

# Introducing R-Instat

For the training course on  
**The Use of Gridded Satellite Data for Climate Services in Africa**

Roger Stern, Danny Parsons, Sebastian Steinmüller  
Statistics for Sustainable Development (Stats4SD)

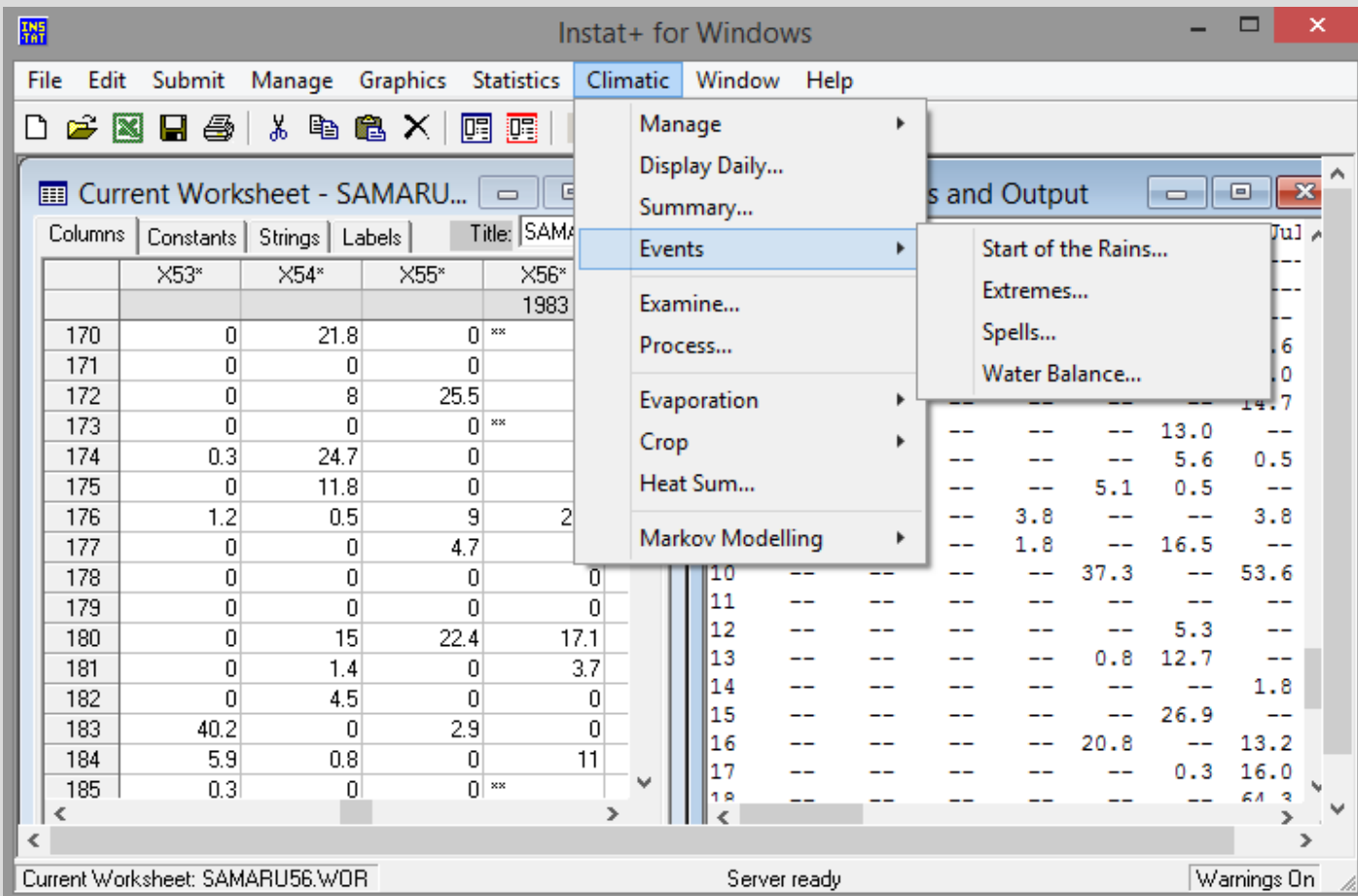
# Background to R-Instat



- Crowd Sourcing campaign 2015
- Identified a gap in statistical software
- Work in Africa and UK
- The R-Instat software will be offered for the first time
- R-Instat
  - Based on the R statistical system
  - With the same ideas as the original Instat

# The Original Instat

- Simple Statistics package
- Easy to use
- With a special climatic menu.
- Still used (2017) by many National Met Services
- Free to download
- Includes “tailored products”
  - Start of the rains
  - Length of the rainy season
  - Risk of long dry spells
- Soon to be replaced by R-Instat



# The R Statistical System

The screenshot displays the RStudio interface. The editor window shows R code for adding columns to a data object and generating a boxplot. The Environment pane lists variables like 'upper\_quar...', 'var\_label', and 'versions'. The Plots pane shows a boxplot of 'Yield' by 'Village'.

```
1165 self$add_columns_to_data(col_name = col_name, col_data = new_col1)
1166
1167
1168 if(keep_attr) {
1169   if(to_type %in% c("numeric", "integer") && signif_figures_label
1170     tmp_attr[[signif_figures_label]] <- NULL
1171 }
1172 self$append_column_attributes(col_name = col_name, new_attr = tr
1173 }
1174 }
1175 self$data_changed <- TRUE
1176 self$variables_metadata_changed <- TRUE
1177 }
1178 )
1179
1180 data_object$set("public", "copy_columns", function(col_names = "") {
1181   <- function(col_names, to_type, factor_values, set_digits, set_decimals, keep_attr, ignore_labels, keep.labels) {
1171:1 <-function>(col_names, to_type, factor_values, set_digits, set_decimals, keep_attr, ignore_labels, keep.labels) {
```

Environment:

