

## 6.2 EVALUATION OF ALTERNATIVE SPATIAL MODELS OF VAPOUR PRESSURE IN CANADA

Daniel W. McKenney<sup>1\*</sup>, M.F. Hutchinson<sup>2</sup>, P. Papadopol<sup>1</sup>, D.T. Price<sup>3</sup>

<sup>1</sup> Canadian Forest Service, Sault Ste. Marie, Ontario

<sup>2</sup> Centre for Resource and Environmental Studies, The Australian National University, Canberra, Australia

<sup>3</sup> Canadian Forest Service, Edmonton, Alberta

### 1 ABSTRACT

Spatial variation in measurements of atmospheric humidity provide a basis for creating interpolated estimates of humidity variables such as vapour pressure. Here we compare several different thin plate spline models of 1961-90 monthly vapour pressure normals observed at Canadian climate stations including a model where observed vapour pressures were first adjusted to sea-level using standard barometric correction. Using a set of 30 stations withheld from the fitting procedure, we performed independent validations of each model. All models in fact had good signals and would generally be considered as acceptable according to standard diagnostics. In general, withheld data could be predicted with mean errors of less than 10 Pa. The best model was a square root transformation of the vapour pressure field and a trivariate spline with a spatially varying dependence on elevation as determined by the data network. The maps generated as the final products used all data points except for three stations identified as potentially anomalous during the model-building phase.

### 2 INTRODUCTION

Atmospheric vapour pressure,  $e_a$ , is one of many climatic parameters of interest to meteorologists, ecologists and agriculturists. Vapour pressure is defined as the pressure exerted on any surface by gaseous water molecules (Linacre et al., 1977). Water vapour obeys Dalton's Law, which states that the total pressure of a mixture of gases equals the sum of the partial pressures exerted by each constituent gas (List 1949). Hence, vapour pressure is the partial pressure exerted by water vapour in a mixture of other atmospheric gases (which together make up "dry air") to yield the total barometric pressure.

Thus vapour pressure is the pressure exerted by moisture in the air, and increases with both atmospheric temperature and moisture content. It is related to relative humidity in that the latter is the ratio of the actual vapour pressure to saturation vapour pressure at the same temperature. Water vapour in the air can be described directly or indirectly in several ways including:

- 1) Water vapour pressure (the pressure exerted on any surface by gaseous water molecules);
- 2) Vapour pressure deficit (the difference between the saturation water-vapour pressure and the actual water-vapour pressure);
- 3) Relative humidity (the ratio of the actual water-vapour pressure to the saturation water-vapour pressure at dry-bulb temperature);
- 4) Absolute humidity, or vapour density (the mass concentration of water vapour per unit volume of air);
- 5) Mixing ratio (the ratio of the masses of water vapour to dry air, respectively);
- 6) Specific humidity (the ratio of the masses of water vapour to moist air, respectively);
- 7) Dewpoint (the temperature to which the air must be cooled for the water vapour content to reach saturation).

Saturation is a dynamic equilibrium state where water is constantly evaporating and condensing. The intensity of this process is measured directly by vapour pressure and indirectly by temperature. At saturation, air holds the maximum amount of water vapour possible. Hence, at saturation, the dew point temperature equals the air temperature. Further, the relative humidity is 100%, and the vapour pressure is equal to the saturation vapour pressure.

Water vapour is generally introduced into the bottom of the atmosphere by evaporation from a moist surface. It becomes vertically distributed by turbulent mixing processes including free convection of warm surface air and forced advection due to wind, as well as synoptic-scale circulation of air masses (see Stull, 1988 for details).

The unit measurement of vapour pressure is the kilopascal: one kilopascal is equal to 10 millibars, 0.2953 inches of mercury, or 0.145 pounds per square inch. Accurate measurement of atmospheric humidity is as much art as science (Weast, 1973), with the preferred instrument being a high quality aspirated psychrometer, although for most practical purposes dew-point hygrometers are considered more reliable.

Vapour pressure deficit (VPD), defined as  $e^*(T_a) - e_a$ , where  $e^*(T_a)$  denotes saturation vapour pressure at air temperature  $T_a$ , is another measure of atmospheric humidity often used to model evaporation of water from moist surfaces, because it approximates

---

\* Corresponding author address: Daniel McKenney,  
Canadian Forest Service, Sault Ste. Marie,  
Ontario dmckenne@nrcan.gc.ca

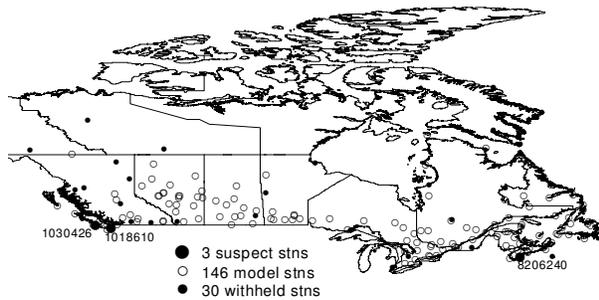


Figure 1. Geographic locations of 176 vapour pressure stations: 146 stations used for surface fitting and 30 stations withheld for validation. All but three suspect stations were used in the final model.

the water potential gradient between saturated air at or just above a moist surface and the vapour pressure of the “free air”. The VPD multiplied by an appropriate transfer coefficient then allows the evaporative flux to be estimated. In practice, however, it is often important to calculate this gradient more precisely by replacing  $T_a$  with the surface temperature. In the case of plants, transpiration is a process that evaporates water molecules from a moist cell surface (at leaf temperature) and transfers them to the free air, although the pathway may be restricted by structures such as stomata embedded in the leaf surface. Hence atmospheric vapour pressure is often used as an important driving variable for many models of plant transpiration and canopy evapotranspiration.

