Synthesis of incoming shortwave radiation for hydrological simulation

Kevin Shook and John Pomeroy

ABSTRACT

Over the past decade, measurements of incoming shortwave radiation and of its proxy, daily bright sunshine hours, have become scarce in western Canada and elsewhere. As these data are critical for computing net radiation for snowmelt, evaporation, soil thaw and other components of the hydrological cycle, other means to estimate incoming shortwave radiation are needed. National Centers for Environmental Prediction and North American Regional Reanalysis atmospheric model reanalysis estimates of daily incoming shortwave radiation (Q_{si}), as well as the results of simplified semi-empirical calculations, were compared with measurements to determine their usefulness for hydrological calculations. It was found that all of the daily estimates show considerable bias and scatter compared with measurements. The best estimated values were produced by the semi-empirical Annandale method, particularly for simulations over the period of spring snowmelt. As many models require hourly radiation data, a method is presented for rescaling simulated daily Q_{si} data to the hourly values required.

Key words | energy balance, hydrological modelling, reanalysis products, snowmelt, solar radiation

Kevin Shook (corresponding author)
John Pomeroy
Centre for Hydrology,
University of Saskatchewan
E-mail: *kevin.shook@usask.ca*

INTRODUCTION

Estimates of net radiation are required to calculate coupled mass and energy balance components of the hydrological cycle such as snowmelt, sublimation, evapotranspiration, and soil thaw. The lack of simple methods to estimate net radiation has encouraged empirical methods, such as temperature-index snowmelt and evaporation models, to persist in operational use long after the development of energy balance methods. Empirical approaches have large uncertainties when considering energy inputs to hydrological systems, and may not be valid under conditions of changing climate or land use. The retention of empirical methods has contributed to a divide between hydrological science, which commonly employs physically based coupled mass and energy calculations, and hydrological practice which employs empirical calculations. This methodological divide has slowed the adoption of recent physically based calculations to problems such as predicting streamflow in ungauged basins.

As net radiation is seldom measured and is difficult to extrapolate or to scale directly from a point to a basin, incoming shortwave radiation, Q_{si}, is commonly employed to calculate net radiation using surface albedo and longwave radiation estimation procedures. The magnitude of Q_{si} is highly dynamic, varying hourly, daily and seasonally, and provides much of the radiative energy incident to the Earth's surface. For many land cover classes, albedo can be determined from point measurements, estimated from tables in many textbooks (e.g. Oke 1987) or modelled (e.g. Gray & Landine 1987) and used to estimate net shortwave radiation from Q_{si}. As Q_{si} provides information on atmospheric transmittance, empirical methods exist to estimate net longwave radiation from shortwave radiation (e.g. Granger & Gray 1990) or to estimate incoming longwave radiation from shortwave transmittance (Sicart *et al.* 2006). Hours of direct sunshine, measured with commonly employed sunshine recorders, were compared with maximum possible hours of direct sunshine and used to estimate

doi: 10.2166/nh.2011.074

incoming shortwave radiation using empirical formulae tested and extended by Penman (1948) and Brutsaert (1982) amongst others for many years. Unfortunately, the recent decline of the radiation monitoring network in Canada has resulted in very few sunshine recorders being in use and the only active shortwave radiation station operated by Environment Canada in the Prairie Provinces is the Bratt's Lake Observatory near Regina, Saskatchewan, which is part of the Baseline Surface Radiation Network (www.bsrn.awi.de). The location of the Bratt's Lake Station is shown in Figure 1.

Since approximately 2005, the Environment Canada solar radiation data have been supplemented by a network of monitoring stations operated by Alberta Agriculture and Rural Development. Although the density of these stations is impressive, they are only located in Alberta, leaving Saskatchewan and Manitoba without coverage, and as the sites were only established relatively recently, they cannot be used for analyses of historical events. These sites are only located within the agricultural regions of Alberta, omitting the mountains and boreal forests. The locations of the Alberta sites having solar radiation data are also shown in Figure 1.

The Canadian prairies are adjacent to the northern U.S. states, which have very high densities of solar radiation stations. For example, the states of Idaho, Montana, North Dakota, and Minnesota have, respectively, 21, 20, 11 and 76 stations measuring solar radiation (National Renewable Energy Laboratory 2007). The locations of the U.S. stations close to the border are plotted in Figure 1.

As with all point measurements, the representativeness of the US solar radiation data decline over distance. Bland & Clayton (1994) found that in Wisconsin, for radiometers spaced 270 km apart, the error in daily global solar radiation was 4.6 MJ/m^2 (53 W/m^2) for 80% of the area on 90% of

