The Australian Rainfall & Runoff Datahub

Mark Babister[1,2], Aaron Trim[1], Isabelle Testoni[1], Monique Retallick[1]

[1]WMAwater, Sydney, Australia
E-mail: babister@wmawater.com.au

[2]Visiting fellow, University of Technology, Sydney, Australia

Abstract

The Australian Rainfall and Runoff (ARR) revision projects have produced a large number of spatial design inputs that practitioners need to access in order to undertake design flood estimation. These inputs will be updated as improvements in terms of data record and methodology are made or anomalies are addressed. The ARR data hub www.data.arr.org.au was created to provide a one stop shop for practitioners to access current inputs in a simple easy manner. The online data hub has the advantage of documenting the version of the data used and allowing improved reproducibility of past results. This new approach represents a significant shift in practice with practitioners accessing data at the start of a study and software vendors not embedding datasets within their software platform. This paper documents the data hub and the methods used to sample the data set.

1. INTRODUCTION

Australian Rainfall and Runoff (ARR) is historically a guideline document, series of maps and table containing design inputs and parameters. In the current edition, ARR (Ball et al, 2016) has evolved into a guideline document, spatial datasets and enabling software. This paper documents the online data hub software that allows easy access to the spatial data sets.

All the design input that are typically used in flood estimation were updated in ARR 2016, and the data hub was seen as an effective way of accessing data from a single source. The datahub also removed the need to read parameters from maps which has been shown to be a minor source of inconsistencies. The most important aspect is that it allows practitioners to access the latest design inputs and for updates to be easily delivered.

2. REGIONAL APPROCHES

Historically ARR has used design inputs that are based on large regions with distinct changes at the region boundaries. In many case the regions are based on jurisdictional boundaries rather than catchment characteristics or meteorological similarity. Region of influence style approaches allow hard boundaries to be removed and national spatial data sets allow some parameters to be estimated at any location. Some inputs recommended in ARR 1987 (Pilgrim, 1987), like areal reduction factors, were based on overseas data.

The Australian Hydrological Geospatial Fabric (Geofabric) (BoM, 2012) is a specialised Geographic Information System data set that allows practitioners to determine the basin, catchment and river system at any point in Australia. On a river system it easy to determine the upstream catchment area and stream network.

The design input used in ARR are either based on regions or have been calculated based on locations at a predetermined resolution. In some cases the approach spatially smooths the data while in other cases output is spatially smoothed.

3. DESIGN INPUTS

The new design input from ARR are listed in Table 1 along with the regionalisation approach and the data resolution.

Table 1: Design Inputs

Design Input	Spatial Approach	Resolution or Source
River Region	Region	Geofabric
Intensity Frequency Duration	Location	0.025 degrees from regionalised pluviographs and then gridded with a spline.
Areal Reduction Factors	Region	
Losses	Location	Spatially averaged at 0.15 degree
Point Temporal Patterns	Region	Hydrologic enforced NRM regions
Areal Temporal Patterns	Region but borrowing from neighbours often needed	
Pre Burst	Location	0.1 degrees
Climate Change Factors	Region	NRM regions
Baseflow	Location	Based on main stream upstream catchment

4. OUTPUTS FROM THE DATAHUB

There are ten (10) design inputs that can be extracted from the datahub website. These outputs are explained in Table 3.

Table 2. Descriptions of Outputs from the Datahub

Output	Type of Output
River Region	Provides the river region for the location input. If a shapefile has been input the region of greatest intersection with the first polygon in the uploaded layer for all regions is shown.
BOM IFD Depths	http://www.bom.gov.au/water/designRainfalls/revised-ifd/ calculated from ARR Book 2 – Chapter 3
ARF Parameters	ARR Book 2 – Chapter 4. The Short Duration ARF equation is taken from equation 2.4.1. Long Duration gives the a to i parameters from Table 2.4.2 for equation 2.4.4.
Storm Losses	ARR Book 5 – Chapter 3. Initial Loss and Continuing Loss values for the grid cell the point is contained within. If a catchment shapefile is uploaded it will take the proportional average of the intersected losses grid cells.
Temporal Patterns	ARR Book 2 – Chapter 5. The temporal pattern region the catchment or point is located within. Provides a link to download a zip file containing the temporal pattern data in a CSV.
Areal Temporal Patterns	ARR Book 2 – Chapter 5. The temporal pattern region the catchment or point is located within. Provides a link to download a zip file containing all of the areal pattern data in a CSV.
Median Preburst Depths and Ratios	ARR Book 2 – Chapter 5. Table of median preburst depth and ratios for standard AEPs and durations.