Here, we report the development of a high resolution spatial vapour pressure data set suitable for driving spatialized process models applied to Canada. Rasmusson (1967) provides discussion and maps of the spatial nature of water vapour in North America. Castellvi et al. (1996, 1997) describes some methods to develop regional vapour pressure deficit models. Castellvi et al. developed models for estimating daily vapour pressure deficit at weather stations in which only temperature and precipitation are measured. There are a range of expressions for estimating saturation vapour pressure as a function of temperature at standard pressure. They include the Goff-Gratch (1946) formulations as presented in the Smithsonian Tables (List, 1949) or one of the many empirical equations which have appeared in the literature (e.g., Tetens 1930; Murray, 1967; Richards, 1971; Tabata, 1973; Campbell 1977; Lowe, 1977; Rasmussen, 1978), all of which use the Goff-Gratch formulation as a standard for comparison. The Goff-Gratch approach is generally considered the most accurate, suitable for generating tables, but it is computationally inefficient when used for “on-the-fly” calculations, as in process models. A complete survey of those equations appears in Sargent (1980). The 6<sup>th</sup> order polynomials of Lowe (1977) are now generally accepted to provide excellent agreement

with the Clausius-Clapyron and Goff-Gratch equations while being computationally much faster than the latter.

Our objective is to describe the development of a national monthly mean vapour pressure (VP) model for Canada and hence provide some baseline vapour pressure reference data for Canada away from the stations for which it is calculated. This VP field is available as one of several climate fields available from the Canadian Forest Service, Landscape Analysis and Applications Section – Sault Ste. Marie ([http://www.glf.cfs.nrcan.gc.ca/landscape/climate\\_models\\_e.html](http://www.glf.cfs.nrcan.gc.ca/landscape/climate_models_e.html)). It was derived using the ANUSPLIN thin plate smoothing spline algorithms of Prof. Michael Hutchinson, Centre for Resource and Environmental Studies at The Australian National University (<http://cres.anu.edu.au/>) and station data from Environment Canada’s Climate Data Centre in Downsview, Ontario.

### 3 DATA AND METHODS

Mean hourly data by month from 176 stations were available from Meteorological Service of Canada’s Monthly Climate data base. Figure 1 shows the locations of the stations including those withheld from the initial model-building phase for validation purposes. The data are a synthesis of hourly vapor pressure values derived from dew point temperature through conversion tables. For air temperatures colder than  $-37^{\circ}$  C where dew-point temperatures were not available, a value of 10 Pascal (Pa) was assumed. The monthly values were averaged for the 1961 to 1990 period using any station with five or more years of records, regardless of missing year-month totals or counts. The values range from 0.07 kPa in winter (northern Arctic) to 1.78 kPa in summer (southern B.C.).

Table 1 illustrates, from 5 selected stations, the seasonal changes in vapor pressure across Canada, which are very much influenced by air streams from three distinct sources: the Pacific, Atlantic and Arctic Oceans. Over land, atmospheric humidity is invariably higher in summer, because the warmer air can hold more moisture and higher surface temperatures increase evaporation and transpiration from lakes, rivers and vegetation. Hence, for much of the country, the humidity of air near the surface is influenced by a combination of synoptic processes and seasonal changes in vegetation, as well as proximity to inland water bodies.

In winter, atmospheric humidity is highest on the two coasts, particularly the Pacific, with values of 0.67 kPa in Vancouver BC and 0.40 kPa in Halifax. In the mid-continental regions, the low temperatures, often due to outflows of frigid Arctic keep the interior humidity very low, with values of 0.16 and 0.18 kPa in Winnipeg and Saskatoon respectively. Humidity increases rapidly in spring, and high throughout the country recording values of 1.44 kPa and 1.66 kPa on the east and west coasts respectively and 1.57 kPa in Portage La Prairie. In October as surface temperatures decrease and after deciduous vegetation has shed its foliage, vapor pressure decreases more rapidly in the interior and slower on the Pacific and Atlantic coasts.

Index	Longitude	Latitude	Elevation m	Jan kPa	Apr kPa	Jul kPa	Oct kPa	Mean Annual kPa	Station Name	Province
1108447	-123.1667	49.1833	3	0.67	0.86	1.44	1.06	1.02	VANCOUVER A	BC
3012205	-113.5833	53.3000	715	0.22	0.49	1.29	0.57	0.63	EDMONTON INT A	ALTA
4057120	-106.6833	52.1667	501	0.18	0.51	1.29	0.56	0.62	SASKATOON	SASK
5012320	-98.2670	49.9000	270	0.16	0.54	1.57	0.67	0.71	PORTAGE LA PRAIRIE A	MAN
7025250	-73.7500	45.4667	31	0.25	0.61	1.73	0.86	0.87	DORVAL A MONTREAL	QUE
8205090	-63.5000	44.6333	51	0.40	0.63	1.60	0.99	0.90	DARTMOUTH A HALIFAX	NS

Table 1. Vapour pressure values for 5 selected stations across Canada by season.

Of the 176 available stations, 25% have between 5-10 years of record, 11% have between 11-14 years of record, 15% had between 15-19 years of record and 49% have more than 20 years of record length.

