NOTES AND CORRESPONDENCE

Improving 30-Day Great Lakes Ice Cover Outlooks*

RAYMOND ASSEL

NOAA/Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan

SHELDON DROBOT

Board on Atmospheric Sciences and Climate, The National Academies, Washington, D.C.

THOMAS E. CROLEY II

NOAA/Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan

10 October 2003 and 3 March 2004

ABSTRACT

Prediction of Great Lakes ice cover is important for winter operations and planning activities. Current 30-day forecasts use accumulated freezing degree-days (AFDDs) to identify similar historical events and associated ice cover. The authors describe statistical models that relate future ice cover to current ice cover, AFDDs, and teleconnection indices, available on the day the forecast is made. These models are evaluated through Monte Carlo simulation and assess the potential of a perfect AFDD forecast in a regression between ice cover and AFDDs between the forecast date (first day of month) and the date for which the forecast is made (first day of next month).

1. Introduction

United States and Canadian federal agencies use 30-day forecasts of ice conditions, in the form of ice charts, as an aid in planning winter operations. The U.S. Navy–National Oceanic and Atmospheric Administration (NOAA) National Ice Center (NIC) issues 30-day forecasts of Great Lake ice conditions (http://www.natice.noaa.gov/pub/great_lakes/) on the first and fifteenth of every month from December through March. Thirty-day air temperature forecasts are used to calculate accumulated freezing degree-days (AFDDs), and then databases of AFDDs and historical ice charts are used to find analogs (Snider 1974). Empirical statistical models are developed here to make 30-day forecasts of beginning of month (BOM) lake-averaged ice cover. Separate models are built for each Great Lake (Superior, Michigan, Huron, Erie, and Ontario) and each BOM date (1 January, 1 February, and 1 March). Predictor variables, BOM ice cover datasets, and model types are discussed. Monte Carlo simulations are used identify the best model for each lake each month.

2. Data

A 30-winter time series of daily lake-averaged ice concentration for each Great Lake is available on the Internet (Assel 2003a). These data are used to estimate ice cover on BOM dates of 1 January, 1 February, and 1 March for each winter from 1973 to 2002. Predictor variables include 1) the previous BOM lake-averaged ice cover, 2) monthly lake-averaged AFDDs (Assel 2003b), and 3) teleconnection indices [tropical–North American (TNH) index, North Atlantic Oscillation (NAO) index, east Atlantic–western Russian (EAWR) index, Southern Oscillation index (SOI), and the polar–eurasian (POL) index], obtained from the NOAA Climate Prediction Center, available online at http://www.cpc.ncep.noaa.gov/products/MD_index.html. Teleconnection indices have significant correlations with annual maximum Great Lakes ice cover and winter severity (Assel and Rodionov 1998; Rodionov and Assel 2000; Rodionov et al. 2001). We obtained monthly teleconnection data for November, December, and January from 1972 through 2002. We did

* Great Lakes Environmental Research Laboratory Contribution Number 1308.

Corresponding author address: Raymond Assel, NOAA/Great Lakes Environmental Research Laboratory, 2205 Commonwealth Blvd., Ann Arbor, MI 48105-2945.
E-mail: ray.assel@noaa.gov

not utilize February data since they are not available in time to forecast BOM March ice conditions.

3. Methods

a. Forecast development

We developed four forecast models of BOM ice cover with a 30-day lead. The first model is the climatological model (C), which predicts the BOM ice cover as the long-term value:

$$I_m = C_m, \tag{1}$$

where I_m is predicted BOM ice cover, and C_m is the long-term mean on month m. The climatological model is the baseline technique for assessing forecast improvements of other methods.

The second model is the anomaly propagation model (AP), which predicts the BOM ice cover by preserving the anomaly from the preceding month:

$$I_m = C\left(\frac{O_{m-1}}{C_{m-1}}\right), \tag{2}$$

where O_{m-1} and C_{m-1} are the previous BOM observation and long-term climatological mean, respectively. Since there is no reported BOM December ice cover, anomaly propagation cannot be used for a January forecast.

