ELSEVIER

Journal of Hydrology 184 (1996) 337–354

Journal of Hydrology

Regional models of potential evaporation and reference evapotranspiration for the northeast USA

Neil M. Fennessey[a,*], Richard M. Vogel[b]

[a]Department of Civil and Environmental Engineering, University of Massachusetts, North Dartmouth, MA 02747, USA
[b]Department of Civil and Environmental Engineering, Tufts University, Anderson Hall, Medford, MA 02155, USA

Received 22 August 1994; accepted 29 October 1995

Abstract

Multivariate regression models of average monthly reference crop evapotranspiration, $E_t(\tau)$, and potential evaporation, $E_p(\tau)$, are developed for the northeast USA. Average monthly values of daily E_t and E_p are estimated from the Penman–Monteith equation using monthly climate data from 34 National Oceanic and Atmospheric Administration (NOAA) First Order weather observatories and specific reference vegetation characteristics. The periodic seasonal behavior of $E_t(\tau)$ and $E_p(\tau)$ and average monthly temperature are approximated by Fourier series functions. Regional regression relationships are then developed which relate the $E_t(\tau)$ and $E_p(\tau)$ Fourier series coefficients with the average monthly temperature Fourier coefficients, station longitude and station elevation. The regional $E_t(\tau)$ model is shown to be an improvement over the Linacre method (Linacre, 1977, *Agric. Meteorol.*, 18: 409–424) and the Hargreaves and Samani method (Hargreaves and Samani, 1985, *Appl. Eng. Agric.*, 1(2): 96–99), and the regional $E_p(\tau)$ model is shown to be an improvement over the Linacre method.

1. Introduction

The Penman–Monteith equation (Monteith, 1965) is an accepted method for describing reference crop evapotranspiration E_t: the atmosphere's near-earth surface demand for water vapor above a well-watered reference vegetation. The primary disadvantage of the Penman–Monteith equation is its data requirements, which include net solar radiation, windspeed, dewpoint temperature, air temperature and vegetation-specific parameters.

* Corresponding author.

PII 0022-1694(95)02980-X

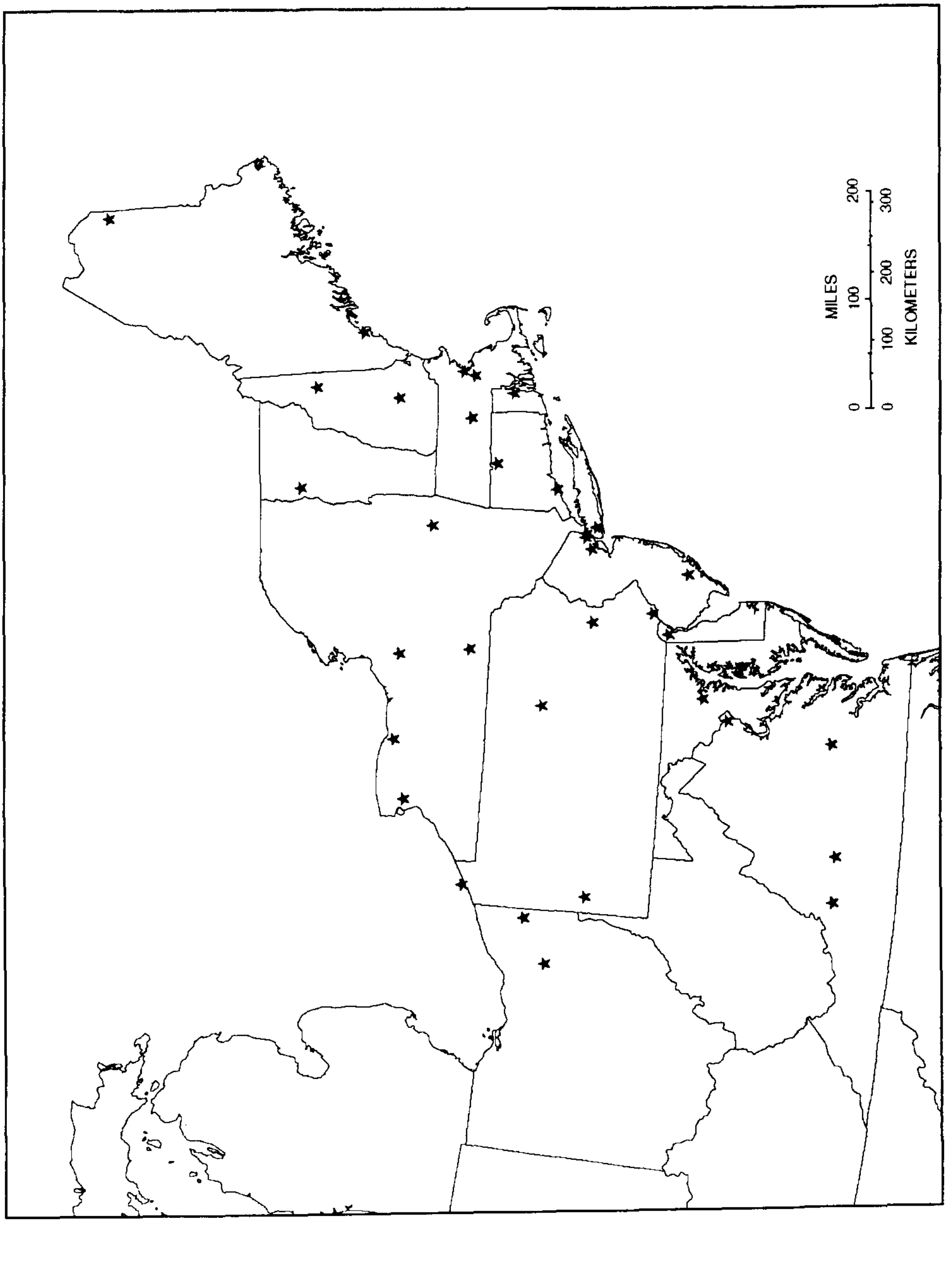

Fig. 1. Location of the NOAA First Order observatories.

These climate data are generally only available at National Oceanic and Atmospheric Administration (NOAA) First Order observatories, typically located at major metropolitan airports and staffed by trained weather observers. For example, Fig. 1 illustrates the locations of the 34 NOAA First Order observatories used in this study. In contrast, local average monthly temperature data are far more readily available from other Federal or State supported networks (over 1000 stations in New England, New York, New Jersey and Pennsylvania). As E_t can only readily be estimated at First Order observatories, a method to estimate E_t using local temperature and other readily available site-specific data would be useful in a broad range of water resources studies including reservoir yield and climate change impact analyses.

