7.6 COMPARISON OF GEO-REFERENCED GRIDS OF 1961 – 1990 CANADIAN TEMPERATURE AND PRECIPITATION NORMALS

Ewa J. Milewska[1]*, R. Hopkinson[2] and A. Niitsoo[1]

[1]Meteorological Service of Canada, Toronto, Ontario
[2]Prairie and Northern Region, Regina, Saskatchewan

1. INTRODUCTION

Four sets of grids of normals (1961-1990 averages) of monthly temperature maximum and minimum, and total precipitation were created for Canada. Evenly spaced, geo-referenced grids are better suited for certain purposes than original observations from the irregularly distributed stations. Grids provide better spatial coverage for regional climate change and variability studies; they can be easily integrated into Geographic Information Systems, where they are used to build environmental models for forestry and agriculture, to study climate change scenarios and impacts, to calculate water budgets, etc. Fields of temperature and precipitation normals – the two primary climatological elements – constitute a very important reference field for this kind of scientific research and applications. There are many interpolation techniques and approaches to gridding and mapping; in this case four academic and government institutions used the following schemes:

- SQUARE-GRID technique based on multivariate regression model (University of Waterloo, Ontario);
- PRISM, or Parameter-elevation Regressions on Independent Slopes Model (Oregon State University, Oregon);
- ANUSPLIN model based on thin plate smoothing splines (Canadian Forest Service in Sault Ste. Marie, Ontario);
- IDW, or Inverse Distance Weighting, a type of weighted average interpolator (Prairie and Northern Region, Saskatchewan).

Each approach is based on specific assumptions and often requires additional variables, such as elevation (Digital Elevation Model - DEM). Each of the method have certain known advantages and caveats and can produce different results, for example, some methods perform better in the mountainous or data sparse regions and some put a lot of emphasis on topography alone (Milewska and Hogg, 2001). The purpose of this study is to summarize major characteristics of each method and ensuing grid sets, and then compare the grids in order to establish major biases and differences between them. The intercomparison is performed for southeastern British Columbia (BC) and the Prairie provinces: Alberta (AB), Saskatchewan (SK) and Manitoba (MB), because to date PRISM grids cover only Western Canada. Due to a multitude of very diverse physiographic and topographic features at various dimensional scales, e.g. Rocky Mountains versus small local valleys, or lone hills in the Prairies, this area, especially BC and AB, presents great challenges for interpolation of climatological elements. The interpolation challenge in northern Saskatchewan and northern Manitoba is the sparsity of data.

We hope that this intercomparison will help users to make informed decision about the suitability of the grids for their particular purposes. Only a few examples of preliminary results can be shown here; a full set of intercomparison grids and other details will be presented elsewhere. Since not all grids were given distinct names, for clarity and consistency *the names of the interpolation methods* will be used further in the paper to identify each grid.

2. DATA AND METHODOLOGY

SQUARE-GRID technique was developed in Canada for hydrometric planning purposes (Solomon, 1968). Multivariate regression of climatological station normals was performed by a team of researchers from University of Waterloo, Ontario, on station geographic coordinates and elevation, local slope, regional slope and a number of other derived physiographic parameters, such as distance to ocean, barrier height and shield effect (Seglenieks and Soulis, 2000). The country was split into four regions and the most significant parameters were identified for each region and season. The residuals were then kriged to provide additional local detail. In additon to familiarity with multivariate regression, this method requires knowledge of dominant physiographic features across the country.

PRISM was originally developed by Chris Daly at Oregon State University (Daly et al., 1994). PRISM uses station data and DEM to calculate linear

Corresponding author address: Ewa J. Milewska, Climate Research Branch, Meteorological Service of Canada, 4905 Dufferin Street, Downsview, ON, M3H 5T4; e-mail: Ewa.Milewska@ec.gc.ca

parameter-elevation relationship, which changes locally with elevation as dictated by data points. In addition, the model incorporates spatially and temporally varied inversions and boundary layers to account for special cases that do not fit the generalized regression equations of precipitation increase and temperature decrease with rising elevation, for example, winter temperature inversion in mountain valleys. The model recognizes that in complex terrain, climatic patterns are defined by topographic features, or "facets" of various barrier, slope and aspect characteristics at a range of scales, and creates zones of different climatological regimes. This way it can accommodate abrupt rain shadows on the leeward slopes of mountain ranges and very sharp gradients in temperature along the coast. In the regions where terrain is not an important factor in precipitation amount and pattern, e.g. flat and or gently rolling Prairies, the heights of individual grids are compared against background terrain and categorized according to their effectiveness in enhancing precipitation. Recently PRISM was chosen to produce a new Climate Atlas of United States (Plantico et al., 2000). It is our understanding that currently only Oregon State University has the expertise and capability to model environmental variables using PRISM.

