# Quality control: practice

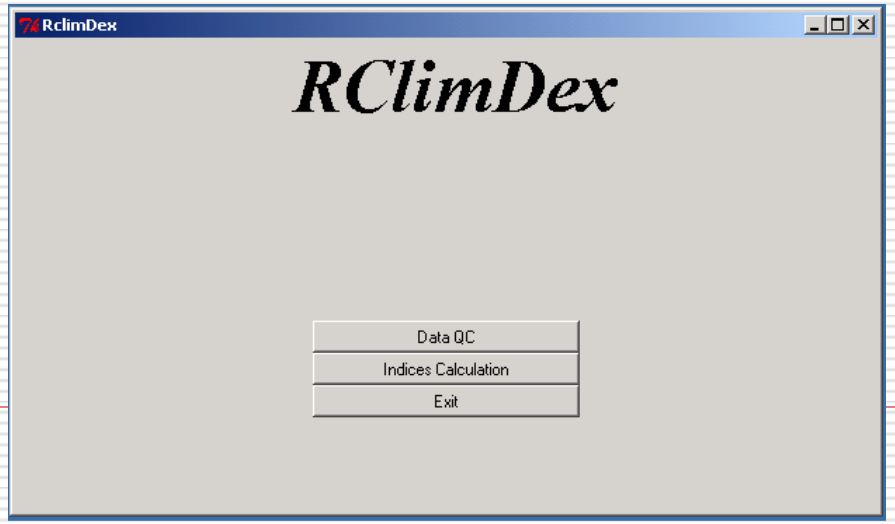
By Lucie Vincent

Climate Research Division, Environment Canada, Toronto, Canada

Capacity building workshop on Data Rescue and Climate Change Indices
The University of the West Indies, Mona, Jamaica
7-10 May 2012

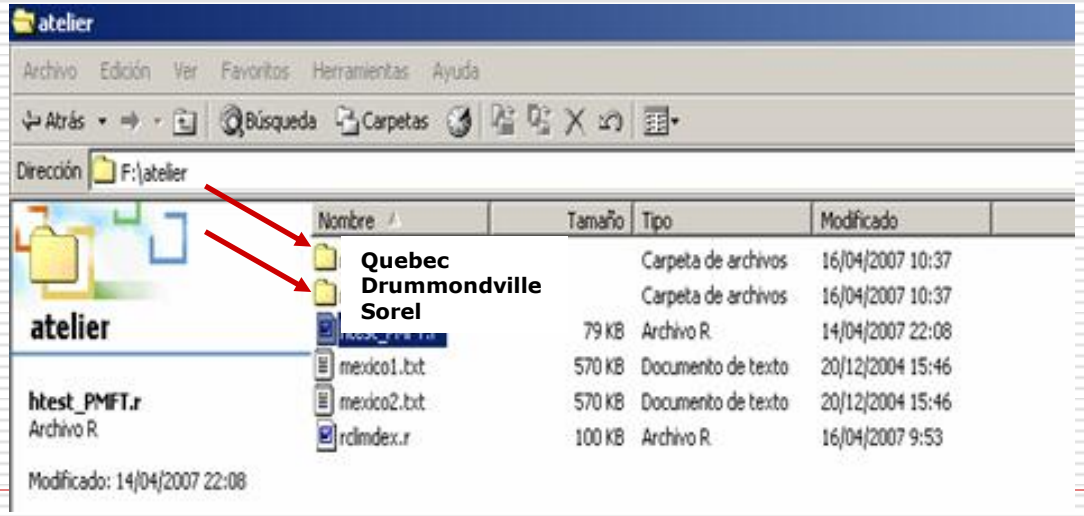
# RClimDex for Quality Control

- RClimDex works under R:R is free and more robust than Excel for statistical analyses
- RClimdex is also used for the computation of climate change indices
- Developed by Xuebin Zhang et Feng Yang of Environment Canada to be used in the ETCCDI workshops



# First

- Create a directory for each station
- Example: create 3 directories for 3 stations: Québec, Drummondville and Sorel
- Install R



