On the Relationship between the Quality and Value of Forecasts in the Generalized Cost–Loss Ratio Situation

YIN-SHENG CHEN*, MARTIN EHRENDORFER AND ALLAN H. MURPHY

Department of Atmospheric Sciences, Oregon State University, Corvallis, OR 97331

(Manuscript received in final form 24 December 1986)

ABSTRACT

This paper investigates the relationship between the quality and value of forecasts in the context of a generalized N-action, N-event model of the cost–loss ratio situation. The forecasts of interest are imperfect categorical forecasts, calibrated according to past performance and represented by multidimensional sets of conditional and predictive probabilities. Forecast quality is measured by the ranked probability score (RPS), a natural measure of the accuracy of forecasts in the context of this model. The measure of value is the difference between the expected expense associated with climatological information and the expected expense associated with imperfect forecasts. Thus, climatological and perfect information define lower and upper bounds, respectively, on the quality and value of the imperfect forecasts.

Quality–value relationships are explored in the three-action, three-event situation, using brute force and mathematical programming methods. Numerical results are presented for several specific cases. In all cases, the relationships are described by envelopes of values rather than by single-valued functions, indicating that a range of forecast value is generally associated with a given level of forecast quality (and vice versa). The existence of these envelopes reveals two important deficiencies in scalar (i.e., one-dimensional) measures of forecast quality, such as the RPS, when they are used as surrogates for measures of value: 1) these quality measures generally provide only imprecise estimates of forecast value and 2) increases in forecast quality, as reflected by such measures, may actually be associated with decreases in forecast value.

1. Introduction

The ultimate measure of the performance of forecasting systems is the value of the forecasts to actual and potential users. Primarily because of difficulties encountered in determining the value of forecasts, meteorologists have usually characterized performance in this context in terms of quality rather than value. However, since the relationship between forecast quality and forecast value is inherently nonlinear (e.g., see Katz and Murphy, 1987), the former is generally not an adequate surrogate for the latter. Thus, it is of some importance to investigate such quality–value relationships in a variety of decision-making situations. Knowledge of the relationship between the quality and value of forecasts can be particularly useful when attempting to estimate the incremental benefits associated with potential improvements in forecast quality.

Recent studies of quality–value relationships for weather or climate forecasts have included investigations of the following prototype and real-world decision-making problems: (i) the cost–loss ratio situation (Katz and Murphy, 1987; Murphy et al., 1985); (ii) the fruit-frost problem (Katz et al., 1982); (iii) the fallowing/planting problem (Brown et al., 1986); and (iv) a choice-of-crop problem (Wilks and Murphy, 1986). The nonlinear relationship between forecast quality and forecast value was evident in the results of these studies. It is of particular importance here to note that forecast quality was characterized in terms of a single parameter (e.g., forecast variance) in each study. As a result, the respective quality–value relationships were represented by single-valued functions. Frequently, however, two or more parameters are required to describe forecast quality unambiguously, and it is also of interest to study quality–value relationships in these more general situations.

The primary purpose of this paper is to investigate the relationship between the quality and value of forecasts in the context of a generalized model of the cost–loss ratio situation. This model was originally described by Epstein (1969), and it was recently used by Murphy (1985) to study the use and value of various types of forecasts. Although the model involves N-actions and N-events in general, it is sufficient for the purposes of this paper to consider the situation in which $N = 3$. The model of the three-action, three-event situation is briefly described in section 2. This section also defines the relevant meteorological information and presents expected expense expressions associated with the use of such information. Measures of forecast quality and forecast value are defined in section 3. Quality–value relationships are investigated in section 4, and this section also includes examples of these relationships in several specific cases. Section 5 contains a brief summary, a discussion of some implications and limitations

* On leave from Nanjing Institute of Meteorology, Nanjing, P.R.C.

of the results, and several suggestions regarding future work in this area.

2. Generalized cost–loss ratio model: three-action, three-event situation

In this situation the decision maker must select one of three admissible actions: A_1—full protection, A_2—partial (half) protection, and A_3—no protection. The relevant weather conditions are described in terms of three mutually exclusive and collectively exhaustive events: W_1—completely adverse weather, W_2—partially (half) adverse weather, and W_3—no adverse weather. This three-action, three-event situation involves nine ($=3^2$) consequences, and the impacts of these consequences on the decision maker are measured in terms of costs of protection and/or losses which may be incurred if protection is inadequate. Under the assumptions that the cost of protection decreases linearly from C if A_1 is taken to zero if A_3 is taken and that the loss incurred is zero if A_i is taken and W_j occurs ($i \leqslant j$) and otherwise ($i > j$) increases linearly to L when A_3 is taken and W_1 occurs, the expenses (to the decision maker) can be expressed as follows:

$$E_{ij} = \begin{cases} (3-i)(C/2), & i \leqslant j \\ (3-i)(C/2) + (i-j)(L/2), & i > j \end{cases} \tag{1}$$

