

Données climatiques

Acquisition, Interprétation et
Manipulation

Trevor James Smith

Spécialiste Plateformes climatiques, données et opérations
Sciences du climat et services climatiques

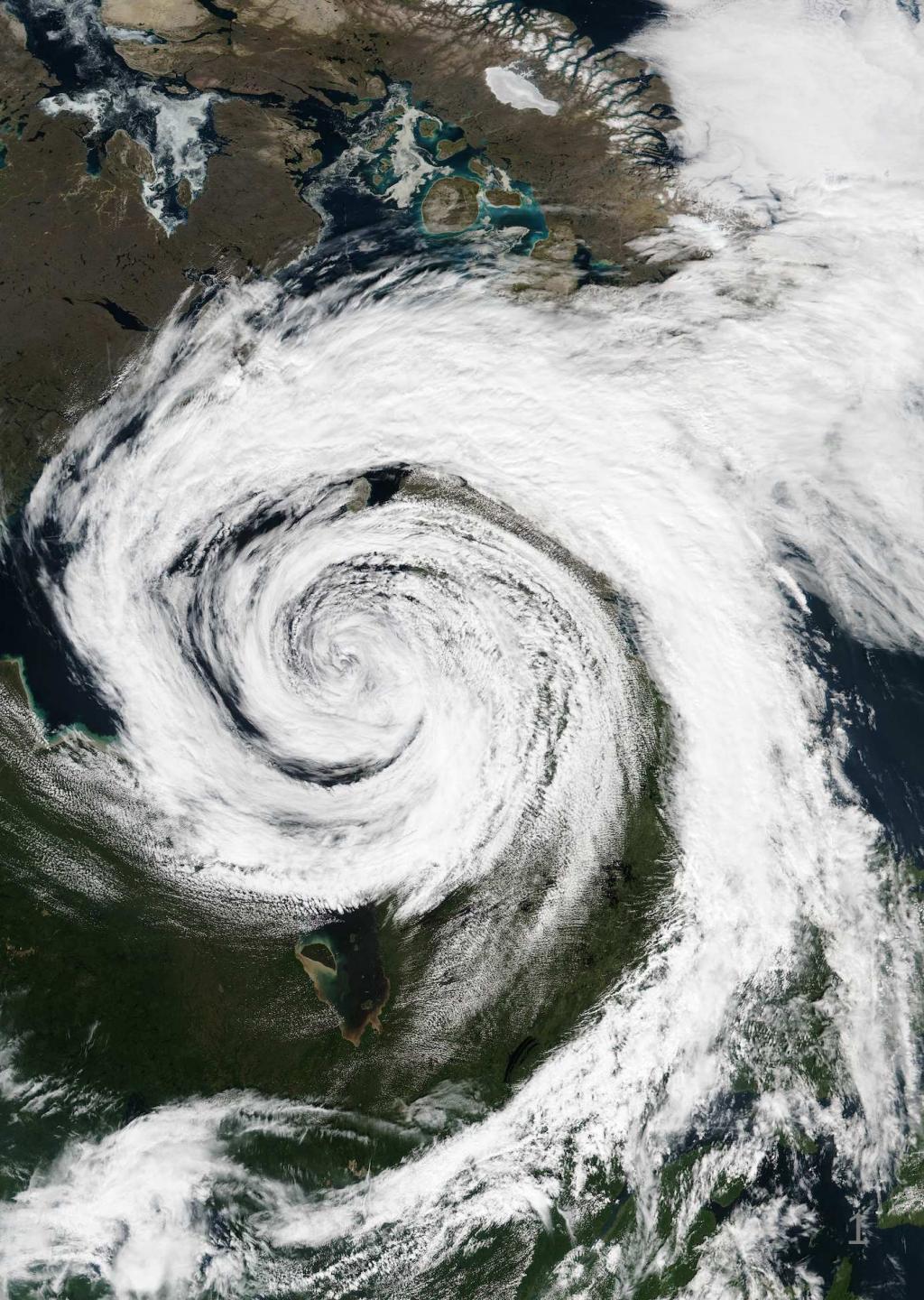
École d'été en science du climat et des changements climatiques
Domaine Forget, Sainte Irénée, Québec
30 mai, 2024



 Ouranos

Objectifs du cours

- Connaissance de base de l'historique de la collecte de données climatiques, du fonctionnement des différents instruments et de leurs incertitudes.
- Vue d'ensemble des données disponibles pour faire des études climatiques.
- Connaissance des forces et faiblesses des différents types de jeux de données.
- Compréhension de la taille considérable de certaines des bases de données et des défis à les analyser.



Qui suis-je ?



Trevor James Smith



github.com/Zeitsperre

- Développeur/packageur/mainteneur des logiciels de recherche scientifiques
- M.Sc. en Géographie, environnement et urbanisme de l'Université Concordia
 - Impacts des changements climatiques sur les vignobles en sud du Québec
- 10+ années d'expérience en utilisant Python, Linux, GIS, et autres technologies associées

C'est quoi Ouranos ?

- Organisation à but non-lucratif, établi 2002 à Montréal, Québec
 - Crée en réponse à la tempête de verglas massif de janvier 1998
- Plans d'adaptations aux changements climatiques
- Source/producteur des données des modèles climatiques
- Services d'informations climatiques



Regional Climatology and Adaptation to Climate Change

Ouranos is a collaborative innovation hub enabling Quebec society to better adapt to an evolving climate. The consortium brings together upwards of 450 researchers, experts, practitioners and decision-makers from an array of disciplines, collectively working on numerous applied research programs and projects.

SEARCH A PROJECT OR PUBLICATION

EXPLORE OUR PROGRAM



Contenu du cours

1. Historique des instruments d'observations météorologiques
2. Fonctionnement des instruments d'observations météorologiques
3. Données observées
4. Données de réanalyses et de prévisions
5. Données de modèles climatiques
6. Données dérivées
7. Formats habituels des données climatiques et outils d'analyses
8. Les plateformes et services climatiques

1. Historique des instruments d'observations météorologiques.

Pour en savoir plus :

- Rétrospective de NOAA : (<https://vlab.noaa.gov/web/nws-heritage>)

1743

Benjamin Franklin fait des observations des mouvements de tempêtes et patrons météorologiques.



1767

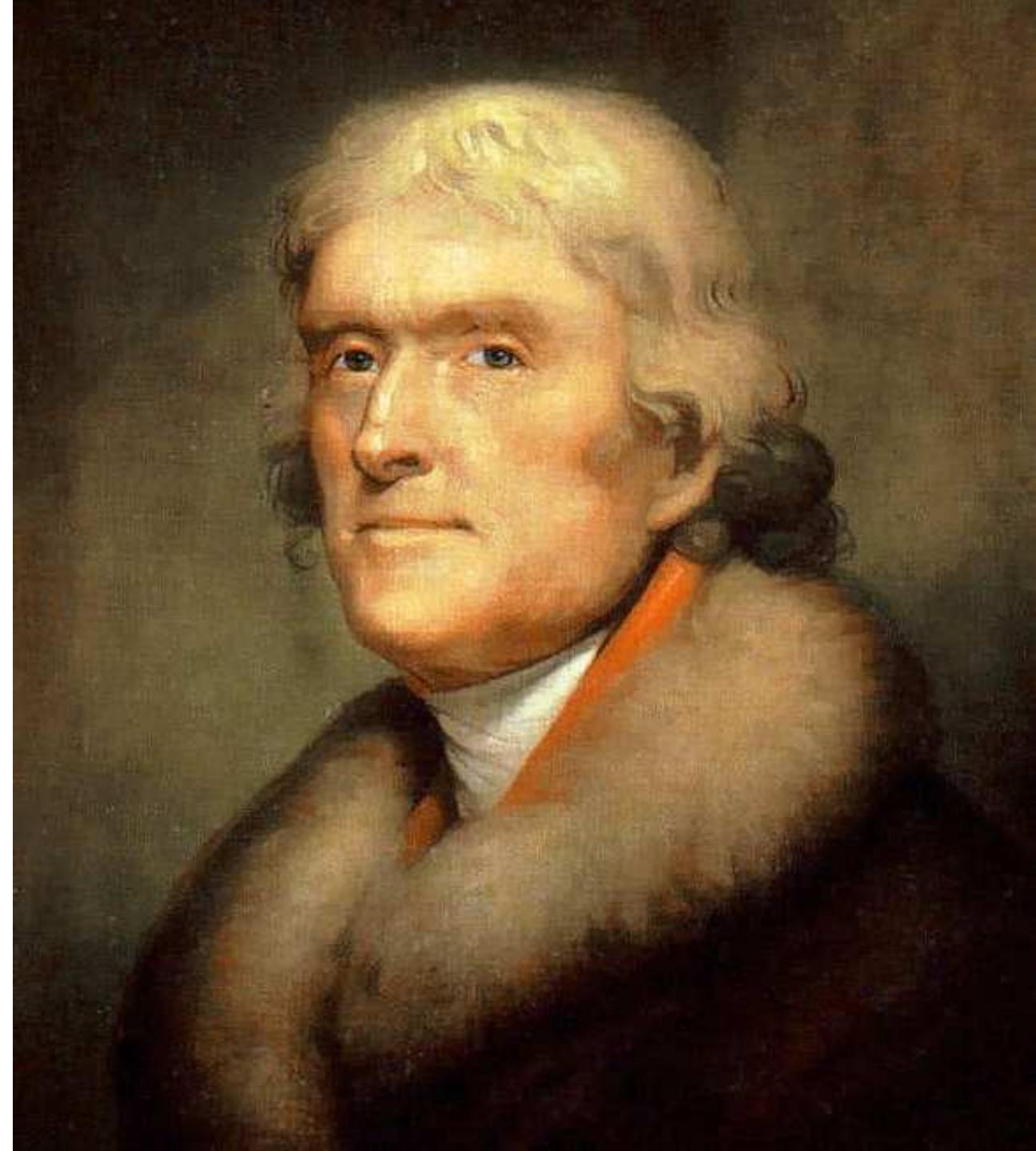
**Georges Washington
commence un journal
d'observations
météorologiques.**

1776 à 1800

Thomas Jefferson recrute des observateurs météo bénévoles en Virginie.

Le réseau de bénévoles s'étend dans 5 états :

- Massachusetts, Pennsylvanie, Connecticut, New York, Caroline du Nord





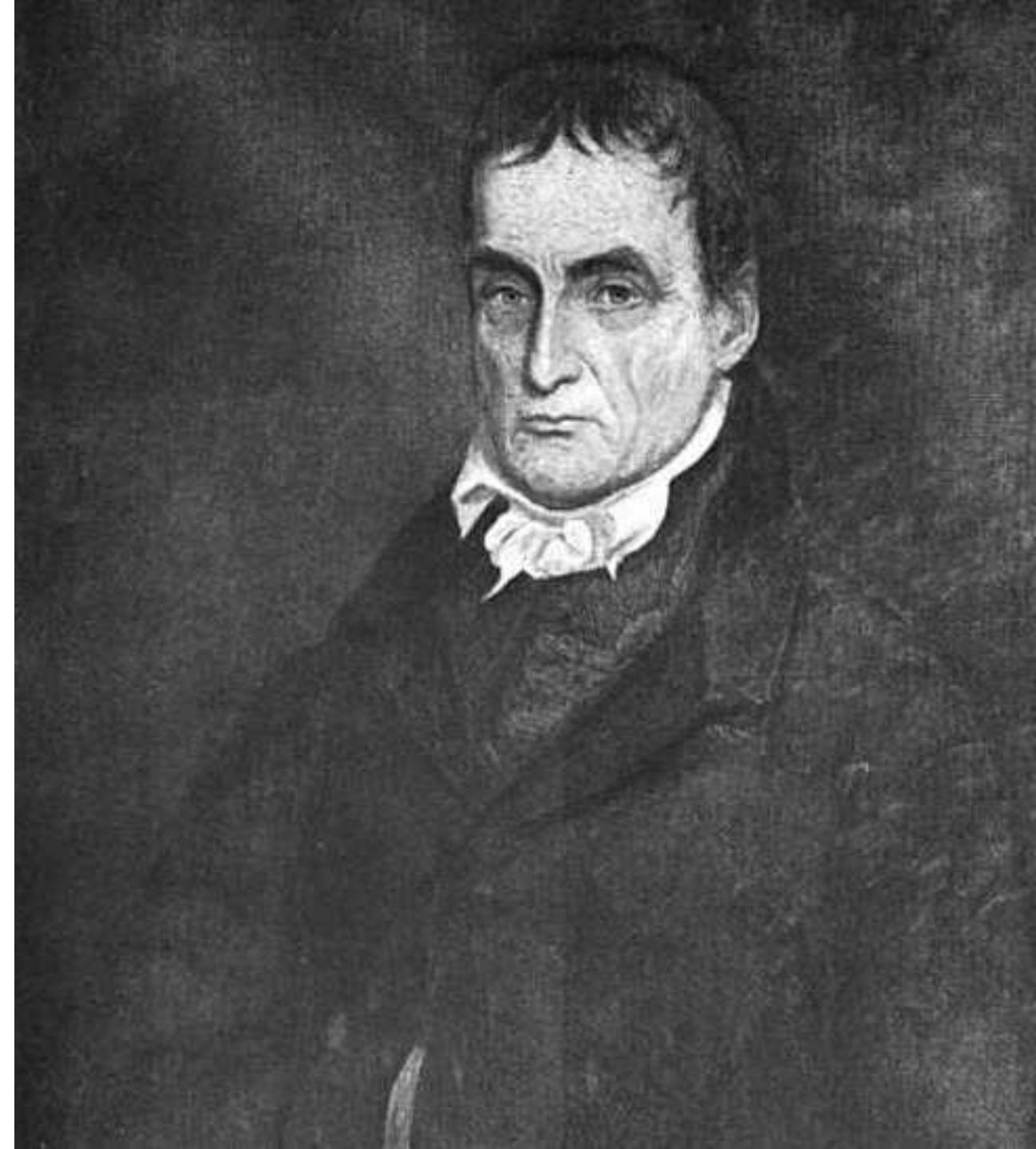
1803 / 1804

L'expédition de Lewis et Clark vers l'ouest américain fait des observations météorologiques régulières

- <https://lewisandclarkjournals.unl.edu/item/lc.jrn.1803-08-30>

1814

James Tilton ordonne les observations météorologiques dans les postes militaires.



1848 à 1860

Joseph Henry inaugure un réseau télégraphique d'observateurs de 150 bénévoles.

500 stations fournissent des rapports télégraphiques journaliers

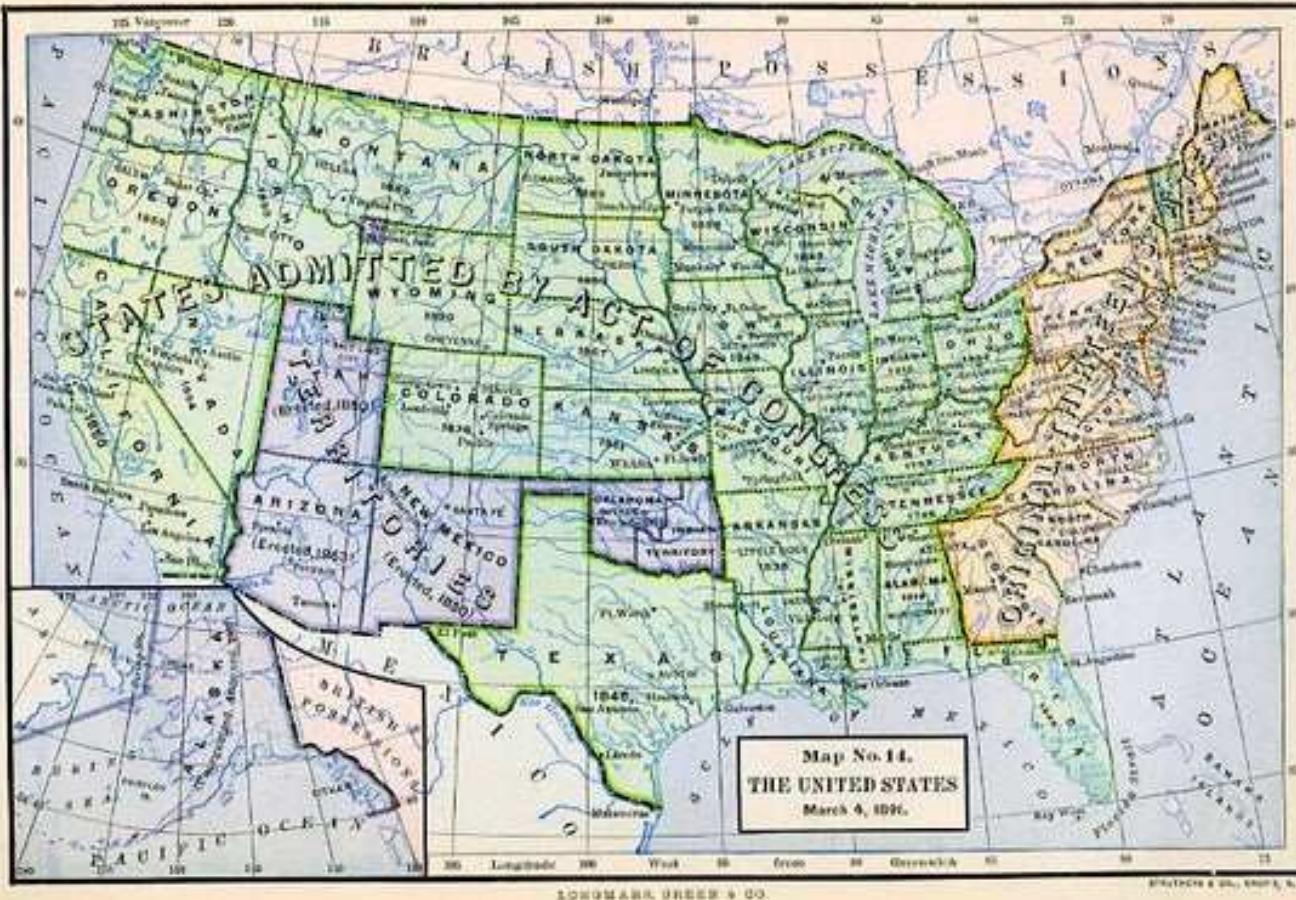
1870

Ulysses S. Grant (à gauche)
autorise un service météo
national sous la direction de
Albert J. Myer (à droit).



1891

Le réseau
d'observateurs météo
bénévoles contenait
plus de 2000 stations
autour du pays.



Données climatiques : Acquisition, interprétation et manipulation

1898

**Le U.S. Weather Bureau
expérimente avec des cerf-
volants.**





1909

Début de l'utilisation de ballons.

Données climatiques : Acquisition, interprétation et manipulation

1926 à 1934

L'aviation change le rôle du Weather Bureau qui doit fournir des services météorologiques à l'aviation civile.





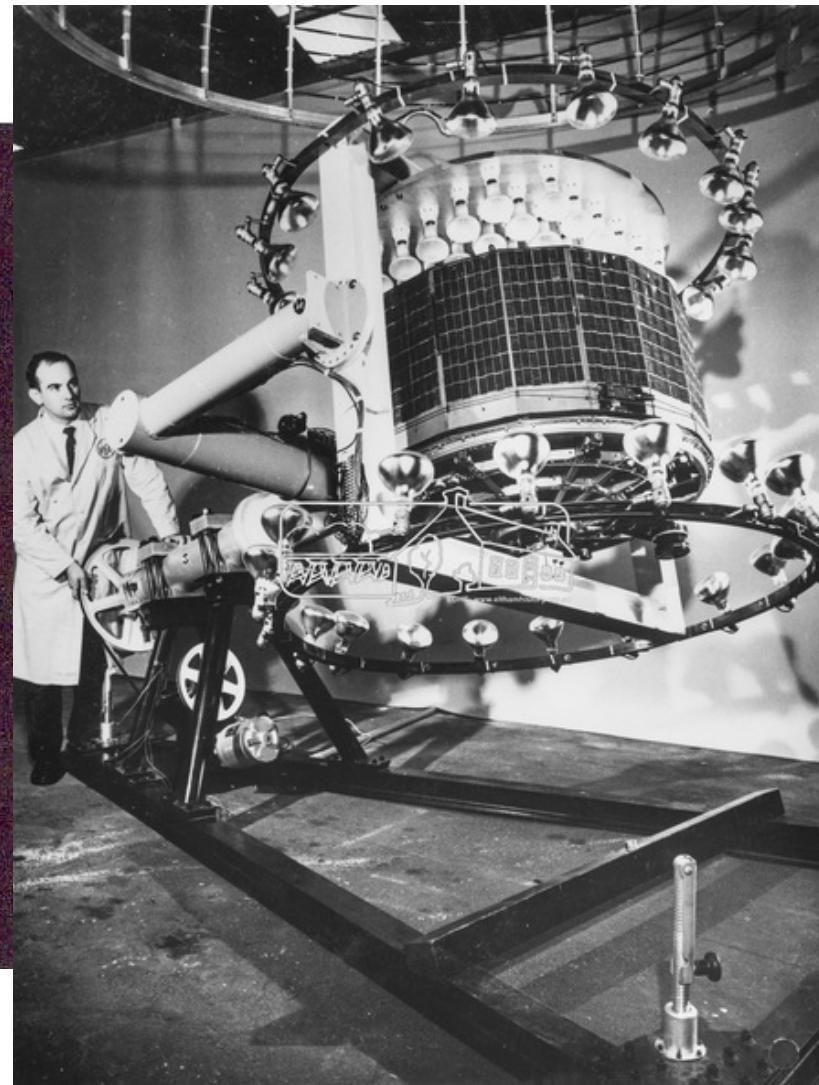
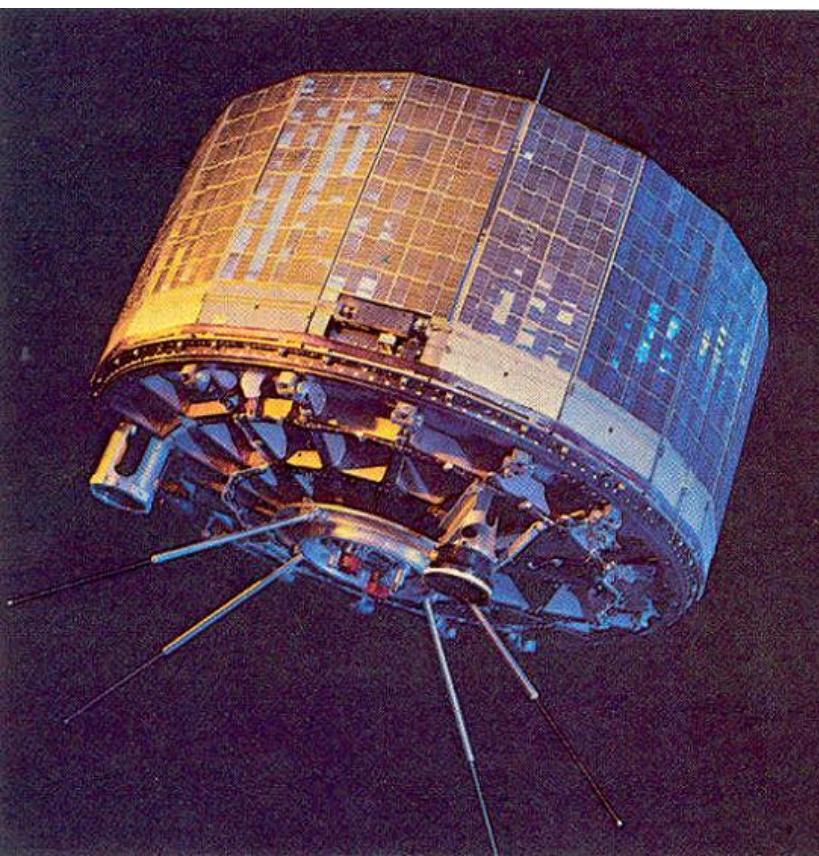
1939

**L'utilisation de radiosondes
remplace les observations par
avions.**

1950

**Les technologies radar
développées durant la 2e
guerre mondiale sont adaptées
aux observations
météorologiques.**





1960

Les premiers satellites d'observations météorologiques sont lancés.

- TIROS-I (à gauche)
- TIROS-II (à droit)

Photos : NASA

2. Fonctionnement des instruments d'observations météorologiques

Station d'observation météo

École d'été en sciences du climat 2024

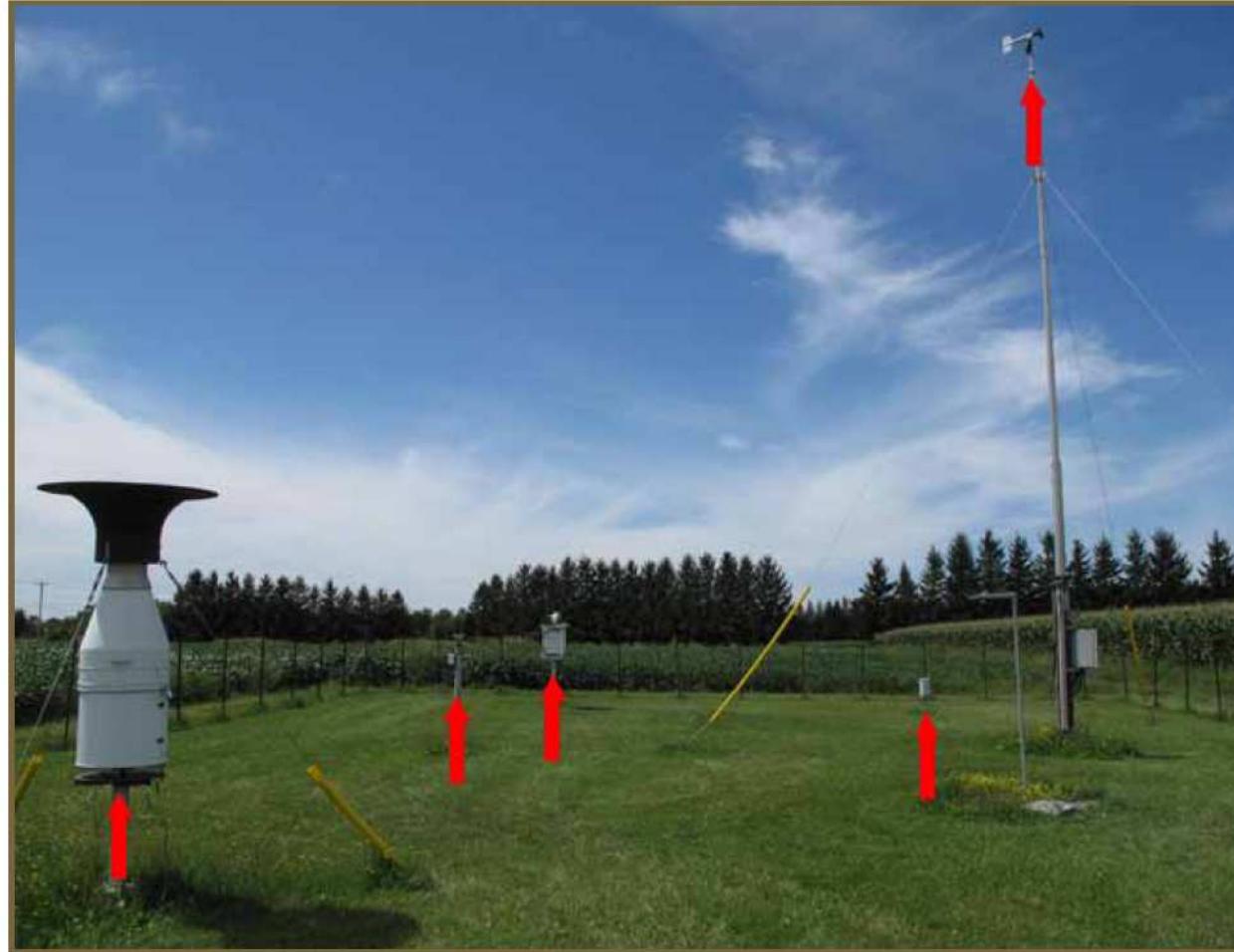


FIGURE 2. STATION MÉTÉOROLOGIQUE (DE GAUCHE À DROITE : PLUVIOMÈTRE À PESÉE, PYRANOMÈTRE, ABRI MÉTÉOROLOGIQUE, PLUVIOMÈTRE À AUGET ET ANÉMOMÈTRE À UNE HAUTEUR DE 10 M)

Photo : Marie-Pier Lepage

Station d'observation météo

Observation « **in situ** »

- Observation obtenue par contact direct avec l'objet en question.

Stations à observateurs : un ou deux temps d'observations par jour.

Au Québec, les heures d'observations sont 8h et 18h heure normale de l'est.

Stations automatiques : observations possibles à toutes les heures.

	année	mois	jour
DATE			
PRÉCIPITATION	08:00 (HNE)	18:00 (HNE)	
Pluie (pluviomètre)	,	,	,
Neige (table à neige)	,	,	,
1° Début	:	:	:
1° Fin	:	:	:
2° Début	:	:	:
2° Fin	:	:	:
Neige au sol (échelle)			
TEMPÉRATURE			
Maximum	,	,	,
Minimum	,	,	,
Instantanée	,	,	,
NÉBULOSITÉ			
VENT			
Direction			
Vitesse			
AUTRES PHÉNOMÈNES	(↓)	(↓)	
Brouillard-Brume			
Giboulée			
Grêle			
Orage-Tonnerre			
Poudrerie			
Vent violent ($\geq 62 \text{ km/h}$)			
Verglas			
Visibilité $\leq 400 \text{ m}$			
NIVOMÈTRE (Nipher)			
Pluie	,	,	,
Neige	,	,	,
Équivalent en eau	,	,	,

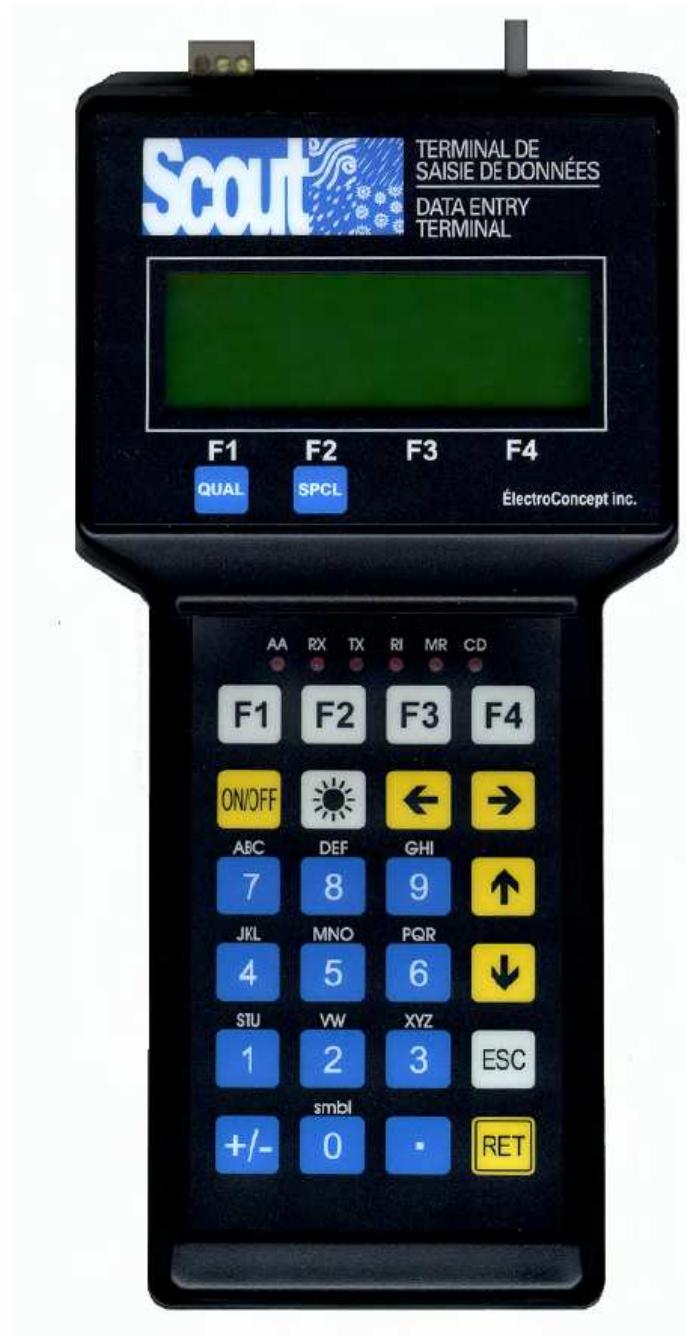


Figure 5.1 Le carnet de l'observateur, relevé biquotidien

Outils de station d'observation météo

source: MELCC

Thermomètre

La température minimum et maximum de l'air est mesurée à l'aide d'une combinaison d'un thermomètre à mercure et thermomètre à alcool.

Pour les stations automatiques, une thermistance est utilisée.

Au Québec, les mesures des observateurs se font à une précision d'un demi-degré Celsius

Pluviomètre

Données climatiques : Acquisition, interprétation et manipulation



Pluviomètre



Table à neige





Nivomètre à écran de « Nipher »

Données climatiques : Acquisition, interprétation et manipulation

« Single Alter » et « DFIR »

École d'été en sciences du climat 2024



Source : Landolt and Rasmussen (2007) Three-dimensional airflow analysis inside snow gauge shielding to determine snow gauge collection efficiencies

« Sous-captation »

La **sous-captation** est un phénomène où le pluviomètre ou nivomètre sous-estime la quantité de précipitation tombée en période de grands vents ou lors de fortes tempêtes.

Au Québec, les mesures se font à une précision d'un dixième de millimètre.

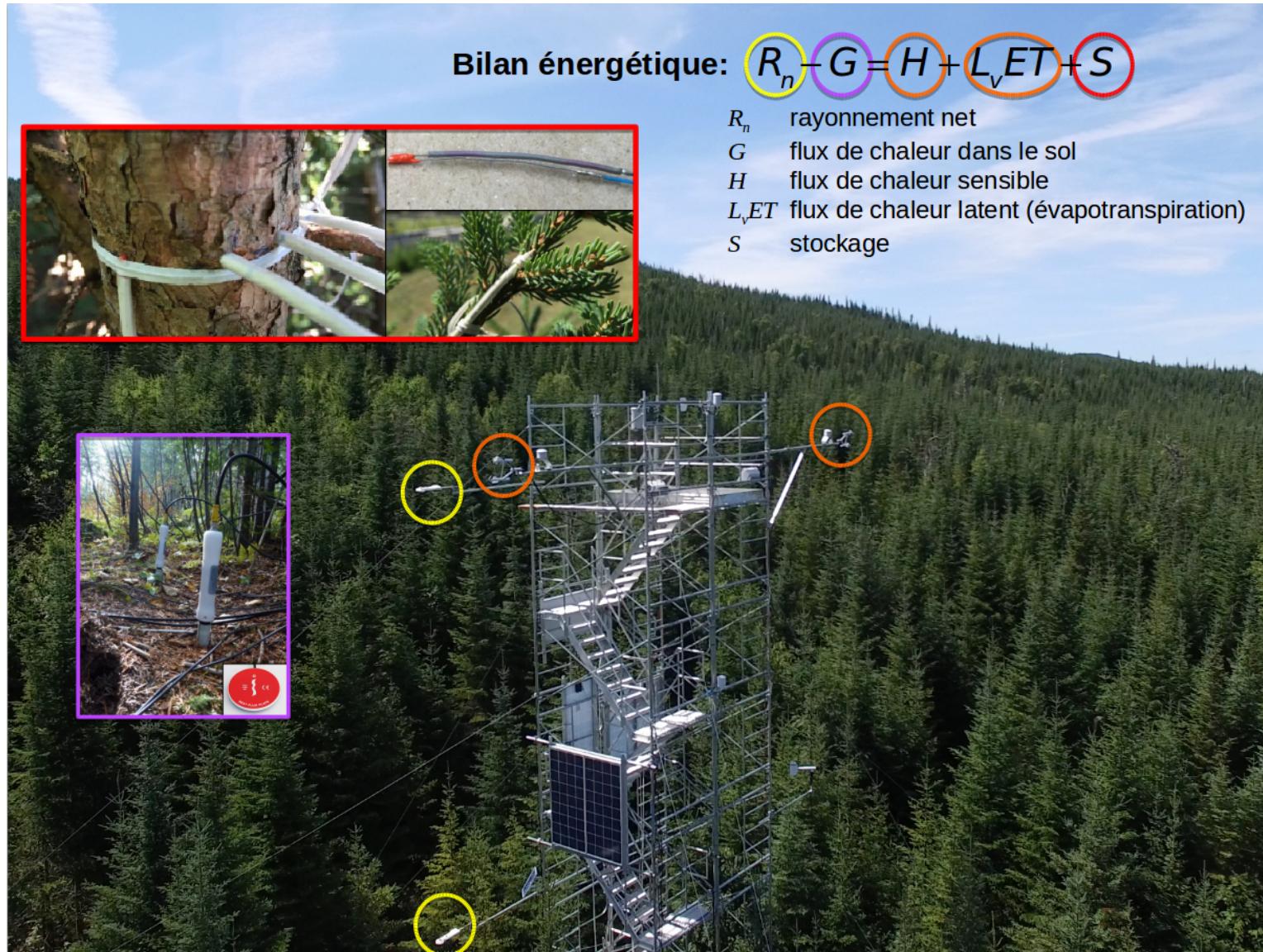


« Sous-captation »



Site expérimental « BEREV »

École d'été en sciences du climat 2024

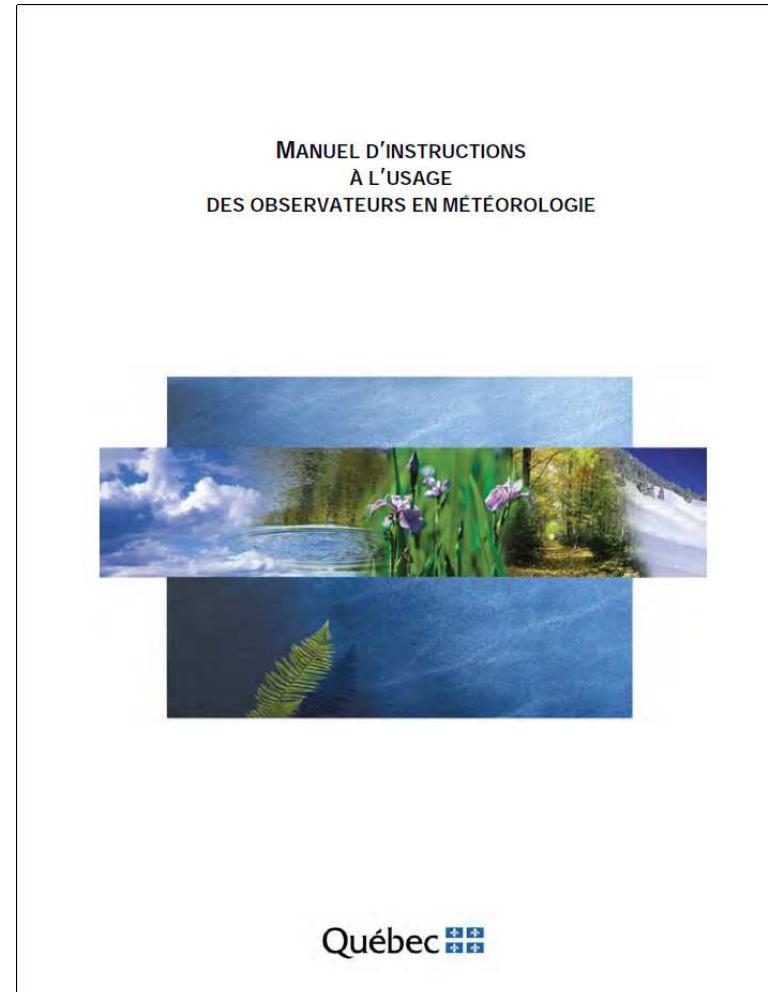
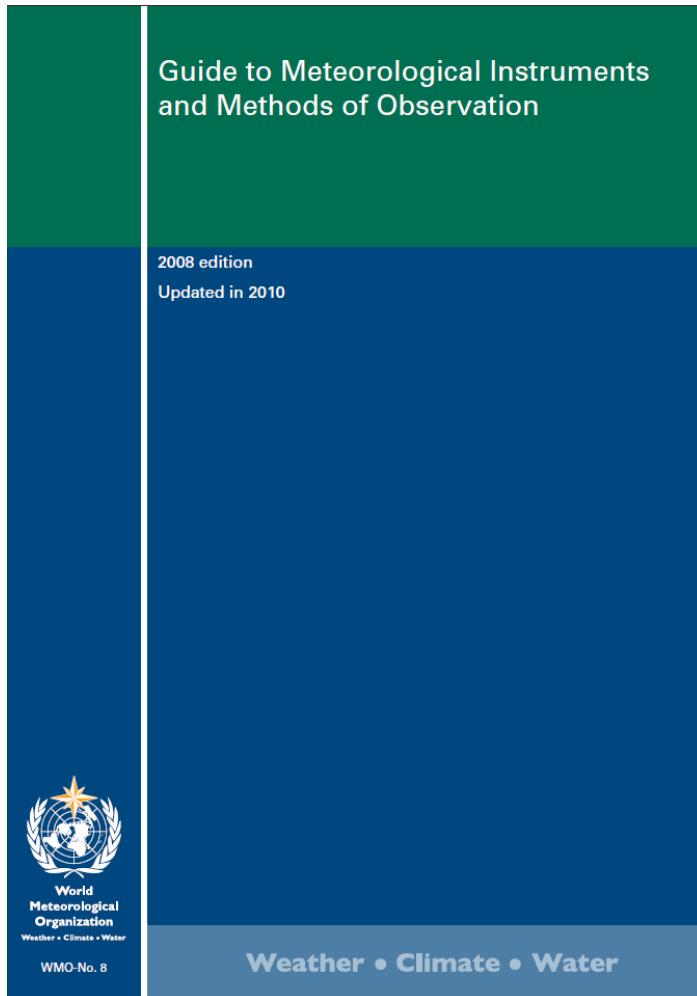


Source : Bassin expérimental du ruisseau des Eaux-Volées (BEREV) de la Forêt Montmorency ([lien](#))

Échelle à neige



Standards des stations d'observation météo



National Weather Service
Observing Handbook No. 2

Cooperative Station Observations

Observing Systems Branch
Office of Systems Operations
Silver Spring, Md.
July 1989

U.S. DEPARTMENT OF COMMERCE
Robert Mosbacher, Secretary
National Oceanic and Atmospheric Administration
William E. Evans, Under Secretary and Administrator
National Weather Service
Elbert W. Friday, Jr., Director

Stations d'observation météo *indépendants*



Photo : UQAM

Source : Station de météo UQAM (<https://eos.meteo.mcgill.ca/stations/4/live-data>)

Radiosonde

La **radiosonde** est composée d'un ensemble d'instruments d'observations attachés à un ballon qui s'élève dans l'atmosphère.

