UQAC
Université du Québec
à Chicoutimi

Modélisation de la période de reproduction du doré jaune (*Sander vitreus)* dans un contexte de changements climatiques au Québec

par Anna-Ève Lafrenière

Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de maîtrise ès sciences (M. Sc.) en ressources renouvelables

Québec, Canada

RÉSUMÉ

La fraie du doré jaune (*Sander vitreus*) constitue une étape déterminante de son cycle de vie, durant laquelle les géniteurs se rassemblent en fortes densités sur des sites de reproduction restreints, généralement situés dans des secteurs facilement accessibles par les pêcheurs. Cette concentration les rend particulièrement vulnérables au prélèvement par la pêche sportive. Au Québec, les facteurs environnementaux qui régissent le moment de ces rassemblements demeurent méconnus, ce qui complexifie la mise en place de mesures de protection adaptées à cette période critique. Cette situation est d'autant plus préoccupante dans un contexte de changements climatiques, susceptibles de modifier les conditions environnementales et, par conséquent, la dynamique reproductive du doré jaune. Il est donc essentiel de mieux comprendre les facteurs environnementaux qui régissent les rassemblements durant la pré-fraie et la fraie pour améliorer la gestion de l'espèce. Cette étude vise à identifier et quantifier l'influence des variables environnementales modulant l'initiation et la cessation de la fraie du doré jaune au Québec, ainsi qu'à développer un modèle prédictif des dates de début et de fin de la reproduction selon les conditions environnementales actuelles et futures. Pour ce faire, 484 observations de géniteurs sur les sites de fraie, issues de la littérature scientifique publiée et de la littérature grise, ont été combinées à des variables environnementales actuelles et à des projections climatiques futures. L'utilisation d'algorithmes d'apprentissage automatique, notamment les forêts d'arbres de décision, a permis de mettre en évidence le rôle prépondérant de la température dans la régulation de la fraie. L'accumulation thermique printanière, représentée par les degrés-jours de croissance au-dessus de 5 °C (DJC5) cumulés au 15 avril, apparaît comme la principale variable prédisant les dates de début et de fin de la reproduction. Les autres variables environnementales, soit la photopériode (estimée par la latitude), l'élévation et les précipitations printanières, présentent une contribution plus limitée et agissent principalement comme des modulateurs des conditions thermiques. Dans un contexte de changements climatiques, un devancement généralisé de la période de fraie à l'horizon 2071-2100, de l'ordre de 7 à 12 jours sous les scénarios SSP2-4.5 et SSP5-8.5, est projeté, bien que l'ampleur de ce déplacement varie en fonction de la latitude. Les résultats montrent également que ces changements ne se limitent pas à un simple décalage des dates de fraie, mais s'accompagnent de modifications de la durée de la période de reproduction, qui tend à s'allonger dans les régions méridionales et à se contracter dans les régions nordiques. Dans l'ensemble, cette étude confirme le rôle central de la température dans l'initiation et la cessation de la fenêtre reproductive du doré jaune et fournit des outils permettant de bonifier la gestion de l'espèce au Québec, afin de mieux l'adapter aux réalités locales des populations et aux conditions climatiques futures.

ABSTRACT

Spawning in walleye (*Sander vitreus*) represents a critical stage of its life cycle, during which spawners aggregate at high densities on restricted and often easily accessible spawning sites. This aggregation makes them particularly vulnerable to harvest by recreational fishing. In Quebec, the environmental factors governing the timing of these aggregations remain poorly understood, complicating the implementation of protection measures adapted to this critical period. This issue is further exacerbated in the context of climate change, which may alter environmental conditions and, consequently, the reproductive dynamics of walleye. In this context, a thorough understanding of the environmental drivers of spawning is essential to improve species management. This study aimed to identify and quantify the influence of environmental variables regulating spawning initiation and termination in walleye in Quebec, and to develop predictive models of spawning onset and end dates under current and future environmental conditions. To this end, 484 observations of spawners at spawning sites, derived from both peer-reviewed and grey literature, were combined with current environmental variables and future climate projections. The use of machine learning algorithms, particularly Random Forest, highlighted the predominant role of temperature in regulating spawning dynamics. Spring thermal accumulation, represented by growing degree days above 5 °C (DJC5) accumulated as of April 15, emerged as the main variable allowing prediction of spawning onset and termination dates. Other environmental variables, including photoperiod (estimated by latitude), elevation, and spring precipitation, showed more limited contributions and primarily acted as modulators of thermal conditions. Under future climate scenarios, a generalized advancement of the spawning period is projected for 2071-2100, ranging from approximately 7 to 12 days under SSP2-4.5 and SSP5-8.5, although the magnitude of this shift varies with latitude. Results also indicate that these changes are not limited to a simple shift in spawning dates but are accompanied by changes in spawning duration, which tends to increase in southern regions and decrease in northern regions. Overall, this study confirms the central role of temperature in the initiation and termination of the reproductive window of walleye and provides tools to improve species management in Quebec, allowing for better adaptation to local population dynamics and future climatic conditions.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

%IncMSE	Per cent Increase in Mean Squared Error
ALE	Accumulated local effects
CanDCS-M6	Canadian Downscaled Climate Scenarios-Multivariate dataset
CMIP5	Coupled Model Intercomparison Project Phase 5
CMIP6	Coupled Model Intercomparison Project Phase 6
COSEWIC	Committee on the Status of Endangered Wildlife in Canada
CREAE	Chaire de recherche sur les espèces aquatiques exploitées
DJC5	Degrés-jours de croissance au-dessus de 5°C
FAO	Food and Agriculture Organization
FSC	Food, Social and Ceremonial
GAMM	Generalized Additive Mixed Models
GBM	Gradient Boosting Machine
GDD5	Growing Degree Days above 5 °C
IncNodePurity	Increase in Node Purity
LightGBM	Light Gradient Boosting Machine
MELCCFP	Ministère de l'Environnement, de la Lutte contre les Changements Climatiques, de la Faune et des Parcs
mm	Millimeters
MSE	Mean squared error
PDPs	partial dependence plots
RCP8.5	Representative Concentration Pathway 8.5
SSP2-4.5	Shared Socioeconomic Pathways 2-4.5
SSP5-8.5	Shared Socioeconomic Pathways 5-8.5
UQAC	Université du Québec à Chicoutimi
XGBoost	EXtreme Gradient Boosting

REMERCIEMENTS

Au-delà de la maîtrise, ces remerciements sont attribués aux personnes ayant exercé une influence importante dans l'ensemble de mon parcours universitaire. Mes remerciements vont d'abord à mon directeur de maîtrise, Olivier Morissette. Je te remercie d'avoir été mon mentor tout au long de mon parcours universitaire et de m'avoir fait bénéficier de ton expertise et de tes conseils. Sans ton soutien, je n'aurais pas eu la chance de vivre toutes ces opportunités enrichissantes, tant académiques que professionnelles, qui ont fait de moi la biologiste que je suis. À vrai dire, ton arrivée à la CREAE a été un cadeau inestimable et je me considère extrêmement chanceuse pour cela.

Merci à mon codirecteur, Pascal Sirois, et aux membres du personnel du laboratoire de la CREAE, particulièrement à Anne-Lise et Sonya. Vos conseils avisés m'ont permis de m'améliorer, en plus de rendre les présentations moins intimidantes.

Merci à Yves Paradis et à Geneviève Ouellet-Cauchon (MELCCFP) pour votre dévouement et votre disponibilité. Vous avez grandement contribué à la réussite de mon projet grâce à nos échanges formateurs, à votre implication et à la générosité de votre temps. J'en profite également pour remercier Kim Ménard (MELCCFP), qui a réalisé un travail colossal en triant et en identifiant les informations pertinentes à mon jeu de données. Tu m'as permis de me plonger rapidement au cœur du sujet et j'en suis très reconnaissante. Travailler avec vous fut un réel plaisir.

J'aimerais remercier l'équipe de la Direction de la gestion de la faune de l'Outaouais, Carolane, Tyler, Ève et Vincent, de m'avoir accueillie sur le terrain. Ces deux semaines ont été très enrichissantes et formatrices, et j'en conserve de très bons souvenirs.

À l'approche de la fin de mes études et de ce tournant de ma vie d'étudiante, je tenais à remercier ma collègue et second mentor, Karine Gagnon. Merci de ta confiance, de tes conseils et de ton précieux temps. Tu m'as permis de participer à une multitude de projets qui ont bonifié ma vision de la gestion de la faune, en plus d'avoir été un modèle de femme ambitieuse et déterminée, auquel j'aspire.

Évidemment, ces remerciements seraient incomplets sans la mention de toutes les personnes ayant cru en moi et m'ayant encouragée tout au long de mon parcours académique. Un grand merci à ma précieuse Maude. Grâce à toi, mon parcours académique aura été rempli de fous rires, d'échanges enrichissants et de souvenirs mémorables. Merci à ma famille et mes amis de m'avoir accompagnée et appuyée toutes ces années. Votre fierté à mon égard m'a donné la force de me dépasser et la confiance nécessaire pour avancer. Vous avez été un pilier essentiel tout au long de ce parcours, et je vous en serai toujours reconnaissante.

Je conclus ces remerciements en exprimant ma gratitude envers la Fondation Héritage Faune pour le soutien offert par l'entremise de la bourse d'études Pierre-Latraverse, ainsi que pour la visibilité que vous avez accordée à mon projet.

AVANT-PROPOS

Ce mémoire a été réalisé dans le cadre de la maîtrise en ressources renouvelables de l'Université du Québec à Chicoutimi et a été rédigé selon la structure d'un mémoire sous forme d'article scientifique en anglais. L'introduction et la conclusion générale sont présentées en français et offrent une mise en contexte exhaustive de l'étude, de sa portée et de ses retombées, ainsi que ses limitations. Le chapitre intitulé *Modeling the spawning phenology of walleye (Sander vitreus) across Québec* est, quant à lui, rédigé en anglais. La méthodologie, ainsi que les résultats et leur interprétation y sont présentés.

Ce travail a été exécuté sous la direction d'Olivier Morissette, cotitulaire de la Chaire de recherche sur les espèces aquatiques exploitées (CREAE) et la codirection de Pascal Sirois, directeur de la CREAE. Il n'aurait toutefois pas été possible sans la participation active de Yves Paradis (MELCCFP) et de Geneviève Ouellet-Cauchon (MELCCFP) qui ont supervisé et rendu accessibles les données utilisées dans cette étude, lesquelles ont été préalablement compilées par Kim Ménard (MELCCFP).

1. INTRODUCTION GÉNÉRALE

1.1 Mise en contexte

L'exploitation des ressources halieutiques constitue une composante importante de l'histoire évolutive et de la survie des populations humaines. Initialement pratiquée à des fins de subsistance, la pêche est progressivement devenue un vecteur d'échanges et de commerce dans plusieurs régions du monde, contribuant à la dispersion humaine, à la structuration des économies locales et, dans certains contextes, à l'établissement des premières sociétés complexes (Sahrhage et Lundbeck 1992; Pitcher et Lam 2014). Encore aujourd'hui, les pêches sportives et commerciales représentent une activité sociale, culturelle et économique majeure à l'échelle mondiale (FAO 2024). Au Canada, les pêches commerciales ciblant les poissons, tant dulcicoles que marins, ont généré près de 591 millions de dollars en valeur brute des débarquements en 2023 et soutenu près de 65 000 emplois, jouant ainsi un rôle important dans l'économie du pays (Pêches et Océans Canada 2025).

Dans les eaux canadiennes, on dénombre plus d'un millier d'espèces de poissons, dont 11 % se trouvent exclusivement en eau douce et 4 % fréquentent à la fois les milieux dulcicoles et marins (Conseil canadien pour la conservation des espèces en péril 2006). Le Québec, deuxième province présentant la plus grande diversité d'espèces de poissons, en compte à lui seul près de 120, dont une vingtaine sont principalement ciblées par la pêche sportive (Conseil canadien pour la conservation des espèces en péril 2006; Gouvernement du Québec 2026). Parmi celles-ci, le doré jaune (*Sander vitreus*) constitue l'espèce la plus recherchée par les pêcheurs canadiens et la seconde espèce la plus prisée au Québec après l'omble de fontaine (*Salvelinus fontinalis*) (Pêches et Océans Canada 2019). Il occupe également une place importante au sein de plusieurs communautés autochtones du Québec,

pour lesquelles il représente une ressource de subsistance et revêt une valeur culturelle et sociale cruciale (Pêches et Océans Canada 2019).

Au début des années 2000, le déclin de la qualité de la pêche et du potentiel reproducteur des populations québécoises de doré jaune, tant en eaux intérieures que dans le fleuve Saint-Laurent, a été attribué à une pression de pêche excessive exercée sur l'espèce. Les inventaires normalisés réalisés par le ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP) entre 1985 et 2008 ont notamment révélé une diminution de 21 % de la masse moyenne des dorés et une surexploitation de l'espèce dans 30 % des plans d'eau échantillonnés. Une situation comparable a été observée dans le fleuve Saint-Laurent, où le doré jaune a été considéré « à risque » ou « en déclin » dans la plupart des secteurs inventoriés (Arvisais *et al.* 2012). Ce constat a conduit à l'instauration, en 2011, de nouvelles modalités de gestion dans le cadre du *Plan de gestion du doré 2011-2016*, reconduit jusqu'en 2026, afin de rétablir la structure des populations et l'abondance des reproducteurs (Arvisais *et al.* 2012; Arvisais *et al.* 2016).

À l'échelle provinciale, les retombées de ces mesures sur le rétablissement des populations se sont révélées variables. Dans certains plans d'eau, les inventaires du MELCCFP ont mis en évidence une amélioration de la qualité de la pêche et une augmentation de l'abondance des reproducteurs (Arvisais *et al.* 2016; Nadon 2022b; Bélanger 2024), alors que dans d'autres, ces modalités n'ont entraîné aucune amélioration notable de l'état global des populations (Bélanger 2022; Nadon 2022a; Bélanger 2023). Dans le fleuve Saint-Laurent, l'instauration d'une gamme de tailles exploitées (381-545 mm) n'a pas permis d'enrayer le déclin de l'abondance des grands dorés jaunes (> 545 mm), ni la diminution de la taille des géniteurs et l'augmentation de la mortalité annuelle totale observées depuis 2001. La croissance, la condition corporelle et la taille à maturité des

femelles ont également continué de diminuer depuis la mise en place de cette réglementation (Mainguy *et al.* 2025).

Outre l'exploitation par la pêche, le recrutement du doré jaune peut varier considérablement sous l'effet de facteurs abiotiques et biotiques influençant notamment la qualité de son habitat (Mainguy *et al.* 2025). Le doré jaune occupe un vaste continuum environnemental, mais performe généralement mieux dans les milieux mésotrophes frais et bien oxygénés (Bozek *et al.* 2011a). Sa vision scotopique le rend particulièrement sensible à la lumière, ce qui explique ses meilleures performances dans les lacs présentant une plus faible transparence de l'eau, correspondant à des profondeurs de disque de Secchi de 1 à 3 m (Lester *et al.* 2004). Des changements de turbidité et de température peuvent donc induire une modification de l'utilisation des habitats vers des milieux moins propices sur les plans alimentaire ou biotique (Lester *et al.* 2004; Hansen *et al.* 2019; Hansen *et al.* 2022; Mahlum *et al.* 2023). L'acidification des plans d'eau constitue un autre facteur susceptible d'affecter sa démographie, en entraînant une mortalité accrue des œufs et des larves ou une cessation de l'activité de fraie à un pH inférieur à 6 (Beamish 1976; Smith *et al.* 2022). Enfin, la diminution de l'abondance des espèces proies (Hansen *et al.* 1998; Smith *et al.* 2022) ainsi que la présence d'espèces envahissantes ont été associées à une réduction du recrutement et de l'abondance de plusieurs populations de doré jaune (Hansen *et al.* 2020; Hansen *et al.* 2022)

La survie des premiers stades de vie est également influencée par les activités anthropiques. L'utilisation des terres riveraines et des secteurs terrestres environnants à des fins agricoles ou forestières peut notamment accroître la sédimentation dans les plans d'eau, réduisant la disponibilité des habitats de fraie optimaux ou provoquant l'anoxie des œufs par leur ensevelissement sous des sédiments fins (Nate *et al.* 2003; Raabe *et al.* 2020). Cette

survie est aussi affectée par les fluctuations de température et de précipitations associées aux changements climatiques (Hansen *et al.* 2019). Les asynchronies trophiques, l'allongement de la période d'incubation augmentant les risques de prédation, l'exondation des œufs et le transport des larves vers des habitats sous-optimaux constituent quelques exemples des effets potentiels de ces changements (Hansen *et al.* 1998; Zhao *et al.* 2009; Barta *et al.* 2024). Ces multiples pressions abiotiques et biotiques soulignent l'importance d'instaurer une gestion étroite de l'espèce, dont le déclin pourrait compromettre l'équilibre des écosystèmes dulcicoles (Colby *et al.* 1987; Thao *et al.* 2016).

À titre de prédateur primaire, le doré jaune exerce un contrôle descendant sur les niveaux trophiques inférieurs, contribuant à réguler leur abondance et à maintenir l'équilibre des communautés ichtyennes (Lyons et Magnuson 1987). Lorsque ses populations déclinent, le relâchement de cette régulation peut favoriser l'augmentation d'espèces fourragères ou compétitrices et modifier la dynamique du réseau trophique (Colby *et al.* 1987; Thao *et al.* 2016). Par exemple, une augmentation de l'abondance de la perchaude a été observée à la suite du déclin du doré jaune dans l'étude de Colby *et al.* (1987) et de Thao *et al.* (2016). Bien que cette espèce constitue souvent une proie importante, une forte abondance de perchaudes adultes peut accroître la prédation sur le doré, notamment sur les œufs et les individus d'âge-0, en plus d'occasionner une compétition alimentaire avec les juvéniles, instaurant une rétroaction écologique défavorable pour le doré (Colby *et al.* 1987; Hansen *et al.* 1998; Roseman *et al.* 2006). Des déséquilibres dans les ratios prédateurs-proies favorisant certaines espèces compétitrices, notamment les centrarchidés (Quist *et al.* 2003; Embke *et al.* 2022) ou l'alose à gésier (*Dorosoma cepedianum*) (Quist *et al.* 2004), ont également été associés à des effets négatifs sur le recrutement du doré jaune. Le déclin de l'espèce ne représente donc pas uniquement une perte halieutique ou économique, mais peut entraîner

des modifications durables du fonctionnement des écosystèmes. Dans ce contexte, une gestion intégrée axée sur les facteurs de pression contrôlables, comme l'exploitation par la pêche, apparaît essentielle pour limiter les risques de déclins des populations (Post 2013).

Dans la province, la réglementation de la pêche sportive et commerciale varie selon des zones de pêche (n = 29) pouvant couvrir de vastes territoires (Arvisais *et al.* 2012). Ainsi, elle ne reflète pas toujours les réalités locales des différentes populations, ce qui peut entraîner un chevauchement entre l'ouverture de la pêche et la période de fraie du doré jaune (Gendron 2009; Paquin et Brodeur 2023). La protection des géniteurs, notamment par le déplacement de la saison de pêche en dehors de la période de fraie, constitue une modalité de gestion répandue. Son efficacité demeure néanmoins débattue au sein de la communauté scientifique, en raison de l'intensification de l'effort de pêche qu'elle peut entraîner sur une période plus restreinte (Heino 1998; Arvisais *et al.* 2012; Overzee et Rijnsdorp 2014). Elle permet toutefois de réduire les risques de surexploitation chez les espèces formant de grandes agrégations de fraie, en diminuant la mortalité par la pêche des individus plus âgés, dont la contribution reproductive est généralement plus élevée (Heino 1998; Venturelli *et al.* 2010; Overzee et Rijnsdorp 2014). Chez le doré jaune, la fraie se déroule en zones peu profondes, souvent à moins d'un mètre, où les individus se rassemblent en denses bancs (Gendron 2009; Paquin et Brodeur 2023). Durant cette période, ils deviennent facilement repérables et particulièrement vulnérables à la récolte, ce qui peut entraîner une mortalité importante du stock reproducteur (Gendron 2009; Paquin et Brodeur 2023; Alglave *et al.* 2024). Bien que ce phénomène soit documenté, les facteurs environnementaux déterminant la période de rassemblement des adultes sur les sites de reproduction sont méconnus au Québec, ce qui limite la capacité d'appuyer scientifiquement la modification des dates d'ouverture de la pêche (Dixon *et al.* 2020).