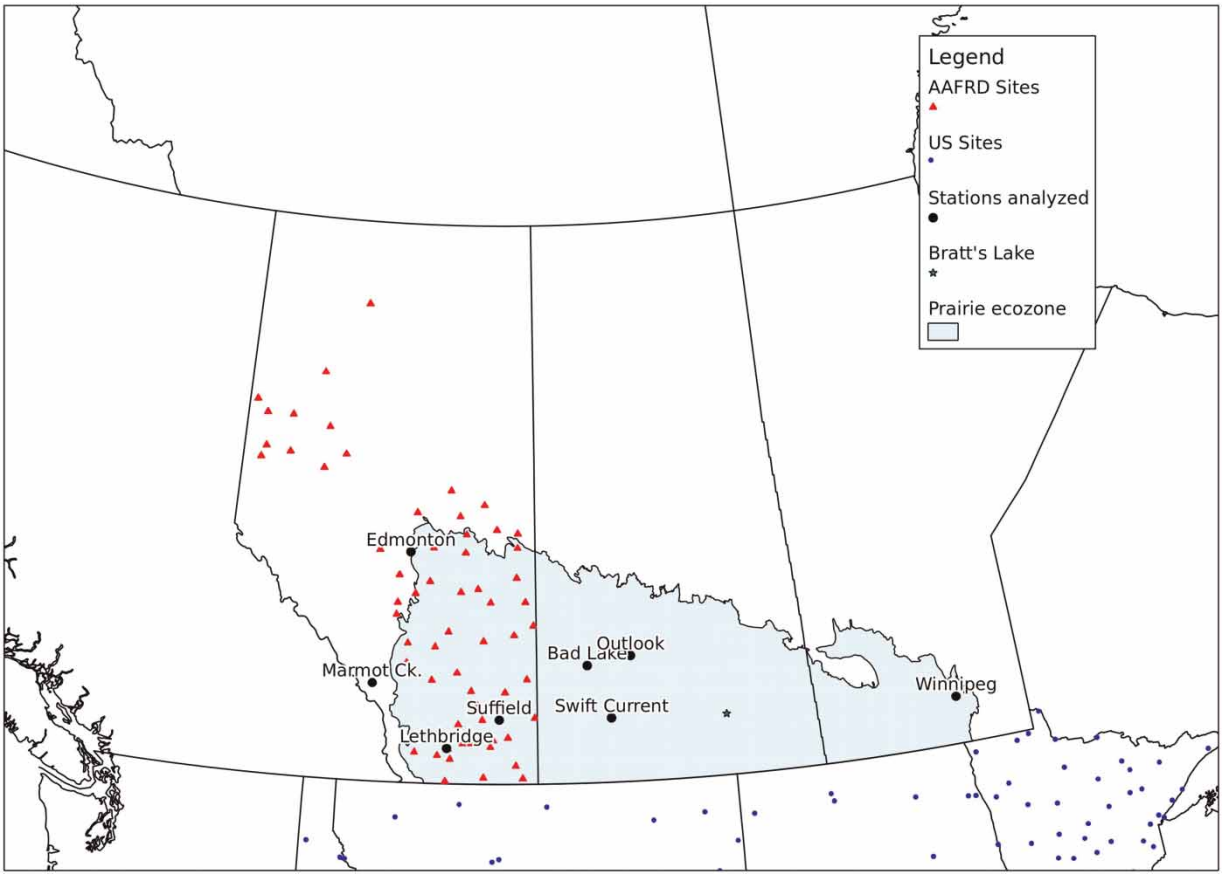

Figure 1 | Sites having solar radiation measurements in Canadian prairies and northern United States.

the days analyzed. Further studies are required to determine the regions in Canada where the US station radiation data will have smaller errors than alternative methods of estimating Q_{si}.

Given the lack of directly measured values, the importance of the data, and the difficulty and uncertainty of using remotely sensed atmospheric transmittance data, it is necessary to find other ways of determining shortwave radiation. The overall objective of this study is to determine the suitability of estimated Q_{si} for hydrological modelling in western Canada. Estimates of daily incoming shortwave radiation from atmospheric reanalysis model results and semi-empirical calculations were compared with measurements. Temporal downscaling methods were then employed to rescale estimated daily radiation to the shorter time scales measured by instruments and used by hydrological models.

SOURCES OF ESTIMATED DAILY Q_{si}

All of the data sources evaluated were capable of producing estimates of daily Q_{si}, designated as Q_{siD}. Although hourly estimates of Q_{si} (i.e. Q_{siH}) are often required for hydrological modelling, these values are often downscaled from daily estimates.

Reanalysis data sets

Reanalysis data sets are produced through (a) assimilating measurements of atmospheric variables into an atmospheric model, (b) running the model for a single time step, and (c) outputting the computed values (Mesinger *et al*. 2006). Thus the data are reanalysed from measured data by reproducing the physics of the atmosphere. The North American Regional Reanalysis (NARR) and National Centers for Environmental Prediction (NCEP) datasets described below use the Eta mesoscale model developed by the U.S. National Meteorological Center (Black 1994). The Eta model does not use measured incoming shortwave radiation data as an input, as the shortwave radiation at the surface is entirely calculated (Shafran *et al*. 2004). Therefore the NARR and NCEP datasets are not affected by the scarcity of measured Q_{si} data in Canada, although they are affected by the coarse spacing of other Canadian meteorological data.

NARR

NARR is a high-resolution (32 km) high-frequency (three-hour) dataset of atmospheric variables, produced by the U.S. NCEP (Mesinger *et al*. 2006). The data originally spanned the period 1979–2003, but have been available on a near-real-time basis since 2003. The data are freely available at www.emc.ncep.noaa.gov/mmb/rreanl. All NARR variables are available as either three-hour or daily values.

NCEP

The NCEP reanalysis project is similar to NARR. However, NCEP provides data over longer periods of time (1948–present) over a larger area (worldwide) at coarser temporal (six-hour) and spatial (approximately 210 km) resolutions. As with the NARR outputs, the NCEP variables are also available as daily values. NCEP reanalysis data are provided free of charge at the U.S. National Oceanographic and Atmospheric Administration web site at http://www.cdc.noaa.gov/. As with the NARR data, daily NCEP values are also available. Both NARR and NCEP data sets are based on universal time.

For ease of data handling, it is much more convenient to use the NARR and NCEP daily values rather than the sub-daily values. Because both NARR and NCEP datasets are based on universal time the daily values will span the daily divisions of local time. However, the magnitude of the daily shortwave radiation flux variable 'dswrf' between 03:00–06:00 is typically very small. Therefore, the mean daily fluxes will be very similar, whether calculated from 00:00–24:00 (NARR daily values) or from 06:00–06:00 (shifted to be closer to local time), as shown in Figure 2. All subsequent analyses of NARR and NCEP data were performed using daily values, as these are the most convenient for input to existing energy balance calculations that rely on daily radiation (e.g. Gray & Landine 1987; Granger & Gray 1990).

Both NARR and NCEP are gridded data sets which use average conditions inside each grid cell. The objective of this study was to examine alternatives to the use of Q_{si} values determined from individual radiometers. Therefore the

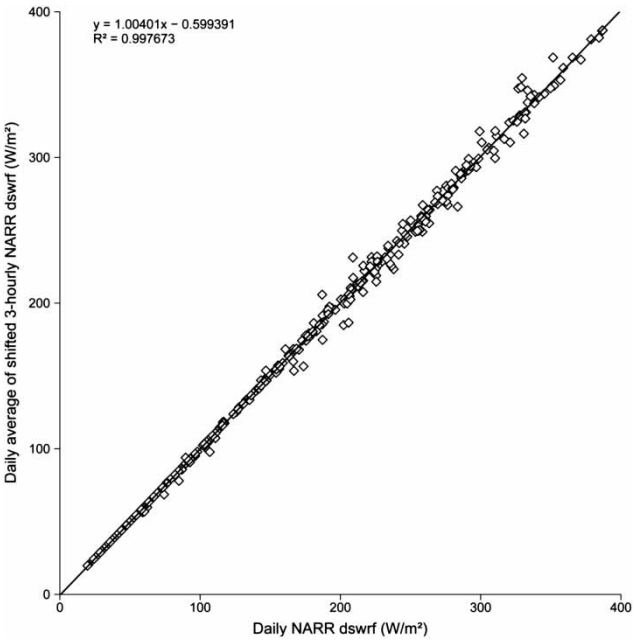

Figure 2 | Daily downward shortwave radiation flux (dswrf) calculated from three-hour data shifted backwards six hours versus daily values. Data for Edmonton, AB (2000).

NARR and NCEP data were evaluated against readings from radiometers.

Semi-empirical calculations

Variation in the magnitude of incoming shortwave radiation at a given location is due to two factors. Firstly, the shortwave radiation at the top of the Earth's atmosphere (extraterrestrial shortwave) varies due to the Sun's position relative to the Earth (which is a function of the latitude, date, and time of day). Secondly, there is temporal and spatial variation in the effects of absorption and scattering of radiation by the atmosphere on shortwave transmittance.