Permet d'obtenir la distribution verticale de température, pression et humidité jusqu'à *30 Km d'altitude*.

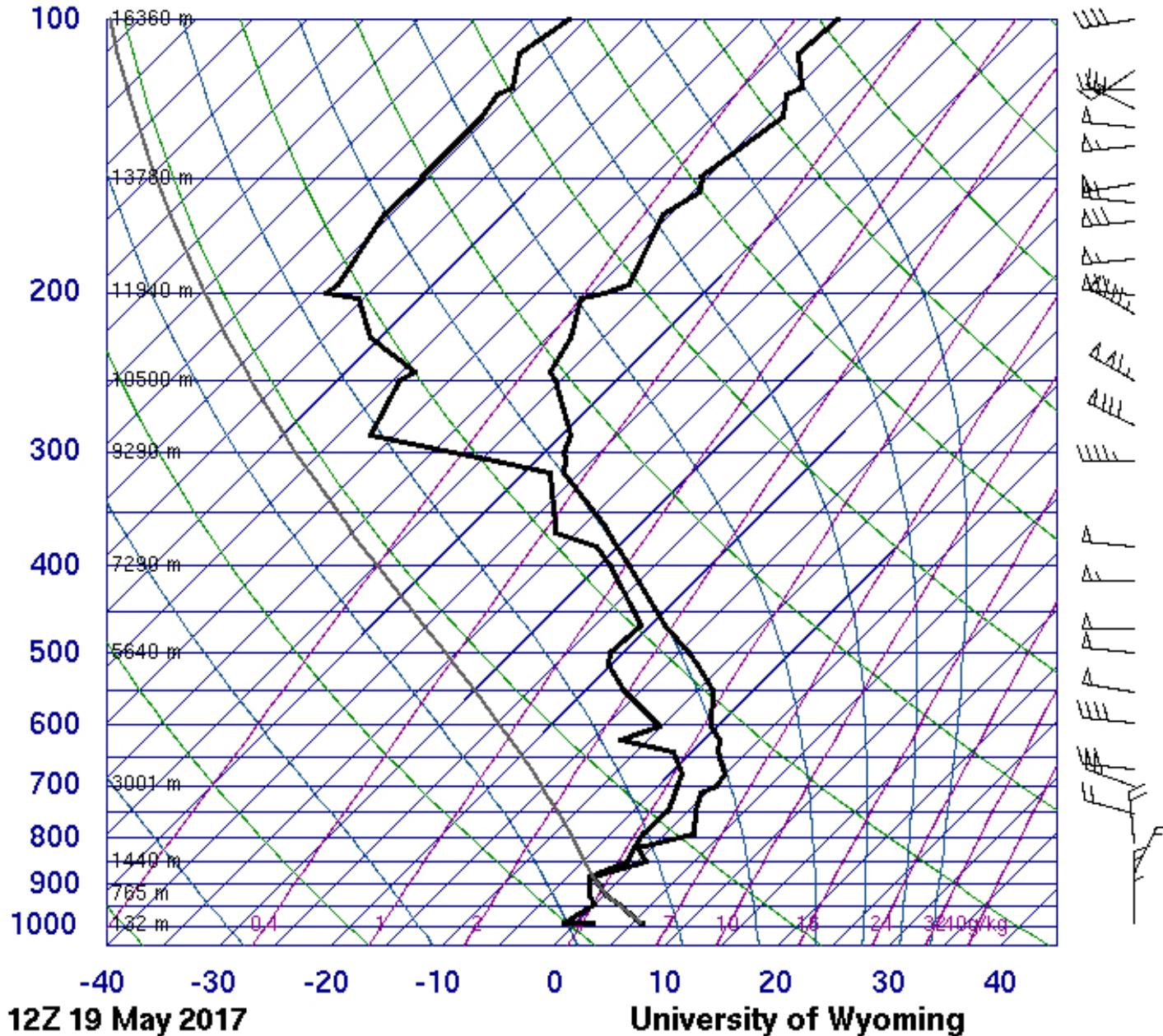
Avec un équipement de localisation au sol, il est possible d'obtenir le profil vertical de vent.

Radiosonde

Données climatiques : Acquisition, interprétation et manipulation



71722 WMW Maniwaki



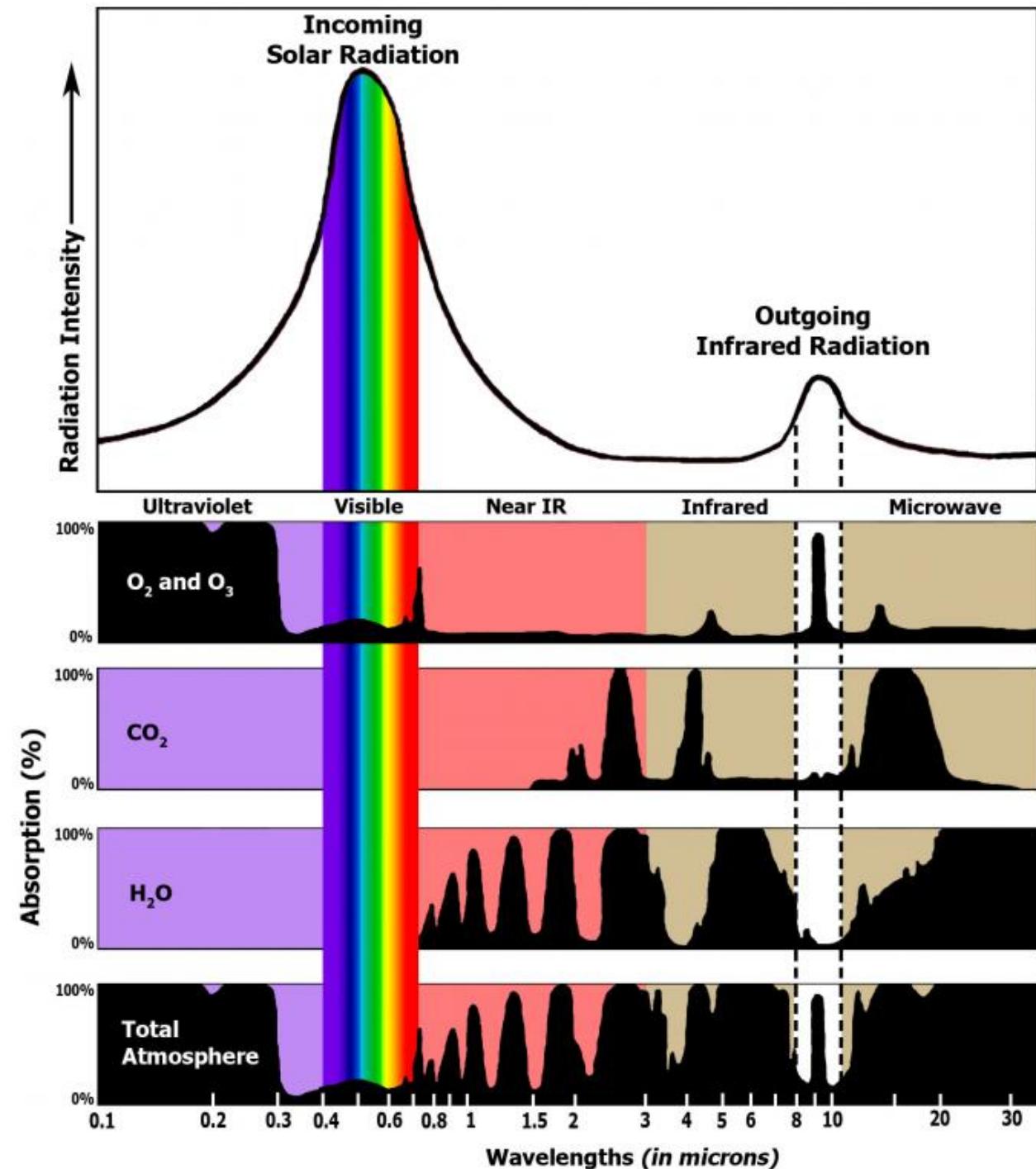
SLAT	46.30
SLON	-76.01
SELV	189.0
SHOW	15.40
LIFT	21.79
LFTV	21.96
SWET	88.00
KINX	10.50
CTOT	13.20
VTOT	14.90
TOTL	28.10
CAPE	0.00
CAPV	0.00
CINS	0.00
CINV	0.00
EQLV	-9999
EQTV	-9999
LFCT	-9999
LFCV	-9999
BRCH	0.00
BRCV	0.00
LCLT	271.7
LCLP	905.2
MLTH	279.6
MLMR	3.84
THCK	5508.
PWAT	19.62

Données de radiosonde

Données climatiques : Acquisition,
interprétation et manipulation

Télédétection

- Différence entre les bandes spectrales ou l'intensité ou fréquence des signaux
- Utile pour identifier la couverture des sols, la distribution spatiale des émissions ou d'espèces atmosphériques
- Essentielle pour générer des topographiques spatiales



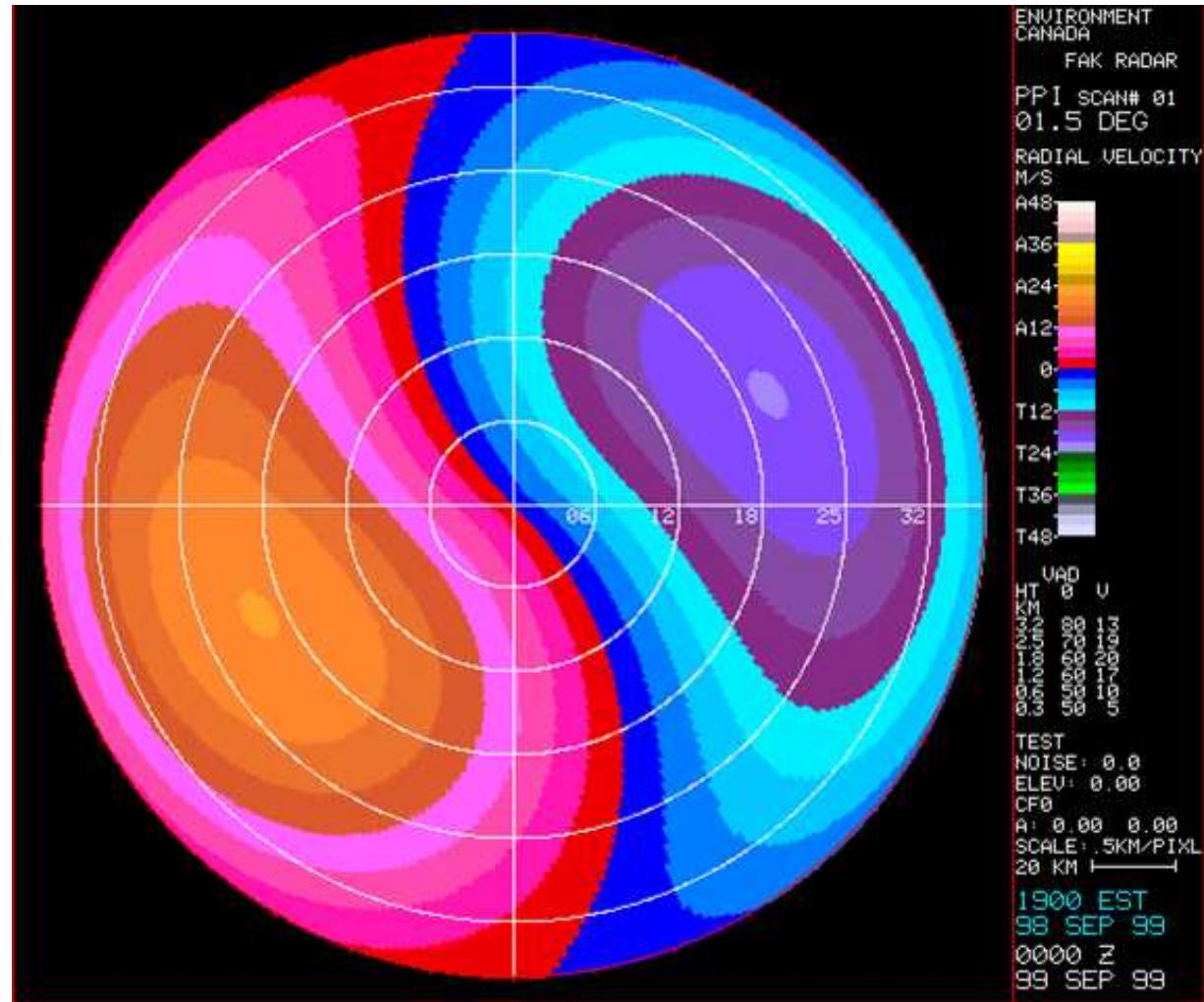
Radar - Observation de télédétection **in situ**

Obtenue sans contact avec l'objet en question

- Portée d'environ **200 Km** autour de l'installation

Trois composantes majeures :

- **un émetteur d'ondes radioélectriques,**
- **une antenne, et**
- **un receveur d'onde**



Installations de radar

- Photo : **Observatoire de J.S. Marshall, Sainte-Anne-de-Bellevue, Québec**
- Pour en savoir plus:
 - [Limite de visibilité du radar de l'université McGill](#)



Installations de radar



Relation Z-R

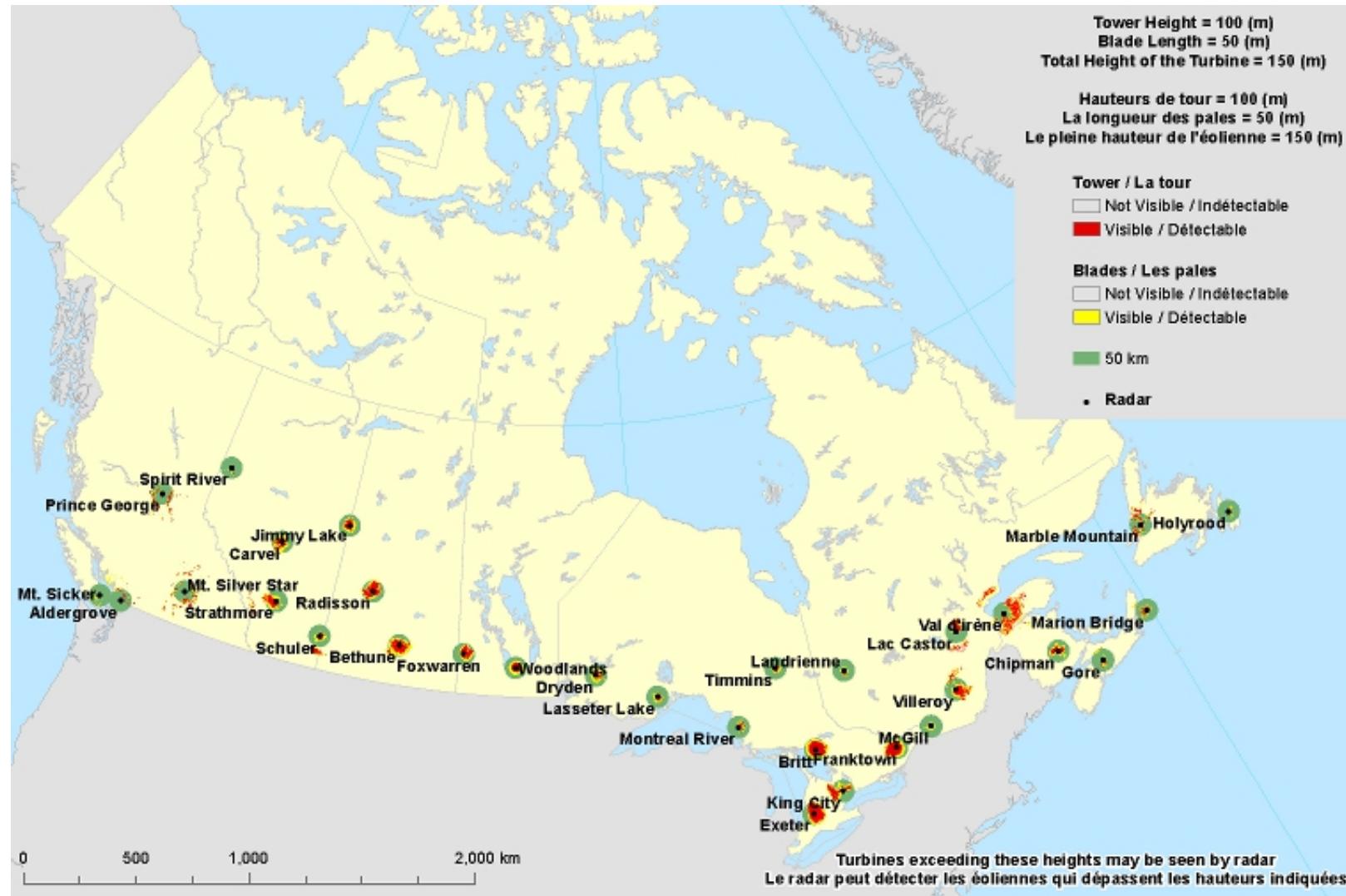
La puissance reçue est reliée au facteur de réflectivité d'un volume de particules.

Le facteur de réflectivité est dépendant de la distribution de la taille des gouttes.

Dans sa version la plus simple, la relation entre facteur de réflectivité et précipitation est donnée par la relation **Z-R**.

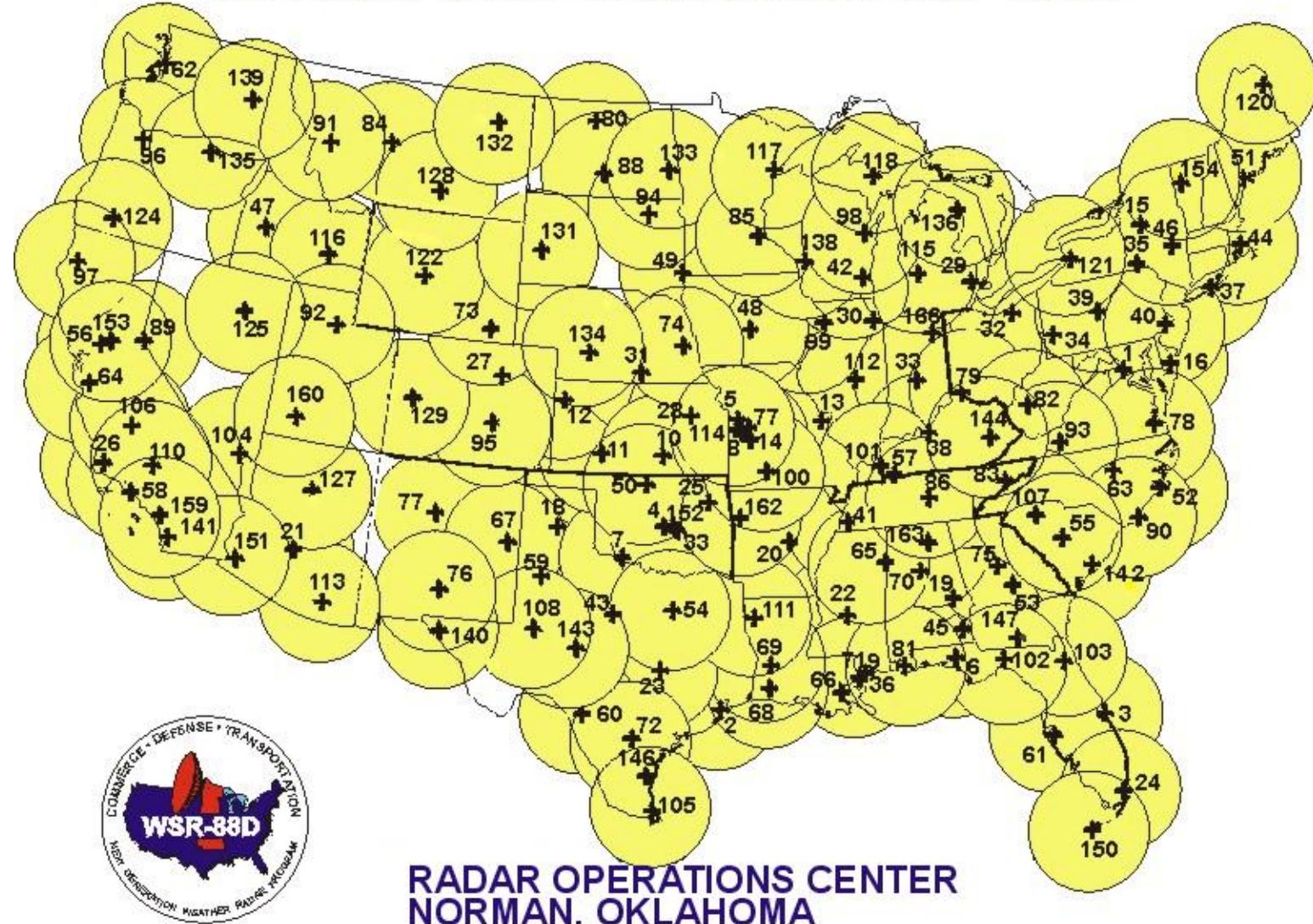
- $Z = 200R^{1.6}$ pour la pluie (Z en mm^6/mm^3 , R en mm/h)
- $Z = 200R^2$ pour la neige

Installations de radar canadiens



COMPLETED WSR-88D INSTALLATIONS WITHIN THE CONTIGUOUS U.S.

Installations de radar américains



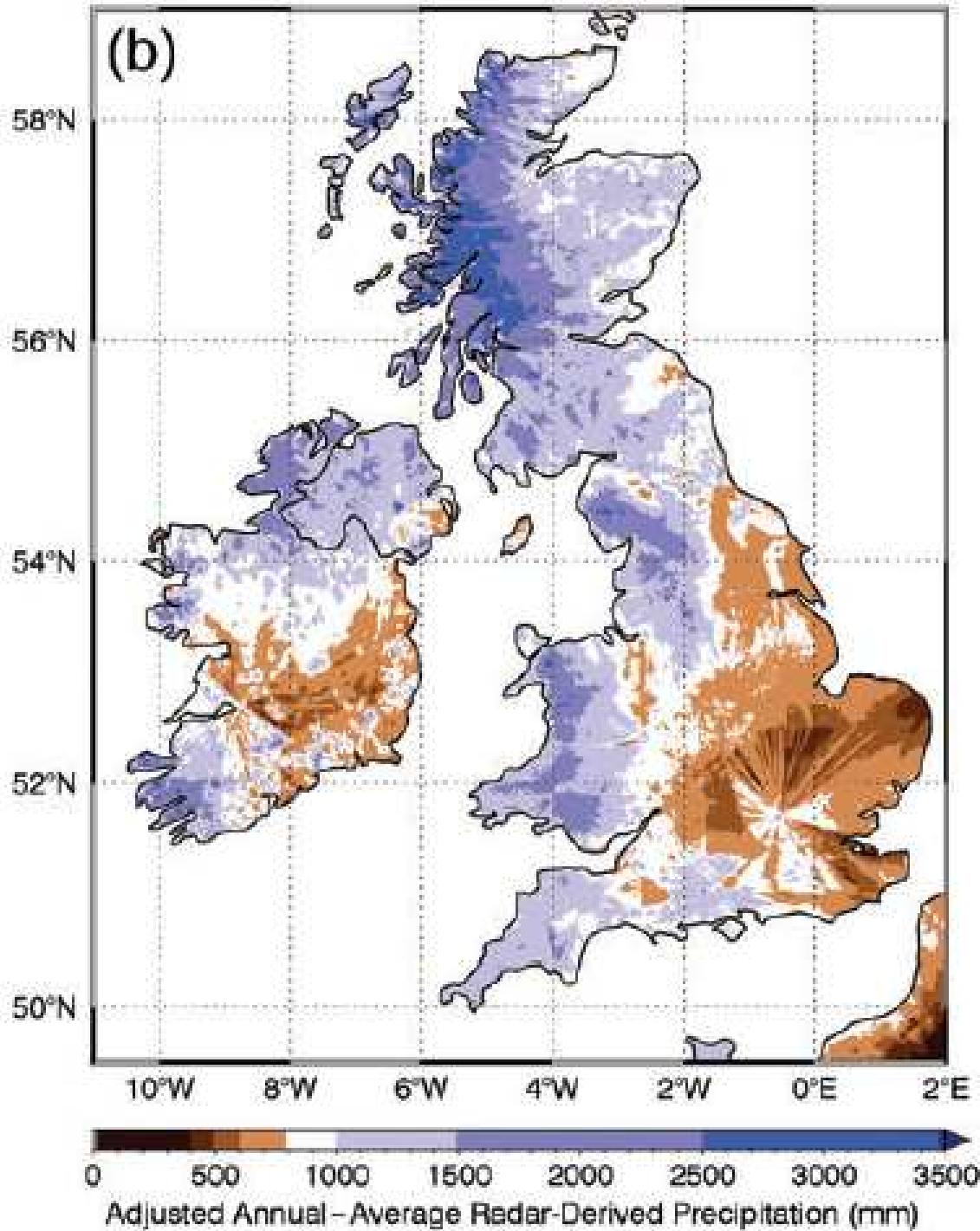
Données de radar

Résolution temporelle de 5 à 10 min.

Résolution spatiale de 1 degré en azimuth et en élévation, ~100 m sur l'axe de l'onde émise.

Une image de 200 km → 3 millions points de grille

Plusieurs téraoctets (To) de données produites par jour



Radar composé / « Composite Radar »

Source : Fairman et al. (2015) A Radar-based rainfall climatology of Great Britain and Ireland
(<https://doi.org/10.1002/wea.2486>)

Radar et prévision immédiat

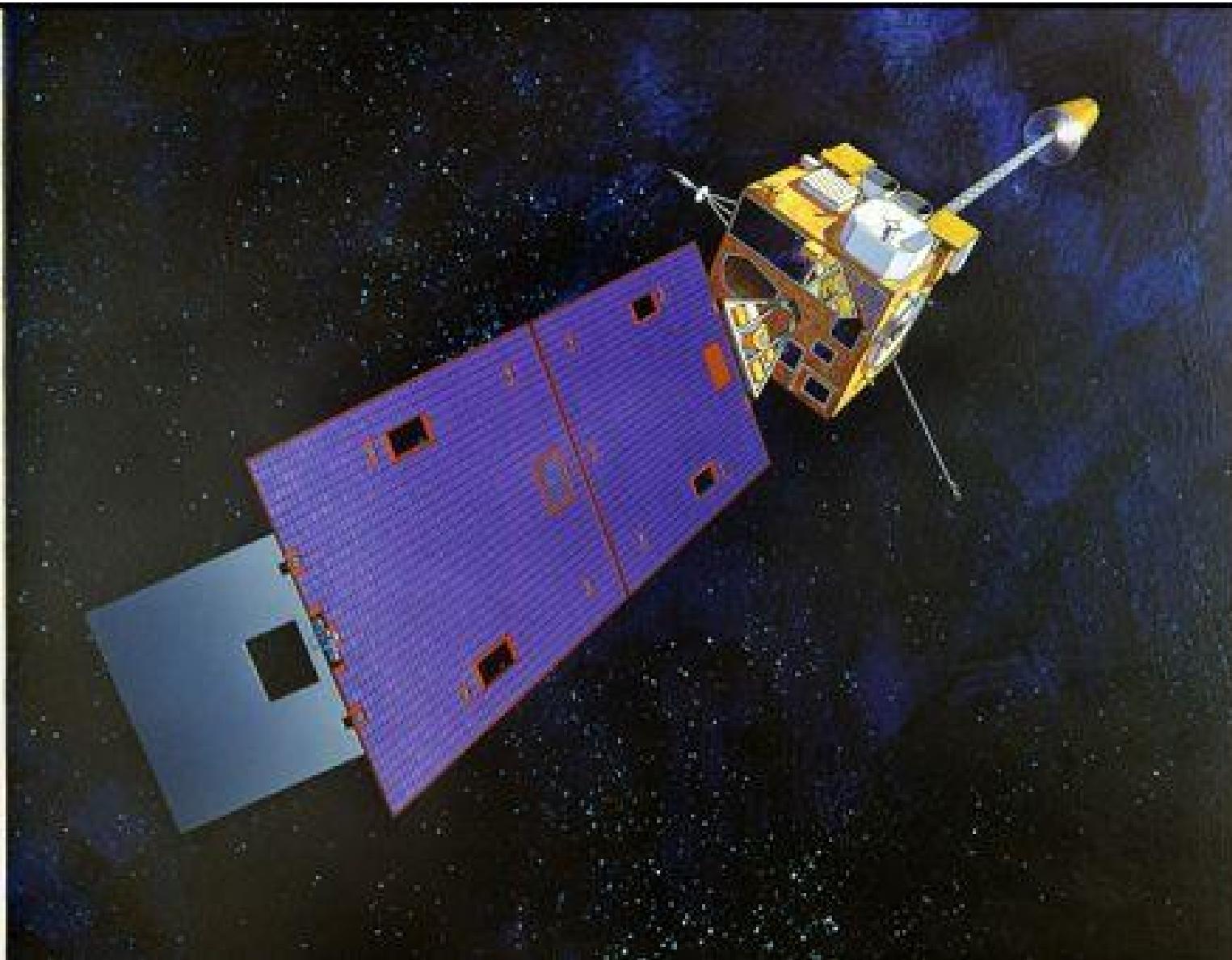
« Nowcasting »



Photo : Université McGill

Source : McGill Radar Nowcasting ([site web](#))



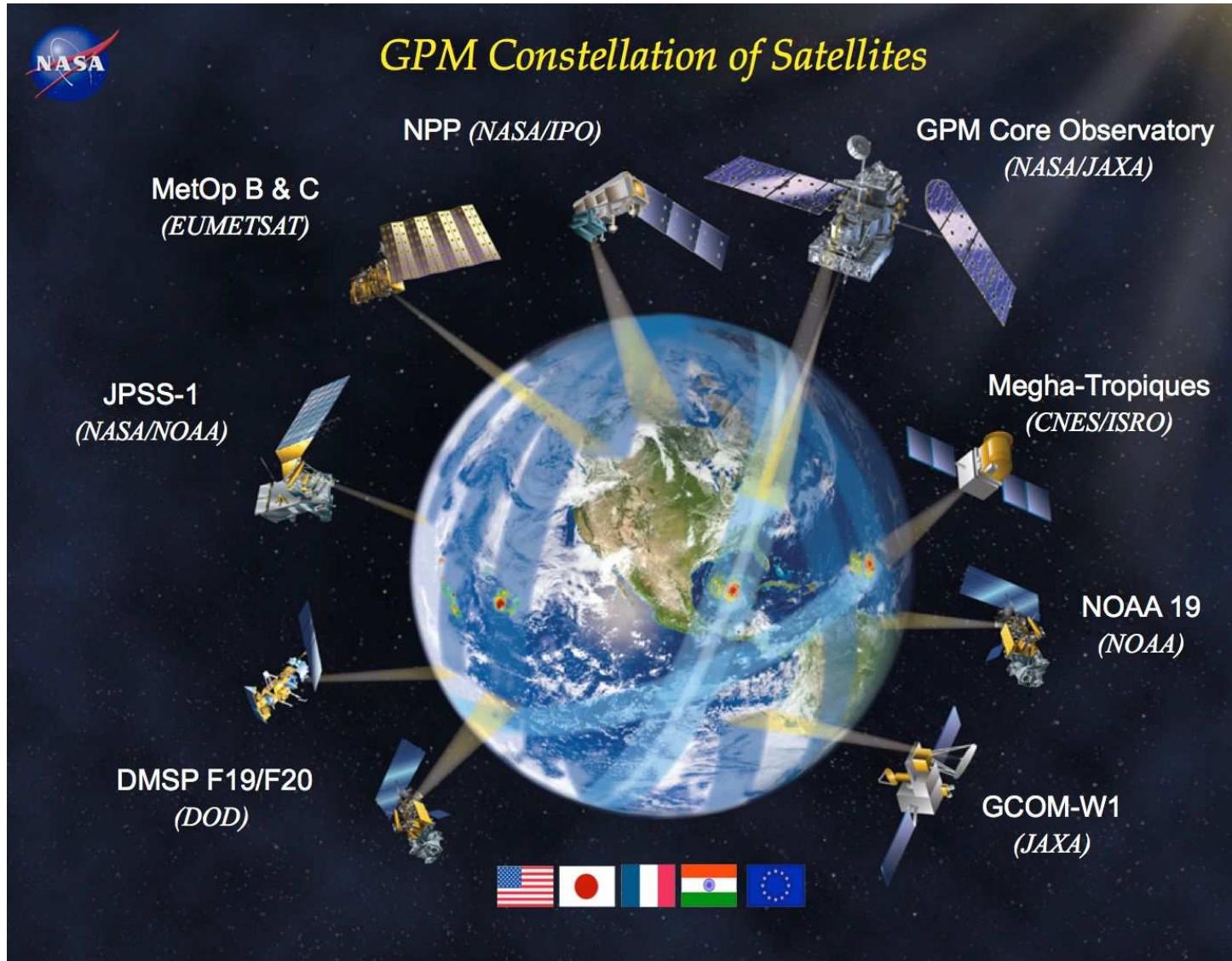


Satellites

Source : NASA

Données satellitaires

- Coordination à l'international
- Plusieurs « missions » en opération en même temps



Types des satellites

Également de la famille d'observation par « **télédétection** »

Deux types d'orbites : **géostationnaires** et **à orbite polaire**

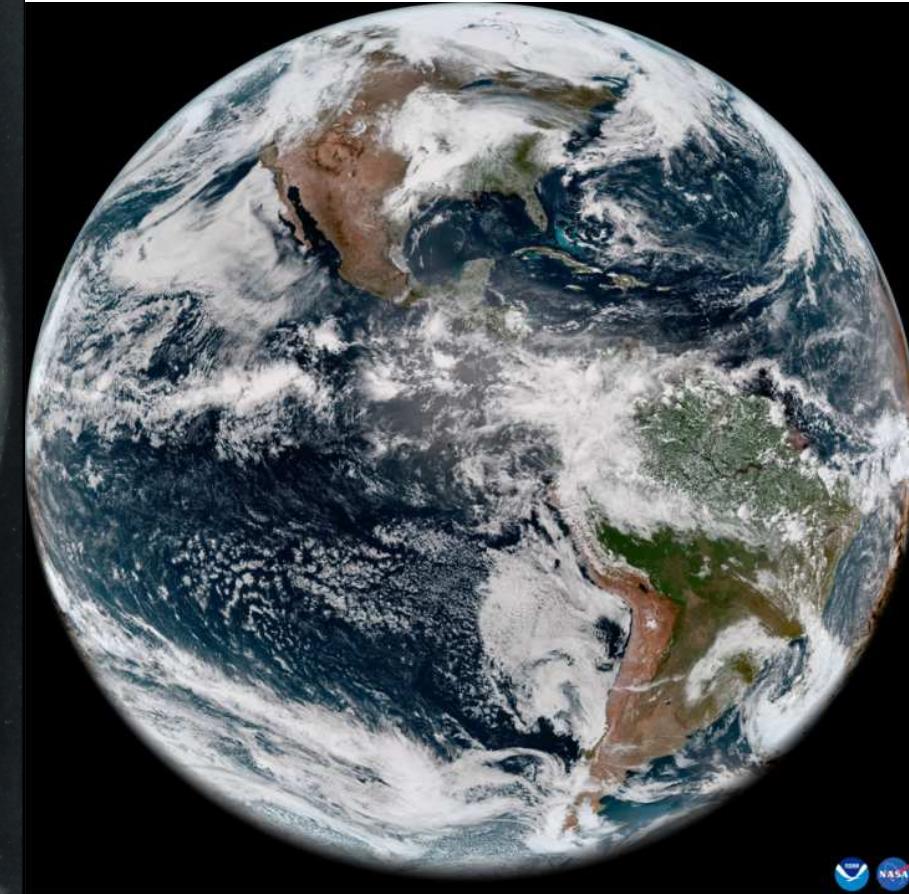
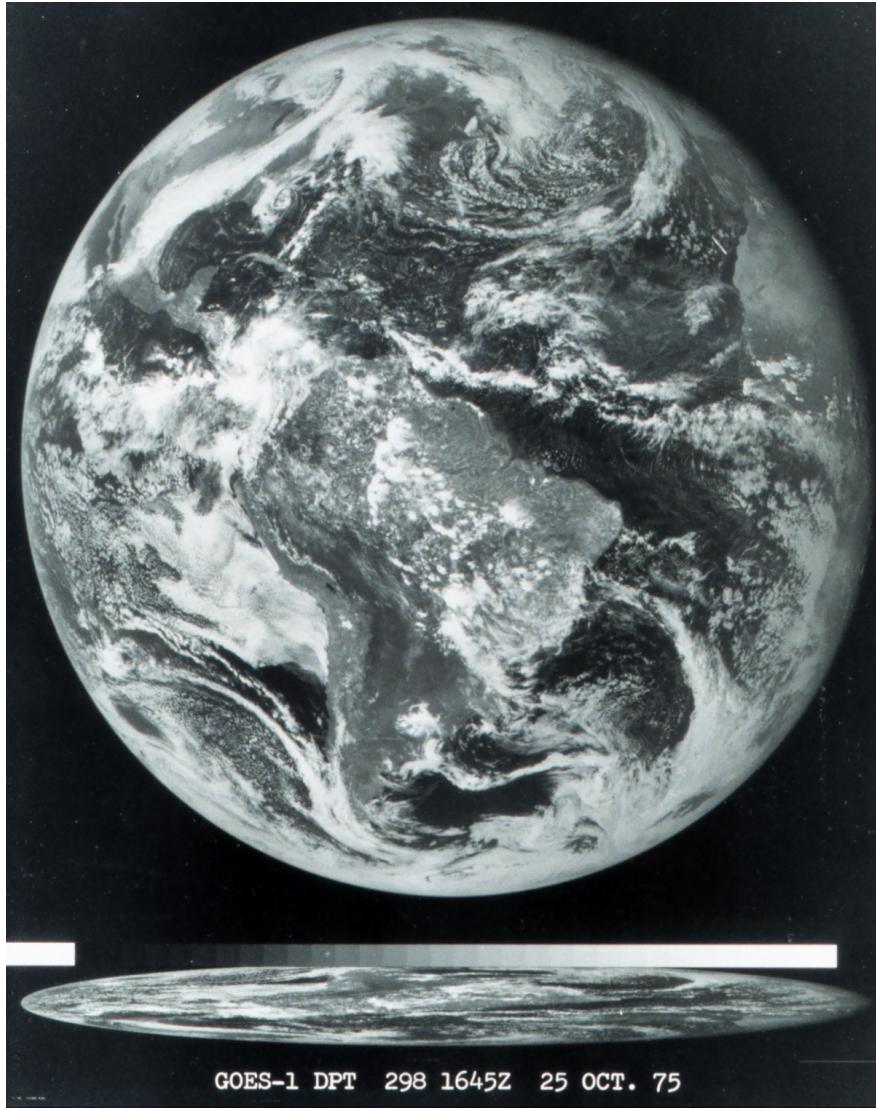
- Géostationnaire :
 - *36,000 Km au dessus de la terre* fixé au dessus du même point terrestre en permanence
- À orbite polaire :
 - *~850 Km au dessus de la terre*
 - *~14 orbites par jour*

Mission Satellite de NASA

« GOES »

GOES-1 (1975)

GOES-18 (2022)



Crédits des images : NASA, NOAA,
JMA

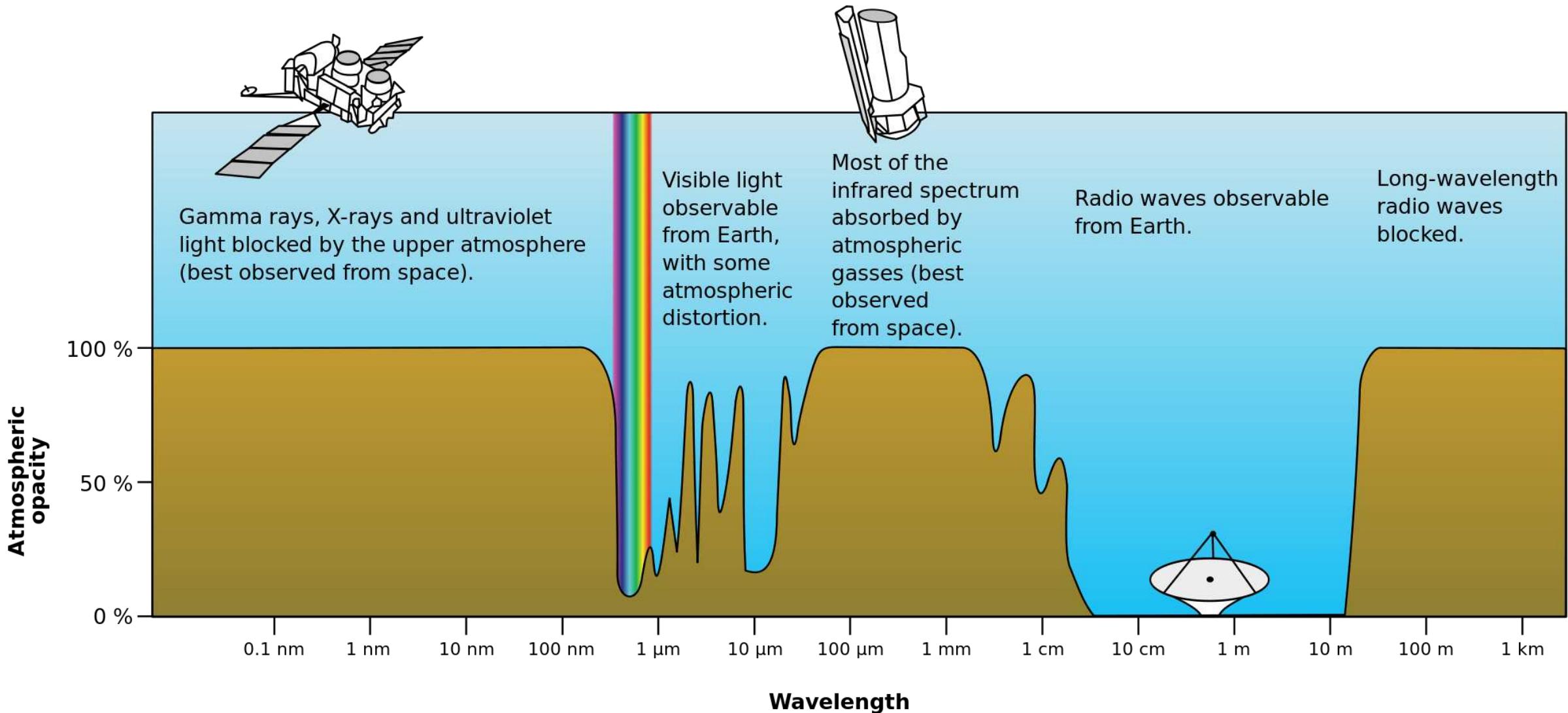
Satellites - Informations

Équipé d'un capteur photographique multispectres et sondeur,

Information obtenue sur :

- Hauteur et épaisseur de nuages,
- Température de surface de l'océan,
- Étendue de la neige et glace de mer,
- Hauteur de la surface de l'océan,
- etc...

