Multisite Rainfall Downscaling (MRD) Help

Table of Contents

1. Overview

The Multisite Rainfall Downscaling (MRD) is a tool that conditionally simulates rainfall occurrence and amount by using the modified Markov model-kernel density estimate (MMM-KDE) (Mehrotra and Sharma, 2010). The conditioning variables include previous day rainfall/state, aggregated wetness state over a period of time and atmospheric variables. By ignoring atmospheric variables, model reduces to standard multisite rainfall simulator (MRS) (Mehrotra et al., 2015). The package provides a stochastic modeling framework for multisite downscaling of daily rainfall incorporating low-frequency variability in the simulations. This framework simulates daily rainfall occurrences and amounts as a two stage process at each individual location. The rainfall occurrence is modeled with a Markov chain conditional on the rainfall occurrence of previous days, wetness state over a predefined past aggregated days and atmospheric variables. The rainfall amounts on the wet days are simulated using a kernel-density estimation procedure conditional on the previous day's rainfall and atmospheric variables.

The spatial dependence across stations is simulated by making use of spatially correlated random numbers. Spatial correlation in the random numbers is introduced on the basis of the at-site observed cross correlations in rainfall occurrences and amounts. The downscaling framework is presented here in the form of an interactive tool and is named as Multisite Rainfall Downscaling (MRD).

2. Features

The MRD has been tested using data from 30 raingauge stations around Sydney, Australia. Additionally, reanalysis and regional climate model outputs have been used as conditioning atmospheric variables. The analyses of the testing results show that MRD adequately captures daily and aggregated long time period rainfall characteristics at individual locations including the spatial distribution of rainfall over the region.

Features of MRD include:

- Displays input time series and stochastically simulated/downscaled data graphically.

- Option to generate a matrix for each day to be used to generate spatially correlated random numbers. It requires to be generated only once for a given dataset and stations.

    - Displays the statistics of observed and simulated/downscaled data with tabulated values and plots (i.e. boxplots for the statistics at any individual station, and qq plots for the mean statistics of all the stations).

    - All the tabulated statistics and plots can be saved to a file of various formats.

3. Data requirements

3.1 Input data files

MRD requires a continuous time series of daily rainfall and/or atmospheric data as input data.
The supported format of input data file is written in **simple text format** (with file extension of .dat, .data or .txt).
For rainfall, the file should follow the following sequences (Figure 1a).
The first three rows record the station number, station name and station number index. The first three columns indicate the year, month and day and are followed by the rainfall values at all the stations under consideration.

	StnNumber		70080	63039	63033	63036	63095	63224	70131	70002	70055	68007
	STN Name		TARALGA	KATOOMBA	GURNANG	OBERON	YRRNDRIE	LITHGOW	WOODHSE	BANNABY	GOLBRN	CAMDEN	HIRANGE
Y	M	D	1	2	3	4	5	6	7	8	9	10	11
1979	1	1	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	2	.00	.00	.00	.00	.10	.00	.00	.00	.00	.00	.00
1979	1	3	.20	3.00	.00	.00	2.80	2.80	.00	.00	.20	3.40	.00
1979	1	4	.00	.20	.00	.00	.10	.60	.00	.00	.00	.00	.00
1979	1	5	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	6	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	7	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	8	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	2.40
1979	1	9	.00	.00	.00	.00	2.39	.00	.00	.00	.00	.00	.00
1979	1	10	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	11	.00	.00	.00	.00	.00	.00	.20	.00	.20	.00	.00
1979	1	12	.00	.00	.00	.00	.20	.00	.20	.00	.00	.00	.00
1979	1	13	.00	.80	.00	.00	.30	.00	.00	.00	.00	.00	1.00
1979	1	14	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	15	.00	.00	.00	.00	.10	.00	.00	.00	.00	.00	.00
1979	1	16	.00	.80	.00	.00	.20	.00	.00	.00	.00	.00	.00
1979	1	17	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
1979	1	18	5.20	9.20	6.94	6.81	4.40	5.40	4.00	3.00	1.20	9.00	3.20

Figure 1a. The first 6 rows of a sample rainfall data file ("rain.dat" is used for demonstration purpose).

If the option of downscaling is selected i.e. atmospheric variables are also included then information on the atmospheric variables is also needed to be included in the 'data.dat' file. This information is

provided through three files and their names are to be included at the end of 'data.dat' file. The first file contains the total number of atmospheric variables at each station location identified as significant for the rainfall downscaling at that location. It is named as 'pred_all.dat' in the sample data file. It also includes the names of current (calibration) and future (application) climates data file at that station location. These files have data structure very similar to the rainfall data file as shown in Figure 1b. These files contain daily values of all identified atmospheric variables. For current climate the time period coincides with the rainfall time period while for future time period it could be of different time periods and years.

The second and third files contain the details of conditioning atmospheric variables for each station and season for rainfall occurrences and amounts, respectively. These identified atmospheric variables may vary with season and rainfall occurrence and amount processes and are subsets of the atmospheric variables included in the first file for each station location. These files are named as 'pred_occr.dat' and 'pred_amt.dat' in the sample 'data.dat' file provided with the package. If identified variables do not change across the stations, it is possible to keep the same file for different station locations as found in the example included in the package.

Year	Month	Day	GPH@925	TDP@850	EPT@700	UWND@850	VWND@700	VV@850	EW Grad of MSLP
1979	1	1	800.51	11.06228	311.3924	-6.89	-0.38	0.02666	-0.39996
1979	1	2	813.99	13.96441	311.421	-9.25	1.64	-0.00418	-0.57001
1979	1	3	822.51	19.98273	311.0858	-12.64	1.21	0.01684	-0.35999
1979	1	4	827.27	22.64675	310.9522	-9.71	2.2	-0.03794	-0.23004
1979	1	5	825.64	22.26785	310.9011	-8.73	2.23	0.00339	-0.56995
1979	1	6	821.76	15.06361	312.5382	-7.3	-5.89	-0.00589	-0.58002
1979	1	7	791.31	12.42099	311.9411	-5.1	-5.94	-0.00392	-0.5
1979	1	8	778.62	12.60832	312.3056	-2.76	-3.49	-0.07569	-0.58002

Figure 1b. The first 6 rows of a sample atmospheric data at a given location ("WRF_re.DAT" is used for demonstration purpose).

In addition to rainfall data, user is required to provide the station-distance information in a file in a simple text format. First line is distance of station 1 with all stations; second line is distance of station 2 with all and so on. As distance of station with itself is zero, this matrix will have all diagonal values as zero. Station sequence will be the same as user would have in the rainfall data file. A typical distance file for five stations will look like:

0 120 80 40 15

120 0 60 72 12

80 60 0 59 8

40 72 59 0 96

15 20 8 96 0

Figure 2. A sample data format of station-distance file for five stations ("dist.dat".is used for demonstration purpose)

The data in both files is in free format i.e. only separated by spaces or comas.