ANUSPLIN model is based on the original thin plate surface fitting technique described by Wahba (1990). M. Hutchinson of Australian National University, Canberra, developed the software and application of the method to mapping climate variables (Hutchinson, 1995). This method fits a surface described by mathematical functions into the data points. The degree of smoothing is optimized objectively by minimizing the predictive error of the fitted function as measured by generalized cross-validation. It incorporates additional dependencies on elevation, in addition to the usual dependence on longitude and latitude. Hutchinson (1991) observed highest errors in the situations of inversed local lapse rates or in proximity to large water bodies, and suggested that appropriate spatial scaling of elevation and aspect effects, such as included in PRISM, clearly deserves further investigation (Hutchinson, 1995). Even though there is a capability to include more dependencies, only three variables are still routinely used in the practical applications (trivariate ANUSPLIN). Dan McKenney (1996) and his team from Canadian Forest Service in Sault Ste. Marie, Ontario, have applied the thin plate smoothing splines method to produce new DEM and Canadian national climate surfaces. The method is universally easy to apply and was used to map climate of various regions and continents: Australia, New Zealand, Africa, China and Europe.

IDW (Isaaks and Srivastava, 1989) is a simple method that performs interpolation using nearby stations. It gives more weight to the closest stations and less to those that are farthest away. The weights are inversely proportional to any power of the distance. With the larger exponents, the closest stations receive larger percentage of the total weight. The choice of the exponent is arbitrary, but traditionally the most common choice of the exponent has been '2', simply because it is very efficient in computations. Ron Hopkinson, Prairie and Northern Region, Regina, Saskatchewan, used an inverse square distance weighting scheme to compute Canada Gridded Climate Data for each month (Hopkinson, 2000). Two passes were made through the data. The initial pass used all data within 70.8 km of a grid point to calculate the weighted (representative) value at the grid point. The intent was to retain in the gridded estimate as much of the spatial variability in the station data as possible. Only in the data sparse areas, stations at the greater distance away were used. This simple method is easy to apply, but it does not attempt to model any physical dependencies, and it does not take into account topographical effects. It gives the best results in station dense areas and smooth terrain.

The traditional way to assess accuracy of the grids would be to perform cross-validation, where grids are re-computed repeatedly with one station or more removed from the set. The grid point values are then compared to the value from the closest station that was temporarily excluded from gridding. The other way would be to compare grid point values to the observations from the independent set of stations. These stations should have never been part of the original set that was used to create the grids. Unfortunately, iterative re-computation of grids is not viable in at least two cases and the search for the independent data set did not render any suitable stations. In fact, the original data set was often maximized by including all seasonal stations, as in an example of the summer only stations in Alberta in a case of PRISM and IDW. Other station networks that were considered were either too small, too short in time, or observations were of questionable quality and would require too much intervention to make them compatible. In the end, it was decided to select one grid as a reference and compare the rest of the grids to it. This approach gives relative differences in the case of temperature and relative ratios in the case of precipitation. It is useful in locating areas of the largest differences or disagreements between the grids, which could be symptomatic of some underlying problems. After comparing interpolation methods, PRISM was chosen arbitrarily as a reference grid. This model seems to be the most comprehensive and exhaustive in its approach, as it takes into account most climatological, physiographic and topographical variables and models their influence on the wide range of three-dimensional scales. This obviously does not necessarily mean that PRISM grids are automatically superior to the others. In certain regions and in data dense areas, very simple methods can often produce very good results.

