



Rapport d'ingénieure : Prédiction de la température de l'eau dans les rivières en fonction de paramètres hydroclimatiques

Auteure

Gabrielle COGNOT

Encadrement

Kévin DUPLAN (CD31)

Mohamed SAADI (IMFT)

Hélène ROUX (IMFT)

Remerciements

Merci à toutes les personnes qui ont permis à ce stage de fin d'études de se dérouler dans les meilleures conditions possibles. Tout d'abord, je remercie Kévin Duplan, Olivier Louis et Vincent Ribot du service ressource eau au conseil départemental de Haute-Garonne pour la proposition du projet, l'intérêt qu'ils y ont porté et le temps qu'ils m'ont accordé pour tenir des réunions constructives malgré des contraintes d'emploi du temps chargées, ainsi que Laurent Labbouz du bureau d'études Eaucéa pour l'obtention et le prétraitement des données de thermie de rivière. Je remercie également Mohamed Saadi et Hélène Roux, mes tuteurs de stage à l'Institut de Mécanique des Fluides de Toulouse (IMFT), ainsi que Ludovic Cassan. Merci aux membres du pôle échohydraulique et aux personnels permanents et non-permanents de l'IMFT pour m'avoir accueillie ces six derniers mois.

Résumé

La température des cours d'eau est un paramètre d'intérêt écologique majeur, gouvernant des processus de l'échelle microscopique à l'échelle des communautés aquatiques, mais aussi d'intérêt économique en raison de son impact sur la production d'eau potable et d'autres processus industriels. Dans un contexte de changement climatique, différents acteurs, chercheurs comme gestionnaires de la ressource, portent un intérêt à la surveillance et la prévision des températures estivales. Ce stage, dans le cadre du Projet de Territoire Garon'Amont (PTGA) piloté par le conseil départemental de Haute-Garonne, vise à étudier sur ce territoire la possibilité de l'utilisation de réseaux de neurones récurrents à mémoire long et court terme (LSTM) pour la simulation de température de l'eau en fonction de différentes variables prédictives, à savoir la température de l'air, le débit, les précipitations et l'évapotranspiration potentielle. Les résultats obtenus sont comparables pour des modèles locaux et régionaux et semblables à la littérature ($MAE=0,8$ °C). Les bonnes performances de modèles régionaux se basant seulement sur des forçages climatiques ouvrent la voie à une application de tels modèles sur l'ensemble du réseau hydrographique de la Garonne amont et à une connaissance approfondie de ses températures futures.

Extended abstract

Introduction

Stream temperature is a key parameter for river ecology at all living scales (Caissie, 2006; Webb *et al.*, 2008), from chemical, physical and biological processes to the life and death of aquatic fauna, especially poikilotherm organisms (Rivière *et al.*, 2021; Larnier, 2010). Summer stream temperatures therefore determine the survival of several aquatic species such as trouts, and are of economical interest to man, as high temperatures undermine drinking water production and its safety, electricity production from nuclear power plants, and other industrial processes. In addition, winter stream temperature is of ecological concern, enabling invasive species to thrive. Multiple water and energy fluxes drive stream temperature, that are either at the basin scale or at the reach scale. These drivers are well described in Leach *et al.* (2023), are dependent on each other, and are sensitive to change (Ouellet *et al.*, 2020), as shown by increasing trends over multiple streams (Larnier, 2010; Hannah et Garner, 2015; Seyedhashemi *et al.*, 2022).

Thus, in a rapidly changing climate, accessing to a fine comprehension of the evolution of this variable, spatially as well as temporally, is crucial. Unfortunately, only few rivers are monitored, either in France or globally (Gallice *et al.*, 2015; Rahmani *et al.*, 2021b; Rajesh et Rehana, 2022) and many of them for only for parts of the year (Beaufort *et al.*, 2020). Despite stream temperature being typically measured as part of routine sampling for other purposes, its values are often questionable (Dugdale *et al.*, 2017; Hannah et Garner, 2015). Inconsistencies are to expect in the data and more standardized networks or the use of teledetection methods (Petitjean, 2022; Dugdale *et al.*, 2017) could improve data quality.

Many modeling approaches have been adopted to overcome this severe lack of data. The most basic one is to consider the more available air temperature as a proxy for stream water temperature, as used in Wanders *et al.* (2019) and stated in Seyedhashemi *et al.* (2022). Alternatively, process-based models have been implemented to physically interpolate the existing measurements and project the future evolution of stream water temperature (see for example Seyedhashemi *et al.* (2022)). Among the data-driven approaches powered by Artificial Intelligence (AI), neural networks have gained popularity in the latest years in multiple scientific fields as well as in water sciences, as reported in Shen (2018).

A particular Recurrent Neural Network (RNN) named Long Short-Term Memory (LSTM, Hochreiter et Schmidhuber (1997)) is one of the most suitable architectures for the tasks of predictions involving chronological data (or time series). These networks share the ability to learn, with little knowledge on physics, dependencies linking climatic forcings and water temperature temporal patterns, as supported by Rahmani *et al.* (2021a). Thus, the much higher availability, with high resolution, of the former motivated us to explore their abilities to provide reliable estimations, in a nearer (forecast) or further (climate projections) future, of stream temperature even at ungauged streams.

In this work we focus on the upper Garonne river in France, as part of Projet de Territoire Garonne Amont (PTGA), a water management project overseen by the Haute-Garonne departmental council. Using a set of 40 gauge stations, we aimed at answering the following questions: How do the hyperparameters of the LSTM model (especially the lookback) impact the quality of the estimated stream water temperature? To what extent does the discharge data improve the performances of the LSTM? In case of unavailable discharge measurements, can they be replaced with catchment-scale climate forcings (precipitation, potential evapotranspiration)?

Data

Time series of daily water temperature over 40 stations of the upper Garonne river and its tributaries from their springs to the city of Toulouse were collected from 5 different sources and span from less than 1 year (for 7 of them) to 31 complete years. These time series cover more than 10 years for only 18 stations, and the rate of missing data is 21% on average over the dataset". Input data consists of local air temperature T_{air} from Météo-France's Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige (SAFRAN ; Vidal *et al.* (2010)) database, discharge Q from Hydroportail, formerly known as Banque Hydro (Dufeu *et al.*, 2022), basin-integrated precipitation P from Météo-France's COmbinaison en vue de la Meilleure Estimation de la Precipitation HORaiRE (COMEPHORE ; Tabary *et al.* (2012)) and basin-averaged evapotranspiration ETP (calculated with the formula from Oudin *et al.* (2005), using basin-averaged SAFRAN air temperature).

Before testing our models, we conducted a preliminary analysis of the multiannual averages of the maximum annual temperatures smoothed over windows of 7 days and 30 days. This analysis revealed a strong spatial gradient in average maximum temperatures, with lower temperatures (less than 17 °C on mean over 7 days and only occasional peaks above 19 °C) upstream of the Garonne catchment and higher temperatures downstream at Toulouse (up to 25 °C on mean over 7 days and 60-80 days above 19 °C). A Mann-Kendall seasonal test was applied over the 18 water temperature time series spanning more than 10 years (including 3 that actually had a little less than 10 years), highlighting significant ($p < 0.05$) increasing trends over 16 of them. Ranging from +0.2 °C/decade for warmer waters downstream to +0.5 and up to +1.3 °C/decade upstream, these trends suggest a high sensitivity of the fresh Pyrénées rivers and their ecosystems to climate warming. These trends reflect similar trends from local air temperature over the same years.

Methodology

To carry out this study, we implemented our LSTM RNN models using the opensource Python library *PyTorch* (Meta, Paszke *et al.* (2019), version 2.0.0). These models function in a similar way to all neural networks, where information passes through a network of a certain number of layers (here only 1 was used) made of a certain number of nodes (here structures of *neurons* called *cells* that work as detailed in Kratzert *et al.* (2018) and Feng *et al.* (2020)) we defined as 64, 128 or 256. Each neuron has a weights and biases matrix of parameters that need to be tuned first. Overall structure and training parameters are called hyperparameters to distinguish them from these trainable parameters. As a crucial hyperparameter, long-term dependencies are learnt by feeding time series of a certain length called *lookback* instead of unique time points, as the network learns short-term patterns by feeding the output of the last time step to the *recurrent* cell.

Training requires feeding *batches* of data one by one through the network, calculating the sum of errors during this step, and backpropagate this error by a *learning rate*. One such cycle, called an *epoch*, is repeated until a set number of epochs runs out, or until a criterion is reached. Each epoch, cells had a set probability (or *dropout rate*=0.2) to be disconnected and not be updated, which helps for better convergence. For this training, from each station, the first 70 % of the data was used. Half of the remainder was used to monitor progress as validation data. The rest was kept for evaluation.

Here, several models were trained. Local models as well as regional models were created, using either a different model for each station or a model that is trained on every station. Hyperparameters were chosen out of a set of (1) 14,30 or 60 (or 7 for the regional models) lookback days, (2) a batch size of 64 or 256, (3) 64, 128 or 256 cells. As for the predictive variables, they were chosen as either T_{air} only, $T_{air,min}$ & $T_{air,max}$, T_{air} & Q or T_{air} & P & ETP . Water temperature from the previous days was purposely not used in these models,

enabling future development for prediction on unmonitored sites. All combinations of the above model types, hyperparameters and predictive variables were tested.

LSTM Models were trained by minimizing mean square error (MSE) until either a counter of 1000 epochs had been reached or until the validation MSE had stopped decreasing for 20 epochs. Results were evaluated by mean absolute error. To evaluate the ability of the LSTM model to reproduce high temperature values, we also estimated the mean absolute error using only the highest 25% temperatures of the test time series and the precision regarding the prediction of surpassing the threshold of 19°C.

One of our stations, the Bazacle in Toulouse, which is the most downstream station, is of high interest because of the presence of a major drinking water plant collecting river water nearby and because of its characteristic location in a densely populated area. For this station, a linear model using minimum and maximum forecasted local air temperatures as well as discharge, is used from June to September and its performances were compared to ours.

Results and discussions

The results of the tested LSTM models were satisfactory. The found MAE were much higher than those of Feigl *et al.* (2021) (0.4 °C for the Danube) but in consistency with the ones found for various machine learning models over seven rivers in Europe and the USA, ranging from 0.4 °C to 1.2 °C, by Zhu *et al.* (2019b). The results of the parametric test showed that, regarding MAE, better hyperparameters included a lookback of 60 days, a larger batch size of 256 and a lower number of cells of 128. The latter two had little impact on model performances, contrary to the findings of Feigl *et al.* (2021). Overall, better performances in terms of MAE on the test datasets were obtained with (T_{air}, Q) or (T_{air}, P, ETP) as input variables, which highlights the importance of catchment-scale variables for the prediction of stream water temperature in addition to point-scale air temperature. Precision regarding the 19 °C threshold did not prove any use analysing the results, its distributions almost identical for every hyperparameter and set of variables tested as this threshold is rarely passed for most of the stations. These parameter sets gave a median MAE of 0.7 °C and 0.8 °C for local models and 0.8 °C and 0.9 °C for regional models for (T_{air}, Q) and (T_{air}, P, ETP) . Models based on (T_{air}, P, ETP) had the same median but a shorter range of distribution, thus better confidence, than models based solely T_{air} .

These parameters were tested on the Bazacle station (the Garonne at Toulouse, data from Migrateurs Dordogne Garonne Charente Seudre (MIGADO)) from 21-11-2015 to 01-01-2019. Surprisingly, the results showed better performances of regional models with a MAE of 1.1 °C for both of them versus a MAE of 1.4 °C (using Q) and 1.3 °C (using P and ETP) for local models, despite this station not being representative of the study area. 25%-highest temperatures were less well reproduced, with MAE of 1.5 to 2.0 °C.

The simple forecasting model previously used showed great performances, with a MAE of 0.4 °C for a next-day forecast that doesn't get worse if only the 25%-highest values are assessed. This can be explained by the use of the water temperature from the previous days as input variables, which makes use of the high autocorrelation in the water temperature time series. However, this limits their use for stations where no record of water temperature is available. Indeed, as water holds a lot of inertia, using the temperature from the current day as forecast for the next is highly accurate, with a MAE of 0.5°C (0.6°C on 25%-highest values).

Conclusion

Our study has shown that using precipitation and evapotranspiration data improved performances compared to using only air temperature and gave promising results, extending the applicability of our approach to stations

with measurement of neither water temperature nor discharge. Using discharge data further improved model performances when combined with air temperature data. Further work could include comparing the models used here to benchmarking models and incorporating more relevant basin attributes, such as riparian vegetation, to the model to analyze their impact on stream water temperature and its dynamics. The good performances of climate forcing-only based regional models could open the way to a more comprehensive mapping of the temperatures to expect in the future over the upper Garonne basin without the need of gauging stations.

Table des matières

Introduction	10
1 Synthèse bibliographique	11
1.1. La température de l'eau de rivière	11
1.1.1. Importance pour le vivant et pour l'Homme	11
1.1.2. Un paramètre aux sources de variabilités nombreuses et sensible aux changements environnementaux	12
1.2. Difficultés récurrentes dans la mesure et intérêt de la modélisation	14
1.2.1. Des réseaux de mesure à la modélisation	14
1.2.2. Modèles à base physique	15
1.2.3. Modèles basés sur les données ou modèles statistiques ; réseaux de neurones	16
1.3. Synthèse et positionnement du stage	17
2 Cas d'étude et données	18
2.1. Zone d'étude : la Garonne amont	18
2.2. Données collectées	19
2.2.1. Observations de température de l'eau	19
2.2.2. Entrées pour les modèles	20
2.3. Analyse et critique de données	21
2.3.1. Variabilité spatiale des indicateurs sur la Garonne amont	22
2.3.2. Tendances observées dans l'évolution temporelle	24
3 Méthodes	28
3.1. Les modèles LSTM	28
3.1.1. Des neurones aux cellules	28
3.1.2. Réseaux de neurones et apprentissage	29
3.2. Implémentation et métriques de performance	31
3.3. Modèles construits	32
3.3.1. Modèles LSTM	32
3.3.2. Modèle linéaire pour une station	33
4 Résultats et discussions	35
4.1. Hyperparamètres et variables prédictives optimaux	35
4.1.1. Tests des hyperparamètres	35
4.1.2. Choix des variables et du type de modèle	39
4.2. Illustration sur une station : le Bazacle à Toulouse	39
4.2.1. Modèles LSTM	39
4.2.2. Comparaison avec le modèle de prévision	42
5 Conclusion et perspectives	44
Références	45
Annexes	50
A Commentaires sur le critère de précision	50
B La Garonne à Valentine	52
C Outil de prévision opérationnel du CD31	54

D Utilisation de la température au jour précédent : exemple du Bazacle 55

Liste des abréviations récurrentes

BV	Bassin Versant
CD31	Conseil Départemental de la Haute-Garonne
CNPE	Centre Nucléaire de Production d'Electricité
COMEPHORE	COmbinaison en vue de la Meilleure Estimation de la Precipitation HOraiRE
DCE	Directive Cadre sur l'Eau
DL	<i>Deep Learning</i> ou apprentissage profond
EDF	Électricité De France
ETP	Evapo-Transpiration Potentielle
IA	Intelligence Artificielle
IMFT	Institut de Mécanique des Fluides de Toulouse
LSTM	<i>Long Short-Term Memory</i>
MAE	<i>Mean Absolute Error</i> ou erreur moyenne absolue
MIGADO	Migrateurs Garonne Dordogne Charente Seudre
ML	<i>Machine Learning</i> ou apprentissage automatique
MSE	<i>Mean Square Error</i> ou erreur quadratique moyenne
PTGA	Projet de Territoire Garon'Amont
RMSE	<i>Root Mean Square Error</i> ou racine de l'erreur quadratique moyenne
RNN	<i>Recurrent Neural Networks</i> ou réseaux de neurones récurrents
RSA31	Réseau de Suivi et d'Alerte 31
TIGRE	Thermie en rivière : analyse Géostatistique et description de REgime
SAFRAN	Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige

Introduction

La température de l'eau dans un cours d'eau ¹ est un paramètre prépondérant pour le fonctionnement écologique de celui-ci et pour toutes les échelles du vivant. Ses valeurs estivales extrêmes sont déterminantes pour la survie de nombreuses espèces aquatiques, ainsi que pour certains usages humains liés aux prélèvements dans les eaux de surface, comme la production d'eau potable. Il est donc important de disposer d'une fine connaissance de ce paramètre à la fois dans l'espace et dans le temps afin de mieux quantifier son évolution future dans un contexte de fort dérèglement climatique. Face à ces enjeux, des réseaux de mesure existent, mais présentent d'importantes lacunes dans leurs dimensions à la fois spatiales et temporelles. Pour pallier ces manques, il existe différentes approches de modélisation, dont de nombreuses basées sur la donnée, parmi lesquelles gagnent en popularité depuis plusieurs années les réseaux de neurones. En effet, ces réseaux de neurones ont le potentiel d'apprendre, sans beaucoup de connaissances sur la physique du système, la relation entre la température de l'eau en un point donné et les forçages climatiques en ce point ainsi qu'à l'échelle du bassin versant drainé par ce point. La forte disponibilité de données de forçage climatique à haute résolution nous motive à explorer cette voie afin de reconstituer et projeter dans le futur proche (prévisions) ou plus lointain (projections climatiques) l'évolution de la température de l'eau dans des cours d'eau même non instrumentés (ou non jaugés).

L'objectif de ce stage est de répondre aux questions suivantes :

1. Quel est l'apport des réseaux de neurones et en particulier les réseaux *Long Short-Term Memory* (LSTM) pour la prédiction de température de l'eau de rivière par rapport à un modèle plus simple (type régression multilinéaire) ? Quels paramètres permettent d'obtenir les meilleures performances, et en particulier, quelle durée d'observations ?
2. Quelle est la valeur ajoutée du débit par rapport à la température de l'air seule dans ce type de modèle ?
3. Au cas où l'on aurait pas accès aux mesures de débit, pourrait-on le remplacer par les forçages de précipitations et d'évapotranspiration à l'échelle du bassin versant en amont de la station ?

Le rapport s'articule en cinq parties : tout d'abord sera présenté le contexte scientifique de ce stage au travers d'un résumé bibliographique. Ensuite, sera présenté le cas d'étude et les données utilisées. Puis sera détaillée la méthodologie employée. Enfin, les principaux résultats seront présentés et analysés avant de conclure dans une dernière partie.

¹Par la suite, le terme "température", sauf précision, se réfère à la "température de l'eau de rivière".

1. Synthèse bibliographique

1.1. La température de l'eau de rivière

1.1.1. Importance pour le vivant et pour l'Homme

La température de l'eau est un paramètre essentiel pour le fonctionnement écologique des cours d'eau, contrôlant les processus chimiques, physiques et biologiques (Caissie, 2006; Webb *et al.*, 2008). Parmi ces processus, on retrouve les cycles des nutriments (azote et carbone) et la disponibilité de l'oxygène dissous (Leach *et al.*, 2023; Rivière *et al.*, 2021). Les organismes aquatiques, et plus particulièrement ceux poikilothermes² sont sensibles à la température du milieu et leur métabolisme, leur développement ainsi que leur distribution spatiale le long d'un cours d'eau sont soumis à celle-ci (Ouellet *et al.*, 2020; Larnier, 2010), se répartissant selon leur affinité sur le gradient amont-aval de température. La température de l'eau est ainsi un meilleur prédicteur de la structuration des communautés de poissons que la température de l'air souvent utilisée comme substitution selon Picard *et al.* (2022). Une température de l'eau élevée est aussi favorable au développement de certaines cyanobactéries potentiellement nocives (Cha *et al.*, 2017).

Outre la fonctionnalité des écosystèmes, dont la valeur ne saurait être estimée, la température de l'eau de rivière a une incidence sur diverses activités humaines à caractère économique, récréatif (pêche) comme industriel (production d'eau potable, production d'électricité, autres processus industriels ; voir par exemple Rivière *et al.* (2021). Webb *et al.* (2008) relèvent que si de nos jours l'attention des Hommes se porte principalement sur la surveillance des températures élevées en raison de ces implications écologiques et économiques, autrefois la problématique de la prise en glace des axes de navigation continentaux majeurs a pu motiver l'installation des premières stations de mesure comme en Autriche à la fin du XIX^{ème} siècle. Les pêcheurs, en dépendant et en observant l'abondance des ressources piscicoles, sont parmi les premiers à observer l'impact de la température d'un cours d'eau dans leur activité, et représentent des acteurs-clé dans l'obtention de cette donnée, via les fédérations départementales de pêche et de protection du milieu aquatique et l'association Migrateurs Garonne Dordogne Charente Seudre (MIGADO). Du côté de la production d'eau potable, la température de l'eau a jusque récemment été limitante, un seuil de prélèvement remplacé par une recommandation en début d'année 2023 étant auparavant fixé à 25 °C par l'Agence Régionale de Santé (ARS). A titre d'exemple, en Haute-Garonne, 91% de l'eau potable provient des eaux de surface (contre 33% pour la moyenne nationale française), et ces dernières années, l'alimentation en eau potable de la métropole toulousaine aurait été plus souvent menacée par cet impératif sur la production que par une réelle indisponibilité volumétrique selon le rapport annuel 2021-2022 du Projet de Territoire Garon'Amont (PTGA) (Labbouz, 2022). Enfin, la production d'électricité par un Centre Nucléaire de Production d'Electricité (CNPE) nécessite d'importantes quantités d'eau de refroidissement, la plupart du temps captée puis restituée par la centrale dans un cours d'eau proche. Cette eau restituée est soumise à réglementation de température selon l'arrêté du 2 février 1998 relatif aux prélèvements et à la consommation d'eau ainsi qu'aux émissions de toute nature des Installations Classées pour la Protection de l'Environnement (ICPE) soumises à autorisation (ATEP9870017A, consulté le 22/06/2023).

Pour souligner l'importance de la température de l'eau dans le suivi de la santé des cours d'eau, plusieurs

²Organismes ne régulant pas leur propre température (à «sang froid») : poissons, amphibiens, reptiles

statistiques sont considérées afin de caractériser le régime thermique de ceux-ci. En particulier, on considère généralement deux températures - seuil : 19 °C et 25 °C. Ces températures correspondent au seuil de stress physiologique et au seuil léthal pour la truite, pour lesquelles on examine 1) le nombre de jours dans l'année où ces températures sont dépassées et 2) si ces seuils sont dépassés plusieurs jours consécutifs, grâce à des indicateurs répandus comme la moyenne inter-annuelle des températures moyennées sur 30 jours glissants maximales tels que chez Beaufort *et al.* (2020) et dans le rapport annuel pour le PTGA (Labbouz, 2022).

