

Interpolation de corrections météorologiques combinant les données GRIB, Météocode et Nowcast

Marc-Élie Lapointe & Pascal Vincent



Environnement
Canada

Environment
Canada

L'apprentissage statistique

en quelques mots

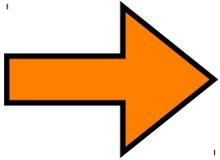
- Apprendre une *fonction* effectuant des **prédictions**
- En se basant sur un **grand ensemble d'exemples** (base de données)
- Capable de **généraliser** à de nouveaux cas (pas vus dans les données)

Données d'EC présentement stockées:

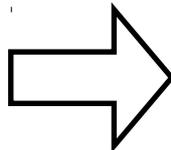
Météocode: prédictions de **très haute qualité** (corrigées à la main par météorologue) **à faible résolution** (zone code).

GRIB: prédictions **brutes du modèle** à **très haute résolution**.

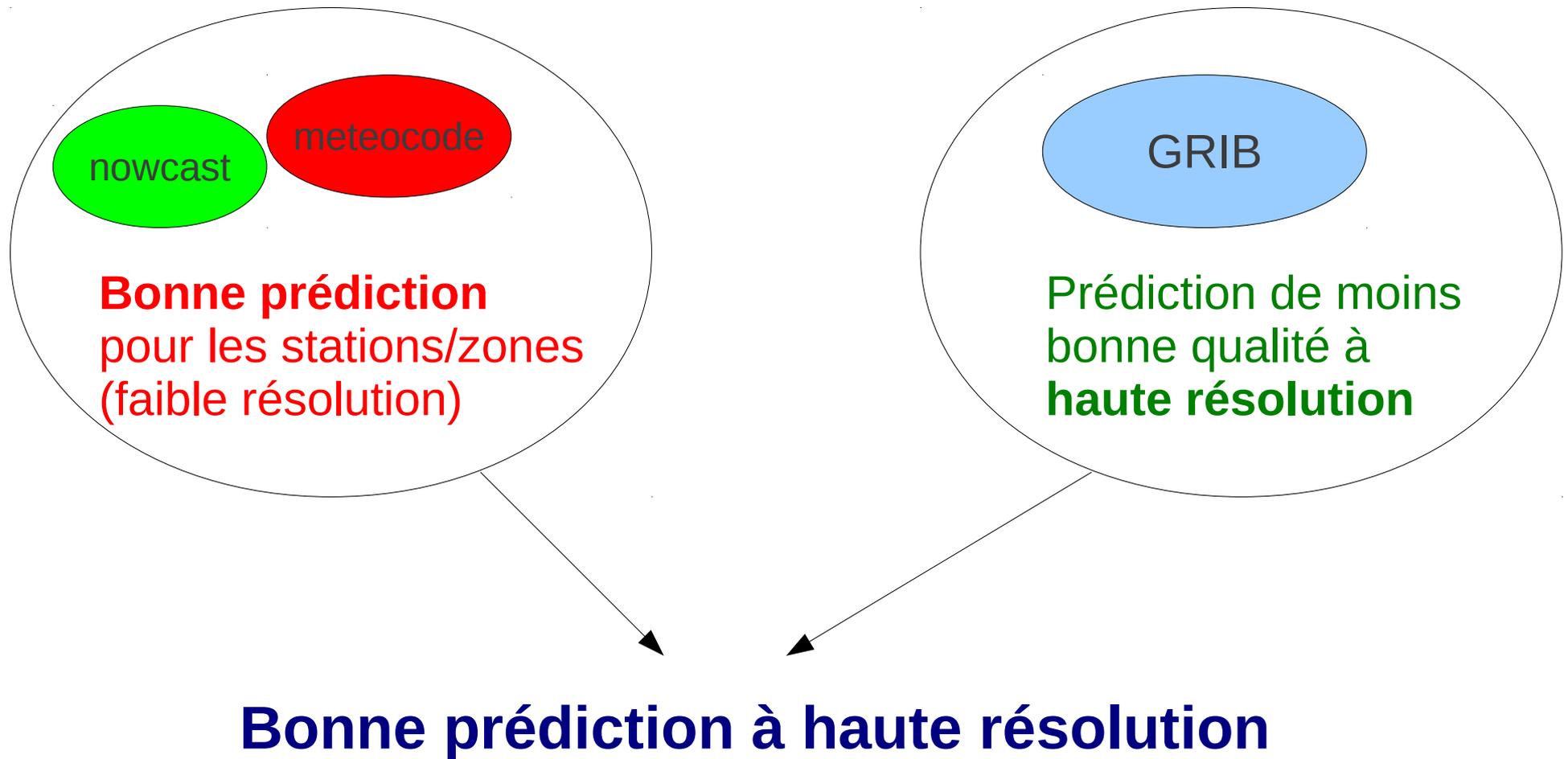
Nowcast: **observations + bonnes prédictions court terme** à **moyenne résolution** (stations météorologiques)



