Hydrology in a Changing World: Environmental and Human Dimensions
Proceedings of FRIEND-Water 2014, Montpellier, France, October 2014 (IAHS Publ. 363, 2014).

367

Standardized precipitation-evapotranspiration index (SPEI): Sensitivity to potential evapotranspiration model and parameters

JAMES H. STAGGE[1], LENA M. TALLAKSEN[1], CHONG-YU XU[1] & HENNY A. J. VAN LANEN[2]

1 Department of Geosciences, University of Oslo, PO Box 1047 Blindern, N-0316 Oslo, Norway
j.h.stagge@geo.uio.no
2 Hydrology and Quantitative Water Management Group, Wageningen University, the Netherlands

Abstract The Standardized Precipitation-Evapotranspiration Index (SPEI), a variant of the WMO-recommended Standardized Precipitation Index (SPI), has significant potential as a meteorological drought index because it uses a more comprehensive measure of water availability, climatic water balance. However, inclusion of PET, a derived term, requires rigorous testing before the index gains wide acceptance. This study addresses whether the SPEI differs significantly from the SPI and tests its sensitivity to the choice of PET method by first comparing derived PET and then SPEI/SPI across 3950 gridded land cells in Europe using five commonly used PET methods with different complexity and input requirements. The SPEI was found to differ significantly from the SPI and the resulting PET and SPEI values found to group according to the PET radiation term. The mass transfer term, which integrates wind speed and humidity/pressure, was found to have a secondary effect on PET and no detectable effect on SPEI.

Key words drought index; sensitivity; drought; meteorological drought; potential evapotranspiration; standardized precipitation index; SPI; SPEI

INTRODUCTION

Drought indices are vital to objectively quantify and compare drought severity, duration, and extent across regions with varied climatic and hydrologic regimes. The Standardized Precipitation Index (SPI), outlined in McKee *et al.* (1993) and Guttman (1999), measures normalized anomalies in precipitation and has been recommended as a key drought indictor by the World Meteorological Organization (WMO, 2006) and a universal meteorological drought index by the Lincoln Declaration on Drought (Hayes *et al.*, 2011). In this way, accumulated precipitation can be compared objectively across locations with different climatology and highly non-normal precipitation distributions.

Despite this widespread acceptance, the SPI does not account for atmospheric conditions other than precipitation that may affect drought severity, such as temperature, wind speed, and humidity. To address this, the Standardized Precipitation-Evapotranspiration Index (SPEI) was developed by Vicente-Serrano *et al.* (2010). The SPEI utilizes a similar form, but instead normalizes anomalies in accumulated climatic water balance, defined as the difference between precipitation and potential evapotranspiration. This modification retains the simplicity of calculation, multi-temporal nature, and statistical interpretability of the SPI, while providing a more comprehensive measure of water availability that includes a broader measure of climatic conditions.

Inclusion of potential evapotranspiration (PET, McMahon *et al.*, 2013), a derived term, in the SPEI highlights the need for rigorous testing to determine if this inclusion produces significantly different index values than the already accepted SPI and to identify the sensitivity of the SPEI to different PET formulations. Initial validation by Vicente-Serrano *et al.* (2010) focused on 11 points globally and used only the Thornthwaite equation (Thornthwaite, 1948), justifying the choice of PET method by its low input requirements and a study by Mavromatis (2007), which concluded that simple or complex PET methods make little difference when indices are calculated. Vicente-Serrano *et al.* (2010) concluded that little differences exist between the SPEI and SPI using historical data for 11 stations, but found the indices diverge when a progressive temperature increase of 2° or 4°C is added to the historical series, designed to simulate climate change.

This study addresses both whether the SPEI differs significantly from the SPI and how sensitive this index is to choice of PET method by comparing first PET and then SPEI/SPI across 3950 gridded land cells in Europe using five commonly used PET methods.

DATA AND METHODS

Climate data

All climate estimates are based on the Watch Forcing Dataset (WFD), a gridded historical climate dataset based on ERA-40 reanalysis with $0.5° \times 0.5°$ resolution (Weedon *et al.*, 2010). The WFD consists of sub-daily forcing data spanning the time period 1 January 1958 to 31 December 2001 and employs bias-correction for temperature and precipitation based on CRU-TS2.1 and GPCCv4 observations. Other climate variables, which include wind speed (10 m), surface pressure (10 m), specific humidity, net long-wave surface radiation, and short-wave downwards surface radiation, have not been bias corrected (Haddeland *et al.*, 2011). For the purpose of this research, the European extent is defined as the region between $34°–72°$ N latitude and $–13°–32°$ E longitude.

