MODULE 7.3

STATISTICAL METEOROLOGY ORIENTATION

TABLE OF CONTENTS

1. Perfect Prog and MOS Techniques

1.1 Introduction

Meteorology and statistics seem to have gone hand in hand with one another for a very long time. Forecasters use climatology as a reference level when preparing their forecasts and statistical forecasts have long been used in the preparation of weather forecasts. Prior to the advent of Numerical Weather Prediction (NWP), classical relationships were developed using present and samples of past meteorological observations as predictors to generate the predictands which were the future weather conditions.

NWP opened up new horizons in statistical forecasting by providing forecast variables that could be used as predictors. This led to the development of sophisticated statistical techniques to utilise these data. Two of these in use at the Canadian Meteorological Centre (CMC) are the Perfect Prog (PP) and Model Output Statistics (MOS) schemes which will be discussed later. As the NWP models improved in their horizontal and vertical resolution and as their physical parameterization of sub model scale phenomena improved, the available existing predictors improved and new ones became available.

Statistical techniques generally forecast weather elements such as temperature, wind, cloud cover and precipitation. Modern NWP models also forecast weather elements either directly or derived from basic model variables through diagnostic and other parameterization techniques. Why then do we still require statistical methods. One of the reasons for this is that there are some variables such as Probability of Precipitation (PoP) that are not forecast by deterministic NWP models. Another is that statistical methods, which can be considered a form of post-processing of the model output, adds value to the model output by improving the performance and quality of the forecasts. Through screening processes for individual sites, predictors are chosen which often reflect local forcing conditions such as topography, thus enhancing the ability of the system as a whole to forecast local conditions. The value added to the model forecast is illustrated in the following two verification charts for the temperature forecasts for Lytton, British Columbia where the benefits of statistical techniques are clearly illustrated.

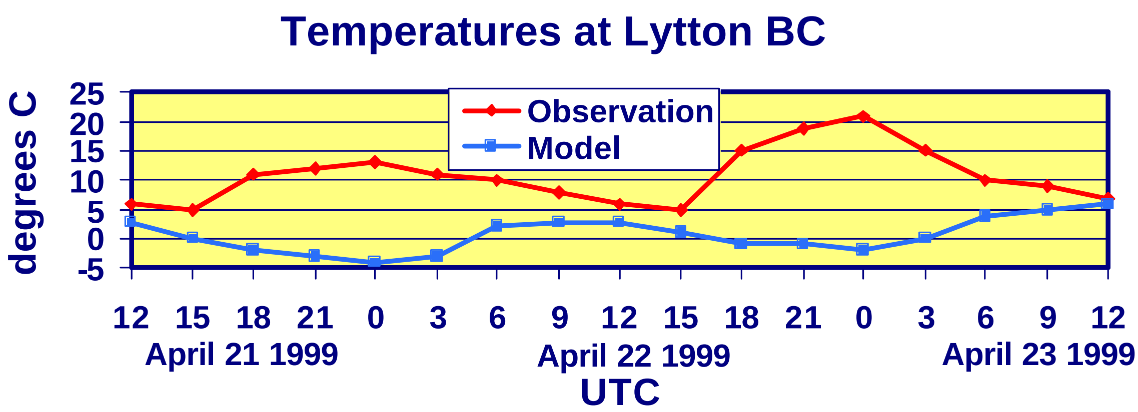
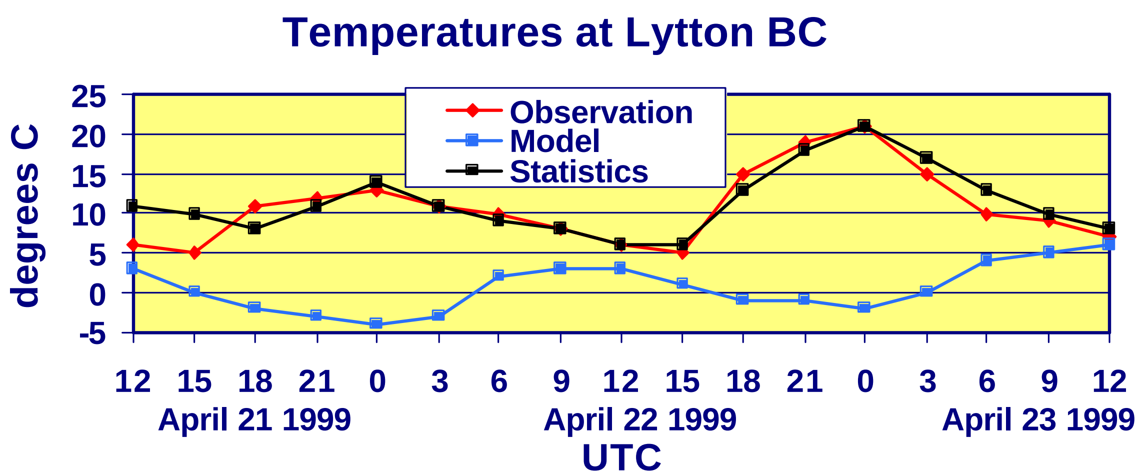


Figure 1. Comparison of direct model temperature forecasts with the verifying observations

Figure 2. Same data set as Figure 1 but with the addition of statistical forecasts



Additional benefits of objective statistical forecasting methods are that they can be considered a way of quantifying experience. New forecasters or experienced forecasters, when called upon to produce forecasts for unfamiliar areas, may not be well acquainted with the local characteristics of the weather or climate and can rely on statistical predictions to fill the void. In a sense this provides a form of instant training for forecasters. Forecasts generated by objective statistical schemes provide a first cut at the final forecast product and in some cases are considered good enough to provide the final product. This is done either directly or through feeding expert systems that use the best possible guidance to generate weather forecast products, tailored to specific users.

1.2 Methods

Irrespective of the method, a statistical formula or equation will express the relationship between the independent variables (predictors) and the dependent variables (predictands) mathematically. It is important when using products derived from statistical formulae to understand how, under what conditions these relationships were derived, and any shortcomings the mathematical technique might have.

The process of producing a statistical formula for a certain dependent variable starts with a data set which comes from actual weather cases. Long term data sets are required to provide a representative sample and insure the stability of the statistics. In general there are three main sources of predictors: observations and variables derived from observations, climatic terms and numerical model output. From these, predictors that are chosen which must have meteorological significance. This means that they should be expected to have some physical relationship to the predictand and their choice should be defensible on physical grounds. Predictors, whether or not they come from NWP models must be available in forecast form with a sufficient degree of accuracy to perform well in the statistical system. Screening or selection processes are used to find those predictors which have the highest correlation to the predictand and therefore do the best job of expressing the values of the predictand and explaining as much as possible the variance of the predictand in the development sample.

Development data are very often stratified or grouped for meteorological and climatological reasons in order to select predictors. One grouping or statification may be by geography. It is reasonable to expect, for example, that the screening process will pick different predictors near the surface in mountainous areas than over flat terrain. Another example is the difference in maritime versus continental climates which is significant and it is therefore reasonable to expect that better results would be obtained by developing seperate sets of equations for each of these areas.

It is also very important to take seasonal variations into account when developing statistical relationships. This is particularly essential in Canada where the difference between winter and summer climates is quite large. Data are therefore often grouped and equations developed by season. The length of a season, for statistical purposes, depends on the weather element we wish to forecast and also, as we shall see later, on limitations imposed by the size of the development sample. In some cases, the annual cycle is simply approximated by two six month seasons, winter and summer, and in others there may be as many as twelve one month periods.

1.2.1 Classical Approach

The so-called classical approach is useful mainly in short range prediction. It uses presently observed or recent values of a weather parameter to predict the future value of this or some other weather parameter.

As an example, if you wished to predict the overnight minimum temperature, you could use the drop in temperature between, say, the 3 p.m. temperature and the sunset temperature as your guideline. You could derive statistics for each month of the year to relate this observed temperature drop to the expected overnight drop. You might notice some relationship between sunrise dewpoint and afternoon cumulus cover. Obviously, you could use more than one predictor to improve your technique. There are a number of possible relationships that could be used for classical statistical prediction. An interesting point to keep in mind is that you don't have to use data at the station for which you are forecasting. You might find you can use the maximum temperature at another station on the previous day as a predictor of maximum temperature at your target station today. The time lag between the predictors and the predictand in the development sample becomes the projection time when the statistical relationship is used in forecast mode. In forecast mode, observed predictors are used to provide the forecast value of the predictand. No NWP models are involved in the process.

The advantage of the classical approach is that you use actual values of your predictor rather than a forecast value. It is most useful with conditions which would favour persistence. It is also inexpensive to use and does not require significant computing power to be used operationally since no NWP models are involved.

The disadvantage of this technique is that it could only reasonably work in the short term, since it is mostly based on persistence. There can be many changes in the atmospheric conditions which will affect the value of your predictand, and the present value of the predictor may not be representative. A fast-moving system would produce problems for this approach. Also, the classical technique requires a large data base to provide stable and reliable results.

Note that although the classical method is generally used for very short range forecasts, in a very general form classical techniques are also used for very long range forecasts such as seasonal predictions where there is very little skill in the numerical progs.

1.2.2 Perfect prog approach or method (PPM)

Perfect prog techniques were first formulated over forty years ago and have been in use at the CMC since 1968. Even in the light of predictability considerations, we recognise that numerical progs are not perfect. The term “Perfect Prog” therefore appears unrealistic, but what it refers to is an estimate of what we could expect if the numerical forecasts were correct. Predictors are chosen from observations and objective analyses, i.e. Perfect Progs, and the resultant regression equations are then applied to forecast values of these predictors.

The technique is developed in a similar way to the classical approach, but it can use values of the predictor which occur at the same (or similar) time as the predictand. All the predictors are either observed values and/or extracted from analyses. Although the equations are dependant on valid time,, such as 00 UTC or 12 UTC, they are independent of projection time. The same relationship that is applied to the 24 hour forecast is also applied to any other valid projection time such as the 48, 72, 96 hour forecasts. The use of observed values as predictors effectively brings about a mix of the classical approach with a pure Perfect Prog approach. As in the pure classical approach, no models are involved in the development of the statistical relationships between the predictors and the predictands.