The third model is the observational linear regression model (OLR), which predicts BOM ice cover from regressions on observed ice cover, AFDD, and teleconnection data:

$$I_m = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n, \tag{3}$$

where β_0 is the intercept; β_1 is the regression coefficient of the first independent variable, X_1; and β_n is the regression coefficient of the nth independent variable, X_n. Initially, we correlated BOM ice cover data with all observed AFDD and teleconnection data available at the time of forecast. For instance, we correlated BOM February ice cover with all AFDD and teleconnection data from December and BOM January ice cover, all of which are available on 1 January when the BOM February forecast is made. Next, we entered all variables significantly correlated with the BOM ice cover into a stepwise regression for selection of a model. Finally, we tested the models to ensure that (a) the overall fit of the regression equation is significant; (b) the individual regression coefficients are significant; (c) there is a linear relationship between the independent and dependent variables; (d) there is residual normality; (e) there is constant error variance; and (f) the independent variables are not significantly correlated with one another.

The fourth model is the perfect AFDD linear regression model (PLR), which is mathematically similar to the OLR model, except that X_1 is the observed value of the AFDD for the month between the forecast issue data and the forecast date. We aim to assess how well perfect predictions of the upcoming month's AFDDs could improve the model.

b. Forecast evaluation with Monte Carlo simulations

We broke our historical dataset (observed ice cover, AFDD, and teleconnection indices) into a parameter-estimation dataset (two-thirds of the total) and an evaluation dataset (remaining one-third). We estimated model parameters from the first dataset and error measures from the second dataset to avoid cross-validation problems. We repeated this process 1000 times in a ''Monte Carlo'' approach in which the data selection was randomly generated. We computed model error E_i and forecast skill S_i for each Monte Carlo sample i:

$$E_i = \frac{3}{T}\sum_{m\epsilon T_i} \mathrm{abs}(I_{m,i} - O_m), \tag{4}$$

where T is the length of the total dataset, T_i is the set of all months randomly chosen to be in Monte Carlo simulation i, $I_{m,i}$ is the predicted BOM ice cover for month m from model parameters estimated in Monte Carlo simulation i, and O_m is the observed BOM ice cover for month m. Skill compares a model with climatology (Wilks 1995):

$$S_i = \frac{R_i - E_i}{R_i}100\%, \tag{5}$$

where R_i is the reference error associated with climatology [computed from (4) with $I_{m,i} = C_m$]. If $E_i = 0$, then the skill score is 100%, the maximum value. If $E_i = R_i$, then the skill score is 0%, indicating no improvement over climatology. If $E_i > R_i$, then the skill score is negative, and the method is worse than climatology. If $0 < E_i < R_i$, then the skill score is positive, and the method is better than climatology. The following model discussion and Figs. 1 and 2 use the *mean* model error and *mean* forecast skill over the 1000 simulations.

4. Results

a. Variability in BOM conditions

On average, Lake Ontario has the lowest mean BOM ice cover for all months, and Lake Erie has the highest (Table 1). Lake Ontario has the lowest interannual variability in BOM ice cover, as expressed by the standard deviation; Lake Erie has the highest interannual variability for BOM ice cover for January and February, while Lake Superior has the highest for March.

Ice cover is related to air temperature (AFDDs), wind conditions, and heat storage capacity of the lake; high winds can produce upwelling of warm water and break up ice cover. The higher interannual variability in Lake Erie BOM ice cover is related to its low heat storage capacity (lowest lake volume, Table 1), making it most responsive to interannual atmospheric variations. High interannual variability in Lake Superior BOM March

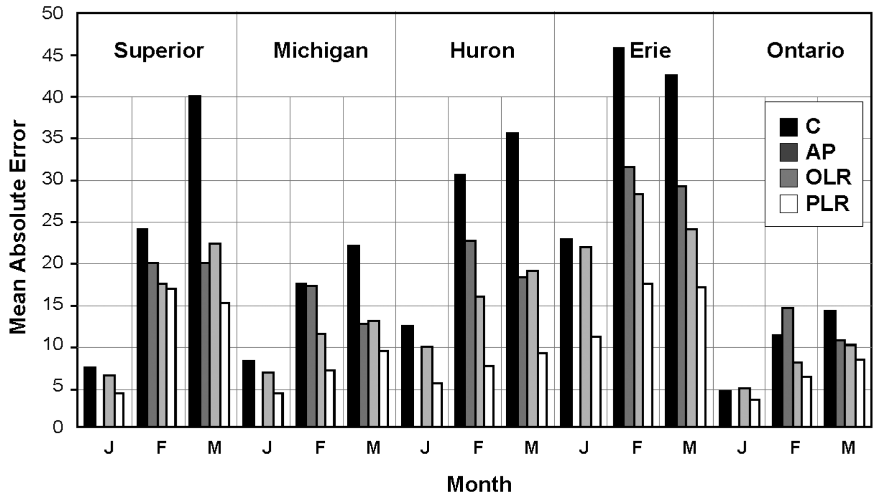

FIG. 1. Mean absolute error (MAE) (%) for prediction schemes (lower MAE indicates better models).