Other Preburst Depths and Ratios	As above, but for the 10%, 25%, 75% and 90% distributions
Interim Climate Change Factors	A table of temperature increases and percentage increase in rainfall for a set of forecast years and RCP 4.5, 6 and 8.5 emissions schemes (CSIRO and BoM, 2015)
Baseflow Factors	ARR Book 5 – Chapter 4. Data for the baseflow factors in a region, including area and r3, r1 runoff factors. If a catchment shapefile is uploaded the baseflow region with the greatest intersection is used. It is advised to use a point at the catchment outlet when collecting this data.

5. SAMPLING APPROACHES

The data hub allows inputs to be sampled at the any point In Australia by the use of latitude and longitude or a catchment shape file. Point inputs are generally used on small catchments and located at the catchment centroid except for baseflow which is calculated based upon the outlet location. Where a shape file is used the catchment average value is returned, though on large catchments the practitioner could sample at a number of point values so the catchment average value can be spatially distributed.

Table 3 below list the sampling approaches used when shape files are used.

Table 3: Sampling Approaches

Data	Source	Catchment Polygon value format
River Region	Geofabric shapefile (BoM, 2012)	Maximum intersection with catchment polygon
BOM IFD Depths	No file, links to BoM website	N/A
ARF Parameters	Long duration ARF shapefile (Podger et al 2015a)	Maximum intersection with catchment polygon
Storm Losses	Initial loss and Continuing Loss grids (Hill et al, 2016)	Proportional average of grid cells intersecting with catchment polygon
Temporal Patterns	Temporal Patterns Regions shapefile and Temporal patterns (as csv) (Loveridge et al, 2015b and Testoni et al, 2016)	Maximum intersection with catchment polygon
Areal Temporal Patterns	Areal Patterns (as csv) download (Podger et al, 2016)	N/A
Median Preburst Depths and Ratios	Preburst depths netCDF file (Loveridge et al, 2015a)	Average of grid cells with centroids contained within catchment polygon
Other Preburst Depths and Ratios	Preburst depths netCDF file (Loveridge et al, 2015a)	Average of grid cells with centroids contained within catchment polygon
Interim Climate Change Factors	NRM Regions shapefile (Bates et al, 2015 and CSIRO and BoM, 2015)	Maximum intersection with catchment polygon
Baseflow Factors	Baseflow Factors shapefile (Murphy et al, 2011a,b)	Maximum intersection with catchment polygon. It is advised not to use a polygon for this output and instead use

		a point at the catchment outlet.

6. INTERFACE

For spatial files the input file format needs to be a shapefile in latitude and longitude projection (not Map Grid Australia) to work with the hub, and all component files of a shapefile must be uploaded including the shx and dbf files.

