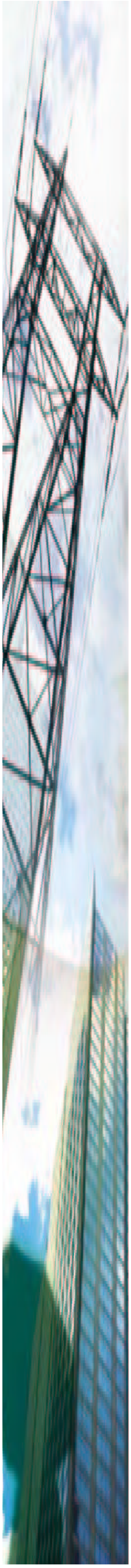


BETWIXT
Built EnvironmenT: Weather scenarios for investigation of Impacts and eXTremes

BETWIXT Technical Briefing Note 4
Version 1, June 2004

VALIDATION OF THE CRU DAILY WEATHER GENERATOR

M. WATTS, C.M. GOODESS* AND P.D. JONES

Climatic Research Unit, University of East Anglia
Norwich, NR4 7TJ

*Corresponding author: c.goodess@uea.ac.uk

1. THE CRU DAILY WEATHER GENERATOR

The CRU daily weather generator was initially developed by Jones and Salmon (1995) and has been substantially modified as part of the BETWIXT project (Watts *et al*., 2004) in order to construct climate scenarios for use in the Building Knowledge for a Changing Climate (BKCC) programme.

Precipitation is the fundamental, primary variable in the weather generator, from which all the other variables are derived using regression relationships or subsequent direct calculation (Table 1). A first-order Markov chain model (Richardson, 1981) is used. A continuous distribution is used for precipitation, making this an infinite state model. Once precipitation has been generated, the secondary variables (minimum and maximum temperature, vapour pressure, wind speed and sunshine duration) are generated. Finally, relative humidity and reference potential evapotranspiration (PET) are calculated from the generated variables. The methods used to generate the primary and secondary variables and to calculate the latter two variables are described in BETWIXT Technical Briefing Note 1 (Watts *et al*., 2004).

Table 1: Weather variables produced by the daily CRU weather generator.

Primary generated variable: Precipitation (mm)
Secondary generated variables: Minimum temperature (degrees C) Maximum temperature (degrees C) Vapour pressure (hPa) Wind speed (ms^{-1}) Sunshine duration (hours)
Calculated variables: Relative humidity (%) Reference potential evapotranspiration (mm day^{-1})

A suitably long (i.e., at least 20 years) daily time series of observed meteorological data must be available in order to calibrate or train the weather generator for each station location, i.e., in order to calculate the weather generator variables for each site. For BETWIXT, 11 such sites have been identified in consultation with BKCC partners (Table 2). Daily data for these station sites were obtained from the British Atmospheric Data Centre and used to calculate the weather generator parameters. These data have also been re-formatted into the same format as the CRU weather generator output and are available for use by academic BKCC partners from the BETWIXT web site.

Table 2: The 11 BETWIXT stations.

	Latitude	*Longitude*	*Years*
Abbotsinch/Paisley	55.86	-4.43	1961-1985
Bradford	53.82	-1.77	1961-1990
Coltishall	52.77	1.35	1963-1980
Elmdon	52.45	-1.73	1961-1990
Eskdalemuir	55.32	-3.20	1961-1990
Gatwick	51.15	-0.18	1961-1990
Heathrow	51.48	-0.45	1961-1990
Hemsby	52.68	1.68	1961-1987
Herstmonceaux*****	50.94	0.19	*****
Ringway	53.35	-2.28	1961-1990
Yeovilton	51.00	-2.63	1965-1990

** Data for Herstmonceaux requires further processing and re-formatting before training the weather generator for this station.*

Wherever data availability permits, a standard training period of 1961-1990 has been used to calibrate the weather generator. However, shorter training periods had to be used for four stations (Table 2) – the shortest being 17 years for Coltishall. The weather generator is stochastic, which means that, once the parameters have been calculated, it can be run for any length of time. For all validation simulations, the weather generator is run for 30 years. The simulated time series should have the same distribution and statistical characteristics as the training period, but the simulated years do not correspond to a ‘real’ calendar year, i.e., there is no day-by-day or year-by-year correspondence between the observed and simulated time series. Thus in the output files the years are numbered 0001 to 0030 rather than 1961 to 1990.

A different sequence of random numbers is produced each time the weather generator is run, hence different daily time series are produced each time. Thus it is important to use output from multiple runs when validating performance. Initially the weather generator was run 1000 times for each simulation set (e.g., Figure 1 in BETWIXT Technical Briefing Note 1). However, sensitivity studies indicate that similar variability is obtained across 100 runs as across 1000 runs. Thus, in order to reduce computational time and data volumes, the weather generator is run 100 times, i.e., each validation simulation set consists of 100 30-year long simulations.

Three sets of validation plots are presented in this technical briefing note: (a) precipitation and temperature (Section 2 and Appendix 1); (b) sunshine, wind speed, vapour pressure and reference PET (Section 3 and Appendix 2); and (c) temperature and precipitation extreme events (Section 4 and Appendix 3). These plots compare observed (blue) and simulated (red) variables. In each case, the mean of the 100 weather generator simulations is shown (red dots), together with the plus/minus two standard deviation range (red vertical lines and bars) calculated across the 100 simulations. Since there is only one realisation of the observations, these are indicated by a single blue cross. The performance of the weather generator is considered to be good whenever the observed value falls within the simulated range, and very good whenever the observed value is very close to the mean of the simulated series.

2. PRECIPITATION AND TEMPERATURE

The first set of validation plots (Appendix 1) is for precipitation (the primary generated variable) and maximum and minimum temperature (secondary generated variables). In each case, mean values for each half-month are shown.

The first parameter shown is proportion of dry days, an indicator of precipitation occurrence, followed by the mean wet day precipitation, an indicator of precipitation intensity. Both parameters are well simulated for all sites and half months, i.e., the observed value falls within the simulated range in the majority of cases. The observed seasonal cycle is successfully captured, although the weather generator has a slight tendency to underestimate summer dryness at the drier southern stations (e.g., Yeovilton).

As noted in the previous section, the stochastic nature of the weather generator means that there is no day-by-day or year-by-year correspondence between the observed and simulated time series, thus correlation is not a useful statistic for validation. However, interannual variability of half-monthly precipitation totals (the third panel in the Appendix 1 plots) provides a useful indicator of how well year-to-year variability is reproduced by the weather generator. With a single exception (the second half of August at Bradford), the observed value always falls within the simulated range, although the latter is fairly large.

The bottom two panels in the Appendix 1 plots show mean minimum and maximum temperature respectively. The seasonal cycle of both variables is very well simulated at all stations. The regression equations used to generate temperature (equations 3.1 to 3.3 in BETWXT Technical Briefing Note 1) include a random element term. However, the variability across the 100 simulations is considerably less than for precipitation.

3. SUNSHINE, WIND SPEED, VAPOUR PRESSURE AND PET

The second set of validation plots (Appendix 2) shows mean values for each half-month for the remaining secondary variables (i.e., sunshine hours, wind speed and vapour pressure) and calculated reference PET. The other calculated variable, relative humidity, is not shown as this is calculated from vapour pressure (equations 4.1 and 4.2 in BETWIXT Technical Briefing Note 1).

At all stations, the weather generator slightly overestimates mean sunshine hours in winter and autumn. There is also a slight tendency to underestimate the summer peak in mean sunshine hours at southernmost stations, i.e., Gatwick, Heathrow and Yeovilton. Overall, however, the shape of the seasonal cycle is well simulated.

The weather generator overestimates mean wind speeds in winter and autumn when the observed values lie just below the simulated range. This tendency is also evident in spring and summer, particularly at more northerly stations such as Eskdalemuir, although to a lesser extent.

Vapour pressure is well simulated at all stations.

Reference PET is also well simulated, although there is a slight tendency to underestimate the summer peak at southernmost stations (i.e., Gatwick, Heathrow and Yeovilton) – a tendency also noted in sunshine hours.

4. PRECIPITATION AND TEMPERATURE EXTREME EVENTS

In the previous two sections it has been shown that the CRU daily weather generator generally performs well with respect to the reproduction of mean climate. However, for many of the impact sectors being investigated as part of the BKCC programme, the main concern is potential future changes in the frequency and intensity of extreme events.

The important precipitation and temperature extremes depend on the particular impact sector being studied and stakeholder requirements with respect to information about extremes also vary widely (Goodess *et al*., 2003). These divergent needs are reflected in the STARDEX Diagnostic Extremes Indices Software developed as part of the European-Union funded STARDEX project on 'STAtistical and Regional dynamical Downscaling of EXtremes for European regions' (http://www.cru.uea.ac.uk/cru/projects/stardex/). This publicly-available software package calculates 57 different indices of extreme precipitation and temperature, identified after consultation with end users. Here, six of the STARDEX indices (Table 3) which are likely to be relevant to impacts of climate change on the built environment are used to validate the CRU daily weather generator.

Table 3:Definitions of extreme events used for validation.

BETWIXT description	***STARDEX name***	***Definition***
Fraction of total precipitation from intense events	pf95	Fraction of total precipitation above the annual 95th percentile value
Maximum number of consecutive dry days	pxcdd	Maximum number of consecutive dry days
Number of "Hot days"	txf90	Number of days when maximum temperature is greater than the 90th percentile value
Heatwave duration	txhwd	Cumulative count of number of consecutive days when maximum temperature exceeds the 90th percentile value for more than 5 days (NB the first 5 days are not counted in the index)
Number of "Warm nights"	tnf90	Number of days when minimum temperature is greater than the 90th percentile value
Number of "Cold nights"	tnf10	Number of days when minimum temperature is less than the 10th percentile value

The first index 'Fraction of total precipitation from intense events' is an indicator of the intensity of extreme precipitation events (i.e., the wettest 5% of events), while the second precipitation-based index 'Maximum number of consecutive dry days' is an indicator of the persistence of dry, i.e., potential drought, conditions. The other four indices describe different characteristics of the extreme temperature regime, based on the warmest/coldest 10% of events. They have been chosen because of their relevance to thermal comfort in urban areas and buildings, and to heating/cooling energy requirements, for example.

The weather generator tends to overestimate the fraction of total precipitation from intense events, although the observed values never fall far outside the simulated range. At some stations, such as Hemsby and Ringway, this index is well simulated in all seasons. At the majority of stations, it has a summer maximum and this is correctly reproduced by the weather generator. However, at Heathrow, this summer peak is underestimated.

The maximum number of consecutive dry days is consistently underestimated by the weather generator. In a number of cases, the observed value exceeds the simulated range by 5 days or more. This underestimation of the persistence of dry (and wet) days is an inherent problem of stochastic weather generators (Gregory *et al*., 1993; Wilks and Wilby, 1999; Goodess, 2000; Wilby and Wigley, 2000). It is often associated with the underestimation of variance, hence it is encouraging that interannual variability of precipitation is reasonably well simulated by the CRU daily weather generator (see Section 2).

At a number of stations, the weather generator tends to overestimate the number of "Hot days" and Heatwave duration in winter (and to a lesser extent in autumn) and to underestimate these indices in summer (and to a lesser extent in spring). Over the year as a whole, these indices are more frequently overestimated than underestimated. In general, the number of "Warm nights" and "Cold nights" (which are based on minimum temperature) are rather better simulated than the indices based on maximum temperature. Both night-time indices are well simulated at stations such as Abbotsinch, Bradford, Heathrow and Ringway. However, these indices do tend to be underestimated (except at Coltishall) – most noticeably at Yeovilton.

5. CONCLUDING REMARKS

The figures presented in the appendices to this technical briefing note indicate that mean values are generally well and in most cases, very well simulated, by the CRU daily weather generator – for the primary, secondary and calculated variables. Performance is somewhat less good with respect to extreme events. However, the indices of extremes chosen are based on 90^{th} and 95^{th} percentile values and thus provide a severe test of how well the model reproduces the tails of the distributions. In general, given the severity of the test, the model is considered to perform reasonably well with respect to extremes.

The only consistently poor performance of the weather generator is with respect to the maximum number of consecutive dry days, with is underestimated in all seasons, at all stations. As noted in Section 4, this is an inherent problem of stochastic weather generators. A number of solutions to this problem have been proposed, including using a higher-order Markov Chain (i.e., precipitation occurrence is dependent on events over several days prior to the day being simulated rather than just the previous day) and the use of an inflation or expansion factor to increase persistence. However, these all have disadvantages, such as increasing the number of model parameters in the case of higher-order models, and tending to be somewhat arbitrary with unpredictable behaviour in the case of inflation factors (see brief review in Section 7.3.5 of Goodess, 2000) and were not considered to be appropriate solutions for implementation in the CRU daily weather generator.

Overall, however, the performance of the CRU daily weather generator is considered to very satisfactory, and robust across the range of UK climate regimes for which it has been tested (i.e., from wet northern sites such as Eskdalemuir to dryer/warmer southern sites such as Hemsby and Yeovilton). Thus it is concluded that it is an appropriate tool for the needs of the BKCC programme, i.e., for the construction of high spatial/temporal resolution state-of-art climate scenarios for selected case-study locations which are consistent with the UKCIP02 scenarios (Hulme *et al*., 2002). The construction of these scenarios and the projected changes will be described in another BETWIXT technical briefing note.

REFERENCES

Goodess, C.M., 2000: *The Construction of Daily Rainfall Scenarios for Mediterranean Basins Using a Circulation-type Approach to Downscaling*, PhD Thesis, University of East Anglia, available from: http://www.cru.uea.ac.uk/~clareg/clare.htm

Goodess, C.M., Osborn, T.J. and Hulme, M., 2003: *The Identification and Evaluation of Suitable Scenario Development Methods for the Estimation of Future Probabilities of Extreme Weather Events*., Tyndall Centre Technical Report 4, available from: http://www.tyndall.ac.uk/research/theme3/summary_it1_16.shtml

Gregory, J.M., Wigley, T.M.L. and Jones, P.D., 1993: ‘Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals’, *Climate Dynamics*, **8**, 299-310.

Hulme, M., Jenkins, G.J., Lu, X., Turnpenny, J.R., Mitchell, T.D., Jones, R.G., Lowe, J., Murphy, J.M., Hassell, D., Boorman, P., McDonald, R. and Hill, S., 2002: *Climate Change Scenarios for the United Kingdom: The UKCIP02 Scientific Report*, Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK. 120pp.

Jones, P.D. and Salmon, M., 1995: *Development and Integration of a Stochastic Weather Generator into a Crop Growth Model for European Agriculture*, MARS Project, Final Report to Institute of Remote Sensing Applications, Agricultural Information Systems (ISPRA)., Contract No. 5631-93-12, ED ISP GB.

Richardson, C.W., 1981: ‘Stochastic simulation of daily precipitation, temperature, and solar radiation’, *Water Resources Research*, **17**, 182-190.

Watts, M., Goodess, C.M. and Jones, P.D., 2004: *The CRU Daily Weather Generator*, BETWIXT Technical Briefing Note 1, available from: http://www.cru.uea.ac.uk/cru/projects/betwixt/.

Wilby, R.L. and Wigley, T.M.L., 2000: ‘Precipitation predictors for downscaling observed and general circulation model relationships’, *International Journal of Climatology*, **20**, 641-661.

