An application of HOMER for homogenising Ireland's monthly precipitation station network

John Coll[a], Mary Curley[b], Séamus Walsh[b], John Sweeney[a]

a. Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Ireland

b. Met Éireann, Glasnevin Hill, Dublin, Ireland

Coll: p.m. 03/03/14

Introduction

Motivation: Climate change studies based only on raw long-term data are potentially flawed due to the many breaks introduced from non-climatic sources, consequently quality controlled and homogenised climate data is desirable for basing climate related decision making on. Seasonal cycles of precipitation in Ireland and the UK are projected to become more marked as the climate changes, and regional extremes in summer dry spells and winter precipitation have been recorded in recent years. Therefore to analyse and monitor the evolution of precipitation patterns across Ireland, quality controlled and homogenous climate series are needed.

Aims and objectives: The primary aim of this study is to compare the reference station networks identified within HOMER with a complementary approach using correlation and other statistical measures for the series in combination with spatial scrutiny in a Geographical Information System (GIS). A secondary aim is to summarise results using the HOMER software to homogenise monthly mean precipitation (Precip) totals for Ireland for the 1941 – 2010 period for an initial set of 88 stations from the Met Éireann monthly precipitation station database.

Data

- Rainfall has been measured in Ireland since the early nineteenth century with a peak of over 800 rainfall stations in the late 1950s, and currently rainfall is recorded at synoptic and climatalogical weather stations; in addition, there is a wide network of voluntary rainfall observers (Walsh, 2012). The selected stations for the work described here are distributed across the country, but more spatial clustering of the available series is apparent in the east (Figure 1). Stations elevations were within the range of 5 – 404 m above sea level (a.s.l.) with a mean elevation of ~84 m.

Methods

- A Pearson cross-correlation was done as a standard exercise for the 88 station series, although the results are not reported here. However, and as a more refined exercise, reference series were also produced using the first difference correlation coefficients for the series, this followed suggestions in, e.g.; Peterson *et al*, 1998 and Štěpánek and Mikulová, 2008.
- HOMER was then applied to these 88 longer station series where contiguous intact monthly records ranged from ~40 to 71 years (1941 – 2010). Reference network results from HOMER were compared with first difference correlation (FDC) results in a parallel statistical approach, and scrutinised spatially in GIS.

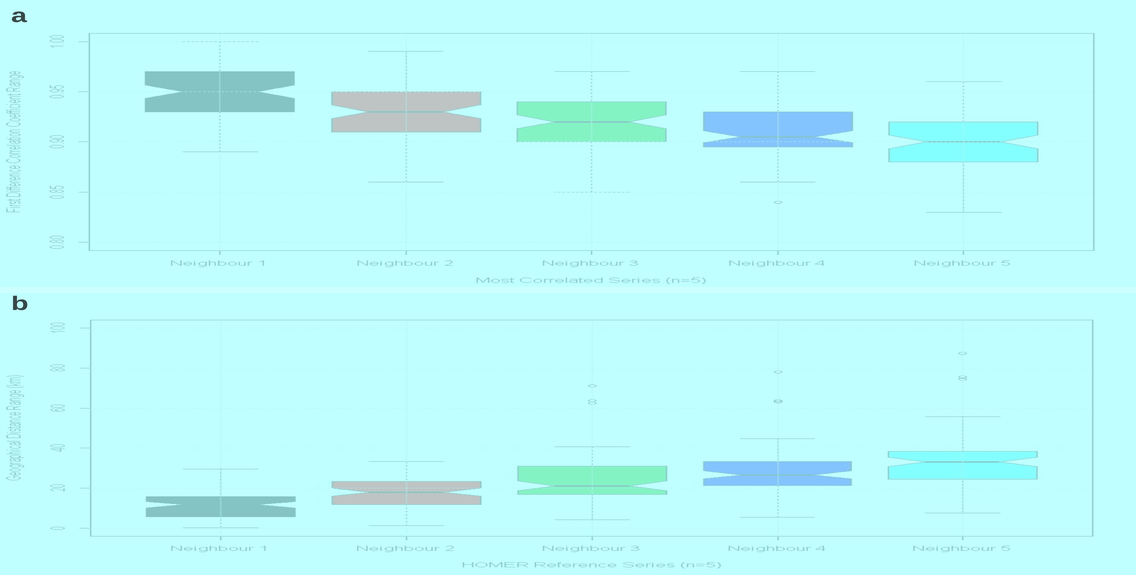

Figure 2a) Waisted box plots summarising the first difference correlation coefficient range for the base series and their five most closely correlated neighbours. The plots describe a summary analysis for 88 station and 440 potentially available neighbour series. b) The same summary plots for the HOMER-derived geographical distance range for the base series and their five nearest neighbours. Boxes: interquartile range; whiskers 5th and 95th percentiles.