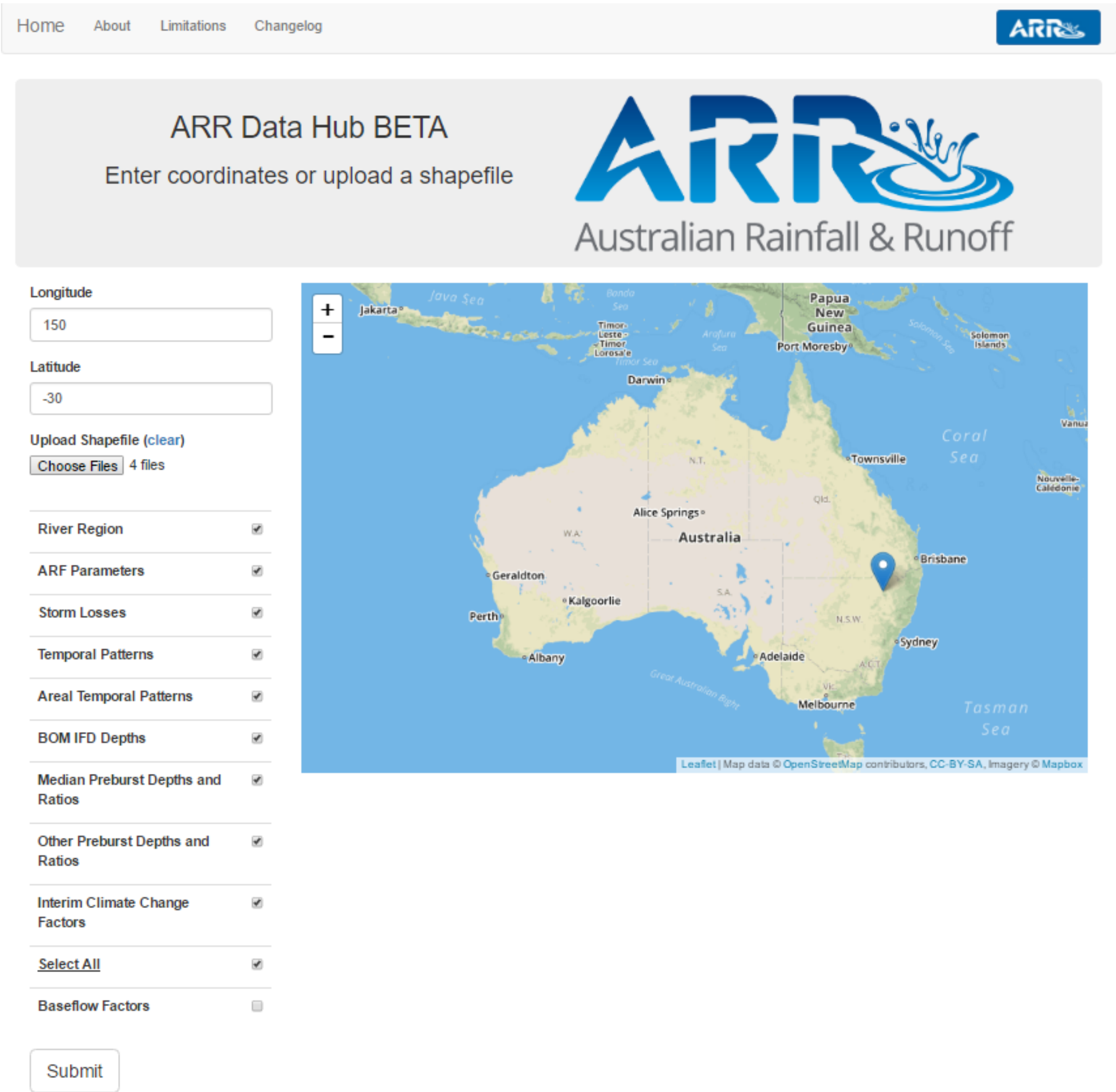

Figure 1. screen shot of the ARR data hub landing page.

API requests can also be sent to the server directly through the URL. The results of this style of request yield a text file similar to the one downloadable on the regular results page, with the temporal pattern increments also displayed at the end of the file. An example URL request to obtain data at (-29.573, 152.648) in latitude and longitude for River Regions and Temporal Patterns is as follows:

http://data.arr.org.au/?lon_coord=152.648&lat_coord=-29.573&River_Region_SinglePart.sqlite=on&TP_Regions.sqlite=on

The output interface allows the user to compare the point or shape file to the underlying data resolution for each parameter. This allows the user to interactively check if any parameters are near a boundary. Figure 2 show an example results page with the location of interest close to a boundary.