($i, j = 1, 2, 3$) (see Epstein, 1969; Murphy, 1985). It is generally convenient to transform these expenses into standardized expenses $E'_{ij} = E_{ij}/L$, in which case

$$E'_{ij} = \begin{cases} [(3-i)/2](C/L), & i \leqslant j \\ [(3-i)/2](C/L) + (i-j)/2, & i > j \end{cases} \tag{2}$$

($i, j = 1, 2, 3$). The admissibility of the three actions implies that $0 < C < L < \infty$. Thus, $0 < C/L < 1$ and $0 \leqslant E'_{ij} \leqslant 1$. The basic (standardized) expense tableau is depicted in Table 1.

Three types of meteorological information are considered here: climatological information, imperfect forecasts, and perfect information. Let $X_j = 1$ if W_j occurs and $X_j = 0$ otherwise ($j = 1, 2, 3$). Then climatological information consists simply of the set of three probabilities $p_j = \Pr(X_j = 1)$ ($p_j \geqslant 0$, $\sum_j p_j = 1$; $j = 1, 2, 3$). Imperfect forecasts are assumed to represent categorical forecasts of the events, where $F_l = 1$ if the lth event is forecast and $F_l = 0$ otherwise ($l = 1, 2, 3$).

TABLE 1. The basic expense tableau for the three-action, three-event generalized cost–loss ratio situation, with standardized expenses E'_{ij} ($i, j = 1, 2, 3$).

	Events		
Actions	W_1	W_2	W_3
A_1	C/L	C/L	C/L
A_2	$C/2L + 1/2$	$C/2L$	$C/2L$
A_3	1	1/2	0

These forecasts can be completely characterized in terms of conditional probabilities p_{jl}, where $p_{jl} = \Pr(X_j = 1 | F_l = 1)$ ($p_{jl} \geqslant 0$, $\sum_j p_{jl} = 1$; $j, l = 1, 2, 3$), and predictive probabilities π_l, where $\pi_l = \Pr(F_l = 1)$ ($\pi_l \geqslant 0$, $\sum_l \pi_l = 1$; $l = 1, 2, 3$). Climatological information corresponds to the limiting case of imperfect forecasts for which $p_{jl} = p_j$ for all l ($j, l = 1, 2, 3$). Perfect information, on the other hand, corresponds to the limiting case of imperfect forecasts for which $p_{jl} = 1$ if $j = l$ and $p_{jl} = 0$ otherwise (and necessarily $\pi_j = p_j$) ($j, l = 1, 2, 3$). Climatological and perfect information are of interest because they represent lower and upper bounds, respectively, on the quality of imperfect forecasts. It is also important to recognize that certain relationships exist among the climatological, conditional, and predictive probabilities. Specifically, according to the definition of conditional probability,

$$p_j = \sum_{l=1}^{3} \pi_l p_{jl} \tag{3}$$

($j = 1, 2, 3$).

The decision maker is assumed to choose the action that minimizes his/her expected expense, where expected expense is the probability-weighted average of the relevant expenses (costs and/or losses). For example, in the case of climatological information, the expected expense associated with action A_i is $\mathrm{EE}(A_i)$, where

$$\mathrm{EE}(A_i) = [(3-i)/2](C/L) + \frac{1}{2}\sum_{j=1}^{i}(i-j)p_j \tag{4}$$

($i = 1, 2, 3$). This decision criterion implies that the decision maker will prefer action A_i to the other two actions when

$$\sum_{j=1}^{i-1} p_j < C/L < \sum_{j=1}^{i} p_j \tag{5}$$

(see Epstein, 1969; Murphy, 1985).

Expected expense expressions for climatological information, imperfect forecasts, and perfect information in the framework of the generalized cost–loss ratio model were presented by Murphy (1985). They were based on the assumption that the decision maker adopts the information (or forecasts) as the sole basis for choosing the optimal action. Specifically, if EC_3 denotes the expected expense associated with climatological information in the three-action, three-event situation, then $\mathrm{EC}_3 = \mathrm{EE}(A_i)$; that is, from (4),

$$\mathrm{EC}_3 = [(3-i)/2](C/L) + \frac{1}{2}\sum_{j=1}^{i}(i-j)p_j, \tag{6}$$

where the numerical value of the index i is defined by the inequality in (5). If EF_3 denotes the expected expense associated with imperfect forecasts in this situation, then

$$\mathrm{EF}_3 = \sum_{l=1}^{3} \pi_l \left\{ [(3-k_l^*)/2](C/L) + \frac{1}{2}\sum_{j=1}^{k_l^*}(k_l^* - j)p_{jl} \right\}, \tag{7}$$

