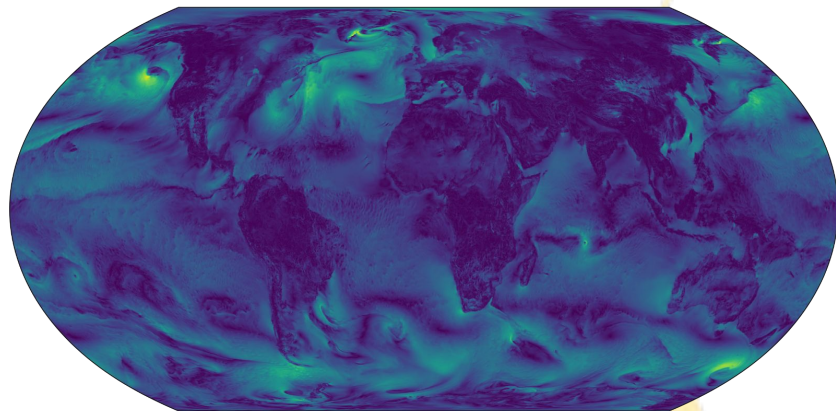


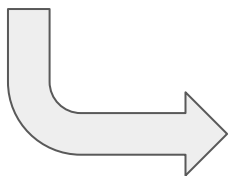


Améliorer les prévisions météo à l'aide de Transformers

Masterclass du CRIM
David Landry
2023-01-27



- Formation en intelligence artificielle de l'Université Laval
- Anciennement employé du Service Météorologique Canadien comme programmeur scientifique
- Employé du CRIM à titre de spécialiste du développement expérimental en IA



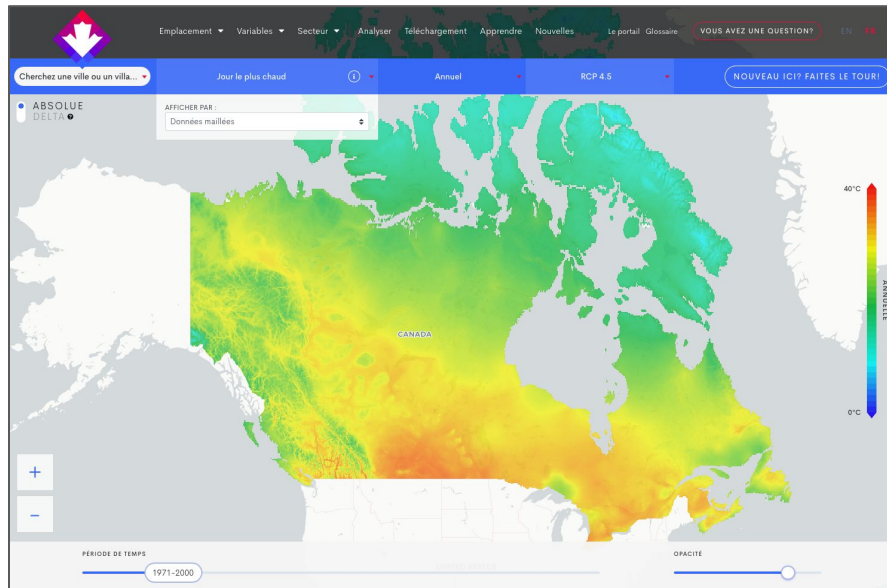
Point de vue très applicatif sur le sujet

À propos du CRIM

- Spécialiste du développement expérimental auprès de la PME québécoise
- Grande expérience en traitement de données géospatiales
 - Interopérabilité, plateformes
 - <https://donneesclimatiques.ca>
- Plus récemment, application de l'IA aux données des sciences de la terre



Environment and
Climate Change Canada
Environnement et
Changement climatique Canada



LA PRESSE Aujourd'hui, jeudi 20 janvier 2022 Montréal -15°C Suivez La Presse Je soutiens La Presse

ACTUALITÉS INTERNATIONAL DÉBATS CONTEXTE AFFAIRES SPORTS AUTO ARTS CINÉMA SOCIÉTÉ GOURMAND VOYAGE MAISON

Chroniques Économie Marchés Entreprises Techno Médias Finances personnelles PME Portfolio Têtes d'affiche

Météo au Canada en 2021

Plus de deux milliards en dommages selon des assureurs




PHOTO MARTIN TREMBLAY ARCHIVES LA PRESSE

Les conditions météorologiques ont provoqué des inondations historiques en Colombie-Britannique en 2021. 15 000 personnes ont été forcées de quitter leur maison inondée. Le montant des dommages y a été de plus de 500 millions de dollars.

(Toronto) La météo agitée partout au Canada a signifié que l'an dernier, les dommages aux biens assurés attribuables aux phénomènes météorologiques violents ont atteint 2,1 milliards, selon des données de Catastrophe Indices and Quantification dévoilées mardi par le Bureau d'assurance du Canada (BAC).

Mis à jour le 18 janvier

LA PRESSE CANADIENNE

Les montants des dommages les plus élevés de 2021 ont été, dans l'ordre, les inondations du 13 novembre en Colombie-Britannique (515 millions), la tempête de grêle du 2 juillet à Calgary (500 millions), la tempête de vent du 11 décembre dans l'est du pays (152 millions), la tempête du 12 janvier dans l'ouest du Canada (134 millions), et les tempêtes du 22 juillet dans les Prairies (120 millions).

AFFAIRES LES PLUS CONSULTÉS DERNIÈRE HEURE

01 CHRONIQUES
Logements
Incapable de remplir des HLM vides en région
Publié à 6h30

02 ÉCONOMIE
L'inflation à 4,8 % en décembre
Les taux d'intérêt sur le point de redécoller
Publié à 6h00

