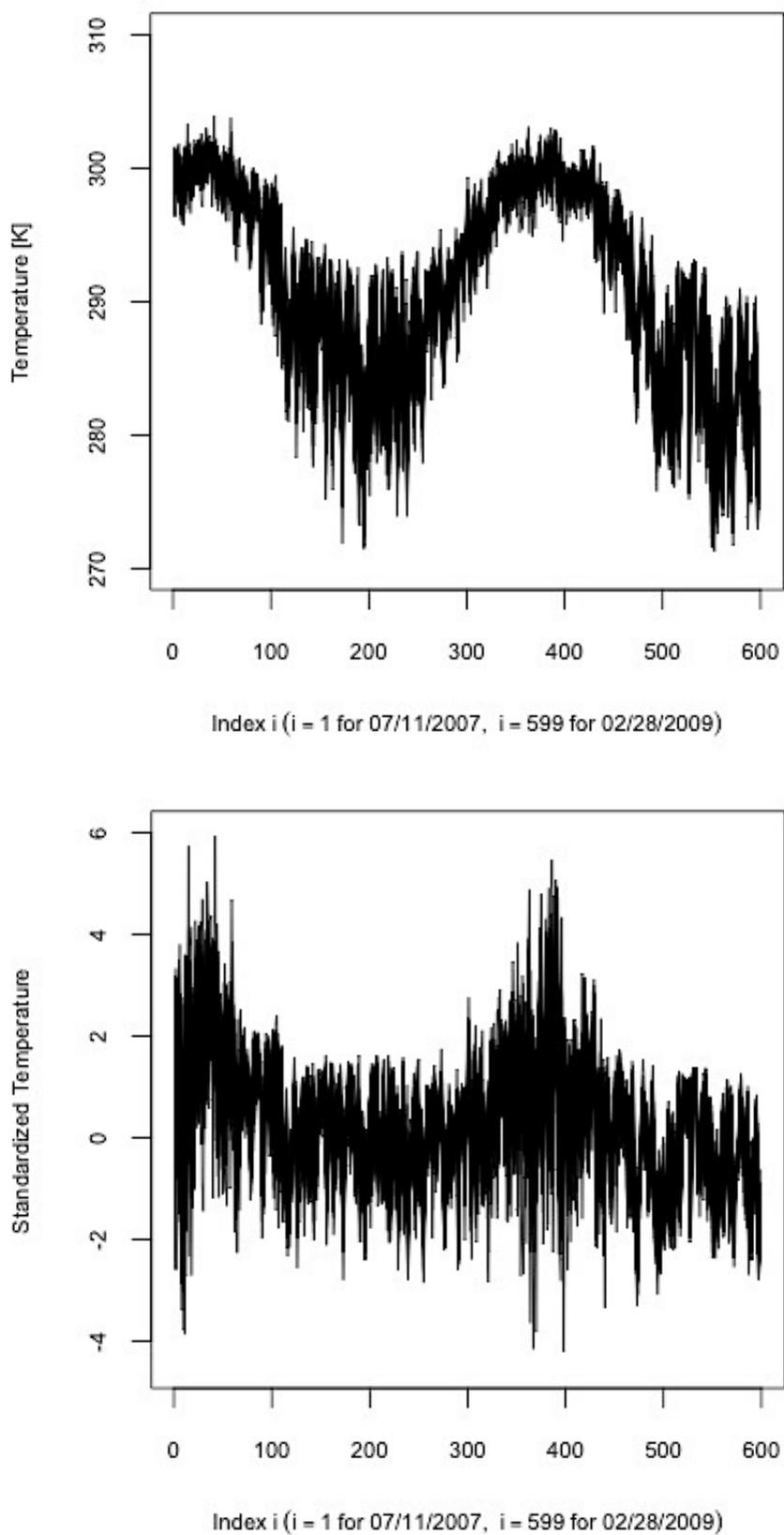
Figure 3.1. Original (top) and standardized (bottom) CMC forecasts consisting of 20 ensemble members, with lead time 108 h. Standardization brings the ensemble forecasts to near margin stationarity.

3.1.4 Seasonality

As is the case for the NCEP standardized ensemble data, the standardized CMC ensemble data exhibit a pronounced seasonal trend. Specifically, forecasts for days in April through September exhibit greater central tendency and larger spread than do forecasts for days in the remaining months of the year. Thus, the year is divided into warm season (April – Sept.) and cool season (Oct. – March.) as in Lee (2010). The division of available data is as follows:

Warm Season – 7/11/2007 to 09/30/2007, 04/01/2008 to 09/30/2008

Cool Season – 10/01/2007 to 03/31/2008, 10/01/2008 to 02/28/2009

Thus, of a total of 599 days of forecasts, 265 fall in the warm season and 334 days fall in cool season. However, a forecast with lead time greater than 12 hours corresponds to a predictand realization on a subsequent day. Therefore, a set of forecasts with lead time $12 + 24n$ ($n = 1, 2, ..., 15$) has a reduced number of days in the joint sample of ensemble forecasts and predictand. Table 3.1 shows the number of matching days in the warm (N_{warm}) and cool (N_{cool}) seasons for lead times 12 h, 60 h, 108 h, 156 h, 204 h, 252 h, 300 h, and 348 h.

Table 3.1. Number of matching days between predictand and ensemble forecasts for warm and cool seasons, for lead time $12 + 24n$, n = 0, 2, 4, ..., 14.

	Lead Time [hours]							
	12	60	108	156	204	252	300	348
N_{warm}	265	263	261	259	257	255	253	251
N_{cool}	334	334	334	334	334	334	334	334

3.2 Normality of Ensemble Members

This section examines the distributions of the predictand and ensemble members. First, for each season and lead time, the empirical distribution function of each variate is plotted using the meta-Gaussian plotting positions. Each variate appears to follow a Gaussian distribution. The least-squares method is used to fit a Gaussian distribution for each variate. A consistently low maximum absolute difference (MAD) between the empirical and hypothesized distributions is

consistent with the hypothesis that the predictand and each CMC ensemble member closely follows a Gaussian distribution. In Figure 3.2, the empirical distribution and fitted Gaussian distribution is plotted for the predictand and each of two ensemble members, with lead time 108 hours.

The distributions of the ensemble members exhibit various biases when compared to the distribution of the predictand. The distribution of the predictand superposed with the distributions of two ensemble members is shown in Figure 3.3 and Figure 3.4. In the warm season, the distributions have large variation in the lower tails at short lead times, and larger variation in the upper tails at long lead times. In the cool season, the distributions diverge in the lower tails, and converge at higher forecast values. The variety of distributional differences may be due to the unique combination of parameters and parameterizations corresponding to each ensemble member's underlying GEM model.

Next, the distributions of the ensemble members are compared for each lead time and season. The empirical distribution functions of all 20 ensemble members are superposed in Figure 3.5 (warm season) and Figure 3.6 (cool season), for lead times 12, 108, 204, and 300 hours. In the warm season, the empirical distribution functions of the ensemble members vary substantially, and thus the ensemble members are not likely identically distributed. In the cool season, the empirical distribution functions appear approximately identical at short lead times, but spread substantially as lead time increases. Perhaps none of the ensemble members are truly identically distributed because each is output from a unique GEM model.
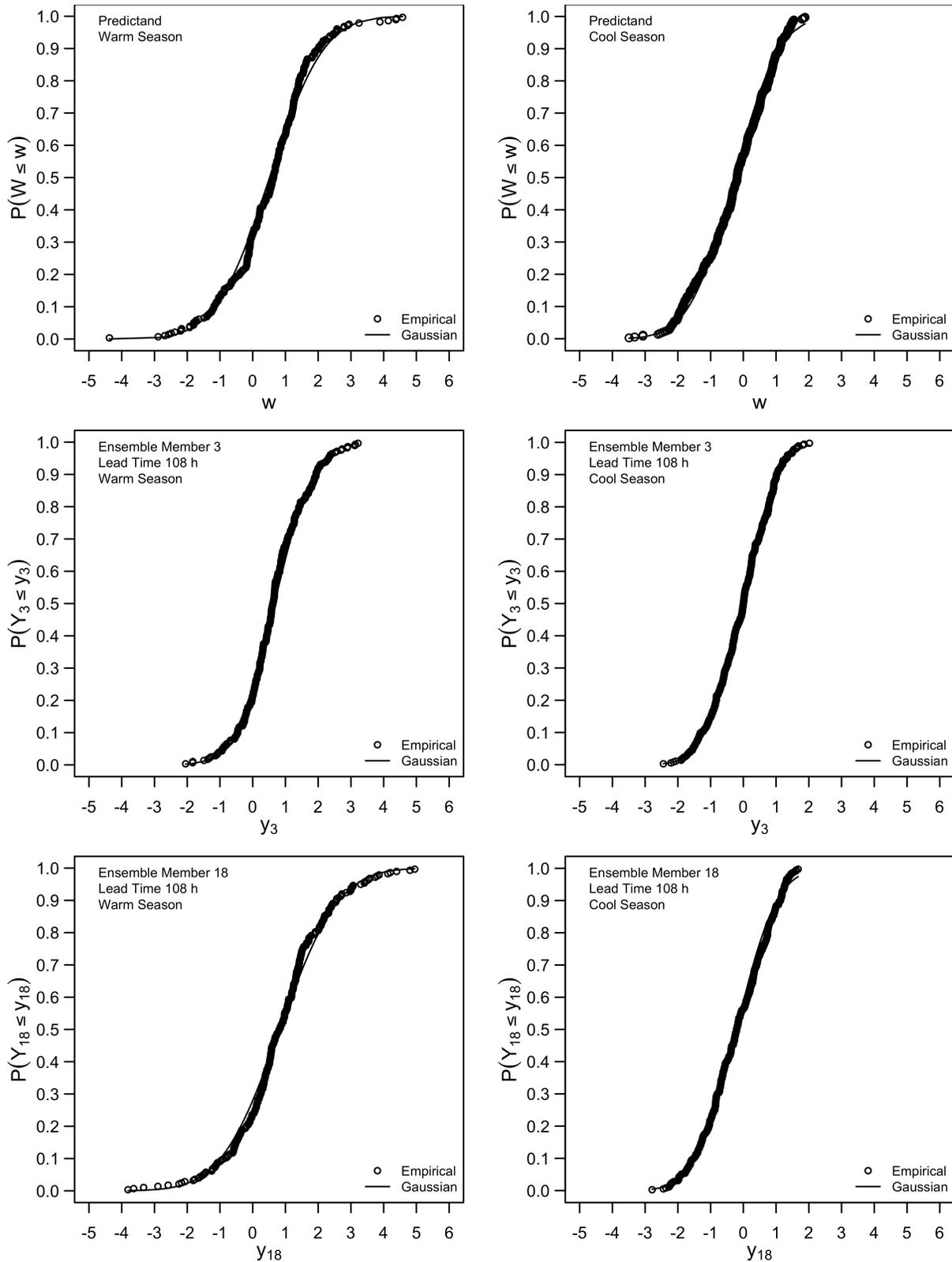
Figure 3.2. Empirical distribution and fitted Gaussian distribution of predictand W and two ensemble members, Y_3 and Y_{18}, with lead time 108 h in warm and cool seasons. Each variate appears to be normally distributed in each season.
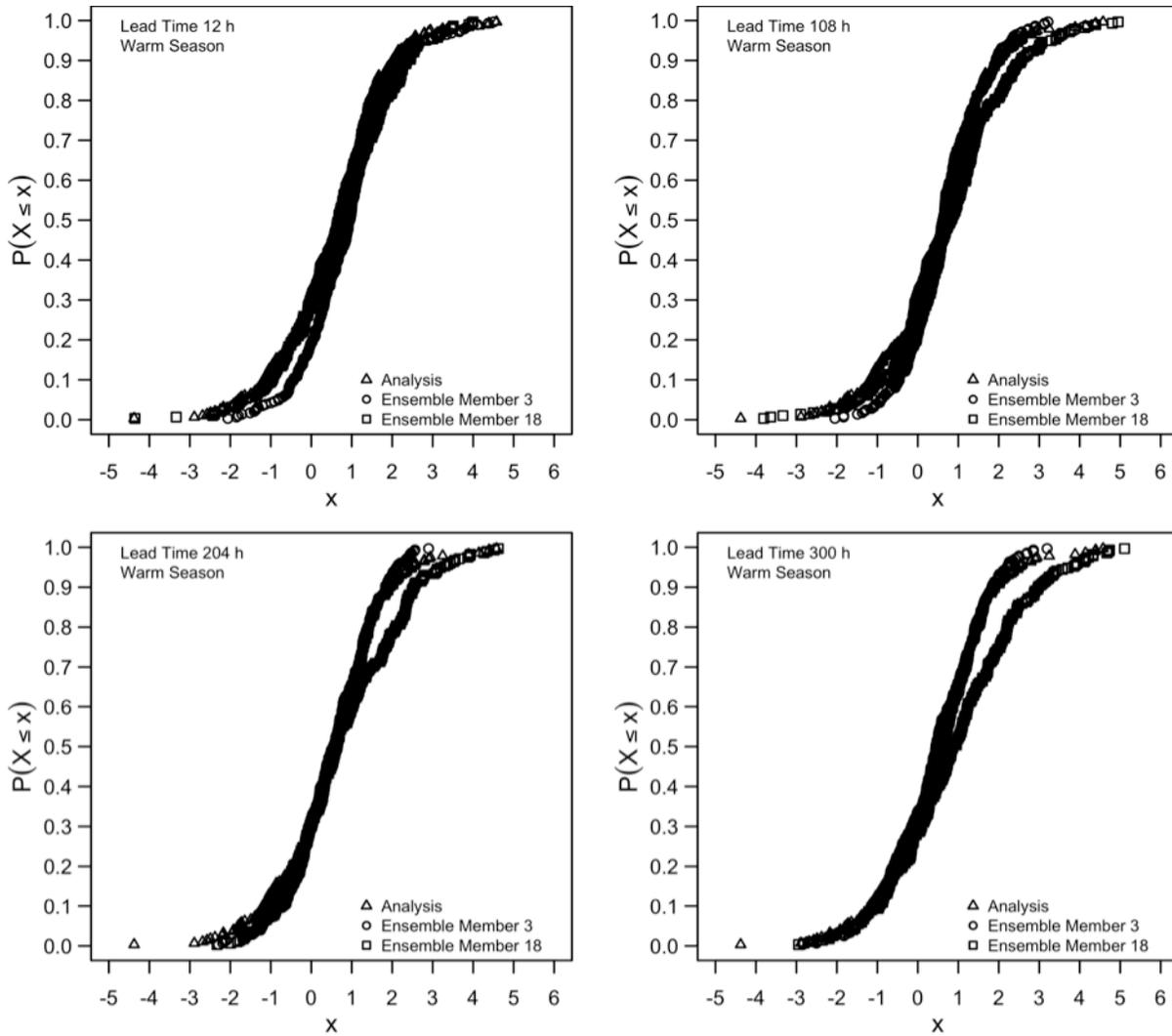
Figure 3.3. Empirical distributions of predictand W and two ensemble members, Y_3 and Y_{18}, with lead times 12 h, 108 h, 204 h, and 300 h, in the warm season. For lead time 12 h, the distributions appear to deviate at the lower tails and merge toward the upper tails. Conversely, for lead times 108 h, 204 h, and 300 h, the distributions appear to deviate at the upper tails and merge toward the lower tails.
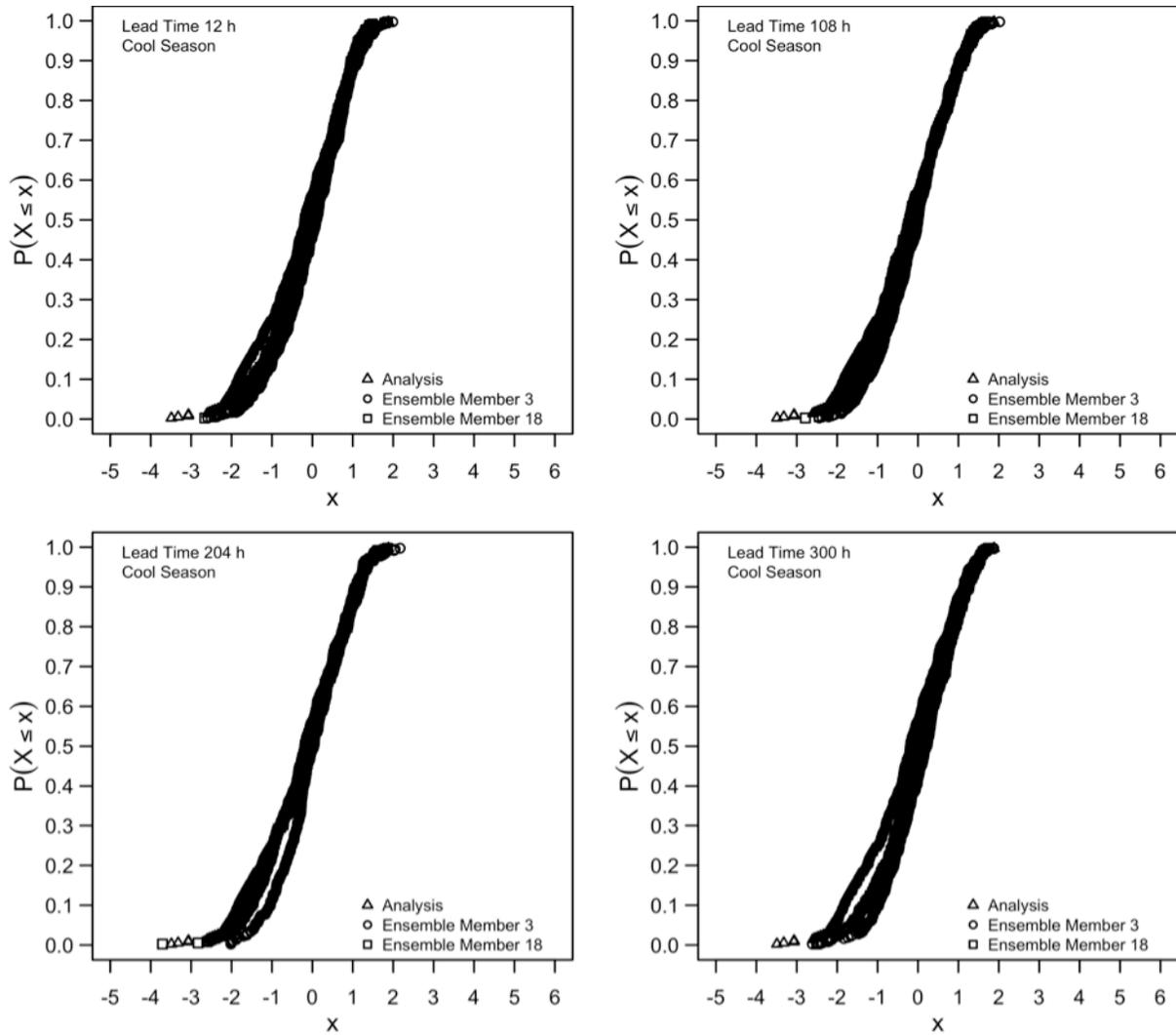
Figure 3.4. Empirical distributions of predictand W and two ensemble members, Y_3 and Y_{18}, with lead times 12 h, 108 h, 204 h, and 300 h, in the cool season. The distributions appear to deviate substantially at the lower tails, and gradually merge with increasing realization x.
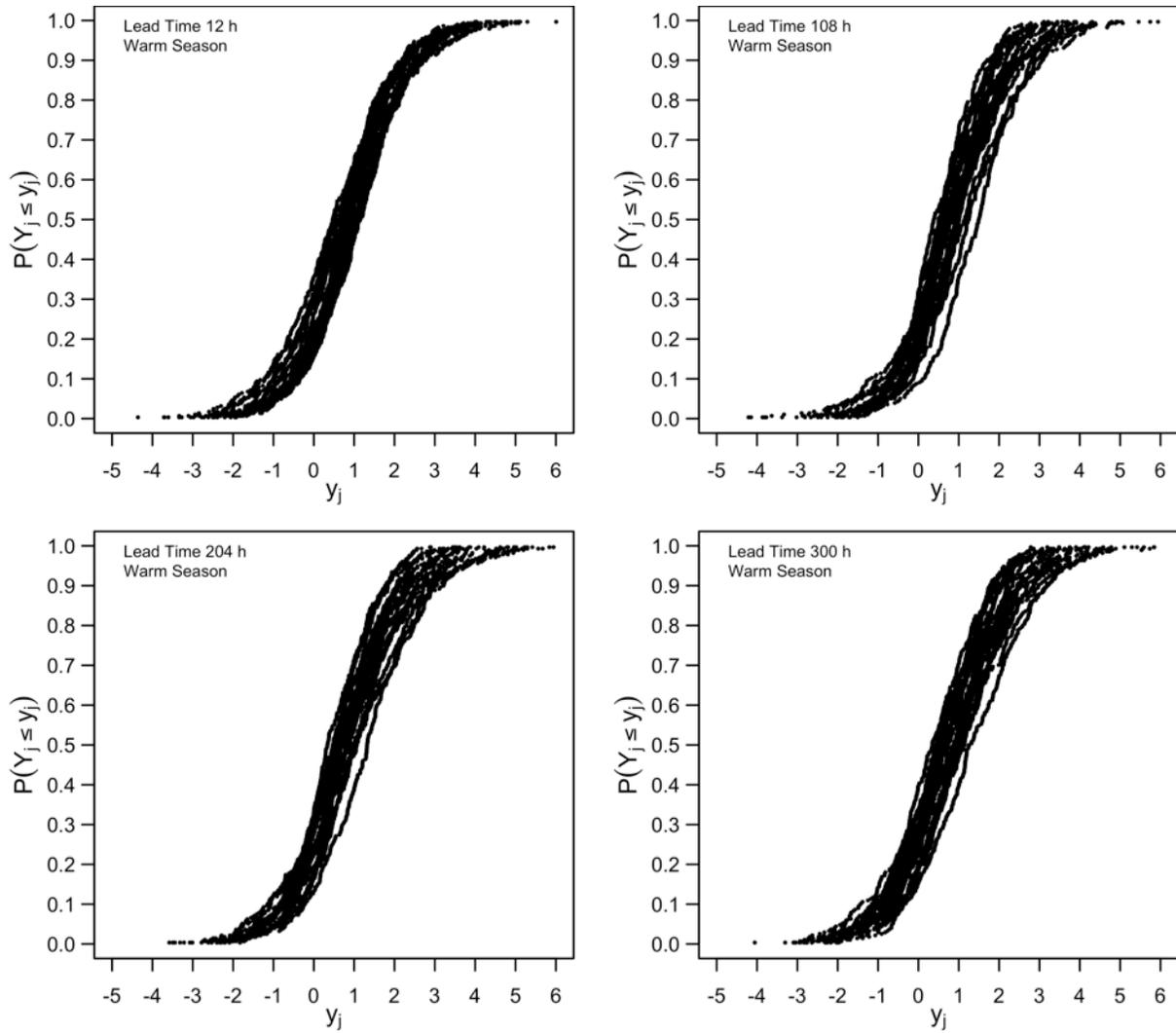
Figure 3.5. Empirical distributions of all 20 ensemble members for lead times 12 h, 108 h, 204 h, and 300 h in the warm season. The ensemble members do not appear to be identically distributed.
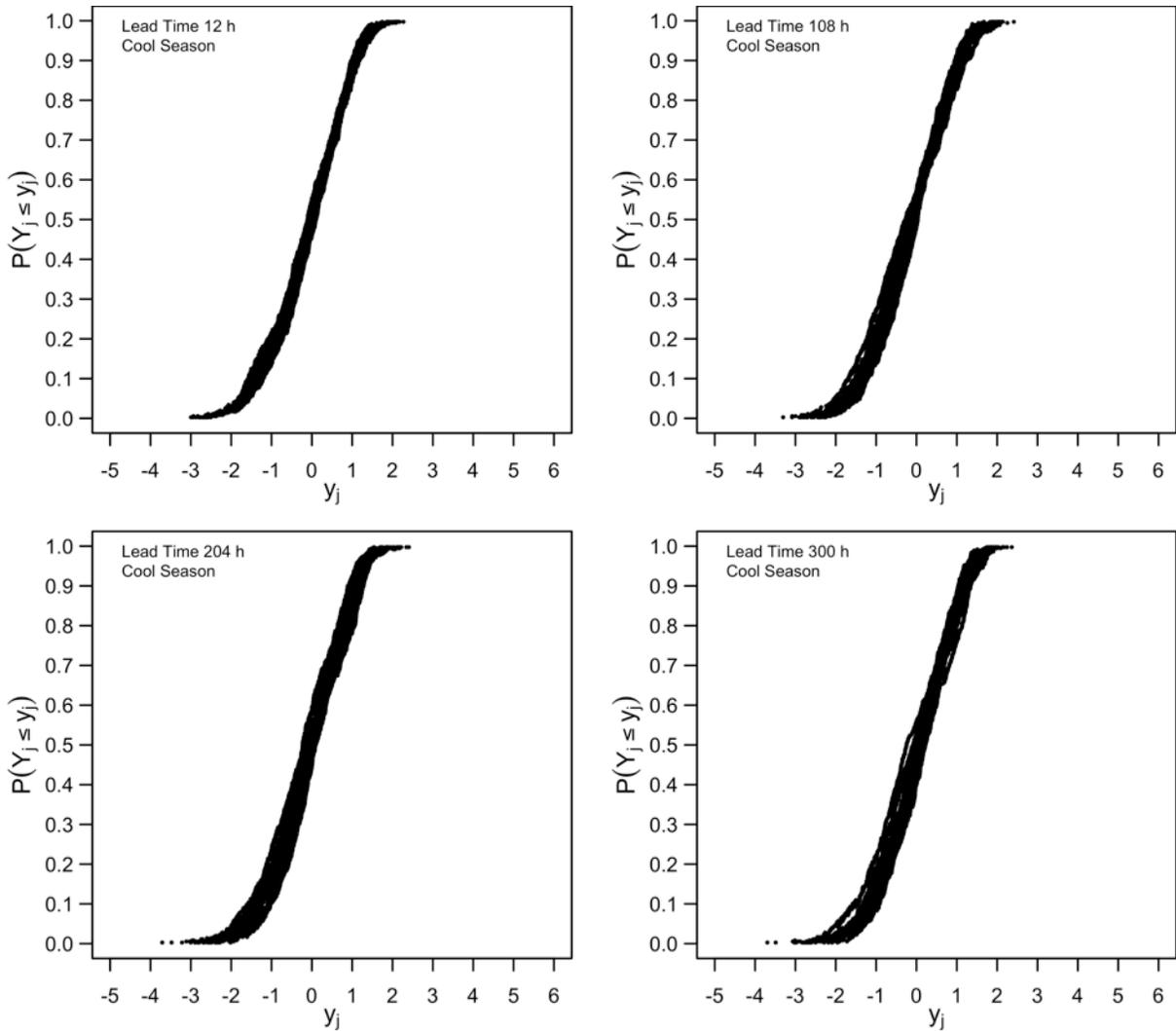
Figure 3.6. Empirical distributions of all 20 ensemble members for lead times 12 h, 108 h, 204 h, and 300 h in the cool season. The ensemble members appear to be approximately identically distributed at lead times 12 h and 108 h, and do not appear to be identically distributed at lead times 204 h and 300 h.

3.3 Correlation between Ensemble Members

This section examines the degree to which the ensemble members are pairwise correlated. The correlation between any two predictors X_i and X_j is given by the following formula:

$$\text{Cor}(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\text{Var}^{1/2}(X_i)\text{Var}^{1/2}(X_j)}. \tag{3.5}$$

Correlation matrices for lead times 12 h, 108 h, 204 h, and 300 h, for warm and cool seasons, are shown in Tables 3.3–6. Each table contains the pairwise correlations among all 20 ensemble members and the low resolution control forecast. Unsurprisingly, the correlation between each pair of variates generally decreases as lead time increases.

3.4 Conditional Correlation between Ensemble Members

The stochastic dependence between a pair of predictors (X_i, X_j) and the predictand W conforms approximately to the normal-linear model, which is detailed in Chapter 5. Under this model, the conditional correlation between any two predictors X_i and X_j, given realization w of predictand W, is given by the following formula:

$$\text{Cor}(X_i, X_j| W = w) = \frac{\text{Cov}(X_i, X_j| W=w)}{\text{Var}^{1/2}(X_i | W=w)\text{Var}^{1/2}(X_j| W=w)} \tag{3.6.1}$$

$$= \frac{\sigma_{ij}}{\sigma_i \sigma_j}; \tag{3.6.2}$$

the estimators σ_{ij} and σ_j are defined by

$$Var(X_i | W = w) = \sigma_i{}^2, \tag{3.7.1}$$

$$Cov(X_i, X_j| W = w) = \sigma_{ij}, \tag{3.7.2}$$

and their values are calculated using the residuals $\theta_i(n)$ from the normal-linear model (Section 5.2.1) as follows:

$$\sigma_i^2 = \frac{1}{N}\sum_{n=1}^{N}\theta_i^2(n), \tag{3.8.1}$$

$$\sigma_{ij} = \frac{1}{N}\sum_{n=1}^{N}\theta_i(n)\theta_j(n). \tag{3.8.2}$$

The conditional correlation matrices are given for the same lead times and seasons in Tables 3.7–10. The correlation between two ensemble members is reduced when conditioned on a realization of the predictand. For shorter lead times, the ensemble members seem to be more correlated in the cold season, and for longer lead times, the ensemble members seem to be more correlated in the warm season.

The average correlation and average conditional correlation are shown in Table 3.2 for each of the lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons.

Table 3.2. Average correlation and average conditional correlation between pairs of ensemble members for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons. The average conditional correlations are well above zero. Thus, for any lead time and season, the ensemble members cannot be considered a random sample.

Lead Time:	12 h		108 h		204 h		300 h	
Season:	Warm	Cool	Warm	Cool	Warm	Cool	Warm	Cool
Correlation	0.7816	0.9606	0.5046	0.7745	0.3963	0.3632	0.3743	0.1821
Conditional Correlation	0.4079	0.5602	0.3384	0.5143	0.3471	0.2870	0.3457	0.1679

Table 3.3. Correlation matrices for lead time 12 h, in warm and cool seasons.

