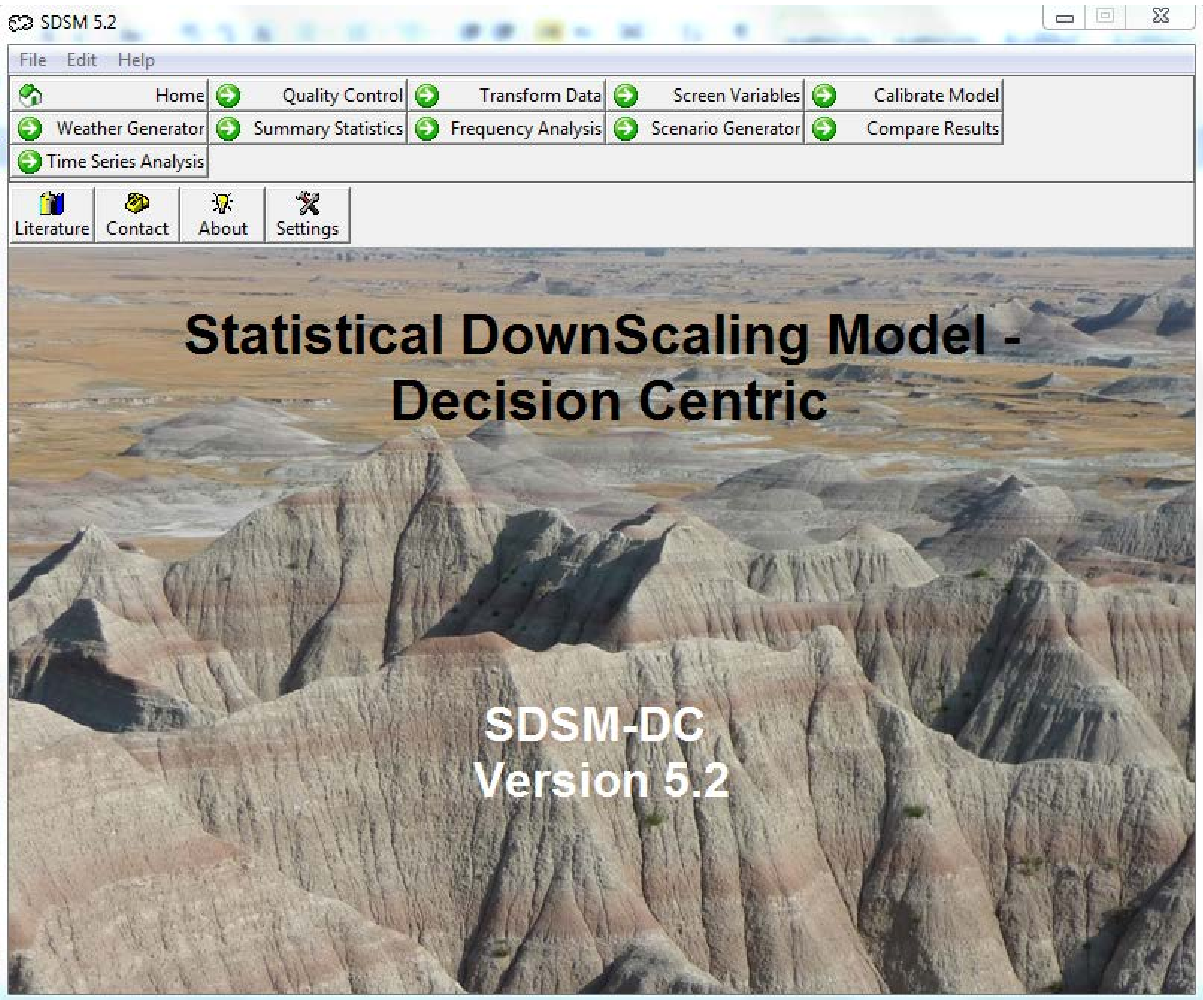
Statistical DownScaling Model – Decision Centric (SDSM-DC) Version 5.2 Supplementary Note

6 March 2015

Robert L. Wilby[1] and Christian W. Dawson[2]

[1] *Department of Geography, Loughborough University, Loughborough, LE11 3TU, UK*

[2] *Computer Sciences Department, Loughborough University, Loughborough, LE11 3TU, UK*



1. Introduction

This note is to advise Users of two new developments: 1) upgrade of SDSM-DC to enable more stringent testing and cross-validation of downscaling model skill; and 2) creation of a data portal and daily downscaling predictor set for the period 1948-2014. Further details and example applications of these new tools are provided below.