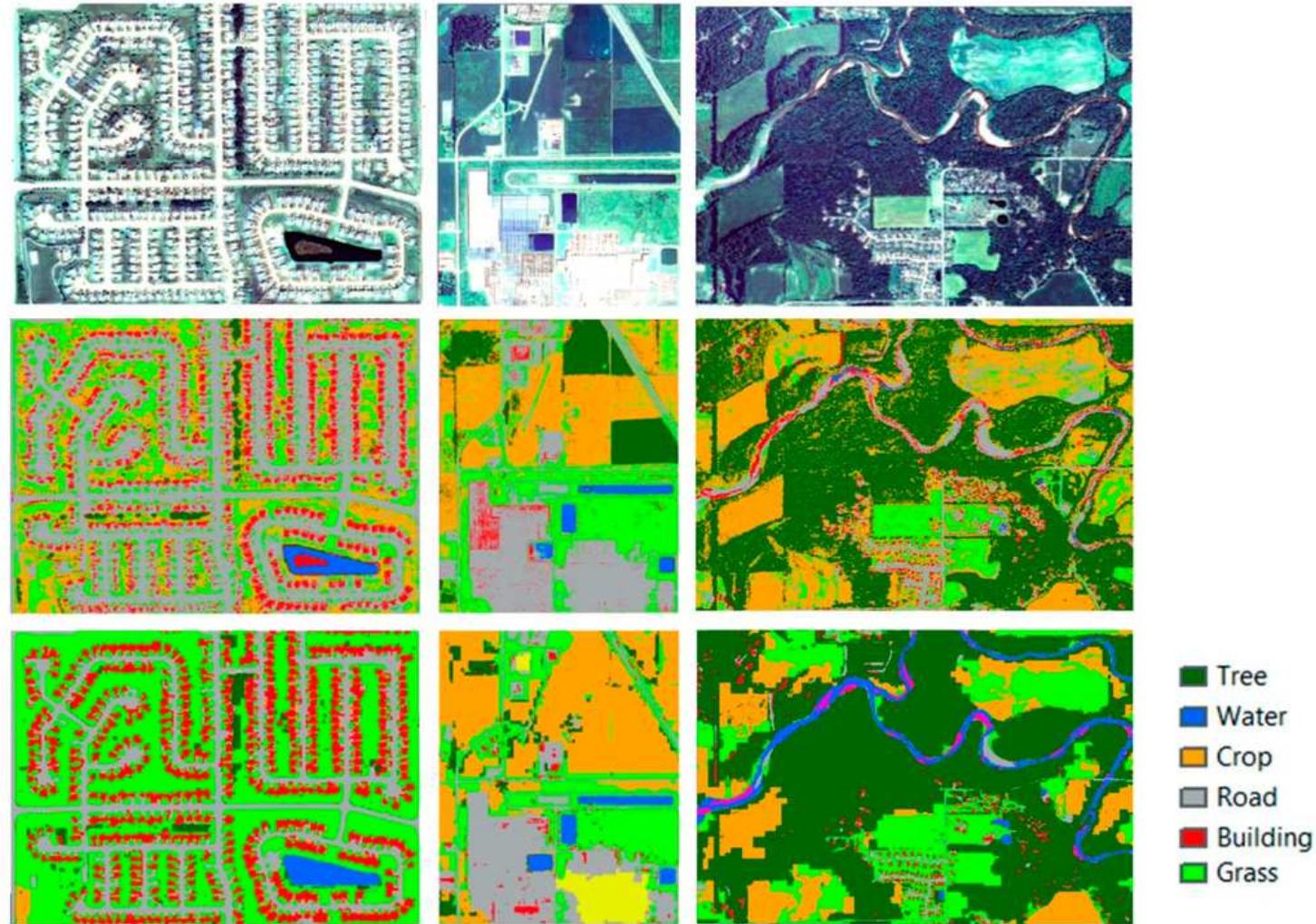
Satellites - Aspects optiques

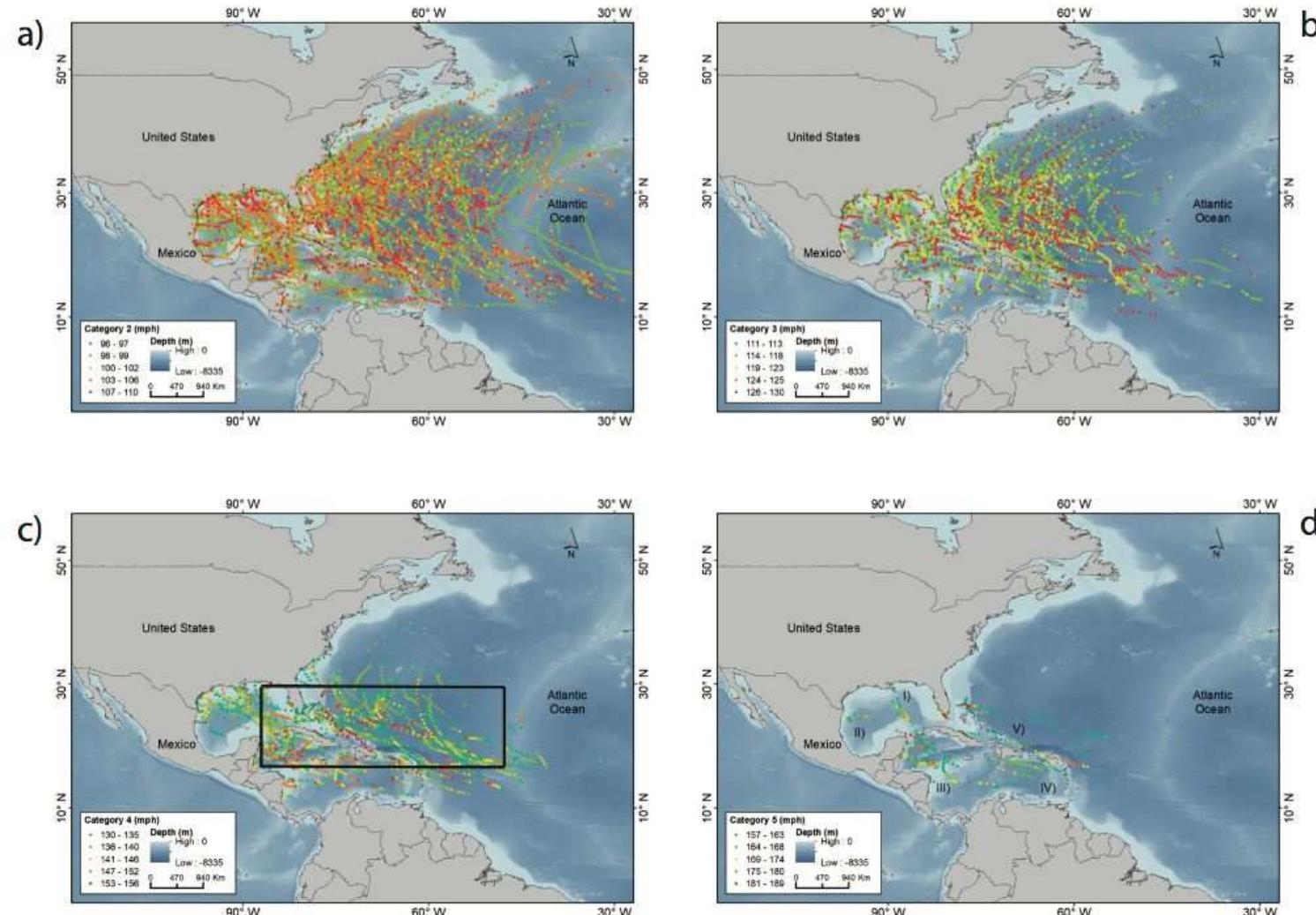


Satellites

Imagerie aérienne

Source : Li (2014): Object-based land-cover mapping with high resolution aerial Photography at the county scale in midwestern USA
(<https://doi.org/10.3390/rs61111372>)





Satellites

Les ouragans et apprentissage machine (« Machine Learning »)

Source : Herrera et al. (2022) Predicting Atlantic Hurricanes Using Machine Learning
(<https://doi.org/10.3390/atmos13050707>)

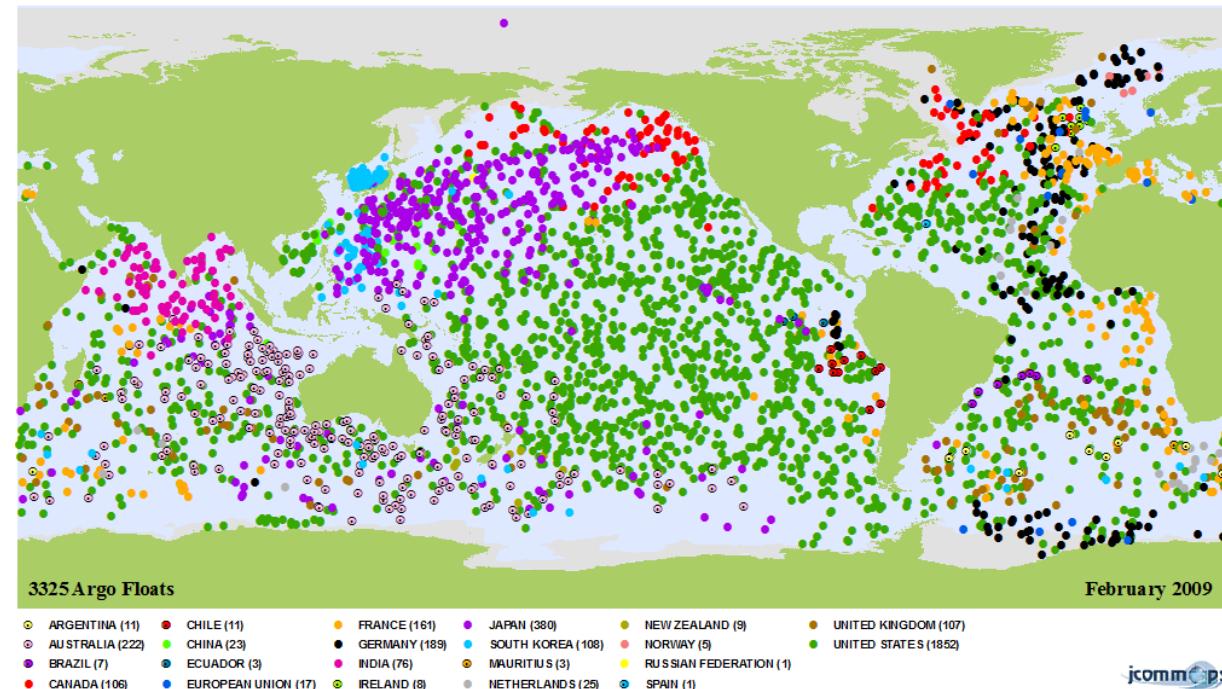
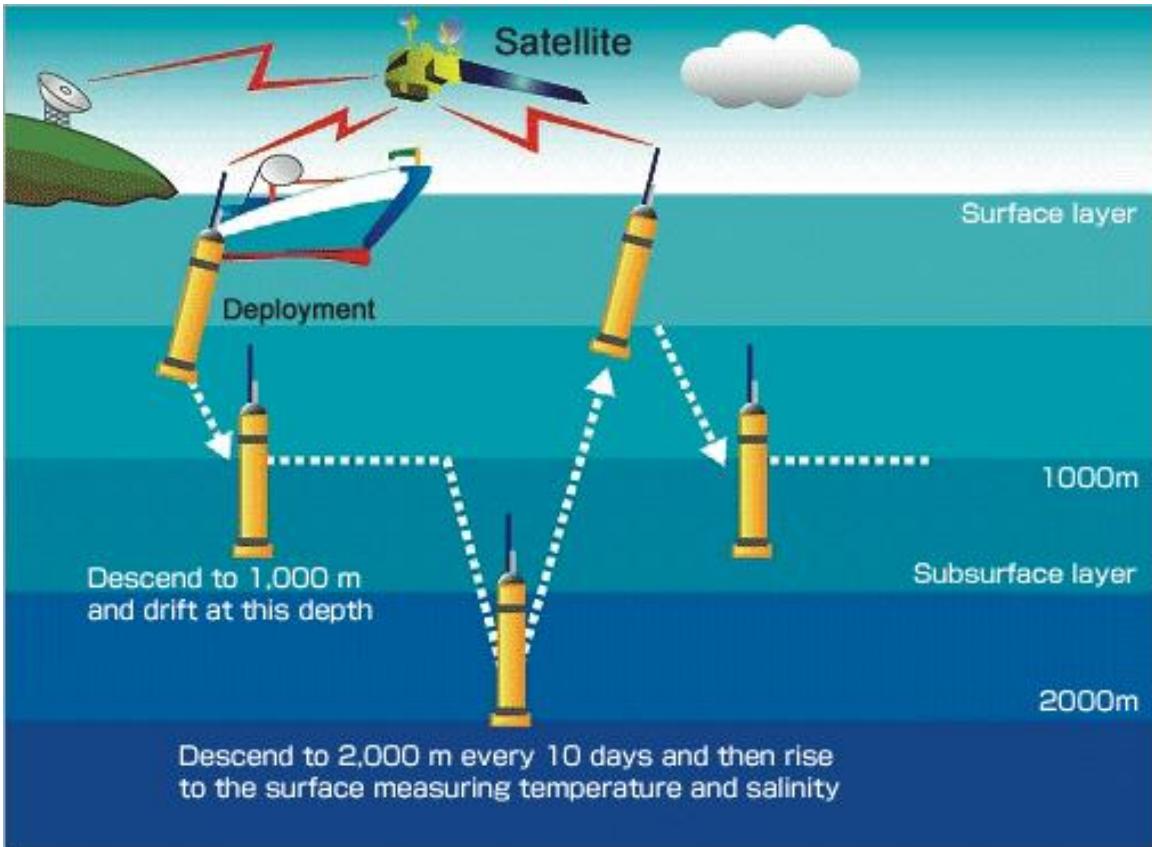
Satellites en opération

Couverture globale de satellite géostationnaire :

- **GMS Himawari 8** (japonais : **JMA**)
 - l'est asiatique et l'océan pacifique.
- **GOES West et GOES East** (américains : **NASA**)
 - l'océan pacifique, les Amériques, l'océan atlantique.
- **METEOSAT** (européen : **EUMETSAT**)
 - Europe et Afrique.
- **INSAT** (indien : **ISRO**)
 - Asie et l'océan Indien.
- Les satellites en orbites polaires ont des missions plus spécialisées
 - e.g. **LANDSAT**, **SRTM**, **Terra**, **Aqua**, etc.

Données des courants / océans

Bouée Argos



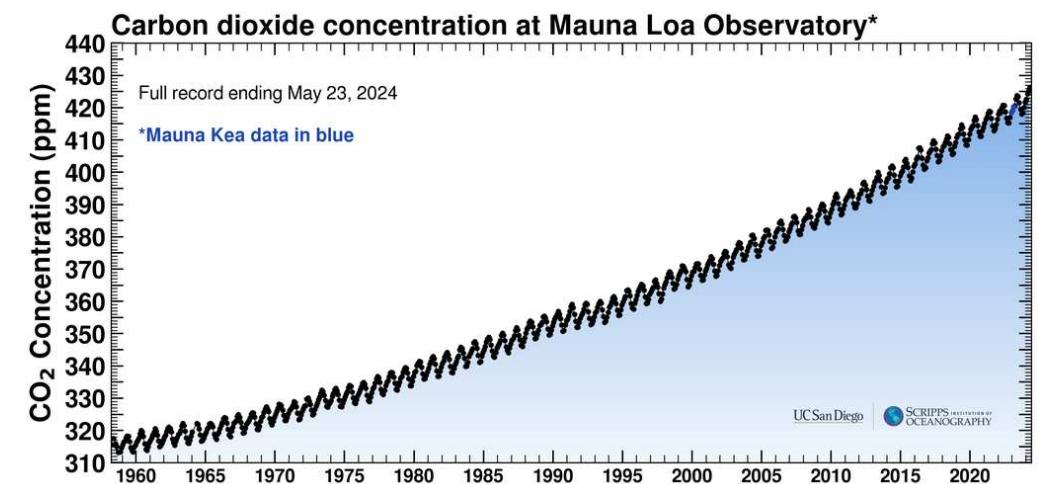
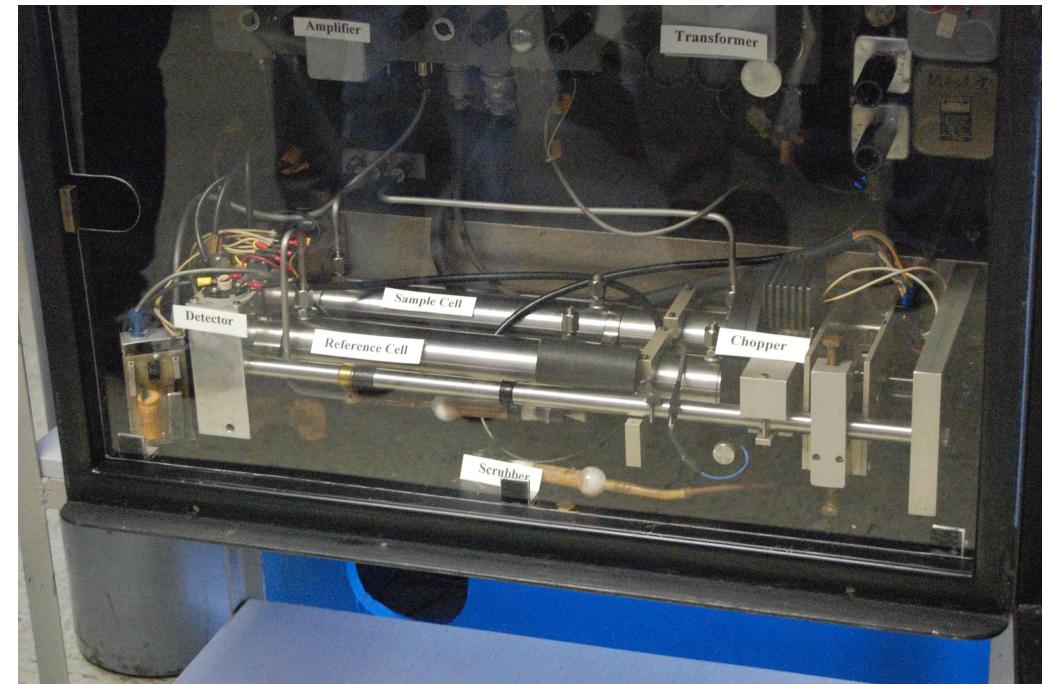
Données des rivières

Stations de jaugeage



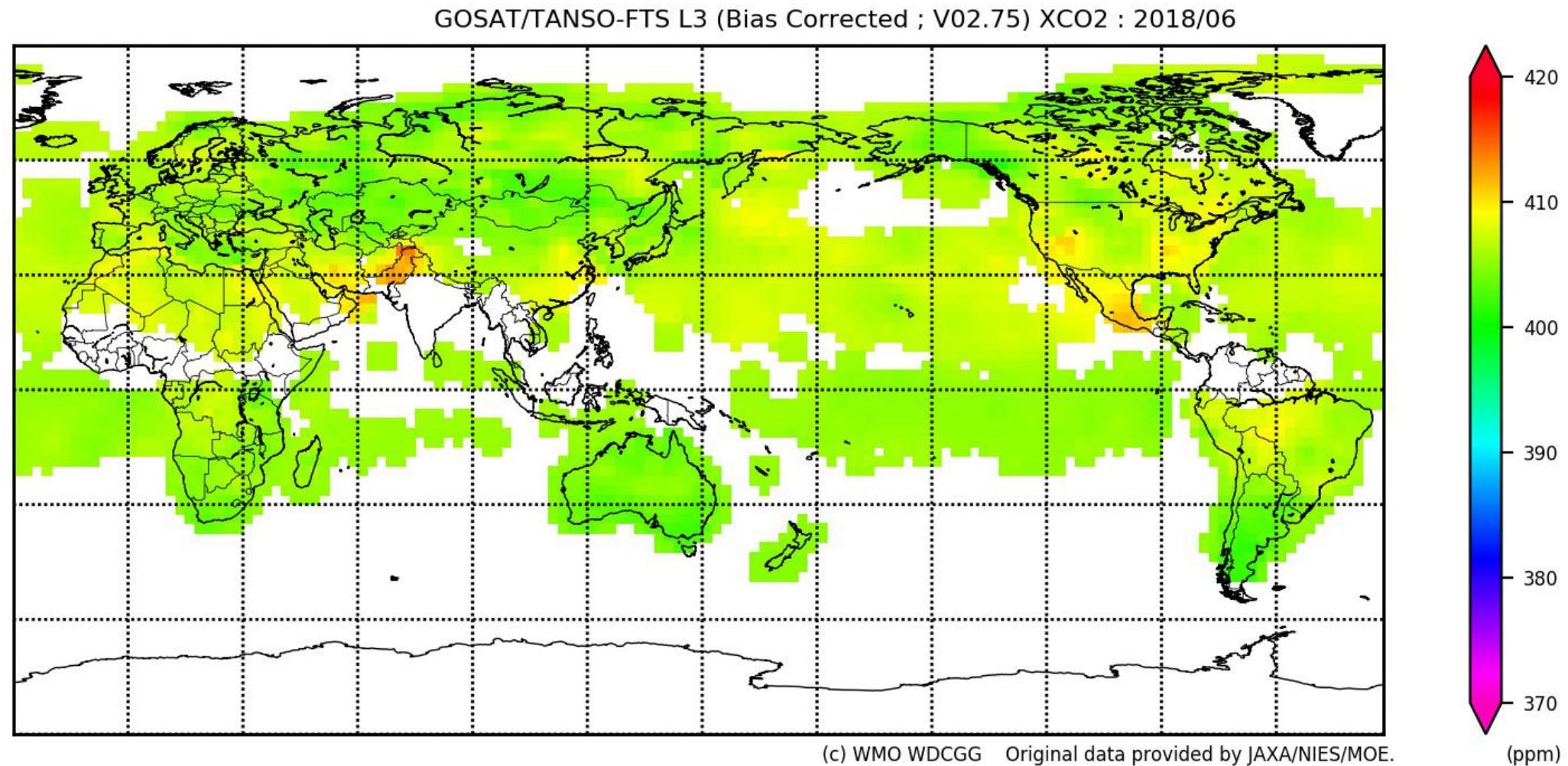
Mesures de CO₂ in situ

Spectrophotomètre - Courbe de Keeling



Source : Scripps Institution of Oceanography (UCS) (<https://keelingcurve.ucsd.edu/>)

Mesures de CO₂ par satellite



3. Données observées

Données de stations

École d'été en sciences du climat 2024

REPORT OF METEOROLOGICAL OBSERVATIONS taken at												for	
Date Year.	Time	Local Time	Barometric				Wind or Wind A.		Wind Data				Humid. Pressure
			Station	Altitude	Corrected for Alt.	Reduced to Sea Level	Reduced to Standard	Station	Corrected	Reduced to Standard	Wind	Wind	
Latitude	142	29.912	57	29.995	29.995	29.995	29.995	49.8	24.6	5.0	100	50	
	4	29.966	52	29.979	29.979	29.979	29.979	87.4	42.5	4.9	120	50	
	7	29.965	53	29.978	29.978	29.978	29.978	88.3	41.0	5.0	135	54	
	10	29.960	54	29.913	29.913	29.913	29.913	88.7	40.0	8.7	150	41	
	28	29.911	55	29.844	29.773	29.971	29.971	89.0	41.0	9.0	180	43	
	4	29.772	57	29.785	29.777	29.776	29.776	57.5	48.2	8.3	172	45	
	7	29.751	53	29.711	29.706	29.706	29.706	48.7	42.7	5.8	200	61	
	10	29.745	53	29.715	29.693	29.713	29.713	47.0	44.7	5.0	206	52	
	213	29.744	53	29.717	29.692	29.712	29.712	47.0	44.7	3.3	206	52	

Données de stations canadiens

Les archives débutent vers les années 1840

Les données de stations d'Environnement Canada sont disponibles en tableaux mensuels :

- <https://climat.meteo.gc.ca>

Pour un usage plus comprehensif, le Service météorologique du Canada (SMC) offre quelques manières d'accéder aux données observées gratuitement

- <https://eccc-msc.github.io/open-data/>

Sommaires de données de stations canadiens

Source: Environnement Canada

Daily Climatological Data (DLY02, DLY04) - (Note 16)

Element number, units, code, description, flags and notes for DLY02 and DLY04 weather elements

Element	Units	Code	Description	Flags	Note
001	0.1 deg C		Daily Maximum Temperature	E	
002	0.1 deg C		Daily Minimum Temperature	E, N, Y	
003	0.1 deg C		Daily Mean Temperature	E	
004	%		Daily Max. Relative Humidity		
005	%		Daily Min. Relative Humidity		
023	10's of deg (See 157)		Direction of Extreme Gust (16 pts) to Dec. 1976	S, E	5, 34
024	km/h		Speed of Extreme Gust	S, E	
025			UTC Hour of Extreme Gust (Earliest)		41
157	10's of deg (see 023)		Direction of Extreme Gust (36 pts) from Jan. 1977	S, E	

006-009 6 hrly. Precipitation ending

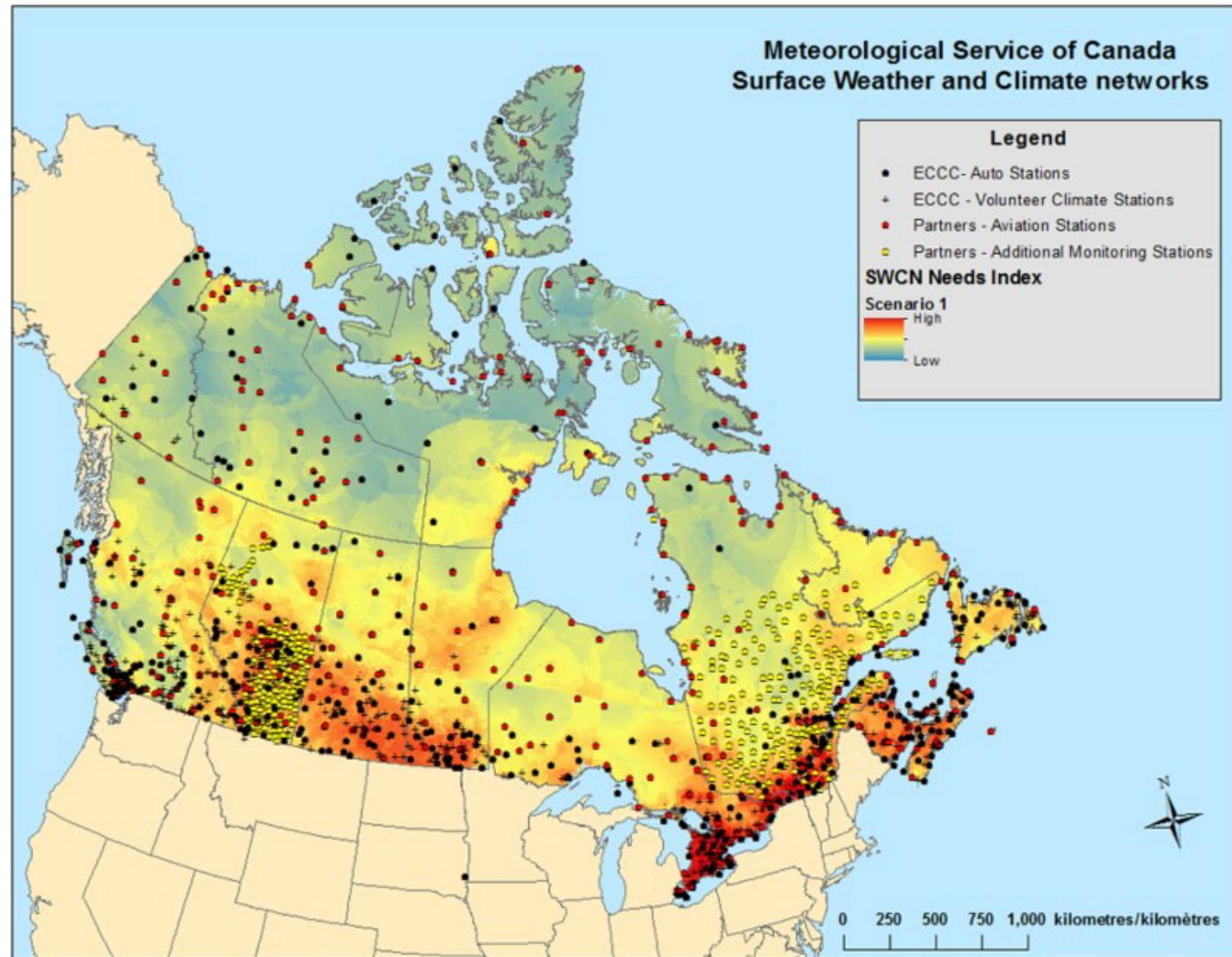
006	0.1mm		1200 UTC	I	15
007	0.1mm		1800 UTC	I	15
008	0.1mm		0000 UTC	I	15
009	0.1mm		0600 UTC	I	15
010	0.1mm		Total Rainfall	E,I,C,L,A,E	
011	0.1cm		Total Snowfall	E,I,C,L,A,E	37
012	0.1mm		Total Precipitation	E,I,C,L,A,E	37
013	cm		Snow on the Ground	E, I	8, 37

014-022 Day with... (DLY04 - also available for DLY44, elements 014-017)

014		1=Yes, 0=No	Thunderstorms		14
015		1=Yes, 0=No	Freezing Rain or Drizzle		14
016		1=Yes, 0=No	Hail		14
017		1=Yes, 0=No	Fog or Ice Fog		14

Données de stations canadiens

École d'été en sciences du climat 2024



Source : Mekis et al. (2016) An Overview of Surface-Based Precipitation Observations at Environment and Climate Change Canada (<https://doi.org/10.1080/07055900.2018.1433627>)

Données de stations du MELCCFP

- *Les ministres changent leurs noms/acronyms après tous les élections*
 - c'est ennuyeux !