Correlation Matrix Lead Time 12 h Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.80	0.72	0.70	0.65	0.71	0.69	0.69	0.72	0.78	0.76	0.69	0.64	0.72	0.79	0.69	0.73	0.68	0.78	0.73	0.68
y_1		1.00	0.69	0.69	0.64	0.66	0.63	0.65	0.72	0.68	0.74	0.67	0.60	0.66	0.76	0.68	0.74	0.65	0.71	0.71	0.64
y_2			1.00	0.83	0.85	0.85	0.72	0.85	0.84	0.71	0.88	0.84	0.83	0.88	0.68	0.83	0.81	0.85	0.73	0.88	0.86
y_3				1.00	0.83	0.86	0.71	0.81	0.85	0.72	0.83	0.88	0.83	0.86	0.71	0.82	0.84	0.85	0.69	0.84	0.88
y_4					1.00	0.88	0.71	0.87	0.83	0.66	0.86	0.83	0.88	0.90	0.65	0.85	0.78	0.90	0.69	0.87	0.86
y_5						1.00	0.75	0.87	0.85	0.70	0.89	0.86	0.86	0.90	0.66	0.85	0.81	0.89	0.74	0.91	0.86
y_6							1.00	0.68	0.73	0.70	0.71	0.74	0.66	0.75	0.69	0.63	0.67	0.68	0.72	0.71	0.73
y_7								1.00	0.83	0.68	0.87	0.81	0.85	0.87	0.64	0.88	0.79	0.87	0.73	0.89	0.84
y_8									1.00	0.74	0.86	0.89	0.78	0.84	0.71	0.83	0.86	0.84	0.75	0.85	0.87
y_9										1.00	0.72	0.73	0.66	0.72	0.68	0.70	0.70	0.70	0.77	0.72	0.72
y_{10}											1.00	0.83	0.83	0.89	0.68	0.86	0.83	0.87	0.74	0.92	0.82
y_{11}												1.00	0.79	0.86	0.70	0.80	0.84	0.85	0.74	0.85	0.90
y_{12}													1.00	0.88	0.61	0.83	0.76	0.87	0.67	0.85	0.84
y_{13}														1.00	0.67	0.85	0.80	0.89	0.74	0.89	0.88
y_{14}															1.00	0.63	0.71	0.68	0.69	0.66	0.69
y_{15}																1.00	0.82	0.85	0.70	0.87	0.80
y_{16}																	1.00	0.78	0.72	0.82	0.83
y_{17}																		1.00	0.70	0.89	0.85
y_{18}																			1.00	0.72	0.76
y_{19}																				1.00	0.84
y_{20}																					1.00

Correlation Matrix Lead Time 12 h Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.98	0.97	0.96	0.96	0.96	0.96	0.98	0.98	0.98	0.98	0.96	0.96	0.96	0.96	0.97	0.97	0.96	0.98	0.97	0.96
y_1		1.00	0.97	0.95	0.95	0.95	0.95	0.96	0.96	0.96	0.97	0.95	0.95	0.95	0.95	0.96	0.96	0.95	0.97	0.96	0.95
y_2			1.00	0.96	0.96	0.96	0.95	0.97	0.97	0.96	0.97	0.96	0.96	0.96	0.95	0.97	0.96	0.96	0.96	0.97	0.96
y_3				1.00	0.97	0.97	0.96	0.95	0.96	0.95	0.95	0.97	0.97	0.97	0.96	0.95	0.95	0.97	0.96	0.95	0.97
y_4					1.00	0.98	0.97	0.96	0.96	0.95	0.96	0.98	0.97	0.98	0.96	0.96	0.95	0.98	0.95	0.95	0.98
y_5						1.00	0.97	0.95	0.96	0.95	0.96	0.98	0.97	0.98	0.97	0.95	0.95	0.97	0.96	0.96	0.97
y_6							1.00	0.95	0.96	0.95	0.95	0.97	0.97	0.97	0.97	0.95	0.94	0.97	0.95	0.94	0.97
y_7								1.00	0.97	0.96	0.97	0.95	0.96	0.96	0.95	0.97	0.97	0.96	0.96	0.97	0.96
y_8									1.00	0.96	0.97	0.96	0.96	0.96	0.96	0.97	0.96	0.96	0.96	0.97	0.97
y_9										1.00	0.96	0.95	0.95	0.95	0.96	0.95	0.96	0.95	0.97	0.96	0.95
y_{10}											1.00	0.96	0.96	0.96	0.95	0.97	0.97	0.96	0.97	0.97	0.96
y_{11}												1.00	0.97	0.97	0.97	0.95	0.95	0.98	0.96	0.95	0.98
y_{12}													1.00	0.97	0.96	0.96	0.96	0.98	0.96	0.96	0.97
y_{13}														1.00	0.96	0.96	0.95	0.97	0.95	0.95	0.98
y_{14}															1.00	0.95	0.95	0.97	0.95	0.95	0.96
y_{15}																1.00	0.97	0.96	0.96	0.97	0.96
y_{16}																	1.00	0.96	0.96	0.96	0.96
y_{17}																		1.00	0.96	0.96	0.97
y_{18}																			1.00	0.96	0.96
y_{19}																				1.00	0.95
y_{20}																					1.00

Table 3.4. Correlation matrices for lead time 108 h, in warm and cool seasons.

Correlation Matrix　　　　　Lead Time 108 h　　　　　Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.55	0.40	0.41	0.29	0.41	0.44	0.42	0.42	0.50	0.48	0.40	0.38	0.39	0.43	0.32	0.30	0.31	0.57	0.52	0.28
y_1		1.00	0.44	0.44	0.30	0.44	0.42	0.39	0.49	0.48	0.44	0.41	0.40	0.46	0.34	0.31	0.37	0.33	0.48	0.48	0.38
y_2			1.00	0.56	0.57	0.68	0.55	0.63	0.61	0.54	0.65	0.56	0.60	0.66	0.49	0.59	0.50	0.53	0.54	0.73	0.53
y_3				1.00	0.51	0.64	0.50	0.57	0.58	0.47	0.56	0.61	0.58	0.59	0.48	0.45	0.51	0.53	0.51	0.55	0.59
y_4					1.00	0.55	0.49	0.47	0.46	0.40	0.37	0.48	0.58	0.50	0.36	0.43	0.38	0.51	0.40	0.47	0.44
y_5						1.00	0.58	0.62	0.58	0.47	0.65	0.58	0.62	0.71	0.45	0.53	0.49	0.58	0.49	0.71	0.54
y_6							1.00	0.49	0.45	0.51	0.51	0.49	0.55	0.55	0.45	0.40	0.41	0.48	0.48	0.59	0.43
y_7								1.00	0.55	0.46	0.61	0.54	0.55	0.62	0.40	0.55	0.54	0.53	0.49	0.62	0.51
y_8									1.00	0.52	0.62	0.51	0.51	0.57	0.38	0.50	0.58	0.46	0.43	0.60	0.49
y_9										1.00	0.59	0.50	0.52	0.53	0.46	0.42	0.40	0.43	0.50	0.58	0.38
y_{10}											1.00	0.52	0.56	0.60	0.43	0.62	0.44	0.49	0.56	0.69	0.52
y_{11}												1.00	0.59	0.63	0.50	0.42	0.44	0.56	0.44	0.58	0.60
y_{12}													1.00	0.62	0.40	0.48	0.42	0.59	0.51	0.62	0.51
y_{13}														1.00	0.45	0.51	0.53	0.63	0.48	0.64	0.58
y_{14}															1.00	0.35	0.35	0.35	0.43	0.48	0.33
y_{15}																1.00	0.40	0.47	0.48	0.56	0.39
y_{16}																	1.00	0.37	0.36	0.51	0.40
y_{17}																		1.00	0.42	0.52	0.49
y_{18}																			1.00	0.58	0.37
y_{19}																				1.00	0.47
y_{20}																					1.00

Correlation Matrix　　　　　Lead Time 108 h　　　　　Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.86	0.86	0.80	0.78	0.80	0.83	0.86	0.87	0.84	0.85	0.82	0.81	0.81	0.82	0.85	0.82	0.79	0.86	0.85	0.80
y_1		1.00	0.78	0.76	0.75	0.76	0.75	0.80	0.79	0.77	0.83	0.76	0.74	0.78	0.76	0.80	0.77	0.76	0.81	0.79	0.75
y_2			1.00	0.75	0.73	0.75	0.76	0.79	0.81	0.79	0.78	0.77	0.75	0.78	0.78	0.81	0.78	0.73	0.79	0.81	0.74
y_3				1.00	0.82	0.82	0.80	0.79	0.75	0.75	0.73	0.82	0.82	0.86	0.81	0.74	0.72	0.80	0.75	0.73	0.84
y_4					1.00	0.81	0.79	0.73	0.74	0.71	0.75	0.80	0.84	0.81	0.79	0.76	0.68	0.81	0.74	0.78	0.84
y_5						1.00	0.79	0.76	0.74	0.75	0.76	0.82	0.81	0.82	0.83	0.75	0.72	0.80	0.76	0.74	0.83
y_6							1.00	0.74	0.78	0.77	0.72	0.83	0.79	0.81	0.83	0.74	0.72	0.78	0.76	0.74	0.79
y_7								1.00	0.81	0.78	0.79	0.76	0.77	0.76	0.77	0.81	0.78	0.76	0.80	0.78	0.76
y_8									1.00	0.78	0.80	0.78	0.75	0.77	0.78	0.79	0.81	0.77	0.81	0.79	0.79
y_9										1.00	0.76	0.78	0.75	0.77	0.77	0.81	0.76	0.75	0.78	0.78	0.74
y_{10}											1.00	0.71	0.72	0.73	0.73	0.79	0.76	0.74	0.81	0.79	0.75
y_{11}												1.00	0.83	0.83	0.84	0.75	0.71	0.83	0.76	0.75	0.83
y_{12}													1.00	0.83	0.81	0.76	0.71	0.80	0.74	0.75	0.83
y_{13}														1.00	0.81	0.76	0.70	0.80	0.75	0.75	0.84
y_{14}															1.00	0.75	0.73	0.82	0.77	0.75	0.82
y_{15}																1.00	0.75	0.76	0.78	0.81	0.78
y_{16}																	1.00	0.70	0.78	0.75	0.72
y_{17}																		1.00	0.73	0.75	0.81
y_{18}																			1.00	0.77	0.75
y_{19}																				1.00	0.76
y_{20}																					1.00

Table 3.5. Correlation matrices for lead time 204 h, in warm and cool seasons.

Correlation Matrix Lead Time 204 h Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.23	0.26	0.20	0.38	0.37	0.21	0.31	0.31	0.26	0.28	0.28	0.29	0.31	0.35	0.29	0.20	0.36	0.24	0.29	0.27
y_1		1.00	0.37	0.22	0.23	0.34	0.28	0.37	0.31	0.29	0.41	0.25	0.30	0.34	0.25	0.30	0.29	0.29	0.40	0.34	0.23
y_2			1.00	0.37	0.42	0.49	0.38	0.44	0.40	0.33	0.55	0.42	0.45	0.49	0.26	0.41	0.38	0.43	0.35	0.44	0.42
y_3				1.00	0.43	0.57	0.40	0.45	0.41	0.31	0.42	0.49	0.51	0.46	0.34	0.39	0.36	0.50	0.33	0.32	0.47
y_4					1.00	0.50	0.38	0.47	0.39	0.30	0.36	0.57	0.61	0.53	0.38	0.43	0.28	0.60	0.37	0.37	0.57
y_5						1.00	0.37	0.48	0.38	0.38	0.47	0.50	0.59	0.63	0.34	0.41	0.41	0.56	0.44	0.35	0.54
y_6							1.00	0.44	0.41	0.31	0.38	0.39	0.40	0.43	0.35	0.33	0.30	0.44	0.33	0.28	0.29
y_7								1.00	0.45	0.34	0.50	0.51	0.48	0.54	0.35	0.52	0.39	0.56	0.38	0.43	0.46
y_8									1.00	0.31	0.42	0.47	0.37	0.40	0.32	0.45	0.40	0.44	0.32	0.40	0.32
y_9										1.00	0.33	0.32	0.29	0.40	0.22	0.40	0.32	0.37	0.39	0.36	0.27
y_{10}											1.00	0.39	0.42	0.52	0.23	0.51	0.35	0.46	0.28	0.41	0.39
y_{11}												1.00	0.50	0.46	0.40	0.42	0.34	0.57	0.32	0.31	0.52
y_{12}													1.00	0.54	0.36	0.44	0.38	0.60	0.39	0.34	0.53
y_{13}														1.00	0.33	0.40	0.39	0.60	0.46	0.48	0.50
y_{14}															1.00	0.32	0.25	0.36	0.26	0.20	0.34
y_{15}																1.00	0.26	0.45	0.34	0.33	0.39
y_{16}																	1.00	0.38	0.27	0.40	0.33
y_{17}																		1.00	0.41	0.33	0.53
y_{18}																			1.00	0.38	0.29
y_{19}																				1.00	0.28
y_{20}																					1.00

Correlation Matrix Lead Time 204 h Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.47	0.47	0.27	0.40	0.36	0.42	0.44	0.50	0.39	0.44	0.35	0.36	0.35	0.42	0.44	0.50	0.37	0.40	0.40	0.43
y_1		1.00	0.39	0.38	0.47	0.40	0.39	0.42	0.35	0.38	0.46	0.28	0.31	0.38	0.37	0.39	0.35	0.35	0.38	0.37	0.43
y_2			1.00	0.26	0.41	0.32	0.36	0.34	0.35	0.31	0.33	0.24	0.31	0.37	0.35	0.35	0.42	0.38	0.34	0.31	0.33
y_3				1.00	0.40	0.47	0.41	0.33	0.21	0.30	0.27	0.42	0.35	0.36	0.40	0.33	0.27	0.37	0.38	0.30	0.37
y_4					1.00	0.43	0.43	0.42	0.30	0.34	0.35	0.41	0.49	0.42	0.51	0.38	0.34	0.45	0.40	0.38	0.45
y_5						1.00	0.42	0.36	0.27	0.33	0.37	0.35	0.41	0.39	0.42	0.39	0.31	0.40	0.41	0.34	0.43
y_6							1.00	0.34	0.29	0.33	0.42	0.37	0.37	0.34	0.49	0.36	0.33	0.32	0.45	0.33	0.46
y_7								1.00	0.37	0.38	0.33	0.33	0.38	0.35	0.43	0.47	0.37	0.38	0.42	0.37	0.43
y_8									1.00	0.27	0.29	0.25	0.30	0.26	0.31	0.37	0.40	0.29	0.25	0.27	0.32
y_9										1.00	0.30	0.34	0.27	0.27	0.37	0.31	0.34	0.30	0.42	0.33	0.30
y_{10}											1.00	0.30	0.30	0.34	0.39	0.32	0.39	0.30	0.40	0.36	0.48
y_{11}												1.00	0.34	0.32	0.42	0.34	0.34	0.30	0.34	0.36	0.49
y_{12}													1.00	0.38	0.53	0.35	0.32	0.39	0.35	0.26	0.47
y_{13}														1.00	0.45	0.28	0.35	0.35	0.28	0.29	0.46
y_{14}															1.00	0.38	0.40	0.34	0.46	0.36	0.54
y_{15}																1.00	0.36	0.36	0.36	0.37	0.37
y_{16}																	1.00	0.44	0.36	0.35	0.39
y_{17}																		1.00	0.36	0.30	0.39
y_{18}																			1.00	0.36	0.40
y_{19}																				1.00	0.36
y_{20}																					1.00

Table 3.6. Correlation matrices for lead time 300 h, in warm and cool seasons.

Correlation Matrix   Lead Time 300 h   Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.29	0.34	0.25	0.30	0.24	0.24	0.32	0.28	0.21	0.18	0.22	0.24	0.31	0.24	0.32	0.26	0.25	0.32	0.21	0.30
y_1		1.00	0.27	0.21	0.23	0.34	0.24	0.34	0.32	0.22	0.26	0.25	0.42	0.37	0.25	0.35	0.28	0.33	0.34	0.36	0.36
y_2			1.00	0.34	0.31	0.39	0.21	0.40	0.30	0.33	0.35	0.27	0.44	0.43	0.25	0.33	0.37	0.35	0.38	0.37	0.43
y_3				1.00	0.49	0.45	0.40	0.48	0.23	0.34	0.25	0.43	0.53	0.41	0.39	0.35	0.32	0.46	0.40	0.22	0.54
y_4					1.00	0.48	0.41	0.43	0.32	0.40	0.25	0.42	0.56	0.53	0.40	0.35	0.27	0.57	0.43	0.35	0.45
y_5						1.00	0.32	0.40	0.38	0.32	0.33	0.48	0.56	0.58	0.44	0.39	0.34	0.51	0.40	0.39	0.49
y_6							1.00	0.33	0.22	0.27	0.31	0.42	0.46	0.42	0.38	0.32	0.27	0.40	0.25	0.23	0.36
y_7								1.00	0.47	0.37	0.36	0.43	0.56	0.47	0.32	0.44	0.33	0.46	0.47	0.41	0.34
y_8									1.00	0.30	0.36	0.41	0.40	0.46	0.25	0.38	0.31	0.38	0.40	0.31	0.30
y_9										1.00	0.27	0.35	0.40	0.45	0.26	0.35	0.29	0.36	0.39	0.30	0.36
y_{10}											1.00	0.28	0.36	0.30	0.21	0.39	0.30	0.32	0.25	0.41	0.32
y_{11}												1.00	0.49	0.53	0.32	0.36	0.35	0.47	0.36	0.31	0.46
y_{12}													1.00	0.57	0.48	0.46	0.37	0.65	0.45	0.42	0.50
y_{13}														1.00	0.44	0.39	0.33	0.59	0.52	0.38	0.51
y_{14}															1.00	0.29	0.25	0.37	0.34	0.20	0.46
y_{15}																1.00	0.33	0.39	0.33	0.47	0.40
y_{16}																	1.00	0.33	0.36	0.28	0.38
y_{17}																		1.00	0.42	0.37	0.45
y_{18}																			1.00	0.36	0.43
y_{19}																				1.00	0.34
y_{20}																					1.00

Correlation Matrix   Lead Time 300 h   Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.24	0.12	0.12	0.20	0.14	0.13	0.13	0.29	0.19	0.15	0.09	0.21	0.18	0.22	0.22	0.21	0.15	0.13	0.15	0.20
y_1		1.00	0.14	0.13	0.18	0.17	0.12	0.02	0.19	0.18	0.08	0.19	0.13	0.14	0.18	0.11	0.15	0.14	0.15	0.16	0.24
y_2			1.00	0.15	0.12	0.17	0.21	0.14	0.16	0.23	0.06	0.15	0.15	0.14	0.11	0.15	0.18	0.23	0.12	0.18	0.15
y_3				1.00	0.22	0.23	0.31	0.21	0.17	0.18	0.18	0.27	0.35	0.21	0.26	0.06	0.23	0.31	0.12	0.11	0.29
y_4					1.00	0.34	0.22	0.09	0.18	0.18	0.17	0.17	0.30	0.26	0.26	0.07	0.18	0.26	0.02	0.15	0.33
y_5						1.00	0.29	0.11	0.13	0.20	0.19	0.23	0.28	0.19	0.30	0.13	0.20	0.29	0.08	0.21	0.33
y_6							1.00	0.17	0.16	0.21	0.12	0.22	0.27	0.29	0.34	0.10	0.19	0.30	0.20	0.15	0.23
y_7								1.00	0.12	0.12	0.13	0.05	0.15	0.15	0.10	0.11	0.17	0.17	0.17	0.08	0.09
y_8									1.00	0.15	0.21	0.10	0.16	0.19	0.24	0.09	0.21	0.24	0.13	0.21	0.18
y_9										1.00	0.11	0.22	0.14	0.22	0.25	0.19	0.18	0.20	0.22	0.20	0.18
y_{10}											1.00	0.11	0.13	0.21	0.20	0.14	0.12	0.19	0.08	0.11	0.11
y_{11}												1.00	0.24	0.28	0.28	0.21	0.12	0.30	0.10	0.10	0.22
y_{12}													1.00	0.27	0.37	0.17	0.15	0.36	0.15	0.12	0.36
y_{13}														1.00	0.29	0.14	0.11	0.34	0.19	0.17	0.19
y_{14}															1.00	0.14	0.16	0.40	0.13	0.16	0.34
y_{15}																1.00	0.15	0.08	0.08	0.12	0.03
y_{16}																	1.00	0.18	0.21	0.16	0.25
y_{17}																		1.00	0.13	0.17	0.38
y_{18}																			1.00	0.05	0.12
y_{19}																				1.00	0.16
y_{20}																					1.00

Table 3.7. Conditional correlation matrices for lead time 12 h, in warm and cool seasons.

Conditional Correlation Matrix Lead Time 12 h Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.61	0.33	0.29	0.16	0.26	0.29	0.26	0.33	0.53	0.43	0.23	0.20	0.31	0.58	0.27	0.37	0.23	0.56	0.34	0.22
y_1		1.00	0.26	0.28	0.16	0.14	0.19	0.21	0.33	0.33	0.40	0.21	0.13	0.16	0.53	0.27	0.41	0.19	0.42	0.30	0.14
y_2			1.00	0.47	0.53	0.46	0.17	0.52	0.45	0.27	0.62	0.44	0.51	0.59	0.28	0.46	0.40	0.54	0.34	0.61	0.53
y_3				1.00	0.48	0.52	0.18	0.43	0.52	0.29	0.44	0.60	0.54	0.53	0.36	0.47	0.51	0.55	0.26	0.49	0.63
y_4					1.00	0.58	0.19	0.62	0.45	0.14	0.57	0.45	0.67	0.65	0.20	0.55	0.33	0.69	0.27	0.59	0.55
y_5						1.00	0.22	0.57	0.46	0.19	0.60	0.48	0.59	0.61	0.19	0.51	0.37	0.63	0.34	0.66	0.53
y_6							1.00	0.10	0.20	0.29	0.16	0.23	0.12	0.25	0.34	-0.03	0.08	0.10	0.36	0.16	0.23
y_7								1.00	0.45	0.21	0.61	0.40	0.59	0.56	0.19	0.64	0.39	0.61	0.36	0.65	0.50
y_8									1.00	0.32	0.51	0.63	0.38	0.43	0.33	0.46	0.56	0.49	0.40	0.49	0.58
y_9										1.00	0.27	0.31	0.20	0.27	0.36	0.25	0.25	0.25	0.51	0.27	0.29
y_{10}											1.00	0.42	0.53	0.60	0.25	0.58	0.47	0.60	0.38	0.72	0.43
y_{11}												1.00	0.41	0.50	0.32	0.36	0.49	0.51	0.35	0.49	0.66
y_{12}													1.00	0.67	0.18	0.54	0.34	0.64	0.26	0.57	0.56
y_{13}														1.00	0.22	0.50	0.34	0.65	0.35	0.63	0.61
y_{14}															1.00	0.18	0.36	0.28	0.39	0.21	0.31
y_{15}																1.00	0.46	0.54	0.30	0.59	0.40
y_{16}																	1.00	0.35	0.35	0.42	0.48
y_{17}																		1.00	0.30	0.64	0.54
y_{18}																			1.00	0.31	0.42
y_{19}																				1.00	0.48
y_{20}																					1.00

Conditional Correlation Matrix Lead Time 12 h Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.74	0.62	0.49	0.43	0.46	0.54	0.70	0.68	0.74	0.67	0.48	0.52	0.47	0.61	0.62	0.69	0.54	0.69	0.66	0.50
y_1		1.00	0.57	0.42	0.40	0.44	0.42	0.55	0.56	0.58	0.59	0.42	0.43	0.42	0.49	0.51	0.57	0.49	0.62	0.51	0.44
y_2			1.00	0.53	0.56	0.50	0.44	0.65	0.58	0.50	0.61	0.52	0.54	0.55	0.49	0.58	0.57	0.59	0.54	0.58	0.50
y_3				1.00	0.67	0.66	0.62	0.52	0.52	0.42	0.43	0.69	0.70	0.73	0.62	0.44	0.54	0.71	0.55	0.49	0.69
y_4					1.00	0.72	0.68	0.54	0.57	0.41	0.51	0.73	0.73	0.77	0.60	0.48	0.47	0.74	0.44	0.47	0.76
y_5						1.00	0.67	0.49	0.51	0.42	0.49	0.74	0.69	0.77	0.65	0.43	0.49	0.69	0.48	0.48	0.70
y_6							1.00	0.43	0.54	0.44	0.41	0.73	0.68	0.71	0.70	0.39	0.45	0.68	0.49	0.41	0.70
y_7								1.00	0.64	0.55	0.64	0.48	0.61	0.52	0.48	0.61	0.67	0.58	0.53	0.60	0.53
y_8									1.00	0.52	0.66	0.59	0.57	0.59	0.55	0.56	0.60	0.56	0.55	0.59	0.62
y_9										1.00	0.51	0.41	0.47	0.46	0.55	0.41	0.54	0.49	0.59	0.50	0.42
y_{10}											1.00	0.48	0.50	0.49	0.50	0.55	0.62	0.55	0.55	0.61	0.49
y_{11}												1.00	0.74	0.73	0.66	0.42	0.50	0.74	0.51	0.46	0.78
y_{12}													1.00	0.72	0.61	0.52	0.57	0.77	0.51	0.53	0.72
y_{13}														1.00	0.64	0.48	0.50	0.72	0.45	0.43	0.75
y_{14}															1.00	0.42	0.48	0.67	0.50	0.48	0.63
y_{15}																1.00	0.61	0.53	0.51	0.57	0.49
y_{16}																	1.00	0.55	0.56	0.61	0.53
y_{17}																		1.00	0.53	0.58	0.73
y_{18}																			1.00	0.56	0.49
y_{19}																				1.00	0.46
y_{20}																					1.00

Table 3.8. Conditional correlation matrices for lead time 108 h, in warm and cool seasons.

Conditional Correlation Matrix   Lead Time 108 h   Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.40	0.16	0.20	0.12	0.17	0.25	0.23	0.20	0.32	0.26	0.18	0.19	0.12	0.26	0.15	0.12	0.12	0.43	0.31	0.04
y_1		1.00	0.20	0.24	0.14	0.21	0.21	0.19	0.30	0.29	0.21	0.19	0.21	0.22	0.15	0.14	0.21	0.14	0.31	0.25	0.17
y_2			1.00	0.38	0.46	0.51	0.36	0.48	0.44	0.33	0.46	0.35	0.46	0.47	0.32	0.47	0.35	0.37	0.35	0.58	0.35
y_3				1.00	0.40	0.48	0.31	0.43	0.42	0.27	0.37	0.46	0.45	0.41	0.33	0.30	0.38	0.40	0.34	0.35	0.44
y_4					1.00	0.44	0.37	0.35	0.33	0.26	0.20	0.36	0.49	0.36	0.23	0.32	0.27	0.41	0.26	0.33	0.31
y_5						1.00	0.41	0.47	0.39	0.24	0.46	0.40	0.49	0.55	0.27	0.40	0.34	0.44	0.28	0.56	0.36
y_6							1.00	0.31	0.24	0.33	0.30	0.29	0.40	0.34	0.28	0.24	0.25	0.33	0.30	0.41	0.24
y_7								1.00	0.39	0.26	0.46	0.38	0.41	0.46	0.23	0.43	0.43	0.40	0.32	0.47	0.35
y_8									1.00	0.33	0.45	0.32	0.35	0.38	0.19	0.37	0.47	0.29	0.22	0.42	0.31
y_9										1.00	0.39	0.31	0.36	0.31	0.29	0.26	0.23	0.26	0.32	0.38	0.16
y_{10}											1.00	0.30	0.39	0.37	0.24	0.51	0.27	0.31	0.38	0.52	0.32
y_{11}												1.00	0.45	0.45	0.35	0.26	0.29	0.43	0.24	0.39	0.46
y_{12}													1.00	0.48	0.24	0.36	0.29	0.48	0.36	0.48	0.36
y_{13}														1.00	0.25	0.36	0.38	0.50	0.26	0.43	0.41
y_{14}															1.00	0.20	0.21	0.19	0.26	0.30	0.15
y_{15}																1.00	0.28	0.35	0.34	0.43	0.23
y_{16}																	1.00	0.23	0.20	0.37	0.24
y_{17}																		1.00	0.25	0.35	0.34
y_{18}																			1.00	0.40	0.16
y_{19}																				1.00	0.25
y_{20}																					1.00

Conditional Correlation Matrix   Lead Time 108 h   Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.61	0.61	0.55	0.49	0.57	0.60	0.65	0.67	0.60	0.59	0.59	0.58	0.57	0.58	0.60	0.56	0.54	0.66	0.61	0.57
y_1		1.00	0.43	0.48	0.44	0.48	0.42	0.51	0.51	0.45	0.58	0.46	0.44	0.50	0.45	0.47	0.45	0.47	0.54	0.48	0.45
y_2			1.00	0.46	0.40	0.47	0.46	0.50	0.55	0.51	0.45	0.50	0.45	0.52	0.50	0.51	0.50	0.41	0.50	0.54	0.45
y_3				1.00	0.64	0.64	0.60	0.56	0.48	0.48	0.43	0.64	0.66	0.73	0.61	0.43	0.43	0.61	0.48	0.43	0.68
y_4					1.00	0.63	0.57	0.42	0.45	0.38	0.46	0.60	0.68	0.61	0.57	0.47	0.32	0.62	0.44	0.52	0.68
y_5						1.00	0.58	0.52	0.47	0.49	0.51	0.64	0.63	0.65	0.66	0.47	0.45	0.61	0.51	0.45	0.67
y_6							1.00	0.44	0.54	0.51	0.37	0.66	0.57	0.62	0.64	0.41	0.41	0.56	0.48	0.42	0.57
y_7								1.00	0.58	0.50	0.52	0.49	0.53	0.50	0.51	0.54	0.53	0.51	0.56	0.50	0.50
y_8									1.00	0.51	0.53	0.53	0.49	0.51	0.52	0.51	0.57	0.53	0.58	0.52	0.57
y_9										1.00	0.43	0.52	0.48	0.51	0.50	0.55	0.46	0.48	0.51	0.49	0.46
y_{10}											1.00	0.37	0.40	0.41	0.40	0.49	0.45	0.46	0.56	0.49	0.47
y_{11}												1.00	0.66	0.66	0.68	0.44	0.39	0.66	0.50	0.44	0.67
y_{12}													1.00	0.65	0.62	0.48	0.41	0.62	0.46	0.46	0.67
y_{13}														1.00	0.60	0.47	0.39	0.60	0.47	0.45	0.69
y_{14}															1.00	0.42	0.44	0.64	0.51	0.46	0.64
y_{15}																1.00	0.43	0.48	0.48	0.54	0.52
y_{16}																	1.00	0.40	0.52	0.44	0.44
y_{17}																		1.00	0.46	0.48	0.64
y_{18}																			1.00	0.47	0.49
y_{19}																				1.00	0.49
y_{20}																					1.00

Table 3.9. Conditional correlation matrices for lead time 204 h, in warm and cool seasons.