where the numerical value of the index k_l^* is defined by the inequality

$$\sum_{j=1}^{k_l^*-1} p_{jl} \leq C/L < \sum_{j=1}^{k_l^*} p_{jl} \quad (8)$$

($l = 1, 2, 3$). Finally, if EP_3 denotes the expected expense associated with perfect information in this situation, then

$$EP_3 = (C/2L) \sum_{j=1}^{3} (3-j)p_j. \quad (9)$$

Since climatological and perfect information represent lower and upper bounds, respectively, on the quality of imperfect forecasts, it can be shown that $0 \leqslant EP_3 \leqslant EF_3 \leqslant EC_3 \leqslant 1$. For a more detailed discussion of these expense expressions and other aspects of the generalized cost–loss ratio situation, refer to Murphy (1985).

3. Measures of quality and value

a. Measure of quality

We are concerned here with the identification of a suitable measure of the quality of imperfect forecasts. First, however, it may be instructive to discuss briefly the determinants of forecast quality in this context. As noted in section 2, the imperfect forecasts considered in this paper are completely characterized by the conditional probabilities (i.e., the p_{jl}) and the predictive probabilities (i.e., the π_l). Thus, in a three-action, three-event situation, forecast quality is determined by nine conditional probabilities and three predictive probabilities. Of course, certain relationships exist among these probabilities [e.g., $\sum_j p_{jl} = 1$ and $\sum_l \pi_l = 1$ ($j, l = 1, 2, 3$); see also (3)], thereby reducing the number of *independent* parameters involved. Moreover, in some situations, it may be reasonable to specify the predictive probabilities (i.e., to consider the π_l as given). Nevertheless, the set of determinants of forecast quality is, in general, multidimensional (i.e., it contains two or more independent parameters). Further discussion of this issue will be postponed until section 5.

It is traditional and convenient to utilize a scalar (or one-dimensional) measure to assess forecast quality. In this regard, Epstein (1969) formulated a measure of the quality of probabilistic forecasts—the ranked probability score (RPS)—within the context of the generalized cost–loss ratio situation. Thus, the RPS, which can also be shown to represent the mean square error of cumulative forecasts (see Murphy, 1971), is a natural measure of forecast accuracy in this context. For a three-action, three-event situation, the expected RPS can be defined in terms of the conditional and predictive probabilities as follows (Murphy, 1985):

$$RPS_3 = \sum_{l=1}^{3} \sum_{j=1}^{3} \pi_l p_{jl} \sum_{k=1}^{3} \left(\sum_{h=1}^{k} p_{hl} - \delta_k \right)^2, \quad (10)$$

where $\delta_k = 1$ if $k \geqslant j$ and $\delta_k = 0$ otherwise. RPS_3 in (10) ranges from zero for perfect forecasts [$p_{jl} = 1(0)$ for $j = (\neq)\ l$ and $\pi_j = p_j$] to $p_1(1 - p_1) + p_3(1 - p_3)$ for climatological information ($p_{jl} = p_j$ for all l). The latter attains its maximum value of one-half when $p_1 = p_3 = 1/2$ (and necessarily $p_2 = 0$).

b. Measure of value

The value of imperfect forecasts depends on the value of the information consulted by the decision maker in the absence of these forecasts. In this context, it is reasonable to assume that climatological information would always be available to the decision maker. Thus, we define the value of imperfect forecasts in terms of the reduction in expected expense associated with these forecasts vis-a-vis the expected expense associated with climatological information. Therefore, if VF_3 denotes the expected value of the forecasts, then

$$VF_3 = EC_3 - EF_3, \quad (11)$$

and, from (6) and (7),

$$VF_3 = [(3-i)/2](C/L) + \frac{1}{2} \sum_{j=1}^{i} (i-j)p_j - \sum_{l=1}^{3} \pi_l \left\{ [(3-k_l^*)/2](C/L) + \frac{1}{2} \sum_{j=1}^{k_l^*} (k_l^* - j)p_{jl} \right\}, \quad (12)$$

where the numerical values of the indices i and k_l^* are determined by the inequalities in (5) and (8), respectively. Analogously, if VP_3 denotes the expected value of perfect information, then

$$VP_3 = EC_3 - EP_3, \quad (13)$$

and, from (6) and (9),

$$VP_3 = [(3-i)/2](C/L) + \frac{1}{2} \sum_{j=1}^{i} (i-j)p_j - \frac{1}{2}(C/L)(2p_1 + p_2), \quad (14)$$

where the numerical value of the index i is determined by the inequality in (5). Since $0 \leqslant EP_3 \leqslant EF_3 \leqslant EC_3 \leqslant 1$, it follows that $0 \leqslant VF_3 \leqslant VP_3 \leqslant 1$.