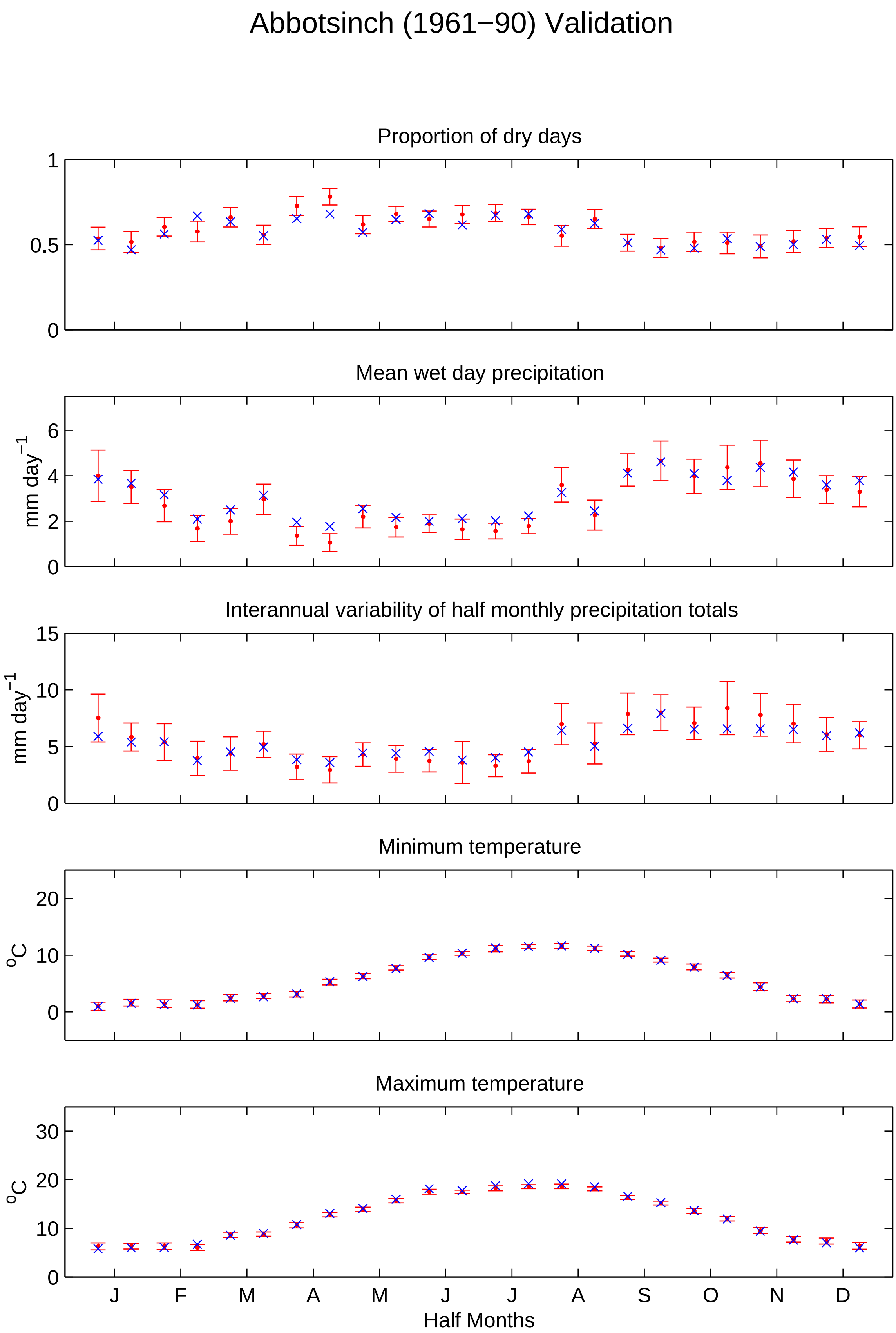
Wilks, D.S. and Wilby, R.L., 1999: ‘The weather generation game: a review of stochastic weather models’, *Progress in Physical Geography*, **23**, 329-357.

APPENDIX 1: VALIDATION PLOTS FOR PRECIPITATION AND TEMPERATURE

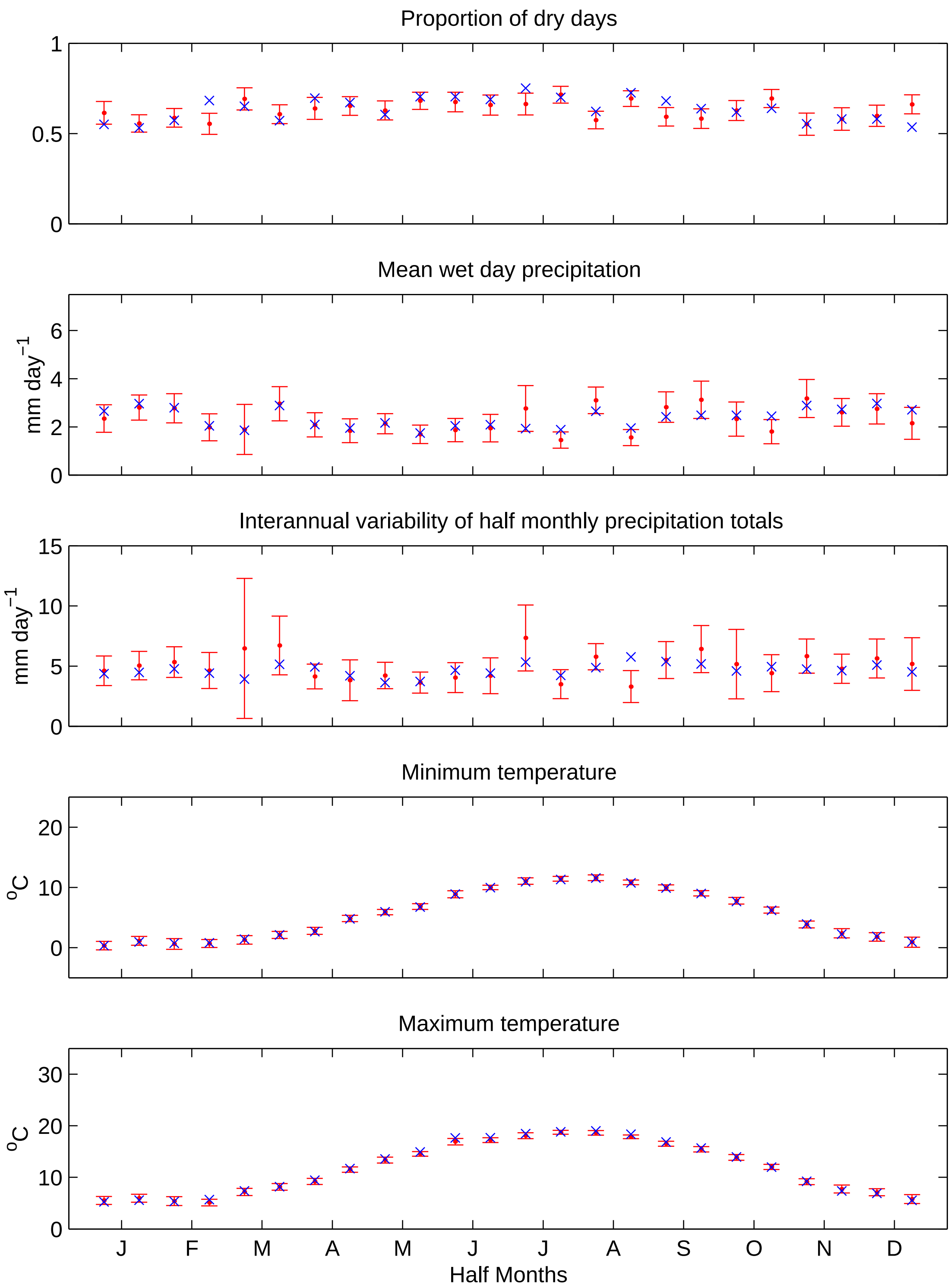
Observed (blue) and simulated (red) values for each half month for the following variables:

- Proportion of dry days
- Mean wet day precipitation (mm day^{-1})
- Interannual variability of half monthly precipitation totals (mm day^{-1})
- Minimum temperature (^{o}C)
- Maximum temperature (^{o}C)

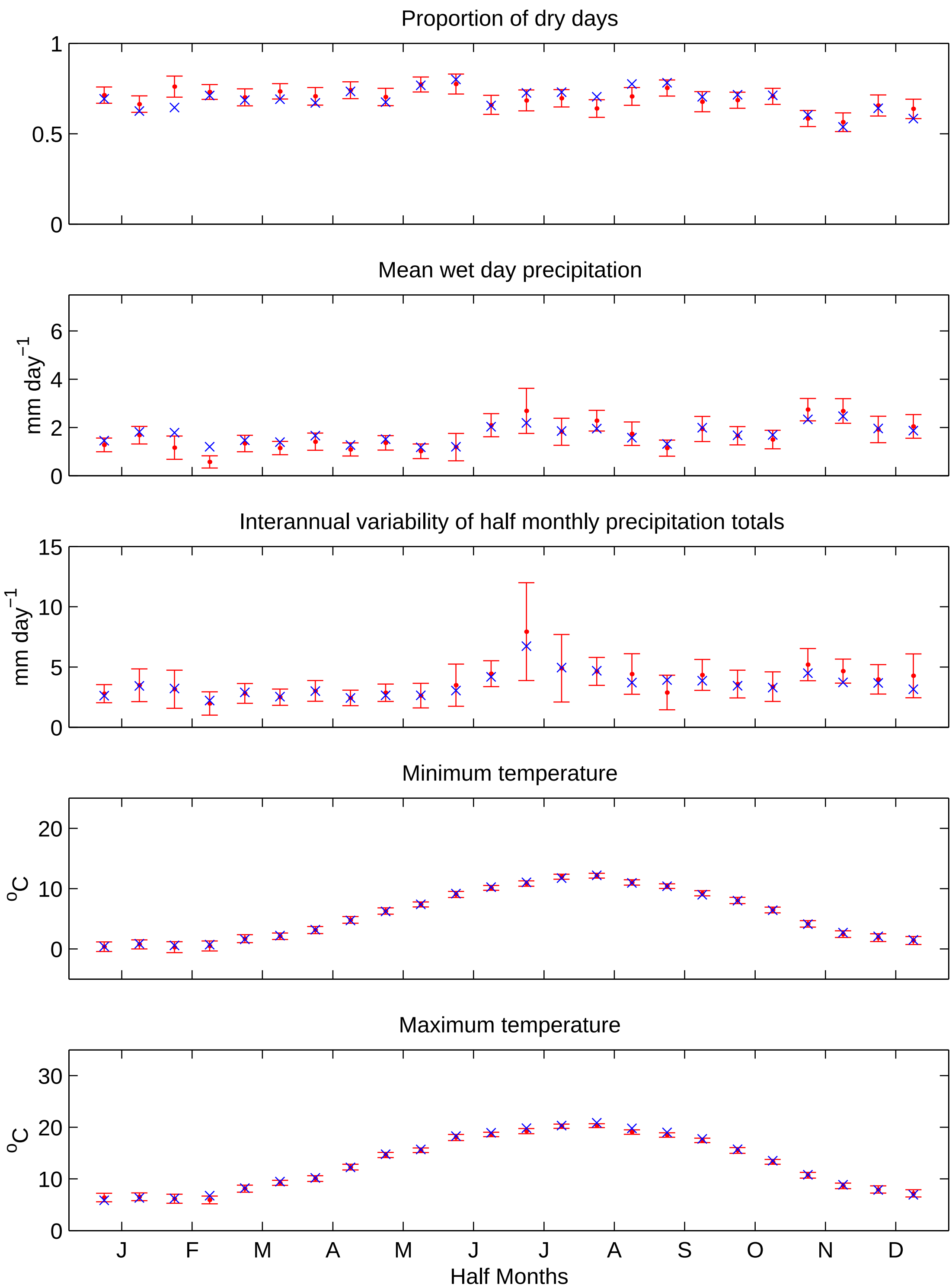
Observed values are the mean for the period shown in Table 2 (i.e., usually 1961-1990). The simulated values are the mean of 100 30-year weather generator runs (red dots). The red lines and bars show the variability of the 100 series (plotted as plus/minus two standard deviations around the mean).

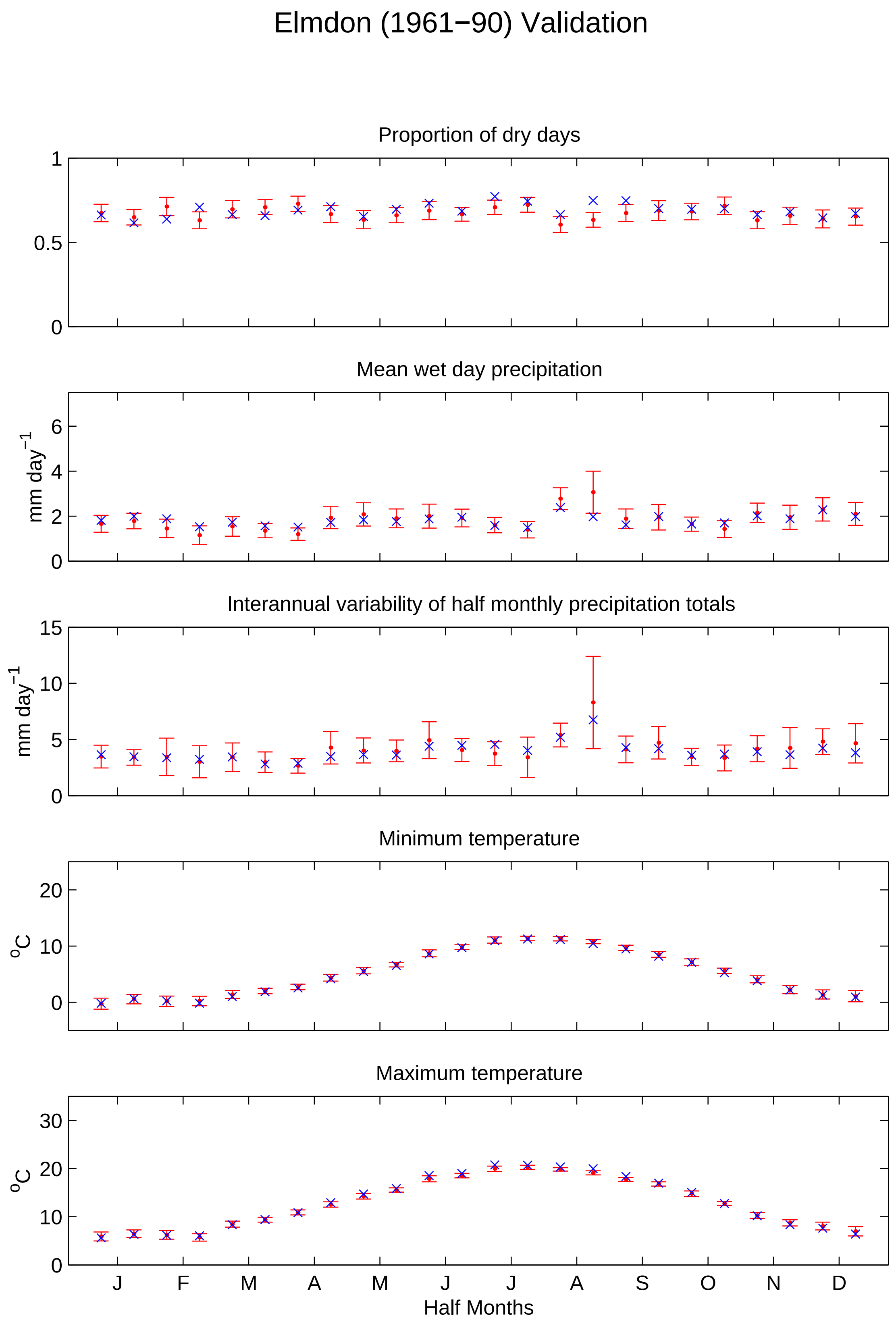

Abbotsinch (1961−90) Validation
Proportion of dry days
1
0.5
0
Mean wet day precipitation
mm day−1
6
4
2
0
Interannual variability of half monthly precipitation totals
mm day−1
15
10
5
0
Minimum temperature
oC
20
10
0
Maximum temperature
oC
30
20
10
0
J F M A M J J A S O N D
Half Months