ANUSPLIN is a suite of Fortran programs for applying thin plate spline data smoothing techniques to noisy multi-variate data. It is now quite widely used as an interpolation technique for climate data in Australia, China and parts of southeast Asia, South America, Africa, Europe and New Zealand and Canada (e.g. Hutchinson, 1995; 1998a; 1998b; New et al 2002; Price et al. 2000; McKenney et al. 2001). Thin plate splines can be viewed as a generalization of a multivariate linear regression model in which the parametric model is replaced by a “suitably” smooth nonparametric function. It becomes a spatially continuous interpolation methodology when position variables (e.g. latitude and longitude) are used as independent variables in a model. A comprehensive description of thin plate smoothing splines (not to be confused with univariate splines) has been given by Wahba (1990). A general model for a thin plate spline function  $f$  fitted to  $n$  data values  $z_i$  at positions  $x_i$  is given by (Hutchinson, 1995):

$$z_i = f(x_i) + \varepsilon_i \quad (i = 1, \dots, n) \quad [1]$$

where the  $x_i$  typically (although not necessarily) represent longitude, latitude and a suitably scaled elevation. The  $\varepsilon_i$  are zero mean random errors that account for measurement error as well as deficiencies in the spline model, such as local effects below the resolution of the data network. Importantly for the case considered here, the  $\varepsilon_i$  include errors due to the shortness of record of half of the data points. Hutchinson (1995) presents an analysis of these errors. The degree of smoothing in the fitted surfaces is determined by minimizing the Generalized Cross Validation (GCV) statistic, a measure of the predictive error of the surface. It is calculated by implicitly removing each data point and summing the square of the difference of each omitted data point from a surface fitted to all the remaining data points (see Hutchinson

and Gessler, 1994). The procedure also provides an estimate of the variance of the  $\varepsilon_i$ .

Model assessments are generally done through examination of automatically generated model diagnostics and in some cases withholding data from initial models and comparing estimated versus observed values at these locations. Three automatically generated diagnostics are described here. The SIGNAL is the degrees of freedom of the fitted spline and varies between zero and the number of stations used for interpolation. Hutchinson and Gessler (1994) suggest that the signal should generally not be greater than about half the number of data points. Models with a good signal provide a balance between data smoothing and exact interpolation. Models with a poor signal are typically closer to an exact interpolation and can result in steeper gradients between stations. Exact interpolation also implies no source data errors, which is normally an unrealistic conclusion. The square root of the GCV (RTGCV) is a measure of the predictive error of the surface. It is a robust, but somewhat conservative estimate of overall prediction error, since it includes the errors in the data as given by the variance of the  $\varepsilon_i$ . The root mean square model error (RTMSE) is an estimate of standard error after the estimated data error, as ascribed to the  $\varepsilon_i$ , has been removed. Since the  $\varepsilon_i$  include both model error and data error, the true error of the fitted model lies between the RTGCV and the RTMSE. Both RTGCV and RTMSE can be less accurate estimates of overall error when the data points are very unevenly distributed as shown in Figure 1.

If data are withheld from the model fitting process ANUSPLIN can also generate a number of comparative statistics for the withheld data. These statistics are independent of model assumptions, but do depend on how the withheld data are chosen. Two reported here are the Mean Error (ME) and the Root Mean Square Error (RMSE). The Mean Error is an indicator of bias in the model. The Root Mean Square Error is sensitive to the size of outliers and is an indicator of the magnitude of extreme errors (i.e., lower RMSE indicates greater central tendency and generally smaller extreme errors).

If the withheld data are chosen randomly (not the case here) the RMSE will be similar to the RTGCV.

An important feature of ANUSPLIN is its capacity to generate Bayesian standard error estimates at specific point locations or for maps (see Hutchinson, 1995). These error estimates allow for the density of the supporting data points. We map these estimates for the final selected model. A complete description of the statistics associated with thin plate smoothing splines can be found in the ANUSPLIN manual (see also Hutchinson, 1995). For some models we used a square root transformation of the vapour pressure field in the surface fitting procedure. This is useful in reducing the positive skew in naturally positive data such as vapour pressure and precipitation. Hutchinson (1998a) found the square root transformation to reduce interpolation error in precipitation analyses.

Our national vapour pressure model was completed following experimentation with several alternative ANUSPLIN model formulations and investigation into certain data fidelity questions that arose during the initial analysis phase. The models reported here were as follows:

- Model 1 – a bivariate model (latitude and longitude –  $x, y$ );
- Model 2 – a bivariate model with elevation as a covariate (implying a constant, non-spatially varying lapse rate);
- Model 3 – a trivariate spline incorporating a spatially varying dependence on elevation ( $x, y, z$ );
- Model 4 – a trivariate spline ( $x, y, z$ ) with a square root transformation of the vapour pressure values;
- Model 5 – identical with Model 4 but using a data set from which 3 suspect stations were removed.
- Model 6 – a bivariate spline where the effect of elevation was removed by bringing the observed vapour pressure values to a notional sea level pressure prior to the surface fitting.

Model 6 was an attempt to determine whether the empirical fit of elevation-dependence achieved by ANUSPLIN was sufficient to account for an effect that could be explained by physical laws. Observing the Ideal Gas Law, each monthly vapour pressure datum was multiplied by an expression used to estimate the decrease in barometric pressure with elevation, using an equation provided by R.L. Snyder ([http://biomet.ucdavis.edu/conversions/humidity\\_conversion.htm](http://biomet.ucdavis.edu/conversions/humidity_conversion.htm))

$$e_a(0) = e_a(z) / [(1.0 - 0.0065z/293.0)^{5.26}] \quad [2]$$

where  $e_a(z)$  was the vapour pressure reported at climate station elevation  $z$ .

Besides assessments of the model diagnostics, models were evaluated by withholding 30 stations from initial runs. Following previous convention (Hutchinson, 1995) the validation data were selected using SELNOT, part of the ANUSPLIN package. SELNOT identifies locations (data points) that maximize the Euclidean distance between the positions. Normally SELNOT is used to select stations (knots) for thin plate spline models with very large datasets. In this case the

SELNOT algorithm was used to withhold 30 stations from the original dataset. These tend to be the most remote or extreme positions and hence generally critical to include in spatial interpolation models. Withholding these stations is therefore an exacting test and the Root Mean Square Error of these stations can be expected to be somewhat larger than the RTGCV. Locations of these and the remaining stations used for surface fitting are shown in Figure 1. Vapour pressure values were estimated at each of the locations of the 30 withheld stations and compared to the observed values plotted for each month. These residuals are reported in terms of mean errors, root mean square error and as a percentage of the root mean square error calculated over the mean of validation data. The latter metric is commonly used to help evaluate precipitation models (e.g., Price et al. 2000).