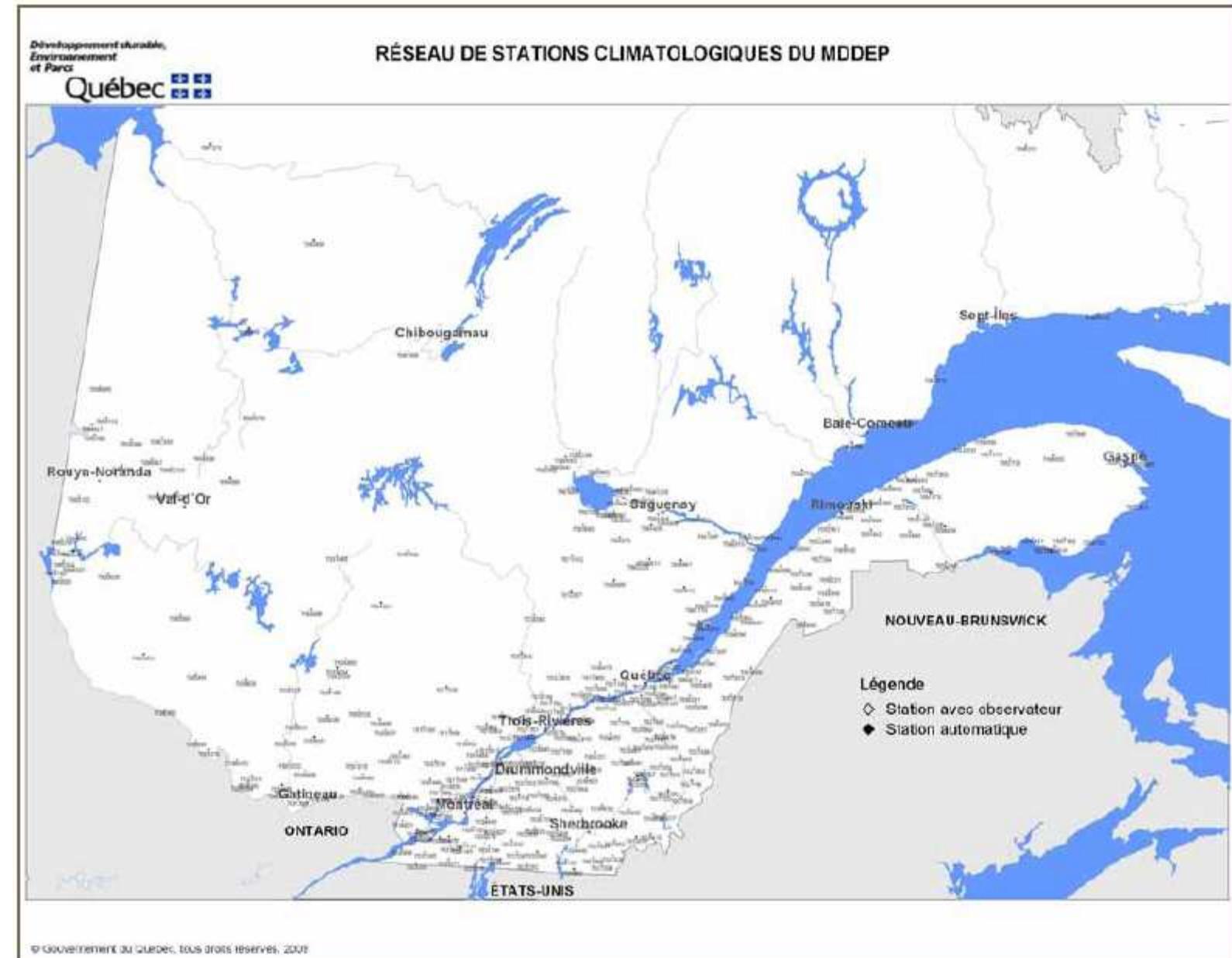
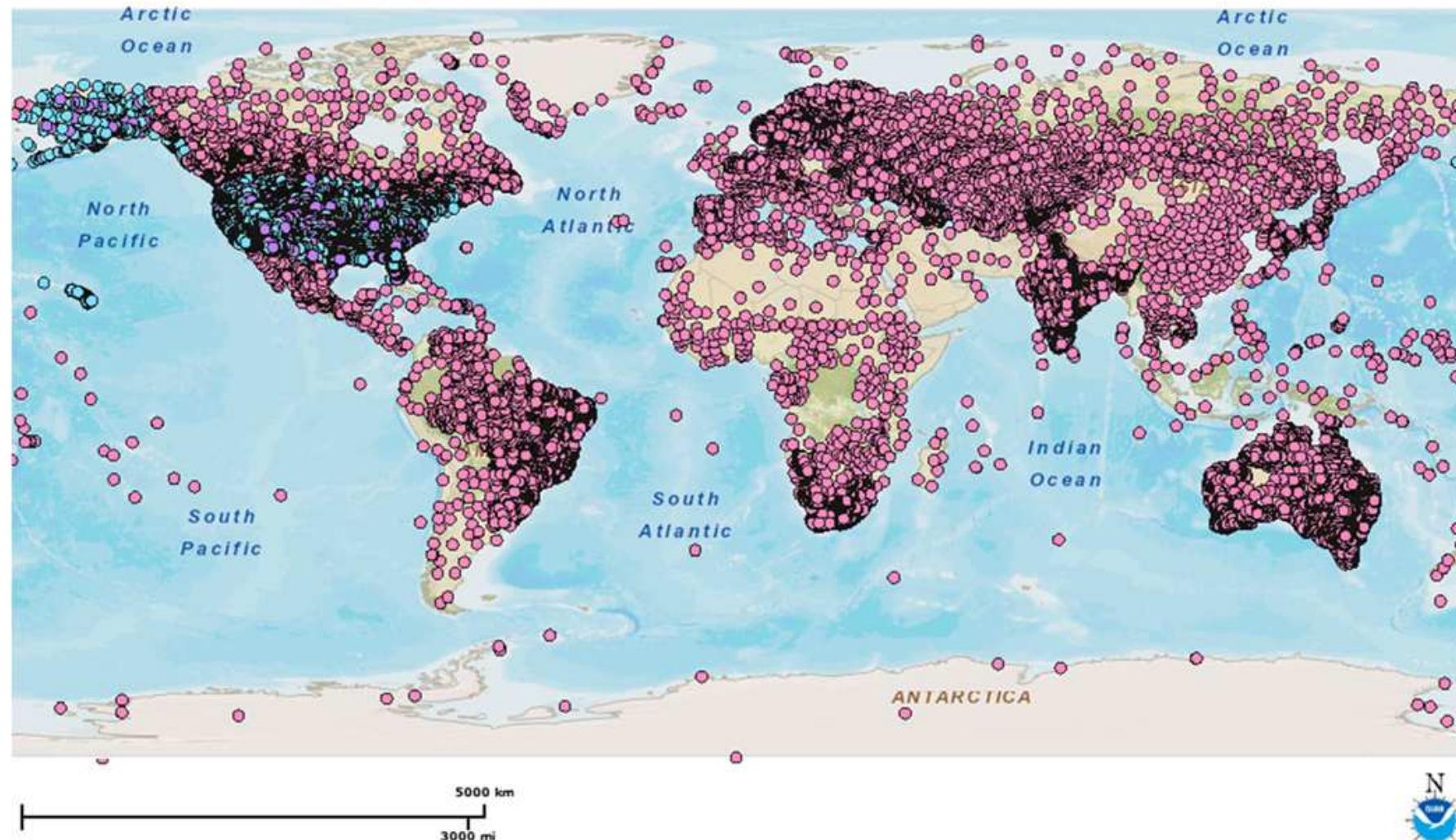


FIGURE 1. RÉSEAU DE STATIONS MÉTÉOROLOGIQUES DU MDDEP

Source : Ministère du Développement durable, de l'Environnement et des Parcs (MDDEP). Surveillance du climat, réseau et paramètres, sur le site du MDDEP <http://www.mddep.gouv.qc.ca/climat/surveillance/index.asp> (consulté le 12 août 2010)

Données de stations globales

École d'été en sciences du climat 2024



Menne et al. (2012): Global Historical Climatology Network - Daily (GHCN-Daily), Version 3
(<https://doi.org/10.7289/V5D21VHZ>)

Données de stations - Conclusions

Forces

- Information locale à partir d'instruments
- Représentatif de l'échelle à laquelle l'humain perçoit la météo
- Records les plus anciens des valeurs « vérifiées sur le terrain » (« *ground-truthed* »)

Faiblesses

- Possibilité d'erreurs par les observateurs
- Peut être affecté par la proximité de perturbations (chaleur urbaine, masse d'eau)
- Déplacement et fermeture de stations
- Discontinuités temporelles
- Couverture spatiale très hétérogène

Données homogénéisées

Ajustement des données brutes de stations à l'aide de connaissances sur les erreurs des différents instruments, le déplacement de stations, etc.

Utilisation de données de stations avoisinantes pour combler les données manquantes.

Détection et correction des sauts dans les séries à l'aide de méthodes statistiques

- *régression linéaire, test de Student, test de Mann-Kendall, etc.*

Données homogénéisées

Environnement Canada est le principal pourvoyeur de ce type de données au pays via les travaux des chercheurs en climat et météo

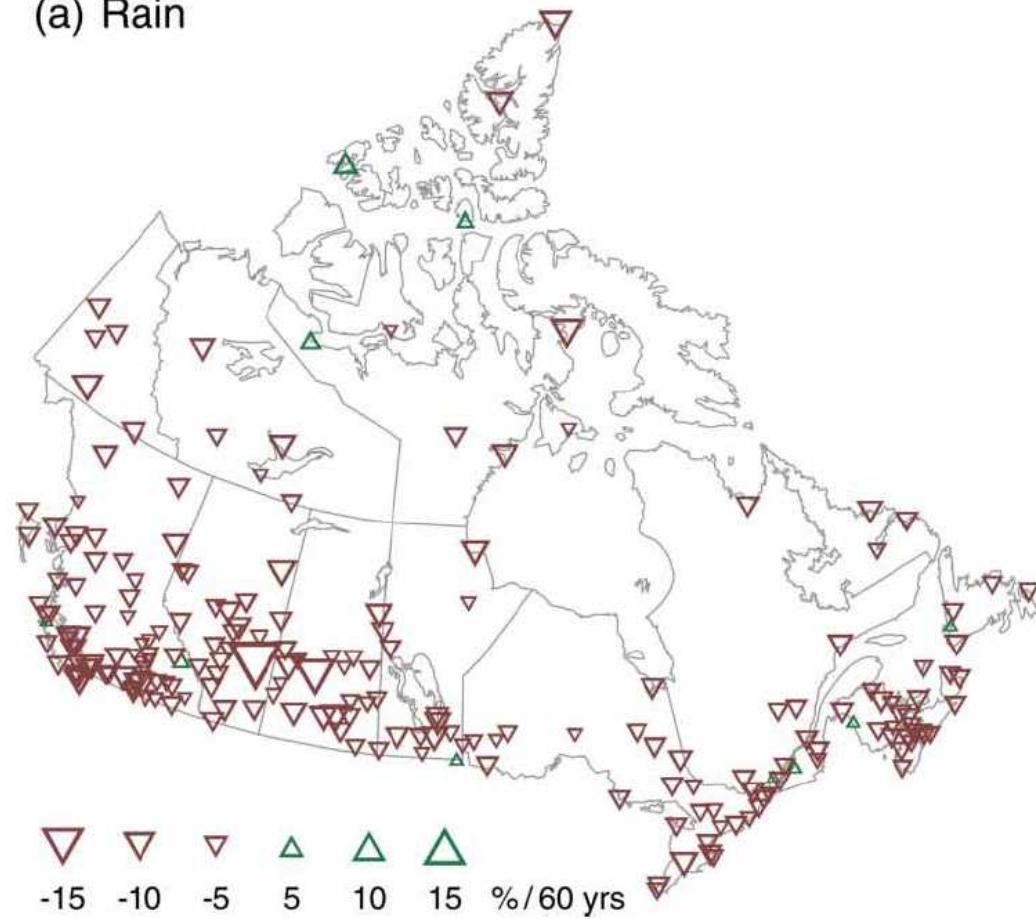
- <http://www.ec.gc.ca/dccha-ahcccd/>

Plus de 400 stations homogénéisées pour les *températures minimum, maximum et la précipitation.*

- **Wan et al. 2007; Wan et al. 2010 (Génération 1)**
- **Mekis et Vincent 2011; Vincent et al. 2012. (Génération 2)**
- **Vincent et al. 2020 (Génération 3)**

Données homogénéisées

(a) Rain



(b) Snow

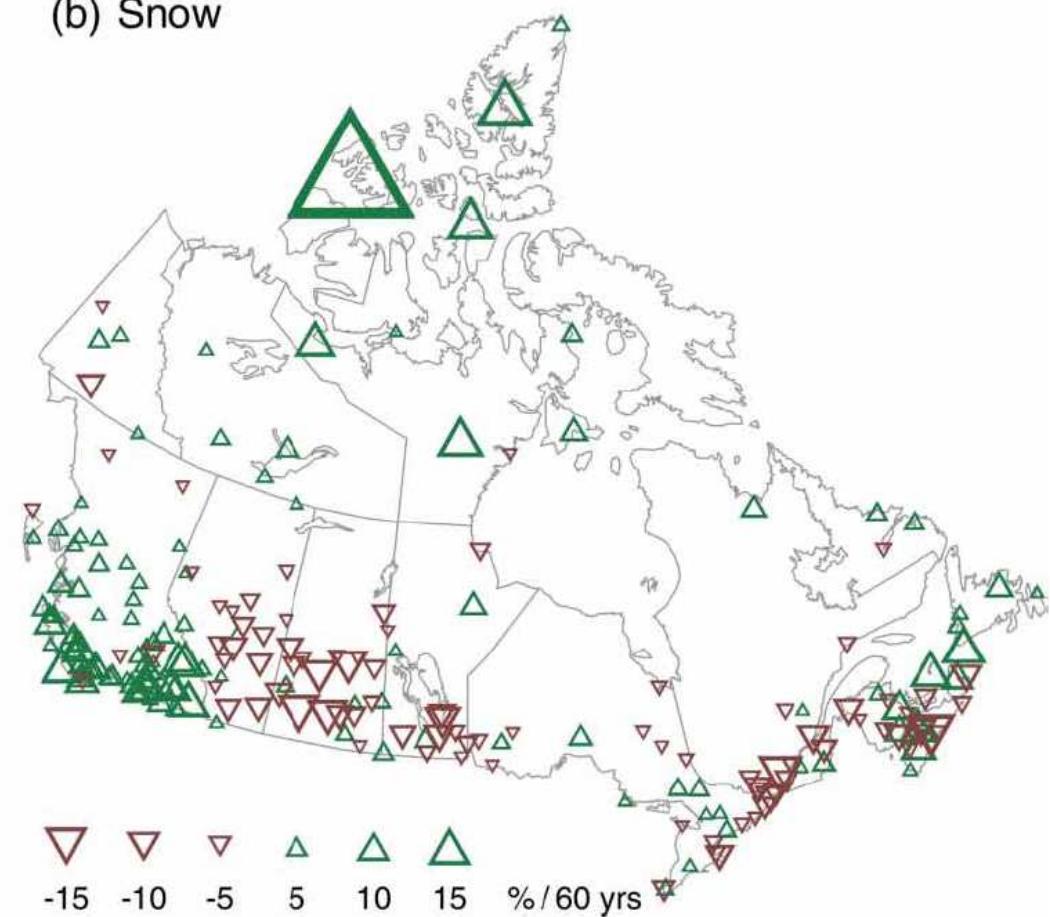


Fig. 6 Difference in the trends in annual total rainfall and snowfall (expressed as a percentage) before and after adjustments for 1950–2009. Panels a) and b) include rain-gauge, snow and trace adjustments. The size of the triangle is proportional to the magnitude of the change in trends.

Mekis & Vincent 2011. An Overview of the Second Generation Adjusted Daily Precipitation Dataset for Trend Analysis in Canada (<https://doi.org/10.1080/07055900.2011.583910>)

Données homogénéisées - Conclusions

Forces

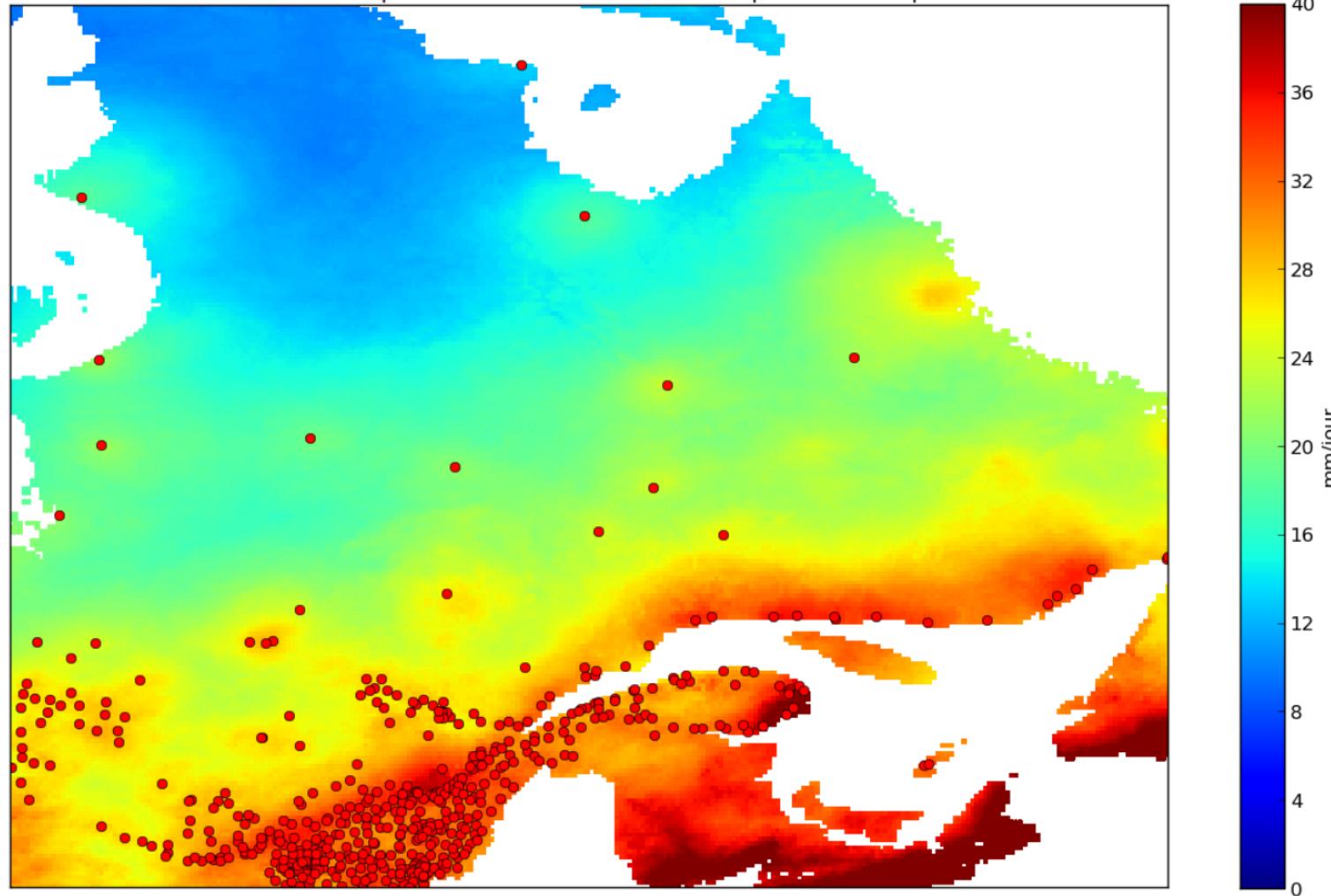
- Information locale
- Représentatif de l'échelle à laquelle l'humain perçoit la météo
- Séries (plus) continues
- Corrigé pour donner des tendances plus représentative de la réalité

Faiblesses

- Moins de stations disponibles
- Couverture spatiale très hétérogène
- Peu d'information sur l'incertitude des méthodes d'homogénéisation
- Seulement température et précipitation disponible

Données interpolées

Quantile 99.5 pour la précipitation journalière NRCAN (1971-2000)
et stations avec plus de 25% de couverture temporelle sur la période



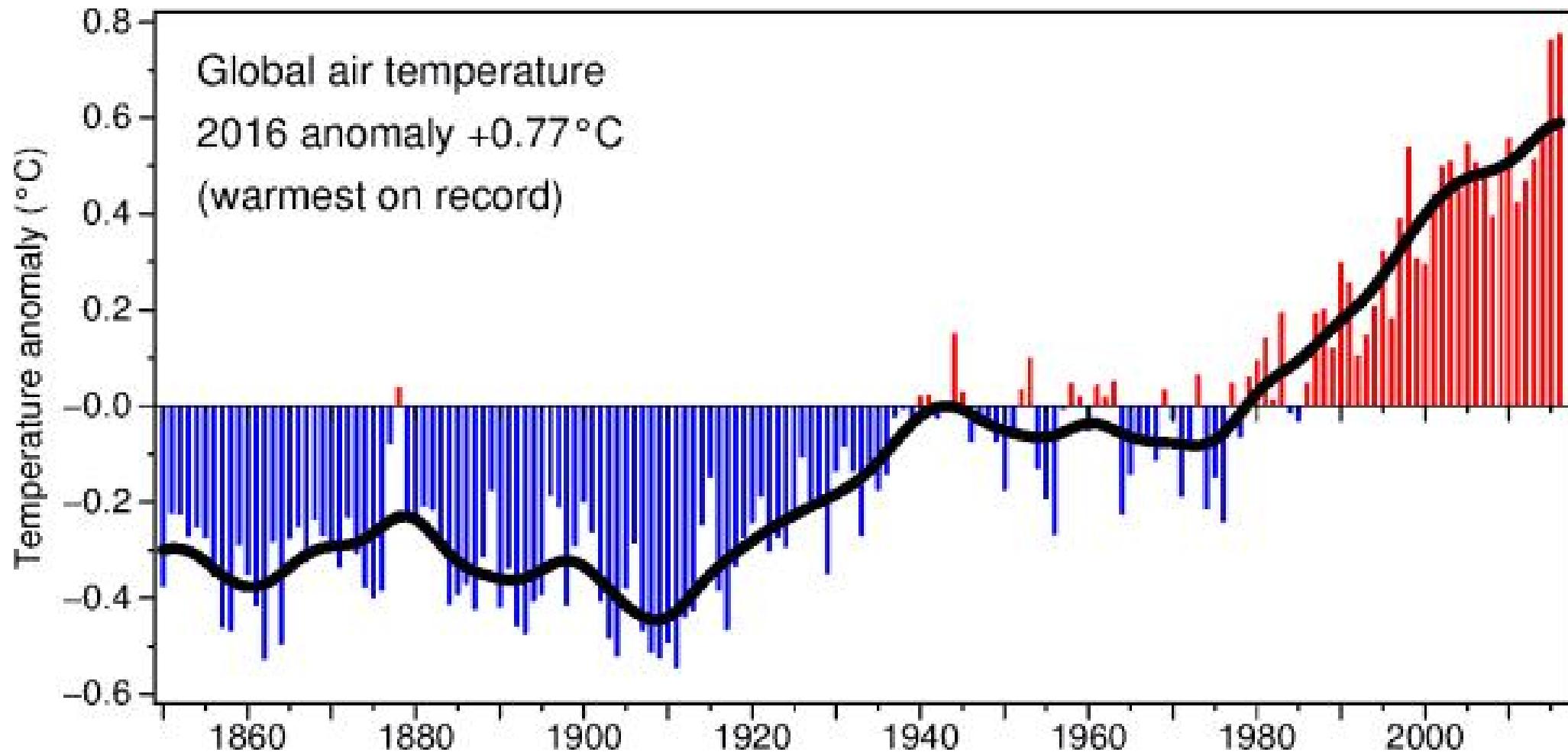
Données interpolées - Méthodes

Interpolation de données de stations (brutes ou homogénéisées) sur une grille.

Plusieurs méthodes possibles :

- **Bilinéaire**
- **Interpolation multivariée** (Gandin 1963)
- « **Inverse distance weighting** » (Shepard 1968)
- « **Thin-plate splines** » (Duchon 1976, Wahba 1979)
- **Krigage** (Matheron 1960), etc.

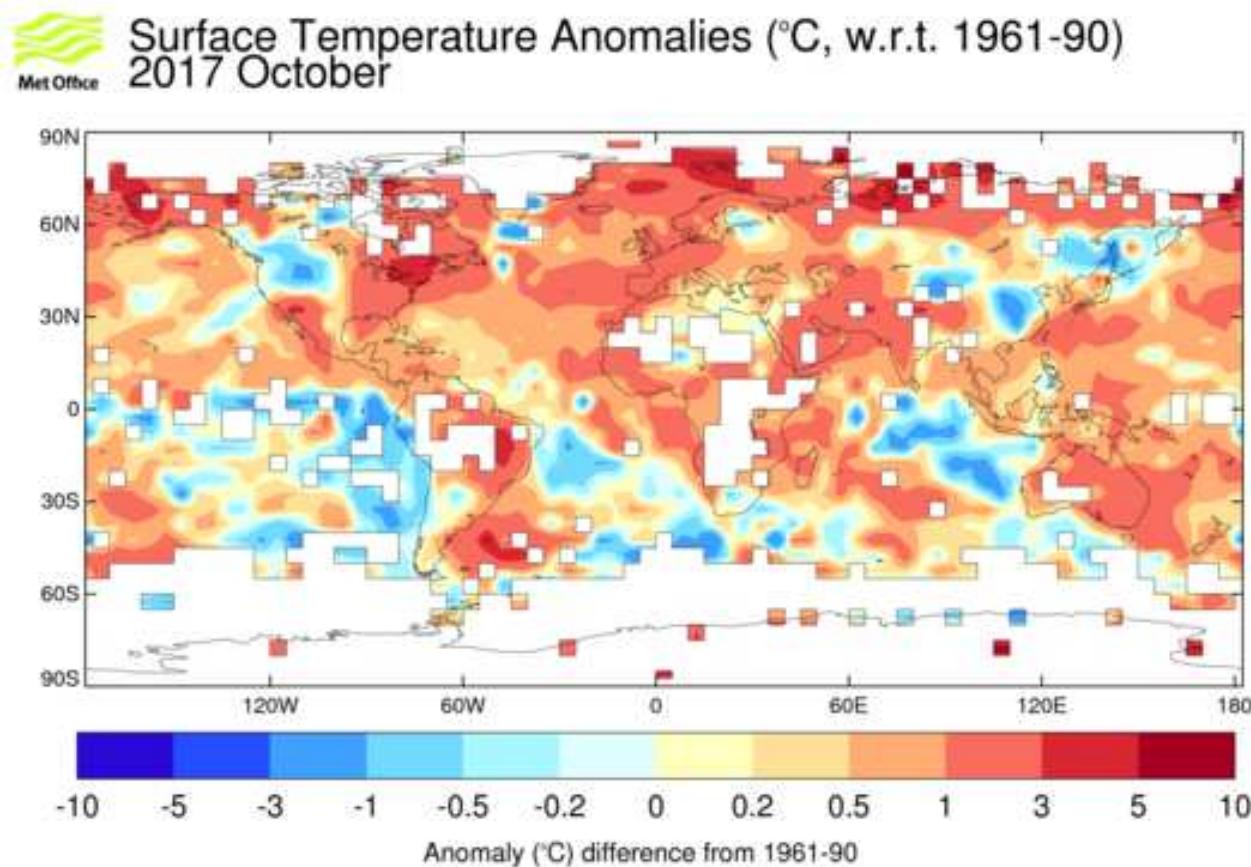
Données interpolées - Globales



Données interpolées - Globales

Le CRU produit plusieurs produits d'interpolations (**CRUTEM**, **HadCRUT**, **HadSST**...)

- Résolution temporelle mensuelle *à partir de 1850.*
- Résolution spatiale de ~500 km *sur le globe*
- Données direct - **HadCRUT4**



Source : Jones et al. (1999)

(<https://doi.org/10.1029/1999RG900002>)

Données interpolées - Canada

CanGRD est une interpolation des données homogénéisées d'Environnement Canada avec une interpolation optimale.

- Résolution temporelle mensuelle à *partir de 1900*.
- Résolution spatiale de ~50 km sur le territoire canadien.

Ressources Naturelles Canada produit une grille interpolée en utilisant la méthode ANUSPLIN (**NRCANMet**)

- Résolution temporelle journalière à *partir de 1950*.
- Résolution spatiale de ~10 km sur le territoire canadien.
- ANUSPLIN : **Hutchinson, M. F. (1995)**
(<https://doi.org/10.1080/02693799508902045>)

Données interpolées - Québec

Produit de la MELCCFP

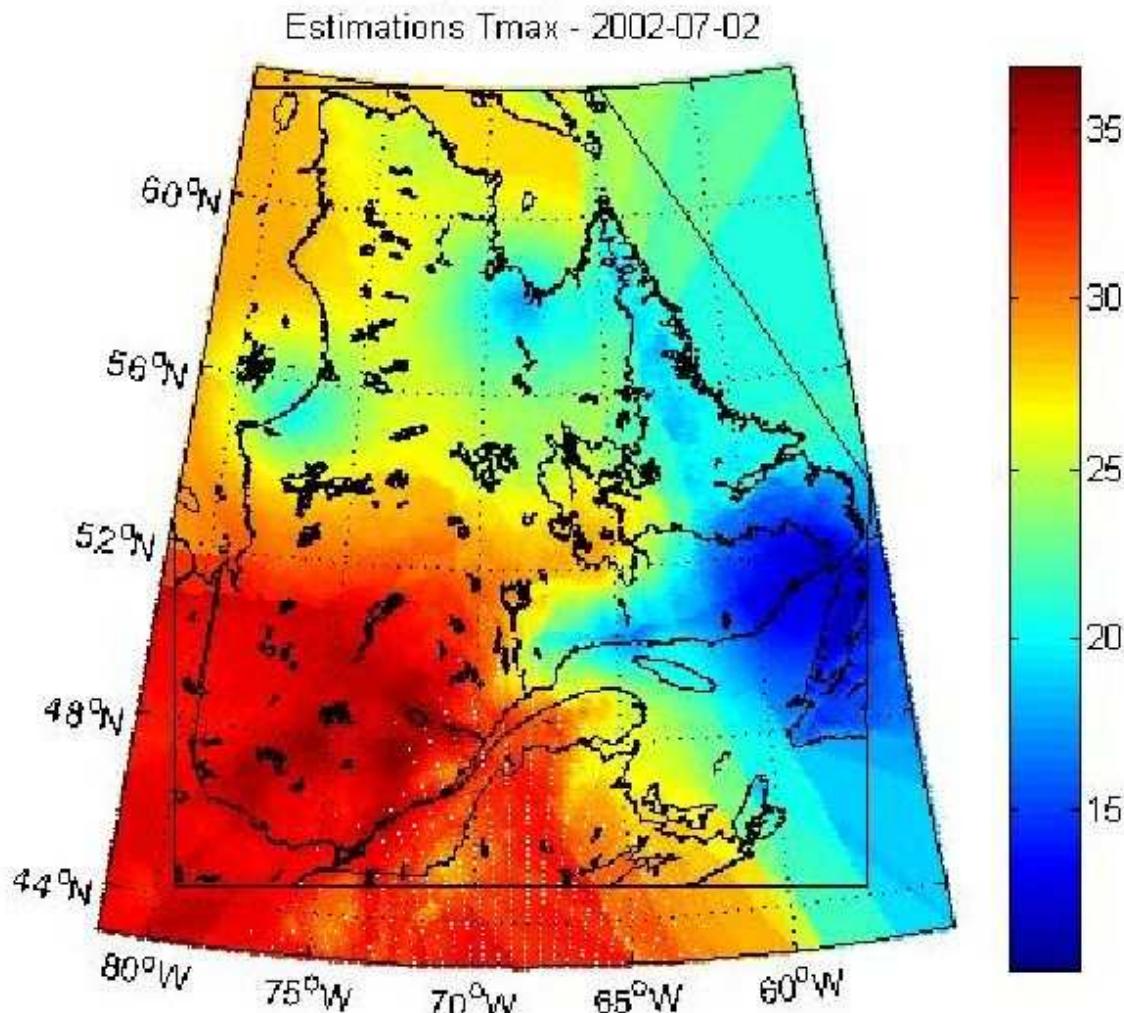
« *Grille climatique quotidien* »

- Méthode interpolation : Krigeage
- Résolution temporelle journalière à partir de 1961.
- Résolution spatiale de ~10 km sur le territoire de Québec.
- **Données non-public, sauf pour les chercheurs**

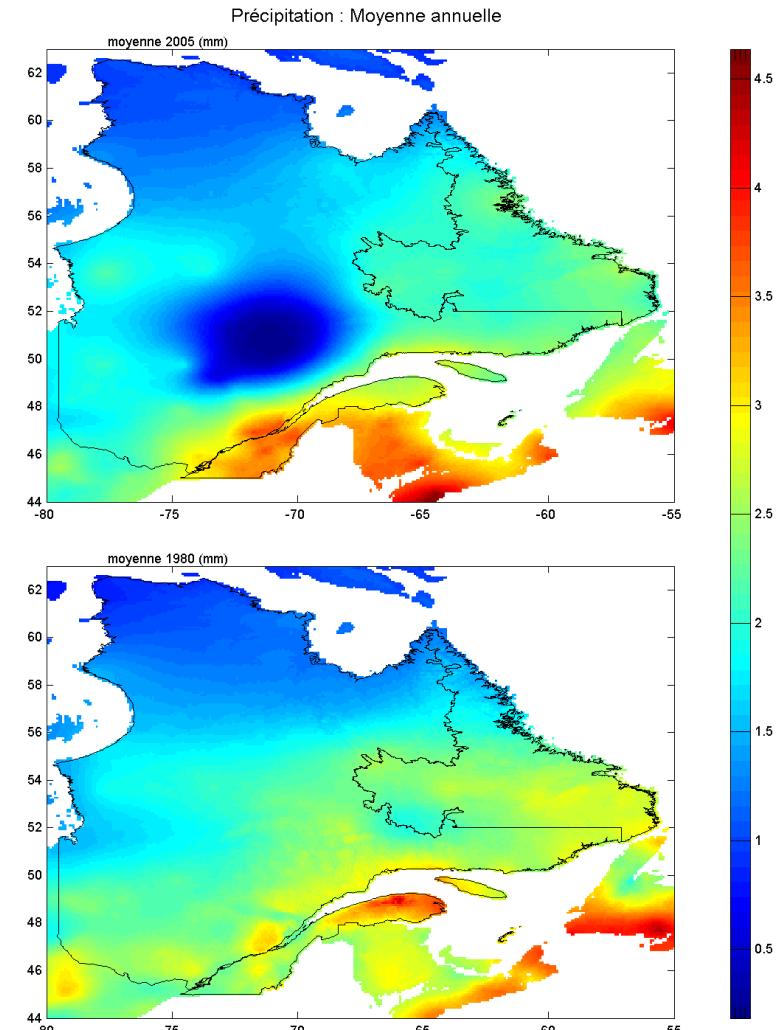
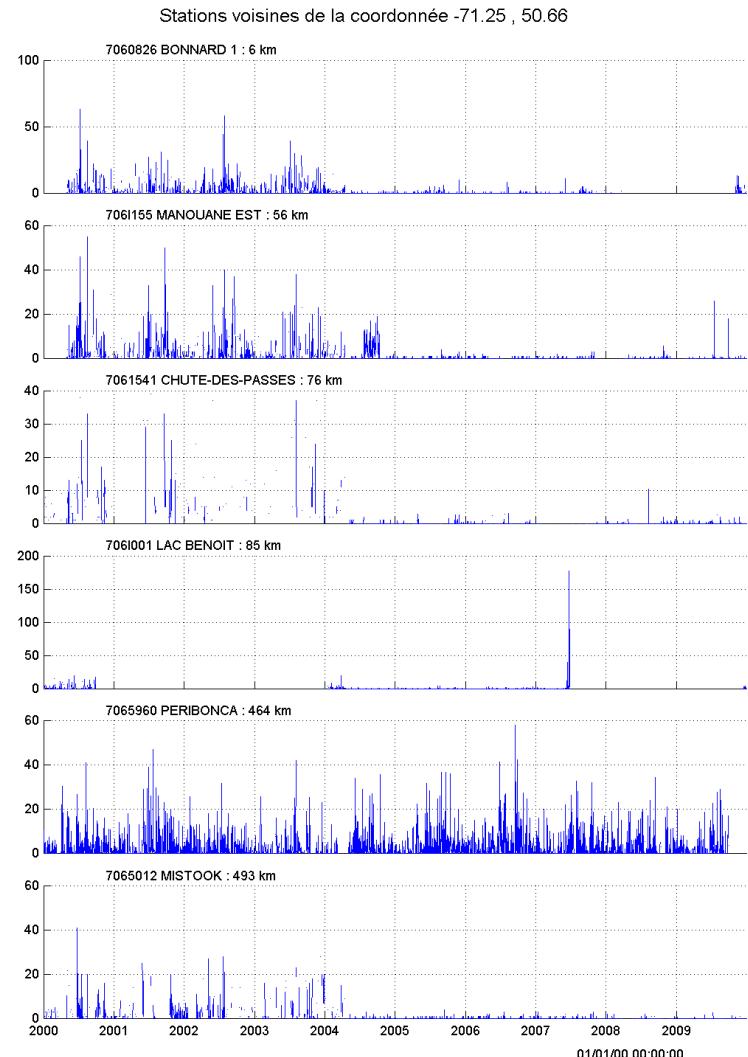
Source : Service Info-Climat (MELCCFP)

(<https://www.environnement.gouv.qc.ca/climat/surveillance/index.asp>

p)



Attention - Propagation des erreurs



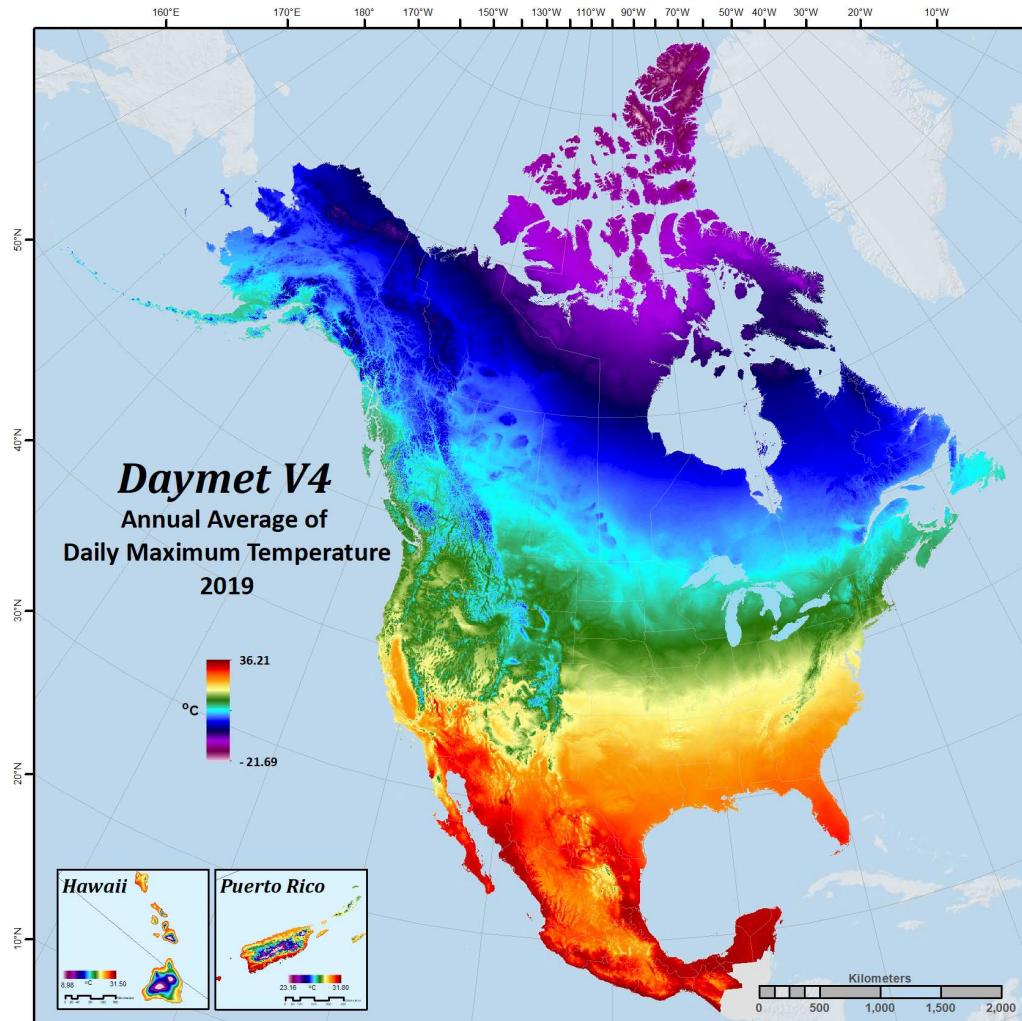
Données interpolées – Amérique du Nord

« DayMET » : Interpolation de NASA

- Résolution temporelle quotidien à partir de 1980
- Résolution spatiale de ~1 km sur Amérique du Nord

Intégration des données :

- Aux stations (États Unis, Canada, Mexique)
- Satellitaires (NASA SRTM)



Données interpolées - Conclusions

Forces

- Couverture spatiale complète à haute résolution.
- Séries temporelles continues.

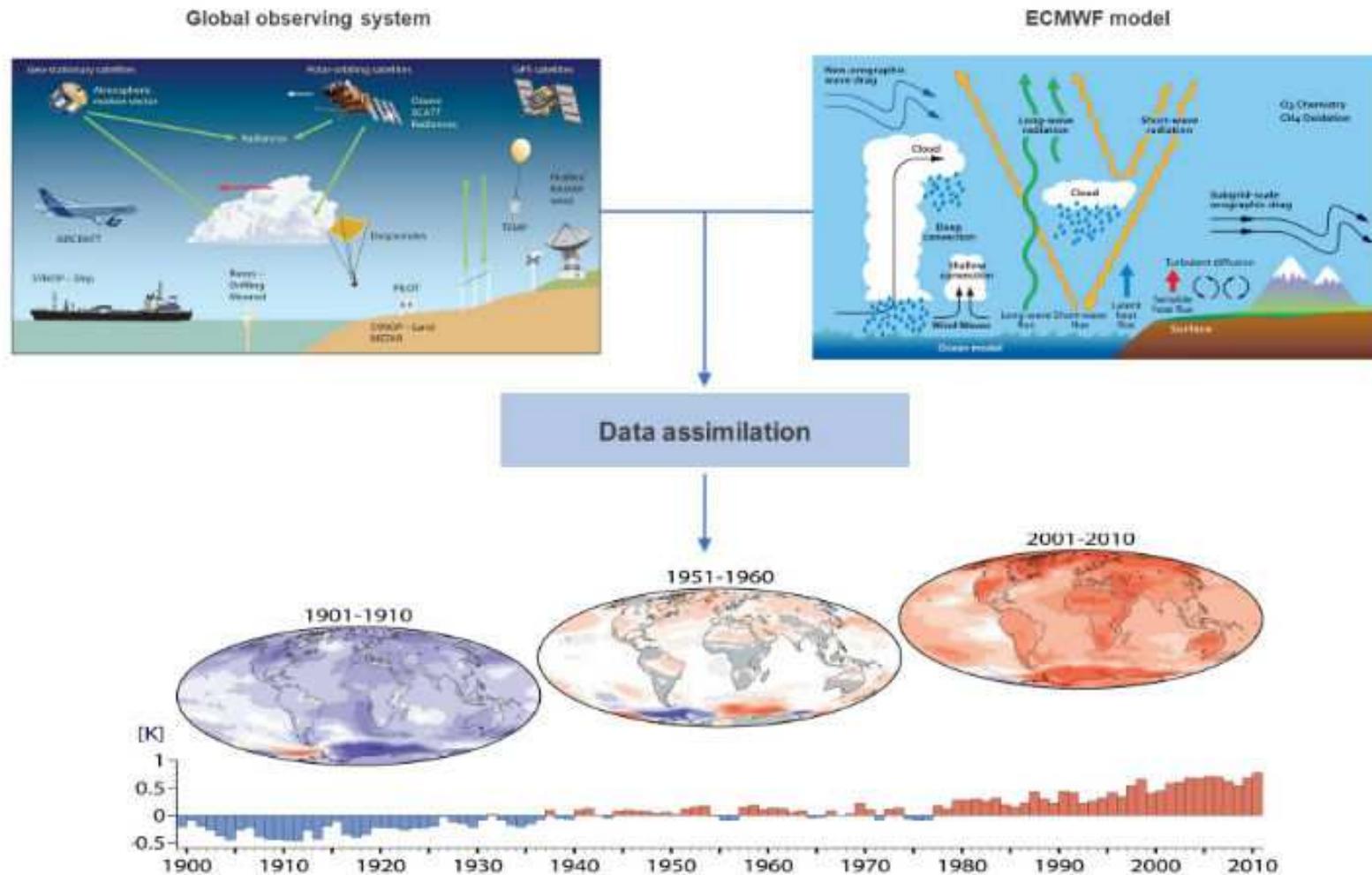
Faiblesses

- Les données dans les régions à faible densité de stations sont de moindre qualité.
- Peu d'information sur les incertitudes liées aux méthodes d'interpolation.
- Les erreurs dans les données de stations utilisées se propagent dans les données interpolées.

