Objective Guidance for 1- and 2-Day Mesoscale Forecasts of Lake-Effect Snow

W.A. BURROWS

Atmospheric Environment Service
Meteorological Services Research Branch
4905 Dufferin Street
Downsview, Ontario M3H 5T4, Canada

ABSTRACT

Lake-effect-snow-possible (LESP) days were identified for each of 29 climatological stations in the lee of Lake Huron and Georgian Bay for the November to April winters of 1984-1988 using output from 0-24 hour and 24-48 hour forecasts by the Canadian Meteorological Center (CMC) operational spectral Numerical Weather Prediction (NWP) model. Twenty-four hour observed snow amounts on LESP days were separated into 5 ordered categories. The distribution of snowfall on LESP days for the aggregate of the 29 stations is centered on category 3, but the distributions for individual stations are centered on categories 1 or 2. This suggests lake-effect snow occurrence and amounts are likely to be over-forecast by meteorologists for even relatively small public forecast areas.

A recent non-parametric classification procedure known as "Classification and Regression Trees (CART)" was used to classify the categorical snowfalls by predictor threshold values in decision trees. Predictors were designed from the physics of lake-effect snow formation and calculated from the NWP model output data on a 63 km interpolation grid. Predictors most frequently used by CART to create the decision trees involved low-level divergence (convergence) at nearly every station, followed by predictors related to air-water temperature difference. The method shows considerable promise for timely many-site production of objective operational mesoscale guidance for 1 and 2, day forecasts of 24-hour lake-effect snow accumulation.

1. INTRODUCTION

Lake-effect snowfall is one of the challenging and important mesoscale forecast problems around the Great Lakes. On occasion a great deal of snow falls in a 24 hour period over small areas, with wide variation in amounts over relatively small distances. It has been estimated that as much as one-half the annual snowfall around the shores of Lake Michigan is due to lake effects (Braham and Dungey, 1984). A similar or even greater fraction appears to be the case around Lake Huron and Georgian Bay as well (Figure 1).

Peace and Sykes (1966) concluded that "while the formation of lake-effect bands is caused by heating of the air by a warm lake, the location and movement of snow bands are controlled by winds aloft". Thus lake-effect snow, while it is a mesoscale phenomenon, is strongly controlled by characteristics of the synoptic scale environment in which it occurs. This has enabled meteorologists to predict the occurrence of lake-effect snow in

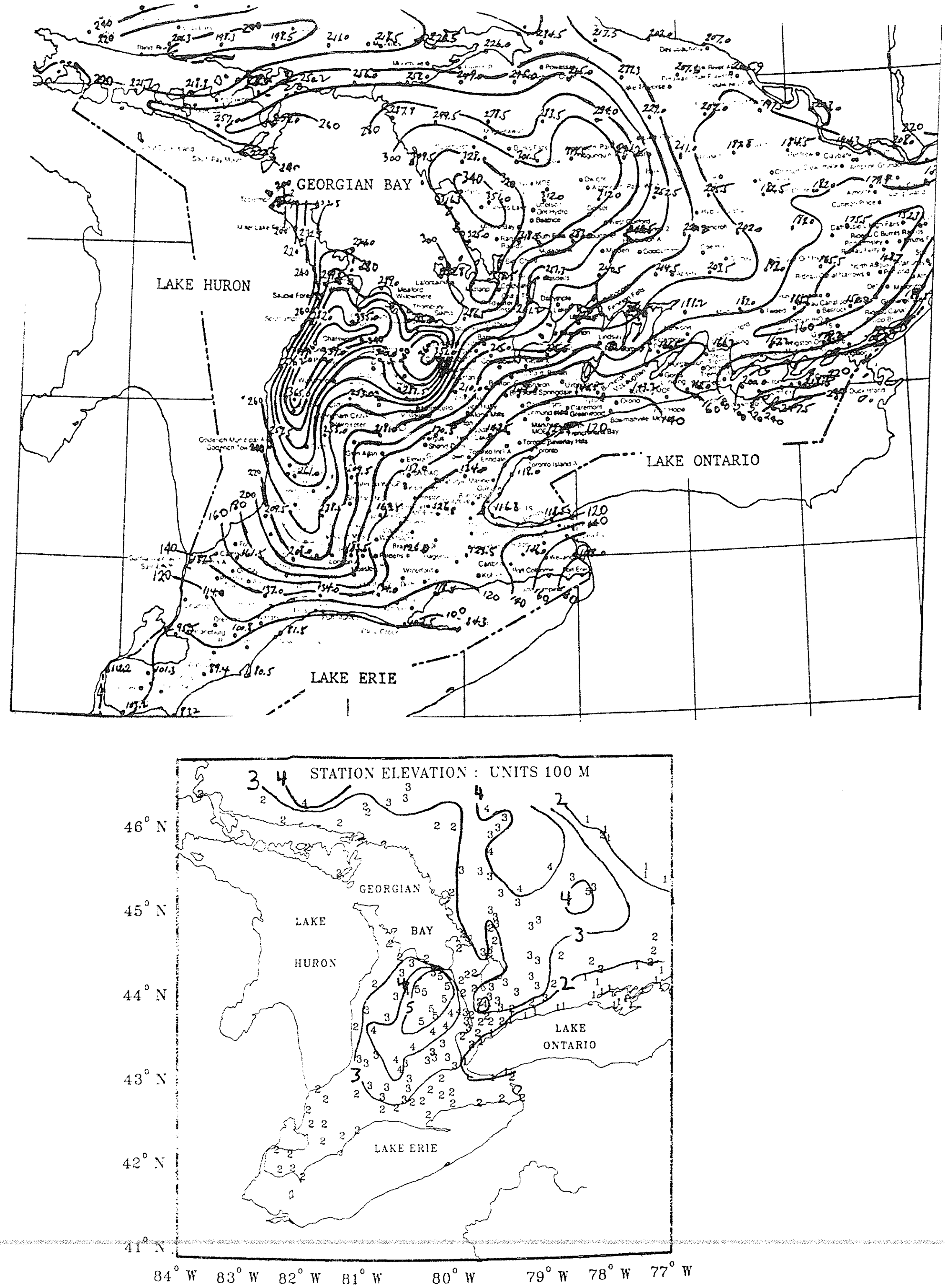

Figure 1: Top: mean annual snowfall (cm) for 1951-1980 for southern Ontario, from Crowe(1985), reproduced in Burrows (1990). Bottom: southern Ontario elevation in 100's of meters.