Conditional Correlation Matrix   Lead Time 204 h   Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.18	0.21	0.13	0.32	0.31	0.16	0.24	0.24	0.19	0.20	0.22	0.21	0.24	0.30	0.23	0.15	0.29	0.19	0.24	0.22
y_1		1.00	0.34	0.18	0.19	0.30	0.26	0.33	0.27	0.25	0.38	0.21	0.25	0.30	0.21	0.26	0.26	0.25	0.38	0.31	0.19
y_2			1.00	0.33	0.38	0.45	0.35	0.40	0.35	0.28	0.51	0.38	0.40	0.45	0.20	0.37	0.35	0.38	0.31	0.41	0.38
y_3				1.00	0.38	0.53	0.37	0.40	0.35	0.25	0.36	0.45	0.46	0.41	0.29	0.34	0.32	0.45	0.29	0.27	0.44
y_4					1.00	0.45	0.34	0.41	0.32	0.23	0.28	0.53	0.56	0.48	0.33	0.38	0.23	0.55	0.33	0.32	0.53
y_5						1.00	0.34	0.43	0.31	0.32	0.41	0.45	0.54	0.59	0.29	0.35	0.37	0.51	0.40	0.31	0.50
y_6							1.00	0.40	0.37	0.27	0.34	0.36	0.36	0.39	0.31	0.30	0.27	0.41	0.30	0.25	0.25
y_7								1.00	0.39	0.27	0.44	0.46	0.41	0.48	0.28	0.47	0.34	0.50	0.34	0.38	0.41
y_8									1.00	0.24	0.34	0.42	0.28	0.33	0.25	0.39	0.36	0.36	0.27	0.35	0.26
y_9										1.00	0.25	0.27	0.21	0.33	0.15	0.34	0.27	0.30	0.35	0.32	0.22
y_{10}											1.00	0.32	0.33	0.45	0.15	0.46	0.30	0.38	0.22	0.35	0.34
y_{11}												1.00	0.44	0.41	0.35	0.37	0.30	0.52	0.28	0.27	0.48
y_{12}													1.00	0.47	0.29	0.37	0.33	0.54	0.34	0.28	0.49
y_{13}														1.00	0.26	0.33	0.34	0.54	0.41	0.44	0.45
y_{14}															1.00	0.26	0.20	0.30	0.21	0.14	0.29
y_{15}																1.00	0.21	0.39	0.29	0.28	0.35
y_{16}																	1.00	0.34	0.23	0.37	0.30
y_{17}																		1.00	0.36	0.27	0.49
y_{18}																			1.00	0.34	0.25
y_{19}																				1.00	0.24
y_{20}																					1.00

Conditional Correlation Matrix   Lead Time 204 h   Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.36	0.38	0.17	0.31	0.27	0.31	0.35	0.42	0.32	0.33	0.24	0.28	0.25	0.33	0.33	0.42	0.29	0.29	0.29	0.31
y_1		1.00	0.30	0.31	0.40	0.33	0.29	0.33	0.26	0.31	0.37	0.17	0.23	0.30	0.29	0.28	0.25	0.27	0.28	0.27	0.33
y_2			1.00	0.18	0.34	0.25	0.27	0.25	0.27	0.24	0.23	0.15	0.24	0.30	0.27	0.25	0.34	0.32	0.25	0.22	0.22
y_3				1.00	0.34	0.42	0.35	0.27	0.13	0.24	0.18	0.37	0.30	0.30	0.35	0.25	0.20	0.32	0.31	0.22	0.30
y_4					1.00	0.38	0.36	0.35	0.21	0.28	0.26	0.34	0.45	0.35	0.46	0.30	0.26	0.40	0.32	0.30	0.38
y_5						1.00	0.35	0.29	0.19	0.28	0.29	0.29	0.36	0.33	0.37	0.32	0.24	0.34	0.34	0.27	0.36
y_6							1.00	0.26	0.20	0.26	0.33	0.29	0.31	0.25	0.42	0.26	0.24	0.24	0.37	0.23	0.38
y_7								1.00	0.29	0.32	0.23	0.25	0.33	0.28	0.37	0.40	0.30	0.32	0.35	0.29	0.34
y_8									1.00	0.20	0.18	0.16	0.24	0.17	0.23	0.28	0.32	0.22	0.15	0.18	0.22
y_9										1.00	0.22	0.28	0.22	0.20	0.31	0.23	0.27	0.24	0.36	0.26	0.22
y_{10}											1.00	0.20	0.22	0.25	0.32	0.20	0.30	0.22	0.30	0.26	0.38
y_{11}												1.00	0.29	0.24	0.36	0.25	0.26	0.23	0.26	0.28	0.42
y_{12}													1.00	0.32	0.49	0.28	0.26	0.34	0.28	0.18	0.41
y_{13}														1.00	0.39	0.18	0.28	0.28	0.19	0.20	0.39
y_{14}															1.00	0.30	0.33	0.28	0.39	0.29	0.48
y_{15}																1.00	0.27	0.28	0.26	0.27	0.26
y_{16}																	1.00	0.38	0.27	0.26	0.30
y_{17}																		1.00	0.29	0.23	0.32
y_{18}																			1.00	0.26	0.30
y_{19}																				1.00	0.26
y_{20}																					1.00

27

Table 3.10. Conditional correlation matrices for lead time 300 h, in warm and cool seasons.

Conditional Correlation Matrix   Lead Time 300 h   Warm Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.27	0.33	0.23	0.28	0.21	0.22	0.30	0.26	0.18	0.16	0.20	0.21	0.29	0.22	0.30	0.24	0.23	0.29	0.18	0.28
y_1		1.00	0.23	0.16	0.19	0.29	0.20	0.31	0.28	0.16	0.22	0.21	0.38	0.32	0.21	0.29	0.23	0.29	0.29	0.31	0.32
y_2			1.00	0.32	0.29	0.36	0.19	0.38	0.28	0.30	0.34	0.25	0.42	0.40	0.22	0.31	0.35	0.33	0.36	0.35	0.40
y_3				1.00	0.47	0.42	0.38	0.46	0.21	0.30	0.23	0.41	0.50	0.38	0.37	0.32	0.29	0.44	0.37	0.18	0.51
y_4					1.00	0.46	0.40	0.41	0.29	0.38	0.23	0.40	0.54	0.51	0.38	0.32	0.24	0.56	0.40	0.32	0.43
y_5						1.00	0.30	0.38	0.35	0.28	0.31	0.46	0.53	0.56	0.41	0.35	0.31	0.48	0.37	0.35	0.46
y_6							1.00	0.32	0.20	0.24	0.29	0.40	0.44	0.40	0.36	0.29	0.25	0.38	0.23	0.20	0.34
y_7								1.00	0.45	0.34	0.35	0.41	0.54	0.45	0.30	0.42	0.31	0.44	0.45	0.38	0.31
y_8									1.00	0.26	0.34	0.39	0.37	0.43	0.23	0.35	0.28	0.36	0.37	0.28	0.27
y_9										1.00	0.24	0.32	0.36	0.41	0.23	0.30	0.25	0.33	0.35	0.25	0.32
y_{10}											1.00	0.26	0.34	0.28	0.19	0.37	0.28	0.30	0.22	0.39	0.30
y_{11}												1.00	0.47	0.51	0.30	0.33	0.33	0.45	0.33	0.28	0.44
y_{12}													1.00	0.54	0.45	0.43	0.34	0.63	0.41	0.38	0.47
y_{13}														1.00	0.41	0.35	0.29	0.57	0.49	0.34	0.48
y_{14}															1.00	0.26	0.22	0.35	0.31	0.16	0.44
y_{15}																1.00	0.29	0.36	0.29	0.43	0.36
y_{16}																	1.00	0.30	0.33	0.23	0.34
y_{17}																		1.00	0.39	0.33	0.42
y_{18}																			1.00	0.31	0.40
y_{19}																				1.00	0.30
y_{20}																					1.00

Conditional Correlation Matrix   Lead Time 300 h   Cool Season

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_0	1.00	0.21	0.08	0.08	0.18	0.12	0.10	0.12	0.28	0.16	0.12	0.07	0.19	0.15	0.19	0.20	0.19	0.13	0.12	0.13	0.18
y_1		1.00	0.10	0.09	0.16	0.15	0.09	0.00	0.17	0.15	0.05	0.17	0.11	0.11	0.15	0.08	0.13	0.11	0.14	0.14	0.23
y_2			1.00	0.12	0.11	0.15	0.19	0.13	0.14	0.21	0.04	0.13	0.13	0.12	0.08	0.12	0.16	0.21	0.11	0.16	0.13
y_3				1.00	0.21	0.22	0.29	0.20	0.15	0.16	0.15	0.25	0.33	0.19	0.24	0.03	0.22	0.29	0.11	0.10	0.27
y_4					1.00	0.33	0.20	0.08	0.17	0.17	0.16	0.16	0.29	0.25	0.24	0.05	0.17	0.25	0.01	0.14	0.32
y_5						1.00	0.28	0.10	0.12	0.18	0.17	0.22	0.27	0.18	0.29	0.12	0.19	0.27	0.07	0.20	0.32
y_6							1.00	0.17	0.15	0.20	0.10	0.20	0.26	0.27	0.32	0.08	0.18	0.29	0.19	0.14	0.22
y_7								1.00	0.11	0.12	0.12	0.04	0.15	0.14	0.10	0.10	0.16	0.17	0.17	0.07	0.08
y_8									1.00	0.13	0.19	0.09	0.14	0.17	0.22	0.07	0.20	0.23	0.13	0.20	0.17
y_9										1.00	0.09	0.20	0.13	0.21	0.24	0.17	0.17	0.18	0.21	0.18	0.17
y_{10}											1.00	0.09	0.11	0.19	0.19	0.12	0.11	0.17	0.07	0.10	0.10
y_{11}												1.00	0.22	0.27	0.26	0.20	0.11	0.28	0.09	0.09	0.21
y_{12}													1.00	0.26	0.36	0.16	0.14	0.35	0.14	0.11	0.35
y_{13}														1.00	0.27	0.12	0.10	0.32	0.18	0.16	0.18
y_{14}															1.00	0.12	0.15	0.38	0.12	0.15	0.33
y_{15}																1.00	0.13	0.06	0.07	0.11	0.02
y_{16}																	1.00	0.17	0.21	0.16	0.24
y_{17}																		1.00	0.12	0.16	0.37
y_{18}																			1.00	0.04	0.12
y_{19}																				1.00	0.15
y_{20}																					1.00

4. BAYESIAN PROCESSOR OF ENSEMBLE

4.1 Theory

In the Bayesian processor of ensemble (BPE) introduced in Section 2.2, the family of likelihood functions f of predictand W, conditional on the realization of ensemble $\mathbf{y}$, can be factorized as:

$$f(\mathbf{y}|w) = f_{20}(y_{20}|y_{19}, \dots, y_1, w) f_{19}(y_{19}|y_{18}, \dots, y_1, w) \cdots f_2(y_2|y_1, w) f_1(y_1|w). \quad (4.1)$$

Since the ensemble members do not constitute a random sample and are not stochastically independent conditional on $W = w$ (see Section 3.3), the likelihood function cannot be further simplified. The challenge in processing the ensemble empirically is that the conditional density functions in (4.1) are unknown. Lee (2010) suggests searching for an ensemble summary statistic (X, T) that is as informative as the vector $\mathbf{Y}$ of all 20 ensemble members, where X is a predictor of central tendency of W, and T is a predictor of uncertainty about W. Using realization (x, t) of (X, T) in place of realization $\mathbf{y}$ simplifies the BPE to the following form (Krzysztofowicz 1983, 2010):

$$\phi(w|x, t) = \frac{f(x, t|w)\, g(w)}{\kappa(x, t)} = \frac{f_1(x|t, w) f_2(t|w)\, g(w)}{\kappa_1(x|t)\kappa_2(t)}, \quad (4.2)$$

$$\kappa(x, t) = \kappa_1(x|t)\kappa_2(t) = \int f(x, t|w)g(w)dw = \int f_1(x|t, w)f_2(t|w)g(w)dw. \quad (4.3)$$

The Gaussian-Gamma model (validated in Chapter 7) has the structural assumption that T and W are stochastically independent (Section 7.4.1), which implies $f_2(t|w) = \kappa_2(t)$. Under this assumption, the BPE simplifies to (Krzysztofowicz 2010):

$$\phi(w|x,t) = \frac{f(x,t|w)\,g(w)}{\kappa(x,t)} = \frac{f_1(x|t,w)g(w)}{\kappa_1(x|t)}, \tag{4.4}$$

$$\kappa_1(x|t) = \int f_1(x|t,w)g(w)dw. \tag{4.5}$$

4.2 Gaussian-Gamma Model

In this context, X is an aggregate predictor calculated via the normal-linear prediction model (Section 5.2.3). The components of the Gaussian-Gamma BPE (4.4) – (4.5) are modeled as follows (Krzysztofowicz 2010). The prior density function g is normal with mean M and variance S^2:

$$g(w) \approx N(M, S^2). \tag{4.6}$$

The reciprocal of T is assumed to follow a gamma distribution with scale parameter α and shape parameter β:

$$\frac{1}{T} \sim Gamma(\alpha, \beta). \tag{4.7}$$

Conditional density function f_1 is normal with mean and variance parameters equal to, respectively, the mean and variance of X conditional on $T = t$ and $W = w$:

$$f_1(x|t,w) \approx N\big(E(X|t,w), Var(X|t,w)\big). \tag{4.8}$$

Conditional mean $E(X|t,w)$ and conditional variance $Var(X|t,w)$ are calculated from elements of the normal-linear likelihood model (Section 5.2.1) and parameters α and β of the gamma distribution:

$$E(X|t,w) = aw + b, \tag{4.9}$$

$$Var(X|t,w) = v^2 t, \tag{4.10}$$

$$v^2 = \alpha(\beta - 1)\sigma^2. \tag{4.11}$$

From the above three equations, conditional density function f_1 is:

$$f_1(x|t, w) \approx N(aw + b, v^2 t). \tag{4.12}$$

Expected density function κ_1 is also normal, and is used to validate the Gaussian-Gamma model in Section (7.4):

$$\kappa_1(x|t) \approx N(aM + b, \, a^2 S^2 + v^2 t). \tag{4.13}$$

Lastly, posterior density function ϕ, an output of the BPE, is specified as follows:

$$\phi(w|x, t) \approx N(E(W|x, t), \, Var(W|x, t)), \tag{4.14}$$

$$E(W|x, t) = \frac{aS^2}{aS^2 + v^2 t} x + \frac{Mv^2 t - abS^2}{aS^2 + v^2 t}, \tag{4.15}$$

$$Var(W|x, t) = \frac{S^2 v^2 t}{aS^2 + v^2 t}. \tag{4.16}$$

From (4.15), the posterior mean of W is a linear function of x and a nonlinear function of t. From (4.16), the posterior variance of W is a nonlinear function of t. While the posterior variance does not depend on x directly, it does depend on signal-to-noise ratio $|a|/\sigma$ of X, since v^2 is a function of residual variance σ^2. Because realizations x and t vary with each forecast, the BPE implies that the posterior distribution of W varies with each forecast as well, according to changes in both predictors, of central tendency and of uncertainty.

4.3 Sufficient Statistics

4.3.1 Definition

As efficiency is a primary concern in the processing of any global ensemble forecast, it is advantageous to find a summary statistic that contains as much information about the predictand as a vector of all ensemble members. Such a summary statistic is called a sufficient statistic, and is defined formally as follows (DeGroot 1970):

Let ξ be any function of $\mathbf{y}$. Then such a function ξ is a sufficient statistic for the family $\{f(\cdot\,|w): all\ w\}$ of conditional density functions of $\mathbf{Y}$ if two posterior density functions are identical, $\phi(\cdot\,|\mathbf{y}_1) = \phi(\cdot\,|\mathbf{y}_2),$ for any prior density function g and any two ensembles $\mathbf{y}_1$ and $\mathbf{y}_2$ such that $\xi(\mathbf{y}_1) = \xi(\mathbf{y}_2)$.

4.3.2 Approximately Sufficient Statistics

The NCEP's 20-member ensemble for the temperature variate was found to have an *approximately* sufficient statistic $\xi(\mathbf{y}) = (x, t)$ (i.e., a statistic with consistently comparable informativeness to that of the full ensemble), consisting of two elements: ensemble mean x (a measure of central tendency) and ensemble range t (a measure of uncertainty) (Lee 2010). It is hypothesized that similarly functioning summary statistics exist for the CMC ensemble.

Some 19 ensemble statistics are considered (see Section 5.1) as predictors of central tendency, or as predictors of uncertainty. As is the case of the NCEP ensemble, the CMC ensemble members do not constitute a random sample; as such, they need not be equally informative, and in combinations some may be extraneous. Therefore, the informativeness of various combinations of ensemble members and ensemble statistics must be tested empirically.

5. PREDICTORS OF CENTRAL TENDENCY: MAIN RESULTS

5.1 The Ensemble Statistics

In order to find approximately sufficient statistics of the ensemble, several summary statistics are hypothesized and calculated (Table 5.1). The mean, median, mode, midrange, upper mean, lower mean, mean of majority, and mean of minority are potential predictors of central tendency. The range, the width of a central credible interval, variance, coefficient of skewness, and coefficient of kurtosis are potential predictors of uncertainty about W. The mode is estimated from the mean and median (Bulmer 1979), rather than observed directly. The midrange is the midpoint between the minimum ensemble member and the maximum ensemble member. The ensemble midrange divides the ensemble members into two groups. The upper mean is the mean of ensemble members that lie above the midrange, while the lower mean is the mean of ensemble members that lie below the midrange. An ensemble member that is exactly the value of the midrange is incorporated into the calculation of both the upper mean and lower mean. The mean of majority is the mean of the majority of ensemble members that lie on one side of the midrange. The mean of minority, likewise, is the mean of the minority of ensemble members that lie on one side of the midrange. The range is the difference between the maximum ensemble member and the minimum ensemble member. The width of the p-probability central credible interval is the difference between ensemble members having appropriate ranks. The coefficient of skewness is the third moment divided by the standard deviation cubed, and measures the degree to which a density function is skewed. The coefficient of kurtosis is the fourth moment divided by the standard deviation raised to the fourth power, and measures the relative peakedness of a density function.

Table 5.1. Notation and formula for each ensemble summary statistic, hypothesized to be a predictor of either the central tendency of, or uncertainty about, predictand W.

Name	Notation	Formula
Mean	$\bar{y}$	$\frac{1}{J}\Sigma_{j=1}^{J}y_j$
Median	$y_{0.5}$	$\frac{1}{2}(y_{(J/2)}+y_{(J/2+1)})$ when J is even; $y_{((J+1)/2)}$ when J is odd
Mode	y_M	$\bar{y}+3(y_{0.5}-\bar{y})$
Midrange	y_R	$\frac{1}{2}\left[\max_j\{y_j\}+\min_j\{y_j\}\right]$
Upper mean	$\bar{y}_U$	Mean of $\{y_j: y_j \geq y_R\}$
Lower mean	$\bar{y}_L$	Mean of $\{y_j: y_j \leq y_R\}$
Mean of majority	$\bar{y}_M$	Mean of the majority of y_j on one side of y_R
Mean of minority	$\bar{y}_m$	Mean of the minority of y_j on one side of y_R
Maximum	$y_{(J)}$	$\max_j\{y_j\}$
Minimum	$y_{(1)}$	$\min_j\{y_j\}$
Range	r	$\max_j\{y_j\}-\min_j\{y_j\}$
Width of central credible interval	t_p	$y_{(J-k+1)}-y_{(k)}$, $\quad k=2,3,4,5,6$ $p=(J-2k+2)/J$, $\quad p=0.9,0.8,0.7,0.6,0.5$
Variance	d^2	$\frac{1}{J}\sum_{j=1}^{J}(y_j-\bar{y})^2$
Coefficient of skewness	cs	$\dfrac{\frac{1}{J}\Sigma_{j=1}^{J}(y_j-\bar{y})^3}{d^3}$
Coefficient of kurtosis	ck	$\dfrac{\frac{1}{J}\Sigma_{j=1}^{J}(y_j-\bar{y})^4}{d^4}$

5.2 Predictors from the Ensemble

5.2.1 Methodology for a Single Predictor

To model the stochastic dependence between the predictand and each predictor, Lee chooses a normal-linear likelihood structure:

$$X = aw + b + \Theta, \tag{5.1}$$

where X is any predictor constructed from the ensemble members, w is a realization of predictand W, and Θ is a normal variate with zero mean and constant variance σ^2. The informativeness of each predictor is determined by the parameters a and σ using the informativeness score (IS), derived by Krzysztofowicz (1992):

$$\text{IS} = \left(\frac{\sigma^2}{a^2 S^2} + 1\right)^{-1}. \tag{5.2}$$

The IS varies from 0 (no information) to 1 (perfect information). The measure is interpretable as the squared correlation coefficient between X and W, or as the relative reduction in variance of predictand W, from prior variance $Var(W) = S^2$ to posterior variance $Var(W|X = x)$:

$$\text{IS} = [Cor(X, W)]^2 = \frac{Var(W) - Var(W|X = x)}{Var(W)}. \tag{5.3}$$

After parameters a, b, and σ are estimated for each predictor, the normal-linear structure must be validated. Accordingly, a predictor will be discarded if either 1) it does not exhibit a linear dependence on the predictand, or 2) the residual exhibits heteroscedasticity or non-normal distribution. Further, a predictor will be discarded if coefficient a is not significantly different from zero, as determined by a pre-specified p-value threshold of 0.07. As with the NCEP

ensemble, the above model construction, validation, and evaluation will be performed for every hypothesized predictor *X* of *W* constructed from the CMC ensemble.

5.2.2 Individual Ensemble Members as Single Predictors

For each lead time and season, the IS of each ensemble member is calculated, and the ensemble members are sorted from most informative to least informative (from highest IS to lowest IS). For a fixed rank and season, the IS decreases as lead time increases. This observation is consistent with the expectation that ensemble forecasts become less informative with increasing lead time. For lead times 12 h, 60 h, 108 h, 156 h, 204 h, and 252 h, the IS's are generally greater in the cool season than in the warm season. For lead times 300 h and 348 h, the IS's—while all close to zero—are generally greater in the warm season than in the cool season.

Table 5.2 shows the highest three IS's and lowest three IS's for four lead times, in warm and cool seasons. The order of informativeness of the ensemble members appears to be random across lead times and seasons. Indeed, this observation is consistent with the profile plots of ensemble member ranks in Figure 5.1. The multitude of steep lines between consecutive lead times in each plot indicate large relative changes in informativeness. Thus, in the large, the ensemble members appear to be indistinguishable in terms of informativeness.

Table 5.2. The IS of the three most informative ensemble members and three least informative ensemble members, for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons.

Lead Time:	12 h				108 h			
Season:	Warm		Cool		Warm		Cool	
	Predictor	IS	Predictor	IS	Predictor	IS	Predictor	IS
Highest 3 IS	y_5	0.6293	y_{15}	0.9138	y_{13}	0.2487	y_1	0.5747
	y_{13}	0.6122	y_{10}	0.9097	y_{19}	0.2403	y_2	0.5668
	y_{11}	0.5964	y_{18}	0.9091	y_{10}	0.2399	y_{15}	0.5623
Lowest 3 IS	y_{18}	0.3752	y_{12}	0.8812	y_{15}	0.1044	y_{20}	0.4363
	y_1	0.3576	y_6	0.8810	y_{16}	0.1043	y_{17}	0.4354
	y_{14}	0.3411	y_{14}	0.8749	y_4	0.0892	y_5	0.4319

Lead Time:	204 h				300 h			
Season:	Warm		Cool		Warm		Cool	
	Predictor	IS	Predictor	IS	Predictor	IS	Predictor	IS
Highest 3 IS	y_{12}	0.0839	y_{20}	0.1349	y_1	0.0518	y_1	0.0463
	y_{10}	0.0811	y_1	0.1277	y_{19}	0.0418	y_2	0.0295
	y_{13}	0.0749	y_{10}	0.1270	y_{15}	0.0399	y_3	0.0272
Lowest 3 IS	y_{16}	0.0286	y_3	0.0644	y_2	0.0167	y_{16}	0.0078
	y_6	0.0211	y_9	0.0606	y_6	0.0144	y_7	0.0034
	y_1	0.0188	y_{12}	0.0556	y_{10}	0.0134	y_{18}	0.0027
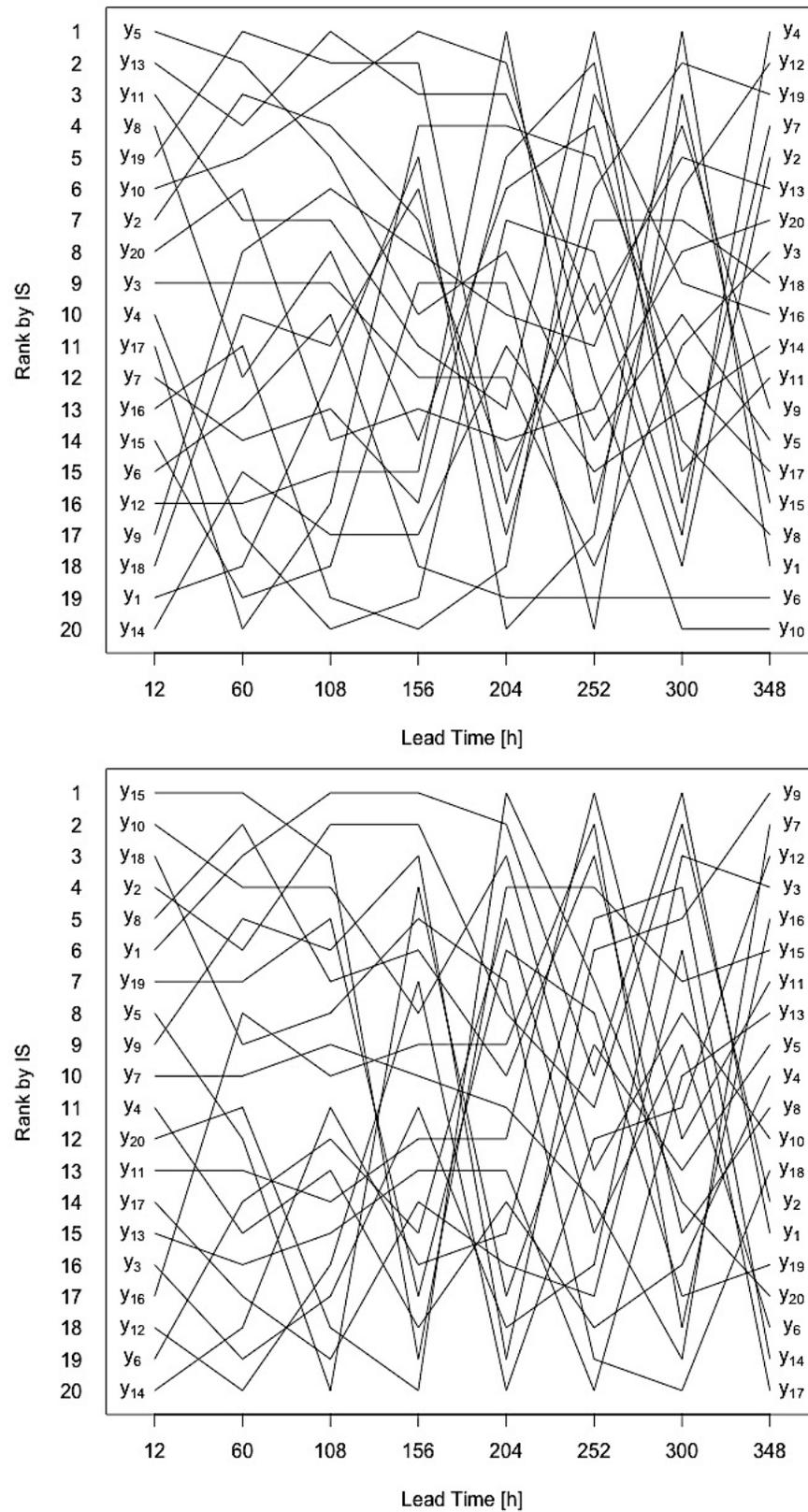
Figure 5.1. Profile plots of the rank of ensemble member by informativeness for each lead time, in the warm season (top) and cool season (bottom). The ranks fluctuate rapidly, indicating that in the large, the ensemble members are indistinguishable.