03 ENTREPRISES
Guru sacrifie la rentabilité à court terme pour promouvoir sa marque
Publié à 12h29

AFFAIRES LES PLUS CONSULTÉS AUJOURD'HUI

01 ENTREPRISES
Les quinqualliers s'opposent au passeport vaccinal
Publié hier à 13h03

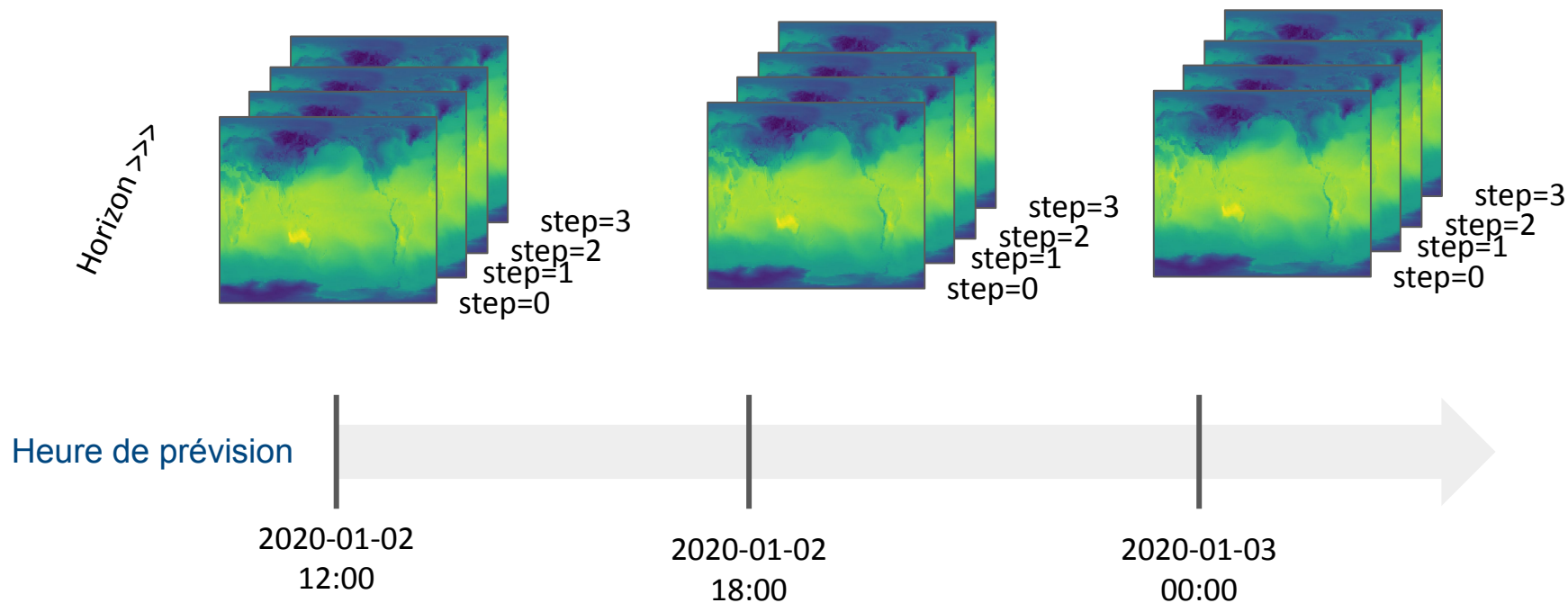
02 ÉCONOMIE
L'inflation à 4,8 % en décembre

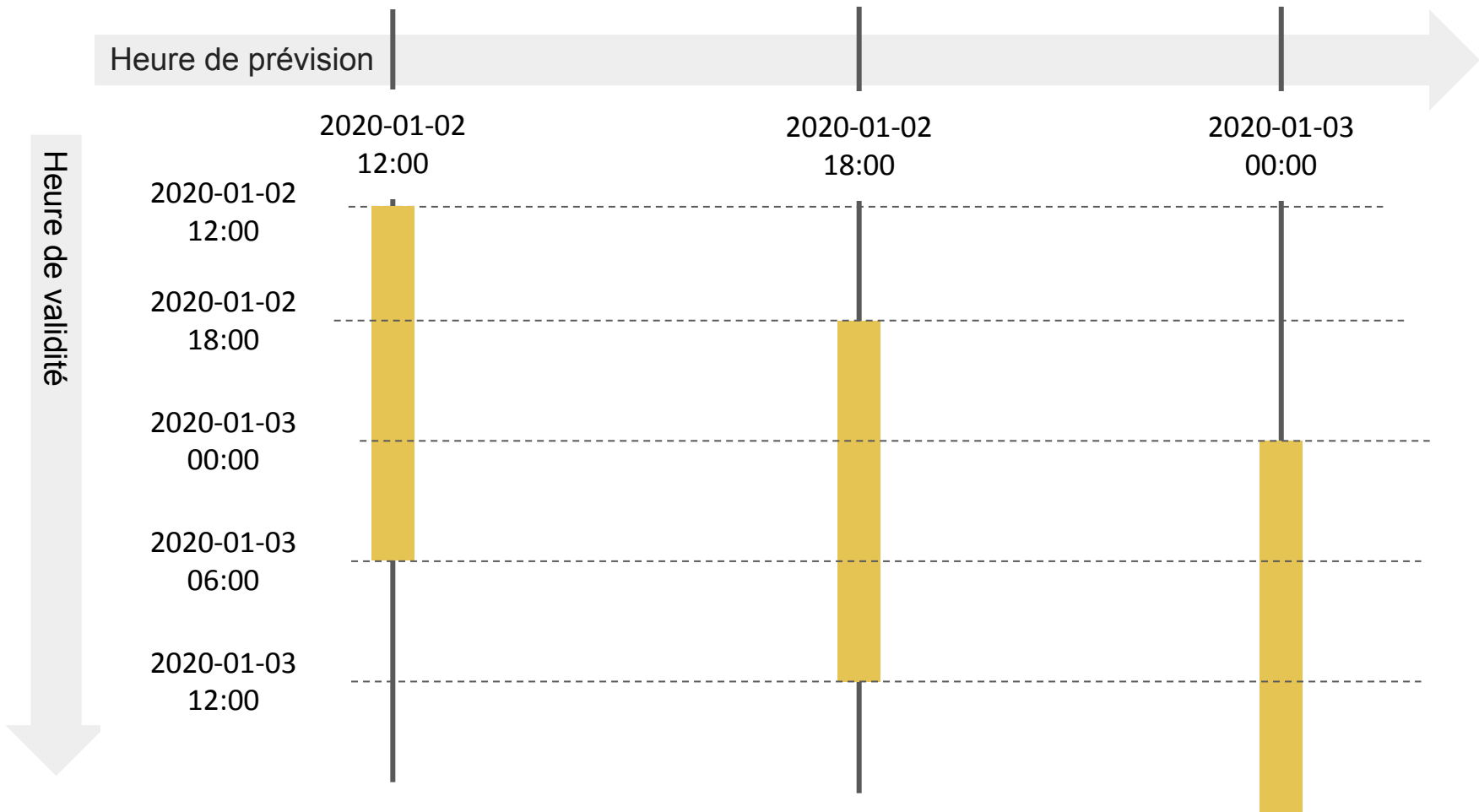
- «[...] invite les autorités canadiennes à faire de la **résilience au climat** une priorité»
- « il faut financer et adopter des mesures immédiatement pour protéger les Canadiens contre l'aggravation des **phénomènes météorologiques violents.** »

- Introduction du post-traitement de prévisions météorologiques
- Les *Transformers* pour le post-traitement
 - Bref retour sur les *Transformers*
 - Pourquoi?
 - Comment?
- Résultats
- Analyse des représentations
- Conclusion

La prévision météo et le post-traitement statistique

Anatomie d'une prévision météo





Les prévisions contiennent...