It is possible to estimate Q_{si} directly, using calculations which incorporate theoretical components (to predict the Sun's position relative to a given location) and empirical relationships (which estimate the effects of the Earth's atmosphere). Q_{si} can be expressed as the sum of direct beam (Q_{dr}) and diffuse sky (Q_{df}) radiation as

$$Q_{si} = Q_{dr} + Q_{df} \tag{1}$$

Garnier & Ohmura (1968) demonstrated that the daily clear-sky direct-beam incoming shortwave radiation (Q_{droD}) could be calculated for any location on Earth using the equation

$$Q_{droD} = \frac{I_0}{r^2} \int_{H_1}^{H_2} p^m \cos(X \wedge S) \mathrm{d}H \tag{2}$$

where I_0 = solar constant (1.35 kW/m^2), r = radius vector of Earth's orbit, p = atmospheric transmissivity along the zenith, m = optical air mass (distance through atmosphere/depth of atmosphere at the zenith), H_1 = time that Sun begins shining on surface, H_2 = time that Sun ends shining on surface, X = unit vector normal to the surface of the Earth at a given location, S = unit vector defining the Sun's position in the sky, and H = hour angle measured from solar noon.

The symbol $\wedge$ refers to the cross-product of the unit vectors.

Note that by using times other than local sunrise and sunset for H_1 and H_2, the total incoming radiation may be calculated for any time period.

The daily, clear-sky diffuse shortwave radiation (Q_{dfoD}) in MJ/(m^2 day) for the prairies can be estimated from the relationship (Granger & Gray 1990):

$$Q_{dfoD} = \frac{3.5(P/P_0)\cos(X \wedge S)}{p} + 0.45 \sin\left[\frac{(172 - d)\, 2\pi}{365}\right] \tag{3}$$

where

P = air pressure at the Earth's surface at location of calculation, P_0 = standard air pressure at sea level (same units as P), p = transmissivity of atmosphere, and d = day of year.

The value of P/P_0 can be estimated from

$$\frac{P}{P_0} = \left(\frac{288 - 0.065\ \mathrm{Alt}}{288}\right)^{5.256} \tag{4}$$

where Alt = site altitude in metres above sea level.

At solar noon,

$$\cos(X \wedge S) = \cos\delta \cos\varphi \tag{5}$$

where φ = latitude of given location, and δ = Sun's declination, which can be estimated from

$$\delta = 0.4093 \sin\left[\frac{(d - 81)\, 2\pi}{365}\right] \tag{6}$$

The transmissivity is affected by the mass of air along the Sun's zenith, its humidity and by the impurities that the air contains. Granger & Gray (1990) found that the seasonal variation in the value of p could be expressed as

$$p = 0.818 - 0.064 \sin\left[\frac{(d-90)2\pi}{365}\right] \quad (7)$$

which yields estimates of p between 0.75 and 0.88.

Equations (1)–(3) are time-consuming to solve, and have been simplified for horizontal surfaces (Walter *et al.* 2005) as

$$Q_{siD} = S_o \tau_D \quad (8)$$

where Q_{siD} = mean daily Q_{si}, τ_D = daily atmospheric transmittance, and S_o = extraterrestrial shortwave radiation incident to a horizontal plane at the top of the atmosphere, as given by

$$S_o = \frac{I_0}{\pi}[\arccos(-\tan(\delta)\tan(\varphi))\sin(\varphi)\sin(\delta) + \cos(\varphi)\cos(\delta)\sin(\arccos(\tan(\delta)\tan(\varphi)))] \quad (9)$$

Although the atmospheric transmittance is easily estimated for clear-sky conditions, the effects of clouds are more difficult to quantify. Granger & Gray (1990) give methods for estimating the direct-beam and diffuse radiation for cloudy days, but these methods require the fraction of each day experiencing bright sunshine calculated from the daily sunshine hours, which is no longer measured in most of Canada.

Ball *et al.* (2004) evaluated four semi-empirical models for determining Q_{si} at locations throughout North America which spanned 23° of latitude and 42° of longitude. While some of the models evaluated were based solely on air temperatures, the analyses also included the model of Thornton & Running (1999) which includes the effects of humidity. Ball *et al.* (2004) concluded that all of the models, despite greatly differing in their requirements of data and computational effort, produced similar results. Operationally, use of the Thornton and Running model is made difficult by its requirement for site-specific constants based on antecedent precipitation. Therefore, as the objective of this research is to determine the usefulness of simple methods for determining Q_{si}, the analyses of semi-empirical values was restricted to methods which only require measurements of air temperature.

CAMPBELL–BRISTOW–WALTER METHOD

Bristow & Campbell (1984) developed a simple relationship between τ_D and the difference between daily minimum and maximum air temperatures:

$$\tau_D = A[1 - \exp(-B\Delta T^C)] \quad (10)$$

where A, B, and C are constants, and ΔT is range of daily air temperatures (°C).

Walter *et al.* (2005) gave a simplified version of the Campbell–Bristow model using values for A, B, and C that were developed by other researchers: $A = 0.75$, $B = b/S_o30$, and $C = 2$. $S_o30 = S_o$ computed for the date 30 days prior to the date of simulation. It can be calculated using Equation (8).

Walter *et al.* (2005) used two simple equations for estimating the value of b:

$$b = 0.282\varphi^{-0.431} \text{ for summer} \quad (11)$$

and

$$b = 0.170\varphi^{-0.979} \text{ for winter} \quad (12)$$

where φ is the latitude of the location in radians. Summer is defined as extending from 90 days preceding the summer solstice to 90 days following. Winter is the rest of the year.

Annandale *et al.* (2001) developed a simple empirical method for estimating τ_D. Like the Campbell–Bristow and Walter methods, it is based on the daily range of air temperatures, but the Annandale model is much easier to calculate as it does not require the computation of S_o30. In addition to being complex to calculate as shown in Equation (9), S_o30 is temporally out of sync with the other calculations making it inconvenient to use in spreadsheets and script and macro languages having a fixed time step. The Annandale model, shown in Equation (13), explicitly incorporates the effects of altitude on transmittance and so

has promise for mountain applications.

$$\tau_D = k_{Rs}(1 + 2.7 \times 10^{-5} \text{Alt}\, \Delta T^{0.5}) \quad (13)$$

where k_{RS} = adjustment coefficient, 0.16 for interior locations, 0.19 for coastal regions, and Alt = site altitude (m).

REGRESSION ANALYSES

To determine the usefulness of simulated values as replacements for measurements of Q_{siD}, the simulated Q_{si} values were compared with measured values using simple linear regressions. All analyses were performed using the open-source statistical language R (R Development Core Team 2008).

Autocorrelation of residuals

All of the solar radiation data, whether measured, reanalysed, or calculated from semi-empirical relationships, are highly autocorrelated, due to the time dependency of both S_o and τ_D. Differences in the autocorrelation between the measured and calculated τ_D datasets may lead to the presence of autocorrelation in the residuals of regressions, which is generally considered to be undesirable because it leads to error in the regression coefficients, and underestimation of the standard error of the regression coefficients (Chatterjee & Hadi 2006). The presence of autocorrelation in the residuals of a regression may be diagnosed by the Durbin–Watson test (Durbin & Watson 1951) of the correlation coefficient (ρ) among adjacent residuals. The null hypothesis of the test (H_0: $\rho = 0$) is that the residuals are uncorrelated, the alternative hypothesis (H_1: $\rho > 0$) is that correlation exists. All of the regressions of simulated Q_{siD} datasets against the measured values using ordinary least-squares had residuals which showed autocorrelation at the 5% level.