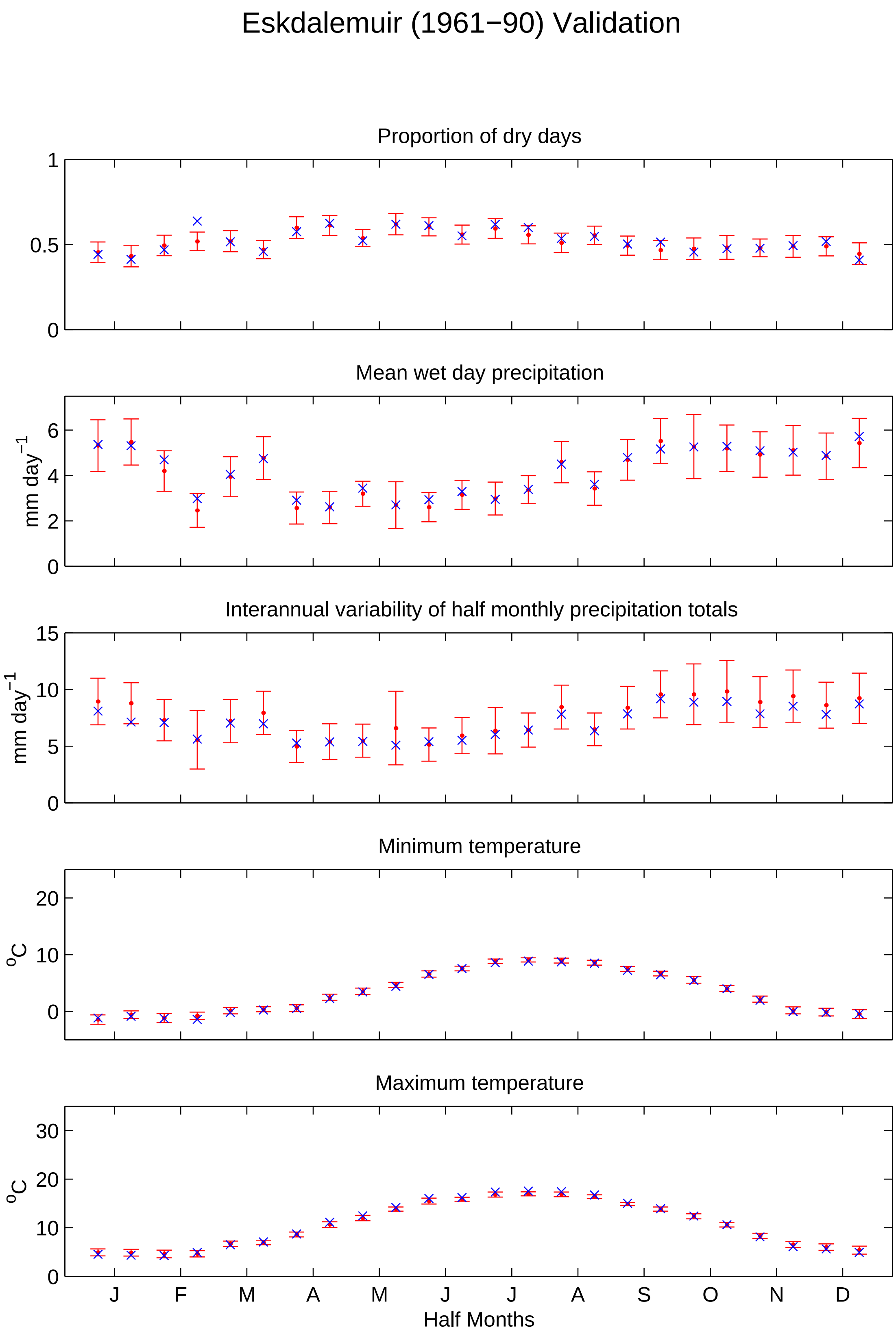
Bradford (1961−90) Validation

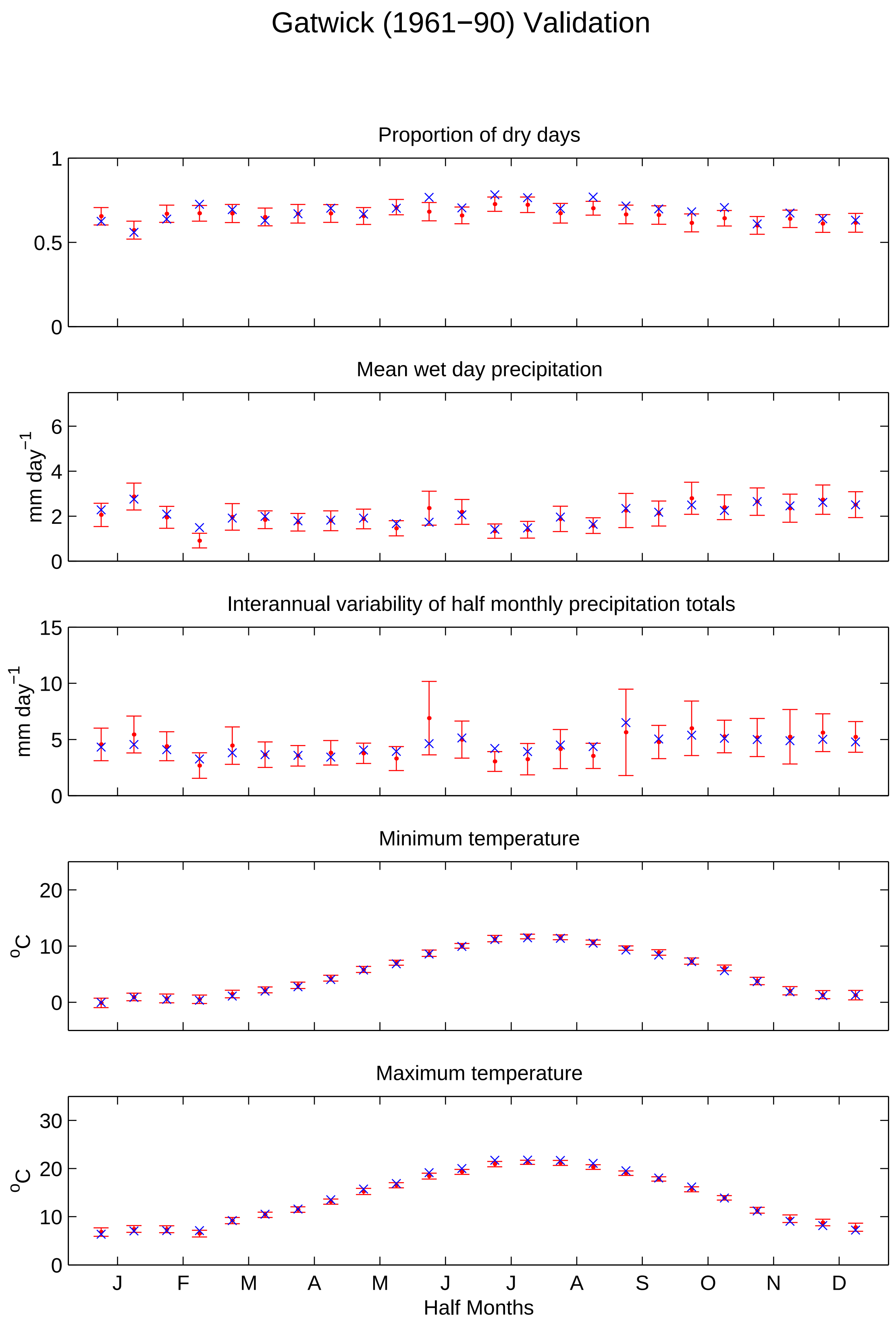


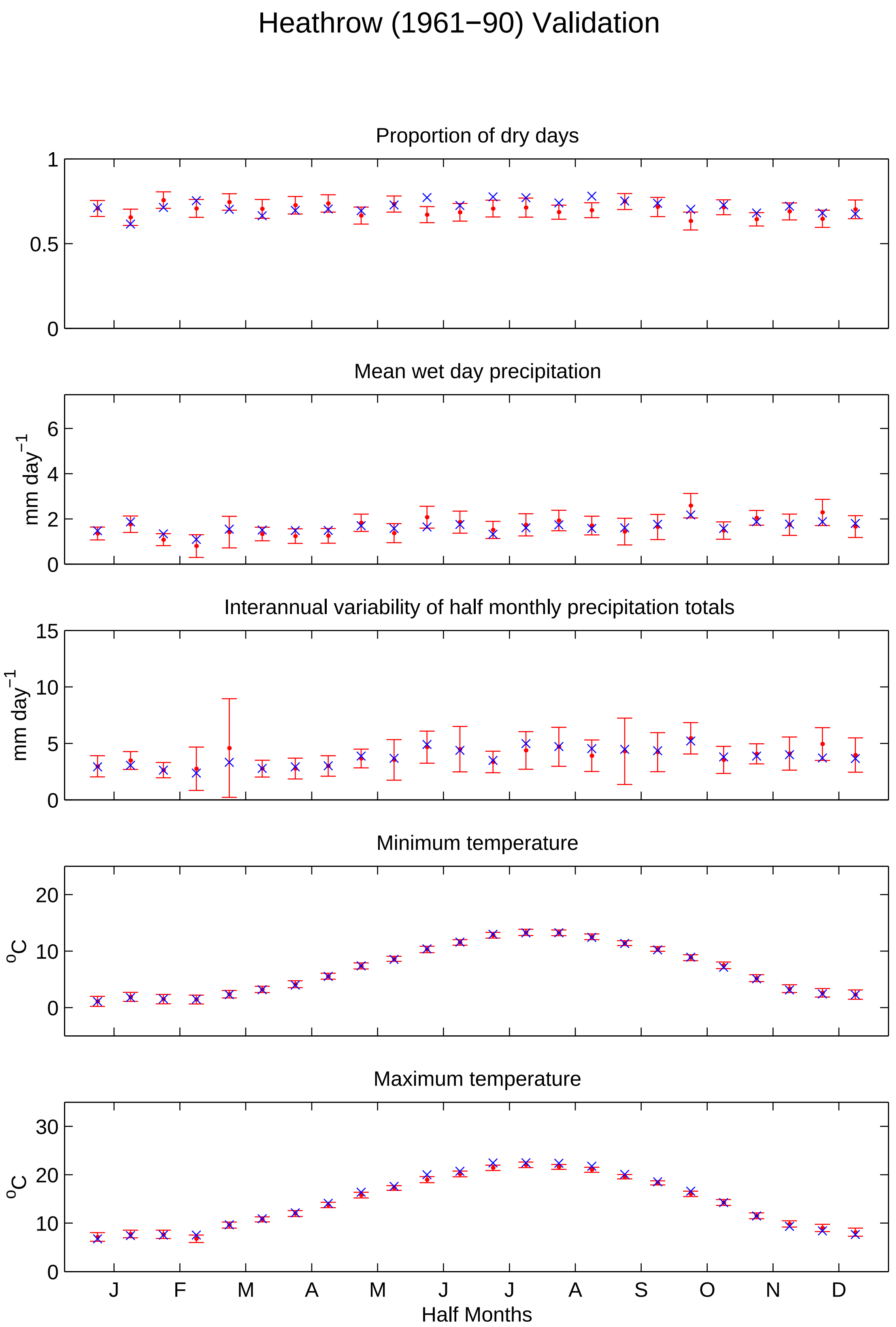
Coltishall (1961−90) Validation

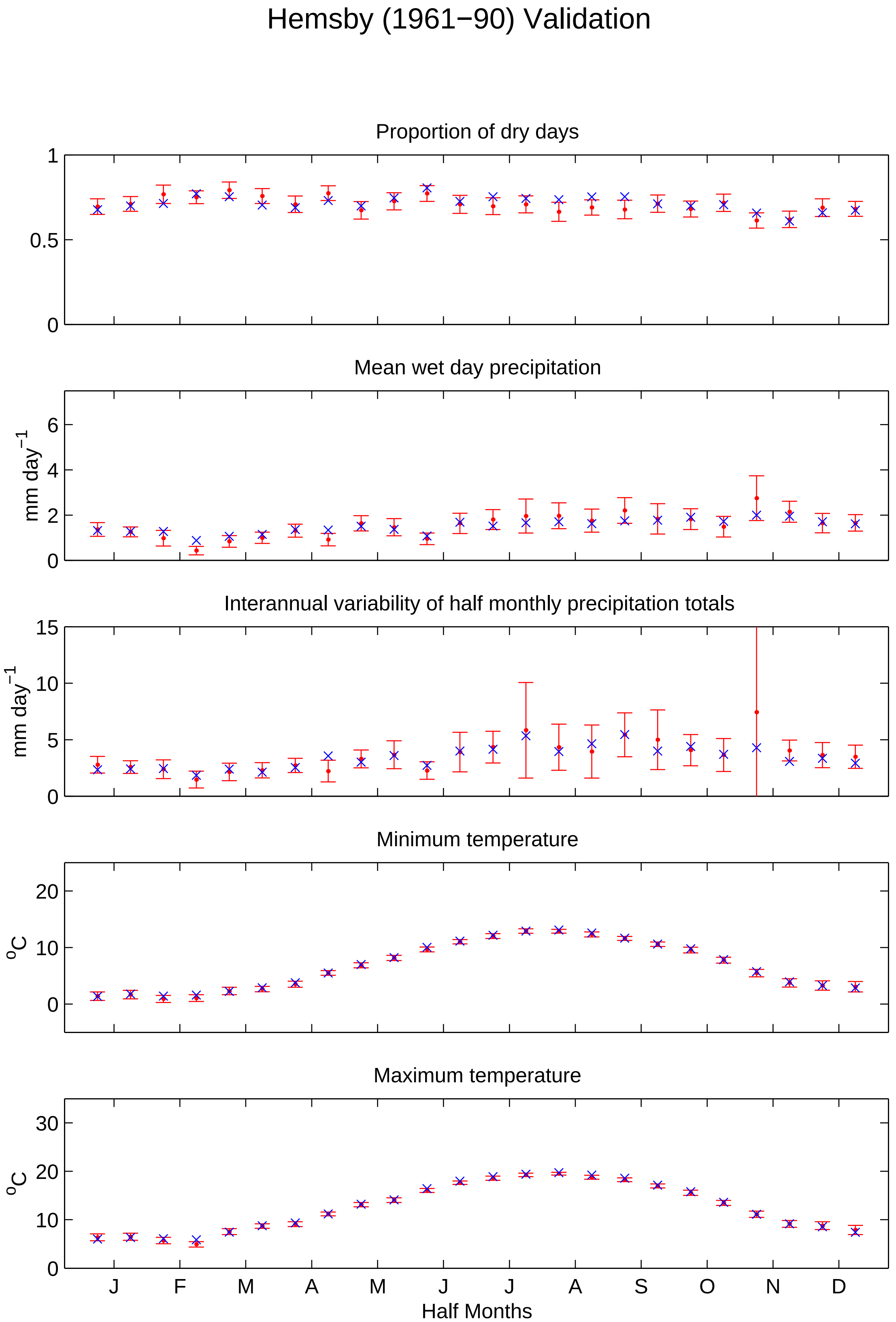


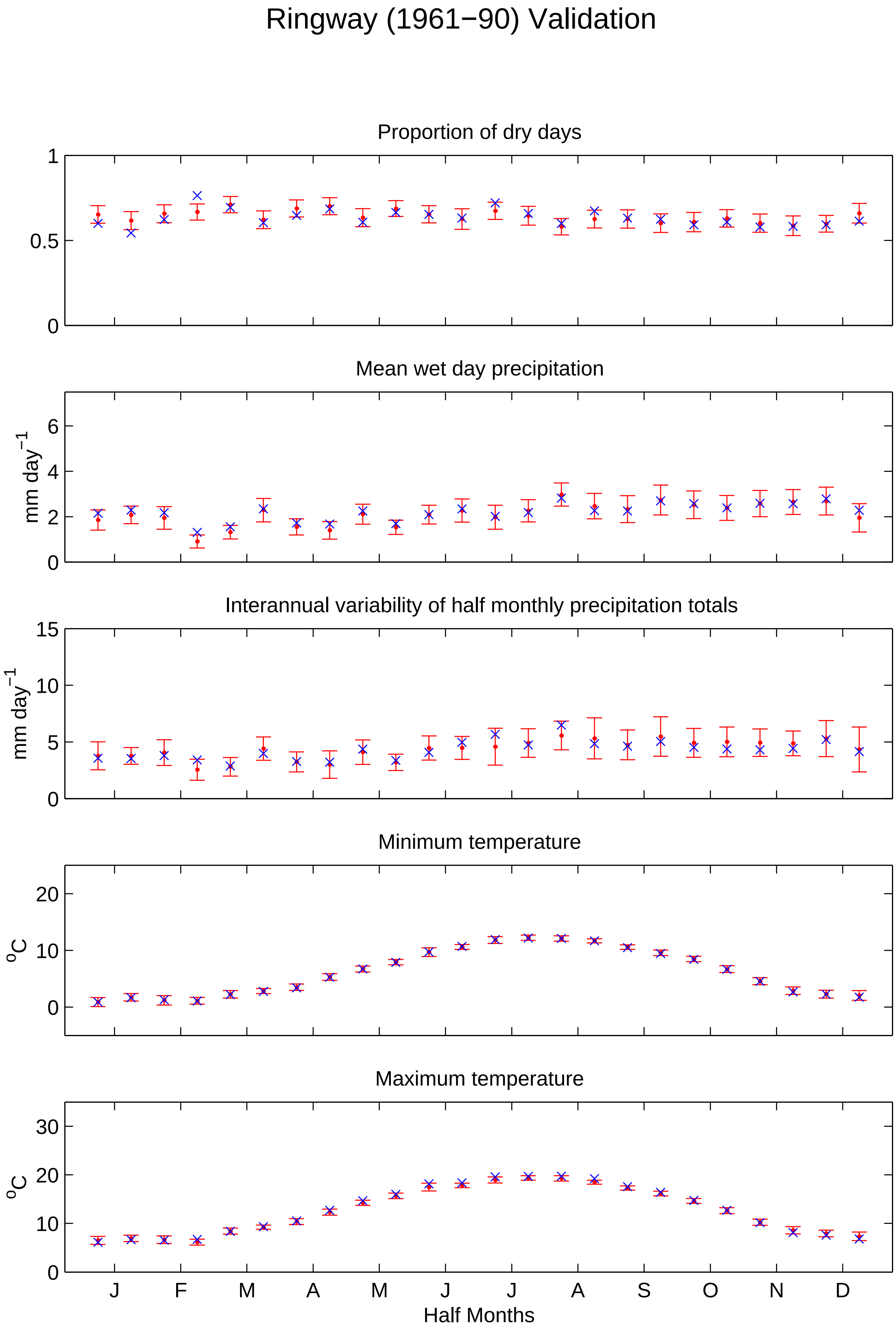

Elmdon (1961−90) Validation
Proportion of dry days
1
0.5
0
Mean wet day precipitation
mm day−1
6
4
2
0
Interannual variability of half monthly precipitation totals
mm day−1
15
10
5
0
Minimum temperature
oC
20
10
0
Maximum temperature
oC
30
20
10
0
J F M A M J J A S O N D
Half Months