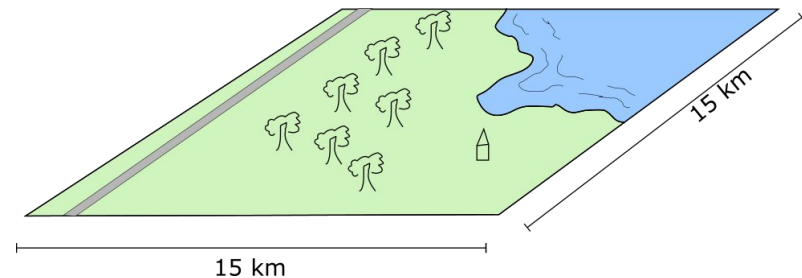
- Plusieurs variables (température, pression, précipitations, nuages...)
 - Le choix du sous-ensemble des variables à fournir au post-traitement est sujet à discussions
- Certaines variables ont plusieurs niveaux verticaux (température @ 2m, température @ 10m, température @ 1000m, ...)
- Les variables sont donc réparties sur une grille 5D (latitude, longitude, niveau vertical, temps de prévision, *step*)
- Ce nombre de dimensions est important et nécessite du travail spécifique lors de l'application des méthodes d'IA

Post-traitement statistique

- Nos prévisions météorologiques sont produites par des systèmes de prévisions numériques
- Malheureusement, les sorties de ces modèles ne sont pas toujours fidèles à ce qui est observé sur le terrain
 - Résolution insuffisante pour reproduire tous les phénomènes pertinents à la prévision, ou autres problèmes de paramétrisation
 - Problèmes dans la création de l'état météo initial (l'assimilation)
- Est-il possible d'apprendre de nos erreurs du passé pour corriger les sorties des modèles de prévision numérique?

Pourquoi a-t-on besoin d'un post-traitement?

- Compromis de la grille
- Effets locaux non-résolus
- Le problème des conditions initiales (assimilation)



NOAA Meteorological Development Laboratory, What is MOS? <https://vlab.noaa.gov/web/mdl/mos>

- Trois méthodes évaluées
 - Correction linéaire
 - *Random forest*
 - Réseau de neurones simple
- Les trois méthodes performant de façon similaire, indiquant qu'on a peut-être atteint la limite de l'information à extraire de cette configuration
- Des gains de 10-15% sont à espérer pour une prévision de température

Technical Memo



896

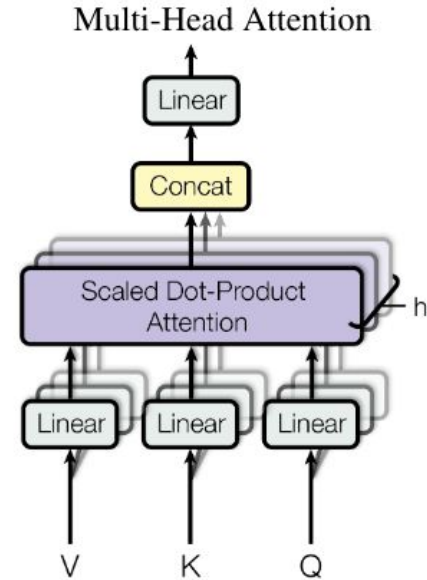
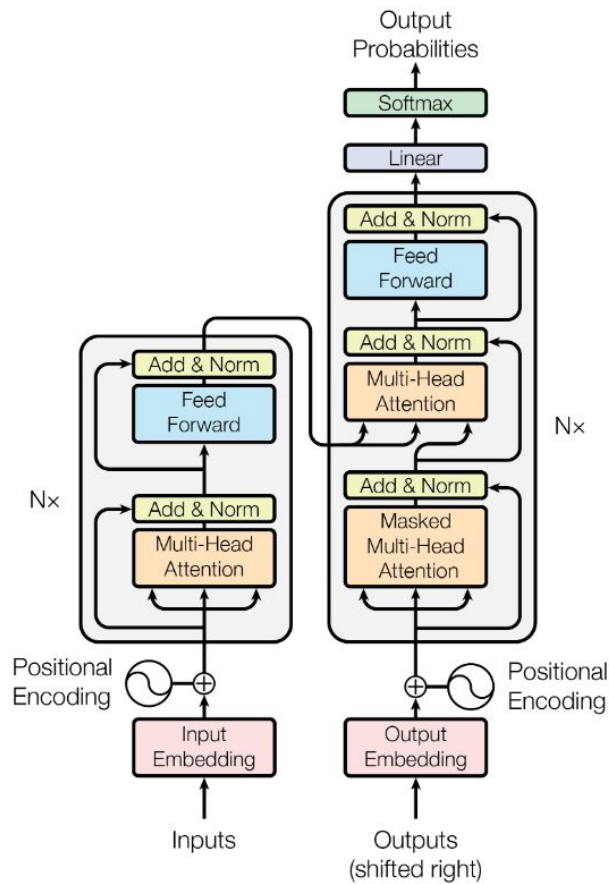
Statistical modelling of 2m temperature and 10m wind speed forecast errors

Zied Ben Bouallègue, Fenwick Cooper, Matthew Chantry, Peter Düben, Peter Bechtold, Irina Sandu
(Forecast & Research Department)

April 2022 - Manuscript submitted to JAWR

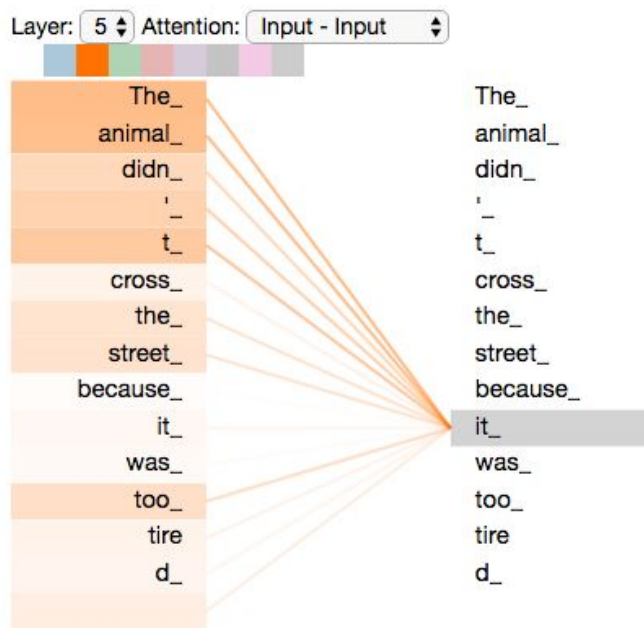
Transformers

- D'abords introduits pour des problèmes de traitement de la langue naturelle tels que la traduction automatique
- Ils sont derrière les *foundation models* tels que GPT-3
- Leurs conséquences sont toujours en train de percoler dans beaucoup de domaines de l'IA (vision par ordinateur, sciences de la terre)



Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, 2017: Attention is All you Need. *Advances in Neural Information Processing Systems*, Vol. 30 of, Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html> (Accessed May 17, 2022).