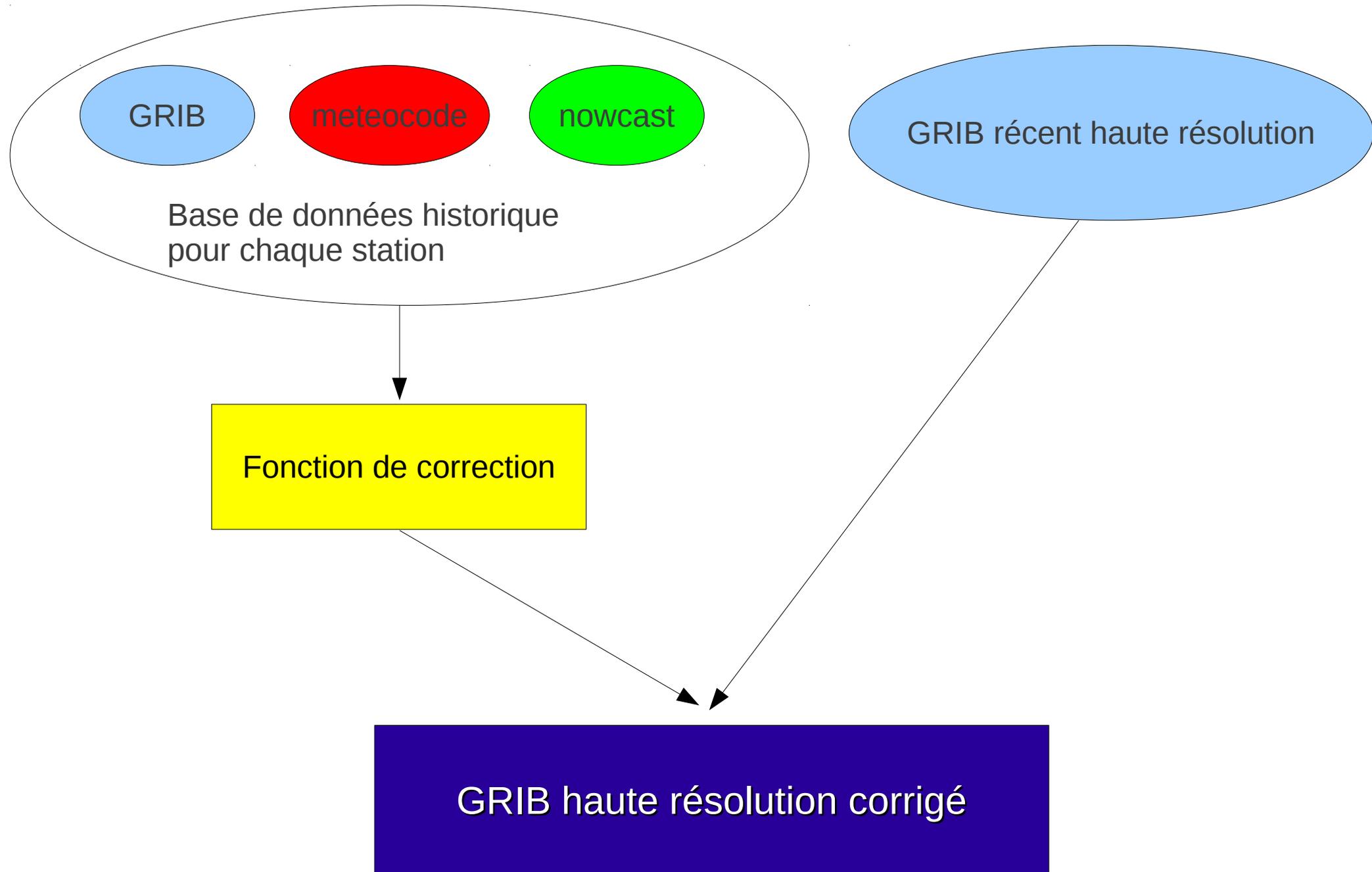
Possibilité de combiner deux des trois: apprendre à "interpoler" les **corrections** des experts, ou les **observations**, sur la grille haute résolution.

Différents algorithmes d'apprentissage envisageable (processus Gaussiens, réseaux de neurones, etc...)  **Recherche!**

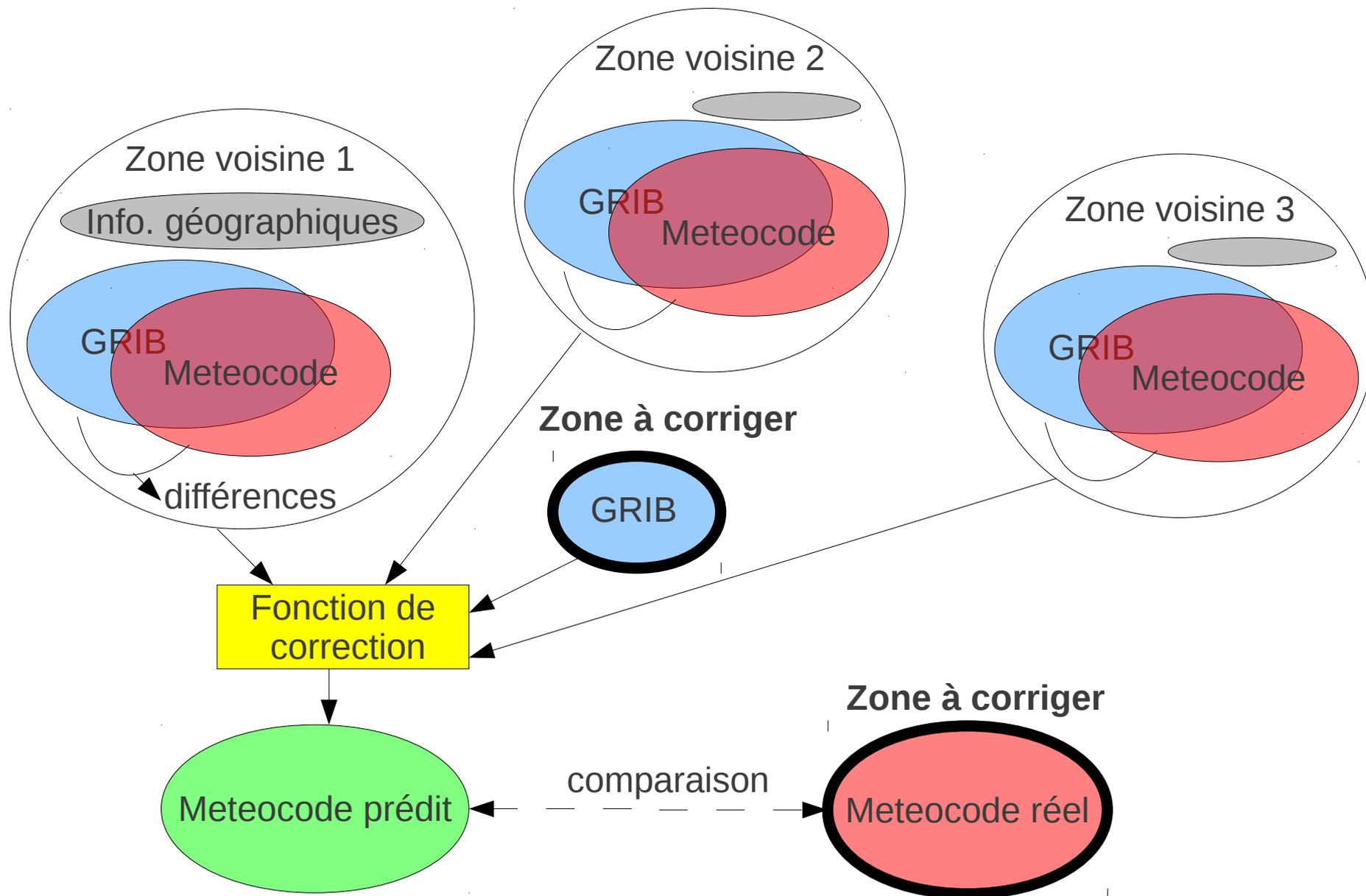
Le problème de correction:



Comment obtenir une bonne prédiction à haute résolution:



Exemple: fonction de correction des trois voisins



On cherche la fonction qui **minimise** la différence entre le meteocode prédit et le vrai meteocode

Les différentes fonctions que met en place l'interface web (expérimentale)

1. **Collecter** régulièrement les données et les **stocker** dans une base de données
2. **Visualiser les données** qui sont dans la base de données et les données qui sont dans les fichiers GRIB récents.
3. **Créer** des tables CSV à partir de la base de données et prétraiter ces tables afin d'obtenir un **dataset python** préformaté pour un algorithme d'apprentissage.
4. **Importer des modèles appris** de correction et **tester** ces différents modèles

Étape 1: collecte d'une historique de données

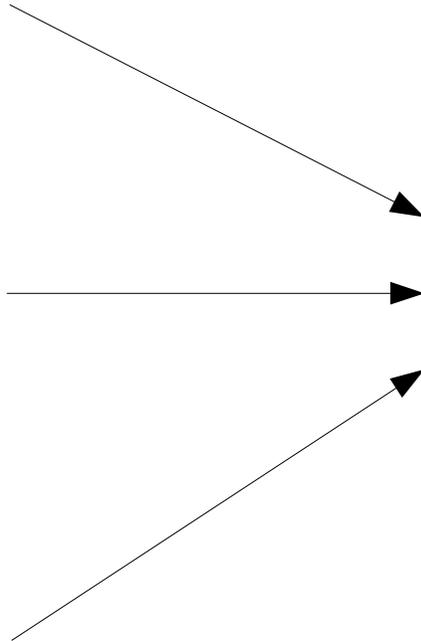
Météocode
XML
journalier

Rapport
GRIB
journalier

SCRIBE
Nowcast
horaire

Base de donnée
SQL
historique

Pour chaque station
météorologique d'EC



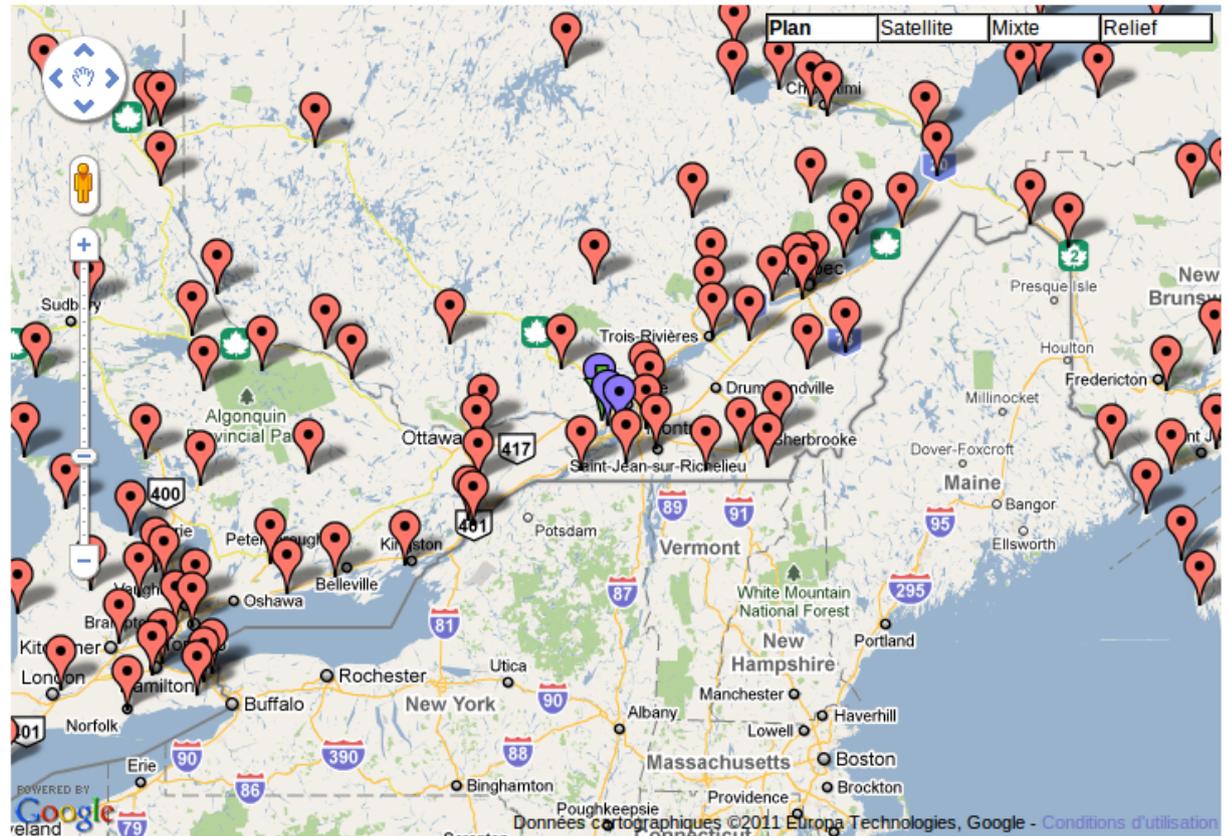
Visualiser les données grib, meteo code et nowcast pour différentes zones et/ou stations ainsi que des données grib partout