An example might be using 1000-500 hPa thickness at 00UTC as a predictor for the maximum temperature of the day. In this case, the predictor does not have to be observed in advance of the predictand (since 00UTZ is past maximum heating time over much of Canada). However, the data set used to derive the mathematical relationship consists of actually recorded weather data. Once the statistical formula has been created, the forecasts of maximum temperature can be made by using each of the forecast values of the 1000-500 hPa thickness which the model predicts for 00UTC. In practice, there would be more than one predictor used to come up with the predictand.

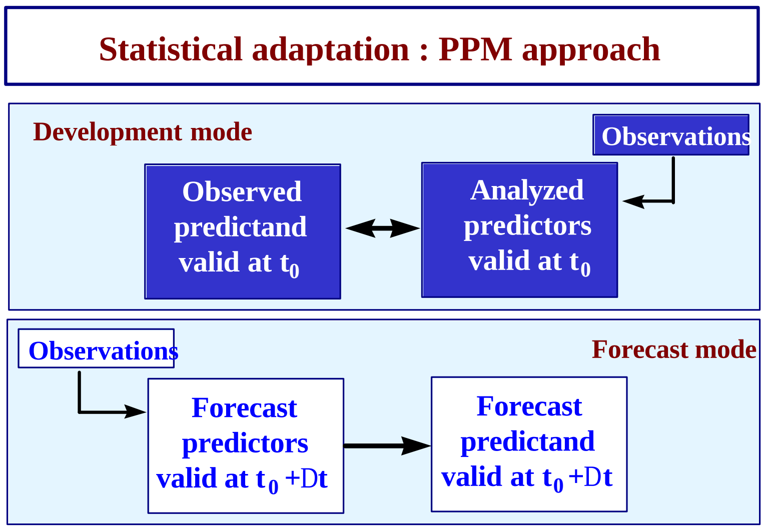


Figure 3. Graphical Summary of the PPM approach

The disadvantage of the Perfect Prog approach is that you are dependent upon forecast values of the predictors. This statistical approach cannot correct any of the model errors (systematic or non-systematic) since only analysed or observed predictors are involved in the development of the statistical relationships. Whatever error the model makes in predicting these parameters will go into the error of your predictand. In other words, the less "perfect" the prog is, the less perfect the statistical forecast will be.

The advantage of this approach over the classical approach is that you can use the technique any time in the future, not just short-range. It also has access to a wider range of potential predictors. Another advantage of Perfect Prog is that it is possible to build large development data sample to ensure the stability and significance of the statistical relationships.

1.2.3 Model Output statistics (MOS/UMOS)

With increasing reliance on the output of computer models for forecast parameters, MOS became a more and more useful technique. The difference with MOS is how it generates its original set of statistics. Actual weather cases are used, along with observed values of the predictand. But here, the correlation is made between the observed predictand and forecast values of the predictors generally valid at the same time. This is in contrast to the classical approach which does not use numerical forecasts at all and the perfect prog approach which uses numerical progs only in the operational application of the equations. Since the predictors consist of forecast parameters whose quality is obviously not fixed throughout a NWP model's integration, MOS relationships are dependant on both valid and projection times.

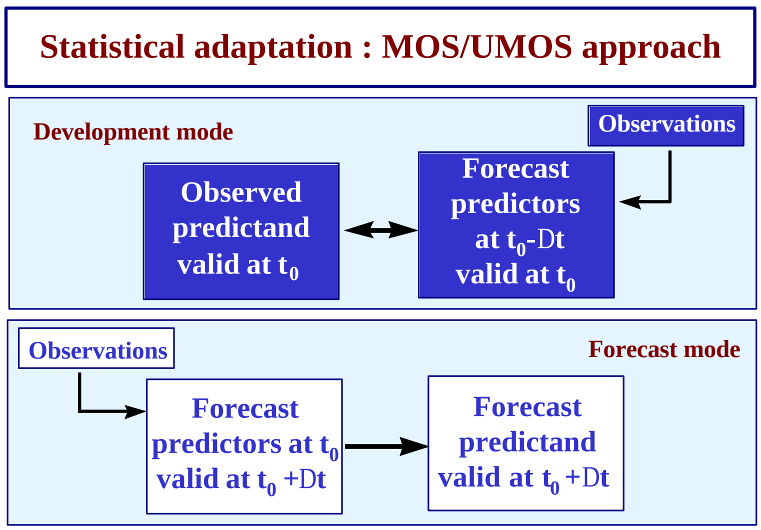


Figure 4. Summary of the MOS approach

To use our previous example, the technique would correlate observed values of the maximum temperature and values of the 1000-500 hPa thickness at 00Z that the model predicted on that occasion. The projection time of the forecast predictors used in the development of the statistical relationships, becomes the projection time of the forecast predictand when used in forecast mode.

The advantage of this approach is that the bias the model has in forecasting the predictor is removed. For example, the model you are using may under forecast values of the thickness, but this would not matter in the operational use of the technique. You are using the predictor in exactly the same way it was used in generating the original equation.

Generally, MOS is more reliable but will tend toward the climatology of the development sample with increasing projection time. This is because, by construction, the process discerns that the reliability of predictors decreases with projection time. Perfect Prog forecasts, on the other hand, are generally sharper in the sense that at longer projection times they are less likely to forecast near the climatological mean and will forecast extremes in the predictand more often than MOS. Reliability, as the term is used here, refers to the ability to forecast at the same frequency as the event is observed. This is mostly applicable to probabilistic forecasts such as probability of precipitation (PoP). For example, a probabilistic forecast system is said to be reliable when the relative frequency of individual categories of occurrence of an event will equal the forecast relative frequency of those categories. In other words, taking all probability forecasts of x% together, the observed occurrence of the events should be x% if the system is reliable.

The disadvantage of MOS is that it is dependent upon the characteristics of the model. You must have the model running for some time in order to generate the necessary equations; the more data you have, the better the equations should work. If you make any changes at all in the model, the MOS equations would, in principle, no longer apply. In practice, this turns out to be a serious problem for MOS systems. Numerical models are anything but static and so data samples tend to be small. While major model changes do not occur that often, improvements to the driving analyses and physics packages occur with regularity and, although beneficial to the driving model, can negatively impact the statistics system. Recompilations of the statistical relationships are required involving building a sample data set by archiving model output. This can lead to a significant lag between model change and implementation of updated MOS equations. This problem led to a gradual phasing out of the use of MOS for a time at the CMC.

1.2.3.1 UMOS

Updatable Model Output Statistics (UMOS) is the same as MOS. The term refers to a strategy and procedure developed to compensate for the difficulties discussed above. The key word is “updatable”. Unlike the traditional methods where data are collected over time and equations developed from the resultant data set and then used with little or no changes for a relatively long period of time, UMOS involves regular updating of the statistical relationships. At CMC there is an automatic daily preparation of the data for statistical processing followed by a weekly redevelopment of the equations to be used to forecast weather elements. The set-up can be adjusted such that the equations can be updated more or less frequently.

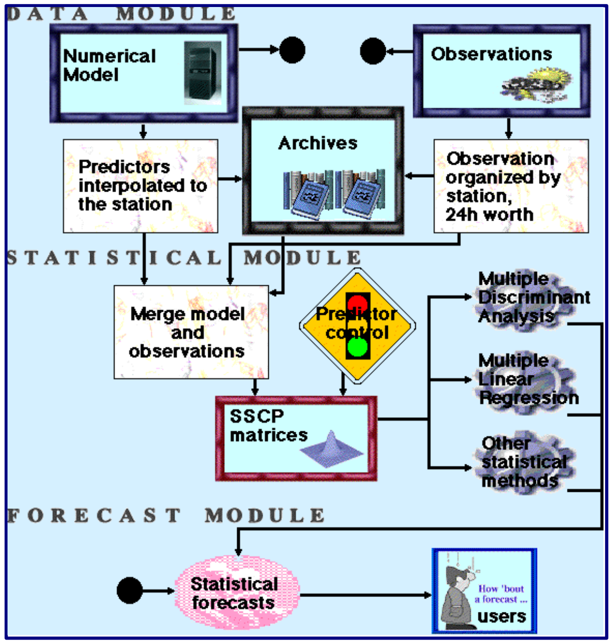


Figure 5. Graphical representation of the UMOS technique

To ensure a smooth transition across model changes, a weighting scheme is employed using data prior to and after the change to develop the weekly equations. This scheme is designed to emphasise data from the new or modified model while they are relatively scarce while using data from prior to the change to ensure a sufficiently large sample size to ensure stable statistical relationships.

PPM and traditional MOS systems generally have more than one set of equations developed from data sets representing periods or seasons throughout the year. In some cases, there are only two, one based on summer data and one based on winter data. Sets of equations are then rotated in the forecast system throughout the year depending on the period for which they were derived. The resultant transitions can be abrupt. The UMOS blending procedure is also used to smooth the transition between seasons. However in this case, data from the previous year are blended with current data.

On September 15, 1998 a significant change occurred in the Canadian regional implementation of the GEM (Global Environmental Multi-scale) model. The UMOS system was tested following the change and Table 1 provides a good illustration of how the data were merged and the effect on the weekly forecast equations. As can be seen by examining the coefficients of the individual predictors, the equations stabilised quite rapidly.