broad regions to the lee of large water bodies with good success by using output from operational NWP models. However, the mesoscale details of the snowfall, such as band movement and snow amounts at specific points are often not well forecast. There is little objective operational guidance available to forecasters for mesoscale lake-effect snow prediction, likely because the scale on which lake-effect snow occurs is well below the resolution of current operational NWP models. Present forecast procedures in operational weather offices generally rely on subjective manual or semi-computerized interpretation of model output fields using decision tree methods and accumulated office experience (e.g. Dockus, 1985; Niziol, 1987; Murphy, 1989). These methods, while useful for specific sites, are too time-consuming to be simultaneously applied in an operational forecast mode for a large number of sites covering broad, complicated areas. Research studies using mesoscale models have appeared in the literature (e.g. Hjelmfelt, 1990; Lavoie, 1972), but these have not to this author's knowledge, led to production of regular objective forecast guidance in the field. A method for generating objective forecast guidance for lake-effect snow using an operational NWP model was developed by Dewey (1979a,b), but was never implemented.

A growing number of public and private sector users are interested in timely, accurate many-site forecast guidance for snow amounts within public forecast regions. Aware of this, Burrows (1990) studied the feasibility of producing objective mesoscale forecast guidance for 24-hour lake-effect snow amounts for 1-Day and 2-Day projection times for climatological stations in the southern Georgian Bay region. "Perfect Prog (PP)" forecasts of 24-hour snow amount in 5 categories were produced using multiple discriminant analysis (MDA). Using the MDA probabilities as predictors, the forecasts were tuned with a statistical classification procedure known as Classification and Regression Trees (CART) developed recently by Brieman et al (1984). CART establishes rules by which categorical or continuous predictands can be classified in a decision tree by threshold values of user-specified predictors. It is non-parametric (makes no assumptions about the statistical distribution of the predictors), and will use both single predictors and linear combinations of predictors offered to it.

It was suggested in the above work that while the tuned PP statistical forecasts lacked the accuracy to be useful as guidance, there was considerable promise that better forecasts could be generated if the CART procedure were more fully employed and if a better predictor set were designed. The latter would be possible if NWP model output data were used for derivation of predictors instead of analyzed observed data (i.e. "Model Output Statistics (MOS) (Glahn and Lowry, 1972)" approach instead of PP approach). This indeed turned out to be the case, and the resulting method for producing mesoscale lake-effect snow forecast guidance will be implemented experimentally for use at the Ontario Weather Center. This paper describes the development of the forecast system.

2. PREDICTAND

Daily snowfall data for the winters of 1984-1988 (1 November to 30 April, except to 25 March 1987 and to 8 April 1988) were gathered for 29 climatological stations to the lee of Lake Huron and Georgian Bay shown in the left panel of Figure 2. (The right panel is the interpolation grid for NWP model data and is discussed in Section 3). A climatological day at these stations is defined as 8am today to 8am tomorrow. The total number of days without missing data during the study period ranged from a low of 600 to a high of 643, with about 633 days available at many stations. The data set was screened at each station to separate out "lake-effect snow possible" (LESP) days. These were defined as at least 12 hours in a 24 hour period when the CMC NWP model forecast a local off-water 850 mb wind fetch and either a lake-850 mb temperature difference of at least 13 degrees C (dry adiabatic), or 8 to < 13 degrees with upward vertical motion at 700 mb in the middle of Lake Huron (which is always free of ice). Water temperature was taken from time-mean charts published in Saulesleja (1986). Snow amount in cm at each station was divided into 5 categories: 1= 0-trace, 2= >trace-5, 3= >5-12.5, 4= >12.5-22.5, 5= >22.5.

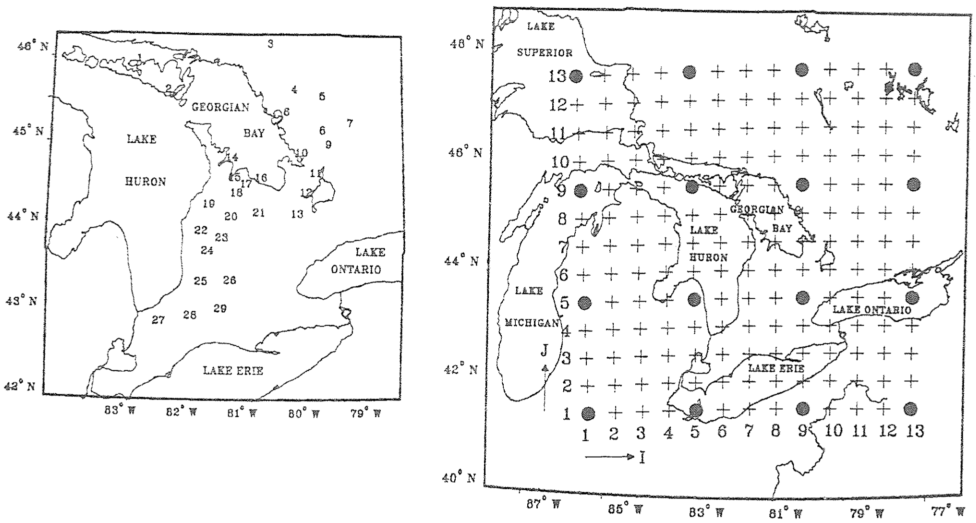

Figure 2: Left panel: Location of climatological stations included in this study. Right panel: NWP model output data grid points (solid circles), interpolation grid points for calculation of predictors (crosses).

The top left panel of Figure 3 shows the fraction of days at each station that qualified as LESP days when classified with Day 1 NWP forecast data (units days per thousand, e.g. 455 means .455). Results are what would be expected, considering the prevailing west to northwest winds over this region. The number of LESP days varies with the fetch direction at each station. The maximum fraction of LESP days occurs to the lee of central Lake Huron, with fewest occurrences at the northern and southern extremities of Lake Huron, and occurrences decreasing with distance inland. Table 1 shows the highest snow category that was observed at any of the 29 stations when a LESP day was observed at any one of them. Little difference is seen between the two projection times. This suggests the NWP model forecasts of the large scale flow in this region for 0-24 hours and 24-48 hours projection times tend to be consistent for LESP days. A normal distribution *centered on category 3* is seen for LESP days identified with both 00-24 hour (Day 1) and 24-48 hour (Day 2) NWP model forecast data.

