P2.5

ENCOURAGING STUDENT INVOLVEMENT THROUGH USE OF BASIC PROBABILITY FORECASTING GAMES

Steven B. Newman
Central Connecticut State University, New Britain, Connecticut

1. INTRODUCTION

Students in introductory Meteorology courses, or in Earth Science curricula that include a weather unit, all seem to want to learn one thing before all others—how to forecast the weather. It's a skill that impresses their parents and friends, and makes them feel "important." They are usually disappointed when told that forecasting is a skill that can only be learned after one has a basic understanding of the atmosphere, and the factors that control the daily weather. Thus in most introductory level university courses, the basics of forecasting are taught later in the course, (and not with any real depth) and actual forecasting usually involves a laboratory exercise in which a scenario (often from a real weather event of the past) is presented, and students are asked to forecast the next 24 hours of weather at a given city or cities. The results are often disappointing for both students and instructors, as they really don't get a feel for what operational forecasting involves.

In upper level meteorology courses, daily forecasting becomes part of the curriculum. Many students look forward to spending time in the Weather Center pouring over National Weather Service (NWS) charts and endless printouts of numerical model data, preparing a daily forecast. In order to assure that students actually participate in forecasting every day, some schools have made the forecasting exercise into a "competition," or "game," in which the students compete against one another, as well as faculty members that wish to participate. Often, there is some sort of incentive to prepare the most skillful forecasts over the course of an entire semester (usually a free lunch provided by the supervising faculty member) for any students that outperform the faculty member.

* *Corresponding author address:* Steven B. Newman, Department of Physics and Earth Sciences, Central Connecticut State University, New Britain, Connecticut 06050; e-mail: newman@mail.ccsu.edu

2. PROBABILITY FORECASTING

Local weather forecasts prepared for public dissemination must include certain elements as proscribed by the National Weather Service (Chaston, 1992). These include, temperature, sky condition, probability of precipitation, wind direction and speed, etc. Without an in-depth understanding of weather elements and how they are affected by upper level features, students in introductory level courses cannot prepare forecasts of this nature. However, if the basic elements of forecasting are reduced to the probability of certain events either happening or not happening, then students can be taught at an early stage of the course, how to interpret the guidance from the National Weather Service to make such forecasts, in many cases with some degree of success. The competitive aspect of the "game" encourages student involvement, and also serves to form a much closer bond between student and faculty member, as there is considerable interaction between them outside the formal classroom.

The easiest (to teach) form of probability forecasting uses the method first developed at the Massachusetts Institute of Technology (MIT) by Sanders (1967), and used for many years at the University at Albany (Bosart, 1983). This type of forecast involves predicting the probability of two events happening at an official NWS location during each of the next four 24-hour periods, respectively. The events being predicted are a) the likelihood that the minimum temperature during each 24-hour period will be below the normal minimum (as defined by the NWS); and b) the likelihood that there will be at least .01 inch of measurable precipitation during each 24-hour period. The normal high and low temperatures for most NWS stations can be found on the websites of those stations. Official verifications are from observations made at the NWS station. Scoring is done using the scoring method described by Brier (1950).

3. GAME METHODOLOGY AND SCORING

Each day, students are asked to forecast for the next four 24-hour periods. The forecast is given as the probability (chances in 10) that the minimum temperature for each 24-hour period will be below the normal minimum for that day, and that there will be measurable (i.e., ≥ 0.01 inches) precipitation for each 24-hour period. If the forecaster believes that each event is certain to occur, the forecast would be a "10." If the forecaster believes that there is no possibility of the event occurring, then the forecast would be a "0." Forecasts between "0" and "10" are written such that they reflect the degree of uncertainty of the forecaster. Verification for each day is either a "10" if the event occurs, or a "0" if the event does not occur. Forecasters accumulate "error points" equal to the square of the difference between the forecast number and the verification. A forecaster may receive up to 100 error points for each forecast. The fewer error points accumulated, the better the forecast. A CONSENSUS sheet and a CLIMATOLOGY sheet are kept for each week, and at week's end, each forecaster's sheet is scored relative to CONSENSUS and CLIMATOLOGY. CONSENSUS is merely the average of all forecasts submitted each day. CLIMATOLOGY (or CLIMO) is a persistence-based forecast which is unique to any specific location and must be determined by studying many years of temperature and precipitation data. Newman (1991) has developed such a CLIMATOLOGY for both interior and coastal Southern New England locations, and has described how such a climatology can be compiled for any location for which a sufficient database exists.