ice cover is likely due to its northern location and to variations in atmospheric conditions in February. During winters, when enough heat has been extracted from the lake by the end of January, extensive ice formation can occur if low air temperatures persist in February and winds are relatively calm. Lake Ontario's low interannual variability in BOM ice cover is due to the combination of low AFDDs; only Lake Erie has lower AFDDs and high lake volume (Table 1), over 3 times that of Lake Erie.

b. Forecasting BOM conditions

The climatology model performs better for lakes with lower interannual variability. Thus, the climatology model error is lowest for Lake Ontario and highest for Lake Erie (Fig. 1). As noted above, atmospheric conditions exert a greater influence over Lake Erie than they do over Lake Ontario because of differences in heat storage capacity. Since the interannual variability increases from January to March in all lakes, model errors also increase for all lakes except Lake Erie, where the interannual variability is at its maximum in February. As the reference model, the mean forecast skill for the climatological model is 0 for all lakes and all months (Fig. 2).

The BOM ice cover data begins in January, so the AP model only applies for February and March. Excluding Lake Ontario, AP model errors are lower than climatology in February and, in many cases, are much lower (Fig. 1). The AP model is also better than climatology in March for all lakes. The forecast skill is larger for March than for February (Fig. 2), suggesting ice conditions from the previous month are more im-

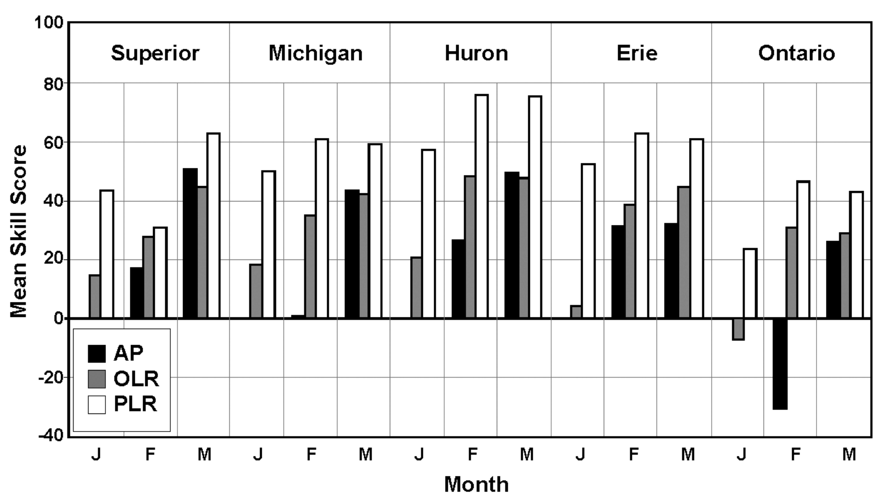

FIG. 2. Mean skill score (%) for prediction schemes (higher skill scores indicate better models).

TABLE 1. Great Lakes characteristics.

	Superior	Michigan	Huron	Erie	Ontario
Mean depth (m)	148	85	59	19	86
Volume (km^3)	12 100	4920	3540	484	1,640
*Latitude (°N)	47.75	44.00	45.50	42.13	43.50
**Feb ice cover (%)	29.7	24.7	43.7	57	14.1
**Feb ice cover std dev	22.4	13.8	21.1	36.6	10.5
**Feb AFDDs	588	332	329	206	239

* Taken from Assel and Rodionov (1998).
** February illustrates the relative relationships in these variables among the Great Lakes. These relationships remain similar for other months.

portant for forecasting March ice conditions than they are for forecasting February ice conditions.

The OLR model has lower errors than climatology in 14 of 15 cases (Fig. 1) and skill scores greater than 0 (Fig. 2). Of these 14 cases, 7 are improvements over the AP model, and 5 have no AP model for comparison. Overall, the OLR models are superior in 11 of the 15 prediction cases, the AP model is the best for 3 cases, and the C model is best in 1 case.

For BOM January predictions, the TNH and the November NAO indices are the only significant predictors. The November TNH statistically exerts more influence than does the NAO. The TNH pattern affects the strength and position of upper airflow near the Great Lakes (Assel and Rodionov 1998). During positive TNH phases, meridional circulation dominates, leading to cooler temperatures and more ice cover (Rodionov and Assel 2000). Positive phases of the November NAO are related to lower ice covers on 1 January. Although positive phases of the NAO are linked to cooler eastern North American temperatures, we suggest that windier conditions may increase mixing in the Great Lakes, inhibiting ice cover.