Maps of the fitted surfaces and their standard errors were generated using an ~ 10km digital elevation model (DEM) developed from the 1:250,000 scale national topographic data. (also available at: [http://www.glf.cfs.nrcan.gc.ca/landscape/topographic\\_models\\_e.html](http://www.glf.cfs.nrcan.gc.ca/landscape/topographic_models_e.html))

## 4 RESULTS AND DISCUSSION

Table 2 lists the Mean Error (ME) and Root Mean Square Error (RMSE) values obtained as the differences between observed and interpolated values for the 30 stations withheld from the analyses. The bivariate and trivariate thin plate splines (Model 1, 2, 3 and 6) were inferior to the splines with the square root transformation (Model 4 and 5) in terms of their ME and RMSE. The square root transformation is appropriate because vapour pressure values are always positive and larger values tend to have larger errors. During the winter (December, January, February) values of ME ranged from 0.003 to 0.009 for Models 1 and 2, from -0.004 to -0.008 for Model 3, and from 0.001 to 0.002 for Models 4 and 5. The RMSE values ranged from 0.039 and 0.047 to 0.060 for Models 1 and 2, from 0.052 to 0.064 for Model 3 and from 0.032 to 0.039 for Models 4 and 5. During the summer (June to August) the RMSE values for Models 4 and 5 ranged from 0.066 to 0.077 compared to 0.087–0.099 for Model 1, 0.099–0.114 for Model 2 and 0.080–0.093 for Model 3.

An assessment of model accuracy was also made by calculating, for each month, the RMS residual of the validation data as estimated from the fitted surfaces not including these data. The RMS residuals were expressed as a percentage (PRMS) of the mean vapour pressure for the month over the validation data set. The superiority of Models 4 and 5 was emphasized by this metric presented in Table 3. The inclusion of elevation as a variable in Models 3, 4 and 5 generally improved the model fits compared to Models 1 and 2, as would be expected. Model 6, which adjusted for elevation effects prior to the surface fitting, also performed better than Models 1 and 2, and generally better than Model 3. For all six models the PRMS is higher in the winter than in the summer months. This might be explained by seasonal changes in continentality effects between coastal areas and central Canada.

Month	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	ME	RMSE										
Jan	0.004	0.056	0.000	0.049	-0.006	0.064	0.002	0.039	0.002	0.039	0.008	0.038
Feb	0.003	0.048	-0.002	0.042	-0.008	0.061	0.001	0.033	0.001	0.032	0.007	0.030
Mar	0.010	0.047	0.006	0.037	0.001	0.052	0.009	0.027	0.009	0.026	0.015	0.032
Apr	0.013	0.047	0.008	0.032	0.004	0.041	0.010	0.027	0.010	0.027	0.014	0.035
May	0.008	0.066	0.001	0.051	-0.005	0.060	0.002	0.043	0.002	0.043	0.010	0.051
Jun	-0.007	0.099	-0.013	0.090	-0.017	0.093	-0.009	0.077	-0.009	0.077	-0.003	0.082
Jul	-0.021	0.096	-0.031	0.114	-0.019	0.089	-0.012	0.077	-0.012	0.077	-0.017	0.096
Aug	-0.007	0.087	-0.016	0.092	-0.015	0.080	-0.007	0.066	-0.007	0.066	-0.001	0.084
Sep	0.004	0.068	-0.005	0.064	-0.010	0.067	-0.001	0.051	-0.001	0.051	0.006	0.067
Oct	0.010	0.049	0.002	0.043	-0.006	0.053	0.001	0.038	0.001	0.038	0.012	0.048
Nov	0.004	0.044	-0.004	0.038	-0.010	0.050	-0.004	0.035	-0.004	0.035	0.006	0.042
Dec	0.009	0.039	0.004	0.042	-0.004	0.052	0.002	0.032	0.002	0.032	0.012	0.041

Table 2. Summary of Mean Error (ME) and Root Mean Square Error (RMSE) for 30 stations withheld from analyses.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Model 1	22.4	17.4	14.5	10.6	9.8	10.6	10.5	9.6	9.8	9.7	13.8	19.8
Model 2	19.4	15.3	11.4	7.0	7.7	9.5	9.8	7.9	6.9	6.6	8.8	14.3
Model 3	25.1	22.2	16.0	9.1	9.1	9.8	7.6	6.9	7.3	8.4	11.9	17.6
Model 4	15.2	11.8	8.4	6.1	6.5	8.1	6.7	5.7	5.6	5.9	8.0	10.7
Model 5	15.2	11.8	8.3	6.1	6.4	8.2	6.6	5.7	5.5	5.9	8.2	10.9
Model 6	14.9	10.7	9.3	7.4	7.4	8.4	8.0	7.0	7.0	7.1	9.3	13.5

Table 3. Percentage Root Mean Square Error for the 30 withheld stations.