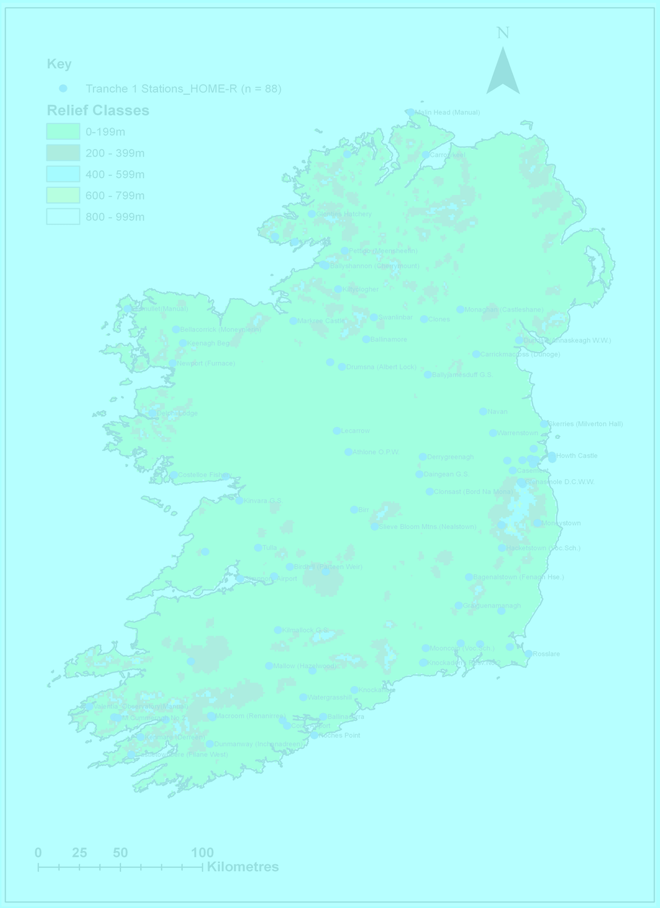

Figure 1: Annotated map of the island of Ireland showing the selected Met Éireann monthly station locations. Precipitation stations which have been homogenised using the HOMER algorithm are denoted by blue circles. Upland areas are represented by graded brown, blue and yellow shading.

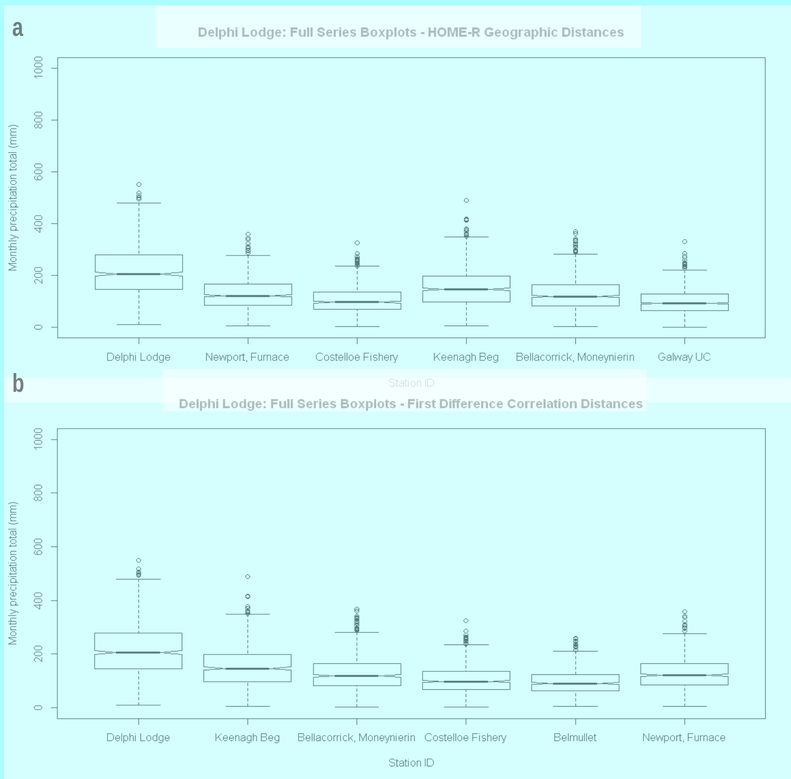

Figure 3: Box plot comparison of the base Delphi Lodge and neighbour series derived via; a) geographical distance in HOMER, and b) first difference correlation for the nearest five neighbours. Boxes: interquartile range; whiskers 5th and 95th percentiles.