The first objective of this study was the development of a multivariate regional regression model which may be used to estimate $E_t(\tau)$, $\tau = 1,12$, the average monthly evapotranspiration rate at any location in the northeastern USA for a reference vegetation. The second objective was to develop a similar regional model for monthly average potential evaporation, $E_p(\tau)$, $\tau = 1,12$, the atmosphere's demand for water vapor above a free water surface. In this study, $E_t(\tau)$ is estimated for a reference grass of 12 cm height. The study concludes with a comparison of the methods developed here for estimating $E_t(\tau)$ and $E_p(\tau)$ with other temperature-based models.

2. Reference crop evapotranspiration

Reference crop evapotranspiration is the atmosphere's near-earth surface demand for water vapor from a moist, well-watered soil completely covered by a transpiring but dry vegetation canopy. The reference vegetation has specific physical characteristics. The Penman–Monteith equation (Monteith, 1965) remains the model of choice when estimating actual evapotranspiration (see Jensen et al. (1990) and Shuttleworth (1993)).

The Penman–Monteith equation describes the vertical (one-dimensional) latent heat flux, λE_t:

$$\lambda E_t = \frac{\Delta(R_n - G_t) + \dfrac{\rho C_p[e^\circ(z) - e(z)]}{r_a}}{\Delta + \gamma\left(1 + \dfrac{r_c}{r_a}\right)} \tag{1}$$

where λ is the latent heat of vaporization, E_t is the reference crop evapotranspiration, Δ is the gradient of the saturation vapor pressure–temperature function, R_n is the net radiation, G_t is the vertical heat energy flux within the soil, ρ is the air density, C_p is the specific heat of the air at constant pressure, $e^\circ(z)$ is the saturated vapor pressure of the air measured at height z, $e(z)$ is the vapor pressure of the air measured at height z, r_a is the aerodynamic resistance to water vapor diffusion into the atmospheric boundary layer, r_c is the vegetation's canopy resistance to water vapor transfer, and γ is the psychrometric constant. Further details associated with Eq. (1) have been described by Jensen et al. (1990), Fennessey (1994) and Fennessey and Kirshen (1994), and elsewhere.

2.1. Vegetation data

Several vegetation-specific variables are required by the Penman–Monteith equation. In this study, r_a, the aerodynamic resistance (see, e.g. Brutsaert (1984) or Jensen et al. (1990, p. 93)) is determined by

$$r_a = \frac{\ln\left(\frac{z_w - d_o}{z_{om}}\right)\ln\left(\frac{z_p - d_o}{z_{ov}}\right)}{\kappa^2 U(z_w)} \tag{2}$$

where z_w and z_p are the heights of the windspeed anemometer and humidity psychrometer, respectively, d_o is the zero-windspeed plane displacement height, z_{om} is the roughness length for momentum transfer, z_{ov} is the roughness length for vapor transfer, κ is von Kármán's constant and $U(z_w)$ is the windspeed measured at height z_w. Eq. (2) is valid for a neutrally buoyant atmosphere (no free convection), which is a reasonable assumption for estimates of reference crop evapotranspiration at time scales of 1 day or greater.

For a reference vegetation, we use a grass of height $h_c = 12$ cm with a canopy resistance, r_c, of 70 s m^{-1}. Jensen et al. (1990) suggested $d_o = 0.67h_c$, $z_{om} = 0.123h_c$, $z_{ov} = 0.1z_{om}$ and $\kappa = 0.41$. Finally, we use surface albedo of 0.23, as suggested by Shuttleworth (1993), which is necessary to estimate the net radiation, R_n, in Eq. (1) (see Fennessey and Kirshen, 1994).

3. Regional regression model for reference crop evapotranspiration

Our objective was to develop a procedure for obtaining an accurate estimate of $E_t(\tau)$, the average monthly value of the daily reference crop evapotranspiration rate, using readily available independent variables. This would allow one to obtain good estimates of $E_t(\tau)$ anywhere in the northeastern USA, unlike the current situation, which only allows for good estimates of $E_t(\tau)$ in the 'local' vicinity of major metropolitan airports. Spatial interpolation of $E_t(\tau)$ estimates derived from NOAA observatory data is not an attractive way to estimate site-specific $E_t(\tau)$ because the topography of the northeast USA between adjacent stations may range from coastal plains to mountains of 1900 m or more.

3.1. Regional climate data

Estimates of the average monthly reference crop evapotranspiration using the Penman–Monteith equation were obtained using climate data from 34 NOAA First Order observatories and the procedures described by Fennessey and Kirshen (1994). The data are based upon average monthly values of the daily climate variables derived from the 'Normal' 1951–1980 record period. These long-term normal climate data are available in the Local Climatological Data Annual Summary with Comparative Data, published by the National Climatic Data Center (NCDC), Ashville, NC. The locations of these observatories are shown in Fig. 1 and listed in Table 1. These data include either observed cloud cover and/ or per cent of possible sunshine, relative humidity, windspeed and air temperature. The station air pressure was also used if it was reported. The NOAA observatories record

Table 1
NOAA First Order climate observatories

Gage	Location	NOAA id.	Latitude	Longitude	Elevation
1	Portland, ME	14764	43°39′	70°18′	13
2	Caribou, ME	14607	46°52′	68°01′	190
3	Concord, NH	14745	43°12′	71°30′	104
4	Mt. Washington, Gorham, NH	14755	44°16′	71°18′	1909
5	Burlington, VT	14742	44°28′	73°09′	101
6	Blue Hill Obs., Milton, MA	14753	42°13′	71°07′	192
7	Boston, MA	14739	42°22′	71°02′	5
8	Worcester, MA	94746	42°16′	71°52′	301
9	Hartford, CT	14740	41°56′	72°41′	52
10	Bridgeport, CT	94702	41°10′	73°08′	2
11	Providence, RI	14765	41°44′	71°26′	16
12	Syracuse, NY	14771	43°07′	76°07′	125
13	Rochester, NY	14768	43°07′	77°40′	167
14	La Guardia Field, New York, NY	14732	40°46′	73°54′	3
15	JFK Airport, New York, NY	94789	40°39′	73°47′	4
16	Central Park, New York, NY	94728	40°47′	73°58′	40
17	Buffalo, NY	94728	42°56′	78°44′	215
18	Binghampton, NY	04725	42°13′	75°59′	485
19	Albany, NY	14735	42°45′	73°48′	84
20	Newark, NJ	14734	40°42′	74°10′	2
21	Atlantic City, NJ	93730	39°27′	74°34′	20
22	Williamsport, PA	14778	41°15′	76°55′	160
23	Pittsburgh, PA	94823	40°30′	80°13′	347
24	Philadelphia, PA	13739	39°53′	75°15′	2
25	Erie, PA	14860	42°05′	80°11′	223
26	Allentown, PA	14737	40°39′	75°26′	118
27	Baltimore, MD	93721	39°11′	76°40′	45
28	Richmond, VA	13740	37°30′	77°20′	50
29	Roanoke, VA	13741	37°19′	79°58′	350
30	Lynchburg, VA	13733	37°20′	79°12′	281
31	Wilmington, DE	13781	39°40′	75°36′	23
32	National Arpt., Washington, DC	13743	38°51′	77°02′	3
33	Youngstown, OH	14852	41°15′	80°40′	359
34	Akron, OH	14895	40°55′	81°26′	368