The temporal correlation of the measured and simulated data sets was removed by randomly-ordering the pairs of values. When tested, the null hypothesis of the Durbin–Watson test was accepted for all of the re-ordered data sets. Despite this, there was no difference between the regression coefficients obtained using the original or the scattered values.

Study locations

Measured Q_{si} datasets for locations in Alberta, Saskatchewan, and Manitoba are listed in Table 1, and plotted in Figure 1. The ecozones are as specified by Marshall *et al.* (1996). Note that the prairie sites and the Edmonton site are considered to be plains sites.

Although measured Q_{si} data preceding 1979 existed at some of the sites, these values were not included in the analyses as they precede the NARR values. The measured data for all sites other than Marmot Creek were obtained from the Data Access Integration (DAI) web portal created by the Global Environmental and Climate Change Centre (GEC3), the Ouranos Consortium on regional climate change and impacts, the Adaptation and Impacts Research Division (AIRD) of Environment Canada, and the Drought Research Initiative (DRI, www.drinetwork.ca). The Marmot Creek data were measured by the Centre for Hydrology, University of Saskatchewan using a Kipp and Zonen CNR1 radiometer.

Table 1 | Locations of study datasets

Station	Latitude (°N)	Longitude (°W)	Elevation (m)	Ecozone	Years
Edmonton, Alberta	53.55	114.11	723	Boreal plain	1979–2005
Marmot Creek, Alberta	50.94	115.14	1,436	Montane Cordillera	2005–2006
Suffield, Alberta	50.27	111.18	770	Prairie	1979–1990
Lethbridge, Alberta	49.70	112.78	929	Prairie	1990–1997
Outlook, Saskatchewan	51.48	107.05	541	Prairie	1988–1995
Bad Lake, Saskatchewan	51.32	108.42	637	Prairie	1979–1986
Swift Current, Saskatchewan	50.27	107.73	825	Prairie	1979–2000
Winnipeg, Manitoba	49.92	97.23	239	Prairie	1979–2000

The air temperatures required for the Annandale and CBW calculations were measured at Marmot Creek by the Centre for Hydrology using shielded Vaisala HMP45C hygrothermometers. For all other sites, daily minimum and maximum air temperatures were obtained from the Canadian National Climate Data and Information Archive at www.climate.weatheroffice.ec.gc.ca, and follow WMO standard protocols.

Regression results

Plains sites

To simplify comparisons, the coefficients of all of the regressions for the plains sites (i.e. all sites other than Marmot Creek) are averaged and tabulated by calculation method in Table 2. As an illustration of typical results, Figure 3 plots all four simulation methods against measured values, with least-squares fitted lines, for Lethbridge.

Evidently, the Annandale method produces regressions having the smallest standard error, followed by the NARR data, the CBW method, and finally the NCEP data.

Mountain site

The regression coefficients for the mountain site (Marmot Creek) are listed in Table 3. In all cases, the standard error of the residuals is greater, and the value of R^2 is smaller, than the equivalent mean value listed in Table 2. As with the plains sites, the Annandale method produced better results than did the CBW method, but the NARR regressions were the best, having the greatest magnitude of R^2, and the smallest standard errors of the residuals.

Table 2 | Mean values of regression coefficients for plains sites

Method	Intercept (W/m^2)	Slope	R^2	Std. error of residual (W/m^2)
Annandale	17.7	0.94	0.84	40.4
CBW	13.9	0.91	0.81	43.0
NARR	37.5	0.96	0.83	42.6
NCEP	77.5	0.76	0.60	59.7
Mean	36.6	0.89	0.77	46.4

The poorer results of all of the mountain site regressions are not due to terrestrial blocking or reflection as the sky-view factor at this location is approximately 0.99. The relatively poor performance of the Annandale and CBW models may be due to their inability to account for the effects of complex mountain weather systems on transmittance. The Marmot site is located in a valley, where the weather is dominated by mesoscale advection and cold air drainage at night (Helgason 2009). Therefore, air temperatures at ground level are often not representative of the state of the local atmosphere, and NARR's physically based simulations would be expected to better estimate atmospheric transmittance.

SEASONAL REGRESSIONS

Solar radiation is primarily of interest to hydrology during the seasons of snowmelt and evapotranspiration. On the Canadian prairies, snowmelt typically occurs during March and/or April, and almost all of the energy required for the melt process is supplied by solar radiation. Evaporation from small prairie wetlands occurs during the open-water season, which typically runs from May to October. Transpiration from plants is restricted to their growing season. The Canadian prairies are dominated by the agriculture of grasses and small grains, whose growing season generally runs from May to August, with the peak evapotranspiration rates occurring during July (Armstrong *et al*. 2008).

Regressions were calculated for the periods of March–April (spring) and May–August (summer) for all of the plains data sets. The coefficients of the seasonal regressions are listed in Table 4. Compared with the coefficients of the regressions of the complete data sets listed in Table 2, the values of R^2 for the spring and summer regressions are much smaller, the slopes and intercepts are greatly different, and the standard errors of the residuals are only slightly changed. These seemingly contradictory behaviours are illustrated in Figure 4, which plots the Annandale calculated Q_{siD} against the measured values for Winnipeg for all three time periods. The spring and summer values occupy distinct regions of the plot, and therefore have very distinct regressions. Each seasons' points are also concentrated in a particular portion of the plot, resulting in reduced values of R^2.