Attention



$$\text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) V$$

Alammar, J., 2018: The Illustrated Transformer. <https://jalammar.github.io/illustrated-transformer/> (Accessed January 9, 2023).

Pourquoi les *Transformers*?

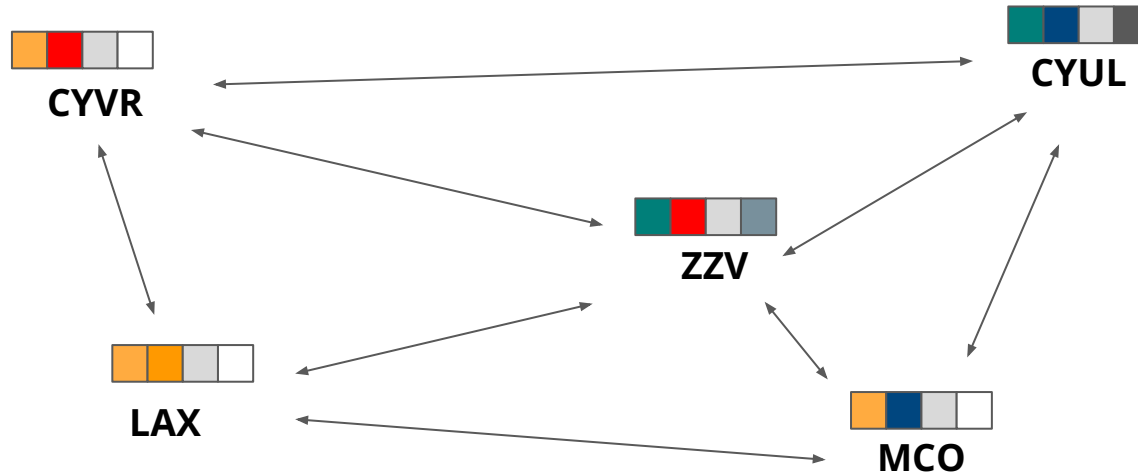
Pourquoi les *Transformers*?

- Apprentissage de représentations tout-usage -- permet l'utilisation de l'apprentissage par transfert
- Exploitation des relations distantes
- Plus grande robustesse aux données manquantes

Vu ces possibilités, on cherche à faire une première étape dans l'utilisation des *transformers* sur des données météorologiques en les appliquant au problème de post-traitement.

Pourquoi l'attention est-elle un bon mécanisme pour le post-traitement?

- Apprentissage par transfert possible via l'embedding de stations
- Plus de robustesse aux données manquantes (travaux futurs)
 - On peut envisager de remplacer les valeurs d'une données manquante, similairement à ce que font les modèles de langue
- Capacité de représentation très grande (travaux futurs)
 - On peut envisager un seul système pour traiter plus de variables en sortie, possiblement plusieurs modèles

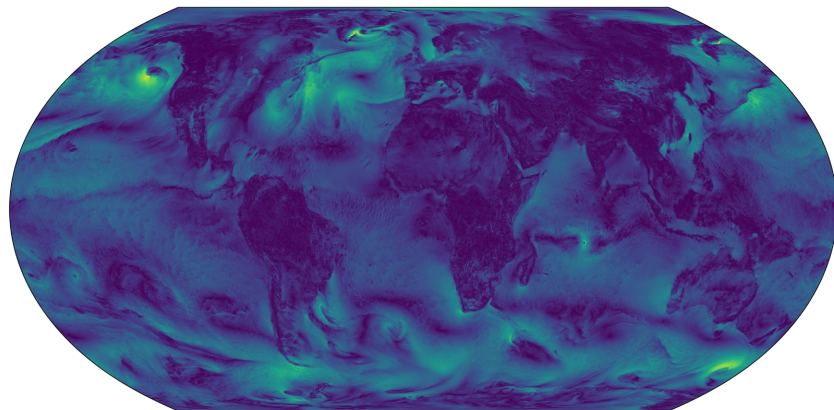


Le réseau d'attention apprend à partager de l'information entre les stations grâce à la représentation apprise

Expériences

1 - Post-traitement

- Les prévisions sont issues du **Global Deterministic Prediction System (GDPS)** canadien
 - Modèle global à résolution d'environ 15km
 - Seules les prévisions au Canada et aux États-Unis sont utilisées
- Les observations sont issues du réseau METAR via le [Iowa Environmental Mesonet](#) de l'université Iowa State
 - Peu de contrôle de qualité
 - Jeu de données d'observations canadiennes avec contrôle de qualité?
- Les données GDPS sont interpolées pour chaque station avec le plus proche voisin dans la grille

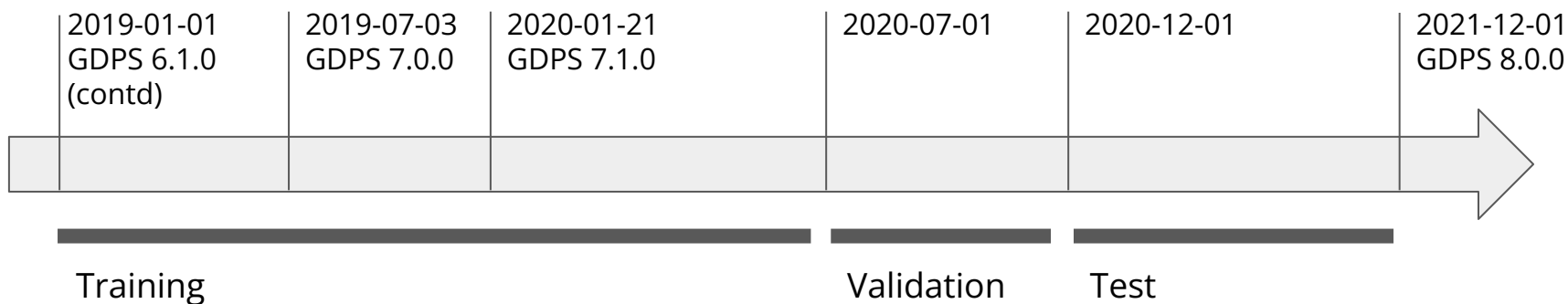


Issus des sorties du modèles GDPS

Variable	Niveau vertical
Albedo	-
Hauteur géopotentielle	850, 500 hPa
Point de rosée	2m
Pression	850, 500 hPa
Température	2m, 850 et 500 hPa
<i>Thickness</i>	-
Vent (vitesse, u, v)	10 m
Vent (u,v)	500 hPa