1.1.2. Un paramètre aux sources de variabilités nombreuses et sensible aux changements environnementaux

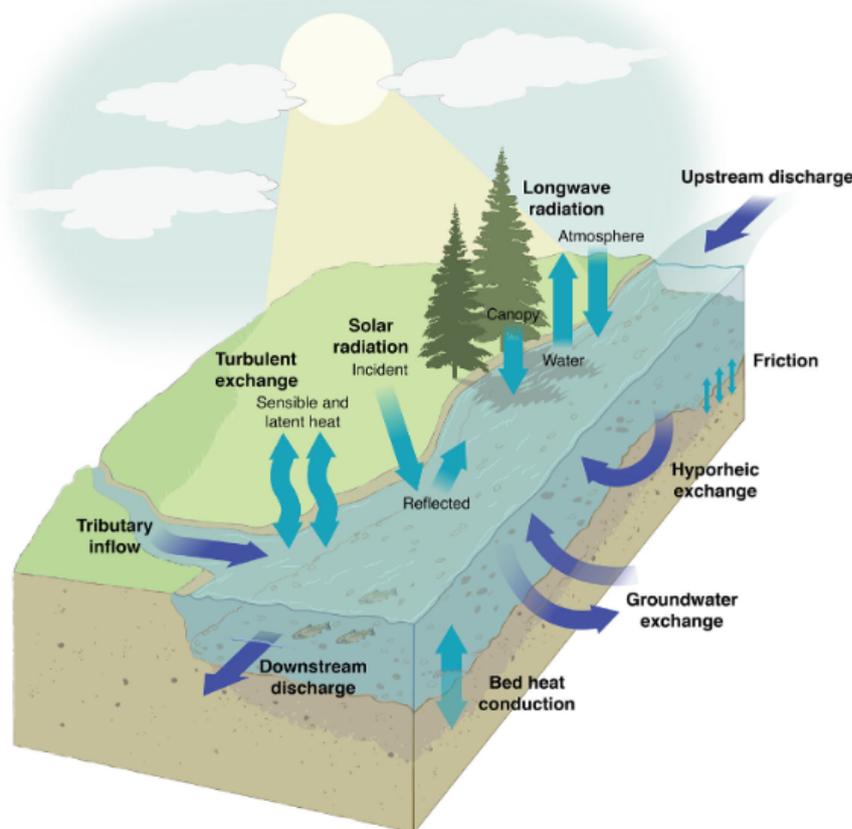


FIGURE 1.1.1 : Échanges influant sur la température de l'eau de rivière. Schéma de Leach *et al.* (2023).

La température d'un cours d'eau en un point est l'expression du bilan d'eau et d'énergie non seulement à l'échelle du tronçon, mais également de tout le bassin versant en amont de ce point (Leach *et al.*, 2023). Etant dépendante de nombreux échanges d'eau et d'énergie ponctuels (l'apport frais d'une source par exemple) comme diffus à l'échelle du bassin versant (le rayonnement solaire), elle est sensible à des sources de variabilité souvent interdépendantes, auxquelles il n'est pas toujours aisé d'imputer les évolutions observées (Hannah et Garner, 2015). De plus, la température de l'eau est susceptible d'être très hétérogène y compris localement, dépendamment de la morphologie du lit et de l'ombre projetée par la ripisylve. Cette dernière diminue sensiblement la température des petits cours d'eau, mais a peu d'effet sur les cours d'eau plus larges qui ne peuvent être totalement ombragés (Loicq *et al.*, 2018; Seyedhashemi *et al.*, 2022).

Sur les grands cours d'eau, les échanges thermiques eau-air sont largement prépondérants dans la température de l'eau (Dugdale *et al.*, 2017). Ainsi, bien que pour une étude sur la Seine les températures de l'air journalières maximale, minimale et moyenne soient prépondérantes, il est supposé que d'autres paramètres climatiques puissent avoir une importance non négligeable dans le cas des cours d'eau d'ordre inférieur et moins impactés par les ouvrages hydrauliques (Rivière *et al.*, 2021).

Le débit, en modulant la capacité thermique de l'eau, joue également sur la température de l'eau, particulièrement lorsque celui-ci (ou la profondeur de l'eau) est faible, à l'étiage ainsi que dans les têtes de bassin (Zhu *et al.*, 2019b), mais a peu d'importance lorsque le débit est élevé (Gu *et al.*, 1998). Cette relation entre le débit à l'étiage et la température est par ailleurs étudiée afin de définir les débits écologiques³ de manière plus adéquate par Ferchichi *et al.* (2023). Les variations du régime de précipitations à l'échelle du bassin versant, et ainsi du régime de débit, ont aussi un impact sur les températures des cours d'eau. De la même manière, en influant sur la relation pluie-débit, un changement d'occupation du sol comme la déforestation, entraîne une modification du débit ruisselé ainsi que de sa température (Caissie, 2006).

La température de l'air et le débit étant les deux paramètres gouvernant principalement la température de l'eau, il est observé que leurs variations dans un contexte de changement climatique s'y répercutent (Ouellet *et al.*, 2020). Un réchauffement des cours d'eau est attendu en réponse au réchauffement de l'air, affectant les écosystèmes aquatiques terrestres de l'échelle microscopique à macroscopique. Une variation de régime de débit pourrait exacerber les tendances estivales d'augmentation si les débits diminuent et avec eux la capacité thermique de la rivière. En raison d'hivers plus doux et de la construction de retenues face à une pression anthropique sur la ressource en eau accrue agissant comme *hotspots* pour les espèces invasives, ces dernières pourraient se répandre dans le milieu naturel avec plus de facilité, exerçant une compétition et une prédation sur les espèces natives, lesquelles seront en concurrence entre elles-mêmes, déplacées par la modification de leur habitat (Rahel et Olden, 2008). Cependant les observations de température de rivière sont rares et cela freine l'analyse de tendances. Pour la Loire, Seyedhashemi *et al.* (2022) ont attribué l'augmentation de la température de l'eau de 1963 à 2019 à ce changement climatique et cette augmentation dépasse celle de l'air, plus rapidement sur le sud du bassin versant et sur les sections larges. Pour la Garonne à Toulouse, Larnier (2010) a estimé une augmentation en moyenne estivale de 2.8 °C et en moyenne annuelle de 0.8°C entre 1978 et 2005. Pour le Royaume-Uni, Hannah et Garner (2015) ont établi une augmentation tendancielle des températures de rivière entre 1990 et 2006 d'en moyenne 0.3 °C par décennie, malgré l'absence de consensus sur l'attribution de ces augmentations au changement climatique en raison de la complexité de la dynamique de la température de rivière, fruit de la réponse climatique et hydrologique modulée par les propriétés du bassin versants et les pressions anthropiques. Il est globalement de plus en plus probable qu'elle augmente en raison de modification de ces facteurs cités précédemment selon Ouellet *et al.* (2020).

Une partie de la variabilité de la température n'est pas imputable directement aux forçages atmosphériques. Ponctuellement, des rejets d'eau réchauffée par une installation industrielle ou une centrale nucléaire peuvent provoquer un échauffement de l'eau, notamment lors des étiages estivaux (Caissie, 2006). La température de l'eau en surface est aussi contrôlée par les échanges avec les eaux souterraines (voir la Figure 1.1.1). Ces eaux n'étant pas soumises aux échanges de chaleur en surface, elles présentent une température beaucoup plus stable, réchauffant le cours d'eau l'hiver et le refroidissant l'été, ce qui peut expliquer des différences de variabilité temporelle en fonction de la lithologie du cours d'eau (Hannah et Garner, 2015) ou de l'artificialisation de

³Une définition des débits écologiques est la suivante : les débits écologiques sont considérés dans le contexte de la Directive Cadre sur l'Eau (DCE) comme « un régime hydrologique compatible avec la réalisation des objectifs environnementaux de la DCE dans les masses d'eau naturelles de surface telles que mentionnées à l'article 4(1).

celui-ci lorsque le lit est privé de son matelas alluvial. La présence de retenues d'eau peut également perturber le régime thermique des rivières par rapport à des cours d'eau naturels (Auffray *et al.*, 2023), en influant directement sur la température de l'eau lâchée et indirectement par la modification du régime de débit (Zhu *et al.*, 2019a,b; Rahmani *et al.*, 2021b).

Ainsi, chercheurs comme gestionnaires dans des domaines-métiers variés de la biologie à l'hydrologie, l'ingénierie et la gestion de bassin versant sont concernés par l'impact des changements climatiques et environnementaux sur ce paramètre qu'est la température de l'eau de rivière, qui est à prendre en compte dans les études d'impact environnemental et pour une gestion efficace des ressources piscicoles (Caissie, 2006; Leach *et al.*, 2023).

1.2. Difficultés récurrentes dans la mesure et intérêt de la modélisation

1.2.1. Des réseaux de mesure à la modélisation

Les mesures effectuées afin de surveiller et caractériser la température de l'eau des rivières et leur dépendance aux facteurs présentés ci-dessus présentent la plupart du temps des lacunes dans deux dimensions : l'espace et le temps. Les archives de température de l'eau sont souvent à faible résolution et/ou de courte durée, rendant ardue l'extraction de liens cause-effet, de tendances ainsi que leur interprétation, comme rapporté sur la Garonne amont (Rapport annuel du PTGA; Labbouz, 2022). A l'échelle nationale, le nombre de stations ayant des observations sur des périodes concomitantes ne permet pas d'établir des statistiques robustes à long terme selon Beaufort *et al.* (2020). Bien qu'il s'agisse d'un paramètre estimé lors de routines d'échantillonnage de la qualité de l'eau et d'études écologiques, la qualité des mesures est très variable (Dugdale *et al.*, 2017; Hannah et Garner, 2015). La plupart des rivières, à l'échelle du territoire français comme à l'échelle globale, ne sont pas instrumentées, ce qui est relevé par de nombreux auteurs (Gallice *et al.*, 2015; Rahmani *et al.*, 2021b; Rajesh et Rehana, 2022) et pour celles qui le sont, beaucoup ne sont pas instrumentées toute l'année ou sont victimes de défaillance des sondes, provoquant l'apparition de valeurs aberrantes dans le signal souvent non détectées lors du processus d'acquisition (Beaufort *et al.*, 2020; Labbouz, 2022). La mise en place de réseaux de capteurs de façon standardisée et les méthodes de télédétection (Petitjean, 2022; Dugdale *et al.*, 2017) pourraient pallier une partie de ce problème.

En raison de ce manque de mesures, la température de l'air et ses variations portant la majeure partie de la dynamique de celle de l'eau sont biens souvent utilisées comme proxy, comme dans le modèle biogéochimique RIVERSTRAHLER dans l'étude de Ducharme (2008), ou comme approximation de la température de l'eau pour le calage du modèle global de Wanders *et al.* (2019). Des relations liant la température de l'eau à celle de l'air sont couramment utilisées (Bustillo *et al.*, 2014; Leach *et al.*, 2023), la seconde étant simple et peu chère à mesurer contrairement aux échanges d'énergie. Ces modèles ne peuvent pas représenter avec une précision exacte la température de l'eau, de par la nature spatialement ponctuelle des observations que l'on compare à une simulation supposée moyenne et mélangée verticalement (Wanders *et al.*, 2019).

Des approches de modélisation permettent alors de reconstituer la variabilité spatio-temporelle avec une résolution fine de la température de l'eau de rivière, de réaliser des prédictions de température, de tester des hypothèses permettant de comprendre les phénomènes physiques, comme l'impact de modifications telle que l'implantation d'une retenue (Seyedhashemi *et al.*, 2022) ou un changement d'occupation du sol, et de réaliser des projections futures de scénarios climatiques. Ces approches se divisent en deux familles : des modèles à base physique ou à représentation explicite, dérivant des processus d'échange d'énergie, et des modèles statistiques

ou à représentation implicite, basés sur les liens mathématiques entre les observations de la variable cible et des variables prédictives qui pourraient l'influencer sans recourir à des connaissances sur la physique du système étudié. La plupart des modèles recensés par Gallice *et al.* (2015) ont pour but d'estimer comme métrique des valeurs agrégées sur l'année telle que la température moyenne hebdomadaire maximale pour un cours d'eau, estimant ainsi une description du régime thermique plutôt que des valeurs précises, comme dans le projet Thermie en rivière : analyse Géostatistique et description de REgime (TIGRE) (Beaufort *et al.*, 2020). Cela étant, de nombreuses approches, à base physique comme statistique, permettent d'accéder à une résolution plus fine, de l'ordre de la journée (ce qui nous intéresse par la suite) (Feigl *et al.*, 2021; Rahmani *et al.*, 2021b,a) et plus rarement, de l'heure (Stajkowski *et al.*, 2020). L'emprise spatiale de ces modèles peut être à l'échelle d'un cours d'eau comme Larnier (2010), nationale comme le projet TIGRE, ou globale comme Wanders *et al.* (2019).

1.2.2. Modèles à base physique

Un modèle à base physique ou déterministe ou à représentation explicite est un modèle représentant directement des processus gouvernant le phénomène étudié via des équations issues de la physique et généralement vérifiées en laboratoire. Pour ce faire, le domaine étudié est discrétisé en des éléments géométriques sur lesquels les paramètres du domaine doivent être renseignés (approche distribuée). Dans le cas de la température de rivière, les processus qui doivent être représentés relèvent de deux catégories : les échanges d'énergie et les échanges en eau. La Figure 1.1.1 récapitule les échanges identifiés comme gouvernant la température de l'eau : des échanges en surface (rayonnements à longueurs d'onde courte et longue et atténuation par l'ombre projetée par la végétation, flux de chaleur latente et sensible), des échanges au niveau du lit (conduction, frottement, échanges nappe-rivière et échanges hyporhéiques), et des échanges par advection (débit du cours d'eau et apport des affluents) .

Des synthèses bibliographiques comme Dugdale *et al.* (2017) et Ouellet *et al.* (2020) recensent un certain nombre de tels modèles plus ou moins complexes. Le bilan d'énergie impliquant la température de l'eau de rivière peut être réalisé à l'échelle d'un tronçon de cours d'eau comme la Garonne (Larnier, 2010), régionale pour la Loire (Seyedhashemi *et al.*, 2022) ou globale (van Vliet *et al.*, 2012; Wanders *et al.*, 2019) selon le(s) objectif(s) de l'étude. D'un point de vue des performances, le modèle de Beaufort *et al.* (2016) Temperature-NETwork (T-NET) obtient une *Root Mean Square Error* ou racine de l'erreur quadratique moyenne (RMSE) moyenne de 1,6 °C pour la reproduction de chroniques journalières sur la période 2008-2012 et sur 128 stations situées dans le bassin de la Loire. Le modèle de Toffolon et Piccolroaz (2015) *air2stream*, souvent cité comme *benchmark* dans d'autres études de modélisation, simule la température journalière de l'eau en fonction de celle de l'air et du débit grâce à un bilan d'eau et d'énergie avec des paramètres calés sur les observations, avec une RMSE inférieure à 1°C pour trois cours d'eau suisses sur les périodes 2010-2012 et 2004-2013.

Ce type de modèle présente l'avantage de permettre d'explorer la réponse d'un système et sa sensibilité aux différents flux d'eau et d'énergie (Dugdale *et al.*, 2017). Ces informations peuvent ensuite être utilisées dans le développement de modèles statistiques à plus large échelle en renseignant sur la pertinence des paramètres et des phénomènes intégrés, comme relevé par Rivière *et al.* (2021). En effet Leach *et al.* (2023) établissent que les processus physiques sont aujourd'hui relativement bien connus. Certains modèles intègrent les effets de la ripisylve parmi des paramètres statiques du cours d'eau comme Loicq *et al.* (2018) et Seyedhashemi *et al.* (2022).

Cependant, un modèle à base physique exige une description géométrique fine du tronçon de rivière, ainsi que des temps de calcul qui peuvent devenir importants (Larnier, 2010), rendant ainsi prohibitive leur utilisation

dans un cadre opérationnel de prévision et/ou pour des cours d'eau pauvres en observations. Les paramétrisations peuvent aussi devenir difficiles à caler. De plus, il nécessite une connaissance fine de la géométrie du cours d'eau et de ses propriétés géopédologiques, laquelle n'est pas toujours disponible de manière homogène sur un nombre suffisant de points, malgré la disponibilité de plus en plus importante de la donnée.

1.2.3. Modèles basés sur les données ou modèles statistiques ; réseaux de neurones

Contrairement aux modèles à base physique, les modèles statistiques, dits de régression, explorent directement les liens entre les variables explicatives (climatiques par exemple) et la température de l'eau de rivière sans représentation explicite des mécanismes physiques. Ces modèles sont couramment utilisés en particulier pour les applications à large échelle (Seyedhashemi *et al.*, 2022). Ceux-ci visent à optimiser un certain nombre de paramètres en minimisant une certaine fonction objectif, c'est à dire une mesure de l'erreur commise par le modèle (par ex. *Mean Square Error* ou erreur quadratique moyenne (MSE)). Différents types d'approches, regroupées sous le nom d'Intelligence Artificielle (IA) et plus particulièrement de *Machine Learning* ou apprentissage automatique (ML), existent, de la simple régression linéaire, à d'autres approches plus complexes, parmi lesquelles des concepts de réseaux de neurones et de *Deep Learning* ou apprentissage profond (DL). Ceux-ci, bien qu'ayant émergé dans les années 1940 à 1960, regagnent en popularité depuis une vingtaine d'années (Cardon *et al.*, 2018) dans diverses disciplines scientifiques dont les sciences de l'eau (pour une synthèse approfondie, se référer au travail de Shen (2018)).

Parmi les modèles de DL, on trouve les *Recurrent Neural Networks* ou réseaux de neurones récurrents (RNN). Ce sont des réseaux de neurones particuliers prenant en compte des séquences de données dans leur prédiction, ce qui les rend particulièrement adaptés à la prévision de texte ou de séries temporelles comme la température de l'eau. Deux modèles sont le plus couramment utilisés : les *Gated Recurrent Unit* (GRU, Cho *et al.* (2014)) et les *Long Short-Term Memory* (LSTM) (Hochreiter et Schmidhuber, 1997). Les LSTM, comme leur nom l'indiquent, sont capables de retenir et utiliser grâce à une structure en «cellules» une information proche (*Short-Term Memory*) et une information éloignée dans le temps (*Long-Term Memory*). Ils ont été créés afin de répondre au problème de disparition/explosion du gradient lors de l'ajustement des paramètres (ou poids) pendant l'entraînement du modèle. Leur fonctionnement sera détaillé en 3.1..

Les réseaux LSTM) ont déjà été utilisés avec succès dans divers domaines des hydrosciences, notamment dans le cadre de la simulation de débit (Kratzert *et al.*, 2018; Hashemi *et al.*, 2022) et de qualité d'eau dans un lac (Barzegar *et al.*, 2020). Pour la simulation de température de rivière, Feigl *et al.* (2021) les ont comparés avec 5 autres modèles de ML courants (régression linéaire *step-wise*, *eXtreme Gradient Boosting* (XGBoost), *Feed-Forward Neural Networks* (FNNs), GRU) et 2 modèles *benchmark* (régression linéaire et *air2stream*). Cette comparaison est effectuée sur 10 bassins versants autrichiens de 200 à 96 000 km² et en utilisant des jeux de variables prédictives journalières différents parmi les forçages météorologiques (température de l'air, rayonnement solaire et précipitations) et le débit. Les résultats montrent une performance supérieure de tous ces modèles ML sur ces derniers, les RNN (dont LSTM) étant les meilleurs sur les plus grands bassins versants.

Trois autres études de thermie de rivière s'intéressant aux LSTM sont relevées :

- Huang *et al.* (2023) utilisent les LSTM afin de reconstruire les tendances à long terme de cours d'eau du bassin du lac Dongting (Chine), à partir d'observations de température de l'eau sur la période 2007-2020 et de données de température de l'air, d'évapotranspiration et de débit de 1960 à 2020. Le modèle est testé sur 2007-2010 avec une RMSE moyenne entre 1,21 °C et 2,35°C.

- Rahmani *et al.* (2021a) proposent sur des bassins versants bien instrumentés des États-Unis une RMSE médiane de 0,69 °C en utilisant des débits observés et des forçages météorologiques, 0,81 °C avec des débits simulés et des forçages météorologiques, et 0,86 °C avec ces derniers uniquement.
- Rahmani *et al.* (2021b) ont testé la même approche sur des cours d'eau aux observations lacunaires et ont montré une bonne capacité de ces réseaux à reconstituer la température, en obtenant une RMSE médiane de 1,13 °C.

Cependant, une critique répandue à l'encontre des approches par DL est qu'elles ne peuvent pas améliorer la compréhension d'un système, car il s'agit de boîtes noires (Shen, 2018), critique généralement associée aux modèles statistiques. En effet, ce type de modèle ne permet pas d'évaluer la contribution de tel ou tel phénomène dans le système étudié. Il peut néanmoins renseigner sur l'importance des variables utilisées, dont le choix et la qualité nécessitent une réflexion préalable afin d'obtenir des performances optimales. Selon Hashemi *et al.* (2022), il est de notre compétence non pas nécessairement de discuter de la formulation mathématique de ces modèles, mais de les utiliser de la meilleure des manières à la lumière de la connaissance des processus. De plus, ces modèles sont calés sur les observations, et on ne peut pas affirmer qu'ils seront capables de produire des résultats de qualité égale pour des conditions différentes de celles sur lesquelles ils ont été entraînés, particulièrement dans le cadre d'un changement global qui pourrait modifier la relation entre les variables prédictives et la variables cible. L'écueil de la nécessité de longues séries de données pour réaliser l'entraînement du modèle concerne particulièrement les réseaux de neurones, lesquels contiennent un nombre trop important de paramètres par rapport au reste des approches de régression statistique. Ils nécessitent donc l'instrumentation des sites sur le long terme, ce qui restreint les zones d'étude possibles afin de qualifier leurs performances.

1.3. Synthèse et positionnement du stage

Ainsi, la température est importante pour la diversité aquatique et impacte plusieurs activités humaines vitales, comme l'eau potable et la production d'énergie ; toutefois il y a un manque important de mesures pour son suivi, d'où l'intérêt de la modélisation. Il existe un certain nombre de projets en cours ou achevés qui l'étudient à l'échelle de plusieurs bassins versants métropolitains comme la Loire (Moatar et Gailhard, 2006; Bustillo *et al.*, 2014; Seyedhashemi *et al.*, 2021), la Seine (Ducharne, 2008; Rivière *et al.*, 2021), la Garonne (Larnier, 2010), et même à l'échelle nationale avec le projet TIGRE (Beaufort *et al.*, 2020; Moatar, 2021). Le Conseil Départemental de la Haute-Garonne (CD31) pilote le PTGA, dont l'une des actions vise à suivre et caractériser finement dans l'espace et projeter dans le temps la température des cours d'eau de Haute-Garonne. Pour ce faire, le manque de données nécessaires à la caractérisation fine des cours d'eau nous pousse à explorer la voie des modèles statistiques, en particulier les LSTM. Nous tenterons de répondre aux questions suivantes :

- Quel est l'apport des réseaux de neurones et en particulier les réseaux LSTM pour la prédiction de température de l'eau de rivière par rapport à un modèle plus simple (type régression multilinéaire) ? Quels paramètres permettent d'obtenir les meilleures performances, et en particulier, quelle durée d'observations ?
- Quelle est la valeur ajoutée du débit par rapport à la température de l'air seule dans ce type de modèle ?
- Au cas où l'on aurait pas accès aux mesures de débit, pourrait-on le remplacer par les forçages de précipitations et d'évapotranspiration à l'échelle du bassin versant en amont de la station ?