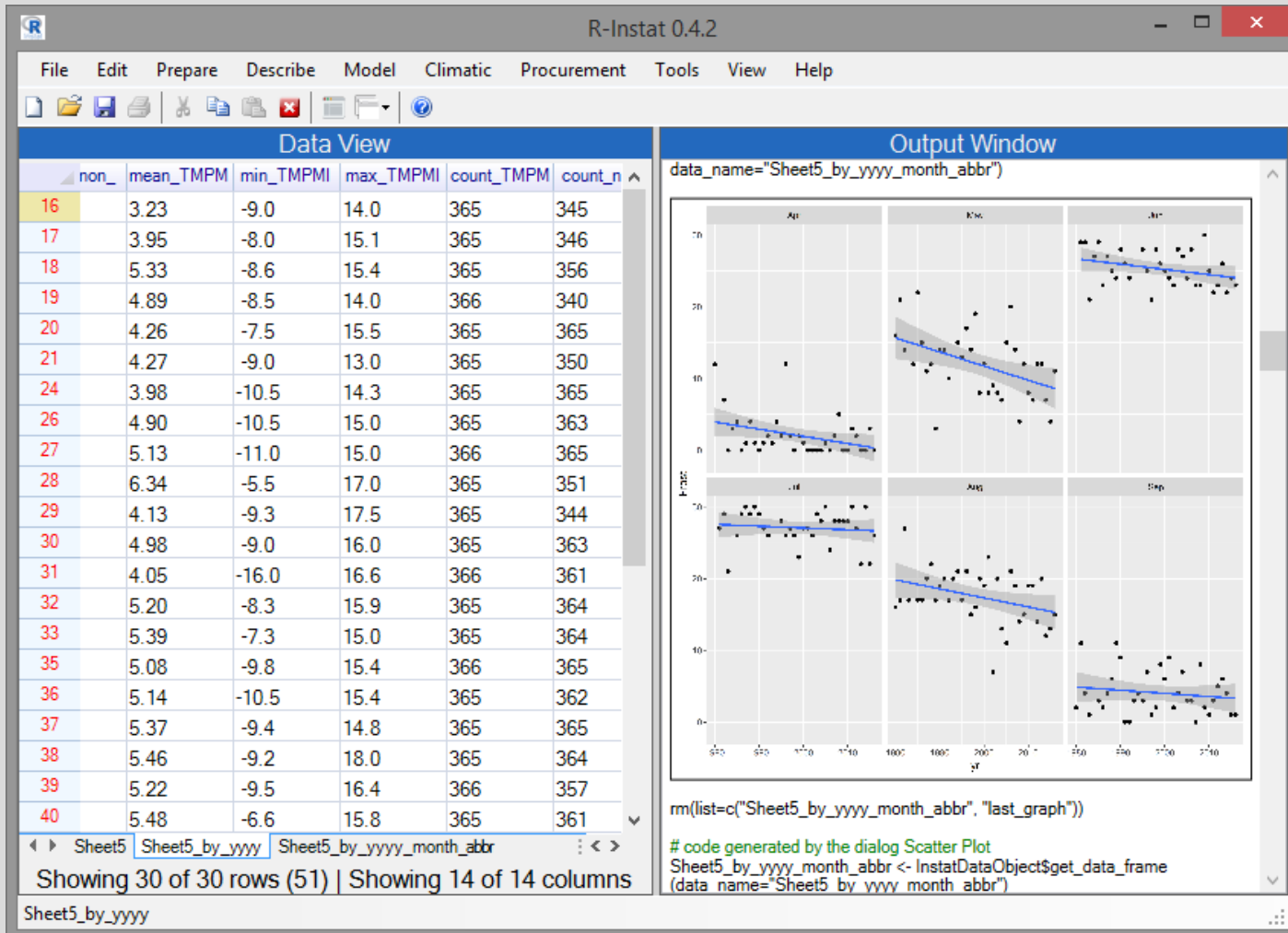
Variable	Value
upper_quar...	"upper_quartile"
var_label	"summary_var"
versions	chr [1:231] "1.4-5" "1.12" "1..."
wet_build_l...	"wet_bulb"
wind_dirac...	"wind_direction"
wind_speed...	"wind_speed"
year_label	"year"
year_month...	"year_month"

Boxplot Data (Estimated):

Village	Yield (Approx. Median)
KESEN	28
NANDA	45
NIKO	30
SABEY	48

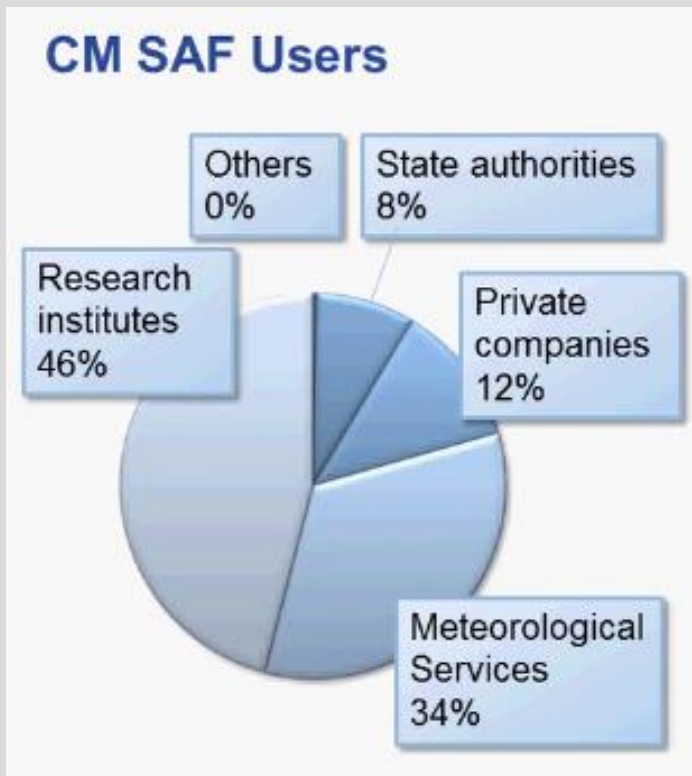
- Very powerful statistical system
- Free and open source
- RStudio makes it easier to use
- Use through commands or scripts
- Over 5,000 extra packages add to the facilities
- Including CM SAF package used in this workshop
  - Together with special scripts for specific analyses