Table 4 lists the SIGNAL, surface mean, RTGCV and RTMSE for each model for all 176 stations (no data were withheld from these models). The residual degrees of freedom are slightly in excess of half the number of data points for Models 3 and 4, and less than half the number of data points for Model 1. For all models, the lowest signal is recorded in May and October and the highest in June, July and August, although model 4 is more balanced. As expected, the RTGCV values in Table 4 are somewhat smaller than the RMSE validation errors in Table 2. The RTGCV varies between .023 to .061 kPa for Model 1 and Model 2, .022 to .052 kPa for Model 3, and .014 to .023 kPa for Model 4 and Model 5. The signals are slightly better for Model 5 than Model 4. In this application, all models appear to have performed reasonably well. The RTMSE varies from 0.011 to 0.030, 0.029 and 0.026 kPa for Models 1,2 and 3 respectively and from 0.006 to 0.011

for Model 4. The RMSEs of Model 5 are similar to Model 4. These results in conjunction with the withheld data analysis (Tables 2 and 3) confirm Model 5 slightly outperforms Models 4 and 6, and hence is our preferred model.

Table 5 shows the estimated and observed values for the three potentially anomalous stations withheld from the final Model 5. The greatest differences are in the summer months, although for most months the estimated values are quite close to the observed.

Figure 2 shows the reduction in the data spread from the fitted function for each month for the 30 withheld stations for Model 5 (see also table 2). The axes are the same for all months. As mentioned previously, the use of SELNOT to select the withheld stations ensures sufficient spread in the range of data values for each month. A random selection could result in a geographic bias in the withheld data.

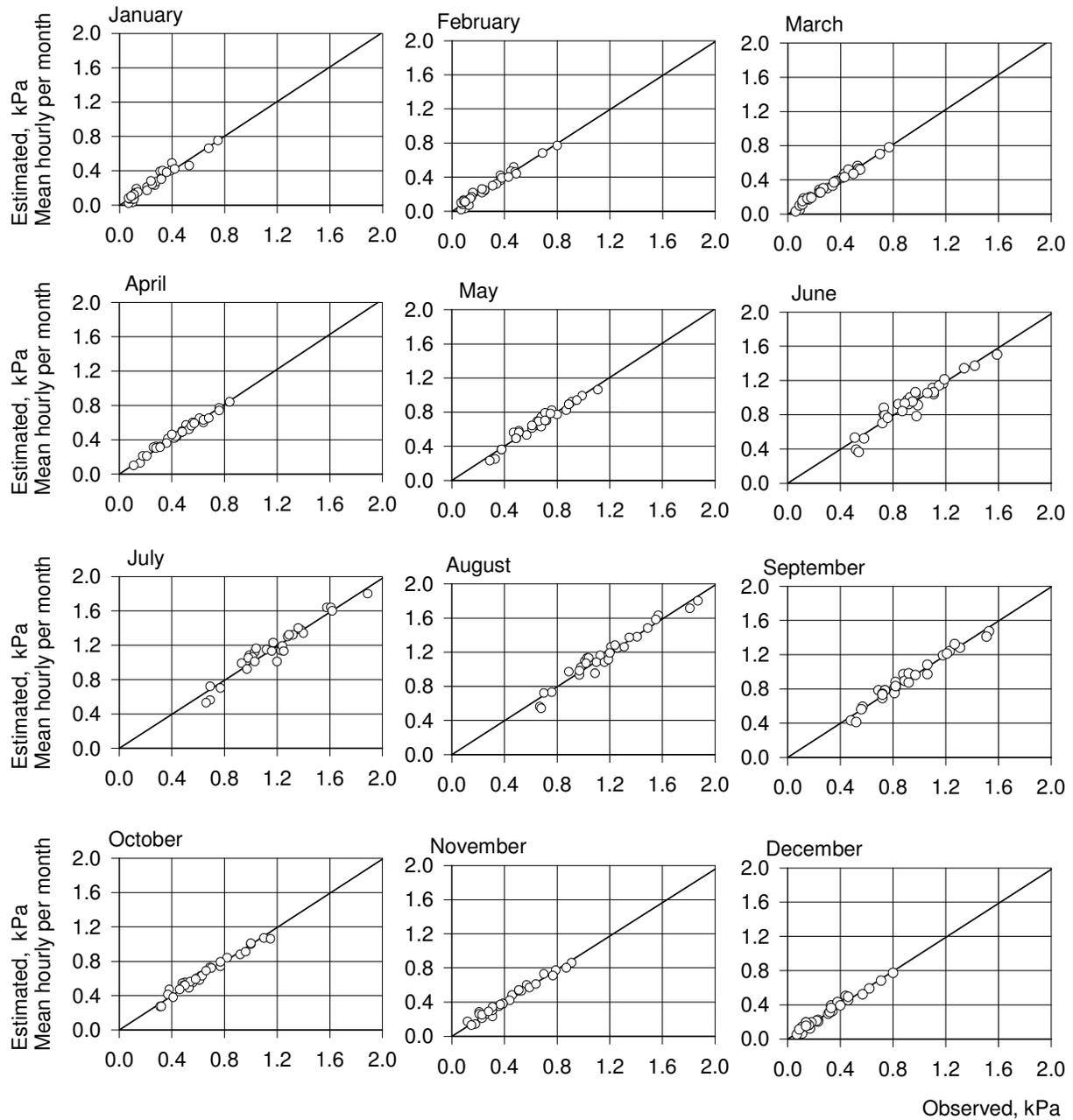


Figure 2. Reduction in data spread for 30 withheld stations (observed versus estimated) from Model 5 – square root transformation with 3 suspected stations removed

