

Le contrôle de qualité: pratique

Par Enric Aguilar¹ et Lucie Vincent²

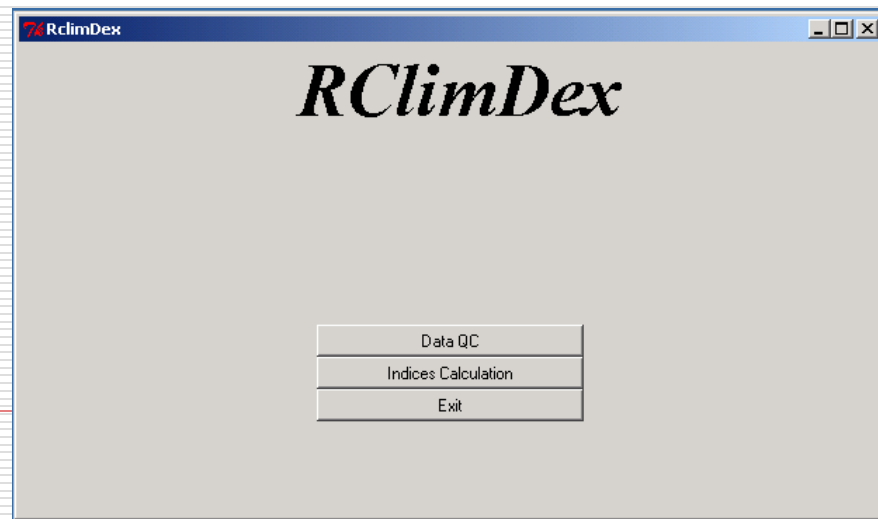
¹C3-URV, Tarragona, Spain

²Climate Research Division, Environment Canada, Toronto, Canada

Séminaire sur l'homogénéisation des données climatiques
et les indices des changements climatiques
CMM/CCI/VAR/JCOMM ETCCDI,
octobre 2009, Vacoas, Île Maurice

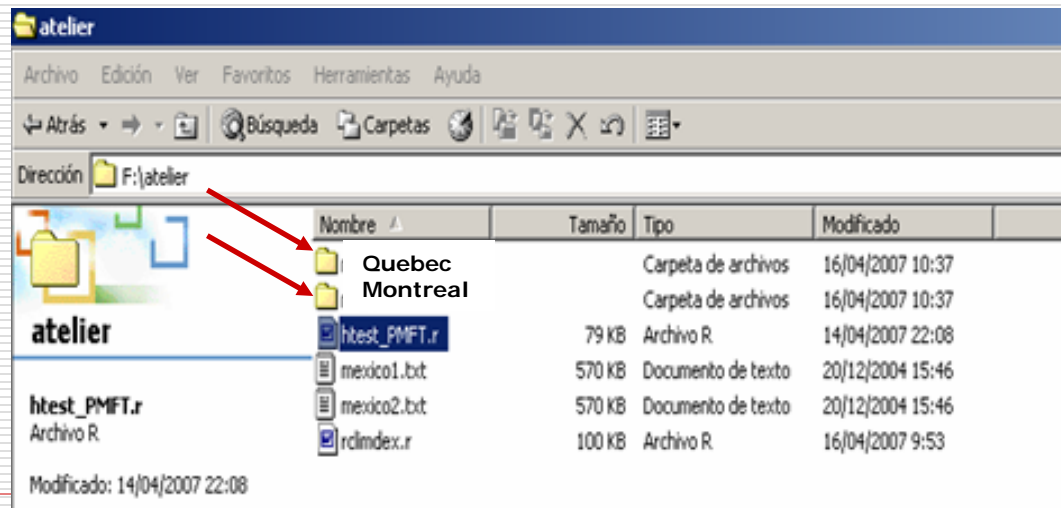
Utiliser RClimDex pour faire le contrôle de la qualité

- RClimDex fonctionne sous l'environnement R parce que R est disponible gratuitement et qu'il est plus robuste qu'Excel pour les analyses statistiques
- RClimdex est utilisé pour le contrôle de la qualité et aussi pour le calcul des indices climatiques
- Il a été développé par Xuebin Zhang et Feng Yang d'Environnement Canada pour utilisation dans les séminaires ETCCDI



Avant de commencer

- Il faut créer un répertoire pour chaque station et placer les fichiers d'entrée dans chaque répertoire
- Dans cet exemple, on a 2 stations: Québec et Montréal
- RClimDex va créer des fichiers de sortie dans chaque répertoire




Fichiers d'entrée

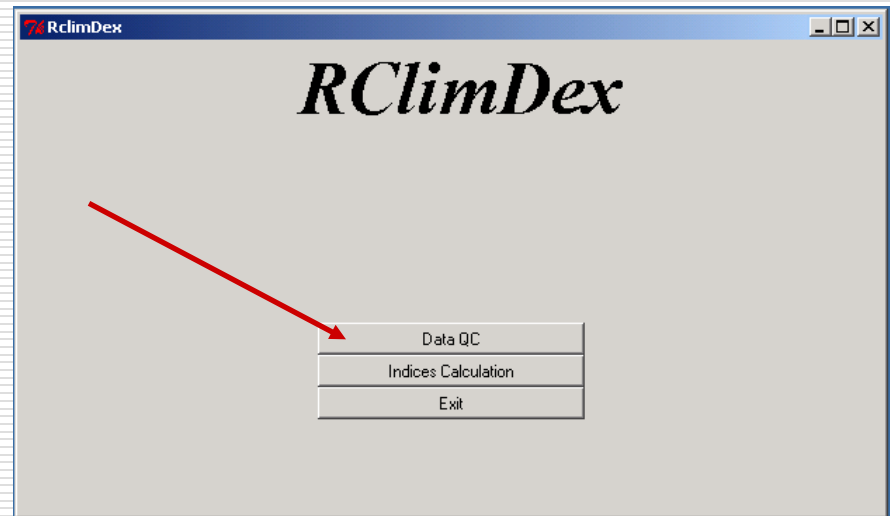
- Il faut s'assurer que les fichiers d'entrée ont le bon format:
année, mois, jour, préc (mm),
tmax (°C), tmin (°C)
- Valeur manquante : -99.9
- Un fichier par station