relative humidity four times daily and the Annual Summary tabulates the 1951–1980 average of each value every month. In this study, the average of those four daily relative humidity values was used to estimate a single average daily relative humidity.

3.2. Fourier series approximation of $E_t(\tau)$

Periodic data can be approximated by a continuous Fourier series function. For this study, we used the Fourier series formulation reported by Bloomfield (1976) to approximate $E_t(\tau)$. The Fourier series approximation is

$$E_{t,f}(\tau) \approx E_{ta} + \sum_{k=1}^{m} a_k \cos\left[\left(\frac{\pi k \tau}{6}\right) + b_k \sin\left(\frac{\pi k \tau}{6}\right)\right] \tag{3}$$

where τ is the month, $\tau = 1,12$; $E_{t,f}(\tau)$ denotes the Fourier series approximation to the average monthly value of the daily reference crop evapotranspiration during the τth month; E_{ta} is the annual average daily reference crop evapotranspiration rate; k is the summation index for the kth harmonic; m equals the total number of harmonics required to accurately approximate $E_t(\tau)$ derived from the Penman–Monteith equation. We determined that $m = 2$ harmonics results in values of $E_{t,f}(\tau)$ which closely approximate $E_t(\tau)$.

The Fourier series coefficients for the kth harmonic of $E_{t,f}(\tau)$, a_k and b_k, are estimated using

$$a_k = \frac{1}{6}\sum_{\tau=1}^{12}[E_t(\tau) - E_{ta}]\cos\left(\frac{\pi k \tau}{6}\right) \tag{4a}$$

$$b_k = \frac{1}{6}\sum_{\tau=1}^{12}[E_t(\tau) - E_{ta}]\sin\left(\frac{\pi k \tau}{6}\right) \tag{4b}$$

3.3. *Regional regression model for* $E_t(\tau)$

The near-earth surface air temperature is an indicator of the planetary boundary layer heat and moisture fluxes and the surface energy balance. Therefore, we attempted to estimate $E_t(\tau)$ using multivariate regression with air temperature likely to be the most important candidate independent variable. As noted by Fennessey and Kirshen (1994), equations which describe many of the individual Penman–Monteith equation variables in Eq. (1) are temperature dependent.

Regional regression equations were developed which describe the five reference crop evapotranspiration Fourier coefficients of a two harmonic form of Eq. (3): E_{ta}, a_1, b_a, a_2 and b_2. The candidate independent variables include station decimal latitude, longitude and elevation, average monthly temperature and the average annual temperature. Owing to significant intercorrelation among monthly values of temperature (range 0.932–0.997), we chose to fit a Fourier series to monthly temperatures rather than face parameter estimation difficulties associated with multicollinearity (see Hirsch et al., 1993).

Fourier series were fitted to monthly average temperatures, $T(\tau)$, $\tau = 1,12$, shown by

$$T_f(\tau) \approx T_a + \sum_{k=1}^{m}\left[c_k\cos\left(\frac{\pi k \tau}{6}\right) + d_k\sin\left(\frac{\pi k \tau}{6}\right)\right] \tag{5}$$

where $T_f(\tau)$ is the Fourier series approximation of the monthly averaged daily temperature (in °C) for the τth month and T_a is the annual average daily temperature. Like $E_{t,f}(\tau)$, we limited $T_f(\tau)$ to the first two harmonics, hence the five $T_f(\tau)$ Fourier series coefficients are described by T_a, c_1, d_1, c_2 and d_2 using

$$T_a = \frac{1}{12}\sum_{\tau=1}^{12} T(\tau) \tag{6a}$$

$$c_1 = \frac{1}{6}\sum_{\tau=1}^{12}[T(\tau) - T_a]\cos\left(\frac{\pi \tau}{6}\right) \tag{6b}$$

Table 2
Coefficients for the Eq. (7) regional regression equations required to estimate the reference crop evapotranspiration

θ_t	Model coefficients							
	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7
E_{ta}	1.25	0.0	0.0	0.118	0.0	0.0	0.0	0.0
a_1	1.62	−0.00843	0.0	−0.0860	0.191	0.0	−0.564	0.0
b_1	2.59	−0.0137	0.0	−0.0318	0.0	0.269	0.0	−0.190
a_2	−0.637	0.00761	−0.000095	0.0	−0.0573	0.0374	0.404	0.0
b_2	−0.708	0.00409	−0.000131	0.0	−0.0403	−0.039	0.271	0.205

$$d_1 = \frac{1}{6}\sum_{\tau=1}^{12}[T(\tau) - T_a]\sin\left(\frac{\pi\tau}{6}\right) \tag{6c}$$

$$c_2 = \frac{1}{6}\sum_{\tau=1}^{12}[T(\tau) - T_a]\cos\left(\frac{\pi\tau}{3}\right) \tag{6d}$$

$$d_2 = \frac{1}{6}\sum_{\tau=1}^{12}[T(\tau) - T_a]\sin\left(\frac{\pi\tau}{3}\right) \tag{6e}$$

Even though the monthly average temperature values, $T(\tau)$, are highly collinear, the correlation matrix among the five Fourier coefficients of $T_f(\tau)$ are much smaller, ranging from 0.407 to 0.828.

Regional equations for dependent variables E_{ta}, a_1, a_2, b_1 and b_2 are developed using the independent variables station decimal longitude, station elevation, T_a, c_1, c_2, d_1 and d_2. The station latitude is highly correlated (r = 0.920) with d_1; hence, latitude was dropped from further consideration. The final regression equations take the form

$$\begin{aligned}\theta_t &= e_0 + e_1\mathrm{Long} + e_2\mathrm{Elev} + e_3T_a + e_4c_1 \\ &\quad + e_5d_1 + e_6c_2 + e_7d_2\end{aligned} \tag{7}$$

where Long is the site longitude (in decimal degrees), Elev is the site elevation (in meters), and θ_t denotes the dependent variable E_{ta}, a_1, b_1, a_2 or b_2.