# R-Instat

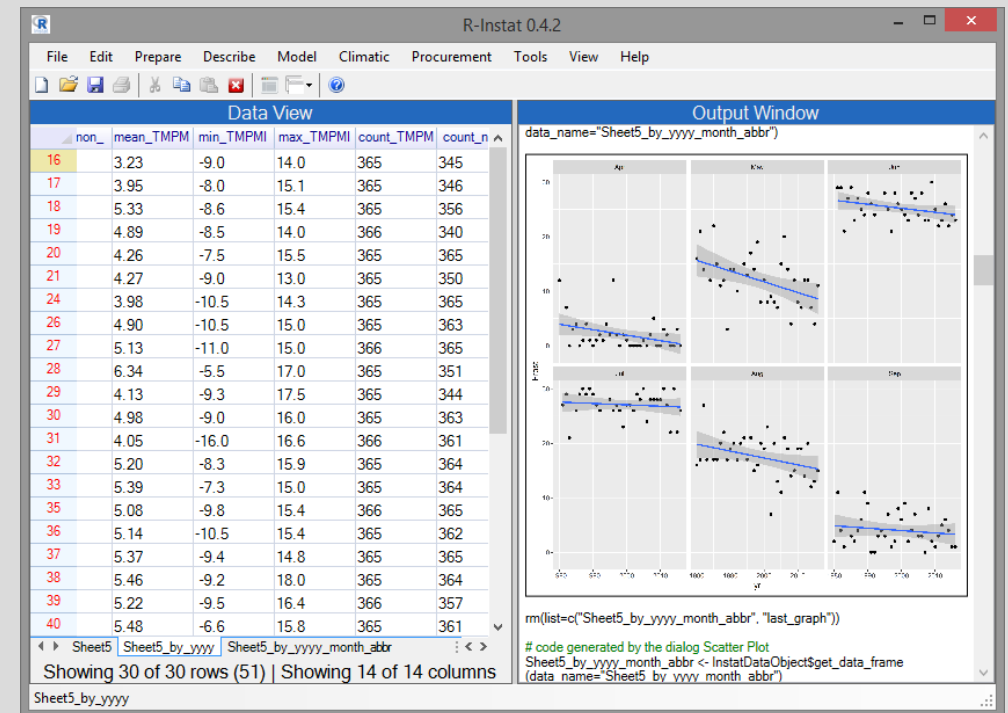


- Menu-driven front end to R
- Designed to make R really easy to use
- Particularly for those who already use a spreadsheet.
- Like Instat it is a general statistics package
- With a special climatic menu.
- Free to download

# In this workshop



- CM SAF will be used to manage and analyse the satellite data



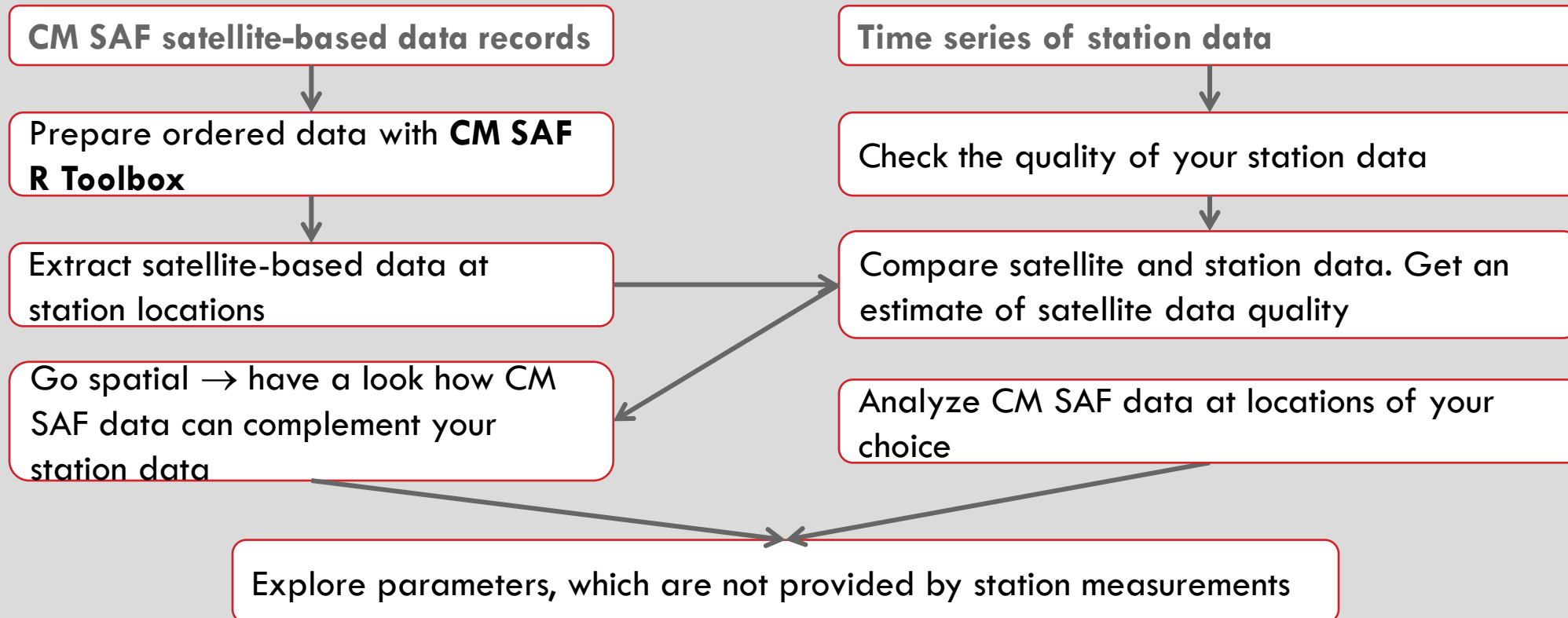
- R-Instat will support the comparison of the satellite data with the data from the ground stations
- And could do further analyses of the satellite data

