

Extrapolation spatio-temporelle des températures des cours d'eau en temps présent et futur à l'échelle nationale

Marion Moussay¹, André St-Hilaire², Florentina Moatar¹

¹INRAE, UR RiverLy, Centre Lyon-Grenoble Rhône Alpes, Villeurbanne, France

²INRS-ETE, Québec, Canada

marion.moussay@inrae.fr, andre.st-hilaire@inrs.ca, florentina.moatar@inrae.fr

RÉSUMÉ

Le réchauffement de la température des cours d'eau est un fait indéniable du fait du changement climatique et des activités humaines ¹. Cependant, en France, nous ne disposons pas encore de réseau national de suivi en continu de la température des cours d'eau, représentatif de l'ensemble des conditions et pérenne afin de pouvoir quantifier ce réchauffement sur l'ensemble du réseau hydrographique et pour un grand nombre de métriques compatibles avec les exigences des communautés biologiques ². Suite aux différentes sécheresses et au manque d'information sur leurs conséquences sur les écosystèmes aquatiques, les autorités publiques s'intéressent de plus en plus aux suivis de température de l'eau afin de mieux gérer les ressources en eau. Dans ce travail nous allons montrer une méthodologie nouvelle permettant de reconstituer à partir des bases de données hétérogènes, des chroniques, des métriques et des régimes de températures sur le territoire français. Une attention particulière sera portée d'une part aux stations de suivis piscicoles ³, ce qui permettra une analyse de l'évolution des tendances sur les 40 dernières années des communautés en lien avec les conditions thermiques des rivières et l'utilisation d'indicateurs d'habitat directs dans les modèles de distribution d'espèces ⁴. D'autre part, la caractérisation des régimes (naturels et altérés) pourrait aboutir à la construction d'un indicateur de qualité thermique. Le portail thermie-rivières.inrae.fr de valorisation des données thermie porté par INRAE sera ainsi enrichi avec des nouveaux indicateurs et nouvelles fonctionnalités, suite à ce travail.

ABSTRACT

The warming of river temperatures is an undeniable fact due to climate change and human activities¹. However, in France, we do not yet have a national network for continuous monitoring of river temperature, representative of all conditions in order to be able to quantify this warming over the entire hydrographic network and for a large number of metrics compatible with biological studies². Following various droughts and the lack of information on their consequences on aquatic ecosystems, public authorities are becoming increasingly interested in water temperature monitoring in order to better manage water resources. In this work, we will demonstrate a novel methodology for reconstructing temperature time-series, metrics and regimes across France, based on heterogeneous databases. Particular attention will be paid to fish monitoring stations³, enabling analysis of community trends over the last 40 years in relation to river thermal conditions and the use of direct habitat indicators in species distribution models ⁴. On the other hand, characterization of regimes (natural and altered) could lead to the construction of a thermal quality indicator. The INRAE website "thermie-rivières.inrae.fr" will be enhanced with new indicators and new functionalities as a result of this work.

KEYWORDS

Anomalies detection (détection des anomalies), deep learning (apprentissage profond), heterogeneous data (données hétérogènes), thermal biological quality index (indice biologique de qualité thermique), thermal regimes (régimes thermiques), streamwater temperature forecasting (prévision de la température des cours d'eau)

¹ Florentina Moatar et Joël Gailhard., *Comptes Rendus. Géoscience* 338, n° 5 (2006): 319-28. <https://doi.org/10.1016/j.crte.2006.02.011>; Hanieh Seyedhashemi et al., *Science of The Total Environment* 766 (20 avril 2021): 142667. <https://doi.org/10.1016/j.scitotenv.2020.142667>.

² Aurélien Beaufort et al., *Hydrology and Earth System Sciences* 26, n° 13 (7 juillet 2022): 3477-95. <https://doi.org/10.5194/hess-26-3477-2022>.