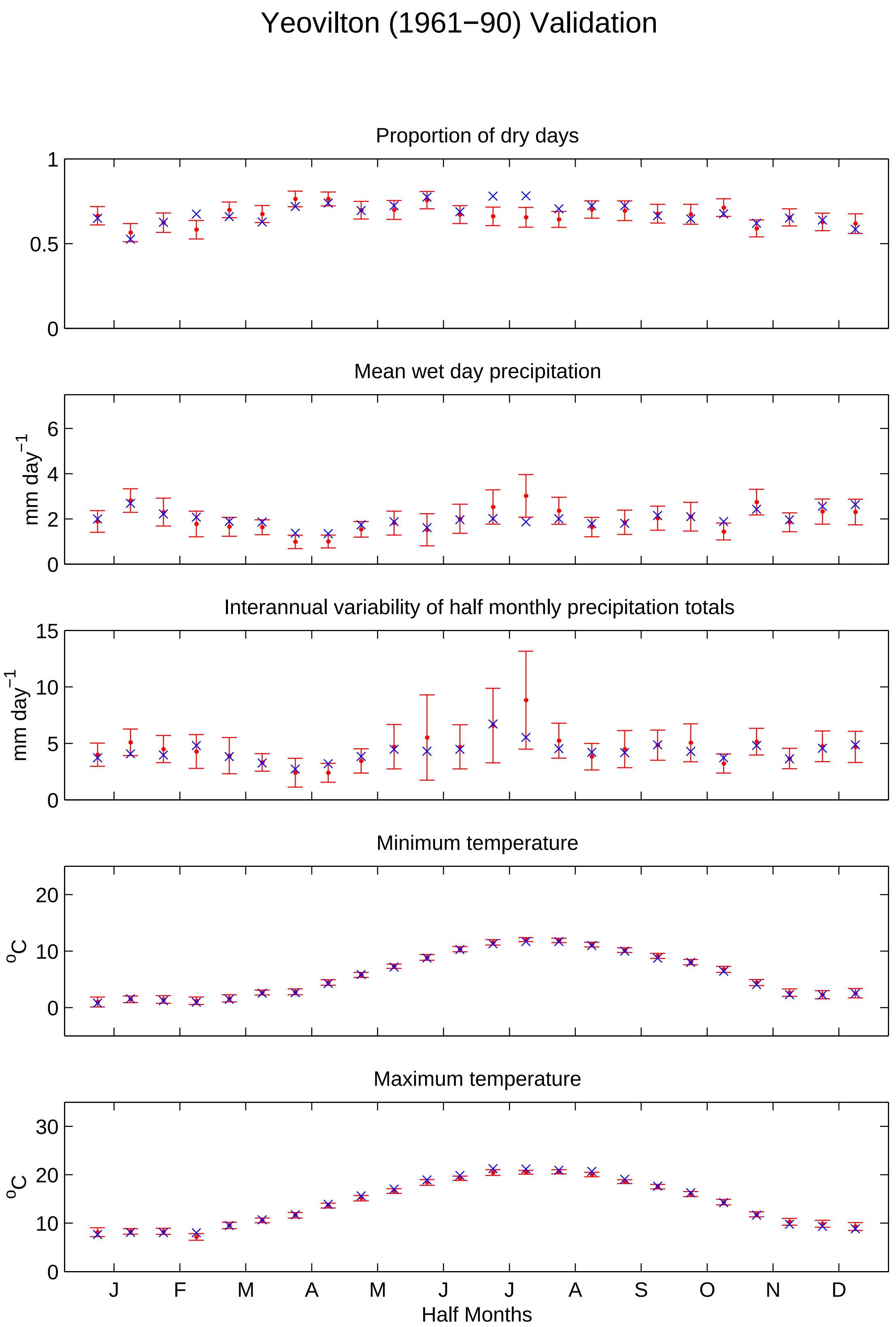

Eskdalemuir (1961−90) Validation
Proportion of dry days
Mean wet day precipitation
mm day−1
Interannual variability of half monthly precipitation totals
mm day−1
Minimum temperature
oC
Maximum temperature
oC
J F M A M J J A S O N D
Half Months


Gatwick (1961−90) Validation
Proportion of dry days
Mean wet day precipitation
mm day−1
Interannual variability of half monthly precipitation totals
mm day−1
Minimum temperature
oC
Maximum temperature
oC
J F M A M J J A S O N D
Half Months


Heathrow (1961−90) Validation
Proportion of dry days
Mean wet day precipitation
mm day−1
Interannual variability of half monthly precipitation totals
mm day−1
Minimum temperature
oC
Maximum temperature
oC
J F M A M J J A S O N D
Half Months


Hemsby (1961−90) Validation
Proportion of dry days
1
0.5
0
Mean wet day precipitation
mm day−1
6
4
2
0
Interannual variability of half monthly precipitation totals
mm day−1
15
10
5
0
Minimum temperature
oC
20
10
0
Maximum temperature
oC
30
20
10
0
J F M A M J J A S O N D
Half Months


Ringway (1961−90) Validation
Proportion of dry days
1
0.5
0
Mean wet day precipitation
mm day−1
6
4
2
0
Interannual variability of half monthly precipitation totals
mm day−1
15
10
5
0
Minimum temperature
oC
20
10
0
Maximum temperature
oC
30
20
10
0
J F M A M J J A S O N D
Half Months



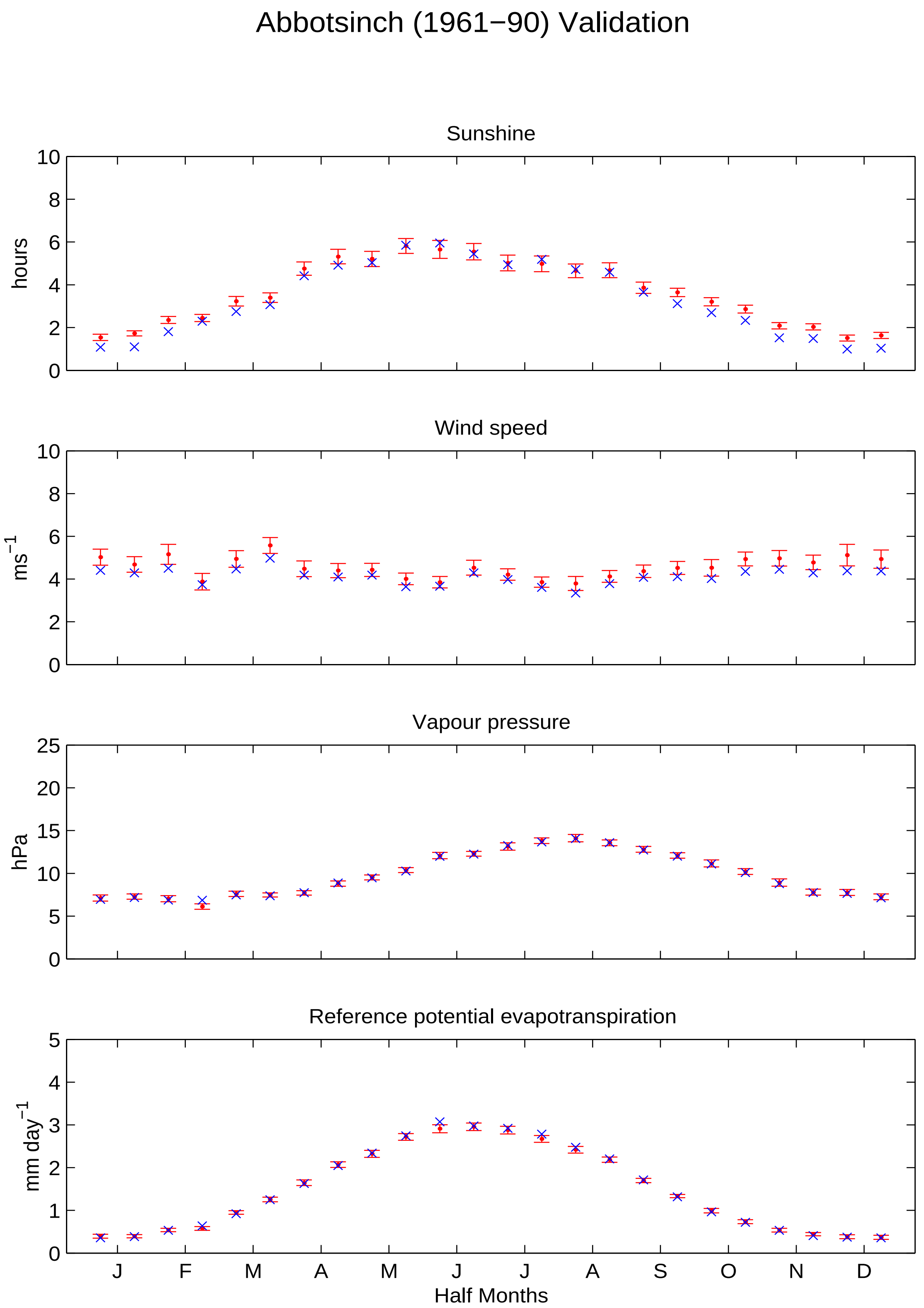
Yeovilton (1961−90) Validation
Proportion of dry days
Mean wet day precipitation
mm day−1
Interannual variability of half monthly precipitation totals
mm day−1
Minimum temperature
oC
Maximum temperature
oC
J F M A M J J A S O N D
Half Months

APPENDIX 2: VALIDATION PLOTS FOR SUNSHINE, WIND SPEED, VAPOUR PRESSURE AND REFERENCE PET

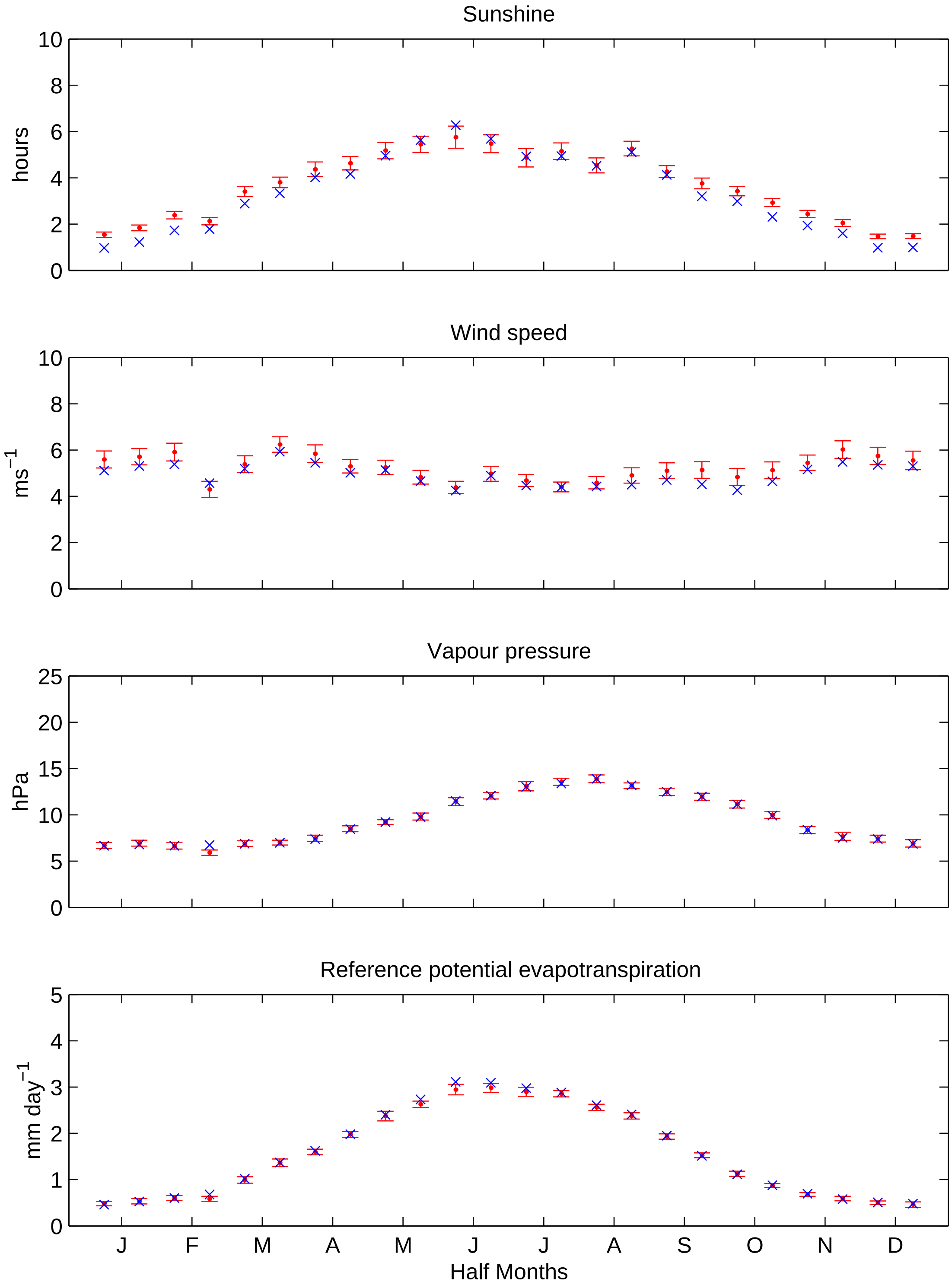
Observed (blue) and simulated (red) values for each half month for the following variables:

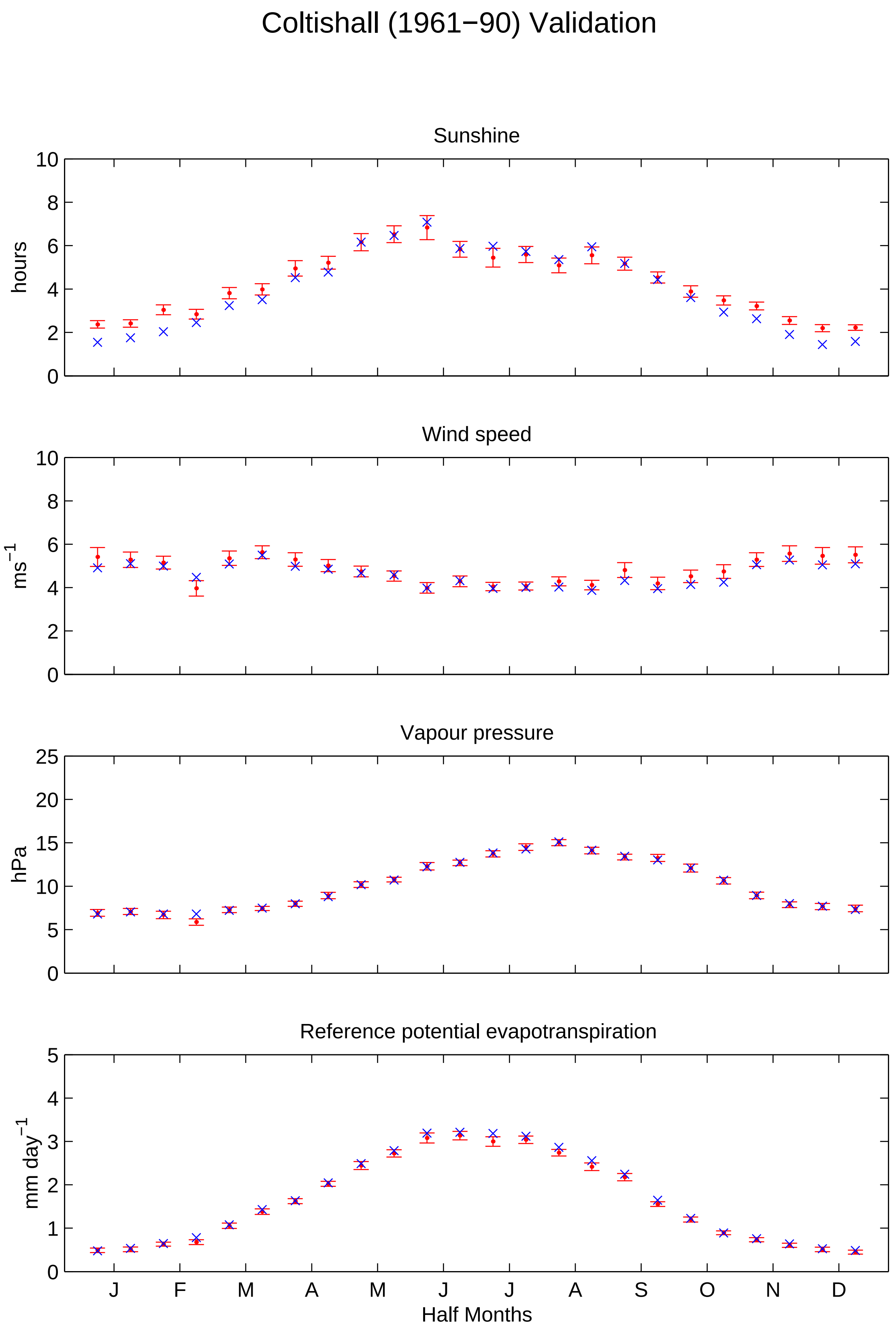
- Sunshine (hours)
- Wind speed (ms^{-1})
- Vapour pressure (hPa)
- Reference potential evapotranspiration (mm day^{-1})

