



SUIVI ET MODÉLISATION DU POTENTIEL HYDRIQUE DU SOL DANS UN
CONTEXTE DE STRESS CLIMATIQUES : LE CAS D'UNE ÉRABLIÈRE À
BOULEAU JAUNE À LA MARGE NORDIQUE DE SA DISTRIBUTION

MÉMOIRE PRÉSENTÉ COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN TECHNOLOGIE DE L'INFORMATION

PAR
BLANDINE COURCOT

SEPTEMBRE 2023



<https://r-libre.teluq.ca/3036>

TABLE DES MATIÈRES

LISTE DES TABLEAUX	iv
LISTE DES FIGURES	v
REMERCIEMENTS	xii
RÉSUMÉ	xiv
INTRODUCTION	1
CHAPITRE I CONTEXTE	6
1.1 La forêt comme acteur central	6
1.1.1 Quelques chiffres	6
1.1.2 Diversité des régions forestières du Canada	8
1.1.3 Rôle clé des forêts	8
1.1.4 Impacts des changements climatiques pour le Canada	12
1.1.5 Quels sont les impacts potentiels de ces changements et cycles climatiques sur les forêts?	18
1.2 Événements extrêmes et nouvelle normalité	26
1.3 Le cas des sécheresses et les conséquences sur les forêts	32
1.3.1 Impact des sécheresses sur les forêts	36
CHAPITRE II CADRE THÉORIQUE ET PROBLÉMATIQUE	40
2.1 La triade {sol, eau, potentiel hydrique du sol}	40
2.1.1 Le sol et l'eau	40
2.1.2 Le potentiel hydrique du sol (Ψ)	43
2.1.3 Mesure du potentiel hydrique du sol	46
2.2 Le stress hydrique	52
2.2.1 Une cascade d'événements	52
2.2.2 Boucles de rétroactions positives	54
2.2.3 Différents niveaux de stress hydriques	55

2.2.4	Mécanismes adaptatifs	57
2.2.5	Mémoire face aux stress hydriques	58
2.3	Angle d'étude : les séries temporelles	62
2.4	Problématique centrale et formulation des questions de recherche	63
CHAPITRE III MÉTHODOLOGIE		65
3.1	Partie expérimentale : SmartForests Canada	65
3.2	Ressources et gestion des données	70
3.3	Partie modélisation : les différentes phases du projet	71
3.3.1	Les variables d'intérêt	71
3.3.2	Sept phases d'analyse	73
CHAPITRE IV RÉSULTATS ET ANALYSE		94
4.1	Résumé des résultats obtenus et réponses aux questions de recherche	94
4.1.1	Phase 1 : préparation des données	94
4.1.2	Phase 2 : sélection des variables prédictives	107
4.1.3	Phase 3 : analyse des changements d'état du potentiel hydrique du sol	111
4.1.4	Phase 4 : exploration des séries temporelles	115
4.1.5	Phase 5 : tests et validations des modèles prédictifs ARIMAX	120
4.1.6	Phase 6 : exploration d'un autre type de modèle	135
4.2	Application : étude de périodes de sécheresse-flash entre 2017 et 2020	137
4.2.1	Potentiel hydrique et température du sol	138
4.2.2	Nouvelle stratégie d'entraînement du modèle ARIMAX	147
4.3	Ce qu'il faut retenir	151
4.4	Discussion et perspectives	153
CONCLUSION		165

LISTE DES TABLEAUX

Tableau		Page
1.1	Changements anticipés au niveau des forêts sous l'influence des changements climatiques (d'après les modélisations réalisées par Ouranos).	20
1.2	Sensibilité à la sécheresse de certaines essences, d'après Duchesne <i>et al.</i> (2020).	38
3.1	Surface basale totale (SB) en $[m^2 ha^{-1}]$ et contributions de chaque espèce d'arbre à cette surface dans chaque bloc (données de 2018).	69
3.2	Contribution moyenne de chaque espèce d'arbre à la masse totale de la litière dans chaque bloc (données de 2018). L'indice foliaire (LAI) s'exprime en $[m^2 m^{-2}]$	70
3.3	Extrait de la base de données journalières. Par défaut, Matlab affiche les données avec un format de 4 chiffres après la virgule (format <i>short</i>).	79
3.4	Description des champs de la base de données journalières.	80
3.5	Description des paramètres du modèle ARIMAX.	85
4.1	Coefficients d'asymétrie calculés pour la sapinière à érable (MW), l'érablière à bouleau (HB) et l'érablière à hêtre (HB), en 2017 et 2020.	118
4.2	Pourcentage de prédictions exactes de l'état du potentiel hydrique du sol, pour la sapinière à érable, année 2020.	130
4.3	Corrélations croisées pour les variables exogènes suivantes : température de l'air, température du sol et déficit de pression de vapeur.	131
4.4	Pourcentage de prédictions exactes de l'état du potentiel hydrique du sol, pour la sapinière à érable (année 2019).	134

LISTE DES FIGURES

Figure	Page
0.1 Réseau formé par les mots clés associés aux changements climatiques et aux données massives. Sept groupes encodés par des couleurs différentes apparaissent. Neuf mots clés ont été extraits pour plus de clarté. Figure extraite et adaptée de Sebestyén <i>et al.</i> (2021) . . .	2
0.2 Répartition des sites expérimentaux du projet SmartForests Canada. Figure extraite de Pappas <i>et al.</i> (2022).	4
1.1 Régions forestières du Canada, Ressources naturelles Canada, Service canadien des forêts (2007).	8
1.2 Composition des forêts du Canada, Ressources naturelles Canada, Service canadien des forêts (2020).	9
1.3 Émissions nettes de CO ₂ équivalent dans les forêts aménagées du Canada, entre 1990 et 2020. Reproduit à partir de Ressources naturelles Canada, <i>Comment les perturbations façonnent-elles les forêts du Canada?</i>	11
1.4 Variations projetées de la température (A) et des précipitations annuelles (B) d’ici la fin du XXI ^e siècle (horizon 2100). Ressources naturelles Canada.	13
1.5 (A) Projections des températures en °C et (B) des précipitations en mm pour la région des Laurentides (QC). La SBL se situe au sud de la région. Pour (A) et (B), la période de référence est représentée à gauche (1981-2010) et les projections sont sur la carte représentée à droite (2041-2070). Le scénario RCP8.5 a été utilisé pour permettre une comparaison avec les résultats de la figure 1.4. Pour les précipitations, l’échelle de gradient de couleurs va du rouge (950 mm) vers le bleu (1200 mm). Les résultats ont été obtenus à partir de la plateforme <i>Portraits climatiques</i> de Ouranos, version 1.3 (url : https://www.ouranos.ca/en/climate-portraits)	15
1.6 Évolution de la température moyenne annuelle à Saint-Hippolyte (QC). Position : 45.933°N, 74,016°W Trois scénarios sont testés (sur la période 2005-2100) : RCP2.6 (courbe bleue), RCP4.5 (courbe verte) et RCP8.5 (courbe rouge). La courbe noire représente les données historiques entre 1950 et 2005 (échelle de l’axe des abscisses : de 1950 à 2100, avec un pas de 10 ans). Le graphique a été généré à partir du portail de DonneesClimatiques.ca	16

1.7	(A) Répartition de l'érable rouge (données de 1971 à 2000). (B) Probabilité de présence de l'érable rouge pour 2041-2070, même échelle de probabilité que (A) et scénario RCP2.6, (C) scénario RCP8.5. Résultats obtenus à partir du modèle MaxEnt de Ressources naturelles Canada (url : http://planthardiness.gc.ca).	22
1.8	(A) Répartition du bouleau jaune (données de 1971 à 2000). (B) Probabilité de présence du bouleau jaune pour 2041-2070, même échelle de probabilité que (A) et scénario RCP2.6, (C) scénario RCP8.5. Résultats obtenus à partir du modèle MaxEnt de Ressources naturelles Canada (url : http://planthardiness.gc.ca).	23
1.9	(A) Répartition du hêtre à grandes feuilles (données de 1971 à 2000). (B) Probabilité de présence de l'hêtre pour 2041-2070, même échelle de probabilité que (A) et scénario RCP2.6, (C) scénario RCP8.5. Résultats obtenus à partir du modèle MaxEnt de Ressources naturelles Canada (url : http://planthardiness.gc.ca).	24
1.10	Différents cas de variabilité à considérer en lien avec les changements climatiques (figure extraite et adaptée de Rosenzweig et Hillel (2008), p.13)	25
1.11	Représentation graphique des changements sur une forêt saine en fonction du stress des arbres et des perturbations que subit la forêt (figure extraite de Trumbore <i>et al.</i> (2015)). Trois états sont définis (sain, instable, non-sain), se rattachant à trois états de résilience (élevé, faible, inexistant). Lorsque la résilience est élevée, le système a tendance à rester dans son état actuel. Quand celle-ci est faible, il aura tendance à basculer plus facilement vers un autre état. La zone en vert illustre une zone de stress faible. La zone orange représente la zone critique.	28
1.12	Analogie avec la boule qui roule permettant de schématiser le concept de résilience écologique, au sens de Holling (1996) (figure extraite et adaptée de Capdevila <i>et al.</i> (2021)).	30
1.13	Suivi des sécheresses au Canada : situation en juin 2020 (ressource accessible en ligne) et comparaison entre juin 2020 et juin 2019 (d'après l'outil de surveillance des sécheresses au Canada, Gouvernement du Canada).	35
2.1	Répartition des eaux de précipitations dans le sol (figure inspirée de Duchaufour <i>et al.</i> (2020), p.101).	42
2.2	Relation sol-eau et phénomène d'hystérésis (figure extraite et adaptée de Or et Wraith (2001), p.78). Le potentiel hydrique est considéré comme étant positif, tel que décrit dans le référentiel que nous avons choisi.	47

2.3	Sonde granulométrique matricielle de la marque Watermark, extraite de Irmak <i>et al.</i> (2016)	48
2.4	Potentiel hydrique du sol en fonction de la résistance, pour deux valeurs de températures fixées. En bleu T=24 °C et en vert T=10 °C (graphique réalisé avec Matlab).	49
2.5	Actions combinées du climat, du sol et de la végétation sur le déficit hydrique du sol (A) et cascade d'événements menant à un état de stress hydrique (B). Figure adaptée de Rodríguez-Iturbe et Porporato (2005), p.87. Traduction libre.	53
2.6	Liens entre la demande évaporative, l'évapotranspiration et le déficit hydrique (A) et description des mécanismes reposant sur des boucles de rétroaction positive (B). Figure adaptée de Pendergrass <i>et al.</i> (2020) et issue de Lukas <i>et al.</i> (2007).	54
2.7	Mise en évidence de différents niveaux de stress hydrique, d'après Kögler et Söffker (2020).	56
2.8	Représentation des concepts centraux de notre étude au moyen d'un organigramme.	64
3.1	Emplacement de la SBL (QC, Canada), d'après Bélanger <i>et al.</i> (2021). Sur la carte de gauche, les distributions de l'érable à sucre (<i>Acer saccharum</i>) et du hêtre à grandes feuilles (<i>Fagus grandifolia</i>) sont représentées en orange et en gris, respectivement. La limite nordique de la forêt tempérée est en rouge, sur la carte de droite.	66
3.2	Organisation spatiale des secteurs (site 1 à 3) et des blocs d'échantillonnage à la SBL (bloc 1 à 8). Les 32 stations sont représentées par des losanges. Crédit : Simon Lebel-Desrosiers.	68
3.3	Premier exemple d'une série temporelle : température du sol mesurée entre le 01 mai 2020 minuit et le 02 mai 2020 minuit inclus (moyenne par heure). La série est constituée de 25 points temporels.	74
3.4	Deuxième exemple d'une série temporelle : température du sol mesurée entre le 01 mai 2020 et le 15 mai 2020 (moyenne journalière). La série est constituée de 15 points temporels.	74
3.5	Illustration simple d'une chaîne de Markov à temps discret avec la matrice des probabilités de transition pour passer d'un état à un autre.	82
3.6	Interprétation de la matrice de confusion : exemple pour la prédiction de l'état du potentiel hydrique du sol.	88
3.7	Calcul des différentes métriques caractérisant la performance du modèle.	89
3.8	Représentation du réseau utilisé pour l'approche NARX.	90
3.9	Stratégie mise en place dans l'étude des périodes de sécheresse-flash.	92

3.10	Résumé de la méthodologie mise en place, proposé en anglais (en vue d’une insertion dans une publication éventuelle).	93
4.1	Répartition initiale des observations par peuplement étudié, entre 2017 et 2020, avant la première phase méthodologique de nettoyage des données.	95
4.2	Potentiel hydrique du sol moyenné (par heure) sur les deux sondes indépendantes (nommées C et D), avec les incertitudes associées (station01, année 2020). Pour chaque point, il s’agit de la moyenne des deux sondes.	96
4.3	Potentiel hydrique du sol mesuré par la sonde C (en noir) et la sonde D (en vert) pour la station01 en 2020.	97
4.4	Potentiel hydrique du sol mesuré en [kPa] à différentes heures de la journée, à l’échelle de la station01. Le potentiel a été moyenné sur les deux sondes de mesure (C et D). L’axe des abscisses représente un point temporel par jour, entre juin et octobre 2020.	98
4.5	Mesures avant régularisation (en bleu) et après régularisation (en orange, trait en pointillé). Données échantillonnées sur les deux sondes de température et de potentiel hydrique du sol, sur une heure (exemple pour la station01, année 2020).	100
4.6	Exemple du potentiel hydrique du sol ajusté à la température du sol (peuplement MW, année 2020).	101
4.7	Comparaison entre l’humidité relative générée par BioSIM (en orange) et l’humidité relative mesurée par la station météorologique de la SBL à 2 m et à 20 m de hauteur (en bleu) en 2020. . .	102
4.8	Comparaison entre la température de l’air générée par BioSIM (en orange) et la température de l’air mesurée par la station météorologique de la SBL à 2 m et à 20 m de hauteur (en bleu) en 2020.	103
4.9	Évolution du contenu volumétrique en eau dans le sol [%] , à différentes profondeurs (10 cm, 21 cm, 40 cm, 58 cm) et précipitations générées avec BioSIM (année de référence : 2020).	104
4.10	Comparaison des précipitations [mm] entre 2017 et 2020 pour les mois de mai à octobre.	105
4.11	Évolution temporelle du potentiel hydrique du sol moyenné sur une journée, entre 2017 et 2020, pour la sapinière à érable (MW), l’érablière à bouleau (HW) et l’érablière à hêtre (HB).	106
4.12	Importance des variables exogènes dans la prédiction du potentiel hydrique du sol pour les trois peuplements étudiés.	108
4.13	Probabilités de transition d’état (notée P) et temps attendu pour atteindre chaque état (en jours) pour la sapinière à érable (MW) . Comparaison entre 2019 et 2020.	113

4.14	Probabilités de transition d'état pour la sapinière à érable (MW) et pour l'érablière à hêtre (HB) en 2020.	114
4.15	Comparaison des temps attendus pour atteindre les états <i>Medium</i> et <i>High</i> du potentiel hydrique du sol, pour la sapinière à érable (MW) et l'érablière à hêtre (HB) en 2020.	114
4.16	Représentation de l'autocorrélation pour la sapinière à érable (MW) de 2017 à 2020. En abscisse sont représentés les décalages en jours (entre -15 jours et 15 jours) et en ordonnées, la valeur des coefficients d'autocorrélation (non normalisés).	116
4.17	Différence entre les coefficients d'autocorrélation obtenus en 2020 et en 2017, pour les trois peuplements étudiés (MW, HW et HB). En abscisse sont représentés les décalages en jours (entre -15 jours et 15 jours) et en ordonnées, la différence entre les coefficients.	117
4.18	Schéma résumant les différentes corrélations-croisées entre le potentiel hydrique du sol et les variables exogènes du modèle. Les calculs ont été menés sur les données de 2020, pour chaque peuplement considéré. Le coefficient maximal de corrélation-croisée est représenté par une couleur plus foncée.	119
4.19	Modèle autorégressif d'ordre 1 : AR(1).	121
4.20	Modèle de moyenne mobile d'ordre 1 : MA(1).	122
4.21	Cartes thermiques du pourcentage de prédictions exactes, pour deux peuplements. En abscisse, l'année de prédiction et en ordonnées, l'année d'estimation du modèle. Les pourcentages sont représentés à l'aide d'une échelle de gris et les valeurs obtenues sont précisées en rouge.	124
4.22	Modèle ARIMAX pour l'érablière à hêtre (HB) entraîné sur les données de 2020 et prédictions des données de 2019. Différentes tailles d'horizon prédictif sont testés : 60 jours, 160 jours et 60 jours mais translatés de 60 jours.	126
4.23	Pourcentage d'exactitude des prédictions en fonction de la taille de la fenêtre de prédiction. Test réalisé sur la sapinière (prédiction 2019).	127
4.24	Score-F1 en fonction de la taille de la fenêtre de prédiction (en jours). Test réalisé sur la sapinière (prédiction 2019).	127
4.25	Influence de la taille (en nombre de jours) de la fenêtre d'initialisation du modèle : prédictions réalisées pour 2019 pour la sapinière à érable (MW). Deux horizons prédictifs ont été testés : du jour 31 au jour 95 (en gris) et du jour 51 au jour 115 (en orange).	128
4.26	Modèle ARIMAX pour la sapinière à érable entraîné sur les données de 2018 ou 2019 et prédiction du potentiel hydrique mesuré en 2020 (référence en bleu).	129

4.27	Modèle ARIMAX pour la sapinière à érable entraîné sur les données de 2018 ou 2020 et prédiction du potentiel hydrique de 2019. Les résidus obtenus pour le cas (b) sont représentés avec le même code couleur.	133
4.28	Ajustement linéaire des données observées (2019) en fonction des données prédites par le modèle ARIMAX(14,1,4) de la sapinière à érable (troisième fenêtre de prédiction, modèle estimé en 2020). .	134
4.29	Modèle ARIMAX pour la sapinière entraîné sur les données de 2020 et prédiction du potentiel hydrique de 2019 (fenêtres de 40 jours). .	134
4.30	Résultat obtenu pour le modèle NARX, tel que décrit dans la partie 3.	136
4.31	Modèle NARX entraîné sur toutes les données de potentiel hydrique mesurées en 2020 (sapinière à érable), sauf sur les 20 derniers points temporels qui ont servi à la validation.	136
4.32	Première sécheresse-flash de 2020 (mai) : potentiel hydrique et température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune). . .	142
4.33	Deuxième sécheresse-flash de 2020 (juin) : potentiel hydrique et température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune). . .	143
4.34	Troisième sécheresse-flash de 2020 (septembre) : potentiel hydrique et température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune).	144
4.35	Évolution de la température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune).	145
4.36	Coefficients de corrélation-croisées (normalisés) entre le potentiel hydrique et la température du sol, calculés en utilisant (a) toutes les données de 2020 de l'érablière à hêtre et (b) les données extraites de la deuxième sécheresse-flash de 2020, pour HB. En abscisse sont représentés les décalages ou <i>lag</i> (en jours) et en ordonnées les coefficients de corrélation-croisée normalisés.	146
4.37	Arbre de décision obtenu pour la sapinière à érable (MW), deuxième sécheresse de 2020. Les températures sont en [°C].	147
4.38	Arbre de décision obtenu pour l'érablière à hêtre (HB), deuxième sécheresse de 2020. Les températures sont en [°C].	147
4.39	Concaténation des données du potentiel hydrique (en bleu) et de la température du sol (en pointillé orange) durant les sécheresses-flash de 2017 à 2020, en considérant 16 jours avant et après chaque période. L'axe des abscisses représente le temps fictif en jours. . .	149

4.40	Modèle ARIMAX(14,1,4) pour la sapinière entraîné sur les données concaténées des sécheresses-flash de 2017 à 2019 et prédiction du potentiel hydrique de 2020.	150
4.41	Nouveau modèle ARIMAX(6,1,1) pour la sapinière entraîné sur les données concaténées des sécheresses-flash de 2017 à 2019 et prédiction du potentiel hydrique de 2020.	150
4.42	Trois scénarios synthétiques pour générer les variables exogènes, d'après Ficklin et Novick (2017).	161
4.43	L'information totale d'une observation (cercle noir en B) est composée de l'information nouvelle et de l'information redondante. La prédictibilité réalisée est le degré d'exploitation de l'information active de la série temporelle par le modèle. Le modèle ARIMA (cercle en vert) et le modèle EDM (cercle en bleu) exploitent davantage cette information active et donnent un RMSE plus faible (en C). Figure extraite de Pennekamp <i>et al.</i> (2019).	164

REMERCIEMENTS

C'est dans une perspective d'éducation tout au long de la vie que j'ai entrepris en novembre 2021 ce programme de maîtrise en technologie de l'information, profil science des données. Pouvoir continuer à me former, me confronter à de nouvelles façons d'appréhender les problèmes, devoir me remettre en question dans ma démarche scientifique et de recherche, rencontrer des personnes passionnées et passionnantes, c'est une opportunité d'apprentissage qui fut transformationnelle, au sens de l'andragogue Jack Mezirow.

Cette opportunité, je la dois au Pr Nicolas Bélanger, mon co-directeur, qui m'a encouragée et soutenue dans ma démarche initiale. Il m'a permis de rencontrer le Pr Daniel Lemire qui a également accepté de m'accompagner dans mon cheminement à titre de directeur.

Je tiens à les remercier sincèrement pour leur immense générosité dans leur partage permanent de connaissances et d'expertises qui m'a aidée à explorer deux domaines : la science des données et les sciences de l'environnement. Ce fut un immense privilège de les avoir comme mentors et je leur en suis très reconnaissante. Je les remercie également pour leur soutien financier à travers leurs organismes de subvention, en particulier pour le Pr Bélanger : le *Conseil de recherches en sciences naturelles et en génie du Canada*, programme de subvention à la découverte (RGPIN 2015-03699 et 2020-04931), les *Fonds des leaders John-R.-Evans* (35370), la *Fondation canadienne pour l'innovation* (36014) ; pour le Pr Lemire : le *Conseil de recherches en sciences naturelles et en génie du Canada* (RGPIN-2017-03910).

Le DOT-Lab est un endroit dynamique et bienveillant, et je remercie les professeur-es, les étudiant-es et les personnes professionnelles qui le composent !

En plus d’abriter des laboratoires de recherche de grandes qualités, humaine et scientifique, l’Université TÉLUQ a à coeur de soutenir les apprenant.es aux parcours de vie variés qui ne sont pas tous linéaires. Ce fut ainsi agréable de me sentir soutenue dans mes apprentissages.

Je remercie également chaleureusement la Pre Gisèle Trudel, titulaire de la Chaire de recherche MÉDIANE, et Marc-André Cossette, artiste et candidat au doctorat à l’Université Concordia, pour m’avoir permis de voir les données de *SmartForests* avec un regard différent à travers le projet *Devenir-Hêtre* présenté en mai 2023 à la Fondation Grantham pour l’art et l’environnement. Nos conversations furent très inspirantes et elles ont nourri mes réflexions.

Je voudrais exprimer ma gratitude au Pr Wassim Bouachir (président du jury, Université Téléq) et au Pr François Courchesne (évaluateur externe, Université de Montréal) pour leur lecture attentive et leurs commentaires.

Enfin, je remercie ma famille pour leur soutien au quotidien.

Il me reste encore beaucoup de choses à explorer, mais cette expérience m’a permis de construire de solides fondations à l’interface entre deux sciences qui s’enrichissent mutuellement.

RÉSUMÉ

Dans un contexte de changements climatiques globaux, une question a guidé notre étude à l'interface entre la science des données et les sciences de l'environnement : quelle serait l'évolution temporelle de notre variable d'intérêt - le potentiel hydrique du sol - dans une future nouvelle normalité climatique des érablières à bouleau jaune à la marge nordique de leur distribution ?

Notre questionnement a eu pour cadre expérimental le projet pancanadien *Smart-Forests*, et en particulier le site de la Station de biologie des Laurentides (SBL) qui a permis de cibler trois peuplements : une sapinière à érable (MW), une érablière à bouleau (HW) et une érablière à hêtre (HB).

En réfléchissant à ce glissement vers une nouvelle normalité climatique, nous avons explicité la notion de *tendance* et d'*événements extrêmes* en identifiant également comment ce basculement vers un nouvel état de fond pouvait être caractérisé. Ceci nous a alors amené à clarifier la notion de *résilience écologique*, i.e. la capacité à résister aux changements. Ainsi, dans un contexte de sécheresse à déficit de précipitations nous nous sommes demandés si certains peuplements étaient plus résilients que d'autres.

En exploitant la temporalité des millions de données qui ont été collectées à la SBL entre 2017 et 2020, nous avons développé et testé un modèle prédictif du potentiel hydrique du sol de type autorégressif à moyenne mobile qui intègre les variables météorologiques (ARIMAX). De plus, pour les trois peuplements d'intérêt, nous avons analysé les changements de régime de leur potentiel hydrique en exploitant les chaînes de Markov à temps discret.

À l'issue de ce projet, nous proposons donc un modèle novateur qui permet de prédire l'évolution du potentiel hydrique du sol avec une très bonne exactitude sur un horizon de taille maximale de 60 jours, grâce à la connaissance des variables météorologiques antérieures. Ce modèle a été adapté à l'étude des sécheresse-flash durant lesquelles les arbres sont soumis à des stress hydriques.

Pendant ces périodes d'assèchement, nous mettons en évidence le comportement spécifique de l'érablière à hêtre qui montre une meilleure régulation de la température de son sol et un maintien de son potentiel hydrique autour de valeurs moins élevées que celles observées pour la sapinière et l'érablière à bouleau.

Enfin, nos analyses suggèrent la tendance d'un rapprochement plus imminent de la sapinière à érable vers un changement de régime hydrique.

L'utilisation originale des séries temporelles pour développer des modèles prédictifs en écologie forestière s'avère ainsi prometteuse.

INTRODUCTION

Ce projet s'insère tout d'abord dans une réalité et un futur climatiques étudiés : durant son sixième cycle d'observations, entre octobre 2018 et septembre 2019, le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) a estimé que les activités humaines ont provoqué un réchauffement planétaire d'environ 1°C au-dessus des niveaux qualifiés de pré-industriels, avec une fourchette probable allant de 0.8°C à 1.2°C (Masson-Delmotte *et al.*, 2018). Leur rapport souligne également qu'il est probable que le réchauffement planétaire atteindra 1.5°C entre 2030 et 2052 s'il continue d'augmenter au rythme actuel, tel que mesuré pendant ce cycle d'observations (p.6). Durant la période comprise entre 2006 et 2015, 20 à 40 % de la population mondiale a déjà été confronté à un réchauffement de plus de 1.5°C par rapport à la période de référence, et ce pendant au moins une saison. Des impacts de ces changements climatiques sont déjà observables sur les écosystèmes, au niveau de leur structure, des espèces qui y évoluent, mais aussi sur la temporalité des processus phénologiques qui s'y déroulent. En définissant des motifs de préoccupations (MdP), les experts du GIEC ont ainsi pu leur associer des niveaux de risques en fonction des variations de la température moyenne à la surface du globe, toujours par rapport aux niveaux pré-industriels. Pour une augmentation de 1.5°C , il existe un risque élevé de voir une augmentation de l'incidence des phénomènes météorologiques qualifiés d'extrêmes (MdP2), et un risque modéré à élevé d'assister à une augmentation du taux de mortalité des arbres, selon le scénario de forçage radiatif considéré.

Ce projet s'insère ensuite dans une tendance : l'utilisation des données massives (*big data*) en sciences de l'environnement.

Pour illustrer cette tendance, Sebestyén *et al.* (2021) ont étudié 442 articles publiés entre 2012 et 2020 et accessibles sur la base de données Scopus, en associant les mots clés *climate change* et *big data*. En utilisant une analyse par clusters des thèmes abordés par ces articles, les auteurs ont mis en évidence sept groupes de sujets en lien avec les changements climatiques et les méthodes associées aux données massives (figure 0.1).

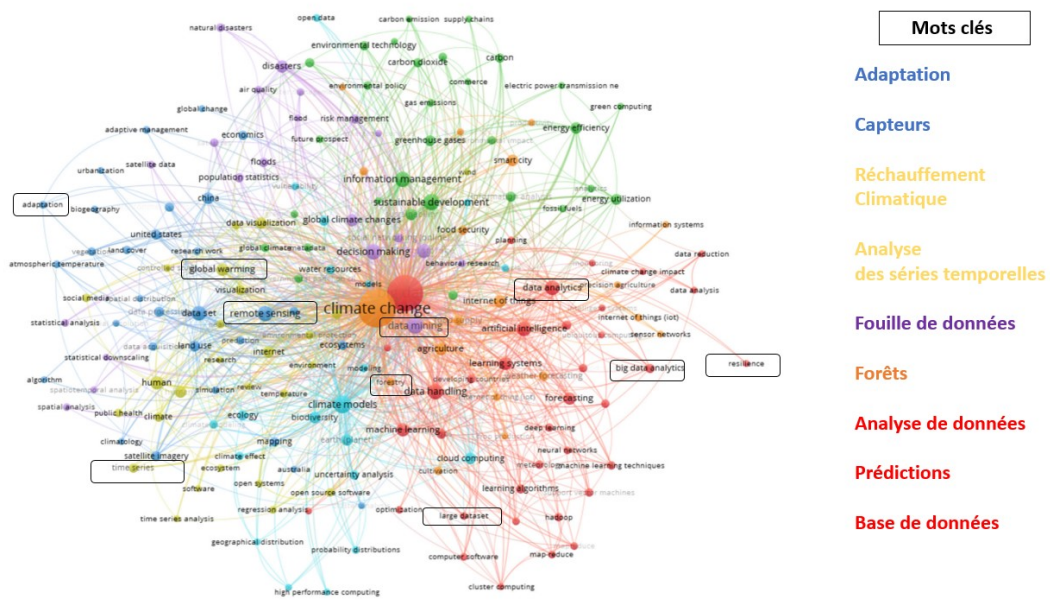


FIGURE 0.1 : Réseau formé par les mots clés associés aux changements climatiques et aux données massives. Sept groupes encodés par des couleurs différentes apparaissent. Neuf mots clés ont été extraits pour plus de clarté. Figure extraite et adaptée de Sebestyén *et al.* (2021)

De nombreux thèmes mis en évidence par leur analyse font écho à ceux que nous aborderons : l'origine des données (cluster bleu et cluster orange), les méthodes utilisées (cluster rouge et cluster violet) et le cadre choisi (cluster jaune). Deux aspects qui ressortent de la figure 0.1 sont la diversité des méthodes et la complexité des liens entre les données massives et les sciences de l'environnement, formant alors un domaine prismatique aux facettes multiples. Cela implique donc la nécessité de choisir parmi de nombreux angles d'étude possibles.

Pour naviguer à travers ce réseau, et permettre la rencontre de ces deux sciences, un cadre d'analyse a été pensé, en s'appuyant sur quatre piliers : l'observation, la compréhension, la prédiction et l'optimisation (Hassani *et al.*, 2019). Un plan de route a également été proposé pour mieux baliser cette approche transdisciplinaire et « pour ne pas sombrer au milieu des données » (Blair *et al.*, 2019; Blair, 2020; Blair *et al.*, 2021). Dix défis sont soulevés par Blair *et al.* (2019) et ils pourraient être englobés dans les enjeux suivants : défi lié à la nature des données environnementales (incertitudes, échelles spatiales et temporelles); défi lié à la complexité des processus étudiés (boucles rétroactives, événements extrêmes); défi face aux méthodes existantes et aux modèles hybrides à développer (reposant sur des données); et défi lié à la collaboration entre disciplines.

Ce projet s'insère enfin dans une vision : *SmartForests Canada* (Pappas *et al.*, 2022). Cette expérimentation pancanadienne a été mise en place pour répondre à cette réalité climatique évoquée plus haut, en exploitant ce que la technologie peut offrir.

La figure 0.2 présente la répartition des sites dédiés au projet des forêts connectées, à travers le Canada. En rouge, le site de la Station de biologie des Laurentides (SBL) qui sera l'épicentre de notre travail. La particularité de SmartForests est de proposer une approche granulaire : à l'échelle des peuplements d'Amérique du Nord, mais aussi à l'échelle individuelle de l'arbre, avec des informations recueillies au niveau des feuilles, du tronc, des racines et du sol. La collecte des observations a donc la particularité de se faire à différentes échelles spatiales, mais aussi temporelles, en s'étalant sur le long terme.

Trois objectifs principaux de SmartForests ont été définis¹. Le premier objectif consiste en la mise en place d'un réseau de surveillance forestière de haute précision, à travers un gradient de types de forêts et de conditions environnementales, afin

1. <https://smartforest.uqam.ca>

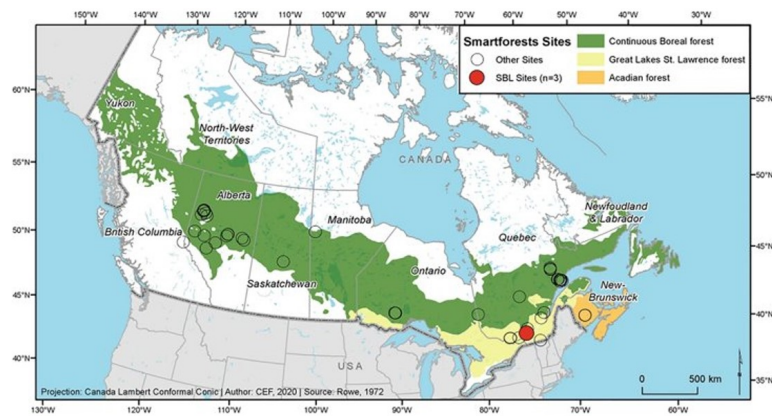


FIGURE 0.2 : Répartition des sites expérimentaux du projet SmartForests Canada. Figure extraite de Pappas *et al.* (2022).

d'examiner les effets des changements globaux sur la dynamique complexe des forêts. Le deuxième objectif vise à synthétiser les données multivariées collectées, dans le but d'évaluer le fonctionnement et la résilience des écosystèmes ciblés. Le troisième objectif est l'assimilation des données pour aider à la compréhension des processus. Cette compréhension permettra alors de guider le développement de la prochaine génération de modèles de simulation de croissance forestière et éclairer l'élaboration de politiques d'aménagement et de gestion des forêts.

Notre projet de recherche se situe dans le scope des objectifs (2) et (3). Il s'appuie sur l'obtention de données, massives, de qualité.

C'est donc aussi, et surtout, à la suite d'un travail expérimental approfondi mené sur le terrain par l'équipe du Pr Bélanger, que ce projet s'insère, et sur lequel il repose.

Dans la mise en place et le déroulement de notre projet, un des défis, comme le soulignait Blair *et al.* (2019), est d'évoluer dans la zone commune à la science des données et à l'écologie. Par *écologie*, nous entendons, dans le cadre de notre projet,

les domaines ayant trait à la science de la forêt, la science du sol, et l'hydrologie. Quant au terme *science des données*, il se réfère à la science qui a pour but de construire du sens à partir de données complexes. Une *donnée* fait référence à une mesure, une information dont l'origine, dans notre contexte d'étude, provient d'un très grand nombre de capteurs installés, nous le verrons, à des endroits spécifiques de la SBL.

Le travail qui sera présenté a été réalisé dans le cadre d'une maîtrise en technologie de l'information, avec une dominante en science des données. L'accent sera donc mis sur les méthodes utilisées et exploitées, mais dans le but de répondre à des questions d'ordre environnemental.

La question centrale, qui nous a guidé, est la suivante : *Comment modéliser le potentiel hydrique du sol dans un contexte de stress climatiques, et en particulier pour l'érablière ?*

Avant d'arriver au *comment*, nous expliquerons le *pourquoi* en détaillant le contexte du projet qui a pour acteur central la forêt, de plus en plus contrainte à des événements extrêmes, et en particulier les sécheresses. Cette mise en contexte amènera les cadres théoriques de notre étude. Tout d'abord, la triade {sol, eau, potentiel hydrique du sol} sera définie, en clarifiant notre variable d'intérêt - le potentiel hydrique du sol. Nous affinerons ensuite ce que nous entendons par stress climatiques, en nous limitant spécifiquement aux stress hydriques. Nous montrerons alors la nécessité de considérer dans notre modèle *le temps* comme une variable explicite, nous amenant à choisir les séries temporelles comme angle d'étude. Ce choix méthodologique sera justifié et la démarche mise en place sera détaillée.

Nous présenterons également l'étude qui vise les sécheresses dites flash. Cette étude repose sur les analyses développées dans le cadre du projet et permet d'illustrer comment la science des données et les sciences environnementales peuvent interagir.

CHAPITRE I

CONTEXTE

1.1 La forêt comme acteur central

1.1.1 Quelques chiffres

Pour commencer, attardons-nous sur une définition : qu'appelons-nous *une forêt* ? Selon l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), une forêt est « une terre occupant une superficie de plus de 0.5 hectare avec des arbres atteignant une hauteur supérieure à 5 mètres et un couvert forestier de plus de 10 pour cent, ou avec des arbres capables d'atteindre ces seuils *in situ* » (Ressources naturelles Canada, 2020).

À l'échelle mondiale, les forêts couvrent 31 % de la superficie terrestre, la moitié étant relativement intacte et plus d'un tiers est composé de forêts primaires, c'est-à-dire n'ayant jamais subi d'interventions humaines (FAO et PNUE, 2020).

À l'échelle du Canada, au-delà de cette définition, les forêts sont au cœur de l'identité du pays. Elles représentent à la fois un pilier écologique, une ressource économique¹ et un joyau culturel. Omniprésentes, les terres forestières occupent près de 38 % de la superficie du pays – ce qui, avec 362 millions d'hectares, compte pour 9 % des forêts mondiales – abritant une grande diversité de plantes, d'animaux et de microorganismes tous tributaires de cet habitat forestier (Ressources naturelles

1. En 2019, le secteur forestier représentait 1.1 % du PIB canadien, selon le rapport annuel 2020 de Ressources naturelles Canada (p.51).

Canada, 2022b). Ajoutons également à ces chiffres les arbres qui ne répondent pas à la définition de forêt et qui néanmoins constituent le paysage non forestier du Canada avec plus de 50 millions d’hectares d’arbres hors forêt. Plus de 90 % des forêts du pays se trouvent sur des terres publiques. Il est intéressant également de souligner que la superficie forestière du Canada demeure assez stable. Depuis 1990, moins de 1 % des terres forestières du pays ont été converties à un usage non forestier, et les secteurs qui contribuent le plus au déboisement dans le pays sont l’agriculture et l’extraction minière, pétrolière et gazière (Ressources naturelles Canada, 2020). Le taux de déboisement était d’environ 34 300 hectares en 2018, mais il faut préciser que ce taux ne considère pas les zones d’exploitation forestière. Bien que cette superficie soit à mettre en perspective avec les 13 millions d’hectares de forêts qui disparaissent chaque année dans le monde, dont 53 % en Amazonie², les chiffres avancés par le Canada sont néanmoins à nuancer, car en différenciant la déforestation de l’exploitation forestière, le pays est moins transparent sur ses pratiques forestières réelles. C’est ainsi que l’organisme américain *National Resources Defense Council* (NRDC) souligne cette ambiguïté dans les données, et en s’appuyant sur les analyses de *Global Forest Watch*, avance que les coupes de forêts primaires et intactes au Canada s’élèveraient en 2021 à 550 000 hectares de boisés, plaçant alors le pays derrière la Russie et le Brésil (Sims, M., Potapov, P., Goldman, L., 2022). Cependant, dans un contexte d’exploitation forestière, les sites sont généralement reboisés ; une réalité que les données de *Global Forest Watch* ne reflètent pas. En 2020, 426 000 hectares de terres forestières provinciales au Canada ont été régénérés (Ressources naturelles Canada, 2022b).

2. Source : <https://envol-vert.org>

1.1.2 Diversité des régions forestières du Canada

D'après la figure 1.1, 77% des forêts du pays se trouvent dans la zone boréale, mais d'autres types de couverts se répartissent sur le territoire.

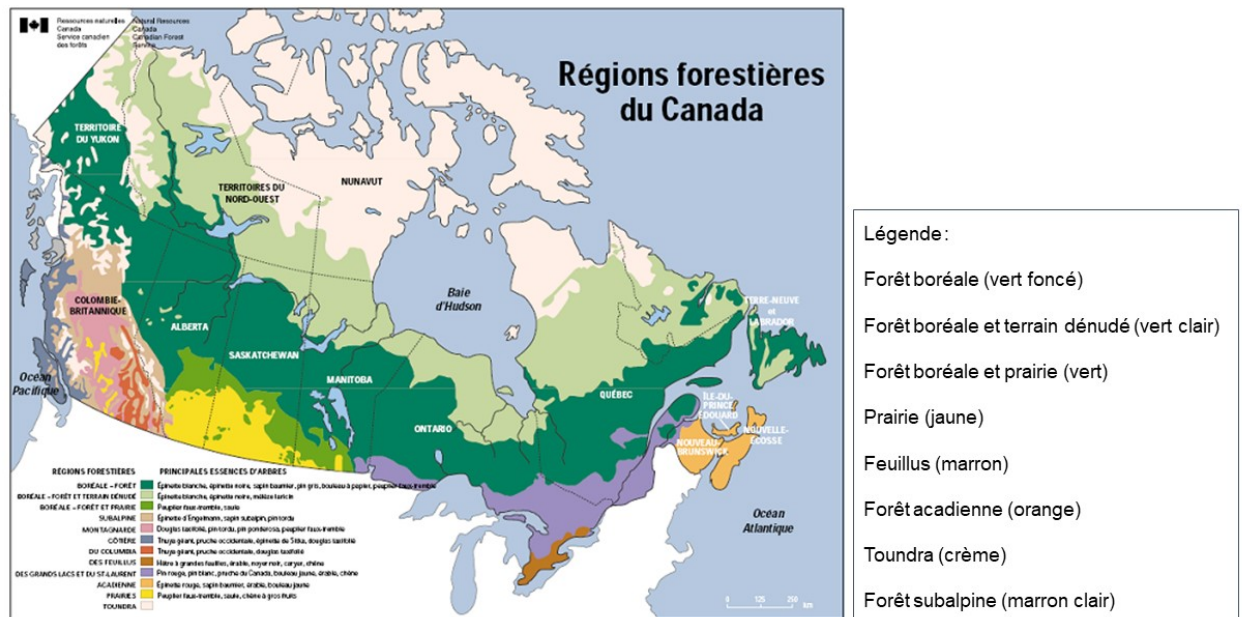


FIGURE 1.1 : Régions forestières du Canada, Ressources naturelles Canada, Service canadien des forêts (2007).

La diversité des régions forestières implique donc une grande variété d'essences, comme l'illustre la figure 1.2. Plus de la moitié est composée de conifères, comme les épinettes ou les sapins, mais se trouvent également des types forestiers mixtes (conifères et feuillus) ou uniquement de feuillus.

1.1.3 Rôle clé des forêts

Les forêts ont un rôle central au niveau de la biodiversité. Elles permettent notamment la purification de l'air et de l'eau, et elles aident à contrôler les inondations.

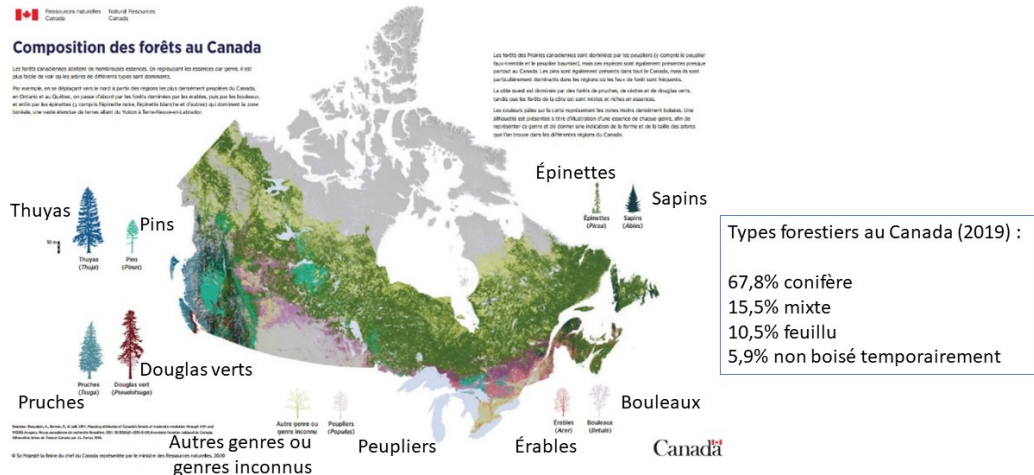


FIGURE 1.2 : Composition des forêts du Canada, Ressources naturelles Canada, Service canadien des forêts (2020).

Dans la régulation du climat, les forêts constituent une partie importante du cycle du carbone. En effet, elles stockent et libèrent du carbone suivant « un processus dynamique de croissance, de décomposition, de perturbation et de renouvellement » (Ressources naturelles Canada, 2022a). Un de leurs rôles majeurs, à l'échelle mondiale, est d'agir comme puits de carbone, c'est-à-dire de capteur, sous l'effet combiné de la fertilisation par le dioxyde d'azote et de l'augmentation dans certaines régions de superficies boisées. Au cours des 40 dernières années, les forêts ont absorbé le quart du dioxyde de carbone (CO₂) émis par les activités humaines (Ressources naturelles Canada, 2022a). Même si de nombreuses questions demeurent par rapport aux mécanismes du puits de carbone observé, qui font intervenir des processus variés – à des échelles multiples – que ce soit sur le plan physiologique des arbres ou par rapport à la disponibilité en eau qui est impactée par les changements climatiques, une chose est sûre : la déforestation contribue de façon significative aux émissions de CO₂, en particulier en Amazonie (Gatti *et al.*,

2021).

Définissons plus précisément les notions de puits et de source de carbone.

Une forêt est considérée comme une source de CO_2 dans le cas où elle rejette plus de CO_2 qu'elle n'en absorbe. Le CO_2 forestier est rejeté, par exemple, quand les arbres brûlent ou lorsqu'ils se décomposent après leur mort. Inversement, elle sera considérée comme un puits de CO_2 si la quantité rejetée est inférieure à celle absorbée. La photosynthèse est le mécanisme qui contribue à l'absorption de CO_2 . Ce dernier est alors assimilé pour fabriquer la biomasse (troncs, racines, feuilles, etc.) et s'accumule dans les sols par l'entremise de la matière organique morte (Ressources naturelles Canada, 2022a). En considérant ainsi la différence entre la productivité primaire nette provenant de la photosynthèse et la respiration hétérotrophe, le bilan net de ces échanges va alors déterminer si une forêt est une source ou un puits de CO_2 .

Les forêts emmagasinent donc de grandes quantités de carbone, à travers les arbres et les sols, mais du carbone est également relâché dans l'atmosphère par les microorganismes qui décomposent les arbres en dégradation. Les incendies accélèrent ce processus en augmentant l'émission de CO_2 , mais aussi de méthane (CH_4) et d'oxyde nitreux (N_2O) – de forts gaz à effet de serre (GES). Ces GES contenus naturellement dans l'atmosphère ont la capacité de retenir dans les basses couches de l'atmosphère une partie du rayonnement infrarouge émis par la surface de la Terre qui est réchauffée par le soleil. Ces gaz agissant comme une serre permettent de maintenir une température viable. Mais une trop grande quantité de ces gaz implique un réchauffement trop important de la planète, phénomène qui est observable depuis la révolution industrielle.

Ainsi, dans ce bilan, il faut alors considérer les plantations d'arbres, la déforestation, les épidémies d'insectes et les incendies. Les données récoltées chaque année par Ressources naturelles Canada permettent de dresser un bilan annuel des émissions

ou absorptions en Mt (mégatonne) équivalent CO₂ (figure 1.3).

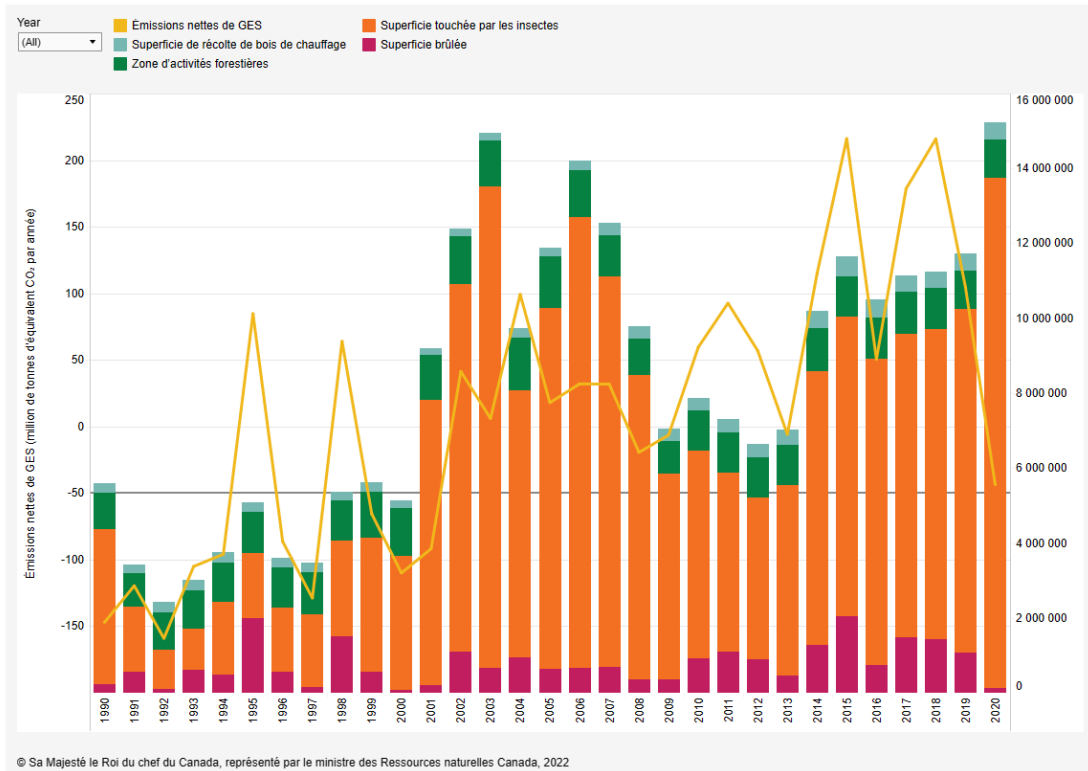


FIGURE 1.3 : Émissions nettes de CO₂ équivalent dans les forêts aménagées du Canada, entre 1990 et 2020. Reproduit à partir de Ressources naturelles Canada, *Comment les perturbations façonnent-elles les forêts du Canada ?*

Nous constatons au regard de ces relevés que les forêts canadiennes entre 1990 et 2001 ont été dans l'ensemble des puits de carbone, sauf en 1995 et 1998 où elles se sont comportées comme des sources. Ces années semblent être les années pour lesquelles une plus grande superficie a été perturbée par des feux. À partir de 2001, les perturbations liées aux insectes ont été prépondérantes. En effet, depuis 1999, le dendroctone du pin ponderosa a causé la mort d'arbres sur environ 10 millions d'hectares, en Colombie-Britannique, augmentant ainsi les émissions de CO₂ (Ressources naturelles Canada, 2021). Depuis 2001, les émissions ont été majoritairement supérieures et les données plus récentes des forêts de cette province ont mis en évidence une élévation des émissions nettes en équivalent CO₂, avec

+165.4 millions de tonnes en 2019. Les raisons principales de ces émissions furent 1 million d'hectares de superficie brûlée et 9.1 millions d'hectares perturbés par les insectes et la décomposition du bois mort qui n'a pas été ramassé (Ressources naturelles Canada, 2021).

Ces perturbations, qui peuvent être abiotiques (provenant de composantes non vivantes de l'environnement) ou biotiques ont donc un impact important sur le bilan de carbone des forêts. Or, l'apparition et la récurrence des ces perturbations sont liées aux conditions climatiques. Pour le Canada, comment ces conditions vont-elles évoluer dans un contexte de changements climatiques ?

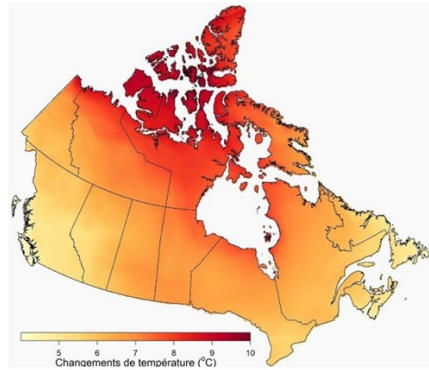
1.1.4 Impacts des changements climatiques pour le Canada

- À l'échelle du Canada

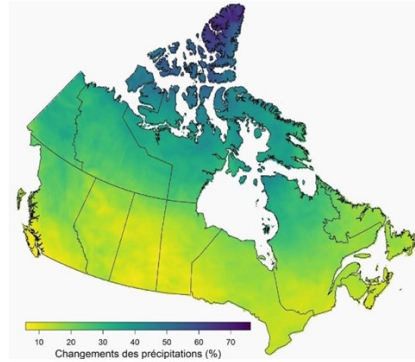
À l'échelle du Canada, la température annuelle moyenne a connu une hausse de 1.7°C entre 1948 et 2016, avec un réchauffement plus important dans le nord, en particulier durant l'hiver (hausse de 2.3°C durant la même période) (Environnement et Changement climatique Canada, 2017b).

Selon les modèles prédictifs RCP *Representative Concentration Pathway*, cette tendance va se poursuivre et toutes les provinces du Canada vont connaître une augmentation de leurs températures moyennes au cours du siècle. Le scénario de forçage radiatif le plus pessimiste (RCP8.5) prévoit une augmentation des températures plus prononcée dans le nord du Canada, allant jusqu'à 10°C. Les augmentations les plus faibles seront pour la côte sud-ouest de la Colombie-Britannique et pour les provinces de l'Atlantique (figure 1.4 (A)) (Environnement et Changement climatique Canada, 2017b).

Selon le même scénario RCP8.5, les précipitations augmenteraient de 8% jusqu'à plus de 70% dans les régions nordiques (figure 1.4 (B)) (Environnement et Chan-



(A) Illustration des variations projetées de la température annuelle d'ici la fin du XXI^e siècle pour l'ensemble de la masse continentale du Canada, par rapport à la période de référence 1996-2005. Le scénario d'émission utilisé est le RCP8.5.



(B) Variations projetées des précipitations annuelles. (Même période de référence et même scénario).

FIGURE 1.4 : Variations projetées de la température (A) et des précipitations annuelles (B) d'ici la fin du XXI^e siècle (horizon 2100). Ressources naturelles Canada.

gement climatique Canada, 2017a). Les augmentations les plus faibles seraient, comme pour les températures, sur la côte sud-ouest de la Colombie-Britannique, et sur les Prairies. Les variations observées pour les précipitations dépendent des saisons, contrairement aux températures qui augmenteraient pour toutes les saisons. L'influence géographique serait également plus importante dans le cas des précipitations.

Néanmoins, malgré la hausse prédite des précipitations, celle-ci ne serait pas partout compensée par la hausse de l'évapotranspiration et plusieurs régions connaîtraient alors des déficits hydriques.

- À l'échelle des Laurentides (Québec)

Certains modèles prédictifs sont développés à des échelles géographiques plus petites. Au Québec, grâce au Consortium sur la climatologie régionale et l'adaptation

aux changements climatiques (Ouranos), il est possible d'utiliser en ligne le modèle régional canadien du climat (MRCCS, version 5) pour obtenir les projections de certains paramètres.