## CM SAF R TOOLBOX

- Prepare CM SAF NetCDF data, from tar-file to ready-to-use NetCDF
- Analysis and visualization of CM SAF spatial data
- Quick and easy-to-use functions

## R-Instat

- Click and point (menu based) general statistics software
- Special climatic menu to analysis historical records
- Many dialogs for further analysis e.g. comparison between satellite and station data



# R-Instat easy to use?

maximum - Micros...

	A	B	C	D	F	G	H	I
1	station	long	lat	elevat	yyyy	mm	dd	tmax
2	OXBOW	28.37	-28.43	2600	1961	1	1	15.5
3	OXBOW	28.37	-28.43	2600	1961	1	2	15.6
4	OXBOW	28.37	-28.43	2600	1961	1	3	15
5	OXBOW	28.37	-28.43	2600	1961	1	4	17
6	OXBOW	28.37	-28.43	2600	1961	1	5	16.8
7	OXBOW	28.37	-28.43	2600	1961	1	6	17.7
8	OXBOW	28.37	-28.43	2600	1961	1	7	19.2
9	OXBOW	28.37	-28.43	2600	1961	1	8	19.6
10	OXBOW	28.37	-28.43	2600	1961	1	9	21.6
11	OXBOW	28.37	-28.43	2600	1961	1	10	22.4
12	OXBOW	28.37	-28.43	2600	1961	1	11	21.3
13	OXBOW	28.37	-28.43	2600	1961	1	12	21.1

\*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

	DateD	Year	Mon	Day	Rain	var
1	01-Jan-1952	1952	1	1	1.3	
2	02-Jan-1952	1952	1	2	.0	
3	03-Jan-1952	1952	1	3	.0	
4	04-Jan-1952	1952	1	4	.0	
5	05-Jan-1952	1952	1	5	.0	
6	06-Jan-1952	1952	1	6	.0	
7	07-Jan-1952	1952	1	7	.0	
8	08-Jan-1952	1952	1	8	.0	
9	09-Jan-1952	1952	1	9	.0	
10	10-Jan-1952	1952	1	10	.0	
11	11-Jan-1952	1952	1	11	14.2	
12	12-Jan-1952	1952	1	12	.0	
13	13-Jan-1952	1952	1	13	.0	
14	14-Jan-1952	1952	1	14	.0	

Instat+ for Windows

File Edit Submit Manage Graphics Statistics Climatic Window Help

Current Worksheet - SAMARU...

Columns Constants Strings Labels Title: SAMARU...

Visible: 5 of 5 Variables

Server ready Warnings On

- Manage
  - Display Daily...
  - Summary...
- Events
  - Start of the Rains...
  - Extremes...
  - Spells...
  - Water Balance...
- Examine...
- Process...
- Evaporation
  - ...
- Crop
  - ...
- Heat Sum...
- Markov Modelling
  - ...



# Similar to a spreadsheet for data analysis?

✓ Name at the top of column

✗ Data start at row 1

✗ Results in a different window

✓ Data in “columns” – one type of data only per column

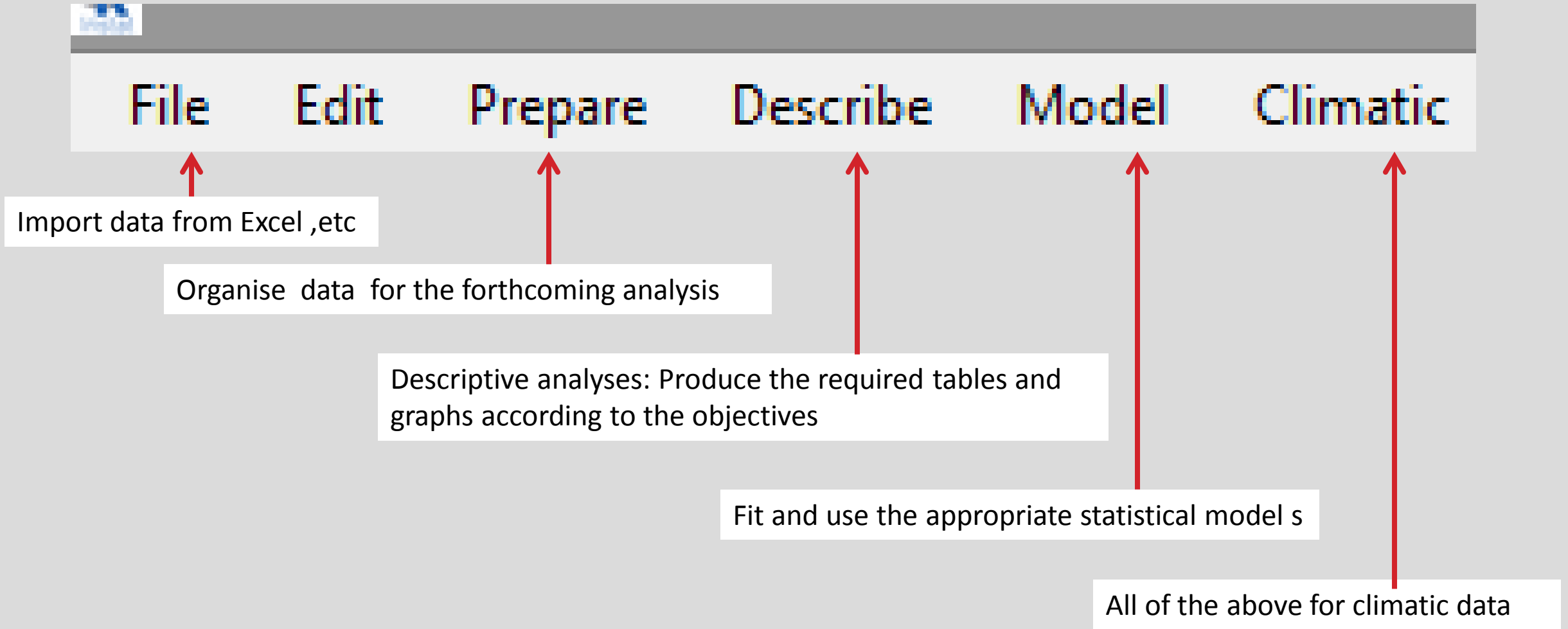
✓ Multiple sheets (data frames)