Month	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	Signal	RTGCV	RTMSE															
Jan	58.1	0.028	0.013	61.3	0.026	0.012	95.5	0.027	0.014	81.5	0.023	0.011	82.2	0.023	0.012	49.1	0.024	0.011
Feb	56.8	0.023	0.011	61.1	0.020	0.010	94.0	0.022	0.011	87.1	0.018	0.009	88.0	0.018	0.009	48.4	0.019	0.009
Mar	44.6	0.027	0.012	45.0	0.023	0.010	72.1	0.024	0.012	66.6	0.018	0.008	67.8	0.017	0.009	39.0	0.019	0.008
Apr	61.1	0.027	0.013	60.6	0.022	0.010	94.4	0.022	0.011	83.1	0.014	0.007	86.6	0.014	0.007	60.2	0.016	0.008
May	49.2	0.037	0.017	46.2	0.031	0.014	69.0	0.031	0.015	67.4	0.017	0.008	68.7	0.016	0.008	47.4	0.018	0.008
Jun	65.4	0.051	0.025	61.5	0.046	0.022	114.9	0.044	0.021	115.7	0.021	0.009	106.7	0.020	0.009	65.1	0.022	0.010
Jul	74.1	0.061	0.030	80.4	0.053	0.026	104.9	0.052	0.026	102.4	0.022	0.011	97.0	0.021	0.010	73.3	0.023	0.011
Aug	56.8	0.058	0.027	60.6	0.048	0.023	159.5	0.044	0.013	155.2	0.019	0.006	145.7	0.018	0.006	55.5	0.021	0.010
Sep	49.9	0.048	0.022	47.8	0.038	0.017	86.7	0.038	0.019	82.6	0.018	0.008	82.3	0.018	0.009	47.0	0.020	0.009
Oct	47.9	0.038	0.017	44.0	0.031	0.013	69.3	0.033	0.016	66.4	0.018	0.008	66.4	0.018	0.009	43.7	0.019	0.008
Nov	53.7	0.033	0.015	50.3	0.027	0.012	79.5	0.028	0.014	83.1	0.019	0.009	81.8	0.018	0.009	58.2	0.020	0.009
Dec	60.7	0.030	0.014	63.7	0.026	0.012	96.6	0.026	0.013	92.1	0.021	0.010	90.6	0.021	0.010	61.1	0.023	0.011

Model 1 – Bivariate model – function of longitude and latitude only (176 stations).

Model 2 – Covariate model – function of longitude latitude, elevation as a covariate.

Model 3 – Trivariate model – function of longitude, latitude and elevation (176 stations).

Model 4 – Trivariate model – function of longitude, latitude and elevation with square root transformation of dependent variable and covariance file (176 stations).

Model 5 – Trivariate model – function of longitude, latitude and elevation with three suspected stations removed (173 stations).

Model 6 – Bivariate model – vapour pressure values adjusted to sea level prior to surface fitting.

Table 4. ANUSPLIN diagnostics for six models using all data points.

Month	Station 1018610 Victoria Gonzales Hts., B.C.		Station 1030426 Amphitrite Point, B.C.		Station 8206240 Western Head, N.S.	
	Observed	Estimated	Observed	Estimated	Observed	Estimated
Jan	0.72	0.68	0.82	0.74	0.42	0.42
Feb	0.79	0.73	0.79	0.78	0.40	0.40
Mar	0.77	0.75	0.84	0.77	0.49	0.49
Apr	0.85	0.83	0.88	0.84	0.63	0.66
May	1.00	1.00	1.06	1.00	0.85	0.94
Jun	1.18	1.18	1.23	1.19	1.10	1.30
Jul	1.29	1.32	1.38	1.34	1.38	1.64
Aug	1.33	1.34	1.45	1.41	1.44	1.66
Sep	1.25	1.22	1.31	1.28	1.34	1.35
Oct	1.07	1.02	1.12	1.07	1.06	1.00
Nov	0.88	0.82	0.92	0.88	0.79	0.75
Dec	0.77	0.72	0.76	0.78	0.53	0.51

Table 5. Estimated and observed values for three stations withheld from Model 5 (n=173)

Figure 3 is a visualization resolved on an ~10km grid resolution of each of the monthly surfaces. Seasonal patterns are quite evident with more complex patterns and higher values arising in the summer months. Higher values are evident for the fall, winter and spring months along the southern east and west coasts. Late spring, summer and fall months have higher vapour pressure values in the interior parts of the country. The highest values occur in July in the corridor between southern Manitoba and southern Ontario extending into the Maritime Provinces, due to the combined influences of the Great Lakes, the Corn Belt and possibly the storm tracks up from the Gulf and/or the Caribbean.

Figure 4 provides the 95% model prediction error estimates for Model 5. The error estimates are generally highest in the summer months when values are higher and when greater spatial variation would be expected.

## 5 CONCLUDING COMMENTS

Several thin plate spline models of vapour pressure were developed for 1961-1990 monthly means calculated at 176 locations distributed across Canada. All models had reasonable signals with errors (estimated from 30 withheld stations) ranging from as low as 6% to at worst 25% for January for one of the models. The model selected for mapping had errors ranging from just 5.5% in September to 15.2% in January. The examples show that the diagnostic procedures can detect subtle information about the nature of the spatial variation of the data. Three stations were removed from the final model that were short records and potentially anomalous. Applying the square root transformation to the data appears to have played a significant role in reducing interpolation error over all months. Hutchinson

(1998a) found the square root transformation to provide a similar improvement in model accuracy when analyzing precipitation data. The final model that was mapped included all available stations except these three stations. The 95% Bayesian standard error estimates were also mapped and should provide useful insights to potential users.

Further applications of thin plate splines to Canadian climate data are underway including research on the robustness of such models at different time steps and to other variables.

## 6 ACKNOWLEDGEMENTS

Thanks to Kathy Campbell for assistance in the preparation of this manuscript.

