



Réseaux de neurones et prévision des températures

André DAUPHINÉ¹✉

¹ Université Côte d'Azur, Nice, France

Résumé : Les climatologues établissent des prédictions à l'aide de modèles statistiques et par simulation de modèles physiques numériques. Cet article propose d'explorer l'efficacité d'un modèle d'intelligence artificielle pour prévoir des températures moyennes. Ce modèle, un réseau de neurones récurrents est appliqué à des séries temporelles de températures de la station de Nice (France) et de Cluj-Napoca (Roumanie).

Mots-clés : apprentissage automatique, réseau de neurones récurrents, températures mensuelles et par semaine

Abstract. Climatologists make predictions using statistical models and by simulating numerical physical models. This article explores the effectiveness of a machine learning model. This model, a recurrent neural network, is applied to temperature time series from the Nice station (France) and the Cluj-Napoca station (Romania).

Keywords : machine learning, recurrent neural network, monthly and weekly temperatures

Introduction

Cet article a pour objectif de réaliser des prévisions de températures à l'aide d'un modèle d'apprentissage automatique. Un modèle qui s'applique aux deux stations de Nice et de Cluj-Napoca, représentatives d'un climat méditerranéen et d'un climat semi continental. Comme dans toutes les disciplines où l'expérimentation s'avère impossible, les connaissances en climatologie progressent grâce à l'emploi de modèles. Pour répondre à un questionnement, le climatologue choisit une des quatre formes de modélisation disponibles. Les modèles physiques numériques ont pour objectif principal de répondre à la question *Pourquoi* (Emanuel K., 2019 ; Gentleman A., Rood R. B., 2016, McGuffie K. Henderson-Sellers A., 2014). Répondre à la question *Comment* s'obtient avec les modèles statistiques (Privalesky V., 2021 ; Wei W. W. S., 2005). Quant

CORRESPONDENCE:

✉ 186B, avenue de Pessicart 06100 Nice, France

✉@ a.dauphine@orange.fr

ARTICLE HISTORY:

Received: 14.04.2025

Received in final form: 10.06.2025

Accepted: 15.06.2025

aux modèles d'apprentissage automatique, dits d'intelligence artificielle, leur emploi semble recommandé pour réaliser des prévisions et fournir des aides à la décision. Enfin, le rôle principal des modèles graphiques et cartographiques consiste à illustrer les résultats obtenus avec les trois autres types de modèles.

Certes, tous ces modèles débordent leur compétence principale et ils permettent aussi de répondre en partie à d'autres questions. Par exemple, les modèles physiques numériques servent à établir des projections, projections qu'exploitent les rapports du GIEC. Cependant, malgré d'indéniables progrès, la plage d'incertitude issue de ces modèles demeure stable depuis une vingtaine d'années. En témoigne les études comparées. En effet, les analyses réalisées dans le cadre de CMIP6 montrent que la plage d'incertitude relative aux températures simulées par les plus récents modèles n'a pas diminué par rapport aux résultats comparés dans les programmes CMIP 4 ou CMIP5. Pourtant, les modèles physiques récents sont bien plus complets. Ce paradoxe s'explique très simplement. En joignant des sous-modèles pour prendre en compte de nouveaux facteurs jugés non négligeables, les modélisateurs ajoutent des imprécisions et des incertitudes. Ces limites compensent les acquis de connaissance obtenus par l'agrégation de ces nouveaux mécanismes qui influencent l'évolution du climat. De même, le reproche souvent fait aux modèles statistiques de se cantonner dans les seules description et compréhension d'un phénomène climatique mérite d'être nuancé. Ainsi, reconnaître le caractère stochastique d'une tendance signifie que cette tendance résulte de plusieurs causes. Et démontrer le caractère multifractal d'une série temporelle quelconque indique que les valeurs prises par cette variable sont le produit d'interactions qui agissent à différentes échelles d'espace et de temps (Laguës M., Lesne A., 2006). Même certaines représentations graphiques possèdent un pouvoir de compréhension, comme le montre les graphiques de récurrence (Webber Ch. L. and Marwan N., 2015).

Par ailleurs, certains scientifiques n'hésitent pas à recourir conjointement à plusieurs types de modèles. Les données de simulation des grands modèles numériques deviennent souvent des entrées pour des modèles statistiques. Un autre exemple : dans des articles récents, Sh. Lovejoy (2022) propose d'investir dans des modèles FEBE, des modèles simplifiés de bilan énergétique, mais fractionnaires. Il estime que de tels modèles fourniront de meilleures prévisions supérieures à celles des grands modèles numériques.