2. SDSM-DC 5.2

Likely non-stationarity of predictor-predictand relationships has long been recognised as a limitation of all downscaling techniques (Wilby, 1997; Schmith, 2008). This is because atmospheric circulation, moisture and temperature properties may not be the only factors determining local weather conditions. For instance, changes in land properties due to rapid urbanisation may be just as consequential to near surface temperature trends as regional airflow patterns. Likewise, any abrupt or structural changes in the downscaling relationship (for example due to relocation of a meteorological station, or enforcement of air quality standards in the case of ozone downscaling) can also be problematic (Estrada et al., 2013).

Although there is no means of determining accuracy of downscaling relationships for future periods, it is possible to assess model stability *within* observed data by means of cross-validation. The well-known *k*-folds cross-validation technique involves dividing data into *k* equal-size subsamples and sequentially fitting to one sub-sample and testing against the remainder (e.g. Bedia et al., 2013; Casanueva et al., 2014). The simplest *k*-fold validation is when *k*=2. This is the conventional split-sample test when one half of the data is used to calibrate (e.g. 1961-1980) and the other to test the model (e.g. 1981-2000). The two halves can then be swapped so the latter is next used to calibrate (i.e. 1981-2000) and the former to validate (i.e. 1961-1980).

The lower the value of *k* the smaller the sub-sample of data used to calibrate the model. The corollary is that a larger portion of the information content is not used to estimate model parameters. This can be overcome by using a higher value of *k*. For example, if *k*=10 then 90% of the data are used to calibrate the model whilst the other 10% is reserved for comparison with model predictions. The process is repeated 10 times until every 10% block of data has been predicted. The final step is to concatenate all predicted sub-samples then compare with observations to obtain an overall measure of model performance when tested against data not used for model calibration.

The *k*-folds cross-validation technique has been automated in SDSM-DC 5.2. Within the *Calibrate Model* screen it is possible to click the Cross Validation box and select *k* in the range 2 to 10. This divides available data (excluding missing values) within the chosen date range into *k* sub-samples. Figure 1 shows the screen when a 40-year (1961-2000) daily mean temperature series is downscaled using six predictor variables calibrated and cross-validated using *k*=10 sub-samples.

Three measures of calibration fit and validation skill are provided: the fraction of explained variance (Pearson RSquared); the standard error of the estimate (SE); and the Durbin-Watson (to detect presence of autocorrelation in model residuals, where a value near 2 indicates no autocorrelation, 0 and 4 are positive and negative autocorrelation respectively). In addition, Spearman rank correlation and the bias are given for the cross-validated model. Spearman is helpful when data are non-normal; bias is simply the difference between the observed and expected means (in native units).

The *Calibration Results* for the example show modest reductions in explained variance when comparing calibration fit (to all data) versus cross-validated skill (Figure 2). Similarly, the standard error and Durbin-Watson metrics are only weakly affected (the latter showing test values suggesting insignificant positive autocorrelation in residuals). As expected, a model tested using a split-sample (*k*=2) yields poorer metrics than one cross-validated with a larger number of sub-samples (*k*=10) (Figure 3a). The model error grows when there are more missing values, the data are of poorer quality, and/or overall predictability is lower, as in the case of daily wet-day amounts (Figure 3b).