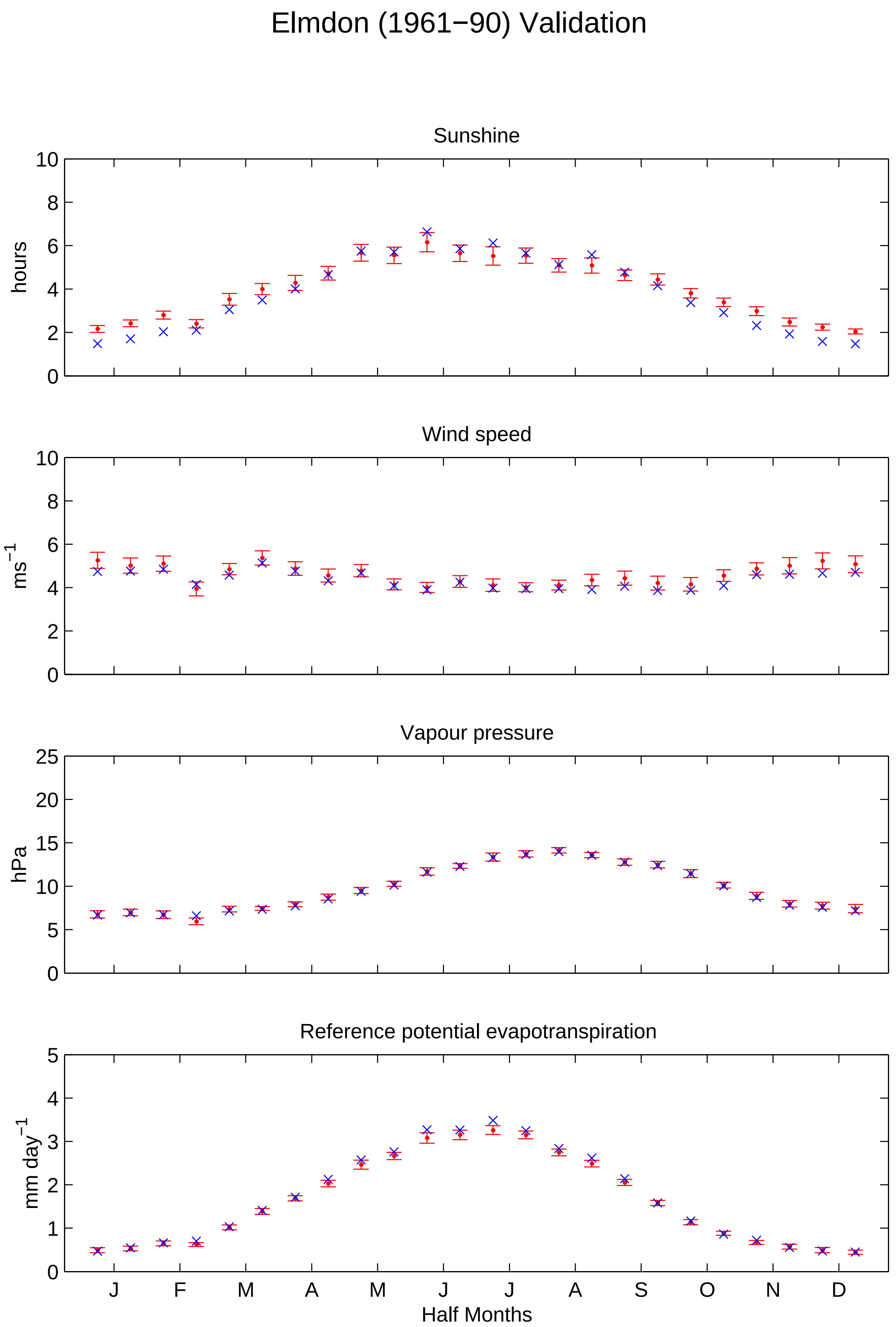
Observed values are the mean for the period shown in Table 2 (i.e., usually 1961-1990). The simulated values are the mean of 100 30-year weather generator runs (red dots). The red lines and bars show the variability of the 100 series (plotted as plus/minus two standard deviations around the mean).

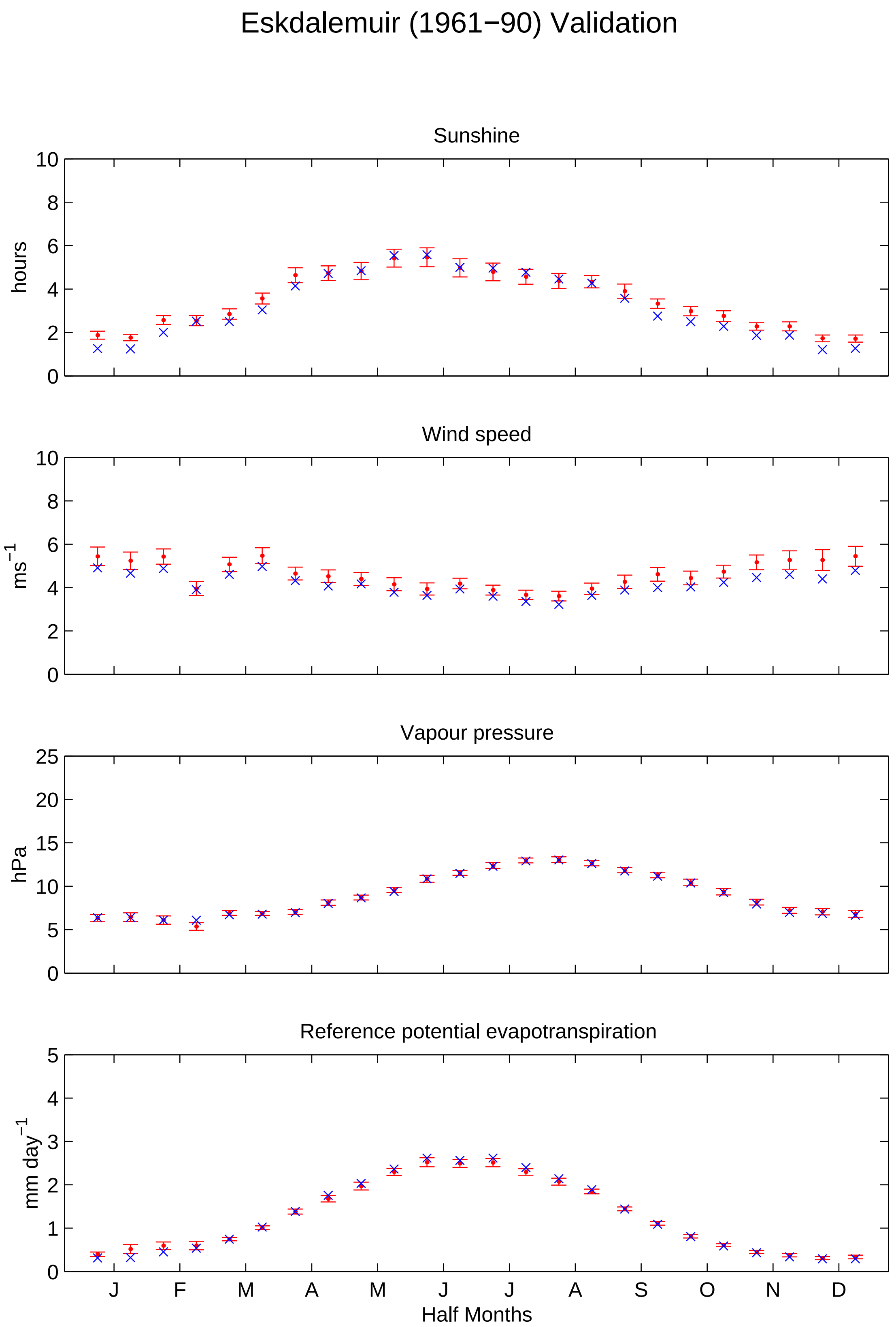

Abbotsinch (1961−90) Validation
Sunshine
hours
10
8
6
4
2
0
Wind speed
ms−1
10
8
6
4
2
0
Vapour pressure
hPa
25
20
15
10
5
0
Reference potential evapotranspiration
mm day−1
5
4
3
2
1
0
J F M A M J J A S O N D
Half Months

Bradford (1961−90) Validation

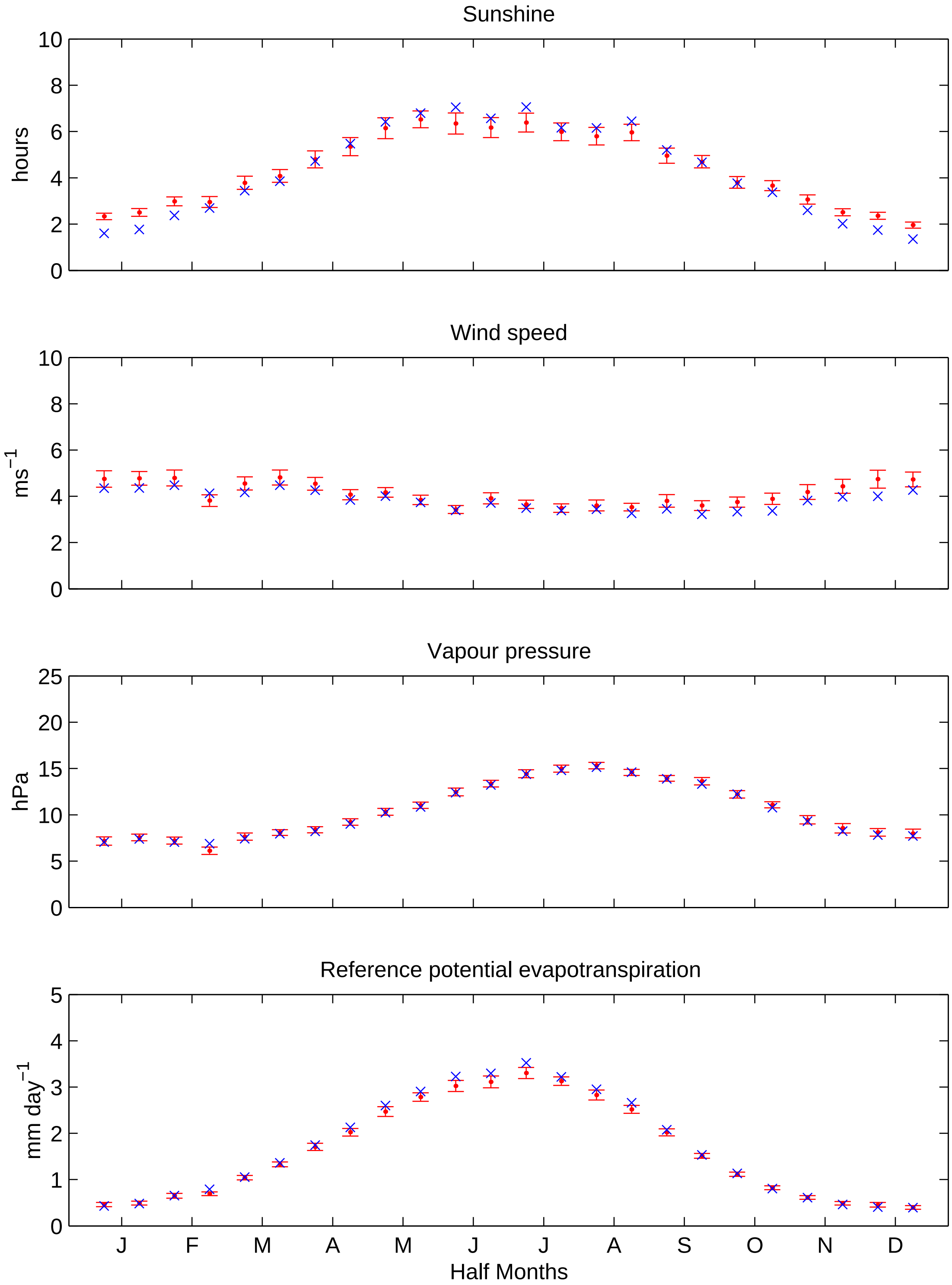



Coltishall (1961−90) Validation
Sunshine
hours
Wind speed
ms−1
Vapour pressure
hPa
Reference potential evapotranspiration
mm day−1
J F M A M J J A S O N D
Half Months

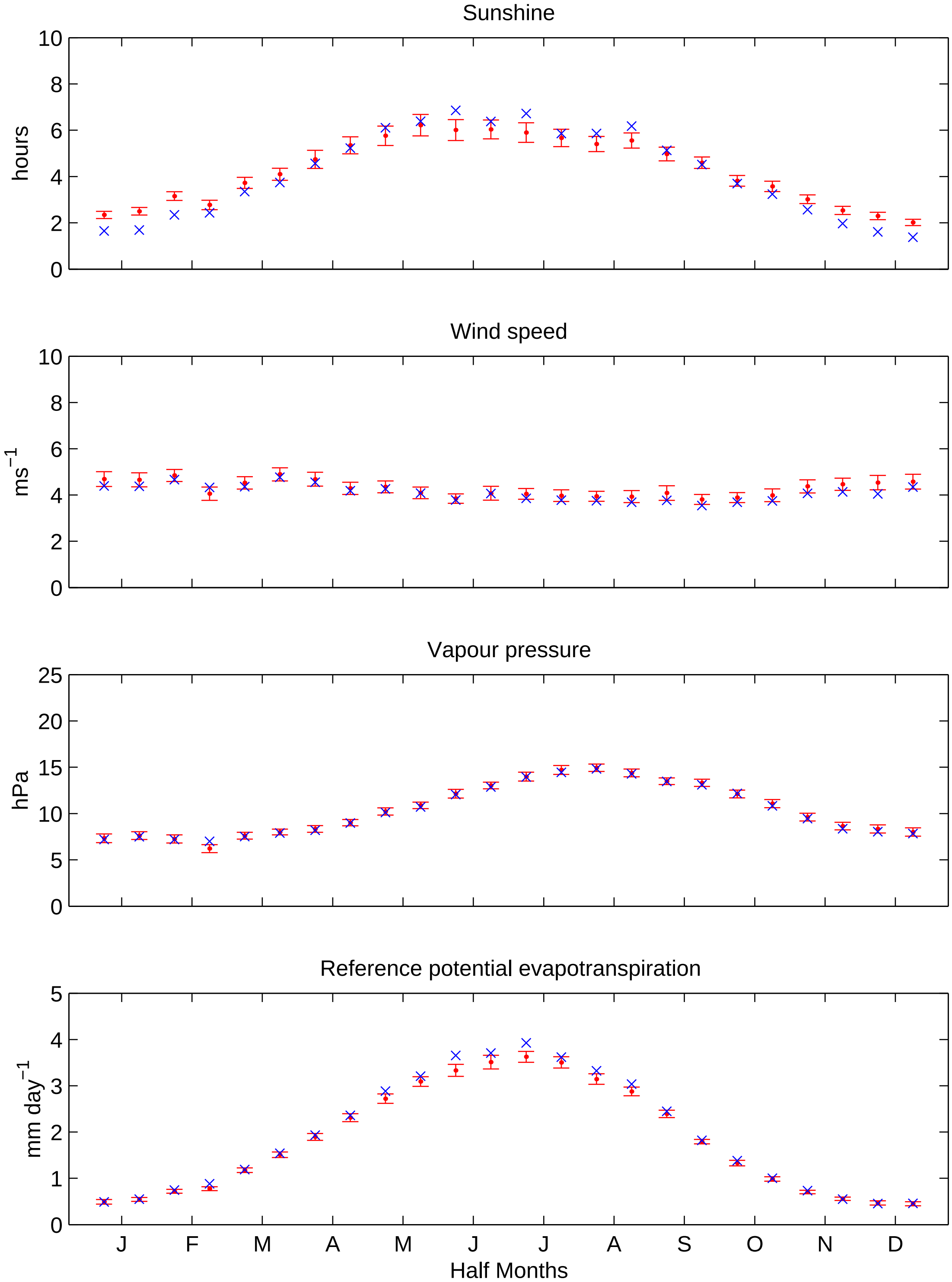

Elmdon (1961−90) Validation
Sunshine
hours
Wind speed
ms−1
Vapour pressure
hPa
Reference potential evapotranspiration
mm day−1
J F M A M J J A S O N D
Half Months