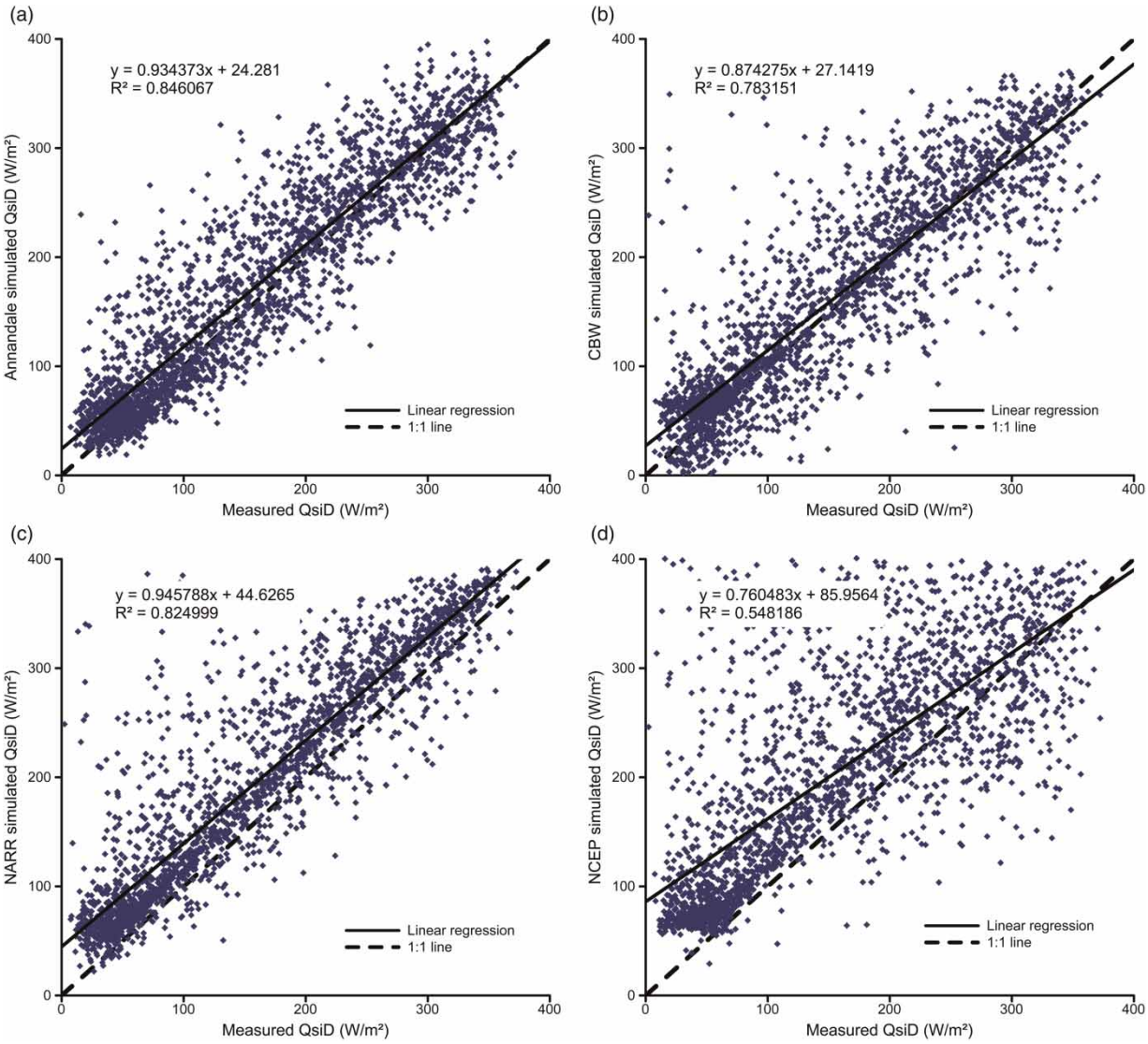

Figure 3 | Simulated Q_{siD} by (a) Annandale, (b) CBW, (c) NARR, and (d) NCEP versus measured for Lethbridge.

Table 3 | Regression constants for mountain site (Marmot Creek)

				Confidence levels				
				Intercept		Slope		
Method	**Intercept (W/m²)**	**Slope**	**R^2**	**2.5% (W/m²)**	**97.5% (W/m²)**	**2.5%**	**97.5%**	**Std. error of residual (W/m²)**
Annandale	22.2	0.90	0.76	14.0	30.3	0.85	0.94	48.0
CBW	35.7	0.85	0.78	25.1	46.3	0.91	0.62	60.1
NARR	4.8	0.94	0.80	−5.2	14.7	0.90	0.99	46.0
NCEP	0.9	0.71	0.57	−11.8	13.5	0.65	0.77	58.7
Mean	15.9	0.85						53.2

Table 4 | Mean coefficients of seasonal linear regressions of measured against simulated Q_{siD}

	Method	Intercept (W/m^2)	Slope	R^2	Std. error of residual (W/m^2)
Spring	Annandale	50.8	0.70	0.55	40.0
	CBW	36.3	0.82	0.52	49.5
	NARR	84.0	0.68	0.52	40.7
	NCEP	179.5	0.33	0.19	43.8
	Mean	87.6	0.60	0.40	43.5
Summer	Annandale	147.9	0.51	0.53	35.4
	CBW	61.4	0.74	0.53	53.0
	NARR	156.6	0.55	0.45	45.0
	NCEP	276.6	0.04	0.01	55.2
	Mean	160.6	0.50	0.40	47.2

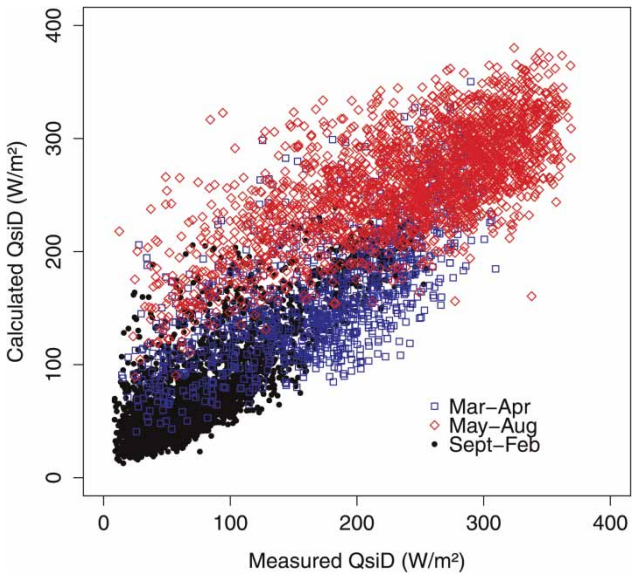

Figure 4 | Annandale calculated Q_{siD} vs measured Q_{siD} (W/m^2) for Winnipeg during snowmelt (Mar–Apr), growing (May–Aug) and dormant (Sept–Feb) seasons.

Total error during snowmelt period

All of the analyses presented above are based on individual daily values. On the Canadian prairies, where snowmelt generally occurs over a short period, generally in the order of several days (Granger & Gray 1990), the error in the total solar energy modelled will be smaller than for a single day, but will not decrease to the same extent as over the relatively longer period of evapotranspiration.

The RMS error of the total solar energy was determined for each of the estimation methods for melt periods of between 1 and 30 days over the melt seasons of each year. For each year, errors were computed for all possible values of each melt period and were converted to fractions of the total solar radiation energy. The average RMS errors for each estimation method are plotted against the number of days simulated in Figure 5. On average, the Annandale method yields the smallest RMS error when summed over any number of days. Surprisingly, the CBW data had the second-smallest RMS errors, despite their regressions having poorer values of R^2 than did the NARR regressions. Cumulative errors in the Annandale and CBW atmospheric transmittance models declined with a longer simulation period whilst those from the reanalysis datasets did not, suggesting the transmittance models might be more suitable than reanalysis data for hydrological applications.

The discrepancy between the error of the regression and that of the cumulative values is explained by plots of the kernel densities (which are analogous to histograms) of the measured and simulated values for Edmonton in Figure 6(a). The Annandale and CBW data both under- and over estimate the measured distribution, while the NARR and NCEP data consistently overestimate the measured values. The same distributions are shown in Figure 6(b) which plots the quantiles of the simulated Q_{siD} values against the corresponding quantiles of the measured Q_{siD}.