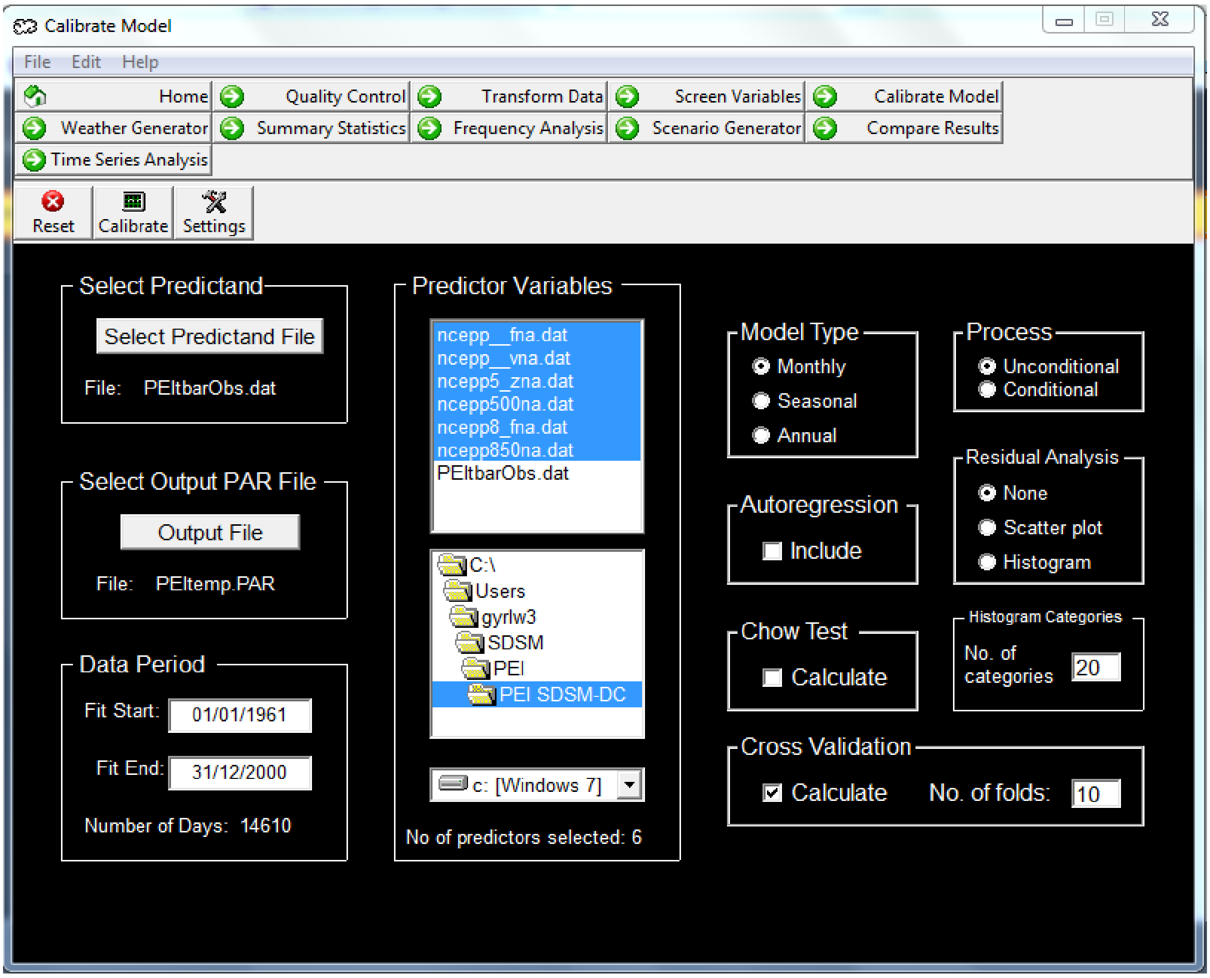
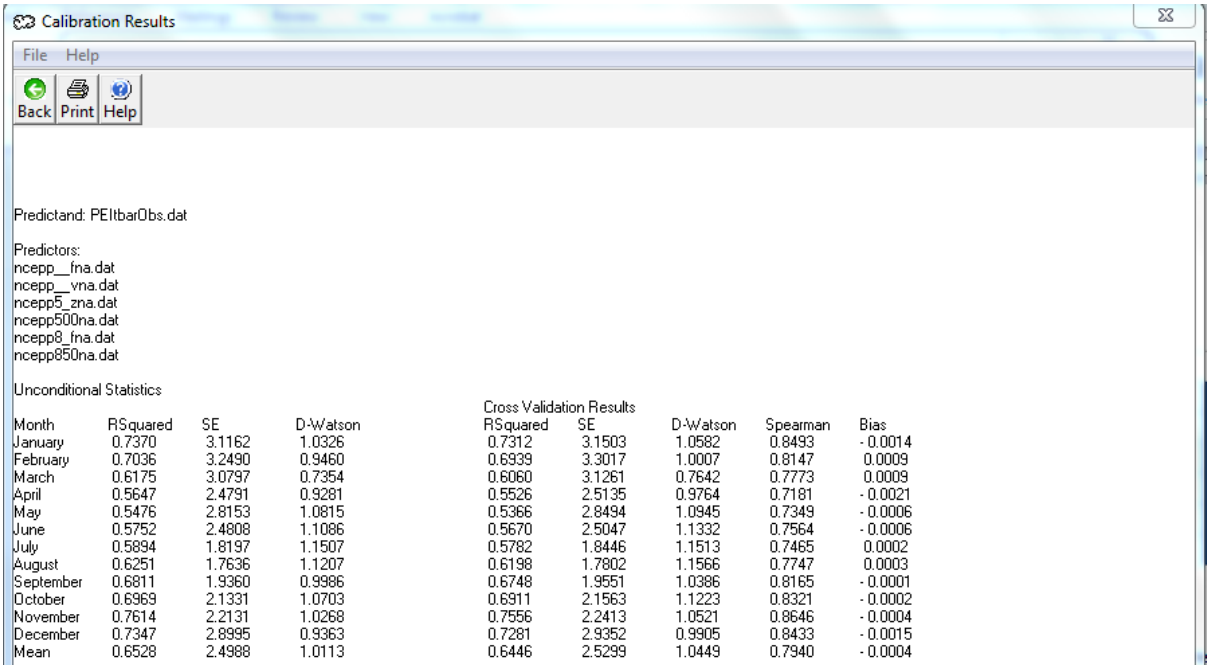