3.2 Input parameters

Parameter Names	Comments
Number of years (ob.)	Number of years of rainfall data from observations used for the development of model and estimation of model parameters. Maximum is 30.
Start year (ob.)	Start year of observations. Can take any reasonable value.
Number of stations	Number of stations used in simulation. Minimum is 1 and maximum is 30.
band	Indicating the width of moving window (i.e. band*2+1). Minimum is 1 and maximum is 21.
lag	Time lags considered by the Markovian process to generate the conditional probabilities based on rainfall state of previous time step(s). Can take either 0 or 1.

local	Representing the influence of local neighborhood during rainfall amount generation. Can take any number from 0, 1 and 2. If 0 is taken, then local influence is ignored. If 1 is taken, local wetness is considered as a conditioning variable. If 2 is taken, selected data having similar wetness is used in kernel density estimation.
nLon	Number of time periods. Maximum allowed is 3.
Number of previous lags for nLon	Number of previous time lags (in days) for nLon. Maximum is 900.
Number of years (sim.)	Number of years of simulations to be generated. Can take any value between 1 and 150.
Start year (sim.)	Start year of simulation generated. Can take any reasonable value.
Number of simulations	Number of simulations generated. Maximum is 100.
Simulate rainfall occurrence only	If ticked, the rainfall amount will not be simulated.
Include spatial dependence matrix	If unchecked, spatial dependence is ignored.
Include atmospheric variables	If unchecked, simple rainfall generation, no downscaling

The input parameters and the name of files can be either provided through the interactive graphical user interface (i.e. Figure 10) or through the modification of the file "data.dat". The value of each parameters need to be **tab delimited**. It always helps to keep the original sample data file at a safe location and then modify it to suit individual requirements.

4. Technical specifications

Operating system:
Windows 7

Essential applications:

R 3.0.2 (download from http://cran.r-project.org/bin/windows/base/)

5. Running Multisite Rainfall Simulator/Downscaling (MRD)

5.1 Getting started

1. Download the zipped folder "MRD", right click on it and select **Extract All…** as highlighted in Figure 2.

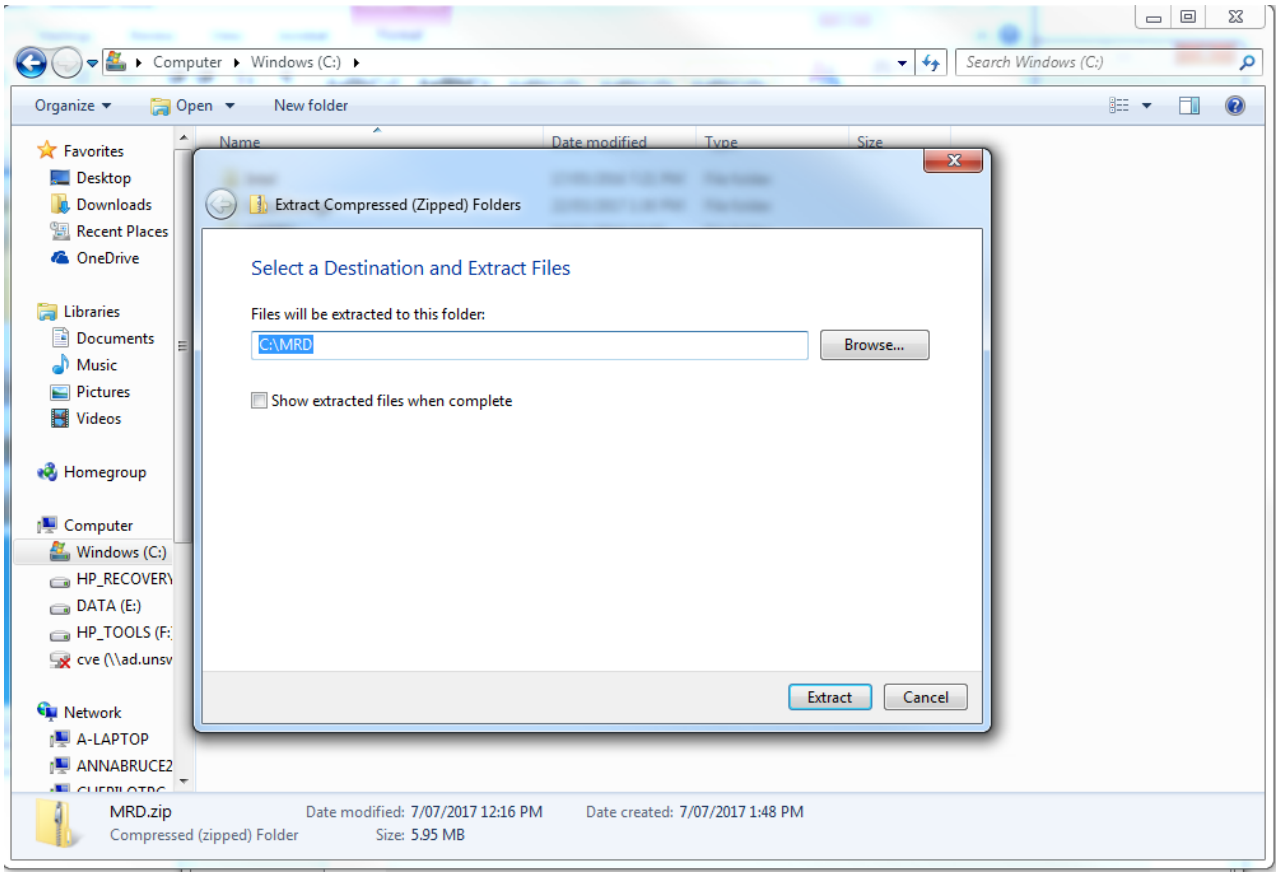


Figure 2. Extract All.