1.2 Calendrier de la reproduction

Dans la littérature, le cycle reproducteur du doré jaune est étroitement lié à la température et à la photopériode (Malison et Held 1996). Il s'amorce dès le début de l'automne, lorsque la diminution de la photopériode initie la gamétogenèse. Ce processus se poursuit durant l'hiver et requiert une période minimale de trois à cinq mois pendant laquelle la température de l'eau doit demeurer inférieure à 10 °C afin de permettre le développement optimal des gamètes (Hokanson 1977; Malison *et al.* 1994; Malison et Held 1996; Feiner et Höök 2015; Firkus *et al.* 2024). Chez les mâles, la gamétogenèse est généralement complétée vers le milieu de l'hiver, tandis que chez les femelles, elle se poursuit jusqu'à l'approche de la fraie (Malison *et al.* 1994; Malison et Held 1996).

La fraie survient dès les premiers réchauffements printaniers des eaux et varie selon la latitude et les caractéristiques locales des plans d'eau (Bozek *et al.* 2011a; Feiner et Höök 2015). À l'extrémité sud de son aire de répartition nord-américaine, elle peut débuter dès la fin janvier, alors qu'aux latitudes plus nordiques elle est observée jusqu'à la fin juin (Hokanson 1977). La température de l'eau est généralement reconnue comme le principal facteur déclenchant les migrations reproductives et le comportement de fraie (Schneider *et al.* 2010; Bozek *et al.* 2011a; Arvisais *et al.* 2012; Feiner et Höök 2015). Bien que les valeurs rapportées varient selon les études, la fraie survient le plus souvent à des températures de l'eau comprises entre 4 et 11 °C, soit des températures favorisant la viabilité des gamètes, la fécondité et l'incubation des œufs (Koenst et Smith 1976; Hokanson 1977; Schneider *et al.* 2010; Arvisais *et al.* 2012; Feiner et Höök 2015; Feiner *et al.* 2022). Généralement, la période de fraie s'étend sur une à deux semaines et prend fin lorsque la température de l'eau atteint ou dépasse 15 °C (Priegel 1970; Malison et Held 1996; Feiner *et al.* 2025).

De manière générale, les mâles sont les premiers à arriver et les derniers à quitter les sites de fraie (Bade *et al.* 2019; Elliott *et al.* 2022). Cette stratégie leur confère un avantage reproductif en leur permettant de fertiliser les œufs de plusieurs femelles sur une période prolongée (Bozek *et al.* 2011a; Bade *et al.* 2019; Elliott *et al.* 2022). À l'inverse, les femelles libèrent leurs œufs sur une période plus restreinte et quittent rapidement les frayères. Ce comportement a été attribué à l'investissement énergétique plus élevé qui est requis pour la production des gamètes chez les femelles, les incitant à rejoindre plus rapidement des zones d'alimentation plus productives (Wang *et al.* 2007; Hayden *et al.* 2014; Bade *et al.* 2019). Dans le lac Ontario, par exemple, des mâles ont été observés entrant en moyenne cinq jours plus tôt dans les rivières que les femelles et les quittaient en moyenne sept jours plus tard (Elliott *et al.* 2022).

Plusieurs auteurs soulignent l'importance d'une compréhension approfondie de la phénologie reproductive des espèces exploitées afin d'en assurer une gestion optimale et proactive, particulièrement dans un contexte de changements climatiques (Dixon *et al.* 2020; Feiner *et al.* 2022; Feiner *et al.* 2025; Mainguy *et al.* 2025). L'étude de la phénologie est d'ailleurs considérée comme « un des moyens les plus simples de suivre les changements écologiques des espèces en réponse aux changements climatiques » (Walther *et al.* 2002). Cependant, peu d'études ont documenté la phénologie reproductive du doré jaune malgré l'abondante littérature scientifique portant sur cette espèce. Les études existantes ont principalement été réalisées dans la portion méridionale de son aire de répartition, et, dans la plupart des cas, le pic de fraie a été associé à la date de retrait des glaces, un indicateur climatique corrélé à la température de l'air et, indirectement, à la température de l'eau (Schneider *et al.* 2010; Dixon *et al.* 2020; Feiner *et al.* 2022; Barta *et al.* 2024). Des études récentes ont toutefois mis en évidence des asynchronies entre ces métriques, suggérant que

des facteurs autres que la température pourraient contribuer à la régulation du calendrier de fraie (Barta *et al.* 2024). La photopériode a notamment été identifiée comme un déterminant majeur du pic de fraie au lac Escanaba entre 1957 et 2019 (Feiner *et al.* 2025), contrastant avec les travaux antérieurs qui attribuaient un rôle dominant au réchauffement de l'eau (Schneider *et al.* 2010; Bozek *et al.* 2011a).

Selon les projections climatiques issues du scénario RCP8.5 du CMIP5, les températures annuelles moyennes devraient augmenter de plus de 2 °C d'ici le milieu du siècle et dépasser 4 °C d'ici sa fin (Hansen *et al.* 2006; IPCC 2014; Poesch *et al.* 2016). Les régions de hautes latitudes, dont le Canada et le nord des États-Unis, devraient connaître un réchauffement plus marqué, avec une hausse des températures annuelles moyennes pouvant excéder 6 °C à la fin du siècle (Hansen *et al.* 2006; IPCC 2014; Poesch *et al.* 2016). Si la reproduction du doré jaune est fortement dépendante des signaux photopériodiques, elle pourrait alors survenir dans des conditions environnementales moins favorables, dans la mesure où celle-ci continuerait de diverger des signaux saisonniers associés au climat (Bozek *et al.* 2011a; Barta *et al.* 2024). Par exemple, le succès de fécondation et l'incubation des œufs sont fortement dépendants de la température de l'eau (Koenst et Smith 1976; Hokanson 1977). La plage thermique optimale pour l'incubation se situe entre 9 et 15 °C, avec un succès d'éclosion maximal autour de 15 °C (Koenst et Smith 1976; Hokanson 1977; Engel *et al.* 2000). Bien que les embryons tolèrent certaines fluctuations thermiques, une exposition prolongée à des températures supérieures à 19 °C a été associée à une forte mortalité embryonnaire (Koenst et Smith 1976; Schneider *et al.* 2002). Ces éléments mettent en évidence la nécessité d'améliorer notre compréhension de la phénologie reproductive de l'espèce afin de soutenir une gestion adaptée aux changements climatiques.

1.3 Méthode préconisée

Outre la hausse globale des températures prévue, les changements climatiques devraient accroître la variabilité interannuelle des températures, modifier les régimes de précipitations et augmenter la fréquence des événements météorologiques extrêmes (IPCC 2014). Une compréhension approfondie des facteurs environnementaux modulant la fraie du doré jaune apparaît donc essentielle à la gestion durable de l'espèce (Alglave *et al.* 2024). Dans ce contexte, l'apprentissage automatique constitue une approche de plus en plus utilisée en écologie pour analyser les interactions entre les espèces et leur environnement (Branco *et al.* 2023; Pichler et Hartig 2023). Comparativement aux méthodes statistiques traditionnelles, ces algorithmes présentent plusieurs avantages, notamment en raison de leur capacité à exploiter des jeux de données biologiques complexes comprenant un grand nombre de prédicteurs (Brownscombe *et al.* 2020; Pichler et Hartig 2023). Ils permettent de modéliser des relations non linéaires et des interactions de haute dimension, tout en intégrant efficacement des variables continues et catégorielles (Cutler *et al.* 2007; Thessen 2016). Certaines approches sont également robustes face aux données manquantes, une situation fréquente dans les jeux de données écologiques (Thessen 2016; Scowen *et al.* 2021). Les avancées technologiques récentes ont par ailleurs amélioré l'accessibilité de ces méthodes, si bien qu'elles sont désormais largement utilisées en modélisation spatiale ainsi qu'en gestion et conservation de la faune (Thessen 2016; Effrosynidis *et al.* 2020; Flores *et al.* 2024; Paquette *et al.* 2026).

Ces algorithmes peuvent ainsi être utilisés pour explorer la structure des données ou établir des relations entre des variables explicatives et une variable réponse afin de produire des prédictions pour de nouvelles observations (Bonaccorso 2018; Pandey *et al.* 2019). Ils reposent généralement sur une phase d'entraînement, au cours de laquelle le modèle apprend

à partir de données connues, suivie d'une phase de validation utilisant des données indépendantes permettant d'évaluer sa performance (Pandey *et al.* 2019; Pichler et Hartig 2023). Ils permettent de réaliser des tâches de classification ou de régression et, dans plusieurs cas, de quantifier l'importance relative des variables explicatives ainsi que leur effet sur la variable réponse (Thessen 2016). Leur utilisation apparaît donc particulièrement adaptée pour quantifier l'influence des variables environnementales et modéliser les dates de rassemblement des dorés adultes sur les sites de fraie au Québec.

1.4 Objectifs du travail

Dans cette étude, nous visons à documenter la phénologie reproductive du doré jaune à l'échelle de son aire de répartition québécoise à l'aide d'algorithmes d'apprentissage automatique. Spécifiquement, (1) les dates caractérisant la période de fraie du doré seront mises en relation avec des variables environnementales, afin de quantifier leurs influences respectives. La période de reproduction du doré jaune sera (2) modélisée sur l'ensemble de son aire de répartition québécoise, grâce à l'utilisation d'algorithmes d'apprentissage automatiques entraînée avec les variables environnementales influentes. Finalement, (3) l'effet des changements climatiques sur la période de fraie du doré jaune sera analysé, grâce à l'utilisation de scénarios climatiques disponibles.

2. MODELING THE SPAWNING PHENOLOGY OF WALLEYE (SANDER VITREUS) ACROSS QUÉBEC

2.1 Introduction

Walleye (*Sander vitreus*) is a fish species native to North America whose distribution spans a broader latitudinal range than that of other freshwater predatory fish (Bozek *et al.* 2011a; Spanou *et al.* 2024). Owing to its trophic position, its presence plays a crucial role in maintaining the balance of freshwater ecosystems, where it exerts top-down control on lower trophic levels (Lyons et Magnuson 1987). Throughout its distribution range, the species supports important commercial, recreational, and cultural fisheries (Sass 2025). In Canada, walleye constitutes the most sought-after freshwater species in commercial fisheries as well as the most sought-after species among recreational anglers across all species, accounting for 26% of total catches in 2015 (Pêches et Océans Canada 2019; Pêches et Océans Canada 2026). In the province of Quebec, the species is subject to intensive exploitation, as it represents the second most targeted species in recreational fisheries while also constituting a central component of the cultural practices and Food, Social and Ceremonial (FSC) Fisheries of several Indigenous communities (Arvisais *et al.* 2012). Consequently, walleye is exposed to high fishery removals rates that may threaten the sustainability of its stocks (Post 2013).

Between 1995 and 2008, Quebec's Ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP) documented a decline in reproductive potential together with an increase in walleye mortality rates in inland waters and in the St. Lawrence River, suggesting a situation of overexploitation (Arvisais *et al.*

2012). These observations led to the implementation, in 2011, of new management measures under the *Plan de gestion du doré 2011-2016*, later extended to 2026, including the introduction of protected slot limits and minimum length restrictions to restore population structure and spawning stock abundance (Arvisais *et al.* 2012; Arvisais *et al.* 2016). At the provincial scale, outcomes of this management plan were variable. Although overall population status improved in some waterbodies (Arvisais *et al.* 2016; Nadon 2022b; Bélanger 2024), the implemented measures did not produce the expected effects in several others populations (Bélanger 2022; Nadon 2022a; Bélanger 2023). In the St. Lawrence River, these measures failed to halt declines in spawner abundance and size, as well as the increase in total annual mortality observed since 2001, highlighting the need to reassess current walleye management strategies in Quebec (Mainguy *et al.* 2025).

Protection of spawning adults during the spawning period, through adjustments to fishing seasons, constitutes a widely used management approach. This measure reduces the risk of overexploitation in species forming large spawning aggregations by decreasing fishing mortality on older individuals, whose reproductive contribution is generally higher (Heino 1998; Venturelli *et al.* 2010; Overzee et Rijnsdorp 2014). Walleye spawning occurs in shallow areas, often at depths of less than one meter, where individuals form dense spawning aggregations for several days (Malison et Held 1996; Gendron 2009; Paquin et Brodeur 2023). During this period, fish become more easily detectable and particularly vulnerable to harvest, potentially resulting in substantial spawning stock mortality when fishing pressure is high (Paquin et Brodeur 2023; Alglave *et al.* 2024).

Alike other jurisdictions, recreational and commercial fishing regulations vary among fishing zones in the province of Québec, some encompassing extensive geographic areas (Arvisais *et al.* 2012). As a result, regulations do not always reflect local population

phenology, which may lead to overlap between fishing seasons and the walleye spawning period (Gendron 2009; Paquin et Brodeur 2023; Alglave *et al.* 2024). Although this phenomenon has been documented, environmental factors determining the period when adults aggregate at spawning sites remain poorly documented in most walleye range, especially in its northern extremity (e.g., Québec, Ontario), limiting the ability to scientifically support adjustments to fishing opening dates (Dixon *et al.* 2020).

Several authors have emphasized the importance of a thorough understanding of the reproductive phenology of exploited species to ensure optimal and proactive management (Dixon *et al.* 2020; Feiner *et al.* 2022; Feiner *et al.* 2025; Mainguy *et al.* 2025). Despite the extensive scientific literature on walleye, few studies have documented its reproductive phenology. Existing studies have primarily been conducted in the southern portion of its distribution range. In most studies, peak spawning activity has been associated with ice-off date, a climatic indicator correlated with air temperature and, indirectly, with water temperature (Schneider *et al.* 2010; Dixon *et al.* 2020; Feiner *et al.* 2022; Barta *et al.* 2024). Recent studies have nevertheless identified mismatches between these metrics, suggesting that factors other than temperature may contribute to regulating spawning timing (Barta *et al.* 2024). Photoperiod has notably been identified as a major determinant of peak spawning in Escanaba Lake between 1957 and 2019 (Feiner *et al.* 2025), contrasting with earlier studies that attributed a dominant role to water warming (Schneider *et al.* 2010; Bozek *et al.* 2011a). These inconsistencies highlight the need for a better understanding of walleye reproductive phenology, particularly given that potential dependence on photoperiodic cues may result in unfavorable consequences for recruitment (Feiner *et al.* 2022; Barta *et al.* 2024; Feiner *et al.* 2025).

In this study, we aim to document walleye reproductive phenology across its distribution range in Quebec using machine learning algorithms. Specifically, (1) the start and end dates characterizing the walleye spawning period will be related to environmental variables to quantify their respective influences. The reproductive period of walleye will (2) be modeled across its Quebec distribution range using machine learning algorithms trained on influential environmental variables. Finally, (3) the effects of climate change on the walleye spawning period will be assessed using available climate scenarios.

2.2 Materials and methods

2.2.1 Sampling sites and biological data sources

The study area includes water bodies within the native range of walleye in Québec, ranging from 45°N and 55°N. Most data were derived from published scientific literature and governmental grey literature, including reports from research projects and surveys (n = 478). These observations were collected between 1942 and 2024, although fewer than 30% observations predate 1990. A small number of additional records (n = 6) were extracted from scientific publications and reports by the Québec national hydroelectricity company, Hydro-Québec, environmental office and cover a period from 2000 to 2012 (Figure 1).

Data on presence of spawning adults were mostly obtained by visual surveys on spawning sites or by capture of spawners. When individuals were captured, sexual maturation stage was assessed either by abdominal pressure and partial extrusion of gametes according to the Nikolsky (1963) scale (1 = immature, 4 = mature, 5 = spawning, 6 = post-spawning), or by visual identification of gonadal development stage following dissection. The database comprises 484 observations of spawning adults at spawning sites and includes three observation types: “complete” observations (n = 97), associated with spawning onset and end

dates; “partial” observations (n = 124), associated with either a start or an end date; and “single” observations (n = 263), not associated with spawning onset or end dates, representing only sighting of spawning adults.

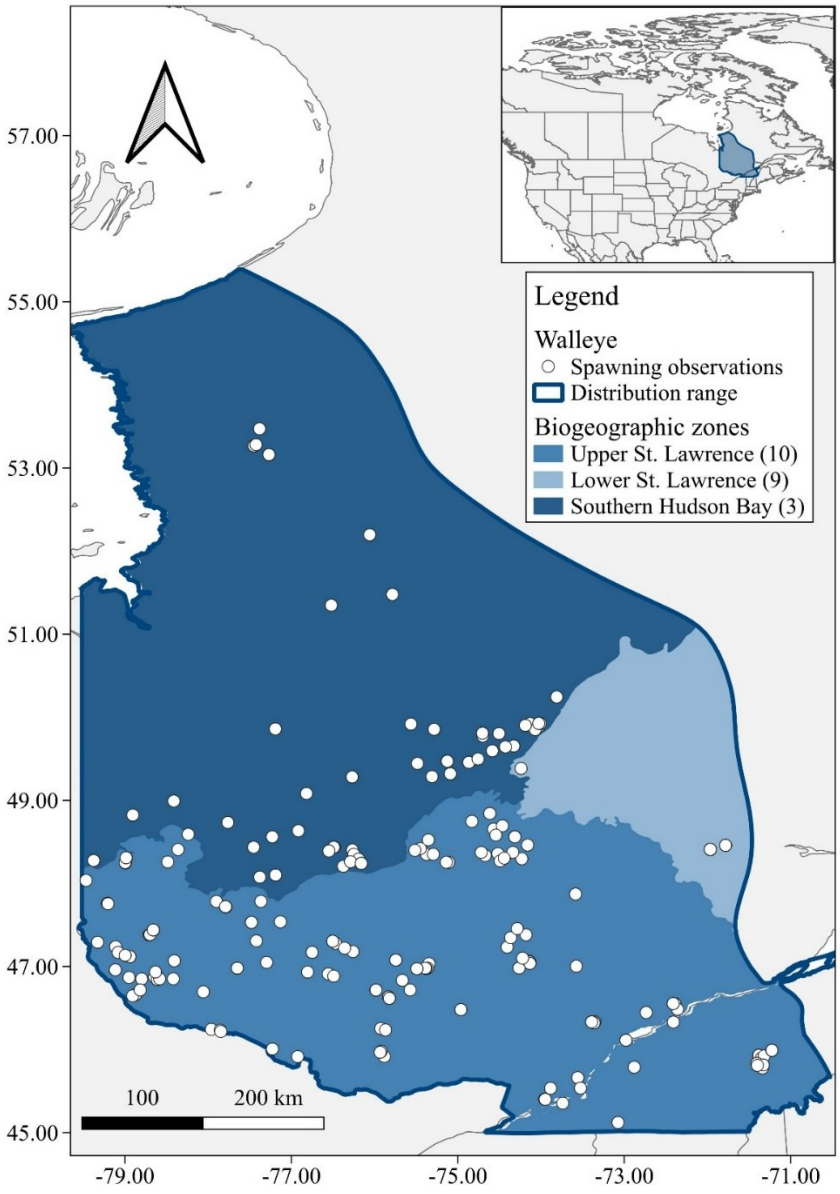


Figure 1. Sampling sites for walleye spawning observations across the Québec range, including the Canadian freshwater biogeographic zones defined by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) (2007), with zone numbers shown in parentheses.