✗ Just a window onto the data

The screenshot shows the R-Instat 0.4.2 interface. The 'Data View' window displays a table with columns: State (f), Date (D), year, mont, day, doy, RR, and Tn. The 'Output Window' displays a grid of 12 scatter plots for months Jan through Dec, showing rainfall (mm) on the y-axis and year (yr) on the x-axis. Red arrows point from the text boxes to specific elements in the software interface.

	State (f)	Date (D)	year	mont	day	doy	RR	Tn
1	Saltpond	1944-01-01	1944	Jan	1	1	0.0	NA
2	Saltpond	1944-01-02	1944	Jan	2	2	0.0	NA
3	Saltpond	1944-01-03	1944	Jan	3	3	0.0	NA
4	Saltpond	1944-01-04	1944	Jan	4	4	0.0	NA
5	Saltpond	1944-01-05	1944	Jan	5	5	0.0	NA
6	Saltpond	1944-01-06	1944	Jan	6	6	0.0	NA
7	Saltpond	1944-01-07	1944	Jan	7	7	0.0	NA
8	Saltpond	1944-01-08	1944	Jan	8	8	0.0	NA
9	Saltpond	1944-01-09	1944	Jan	9	9	0.0	NA
10	Saltpond	1944-01-10	1944	Jan	10	10	0.0	NA
11	Saltpond	1944-01-11	1944	Jan	11	11	0.0	NA
12	Saltpond	1944-01-12	1944	Jan	12	12	0.0	NA
13	Saltpond	1944-01-13	1944	Jan	13	13	39.6	NA
14	Saltpond	1944-01-14	1944	Jan	14	14	0.0	NA
15	Saltpond	1944-01-15	1944	Jan	15	15	0.0	NA
16	Saltpond	1944-01-16	1944	Jan	16	16	0.0	NA
17	Saltpond	1944-01-17	1944	Jan	17	17	0.0	NA

# The menus in R-Instat



# Example: Prepare – Annual totals from daily data

The screenshot shows the RStudio interface. The 'Prepare' menu is open, and 'Column: Reshape' is selected. The 'Data View' window shows a table with columns 'month\_abbrev' and 'day\_in\_month'. The table contains data for the month of January, with 'day\_in\_month' ranging from 1 to 5.