Retenons enfin que tous les modèles souffrent d'imprécisions et d'incertitudes (Dauphiné A., 2025). Cependant, dès la fin du XIXe siècle, le statisticien G. Box émettait la réserve suivante : « Tous les modèles sont faux, mais certains sont utiles ». Cette phrase est souvent reprise par les idéologues qui méprisent la démarche scientifique. Certes, tout scientifique doit accepter cette idée, mais sous une formulation nettement plus nuancée. Il est vrai que tous les modèles sont imprécis et incertains, néanmoins les modèles scientifiques possèdent une part de vérité. Ils lèvent le voile sur une partie

de la réalité. Ils sont donc, sauf exception, toujours utiles. C'est grâce aux modèles que la science progresse, en réduisant les imprécisions et les incertitudes, et parfois même en proposant un nouveau modèle. Certes, il n'y a pas de science exacte, en témoigne la climatologie, une science jeune, où beaucoup d'imprécisions et d'incertitudes perdurent malgré les efforts des modélisateurs. Les imprécisions concernent les données et les paramètres inclus dans un modèle, tandis que les incertitudes touchent à la logique même du modèle, à sa forme mathématique ou à sa traduction en langage informatique.

1. Quatre étapes pour construire un modèle d'apprentissage automatique

L'élaboration d'un modèle d'apprentissage automatique passe par quatre étapes : définir un questionnement précis, construire deux ensembles de données, d'apprentissage et de test, choisir un modèle, le traduire en langage informatique et le tester.

1.1. Un objectif : la prévision des températures en un lieu donné

Dans cet article, nous abordons la prévision des températures aux deux échelles du mois et de la semaine. L'utilité de ces prévisions est reconnue pour la production agricole, la consommation d'énergie ou la fréquentation touristique. En effet, les acteurs du tourisme, des transports, des risques et d'activités agricoles souhaitent anticiper les écarts thermiques. De fortes chaleurs ou un mois plus froid sont-ils prévisibles ?

Pour réaliser de telles prévisions avec un modèle d'apprentissage automatique, le géographe privilégiera deux approches. D'une part, il peut réduire son champ de recherche au traitement d'une seule variable, donc la température mensuelle ou par semaines. Dans ce cas, il est guidé par le seul principe de récurrence, car chaque moyenne du mois de juillet s'apparente, malgré un écart variable, à celle des mois de juillet des années précédentes. Dans ce type d'étude, qui ignore la recherche des mécanismes sous-jacents, seule la qualité des projections compte et retient l'attention du décideur. D'autre part, il peut envisager de prévoir ces températures en prenant en compte l'impact des télé connexions. Une seconde approche écartée dans cet article, mais abordée dans l'ouvrage consacré à l'apprentissage automatique en géographie (Dauphiné A., 2025).

1.2. Importer et bâtir les ensembles de données d'apprentissage et de test

Pour élaborer un modèle d'apprentissage automatique, il est d'abord nécessaire de récupérer un ou plusieurs ensembles de données. Le premier ne compte que les températures mensuelles moyennes. Ainsi, pour cet exercice, les données de températures retenues concernent les stations de Nice et de Cluj-Napoca. Ces données

s'obtiennent facilement avec le logiciel *Wolfram Mathematica*. Il offre une gigantesque base de connaissances, notamment des informations climatiques qui proviennent de GISS4. Les températures maximales et minimales pour chaque mois sont récupérées de 1975 à 2024, soit 50 ans. Ainsi, les deux séries mensuelles des températures maximales et minimales comptent 600 données tandis que les séries par semaine en comptent 2646. Le second jeu est composé des températures moyennes non plus par mois mais par semaine. Et cela pour chaque station.

Chacun de ces deux jeux de données est divisé en deux sous-ensembles. Le premier, de 1975 à 2021, constitue l'ensemble d'apprentissage. Le deuxième, s'étend de 2022 à 2024, soit 36 données mensuelles. Il sert de données tests pour vérifier la qualité du modèle d'apprentissage automatique. Pour les températures moyennes par semaine, le fait de disposer de données plus nombreuses, nous avons élargi l'ensemble des données test à un peu plus de 4 ans, soit 211 données. Comme le chercheur dispose des températures réellement observées entre 2022 et 2024 ou 2021 et 2024, il peut comparer la courbe des températures réellement observées et celles que le modèle simule.

1.3. Choisir un modèle d'apprentissage automatique

Pour réaliser des prévisions de futures températures, le géographe peut emprunter plusieurs modèles d'apprentissage automatique. Cependant, la plupart des ouvrages scientifiques mettent en avant la qualité des réseaux de neurones récurrents pour traiter des séries temporelles.