Figure 1 SDSM-DC 5.2 *Calibrate Model* screen showing automated cross-validation option (lower right corner).



Predictand: PEltbarObs.dat

Predictors:
ncepp__fna.dat
ncepp__vna.dat
ncepp5_zna.dat
ncepp500na.dat
ncepp8_fna.dat
ncepp850na.dat

Unconditional Statistics

				Cross Validation Results				
Month	RSquared	SE	D-Watson	RSquared	SE	D-Watson	Spearman	Bias
January	0.7370	3.1162	1.0326	0.7312	3.1503	1.0582	0.8493	-0.0014
February	0.7036	3.2490	0.9460	0.6939	3.3017	1.0007	0.8147	0.0009
March	0.6175	3.0797	0.7354	0.6060	3.1261	0.7642	0.7773	0.0009
April	0.5647	2.4791	0.9281	0.5526	2.5135	0.9764	0.7181	-0.0021
May	0.5476	2.8153	1.0815	0.5366	2.8494	1.0945	0.7349	-0.0006
June	0.5752	2.4808	1.1086	0.5670	2.5047	1.1332	0.7564	-0.0006
July	0.5894	1.8197	1.1507	0.5782	1.8446	1.1513	0.7465	0.0002
August	0.6251	1.7636	1.1207	0.6198	1.7802	1.1566	0.7747	0.0003
September	0.6811	1.9360	0.9986	0.6748	1.9551	1.0386	0.8165	-0.0001
October	0.6969	2.1331	1.0703	0.6911	2.1563	1.1223	0.8321	-0.0002
November	0.7614	2.2131	1.0268	0.7556	2.2413	1.0521	0.8646	-0.0004
December	0.7347	2.8995	0.9363	0.7281	2.9352	0.9905	0.8433	-0.0015
Mean	0.6528	2.4988	1.0113	0.6446	2.5299	1.0449	0.7940	-0.0004

Figure 2 SDSM-DC 5.2 *Calibrate Results* screen showing skill metrics when an unconditional model is fit against all data (three left hand columns) compared with k=10 cross-validation sub-samples (right columns).

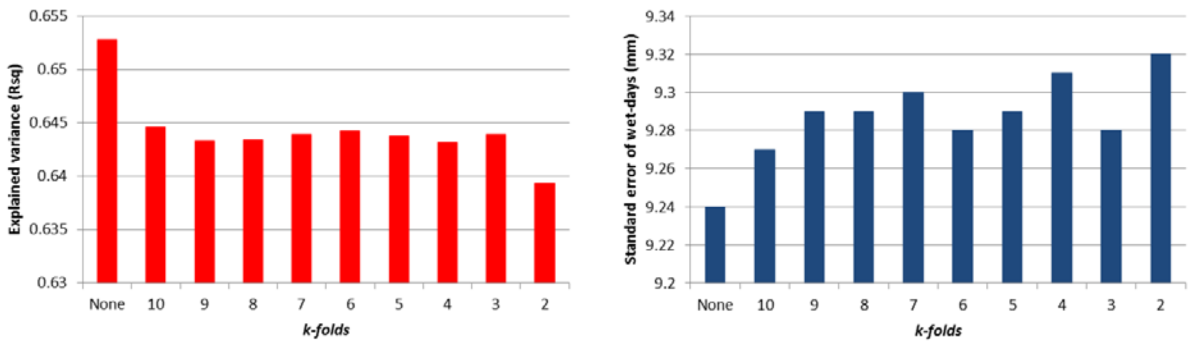


Figure 3 Cross-validation (with *k*-folds) compared with calibration against the full data (no folds) for daily mean temperature with <1% missing values (left) and daily wet-day amounts with 36% missing values (right).