2. Cas d'étude et données

Ce chapitre présente le cadre dans laquelle s'inscrit l'étude ainsi que les données collectées, qui seront ensuite analysées.

2.1. Zone d'étude : la Garonne amont

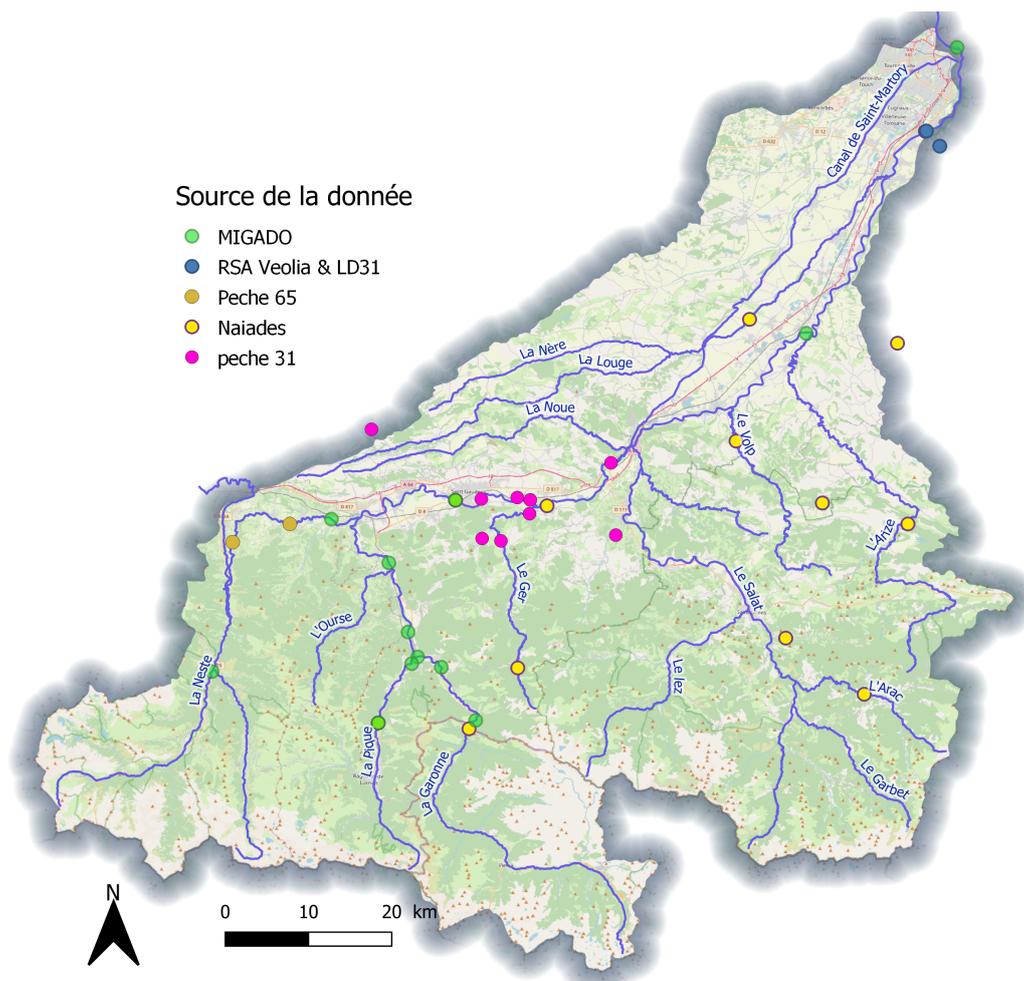


FIGURE 2.1.1 : Carte des 40 stations collectées pour l'étude.

Les données collectées pour le stage l'ont été dans le cadre du Projet de Territoire Garon'Amont (PTGA)⁴. Le PTGA est un projet de gestion de l'eau à l'échelle du territoire amont de la Garonne (Figure 2.1.1) piloté par le CD31 à la suite d'une concertation avec les habitants et les acteurs locaux entre 2019 et 2020. De cette concertation ont émergé quatre grands thèmes ("Économiser l'eau", "Stocker l'eau", "Pacte de gouvernance" et "Aménager le territoire"), découpés en 32 actions portant sur la période 2021-2025. Le stage s'inscrit dans le contenu de l'action D 1.2 sur la constitution et l'exploitation d'un observatoire thermique des cours d'eau et des nappes alluviales afin d'acquérir une meilleure connaissance collective de ce paramètre et d'en suivre les tendances.

⁴<https://garonne-amont.fr/>

A l'avenir, grâce à ce projet, 11 stations de mesure supplémentaires vont être ajoutées sur des cours d'eau du bassin versant de la Garonne amont jusqu'ici non instrumentés pour compléter le réseau en place (Rapport 2021-2022 du PTGA ; Labbouz (2022)). Ces stations permettront d'alimenter de futures études, et élargiront le périmètre de surveillance à des cours d'eau auparavant non instrumentés, le réseau actuel étant inégalement réparti sur le réseau hydrographique et couvrant majoritairement la Garonne.

Pour le territoire concerné par le PTGA (Figure 2.1.1), l'enjeu majeur est l'alimentation en eau potable, dont la température de prélèvement ne doit en théorie pas dépasser 25°C⁵. Ce critère est en partie technique, en raison d'une plus grande difficulté du traitement, car la température influe sur l'efficacité de la floculation et sur l'oxydation et la gestion des sous-produits du traitement de l'eau, mais est surtout sanitaire avec un risque de développement bactérien dans les réseaux de distribution et le relargage de chlorure de vinyle monomère dans les réseaux PVC antérieurs à 1980, d'après le rapport annuel 2021-2022 du PTGA.

Le territoire dans lequel se place l'étude couvre le bassin versant de la Garonne à l'amont de sa confluence avec l'Ariège. Son contexte géographique et hydro-climatique est contrasté, puisqu'il s'étend des Pyrénées au relief important (18 sommets ont une altitude supérieure à 3000 m en Haute-Garonne⁶) drainé par des cours d'eau forestiers et torrentiels à la basse plaine de la Garonne jusqu'à Portet-sur-Garonne à une altitude de 142 m⁷. La Garonne amont compte de nombreuses retenues destinées à l'hydroélectricité et au soutien d'étiage. Il est divisé en trois grandes parties : une zone montagneuse, au Sud ; une zone de piémont ; puis un parcours de plaine. Le climat est montagnard au sud, océanique avec une influence méditerranéenne au nord⁸. Les cours d'eau et les zones humides associées y abritent une importante biodiversité végétale et animale. Cette dernière est sensible à la température.

2.2. Données collectées

2.2.1. Observations de température de l'eau

Les données de température à pas de temps horaire de l'eau pour les stations entrant dans le périmètre du PTGA ont été compilées et prétraitées par le bureau d'étude Eaucéa (Labbouz, 2022) et comptent 40 stations réparties de manière hétérogène sur le territoire, la moitié se situant sur la Garonne même (20, voir Figure 2.1.1). Ces stations et les données associées sont gérées par divers organismes : le Réseau de Suivi et d'Alerte 31 (RSA31), l'association MIGADO, des fédérations de pêche départementales, ainsi que la base de données Naiades. Cette dernière est alimentée par l'Agence de l'Eau Adour-Garonne⁹. Pour celles gérées par MIGADO, certaines sont également disponibles au pas de temps journalier sur des périodes plus étendues. Les mesures sont réalisées avec des sondes de plusieurs modèles selon les stations et leur gestionnaire, y compris pour la même station lorsque des remplacements ont pu avoir lieu.

Les données seront ensuite utilisées en moyenne sur la journée. Les périodes de disponibilité (Tableau 2.1) s'étendent de 1991 à 2021, couvrant pour 13 stations moins de 5 ans (dont 7 couvrant moins d'un an), pour 12 entre 5 et 10 ans, pour 7 entre 10 et 15 ans et pour 7 entre 21 et 23 ans. Une seule station présente 31 ans de

⁵ Auparavant obligation soumise à dérogation pour la métropole (annexe 13-3 du décret n° 2003-462 du 21 mai 2003 relatif aux dispositions réglementaires des parties I, II et III du code de la santé publique), il s'agit aujourd'hui et depuis début 2023 d'une simple recommandation pour l'ensemble du territoire français.

⁶ <https://www.hautegaronnetourisme.com/des-emotions-pour-de-vrai/la-tres-haute-garonne/>

⁷ Le territoire inclut également Toulouse et la station du Bazacle.

⁸ https://fr.wikipedia.org/wiki/Climat_de_la_Haute-Garonne

⁹ La base de données Naiades est nationale et alimentée par les Agences de l'Eau, les Offices de l'Eau et la DREAL (pour les outre-mer), l'INRAE et l'OFB.

TABLE 2.1 : Répartition des stations par longueur de données de température de rivière disponibles. 7 stations comptent moins d'un an de mesures.

Durée (a)	Nombre de stations
0-5	13
5-10	12
10-31	15

mesures (le Bazacle à Toulouse).

Ces données ont fait l'objet de pré-traitements par le bureau d'études qui les a rassemblées. Ces traitements comportent :

- La conversion manuelle au format tableur.
- La suppression des doublons d'enregistrement.
- La suppression de valeurs aberrantes : valeurs hors de l'intervalle 0-35 °C, variations supérieures à 2 °C en une heure et à 5 °C autour de la moyenne en 12 h.
- La mise des données au format horaire et journalier.

Il est probable que des erreurs de mesure persistent dans les chroniques, dues à une dérive ou à une exondation de la sonde qui n'auraient pas été détectées, toutefois ces erreurs sont supposées minimales. La méthode de mesure et l'hétérogénéité locale du champ de température peuvent également engendrer des différences sur un même site, comme entre les deux stations de MIGADO et du RSA31 au Bazacle. D'autre part, il faut garder à l'esprit la faillibilité des mesures, qu'elle provienne de l'incertitude liée au modèle de sonde dont on n'a pas toujours connaissance, d'une erreur de mesure, ou d'une erreur dans la chaîne d'obtention et de traitement. Les chroniques ne sont presque jamais complètes et pourraient ainsi difficilement raisonnablement l'être, des mesures manquantes ou erronées portant le pourcentage de disponibilité moyen à 79,3 % dans les chroniques journalières, des saisons parfois entières manquant à l'appel.

2.2.2. Entrées pour les modèles

En nous appuyant sur la bibliographie et les résultats produits par le bureau d'études Eaucéa, nous avons choisi quatre variables explicatives principales pour notre étude : la **température de l'air**, dont l'importance serait prépondérante, et le **débit**. Alternativement, ce débit pourrait être reconstitué à partir des signaux de **précipitations** et d'**évapotranspiration potentielle** à l'échelle du bassin versant, et ce en négligeant les processus d'infiltration et les processus hydrologiques d'origine purement anthropique, comme les prélèvements, les rejets et les opérations de soutien d'étiage. Dans une perspective opérationnelle, la température de l'air moyenne pourra également être remplacée par les valeurs maximales et minimales, puisque ce sont les données mises à disposition par Météo-France pour la prévision au CD31.

Les données de température de l'air sont issues des réanalyses du Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige (SAFRAN) (Météo-France; Vidal *et al.* (2010)), sur des mailles carrées de 8 km de côté, du 1^{er} août 1989 au 1^{er} août 2020 dont les valeurs sont prises aux centres des mailles les plus proches des localisations des stations. Il s'agit donc de températures locales. Les stations de débit de l'HydroPortail (anciennement Banque Hydro; Leleu *et al.* (2014); Dufeu *et al.* (2022))¹⁰ ne coïncidant pas

¹⁰<https://hydro.eaufrance.fr/>

avec celles de température, pour chaque station, la station la plus proche géographiquement est sélectionnée, si possible sans confluence importante entre les deux. Ce choix est suffisant pour notre cas car les modèles LSTM que nous avons testés exploitent plutôt la dynamique du signal, et non pas ses valeurs absolues. Les données de ces stations sont disponibles pour les identifiants hydrométriques au sens du référentiel Sandre¹¹ commençant par O0 et O1 (correspondant à la Garonne en amont de sa confluence avec l'Ariège, et à l'Ariège) du 1^{er} janvier 1990 au 1^{er} janvier 2021.

Des températures moyennes à l'échelle du Bassin Versant (BV) sont calculées à partir des données SAFRAN et à l'aide de polygones de Thiessen¹² puis utilisées pour estimer des lames journalières moyennes d'Évapotranspiration Potentielle (ETP) selon la formule de Oudin *et al.* (2005). Les données d'ETP sont ainsi disponibles du 2 août 1989 au 31 juillet 2020. Les données de précipitation sont obtenues de manière similaire en délimitant les bassins versants topographiques en amont de chacune des stations et en intégrant les données de la lame d'eau COMbinaison en vue de la Meilleure Estimation de la Precipitation HOraiRE (COMEPHORE) (Météo-France; Tabary *et al.* (2012)). Elles sont disponibles du 1^{er} janvier 1997 au 31 décembre 2018. Ces deux dernières données sont estimées à l'échelle du bassin versant et non localement.

Le Tableau 2.2 récapitule les données collectées.

TABLE 2.2 : Récapitulatif des données.

Variable	Notation	Source	Résolution	Disponibilité	Type de donnée
Température de l'eau	T_{eau}	MIGADO, Naïades, RSA31, Peche31, Peche65	Horaire & journalière	Selon station (voir Annexe ??)	Ponctuelle
Température de l'air	T_{air}	SAFRAN (Météo-France); Vidal <i>et al.</i> (2010)	Horaire, grille 8 km de côté	01/08/1989-01/08/2020	Ponctuelle
Débit	Q	Hydroportail (Auparavant Banque Hydro; Leleu <i>et al.</i> (2014))	Horaire	01/01/1990-01/01/2021	Ponctuelle
Précipitations	P	COMEPHORE (Météo-France); Tabary <i>et al.</i> (2012)	Horaire	01/01/1997-31/12/2018	Moyenne sur BV
Évapotranspiration potentielle	ETP	Formule de Oudin <i>et al.</i> (2005) avec SAFRAN (Météo-France)	Journalier & horaire (redistribution de l'estimation journalière)	02/08/1989-31/07/2020	Moyenne sur BV

2.3. Analyse et critique de données

Après réception des données, quelques statistiques et indicateurs permettant de les caractériser et de définir l'état du territoire étudié en terme de thermie des cours d'eau ont été calculés :

¹¹Service d'administration nationale des données et référentiels sur l'eau, <https://www.sandre.eaufrance.fr/missions-et-organisation-du-sandre>

¹²<http://geoconfluences.ens-lyon.fr/glossaire/polygones-de-thiessen-ou-de-voronoi>

- Les moyennes interannuelles des 30 jours et 7 jours les plus chauds de chaque année (représentatives des extrêmes chauds, en se référant au projet TIGRE (Beaufort *et al.*, 2020). Ces indicateurs seront nommés MTw30J et MTw7J par la suite.
- Le nombre moyen, par année, de dépassement des seuils 19 °C et 25 °C. Pour cet indicateur, nous avons exclu les années pour lesquelles plus de 10 jours manquaient entre juin et septembre inclus.
- L'existence d'une tendance significative sur le long terme ¹³ à l'aide d'un test de Mann-Kendall adapté aux séries avec une variabilité saisonnière (bibliothèque Python *pymannkendall*; Hussain et Mahmud (2019)). Seules les 18 stations disposant de plus de 10 ans de mesures sont considérées. Cette tendance est également calculée sur la température de l'air, dont l'évolution devrait se refléter sur celle de l'eau, et ce sur la même période que celle de l'eau mais aussi sur la période 1989-2020.

2.3.1. Variabilité spatiale des indicateurs sur la Garonne amont

Les cartes obtenues pour les deux premières statistiques sont visibles sur la Figure 2.3.1. Ces indicateurs permettent de retrouver la répartition des stations selon les trois zones décrites par le relief du territoire : une zone fraîche, dans les Pyrénées, où la température ne dépasse en moyenne jamais 19 °C et où la température moyenne maximale sur 30 jours ne dépasse pas 17 °C; une zone de transition, où ce seuil de 19 °C est dépassé quelques fois par an, et une zone plus chaude, vers la plaine, où ce seuil est dépassé plus de 60 jours par an et où le seuil de 25 °C est dépassé en moyenne tous les ans. En raison des hypothèses préalables au calcul, les températures extrêmes chaudes sont plus basses que celles obtenues dans le rapport annuel 2021-2022 du PTGA (Labbouz, 2022) : le bureau d'études Eaucéa écarte les années pour lesquelles moins de 90 % des données sont disponible, alors qu'ici seules les années dont plus de 30 jours manquent entre juillet et septembre sont écartées.

¹³Ce «long terme» est tout relatif, une période de 30 ans minimum étant généralement considérée pour établir ce type de statistique.

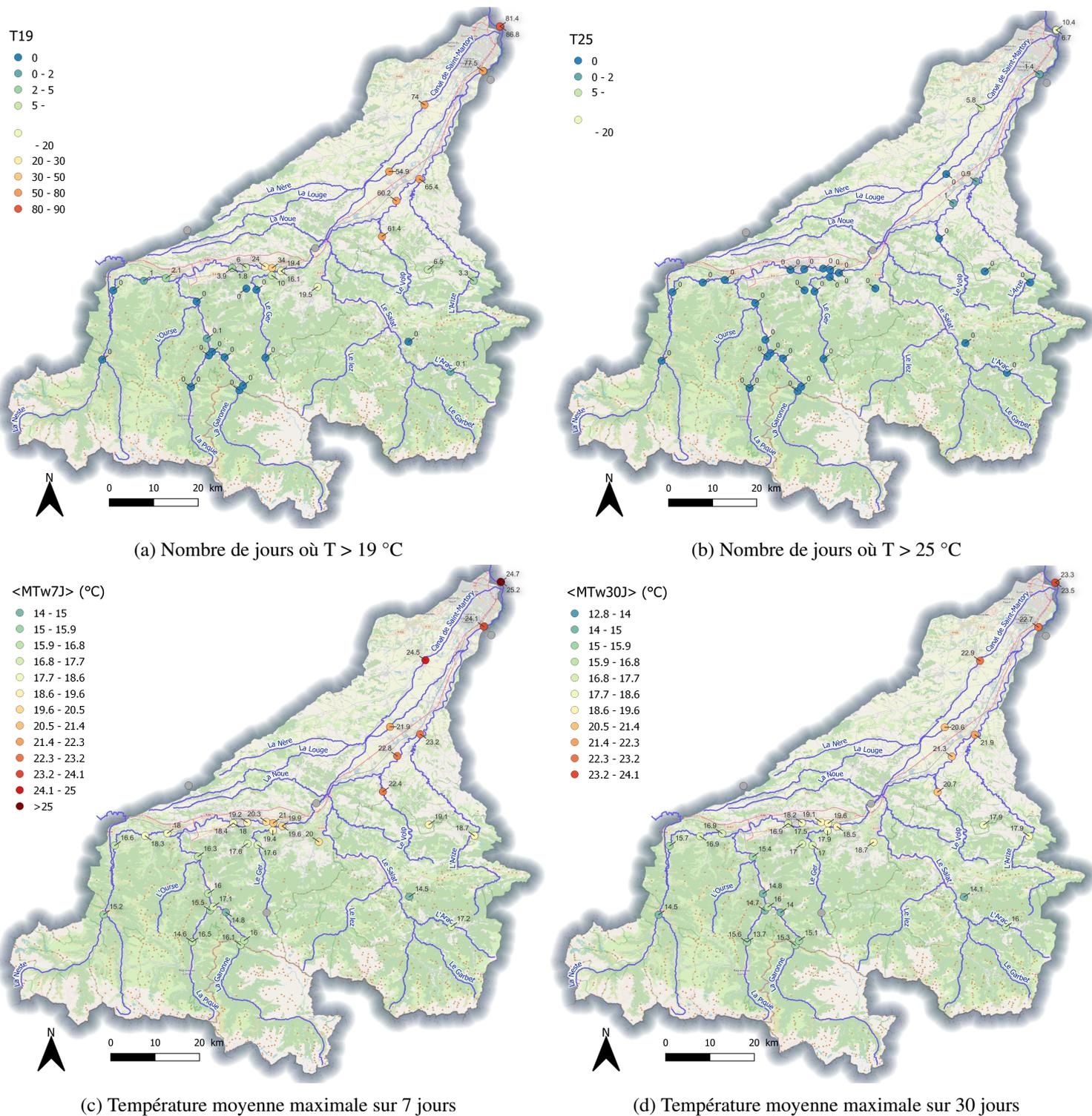


FIGURE 2.3.1 : Moyennes interannuelles (les années et le nombre d'années varient d'une station à l'autre).

2.3.2. Tendances observées dans l'évolution temporelle

Les tendances révélées par le test de Mann-Kendall sont calculées pour 18 stations sur 40 et sont significatives pour 16 d'entre elles, où elles sont à l'augmentation (Table 2.3 et Figure 2.3.2). Celles-ci sont d'autant plus élevées vers l'amont. À l'aval, les tendances sont toujours présentes mais plus faibles (+ 0,2-0,3 °C/10a) à non significatives (pour le Volp à Plan et la Louge à Fousseret).

Elles sont conformes à ce qui a pu être établi pour la France métropolitaine :

- Pour la Garonne entre Toulouse et Malause (tronçon aval), une augmentation estivale de 2,8 °C et d'une augmentation de la moyenne annuelle de 0,8 °C entre 1978 et 2005 (Larnier, 2010).
- Pour la Loire, entre 1977 et 2019, une augmentation à hauteur de 0,1 à 0,6 °C par décennie (Seyedhashemi *et al.*, 2022).
- À l'échelle nationale selon un rapport d'Électricité De France (EDF) (Maire, 2021), ces tendances sont perceptibles sur les mesures au voisinage de 11 CNPE sur tous les grands bassins versants français sauf celui de la Garonne, à hauteur de 0,8 °C/décennie pour la période avril-juin depuis 1980.

Cependant le rapport annuel 2021-2022 du PTGA (Labbouz, 2022) ne parvient pas à extraire de tendance linéaire statistiquement significative pour la seule période 1994-2020.