While the ensemble members can be ranked by informativeness score within each lead time and season, the IS of the most informative member may not be significantly different from the IS of the least informative member. To address this possibility, the difference between the maximum IS and minimum IS was tested for statistically significant difference from zero, within each lead time and season, according to the following hypothesis test:

Let $r = \mathrm{Cor}(Y_i, Y_j)$, where Y_i and Y_j are the most informative and least informative ensemble members respectively, and let $\max \mathrm{IS}$ and $\min \mathrm{IS}$ be their respective informativeness scores. Let the null hypothesis be that the difference $(\max \mathrm{IS} - \min \mathrm{IS})$ is zero, and let the alternative hypothesis be that this difference is greater than zero. Williams' test statistic T is defined according to Williams (1959):

$$T = (\max \mathrm{IS} - \min \mathrm{IS}) \sqrt{\frac{(N-1)(1+r)}{2\left(\frac{N-1}{N-3}\right)|R| + \overline{\mathrm{IS}}^2(1-r)^3}},$$

where $|R|$ is the determinant of the following matrix:

$$R = \begin{bmatrix} 1 & \max \mathrm{IS} & \min \mathrm{IS} \\ \max \mathrm{IS} & 1 & r \\ \min \mathrm{IS} & r & 1 \end{bmatrix},$$

$\overline{\mathrm{IS}} = (\max \mathrm{IS} + \min \mathrm{IS})/2$, and N is the number of days (sample size) for a specified lead time and season. T is assumed to follow the t distribution with $N - 3$ degrees of freedom. The p-values for the one-sided t-test are shown in Table 5.3 for all lead times in both warm and cool seasons. For lead times 12 h, 60 h, 108 h, and 156 h, the null hypothesis is rejected in each season, i.e., the difference in IS is statistically significant. However, for lead times 204 h, 252 h, 300 h, and 348 h, the p-values are high, and thus it is inconclusive as to whether the differences in informativeness among the ensemble members are significantly greater than zero.

Table 5.3. Results of the statistical hypothesis test for the significance of the difference between the IS of the most informative and least informative ensemble member, within each lead time and season. The p-values in the table are obtained from the one-sided *t*-test using Williams' test statistic (Williams 1959). A p-value less than 0.07 is chosen to infer statistical significance. For each of lead times 12 h, 60 h, 108 h, and 156 h, the IS difference is significantly greater than zero.

Lead Time	Warm	Cool	Lead Time	Warm	Cool
12 h	0.0000	0.0000	60 h	0.0003	0.0000
108 h	0.0044	0.0000	156 h	0.0422	0.0063
204 h	0.1908	0.0801	252 h	0.2688	0.1908
300 h	0.3095	0.2711	348 h	0.2840	0.3829

5.2.3 Methodology for a Combination of Predictors

In addition to individual ensemble members and summary statistics, Lee (2010) evaluates the informativeness of combinations of these predictors. A vector of predictors may be more informative than the most informative predictor. To assess the informativeness of a vector of I predictors, Lee uses the following aggregation method, which is valid under the normal-linear prediction model:

$$W = \sum_{i=1}^{I} c_i x_i + c_0 + \Xi, \tag{5.4}$$

where the residual Ξ is a normal variate with zero mean and variance τ^2. Once again, a model with one or more coefficients c_i ($i = 1, ..., I$) not significantly different from zero is discarded. An accepted model, then, defines a combined predictor X, which takes the form

$$X = \sum_{i=1}^{I} c_i x_i + c_0, \tag{5.5}$$

and its IS is calculated via the normal-linear likelihood model (5.1). This IS, determined again by the parameters a, σ, and S, indicates the informativeness of the linear combination of predictors $(X_1, ..., X_I)$. Lee notes that the sample of the predictand W is used twice, first to estimate aggregation coefficients c_i ($i = 1, ..., I$), and again to estimate the parameters (a, b, σ) of the normal-linear likelihood model for the single predictor X. Once all hypothesized predictors from

the CMC ensemble are validated and evaluated, they will be compared to the predictors constructed through Lee's analysis of the NCEP ensemble.

5.2.4 Optimal Combination of Ensemble Members

Lee (2010) describes a procedure by which to search for the most informative combination of ensemble members within each lead time and season. First, the IS of each individual ensemble member is calculated. Next, the two most informative ensemble members are selected and each is combined with every other member. Among pairs whose coefficients in normal-linear prediction model (5.4) are all significantly different from zero, the two most informative pairs are selected and combined with each remaining ensemble member to form triplets. The procedure continues likewise to select the two most informative four-tuplets, five-tuplets, and so on, until a number of ensemble members is reached at which the normal-linear prediction model for each combination contains at least one coefficient that is not significantly different from zero. When this stopping criterion is met, the "optimal" combination is one having the highest IS thus far and all coefficients significantly different from zero. Since this procedure does not calculate the IS for all possible combinations of ensemble members, it is suboptimal. However, it offers a reasonable compromise between proximity to an optimal solution and computational expense.

Table 5.4 shows the IS of the most informative four combinations of ensemble members for four lead times and two seasons, as validated under the normal-linear likelihood model, along with the IS of the most informative individual member. For each lead time and season, the optimal combination of ensemble members is more informative than either the most informative ensemble member or any ensemble statistic (Table 5.5). However, a single ensemble member does not consistently enter into the most informative combination of members, as can be inferred from a barplot of the frequency with which each ensemble member enters the optimal combination (Figure 5.2).

Table 5.4. The IS of the most informative four combinations of ensemble members, along with the IS of the most informative ensemble member, for lead times 12 h, 108 h, 204 h, and 300 h and for two seasons. There does not appear to be a member that is consistently in the most informative combination.

Lead Time	Warm Season Predictors (y)						IS	Cool Season Predictors (y)						IS
12 h	5	6	16	15	2	11	**0.7450**		15	9	5	18	10	**0.9419**
	5	6	16	15	2	8	0.7424		15	9	5	18	2	0.9416
		5	6	16	15	11	0.7403		15	9	5	10	1	0.9414
		5	6	16	15	2	0.7393		15	9	5	10	2	0.9414
						5	0.6293						15	0.9138
108 h	13	10	1	2	11	6	**0.3803**		1	2	15	10	6	**0.6810**
		13	10	1	2	11	0.3743		1	2	15	10	19	0.6801
		13	10	1	2	9	0.3743		1	2	15	10	9	0.6799
		13	10	1	2	6	0.3741		1	2	15	6	19	0.6791
						13	0.2487						1	0.5747
204 h			12	10	9	8	**0.1425**	20	15	10	2	19	8	**0.2582**
			12	10	9	14	0.1410	20	15	10	2	19	1	0.2574
				12	10	8	0.1330	20	15	10	2	1	8	0.2553
				12	10	9	0.1322		20	15	10	2	19	0.2500
						12	0.0839						20	0.1349
300 h				1	9	19	**0.0843**			1	2	3	15	**0.0898**
				1	9	15	0.0825				1	2	3	0.0816
				1	9	16	0.0812				1	2	10	0.0782
				1	19	16	0.0792				1	3	15	0.0768
						1	0.0518						1	0.0463

Of the 16 optimal combinations (eight lead times in two seasons), an ensemble member appears at most in eight (members y_9 and y_{10}). Members y_4, y_{13}, y_{14}, y_{17}, and y_{20} enter into the optimal combination two or fewer times. This marked difference in the frequency of entrances into the optimal combination of members motivates the exploration of a possible reduction of the ensemble, conducted in Chapter 6.
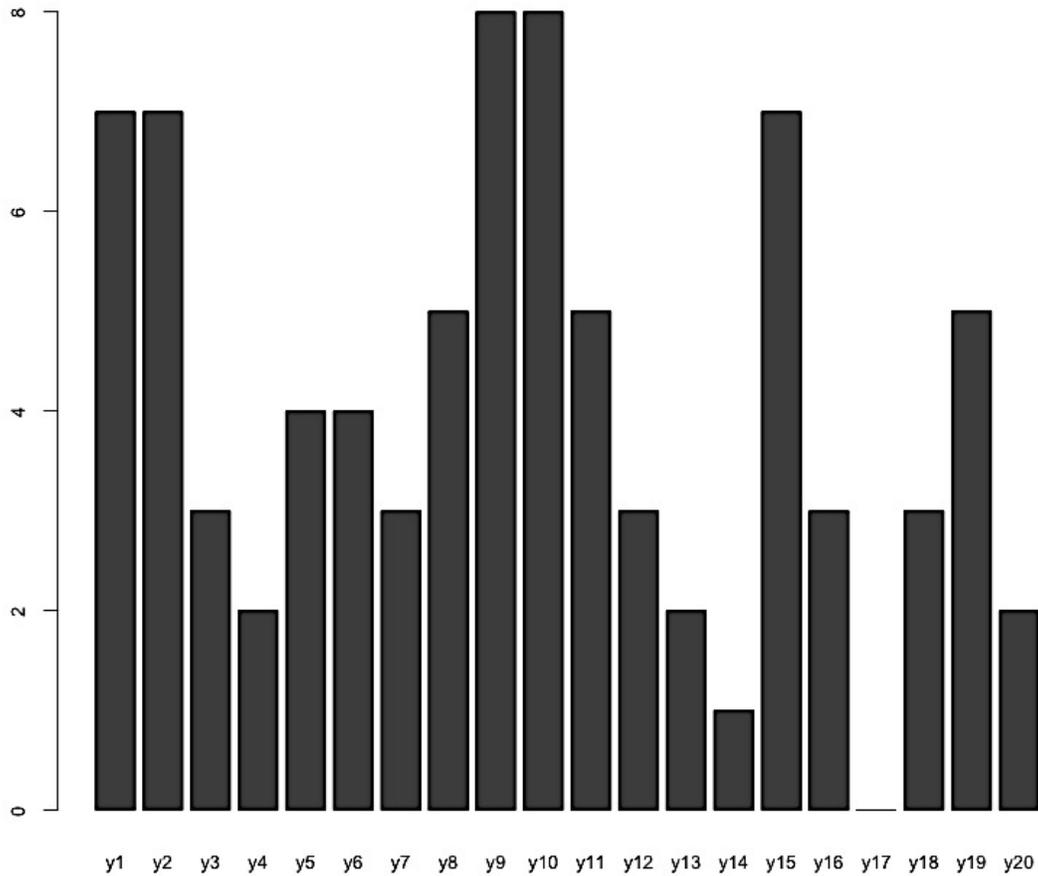
Figure 5.2. Frequency with which each ensemble member enters the optimal combination. [A member entering every combination would have a frequency of 16 (8 lead times × 2 seasons).]

5.2.5 Ensemble Statistics as Single Predictors

In addition to combinations of ensemble members, informativeness is evaluated for each ensemble statistic that is hypothesized to predict central tendency. Table 5.5 shows the IS of five such summary statistics—mean, median, midrange, upper mean, and lower mean—for four lead times. Figure 5.3 shows a profile plot of the IS of these statistics for eight lead times, in warm and cool seasons. The mean is the most informative statistic for 11 of the 16 combinations of lead time and season. Further, the IS of the ensemble mean is always greater than that of the most informative ensemble member. No ensemble statistic among the mode, midrange, majority mean, and minority mean is ever the most informative.

Table 5.5. The IS of the mean, median, midrange, upper mean, and lower mean, for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons. Overall, the mean is the most informative ensemble statistic.

12 h				108 h			
Warm		Cool		Warm		Cool	
Predictor	IS	Predictor	IS	Predictor	IS	Predictor	IS
mean	**0.7149**	mean	0.9359	mean	**0.3563**	mean	**0.6445**
median	0.7128	median	0.9328	median	0.3275	median	0.6188
midrange	0.6099	midrange	0.9375	midrange	0.2947	midrange	0.6361
upper mean	0.6562	upper mean	0.9217	upper mean	0.2469	upper mean	0.6108
lower mean	0.4204	lower mean	**0.9378**	lower mean	0.2843	lower mean	0.5990

204 h				300 h			
Warm		Cool		Warm		Cool	
Predictor	IS	Predictor	IS	Predictor	IS	Predictor	IS
mean	0.1162	mean	**0.2410**	mean	0.0683	mean	0.0676
median	**0.1225**	median	0.2271	median	0.0632	median	**0.0751**
midrange	0.0795	midrange	0.2061	midrange	0.0645	midrange	0.0310
upper mean	0.0735	upper mean	0.2240	upper mean	**0.0737**	upper mean	0.0204
lower mean	0.0726	lower mean	0.1800	lower mean	0.0320	lower mean	0.0182

5.2.6 Combination of Ensemble Statistics

While the upper mean, lower mean, majority mean, and minority mean sometimes exhibit substantial informativeness as individual predictors, each statistic only describes a subset of the ensemble members. Therefore, it is hypothesized that these statistics may prove informative in combination. In addition, the coefficient of skewness is hypothesized to enhance informativeness when combined with another statistic.

Each possible pair among nine ensemble statistics—mean, median, mode, midrange, upper mean, lower mean, majority mean, minority mean, and coefficient of skewness—was combined via the normal-linear prediction model, and the IS of the combined predictor was calculated. A pair was discarded if the prediction model or likelihood model contained any coefficient that was not significantly different from zero. Table 5.6 shows the IS of the four most informative pairs of statistics for four lead times. A pair is ranked and shown only if its model coefficients pass all hypothesis tests and its associated IS is higher than the IS of the mean.
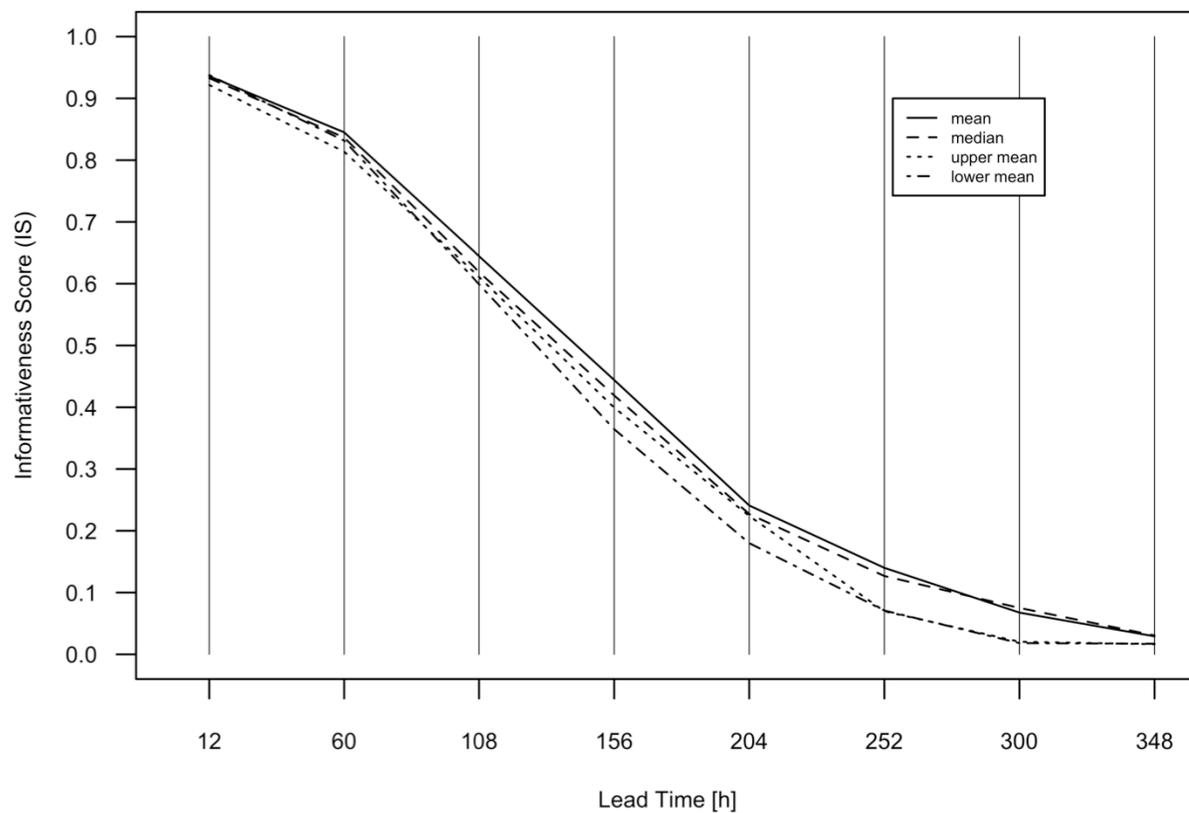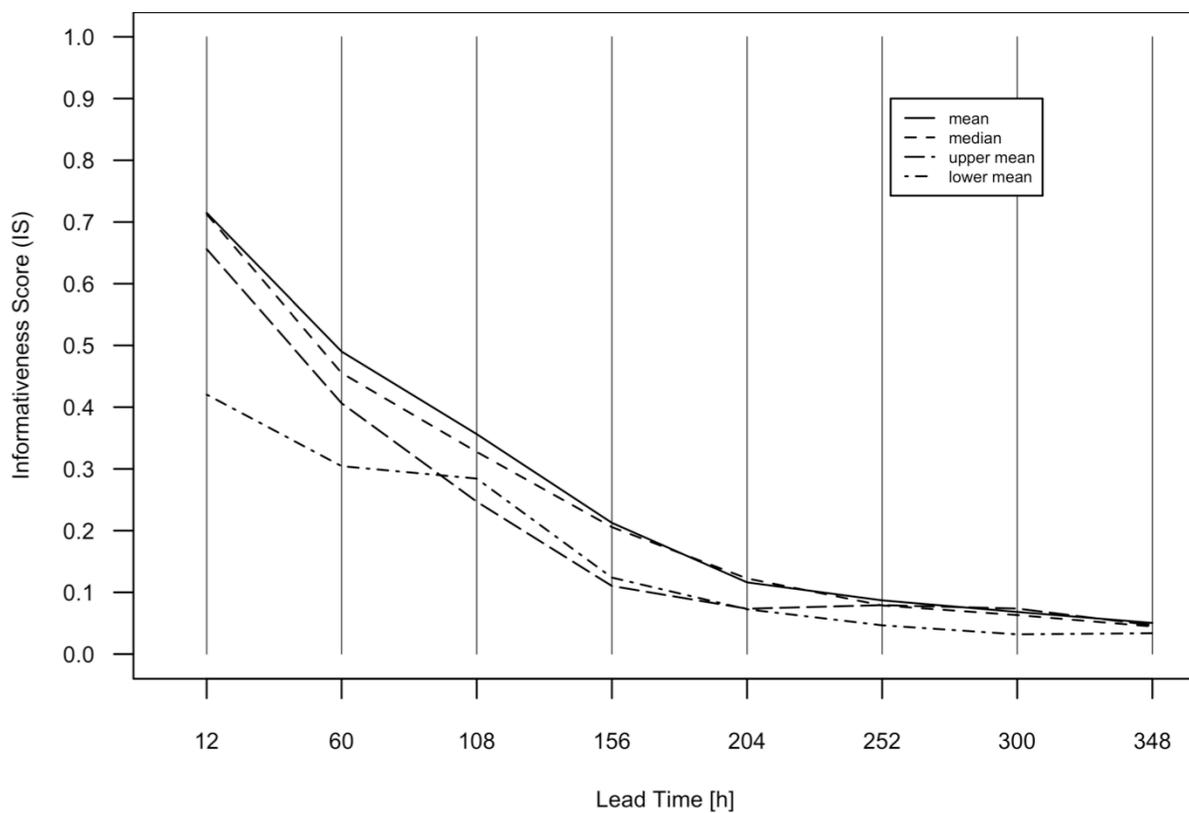
Figure 5.3. Profile plots showing the IS of four ensemble statistics—mean, median, upper mean, and lower mean—at lead times 12 h, 60 h, 108 h, 156 h, 204 h, 252 h, 300 h, and 348 h, in the warm season (top) and cool season (bottom). The ensemble mean is consistently either the most informative statistic, or has an IS near that of the most informative statistic.

Table 5.6. The four most informative pairs of ensemble statistics for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons. There always exists a pair of ensemble statistics that is more informative than the mean. However, the most informative pair varies with lead time and season.

Lead Time	Warm Season		IS	Cool Season		IS
	Predictors			Predictors		
12 h	upper mean	mean	**0.7316**	mode	lower mean	**0.9423**
	median	upper mean	0.7202	median	lower mean	0.9422
	median	lower mean	0.7178	mean	lower mean	0.9417
	median	mode*	0.7172	lower mean	majority mean	0.9415
		mean	0.7149		mean	0.9359
108 h	upper mean	mean	**0.3801**	mean	majority mean	**0.6565**
	lower mean	mean	0.3707	median	mean	0.6549
	lower mean	median	0.3647			
	median	mean	0.3629			
		mean	0.3563		mean	0.6445
204 h	mean	upper mean	**0.1271**	median	upper mean	**0.2499**
	mode	lower mean	0.1240	mean	upper mean	0.2489
	midrange	mode	0.1184	mode	upper mean	0.2467
	mode	minority mean	0.1175	majority mean	upper mean	0.2441
		mean	0.1162		mean	0.2410
300 h	midrange	lower mean	**0.0855**	upper mean	mean	**0.0896**
				mean	minority mean	0.0803
				midrange	minority mean	0.0753
				cs	midrange	0.0739
		mean	0.0683		mean	0.0676

* Since the estimate of the mode is obtained from the mean and median, the same IS is obtained for statistic pairs (mean, median) and (mean, mode). The three statistics are thus linearly dependent.

For each lead time and season, there exists a pair of ensemble statistics that is more informative than the mean. No particular pair of statistics has a consistently higher IS than that of the mean. However, the combination of mean and upper mean yields one of the highest IS's in the warm season for lead times 12 h, 60 h, 108 h, 156 h, and 204 h, and in the cool season for lead times 204 h and 300 h.

The highest IS values in Tables 5.6 and 5.4 are not much different. Consequently, if a global user prefered to use ensemble statistics over combinations of ensemble members, then it would be reasonable to use the ensemble mean as the primary predictor, enhanced with the upper mean whenever such a combination is more informative.

5.2.7 Enhancing Ensemble Mean with the Maximum or Minimum

Lee (2010) hypothesized that within the NCEP ensemble, the mean could be enhanced with the minimum or maximum ensemble member, or perhaps a combination of them. Likewise, the informativeness of every combination from the set {min, max, mean} was evaluated. Table 5.7 shows the IS of the minimum, maximum, mean, and optimal combination (if it exists) for each lead time and season. As before, predictors were discarded if any coefficient in models (5.1) or (5.4) was not significantly different from zero.

The ensemble minimum and maximum—in combination with the mean or with one another—is more informative than the mean alone only about 40% of the time. However, for shorter lead times, combining the mean with an extreme ensemble member is both beneficial and intuitive: IS is substantially improved by enhancing the mean with (i) the maximum ensemble member in the warm season for lead times up to 156 h, and (ii) the minimum ensemble member in the cool season for lead times up to 60 h. A global user is advised to employ these enhancements, for the specified lead times and seasons, whenever possible.

5.2.8 Enhancing Ensemble Mean with Ensemble Members

In Section 5.2.4, it was concluded that for each lead time and season, the optimal combination of ensemble members is more informative than the ensemble mean. It is hypothesized that the ensemble mean can be further enhanced by a combination of ensemble members. The search for such an enhanced combination proceeds similarly to the procedure described in Section 5.2.4. First, the ensemble mean is fixed as the first predictor in the combination, and is combined with every ensemble member via (5.4) − (5.5). Each of the two most informative mean-member pairs is selected and combined with every remaining ensemble

member to form a triplet. This process continues (as in Section 5.2.4) to form four-tuplets, five-tuplets, and so on, until a combination size is reached at which the normal-linear prediction model for each combination contains at least one coefficient that is not significantly different from zero. Table 5.8 shows the IS of the optimal combination of mean and ensemble members for each lead time and season.

Table 5.7. The IS of the minimum ensemble member, maximum ensemble member, and ensemble mean is shown for each lead time and season. The IS of the optimal combination of these three predictors is shown if it exists.

Lead Time	Warm Season Predictors			IS	Cool Season Predictors			IS
	Predictors			IS	Predictors			IS
12 h	min			0.3308	min			0.9376
	max			0.5411	max			0.8944
	mean	min	max	**0.7414**	mean	min		**0.9454**
			mean	0.7149			mean	0.9359
60 h	min			0.2547	min			0.8084
	max			0.2868	max			0.7689
	mean	max		**0.5244**	mean	min		**0.8492**
			mean	0.4901			mean	0.8451
108 h	min			0.1937	min			0.5728
	max			0.1516	max			0.5480
	mean	max		**0.4026**	min	max		0.6362
			mean	0.3563			mean	**0.6445**
156 h	min			0.0908	min			0.3266
	max			0.0931	max			0.3151
	mean	max		**0.2477**	min	max		0.4107
			mean	0.2126			mean	**0.4441**
204 h	min			0.0383	min			0.1489
	max			0.0648	max			0.1611
	min	max		0.0797	min	max		0.2116
			mean	**0.1162**			mean	**0.2410**
252 h	min			0.0235	min			0.0709
	max			0.0715	max			0.0342
					min	max		0.0812
			mean	**0.0870**			mean	**0.1400**
300 h	min			0.0307	min			0.0294
	max			0.0580	max			0.0114
	min	max		0.0657	mean	max		**0.0765**
			mean	**0.0683**			mean	0.0676
348 h	min			0.0222	min			0.0096
	max			0.0454	max			0.0175
			mean	**0.0504**			mean	**0.0292**

Table 5.8. The optimal combination of mean and ensemble members and its IS for lead times 12 h, 60 h, 108 h, 156 h, 204 h, 252 h, 300 h, and 348 h in warm and cool seasons.

Warm Season

Lead Time	Predictors (y)								IS
12 h				mean	6	18	12	14	0.7454
								mean	0.7149
60 h			mean	15	12	7	5	17	0.5512
								mean	0.4901
108 h				mean	12	15	4	16	0.3885
								mean	0.3563
156 h					mean	7	16	4	0.2543
								mean	0.2126
204 h									
								mean	0.1162
252 h					mean	15	12	6	0.1156
								mean	0.0870
300 h							mean	1	0.0803
								mean	0.0683
348 h							mean	10	0.0599
								mean	0.0504

Cool Season

Lead Time	Predictors (y)								IS
12 h	mean	15	17	16	12	14	13	7	0.9449
								mean	0.9359
60 h	mean	3	12	15	7	8	1	10	0.8783
								mean	0.8451
108 h				mean	1	2	15	10	0.6800
								mean	0.6445
156 h					mean	1	17	4	0.4743
								mean	0.4441
204 h						mean	14	5	0.2559
								mean	0.2410
252 h				mean	4	18	13	17	0.1887
								mean	0.1400
300 h							mean	1	0.0820
								mean	0.0676
348 h							mean	17	0.0411
								mean	0.0292

Table 5.9. Comparison between the IS of the optimal combination of ensemble members and the IS of the optimal combination of mean and ensemble members, for all eight lead times in warm and cool seasons. The optimal combination of mean and members is the more informative predictor in 10 of 16 cases.

Lead Time	Warm Season		Cool Season	
	Members	Mean + Members	Members	Mean + Members
12 h	0.7450	**0.7454**	0.9419	**0.9449**
60 h	0.5442	**0.5512**	0.8782	**0.8783**
108 h	0.3803	**0.3885**	**0.6810**	0.6800
156 h	0.2528	**0.2543**	0.4739	**0.4743**
204 h	**0.1425**	0.1162	**0.2582**	0.2559
252 h	0.1058	**0.1156**	**0.1955**	0.1887
300 h	**0.0843**	0.0803	**0.0898**	0.0820
348 h	0.0574	**0.0599**	0.0322	**0.0411**

Table 5.9 compares the IS of the optimal combination of mean and ensemble members with the IS of the optimal combination of the ensemble members alone. The combination of mean and members is more informative in 10 of 16 cases, but the gain in IS in these cases is usually minimal. This suggests that the ensemble mean cannot, consistently and substantially, be enhanced by ensemble members.

5.2.9 Comparison of Predictors

Among the ensemble statistics tested for informativeness, the ensemble mean was concluded to be the most informative overall. At each lead time and in each season, the ensemble mean is always more informative than the most informative individual member; further, a combination of ensemble members can be found with an IS higher than the IS of the ensemble mean. Therefore, the following relations exist among these predictors for each lead time:

IS(optimal combination of members) > IS(mean) > IS(most informative member).

However, the difference in IS is not always significant. Table 5.10 shows the IS of the most informative member, the ensemble mean, and the optimal combination of members for each lead time, in warm and cool seasons. For a particular lead time and season, an IS of a predictor is

underlined if it is at least 5% greater than the IS of the predictor to its left. The IS of the mean is at least 5% greater than the IS of the best member in 14 of 16 cases. The IS of the optimal combination of members is at least 5% greater than the IS of the mean in 13 of 16 cases. These results suggest that the mean is indeed preferable to the most informative member, and, computing resources permitting, the optimal combination of members is preferable to the mean.

In addition, combining the ensemble mean with an extreme ensemble member is recommended at short lead times, specifically: 1) In the warm season, for lead times less than or equal to 156 h, the mean should be enhanced with the maximum ensemble member, and 2) in the cool season, for lead times 12 h and 60 h, the mean should be enhanced with the minimum ensemble member. An asterisk next to an IS of the mean in Table 5.10 indicates that its recommended combination with the minimum or maximum member achieves an IS even greater than that of the optimal combination of members.

Table 5.10. The IS of the best ensemble member, the ensemble mean, and the optimal combination of members, for eight lead times in warm and cool seasons. In all cases, the optimal combination of members is more informative than the mean, which is more informative than the best member.

Lead Time	Warm Season			Cool Season		
	Best Member	Mean	Optimal Combination of Members	Best Member	Mean	Optimal Combination of Members
12 h	0.6293	0.7149	0.7450	0.9138	0.9359*	0.9419
60 h	0.3682	0.4901	0.5442	0.8140	0.8451	0.8782
108 h	0.2487	0.3563*	0.3803	0.5747	0.6445	0.6810
156 h	0.1580	0.2126	0.2528	0.3374	0.4441	0.4739
204 h	0.0839	0.1162	0.1425	0.1349	0.2410	0.2582
252 h	0.0583	0.0870	0.1058	0.0691	0.1400	0.1955
300 h	0.0518	0.0683	0.0843	0.0463	0.0676	0.0898
348 h	0.0458	0.0504	0.0574	0.0216	0.0292	0.0322

5.3 Control Forecast

5.3.1 Combination of Control Forecast and Ensemble Statistics

The CMC control forecast (CF) is generated using a GEM model as its dynamical core; a unique GEM model defined by changes to the physical parameterizations of the control forecast generates each ensemble member. The nature of this control member is notably different from that of the NCEP control member: The NCEP ensemble forecast system generates ensemble forecasts through direct perturbations of the control forecast. Thus, the CMC control forecast may offer additional information that is not captured by the 20-member ensemble. In the warm season across lead times, the variation in the IS of the control member makes it indistinguishable among the ensemble members. However, in the cool season, the control member is more informative than any of the ensemble members for the first six lead times. This suggests that the control member may add information when combined with ensemble statistics.