2. On the Extract Compressed (Zipped) Folder window, click **Extract**.

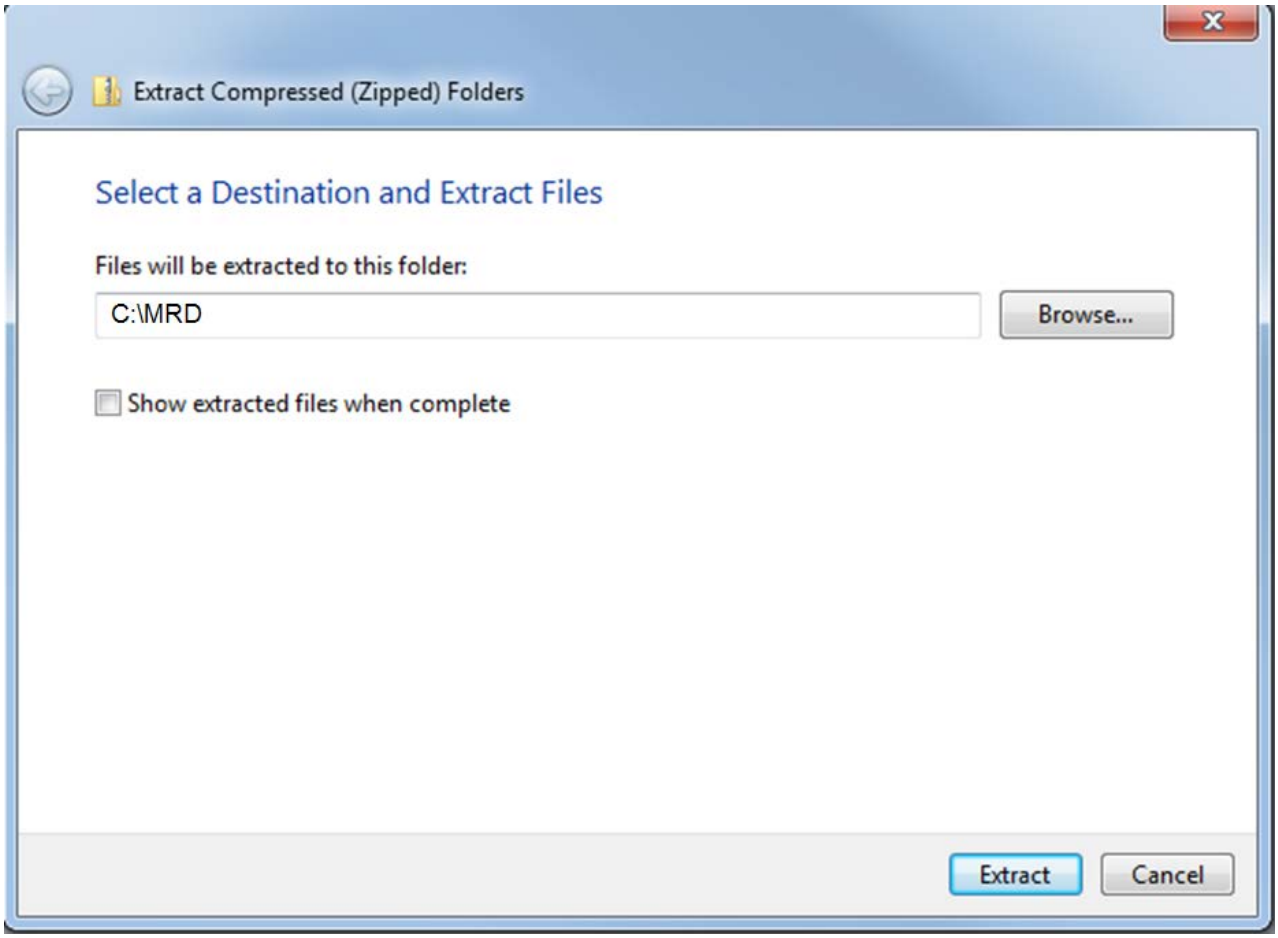


Figure 3. Extract Compressed (Zipped) Folder window.

3. Open R, click **File** on the menu bar, and click **Change dir…** as highlighted in Figure 4.

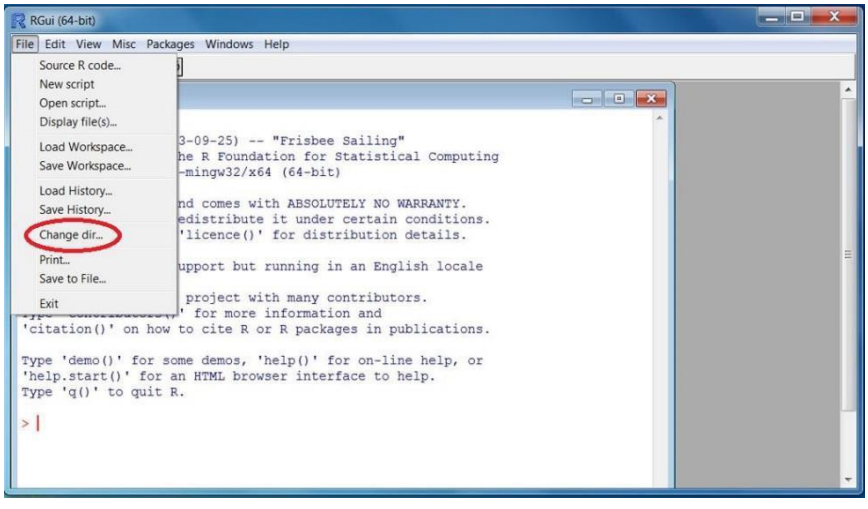


Figure 4. Change directory selection.