Table 2 shows the coefficients e_0 to e_7 for the regional equations. Table 3 documents the

Table 3
Student t-ratios and adjusted coefficient of determination, R^2, of the five regional regressions described by Eq. (7)

θ_t	Student t-Ratios[a]								R^2
	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7	
E_{ta}	23.53	–	–	23.02	–	–	–	–	94.1
a_1	5.02	−2.42	–	−14.68	10.95	–	−5.33	–	91.8
b_1	7.32	−4.33	–	−9.31	–	11.67	–	−3.84	83.7
a_2	−4.24	6.48	−6.30	–	−8.57	2.40	7.60	–	90.0
b_2	−2.46	2.29	−7.86	–	−2.55	−2.23	4.42	4.19	84.8

[a] The critical value of the t-ratio, using a 5% significance level, is $t_{0.025,32}$ = 2.03.

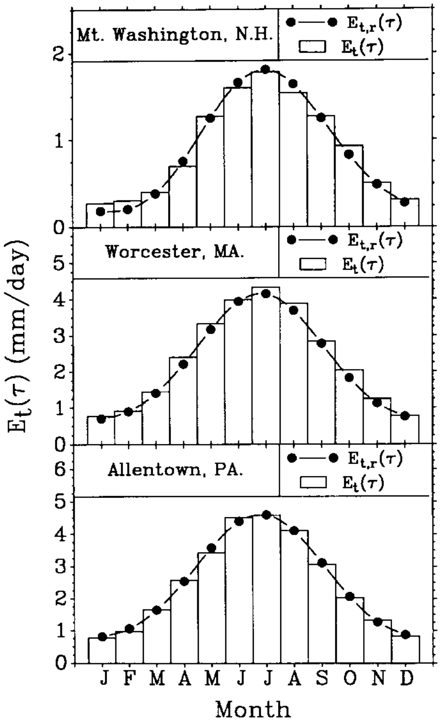

Fig. 2. Comparison between Penman–Monteith estimated $E_t(\tau)$ and regional regression model $E_{t,r}(\tau)$ reference crop evapotranspiration at three NOAA observatories.

Student t-ratio for each coefficient and each regression equation's adjusted coefficient of determination, R^2. It should be noted that R^2 cannot be computed if the intercept coefficient, e_o, equals zero. All computed t-ratios in Table 3 were larger than $T_{crit} = t_{0.025,32} = 2.03$, hence we conclude that all coefficients are significantly different from zero at the $\alpha = 0.05$ significance level. In addition, all model residuals were found to be approximately normally distributed, using a 5% level normal probability plot correlation coefficient hypothesis test (see Stedinger et al. (1993)).

The final regional model for $E_t(\tau)$, denoted $E_{t,r}(\tau)$, is

$$E_{t,r}(\tau) = E_{ta} + a_1\cos\left(\frac{\pi\tau}{6}\right) + b_1\sin\left(\frac{\pi\tau}{6}\right) + a_2\cos\left(\frac{\pi\tau}{3}\right) + b_2\sin\left(\frac{\pi\tau}{3}\right) \qquad (8)$$

where the model coefficients are obtained using the regional regression equations in Eq. (7) with model coefficients summarized in Table 2.

The overall goodness-of-fit associated with the regional regression $E_{t,r}(\tau)$ is summarized in Fig. 2 and Fig. 3. Fig. 2 compares direct application of the Penman–Monteith equation $E_t(\tau)$ and the regression estimator $E_{t,r}(\tau)$ for three NOAA observatories. The upper figure corresponds to the Mt. Washington Observatory, where some of the worst weather in the