	N_2 (%)	N_1(%)	Weekly Equation	RV (%)
Equation 1 (Dec 13 1998)	100	0	.394HR700+.234PR-.399U700/dy+1185.7	34.93
Equation 2 (Dec 23 1998)	81	19	.259PR+.397HR700-.375U700/dy+1114.9	36.65
Equation 3 (Dec 31 1998)	79	21	.259PR+.406HR700-.390U700/dy+1159.9	32.21
Equation 4 (Jan 08 1999)	76	24	.249PR+.415HR700-.388U700/dy+1151.6	38.00
Equation 5 (Jan 14 1999)	74	26	.265PR+.428HR700-.382U700/dy+1133.5	39.62
Equation 6 (Jan 25 1999)	71	29	.275PR+.448HR700-.376U700/dy+1116.2	41.37
Equation 7 (Feb 01 1999)	69	31	.265PR+.446HR700-.385U700/dy+1141.1	41.00
Equation 8 (Feb 09 1999)	67	33	.267PR+.426HR700-.368U700/dy+1091.8	39.04
Equation 9 (Feb 16 1999)	65	35	.273PR+.416HR700-.372U700/dy+1103.8	39.17
Equation 10 (Mar 02 1999)	61	39	.268PR+.421HR700-.352U700/dy+1045.9	38.04
Equation 11 (Mar 09 1999)	59	41	.271PR+.419HR700-.355U700/dy+1052.4	39.07
Equation 12 (Mar 16 1999)	57	43	.270PR+.411HR700-.355U700/dy+1053.8	38.98
Equation 13 (Mar 26 1999)	55	45	.270PR+.408HR700-.342U700/dy+1014.2	38.48

Table 1. Weekly PoP regression equations for Toronto for a 06 - 12 hour forecast period updated during the test. N2 and N1 represent the percentage of data based on the old and new models respectively, RV is the reduction of variance of the fitting equations, PR is the total precipitation amount, HR700 is the relative humidity at 700 hPa, U700/dy is the negative part of the vorticity at 700 hPa.

1.2.4 Analogue Approach

The analogue technique is based on finding historical weather patterns that are similar or analogous to the current patterns or to patterns forecast by a numerical model. These analogue cases are then used to generate forecasts based on the assumption that the atmosphere is reproducible and that similar weather patterns and their associated weather conditions will evolve in a similar way.

There are sophisticated pattern matching techniques available however a simple method is used in which analogues are found by statistical matching of the pattern at both 1000 and 500 hPa to cases from a historical data base. Both levels are used to ensure vertical consistency. This is done for each projection time. The selection strategy is based on correlations of the height fields to the current situation and also using the S1 score (see Verification Module), normally used for forecast verification, to look at the similarity of the gradients in addition to the flow pattern between the two height fields. At CMC the twenty best analogues are chosen. Once the analogues are picked, weather element forecasts including probabilities are then based on the weather elements associated with the analogue cases. The analogue technique is used for medium range forecast only.

1.2.5 Ensemble Prediction Systems (EPS)

Ensembles are sets of forecasts all valid at the same time. As a general rule, these different forecasts may come from many different sources. An example could consist of CMC forecasts together with forecasts from other NWP centres around the world running global models. Another possibility are forecasts with different projection times but valid for the period in question. The CMC Ensemble Prediction System is a 16-member ensemble forecast system. Each day 16 "perturbed" 10-day weather forecasts are performed as well as an unperturbed 10-day forecast. Currently CMC performs 8 forecasts with the global spectral model and 8 forecasts with the GEM model. The 16 model runs have different physics parametrizations, data assimilation cycles and sets of perturbed observations. Observations are perturbed within the limits of observational error. Boundary conditions such as sea surface temperature, albedo and roughness length have been perturbed as well.

Each member of the ensemble has its own set of weather elements for each projection time and these can be combined statistically to produce weather element and probability forecasts.

1.3 Caveats and tips on using statistical forecast products

The following are some general things to consider when doing an evaluation of statistical output. Additional details more specific to individual products will be presented later in the discussion of CMC operational products.

i) An assessment of Model errors. Model errors can be classified into two types:

a) Systematic Model Errors: i.e. biases such as too fast, too slow, too warm, too wet, etc. With MOS approach little or no correction should be applied to correct systematic model errors since MOS has already been adjusted to account (at least partially) for the average systematic error. It is valid to apply corrections to other forecast approaches.

b) Non-systematic Model Errors: i.e. day-to-day departures from the mean error, for example, errors as diagnosed by studies of differences in the model's behaviour in specific synoptic situations. Be alert to rapidly moving systems and tight gradients in the forecast values, a slight shift of the weather producing system could cause a large difference with respect to the forecast output. When possible, the various prog approaches should be modified to reflect these errors.

ii) An assessment of predictors. Not all statistical guidance are derived from single station data. Some use regional equations in which several stations being grouped together to increase the sample size. Hence, in the latter cases, local effects must be considered, e.g. land/sea breeze, mountain winds, fog, soil moisture, sharp fronts, low stratus.

iii) Scepticism of rare events or anomalous situations. Statistical databases are too small to accurately describe these rare events such as tropical storms for example. Continued anomalous situations such as prolonged hot/cold, wet/dry spells where persistence is expected to prevail are also a problem because of the small size of the database. However, there are circumstances where statistical techniques can forecast extreme events.

iv) An assessment of how upstream events are being handled. Watch the edges of a spreading or moving meteorological event. For example, if it is raining at a station with low probability of precipitation, it may also rain at other stations with the same PoP value.

v) An assessment of how the data upon which the equations are base are stratified

vi) An evaluation of unusual predictors and/or post-processing in the regression equations.

vii) An assessment of the time-of-day difference, e.g. the forecast event might occur at times other than the normal time - abnormal temperature cycle, for example.

viii) An assessment of the forecast period. Beyond 36 hours, MOS will in general reflect the climatological mean of the development sample, while the Perfect Prog predictors will only do so if the driving model does. (MOS is more reliable, but Perfect Prog is sharper)

ix) Knowledge of the predictors used, e.g. if last night's minimum temperature is a predictor for tonight's minimum temperature, beware of a changing airmass where abnormal temperature trends are taking place. Statistical output may not be based on the latest reports and may also have a built in lag time.

1.4 Comparative summary of the Perfect Prog and Model Output Statistics Approaches

Model Output Statistics	**Perfect Prog**
partially takes into account systematic errors of driving model	does not take into account systematic errors of driving model
less likely to be affected by model systematic error	likely to be affected by model systematic error
takes into account driving model reliability	does not take into account driving model reliability
model dependent	model independent
without a UMOS system, equations must be redeveloped when driving model is changed	equations should be redeveloped due to evolving climatology
sophisticated predictors	conventional predictors
small sample	large sample
less stable statistical equations	more stable statistical equations
less likely to forecast extremes at longer projection times	more likely to forecast extremes
better reliability	lesser reliability
lesser sharpness	better sharpness
equations required for each projection time	need to develop equations only once per location
makes best use of derived predictors not available to the PPM	more limited availability of predictors
costly	less costly

TABLE 2. The advantages and disadvantages of the Perfect Prog and Model Output Statistics systems

2. Statistical Models

2.1 Multiple Linear Regression (MLR)

2.1.1 Introduction

Regression, and specifically Multiple Linear Regression, has traditionally been the model most used in developing objective statistical forecasts. As in all statistical regression schemes there is a dependent sample of data from which regression equations are developed relating a predictand from within the sample to one or more predictors within the sample. These equations are then applied to current and future values of the predictors to produce forecasts.

Figure 6 illustrates a simple case in which there is only one predictor having a linear relationship to the predictand.

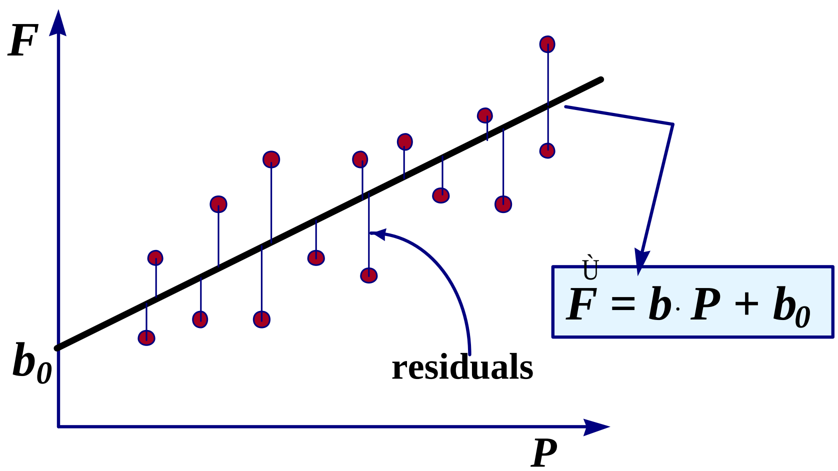


Figure 6. A linear relationship where F is the predictand and P is the Predictor. The circumflex indicates that this is an estimate of F

The more general appearance of a Multiple Linear Regression equation with many predictors takes the following form.

$F = b_0 + b_1P_1 + b_2P_2 + b_3P_3 + \ldots + b_nP_n$

Here, "F" is the value of the predictand which the equation predicts (say, maximum temperature) and " b_0" is a constant. The various P's are the values of the predictors used (an example may be the value of the 1000-500 hPa thickness). The b's in the equation represent coefficients which modify the value that the predictor takes. The coefficients and the constant are determined through a least squares technique which determines the best fit, the line in the above illustration, of the predictor data to the predictand values by minimizing the sum of the squared residuals.

The term “multiple” is used since there can be more than one predictor and the equation is linear since all the terms are linear, that is, there are no higher powers of any of the “P “’s and there are no cross products.

2.1.2 Predictor selection

Each predictor contributes to the determination of the value of the forecast predictand however, they must be carefully chosen. Predictors must be physically related to the predictand, but there are usually many predictors that have a relationship with the predictand. In some cases there may be hundreds. Why not use them all to get the best possible forecast? In fact, this is not mathematically or computationally practical nor is it necessary. An important consideration in the choice is that observational predictors must be reliably available in near real time and forecast predictors must be available with sufficient accuracy to be useful. Another consideration is that if we consider a data sample of N cases from which we wish to compute the regression coefficients, there must be N values of each predictor in the sample so that all coefficients are based on the same number of cases otherwise computational difficulties will arise. This eliminates potential predictors that may be frequently missing in historical data records. The size of the available data sample also limits the number of predictors that can be used. The more variables, the larger the data sample must be to develop stable equations. A stable equation is one that will have a good relationship to independent samples of data such as when it is used to make forecasts. If the sample size is too small for the number of predictors, then we will "overfit" the data. This means that the relationship that we have developed will fit the sample very well but will not hold up on other independent samples. This would be like trying to fit noise or trying to explain variances that cannot be explained.