Table 1: The highest snow category (Cat) observed when a LESP day was identified at any of the 29 stations, shown as a percent of total LESP days for each category. Results are for LESP days determined by Day 1 and Day 2 NWP model forecast data.

Cat	%Day 1	%Day 2
1	06.8	07.0
2	24.9	22.4
3	33.3	35.8
4	22.7	22.7
5	12.3	12.1

A different story emerges from the results for *individual* stations. The remaining panels in Figure 3 show the fraction of LESP days when snowfall occurred in categories 1-5 at each station. For the aggregate of stations it was seen in Table 1 that snow was observed by at least 1 of the 29 stations on about 93% of the LESP days. However, Figure 3 shows that on roughly 25-65% of the LESP days at *individual* stations no snow fell (CAT 1) or if snow did occur, snowfall was not more than 5 cm (category 2) on roughly another 30-50% of the LESP days (CAT 2). Thus the maximum number of snow occurrences on LESP days for

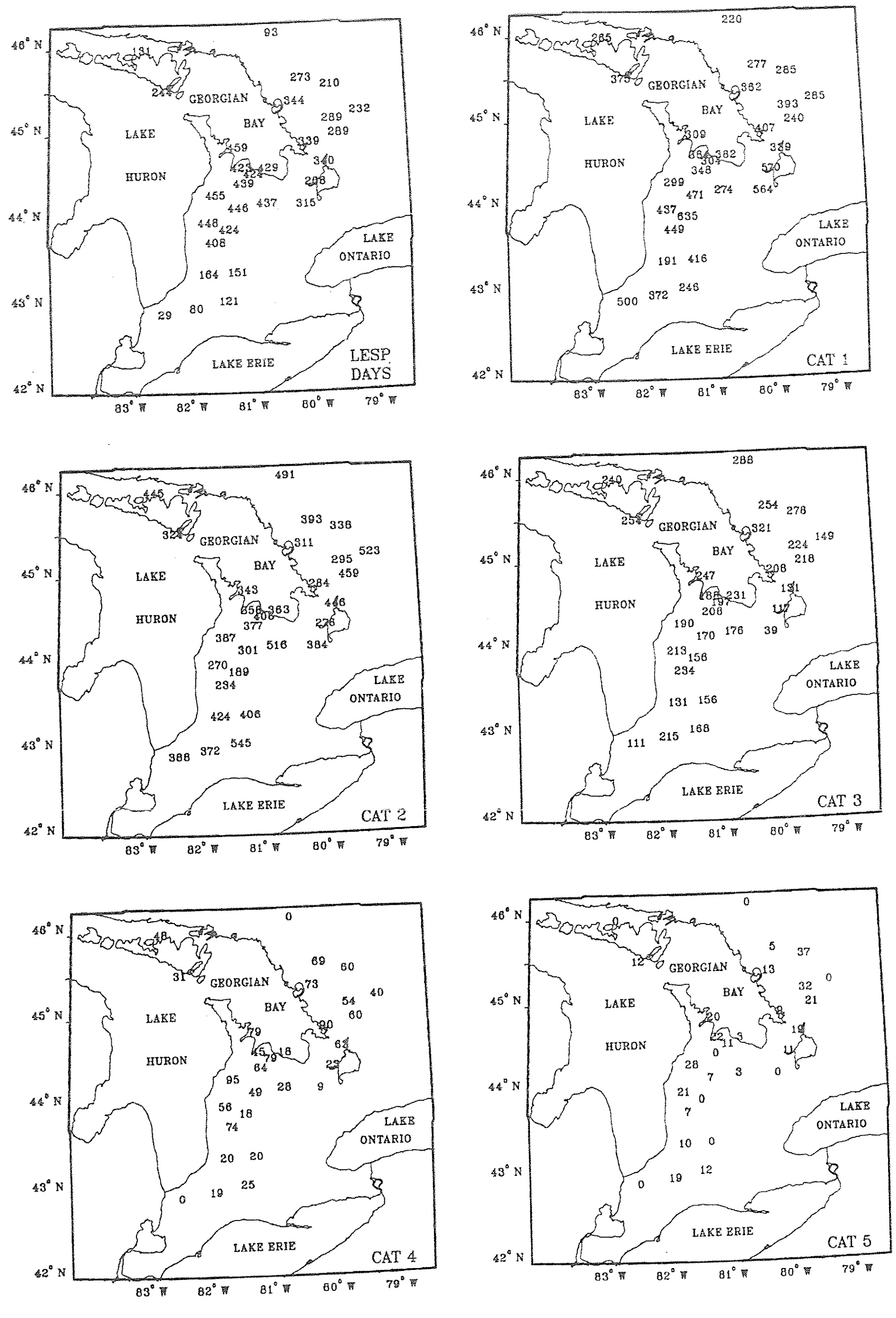

Figure 3: LESP DAYS: number of days that "lake effect snow possible" (LESP) days were identified with Day 1 (0-24 hr) NWP model forecast data; CAT 1, ..., CAT 5 number of LESP days when 24-hour observed snow amount was in category 1 (0 - trace), or category 2 (>trace-5cm), or category 3 (>5-12.5 cm), or category 4 (>12.5-22.5 cm), or category 5 (>22.5 cm). Units days per thousand days (e.g. 455 means .455).

individual stations is centered on *categories 1 or 2*, even in the heart of the "snowbelt" region to the lee of central Lake Huron and central Georgian Bay. (More than 90% of the snow occurrences on LESP days were in these two categories in the highly populated region to the lee of southern Georgian Bay). Heavy snowfalls on LESP days at individual stations are relatively rare: 0-10% for category 4 snowfalls and 0-4% for category 5 snowfalls. The most occurrences of category 5 snow on LESP days at any station was only 8 in the 4 years.

The above results imply that *over-forecasting* the amount and occurrence of snow in lake-effect storms for individual stations and small areas is likely to be a problem for even relatively small public forecast areas. This is because the maximum snowfall will occur over a small portion of a region, but a forecaster will tend to predict the maximum amount everywhere to be safe.