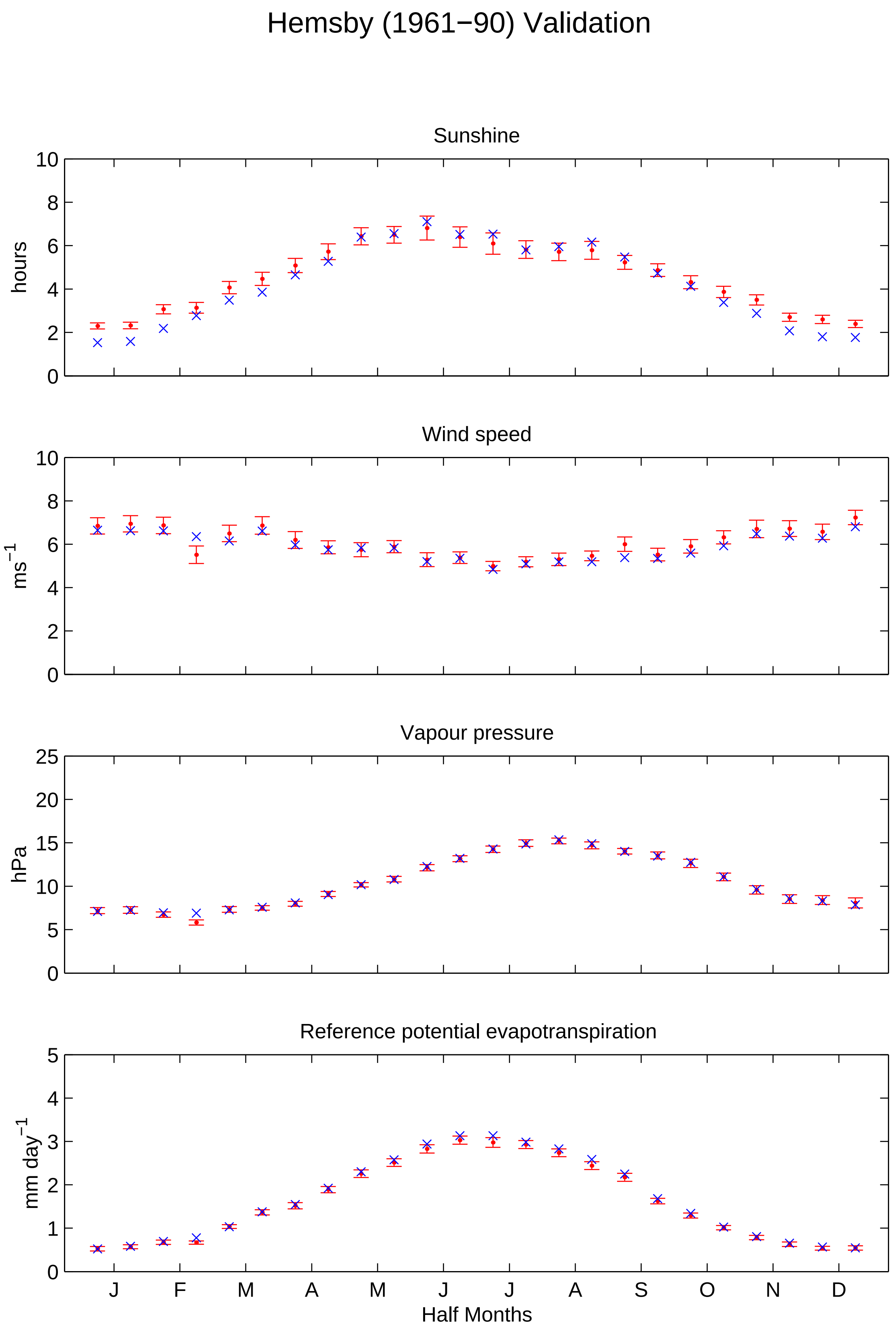

Eskdalemuir (1961−90) Validation
Sunshine
hours
Wind speed
ms−1
Vapour pressure
hPa
Reference potential evapotranspiration
mm day−1
J F M A M J J A S O N D
Half Months

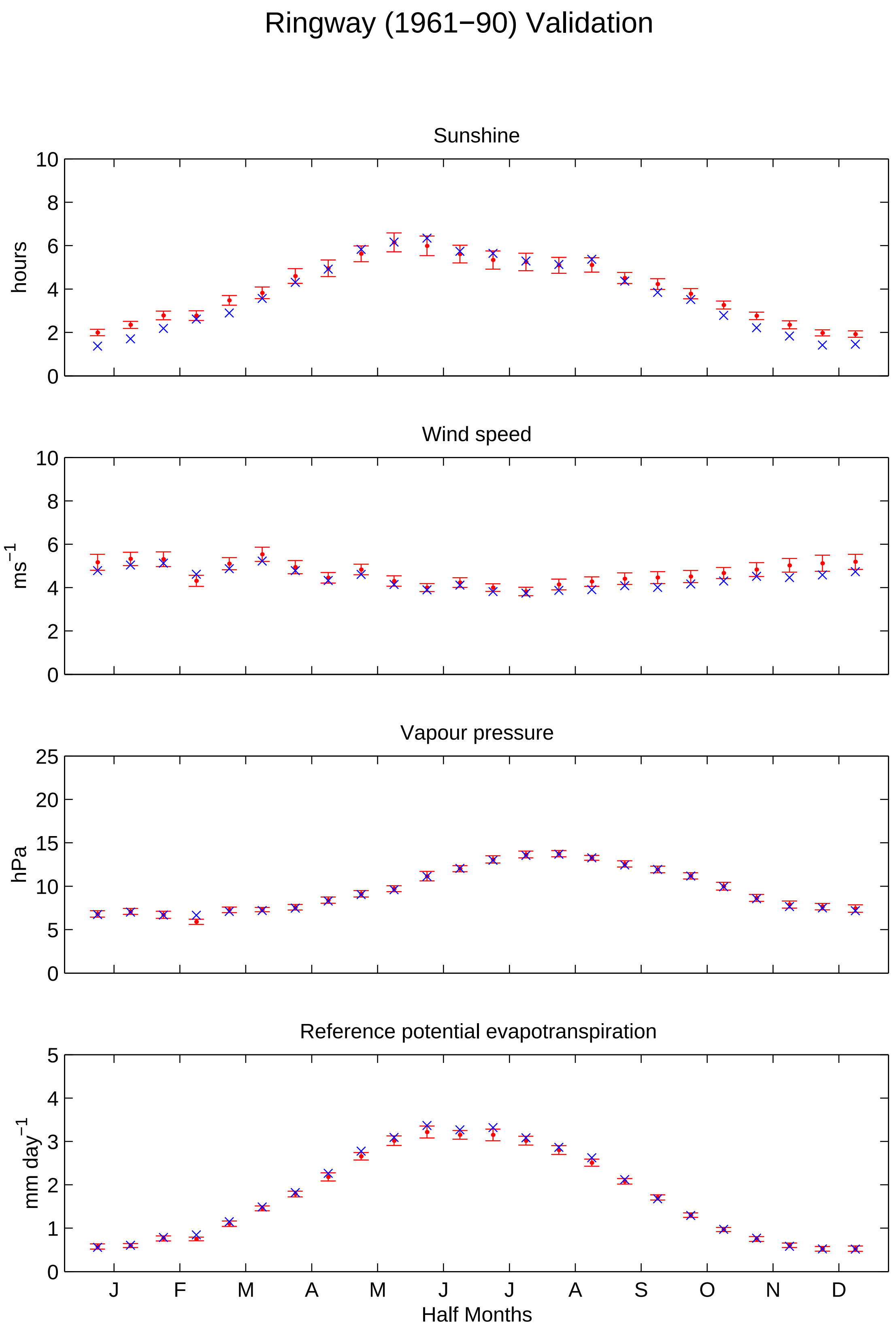
Gatwick (1961−90) Validation

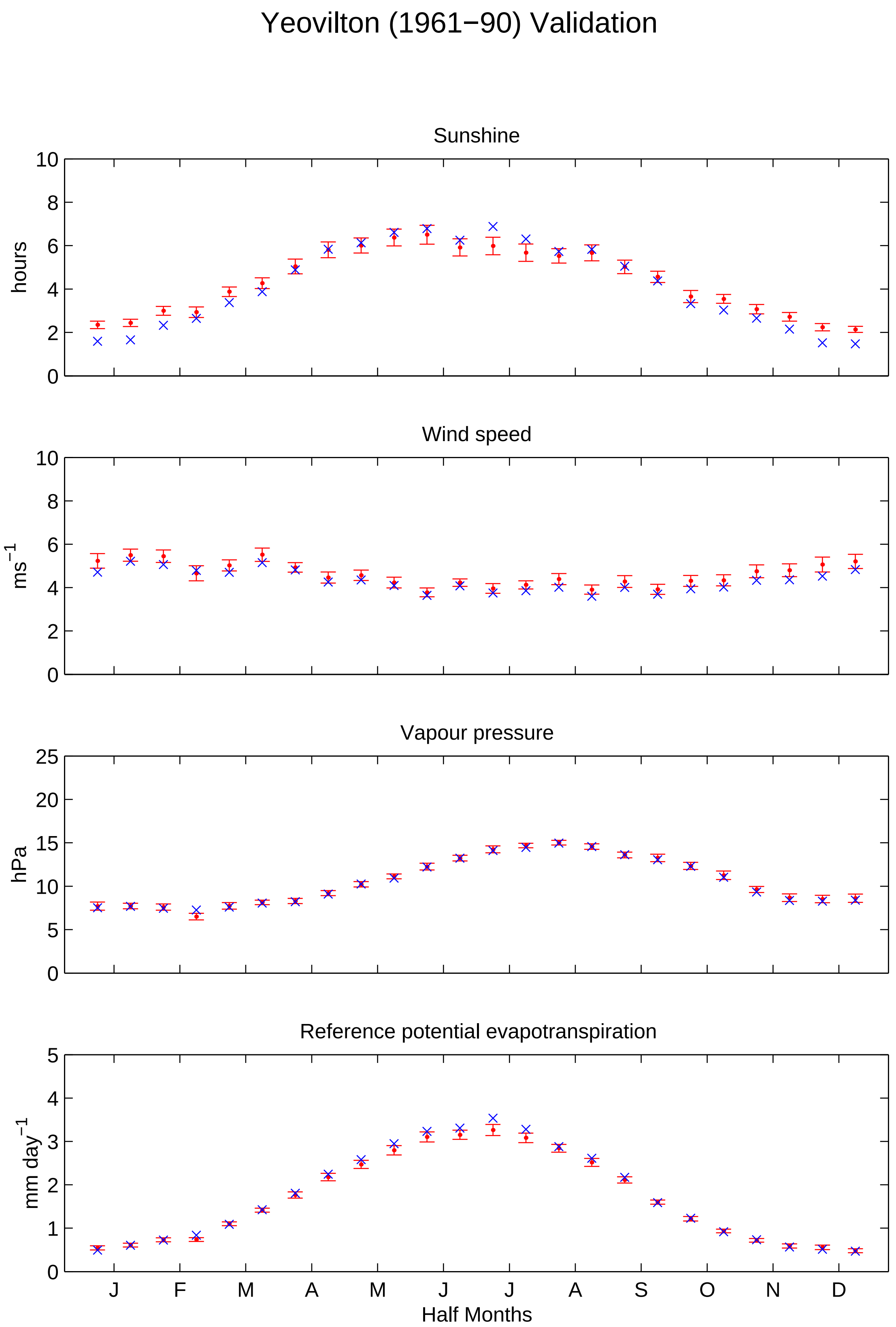


Heathrow (1961−90) Validation




Hemsby (1961−90) Validation
Sunshine
hours
10
8
6
4
2
0
Wind speed
ms−1
10
8
6
4
2
0
Vapour pressure
hPa
25
20
15
10
5
0
Reference potential evapotranspiration
mm day−1
5
4
3
2
1
0
J F M A M J J A S O N D
Half Months


Ringway (1961−90) Validation
Sunshine
hours
Wind speed
ms−1
Vapour pressure
hPa
Reference potential evapotranspiration
mm day−1
J F M A M J J A S O N D
Half Months

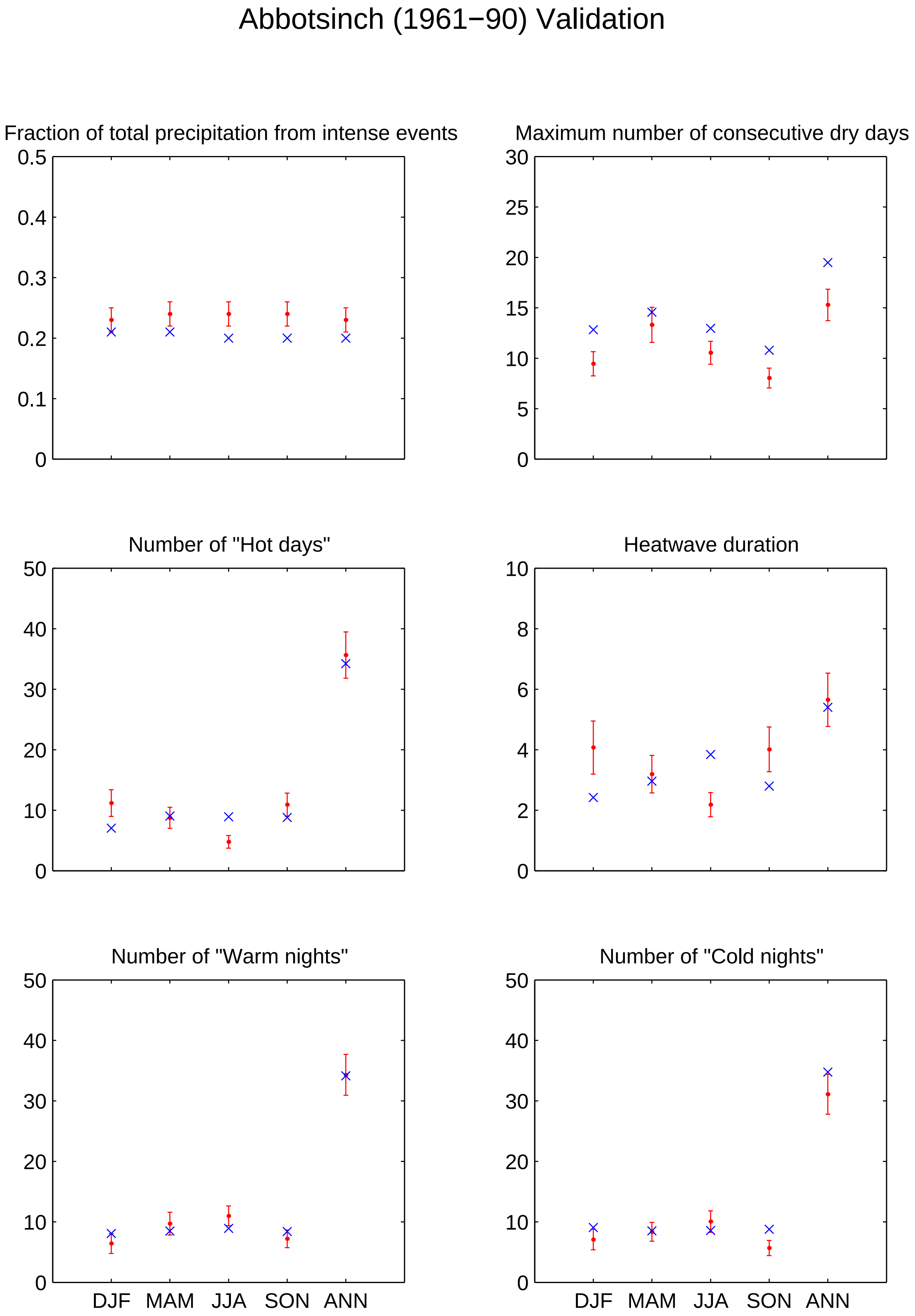

Yeovilton (1961−90) Validation
Sunshine
hours
10
8
6
4
2
0
Wind speed
ms−1
10
8
6
4
2
0
Vapour pressure
hPa
25
20
15
10
5
0
Reference potential evapotranspiration
mm day−1
5
4
3
2
1
0
J F M A M J J A S O N D
Half Months

APPENDIX 3: VALIDATION PLOTS FOR PRECIPITATION AND TEMPERATURE EXTREME EVENTS

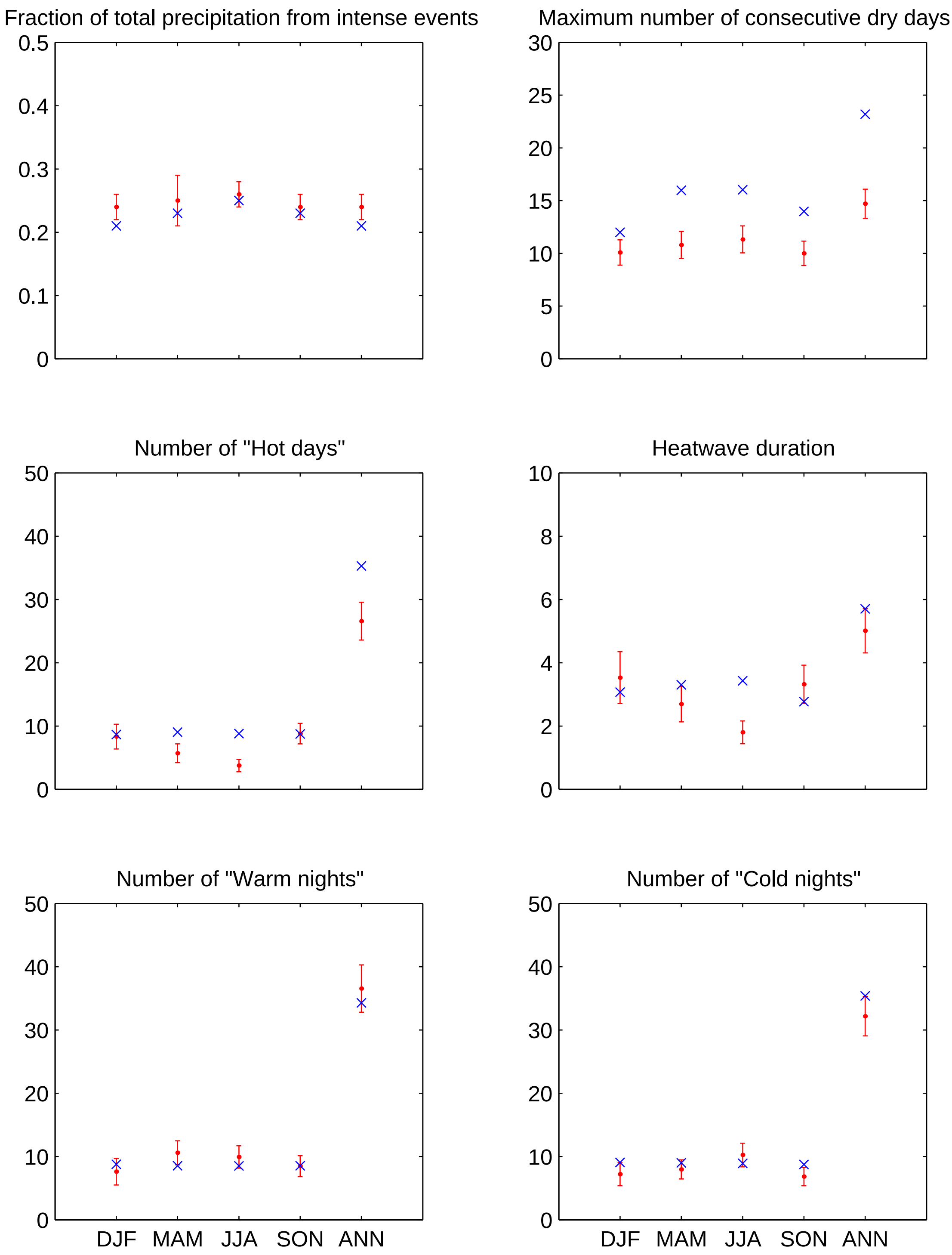
Observed (blue) and simulated (red) values for winter (DJF), spring (MAM), summer (JJA), autumn (SON) and the year as a whole (ANN) for the following variables (which are defined in Table 3):

