Reply

CHRISTOPH C. RAIBLE, GEORG BISCHOF, KLAUS FRAEDRICH, AND EDILBERT KIRK

Meteorologisches Institut, Universität Hamburg, Hamburg, Germany

(Manuscript received 24 May 2001, in final form 2 August 2001)

1. Introduction

Short-term weather predictions provide future information on a significant set of local weather variables, which include extreme events and forecast errors, and require, a posteriori, comprehensive verification. Linear empirical forecast models, which Leslie and Speer (2001, hereafter LS01) refer to, play, in general, the minor role of supporting the highly advanced numerical weather prediction (NWP) systems.

2. Data and globally linear reference forecasts

Like most linear empirical forecast models, the Raible et al. (1999, hereafter RB99) schemes deliver forecasts for a significant subset of local weather variables: the near-surface air temperature and the probability of precipitation (PoP). Both predictors serve as an example to demonstrate the applicability of empirical forecast models. Input data are those that are immediately available to the public and can be extracted automatically from the Internet; higher-resolution data required to forecast extremes etc. are, in general, *not* freely available. That is, NWP forecasts for energy suppliers or other enterprises only cannot simply be provided due to national weather service restrictions. However, this limited data supply also poses a challenge for student forecasters interested in empirical forecasting, forecast verification, and local weather. Therefore, our forecasts focus on a region dominated by weather systems near the tail end of the North Atlantic storm track. Only if, as an added value, suitable NWP output is at hand, which, in the continental European environment, is provided by national weather services, an optimal combination of both forecast schemes is possible. In summarizing, the RB99 forecast system is not guided by the same goal as LS01, whose aim is to deliver forecasts for public or private enterprise by utilizing the latest in NWP development with the highest space and time resolution possible. Nevertheless, it is possible to implement the RB99 schemes to high-resolution data in space and time, if this information is available.

Many empirical forecast schemes are a variant of the generalized equivalent Markov system (Miller 1984), which is globally linear in phase space and has been advocated by Miller since 1964. Applying those or similar techniques, RB99 provides forecasts satisfying some of the short-term prediction goals.

- A local weather variable subset: Temperature anomalies from the ensemble-averaged diurnal cycle capture the Markovian contributions to local weather variability and lead to a surprisingly good model performance. Using the observed low cloud states leads to considerable improvements of PoP forecasts. An extension to other variables is possible.
- A comprehensive verification scheme for these forecasts (Murphy and Winkler 1987; Murphy et al. 1989) is employed, which goes beyond a rudimentary verification often used in the context of forecast performance. Besides the classical measures of accuracy, like the root-mean-square error and the half Brier score, which directly compare forecasts and observations, the verification scheme in RB99 includes joint and conditional distributions leading to reliability and sharpness diagrams of the forecasts versus observations. Skill scores using persistence and climatology as benchmarks are classical standards of comparison. It enables current systems to be easily compared with forecast experiments of former (reviewed) studies.
- A model building process was established before selecting the globally linear statistical forecast scheme: nonlinear empirical models were tested prior to the RB99 analysis. However, application of such techniques, for example, Euclidean metric analogs and a publicly available feed-forward neural network, supported the conjecture that empirical forecasts by multivariate autoregression type (or globally linear) models are not very easily beaten. More precisely, under the same data constraints, the analog approach was not able to produce the level of accuracy achieved by the linear Markovian methods, whereas the neural net led to comparable results (http://dome.dkrz.de/simon).

Corresponding author address: Klaus Fraedrich, Meteorologisches Institut, Universität Hamburg, Bundesstr. 55, D-20146, Hamburg, Germany.
E-mail: fraedrich@dkrz.de

TABLE 1. The rms error of the 24-h "maximum" and "minimum" temperature (1200 and 0600 UTC) forecasts for Hamburg (1994–95).

Model	Rms error (°C)	
	Max	Min
NWP (direct output)	3.34	2.69
Empirical model (RB-99)	2.95	2.76
Combination	2.74	2.38

Even if the outcome of the model building process had been in favor of the neural net, the globally linear system would have been selected as the empirical reference forecast system for future model development and comparison.