In section 4c, we describe the results of an investigation of the relationship between the quality and value of imperfect forecasts. In order to limit the number of different cases that must be considered in the quality–value context, we will assume that the indices k_l^* take on a particular set of values; namely, $k_l^* = l$ ($l = 1, 2, 3$). From (8), this assumption places certain bounds on the values of the conditional probabilities p_{jl} and the cost–loss ratio C/L. Specifically, it implies that $p_{11} > C/L$, $p_{12} < C/L$, $p_{12} + p_{22} > C/L$, and $p_{13} + p_{23} < C/L$. Since many decision-making situations involve relatively small cost–loss ratios (e.g., $0 < C/L < 1/3$), these conditions on the p_{jl} include those cases in which forecast quality attains relatively high levels [i.e., in which p_{jj} is large relative to p_{jl} $(l \neq j)$]. Under these assumptions, the expression for VF_3 in (12) can be rewritten in a simplified form. Specifically,

$$VF_3 = [(3-i)/2](C/L) + \frac{1}{2}\sum_{j=1}^{i}(i-j)p_j - \pi_1(C/L) - (\pi_2/2)[(C/L)+p_{12}] - (\pi_3/2)(2p_{13}+p_{23}). \quad (15)$$

4. Relationships between quality and value

a. Basic considerations

The measure of value VF_3 in (15) depends on the numerical value of the index i, which relates to the optimal action associated with climatological information. To facilitate the investigation of quality–value relationships, we will assume that the optimal action for climatological information is A_2 (i.e., that $i = 2$). This assumption is equivalent assuming that $p_1 < C/L$ and $p_1 + p_2 > C/L$ [see (5)]. Under these additional conditions, (14) and (15) become

$$VP_3 = \frac{1}{2}[(1-2p_1-p_2)(C/L)+p_1] \quad (16)$$

and

$$VF_3 = (C/L)\left[\frac{1}{2} - \pi_1 - (\pi_2/2)\right] + \frac{1}{2}p_1 - \frac{1}{2}\pi_2 p_{12} - \pi_3 p_{13} - \frac{1}{2}\pi_3 p_{23}, \quad (17)$$

respectively.

At this point, it is also convenient to rewrite RPS_3 in (10) in the following form:

$$RPS_3 = \mathbf{a}'\mathbf{p} + \mathbf{p}'\mathbf{B}\mathbf{p}, \quad (18)$$

where $\mathbf{p}' = (p_{11}, p_{21}, p_{12}, p_{22}, p_{13}, p_{23})$, $\mathbf{a}' = (2\pi_1, \pi_1, 2\pi_2, \pi_2, 2\pi_3, \pi_3)$,

$$\mathbf{B} = (-1)\begin{pmatrix} 2\pi_1 & \pi_1 & 0 & 0 & 0 & 0 \\ \pi_1 & \pi_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\pi_2 & \pi_2 & 0 & 0 \\ 0 & 0 & \pi_2 & \pi_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\pi_3 & \pi_3 \\ 0 & 0 & 0 & 0 & \pi_3 & \pi_3 \end{pmatrix}, \quad (19)$$

and a prime denotes transposition. RPS_3 in (18) depends on six conditional probabilities (as well as three predictive probabilities). However, it is possible to use the relationships among the climatological, conditional, and predictive probabilities in (3) to eliminate p_{11} and p_{12} from this expression. It should also be noted that the matrix $\mathbf{B}$ is symmetric and negative definite; that is, RPS_3 is a strictly concave function.

To facilitate further the comparison of the expressions for VF_3 and RPS_3 in (17) and (18), respectively, and the eventual investigation of quality–value relationships, we will also assume that the p_j $(j = 1, 2, 3)$ and the π_l $(l = 1, 2, 3)$ are given. Then it can be seen that VF_3 depends on a set of three parameters (assuming that C/L is also specified)—namely, p_{12}, p_{13}, and p_{23}—and that RPS_3 depends on a set of four parameters (assuming that p_{11} and p_{21} are eliminated)—namely, p_{12}, p_{22}, p_{13}, and p_{23}. Thus, VF_3 and RPS_3 depend on overlapping sets of parameters, which implies the existence of a complex nonlinear relationship between these two quantities.

Further consideration of this relationship also suggests intuitively that it is not a one-to-one (or single-valued) relationship. That is, a single value of VF_3 is not associated with a given value of RPS_3 (and vice versa). For example, when VF_3 is fixed (and even when p_{12}, p_{13} and p_{23} are also fixed), RPS_3 can still vary depending on the numerical value of p_{22}. Moreover, recall that no scalar (or one-dimensional) measure such as RPS_3 can *uniquely* characterize forecast quality since quality depends on a multidimensional set of parameters in this context. As a result, a fixed value of RPS_3 is generally associated with a set of possible values of the parameters p_{12}, p_{22}, p_{13}, and p_{23}, thereby yielding a range of values of VF_3. Thus, the relationship between RPS_3 and VF_3 would be expected to be characterized by an *envelope* of values, with a range of values of VF_3 associated with a fixed value of RPS_3 (and vice versa).