An example of a scored sheet for one week is given in Figure 1. Looking at the scoring summary near the bottom of the sheet, we see that during this particular week in December, 2001, the forecaster was 22.8% more skillful than CONSENSUS for temperature, but 3.3% less skillful for precipitation. He was 28.1% more skillful than CLIMO for temperature, but 15.9% less skillful for precipitation. For the entire semester, up to this point, the forecaster was 46.8% more skillful than CLIMO at forecasting the minimum temperature, but 13.6% less skillful at forecasting the probability of precipitation. The last two numbers on the sheet indicate the combined skill of the forecaster, for both temperature and precipitation, vs. CLIMO (on the left side) and CONSENSUS (on the right side). This forecaster shows an overall skill of 32.2% vs. CLIMO, and –3.9% vs. CONSENSUS (i.e., the CONSENSUS forecast is, on average, 3.9% more skillful than the forecaster's). Figures 2 and 3 are the

Forecast for 00Z to 00Z Period			Prob. of Min T < Normal Min					Prob. of Precip ≥ 0.01"			
Verify	Start		CLIMO	NORMS	P_F	P_O	$(E_{O-F})^2$	$(E_{O-F})^2$	P_O	P_F	CLIMO
TUE 4	MON	24		25	0	0	0	0	0	0	
WED 5	TUE	24		24	0	0	0	64/64	10	2	
	MON	48			0	0	0	49/113	10	3	
THU 6	WED	24	3		0	0	0	16/129	0	4	3
	TUE	48			0	0	0	36/165	0	6	
	MON	72			0	0	0	36/201	0	6	
FRI 7	THU	24			0	0	0	16/217	0	4	
	WED	48	4		0	0	0	36/253	0	6	4
	TUE	72			0	0	0	36/289	0	6	
	MON	96			3	0	9/9	16/305	0	4	
SAT 8	FRI	24		23	0	0	9	49/354	10	3	
	THU	48			2	0	4/13	64/418	10	2	
	WED	72	4		1	0	1/14	49/467	10	3	4
	TUE	96			3	0	9/23	49/516	10	3	
SUN 9	FRI	48			1	10	81/104	1/517	10	9	
	THU	72			4	10	36/140	36/553	10	4	
	WED	96	4		3	10	49/189	36/589	10	4	4
MON 10	FRI	72		22	4	10	36/225	4/593	0	2	
	THU	96			4	10	36/261	4/597	0	2	
TUE 11	FRI	96			4	0	16/277	4/601	0	2	
WEEKLY TOTALS					Cl 385	Co 359	F 277	F 601	Co 582	Cl 521	
WEEKLY REDUCTIONS						(Cl-F)/Cl +28.1	(Co-F)/Co +22.8	(Co-F)/Co -3.3	(Cl-F)/Cl -15.4		
PREVIOUS WEEK TOTALS					3670	1955	1879	2278	1952	2911	
NEW CUM. TOTALS					4055	2314	2156	2879	2534	3432	
CUMULATIVE REDUCTIONS						(Cl-F)/Cl +46.8	(Co-F)/Co +6.8	(Co-F)/Co -13.6	(Cl-F)/Cl +16.1		
COMBINED CUMULATIVE REDUCTIONS							(Cl-F)/Cl +32.7	(Co-F)/Co -3.9			
NAME: MESSIER								WEEK OF: 03 DEC 01			

Fig. 1 Example of a Scored Weekly Forecast Sheet

CONSENSUS and CLIMO sheets for the same week.