3. PREDICTORS

1200 UTC NWP model forecasts of wind, temperature, geopotential height, and vertical motion at 6-hour intervals from 0 to 48 hours projection time were interpolated with a bi-cubic spline algorithm to a mesoscale grid of about 63 km spacing over southern Ontario (shown in the right panel of Figure 2). The solid circles are locations of the NWP model data (about 254 km resolution), the crosses are interpolated points. Data on the 63 km interpolation grid can be thought of as representing the "between-grid-point" variation of data on a 254 km scale, thus shoreline-water differential frictional effects on boundary layer winds be poorly represented. We are therefore dealing with relations between mesoscale lake-effect snow formation and upper air controls on the much larger synoptic scale.

Predictors known or expected to be physically related to formation of lake-effect snow were designed and calculated on the 63 km grid. Comprehensive recent discussions of the meteorological parameters and features important in the process of lake-effect snow formation can be found in Dockus (1985), Niziol (1987), Murphy (1989), and Hjelmfelt (1990). A total of 129 separate potential predictors were designed to accommodate the 24-hour period of the predictand. These are derived from physical parameters shown in Tables 2 and 3. Lake water temperature and mean ice cover were estimated as space averages surrounding the grid points from time-mean charts for periods of variable length during the winter season in Saulesleja (1986). Low level divergence was calculated from interpolated east-west and north-south (u and v) wind components by calculating local small-scale derivatives at each interpolation grid point.

4. CART

CART was used to find predictand classification trees for each station using Day 1 and Day 2 NWP model forecasts of predictors. Trees were created with the basic 129 predictors for 28 of the 29 stations (station 27 had only 18 LESP days). For stations where only 1 or 2 cases of the rarest category occurred, those observations were classified at the next lowest category before CART was run. Since it is a recent development and is not widely known in meteorological circles, a brief description of CART and its use in this study follows. For greater detail the reader is asked to refer to Brieman et al (1984).

CART is given a data base consisting of predictand and predictor values which it uses to establish a decision tree that classifies the predictand. Computer output from the CART program for Day 1 at station 19 (Paisley, Ontario) is shown in Figure 4. The predictand was snow category (1-5) on LESP days. The "Tree Sequence" summary shows CART first found a tree that classified all the data perfectly (Tree 1 with 65 Terminal Regions), then began "pruning" "nodes" in the "weakest links" up from the bottom of the tree until 1 tree remained (Tree 12). Tree 12 would assign the "initial class assignment", which is the category which gives the lowest misclassification cost (.614) if all data were classified as a single category (2 here). This will be the most common category if unit

Table 2: Physical parameters important for lake-effect snow production. Operations "*av, mx, mn, ch*" mean respectively: average value, maximum value, minimum value, and change between end and beginning, in a 24-hour period. Operation N6 is the number of 6-hour times within a 24-hour period (0-5) that a specified condition applies. Location identifiers such as "65, 68, 88" denote points (I=6,J=5), (I=6,J=8), (I=8,J=8) in Figure 2, respectively. "loc" denotes station location.

1. lake - 850 mb temperature difference. *(av, mx, mn at 68)*
2. 500 mb temperature. *(mn at 68)*.
3. ice cover. (percent).
4. lake - 850 mb temperature difference plotted against lake - 700 mb temperature difference [in Figure 5 of Niziol (1987), define an index: 0=outside graph, 1=conditional, 2=moderate, 3=extreme]. *(av, mx, mn, ch, N6 of index value at 68)*.
5. 1000 mb wind direction. *(av, ch at 65, 68, 88)*.
6. 850 mb wind direction. *(av, ch at 65, 68, 88)*.
7. 700 mb - 1000 mb wind direction difference. *(av, mx, mn, ch, N6 at 65, 68, 88)*.
8. 700 mb - 1000 mb wind direction difference in two ranges: 0-30 degrees, 30-60 degrees. *(N6 at 65, 68, 88)*
9. 850 mb wind speed. *(av at 65, 68, 88)*.
10. 1000 mb wind speed in 12 direction segments 360-30 degrees, 30-60 degrees, ..., 330-360 degrees. *(av at loc)*.
11. simultaneous positive vorticity advection at 500 mb, 700 mb, 850 mb. *(N6 at 68)*.
12. 500 mb advection of absolute vorticity. *(av, mx at 68)*.
13. 700 mb vertical velocity. *(av, mx, mn, ch 65, 68, 88)*.
14. 700 mb temperature advection. *(av, mn at 68)*.
15. 1000 mb divergence within lines and areas defined in Table 3. *(av)*
16. 1000 mb divergence within lines and areas defined in Table 3. *(N6* of min for all J=1-13 at each I within a specified line or area, summed over each I within that line or area)
17. [700 mb + 1000 mb] divergence within lines and areas defined in Table 3. *(av)*.
18. 700 mb wind maximum for all J=1-13 in Figure 2 occurs within strips [I=2-4], [I=4-6], [I=6-8]. *(N6 in each strip)*.
19. 500 mb low center vicinity of James Bay. (yes or no).
20. 850 mb wind direction at 4,9 is between 280 degrees and 340 degrees and direction at 4,9 minus direction at 9,7 is at least 20 degrees. *(av, N6)*.
21. times "fetch importance index" (1 to 3) defined arbitrarily in 4 local direction sectors for each station from 850 mb wind direction]. *(N6 of index values)*.

Table 3: I and J values in Figure 2 for horizontal and vertical space averages in divergence predictors (Pred) in Table 2. Div01 - 13 are averaged over ranges of I and J in Figure 2 (e.g. Div01 is calculated over the area I= 4 to 12, J=4 to 10, and Div06 along the line I=4 to 9, J=8); Div14 - 18 are calculated over diagonal lines (e.g. Div14 is calculated from the points I=4,J=9 to I=6,J=8).