Accueil Visualiser Tester Tables Dataset Créer table Prétraiter Comparer Importer Créer dataset Admin

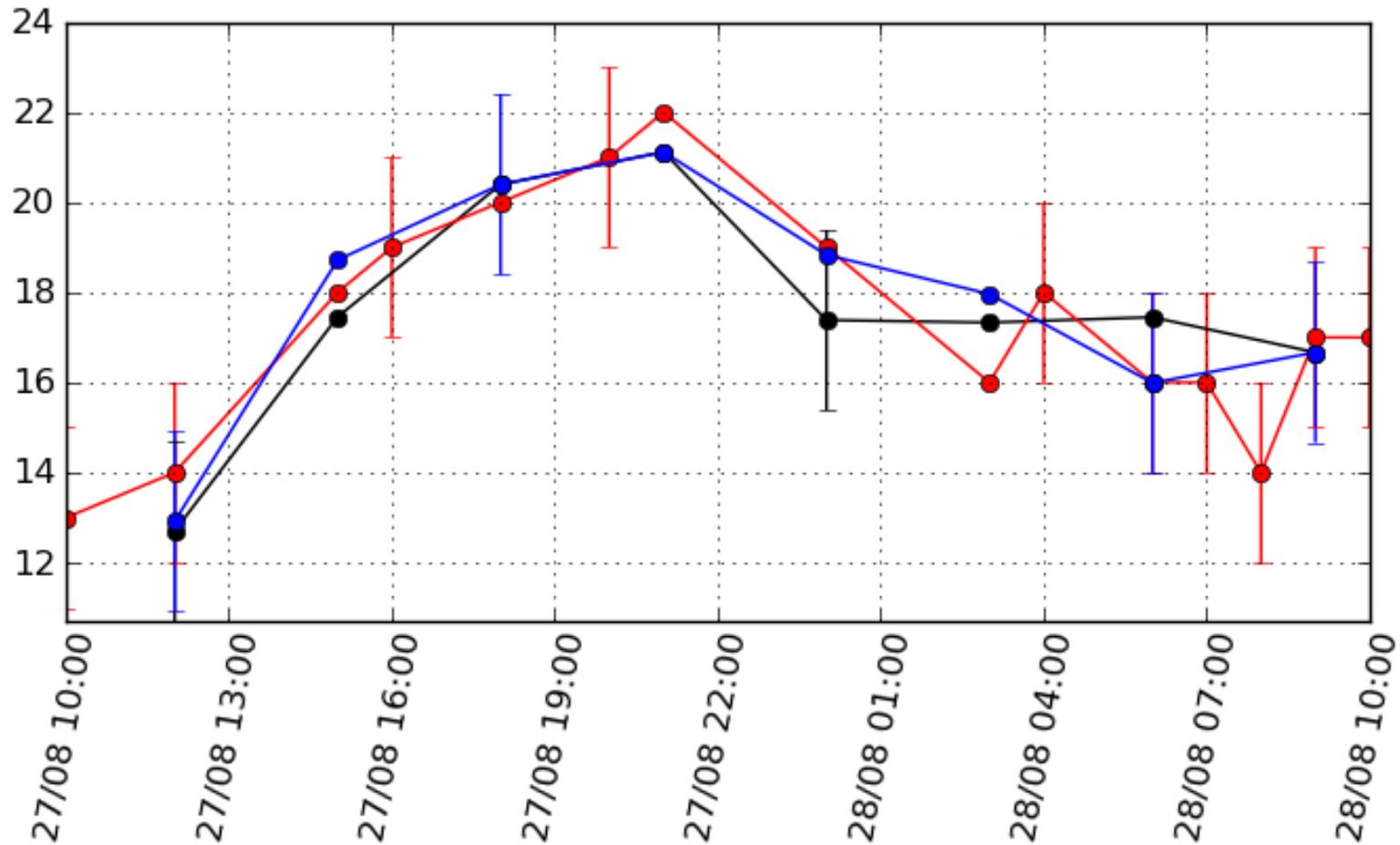
Cette page sert à obtenir et comparer différentes prévisions météo en créant deux graphiques. Le premier contient les données demandées et le deuxième contient les données qui on mener à la correction (jaune) de ce qui est appelé récent(vert). C'est deux informations sont liées à l'emplacement du marqueur vert sur la map. Toutes les autres informations sont liées à la station choisie et à la zone choisie qui y correspond. Soyez patient, la map est très lente du au grand nombre de station et la création des graphiques peut prendre quelques minutes.