To provide an explicit illustration of the nature of this relationship, consider the case in which $p_1 = \pi_1 = 0.1$, $p_2 = \pi_2 = 0.3$, $p_3 = \pi_3 = 0.6$, and $C/L = 0.3$ (and assuming that all other previously discussed conditions also hold). Further, suppose that RPS_3 is fixed at 0.199. Then, it can be shown (see section 4c) that, for this value of RPS_3, VF_3 can range from 0.036 to 0.064 ($VP_3 = 0.125$ in this case). These "limits" on VF_3 are associated with the following sets of parameter values: $(p_{12}, p_{13}, p_{22}, p_{23}) = (0.000, 0.001, 0.410, 0.295)$ and $(p_{12}, p_{13}, p_{22}, p_{23}) = (0.035, 0.038, 0.745, 0.109)$, respectively. Thus, although these sets of conditional probabilities yield the same value of RPS_3, they lead to quite different values of VF_3. Similarly, for $VF_3 = 0.060$, RPS_3 can range from 0.153 to 0.209 in this case.

b. Methods of determining quality–value envelopes

Two methods have been employed to investigate the relationship between RPS_3 and VF_3. First, we have used a brute force approach in which RPS_3 and VF_3 have been computed for all possible values of the parameters p_{12}, p_{13}, p_{22}, and p_{23}. Specifically, since these parameters are all defined on the closed interval [0, 1], it is possible to compute values of the measures over a relatively dense four-dimensional grid of parameter values. In this regard, we have made such computations based on grid "steps," in probability terms, of both 0.010 and 0.005. Given the pairs of values of RPS_3 and VF_3 for this grid of parameter values, we can then find maximum and minimum values of VF_3 for a specific value of RPS_3, or vice versa. The "curves" defined by these maxima and minima determine the boundaries of the quality–value envelope.

Alternatively, the problem of finding the quality–value envelope can be viewed as a mathematical programming problem of maximizing or minimizing a quadratic function—namely, RPS_3—subject to certain linear constraints. In this regard, it can be readily seen that RPS_3 in (18) is a quadratic function of the p_{jl}. The linear constraints can be characterized as follows: (i) the value of the forecasts as defined by VF_3 in (17) is constant; (ii) the conditional probabilities satisfy the inequalities specified by the numerical values of the indices k_l^* (l = 1, 2, 3) and the cost–loss ratio C/L described in (8); (iii) the conditional probabilities satisfy the relationships among the climatological, conditional, and predictive probabilities described in (3); and (iv) the conditional probabilities for a specific forecast $F_l = 1$ sum to unity (i.e., $\sum_j p_{jl} = 1; j, l = 1, 2, 3$).

To solve this optimization problem, we have employed software associated with the Multi-Purpose Optimization System developed by Northwestern University (Cohen and Stein, 1978). Specifically, we used an algorithm known as BEALE, which is described in Beale (1968). Unfortunately, the optimization problem is properly posed only for the case of maximizing RPS_3, since it is a concave function. As a result, the BEALE algorithm will generally not yield a global optimum for the case of minimizing RPS_3. Nevertheless, the mathematical programming approach provides a means of validating the results obtained from the brute force method for the maximum value of RPS_3 (for fixed VF_3).

c. *Examples of quality–value envelopes*

The basic three-action, three-event situation of concern can be described in terms of the values of the parameters i and k_l^* (l = 1, 2, 3). As noted previously, we have assumed here that $i = 2$ (i.e., that A_2 is the optimal action associated with climatological information) and that $k_1^* = 1$, $k_2^* = 2$, and $k_3^* = 3$ [which imply certain conditions on the p_{jl} and on C/L; see (8)]. To determine the quality–value envelope it is also necessary to specify the values of the climatological probabilities p_j (j = 1, 2, 3), the predictive probabilities π_l (l = 1, 2, 3), and the cost–loss ratio C/L. It is convenient to assume that the predictive and climatological probabilities are equal [i.e., that $\pi_j = p_j$ (j = 1, 2, 3)]. In effect, we are assuming that the relative frequency of use of the jth forecast is equal to the relative frequency of occurrence of the jth event. This assumption seems quite reasonable (e.g., Katz and Murphy, 1987); in any case, it can be relaxed when appropriate (see section 5).