Table 5.11 shows the IS of combinations of the control forecast and each ensemble statistic for four lead times. The IS of the ensemble mean and control forecast as individual predictors are shown for reference. For four of eight lead times in the warm season, and for seven of eight lead times in the cool season, there exists a combination of ensemble statistic and control member that is more informative than the ensemble mean and control member individually. The best statistic to be paired with the control member varies with lead time and season, and the improvement in IS over that of the ensemble mean is sporadic. Because of these shortcomings, the control-statistic pair is not an acceptable predictor.

Table 5.11. The IS of combinations of the control forecast and each ensemble statistic, for lead times 12 h, 108 h, 204 h, and 300 h. A missing IS indicates an insignificant result.

Lead Time	Warm Season			Cool Season		
	Predictors		IS	Predictors		IS
12 h	CF	mean		CF	mean	0.9397
	CF	median		CF	median	0.9389
	CF	mode	0.6717	CF	mode	0.9362
	CF	midrange		CF	midrange	0.9411
	CF	upper mean		CF	upper mean	0.9358
	CF	lower mean	0.4562	CF	lower mean	**0.9414**
	CF	majority mean	0.6641	CF	majority mean	0.9331
	CF	minority mean	0.4152	CF	minority mean	0.9334
	CF	cs	0.4031	CF	cs	
		mean	**0.7149**		mean	0.9359
		CF	0.3726		CF	0.9276
108 h	CF	mean	**0.3705**	CF	mean	0.6585
	CF	median	0.3493	CF	median	0.6460
	CF	mode	0.2856	CF	mode	0.6292
	CF	midrange	0.3171	CF	midrange	**0.6617**
	CF	upper mean	0.2844	CF	upper mean	0.6570
	CF	lower mean	0.3121	CF	lower mean	0.6508
	CF	majority mean	0.2766	CF	majority mean	0.6287
	CF	minority mean	0.2263	CF	minority mean	0.6502
	CF	cs		CF	cs	
		mean	0.3563		mean	0.6445
		CF	0.1647		CF	0.6234
204 h	CF	mean	0.1242	CF	mean	0.2541
	CF	median	**0.1303**	CF	median	0.2432
	CF	mode	0.1165	CF	mode	0.2127
	CF	midrange	0.0947	CF	midrange	0.2422
	CF	upper mean	0.0903	CF	upper mean	**0.2568**
	CF	lower mean	0.0899	CF	lower mean	0.2294
	CF	majority mean	0.0972	CF	majority mean	0.2074
	CF	minority mean	0.0638	CF	minority mean	0.1879
	CF	cs		CF	cs	
		mean	0.1162		mean	0.2410
		CF	0.0486		CF	0.1619
300 h	CF	mean		CF	mean	0.0792
	CF	median		CF	median	**0.0869**
	CF	mode		CF	mode	0.0831
	CF	midrange		CF	midrange	0.0535
	CF	upper mean		CF	upper mean	0.0471
	CF	lower mean		CF	lower mean	
	CF	majority mean		CF	majority mean	0.0783
	CF	minority mean		CF	minority mean	
	CF	cs		CF	cs	0.0586
		mean	**0.0683**		mean	0.0676
		CF	0.0123		CF	0.0382

5.3.2 Combination of Control Forecast and Ensemble Members

Since the control forecast as an individual predictor exhibits average informativeness in the warm season and exceptional performance in the cool season, it may also enhance a combination of ensemble members. The search for the optimal combination of control forecast and ensemble members is essentially the same as the search for the optimal combination of ensemble mean and ensemble members. The control member, as the first predictor in the combination, is combined with every ensemble member to form pairs. Each of the most informative two pairs is then combined with every remaining ensemble member to form triplets. The search continues similarly until, for a particular combination size, every normal-linear prediction model contains a coefficient that is not significantly different from zero.

Table 5.12 shows the IS of the optimal combination of the control forecast and ensemble members for each lead time and season, alongside the IS of the optimal combination of ensemble members. The control member offers improvement in IS for four of eight lead times in the warm season, and for six of eight lead times in the cool season.

Table 5.13 compares the combination of control forecast and ensemble members, and the combination of control forecast and ensemble mean. The combination of control forecast and ensemble mean yields an unaccepted prediction model in five of 16 cases; in all other cases, the optimal combination of control forecast and ensemble members is more informative than the combination of control forecast and mean. The sporadic improvement in IS exhibited by the addition of the control forecast suggests that, in combination with ensemble members and the mean, the control forecast and ensemble members are indistinguishable.

5.3.3 Combination of Control Forecast, Ensemble Mean, and Ensemble Members

The optimal combination of ensemble members is always more informative than the mean. Further, the ensemble mean and control forecast individually are enhanced by ensemble members at various lead times. Perhaps, then, a combination of control forecast and ensemble members can be enhanced by the ensemble mean. The search for such a combination is conducted as in Section

5.3.2, except that the control forecast and the ensemble mean are fixed as the first two predictors. The combination of control, mean, and members yields an unaccepted prediction model in four of 16 cases; moreover, the infrequent improvement in IS that this combination offers (over the combination of control and members) is insubstantial (Table 5.13).

Table 5.12. The IS of the optimal combination of ensemble members, and the IS of the optimal combination of control forecast and ensemble members, for eight lead times in warm and cool seasons. The optimal combination of control forecast and members is the more informative predictor in 10 of 16 cases.

	Warm Season		Cool Season	
Lead Time	Members	CF + Members	Members	CF + Members
12 h	**0.7450**	0.7255	0.9419	**0.9427**
60 h	0.5442	**0.5745**	0.8782	**0.8790**
108 h	0.3803	**0.3867**	0.6810	**0.6844**
156 h	0.2528	**0.2677**	0.4739	**0.4803**
204 h	0.1425	**0.1507**	0.2582	**0.2589**
252 h	**0.1058**	0.0944	**0.1955**	0.1851
300 h	**0.0843**	0.0219	0.0898	**0.0980**
348 h	**0.0574**	0.0326	**0.0322**	0.0070

Table 5.13. The IS of the optimal combination of control forecast and members, combination of control forecast and mean, and optimal combination of control forecast, mean, and members, for eight lead times in warm and cool seasons. The optimal combination of control forecast and members is the most informative predictor (or the only predictor with all coefficients significantly different from zero) in 11 of 16 cases. The optimal combination of control forecast, mean, and members is the most informative predictor in five of 16 cases.

	Warm Season			Cool Season		
Lead Time	CF + Members	CF + Mean	CF + Mean + Members	CF + Members	CF + Mean	CF + Mean + Members
12 h	**0.7255**	--	0.7218	0.9427	0.9397	**0.9453**
60 h	**0.5745**	0.5314	0.5728	0.8790	0.8576	**0.8791**
108 h	**0.3867**	0.3705	0.3838	**0.6844**	0.6585	0.6788
156 h	0.2677	0.2429	**0.2790**	**0.4803**	0.4512	0.4789
204 h	**0.1507**	0.1242	0.1242	0.2589	0.2541	**0.2608**
252 h	0.0944	--	--	0.1851	0.1529	**0.1896**
300 h	0.0219	--	--	**0.0980**	0.0792	0.0911
348 h	0.0326	--	--	0.0070	--	--

5.3.4 Summary for Control Forecast

Based on the results of this section, the following relations exist among the control forecast, ensemble mean, and optimal combination of ensemble members, for each lead time and season:

IS(optimal combination of members) > IS(mean) > IS(control forecast).

Pairing the control forecast with an ensemble statistic showed infrequent improvement over the ensemble mean alone. Likewise, the comparison of the optimal combination of control forecast and ensemble members with the optimal combination of ensemble members was inconclusive. Therefore, the CMC's control forecast is not recommended as a predictor of central tendency.

5.4 Summary and Recommendations

This section compares predictors of central tendency hypothesized and tested in this chapter for which the prediction model is accepted in every season and lead time: ensemble mean, control forecast, individual ensemble members, optimal combination of members, optimal combination of mean and members, optimal combination of control and members, ensemble statistics, pairs of ensemble statistics, minimum ensemble member, and maximum ensemble member.

Based on all test results, the final choice boils down to the control forecast, ensemble mean, combination of mean and extreme ensemble member (for shorter lead times), optimal combination of ensemble members, optimal combination of mean and ensemble members, and optimal combination of the control forecast and ensemble members. Table 5.14 shows the IS of each of these predictors for eight lead times, in warm and cool seasons. Figure 5.4 contains profiles of the IS range of ensemble members along with the IS of the mean, control forecast, optimal combination of members, and optimal combination of the control forecast and members. While the optimal combination of the control forecast and members is most often the most informative predictor, its IS drops well below the IS of the optimal combination of members for longer lead times in the warm season.

Table 5.14. The IS comparison of control forecast, ensemble mean, combination of mean and extreme ensemble members, optimal combination of ensemble members, optimal combination of mean and ensemble members, and optimal combination of control forecast and ensemble members, for all eight lead times in warm and cool seasons. The highest IS for each lead time and season is boldfaced. Among the optimal combination of members, optimal combination of mean and members, and optimal combination of control forecast and members, an IS is underlined if it is not significantly different from the IS of the most informative of the three, according to a one-sided *t*-test (Section 5.2.2).

Lead Time	Warm Season					
	CF	Mean	Mean + Max	Members	Mean + Members	CF + Members
12 h	0.3726	0.7149	0.7323	<u>0.7450</u>	**0.7454**	0.7255
60 h	0.3373	0.4901	0.5244	0.5442	0.5512	**0.5745**
108 h	0.1647	0.3563	**0.4026**	<u>0.3803</u>	0.3885	<u>0.3867</u>
156 h	0.1576	0.2126	0.2477	<u>0.2528</u>	<u>0.2543</u>	**0.2677**
204 h	0.0486	0.1162	--	<u>0.1425</u>	<u>0.1162</u>	**0.1507**
252 h	0.0296	0.0870	--	<u>0.1058</u>	**0.1156**	<u>0.0944</u>
300 h	0.0123	0.0683	--	**0.0843**	<u>0.0803</u>	<u>0.0219</u>
348 h	0.0160	0.0504	--	<u>0.0574</u>	**0.0599**	<u>0.0326</u>

Lead Time	Cool Season					
	CF	Mean	Mean + Min	Members	Mean + Members	CF + Members
12 h	0.9276	0.9359	**0.9454**	0.9419	0.9449	0.9427
60 h	0.8418	0.8451	0.8492	<u>0.8782</u>	<u>0.8783</u>	**0.8790**
108 h	0.6234	0.6445	--	<u>0.6810</u>	<u>0.6800</u>	**0.6844**
156 h	0.3859	0.4441	--	<u>0.4739</u>	<u>0.4743</u>	**0.4803**
204 h	0.1619	0.2410	--	<u>0.2582</u>	<u>0.2559</u>	**0.2589**
252 h	0.0800	0.1400	--	**0.1955**	<u>0.1887</u>	<u>0.1851</u>
300 h	0.0382	0.0676	--	<u>0.0898</u>	<u>0.0820</u>	**0.0980**
348 h	0.0070	0.0292	--	<u>0.0322</u>	**0.0411**	<u>0.0070</u>
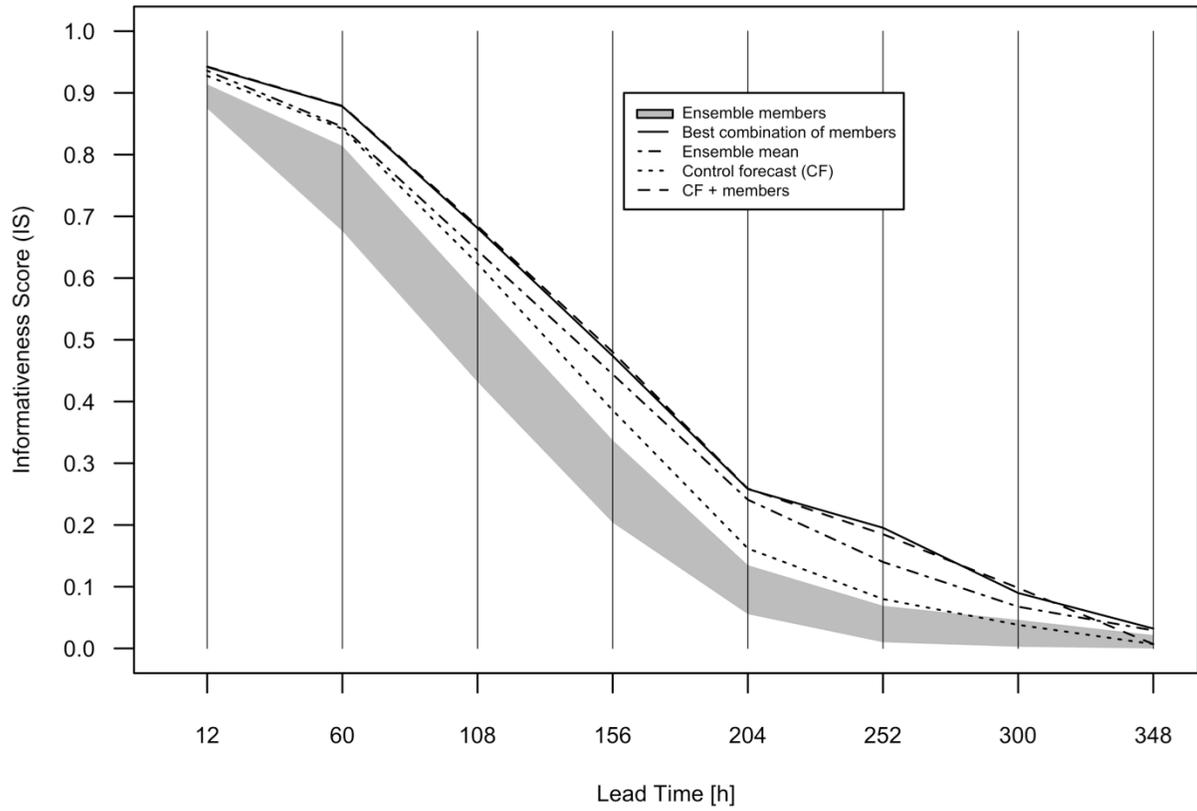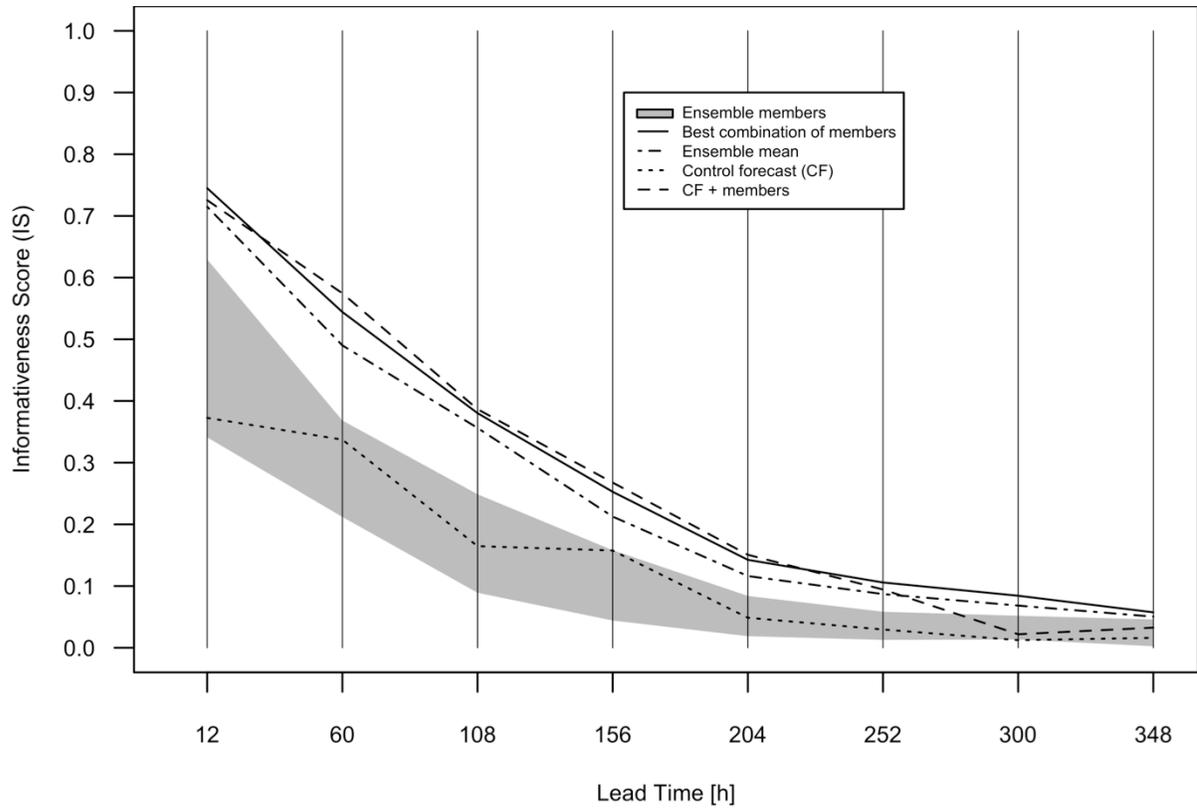
Figure 5.4. Profiles showing the range of the IS of the individual ensemble members, along with the IS of the ensemble mean, control forecast, optimal combination of ensemble members, and optimal combination of control forecast and ensemble members.

From Table 5.14 and Figure 5.4, the following conclusions can be drawn about predictors of central tendency from the CMC ensemble:

1. The most consistently informative ensemble statistic (among nine tested) is the ensemble mean. The ensemble mean or control forecast enhanced by ensemble members is always more informative than the ensemble mean alone, albeit to varying degrees.

2. Out of the 16 cases, the most informative predictor is (i) the optimal combination of the control forecast and ensemble members in eight cases, (ii) the optimal combination of the mean and ensemble members in four cases, (iii) the optimal combination of ensemble members in two cases, (iv) and the combination of mean and maximum or minimum member in two cases.

3. For each lead time and season, the IS of the optimal combination of members is greater than the IS of the mean, which is greater than the IS of the control forecast. Also, in most cases, the IS of the optimal combination of members, the IS of the optimal combination of the mean and members, and the IS of the optimal combination of the control forecast and members are not significantly different from IS of the most informative of the three predictors.

4. The following relations exist among the optimal combination of ensemble members, the optimal combination of ensemble mean and ensemble members, and the optimal combination of control forecast and ensemble members:

$$IS(members) \sim IS(mean + members) \sim IS(CF + members).$$

5. For a sophisticated user, any of the above three optimal combinations is the recommended predictor of central tendency.

6. For a global user, perhaps without the computational resources to search for optimal combinations in real time, the ensemble mean—when applicable, enhanced by the maximum or minimum ensemble member—is the recommended predictor of central tendency.

6. PREDICTORS OF CENTRAL TENDENCY: OTHER RESULTS

6.1 Informativeness of Ensemble Mean

6.1.1 Mean of Combinations of Members

This section explores the informativeness of the ensemble mean when calculated from various combinations of ensemble members. In the analysis so far, the mean of all 20 ensemble members has been evaluated for its informativeness. However, since the informativeness of individual ensemble members varies with lead time and season, it may be the case that the mean of only a few members is required to obtain an IS as high as the IS of the 20-member mean. As in Lee (2010), this possibility is tested by calculating the mean of every possible k-tuplet of ensemble members, for $k \in \{1, 2, ..., 20\}$. The IS of each mean when $k = 1$ is the IS of each individual ensemble member. The maximum number of combinations occurs at $k = 10$. As k increases, the IS's converge to a single number, the IS of the 20-member mean.

Table 6.1 shows the maximum IS among means of combinations, along with the members used to calculate the mean, for four lead times in warm and cool seasons. The number of ensemble members whose mean yields the maximum IS ranges from 6 to 13. There is little consistency in which ensemble members appear in the group whose mean yields the highest IS. While the maximum number of combinations of k-tuplets occurs when k is 10, the number of ensemble members whose mean has the maximum IS is less than 10 in six of the eight cases shown.

Figure 6.1 shows the IS of the mean of every combination of ensemble members for lead time 108 h in the warm season. In the first column, the 20 points plotted are the IS of each ensemble member as a single predictor. The maximum IS occurs at a combination size of 13 members, although the maximum IS for a combination of 10–12 members comes very close. While the maximum IS, 0.3876, is substantially greater than the IS of the 20-member mean, 0.3563, the mean of a *randomly chosen* set of 13 members could have an IS much lower than the IS of the 20-member mean. Further, considering that the number of possible 13-member

Table 6.1. Maximum IS among the mean of every possible combination of ensemble members, for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons. The number of members yielding the maximum IS of their mean ranges from 6 to 13.

	Warm Season		
Lead Time	# of members	IS	Ensemble Members
12 h	7	0.7391	2, 5, 6, 11, 13, 15, 16
108 h	13	0.3876	1, 2, 5, 6, 8, 9, 10, 11, 13, 14, 18, 19, 20
204 h	6	0.1472	8, 9, 10, 12, 13, 14
300 h	6	0.0912	1, 9, 14, 15, 16, 19

	Cool Season		
Lead Time	# of members	IS	Ensemble Members
12 h	9	0.9424	1, 2, 5, 6, 9, 10, 15, 18, 19
108 h	7	0.6812	1, 2, 6, 10, 15, 16, 19
204 h	11	0.2725	1, 2, 6, 8, 10, 11, 13, 15, 18, 19, 20
300 h	6	0.0926	1, 2, 3, 10, 14, 15

Table 6.2. The smallest number of ensemble members whose mean can have a higher IS than that of the 20-member mean, for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons. This number ranges from 2 to 5.

	Warm Season	Cool Season
Lead Time	Number of Members	Number of Members
12 h	3	3
108 h	4	2
204 h	2	5
300 h	2	3

Table 6.3. The smallest number of ensemble members whose mean is guaranteed to have a higher IS than that of the most informative ensemble member, for lead times 12 h, 108 h, 204 h, and 300 h, in warm and cool seasons. This number ranges from 7 to 16.

	Warm Season	Cool Season
Lead Time	Number of Members	Number of Members
12 h	8	10
108 h	8	12
204 h	12	7
300 h	15	16
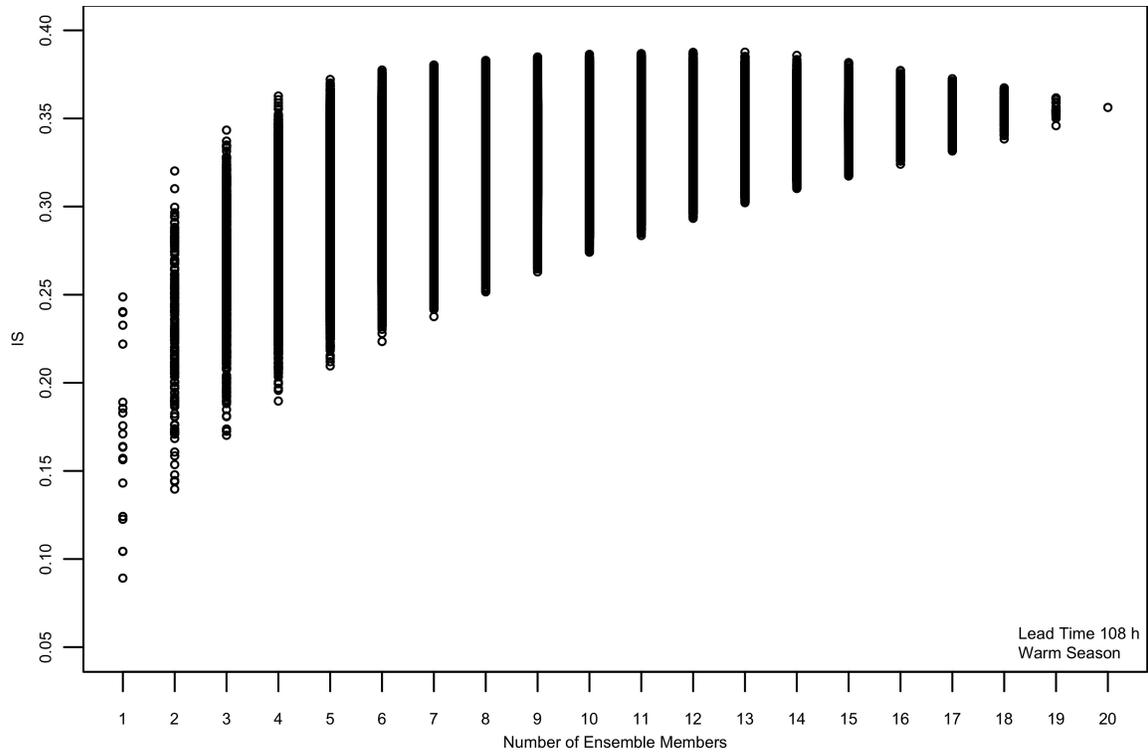
Figure 6.1. The IS of the mean of every combination of ensemble members for lead time 108 h in the warm season.
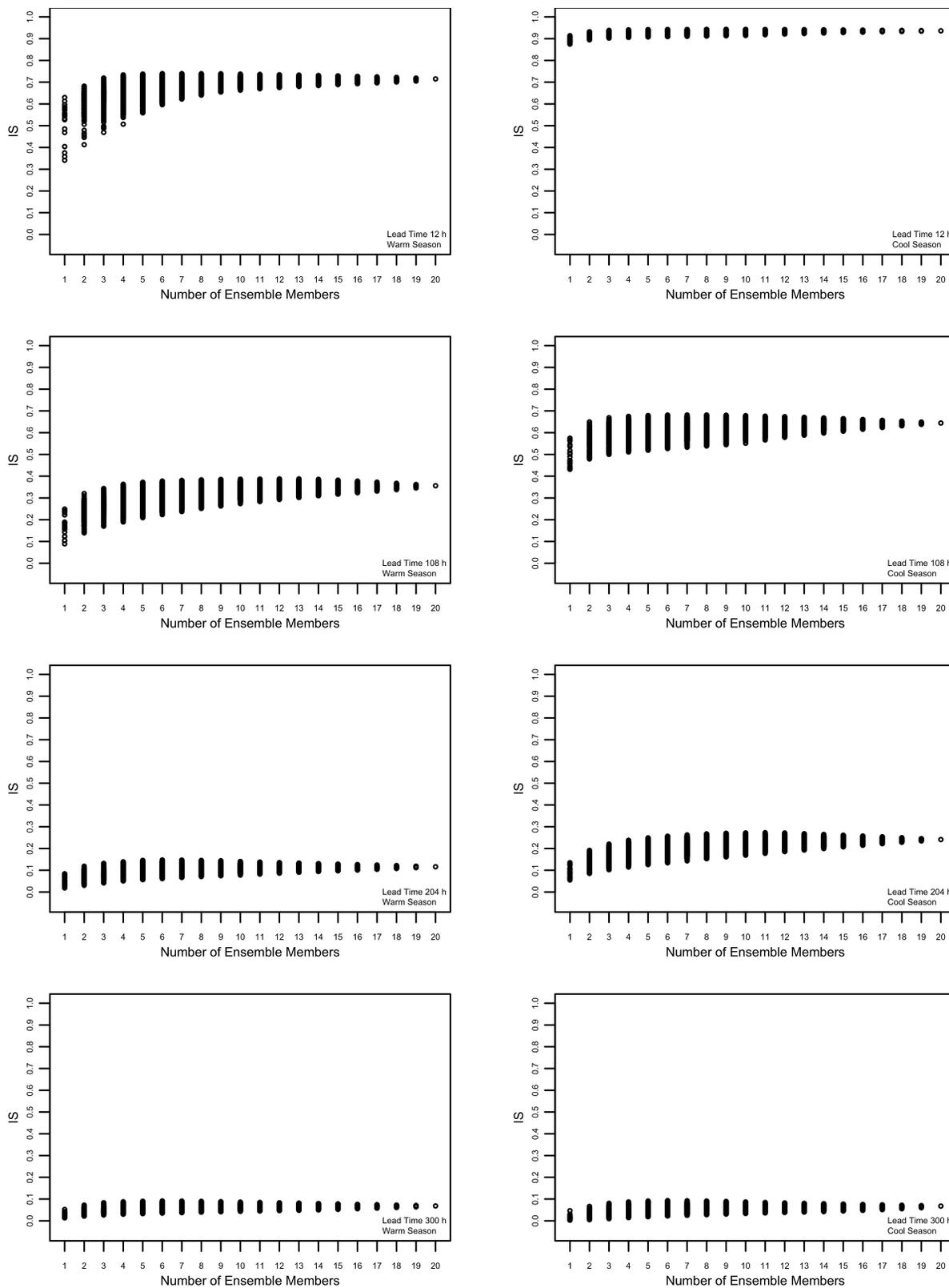
Figure 6.2. The IS of the mean of every combination of ensemble members for lead times 12 h, 108 h, 204 h, and 300 h, in each season.

combinations of 20 ensemble members is a very large number (77,520), searching for the most informative set of 13 ensemble members would be computationally too expensive to justify the gain in IS.

A more practical approach, then, is to identify smaller combinations of ensemble members that can improve upon the IS of the ensemble mean. Table 6.2 shows the smallest number of ensemble members such that the mean of their combination yields an IS at least as high as the IS of the 20-member mean. The mean of as few as 2–5 ensemble members can have an IS greater than that of the ensemble mean.

Also of interest is a mean of fewer than 20 ensemble members that is guaranteed to be more informative than the best individual ensemble member. Table 6.3 shows the smallest number of ensemble members such that the mean of any combination yields an IS that is higher than the IS of the most informative member. This number ranges from 7 to 16. For lead time 108 h in the warm season, for example, the mean of any combination of eight members yields an IS greater than that of any individual ensemble member.