2.2.2 Environmental data

2.2.2.1 Current and historical data

Twenty environmental variables were collected and linked to spawning observations to build models estimating spawning period onset, end, and duration (Appendix A). These variables were selected based on their accessibility and their potential influence on the species' life cycle (Bozek *et al.* 2011a; Feiner *et al.* 2025). All environmental variables were obtained from open-access data sources, including WorldClim (Fick et Hijmans 2017), Environment and Climate Change Canada and Ouranos (Gasset *et al.* 2021), the Info-Climat Service of the MELCCFP (2025) and the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) (2007), with the exception of water temperature (°C) and geographic coordinates (latitude and longitude), which were measured directly at sampling sites. The MELCCFP Info-Climat Service (2025) provided daily air temperature data across Québec through 2024. Derived from interpolation of multiples meteorological stations (approx. 550 active stations in 2020, Info-Climat, pers comm.), these data were used to model spawning onset and end dates of walleye across its range for the 1991-2020 climate period, representing current climatic conditions.

The database also includes variables that are constant across years, such as photoperiod, solar radiation, and elevation, as well as yearly variables, including mean air temperature, the date of the last frost-free day, and total spring precipitation (mm). Among the yearly variables, cumulative growing degree-days above 5 °C (GDD5) were calculated from mean daily air temperatures. The 5 °C thermal threshold was selected in accordance with the known variable describing walleye's growing season (Venturelli *et al.* 2010; Uphoff *et al.* 2013). GDD5 were then accumulated from January 1 to April 15 and from January 1 to June 30,

respectively corresponding to the minimum and maximum spawning dates in the entire range of the study.

2.2.2.2 Climate scenarios data

To inform the prediction of future spawning period, environmental variables derived from climate scenarios were obtained from the Environment and Climate Change Canada CanDCS-M6 dataset, which is based on an ensemble of 31 global climate models from the CMIP6 program (Gouvernement du Canada 2025). In this study, the ensemble-mean air temperature (°C) was used to represent the SSP2-4.5 and SSP5-8.5 scenarios. The SSP2-4.5 scenario represents an intermediate emissions pathway characterized by the absence of major abrupt changes and is generally considered the most plausible. It projects a global mean warming of approximately 2.7 °C for the period 2081-2100 (Lepousez et Aboukrat 2022). In contrast, the SSP5-8.5 scenario assumes the absence of effective mitigation policies and the continuation of current energy consumption trends. Considered unlikely in the long term, it projects a global mean warming of approximately 4.4 °C for the same period (Lepousez et Aboukrat 2022). This very high concentration scenario is often used as a proxy for post-2100 climate conditions or to explore highly unlikely extreme climate outcomes (Ouranos 2024). The 1991-2020 reference period was extracted under the SSP2-4.5 scenario to compare current and future climatic conditions, as the Environment and Climate Change Canada data server provides historical data covering the period 1950-2014 (Gouvernement du Canada 2025). Projected years from 2015 to 2020 were therefore obtained from SSP2-4.5 simulations, as this scenario is considered the most plausible.

2.3 Data analysis

2.3.1 Data standardization

To associate all observations with spawning start and end dates, a data standardization step was conducted. For this purpose, a generalized additive mixed model (GAMM) with a gamma distribution and a log link function was fitted to complete observations (n = 97), using spawning duration (in days) as the response variable and incorporating environmental variables derived from sampling sites and open-data servers. The model was implemented using the *mgcv* package in R software (R Core Team 2025). GAMMs are widely used in ecology, as they allow nonlinear relationships between predictors and response variables to be modeled using smoothing functions, while incorporating random effects to account for hierarchical or spatial structures (Pedersen *et al.* 2019). These models are relatively simple and interpretable; however, interpretation of individual effects may become more challenging in the presence of collinearity among explanatory variables (Martin *et al.* 2021; Lai *et al.* 2024). Collinearity among predictors was therefore first evaluated using a Spearman correlation matrix, applying a threshold of $|\rho| < 0.7$ as suggested by Dormann *et al.* (2013) (Appendix B). Predictor relationships were subsequently examined using pairwise concurvity estimates among GAMM smooth terms, based on a threshold of < 0.5 as proposed by Ramsay *et al.* (2003) (Appendix C).

Selected environmental variables included photoperiod at spawning onset (Onset_Photoperiod; hours), Elevation (m), the fraction of the lunar surface illuminated at midnight at spawning onset (Onset_Moon; %), maximum or minimum water temperature (Max_Twater or Min_Twater; °C), modeled maximum or minimum air temperature (Max_Tair or Max_Tmin; °C), as well as COSEWIC Canadian freshwater biogeographic

zones (Biogeographic) and watersheds (levels 2 or 1) included as random-effect terms. Final model selection was performed by comparing candidate models using the Akaike information criterion (AIC) (Apendix D). Among models with comparable AIC values ($\Delta AIC < 2$), the selected model (#3) included the fewest predictors, namely s(Max_Twater), s(Max_Tair), Onset_Photoperiod, and s(Biogeographic) as a random effect. Assumptions of normality and heteroscedasticity were evaluated through visual inspection of residuals using the *gratia* package, while predictor effects and relative importance were visualized using the *visreg* and *gam.hp* (Lai et al. 2024) packages in R software (R Core Team 2025).

The final model was used to estimate spawning duration, defined as the number of days during which spawners can be observed at a spawning site. For partial observations, the missing date (start or end of spawning) was estimated by adding or subtracting the modeled spawning duration from the observed date. For single observations, the observed date was assumed to correspond to mid-spawning and start and end dates were calculated by subtracting or adding half of the modeled duration, respectively.

2.3.2 Spawning modeling

Machine learning algorithms were used to model walleye spawning onset and termination because of their ability to handle complex biological datasets with many predictors (Brownscombe *et al.* 2020; Pichler et Hartig 2023). In particular, the Random Forest (RF) algorithm of Breiman (2001) (implemented using the *ranger* package in R ; Wright et Ziegler (2017)) and EXtreme Gradient Boosting (XGBoost) (implemented using the *xgboost* package in R; Chen et Guestrin (2016)) were used.

Random Forest is one of the most used machine learning algorithms in biological data analysis due to its strong predictive performance, robustness to overfitting, and lack of

reliance on parametric assumptions (Breiman 2001; Fox *et al.* 2017; Kang et Yoon 2024). It also exhibits strong tolerance to noise, outliers, and missing data, and remains effective when the number of predictor variables is high relative to the number of observations (Fox *et al.* 2017; Flores *et al.* 2024; Kang et Yoon 2024; Spanou *et al.* 2024). It therefore represents an attractive alternative to more traditional methods, which often require assumptions of normality and predictor independence to be satisfied (Fox *et al.* 2017; Spanou *et al.* 2024). Compared with other machine learning algorithms, Random Forest is also known to be relatively insensitive to hyperparameter choices, with default values performing well on most datasets, making it more intuitive to implement (Fox *et al.* 2017).

Random Forest is an ensemble model based on the bagging principle, relying on multiple uncorrelated decision trees whose final output is obtained by averaging predictions across trees. This process produces relatively accurate, robust, and stable predictions (Fox *et al.* 2017; Genuer et Poggi 2020; Pichler et Hartig 2023). In contrast, XGBoost is also an ensemble model but based on the boosting principle, in which decision trees are built sequentially, each new tree attempting to explain what was not captured by previous trees (Pandey *et al.* 2019; Pichler et Hartig 2023). Due to these distinct learning strategies, the joint use of Random Forest and XGBoost allows comparison of algorithm performance on the same dataset and helps identify an optimal final predictive model.

XGBoost is a relatively recent optimized extension of the Gradient Boosting Machine (GBM) algorithm, widely recognized for strong predictive performance and computational efficiency. It relies on several optimization strategies designed to accelerate tree construction, including histogram-based approaches that reduce computation time while maintaining high accuracy (Chen et Guestrin 2016; Bentéjac *et al.* 2021; Sibindi *et al.* 2022). XGBoost also incorporates multiple regularization and tree-complexity control mechanisms that help limit

overfitting compared with other GBM extensions (e.g., LightGBM; Ke *et al.* (2017) (Chen et Guestrin 2016; Bentéjac *et al.* 2021; Sibindi *et al.* 2022). Despite these mechanisms, XGBoost performance remains strongly dependent on hyperparameter selection, making it more sensitive to hyperparameter tuning than other ensemble methods such as Random Forest (Kavzoglu et Teke 2022; Bergen *et al.* 2023). Hyperparameter optimization is therefore an essential step for maximizing predictive performance and ensuring fair comparison among evaluated models (Arnold *et al.* 2024). Despite this increased complexity, several studies have shown that XGBoost can outperform Random Forest in predictive performance, supporting its inclusion in this study (Effrosynidis *et al.* 2020; Bergen *et al.* 2023; Kang et Yoon 2024).

Although machine learning algorithms such as Random Forest and XGBoost are generally robust to collinearity among predictors, removing redundant or weakly informative variables is common practice which may improve computational efficiency and, in some cases, model performance (Strobl *et al.* 2007; Effrosynidis *et al.* 2020; Martin *et al.* 2021). Accordingly, highly correlated predictors were removed prior to modeling using a Spearman correlation matrix, applying a threshold of $|\rho| > 0.8$. The final predictors retained for modeling were elevation (Elevation), COSEWIC Canadian freshwater biogeographic zones (Biogeographic), cumulative growing degree-days above 5 °C on 15 April (GDD5_15April) and 30 June (GDD5_30June), latitude (Latitude), and total spring precipitation (Spring_Precipitation) (Appendix A).

In addition to machine learning models, generalized additive mixed models (GAMMs) were also fitted using the same predictors and the same training set, consisting of 80% of the data randomly selected (set.seed(123)), to ensure comparability of results across approaches. To reduce the risk of overfitting and improve model robustness, repeated fivefold cross-

validation (5 × 5) was applied to the training set (Effrosynidis *et al.* 2020). With the exception of GAMMs, model hyperparameters were tuned individually using a grid search implemented via the *dials* package in R, with the objective of minimizing root mean squared error (RMSE; Appendix E) (Martin *et al.* 2021; R Core Team 2025). The presence of overfitting or underfitting was assessed by comparing training and testing performance (RMSE and R^2), as well as through visual inspection of plots comparing predicted and observed values (Appendix F and G) (Breiman 1996; Bonaccorso 2018; Kavzoglu et Teke 2022). Because the initial XGBoost model obtained from the grid search showed marked overfitting, the hyperparameter grid was narrowed toward more regularized configurations by limiting tree complexity and strengthening split criteria, which improved its performance and generalization performance.

Final model performance was evaluated using the remaining portion of the data (test set = 20%), from which RMSE and the coefficient of determination (R^2) were computed using the *caret* package in R (R Core Team 2025). These two metrics are widely used in the literature to compare predictive performance across algorithms (Effrosynidis *et al.* 2020; Martin *et al.* 2021). Final model selection was based on RMSE and R^2 values, as well as computation efficiency. The selected model was then applied to predict spawning onset and end dates across the Québec range of walleye using spatially extracted environmental predictors. The resulting raster maps had a spatial resolution of 0.1° × 0.1°, corresponding to the lowest resolution among the data sources used.

2.3.3 Model selection and analysis

The relative importance of environmental variables used in the Random Forest models was visualized using the *varImpPlot* function from the *randomForest* package in R (R Core

Team 2025). This function provides two complementary variable-importance measures for each predictor: %IncMSE and IncNodePurity (Dewi et Chen 2019). The %IncMSE metric relies on data not used to build each tree (out-of-bag; OOB) and evaluates the predictive contribution of a variable by quantifying the average increase in prediction error when its values are permuted (Breiman 2001; Dewi et Chen 2019). Thus, higher %IncMSE values indicate a greater contribution to model predictive performance. The second measure, IncNodePurity, reflects the contribution of variables to increasing node purity during the construction of decision trees (Dewi et Chen 2019). In regression models, this measure is based on the reduction of mean squared error (MSE) associated with successive tree splits, with higher values indicating an increased ability to partition observations into more homogeneous groups with respect to the response variable (Dewi et Chen 2019).

Predictor effects on the response variable were then explored using accumulated local effects (ALE) curves (Apley et Zhu 2020), calculated with the *iml* package in R (R Core Team 2025). This approach allows visualization of changes in model-predicted values as predictor values increase, while accounting for dependencies among explanatory variables (Gkolemis *et al.* 2022; Dyba 2024). Although ALE curves are more sensitive when applied to small datasets, they offer an advantage over more classical methods such as partial dependence plots (PDPs), as they are less affected by interactions and collinearity among predictors and rely solely on combinations observed in the training set (Gkolemis *et al.* 2022; Dyba 2024).

2.4 Results

2.4.1 Model selection and analysis

Although Random Forest and XGBoost showed comparable performance overall, Random Forest performed better in modeling spawning end dates, whereas XGBoost performed better for spawning onset dates (Table 1). Because its execution time was approximately three times shorter, the Random Forest algorithm was selected as the primary model to document and model the spawning period of walleye in Québec.

Table 1. Comparison of predictive performance among models (RandomForest, XGBoost and GAMM) for spawning timing (onset and end) of walleye in Québec

Model	Onset of spawning			End of spawning			Mean		
	RMSE	R^2	Time (s)	RMSE	R^2	Time (s)	RMSE	R^2	Time (s)
RandomForest	5.48	0.72	721.53	5.03	0.77	453.74	5.26	0.75	587.64
XGBoost	5.30	0.74	2230.06	5.14	0.75	2300.5	5.22	0.75	2265.28
GAMM	6.64	0.59	15.88	5.88	0.67	12.95	6.26	0.63	14.42

2.4.2 Spawning duration

The GAMM including maximum water temperature (°C), maximum air temperature (°C), photoperiod at spawning onset (hours), and biogeographic zones as a random effect explained 49.9% of the observed deviance in spawning duration in walleye populations sampled across Québec (Table 2). All predictors included in the model were statistically significant ($p < 0.05$). In terms of relative contribution to deviance explained, maximum water temperature was the dominant predictor (48.6%), followed by maximum air temperature (28.4%), photoperiod (14.7%), and biogeographic zones (8.3%) (Table 2).

Table 2. Effects of environmental predictors on walleye spawning duration in Québec (GAMM results and deviance explained)

Variable	**Estimate**	**Std. Error**	**t-value**	**Pr(>|t|)**	**% Deviance explained**
Onset_Photoperiod	-0.155	0.066	-2.335	0.022	14.7
Variable	**edf**	**Reference df**	**F**	**p-value**	**-**
s(Max_Twater)	1.939	2.441	19.293	< 2e-16	48.6
s(Max_Tair)	1.874	2.374	13.124	5.55E-06	28.4
s(Biogeographic) = "re"	1.686	2	4.298	0.00763	8.3
	Deviance explained (%) = 49.9			**R^2 adj. = 0.408**	

Partial effects showed that relationships between predictors and the response were generally linear. Maximum air temperature showed a marked negative effect, particularly beyond approximately 15-18 °C, though this portion of the predictor range was supported by fewer observations. In contrast, maximum water temperature showed a positive effect, with the estimated effect increasing when temperatures exceeded roughly 13-18 °C. The estimated linear effect of photoperiod at spawning onset was negative, indicating shorter spawning durations under longer day-length conditions, with observed photoperiod values ranging from approximately 13.5 to 16.5 hours. Finally, the random effect associated with biogeographic zones revealed significant differences among regions ($p = 0.008$). Biogeographic zone 3, corresponding to southern Hudson Bay and James Bay, was associated with longer spawning duration than in other zones (Figure 2).

Partial effects of maximum water temperature showed an increase in spawning duration with increasing maximum water temperature. Climatic conditions associated with observations recorded at a maximum water temperature of at least 15 °C were compared with those observed at temperatures below 15 °C using Student's t-tests with Bonferroni correction (Armstrong 2014). Only cumulative GDD5 values as of 30 June were significantly

higher when maximum water temperature was at least 15 °C (p = 0.003), whereas no significant differences were detected for the other climatic variables tested (p > 0.05; Appendix H).

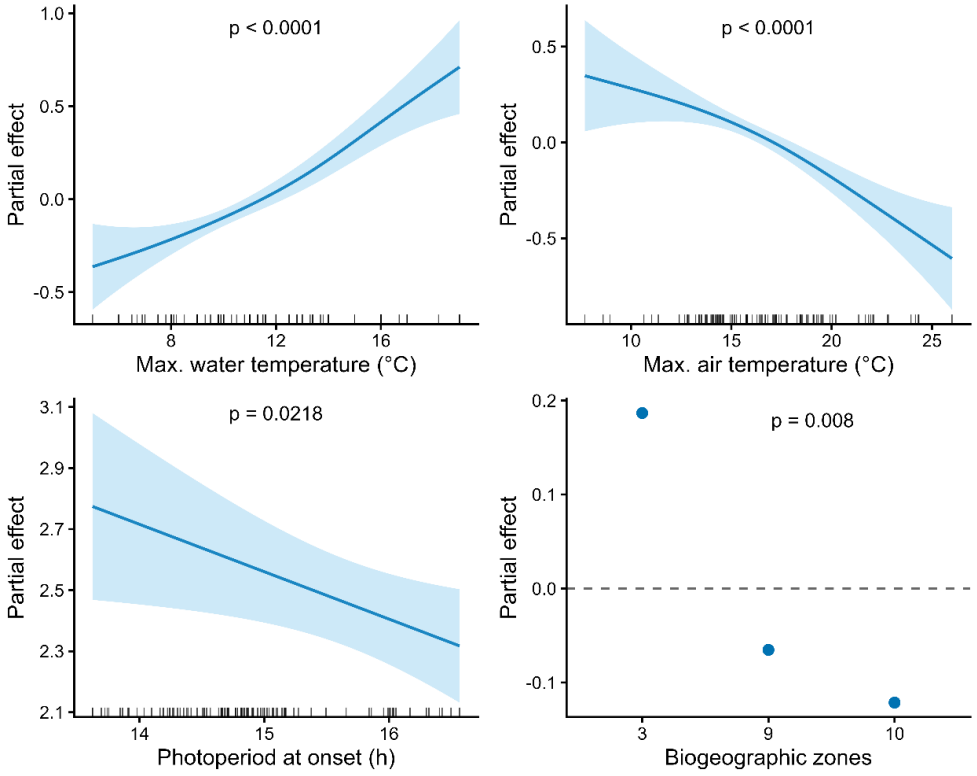


Figure 2. Effects of environmental variables on walleye spawning duration inferred from GAMM analysis.

2.4.3 Spawning onset and end

Models predicting spawning start and end dates showed comparable performance, with an RMSE of 5.48 days and an R^2 of 0.72 for spawning onset, and an RMSE of 5.03 days and an R^2 of 0.77 (Table 1). In both cases, the %IncMSE and IncNodePurity metrics identified GDD5_15April as the most important predictor in terms of both predictive performance and model structure. In the spawning onset model, this variable represents the main explanatory factor, as latitude, elevation, GDD5_30June, and spring precipitation exert a similar and

markedly weaker influence. Conversely, in the spawning end model, the importance of GDD5_15April is closely followed by elevation, latitude, GDD5_30June, and spring precipitation, in decreasing order of importance. In both models, biogeographic zones show low and comparable importance from both performance and structural perspectives. From a structural perspective, latitude and GDD5_30June exert similar influence in the spawning end model, whereas their respective contributions are reversed and overall weaker in the spawning onset model. Finally, elevation contributes moderately to model structure in both cases (Appendix I).