Pred	I	J	Pred	I	J	Pred	I	J	Pred	I	J
Div01	4-12	4-10	Div06	4-9	8	Div11	9-12	4-6	Div14	4,9	6,8
Div02	5-7	4	Div07	4-9	9	Div12	9-12	7-9	Div15	5,8	6,6
Div03	5-7	5	Div08	4-11	10	Div13	4-5	9	Div16	5,6	6,5
Div04	4-7	6	Div09	7-9	3-5				Div17	7,9	8,8
Div05	4-9	7	Div10	7-9	6-8				Div18	7,8	8,7

misclassification costs are assigned. The final tree selected is Tree 5, the one which had the minimum "cross-validated cost" relative to Tree 12. If no tree structure were found with a cross-validated cost less than 1, then Tree 12 would have been selected.

TREE SEQUENCE

TREE	TERMINAL NODES	CROSS-VALIDATED RELATIVE COST	RESUBSTITUTION RELATIVE COST	COMPLEXITY PARAMETER
1	65	0.97 +/- 0.047	0.00	0.000E+00
2	35	0.91 +/- 0.048	0.17	0.353E-02
3	28	0.91 +/- 0.048	0.25	0.703E-02
4	19	0.90 +/- 0.048	0.37	0.820E-02
5*	13	0.89 +/- 0.048	0.47	0.105E-01
6	11	0.91 +/- 0.048	0.51	0.123E-01
7	8	0.95 +/- 0.048	0.58	0.140E-01
8	5	0.92 +/- 0.048	0.67	0.176E-01
9	4	0.92 +/- 0.048	0.71	0.281E-01
10	3	0.94 +/- 0.048	0.77	0.351E-01
11	2	0.96 +/- 0.047	0.86	0.562E-01
12	1	1.00 +/- 0.047	1.00	0.842E-01

INITIAL MISCLASSIFICATION COST = 0.614
INITIAL CLASS ASSIGNMENT = 2

CLASSIFICATION TREE DIAGRAM

```
                                          I
                +-------------------------1-------------------------+
                I                                                   I
+---------------2---------------+                           +---10----+
I                               I                           I         I
+--3--+               +-------------4-----------+       +---11---+
I     I               I                         I       I        I
          +------5------+           +----8----+    +-12--+
          I             I           I         I    I     I
       +--6--+       +--7--+     +--9--+
       I     I       I     I     I     I

TERMINAL REGIONS
1    2     3     4     5     6     7     8     9     10     11     12     13
```

Figure 4: Tree sequence summary and classification tree diagram for Tree 5 found by CART for LESP days classified with Day 1 (0-24 hr) NWP model forecast data for Station 19 in Figure 2.

The cross-validated relative cost (.89 +/-.048 for Tree 5) can be thought of as a "skill score with respect to climatology" for each tree. A ten-fold cross-validation method was used here to evaluate the cost of each tree when searching for the optimally-pruned tree. This works by splitting the four-year "learning sample" into ten "cross-validation sub-samples", each containing 90% of the learning sample and about the same distribution of data in the categories. For every "complexity value" (a number inversely proportional to the number of nodes) reached by the main tree as it is pruned, CART takes each cross-validation sub-sample and grows "auxiliary trees" up to the complexity value reached by the main tree. It then evaluates the misclassification cost of each auxiliary tree by running down the tree the 10% of data not contained in its cross-validation sub-sample and storing the results in a "cross-validation classification matrix". *Cross-validation approximates a test of the main decision tree with an independent data set in a conservative manner* (misclassification costs are over-estimated because the full learning sample is not used to grow the auxiliary trees). It is recommended for small data sets or data sets that are sparse in some categories, such as categories 4 and 5 in this study. (For data sets with relatively large numbers of cases in each category, one can set aside a specified fraction of the data as an independent set on which to evaluate the misclassification cost of the main tree as it is grown).

Three tree construction rules for splitting the samples in the nodes were tested: "gini, twoing, and ordered twoing". The reader is referred to Brieman et al (1984) for a full explanation of these terms. "Ordered twoing", which treats the decision in each node as a choice between two ordered classes when splitting cases to the left or to the right, seemed appropriate here since snow amount increases with category number.

Table 4 shows the decision rules CART found for segregating cases in the nodes, and the data populations in the Terminal Regions (final classification nodes), for the classification tree diagram in Figure 4. In Node 1, events with greater snow amounts overall are sent left into Node 2 while events with smaller overall snow amounts are sent right into Node 10. The test for sending data left or right in Node 1 is based on a linear combination of three predictors. The rules are physically realistic. A leftward split into Node 2 occurs for events where, relative to all the cases, the 700 mb-1000 mb wind direction change over southern Lake Huron between is 0 and 30 degrees for a relatively large portion of a 24-hour period, 24-hour average low level convergence over southern

<u>Table 4:</u> First part: Decision tree rules found by CART for splitting the nodes in the Day 1 learning sample for station 19 for the classification tree diagram shown in Figure 4. Classes assigned to cases going left or right are shown for each node, with final classification categories in Terminal Regions highlighted in bold. Predictors are coded to follow the explanation in Tables 2 and 3. For example: #8/0-30/*N6*/<u>68</u> means predictor number 8, 0-30 degree range, operation *N6*, calculated at grid point <u>68</u>; #15/02/*Av* means predictor number 15, Div02 in Table 3, operation *Av*. Units of some quantities are shown, divergence units are 10^{-6} s^{-1}, vorticity advection units are 10^{-11} s^{-2}. Second part: data populations in terminal regions in Figure 4.

<u>Decision Rules:</u>

Node 1:	L=2 R=1	split left if { -.613*[#8/0-30/N6/68] +.553*[#15/02/Av] -.563*[#1/Mx] }	<= -12.0
Node 2:	L=2 R=3	split left if { -.251*[#15/02/Av] + .585*[#1/Mx] -.222*[#10/270-300] + .738*[#11] }	<= 14.8
Node 3:	L=2 R=3	split left if [#7/Mn/68]	<= 51.3 degrees
Node 4:	L=4 R=3	split left if [#5/Av/68]	<= 306 degrees
Node 5:	L=3 R=4	split left if [#2]	<= -40.7 degrees C
Node 6:	L=3 R=5	split left if [#10/270-300]	<= 2.15 ms^{-1}
Node 7:	L=4 R=3	split left if [#10/270-300]	<= 10.9 ms^{-1}
Node 8:	L=2 R=3	split left if [#13/Ch/65]	<= 3.05 mb hr^{-1}
Node 9:	L=2 R=3	split left if [#14/Av/68]	<= .05 degrees C hr^{-1}
Node 10:	L=1 R=2	split left if [#1/Av]	<= 20.1 degrees C
Node 11:	L=1 R=2	split left if [#20/Av]	<= 32.0 degrees
Node 12:	L=2 R=1	split left if [#13/Mn/68]	<= -8.15 mb hr^{-1}