Station:	Aéroport int. de Mont ▾
Zone:	Montréal métropolitain ▾
Zone 1:	Vaudreuil - Soulange ▾
Zone 2:	Montréal métropolitain ▾
Zone 3:	Laurentides - r71.6 ▾
Caractéristique:	air_temp ▾
Date:	2010-08-27 10 UTC
Fenetre temporelle:	1 jour(s)
Delai de prevision:	0 heure(s)
Coordonnee geographique:	45.5694051,-73.8009789
Grib de la station	<input checked="" type="checkbox"/>
Grib00 de la station	<input checked="" type="checkbox"/>
Grib du marqueur vert	<input checked="" type="checkbox"/>
Meteocode de la zone	<input checked="" type="checkbox"/>
Nowcast de la station	<input checked="" type="checkbox"/>
Reference:	meteo code ▾
<input type="button" value="Soumettre"/> <input type="button" value="←"/> <input type="button" value="→"/>	

Distance(km)	Stations les plus proche	
8.8785814221	Sainte-Anne-de-Bellevue	WVQ
11.7584291808	Montréal-Trudeau Int'l Airport	YUL
21.0312318966	Montréal-Mirabel Int'l Airport	YMX



Comparaison des prevision hautes resolution et meteoceode



Étape 2: création de table CSV (infos pertinentes pour faire une prédiction)

Base de donnée
SQL
historique

Onglet créer table

Table CSV

Table CSV:

entrées

cible

Zone 1: GRIB, corrections voisins, donnée géographique, meteocode
Zone 2: GRIB, corrections voisins, donnée géographique, meteocode
Zone 3: GRIB, corrections voisins, donnée géographique, meteocode

⋮

(ou nowcast)

Création d'une table CSV

Créateur de table csv

Accueil Visualiser Tester Tables Dataset Créer table Prétraiter Comparer Importer Créer dataset Admin

Ce formulaire crée et ajoute à la base de données une table csv source qui peut soit être prétraitée, soit être directement présentée à un algorithme d'apprentissage. Chaque ligne de la table contient toutes les caractéristiques météorologiques choisies pour un endroit en particulier à un moment en particulier ainsi que les caractéristiques de ses voisins à ce même moment. La latitude, la longitude et l'altitude de chaque endroit et voisin sont aussi disponibles. Si vous choisissez la localisation par station, les endroits seront des stations, les données meteo code seront celles de la zone qui contient cette station et les voisins seront des stations. Si vous choisissez la localisation par zone, les endroits seront la première station de chaque zone et les voisins seront des stations mais en regardant les zones les plus près.

Cliquez [ici](#) pour voir un exemple de création de dataset du début à la fin pour un problème donné.

Province

Alberta:	<input type="checkbox"/>
British_Columbia:	<input type="checkbox"/>
Manitoba:	<input type="checkbox"/>
New_Brunswick:	<input type="checkbox"/>
Northwest_Territories:	<input type="checkbox"/>
Nova_Scotia:	<input type="checkbox"/>
Newfoundland:	<input type="checkbox"/>
Nunavut:	<input type="checkbox"/>
Ontario:	<input type="checkbox"/>
Prince_Edward_Island:	<input type="checkbox"/>
Quebec:	<input type="checkbox"/>
Saskatchewan:	<input type="checkbox"/>
Yukon:	<input type="checkbox"/>

Caractéristiques météos

Meteocode		Grib		Nowcast	
Accum:	<input type="checkbox"/>	CMC_reg_PRATE_SFC_0:	<input type="checkbox"/>	NW_ACC_QP:	<input type="checkbox"/>
Cloud_cover:	<input type="checkbox"/>	CMC_reg_TCDC_SFC_0:	<input type="checkbox"/>	NW_SKY:	<input type="checkbox"/>
Ceiling:	<input type="checkbox"/>	null		NW_CIG:	<input type="checkbox"/>
Precip_prob:	<input type="checkbox"/>	null		NW_POP:	<input type="checkbox"/>

Tables stockées et accessibles pour téléchargement

meteocode

[connectez-vous | mot de passe perdu?]

Listes de tables csv

Accueil Visualiser Tester Tables Dataset Comparer

Vous trouverez ici des tables CSV qui ont été créées par les utilisateurs. La première ligne de la table est toujours une légende et les lignes suivantes contiennent les informations météorologiques et géographiques reliés à un instant précis.

Table source

Fichier	Description	Créateur	Date	#lignes	#colonnes	comp
Exemple_1_0.csv	Table de l'exemple 1	Marc-Elie	2010-12-21 13:23:24	488	34	comparé!
qcMCGR_temp0.csv	qc, air_temp, grib et mc, 3 voisin	Marc-Elie	2010-08-17 16:26:36	160	34	comparé!
PE_air_temp0.csv	pe, TMP, TEMP_T, 2 jours, clean	Marc-Elie	2010-08-12 16:23:56	76	13	comparé!
qcTEMP0.csv	qc, air_temp, 1 jour, gr,nw	Marc-Elie	2010-08-16 08:24:30	547	20	comparé!
qcFullGribTEMP0.csv	qc, full grib vers TEMP_T, 1 jour	Marc-Elie	2010-08-16 17:48:12	237	60	comparé!
peFullGribTEMP0.csv	pe, full grib vers TEMP_T	Marc-Elie	2010-08-17 09:38:23	16	60	comparé!
qc_MCGR6V_temp0.csv	qc, mc et gr, air_temp, 6 voisin, 4 jour	Marc-Elie	2010-08-19 10:31:55	1803	55	comparé!
qc_temp_6V_18_080.csv	qc, air_temp, grib et mc, 6v, 18/08	Marc-Elie	2010-08-19 13:21:15	854	55	comparé!
qc_temp_10V0.csv	qc, air_temp, mc et gr, 10 voisin	Marc-Elie	2010-08-19 15:40:46	1563	97	comparé!
qc_temp_10V_17_08_0.csv	qc, air_temp, mc et gr, 10 voisin autre date	Marc-Elie	2010-08-19 15:44:37	1094	97	comparé!