Accumulated local effects (ALE) curves revealed generally consistent relationships between predictors and spawning timing. In both models, GDD5_15April and GDD5_30June showed negative trends, indicating that higher GDD5 values were associated with earlier spawning dates. A latitudinal gradient in spawning timing was also observed, with populations located at higher latitudes associated with later spawning dates. In both models, a nonlinear relationship between elevation and spawning timing was identified, with earlier dates occurring at intermediate elevations (approximately 300-350 m), whereas lower or higher elevations were associated with later dates. Biogeographic zones also showed differential effects, with zone 3 (southern Hudson Bay-James) associated with generally earlier spawning dates compared with other zones. Finally, spring precipitation showed similar responses across spawning phases. For spawning onset, intermediate values (200-250 mm) were associated with relatively later dates, whereas the spawning end model highlighted a more complex response characterized by multiple changes in effect between 200 and 275 mm (Appendix J).

2.4.3.1 Modeling under current climate conditions

Across Québec, model-predicted spawning onset ranged from approximately April 23 to May 13, while spawning end ranged from May 15 to June 1. Although spawning duration remained relatively constant across the latitudinal gradient, a slight increase was observed beyond 52°N, coinciding with stabilization of spawning onset between 52-55°N. Between 45-48°N, spawning onset showed stronger latitudinal variation compared with spawning end, which showed a more gradual latitudinal pattern (Figure 3). In this region, the maps revealed substantial variation in spawning dates within the same latitudinal band, particularly in spawning onset. Notably, populations located in southern regions and in the St. Lawrence Lowlands near Montreal and along the Ottawa River exhibited generally earlier spawning dates. Earlier spawning periods were also observed in certain inland areas, notably in the Mont-Laurier depression between Rouyn-Noranda and Montreal and in the Lake Saint-Jean Lowlands near Saguenay, reflecting the influence of elevation (Figure 4).

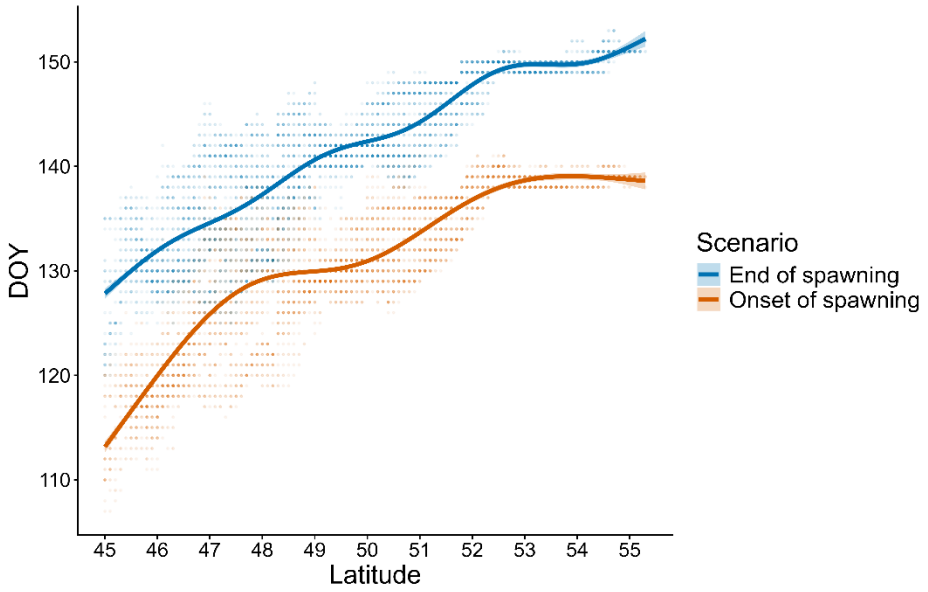


Figure 3. Latitudinal patterns in modeled spawning start and end dates of walleye in Québec for the 1991-2020 period.

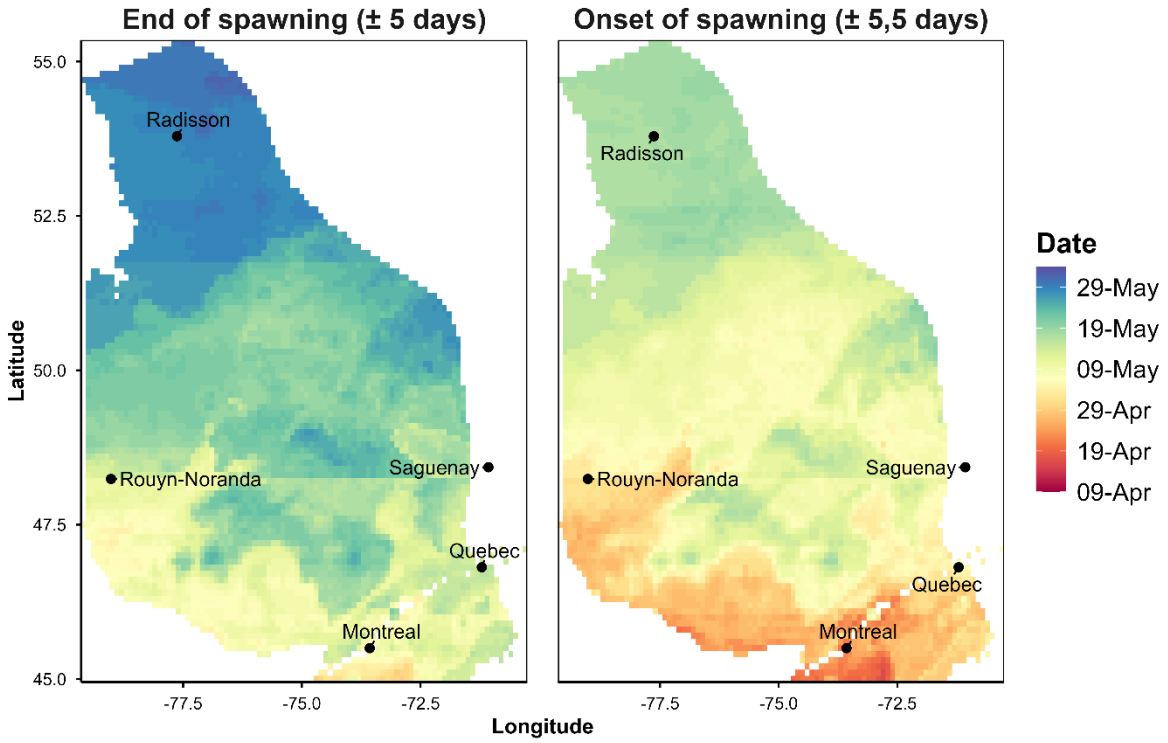


Figure 4. Modeled spawning onset (left) and end (right) dates of walleye across its Québec range (1991-2020) at 0.1° x 0.1° spatial resolution.

2.4.3.2 Modeling under future climate scenarios

Modeled spawning start and end dates under the SSP2-4.5 scenario for the 1991-2020 period showed spatial and latitudinal patterns comparable to those obtained from historical Info-Climat data (Figures 4 and 5). Dates predicted under SSP2-4.5 appeared slightly later for both spawning onset and end based on visual analysis of the maps, potentially reflecting differences between the modeled environmental predictors used for future scenarios and the observed data used for the reference period.

Comparison of future projections (2071-2100) with current conditions (SSP2-4.5 - 1991-2020) indicated a marked advancement of the spawning period under both scenarios tested. On average, spawning onset was projected to happen approximately 8 days earlier

under SSP2-4.5 and 12 days earlier under SSP5-8.5, while spawning end advanced by about 7 and 12 days, respectively. This advancement was not uniform along the latitudinal gradient under either scenario. For example, populations located south of 48°N showed a more pronounced shift toward earlier spawning under both scenarios compared with the reference period. In contrast, populations north of 52°N showed little change under SSP2-4.5, whereas a stronger advancement was observed under SSP5-8.5 (Figure 5).

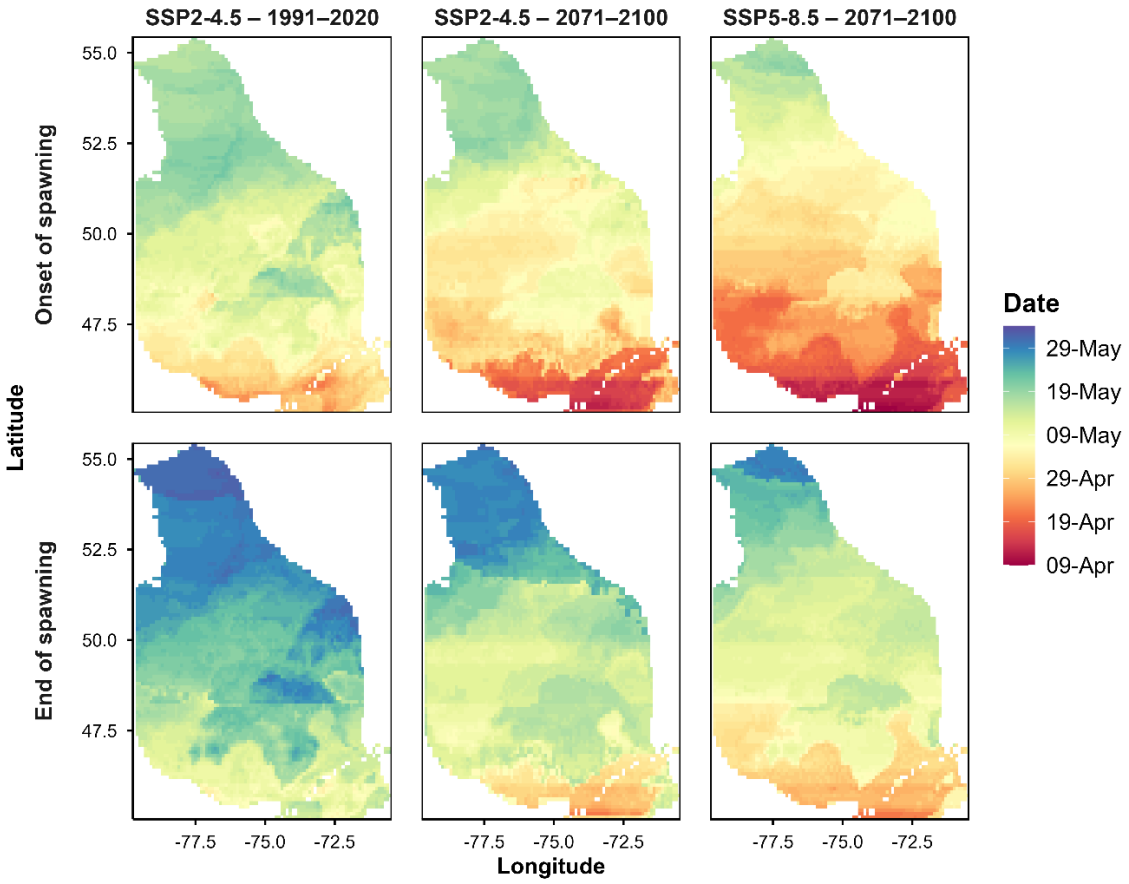


Figure 5. Modeled spawning onset (top) and end (bottom) dates of walleye across its Québec range under SSP2-4.5 (1991-2020, 2071-2100) and SSP5-8.5 (2071-2100) scenarios at 0.1° x 0.1° spatial resolution.

In addition to observations from the spatial analysis, differences in spawning onset and termination dates between the reference period (1991-2020) and the future period (2071-2100) under both scenarios revealed contrasting latitudinal patterns. Populations located at the southernmost extent of the province exhibited the greatest advances, reaching approximately 15 days. Under the SSP2-4.5 scenario, between 46°N and 52°N, advancement of spawning onset remained relatively stable (~7 days), whereas advancement of spawning termination was more variable (~5 to 10 days), particularly around 47°N, as indicated by the dispersion of observations around the trend curves. Under the SSP5-8.5 scenario, between 46°N and 50°N, advancement of spawning onset ranged from 10 to 15 days, while spawning termination remained more stable (~10 days). Beyond approximately 52°N, the magnitude of advancement decreased and became nearly negligible around 53°N (SSP2-4.5) and 55°N (SSP5-8.5). At these latitudes, the spawning period was shorter, with spawning onset showing less advancement, or even slight delays, relative to spawning termination (Figure 6).

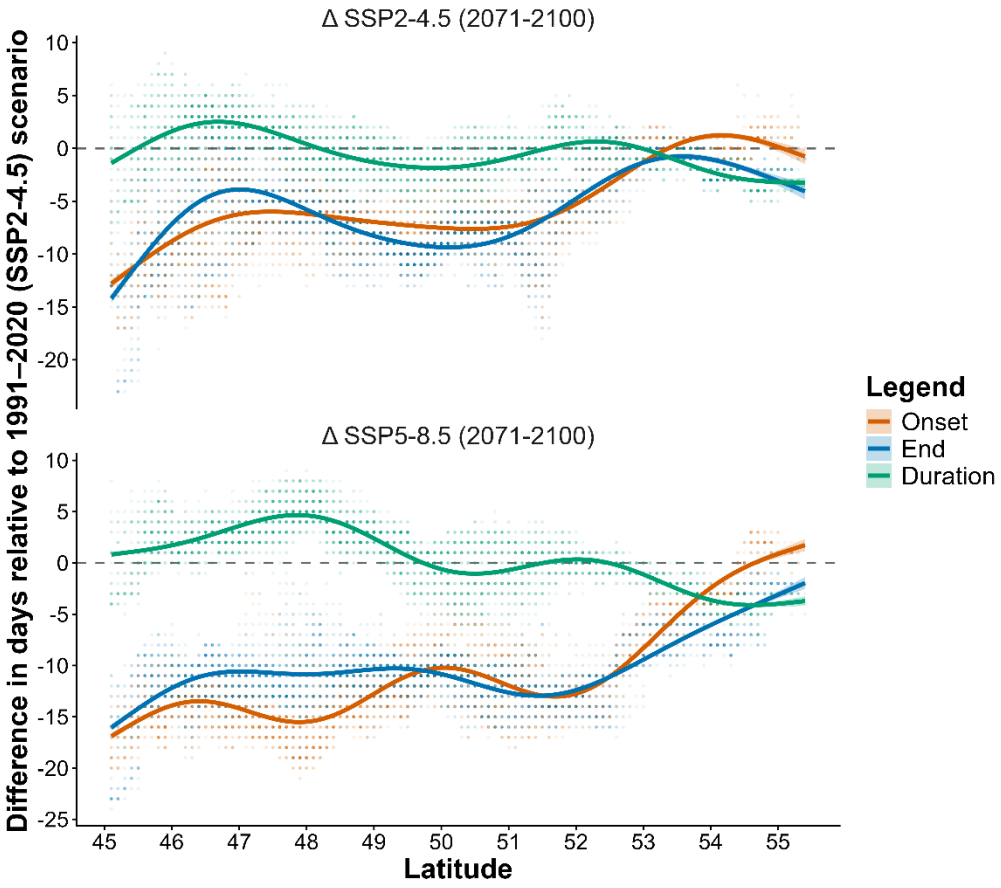


Figure 6. Latitudinal patterns in modeled spawning onset, end, and duration of walleye in Québec under SSP2-4.5 (1991-2020 and 2071-2100) and SSP5-8.5 (2071-2100).

2.5 Discussion

The overall objective of this study was to document the influence of environmental variables on the spawning period of walleye in Quebec. As suggested by the literature, temperature emerged as the determining factor in the initiation, duration, and cessation of the species' spawning period. Our results indicated a negative effect of cumulative DJC5 from January 1 to April 15 on spawning onset, with this variable identified as the dominant predictor. Cumulative DJC5 as of April 15 also emerged as the most influential variable for spawning termination, although this response was also associated with the combined effect of other environmental factors. To our knowledge, this study is among the first to have

developed a predictive model of spawning onset and termination dates for walleye across its entire distribution range in Quebec. In total, 484 observations of spawners on spawning sites, derived from both published scientific and grey literature, were used to train the machine learning models. Combined with current environmental variables and future climate projections, these models represent an innovative management tool for the species, notably in the context of recreational fishing openings and climate change.

2.5.1 Performance of machine learning algorithms

Large-scale observational biological studies, such as the present one, are generally characterized by high data variability due to the large number of factors influencing the response variables that are not systematically accounted for, as well as the inherent noise and stochasticity of natural processes (Møller et Jennions 2002; Leathwick *et al.* 2006; Martin *et al.* 2021). As a result, ecological relationships between environmental predictors and the response variable can be particularly difficult to model. By aggregating many relatively simple models, ensemble methods provide increased flexibility and the ability to capture complex interactions among predictors (Leathwick *et al.* 2006; Kosicki 2020). In several ecological studies, they have demonstrated superior predictive performance compared with more traditional approaches such as GAMMs, which rely on fitting a single parsimonious model (Leathwick *et al.* 2006; Navarro *et al.* 2019; Kosicki 2020; Martin *et al.* 2021). Our results confirm that ensemble models, particularly Random Forest and XGBoost, better represented the structure of the ecological relationships observed in this study, exhibiting higher predictive performance than GAMM (Table 1).

Random Forest and XGBoost were selected due to their widespread use in ecology and their distinct learning principles (Pichler et Hartig 2023; Cao *et al.* 2025). However, their

relative performance is known to depend strongly on dataset characteristics (Opitz et Maclin 1999; Dietterich 2000; De'ath 2007; Cao *et al.* 2025). In the present study, the evaluation metrics used (RMSE and R^2) did not discriminate between the two approaches, indicating comparable predictive performance. This similarity may be explained by the structure of the relationships observed between predictors and the response variable. Indeed, accumulated local effects (ALE) revealed relatively simple relationships close to linearity. This suggests that the enhanced capacity of XGBoost to model highly nonlinear and complex structures did not translate into performance gains, while Random Forest appeared sufficient to capture the dominant structure of the data (Dietterich 2000; De'ath 2007; Elith *et al.* 2008).

Moreover, the high variability observed in the biological data used may also have reduced the advantages of boosting approaches. Due to their sequential learning process, these methods iteratively fit residuals, including those associated with atypical observations, which can lead to partial modeling of noise (Dietterich 2000; Elith *et al.* 2008). Although XGBoost incorporates regularization mechanisms that improve its tolerance to such conditions, it generally remains more sensitive to variability than Random Forest, which relies on independent models and random predictor selection (Opitz et Maclin 1999; Dietterich 2000; Geurts *et al.* 2006; Cao *et al.* 2025).

Finally, algorithm-specific hyperparameters may also have contributed to the similarity in performance. XGBoost relies on a larger number of hyperparameters to control overfitting and model complexity (Chen et Guestrin 2016; Cao *et al.* 2025). While this flexibility represents a theoretical advantage, it also makes the model more sensitive to hyperparameter tuning and more difficult to optimize than Random Forest, which involves fewer hyperparameters and is generally more robust to their specification (Thessen 2016; Kavzoglu et Teke 2022; Bergen *et al.* 2023; Cao *et al.* 2025). Hyperparameter optimization for

XGBoost was performed using grid search and subsequently refined over a reduced grid to minimize RMSE. The process was continued until performance gains became negligible or were accompanied by overfitting. Despite this procedure, it is possible that XGBoost was not optimally tuned, thereby limiting the expected performance gains. In addition, the larger number of hyperparameters, and thus combinations to evaluate, explains why XGBoost required longer training times without providing notable predictive benefits in this context.

Overall, Random Forest and XGBoost demonstrated comparable predictive performance in this study. The lack of superiority of XGBoost is likely explained by dataset characteristics, particularly the relative simplicity of the relationships between predictors and the response variable and the inherent variability of biological data. In this context, the increased flexibility of boosting approaches did not result in additional performance gains. Conversely, the robustness of Random Forest to noise, combined with its lower sensitivity to parameter tuning and reduced computational cost, makes it a suitable approach for this type of data. Accordingly, the Random Forest algorithm was retained for modeling spawning onset and termination dates.