The linearity of the NARR plot in Figure 6(b), is caused by the good agreement of the shapes of the NARR and measured Q_{siD} distributions, which is also visible in the

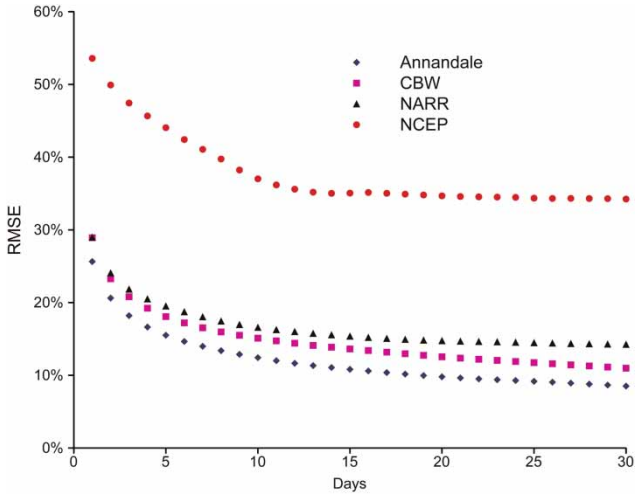

Figure 5 | Mean total RMS error (percent) versus number of days for estimation of Q_{siD} during March and April.

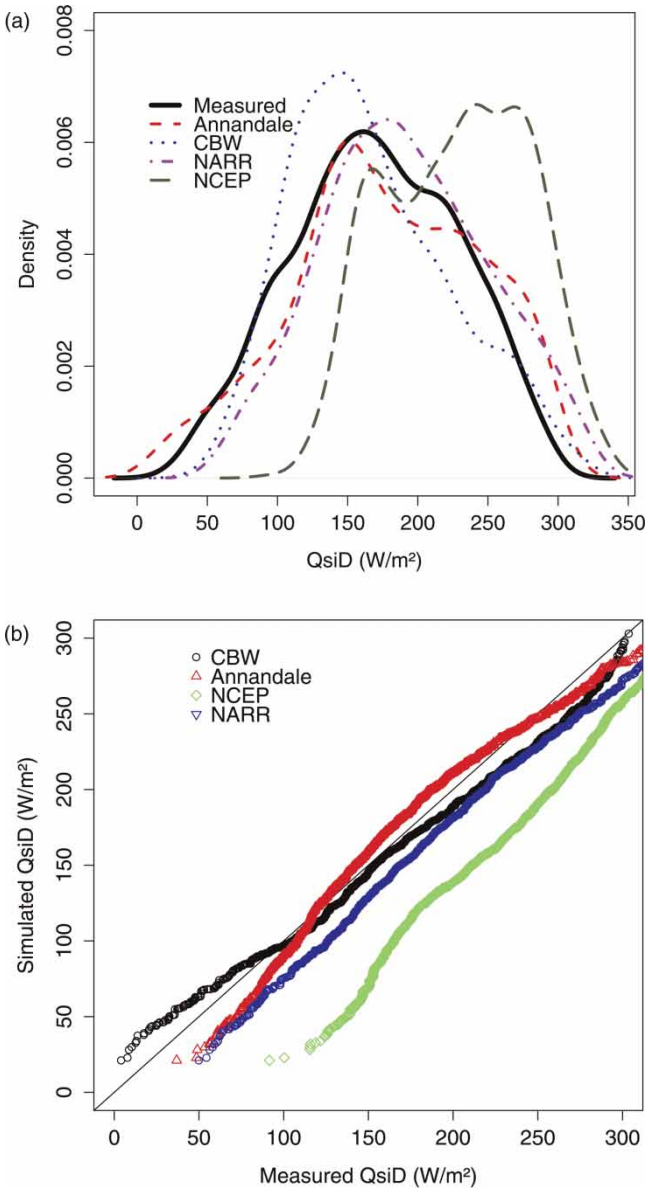

Figure 6 | Plots describing frequency distributions of Edmonton measured and simulated Q_{siD}, (a) Kernel density plots of Edmonton spring measured, Annandale, CBW, NARR, and NCEP Q_{siD}, (b) Quantile–quantile plots of Edmonton spring Annandale, CBW, NARR, and NCEP Q_{siD}, against measured quantiles. The line is 1:1.

quantile plots of Figure 6(a). Therefore, it is possible that the offset of the NARR data could be corrected by a simple linear transformation. However, the mean differences between the NARR and measured data are highly variable, as shown by the quantile–quantile plots of ten sets of spring NARR values for Swift Current in Figure 7. The Swift Current NARR data are generally smaller than the measured values, while the Edmonton NARR data were

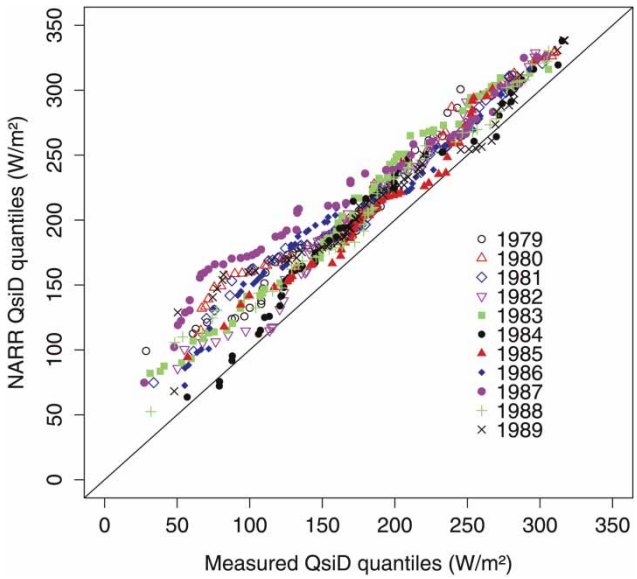

Figure 7 | Quantiles of simulated versus measured Q_{siD} for Swift Current, by year.

larger than the corresponding measured values. Over all datasets, the mean difference between NARR and measured values is 24 W/m^2, but as the standard deviation of the differences is 21 W/m^2 it is difficult in practice to adjust values. Any adjustment is very likely to make the agreement between NARR and measured values poorer.

INTERPOLATING HOURLY VALUES OF Q_{si}

Because the incoming shortwave radiation flux varies continuously from sunrise to sunset, most recent physically based hydrological calculations require sub-daily values of incoming shortwave radiation to reproduce temporal variability of melt, sublimation and evaporation throughout each day. Where multiple-hour data are available, as in the case of the NARR and NCEP datasets, the values may be linearly interpolated. The Annanadale and CBW methods, which only produce daily Q_{si} values, can be interpolated using calculated values for the hourly clear-sky incoming shortwave radiation, Q_{sioH}.

From Equation (1), Q_{sioH} can be expressed as the sum of hourly clear-sky direct beam (Q_{droH}) and hourly clear-sky diffuse (Q_{dfoH}) radiation:

$$Q_{sioH} = Q_{droH} + Q_{dfoH} \quad (14)$$

The value of Q_{droH} can be calculated by solving Equation (2) over the limits $h-1$ to h, where h is the hour of the day.