Table prétraitée

Fichier	Parent	Description	Créateur	Date	#lignes	#colonnes
Exemple_1_00.csv	Exemple_1_0.csv	Table prétraitée de l'exemple 1	Marc-Elie	2010-12-21 13:24:55	489	21
PE_air_temp00.csv	PE_air_temp0.csv	pe, grib vers nowcast, air_temp	Marc-Elie	2010-08-12 16:24:59	53	8
qcTEMP00.csv	qcTEMP0.csv	cible:nowcast, air temp	Marc-Elie	2010-08-16 08:30:13	298	8
qcFullGribTEMP00.csv	qcFullGribTEMP0.csv	qc, full grib vers TEMP_T	Marc-Elie	2010-08-16 17:57:40	238	18
peFullGribTEMP00.csv	peFullGribTEMP0.csv	pe, full grib vers TEMP_T	Marc-Elie	2010-08-17 09:39:30	17	18
qcMCGR_temp00.csv	qcMCGR_temp0.csv	qc, gr vers mc air_temp, delta, avec voisin	Marc-Elie	2010-08-17 16:35:04	161	21
qc_MCGR6V_temp00.csv	qc_MCGR6V_temp0.csv	qc, air_temp, grib vers mc, 6 v, delta, 4 jours	Marc-Elie	2010-08-19 10:32:54	1804	33
qc_temp_6V_18_0800.csv	qc_temp_6V_18_080.csv	qc, air_temp, grib vers mc, 6 v, delta, 3 jours	Marc-Elie	2010-08-19 13:25:04	855	33
qc_temp_10V00.csv	qc_temp_10V0.csv	qc, air_temp, mc vers gr, 10 voisin- 05/08/2010	Marc-Elie	2010-08-19 15:46:01	1564	57
qc_temp_10V_17_08_00.csv	qc_temp_10V_17_08_0.csv	qc, air_temp, mc vers gr, 12 voisin- 17/08/2010	Marc-Elie	2010-08-19 15:46:35	1095	57

Étape 3: prétraitements (et préformattage pour l'entraînement de prédicteur par algorithmes d'apprentissage)



meteocode

Welcome Marc-Elie [déconnectez-vous | profile | mot de passe]

Accueil projet meteocode

Accueil Visualiser Tester Tables Dataset **Créer table** **Prétraiter** Comparer Importer **Créer dataset** Admin

Sur ce site vous trouverez différentes parties du projet de recherche meteocode. Le but de ce projet est d'obtenir des prévisions météorologiques 00-48 heures de haute qualité à très grande résolution. À terme, le projet de recherche permettrait au publique d'avoir accès à des prévisions très près de chez lui qui sont de très bonnes qualité. Pour ce faire nous avons accès à plusieurs sources de données d'environnement Canada qui sont disponibles sous leur onglets respectifs sur le site d'EC: <http://dd.weatheroffice.ec.gc.ca/>.

Nous stockons présentement trois types de données dans notre base de donnée: les prévisions brutes qui sortent des modèles d'EC en haute résolution (GRIB), les prévisions corrigées par les météorologues (meteocode) et les observations de ce qui c'est réellement produit (nowcast). Nous recueillons les données GRIB et nowcast pour chaque station météorologique du Canada et les meteocode seulement au Québec.

- Pour visualiser les données cliquer sur l'onglet [visualiser](#).
- Pour avoir accès aux tables CSV déjà créé cliquer sur l'onglet [table](#).
- Pour comparer deux types de données cliquer sur l'onglet [comparaison](#).
- Pour tester les différents modèles de corrections disponibles cliquer sur l'onglet [tester](#).

Pour [créer](#) une nouvelle table, [prétraiter](#) une table source, [créer](#) un dataset ou [importer](#) un nouveau modèle vous devez être connecté.

Cliquez [ici](#) pour voir un exemple de création de dataset du début à la fin pour un problème donné.

Datasets prétraités accessibles pour téléchargement

meteocode

Welcome Marc-Elie [[déconnectez-vous](#) | [profile](#) | [mot de passe](#)]

Liste de datasets

[Accueil](#) [Visualiser](#) [Tester](#) [Tables](#) [Dataset](#) [Créer table](#) [Prétraiter](#) [Comparer](#) [Importer](#) [Créer dataset](#) [Admin](#)

Vous trouverez ici des datasets qui ont été créé par les utilisateurs.

Datasets

Fichier	Description	Créateur	Date	#x	#y	#exemples	table	localisation	public	del
qc_hasard	qc hasard	Marc-Elie	2010-11-28 15:31:09	5	1	237	qcFullGribTEMP00.csv	hasard	makePrivate	supprimer
qc_geo	qc geo test	Marc-Elie	2010-11-28 15:36:38	5	1	237	qcFullGribTEMP00.csv	geographique	makePrivate	supprimer
Exemple_1_temp_3voisins	Dataset de l'exemple 1	Marc-Elie	2010-12-21 13:26:08	16	1	488	Exemple_1_00.csv	hasard	makePrivate	supprimer

Étape 4: recherche d'algorithmes

En cours...