Potential evapotranspiration equations

Care was taken to select the most commonly used PET methods that span the range of equation types, including temperature-based, radiation-based, and combination equations (Table 1, McMahon, 2013). In order of complexity, these models are the Thornthwaite (Thornthwaite, 1948), Hargreaves (Hargreaves & Samani, 1985), Penman-Montieth with Hargreaves radiation term referred to here as P-M (Hargreaves) (Allen *et al.*, 1998), Priestley-Taylor (Priestley & Taylor, 1972), and FAO-56 Penman-Montieth referred to here as P-M (FAO-56) (Allen *et al.*, 1998). The FAO-56 reference crop definition (Allen *et al.*, 1998) is used for all PET models except for the Thornthwaite equation, which uses the original Thornthwaite (1948) definition. The term "PET" is used across all models to maintain consistency with the proposed definition of SPEI (Vicente-Serrano *et al.*, 2010).

Of these methods, the Thornthwaite is unique because it is based on an empirical relationship between average monthly temperature and potential evapotranspiration. Despite its differences and noted limitations (Jensen, 1973; Amatya *et al.*, 1995), the Thornthwaite method was included because it is commonly used and was cited in the original SPEI methodology (Vicente-Serrano *et al.*, 2010). The remainder of the PET methods follow the general form:

$$PET = \frac{\Delta \cdot R_n + \gamma \cdot \text{"mass transfer term"}}{\Delta + \gamma} \tag{1}$$

where R_n represents net radiation, γ is the psychrometric constant, and Δ is the slope of the saturation-vapour-pressure *vs* temperature curve at the given air temperature. Hargreaves equation uses the daily difference between T_{max} and T_{min} as a proxy to estimate net radiation (Hargreaves & Samani, 1985), and simplifies the mass transfer term with a constant. The P-M (Hargreaves) equation uses an identical radiation estimate, but evaluates mass transfer using wind speed from the WFD. The Priestley-Taylor and P-M (FAO56) both use net radiation directly from the WFD, though the Priestley Taylor simplifies mass transfer with a constant and the full Penman-Monteith model (P-M FAO56) uses WFD wind speed and atmospheric conditions to calculate this term.

Table 1 Summary of potential evapotranspiration equations.

PET Groups	Group 1 Empirical	Group 2 Temp-Proxy Radiation		Group 3 Observed Radiation	
PET Model	Thornthwaite	Hargreaves	P-M (Hargreaves)	Priestley- Taylor	P-M (FAO56)
Mean temp	X	X	X	X	X
Min/Max temp		X	X		
Wind speed			X		X
Surface pressure					X
Specific humidity					X
Radiation	T_{mean}	$T_{max}–T_{min}$	$T_{max}–T_{min}$	WFD	WFD
Reference	Thornthwaite (1948)	Hargreaves-Samani (1985)	Allen *et al.* (1998)	Priestley-Taylor (1972)	Allen *et al.* (1998)

PET was estimated at the daily time step for all methods and summed to create monthly potential evapotranspiration except for the Thornthwaite method, which calculated monthly PET directly.

SPI and SPEI Calculation

SPI is computed by summing precipitation over *n* months (or days), where *n* is typically 1, 2, 3, 6, 9, 12, or 24 months, and fitting these accumulated precipitation values to a parametric statistical distribution from which non-exceedence probabilities are transformed to the standard normal distribution ($\mu = 0$, $\sigma = 1$) (McKee *et al.*, 1993; Guttman, 1999; Lloyd-Hughes & Saunders, 2002). SPEI is calculated in a similar fashion, but instead sums climatic water balance, defined as the difference between precipitation and PET (Vicente-Serrano *et al.*, 2010).