For BOM February predictions, the BOM January ice cover from each lake is the most important predictor for all lakes, excepting Lake Ontario, where only the December EAWR is valuable. Positive phases of the December EAWR pattern are linked to lower ice covers on 1 February. The December SOI is valuable in predictions of Lake Superior and Lake Huron ice cover. Rodionov and Assel (2000) demonstrate that positive SOI phases (El Niño) are linked to warmer winters in the Great Lakes region, leading to reduced ice extent, and vice versa.

Finally, for BOM March, the BOM February ice cover is most important for all lakes except Lake Erie, where the AFDDs from January are more important. Over Lake Michigan, the January POL also is relevant. Rodionov and Assel (2000) suggest that the POL index is the most important teleconnection in determining mean basinwide ice conditions; during positive POL phases, the polar vortex is strengthened, leading to more zonal flow over the eastern United States, which in turn is related to warmer winter temperatures and less ice cover.

The PLR mean model errors are much lower than for any other model on all lakes, and skill scores are higher (Figs. 1 and 2). Similar to the OLR models, all PLR models are statistically significant at the 5% level and do not violate any of the assumptions for regression analysis. As with the other models, error is highest for Lake Erie and lowest for Lake Ontario. For BOM January forecasts, only an accurate prediction of December AFDDs is needed for all lakes, excepting Lake Michigan, where the November TNH also retains value. For BOM February, the January AFDDs are useful on all lakes, BOM January ice cover is relevant on all lakes except Ontario, and the December SOI remains valuable for Lake Superior. For BOM March, only February AFDDs and BOM February ice cover are needed for each lake.

5. Concluding remarks

If perfect forecasts of the upcoming month's AFDDs were available, the forecast equations for most months would need fewer parameters and the error would be lower for all months. Thus, as numerical weather models improve their accuracy of 30-day forecasts, analysts should consider utilizing these predictions more rigorously in the 30-day ice forecast.

While it is desirable to compare the 30-day forecasts made in this paper with past NIC 30-day forecasts, such an effort is beyond the scope of this study. Besides, NIC 30-day forecasts are not available for comparison except for the last few winters. The NIC should archive the 30-day Great Lakes ice forecasts and 30-day air temperature forecasts used for the ice forecasts for future studies, including forecast validation and comparison with alternative forecast methods.

The work summarized in this paper is given in greater depth in Assel et al. (2004).

REFERENCES

Assel, R. A., 2003a: *An Electronic Atlas of Great Lakes Ice Cover: Winters 1973–2002.* NOAA/Great Lakes Environmental Research Laboratory, Ann Arbor, MI, CD-ROM. [Available online at http://www.glerl.noaa.gov/data/ice/atlas/.]

——, 2003b: Great Lakes monthly and seasonal accumulations of freezing degree days—Winters 1898–2002. NOAA Tech. Memo.

GLERL-127, NOAA/Great Lakes Environmental Research Laboratory, Ann Arbor, MI, 36 pp. [Available online at ftp://ftp.glerl.noaa.gov/publications/tech_reports/glerl-127/tm-127.pdf.]

——, and S. Rodionov, 1998: Atmospheric teleconnections for annual maximal ice cover on the Laurentian Great Lakes. *Int. J. Climatol.,* **18,** 425–442.

——, T. E. Croley II, and S. Drobot, 2004: Improving monthly Great Lakes ice cover outlooks. NOAA Tech. Memo. GLERL-129, NOAA/Great Lakes Environmental Research Laboratory, Ann Arbor, MI, 27 pp.

Rodionov, S., and R. A. Assel, 2000: Atmospheric teleconnection patterns and severity of winters in the Laurentian Great Lakes basin. *Atmos.–Ocean,* **38,** 601–635.

——, ——, and L. R. Herche, 2001: Tree-structured modeling of the relationship between Great Lakes ice cover and atmospheric circulation patterns. *J. Great Lakes Res.,* **27,** 486–502.

Snider, C. R., 1974: Great Lakes ice forecasting. NOAA Tech. Memo. NWS OSD 1, Ocean Services Division, Silver Spring, MD, 106 pp. [Available from the National Technical Information Service, U.S. Department of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, VA 22121.]

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.