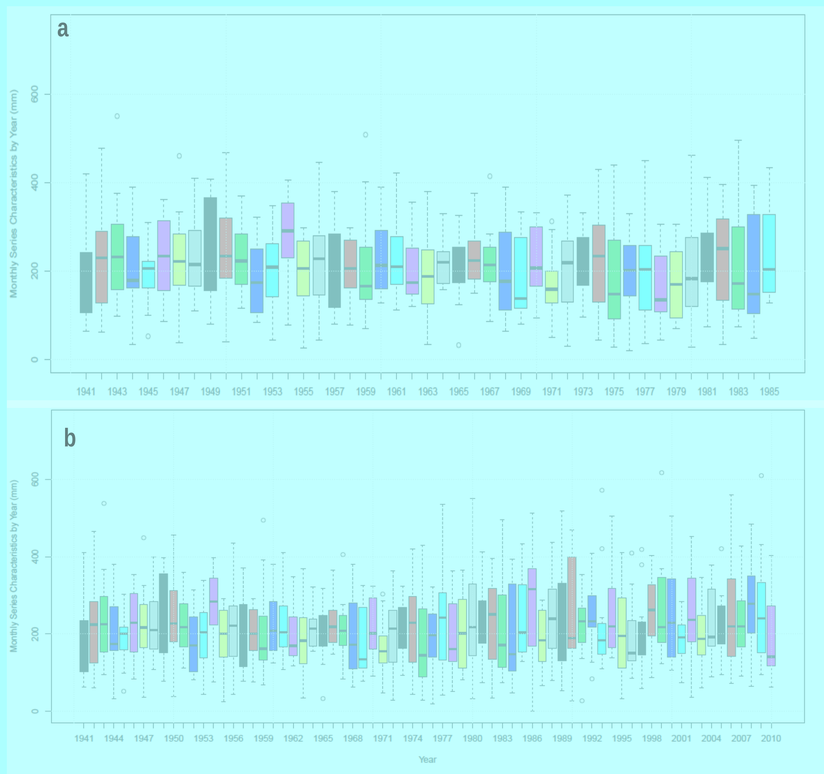

Figure 4: Time series box plots of a) the available Delphi Lodge annual series (1941-1985), and b) of the HOMER homogenised data for the same series extended with reference to the neighbour data (1941-2010). Boxes: interquartile range; whiskers 5th and 95th percentiles.

Results and discussion

1. HOMER Geographical Distance and First Difference Correlation Compared

- For both approaches, the statistical properties of the base and derived neighbour series are very similar. First difference correlation coefficients ranged from r = 0.93 – 0.97 for the reference series identified (Figure 2a), whereas geographical distances for the HOMER-derived series were 9.50 km – 24.29 km (Figure 2b). For all 88 of the candidate base series, results obtained between both methods identified largely similar reference stations and series.
- Overall therefore, both HOMER and FDC returned climatologically coherent and statistically similar neighbour series, and show that HOMER geographical distance selections identify statistically coherent neighbour series overlapping those provided by FDC. For the western case study example outlined here both approaches identified four reference series in common; Keenagh Beg, Bellacorrick, Costelloe Fishery and Newport (Figure 3). For both sets of results, the statistical properties of the base and neighbour series are largely coherent, although Delphi Lodge is a wetter station than the neighbours. First difference correlation coefficients ranged from r = 0.88 – 0.92, whereas geographical distances were 27.57 km – 48.35 km.

2. Homogenisation results

- HOMER accurately detected breaks in 11 of the 88 monthly series, of which 6 were confirmed from the metadata. Based on the neighbour series selected on the basis of geographical distance, and after the metadata checks and user interaction with the software, HOMER provided a new and homogenised output for the focal base series based on the PRODIGE comparison technique. An examples is provided here for one case study series (Delphi Lodge) where the original series have been extended using HOMER-derived values based on the neighbour series (Figure 4) . The data are presented as annual time series box plots to highlight that the HOMER-derived missing values are statistically coherent with the data range and other characteristics of the original series. In the example here, the 1986-2010 segment for Delphi Lodge, the additions are HOMER-derived extensions based on neighbour station data (Figure 4).

Conclusions and future prospects

- FDC was used to critically evaluate HOMER results output in a parallel statistical computing framework. HOMER has accurately detected breaks in 11 from the 88 monthly precipitation of which 6 were confirmed from the metadata. Results show that both approaches yield valid and statistically similar corresponding neighbour series.
- Ongoing analysis extending the functionality HOMER offers indicate excellent prospects for the work going forward and the ongoing work remains novel and is certainly the first such application for Irish monthly precipitation series.

References

Peterson TC, Easterling DR, Karl TR, Groismanm P and others (1998). Homogeneity Adjustments of *in Situ* Atmospheric Climate Data: A Review," *International Journal of Climatology*, 18 (13), 1998, 1493-1517.

Štěpánek P, and Mikulová K (2008). Homogenisation of air temperature and relative humidity monthly means of individual observation hours in the area of the Czech and Slovak Republic. In: *5th Seminar for Homogenisation and Quality Control in Climatological Databases.* Hungarian Met. Service, Budapest; 147-163.

Walsh, S (2012). *A summary of climate averages for Ireland 1981-2010*. Met Éireann Climatological Note 14. Dublin. 16pp.

Contact: john.coll@nuim.ie

Acknowledgements: We thank the Irish Environmental Protection Agency (EPA) for the project funding and support. This research was supported under grant 2012-CCRP-FS.11.