In a sense, this analysis of the mean is similar to the bootstrap, except that the resampling of ensemble members is exhaustive. The results show the uncertainty about the informativeness score of the mean when only a subset of ensemble members can be used.

For comparison, Figure 6.2 shows the IS of the mean of every combination of members for four lead times in two seasons.

6.1.2 Reduction of Ensemble Size

It may be the case that the mean of a fixed set of fewer than 20 ensemble members is generally more informative than the mean of all 20 members. It is computationally too expensive to compare the IS of the mean of every fixed subset of ensemble members to the IS of the 20-member mean. However, the frequency with which a specific ensemble member enters an optimal combination may reflect the member's relative contribution to the informativeness of the ensemble mean. This idea is motivated by Figure 5.2, which shows the frequency with which each ensemble

member enters the optimal combination, over eight lead times in two seasons (16 cases). While an ensemble member enters into the optimal combination up to eight times (members y_9 and y_{10}), members y_4, y_{13}, y_{14}, y_{17}, and y_{20} enter into the optimal combination two or fewer times. Perhaps the exclusion of these members from the ensemble would generally increase the informativeness of the ensemble mean. Moreover, choosing a higher frequency threshold, and thus further reducing the ensemble size, would perhaps further improve the IS of the ensemble mean.

To help explore these possibilities, Table 6.4 compares the IS of the 20-member ensemble mean to the IS of the mean of three fixed subsets of members, for each lead time and season. Each subset—consisting of 15, 10, or 5 members—is selected by excluding all ensemble members that enter into the optimal combination, respectively, (i) two or fewer times, (ii) three or fewer times, or (iii) five or fewer times. In 13 of 16 cases, the 15-member mean is more informative than the 20-member mean; further, any decrease in IS from the 20-member mean to the 15-member mean is insubstantial. In 12 of 16 cases, the 10-member mean is more informative than the 20-member mean; any decrease in IS due to the reduction of ensemble size is, again, insubstantial. The 5-member mean is more informative than the 20-member mean in only four of 16 cases, and sometimes the decrease in IS from the 20-member mean to the 5-member mean is very large. The 15-member mean is more informative than the 10-member mean in 12 of 16 cases.

These results suggest that in general, the CMC could benefit in terms of both informativeness and practicality by using the 15-member mean, or even the 10-member mean, as the predictor of central tendency. It is not advisable, however, to use the five-member mean as the predictor of central tendency, as this reduction of ensemble size can greatly reduce the informativeness of the ensemble mean.

Table 6.4. The IS of the mean of all 20 ensemble members, and the IS of the mean of fixed subsets of 15, 10, and five ensemble members, for each lead time and season. Each of the three subsets was selected by excluding all ensemble members that enter into the optimal combination, respectively, (i) two or fewer times, (ii) three or fewer times, and (iii) five or fewer times. The greatest IS for each lead time and season is boldfaced. An IS is underlined if it is greater than the IS of the 20-member mean.

	Warm Season			
Lead Time	20 Members	15 Members	10 Members	5 Members
12 h	0.7149	<u>0.7155</u>	**<u>0.7246</u>**	0.6553
60 h	0.4901	<u>0.4923</u>	**<u>0.4945</u>**	0.4429
108 h	0.3563	0.3539	**<u>0.3636</u>**	0.3293
156 h	0.2126	<u>0.2179</u>	**<u>0.2349</u>**	<u>0.2223</u>
204 h	**0.1162**	0.1090	0.1010	0.0889
252 h	0.0870	**<u>0.0872</u>**	0.0799	0.0585
300 h	0.0683	**<u>0.0704</u>**	<u>0.0696</u>	0.0669
348 h	**0.0504**	0.0474	0.0398	0.0309

	Cool Season			
Lead Time	20 Members	15 Members	10 Members	5 Members
12 h	0.9359	**<u>0.9549</u>**	<u>0.9412</u>	<u>0.9385</u>
60 h	0.8451	**<u>0.8821</u>**	<u>0.8655</u>	<u>0.8664</u>
108 h	0.6445	**<u>0.7131</u>**	<u>0.6655</u>	<u>0.6768</u>
156 h	0.4441	**<u>0.5104</u>**	<u>0.4612</u>	0.4304
204 h	0.2410	**<u>0.2951</u>**	<u>0.2506</u>	0.2250
252 h	0.1400	**<u>0.1932</u>**	<u>0.1568</u>	0.1199
300 h	0.0676	**<u>0.1200</u>**	<u>0.0739</u>	<u>0.0815</u>
348 h	0.0292	**<u>0.0773</u>**	0.0269	0.0253

6.2 Constructed Predictors

6.2.1 Weighted Combination of Mean of Majority and Mean of Minority

This section describes the construction of predictors based on the relationships among the coefficient of skewness (cs), coefficient of kurtosis (ck), mean of majority ($\bar{y}_M$), and mean of minority ($\bar{y}_m$). (See Table 5.1.) The coefficient of skewness is large in absolute value if a density functionis highly skewed. The coefficient of kurtosis measures the relative peakedness of a density function, and is larger for a more peaked density function. The ck is bounded from below by 1. If

the density function of variate X is highly skewed or highly peaked (i.e., if $|cs|$ or ck is high), then the degree of certainty about X is greater. That is, many realizations in the random sample of X fall within a narrow region of the sample space, making the mean of majority a good predictor. Lee (2010) suggests constructing predictors that assign weights to $\bar{y}_M$ and $\bar{y}_m$ using exponential functions of $|cs|$ and ck. The first set of predictors gives more weight to $\bar{y}_M$ (and less weight to $\bar{y}_m$) as $|cs|$ increases:

$$x_1 = \left(1 - \frac{1}{2}e^{-|cs|}\right)\bar{y}_M, \qquad (6.1.1)$$

Set 1

$$x_2 = \frac{1}{2}e^{-|cs|}\bar{y}_m. \qquad (6.1.2)$$

The second set of predictors shifts ck to the origin and scales it by 0.2, which makes its range comparable to that of $|cs|$:

$$x_1 = \left(1 - \frac{1}{2}e^{-0.2(ck-1)}\right)\bar{y}_M, \qquad (6.2.1)$$

Set 2

$$x_2 = \frac{1}{2}e^{-0.2(ck-1)}\bar{y}_m. \qquad (6.2.2)$$

Each set of constructed predictors is combined and validated via the normal-linear prediction model, and the informativeness of the combined predictor is assessed via (5.1) − (5.2). Table 6.5 shows the IS of Set 1 and Set 2 for all lead times in both seasons, with the IS of the ensemble mean for reference. Set 1 has a higher IS than the mean has in 11 of 16 cases, while Set 2 has a higher IS than the mean has in four of 16 cases. While Set 1 is usually more informative than the mean, the IS gain is not consistently substantial enough to justify the cost of constructing the predictors.

6.2.2 Weighted Combination of Upper Mean and Lower Mean

Lee (2010) also considers predictors that assign weights to the upper mean ($\bar{y}_U$) and lower mean ($\bar{y}_L$). When $cs < 0$, the ensemble members are clustered toward the right of the sample space, and $\bar{y}_U$ is hypothesized to be a better predictor than $\bar{y}_L$. Likewise, when $cs > 0$, $\bar{y}_L$ is hypothesized to be the better predictor. The third set of constructed predictors reflects this hypothesis:

Set 3

$$x_1 = e^{-cs}\,\bar{y}_U,$$ (6.3.1)

$$x_2 = e^{cs}\,\bar{y}_L.$$ (6.3.2)

The fourth set of constructed predictors uses the exponential functions from Set 1. This set is split into cases determined by the sign of cs:

If $cs < 0$, then

$$x_1 = \left(1 - \frac{1}{2}e^{cs}\right)\bar{y}_U,$$ (6.4.1)

$$x_2 = \frac{1}{2}e^{cs}\,\bar{y}_L.$$ (6.4.2)

Set 4

If $cs > 0$, then

$$x_1 = \frac{1}{2}e^{-cs}\,\bar{y}_U,$$ (6.5.1)

$$x_2 = \left(1 - \frac{1}{2}e^{-cs}\right)\bar{y}_L$$ (6.5.2)

As with Set 1 and Set 2, each of the third and fourth sets of constructed predictors is combined and validated via the normal-linear prediction model and its IS is evaluated via (5.1) – (5.2). Table 6.6 shows the IS of Set 3 and Set 4, again with the IS of the ensemble mean. Set 3 is substantially less

informative than the mean, except for long lead times in the warm season. Set 4 is more informative than the mean in eight of 16 cases. Each of the two sets offers IS improvement sporadically, and yields an invalid prediction model in two of 16 cases. Thus, neither set is justifiable as a predictor of central tendency.

Table 6.5. The IS of the first two sets of constructed predictors—weighted averages of the majority mean and minority mean—for eight lead times, in warm and cool seasons. A bold IS indicates the highest IS in each lead time and season; an underlined IS indicates an IS that is higher than that of the ensemble mean. Neither set is substantially more informative than the mean.

Lead Time	Warm Season			Cool Season		
	Set 1	Set 2	Mean	Set 1	Set 2	Mean
12 h	**0.7188**	0.7004	0.7149	0.9353	**0.9370**	0.9359
60 h	**0.4904**	0.4840	0.4901	**0.8454**	0.8446	0.8451
108 h	0.3513	0.3501	**0.3563**	**0.6506**	0.6470	0.6445
156 h	0.2090	0.1968	**0.2126**	0.4451	**0.4524**	0.4441
204 h	**0.1184**	0.1070	0.1162	**0.2478**	0.2474	0.2410
252 h	0.0821	0.0830	**0.0870**	0.1359	0.1319	**0.1400**
300 h	**0.0686**	0.0674	0.0683	0.0703	--	0.0676
348 h	**0.0539**	0.0514	0.0504	**0.0314**	0.0250	0.0292

Table 6.6. The IS of the third and fourth sets of constructed predictors—weighted averages of the upper mean and lower mean—for eight lead times, in warm and cool seasons. A bold IS indicates the highest IS in each lead time and season; an underlined IS indicates an IS that is higher than that of the ensemble mean. Neither set is substantially more informative than the mean.

Lead Time	Warm Season			Cool Season		
	Set 3	Set 4	Mean	Set 3	Set 4	Mean
12 h	0.5919	**0.7206**	0.7149	0.8128	0.9353	**0.9359**
60 h	0.4347	**0.5011**	0.4901	0.8027	0.8438	**0.8451**
108 h	0.2167	**0.3665**	0.3563	0.4539	0.6392	**0.6445**
156 h	0.1578	0.2118	**0.2126**	0.3057	**0.4450**	0.4441
204 h	0.0674	**0.1212**	0.1162	0.1977	**0.2549**	0.2410
252 h	**0.0882**	0.0827	0.0870	0.1217	0.1329	**0.1400**
300 h	**0.0721**	0.0712	0.0683	--	--	0.0676
348 h	**0.0597**	0.0539	0.0504	--	--	0.0292

7. PREDICTORS OF UNCERTAINTY

7.1 Methodology

The premise of ensemble forecasting is that an ensemble $\mathbf{y}$ contains not only information about the central tendency, but also information about total uncertainty (i.e., the aggregate of uncertainties due to the NWP model, parameters, observation, sampling, and initial condition) (Kolczynski et al. 2011). The uncertainty about predictand W given a reasonable predictor X is quantified by variance σ^2 of the residual variate Θ in the likelihood model. When an ensemble provides predictor of uncertainty T in addition to predictor of central tendency X, recall from Section 4.1 that the BPE takes the form (Krzysztofowicz 1983):

$$\phi(w|x,t) = \frac{f(x|t,w)g(w)}{\kappa(x|t)}, \tag{7.1}$$

$$\kappa(x|t) = \int f(x|t,w)g(w)dw. \tag{7.2}$$

Each statistic T—the standard deviation, range, or some other predictor—is tested according to the Bayesian Gaussian-Gamma model introduced in Section 4.2 (Krzysztofowicz 1983). Under this model, conditional on realization t of T, residual Θ has variance v^2t; parameter v^2 is calculated from unconditional variance σ^2 and the parameters of a gamma distribution. The reciprocal $1/T$ is assumed to follow a gamma distribution with parameters α and β ($\alpha > 0, \beta > 1$), each of which is estimated using the method of moments:

$$m = E\left(\frac{1}{T}\right) = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{t(n)}, \tag{7.3}$$

$$s^2 = Var\left(\frac{1}{T}\right) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{t(n)} - m\right)^2, \tag{7.4}$$

$$\alpha = \frac{s^2}{m}, \tag{7.5}$$

$$\beta = \frac{m^2}{s^2}. \tag{7.6}$$

Then, v^2 is derived theoretically:

$$v^2 = \alpha(\beta - 1)\sigma^2. \tag{7.7}$$

Since residual Θ has zero mean, the value of v^2 is the slope of the linear regression of Θ^2 on T (with zero intercept):

$$E(\Theta^2|T = t) = v^2 t. \tag{7.8}$$

This fact will be used in Section 7.4.3 to validate the Gaussian-Gamma model empirically.

7.2 Candidate Predictors

Table 7.1 contains a list of candidate predictors of uncertainty. The primary candidates, selected in accordance with Lee (2010), are the standard deviation, range, widths of several central credible intervals, and the coefficient of kurtosis shifted to the origin. Also included are various combinations of the range and shifted coefficient of kurtosis. These combinations are hypothesized in order to explore potential correlation structures, in both sign and magnitude, between θ^2 and a combination of r and $ck - 1$. The linear combination of r and $ck - 1$ is constructed via linear regression:

$$E(\Theta^2|r, ck - 1) = b_0 + b_1 r + b_2(ck - 1), \tag{7.9}$$

$$t = b_1 r + b_2(ck - 1). \tag{7.10}$$

Table 7.1. Candidate predictors of uncertainty.

T	notes
d	standard deviation
r	range
t_p	width of central credible interval ($p = 0.9, 0.8, 0.7, 0.6, 0.5$)
$ck - 1$	coefficient of kurtosis, shifted to the origin
$r/(ck - 1)$	
$r(ck - 1)$	
$b_1 r + b_2(ck - 1)$	

7.3 Evaluation of Predictors

A good predictor of uncertainty T is one that is highly correlated with squared residual variate Θ^2. This analysis uses the ensemble mean as predictor of central tendency X in the normal-linear likelihood model, which determines Θ^2. The correlation of Θ^2 with each of the 11 candidate predictors (Table 7.1) is calculated, and parameters α and β of the gamma model are estimated. The best predictor for each lead time and season, shown in Table 7.2, is the candidate most highly correlated with Θ^2 that is also empirically consistent with the assumptions of the Gaussian-Gamma model. (Empirical validation of the Gaussian-Gamma model is performed in Section 7.4.)

The width of the 90% central credible interval ($t_{0.9}$) is most frequently the best predictor of uncertainty (six of 16 cases). In three cases, the width of the 50% central credible interval ($t_{0.5}$) is the best predictor, while the ensemble range (r) is the best predictor in two cases. The ensemble standard deviation d, initially thought to be a good predictor of the posterior variance of W, is of particular interest. Notably, there is no lead time and season for which d is the best predictor. A more detailed summary of the correlations, with emphasis on d, is given in Table 7.3. This table reveals that d is virtually uninformative at longer lead times. Furthermore, $t_{0.8}$ and $r/(ck-1)$ are each among the top two predictors in five of 16 cases. Overall, then, the most consistent predictors of uncertainty are $t_{0.9}$, $t_{0.8}$, and $r/(ck-1)$.

Table 7.2. The best predictor of uncertainty, along with its correlation with the squared residual variate, for each lead time and season. The correlation with the highest absolute value in each season is boldfaced.

	Warm Season		Cool Season	
Lead Time	T	$\mathrm{Cor}(T,\Theta^2)$	T	$\mathrm{Cor}(T,\Theta^2)$
12 h	r	0.2574	$t_{0.9}$	0.1639
60 h	$t_{0.9}$	**0.3408**	$b_1 r + b_2(ck-1)$	0.2417
108 h	$t_{0.9}$	0.1969	$t_{0.9}$	**0.2503**
156 h	$t_{0.9}$	0.0509	$t_{0.7}$	0.1321
204 h	$t_{0.5}$	0.1350	$r/(ck-1)$	-0.0806
252 h	$t_{0.5}$	0.1577	$t_{0.5}$	-0.2195
300 h	$r/(ck-1)$	0.0708	$t_{0.9}$	-0.1051
348 h	r	0.0437	$b_1 r + b_2(ck-1)$	0.1592

Table 7.3. Predictors of uncertainty with the two highest correlations for each lead time and season.

Season	Lead Time	Best T	Best Cor	Second Best T	Second Best Cor	Std. Dev. d	Std. Dev. Cor	$\lvert Cor(best)\rvert - \lvert Cor(d)\rvert$
Warm	12 h	r	0.2574	d	0.2220		0.2220	0.0353
	60 h	$t_{0.9}$	**0.3408**	$t_{0.8}$	**0.3340**		0.3179	0.0228
	108 h	$t_{0.9}$	0.1969	$r/(ck-1)$	0.1746		0.1605	0.0364
	156 h	$t_{0.9}$	0.0509	$t_{0.6}$	0.0482		0.0471	0.0038
	204 h	$t_{0.5}$	0.1350	$t_{0.6}$	0.1035		0.0862	0.0488
	252 h	$t_{0.5}$	0.1577	$ck-1$	-0.1499		0.0352	0.1225
	300 h	$r/(ck-1)$	0.0708	$t_{0.8}$	0.0530		0.0201	0.0507
	348 h	r	0.0437	$t_{0.7}$	-0.0407		0.0065	0.0371
Average Abs. Val.			0.1566		0.1407		0.1120	0.0447
Cool	12 h	$t_{0.9}$	0.1639	$b_1 r + b_2(ck-1)$	0.1547		0.1320	0.0320
	60 h	$b_1 r + b_2(ck-1)$	0.2417	r	**0.2374**		0.2173	0.0244
	108 h	$t_{0.9}$	**0.2503**	$t_{0.8}$	0.2341		0.2300	0.0203
	156 h	$t_{0.7}$	0.1321	$t_{0.8}$	0.1280		0.0981	0.0340
	204 h	$r/(ck-1)$	-0.0806	$t_{0.8}$	-0.0702		0.0563	0.0244
	252 h	$t_{0.5}$	-0.2195	$r/(ck-1)$	-0.1801		0.1021	0.1174
	300 h	$t_{0.9}$	-0.1051	$r/(ck-1)$	-0.0872		0.0673	0.0378
	348 h	$b_1 r + b_2(ck-1)$	0.1592	$ck-1$	0.1589		0.0378	0.1214
Average Abs. Val.			0.1691		0.1563		0.1176	0.0514

7.4 Validation of the Gaussian-Gamma Model

7.4.1 Stochastic Independence

According to an assumption of the Gaussian-Gamma model, predictor of uncertainty T should be stochastically independent of predictand W. Figure 7.1 shows scatterplots of t versus w for the best predictor for each of lead times 12 h, 108 h, 204 h, and 300 h, in both seasons. Since W is normally distributed while T is not, the correlation $Cor(T, W)$ is not a measure of stochastic dependence between the predictors. Thus, independence is confirmed visually. Still, the correlations between T and W are near zero.
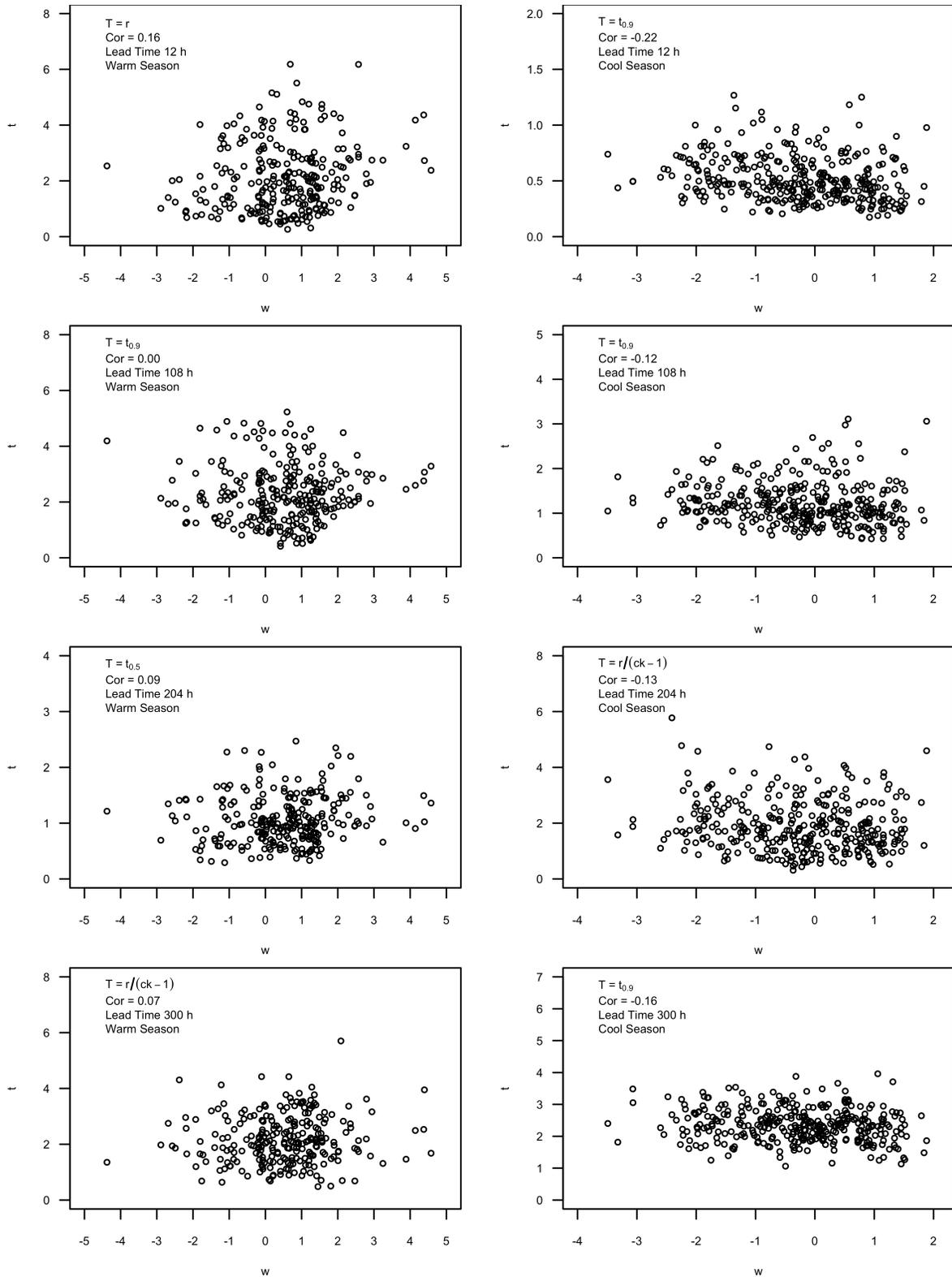
Figure 7.1. Scatter plot of t vs. w, for best predictor of uncertainty t, for lead times 12 h, 108 h, 204 h, and 300 h in each season. Stochastic independence of t and w is evident.

7.4.2 Gamma Distribution

The reciprocal $1/T$ is assumed to follow a gamma distribution. As described in Section 7.1, distribution parameters α and β are estimated using the method of moments. The values of α and β for the best predictor for each lead time and season are shown in Table 7.4. The empirical and parametric distribution functions of the reciprocal of the best predictor, for four lead times in each season, are shown in Figure 7.2. The gamma distribution appears to fit the empirical distribution well in each case.

To numerically check the fit of the gamma distribution, the two-sided Kolmogorov-Smirnov test (K-S test) is performed on the distribution of the reciprocal of the best predictor. The K-S test statistic is the maximum absolute difference between the empirical and parametric distributions. The critical value is dependent on the sample size and significance level. The null hypothesis and alternative hypothesis for the test are, respectively, as follows:

H_0: The empirical data come from the estimated gamma distribution.

H_1: The empirical data come from a distribution other than the estimated gamma

distribution.

The results of the K-S test at significance level 0.05, along with the test statistic, critical value, and p-value are shown in Table 7.5. At the chosen significance level, result $H = 0$ indicates that the null hypothesis is not rejected, while result $H = 1$ indicates a rejection of the null hypothesis. The test passes in only 50% of the cases. However, of the cases in which the conclusion is that the empirical and parametric distributions differ, three would have passed at a significance level of 0.01. Considering that the method of moments, a simple parameter estimation technique, was used to fit the parametric distributions, a better fit to the empirical data would most certainly be obtained using a more sophisticated distribution fitting method. Still, the relatively consistent fit of the estimated gamma distributions validates the distribution assumption of the Gaussian-Gamma model.
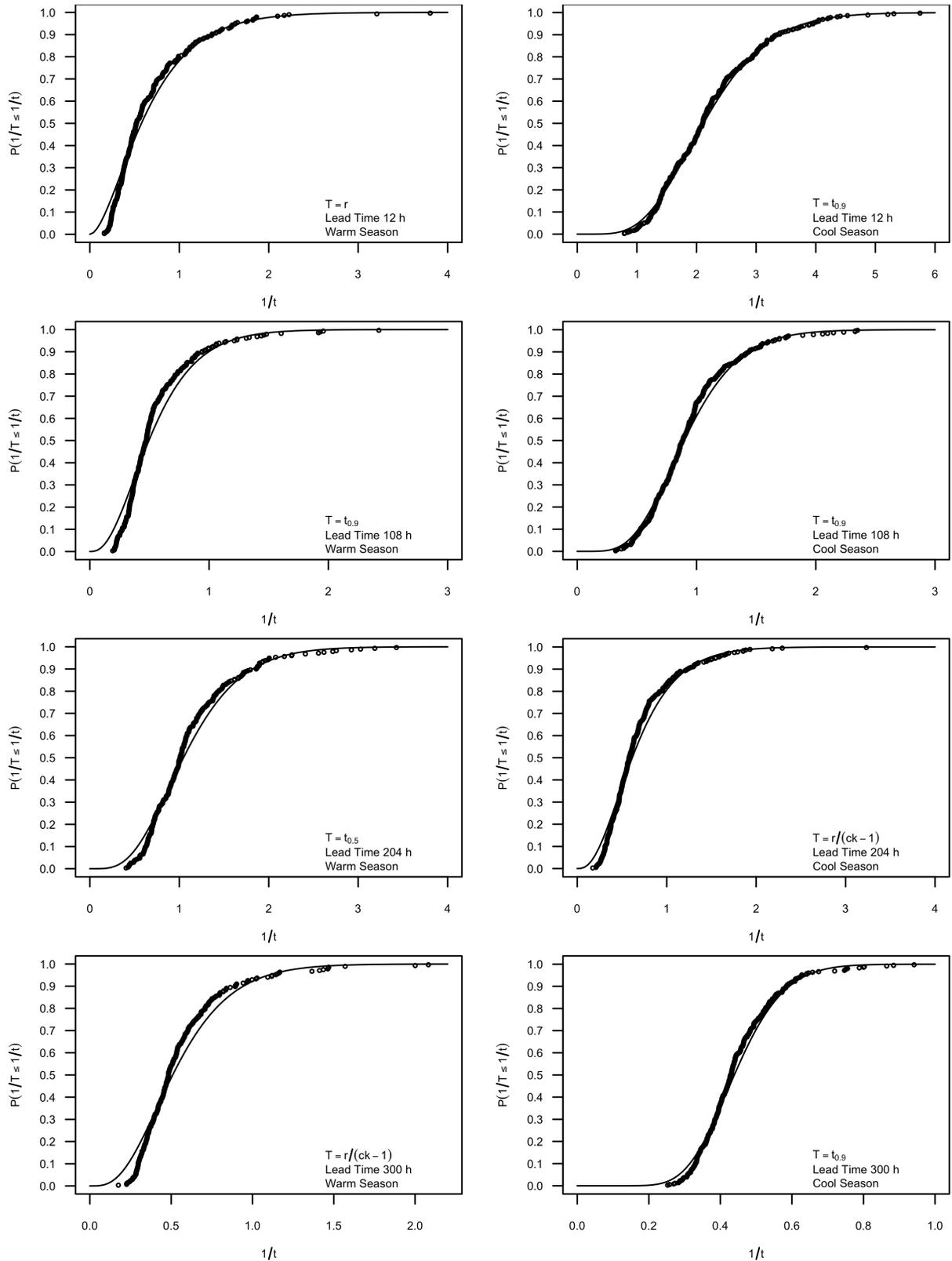
Figure 7.2. Empirical and estimated parametric gamma distribution functions of $1/T$, for the best predictor of uncertainty T, for lead times 12 h, 108 h, 204 h, and 300 h, in each season.

Table 7.4. Estimated parameter values of the gamma distribution, using the best predictor of uncertainty for each lead time and season.

Lead Time	Warm Season			Cool Season		
	T	α	β	T	α	β
12 h	r	0.3605	1.9203	$t_{0.9}$	0.3366	6.6221
60 h	$t_{0.9}$	0.2859	2.3923	$b_1 r + b_2(ck-1)$	3.0865	4.3999
108 h	$t_{0.9}$	0.1925	2.8946	$t_{0.9}$	0.1444	6.5376
156 h	$t_{0.9}$	0.0839	5.5383	$t_{0.7}$	0.2301	4.6467
204 h	$t_{0.5}$	0.2390	4.7086	$r/(ck-1)$	0.2356	2.9348
252 h	$t_{0.5}$	0.2277	4.6402	$t_{0.5}$	0.2908	4.0066
300 h	$r/(ck-1)$	0.1484	3.7581	$t_{0.9}$	0.0267	16.7653
348 h	r	0.0220	12.3064	$b_1 r + b_2(ck-1)$	1.9225	7.1662

Table 7.5. K-S test results at significance level 0.05, with test statistic, critical value, and p-value, using the best predictor of uncertainty for each lead time and season. The sample size ranges from 251 to 265 for the warm season, and is 334 for the cool season.

