**RESEARCH LOG**

**Lachlan Stoney**

**Work for Russell Drysdale**

**MUNDRABILLA DATA**

Obtained daily and monthly rain data for mundrabilla (1901-2016) from:

http://www.bom.gov.au/jsp/ncc/cdio/weatherData/av?p\_nccObsCode=139&p\_display\_type=dataFile&p\_stn\_num=011008

stored in rus/nullabor/mundrabilla/

3 files: daily, monthly1, monthly2 (not actually named like that).

data format = .csv

monthly 2 has a slightly different format (12 months per row, annual rainfall included).

stored in /rus/nullabor/mundrabilla/original\_data/

Mundrabilla rainfall data has a GAP from 1971 to 1984, and some missing data in 2012 (January and 2012 total rainfall amount).

Also downloaded rainfall data for the town EUCLA. Starts at 1876.

stored in /rus/nullarbor/eucla/

**OPENING THE RAIN DATA**

Matlab command csvread() only works on files with purely numerical data, but the BOM data has some text.

Wrote AWK script **rain\_filter.sh** to send relevant .csv data (the important numerical stuff from 1984-2015) to new .dat files

Usage: "bash rain\_filter.sh" (if script and .csv files are in the correct directories)

Currently this script creates .dat files in /rus/nullabor/mundrabilla/filtered\_data/

i.e. one file for January rainfall, Feb rainfall, yearly rainfall, yearly rainfall etc.

Wrote a line in this script to add an entry for Jan 2012 (with missing data). Otherwise there is a gap.

Configure this script to filter for particular months, etc.

Plan to use this script to select only summer rainfall days with > 5mm.

Wrote matlab script **read\_rain\_data.m** to read the new .dat files , test some plots etc. Basic.

**GRIDDED DATA**

**NCEP2** data (Jan 1979-July 2015) downloaded from:

ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis2.derived/gaussian\_grid/skt.sfc.mon.mean.nc

^ skin temperature. Stored at /rus/ncep2/SST/

<ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis2.derived/surface/mslp.mon.mean.nc>

^ mean sea level pressure. Stored at /rus/ncep2/MSLP/

Land values were masked for the skin temperature data. Missing value attribute added to the SST and MSLP files. See the command history with:

ncdump -h /rus/ncep/SST/SST.nc (this is the final modified file)

**NCEP1** data (Jan 1948 – March 2016) downloaded from:

ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis.derived/surface/slp.mon.mean.nc

^ sea level pressure. Stored at /rus/ncep1/MSLP1

Data was cut down to the Jan 1948 - Dec 2015 time period with the command:

ncks -F -d time,1,816 slp.mon.mean.nc slp.nc

slp.nc is the final modified file

ftp://ftp.cdc.noaa.gov/Datasets/ncep.reanalysis.derived/other\_gauss/ulwrf.ntat.mon.mean.nc

^ upward longwave radiation flux (Same as OLR, but model-calculated). Stored at /rus/ncep1/ULRF/

Data was cut down to the Jan 1948 - Dec 2015 time period with the command:

ncks -F -d time,1,816 slp.mon.mean.nc ulrf.nc

ulrf.nc is the final modified file

NOAA interpolated **OLR** data (June 1974- December 2013) downloaded from:

<ftp://ftp.cdc.noaa.gov/Datasets/interp_OLR/olr.mon.mean.nc>

^ outgoing longwave radation. Stored at /rus/OLR/

**ERSST** data (Jan 1854 – March 2016) downloaded from:

ftp://ftp.cdc.noaa.gov/Datasets/noaa.ersst/sst.mnmean.v4.nc

^ sea surface temperature. Stored at /rus/ERSST. Data was cut down to the Jan 1903 – Dec 2015 time period with the command:

ncks -F -d time,577,1944 sst.mnmean.v4.nc ersst.nc

There is also ERSST data going back to 1876 in the directory /rus/ERSST\_1876/ . This is for calculating the correlations with Eucla rainfall data. And I downloaded ERSST v3 for another test, stored in /rus/ERSST\_v3/ - This data includes satellite measurements, unlike v4.

For each of the downloaded files above (or the modified version thereof), I created monthly,yearly,seasonal data with the script **reformat\_field.sh**

This calls other scripts: month.sh, year.sh,summer.sh,winter.sh

Example: “bash reformat\_field.sh SST.nc” creates one SST file for each month (i.e. every january from 1984-2015 is in the january file rus/NCEP2/SST/month\_1SST.nc), and for summer, winter, and yearly. The new files may have different lengths, since there may be 37 Januaries in the data set but only 36 Decembers.