All four techniques used monthly minimum and maximum temperatures from the National Climate Archive, but they differed in data pre-processing. SQUARE-GRID, ANUSPLIN and PRISM first computed normals at the stations and then performed gridding. The constraint of having at least 15-20 years of data is usually assumed and missing data are not estimated. IDW did the reverse; it first did gridding for each month, using all available station reports, and then computed normals at the grid points. No adjustments were made to the data to account for instrumental or other known biases. For precipitation, SQUARE-GRID did not use monthly records from the archive, as did the other three methods. Instead, it used stations from the rehabilitated data set, which represents a much smaller subset of the stations that are stored in the archives. Fewer than five hundred stations across Canada were rehabilitated, which means that observations of precipitation were adjusted to account for replacement of rain gauges, trace precipitation, and regionally variable snow water equivalent ratio (Mekis and Hogg, 1999). For example, the standard Type B rain gauge that was introduced in the 1970s had different undercatch, evaporation and funnel loss that the one previously used. The resultant rehabilitated precipitation is substantively altered from the original archive data so comparisons are presented for information only. It is not possible to attribute characteristics that arise from the gridding technique as opposed to differences in the underlying precipitation data. The resolution of grids also differs from one model to another: SQUARE-GRID – 1' (arc minutes) or 5' (1' grid was used in intercomparison), PRISM – 2.5', ANUSPLIN – 5', and IDW – 30'. As already mentioned, PRISM's and IDW's analyses also incorporated seasonal stations.

3. RESULTS

In general, all four methods rendered similar results across much of the plains area, which is characterized by high station density and smooth, low topographic relief. Temperatures there agree within one degree Celsius, and precipitation within a few percent with the exception of the SQUARE-GRID.

Major discrepancies between grids are discussed on a case-by-case basis. Original digital maps of differences for temperatures and ratios for precipitation are in color. Grayscale paper prints cannot distinguish between opposite biases, e.g. for extreme negative or positive biases both dark blue and dark red show as black. They can still illustrate relatively well the extent of major patterns and can be used in conjunction with the description of the results.

3.1 Maximum Temperature

IDW vs. PRISM Tmax:

1) In the main cold months, November to March, there is a strong negative bias, when IDW produced lower temperatures over the higher terrain of northern Alberta and in January to March in northern SK and MB. This may be a reflection of implied temperature inversion conditions in PRISM at modest elevations associated with the strong Arctic inversion. This inversion is not reflected in the observations, which are taken primarily at low elevations.

2) For the months February to August, there is a strong positive bias in northern Manitoba on the shore of Hudson Bay around Nelson River. This may be a combination of two effects: a lack of data in this region – interpolation over larger distances increases uncertainty, especially in IDW; and implied influence of Hudson Bay by PRISM, as suggested by the shape of the pattern over the low lying Nelson River delta. While the infiltration of cold air is possible during the ice-free period July-August, it seems that PRISM might have otherwise exaggerated the influence of the Bay, which is frozen solid for the most of the year. Then again, cold air might be pooling in these lowlands.

3) For all months, the IDW maximum temperatures are warmer over the mountains of southeastern BC. The IDW gridded values have no elevation factor built in other than what is related to the data, which is usually observed in valley bottoms.

4) IDW gridded maximum temperatures tend to be colder over Lake Winnipeg during the main summer months of June to August. This is because in PRISM the reporting station in the centre of the Lake was intentionally removed from analysis as not representative of the inland temperature over the area. Admittedly, this practice of removing stations on the lake is open to debate: these high resolution grids are able to discriminate between land and large lakes and the lake grid points should represent marine microclimate. In addition, numerous lakes in northern parts of Prairie provinces are likely to influence and moderate continental climate there during the ice-free season.

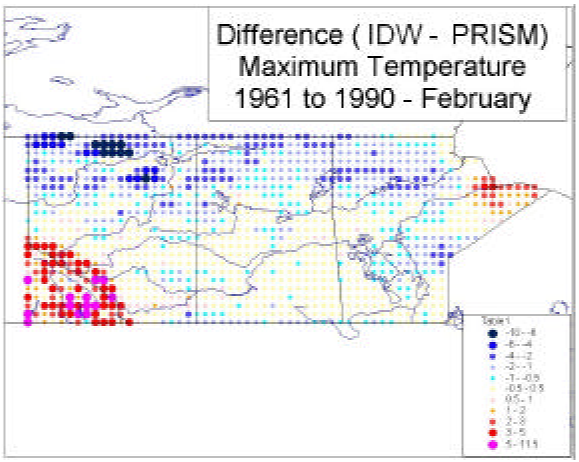

Figure 1.