Nous avons testé ce modèle pour les prévisions de températures et de précipitations³, pour la région des Laurentides (QC) dans laquelle se trouve la SBL, en utilisant le même scénario d'émission que pour les résultats canadiens présentés précédemment et obtenus par Ressources naturelles Canada. La figure 1.5 présente les prédictions obtenues pour la période allant de l'année 2041 à 2070.

Il faut préciser que la résolution spatiale utilisée dans le modèle d'Ouranos est de 60 km^2 (ce qui correspond à la taille d'un pixel sur la carte). De plus, les résultats obtenus à la figure 1.5 ne montrent pas les variations des températures et des précipitations, mais comparent les moyennes obtenues entre 1981 et 2010 aux moyennes modélisées pour 2041-2070 (figure 1.5 (A) et (B)).

Pour les températures, à l'échelle des Laurentides, nous constatons la même tendance que celle observée avec un modèle national. Le nord de la région, qui a connu une température moyenne entre 2°C et 3°C pendant la période de référence, aura une température moyenne qui tendra davantage vers 6°C . Cette augmentation sera également perceptible au sud de la région.

Le modèle régional d'Ouranos indique que les précipitations seront, en moyenne, plus importantes dans les Laurentides sur la période 2041-2070, en comparaison avec la période de référence 1981-2010. Les précipitations ne se référant pas seulement à la pluie, mais considèrent également la neige, la pluie verglaçante, la grêle et la bruine, leurs variations doivent donc englober les variations qui peuvent être observées durant l'hiver, tenant compte alors des éventuels déficits qui peuvent être prédits durant ces périodes.

3. <https://www.ouranos.ca/en/climate-portraits>

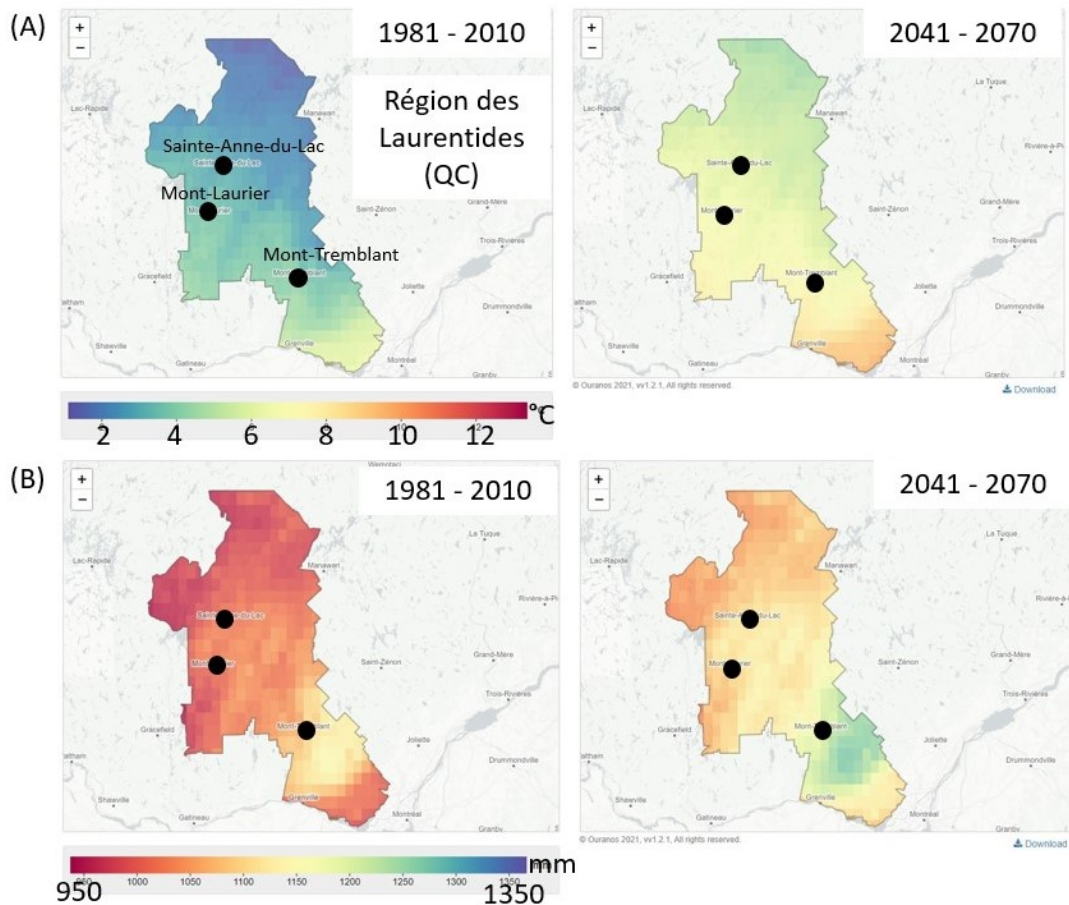


FIGURE 1.5 : (A) Projections des températures en °C et (B) des précipitations en mm pour la région des Laurentides (QC). La SBL se situe au sud de la région. Pour (A) et (B), la période de référence est représentée à gauche (1981-2010) et les projections sont sur la carte représentée à droite (2041-2070). Le scénario RCP8.5 a été utilisé pour permettre une comparaison avec les résultats de la figure 1.4. Pour les précipitations, l'échelle de gradient de couleurs va du rouge (950 mm) vers le bleu (1200 mm). Les résultats ont été obtenus à partir de la plateforme *Portraits climatiques* de Ouranos, version 1.3 (url : <https://www.ouranos.ca/en/climate-traits>)

- À l'échelle de Saint-Hippolyte (QC)

Nous pouvons enfin affiner ces résultats, en obtenant des projections pour la ville de Saint-Hippolyte (QC) où se situe la SBL.

Pour cela, nous avons utilisé le modèle disponible sur le site de *DonneesClima-*

tiques.ca⁴. La figure 1.6 présente l'évolution de la température annuelle moyenne⁵. Si nous regardons les valeurs médianes données par le modèle, selon les scénarios

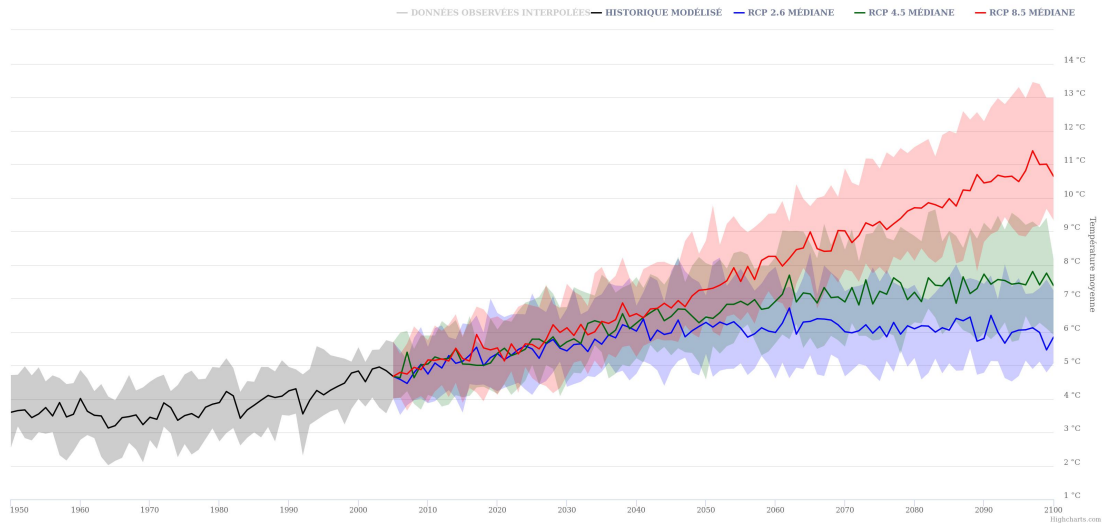


FIGURE 1.6 : Évolution de la température moyenne annuelle à Saint-Hippolyte (QC). Position : 45.933°N, 74,016°W Trois scénarios sont testés (sur la période 2005-2100) : RCP2.6 (courbe bleue), RCP4.5 (courbe verte) et RCP8.5 (courbe rouge). La courbe noire représente les données historiques entre 1950 et 2005 (échelle de l'axe des abscisses : de 1950 à 2100, avec un pas de 10 ans). Le graphique a été généré à partir du portail de DonneesClimatiques.ca

RCP2.6, RCP4.5 et RCP8.5, en 2030, les températures moyennes annuelles projetées seront de 5.8°C, 7.4°C et 10.6°C, respectivement. Entre 2000 et 2010, cette moyenne annuelle se situait autour de 4.9°C (Collin *et al.*, 2017a).

Les changements futurs des températures et des précipitations seront principalement déterminés par les émissions de GES qui sont pour la majorité d'origine anthropique.

4. Données climatiques accessibles en ligne : <https://donneesclimatiques.ca>
Données.Climatiques.ca est le fruit d'un travail de collaboration entre le Consortium sur les impacts climatiques du Pacifique (CICP), Ouranos Inc, le Centre climatique des Prairies (CCP), Environnement et Changement climatique Canada (ECCC), le Centre de Recherche Informatique de Montréal (CRIM) et Habitat7

5. Pour la modélisation réalisée en ligne : source en ligne
 Se référer également à (Charron, 2016).

Cependant, il est nécessaire de considérer également l'influence des variabilités naturelles du climat sur ces changements, elles-mêmes perturbées par l'activité humaine.

En effet, le climat est soumis à des variations naturelles qui oscillent d'un état à un autre. Ces cycles sont connus et ils peuvent être suivis, voire prédits.

Les oscillations qui touchent le Canada ne sont pas toutes identiques à celles que connaissent d'autres parties du monde, mais celles affectant les océans Pacifique et Indien entraînent des répercussions sur le climat du pays (E. Bush and D.S. Lemmen, 2019). Bush et Lemmen, dans leur rapport sur le climat changeant du Canada, rappellent qu'*El Niño Southern Oscillation* (ENSO) est une variation quasi périodique de la température à la surface de l'océan Pacifique, impactant aussi les vents. Les phases El Niño, plus chaudes, ont pour effet sur le Canada des hivers plus doux et des conditions plus sèches. Pendant les phases La Niña, plus fraîches, les températures sont plus basses que les normales et les précipitations ont tendance à être plus importantes. Le début 2022 est marqué par une période La Niña, selon les suivis réalisés par le Bureau météorologique australien (BOM). S'ajoutent à ces effets, ceux du dipôle de l'océan Indien (DOI) qui est défini par la différence des températures de mer (en surface) entre la partie tropicale ouest et est de l'océan Indien⁶. Un dipôle négatif en interaction avec une phase La Niña aura tendance à augmenter davantage les précipitations.

Une autre variabilité de la température de l'océan Pacifique – touchant les latitudes plus au nord du Pacifique – est l'oscillation décennale du Pacifique (ODP). Sa variation est plus irrégulière, avec des périodes d'oscillation allant de quelques années à plusieurs décennies. Ses effets sont similaires à ceux de El Niño (températures plus chaudes en hiver pour le Canada).

6. Gouvernement de l'Australie. Bureau of Meteorology. The Indian Ocean Dipole (IOD), [En Ligne]. <http://www.bom.gov.au/climate/enso/history/ln-2010-12/IOD-what.shtml>

Le pays est aussi soumis à l'oscillation arctique (OA) ou mode annulaire boréal, qui explique plus fortement les variabilités de la température et de la pression atmosphérique au nord du 20^e parallèle nord. Ce mode n'a pas de périodicité bien définie. Sa phase positive semble impliquer des températures plus élevées dans l'ouest canadien, et des températures plus froides à l'est et au nord. Influant également les vents venant de l'ouest, l'OA a un impact sur l'emplacement du passage des tempêtes au-dessus de l'Atlantique Nord (E. Bush and D.S. Lemmen, 2019).

Du côté de l'Atlantique, l'oscillation atlantique multidécennale (OAM) est un mode qui va régir la température à la surface de l'océan Atlantique Nord. Sa période est plus longue, de 60 à 80 ans. L'OAM va avoir davantage de répercussions sur l'activité des ouragans et l'intensité des pluies du côté de l'Atlantique Nord.

Les effets de ces variations naturelles du climat sont donc à ajouter aux effets induits par les changements climatiques, accentuant ainsi les conséquences potentielles sur les forêts.

1.1.5 Quels sont les impacts potentiels de ces changements et cycles climatiques sur les forêts ?

Des effets cumulatifs liés aux changements climatiques sont visibles sur les forêts du Canada. Nous l'avons souligné, la fréquence et la gravité des perturbations naturelles comme les feux, les sécheresses et les épidémies d'insectes vont augmenter, mais de façon inégale sur le territoire. La concentration en CO₂ augmentant (voir figure 1.3), cela impactera la distribution et l'abondance relative des essences forestières (Forget *et al.*, 2003; Iverson *et al.*, 2019a,b). Depuis 2006, des effets ont déjà été observés au Québec, tels qu'un prolongement de la saison de croissance des arbres et un débourrement hâtif de certaines essences comme l'érable à sucre, mais aussi un déplacement de la migration des espèces vers des régions plus fraîches, surtout mis en évidence par des approches de modélisation (Ministère des forêts

de la faune et des parcs (2021), p.1), (Boisvert-Marsh *et al.*, 2014, 2022).

Les modèles utilisés, notamment le modèle régional d'Ouranos, permettent d'anticiper les principaux changements climatiques qui entraîneront des conséquences directes sur les forêts du Québec, et ce, selon leurs sous-domaines bioclimatiques. Précisons qu'un domaine bioclimatique est un territoire caractérisé par la nature de la végétation qui, à la fin des successions, couvre les sites où les conditions pédo-logiques, de drainage et d'exposition sont représentées par des données moyennées (Ministère des forêts de la faune et des parcs , 2021). C'est en quelque sorte une zone moyenne représentative d'un territoire donné.

Nous avons résumé les changements anticipés par le modèle d'Ouranos dans le tableau 1.1.

La spécificité de chaque sous-domaine, peuplé d'essences différentes, va moduler les effets induits par les changements climatiques. Certains seront positifs, mais jusqu'à quel seuil ?

Les écosystèmes situés aux limites plus nordiques seront plus touchés par les hausses de températures. Les précipitations totales resteront stables, mais associées à une température moyenne plus élevée, et cela aura pour conséquence une augmentation du risque de sécheresse. En forêt boréale, il y aura des effets positifs sur la productivité, mais à court terme. Au-delà d'une certaine température, l'impact sera négatif et la croissance diminuera (D'Orangeville *et al.*, 2018a). Comme le montrent D'Orangeville *et al.* (2018a), toutes les essences ne répondent pas de la même façon à une augmentation des températures et leur optimum varie.

Le problème d'adaptation des essences aux changements climatiques est un enjeu important. Certaines vont migrer vers le nord, d'autres vont au contraire en bénéficier. Le rapport sur les forêts du Canada (2020) explique que l'anticipation de ces effets pour les forêts peut passer par la mise en valeur des espèces qui ont le plus de capacités à résister aux changements ou par la promotion des conditions

TABLEAU 1.1 : Changements anticipés au niveau des forêts sous l'influence des changements climatiques (d'après les modélisations réalisées par Ouranos).

Variable ou phénomène	Changements anticipés
Température moyenne	Augmentation dans tous les sous-domaines bioclimatiques du Québec.
Précipitations totales moyennes	Stables dans tous les sous-domaines.
Précipitations totales stables associées à une température moyenne plus élevée	Risque accru de sécheresse. Plus grande récurrence de la hausse des débits de pointe (maximaux) dans certains bassins versants.
Croissance des arbres	La saison de croissance se prolonge (mais plus marquée dans la région de la pessière à mousses).
Période de gel/dégel	Les événements de gel/dégel augmentent pendant l'hiver dans les sous-domaines plus au nord, avec une diminution au printemps et à l'automne.
Écozones de végétations	Déplacement des écozones vers le nord. Ceci entraînera des conséquences sur les espèces fauniques forestières. Certaines espèces ne pourront plus migrer vers des zones plus froides, en raison d'obstacles. Il faudra alors créer des corridors écologiques pour faciliter ces migrations. Mésadaption de certaines essences aux nouvelles conditions. Apparition d'espèces exotiques envahissantes.

qui vont favoriser la résilience des écosystèmes forestiers.

Pour conclure cette partie et dans le but d'illustrer les conséquences de ces changements climatiques sur les espèces forestières, nous avons utilisé le modèle MaxEnt de Ressources naturelles Canada⁷. C'est leur modèle le plus récent qui s'appuie sur

7. Source : modèle MaxEnt

l'apprentissage machine pour déterminer la probabilité de présence d'une espèce végétale dans un lieu donné. Il repose sur six variables climatiques :

- La température annuelle moyenne
- La température maximale du mois le plus chaud
- La température minimale du mois le plus froid
- Les précipitations annuelles
- Les précipitations du mois le plus chaud
- Les précipitations du mois le plus froid

MaxEnt fait appel à l'optimisation de l'entropie (**Maximum Entropie**). Nous l'avons appliqué, en ligne, sur le site mis à disposition par Ressources naturelles Canada. Nous avons choisi d'évaluer la probabilité de présence de différentes essences présentes à la SBL pour la période 2041-2070. Nous présentons les résultats obtenus dans le cadre des scénarios RCP2.6 et RCP8.5, du plus optimiste au plus pessimiste (figure 1.7).

La comparaison des figures 1.7 (A) et 1.7 (B) met en évidence l'apparition de l'érable rouge (*Acer rubrum*) vers des zones plus au nord que celles où il était répertorié entre 1971 et 2000. Les deux scénarios présentent des zones de migration similaires, mais dans le cas du scénario le plus pessimiste, à la figure 1.7 (C), la probabilité de présence y est plus élevée.

À la figure 1.8, les mêmes phénomènes sont observables pour le bouleau jaune (*Betula alleghaniensis*), mais avec une migration plus accentuée vers les régions plus au nord associée à une perte de l'essence dans les régions situées proches des grands lacs.

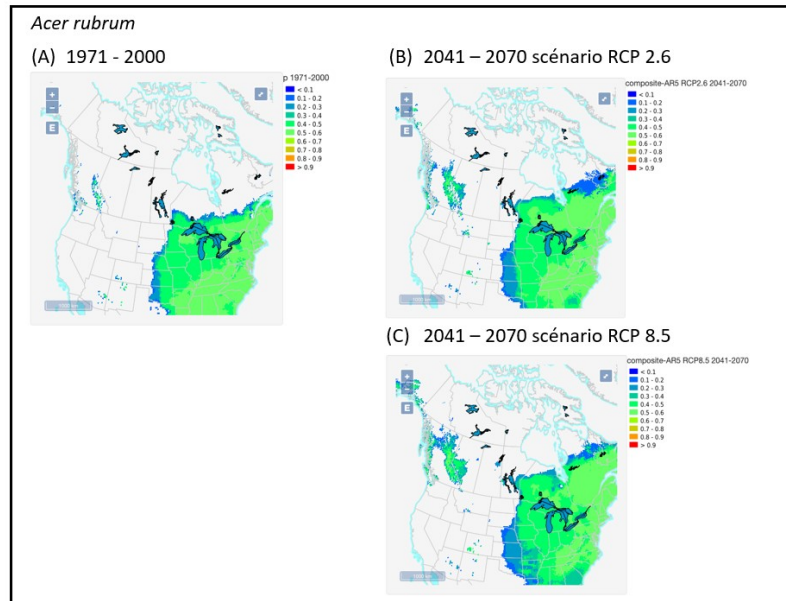


FIGURE 1.7 : (A) Répartition de l'érable rouge (données de 1971 à 2000). (B) Probabilité de présence de l'érable rouge pour 2041-2070, même échelle de probabilité que (A) et scénario RCP2.6, (C) scénario RCP8.5. Résultats obtenus à partir du modèle MaxEnt de Ressources naturelles Canada (url : <http://planthardiness.gc.ca>).

Pour le hêtre à grandes feuilles (*Fagus grandifolia*), les projections représentées sur la figure 1.9 montrent également une migration de l'espèce vers le nord, avec une perte d'espace dans les régions au sud-ouest des grands lacs ; leur probabilité de présence y devient en effet inférieure à 0.1, ce résultat étant accentué pour le scénario RCP8.5. Cependant, nous pouvons noter l'apparition de nouvelles zones propices au hêtre, à l'ouest du Canada, dans des régions aux latitudes plus élevées.

Ainsi, ces projections permettent de visualiser le déplacement des essences forestières vers le nord sous l'effet des changements climatiques. Cependant, comme l'ont montré Collin *et al.* (2017a), des facteurs non climatiques doivent également être pris en considération dans l'étude migratoire des espèces. Les auteurs mettent en évidence en particulier l'interaction entre les espèces et leur capacité à s'acclimater aux conditions des sols, capacité qui évolue également sous l'augmentation

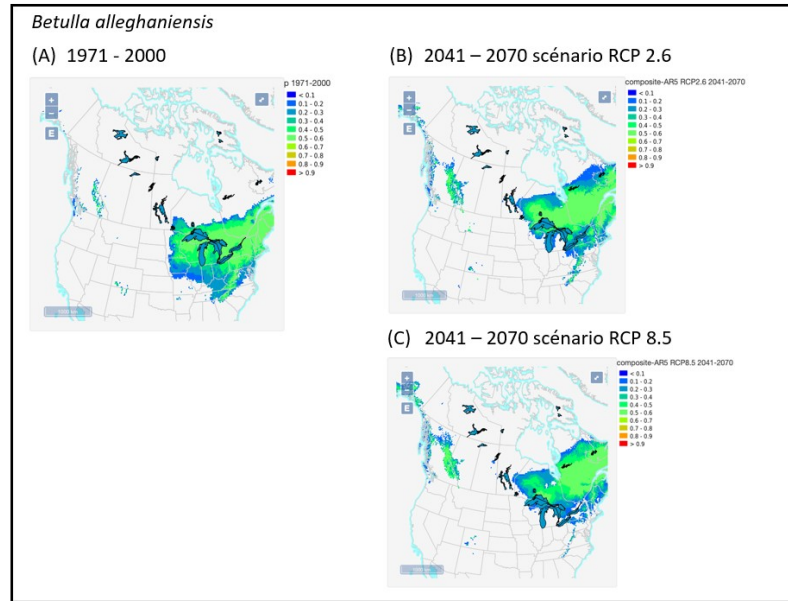


FIGURE 1.8 : (A) Répartition du bouleau jaune (données de 1971 à 2000). (B) Probabilité de présence du bouleau jaune pour 2041-2070, même échelle de probabilité que (A) et scénario RCP2.6, (C) scénario RCP8.5. Résultats obtenus à partir du modèle MaxEnt de Ressources naturelles Canada (url : <http://planthardiness.gc.ca>).

des températures. Or, les modèles utilisés, en ne considérant souvent que des enveloppes climatiques, ne prennent généralement pas ces aspects en considération.

De nombreuses variabilités sont présentes en hydroclimatologie. Nous évoquons les variations naturelles cycliques dues à différentes oscillations climatiques qui affectent les températures et les précipitations. Ces variations naturelles sont également impactées par l'activité humaine et leur période d'oscillation s'en trouvera bouleversée. Il n'est pas encore bien clair quels seront les impacts des changements climatiques sur ces périodes d'oscillation, certaines études prévoyant une augmentation de la fréquence des épisodes de type El Niño sous des scénarios différents de changements climatiques, d'autres réfutent ces résultats en montrant une variabilité inhérente à ces phénomènes oscillatoires (Alizadeh *et al.*, 2022).

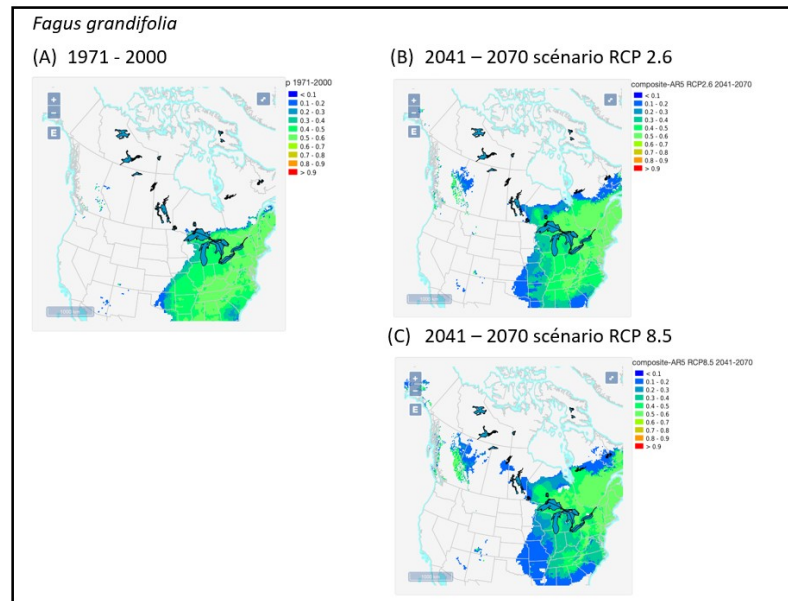
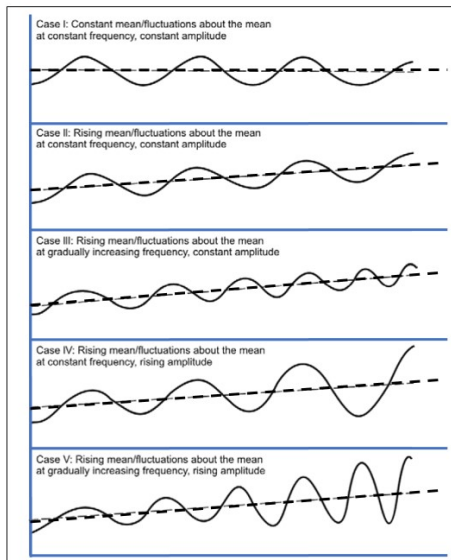


FIGURE 1.9 : (A) Répartition du hêtre à grandes feuilles (données de 1971 à 2000). (B) Probabilité de présence de l’hêtre pour 2041-2070, même échelle de probabilité que (A) et scénario RCP2.6, (C) scénario RCP8.5. Résultats obtenus à partir du modèle MaxEnt de Ressources naturelles Canada (url : <http://planthardiness.gc.ca>).

Ce qui ressort de ces débats est la nécessité de bien définir le terme de *variabilité climatique* qui se réfère aux fluctuations temporelles du climat - ou plus précisément des variables le caractérisant - autour des valeurs moyennes - tandis que le terme de *changement climatique* désigne un glissement global des conditions moyennes du climat, dans une région donnée (Rosenzweig et Hillel, 2008).

La figure 1.10 extraite du livre de Rosenzweig et Hillel (2008) illustre les différents cas à considérer dans un contexte de variabilité. Le cas (1) représente un climat moyen constant, présentant des variations régulières autour de la valeur moyenne (en pointillée sur le graphique), et ces variations sont de fréquence et d’amplitude constantes. Le cas (2) considère un changement graduel du climat moyen, où les fluctuations sont croissantes, avec des variations autour de la valeur moyenne qui se font toujours à fréquence et amplitude constantes. Néanmoins, nous observons



Cas 1 : Fluctuations constantes autour de la valeur moyenne, à fréquence et amplitude constantes.

Cas 2 : Fluctuations croissantes autour de la valeur moyenne, à fréquence et amplitude constantes.

Cas 3 : Fluctuations croissantes autour de la valeur moyenne, à une **fréquence augmentant graduellement** et à une amplitude constante.

Cas 4 : Fluctuations croissantes autour de la valeur moyenne, à une fréquence constante et à une **amplitude croissante**.

Cas 5 : Fluctuations croissantes autour de la valeur moyenne, à une **fréquence augmentant graduellement** et à une **amplitude croissante**.

Modes de la variabilité climatique en lien avec les changements climatiques (Rosenzweig, 2008)

FIGURE 1.10 : Différents cas de variabilité à considérer en lien avec les changements climatiques (figure extraite et adaptée de Rosenzweig et Hillel (2008), p.13)

que les maxima et les minima des variations augmentent avec le temps. Le cas (3) illustre des fluctuations croissantes autour de la valeur moyenne, la fréquence augmentant graduellement, mais l'amplitude reste constante. Lorsque le climat moyen change, les fluctuations autour de la valeur moyenne voient alors leur amplitude augmenter, à fréquence constante, impliquant des valeurs maximales de plus en plus hautes au court du temps (cas 4). Quand la fréquence des fluctuations augmentent également, telle que représentée par le cas (5), le climat devient de plus en plus instable et ses manifestations de plus en plus extrêmes.

Stevenson *et al.* (2022) mettent en avant la nécessité de considérer différentes échelles temporelles pour appréhender ces variabilités. À celles-ci s'ajoutent les tendances plus grandes influencées par les changements climatiques, ce que nous nommons le *glissement* global des phénomènes climatiques, comme par exemple une augmentation de l'humidité ou une aridification, selon les régions du monde. Il est attendu que ces tendances de fond dominant les variabilités internes qui

régissent le risque de sécheresses régionales (Stevenson *et al.*, 2022). L'enjeu est donc d'étudier l'influence de paramètres locaux dans un contexte hydroclimatique globalement changeant où les événements extrêmes voient leur fréquence augmenter. Or, dans le cas des sécheresses, il devient plus difficile de définir ce type d'événement dans un contexte où l'état de fond se déplace continuellement. Stevenson *et al.* (2022) le formulent ainsi : « *How one should properly define a drought event, in a situation where the background is shifting?* »

1.2 Événements extrêmes et nouvelle normalité

Cette réflexion sur les lignes de base en constante évolution (*continuously shifting baseline*) est importante lorsque sont abordées les notions d'événements extrêmes. Tout d'abord, parler d'*événements* implique des phénomènes ponctuels, dont l'analyse ne sera plus centrée sur les tendances (Jentsch *et al.*, 2007). Quand ces événements se font de plus en plus récurrents et longs, nous passons d'une considération temporelle discrète à une approche continue (Stevenson *et al.*, 2022). La question n'est plus : quelle sera la fréquence des périodes de sécheresse extrême ou des périodes pluvieuses à venir ? Mais plutôt : quand reconnaissons-nous une période de sécheresse ou de pluie comme un changement dans ce qui est considéré comme normal ?

Notre projet se place dans le prolongement de cette réflexion, car sa motivation première est de réfléchir à la *nouvelle normalité* du potentiel hydrique du sol, sous contraintes climatiques.

Ce passage du discret au continu, sous l'effet des changements climatiques, implique alors une autre question : comment identifier ce basculement vers un nouvel état de fond ?

Grâce à une approche calculatoire s'appuyant sur le concept de temps d'émergence (*time-of-emergence*), Hawkins et Sutton (2012) ont essayé de déterminer quand

un scénario de hautes émissions de GES prédit que les conditions normales atteindront ou dépasseront le seuil d'une méga-sécheresse/pluie actuelle. Le moment de l'émergence est défini comme l'année au cours de laquelle la moyenne de l'ensemble dépasse le seuil pertinent défini et n'y revient plus. Ainsi, quand le changement est de magnitude suffisante par rapport à la variabilité de la ligne de base, nous pouvons nous assurer d'un changement significatif.

À une échelle plus petite, la variabilité naturelle s'avère plus élevée, impliquant alors une détection plus difficile des changements. Or, un des enjeux est de pouvoir anticiper ces changements dans les régimes actuels des systèmes climatiques. Pouvons-nous déterminer des signes avant-coureurs (*early warning signal*) d'un basculement d'un état vers un autre état ? Il devient en effet crucial de pouvoir anticiper ces bifurcations, ou changements de régime, surtout lorsque que des changements même faibles peuvent avoir des conséquences importantes sur les forêts, et particulièrement lorsque les conditions sont proches d'un seuil critique (Trumbore *et al.*, 2015).

Trumbore *et al.* (2015) se posent d'ailleurs la question : « *How much disturbance or stress is too much ?* ». Pour y répondre, il faut tout d'abord définir différents états à l'échelle de la forêt : un état sain (*healthy*), un état instable (*unstable*), et un état non-sain (*unhealthy*) (figure 1.11). Chaque transition d'un état à un autre va caractériser un état de résilience du système *forêt* : élevé, faible ou inexistant. La zone orange sur la figure 1.11 est la zone critique, car la transition d'un état sain à un état non-sain passe par cet état de transition instable qui dépend de l'état de résilience du système. Le défi est de savoir si le passage se fera de façon abrupte ou en suivant un déclin linéaire, et si ce changement a lieu, des signes avant-coureurs qu'un seuil est sur le point d'être franchi peuvent-ils être détectés ?

D'après la FAO, pour définir la santé et la vitalité d'une forêt, il faut se reposer sur une analyse combinée des facteurs de perturbations abiotiques et biotiques et de leurs impacts sur la croissance et la survie des arbres. La difficulté, nous l'avons

soulignée plus haut, est là aussi de déterminer quand la fréquence, l'étendue et l'amplitude des stressseurs - souvent couplés entre eux - dépassent les variabilités naturelles et affectent la trajectoire de récupération de la végétation, à l'échelle du paysage et à l'échelle régionale (Trumbore *et al.*, 2015).

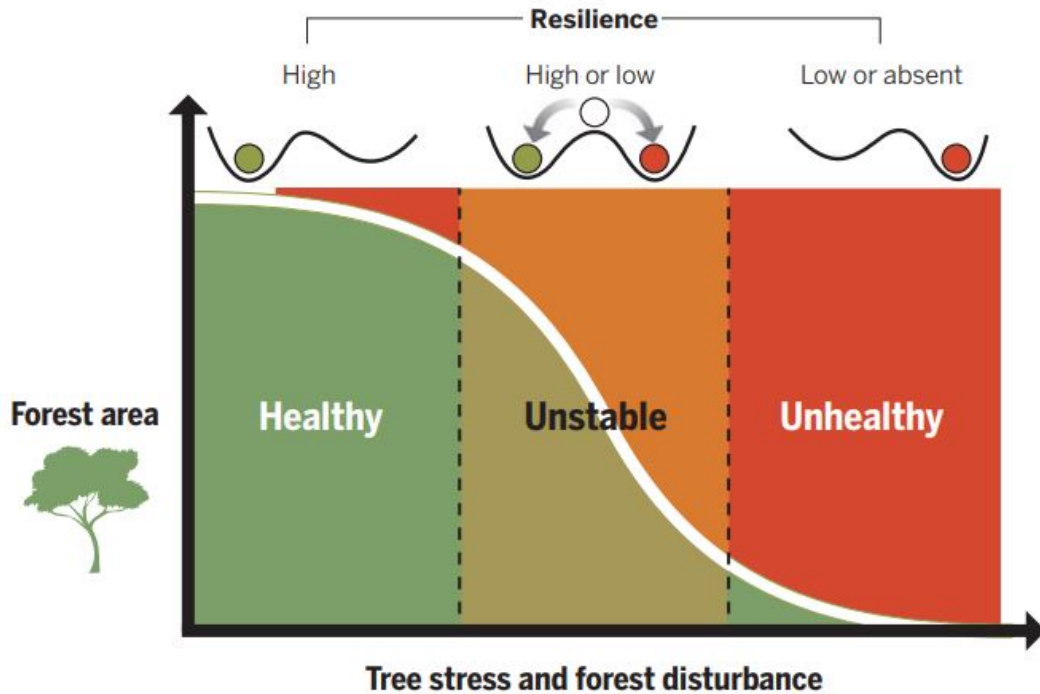


FIGURE 1.11 : Représentation graphique des changements sur une forêt saine en fonction du stress des arbres et des perturbations que subit la forêt (figure extraite de Trumbore *et al.* (2015)). Trois états sont définis (sain, instable, non-sain), se rattachant à trois états de résilience (élevé, faible, inexistant). Lorsque la résilience est élevée, le système a tendance à rester dans son état actuel. Quand celle-ci est faible, il aura tendance à basculer plus facilement vers un autre état. La zone en vert illustre une zone de stress faible. La zone orange représente la zone critique.

Nous avons repris le terme de *résilience* utilisé par Trumbore *et al.* (2015). Holling (1996) fut l'un des premiers à considérer le mot dans un contexte écologique en le définissant comme une résistance au changement, i.e. l'amplitude d'une perturbation qu'un système peut absorber avant de basculer d'un état à un

autre. Le concept est utilisé dans de nombreux domaines, de la physique à la médecine. Quelque soit le contexte, le dénominateur commun de la définition du mot *résilience* est la capacité d'un système - d'une entité - à résister au changement (Capdevila *et al.*, 2021). La définition de Holling (1996) met l'accent sur la notion de *résistance* au changement ; ce qui est différent de la considération en ingénierie, qui considère la résilience comme le processus de *récupération* d'un système suite à une perturbation, et l'intègre donc comme une composante de la stabilité d'un système, qui ne peut qu'être unique si nous faisons l'hypothèse qu'un système ne peut avoir qu'un seul régime stable. La vision écologique de la résilience, en s'apparentant à la résistance au changement, permet de considérer un système comme ayant une multitude d'états. Cette distinction entre la vision de l'écologie et de l'ingénierie est importante car au fil du temps, en écologie, les deux définitions se sont mélangées, apportant un flou autour du mot *résilience*. Standish *et al.* (2014) clarifient la définition, en séparant la définition de Holling (1996) qui se réfère donc à la capacité d'un écosystème à tolérer les perturbations tout en persistant dans son état actuel, de celle de Pimm (1984) qui a été le premier à lier le terme de résilience au temps mis par un écosystème à revenir à un état de pré-perturbation. La première définition est nommée résilience écologique et la seconde, résilience d'ingénierie. Celle-ci est associée au concept de récupération. D'autres définitions ont été formulées autour des notions de *persistance* et de *perturbation* (Delettre, 2021).

Ainsi, les trois courbes en noir sur la figure 1.11 schématisent les différents états qui peuvent exister en s'appuyant sur la représentation de l'énergie potentielle d'une boule qui roule. En adaptant la figure explicative proposée par Capdevila *et al.* (2021), nous illustrons cette analogie (figure 1.12). Le système (a), avec un bassin d'énergie plus haut et plus profond que le système (b), possède une résilience - écologique - élevée, capable en effet de résister davantage aux perturbations, avec un point de basculement plus difficile à atteindre que pour le système (b).

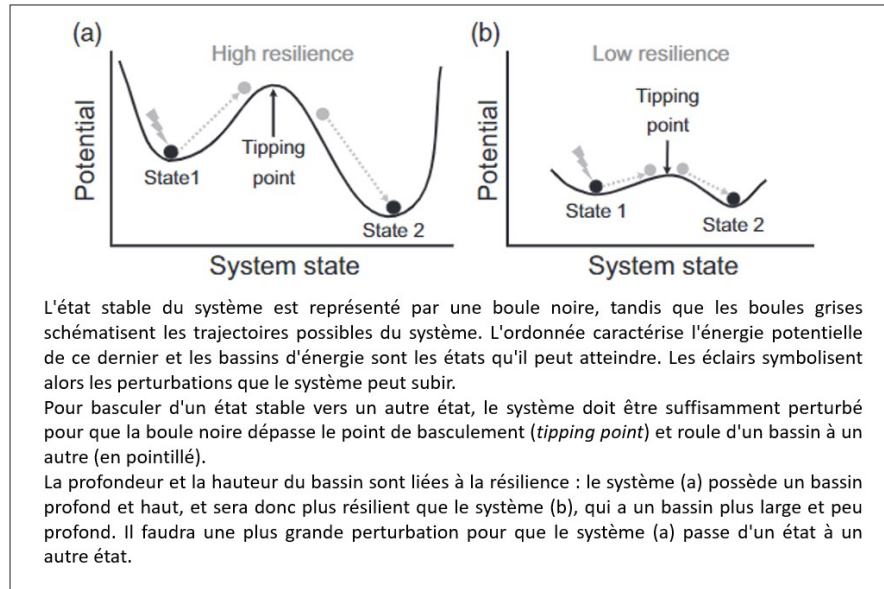


FIGURE 1.12 : Analogie avec la boule qui roule permettant de schématiser le concept de résilience écologique, au sens de Holling (1996) (figure extraite et adaptée de Capdevila *et al.* (2021)).

Voilà ce qu'il en est, sur le plan théorique. Cependant, Scheffer (2009), pionnier dans l'étude de la stabilité des systèmes complexes, en fait le constat dans son livre *Critical transitions in Nature and Society* (p.260) : « *Predicting critical transitions is particularly difficult* » .

Des méthodes ont été développées pour appréhender ces transitions critiques, et elles ont permis de mettre en évidence l'existence de métriques capables de les anticiper (Scheffer *et al.*, 2009; Dakos *et al.*, 2012). Wissel (1984) a montré qu'un système qui s'approche d'une transition critique, lorsqu'il est soumis à une faible perturbation, reviendra plus lentement à son état d'équilibre. Ce ralentissement a été nommé *ralentissement critique* (van Nes et Scheffer, 2007). Six mécanismes peuvent mener à ce ralentissement, avec ou sans passage par un point de bifurcation, et des indicateurs ont été définis pour permettre leur détection (Dakos *et al.*, 2015). Parmi ces indicateurs, nous pouvons citer, dans le domaine temporel, l'analyse de

l'autocorrélation (Dakos *et al.*, 2012) et de la variance (Dakos *et al.*, 2012; Seekell *et al.*, 2011), l'augmentation des coefficients d'asymétrie (Guttal et Jayaprakash, 2008) et d'aplatissement (Biggs *et al.*, 2009), respectivement *skewness* et *kurtosis*. En exploitant l'autocorrélation, Forzieri *et al.* (2022) ont mis en évidence que 23 % des forêts intactes non perturbées a déjà atteint un seuil critique et que cette proportion est en train de vivre une dégradation de sa capacité de résilience, ce qui la rend plus sensible aux perturbations externes. L'analyse de la cohérence spatiale peut également être réalisée, et des signes avant-coureurs de ralentissement critique mis en évidence, comme par exemple pour l'étude de la désertification d'écosystèmes (Kéfi *et al.*, 2014).

Ces indicateurs, précisent Dakos *et al.* (2015), ne sont pas des outils de prédiction en tant que tels. En théorie, exactement au moment de la transition critique, l'autocorrélation du système est égale à 1 et sa variance explose à l'infini : en d'autres termes, l'état du système devient de plus en plus proche de son état passé. Or, c'est la différence entre la valeur attendue à la transition et la valeur actuelle qui permet de quantifier si le système écologique est proche d'un basculement. Mais cela n'est vrai que si la transition arrive au moment exact de la bifurcation. Pour des systèmes stochastiques, la transition peut se produire avant cette bifurcation. Ainsi, pour augmenter le pouvoir prédictif de ces indicateurs, il faut les mesurer à travers le temps et évaluer s'ils s'écartent de la ligne de base qui a été établie.

De plus, il faut prendre en considération que deux types de point de basculement (*tipping point*) existent : un système peut en effet basculer vers un autre état à cause d'un changement dû à des conditions externes, ou à cause d'un changement de son propre état. Le premier type de point de basculement est plus difficile à anticiper par les indicateurs de résilience (van Nes *et al.*, 2016).

Proposant un cadre d'analyse, Dakos *et al.* (2015) concluent néanmoins que le fait de savoir le risque d'une transition est une condition nécessaire mais pas suffisante pour l'éviter. Souvent, les indicateurs sont étudiés rétrospectivement, ou

les signes avant-coureurs émergent trop tard, voire ne sont pas présents (Boerlijst *et al.*, 2013). Enfin, ils peuvent nous dire que quelque chose va se produire, qu'un changement d'état du système aura lieu, mais ils ne nous disent pas quel sera ce nouveau système. Ainsi, il sera nécessaire de remettre ces outils mathématiques - ses signes avant-coureurs - dans le contexte écologique propre à chaque système étudié.

Ce type d'approche a été utilisé dans des contextes écologiques forestiers, en particulier pour comprendre la distribution du couvert de la forêt boréale et les transitions que celle-ci pourrait connaître (Hirota *et al.*, 2011; Scheffer *et al.*, 2012). Liu *et al.* (2019) ont également exploité ces outils pour montrer que la perte de résilience, au cours du temps, permettait de prédire la probabilité du taux de mortalité d'une forêt située en Californie, région qui a subi deux périodes de sécheresse entre [2007 ;2009] et [2012 ;2015]. La prédiction a permis d'anticiper la perte du couvert végétal induite par les sécheresses.

Dans ce contexte d'événements extrêmes, nous allons nous intéresser en particulier aux sécheresses, en mettant en lumière leurs conséquences sur les forêts.

1.3 Le cas des sécheresses et les conséquences sur les forêts

Avant de nous tourner vers le concept de sécheresse, arrêtons-nous sur le mot *perturbation*. Il peut être défini comme survenant quand « des forces potentiellement destructrices sont appliquées sur un espace d'habitat occupé par une population, une communauté ou un écosystème » (Lake, 2008); ces forces pouvant être de nature physique, chimique, biologique ou une combinaison des trois. C'est la mesure des conséquences induites qui permettent d'évaluer la puissance des perturbations, mais cette approche rend difficile les comparaisons entre événements. En ce sens, les perturbations - selon Lake (2008) - devraient être évaluées par leurs propriétés abiotiques qui doivent être dissociés des réponses biotiques qui ont lieu suite à ces perturbations. Ainsi, une sécheresse devrait être caractérisée par la durée

de précipitations extrêmement faibles, l'intervalle moyen de leur récurrence, leur délimitation spatiale et leur caractère prédictif. Il est ainsi nécessaire de faire la distinction entre la perturbation et la réponse du système.

Trois types de perturbations sont proposées par Lake (2008) :

1. Pulse : semblable à un pic de Dirac, c'est un événement défini très précisément dans le temps.
2. Press : événement qui augmente en intensité pour atteindre une pression constante (au sens de force).
3. Ramp(e) : augmentation constante de la force au cours du temps, sans point final clair.

Selon lui, les sécheresses représentent la troisième catégorie de perturbation, et peuvent être classées, d'après Wilhite et Glantz (1985), par rapport à :

1. des considérations météorologiques (absence de précipitations)
2. des considérations hydrologiques (manque de ressources hydriques)
3. leurs conséquences en agriculture (diminution de l'hydratation des sols)
4. leurs conséquences socio-économiques (réponse à la demande en eau)

Mishra et Singh (2010) en ajoutent une cinquième, en lien avec la diminution des réserves des nappes phréatiques.

Une grande variété de définitions existe dans la littérature associée au thème des sécheresses. Slette *et al.* (2019) ont étudié 564 articles et ont conclu que seulement 32 % les définissent de façon précise dans les contextes relatifs à chaque étude, et que 50 % des analyses sont réalisées dans les intervalles de variabilité climatique. Il peut être alors nécessaire de caractériser le degré de sévérité des sécheresses, au moyen d'indices qui peuvent refléter leur intensité, leur durée et leur étendue spatiale (Svoboda et Fuchs, 2017). Les échelles temporelles généralement utilisées sont de l'ordre de l'année ou du mois. Le calcul de ces indices diffère, mais dans

la majorité des cas celui-ci s'appuie sur les précipitations. Une standardisation dans la caractérisation des sécheresses permet alors une meilleure comparaison des études. Les États-Unis ont ainsi créé un système national de suivi des sécheresses à travers le *National Integrated Drought Information System*⁸ dont l'objectif est de développer un outil d'anticipation des sécheresses à l'échelle du pays (*a drought early warning signal*).

Au Canada, Agriculture et Agroalimentaire Canada met à disposition un suivi des sécheresses qui permet d'obtenir des rapports statiques des conditions observées. Le Gouvernement fédéral propose également un outil en ligne qui permet de suivre l'évolution des sécheresses. À titre d'exemple, nous avons compilé les mesures obtenues pour le mois de juin 2020, sur la figure 1.13. Cinq états de sécheresse (*drought*) peuvent être répertoriés : de anormalement sec (niveau D0) à sécheresse exceptionnelle (niveau D4). Au Québec, en juin 2020, les trois premiers niveaux ont été atteints. La province a connu cette année là un printemps exceptionnellement chaud et sec (Bélanger *et al.*, 2021).

Plus récemment, au concept de méga-sécheresse, désignant de longues périodes de sécheresse s'étalant parfois sur 20 ans, s'est opposé le concept de sécheresse-flash (*flash drought*) qui se caractérise par un déclenchement soudain et une intensification rapide des conditions de sécheresse, à l'échelle de la semaine ou du mois (Otkin *et al.*, 2018).

Deux définitions ont été avancées pour caractériser ce type d'événements (Pendergrass *et al.*, 2020) :

1. Période exhibant une augmentation de plus de 50 % de l'indice EDDI (*Evaporative Demand Drought Index*), sur deux semaines, maintenue pendant au moins deux autres semaines.

8. Source : <https://www.drought.gov/about>

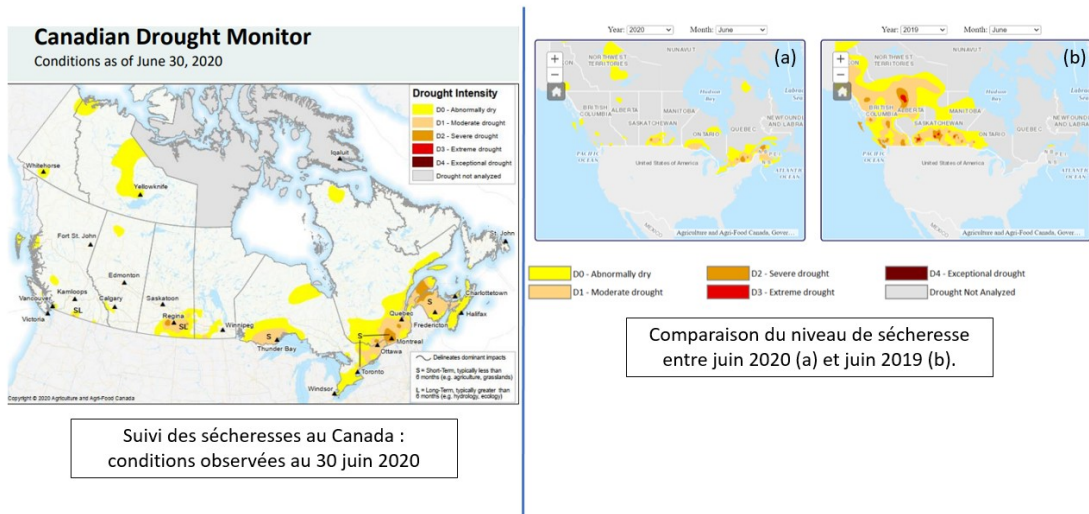


FIGURE 1.13 : Suivi des sécheresses au Canada : situation en juin 2020 (ressource accessible en ligne) et comparaison entre juin 2020 et juin 2019 (d’après l’outil de surveillance des sécheresses au Canada, Gouvernement du Canada).

2. Changement de deux catégories dans la métrique de l’USDM (*US drought monitor*) évoquée plus haut. Ce changement doit également se faire sur deux semaines et être maintenu pendant au moins deux semaines.

La fenêtre temporelle utilisée dans les définitions est donc relativement courte. De plus, la première définition s’appuie sur l’indice EDDI qui est défini à partir de l’évapotranspiration (ET) et de la demande évapotranspirative de l’atmosphère (E_0). Cet indice met en évidence l’anomalie journalière de la demande évapotranspirative à un endroit donné, agrégée sur une fenêtre temporelle spécifique. La condition d’humidité des sols est ainsi considérée indirectement dans son calcul, qui repose sur la température de l’air, l’humidité, la vitesse du vent et la radiance solaire (Hobbins *et al.*, 2016; McEvoy *et al.*, 2016).

Le déclenchement soudain de ces sécheresses-flash est un défi pour les approches prédictives. Pourtant, il existe un vrai enjeu car ce type de perturbation provoque

des dégâts sur les écosystèmes, et en particulier sur les arbres.

1.3.1 Impact des sécheresses sur les forêts

L'augmentation des températures due aux changements climatiques ne causera pas forcément des sécheresses, mais il est attendu que lorsque celles-ci auront lieu, il est probable qu'elles s'installent plus rapidement et qu'elles soient plus intenses (Trenberth *et al.*, 2014). Les sécheresses-flash à répétition vont bouleverser la temporalité de récupération des écosystèmes en empêchant l'activation de mécanismes d'adaptation (Rennenberg *et al.*, 2006), voire auront plus d'impacts sur le taux de mortalité des forêts que des sécheresses plus longues ; la fréquence ayant plus d'effets délétères que l'intensité (Allen *et al.*, 2015; Sánchez-Pinillos *et al.*, 2022).

Une augmentation des taux de mortalité des arbres à cause des sécheresses et du réchauffement du climat a été observée à l'échelle de la planète (Allen *et al.*, 2010; Choat *et al.*, 2012; Senf *et al.*, 2020). En analysant des données recueillies en Australie entre 2010 et 2011, Evans et Lyons (2013) ont montré que le manque d'eau était la première cause de mortalité des arbres, alors que de nombreuses études, soulignent-ils, n'avaient pas réussi à faire de lien clair entre de forts stress hydriques induits par les sécheresses et la réponse physiologique des arbres. En effet, souvent était avancée la régulation faite par les stomates - pores situés dans le tissu épidermique des feuilles - pour permettre de limiter l'assimilation du carbone et la perte en eau par les feuilles pendant les sécheresses, protégeant ainsi les arbres. En Europe, Senf *et al.* (2020) ont démontré l'effet non-linéaire entre la disponibilité en eau et la mortalité des arbres. Au Canada, la mortalité des arbres a augmenté en moyenne de 4.7 % entre 1963 et 2008 pour la forêt boréale, avec un taux plus élevé dans l'ouest du pays comparativement à l'est (4.9 % par an *vs* 1.9 % par an), et il a été montré que le stress hydrique créé par les sécheresses régionales était le

facteur prépondérant dans cette augmentation de la mortalité (Peng *et al.*, 2011).

Qu'en est-il alors de la capacité de résilience des arbres face aux sécheresses ?

Elle est influencée par de nombreux facteurs. À court terme, suite à une sécheresse induite, ce sont le degré de sévérité de celle-ci et les caractéristiques à l'échelle de la parcelle d'étude (diversité des espèces, densité des arbres et leur surface basale) qui ont la plus grande influence sur la résilience des arbres (Lloret *et al.*, 2022). Cette équipe du CREAF (Barcelone, Espagne) en a conclu que les espèces qui vivaient dans des conditions qui étaient proches de leur optimum climatique étaient plus vulnérables aux épisodes de sécheresses, en comparaison aux populations d'arbres qui se situaient à la limite de leur zone optimale et qui pouvaient alors mettre en place des mécanismes leur permettant de mieux résister à ces changements. En effet, leur structure demandent moins de ressources.

Contrairement à l'agriculture, il est difficile d'établir pour les arbres des seuils de rupture hydrologique (*hydraulic failure*), car ces seuils varient d'un peuplement à l'autre. Sur 226 espèces de forêt étudiées par Choat *et al.* (2012), sur 81 sites mondiaux, 70 % opéraient avec une marge de sécurité (*hydraulic safety margins*) très mince pour faire face aux conséquences induites par les stress hydriques⁹. Cela correspond à une valeur du potentiel hydrique - indice reflétant la disponibilité en eau - inférieure à 1 Mégapascal (10 bar). Or, cette marge est indépendante des précipitations moyennes annuelles, expliquant ainsi pourquoi un déclin des forêts face aux sécheresses n'était pas seulement observable dans les régions arides, mais également dans les forêts humides.

Chaque essence d'arbre répond de façon spécifique aux sécheresses selon les conditions biotiques et abiotiques du site d'étude (Liu *et al.*, 2019; Anderegg *et al.*, 2020).

9. Cette marge de sécurité (HSM) s'exprime comme la différence entre la valeur saisonnière minimale du potentiel hydrique du xylem et la valeur P₅₀ ou PLC₅₀ qui correspond à un pourcentage de perte de conductivité (*per cent loss of xylem conductivity*) égal à 50 % (Anderegg *et al.*, 2016). Le potentiel hydrique du sol est alors critique.