- A field representation based on continuous curvature splines in tension has been adopted from geophysical applications (Smith and Wessel 1990, publicly available software). Mapping temperature or PoP forecasts with external parameters like geography and/or vegetation fields by kriging (Hudson and Wackernagel 1994) was challenging but, at the time, not considered to fit the goals of the RB99 paper.

In summarizing, the aim of RB99 was to provide an empirical reference forecast system for both temperature and PoP in the continental European area where, to the authors knowledge, it has not been established; all input data to run the model, the model building procedures, and the field representations are freely obtained and easily reproducible. Other purposes were to employ comprehensive verification measures developed in the late 1980s and, in addition, to optimally combine numerical and empirical products. The empirical model is built on Miller (1964) using a Markov chain and regression mix not unlike Leslie and Miller (1984, 1985) providing only PoP forecasts for Australian capitals. Finally, we wish to demonstrate in the remaining sections further applications of the RB99 system and recently published developments in statistical predictions.

3. Nonlinear deterministic and linear statistical forecasts

The RB99 model is employed in 24-h forecasts of the minimum and maximum temperature, which are represented by the 0600 and 1200 UTC temperatures, and combined with predictions of the German Weather Service Europa Modell (NWP, developed before 1993 with about 55-km resolution). The results are shown in Table 1 for the Hamburg weather station during the independent verification period 1994–95. In addition, two case studies for the winter and summer season are presented to test the ability to forecast *extreme* weather situations and the diurnal cycle in summer (Figs. 1a and 1b). The following results are noted:

- The rms error of the optimal combination shows an improvement of the NWP minimum (maximum) temperature forecasts by 12% (18%) for Hamburg. Though significant (above 95%), this improvement is relatively small, which may be due to a variety of possible causes: (i) the local model predictability may be substantially small due to the location of the weather station at the tail end of the North Atlantic storm track; (ii) 18–24-h forecasts reach the limits for empirical systems exploiting the weather's Markovian properties; note that 12-h combination forecasts for all temperatures improve the NWP by 34%.
- The direct model output of NWP (dashed) and the correction achieved by the combination (dashed–dotted) is shown for a week of 6-hourly 24-h forecasts in summer and in winter. The winter case is an event with extreme low temperatures for 4 days (−2.2°C is the long-term mean minimum temperature in January). NWP predicts too cold temperatures, the linear method (dotted) predicts warmer than observed (full line), and the combination is able to reproduce this cold spell sufficiently well. The summer case shows a pronounced diurnal cycle.