None of the previous equations provides a method for estimating the hourly clear-sky diffuse (Q_{dfoH}) radiation. However, as Equations (2) and (3) indicate that Q_{dfoD} and Q_{droD}, and by extension, Q_{droH}, vary according to $\cos(X \wedge S)$, it is assumed that Q_{dfoH} can be estimated by proportionality from

$$Q_{\mathrm{dfoH}} = Q_{\mathrm{dfoD}}\left(\frac{Q_{\mathrm{droH}}}{Q_{\mathrm{droD}}}\right) \tag{15}$$

Having computed Q_{sioH}, the value of Q_{siH} can be computed from

$$Q_{\mathrm{siH}} = Q_{\mathrm{sioH}} R_{Q\mathrm{sioH}} \tag{16}$$

where $R_{Q\mathrm{sioH}}$ can be defined as

$$R_{Q\mathrm{sioH}} = \frac{Q_{\mathrm{siH}}}{Q_{\mathrm{sioH}}} \tag{17}$$

The value of $R_{Q\mathrm{sioH}}$ is unknowable from daily data. Because the effects of clouds vary over time, the value of $R_{Q\mathrm{sioH}}$ will vary from hour to hour during any given day. This effect is assumed to be nearly random, and therefore on any given day,

$$\overline{R_{Q\mathrm{sioH}}} \approx \frac{Q_{\mathrm{siD}}}{Q_{\mathrm{sioD}}} \tag{18}$$

where the bar denotes averaging over a day. Ignoring the hour-to-hour variation in $R_{Q\mathrm{sioH}}$ allows Equation (15) to be rewritten as

$$Q_{\mathrm{siH}} \approx \frac{Q_{\mathrm{sioH}} Q_{\mathrm{siD}}}{Q_{\mathrm{sioD}}} \tag{19}$$

The ratio $R_{Q\mathrm{sioHD}}$ can be defined as

$$R_{Q\mathrm{sioHD}} = \frac{Q_{\mathrm{sioH}}}{Q_{\mathrm{sioD}}} \tag{20}$$

By fitting an empirical function (R) of hour number (h) and day number (d) such that

$$R(h,d) \approx R_{Q\mathrm{sioHD}} \tag{21}$$

the value of $R_{Q\mathrm{sioHD}}$ can be estimated for any hour and date, thereby allowing calculation of Q_{siH}.

Figure 8 plots hourly values of $R_{Q\mathrm{sioHD}}$ computed for latitudes of 49°N and 55°N which roughly define the southern and northern extents of the settled Prairie Provinces. Evidently, $R_{Q\mathrm{sioHD}}$ is only very weakly dependent on latitude. Therefore $R_{Q\mathrm{sioHD}}$ values for 52°N were used for the purpose of determining $R(h,d)$, which is computed from

$$R(h,d) = \max\left[\sin\left(\frac{c1\pi h}{24} + c2\right)c3, 0\right] \tag{22}$$

where $c1$, $c2$, $c3$ = functions of day number.

A small problem with the formulation of Equation (22) is that the sine function can produce spurious values of $R(h, d)$ for winter months at times when the Sun is actually below the horizon. This can be avoided by limiting the value of $R(h,d)$ to zero between the hours of 22:00 and 03:00.

Values of $c1$, $c2$, and $c3$ were determined iteratively for the first day of each month. By continuity, the mean value of $R(h, d)$ over each day should be 1. Therefore the solutions of Equation (22) were constrained so that the daily mean of $R(h, d)$ computed using the coefficients was within the range of 0.99–1.01. Second-order polynomial

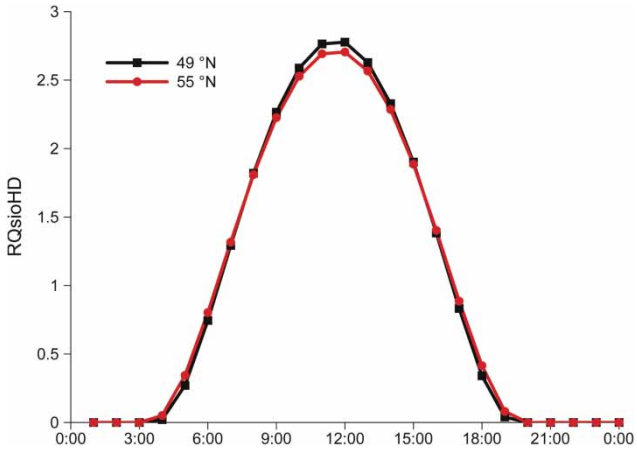

Figure 8 | Hourly values of $R_{Q\mathrm{sioHD}}$ (the ratio of Q_{sioH} to Q_{sioD}), computed for latitudes 49° and 55°N, for day number 172 (June 21, for non-leap year).

equations were fitted to plots of $c1$, $c2$, and $c3$ vs d. In all cases the correlation coefficients of the fit exceeded 0.995.

For days numbered 1 through 305 (Jan. 1 through Nov. 1, for a non-leap year) the value of $c3$ is well estimated by the second-order polynomial

$$c3 = 8.840749 \times 10^{-5}\, d^2 - 0.030195\, d + 5.33384 \qquad (23)$$

Although Equation (23) does not work well for days 305 through 366, values of Q_{si} are rarely required for this period. If necessary, the value for day 304 can be used as a constant for this interval.

The second-order polynomial giving $c2$ as a function of d can be replaced by simple linear function of $c3$:

$$c2 = -0.963832\, c3 + 7.86346 \qquad (24)$$

Similarly, $c1$ is also a linear function of $c2$:

$$c1 = -0.658542\, c2 + 5.18605 \qquad (25)$$

The coefficients of Equations (23) and (24) were determined through linear regressions for which the correlation coefficients exceeded 0.999.

Figure 9 demonstrates that the mean daily value of $R(h, d)$ is generally within ±1% of 1 over the year, except for very late in the year, as described above.

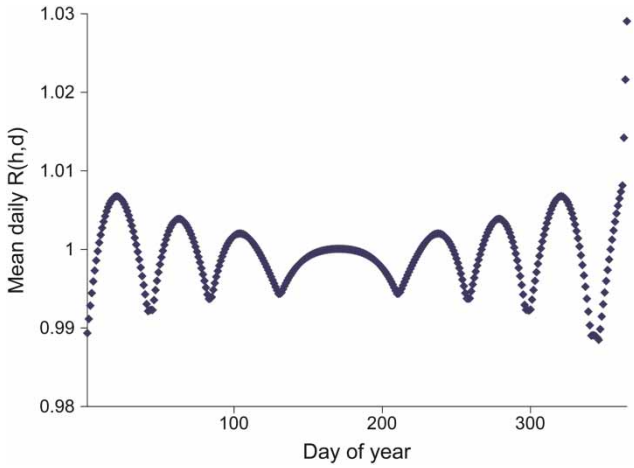

Figure 9 | Mean daily value of $R(h,d)$ versus day of year.