ANUSPLIN vs. PRISM Tmax:

1) Similar effect as described in <u>IDW vs. PRISM Tmax: 1)</u> for the period from November to April. Trivariate ANUSPLIN does not recognize the winter Arctic inversion over the hills of northern AB.

2) Similar to <u>IDW vs. PRISM Tmax: 2)</u> for all months of the year.

3) In southeastern BC, southwestern section appears slightly cooler and the northeastern ranges warmer.

4) In January, there are numerous areas across the southern Prairies where the ANUSPLIN values are warmer than the corresponding PRISM values. As in <u>ANUSPLIN vs. PRISM Tmax 1)</u>, the reason could be associated with the inverse lapse rate that PRISM models over small hills, and that ANUSPLIN does not.

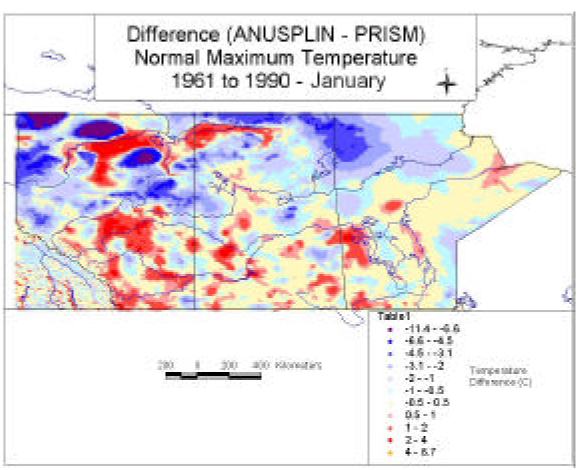

Figure 2.

SQUARE-GRID vs. PRISM Tmax:

1) Similar to <u>IDW vs. PRISM Tmax: 1) and ANUSPLIN vs. PRISM Tmax: 1)</u>, except that there is a pronounced *warm* bias in northern SK and MB in November to January.

2) Similar to <u>2) in both IDW vs. PRISM Tmax and ANUSPLIN vs. PRISM Tmax</u>.

3) In southeastern BC valleys appear a few degrees colder and ridges a few degrees warmer.

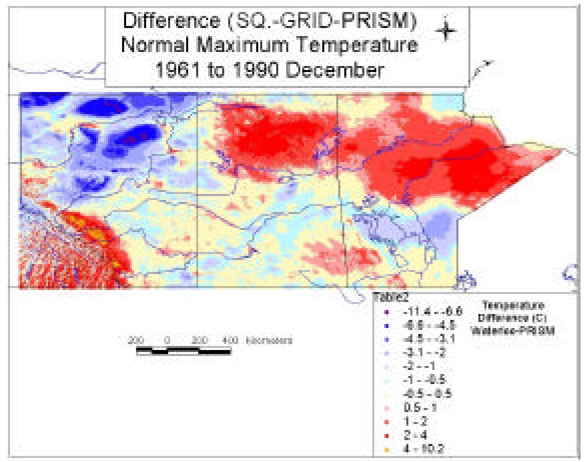

Figure 3.

3.2 Minimum Temperature

IDW vs. PRISM Tmin:

1) Similar to <u>IDW vs. PRISM Tmax: 1)</u>. Overall the agreement is poorer in the northern regions in the data sparse area.

2) Similar to <u>IDW vs. PRISM Tmax: 2)</u>, but only for April to July.

3) Similar to <u>IDW vs. PRISM Tmax: 3)</u>.

4) Minimum temperatures over Lake Winnipeg are warmer in IDW from August to October for the reason explained in <u>IDW vs. PRISM Tmax: 4)</u>. Otherwise, for the plains region, the differences between the gridded minimum temperatures are not excessive for all months.

ANUSPLIN vs. PRISM Tmin:

The results are very similar to the ones described for <u>ANUSPLIN vs. PRISM Tmax</u>.

SQUARE-GRID vs. PRISM Tmin:

The results are very similar to the ones described for <u>SQUARE-GRID vs. PRISM Tmax</u>.