En prenant comme variable d'intérêt la croissance radiale des arbres, Duchesne *et al.* (2020) ont quantifié la sensibilité à la sécheresse de certaines essences de forêts boréales et mixtes durant des périodes montrant des déficits hydriques historiques sur une période de 1961 à 2000 et sur six domaines bioclimatiques. Nous avons colligé dans le tableau 1.2 les résultats obtenus.

TABLEAU 1.2 : Sensibilité à la sécheresse de certaines essences, d'après Duchesne *et al.* (2020).

Sensibilité élevée	Sensibilité moyenne	Sensibilité faible
pin rouge	pin gris	épinette noire
pin blanc	pruche du Canada	sapin baumier
thuya occidental	pépinette blanche	bouleau à papier
peuplier à grandes dents	peuplier faux-tremble	épinette rouge

Deux paramètres ont une forte influence sur la sensibilité des arbres face aux sécheresses : la temporalité de celle-ci et l'évapotranspiration. En effet, si la sécheresse intervient en juin alors des baisses de croissance sont observées pour la plupart des espèces (D'Orangeville *et al.*, 2018b). Ces deux facteurs semblent plus importants que les caractéristiques des sols et du peuplement.

Comme le mettent en avant Fu *et al.* (2022), pour évaluer les impacts des sécheresses sur les écosystèmes, et pour améliorer les modèles prédictifs climatiques, il apparaît *impératif* de quantifier le seuil critique de la teneur en eau du sol par rapport aux stress hydriques que peuvent subir ces écosystèmes. Ce seuil représente le moment où l'évapotranspiration commence à diminuer à cause d'un déficit en humidité du sol, celle-ci étant définie comme la quantité d'eau stockée dans un sol non saturé, rapportée à une unité de volume (Seneviratne *et al.*, 2010; Feldman *et al.*, 2019). Les températures de surface, plus élevées, vont alors impliquer un air plus sec et une augmentation du déficit de pression de vapeur (*vapor pressure*

deficit ou VPD) ayant des conséquences sur l'absorption du CO₂¹⁰.

Pour mieux comprendre cette cascade de processus dynamiques, avec ses constituants et ses phénomènes physiques souvent couplés, nous allons concentrer notre réflexion sur la variable centrale de notre projet : la triade {sol, eau, potentiel hydrique du sol}.

10. Le VPD sera explicité au chapitre 3.

CHAPITRE II

CADRE THÉORIQUE ET PROBLÉMATIQUE

2.1 La triade {sol, eau, potentiel hydrique du sol}

Le fonctionnement et de nombreuses propriétés physiques, chimiques et biologiques des sols sont conditionnés par la structure et la texture de ces derniers, c'est-à-dire, par le mode d'assemblage des particules qui les composent, à l'échelle macro- et microscopique. Cette structure est déterminée par la répartition et l'organisation dans l'espace des matières solides, mais aussi dans les vides associés, occupés par de l'eau ou de l'air (Duchaufour *et al.* (2020), p.77).

2.1.1 Le sol et l'eau

Le sol est composé en partie de pores qui vont constituer un réseau de vides. Cet espace a un rôle prépondérant dans la circulation de l'eau et il peut notamment être caractérisé par différents types de pores :

- des pores grossiers (diamètre $\geq 50 \mu\text{m}$), normalement occupés par de l'air après ressuyage (*drainage*) rapide des pluies qui s'écoulent sous l'effet de la gravité ;
- des pores moyens (diamètre compris entre $10 \mu\text{m}$ et $50 \mu\text{m}$), qui se ressuient très progressivement, et qui sont occupés en alternance par de l'air ou de l'eau ;

- des pores fins (diamètre compris entre $0.2 \mu\text{m}$ et $10 \mu\text{m}$), qui retiennent l'eau dite capillaire, absorbable par les racines ;
- des pores très fins (diamètre $\leq 0.2 \mu\text{m}$), normalement occupés par de l'eau dite liée, non absorbable par les racines.

La porosité d'un sol, en d'autre terme le réseau des vides, va avoir un rôle sur la circulation des fluides - air et eau - facilitant ou empêchant les possibilités de flux et de réserves hydriques. Elle assure donc aux végétaux leur alimentation en eau et la respiration de leurs racines. Les pores grossiers favorisent un écoulement rapide de l'eau, mais d'autres pores, plus petits, vont au contraire la retenir pour ainsi constituer des réserves qui pourront ensuite être utilisées. Enfin, les pores les plus fins contiennent de l'eau qui n'est pas utilisable par les plantes car elle est chimiquement liée.

L'équilibre entre la teneur en eau et en air dans ce système constitué par les pores est primordial pour les végétaux, car en l'absence d'air, le milieu devient anoxique, impliquant alors des conséquences pour la vie dans le sol au niveau racinaire et au niveau des être vivants qui s'y retrouvent (Duchaufour *et al.* (2020), p.100). L'eau, quant à elle, intervient directement et indirectement dans la nutrition des végétaux, en véhiculant des éléments nutritifs dissous. C'est elle aussi qui conditionne la plupart des processus de formation des sols. Elle provient des précipitations, de l'irrigation et des nappes phréatiques.

L'eau des précipitations qui tombe à la surface des sols se répartit en diverses fractions, donnant naissance à différentes formes d'eau, dont l'eau gravitaire et les eaux capillaire et liée, citées plus haut. Sans rentrer dans les détails de cette répartition, il est important de comprendre leur spécificité pour comprendre ce qui va suivre.

La figure 2.1 illustre cette répartition en mettant en évidence trois fractions principales. Tout d'abord, il y a l'eau appelée *eau de ruissellement* qui se retrouve

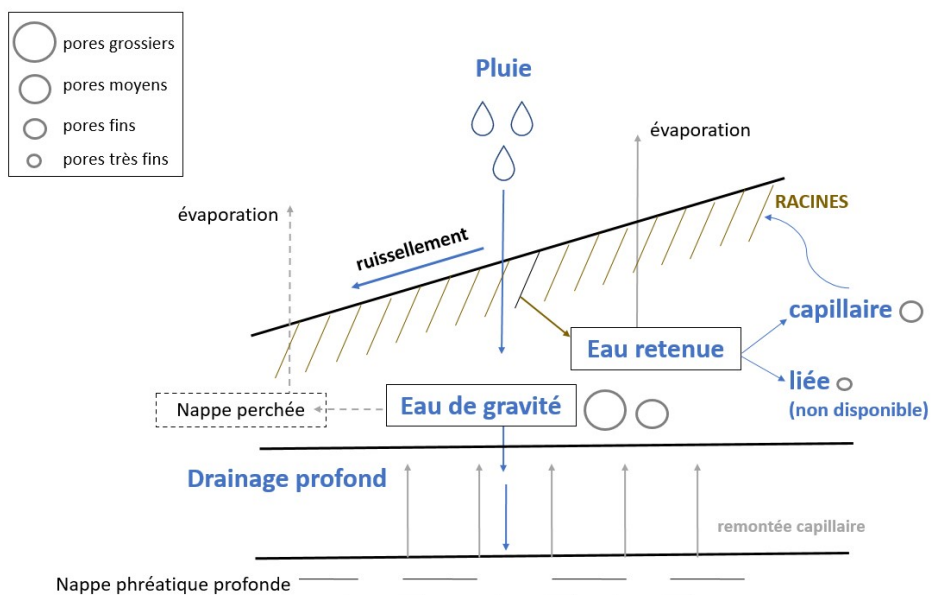


FIGURE 2.1 : Répartition des eaux de précipitations dans le sol (figure inspirée de Duchaufour *et al.* (2020), p.101).

à la surface du sol, circulant parallèlement à ce dernier. Le ruissellement n'est pas constant et affecte surtout les surfaces de forte pente et dénudées de végétation, soumises à des pluies importantes et surtout plus intenses. Les sols qui contiennent des matériaux et/ou des horizons peu perméables, comme l'argile ou les horizons à ortstein fortement cimentés, mais également les sols compactés ou gelés sont sensibles à l'eau de ruissellement.

Il existe également l'*eau de gravité* ou *gravitaire*, entraînée par la pesanteur, qui circule dans les pores grossiers et moyens, le plus souvent verticalement. Cette eau de gravité peut se diviser en deux sous-catégories, en fonction du type de pores qu'elle traverse. En effet, si elle circule à travers les pores qui ont un diamètre plus élevé, elle sera appelée à *écoulement rapide* et cet écoulement aura lieu seulement quelques heures après les pluies. L'eau qui descend plus lentement, sur plusieurs semaines, à travers les pores moyens, va alimenter le drainage profond si le sol est perméable. Cette eau de gravité sera dite à *écoulement lent*. Nous pouvons observer

sur la figure 2.1 l'existence d'une *nappe perchée* qui est souvent temporaire, car elle s'assèche pendant les périodes chaudes. Quand la perméabilité des sols décroît et ne permet plus l'évacuation des eaux de gravité, tous les pores sont occupés par de l'eau, dont une partie peut circuler très lentement, latéralement, formant alors cette nappe perchée, moins profonde et de moins grand volume que la nappe phréatique. Elle est souvent induite par une couche de sol cimentée ou indurée. Enfin, la troisième fraction des précipitations se retrouve sous forme d'*eau retenue*, contenue dans les pores fins et très fins. Les forces capillaires et d'adsorption sont dans ce cas suffisamment élevées pour s'opposer aux forces de gravité. Une partie de cette eau est absorbable par les racines, c'est l'*eau capillaire* qui occupe les pores fins ; tandis qu'une autre partie, l'*eau liée* ou *adsorbée* ne peut être absorbée par le système racinaire car elle est, comme son nom l'indique, chimiquement et fortement liée.

Ainsi, pour les végétaux, deux facteurs se détachent : la disponibilité en eau dans le sol qui repose sur sa structure et son réseau de pores, et leur capacité à extraire cette eau. Ce qui se joue alors sur le plan physique entre l'eau liquide et le sol solide peut être décrit par des interactions de courte portée - des forces attractives entre deux dipôles, de type Van der Waals - et des interactions électrostatiques de plus longue portée (Whalley *et al.*, 2013). Pour pouvoir étudier l'importance de ces facteurs et pour comprendre les conséquences des stress hydriques sur les végétaux, et en particulier dans notre contexte sur les arbres, il faut pouvoir décrire quantitativement le statut hydrique du sol et être capable de le mesurer.

2.1.2 Le potentiel hydrique du sol (Ψ)

Le potentiel hydrique du sol, noté Ψ , est la grandeur physique qui permet de relier l'eau et le sol : il caractérise en effet l'état énergétique de l'eau dans le sol, et cet état permet alors de déterminer la capacité de l'arbre à extraire l'eau par le biais

de son système racinaire. Cette grandeur est mesurable au moyen de différentes sondes et elle s'exprime en unité de pression [kPa] (Or et Wraith, 2001).

Le potentiel hydrique total Ψ_t résulte de différentes contributions énergétiques, d'origines différentes :

- du potentiel gravitaire qui permet l'infiltration de l'eau de pluie par gravité (avec $\Psi_g \geq 0$)
- du potentiel matriciel qui définit la force de rétention de l'eau par le sol (avec $\Psi_m \leq 0$)
- du potentiel osmotique qui a son origine dans les sols salins (Ψ_π)

On écrira alors : $\Psi_t = \Psi_g + \Psi_m + \Psi_\pi$.

En général, sauf dans des contextes de salinité élevée, Ψ_π est nul. Il faut noter que Ψ_g et Ψ_m sont antagonistes, de signes opposés. Le potentiel gravitaire joue un rôle important quand les pores sont saturés en eau. Quand le sol est non saturé, c'est le potentiel matriciel qui intervient. Dans notre contexte, nous allons considérer que le potentiel matriciel représente en majorité le potentiel hydrique du sol, et par abus de langage, appelé Ψ_m , le potentiel hydrique du sol. Dans la littérature, les deux termes sont utilisés de façon identique.

Nous l'avons souligné en évoquant l'importance du réseau de pores du sol, l'accessibilité en eau dépend de la granulométrie du sol. L'eau retenue, i.e. capillaire et adsorbée, est d'autant moins abondante que la granulométrie est grossière, pour un niveau de développement structural comparable (Duchaufour *et al.* (2020), p.103). Néanmoins, cette relation atteint un seuil au-delà duquel la trop grande finesse de la texture va au contraire créer des conditions de faible perméabilité et encourager le ruissellement au détriment de la rétention d'eau. Il est ainsi possible de définir une courbe de rétention en eau, caractérisée par le potentiel, en fonction de la nature du sol.

Il faut différencier la teneur en eau dans le sol (*soil water content* SWC) du potentiel hydrique du sol (*soil matric potential* SMP ou *soil water potential* SWP). La teneur en eau dans le sol indique la quantité d'eau présente dans le sol, mais ne donne pas d'indice sur la quantité d'eau réellement accessible à l'arbre. C'est le potentiel hydrique qui permet de savoir quelle force l'arbre doit exercer pour extraire les molécules d'eau contenues dans le sol.

Si le sol devient de plus en plus sec, l'arbre devra exercer une force plus grande pour extraire les molécules d'eau contenues dans les pores fins, une fois que l'eau des pores plus grossiers ait été extraite. Dans les pores fins, nous l'avons vu, les molécules d'eau et les particules du sol créent des liaisons chimiques fortes. Le potentiel matriciel va donc augmenter négativement, de façon graduelle, à mesure que le sol s'assèche. Plus Ψ_m est négatif, plus le sol est sec. Inversement, lorsque Ψ_m tend vers zéro, le sol est humide et va donc impliquer une force d'extraction requise par les racines plus faible.

Il est important de souligner ici que pour des raisons pratiques, liées à la mesure de la grandeur, le signe négatif du potentiel matriciel est généralement ignoré (Irmak *et al.*, 2016). Il est aussi plus intuitif d'utiliser des valeurs positives pour exprimer une *tension* (Or et Wraith (2001), p.61).

Nous avons choisi d'exprimer le potentiel matriciel du sol avec des valeurs positives. Ces valeurs positives représentent ainsi la tension exercée par la plante pour extraire l'eau du sol. Ainsi, dans notre référentiel, le potentiel hydrique du sol *augmentera* quand le sol s'asséchera (il deviendra en réalité encore plus négatif), et il *diminuera* quand le sol s'humidifiera (il se rapprochera dans ce cas de zéro).

Il faut néanmoins ajouter que le potentiel et la teneur en eau du sol ne sont pas uniquement fonction de la taille des pores et des propriétés physiques de l'interface air-eau-sol. Les propriétés minéralogiques du sol sont également à considérer, telle que la teneur en particules argileuses comme les phyllosilicates à la structure

atomique lamellaire. Ces particules ont une influence sur la capacité de rétention et de stockage de l'eau du sol.

Une autre relation peut être obtenue en considérant les deux chemins possibles suivants : (1) il est possible de partir d'un sol initialement saturé en eau et y appliquer une pression pour le rendre alors non-saturé ; (2) inversement, un sol sec peut devenir de plus en plus humide, jusqu'à saturation. Ces deux chemins sont illustrés respectivement en rouge et en vert à la figure 2.2. Ils suivent, pour l'un, la courbe d'assèchement, pour l'autre, la courbe d'humidification. Or, nous observons que les deux chemins ne produisent pas la même relation entre le potentiel hydrique et la teneur en eau : un phénomène d'hystérésis régit ainsi la relation entre le sol et l'eau. Pour une valeur donnée du potentiel hydrique du sol, la teneur en eau située sur la courbe d'assèchement est supérieure à celle située sur la courbe d'humidification. La réaction du système {sol, eau, potentiel hydrique du sol} n'est donc pas totalement réversible. De plus, l'existence de ce phénomène d'hystérésis présage des relations non-linéaires au sein de notre triade d'intérêt.

2.1.3 Mesure du potentiel hydrique du sol

Dans le cadre de notre projet, la sonde granulométrique matricielle (*granular matrix sensor*) Watermark¹ a été choisie pour mesurer le potentiel hydrique du sol.

Différents types de sonde reposant sur des technologies différentes existent. Selon le type de sonde utilisé, la précision et l'ordre de grandeur du potentiel mesuré sont différents. Les sondes granulométriques permettent d'obtenir des valeurs de l'ordre de -200 kPa à 0 kPa. Ces sondes ont une exactitude inférieure à celle obtenue avec d'autres sondes, telles que les sondes WP4C² qui permettent d'obtenir des valeurs

1. GMS, Watermark Soil Moisture Sensor, Irrometer Co., Riverside, CA.

2. Voir le site internet de Wafer Sensor Inc.

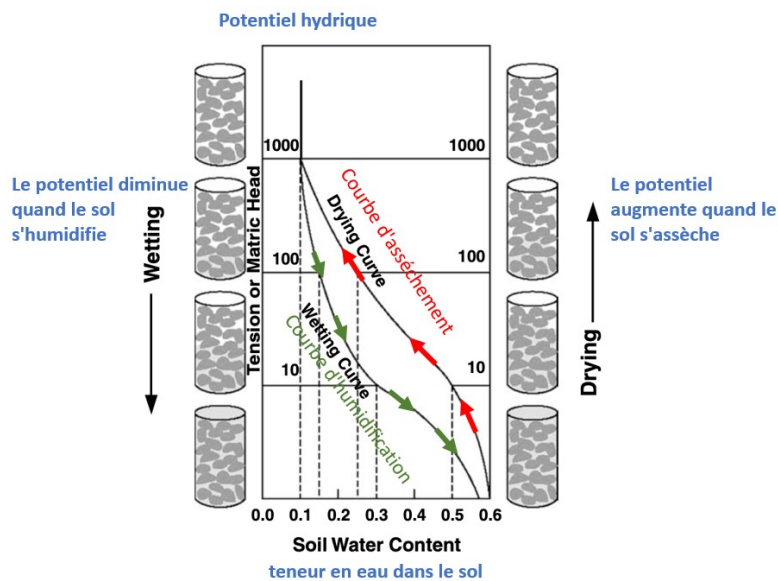


FIGURE 2.2 : Relation sol-eau et phénomène d’hystérésis (figure extraite et adaptée de Or et Wraith (2001), p.78). Le potentiel hydrique est considéré comme étant positif, tel que décrit dans le référentiel que nous avons choisi.

de l’ordre de grandeur du MPa. Cependant, ces dernières sont plus adaptées à des mesures réalisées en laboratoire. En effet, le temps de lecture est de 10 à 15 minutes et les dimensions de l’appareil (24.1 x 22.9 x 8.9 cm) ne permettent pas une utilisation facile sur le terrain, surtout dans des sols à forte pierrosité.

De plus, dans notre contexte, la sonde Watermark a été choisie car l’objectif était d’installer un grand nombre de sondes pour capturer la variabilité sur un large dispositif de recherche. Le coût d’une sonde a aussi été une motivation dans la sélection du modèle car, dans ce contexte, la quantité était un facteur de choix. Ce qui nous importait dans le cadre de notre projet, ce sont les variations du potentiel hydrique. Les sondes Watermark permettent la mesure de ces variations. L’échantillonnage aux 15 minutes des mesures, ainsi que la possibilité d’en utiliser un grand nombre assez facilement sur le terrain pour couvrir une large gamme de conditions et mieux capturer les variations spatiales et temporelles, compensent

en bonne partie leur moins bonne exactitude en comparaison avec d'autres sondes. Enfin, les sondes granulométriques choisies ont été testées et validées à multiples occasions et dans d'autres contextes qu'en milieu forestier (Shock *et al.*, 2013),(Irmak *et al.*, 2016, 2022).

Nous décrivons ici succinctement la sonde de type Watermark utilisée pour mesurer le potentiel hydrique du sol. La figure 2.3 extraite de Irmak *et al.* (2016) illustre l'aspect général de la sonde. Les électrodes présentes dans la sonde sont incorporées

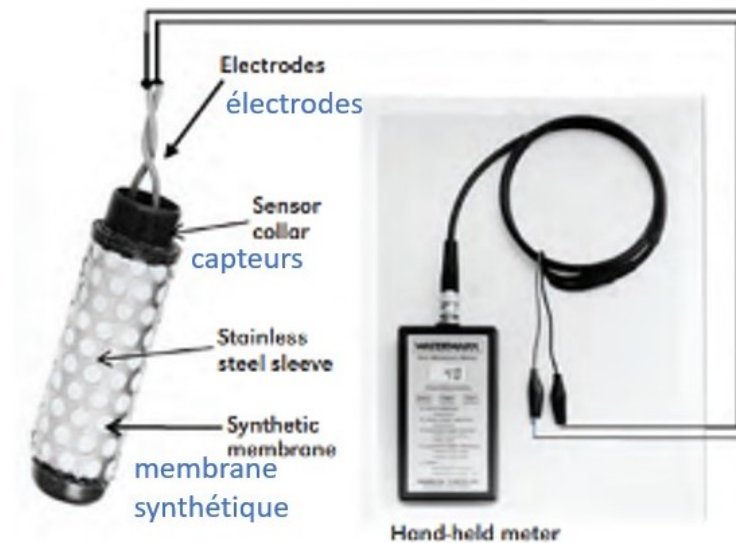


FIGURE 2.3 : Sonde granulométrique matricielle de la marque Watermark, extraite de Irmak *et al.* (2016)

dans un matériau granulaire, i.e. constitué d'un grand nombre de particules solides distinctes, au-dessus d'une plaquette de gypse ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$). Sous la plaquette, un espace matriciel permet à l'eau d'entrer et de sortir de la sonde. Le gypse dissous dans l'eau est un bon conducteur. Ainsi, quand la sonde contient beaucoup d'eau – quand le sol contient lui-même beaucoup d'eau – le courant électrique à travers la sonde se propage bien. Quand le sol commence à sécher, la sonde devient moins

humide, et la résistance au courant augmente. La sonde granulométrique mesure directement cette résistance électrique R , à l'intérieur de la matrice granulaire, et la transforme en tension Ψ grâce à l'équation développée par Shock *et al.* (1998) et reprise dans le rapport technique du constructeur Irrometer³. Cette équation se définit à partir de la résistance et de la température du sol, respectivement R et T_{sol} :

$$\Psi(R, T) = \frac{-3,213R - 4,093}{1 - 0,009733R - 0,01205T_{sol}}$$

avec Ψ en [kPa], R en [$k\Omega$] et T_{sol} en [$^{\circ}C$].

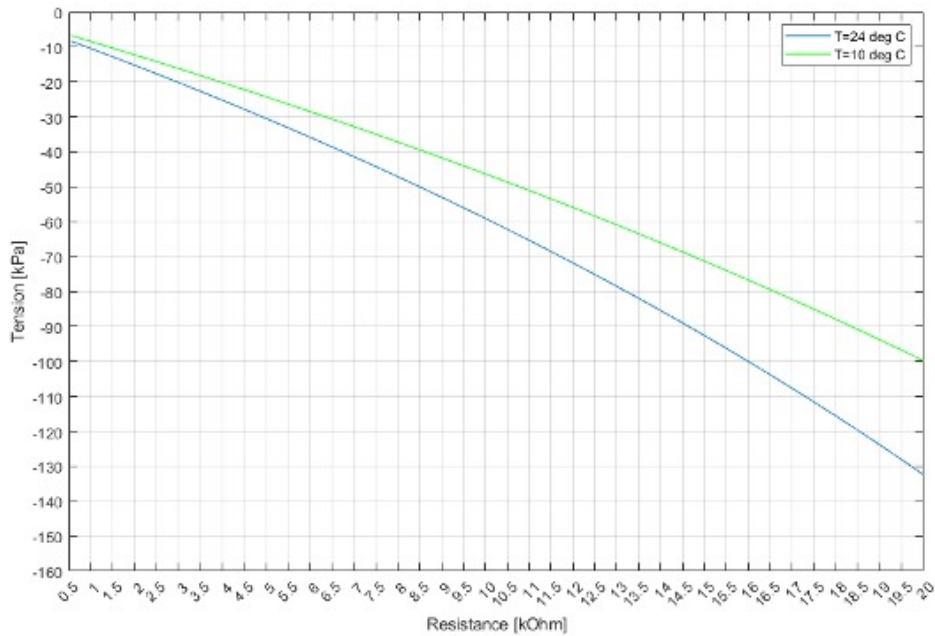


FIGURE 2.4 : Potentiel hydrique du sol en fonction de la résistance, pour deux valeurs de températures fixées. En bleu $T=24^{\circ}C$ et en vert $T=10^{\circ}C$ (graphique réalisé avec Matlab).

La figure 2.4 permet de visualiser l'influence de la température du sol sur la mesure

3. Rapport technique accessible en ligne

du potentiel hydrique. Nous avons choisi à titre d'exemple une température de 24 °C et une température de 10 °C. La relation n'étant pas tout à fait linéaire, la différence se fait sentir davantage quand le sol devient sec et que la résistance augmente. La documentation technique du constructeur Irrrometer souligne que la calibration proposée pour obtenir la correspondance entre la résistance, mesurée en $k\Omega$, et le potentiel hydrique en kPa, a été faite par défaut pour une température du sol égale à 24 °C. Le constructeur préconise d'adapter l'équation pour obtenir des mesures plus exactes du potentiel hydrique. Faisant écho aux recommandations du constructeur, Irmak *et al.* (2016) ont également mis en évidence la nécessité d'ajuster le potentiel hydrique aux températures du sol, en l'augmentant de 1 % pour chaque degré inférieur à 70 °F (ce qui correspond à 21 °C), ou en le diminuant de 1 % pour chaque degré supérieur à 70 °F. Ils proposent l'équation suivante pour l'ajustement du potentiel hydrique du sol mesuré :

$$\Psi_{adj} = \Psi - (T_{sol} - 70 \text{ } ^\circ F) \times 0,01 \times \Psi.$$

La température du sol T_{sol} s'exprime en [°F] et Ψ , Ψ_{adj} en [kPa].

Entre mai et novembre, les températures du sol mesurées à la SBL se situent généralement dans l'intervalle [5;15] °C, en dessous de la température de 21 °C suggérée par Irmak *et al.* (2016). *Ainsi, pour ce projet, nous avons opté pour un ajustement de chaque valeur du potentiel hydrique par rapport à la température du sol mesurée au même moment.*

La température d'un sol est gouvernée par la nature de ses constituants et leur organisation, mais aussi par sa teneur en eau. Il faut également ajouter l'influence de facteurs météorologiques externes, ainsi que les échanges de chaleur entre le sol et l'atmosphère (Duchaufour *et al.* (2020), p.123). Ces flux de chaleur vont être modulés par la quantité d'énergie solaire reçue à la surface du sol. Cette énergie

est mesurée à travers l'irradiance solaire (anglicisme provenant du terme *solar irradiance*) et s'exprime en $[\text{W m}^{-2}]$. Cette grandeur dépend de la saison considérée, de la latitude, de l'angle d'incidence des rayons solaires, de la couverture nuageuse, de l'albédo de la surface du sol (son pouvoir réfléchissant), et enfin de la nature de l'interface sol-atmosphère qui est définie par le type de la litière forestière, du manteau neigeux, etc. Lorsqu'elle est intégrée sur une période de temps, le terme d'irradiation solaire est utilisé.

La teneur en eau du sol conditionnera son comportement thermique. La compréhension physique de ces phénomènes de flux de chaleur se fait principalement à travers deux grandeurs aux effets antagonistes : la capacité thermique du sol C_p qui caractérise sa capacité à stocker de la chaleur, et sa conductivité thermique qui va moduler sa capacité à conduire la chaleur.

La capacité thermique correspond à la quantité de chaleur nécessaire pour augmenter 1 g du corps considéré de 1 degré Kelvin. L'eau, contrairement à l'air, a une capacité thermique élevée, égale à $4.18 \text{ J g}^{-1} \text{ K}^{-1}$. Ainsi, cela demandera davantage d'énergie pour faire varier la température d'un sol chargé en eau, comparativement à l'air qui s'échauffe facilement pendant le jour, et se refroidit rapidement pendant la nuit. Un sol humide va donc mieux résister aux changements rapides de température.

Ainsi, considérant ces deux phénomènes, si un sol est humide alors il va mieux conduire la chaleur (augmentation de sa conductivité thermique), mais si la concentration en eau augmente, la capacité thermique du sol va également devenir décisive et va ralentir les variations de température. Dans le cas où la température du sol augmente, le stockage de chaleur va être favorisé par un sol humide, pour assurer une stabilité thermodynamique régie par le deuxième principe de la thermodynamique. Selon la nature du sol, la conductivité et la capacité thermique vont varier différemment (Abu-Hamdeh, 2003; Arkhangelskaya et Lukyashchenko, 2018). La teneur en eau d'un sol exercera une influence directe sur sa température, cette

dernière étant par conséquent une variable que nous intégrerons à notre étude.

Nous avons mis en évidence ci-dessus l'importance de l'équilibre {sol, eau, potentiel hydrique du sol}, et isolé la variable principale de notre projet. Au chapitre 1, nous avons défini plus précisément ce que nous entendions par sécheresse et sécheresse-flash en exposant leurs conséquences à l'échelle des forêts. Dans un contexte où l'eau est en déficit, nous explorons maintenant les conséquences des stress hydriques au niveau physiologique des végétaux.

2.2 Le stress hydrique

D'après le dictionnaire Larousse, le stress est l'état réactionnel d'un organisme soumis à une agression brusque. Ainsi, le stress hydrique se réfère au stress subi par des végétaux placés dans un environnement où la quantité d'eau qu'ils transpirent est supérieure à la quantité qu'ils prélèvent.

Utiliser le mot *stress* implique plusieurs notions. Tout d'abord, la notion de seuil, en particulier pour les variables météorologiques dont les valeurs vont être liées à la quantification de différents niveaux de stress. Il y a ensuite la notion de durée, car les conséquences de stress hydriques seront plus ou moins importantes selon le nombre de jours pendant lesquels les conditions de stress seront appliquées. La notion de dommage causé est aussi à considérer. Pour mieux anticiper ces dommages, il s'avère fondamental d'éluder les mécanismes physiologiques déclenchés par les végétaux soumis à ces stress, mais aussi ceux pour y résister.

2.2.1 Une cascade d'événements

Que ce soit au niveau individuel ou au niveau d'un écosystème entier, deux processus fondamentaux régissent la réponse des végétaux face aux actions combinées du climat, du sol et de la végétation (figure 2.5 (A)).

Le premier est centré sur la dynamique hydrique du sol, le second régule les impacts des déficits hydriques au niveau physiologique. Comme le soulignent Rodríguez-

Iturbe et Porporato (2005), les conditions favorables aux déficits hydriques ne sont pas tant dues aux manques de précipitations qu'à des conditions d'accessibilité en eau à travers le sol. Pour assurer leur croissance et leur survie, les végétaux ont besoin d'un niveau adéquat d'eau. Un déficit en eau a des conséquences sur le potentiel hydrique, entraînant alors une cascade d'événements qui mène à un état de stress hydrique, impliquant des dommages qui peuvent être irréversibles. Nous l'avons exposé précédemment et schématisé sur la figure 2.5 (B).

Si nous revenons en amont de cette chaîne d'événements, nous pouvons revenir aux interactions climat-sol, et isoler les liens entre le déficit hydrique, la demande évaporative et l'évapotranspiration.

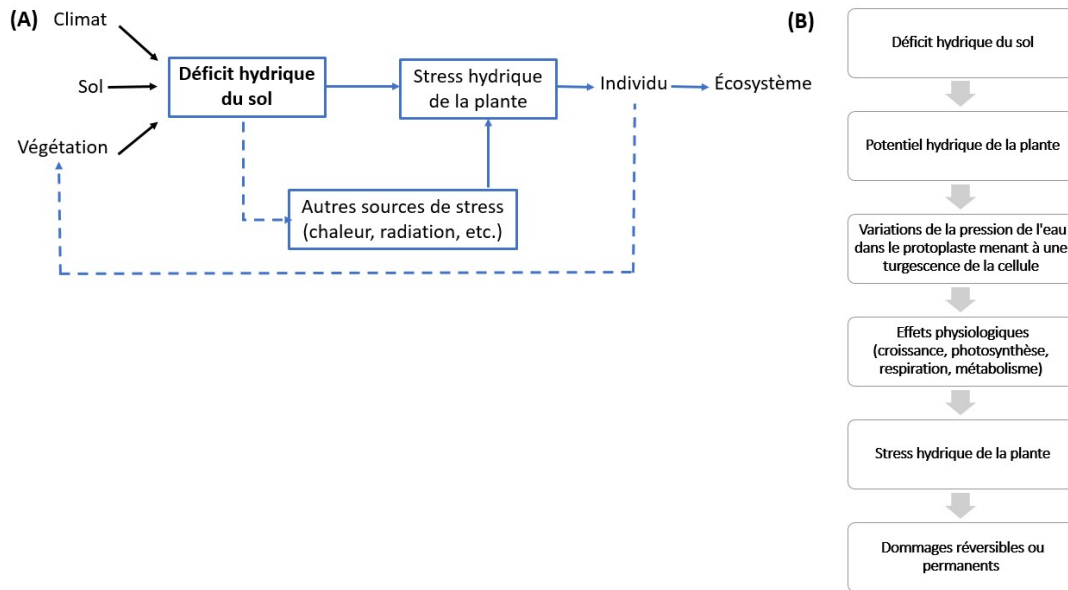


FIGURE 2.5 : Actions combinées du climat, du sol et de la végétation sur le déficit hydrique du sol (A) et cascade d'événements menant à un état de stress hydrique (B). Figure adaptée de Rodríguez-Iturbe et Porporato (2005), p.87. Traduction libre.

2.2.2 Boucles de rétroactions positives

Pour clarifier les mécanismes, évoqués au chapitre 1, qui lient la demande évaporative et l'évapotranspiration aux déficits hydriques, nous avons repris et adapté la figure de Pendergrass *et al.* (2020) issue de Lukas *et al.* (2007). La figure 2.6 (A) illustre l'évolution de la demande évaporative (ET_0), de l'évapotranspiration (ET) et de la teneur en eau du sol pour un sol humide et pour un sol sec. La figure 2.6 (B) met en parallèle les boucles de rétroaction positive qui régissent les phénomènes observés.

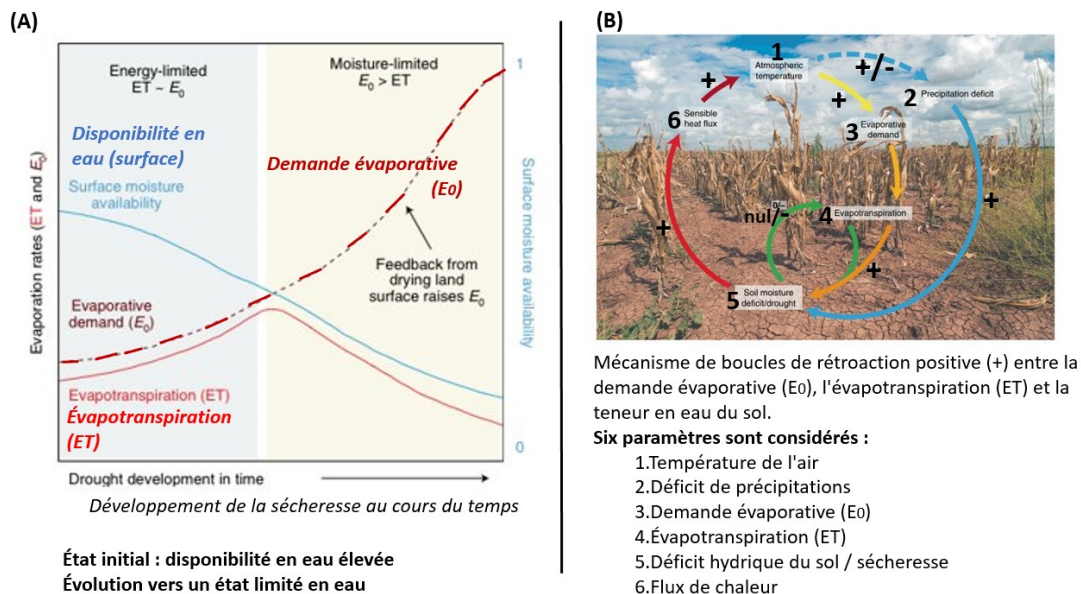


FIGURE 2.6 : Liens entre la demande évaporative, l'évapotranspiration et le déficit hydrique (A) et description des mécanismes reposant sur des boucles de rétroaction positive (B). Figure adaptée de Pendergrass *et al.* (2020) et issue de Lukas *et al.* (2007).

Partons d'un sol humide, représenté par la zone en bleu sur la figure 2.6 (A) : il y a alors une variation concomitante de la demande évaporative et de l'évaporation (courbes en rouge). Ainsi, quand la demande évaporative de l'atmosphère augmente, alors l'évapotranspiration augmente également (deux boucles de rétroaction

positive représentées en jaune et orange sur la figure 2.6 (B)).

Quand l'évapotranspiration devient élevée et qu'il ne pleut pas (déficit de précipitations, courbes en bleu A et B), la teneur en eau du sol n'est plus suffisante pour fournir de l'eau à l'évapotranspiration (boucle de rétroaction positive en vert). L'eau devient donc le facteur limitant. Dans ces conditions, l'évapotranspiration se met à diminuer, telle que représentée par la courbe (A), ce qui implique l'augmentation du flux de chaleur (boucle de rétroaction positive en rouge) qui va augmenter la température de l'air (boucle en rouge foncé), mais aussi le VPD, et donc, par un autre mécanisme de rétroaction positive, augmenter la demande évaporative (boucle en jaune). Une demande évaporative anormalement élevée peut mener à un stress hydrique à la surface du sol, et ultimement à une sécheresse. Quand cet état s'est installé, la surface étant sèche, l'air au-dessus de celle-ci va se réchauffer et devenir plus sec, augmentant alors davantage la demande évaporative.

Les conséquences pour les végétaux, et en particulier les arbres, soumis à ces stress hydriques dépendent de l'intensité et de la durée du déficit hydrique.

2.2.3 Différents niveaux de stress hydriques

Un déficit hydrique ne va pas forcément impliquer un état de stress. Kögler et Söffker (2020) ont isolé trois zones de stress hydrique qui dépendent de l'intensité et de la durée du déficit (figure 2.7).

La première zone présente un faible déficit hydrique, de courte durée, se situant sous le seuil dit *de réponse*. Dans cette zone, la plante ne met pas en place de mécanisme adaptatif. Lorsque le déficit hydrique augmente et dépasse ce premier seuil, la plante est alors confrontée à un stress qualifié de moyen (zone 2 sur la figure 2.7). Jusqu'à une certaine valeur de ce déficit - qui varie selon les espèces - appelé seuil *de récupération*, la plante subit ce que les auteurs nomment un stress positif. Ce stress va impliquer une réponse permettant à la plante de s'adapter à ce

Trois zones mises en évidence :

(1) Zone sans stress hydrique, sous le seuil de réponse, pas de réponse adaptative mise en place.

(2) Zone de stress moyen, entre les seuils de réponse et de récupération (zone de stress positif): réponse adaptative, pas de dommages.

(3) Zone de stress élevé, au-dessus du seuil de récupération : réponse adaptative (mémoire) avec des dommages pouvant causer la mort du végétal.

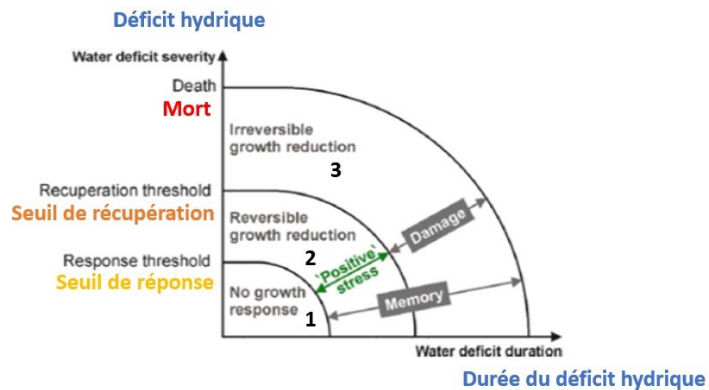


FIGURE 2.7 : Mise en évidence de différents niveaux de stress hydrique, d'après Kögler et Söffker (2020).

déficit hydrique sans entraîner de dommages. Quand ce seuil de récupération est dépassé, le déficit hydrique devient plus élevé et la plante n'arrive plus à mettre en place de mécanismes adaptatifs suffisamment efficaces et les dommages qu'elle subit deviennent irréversibles, pouvant même mener à sa mort.

L'objectif de cette approche, testée sur des plants de maïs, est de pouvoir définir des zones plus fines qui vont au-delà du moment où le stress est initié, en mettant en avant des intervalles hydriques qui peuvent être pertinents à exploiter et qui restent bénéfiques pour les végétaux. Kögler et Söffker (2020) font le parallèle avec le modèle de surcompensation développé dans le domaine du sport, lequel repose sur des séquences d'entraînement et de repos (Jakovlev et Pickenhain, 1977). En effet, un stimulus, alors considéré comme un stresseur, a pour effet initial une fatigue du muscle sollicité. Après une période de récupération, le système développe un niveau d'efficacité supérieure. Si le stimulus est répété durant une période spécifique, le niveau ainsi atteint peut être maintenu (effet d'entraînement), sinon, il revient à son niveau initial.

La paramétrisation expérimentale de ces seuils reste un enjeu majeur qui dépend

de chaque espèce végétale considérée, et le défi est de pouvoir détecter les déficits hydriques à un stade précoce du manque d'eau. Nous l'évoquons avec le concept de seuil de rupture hydrologique et l'étude de Choat *et al.* (2012) (chapitre 1). Il faut également considérer la temporalité des réponses des plantes soumises à des stress hydriques qui se font à court, moyen et long termes (Naikwade, 2023) et selon la nature des processus adaptatifs impliqués dans la réponse. Face à des stress hydriques, les plantes vont moduler leur métabolisme, leurs caractéristiques morphologiques et l'expression de leurs gènes (Janiak *et al.*, 2018).

De façon générale, face à ce type de stress, trois mécanismes peuvent être mis en place (Byeon *et al.*, 2023) : la fuite, l'évitement et la tolérance (*escape, avoidance, tolerance*).

2.2.4 Mécanismes adaptatifs

Sans approfondir leurs aspects physiologiques, nous allons préciser les trois mécanismes qui ont été avancés.

Le premier, appelé mécanisme de fuite, repose sur un cycle de vie rapide de la plante, qui va généralement être complété avant l'installation de la sécheresse.

Le second, l'évitement, exploite la capacité de la plante à maintenir sa teneur en eau pendant la sécheresse, par exemple à travers son système racinaire qui présente des racines plus profondes ou en fermant ses stomates. La fermeture des stomates permet de maintenir un potentiel hydrique plus faible (i.e. un niveau d'eau plus élevé) pendant les périodes de stress, mais elle implique un état physiologique stagnant : la transpiration et la photosynthèse diminuent alors dès un stress modéré, ce qui mène à une privation de carbone (McDowell *et al.*, 2008).

Les espèces isohydriques, comme le pin gris, vont adopter ce type de stratégie (Léger-Beaulieu, 2018). Les espèces anisohydriques, comme l'épinette, vont laisser leurs stomates ouverts pour continuer à puiser l'eau dans le sol. Le potentiel hydrique va donc augmenter (i.e. qu'il va devenir encore plus négatif) tandis que

les activités biologiques de la plante vont être maintenues.

Enfin, le troisième mécanisme possible face à un stress hydrique, la tolérance, va permettre à la plante de maintenir une homéostasie hydrique dans ses cellules, en régulant l'expression de certains gènes. Ces gènes participent notamment aux ajustements osmotiques en ciblant la modification des parois cellulaires. De plus, l'augmentation dans les osmolytes (petites molécules de matière organique, solubles dans la solution intracellulaire) de la concentration en sucres solubles et l'augmentation de certains acides aminés comme la proline vont permettre la diminution du potentiel hydrique (la tension devient moins forte) et améliorer ainsi la tolérance aux sécheresses de la plante.

C'est ainsi que des mécanismes moléculaires vont permettre à certains végétaux de devenir plus efficaces à résister aux stress hydriques. S'ajoutent à ces mécanismes des transformations au niveau de l'expression génétique des végétaux. Cette voie d'adaptation repose sur la mémoire épigénétique de ces derniers qui permet une modification de leur phénotype suite à des interactions avec l'environnement (Amaral *et al.*, 2020).

Ce concept de *mémoire* en écologie apparaît dans un contexte de réponses à des perturbations et à la formation d'écosystèmes résilients (Johnstone *et al.*, 2016; Itter *et al.*, 2019). En effet, il a été montré que les écosystèmes peuvent conserver une mémoire des conditions passées auxquelles ils ont été confrontés (Bruce *et al.*, 2007).

2.2.5 Mémoire face aux stress hydriques

À l'échelle des forêts, Anderegg *et al.* (2015) ont mis en évidence l'existence de cette mémoire face à des stress hydriques, lors de sécheresses, en analysant 1338 sites forestiers répartis à travers les États-Unis et l'Europe. L'étude des impacts de sécheresses sur la croissance des arbres a permis de conclure à l'existence d'effets hérités (*legacy effects*) qui sont plus élevés chez des essences qui ont une plus faible

marge de sécurité hydrologique (HSM). La définition de cette marge de sécurité a été amenée au chapitre 2. Elle se rapporte à la différence, en kPa, entre la valeur saisonnière du potentiel hydrique du xylem et la valeur P_{50} et P_{90} , respectivement pour les gymnospermes et les angiospermes. Lorsque la valeur P_{50} est élevée, la marge de sécurité devient faible et la plante est plus susceptible d'atteindre un état qui peut avoir un impact négatif sur le xylem qui assure le transport de l'eau. La cavitation du xylem qui correspond à une rupture des colonnes d'eau dans ses vaisseaux permet alors le remplissage de ces derniers par de l'air, pouvant causer une embolie gazeuse, qui rend impossible le transport de l'eau (Granier et Cochard, 2023).

L'importance des événements passés semble avoir des conséquences plus intenses et plus longues sur les gymnospermes (comme les feuillus) soumis à des stress hydriques, en comparaison aux angiospermes (DeSoto *et al.*, 2020) qui résistent mieux aux sécheresses (Urli *et al.*, 2013). La mémoire des arbres par rapport aux variations climatiques est critique dans leur réponse, en terme de croissance, et la prendre en compte dans les modèles peut influencer les prédictions sur la récupération des arbres à la suite d'une sécheresse (Peltier *et al.*, 2022).

En amont de ce concept de mémoire écologique se situe l'idée d'amorçage (*priming*), i.e. que les expositions antérieures à des stress vont renforcer la résistance aux expériences futures, en activant notamment des mécanismes plus intenses et/ou plus rapides. Pour des stress abiotiques, comme les stress hydriques, le terme d'acclimatation (*acclimation* ou *hardening*) sera utilisé (Bruce *et al.*, 2007). Il est très proche de l'idée de zone de récupération introduite par Kögler et Söffker (2020). La durée entre l'événement d'amorçage et les expositions subséquentes pendant laquelle la réponse de la plante aura lieu est importante. La plante va alors stocker les informations relatives aux expositions passées (Trewavas, 2003), ce qui peut s'apparenter à une forme de mémoire, ou, pour éviter la connotation

anthropomorphe du terme, une empreinte du stress (*stress imprint*), telle que nommée par Bruce *et al.* (2007). Cette empreinte est définie comme une modification biochimique ou génétique du végétal qui intervient après une exposition au stress et qui va permettre une réponse différente et mieux adaptée aux stress futurs. Une adaptation des arbres sur le long terme à des conditions plus sèches pourrait les aider à mieux récupérer, et plus rapidement (Yao *et al.*, 2022). C'est ce que Li *et al.* (2021) ont montré en étudiant des genévriers du Plateau Tibétain. De plus, selon le degré de stress et les processus physiologiques impliqués, la vitesse de récupération va varier. Par exemple, après hydratation, le potentiel hydrique des feuilles va diminuer après quelques jours (il devient donc moins négatif), mais les variables liées aux échanges gazeux vont être décalés de plusieurs jours (Ruehr *et al.*, 2019).

Cette mémoire des écosystèmes face aux événements passés façonne la résilience des végétaux (Johnstone *et al.*, 2016). Plus la fréquence des sécheresses augmente, plus les écosystèmes semblent avoir une grande résilience écologique (Yao *et al.*, 2022). Cette conclusion a été confirmée par Cole *et al.* (2014) pour la forêt tropicale. Ce résultat est en phase avec l'existence d'une zone de stress positif mise en évidence à la figure 2.7.

Il faut également considérer la mémoire hydrique du sol (*soil moisture memory*), contrôlée en partie par l'évapotranspiration et la demande évaporative (Koster et Suarez, 2001), mais aussi par les communautés microbiennes du sol qui affectent le carbone et l'azote disponibles (Jacques *et al.*, 2021). Les sécheresses périodiques contribuent ainsi à la mémoire écologique des sols, qui peuvent également devenir plus résilients face à des stress hydriques (Canarini *et al.*, 2021). Orth et Seneviratne (2012) ont quantifié cette mémoire hydrique du sol sur cinq sites européens. Selon les saisons, l'effet de mémoire - que les auteurs traduisent en terme d'évolution de la corrélation au cours du temps - pouvait durer jusqu'à 40 jours. Cet effet s'est avéré

plus important lors de conditions sèches sur des sites humides ; la teneur en eau initiale et les précipitations accumulées suite à l'épisode de sécheresse semblaient être les variables ayant le plus d'influence dans le modèle proposé par les deux auteurs.

De nombreux paramètres interviennent donc dans la définition et l'étude de la mémoire écologique. Dans le cas spécifique de la mémoire au stress hydrique, Jacques *et al.* (2021) ont développé un cadre conceptuel visant à englober tous les processus qui ont lieu au niveau du sol, des cellules et des gènes des plantes. L'originalité de ce cadre réside également dans la considération temporelle des événements, en incorporant des phases d'adaptation, de récupération, voire d'oubli du stress. Citant Sedaghatmehr *et al.* (2019), Jacques *et al.* (2021) rappellent que dans le cas d'un stress liée à la chaleur, il faut 4 jours pour effacer l'impact immédiat du stress.

Pour reprendre la question de Bruce *et al.* (2007), une fois les mécanismes de mémoire appréhendés, il devient essentiel pour les écosystèmes de se demander comment effacer ces empreintes de stress qui s'accumulent au cours du temps, pour certaines espèces de façon bénéfique, pour d'autres plus sensibles, de façon létale.

Nous avons défini dans ce chapitre la variable centrale de notre projet, le potentiel hydrique du sol, mais aussi isolé les variables exogènes au modèle qui sera développé. De plus, l'importance de la temporalité des mécanismes mis en place par les végétaux pour faire face à des stress hydriques a été mis en évidence à travers, notamment, ce concept de mémoire écologique.

D'un point de vue méthodologique, il faut donc être capable de traduire mathématiquement la dépendance temporelle qui existe entre nos variables d'intérêt. Une analyse statique des données relatives au potentiel hydrique du sol ne permettrait pas d'expliquer des interactions qui se déploient à des échelles temporelles diffé-

rentes. C'est ainsi qu'un angle d'étude reposant sur l'analyse de séries temporelles a été choisi.

2.3 Angle d'étude : les séries temporelles

Pour saisir l'évolution temporelle du potentiel hydrique du sol à travers le temps, il nous a semblé fondamental de considérer une approche d'analyse des données qui considère le temps comme une variable explicative. D'un point de vue écologique, nous pourrions néanmoins nuancer notre considération. En effet, si nous prenons le temps au sens mécanistique, son passage - en l'absence de tout mécanisme - n'induit aucun changement du sol (Jenny, 1980). Ainsi, il pourrait s'avérer moins décisif de l'inclure spécifiquement dans les analyses. En définissant les mesures collectées comme des séries temporelles, nous faisons l'hypothèse que des changements ont lieu au cours du temps, en nous appuyant sur une temporalité de l'hydrologie des sols moins longue que celle de leur genèse.

Une série temporelle est une séquence d'observations prise de façon séquentielle dans le temps (Box *et al.*, 2008). La particularité intrinsèque des séries temporelles est que, généralement, les observations adjacentes sont dépendantes. La nature de cette dépendance entre les observations est au coeur de l'analyse des séries temporelles. Cette approche nécessite alors l'utilisation de méthodes adaptées qui reposent sur des analyses statistiques spécifiques. À partir de mesures arrangées en ordre chronologique, il est possible d'analyser des tendances passées pour prédire des comportements futurs.

Historiquement, le statisticien anglais George Box a permis le développement d'un modèle d'analyse dédié aux séries temporelles : la démarche de Box-Jenkins. Notre méthodologie reposant en partie sur ce modèle, nous le détaillerons au chapitre 3.

2.4 Problématique centrale et formulation des questions de recherche

Nous pouvons résumer les concepts centraux définis dans ce chapitre au moyen d'un organigramme (figure 2.8). Au coeur de notre étude : le potentiel hydrique du sol.

Au regard du contexte général que nous avons précisé et du cadre théorique dans lequel s'insère notre projet, nous pouvons alors établir la problématique suivante :

Dans un contexte de changements globaux, quelle serait l'évolution temporelle du potentiel hydrique du sol quand la nouvelle normalité climatique des érablières à bouleau jaune au Québec sera atteinte ?

Cinq questions peuvent être formulées pour tenter de répondre à notre problématique.

1. Quelles sont les variables climatiques considérées dans notre étude qui ont des conséquences importantes sur l'évolution temporelle du potentiel hydrique du sol mesuré dans la forêt de la SBL entre 2017 et 2020 ?
2. Existe-t-il différents états de ce potentiel hydrique du sol et quelles sont les fréquences et les probabilités associés aux changements d'états ainsi mis en évidence ?
3. Est-ce qu'un modèle prédictif de type ARIMAX, développé avec des données provenant de la SBL entre 2017 et 2020, permet de prédire l'évolution temporelle du potentiel hydrique du sol de ce même site en considérant l'influence de variables climatiques exogènes ?
4. Pouvons-nous mettre en évidence des différences, entre divers types de couverts forestiers étudiés, dans les variations du potentiel hydrique du sol dans un contexte spécifique de stress climatiques comme le stress hydrique ?
5. Dans ce contexte de stress hydrique, sommes-nous capables de prédire l'évo-

lution temporelle du potentiel hydrique du sol du site d'intérêt ?

L'objectif de notre projet est de pouvoir développer, tester et valider un cadre d'analyse reposant sur les données collectées à la SBL.