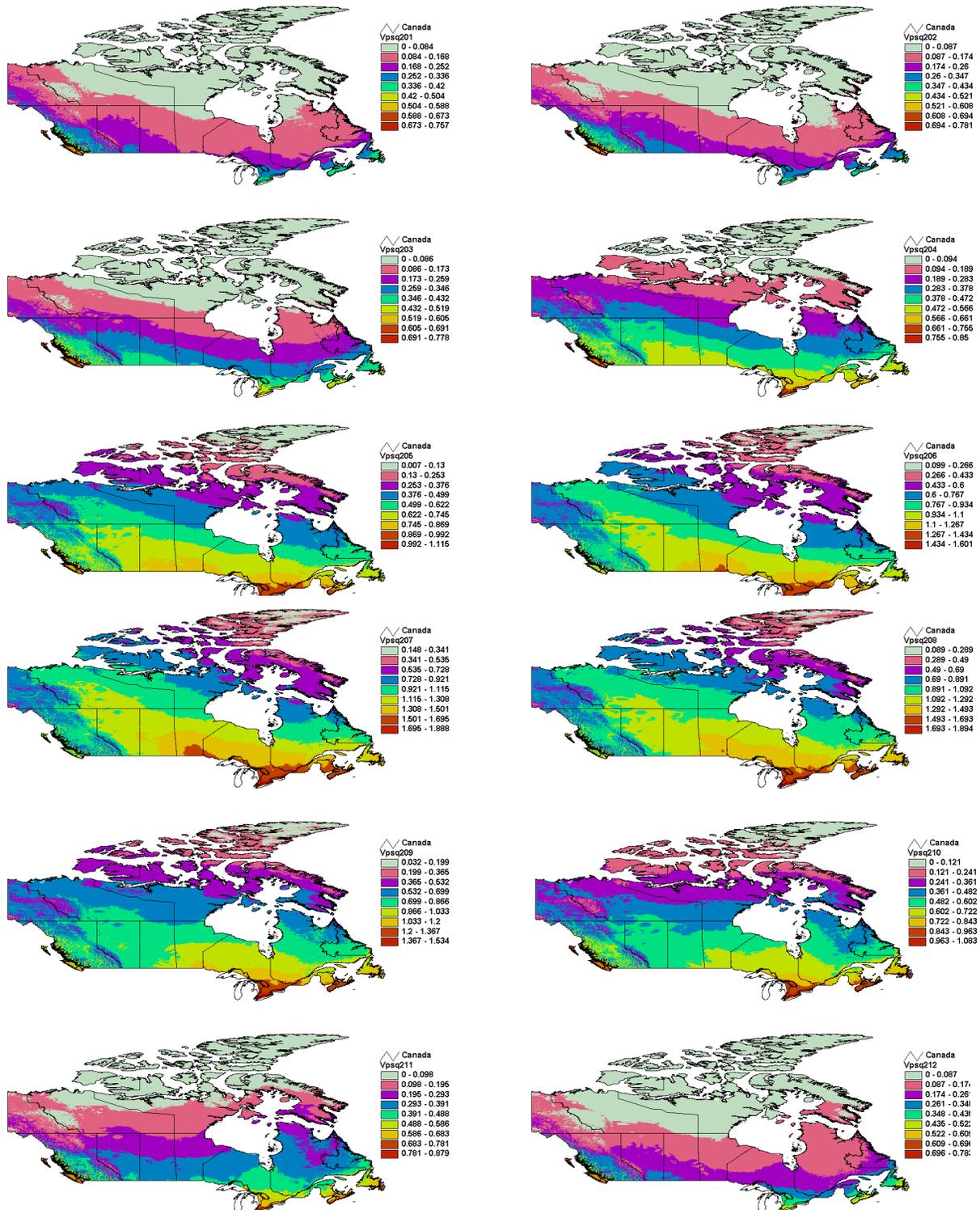


Figure 3. Monthly mean vapour pressure maps from Model 5 (see vapour pressure link at: [http://www.glf.cfs.nrcan.gc.ca/landscape/climate\\_models\\_e.html](http://www.glf.cfs.nrcan.gc.ca/landscape/climate_models_e.html) for higher resolution maps)