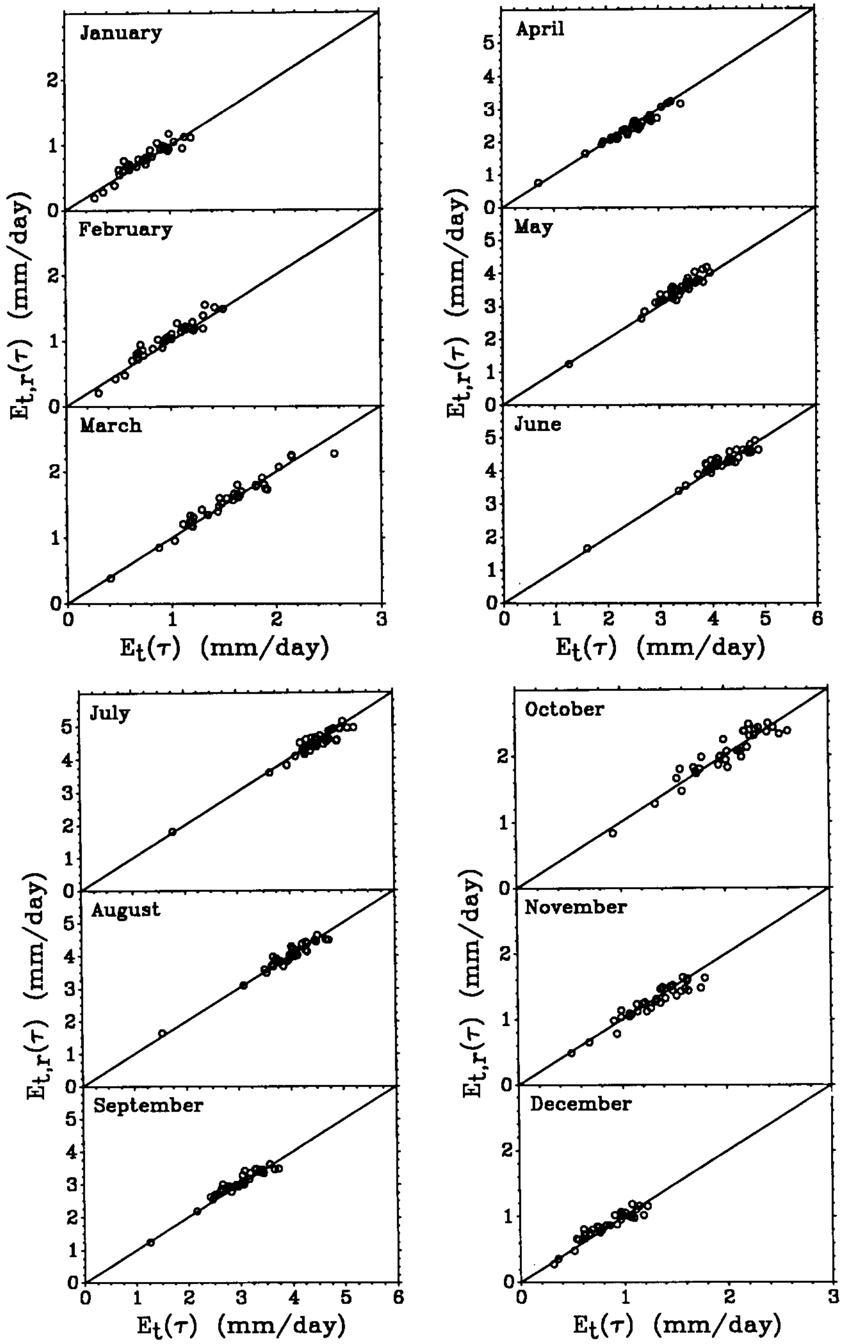

Fig. 3. Comparison between Penman–Monteith estimated $E_t(\tau)$ and regional regression model $E_{t,r}(\tau)$ reference crop evapotranspiration by month.

world has been observed, including the highest windspeed ever recorded. The agreement between $E_t(\tau)$ and $E_{t,r}(\tau)$ for the three sites illustrated in Fig. 2 is indicative of the agreement obtained overall, for the remaining 31 NOAA observatories. Fig. 3 illustrates the overall goodness-of-fit of $E_t(\tau)$ vs. $E_{t,r}(\tau)$ by month for all 34 NOAA observatory locations used to develop the regression equations.

3.4. Comparison with other studies

It is useful to examine how well the regression model $E_{t,r}(\tau)$ compares with other temperature-based models of reference crop evapotranspiration. Because of their simplicity, temperature models for estimating reference crop evapotranspiration E_t have been used for many years. As a result of comments by Jensen et al. (1990), one of the most popular methods, the Thornthwaite (1948) method, is not included in this comparison. We compare $E_{t,r}(\tau)$ with the Linacre (1977) model, $E_{t,L}(\tau)$, and a more recent approach, the Hargreaves method, $E_{t,H}(\tau)$, described by Hargreaves and Samani (1985) and Hargreaves (1994).

The model described by Linacre (1977) approximates the monthly averaged daily Penman (1948) method evapotranspiration rate. The Linacre model, $E_{t,L}(\tau)$, depends upon average monthly maximum and minimum temperature (from which average monthly temperatures are estimated and therefore widely available), latitude and elevation. This method is valid for regions where the average monthly precipitation is at least 5 mm month^{-1} and the average difference between the mean monthly air and dewpoint temperature is at least 4°C, conditions that are met in the northeast USA.

Shuttleworth (1993) recommended that if a temperature method must be used to estimate E_t, owing to the lack of the necessary data required for the Penman–Monteith equation, one should employ the method described by Hargreaves and Samani (1985). According to Jensen et al. (1990, p. 102) Hargreaves and Samani (1985) developed their model from 8 years of cool-season alta fescue grass grown near Davis, CA. The Hargreaves and Samani $E_{t,H}(\tau)$ method requires site latitude, and mean maximum and minimum monthly temperatures.

A 1992 study discussed by Hargreaves (1994) was conducted by the Centre Commun de Recherche, Commission des Communautes Europeennes (CCREEC) using synoptic climate data and lysimeter E_t data. They chose to use the Penman (1948) equation as the standard method. As discussed by Hargreaves (1994), climate data from various locations were used to compare the predictive power of nine E_t models, each of which is less data intensive than the Penman method, against the Penman equation prediction. Hargreaves (1994) reported that the Hargreaves and Samani (1985) method was selected by the CCREEC as the $E_t(\tau)$ model with results closest to the Penman prediction.

3.5. Goodness-of-fit comparison

This section compares how well the regression model, $E_{t,r}(\tau)$, the Hargreaves and Samani model, $E_{t,H}(\tau)$, and the Linacre model, $E_{t,L}(\tau)$, approximate the Penman–Monteith method estimated $E_t(\tau)$. We measure the goodness-of-fit using the monthly bias, BIAS,

Table 4
A comparison of monthly BIAS and r.m.s.e. for the three estimators $E_{t,r}(\tau)$, $E_{t,H}(\tau)$ and $E_{t,L}(\tau)$

τ	Month	BIAS[w]		
		$w = E_{t,r}(\tau)$ (mm day^{-1})	$w = E_{t,H}(\tau)$ (mm day^{-1})	$w = E_{t,L}(\tau)$ (mm day^{-1})
1	January	0.00	−0.20	−0.54
2	February	0.04	−0.12	−0.55
3	March	−0.01	−0.03	−0.24
4	April	−0.08	0.25	−0.37
5	May	0.10	0.60	1.05
6	June	0.04	0.59	1.69
7	July	−0.07	0.50	2.26
8	August	0.01	0.44	2.45
9	September	0.08	0.37	2.16
10	October	0.00	0.09	1.39
11	November	−0.04	−0.18	0.53
12	December	0.02	−0.22	−0.27
τ	Month	r.m.s.e.[w]		
		$w = E_{t,r}(\tau)$ (mm day^{-1})	$w = E_{t,H}(\tau)$ (mm day^{-1})	$w = E_{t,L}(\tau)$ (mm day^{-1})
1	January	0.10	0.23	0.57
2	February	0.11	0.17	0.59
3	March	0.11	0.15	0.35
4	April	0.13	0.31	0.46
5	May	0.17	0.65	1.08
6	June	0.18	0.70	1.70
7	July	0.19	0.66	2.25
8	August	0.15	0.60	2.43
9	September	0.15	0.52	2.15
10	October	0.14	0.28	1.39
11	November	0.11	0.25	0.56
12	December	0.10	0.25	0.35

and monthly root mean square error, r.m.s.e., respectively, across the 34 sites:

$$\mathrm{BIAS}[w(\tau)] = \frac{1}{34}\sum_{j=1}^{34}\left[w(\tau)_j - E_t(\tau)_j\right] \tag{9}$$

$$\mathrm{r.m.s.e.}[w(\tau)] = \left\{\frac{1}{34}\sum_{j=1}^{34}\left[w(\tau)_j - E_t(\tau)_j\right]^2\right\}^{1/2} \tag{10}$$

where $W(\tau)$ equals $E_{t,r}(\tau)$, $E_{t,H}(\tau)$ or $E_{t,L}(\tau)$.