Selection or screening techniques have been developed to select the predictors used in MLR. One of these is forward selection. The first step in this process is to select the variable that has the highest correlation with the predictand. This is, of course, the one that gives the highest reduction of variance for the predictand. If we were to stop here we would have a simple relationship such as $F = b_0 + b_1P_1$. The next step is to compute all the multiple correlations involving two predictors where P_1 is one of them and then redoing the least squares procedure to generate a new equation. We now have a relationship involving two predictors. The process is continued adding one predictor at a time. At each step we calculate the total reduction of variance and the process is stopped when adding predictors ceases to contribute any meaningful addition to the reduction of variance.

Backward elimination is a process in which all predictors are used to generate the regression equation and then determining which one contributes the least, again using reduction of variance as an indicator, and eliminating it. The equation is then recomputed. The process is repeated until removal of the least helpful variable becomes significant.

Forward stepwise selection is a process that combines forward selection with backward elimination. In this method, at some point in the forward selection process, when adding an additional predictor we test each of the previously chosen predictors for significance. This is done using a standard statistical significance test such as the "Student-t" or F test. It may be that one of the N predictors already selected is no longer significant in an N+1 variable equation and if this is the case it is eliminated. This is therefore a forward-backward process and is a method of looking at combinations of predictors without looking at all possible combinations.

At the Canadian Meteorological Centre, a variation of forward stepwise selection known as "forward stepwise predictor selection with a backward glance" is done. It is normally started by picking the predictor which shows the highest correlation with the predictand. Then without using that predictor, we pick as the second predictor the one that correlates best with the predictand (after what is explained by the previously selected predictor has been removed). A significance test (usually the F-test) is applied to make sure that the new predictor explains a significant portion of the remaining total variance of the predictand. The process is then repeated. It is possible that a predictor that was previously selected is thrown out after a new one has been added if the significance test says that it is not needed anymore. The predictor selection stops when the minimum variance explained by adding a new predictor is not met. This threshold value is adjustable according to the problem we want to solve.

Normally it is better not to exceed something like 8-10 predictors, because experience has shown that the system is then over-fitting the data, which leads to unstable equations that can give us unreliable and sometimes unrealistic forecasts.

2.1.3 Binary variables in regression/Probability Forecasting

Although the discussion up to now has implied continuous variables such as temperature, in meteorology, many events we observe or wish to forecast are not continuous. Thunderstorms either occur or they do not. There will or will not be precipitation in a given time period. Other events may be categorical such as precipitation type which may be liquid, freezing or frozen. Still other events are catagorized numerically such as cloud amount which may be reported in eighths or tenths of total sky cover. Snowfall amounts can be catagorized according to some scale such as none, 0 to less than 5 centimetres, 5 to less than 10 centimetres, 10 to less than 20 centimetres and 20 centimetres and above.

Regression with a binary variable as predictand can be used to estimate the probability of that event occurring. The regression equation is arrived at exactly as with continuous variables since there were no assumptions made on the distribution of the variables in MLR. If a “1” is used to represent the occurrence of the event in question and a “0” is used to represent the non occurrence then the equation will produce a number which can be interpreted as a probability of occurrence. It is essentially an estimate of the relative frequency of the predictand for those times when the predictors have a particular combination of values. It should be noted that there is nothing in MLR that constrains a computed probability to be within the range of 0 to 100%. In practice, values do not generally fall far outside this range and can be constrained through truncation.

The two possibility case illustrated in the previous paragraph is a simple case of what is known as Regression Estimation of Event Probabilities or REEP. For completeness, a more general application can be illustrated using the snowfall categories mentioned above. These four categories are exhaustive and mutually exclusive. This means that all possible occurrences are covered and an event will fall into one and only one of the categories. We can therefore treat this problem as four separate events and develop probability of occurrence equations for each of these events. The equations would have the same predictors and the sum of the four probability estimates should add up to 100%. Since this is the case, one of the four equations is redundant since we can add up three of the results and subtract from 100% to get the fourth. In practice, the use of REEP, is most often limited to binary predictands but it is used at times for multi-category predictands. Multiple Discriminant Analysis (MDA) (See Section 2.2) is a technique that is preferred for multi-category predictands

Binary predictors are also used in regression. As for predictands, there is nothing in the mathematics of regression that would rule out the use of binary variables. Binary predictors are important when we actually have a binary value, such as snow cover, to start with or when we know that the relationship between a predictand and a continuous predictor is non-linear. One way this non-linearity can be addressed is by creating two or more binary predictors from the original variable. Although some information is lost in the process, we can effectively account for highly non-linear relationships.

Unlike predictands, it is not necessary to assign the values of “1” and “0” to binary predictors. It is necessary, however, that they be different. The coefficients that are derived for the regression equations will adapt to the values chosen. In REEP, ones and zeroes as values for binary predictors are commonly used when there are thresholds in the distribution of a predictor such that useful results will not be obtained by using the original numerical values. An example of this could be cloud height as a predictor where only the occurrence of low ceilings is useful in forecasting the predictand in question. In such a case, the use of cloud height as a continuous variable would not give acceptable results but we would achieve meaningful results by assigning a value of “1” to all cloud heights below 500 feet and assigning a value of “0” to all other heights. In the resultant equation this predictor will contribute nothing to the result for high clouds and the value of its coefficient for low clouds. Relative humidity is another parameter which lends itself to this approach. In fact there are many situations in meteorology where such thresholds exist.

An important application where REEP and binary variables are used is in forecasting probability of precipitation and precipitation amounts above certain thresholds. At the Canadian Meteorological Centre, REEP is used to forecast probability of precipitation (PoP) in the perfect prog system over 6 and 12-h periods. The 12-h PoP are for the 0.2mm (trace), 2 and 10 mm levels.

2.2 Multiple Discriminant Analysis (MDA)

2.2.1 Introduction

Like MLR, MDA is a multivariate linear statistical technique. It can be considered a type of regression analysis that classifies the dependent variable into discrete groups based on two or more continuous independent variables. The intent is to discriminate, that is to distinguish between categories of the predictand and in fact is best suited for predictands that can be distinctly catagorized into two or more well defined mutually exclusive and exhaustive groups. MDA is therefore used only for categorical predictands. It is applicable to predictands that have non normal distributions such as precipitation type, clouds and categorical ceilings.

Discriminant analysis is a process where relationships are sought to maximize the ability to make the distinction between groups or categories. The output of MDA is a set of probabilities of category membership. MDA forecasts are sharper than REEP in that the process is more likely to produce forecasts near 0 and 100%.

The choice of the predictand that we wish to forecast bears some consideration. The categories should be distinct enough so that we can easily distinguish between them. For example, it is unlikely that MDA can discriminate between a 100 feet ceiling and a 200 feet ceiling. It can however discriminate between IFR, MVFR and VFR flight conditions. Another example is precipitation type. It is doubtful whether MDA can meaningfully differentiate between freezing rain and ice pellets, but it can discriminate between snow, rain and freezing/frozen precipitation.

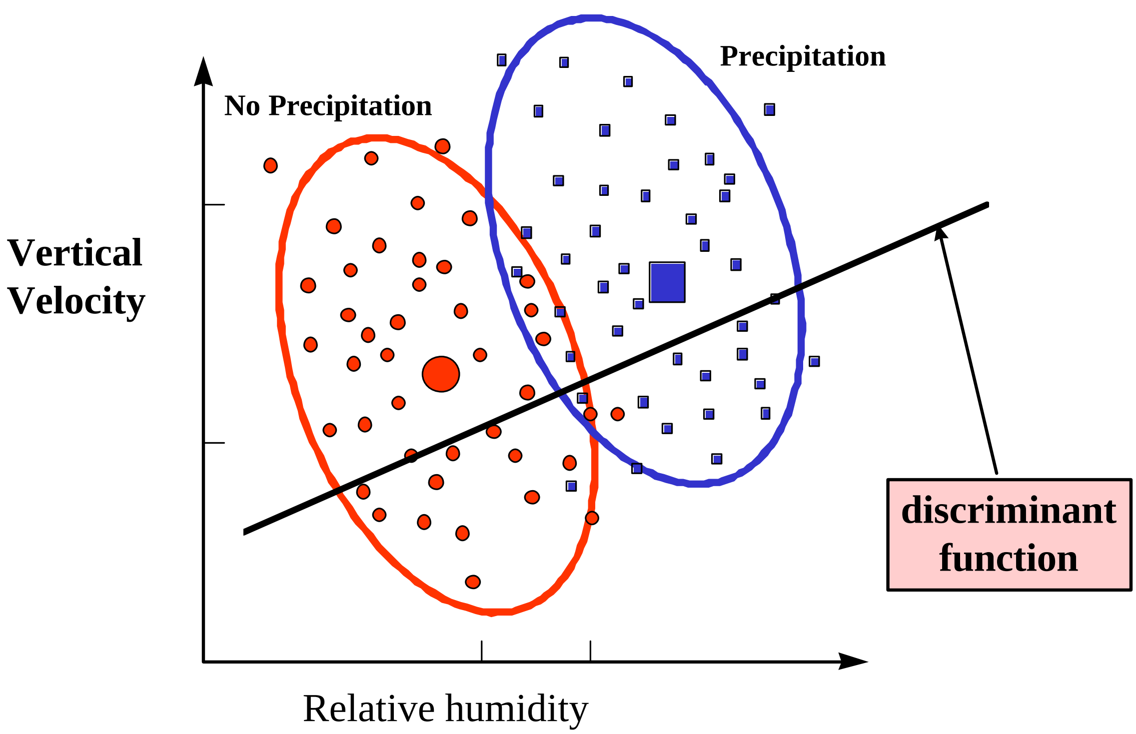


Figure 7. Schematic representation of a scatter plot of precipitation and non precipitation events with respect to vertical velocity and relative humidity

Figure 7 represents a hypothetical case of precipitation events (boxes) and non precipitation events (circles). Ellipses are drawn to enclose a specific percentage of each data set. The large box and circle indicate the category average of the two variables. When using such a diagram for prediction purposes, the choice is easy when the combination of the two variables falls into one of the ellipses outside of the overlap area. In that case the situation is unambiguous. There is ambiguity if the combination falls into the overlap .This ambiguity would be minimized if the category means were farther apart and if the dispersion or variance within each group was smaller. As it is, if we project the data points onto the axes, we can see that the overlap area is relatively large as evidenced by the tick marks on the axes. If a line is drawn parallel to the line joining the two category means and the data points are then projected onto this line, the projection of the data onto that line would produce sharper within group distributions and the overlap would be minimized. Such a line is an example of a discriminant function.