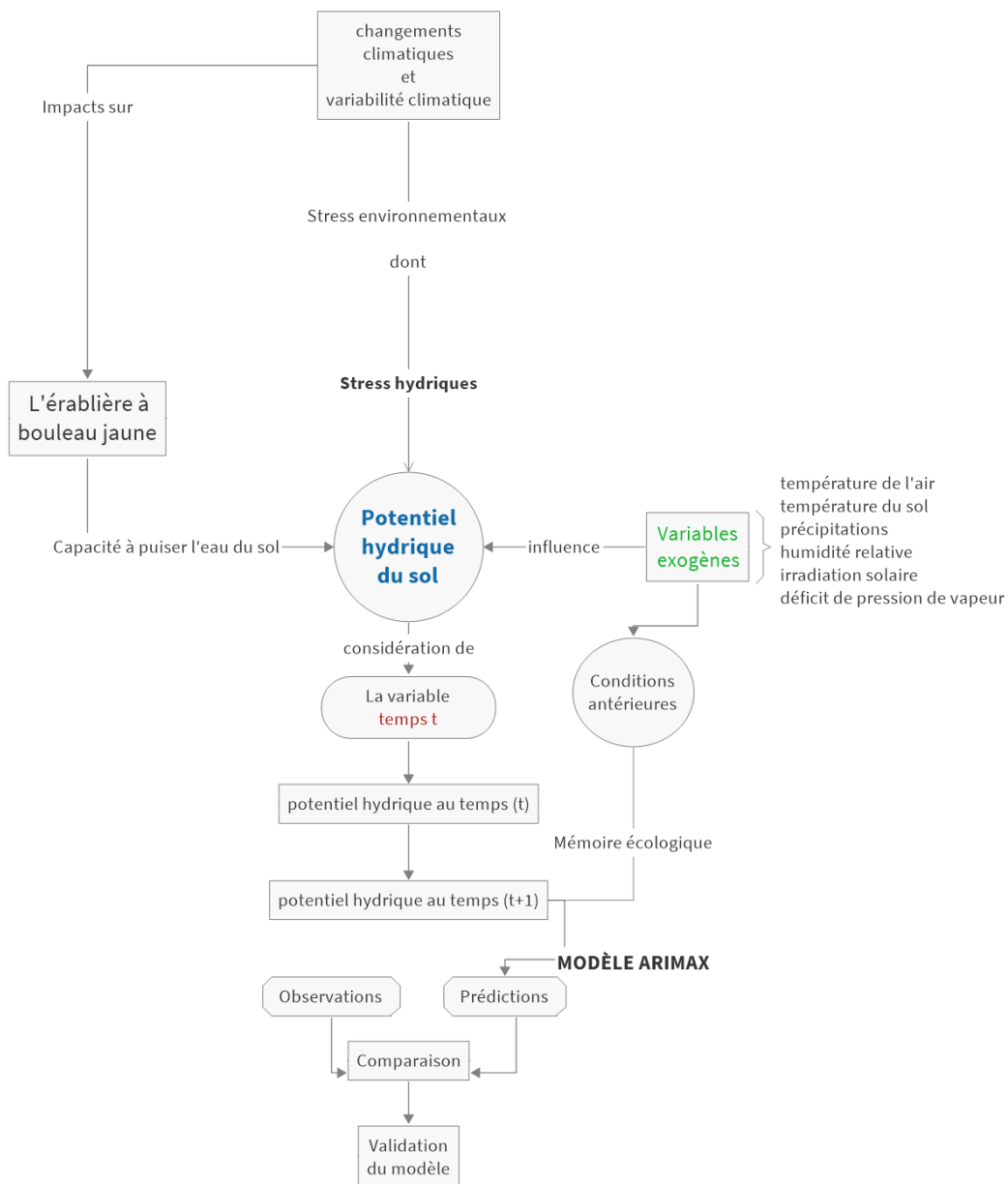


FIGURE 2.8 : Représentation des concepts centraux de notre étude au moyen d'un organigramme.

CHAPITRE III

MÉTHODOLOGIE

3.1 Partie expérimentale : SmartForests Canada

Nous l'avons exposé dans l'introduction, notre projet de recherche s'insère dans le contexte expérimental du projet pancanadien *SmartForests Canada*. Les données initiales qui ont été exploitées proviennent de la Station de biologie des Laurentides (SBL) qui se situe à St. Hippolyte, à 80 km au nord de Montréal (45°59'N ; 74°0'W). La figure 3.1 illustre sa position géographique et la distribution des peuplements d'arbres majoritaires sur le site d'étude et représentatifs de la forêt mixte.

Les arbres qui composent le site expérimental sont : l'érable à sucre (*Acer saccharum*), l'érable rouge (*Acer rubrum*), le hêtre américain ou à grandes feuilles (*Fagus grandifolia*), le bouleau jaune (*Betula alleghaniensis*), le bouleau blanc (*Betula papyrifera*), le peuplier à grandes dents (*Populus grandidentata*), le cèdre blanc (*Thuja occidentalis*), le pin blanc (*Pinus strobus*), l'érable de Pennsylvanie (*Acer pensylvanicum*) et l'épinette rouge (*Picea rubens*), mais aussi des espèces qui se retrouvent généralement dans la forêt boréale comme le sapin baumier (*Abies balsamea*) et l'épinette blanche (*Picea glauca*).

D'après les résultats obtenus avec le modèle BioSIM (Régnière et Saint-Amand, 2017), la température annuelle moyenne à la SBL entre 2003 et 2013 a été de 4.9°C et les précipitations moyennes sur la même période de 1270 mm, dont 30% de neige (Bélanger *et al.*, 2021). Les sols répertoriés à la SBL se sont développés à

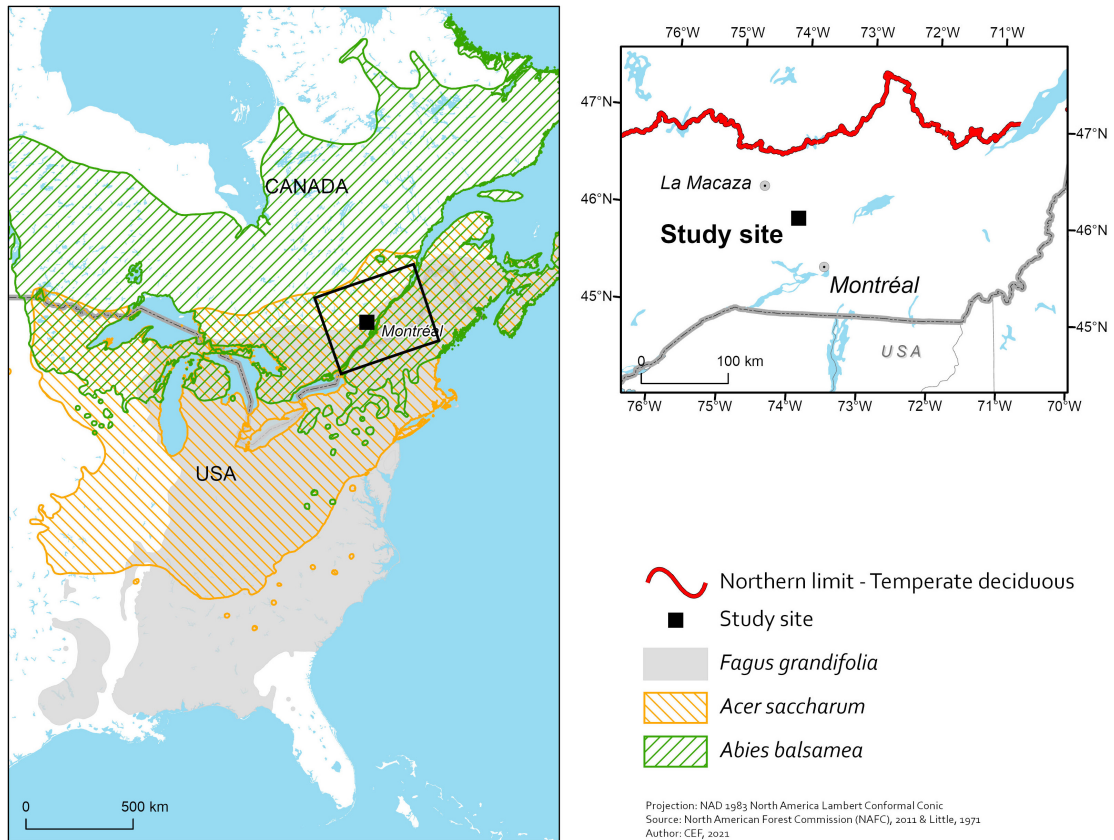


FIGURE 3.1 : Emplacement de la SBL (QC, Canada), d'après Bélanger *et al.* (2021). Sur la carte de gauche, les distributions de l'érable à sucre (*Acer saccharum*) et du hêtre à grandes feuilles (*Fagus grandifolia*) sont représentées en orange et en gris, respectivement. La limite nordique de la forêt tempérée est en rouge, sur la carte de droite.

partir de till glaciaire composé principalement d'anorthosite locale et de roches felsiques du Bouclier Canadien (Bélanger *et al.*, 2012). D'après la classification des sols (Soil Classification Working Group, 1998), ils sont de type podzol orthique humo-ferrique et ferro-humique à la texture sableuse et au pH acide (humus de type moder).

La SBL se situe à la limite nordique de l'érablière à bouleau jaune (*Betula alleghaniensis* Britton), ce qui positionne ainsi notre étude dans ce domaine bioclimatique spécifique.

Le terrain expérimental du projet SmartForests s'étale sur une zone de 18 ha qui comporte 32 *stations*¹ ou *parcelles* de 3 m sur 3 m. Ces 32 stations se répartissent sur trois secteurs géographiques, chacun comportant 8 blocs de 4 stations (Figure 3.2). Chaque bloc, d'une surface minimale de 0.5 ha, est composé d'arbres similaires (Bélanger *et al.*, 2021).

Ce design expérimental permet d'isoler trois peuplements principaux :

1. la sapinière à érable (*Mixedwood* ou MW)
2. l'érablière à bouleau (*Hardwood* ou HW)
3. l'érablière à hêtre (*Hardbeech* ou HB)

Les peuplements seront par la suite désignées par leur abbréviation (MW, HW, HB). La catégorisation d'une station s'est faite par rapport à différents critères. Ainsi, une station a été classée MW lorsque sa surface basale totale était constituée de plus de 20 % de sapin baumier (bloc 1, bloc 4 et bloc 7). Pour différencier les stations de peuplements HW et HB, la composition de la litière a été utilisée. Ainsi, quand 20 % de la masse totale de sa litière provenait du hêtre, la station était classée HB (bloc 3, bloc 5 et bloc 8), tandis que les stations HW en présentaient moins de 20 % (bloc 2 et bloc 6). Les méthodes qui ont permis cette classification sont détaillées dans Bélanger *et al.* (2021) dont nous avons repris les tableaux 3.1 et 3.2 qui résument les contributions des différentes espèces selon les deux critères choisis, la surface basale totale définie en $[\text{m}^2 \text{ha}^{-1}]$ et la composition de la litière en $[\text{m}^2 \text{m}^{-2}]$.

Chaque station est équipée de deux sondes mesurant la température du sol (Spectrum Technologies, United States) et de deux sondes mesurant le potentiel hydrique du sol (Irrometer 200SS-5, Watermark, United States). Elles ont été positionnées à 10 cm de profondeur et connectées au même enregistreur de données (WatchDog

1. Nous avons privilégié pour la suite le terme de *station*.

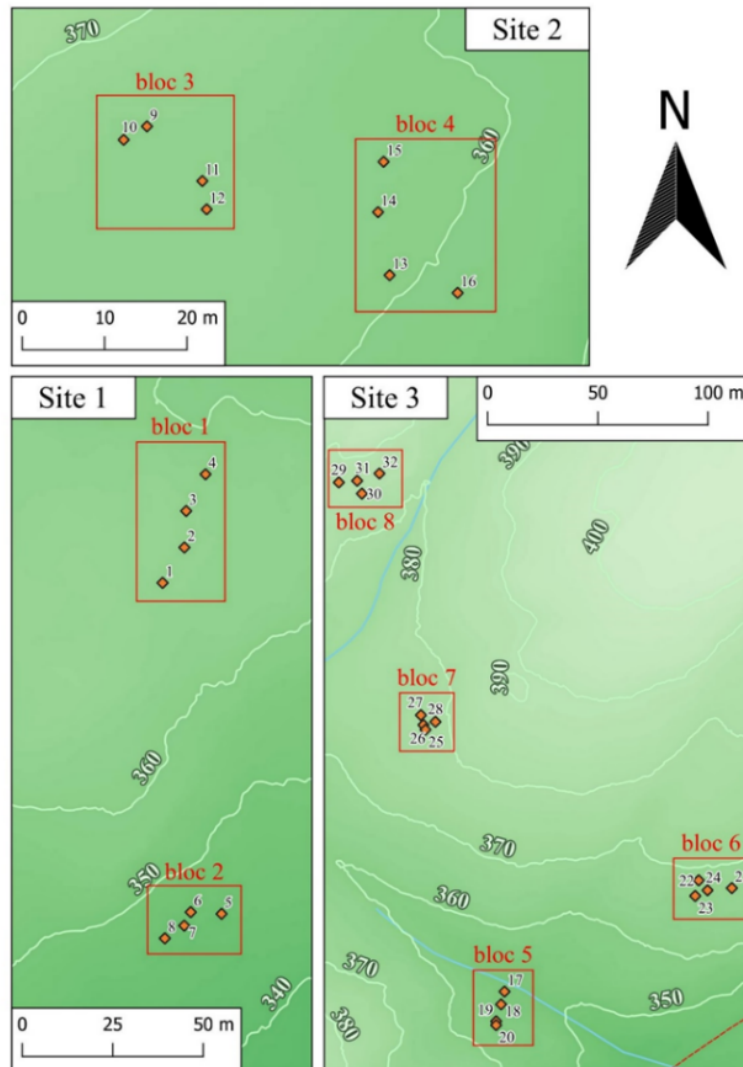


FIGURE 3.2 : Organisation spatiale des secteurs (site 1 à 3) et des blocs d'échantillonnage à la SBL (bloc 1 à 8). Les 32 stations sont représentées par des losanges. Crédit : Simon Lebel-Desrosiers.

1650 Micro Station, Spectrum Technologies). Cette profondeur a été choisie, car elle est centrale à la zone racinaire, mais aussi, car elle correspond à la profondeur des câbles chauffants qui ont été installés dans le sol de certaines stations (stations 2, 3, 5, 7, 10, 11, 13, 16, 18, 19, 21, 24, 26, 28, 29, et 32). Les sondes peuvent ainsi suivre l'effet du câble chauffant.

TABLEAU 3.1 : Surface basale totale (SB) en [m² ha⁻¹] et contributions de chaque espèce d'arbre à cette surface dans chaque bloc (données de 2018).

zone	bloc	peuplement	SB	Contribution [%]							
				AS	AB	BA	BP	FG	AR	PG	AP
1	1	MW	35.5	11.3	29.6	0.00	32.4	0.00	21.1	2.80	2.80
1	2	HW	35.5	35.2	9.90	0.00	32.4	0.00	19.7	2.80	0.00
2	3	HB	29.0	87.9	0.00	6.90	0.00	5.20	0.00	0.00	0.00
2	4	MW	35.5	45.1	29.6	2.80	2.80	2.80	11.3	0.00	5.60
3	5	HB	22.5	51.1	8.90	17.8	0.00	20.0	0.00	0.00	2.20
3	6	HW	31.0	54.8	6.50	11.3	14.5	12.9	0.00	0.00	0.00
3	7	MW	40.5	58.0	25.9	4.90	8.60	0.00	2.50	0.00	0.00
3	8	HB	34.0	76.5	0.00	5.90	2.90	14.7	0.00	0.00	0.00

N.B. MW correspond à la sapinière à érable, HW à l'érablière à bouleau et HB à l'érablière à hêtre, AS est *Acer saccharum*, AB est *Abies balsamea*, BA est *Betula alleghaniensis*, BP est *Betula papyrifera*, FG est *Fagus grandifolia*, AR est *Acer rubrum*, PG est *Populus grandidentata*, AP est *Acer pensylvanicum*. La somme totale des contributions de AS, AB, BP, BA, FG, AR, PG et AP est de 100%.

Le potentiel hydrique du sol a été mesuré par deux sondes distinctes, avec la même fréquence, toutes les 15 minutes, pour les mesures effectuées entre 2017 et 2020. La température du sol a été mesurée avec le même protocole. Les mesures ont été récoltées grâce à l'enregistreur de données, généralement entre mai et octobre.

À partir de 2020, une station météorologique complète de Campbell Scientific a été installée à la SBL. Ainsi, pour obtenir les données météorologiques antérieures, et en particulier à partir de 2017, l'outil logiciel BioSIM a été exploité (Régnière et Saint-Amand, 2017). D'après la description fournie par les concepteurs du modèle, BioSIM permet de fournir les données météorologiques en interpolant les données de stations météorologiques proches de la zone d'intérêt et en les ajustant pour les différences d'élévation, de latitude et de longitude à l'aide de gradients régionaux².

2. Les paramètres utilisés dans notre contexte sont décrits dans le rapport technique : Courcot, B. (2023) *Rapport technique du mémoire présenté comme exigence partielle de la maîtrise en technologie de l'information*. Université TÉLUQ

TABLEAU 3.2 : Contribution moyenne de chaque espèce d'arbre à la masse totale de la litière dans chaque bloc (données de 2018). L'indice foliaire (LAI) s'exprime en $[m^2 m^{-2}]$.

zone	bloc	peuplement	LAI	Contribution [%]							
				Feuilles	Aig.	AS	Betula	FG	AR	PG	AP
1	1	MW	3.20	54.1	45.9	1.14	21.3	0.09	31.3	0.00	0.25
1	2	HW	2.58	98.2	1.83	59.0	8.26	0.00	20.5	10.4	0.00
2	3	HB	2.66	99.9	0.08	59.5	10.6	21.4	5.22	1.22	2.00
2	4	MW	2.72	72.7	27.3	44.4	16.0	9.87	0.00	0.01	2.37
3	5	HB	3.01	98.4	1.64	31.4	35.0	29.0	2.36	0.30	0.32
3	6	HW	3.03	97.9	2.08	37.1	28.2	14.9	11.3	3.33	3.07
3	7	MW	2.93	90.8	9.23	36.4	3.02	1.39	48.2	1.39	0.39
3	8	HB	2.62	100	0.00	36.4	2.48	31.5	29.6	0.00	0.00

N.B. MW correspond à la sapinière à érable, HW à l'érablière à bouleau et HB à l'érablière à hêtre. LAI est l'indice foliaire, les feuilles se réfèrent à toutes les feuilles, les aiguilles (Aig.) sont principalement de Abies balsamea, AS est Acer saccharum, AB est Abies balsamea, Betula est Betula alleghaniensis et Betula papyrifera, FG est Fagus grandifolia, AR est Acer rubrum, PG est Populus grandidentata, AP est Acer pensylvanicum. La somme totale des contributions des feuilles et des aiguilles ou des aiguilles et de AS, Betula, FG, AR, PG et AP, par ligne, est de 100%.

3.2 Ressources et gestion des données

Les analyses ont été effectuées avec le logiciel Matlab R2022a. Ce choix a été motivé par un cours sur la fouille des données suivi dans le cadre du programme d'étude choisi. Dans la continuité de cette formation, le choix s'est porté sur ce logiciel que nous avons eu l'occasion d'utiliser dans un autre contexte de traitement d'images. Il aurait pu se porter sur le langage R, qui dispose de nombreux paquets de fonctions développés pour la science des données, et a l'avantage d'être ouvert et accessible gratuitement. Un des inconvénients est néanmoins le maintien de ces paquets qui ne sont pas toujours mis à jour, entraînant donc une obsolescence du code développé.

Quel que soit le langage ou le logiciel choisi, nous avons préféré mettre l'emphasis

sur la démarche méthodologique mise en place et la justification de celle-ci, dans notre contexte d'étude.

Les données brutes provenant des enregistreurs de données (*data loggers*) installés à la SBL ont été récupérées à quelques reprises entre mai et octobre, et elles ont été ensuite mises au format accessible au logiciel Matlab.

Pour les quatre années d'intérêt, de mai 2017 à octobre 2020, ce sont 1 696 473 observations brutes mesurées et collectées à la SBL, par les équipes des laboratoires iForêt³ et DOT-Lab⁴, qui ont été exploitées.

3.3 Partie modélisation : les différentes phases du projet

3.3.1 Les variables d'intérêt

Certaines variables météorologiques sont considérées comme des variables exogènes à notre problème, i.e. des variables dont la mesure est déterminée en dehors du modèle qui sera développé.

Ont été considérées comme variables exogènes :

- la température de l'air [$^{\circ}\text{C}$]
- les précipitations [mm]
- l'humidité relative [%]
- l'irradiance solaire [W m^{-2}]
- la température du sol [$^{\circ}\text{C}$]

Les valeurs journalières des quatre premières variables ont été générées avec le logiciel BioSIM. La température du sol a été mesurée toutes les 15 minutes, avec

3. Laboratoire iForêt

4. DOT-Lab

une sonde Spectrum Technologies (United States).

Il est possible de définir des paramètres intégrateurs qui permettent d'incorporer certaines variables météorologiques disponibles. Le déficit de pression de vapeur (*vapor pressure deficit* ou VPD) est un paramètre pertinent dans le cadre de notre étude au regard des variables météorologiques accessibles. Il se définit comme la différence entre la quantité d'eau que l'atmosphère pourrait contenir (100 % d'humidité) à une température donnée⁵ et la quantité réellement présente dans l'atmosphère. Il correspond schématiquement à la place potentiellement disponible dans l'air pour davantage de vapeur d'eau à une température donnée.

Nous avons calculé le déficit de pression de vapeur (VPD) à partir des données modélisées par le logiciel BioSIM.

En se reposant sur sa définition, ce paramètre se définit donc comme la différence entre la pression de vapeur saturante de l'air (SVP) et sa pression partielle de vapeur actuelle (VPA). Il peut s'exprimer à l'échelle d'une journée (*daily*) :

$$VPD_{daily} = SVP - VPA.$$

Nous définissons alors SVP à partir de la température moyenne journalière T_{mean} , qui correspond à T_{air} moyennée sur une journée, en utilisant l'équation de Tetens (1930), reformulée par Monteith et Unsworth (2007) pour des températures supérieures à 0 °C :

$$SVP = 0.61078 \times \exp\left(\frac{17.27 * T_{\text{air}}}{(T_{\text{air}} + 237.3)}\right).$$

Pour des températures inférieures à 0 °C, les coefficients varient et la SVP se définit

5. C'est la capacité hygrométrique.

alors ainsi (Murray, 1967) :

$$SVP = 0.61078 \times \exp^{\frac{21,875 * T_{air}}{(T_{air} + 265,5)}}.$$

La pression de vapeur actuelle se définit quant à elle à partir de la SVP et de l'humidité relative (RelH), telle que :

$$VPA = SVP \times \frac{RelH}{100}.$$

Précisons les unités des variables dans les formules définies ci-dessus : VPA, SVP et VPD en [kg Pa]; T_{air} en [°C] et RelH en [%].

3.3.2 Sept phases d'analyse

Avant de détailler davantage la méthodologie, définissons plus en détail la notion de *série temporelle*, qui est centrale dans notre approche. Au chapitre 2, nous avons évoqué la nécessité de considérer le temps comme une variable explicative. En effet, nos données représentent des mesures réalisées au moyen de sondes qui sont programmées pour enregistrer une mesure toutes les 15 minutes. Ainsi, chaque entrée de notre base de données initiale est caractérisée par un horodatage (*timestamp*) qui marque la dimension temporelle de la mesure, précisant alors quand celle-ci a été ajoutée dans la base de données. La date et l'heure de la mesure, souvent combinées ensemble dans un format unique numérique, deviennent donc un champ spécifique qui va être exploité dans l'analyse.

Chaque ligne dans la base de donnée possède son horodatage et les autres variables peuvent par conséquent être représentées en fonction de cette dimension temporelle. La succession des mesures devient une *série temporelle* ou *chronologique* de mesures discrètes. À partir de cette succession de points, il est possible d'effectuer un échantillonnage des données, en les moyennant par exemple sur une heure

(figure 3.3) ou sur une journée (figure 3.4). Les deux figures illustrent l'évolution de la température du sol en fonction du temps, au cours d'une journée et de 15 jours.

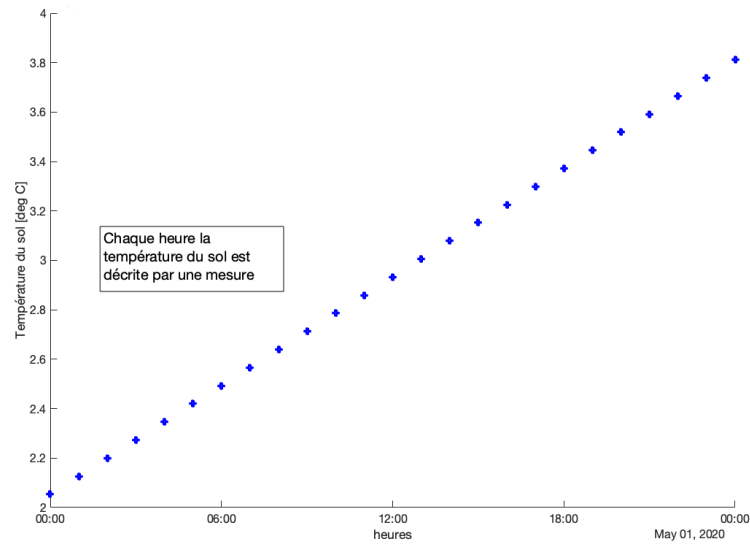


FIGURE 3.3 : Premier exemple d'une série temporelle : température du sol mesurée entre le 01 mai 2020 minuit et le 02 mai 2020 minuit inclus (moyenne par heure). La série est constituée de 25 points temporels.

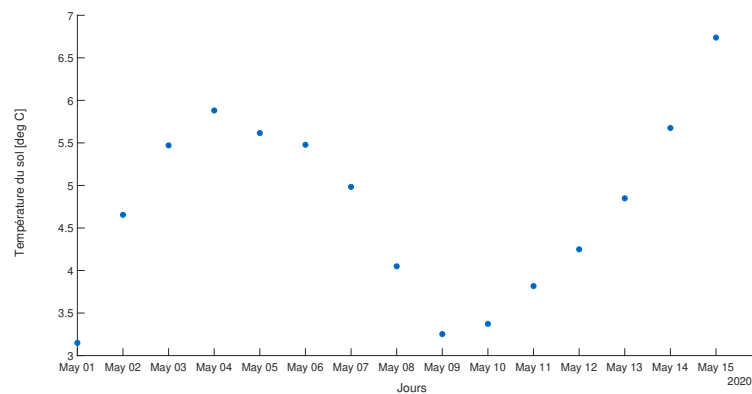


FIGURE 3.4 : Deuxième exemple d'une série temporelle : température du sol mesurée entre le 01 mai 2020 et le 15 mai 2020 (moyenne journalière). La série est constituée de 15 points temporels.

L'avantage de cette *temporalisation* des données est qu'il est possible d'analyser les phénomènes en changeant d'échelle temporelle. Une analyse plus fine permet d'isoler la variation de variables d'intérêt au cours d'une même journée, tandis que des variations étudiées à l'échelle des saisons peuvent amener une compréhension différente des processus dynamiques.

La manipulation de la composante temporelle des données à travers les horodatages est à considérer avec précaution, car des contraintes d'heures d'été ou d'hiver, et des formats spécifiques des enregistreurs de données sont des sources d'erreur qui interviennent au début de l'analyse, plus spécifiquement pendant la préparation des données. Il est alors important de visualiser les données en variant les échelles temporelles pour s'assurer de leur cohérence.

Cette validation constante des données a été faite à chaque étape clé de la méthodologie. Les graphiques générés sont accessibles dans le rapport technique qui a été rédigé en complément de ce mémoire. Chaque phase de l'analyse y a été détaillée, et les résultats obtenus y sont également présentés.

La méthodologie développée repose sur sept phases.

1. Préparation des données ;
2. Sélection des variables prédictives ;
3. Analyse des changements d'état du potentiel hydrique du sol ;
4. Exploration des séries temporelles avec la méthode de Box-Jenkins (Box *et al.*, 2008) ;
5. Test et validation de modèles prédictifs de type ARIMAX (*autoregressive integrated moving average model with exogenous variables*) ;
6. Exploration de modèles prédictifs de type NARX (*non linear autoregressive model with exogenous variables*) ;
7. Application à l'étude de périodes de sécheresse-flash.

La dernière phase est une application de la démarche dans le cadre de l'étude des périodes de sécheresse-flash qui ont été mises en évidence entre 2017 et 2020. Elle est présentée au chapitre 4 et permet de répondre à la question de recherche QR4.

3.3.2.1 Préparation des données

La phase de préparation des données est détaillée dans le rapport technique associé au projet (Courcot, 2023).

- Suppression des données aberrantes concernant les mesures issues des deux sondes de température du sol désignées respectivement par les lettres A et B. Les périodes d'intérêt, pour chaque année considérée, s'étalent de mai à novembre. Ainsi les critères de suppression sont : températures inférieures à -20°C et supérieures à 46°C ;

- Ajustement des mesures du potentiel hydrique du sol par rapport à la température du sol, selon les recommandations de Irmak *et al.* (2016) ;

- Échantillonnage des données de la SBL pour permettre de réduire le rapport signal/bruit et la variabilité des mesures entre les deux sondes utilisées par mesurandes : température du sol (sondes désignées par A et B) et potentiel hydrique du sol (sondes désignées par C et D) ;

Un échantillonnage sur 1 heure et un autre sur 1 journée ont été effectués par défaut, mais toutes les fréquences d'échantillonnage sont accessibles. Ces deux échantillonnages ont été réalisés pour chaque station individuelle, mais aussi pour chaque bloc et enfin, pour chaque peuplement. Dans chaque cas, un potentiel hydrique moyen a été obtenu, ainsi que l'incertitude calculée pour chaque mesure. Il en a été de même pour la température ;

- Régularisation de chaque mesure échantillonnée grâce à une interpolation

de type BSpline cubique pour obtenir une mesure à chaque point temporel. L'objectif ici est d'obtenir des séries temporelles régulières sans données manquantes ;

- Lissage des données avec un filtre de type Savitzky-Golay (polynôme de degré 6) associé à une moyenne mobile sur 3 jours (centrée) (Savitzky et Golay, 1964). Ce lissage effectué après le moyennage des données est une stratégie à privilégier car il permet de réduire le bruit d'un facteur égal à la racine carrée du nombre de mesures (Savitzky et Golay, 1964). De plus, le filtre passe-bas sélectionné permet de ne pas sacrifier les valeurs extrêmes dans les données ;
- Échantillonnage sur une heure et sur une journée des données météorologiques issues de la station météo de la SBL qui sont mesurées toutes les 30 minutes, et régularisation des séries temporelles pour permettre une synchronisation des données avec les données de potentiel hydrique et de température du sol ;
- Sélection pour chaque année étudiée d'un intervalle temporel, qui est fixé par les données de la station qui présente le plus petit intervalle d'observation.

Les intervalles sélectionnés sont : du 11 mai au 29 octobre pour l'année 2017, du 24 mai au 7 novembre pour 2018, du 8 mai au 10 novembre pour 2019 et du 1^{er} mai au 23 octobre pour 2020.

Ces intervalles de temps harmonisés ont permis de concaténer, par rapport à chaque point temporel, les données ainsi préparées de la SBL et les données provenant de la station météo qui n'étaient pas synchronisées entre elles. Le fait d'avoir des séries temporelles régulières commençant et finissant au même moment a permis cette concaténation.

Au cours de la dernière étape de cette première phase méthodologique, nous avons également généré avec le modèle BioSIM les données météorologiques pour les

années 2017, 2018 et 2019 qui n'étaient pas accessibles à travers la station météo de la SBL. Les données de l'année 2020 ont également été générées pour permettre une comparaison avec celles de la station météo. Nous avons implémenté des fonctions qui permettent d'utiliser les données extraites de BioSIM dans notre pipeline d'analyse développé sous Matlab.

Les données météo générées par BioSIM ont une fréquence d'échantillonnage d'une heure, mais aussi journalière. Les deux fréquences ont été extraites, et les données ont été également régularisées et synchronisées sur les mêmes intervalles de temps d'intérêt, permettant ainsi de les concaténer avec les données du potentiel hydrique du sol ainsi que celles de la température du sol.

Une validation des données météo générées par le logiciel BioSIM a été réalisée grâce à une comparaison avec les données recueillies par la station météo de la SBL en 2020.

Enfin, les périodes de sécheresse-flash, périodes sans précipitation de 8 jours ou plus, pendant lesquelles le potentiel hydrique augmente, ont également été mises en évidence et les données correspondantes ont été extraites pour permettre des analyses plus précises de ces périodes spécifiques. Ce choix de 8 jours s'est fait par rapport à l'étude réalisée par Fu *et al.* (2022) dans laquelle les auteurs considèrent une période de 10 jours consécutifs sans précipitation durant laquelle l'humidité du sol diminue. Dans notre contexte, une période plus courte a permis d'intégrer davantage de périodes d'assèchement.

Cette première phase a permis de préparer les données initiales pour les analyses des séries temporelles. En effet, l'objectif était d'obtenir des données provenant de différentes sources de mesure qui soient synchronisées dans le temps.

La base de données journalières pour chaque peuplement (MW, HW, HB) qui a été exploitée par la suite (tableau 3.3) comporte ainsi les champs résumés au tableau 3.4. Nous pouvons souligner que la précision des mesures collectées est

différente de celles générées par le modèle BioSIM. De plus, par défaut, l’affichage sous Matlab se fait avec quatre chiffres après la virgule, ce qui explique le format des données obtenues dans la base de données nettoyée.

TABLEAU 3.3 : Extrait de la base de données journalières. Par défaut, Matlab affiche les données avec un format de 4 chiffres après la virgule (format *short*).

dateCollectedSampled	T_{sol}	Ψ	T_{air}	Prcp	RelH	SRad	VPD
01-May-2020	3.1494	12.3950	9.7000	2.5000	81.7000	190.0000	0.1401
02-May-2020	4.6544	12.1365	7.6000	1.1000	58.3000	260.5000	0.3550
03-May-2020	5.4710	12.2721	11.4000	0.8000	69.0000	219.9000	0.2794
04-May-2020	5.8811	12.5059	6.5000	0.0000	61.2000	184.4000	0.3598
05-May-2020	5.6159	12.2829	4.8000	0.0000	52.5000	304.0000	0.4324
06-May-2020	5.4779	12.0331	5.8000	0.0000	45.7000	338.5000	0.4896
07-May-2020	4.9828	12.4315	5.6000	0.5000	45.8000	258.8000	0.4722
08-May-2020	4.0494	13.8908	1.2000	0.1000	53.5000	206.1000	0.3794
09-May-2020	3.2530	15.2500	-0.9000	0.0000	45.7000	218.3000	0.4189
10-May-2020	3.3731	16.1455	2.4000	0.0000	48.3000	251.7000	0.4022

TABLEAU 3.4 : Description des champs de la base de données journalières.

Description	Nom	Nom dans le code	Unité ou format
horodatage	$date$	dateCollectedSampled	DD-MM-YYYY
date en jour julien	$date_{julien}$	dateJulianYear	entier
température du sol moyennée sur une journée	T_{sol}	mean_temp_sampled	[°C]
potentiel hydrique du sol moyenné sur une journée	Ψ	mean_potential_sampled	[kPa]
état du potentiel	$etat_{\Psi}$	StateSM	{low, medium, high, very high}
température de l'air moyenne journalière	T_{air}	T _{air}	[°C]
température de l'air maximale journalière	T_{max}	T _{max}	[°C]
température de l'air minimale journalière	T_{min}	T _{min}	[°C]
température du point de rosée	T_{dew}	T _{dew}	[°C]
humidité relative	$RelH$	RelH	[%]
précipitations totales	$Prcp$	Prcp	[mm]
irradiance solaire	$SRad$	SRad	[W m ⁻²]
déficit de pression de vapeur calculé	VPD	VPD	[kPa]

3.3.2.2 Sélection des variables prédictives

L'objectif de cette phase d'analyse était de répondre à la question de recherche QR1.

Une variable catégorielle a été introduite pour décrire l'état du potentiel hydrique du sol : StateSM (*state of the soil moisture*). Quatre niveaux de tension ont été définis dans un premier temps : faible (*low*), moyen (*medium*), élevé (*high*) et très élevé (*very high*). Des valeurs de potentiel hydrique (en kPa) ont été fixées pour délimiter chaque niveau. Pour chaque peuplement et pour les quatre années d'intérêt, une étude préalable des courbes de densité du potentiel hydrique a été réalisée et la répartition des valeurs nous a permis de dégager ces seuils. Des seuils d'irrigation mis en évidence en agriculture nous ont également aiguillé (Shock et Wang, 2011). En effet, selon les études citées par Shock et Wang (2011) qui utilisent le même type de sondes, positionnées entre 15 cm et 30 cm de profondeur, pour les pommes de terre il est conseillé d'irriguer lorsque le potentiel hydrique du sol se situe entre 20 kPa et 60 kPa. Pour les choux-fleurs cultivés en sol sableux, le seuil d'irrigation est entre 20 kPa et 40 kPa, tandis que pour le céleri, il se situe autour de 10 kPa. Nous l'avons souligné la nature du sol a une influence sur ces seuils.

Pour sélectionner les variables qui permettent d'améliorer la prédiction de l'état du potentiel hydrique du sol, nous avons exploité les arbres de décision. Ce sont des méthodes de classification qui reposent sur un apprentissage supervisé. L'algorithme CART (*classification and regression tree*) a été utilisé (Breiman *et al.*, 1984).

Pour affiner les résultats obtenus et surmonter l'aspect statique de ce type d'approche, une analyse de la causalité au sens de Granger entre le potentiel hydrique du sol et les variables exogènes du modèle a été menée (Granger, 1969).

3.3.2.3 Analyse des changements d'état du potentiel hydrique du sol

Le potentiel hydrique du sol étant décrit par sa valeur à chaque point temporel, mais aussi par son état (faible, moyen, élevé, très élevé), nous avons analysé comment cet état évolue au cours du temps, en déterminant les probabilités de transition d'un état à un autre et le nombre de jours espérés pour les atteindre. Pour ce faire, les séries temporelles ont été exprimées au moyen de chaînes de Markov à temps discret. La notion de chaîne permet de modéliser des phénomènes dynamiques aléatoires dans lesquels le passé n'intervient que dans le dernier instant de la chaîne (L'Écuyer, 2020). En effet, un processus de Markov est un processus stochastique (aléatoire) qui possède la propriété de Markov. Cette propriété se définit ainsi : l'information utile pour la prédiction du futur est entièrement contenue dans l'état présent du processus et n'est pas dépendante des états précédents.

De façon simple, nous pouvons illustrer une chaîne de Markov à temps discret en considérant une pièce de monnaie non-truquée qui possède deux états : Pile et Face (figure 3.5).



FIGURE 3.5 : Illustration simple d'une chaîne de Markov à temps discret avec la matrice des probabilités de transition pour passer d'un état à un autre.

Il est alors possible de passer de l'état Pile à l'état Face avec une probabilité égale à 0.5. Mais il est également possible de passer de l'état Pile à l'état Pile avec la même probabilité. Ainsi, nous pouvons définir toutes les transitions possibles et

leur probabilité associée. Ces probabilités constituent la matrice des probabilités de transition.

Cette démarche a été appliquée au potentiel hydrique du sol afin de caractériser les probabilités des différents changements d'état qui sont observés dans les données. L'objectif était de répondre à la question de recherche QR2.

3.3.2.4 Méthode de Box-Jenkins

La méthode de Box-Jenkins a été appliquée pour obtenir les modèles prédictifs ARIMA du potentiel hydrique du sol afin de répondre à la question de recherche QR3. Cette démarche d'analyse des séries temporelles repose sur cinq étapes.

Avant de les détailler, il faut préciser que l'objectif des modèles ARIMA est de déterminer chaque valeur d'une série temporelle en fonction des valeurs qui la précède.

Ainsi, puisque nous avons considéré le potentiel hydrique du sol Ψ en fonction du temps (t), il s'agit de déterminer $\Psi(t)$ en fonction des valeurs de Ψ au temps ($t-1$), ($t-2$), etc.

L'hypothèse centrale repose donc sur le fait que les points d'une série temporelle donnée sont fonction des points précédents, mais aussi de leurs erreurs.

Chaque point de la série temporelle peut alors être représenté par :

- (a) un processus dit autorégressif qui correspond mathématiquement à la somme pondérée d'un ensemble p de points précédents, plus un terme aléatoire d'erreur ;
- (b) un processus de moyenne mobile qui suppose que chaque point est fonction des erreurs entachant les q points précédents, plus sa propre erreur ;
- (c) un processus de différenciation d'ordre d qui suppose dans le cas où d est égal à 1 que chaque point présente une différence constante avec le point précédent. Dans le cas où d est égal à 2, c'est une différence de différence qui est considérée.

Si nous traduisons en équation ces trois représentations, cela donne pour le poten-

tiel hydrique du sol $\Psi(t)$:

(a)

$$\Psi(t) = \Phi_1 \Psi_{t-1} + \Phi_2 \Psi_{t-2} + \dots + \Phi_p \Psi_{t-p} + \epsilon(t)$$

(b)

$$\Psi(t) = -(\Theta_1 \epsilon_{t-1} + \Theta_2 \epsilon_{t-2} - \dots + \Theta_q \epsilon_{t-q}) + \epsilon(t).$$

Par convention, le signe négatif a été choisi par Box et Jenkins. Cependant, certains auteurs et certains logiciels, et en particulier Matlab, ont opté pour une définition des paramètres relatifs à la moyenne mobile qui est positive (Nau, 2020).

(c) différenciation d'ordre 1 :

$$\Delta_t \Psi = \Psi_t - \Psi_{t-1}$$

différenciation d'ordre 2 :

$$\Delta_t \Psi = (\Psi_t - \Psi_{t-1}) - (\Psi_{t-1} - \Psi_{t-2}).$$

Un modèle ARIMA combine alors les trois équations, et la détermination du modèle consiste à trouver les paramètres Φ_i et Θ_j qui permettent d'aboutir à une représentation de l'évolution temporelle du potentiel hydrique du sol. La méthode de Box-Jenkins aide à la mise en place d'une méthodologie ayant pour finalité l'obtention d'un tel modèle. C'est la démarche que nous avons choisie.

La première étape de la méthode de Box-Jenkins consiste à vérifier la stationnarité des séries temporelles du potentiel hydrique du sol. Pour cela, le test de Philips-Perron est appliqué à chaque série annuelle. Si la stationnarité n'est pas vérifiée, alors il faut différencier la série et appliquer de nouveau le test.

Le type de processus suivi par la série temporelle est ensuite déterminé par le processus de type autorégressif (AR) et le processus de type moyenne mobile (MA).

La troisième étape consiste à déterminer le degré p du processus AR et le degré q du processus MA. Si la série temporelle doit être différenciée pour assurer sa stationnarité, alors il faut ajouter le degré d de l'intégration (I). En effet, il est avantageux d'incorporer l'intégration de la série dans le processus car le modèle obtenu, dans le cas où une dérivation a été faite, ne devra pas être explicitement intégré pour obtenir le résultat final. Précisons que (p,d,q) sont des entiers positifs. Enfin, si k variables exogènes sont ajoutées pour améliorer la prédiction de la série temporelle étudiée, il faut les définir explicitement dans le modèle. Ces variables s'expriment sous la forme d'une combinaison linéaire de type $\sum_k \alpha_k X_k$, qui est ajoutée aux équations du modèle ARIMA.

Le tableau 3.5 résume les paramètres utilisés dans le modèle ARIMAX.

TABLEAU 3.5 : Description des paramètres du modèle ARIMAX.

Description	Nom	Nom dans le code
ordre du processus autorégressif (AR)	p	p
ordre du processus de moyenne mobile (MA)	q	q
degré de différenciation (I)	d	d
variables exogènes	X_k	X

3.3.2.5 Test et validation du modèle prédictif de type ARIMAX

Le modèle de type ARIMAX(p,d,q) est alors évalué sur des données appelées données d'entraînement. La qualité de l'ajustement est déterminée au moyen du critère d'information de Akaike (AIC) et du critère bayésien d'information (BIC) (Akaike, 1974; Schwarz, 1978).

La validation du modèle ARIMAX(p,d,q) est ensuite réalisée en utilisant des données différentes des données d'entraînement, appelées données de test. L'objectif de l'étape de validation est de pouvoir prédire des données passées, réelles, mais qui n'ont pas servi à la phase d'apprentissage. Cela constitue la cinquième phase de la méthodologie. L'étape de validation est primordiale dans la méthodologie.

Il y a bien sûr la validation du code implémenté qui se fait à l'aide des outils de debugage accessibles dans Matlab. Il y a ensuite la validation des modèles obtenus. Il faut, nous l'avons souligné, que chaque modèle soit entraîné sans être contaminé par les données qui serviront à la validation, c'est un pré-requis fondamental et il faut s'en assurer tout au long du développement en préparant soigneusement les différents échantillons de données.

Lorsque l'entraînement du modèle est terminé, celui-ci peut être utilisé pour prédire des données connues qui n'ont pas servi à la phase d'apprentissage. En définissant pour chaque point temporel la différence entre la valeur observée et la valeur ajustée, appelée *résidu*, cela permet d'évaluer la cohérence des résultats. En visualisant ces résidus grâce à une représentation graphique, il est alors possible d'avoir une première validation du modèle en vérifiant la faible valeur de chaque résidu. De plus, il faut vérifier que les résidus ne sont pas corrélés entre-eux au cours du temps. En traçant leur fonction d'auto-corrélation, cela permet de le valider. Le test statistique de Greta Ljung et George Box peut également être réalisé à cette

fin⁶.

Une représentation matricielle appelée *matrice de confusion* est également un moyen efficace de valider le modèle obtenu. La figure 3.6 schématise cette matrice et explique, dans notre contexte, comment l'interpréter en particulier lors de la prédiction de l'état du potentiel hydrique du sol {Small, Medium, High}. La représentation qui a été adoptée sur le schéma 3.6 correspond à la représentation utilisée par défaut dans Matlab.

Nous avons également calculé différentes métriques qui caractérisent la performance du modèle testé. Ces indicateurs se définissent à partir du nombre d'états correctement prédits, i.e. lorsque l'état prédit correspond à l'état observé (vrai positif et vrai négatif), et à partir des états qui ne sont pas correctement prédits (faux positif et faux négatif).

L'objectif est de viser une exactitude⁷ élevée (en pourcentage) et de diminuer le nombre d'états incorrectement classés par le modèle. Il est également important de réfléchir aux implications des mauvaises prédictions. Cette réflexion est propre à chaque contexte. Par exemple dans notre cas, si un état est prédit comme *High* alors qu'il a été observé comme *Small* (faux négatif) quelle serait la conséquence ? On pourrait irriguer à tort la station considérée. Inversement, si un état est prédit comme *Small* alors qu'il a été observé comme *High* (faux positif) alors le stress hydrique que vont subir les arbres ne sera pas considéré. Faut-il optimiser un modèle qui va privilégier le premier ou le second cas décrit ? La réalité expérimentale nous guiderait davantage vers un modèle qui aurait tendance à avoir moins de faux positifs pour ne pas passer à côté de stress hydriques éventuels. Nous visons un modèle qui aura une meilleure précision par rapport à sa sensibilité ; en d'autres

6. Se référer au rapport technique pour plus de détails.

7. Nous utiliserons par la suite le terme d'*exactitude*, au sens du mot anglais *accuracy*, pour désigner la proportion de prédictions exactes ou *correctes* du modèle. Pour comparer deux modèles, nous parlerons alors de *précision*.

termes, la qualité de la prédiction est préférée à sa quantité.

L'état *Medium* du potentiel hydrique agit comme une zone de transition entre les valeurs basses et élevées de celui-ci. Nous nous attendons à une plus faible exactitude des prédictions pour cet état, car les valeurs fixées pour définir les seuils de transition d'un état à un autre nécessiteraient une optimisation spécifique.

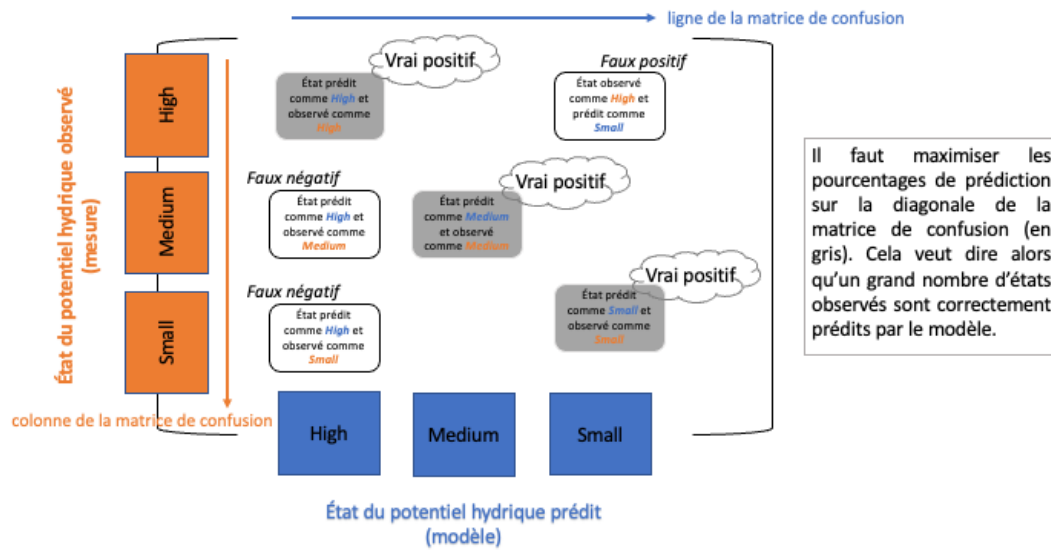


FIGURE 3.6 : Interprétation de la matrice de confusion : exemple pour la prédiction de l'état du potentiel hydrique du sol.

Exactitude de la prédiction (*accuracy*)

$$Accuracy = (VP + VN) / (VP + FP + VN + FN) = \text{Nb états correctement prédits} / \text{Nb observations}$$

Somme sur la diagonale de la matrice de confusion

Précision de la prédiction (*precision*)

$$Precision = VP / (VP + FP)$$

Terme sur la diagonale Somme des termes sur les lignes de la matrice

Sensibilité de la prédiction (*recall*)

$$Recall = VP / (VP + FN)$$

Somme des termes sur les colonnes de la matrice

Spécificité de la prédiction (*specificity*)

$$Specificity = VN / (VN + FP)$$

Score F1 (*F1-score*)

$$F1\text{-score} = 2 * precision * recall / (precision + recall)$$

Objectif :

- augmenter le nombre de vrais positifs (VP) et de vrais négatifs (VN)
- diminuer le nombre de faux positifs (FP) et de faux négatifs (FN)

Signe de sur-ajustement du modèle
➤ précision élevée, faible sensibilité

Il faut viser un équilibre précision/sensibilité

FIGURE 3.7 : Calcul des différentes métriques caractérisant la performance du modèle.

La dernière étape de la méthode de Box-Jenkins peut consister en la prédiction du potentiel hydrique du sol sur des horizons futurs. Lorsque des variables exogènes sont ajoutées au modèle, il faut alors s'appuyer sur des données futures, soit générées par d'autres modèles prédictifs (analogues synthétiques), soit issues d'analogues climatiques représentant le climat futur de la zone étudiée. Dans notre cas, cet analogue aurait pu être similaire au climat d'une région plus au sud de la SBL, comme la région du Maine aux États-Unis. Cependant, nous n'avons pas développé cette étape.

3.3.2.6 Exploration de modèles prédictifs de type NARX

La sixième étape de notre méthodologie consiste à tester une autre approche de modèle reposant sur les réseaux de neurones à propagation avant (*feedforward time delay neural network*). Cette approche de type NARX (*Non-linear autoregressive*

with exogeneous variables) exploite un réseau neuronal dit récurrent. Ce type de réseaux permet de capturer la dynamique de systèmes complexes. Il a été choisi car il est possible d'intégrer des variables exogènes dans le modèle prédictif de séries temporelles, et il est adapté à la modélisation de système dynamique non linéaire.

L'aspect récurrent du réseau neuronal est similaire au concept d'autorégression décrit dans le modèle ARIMA. En effet, les valeurs passées du potentiel hydrique du sol sont exploitées pour prédire la valeur au temps t .

La figure 3.8 illustre le réseau utilisé et précise les paramètres qui ont été choisis.

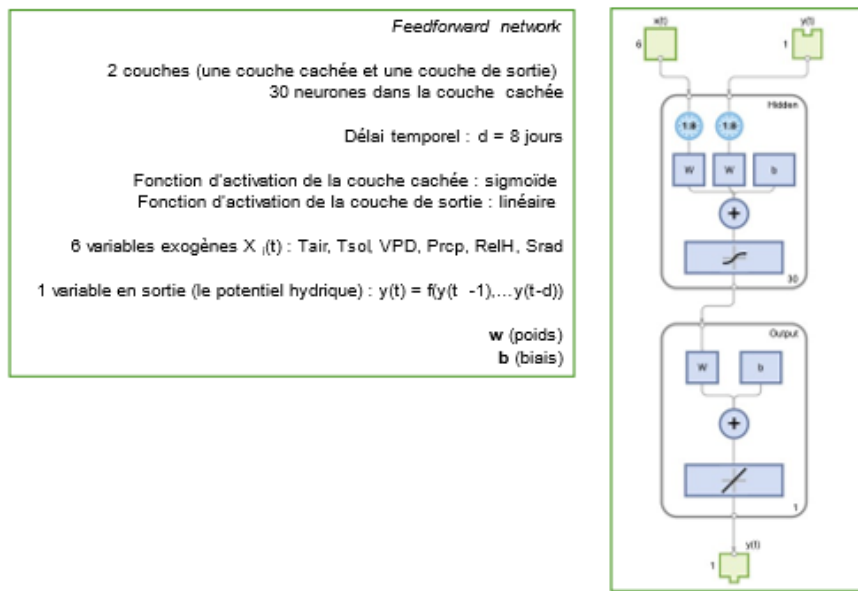


FIGURE 3.8 : Représentation du réseau utilisé pour l'approche NARX.

Ainsi, les paramètres optimaux trouvés sont les suivants :

- 2 couches : une couche cachée et une couche de sortie ;
- 30 neurones dans la couche cachée ;
- Décalage de 8 jours introduit ($d = 8$ jours) ;

- Algorithme de Levenberg-Marquardt pour l'optimisation des poids w et des biais b ;
- 80 % des données d'entrée sont utilisées pour l'entraînement (*training*), 10 % pour la validation et 10 % pour le test.

3.3.2.7 Application à l'étude de périodes de sécheresse-flash

Pour la dernière phase méthodologique, les modèles ARIMAX, développés pour chacun des peuplements étudiés, ont été utilisés dans le cas spécifique des périodes de sécheresse-flash à déficit de précipitations. Les modèles ont été entraînés sur des données correspondant à ces périodes, en incluant 16 jours avant et après chaque épisode de sécheresse. L'objectif de cette phase était d'explorer la capacité des modèles à prédire le potentiel hydrique du sol dans un contexte de stress hydrique (question de recherche QR5). La figure 3.9 résume la stratégie qui a été développée pour modéliser le potentiel hydrique du sol durant ces périodes, en considérant quatre variables exogènes. Les sept périodes de sécheresse-flash qui ont été mises en évidence entre 2017 et 2020 ont été ciblées. Pour chaque peuplement étudié, un modèle ARIMAX a été entraîné sur les données relatives aux sécheresses entre 2017 et 2019. Chaque modèle a été ensuite validé en prédisant le potentiel hydrique du sol pour l'année 2020.

La figure 3.10 résume les six premières phases méthodologiques.