2.5.2 Effects of environmental variables on the reproductive window

2.5.2.1 Thermal cues

One of the main objectives of this study was to identify and quantify the relative importance of environmental variables influencing the initiation and cessation of the spawning period of walleye in Quebec. In this regard, growing degree days above 5 °C (DJC5), accumulated as of April 15, emerged as the primary predictor in both the spawning onset and termination models. DJC5 calculated from air temperature provide a robust indicator of thermal energy accumulation and a reliable proxy for water degree days (Honsey

et al. 2019), thereby indirectly capturing the thermal signal to which ectothermic organisms respond (Neuheimer et Taggart 2007). In contrast to DJC5 accumulated as of June 30, the use of April 15 as a threshold date likely better reflects spring thermal conditions specific to each year. Because Quebec's climate generally remains below 5 °C during April, this approach provides a measure of early spring thermal conditions that is independent of subsequent short-term warming or cooling events that may affect spawning timing (Ministère du Développement durable de l'Environnement de la Faune et des Parcs 2012). DJC5 accumulated as of April 15 may therefore provide a more informative measure of the thermal conditions preceding spawning.

In general, spawning initiation is associated with the first spring warming of water or with ice-out date, a climatic indicator correlated with air temperature and, indirectly, water temperature (Hokanson 1977; Malison et Held 1996; Schneider *et al.* 2010; Feiner *et al.* 2022). As these two metrics are rarely available at large spatial and temporal scales, the results of this study highlight the relevance of DJC5 accumulated as of April 15 and June 30 as integrative proxies of spawning initiation. Thus, our findings are consistent with the existing literature emphasizing the central role of temperature in the reproductive processes of percids (Hokanson 1977; Migaud *et al.* 2002; Lappalainen *et al.* 2003; Kayes et Calbert 2009; Fontaine *et al.* 2015).

In the spawning duration model, water and air temperature accounted for 49% and 28% of the explained variance, respectively, confirming the predominant role of temperature in regulating the pace of reproductive activity. Spawning generally occurs at water temperatures between 4 and 11 °C, conditions favorable for gamete viability, fecundity, and embryonic development (Koenst et Smith 1976; Hokanson 1977; Schneider *et al.* 2010; Arvisais *et al.* 2012; Feiner et Höök 2015; Feiner *et al.* 2022). Consequently, higher spring water

temperatures are typically associated with shorter spawning periods (Rawson 1957). Unexpectedly, the partial effects associated with maximum water temperature on spawning duration reveal a pattern that appears inconsistent with the literature, as well as with the results obtained for maximum air temperature, with spawning duration increasing at maximum water temperatures above 13 °C (Figure 2). This relationship should, however, be interpreted with caution. Analysis of observations associated with water temperatures ≥ 15 °C indicates that they correspond to significantly higher DJC5 accumulated as of June 30 ($p = 0.002964$; Appendix H), reflecting a context of early seasonal warming. The absence of differences between water temperature classes (< 15 °C vs ≥ 15 °C) for other thermal variables, (maximum air temperature and minimum air and water temperatures; APPENDIX H) also suggests the occurrence of cooling events during spawning, maintaining water temperature within a favorable range and thereby contributing to extended spawning duration (Bozek *et al.* 2011a). Accordingly, the effects associated with maximum water temperature should be interpreted in conjunction with the other indicators of seasonal warming, which more accurately reflect the dominant thermal conditions during spawning. Overall, these results confirm the structuring role of temperature in the initiation, duration, and cessation of spawning in walleye.

2.5.2.2 Photoperiod and latitudinal cues

Latitude, used as a proxy for the spatial variation in photoperiod, did not emerge as a major predictor in the spawning onset and termination models. Accumulated local effects showed an approximately linear relationship between latitude and both spawning onset and termination dates, consistent with the latitudinal thermal gradient commonly observed in percids (Lappalainen *et al.* 2003; Bozek *et al.* 2011a). In the spawning duration model,

photoperiod accounted for 14.7% of the explained deviance and showed a negative relationship with spawning duration (Table 2). This pattern has been reported in other studies and has been attributed to rapidly warming thermal conditions during late spawning events (Feiner *et al.* 2022; Feiner *et al.* 2025). These results suggest that the effects of latitude and photoperiod on spawning phenology are closely associated with seasonal thermal dynamics.

To our knowledge, only one study has identified photoperiod as the primary determinant of spawning peak in walleye, in contrast with the results obtained here (Feiner *et al.* 2025). In the present study, photoperiod had a limited contribution to spawning duration, and its relationship with spawning phenology may partly reflect the underlying latitudinal thermal gradient. This discrepancy may reflect local adaptation to photoperiodic cues, it may also be related to the stocking history of the studied system, where the species was introduced between 1933 and 1942 (Kempinger et Churchill 1972). Stocked populations may exhibit regulatory mechanisms distinct from those observed in natural populations (Claussen et Philipp 2022). For instance, Barta *et al.* (2024) reported that partially or fully stocked populations of walleye showed no temporal trend in spawning peak, in contrast to natural populations, which exhibited an advancement between 1940 and 2020 associated with earlier ice-out. Similar patterns have been reported in other taxa, particularly salmonids, where stocked individuals display reproductive patterns that differ from those of native populations (Stefanik et Sandheinrich 1999; Austin *et al.* 2021).

In Quebec, stocking of walleye has been relatively limited over the past fifteen years, and its effects on populations have been minimal (Geneviève Ouellet-Cauchon, personal communication). The predominance of natural populations in the present dataset and the regional scale of this study may partly explain the limited contribution of photoperiod observed here. In this context, the importance of photoperiod reported by Feiner *et al.* (2025)

may reflect a response specific to the single population they studied, or one characteristic of stocked populations, rather than a mechanism that can be generalized across the species. This interpretation remains hypothetical and would require targeted analyses on stocked populations.

2.5.2.3 Other environmental variables

Similar to latitude, the other environmental variables included in the spawning onset and termination models exhibited an influence secondary to degree-day, generally interpretable through their relationship with climatic conditions. Elevation represented the second most important predictor in the spawning termination model, followed by latitude, whereas in the spawning onset model, the remaining predictors, excluding biogeographic zones, showed comparable and secondary effects. In both Random Forest models, accumulated local effects indicated later spawning at both low (< 250 m) and high (> 350 m) elevations, the latter being predominantly associated with higher-latitude regions (Appendix J). This pattern is consistent with the combined influence of altitudinal and latitudinal thermal gradients. Indeed, colder temperatures observed at higher elevations and latitudes slow spring warming, which may delay spawning initiation, extend its duration, and postpone its termination (Webb et McLay 1996; Riedl et Peter 2013).

Spring precipitation likely exerts a similarly indirect influence by modulating the thermal and hydrological conditions of aquatic systems (Caissie 2006; Webb et Nobilis 2007). Few studies have identified precipitation as a direct determinant of walleye spawning, and no significant relationship has been observed in some systems (Bozek *et al.* 2011b; Feiner *et al.* 2025). In this study, the use of total spring precipitation, including both solid and liquid forms, complicates the interpretation of predictor effects and may explain the

absence of clear trends (Appendix J). Indeed, these two forms of precipitation exert opposing effects on the thermal conditions of water bodies. Rainfall, more frequent at lower latitudes and elevations, may promote earlier ice breakup (Chen et She 2020). Conversely, substantial snow cover, due to its insulating properties and albedo effect, limits heat exchange at the air–ice interface and delays ice breakup. Reduced snow cover may therefore promote earlier warming of aquatic systems (Chen et She 2020). Taken together, these findings suggest that precipitation may primarily influence spawning indirectly through its effects on the thermal regimes of water bodies.

Finally, biogeographic zones contributed little to model performance across all models, requiring cautious interpretation of their effects. The limited number of observations in certain zones, particularly zone 9 (Lower St. Lawrence) and northern regions of zone 3 (Southern Hudson Bay), may have reduced the ability of the models to detect potential evolutionary adaptations of populations. Overall, only zone 3 exhibited a pattern distinct from the other zones. However, it remains difficult to determine whether this pattern reflects evolutionary adaptation or is primarily attributable to the climatic conditions specific to this northern region, which may have slowed spring warming and contributed to a longer spawning duration. With respect to this predictor in the Random Forest models, results suggest generally earlier spawning in this zone, potentially reflecting a distinct evolutionary response compared with other zones. However, the marginal importance of this predictor in the models limits the strength of this interpretation. Additional data would be required to better characterize its influence.

2.5.3 Summary

Overall, the results of this study highlight the predominant role of temperature in regulating the different components of walleye spawning, including its initiation, duration, and cessation. Variables derived from thermal accumulation, such as DJC5, proved particularly effective in capturing this thermal signal, despite the absence of direct water temperature measurements. In contrast, latitude and photoperiod appeared to play secondary roles, with their effects likely reflecting the influence of the latitudinal thermal gradient rather than direct effects on spawning phenology. The spawning termination model, however, was characterized by a greater contribution of elevation and latitude, suggesting a more complex response linked to geography and local context. Other environmental variables, including precipitation and biogeographic zones, showed relatively weak effects, and their interpretation in the context of this study remains limited, but may partly reflect their association with regional gradients. Overall, although multiple predictors contributed to explaining the observed variability, the results consistently point to a dominant role of temperature in shaping the spawning dynamics of walleye in its northern distribution range.

2.5.3.1 Effects of climate change on the reproductive window

Because temperature emerged as the dominant driver of spawning phenology, projected climate change provides an opportunity to assess how future changes in thermal conditions may alter the timing and duration of reproduction across Québec. The analysis of projected trends in the spawning period of walleye for the 2071-2100 horizon, under the SSP2-4.5 and SSP5-8.5 scenarios, should therefore be interpreted considering the relationships identified between environmental variables and the different components of spawning. Overall, the results indicate an earlier date of spawning onset under both climate

scenarios, although the magnitude of this shift varies with latitude and emission scenario. This general pattern is consistent with the structuring role of temperature highlighted in this study and with overall spring warming leading to earlier attainment of the thresholds required to trigger reproduction (Ouranos 2015).

In the southern portion of the province, populations are expected to exhibit the most pronounced advances. This response can be explained by the combination of higher thermal conditions, increased winter and spring precipitation in the form of rain, and reduced snow cover, all of which promote more rapid and earlier warming of water bodies (Ouranos 2015; Chen et She 2020). However, spawning termination shows a less pronounced advance than initiation, resulting in an overall extension of the reproductive period. This pattern is consistent with the greater contribution of elevation in the spawning termination model and with the characteristic topography of southern latitudes (Appendix I). The presence of higher-elevation areas may locally slow spring warming, maintaining favorable thermal conditions over a longer period and thereby delaying spawning termination (Webb et McLay 1996; Riedl et Peter 2013).

At higher latitudes, projected trends appear more contrasted and reflect the importance of interactions between temperature and precipitation. In a warming context where winter and spring temperatures remain sufficiently low for precipitation to still occur as snow, an increase in maximum snowpack is expected (Ouranos 2015). This greater snow cover may delay the warming of water bodies and thus postpone spawning onset. However, once snow cover disappears, thermal conditions may evolve rapidly toward warm summer-like conditions due to higher air temperature (Feiner *et al.* 2022). This accelerated warming is likely to reduce the duration over which water temperatures remain within the optimal range for reproduction, resulting in a compression of the spawning period, as observed in the results

(Figure 7). Thus, projected changes are not limited to earlier spawning dates but also involve variations in spawning duration, which differ according to latitude and the topographic characteristics of the province.

2.5.4 Limitations

At the provincial scale, abiotic variables generally represent good predictors of the observed variance (Alofs et Jackson 2015). However, the use of air temperature as a proxy for water temperature relies on the assumption that water bodies within a given region respond similarly to a given level of climate warming. Lake-specific characteristics such as surface area, water clarity, depth, wind exposure and groundwater resurgence can modulate thermal responses (Fang et Stefan 2009; Read *et al.* 2014; O'Reilly *et al.* 2015). Hansen *et al.* (2016) notably highlighted a degree of thermal resilience in certain water bodies, attributed to their morphometric properties. The integration of such variables was not possible in this study due to its spatial extent, the large number of systems considered, and the lack of availability of these data at broad scales. The observed trends should therefore be interpreted with caution, as they do not account for thermal heterogeneity among systems. Additional limitations related to data structure, including the potential influence of stocking as well as the limited number of observations in certain regions of the province, have also been discussed and should be considered when interpreting the results.

Beyond these constraints, several biotic factors likely influencing the reproductive window of walleye could not be included. Population age structure may influence spawning duration, as some studies have reported staggered arrival of individuals at spawning sites, with males followed by older females and then younger ones (Bodaly 1980; Lappalainen *et al.* 2003). Populations with a broader age structure may therefore exhibit longer spawning

periods. Female body condition represents another factor potentially influencing reproductive phenology. Although poorly documented in this species, it may delay or prevent spawning in individuals lacking sufficient energy reserves to complete gamete maturation (Henderson et Morgan 2002; Jørgensen *et al.* 2006). Targeted studies would be required to better quantify the influence of these biotic factors on spawning initiation, duration, and termination. Despite these limitations, the models developed provide relevant tools for characterizing general trends in the spawning period of walleye at the scale of Quebec, as well as their potential evolution under different climate scenario

3. CONCLUSION GÉNÉRALE

L'objectif général de cette étude était de documenter l'influence des variables environnementales sur la période de fraie du doré jaune au Québec. À notre connaissance, il s'agit de la première étude ayant permis de développer un modèle prédictif des dates de début et de fin de la fraie de l'espèce à l'échelle de son aire de répartition québécoise. Notamment, l'utilisation d'une forêt d'arbres de décision a permis d'identifier et de quantifier les prédicteurs environnementaux modulant l'initiation et la cessation de la fraie, en plus de capter leur influence respective sur la variable réponse.

Les résultats obtenus mettent en évidence un contrôle dominant des conditions thermiques sur l'ensemble des composantes de la fraie. L'accumulation thermique printanière, représentée par les degrés-jours de croissance calculés au-dessus de 5 °C (DJC5) et cumulés au 15 avril, apparaît notamment comme le principal prédicteur des dates de début et de fin de la reproduction. Moins sensibles aux épisodes de chaleurs extrêmes que les DJC5 cumulés au 30 juin, les DJC5 au 15 avril ont permis de mieux refléter les conditions thermiques printanières des systèmes étudiés. Plus précisément, les résultats indiquent qu'un réchauffement printanier hâtif entraîne un déclenchement précoce de la fraie, tandis qu'un printemps tardif est associé à un début retardé et à une fraie écourtée, en raison d'un réchauffement plus rapide des conditions thermiques (Feiner *et al.* 2022; Feiner *et al.* 2025). Ces observations confirment le rôle central de la température dans l'initiation et la cessation de la fenêtre reproductive du doré jaune (Hokanson 1977; Malison et Held 1996; Schneider *et al.* 2010; Feiner *et al.* 2022).

Les autres variables environnementales intégrées aux modèles, notamment la photopériode (estimée par la latitude), l'élévation et les précipitations printanières totales,

présentent une contribution plus limitée et semblent principalement agir comme des indicateurs ou des modulateurs des conditions thermiques, plutôt que comme des facteurs déclencheurs directs. Dans cette étude, la latitude, utilisée comme proxy de la variation spatiale de la photopériode, n'a pas émergé comme un déterminant direct de la fraie. Son influence semble plutôt refléter celle du gradient thermique latitudinal. Ce résultat suggère que les variations spatiales observées à l'échelle du Québec sont principalement associées aux dynamiques climatiques régionales. De même, l'influence de l'élévation et des précipitations semble s'expliquer par leur rôle dans la modulation du réchauffement printanier, notamment à travers leurs effets sur la dynamique du réchauffement printanier et sur les conditions d'enneigement (Webb et McLay 1996; Riedl et Peter 2013; Chen et She 2020). Dans l'ensemble, ces observations confirment que, malgré la diversité des variables considérées, la température demeure le facteur structurant de la dynamique de fraie du doré jaune au Québec (Hokanson 1977; Malison et Held 1996; Schneider *et al.* 2010; Feiner *et al.* 2022).

Quant aux projections climatiques, elles ont indiqué un devancement généralisé de la période de fraie à l'horizon 2071-2100, sous les scénarios SSP2-4.5 et SSP5-8.5, bien que l'ampleur de ce déplacement varie selon la latitude et le scénario d'émissions. Ce patron général est cohérent avec le rôle structurant de la température mis en évidence dans cette étude et d'un réchauffement global des températures printanières se traduisant directement par une atteinte plus précoce des seuils nécessaires au déclenchement de la reproduction (Ouranos 2015). Les résultats montrent toutefois que ces changements ne se limitent pas à un simple décalage des dates de fraie, mais s'accompagnent également de modifications de la durée de la période de reproduction. Dans les régions méridionales, le début de la fraie se déplace plus rapidement que sa fin, ce qui se traduit par un allongement de la période de

reproduction. À l'inverse, dans les régions nordiques, une contraction de la période de fraie est observée, caractérisée par un décalage plus faible du début et un devancement plus prononcé de la fin. Ces contrastes mettent en évidence l'importance de la topographie et des précipitations printanières dans la modulation du rythme de réchauffement printanier, influençant ainsi la durée de la fraie.

Certaines limites aux modèles doivent toutefois être considérées. L'utilisation de la température de l'air comme métrique de la température de l'eau ne permet pas de tenir compte de l'hétérogénéité des réponses thermiques entre les plans d'eau face aux changements climatiques. De plus, le faible nombre d'observations dans certains secteurs, notamment dans la zone biogéographique 9 (Bas-Saint-Laurent), à l'extrémité nord de l'aire d'étude ou dans des systèmes complexes comme le fleuve Saint-Laurent, a pu limiter la détection d'adaptations locales. Enfin, certains facteurs biotiques susceptibles d'influencer la fraie, tels que la structure d'âge des populations ou la condition corporelle des individus, n'ont pas été intégrés aux modèles.

Les résultats de cette étude présentent un intérêt direct pour la gestion du doré jaune au Québec. Cette espèce étant particulièrement vulnérable lors de la fraie, en raison de la concentration de géniteurs dans des secteurs restreints, un chevauchement entre la période de fraie et l'ouverture de la pêche sportive peut entraîner la mortalité de nombreux géniteurs avant qu'ils n'aient contribué au renouvellement du stock. Dans cette optique, les modèles développés constituent des outils permettant une gestion proactive adaptée aux réalités locales des populations. Ils peuvent notamment servir d'assise scientifique à l'établissement de dates d'ouverture de la pêche sportive maximisant la protection des rassemblements de reproducteurs. Dans un contexte de changements climatiques, cette protection devient d'autant plus importante. La combinaison de pressions supplémentaires, telles que les

événements météorologiques extrêmes, les espèces envahissantes, les activités anthropiques et les risques d'asynchronies trophiques, accentue les risques pour la pérennité des populations. La mise en place d'une gestion intégrée axée sur les facteurs de pression contrôlables devient ainsi essentielle afin de limiter les risques de déclin. Ces éléments soulignent la nécessité de poursuivre les travaux afin de mieux comprendre les interactions entre facteurs abiotiques et biotiques. Pour de futures recherches, l'ajout d'observations dans les secteurs sous-représentés, ainsi que l'intégration de variables permettant de distinguer les systèmes complexes, constitue des solutions accessibles afin d'améliorer la performance et la précision des modèles.

LISTE DE RÉFÉRENCES

Alglave B, Olmos M, Casemajor J, Etienne M-P, Rivot E, Woillez M et Vermard Y. 2024. Investigating fish reproduction phenology and essential habitats by identifying the main spatio-temporal patterns of fish distribution. ICES Journal of Marine Science, 81 : 1563-1574.

Alofs KM et Jackson DA. 2015. The abiotic and biotic factors limiting establishment of predatory fishes at their expanding northern range boundaries in Ontario, Canada. Global Change Biology, 21 : 2227-2237.

Apley DW et Zhu J. 2020. Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society, 82 : 1059-1086.

Armstrong RA. 2014. When to use the Bonferroni correction. Ophthalmic and Physiological Optics, 34 : 502-508.