For this study, SPI and SPEI were calculated at the monthly resolution, using a 6-month accumulation period. The 6-month period was chosen because it represents a seasonal estimate of drought. Additional accumulation periods may be evaluated as part of future studies. Index values were fitted by maximum likelihood estimation and normalized using the two parameter gamma distribution for the SPI and generalized extreme value (GEV) distribution for SPEI, following the recommendations outlined in Stagge *et al.* (2013). Also, as in Stagge *et al.* (2013), index values are limited to the range between –3 and 3 to ensure reasonableness. All normalization was performed relative to the reference period 1970–1999, in accordance with WMO standard reference periods.
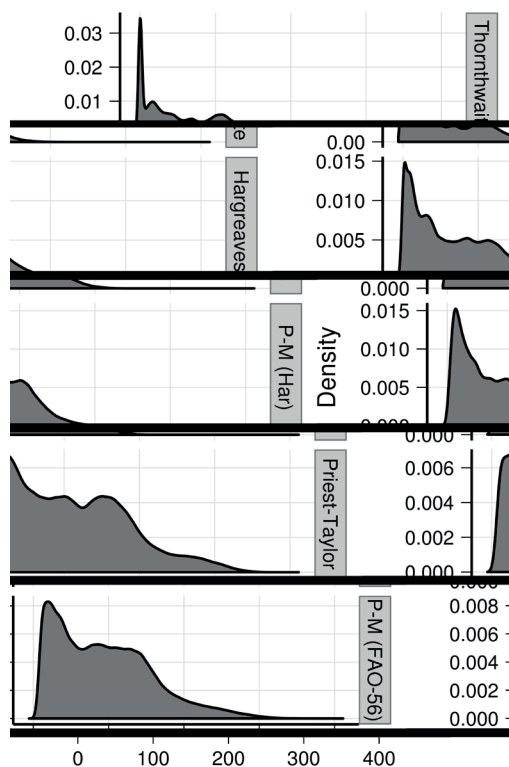

Fig. 1 Distribution of monthly potential evapotranspiration across PET equations. Equations ordered by increasing complexity.

Paired Wilcoxon rank-sum statistic

Paired testing is used to determine systematic bias between PET and SPI/SPEI values at each location and time step. Because the distribution of PET differences ($\Delta_{PET1-PET2}$) is highly skewed, the paired Wilcoxon rank-sum test is used as a nonparametric alternative to the paired t-test to determine significant differences between the PET methods. This test ranks paired differences, $\Delta_{PET1-PET2}$, and compares the sum of rank orders, incorporating magnitude as well as sign.

RESULTS

Potential evapotranspiration comparison

Comparison of PET estimates form three distinct PET groups: (1) the Thornthwaite equation, (2) Hargreaves and P-M (Hargreaves), and (3) the Priestley-Taylor and P-M (FAO-56) (Fig. 1). This demonstrates the importance of the radiation term in equation (1) relative to mass transfer, as the groups correspond to different radiation forms, namely (1) empirical radiation, (2) temperature-proxy radiation, and (3) observed radiation. Overall, PET values follow the order: observed radiation > temperature-proxy radiation > empirical radiation. This general order is consistent with prior studies on PET magnitude (Xu & Singh, 2001; Weiß & Menzel, 2008). The sharp peak at zero evapotranspiration in the Thornthwaite equation (Fig. 1) is caused by the assumption of zero evapotranspiration at temperatures below freezing.
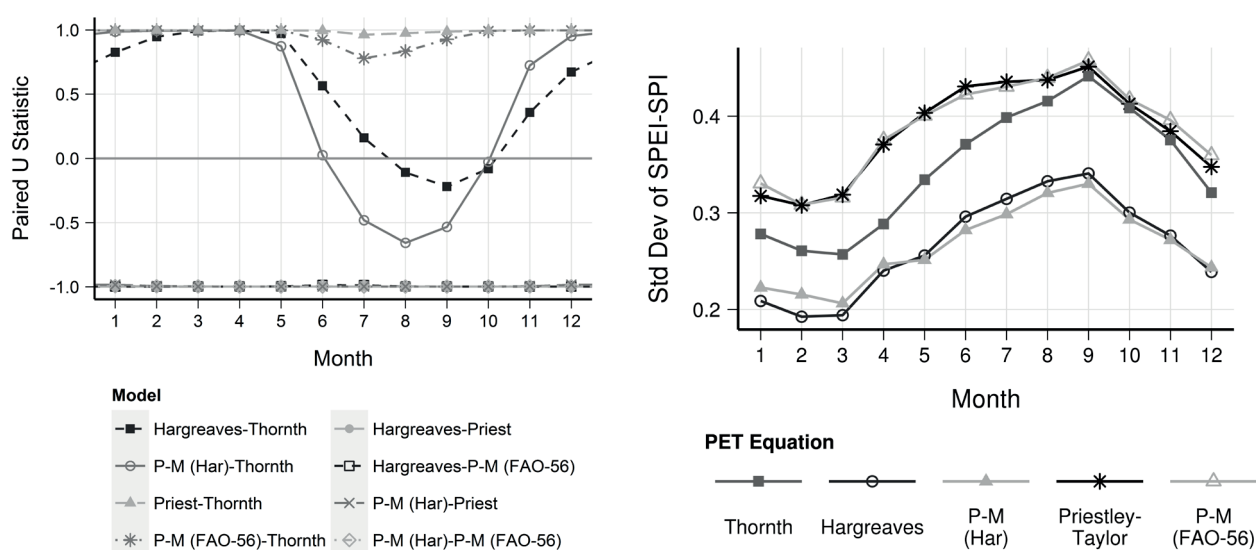