LT	Warm Season				Cool Season			
	H =	Test Stat.	Critical Value	p-value	H =	Test Stat.	Critical Value	p-value
12	1	0.1092	0.0828	0.0033	0	0.0400	0.0738	0.6433
60	1	0.1001	0.0831	0.0095	1	0.0751	0.0738	0.0440
108	1	0.1014	0.0834	0.0087	0	0.0617	0.0738	0.1508
156	1	0.0896	0.0837	0.0293	0	0.0729	0.0738	0.0549
204	0	0.0780	0.0840	0.0832	1	0.0780	0.0738	0.0327
252	0	0.0832	0.0843	0.0555	1	0.1144	0.0738	0.0003
300	1	0.1074	0.0846	0.0054	0	0.0703	0.0738	0.0704
348	0	0.0249	0.0849	0.9967	0	0.0402	0.0738	0.6371

7.4.3 Linearity

Under the Gaussian-Gamma model, Θ^2 and T are assumed to be linearly dependent—i.e., $E(\Theta^2|T=t) = v^2 t$, where v^2 is derived theoretically as in (7.7). To check this relationship, the line $\theta^2 = v^2 t$ is compared graphically to a zero-intercept regression line fit to the points (t, θ^2). The assumption is empirically validated if the slopes of the two lines are approximately equal. In Figure 7.3, using the best predictor for each of four lead times in both seasons, the line $\theta^2 = v^2 t$ (solid) and regression line (dashed) are shown along with a scatter plot of points (t, θ^2). The
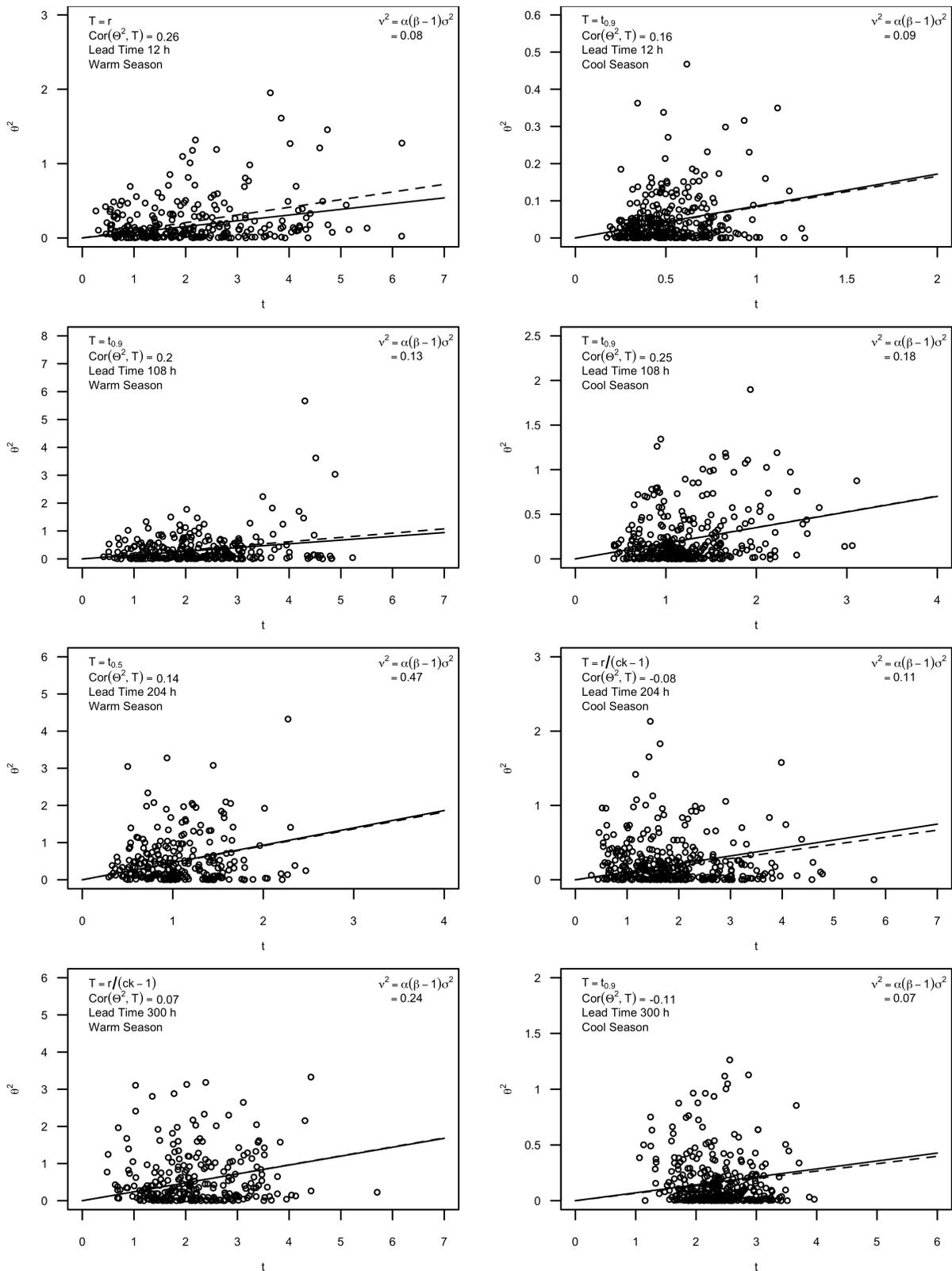
Figure 7.3. Scatter plot of points (t, θ^2) with line $\theta^2 = v^2 t$ (solid) and fitted regression line (dashed), for best predictor of central tendency in lead times 12 h, 108 h, 204 h, and 300 h in both seasons.
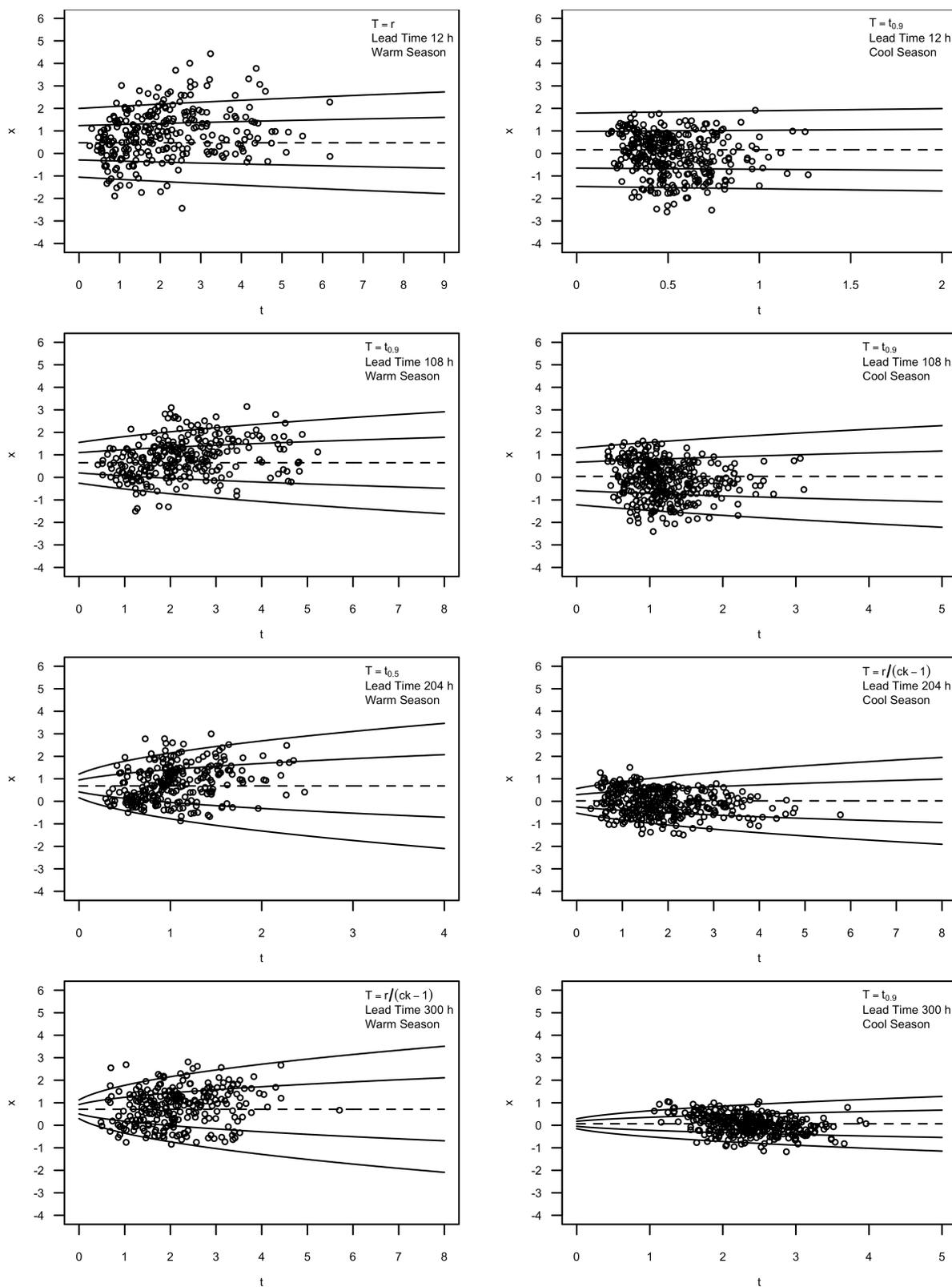
Figure 7.4. Scatter plot of points (t, x) with line $E(X|t)$ (dashed) and curves indicating ±1, ±2 standard deviations (solid), for the best predictor of uncertainty for lead times 12 h, 108 h, 204 h, and 300 h in each season. Predictor of central tendency x is the ensemble mean.

theoretically derived line lies very close to the regression line in all cases. Therefore, the theoretical relationship (7.7) under the Gaussian-Gamma model is validated. (This does not validate the linearity of the dependence structure itself. The scatterplots in Figure 7.3 do not suggest an alternative model, however. What they do suggest is that the predictor uncertainty is weak.)

7.4.4 Heteroscedasticity

As stated in Section 4.2, the distribution of predictor of central tendency X given t is normal with mean $aM + b$ and variance $a^2S^2 + v^2t$. Thus, while $E(X|t)$ is independent of t, $Var(X|t)$ grows with t. Using best predictor t for each of four lead times in each season, Figure 7.4 shows a scatter plot of points (t, x), along with $E(X|t)$ (dashed line) and curves indicating ± 1 and ± 2 standard deviations (solid lines). (As stated earlier, x denotes the ensemble mean.) The normal distribution of X implies that approximately 96% of the points in each plot should lie within two standard deviations of the expected value. The empirical data confirm this implication in each scatter plot.

7.5 Summary and Recommendations

Predictors of uncertainty from the CMC ensemble are hypothesized and evaluated according to the correlation of each with squared residual variate Θ^2, with the CMC ensemble mean as the predictor of central tendency. The hypothesized predictors include 11 ensemble statistics: standard deviation, range, widths of central credible intervals (five), shifted coefficient of kurtosis, and three combinations of range and shifted coefficient of kurtosis. Based on the analysis in this chapter, the following six conclusions and recommendations are made:

1. The best predictor of uncertainty varies with lead time and season, with no apparent trend.

2. Consistently among the best predictors are the width of the 90% and 80% central credible intervals and the quotient of the range and shifted coefficient of kurtosis. Other best predictors include the ensemble range, widths of 50% and 70% central credible intervals, and the linear combination of range and shifted coefficient of kurtosis. However, none of the predictors of uncertainty are very informative.

3. The ensemble standard deviation is never the best predictor of uncertainty, and is uninformative for lead times greater than 204 h in the warm season, and at lead time 348 h in the cool season.

4. The four assumptions of the Gaussian-Gamma model are empirically validated.

5. It is recommended for a global user to use the 90% central credible interval as the predictor of uncertainty in the BPE.

6. It is recommended for a sophisticated user to use the best predictor of uncertainty for each lead time and season (Table 7.2) in the BPE.

8. COMPARISON OF ENSEMBLES

8.1 Sample Size and Standardization

In this chapter, various properties of the predictors of central tendency and uncertainty from the NCEP and CMC ensembles are compared. Specifically, select results from the analysis of the CMC ensemble in Chapters 3-7 are compared to the results from the analysis of the NCEP ensemble from Lee (2010). It is important to note the difference in sample sizes: The NCEP ensemble has 704 days of complete forecasts (03/28/2007 – 02/28/2009) while the CMC ensemble has 599 days of complete forecasts (07/11/2007 – 02/28/2009). The NCEP and CMC ensemble forecasts thus overlap on 599 days. All numerical results in the following comparisons are based on the entire samples.

The NCEP and CMC ensembles are both standardized using the prior mean and standard deviation, and divided into warm and cool seasons as described in Section 3.1. Standardization is found to bring both sets of ensemble forecasts to near margin stationarity. Further, each member from the NCEP and CMC ensembles appears to follow a Gaussian distribution (Lee 2010).

8.2 Comparison of Correlation between Ensemble Members

The average correlation between pairs of ensemble members in each ensemble is shown in Table 8.1 for lead times 12 h, 108 h, 204 h, and 300 h in both seasons. The average correlation between pairs of ensemble members decreases with lead time in each ensemble. In each case, pairs of NCEP ensemble members are more highly correlated on average than are pairs of CMC ensemble members.

Table 8.1. In each ensemble, the average correlation between pairs of ensemble members, for each of four lead times, in warm and cool seasons.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	0.7816	**0.9859**	0.9606	**0.9957**
108 h	0.5046	**0.8836**	0.7745	**0.8666**
204 h	0.3963	**0.6089**	0.3632	**0.5421**
300 h	0.3743	**0.3814**	0.1821	**0.2348**

8.3 Comparison of Conditional Correlation between Ensemble Members

In each ensemble, conditional correlations between pairs of ensemble members, given realization w of predictand W, were calculated using the formulae given in Section 3.4. Table 8.2 shows the average conditional correlation between pairs of members in each ensemble, for lead times 12 h, 108 h, 204 h, and 300 h in warm and cool seasons. Pairs of members in each ensemble have nonzero conditional correlations, and thus each ensemble is not a random sample.

The conditional correlations of the two ensembles exhibit four primary properties: (i) In all eight cases, the average conditional correlation is less than the average correlation; this is a positive property, since it is the conditional correlation that should be zero if the ensemble is to constitute a random sample. (ii) The average conditional correlations decrease with lead time in each season, in each ensemble. (iii) In 7 of 8 cases, the average conditional correlation between pairs of CMC ensemble members is lower than the average conditional correlation between pairs of NCEP ensemble members. Thus, the CMC ensemble more so exhibits the conditional correlation structure of a random sample than does the NCEP ensemble. (However, the NCEP ensemble members are much closer to having identical distributions than are the CMC ensemble members.) (iv) Considering a conditional correlation less than 0.35 to be rather weak, the CMC ensemble members approximately exhibit conditional independence for lead times 108 h, 204 h, and 300 h in the warm season, and for lead times 204 h and 300 h in the cool seasons; the NCEP ensemble members approximately exhibit conditional independence for only lead time 300 h in each season.

Table 8.2. In each ensemble, the average conditional correlation between pairs of ensemble members, for each of four lead times, in warm and cool seasons.

	Warm Season		Cool Season	
Lead Time	CMC	NCEP	CMC	NCEP
12 h	0.4079	**0.9086**	0.5602	**0.9098**
108 h	0.3384	**0.7723**	0.5143	**0.6266**
204 h	0.3471	**0.5191**	0.2870	**0.4528**
300 h	**0.3457**	0.3295	0.1679	**0.2095**

8.4 Comparison of Predictors of Central Tendency

8.4.1 Comparison of Individual Ensemble Members

The informativeness of each predictor of central tendency from the NCEP ensemble and the CMC ensemble is evaluated via the normal-linear model detailed in Section 5.2 (Lee 2010). Table 8.3 shows the IS of the most informative member in each ensemble, for each lead time and season. The IS of the best NCEP ensemble member is consistently greater than the IS of the best CMC ensemble member.

The ensembles can also be compared in terms of average informativeness of individual members. Table 8.4 shows the mean and standard deviation of the IS's of 20 individual members in each ensemble, for each lead time and season. In each case, the average IS from the NCEP ensemble is greater than the average IS from the CMC ensemble. (However, the relation between the IS standard deviations varies with lead time and season.) Further, for a fixed average IS achieved by the CMC ensemble, the NCEP ensemble on average can achieve the same IS at a longer lead time. For instance, at lead time 156 h in the warm season, the NCEP ensemble achieves an average IS of 0.2403, which is close to the average IS of 0.2848 achieved by the CMC ensemble at lead time 60 h; therefore, the use of the NCEP forecasts instead of the CMC forecasts provides, on average, a "gain" of about four days in the warm season.

Table 8.3. The IS of the most informative member in each ensemble, for each lead time and season.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	0.6293	**0.7834**	0.9138	**0.9477**
60 h	0.3682	**0.5388**	0.8140	**0.8516**
108 h	0.2487	**0.4111**	0.5747	**0.6474**
156 h	0.1580	**0.2846**	0.3374	**0.4135**
204 h	0.0839	**0.1522**	0.1349	**0.1816**
252 h	0.0583	**0.1174**	0.0691	**0.0910**
300 h	0.0518	**0.0753**	0.0463	**0.0755**
348 h	0.0458	**0.0579**	0.0216	**0.0534**

8.4.2 Comparison of Ensemble Means

For both the NCEP and CMC ensembles, the mean is consistently the most informative single predictor of central tendency. Table 8.5 compares the IS of the NCEP ensemble mean and CMC ensemble mean for each lead time and season. The NCEP ensemble mean is more informative than the CMC ensemble mean in all cases, although the difference in IS is not always substantial. At lead time 204 h in the warm season, the NCEP ensemble mean has an IS of 0.2040, which is close to the IS of the CMC ensemble mean, 0.2126, at lead time 156 h. Likewise, at lead time 300 h in the warm season, the NCEP ensemble mean has an IS of 0.1277, which is close to the IS of the CMC ensemble mean, 0.1162, at lead time 204 h. Thus, the use of the NCEP ensemble mean over the CMC ensemble mean provides a gain of about two to four days.

Table 8.4. Average IS of 20 members in each ensemble, for each lead time and season. The standard deviation of IS's in each case is in parentheses.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	0.5187 (0.0860)	**0.7700** (0.0064)	0.8944 (0.0112)	**0.9442** (0.0020)
60 h	0.2848 (0.0512)	**0.5089** (0.0161)	0.7534 (0.0434)	**0.8425** (0.0064)
108 h	0.1708 (0.0482)	**0.3677** (0.0228)	0.4914 (0.0466)	**0.6022** (0.0308)
156 h	0.1003 (0.0284)	**0.2403** (0.0211)	0.2534 (0.0306)	**0.3490** (0.0315)
204 h	0.0492 (0.0198)	**0.1230** (0.0203)	0.0930 (0.0237)	**0.1417** (0.0243)
252 h	0.0345 (0.0136)	**0.0791** (0.0196)	0.0381 (0.0181)	**0.0649** (0.0174)
300 h	0.0276 (0.0105)	**0.0516** (0.0158)	0.0161 (0.0100)	**0.0301** (0.0181)
348 h	0.0224 (0.0096)	**0.0348** (0.0107)	0.0068 (0.0055)	**0.0199** (0.0133)

Table 8.5. The IS of the ensemble mean of each ensemble, for each lead time and season. In all cases, the NCEP ensemble mean is more informative than the CMC ensemble mean.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	0.7149	**0.7857**	0.9359	**0.9487**
60 h	0.4901	**0.5354**	0.8451	**0.8716**
108 h	0.3563	**0.4257**	0.6445	**0.7012**
156 h	0.2126	**0.3137**	0.4441	**0.4829**
204 h	0.1162	**0.2040**	0.2410	**0.2541**
252 h	0.0870	**0.1589**	0.1400	**0.1615**
300 h	0.0683	**0.1277**	0.0676	**0.1031**
348 h	0.0504	**0.0970**	0.0292	**0.0792**

Table 8.6. IS of the optimal combination of ensemble members from each ensemble, for each lead time and season. In all but two cases, the combination of members from the NCEP ensemble has the higher IS.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	0.7450	**0.7897**	0.9419	**0.9502**
60 h	0.5442	**0.5506**	**0.8782**	0.8750
108 h	0.3803	**0.4465**	0.6810	**0.7193**
156 h	0.2528	**0.3318**	0.4739	**0.5011**
204 h	0.1425	**0.2187**	0.2582	**0.2834**
252 h	0.1058	**0.1805**	**0.1955**	0.1704
300 h	0.0843	**0.1382**	0.0898	**0.1438**
348 h	0.0574	**0.1021**	0.0322	**0.0942**

8.4.3 Comparison of Optimal Combinations of Ensemble Members

The most informative combination of ensemble members is searched for in each ensemble using the procedure described in Section 5.2.4 (Lee 2010). The IS of the optimal combination of members for each ensemble, lead time, and season is shown in Table 8.6. In 14 of 16 cases, the combination of members from the NCEP ensemble has the higher IS. In the warm season, the difference in IS is consistent and grows with lead time. For lead times 108 h, 156 h, and 204 h in the warm season, the optimal combination of NCEP ensemble members provides a gain of about one day over the optimal combination of CMC ensemble members; for lead times 252 h, 300 h,

and 348 h in the warm season, the optimal combination of NCEP ensemble members provides a gain of about two days over the optimal combination of CMC ensemble members.

Also of interest are the sizes of optimal combinations of members. Table 8.7 shows the number of members constituting the optimal combination from each ensemble, for each lead time and season. In nine of 16 cases, the optimal combination from the CMC ensemble consists of more members than does the optimal combination from the NCEP ensemble. In three cases, the sizes of the optimal combinations are equal. The fact that the optimal combination of CMC members is frequently larger is consistent with the lower conditional correlations observed between pairs of CMC ensemble members.

Table 8.7. Number of members in optimal combination from each ensemble, for each lead time and season. In 12 of 16 cases, the optimal combination from the CMC ensemble consists of at least as many members as does the optimal combination from the NCEP ensemble.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	**6**	2	**5**	2
60 h	**8**	2	**8**	4
108 h	**6**	3	5	5
156 h	**5**	4	**6**	5
204 h	4	**5**	**6**	5
252 h	4	4	**8**	5
300 h	3	**4**	4	4
348 h	2	**4**	2	**4**

8.4.4 Comparison of Control Forecasts

As noted in Section 5.3.1, the CMC control forecast (CF) is generated using a GEM model as its dynamical core, the parameters of which are altered to define the dynamical core of each ensemble member. On the other hand, the NCEP forecast system generates ensemble members through direct perturbations of the control forecast. The difference in the nature of the control forecasts may contribute to differences in their informativeness. Table 8.8 shows the IS of each ensemble's control forecast, for each lead time and season. In all cases, the NCEP's control forecast has the higher IS. The IS differences are notably large in the warm season.

8.4.5 Comparison of Combinations of HR/CF and Ensemble Members

Unlike the CMC, the NCEP produces a high-resolution forecast (HR). Lee (2010) recommends that sophisticated users combine the high-resolution forecast with the ensemble members whenever possible. In the analysis of the CMC ensemble, the control forecast was used as a substitute for HR in combination with ensemble members (Section 5.3.2). Table 8.9 compares the IS of the optimal combination of ensemble members and control forecast from the CMC to the IS of the optimal combination of ensemble members and high-resolution forecast from the NCEP, for each lead time and season. In 15 of 16 cases, the combination of members and high-resolution forecast from the NCEP is the more informative predictor. Again, the IS differences between the two ensembles are pronounced in the warm season: for lead times 60 h, 108 h, 156 h, 204 h, and 252 h, the use of the optimal combination of the NCEP's high-resolution forecast and ensemble members provides a gain of about 1.5–2 days over the use of the optimal combination of the CMC's control forecast and ensemble members.

Table 8.8. The IS of the control forecast from each ensemble, for each lead time and season. The control forecast from the NCEP ensemble has the higher IS in each case.

Lead Time	Warm Season		Cool Season	
	CMC	NCEP	CMC	NCEP
12 h	0.3726	**0.7820**	0.9276	**0.9491**
60 h	0.3373	**0.5240**	0.8418	**0.8702**
108 h	0.1647	**0.4197**	0.6234	**0.6787**
156 h	0.1576	**0.2944**	0.3859	**0.3933**
204 h	0.0486	**0.1789**	0.1619	**0.1709**
252 h	0.0296	**0.0717**	0.0800	**0.0835**
300 h	0.0123	**0.0427**	0.0382	**0.0547**
348 h	0.0160	**0.0435**	0.0070	**0.0150**

Table 8.9. Comparison of IS of the optimal combination of ensemble members and control forecast from the CMC to the IS of the optimal combination of ensemble members and high resolution forecast from the NCEP. In all but one case, the combination from the NCEP ensemble has the higher IS.

Lead Time	Warm Season		Cool Season	
	CMC (CF + Members)	NCEP (HR + Members)	CMC (CF + Members)	NCEP (HR + Members)
12 h	0.7255	**0.9087**	0.9427	**0.9708**
60 h	0.5745	**0.7330**	0.8790	**0.9015**
108 h	0.3867	**0.5344**	0.6844	**0.7348**
156 h	0.2677	**0.3614**	0.4803	**0.5248**
204 h	0.1507	**0.2321**	0.2589	**0.3091**
252 h	0.0944	**0.2047**	**0.1851**	0.1776
300 h	0.0219	**0.0765**	0.0980	**0.1401**
348 h	0.0326	**0.0643**	0.0070	**0.0191**

8.5 Comparison of Predictors of Uncertainty

In this section, comparisons are made between the best predictors of uncertainty from the NCEP and CMC ensembles. A good predictor of uncertainty is one that is highly correlated with squared residual variate Θ^2 from the normal-linear likelihood model. Table 8.10 shows the summary statistic that is most highly correlated with Θ^2, in each ensemble and for each lead time and season. Among all 32 cases (2 ensembles $\times$ 8 lead times $\times$ 2 seasons), the ensemble range and width of the 90% central credible interval are most frequently the best predictors of uncertainty. Only for lead time 12 h in the warm season and for lead time 108 h in the cool season do the two ensembles have the same best predictor of uncertainty.

Table 8.11 gives the correlation between the best predictor of uncertainty and Θ^2 for each lead time and season. In 11 of 16 cases, the best predictor of uncertainty from the NCEP ensemble is more highly correlated with Θ^2.

Table 8.10. The most informative predictor of uncertainty from each ensemble, for each lead time and season. Between the two ensembles, the ensemble range is most frequently the best predictor of uncertainty.

	Warm Season		Cool Season	
Lead Time	CMC	NCEP	CMC	NCEP
12 h	r	r	$t_{0.9}$	$r(ck-1)$
60 h	$t_{0.9}$	r	$b_1 r + b_2(ck-1)$	$t_{0.9}$
108 h	$t_{0.9}$	r	$t_{0.9}$	$t_{0.9}$
156 h	$t_{0.9}$	r	$t_{0.7}$	r
204 h	$t_{0.5}$	r	$r/(ck-1)$	$ck-1$
252 h	$t_{0.5}$	r	$t_{0.5}$	$r(ck-1)$
300 h	$r/(ck-1)$	$t_{0.7}$	$t_{0.9}$	$ck-1$
348 h	r	$t_{0.5}$	$b_1 r + b_2(ck-1)$	$ck-1$

Table 8.11. The correlation of the squared residual variate with the best predictor of uncertainty from each ensemble, for each lead time and season. In 11 of 16 cases, the best predictor of uncertainty from the NCEP ensemble is more highly correlated.

	Warm Season		Cool Season	
Lead Time	CMC	NCEP	CMC	NCEP
12 h	**0.2574**	0.1919	0.1639	**0.2274**
60 h	**0.3408**	0.2931	**0.2417**	0.1085
108 h	0.1969	**0.2545**	**0.2503**	0.1888
156 h	0.0509	**0.0707**	0.1321	**0.1691**
204 h	0.1350	**0.2628**	**-0.0806**	0.0746
252 h	0.1577	**0.1969**	-0.2195	**0.3226**
300 h	0.0708	**0.1051**	-0.1059	**0.1532**
348 h	0.0437	**0.1559**	0.1592	**0.2034**

8.6 Summary of Comparisons

Predictors of central tendency and uncertainty from the NCEP and CMC ensembles are compared. In general, the NCEP ensemble yields more informative predictors of central tendency. In the warm season especially, the use of predictors of central tendency from the NCEP ensemble, instead of from the CMC ensemble, offers potential for substantial gain in lead time. The ensemble that yields the most informative predictor of uncertainty varies with lead time and season. Between the two ensembles, the most informative predictors of uncertainty are the

ensemble range and width of the 90% central credible interval. The ensemble differences summarized in this section motivate the hypothesis that combining members or summary statistics from both ensembles will yield predictors that are more informative than the predictors from either ensemble separately.

9. ENSEMBLE FUSION

9.1 Approach

When a 20-member ensemble forecast is available from the CMC as well as the NCEP, it may be the case that informative predictors of central tendency and uncertainty can be constructed by combining the two forecasts. Possibilities for such a predictor include (i) a summary statistic from all 40 ensemble members, and (ii) a combination of summary statistics from each ensemble. This section explores hypothesized predictors of central tendency and uncertainty that are constructed from the joint sample of ensemble forecasts. (The joint sample contains 599 days of ensemble forecasts, from 07/11/2007 through 02/28/2009, stratified into the warm season (265 days) and the cool season (334 days).) In this chapter, for sake of comparison, a predictor of central tendency from either ensemble individually is evaluated using a matching sample for each season. This change will only affect the evaluation of the NCEP predictors of central tendency in the warm season, since while the full sample from the NCEP contains 370 days of forecasts in the warm season, only 265 days of forecasts will be used in order to match the size of the joint sample in the warm season.

9.2 Predictors of Central Tendency from Two Ensembles

9.2.1 Combination of Ensemble Means

Perhaps a combination of the two ensemble means is more informative than either ensemble mean separately. The CMC ensemble mean $\bar{y}_C$ and the NCEP ensemble mean $\bar{y}_A$ are combined via the normal-linear prediction model:

$$W = b_1 \bar{y}_C + b_2 \bar{y}_A + b_0 + \Xi, \tag{9.1.1}$$

$$X = b_1 \bar{y}_C + b_2 \bar{y}_A + b_0. \tag{9.1.2}$$

Such a combined predictor is discarded if either b_1 or b_2 is not significantly different from zero.