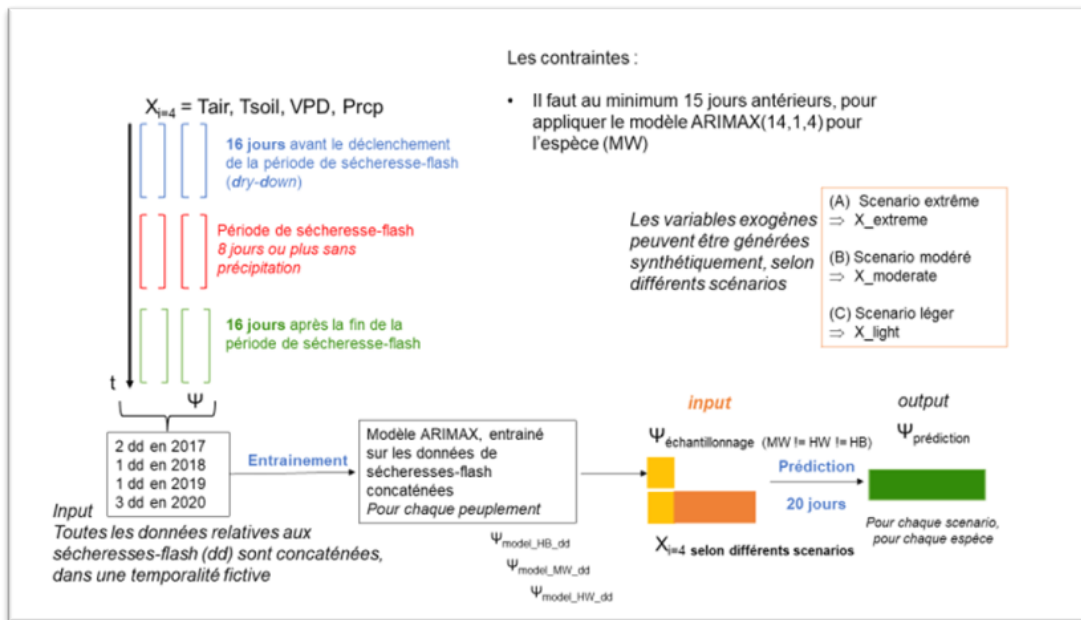


FIGURE 3.9 : Stratégie mise en place dans l'étude des périodes de sécheresse-flash.

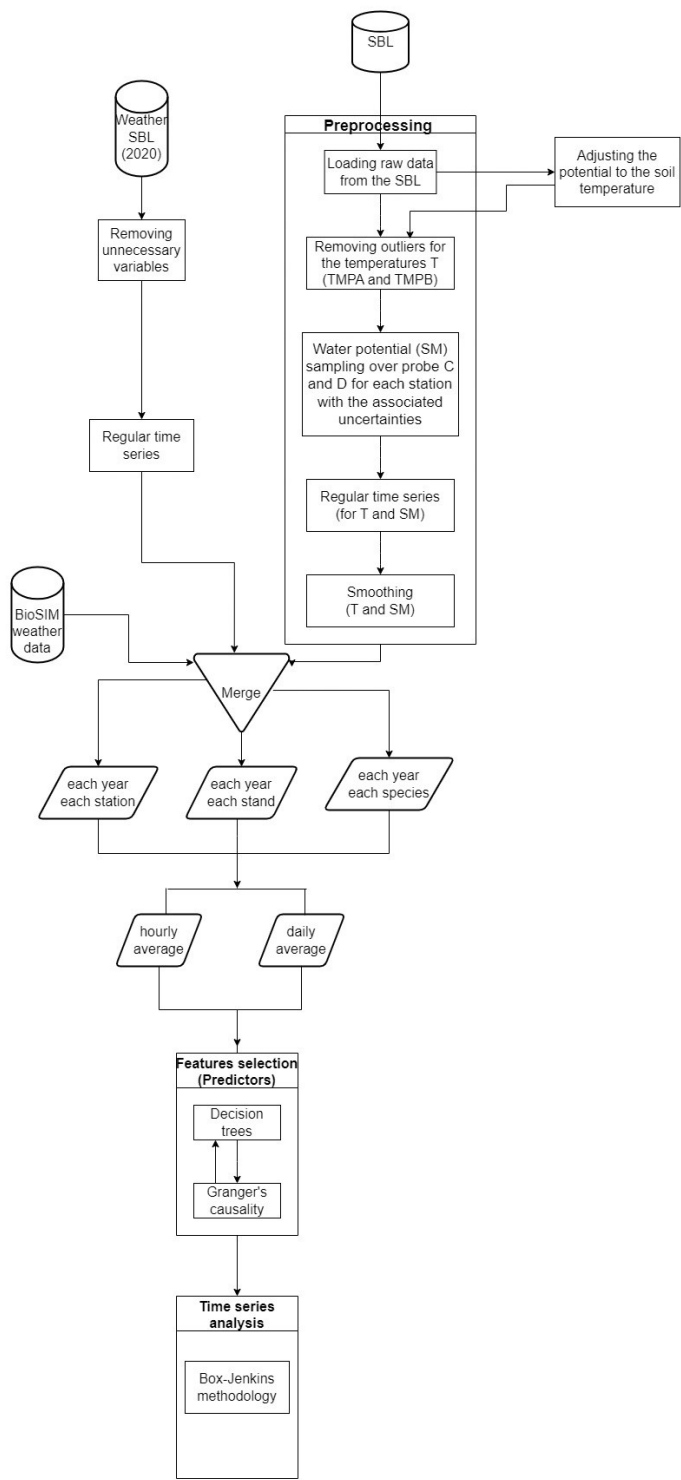


FIGURE 3.10 : Résumé de la méthodologie mise en place, proposé en anglais (en vue d'une insertion dans une publication éventuelle).

CHAPITRE IV

RÉSULTATS ET ANALYSE

Nous allons tout d’abord reprendre les questions de recherche qui ont été définies au chapitre 2, et proposer un résumé des résultats qui ont été obtenus.

La seconde partie de ce chapitre présentera les résultats spécifiques à l’étude des périodes de sécheresse-flash. En ciblant plus spécifiquement des périodes de déficit de précipitations responsables de sécheresses-flash, nous nous sommes intéressés aux différences entre une érablière à bouleau, une érablière à hêtre et une sapinière à érable face à ce type de stress, courts mais répétés¹.

4.1 Résumé des résultats obtenus et réponses aux questions de recherche

4.1.1 Phase 1 : préparation des données

Commençons par souligner que sur les 1 696 473 observations brutes collectées à la SBL entre 2017 et 2020, la répartition des données par peuplement étudié est homogène. Néanmoins, comme le peuplement *hardwood* (HW) ne comporte que deux blocs expérimentaux, contrairement aux peuplements *mixedwood* (MW) et *hardbeech* (HB) qui en ont trois, cela a introduit un léger déséquilibre dans le

1. Courcot, B, Bélanger, N et Lemire, D. (2023, 8 mai). L’expansion du hêtre dans l’érablière à bouleau jaune peut-elle s’expliquer par l’entremise du potentiel hydrique du sol? Aperçu d’une variable inexplorée. 90e congrès de l’ACFAS, Montréal, Québec <https://www.acfas.ca/evenements/congres/programme/90/600/632/c>

nombre d'observations par peuplement (figure 4.2).

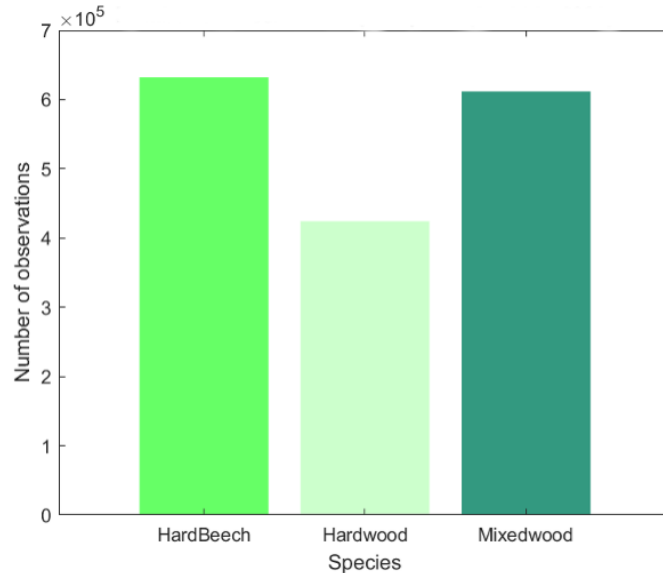


FIGURE 4.1 : Répartition initiale des observations par peuplement étudié, entre 2017 et 2020, avant la première phase méthodologique de nettoyage des données.

Ce qui ressort de cette première phase exploratoire, c'est une grande variabilité des données initiales. Ce constat est un enjeu qui ressort en science des données, et en particulier en sciences de l'environnement (Blair *et al.*, 2019). Dans notre contexte, nous avons mis en évidence des variabilités entre les mesures effectuées par les différentes sondes de température et de potentiel hydrique. La figure 4.2 illustre cette variabilité inter-sondes pour la mesure du potentiel hydrique du sol en représentant à titre d'exemple l'évolution temporelle du potentiel de la station01 en 2020. Les incertitudes ont été calculées et ajoutées au graphique.

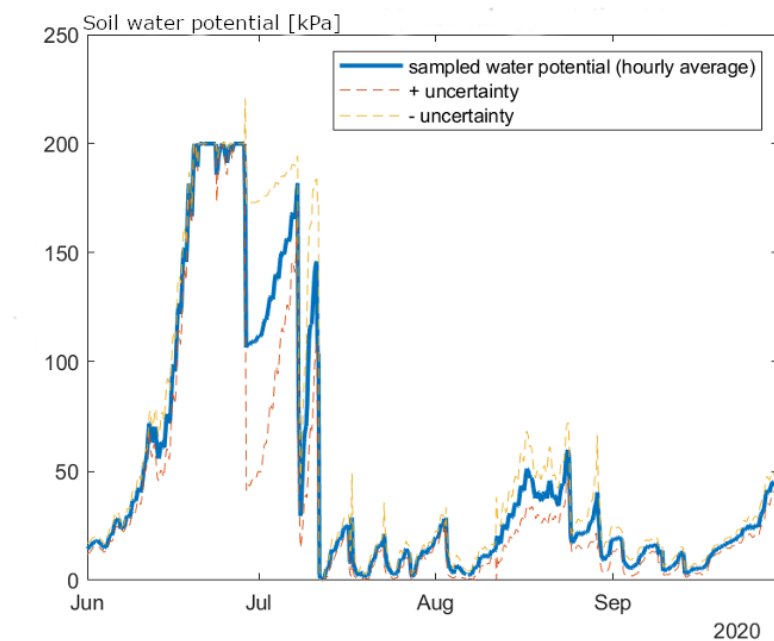


FIGURE 4.2 : Potentiel hydrique du sol moyenné (par heure) sur les deux sondes indépendantes (nommées C et D), avec les incertitudes associées (station01, année 2020). Pour chaque point, il s’agit de la moyenne des deux sondes.

En analysant alors chaque sonde, il est possible de repérer les périodes problématiques et de vérifier sur le terrain le bon fonctionnement des appareils de mesure. Dans le cas des mesures du potentiel hydrique de la station01, en 2020, en traçant Ψ_C et Ψ_D nous pouvons mettre en évidence des différences entre les deux sondes (figure 4.3), expliquant ainsi les incertitudes plus grandes à certaines périodes, en particulier en juillet 2020 (figure 4.2). Dans le rapport technique (Courcot, 2023, p.144), nous avons détaillé le calcul des différents types d’incertitude auxquelles nous sommes confrontés :

- l’incertitude dite de type A, liée à la mesure (erreur aléatoire) ;
- l’incertitude dite de type B, liée à l’appareil de mesure, et qui comprend l’incertitude due à l’erreur de justesse de l’appareil et celle liée à sa résolution, i.e. la plus petite graduation disponible pour la mesure.

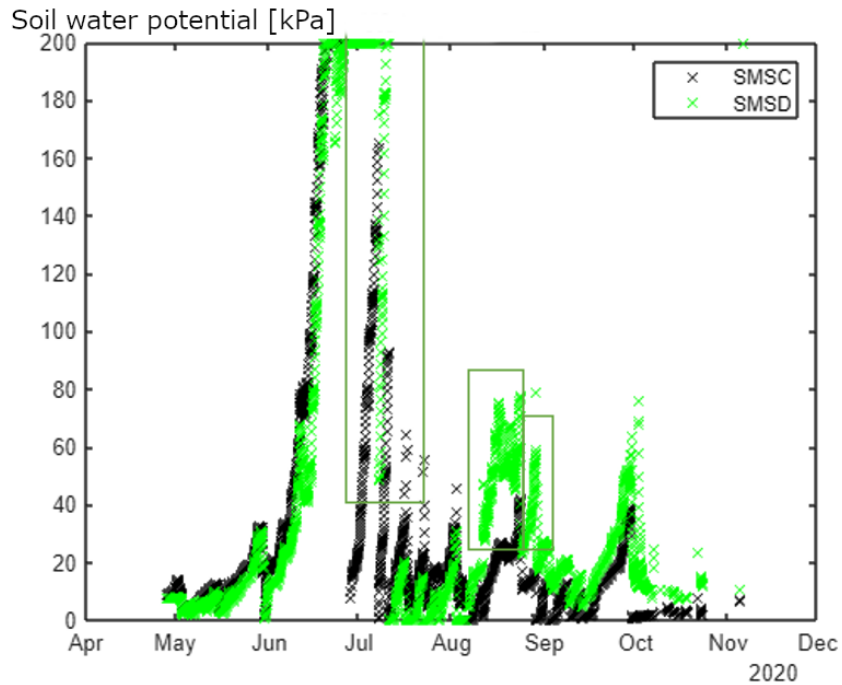


FIGURE 4.3 : Potentiel hydrique du sol mesuré par la sonde C (en noir) et la sonde D (en vert) pour la station01 en 2020.

Ce sont ces incertitudes qui sont représentées en pointillé sur la figure 4.2. Cette variabilité dans les données est donc due aux appareils de mesure, également parfois à des erreurs humaines ou de logistiques, et aussi à la variabilité spatiale des propriétés des sols qui peut induire des réponses différentes pour deux sondes identiques mais situées sur des lieux différents même si la distance les séparant est petite. Une analyse détaillée de la répartition des mesures de potentiel hydrique du sol pour chaque sonde utilisée a été réalisée (Courcot, 2023, p.18). Pour pallier cette variabilité, notre stratégie d'analyse s'est reposée sur deux spécificités expérimentales :

- une fréquence d'échantillonnage très haute qui correspond à une période de 15 minutes ;

- la possibilité d'augmenter l'échelle d'analyse : de la parcelle au peuplement.

Nous avons développé notre pipeline d'analyse pour pouvoir mener facilement les analyses aux trois échelles accessibles : au niveau de la parcelle, du bloc et du peuplement.

Pour la suite des analyses, nous avons fait le choix d'utiliser la moyenne journalière du potentiel hydrique du sol. Nous avons en effet montré que l'heure du jour n'impliquait pas de grandes variabilités dans la mesure (figure 4.4).

Les sondes de mesure utilisées ne permettent pas de mettre en évidence les variabilités diurnes et nocturnes du potentiel hydrique du sol, nous pouvons ainsi utiliser la moyenne journalière réalisée à partir de l'échantillonnage horaire.

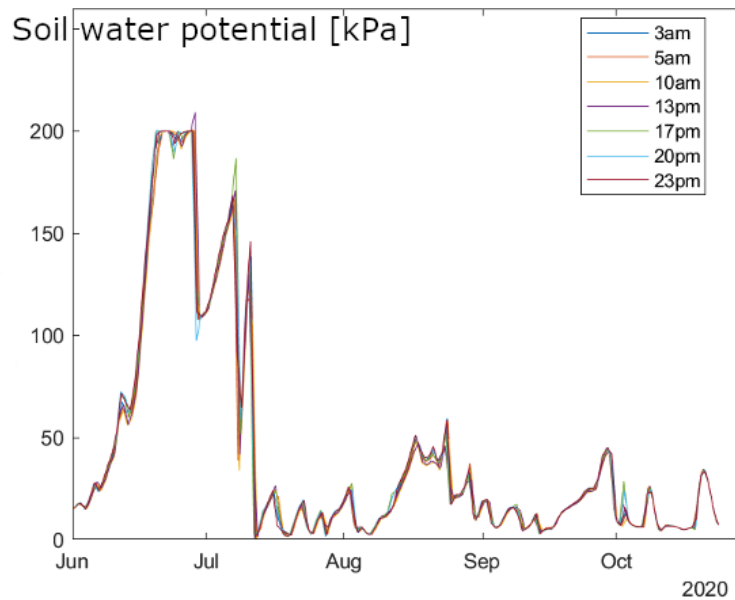


FIGURE 4.4 : Potentiel hydrique du sol mesuré en [kPa] à différentes heures de la journée, à l'échelle de la station01. Le potentiel a été moyenné sur les deux sondes de mesure (C et D). L'axe des abscisses représente un point temporel par jour, entre juin et octobre 2020.

Il est intéressant de noter que dans la littérature, lorsque le potentiel hydrique des feuilles est utilisé à la place du potentiel hydrique du sol, il est alors souvent

mesuré avant l'aube. En effet, comme le rappellent Coolong *et al.* (2012), il est généralement accepté que les potentiels hydriques du sol et des feuilles sont à l'équilibre à l'aube. Dans de nombreuses études le potentiel hydrique des feuilles est utilisé. Or, même si les deux mesures sont corrélées, il n'est pas conseillé de s'appuyer sur le potentiel hydrique des feuilles pour caractériser l'état hydrique du sol (Coolong *et al.*, 2012).

Enfin, nous avons fait le choix de régulariser les données en remplaçant les données manquantes par des valeurs interpolées (Courcot, 2023, p.28). Cette interpolation par BSpline cubique a été réalisée pour les mesures de température du sol (figure 4.5a) et pour les mesures du potentiel hydrique du sol (figure 4.5b).

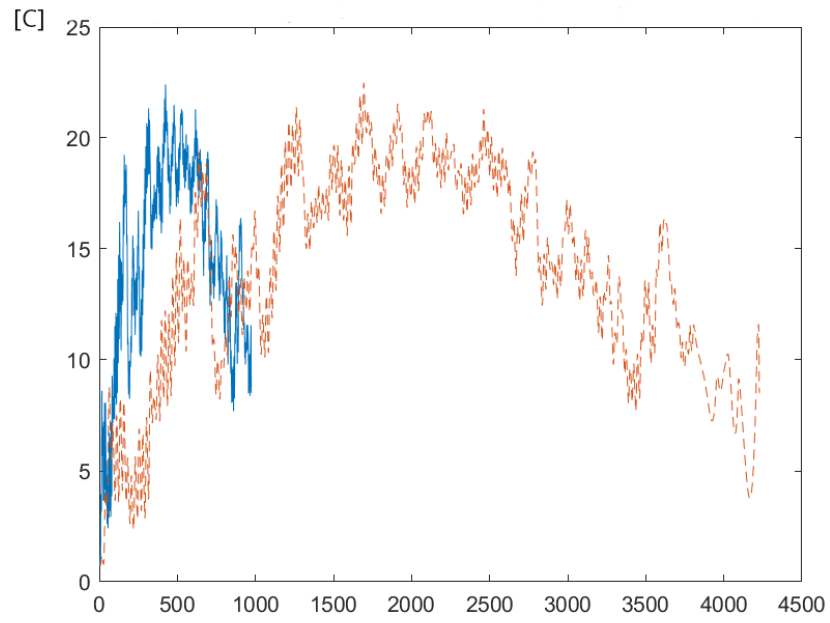
Sur la figure 4.5 sont représentées en bleu les données originales, échantillonnées sur les deux sondes de mesure respectives sur une heure. Nous constatons, pour la température (figure 4.5a) et le potentiel hydrique du sol (figure 4.5b), que les données originales comptent environ 1000 points temporels : des jours sont manquants. Après l'interpolation, nous obtenons 4227 points temporels, représentés en orange. Nous constatons que l'évolution des variables au cours du temps est reproduite après l'interpolation.

Cette interpolation est possible car l'échantillonnage des données se fait sur un intervalle de temps très court (de 15 minutes).

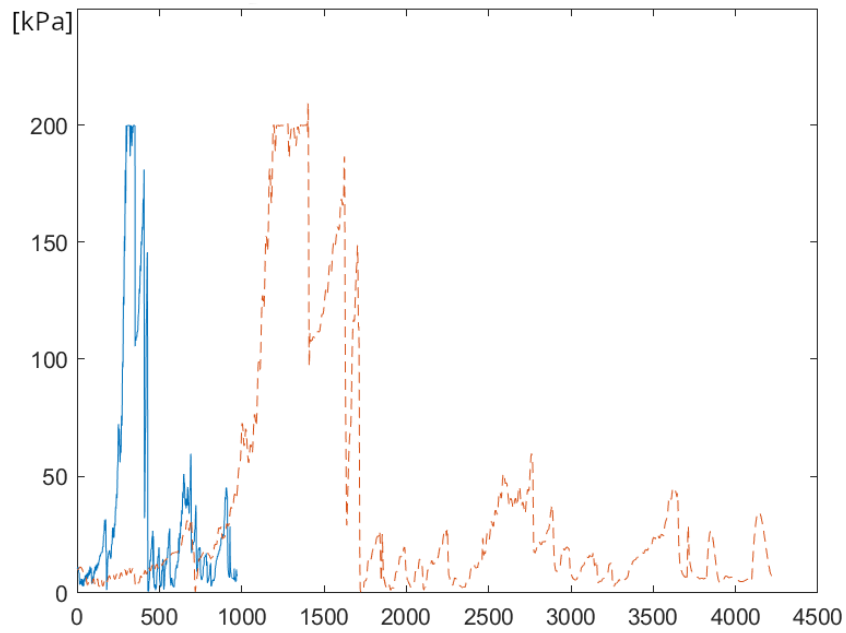
L'obtention de séries temporelles dites régularisées est une condition importante pour la suite de l'analyse.

Nous avons également appliqué un filtre passe-bas afin de lisser les données, c'est-à-dire d'augmenter le rapport signal sur bruit.

Nous avons testé plusieurs filtres et le filtre Savitzky-Golay (SG) s'est avéré être le plus performant. En effet, associé à une moyenne mobile sur trois jours, un filtre SG s'appuyant sur un polynôme de degré six a permis de ne pas tronquer les pics du potentiel hydrique du sol (Courcot, 2023, p.29).



(a) Mesures de la température du sol [°C]



(b) Mesures du potentiel hydrique du sol [kPa]

FIGURE 4.5 : Mesures avant régularisation (en bleu) et après régularisation (en orange, trait en pointillé). Données échantillonnées sur les deux sondes de température et de potentiel hydrique du sol, sur une heure (exemple pour la station01, année 2020).

Il faut de plus ajouter que l'ajustement du potentiel hydrique à la température du sol, bien qu'il n'entraîne pas une grande différence dans les mesures, est préférable pour suivre les recommandations de Irmak *et al.* (2016). À titre d'exemple, la figure 4.6 met en perspective le potentiel hydrique du sol sans ajustement (en pointillé) et avec ajustement pour le peuplement MW en 2020.

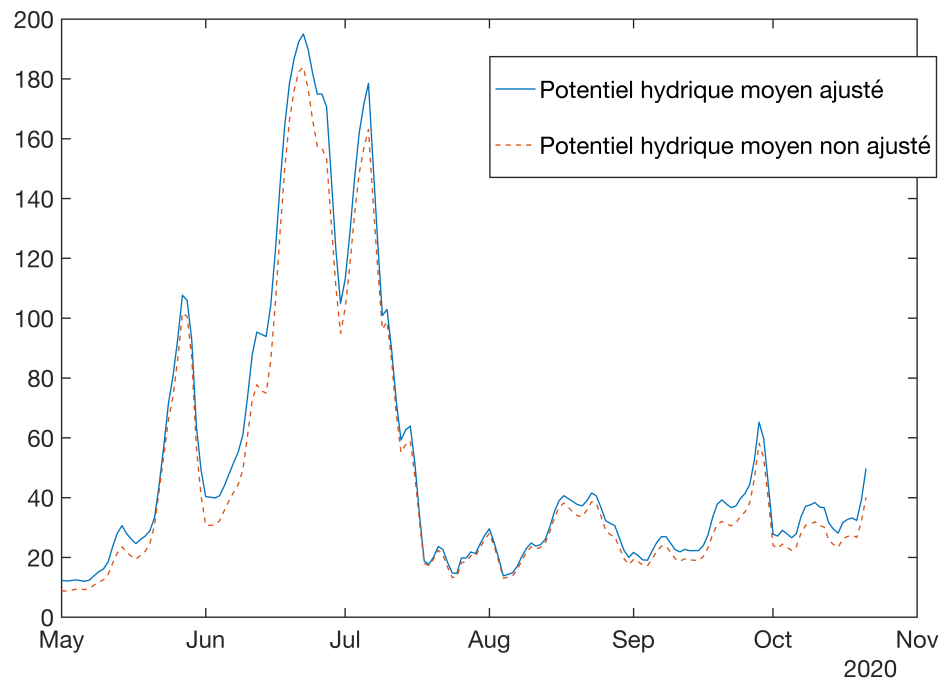


FIGURE 4.6 : Exemple du potentiel hydrique du sol ajusté à la température du sol (peuplement MW, année 2020).

Cette première phase d'analyse a permis également la préparation et la synchronisation des données météorologiques avec les mesures de potentiel hydrique et de température du sol.

La comparaison entre les données générées par le modèle BioSIM (Régnière et Saint-Amand, 2017) pour l'année 2020 et les données issues la même année de la station météorologique installée à la SBL a mené à plusieurs constats.

(1) Les précipitations ne sont pas fournies par les données de la station météo de la SBL.

(2) L'humidité relative de l'air générée par BioSIM (en orange) est plus basse que celle relevée à la SBL (en bleu), mais les mêmes tendances sont observées (figure 4.7).

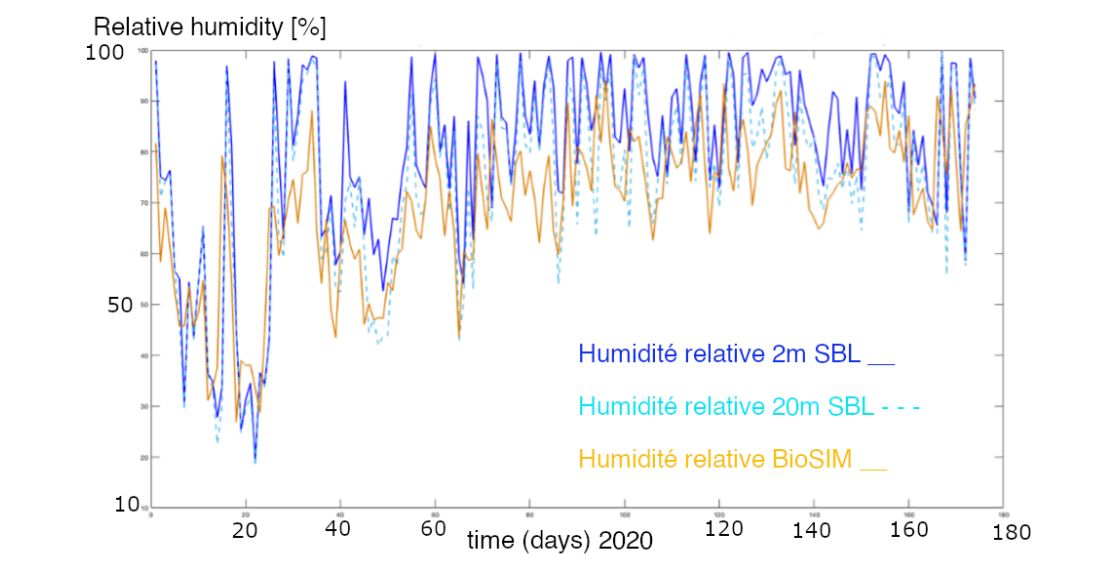


FIGURE 4.7 : Comparaison entre l'humidité relative générée par BioSIM (en orange) et l'humidité relative mesurée par la station météorologique de la SBL à 2 m et à 20 m de hauteur (en bleu) en 2020.

(3) La température de l'air générée par BioSIM (en orange) suit les mêmes tendances que les mesures faites à la SBL (en bleu), mais elle est toujours plus élevée (figure 4.8).

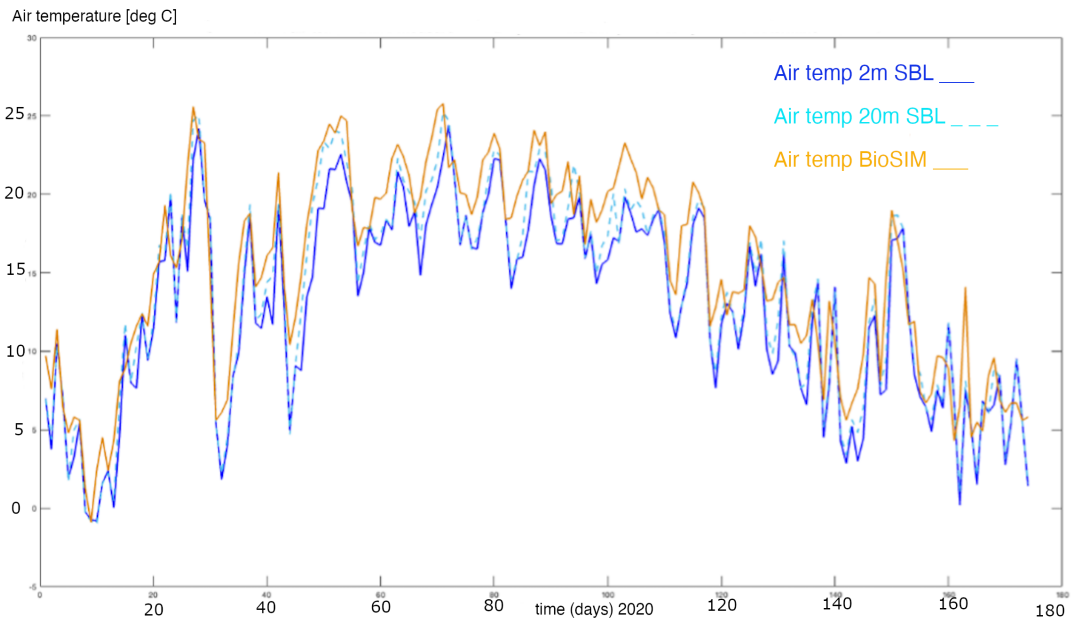


FIGURE 4.8 : Comparaison entre la température de l'air générée par BioSIM (en orange) et la température de l'air mesurée par la station météorologique de la SBL à 2 m et à 20 m de hauteur (en bleu) en 2020.

(4) Le graphique montrant l'évolution de la quantité d'eau dans le sol mesurée à différentes profondeurs (par la station météo de la SBL) avec sur le même graphique les précipitations (en orange) obtenues avec BioSIM permet de valider l'effet des précipitations. En effet, nous constatons qu'après chaque pluie, le contenu volumétrique en eau augmente, parfois faiblement, mais un pic est observable après chaque événement (figure 4.9).

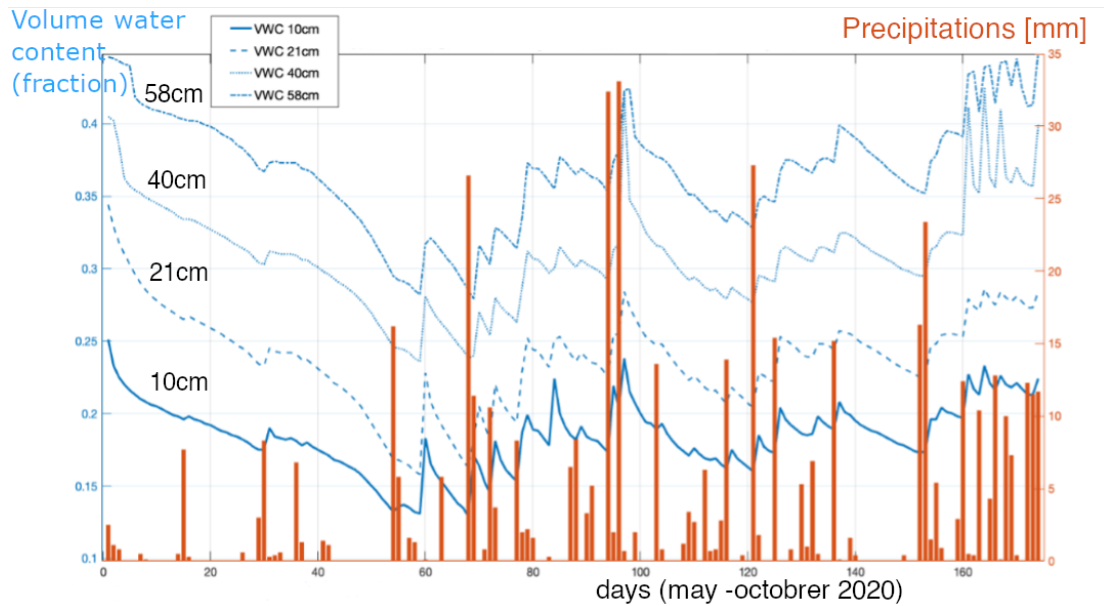


FIGURE 4.9 : Évolution du contenu volumétrique en eau dans le sol [%] , à différentes profondeurs (10 cm, 21 cm, 40 cm, 58 cm) et précipitations générées avec BioSIM (année de référence : 2020).

Enfin, cette première phase méthodologique a permis de mettre en évidence la spécificité de l'année 2020 par rapport à la quantité de précipitations reçues entre mai et novembre, en comparaison aux trois autres années d'intérêt (figure 4.10). En effet, les mois de mai, juin et septembre ont été particulièrement secs par rapport aux autres années, alors que le mois d'août a été plus arrosé et que le mois de juillet a été dans la moyenne des précipitations observées en 2017, 2018 et 2019.

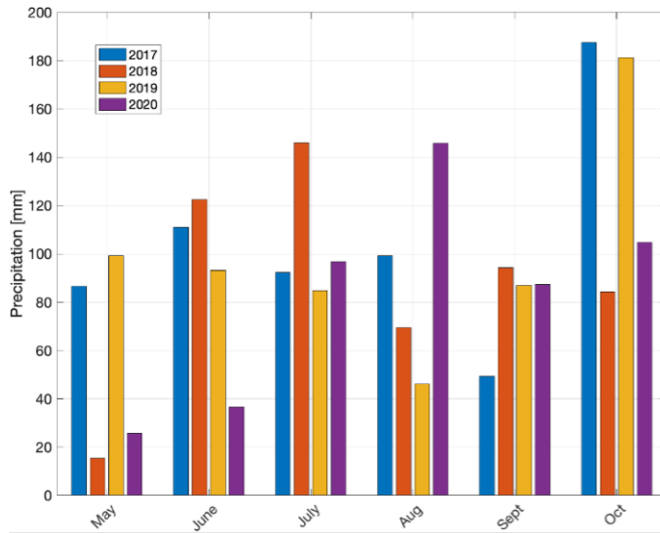
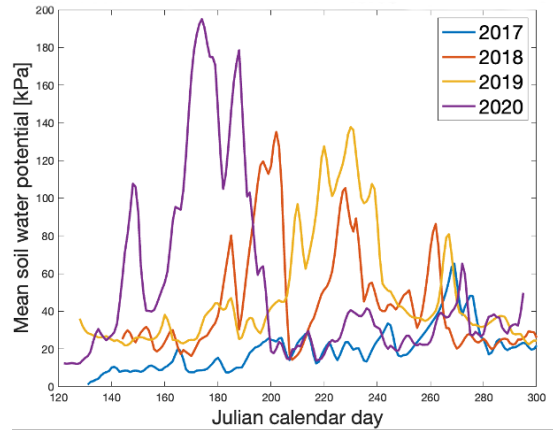


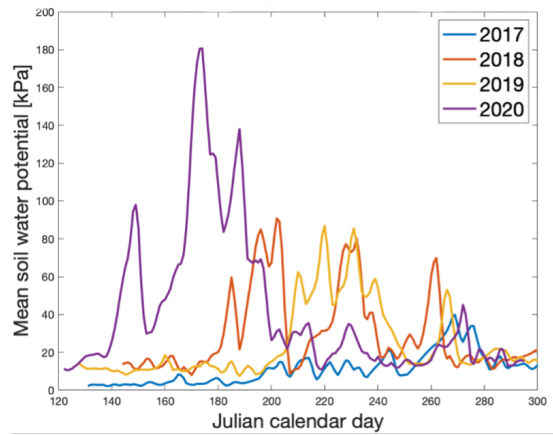
FIGURE 4.10 : Comparaison des précipitations [mm] entre 2017 et 2020 pour les mois de mai à octobre.

Pour conclure, à l'issue de cette première phase méthodologique, nous avons obtenu des séries temporelles régularisées, lissées et synchronisées entre elles, pour chaque peuplement et chaque année que nous avons considérés. Ainsi, l'évolution temporelle du potentiel hydrique moyenné sur une journée et ajusté à la température du sol peut être visualisée pour la sapinière à érable (figure 4.11a), pour l'érablière à bouleau (figure 4.11b) et l'érablière à hêtre (figure 4.11c).

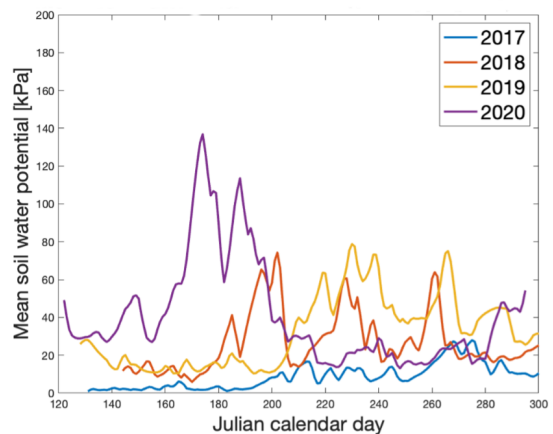
L'année 2020 se détache pour les trois peuplements : le potentiel hydrique du sol atteint des valeurs élevées qui se démarquent des autres années d'intérêt. Nous constatons également que pour l'année 2017 l'augmentation du potentiel hydrique arrive vers l'automne ; même si la figure 4.10 montre des précipitations plus faibles en septembre 2017 et que ce mois a connu une vague de chaleur après l'équinoxe du 22 septembre, l'amplitude du potentiel hydrique reste plus faible que les autres années. Ainsi, une période de latence entre l'installation des sondes et la production de données fiables semble nécessaire pour que les sondes de mesure se stabilisent et donnent des données exploitables.



(a) Peuplement MW



(b) Peuplement HW



(c) Peuplement HB

FIGURE 4.11 : Évolution temporelle du potentiel hydrique du sol moyenné sur une journée, entre 2017 et 2020, pour la sapinière à érable (MW), l'érablière à bouleau (HW) et l'érablière à hêtre (HB). 106

4.1.2 Phase 2 : sélection des variables prédictives

Des méthodes de classification utilisant les arbres de décision, affinées par l'utilisation de la causalité de Granger, ont été considérées afin de répondre à la première question de recherche relative aux variables capables de prédire le potentiel hydrique du sol Ψ . La démarche et les résultats obtenus ont été détaillés dans le rapport technique.

La phase 2 de la méthodologie nous a permis ainsi de répondre à la question de recherche QR1 : *Quelles sont les variables climatiques considérées dans notre étude qui ont des conséquences importantes sur l'évolution temporelle du potentiel hydrique du sol mesuré à la Station Biologique des Laurentides entre 2017 et 2020 ?*

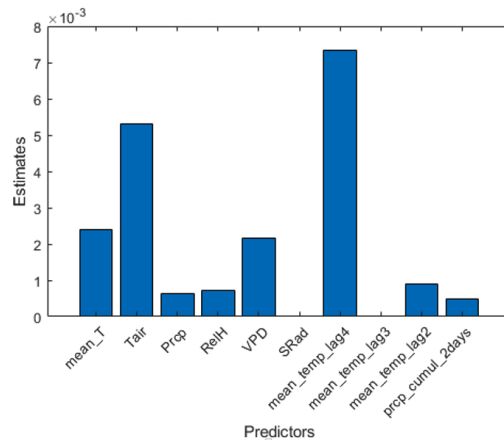
Nous avons montré l'importance de trois variables dans la prédiction du potentiel hydrique du sol :

- Température de l'air
- Température du sol, considérée avec un décalage de 4 jours
- Déficit de pression de vapeur (VPD)

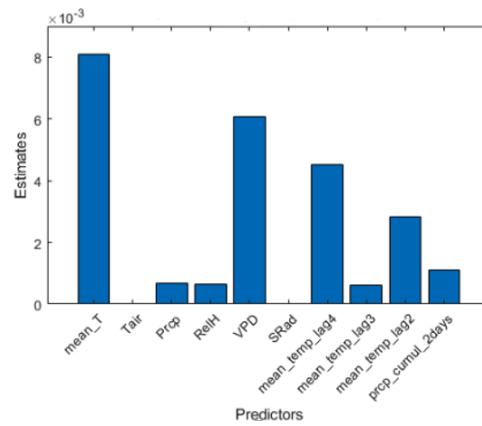
La figure 4.12a donne un exemple du résultat obtenu pour le peuplement MW, avec l'algorithme CART et l'arbre entraîné avec les données de 2020. Ce résultat est comparé aux résultats obtenus avec le même algorithme pour les peuplements HW (figure 4.12b) et HB (figure 4.12c).

L'axe des ordonnées des trois graphiques de la figure 4.12 représente l'importance des variables prédictives du potentiel hydrique du sol. Cette valeur se définit à partir du risque associé à chaque noeud de l'arbre de décision. Ce risque est calculé en utilisant la probabilité associée au noeud considéré et son impureté (index de Gini)².

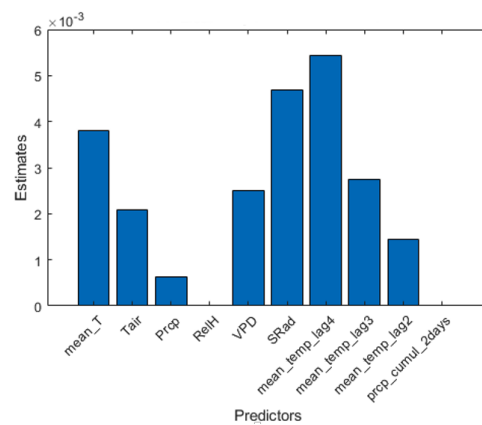
2. Se référer à l'Annexe V. du rapport technique (Courcot, 2023).



(a) Sapinière à érable



(b) Érablière à bouleau



(c) Érablière à hêtre

FIGURE 4.12 : Importance des variables exogènes dans la prédiction du potentiel hydrique du sol pour les trois peuplements étudiés.

Suite à l'analyse de la causalité de Granger entre le potentiel hydrique et les variables exogènes, nous avons introduit les variables relatives à la température moyenne du sol considérées sur un décalage cumulé de quatre, trois et deux jours. Pour la sapinière à érable, nous constatons que la température du sol considérée sur quatre jours a un pouvoir prédictif du potentiel hydrique supérieur à la température moyenne du sol considérée sur un seul jour. Le résultat confirme celui obtenu après l'analyse de la causalité au sens de Granger. Nous le rappelons, ce type de causalité permet de déterminer, non pas la corrélation entre les variables, mais quelles séries temporelles causent une variation de la série temporelle du potentiel hydrique, et combien de points temporels sont nécessaires pour permettre une meilleure prédiction de cette même série. En d'autres termes, nous avons montré que la connaissance des quatre derniers points temporels de la température du sol, par rapport au temps t , permet une meilleure prédiction du potentiel hydrique du sol pour ce même temps t . Notre unité de temps étant rapportée à une journée, les quatre derniers points temporels font référence à quatre jours.

Les arbres de décision considèrent les données de manière statique. Nous avons donc créé artificiellement de nouvelles variables qui représentent une certaine temporalité de la température du sol. Cette approche s'avère concluante puisque nous retrouvons la même conclusion que celle faite à la suite de l'analyse de la causalité de Granger.

Pour la sapinière à érable, la température de l'air et le déficit de pression de vapeur (VPD) ressortent aussi comme des variables importantes dans la prédiction du potentiel hydrique du sol.

Pour l'érablière à bouleau, ces trois variables sont également importantes mais le décalage (antérieur) cumulé sur quatre jours semble avoir moins d'impact sur la température du sol. De même, le VPD semble être important dans la prédiction de $\Psi(t)$.

Pour l'érablière à hêtre, le décalage antérieur de quatre jours pour T_{sol} permet

une meilleure prédiction du potentiel hydrique. Pour nous assurer de l'absence d'un raisonnement circulaire par rapport à l'ajustement du potentiel hydrique à la température du sol, les analyses ont également été faites avec le potentiel non-ajusté (Courcot, 2023, p.63). Les résultats obtenus ne diffèrent pas.

L'irradiance solaire (SRad) qui ne ressortait pas pour MW et HW apparaît avoir davantage d'importance pour ce peuplement dans la prédiction de $\Psi(t)$.

Soulignons que le VPD est fortement corrélé à l'humidité relative (RelH) et à l'irradiance solaire, avec des coefficients de corrélation respectifs de -0.81 et 0.77 (Courcot, 2023, p.46). Ainsi, leur pouvoir prédictif est atténué lorsque le VPD est considéré par les arbres de décision. Ce constat est vérifié lorsque nous testons un algorithme de classification CART qui permet la sélection préalable des variables prédictives ; le VPD devient alors le prédicteur le plus important par rapport à RelH et SRad.

Enfin, pour les trois peuplements les précipitations ne semblent pas avoir une forte importance prédictive, mais elles ressortent néanmoins dans l'analyse avec les arbres de décision, lorsque les prédicteurs ne sont pas sélectionnés au préalable (figure 4.12).

Ce résultat est de prime abord contre-intuitif, car nous nous attendions à un pouvoir prédictif des précipitations qui soit élevé. Or, ce n'est pas le cas. Plusieurs arguments peuvent expliquer ce résultat.

(1) Les précipitations ont été calculées à partir du modèle BioSIM, qui interpole les mesures en considérant des stations météorologiques proche de la SBL. Or, les précipitations ont un effet très localisé qui n'est pas suffisamment pris en compte dans l'approche choisie. Le fait que la station météorologique installée à la SBL ne mesurait pas en 2020 les précipitations n'a pas permis une comparaison entre les mesures calculées et celles réellement collectées à la SBL.

(2) Le potentiel hydrique est mesuré grâce à deux sondes enfouies à 10 cm dans le sol. En prenant en compte l'interception des précipitations par les feuilles et l'évaporation de l'eau à leur surface, la quantité de précipitations qui rentre en contact avec le sol est inférieure à celle mesurée par la station météorologique et varie également beaucoup. L'importance des précipitations est donc diminuée par la mesure imprécise de celles-ci.

Soulignons que les variations de la quantité de précipitations dépendent de la nature de la canopée et du positionnement des parcelles. Cela s'est avéré un facteur important sur la décomposition de la litière de l'érable à sucre à la SBL (Bélanger *et al.*, 2019).

(3) Nous pouvons avancer la possibilité que l'influence des précipitations soit masquée par l'éventuelle interférence d'une autre variable elle-même influencée par celles-ci.

Ainsi, inclure les précipitations comme variable permettant une meilleure prédiction du potentiel hydrique du sol n'est pas concluant. Cependant, pour les autres phases de l'analyse nous avons continué à les intégrer comme variable exogène, car les ajouter n'impacte pas le modèle prédictif. Celui-ci n'est pas amélioré, mais il n'est pas dégradé (se référer au rapport technique). Donc dans l'hypothèse d'une meilleure mesure des précipitations dans le futur, cela permettra d'ajuster le modèle en conséquence.

4.1.3 Phase 3 : analyse des changements d'état du potentiel hydrique du sol

La seconde question de recherche qui a été définie est la suivante : *Pouvons-nous définir différents états de ce potentiel hydrique du sol et déterminer les fréquences et les probabilités associés aux changements d'états ainsi mis en évidence ?*

Grâce aux chaînes de Markov à temps discret, il est possible d’expliciter des systèmes dynamiques qui sont régis par des processus stochastiques. En exploitant la temporalité des données du potentiel hydrique du sol et en nous appuyant sur le postulat de la propriété de Markov, nous avons pu analyser les transitions qui ont eu lieu durant chaque année d’intérêt. Cette analyse s’est faite en calculant les probabilités associées à chaque changement d’état observé et en évaluant le temps espéré pour atteindre ces états. Pour cela, dans la continuité de la deuxième phase méthodologique, trois états du potentiel ont été définis en fixant des seuils arbitraires, ne pouvant nous référer à des études similaires. Ces seuils, nous l’avons précisé, ont été évalués en nous appuyant sur l’évolution de la densité du potentiel hydrique entre 2017 et 2020, et ils ont été ajustés à la suite de l’analyse utilisant les arbres de décision. Nous avons également considéré les normes agricoles pour l’irrigation des cultures (Shock et Wang, 2011), normes qui n’existent pas en foresterie et qui rendent donc le choix de nos seuils arbitraires. Les résultats ont été obtenus avec les seuils suivants :

- état *Small* : Ψ compris entre 0 et 40 kPa (exclu)
- état *Medium* : Ψ compris entre 40 et 100 kPa (exclu)
- état *High* : Ψ supérieur ou égal à 100 kPa

D’une année à l’autre, nous avons pu isoler des différences dans les transitions d’état au sein d’un même peuplement. En particulier pour la sapinière à érable en 2019, il fallait 81 jours pour atteindre un état élevé de potentiel hydrique du sol en partant d’un état moyen (figure 4.13). En 2020, pour la même transition, il ne fallait plus que 63.5 jours. De plus, cette transition était plus probable en 2020 qu’en 2019, même si ces probabilités demeurent faibles au regard de celles des transitions *intra-état*.

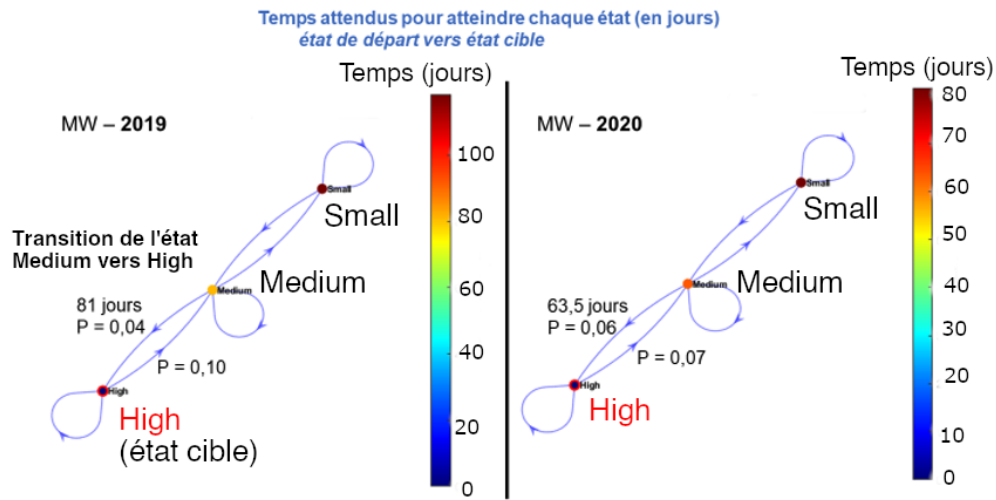


FIGURE 4.13 : Probabilités de transition d'état (notée P) et temps attendu pour atteindre chaque état (en jours) pour la sapinière à érable (MW) . Comparaison entre 2019 et 2020.

Cette approche a permis également de mettre en évidence des différences entre la sapinière (MW) et l'érablière à hêtre (HB) dans les transitions Medium \rightarrow High et High \rightarrow Medium (figure 4.14). En 2020, il fallait moins de jours à la sapinière à érable, comparativement à l'érablière à hêtre, pour basculer d'un potentiel considéré comme moyen vers un potentiel élevé qui caractérise l'existence d'un stress hydrique ; et ce basculement a été plus probable pour la sapinière. De plus, pour cette même année, la sapinière à érable est restée plus longtemps dans un état élevé du potentiel que l'érablière à hêtre (14 jours *vs* 6 jours).

Small, Medium et High : [0;40[, [40;100[et supérieur ou égal à 100 kPa

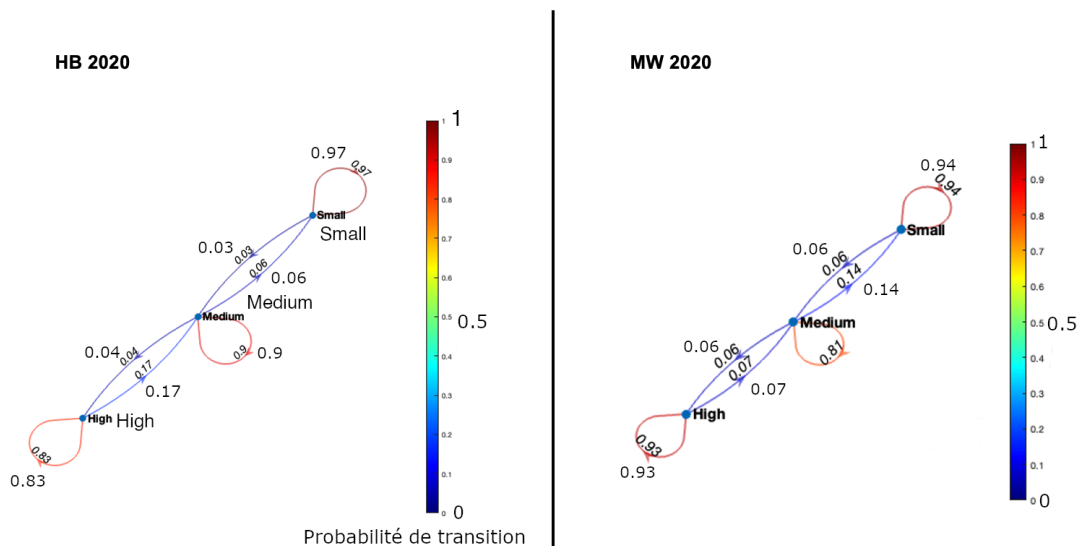


FIGURE 4.14 : Probabilités de transition d'état pour la sapinière à érable (MW) et pour l'érablière à hêtre (HB) en 2020.

Small, Medium et High : [0;40[, [40;100[et supérieur ou égal à 100 kPa

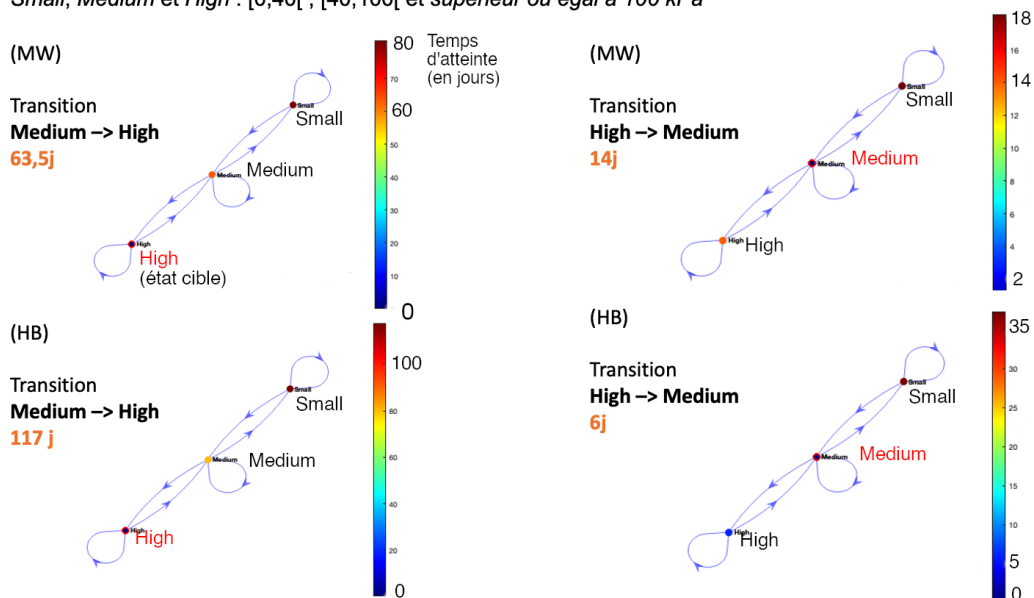


FIGURE 4.15 : Comparaison des temps attendus pour atteindre les états *Medium* et *High* du potentiel hydrique du sol, pour la sapinière à érable (MW) et l'érablière à hêtre (HB) en 2020.