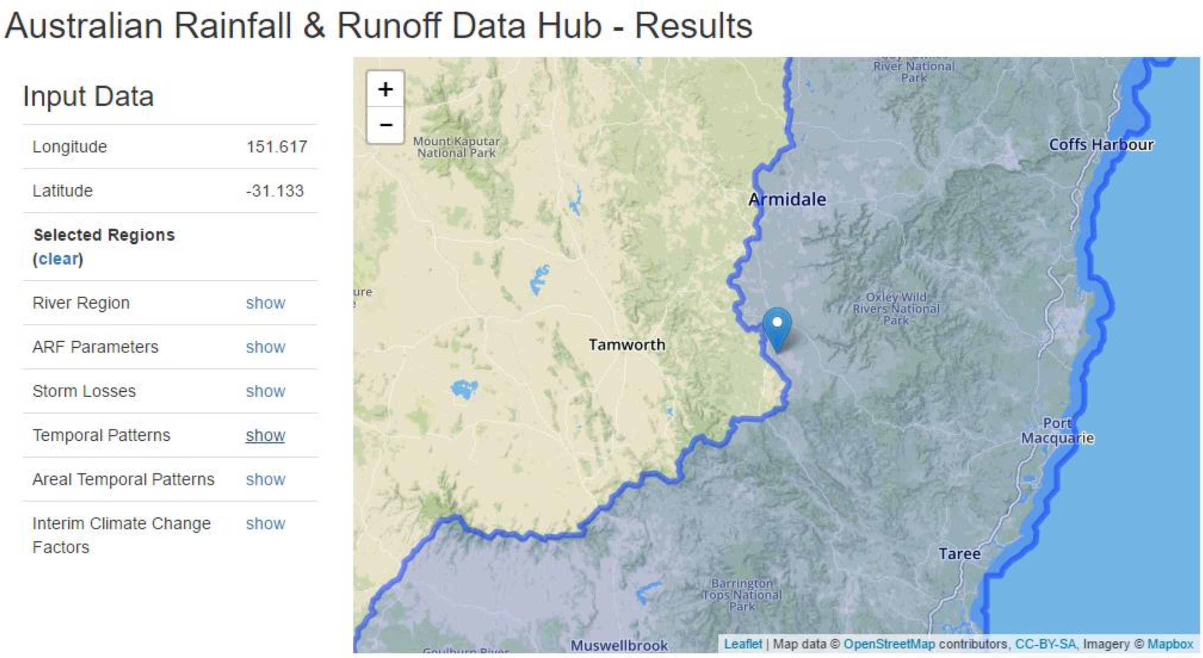

Figure 2. ARR Datahub results page

As well as the web based output that is deign to work on a range of devises a text based format has been provided that software can read and PDF format that can included in reports

As version control is a key aspect to moving to an online platform, the datahub has made it as easy as possible to check which version of the data was accessed. Figure 3 shows an example output of the design losses information with the meta-data on the right hand side. These shows the time at which the data was accessed through the site and the version number. The meta-data is also shown on the PDF output to be put into reports so that a practitioner can reproduce the same results with the right version.

Storm Losses

Initial Losses (mm)	54.0
Continuing Losses (mm/h)	0.4

Layer Info

Time Accessed	14 September 2016 05:42PM
Version	2016_v1

Figure 3. Example output with meta-data

7. CHANGES TO INCREASE COMPUTATIONAL SPEED

It was initially proposed to read the raw data formats of the underlying data sets but changes were made to increase speed, particularly for user uploaded shape files as this required multiple calculations.

Data set shapefiles were converted into SQLite databases with each polygon converted into a database row. The first columns of the database contain the attributes of the shapefile, and the final column contains the polygon data stored in Well-Known Text (WKT) format. This allowed the polygons with centroids within a particular range (+- 10%) to be quickly retrieved which speed up the analysis of the intersection between points/polygons and the target layer.

NetCDF files were converted to HDF5 files through h5py library formatting. This greatly increased the read speed of the files while remaining in the same multi-dimensional array format.

The formats used are presented in Table 4 below

Table 4

Data	Source	Format
River Region	River_Region shapefile	SQLite
BOM IFD Depths	No file, links to BoM website	N/A
ARF Parameters	Final_ARF_Regions shapefile	SQLite
Storm Losses	Combined_Losses_Final shapefile (IL + CL shapefile combined)	SQLite
Temporal Patterns	TP_Regions shapefile, TemporalPatternsRegions zip files	SQLite, CSV
Areal Temporal Patterns	ArealPatterns.zip download	CSV
Median Preburst Depths and Ratios	Preburst depths netCDF file	HDF5
Other Preburst Depths and Ratios	Preburst depths netCDF file	HDF5
Interim Climate Change Factors	NRMRegions shapefile (BoM)	SQLite
Baseflow Factors	Baseflow_Factors shapefile	SQLite

8. CONCLUSIONS

The ARR data hub provides a one stop efficient way of accessing all of the design inputs and documenting the versions used. It will allow software providers and practitioners to access information in a simple manner and provides a platform for updates to be easily rolled out.