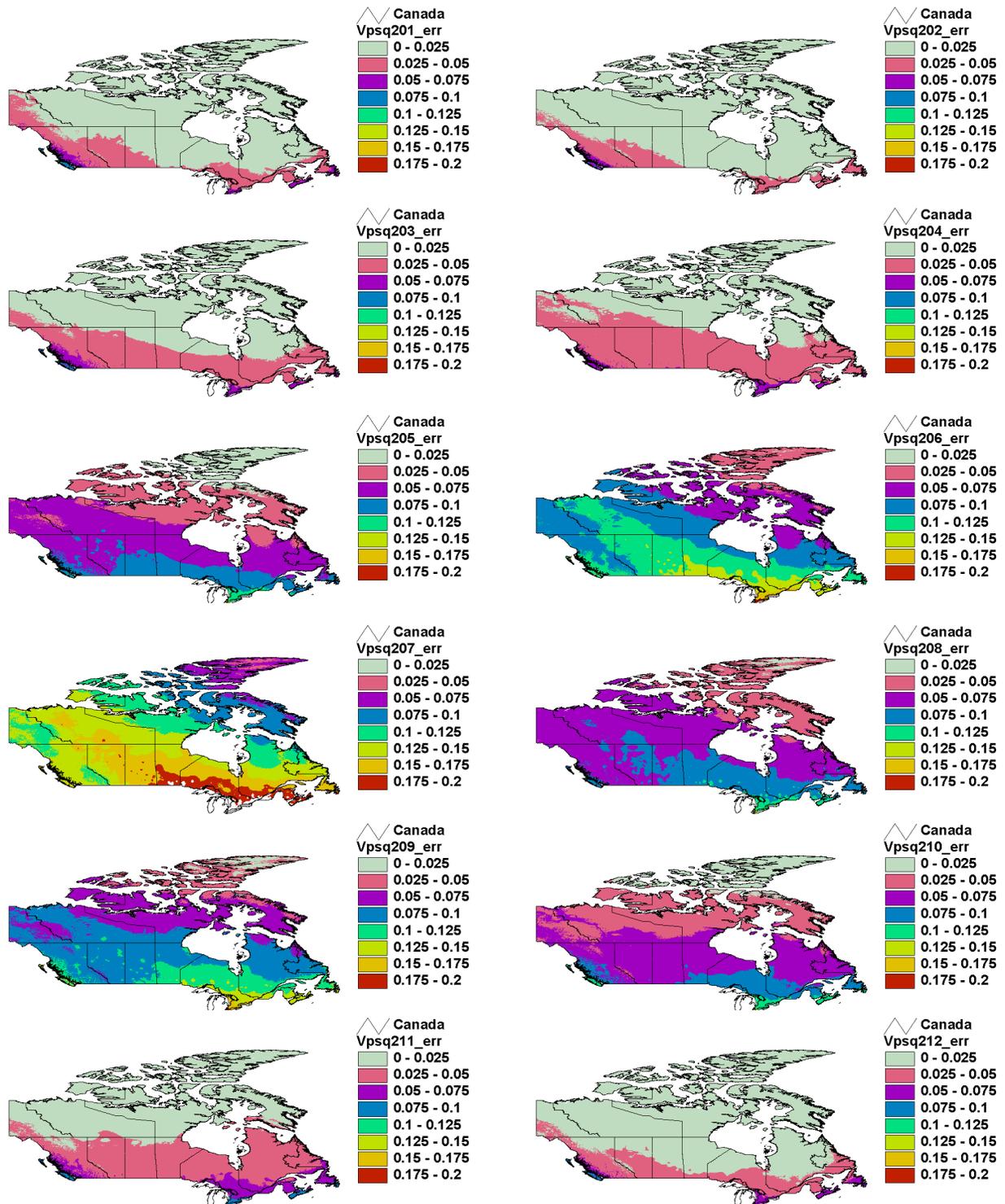


Figure 4. Model 5, 95% prediction error estimates.

## 7 REFERENCES

- Campbell, G.S. 1977. An Introduction to Environmental Biophysics. Springer-Verlag, New York, New York, USA.
- Castellvi, F., Perez, P.J., Villar, J.M., Rosell, J. I. 1996. Analysis of methods for estimating vapour pressure deficits and relative humidity. *Agricultural and Forest Meteorology* **82**, 29-45
- Castellvi, F., Perez, P.J., Stockle, C.O., Ibanez, M. 1997. Methods of estimating vapour pressure deficit at a regional scale depending on data availability. *Agricultural and Forest Meteorology* **87**, 243-252
- Goff, J.A., Gratch, S. 1946. Low-pressure properties of water from  $-160^{\circ}$  to  $212^{\circ}$  F. *Trans. Amer. Soc. Heat. Vent. Eng.*, **52**, 95-121.
- Hutchinson, M. F., 1995. Interpolating mean rainfall using thin plate smoothing splines.. *International Journal of GIS*, **9**, 305-403.
- Hutchinson, M. F., 1998a. Interpolation of rainfall data with thin plate smoothing splines: I two dimensional smoothing of data with short range correlation. *Journal of Geographic Information and Decision Analysis* **2**(2), 152-167.
- Hutchinson, M. F., 1998b. Interpolation of rainfall data with thin plate smoothing splines: II analysis of topographic dependence. *J. Geographic Information and Decision Analysis* **2**(2), 168-185.  
<http://publish.uwo.ca/~imalczew/gida>
- Hutchinson, M. F., Gessler P.E., 1994. Splines—more than just a smooth interpolator. *Geoderma* **62**, 45-67.
- Linacre, E.T., Hobbs J.E., 1977. The Australian Climatic Environment. Jacaranda Wiley.
- List, R.J., 1949. Smithsonian Meteorological Table, 6<sup>th</sup> ed. Smithsonian Institution Press, 350 pp.
- Lowe, P.R., 1977. An approximating polynomial for the computation of saturation vapor pressure. *J. Appl. Meteor.*, **16**, 100-103.
- Murray, F.W., 1967. On the computation of saturation vapour pressure. *J. Appl. Meteor.*, **6**, 203-204.
- New, M., Lister, D., Hulme, M., Makin, I., 2002. A high-resolution data set of surface climate over global land areas. *Climate Research*, **21**(1), 1-25.
- Price, D.T., McKenney, D.W., Nalder, I.A., Hutchinson, M.F., Kestevan, J.L. 2000. A comparison of statistical and thin-plate spline methods for spatial interpolation of Canadian monthly mean climate data. *Agricultural and Forest Meteorology*, **101**, 81–94.
- Rasmussen, L.A., 1978. On the computation of saturation vapor pressure. *J. Appl. Meteor.*, **17**, 1564-1565.
- Rasmusson, M.E. July 1967. Atmospheric water vapour transport and the water balance of North America. *Monthly Weather Review* **95**(7), 403-425.
- Richards, J.M., 1971. A simple expression for the saturation vapour pressure of water in the range of  $-50^{\circ}$  C to  $140^{\circ}$  C. *Brit. J. Appl. Phys.*, **4**, L15-L18.
- Sargent, G.P., 1980. Computation of vapor pressure, dew point and relative humidity from dry- and wet-bulb temperatures. *Meteor. Mag.*, **109**, 238-224.
- Stull, R.B. 1988. An Introduction to Boundary Layer Meteorology. Kluwer. Dordrecht, The Netherlands, 665 pp.
- Tetens, O. 1930. Über einige meteorologische Begriffe. *Zeitschrift Geophysik*, **6**, 297-309.
- Weast, R. C., 1973: 1973 Handbook of Chemistry and Physics. CRC Press, 2350 pp.