4. Données de réanalyses et prévisions

Expériments de « Hindcasting » et « Forecasting »

Réanalyses



Source : ECMWF

Vidéo : <https://www.youtube.com/watch?v=FAGobvUGI24>

Historique des réanalyses

Première génération (~1990) : **NCEP1, ERA-15**

Deuxième génération (~2000) : **NCEP2, NARR, JRA-25, ERA-40, ERA-interim**

Troisième génération (~2008) : **CFSR, 20CR, MERRA, JRA55, ERA20C**

Quatrième génération (~2014) : **CFSv2, MERRA2, ERA-20CM**

et ça continue... (~2019 et après) : **ERA5, ERA5-Land, RDRS**

Données des réanalyses

Réanalyses modernes :

- **Résolutions spatiales entre ~25 à 10 Km (ou moins!)**
- **Résolution temporelle d'archivage jusqu'aux heures**

Plusieurs Go de données par mois/niveau/variable.

- *16 octets * 24 heures * ~30 jours * 12 mois * (360 / résolution longitude) * (180 / résolution latitude) * nombres de niveaux en altitude * etc... = ENORME*

Quelques exemples :

- **ERA-Interim (ancien)** : 30 ans de réanalyses pour une vingtaine de variables
 - ~7 To de données en totale
- **ERA5** : 11 To de nouvelles données par mois
 - **ERA5-Land** : Résolution encore plus fine (~7 fois plus fine qu'ERA5)

Réanalyses appliquées

La comparaison de réanalyses et l'évaluation de leurs forces et faiblesses respectives est encore très anecdotiques.

Plusieurs avancements dans l'assimilation des données, mais...

- **Les incertitudes associées au jeu de données de référence sont souvent ignorées lors de la validation !**

Pour en savoir plus :

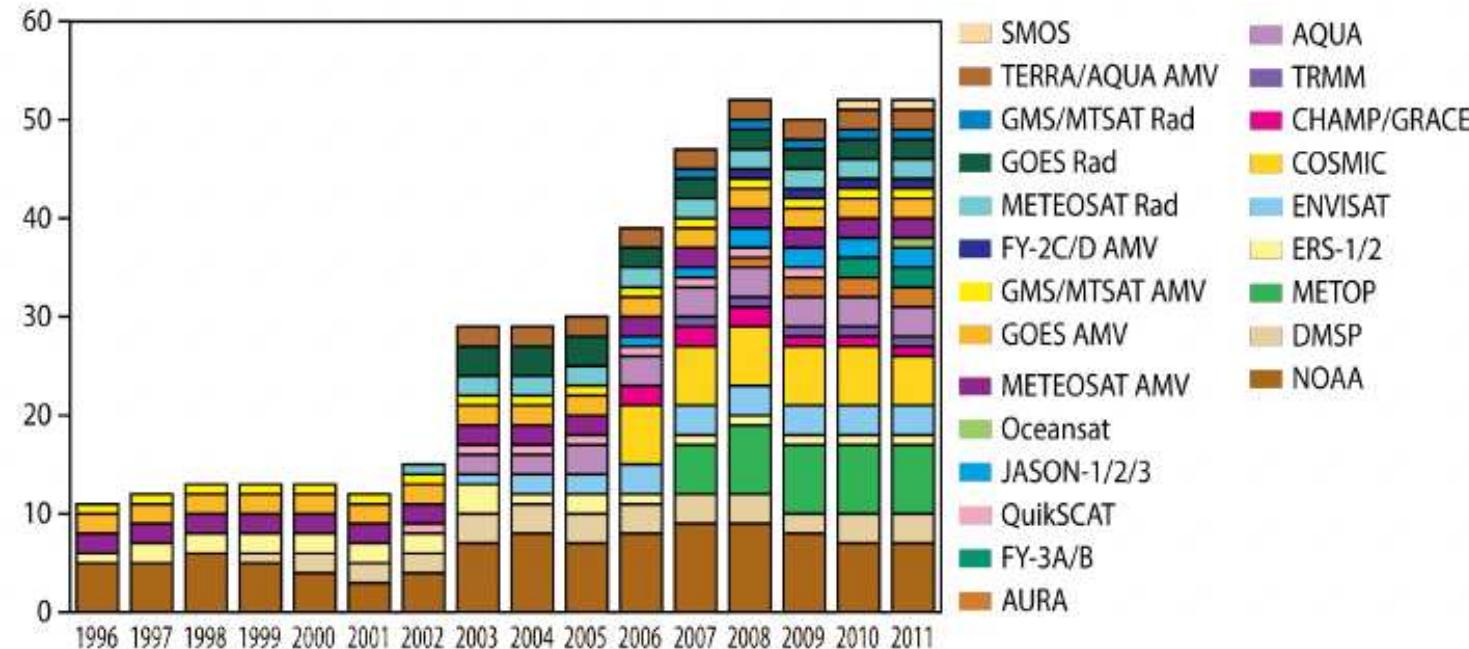
- <https://reanalysis.org/>
- <https://climatedataguide.ucar.edu>

Intégration des données

Augmentation du nombre des données des sources satellitaires intégrés dans la production des réanalyses ECMWF de 1996 to 2011.

Pour en savoir plus :

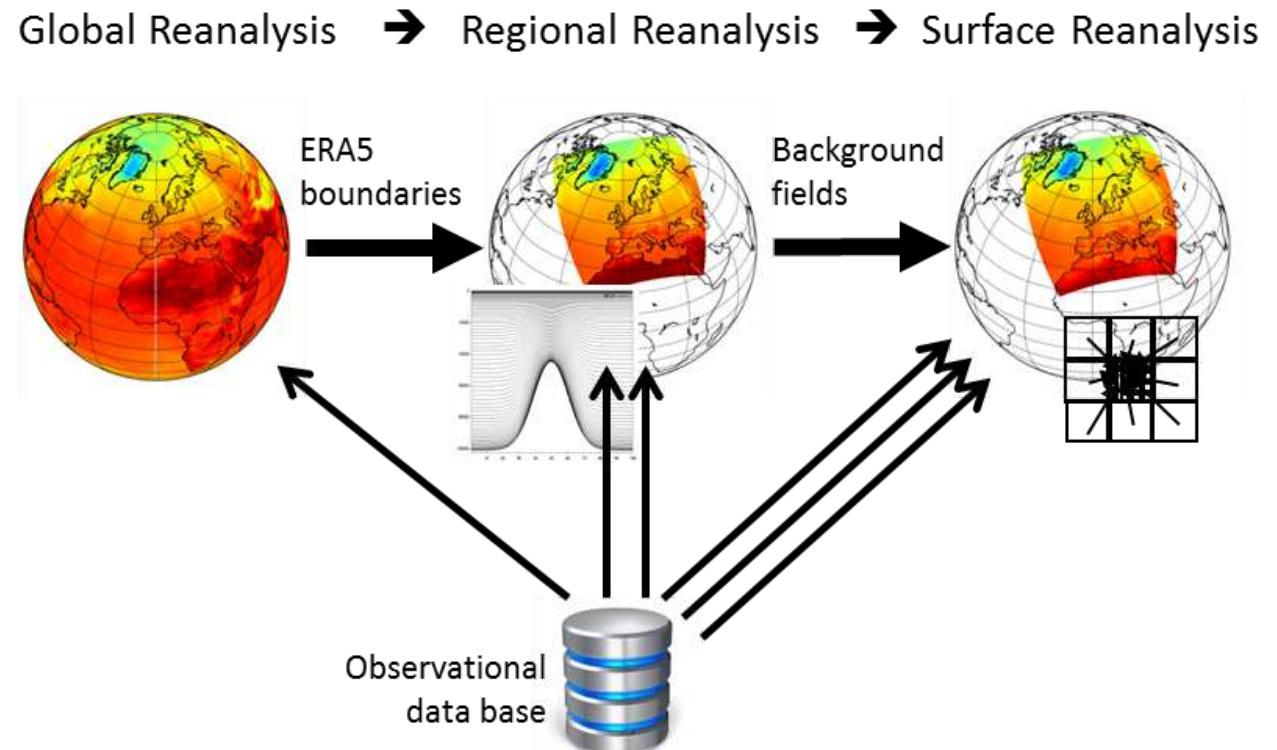
- [Simmons. 2017](#)



Réanalyses « ré-injectées »

Jeu de donnée hybride à base de :

- Modèles typiquement utilisées pour générer des prévisions
- Données observées intégrées
- Piloté par une autre réanalyse



Source : CERRA (Copernicus)

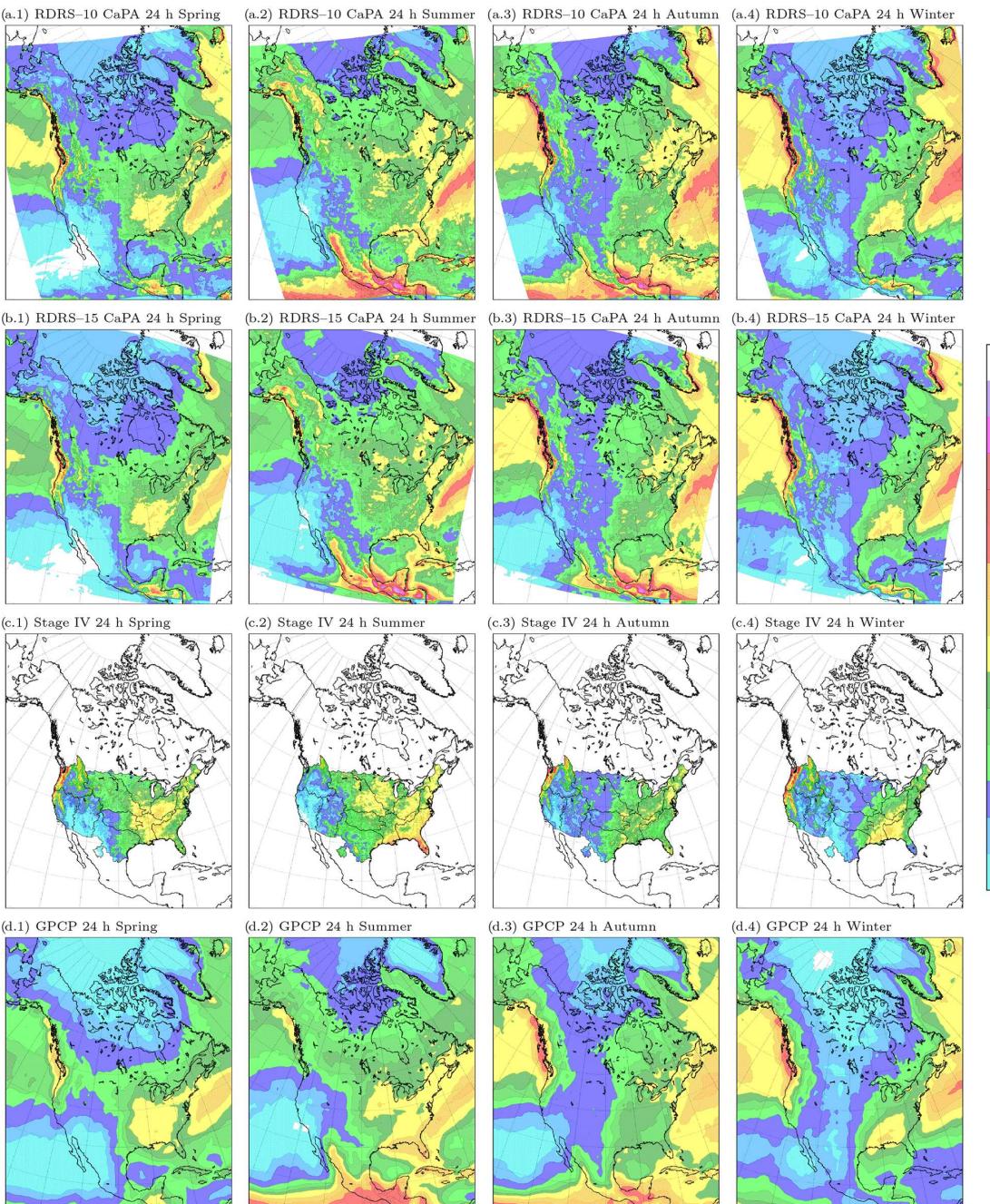
(<https://climate.copernicus.eu/copernicus-regional-reanalysis-europe-cerra>)

Réanalyses « ré-injectées »

RDRS (Environnement Canada)

- Modèle de prévision « **GEM** »
- Données de stations **RMCQ**,
MELCCFP
- Piloté par réanalyses **ERA-interim** ou
ERA5

Source : Gasset et al. (2022) A 10 km North American precipitation and land-surface reanalysis based on the GEM atmospheric model.
(<https://doi.org/10.5194/hess-25-4917-2021>)



Réanalyses

- Plusieurs produits disponibles
- Forces et faiblesses avec certains produits dépendamment la phénomène
- Débat continue entre les producteurs et utilisateurs

Source : Sabarly, et al. (2016)

(<https://doi.org/10.1175/JHM-D-15-0093.1>)

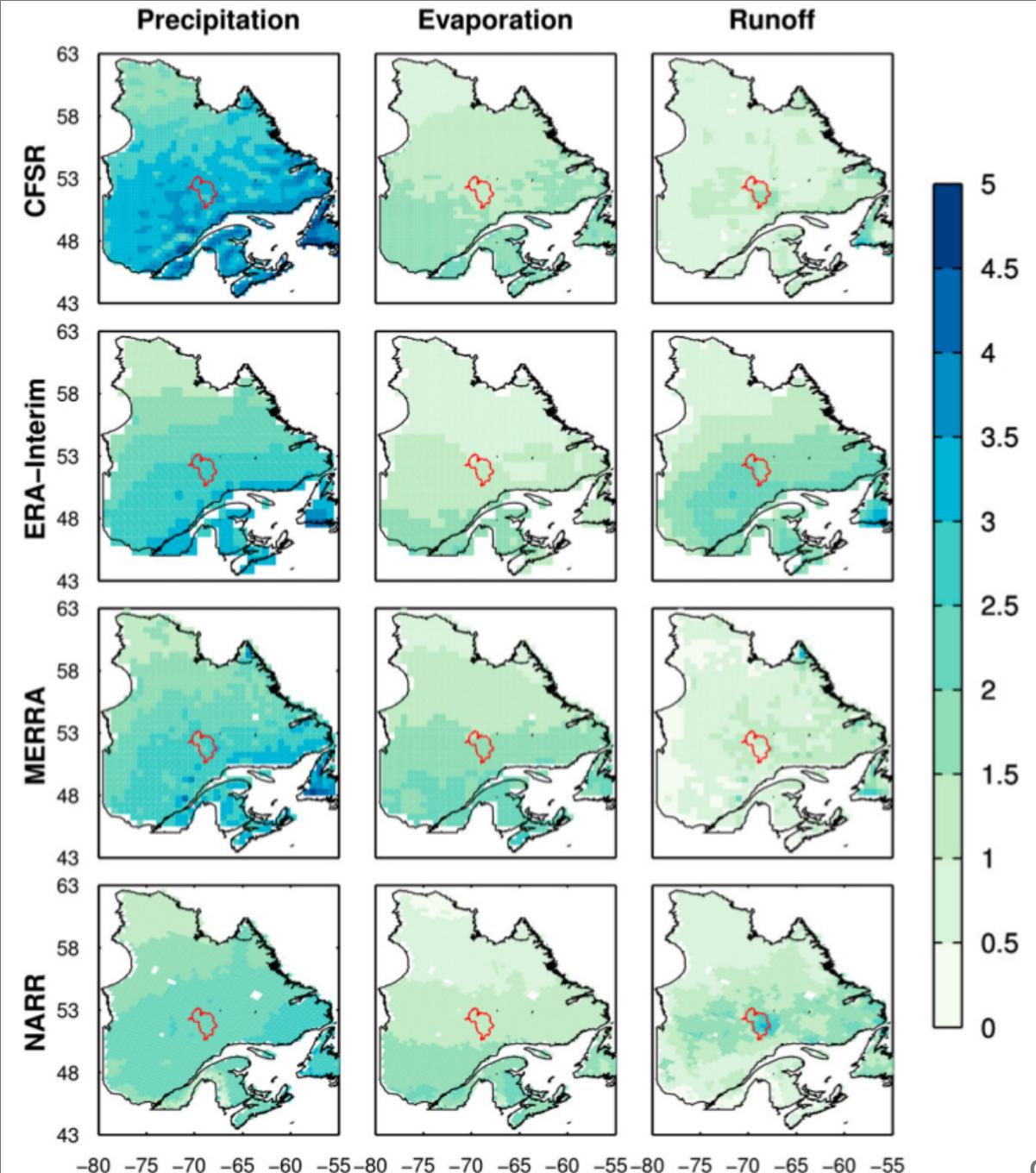
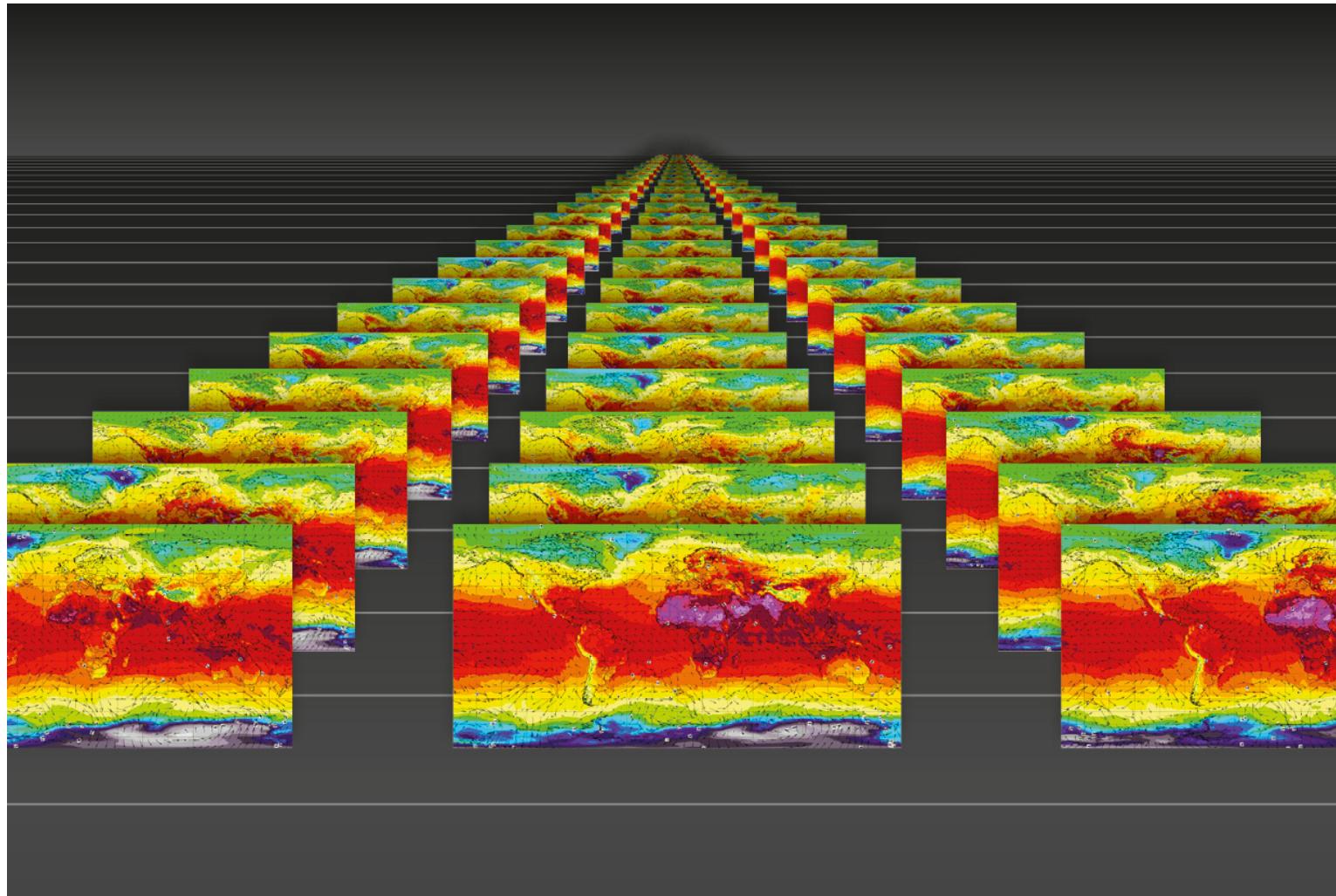


FIG. 3. Long-term mean precipitation, evaporation, and runoff (mm day^{-1}) of CFSR, ERA-Interim, MERRA, and NARR over the province of Quebec. The red line represents the contour of the river basin Manic-5 in which Manicouagan Lake is located.

Modèles de prévision



Modèles de prévision

Modèles utilisés pour faire des prévisions pour (souvent) les *14 jours à venir*.

Le service météorologique du Canada roule 20 configurations du modèle « **GEM** »

- *2 * jour*, sur un horizon de *16 jours*

Archivage de durée limitée et rarement utilisé pour des études climatiques.

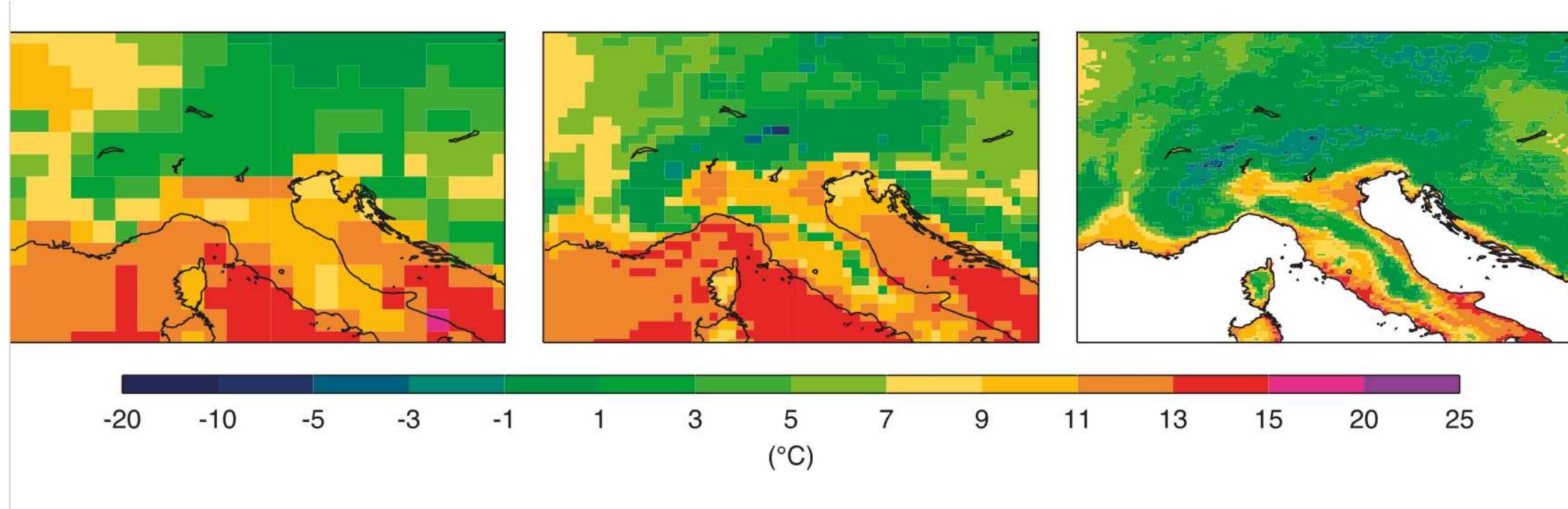
Intégration des données observées de plusieurs sources pour piloter les prévisions

Prévisions en temps réelle



Source : Ventusky (<https://www.ventusky.com/?p=47.515;-70.201;10&l=temperature-2m>)

À côté - Résolutions



- **Résolutions plus fines : Choix des variables ↓, tailles des domaines ↓**
- **Résolutions plus grossières : Choix des variables ↑, tailles des domaines ↑**

Réanalyses - Conclusions

Forces

- Couverture spatiale complète.
- Séries continues à haute résolution temporelle
- Cohérence physique entre les variables.
- Utilise un maximum de données observées pour initialiser le modèle plusieurs fois par jour.
- Variables souvent beaucoup plus intéressantes.

Faiblesses

- Plusieurs phénomènes encore paramétrés (p. ex. précipitation).
- Les modèles de réanalyses / prévisions ont des différences majeures entre elles.
- Les normes relatives aux variables ne sont pas toujours cohérentes entre les jeux des données réanalyses / prévisions et les modèles climatiques.

5. Données des modèles climatiques

Modèles du climat

Modèles utilisés pour faire de longues séries climatiques, soit des possibilités climatiques passées ou des projections futures.

Archivage de longue durée pour une sélection de variables climatiques jusqu'à des fréquences horaires.

Expériences fortement guidées par les scénarios développés par le **GIEC** (le « **IPCC** »)

- **AR/AR4 → SRES; AR5 → RCP; AR6 → SSP**

Modèles du climat global (GCM)

Les modèles du climat modernes ont une résolution spatiale allant jusqu'à **50 Km*** et une résolution temporelle d'archivage jusqu'aux heures.

- *Ce chiffre peut descendre, dépendamment le modèle / expériment

La base de données CMIP5 récupérée à Ouranos fait environ *20 To* pour quelques variables, ce qui représente seulement une infime partie de la base de données CMIP5 complète.

Données CMIP6 : Pas raisonnable de télécharger toute la base

- *8 ou 16 octets * expérience * modèle * réalisation * niveaux * pas de temps * projet*
* etc... = **ENORME**

Coordination des modèles globales

WCRP : « World Climate Research Program »

- CMIP : « Coupled Model Intercomparison Project »

Standardisation de la priorité des expériences évaluées

Standardisation du format / métadonnées de données

- Vocabulaire Contrôlée

Établissement de portail de données sur le web

- ESGF : « Earth System Grid Federation »

Présentement : CMIP6 (utilisation)

Structure organisationel

École d'été en sciences du climat 2024

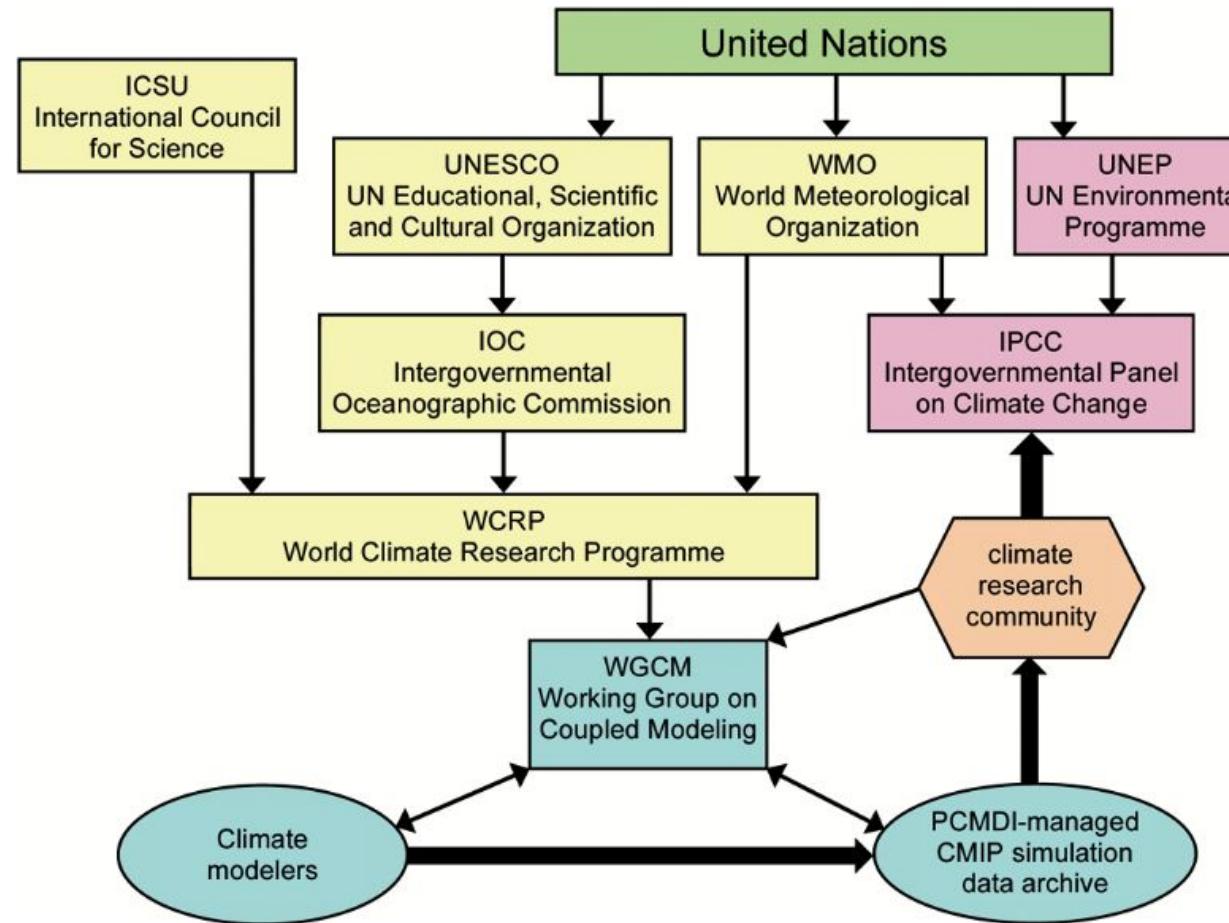
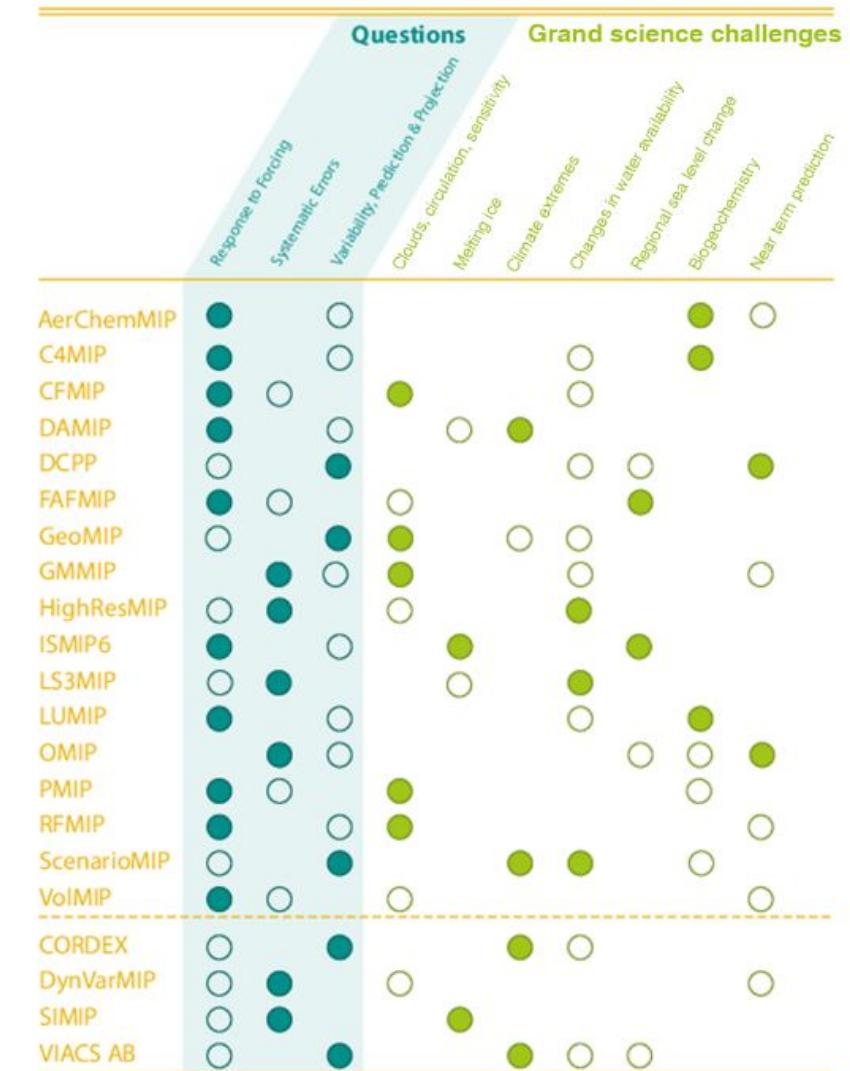
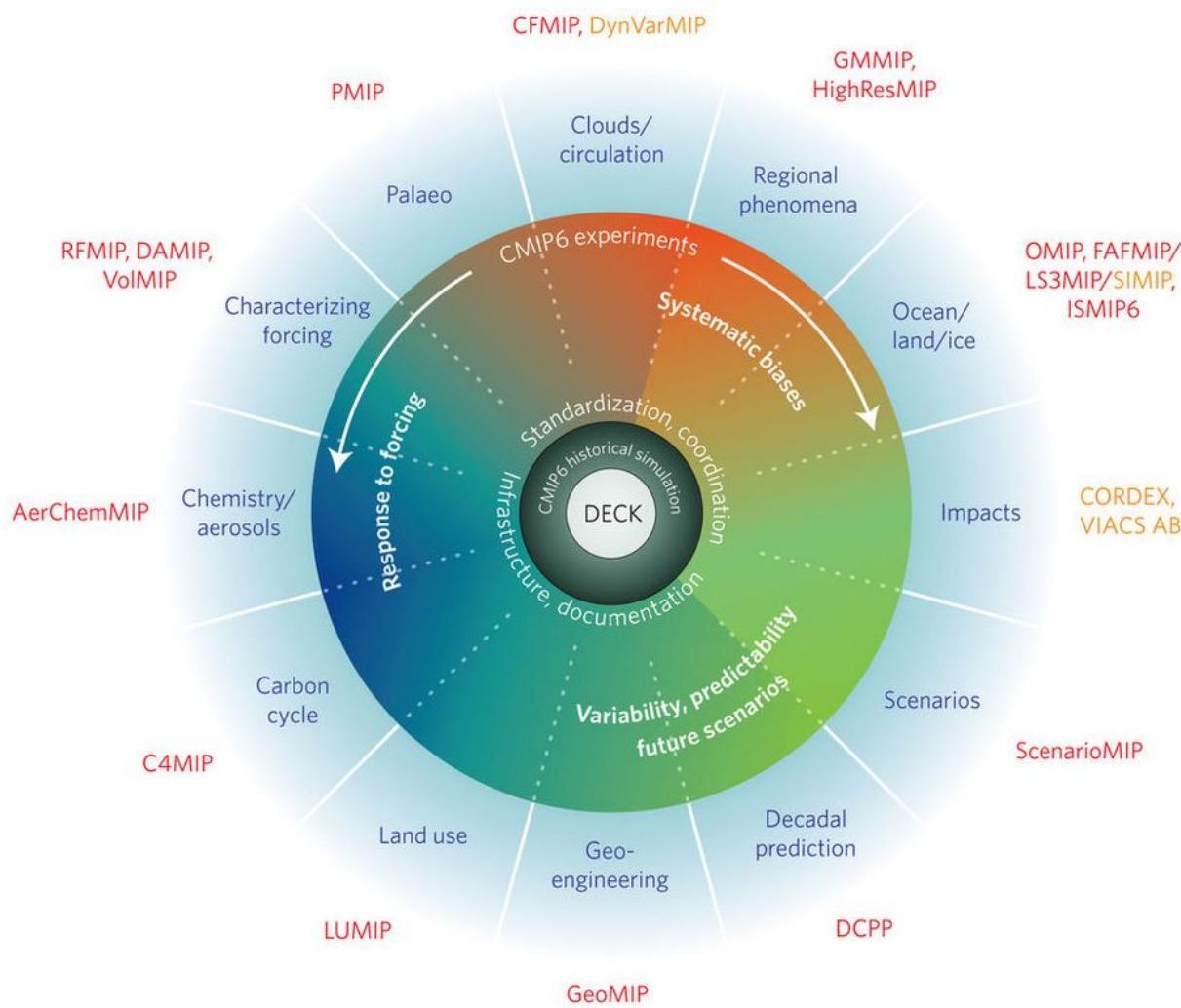


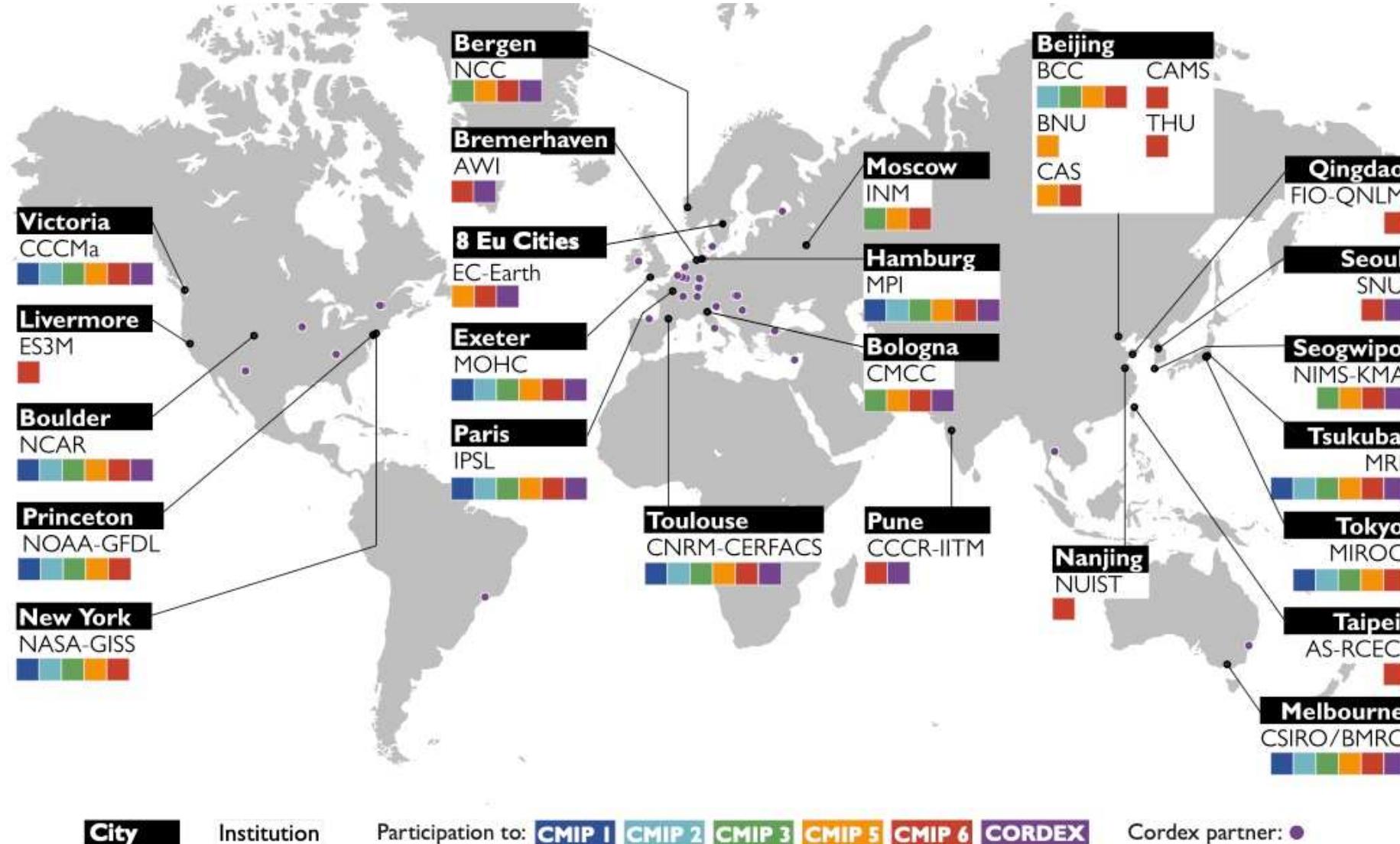
FIG. 1. The relationship of CMIP5 to organizations established to coordinate climate research activities internationally and to the IPCC, the modeling centers, and the climate research community.