There should be a copies of these shell scripts in each data folder. They will be identical except for reformat\_field.sh, which will be configured for the particular data type (e.g. reformat\_field\_ersst.sh is configured for ersst data).

The reason the data was split up this way is that the spatial correlation (see next section) can be performed with one rainfall data file (e.g. jan\_filtered.dat) and a corresponding gridded data file (e.g. month\_1ersst.nc). I also did it this way, rather than defining the time series for each month and season inside matlab, because the netcdf files can be inspected for consistency in ncview (i.e. check that the dates are correct and inspect missing values, etc.). Additionally, it saves time (but takes more disk space) to have the monthly and seasonal data already defined before the matlab script starts running.

**CORRELATIONS**

Wrote matlab function s**patial\_correlation.m**

Accepts one rain fall file (i.e. from /rus/nullabor/mundrabilla/filtered\_data/ ) and one gridded data file (e.g. from /rus/ncep/SST/). Both need to have same number of time steps. The code needs to be modified a little, depending on which ncep data is being read (change variables 'lines' and 'P'). *It is crucial to get this right – sometimes the script will run even if the wrong data is being read*.

Usage: spatial\_correlation(rainfile,gridded\_file,type,gridded\_variable)

This function will spit out a correlation field for the mundrabilla rainfall time series vs the gridded data. Currently the plots only show the correlation values that are statistically significant.

The code can be configured to adjust the nature of the correlation calculation. e.g. detrending the data, first difference calculations, Pearson's correlation vs Rank correlation, statistical significance. Currently the script is configured to first calculate first differences on both the rain and gridded data, and then calculate Spearman correlations.

spatial\_correlation.m uses the plotting routines in m\_map (downloaded from <https://www.eoas.ubc.ca/~rich/map.html> and stored in rus/m\_map). Had some trouble with m\_map only showing contour data for positive longitudes. Not a problem if you focus on Australian region.

**Issues with spatial\_correlation.m**:

Sometimes you get an error message like:

“Warning: Rank deficient, rank = 0, tol = inf.”

This has something to do with nan\_detrend not handling zero values (maybe dividing by zero). Although I am not currently using nan\_detrend. a work around is to add a large arbitary constant to the gridded data. The covariance of two time series is independent of the absolute values or the value of the mean, so adding an arbitrary constant doesn't matter. I tested this; I get identical plots if I add a constant or not (assuming the error message doesn't appear in the first place).

May need to add /rus/ and subdirectories to matlab path, otherwise custom functions might not work.

**Notes on correlation results**

Overall, the summer and yearly correlations are quite weak, or at least not statistically significant. Even statistically significant random correlations can pop up with probability=p. This is inevitable when many tests are being performed (i.e. at each grid point, and for each month and variable). Lack of summer significance possibly due to sporadic summer rainfall in the nullabor, and lack of “negative rain” associated with low SST, high pressure, high OLR, etc. Winter correlations show statistically significant patterns that are consistent with NW cloud-bands (using ncep1 mslp,ulrf, and ersst). Individual months don't show much significance. Perhaps the rainfall is dominated by short term weather fluctuations that aren't captured in the monthly data sets.

Some preliminary plots in /rus/nullabor/plots/ ...compared these with similar plots using the NINO3.4 index instead of the mundrabilla rain record. This gives the expected ENSO pattern with high and significant correlations. See plots labelled 'nino\_test.png' etc. This test can be run inside spatial\_correlation.m

Repeated the seasonal correlations with rainfall data from Eucla (1876-present). Similar results: north west trough patterns clear in winter. Not much in summer.

Wrote the matlab function **correlation\_test.m** to calculate the correlation field for two renanalysis files. This can be tested against

cdo -b 32 timecor file1 file2

Tested both methods with the files SST.nc and pres.sfc.mon.mean.nc and they gave the same results. Need to re-check this.

**SYNOPTIC TYPING**

Wrote awk script **high\_rainfall\_days.sh** to filter the daily mundrabilla rain for days with summer rainfall exceeding 5mm. Currently only considering measurements that have been quality checked. Between 1979 and 2016 there are 180 odd such days. The time stamp refers to the measurement made at 9am of that day, of the previous 24hrs of rain.  
  
For 1999-present, can find archived synoptic charts at: <http://www.bom.gov.au/australia/charts/archive/>

radar loops at Eucla, which is near Mundrabilla:

http://www.theweatherchaser.com/radar-loop/IDR453-eucla

satellite imagery:

<http://www.australiasevereweather.com/links/ozsatpic.htm>

<http://www.ncdc.noaa.gov/gibbs/>

nice Nw cloudband on Jan 1st 2006, march 18th 2000 (TC olga)