TABLE 2.3 : Tendances linéaires pour les stations présentant plus de 10 années d'observations.

Station	Nombre d'années	Tendance
La Garonne au Bazacle (MIGADO)	31	+ 0,2 °C/10a
La Garonne au Bazacle (RSA-CD31)	22	+ 0,75 °C/10a
La Garonne à Carbonne	23	+ 0,3 °C/10a
La Garonne à Valentine (MIGADO)	13	+ 1,3 °C/10a
La Garonne à Loures-Barousse	22	+ 0,4 °C/10a
La Neste à Aventignan	21	+ 0,7 °C/10a
La Pique à Cier-de-Luchon	14	+ 1,1 °C/10a
La Pique à Marignac	10	+ 1,0 °C/10a
L'Arac à Biert	12	+ 0,5 °C/10a
L'Arize au Mas-d'Azil	12	+ 0,6 °C/10a
La Garonne à Labarthe-Inard	12	+ 0,8 °C/10a
La Garonne à Valentine (Naiades)	12	+ 1,2 °C/10a
La Louge en aval de Fousseret	10	non
Le Montbrun en amont de Montbrun-Bocage	10	+ 0,2 °C/10a
Le Nert à Rivèrenert	12	+ 0,6 °C/10a
Le Volp au Plan	10	non
La Garonne à Montespan	22	+ 0,17 °C/10a
La Garonne à Saint-Julien	22	+ 0,46 °C/10a

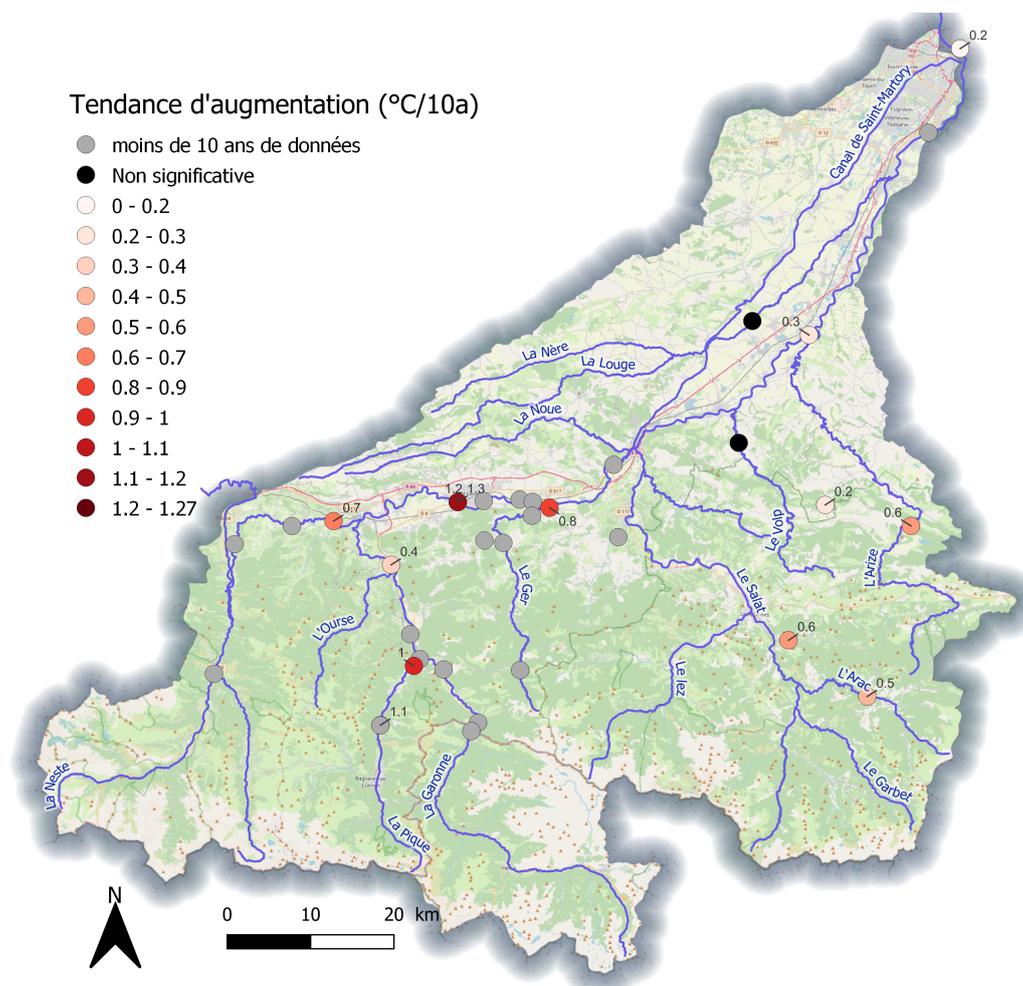


FIGURE 2.3.2 : Carte des tendances calculées par le test de Mann-Kendall désaisonnalisé sur la température de l'eau des 18 stations.

Toutefois, si on observe bien des tendances à l'augmentation allant de 0,18 ou 0,75 °C/10a à Toulouse jusqu'à 1,3 °C/10a à Valentine ¹⁴, sur des séries longues de 10 à 30 ans et différentes pour chaque station, ces observations ne sont pas suffisantes afin d'attribuer fiablement ces tendances au changement climatique. Ces chroniques ne sont pas d'une longueur suffisante et ne permettent pas de quantifier une tendance à long terme et il se pourrait que les tendances soient liées à la coupe ou la pousse de la ripisylve locale. De plus, ces augmentations sont à relativiser notamment par rapport à une éventuelle défaillance non détectée des sondes, un effet local lié au courant, ou une stratification thermique, comme on peut l'observer pour les deux stations au Bazacle dont le comportement est inexplicablement différent et appelle à la prudence quant au protocole de mesure et à la qualité de ces dernières servant de point de départ à l'analyse.

Pour mettre en perspective ces tendances obtenues pour la température de l'eau, nous avons également calculées des tendances sur la température de l'air extraites à partir des données SAFRAN aux localisations des stations. Dans un premier temps, nous avons restreint le calcul aux mêmes périodes de disponibilité de la température de l'eau (court terme), ensuite nous avons étendu le calcul à toute la chronique de la température de l'air (long terme). Sur la Figure 2.3.3, les tendances à court terme de la température de l'eau et de l'air sont plutôt similaires, la température de l'eau tendant à se réchauffer plus vite que celle de l'air. En revanche, les tendances long-terme sur la température de l'air sont en augmentation significative pour seulement une poignée

¹⁴Les tendances pour les stations gérées par MIGADO et Naïades à Valentine, géographiquement distinctes, concordent.

de stations. Ces résultats suggèrent que l'on n'aurait pas de tendance significative sur la température de l'eau pour ces stations même si l'on avait des chroniques aussi longues que celles de la température de l'air.

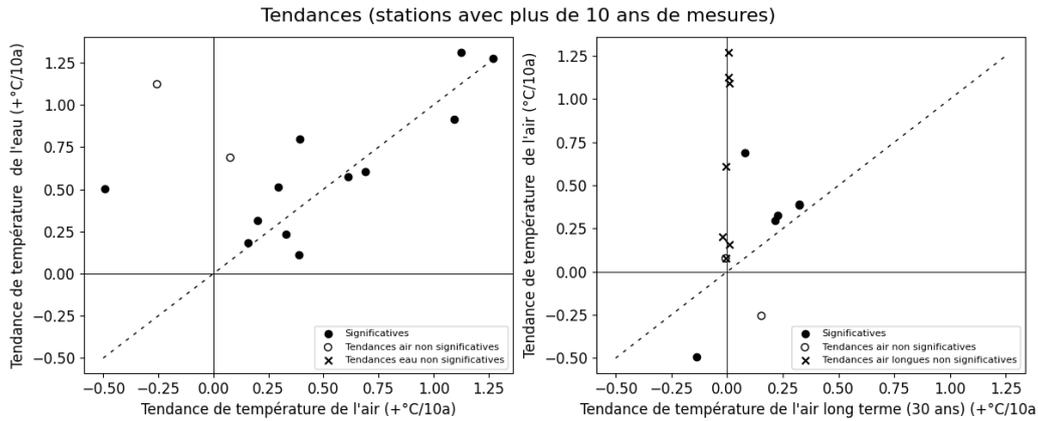


FIGURE 2.3.3 : A gauche, les tendances de température de l'eau en fonction de celles pour l'air sur la même période "courte". A droite, les tendances de température de l'air sur le temps "court" en fonction de celles sur le temps "long" (toute la longueur des données SAFRAN disponibles).

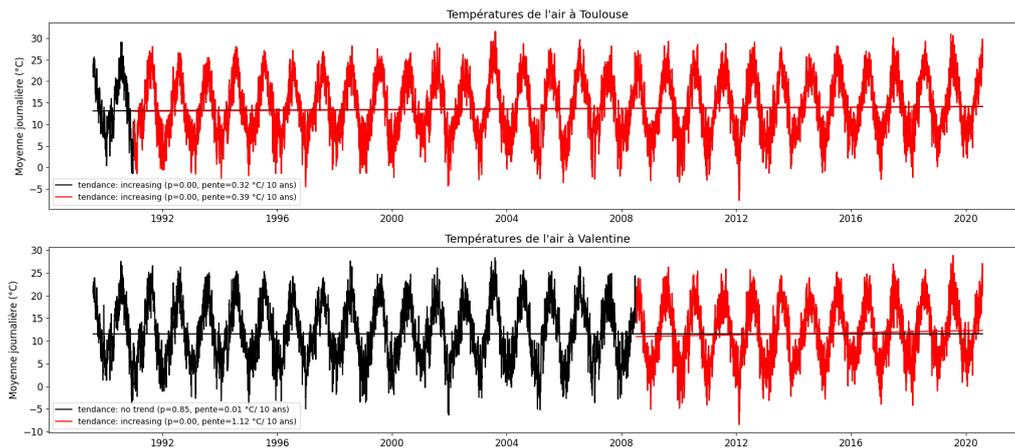


FIGURE 2.3.4 : En noir, les données de température de l'air SAFRAN 1989-2020. En rouge, les données de température de l'air SAFRAN «courtes» correspondant à la période à laquelle les observations de température de l'eau sont disponibles. En haut : Toulouse, en bas : Valentine.

Du côté des tendances de température de l'air, il est à noter que même si l'on parvient à extraire des tendances d'augmentation significatives à court terme, sur la période 1989-2020 une bonne moitié des tendances deviennent non significatives et plus proches de zéro (Figure 2.3.3, à droite). Cette absence peut être due à la méthode ou la durée largement insuffisante utilisée pour estimer les tendances (10 à 30 ans sont trop peu), ou à une oscillation du climat sur ces quelques années, mais ne peut en aucun cas remettre en question l'existence d'un changement de celui-ci. La figure 2.3.4 montre les températures de l'air à Toulouse et à Valentine. A Toulouse, leur tendance est significative que l'on prenne les 30 ans de données SAFRAN ou celles sur lesquelles les températures de l'eau sont disponibles, et la pente est presque égale (la durée est presque identique également). A Valentine, la période de disponibilité des températures de l'eau est plus faible et, malgré l'absence de tendance sur 30 ans, met en évidence une forte augmentation de la température de l'air (+1,12 °C/décade, similaire à l'augmentation de celle de l'eau) sur les 13 dernières années.

Ces tendances sont toujours à analyser avec précaution, comme recommandé par le NCARS (2014) : il faut

garder à l'esprit l'incertitude associée au réseau de mesure, les erreurs associées aux stations, et l'introduction possible de discontinuités, qui sont ici toutes les problématiques liées aux stations exposées dans le chapitre précédent. La non-stationnarité du système climatique et la possibilité d'extraire des tendances significatives différentes (voire négatives) pour une sous-période des observations plutôt qu'une autre, et ainsi l'influence de la période de calcul, peuvent être responsables de la disparition ou de l'apparition des tendances.

3. Méthodes

3.1. Les modèles LSTM

L'intérêt dans ce stage a été porté sur une méthode de ML utilisant des RNN LSTM. Comme établi dans la synthèse bibliographique, ce type de modèle jouit d'une popularité récente appuyée par des performances encourageantes dans le domaine des sciences de l'eau particulièrement pour la modélisation hydrologique et thermique de bassins versants en l'absence d'observations (Shen, 2018; Rahmani *et al.*, 2021b).

3.1.1. Des neurones aux cellules

Tout réseau de neurones est composé d'un certain nombre d'unités simples (Figure 3.1.1) organisés selon une certaine structure.

Ces neurones représentent chacun une série d'opérations mathématiques : la multiplication du signal X par une matrice de **poids** W à laquelle un **biais** b est ajouté, puis une **fonction d'activation** σ est appliquée ou non au résultat. Dans le cas d'un RNN, une multiplication supplémentaire est ajoutée, celle de la sortie du neurone Y au pas de temps précédent par de nouveaux poids, ce qui représente une rétention d'information à court terme sous forme d'un état interne souvent noté h .

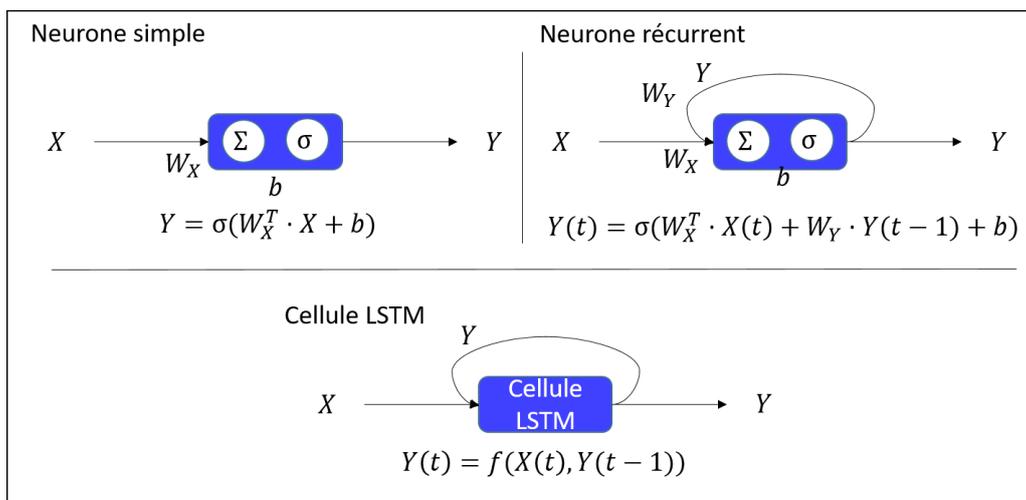


FIGURE 3.1.1 : Schéma des structures unitaires d'un réseau de neurones.

Une cellule LSTM (Hochreiter et Schmidhuber, 1997) est un cas particulier de RNN permettant de s'affranchir d'un problème rencontré ceux-ci, la disparition (ou l'explosion) du gradient (voir Bengio *et al.* (1994)) ne permettant pas de retenir l'information de plus de 10 pas de temps précédents.

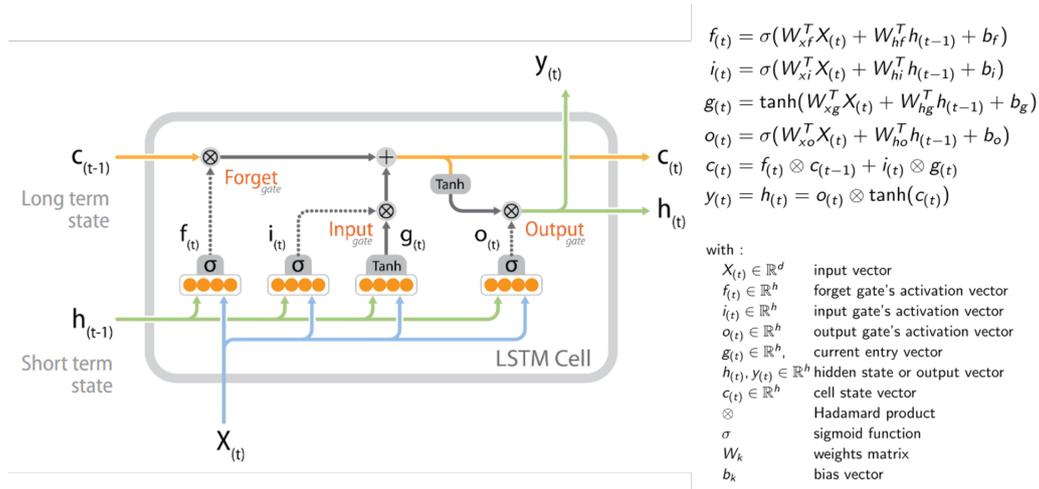


FIGURE 3.1.2 : Schéma d'une cellule LSTM. Repris des supports de la formation Fidle-CNRS (<https://gricad-gitlab.univ-grenoble-alpes.fr/talks/fidle>).

Ces cellules sont composées de neurones organisés dont le fonctionnement est notamment détaillé par Kratzert *et al.* (2018) et Feng *et al.* (2020), intégrant une rétention d'information à long terme supplémentaire (Figure 3.1.2). Dans ces réseaux, un deuxième état (noté c sur le schéma) et deux portes permettant pour la porte d'oubli d'en oublier une partie et pour la porte d'entrée d'utiliser une partie de cet état interne dans la prédiction actuelle.

3.1.2. Réseaux de neurones et apprentissage

Les réseaux de neurones fonctionnent tous sur le même principe : une information d'entrée (*input*) traverse un **réseau**, structure complexe composée d'un certain nombre de neurones (ou ici, cellules) connectés entre eux, sur une certaine épaisseur. En sortie de ce réseau, l'information (*output*) peut être comparée à celle attendue grâce à une métrique estimant l'erreur.

Avant de procéder à l'entraînement d'un modèle, les données sont divisées en trois lots : des données servant à cet entraînement, des données servant à la validation conjointe, et un lot gardé de côté pour le test afin qu'il n'aie jamais été vu par le modèle lors de celui-ci.

Le lot d'apprentissage est divisé en paquets appelés *batches*. Lors de l'apprentissage, chacun de ces paquets est passé à travers le réseau, ce qui permet de calculer une erreur cumulée. Ensuite, cette erreur est rétropropagée à travers le réseau afin d'ajuster les poids et les biais (les **paramètres** du modèle) d'un certain **taux d'apprentissage** (*learning rate*) par un algorithme de rétropropagation dont on définit l'**optimiseur**. Séparer les données en *batches* permet à l'optimisation d'explorer un maximum de directions et de trouver de meilleurs paramètres que si toutes les données lui avaient été fournies en une fois.

Un cycle propagation-rétropropagation est aussi appelé une **époque**, répétée autant de fois que nécessaire afin d'obtenir des paramètres pour lesquels la performance est la plus satisfaisante possible. A la fin de chaque époque, l'erreur est calculée sur le lot réservé pour la validation. Si elle ne s'améliore plus ou régresse (choix d'un point d'arrêt), ou si le nombre d'époques choisi par défaut est atteint, le réseau est prêt à être évalué sur des données qu'il n'a jamais rencontrées. En effet, plus les époques passent, plus le modèle est adapté au jeu de données d'entraînement. Cependant, le modèle atteint une erreur minimale, puis réaugmente. Ce phénomène s'appelle l'*overfitting* ou surapprentissage et est un défi majeur de l'apprentissage automatique (Bejani et Ghaate, 2021).

Les Figures 3.1.3 et 3.1.4 montrent des schémas simples de la structure d'un réseau de neurones comportant une couche de cellules LSTM et de l'entraînement d'un modèle avec en rouge, des paramètres appelés hyperparamètres qui peuvent être variés par la suite du travail :

- **Variables d'entrée** : les différentes variables explicatives à tester.
- **Nombre de cellules** : nombre de cellules dans la couche LSTM.
- **Taux d'oubli** (ou *dropout rate*) : Probabilité de désactiver un neurone pendant un passage d'entraînement, évitant ainsi un surapprentissage (Srivastava *et al.*, 2014).
- **Batch size** : nombre d'éléments dans les *batches*.
- **Choix de la fonction coût** : choix de la métrique d'erreur à optimiser (par exemple, MSE).
- **Optimiseur** : algorithme utilisé pour la rétropropagation.
- **Taux d'apprentissage** (ou *learning rate* : taux d'actualisation des paramètres du modèle à chaque époque.

Dans le cas d'un réseau LSTM, un paramètre supplémentaire nous intéresse plus particulièrement : la **longueur de séquence prédictive** (ou *lookback*). Ce paramètre fixe la durée des observations précédant un point à simuler : par exemple, on peut utiliser l'évolution des paramètres d'entrée sur une semaine pour prédire le jour suivant. Le *lookback* pourrait porter un sens physique, puisqu'il porte l'information d'inertie de la masse d'eau et du régime thermique du cours d'eau.

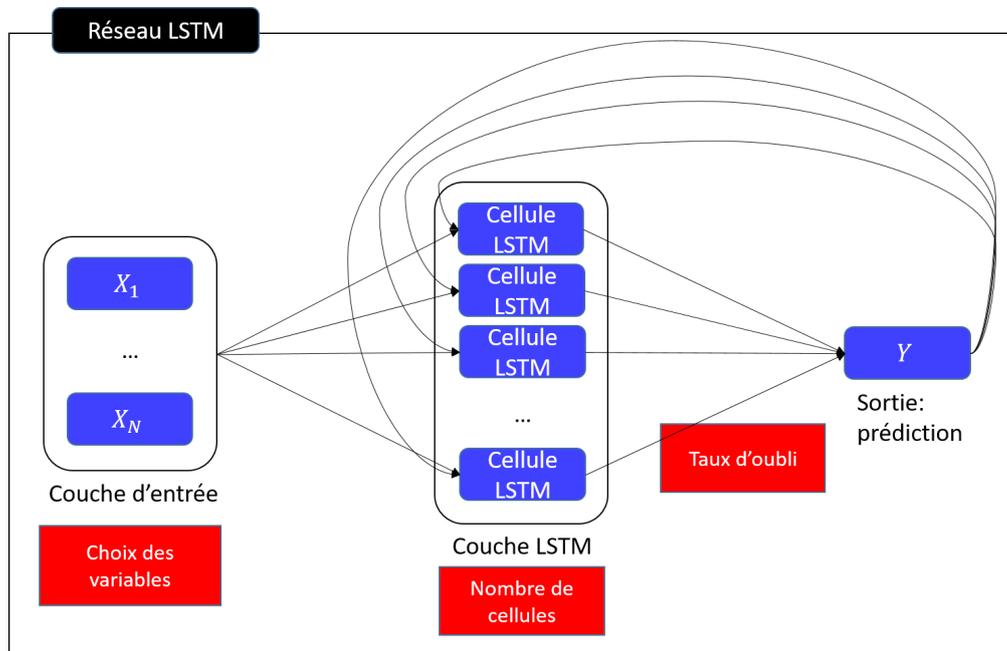


FIGURE 3.1.3 : Schéma d'un réseau de neurones LSTM.