# Input data file

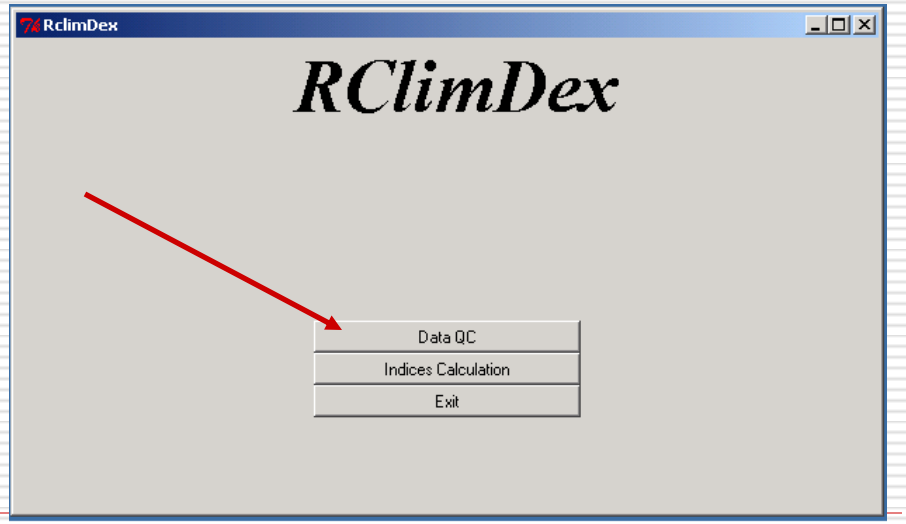
- Data format (file.txt):
  year, month, day, rain (mm), tmax (°C), tmin (°C)
- Missing values : -99.9
- One file by station
- Correct # of days in each month (e.g. no Apr 31 w -99.9)

| Year | Month | Day | Rain | Tmax | Tmin |
|---|---|---|---|---|---|
| 1949 | 03 | 19 | 0.0 | 25.5 | -99.9 |
| 1949 | 03 | 20 | 0.0 | 30.0 | 12.0 |
| 1949 | 03 | 21 | 0.0 | 30.0 | 5.0 |
| 1949 | 03 | 22 | 0.0 | 32.0 | 7.5 |
| 1949 | 03 | 23 | 0.0 | 32.5 | 10.5 |
| 1949 | 03 | 24 | 0.0 | 32.5 | 10.5 |
| 1949 | 03 | 25 | 0.0 | 31.0 | 10.0 |
| 1949 | 03 | 26 | 0.0 | 30.5 | 9.0 |
| 1949 | 03 | 27 | 0.0 | 31.5 | 7.5 |
| 1949 | 03 | 28 | 0.0 | 29.0 | 9.0 |
| 1949 | 03 | 29 | 0.0 | 28.5 | 7.0 |
| 1949 | 03 | 30 | 0.0 | 29.0 | 7.0 |
| 1949 | 03 | 31 | 0.0 | 29.0 | 8.0 |
| 1949 | 04 | 01 | 0.0 | 28.5 | 10.0 |
| 1949 | 04 | 02 | 0.0 | 27.0 | 5.0 |
| 1949 | 04 | 03 | 0.0 | 26.0 | 7.0 |
| 1949 | 04 | 04 | 0.0 | 26.0 | 6.0 |
| 1949 | 04 | 05 | 0.0 | 26.0 | 7.0 |
| 1949 | 04 | 06 | 0.0 | 28.0 | 5.5 |
| 1949 | 04 | 07 | 0.0 | 31.0 | 12.0 |
| 1949 | 04 | 08 | 0.0 | 32.0 | 10.0 |
| 1949 | 04 | 09 | 0.0 | 31.5 | 7.0 |
| 1949 | 04 | 10 | 0.0 | 31.5 | 8.0 |
| 1949 | 04 | 11 | 0.0 | 32.0 | 9.0 |
| 1949 | 04 | 12 | 0.0 | 31.5 | 12.0 |
| 1949 | 04 | 13 | 0.0 | 31.0 | 8.0 |
| 1949 | 04 | 14 | 0.0 | 32.0 | 7.0 |
| 1949 | 04 | 15 | 0.0 | 34.0 | 9.0 |
| 1949 | 04 | 16 | 0.0 | 30.0 | 10.0 |
| 1949 | 04 | 17 | 0.0 | 32.5 | 8.0 |
| 1949 | 04 | 18 | 0.0 | 32.0 | 9.0 |
| 1949 | 04 | 19 | 0.0 | 31.0 | 8.0 |
| 1949 | 04 | 20 | 0.0 | 34.5 | 8.0 |
| 1949 | 04 | 21 | 0.0 | 30.0 | 12.0 |
| 1949 | 04 | 22 | 0.0 | 31.0 | 13.0 |