Dans les réseaux de neurones classiques, l'information circule dans un seul sens, des entrées vers les sorties. Elle passe de couche en couche. De ce fait, les données sont traitées successivement, et de manière indépendante. Leurs relations ne sont pas prises en compte. Ainsi, dans la phrase « Je vais jouer au football », le réseau analyse chaque mot séparément, hors de son contexte, et non pas un mot par rapport aux autres mots. Et, pour une série des températures, chaque valeur mensuelle est l'objet d'un calcul simple, sans tenir compte des températures relevées les mois précédents. Dans un réseau récurrent, au contraire, chaque mot tient compte des mots précédents et chaque température mensuelle dépend aussi des températures des mois précédents. Grâce à cette mémoire, ce type de réseau est très efficace pour le traitement de séquences, donc pour analyser des séries temporelles ou la succession de mots dans une phrase. Avec cette nouvelle architecture, proposée lors des vingt dernières années, les spécialistes construisent plusieurs algorithmes. Ils en privilégient trois. Dans les réseaux récurrents basiques, dits Vanilla, les couches se succèdent les unes après les autres. Ces réseaux constituent un simple outil pédagogique. Les LSTM (*Long Short Term Memory*) prennent en compte plusieurs mémoires dans les données, à court ou à long terme. Ces réseaux sont fondés sur la notion de porte (*gate*). Une porte d'oubli provoque l'évacuation de données devenues inutiles pour construire le modèle. Il en

va de même des GRU (*Gated Recurrent Unit*). De plus, chacun de ces algorithmes offre une large panoplie d'options. Et comme la majorité des réseaux de neurones récents, ces réseaux récurrents peuvent compter de nombreuses couches. Quand le chercheur envisage un tel modèle, les données de l'ensemble d'apprentissage s'ordonnent selon le schéma suivant :

$$\{ \text{température } t-x, \dots, \text{température } t-1 \} \rightarrow \text{température } t0$$

1.4. Un modèle informatique en langage *Wolfram Mathematica*

Pour élaborer un modèle d'apprentissage automatique, le géographe choisit un logiciel. Trois catégories de logiciels sont disponibles. Les premiers, les logiciels spécialisés proviennent de grandes entreprises. Ainsi, Google développe depuis 2015 le logiciel open source *TensorFlow*. Ces logiciels adaptés pour résoudre une seule tâche présentent un intérêt pour une recherche en thèse. Mais, le géographe doit alors se plonger dans d'autres logiciels, notamment pour inclure ses résultats dans un contexte géographique, tel un SIG. Une seconde catégorie comprend les logiciels fondés sur un langage, comme *C* ou *Python*. Malgré l'apport de bibliothèques prêtes à l'emploi, telle *PyTorch* avec le langage *Python*, cette solution présente un grave défaut. Les programmes à écrire pour concrétiser une recherche géographique sont relativement longs.

En fait trois logiciels généralistes procurent plus d'avantages : *R*, *Matlab* et *Mathematica*. Le premier, *R*, offre deux avantages : la gratuité et une très vaste communauté de chercheurs qui l'enrichit sans cesse. Mais, ce logiciel n'est pas tout en un. De ce fait, le géographe doit parfois jongler avec de très nombreux paquetages imbriqués. Ce qui constitue une potentielle source d'erreurs. En outre, les aides au logiciel sont succinctes hormis une très abondante bibliothèque d'ouvrages en langue anglaise. Le logiciel *Matlab*, le préféré des disciples des sciences dites dures reste un des plus complets avec ses toolbox, ses aides de toute nature et sa communauté d'utilisateurs.

Cependant, nous préférons *Wolfram Mathematica*. D'abord, avec plus de 7000 instructions, il couvre toutes les formes de modélisations géographiques : statistiques et processus stochastiques, géostatistiques, automates cellulaires, systèmes multi-agents, équations différentielles ordinaires et partielles, traitements d'images, théorie des graphes, cartographie... Toutes ces instructions sont directement accessibles sans recours à des programmes externes. De plus, ce langage est sans doute le plus concis. Un programme de plus de deux pages dans un autre langage nécessite moins d'une demi-page en langage *Wolfram Mathematica*. En outre les aides relatives à chaque instruction facilitent grandement la programmation. De nombreux livres, plus de deux cents, complètent ces aides. Deux d'entre eux traitent de l'apprentissage automatique (Bernard E., 2021 ; Villalobos Alva J., 2021). Ajoutons que plus de 13 000 projets de démonstration et des sites Internet apportent les réponses d'une large communauté de

chercheurs pour résoudre certaines difficultés que le géographe peut rencontrer. De plus, des dizaines de cours sur l'apprentissage automatique en accès libre guident le chercheur. Ils contiennent des programmes rapidement transformés pour répondre à des questionnements géographiques. En outre, des ponts existent entre ce logiciel, le logiciel *R*, et divers langages généralistes, dont *Python*. Et, depuis la version 13, le logiciel *Wolfram Mathematica* importe facilement des réseaux de neurones élaborés dans des centres de calculs avec des logiciels spécialisés, comme *TensorFlow*.