To evaluate more fairly the BIAS and r.m.s.e. between $E_{t,r}(\tau)$ and $E_t(\tau)$, 33 separate sets of the five regional regression equations for E_{ta}, a_1, a_2, b_1 and b_2 were developed where each set does not include one of the 34 NOAA sites. Comparisons between $E_t(\tau)$ and $E_{t,r}(\tau)$ are more fairly made because the regression estimate $E_{t,r}(\tau)$ for a particular site is not based on temperature, longitude and elevation data from that site. Instead, the model employed to estimate $E_{t,r}(\tau)$ at that site is based on data from the remaining 33 NOAA sites. Thus, the r.m.s.e.[$E_{t,r}(\tau)$] is analogous to the PRESS statistic, used for validation of regression equations (see Hirsch et al., 1993).

Table 4 documents the BIAS and r.m.s.e. associated with the regression estimator $E_{t,r}(\tau)$, with the Hargreaves and Samani estimator, $E_{t,H}(\tau)$, and Linacre estimator $E_{t,L}(\tau)$. Clearly, the regression model $E_{r,r}(\tau)$ provides a very good approximation to $E_t(\tau)$. Practically speaking, even the largest BIAS of 0.1 mm day^{-1} is essentially immeasurable. These results also suggest that $E_{t,r}(\tau)$ is superior to $E_{t,H}(\tau)$ and $E_{t,L}(\tau)$ during all months of the year for locations in the northeast USA.

4. Potential evaporation

Potential evaporation, $E_p(\tau)$, is simply the atmosphere's near-earth surface demand for water vapor above a well-moistened surface including open water. An accepted approach to estimating evaporation for a water body is to evaluate its energy budget (Bras, 1991). If the net energy contribution from precipitation, streamflow and other releases from a reservoir are accounted for, the Penman–Monteith equation would describe the actual evaporation rate where the canopy resistance, r_c, equals zero, as discussed by Shuttleworth (1993).

The potential evaporation rate, E_p, is described by

$$\lambda E_p = \frac{\Delta(R_n - G_p) + \dfrac{\rho C_p[e^{\circ}(z) - e(z)]}{r_a}}{\Delta + \gamma} \tag{11}$$

Eq. (11) was referred to as the modified Penman equation by Eagleson (1970), who cited the work of Van Bavel (1966). Monteith (1965, p. 208) attributed Eq. (11) to Penman (1948).

In this study, the water body heat flux, G_p, which is analogous to the soil heat flux, G_t, in Eq. (1), was set equal to zero. This assumption represents a free water surface as a thin film of water devoid of any significant heat storage capacity but having the albedo of a lake or pond. Because each water body has a unique heat capacity, no attempt is made to generalize this characteristic, which depends upon the volume, surface area and depth of the water body among other factors.

Monthly values of surface albedo reported by Brest (1987) and discussed by Fennessey and Kirshen (1994) are used. The canopy resistance term of Eq. (1), r_c, is equal to zero in the case of free surface evaporation and therefore is not shown in Eq. (11).

The formulation of the aerodynamic resistance term, r_a, is the same as that shown in Eq. (2); however, the values of z_{om} and z_{ov} for a water surface are different from those for a reference vegetation. This is because a water surface is aerodynamically smoother than a vegetation surface. For this study, estimates of z_{om} and z_{ov} over fresh water are made using the following approach.

From Brutsaert (1984, p. 58), and rearranging terms, the mean windspeed at elevation z above the surface is described by

$$\frac{\bar{U}(z_w)}{U_*} = \frac{1}{\kappa}\ln\left(\frac{z_w}{z_{om}}\right) \quad \text{for } z_w \gg z_{om} \tag{12}$$

where U_* equals the friction velocity.

From Brutsaert (1984, p. 63) and rearranging terms,

$$\frac{\bar{U}(z_w)}{U_*} = C_p\left(\frac{z_w - d_o}{z_o}\right)^m$$

where for over water, his dimensionless parameter $C_p = 6/7m$ and $m \approx 1/7$, and z_o equals the aerodynamic roughness length.

Typically, lake and sea surface aerodynamic studies measure windspeed at the z_w = 10 m (1000 cm) height. From Brutsaert (1984, p. 114, Table 5.1), for 'large water surfaces', z_o = 0.01–0.06 cm where smaller values correspond to shallow water bodies and larger values correspond to the open ocean. We assume that z_o = 0.01 cm is probably more representative of lakes and ponds in the northeast USA than z_o = 0.06 cm. The zero windspeed displacement plane over water, d_o, equals zero. Substituting these values into the above,

$$\begin{aligned}\frac{\bar{U}(z_w)}{U_*} &= 6\left(\frac{1000}{0.01}\right)^{\frac{1}{7}} \\ &\approx 30\end{aligned} \tag{13}$$

Substituting the above into Eq. (12), with k = 0.41 and z_w = 1000 cm, results in z_{om} = 0.003 cm. From Brutsaert (1984, p. 122), for an aerodynamically smooth surface, $z_{ov} = 4.6z_{om} \approx 5z_{om}$ = 0.015 cm.

There are those who advocate an empirical wind function to estimate r_a over water (e.g. see Shuttleworth (1993), who cited Penman (1948)). However, Brutsaert (1984) stated that Eq. (12) is adequate when used to estimate E_p over periods of a day or longer. For these longer time scales, the required condition that the atmosphere be stable and neutrally buoyant is most likely to be true. Shorter time scales require that r_a be corrected for atmospheric instability owing to forced convection.

5. A regional model for potential evaporation

As with the regional model developed for $E_t(\tau)$, a two harmonic Fourier series was fitted to $E_p(\tau)$, resulting in the approximation $E_{p,f}(\tau)$. Again, individual regional regression equations were developed for the five $E_{p,f}(\tau)$ Fourier coefficients using NOAA station

Table 5
Coefficients for the Eq. (15) regional regression equations required to estimate the potential evaporation

θ_p	Model coefficients							
	h_0	h_1	h_2	h_3	h_4	h_5	h_6	h_7
E_{pa}	2.78	−0.0163	0.000492	0.156	0.0	0.0	0.414	0.0
f_1	0.0	−0.0141	0.000200	−0.0520	0.0990	0.0	−0.591	0.0
g_1	0.962	−0.00908	0.0	−0.0187	−0.0570	0.192	0.0	0.0
f_2	0.0	0.0104	−0.000200	−0.0155	−0.0451	0.101	0.515	0.0
g_2	0.0	0.00606	−0.000246	−0.0142	0.0	0.0	0.235	0.075

Table 6
Student t-ratios and adjusted coefficient of determination, R^2, of the five regional regressions described by Eq. (15)