3. Downscaling predictor portal

Users have previously accessed daily downscaling predictors from a web-portal maintained by Environment Canada. This service was discontinued in 2014.

To facilitate convenient, global usage of SDSM-DC, a new data portal has been incorporated within the SDSM web-site (Figure 4): http://co-public.lboro.ac.uk/cocwd/SDSM/data.html

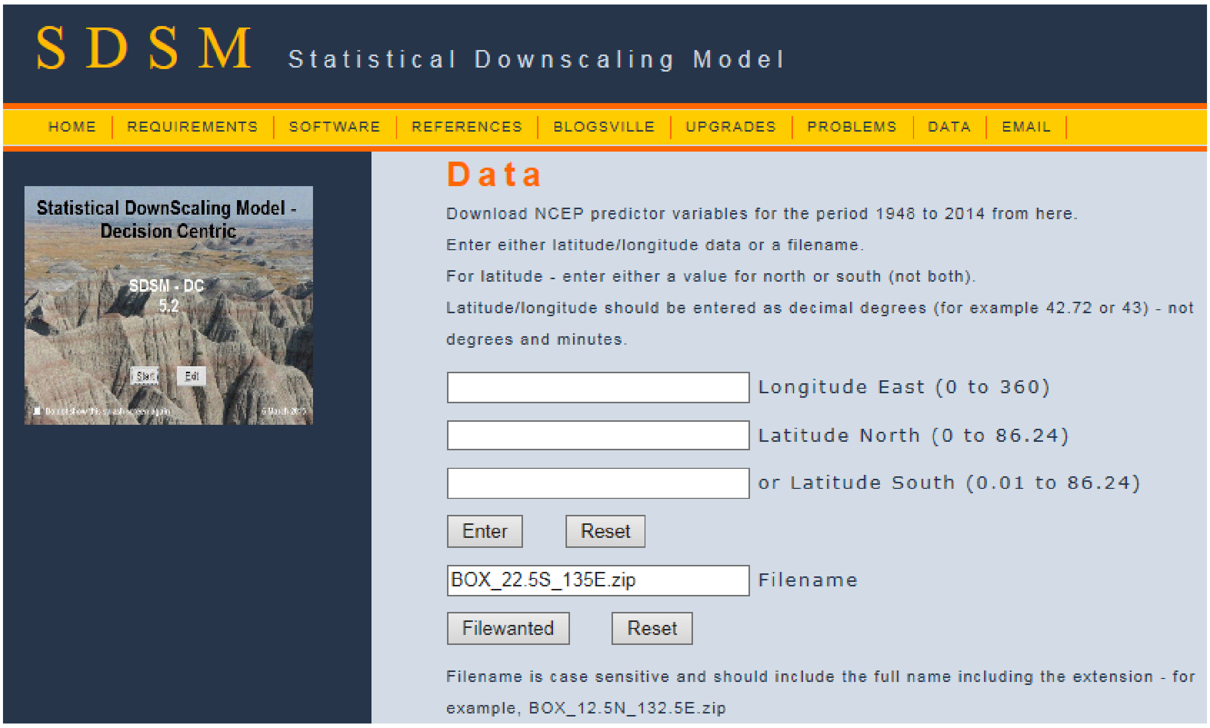


Figure 4 The SDSM predictor variable portal.

Users can search for data by entering the longitude and latitude (decimal degrees) of the study site(s) or by entering a known grid-box, following the file nomenclature (e.g. BOX_12.5N_132.5E.zip). All daily predictor variables are delivered as individual, single column ASCII files, bundled by 2.5° x 2.5° grid-box, and zipped. Predictor variables are available globally except for the poles.