Arnold C, Biedebach L, Küpfer A et Neunhoeffer M. 2024. The role of hyperparameters in machine learning models and how to tune them. Political Science Research and Methods, 12 : 841-848.

Arvisais M, Paradis Y et Thibault I. 2016. Plan de gestion du doré au Québec 2016-2026. Ministère des Forêts, de la Faune et des Parcs, Direction de l'expertise sur la faune aquatique, Québec, 12 p.

Arvisais M, Nadeau D, Legault M, Fournier H, Bouchard F et Paradis Y. 2012. Plan de gestion du doré au Québec 2011-2016. Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs, Direction générale de l'expertise sur la faune et ses habitats et Direction de la faune aquatique, Québec, 73 p.

Austin CS, Essington TE et Quinn TP. 2021. In a warming river, natural-origin Chinook salmon spawn later but hatchery-origin conspecifics do not. Canadian Journal of Fisheries and Aquatic Sciences, 78 : 68-77.

Bade AP, Binder TR, Faust MD, Vandergoot CS, Hartman TJ, Kraus RT, Krueger CC et Ludsin SA. 2019. Sex-based differences in spawning behavior account for male-biased harvest in Lake Erie walleye (*Sander vitreus*). Canadian Journal of Fisheries and Aquatic Sciences, 76 : 2003-2012.

Barta ME, Sass GG, Reed JR, Cichosz TA, Shultz AD, Luehring M et Feiner ZS. 2024. Lagging spawning and increasing phenological extremes jeopardize walleye (*Sander vitreus*) in north-temperate lakes. Limnology and Oceanography Letters, 9 : 229-236.

Beamish RJ. 1976. Acidification of lakes in Canada by acid precipitation and the resulting effects on fishes. Water, Air, and Soil Pollution, 6 : 501-514.

Bélanger M. 2022. État de situation du doré jaune au lac Hébert - Bilan des inventaires de 2008 et 2020. Ministère des Forêts, de la Faune et des Parcs, Direction de la gestion de la faune de l'Abitibi-Témiscamingue, 8 p.

Bélanger M. 2023. État de situation du doré jaune au lac Joannès - Bilan des inventaires de 2011 et 2022. Ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs, Direction de la gestion de la faune de l'Abitibi-Témiscamingue, 8 p.

Bélanger M. 2024. État de situation du doré jaune au réservoir Dozois - Bilan des inventaires de 2010 à 2023. Ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs, Direction de la gestion de la faune de l'Abitibi-Témiscamingue, 9 p.

Bentéjac C, Csörgő A et Martínez-Muñoz G. 2021. A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54 : 1937-1967.

Bergen S, Huso MM, Duerr AE, Braham MA, Schmuecker S, Miller TA et Katzner TE. 2023. A review of supervised learning methods for classifying animal behavioural states from environmental features. Methods in Ecology and Evolution, 14 : 189-202.

Bodaly RA. 1980. Pre- and post-spawning movements of walleye, *Stizostedion vitreum*, in Southern Indian Lake, Manitoba. Minister of Supply and Services Canada, Department of Fisheries and Oceans, Winnipeg, 30 p.

Bonaccorso G. 2018. Machine learning algorithms : popular algorithms for data science and machine learning. Packt Publishing, Birmingham, 501 p.

Bozek M, Baccante D et Lester N. 2011a. Walleye and sauger life history. Dans : Barton BA éd. Biology, management, and culture of walleye and sauger. American Fisheries Society, Portland, p. 233-301.

Bozek M, Raabe J et Haxton T. 2011b. Walleye and sauger habitat. Dans : Barton BA éd. Biology, management, and culture of walleye and sauger. American Fisheries Society, p. 133-197.

Branco VV, Correia L et Cardoso P. 2023. The use of machine learning in species threats and conservation analysis. Biological Conservation, 283 : 110091.

Breiman L. 1996. Heuristics of instability and stabilization in model selection. The Annals of Statistics, 24 : 2350-2383.

Breiman L. 2001. Random Forests. Machine Learning, 45 : 5-32.

Brownscombe JW, Griffin LP, Morley D, Acosta A, Hunt J, Lowerre-Barbieri SK, Adams AJ, Danylchuk AJ et Cooke SJ. 2020. Application of machine learning algorithms to identify cryptic reproductive habitats using diverse information sources. Oecologia, 194 : 283-298.

Caissie D. 2006. The thermal regime of rivers: a review. Freshwater Biology, 51 : 1389-1406.

Cao Y, Schartel TE, Houghton DC, Ross J et Infante DM. 2025. Evaluating machine learning algorithms for accuracy, stability, and among-predictors discriminability in modeling species-richness across ten datasets. Ecological Informatics, 90 : 103323.

Chen T et Guestrin C. 2016. XGBoost: a scalable tree boosting system (Article). Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13 août 2016, San Francisco.

Chen Y et She Y. 2020. Long-term variations of river ice breakup timing across Canada and its response to climate change. Cold Regions Science and Technology, 176.

Claussen JE et Philipp DP. 2022. Assessing the role of supplementation stocking: a perspective. Fisheries Management and Ecology, 30 : 583-591.

Colby PJ, Ryan PA, Schupp DH et Serns SL. 1987. Interactions in north-temperate lake fish communities. Canadian Journal of Fisheries and Aquatic Sciences, 44 : s104-s128.

Committee on the Status of Endangered Wildlife in Canada (COSEWIC). 2007 (mis à jour le novembre 2023). National freshwater biogeographic zones. Consulté le 25 janvier 2026, https://www.cosewic.ca/index.php/fr/rapports/preparation-rapports-situation/lignes-directrices-reconnaitre-unites-designables.html

Conseil canadien pour la conservation des espèces en péril. 2006. Les espèces sauvages 2005: la situation des espèces au Canada. Gouvernement du Canada, 161 p.

Cutler DR, Edwards Jr. TC, Beard KH, Cutler A, Hess KT, Gibson J et Lawler JJ. 2007. Random Forests for classification in ecology. Ecology, 88 : 2783-2792.

De'ath G. 2007. Boosted trees for ecological modeling and prediction. Ecology, 88 : 243-251.

Dewi C et Chen R-G. 2019. Random forest and support vector machine on features selection for regression analysis. International Journal of Innovative Computing, Information and Control, 15 : 2027–2037.

Dietterich TG. 2000. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. Machine Learning, 40 : 139-157.

Dixon HJ, Harrison G, Lister AL et MacLatchy DL. 2020. Reproductive status of walleye (*Sander vitreus*) and lake whitefish (*Coregonus clupeaformis*) in two large, shallow Canadian subarctic lakes. Environmental Biology of Fishes, 103 : 1119-1136.

Dormann C, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Diekötter T, García Márquez J, Gruber B, Lafourcade B, Leitão P, Münkemüller T, McClean C, Osborne P, Reineking B, Schröder B, Skidmore A, Zurell D et Lautenbach S. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36 : 27-46.

Dyba K. 2024. Explanation of the influence of geomorphometric variables on the landform classification based on selected areas in Poland. Scientific Reports, 14 : 5447.

Effrosynidis D, Tsikliras A, Arampatzis A et Sylaios G. 2020. Species distribution modelling via feature engineering and machine learning for pelagic fishes in the Mediterranean Sea. Applied Sciences, 10 : 8900.

Elith J, Leathwick JR et Hastie T. 2008. A working guide to boosted regression trees. Journal of Animal Ecology, 77 : 802-813.

Elliott CW, Ridgway MS, Brown E et Tufts BL. 2022. Spatial ecology of Bay of Quinte walleye (*Sander vitreus*): annual timing, extent, and patterns of migrations in eastern Lake Ontario. Journal of Great Lakes Research, 48 : 159-170.

Embke H, Carpenter S, Isermann D, Coppola G, Beard TD, Lynch A, Sass G, Feiner Z et Vander Zanden J. 2022. Resisting ecosystem transformation through an intensive whole-lake fish removal experiment. Fisheries Management and Ecology, 29 : 364-377.

Engel S, Hoff MH et Newman SP. 2000. Walleye fry hatching, diet, growth and abundance in Escanaba Lake, Wisconsin, 1985-1992. Wisconsin Department of Natural Resources, 17 p.

Fang X et Stefan H. 2009. Simulations of climate effects on water temperature, dissolved oxygen, and ice and snow covers in lakes of the contiguous U.S. under past and future climate scenarios. Limnology and Oceanography, 54.

FAO. 2024. La situation mondiale des pêches et de l'aquaculture en 2024 - La transformation bleue en action. Rome.

Feiner ZS et Höök TO. 2015. Environmental biology of percid fishes. Dans : Kestemont P*, et al.* éds. Biology and Culture of Percid Fishes: Principles and Practices. Springer Netherlands, Dordrecht, p. 61-100.

Feiner ZS, Shaw SL et Sass GG. 2025. Understanding shifting cues for walleye spawning phenology and recruitment in a changing climate. Canadian Journal of Fisheries and Aquatic Sciences, 82 : 1-10.

Feiner ZS, Dugan HA, Lottig NR, Sass GG et Gerrish GA. 2022. A perspective on the ecological and evolutionary consequences of phenological variability in lake ice on north-temperate lakes. Canadian Journal of Fisheries and Aquatic Sciences, 79 : 1590-1604.

Fick SE et Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37 : 4302-4315.

Firkus T, Branville C, Neibauer J, Hartleb C, Holmes K, Hauser E et Fischer G. 2024. Induction of out-of-season spawning in an intensively reared walleye (*Sander vitreus*) broodstock. Aquaculture, Fish and Fisheries, 4.

Flores A, Wiff R, Donovan CR et Gálvez P. 2024. Applying machine learning to predict reproductive condition in fish. Ecological Informatics, 80 : 102481.

Fontaine P, Wang N et Hermelink B. 2015. Broodstock management and control of the reproductive cycle. Dans : Kestemont P*, et al.* éds. Biology and culture of percid fishes: Principles and practices. Springer Netherlands, Dordrecht, p. 103-122.

Fox EW, Hill RA, Leibowitz SG, Olsen AR, Thornbrugh DJ et Weber MH. 2017. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. Environmental Monitoring and Assessment, 189 : 316.

Gasset N, Fortin V, Dimitrijevic M, Carrera ML, Bilodeau B, Muncaster R, Gaborit É, Roy GP, Pentcheva N, Bulat M, Wang X, Pavlovic R, Lespinas F et Khedhaouiria D. 2021. A 10 km North American precipitation and land surface reanalysis based on the GEM atmospheric model. Hydrology and Earth System Sciences, 24 : 4917-4945.

Gendron M-H. 2009. Synthèse des connaissances sur le doré jaune (*Sander vitreus*) et la lotte (*Lota lota*) du lac Saint-Jean. Corporation de LACtivité Pêche Lac-Saint-Jean, Dolbeau-Mistassini, 88 p.

Genuer R et Poggi J-M. 2020. Random Forests. Dans : Genuer R et Poggi J-M éds. Random Forests with R. Springer International Publishing, Cham, p. 33-55.

Geurts P, Ernst D et Wehenkel L. 2006. Extremely randomized trees. Machine Learning, 63 : 3-42.

Gkolemis V, Dalamagas T et Diou C. 2022. DALE: differential accumulated local effects for efficient and accurate global explanations. Proceedings of Machine Learning Research, 189 : 375-390.

Gouvernement du Canada. 2025. Téléchargez les scénarios climatiques statistiquement réduits de CMIP6 (CanDCS-U6 & CanDCS-M6). Consulté le 26 janvier 2026, https://climate-scenarios.canada.ca/?page=CanDCS6-data

Gouvernement du Québec. 2026 (mis à jour le 20 février 2026). Fiches des espèces fauniques. Consulté le 24 février 2026, https://www.quebec.ca/agriculture-environnement-et-

ressources-naturelles/faune/animaux-sauvages-quebec/liste-des-especes-fauniques?id=23879&tx_solr%5Bq%5D=&tx_solr%5Bfilter%5D%5B%5D=species_groups%3AEsp%C3%A8ce%20p%C3%AAch%C3%A9e&tx_solr%5Bsort%5D=alphaAsc%20asc&tx_solr%5Bpage%5D=3

Hansen GJA, Read JS, Hansen JF et Winslow LA. 2016. Projected shifts in fish species dominance in Wisconsin lakes under climate change. Global Change Biology, 23 : 1463-1476.

Hansen GJA, Winslow LA, Read JS, Treml M, Schmalz PJ et Carpenter SR. 2019. Water clarity and temperature effects on walleye safe harvest: an empirical test of the safe operating space concept. Ecosphere, 10 : e02737.

Hansen GJA, Ahrenstorff TD, Bethke BJ, Dumke JD, Hirsch J, Kovalenko KE, LeDuc JF, Maki RP, Rantala HM et Wagner T. 2020. Walleye growth declines following zebra mussel and *Bythotrephes* invasion. Biological Invasions, 22 : 1481-1495.

Hansen GJA, Ruzich JK, Krabbenhoft CA, Kundel H, Mahlum S, Rounds CI, Van Pelt AO, Eslinger LD, Logsdon DE et Isermann DA. 2022. It's complicated and it depends: A review of the effects of ecosystem changes on walleye and yellow perch populations in North America. North American Journal of Fisheries Management, 42 : 484-506.

Hansen J, Sato M, Ruedy R, Lo K, Lea DW et Medina-Elizade M. 2006. Global temperature change. Proceedings of the National Academy of Sciences, 103 : 14288-14293.

Hansen MJ, Bozek MA, Newby JR, Newman SP et Staggs MD. 1998. Factors affecting recruitment of walleyes in Escanaba Lake, Wisconsin, 1958–1996. North American Journal of Fisheries Management, 18 : 764-774.

Hayden TA, Holbrook CM, Fielder DG, Vandergoot CS, Bergstedt RA, Dettmers JM, Krueger CC et Cooke SJ. 2014. Acoustic telemetry reveals large-scale migration patterns of walleye in Lake Huron. PLoS One, 9 : e114833.

Heino M. 1998. Management of evolving fish stocks. Canadian Journal of Fisheries and Aquatic Sciences, 55 : 1971-1982.

Henderson BA et Morgan GE. 2002. Maturation of walleye by age, size and surplus energy. Journal of Fish Biology, 61 : 999-1011.

Hokanson KEF. 1977. Temperature requirements of some percids and adaptations to the seasonal temperature cycle. Journal of the Fisheries Research Board of Canada, 34 : 1524-1550.

Honsey AE, Venturelli PA et Lester NP. 2019. Bioenergetic and limnological foundations for using degree-days derived from air temperatures to describe fish growth. Canadian Journal of Fisheries and Aquatic Sciences, 76 : 657-669.

IPCC. 2014. Climate change 2014 : impacts, adaptations, and vulnerability. Part B : regional aspects. . Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, 688 p.

Jørgensen C, Ernande B, Fiksen Ø et Dieckmann U. 2006. The logic of skipped spawning in fish. Canadian Journal of Fisheries and Aquatic Sciences, 63 : 200-211.

Kang H et Yoon SC. 2024. Enhancing length at first maturity estimation using machine learning for fisheries resource management: a case study on small yellow croaker (*Larimichthys polyactis*) in South Korea. Fishes, 9 : 373.

Kavzoglu T et Teke A. 2022. Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). Bulletin of Engineering Geology and the Environment, 81 : 201.

Kayes T et Calbert H. 2009. Effect of photoperiod and temperature on the spawning of yellow perch (*Perca flavescens*). Proceedings of the World Mariculture Society, 10 : 306-316.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q et Liu T-Y. 2017. LightGBM: a highly efficient gradient boosting decision tree. 31st Conference on Neural Information Processing Systems, Long Beach.

Kempinger JJ et Churchill WS. 1972. Contribution of native and stocked walleye fingerlings to the anglers' catch, Escanaba Lake, Wisconsin. Transactions of the American Fisheries Society, 101 : 644-649.

Koenst WM et Smith LL. 1976. Thermal requirements of the early life history stages of walleye, *Stizostedion vitreum vitreum*, and sauger, *Stizostedion canadense*. Journal of the Fisheries Research Board of Canada, 33 : 1130-1138.

Kosicki JZ. 2020. Generalised additive models and random forest approach as effective methods for predictive species density and functional species richness. Environmental and Ecological Statistics, 27 : 273-292.

Lai J, Tang J, Li T, Zhang A et Mao L. 2024. Evaluating the relative importance of predictors in generalized additive models using the gam.hp R package. Plant Diversity, 46 : 542-546.

Lappalainen J, Dörner H et Wysujack K. 2003. Reproduction biology of pikeperch (*Sander lucioperca* (*L.*)) – a review. Ecology of Freshwater Fish, 12 : 95-106.

Leathwick JR, Elith J, Francis M, Hastie T et Taylor P. 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. Marine Ecology Progress Series, 321 : 267-281.

Lepousez V et Aboukrat M. 2022. Les scénarios SSP : décryptage et recommandations d'utilisation pour une démarche d'adaptation au changement climatique. Consulté le 29 janvier 2026, https://www.carbone4.com/publication-scenarios-ssp-adaptation

Lester NP, Dextrase AJ, Kushneriuk RS, Rawson MR et Ryan PA. 2004. Light and temperature: key factors affecting walleye abundance and production. Transactions of the American Fisheries Society, 133 : 588-605.

Lyons J et Magnuson JJ. 1987. Effects of walleye predation on the population dynamics of small littoral-zone fishes in a northern Wisconsin lake. Transactions of the American Fisheries Society, 116 : 29-39.

Mahlum S, Vitense K, Corson-Dosch H, Platt L, Read JS, Schmalz PJ, Treml M et Hansen GJA. 2023. Connecting habitat to species abundance: the role of light and temperature on the abundance of walleye in lakes. Canadian Journal of Fisheries and Aquatic Sciences, 80 : 273-286.

Mainguy J, Paradis Y et Moral RdA. 2025. Limited demographic effects one decade after implementation of a harvest-slot length limit for walleye (*Sander vitreus*) in the St. Lawrence River, Québec, Canada. Fisheries Management and Ecology, 32 : 388-400.

Malison JA et Held JA. 1996. Reproduction and spawning in walleye (*Stizostedion vitreum*). Journal of Applied Ichthyology, 12 : 153-156.

Malison JA, Procarione LS, Barry TP, Kapuscinski AR et Kayes TB. 1994. Endocrine and gonadal changes during the annual reproductive cycle of the freshwater teleost, *Stizostedion vitreum*. Fish Physiology and Biochemistry, 13 : 473-484.

Martin B, González-Arias J et Vicente-Vírseda J. 2021. Machine learning as a successful approach for predicting complex spatio–temporal patterns in animal species abundance. Animal Biodiversity and Conservation, 44 : 289-301.

Migaud H, Fontaine P, Sulistyo I, Kestemont P et Gardeur J-N. 2002. Induction of out-of-season spawning in Eurasian perch *Perca fluviatilis*: effects of rates of cooling and cooling durations on female gametogenesis and spawning. Aquaculture, 205 : 253-267.

Ministère de l'Environnement de la Lutte contre les changements climatiques de la Faune et des Parcs. 2025. Service Info-Climat. Consulté le 21 février 2025, https://www.environnement.gouv.qc.ca/climat/surveillance/produits.htm

Ministère du Développement durable de l'Environnement de la Faune et des Parcs. 2012. Température moyenne - Moyenne avril basée sur les normales 1981-2010. Consulté le 31 mars 2026, https://www.environnement.gouv.qc.ca/climat/normales/cartes/temperature-moyenne.pdf

Møller A et Jennions MD. 2002. How much variance can be explained by ecologists and evolutionary biologists? Oecologia, 132 : 492-500.

Nadon L. 2022a. État de situation du doré jaune au réservoir aux Sables - Bilan des inventaires de 2011 à 2020. Ministère des Forêts, de la Faune et des Parcs, Direction de la gestion de la faune de Lanaudière et des Laurentides, 10 p.