Quality–value envelopes for three cases are presented here: (a) *Case A:* $p_1 = \pi_1 = 0.10$, $p_2 = \pi_2 = 0.30$, $p_3 = \pi_3 = 0.60$, $C/L = 0.30$; (b) *Case B:* $p_1 = \pi_1 = 0.10$, $p_2 = \pi_2 = 0.30$, $p_3 = \pi_3 = 0.60$, $C/L = 0.15$; (c) *Case C:* $p_1 = \pi_1 = 0.05$, $p_2 = \pi_2 = 0.20$, $p_3 = \pi_3 = 0.75$, $C/L = 0.15$. The quality–value envelope for Case A is depicted in Fig. 1. This envelope was determined by the brute force method by finding the maximum and minimum values of RPS_3 for each value of VF_3 in the set 0.000 (0.001) 0.125. (Note that $VP_3 = 0.125$ in this case.) The maximum values of RPS_3 obtained by this method were validated by means of the BEALE algorithm.

The quality–value envelope in Fig. 1 exhibits some interesting features. First, as expected, value (i.e., VF_3) generally increases as quality increases (i.e., as RPS_3 decreases). Second, the shapes of the loci of points describing the maximum and minimum values of RPS_3 are quite different. The former evidently define a single smooth curve, whereas the latter appear to consist of four smooth curves joined at three cusps. In general, the shapes of the boundaries of the quality–value envelope are determined by the conditions (on the p_{jl}) associated with the case in question.

As a result of the differences in shapes between the two boundaries, the quality–value envelope itself is irregular in appearance. For very small values of RPS_3 (i.e., $RPS_3 \leq 0.05$, indicating almost perfect forecasts), the range of values of VF_3 is quite limited. For larger values of RPS_3, however, VF_3 possesses a range of values exceeding 0.015 from the minimum to the maximum. Moreover, for some values of RPS_3, this range of VF_3 values exceeds 0.025, or 20% of the value of perfect information. Thus, knowledge of the quality of the forecasts, as reflected by RPS_3, frequently will not provide a very precise estimate of the value of the forecasts.

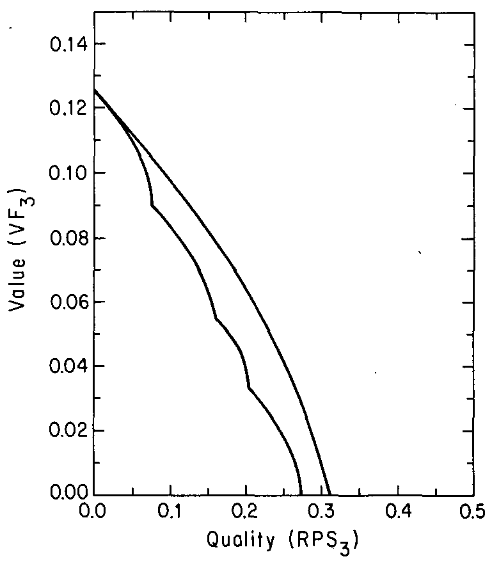

FIG. 1. Quality–value envelope in the basic situation ($i = 2$; $k_1^* = 1$, $k_2^* = 2$, $k_3^* = 3$) when $p_1 = \pi_1 = 0.1$, $p_2 = \pi_2 = 0.3$, $p_3 = \pi_3 = 0.6$, and $C/L = 0.3$.

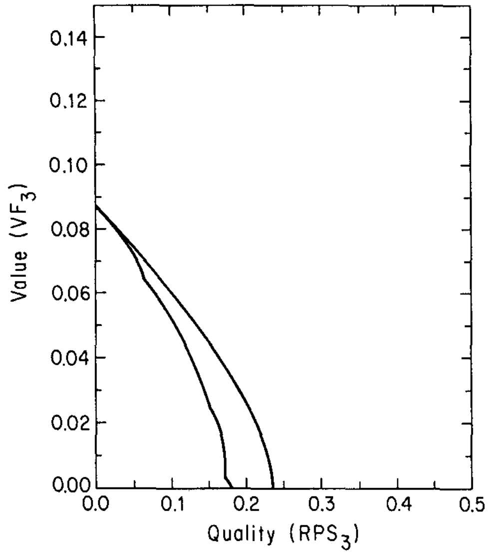

FIG. 2. As in Fig. 1, except $C/L = 0.15$.

Quality–value envelopes in Cases B and C are depicted in Figs. 2 and 3, respectively. These envelopes were also obtained by the brute force method, with subsequent validation of the maximum values of RPS_3 using the BEALE algorithm. Although the values of VP_3 and VF_3 are smaller in these cases—undoubtedly due to the smaller value of C/L—the quality-value envelopes possess the same general appearance as in Case A. Namely, the loci of maximum values of RPS_3 represent a single smooth curve, whereas the loci of minimum values of RPS_3 consist of four smooth curves that are joined at three cusps. As in Case A, VF_3 in these two cases possesses an appreciable range of values—vis-a-vis VP_3—for all but the smallest values of RPS_3. In this regard, the range of VF_3 in Case B exceeds 30% of VP_3 for large values of RPS_3. Moreover, the width of the range of values of VF_3 generally decreases as RPS_3 decreases in this case. On the other hand, the width of the range of VF_3 values in Case C attains its maximum value for intermediate values of RPS_3. In general, the results for Cases B and C support the conclusion (drawn from the results for Case A) that RPS_3 will frequently provide an imprecise estimate of VF_3.