4. Select the extracted folder “MRD” as highlighted in Figure 5, and click **OK**.

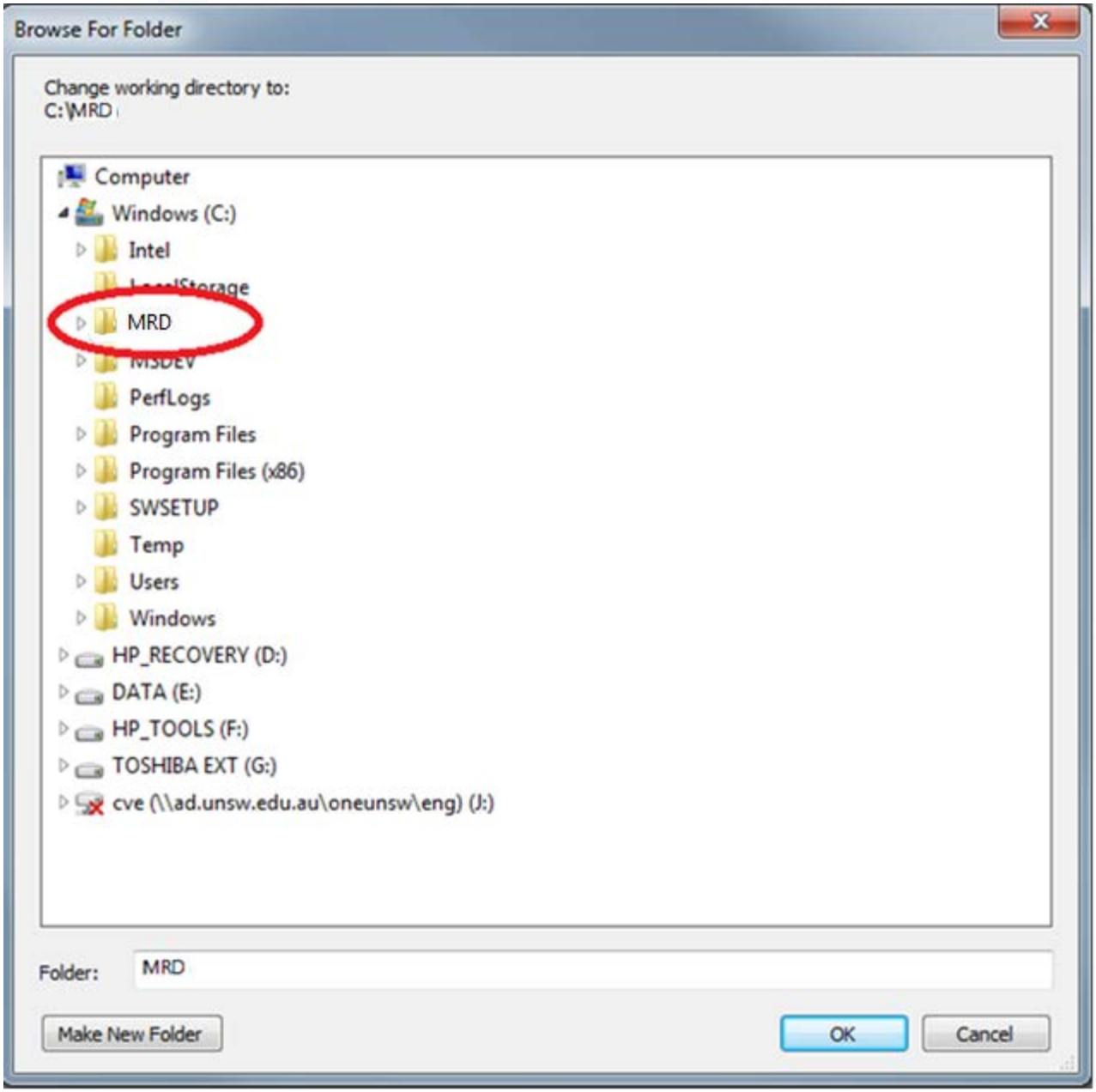


Figure 5. Browser for Folder window.

5. Click **File** on the menu bar, click **Source R code…** as highlighted in Figure 6, and click **OK**.

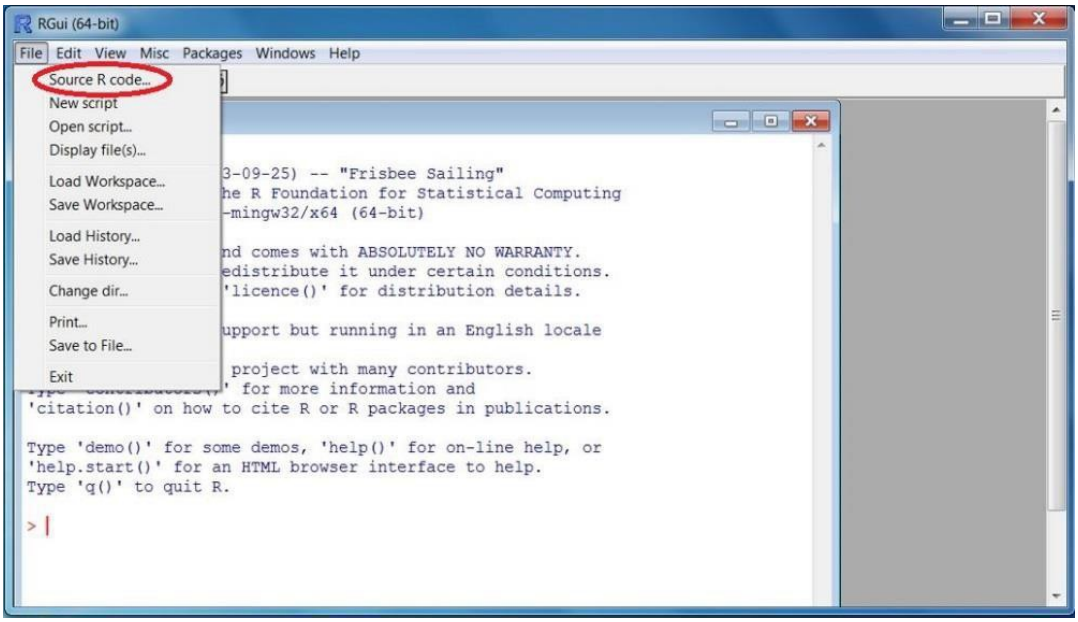


Figure 6. Source R code selection.

6. Select “simulator.r” as highlighted in Figure 7, and click **Open**.

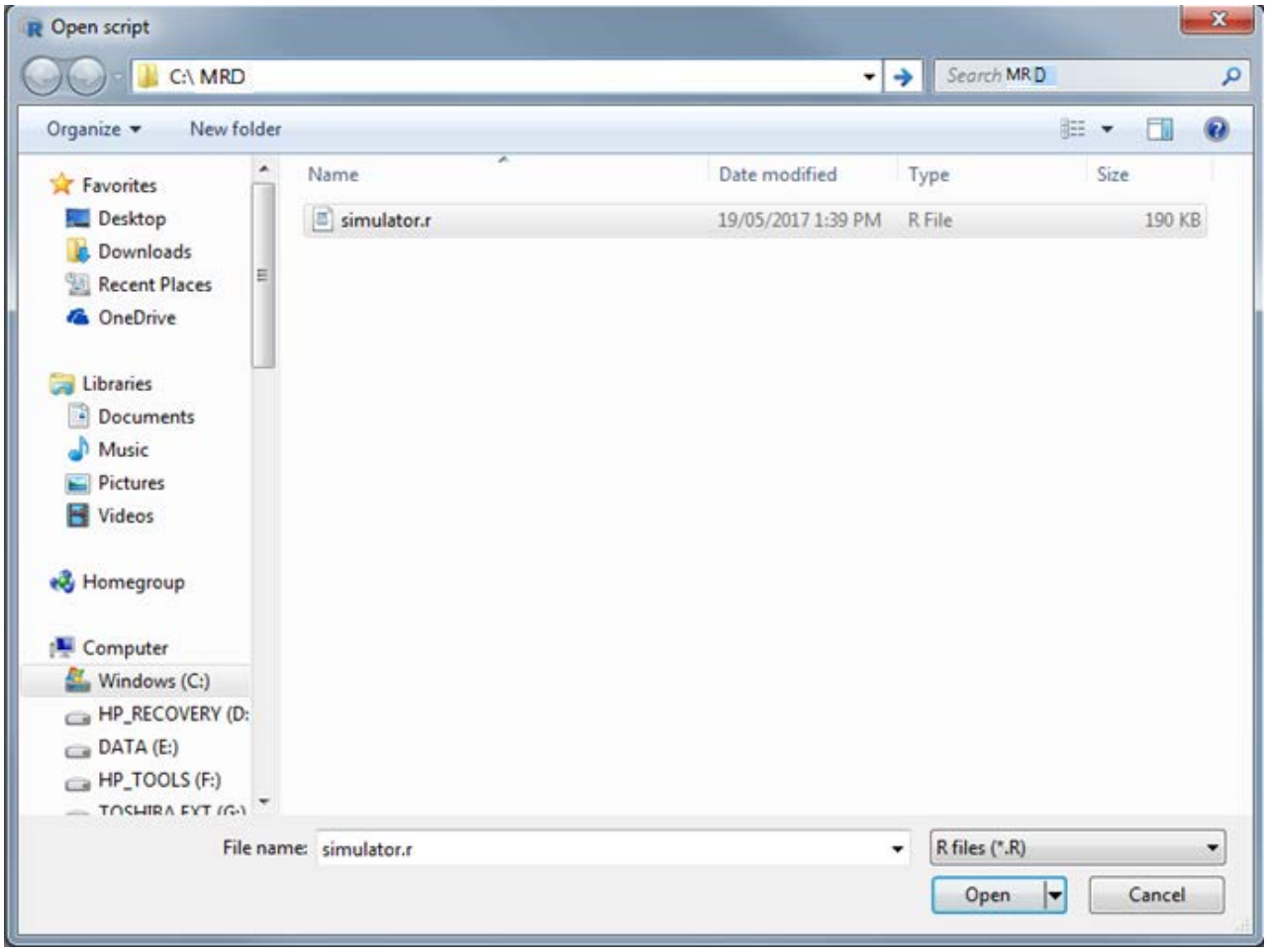


Figure 7. Select file to source window.