In general there is more than one discriminant function involved. We would normally seek N-1 discriminant functions for N groups or categories. The job of each function is to maximize the between group separation and minimize the within group dispersion. In fact, discriminant functions are obtained by maximizing the between group to within group dispersion ratio. Functions are independent in a statistical sense, meaning that the values of any one function are uncorrelated with the values of another.

2.2.2 Predictor selection

Predictor selection is as important for MDA as it is for MLR. There must, however, be a prescreening where the knowledge of Meteorology is used to decide which predictors to include in the process. There should be a physical relationship to the predictand. In MDA this means that there should be a good basis for the separation of predictand categories. The predictors must be reliable in that there is not too much variation not related to the predictand for similar weather conditions. In other words, the predictor values should be highly repeatable. Predictors should be available in forecast form with a sufficient degree of accuracy to be useful.

Objective screening procedures are used for predictor selection. As in all screening procedures, a test statistic is required to measure the strength of the predictor for the intended purpose and to determine if the addition of a new predictor adds significant value to ones previously chosen. In MDA the test statistic is the Mahalanobis distance. This measures the overall discriminant power of the predictors by determining the separation of the group means while taking into account the variance within the groups. Screening procedures are very analogous to those used in MLR except that the Mahalanobis distance is used as the selection criteria instead of the variance explained to choose the first predictor. A forward or forward selection procedure is used until no predictor increases the Mahalanobis statistic by a significant amount. At CMC a forward selection procedure similar to that used in MLR is used. One consideration, in choosing the threshold at which the procedure is stopped, is the problem at hand and another is the sample size. As in MLR, overfitting can be a problem with too many predictors. This would produce equations which are not stable.

2.2.3 Discriminant Functions

MDA is normally a two step procedure. First, discriminant functions are chosen to sharpen and concentrate the discriminative information into new predictors that are linear combinations of the original predictors. As we have seen the predictors are chosen because they have some discriminating ability to start with. These functions take the following form .

$$Y_j = b_{1j} P_1 + b_{2j} P_2 + b_{3j} P_3 + \ldots + b_{nj} P_n$$

Where Y_j is the j^{th} the discriminant function, The P's are the original predictors and the b's are constant coefficients. The expression is then applied to all the values of the P's in the sample to calculate values of Y. This is usually referred to transforming the variables to discriminant function space.

The new variables are then used, that is they become the predictors, in the second step which is to determine the probability of event occurrence for each category. The advantage of this is that we now have the predictive power of a large number of predictors available in a smaller number of variables in which the discriminant or predictive power is concentrated.

2.2.4 Probability Forecasting

There are two general ways of obtaining probabilities, parametric and non parametric methods.

The non-parametric method makes no assumption about the distribution of the data in the dependant sample. Forecast predictors are first transformed into discriminant function space. The distance from each point in the dependant sample to each function is then computed. This is not a true distance but one which is weighted or skewed by discriminant function to account for the fact that the first function provides the most discrimination, the second provides the next largest amount, and so on. A neighbourhood is defined by selecting a maximum distance. Probabilities are assigned by the proportion of dependant sample points for each category within the neighbourhood. The size of the neighbourhood must be carefully chosen. It must be large enough to contain a representative sample of points from each category but not too large. If it is too large the results will be too smooth.

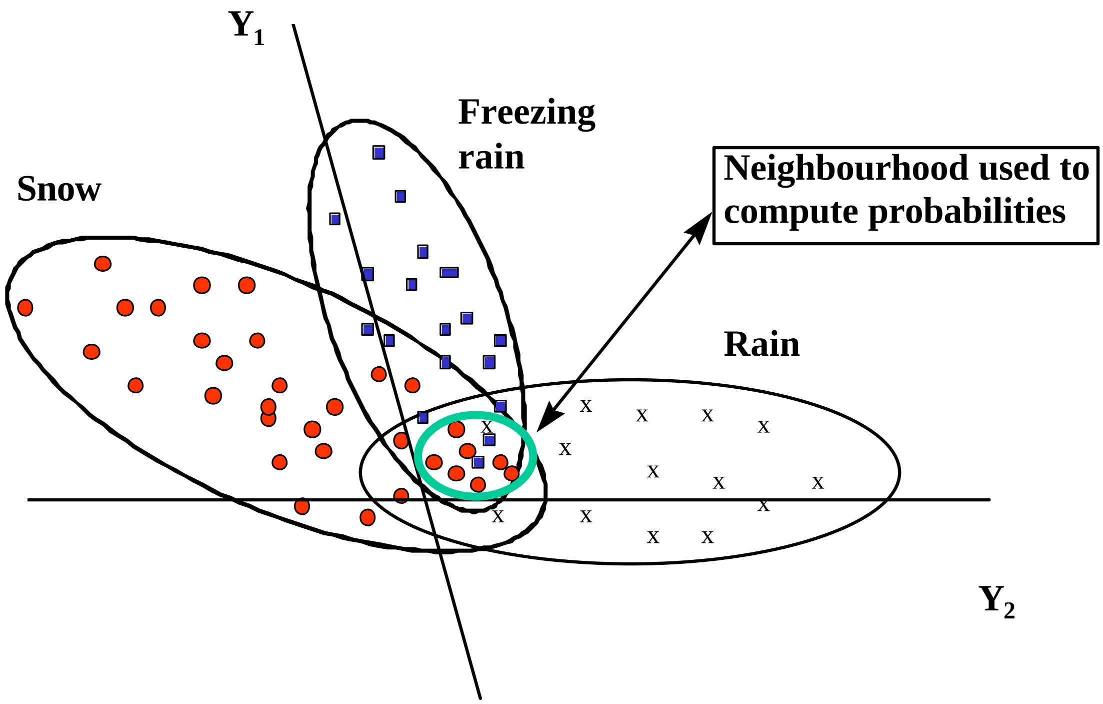


Figure 8. Schematic illustration of a neighbourhood

Figure 8 is an illustration of the use of neighbourhoods. It is a completely hypothetical case. Precipitation, catagorized into three groups, is plotted with respect to two discriminant functions. Discriminant functions are not necessarily orthogonal and the lines Y_1 and Y_2, in this illustration were deliberately drawn to illustrate this. There are ten events within the neighbourhood. Seven of these are snow, two are freezing rain and there is one rain event. Probabilities of 70% snow, 20% freezing rain and 10% rain would therefore be assigned. This illustration is a simple case involving two functions and therefore can be illustrated in two dimensions. In general for N discriminant functions, neighbourhoods have N dimensions. The Y's are not static but are functions of the predictors and are calculated anew for each forecast.

Unlike the non-parametric method the parametric method does not require the use or storage of the dependant sample. This is because the data in the dependant sample are assumed to have specific, usually normal, distributions. Another important assumption is that the within group dispersion is the same for all groups or categories. This means that the variance of the dependant sample discriminant function values must be the same for all groups. Differences in size of the categories are not a problem. If we were looking at rain versus no rain situation , there would normally be a lot more no rain than rain cases. The assumption about the distribution of the predictor simply requires that the standard deviation of the predictors be the same for both categories.

Probabilities are arrived at through the use of Baye's Theorem which states that the probability of group membership is the a priori probability of group member ship times the likelihood divided by the sum of all the products of a priori probability group membership and likelihood's for all the groups. The a priori probability is simply the relative frequency of occurrence in the dependant sample and the likelihood's are obtained from predictor values. As before the predictors are transformed into discriminant function space. The normal distribution means and variances from the dependant sample are then used to estimate the likelihood's.

2.3 Considerations in choice of statistical model

There is a large number of statistical models used in various scientific and engineering applications. Although MLR and MDA are the current statistical models of choice in weather forecasting, there are other models that are seeing some use or are under investigation for their application in operational meteorology. These include Classification and Regression Trees (CART), Canonical Correlation Analysis (CCA), neural networks, Kalman Filters and fuzzy logic.

The choice between MLR and MDA depends on the predictand itself. Normally MLR is well suited for predictands that are continuous and that have a climatological frequency distribution that is close to normal such as temperatures. Wind can also be a good candidate for MLR and probability of precipitation (in the REEP sense). Although REEP can be used, MDA is a better choice for forecasting precipitation type. Predictands, whose distribution is far from normal, such as cloud amounts are better forecast by MDA techniques. Clouds have a U-shape frequency distribution which is considerably different from a normal distribution. MLR cloud forecasts will have a normal distribution and the problem will be that the frequency of middle cloud values (4-7 tenths of sky cover) will be largely over forecast. The system will also largely underestimate the frequency of the extreme values.

Multiple Linear Regression shares with many other regression techniques the problem that it works best under conditions which are close to the long term average. Because of how the forecast equations are arrived at, MLR tends to pull forecasts close to the climatic mean of the development sample and tends to underforecast "extreme" events although it can do so at times. When sharpness is required to better forecast extremes, MDA may be the better choice.