Dataset

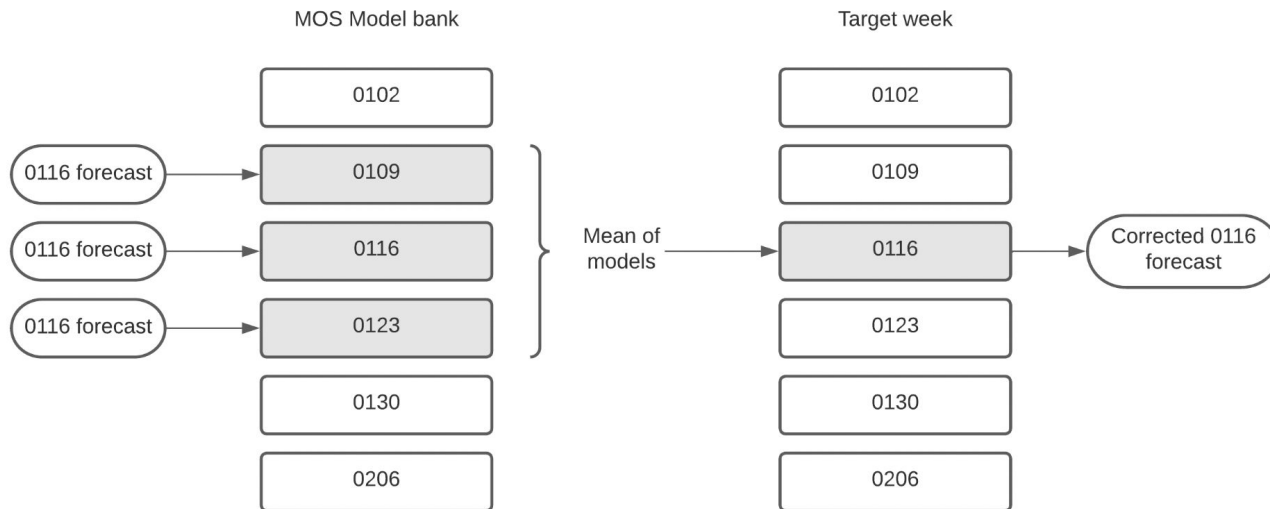
- Les stations METAR sont filtrées selon la disponibilité des données
- Les prévisions GDPS sont interpolées aux stations via le plus proche voisin



- *Model Output Statistics*
- Régression linéaire multiple qui utilise des champs de sortie du modèle comme prédicteurs, et qui prédit l'erreur du modèle pour le champ visé (température @ 2m dans notre cas)
- Quelques décisions à prendre
 - Quantité de modèle linéaires appris (un seul? un par station? un par station par pas de temps? un par station par pas de temps par passe?)
 - Quantité de paramètres à inclure dans la régression linéaire (pour éviter le surapprentissage)
- Pour ce jeu de données, la meilleure configuration que nous avons trouvé est d'avoir de nombreux modèles organisés en fenêtre roulante, avec un seul prédicteur

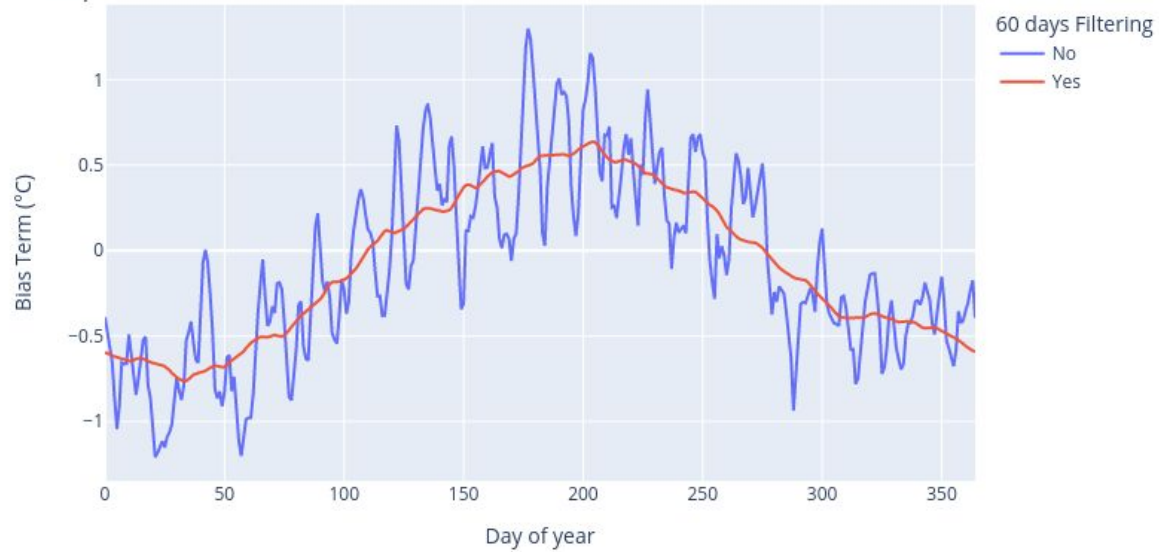
Modèle linéaire

How the rolling window works



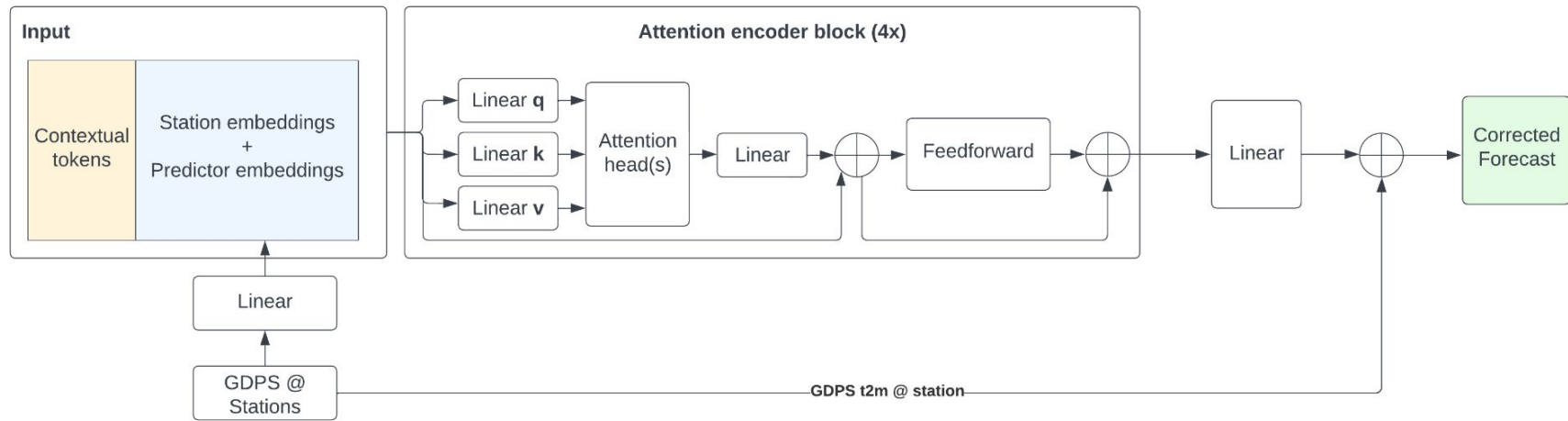
Configuration de la fenêtre roulante
Taille de fenêtre optimale pour nous: 60 jours