A graphical user interface will be generated as shown in Figure 8.

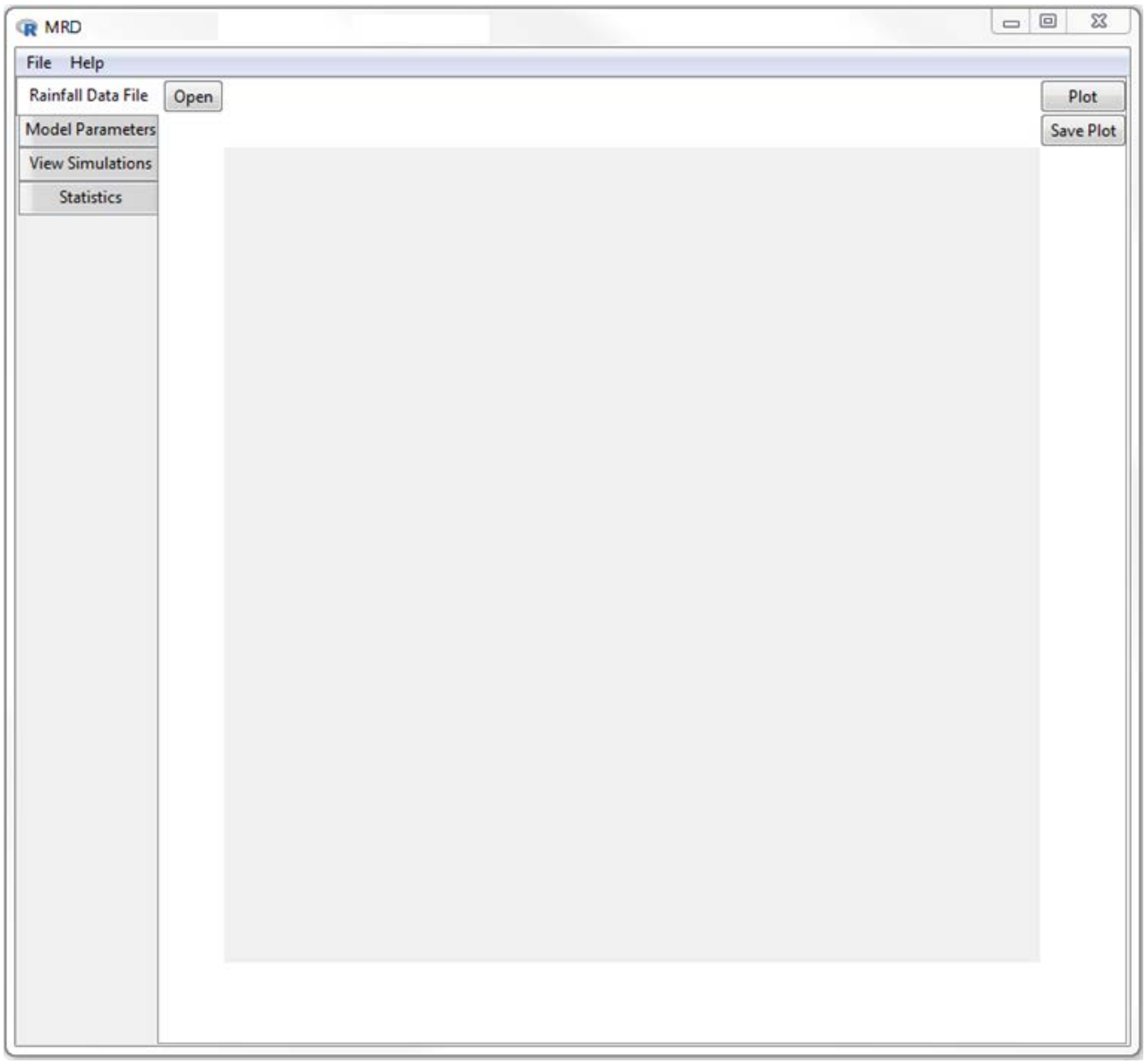


Figure 8. Initial view of the graphical user interface.

5.2 Steps in running the MRD

1. Click the "Open" button to select a daily observation rainfall data file with the format described in section 3.1. For the purpose of demonstration, the file "rain.dat" can be selected as the input observational rainfall data as shown in Figure 9.
2. Click the "Plot" button to plot the monthly rainfall of the input data.
3. Click the "Save Plot" button to save the graph into various formats. (Step 2 and 3 is optional.)

4. Click the “Model Parameters” tab to input parameters and then click the “Estimate” button (Figure 10). If the user wants to use their own data file, the corresponding distance file need to be provided by the user (Figure11) and should be of the same format as the file “dist.dat”.The estimation will take about 10 minutes or so for 30 stations, although exact time will vary depending on the number of simulations and the computer processor.

5. Click the “View Simulations” tab to view and save each of the replicates (Figure 12).

6 Click the “Statistics” tab to select a statistic option to view as either tabulated values or plots (Figure 13 & 14). Each statistic view can be saved into a file.