A limitation of the MLR technique is that there is an assumption that the predictand has a linear relationship to the predictors, but it may not. Perhaps the predictand varies as the square or the logarithm of some predictor. The process of Multiple Linear Regression "forces" the two parameters to assume a linear relationship. However, it is possible up to some point to linearize the relationship between the predictor and the predictand and get meaningful results.

3. Quality Control

3.1 Forecast Consistency

Irrespective of whether we are basing our forecasting system on the PP or MOS system or whether we use MLR, MLR with REEP, MDA or any other statistical model, all results of statistical analyses have to be checked carefully for coherence. Statistical procedures will produce equations and some indication of the quality of the equations in terms of the percentage of variance explained in the dependent sample, but there is no meteorology involved, other than the initial selection of predictors presented to the process. It is important to verify that each chosen predictor has a valid physical relationship with the predictand. If a predictor is chosen that violates this, then it should be eliminated and the equations redone. Another possibility is when a valid predicator is used in an inappropriate or unexplainable fashion. An example of this would be the choice the 850-500 thickness in an equation to forecast surface temperature, but with a negative coefficient. This is a strong indication that the statistics is over-fitting the data, and here again the equation should be redone without that predictor.

We must also ensure consistency between all different weather elements based on statistics. Since the equations are generated independently, no consistency is ensured. For example, the PoP over 6 and 12 hours may very well be inconsistent in that there is a 100% PoP forecast for a given 6 hour period while a 12 hour PoP that includes that 6 hour period is less that 100%. Spot time temperature forecasts are derived independently from the maximum and minimum for the day. They need to be checked to ensure that they are within these limits. Another example is the temperature forecast which does not know whether the PoP produced at the same time is high or low. This can be a problem since occurrence of precipitation has an impact on the temperature.

Consistency in space and time must also be addressed. In MOS, there are separate equations for each projection time and in the PP system there are separate equations for each observation time. When separate equations are developed independently carrying out the screening procedure each time, there is no guarantee that equations for the same forecast element will use the same predictors each projection time nor is there any guarantee that forecasts for neighbouring stations will use the same predictors for the same valid times. In fact, in some cases, local conditions would dictate otherwise.

We rely in part on the internal consistency of the predictors provided by the driving model. The only way, however, to ensure consistency is to develop everything at once all together. This is not possible so there must be some analysis and quality control procedures used to ensure consistency.

3.2 Post Processing Techniques

In addition to correcting forecasts to ensure consistency, there is other post processing that is carried out. Post processing, in this context, refers to any systematic procedure of altering the output of statistical models before it is sent to the users. MOS and Perfect Prog systems, MLR , MDA and other techniques each have their limitations together with desirable and undesirable attributes. Post processing techniques are employed to try and correct for the undesirable characteristics. Different techniques are used to address different problems.

3.2.1 Inflation

Inflation is a means of altering the forecasts such that the extremes of the distribution are forecast more often. It has the effect of pushing the forecast away from the mean. Regression equations tend to forecast toward the mean, resulting in a narrow distribution. Inflation is used to try and make the distribution of the forecasts match the distribution of the observed weather element. Inflation is also employed when extremes can be significant events. An example of this is wind forecasts. Strong winds are considerably rarer than moderate winds but are relatively more important when they do occur.

3.2.2 Deflation

Deflation, as the name implies, is the opposite of inflation. Its aim is to reduce the spread of the predictand and move it closer to the mean. It is often employed in perfect prog techniques, which apply the same relationships at all projection times, to account for decreasing accuracy of the driving model with increasing projection times. A positive attribute of PP is sharpness or the ability to forecast extremes. Forecasting extremes, however, based on extreme predictors which may not be accurate will produce false alarms and is certainly not desirable. Applying a temperature anomaly reduction, for example, which increases with projection times is one form of deflation.

3.2.3 Bias Correction

Bias correction is simply correcting for the mean error in the forecast. The bias must be determined from an independent sample since there is no bias in the dependant sample from which the equations were derived. The biases are usually stratified by projection time, season, geographical area, size and sign of the anomaly and other factors. It is important that the sample used to determine the biases be large enough to be representative of the population of forecasts. Bias correction can only be applied, therefore, once a sufficient number of forecasts has been made and verified. Once determined, the bias correction is applied to all future forecasts.

Bias correction and inflation/deflation can be done in a single step through the use of regression. When done this way, the technique is referred to as calibration.

3.2.4 Real time error feedback methods

Unlike bias correction, real time error feedback only uses recent history of the forecasts and observations to compute the errors. There is an assumption of persistence of errors at least over the period during which error data are collected. This period is normally smaller than for bias correction but must be large enough to be statistically significant. It is a recursive technique which continuously updates itself and thus can adjust to errors that are correlated in time. The error feedback can be as simple as subtracting the bias or a fraction of it over the last few days or the correction can be determined by assigning weights to the errors with the most recent having the largest weights. More sophisticated techniques can be used if it is known that the error is correlated to the predictand or its anomaly. In such a case, regression techniques can be used to determine the correction.

The Kalman filter is another technique that can be used for real time error feedback. It essentially models the errors based on prechosen parameters. One such parameter could be the anomaly or departure from the normal of the forecast. It then establishes the error characteristics of the forecast based on the model and the results of recent forecasts and then applies these to the new forecast. The Kalman filter essentially diagnoses systematic errors. It is a recursive technique which is continuously updating itself and is also based on the persistence of errors. Unlike other error feedback techniques which are basically maintenance free once implemented, Kalman filters require tuning and need to be restarted once in a while.

3.2.5 Model Blending

This is a different approach to post processing than what has been discussed above. Based on the assumption that different sources or forecasts of the same element contribute some new information about the prediction, the combination of forecasts from different sources can result in an overall result which is better than that of any one of its contributors by themselves. Different sources may be products of the same statistical techniques but driven by different Numerical Prediction Models. A more likely source of different forecasts would be direct model outputs, MOS and PP forecasts.

Model blending can be done in several ways. The simplest is to average them but this would generally just have a smoothing effect. It is much better to use a flexible approach in which rules or contingency tables are developed to choose one of the forecasts in any given situation. The rules must take into account known strengths and weaknesses of each of the component techniques. Alternatively, a linear combination of the forecasts can be done using weights based on physical reasoning and/or verification results. This can be carried one step further by using regression with the component forecasts as predictors. No matter how the blending is done, the objective is to maximize the use of available predictive information and minimize the effects of the deficiencies of the statistical models.

3.2.6 Best Category Selection

Probabilistic forecasts for multi-category predictands lead to the problem of determining which of the possible categories should be chosen when the users want to make a decision. There are practical reasons why we would go to the trouble of determining the probabilities of the individual categories and then turn the result into a categorical forecast. One of these is that the final result that the forecaster is required to produce is categorical. Another is to simplify the assessment of a large number of probability forecasts for many neighbouring stations at many projection times. Another reason may be to help the user adjust the forecast relative to climatology. An example of this is making a decision on a forecast probability of 70% when the climatology of the occurrence of that event is 5% relative to the decision that would be made if the climatology was 50%.

The choice of a decision strategy depends on the situation and the use or objective of the forecast. There are several options; four of which are presented here. The simplest option, which is the Maximum Probability Model, is to choose the category with the highest probability. In doing so, rare events would tend to be ignored. The Climatology Model chooses that category with the highest positive difference between the forecast probability and the climatological frequency. This technique has the characteristic of overforecasting the rare or less common events at the expense of a higher number of false alarms. The Unit Bias Model sets probability thresholds such that each category is forecast with the same frequency at which it is observed. It is designed to neither over forecast nor under forecast extreme events. The Maximum Threat Model is one designed to maximize the Threat score. Its intention is to allow some forecasting of rare events but more conservatively than the Climatology Model.

4. CMC Operational Statistical Products

4.1 Introduction

The Canadian Meteorological Centre currently operates a single Numerical Prediction Model which is the Global Environmental Multiscale (GEM) model. GEM is a global model with a variable grid which can be rotated and the grid spacing can be non uniform when required. This allows the same model with different grid configurations to be used for different applications. Operational applications presently include short-range regional forecasting, medium-range global forecasting, and data assimilation cycles.

GEM is formulated in terms of the hydrostatic primitive equations with a terrain following pressure vertical coordinate (*eta*) and is interfaced with a full complement of physical parametrizations to take sub-scale phenomena into account.

Global medium and long-range forecasts are produced by a global version of GEM with the following properties: (properties in force December 2000)

- uniform lat-lon grid with 0.9 degree resolution,

- 28 *eta* vertical levels,
- fed by a global variational analysis data assimilation scheme,
- 72 hour forecasts at 12 UTC,
- 240 hour forecasts at 00 UTC,
- 360 hour forecasts at 00 UTC on Saturday.

Short-range regional forecasts are produced by a regional configuration of GEM with the following characteristics: (properties in force December 2000)

- non-uniform horizontal grid with 24km resolution over a central window covering North America and adjacent waters,
- grid rotated at approximately 90 degrees,
- 28 *eta* vertical levels,
- fed by a regional variational data assimilation system,
- 48 hour forecasts twice daily at 00 and 12 UTC.

The two main statistical techniques that are used at CMC, and driven by GEM in both regional and global configurations, are Perfect Prog and updatable MOS (UMOS). Although, at the time of writing (December 2000), the MOS products were not yet fully operational, they are available on the CMC internal web pages and are presented here for completeness. All statistical forecasts are based on Multiple Linear Regression Model (MLR) except for the PoP and opacity forecasts for days 3, 4 and 5 which are based on an analogue approach. Research and development is underway to add Multiple Discriminant Analysis (MDA) techniques and to replace the current days 3, 4 and 5 analogue products by forecasts based on Ensemble Prediction Systems.

4.2 Operational Products

4.2.1 Probability of Precipitation (PoP)

Technique 6-hour PoP:- Probabilities of precipitation occurring in each of the 6 hour periods of the forecast projection time are forecast by both the perfect Prog and UMOS systems. Precipitation is considered to have occurred if the accumulation reaches or exceeds a trace (0.2mm). The predictand is the observed precipitation accumulation, as reported in synoptic observations, converted to binary form, 0 if less that a trace and 1 for a trace and more.