Nadon L. 2022b. État de situation du doré jaune au Grand lac Nominingue - Bilan des inventaires de 2000 à 2020. Ministère des Forêts, de la Faune et des Parcs, Direction de la gestion de la faune de Lanaudière et des Laurentides, 11 p.

Nate NA, Bozek MA, Hansen MJ, Ramm CW, Bremigan MT et Hewett SW. 2003. Predicting the occurrence and success of walleye populations from physical and biological features of northern Wisconsin lakes. North American Journal of Fisheries Management, 23 : 1207-1214.

Navarro M, Hailu A, Langlois T, Ryan KL et Kragt ME. 2019. Determining spatial patterns in recreational catch data: a comparison of generalized additive mixed models and boosted regression trees. ICES Journal of Marine Science, 77 : 2216-2225.

Neuheimer AB et Taggart CT. 2007. The growing degree-day and fish size-at-age: the overlooked metric. Canadian Journal of Fisheries and Aquatic Sciences, 64 : 375-385.

Nikolsky GV. 1963. The ecology of fishes. Academic Press, London, 352 p.

O'Reilly CM, Sharma S, Gray DK, Hampton SE, Read JS, Rowley RJ, Schneider P, Lenters JD, McIntyre PB, Kraemer BM, Weyhenmeyer GA, Straile D, Dong B, Adrian R, Allan MG, Anneville O, Arvola L, Austin J, Bailey JL, Baron JS, Brookes JD, de Eyto E, Dokulil MT, Hamilton DP, Havens K, Hetherington AL, Higgins SN, Hook S, Izmest'eva LR, Joehnk KD, Kangur K, Kasprzak P, Kumagai M, Kuusisto E, Leshkevich G, Livingstone DM, MacIntyre S, May L, Melack JM, Mueller-Navarra DC, Naumenko M, Noges P, Noges T, North RP, Plisnier P-D, Rigosi A, Rimmer A, Rogora M, Rudstam LG, Rusak JA, Salmaso N, Samal NR, Schindler DE, Schladow SG, Schmid M, Schmidt SR, Silow E, Soylu ME, Teubner K, Verburg P, Voutilainen A, Watkinson A, Williamson CE et Zhang G. 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophysical Research Letters, 42 : 10,773-710,781.

Opitz D et Maclin R. 1999. Popular ensemble methods: an empirical study. Journal of Artificial Intelligence Research, 11 : 169-198.

Ouranos. 2015. Synthèse des connaissances sur les changements climatiques au Québec. Ouranos, Montréal, 415 p.

Ouranos. 2024 (mis à jour le 15 janvier 2024). Guide de recommandations scientifiques visant à faciliter l'utilisation de projections climatiques. Consulté le 9 mars 2026, https://www.ouranos.ca/fr/guide-de-recommandations-scientifiques

Overzee H et Rijnsdorp A. 2014. Effects of fishing during the spawning period: implications for sustainable management. Reviews in Fish Biology and Fisheries, 25.

Pandey D, Niwaria K et Chourasia B. 2019. Machine learning algorithms: a review. International Research Journal of Engineering and Technology (IRJET), 06 : 916-922.

Paquette C, Gagné S, Gaudet-Boulay M, Ouellet-Cauchon G, Leclerc V, Arkilanian A, Morissette O, Rodríguez MA, Turgeon K, Taranu ZE, Beisner BE et Fugère V. 2026. Drivers of walleye, brook trout, and lake trout abundance in Québec lakes. Canadian Journal of Fisheries and Aquatic Sciences, 83 : 1-14.

Paquin E et Brodeur P. 2023. Chronologie de la reproduction du doré noir du lac Saint-Pierre de 2005 à 2022. Ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs, Direction de la gestion de la faune de la Mauricie et du Centre-du-Québec, Québec, 14 p.

Pêches et Océans Canada. 2019. Enquête sur la pêche récréative au Canada 2015. Gouvernement du Canada, Ontario, 27 p.

Pêches et Océans Canada. 2025 (mis à jour le 25 septembre 2025). 2024 Info-éclair Pêches canadiennes. Consulté le 24 février 2026, https://www.dfo-mpo.gc.ca/stats/publications/fast-facts-info-eclair/2024/index-fra.html

Pêches et Océans Canada. 2026 (mis à jour le 13 février 2026). 2024 Débarquements du pêches en eau douce. Consulté le 26 février 2026, https://www.dfo-mpo.gc.ca/stats/commercial/land-debarq/freshwater-eaudouce/2024-fra.htm

Pedersen E, Miller D, Simpson G et Ross N. 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. Peer J, 7 : e6876.

Pichler M et Hartig F. 2023. Machine learning and deep learning—A review for ecologists. Methods in Ecology and Evolution, 14 : 994-1016.

Pitcher T et Lam M. 2014. Fish commoditization and the historical origins of catching fish for profit. Maritime Studies, 14.

Poesch MS, Chavarie L, Chu C, Pandit SN et Tonn W. 2016. Climate change impacts on freshwater fishes: a canadian perspective. Fisheries, 41 : 385-391.

Post JR. 2013. Resilient recreational fisheries or prone to collapse? A decade of research on the science and management of recreational fisheries. Fisheries Management and Ecology, 20 : 99-110.

Priegel GR. 1970. Reproduction and early life history of the walleye in the Lake Winnebago region. Wisconsin Department of Natural Resources, Technical bulletin 45, Wisconsin, 105 p.

Quist MC, Guy CS et Stephen JL. 2003. Recruitment dynamics of walleyes (*Stizostedion vitreum*) in Kansas reservoirs: generalities with natural systems and effects of a centrarchid predator. Canadian Journal of Fisheries and Aquatic Sciences, 60 : 830-839.

Quist MC, Guy CS, Bernot RJ et Stephen JL. 2004. Factors related to growth and survival of larval walleyes: implications for recruitment in a southern Great Plains reservoir. Fisheries Research, 67 : 215-225.

R Core Team. 2025. The R Project for Statistical Computing (version 4.5.2). Consulté le 31 octobre 2025, https://www.r-project.org/

Raabe JK, VanDeHey JA, Zentner DL, Cross TK et Sass GG. 2020. Walleye inland lake habitat: considerations for successful natural recruitment and stocking in north central North America. Lake and Reservoir Management, 36 : 335-359.

Ramsay TO, Burnett RT et Krewski D. 2003. The effect of concurvity in generalized additive models linking mortality to ambient particulate matter. Epidemiology, 14 : 18-23.

Rawson DS. 1957. The life history and ecology of the yellow walleye, *Stizostedion vitreum*, in Lac La Ronge, Saskatchewan. Transactions of the American Fisheries Society, 86 : 15-37.

Read JS, Winslow LA, Hansen GJA, Van Den Hoek J, Hanson PC, Bruce LC et Markfort CD. 2014. Simulating 2368 temperate lakes reveals weak coherence in stratification phenology. Ecological Modelling, 291 : 142-150.

Riedl C et Peter A. 2013. Timing of brown trout spawning in Alpine rivers with special consideration of egg burial depth. Ecology of Freshwater Fish, 22 : 384-397.

Roseman EF, Taylor WW, Hayes DB, Jones AL et Francis JT. 2006. Predation on walleye eggs by fish on reefs in western Lake Erie. Journal of Great Lakes Research, 32 : 415-423.

Sahrhage D et Lundbeck J. 1992. A history of fishing. Springer Berlin Heidelberg, Berlin, Heidelberg, 348 p.

Sass GG. 2025. The decline of walleye populations: an ecological tipping point? FACETS, 10 : 1-17.

Schneider JC, Copeland J et Wolgamood M. 2002. Tolerance of incubating walleye eggs to temperature fluctuation. North American Journal of Aquaculture, 64 : 75-78.

Schneider KN, Newman RM, Card V, Weisberg S et Pereira DL. 2010. Timing of walleye spawning as an indicator of climate change. Transactions of the American Fisheries Society, 139 : 1198-1210.

Scowen M, Athanasiadis IN, Bullock JM, Eigenbrod F et Willcock S. 2021. The current and future uses of machine learning in ecosystem service research. Science of The Total Environment, 799 : 149263.

Sibindi R, Mwangi R et Waititu A. 2022. A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices. Engineering Reports, 5.

Smith D, Hilling C, Welsh S et Wellman D. 2022. Differences in population characteristics and modeled response to harvest regulations in reestablished appalachian walleye populations. North American Journal of Fisheries Management, 42.

Spanou DS, Petroudi P, Dimou E, Kokkinos K et Klaoudatos D. 2024. Walleye (*Sander vitreus*, Mitchill 1818) age and sex classification using innovative supervised and unsupervised machine learning and soft computing methodologies. Fisheries Research, 275 : 107031.

Stefanik EL et Sandheinrich MB. 1999. Differences in spawning and emergence phenology between stocked and wild populations of brown trout in southwestern Wisconsin streams. North American Journal of Fisheries Management, 19 : 1112-1116.

Strobl C, Boulesteix A-L, Zeileis A et Hothorn T. 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics, 8 : 25.

Thao C, Mosindy TE et Venturelli PA. 2016. Fish community changes in Shoal Lake, Canada, following the overexploitation of a top predator. Journal of Freshwater Ecology, 31 : 443-450.

Thessen AE. 2016. Adoption of machine learning techniques in ecology and earth science. One Ecosystem, 1.

Uphoff CS, Schoenebeck CW, Hoback WW, Koupal KD et Pope KL. 2013. Degree-day accumulation influences annual variability in growth of age-0 walleye. Fisheries Research, 147 : 394-398.

Venturelli PA, Lester NP, Marshall TR et Shuter BJ. 2010. Consistent patterns of maturity and density-dependent growth among populations of walleye (*Sander vitreus*): application of the growing degree-day metric. Canadian Journal of Fisheries and Aquatic Sciences, 67 : 1057-1067.

Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O et Bairlein F. 2002. Ecological responses to recent climate change. Nature, 416 : 389-395.

Wang HY, Rutherford ES, Cook HA, Einhouse DW, Haas RC, Johnson TB, Kenyon R, Locke B et Turner MW. 2007. Movement of walleyes in Lakes Erie and St. Clair inferred from tag return and fisheries data. Transactions of the American Fisheries Society, 136 : 539-551.

Webb BW et Nobilis F. 2007. Long-term changes in river temperature and the influence of climatic and hydrological factors. Hydrological Sciences Journal, 52 : 74-85.

Webb JH et McLay HA. 1996. Variation in the time of spawning of Atlantic salmon (*Salmo salar*) and its relationship to temperature in the Aberdeenshire Dee, Scotland. Canadian Journal of Fisheries and Aquatic Sciences doi : 10.1139/f96-240.

Wright MN et Ziegler A. 2017. Ranger: a fast implementation of random forests for high dimensional data in C++ and R. Journal of Statistical Software, 77 : 1 - 17.

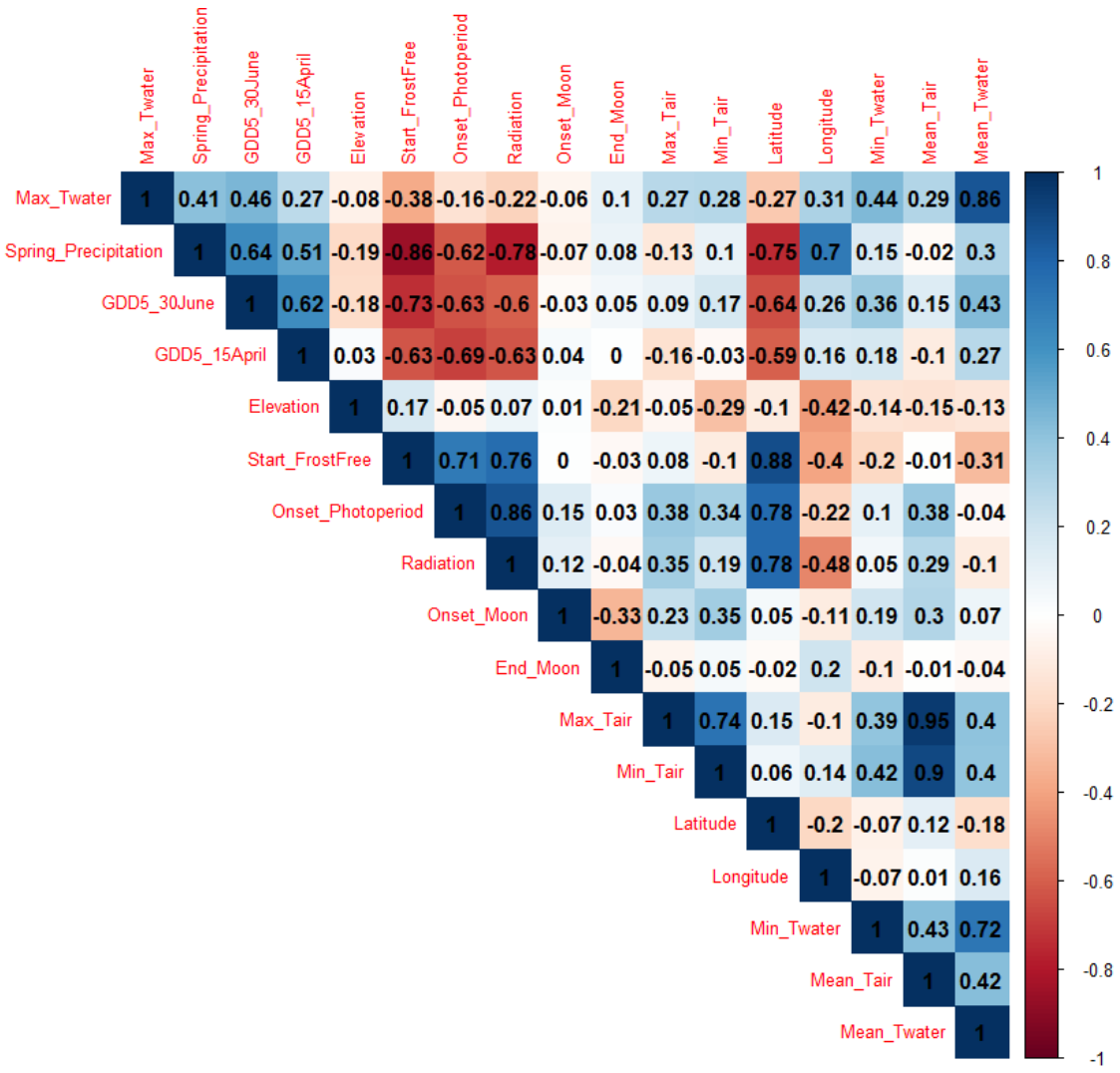
Zhao Y, Jones ML, Shuter BJ et Roseman EF. 2009. A biophysical model of Lake Erie walleye (*Sander vitreus*) explains interannual variations in recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 66 : 114-125.

ANNEXES

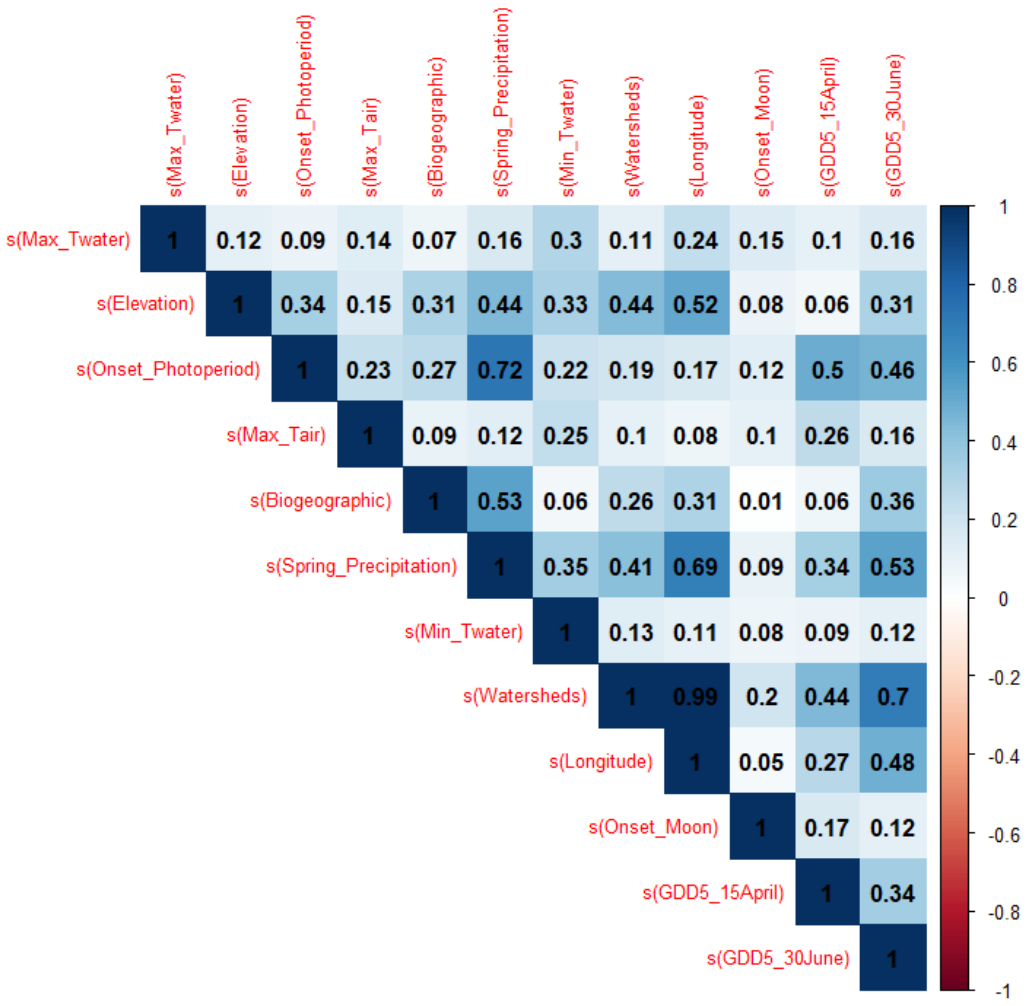
APPENDIX A. Descriptive summary of predictors in the models

Name	Description	Unit	Source
Max_Tair Min_Tair Moy_Tair	Modeled daily air temperature (max/min/mean)	Degrees Celsius (°C)	Service Info-Climat du MELCCFP
Max_Twater Min_Twater Mean_Twater	Daily water temperature (max/min/mean)	Degrees Celsius (°C)	Collected at the sampling site
Elevation	Elevation 30 seconds	Meters (m)	WorldClim
Onset_Photoperiod End_Photoperiod	Daylight duration on the first or last day of spawning	Hours (H)	getSunlightTimes function from the suncalc R package
GDD5_15April GDD5_30June	Cumulative growing degree-days above 5°C by April 15 or June 30	Degrees Celsius (°C)	Calculated from mean daily air temperature
Spring_Precipitation	Total liquid and solid precipitation from March to May	Milimeters (mm)	Ouranos
Start_FrostFree	Start of the period without at least four consecutive frost nights (minimum temperatures below 0 °C)	Day of the year (DOY)	Ouranos
Latitude Longitude	Geographic coordinates	Decimal degrees (DD)	Collected at the sampling site
Biogeographic	National freshwater biogeographic zones	-	COSEWIC
Watershelds	Multiscale watershed levels 1 and 2 in Québec (priority given to level 2)	-	
Onset_Moon End_Moon	Daily fraction of the moon illuminated at midnight (0–1, where 1 represents a full moon)	Percentage (%)	getMoonIllumination function from the suncalc R package
Radiation	Solar radiation	KJ m-2 day-1	WorldClim