All 27 predictor variables originate from the National Center for Environmental Prediction (NCEP) re-analysis (Kalnay et al., 1996). These raw and derived variables describe atmospheric circulation, thickness, stability and moisture content at various levels (near surface, 850 hPa and 500 hPa) using observations assimilated from stations, upper air and satellite measurements between 1948 and 2014 (Table 1). As with the Environment Canada predictor suite all variables (except wind direction and precipitation) are expressed as z-scores using the mean and standard deviation of the baseline period 1961-1990. Precipitation values less than 0.01 mm/day are set to zero.

Daily variable	Code	Daily variable	Code
Precipitation (mm)	prec	Near surface specific humidity	shum
Mean temperature	temp	*Geostrophic airflow velocity*	**_f
Mean sea level pressure	mslp	*Vorticity*	**_z
500 hPa geopotential height	p500	*Zonal velocity component*	**_u
850 hPa geopotential height	p850	*Meridional velocity component*	**_v
Near surface relative humidity	rhum	***Geostrophic wind direction***	**th
Relative humidity at 500 hPa height	r500	*Divergence*	**zh
Relative humidity at 850 hPa height	r850		

Table 1 Daily NCEP predictor variables supplied by the SDSM data portal. **Bold** indicates variables that have **not** been normalised. *Italics* indicate secondary (airflow) variables derived from pressure fields (at near surface, 500 and 850 hPa levels).

The new predictor suite has been quality assured by comparing with the earlier products offered by Environment Canada. Some minor discrepancies are evident because of slight differences in the grid cell sizes: 2.5° latitude x 2.5° longitude (SDSM portal) compared with 2.5° latitude x 3.75° longitude (Environment Canada portal). These differences are most noticeable in wind direction (**th) and divergence (**zh), and found to be greater in data sparse regions (e.g. Central Asia). Note that uncertainty is also introduced due to the method of extrapolating surface pressure to mean sea level pressure and by the representation of orography in the NCEP model. This might be a consideration when downscaling in high altitude (>1000 m) regions (Wang et al., 2013).

The intention is to update the SDSM predictor set annually.

Acknowledgements

Dr Colin Harpham (University of East Anglia) is thanked for assisting with the creation of the new SDSM predictor suite.

A regularly updated bibliography of SDSM studies may be downloaded from here:
http://co-public.lboro.ac.uk/cocwd/SDSM/Bibliography.pdf

References

Bedia, J., Herrera, S., San Martín, D., Koutsias, N. and Gutiérrez, J.M. 2013. Robust projections of Fire Weather Index in the Mediterranean using statistical downscaling. *Climatic Change*, **120**, 229-247.

Casanueva, A., Frías, M.D., Herrera, S., San-Martín, D., Kaminovic, K. and Gutiérrez, J.M. 2014. Statistical downscaling of climate impact indices: testing the direct approach. *Climatic Change*, **127**, 547-560.

Estrada, F., Guerrero, V.M., Gay-García, C. and Martínez-López, B. 2013. A cautionary note on automated downscaling methods for climate change. *Climate Change*, **120**, 263-276.

Kalnay, E., Kanamitsu, M., Kistler, R. *et al.* 1996. The NCEP/NCAR 40–year reanalysis project. *Bulletin of the American Meteorological Society*, **77**, 437–471.

Schmith, T. 2008. Stationarity of regression relationships: Applications to empirical downscaling. *Journal of Climate*, **21**, 4529-4537.

Wang, X.L., Feng, Y., Compo, G.P., Swail, V.R., Zwiers, F.W. and Allan, R.J. 2013. Trends and low frequency variability of extra-tropical cyclone activity in the ensemble of twentieth century reanalysis. *Climate Dynamics*, **40**, 2775-2800.

Wilby, R.L. 1997. Nonstationarity in daily precipitation series: implications for GCM downscaling using atmospheric circulation indices. *International Journal of Climatology*, **17**, 439-454.

Wilby, R.L. and Dawson, C.W. 2012. The Statistical DownScaling Model (SDSM): Insights from one decade of application. *International Journal of Climatology*, **33**, 1707-1719. http://onlinelibrary.wiley.com/doi/10.1002/joc.3544/abstract

Wilby, R.L., Dawson, C.W., Murphy, C., O'Connor, P. and Hawkins, E. 2014. The Statistical DownScaling Model – Decision Centric (SDSM-DC): Conceptual basis and applications. *Climate Research*, **61**, 251-268. http://www.int-res.com/abstracts/cr/v61/n3/p259-276