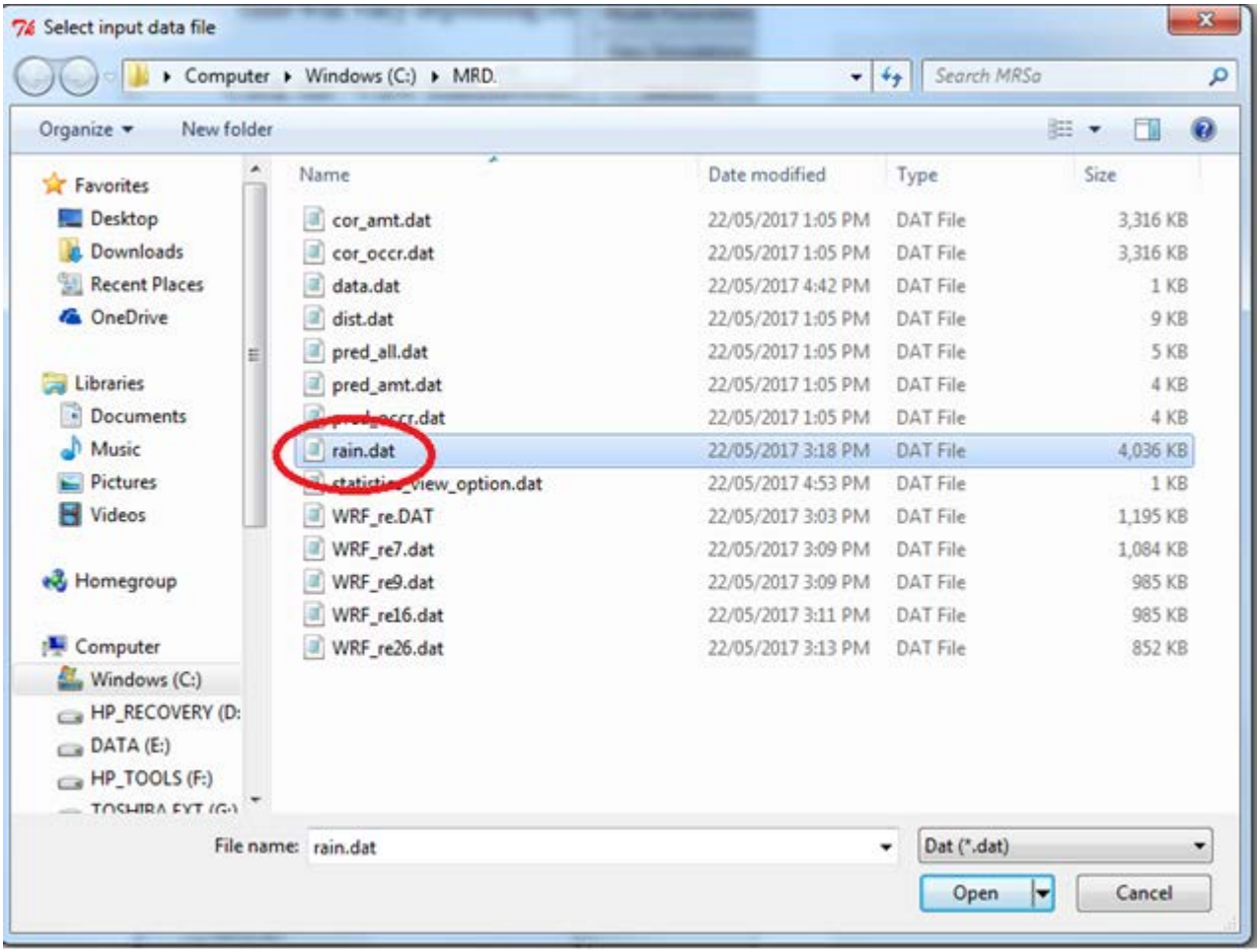


Figure 9. Select input data file.

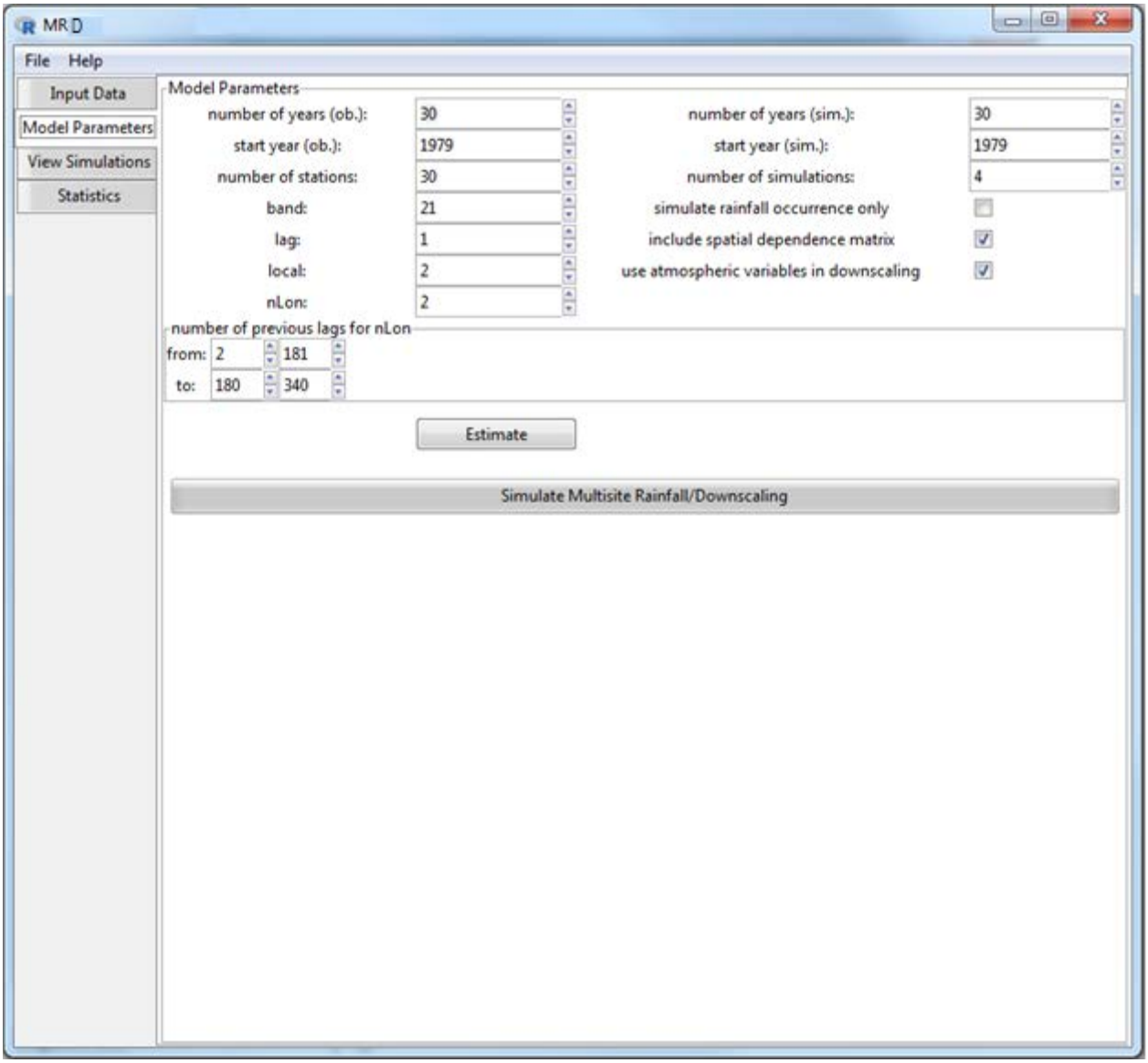


Figure 10. Model Parameters page.