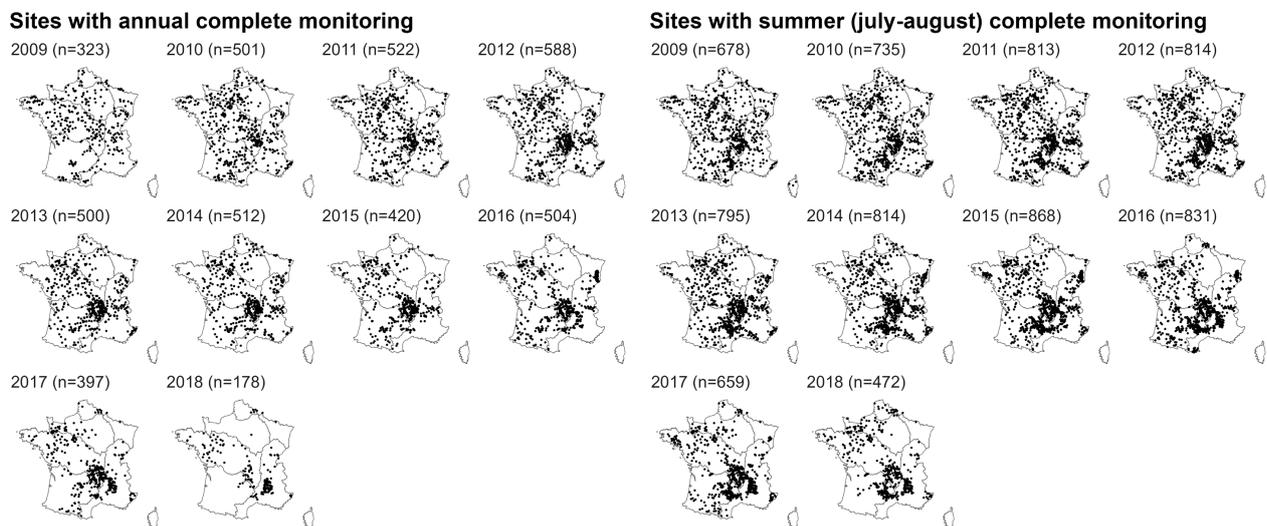
³ Pascal Irz et al., *Knowledge & Management of Aquatic Ecosystems*, n° 423 (2022): 25. <https://doi.org/10.1051/kmae/2022021>.

⁴ Coline Picard et al., *PLOS ONE* 17, n° 9 (22 septembre 2022) : e0274167. <https://doi.org/10.1371/journal.pone.0274167>.

1 OBJECTIFS ET METHODOLOGIE

1.1 Contexte

En 2018, le projet TIGRE a été à l'initiative d'une collecte importante de chroniques horaires ou journalières de températures de l'eau auprès de différents acteurs, avec comme objectif d'analyser la variabilité spatiale des régimes thermiques des cours d'eau à l'échelle nationale et d'extrapoler deux métriques de température estivale (moyennes des 30 et 7 jours consécutifs les plus chauds, Tw30J, Tw7j) le long des réseaux hydrographiques⁵. Cette collecte a mis en évidence l'hétérogénéité des données à l'échelle nationale, dépendante des objectifs de chaque producteur : 1) Stations du Réseau National Thermique mis en place par l'OFB en 2008 (environ 600 stations comportant entre une année et plusieurs années) et DREAL (181 stations) pour connaître l'évolution de la thermie des rivières sous changement climatique ; 2) Stations de suivi réglementaires d'EDF au niveau des centrales de production d'électricité sur les fleuves français (38 stations) ; 3) Stations de suivi en période estivale des Fédérations de Pêche pour caractérisation de l'habitat piscicole (1815 stations) ; 4) Stations des syndicats d'eau, EPTBs, Parcs régionaux ou nationaux pour la gestion des milieux ; 5) Organismes de recherche pour des questions de recherche spécifiques. Cette collecte a permis de constituer une base de données avec 2671 stations dont 1267 avec au moins une année complète de suivi et 2347 avec au moins un été complet⁵.