$\bar{z}_{500}$ = 551 $\bar{T}_{850}$ = -2.0 $\overline{\Delta z}_{5,0}$ = 538

Forecast for 00Z to 00Z Period: Verify	Start		CLIMO	NORMS	Prob. of Min T < Normal Min: P_F	P_O	$(E_{O-F})^2$	Prob. of Precip ≥ 0.01": $(E_{O-F})^2$	P_O	P_F	CLIMO
TUE 4	MON	24		25	0 (0010)	0	0/0	0/0	0	0 (0010)	
WED 5	TUE	24		24	0 (000000001)	0	0/0	81/81	10	1 (21110010)	
	MON	48		↓	0 (1000)	0	0/0	64/145	10	2 (1213)	
THU 6	WED	24			0 (00000)	0	0/0	4/149	0	2 (04031)	
	TUE	48			0 (000000001)	0	0/0	25/174	0	5 (64762571)	
	MON	72			0+ (1100)	0	0/0	16/190	0	4− (1436)	
FRI 7	THU	24			0 (000000011)	0	0/0	16/207	0	4 (43633434)	
	WED	48			0 (00000)	0	0/0	16/223	0	4 (36236)	
	TUE	72			1 (01101102)	0	1/1	25/248	0	3 (66249334)	
	MON	96		↓	2 (2213)	0	4/5	9/257	0	3 (2334)	
SAT 8	FRI	24		23	0 (00100)	0	0/5	25/282	10	5 (73457)	
	THU	48			1 (00001212)	0	1/6	49/331	10	3 (91222244)	
	WED	72			1 (21011)	0	1/7	64/395	10	2 (13232)	
	TUE	96			2 (31212113)	0	4/11	64/459	10	2 (32222322)	
SUN 9	FRI	48			1 (41100)	10	81/92	1/460	10	9 (69909)	
	THU	72			2 (11213413)	10	64/156	36/496	10	4 (32444463)	
	WED	96		↓	1 (03031)	10	81/237	64/560	10	2 (04132)	
MON 10	FRI	72		22	3 (24614)	10	49/286	9/569	0	3 (32233)	
	THU	96			2 (12113413)	10	64/350	9/578	0	3 (42313233)	
TUE 11	FRI	96		↓	3 (24421)	0	9/359	4/582	0	2 (32223)	
WEEKLY TOTALS					Cl 385	Co 359	F 359	F 582	Co 582	Cl 521	
WEEKLY REDUCTIONS						(Cl-F)/Cl +6.8	(Co-F)/Co —	(Co-F)/Co —	(Cl-F)/Cl -11.7		
PREVIOUS WEEK TOTALS					3984	2058	2058	2017	2017	3023	
NEW CUM. TOTALS					4369	2417	2417	2599	2599	3544	
CUMULATIVE REDUCTIONS						(Cl-F)/Cl +44.7	(Co-F)/Co —	(Co-F)/Co —	(Cl-F)/Cl +26.7		
COMBINED CUMULATIVE REDUCTIONS							(Cl-F)/Cl +36.6	(Co-F)/Co —			
NAME: CONSENSUS								WEEK OF: 3 Dec '01			

Fig. 2 CONSENSUS Forecast Sheet for week of 3 December, 2001

In Figure 2, the individual forecasts for each 24-hour period are listed below the line in each box, and the average (or CONSENSUS) forecast, rounded to the nearest whole number, is written above the line. In case of an average being exactly between two numbers, rounding is, by convention, always done to the even number. During this particular week, CONSENSUS showed only a 6.8% improvement over the CLIMATOLOGY forecast on temperature, and was 11.7% less skillful than CLIMATOLOGY on precipitation. However, overall, CONSENSUS had shown a 36.6% improvement on temperature, and a 26.7% improvement on precipitation over CLIMATOLOGY over the course of the semester. In practice, long-term improvements of more than 30-35% are considered to be somewhat skillful forecasts.

Figure 3 shows the CLIMO forecast sheet for the week of 3 December 2001. Each day, CLIMO is entered based on the verifications of the previous day (thus CLIMO is a four-day long, persistence-type forecast, based on what happened during the past 24 hours).

You can see that CLIMO usually is well behind CONSENSUS (as well as most individual forecasters) in both temperature and precipitation forecasting. This is because persistence forecasting becomes less accurate, the farther out ahead such forecasts are made, especially in changeable weather regions such as the Northeast United States. Only a very poor forecaster will fail to improve upon CLIMATOLOGY over an extended period of time.

Forecast for 00Z to 00Z Period: Verify	Start		CLIMO	NORMS	Prob. of Min T < Normal Min: P_F	P_O	$(E_{O-F})^2$	Prob. of Precip ≥ 0.01": $(E_{O-F})^2$	P_O	P_F	CLIMO
TUE 4	MON	24			3	0	9/9	9/9	0	3	
WED 5	TUE	24			3	0	9/18	49/58	10	3	
	MON	48			4	0	16/34	36/94	10	4	
THU 6	WED	24			3	0	9/43	25/119	0	5	
	TUE	48			4	0	16/59	16/135	0	4	
	MON	72			4	0	16/75	16/151	0	4	
FRI 7	THU	24			3	0	9/84	9/160	0	3	
	WED	48			4	0	16/100	16/176	0	4	
	TUE	72			4	0	16/116	16/192	0	4	
	MON	96			4	0	16/132	16/208	0	4	
SAT 8	FRI	24			3	0	9/141	49/257	10	3	
	THU	48			4	0	16/157	36/293	10	4	
	WED	72			4	0	16/173	36/329	10	4	
	TUE	96			4	0	16/189	36/365	10	4	
SUN 9	FRI	48			4	10	36/225	36/401	10	4	
	THU	72			4	10	36/261	36/437	10	4	
	WED	96			4	10	36/297	36/473	10	4	
MON 10	FRI	72			4	10	36/333	16/489	0	4	
	THU	96			4	10	36/369	16/505	0	4	
TUE 11	FRI	96			4	0	16/385	16/521	0	4	
WEEKLY TOTALS					Cl 385	Co 359	F 385	F 521	Co 582	Cl 521	
WEEKLY REDUCTIONS						(Cl-F)/Cl —	(Co-F)/Co -7.2	(Co-F)/Co +10.5	(Cl-F)/Cl —		
PREVIOUS WEEK TOTALS					3984	2058	3984	3023	2017	3023	
NEW CUM. TOTALS					4369	2417	4369	3544	2599	3544	
CUMULATIVE REDUCTIONS						(Cl-F)/Cl —	(Co-F)/Co -80.8	(Co-F)/Co -36.4	(Cl-F)/Cl —		
COMBINED CUMULATIVE REDUCTIONS							(Cl-F)/Cl —	(Co-F)/Co -57.8			
NAME: CLIMO								WEEK OF: 3 Dec '01			