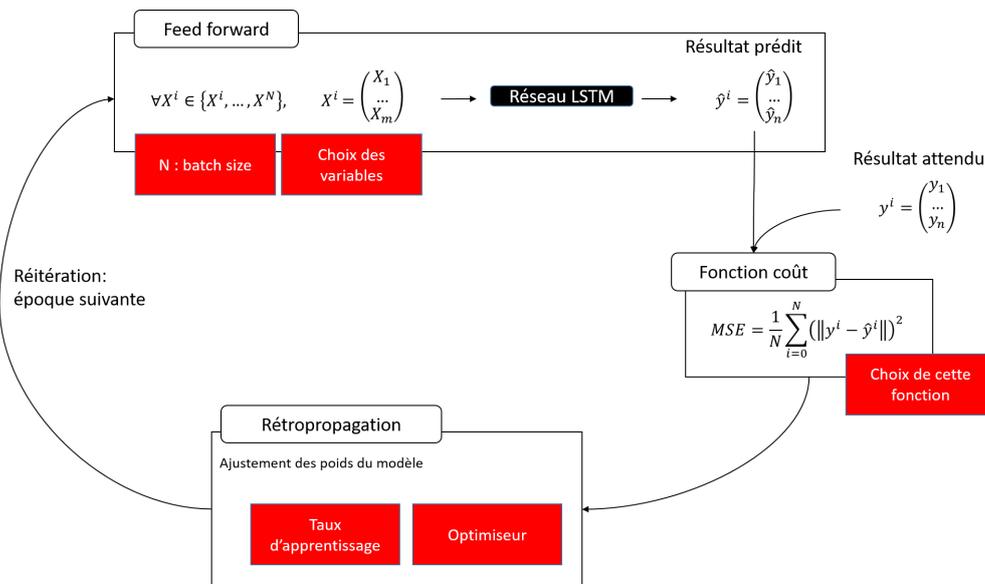


FIGURE 3.1.4 : Fonctionnement de l'apprentissage d'un réseau de neurones.

3.2. Implémentation et métriques de performance

Le travail a été réalisé sous Python à l'exception de quelques traitements avec R et QGIS, et les modèles ont été implémentés grâce à la librairie open source *PyTorch* (Meta, Paszke *et al.* (2019), version 2.0.0), en s'inspirant du code de Kratzert *et al.* (2018) ¹⁵. Cette librairie permet la mise en oeuvre clé en main relativement aisée de la création de modèles complexes et leur apprentissage. Divers traitements utilisent les librairies NumPy (Van Der Walt *et al.*, 2011), Pandas (McKinney, 2010) et Scikit-learn (Pedregosa *et al.*, 2012). Les graphiques sont réalisés à l'aide de la librairie Matplotlib (Hunter, 2007).

Les séries temporelles pour chaque station sont normalisées préalablement au calcul, en utilisant la moyenne et l'écart-type de la période d'entraînement. Avant d'évaluer le résultat, les valeurs sont dénormalisées de la même manière.

La métrique utilisée pour l'apprentissage est la MSE en °C² :

$$MSE = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} (T_{sim} - T_{obs})^2 \in [0, \infty]$$

Avec n_{obs} le nombre d'observations reproduites ¹⁶, T_{obs} la température observée et T_{sim} la température simulée. Cette MSE doit être le plus proche de 0 possible.

La même méthodologie pourrait être appliquée avec une métrique pénalisant plus fortement les erreurs sur les températures élevées pour représenter au mieux ces valeurs.

Des choix ont dû être réalisés afin de pouvoir comparer les performances des modèles. En effet, les périodes de disponibilité des données sont très hétérogènes. Pour pouvoir comparer les performances dans les meilleures conditions, dans l'idéal une période de disponibilité commune à toutes les stations devrait être utilisée, afin de pouvoir analyser d'une station à l'autre les performances pour une même saison. Cela impliquant une perte de

¹⁵Les premiers essais utilisaient des outils similaires, *Keras* (Google, Chollet *et al.* (2015)) et *TensorFlow* (Abadi *et al.*, 2016)

¹⁶La simulation de valeurs manquantes dans les observations n'influe donc pas sur le résultat

données trop importante, il a été choisi de sélectionner, par station, la plus longue période possible contenant les observations pour toutes les combinaisons de variables possibles.

De plus, les données manquantes sont un problème pour l'apprentissage. Elles sont alors remplacées par une interpolation linéaire ou, quand la période manquante est plus longue, par la climatologie. Cela fait que les modèles peuvent apprendre sur des valeurs hautement incertaines.

Afin d'estimer la capacité des modèles à reconstruire les chroniques de température, nous calculons la *Mean Absolute Error* ou erreur moyenne absolue (MAE) en °C :

$$\text{MAE} = \frac{1}{n_{obs}} \sum_{i=1}^{n_{obs}} |T_{sim} - T_{obs}| \in [0, \infty]$$

Cette MAE doit être le plus proche de 0 possible.

L'un des objectifs du réseau de thermie départemental est de produire des prévisions de températures estivales. Il est alors calculé la matrice de confusion relativement à l'observation du dépassement d'un seuil de température donné (19°C, correspondant aux seuil de stress physiologique pour la truite fario) et la précision associée. Cette précision est calculée en estimant les ratios suivants :

- VP (Vrai Positif) : Nombre de jours où $T_{obs} \geq 19 \text{ °C}$ & $T_{sim} \geq 19 \text{ °C}$.
- FP (Faux Positif) : Nombre de jours où $T_{obs} < 19 \text{ °C}$ & $T_{sim} \geq 19 \text{ °C}$.
- VN (Vrai Négatif) : Nombre de jours où $T_{obs} < 19 \text{ °C}$ & $T_{sim} < 19 \text{ °C}$.
- FN (Faux Négatif) : Nombre de jours où $T_{obs} \geq 19 \text{ °C}$ & $T_{sim} < 19 \text{ °C}$.
- A partir de ces quatre grandeurs, on calcule la précision = $\frac{VP + VN}{VP + FP + VN + FN} \in [0, 1]$ qui doit être la plus proche de 1 possible.

Ce seuil étant dans les faits rarement dépassé sur la majeure partie du territoire, nous nous intéressons également à la MAE calculée sur les valeurs de température supérieures à 75% des températures d'un jeu de données.

3.3. Modèles construits

Trois types de modèles ont été implémentés, et récapitulés dans la Table 3.1. La première vise à élaborer des modèles LSTM station par station en optimisant la structure du réseau de neurones et les paramètres d'entraînement. La deuxième vise à tester l'applicabilité d'un modèle LSTM régional nourri par toutes les observations du territoire d'étude et donc adapté à leur comportement moyen. La dernière vise à comparer les performances des deux autres vis-à-vis d'un modèle issu d'une régression linéaire déjà utilisé sur une station particulière.

3.3.1. Modèles LSTM

Afin d'extraire la meilleure paramétrisation de modèle LSTM pour la simulation de température de l'eau, un certain nombre de tests sont lancés, à raison d'un modèle par station et par combinaison d'hyperparamètres définissant l'architecture du modèle. Inutilisables pour les modèles locaux, les cas pour lesquels moins d'un an de données est disponible sont écartés, ce qui porte le nombre de stations à 27 sur les 40 disponibles. 30 sont utilisables pour les modèles régionaux.

TABLE 3.1 : Récapitulatif des tests effectués.

Nom	Description	Espace de paramètres testés : nombre de couches ¹⁷ × look-back × batch size × nombre de cellules × entrées.	Nombre de stations
Local	Un modèle LSTM pour chaque station pour chaque jeu d'hyperparamètres	{1} × {14, 30, 60} × {64, 256} × {64, 128, 256} × { T_{air} , [$T_{air,min}$, $T_{air,max}$], [T_{air} , Q], [T_{air} , P , ETP]}	27
Régional	Un modèle LSTM pour toutes les stations pour chaque jeu d'hyperparamètres	{1} × {7, 14, 30, 60} × {64, 256} × {64, 128, 256} × { T_{air} , [$T_{air,min}$, $T_{air,max}$], [T_{air} , Q], [T_{air} , P , ETP]}	30
Linéaire Bazacle	Modèle de prévision appliqué au Bazacle à Toulouse	Variables : [$T_{air,min}$, $T_{air,max}$, Q , T_{eau}]	1

Ainsi, 72*27 modèles locaux ¹⁸ et 72 modèles régionaux sont contruits, entraînés et évalués. Certains hyperparamètres sont maintenus constants (voir Table 3.2).

TABLE 3.2 : Hyperparamètres constants.

Nom	Valeur
<i>Dropout rate</i>	20%
<i>Learning rate</i>	0,0001
Proportions entraînement-validation-test	70%-15%-15%
Nombre d'époques	1000
Critère d'arrêt	20 époques sans amélioration de MSE sur la validation

3.3.2. Modèle linéaire pour une station

Afin de comparer les résultats des modèles et d'évaluer les performances des modèles LSTM pour la prévision de la température de l'eau, nous avons testé sur la station du Bazacle le modèle linéaire opérationnel calibré et utilisé par Eaucéa pour des prévisions jusqu'à 3 jours dans le rapport sur la modélisation thermique de la Garonne de Labbouz (2023) :

$$T_{eau}(J) = 0.84T_{eau}(J - 1) + 1.152TN(J) + 0.105TX(J) - 0.0035Q(J)$$

Avec J le jour considéré, TN et TX ¹⁹ les températures minimale et maximale de l'air (Eaucéa utilise la station de Blagnac, ici remplacée par les données SAFRAN) et Q le débit moyen, considéré seulement si le débit moyen sur 8 jours est inférieur à $300 \text{ m}^3/\text{s}$. Le modèle n'est valable que de juin à septembre. Les échéances suivantes ($J+1$ à $J+3$) sont prédites en fonction des valeurs des échéances précédentes pour la température de l'eau, et en fonction des prévisions pour les autres variables. Ce modèle nécessite d'avoir accès à la température de l'eau

¹⁸En réalité, un nombre moindre de modèles, le débit n'étant pas disponible en raison de l'absence de station hydrométrique identifiée à proximité et/ou absente dans les données collectées pour certaines stations

¹⁹Ces températures sont définies par convention comme «température la plus basse qui ait été relevée sur le site entre le jour J-1 à 18 h UTC et le jour J à 18 h UTC» et «température la plus élevée qui ait été relevée sur le site entre le jour J à 6 h UTC et le jour J+1 à 6 h UTC». En France elles correspondent sauf cas particuliers à la définition intuitive «température minimale/maximale de 00h00 à 23h59», qui peut donc être utilisée.

en un point donné afin d'être appliqué. La température de l'eau du jour précédent n'a volontairement pas été utilisée dans le développement de nos modèles, afin de conserver la possibilité d'une applicabilité pour des sites non instrumentés (problématique de régionalisation et prévision en sites non jaugés). Les résultats d'un modèle LSTM l'utilisant sont disponibles en annexe D.

4. Résultats et discussions

Ce chapitre présente et discute les principaux résultats de modélisation obtenus, de performances générales sur toutes les stations puis sur une station en particulier (la Garonne au Bazacle à Toulouse), choisie en raison de son intérêt de prévision pour le CD31. Les résultats pour une seconde station (La Garonne à Valentine) sont présentés en annexe B.

4.1. Hyperparamètres et variables prédictives optimaux

4.1.1. Tests des hyperparamètres

Modèles locaux

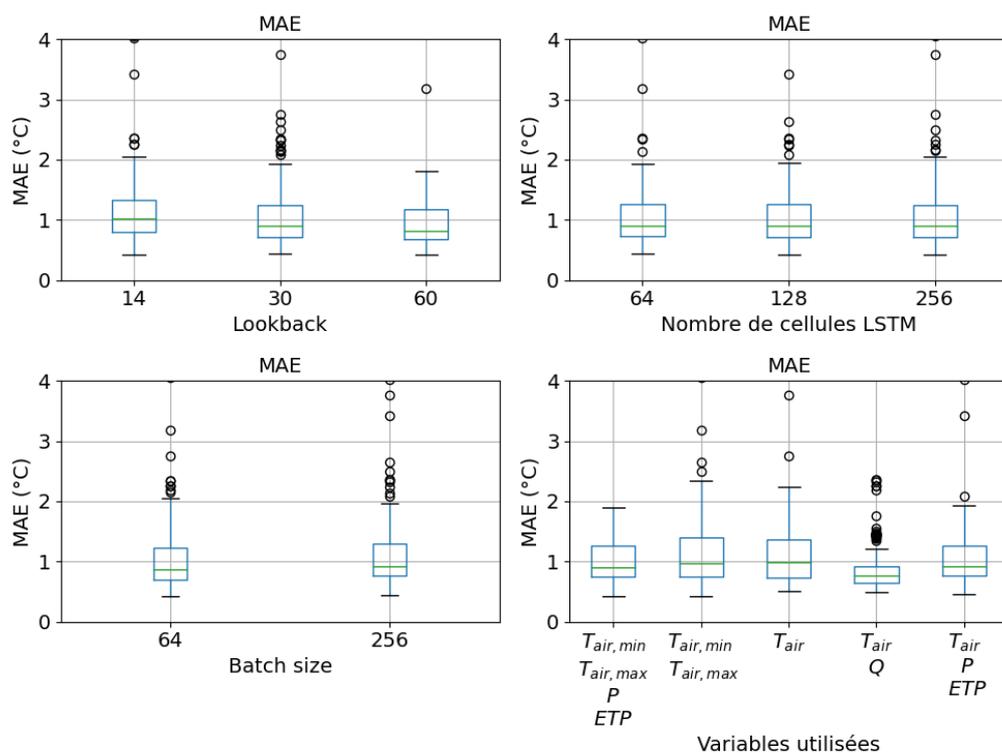


FIGURE 4.1.1 : Distribution des MAE obtenues par les modèles LSTM locaux (modèle par station) en fonction des différentes configurations (Tableau 3.1).

La Figure 4.1.1 présente les performances de tous les modèles locaux testés pour toutes les stations où cela était possible. Les distributions des performances sont peu dépendantes des hyperparamètres "nombre de cellules" et "batch size", se situant en médiane environ à 0.9 °C pour la MAE et 1 pour la précision, la dispersion étant similaire. L'hyperparamètre de *lookback*, en augmentant, fait diminuer la MAE médiane ainsi que la dispersion des performances. Le choix des variables prédictives a lui un impact sur les performances, avec une MAE médiane inférieure à 1 pour les trois jeux de variables intégrant des informations sur l'hydrologie (Q ou P et ETP). La précision médiane ne change pas et les dispersions sont similaires, sauf pour le couple (T_{air}, Q) où les performances sont moins dispersées par rapport à la médiane.

Occurrence des paramètres parmi les 3 meilleurs modèles sur la validation par station

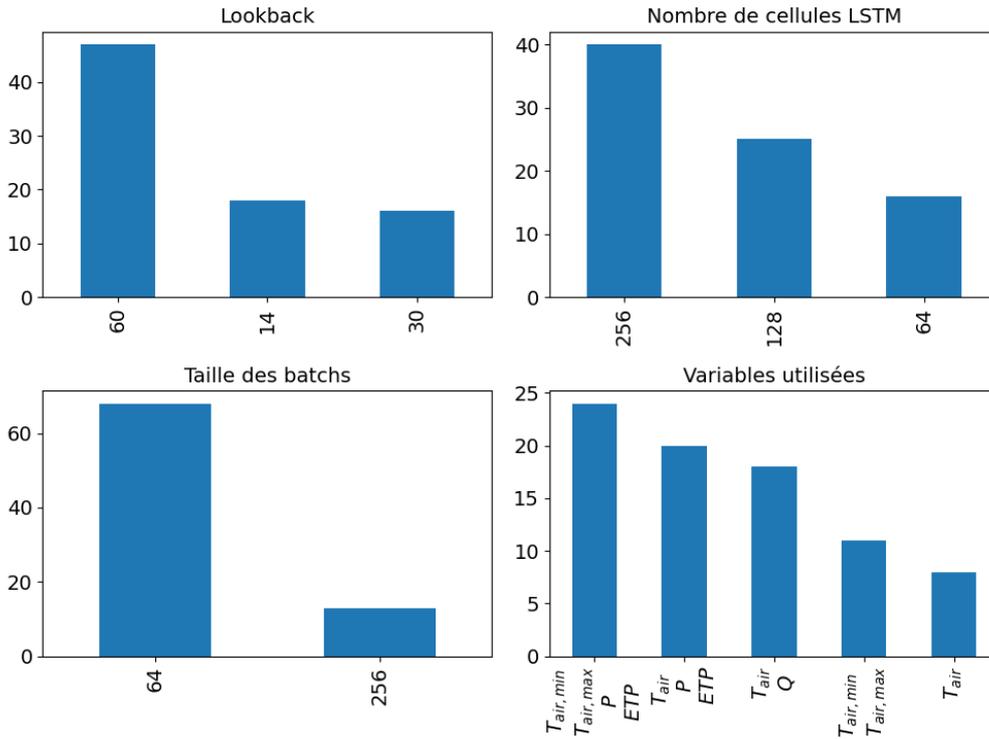


FIGURE 4.1.2 : Hyperparamètres et variables les plus fréquemment rencontrés parmi les 3 meilleurs modèles locaux en MAE par station.

La Figure 4.1.2 présente l'occurrence de chacun des jeux de variables prédictives, après avoir sélectionné pour chaque station uniquement les trois meilleures combinaisons des hyperparamètres testés en terme de performance. Trois modèles sont choisis pour cette tâche, afin de ne pas exclure par mégarde une paramétrisation dont les performances seraient quasi égales au meilleur. Parmi ceux-ci, on trouve une MAE médiane de 0,8 °C, avec un écart-type de 1,0 °C. La précision par rapport au seuil de 19 °C médiane est de 99% avec un écart-type de 5%. Le *lookback* le plus représenté parmi les paramétrisations est celui le plus long proposé (60 jours, 46). Le nombre de cellules le plus représenté est 256 (40) devant 128 (24). Pour la taille des *batchs*, le plus représenté est 64 (63). Pour le choix des variables descriptives, le jeu de variables le plus représenté est $(T_{air,min}, T_{air,max}, P, ETP)$ (24), devant (T_{air}, P, ETP) (19), (T_{air}, Q) (17, combinaison pour laquelle seules 16/27 stations ont pu être testées), $(T_{air,min}, T_{air,max})$ (11) et finalement T_{air} (7). La combinaison (T_{air}, Q) semble ne pas être très intéressante dans ce graphique, mais elle n'a pu être testée que dans une partie des cas, ce qui relativise sa sous-représentation dans cet histogramme final.

Modèles régionaux

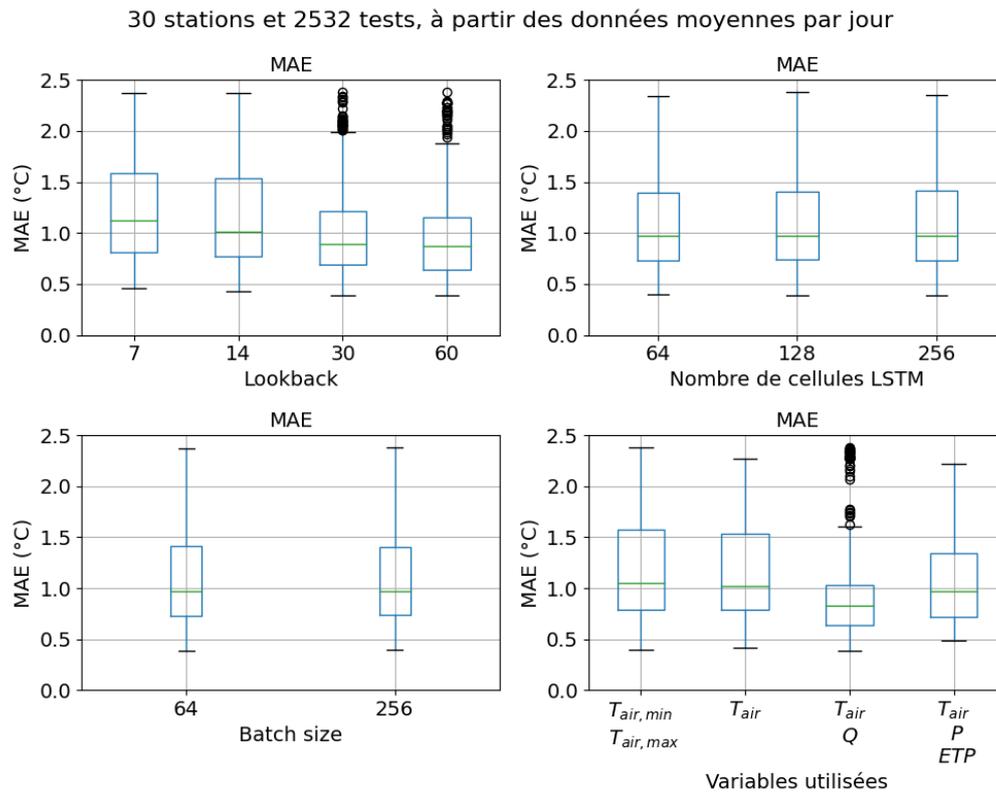


FIGURE 4.1.3 : MAE pour les tests de modèles régionaux

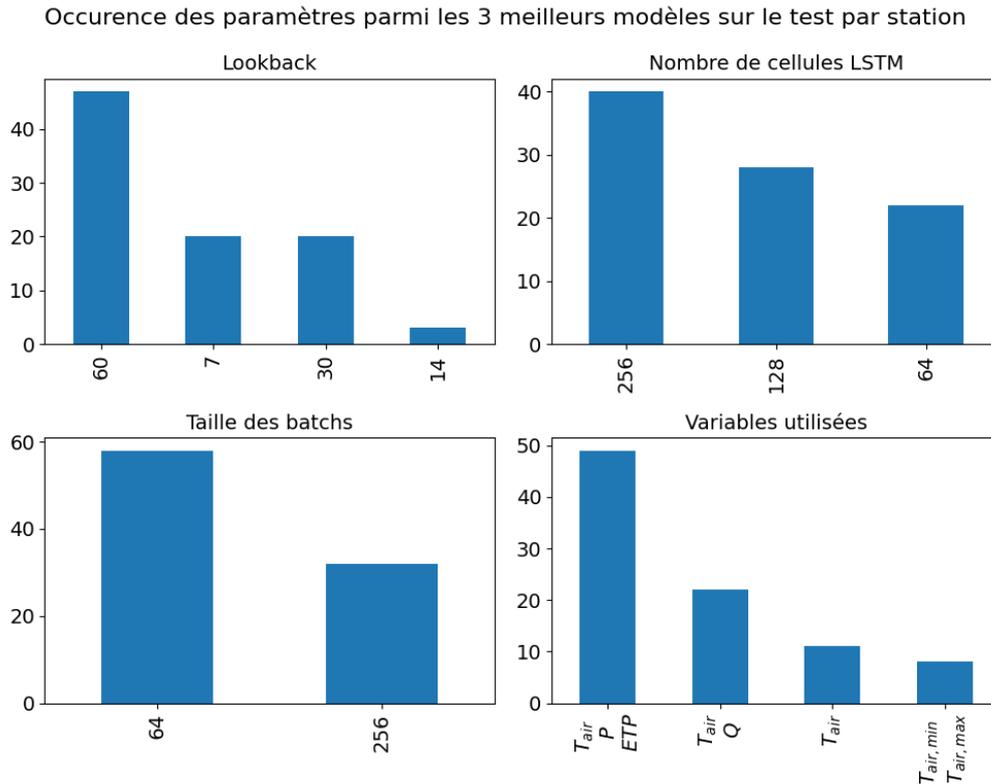


FIGURE 4.1.4 : Hyperparamètres et variables les plus fréquemment rencontrés parmi les 3 meilleurs modèles régionaux en MAE par station.