1.2 Objectifs et méthodologie

Dans ce travail nous réalisons une réactualisation de ces données, permettant ainsi d'intégrer les années de sécheresse récentes (2019, 2020, 2022), ainsi qu'un historique plus important pour quelques stations ayant des chroniques longues (plus de 20 ans d'observations). Dans la première phase du projet TIGRE, les modèles de reconstitution des données journalières ont été relativement simples, prenant en compte seulement la température de l'air et l'incertitude n'a pas été quantifiée. Dans ce travail, nous allons intégrer les débits (modélisés ou mesurés à des stations proches) et autres variables climatiques ou hydrogéologiques permettant ainsi une meilleure analyse des associations débits-températures. Enfin environ 1000 stations ont été écartés de l'analyse du projet TIGRE, car les stations ont été qualifiées comme influencées ou sans données suffisantes. L'analyse des données aux stations influencées permettrait de quantifier ces influences par rapport à des régimes thermiques naturels. Une attention particulière sera portée d'une part aux stations de suivis piscicoles⁶, ce qui permettra une analyse de l'évolution des tendances sur les 40 dernières années des communautés en lien avec les conditions thermiques des rivières et l'utilisation d'indicateurs d'habitat directs dans les modèles de

⁵ Aurélien Beaufort et al., *Hydrological Processes* 34 (2020). <https://doi.org/10.1002/hyp.13608>; Aurélien Beaufort et al., *Hydrology and Earth System Sciences* 26, no 13 (7 juillet 2022) : 3477-95. <https://doi.org/10.5194/hess-26-3477-2022>.

⁶ Pascal Irz et al., *Knowledge & Management of Aquatic Ecosystems*, no 423 (2022) : 25. <https://doi.org/10.1051/kmae/2022021>

distribution d'espèces ⁷. D'autre part, la caractérisation des régimes (naturels et altérés) pourrait aboutir à la construction d'un indicateur de qualité thermique. Le portail *thermie-rivieres.inrae.fr* de valorisation des données thermie porté par INRAE sera ainsi enrichi avec des nouveaux indicateurs et nouvelles fonctionnalités, suite à ce travail. La méthodologie d'analyse et de prédiction de la variabilité spatio-temporelle sera donc adaptée au contexte lacunaire et hétérogène des données. Ainsi nous considérons plusieurs types de modèles pour la reconstitution des chroniques journalières de température, la caractérisation des régimes et la simulation du futur en intégrant autant que possible des variables hydro(géo)logiques: 1) développement de modèles deep-learning (LSTM), pour les stations disposant de chroniques à long terme ou pour des régions disposant d'une densité importante de stations ; 2) développement des modèles empiriques en site suivi mais de façon lacunaire (données disponibles au moins une année ou un été) et non-suivi (par exemple stations piscicoles), en considérant pour certaines stations, un calage ou une validation sur des données basse-fréquence (mensuelles). Enfin, la caractérisation des régimes annuels et la spatialisation des paramètres des régimes pourra permettre une analyse des stations influencées.

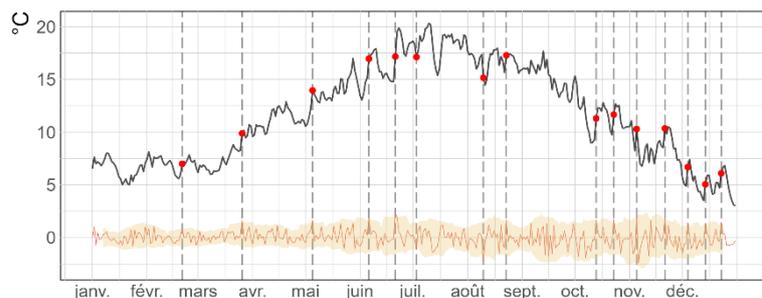
Pour spatialiser les régimes et métriques biologiques, deux méthodes principales sont testées. Une première est un modèle basé sur une fonction gaussienne des températures moyennes journalières, caractérisant synthétiquement la saisonnalité thermique à travers des paramètres écologiquement pertinents (amplitude, date, durée) qui ont été utilisés pour dresser une première cartographie des régimes thermiques des cours d'eau⁸. Les fonctions gaussiennes permettent une estimation explicite des incertitudes liées à la variabilité saisonnière et pourraient être exploitées à la longueur des séries temporelles, ce qui rend possible l'identification plus précisément des régions où les données manquent. Une seconde consiste au calibrage de modèles multi-métriques empirique qui permettent l'analyse des tendances interannuelles et offrent une base pour modéliser les impacts climatiques sur les régimes thermiques des rivières⁹. En intégrant les deux méthodes, il serait possible de développer une approche hybride où les régimes thermiques définis par les fonctions gaussiennes serviraient de base pour calibrer des modèles basés sur l'approche multi-métriques empiriques. La simplicité et la robustesse du modèle gaussien serait couplé avec la capacité des modèles multi-métriques empiriques qui permettent d'intégrer d'autres métriques d'intérêt écologiques, notamment pour caractériser les variabilités à court-terme.