APPENDIX B. Spearman correlation matrix of environmental predictors



APPENDIX C. Concurvity matrix of smooth terms



APPENDIX D. Top five GAMM models explaining spawning duration ranked by AIC

#	Formula	AIC	ΔAIC
1	s(Max_Twater) + s(Max_Tair) + Onset_Photoperiod + Elevation + s(Biogeographic, bs = "re") + s(Watersheds, bs = "re")	506,73	0,00
2	s(Max_Twater) + s(Max_Tair) + Onset_Photoperiod + Elevation + s(Biogeographic, bs = "re")	507,42	0,69
3	s(Max_Twater) + s(Max_Tair) + Onset_Photoperiod + s(Biogeographic, bs = "re")	507,93	1,21
4	s(Max_Twater) + s(Max_Tair) + Onset_Photoperiod + Elevation + Onset_Moon + s(Biogeographic, bs = "re") + s(Watersheds, bs = "re")	508,84	2,11
5	s(Max_Twater) + s(Max_Tair) + Onset_Photoperiod + Elevation + Onset_Moon + s(Biogeographic, bs = "re")	509,46	2,73

APPENDIX E. Hyperparameter tuning grid and selected values for Random Forest and XGBoost models

Random Forest (start and end of spawning)		
Hyperparameter	Range tested	Selected Value
mtry	1-6	2
trees	300-2000	2000
min_n	2-60	2
XGBoost (start and end of spawning)		
Hyperparameter	Range tested	Selected Value
trees	200-2500	1650
min_n	1-80	10
tree_depth	2-10	5
learn_rate	0,001-0,1	Not tuned
loss_reduction	1e-06-0,1	0,00316
mtry	2-7	2

APPENDIX F. Comparison of model performance for predicting spawning onset

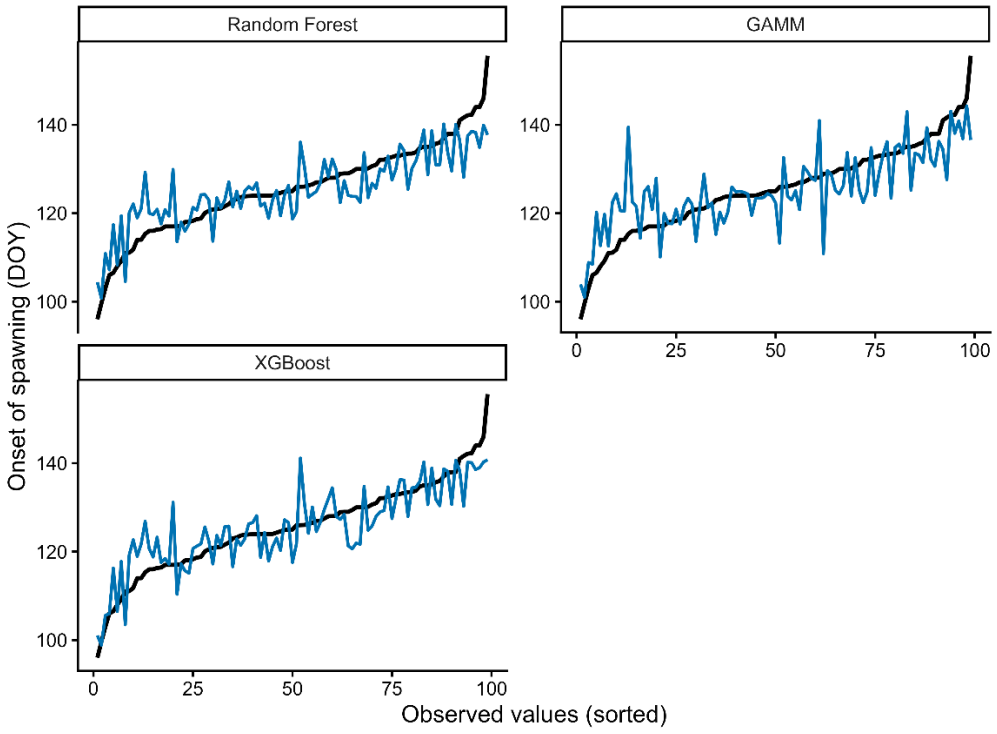


Figure A. Observed (black) vs. predicted (blue) values for spawning onset models.

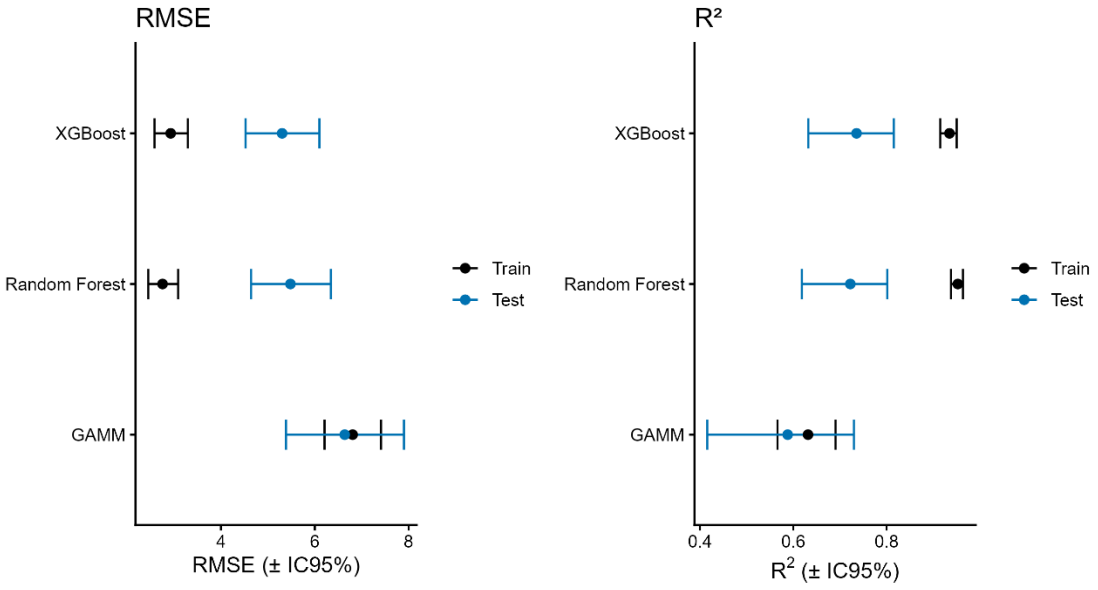


Figure B. Comparison of training (black) and test (blue) performance (RMSE, R^2) for models predicting spawning onset.

APPENDIX G. Comparison of model performance for predicting spawning end

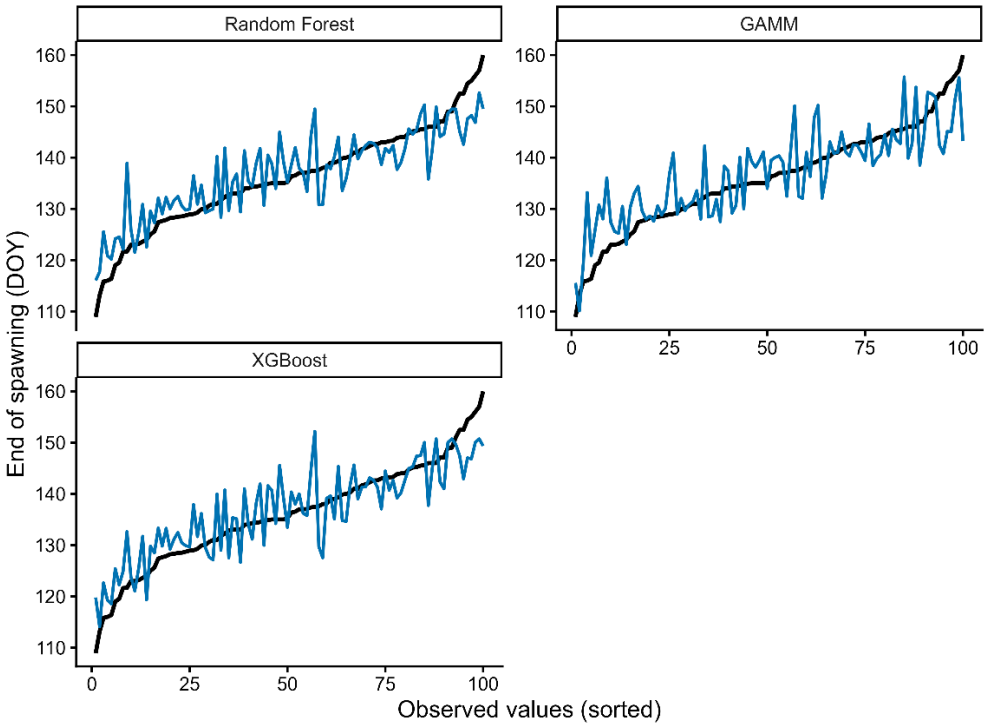


Figure C. Observed (black) vs. predicted (blue) values for spawning end models.

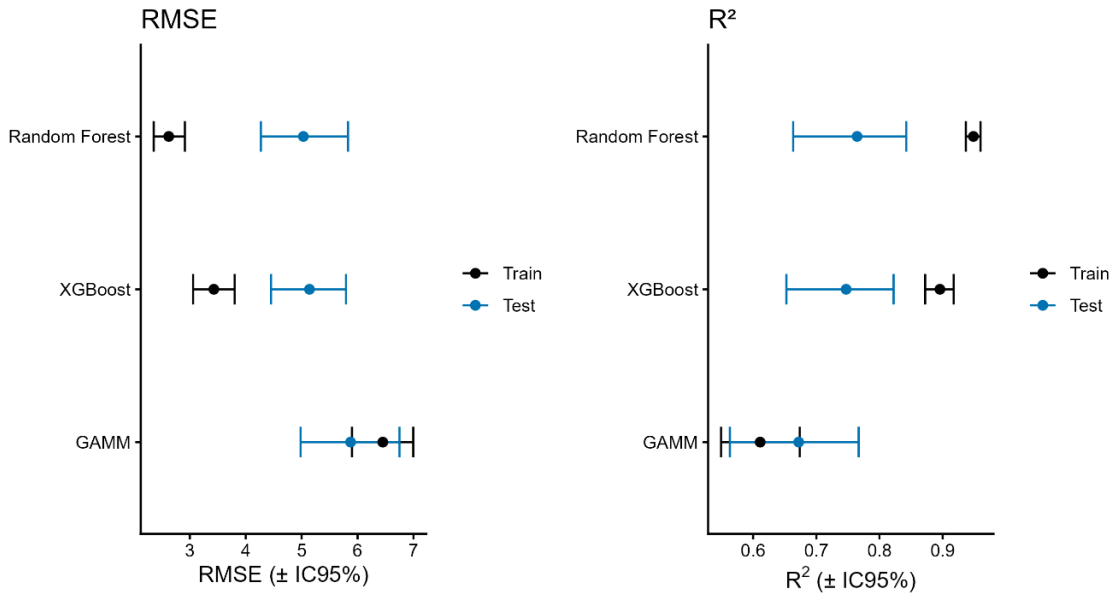


Figure D. Comparison of training (black) and test (blue) performance (RMSE, R^2) for models predicting spawning end.

APPENDIX H. Comparison of climatic conditions associated with complete observations recorded at a maximum water temperature below 15 °C or at least 15 °C

Variables	Max_Twater groups (mean ± SE)		t	df	*p-value*	*p-value* bonferroni
	<15°C	≥15°C				
GDD5_15av	6,93 ± 1,33	8,51 ± 2,31	-0,59	20,94	0,56	1
GDD5_30juin	467,03 ± 13,07	572,74 ± 22,21	-4,10	21,37	0,000494	0,002964
Onset_Photoperiod	16,26 ± 0,38	18,79 ± 0,99	1,07	16,59	0,298	1
Min_Twater	2,41 ± 0,28	4,38 ± 0,82	-2,23	15,44	0,0412	0,2472
Max_Tair	6,35 ± 0,23	7,82 ± 0,62	-2,39	15,84	0,0298	0,1788
Min_Tair	14,90 ± 0,08	14,68 ± 0,19	-2,29	14,86	0,0374	0,2244

APPENDIX I. Importance of environmental predictors in the Random Forest model predicting spawning onset and end

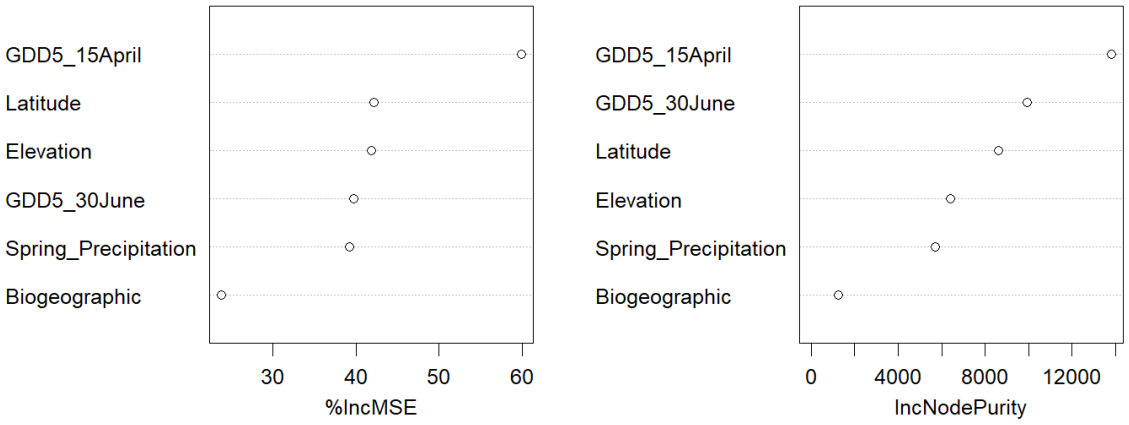


Figure E. Importance of environmental predictors for spawning onset.

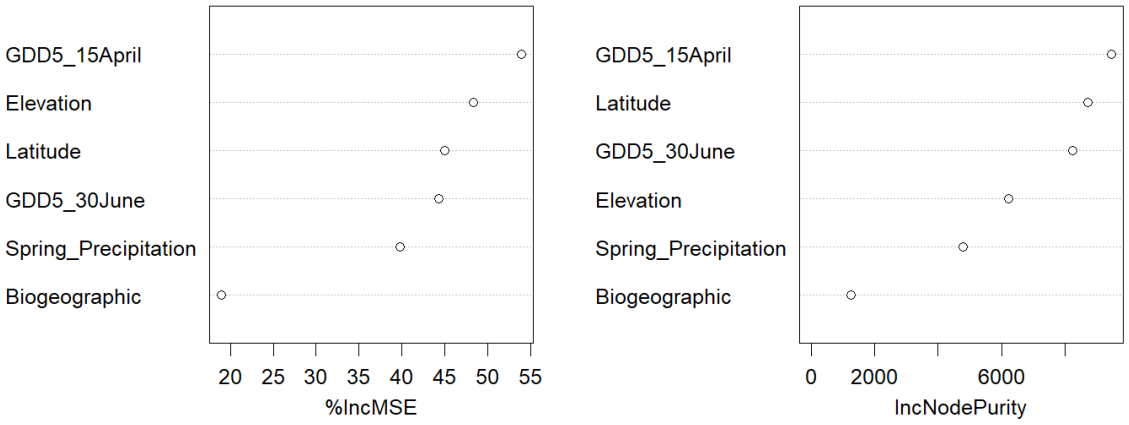


Figure F. Importance of environmental predictors for spawning end.

APPENDIX J. Effects of environmental predictors in the Random Forest model predicting spawning onset and end

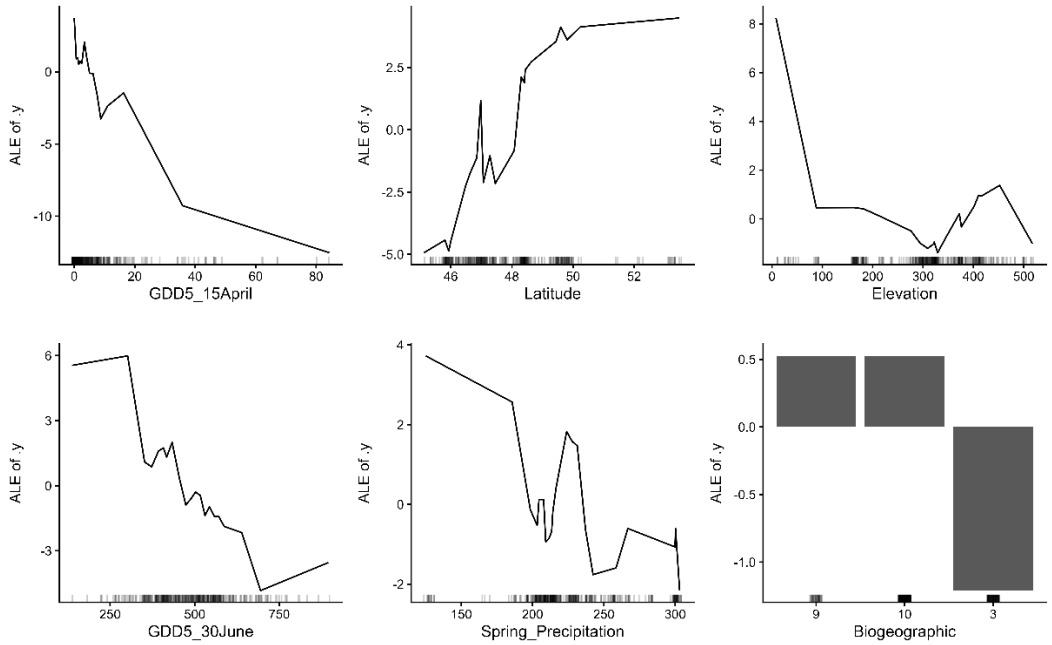


Figure G. Effects of environmental predictors for spawning onset.

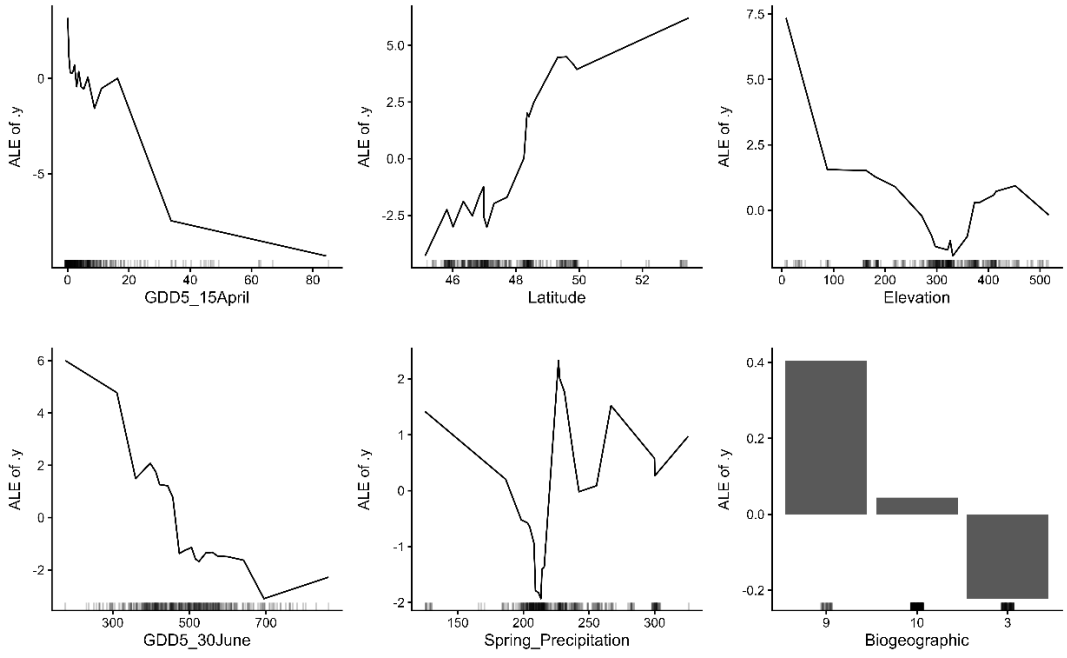


Figure H. Effects of environmental predictors for spawning end.