**Data**

Areal and point rainfall data used in this analysis.

Data used is monthly rainfall records for the UK sourced from the UK Met Office [1].

This data was recently updated as a result of the citizen science project 'Rainfall Rescue' [2] and covers the period January 1836 to December 2021 for each of the ten district regions as defined by the UK Met Office [3].

Downloaded as a .text file for each region and converted to .csv format, then merged into a single file covering all regions and indexed to enable join with shapefiles.

Monthly UK station rainfall data also sourced from the UK Met Office. Not all data from all stations used as length of record for each station is different. Took a period where majority of stations have a complete record (Sep’64 – Aug-16). Gaps in the data for individual months for individual stations filled to create continuous record to avoid issue of NAs when working with the data in R. Gaps filled by taking same month in previous year (t-12) for same station as a ratio of all other stations in previous year and applying that ratio to the current month’s data for all stations

where S is the rainfall for all the other stations

Shapefile for Great Britain district regions sourced from CEGE0042 Tutorial data. Northern Ireland sourced from OSNI Open Data [4]. Combined into a single shapefile.

Shapefile and rainfall data joined in ArcGIS Pro to create a geodatabase feature layer for import into R.

**Exploratory Data Analysis**

Spatial and temporal characteristics of the data analysed in numerous ways

Spatial Analysis

Spatial analysis indicates UK rainfall increases the further north and west

Moran’s I for GB (not N.Ireland as this creates an empty neighbour set) weather stations of 0.516 within range of -0.593 to 1.018, so appears significant. Moran Test and Moran.mc both give p-values <0.05, confirming statistical significance

Lowest local autocorrelation values in S Wales & England SW and England NW & N Wales. These are wet regions in the west with long borders with dry regions in the east, hence low autocorrelation. Highest local Moran’s I values seen in East Anglia, a dry region in the east bordering other dry regions. Also, Scotland N, a wet region in the north-west bordering other wet regions.

Spatial autocorrelation in GB weather station point data using variogram shows scattered result but indication is that rainfall levels at weather stations that are closer are more similar than those further away. No clear results from the directional variograms.

Temporal Analysis

Temporal analysis shows no long-term trend, but seasonality present over a period of 12 months

Temporal autocorrelation is not great, with one month’s rain less strongly correlated to previous month’s rainfall than seen with UK temperatures. Scotland North shows highest temporal autocorrelation with a PMCC of 0.306; Midlands shows lowest PMCC of 0.081 (1 month lag interval).

Looking at annual data for UK weather stations also shows week temporal autocorrelation (PMCC = 0.121), indicating one year’s rainfall is not significantly related to the previous year’s rainfall.

Back to regional data, ACF for Scotland N shows seasonality with statistically significant positive autocorrelation peaks at lags of 12, 24, 36, 48 months, negative peaks at 6, 18, 30, 42 months: non-stationary with peaks at fixed intervals.

PACF for Scotland N region shows significant positive results at lag 12 and 24, negative at lags 5, 6, 18 – again demonstrates seasonality and suggests need to include a seasonal autoregressive term in ARIMA / STARIMA

No statistically significant PACF for annual UK weather station data, suggesting annual rainfall is random from one year to the next (not looked at individual weather stations)

Spatio-Temporal Analysis

Spatio-temporal ACF for GB regions shows similar seasonal pattern as for temporal analysis. No statistically significant STPACF

ST semivariogram point UK weather station point data shows the semivariance increases rapidly with increasing temporal separations and then becomes more complex with some peaks and troughs. No equivalent increase in semivariance with increasing spatial separations, but spatial and temporal separations are not comparable.

**Statistical Modelling of Time Series and Spatio-Temporal Series**

**Artificial Neural Networks**

Form of Supervised Learning – labelled training data used to predict labels for unseen data

References

[1] Met Office (2022). UK and regional series. Available at https://www.metoffice.gov.uk/ research/climate/maps-and-data/uk-and-regional-series. Accessed 04/04/2022

[2] Hawkins, E., Burt, S., McCarthy, M., Murphy, C., Ross, C., Baldock, M., et al (2022) Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists. Geoscience Data Journal, 00, 1– 16. Available from: https://doi.org/10.1002/gdj3.157

[3] Met Office (2022). UK climate districts map. Available at https://www.metoffice.gov.uk/research/climate/maps-and-data/about/districts-map. Accessed 04/04/2022

[4] Open Data NI (2022). OSNI Open Data – 50K Boundaries – NI Outline. Available at https://www.opendatani.gov.uk/dataset/osni-open-data-50k-boundaries-ni-outline. Accessed 04/04/2022