Pour répondre à la question posée un premier programme récupère les données mensuelles de températures de la station de Nice (LFMN) et de Cluj-Napoca (LRCL). Pour récupérer toutes les données, il suffit de remplacer le terme LFMN par LRCL, puis l'option Month par Week. Au total, ce sont 8 jeux de données, qui sont importés de la base de connaissance WeatherData, puis sauvegardés.

```
ClearAll["Global`*"]
Nice1 = WeatherData["LFMN", "MinTemperature", {{1975, 01, 1}, {2024, 12, 31}, "Month"}]
Nice2 = WeatherData["LFMN", "MaxTemperature", {{1975, 01, 1}, {2024, 12, 31}, "Month"}]
niceMin = nice1["Values"] // QuantityMagnitude;
niceMax = nice2["Values"] // QuantityMagnitude;
Export["aaniceMinMonth.dat", niceMin]
Export["aaniceMaxMonth.dat", niceMax]
```

Un second programme élabore alors le modèle, l'applique successivement à chacun des 8 jeux de données. Pour cela, le programme est relancé 8 fois, mais en choisissant un des 8 jeux de données. En sortie diverses informations permettent d'apprécier la qualité du modèle, et illustre la validité des prévisions.

```
ClearAll["Global`*"]
data = Flatten[Import[SystemDialogInput["FileOpen"]]];
dataformat = List /@ Most[#] -> List@Last[#] & /@ (Partition[data, 51, 1]);
training=Take[dataformat,514]
test = Drop[dataformat, 514];
net = NetChain[{{GatedRecurrentLayer[20], LinearLayer[1]}, "Input" -> {50, 1}, "Output" -> 1}
model = NetTrain[net, training]
reel = Flatten[test[All, 2]]
test = Partition[Flatten[test[All, 1]], 50];
prevision = Flatten[Table[model[test[[i]]], {i, 1, 36, 1}]]
Information[model]
ListLinePlot[{reel, prevision}, PlotLegends -> {"reel", "prevision"}]
measure = PredictorMeasurements[prevision, reel]
Dataset[AssociationMap[measure[#, ComputeUncertainty -> True] &,{ "MeanDeviation",
"RSquared", "RejectionRate"}]]
```

Après avoir effacé préalablement la mémoire, une série de températures, obtenue avec le programme précédent, est importée. Ces données sont alors formatées. Ce formatage s'inspire d'une question posée aux utilisateurs de la communauté

Stackexchange par M. Kagaratsch. Chaque valeur de température, mensuelle ou par semaine, dépend des 51 valeurs précédentes. De ces données formatées sont engendrés un ensemble d'apprentissage, nommé training, puis un ensemble de test. Le réseau de neurones récurrents, appelé net, associe deux types de couches, une GatedRecurrentLayer, suivie d'une LinearLayer. Et, en entraînant les données d'apprentissage avec ce réseau de neurones, le programme livre un modèle. Ce modèle, dénommé model, est alors appliqué à l'ensemble des données test. Enfin, pour apprécier la qualité de généralisation, donc le pouvoir prévisionnel du modèle, le programme affiche un premier graphique réalisé avec l'instruction ListLinePlot[]. Il compare les données réelles aux données prédites par le modèle de 2021 à 2024. Puis l'instruction PredictorMeasurements[] détermine une cinquantaine de paramètres numériques et de graphiques, dont le graphique de comparaison. Nous retenons seulement trois indicateurs numériques de qualité, la déviation moyenne, le coefficient de détermination, et le taux de rejet. Enfin, pour appliquer ce modèle aux températures moyennées sur un intervalle d'une semaine, il faut remplacer les valeurs 514 par 2388 et 36 par 211.

2. Analyse critique des résultats.

L'application de ce programme aux 8 jeux de données crée en sortie 16 graphiques et 8 tableaux d'indicateurs. Si tous sont utilisés pour une analyse critique des résultats, la plupart ne sont pas affichés dans cet article.