- Fraction of total precipitation from intense events
- Maximum number of consecutive dry days
- Number of "Hot days"
- Heatwave duration
- Number of "Warm nights"
- Number of "Cold nights"

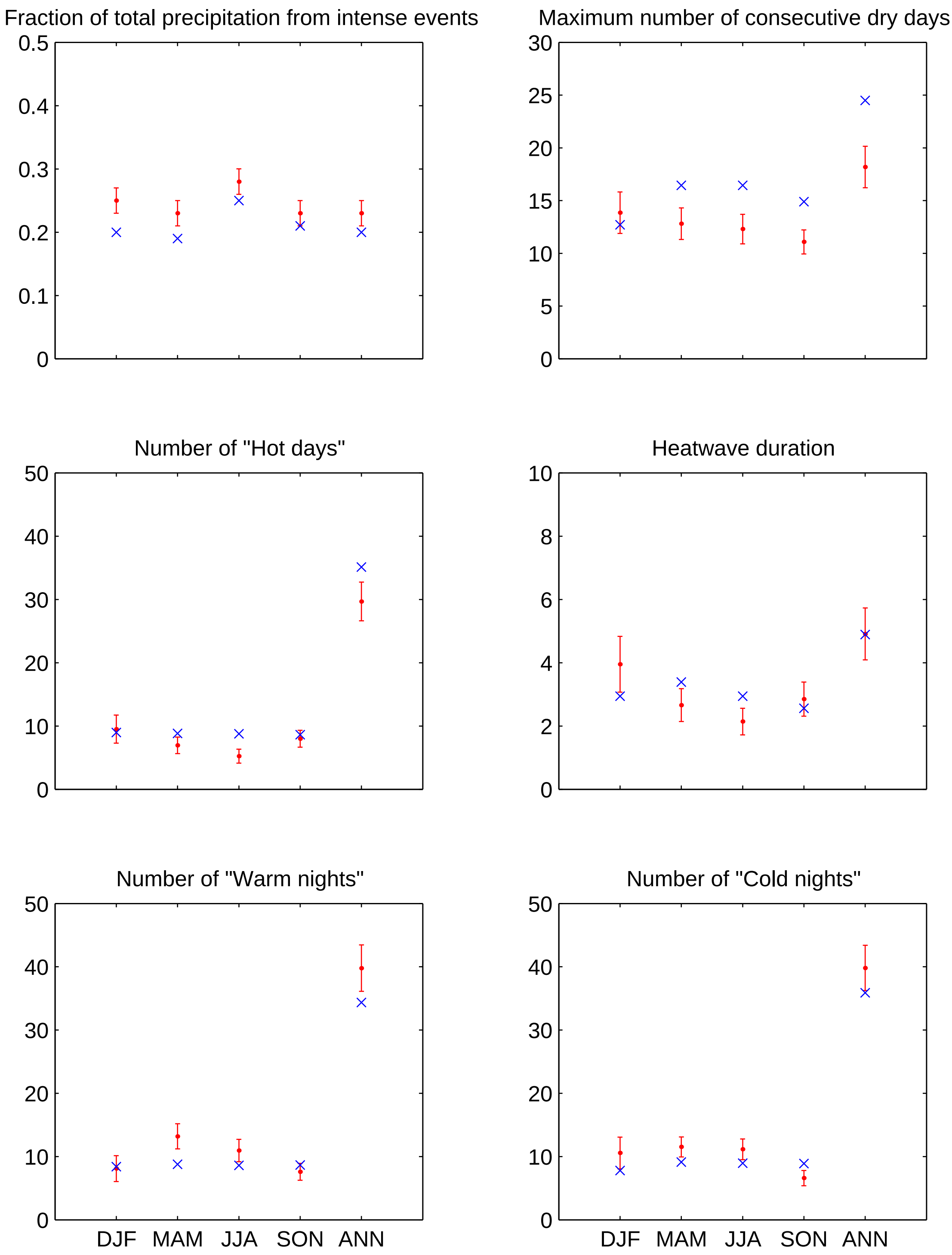
Observed values are the mean for the period shown in Table 2 (i.e., usually 1961-1990). The simulated values are the mean of 100 30-year weather generator runs (red dots). The red lines and bars show the variability of the 100 series (plotted as plus/minus two standard deviations around the mean).


Abbotsinch (1961−90) Validation
Fraction of total precipitation from intense events
Maximum number of consecutive dry days
Number of "Hot days"
Heatwave duration
Number of "Warm nights"
Number of "Cold nights"
DJF MAM JJA SON ANN
DJF MAM JJA SON ANN