Cette analyse a mis en évidence la résistance plus élevée de l'érablière à hêtre face aux changements d'état du potentiel hydrique du sol en comparaison avec la sapinière. Cette résistance aux changements est le signe d'une plus grande résilience écologique, telle qu'elle a été définie au chapitre 1.

Dans le cadre de notre étude, l'utilisation des chaînes de Markov nous a semblé pertinente pour décrire l'évolution du potentiel hydrique du sol. Les résultats obtenus donnent une première tendance quant à une possible plus grande résilience de l'érablière à hêtre face à des stress hydriques. L'enjeu reste encore ici de définir plus précisément les seuils de transition.

Une des principales limites de cette analyse, nous l'avons souligné précédemment, réside dans la propriété de Markov qui stipule que la prédiction du futur est entièrement contenue dans l'état présent du processus. Or, en voulant incorporer dans notre modèle les variables exogènes, nous nous heurtons à un problème conceptuel. En effet, l'étude de la causalité de Granger a montré un décalage de quatre jours entre la variation maximale du potentiel et celle de la température du sol. Ainsi, la nécessité de prendre en compte la corrélation entre les variables au cours du temps est primordiale pour pouvoir décrire les phénomènes physiques observés. À travers le développement d'un modèle autorégressif de type ARIMAX, nous visions cet objectif.

4.1.4 Phase 4 : exploration des séries temporelles

Avant de déterminer les modèles ARIMAX, pour chaque année d'intérêt, les calculs des coefficients d'autocorrélation du potentiel hydrique ont été réalisés³. Entre 2017 et 2020, la valeur des coefficients d'autocorrélation du potentiel hydrique a augmenté pour la sapinière à érable (figure 4.16). Nous montrons qu'il en a été de même pour les deux autres peuplements.

3. Si nous considérons un décalage de 1 jour, nous parlerons alors de l'autocorrélation au *lag-1*.

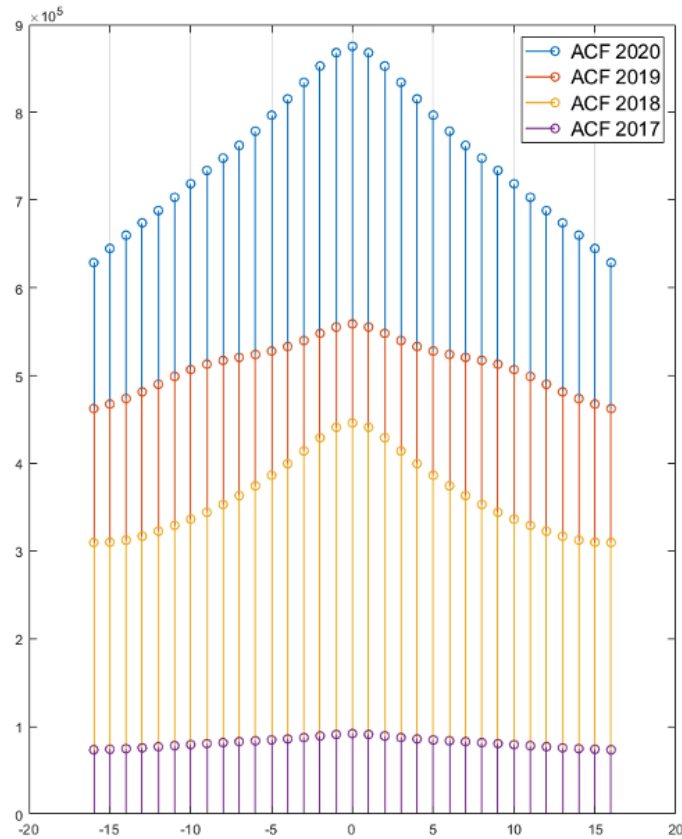


FIGURE 4.16 : Représentation de l'autocorrélation pour la sapinière à érable (MW) de 2017 à 2020. En abscisse sont représentés les décalages en jours (entre -15 jours et 15 jours) et en ordonnées, la valeur des coefficients d'autocorrélation (non normalisés).

Si nous traçons la différence des coefficients obtenus en 2020 et en 2017 pour les trois peuplements, nous mettons en évidence l'augmentation de l'autocorrélation, en particulier celle considérée au *lag-1* (figure 4.17).

La plus forte augmentation entre 2017 et 2020 a été mise en évidence pour la sapinière à érable (en bleu), puis l'érablière à bouleau (en orange) et enfin l'érablière à hêtre (en jaune).

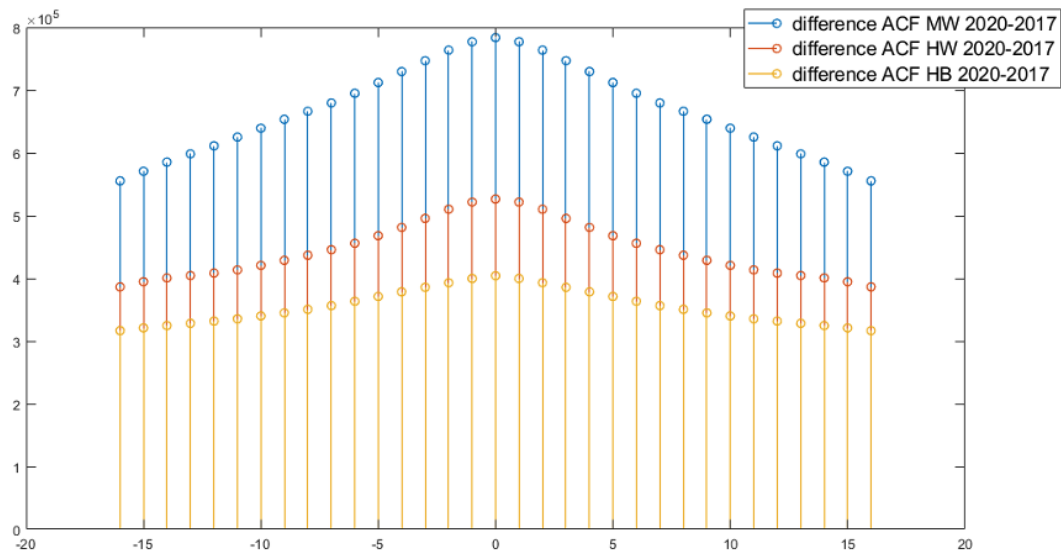


FIGURE 4.17 : Différence entre les coefficients d'autocorrélation obtenus en 2020 et en 2017, pour les trois peuplements étudiés (MW, HW et HB). En abscisse sont représentés les décalages en jours (entre -15 jours et 15 jours) et en ordonnées, la différence entre les coefficients.

Nous l'avons détaillé au chapitre 1, il a été montré que l'augmentation de l'autocorrélation au cours du temps était un signe de rapprochement d'un point de bifurcation. Or, nous avons montré que l'érablière à hêtre a connu la moins grande augmentation de son autocorrélation entre 2017 et 2020, en comparaison avec la sapinière à érable. Nous pourrions ainsi formuler l'hypothèse que la sapinière, entre 2017 et 2020, s'est davantage rapprochée d'un changement de régime de son potentiel hydrique du sol.

Cette tendance est confirmée pour les trois peuplements par une augmentation du coefficient d'asymétrie (*skewness*) entre 2017 et 2020 (tableau 4.1). Au fil des années, les données de potentiel hydrique du sol se répartissent davantage à droite de leur valeur moyenne, exhibant alors une répartition de plus en plus asymétrique, s'éloignant d'une distribution normale. Cette distribution est confirmée par l'analyse des coefficients d'aplatissement (*kurtosis*) qui ont des valeurs supérieures à

TABLEAU 4.1 : Coefficients d’asymétrie calculés pour la sapinière à érable (MW), l’érablière à bouleau (HB) et l’érablière à hêtre (HB), en 2017 et 2020.

	MW	HW	HB
2017	1.3	1.2	0.9
2020	1.6	1.7	1.5

3, signe d’une distribution dite leptokurtique, comportant davantage de valeurs extrêmes.

Cependant, comme le soulignent Dakos *et al.* (2015), il faut être prudent : « (...) *positive trends in both variance and autocorrelation can be observed when there is some memory in the temporal evolution of environmental shocks.* »

Cette mémoire peut se manifester au fil des années, mais aussi à travers des temporalités plus courtes. L’analyse des corrélations-croisées entre le potentiel hydrique et les variables exogènes de notre modèle a mis en exergue la spécificité de chaque peuplement face à cette notion de *mémoire*.

En effet, à l’issue de cette analyse, nous avons montré que chaque peuplement présente des corrélations différentes au cours du temps (figure 4.18). Même si pour les variables, à l’exception de l’humidité relative, les maxima représentés par des couleurs plus foncées se situent à des décalages souvent identiques, les corrélations s’estompent à des jours différents. Cela est le cas pour le VPD, dont l’influence se fait sentir pour la sapinière jusqu’à six jours après la valeur la plus élevée du potentiel. Pour l’érablière à bouleau, celle-ci s’estompe après cinq jour. Pour l’érablière à hêtre, après quatre jours, cette influence a diminué comparativement au jour (t), i.e. quand le décalage (*lag*) est nul. Si nous regardons la température moyenne du sol, elle influence le potentiel hydrique jusqu’à trois jours avant son pic d’intensité. Nous avons mis en évidence une temporalité de quatre jours avec la causalité de Granger, ce qui reste comparable lorsque nous regardons les coefficients

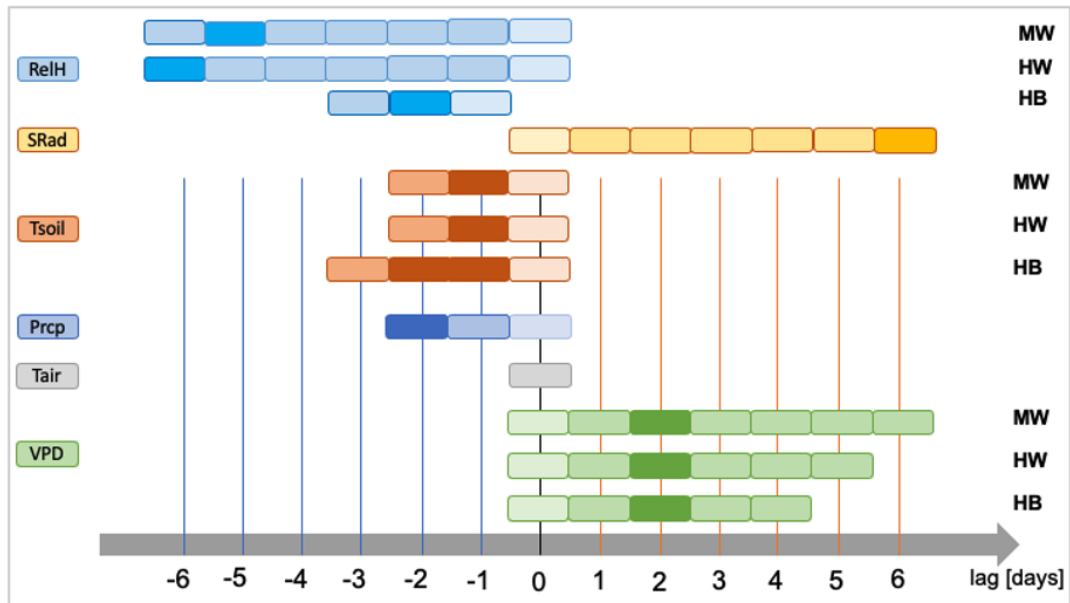


FIGURE 4.18 : Schéma résumant les différentes corrélations-croisées entre le potentiel hydrique du sol et les variables exogènes du modèle. Les calculs ont été menés sur les données de 2020, pour chaque peuplement considéré. Le coefficient maximal de corrélations-croisées est représenté par une couleur plus foncée.

de corrélations-croisées qui restent élevés même cinq jours après. À $(t-4\text{jours})$ le coefficient devient légèrement inférieur à celui obtenu au jour (t) , mais il reste très proche.

L'influence de l'humidité relative est maximale deux jours avant le pic du potentiel hydrique pour l'érablière à hêtre, mais celle-ci est plus longue pour l'érablière à bouleau et la sapinière (six et cinq jours avant, respectivement). Ces deux peuplements présentent des corrélations croisées très comparables, tandis que le potentiel hydrique de l'érablière à hêtre semble se détacher pour certaines variables exogènes. Pour les trois peuplements, le potentiel hydrique et la température de l'air mesurés au jour (t) ont leur corrélations-croisées maximales, ne montrant donc aucune influence antérieure ou postérieure de la température de l'air sur le potentiel hydrique du sol.

Il faut souligner que certaines variables confondantes, qui n'apparaissent pas dans nos questions de recherche, peuvent exercer également une influence sur les résultats obtenus. Par exemple, l'exposition des 32 stations considérées, leur pente, ou le type de sol à la profondeur des sondes peuvent impliquer une variation des propriétés à l'échelle du sol et un effet sur le potentiel hydrique entre les blocs porteurs de différents types de végétation.

La température du sol nous permet cependant de moduler ces variables confondantes. De plus, en considérant la composition en espèces, laquelle est déterminée par les régimes hydrique et thermique, donc liés à la position des stations, nous avons indirectement l'influence de ces variables, dont nous faisons aussi l'hypothèse qu'elles sont distribuées de manière égale afin de minimiser leur impact dans les analyses.

4.1.5 Phase 5 : tests et validations des modèles prédictifs ARIMAX

Rappel de la question de recherche QR3 : *Est-ce qu'un modèle prédictif de type ARIMAX, développé avec des données provenant de la SBL, permet de prédire l'évolution temporelle du potentiel hydrique du sol de ce même site, en considérant l'influence de variables climatiques exogènes ?*

En implémentant la méthode de Box-Jenkins nous avons pu estimer, tester et valider un modèle de type ARIMAX, pour chaque peuplement considéré. Chaque modèle est une solution possible, et non *la* solution espérée. Pour les développer, nous avons mis en place diverses stratégies méthodologiques qui se sont adaptées à la réalité de nos données expérimentales. Sans modèle de référence il nous est impossible de confronter les nôtres. Cependant, l'objectif était de répondre à la question de recherche QR3, et nous pouvons répondre que oui, un modèle de type ARIMAX développé avec les données de la SBL permet de considérer l'influence que

les variables climatiques exogènes ont sur le potentiel hydrique et de les incorporer dans un modèle prédictif.

Nous pouvons résumer les différents résultats que nous avons obtenus à la suite des phases méthodologiques 4 et 5.

(1) Des modèles simples autorégressifs de type AR(1) ou MA(1) ne permettent pas de décrire l'évolution temporelle du potentiel hydrique du sol, comme l'illustrent les figures 4.19 et 4.20. Les valeurs élevées des coefficients AIC et BIC en témoignent. Bien que le modèle AR(1) semble donner un bon ajustement, un décalage temporel est introduit entraînant alors l'obtention de coefficients AIC et BIC élevés. C'est en combinant les deux modèles que ce décalage pourra être corrigé.

$$(1 - \phi_1 L)y_t = \varepsilon_t$$

Paramètres	Valeur
Constant	0
AR{1}	1
Variance	100
AIC	1198
BIC	1204

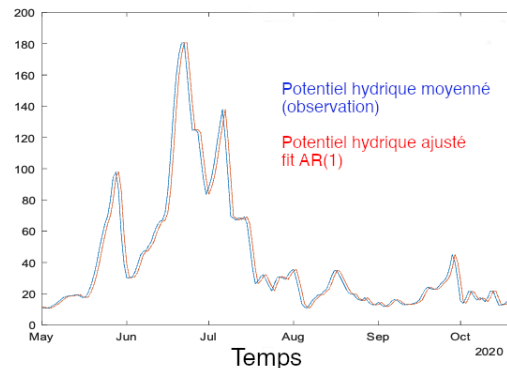


FIGURE 4.19 : Modèle autorégressif d'ordre 1 : AR(1).

$$y_t = (1 + \theta_1 L)\varepsilon_t$$

Paramètre	Valeur
Constant	0
MA{1}	0.95558
Variance	1173.5468
AIC	1674
BIC	1681

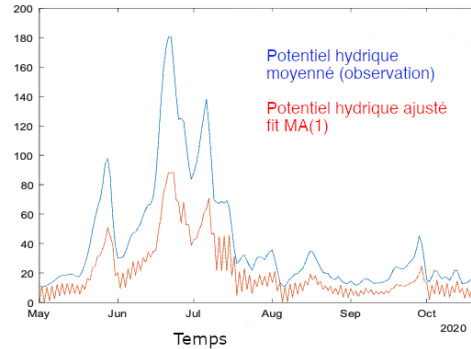


FIGURE 4.20 : Modèle de moyenne mobile d'ordre 1 : MA(1).

(2) Les séries chronologiques considérées ne sont pas stationnaires et il est nécessaire de les dériver pour s'assurer de leur stationnarité et pouvoir ainsi appliquer la démarche de Box-Jenkins.

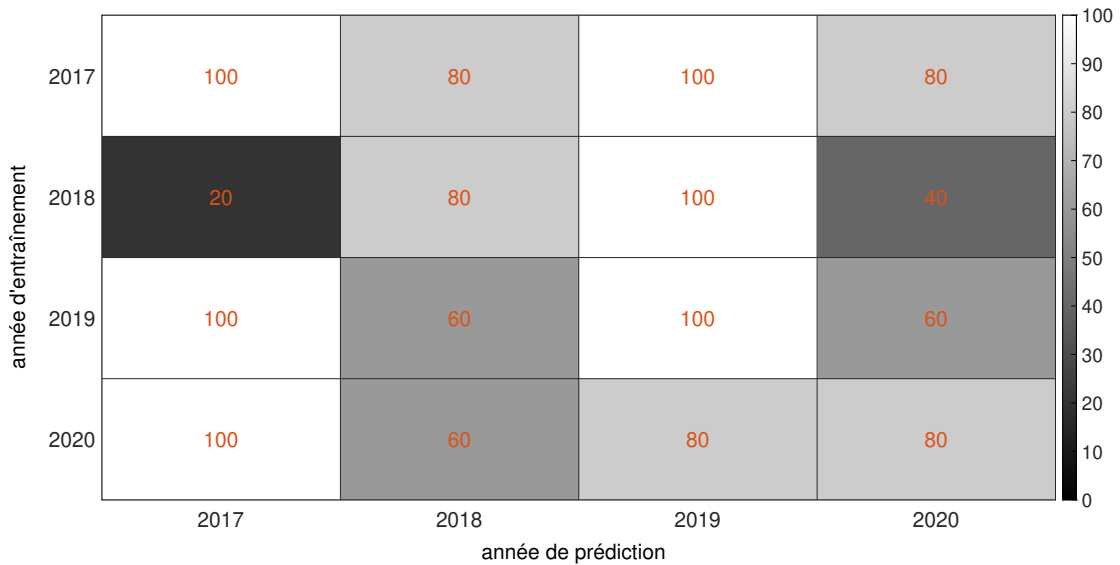
(3) Introduire la variable catégorielle qui caractérise qualitativement l'état du potentiel hydrique du sol (*Small, Medium, High*) permet d'améliorer le pourcentage de prédictions exactes, en comparaison aux résultats obtenus avec les valeurs numériques du potentiel hydrique.

(4) Selon les périodes temporelles considérées l'année d'estimation du modèle qui permet d'obtenir les meilleurs pourcentages de prédictions exactes varie.

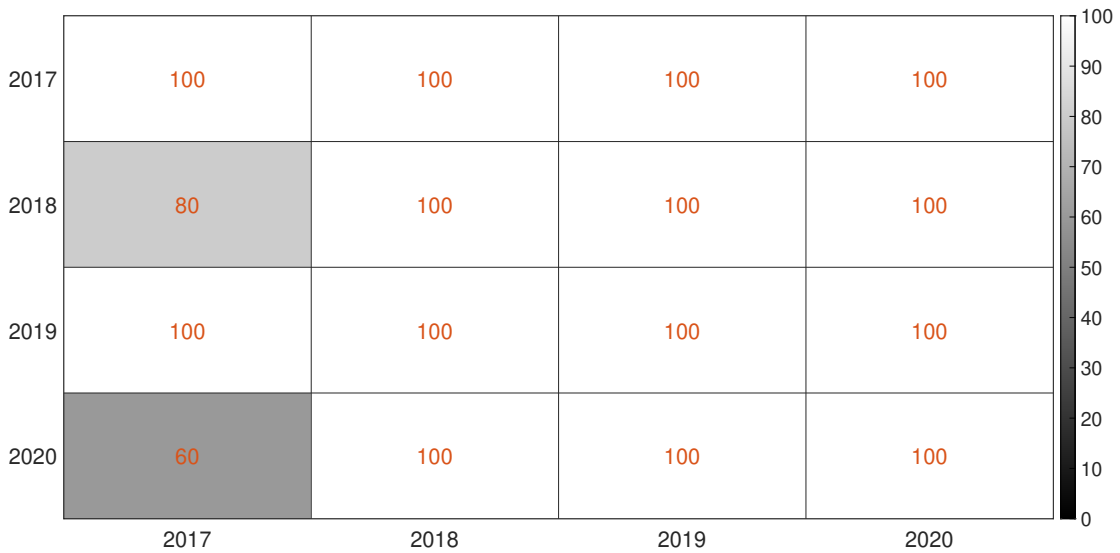
Des tests systématiques réalisés sur toutes les années d'intérêt et pour les trois peuplements ont permis d'obtenir une carte thermique du pourcentage d'exactitude des prédictions obtenues selon tous les cas considérés. Ainsi, pour chaque peuplement, il est possible de visualiser l'impact de l'année d'estimation du modèle sur l'exactitude des prédictions (figure 4.21). En choisissant une fenêtre de prédiction très courte, de cinq jours, nous montrons que les prédictions faites pour l'éraablière à hêtre (figure 4.21b) sont plus performantes que celles réalisées pour la sapinière

à érable (figure 4.21a). Nous avançons l'hypothèse que la série chronologique de la sapinière présente une non-stationnarité locale plus importante que l'érablière à hêtre, ce qui pousse le modèle ARIMAX a ses limites d'application, impliquant alors un pourcentage de prédictions exactes moins élevé.

En comparant les années d'entraînement du modèle, nous pouvons constater que 2019 permet d'obtenir des pourcentages élevés de prédictions exactes des états du potentiel hydrique du sol. Néanmoins, l'année 2020 donne également de très bons résultats. La prédiction de l'année 2017 donne des résultats surprenants, notamment pour la sapinière, nous incitant à la prudence quant à l'analyse de cette année qui est la première année de l'expérimentation. Cela confirme que les sondes pourraient nécessiter un temps de stabilisation avant l'exploitation des données collectées.



(a) Carte pour la sapinière à érable (MW)

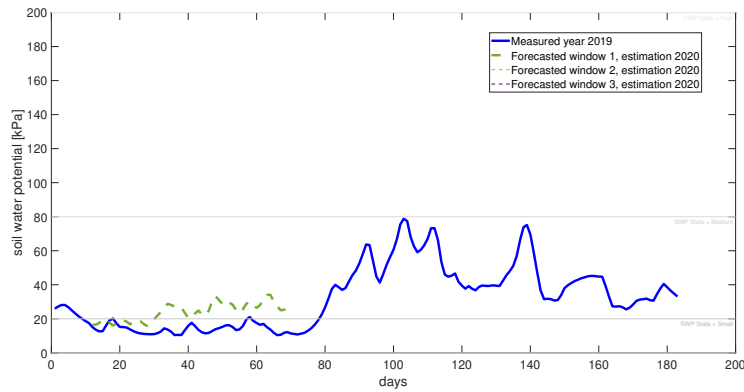


(b) Carte pour l'érablière à hêtre (HB)

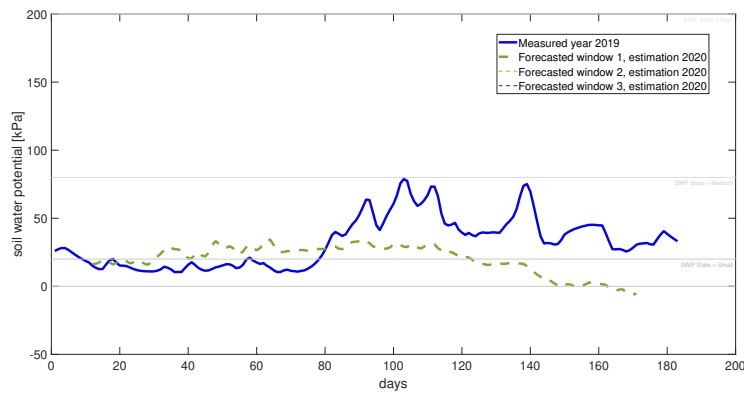
FIGURE 4.21 : Cartes thermiques du pourcentage de prédictions exactes, pour deux peuplements. En abscisse, l'année de prédiction et en ordonnées, l'année d'estimation du modèle. Les pourcentages sont représentés à l'aide d'une échelle de gris et les valeurs obtenues sont précisées en rouge.

(5) Les différents tests réalisés ont montré que l'horizon de prédiction optimal est de l'ordre de 60 jours (exemple à la figure 4.22a avec une fenêtre de 60 jours).

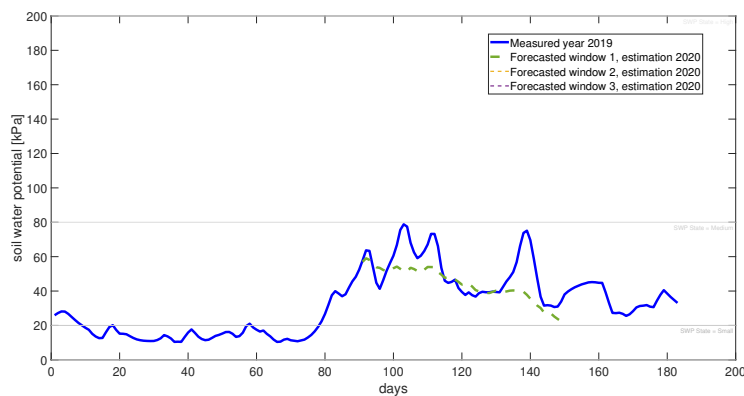
Un horizon supérieur à 60 jours impacte les prévisions, en particulier lorsque l'horizon correspond à toutes les données de l'année (exemple à la figure 4.22b). Si nous considérons à la figure 4.22b le jour 90, nous observons que le potentiel hydrique prédit (en pointillé vert) est inférieur à celui observé (en bleu). Pour les figures 4.22a et 4.22b le modèle prédictif a été initialisé sur 10 jours, ce qui correspond à la durée d'initialisation minimale pour le modèle ARIMAX(9,1,4) de l'érablière à hêtre. Lorsque nous effectuons une translation de la fenêtre de prédiction, pour démarrer celle-ci au jour 90 jusqu'au jour 150, en initialisant cette fois le modèle sur 30 jours, nous constatons que le potentiel hydrique prédit suit davantage l'évolution de celui qui est réellement observé en 2019.



(a) Horizon de 60 jours



(b) Horizon complet (mai-octobre 2019)



(c) Horizon de 60 jours, avec translation.

FIGURE 4.22 : Modèle ARIMAX pour l'érablière à hêtre (HB) entraîné sur les données de 2020 et prédictions des données de 2019. Différentes tailles d'horizon prédictif sont testés : 60 jours, 160 jours et 60 jours mais translatisés de 60 jours.

Nous avons confirmé ce résultat en montrant pour la sapinière que lorsque la taille de la fenêtre de prédiction était supérieure à 60 jours, l'exactitude des prédictions diminuait (figure 4.23) ainsi que le score-F1 (figure 4.24). Les tests ont été faits avec un modèle estimé sur les données de 2020 et les données prédites étaient celles de 2019.

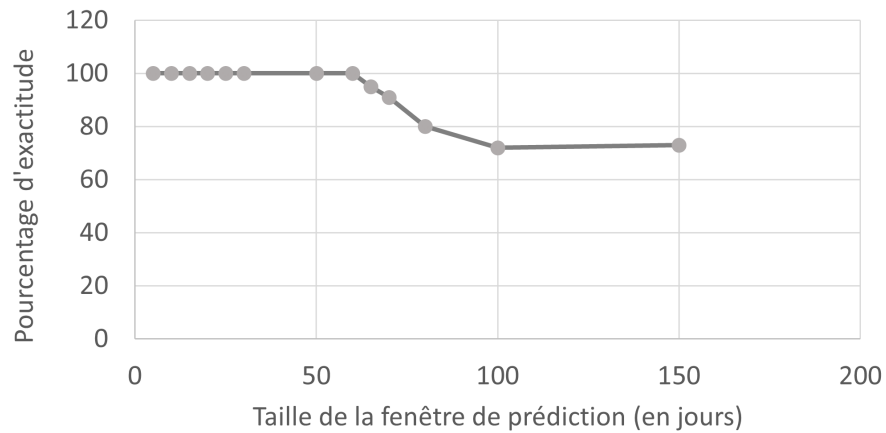


FIGURE 4.23 : Pourcentage d'exactitude des prédictions en fonction de la taille de la fenêtre de prédiction. Test réalisé sur la sapinière (prédiction 2019).

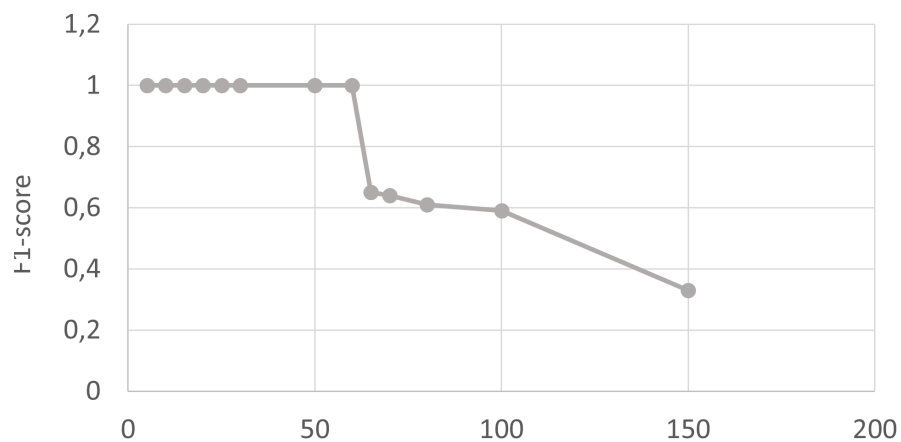


FIGURE 4.24 : Score-F1 en fonction de la taille de la fenêtre de prédiction (en jours). Test réalisé sur la sapinière (prédiction 2019).

Nous avons également évalué l'influence de la taille de la fenêtre de données qui sont utilisées pour initialiser le modèle prédictif. Nous rappelons que par défaut cette fenêtre doit être de taille minimale égale à la somme du coefficient autorégressif (p) et du coefficient de différenciation (d). Pour le peuplement MW, elle doit donc compter au minimum 15 points temporels.

En considérant trois tailles de fenêtre d'initialisation (15, 20 ou 30 jours), nous avons déterminé le pourcentage de prédictions exactes pour deux horizons différents de 2019, du jour 31 au jour 95 et du jour 51 au jour 115, respectivement représentés en gris et en orange sur la figure 4.25.

Nous constatons que le meilleur pourcentage de prédictions exactes est obtenu pour le premier horizon prédictif lorsque la fenêtre est de 15 jours. Le second horizon correspond à une période durant laquelle le potentiel hydrique du sol fluctue autour de valeurs élevées.

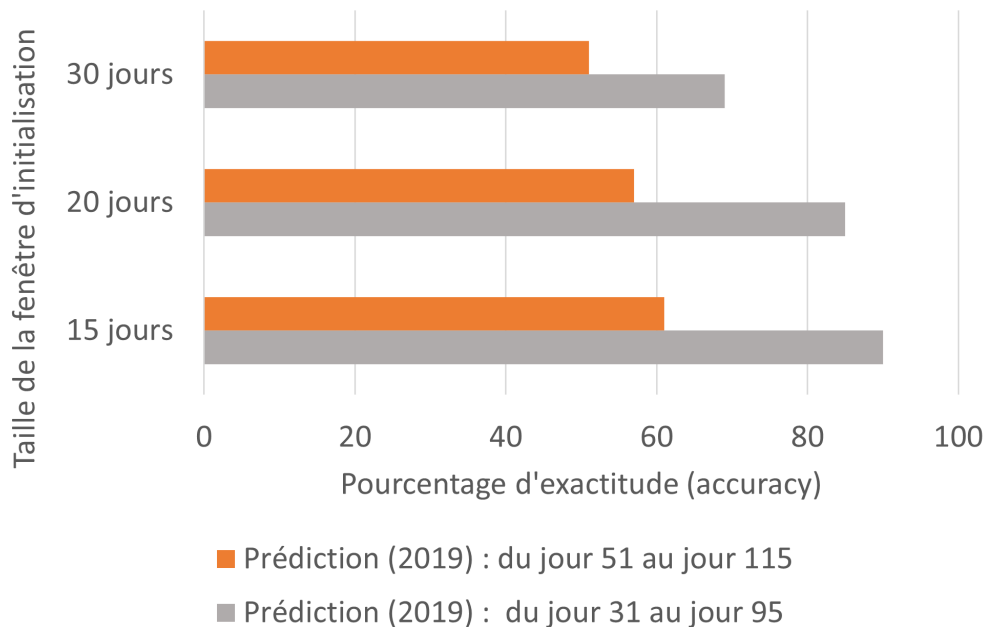
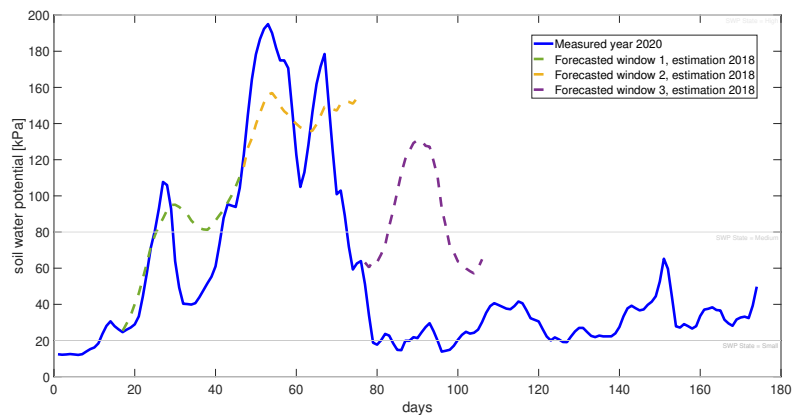
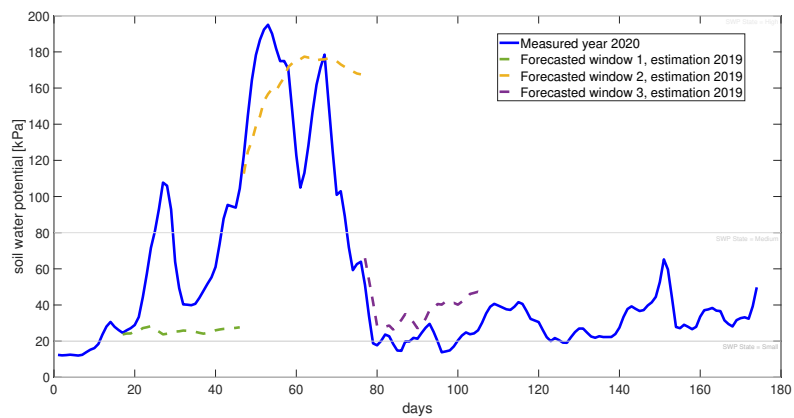


FIGURE 4.25 : Influence de la taille (en nombre de jours) de la fenêtre d'initialisation du modèle : prédictions réalisées pour 2019 pour la sapinière à érable (MW). Deux horizons prédictifs ont été testés : du jour 31 au jour 95 (en gris) et du jour 51 au jour 115 (en orange).

Ainsi, pour la suite des tests réalisés, la fenêtre d'initialisation du modèle prédictif a été fixée à 15 jours pour les peuplements MW et HW, et à 10 jours pour le peuplement HB. La taille des fenêtres de prédiction était égale à 30 jours. Nous présentons les résultats obtenus pour la sapinière. Nous avons prédit le potentiel hydrique du sol observé en 2020 en entraînant le modèle soit avec les données de 2018 (figure 4.26a) soit avec les données de 2019 (figure 4.26b).



(a) Estimation sur les données de 2018



(b) Estimation sur les données de 2019

FIGURE 4.26 : Modèle ARIMAX pour la sapinière à érable entraîné sur les données de 2018 ou 2019 et prédiction du potentiel hydrique mesuré en 2020 (référence en bleu).

Les pourcentages de prédictions exactes, calculés sur la classification des états du potentiel hydrique du sol, sont regroupés dans le tableau 4.2. Ce que nous mettons

TABLEAU 4.2 : Pourcentage de prédictions exactes de l'état du potentiel hydrique du sol, pour la sapinière à érable, année 2020.

fenêtre de prédiction	estimation en 2018	estimation en 2019
fenêtre 1	57	67
fenêtre 2	87	87
fenêtre 3	37	60

en évidence sur la figure 4.26 est que dans le premier cas, le potentiel hydrique prédit correspond relativement bien au potentiel hydrique observé pour les deux premières fenêtres de prédiction. Cependant, pour la troisième fenêtre, le potentiel hydrique prédit correspond à un état *High* alors qu'il est observé autour de 20 kPa. En effet, le pourcentage de prédictions exactes pour cette fenêtre est faible (37%). Nous constatons un résultat opposé lorsque le modèle est estimé en utilisant les données de 2019. En effet, dans ce cas le potentiel hydrique prédit à la première fenêtre de prédiction est plus faible comparativement aux valeurs numériques observées à la même période. Néanmoins, lorsque nous regardons la prédiction sur les états du potentiel hydrique (tableau 4.2), le pourcentage d'exactitude est plus faible en 2018 qu'en 2019, car la diminution du potentiel à la fin de la fenêtre 1 n'est pas prédite correctement sur la figure 4.26a. Le potentiel prédit reste au-dessus du seuil de 80 kPa alors qu'il a en réalité chuté en quelques jours pour atteindre une valeur de 40 kPa. La diminution du potentiel est prédite, mais avec une amplitude plus faible.

Les résultats obtenus pour la seconde fenêtre prédisent un potentiel élevé, comme observé (figure 4.26b). Contrairement au cas (a) pour lequel le modèle a été estimé avec les données de 2018, le résultat obtenu pour la troisième fenêtre en (b) concorde relativement bien avec le potentiel hydrique observé. Or, le pourcentage

d'exactitude n'est que de 67 % car dans cette fenêtre le potentiel hydrique fluctue à la limite de la zone de l'état *Small* et de l'état *Medium*. Cette fluctuation entraîne une mauvaise classification des états prédits qui pourrait être améliorée en changeant la valeur des seuils.

Ainsi, nous pouvons conclure que l'influence des données d'entraînement du modèle ARIMAX est locale. Nous émettons l'hypothèse ici que les variables exogènes *ressenties* durant la période d'initialisation du modèle impacte la performance de la prédiction. Ces résultats nous indiquent qu'il est difficile de sélectionner une année spécifique pour l'entraînement des modèles, et qu'il faudrait plutôt privilégier une sélection locale qui s'appuie sur des critères de similarité des variables exogènes.

Dans une première approche, nous avons calculé à titre indicatif la corrélation croisée entre les variables exogènes, d'une année à l'autre, sur la même période qui a servi à initialiser le modèle ARIMAX (tableau 4.3).

TABLEAU 4.3 : Corrélations croisées pour les variables exogènes suivantes : température de l'air, température du sol et déficit de pression de vapeur.

Années considérées	T _{air}	T _{sol}	VPD
2020 - 2019	12.99	3.94	-0.28
2020 -2018	-2.93	1.60	0.09

Le VPD et les températures de l'air et du sol ont été considérées entre le jour 61 et le jour 76, pour 2018, 2019 et 2020. Pendant cette période, nous observons que pour deux des trois variables la plus forte corrélation croisée avec 2020 a lieu en 2019. Cette approche qui utilise un critère de similarité qui s'appuie sur la corrélation croisée au cours du temps pourrait s'avérer pertinente pour sélectionner la meilleure année pour l'entraînement du modèle ARIMAX, pour une fenêtre de prédiction précise. En évaluant quelles données météorologiques passées se rapprochent le plus des données présentes, sur une période de 15 jours, il serait

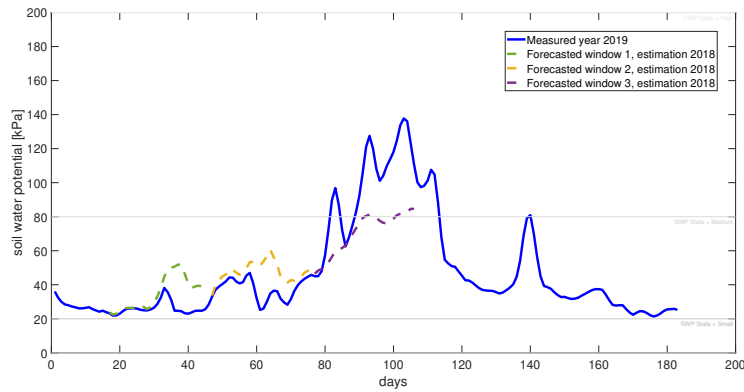
alors possible de sélectionner l'année d'entraînement du modèle.

Un autre exemple est proposé pour illustrer l'impact du choix des données d'entraînement du modèle ARIMAX (figure 4.27). Le même peuplement est considéré, mais le potentiel hydrique est prédit sur trois fenêtres de l'année 2019. Le modèle a été entraîné avec des données de 2018 (figure 4.27a) puis avec les données de 2020 (figure 4.27b).

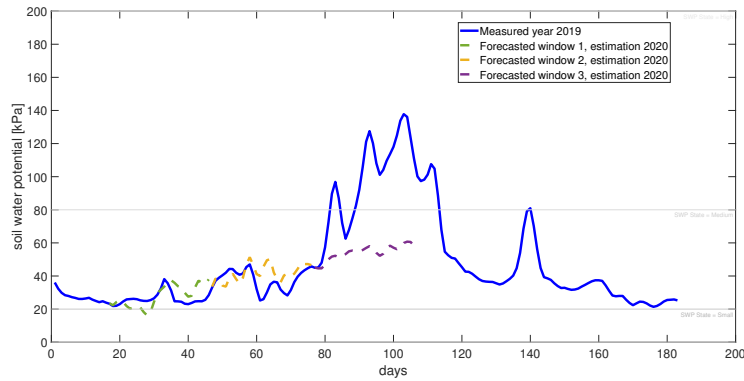
Les pourcentages de prédictions exactes présentés au tableau 4.4 sont élevés pour les fenêtres 1 et 2, dans les deux situations. L'estimation du modèle en utilisant les données du potentiel hydrique du sol mesurées en 2018 permet d'obtenir des meilleures prédictions sur la fenêtre 3. La représentation graphique des résidus souligne la plus grande différence entre les valeurs observées et les valeurs ajustées pour cette fenêtre de prédiction avec le modèle estimé en 2020 (figure 4.27c).

Pour ce même cas, il est possible d'estimer la corrélation entre les valeurs observées et les valeurs prédites par le modèle en déterminant l'ajustement linéaire qui les relie. Dans cet exemple, le coefficient de corrélation noté R^2 , est égal à 0.76 pour la fenêtre 3 (figure 4.28).

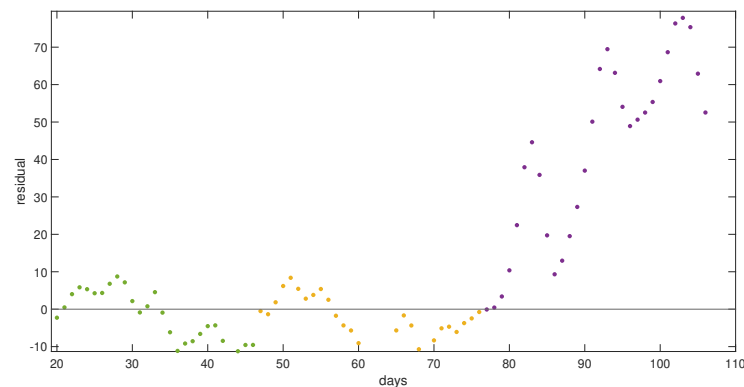
Cette troisième fenêtre a pour spécificité de cibler la période au cours de laquelle le potentiel hydrique du sol connaît en 2019 sa plus grande augmentation (entre le jour 80 et le jour 120). C'est la période estivale. L'objectif est de pouvoir initialiser le modèle ARIMAX alors que le potentiel hydrique est déjà élevé. En augmentant la taille des trois fenêtres de prédiction, de 30 jours à 40 jours, la troisième fenêtre débutera ainsi plus tard. La figure 4.29 met en évidence une meilleure prédiction du potentiel hydrique en comparaison avec la figure 4.27b. Le coefficient de corrélation, tel que défini précédemment, est dans ce cas égal à 0.82.



(a) Estimation sur les données de 2018, 3 fenêtres de prédiction de 30 jours



(b) Estimation sur les données de 2020, 3 fenêtres de prédiction de 30 jours



(c) Résidus (estimation du modèle sur les données de 2020)

FIGURE 4.27 : Modèle ARIMAX pour la sapinière à érable entraîné sur les données de 2018 ou 2020 et prédiction du potentiel hydrique de 2019. Les résidus obtenus pour le cas (b) sont représentés avec le même code couleur.

TABLEAU 4.4 : Pourcentage de prédictions exactes de l'état du potentiel hydrique du sol, pour la sapinière à érable (année 2019).

fenêtre de prédiction	estimation en 2018	estimation en 2020
fenêtre 1	100	90
fenêtre 2	100	100
fenêtre 3	60	30

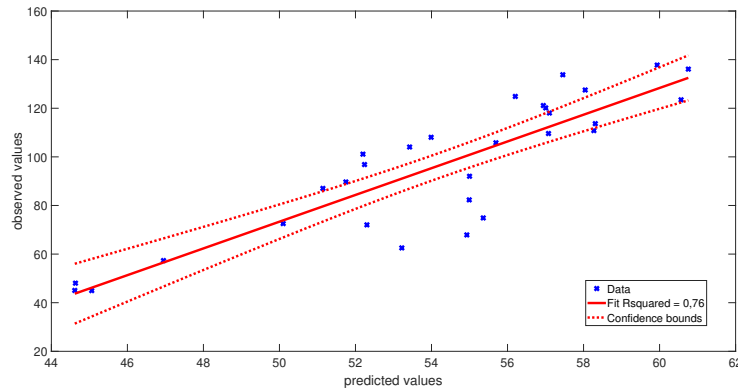


FIGURE 4.28 : Ajustement linéaire des données observées (2019) en fonction des données prédites par le modèle ARIMAX(14,1,4) de la sapinière à érable (troisième fenêtre de prédiction, modèle estimé en 2020).

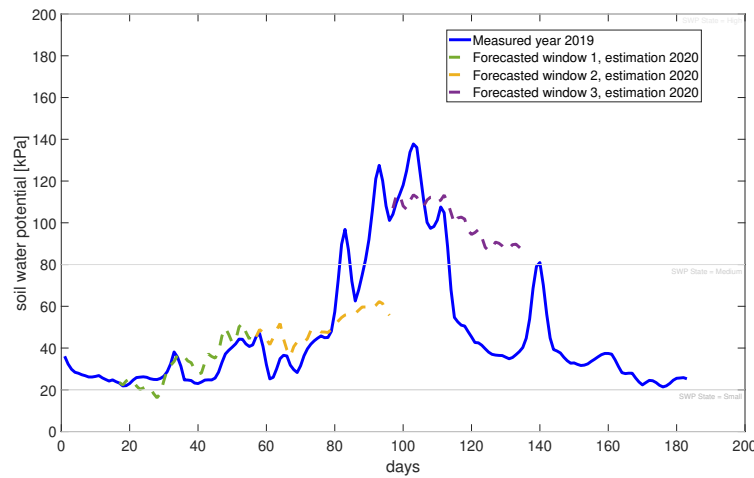


FIGURE 4.29 : Modèle ARIMAX pour la sapinière entraîné sur les données de 2020 et prédiction du potentiel hydrique de 2019 (fenêtres de 40 jours).

Nous l'avons souligné, l'année 2020 a connu trois périodes de sécheresses-flash pendant lesquelles le potentiel hydrique a atteint des valeurs élevées. Les tests que nous avons réalisés mettent en avant la moins bonne performance de la prédiction du potentiel hydrique en 2020. Pour tenter d'y remédier nous pouvons avancer deux solutions : tester un nouveau modèle ou changer la stratégie d'entraînement du modèle ARIMAX développé. Nous avons exploré les deux.

4.1.6 Phase 6 : exploration d'un autre type de modèle

Explorer un autre type de modèle permet de dépasser les contraintes imposées par les données elles-mêmes - nous évoquons la spécificité de l'année 2020 - mais cela permet également de dépasser les limites de l'approche ARIMAX. Ces limites font écho à celles évoquées dans l'analyse exploitant les chaînes de Markov à temps discret : l'effet de mémoire est limité à 15 jours pour la sapinière et l'érablière à bouleau et à 10 jours pour l'érablière à hêtre. De plus, il est nécessaire d'initialiser chaque modèle du même nombre de jours, au minimum. Nous aurions pu augmenter le degré p de la partie autorégressive, mais nos tests n'ont pas montré d'amélioration nette des prédictions. Augmenter la complexité du modèle ne s'est pas avéré être une option intéressante et nous courions le risque d'un sur-ajustement (*over-fitting*) des données. Un compromis a dû être fait.

Une approche différente de type NARX (*non linear autoregressive model with exogenous variables*) qui permet d'inclure également des variables exogènes a été évaluée à titre indicatif (figure 4.30). Néanmoins, la fenêtre temporelle de prédiction ne s'est pas élargie, bien que pour une fenêtre de 20 jours, la prédiction reproduit bien les valeurs observées (figure 4.31). Cela se traduit sur la figure par de faibles valeurs résiduelles.

Ainsi, une perspective prometteuse serait de tester une approche *Long-Short Term Memory* (LSTM) pour s'affranchir de la taille de l'horizon prédictif, en exploitant

une mémoire plus longue des données passées, mais aussi en bénéficiant d'une approche multi-échelles à travers les nombreuses couches de neurones qui peuvent être intégrées.

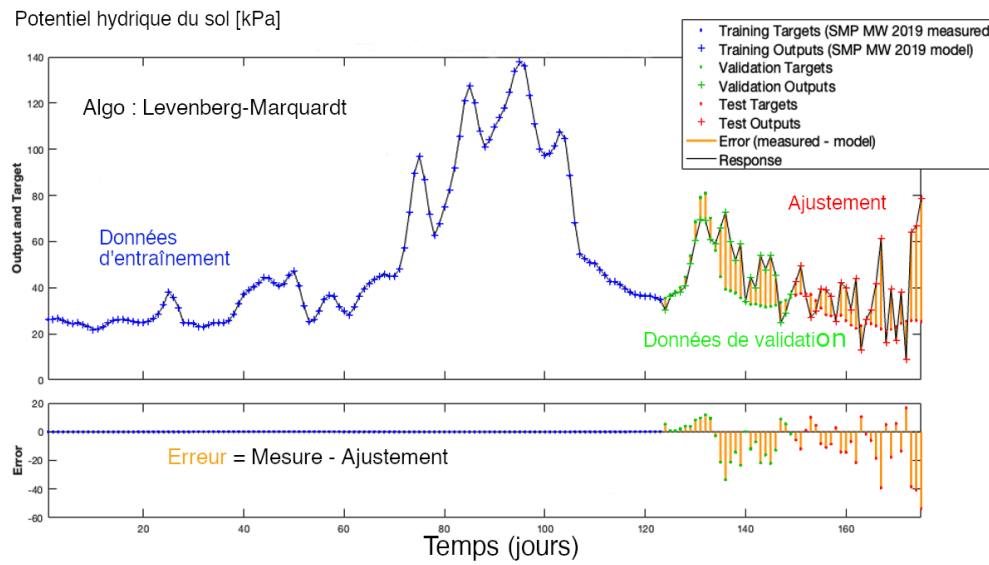


FIGURE 4.30 : Résultat obtenu pour le modèle NARX, tel que décrit dans la partie 3.

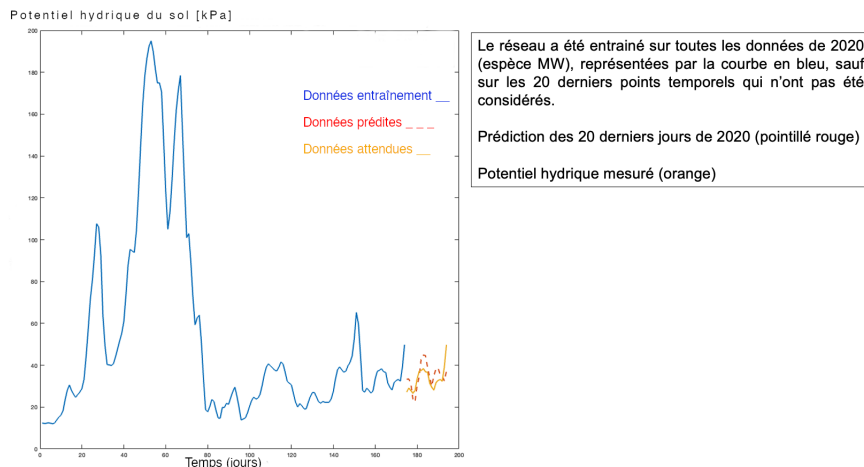


FIGURE 4.31 : Modèle NARX entraîné sur toutes les données de potentiel hydrique mesurées en 2020 (sapinière à érable), sauf sur les 20 derniers points temporels qui ont servi à la validation.

4.2 Application : étude de périodes de sécheresse-flash entre 2017 et 2020

Nous avons défini dans le chapitre 1 le concept de sécheresse-flash (*flash drought*). Une fonction a été implémentée sous Matlab pour pouvoir extraire les données relatives à une période de sécheresse-flash, de k jours, k pouvant être précisé comme paramètre d'entrée de la fonction.

De plus, grâce à cette fonction il est possible d'extraire les données qui précèdent et suivent, d'un certain nombre de jours, chaque période de sécheresse-flash identifiée. Ainsi, nous pouvons obtenir le nombre de périodes de sécheresse-flash de huit jours et plus, pour chaque année considérée. Certaines années ont eu des périodes plus courtes de 6 ou 7 jours. Précisons à nouveau que ce critère de huit jours a été choisi par rapport à l'étude de Fu *et al.* (2022) dans laquelle les auteurs ont exploité les résultats obtenus par McColl *et al.* (2017) et Akbar *et al.* (2018). Le critère de sélection privilégié repose sur l'absence de précipitations durant un certain nombre de jours consécutifs associée à une augmentation du potentiel hydrique, i.e. qu'il devient plus négatif. Ce sont des périodes d'assèchement (*dry-down*).

Nous avons mis en évidence les périodes d'assèchement suivantes :

- Pour 2017 : 2 périodes de 8 jours et plus (du 17 au 26 septembre et du 16 au 23 octobre) ;
- Pour 2018 : 1 période de 8 jours (du 13 au 20 septembre) ;
- Pour 2019 : 1 période de 6 jours (du 21 au 27 août) et 1 période de 7 jours (du 15 au 21 septembre) ;
- Pour 2020 : 3 périodes de 8 jours et plus (du 17 au 25 mai, du 12 au 22 juin, du 18 au 25 septembre).

4.2.1 Potentiel hydrique et température du sol

À partir des données extraites, nous pouvons visualiser l'évolution du potentiel hydrique et de la température du sol, pour chaque peuplement. Par exemple, pour les trois périodes de sécheresse-flash qui ont été mises en évidence en 2020 : en mai (figure 4.32), en juin (figure 4.33) et en septembre (figure 4.34).