5. Summary and conclusion

In this paper we have investigated the relationship between the quality and value of forecasts in the context of a generalized N-action, N-event model of the cost-loss ratio situation. The forecasts of interest are imperfect categorical forecasts, calibrated according to past performance and characterized by multidimensional sets of conditional and predictive probabilities. Climatological and perfect information represent lower and upper bounds, respectively, on the quality and value of these forecasts. The measure of quality is the ranked probability score, a one-dimensional (i.e., scalar) measure of the accuracy of the forecasts. Forecast value is measured by the difference in expected expenses between decisions based on climatological information and decisions based on imperfect forecasts. Both brute force and mathematical programming methods were used to obtain numerical results regarding quality–value relationships for several cases involving the three-action, three-event situation. The results indicate that these quality–value relationships are described by envelopes of values rather than by single-valued functions. That is, a range of forecast value exists for a given level of forecast quality, and vice versa. This range can be quite wide; for example, it exceeds 20–25% of the value of perfect information for some parameter values in each case considered. Thus, the quality of forecasts, as measured by the ranked probability score (or similar verification measures), frequently will provide only an imprecise estimate of forecast value even in those situations in which the quality–value relationship is "known."

At this point, several questions suggest themselves. For example, to what extent are the results presented here representative of the results that would be obtained in cases involving other parameter values and/or in situations involving more than three actions and events? What are the fundamental reasons that quality–value relationships are characterized by envelopes rather than single-valued functions in situations such as those considered in this paper? What are the methodological and practical implications of these results? With respect to the representativeness of the results presented in this paper, it is important to recognize that the quality–value relationships described in section

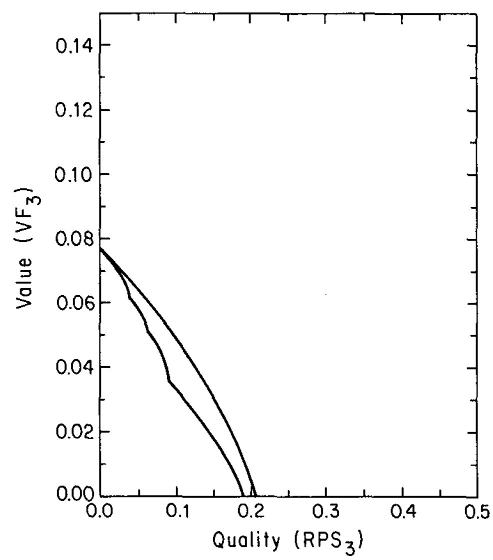

FIG. 3. As in Fig. 1, except $p_1 = \pi_1 = 0.05$, $p_2 = \pi_2 = 0.20$, $p_3 = \pi_3 = 0.75$, and $C/L = 0.15$.

4c relate to a quite limited set of cases (i.e., parameter values). Thus, it is of some interest to determine whether different results would be obtained for other values of the p_j ($j = 1, 2, 3$) and C/L, for cases involving different assumptions concerning the values of i and k_l^* ($l = 1, 2, 3$), or for cases in which the assumption of equality between the p_j and the π_j ($j = 1, 2, 3$) is relaxed. In fact, we have explored quality–value relationships in several cases involving other parameter values and assumptions (results omitted to conserve space). The quality–value relationships in these cases are also described by envelopes, although the characteristics of the envelopes sometimes differ from those of the envelopes depicted in Figs. 1–3. For example, in cases in which $\pi_j \neq p_j$ ($j = 1, 2, 3$), the loci of maximum as well as minimum values of RPS_3 consist of several smooth curves joined at cusps. Nevertheless, in all cases (reported or omitted), a range of values of VF_3 generally exists for each value of RPS_3 (and vice versa).

In all previous studies, quality–value relationships have been described by single-valued functions. Thus, it is important to understand the reasons that these relationships are characterized by envelopes of values (i.e., multivalued functions) in the situation examined herein. As noted in section 3, the quality of the imperfect forecasts considered here is characterized completely in the three-event situation by a set of nine conditional probabilities and three predictive probabilities. Relationships among these probabilities (including the climatological probabilities) reduce the number of independent parameters involved, but this set is still multidimensional (i.e., it contains four parameters, even when the predictive probabilities are assumed to be specified). However, we have chosen, following traditional practices, to measure the quality of the forecasts in terms of a single, scalar (i.e., one-dimensional) measure of performance; namely, RPS_3. As a result, different combinations of parameter values (i.e., values of the conditional probabilities representing various levels of multidimensional quality) yield the same (one-dimensional) value of RPS_3. However, these different combinations of parameter values generally lead to different values of VF_3. Thus, a range of values of VF_3 exists for most values of RPS_3, yielding a quality–value envelope instead of a single-valued function.