<u>Data Population In Terminal Regions:</u>

Terminal Region	Class	Category 1	2	3	4	5
1	2	15	51	11	1	1
2	3	1	0	4	0	0
3	3	0	1	5	0	0
4	5	0	0	0	0	5
5	4	2	7	5	24	1
6	3	1	0	4	0	0
7	2	2	9	2	2	0
8	3	0	0	4	0	0
9	3	0	1	14	0	1
10	2	3	7	1	0	0
11	1	60	19	4	0	0
12	2	1	6	0	0	0
13	2	1	9	0	0	0

Lake Huron was relatively large, and a relatively large 24-hour maximum water-850 mb temperature difference occurred over northern Lake Huron. In Node 2, events with relatively greater low-level convergence, relatively large water-air temperature difference, relatively large local 1000 mb wind speed in the 270-300 degrees sector, and relatively greater positive vorticity advection over northern Lake Huron, are sent right into Node 4. Cases in the remaining nodes are sent to the left or right based on a test of a single variable. The category 4 and 5 snow cases fall into Terminal Region 4 or 5 based on tests of 500 mb temperature, and 1000 mb wind speed and direction.

The learning sample data populations in the terminal regions allow a *statement of confidence* to be made about the classification forecast assigned to the node, and allow for less specific forecasts to be made. For example in Table 4, if the predictor values cause a case to fall down the tree into Terminal Region 4, a forecast of category 5 snow is made with 100% confidence for station 19. However, if the predictor values cause a case to fall down the tree into Terminal Region 11 we can make a forecast of "no snow" with 72% confidence, or "no more than 5 cm of snow" with 95% confidence, or "snowfall will not exceed 12.5 cm" with 100% confidence.

Table 5 shows the learning sample and cross-validation matrices which resulted from segregating the data for station 19 by the rules given in Table 4. Differences in the total number of cases in each category between Days 1 and 2 occurred since the number of days which qualified as LESP days varied due to differences between the NWP model forecasts of the synoptic scale flow for 0-24 hours and 24-48 hours for the same day. The fit of the data in the learning sample by CART shown in Table 5 appears to be very good, although the cross-validation matrices are less impressive but still reasonable. The category with the maximum number of cases in each predicted category matches the correct observed category in all the matrices with only one exception. The high degree of matching was typical for every station for as many categories as CART would fit, and is a desirable attribute.

Table 5: Learning sample and cross-validation sample classification matrices for Day 1 and Day 2 forecasts for station 19 (Paisley, Ontario).

Day 1

		Learning					Cross-Validation				
Observed Class -------->		1	2	3	4	5	1	2	3	4	5
Pre-	1	60	19	4	0	0	47	36	9	0	0
-dic-	2	22	82	14	3	1	28	57	24	11	1
-ted	3	2	2	31	0	1	11	12	17	10	1
	4	2	7	5	24	1	0	5	4	6	3
Class	5	0	0	0	0	5	0	0	0	0	3

Day 2

		Learning					Cross Validation				
Observed Class -------->		1	2	3	4	5	1	2	3	4	5
Pre-	1	81	13	2	0	0	55	24	6	2	0
-dic-	2	11	90	12	5	0	34	56	18	10	1
-ted	3	4	5	39	11	5	6	20	21	9	2
	4	0	0	2	10	0	1	6	8	5	2
Class	5	0	0	0	0	3	0	2	2	0	3

In the majority of cases CART found trees that classified most or all of the observed categories at each station, to or within 1 standard deviation of the lowest cross-validated relative cost. However, CART could not find low-cost trees for some stations which had one very dominant snow category (usually category 1). For these cases a "second attempt", based on an idea described in Burrows(1990) for improving predictand-predictor fits by MDA, was tried in an effort to take advantage of non-linear predictand-predictor relationships that might be present. Ten predictors were chosen from node paths that classified data up to the maximum observed category in trees whose cost was too high (more than one standard deviation above the minimum cost, and often with a large number of nodes). Squares, cubes, natural logarithms, and cross products of the 10 predictors were added to the basic predictor set, then CART was re-run with the enhanced set of 204 predictors. The limit of 10 predictors was arbitrarily imposed to limit the number of predictors in the enhanced set. The result was substantially lowered misclassification rates for several stations.

5. RESULTS

a. Classification Trees

A summary of the results for Day 1 and Day 2 classification trees appears in Table 6. CART found trees which classified all the observed snowfall categories at 14 of the 28 stations (50%) for Day 1 forecasts and 9 of the 28 stations (32%) for Day 2 forecasts. Trees that classified to within 1 category of the maximum observed category were found at 26 of the 28 stations (93%) for Day 1 forecasts and 25 of the 28 stations (89%) for Day 2 forecasts. About 84% of the trees for Day 1 and Day 2 had cross-validated relative costs less than or equal to 1 (i.e. skill as good or better than climatology). This was accomplished with trees having 5-15 nodes at most stations, which is reasonably few. The cross-validated misclassification costs averaged .50 for both Day 1 and Day 2 forecasts, while the misclassification costs for the learning trees were .20 for Day 1 forecasts and .23 for Day 2 forecasts. (The cross-validation misclassification costs may seem high, but we should recall that cross-validation will *over-estimate the true cost* of using the trees with independent data because the auxiliary trees were grown with only 90% of the data. The true costs if the trees were applied to independent data will be somewhere between the cross-validation sample and learning sample misclassification rates, and hopefully are not too far off the latter). As expected, the node-splitting rule that gave the lowest misclassification costs for most stations was "ordered twoing" (32 of 56 trees).

Table 6: For each station (Stn) in Figure 2: maximum category observed (Cats Obsv); maximum number of snow categories classified by CART trees, to within 1 standard deviation of minimum cross-validated relative cost (Cats Fcst); cross-validated relative cost (CV Cost); misclassification costs of auxiliary trees grown from cross-validation samples (CV Miscl Cost) and from the full learning sample (Lrn Miscl Cost).