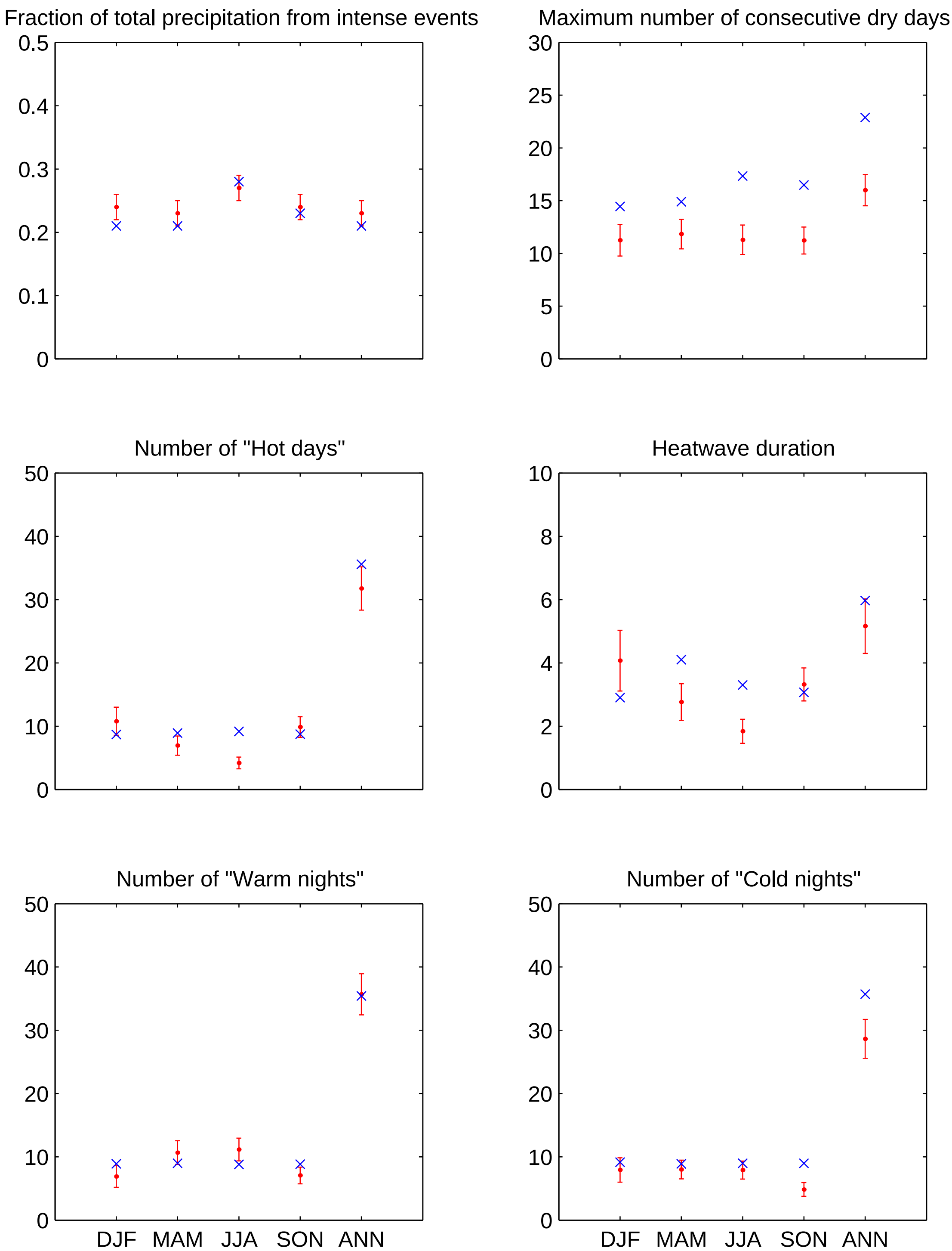
Bradford (1961−90) Validation

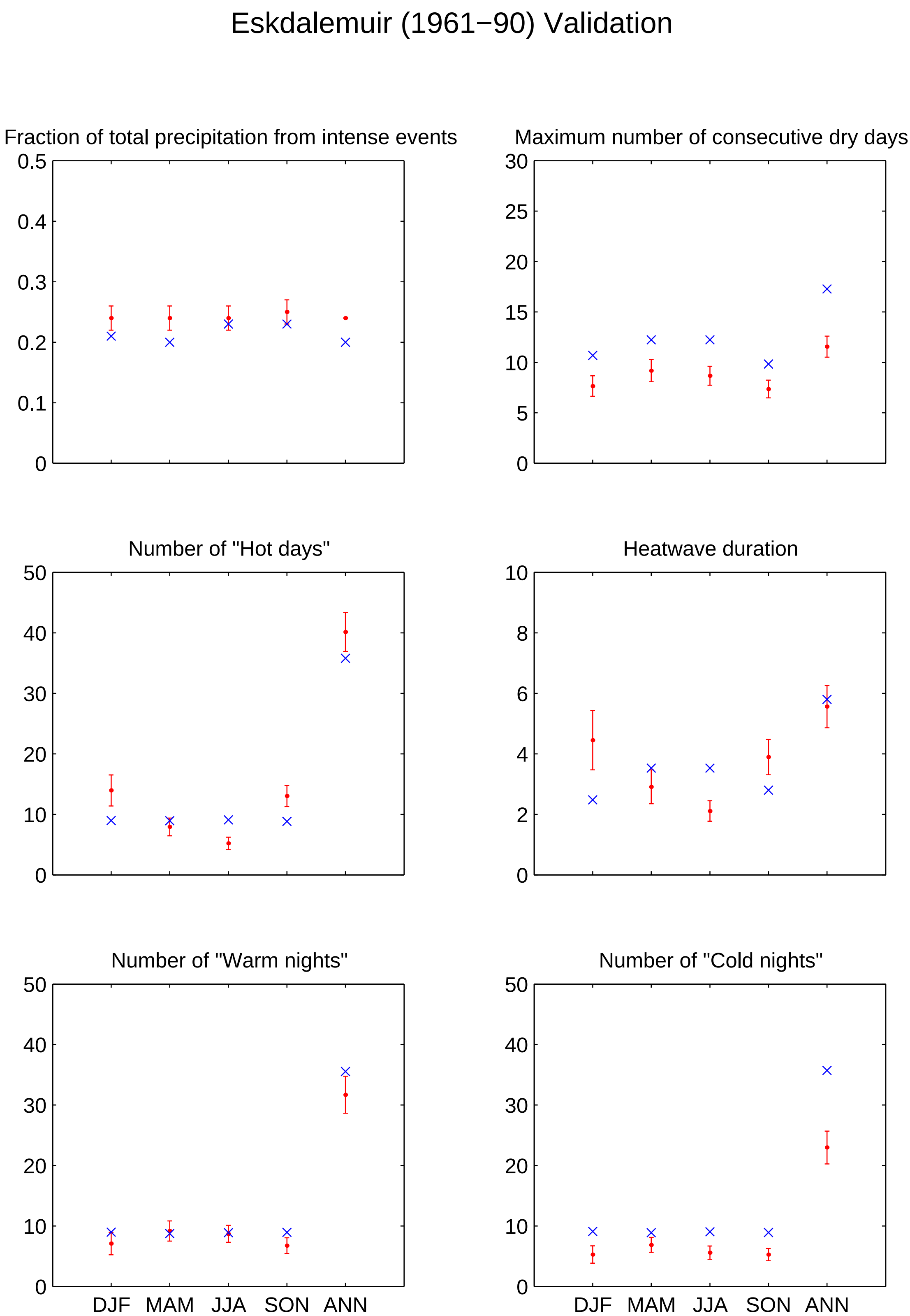


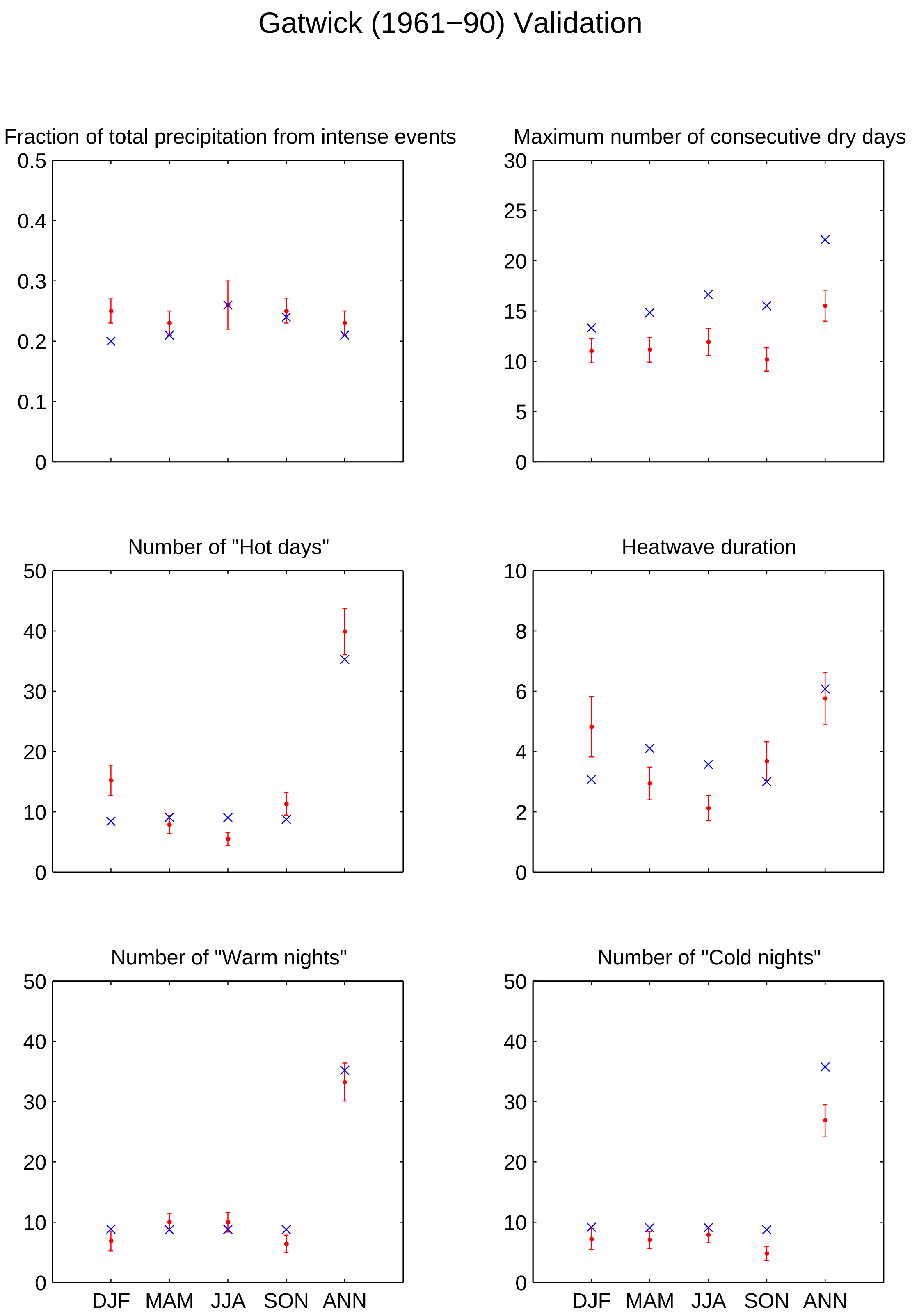
Coltishall (1961−90) Validation



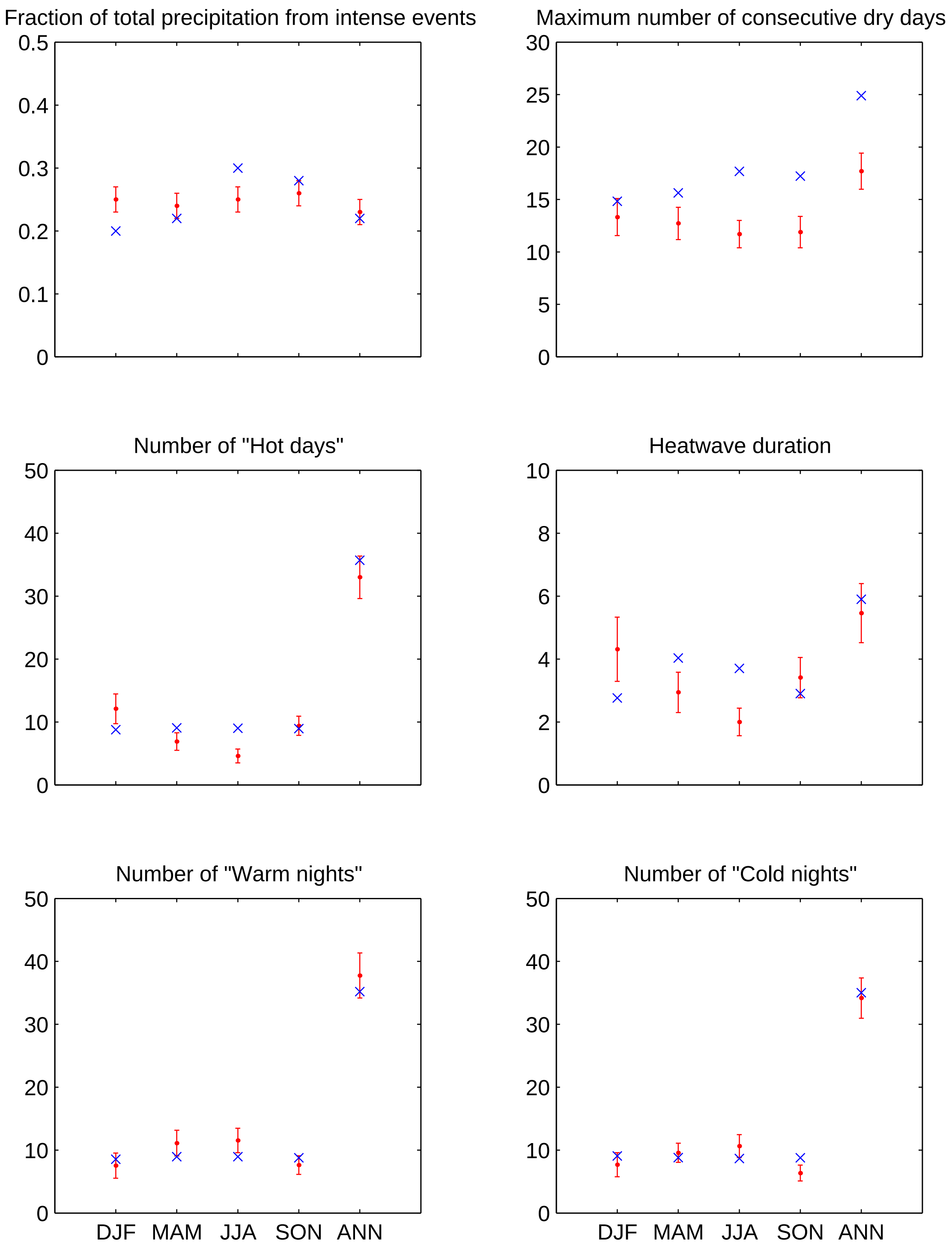
Elmdon (1961−90) Validation

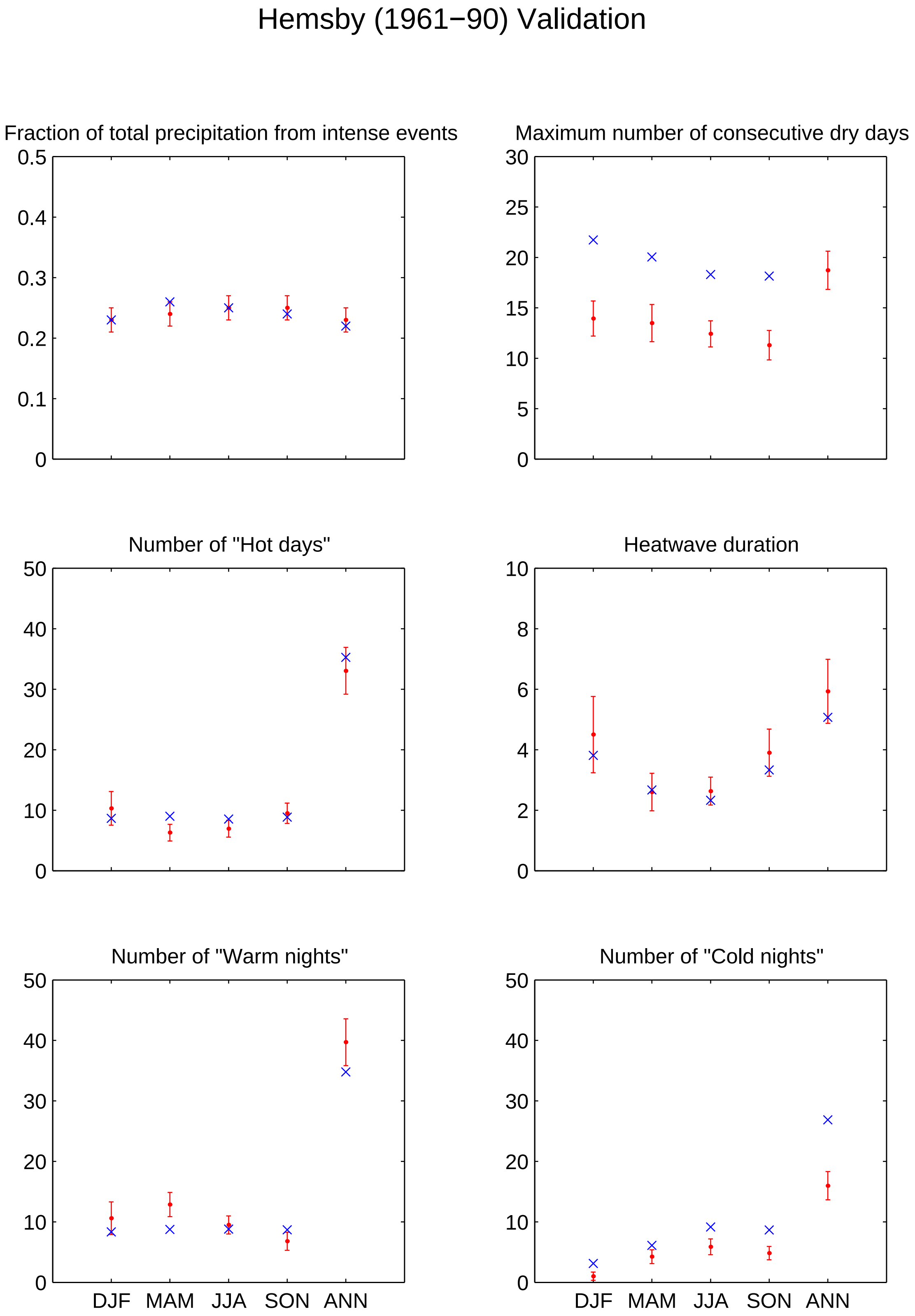


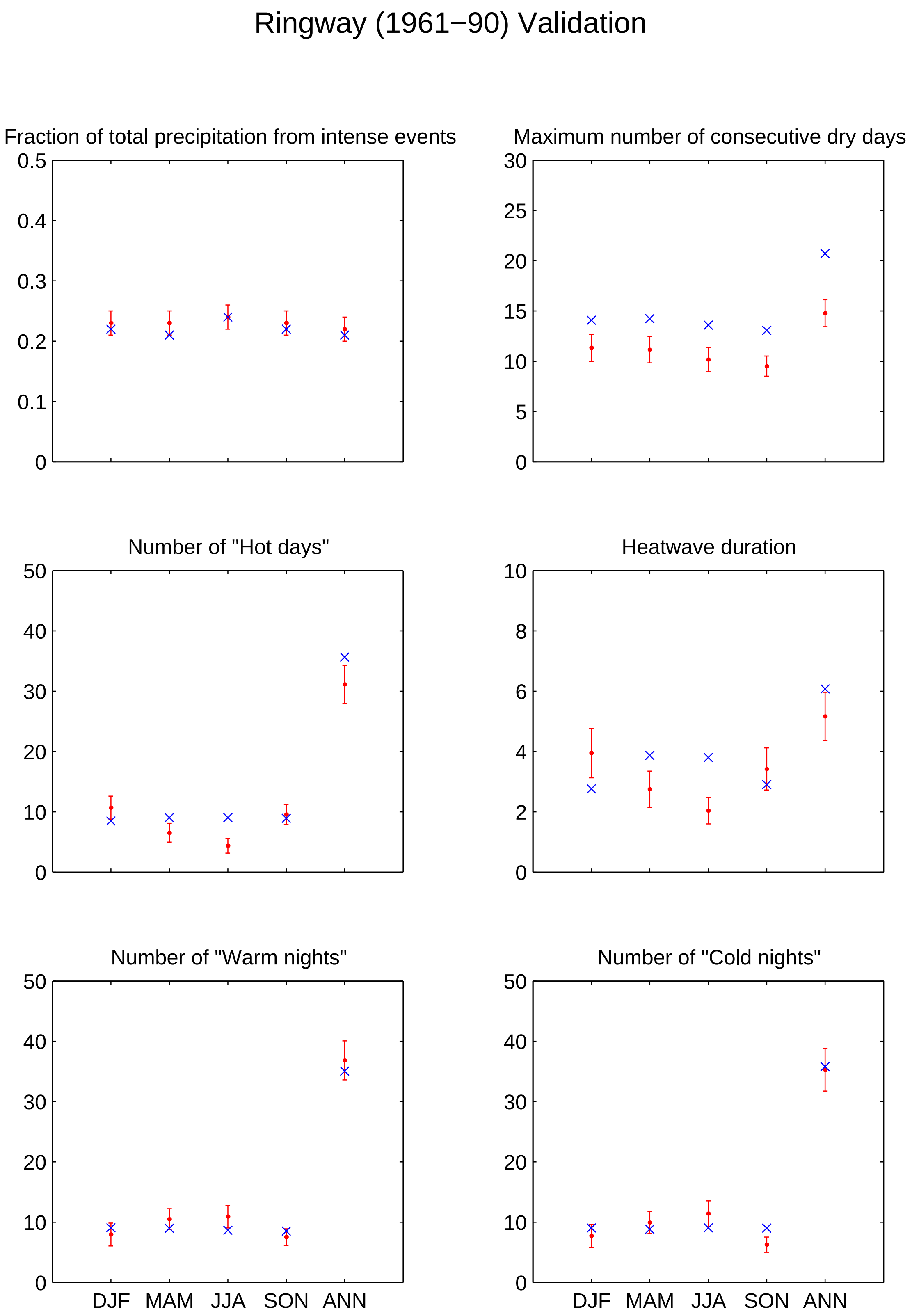

Eskdalemuir (1961−90) Validation
Fraction of total precipitation from intense events
Maximum number of consecutive dry days
Number of "Hot days"
Heatwave duration
Number of "Warm nights"
Number of "Cold nights"
DJF MAM JJA SON ANN
DJF MAM JJA SON ANN


Gatwick (1961−90) Validation
Fraction of total precipitation from intense events
0.5
0.4
0.3
0.2
0.1
0
Maximum number of consecutive dry days
30
25
20
15
10
5
0
Number of "Hot days"
50
40
30
20
10
0
Heatwave duration
10
8
6
4
2
0
Number of "Warm nights"
50
40
30
20
10
0
DJF MAM JJA SON ANN
Number of "Cold nights"
50
40
30
20
10
0
DJF MAM JJA SON ANN

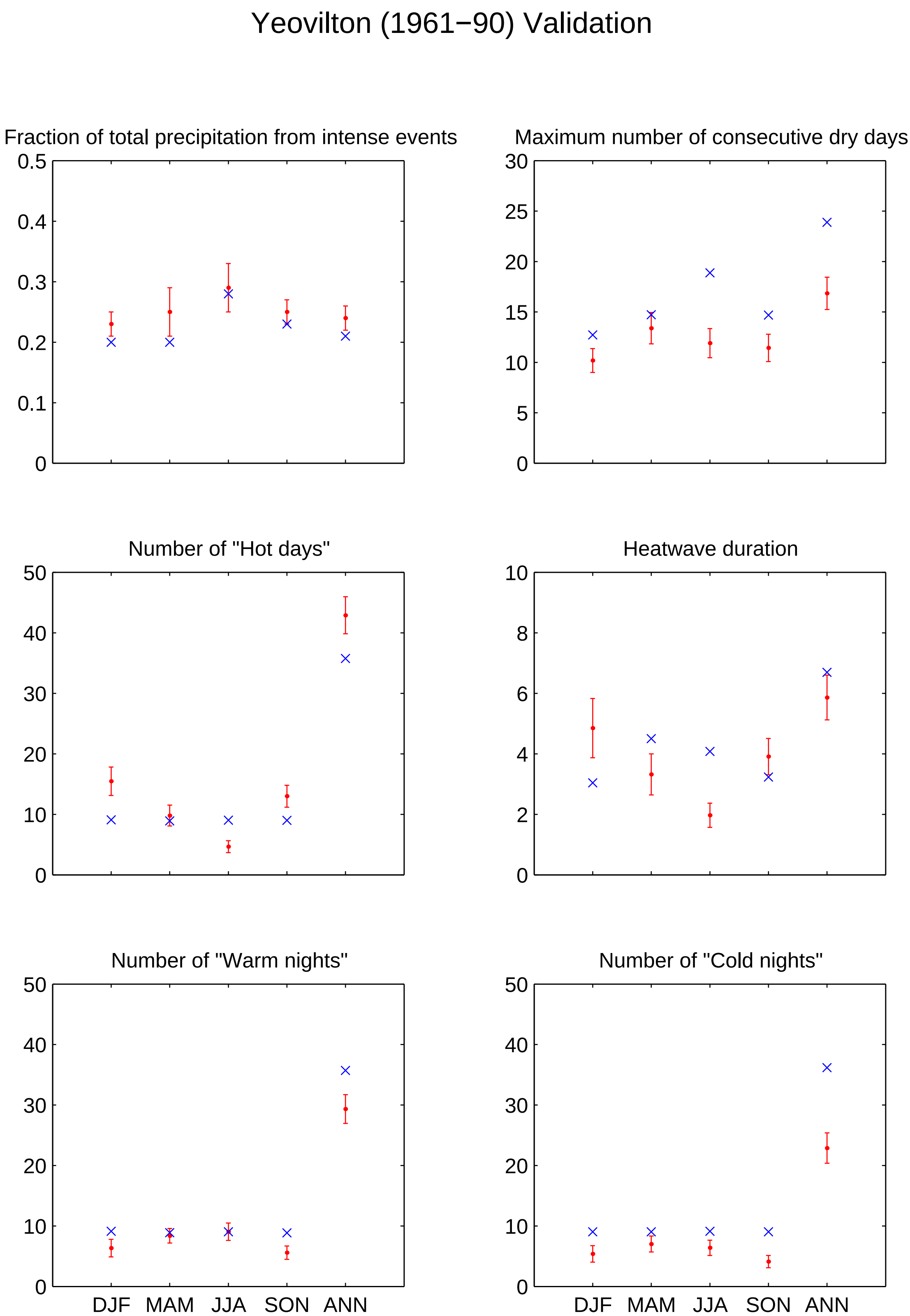
Heathrow (1961−90) Validation




Hemsby (1961−90) Validation
Fraction of total precipitation from intense events
Maximum number of consecutive dry days
Number of "Hot days"
Heatwave duration
Number of "Warm nights"
Number of "Cold nights"
DJF
MAM
JJA
SON
ANN



Ringway (1961−90) Validation
Fraction of total precipitation from intense events
Maximum number of consecutive dry days
Number of "Hot days"
Heatwave duration
Number of "Warm nights"
Number of "Cold nights"
DJF
MAM
JJA
SON
ANN


Yeovilton (1961−90) Validation
Fraction of total precipitation from intense events
Maximum number of consecutive dry days
Number of "Hot days"
Heatwave duration
Number of "Warm nights"
Number of "Cold nights"
0.5
0.4
0.3
0.2
0.1
0
30
25
20
15
10
5
50
40
10
8
6
4
2
DJF
MAM
JJA
SON
ANN