We believe that the results presented in this paper have two important implications with regard to studies of forecast quality, forecast value, and quality–value relationships. First, the arguments set forth in the previous paragraph indicate that no single, scalar measure of performance such as RPS_3 or other traditional verification measures, can completely—or even adequately—describe the quality of forecasts in most situations. From the perspective of forecast verification (i.e., studies of forecast quality), this result suggests that greater attention should be paid to the basic determinants of quality such as the conditional and predictive probabilities. These probabilities constitute elements of the conditional and marginal distributions of forecasts and observations, and verification methods based on such distributions would be consistent with a general framework for forecast verification recently described by Murphy and Winkler (1987).

Second, the results presented here demonstrate the existence of quality–value envelopes (rather than single-valued functions) in multidimensional situations in which forecast quality is characterized by a single, one-dimensional measure of performance. Thus, knowledge of the quality of the forecasts, as determined by such a measure, is inadequate as a measure of value for two reasons: (i) the quality–value relationship is inherently nonlinear (e.g., see Katz and Murphy, 1987) and (ii) the relationship is imprecise due to the existence of a range of forecast value estimates for most forecast quality estimates. In fact, the existence of these envelopes implies that decreases in (expected) value, as measured by VF_3, can actually be associated with decreases in RPS_3, and vice versa!

In view of the importance of quality–value relationships and the limited set of cases considered here, it would be desirable to investigate such relationships in the context of the generalized cost–loss ratio situation under a wider set of conditions. For example, additional studies should be undertaken for cases involving different values of the indices i and k_l^* and of the parameter C/L, as well as for cases in which the assumption of equality between the climatological and predictive probabilities is relaxed. Moreover, since attention was restricted here to a three-action, three-event situation, quality–value relationships should be explored in situations involving more than three actions and events. From a methodological point of view, it would be of interest to investigate the effect of using different measures of quality on the characteristics of the quality–value relationships. (See Murphy and Daan, 1985, for a recent discussion of measures of forecast quality.) For example, do different measures yield different relationships (e.g., envelopes with different shapes), and can the nature of these relationships be used as a criterion for choosing a suitable verification measure? Moreover, can general methods—other than the brute force method—be found that can describe quality–value relationships in situations in which these relationships are multivalued functions (i.e., are represented by envelopes)? These and other studies should provide additional insight into the nature of the complex relationship between forecast quality and forecast value.

Acknowledgments. We would like to thank Jeffrey L. Arthur for his helpful comments regarding the application of mathematical programming algorithms to the optimization problem considered here. This research was supported in part by the National Science Foundation (Division of Atmospheric Sciences) under Grant ATM-8507495.

REFERENCES

Beale, E. M. L., 1968: *Mathematical Programming in Practice.* John Wiley and Sons, 195 pp.

Brown, B. G., R. W. Katz and A. H. Murphy, 1986: On the economic value of seasonal precipitation forecasts: The fallowing/planting problem. *Bull. Amer. Meteor. Soc.,* **67,** 833–841.

Cohen, C., and J. Stein, 1978: *Multi-Purpose Optimization System User's Guide, Version 4.* Northwestern University, Vogelback Computing Center, Manual 320, 154 pp.

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

Katz, R. W., and A. H. Murphy, 1987: Quality/value relationship for imperfect information in the umbrella problem. *The American Statistician,* **41,** (in press).

——, ——, and R. L. Winkler, 1982: Assessing the value of frost forecasts to orchardists: A dynamic decision-making approach. *J. Appl. Meteor.,* **21,** 518–531.

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

——, 1985: Decision making and the value of forecasts in a generalized model of the cost-loss ratio situation. *Mon. Wea. Rev.,* **113,** 362–369.

——, and H. Daan, 1985: Forecast evaluation, *Probability, Statistics, and Decision Making in the Atmospheric Sciences,* A. H. Murphy and R. W. Katz, Eds., Westview Press, 379–437.

——, and R. L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.,* **115,** (in press).

——, R. W. Katz, R. L. Winkler, and W.-R. Hsu, 1985: Repetitive decision making and the value of forecasts in the cost-loss ratio situation: A dynamic model. *Mon. Wea. Rev.,* **113,** 801–813.

Wilks, D. S., and A. H. Murphy, 1986: A decision-analytic study of the joint value of seasonal precipitation and temperature forecasts in a choice-of-crop problem. *Atmos. Ocean,* **24,** 353–368.