Pour les deux premières périodes de sécheresses, nous observons les deux mêmes phénomènes : le potentiel hydrique du sol augmente sur toute la durée des périodes extraites, et la température du sol (moyennée sur une journée) de l'érablière à hêtre reste inférieure à celle mesurée pour la sapinière à érable et l'érablière à bouleau.

Pour la sapinière à érable, le potentiel hydrique atteint des valeurs supérieures à 100 kPa en mai et proche de 200 kPa en juin. Le potentiel hydrique de l'érablière à bouleau se rapproche aussi de ces valeurs. Pour l'érablière à hêtre, le potentiel hydrique augmente également mais avec une pente plus faible, atteignant pour chacune des deux périodes des valeurs maximales de respectivement 40 kPa et 140 kPa. Le 26 mai 2020 des précipitations sont observées (figure 4.32a), mais elles ne semblent pas faire diminuer le potentiel hydrique qui amorce sa descente après le 28 et 29 mai, jours où les précipitations furent plus abondantes. Après la première sécheresse-flash, la sapinière à érable retrouve une valeur de potentiel hydrique supérieure à celle mesurée six jours avant l'événement, dans un état que nous avons caractérisé de *moyen*.

Avant, pendant et après ces deux premières périodes d'assèchement, la température du sol de l'érablière à bouleau reste supérieure à celle des deux autres peuplements. Au mois de mai, la sapinière à érable a sa température du sol qui est très proche de celle mesurée pour l'érablière à hêtre. En revanche, au mois de juin elle suit davantage celle de l'érablière à bouleau, creusant un écart plus important avec la température du sol de l'érablière à hêtre qui demeure la plus basse entre le 12 et

le 22 juin, en comparaison avec les peuplements MW et HW.

En septembre, la température du sol est très similaire pour les trois peuplements bien que le potentiel hydrique du sol de la sapinière augmente davantage ; le 18 septembre il est de 32 kPa tandis que celui des deux érablières est proche de 20 kPa. La sapinière a donc déjà son potentiel hydrique dans un état *moyen* au début de la troisième période de sécheresse-flash, et celui-ci a une augmentation plus prononcée après sept jours d'assèchement que celle observée pour les érablières. À la fin de la période, la température du sol de l'érablière à hêtre est légèrement inférieure à celles de la sapinière et de l'érablière à bouleau, se maintenant également plus basse pendant les 6 jours suivants.

Cette capacité de l'érablière à hêtre à maintenir une température de son sol inférieure à celle mesurée pour les deux autres peuplements étudiés se confirme pendant les sécheresses-flash de 2019 (figure 4.35a) et de 2018 (figure 4.35b).

Notons qu'en juin et en septembre 2020, au début de l'assèchement alors que le potentiel hydrique commence à augmenter sous l'effet des stress hydriques, la température du sol des trois peuplements diminue pendant environ trois jours, avant d'augmenter. Au mois de mai, au début du printemps, nous n'observons pas cette tendance.

Dans l'analyse des corrélations-croisées, nous avons mis en évidence un décalage antérieur de quatre jours entre le maximum de la température du sol et le maximum du potentiel hydrique, signe que la température *dirige* le potentiel hydrique⁴. Cette analyse considérait toutes les données mesurées entre mai et octobre 2020. Nous avons calculé les mêmes coefficients de corrélation-croisée mais en considérant les données de la deuxième période de sécheresse-flash (en juin 2020), en prenant aussi les données six jours avant l'assèchement et six jours après. Les coefficients ont été normalisés pour pouvoir comparer les deux analyses (figure 4.36 pour l'érablière à

4. Stat 510 : Applied time series analysis with R, PennState Eberly College of Science
<https://online.stat.psu.edu/statprogram/stat510>

hêtre). Nous constatons que la température du sol ne présente plus le même effet d'antériorité sur le potentiel hydrique lorsque nous ciblons la sécheresse-flash. En effet, les coefficients de corrélation-croisée avant le temps de référence (*lag-0*) ne sont plus aussi élevés que les coefficients représentés à la figure 4.36a. Les dépendances temporelles entre les variables évoluent localement. Nous pouvons faire la même conclusion en analysant de façon similaire les coefficients de corrélation-croisée entre le potentiel hydrique et la température du sol de la sapinière et de l'érablière à bouleau⁵.

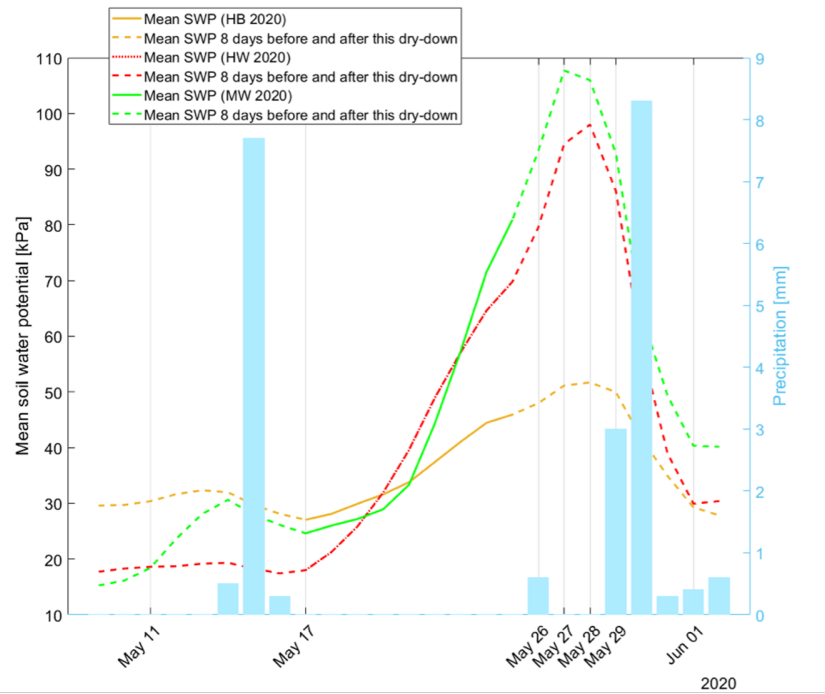
Nous avons donc montré que pendant les périodes de sécheresse-flash, l'érablière à hêtre maintient la température de son sol inférieure à celle mesurée pour la sapinière à érable et pour l'érablière à bouleau. Afin d'explorer davantage le comportement singulier de l'érablière à hêtre face à des stress hydriques, nous avons appliqué les méthodes de classification utilisant les arbres de décision comme nous l'avons fait pour sélectionner les variables prédictives du potentiel hydrique (phase 2 de notre méthodologie).

Nous pouvons ainsi comparer l'arbre entraîné avec les données du peuplement MW (figure 4.37) à celui entraîné avec les données du peuplement HB (figure 4.38), pendant la deuxième période de sécheresse-flash. Nous constatons que pour la sapinière à érable, le basculement vers un état *très élevé* du potentiel hydrique se fait pour une température du sol supérieure ou égale à 12.90 °C. Lorsque nous considérons l'érablière à hêtre, son potentiel hydrique atteint ce même état pour une température du sol supérieure ou égale à 15.94 °C. Pour l'érablière à bouleau, nous trouvons une température du sol supérieure ou égale à 15.10 °C. Précision que dans cette analyse, le seuil *très élevé* représentait une valeur de potentiel hydrique du sol supérieure à 120 kPa.

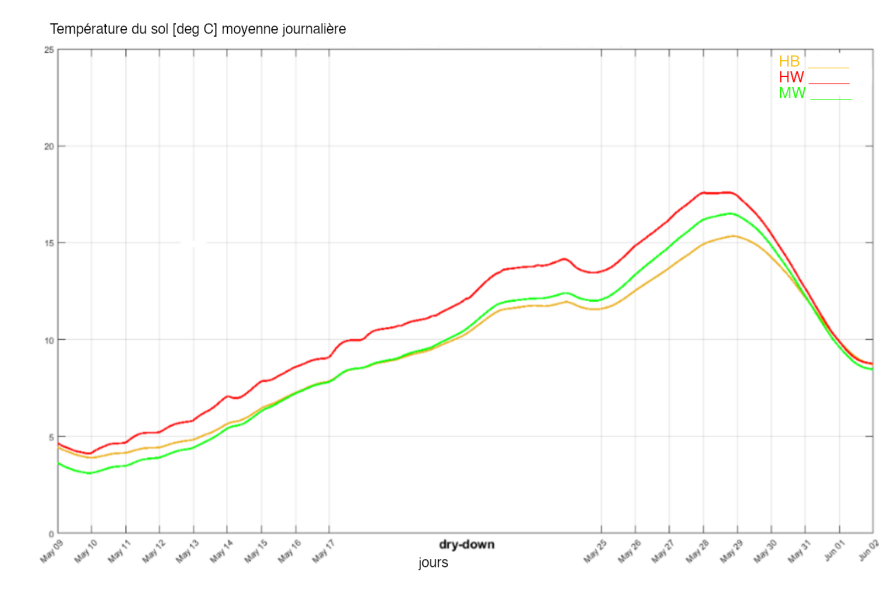
5. Tous les coefficients sont disponibles dans le rapport technique. Les trois périodes de sécheresse-flash de 2020, pour les trois peuplements, ont été analysées.

Dans un contexte de changements climatiques, où pour la région des Laurentides les modèles régionaux prévoient une augmentation de 6 °C sur l'horizon 2041-2070 par rapport aux moyennes de 1981-2010 (voir le chapitre 1), la capacité de l'érablière à hêtre à réguler son potentiel hydrique à travers une température de son sol qui est plus basse, est un atout.

Cette conclusion est une nouvelle hypothèse que nous proposons pour expliquer l'expansion du hêtre à grandes feuilles dans les forêts mixtes du Québec au détriment de l'érable à sucre, depuis les années 1970 (Direction de la recherche forestière, 2017).

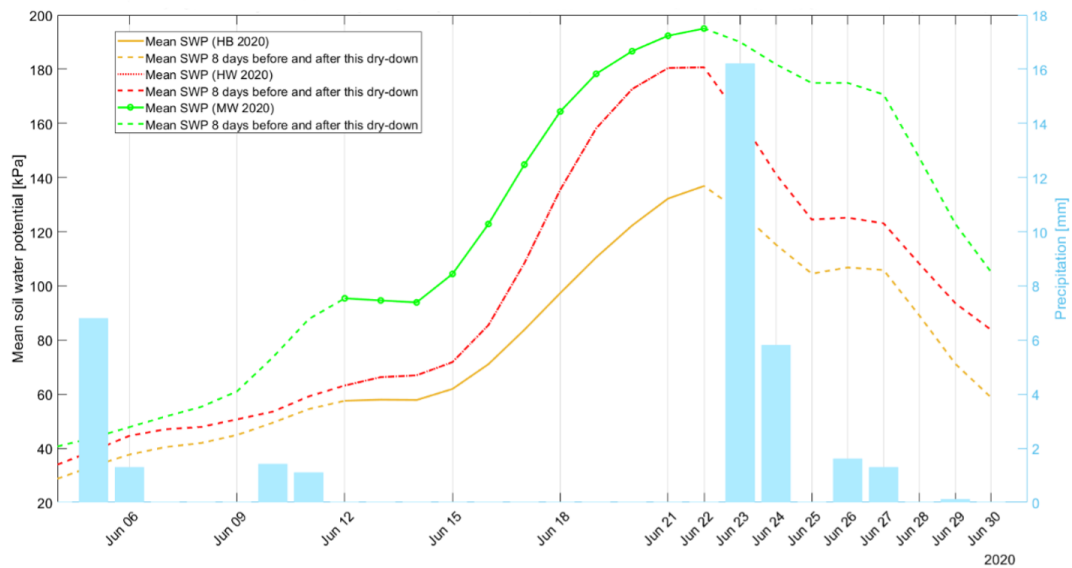


(a) Évolution du potentiel hydrique du sol (*mean soil water potential*) et précipitations

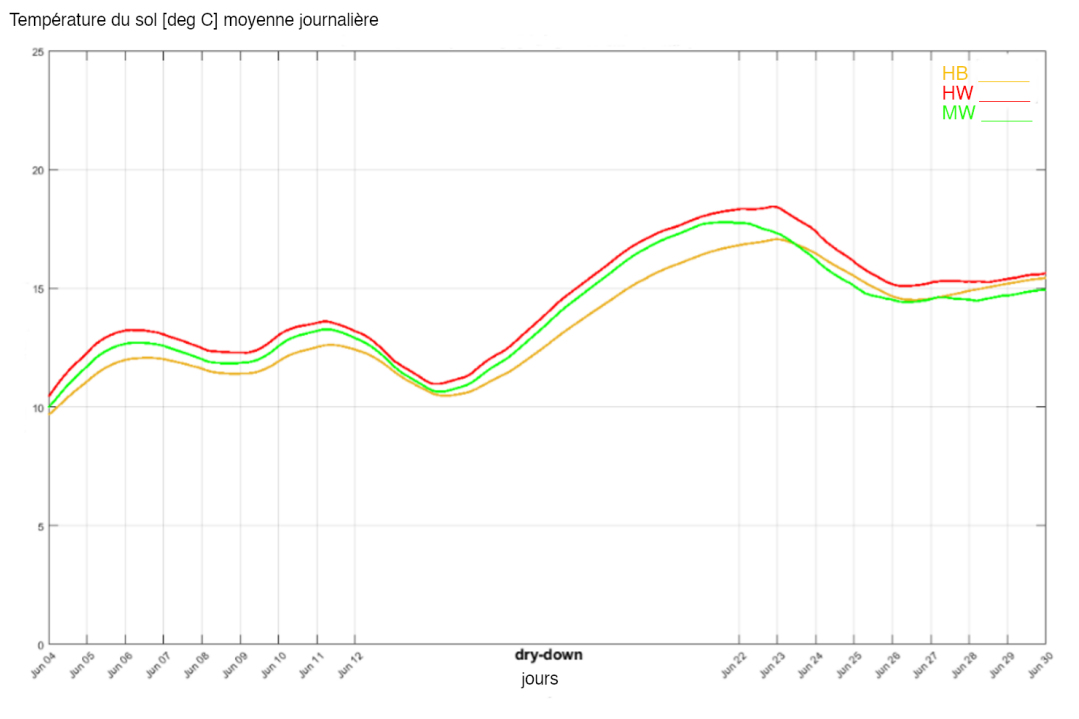


(b) Évolution de la température du sol

FIGURE 4.32 : Première sécheresse-flash de 2020 (mai) : potentiel hydrique et température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune).

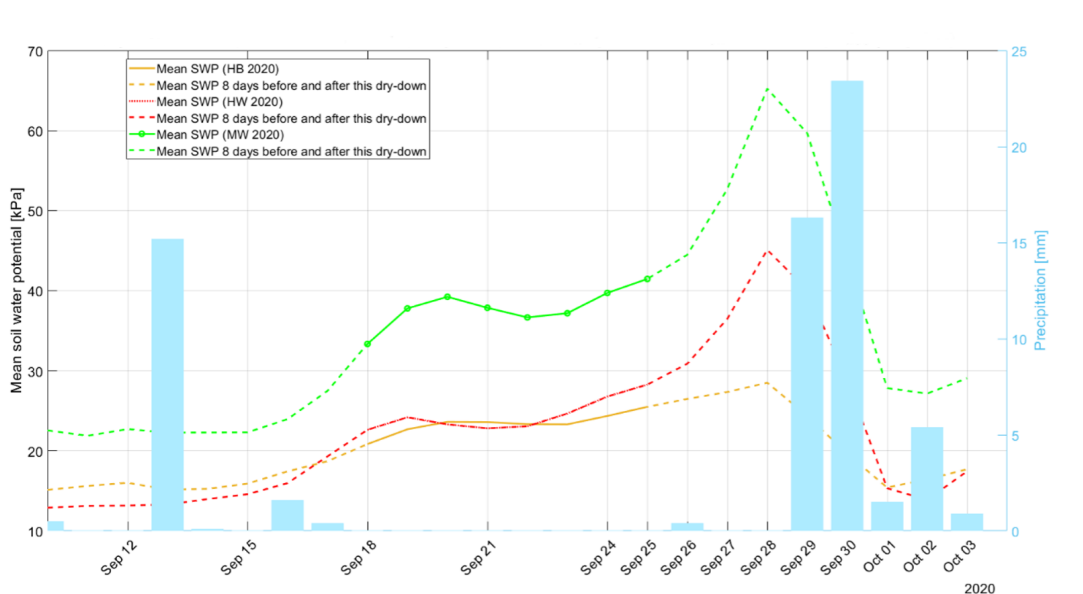


(a) Évolution du potentiel hydrique du sol (*mean soil water potential*) et précipitations

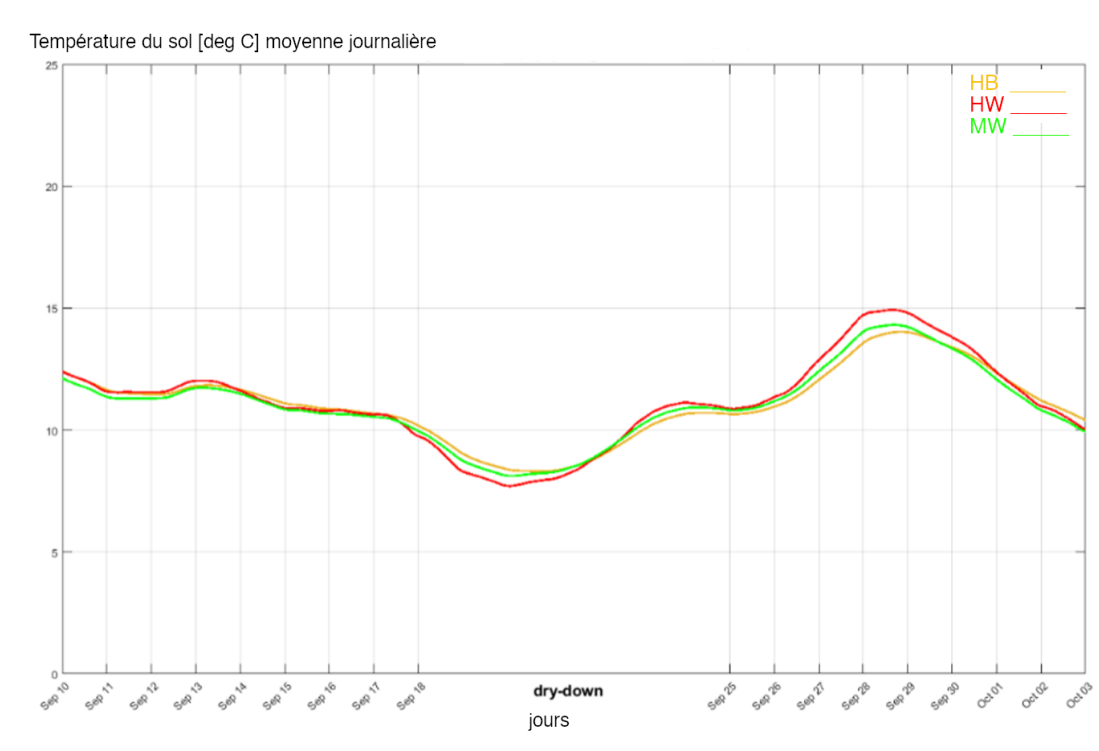


(b) Évolution de la température du sol

FIGURE 4.33 : Deuxième sécheresse-flash de 2020 (juin) : potentiel hydrique et température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune).

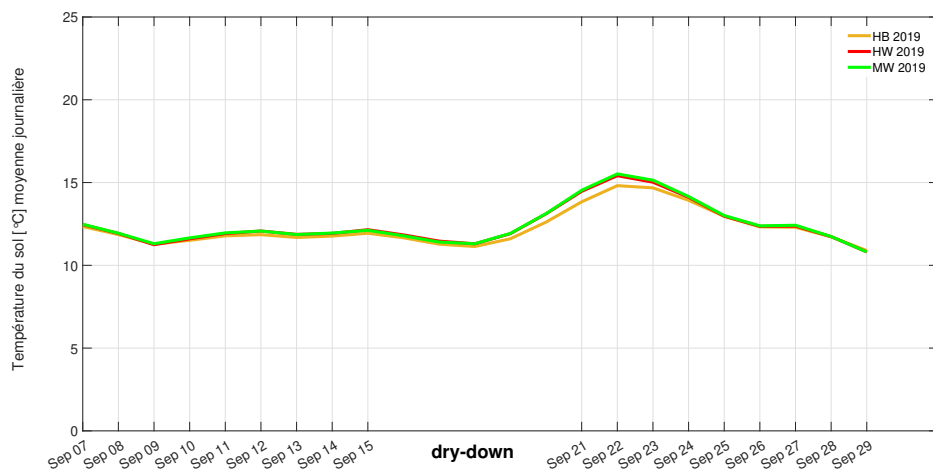


(a) Évolution du potentiel hydrique du sol (*mean soil water potential*) et précipitations

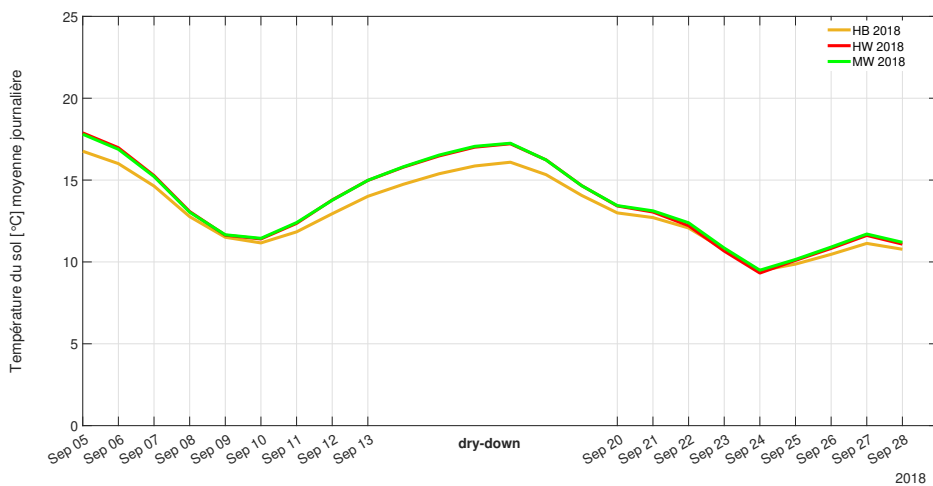


(b) Évolution de la température du sol

FIGURE 4.34 : Troisième sécheresse-flash de 2020 (septembre) : potentiel hydrique et température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune).

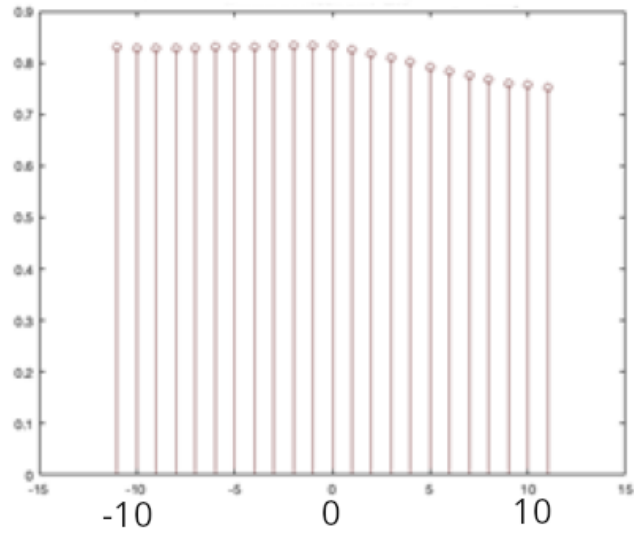


(a) Évolution de la température du sol (sécheresse-flash de 2019)

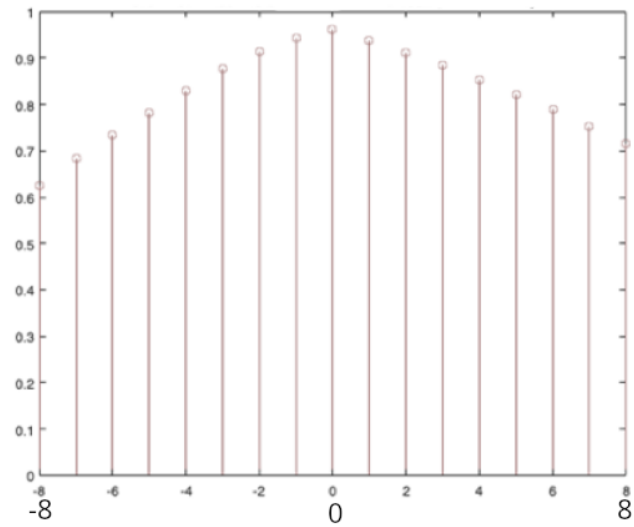


(b) Évolution de la température du sol (sécheresse-flash de 2018)

FIGURE 4.35 : Évolution de la température du sol pour la sapinière à érable (MW en vert), l'érablière à bouleau (HW en rouge) et l'érablière à hêtre (HB en jaune).



(a) De mai à octobre 2020



(b) Deuxième sécheresse-flash (juin 2020)

FIGURE 4.36 : Coefficients de corrélation-croisées (normalisés) entre le potentiel hydrique et la température du sol, calculés en utilisant (a) toutes les données de 2020 de l'érablière à hêtre et (b) les données extraites de la deuxième sécheresse-flash de 2020, pour HB. En abscisse sont représentés les décalages ou *lag* (en jours) et en ordonnées les coefficients de corrélation-croisée normalisés.

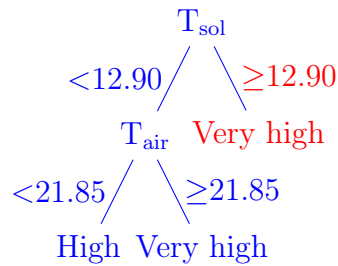


FIGURE 4.37 : Arbre de décision obtenu pour la sapinière à érable (MW), deuxième sécheresse de 2020. Les températures sont en [°C].

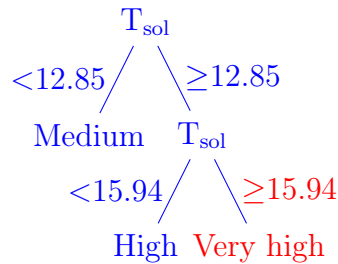


FIGURE 4.38 : Arbre de décision obtenu pour l'érablière à hêtre (HB), deuxième sécheresse de 2020. Les températures sont en [°C].

4.2.2 Nouvelle stratégie d'entraînement du modèle ARIMAX

Les caractéristiques des données mesurées durant les périodes de sécheresse-flash - en particulier un potentiel hydrique du sol élevé - ont nécessité d'explorer une stratégie différente pour entraîner le modèle ARIMAX.

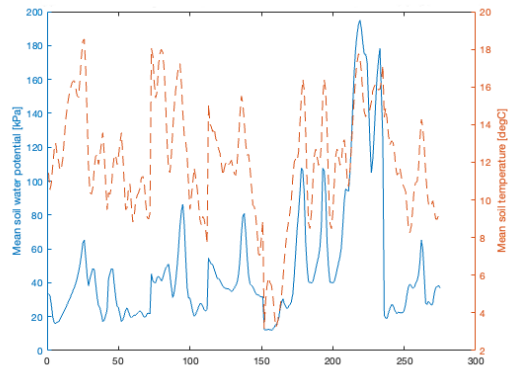
La courte durée de ces périodes ne permet pas de séparer les données afin d'avoir un jeu de données d'entraînement et un jeu de données de validation qui soient dissociés. Pour surmonter ce problème, nous avons concaténé toutes les données relatives aux sécheresses-flash en considérant 16 jours avant le début de chaque période et 16 jours après. La figure 3.9 illustre la stratégie qui a été mise en place.

Ainsi, pour les trois peuplements, nous avons créé une série temporelle fictive du potentiel hydrique en y associant la température du sol mesurée au même moment.

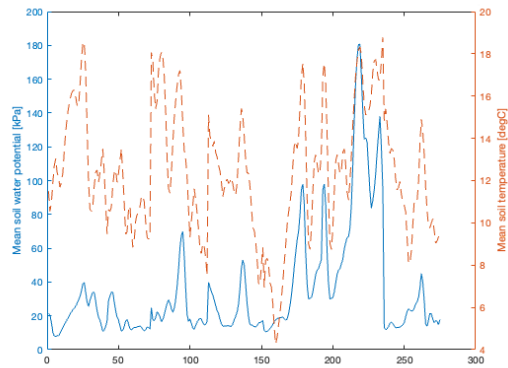
Les deux variables sont représentées pour chaque peuplement sur la figure 4.39. Nous avons également concaténé de la même façon les variables exogènes.

Nous avons alors utilisé les données concaténées de 2017 à 2019 pour estimer le modèle ARIMAX, et nous l'avons validé sur un horizon prédictif correspondant aux données des trois sécheresses-flash de 2020. Nous pouvons visualiser le résultat sur la figure 4.40. Pour la sapinière à érable, avec un modèle ARIMAX(14,1,4) entraîné sur les données des sécheresses-flash qui ont eu lieu entre 2017 et 2019, nous constatons une amélioration de la prédiction du potentiel hydrique du sol en 2020 en comparaison avec la figure 4.26. Si nous reprenons les résultats obtenus lorsque le potentiel hydrique du sol de 2020 était prédit après un entraînement du modèle avec les données de 2018 (figure 4.26a) et avec les données de 2019 (figure 4.26b), nous notons que les valeurs numériques élevées du potentiel hydrique n'étaient pas exactement prédites, même si l'état *élevé* de celui-ci était reproduit. Ce modèle ARIMAX a un coefficient autorégressif p égal à 14 et un coefficient de moyenne mobile q égal à 4. Nous avons optimisé ce modèle sur toutes les données de 2020, pour chaque peuplement. Si nous changeons les deux coefficients en diminuant leur valeur, nous pouvons tester un modèle ARIMAX(6,1,1) qui va considérer une autocorrélation plus courte entre les points de la série temporelle fictive du potentiel hydrique du sol. Nous observons dans ce cas que la prédiction de la troisième fenêtre s'est améliorée (figure 4.41). En effet, l'augmentation du potentiel hydrique est mieux prédite en comparaison avec la figure 4.40.

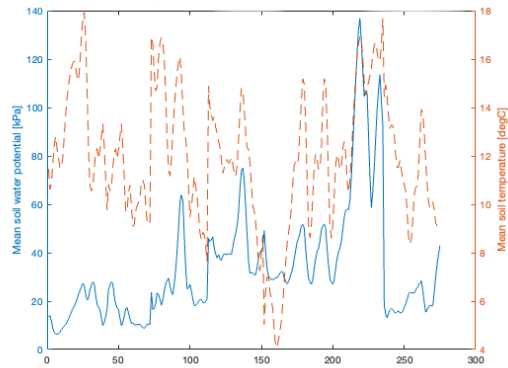
La spécificité des périodes d'assèchement implique donc d'adapter la stratégie d'entraînement du modèle ARIMAX, mais nécessite également une nouvelle optimisation des coefficients autorégressif et de moyenne mobile. Dans un contexte d'événements extrêmes comme les sécheresses-flash, les modèles prédictifs de type ARIMAX doivent ainsi être optimisés à une échelle plus petite, sur des données qui correspondent plus précisément à des stress hydriques.



(a) Potentiel et température du sol (MW)



(b) Potentiel et température du sol (HW)



(c) Potentiel et température du sol (HB)

FIGURE 4.39 : Concaténation des données du potentiel hydrique (en bleu) et de la température du sol (en pointillé orange) durant les sécheresses-flash de 2017 à 2020, en considérant 16 jours avant et après chaque période. L'axe des abscisses représente le temps fictif en jours.

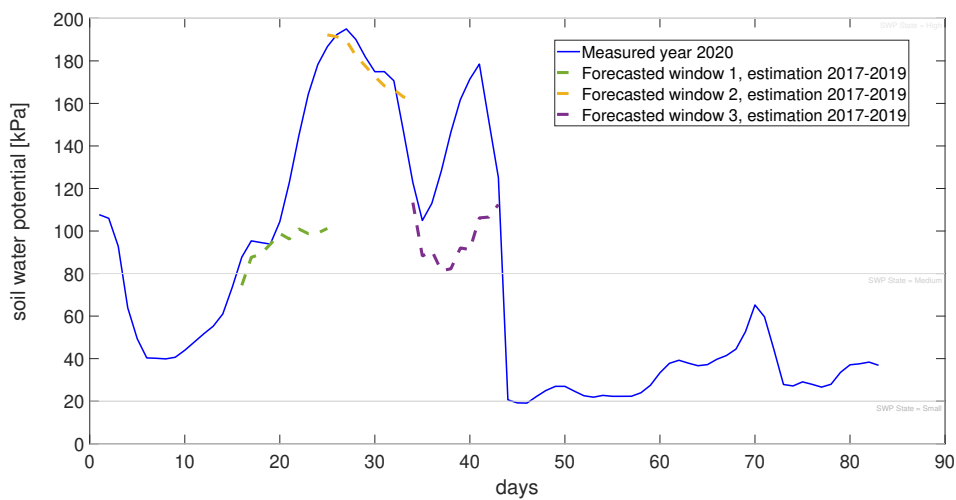


FIGURE 4.40 : Modèle ARIMAX(14,1,4) pour la sapinière entraîné sur les données concaténées des sécheresses-flash de 2017 à 2019 et prédiction du potentiel hydrique de 2020.

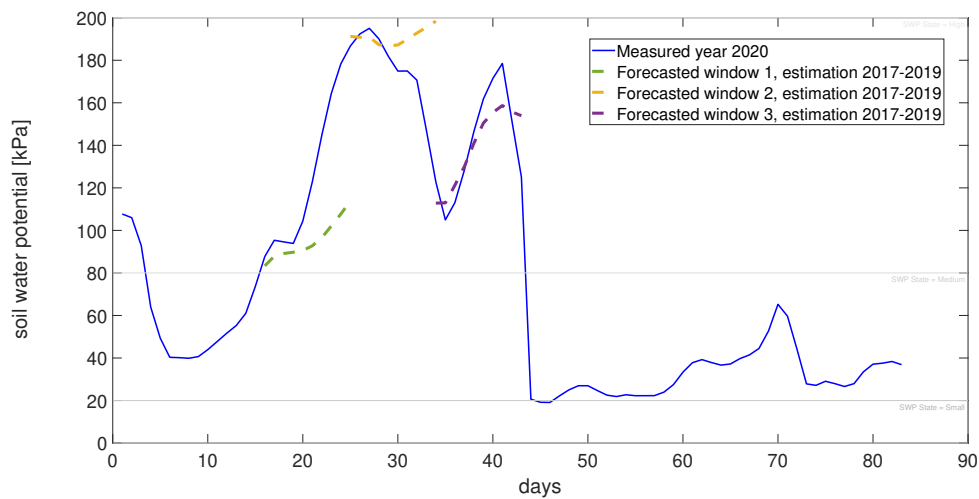


FIGURE 4.41 : Nouveau modèle ARIMAX(6,1,1) pour la sapinière entraîné sur les données concaténées des sécheresses-flash de 2017 à 2019 et prédiction du potentiel hydrique de 2020.

4.3 Ce qu'il faut retenir

Notre démarche méthodologique en sept phases a permis de répondre aux questions de recherche définies au chapitre 3.

Nous pouvons résumer en dix points ce qui a été mis en évidence dans le cadre de notre projet, en considérant les aspects écologiques (points 1 à 5) et les aspects liés au traitement des données (points 6 à 10).

(1) L'érablière à bouleau, qui est le peuplement dominant à la SBL, semble être la norme d'un point de vue écologique par rapport à la disponibilité en eau, avec un potentiel hydrique du sol qui se situe en moyenne entre celui de la sapinière à érable et celui de l'érablière à hêtre.

(2) La sapinière et l'érablière à hêtre représentent quant à eux respectivement les maxima et les minima du potentiel hydrique observés à la SBL entre 2017 et 2020.

(3) L'érablière à hêtre montre un comportement spécifique durant des stress hydriques vécus pendant des périodes de sécheresse-flash. Un effet tampon se met en place, lui permettant de maintenir une température du sol plus basse et un potentiel hydrique moins élevé.

(4) Durant ces périodes d'assèchement, le passage vers un potentiel hydrique très élevé, supérieur à 140 kPa, se fait pour l'érablière à hêtre à une température du sol de 15.90 °C, en comparaison avec la sapinière qui bascule à une température plus basse de 12.90 °C. Dans un contexte de changements climatiques, cet écart de 3 °C pourrait être favorable au hêtre, mais problématique pour le sapin qui se retrouve au sud de sa répartition.

(5) La sapinière à érable montre des signes de vulnérabilité, et semble se rapprocher davantage au fil des années d'un changement de régime de son potentiel hydrique du sol. Cette tendance s'est accentuée entre 2017 et 2020.

(6) Le nettoyage des données constitue une première phase méthodologique très importante. Elle permet la régularisation des données, étape nécessaire à l'analyse des séries temporelles. La synchronisation des mesures de potentiel hydrique et de température du sol avec les données météorologiques est également à ne pas négliger pour s'assurer d'une cohérence entre les données provenant de sources différentes.

Pour cette phase méthodologique, la haute résolution temporelle des données brutes (fréquence d'échantillonnage de 15 minutes ou de 30 minutes) est à privilégier pour les expérimentations futures.

Soulignons que la mise en évidence de la distribution non-normale des données collectées à la SBL peut impliquer des limites dans l'utilisation de certaines lois de probabilité pour les modéliser.

(7) L'ajustement du potentiel hydrique à la température du sol est recommandée.

(8) Un modèle ARIMAX permet une prédiction du potentiel hydrique du sol avec un pourcentage d'exactitude élevé (de 80 % à 100 %) sur un horizon prédictif de 60 jours. Pour l'étude de données plus extrêmes, il est recommandé d'entraîner le modèle sur des données qui ciblent des périodes caractéristiques.

(9) La sensibilité de ce type de modèle autorégressif à moyenne mobile semble être locale rendant difficile la sélection d'une année spécifique pour son entraînement, en particulier dans notre étude où les données annuelles ne couvrent que les mois de mai à octobre. Un indice de similarité pourrait améliorer la phase d'entraînement.

(10) Des ruptures dans les prédictions du potentiel hydrique du sol pourraient être dues à des fluctuations locales de la stationnarité des données, même après une différenciation de degré un. Il faudrait s'assurer d'une stationnarité par morceaux pour limiter ces ruptures.

4.4 Discussion et perspectives

Nous allons apporter un éclairage succinct sur les points (1) à (5) qui seront approfondis sous l'angle écologique dans une publication ultérieure. Notre objectif est de mettre ici davantage l'accent sur les aspects relatifs à la science des données.

Le point (1) rejoint ce qui a été mis en évidence dans l'étude de Bélanger *et al.* (2021) sur la respiration du sol où les résultats ont été inférés par rapport à ces deux peuplements - la sapinière à érable et l'érablière à hêtre - mais également sur le rôle de la luminosité sur l'expansion du hêtre dans les érablières (Collin *et al.*, 2017b).

De nombreux arguments peuvent être avancés au regard de la littérature pour expliquer le point (3). Est-ce que l'effet tampon observé pour l'érablière à hêtre est un phénomène purement abiotique lié au microclimat créé par la canopée et la litière du sol (Collin *et al.*, 2017b) ou un phénomène purement biotique influencé par la fermeture des stomates qui est régulée par la sécrétion d'acide abscisique (ABA), une phyto-hormone liée au stress (Daszkowska-Golec et Szarejko, 2013) ? Ou une combinaison des deux ?

Il faut noter que pendant les sécheresses-flash nous observons que ce pouvoir tampon du hêtre s'atténue au fur et à mesure que les sols se réchauffent. En effet, quand le sol s'assèche, l'eau qui a une chaleur latente spécifique va retenir l'énergie davantage que le sol. Un sol sec va donc surchauffer plus facilement qu'un sol humide (Fu *et al.*, 2022). Le pouvoir tampon atteint ses limites et le différentiel observé entre le potentiel hydrique des trois peuplements devient alors plus petit car il y a moins d'eau dans le sol. C'est ce qui est observé lors de la période d'assèchement de juin 2020 (figure 4.33b).

Le point (4) fait écho à notre analyse exploitant les arbres de décision qui a mis en évidence l'existence de seuils dans les différentes variables étudiées ; ces seuils

variant selon les peuplements. Il est avéré que le degré de stress hydrique subi par les arbres va moduler leur réponse (chapitre 2). Quand le stress est moyen, il a été montré que la densité des stomates augmentait et qu'à partir d'un certain seuil cette densité diminuait (Sun *et al.*, 2014). Mais il faut également considérer l'interaction avec l'azote et le phosphore présents dans le sol qui vont avoir une influence sur la densité des stomates.

Nous n'avons pas pu faire de lien direct entre les stress hydriques et les impacts observés sur les processus biogéochimiques des sols et sur la physiologie des arbres. Il faudrait combiner l'étude du potentiel hydrique du sol aux études sur le transport de la sève brute, sur la respiration des sols et leur activité ionique, mais aussi à celles qui examinent la croissance radiale. Ces variables font l'objet des travaux de maîtrise de Maxence Lemire et de Charlène Laberge au sein du DOT-Lab. L'analyse combinée de tous ces paramètres permettrait de soulever de nouvelles pistes d'analyse.

Enfin, le comportement spécifique du hêtre, que nous avons mis en évidence, s'insère dans une problématique plus large sur le devenir des forêts dans notre région d'étude située au Québec. Quel sera le devenir du hêtre dans un contexte de changements climatiques ? Va-t-il continuer à proliférer dans les érablières comme le suggère la tendance observée depuis les années soixante-dix (Direction de la recherche forestière, 2017) ? Ou son expansion sera régulée par la maladie corticale⁶ dont les zones de dévastation et de destruction s'étendent au Québec et en particulier à l'Ouest de la province (Cale *et al.*, 2017; Direction de la protection des forêts, 2020) ? Les atouts du hêtre face à des stress hydriques ne feront-ils pas de lui un *gentil compagnon* pour l'érable (Nolet, 2023) ?

La migration de certains peuplements qui n'arriveront pas à s'adapter à la *nouvelle normalité* climatique qui s'installe dans le monde, et en particulier au Québec,

6. Maladie causée par l'interaction entre la cochenille du hêtre et deux champignons pathogènes

va bouleverser le paysage forestier actuel. Nous avons montré la vulnérabilité croissante de la sapinière qui semble fragilisée par un climat de plus en plus chaud et sec. Est-ce que cela se traduira par la disparition progressive du sapin des régions plus méridionales du Québec ?

Nous allons à présent ouvrir notre réflexion sous l'angle de la science des données. Nous l'avons mis en avant au fil de ces pages, la particularité de notre projet réside dans l'exploitation de la temporalité des données collectées à la SBL. Afin de développer un modèle prédictif du potentiel hydrique du sol, nous avons suivi la méthodologie de Box *et al.* (2008) qui permet de créer des modèles autorégressifs pour prédire des séries temporelles (*forecasting times series*). Ce sont des approches qui sont applicables à de nombreux domaines, de l'économie à l'écologie, ou encore par exemple dans la prédiction de l'évolution de cas de COVID-19 (Singh *et al.*, 2020). La démarche mathématique de Box-Jenkins a aidé à populariser les modèles ARIMA qui peuvent être utilisés dans des contextes variés en science de l'environnement comme en témoigne l'état de l'art réalisé par Kaur *et al.* (2023). En analysant plus d'une centaine d'articles, les auteurs concluent que la meilleure stratégie est de combiner les approches ARIMA avec d'autres modèles pour obtenir des résultats plus robustes et des modèles capables de mieux appréhender tous les motifs (*patterns*) qui se dégagent d'une même série temporelle (Kaur *et al.*, 2023). Ainsi, l'utilisation de méthodes hybrides pourrait améliorer la prédiction du potentiel hydrique du sol en palliant les difficultés décrites aux points précédents (9) et (10). Par exemple, nous citons l'étude de Singh *et al.* (2020) où le nombre de cas mortels liés à la COVID-19 était prédit grâce à un modèle ARIMA, sur un horizon de 16 jours en s'appuyant sur des données passées couvrant un intervalle de 66 jours. Le fait de combiner une approche par décomposition en ondelettes (*discrete wavelet transformation*) a amélioré de 50 % les résultats prédictifs de la France et des États-Unis et de 80 % les prédictions faites pour l'Italie, l'Espagne

et la Grande-Bretagne.

Cette technique décompose une série temporelle en sous-séries. Nous évoquons le problème de non-stationnarité dans les données de notre étude, cette approche par ondelettes permet d'avoir une décomposition plus fine d'une série. Celle-ci se fait à la fois dans le domaine temporel mais aussi à des échelles spatiales différentes grâce à l'introduction de coefficients propres à chaque sous-série (Richter *et al.*, 2022). La notion de spatialité fait référence ici à différentes localisations au sein d'une séquence temporelle. Il faut noter que la décomposition en ondelettes aide à surmonter le problème de la transformée de Fourier dont les fonctions de base sont localisées en fréquence et non en temps.

Une réflexion plus globale peut être menée sur la prédiction des séries temporelles et sur les méthodes qui s'avèrent les plus performantes. Les méthodes ARIMA ne sont pas nouvelles, et leur degré de complexité demeure relativement faible quand elles sont comparées à d'autres approches comme celles exploitant l'apprentissage profond. Mais y-a-t-il forcément une valeur ajoutée à augmenter la complexité d'un modèle? C'est une question que soulèvent Elsayed *et al.* (2021) et à laquelle ces auteurs tentent de répondre en confrontant des réseaux neuronaux profonds à des méthodes plus simples comme des arbres de régression (*gradient boosting trees*). Leur conclusion est qu'un modèle plus simple mais qui bénéficie d'une ingénierie des variables prédictives (nommée en anglais *feature engineering*) est comparable à des modèles plus sophistiqués.

Dacrema *et al.* (2019) ont poussé l'analyse plus loin en se demandant si des progrès étaient réellement faits en apprentissage machine, mettant alors en avant la difficulté de suivre quelles méthodes représentaient les méthodes de référence. À cette difficulté s'ajoute un enjeu de reproductibilité des études publiées. Ils ont donc testé 18 algorithmes présentés à des conférences réputées. Leur conclusion est que seulement sept algorithmes ont pu être reproduits avec un effort raisonnable, et que parmi eux six ont été dépassés par des méthodes simples de type *k plus*

proches voisins. Un des problèmes soulevés est l'absence d'une ligne de base qui servirait de référence. Dans notre étude, nous l'avons déjà souligné, nous n'avons pas de référence accessible dans la littérature pour pouvoir confronter notre modèle qui est donc un point de départ pour de possibles optimisations.

Néanmoins, nous avons systématiquement évalué l'erreur associée aux prédictions que nous avons réalisées sur les valeurs numériques du potentiel hydrique du sol. Pour la sapinière à érable, par exemple, pour le modèle ARIMAX(14,1,4) entraîné sur les données de 2020, nous avons obtenu une erreur quadratique moyenne (RMSE) allant de 0.38 pour un horizon prédictif de 5 jours à 21.05 pour un horizon de 150 jours (prédiction des données de 2019). Dans la revue de Kaur *et al.* (2023), les performances de modèles ARIMA développés dans le cadre d'études en sciences de l'environnement sont compilées, et nos résultats se situent dans les mêmes ordres de grandeur en terme d'erreur quadratique moyenne. Nous pouvons par exemple citer l'étude de Díaz-Robles *et al.* (2008) où un modèle ARIMA est développé pour prédire la qualité de l'air de la ville de Temuco (Chili). L'erreur quadratique moyenne obtenue est alors de 28.46.

Nous avons soulevé la notion de complexité du modèle, mais il y a aussi la notion d'*explicabilité*. En effet, l'ajout de couches dans les réseaux neuronaux diminue la compréhension que nous pouvons avoir du modèle. Or, si nous voulons faire des liens avec les phénomènes physiques observables, nous devons garder une certaine transparence sur notre modèle et envisager d'utiliser des méthodes qui nous le garantissent.

Un cadre d'analyse reposant sur l'algorithme SHAP (*Shapley Additive exPlanation*) est une piste qui semble intéressante et qui gagne en popularité (Lundberg et Lee, 2017). S'inspirant de la théorie des jeux coopératifs, elle analyse l'importance des variables prédictives utilisées dans un modèle en calculant l'impact de chacune sur la valeur obtenue en sortie. En ce sens, elle ressemble à la démarche que nous

avons utilisée avec les arbres de décision pour évaluer l'importance des variables exogènes dans la prédiction du potentiel hydrique du sol.

Finalement, pour modéliser le potentiel hydrique du sol, les approches d'apprentissage profond ne sont peut-être pas les plus adaptées. En effet, elles auront tendance à privilégier les événements de masse les plus probables introduisant un biais qui va privilégier la classe sur-représentée (Johnson et Khoshgoftaar, 2019).

Quand l'objectif est de classifier des données, la plupart des algorithmes postulent que les classes sont équilibrées. Or, en écologie, lorsque nous faisons face à ces événements extrêmes, les classes tendent à être déséquilibrées. Par exemple, pour le potentiel hydrique du sol, la classe correspondant à un état *élevé* du potentiel est moins représentée ce qui entraîne donc un *déséquilibre* dans les données. Peu d'études évaluent l'utilisation de l'apprentissage profond pour considérer le problème des données déséquilibrées dans des contextes autres que l'analyse d'images (Johnson et Khoshgoftaar, 2019). Le problème pourrait se réduire à une classification entre *événement rare* et *événement non rare* en minimisant les faux positifs, i.e. en s'assurant de bien identifier les vrais événements rares.

Deux possibilités peuvent être explorées : un sur-échantillonnage de la classe minoritaire ou un sous-échantillonnage de la classe majoritaire pour favoriser la sensibilité du classifieur envers la classe minoritaire dans le but d'augmenter le taux de vrais positifs. Cette stratégie est cependant déconseillée car des données importantes pourraient être perdues. Dans le cas d'un sous-échantillonnage aléatoire de la classe minoritaire, il faut au contraire faire attention à ne pas provoquer un sur-ajustement du modèle.

La technique SMOTE (*synthetic minority over-sampling technique*) a été proposée pour combiner un sur-échantillonnage de la classe minoritaire avec un sous-échantillonnage de la classe majoritaire (Chawla *et al.*, 2002). Le sur-échantillonnage est réalisé en générant des données synthétiques appartenant à la classe minoritaire

grâce à un algorithme de type *k plus proches voisins* qui s'appuie sur une distance euclidienne. Les auteurs ont testé leur algorithme sur neuf jeux différents de données débalancées, et ils ont pu montrer la meilleure performance d'une combinaison des deux approches d'échantillonnage.

Cette problématique liée à la prédiction des événements rares nous ramène à notre contexte d'étude décrit au chapitre 1 où nous avons souligné la difficulté de prédire de tels événements.

Martínez-Vilalta *et al.* (2021) ont proposé de mieux caractériser les seuils de rupture hydrologique (décrits au chapitre 1) de différents arbres et d'améliorer l'estimation de la valeur maximale du potentiel hydrique des feuilles (mesuré à l'aube et à midi) qui soit supportable avant cette rupture, et ce pour chaque espèce étudiée. Mettant en avant l'existence de valeurs extrêmes, Martínez-Vilalta *et al.* (2021) ont développé un modèle qui s'appuie sur la théorie des valeurs extrêmes (*extreme value theory* ou EVT). Cette théorie permet de modéliser séparément les données qui se trouvent sur la queue de leur distribution⁷. Deux approches sont couramment utilisées pour sélectionner ces données : la méthode *block maximum* et la méthode *peak-over-threshold* (POT). Dans la première, les données sont divisées en plusieurs blocs pour obtenir ensuite le maximum (ou minimum) de chacun. Dans la seconde, les valeurs supérieures (ou inférieures) à un certain seuil sont extraites. Elle a l'inconvénient de reposer sur le choix arbitraire d'un seuil, ce qui fait écho à l'enjeu que nous soulevions dans notre analyse.

Une alternative que nous trouvons intéressante et nous finirons sur cette possible ouverture méthodologique : la méthode EDM (*empirical dynamic modelling*) de Sugihara *et al.* (2012). Cette approche s'affranchit de tout seuil en proposant une démarche plus holistique qui va considérer toutes les données sans hypothèse

7. Pour plus de détails se référer à de Haan et Ferreira (2006) et Bousquet et Bernardara (2021)

a priori et sans s'appuyer sur des équations⁸. Pour Sugihara, des relations de causalité existent sans forcément montrer des relations de corrélation. Des liens entre les variables d'intérêt peuvent être cachés. En créant à partir des données des attracteurs inspirés de la théorie du chaos, il est possible de décrire la dynamique de systèmes non linéaires. Une corrélation peut apparaître mais sur une courte période de temps que les approches statistiques statiques ne peuvent saisir.

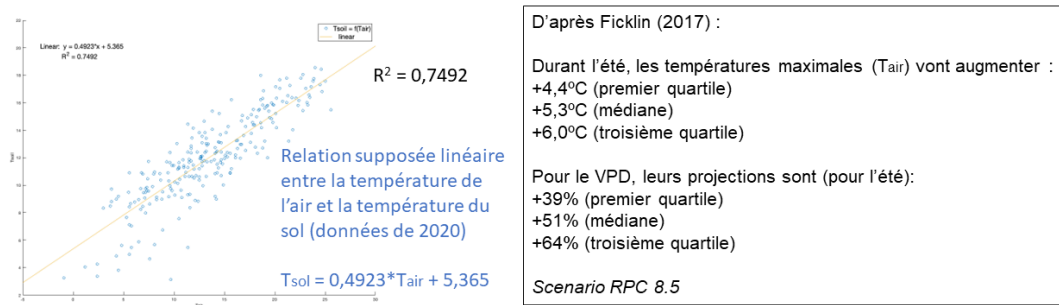
Il existe donc de nombreuses possibilités d'améliorer le modèle prédictif que nous avons développé. Dans une perspective à court terme, nous pourrions garder le modèle actuel, mais explorer différents scénarios climatiques qui auraient une influence sur les variables météorologiques utilisées dans les modèles ARIMAX développés. Nous pourrions utiliser le modèle pour prédire le potentiel hydrique du sol dans un contexte de changements climatiques.

Ainsi, les variables exogènes pourraient être générées synthétiquement, selon trois scénarios définis comme léger, modéré et extrême. À partir de ces trois environnements météorologiques hypothétiques qui reflètent l'impact des changements climatiques, l'objectif serait de prédire l'évolution temporelle du potentiel hydrique du sol sur un horizon de 20 à 40 jours.

En nous appuyant sur l'étude théorique de Ficklin et Novick (2017) qui analysent les projections du VPD en Amérique du Nord durant les étés, le scénario *léger* correspondrait aux prédictions de température de l'air et du VPD du premier quartile du pourcentage d'augmentation entre les valeurs de référence (1979-2013) et leur horizon de prédiction (scénario 25), le scénario *modéré* à celles de la médiane, et le scénario *extrême* à celles du troisième quartile (scénario 75). Les auteurs ont utilisé le modèle GCM (*general circulation model*) pour prédire les augmentations du VPD et de la température de l'air sur la période [2065-2099]. Ils se sont appuyés sur le scénario de forçage radiatif RPC8.5, d'après Sanford *et al.* (2014).

8. <https://deepeco.ucsd.edu/>

La figure 4.42 compile les variables exogènes et leur nouvelle valeur qui pourraient être utilisées, selon les trois scénarios.



	<i>Scenario 75</i>	<i>Scenario médian</i>	<i>Scenario 25</i>
(Tair)_mod [°C]	Tair + 6,0	Tair + 5,3	Tair + 4,4
(Tsol)_mod [°C]	(Tair+6)*0,4923 + 5,365	(Tair+5,3)*0,4923 + 5,365	(Tair+4,4)*0,4923 + 5,365
(VPD)_mod [kPa]	VPD + 64%VPD	VPD + 51%VPD	VPD + 39%VPD
(Prpc)_mod [mm]	(Prpc + 30%Prpc)	(Prpc + 15%Prpc)	(Prpc + 5%Prpc)

FIGURE 4.42 : Trois scénarios synthétiques pour générer les variables exogènes, d'après Ficklin et Novick (2017).