Source : Eyring et al. 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization (<https://doi.org/10.5194/gmd-9-1937-2016>)

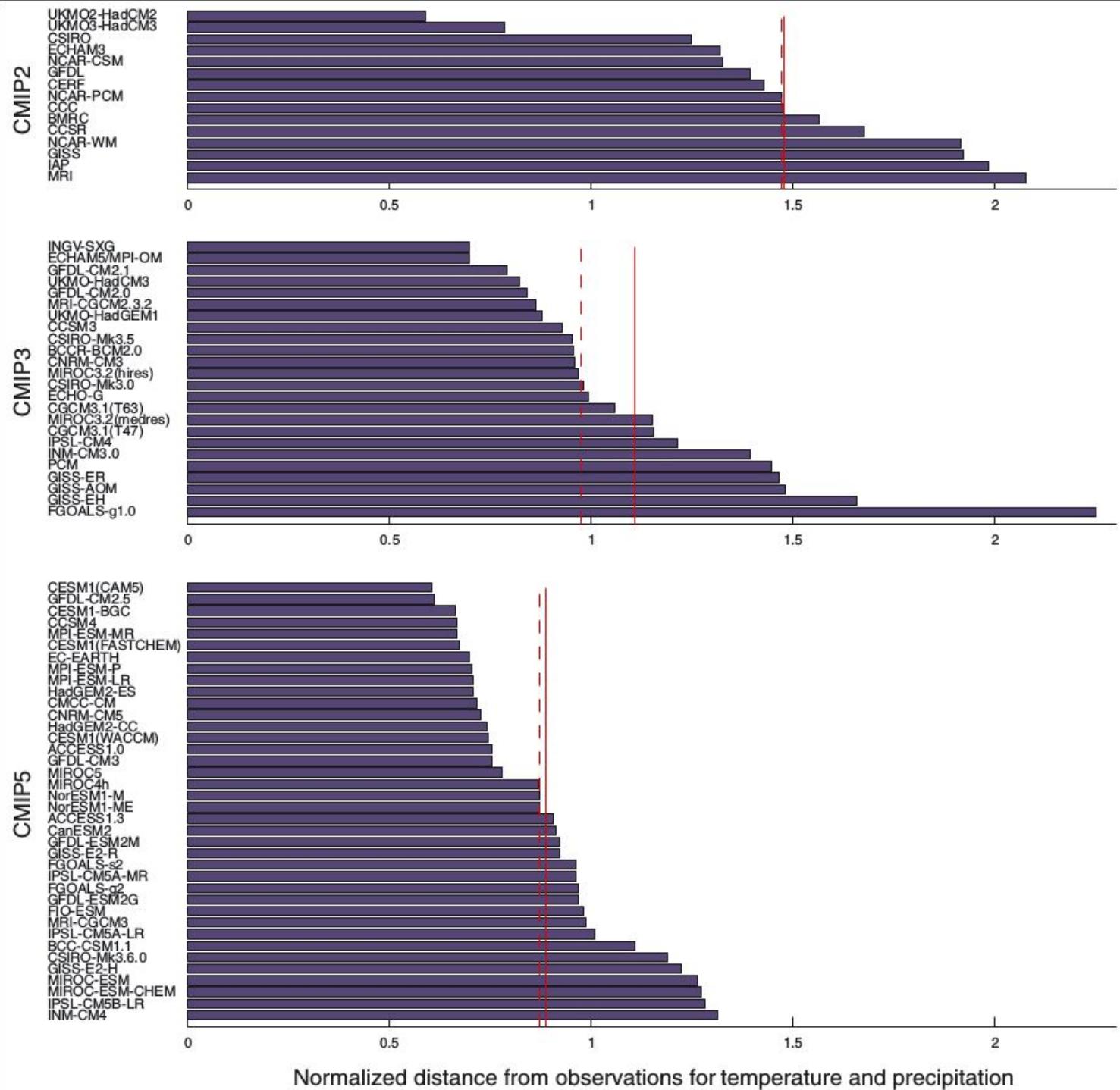
Instituts participants à CMIP

École d'été en sciences du climat 2024



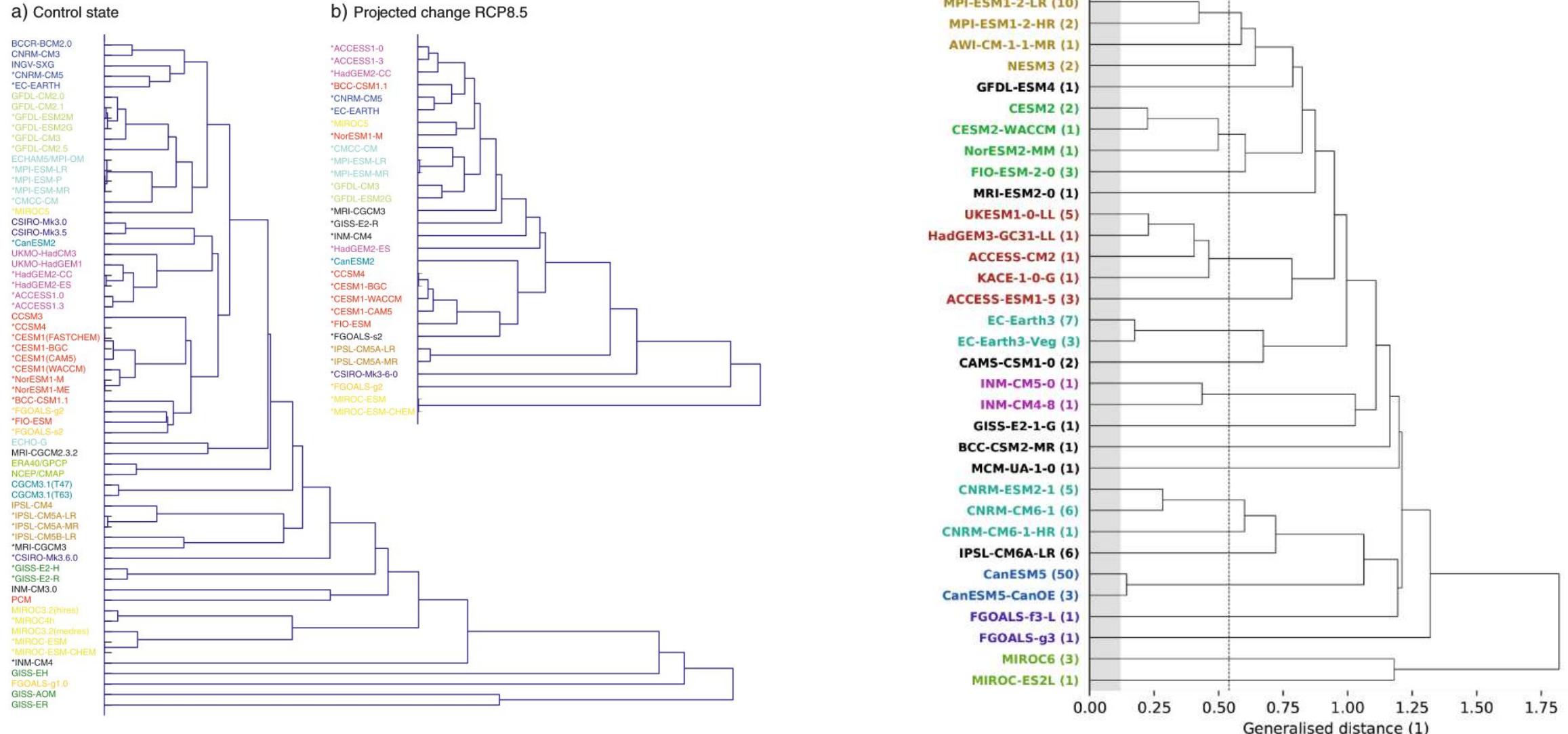
Modèles participants de CMIP2 à CMIP5

NOTE: CMIP4
n'existe pas!



Source : Knutti, et al. (2013) Climate model genealogy: Generation CMIP5 and how we got there
(<https://doi.org/10.1002/grl.50256>)

Généalogies des modèles participants de CMIP5 et CMIP6



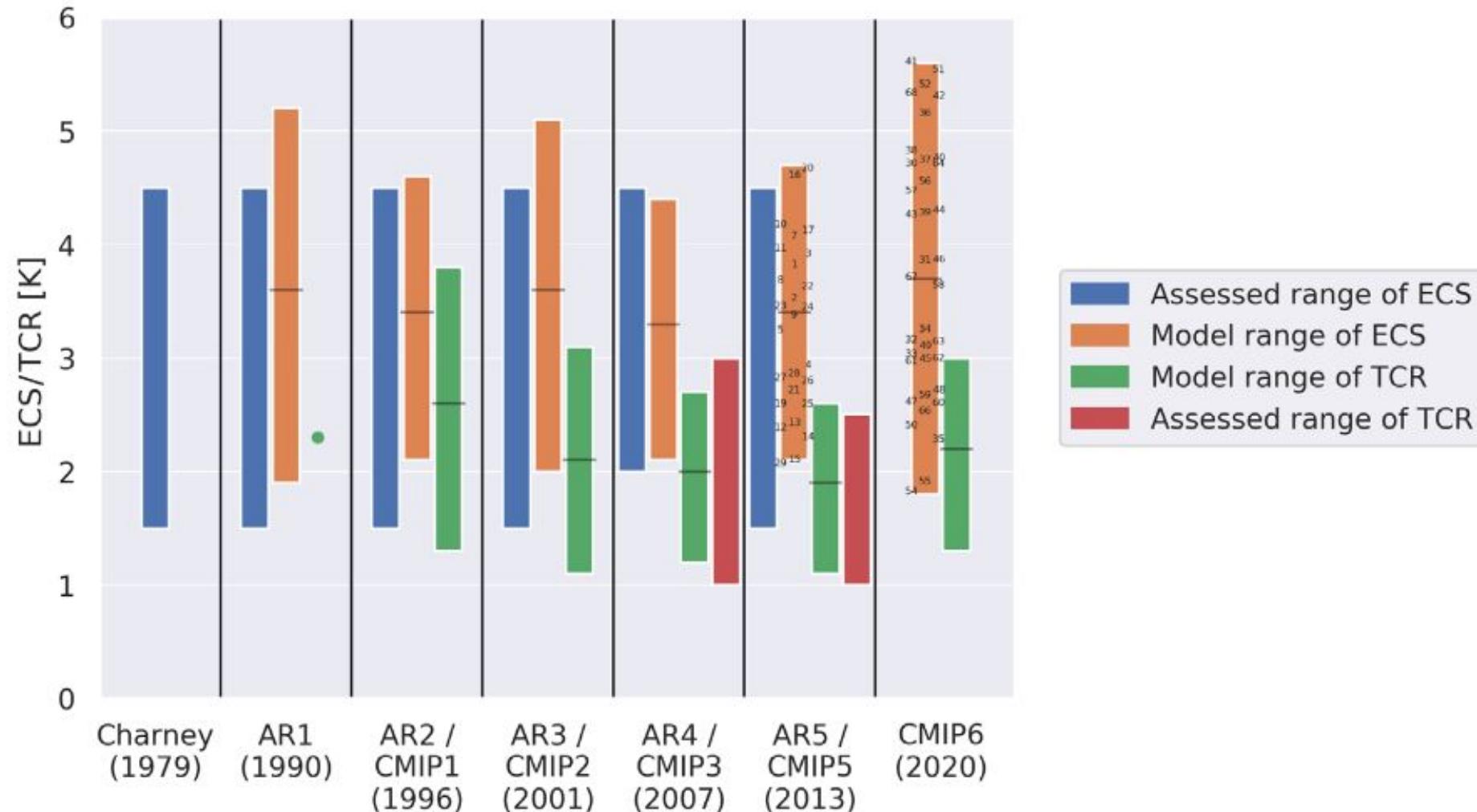
Sources : Knutti et al. (2013) - (<https://doi.org/10.1002/grl.50256>)

Brunner et al. (2020) - (<https://doi.org/10.5194/esd-11-995-2020>)

Comparaison de modèles CMIP

École d'été en sciences du climat 2024

Equilibrium climate sensitivity (gregory method) and transient climate response



Meehl et al. (2020) Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models (<https://doi.org/10.1126/sciadv.aba1981>)

Modèles du climat globales - Conclusions

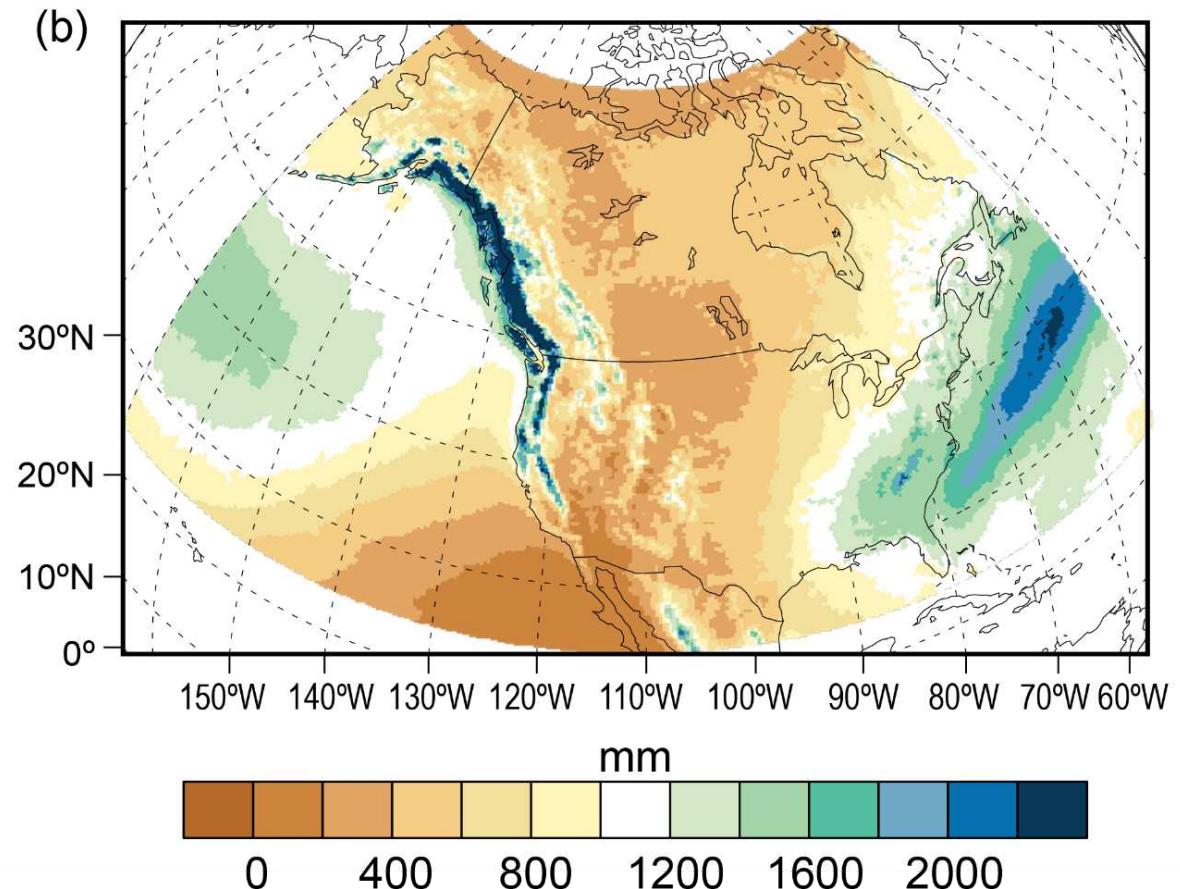
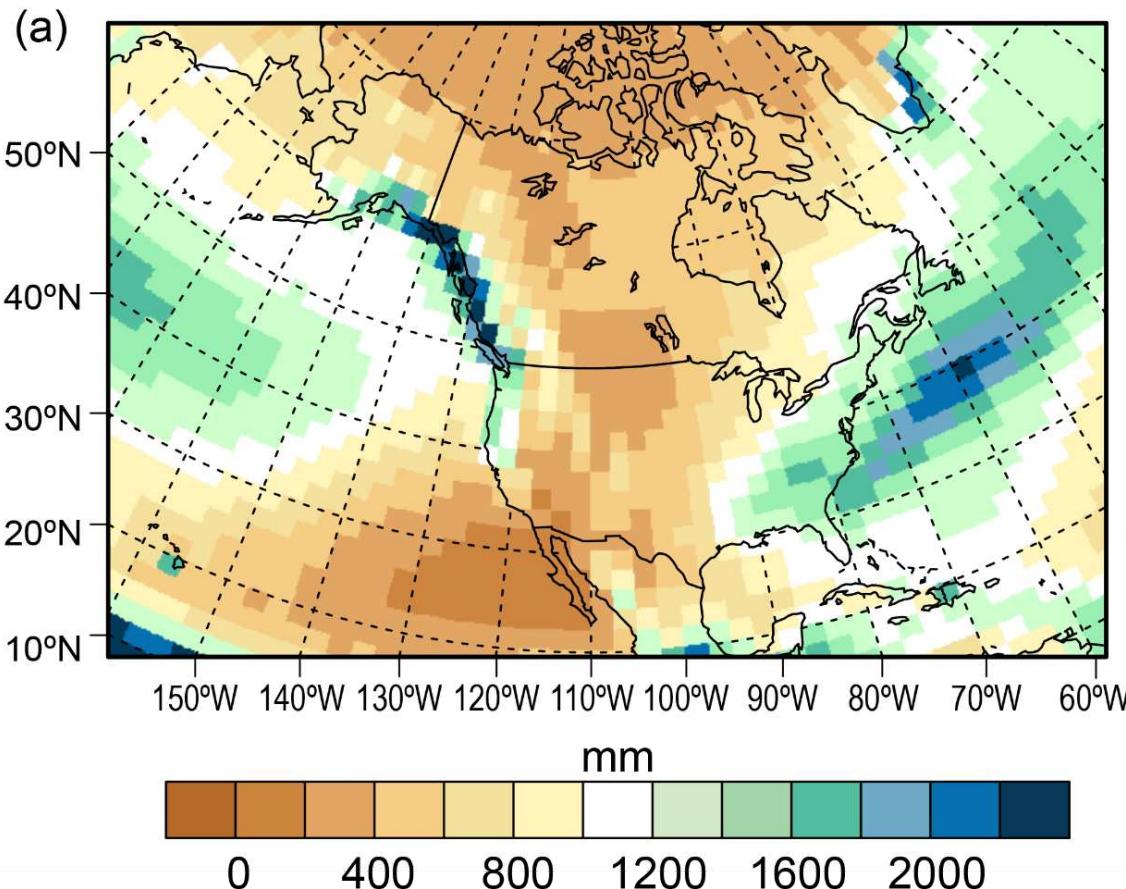
Forces

- Couverture spatiale complète.
- Séries continues.
- Cohérence physique entre les variables.
- Permet de faire des projections dans le futur.
- Standards souvent bien suivis pour le calcul des variables et formats de données

Faiblesses

- Plusieurs phénomènes encore paramétrés (p. ex. types spéciales de la précipitation).
- Les modèles ont des différences majeures entre eux.
 - Analyses des statistiques nécessaires
- Les résolutions spatiales ne supportent pas les analyses aux échelles locales / régionales

Modèles régionaux du climat (RCM)



Source : Kotamarthi, et al. (2021) Global Climate Models. In: Downscaling Techniques for High-Resolution Climate Projections: From Global Change to Local Impacts. Cambridge University Press; 2021:19-39.

Modèles régionaux du climat (RCM) - CORDEX

Tout comme la communauté de modélisation globale, la modélisation régionale a aussi établi des protocoles d'expériences.

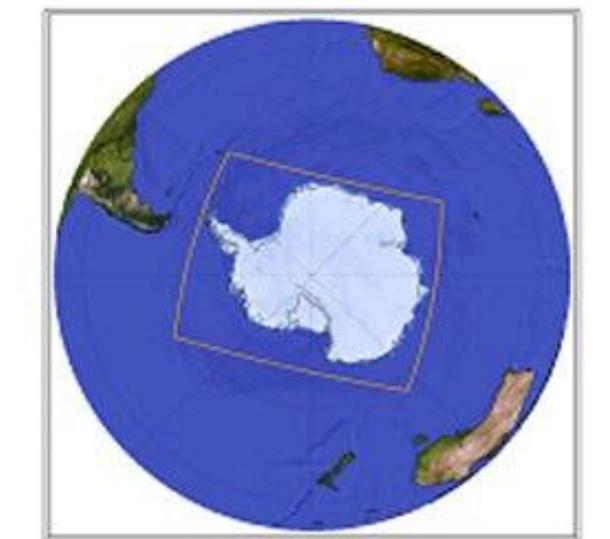
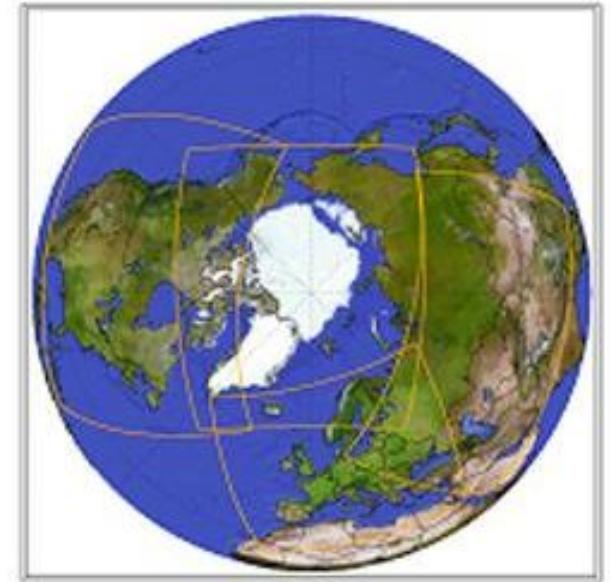
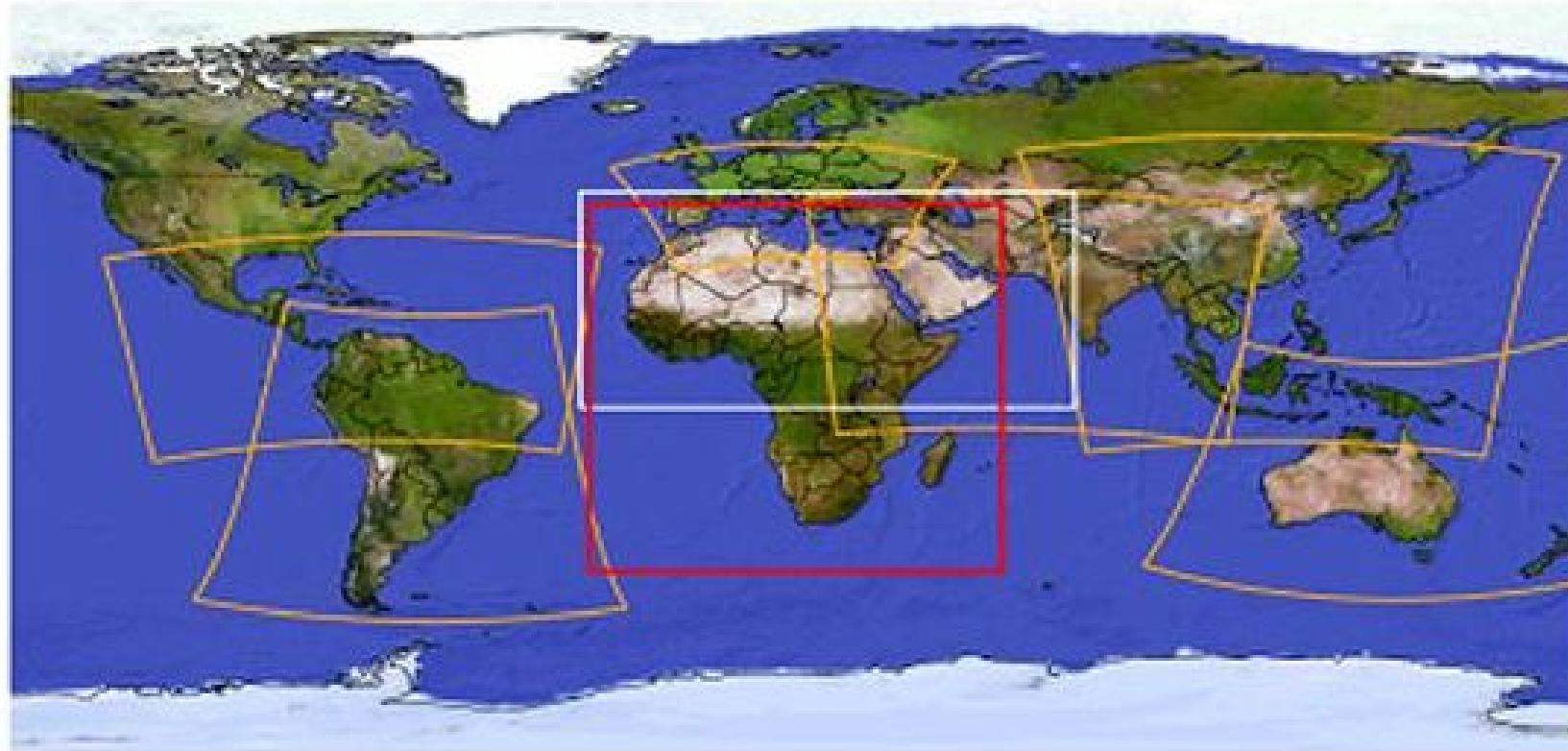
Les projets **PRUDENCE**, **ENSEMBLES** et **NARCCAP** sont les prédecesseurs de **CORDEX** :

- « COordinated Regional Downscaling EXperiment »

CMIP5-CORDEX (passé) et CMIP6-CORDEX (en production)

- Plus que 10 instituts de recherche font partie du projet CMIP6-CORDEX
 - **Ouranos** et **UQAM**

CMIP5-CORDEX



Modèles Régionaux

	RegCM4	WRF	CRCMS -OUR	CRCMS -UQAM	CanRCM4	RCA4	HIRHAM5	RCP	ECS (°C)
ERA-Int	50km 25km	50km 25km	0.22°	0.44° 0.22° 0.11°	0.44° 0.22°	0.44°	0.44°		
HadGEM2-ES	50km 25km	50km* 25km*						8.5	4.6
CanESM2				0.44° 0.22°	0.44° 0.22°	0.44°		4.5	3.7
GEMatm-Can 1†				0.44° 0.22°				8.5	~3.7 †
MPI-ESM-LR	50km* 25km*	50km 25km	0.22°	0.44° 0.22°				4.5	3.6
MPI-ESM-MR				0.44° 0.22°				8.5	3.4
GEMatm-MPI †				0.44° 0.22°				8.5	~3.4 †
EC-EARTH 3‡					0.44°			2.6	
					0.44°	0.44°	4.5	-3.3	
					0.44°	0.44°	8.5		
CNRM-CM5				0.22°				4.5	-3.3
				0.22°				8.5	
GFDL-ESM2M	50km 25km	50km* 25km*	0.22°					8.5	2.4
Institution	Iowa State *NCAR	U Arizona *NCAR	Okanagan	UQAM	CCCma	SMHI	DMI		
Modeler	R. Arritt *M. Bakovskiy	H.-I Chang *M. Bakovskiy	S. Bliner	K. Winger	J. Scinocca	G. Nikulin	O. Christensen		
Access	PoC	PoC	PoC	PoC	CCCma	ESGF	ESGF		

Instituts et responsables

Réanalyse †

Modèles Globaux

CMIP5-
CORDEXDomaine Amérique
du nord

- « NA-
CORDEX »

Couverture des
expériences inégales

Le reseau ESGF

Portail web

Distribué entre
plusieurs centres
de recherche

Source : ESGF-LLNL (<https://esgf-node.llnl.gov/search/esgf-llnl/>)

Hosted by  Department of Energy
Lawrence Livermore National Laboratory

Powered by 

ATTENTION: We have completed the rename of this ESGF server to esgf-node.llnl.gov from pcmdi.llnl.gov. All pcmdi.llnl.gov openids have become esgf-node.llnl.gov for the server name (needed for manual entry in a wget script execution). Several data nodes are still in the process of updating their configuration to account for the server rename. Until complete, there might still be issues in accessing data at some sites. We recommend trying replica data (if available) in the meantime and apologize for the inconvenience.

Welcome, Guest | Login | Create Account

 Lawrence Livermore National Laboratory

You are at the ESGF@DOE/LLNL node | Technical Support

Home Contact Us

This search interface supports data discovery from many ESGF projects on multiple sites.
Please use the "Search All Replicas" checkbox to expand the number of models available.

Project: CMIP5 (52984)

Product: Institute: Model: Experiment: Experiment Family: Time Frequency: Realm: CMIP Table: Ensemble: Variable: Variable Long Name: CF Standard Name: Driving Model: Datanode: obs4MIPs: Instrument: ISIMIP-FT: Impact Model: Impact Sector:

Enter Text: Search Reset Display 10 results per page [More Search Options]

Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)

Search Constraints: CMIP5

Total Number of Results: 52984
-1 2 3 4 5 6 Next >

Please login to add search results to your Data Cart
Expert Users: you may display the search URL and return results as XML or return results as JSON

- project=CMIP5, model=MRI-CGCM3, Meteorological Research Institute, experiment=10- or 30-year run initialized in year 1965, time_frequency=mon, modelingrealm=atmos, ensemble=r8i1p1, version=20120701
Description: MRI-CGCM3 model output prepared for CMIP5 10- or 30-year run initialized in year 1965
Data Node: esgf-data1.diasjp.net
Version: 20120701
Total Number of Files (for all variables): 100
[Show Metadata] [Show Files] [THREDDS Catalog] [WGET Script]
- project=CMIP5, model=MRI-CGCM3, Meteorological Research Institute, experiment=10- or 30-year run initialized in year 1965, time_frequency=mon, modelingrealm=atmos, ensemble=r9i1p1, version=20120701
Description: MRI-CGCM3 model output prepared for CMIP5 10- or 30-year run initialized in year 1965
Data Node: esgf-data1.diasjp.net
Version: 20120701
Total Number of Files (for all variables): 100
[Show Metadata] [Show Files] [THREDDS Catalog] [WGET Script]
- project=CMIP5, model=MRI-CGCM3, Meteorological Research Institute, experiment=10- or 30-year run initialized in year 1965, time_frequency=mon, modelingrealm=land, ensemble=r1i1p1, version=20110915
Description: MRI-CGCM3 model output prepared for CMIP5 10- or 30-year run initialized in year 1965
Data Node: esgf-data1.diasjp.net
Version: 20110915
Total Number of Files (for all variables): 9
[Show Metadata] [Show Files] [THREDDS Catalog] [WGET Script]

ESGF- Download

Gestion des données d'ESGF par ligne de commande

- Je contribue au projet sur GitHub!

Source : ESGF/esgf-download
(<https://esgf.github.io/esgf-download/>)

Introduction

esgpull is a modern ESGF data management tool, bundled with a custom asynchronous interface with the ESGF Search API. It handles scanning, downloading and updating **datasets**, **files** and *queries* from ESGF.

Feature highlight

- Simple syntax for fast data exploration
- Asynchronous download
- Highly configurable

Search datasets

esgpull allows multiple ways for searching ESGF data, with **facet** and **free-text** terms together with **options**.

Facet terms Free-text terms Facets hints Hints Finer grain query Date filter

Query every CMIP6 dataset using facet syntax

```
esgpull search project:CMIP6
```

Rich

dataset	files	size
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.3hr.hfls.gr.v20181203	4	10.2 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.3hr.clt.gr.v20181203	2	7.3 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.cfadLidarsr532.gr.v20181203	2	10.3 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clmisr.gr.v20181203	1	1.5 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clcalipso.gr.v20181203	1	162.6 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clhcalipso.gr.v20181203	1	82.5 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clmcalipso.gr.v20181203	1	94.7 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clcalipso.gr.v20181203	1	1.8 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clccalipso.gr.v20181203	1	136.0 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.clisccp.gr.v20181203	1	1.0 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.jpdftaureliqmodis.gr.v20181203	1	651.6 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.jpdftaureicemodis.gr.v20181203	1	431.2 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.parasolRefl.gr.v20181203	1	981.8 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.E3hrPt.hus.gr.v20181203	144	787.5 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.CFsubhr.hfss.gr.v20181203	1	390.1 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.CFsubhr.hfls.gr.v20181203	1	379.8 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.CFsubhr.mc.gr.v20181203	2	3.9 GiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.CFsubhr.prc.gr.v20181203	1	398.8 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.CFsubhr.huss.gr.v20181203	1	361.7 MiB
CMIP6.CMIP.CNRM-CERFACS.CNRM-CM6-1.amip.r1i1p1f2.CFsubhr.hurs.qn.v20181203	1	347.9 MiB

Table of contents

Feature highlight

Setup

Quickstart

Ensemble du MRCC5

Quelques centres de recherche climatiques développent un ensemble de simulation relié à leur(s) modèle(s).

- Pour faire des tests de sensibilité
- Pour évaluer des améliorations ou ajouts au modèle
- Pour diverses expériences spécifiques
- Pour offrir des nouvelles jeu de données

Ensemble du MRCC à Ouranos

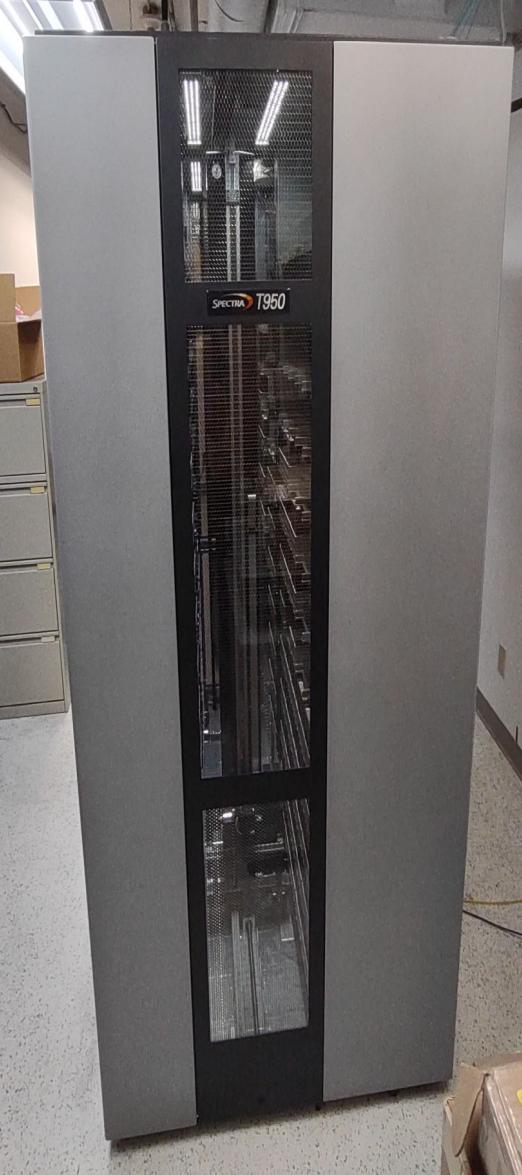
Le MRCC a une **résolution spatiale allant jusqu'à ~12 km et une résolution temporelle d'archivage jusqu'aux heures.**

Conserver une centaine de variables aux $3h$ sur une grille de 12 Km sur l'Amérique du Nord pour une simulation de 100 ans requiert environ **~32 To d'espace disque**.

- **MRCC5-CMIP6 : 2500 années simulées à date → ~800 To d'espace disque**

Ouranos a une système d'archivage créée uniquement pour les sortis du modèle MRCC

Système d'archivage à Ouranos



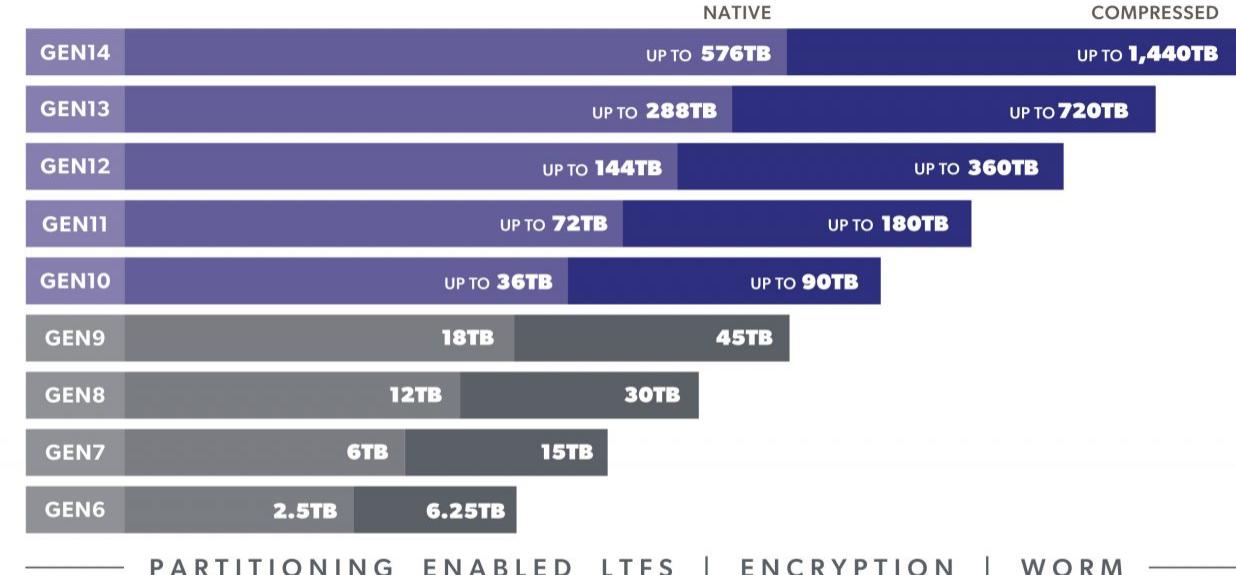
Crédit : Travis Logan et Pascal Bourgault, équipe PCDO

Système d'archivage à Ouranos - Cassettes LTO



LTO ULTRIUM ROADMAP

Addressing your storage needs



NOTE: Compressed capacities assume 2.5:1 compression (achieved with larger compression history buffer).

SOURCE: The LTO Program. The LTO Ultrium roadmap is subject to change without notice and represents goals and objectives only. Linear Tape-Open LTO, the LTO logo, Ultrium and the Ultrium logo are registered trademarks of Hewlett Packard Enterprise Company, International Business Machines Corporation and Quantum Corporation in the US and other countries. Please contact your supplier/manufacturer for more information.