2.1. Un premier résultat encourageant

Soit l'exemple des données de températures mensuelles maximale pour la station de Cluj-Napoca. À cette échelle temporelle, la figure 1 affiche une bonne concordance entre les données réelles et les prévisions.

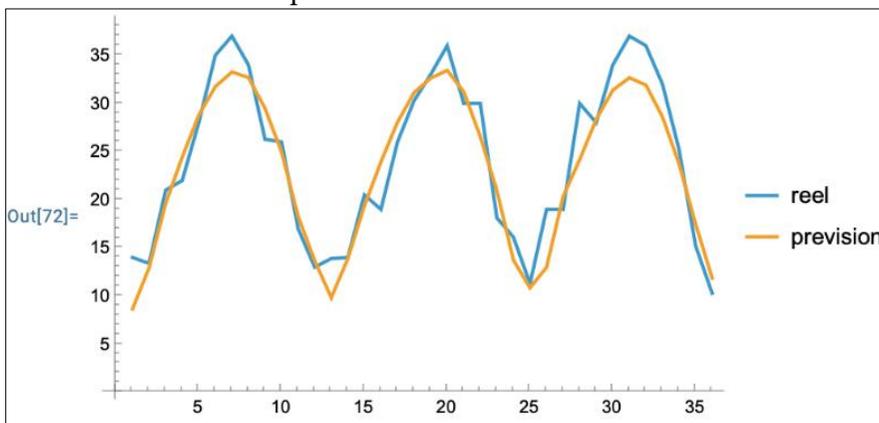


Figure 1. Données test réelles et prévues par le modèle pour les moyennes mensuelles maximales de Cluj-Napoca

Source: auteur

Ce que confirme la figure 2. En effet, les 36 points sont proches de la diagonale. Enfin, le tableau 1 des indicateurs numériques atteste aussi que ce modèle est acceptable en première instance.

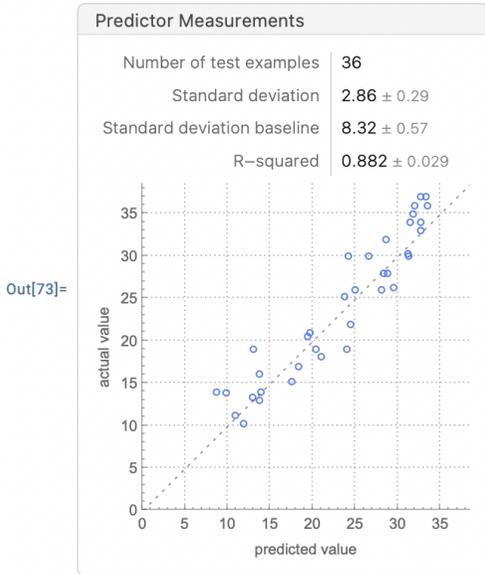


Tableau 1. Indicateurs de qualité du modèle d'apprentissage pour les températures mensuelles minimales à Cluj-Napoca

	Min_Mois
Déviatiion_Moy	2,35
Coef_Détermination	0,88
TxDeRejet	0

Source : auteur

Figure 2. Graphique de comparaison des températures mensuelles maximales réelles et prévues à Cluj-Napoca
Source : auteur

En appliquant cette modélisation, toujours aux données mensuelles moyennes, mais minimales, les résultats sont nettement moins performants. Le coefficient de détermination n'est plus que de 0,62. Et la figure 3 montre un écart important entre les valeurs réelles et les valeurs prédites, notamment pour les mois les plus chauds et les plus froids. Cet écart se creuse même au fil du temps.

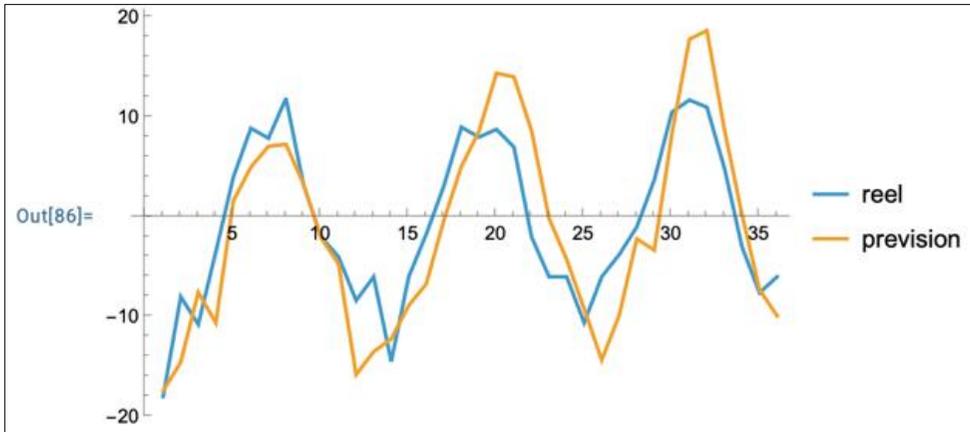


Figure 3. Données test réelles et prévues par le modèle pour les moyennes mensuelles minimales à Cluj-Napoca