2 RESULTATS

2.1 Critique des données

La méthode de détection d'anomalies repose sur les modèles ARIMA, adaptés aux données environnementales pour capturer tendances saisonnières et variations à court terme ¹⁰. Les résidus, calculés comme la différence entre valeurs observées et prédites, servent à identifier les anomalies via une analyse sur une fenêtre glissante définissant des seuils dynamiques ¹¹.

L'approche dynamique permet d'identifier des anomalies contextuelles, mais elle exige un ajustement précis des paramètres clés, notamment la taille de la fenêtre, le niveau de confiance statistique, et le seuil minimal des écarts considérés comme significatifs. Ces trois paramètres influencent directement la sensibilité et la spécificité du



⁷ Coline Picard et al., PLOS ONE 17, no 9 (22 septembre 2022) : e0274167. <https://doi.org/10.1371/journal.pone.0274167>.

⁸ Anik Daigle, Claudine Boyer, et André St-Hilaire., *Journal of Hydrology* 577 (1 octobre 2019) : 123963. <https://doi.org/10.1016/j.jhydrol.2019.123963>.

⁹ Aurélien Beaufort et al., *Hydrology and Earth System Sciences* 26, no 13 (7 juillet 2022) : 3477-95. <https://doi.org/10.5194/hess-26-3477-2022>.

¹⁰ F. Moatar, J. Miquel, et A. Poirel., *Journal of Hydrology* 252, n° 1 (31 octobre 2001) : 25-36. [https://doi.org/10.1016/S0022-1694\(01\)00439-5](https://doi.org/10.1016/S0022-1694(01)00439-5).

¹¹ Catherine Leigh et al., *Science of The Total Environment* 664 (10 mai 2019) : 885-98. <https://doi.org/10.1016/j.scitotenv.2019.02.085>; Amber Spackman Jones, Tanner Lex Jones, et Jeffery S. Horsburgh., *Environmental Modelling & Software* 151 (1 mai 2022) : 105364. <https://doi.org/10.1016/j.envsoft.2022.105364>.

processus de détection et doivent être adaptés aux caractéristiques propres des données. Il n'existe pas de valeur guide universelle pour ces paramètres, et leur choix doit être guidé par une inspection visuelle des résultats.

2.2 Modèles deep-learning pour les stations avec données disponibles à long-terme

Les réseaux LSTM, conçus pour capturer des dépendances temporelles complexes, sont particulièrement adaptés à la prévision des séries temporelles hydrologiques en raison de leur capacité à modéliser des dépendances temporelles complexes dans les données. L'architecture des LSTM comprend des cellules de mémoire qui peuvent maintenir des informations sur de longues périodes, ce qui est essentiel pour modéliser les dynamiques temporelles de la température de l'eau des cours d'eau, influencées par divers facteurs climatiques et environnementaux. Pour cette raison, ils nécessitent des données sur plusieurs décennies pour refléter les dynamiques temporelles.