Fig. 2 Seasonality of the paired $\Delta_{PET1-PET2}$ Wilcoxon rank sum statistic showing the difference between different radiation terms (left). Standard deviation of $\Delta_{SPEI-SPI}$ for the candidate PET methods (right).

Radiation term

When compared pairwise at each location and time step, simulated PET values maintain the relative order based on the radiation term (not shown). PET calculated using observed radiation is consistently higher than PET using temperature-proxy radiation throughout the year and may be explained by the lack of radiation bias correction within the WFD (Haddeland *et al.*, 2011).

The largest difference between the PET groups occurs during the winter, corresponding to consistent paired Wilcoxon U-statistic values of 1 or –1 (Fig. 2). In particular, the Thornthwaite equation predicts significantly lower PET throughout the winter, which is partially explained by a default of zero PET below freezing temperatures. Agreement is highest during the summer when evapotranspiration is highest (Fig. 2), which is reasonable given the focus on the agricultural growth period in PET model development. In particular, between July and October, the Thornthwaite equation predicts higher PET values than temperature-proxy radiation methods (Fig. 2), particularly in colder regions of northern Europe and the Alps, supporting findings that the Thornthwaite equation overestimates PET in the summer and at high altitudes (Amatya *et al.*, 1995).

Mass transfer term

Although the radiation term appears to be the most important contributing factor to differences in PET, there are also within group differences, suggesting the secondary importance of mass transfer. Comparing P-M (FAO-56) with the Priestley-Taylor equation and P-M (Hargreaves) with the Hargreaves equation, the simpler model with a mass transfer approximation produces higher

PET during the summer, while the more complex model produces higher PET during the winter (not shown). This shift is most noticeable among the Group 2 (temperature-proxy radiation) methods and occurs predominantly in coastal regions. During winter, when radiation is lowest, wind speeds and humidity play a larger role in evapotranspiration, particularly where wind speeds are high (Allen *et al.*, 1998; Jabloun & Sahli, 2008).

SPEI comparison

It is important to compare the SPEI with SPI to determine whether inclusion of the PET term produces significantly different index values. Paired analysis shows there is no consistent bias between the indices and that the difference, $\Delta_{SPEI-SPI}$, is normally distributed. This is reasonable, as both indices are normalized to the standard normal distribution ($\mu = 0$, $\sigma = 1$). Because $\Delta_{SPEI-SPI}$ is normally distributed, the standard deviation of $\Delta_{SPEI-SPI}$ is used to quantify the variability between SPEI and SPI index values. When plotted temporally (Fig. 2), the standard deviation shows that the models again group according to their radiation term, with the differences between SPEI and SPI remaining constant throughout the year, following the order: observed radiation based methods > Thornthwaite equation > temperature-proxy radiation (Group 3 > Group 1 > Group 2).