3.3 Precipitation

IDW vs. PRISM P:

1) Systematic negative bias in southeastern BC. IDW interpolates observations from stations that are located in the valleys and report lower amounts than amounts that occur over mountain tops. IDW scheme is not capable of augmenting precipitation amounts in the orographic enhancement zones.

2) Difficult to explain positive bias in eastern Manitoba from September to February.

3) In general, not as good agreement in the north in data sparse areas. Some examples include: a persistent positive bias over the hills southwest of Lake Athabasca in October and January to April; in northern Saskatchewan in November, February and April.

4) A positive bias right along the Canada – US border, especially noticeable in November to March. PRISM included US gauges, however, they underreport in comparison to Canadian gauges. The discontinuity along the border is obvious, even after some gauges were eliminated from the analysis.

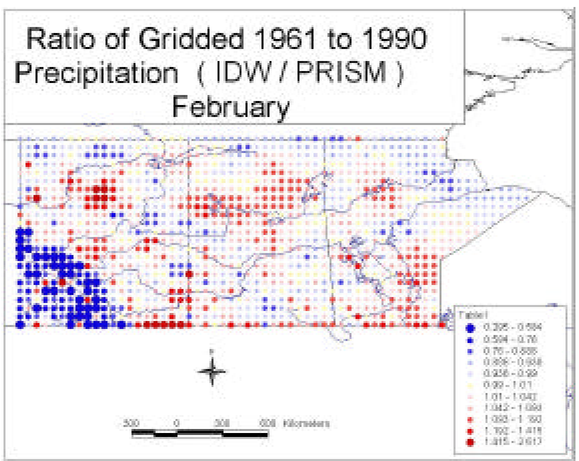

Figure 4.

ANUSPLIN vs. PRISM P:

1) In southeastern BC, there is a positive bias in the southwestern part of the area and negative bias along the Rocky Mountain range. Based on observations from the valleys, ANUSPLIN might not be increasing precipitation enough over the highest mountain ranges.

2) A negative bias around Hudson Bay November to May could be caused by a lack of data there.

3) A widespread negative bias from May to September in northern AB might be a result of not having the summer only stations in ANUSPLIN grids. PRISM reported an increase of precipitation in the north after these stations were added.

4) A wide "band" of "patchy" positive biases from northern Alberta to southeastern Manitoba from October to April. At least some of the "patches", e.g. in Northern Alberta are over the hills, which suggests that ANUSPLIN scheme assigns higher precipitation amounts to this altitudes than PRISM.

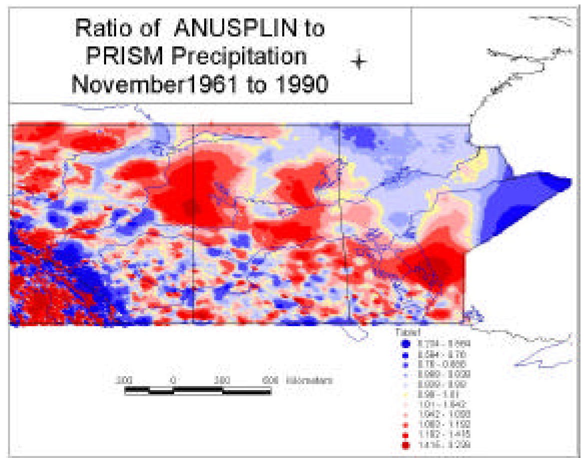

Figure 5.

SQUARE-GRID vs. PRISM P:

SQUARE-GRID's use of rehabilitated precipitation data set may be responsible for the unusually extensive differences.

1) A persistent year-round negative bias over southeastern BC except near US border.

2) A widespread positive bias over southern half of the Prairie provinces, MB near Hudson Bay, and northern AB.

3) A negative bias over northern SK and MB, especially evident from June until January.