	Day 1 (0-24 hours)					Day 2 (24-48 hours)				
Stn	Cats Obsv	Cats Fcst	CV Cost	CV Miscl Cost	Lrn Miscl Cost	Stn	Cats Obsv	Cats Fcst	CV Cost	CV Miscl Cost	Lrn Miscl Cost
1	4	4	1.00	.55	.17	1	4	4	.86	.46	.09
2	5	5	.88	.55	.11	2	5	4	.83	.53	.25
3	3	3	1.07	.54	.14	3	3	3	1.07	.44	.08
4	5	4	.82	.50	.20	4	5	4	.99	.63	.33
5	5	5	.83	.55	.14	5	5	5	.92	.61	.22
6	5	4	.76	.48	.25	6	5	4	.85	.54	.20
7	4	4	1.01	.48	.10	7	4	4	.99	.50	.14
8	5	5	.93	.56	.20	8	5	5	.83	.49	.24
9	5	5	.93	.54	.18	9	5	4	1.00	.54	.28
10	5	4	.90	.53	.17	10	5	4	.93	.52	.33
11	5	5	1.02	.56	.25	11	5	5	1.02	.55	.21
12	5	4	1.03	.44	.26	12	5	4	.90	.39	.15
13	4	3	.94	.41	.19	13	4	3	1.03	.43	.16
14	5	4	.83	.53	.28	14	5	4	.88	.56	.28
15	5	4	.85	.52	.24	15	5	3	.83	.49	.28

Table 6 (contd):

Day 1 (0-24 hours)					
Stn	Cats Obsv	Cats Fcst	CV Cost	CV Mscl Cost	Lrn Mscl Cost
16	5	4	.86	.53	.20
17	5	5	.90	.53	.26
18	4	4	.82	.51	.15
19	5	5	.89	.55	.29
20	5	4	.88	.46	.19
21	4	4	.85	.42	.17
22	5	5	.97	.55	.17
23	4	4	1.11	.40	.12
24	5	4	.84	.46	.34
25	4	3	1.00	.48	.23
26	4	3	.79	.46	.13
27	-	-	-	-	-
28	5	3	.87	.55	.08
29	5	3	.97	.45	.26

Day 2 (24-48 hours)					
Stn	Cats Obsv	Cats Fcst	CV Cost	CV Mscl Cost	Lrn Mscl Cost
16	5	4	.92	.54	.18
17	5	4	.86	.52	.34
18	4	4	.86	.53	.15
19	5	5	.83	.52	.24
20	5	4	.91	.51	.28
21	4	4	1.10	.53	.13
22	5	4	.98	.56	.34
23	4	3	.94	.33	.14
24	5	4	.97	.52	.30
25	4	3	.98	.47	.34
26	4	3	.93	.54	.22
27	-	-	-	-	-
28	5	3	.84	.44	.17
29	5	3	.83	.37	.28

CART gives an ad-hoc ranking of variable importance on a scale of 0-100 after the trees are constructed. *Predictors related to low-level divergence were ranked among the most important predictors more frequently than any other types of predictors at nearly every station.* The location of the line or area of this divergence was frequently upstream from each station, although low-level divergence (convergence) in the vicinity of Lake Ontario *downstream* from stations over southern and central Lake Huron was important as well. The importance of low-level convergence lines to the movement and location of lake-effect cloud bands was noted by Peace and Sykes (1966) and recently by Murphy (1989). The next most frequently picked set of predictors were those measuring the degree of air-water temperature difference, which is well known to be the primary mechanism for formation of lake-effect cloud bands.

b. Areal consistency

Figure 5 shows the observed snowfalls and residuals (observed category minus forecast category) for Day 1 and 2 classifications for a period of heavy lake-effect snowfall (5-6 January 1988). For the residuals a "0" is a perfect result, and a "*" denotes a station where the local 850 wind direction was not off-water for at least 2 of the four 6-hour intervals in a 24-hour period (i.e. non-LESP day). There is a preponderance of "0's" and "+/- 1's" for all the residuals for both Day 1 and Day 2 classifications. Of course, no system is perfect so there are a few 2's and 3's (but no 4's). Overall, the CART classifications for individual LESP days preserved the areal patterns of snowfall rather well in both Day 1 and Day 2 forecasts, even though the classification trees were found separately for each station. There were many other examples which verified this.

6. CONCLUSIONS

The need for detailed forecast guidance and careful assessment of the particulars of the synoptic situation on LESP days is clear. Analysis of the occurrence of snow for a four-year period at 29 stations in the lee of Lake Huron and Georgian Bay on "lake-effect snow possible (LESP)" days suggests that snow occurrence and amounts for individual stations and small areas are likely to be over - forecast in public forecasts, even in the heart of the "snowbelt".