Source : auteur

En outre, contrairement à une hypothèse plausible, les modèles établis à l'échelle de la semaine conservent une qualité satisfaisante. Les coefficients de détermination pour les données maximale (0,85) et minimales (0,80) sont supérieurs à celui fourni pour les données mensuelles minimales (0,62). Cependant, on observe quelques grands écarts, supérieur à 5°C au cœur des saison froide et chaude. Et, la déviation moyenne est de l'ordre de 3°C.

En climat méditerranéen, à Nice, les modèles ont un pouvoir de prédiction légèrement supérieur à ceux de Cluj-Napoca. De plus, les déviations moyennes, sont deux fois moins élevées, de l'ordre de 1,5°C. Pour les températures mensuelles maximales, les deux courbes des températures réelles et simulées par le modèle d'apprentissage automatique sont toujours relativement proches (figure 4).

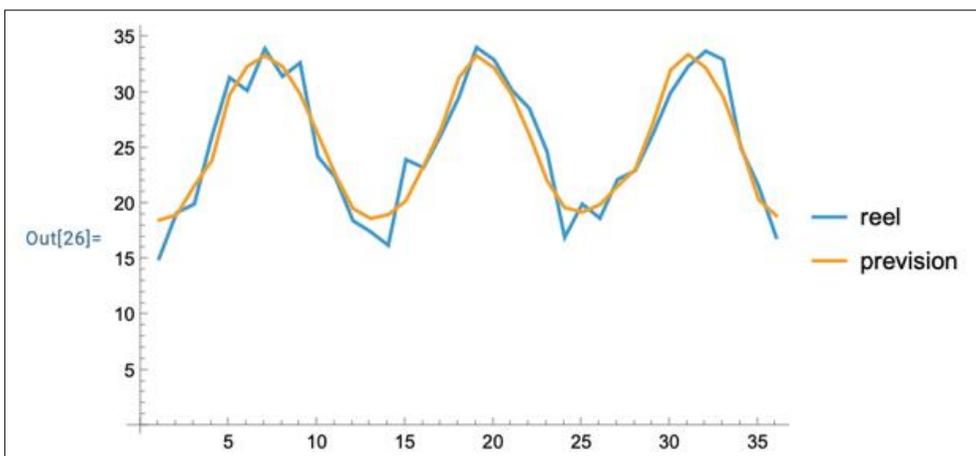


Figure 4. Données test réelles et prévues par le modèle pour les moyennes mensuelles maximales à Nice

Source : auteur

En revanche, ce modèle est nettement moins performant quand il est appliqué aux températures mensuelles moyennes minimales. Le coefficient de détermination n'est que de 0,70. La figure 5 montre que pratiquement toutes les températures prédites sont au-dessus de la diagonale, indiquant que ces températures sont inférieures aux températures réelles.

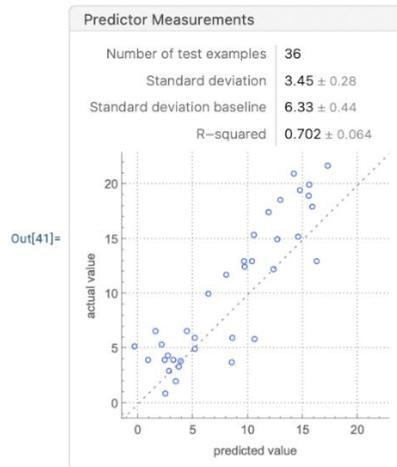


Figure 5. Graphique de comparaison des températures mensuelles maximales réelles et prévues à Nice
Source : auteur

Enfin, à l'échelle de la semaine les modèles de des températures maximales ou minimales sont plus efficaces qu'à Cluj-Napoca. Pour les températures maximales le coefficient de détermination reste supérieur à 0,87. Et surtout, la déviation moyenne est voisine de 1,5. La figure 7 illustre bien cette qualité des prévisions.