Differences between the SPEI-6 and SPI-6 are smallest in February and March (Fig. 2), which correspond to the periods between September–February and October–March, respectively. Similarity between SPEI and SPI during the winter is reasonable, as PET tends to be lowest relative to precipitation during this period, causing SPEI to functionally approach the SPI. By the same principles, the greatest difference occurs in September, which spans the period April–September when PET is highest relative to precipitation. Interestingly, the Thornthwaite equation, which produces the lowest PET values most of the year, generates SPEI values that differ from the SPI more than the temperature-proxy methods (Fig. 2). It is particularly notable that this difference continues through the winter, when the Thornthwaite equation generates zero evapotranspiration in northern latitudes.

Spatially, the differences between the SPEI and SPI are consistent for all PET methods, with the relative magnitude following the order identified above. Correlation of SPEI with SPI is in the range of r = 0.65–0.98, slightly lower, but similar to results from Vicente-Serrano *et al.* (2010), with the greatest differences occurring in three regions: the arid/semi-arid south, the polar north, and a region in eastern Germany/western Poland. Differences between SPEI and SPI at the climatic extremes of Europe are reasonable, considering the relative importance of evapotranspiration in the south and difficulties in modelling evapotranspiration (freezing temperatures and high interannual differences in radiation) for high latitudes. The anomalous region in central Europe warrants further attention in future studies.
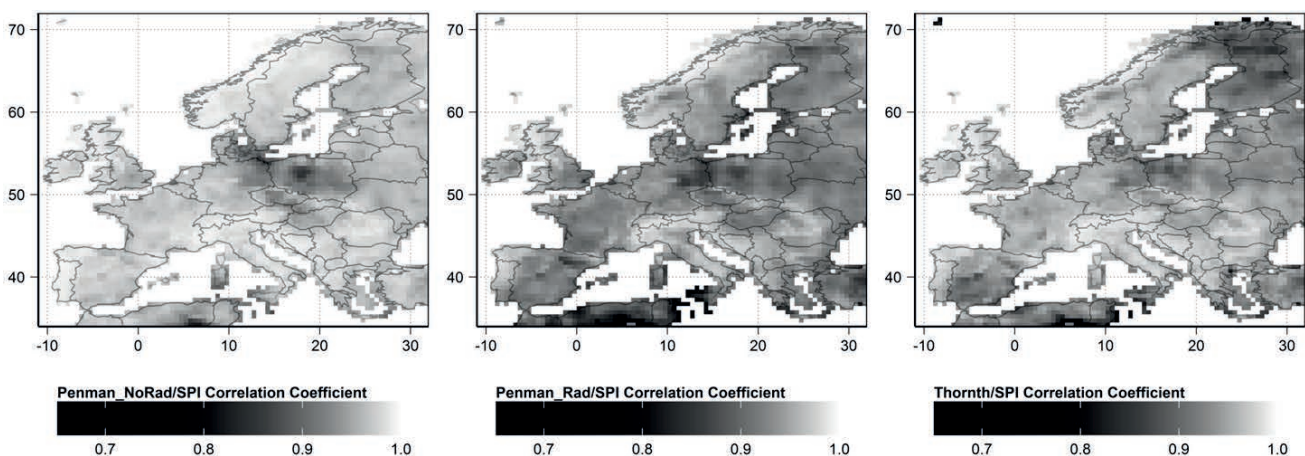

Fig. 3 Temporal correlation of SPEI with SPI. SPEI calculated using P-M (Hargreaves) (left), P-M (FAO-56) (centre) and Thornthwaite equation (right).

The largest difference between SPEI-6 calculated using the five PET equations occurs during the winter and spring, whereas the best agreement occurs during the summer (not shown). As in PET calculations, the Thornthwaite equation stands apart from the remaining equations, particularly during the winter, where it underestimates PET, resulting in higher SPEI values. For the remaining four models, as in the comparison with SPI, there is little consistent bias. However, the clear groupings related to PET radiation (Fig. 3) suggest that there remains a difference between the models with respect to SPEI. Rather than differences in central behaviour, quantified by bias, these differences may be related to extreme SPEI, which would explain the notable differences in standard deviation. More research is needed to evaluate SPEI differences at the extremes.

CONCLUSIONS

The SPEI is an important and useful tool for comparing meteorological drought. However, prior to its widespread acceptance within the hydrologic community, its sensitivity to PET parameters must be identified. This paper concludes that inclusion of PET makes a discernible difference in index values, confirming that SPEI provides a significantly different drought index to the SPI. The greatest differences between SPEI and SPI occur during the summer when PET represents a greater portion of the climatic water balance, and at the climatic extremes of Europe in northern Scandinavia and around the Mediterranean. A small region in central Europe was also detected with poor correlations, warranting future attention.