Qiu et al., (2021) ont démontré l'efficacité des LSTM en les comparant à des modèles air2stream (équations de bilan thermique simplifié), forêts aléatoires (RF) et réseaux neuronaux à rétropropagation (BPNN), à partir de chroniques annuelles de 30 ans pour des stations en Suisse, aux États-Unis et en Chine. Les LSTM ont surpassé les autres modèles avec des RMSE inférieurs à 1,0 °C et des NSE supérieurs à 0,9, tandis que les autres méthodes affichaient des performances inférieures. L'ajout de nouvelles variables explicatives améliore encore la performance des LSTM. Par exemple, il a été montré que l'inclusion du débit dans les modèles réduisait le RMSE médian de 0,17 °C, soulignant l'importance de prendre en compte des facteurs hydrologiques pour affiner les prévisions¹². Enfin, la réflexion sur l'approche à adopter – locale ou régionale – constitue un enjeu central pour l'application des LSTM dans la modélisation hydrologique. Une approche régionale, en intégrant des données de plusieurs bassins, permet aux LSTM de généraliser efficacement tout en réduisant le besoin de calibrer chaque station individuellement, ce qui est particulièrement utile pour des réseaux hydrologiques complexes¹³. Des données d'entraînement diversifiées, issues de différentes régions, renforcent la robustesse des prédictions dans des contextes variés¹⁴. Cependant, les approches locales conservent leur pertinence où des modèles entraînés individuellement pour chaque bassin se révèlent plus performants pour capturer les particularités locales, notamment dans des bassins atypiques ou sous-représentés¹⁵.

Dans ce travail, nous allons tester ces approches sur une quarantaine de stations disponibles en France, ayant des chroniques sur plus d'une dizaine d'années (stations DREAL, EDF, et certaines stations du RNT). Les premiers essais ont montré des performances améliorées notamment en incorporant le débit comme variable explicative par rapport à des méthodes traditionnelles utilisées dans la première phase du projet TIGRE.

3 MISE A DISPOSITION DES INDICATEURS MODELISES – SITE WEB THERMIERIVIERE.INRAE

Le portail thermie-rivieres.inrae.fr de valorisation des données thermie porté par l'INRAE sera enrichi avec de nouvelles métriques et fonctionnalités (exemple : scripts R pour la détection des valeurs aberrantes, pour la caractérisation des régimes et l'extrapolation des chroniques journalières). Il permettra la mise à jour régulière des indicateurs pour certaines stations (OFB, DREAL, autres), la réalisation des analyses en direct (notamment les tendances pour les chroniques avec plus de 10 ans de suivi), le téléchargement de chroniques de données spatiales élaborées (par exemple MTW30J) sur l'ensemble des stations, à l'échelle d'une région ou d'un département ou d'un SAGE possibilité de télécharger les données élaborées sur la base de critères géographiques et temporels allant de l'année et la station à l'intégralité du territoire et de la période d'étude), la proposition de statistiques descriptives réactualisées à partir des chroniques journalières mesurées et reconstituées au droit de stations et le calcul des dépassements des seuils de tolérance d'une vingtaine d'espèces piscicoles.

L'objectif de cette communication par poster est aussi d'interroger la communauté des acteurs présents à IS Rivers sur les besoins en terme de nouvelles fonctionnalités.

Remerciements : Ce travail est financé par l'OFB dans le cadre d'une coopération OFB-INRAE.

¹² Farshid Rahmani et al., *Environmental Research Letters* 16, n° 2 (janvier 2021) : 024025. <https://doi.org/10.1088/1748-9326/abd501>.

¹³ Frederik Kratzert et al., *Hydrology and Earth System Sciences* 23, n° 12 (17 décembre 2019) : 5089-5110. <https://doi.org/10.5194/hess-23-5089-2019>.

¹⁴ Kuai Fang et al., *Water Resources Research* 56 (2020). <https://doi.org/10.1029/2020WR028095>.

¹⁵ Wei Zhi et al., *Nature Climate Change* 13, n° 10 (octobre 2023) : 1105-13. <https://doi.org/10.1038/s41558-023-01793-3>.