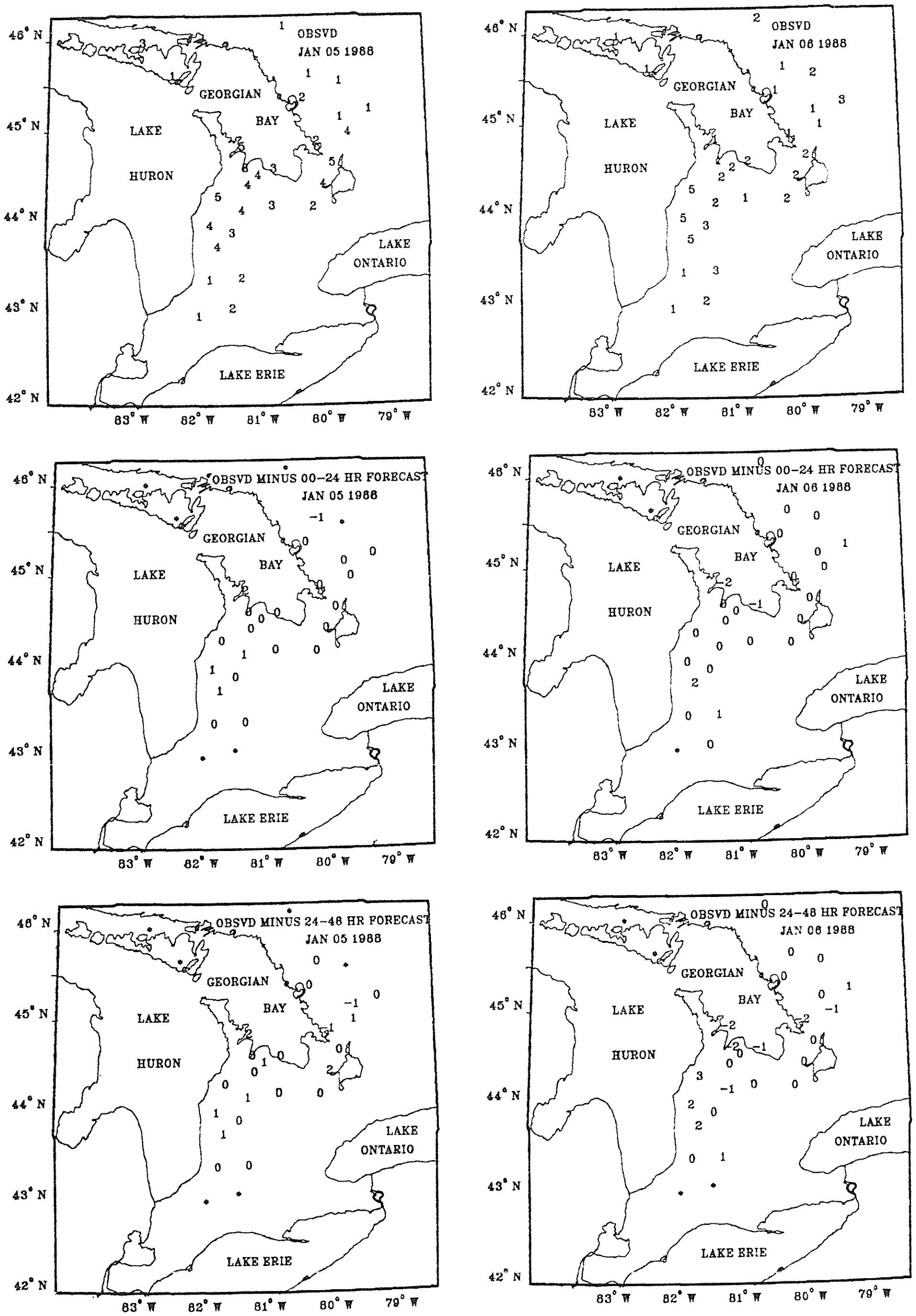

Figure 5: For 5-6 January 1988: Observed snow categories and difference between forecast and observed snow categories for Day 1 (00-24 hr) and Day 2 (24-48 hr) forecasts.

The classification-tree results show the CART-based MOS method outlined here has considerable promise for timely many-site production of useful, areally consistent objective operational mesoscale guidance for 1 and 2 day forecasts of 24-hour lake-effect snow amount. The trees had good skill for both 0-24 hour and 24-48 hour periods in classifying the category of snow occurrence at individual stations, given that a LESP day was expected. The node-splitting decision rules found by CART were in most cases easy to interpret physically and are a useful tool for insight into physical processes important in the formation and prediction of lake effect snow. Once obtained, the trees could be run on a local computer system to make forecasts if the required NWP data were communicated from a central mainframe computer. The trees are easy to use and could be portable to other NWP models by tuning predictor values from model output generated during a period of a few months when two model versions overlap.

An explanation of the CART method and its use here is given in Section 3. A parallel approach to an idea described in Burrows (1990) for including certain non-linear functions of important predictors into the predictor selection process was tried and resulted in significant reductions of misclassification rates for several stations.

Low-level divergence (convergence) induced by events in the upper air, usually located upstream and along lines or zones over the water, is the primary parameter controlling snow amount on days when lake-effect snow is possible.

ACKNOWLEDGEMENTS.

Comments on the manuscript by Dr. S. Venkatesh, Dr. M. Khandekar, and by anonymous reviewers with the Eastern Snow Conference are much appreciated.

REFERENCES

Braham, R.R., and Dungey, M.J., 1984: Quantitative estimates of the effect of Lake Michigan on snowfall. *J. Climate Appl. Meteor.* **23**: 939-949.

Breiman, L., Freidman, J.H., Olshen,R.A. and Stone, C.J., 1984: *Classification and Regression Trees*. Wadsworth & Brooks/Cole, Monterrey, 358pp.

Burrows, W.R., 1990: Tuned perfect prognosis forecasts of mesoscale snowfall for southern Ontario. *J. Geophys. Res.* **95**, No. D3: 2127-2141.

Crowe, R.B., 1985: Effect of carbon dioxide warming scenarios on total winter snowfall and length of winter snow season in southern Ontario, *Rep. 85-19*, p. 18, Can. Clim. Cent., Atmos Environ. Serv., Downsview, Ont., 1985.

Dewey, K.F., 1979a: Lake Erie induced mesosystems-an operational forecast model, *Mon. Wea. Rev.* **107**: 421-425.

----, 1979b: An objective forecast method developed for Lake Ontario induced snowfall systems. *J. Appl. Meteor.* **18**: 787-793.

Dockus, D.A., 1985: Lake-effect snow forecasting in the computer age. *Natl. Wea. Dig.* **10**: 5-19.

Glahn, H.R., and Lowry, D.A. 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.* **11**: 1203-1211.

Hjelmfelt, M.A., 1990: Numerical study of the influence of environmental conditions on lake-effect snowstorms over Lake Michigan. *Mon. Wea. Rev.* **118**: 138-150.

Lavoie, R.L., 1972: A mesoscale numerical model of lake-effect storms. *J. Atmos. Sci.* **29**: 1025-1040.

Murphy, B.P., 1989: Forecasting lake effect snow. Ontario Region Technical Note No. 89-7, Atmospheric Environment Service, Downsview, Ontario, Canada.

Niziol, T.A., 1987: Operational forecasting of lake effect snowfall in western and central New York. *Wea. and Forecasting* **2**: 310-321.

Peace, R.L., and Sykes, R.B., 1966: Mesoscale study of a lake effect snow storm. *Mon. Wea. Rev.* **94**: 495-507.

Saulesleja, A., 1986: *Great Lakes Climatological Atlas*. Environment Canada, Ottawa, 145pp. [Available through Canadian Government Publishing Center.]