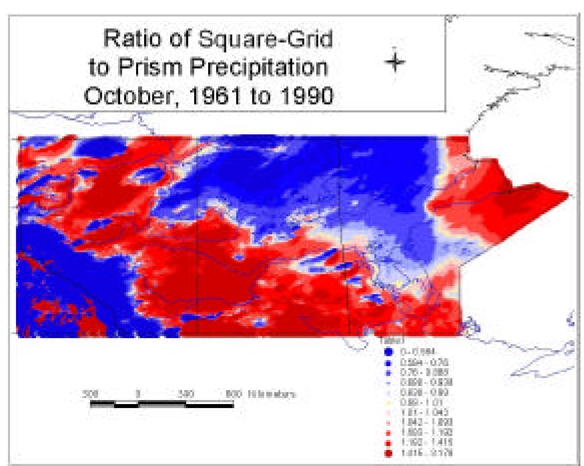

Figure 6.

4. CONCLUSIONS

This preliminary study successfully identified regions of persistent biases in IDW, ANUSPLIN and SQUARE-GRID in comparison to PRISM. Indirectly, especially in cases when all three models exhibited similar behaviour towards PRISM, it was also possible to compare the models among themselves.

The three grids were most divergent from PRISM in the data sparse areas and areas of high, complex terrain. None of the three grids accounted for the presence of strong winter Arctic temperature inversions, which affect temperature over the hills of northern AB. IDW was too warm in the mountains because it only used reports from low elevation stations and did not incorporate the environmental temperature lapse rate. This problem was also reflected in the too low precipitation amounts there, as IDW cannot account for orographic enhancement. The results from all models, including PRISM, seem problematic around Hudson Bay in the Nelson River area. Among all three grids, trivariate ANUSPLIN temperature grids seem to agree the best with PRISM in the high mountains of BC. ANUSPLIN precipitation grids display greater differences. SQUARE-GRID precipitation grids show the largest departure from PRISM grids, which could be related to the fact that rehabilitated data set was used in the SQUARE-GRID case. This set consists of a fewer number of stations, as well precipitation amounts were adjusted to account for changes in instrumentation, etc. Because of this, larger differences can be automatically expected, regardless of which interpolation method is used.

All of the gridded temperature fields generally agree within one degree Celsius across much of the plains area where station density is high and topographic relief is low. Similarly, with the exception of the SQUARE-GRID, the gridded precipitation fields, for all methods, agree within a few percent across plains.

REFERENCES

Daly, C., Neilson, R.P. and D.L. Phillips. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor*. **33**: 140-158.

Hopkinson, R. 2000. Canadian Gridded Climate Data. *WWW* http://www.cics.uvic.ca/climate/data.htm (published 21March 2001)

Hutchinson, M.F. 1991. Climatic analyses in data sparse regions. In: Muchow, R.C. and Bellamy, J.A. (eds) *Climatic Risk in Crop Production*, CAB International, 55-71.

Hutchinson, M. F. 1995. Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geographical Information Systems* **9(4)**: 385-403.

Isaaks, E. H. and R. M. Srivastava. 1989. An introduction to applied geostatistics. *Oxford University Press*: 561 pp.

McKenney, D.W., Mackey, B.G.and R.A. Sims. 1996. Primary databases for forest ecosystem management – examples from Ontario and possibilities for Canada: NatGRID. *Environmental Monitoring and Assessment* **39**: 399-415.

Mekis, E. and W.D. Hogg. 1999. Rehabilitation and analysis of Canadian daily precipitation time series. Atmosphere-Ocean **37(1)**: 53-85.

Milewska, E. and W.D. Hogg. 2001. Spatial representativeness of a long-term climate network in Canada. Atmopshere-Ocean **39(2)**: 145-161.

Plantico, M.S., Goss, L.A., Daly, C. and G. Taylor. 2000. A new U.S. Climate Atlas. *Proc., 12th AMS Conference on Applied Climatology, Amer. Meteorological Soc., Asheville, NC, 8-11 May: 247-248.*

Seglenieks, F. and R. Soulis. 2000. Generation of square grid normals for Canada – Phase I. *U. Waterloo contractor's report to Climate Research Branch, Meteorological Service of Canada, Downsview*.

Solomon, S.I., Denouvilliez, J.P., Chart, E.J., Woolley, J.A. and C. Cadou. 1968. The use of a square grid system for computer estimation of precipitation, temperature, and runoff. *Water Resources Research* **4(5)**: 919-929.

Wahba, G. 1990. Spline Models for Observational Data. *CBMS-NSF Regional Conference Series in Applied Mathematics 59, SIAM, Philadelphia, Pennsylvania*.