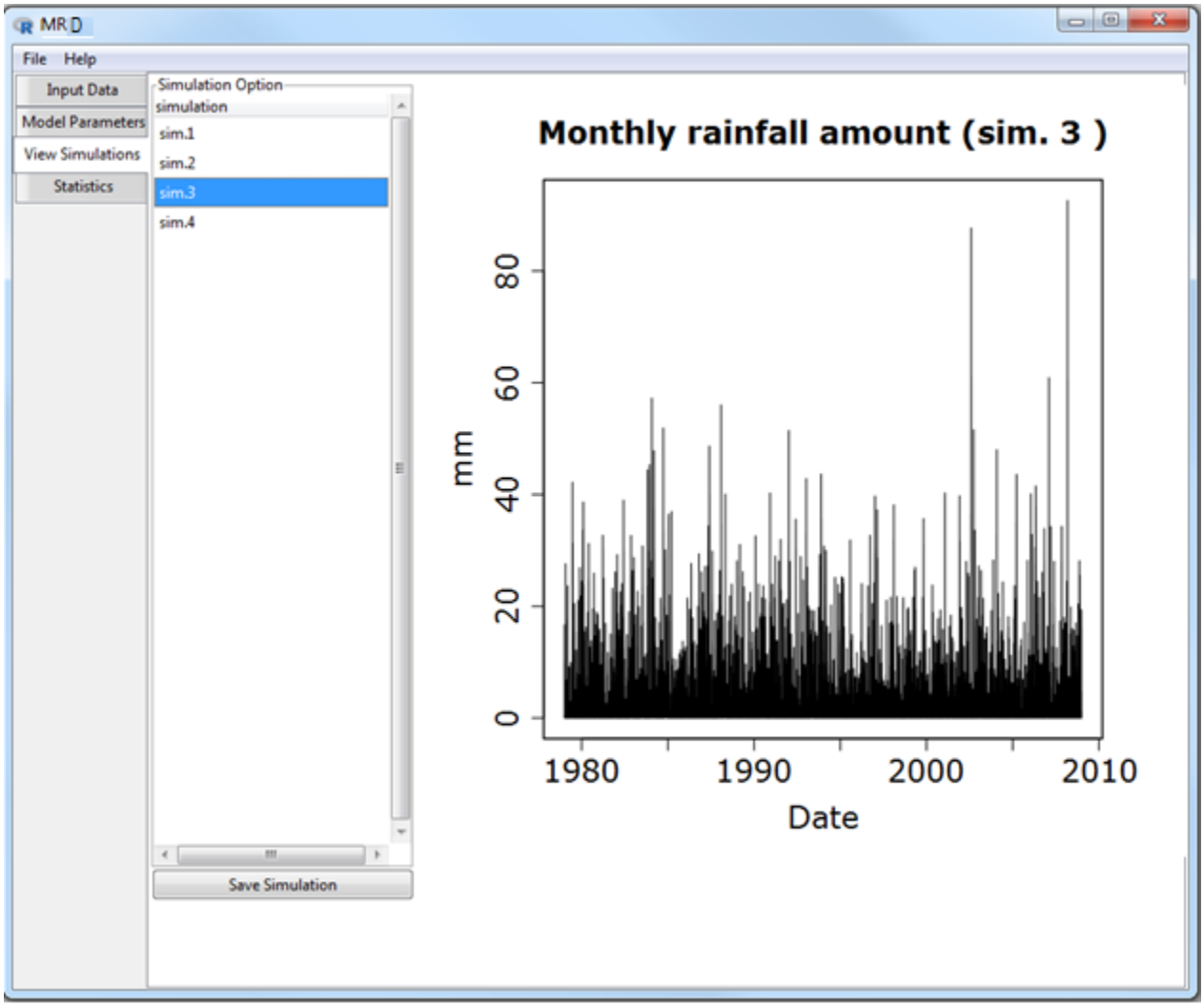


Figure 11. View Simulations page.