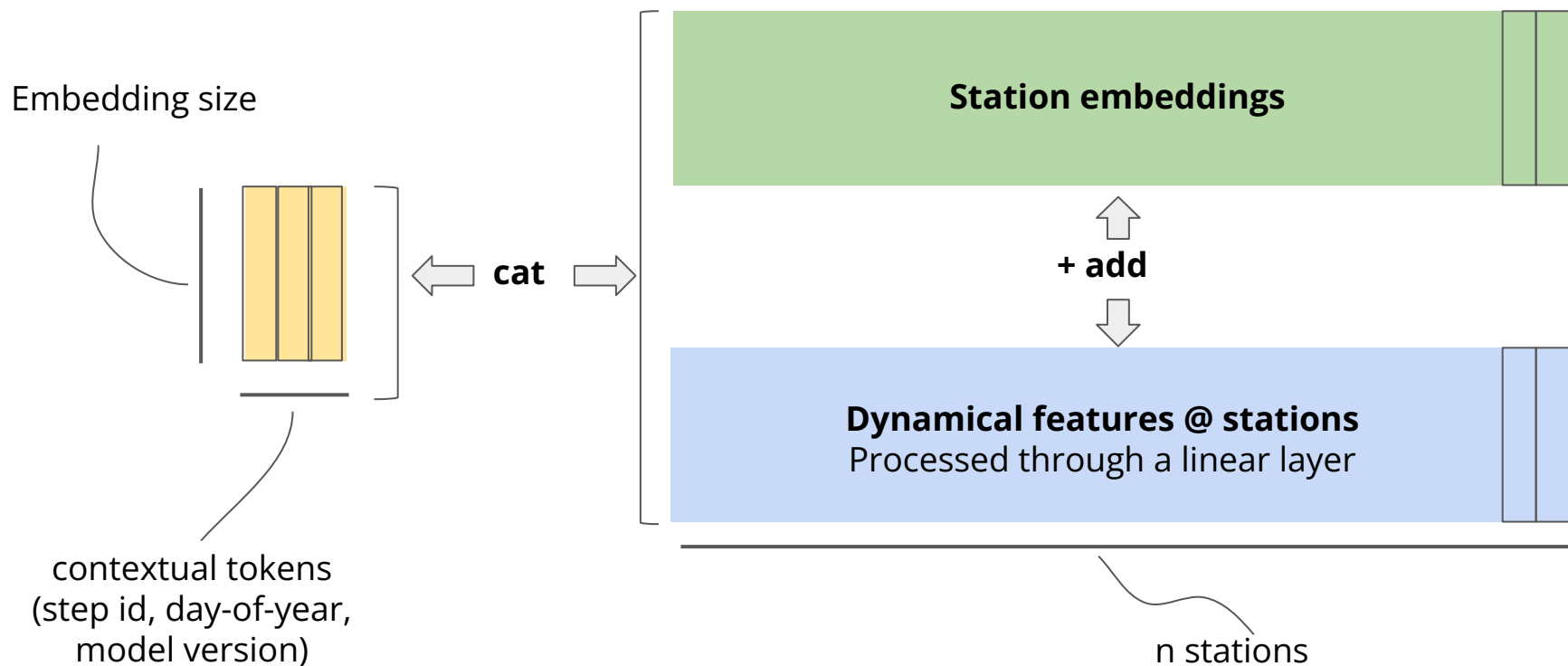
Bias term of MOS post-processing
Station CYUL
Step 192h



- Inspiré directement de l'encodeur de Vaswani et al (2017).
- Les exemples d'entraînements sont fournis *step par step*
 - "Effectue le post-traitement pour cette liste de station, sachant que la prévision est 48 heures dans le futur et qu'elle a été simulée la semaine du 12 mars"
- Travaux futurs: mieux exploiter la composante temporelle

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, 2017: Attention is All you Need. *Advances in Neural Information Processing Systems*, Vol. 30 of, Curran Associates, Inc.
<https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html> (Accessed May 17, 2022).





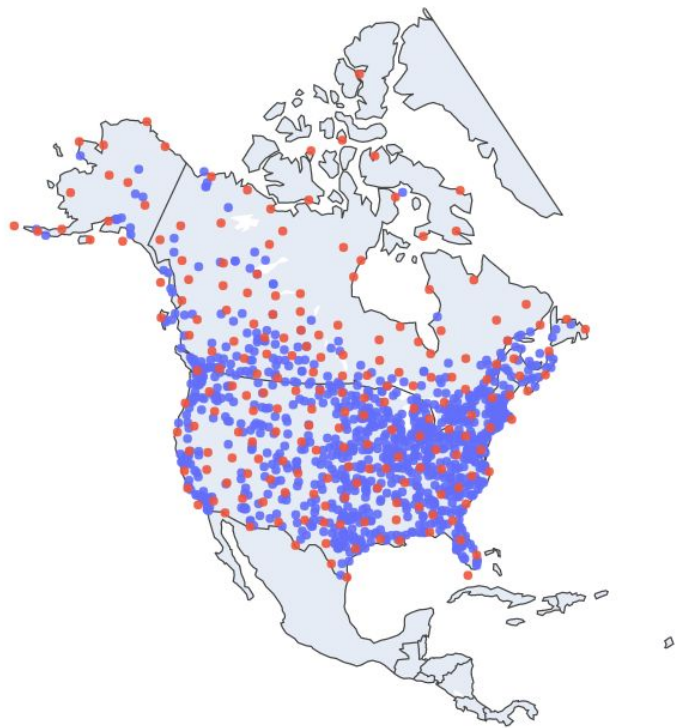
Construction des données d'entrée du transformer

Performance des modèles sur le jeu de données complet

	Toutes les stations	
Modèle	RMSE (°C)	Amélioration (%)
GDPS (sans post-traitement)	3.82	-
MOS	3.40	12
Attention	3.27	17