# Starting RClimDex

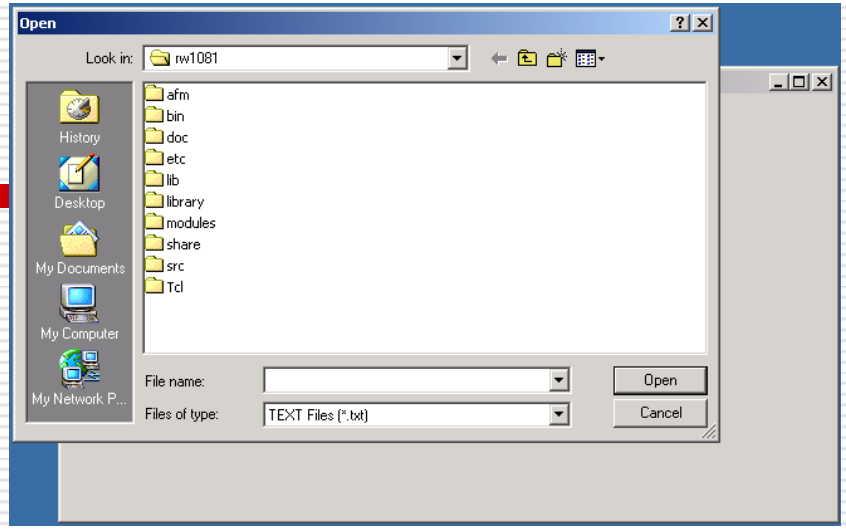
- Launch R 

- Load RClimdex:
  File, Source Code, RClimdex_May2012.r

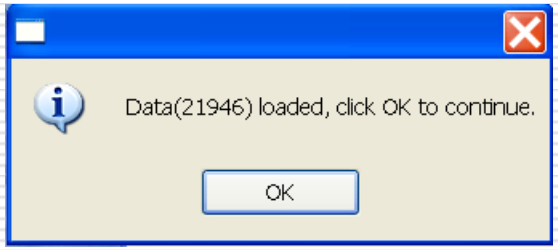


# Procedure

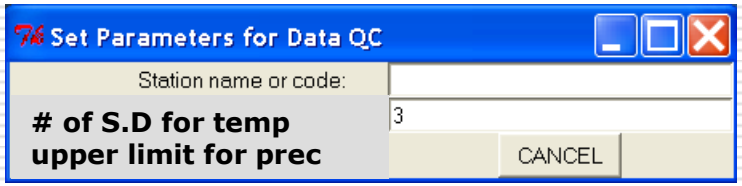
- Load input file
  (data for a station)




- Make sure that data are loaded




- Enter:
  - station name
  - #s.d. for temp: 5 s.d.?
  - upper limit for prec: 300 mm?



# Check 3 files in directory log

- Quebec_ prcpQC.csv:
  - days w prec < 0 mm
  - days w prec > 300 mm (user defined)

| year | month | day | prcp | tmax | tmin |
|---|---|---|---|---|---|
| **1960** | 6 | 26 | -0.3 | 28.2 | 20.9 |
| **1962** | 5 | 17 | 266.8 | 27.8 | 21.7 |
| **1963** | 4 | 22 | 219.2 | 31 | 22 |
| **1963** | 5 | 3 | 207.6 | 27.3 | 22.9 |
| **1964** | 4 | 22 | 210.6 | 25.5 | 21.8 |

# Check 3 files in directory log (continued)

- Quebec_tempQC.csv:
  - days w tmax < tmin

| year | month | day | prcp | tmax | tmin | tmax-tmin |
|---|---|---|---|---|---|---|
| **1963** | 4 | 30 | 0.3 | 20.6 | 21.3 | -0.7 |
| **1973** | 6 | 2 | 88.5 | 30.4 | 31.6 | -1.2 |
| **1976** | 12 | 18 | 18 | 26 | 27 | -1 |
| **1984** | 2 | 28 | 0 | 20.4 | 23.4 | -3 |
| **1984** | 12 | 12 | 48.3 | 30 | 31.9 | -1.9 |
| **1988** | 9 | 29 | 0 | 3 | 20 | -17 |
| **1992** | 12 | 4 | 10 | 23 | 31.5 | -8.5 |
| **1993** | 4 | 27 | 0.4 | 21.5 | 22 | -0.5 |
| **1997** | 5 | 26 | 18 | 31 | 31.5 | -0.5 |
| **1998** | 1 | 19 | 0.3 | 20 | 23.5 | -3.5 |

- Quebec_ tepstd.csv:
  - days w temp outside of 5 s.d. (user defined)

| year | month | day | tmaxlow | tmax | tmaxup | tminlow | tmin | tminup |
|---|---|---|---|---|---|---|---|---|
| **1984** | 2 | 28 | 21.23 | 20.4 | 41.29 | 17.1 | 23.4 | 30.23 |
| **1988** | 9 | 29 | 6.72 | 3 | 50.41 | 12.41 | 20 | 27.64 |
| **1992** | 12 | 4 | 23.06 | 23 | 39.06 | 13.21 | 31.5 | 32.4 |
| **1994** | 7 | 27 | 21.79 | 31 | 34.05 | 8.17 | 31 | 30.76 |

# Check identified values

- Check if:
  - Data is in agreement with neighbour stations
  - Data is part of heat wave, cold wave, wave w ext prec
- Accept, modify or replace data by -99.9
- Keep track of modifications
- Save file under a new name (example: Quebec2.txt)

# *Demonstration*