TESTING THE HOURLY INTERPOLATION

The usefulness of $R(h, d)$ was tested by regressing Annandale Q_{siD} values interpolated hourly against measured Q_{siH} values. The coefficients of the regression, which are summarized in Table 5, show that the hourly regression has a greater value of R^2 than the mean for the Annandale regressions summarized in Table 4. Although the standard error of the residuals of the daily regressions appears to be large, it is increased because the range of Q_{siH} values is much greater than that of the daily values. As shown in Figures 8 and 9, the temporal distribution of $R(h, d)$ over each day is accurate and there is good agreement

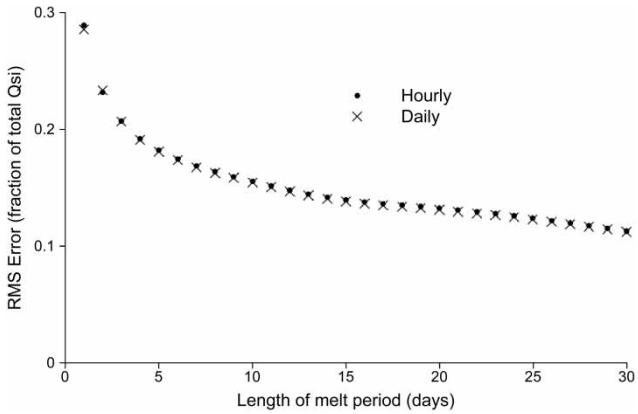

Figure 10 | RMS Error as a fraction of total Q_{si}, for cumulative hourly and daily Winnipeg Annandale estimates.

Table 5 | Regression constants of daily Annandale Q_{siD} interpolated to hourly against measured Q_{siH}. All data are derived from Winnipeg spring data, with all hourly values being greater than zero

				Confidence levels				
				Intercept		Slope		
Regression	Intercept (W/m^2)	Slope	R^2	2.5% (W/m^2)	97.5% (W/m^2)	2.5%	97.5%	Std. Error of Residual (W/m^2)
Hourly	86.4	0.73	0.65	82.3	90.6	0.72	0.74	123.3

between the actual and theoretical daily mean values of $R(h, d)$. Therefore, as hourly values are aggregated in the simulation of snowmelt, their total RMS error will be essentially the same as that of the daily values, as is shown in Figure 10.

SUMMARY AND CONCLUSIONS

The regressions of estimated Q_{siD} against measured values demonstrate that the simple, semi-empirical Annandale method was best able to serve as a replacement for measured individual daily measurements for plains and mountain sites in western Canada. Over the spring melt period, the mean RMS error of the Annandale method was smaller than that of the NARR data, and declined at a greater rate.

The spring NARR data have probability distributions which are very close in shape to those of the measurements, indicating that it may be possible to compensate for the error in the computed values. Unfortunately, such compensation is very difficult in practice as the degree of bias varies spatially and temporally. The NCEP values were disappointing as they showed the poorest fit for measured Q_{siD}. The NCEP Q_{siD} values should only be used as a last resort, when no other data are available.

For methods such as Annandale and Campbell–Bristow–Walter, which are only capable of producing daily Q_{si}, hourly values can be interpolated from daily estimates using a simple equation whose parameters require no data other than the date and the hour. Because the daily mean of the interpolation function is very close to the ideal value of 1, the RMS error of the hourly values is essentially identical to that of the daily values, for periods greater than one day.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support of the Drought Research Initiative (DRI), the Improved Processes and Parameterisation for Prediction in Cold Regions Network (IP3), the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS), NSERC and SGI Canada. This research was performed in its entirety with Free Open Source Software (F.O.S.S.).

REFERENCES

Annandale, J. G., Jovanovic, N. Z., Benadé, N. & Allen, R. G. 2001 Software for missing data analysis of Penman-Monteith reference evapotranspiration. *Irrig. Sci* **21**, 57–67.

Armstrong, R. N., Pomeroy, J. W. & Martz, L. W. 2008 Evaluation of three evaporation estimation methods in a Canadian prairie landscape. *Hydrol. Process* **22**, 2801–2815.

Ball, R. A., Purcell, L. C. & Carey, S. K. 2004 Evaluation of solar radiation prediction models in North America. *Agron. J* **96**, 1500.

Black, T. L. 1994 The new NMC mesoscale eta model: description and forecast examples. *Weather Forecast* **8**, 265–278.

Bland, W. L. & Clayton, M. K. 1994 Spatial structure of solar radiation in Wisconsin. *Agric. Forest Meteorol* **69**, 75–84.

Bristow, K. L. & Campbell, G. S. 1984 On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. Forest Meteorol* **31**, 159–166.

Brutsaert, W. 1982 *Evaporation into the Atmosphere*. Reidel, London.

Chatterjee, S. & Hadi, A. S. 2006 Regression analysis by example. *Technometrics* (Fourth Edin, Vol. 49, p. 375). Hoboken, New Jersey: John Wiley and Sons, Inc.

Durbin, J. & Watson, G. S. 1951 Testing for serial correlation in least squares regression. II. *Biometrika* **38** (1–2), 159–178.

Garnier, B. J. & Ohmura, A. 1968 A method of calculating the direct shortwave radiation income of slopes. *J. Appl. Meteorol* **7**, 796–800.

Granger, R. J. & Gray, D. M. 1990 A net radiation model for calculating daily snowmelt in open environments. *Nordic Hydrol* **21** (4–5), 217–234.

Gray, D. M. & Landine, P. G. 1987 Albedo model for shallow prairie snowcovers. *Can. J. Earth Sci* **24** (9), 1760–1768.

Helgason, W. D. 2009 *Energy Fluxes at the Air–Snow Interface*. PhD Thesis, University of Saskatchewan.

Marshall, I. B., Smith, C. A. & Selby, C. J. 1996 A national framework for monitoring and reporting on environmental sustainability in Canada. *Env. Monitor. Assess* **39**, 25–38.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D. & Shi, W. 2006 North American regional reanalysis. *Bull. Amer. Meteor. Soc* **87**, 343–360.

National Renewable Energy Laboratory 2007 National Solar Radiation Database 1991–2005 Update: User's Manual. Instrumentation. Golden, CO. Retrieved from www.nrel.gov.

Oke, T. R. 1987 *Boundary Layer Climates*, 2nd edition. Methuen, London.

Penman, H. L. 1948 Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. A* **193** (1032), 120–145.

R Development Core Team. 2008 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Shafran, P., Woollen, J., Ebisuzaki, W., Shi, W., Fan, Y., Grumbine, R. & Fennessy, M. 2004 Observational data used for assimilation in the NCEP North American regional reanalysis. Preprints, 14th Conf. on Applied Climatology, Seattle, WA, American Meteorological Society 1.4.

Sicart, J. E., Pomeroy, J. W., Essery, R. L. H. & Bewley, D. 2006 Incoming longwave radiation to melting snow: observations, sensitivity and estimation in Northern environments. *Hydrol. Process* **20** (17), 3697–3708.

Thornton, P. E. & Running, S. W. 1999 An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. Forest Meteorol* **93**, 211–228.

Walter, M. T., Brooks, E. S., McCool, D. K., King, L. G., Molnau, M. & Boll, J. 2005 Process-based snowmelt modelling: does it require more input data than temperature-index modelling? *J. Hydrol* **300**, 65–75.

First received 26 June 2009; accepted in revised form 15 October 2010. Available online July 2011