Expériences

2 - Apprentissage par transfert



The *bootstrap* set of stations which is target by fine-tuning

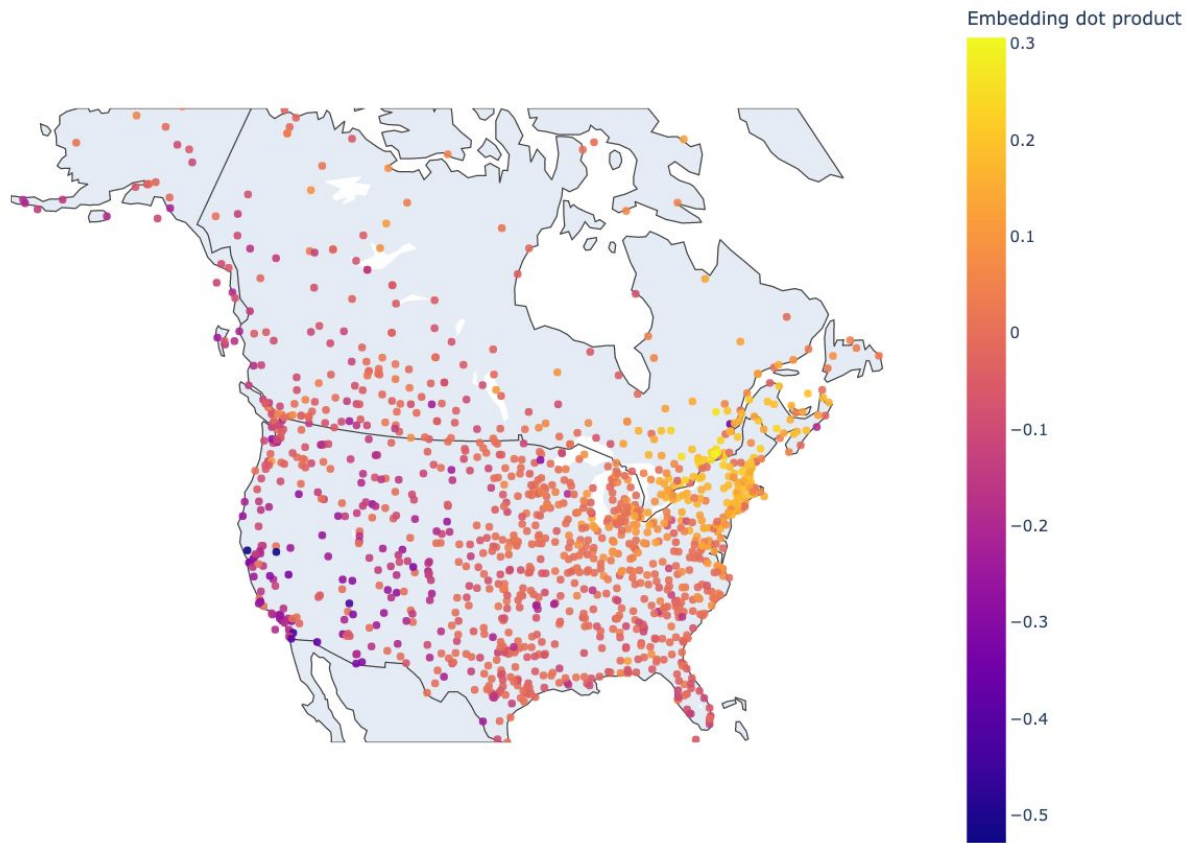
Model performance when trained with respect to station set

	RMSE (°C)		
Model	All stations	Reference set	Bootstrap set (fine-tuning)
Raw GDPS	3.82	3.72	4.25
MOS	3.40	3.27	3.65
Attention	3.27	3.15	3.63



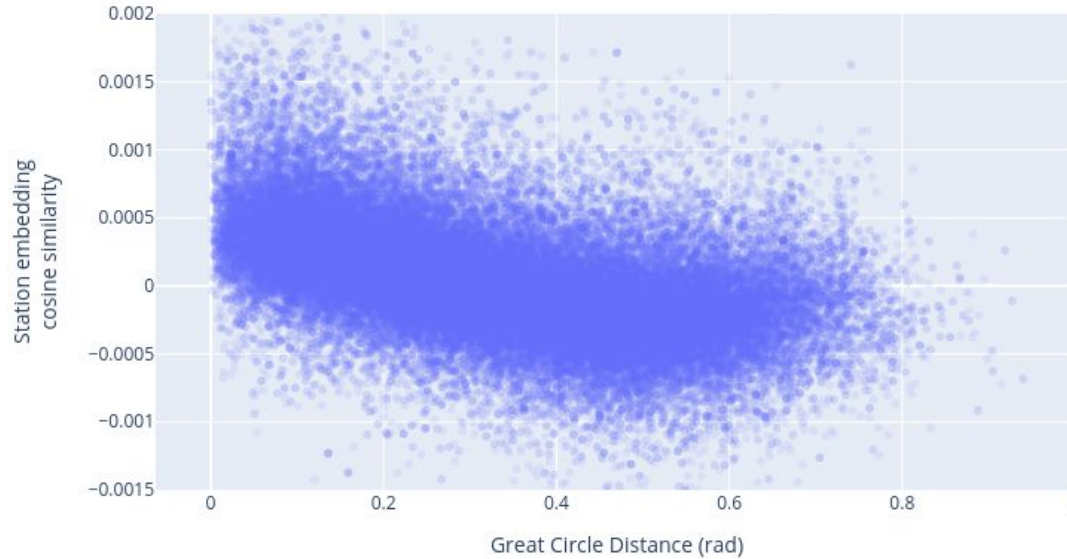
RMSE on Bootstrap stations according to dataset size