θ_p	Student t-ratios[a]								R^{2b}
	h_0	h_1	h_2	h_3	h_4	h_5	h_6	h_7	
E_{pa}	8.48	-3.32	7.05	14.90	–	–	3.05	–	94.8
f_1	–	-3.26	2.32	-4.08	4.54	–	-4.71	–	NA[b]
g_1	4.98	-5.08	–	-8.68	-5.98	14.07	–	–	85.9
f_2	–	6.05	-6.26	-3.41	-4.68	7.59	8.65	–	NA
g_2	–	10.03	-9.52	-3.53	–	–	5.87	2.95	NA

[a] The critical value of the t-ratio, using a 5% significance level, is $t_{0.025,32} = 2.03$.
[b] R^2 cannot be calculated when the intercept term, h_0, equals zero.

longitude, elevation, and the air temperature Fourier coefficients T_a, c_1, c_2, d_1 and d_2. It should be recalled that T_a, c_1, c_2, d_1 and d_2 are the coefficients in Eq. (5) which were estimated using Eq. (6a)–(6e). The regional model of potential evaporation, $E_{p,r}(\tau)$, is

$$E_{p,r}(\tau) = E_{pa} + f_1 \cos\left(\frac{\pi\tau}{6}\right) + g_1 \sin\left(\frac{\pi\tau}{6}\right) + f_2 \cos\left(\frac{\pi\tau}{3}\right) + g_2 \sin\left(\frac{\pi\tau}{3}\right) \tag{14}$$

where the five model coefficients, E_{pa}, f_1, f_2, g_1 and g_2 take the form of

$$\theta_p = h_0 + h_1 \mathrm{Long} + h_2 \mathrm{Elev} + h_3 T_a + h_4 c_1 + h_5 d_1 + h_6 c_2 + h_7 d_2 \tag{15}$$

where θ_p is the dependent variable E_{pa}, f_1, g_1, f_2 or g_2. The coefficients h_0 to h_7 are shown in Table 5 and their corresponding Student t-ratios and adjusted R^2 values are shown in Table 6. Fig. 4 compares $E_p(\tau)$ and $E_{p,r}(\tau)$ month by month, illustrating that $E_{p,r}(\tau)$ provides a good estimate of $E_p(\tau)$ across all sites.

5.1. Comparison with other studies

It is useful to examine how well the regression model $E_{p,r}(\tau)$ compares with another temperature-based models of potential evaporation. Linacre (1977) presented a temperature-based model for estimating potential evaporation $E_{p,L}(\tau)$. As with Linacre's $E_{t,L}(\tau)$ model, $E_{p,L}(\tau)$ requires mean monthly temperature and location latitude and elevation.

5.2. Goodness-of-fit comparison

As with the evaluation between $E_t(\tau)$ with $E_{t,r}(\tau)$, $E_{t,H}(\tau)$ and $E_{t,L}(\tau)$, we employ the monthly bias, BIAS, and monthly root mean square error, r.m.s.e., described by Eq. (9) and Eq. (10), respectively, to compare the modified Penman equation $E_p(\tau)$ against the regression method ($E_{p,r}(\tau)$) and the Linacre method ($E_{p,L}(\tau)$) estimates.

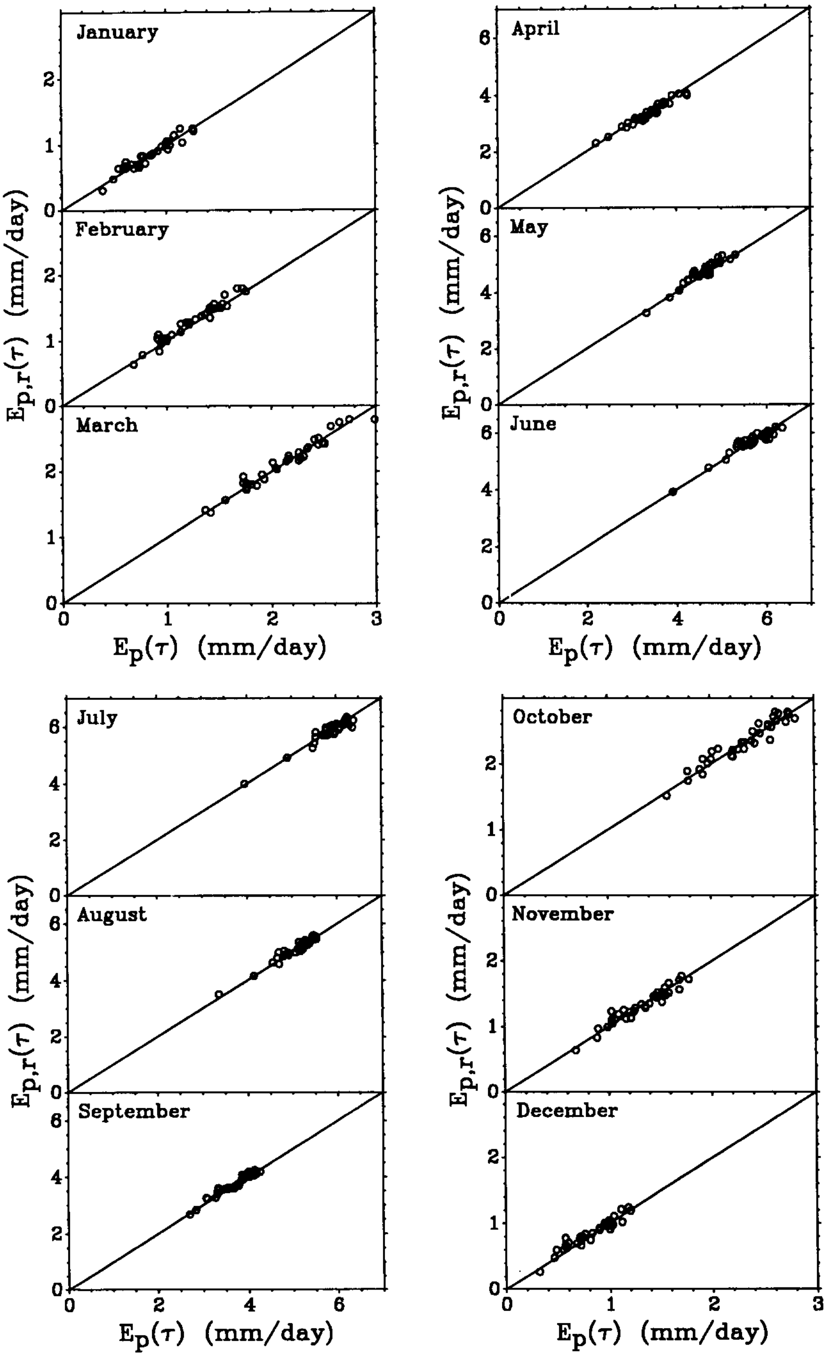

Fig. 4. Comparison between modified Penman estimated $E_p(\tau)$ and regional regression model $E_{t,p}(\tau)$ potential evaporation by month.