The datahub also allows practitioners to document the design inputs used which will greatly aid reproducibility. It uses version control so that many councils and agencies will require practitioners to include summary PDF output in reports.

9. REFERENCES

Loveridge, M., Babister, M. & Retallick, M. 2015(a). Australian Rainfall and Runoff: Revision Project 3 Temporal Patterns of Rainfall – Part 2: Analysis of Pre-burst Rainfall, Report No. P3/S3/013.

Loveridge, M., Babister, M. & Retallick, M. 2015(b). Australian Rainfall and Runoff: Revision Project 3 Temporal Patterns of Rainfall – Part 3: Preliminary Testing of Temporal Pattern Ensembles, Report No. P3/S3/013.

Ball J, Babister M, Nathan R, Weeks W, Weinmann E, Retallick M, Testoni I, (Editors) Australian Rainfall and Runoff: A Guide to Flood Estimation (2016) Commonwealth of Australia

Green, J. Xuereb, K. Johnson, F. Moore, G. The, C. (2012), *The Revised Intensity-Frequency-Duration (IFD) Design Rainfall Estimates for Australia – An Overview*, 34th Hydrology and Water Resources Symposium, Sydney, Australia.

Podger, S., Green, J., Jolly, C., The, C. and Beesley, C. (2015a),Creating long duration areal reduction factors for the new Intensity-Frequency-Duration (IFD) design rainfalls, Proc. Engineers Australia Hydrology and Water Resources Symposium, Hobart, Tasmania, Australia.
Podger, S., Green, J., Stensmyr, P. and Babister, M. (2015b), Combining long and short duration areal reduction factors, Proc. Engineers Australia Hydrology and Water Resources Symposium, Hobart, Tasmania, Australia.

Stensmyr, P., Babister, M. and Retallick, M. (2014), Australian Rainfall and Runoff Revision Project 2: Spatial Patterns of Rainfall: Stage 2 Report, Short Duration Areal Reduction Factors, ARR Report Number P2/S2/019, ISBN 978-085825-9614.

Hill, P., Zhang, J. and Nathan, R. (2016), Project 6: Loss Models for Catchment Simulation, Stage 4A. Report prepared for the Australian Rainfall and Runoff Technical Committee.

Testoni, I. Babister, M. Retallick, M. Loveridge, M. (2016), Regional Temporal Patterns, Accepted for presentation at the 37th Australian Hydrology and Water Resources Symposium, Queenstown, New Zealand.

Podger, S. Babister, M. Brady, P. (2016), Deriving Temporal Patterns for Areal Rainfall Bursts, Accepted for presentation at the 37th Australian Hydrology and Water Resources Symposium, Queenstown, New Zealand.

Bates, B., Evans, J., Green, J., Griesser, A., Jakob, D., Lau, R., Lehmann, E., Leonard, M., Phatak, A., Rafter, T., Seed, A., Westra, S. and Zheng, F. (2015), Development of Intensity-Frequency-Duration Information across Australia - Climate Change Research Plan Project. Report for Institution of Engineers Australia, Australian Rainfall and Runoff Guideline: Project 1. 61p

CSIRO and Bureau of Meteorology (2015), Climate Change in Australia, Projections for Australia's NRM Regions. Technical Report, CSIRO and Bureau of Meteorology, Australia. Retrieved from www.climatechangeinaustralia.gov.au/en [http://www.climatechangeinaustralia.gov.au/en].

Murphy, R.E., Graszkiewicz, Z., Hill, P.I., Neal, B.P., Nathan, R.J. (2011a), Predicting baseflow contributions to design flood events in Australia, Proceedings of the 34th IAHR World Congress, 26 June - 1 July , Brisbane, Australia.

Murphy, R., Graszkiewicz, Z., Hill, P., Neal, B., and Nathan, R. (2011b), Project 7: Baseflow for catchment simulation, Phase 2 - development of baseflow estimation approach. Report prepared for the Australian Rainfall and Runoff Technical Committee by Sinclair Knight Merz, ARR Report Number P7/S2/017, ISBN 978-0-85825-916-4.

Bureau of Meteorology, (2012), *Australian Hydrological Spatial Geospatial Fabric (Geofabric): Topographic Drainage Divisions and River Regions*