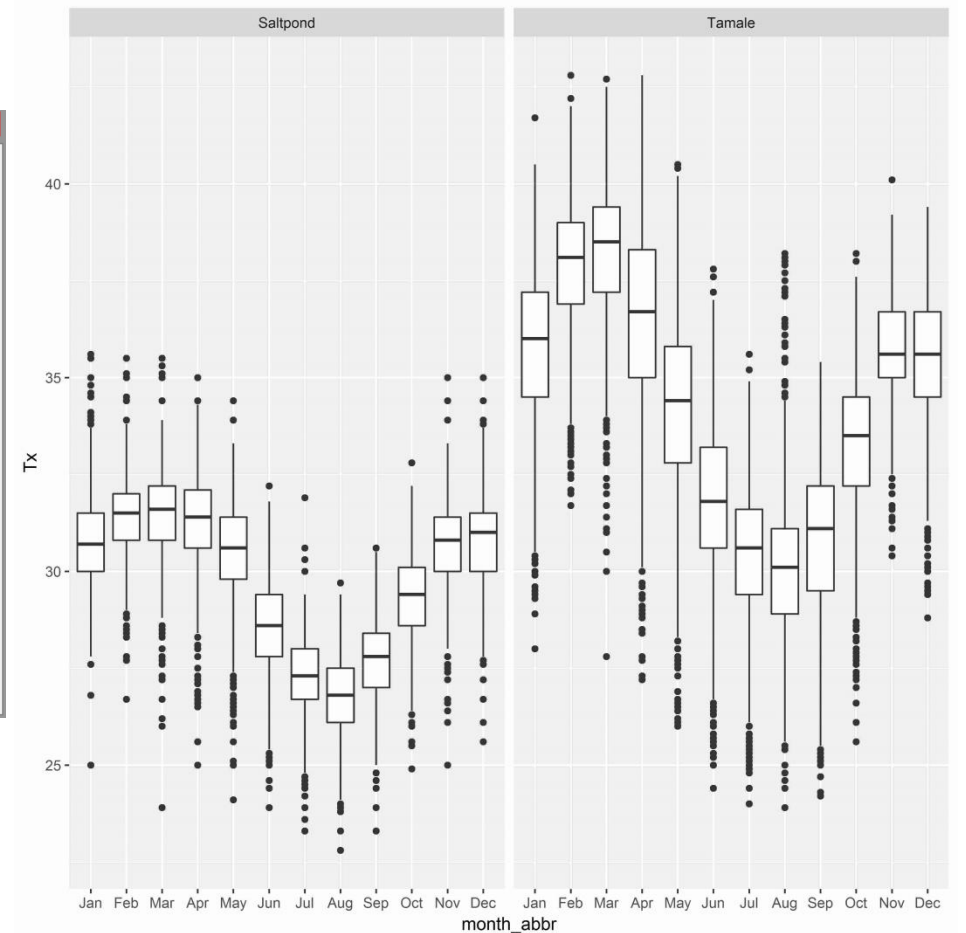
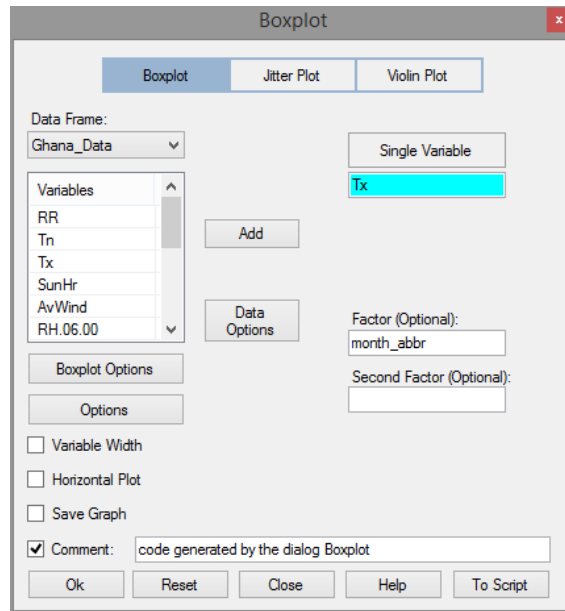
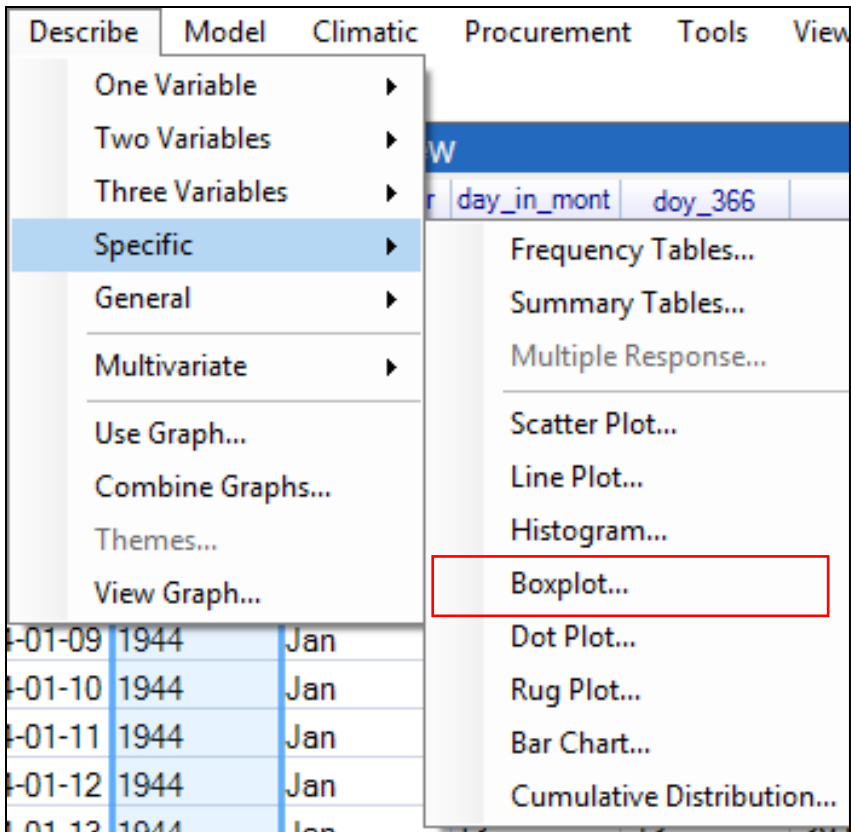
month_abbrev	day_in_month	
Jan	1	1
Jan	2	2
Jan	3	3
Jan	4	4
Jan	5	5

The 'Column Statistics' dialog box is shown. The 'Data Frame' is 'Ghana\_Data'. The 'Variable(s) to Summarize' are 'Ghana\_Data' and 'RR'. The 'By Factor(s)' are 'Ghana\_Data', 'Station', and 'year'. The 'Options' section is checked for 'Store Results in Data', 'Print Results to Output', 'Drop Unused Levels', and 'Omit Missing Values'. The 'Comment' field contains 'code generated by the dialog Column Statistics'.

