EMS Annual Meeting Abstracts
Vol. 14, EMS2017-498, 2017


# Automated reproducible analysis for climatological reports using R

Sandra Spillane
Met Éireann, Climatology and Observations Division, Dublin 9, Ireland (sandra.spillane@met.ie)

Traditionally in Met Éireann, Ireland's meteorological service, the production of climatological periodicals involved a number of manual steps – initially using SQL to access database, analysing through Excel, finally producing monthly/seasonal/annual weather summaries through QuarkXPress as PDF reports. This involved repetitive steps and analysis using knowledge acquired over years by "trial and error" of the data that was stored in the climate database. The methodology has changed very little since 1986 and updates/enhancements were not easily implemented. Furthermore the report cannot be regenerated without completing identical analysis, interpretation, copying and pasting. The need for automated and reproducible research has become a matter of urgency as the knowledge acquired over years vanishes with each staff movement or retirement.

This paper outlines the development of a next generation collaboration tool, AWARE - Automatic Weather Analysis using R for Events, which automates the production of the climatological periodicals, limited only by the availability of observational data and/or interpretative analysis.

R is traditionally used to explore data and produce graphical displays but is capable of conditional text manipulation also. The output is limited only by what the reader might find interesting or appealing. Another historical challenge involves the amount of data available and on how to present this information in a non-confusing and interesting way. A reduction in the number of stations analysed was one strategy used in the past as tools were limited. While using R, as data becomes available – from archive rescue projects or as quality control processes finish, the system can be rerun to produce reports immediately.

AWARE is capable of coding-in the knowledge of staff and the structure of the climate database together with seamless updates from a team of contributors, real-time status reporting/sharing and enhancements with maps, graphs, tables and code-deployed text analysis. Completed using the R package with output in multiple formats (html, pdf, etc). With the final goal being to produce a reproducible climatological report, saving time and ensuring that the best possible analysis is completed on each climate variable in an efficient and appealing manner.