Les modèles régionaux sont entraînés à partir de toutes les données des stations avec plus d'un an de mesures et testés sur chacune de celles-ci.

Comme observé pour les modèles locaux, on observe sur la Figure 4.1.3 une diminution de la MAE médiane en fonction du lookback, avec une amélioration nette (passage sous les 1 °C, à 0.89 et 0.88 °C) comparable aux performances des modèles locaux à 30 et 60 jours, ainsi qu'une dispersion plus faible et comparable. La taille des *batches* et le nombre de cellules ne semblent toujours pas jouer sur les performances. Le meilleur jeu de paramètres en terme de médiane et de dispersion reste T_{air}, Q , suivi de T_{air}, P, ETP .

On relève également (Figure 4.1.4) une majorité de $lookback=60$ parmi les meilleurs hyperparamètres.

Synthèse

Il est donc à retenir comme meilleurs paramètres un $lookback=60$. Le nombre de cellules plus grand et le *batch size* plus petit diminuant sensiblement la vitesse de calcul pour une performance similaire, 128 cellules et *batch size*=256. sont alors de meilleurs paramètres. Par ailleurs, Rahmani *et al.* (2021b) utilisent un *batch size* égal à la moitié de la taille du jeu de données (d'après expérience personnelle), 100 cellules, et un *lookback* de 365 jours ²⁰

Ces hyperparamètres ne sont pas en absolu les meilleurs, aucune optimisation systématique n'ayant été réalisée ici (pour référence, l'étude de Feigl *et al.* (2021) utilise une optimisation bayésienne pour définir ces hyperparamètres), mais sont les meilleurs parmi ceux testés. Les hyperparamètres et les variables prédictives optimales pour les modèles régionaux sont les mêmes que pour des modèles locaux. Ces paramètres ne définissent pas le meilleur modèle pour chacune des stations, mais une structure de modèle et une résolution

²⁰Nous n'avons pas testé de *lookback* si grand.

optimale en moyenne pour toutes les stations.

Par ailleurs, la précision relative au seuil de 19 °C n'a pas démontré d'utilité dans l'analyse de ces résultats, ses distributions étant presque identiques pour chaque hyperparamètre testé, car ce seuil est rarement dépassé pour la plupart des stations (voir Figure 2.3.1). La MAE calculée sur les 25% de températures les plus élevées, légèrement plus élevée, n'apporte pas non plus d'information supplémentaire à la MAE.

4.1.2. Choix des variables et du type de modèle

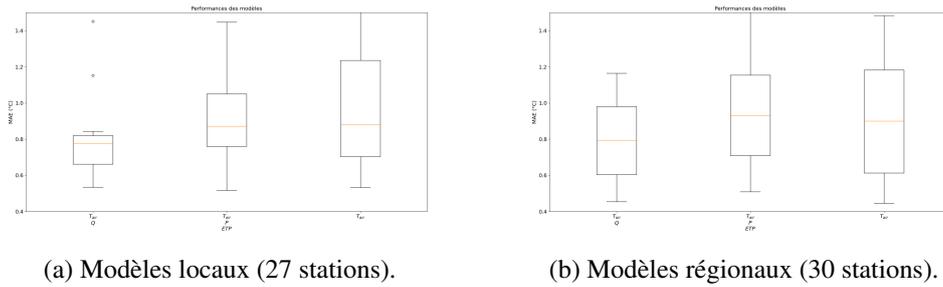


FIGURE 4.1.5 : Distributions des performances des modèles sur toutes les stations en fonction des variables explicatives utilisées.

Les ensembles de paramètres obtenus à la section précédente produisent une MAE médiane de 0,7 °C et 0,8 °C pour les modèles locaux, et de 0,8 °C et 0,9 °C pour les modèles régionaux, respectivement, pour les jeux de variables (T_{air}, Q) et (T_{air}, P, ETP). Ce dernier trio de variables présente la même médiane, mais une plage de distribution plus courte, que les modèles basés uniquement sur T_{air} , ce qui suggère une meilleure fiabilité qui s'ajoute au potentiel d'interpolation spatiale de ce jeu de variables. Ces MAE sont supérieures à celles de Feigl *et al.* (2021) (0,4 °C pour le Danube), mais elles sont cohérentes avec celles obtenues pour divers modèles d'apprentissage automatique sur sept rivières en Europe et aux États-Unis, qui varient de 0,4 °C à 1,2 °C, comme montré dans l'article de Zhu *et al.* (2019b).

Les performances affichées par les modèles régionaux font apparaître la possibilité d'exploiter, par exemple afin de modéliser tous les cours d'eau de Haute-Garonne, un modèle régional dont les variables explicatives seraient uniquement des forçages climatiques T_{air} , P et ETP , en admettant une erreur un peu plus importante qu'un modèle local sur les valeurs les plus fortes. Ce modèle régional pourrait être affiné en définissant par exemple une ou des zones où l'utilisation du même modèle est cohérent, comme la zone amont et la zone aval visibles sur l'analyse de données.

4.2. Illustration sur une station : le Bazacle à Toulouse

Cette partie s'attache à présenter des résultats concrets (performances et séries temporelles) sur une station en particulier, celle du Bazacle à Toulouse. Les débits (station O200001001, la Garonne à Portet-sur-Garonne) ayant été obtenus trop tardivement pour figurer dans les tests généraux ont néanmoins été utilisés ici.

4.2.1. Modèles LSTM

Des modèles LSTM utilisant les hyperparamètres $lookback=60$, $batch\ size=256$ et nombre de cellules=128 ont été utilisés pour visualiser des performances locales de ceux-ci sur la station du Bazacle (Figure 4.2.1, source

de la donnée : MIGADO). Quatre cas ont été choisis : deux modèles locaux calés sur la station et deux modèles régionaux calés sur toutes les stations, et pour chacun, une version utilisant température de l'air et débit et une autre utilisant température de l'air et les lames de précipitation et évapotranspiration à l'échelle du bassin versant.

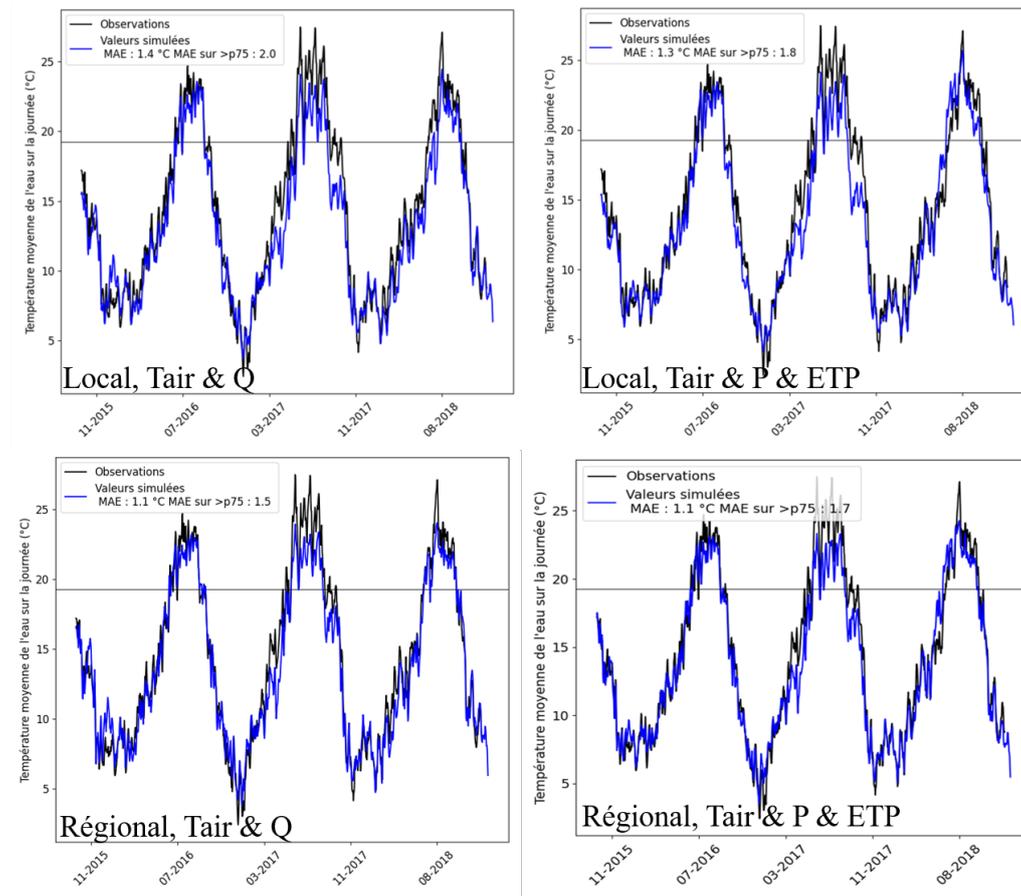


FIGURE 4.2.1 : Chroniques reconstituées pour la Garonne au Bazacle à Toulouse. p75 : température telle que 75 % des mesures sont inférieures à celle-ci.

Nous observons une bonne représentation des dynamiques temporelles, mais une moins bonne représentation des valeurs les plus élevées, et ce quelque soit le modèle. En effet, les modèles sont entraînés en minimisant une erreur moyenne, et sont donc entraînés à représenter des valeurs moyennes, celles-ci étant plus souvent rencontrées.

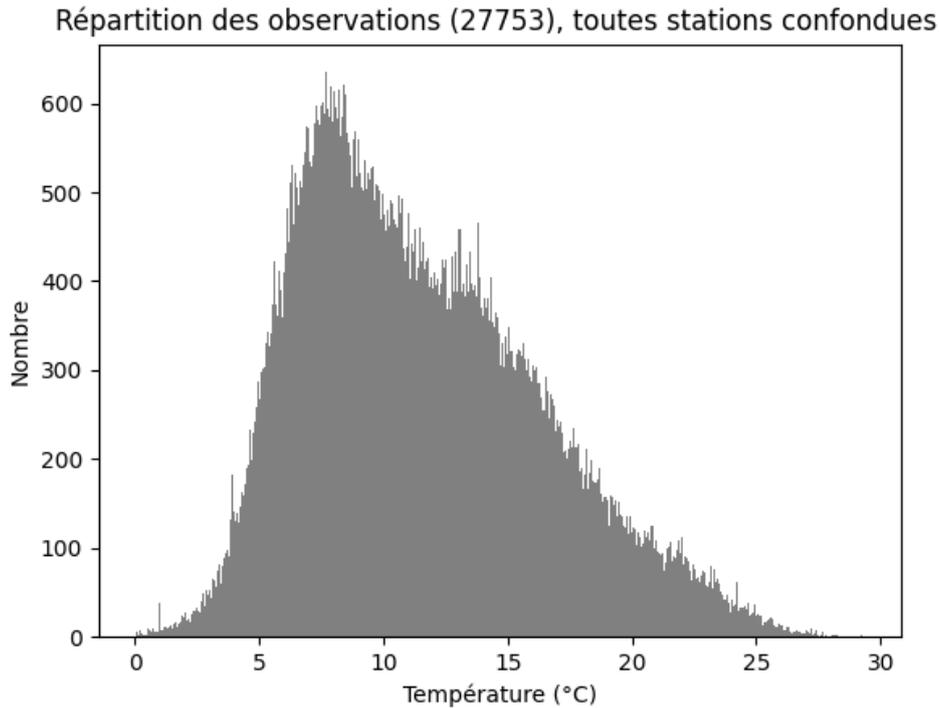


FIGURE 4.2.2 : Distribution des valeurs de température de l'eau dans le jeu de 40 stations.

Les modèles régionaux sont plus performants à cette station que les modèles spécifiques locaux respectifs, ce qui est surprenant. En effet, un modèle régional étant adapté moyennement au territoire, il est attendu qu'un modèle local fournisse de meilleurs résultats. La répartition des observations en figure 4.2.2 montrent que les températures sont majoritairement regroupées entre 5 °C et 15 °C alors que le Bazacle est la station la plus aval et la plus chaude du jeu de données, ce n'est donc pas une station moyenne : les températures supérieures à 19 °C souvent rencontrées (plus de 80 jours par an en moyenne) ne sont pas représentatives du territoire moyen de la Garonne amont. L'annexe B montre les résultats pour ces quatre modèles à la station de Valentine (sur la Garonne également, source de la donnée MIGADO) ; cette fois les modèles locaux aboutissent à de meilleures performances. Toutefois, un modèle régional présente l'avantage de pouvoir être appliqué sur toute la zone, même en l'absence d'observations de température de l'eau.

4.2.2. Comparaison avec le modèle de prévision

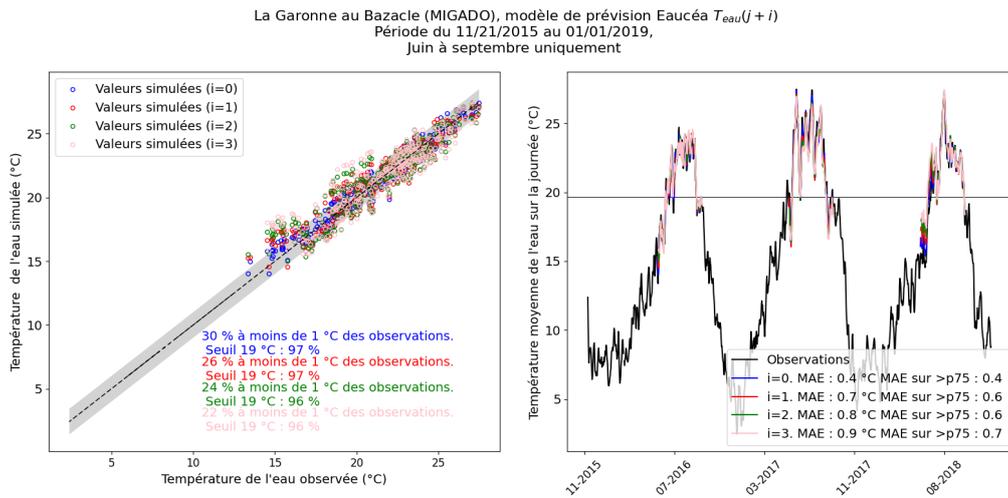


FIGURE 4.2.3 : Performances du modèle opérationnel d'Eaucéa au Bazacle. La prévision à $i=0$ correspond au jour courant.

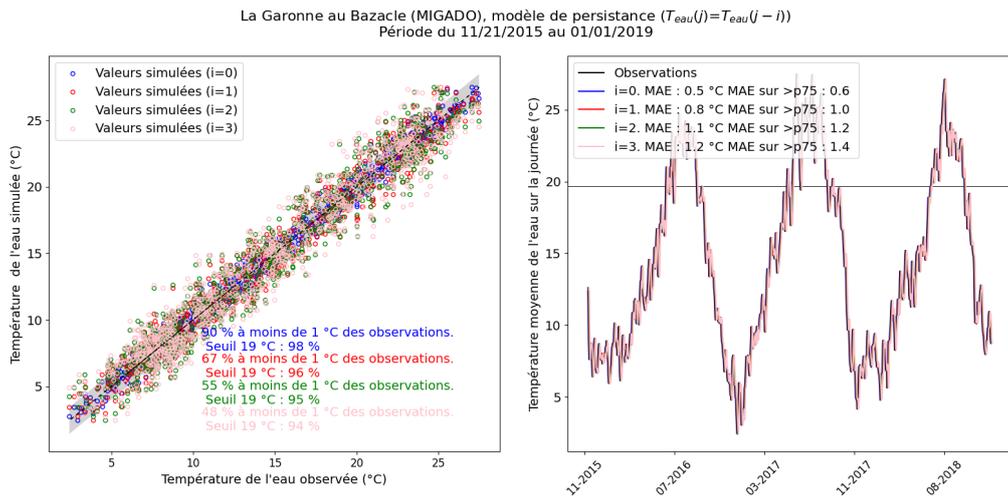


FIGURE 4.2.4 : Résultats pour un modèle naïf de persistance, la prévision de température pour toutes les échéances +0 (jour J), +1, +2 et +3 correspondant à la température moyenne de la veille.

Les résultats du modèle linéaire d'Eaucéa pour le Bazacle sur la même période que les tests effectués sont présentés en figure 4.2.3. On observe que les prévisions sont relativement dispersées par rapport à la droite 1 :1, cependant les MAE sont faibles (0,4 °C pour l'échéance J+1, et toujours inférieures à 1 °C même à une échéance J+3) et sont plus faibles encore pour les valeurs extrêmes (0,7 °C pour l'échéance à J+3).

Ces performances sont encourageantes pour l'objectif opérationnel du CD31 qui a déjà mis en ligne, pour le moment avec un accès restreint, la prévision sur le site <https://www.e-tiage.com/> (voir en Annexe 5.3.1). Elles sont supérieures aux performances des modèles plus complexes LSTM détaillées précédemment, même à une échéance de plusieurs jours. Cela peut vraisemblablement être expliqué par la prise en compte de la température de l'eau du jour précédent dans le calcul, signal portant la plus grande partie de l'information en raison de l'inertie de la masse d'eau, en particulier en plaine. Cette analyse est confirmée par le fait qu'en utilisant un modèle extrêmement simple, dit «de persistance» (Figure 4.2.4), où la prévision à toutes les échéances est prise comme la température moyenne du jour, les performances obtenues sont très similaires à

celles du modèle linéaire d'Eaucéa, la MAE d'un modèle consistant à prendre la température moyenne du jour pour prédire le suivant n'étant que de 0,5 °C.

5. Conclusion et perspectives

Pour conclure, les résultats obtenus pour les modèles locaux comme pour les modèles régionaux sont globalement bons, mais pourraient être améliorés pour obtenir une meilleure représentation des températures estivales, celles dont la prévision nous intéresse (les températures hivernales restent d'un fort intérêt pour la biodiversité, pour lequel le maintien du réseau de suivi se justifie l'hiver, mais ne présentent pas un intérêt en terme de prévision). Les résultats en «mode prévision» seront à comparer avec le modèle plus simple, mais efficace et déjà opérationnel, développé par Eaucéa pour le Bazacle. Un exercice de prévision étant différent d'un exercice de simulation, il est difficile de comparer objectivement les performances d'un modèle destiné à prévoir les valeurs futures en connaissant le passé (comme celui d'Eaucéa) et d'un modèle destiné à retrouver des valeurs lorsqu'elles ne sont pas disponibles, sans utiliser les mesures passées de température de l'eau (comme celui de ce stage).

L'étude a permis de mettre en évidence des performances encourageantes d'un modèle utilisant les forçages sur une longue période, la longueur de séquence prédictive optimale étant de 60 jours. L'intégration du débit améliore sensiblement ces performances, mais celui-ci n'est malheureusement pas toujours disponible, freinant l'applicabilité d'un tel modèle à tout le réseau hydrographique du département. L'utilisation des forçages climatiques uniquement (température de l'air et précipitations) est satisfaisante, mais en implicitant la relation pluie-débit dans le système, impose une réserve pour un bassin versant dont le débit est régulé en période sèche par les opérations de soutien d'étiage, comme dans le cas de la Garonne. En effet, cela induit un découplage entre les dynamiques des précipitations observées (qui seront donc presque nulles l'été) et du débit qui dépend alors presque uniquement de la gestion de la ressource stockée. On pourrait toutefois considérer que l'information «sécheresse» et/ou «gestion» soit contenue dans l'absence prolongée de pluie estivale et/ou les températures de l'eau utilisées pour l'apprentissage. Néanmoins cette approche pourrait être généralisée à l'ensemble des cours d'eau du PTGA afin d'en simuler la température.

Il est observé une erreur souvent plus importante sur les stations dont la durée de disponibilité des mesures, et donc de test est plus importante alors que la période d'entraînement est d'autant plus longue, cela est vraisemblablement dû à la présence de plus de périodes estivales sur ce test, pour lesquelles les températures sont plus difficilement estimées.

Il pourrait également être investiguée la prise en compte de divers attributs statiques comme l'ombre projetée par la ripisylve (par exemple via des données LiDAR (Light Detection And Ranging)) ou d'autres méthodes basées sur l'occupation des sols (Loicq *et al.* (2018)).

Il est observé que l'entraînement de nombreux modèles sur un ordinateur de bureau a tendance à être très lent (de l'ordre de plusieurs dizaines d'heures), ce qui est pénalisant puisque certains tests doivent être réalisés plusieurs fois. Les recherches de paramètres optimaux ont été faites de manière naïve et sur un jeu de paramètres non continu. Il aurait sans doute été judicieux d'avoir recours à une machine plus performante ce qui aurait permis d'utiliser une méthode d'optimisation automatique des hyperparamètres adaptée comme Hyperopt (Bergstra *et al.* (2013)) ou Optuna (Akiba *et al.* (2019)).