The SPEI is most sensitive to the PET radiation term, which separates the observed radiation methods (P-M FAO-56 and Priestley-Taylor) from the temperature-proxy methods (P-M Hargreaves and Hargreaves) and the Thornthwaite equation. In Europe, using the WFD, the observed radiation methods consistently predict the highest PET followed by the temperature-proxy and empirical methods. The mass transfer term, calculated in both Penman-Montieth equations, can make a minor difference in PET, most notably along coastlines with high advection, although this difference was not detected in SPEI comparisons.

Given these findings, the SPEI is recommended as an alternative to SPI to quantify anomalies in accumulated climatic water balance, incorporating potential (reference) evapotranspiration. Of the available PET equations, SPEI calculated using the temperature-proxy or radiation based methods are more internally consistent than with the Thornthwaite equation, which is often significantly different from the other four methods. If data permits, the Hargreaves and the P-M (Hargreaves) equations strike a useful balance between consistency and minimal data requirements, requiring only the addition of minimum/maximum temperature and wind speed, in the case of the P-M (Hargreaves).

REFERENCES

Allen, R., Pereira, L., Raes, D. & Smith, M. (1998) Crop evapotranspiration. *FAO irrigation and drainage paper 56*. FAO, Rome, Italy, 10.

Amatya, D., Skaggs, R. & Gregory, J. (1995) Comparison of methods for estimating REF-ET. *J. Irrig. Drain. Div. ASCE*, 121(6): 427–435.

Guttman, N. B. (1999) Accepting the Standardized Precipitation Index: A Calculation Algorithm. *J. Am. Water Resour. Assoc.* 35(2), 311–322.

Haddeland, I., *et al.* (2011) Effects of climate model radiation, humidity and wind estimates on hydrological simulations. *Hydrol. Earth Syst. Sci. Discuss.* 8(4), 7919–7945.

Hargreaves, G. H. & Samani, Z. A. (1985) Reference crop evapotranspiration from ambient air temperature. *American Society of Agricultural Engineers*, 96–99.

Hayes, M., Svoboda, M., Wall, N. & Widhalm, M. (2011) The Lincoln declaration on drought indices: universal meteorological drought index recommended. *Bulletin of the American Meteorological Society* 92(4), 485–488.

Jabloun, M. & Sahli, A. (2008) Evaluation of FAO-56 methodology for estimating reference evapotranspiration using limited climatic data: Application to Tunisia. *Agricultural Water Management* 95(6), 707–715.

Jensen, M. E. (1973) *Consumptive Use of Water and Irrigation Water Requirements*. New York.

Lloyd-Hughes, B. & Saunders, M. A. (2002) A drought climatology for Europe. *Int. J. Climatol*, 22(13), 1571–1592.

Mavromatis, T. (2007) Drought index evaluation for assessing future wheat production in Greece. *Int. J. Climatol*, 27, 911–924.

McKee, T. B., Doesken, N. J. & Kleist, J. (1993) The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*. American Meteorological Society Boston, MA, pp. 179–183.

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R. & McVicar, T. R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrol. Earth Syst. Sci.* 17, 1331–1363.

Priestley, C. H. B. & Taylor, R .J. (1972) On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* 100(2), 81–92.

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., van Looon, A .F. & Stahl, K. (2013) Pan-European comparison of candidate distributions for climatological drought indices (SPI and SPEI) *Journal of Climate* (submitted).

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical review* 38(1), 55–94.

Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. (2010) A multi-scalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index – SPEI. *J. Climate* 23, 1696–1718.

Weedon, G. P. *et al.* (2010) The WATCH Forcing Data 1958–2001: A Meteorological Forcing Dataset for Land Surface and Hydrological Models, WATCH.

Weiß, M. & Menzel, L. (2008) A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments. *Advances in Geosciences* 18(18), 15–23.

WMO (2006) Drought monitoring and early warning: Concepts, progress and future challenges. *WMO*.

Xu, C. Y. & Singh, V. (2001) Evaluation and generalization of temperature‐based methods for calculating evaporation. *Hydrological Processes* 15(2), 305–319.