Analyse des représentations

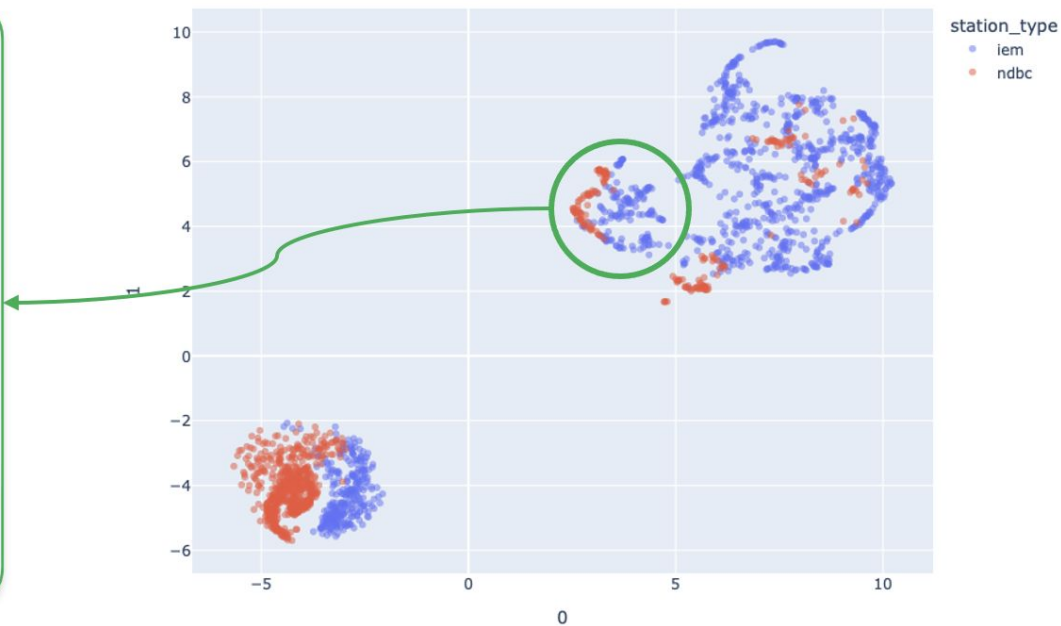
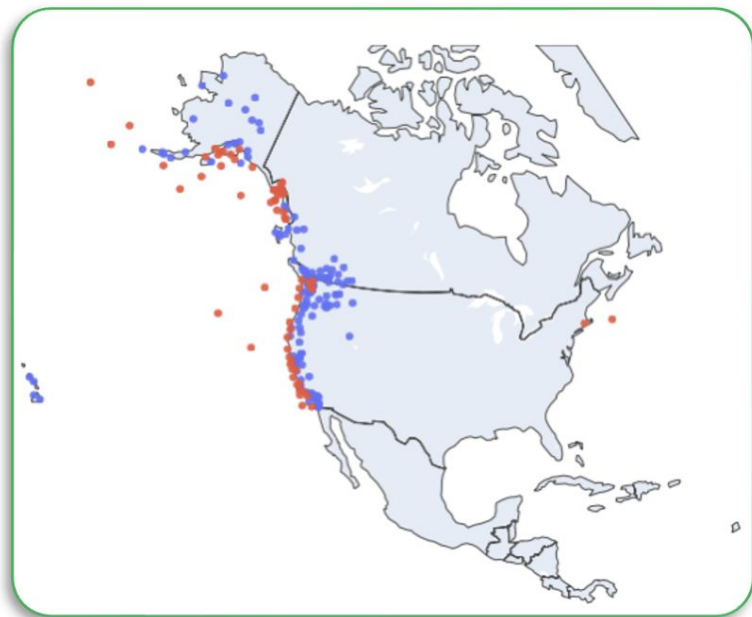


Proximity of the learned station embeddings to the CYUL embedding

Station embedding similarity with respect to their mutual spatial distance
Sample of n=100000 pairs



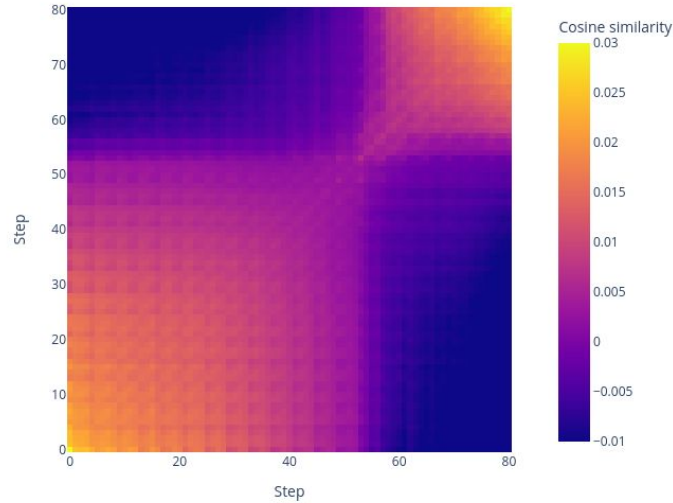
Relation entre la similarité des plongements et la distance géographique entre les stations



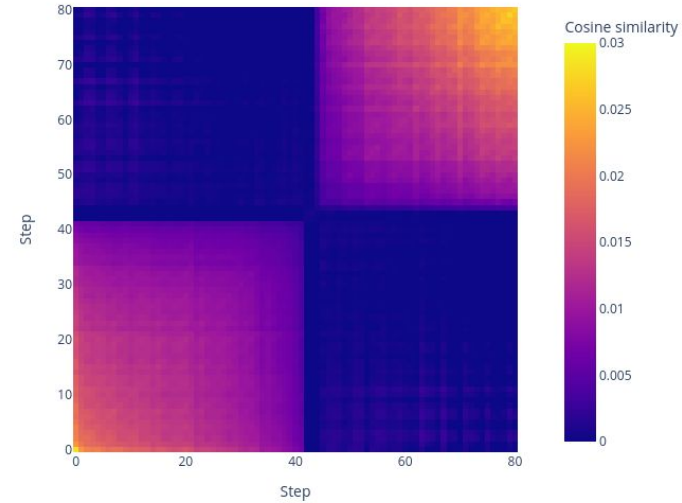
Analyse des *embedding* de stations via t-SNE

Crédits: Luca Vaio

Step embedding self-similarity



Step embedding self-similarity



Auto-similarité du plongement des différentes étapes de prévision, pour différents entraînements

En conclusion

- Les *transformers* ont une architecture appropriée pour le post-traitement statistique de prévisions de météo à des stations. Ils obtiennent des résultats supérieurs à la régression linéaire multiple pour des prévisions de température à la surface.
- Les représentations apprises permettent l'apprentissage par transfert, lors de l'ajout des nouvelles stations.
- Les représentations affichent des caractéristiques que l'on arrive à expliquer a posteriori

- Étudier la robustesse à d'autres changements de modèles de prévisions numériques
 - Ajout de pas de temps
 - Changement de résolution
 - Ajout de biais artificiels
- Améliorer le comportement de convergence du réseau
- Étendre le principe sur grille avec des architectures inspirées des *Visual transformers* (problèmes de résolution)

Remerciements

Le projet SMC01 a été réalisé avec le soutien financier d'Environnement et Changement Climatique Canada



Environment and
Climate Change Canada

Environnement et
Changement climatique Canada

- Vannitsem, S., and Coauthors, 2021: **Statistical Postprocessing for Weather Forecasts: Review, Challenges, and Avenues in a Big Data World**. *Bulletin of the American Meteorological Society*, **102**, E681–E699, <https://doi.org/10.1175/BAMS-D-19-0308.1>.
- Rasp, S., and S. Lerch, 2018: **Neural Networks for Postprocessing Ensemble Weather Forecasts**. *Monthly Weather Review*, **146**, 3885–3900, <https://doi.org/10.1175/MWR-D-18-0187.1>.
- Bouallègue, Z. B., F. Cooper, M. Chantry, P. Düben, P. Bechtold, and I. Sandu, 2022: **Statistical modelling of 2m temperature and 10m wind speed forecast errors**. European Centre for Medium-Range Weather Forecasts,.