Table 7
A comparison of monthly BIAS and r.m.s.e. for the two estimators $E_{p,r}(\tau)$ and $E_{p,L}(\tau)$

τ	Month	BIAS[w]	
		$w = E_{p,r}(\tau)$ (mm day^{-1})	$w = E_{p,L}(\tau)$ (mm day^{-1})
1	January	0.00	−0.63
2	February	0.04	−0.84
3	March	0.00	−0.68
4	April	−0.06	−0.10
5	May	0.10	0.58
6	June	0.02	1.39
7	July	−0.04	2.28
8	August	0.02	2.69
9	September	0.05	2.48
10	October	−0.02	1.72
11	November	−0.01	0.81
12	December	0.02	−0.21
τ	Month	r.m.s.e.[w]	
		$w = E_{p,r}(\tau)$ (mm day^{-1})	$w = E_{p,L}(\tau)$ (mm day^{-1})
1	January	0.08	0.67
2	February	0.08	0.88
3	March	0.08	0.77
4	April	0.12	0.41
5	May	0.16	0.70
6	June	0.16	1.42
7	July	0.17	2.28
8	August	0.13	2.68
9	September	0.12	2.47
10	October	0.10	1.72
11	November	0.08	0.84
12	December	0.08	0.36

Similar to our previous goodness-of-fit evaluation, 33 separate sets of the five regional regression equations for E_{pa}, f_1, f_2, g_1 and g_2 were developed to more fairly estimate the BIAS and r.m.s.e. between $E_p(\tau)$ and $E_{p,r}(\tau)$. Each regression estimator in this evaluation for a particular site, $E_{p,r}(\tau)$, is based on temperature, longitude and elevation data from the 33 other sites.

Table 7 documents the BIAS and r.m.s.e. for each method—$E_{p,r}(\tau)$ and $E_{t,L}(\tau)$. Clearly, the regression model $E_{p,r}(\tau)$ provides the lowest overall BIAS and r.m.s.e. Practically speaking, even the largest single monthly BIAS of 0.09 mm day^{-1} would be difficult to measure.

6. Summary and conclusions

This study developed regional statistical models of estimated monthly average reference crop evapotranspiration, $E_t(\tau)$, where $\tau = 1,12$, and estimated monthly evaporation,

$E_p(\tau)$ based upon the Penman–Monteith method. $E_t(\tau)$ is determined for a reference grass vegetation. Estimates for $E_t(\tau)$ and $E_p(\tau)$ derived from the Penman–Monteith equation require air temperature, relative humidity or dewpoint temperature, windspeed and cloud cover, data which are generally available only at sparsely located NOAA First Order weather observatories. The regional estimators presented here, $E_{t,r}(\tau)$ and $E_{p,r}(\tau)$, require only monthly average air temperature, and the site's longitude and elevation. The models are shown to provide excellent estimates of $E_t(\tau)$ and $E_p(\tau)$ for all locations in the northeastern USA ranging from eastern Ohio through central Virginia to northern Maine, including mountain and coastal areas. $E_{p,r}(\tau)$ is shown to be a significant improvement over the Linacre (1977) temperature method to estimate $E_p(\tau)$. $E_{t,r}(\tau)$ is shown be an improvement over both the Linacre (1977) and Hargreaves and Samani (1985) temperature methods to estimate $E_t(\tau)$ in the northeastern USA.

Acknowledgements

This work was funded in part by the Massachusetts Department of Environmental Protection, Office of Watershed Management and the Division of Water Supply, and in part by the US Environmental Protection Agency Assistance Agreement CR820301 to the Tufts University Center for Environmental Management. The authors express their appreciation for the very helpful review comments by Dr. Dara Entekhabi and three anonymous reviewers, who all contributed to substantially improving the original manuscript.

References

Bloomfield, P., 1976. Fourier Analysis of Time Series: An Introduction. Wiley, New York.

Bras, R.L., 1991. Hydrology; an Introduction to Hydrologic Science. Addison–Wesley, Reading, MA.

Brest, C.L., 1987. Seasonal albedo of an urban/rural landscape from satellite observations. J. Climate Appl. Meteorol., 26(9): 1169–1187.

Brutsaert, W., 1984. Evaporation into the Atmosphere: Theory, History and Applications. D. Reidel, Boston, MA.

Eagleson, P.S., 1970. Dynamic Hydrology. McGraw–Hill, New York.

Fennessey, N.M., 1994. A hydro-climatological model of daily streamflow for the northeast United States. Ph.D. Dissertation, Tufts University, Medford, MA.

Fennessey, N.M. and Kirshen, P.H., 1994. Evaporation and evapotranspiration under climate change in New England. ASCE J. Water Resour. Plann. Manage., 120(1): 48–69.

Hargreaves, G.H., 1994. Defining and using reference evapotranspiration. ASCE J. Irrig. Drainage Eng., 120(6): 1132–1139.

Hargreaves, G.H. and Samani, Z.A., 1985. Reference crop evapotranspiration. Appl. Eng. Agric., 1(2): 96–99.

Hirsch, R.M., Helsel, D.R., Cohn, T.A. and Gilroy, E.J., 1993. Statistical analysis of hydrologic data. In: D.R. Maidment (Editor), Handbook of Hydrology. McGraw–Hill, New York.

Jensen, M.E., Burman, R.D. and Allen, R.G., 1990. Evapotranspiration and irrigation water requirements. ASCE Manual 70. ASCE, New York, 332 pp.

Linacre, E.T., 1977. A simple formula for estimating evaporation rates in various climates, using temperature data alone. Agric. Meteorol., 18: 409–424.

Monteith, J.L., 1965. Evaporation and the environment. The state and movement of water in living organisms. Symp. Soc. Exp. Biol., XIX.

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. Proc. R. Soc. London, Ser. A, 193: 120–148.

Shuttleworth, W.J., 1993. Evaporation. In: D.R. Maidment (Editor), Handbook of Hydrology. McGraw–Hill, New York.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: D.R. Maidment (Editor), Handbook of Hydrology. McGraw–Hill, New York.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Am. Geogr. Rev., 38: 55–94.

Van Bavel, C.H.M., 1966. Potential evaporation: the combination concept and its experimental verification. Water Resour. Res., 2(3): 455–467.