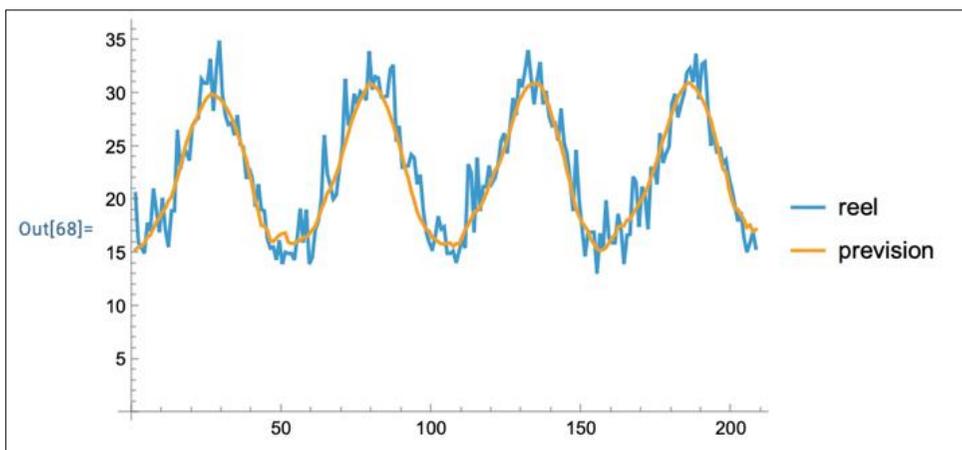


Figure 7. Données test réelles et prévues par le modèle pour les moyennes maximales par semaine à Nice
Source : auteur

En conclusion, le modèle proposé donne des résultats légèrement plus probants pour la station méditerranéenne que pour la station semi-continentale.

2.2. Améliorer et généraliser ce modèle

Pour améliorer ce modèle et obtenir des résultats plus précis qui faciliteraient les prises de décisions des acteurs, plusieurs pistes s'offrent au climatologue. L'accroissement de la série des données est une première solution admise par tous les spécialistes de l'apprentissage automatique. Cependant, cette démarche ne présente pas de garantie dans ce cas précis, car le climat évolue. Ce qui introduirait un biais difficile à corriger.

La seconde voie consiste à tester différents réseaux de neurones récurrents. D'abord, il est possible d'inclure une ou plusieurs différentes couches de neurones. Mais il est tout aussi facile de changer la valeur des hyperparamètres du réseau de neurones que nous avons retenu dans cette recherche. Ce qui revient à modifier, dans les deux cas, la seule ligne d'instruction :

```
net = NetChain[{GatedRecurrentLayer[20], LinearLayer[1]}, "Input" -> {50, 1}, "Output" -> 1]
```

Le lecteur est encouragé à réaliser plusieurs tests. D'une part, il peut modifier la taille du vecteur de sortie de la GatedRecurrentLayer, fixée à 20 dans le programme. Il est possible de remplacer cette valeur et d'améliorer le pouvoir prédictif du modèle. De plus, le climatologue peut puiser dans les différentes options relatives aux deux couches qui charpentent ce modèle.

D'autres généralisations sont envisageables. Quand l'objectif n'est pas uniquement prédictif, et que le géographe souhaite comprendre l'impact sur l'évolution des températures de mécanismes appelés connexions ou oscillations climatiques, il convient d'assembler d'autres ensembles de données d'apprentissage et de test. Pour les climats européens, il convient d'inclure les trois oscillations Nord-Atlantique (NAO), multi décennale (AMO) et ENSO. Ces deux approches conduisent à construire des ensembles différents de données, Pour un mois donné, chaque élément de l'ensemble d'apprentissage se présenterait sous la forme :

$$\{NAO_t, AMO_t, ENSO_t\} \rightarrow Temperature_t.$$

En sortie, outre les prévisions, il est possible d'apprécier l'impact de chaque oscillation sur le rythme thermique. Un exemple de cette approche est exposé dans l'ouvrage qui vient de paraître (Dauphiné A., 2025). L'oscillation AMO impacte plus fortement l'ensemble des valeurs de températures simulées à partir du modèle. Mais, en termes de prévision, les résultats sont moins performants. Cet exercice rappelle qu'il est parfois plus difficile de comprendre un phénomène que de le prévoir. Même si la compréhension et la prévision entretiennent des relations. Le géographe peut envisager une troisième approche qui garderait l'impact des oscillations, mais en

introduisant leur récurrence. Il conviendrait d'élaborer l'ensemble des données d'apprentissage selon le schéma ci-dessous :

$$\{NAO_{t-1}, NAO_t, AMO_{t-1}, AMO_t, ENSO_{t-1}, ENSO_t, Temperature_{t-x} \dots Temperature_{t-1}\} \rightarrow Temperature_t$$

Cet ensemble de données d'apprentissage signifie que la température au temps t dépend des x températures antérieures ainsi que des trois oscillations aux temps t et $t-1$.