Références

- ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J., DEVIN, M., GHEMAWAT, S., GOODFELLOW, I., HARP, A., IRVING, G., ISARD, M., JIA, Y., JOZEFOWICZ, R., KAISER, L., KUDLUR, M., LEVENBERG, J., MANE, D., MONGA, R., MOORE, S., MURRAY, D., OLAH, C., SCHUSTER, M., SHLENS, J., STEINER, B., SUTSKEVER, I., TALWAR, K., TUCKER, P., VANHOUCHE, V., VASUDEVAN, V., VIEGAS, F., VINYALS, O., WARDEN, P., WATTENBERG, M., WICKE, M., YU, Y. et ZHENG, X. (2016). TensorFlow : Large-Scale Machine Learning on Heterogeneous Distributed Systems. arXiv :1603.04467 [cs].
- AKIBA, T., SANO, S., YANASE, T., OHTA, T. et KOYAMA, M. (2019). Optuna : A Next-generation Hyperparameter Optimization Framework. *In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2623–2631, Anchorage AK USA. ACM.
- AUFFRAY, M., SENÉCAL, J.-F., TURGEON, K., ST-HILAIRE, A. et MAHEU, A. (2023). Reservoirs regulated by small dams have a similar warming effect than lakes on the summer thermal regime of streams. *Science of The Total Environment*, 869:161445.
- BARZEGAR, R., AALAMI, M. T. et ADAMOWSKI, J. (2020). Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. *Stochastic Environmental Research and Risk Assessment*, 34(2):415–433.
- BEAUFORT, MOATAR, F. et SAUQUET, E. (2020). Projet TIGRE - Thermie en rivière : Analyse géostatistique et description de régime : Application à l'échelle de la France. Rapport technique, INRAE UR Riverly, Université de Tours GÉHCO.
- BEAUFORT, A., CURIE, F., MOATAR, F., DUCHARNE, A., MELIN, E. et THIERY, D. (2016). T-NET, a dynamic model for simulating daily stream temperature at the regional scale based on a network topology : T-NET model for simulating stream temperature at a regional scale. *Hydrological Processes*, 30(13):2196–2210.
- BEJANI, M. M. et GHATEE, M. (2021). A systematic review on overfitting control in shallow and deep neural networks. *Artificial Intelligence Review*, 54(8):6391–6438.
- BENGIO, Y., SIMARD, P. et FRASCONI, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166.
- BERGSTRA, J., YAMINS, D. et COX, D. (2013). Hyperopt : A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms. pages 13–19, Austin, Texas.
- BUSTILLO, V., MOATAR, F., DUCHARNE, A., THIÉRY, D. et POIREL, A. (2014). A multimodel comparison for assessing water temperatures under changing climate conditions via the equilibrium temperature concept : case study of the Middle Loire River, France : THERMAL MODELS FOR CLIMATE CHANGE STUDIES. *Hydrological Processes*, 28(3):1507–1524.
- CAISSIE, D. (2006). The thermal regime of rivers : a review. *Freshwater Biology*, 51(8):1389–1406.
- CARDON, D., COINTET, J.-P. et MAZIÈRES, A. (2018). La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, n° 211(5):173–220.

- CHA, Y., CHO, K. H., LEE, H., KANG, T. et KIM, J. H. (2017). The relative importance of water temperature and residence time in predicting cyanobacteria abundance in regulated rivers. *Water Research*, 124:11–19.
- CHO, K., van MERRIENBOER, B., GULCEHRE, C., BAHDANAU, D., BOUGARES, F., SCHWENK, H. et BENGIO, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar. Association for Computational Linguistics.
- CHOLLET, F. et AL. (2015). Keras.
- DUCHARNE, A. (2008). Importance of stream temperature to climate change impact on water quality. *Hydrology and Earth System Sciences*, 12(3):797–810.
- DUFEU, E., MOUGIN, F., FORAY, A., BAILLON, M., LAMBLIN, R., HEBRARD, F., CHALEON, C., ROMON, S., COBOS, L., GOUIN, P., AUDOUY, J.-N., MARTIN, R. et POLIGOT-PITSCH, S. (2022). Finalisation de l’opération HYDRO 3 de modernisation du système d’information national des données hydrométriques. *LHB*, 108(1):2099317.
- DUGDALE, S. J., HANNAH, D. M. et MALCOLM, I. A. (2017). River temperature modelling : A review of process-based approaches and future directions. *Earth-Science Reviews*, 175:97–113.
- FEIGL, M., LEBIEDZINSKI, K., HERRNEGGER, M. et SCHULZ, K. (2021). Machine-learning methods for stream water temperature prediction. *Hydrology and Earth System Sciences*, 25(5):2951–2977.
- FENG, D., FANG, K. et SHEN, C. (2020). Enhancing Streamflow Forecast and Extracting Insights Using Long-Short Term Memory Networks With Data Integration at Continental Scales. *Water Resources Research*, 56(9).
- FERCHICHI, H., ST-HILAIRE, A., BERTHOT, L. et CAISSIE, D. (2023). Augmenting environmental flow information with water temperature : case study in Eastern Canada. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 48(2):111–131.
- GALLICE, A., SCHAEFLI, B., LEHNING, M., PARLANGE, M. B. et HUWALD, H. (2015). Stream temperature prediction in ungauged basins : review of recent approaches and description of a new physics-derived statistical model. *Hydrology and Earth System Sciences*, 19(9):3727–3753.
- GU, R., MONTGOMERY, S. et AUSTIN, T. A. (1998). Quantifying the effects of stream discharge on summer river temperature. *Hydrological Sciences Journal*, 43(6):885–904.
- HANNAH, D. M. et GARNER, G. (2015). River water temperature in the United Kingdom : Changes over the 20th century and possible changes over the 21st century. *Progress in Physical Geography : Earth and Environment*, 39(1):68–92.
- HASHEMI, R., BRIGODE, P., GARAMBOIS, P.-A. et JAVELLE, P. (2022). How can we benefit from regime information to make more effective use of long short-term memory (LSTM) runoff models? *Hydrology and Earth System Sciences*, 26(22):5793–5816.
- HOCHREITER, S. et SCHMIDHUBER, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780.

- HUANG, F., QIAN, B. et OCHOA, C. G. (2023). Long-term river water temperature reconstruction and investigation : A case study of the Dongting Lake Basin, China. *Journal of Hydrology*, 616:128857.
- HUNTER, J. D. (2007). Matplotlib : A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95.
- HUSSAIN, M. et MAHMUD, I. (2019). pyMannKendall : a python package for non parametric Mann Kendall family of trend tests. *Journal of Open Source Software*, 4(39):1556.
- KRATZERT, F., KLOTZ, D., BRENNER, C., SCHULZ, K. et HERRNEGGER, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11):6005–6022.
- LABBOUZ, L. (2022). Fiche action D 1.2 - Observatoire thermique des cours d'eau et des nappes alluviales - Rapport annuel 2021-2022. Rapport technique. Eauceá 72 rue Riquet - Bat A 31000 Toulouse 05 61 62 50 68 eaucea@eaucea.fr www.eaucea.fr.
- LABBOUZ, L. (2023). Fiche action D 1.2 - Observatoire thermique des cours d'eau et des nappes alluviales - Modélisation statistique et prévision de la température de l'eau. Rapport technique. Eauceá 72 rue Riquet - Bat A 31000 Toulouse 05 61 62 50 68 eaucea@eaucea.fr www.eaucea.fr.
- LARNIER, K. (2010). *Modélisation thermohydraulique d'un tronçon de Garonne en lien avec l'habitat piscicole : approches statistique et déterministe*. phd.
- LEACH, J. A., KELLEHER, C., KURYLYK, B. L., MOORE, R. D. et NEILSON, B. T. (2023). A primer on stream temperature processes. *WIREs Water*.
- LELEU, I., TONNELIER, I., PUECHBERTY, R., GOUIN, P., VIQUENDI, I., COBOS, L., FORAY, A., BAILLON, M. et NDIMA, P.-O. (2014). La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La Houille Blanche*, (1):25–32.
- LOICQ, P., MOATAR, F., JULLIAN, Y., DUGDALE, S. J. et HANNAH, D. M. (2018). Improving representation of riparian vegetation shading in a regional stream temperature model using LiDAR data. *Science of The Total Environment*, 624:480–490.
- MAIRE, A. (2021). Synthèse du programme de recherche Thermie-Hydrobiologie 2016-2020. Note technique 6125-3313-2021-01078-FR, EDF - LNHE.
- MCKINNEY, W. (2010). Data Structures for Statistical Computing in Python. *In Proceedings of the 9th Python in Science Conference*, pages 51–56, Austin, Texas, USA.
- MOATAR, F. (2021). Projet TIGRE - INRAE.
- MOATAR, F. et GAILHARD, J. (2006). Water temperature behaviour in the River Loire since 1976 and 1881. *Comptes Rendus Geoscience*, 338(5):319–328.
- NCARS (2014). National Center for Atmospheric Research Staff. The Climate Data Guide : Trend Analysis.
- LOUDIN, L., HERVIEU, F., MICHEL, C., PERRIN, C., ANDRÉASSIAN, V., ANCTIL, F. et LOUMAGNE, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? *Journal of Hydrology*, 303(1-4):290–306.

- OUELLET, V., ST-HILAIRE, A., DUGDALE, S. J., HANNAH, D. M., KRAUSE, S. et PROULX-OUELLET, S. (2020). River temperature research and practice : Recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. *Science of The Total Environment*, 736:139679.
- PASZKE, A., GROSS, S., MASSA, F., LERER, A., BRADBURY, J., CHANAN, G., KILLEEN, T., LIN, Z., GIMEL-SHEIN, N., ANTIGA, L., DESMAISON, A., KOPF, A., YANG, E., DEVITO, Z., RAISON, M., TEJANI, A., CHILAMKURTHY, S., STEINER, B., FANG, L., BAI, J. et CHINTALA, S. (2019). PyTorch : An Imperative Style, High-Performance Deep Learning Library. In WALLACH, H., LAROCHELLE, H., BEYGELZIMER, A., d'Alché BUC, F., FOX, E. et GARNETT, R., éditeurs : *proceedings*, pages 8024–8035. Curran Associates, Inc.
- PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., MÜLLER, A., NOTHMAN, J., LOUPPE, G., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M. et DUCHESNAY, (2012). Scikit-learn : Machine Learning in Python. Publisher : arXiv Version Number : 4.
- PETITJEAN, S. (2022). Récolter en cours d'eau des données thermiques de qualité. Rapport technique, Fédération de la Corrèze pour la Pêche et la Protection du Milieu Aquatique.
- PICARD, C., FLOURY, M., SEYEDHASHEMI, H., MOREL, M., PELLA, H., LAMOUREUX, N., BUISSON, L., MOATAR, F. et MAIRE, A. (2022). Direct habitat descriptors improve the understanding of the organization of fish and macroinvertebrate communities across a large catchment. *PLOS ONE*, 17(9):e0274167.
- RAHEL, F. J. et OLDEN, J. D. (2008). Assessing the Effects of Climate Change on Aquatic Invasive Species. *Conservation Biology*, 22(3):521–533.
- RAHMANI, F., LAWSON, K., OUYANG, W., APPLING, A., OLIVER, S. et SHEN, C. (2021a). Exploring the exceptional performance of a deep learning stream temperature model and the value of streamflow data. *Environmental Research Letters*, 16(2):024025. Publisher : IOP Publishing.
- RAHMANI, F., SHEN, C., OLIVER, S., LAWSON, K. et APPLING, A. (2021b). Deep learning approaches for improving prediction of daily stream temperature in data-scarce, unmonitored, and dammed basins. *Hydrological Processes*, 35(11).
- RAJESH, M. et REHANA, S. (2022). Impact of climate change on river water temperature and dissolved oxygen : Indian riverine thermal regimes. *Scientific Reports*, 12(1):9222.
- RIVIÈRE, A., LADET, D., THOMAS, W., LE BRETON, G., DUCHARNE, A. et OUDIN, L. (2021). Projections des températures de l'eau de la Seine à Paris à l'horizon 2100. In *Rapport PIREN SEINE 2021*.
- SEYEDHASHEMI, H., MOATAR, F., VIDAL, J.-P., DIAMOND, J. S., BEAUFORT, A., CHANDESRI, A. et VALLETTE, L. (2021). Thermal signatures identify the influence of dams and ponds on stream temperature at the regional scale. *Science of The Total Environment*, 766:142667.
- SEYEDHASHEMI, H., VIDAL, J.-P., DIAMOND, J. S., THIÉRY, D., MONTEIL, C., HENDRICKX, F., MAIRE, A. et MOATAR, F. (2022). Regional, multi-decadal analysis on the Loire River basin reveals that stream temperature increases faster than air temperature. *Hydrology and Earth System Sciences*, 26(9):2583–2603.
- SHEN, C. (2018). A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resources Research*, 54(11):8558–8593.

- SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. et SALAKHUTDINOV, R. (2014). Dropout : a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958. Publisher : JMLR. org.
- STAJKOWSKI, S., KUMAR, D., SAMUI, P., BONAKDARI, H. et GHARABAGHI, B. (2020). Genetic-Algorithm-Optimized Sequential Model for Water Temperature Prediction. *Sustainability*, 12(13):5374.
- TABARY, P., DUPUY, P., L'HENAFF, G., GUEGUEN, C., MOULIN, L., LAURANTIN, O., MERLIER, C. et SOUBEYROUX, J.-M. (2012). A 10-year (1997–2006) reanalysis of Quantitative Precipitation Estimation over France : methodology and first results. *IAHS Publication*, 351:255–260.
- TOFFOLON, M. et PICCOLROAZ, S. (2015). A hybrid model for river water temperature as a function of air temperature and discharge. *Environmental Research Letters*, 10(11):114011.
- VAN DER WALT, S., COLBERT, S. C. et VAROQUAUX, G. (2011). The NumPy array : a structure for efficient numerical computation. Publisher : arXiv Version Number : 1.
- van VLIET, M. T. H., YEARSLEY, J. R., FRANSSSEN, W. H. P., LUDWIG, F., HADDELAND, I., LETTENMAIER, D. P. et KABAT, P. (2012). Coupled daily streamflow and water temperature modelling in large river basins. *Hydrology and Earth System Sciences*, 16(11):4303–4321.
- VIDAL, J.-P., MARTIN, E., FRANCHISTÉGUY, L., BAILLON, M. et SOUBEYROUX, J.-M. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system : A 50-YEAR HIGH-RESOLUTION ATMOSPHERIC REANALYSIS OVER FRANCE. *International Journal of Climatology*, 30(11):1627–1644.
- WANDERS, N., VLIET, M. T. H., WADA, Y., BIERKENS, M. F. P. et BEEK, L. P. H. R. (2019). High-Resolution Global Water Temperature Modeling. *Water Resources Research*, 55(4):2760–2778.
- WEBB, B. W., HANNAH, D. M., MOORE, R. D., BROWN, L. E. et NOBILIS, F. (2008). Recent advances in stream and river temperature research. *Hydrological Processes*, 22(7):902–918.
- ZHU, S., HEDDAM, S., NYARKO, E. K., HADZIMA-NYARKO, M., PICCOLROAZ, S. et WU, S. (2019a). Modeling daily water temperature for rivers : comparison between adaptive neuro-fuzzy inference systems and artificial neural networks models. *Environmental Science and Pollution Research*, 26(1):402–420.
- ZHU, S., NYARKO, E. K., HADZIMA-NYARKO, M., HEDDAM, S. et WU, S. (2019b). Assessing the performance of a suite of machine learning models for daily river water temperature prediction. *PeerJ*, 7:e7065.

Annexes

A Commentaires sur le critère de précision

Sur 27 stations et 2268 tests, précision par rapport au dépassement du seuil 19°C

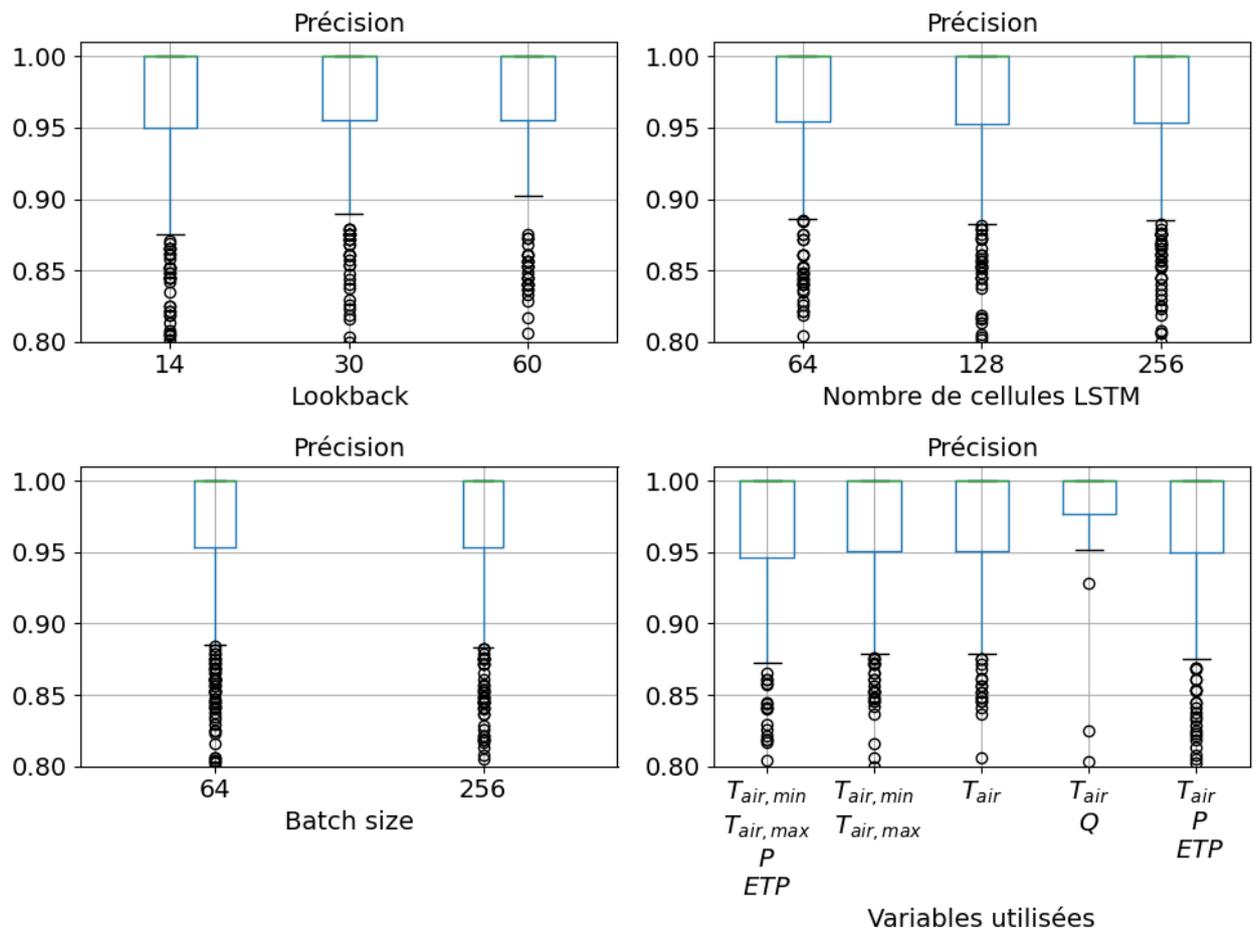


FIGURE 5.1.1 : Précision par rapport au seuil de 19 °C pour les modèles locaux.

Sur 30 stations et 2532 tests, précision par rapport au dépassement du seuil 19°C

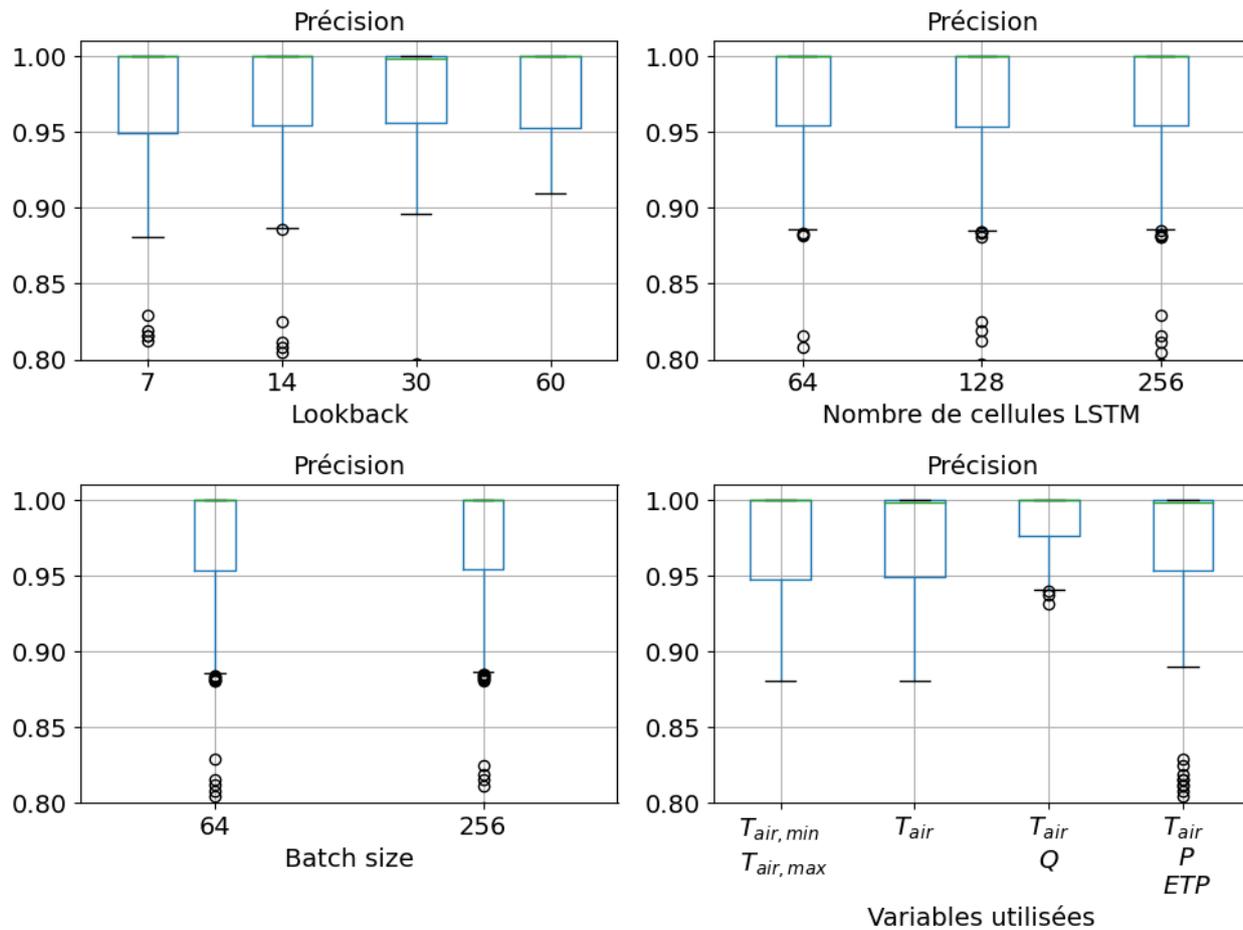


FIGURE 5.1.2 : Précision par rapport au seuil de 19 °C pour les modèles régionaux.

Ces distributions montrent que le critère de précision n'apporte pas d'information à l'étude. En effet, la température de 19 °C n'est rencontrée qu'anecdotiquement sur une bonne partie du territoire (Figure 2.3.1).