**LTO : « Linear Tape-Open » peuvent contenir
jusqu'à 18 To par cassette présentement!**

Données climatiques : Acquisition, interprétation et manipulation

Modèles régionaux du climat - Conclusions

Forces

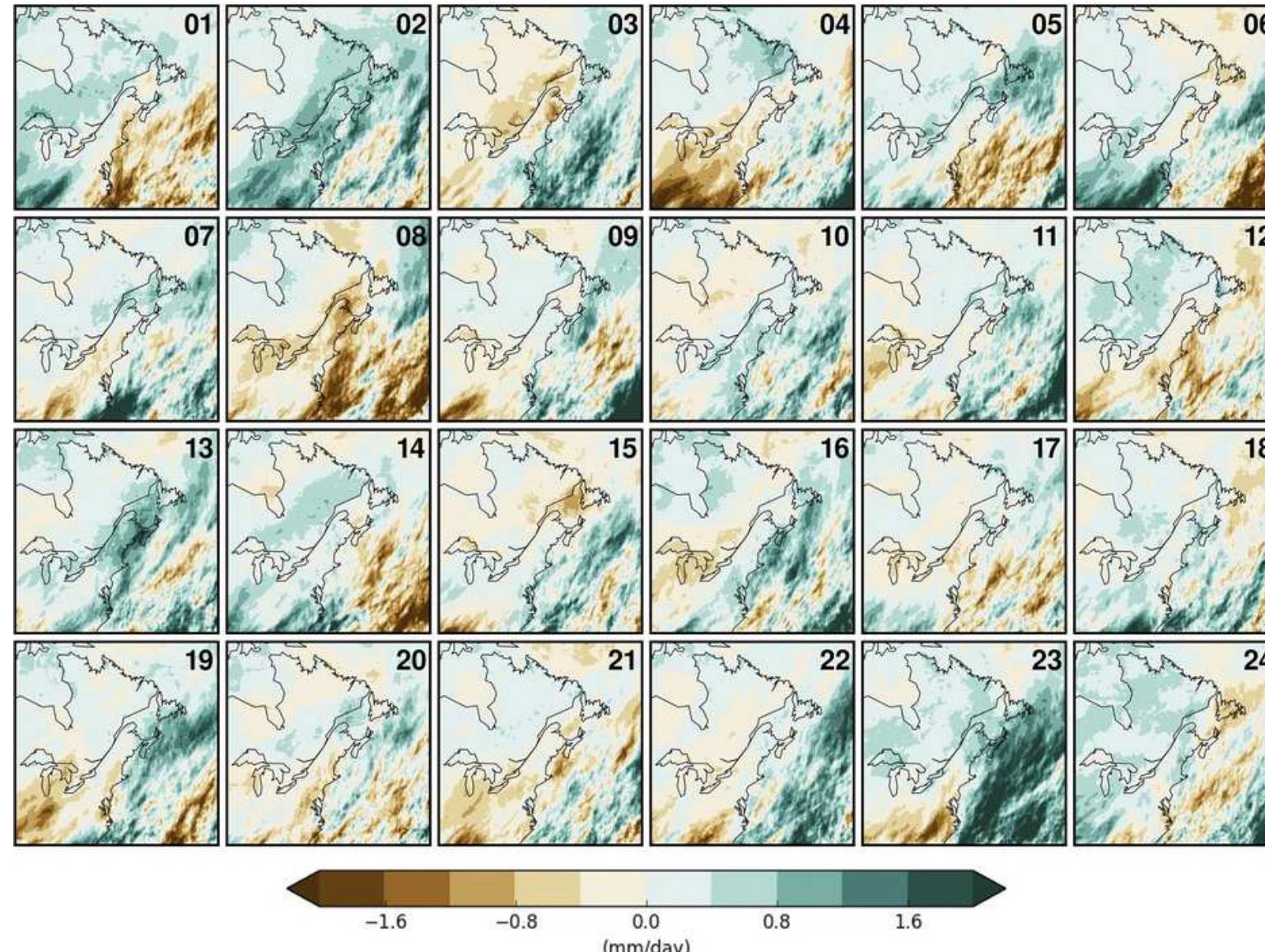
- Couverture spatiale à haute résolution sur un domaine choisi.
- Séries continues.
- Cohérence physique entre les variables.
- Permet de faire des projections dans le futur.

Faiblesses

- Plusieurs phénomènes encore paramétrés (p. ex. formes de précipitation spéciales).
- Les modèles ont des différences majeures entre eux.
- Multiplication des expériences requises pour couvrir l'incertitude des modèles pilotes.
- Coûts énormes en calcul et en stockage / archivage des données

Grands ensembles des modèles

École d'été en sciences du climat 2024



Grands ensembles des modèles

Les grands ensembles consistent à produire plusieurs dizaines de membres d'une même configuration d'un modèle.

Permet de mieux évaluer le rôle de la variabilité naturelle dans la chaîne d'incertitude de la modélisation climatique.

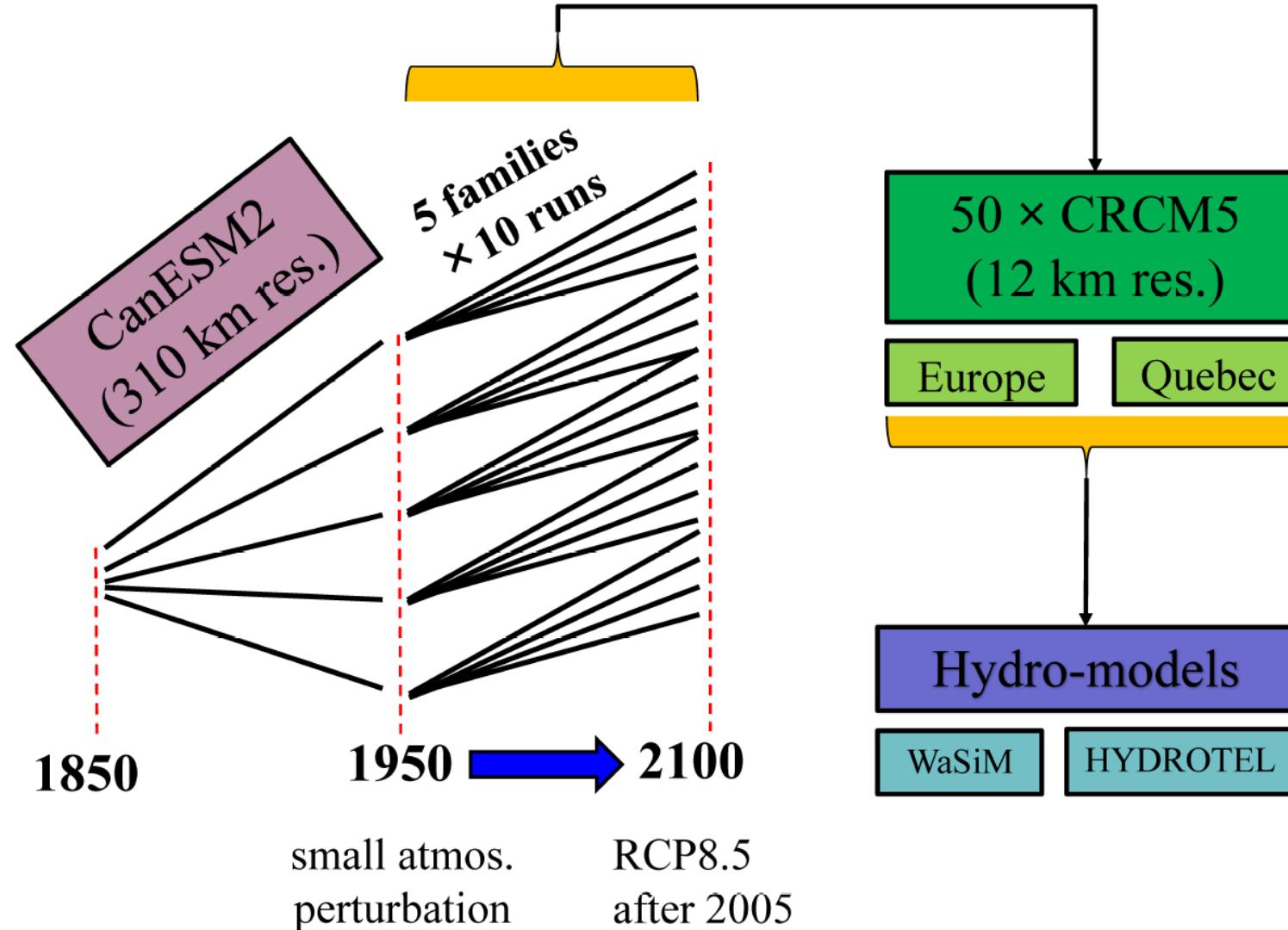
NCAR CESM1 LENS (*40 membres*)

CanESM / CanESM2 Large Ensemble (*50 membres*)

- **CanRCM4-LE → CanLEAD-LE (données post-traitées)**

Ensemble de Simulations « ClimEx »

École d'été en sciences du climat 2024



Source : Leduc et al. (2016) The ClimEx Project: Dynamical downscaling of a GCM large ensemble at very high resolution for Bavaria and Quebec. International Conference on Regional Climate-CORDEX, Stockholm, 17-20 May 2016. Poster.

Calendriers climatiques

Différents modèles climatiques utilisent différents calendriers :

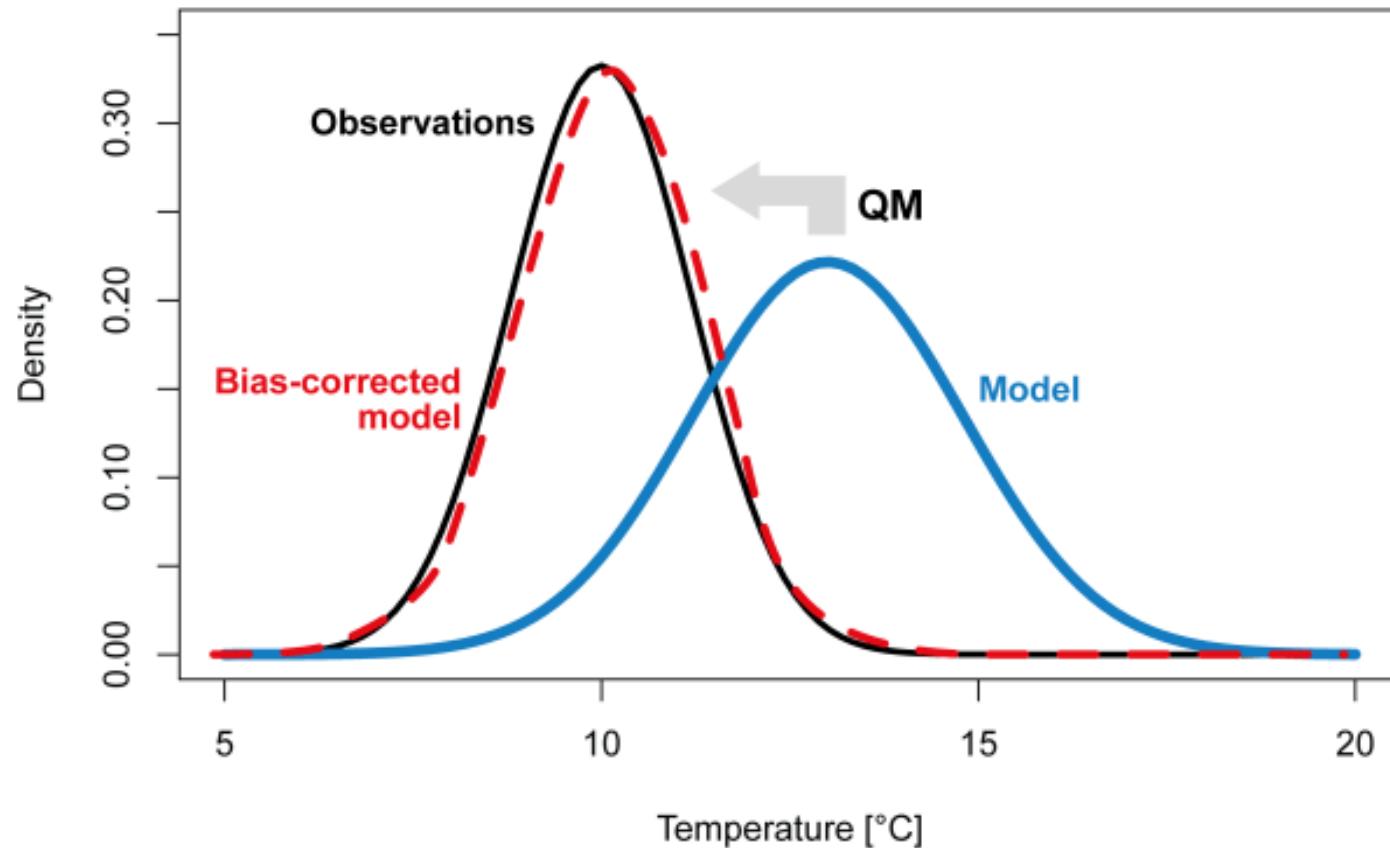
- « *standard* » (ancien nom : « *gregorian* »)
- « *proleptic_gregorian* »
- « *julian* »
- 360 jours (« *360_day* » : modèles **HadGEM** en particulier)
- 365 jours (« *no_leap* »)
- 366 jours (« *all_leap* »)

Pour savoir en plus :

- **Conventions CF** (<http://cfconventions.org/cf-conventions/cf-conventions#calendar>)

Données « post-traitées »

Sortis des modèles ajustées statistiquement pour enlever les biais de modèles et pour supplémenter les analyses du climat générales.



Source : Feigenwinter et al (2018) Exploring quantile mapping as a tool to produce user-tailored climate scenarios for Switzerland ([lien](#))

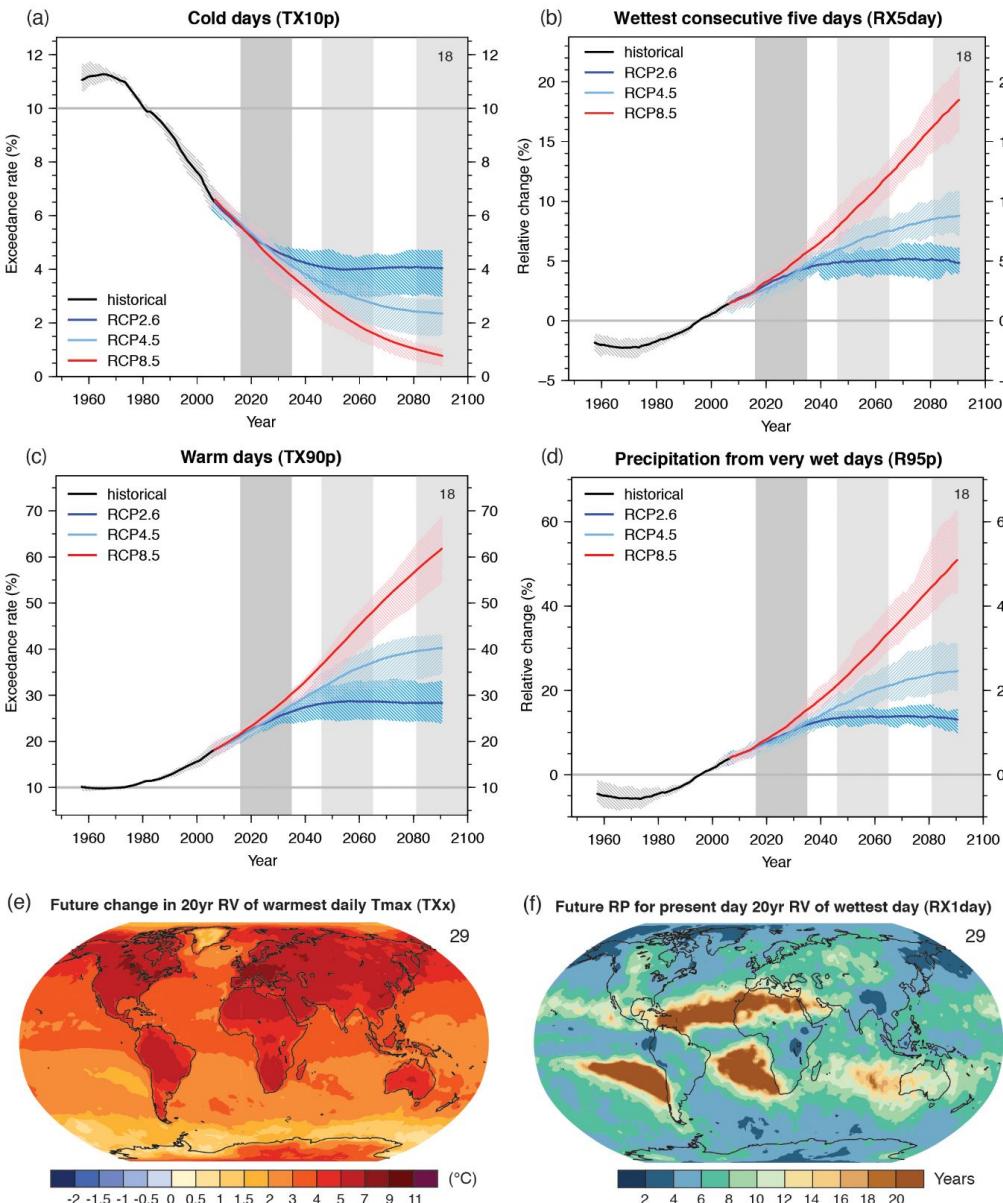
Données « post-traitées »

- Ouranos
 - **ESPO-G** (<https://doi.org/10.1038/s41597-023-02855-z>)
- Climate Impact Lab
 - **GDPCIR** (<https://doi.org/10.5194/gmd-17-191-2024>)
- NASA Center for Climate Simulation (NCCS)
 - **NEX-GDDP** (<https://doi.org/10.1038/s41597-022-01393-4>)

Pour en savoir plus

- **Maraun D. (2016) Correcting Climate Change Simulations - a Critical Review**
(<https://doi.org/10.1007/s40641-016-0050-x>)

6. Données « dérivées »



Données « dérivées »

Transformation de données bruts vers des indices qui décrit des conditions climatiques dans une manière plus accessible / décrivant / pertinente

Aide en faisant les comparaisons des conditions climatiques entre :

- Sources des données; Horizons temporelles; Régions spatiales; etc...

Pour mieux communiquer les informations avec les grands publics

Données « dérivées »

Plusieurs indicateurs climatiques sont calculés à partir de données observées et publiés comme jeu des données

Ressources Naturelles Canada calcule des indicateurs bioclimatiques (p. ex. début de la saison de croissance)

Standard pour plusieurs indices climatiques : European Climate Assessment & Dataset (ECAD)

Pour en savoir plus :

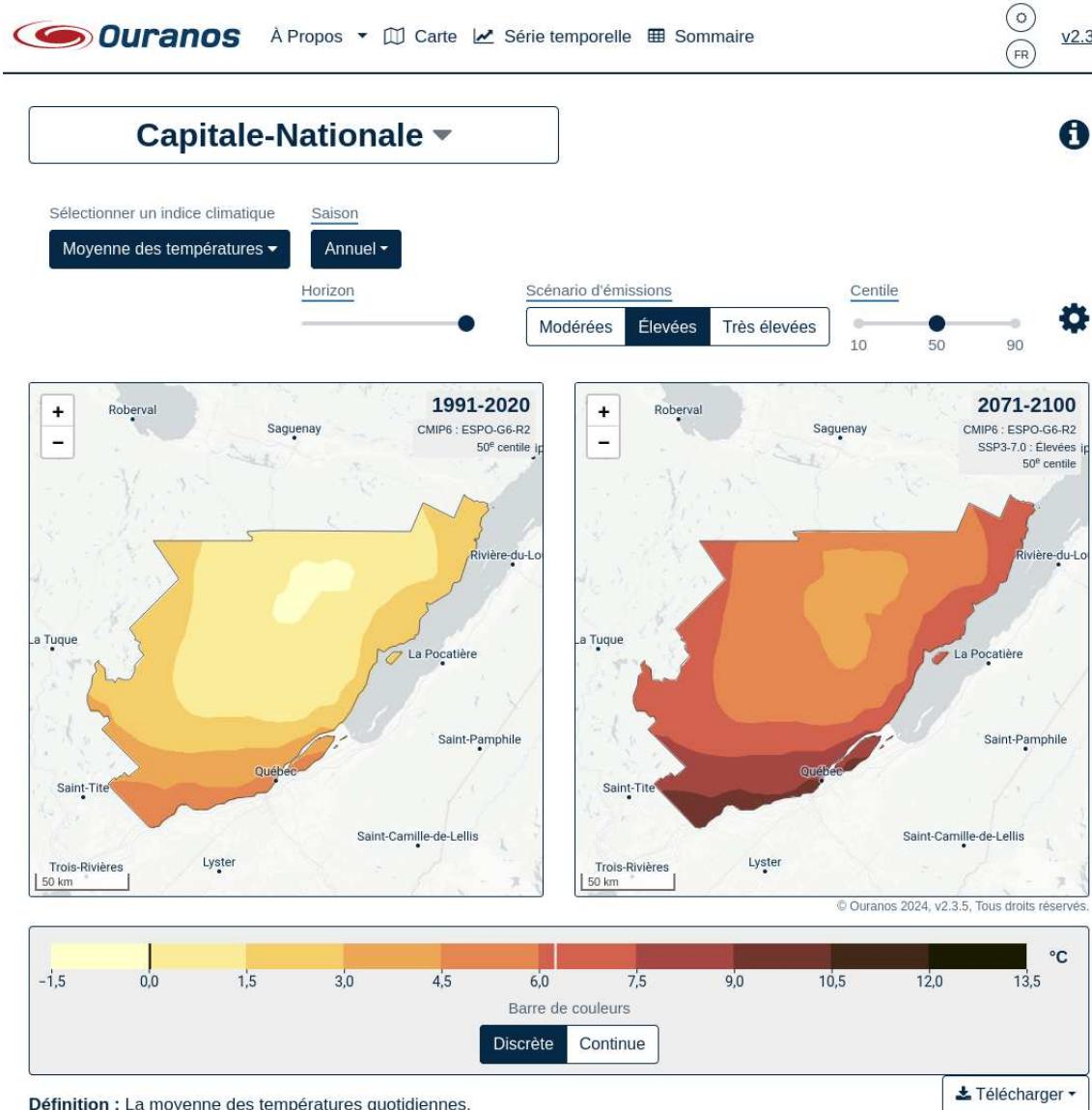
- **ETCCDI Comparaison -**

<https://www.ecad.eu/documents/ETCCDMIndicesComparison.pdf>

Données « dérivées » par Ouranos

Les Portraits climatiques

- Données CMIP5 ou CMIP6 post-traitées
- **25 indices climatiques** pré-calculées
- Domaine de Québec* (et sous-domaines)
- Sommaires disponibles pour téléchargement
- Données sources accessibles par OPeNDAP / PAVICS, etc..



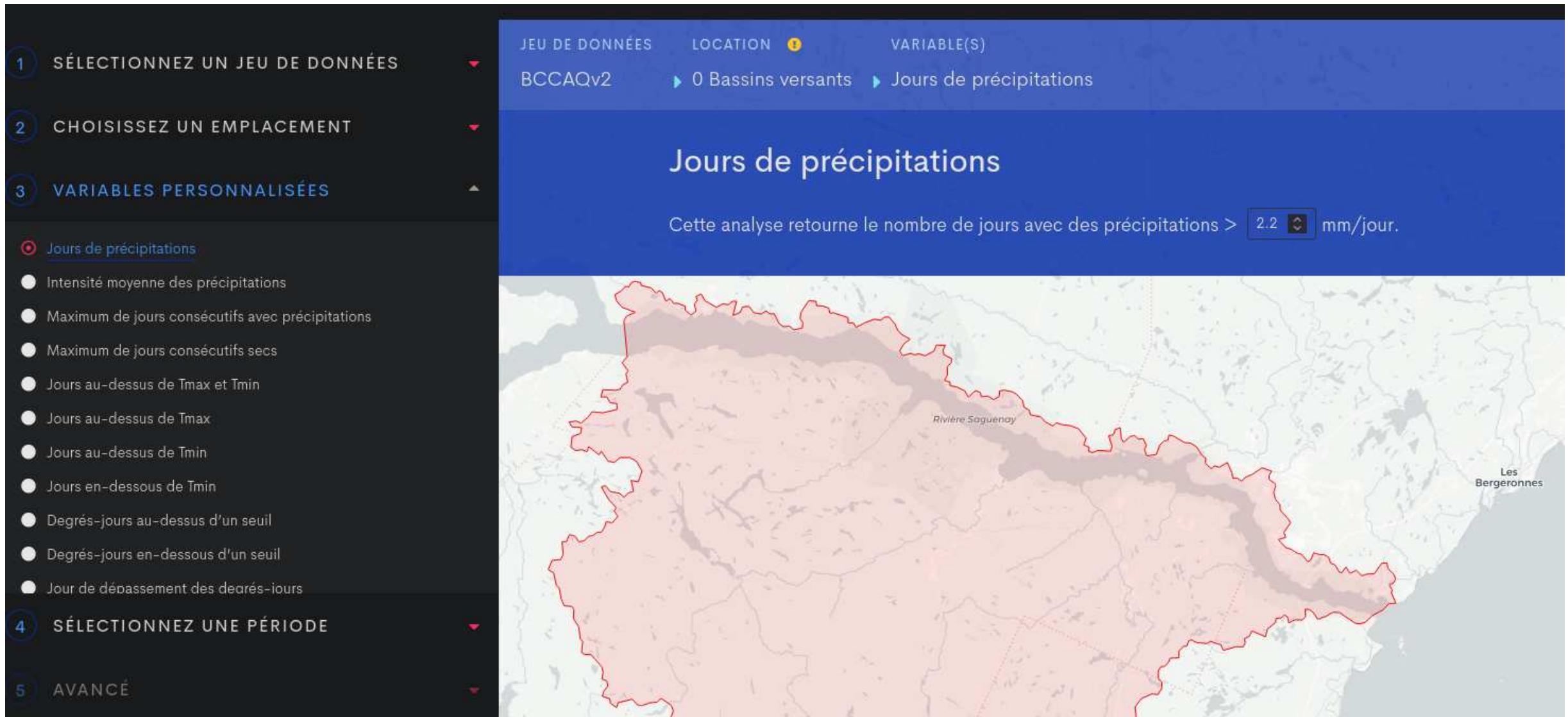
Définition : La moyenne des températures quotidiennes.
Voir la documentation des données.

Données dérivées « sur mesure »

Calculs des indices par requête internet / en-ligne :

- Microsoft Planetary Computer (Global) : <https://planetarycomputer.microsoft.com/>
- Données Climatiques (Canada) : <https://donneesclimatiques.ca/>
- Copernicus Data Store (Global / Europe) : <https://cds.climate.copernicus.eu/>
- PAVICS (Amérique du Nord) : <https://pavics.ouranos.ca/>

Données dérivées sur site-web - [DonneesClimatiques.ca](https://donneesclimatiques.ca/)



Données dérivées - Conclusions

Forces

- Transformation des données vers les informations utiles pour montrer des conditions / changements au grand public
- Facilite la comparaison entre sources / formes différents des données
- Moins d'espace disque nécessaire pour les stocker

Faiblesses

- « Filtrage » des données - enlève plusieurs détails
- Seuils et paramétrages sont très spécifiques de cas à cas – besoin de refaire des calculs
- Outils et algorithmes de calcul complexe nécessaires pour traiter des données énormes
- ~~Produits « one-shot » (Faite une fois) – besoin de refaire des calculs souvent~~

7. Formats habituels des données climatiques et outils d'analyses

Format des données de stations

Format des données de stations

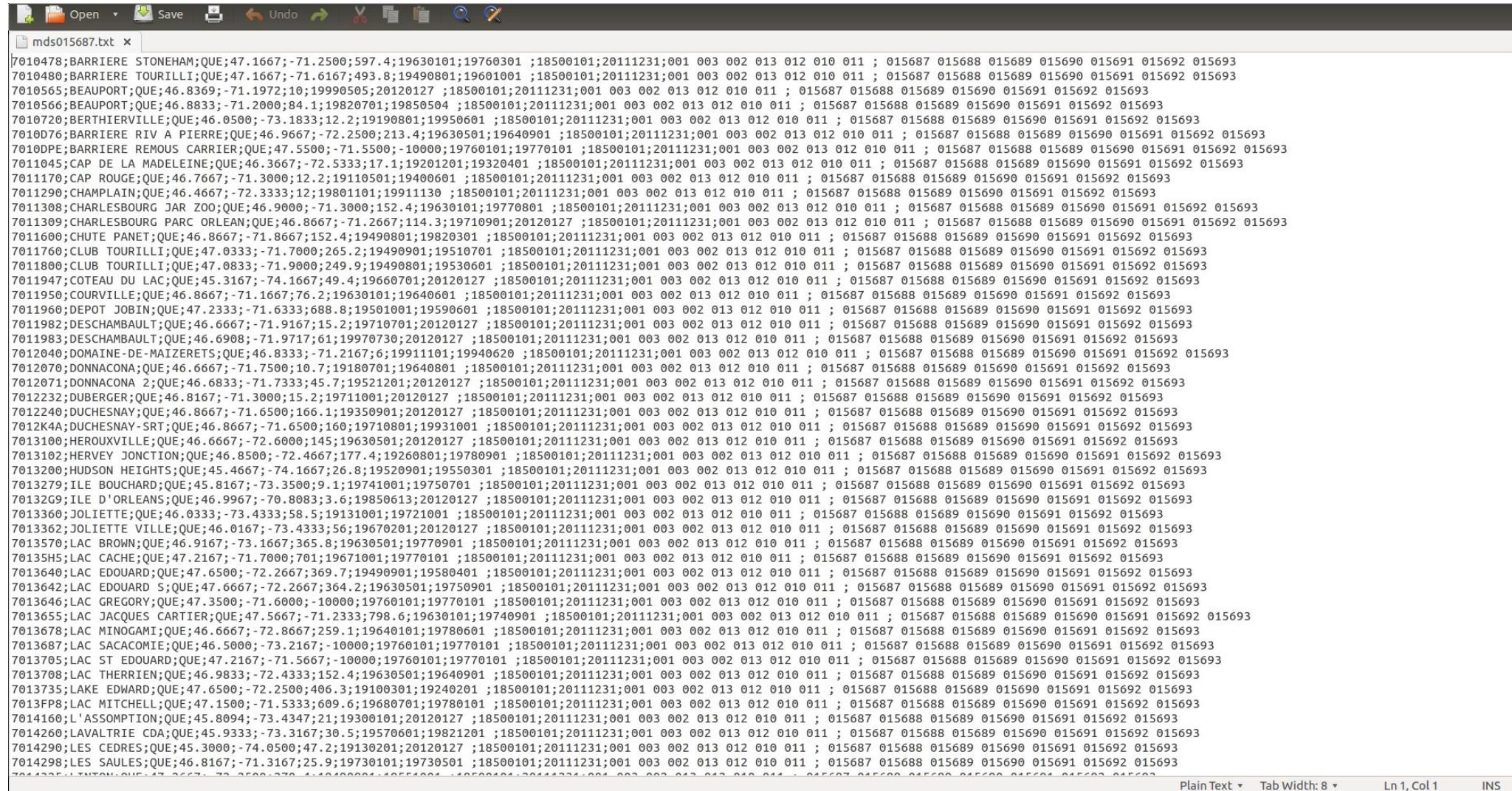
L'information sur les variables, les unités et les “flags” se trouve dans une documentation à part.

Pour les données de précipitations sur le Canada, la base de données contient ~10000 fichiers texte pour un total de ~550 Mo.

Ce format des données est utilisé à l'interne d'ECCC, alors pas facilement accessible aux grands publics.

Format des données de stations

École d'été en sciences du climat 2024



The screenshot shows a Windows-style text editor window with a dark theme. The title bar reads "mds015687.txt x". The menu bar includes "Open", "Save", "Print", "Undo", "Redo", "Find", "Replace", and "Help". The main area contains a large amount of text data, which is a CSV file of climate station records. The columns represent various meteorological parameters and station identifiers. The data starts with:

Station ID	Station Name	Province	Latitude	Longitude	Altitude	Start Date	End Date	Parameter	Value
7010478	BARRIERE STONEHAM	QUE	47.1667	-71.2500	597.4	19630101	19760301	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7010480	BARRIERE TOURILLI	QUE	47.1667	-71.6167	493.8	19490801	19601001	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7010565	BEAUPORT	QUE	46.8369	-71.1972	10	19990505	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7010566	BEAUPORT	QUE	46.8833	-71.2000	84.1	19820701	19850504	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7010720	BERTHIERVILLE	QUE	46.0500	-73.1833	12.2	19190801	19950601	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7010D76	BARRIERE RIV A PIERRE	QUE	46.9667	-72.2500	213.4	19630501	19640901	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7010DPE	BARRIERE REMOUS CARRIER	QUE	47.5500	-71.5500	-10000	19760101	19770101	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011045	CAP DE LA MADELEINE	QUE	46.3667	-72.5333	17.1	19201201	19320401	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011170	CAP ROUGE	QUE	46.7667	-71.3000	12.2	19110501	19400601	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011290	CHAMPLAIN	QUE	46.4667	-72.3333	12	19801101	19911130	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011308	CHARLESBOURG JAR ZOO	QUE	46.9000	-71.3000	152.4	19630101	19770801	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011309	CHARLESBOURG PARC ORLEAN	QUE	46.8667	-71.2667	114.3	19710901	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011600	CHUTE PANET	QUE	46.8667	-71.8667	152.4	19490801	19820301	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011760	CLUB TOURILLI	QUE	47.0333	-71.7000	265.2	19490901	19510701	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011800	CLUB TOURILLI	QUE	47.0833	-71.9000	249.9	19490801	19530601	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011947	COTEAU DU LAC	QUE	45.3167	-74.1667	49.4	19660701	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011950	COURVILLE	QUE	46.8667	-71.1667	76.2	19630101	19640601	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011960	DEPOT JOBIN	QUE	47.2333	-71.6333	688.8	19501001	19590601	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011982	DESCHAMBault	QUE	46.6667	-71.9167	15.2	19710701	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7011983	DESCHAMBault	QUE	46.6908	-71.9717	61	19970730	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7012040	DOMAINE-DE-MAIZERETS	QUE	46.8333	-71.2167	6	19911101	19940620	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7012070	DONNACONA	QUE	46.6667	-71.7500	10.7	19180701	19640801	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7012071	DONNACONA 2	QUE	46.6833	-71.7333	45.7	19521201	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7012232	DUBERGER	QUE	46.8167	-71.3000	15.2	19711001	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7012240	DUCHESNAY	QUE	46.8667	-71.6500	166.1	19350901	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7012K4A	DUCHESNAY-SRT	QUE	46.8667	-71.6500	160	19710801	19931001	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013100	HEROUXVILLE	QUE	46.6667	-72.6000	145	19630501	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013102	HERVEY JONCTION	QUE	46.8500	-72.4667	177.4	19260801	19780901	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013200	HUDSON HEIGHTS	QUE	45.4667	-74.1667	26.8	19520901	19550301	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013279	ILE BOUCHARD	QUE	45.8167	-73.3500	9.1	19741001	19750701	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
701329	ILE D'ORLEANS	QUE	46.9967	-70.8083	3.6	19850613	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013360	JOLIETTE	QUE	46.0333	-73.4333	58.5	19131001	19721001	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013362	JOLIETTE VILLE	QUE	46.0167	-73.4333	56	19670201	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013570	LAC BROWN	QUE	46.9167	-73.1667	365.8	19630501	19770901	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
70135HS	LAC CACHE	QUE	47.2167	-71.7000	701	19671001	19770101	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013640	LAC EDOUARD	QUE	47.6500	-72.2667	369.7	19490901	19580401	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013642	LAC EDOUARD S	QUE	47.6667	-72.2667	364.2	19630501	19750901	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013646	LAC GREGORY	QUE	47.3500	-71.6000	-10000	19760101	19770101	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013655	LAC JACQUES CARTIER	QUE	47.5667	-71.2333	798.6	19630101	19740901	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013678	LAC MINOGAMI	QUE	46.6667	-72.8667	259.1	19640101	19780601	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013687	LAC SACACOMIE	QUE	46.5000	-73.2167	-10000	19760101	19770101	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013705	LAC ST EDOUARD	QUE	47.2167	-71.5667	-10000	19760101	19770101	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013708	LAC THERRIEN	QUE	46.9833	-72.4333	152.4	19630501	19640901	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013735	LAKE EDWARD	QUE	47.6500	-72.2500	406.3	19100301	19240201	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7013FP8	LAC MITCHELL	QUE	47.1500	-71.5333	609.6	19680701	19780101	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7014160	L'ASSOMPTION	QUE	45.8094	-73.4347	21	19300101	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7014260	LAVALTRIE CDA	QUE	45.9333	-73.3167	30.5	19570601	19821201	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7014290	LES CEDRES	QUE	45.3000	-74.0500	47.2	19130201	20120127	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693
7014298	LES SAULES	QUE	46.8167	-71.3167	25.9	19730101	19730501	18500101	20111231;001 003 002 013 012 010 011 ; 015687 015688 015689 015690 015691 015692 015693

Formats « maison »

Données de radiosonde dans le format « FSL »

LEGEND

LINTYP: type of identification line

254 = indicates a new sounding in the output file

1 = station identification line

2 = sounding checks line

3 = station identifier and other indicators line

4 = mandatory level

5 = significant level

6 = wind level (PPBB) (GTS or merged data)

7 = tropopause level (GTS or merged data)

8 = maximum wind level (GTS or merged data)