Divers services météorologiques s'emparent de ces approches et ils proposent d'autres généralisations. Leurs recherches portent sur des types de temps décrits par les températures, les précipitations et les vents. Certains météorologues estiment que ces modèles couplés à l'emploi de photographies satellitaires complèteront ou remplaceront dans le futur les modèles numériques physiques. En effet, de nombreux modèles sont testés dans les grands centres météorologiques. Citons l'exemple du modèle *GraphCast* réalisé avec un réseau de neurones *DeepMind* de Google. Il permet d'établir des prévisions météorologiques pour les dix prochains jours en quelques minutes. Cet apprentissage automatique s'appuie sur des réseaux de neurones à graphes (Graph Neural Networks ou GNN). Les données d'entraînement, d'entrée et de sortie sont en fait un ensemble de données structurées en graphes, dont les nœuds sont des espaces et les arêtes représentent les interactions entre ces espaces. Cependant à ce jour, ces modèles concernent seulement la météorologie ou les émissions de gaz à effet de serre en climatologie.

Conclusion

Pendant longtemps les modèles statistiques et probabilistes furent privilégiés par les climatologues. Les modèles physiques numériques, d'abord utilisés en météorologie, se diffusent dans le champ de la climatologie, notamment dans les études relatives au réchauffement du climat. Mais, cet exemple de traitement par apprentissage automatique montre qu'il est possible de réaliser des prévisions acceptables en utilisant l'intelligence artificielle.

Déclaration de divulgation

Aucun conflit d'intérêts potentiel n'a été signalé par l'auteur.

ORCID

André DAUPHINÉ  0000-0002-9785-8470

References

- Chylek, P., Klett, J., D., Lesins, G., Dubey, M. K. & Hengartner, N. (2014). The Atlantic Multidecadal Oscillation as a dominant factor of oceanic influence on climate. *Geophysical Research Letters*, 41, 1689–1697. <https://doi.org/10.1002/2014GL059274>
- Dauphiné, A. (2025). *Une valse à mille temps, le réchauffement du climat*. Paris : Les Impliqués.
- Dauphiné, A. (2025). *Intelligence artificielle et géographie, avec le langage Wolfram Mathematica*. Éditions Universitaires Européennes.
- Emanuel, K. (2019) The relevance of Theory for Contemporary Research in Atmospheres, Ocean and Climate, *AGU Advances*.
- Gentleman, A. & Rood R. B. (2016), *Demystifying Climate Models*. Berlin : Springer Open.
- IPCC (1990, 1995, 2001, 2007, 2013, 2021). *Climate Change : The Physical Science Basis*.
- Le Treut, H. (2003). Les scénarios globaux de changement climatique et leurs incertitudes, *CR de l'Académie des sciences*.
- Laguës, M. & Lesne, A. (2008). *Invariances d'échelle*. Paris : Belin.
- Lovejoy, Sh. (2019). *Weather, Macroweather and the Climat : Our random yet predictable Atmosphere*. Oxford : Oxford University Press.
- Lovejoy, Sh. (2022). The Future of Climate Modelling : Weather Details, Macroweather stochastic or both?, *Meteorology* 2022, 1(4), 414-449. <https://doi.org/10.3390/meteorology1040027>
- McGuffie, K. & Henderson-Sellers, A. (2014). *Climate Models Primer*. Chichester : Wiley.
- Nichol, J. J. (2021). Machine learning feature analysis illuminates disparity between E3SM climate model and observed climate change, *Journal of Computational and Applied Mathematics*, 395, 2021, 113451. <https://doi.org/10.1016/j.cam.2021.113451>
- Poitou, J., Braconnot, P. & Masson-Delmotte V. (2014). *Le climat : La terre et les hommes*. Les Ulis, Edp sciences.
- Privalesky, V. (2021). *Time Series Analysis in Climatology and Related Sciences*. Berlin : Springer.
- Qasmi, S. (2018). *Sensibilité du climat européen à la variabilité multidéennale de l'Atlantique Nord* [Thèse]. Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, Université de Toulouse.
- Webber, Ch. L., & Marwan, N. (2015). *Recurrence Quantification Analysis*. London: Springer.
- Wei, W. W. S. (2005). *Time Series Analysis: Univariate and Multivariate Methods*. London: Pearson.