Le scénario 25 s'appuierait sur une augmentation de 4.4°C de la température de l'air, ce qui impliquerait, d'après le modèle de Ficklin et Novick (2017), une augmentation de 39% du VPD. La température du sol pourrait être déterminée en évaluant la relation observée en 2020 avec la température de l'air, en faisant l'hypothèse simplificatrice qu'une relation linéaire existe entre les deux variables. L'équation de la droite ainsi obtenue nous permettrait d'interpoler la température du sol. La même équation pourrait être appliquée aux deux autres scénarios, mais avec une température de l'air qui augmente. Pour les précipitations, nous pourrions nous appuyer sur les prédictions détaillées au chapitre 1, à l'échelle du Canada, qui prévoient une augmentation de 8% à 70% selon les régions (Environnement et

Changement climatique Canada, 2017b). Nous pourrions rester dans un intervalle compris entre +5 % et +30 %.

Pour terminer sur le fil rouge que nous avons tenté de déployer à travers notre projet, nous reprenons la question fondamentale que posent Pennekamp *et al.* (2019) : « *A first step towards predicting the future of any system is understanding if the observations of that system contain information about the future, i.e. does the system have a memory ?* » Pour Pennekamp *et al.* (2019), nous pouvons concevoir l'information totale contenue dans chaque observation comme une combinaison d'information provenant d'états passés (de l'information redondante) et d'information uniquement disponible dans l'état présent (de l'information nouvelle). Si le système est hautement redondant et que chaque observation apporte peu d'information nouvelle, la série temporelle résultante est, en théorie, hautement prédictible. Ce qui revient à dire que le système possède une prédictibilité intrinsèque élevée. Les systèmes stochastiques auront tendance à être peu redondants et donc difficilement prédictibles. La perte d'information liée aux mesures imparfaites implique également une perte d'information qui va réduire la prédictibilité. En utilisant l'entropie de permutation introduite par Bandt et Pompe (2002), une métrique capable de caractériser la complexité d'une série temporelle, Pennekamp *et al.* (2019) montrent que celle-ci permet de quantifier la prédictibilité d'un système écologique. En testant alors trois approches prédictives sur une base de données publique (Global Population Dynamic Database), ils évaluent la capacité de chacune à exploiter la structure des données. La figure 4.43 extraite de Pennekamp *et al.* (2019) illustre leur approche. Le modèle ARIMA (en vert) et le modèle EDM (en bleu) donnent une erreur quadratique moyenne plus basse qu'un modèle reposant sur une simple moyenne des données d'entraînement (en rouge).

Il faut retenir de la figure 4.43 que deux leviers sont accessibles pour augmenter la prédictibilité d'un système : diminuer la perte d'information liée aux mesures et

trouver le modèle le plus adapté aux données étudiées.

En considérant le cas d'une érablière à bouleau jaune à la marge nordique de son aire de répartition, le suivi et la modélisation du potentiel hydrique du sol dans un contexte de stress climatiques nous a permis de montrer l'existence chez les trois peuplements étudiés d'une forme de mémoire écologique, ou tout du moins de l'existence de temporalités multiples qui s'influencent mutuellement. Pour poursuivre notre étude, il faudrait ajouter à notre modèle une dimension spatiale qui soit capable d'appréhender la dynamique entre chaque arbre pour mieux réfléchir à l'échelle du peuplement.

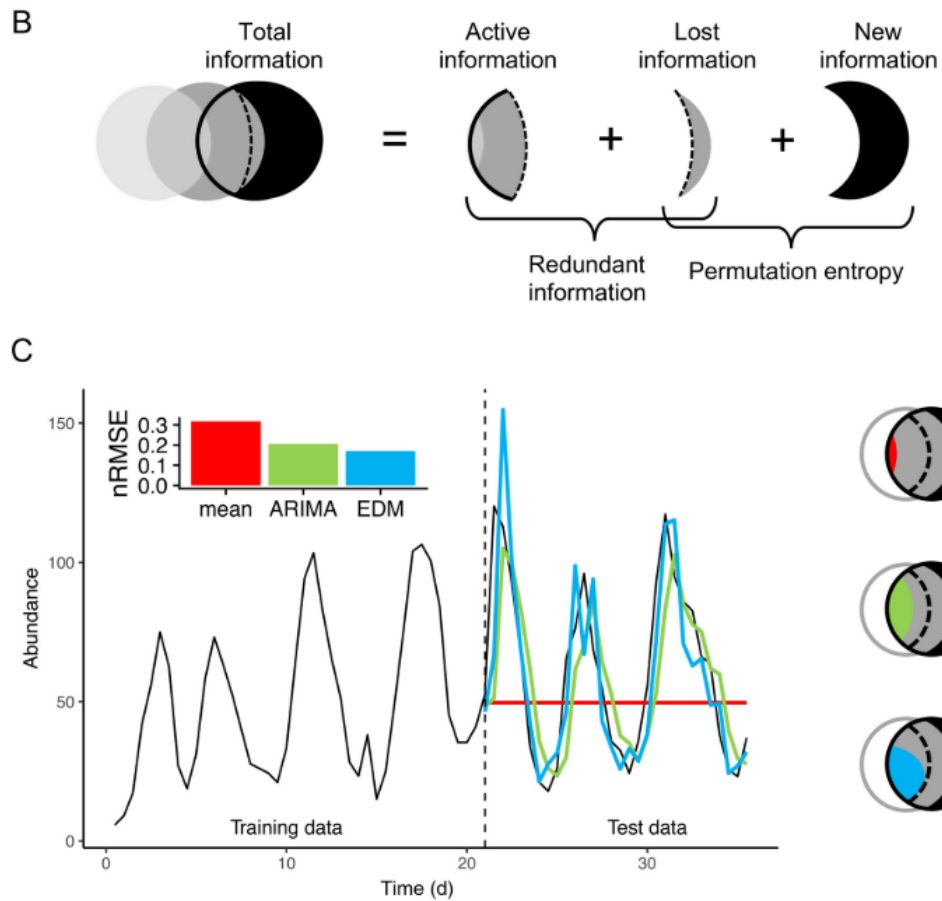


FIGURE 4.43 : L'information totale d'une observation (cercle noir en B) est composée de l'information nouvelle et de l'information redondante. La prédictibilité réalisée est le degré d'exploitation de l'information active de la série temporelle par le modèle. Le modèle ARIMA (cercle en vert) et le modèle EDM (cercle en bleu) exploitent davantage cette information active et donnent un RMSE plus faible (en C). Figure extraite de Pennekamp *et al.* (2019).

CONCLUSION

Prédire, ce n'est pas expliquer. Néanmoins, s'appuyer sur des modèles prédictifs permet d'apporter de nouvelles hypothèses pour guider la mise en place éventuelle de nouveaux dispositifs expérimentaux ou pour explorer plus en profondeur les liens entre les variables étudiées. En mettant en évidence les relations qui existent entre ces variables, il est ensuite possible de clarifier la compréhension des phénomènes physiques. Cela demande une confrontation avec les experts du domaine pour que la structure causale du modèle puisse émerger. Cette garantie permettra au modèle d'être robuste aux changements de contexte, et aux changements à travers le temps (Malik, 2020). Ce projet nous a permis d'appréhender cette dynamique collaborative au coeur de la science des données.

En utilisant différentes approches prédictives, comme les arbres de décision et les modèles autorégressifs à moyenne mobile, nous avons pu apporter de nouvelles hypothèses concernant la plus grande résilience de l'érablière à hêtre face à des stress hydriques en comparaison avec la sapinière à érable. En isolant des périodes dites de sécheresse-flash avec un déficit de précipitations, nous avons confirmé le comportement spécifique de l'érablière à hêtre.

De plus, l'analyse de l'évolution temporelle de l'autocorrélation du potentiel hydrique du sol entre 2017 et 2020 a suggéré pour la sapinière un rapprochement plus imminent d'un changement de régime de son potentiel hydrique. L'augmentation du coefficient d'asymétrie des données a conforté cette tendance.

Notre démarche méthodologique qui considère les données comme des séries chronologiques a permis de mettre en évidence le décalage temporel qui existait dans la corrélation entre le potentiel hydrique du sol et les variables météorologiques, ce qu'une approche statistique descriptive n'aurait pas pu appréhender. Ce type

d'analyse est prometteuse pour étudier l'effet de mémoire écologique que nous avons présenté au chapitre 2. Nous avons également précisé les limites de notre approche méthodologique qui permet une prédiction performante du potentiel hydrique du sol jusqu'à 60 jours. D'autres approches reposant sur des réseaux neuronaux récurrents comme les réseaux à mémoire à court et moyen terme (LSTM) pourraient étendre la taille de l'horizon prédictif. Des méthodes hybrides pourraient aussi améliorer les résultats obtenus.

Nous avons proposé un modèle prédictif que nous avons testé et validé, mais qui demanderait dans une prochaine étape d'être optimisé puis confronté à d'autres approches et dans d'autres contextes d'étude. Comme le souligne Aber (1997) : « *When models cannot fail, we can't learn from them. Models can help frame questions (...). By knowing that a model can fail, we can then have more confidence in it when it does succeed.* »

Enfin, anticiper l'avenir à travers le développement de modèles prédictifs permet de se fixer un point de départ qui pourrait s'apparenter à ce que Nicole Huybens appelle la *forêt souhaitée* (Huybens, 2017). Celle-ci initie alors « *la co-construction d'écosystèmes biodiversifiés et résilients qui permettent la coexistence nature/humanité.* » À travers cette vision partenariale, la technologie déployée par le projet *SmartForests* a la capacité de comprendre la complexité de la forêt grâce à un dialogue constant entre de nombreuses disciplines : de la science des données aux sciences de l'environnement, en incluant également les sciences humaines et sociales.

RÉFÉRENCES

- Aber, J. (1997). Why don't we believe the models? *Bulletin of Ecological Society of America*, 232–233.
- Abu-Hamdeh, N. H. (2003). Thermal properties of soils as affected by density and water content. *Biosystems Engineering*, 86, 97–102. [http://dx.doi.org/10.1016/S1537-5110\(03\)00112-0](http://dx.doi.org/10.1016/S1537-5110(03)00112-0)
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Akbar, R., Gianotti, D. J., McColl, K. A., Haghghi, E., Salvucci, G. D. et Entekhabi, D. (2018). Estimation of landscape soil water losses from satellite observations of soil moisture. *Journal of Hydrometeorology*, 19, 871–889. <http://dx.doi.org/10.1175/JHM-D-17-0200.1>
- Alizadeh, O., Qadimi, M., Zolghadrshojaee, M. et Irannejad, P. (2022). Frequency of different types of el niño events under global warming. *International Journal of Climatology*, 1–13.
- Allen, C. D., Breshears, D. D. et McDowell, N. G. (2015). On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere*, 6, 1–55. <http://dx.doi.org/10.1890/ES15-00203.1>
- Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J. H.,

- Allard, G., Running, S. W., Semerci, A. et Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, 259, 660–684. <http://dx.doi.org/10.1016/j.foreco.2009.09.001>
- Amaral, J., Ribeyre, Z., Vigneaud, J., Sow, M. D., Fichot, R., Messier, C., Pinto, G., Nolet, P. et Maury, S. (2020). Advances and promises of epigenetics for forest trees. *Forests*, 11. <http://dx.doi.org/10.3390/f11090976>
- Anderegg, W., Schwalm, C., Biondi, F., Camarero, J., Koch, G., Litvak, M., Ogle, K., Shaw, J., Shevliakova, E., Williams, A., Wolf, A., Ziaco, E. et Pacala, S. (2015). Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science*, 349, 528–532.
- Anderegg, W. R., Trugman, A. T., Badgley, G., Konings, A. G. et Shaw, J. (2020). Divergent forest sensitivity to repeated extreme droughts. *Nature Climate Change*, 10, 1091–1095. <http://dx.doi.org/10.1038/s41558-020-00919-1>
- Anderegg, W. R. L., Klein, T., Bartlett, M., Sack, L., Pellegrini, A. F. A., Choat, B. et Jansen, S. (2016). Meta-analysis reveals that hydraulic traits explain cross-species patterns of drought-induced tree mortality across the globe. *PNAS*, 113, 5024–5029. <http://dx.doi.org/10.5061/dryad.116j2>
- Arkhangelskaya, T. et Lukyashchenko, K. (2018). Estimating soil thermal diffusivity at different water contents from easily available data on soil texture, bulk density, and organic carbon content. *Biosystems Engineering*, 168, 83–95. <http://dx.doi.org/10.1016/j.biosystemseng.2017.06.011>
- Bandt, C. et Pompe, B. (2002). Permutation entropy : A natural complexity measure for time series. *Phys. Rev. Lett.*, 88, 174102. <http://dx.doi.org/10.1103/PhysRevLett.88.174102>

- Biggs, R., Carpenter, S. R. et Brock, W. A. (2009). Detecting an impending regime shift in time to avert it. *Proc Natl Acad Sci USA*, 106, 826–831. Récupéré de <http://dx.doi.org/10.1073/pnas.0811729106>
- Blair, G. S. (2020). A tale of two cities : Reflections on digital technology and the natural environment. *Patterns*, 1. <http://dx.doi.org/10.1016/j.patter.2020.100068>
- Blair, G. S., Bassett, R., Bastin, L., Beevers, L., Borrajo, M. I., Brown, M., Dance, S. L., Dionescu, A., Edwards, L., Ferrario, M. A., Fraser, R., Fraser, H., Gardner, S., Henrys, P., Hey, T., Homann, S., Huijbers, C., Hutchison, J., Jonathan, P., Lamb, R., Laurie, S., Leeson, A., Leslie, D., McMillan, M., Nundloll, V., Oyebamiji, O., Phillipson, J., Pope, V., Prudden, R., Reis, S., Salama, M., Samreen, F., Sejdinovic, D., Simm, W., Street, R., Thornton, L., Towe, R., Hey, J. V., Vieno, M., Waller, J. et Watkins, J. (2021). The role of digital technologies in responding to the grand challenges of the natural environment : The Windermere accord. *Patterns*, 2. <http://dx.doi.org/10.1016/j.patter.2020.100156>
- Blair, G. S., Henrys, P., Leeson, A., Watkins, J., Eastoe, E., Jarvis, S. et Young, P. J. (2019). Data science of the natural environment : A research roadmap. *Frontiers in Environmental Science*, 7, 121. <http://dx.doi.org/10.3389/fenvs.2019.00121>
- Boerlijst, M. C., Oudman, T. et de Roos, A. M. (2013). Catastrophic collapse can occur without early warning : Examples of silent catastrophes in structured ecological models. *PLoS ONE*, 8. <http://dx.doi.org/10.1371/journal.pone.0062033>
- Boisvert-Marsh, L., Pedlar, J. H., de Blois, S., Squin, A. L., Lawrence, K., McKenney, D. W., Williams, C. et Aubin, I. (2022). Migration-based simulations for

- Canadian trees show limited tracking of suitable climate under climate change. *Diversity and Distributions*, 28, 2330–2348. <http://dx.doi.org/10.1111/ddi.13630>
- Boisvert-Marsh, L., Périé, C. et Blois, S. D. (2014). Shifting with climate ? Evidence for recent changes in tree species distribution at high latitudes. *Ecosphere*, 5. <http://dx.doi.org/10.1890/ES14-00111.1>
- Bousquet, N. et Bernardara, P. (2021). *Extreme value theory with applications to natural hazards : from statistical theory to industrial practice*. Springer. Récupéré de <https://doi.org/10.1007/978-3-030-74942-2>
- Box, G., Jenkins, G. et Reinsel, G. (2008). *Time Series Analysis : Forecasting and Control. 4rd ed.* NJ : Prentice Hall.
- Breiman, L., Friedman, J., Olshen, R. et Stone, C. (1984). *Classification and Regression Trees*. FL : Chapman & Hall.
- Bruce, T. J., Matthes, M. C., Napier, J. A. et Pickett, J. A. (2007). Stressful "memories" of plants : Evidence and possible mechanisms. *Plant Science*, 173, 603–608. <http://dx.doi.org/10.1016/j.plantsci.2007.09.002>
- Byeon, S., Kim, S., Hong, J., Kim, T. K., Huh, W., Kim, K., Lee, M., Lee, H., Kim, S., Park, C., Bhusal, N., Han, A. R., Chandrasekaran, U. et Kim, H. S. (2023). Drought hardening effect on improving transplant stress tolerance in *Pinus densiflora*. *Environmental and Experimental Botany*, p. 105222. <http://dx.doi.org/10.1016/j.envexpbot.2023.105222>
- Bélangier, N., Collin, A., Khelifa, R. et Lebel-Desrosiers, S. (2021). Balsam Fir and American Beech influence soil respiration rates in opposite directions in a Sugar Maple forest near its northern range limit. *Frontiers in Forests and Global Change*, 4 :664584. <http://dx.doi.org/10.3389/ffgc.2021.664584>

- Bélangier, N., Collin, A., Ricard-Piché, J., Kembel, S. W. et Rivest, D. (2019). Microsite conditions influence leaf litter decomposition in sugar maple bioclimatic domain of quebec. *Biogeochemistry*, *145*, 107–126. <http://dx.doi.org/10.1007/s10533-019-00594-1>. Récupéré de <https://doi.org/10.1007/s10533-019-00594-1>
- Bélangier, N., Holmdem, C., Courchesne, F. et Côté, B. (2012). Constraining soil mineral weathering $87\text{Sr}/86\text{Sr}$ for calcium apportionment studies of a deciduous forest growing on soils developed from granitoid igneous rocks. *Geoderma*, *185*, 84–96. <http://dx.doi.org/DOI:10.1016/j.geoderma.2012.03.024>
- Cale, J. A., Garrison-Johnston, M. T., Teale, S. A. et Castello, J. D. (2017). Beech bark disease in North America : Over a century of research revisited. *Forest Ecology and Management*, *394*, 86–103. <http://dx.doi.org/https://doi.org/10.1016/j.foreco.2017.03.031>
- Canarini, A., Schmidt, H., Fuchslueger, L., Martin, V., Herbold, C. W., Zezula, D., Gündler, P., Hasibeder, R., Jecmenica, M., Bahn, M. et Richter, A. (2021). Ecological memory of recurrent drought modifies soil processes via changes in soil microbial community. *Nature Communications*, *12*. <http://dx.doi.org/10.1038/s41467-021-25675-4>
- Capdevila, P., Stott, I., Menor, I. O., Stouffer, D. B., Raimundo, R. L., White, H., Barbour, M. et Salguero-Gómez, R. (2021). Reconciling resilience across ecological systems, species and subdisciplines. *Journal of Ecology*, *109*, 3102–3113. <http://dx.doi.org/10.1111/1365-2745.13775>
- Charron, I. (2016). *Guide sur les scénarios climatiques : Utilisation de l'information climatique pour guider la recherche et la prise de décision en matière d'adaptation*. Ouranos. Récupéré de www.ouranos.ca

- Chawla, N. V., Bowyer, K. W., Hall, L. O. et Kegelmeyer, W. P. (2002). SMOTE : Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <http://dx.doi.org/10.1613/jair.953>. Récupéré de <https://doi.org/10.1613%2Fjair.953>
- Choat, B., Jansen, S., Brodribb, T. J., Cochard, H., Delzon, S., Bhaskar, R., Bucci, S. J., Feild, T. S., Gleason, S. M., Hacke, U. G., Jacobsen, A. L., Lens, F., Maherali, H., Martínez-Vilalta, J., Mayr, S., Mencuccini, M., Mitchell, P. J., Nardini, A., Pittermann, J., Pratt, R. B., Sperry, J. S., Westoby, M., Wright, I. J. et Zanne, A. E. (2012). Global convergence in the vulnerability of forests to drought. *Nature*, 491, 752–755. <http://dx.doi.org/10.1038/nature11688>
- Cole, L. E., Bhagwat, S. A. et Willis, K. J. (2014). Recovery and resilience of tropical forests after disturbance. *Nature Communications*, 5. <http://dx.doi.org/10.1038/ncomms4906>
- Collin, A., Messier, C. et Bélanger, N. (2017a). Conifer presence may negatively affect sugar maple’s ability to migrate into the boreal forest through reduced foliar nutritional status. *Ecosystems*, 20, 701–716. <http://dx.doi.org/10.1007/s10021-016-0045-4>
- Collin, A., Messier, C., Kembel, S. W. et Bélanger, N. (2017b). Low light availability associated with American beech is the main factor for reduced sugar maple seedling survival and growth rates in a hardwood forest of Southern Quebec. *Forests*, 8. <http://dx.doi.org/10.3390/f8110413>
- Coolong, T., Snyder, J., Warner, R., Strang, J. et Surendran, S. (2012). The relationship between soil water potential, environmental factors, and plant moisture status for poblano pepper grown using tensiometer-scheduled irrigation. *International Journal of Vegetable Science*, 18, 137–152.

- Courcot, B., Université T luq (2023). *Suivi et mod lisation du potentiel hydrique du sol dans un contexte de stress hydrique : le cas d'une  rabli re   bouleau jaune   la marge nordique de sa distribution (Notes de recherche du Laboratoire sur la science des donn es)*. , 1–169., R cup r  de <https://r-libre.telug.ca/2977/>
- Dacrema, M. F., Cremonesi, P. et Jannach, D. (2019). Are we really making much progress? A worrying analysis of recent neural recommendation approaches. *CoRR*, *abs/1907.06902*. R cup r  de <http://arxiv.org/abs/1907.06902>
- Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., K fi, S., Livina, V., Seekell, D. A., van Nes, E. H. et Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE*, *7*. <http://dx.doi.org/10.1371/journal.pone.0041010>
- Dakos, V., Carpenter, S. R., van Nes, E. H. et Scheffer, M. (2015). Resilience indicators : Prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B : Biological Sciences*, *370*, 1–10. <http://dx.doi.org/10.1098/rstb.2013.0263>
- Daszkowska-Golec, A. et Szarejko, I. (2013). Open or close the gate - stomata action under the control of phytohormones in drought stress conditions. *Frontiers in Plant Science*, *4*. <http://dx.doi.org/10.3389/fpls.2013.00138>
- de Haan, L. et Ferreira, A. (2006). *Extreme value theory : an introduction*. Ser. Springer series in operations research and financial engineering. R cup r  de <https://doi.org/10.1007/0-387-34471-3>
- Delettre, O. (2021). Identity of ecological systems and the meaning of resilience. *Journal of Ecology*, *109*(9), 3147–3156. <http://dx.doi.org/https://doi.org/10.1111/1365-2745.13655>

DeSoto, L., Cailleret, M., Sterck, F., Jansen, S., Kramer, K., Robert, E. M., Aakala, T., Amoroso, M. M., Bigler, C., Camarero, J. J., Čufar, K., Gea-Izquierdo, G., Gillner, S., Haavik, L. J., Hereş, A. M., Kane, J. M., Kharuk, V. I., Kitzberger, T., Klein, T., Levanič, T., Linares, J. C., Mäkinen, H., Oberhuber, W., Papadopoulos, A., Rohner, B., Sangüesa-Barreda, G., Stojanovic, D. B., Suárez, M. L., Villalba, R. et Martínez-Vilalta, J. (2020). Low growth resilience to drought is related to future mortality risk in trees. *Nature Communications*, 11. <http://dx.doi.org/10.1038/s41467-020-14300-5>

Direction de la protection des forêts, Gouvernement du Québec, ministère des Forêts, de la Faune et des Parcs (2020). *Évolution de la maladie corticale du hêtre au Québec en 2019*. , 1–14., Récupéré de https://mffp.gouv.qc.ca/documents/forets/Evolution-maladie-corticale_2019.pdf

Direction de la recherche forestière, Gouvernement du Québec, ministère des Forêts, de la Faune et des Parcs (2017). *Expansion du hêtre à grandes feuilles et déclin de l'érable à sucre au Québec : portrait de la situation, défis et pistes de solution. Avis scientifique du comité chargé d'étudier l'écologie et la sylviculture des peuplements contenant du hêtre et de l'érable*. , 1–146.

D'Orangeville, L., Houle, D., Duchesne, L., Phillips, R. P., Bergeron, Y. et Kneeshaw, D. D. (2018a). Beneficial effects of climate warming on boreal tree growth may be transitory. *Nature Communications*, 9, 2339–2351.

D'Orangeville, L., Maxwell, J. T., Kneeshaw, D. D., Pederson, N., Duchesne, L., Logan, T., Houle, D., Arseneault, D., Beier, C. M., Bishop, D. A., Druckenbrod, D. L., Fraver, S., Girard, F., Halman, J. M., Hansen, C., Hart, J. L., Hartmann, H., Kaye, M., LeBlanc, D. C., Manzoni, S., Ouimet, R., Rayback, S. A., Rollinson, C. R. et Phillips, R. P. (2018b). Drought timing and local climate determine

- the sensitivity of eastern temperate forests to drought. *Global Change Biology*, 24, 2339 – 2351.
- Duchaufour, P., Faivre, P., Poulenard, J. et Gury, M. (2020). *Introduction à la science du sol* (7 éd.). Dunod.
- Duchesne, L., Domaine, E., Schab, A. et D'Orangeville, L. (2020). Sensibilité de la croissance radiale des arbres en forêt boréale en réponse aux sécheresses ponctuelles et à la variabilité interannuelle du bilan hydrique climatique. numéro 185, Gouvernement du Québec, ministère des Forêts, de la Faune et des Parcs, Direction de la recherche forestière, mémoire de recherche forestière.
- Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G. et Moncada-Herrera, J. A. (2008). A hybrid arima and artificial neural networks model to forecast particulate matter in urban areas : The case of temuco, chile. *Atmospheric Environment*, 42(35), 8331–8340. <http://dx.doi.org/https://doi.org/10.1016/j.atmosenv.2008.07.020>
- E. Bush and D.S. Lemmen. (2019). *Rapport sur le climat changeant du Canada*. Gouvernement du Canada. Récupéré de <https://changingclimate.ca/CCCR2019/fr/chapitre/4-0/>
- Elsayed, S., Thyssens, D., Rashed, A., Schmidt-Thieme, L. et Jomaa, H. S. (2021). Do we really need deep learning models for time series forecasting? *CoRR*, *abs/2101.02118*. Récupéré de <https://arxiv.org/abs/2101.02118>
- Environnement et Changement climatique Canada, Gouvernement du Canada (2017a). *Données et scénarios climatiques : Synthèse des observations et des résultats récents de modélisation (changement dans les précipitations)*. , Récupéré de [175](https://www.canada.ca/fr/environnement-changement-climatique/services/changements-climatiques/centre-</p></div><div data-bbox=)

canadien-services-climatiques/essentiels/tendances-projections/
changements-precipitations.html

Environnement et Changement climatique Canada, Gouvernement du Canada (2017b). *Données et scénarios climatiques : Synthèse des observations et des résultats récents de modélisation (changement dans les températures)*. , Récupéré de <https://www.canada.ca/fr/environnement-changement-climatique/services/changements-climatiques/centre-canadien-services-climatiques/essentiels/tendances-projections/changements-temperature.html>

Evans, B. J. et Lyons, T. (2013). Bioclimatic extremes drive forest mortality in southwest, Western Australia. *Climate*, 1, 28–52. <http://dx.doi.org/10.3390/cli1020028>

FAO et PNUE, FAO et PNUE (2020). *La situation des forêts du monde 2020. Forêts, biodiversité et activité humaine*.

Feldman, A. F., Gianotti, D. J. S., Trigo, I. F., Salvucci, G. D. et Entekhabi, D. (2019). Satellite-based assessment of land surface energy partitioning–soil moisture relationships and effects of confounding variables. *Water Resources Research*, 55, 10657–10677. <http://dx.doi.org/10.1029/2019WR025874>

Ficklin, D. L. et Novick, K. A. (2017). Historic and projected changes in vapor pressure deficit suggest a continental-scale drying of the United States atmosphere. *Journal of Geophysical Research*, 122, 2061–2079.

Forget, E., Drever, R. et Lorenzetti, F. (2003). *Changements climatiques : impacts sur les forêts québécoises, revue de littérature*. Ouranos.

Forzieri, G., Dakos, V., McDowell, N. G., Ramdane, A. et Cescatti, A. (2022).

- Emerging signals of declining forest resilience under climate change. *Nature*, 608, 534–539. <http://dx.doi.org/10.1038/s41586-022-04959-9>
- Fu, Z., Ciais, P., Makowski, D., Bastos, A., Stoy, P. C., Ibrom, A., Knohl, A., Migliavacca, M., Cuntz, M., Šigut, L., Peichl, M., Loustau, D., El-Madany, T. S., Buchmann, N., Gharun, M., Janssens, I., Markwitz, C., Grünwald, T., Rebmann, C., Mölder, M., Varlagin, A., Mammarella, I., Kolari, P., Bernhofer, C., Heliasz, M., Vincke, C., Pitacco, A., Cremonese, E., Foltýnová, L. et Wigneron, J. P. (2022). Uncovering the critical soil moisture thresholds of plant water stress for European ecosystems. *Global Change Biology*, 28, 2111–2123. <http://dx.doi.org/10.1111/gcb.16050>
- Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Domingues, L. G., Cassol, H. L. G., Tejada, G., Aragão, L. E. O. C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A. H., Corrêa, S. M., Anderson, L., Randow, C. V., Correia, C. S. C., Crispim, S. P. et Neves, R. A. L. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595, 388–393. <http://dx.doi.org/10.1038/s41586-021-03629-6>
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, 424–459.
- Granier, A. et Cochard, H. (2023). L'architecture hydraulique, la cavitation et l'embolie des plantes. <https://plantes-et-eau.fr/documentation/etats-et-transferts-hydriques-dans-et-a-travers-la-plante/24-l-architecture-hydraulique-la-cavitation-et-l-embolie-des-plantes/58-l-architecture-hydraulique-la-cavitation-et-l-embolie-des-plantes>. Accessed : 2023-01-31.
- Guttal, V. et Jayaprakash, C. (2008). Changing skewness : an early warning

- signal of regime shifts in ecosystems. *Ecology Letters*, 11, 450–460. <http://dx.doi.org/https://doi.org/10.1111/j.1461-0248.2008.01160.x>
- Hassani, H., Huang, X. et Silva, E. (2019). Big data and climate change. *Big Data and Cognitive Computing*, 3, 1–17. <http://dx.doi.org/10.3390/bdcc3010012>
- Hawkins, E. et Sutton, R. (2012). Time of emergence of climate signals. *Geophysical Research Letters*, 39(1). <http://dx.doi.org/10.1029/2011GL050087>
- Hirota, M., Holmgren, M., Nes, E. H. V. et Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science*, 334(6053), 232–235. <http://dx.doi.org/10.1126/science.1210657>
- Hobbins, M. T., Wood, A., McEvoy, D. J., Huntington, J. L., Morton, C., Anderson, M. et Hain, C. (2016). The evaporative demand drought index. part i : Linking drought evolution to variations in evaporative demand. *Journal of Hydrometeorology*, 17(6), 1745 – 1761. <http://dx.doi.org/10.1175/JHM-D-15-0121.1>
- Holling, C. (1996). Engineering resilience versus ecological resilience. Dans P. Schulze (dir.). *Engineering within Ecological Constraints*, 31–43., Washington DC. National Academy Press.
- Huybens, N. (2017). *La forêt boréale : Relier sciences de la nature et éthique*, Dans *L'espoir malgré tout : L'oeuvre de Pierre Dansereau et l'avenir des sciences de l'environnement*, (p. 189–206). Presses de l'Université du Québec, (1 éd.)
- Irmak, S., Kukal, M. S. et Sharma, K. (2022). Soil moisture heterogeneity and sensor deployment in uniformly managed field with unitextural soil. *Agronomy Journal*, 114(3), 1800–1816. <http://dx.doi.org/https://doi.org/10.1002/agj2.21064>
- Irmak, S., Payero, J., VanDeWalle, B., Rees, J., Zoubek, G., Martin, D., Kranz, W., Eisenhauer, D. et Leininger, D. (2016). Principles and operational characteristics

of watermark granular matrix sensor to measure soil water status and its practical applications for irrigation management in various soil textures. EC783, University of Nebraska-Lincoln Extension.

Itter, M. S., Vanhatalo, J. et Finley, A. O. (2019). EcoMem : An R package for quantifying ecological memory. *Environmental Modelling & Software*, 119, 305–308. <http://dx.doi.org/https://doi.org/10.1016/j.envsoft.2019.06.004>

Iverson, L., Peters, M., Prasad, A. et Matthews, S. (2019a). Analysis of climate change impacts on tree species of the Eastern US : Results of DISTRIB-II modeling. *Forests*, 10(4), 302. <http://dx.doi.org/10.3390/f10040302>

Iverson, L. R., Prasad, A. M., Peters, M. P. et Matthews, S. N. (2019b). Facilitating adaptive forest management under climate change : A spatially specific synthesis of 125 species for habitat changes and assisted migration over the Eastern United States. *Forests*, 10(11). <http://dx.doi.org/10.3390/f10110989>

Jacques, C., Salon, C., Barnard, R. L., Vernoud, V. et Prudent, M. (2021). Drought stress memory at the plant cycle level : A review. *Plants*, 10. <http://dx.doi.org/10.3390/plants10091873>

Jakovlev, N. et Pickenhain, L. (1977). *Sportbiochemie. Sportmedizinische Schriftenreihe der DeutschenHochschule für Körperkultur.*

Janiak, A., Kwasniewski, M., Sowa, M., Gajek, K., Żmuda, K., Kościelniak, J. et Szarejko, I. (2018). No time to waste : Transcriptome study reveals that drought tolerance in barley may be attributed to stressed-like expression patterns that exist before the occurrence of stress. *Frontiers in Plant Science*, 8. <http://dx.doi.org/10.3389/fpls.2017.02212>

Jenny, H. (1980). *The soil resource : origin and behavior.* Springer.

- Jentsch, A., Kreyling, J. et Beierkuhnlein, C. (2007). A new generation of climate-change experiments : events, not trends. *Frontiers in Ecology and the Environment*, 5(7), 365–374. [http://dx.doi.org/https://doi.org/10.1890/1540-9295\(2007\)5\[365:ANGOCE\]2.0.CO;2](http://dx.doi.org/https://doi.org/10.1890/1540-9295(2007)5[365:ANGOCE]2.0.CO;2)
- Johnson, J. M. et Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6. <http://dx.doi.org/10.1186/s40537-019-0192-5>
- Johnstone, J. F., Allen, C. D., Franklin, J. F., Frelich, L. E., Harvey, B. J., Higuera, P. E., Mack, M. C., Meentemeyer, R. K., Metz, M. R., Perry, G. L., Schoennagel, T. et Turner, M. G. (2016). Changing disturbance regimes, ecological memory, and forest resilience. *Frontiers in Ecology and the Environment*, 14, 369–378. <http://dx.doi.org/10.1002/fee.1311>
- Kaur, J., Parmar, K. S. et Singh, S. (2023). Autoregressive models in environmental forecasting time series : a theoretical and application review. *Environmental Science and Pollution Research*, 30, 19617–19641. <http://dx.doi.org/10.1007/s11356-023-25148-9>
- Koster, R. D. et Suarez, M. J. (2001). Soil moisture memory in climate models. *Journal of Hydrometeorology*, 2.
- Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., Seekell, D. A., Scheffer, M., Nes, E. H. V. et Dakos, V. (2014). Early warning signals of ecological transitions : Methods for spatial patterns. *PLoS ONE*, 9. <http://dx.doi.org/10.1371/journal.pone.0092097>
- Kögler, F. et Söffker, D. (2020). State-based open-loop control of plant growth by means of water stress training. *Agricultural Water Management*, 230. <http://dx.doi.org/10.1016/j.agwat.2019.105963>

- Lake, P. S. (2008). *Flow-Generated Disturbances and Ecological Responses : Floods and Droughts*, Dans *Hydroecology and Ecohydrology*, (p. 75–92). John Wiley and Sons, Ltd
- Li, Y., Zhang, Q. B., Fang, O., Mu, Y. M., Jia, H. et Lyu, L. (2021). Recovery time of juniper trees is longer in wet than dry conditions on the Tibetan Plateau in the past two centuries. *Forest Ecology and Management*, 497. <http://dx.doi.org/10.1016/j.foreco.2021.119514>
- Liu, Y., Kumar, M., Katul, G. G. et Porporato, A. (2019). Reduced resilience as an early warning signal of forest mortality. *Nature Climate Change*, 9, 880–885. <http://dx.doi.org/10.1038/s41558-019-0583-9>
- Lloret, F., Jaime, L. A., Margalef-Marrase, J., Pérez-Navarro, M. A. et Batllori, E. (2022). Short-term forest resilience after drought-induced die-off in Southwestern European forests. *Science of the Total Environment*, 806. <http://dx.doi.org/10.1016/j.scitotenv.2021.150940>
- Lukas, J., Hobbins, M. T., Rangwala, I. et EDDI Team (2007). The EDDI user guide (NOAA).
- Lundberg, S. et Lee, S.-I. (2017). A unified approach to interpreting model predictions.
- L'Écuyer, P. (2020). Chaînes de Markov à temps discret, modèles stochastiques, cours IFT-3655. <https://www.iro.umontreal.ca/~lecuyer/ift3655/chaines-Markov.pdf>. Accessed : 2023-02-24.
- Léger-Beaulieu, C. (2018). *Réponses morphologiques et physiologiques des semis de trois espèces de conifères à un gradient de déficit hydrique en serre*. (Mémoire de maîtrise).

- Malik, M. M. (2020). A hierarchy of limitations in machine learning. *CoRR*, *abs/2002.05193*. Récupéré de <https://arxiv.org/abs/2002.05193>
- Martínez-Vilalta, J., Santiago, L. S., Poyatos, R., Badiella, L., de Cáceres, M., Aranda, I., Delzon, S., Vilagrosa, A. et Mencuccini, M. (2021). Towards a statistically robust determination of minimum water potential and hydraulic risk in plants. *New Phytologist*, *232*, 404–417. <http://dx.doi.org/10.1111/nph.17571>
- Masson-Delmotte, V., Zhai, P., Pörtner, H., Roberts, D., Skea, J., Shukla, P., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J., Chen, Y., Zhou, X., Gomis, M., Lonnoy, E., Maycock, T., Tignor, M. et Waterfield, T., Organisation météorologique mondiale (2018). *GIEC, 2018 : Résumé à l'intention des décideurs, Réchauffement planétaire de 1,5 °C, Rapport spécial du GIEC sur les conséquences d'un réchauffement planétaire de 1,5 °C par rapport aux niveaux préindustriels et les trajectoires associées d'émissions mondiales de gaz à effet de serre, dans le contexte du renforcement de la parade mondiale au changement climatique, du développement durable et de la lutte contre la pauvreté.* , 1–32.
- McColl, K., Wang, W., Peng, B., Akbar, R., Gianotti, D. J. S., Lu, H., Pan, M. et Entekhabi, D. (2017). Global characterization of surface soil moisture drydowns. *Geophysical Research Letters*, *44*, 3682–3690. <http://dx.doi.org/10.1002/2017GL072819>
- McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G. et Yezpez, E. A. (2008). Mechanisms of plant survival and mortality during drought : Why do some plants survive while others succumb to drought ? *New Phytologist*, *178*, 719–739. <http://dx.doi.org/10.1111/j.1469-8137.2008.02436.x>

- McEvoy, D. J., Huntington, J. L., Hobbins, M. T., Wood, A., Morton, C., Anderson, M. et Hain, C. (2016). The evaporative demand drought index. part ii : Conus-wide assessment against common drought indicators. *Journal of Hydrometeorology*, 17(6), 1763 – 1779. <http://dx.doi.org/10.1175/JHM-D-15-0122.1>
- Ministère des forêts de la faune et des parcs , Gouvernement du Québec (2021). *Stratégie d'adaptation de la gestion et de l'aménagement des forêts aux changements climatiques.* , 1–41.
- Mishra, A. K. et Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, 391, 202–216. <http://dx.doi.org/10.1016/j.jhydrol.2010.07.012>
- Monteith, J. et Unsworth, M. (2007). *Principles of Environmental Physics* (3 éd.). Elsevier Science.
- Murray, F. (1967). On the computation of saturation vapour pressure. *Journal of Applied Meteorology and Climatology*, 6(1), 203–204.
- Naikwade, P. V. (2023). *Plant responses to drought stress : morphological, physiological, molecular approaches, and drought resistance*, Dans *Plant Metabolites under Environmental Stress*, (p. 1–35). Apple Academic Press
- Nau, R. (2020). Introduction to ARIMA : nonseasonal models. <https://people.duke.edu/~rnau/411home.htm>. Accessed : 2023-03-07.
- Nolet, P. (2023). Le hêtre dans les érablières : Un gentil compagnon ou de la visite envahissante. Récupéré de <https://www.acfas.ca/evenements/congres/programme/90/600/632/c>
- Or, D. et Wraith, J. (2001). *Soil Physics Companion*. CRC Press.

- Orth, R. et Seneviratne, S. I. (2012). Analysis of soil moisture memory from observations in Europe. *Journal of Geophysical Research Atmospheres*, 117. <http://dx.doi.org/10.1029/2011JD017366>
- Otkin, J. A., Svoboda, M., Hunt, E. D., Ford, T. W., Anderson, M. C., Hain, C. et Basara, J. B. (2018). Flash droughts : A review and assessment of the challenges imposed by rapid-onset droughts in the United States. *Bulletin of the American Meteorological Society*, 99, 911–919. <http://dx.doi.org/10.1175/BAMS-D-17-0149.1>
- Pappas, C., Bélanger, N., Bergeron, Y., Blarquez, O., Chen, H. Y. H., Comeau, P. G., De Grandpré, L., Delagrangé, S., DesRochers, A., Diochon, A., D’Oran-geville, L., Drapeau, P., Duchesne, L., Filotas, E., Gennaretti, F., Houle, D., Lafleur, B., Langor, D., Lebel Desrosiers, S., Lorenzetti, F., Man, R., Messier, C., Girona, M. M., Nock, C., Thomas, B. R., Work, T. et Kneeshaw, D. (2022). *Smartforests Canada : A Network of Monitoring Plots for Forest Management Under Environmental Change*, Dans R. Tognetti, M. Smith, et P. Panzacchi (dir.). *Climate-Smart Forestry in Mountain Regions*, (p. 521–543). Springer International Publishing
- Peltier, D. M., Guo, J., Nguyen, P., Bangs, M., Wilson, M., Samuels-Crow, K., Yocom, L. L., Liu, Y., Fell, M. K., Shaw, J. D., Auty, D., Schwalm, C., Anderegg, W. R., Koch, G. W., Litvak, M. E. et Ogle, K. (2022). Temperature memory and non-structural carbohydrates mediate legacies of a hot drought in trees across the southwestern USA. *Tree physiology*, 42, 71–85. <http://dx.doi.org/10.1093/treephys/tpab091>
- Pendergrass, A. G., Meehl, G. A., Pulwarty, R., Hobbins, M., Hoell, A., AghaKou-chak, A., Bonfils, C. J. W., Gallant, A. J. E., Hoerling, M., Hoffmann, D., Kaatz, L., Lehner, F., Llewellyn, D., Mote, P., Neale, R. B., Overpeck, J. T., Sheffield, A.,

- Stahl, K., Svoboda, M., Wheeler, M. C., Wood, A. W. et Woodhouse, C. A. (2020). Flash droughts present a new challenge for subseasonal-to-seasonal prediction. *Nature Climate Change*, *10*, 191–199. <http://dx.doi.org/10.1038/s41558-020-0709-0>. Récupéré de <https://doi.org/10.1038/s41558-020-0709-0>
- Peng, C., Ma, Z., Lei, X., Zhu, Q., Chen, H., Wang, W., Liu, S., Li, W., Fang, X. et Zhou, X. (2011). A drought-induced pervasive increase in tree mortality across Canada’s boreal forests. *Nature Climate Change*, *1*, 467–471. <http://dx.doi.org/10.1038/nclimate1293>
- Pennekamp, F., Iles, A. C., Garland, J., Brennan, G., Brose, U., Gaedke, U., Jacob, U., Kratina, P., Matthews, B., Munch, S., Novak, M., Palamara, G. M., Rall, B. C., Rosenbaum, B., Tabi, A., Ward, C., Williams, R., Ye, H. et Petchey, O. L. (2019). The intrinsic predictability of ecological time series and its potential to guide forecasting. *Ecological Monographs*, *89*. <http://dx.doi.org/10.1002/ecm.1359>
- Pimm, S. (1984). The complexity and stability of ecosystems. *Nature*, *307*, 321–326.
- Rennenberg, H., Loreto, F., Polle, A., Brilli, F., Fares, S., Beniwal, R. S. et Gessler, A. (2006). Physiological responses of forest trees to heat and drought. *Plant Biology*, *8*(5), 556–571. <http://dx.doi.org/https://doi.org/10.1055/s-2006-924084>
- Ressources naturelles Canada, Service canadien des forêts (2020). *L’état des forêts au Canada. Rapport annuel 2020.* , 1–96., Récupéré de <https://scf.rncan.gc.ca/publications?id=40220>
- Ressources naturelles Canada (2021). *Indicateur : émissions et absorptions de carbone dans les forêts.* , Récupéré de <https://www.rncan.gc.ca/nos->

ressources-naturelles/forets/letat-forets-canada-rapport-
annuel/perturbations-forets-canada/16503

Ressources naturelles Canada (2022a). *Le carbone forestier*. , Récupéré de <https://www.rncan.gc.ca/changements-climatiques/changements-climatiques/carbone-forestier/13086>

Ressources naturelles Canada, Service canadien des forêts (2022b). *L'état des forêts au Canada. Rapport annuel 2022*. , 1–96.

Ressources naturelles Canada, Service canadien des forêts (2007). Les forêts du canada : cartes d'atlas canada. <https://www.rncan.gc.ca/cartes-outils-et-publications/cartes/forets/16875>. Accessed : 2023-01-24.

Ressources naturelles Canada, Service canadien des forêts (2020). Composition forestière au canada, carte (11 po x 17 po). <https://scf.rncan.gc.ca/publications?id=40222>. Accessed : 2023-01-24.

Richter, M. M., Sheuli, Kepuska, P. V. et Silaghi, M. (2022). *Chapter 3 : Fundamentals of signal transformations*, Dans *Signal Processing and Machine Learning with Applications*, (p. 88–91). Springer Nature Switzerland AG.

Rodríguez-Iturbe, I. et Porporato, A. (2005). *Plant water stress*, Dans *Ecohydrology of Water-Controlled Ecosystems : Soil Moisture and Plant Dynamics*, (p. 86–116). Cambridge University Press

Rosenzweig, C. et Hillel, D. (2008). *Climate variability and the global harvest : Impacts of El Niño and other oscillations on agro-ecosystems*. Oxford University Press, Inc. <http://dx.doi.org/10.1093/oso/9780195137637.001.0001>

Ruehr, N. K., Grote, R., Mayr, S. et Arneith, A. (2019). Beyond the extreme : Recovery of carbon and water relations in woody plants following heat and

- drought stress. *Tree Physiology*, *39*, 1285–1299. <http://dx.doi.org/10.1093/treephys/tpz032>
- Régnière, J. et Saint-Amand, R. (2017). BioSIM 11. Rapport d'information LAU-X-129, Ressources naturelles Canada, Service canadien des forêts, Centre de foresterie des Laurentides.
- Sanford, T., Frumhoff, P. C., Luers, A. et Gullett, J. (2014). The climate policy narrative for a dangerously warming world. *Nature Climate Change*, *4*, 164–166. <http://dx.doi.org/10.1038/nclimate2148>. Récupéré de <https://doi.org/10.1038/nclimate2148>
- Savitzky, A. et Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*, 1627–1639. doi : 10.1021/ac60214a047, <http://dx.doi.org/10.1021/ac60214a047>
- Scheffer, M. (2009). *Critical Transitions in Nature and Society*. Princeton University Press.
- Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M. et Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, *461*, 53–59. <http://dx.doi.org/10.1038/nature08227>
- Scheffer, M., Hirota, M., Holmgren, M., Nes, E. H. V. et Chapin, F. S. (2012). Thresholds for boreal biome transitions. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 21384–21389. <http://dx.doi.org/10.1073/pnas.1219844110>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

- Sebestyén, V., Czvetkó, T. et Abonyi, J. (2021). The applicability of big data in climate change research : The importance of system of systems thinking. *Frontiers in Environmental Science*, 9 :619092. <http://dx.doi.org/10.3389/fenvs.2021.619092>
- Sedaghatmehr, M., Thirumalaikumar, V. P., Kamranfar, I., Marmagne, A., Masclaux-Daubresse, C. et Balazadeh, S. (2019). A regulatory role of autophagy for resetting the memory of heat stress in plants. *Plant Cell and Environment*, 42, 1054–1064. <http://dx.doi.org/10.1111/pce.13426>
- Seekell, D. A., Carpenter, S. R. et Pace, M. L. (2011). Conditional heteroscedasticity as a leading indicator of ecological regime shifts. *American Naturalist*, 178, 442–451. <http://dx.doi.org/10.1086/661898>
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B. et Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate : A review. *Earth-Science Reviews*, 99, 125–161. <http://dx.doi.org/https://doi.org/10.1016/j.earscirev.2010.02.004>
- Senf, C., Buras, A., Zang, C. S., Rammig, A. et Seidl, R. (2020). Excess forest mortality is consistently linked to drought across Europe. *Nature Communications*, 11. <http://dx.doi.org/10.1038/s41467-020-19924-1>
- Shock, C., Barnum, J. et Seddigh, M. (1998). Calibration of watermark soil moisture sensors for irrigation management. Récupéré de https://www.researchgate.net/publication/228762944_Calibration_of_W_ermark_Soil_Moisture_Sensors_for_Irrigation_Management
- Shock, C., Wang, F., Flock, R., Feibert, E., Shock, C. et Pereira, A. (2013). Irrigation monitoring using soil water tension. EM8900, Oregon State University, Extension Service publication.

- Shock, C. C. et Wang, F.-X. (2011). Soil water tension, a powerful tool for productivity and stewardship. *HortScience*, 46, 178–185. Récupéré de <https://www.researchgate.net/publication/267277609>
- Sims, M., Potapov, P., Goldman, L. (2022). The world's last intact forests are becoming increasingly fragmented. <https://www.globalforestwatch.org/blog/data-and-research/worlds-last-intact-forests-are-becoming-increasingly-fragmented/>. Accessed : 2023-01-24.
- Singh, S., Parmar, K. S., Kumar, J. et Makkhan, S. J. S. (2020). Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. *Chaos, Solitons and Fractals*, 135, 109866. <http://dx.doi.org/https://doi.org/10.1016/j.chaos.2020.109866>
- Slette, I. J., Post, A. K., Awad, M., Even, T., Punzalan, A., Williams, S., Smith, M. D. et Knapp, A. K. (2019). How ecologists define drought, and why we should do better. *Global Change Biology*, 25, 3193–3200. <http://dx.doi.org/10.1111/gcb.14747>
- Soil Classification Working Group. (1998). *The Canadian System of Soil Classification* (3 éd.). Ottawa : Agriculture and Agri-Food Canada Publication and NRC Research Press. Récupéré de <https://sis.agr.gc.ca/cansis/publications/manuals/1998-cssc-ed3/index.html>
- Standish, R. J., Hobbs, R. J., Mayfield, M. M., Bestelmeyer, B. T., Suding, K. N., Battaglia, L. L., Eviner, V., Hawkes, C. V., Temperton, V. M., Cramer, V. A., Harris, J. A., Funk, J. L. et Thomas, P. A. (2014). Resilience in ecology : Abstraction, distraction, or where the action is ? *Biological Conservation*, 177, 43–51. <http://dx.doi.org/https://doi.org/10.1016/j.biocon.2014.06.008>

- Stevenson, S., Coats, S., Touma, D., Cole, J., Lehner, F., Fasullo, J. et Otto-Bliesner, B. (2022). Twenty-first century hydroclimate : A continually changing baseline, with more frequent extremes. *Proceedings of the National Academy of Sciences*, *119*, 1–9.
- Sugihara, G., May, R., Ye, H., hao Hsieh, C., Deyle, E., Fogarty, M. et Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500. <http://dx.doi.org/10.1126/science.1227079>. Récupéré de <https://www.science.org/doi/abs/10.1126/science.1227079>
- Sun, Y., Yan, F., Cui, X. et Liu, F. (2014). Plasticity in stomatal size and density of potato leaves under different irrigation and phosphorus regimes. *Journal of Plant Physiology*, *171*, 1248–1255. <http://dx.doi.org/10.1016/j.jplph.2014.06.002>
- Svoboda, M. D. et Fuchs, B. A. (2017). *Drought and Water Crises : Integrating Science, Management, and Policy*. CRC Press, (second edition éd.).
- Sánchez-Pinillos, M., D’Orangeville, L., Boulanger, Y., Comeau, P., Wang, J., Taylor, A. R. et Kneeshaw, D. (2022). Sequential droughts : A silent trigger of boreal forest mortality. *Global Change Biology*, *28*, 542–556. <http://dx.doi.org/10.1111/gcb.15913>
- Tetens, O. (1930). *Über einige meteorologische Begriffe*. Zeitschrift für geophysik. Friedrich Vieweg & Sohn Akt.- Gesellschaft.
- Trenberth, K. E., Dai, A., Schrier, G. V. D., Jones, P. D., Barichivich, J., Briffa, K. R. et Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, *4*, 17–22. <http://dx.doi.org/10.1038/nclimate2067>
- Trewavas, A. (2003). Aspects of plant intelligence. *Annals of Botany*, *92*, 1–20. <http://dx.doi.org/10.1093/aob/mcg101>

- Trumbore, S., Brando, P. et Hartmann, H. (2015). Forest health and global change. *Science*, *349*(6250), 814–818. <http://dx.doi.org/10.1126/science.aac6759>
- Urli, M., Porté, A. J., Cochard, H., Guengant, Y., Burlett, R. et Delzon, S. (2013). Xylem embolism threshold for catastrophic hydraulic failure in angiosperm trees. *Tree Physiology*, *33*, 672–683. <http://dx.doi.org/10.1093/treephys/tpt030>
- van Nes, E. et Scheffer, M. (2007). Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *The American Naturalist*, *169*(6), 738–747. <http://dx.doi.org/10.1086/516845>
- van Nes, E. H., Arani, B. M., Staal, A., van der Bolt, B., Flores, B. M., Bathiany, S. et Scheffer, M. (2016). What do you mean, ‘tipping point’? *Trends in Ecology and Evolution*, *31*(12), 902–904. <http://dx.doi.org/https://doi.org/10.1016/j.tree.2016.09.011>
- Whalley, W. R., Ober, E. S. et Jenkins, M. (2013). Measurement of the matric potential of soil water in the rhizosphere. *Journal of Experimental Botany*, *64*, 3951–3963. <http://dx.doi.org/10.1093/jxb/ert044>
- Wilhite, D. A. et Glantz, M. H. (1985). Understanding the drought phenomenon : The role of definitions. *Water International*, *10*(3), 111–120. <http://dx.doi.org/10.1080/02508068508686328>
- Wissel, C. (1984). A universal law of the characteristic return time near thresholds. *Oecologia*, *65*, 101–107.
- Yao, Y., Fu, B., Liu, Y., Li, Y., Wang, S., Zhan, T., Wang, Y. et Gao, D. (2022). Evaluation of ecosystem resilience to drought based on drought intensity and recovery time. *Agricultural and Forest Meteorology*, *314*. <http://dx.doi.org/10.1016/j.agrformet.2022.108809>