9 = surface level

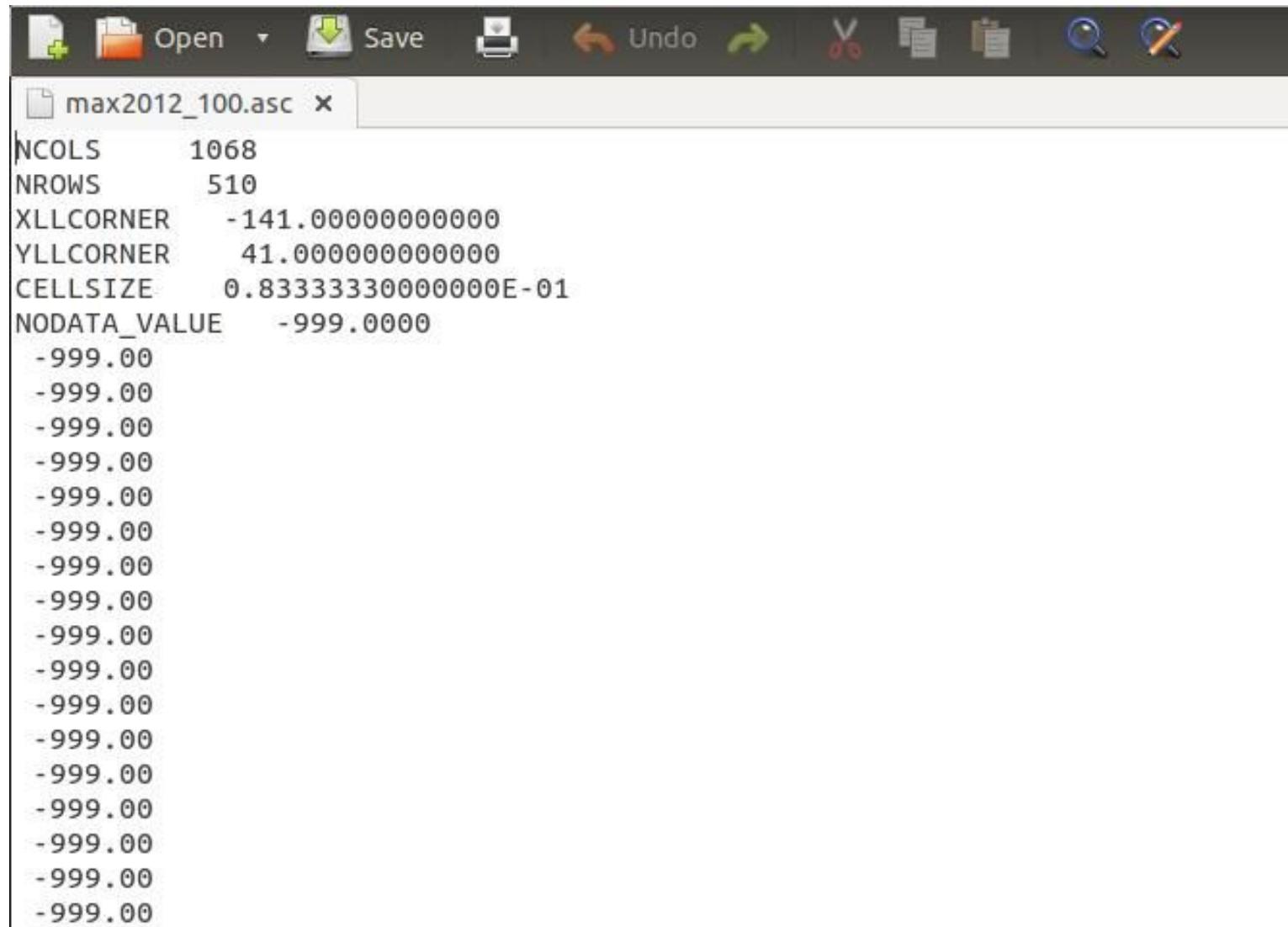
HOUR: time of report in UTC

LAT: latitude in degrees and hundredths

LON: longitude in degrees and hundredths

254	0	24	MAY	2017			
1	4734	71722	46.38N	75.97W	170	2318	
2	70	2190	2060	140	99999	3	
3		YMW			99999	kt	
9	9860	170	176	86	0	0	
4	10000	63	99999	99999	99999	99999	
5	9790	231	182	72	99999	99999	
6	9709	304	99999	99999	315	1	
6	9377	609	99999	99999	240	4	
4	9250	728	142	52	275	4	
6	9046	914	99999	99999	280	8	
6	8721	1219	99999	99999	280	8	
4	8500	1433	78	28	285	9	
6	8094	1828	99999	99999	285	14	
6	7793	2133	99999	99999	275	15	
6	7504	2438	99999	99999	255	21	
6	7225	2743	99999	99999	260	16	
4	7000	2998	-49	-64	260	18	
6	6955	3048	99999	99999	260	18	
5	6820	3201	-67	-72	99999	99999	
5	6760	3270	-73	-96	99999	99999	
5	6700	3340	-47	-187	99999	99999	
6	6435	3657	99999	99999	290	24	
5	6370	3736	-49	-249	99999	99999	
6	5949	4267	99999	99999	285	31	
5	5940	4279	-89	-249	99999	99999	
5	5900	4331	-85	-265	99999	99999	
6	5718	4572	99999	99999	280	33	
5	5500	4871	-107	-377	99999	99999	
6	5496	4876	99999	99999	290	31	
5	5390	5026	-117	-417	99999	99999	
6	5072	5486	99999	99999	265	30	
5	5020	5564	-161	-341	99999	99999	
4	5000	5600	-163	-313	270	27	

Format ArcInfo (Grille ASCII)



Format ArcInfo (Grille ASCII)

Pour une base de données comme les données journalières interpolées de Ressource Naturelle Canada sur une grille à 10 km (1068x510) :

Un fichier par jour de ~5 Mo :

- Pour 62 années → ~22 000 fichiers pour un total de *~110 Go par variable*

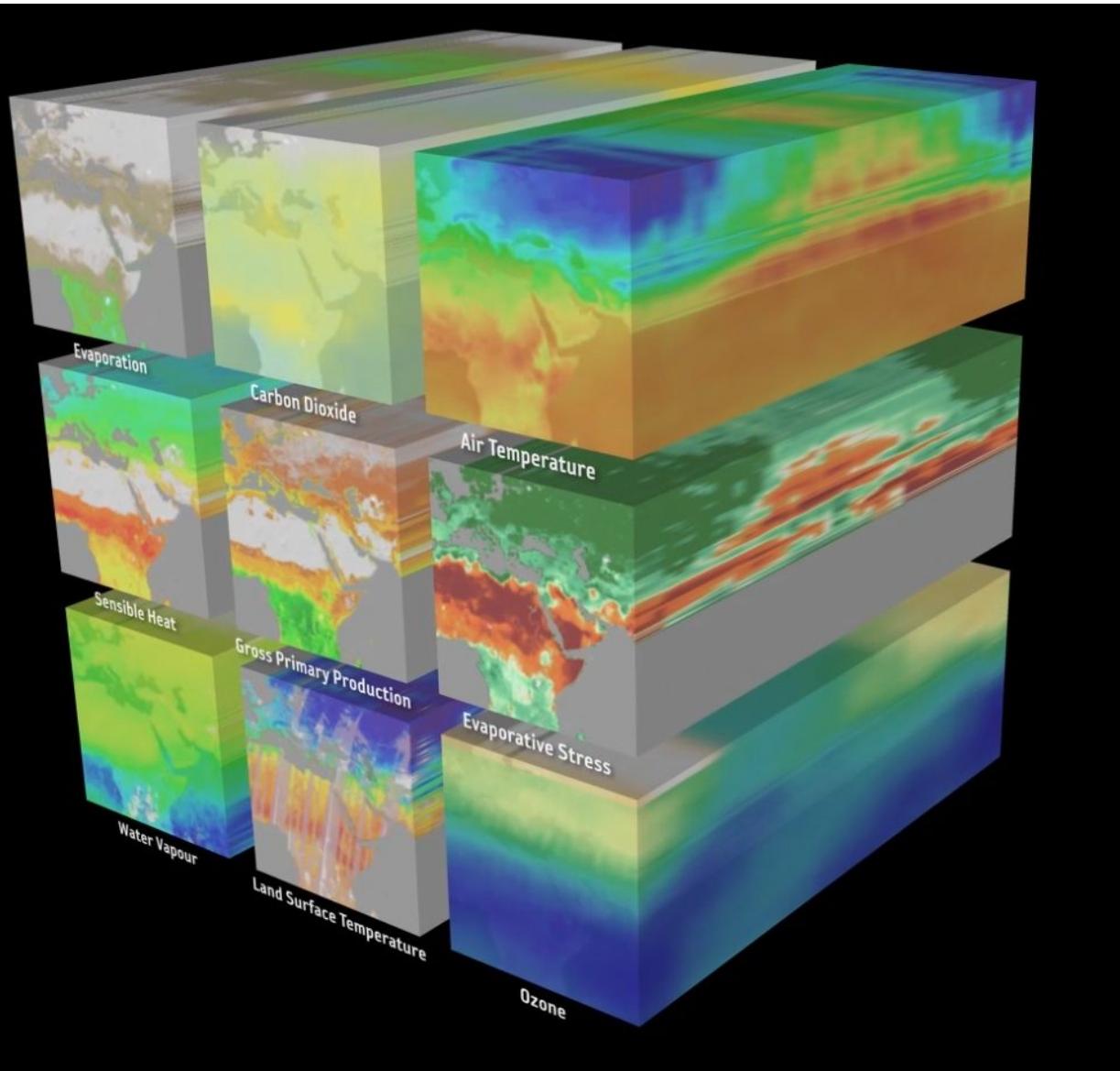
Simple, stable, mais...

- **Aucun compression (gourmand pour l'espace disque)**
- **Aucun métadonnées**

À côté – Les métadonnées

Des données qui décrit / défini les données

- Historique / changements / provenance des opérations
- Créeateur, origine, date de création, versions des outils, modèles et grilles utilisées, standard de métadonnées, unités des valeurs, etc.
- Calendrier(s), fréquences de temps
- Codes de qualité de données



Format « Datacube »

Données avec plusieurs dimensions
(**« N-dimensional matrices »**)

- latitude, longitude, élévation, temps, autres coordonnées, variables, etc...
- Données satellitaires et climatiques sur grille
- **Compression !**
- **Métadonnées intégrés !**

Format « NetCDF »

Ce format permet d'enregistrer des données multidimensionnelles ainsi que les métadonnées qui s'y rattachent.

Algorithmes de compression disponibles lors de l'écriture du fichier.

Développé par UCAR (pas UQAR, malheureusement) et le National Science Foundation (NSF) aux États-Unis

Pour en savoir plus :

- Site web du projet (<https://www.unidata.ucar.edu/software/netcdf>)
- Entrée sur Wikipédia (<https://en.wikipedia.org/wiki/NetCDF>)

Définition du format « NetCDF »

École d'été en sciences du climat 2024

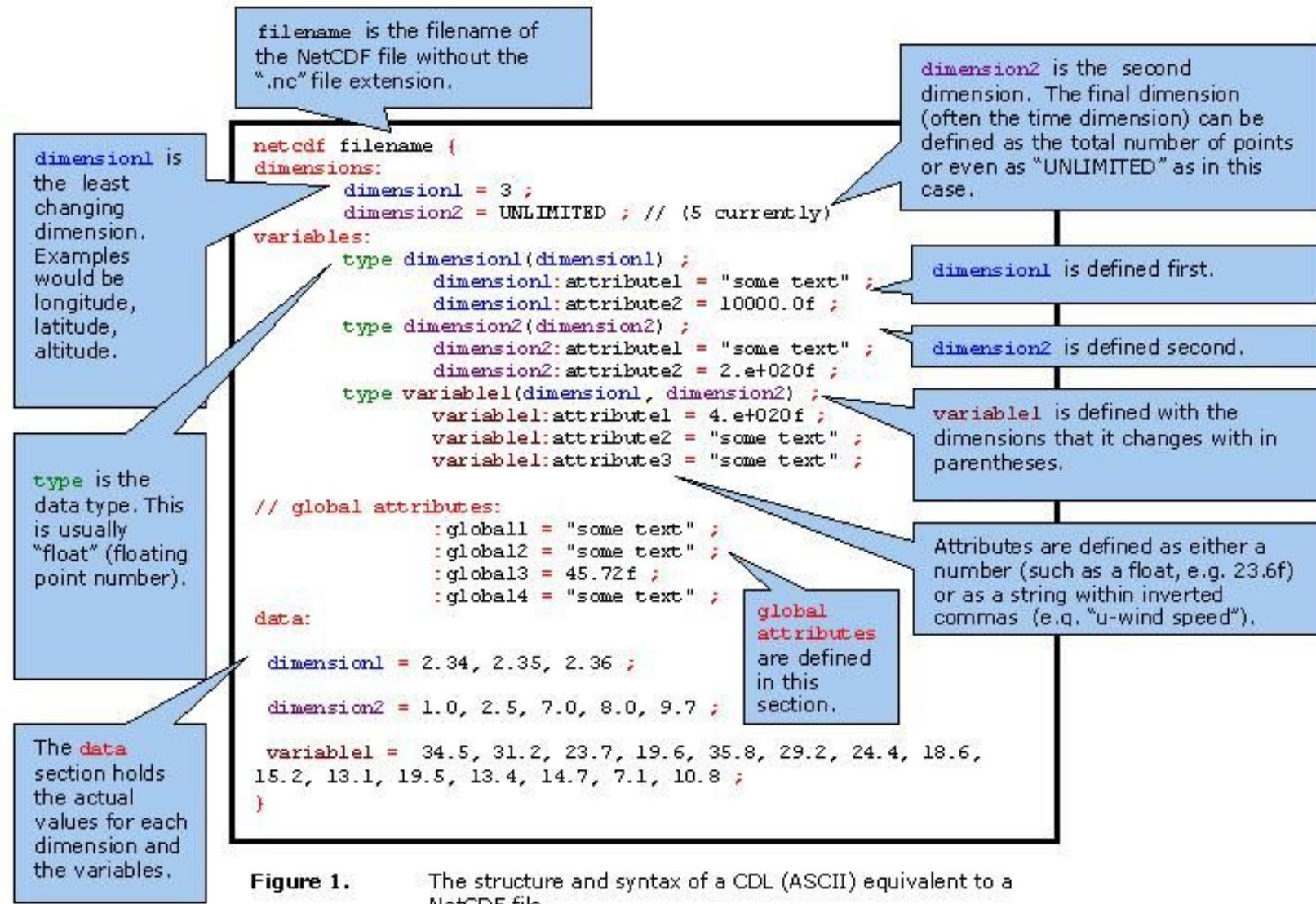


Figure 1. The structure and syntax of a CDL (ASCII) equivalent to a NetCDF file.

Format « NetCDF »

La communauté des sciences du climat a établi une convention à suivre pour bien écrire des données climatiques.

Standards pour les données parmi les climatologues

- **Conventions CF** : <http://cfconventions.org>
- **COARDS (ancien)** :
<https://ferret.pmel.noaa.gov/Ferret/documentation/coards-netcdf-conventions>

Les variables ont des noms standardisés.

La définition des unités temporelles et de la grille géospatiales est standardisée.

À côté - Standardisation des données

Outils de **conversion / gestion / conformité** des données et métadonnées en NetCDF

- En développement pour plusieurs années
- Objectif : en opération pendant hiver 2024

Pour en savoir plus :

- GitHub : [Ouranosinc/Miranda](#)



The screenshot shows the GitHub repository page for 'Miranda'. At the top, there's a large blue globe icon with white latitude and longitude lines. Below it, the word 'Miranda' is written in a sans-serif font. The page has a dark background with light-colored text and buttons. At the top right, there are buttons for 'Versions' (pypi v0.5.0, python 3.7 | 3.8 | 3.9 | 3.10 | 3.11), 'Documentation and Support' (docs failing), and three tabs: 'README' (selected), 'Apache-2.0 license', and 'Security'. Below these are two sections: 'Coding Standards' (code style black, Ruff, pre-commit.ci passed) and 'Development Status' (repo status Active, miranda Testing Suite passing, coverage 18%). A brief description follows: 'Python utilities for climate data collection, conversion, and management.' Two bullet points are listed under 'Features': 'Free software: Apache Software License 2.0' and 'Documentation: <https://miranda.readthedocs.io>'. The page then lists several sections with bullet points: 'Data collection functions for climate and forecast data hosted at:' (ECMWF (ERA5, ERA5-Land, TIGGE), ECCC (Canada) (Monthly Climate Summaries, ECCC GEOAPI - In development), NCAR (CORDEX-NA on AWS)); 'Data conversion for [Climate and Forecasting \(CF\)](#) Variable and Metadata compliance:' (ECMWF (ERA5, ERA5-Land, TIGGE - In Development), ECCC (Canada) (Flat File Observations, Monthly Climate Summaries, Adjusted and Homogenized Climate Data, ECCC GEOAPI - In Development), MELCC (Québec), Hydro-Québec (In Development)); and 'Database structuring and facets validation:' (Simulations: WCRP (CMIP5, CMIP6, CMIP5-CORDEX, CORDEX-ADJUST, ISIMIP, etc.)).

Formats infonuagiques

Services propriétaires de calcul et entreposage :

- « Amazon AWS » (S3)
- « Google Big Query » / « Earth Engine » (Cloud)
- « Apache Spark » (Hadoop)



Protocole ouverte pour servir des données NetCDF et formats similaires en-ligne : « [OPeNDAP](#) »

« Datacube » similaire à NetCDF optimisé pour la téléchargement et lecture en-ligne :
« [Zarr](#) »



Autres formats traditionnels

GRIB « Gridded Binary » : <http://www.grib.us/>

CCC, RPN : formats de données sur grille d'Environnement Canada.

Vectorielle : définition des points, des lignes ou des contours de polygones représentant des phénomènes ou objets spatiaux.

- *Shapefile, GeoJSON, GML, KML ...*
- *PostgreSQL/PostGIS, SQLite, GeoPackage* →



Plateformes et langages de programmation

École d'été en sciences du climat 2024



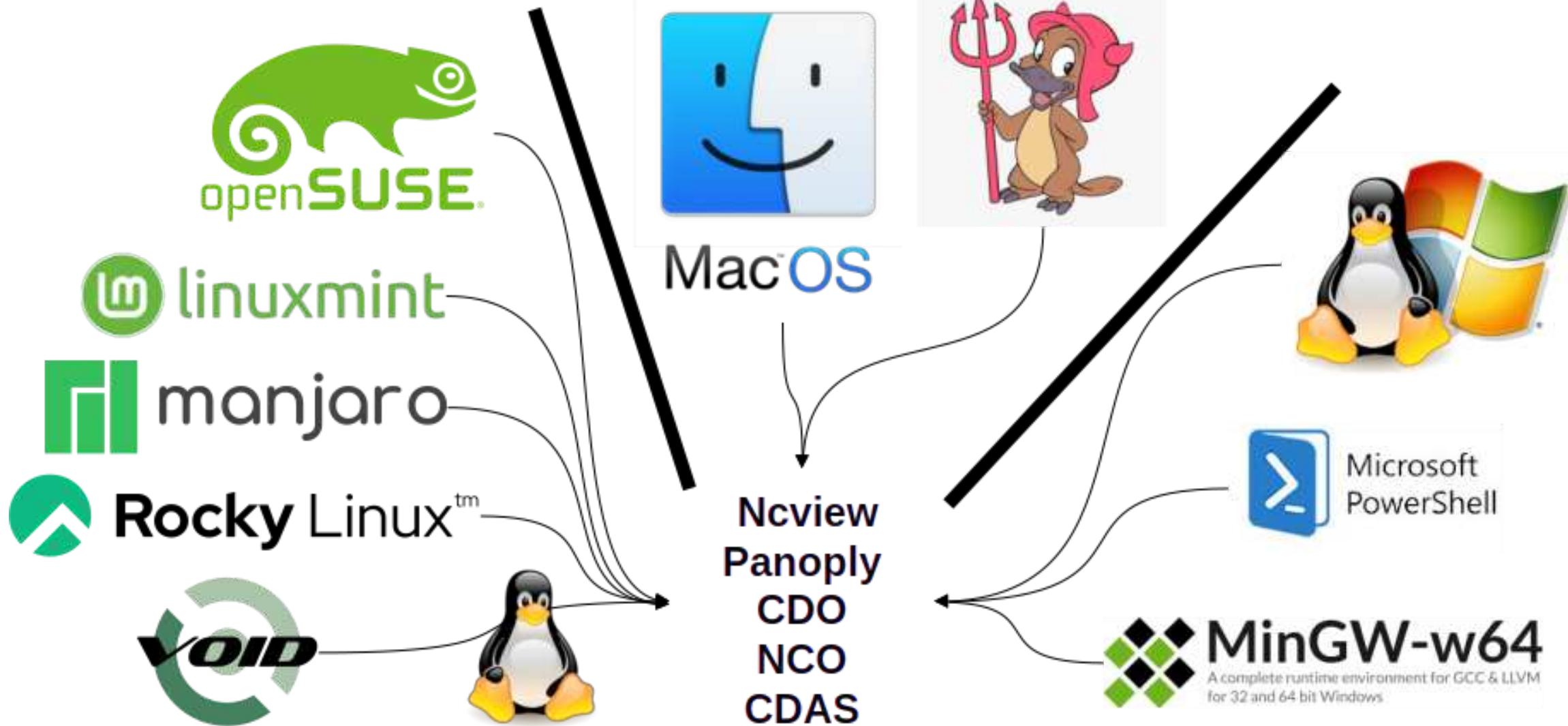
ArcGIS



Interactive/GNU
Data
Language

Outils pour systèmes d'opérations

École d'été en sciences du climat 2024

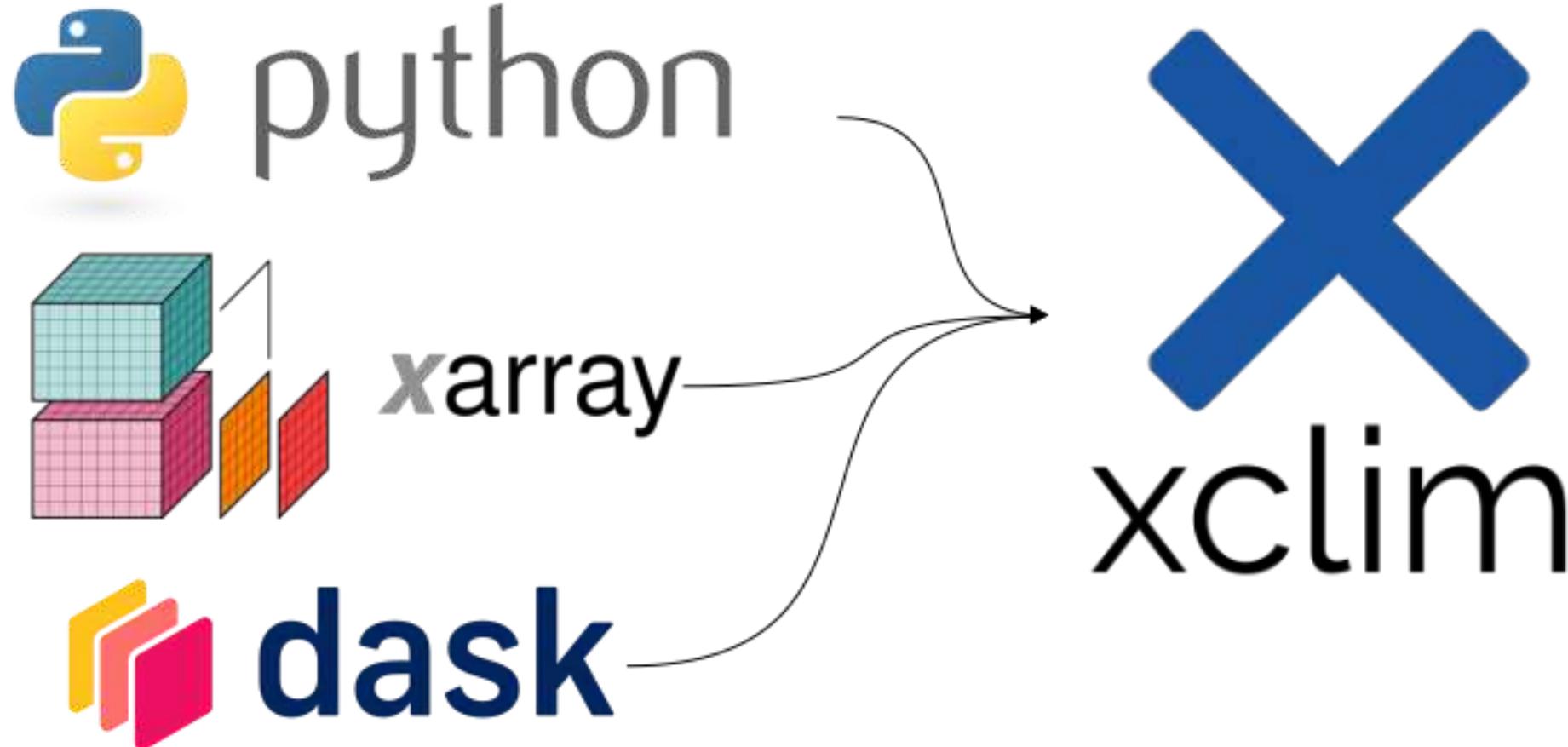


Exemple : `ncvue`

- Un nouveau implémentation de `ncview`

Outils d'analyses créés par Ouranos

École d'été en sciences du climat 2024



Source : Ouranosinc/xclim (<https://github.com/Ouranosinc/xclim>)

Présentation à PyCon Global 2023 (<https://www.youtube.com/watch?v=3zS3IQJ6pb4>) (25 mins)

Bourgault et al. (2023). xclim: xarray-based climate data analytics (<https://joss.theoj.org/papers/10.21105/joss.05415>)

Comment choisir ses données ?

Qu'est-ce qui est **disponible** (ou **accessible**) à la résolution spatiale et temporelle nécessaire ?

Est-ce que le jeu de données **représente bien** le processus à étudier ?

Lors de la validation, est-ce que **l'incertitude** sur les données observées et l'aspect de **la représentativité spatiale et temporelle** est pris en compte ?

Faites-vous une **étude d'impact** ou un projet d'**analyse de processus** ?

8. Les services climatiques

Pont entre la science du climat et la prise de décisions en matière d'adaptation aux changements climatiques.

Les groupes **Scénarios et services climatiques (SSC)** et **Plateformes climatiques, données et opérations (PCDO)** à Ouranos offre des produits synthétisés ou pré-traités, et aussi les rapports informationnels aux partenaires qui n'ont pas l'expertise ou la capacité pour traiter autant de données.

Ouranos offre aussi une plateforme de calcul intégré avec les données climatiques.

Plateforme de calcul « PAVICS »

PAVICS est un laboratoire virtuel facilitant l'analyse de données climatiques. Il permet d'accéder à plusieurs collections de données d'observations, de projections climatiques et de réanalyse. Il offre aussi un environnement de programmation Python permettant d'analyser ces données sans avoir à les télécharger. Cet environnement de travail est constamment mis à jour avec les librairies les plus performantes pour l'analyse de données climatiques. De plus, Ouranos assure un contrôle de qualité sur les données offertes et leurs métadonnées.

- Fichiers et données en format netCDF
- Métadonnées suivant la convention CF
- Environnement d'analyse parallèle (xarray + dask)

Nouvelles

Pour enregistrer : <https://pavics.ouranos.ca/>

Accéder aux données

PAVICS fournit un accès aux données via un serveur THREDDS dédié, donnant accès à une variété croissante de données climatiques et météorologiques, notamment

- Projections futures des changements climatiques
- Données d'observation (aux stations et sur grille)
 - Données de réanalyse
 - Données de prévisions météorologiques

Analyser les données climatiques

PAVICS fournit des méthodes de recherche et d'analyse des données climatiques, offrant des outils qui couvrent une gamme d'étapes courantes d'analyse des données, notamment

- Recherche et accès aux données
- Découpage spatio-temporel
- Calculs d'indices climatiques
 - Statistiques d'ensemble
 - Visualisation de données

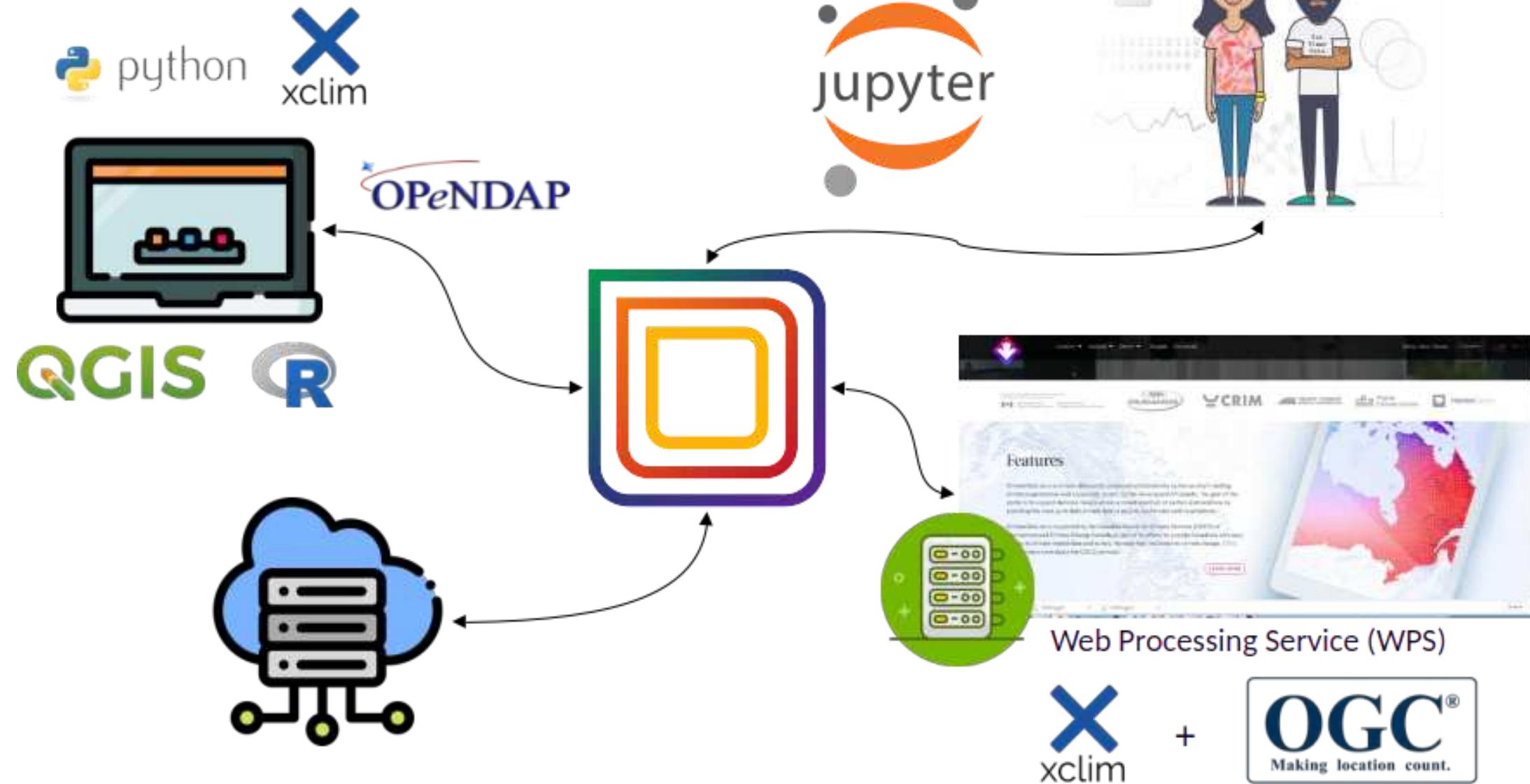
Exécuter des modèles hydrologiques

PAVICS offre une gamme d'outils pour faciliter l'analyse des impacts des changements climatiques sur l'hydrologie. Elle s'appuie pour cela sur Raven, un cadre de modélisation hydrologique qui permet d'assembler différents processus hydrologiques pour créer des modèles sur mesure ou émuler des modèles existants. Ces modèles doivent minimalement être forcés par des séries journalières de températures et de précipitations, et retournent des séries de débits en rivière et les variables d'état du modèle, comme la neige ou l'eau dans le sol.

- Simuler les débits en rivière par quatre modèles globaux: HMETS, GR4J-CemaNeige, HBV-EC et MOHYSE
 - Faire de l'analyse fréquentiel sur les séries de débits
 - Extraire les propriétés physiographiques d'un bassin versant
 - Estimer des paramètres d'un modèle en bassin non-jauge
 - Piloter un modèle hydrologique par les prévisions d'ECCC

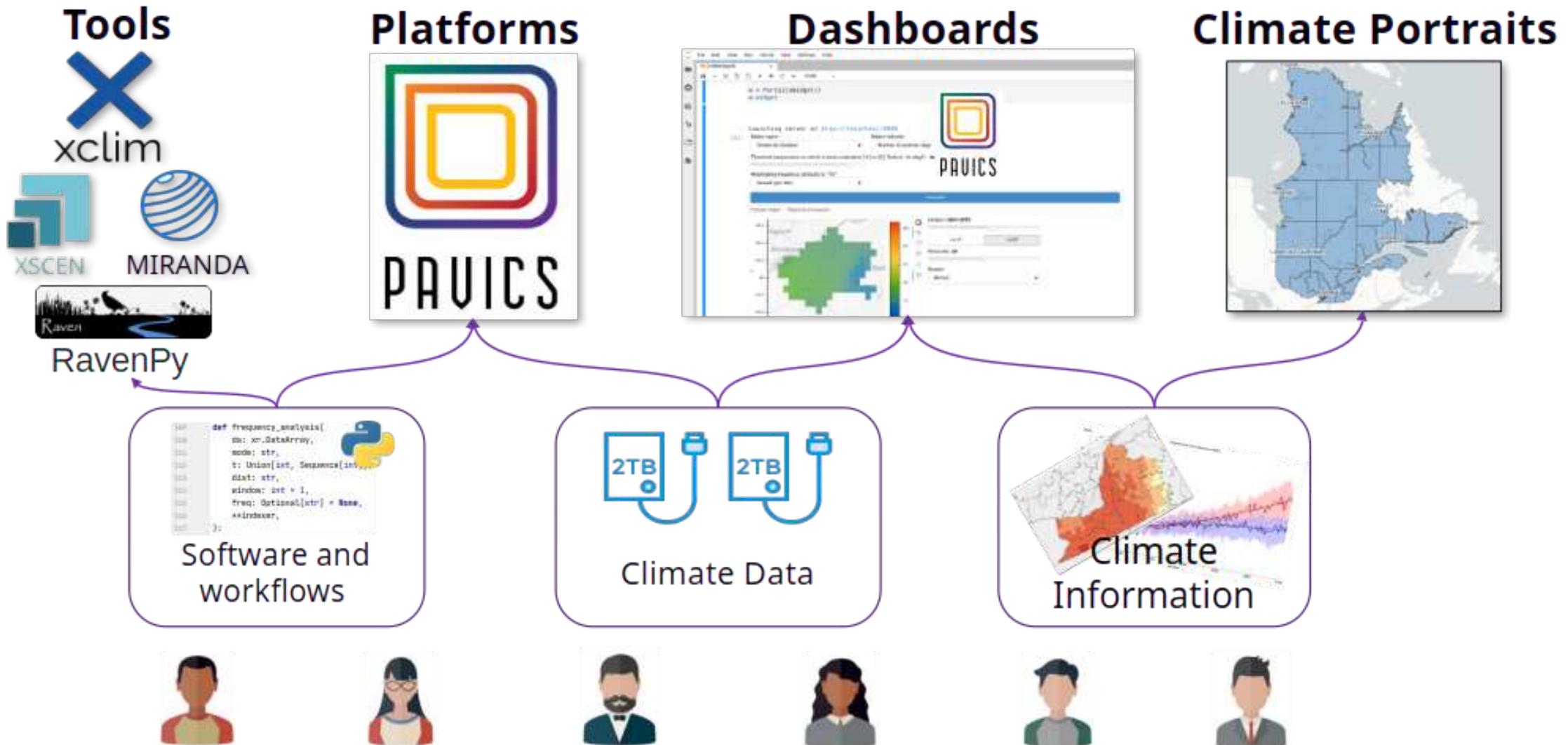
Plateforme de calcul « PAVICS »

École d'été en sciences du climat 2024



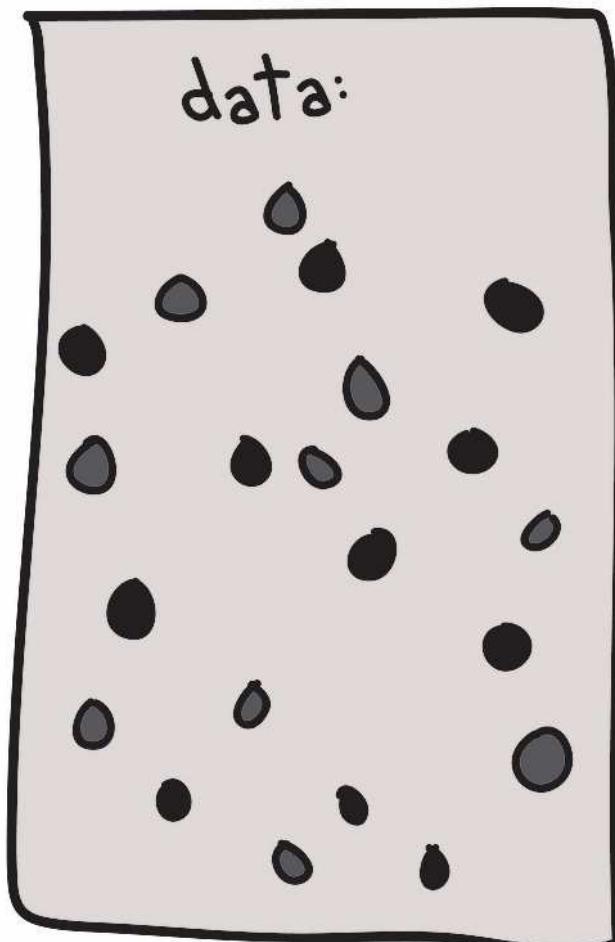
Les services climatiques chez Ouranos

École d'été en sciences du climat 2024



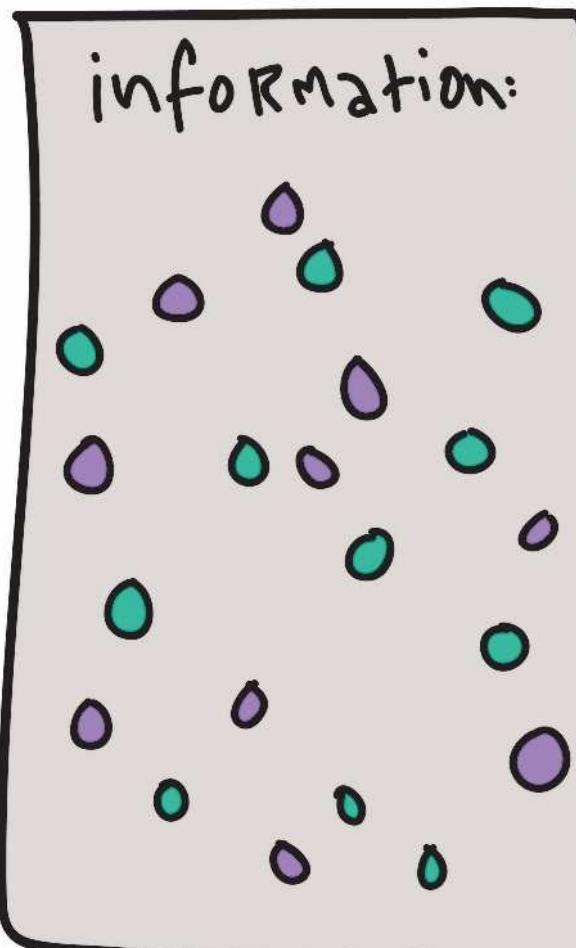
Merci pour votre attention !

École d'été en sciences du climat 2024

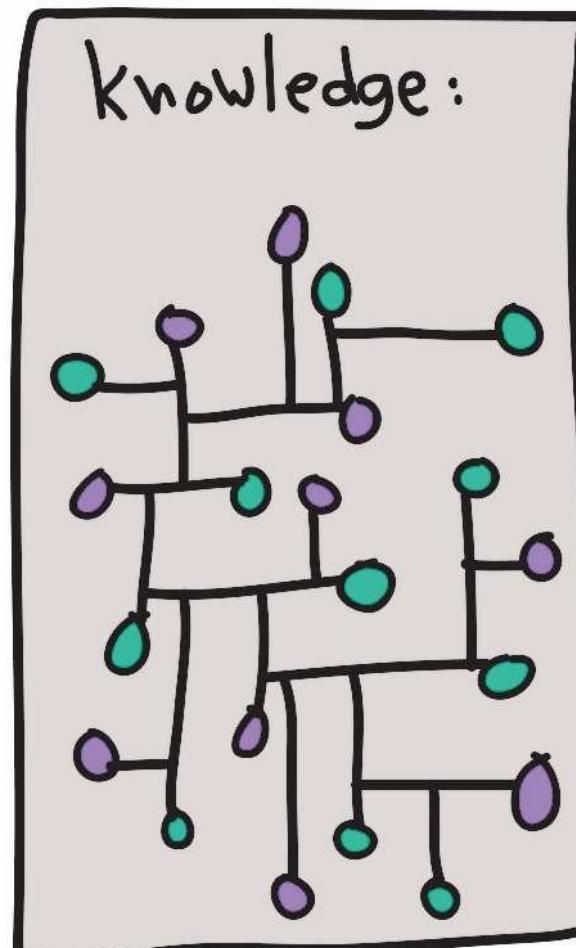


@bestqualitycrab

Diapositives disponibles à : [FIXME: LIEN-ICI](#)



@gapingvoid



@gapingvoid



Trevor James Smith

Spécialiste Plateformes climatiques,
données et opérations

Sciences du climat et services
climatiques

smith.trevorj@ouranos.ca

ouranos.ca