Table 9.1 shows the coefficients and IS of the affine combination of means for each lead time and season. The IS of the combination of means is greater than the IS of either $\bar{y}_C$ or $\bar{y}_A$ when both coefficients are significantly different from zero—at lead times 12 h, 60 h, and 108 h in the warm season, and at all lead times except 348 h in the cool season.

Table 9.1. Coefficients and IS of affine combination of ensemble means, for each lead time and season. An omitted IS indicates a model containing at least one coefficient that is not significantly different from zero. An IS is underlined if it is greater than the IS of either ensemble mean.

	Warm Season		
Lead Time	Coef. CMC	Coef. NCEP	IS
12 h	0.3665	0.7559	0.7940
60 h	0.5565	0.5509	0.5406
108 h	0.4536	0.6531	0.4143
156 h	0.2024	0.7534	--
204 h	0.0602	0.7630	--
252 h	0.0495	0.7923	--
300 h	-0.1327	1.0746	--
348 h	-0.1706	1.1664	--

	Cool Season		
Lead Time	Coef. CMC	Coef. NCEP	IS
12 h	0.4351	0.6736	0.9554
60 h	0.4096	0.6385	0.8831
108 h	0.3654	0.6748	0.7165
156 h	0.4673	0.5589	0.5108
204 h	0.5628	0.4714	0.2956
252 h	0.5358	0.4802	0.1933
300 h	0.4059	0.5316	0.1208
348 h	0.2044	0.5734	--

9.2.2 The Grand Mean

The grand mean—the mean of all 40 ensemble members—has the potential to be more informative than either ensemble mean separately. Table 9.2 shows the IS of the grand mean for each lead time and season. In the warm season, the grand mean is more informative than either

ensemble mean at lead times 12 h, 60 h, and 108 h; in the cool season, the grand mean is more informative than either ensemble mean at all lead times except 348 h.

Table 9.2 The IS of the grand mean for each lead time and season. An IS is underlined if it is greater than the IS of either ensemble mean.

Lead Time	Warm Season	Cool Season
12 h	<u>0.7884</u>	<u>0.9549</u>
60 h	<u>0.5406</u>	<u>0.8821</u>
108 h	<u>0.4130</u>	<u>0.7131</u>
156 h	0.2656	<u>0.5104</u>
204 h	0.1513	<u>0.2951</u>
252 h	0.1262	<u>0.1932</u>
300 h	0.1094	<u>0.1200</u>
348 h	0.0794	0.0773

9.2.3 Optimal Combination of Ensemble Members

A third option for a predictor of central tendency from the joint sample is to search for the most informative combination among all 40 ensemble members (all 20 members from each ensemble). The search procedure is exactly as described in Section 5.2.4, except that at each iteration k, the two most informative k-tuplets are combined with every remaining member of *both* ensembles. Table 9.3 shows the optimal combination among all 40 ensemble members, along with its size and IS. In every case, the optimal combination is more informative than either ensemble mean. Notably, in every case except lead time 12 h in the warm season, the size of the optimal combination among 40 members is at least as great as the size of the optimal combination of members in each ensemble processed separately (see Table 8.7). A possible explanation for the larger combination sizes in the joint sample is that at each iteration k of the search procedure, there is a larger pool of members ($40 - k$ members vs. $20 - k$ members) from which to potentially add information to the current two best combinations.

Table 9.3 Optimal combination among all 40 ensemble members for the joint sample of NCEP and CMC ensembles, for each lead time and season. An IS is underlined if it is greater than the IS of either ensemble mean.

Warm Season				
Lead Time	Combination size	CMC members	NCEP members	IS
12 h	2 + 2 = 4	6, 8	1, 15	<u>0.8149</u>
60 h	7 + 4 = 11	5, 9, 12, 14, 18, 19, 20	2, 4, 11, 14	<u>0.6209</u>
108 h	3 + 3 = 6	1, 9, 10	8, 14, 15	<u>0.4655</u>
156 h	6 + 3 = 9	1, 7, 9, 10, 16, 18	2, 7, 11	<u>0.3533</u>
204 h	1 + 3 = 4	10	9, 17, 18	<u>0.1838</u>
252 h	2 + 5 = 7	11, 15	3, 6, 8, 12, 18	<u>0.1911</u>
300 h	2 + 4 = 6	1, 10	2, 6, 11, 19	<u>0.1730</u>
348 h	2 + 5 = 7	10, 14	2, 6, 10, 17, 20	<u>0.1583</u>

Cool Season				
Lead Time	Combination size	CMC members	NCEP members	<u>IS</u>
12 h	2 + 4 = 6	10, 15	1, 4, 9, 14	<u>0.9613</u>
60 h	5 + 3 = 8	1, 3, 5, 10, 15	2, 9, 19	<u>0.8967</u>
108 h	2 + 4 = 6	1, 9	9, 10, 12, 17	<u>0.7424</u>
156 h	3 + 5 = 8	1, 2, 9	1, 3, 8, 10, 16	<u>0.5446</u>
204 h	2 + 4 = 6	2, 10	5, 11, 13, 16	<u>0.3277</u>
252 h	5 + 3 = 8	8, 9, 11, 15, 16	14, 15, 18	<u>0.2301</u>
300 h	2 + 4 = 6	1, 3	2, 6, 8, 11	<u>0.1679</u>
348 h	1 + 4 = 5	7	7, 9, 10, 16	<u>0.1051</u>

9.2.4 Comparison of Predictors

Table 9.4 compares the IS of the CMC and NCEP ensemble means, the affine combination of ensemble means, the grand mean, and the optimal combination among 40 members for each lead time and season. The optimal combination of members is consistently the most informative predictor of central tendency. However, it is important to note that the membership in the optimal combination varies across the lead times and seasons; it may also be sensitive to sample size. Such instability of the membership would complicate operational implementation. Among four

remaining predictors, the choice that ensures the highest IS is the affine combination of means when it produces a valid linear model (lead times less than or equal to 108 h in the warm season and less than or equal to 300 h in the cool season), and the NCEP mean otherwise; this choice is also easiest to implement operationally.

Table 9.4. The IS of individual NCEP and CMC ensemble means, affine combination of ensemble means, grand mean, and optimal combination among 40 members. An IS is underlined if it is the highest among the IS's in the first four columns of predictors.

| Lead Time | Warm Season | | | | |
	CMC Mean	NCEP Mean	Combination of Means	Grand Mean	Members
12 h	0.7149	0.7748	<u>0.7940</u>	0.7884	**0.8149**
60 h	0.4901	0.4945	<u>0.5406</u>	0.5406	**0.6209**
108 h	0.3563	0.3919	<u>0.4143</u>	0.4130	**0.4655**
156 h	0.2126	<u>0.2702</u>	--	0.2656	**0.3533**
204 h	0.1162	<u>0.1618</u>	--	0.1513	**0.1838**
252 h	0.0870	<u>0.1408</u>	--	0.1262	**0.1911**
300 h	0.0683	<u>0.1413</u>	--	0.1094	**0.1730**
348 h	0.0504	<u>0.1103</u>	--	0.0794	**0.1583**

| Lead Time | Cool Season | | | | |
	CMC Mean	NCEP Mean	Combination of Means	Grand Mean	Members
12 h	0.9359	0.9487	<u>0.9554</u>	0.9549	**0.9613**
60 h	0.8451	0.8716	<u>0.8831</u>	0.8821	**0.8967**
108 h	0.6445	0.7012	<u>0.7165</u>	0.7131	**0.7424**
156 h	0.4441	0.4829	<u>0.5108</u>	0.5104	**0.5446**
204 h	0.2410	0.2541	<u>0.2956</u>	0.2951	**0.3277**
252 h	0.1400	0.1615	<u>0.1933</u>	0.1932	**0.2301**
300 h	0.0676	0.1031	<u>0.1208</u>	0.1200	**0.1679**
348 h	0.0292	<u>0.0792</u>	--	0.0773	**0.1051**

9.3 Predictors of Uncertainty from Two Ensembles

9.3.1 Grand Predictors of Uncertainty

Various summary statistics of the vector of 40 ensemble members are tested as predictors of uncertainty. The grand mean is used as the predictor of central tendency to generate a sample of

Table 9.5. Best two predictors of uncertainty from vector of 40 ensemble members, when the predictor of central tendency is the grand mean, and correlation with Θ^2, for each lead time and season. A correlation is underlined if its absolute value exceeds the absolute value of the correlation of the best predictor of uncertainty from each ensemble, NCEP and CMC (with the ensemble mean of each, respectively, as the predictor of central tendency), processed separately.

Season	Lead Time	Best		Second Best	
		T	Cor	T	Cor
Warm	12 h	$b_1 r + b_2(ck - 1)$	0.2429	r	0.2410
	60 h	$t_{0.7}$	0.3020	$t_{0.9}$	0.2961
	108 h	$r/(ck - 1)$	0.3155	$t_{0.6}$	0.2553
	156 h	$r/(ck - 1)$	0.1624	$t_{0.6}$	0.1508
	204 h	$t_{0.8}$	0.2772	$t_{0.7}$	0.2237
	252 h	$r/(ck - 1)$	0.2762	$t_{0.6}$	0.2454
	300 h	$r/(ck - 1)$	0.1598	$t_{0.6}$	0.1503
	348 h	$t_{0.8}$	0.2060	d	0.1978
Average of Absolute Value			0.2428		0.2200
Cool	12 h	$r/(ck - 1)$	-0.1270	$b_1 r + b_2(ck - 1)$	0.0943
	60 h	$b_1 r + b_2(ck - 1)$	0.2872	r	0.2822
	108 h	$t_{0.9}$	0.2104	$t_{0.7}$	0.2100
	156 h	$t_{0.8}$	0.1613	$t_{0.9}$	0.1520
	204 h	$ck - 1$	0.1305	$r(ck - 1)$	0.1056
	252 h	$b_1 r + b_2(ck - 1)$	0.1226	r	0.1105
	300 h	$ck - 1$	0.1233	$r/(ck - 1)$	-0.1136
	348 h	$b_1 r + b_2(ck - 1)$	0.1682	$ck - 1$	0.1680
Average of Absolute Value			0.1663		0.1545

residuals $\{\theta_j(n) : j = 1, \dots, 40, \ n = 1, \dots, N\}$ from the normal-linear likelihood model, with $N = 265$ for the warm season and $N = 334$ for the cool season. Table 9.5 shows the two predictors of uncertainty T (among the 11 candidate predictors from Table 7.1) that are most highly correlated with squared residual variate Θ^2. While the best predictor varies with lead time and season, the quotient of ensemble range and shifted coefficient of kurtosis is most frequently among the top two predictors of uncertainty. Among the single predictors of uncertainty, the width of the 60% and 80% central credible intervals are the most informative. However, the predictor $r/(ck - 1)$, composed of two predictors of uncertainty, is most frequently the best predictor of

uncertainty. The correlation with Θ^2 of the best predictor only exceeds the correlation with Θ^2 of the best NCEP and CMC predictors of uncertainty in eight of 16 cases.

9.3.2 Combined Predictors of Uncertainty

The best predictors of uncertainty may change when a different predictor of central tendency is used to generate the residuals from the likelihood model. From an operational perspective, it is attractive to combine summary statistics of each ensemble, as opposed to combining all 40 ensemble members. Along these lines, a second predictor of central tendency used to generate Θ^2 is the affine combination of ensemble means described in Section 9.2.1. For a lead time and season for which a combination of ensemble means yields an invalid likelihood model (see Table 9.4), the NCEP ensemble mean alone is used as the predictor of central tendency. For each single candidate predictor of uncertainty (among the ensemble standard deviation, ensemble range, width of a central credible interval, and shifted coefficient of kurtosis), an affine combination of predictors of uncertainty is constructed by performing a bivariate regression of Θ^2 on the same predictor of uncertainty from each ensemble—e.g., the CMC ensemble range and the NCEP ensemble range—and combining them into a single vector via the normal-linear model. However, every regression on this type of hypothesized predictor of uncertainty contains a coefficient that is not significantly different from zero. Thus, an affine combination of the same predictor of uncertainty from each ensemble is an invalid predictor.

There may still exist other "combined" predictors of uncertainty from each ensemble— with an affine combination of means as the predictor of central tendency—that are more informative than the "grand" summary statistics from Table 9.5. To specify such combined predictors, let a single predictor of uncertainty (the ensemble standard deviation, ensemble range, width of a central credible interval, or shifted coefficient of kurtosis) from the CMC ensemble be denoted by t_C, and let the same single predictor of uncertainty from the NCEP ensemble be denoted by t_A. The following six groups of candidate combined predictors of uncertainty from the two ensembles are tested:

1. The CMC predictor alone: t_C

2. The NCEP predictor alone: t_A

3. The arithmetic mean of the two predictors: $\frac{1}{2}(t_C + t_A)$

4. The geometric mean of the two predictors: $\sqrt{t_C \cdot t_A}$

5. The minimum of the two predictors: $\min\{t_C, t_A\}$

6. The maximum of the two predictors: $\max\{t_C, t_A\}$

The additional candidate predictors $r/(ck-1)$, $r(ck-1)$, and $b_1 r + b_2(ck-1)$ are then constructed from the combined single predictors in each of the six groups.

The correlations between each predictor of uncertainty and squared residual variate Θ^2 suggest that no group of predictors, among the six hypothesized, is consistently the most informative. Still, in order to narrow the selection, the six groups are ranked based on the *average* performance of the best predictor of uncertainty across eight lead times.

Table 9.6 shows the mean absolute value, across eight lead times, of the correlation between the best predictor of uncertainty and Θ^2 in each group and season. In the warm season, the best group of predictors of uncertainty, on average, is the group of NCEP predictors; in the cool season, the best group of predictors of uncertainty, on average, is the group consisting of the geometric mean of each pair of individual predictors t_A and t_C. However, the differences between the best group and the second best are very small, and so the arithmetic mean and the geometric mean can be considered nearly equally good for each season, with the geometric mean being slightly dominant. Because using the same group of predictors in each season simplifies the operational implementation, the geometric mean might be preferable.

Table 9.6. The mean absolute value of the correlation between the best predictor of uncertainty and Θ^2, across eight lead times, for each group of candidate "combined" predictors from the two ensembles. The standard deviation of the maximum absolute value of correlation is in parentheses.

	CMC	NCEP	Arithmetic Mean	Geometric Mean	Minimum	Maximum
Warm Season	0.1494 (0.0675)	**0.2183** (0.0899)	0.2123 (0.0636)	0.2137 (0.0727)	0.2016 (0.0892)	0.1973 (0.0548)
Cool Season	0.1683 (0.0962)	0.1745 (0.0630)	0.1878 (0.0780)	**0.1881** (0.0771)	0.1694 (0.0671)	0.1820 (0.0720)

Table 9.7 shows the top two combined predictors of uncertainty and the correlation of each with Θ^2 for each lead time and season, with the affine combination of means as the predictor of central tendency. The predictors in the warm season come from the NCEP ensemble alone, while the predictors in the cool season are the geometric mean of predictors from the CMC and NCEP ensembles. These predictors are more informative than the best predictor of uncertainty from each ensemble (with the ensemble mean as the predictor of central tendency) in nine of 16 cases. The

Table 9.7. Top two new predictors of uncertainty for each lead time and season and the correlation of each with Θ^2, with an affine combination of ensemble means as the predictor of central tendency. In the warm season, each predictor comes from the NCEP ensemble alone; in the cool season, each predictor is the geometric mean of the same predictor from each ensemble. A correlation is underlined if its absolute value exceeds the absolute value of the correlation between the best predictor of uncertainty from each of the two ensembles.

Season	Lead Time	Best		Second Best	
		T	Cor	T	Cor
Warm	12 h	$t_{0.9}$	0.1114	d	0.1056
	60 h	r	**0.3355**	d	0.3082
	108 h	$b_1 r + b_2(ck-1)$	0.2617	r	0.2616
	156 h	$t_{0.9}$	0.1061	r	0.1038
	204 h	$t_{0.9}$	0.3273	$b_1 r + b_2(ck-1)$	**0.3269**
	252 h	$t_{0.5}$	0.2354	d	0.2305
	300 h	$r/(ck-1)$	0.2199	$t_{0.7}$	0.1636
	348 h	$b_1 r + b_2(ck-1)$	0.1490	r	0.1490
Average of Absolute Value			0.2183		0.2062
Cool	12 h	$b_1 r + b_2(ck-1)$	0.2462	r	0.2392
	60 h	$b_1 r + b_2(ck-1)$	0.2372	r	0.2354
	108 h	$t_{0.8}$	**0.2812**	$t_{0.9}$	**0.2720**
	156 h	$t_{0.9}$	0.1934	d	0.1576
	204 h	r	-0.0515	$t_{0.8}$	-0.0472
	252 h	$b_1 r + b_2(ck-1)$	0.2323	$r(ck-1)$	0.2121
	300 h	$ck-1$	0.1411	$t_{0.6}$	-0.1336
	348 h	$b_1 r + b_2(ck-1)$	0.1221	$ck-1$	0.1210
Average of Absolute Value			0.1881		0.1773

linear combination of combined range and combined coefficient of kurtosis (shifted to the origin) is most frequently the best predictor of uncertainty (six of 16 cases). Other predictors that are most frequently in the top two include the combined ensemble range and the combined 90% central credible interval; however, these combined predictors are not consistently more informative than predictors of uncertainty from either ensemble processed separately. Still, for sake of operational simplicity, a global user may prefer to limit the predictors of uncertainty used in the BPE to only the combined ensemble range and combined 90% central credible interval.

9.3.3 Comparison of Predictors

To compare predictors of uncertainty from the two ensembles, using predictor of central tendency X in the BPE, Table 9.8 shows the correlation with Θ^2 of the best grand predictor of

Table 9.8. The correlation with Θ^2 of the best grand predictor of uncertainty, the best combined predictor of uncertainty, and the best predictor of uncertainty from each ensemble processed separately, for each lead time and season. The correlation with the highest absolute value is boldfaced for each lead time and season. Between the grand predictor of uncertainty and combined predictor of uncertainty, the more highly correlated is underlined.

	Warm Season			
Lead Time	Grand	Combined	CMC	NCEP
12 h	<u>0.2429</u>	0.1114	**0.2574**	0.1919
60 h	0.3020	<u>0.3355</u>	**0.3408**	0.2931
108 h	**<u>0.3155</u>**	0.2617	0.1969	0.2545
156 h	**<u>0.1624</u>**	0.1061	0.0509	0.0707
204 h	0.2772	**<u>0.3273</u>**	0.1350	0.2628
252 h	**<u>0.2762</u>**	0.2354	0.1577	0.1969
300 h	0.1598	**<u>0.2199</u>**	0.0708	0.1051
348 h	**<u>0.2060</u>**	0.1490	0.0437	0.1559

	Cool Season			
Lead Time	Grand	Combined	CMC	NCEP
12 h	-0.1270	**<u>0.2462</u>**	0.1639	0.2274
60 h	**<u>0.2872</u>**	0.2372	0.2417	0.1085
108 h	0.2104	**<u>0.2812</u>**	0.2503	0.1888
156 h	0.1613	**<u>0.1934</u>**	0.1321	0.1691
204 h	**<u>0.1305</u>**	-0.0515	-0.0806	0.0746
252 h	0.1226	<u>0.2323</u>	-0.2195	**0.3226**
300 h	0.1233	<u>0.1411</u>	-0.1059	**0.1532**
348 h	<u>0.1682</u>	0.1221	0.1592	**0.2034**

uncertainty (with X equal to the grand mean), the best combined predictor of uncertainty (with X equal to the affine combination of ensemble means), and the best predictor of uncertainty from each ensemble processed separately (with X equal to the ensemble mean.) In 11 of 16 cases, a predictor of uncertainty from the two ensembles (either a grand statistic or combined statistic) is more informative than the predictor of uncertainty from each ensemble processed separately. This suggests that more often than not, the CMC ensemble and NCEP ensemble together provide more information about uncertainty than does either ensemble separately. However, the grand predictor of uncertainty is more informative than the combined predictor of uncertainty in only eight of 16 cases. Therefore, there is no clear choice for the best predictor of uncertainty; not only does the best predictor of uncertainty among the four columns vary with lead time and season, but the best predictor of uncertainty *within* each column also varies with lead time and season. Still, it is recommended to use the combined ensembles (as opposed to either ensemble separately) to construct predictors for the BPE; the gain in informativeness using a predictor of central tendency from the combined ensembles appears to generally outweigh any loss in informativeness using a combined predictor of uncertainty.

9.4 Summary and Recommendations

This chapter evaluates several hypothesized predictors of central tendency and uncertainty from the fusion of the CMC ensemble and NCEP ensemble. Since the informativeness of the predictor of uncertainty depends implicitly on the predictor of central tendency, it is important that the two predictors be considered together for operational use. Analyses are performed of the predictors constructed from the two ensembles combined, as well as the predictors obtained by combining the predictors from each ensemble processed separately. Based on these analyses, the following conclusions and recommendations are made:

1. The optimal combination of members from the two ensembles is the most informative predictor of central tendency; however, the membership of this combination varies with lead time and season.

2. An affine combination of ensemble means, when valid, is always more informative than the grand mean, CMC ensemble mean, and NCEP ensemble mean. When such a combination is invalid, the NCEP ensemble mean is the most informative single predictor of central tendency.

3. Most frequently, with the grand mean as the predictor of central tendency, the best predictor of uncertainty is the quotient of ensemble range and shifted coefficient of kurtosis.

4. Most frequently, with the affine combination of means as the predictor of central tendency, the best predictor of uncertainty is the linear combination of range and shifted coefficient of kurtosis.

5. When the affine combination of means is the predictor of central tendency, the best predictor of uncertainty, on average, is (i) a predictor from the NCEP ensemble alone for the warm season, and (ii) the geometric mean of equivalent predictors of uncertainty, one from the NCEP ensemble and one from the CMC ensemble. However, for sake of operational simplicity, the geometric mean may be the preferred aggregation method in both seasons.

6. A predictor of uncertainty from the two ensembles (either a grand statistic or combined statistic) is more informative than the predictor of uncertainty from each ensemble processed separately about 70% of the time.

7. In terms of correlation with the squared residual variate Θ^2, there is no clear choice between a grand predictor of uncertainty (with the grand mean as the predictor of central tendency), and a combined predictor of uncertainty (with an affine combination of means as the predictor of central tendency).

8. The best predictor of uncertainty from the fused ensembles varies with lead time and season, with no apparent trend.

9. In the BPE, a global user is advised to use an affine combination of ensemble means as the predictor of central tendency when valid (and the NCEP ensemble mean otherwise), and to limit the predictor of uncertainty to either the geometric mean of ensemble ranges or the geometric mean of 90% central credible intervals.

10. In the BPE, a sophisticated user is advised to use the optimal combination of members from the two ensembles as the predictor of central tendency, and to use the best combined predictor of uncertainty—using the geometric mean for aggregation—for each lead time and season.

10. CONCLUSIONS

This research applies and expands upon the methodology of Lee (2010) in order to compare the statistical properties of two ensemble forecasting systems for meteorological variates, and to fuse the two ensembles via a Bayesian processor. This methodology includes the following steps: hypothesizing candidate predictors of central tendency and uncertainty, selecting a likelihood model that captures the stochastic dependence between the predictand and each candidate predictor of central tendency, estimating likelihood parameters, evaluating each predictor of central tendency using the informativeness score, evaluating each predictor of uncertainty according to its correlation with the squared residuals from the likelihood model, and applying the Gaussian-Gamma model for the Bayesian processor of ensemble (BPE) to obtain a posterior distribution function. Applying this methodology to an ensemble yields (i) an approximately sufficient statistic (x, t) to replace vector $\boldsymbol{y}$ of ensemble members in the BPE, and (ii) recommendations for the efficient processing of ensemble forecasts into well-calibrated probability density functions. The methodology is expanded upon by comparing the analyses of two ensemble forecasting systems, and then utilizing the key statistical differences between the ensembles to explore ways of fusing their predictors of central tendency and uncertainty.

As in Lee (2010), this research applies to ensemble forecasts of 2-meter surface temperature at the Savannah, Georgia, ground station, made at 00 UTC for eight lead times: 12 h, 60 h, 108 h, 156 h, 204 h, 252 h, 300 h, and 348 h. While Lee conducts an analysis of a 20-member ensemble from the National Centers for Environmental Prediction (NCEP), this thesis analyzes the 20-member ensemble from the Canadian Meteorological Centre (CMC). Results from the analysis of each ensemble are then compared. Finally, the possibility of fusing the two ensembles is explored.

The sample from the CMC contains 599 complete days of ensemble forecasts made at 00 UTC. Ensemble forecasts made at 12 UTC with lead time 0 h served as realizations of predictand W, the true surface temperature at the time of forecast verification. Various summary statistics are hypothesized as predictors of the central tendency of W, and of uncertainty about W. An empirical

analysis of the CMC ensemble supports the following three findings: (i) The ensemble mean is consistently among the most informative predictors of central tendency, and can be enhanced by the ensemble minimum or maximum at short lead times. (ii) An optimal combination of ensemble members is more informative than the ensemble mean; however, its composition varies with lead time and season. (iii) The best predictor of uncertainty varies across lead times and seasons, but the width of the 90% central credible interval and an amalgam of ensemble range and coefficient of kurtosis are frequently among the best, while the ensemble standard deviation is not.

A comparison of the CMC and NCEP analyses supports the following two findings: (i) The NCEP ensemble yields generally more informative predictors of central tendency than does the CMC ensemble. (ii) The ensemble that yields a better predictor of uncertainty varies with lead time and season, with no apparent trend.

An empirical analysis of a combination of the CMC and NCEP ensembles supports the following four findings: (i) The most informative single predictor of central tendency is the affine combination of ensemble means, when valid, and the NCEP ensemble mean otherwise. (ii) An optimal combination of ensemble members is always more informative than the affine combination of means; however, its composition varies with lead time and season. (iii) The best predictor of uncertainty varies across lead times and seasons with no apparent trend, and is more informative than the predictor of uncertainty from each ensemble processed separately about 70% of the time. (iv) The best predictor of uncertainty is either the arithmetic mean or the geometric mean of the same predictors from each ensemble, although all predictors of uncertainty are generally poor.

Recommendations for operational forecasting centers include the following: (i) A global user of the CMC ensemble is advised to use the ensemble mean as the predictor of central tendency, and the 90% central credible interval as the predictor of uncertainty. (ii) A sophisticated user of the CMC ensemble is advised to use an optimal combination of ensemble members as the predictor of central tendency, and to use the best predictor of uncertainty for each lead time and season. (iii) A global user of the CMC and NCEP ensembles is advised to use as the predictor of

central tendency either the affine combination of means or the NCEP ensemble mean (as appropriate), and to limit the predictor of uncertainty to either the geometric mean of ensemble ranges or the geometric mean of 90% central credible intervals. (iv) A sophisticated user of the two ensembles is advised to use an optimal combination of ensemble members as the predictor of central tendency, and to use the best combined predictor of uncertainty for each lead time and season.

Recommendations for future research on the NCEP and CMC ensemble forecasts include the following: (i) The Gaussian-Gamma model should be validated for processing the sufficient statistics of the fused ensemble forecasts. (ii) New summary statistics of each ensemble, and methods of combining them, could be invented and tested as predictors of central tendency and uncertainty. (iii) Tradeoffs between the informativeness of predictors of central tendency and predictors of uncertainty should be explored, theoretically and empirically, to establish a basis for the joint selection of the two kinds of predictors from a combination of two ensemble forecasts.

REFERENCES

Bulmer, M. G. Descriptive properties of distributions. *Principles of Statistics,* Dover, 45-67, 1979.

Charron, M., Pellerin, G., Spacek, L., Houtekamer, P.L., Gagnon, N., Mitchell, H.L, and Michelin, L. Toward random sampling of model error in the Canadian Ensemble Prediction System, *Monthly Weather Review*, 138:1877-1901, 2010.

Cloke, H.L., Pappenberger, F. Ensemble flood forecasting: a review. *Journal of Hydrology*, 375: 613–626, 2009.

DeGroot, M. H. Section 9.1, Sufficient Statistics. *Optimal Statistical Decisions*, McGraw-Hill, 155-159, 1970.

Gneiting, T., Raftery, A.E., Westveld III, A.H., and Goldman, T. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, 133:1098-1118, 2004.

Kolczynski, W.C., Stauffer, D.R., and Haupt, S.E. Investigation of ensemble variance as a measure of true forecast variance. *Monthly Weather Review*, 139:3954-3963, 2011.

Krzysztofowicz, R. Why should a forecaster and a decision maker use Bayes theorem. *Water Resources Research*, 19(2):327-336, 1983.

Krzysztofowicz, R. Bayesian correlation score: a utilitarian measure of forecast skill. *Monthly Weather Review*, 120:208-219, 1992.

Krzysztofowicz, R. The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249:2-9, 2001.

Krzysztofowicz, R. and Evans, W.B. Probabilistic forecasts from the National Digital Forecast Database. *Weather and Forecasting*, 23:270-289, 2007.

Krzysztofowicz, R. Bayesian processor of ensemble (BPE): A Gaussian-Gamma model. Presented at the 20th Conference on Probability and Statistics in the Atmospheric Sciences, American Meteorological Society, Atlanta, Georgia, 17-21 January 2010.

Lee, N.Y. Sufficient statistics of ensemble for probabilistic weather forecasting (Unpublished master's thesis). University of Virginia, Charlottesville, VA, 2010.

Sivillo, J.K., Ahlquist, J.E., and Toth, Z. An ensemble forecasting primer. *Weather and Forecasting*, 12:809-818, 1997.

Toth, Z. and Kalnay, E. Ensemble forecasting at NMC: the generation of perturbations. *Bulletin of the American Meteorological Society*, 74:2317-2330, 1993.

Toth, Z., Zhu, Y., and Marchok, T. The use of ensembles to identify forecasts with small and large uncertainty. *Weather and Forecasting*, 16:463-477, 2001.

Williams, E. J. The comparison of regression variables. *Journal of the Royal Statistical Society: Series B,* 21:396-399, 1959.