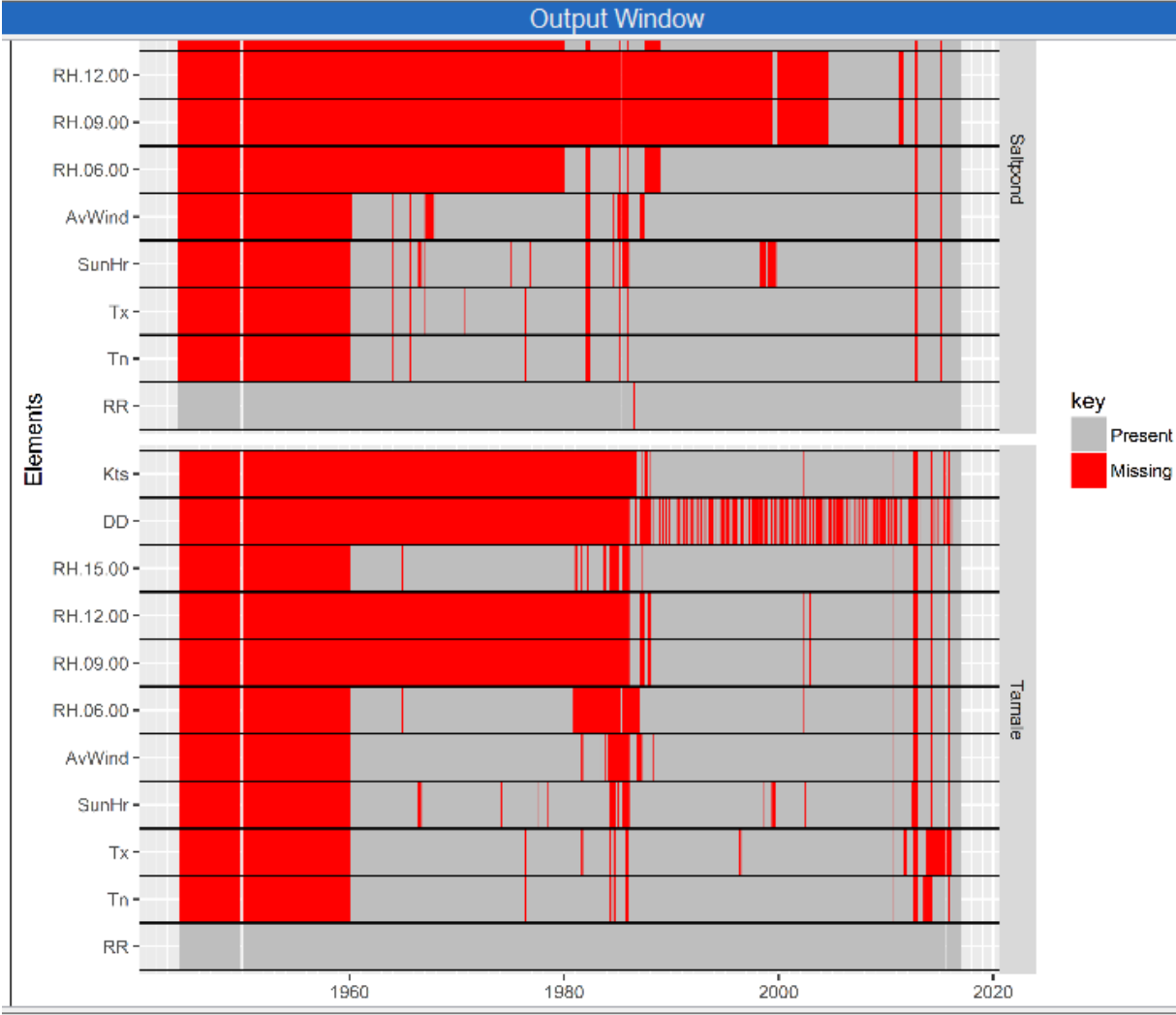
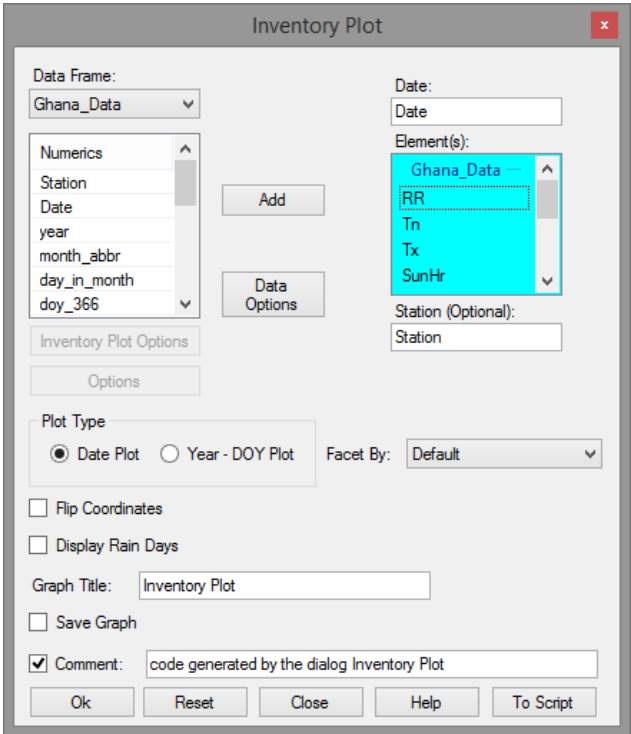
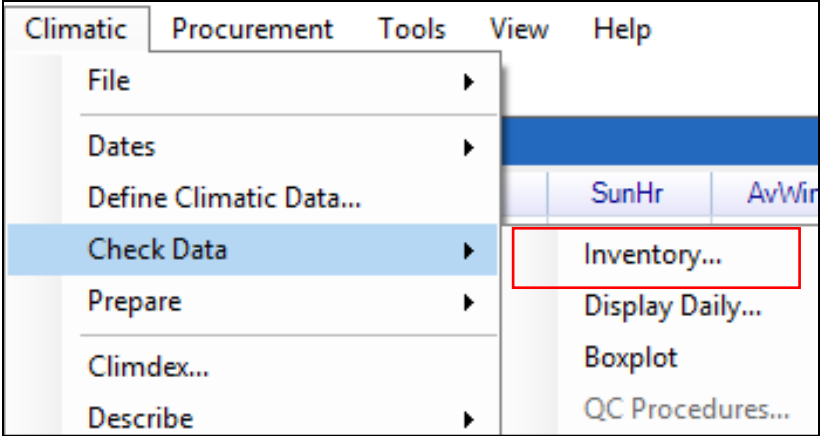
The 'Data View' window shows a table with columns 'Station (f)', 'year (f)', 'count\_n', 'count\_RR', and 'sum\_RF'. The table contains data for the year 1944 to 1960, with 'Station' always being 'Saltpond'.