**Dataset
python**

**Algorithmes
d'apprentissage machine**

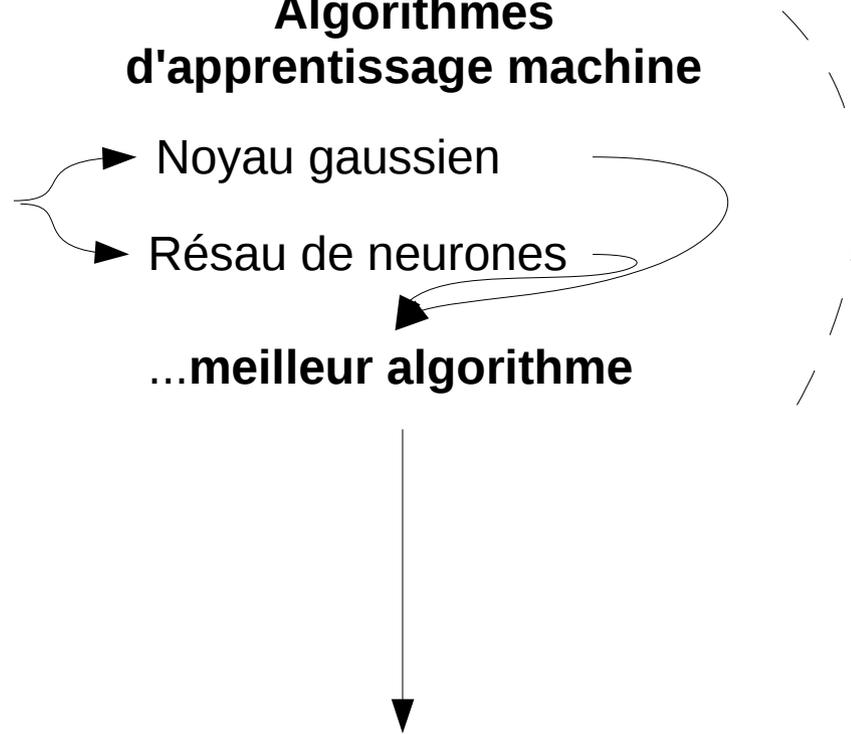
→ Noyau gaussien

→ Réseau de neurones

...meilleur algorithme

Recherche

**Fonction de
correction**



Importer des modèles prédictifs /tester ces modèles

meteocode

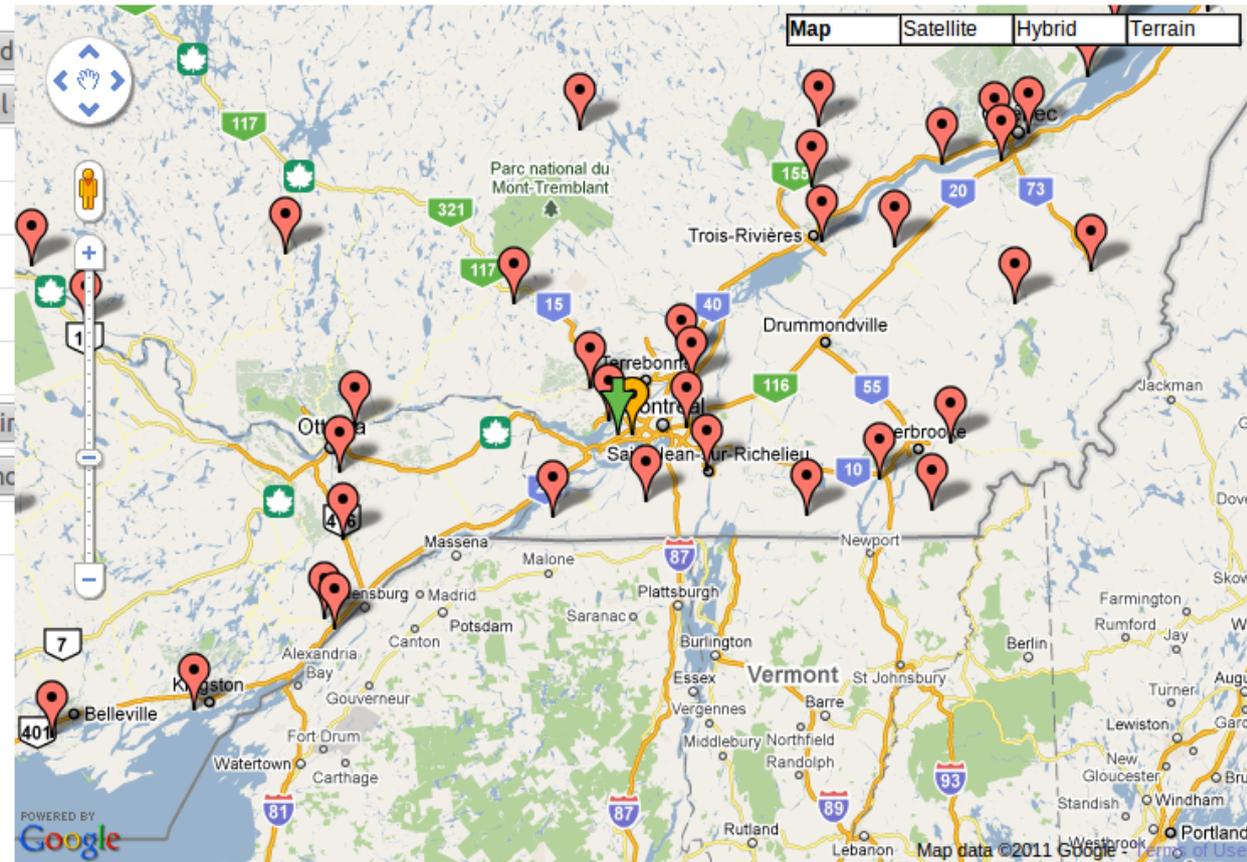
[login | lost password?]