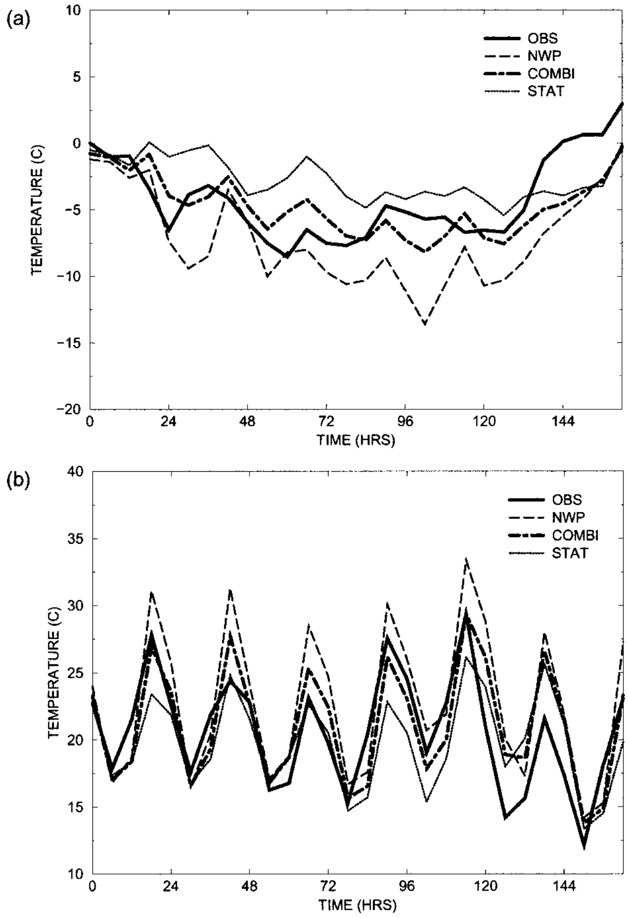

FIG. 1. The 24-h temperature forecasts (in °C) at Hamburg (Germany): (a) a week in winter (1–8 Jan 1995) and (b) in summer (1–8 Jul 1995); NWP model (dashed), RB-99 statistical scheme (STAT, dotted), combination (COMB, dashed–dotted), and observations (OBS, full line).

4. Statistical forecasting: Nonlinear methods and ensemble predictions

Empirical forecasts by statistical and statistical–dynamical methods can be improved in their ability to predict extreme events and forecast errors, which requires application of (i) nonlinear and (ii) ensemble forecasts. Therefore, jointly with the preparation and testing of the RB99 reference predictions, Fraedrich and Rückert (1998) developed a new *nonlinear* forecast scheme for analogs with a self-adapting metric, building on the generically meteorological tradition of Euclidean metric analog forecasts (and not on software available for neural networks, simplex methods, or radial basis functions). This new nonlinear empirical scheme was tested for low-order dynamical systems (Fraedrich and Rückert 1998) and applied, with some success, as an optimized *ensemble* forecast scheme for hurricane tracks (Sievers et al. 2000).

Acknowledgments. The RB99 forecast system was prepared, implemented, and verified as part of a 1996 student workshop at Hamburg University leading to the master theses of the two lead authors; Simon Blessing's participation and engagement in the feed-forward neural network, the construction, testing, and continuing online forecasts are appreciated.

REFERENCES

Fraedrich, K., and B. Rückert, 1998: Metric adaption for analog forecasting. *Physica A,* **253,** 379–393.

Hudson, G., and H. Wackernagel, 1994: Mapping temperature using kriging with external drift: Theory and example from Scotland. *Int. J. Climatol.,* **14,** 77–91.

Leslie, L. M., and M. S. Speer, 2001: Comments on ''Statistical single-station short-term forecasting of temperature and probability of precipitation: Area interpolation and NWP combination.'' *Wea. Forecasting,* **16,** 765–767.

Miller, A. J., and L. M. Leslie, 1984: Short-term single-station forecasting of precipitation. *Mon. Wea. Rev.,* **112,** 1198–1205.

——, and ——, 1985: Short-term single-station probability of precipitation forecasting using linear and logistic models. *Contrib. Atmos. Phys.,* **58,** 517–527.

Miller, R. G., 1964: Regression estimation of event probabilities. Travellers Research Center Tech. Rep. 7411–121, 153 pp.

——, 1984: *GEM: A Statistical Weather Forecasting Procedure.* Short-and Medium Range Weather Prediction Research (PSMP) Publication Series, World Meteorological Organization, Vol. 10, 102 pp.

Murphy, A. H., and R. L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.,* **115,** 1330–1338.

——, B. G. Brown, and Y.-S. Chen, 1989: Diagnostic verification of temperature forecasts. *Wea. Forecasting,* **4,** 485–501.

Raible, C. C., G. Bischof, K. Fraedrich, and E. Kirk, 1999: Statistical single-station short-term forecasting of temperature and probability of precipitation: Area interpolation and NWP combination. *Wea. Forecasting,* **14,** 203–214.

Sievers, O., K. Fraedrich, and C. C. Raible, 2000: Self-adapting analog ensemble predictions of tropical cyclone tracks. *Wea. Forecasting,* **15,** 623–629.

Smith, W. H. F., and P. Wessel, 1990: Gridding with continuous curvature splines in tension. *Geophysics,* **55,** 293–305.