	Station (f)	year (f)	count_n	count_RR	sum_RF
1	Saltpond	1944	366	366	724
2	Saltpond	1945	365	365	644
3	Saltpond	1946	365	365	1001
4	Saltpond	1947	365	365	1017
5	Saltpond	1948	366	366	1089
6	Saltpond	1949	273	273	772
7	Saltpond	1950	365	365	776
8	Saltpond	1951	365	365	1428
9	Saltpond	1952	366	366	1376
10	Saltpond	1953	365	365	699
11	Saltpond	1954	365	365	986
12	Saltpond	1955	365	365	1485
13	Saltpond	1956	365	366	1151
14	Saltpond	1957	365	365	702
15	Saltpond	1958	365	365	749
16	Saltpond	1959	365	365	789
17	Saltpond	1960	366	366	1004

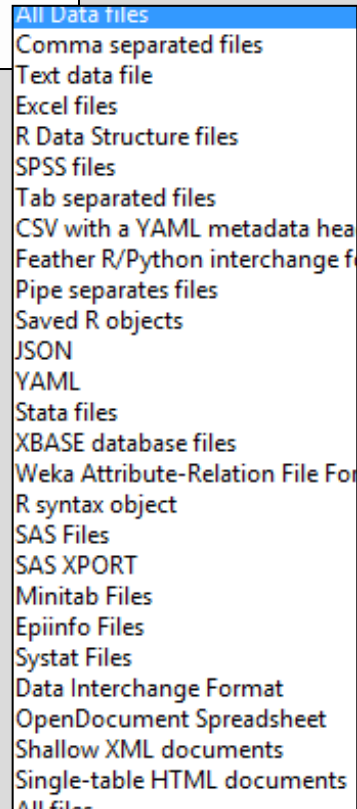
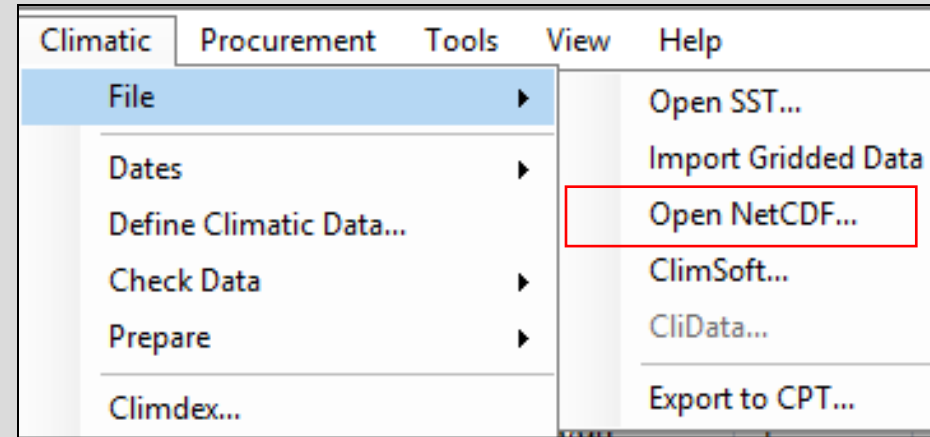
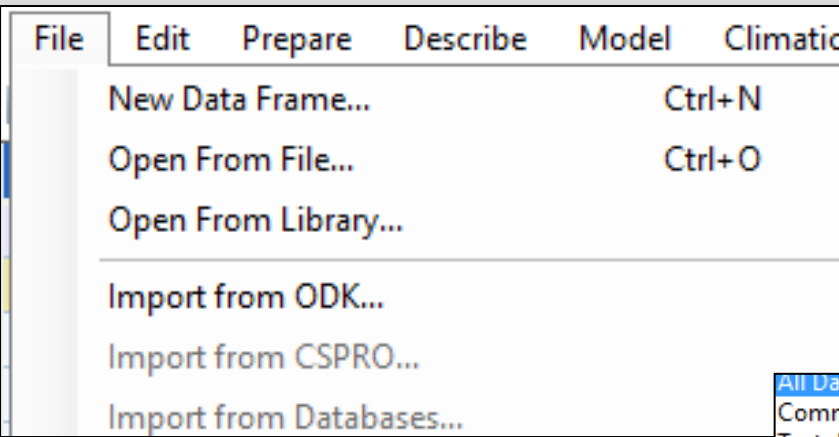
# Example - Describe: Boxplots to check data



# Example - Climatic: Inventory

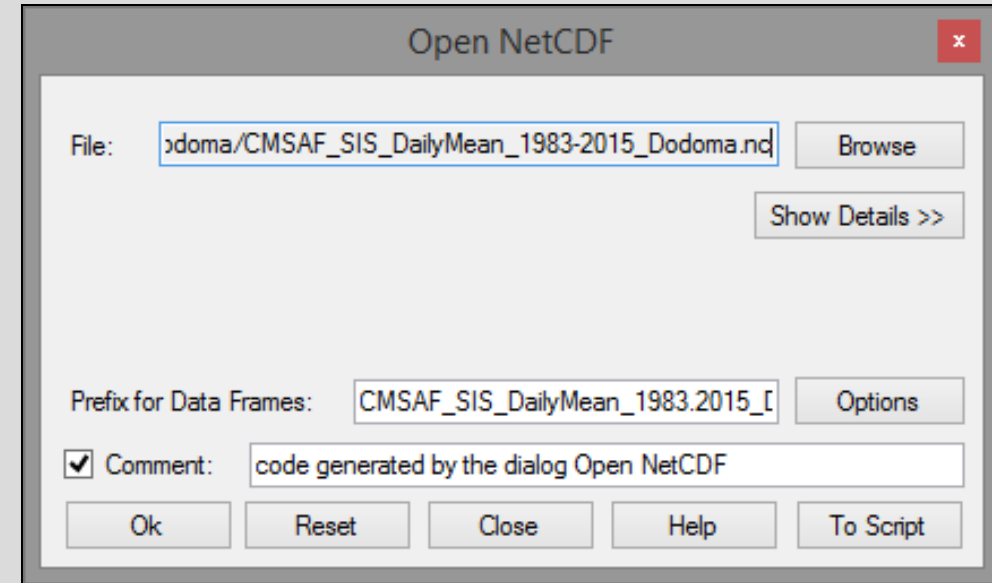
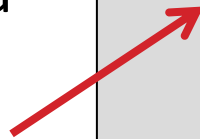


# In this workshop



**File > Open From File**  
for the station data

In many different  
formats



**Climatic > File > Open NetCDF**  
For satellite data in nc files.