B La Garonne à Valentine

Sur les 4 figures suivantes, le graphe de gauche indique la corrélation des simulations aux observations lors du test, la dispersion des simulations, et la précision par rapport au seuil de 19 °C. Le graphe de droite présente la série temporelle reconstituée et la MAE, ainsi que la valeur correspondant aux 25% des températures les plus chaudes sur la période et la MAE associée.

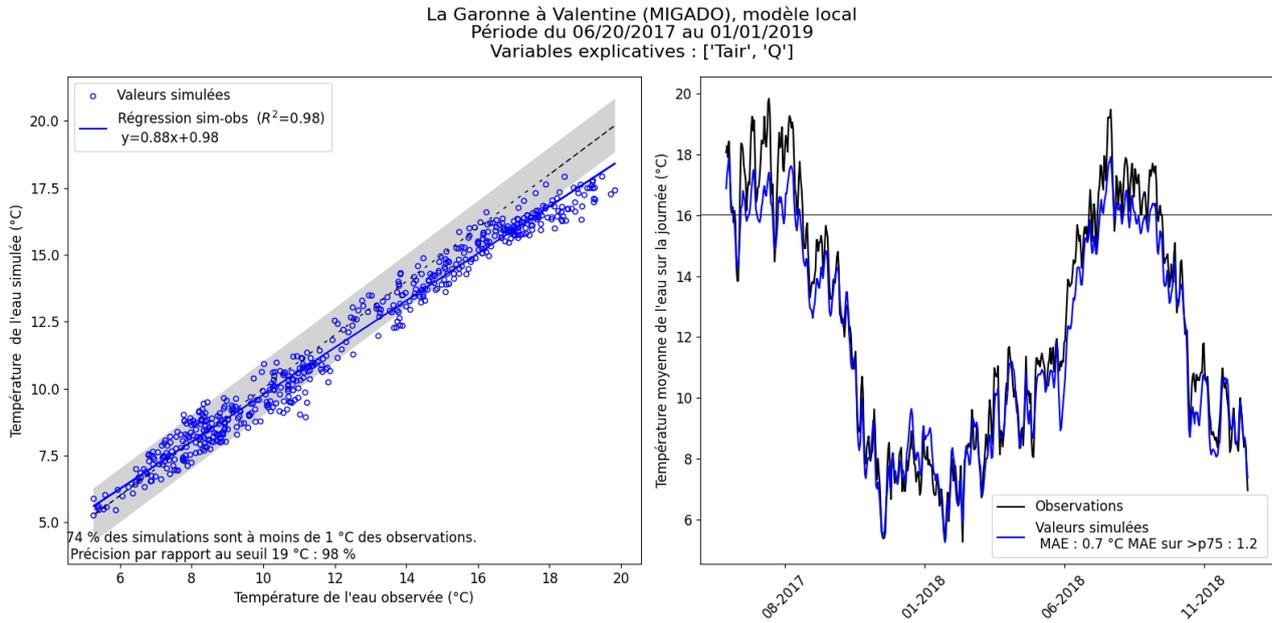


FIGURE 5.2.1 : Performances d'un modèle local utilisant T_{air} et Q .

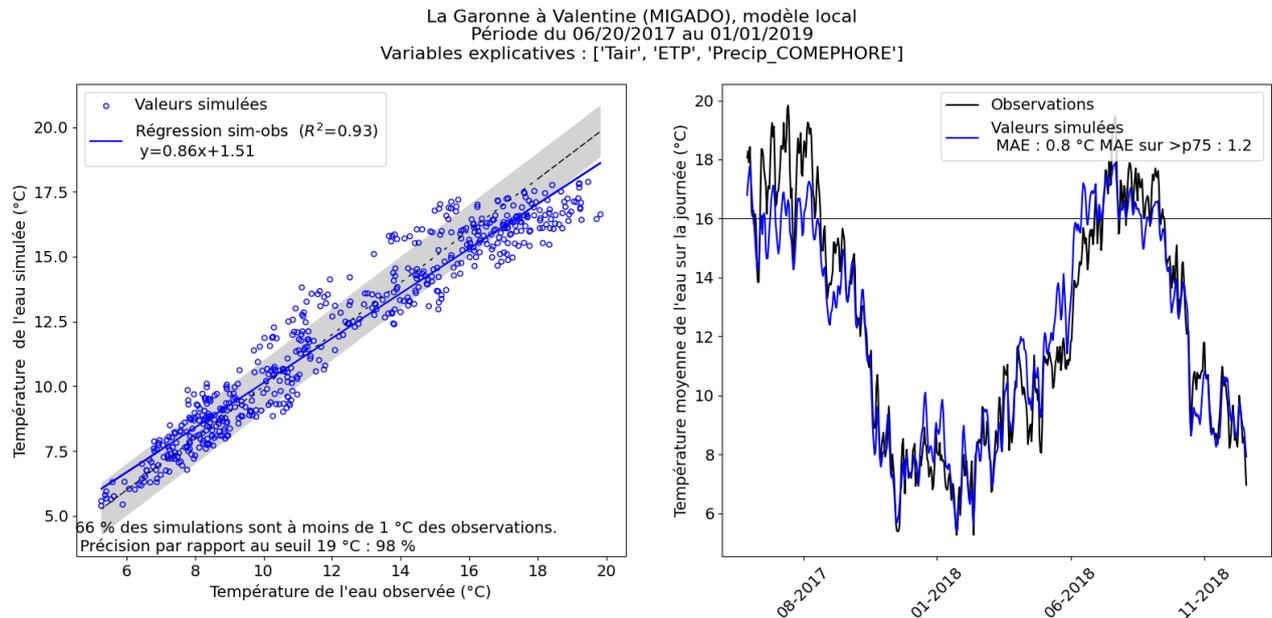


FIGURE 5.2.2 : Performances d'un modèle local utilisant T_{air} , P et ETP .

La Garonne à Valentine (MIGADO), modèle régional
 Période du 06/20/2017 au 01/01/2019
 Variables explicatives : ['Tair', 'Q']

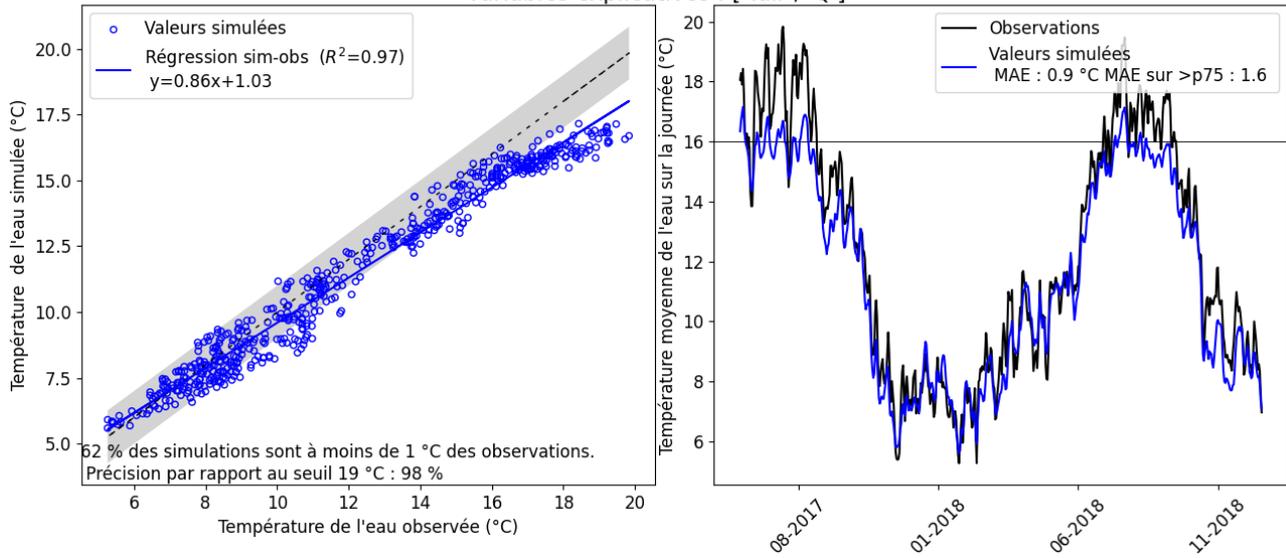


FIGURE 5.2.3 : Performances d'un modèle régional utilisant T_{air} et Q .

La Garonne à Valentine (MIGADO), modèle régional
 Période du 06/20/2017 au 01/01/2019
 Variables explicatives : ['Tair', 'ETP', 'Precip_COMEPHORE']

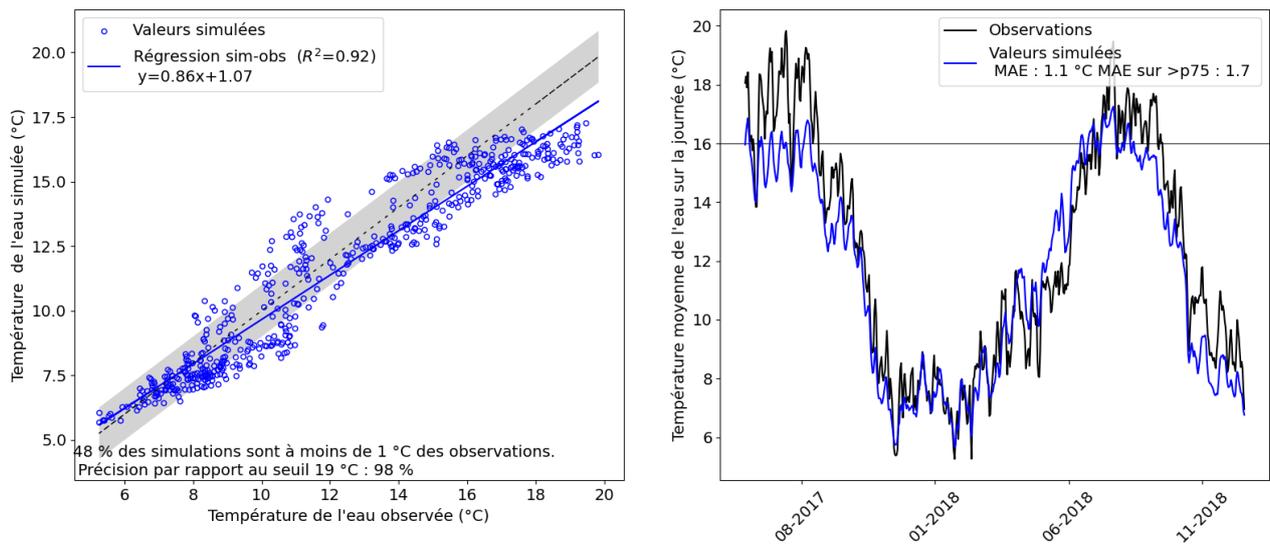


FIGURE 5.2.4 : Performances d'un modèle régional utilisant T_{air} , P et ETP .

C Outil de prévision opérationnel du CD31

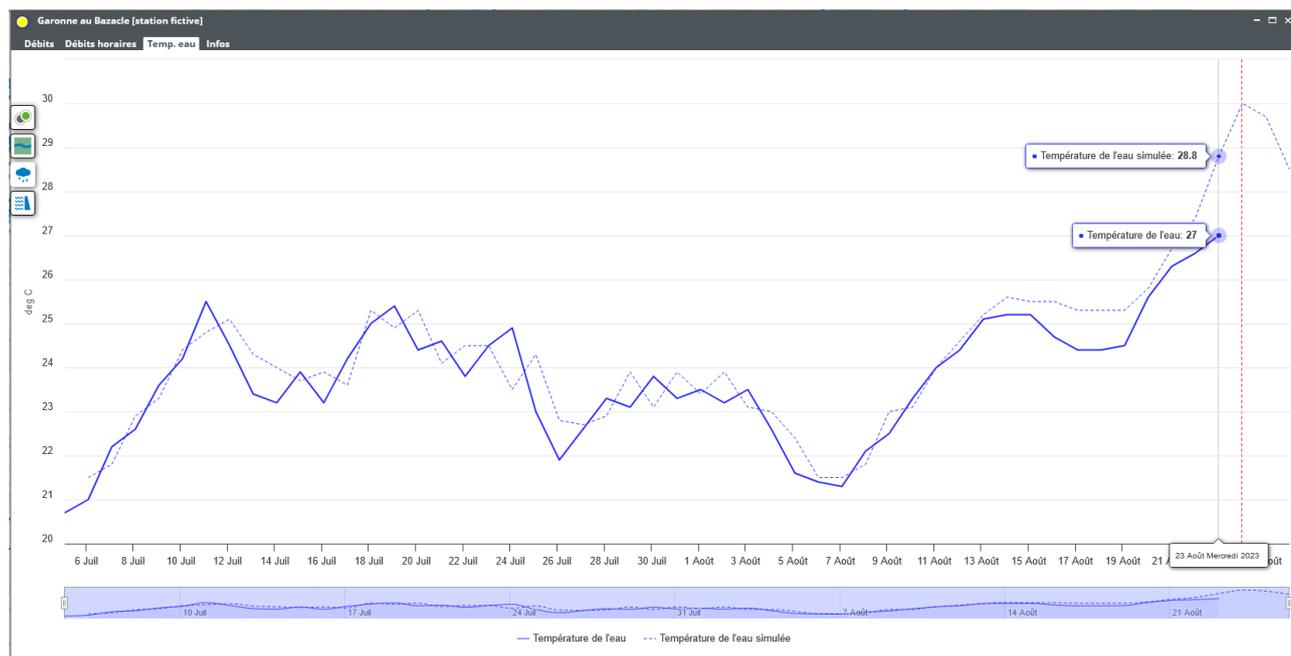


FIGURE 5.3.1 : Capture d'écran du site <https://www.e-tiage.com/>. Vue pour le moment limitée à l'accès du CD31.

Les valeurs prédites pendant la vague de chaleur de la semaine du 20 août 2023 sont très supérieures aux températures réelles (notons par exemple une erreur de 1.8 °C le 23 août). Cela est lié à la forte prise en compte de la température de l'air dans le modèle, qui a atteint de manière exceptionnelle quelques 42,3 °C à Blagnac (source infoclimat, consulté le 24/08/2023), conjuguée avec la rareté de la température observée ce jour-là.

Entre janvier 1991 et mi-juin 2022, la température moyenne sur la journée observée de 27 °C n'avait été dépassée que 14 fois (sur 11490 jours de mesures soit 0,1 %, ou 3798 jours d'été (juin-septembre) soit 0.4%), et 25 °C dépassée 175 fois (soit 1,5 % ou 5% des jours d'été). Ce type de valeur est rarement vu par le modèle sensé les simuler. Ainsi, il est difficile d'obtenir un modèle capable de reproduire fidèlement ces températures spectaculaires, quand bien même il soit centré sur la période estivale.

D Utilisation de la température au jour précédent : exemple du Bazacle

Les températures de l'eau des jours passés n'ont volontairement pas été incluses dans notre étude. Cependant, des modèles utilisant les hyperparamètres optimaux déterminés précédemment ont été testés sur la station du Bazacle. La performance du modèle utilisant seulement T_{eau} est fortement similaire au modèle de persistance (Figure 4.2.4) pour bien plus de complexité.

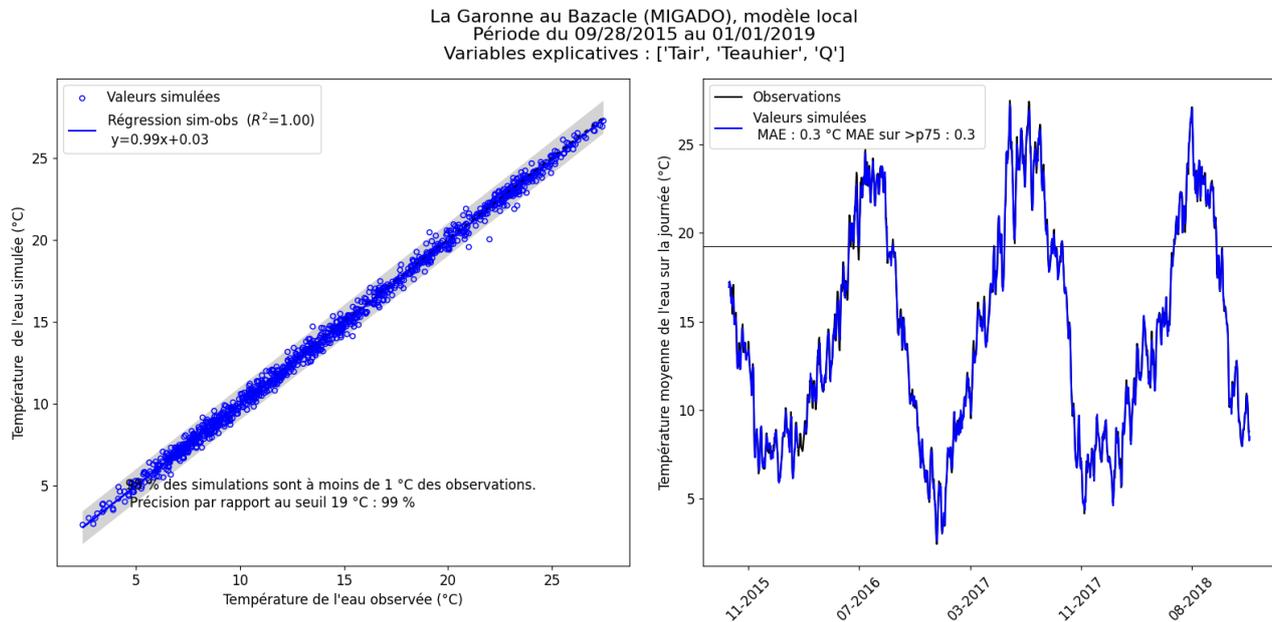


FIGURE 5.4.1 : Utilisation de T_{air} , Q et T_{eau} .

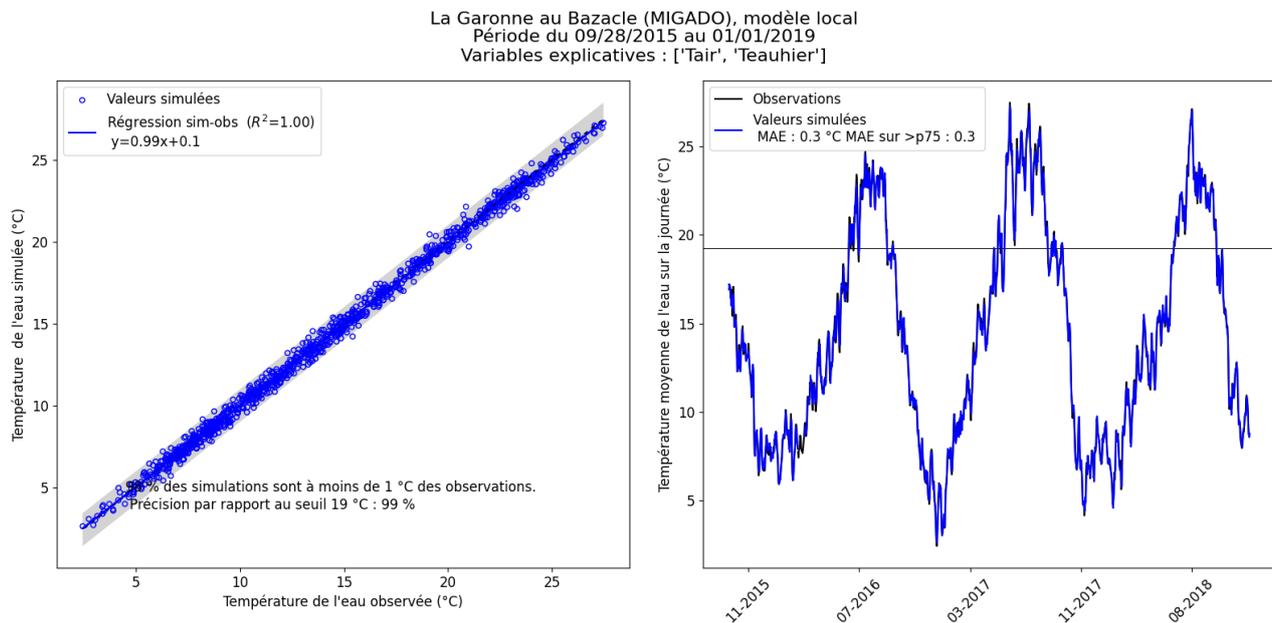


FIGURE 5.4.2 : Utilisation de seulement T_{air} et T_{eau} .

La Garonne au Bazacle (MIGADO), modèle local
Période du 09/28/2015 au 01/01/2019
Variables explicatives : ['Teauhier']

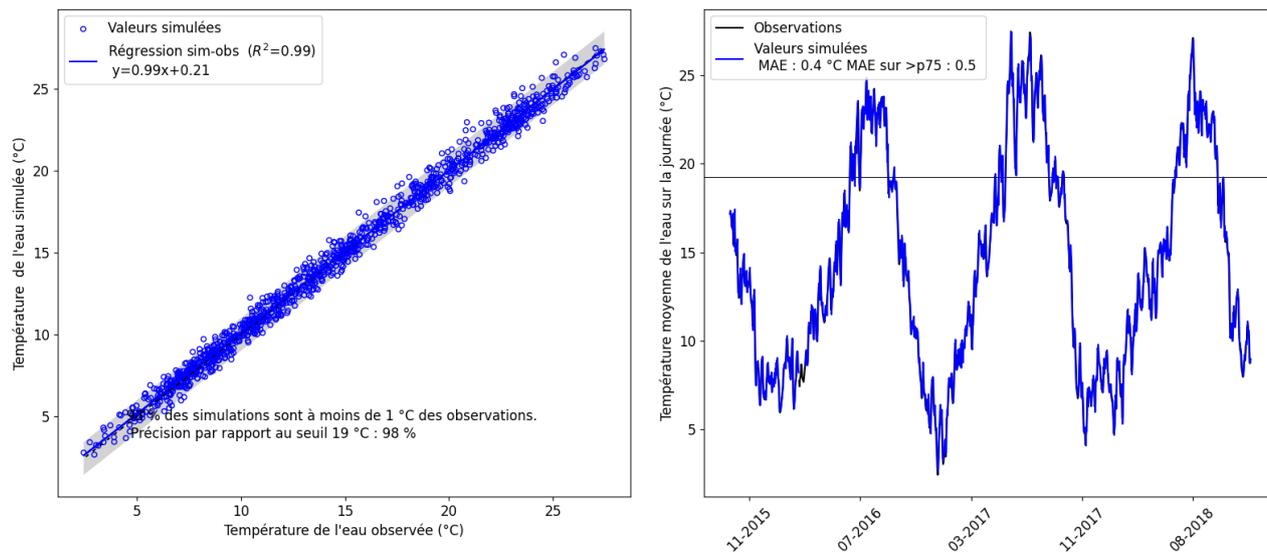


FIGURE 5.4.3 : Utilisation de seulement T_{eau} .