The PP regression equations were developed from 22 years of historical data (1963-1984) stratified into 4 three month seasons. The selection of predictors varies with season, geography and projection time with the seasonal variation being the most significant. As one would expect, humidity either alone or combined with thickness anomalies or thickness advection is an important predictor.

Six hour PoP's are also produced using UMOS. In the UMOS technique, the most important predictor is the model precipitation forecast. As a general rule the first predictor chosen in the UMOS predictor screening is the model forecast of the predictand in question.

Technique 12-hour PoP:- Twelve hours PoP's are forecast for three thresholds, 0.2, 2 and 10 millimeters. The predictand is the total observed precipitation accumulations over two consecutive 6 hour periods converted to binary form for each of the three categories. PP linear regression is used with the same historical or training period as the 6 hour PoP's again stratified into three month seasons with the exception of the 10 mm threshold where there are two 6 month seasons due to the fact that this is a rarer occurrence.

Forecast range:- PP forecasts are prepared for either 6 or 12 hour intervals out to 48 hours based on the regional GEM and out to 72 hours at 12 UTC and 144 hours at 00UTC based on the global run of GEM.

Post processing:- Internal consistency between the 6 and 12 hour PP PoP forecasts based on the regional run is forced according to a rule based system and the adjustments are done using an iterative technique. For each 12 hour period, the two 6 PoP's are combined and compared to the corresponding 12 hour PoP and inconsistencies eliminated. Consistency is also assured between the forecasts for the different threshold levels of 12 hour PoP. Depending on where the forecasts fall compared with predetermined values, they are inflated toward the extreme values (0 or 100%). The end result of this scheme are significantly sharper forecasts without deterioration in skill.

Forecasts based on the global GEM are processed using a similar scheme, except that the biases of the forecasts are first removed. Biases are calculated for the previous 90 days and are site specific. They are updated at each production cycle for each projection time.

There is no post processing applied to the UMOS products.

Caveats:-

As a general rule, when comparing perfect Prog and MOS forecasts one must remember that PP forecasts are sharper while MOS tends to forecasts that are closer to the driving model's climatological mean and is less likely to forecast extremes at longer projection times. This is particularly true for longer projection times. PP forecasts are also more stable because of the longer development sample while MOS is more reliable.

Although model performance has improved considerably over the past decades, one must still be alert for errors in timing, location and system intensity as well as systematic model errors. Model errors are reflected in the statistical forecasts that are derived from them.

Systematic model biases, unless removed in post processing, should be taken into account.

Abnormal circulations will not be handled well since these are not likely to be reflected in the development samples.

There can be problems with boundary layer precipitation

Since regression equations are developed independently for each station, there may be spatial discontinuities.

Interpretation of data in regions of strong gradients, or in areas of rapid change because of rapid movement or development of systems can lead to errors in the statistical forecasts.

There may be discontinuities that occur when the systems "seasons" change in the PP system. The UMOS blending technique smoothes these change periods.

There can be inconsistencies with model quantitative precipitation amount forecasts (QPF).

Forecast Distribution - Bulletins

PP forecasts from the regional model are distributed as part of FXCN09 CWAO1 to 11. These are documented on the CMC internal web site's library under products. Decoding information and a sample message are provided there.

UMOS forecasts are distributed in the FXCN50 bulletin. Forecasts are available for 637 stations.

Global model forecasts are distributed as part of FXCN05 CWAO1 to 11.

Forecast distribution - CMC Internal Web site

Up-to-date FXCN5 and FXCN09 bulletins containing PP PoP forecasts can also be found on the CMC Development Branch's internal web page under weather elements (CMDW) in a menu driven data base.

The MOS products are also available on the CMC Development Branch's internal web page.

4.2.2 Cloud opacity

Technique used:- Cloud opacity is forecast for specific sites in tenths of sky cover as reported in hourly observations. The predictand is time specific and is not averaged over any time period. PP is the technique used and the regression equations were developed from 22 years of historical data (1963-1984) stratified into 4 three month seasons.

Forecast range:- Forecasts are prepared for every three hours out to 48 hours based on the regional GEM and out to 72 hours at 12 UTC and 144 hours at 00UTC based on the global run of GEM.

Post processing:- In the forecasts based on the regional GEM run, internal consistency is ensured between the cloud opacity and corresponding 6 hour PoP forecasts according to a rule based system. Forecasts are inflated towards the extremes, 0 and 10 tenths.

Forecasts from the global run are forced to reproduce the typical U-shaped frequency distribution of the observations by relableling the forecasts according to a set of threshold values calculated by projecting the cumulative observational frequency distribution onto that of the forecast distribution.. The cumulative frequency distributions, both observed and forecast are calculated for the previous 90 days and are updated for each station and projection time at each production cycle.

Caveats:- Caveats are similar to those for PoP forecasts with the addition of the following.

Linear regression does not allow the reproduction of the typical observed U shape frequency distribution for clouds. There is a tendency to forecast near median values and to underestimate the frequency of near extreme values

There are difficulties in forecasting clouds in abnormal circulations which are not represented sufficiently in the dependent sample.

Low clouds (boundary layer) are difficult to forecast.

Forecast distribution - Bulletins

Opacity forecasts from the regional model are distributed as part of FXCN09 CWAO1 to 11. These are documented on the CMC internal web site's library under products. Decoding information and a sample message are provided there.

Global model forecasts are distributed as part of FXCN05 CWAO1 to 11.

Forecast distribution - CMC Internal Web site

Up-to-date FXCN05 and FXCN09 bulletins containing cloud opacity forecasts can also be found on the CMC Development Branch's internal web page under weather elements (CMDW) in a menu driven data base.

4.2.3 Spot Temperatures

Technique:- Spot time temperatures at three hour intervals are forecast by both the perfect Prog and UMOS systems. The predictand is the observed surface temperatures as reported in the hourly observations minus the climatological value.

The PP regression equations were developed from 22 years of historical data (1963-1984). stratified into 6 two month seasons. For temperature forecasts, the main predictor is usually thickness between various levels in the atmosphere.

Three hour spot temperatures are also produced using UMOS. In the UMOS technique, the most important predictor is the model surface temperature forecast. As a general rule the first predictor chosen in the UMOS predictor screening is the model forecast of the predictand in question.

Forecast range:- Forecasts are prepared at every three hours out to 48 hours based on the regional GEM by both the PP and UMOS methods and out to 72 hours at 12 UTC and 144 hours at 00UTC based on the global run of GEM using PP.

Post processing:-

An anomaly reduction scheme is applied to all PP temperature forecasts beyond 18 hours to account for decreasing predictability of model forecast predictors with increasing projection time. This reduction is linear with respect to projection time and is limited to a maximum of 45% of the forecast anomaly.

There is no post processing of the UMOS product.

Caveats:-

Although model performance has improved considerably over the past decades, one must still be alert for errors in timing, location and system intensity as well as systematic model errors. Model errors are reflected in the statistical forecasts that are derived from them.

Systematic model biases, unless removed in post processing, should be taken into account.

Statistical methods have built in weaknesses as indicated by the standard error of estimate. They have a tendency to forecast towards the climatologial mean. Where significant local effects predominate, such as valley, mountain and sea breeze circulations, the forecasts become less reliable.

The diurnal cycle is generally underestimated.

Abnormal situations and circulations will not be handled well since these are not likely to be reflected in the development samples.

Since regression equations are developed independently for each station, there may be spatial discontinuities.

Interpretation of data in regions of strong gradients, or in areas of rapid change because of rapid movement or development of systems can lead to errors in the statistical forecasts.

There may be discontinuities that occur when the systems “seasons” change every two months.

The temperature regression equations were prepared to fit all the data. For example, a maximum of 30^0C may occur with both low and high relative humidity. Sky cover and precipitation are not used explicitly and can be important factors. Given that the model is forecasting the synoptic situation correctly, on a cloudy or rainy day the temperatures are usually forecast too high during the day and too low at night. The reverse is true on sunny days. Temperature forecasts may have to be adjusted accordingly.

Nighttime temperatures should be adjusted upwards in situations where cloud cover and strong winds are expected, especially in above normal temperature regimes.

The valley effect can be underestimated in spite of the fact that predictors are selected to take this into account.

Forecast Distribution - Bulletins

PP spot temperature forecasts from the regional model are distributed as part of FXCN09 CWAO1 to 11. These are documented on the CMC internal web site’s library under products. Decoding information and a sample message are provided there.

Global model forecasts are distributed as part of FXCN05 CWAO1 to 11.

UMOS forecasts are distributed in the FXCN50 bulletin. Forecasts are available for 637 stations.

Forecast distribution - CMC Internal Web site

Up-to-date FXCN5 and FXCN09 bulletins containing PP spot temperature forecasts can also be found on the CMC Development Branch’s internal web page under weather elements (CMDW) in a menu driven data base.

The UMOS products are also available on the CMC Development Branch’s internal web page.

4.2.4 Maximum/Minimum Temperatures.

Technique:- Maximum and minimum temperature forecasts are done for the climatological day (06 to 06 UTC). The predictand is the observed maximum/minimum temperatures as reported in the synoptic observations minus the climatological value. The predictand is therefore the observed temperature anomaly which is then added to the normal at the end to produce the final maximum/minimum temperature forecast

The PP regression equations were developed from 22 years of historical data (1963-1984). stratified into 6 two month seasons.

Forecast range:- Maximum and minimum temperatures are forecast out to day 2 based on the regional GEM run and out to day 3 based on the global GEM run at 12 UTC and to day 6 at 00 UTC.

Post processing:- An anomaly reduction scheme similar to the one used for spot temperature forecasts is used to process the maximum/minimum temperature forecasts based on the regional GEM model. There is a bias correction applied to the forecasts based on the global run of GEM.

Caveats:-

Caveats are basically the same as for spot temperatures with the following additions.

Forecasts are for the climatological day which runs from 06 to 06 UTC. Maximum/minimum temperatures are more closely related to the true local day which varies by latitude and time of year.