Année mois jour préc 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

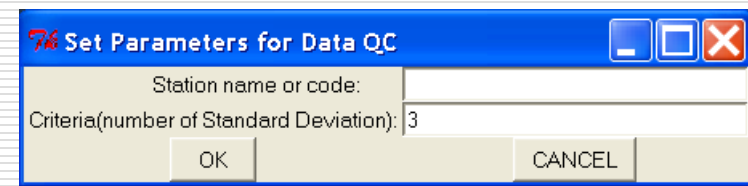
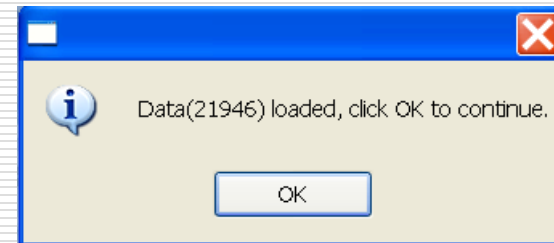
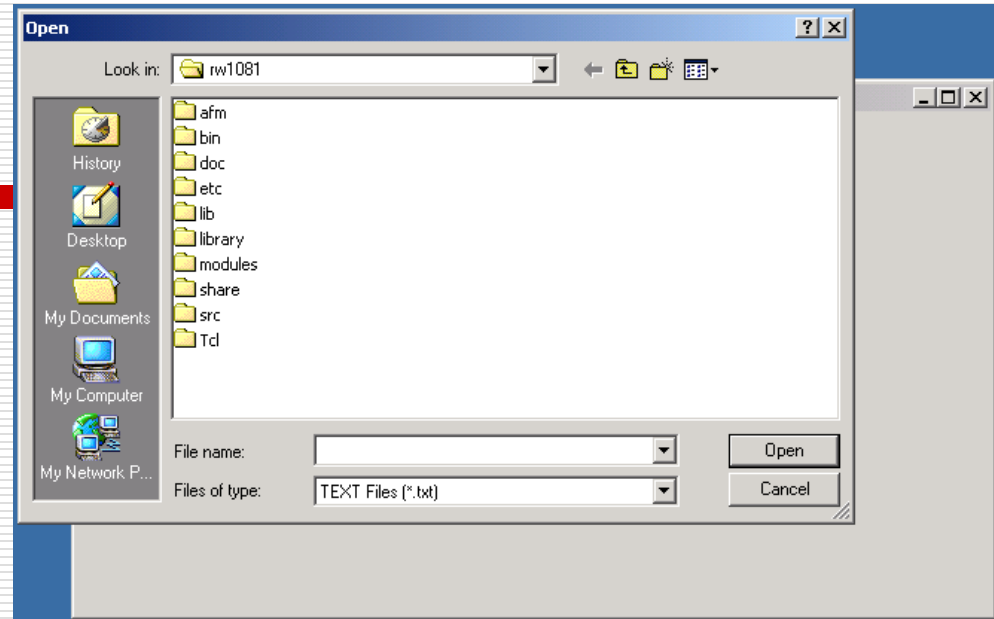
Charger RClimDex

- Démarrer R  Acceso directo a Rgui.Inl
- Charger RClimdex (File, Source code, RClimDex)
- Choisir Data QC



Procédure

- Charger le fichier d'entrée
- S'assurer qu'il est bien chargé
- Fournir le nom de la station et le nombre d'écart-type pour identifier les valeurs extrêmes



4 fichiers de sortie sont créés dans le répertoire log

- **Quebec_nastatistic.csv:**
 - indique le nombre de valeurs manquantes
 - s'il y a trop de valeurs manquantes, le calcul des indices ne se fera pas

TITLE	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANN
PRCP	1900	0	1	0	0	0	0	0	0	0	0	0	0	1
TMAX	1900	0	1	0	0	0	0	0	0	0	0	0	0	1
TMIN	1900	0	1	0	0	0	0	0	0	0	0	0	0	1
PRCP	1901	0	0	0	0	0	0	0	0	0	0	0	0	0
TMAX	1901	0	0	0	0	0	0	0	0	0	0	0	0	0
TMIN	1901	0	0	0	0	0	0	0	0	0	0	0	0	0
PRCP	1902	0	0	0	0	0	0	0	0	0	0	0	0	0

- **Quebec_prcpQC.csv:**
 - indique les jours pour lesquels $\text{prec} < 0$ et $\text{prec} > 200$ mm

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

4 fichiers de sortie sont créés dans le répertoire log

(Suite)

- **Quebec_tempQC.csv:**
 - indique les jours pour lesquels $t_{max} < t_{min}$

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:**
 - indique les jours pour lesquels t_{max} et t_{min} sont des valeurs extrêmes (selon l'écart-type)

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

Corriger les valeurs erronées

- Il faut vérifier chaque valeur identifiée:
 - vérifier si cette valeur est possible avec les stations avoisinantes
 - vérifier s'il y a une vague de chaleur, de froid ou de pluie
- Modifier par une valeur possible ou indiquer une valeur manquante (-99.9)
- Sauvegarder le fichier sous un nouveau nom (exemple: Quebec2.txt)

Démonstration