Testeur de modèle de correction

Accueil Visualiser Tester Tables Dataset Comparer

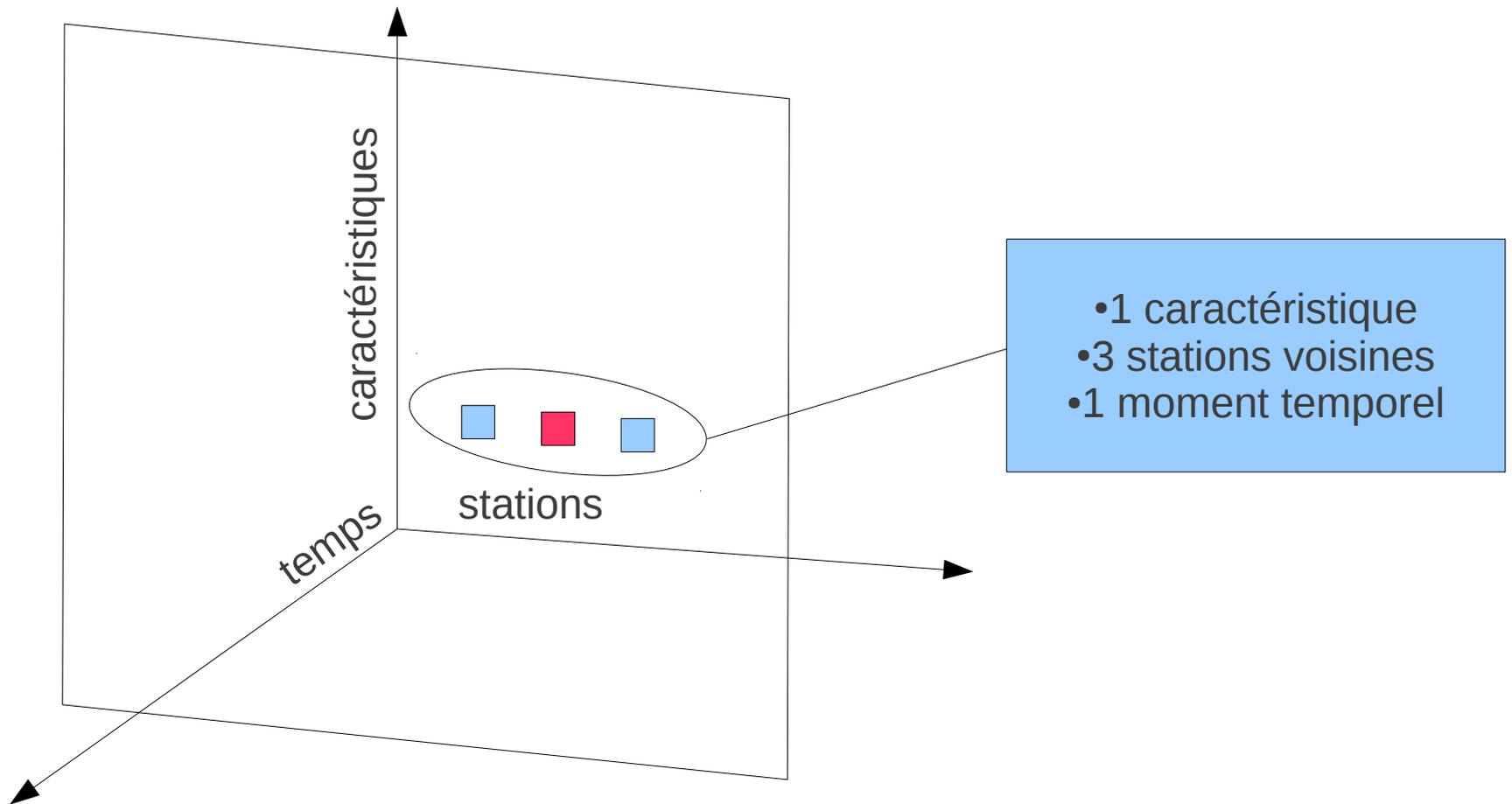
Cette page sert à comparer les corrections des différents modèles. Si vous sélectionnez une station, les meteocode, nowcast et grib00 tenteront d'être récupérées et les RMS calculé par rapport aux grib et aux corrections. Si vous sélectionnez tout autre point sur la map avec le marqueur vert, les grib récent et les corrections seront disponibles. Par contre, aucun RMS ne sera calculé parce qu'il n'y a pas de données de référence disponibles pour ces points.

Station:	Aéroport int. de Montréal-Trud
Zone:	Montréal métropolitain - Laval
Caracteristique:	CMC_reg_TMP_TGL_2
Date:	2011-01-04 14
Fenetre temporelle:	2
Leadtime:	0
Coordonnee géographique:	45.47,-73.75
Model1:	moyenne_3v - moyenne, 3 voisin
Model2:	noyau_6v - moy. pondere par no
	Soumettre <- ->



Avenir...

- Pour l'instant, modèle très simple basé sur une fraction des information disponibles
- **Possibilité de modèles beaucoup plus complexes...**



Avenir...

- Pour l'instant, modèle très simple basé sur une fraction des informations disponibles
- **Possibilité de modèles beaucoup plus complexes...**