Regression equations for maximum/minimum temperatures during the climatological day often use predictors at 12 UTC for the minimum and 00 UTC for the maximum temperature. Synoptic events in which cooling occurs all day or begins early in the day such that the maximum occurs earlier in the day than normal are therefore not well handled.

Forecast Distribution - Bulletins

Forecasts based on the regional GEM are distributed in the FMCN41 to 48 bulletins which contain three forecast temperatures. At 00 UTC, the bulletins contain the maximum for day 1 and the maximum and minimum for day 2. At 12 UTC, the maximum and minimum for day 2 and the minimum for day 3 are transmitted.

Forecasts based on the global Gem run are distributed in the FMCN31 to 38 bulletins which contain four forecast temperatures. At 00 UTC the bulletins contain the maximum for day 1, the maximum/minimum for day 2 and the minimum for day 3. At 12 UTC, they contain the maximum/ minimum temperatures for days 2 and 3.

The FXCN07 CWAO1 to 12 bulletins contain forecasts from the GEM run with maximum/minimum temperatures for days 3, 4 and 5 at 00 UTC and for day 3 only at 12 UTC.

All three bulletin types are documented on the CMC internal web site's library under products. Decoding information and a sample message are provided there.

Forecast distribution - CMC Internal Web site

Up-to-date FMCN31 to 38, FMCN41 to 48 and FXCN07 CWAO1 to 12 bulletins can also be found on the CMC Development Branch's internal web page under weather elements (CMDW) in a menu driven data base.

4.2.5 Surface winds

Technique:- Operational spot forecasts of surface wind direction and speed at 6-hour intervals, from T+6 to T+48 are produced based on the regional configuration of GEM using PP linear regression. Wind forecasts are also produced using UMOS. There is no time or spatial mean calculated from the statistical output. The forecast wind direction and speed are thus interpreted directly as spot direction and speed, at the valid time of the forecast.

There are three predictands. These are the observed West and South components of the wind (U and V) which are used to calculate the wind direction and the wind speed which is forecast directly. The wind direction is presented in tens of degrees with respect to North, and speed is expressed in knots. Winds less than 5 knots are indicated by "LGHT".

The regression equations are based on eight years of data (1974 to 1982), stratified into three seasons, summer, winter and a spring/fall transition period. Data from both 00 and 12 UTC were merged into one sample for the development of the equations and there is therefore one equation per day in the system.

Forecast range:- The system forecasts winds at six hour intervals out to 48 hours**.**

Post processing:- Wind speed forecasts are calibrated on the basis of the regression line of observed versus forecast speeds. The slope of this line is adjusted in order to inflate the forecasts. This corrects the bias and tunes the correction to the characteristics of each station and projection time. Wind direction is treated for bias only. A correction equal to the bias is added to the statistical forecasts after the direction is computed from the U and V components. This is a simpler procedure than trying to correct the U and V components.

Caveats:-

Previous comments in the discussions of other weather element forecast with respect to model errors, difficulties in abnormal circulation and spatial discontinuities apply here as well.

The development data set did not contain any data from 06 and 18 UTC. Therefore these forecasts may be a little less reliable that those for 00 and 12 UTC.

The diurnal variation has been modeled with time dependent predictors but using two points only. In areas where the 00 and 12 UTC observations are near the minima of the diurnal wind curve, the modeled diurnal cycle may not perform reliably.

Since the data were stratified into three seasons, one should be alert for larger errors when the weather situation departs form the normal for the season.

The equations were developed using 1000 hPa as the lowest level with a coarse vertical resolution. Sharp and shallow inversions will not be well handled by the system.

Mesoscale phenomena may not be well handled since the resolution of the dependent data set was 381 kilometers. Such effects as sharp wind shifts will have to be added to the forecasts.

Large errors may occur near the centres of lows, particularly rapidly developing or fast moving systems, since small model errors in the track will produce large wind forecast errors.

Forecast Distribution - Bulletins

Forecasts are transmitted in the FOCN10 CWAO1 to 9 bulletins. These are documented on the CMC internal web site's library under products. Decoding information and a sample message are provided there.

UMOS forecasts are distributed in the FXCN50 bulletin. Forecasts are available for 637 stations.

Forecast distribution - CMC Internal Web site

Up-to-date FOCN10 bulletins can also be found on the CMC Development Branch's internal web page under weather elements (CMDW) in a menu driven data base.

The UMOS products are also available on the CMC Development Branch's internal web page.

4.2.6 24 hour probability of Precipitation - Days 3, 4, 5

Technique:- The system forecasts the probability that precipitation accumulations will reach or exceed the 0.2 mm threshold during each 24 hour period (06 to 06 UTC) over the projection period using an analogue approach. The analogue search is made using the 60, 84 and 108 hour forecasts at both 1000 and 500 hPa from the global GEM run for days 3, 4 and 5. The 24 hour PoP's are produced by using the observed precipitation frequencies for the 20 best analogues within the 28 year (1957 to 1984) database.

Forecast Range:- Days 3, 4 and 5

Post processing:- An anomaly reduction is applied to the forecasts for days 4 and 5. A 10% reduction is applied at day 4 and a 20% reduction is applied at day 5.

Caveats :-

Forecasts are based on a relatively short historical database.

The pattern recognition techniques are somewhat limited.

Model accuracy decreasing with projection time is more of a concern than with shorter range forecasts.

The technique best suited to large scale circulation patterns so there are problems in mesoscale and convective situations.

The 24-h precipitation occurrence is related to the flow pattern at 12 UTC only.

There are no tests on temporal consistency.

Forecast Distribution - Bulletins

Forecasts are transmitted in the FXCN07 CWAO1 to 12 bulletins. Forecasts issued at 12 UTC contain the forecasts for day 3 only. These are documented on the CMC internal web site's library under products. Decoding information and a sample message are provided there.

These forecasts are used to prepare the plain language forecasts for days 3, 4 and 5 transmitted in the FOCN 12 and 13 CWAO1 to 12 bulletins.

Forecast distribution - CMC Internal Web site

Up-to-date FXCN07 and the plain language forecast bulletins can also be found on the CMC Development Branch's internal web page under weather elements (CMDW) in a menu driven data base.

4.2.7 Sky cover - Days 3, 4, 5

Technique:- The system forecasts the total opacity between sunrise and sunset for days 3, 4 and 5 using the analogue approach. Daylight, North of 60N, is defined to be between 7:00 AM and 7:00 PM local time. Averaged cloud opacity at stations for the 20 best analogue dates is used to produce the forecast. If necessary, observed possible sunshine is converted to opacity for this calculation.

Forecast Range:- Days 3, 4 and 5

Post processing:- A very simple relabelling scheme is applied to the sky cover forecasts. It consists of making all forecasts of 3 tenths equal to 2 tenths and all forecasts of 8 tenths equal to 9 tenths. The forecast frequency distribution then matches the observed distribution more closely.

Caveats :-

Forecasts are based on a relatively short historical database.

The pattern recognition techniques are somewhat limited.

Model accuracy decreasing with projection time is more of a concern than with shorter range forecasts.

The technique best suited to large scale circulation patterns so there are problems in mesoscale and convective situations.

The sky cover forecast is related to the flow pattern at 12 UTC only.

There are no tests on temporal consistency.

Forecast Distribution - Bulletins

Forecasts are transmitted in the FXCN07 CWAO1 to 12 bulletins. Forecasts issued at 12 UTC contain the forecasts for day 3 only. These are documented on the CMC internal web site's library under products. Decoding information and a sample message are provided there.

These forecasts are used to prepare the plain language forecasts for days 3, 4 and 5 transmitted in the FOCN 12 and 13 CWAO1 to 12 bulletins.

Forecast distribution - CMC Internal Web site

Up-to-date FXCN07 and the plain language forecast bulletins can also be found on the CMC Development Branch's internal web page under weather elements (CMDW) in a menu driven data base.

Summary CMC Operational Statistical Products as of December 2000			
Weather Element	**Bulletin**	**GEM Configuration**	**Technique**
6 hour PoP	FXCN09	Regional	PP linear regression with inflation
	FXCN5	Global	PP linear regression with error feedback and inflation
12 hour PoP	FXCN09	Regional	PP linear regression with inflation
	FXCN5	Global	PP linear regression with error feedback and inflation
Opacity	FXCN09	Regional	PP linear regression with inflation
	FXCN5	Global	PP linear regression with error feedback
Surface Temperature	FXCN09	Regional	PP linear regression with anomaly reduction
	FXCN5	Global	PP linear regression with anomaly reduction
Max/min Temperatures	FMCN41-48	Regional	PP linear regression with anomaly reduction
	FMCN31-38	Global	PP linear regression with anomaly reduction
	FXCN07 (Days 3, 4, 5)	Global	PP linear regression with anomaly reduction
Surface winds	FOCN10	Regional	PP linear regression with calibration
24 hour PoP (Days 3, 4, 5)	FXCN07	Global	Analogue with anomaly reduction
Sky cover (Days 3, 4, 5)	FXCN07	Global	Analogue with relabelling

Table 3: Summary of CMC operational statistical forecast products as of December 2000

5. Acronyms

Acronyms	
CART	Classification and Regression Trees
CCA	Cannonical Correlation Analysis
CMC	Canadian Meteorological Centre
EPS	Ensemble Prediction System
GEM	Global Environmental Multiscale (model)
hPa	hecto Pascal
IFR	Instrument Flight Rules
MDA	Multiple Discriminant Analysis
MLR	Multiple Linear regression
MOS	Model Output Statistics
MVFR	Marginal Visual Flight Rules
NWP	Numerical Weather Prediction
PoP	Probability of Precipitation
PoPA	Probability of Precipitation Amounts
PPM	Perfect Prog Method
QPF	Quantitative Precipitation Forecast
REEP	Regression Estimation of Event Probabilities
UMOS	Updatable Model Output Statistics
UTC	Universal Time Constant (Greenwich Mean Time)
VFR	Visual Flight Rules