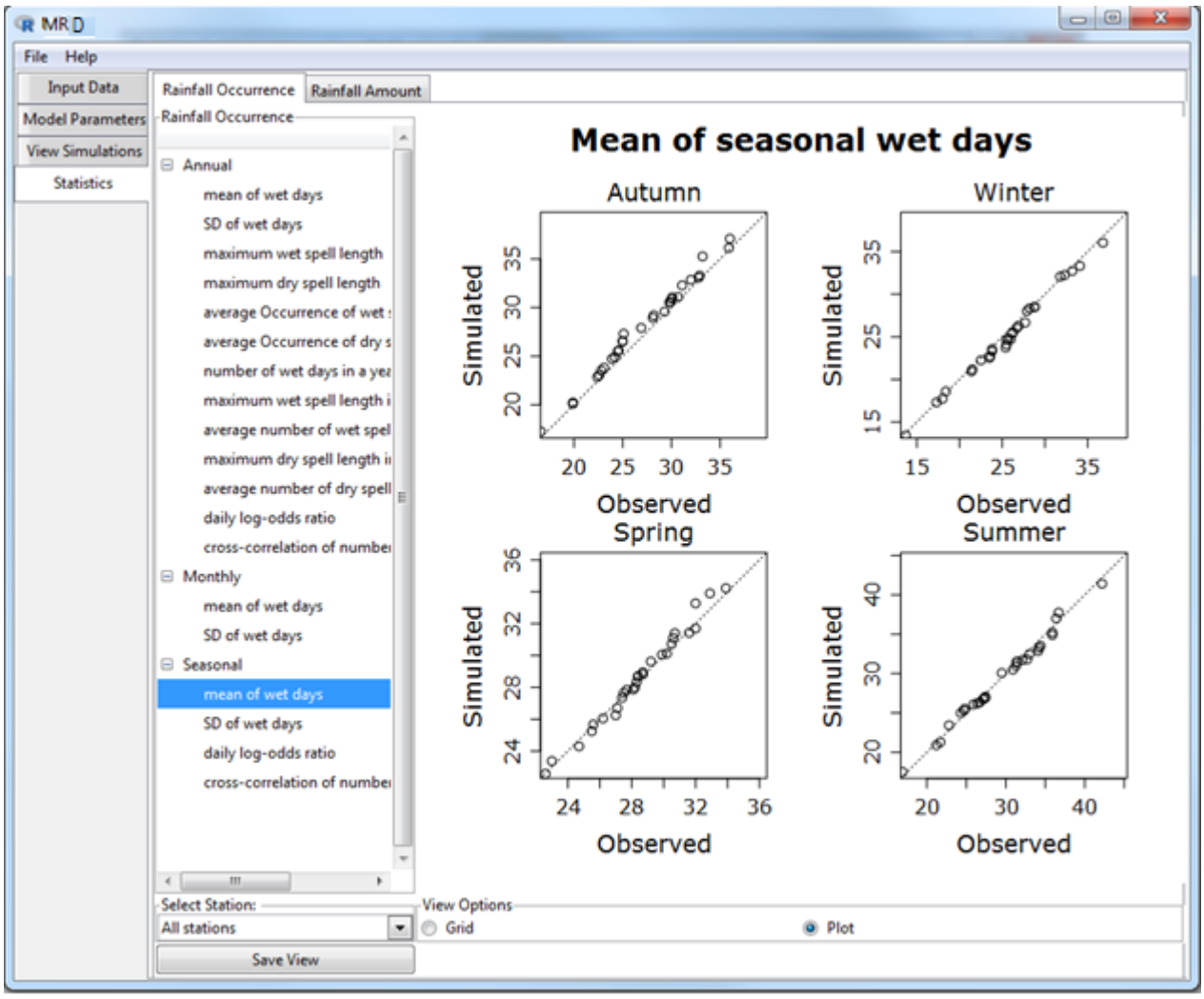
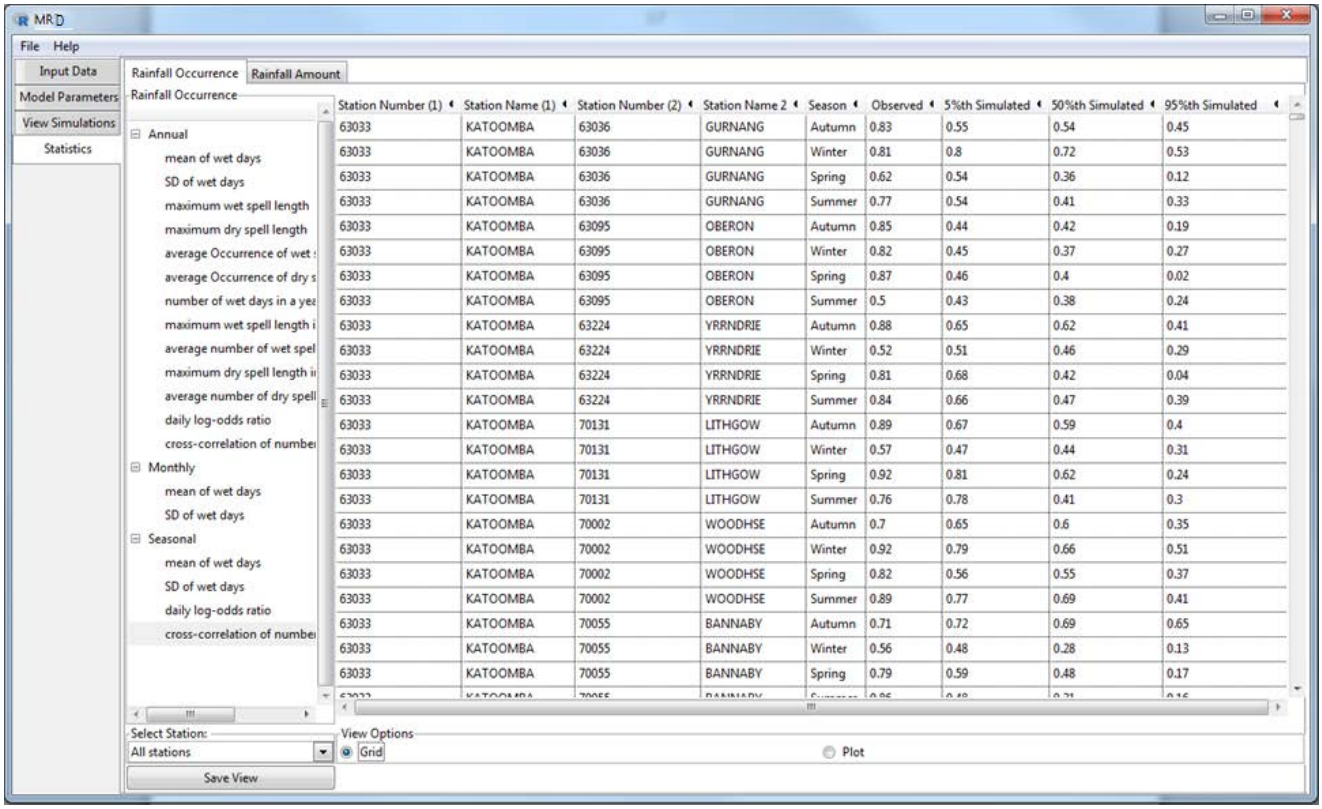


Figure 12. Graphical display of mean of seasonal wet days.



Station Number (1)	Station Name (1)	Station Number (2)	Station Name 2	Season	Observed	5%th Simulated	50%th Simulated	95%th Simulated
63033	KATOOMBA	63036	GURNANG	Autumn	0.83	0.55	0.54	0.45
63033	KATOOMBA	63036	GURNANG	Winter	0.81	0.8	0.72	0.53
63033	KATOOMBA	63036	GURNANG	Spring	0.62	0.54	0.36	0.12
63033	KATOOMBA	63036	GURNANG	Summer	0.77	0.54	0.41	0.33
63033	KATOOMBA	63095	OBERON	Autumn	0.85	0.44	0.42	0.19
63033	KATOOMBA	63095	OBERON	Winter	0.82	0.45	0.37	0.27
63033	KATOOMBA	63095	OBERON	Spring	0.87	0.46	0.4	0.02
63033	KATOOMBA	63095	OBERON	Summer	0.5	0.43	0.38	0.24
63033	KATOOMBA	63224	YRRNDRIE	Autumn	0.88	0.65	0.62	0.41
63033	KATOOMBA	63224	YRRNDRIE	Winter	0.52	0.51	0.46	0.29
63033	KATOOMBA	63224	YRRNDRIE	Spring	0.81	0.68	0.42	0.04
63033	KATOOMBA	63224	YRRNDRIE	Summer	0.84	0.66	0.47	0.39
63033	KATOOMBA	70131	LITHGOW	Autumn	0.89	0.67	0.59	0.4
63033	KATOOMBA	70131	LITHGOW	Winter	0.57	0.47	0.44	0.31
63033	KATOOMBA	70131	LITHGOW	Spring	0.92	0.81	0.62	0.24
63033	KATOOMBA	70131	LITHGOW	Summer	0.76	0.78	0.41	0.3
63033	KATOOMBA	70002	WOODHSE	Autumn	0.7	0.65	0.6	0.35
63033	KATOOMBA	70002	WOODHSE	Winter	0.92	0.79	0.66	0.51
63033	KATOOMBA	70002	WOODHSE	Spring	0.82	0.56	0.55	0.37
63033	KATOOMBA	70002	WOODHSE	Summer	0.89	0.77	0.69	0.41
63033	KATOOMBA	70055	BANNABY	Autumn	0.71	0.72	0.69	0.65
63033	KATOOMBA	70055	BANNABY	Winter	0.56	0.48	0.28	0.13
63033	KATOOMBA	70055	BANNABY	Spring	0.79	0.59	0.48	0.17

Figure 13. Tabulated display of the daily cross-correlation of rainfall in wet days on annual basis.

References:

Mehrotra R., Sharma A. (2010) Development and Application of a Multisite Rainfall Stochastic Downscaling Framework for Climate Change Impact Assessment. Water Resources Research. VOL. 46, W07526, 17 PP., doi:10.1029/2009WR008423.

Mehrotra, R., Li, J., Westra, S., Sharma, A. (2015). "A programming tool to generate multi-site daily rainfall using a two-stage semi parametric model " Environmental Modelling & Software 63: 230–239.