Fig. 3 CLIMO Forecast Sheet for week of 3 December, 2001

4. ANTICIPATED OUTCOMES

Previous experience in my classes has shown that first-time forecasters do not do well, especially for the first few weeks of the game. Once students get the hang of how the forecasting is done, some will quickly demonstrate a great deal of skill in probability forecasting. Others will consistently lag behind CLIMATOLOGY. Interestingly, most of my group of forecasters during the Fall, 2001 semester did well, including some for whom this was a first-time activity. This may have been due to the weather being both unusually warm and dry during the months of October, November and December. When weather conditions differ drastically from climatological norms, forecasters tend to be able to improve upon CLIMATOLOGY to a much greater degree.

The best part of forecasting this way is that it requires less meteorological background to get started than other forecasting exercises. There are no forecasts of sky condition, winds, type of weather, or timing required here. The students are merely forecasting the probability that the temperature will drop below the normal miminum for the next four nights, and the probability that there will be measurable precipitation during each of the next four 24-hour periods. To help them reach their conclusions, you can provide them with forecast model guidance from the NWS website, along with local forecasts from newspapers, TV or the NWS. In this way, students learn to judge the value of computer model generated forecasts, as well as learning to differentiate between those situations in which the computer guidance does well and those in which it does poorly. It only takes a few class sessions at the beginning of the semester to teach students how to read and interpret the NWS guidance messages to use in their own forecasts. Many students become so wrapped up in the game that the first thing they do upon awakening in the morning is log onto their computer to check the overnight low temperature, or look out the window to see if any puddles have formed on the ground (a sure sign that at least 0.01 inches of rain has fallen).

5. CONCLUSIONS AND OTHER OPTIONS

Student forecasters at Central Connecticut State University have been participating in a probability-based forecasting game for many years. The advantages of such a game over a single weather forecasting lab exercise are threefold. First, students can begin learning the basics of forecasting at the very beginning of the semester, even before they have a thorough knowledge of the dynamics of the atmosphere. Second, the daily involvement in forecasting leads to a greater degree of student interest, which can be maintained over the course of the entire semester. Third, the forecast method itself has been shown to demonstrate real skill among forecasters. As students become reach more advanced levels of meteorological background, it is possible to employ alternative probability forecasting methods that will further enhance forecasting skill. Hamill and Wilks (1995) have developed a categorical probability game that involves forecasting daily maximum and minimum temperatures within a specific range (e.g., today's high will be between 45 and 49 degrees). The scoring method is the "ranked probability score" defined by Epstein (1969) and Murphy (1971). This scoring method rewards forecasts where the observed max/min falls within the forecast range. Obviously, a larger forecast range is more likely to hit the verification, but fewer error points are awarded for a smaller, but still correct range. For precipitation, forecasters must divide their probabilities among six categories of precipitation, ranging from zero to greater than an inch. Scoring is done in a method similar to that for temperature. This alternative game is more suitable to students in an advanced course, or those who have previous forecasting experience and training.

6. REFERENCES

Bosart, L.F., 1983: An update on trends of skill of daily forecasts of temperature and precipitation at the State University of New York at Albany. ***Bull. Amer. Meteor. Soc.,*** **64,** 346-354.

Brier, G.W., 1950: Verification of forecasts expressed in terms of probability. ***Mon. Wea. Rev.,*** **78,** 1-3.

Chaston, P., 1992: ***Public Forecasting***, National Weather Service Training Center, Internal Publication, 40 pp.

Epstein, E.S., 1969: A scoring system for probability forecasts of ranked categories. ***J. Appl. Meteor.,*** **8,** 985-987.

Hamill, T.M., and D.S. Wilks, 1995: A probabilistic forecast contest and the difficulty in assessing short-range forecast uncertainty. ***Wea. Forecasting***, **10**, 620-631.

Murphy, A.H., 1971: A note on the ranked probability score. ***J. Appl. Meteor.,*** **10,** 155-156.

Newman, S.B., 1991: A persistence climatology for interior and coastal Southern New England, ***Weather and Forecasting,*** **6**, 111-116.

Sanders, F., 1967: The verification of probability forecasts. ***J. Appl. Meteor.,*** **6,** 756-761.