

Diplôme d'ingénieur d'État

Projet de fin d'études
Filière météorologie
Promotion 2021

Estimation de l'épaisseur de la neige et de l'équivalent en eau sur la terre à partir des micro-ondes par satellite

Réalisé par :

CHAABANI Hamid

Encadrants :

Encadrants internes : Dr. Noureddine SEMANE EHTP
Dr. Salem NAFIRI EHTP
Encadrante externe : Dr. Catherine PRIGENT CNRS / LERMA Observatoire de Paris

Organisme d'accueil :



Laboratoire d'Étude du Rayonnement et de la Matière en Astrophysique

L'observatoire de paris | PSL
61 avenue de l'Observatoire 75014 Paris.
Tel : 0140512379

Année universitaire 2020/2021

TABLE DES MATIÈRES

Table des figures	iv
Liste des tableaux	vii
Table des acronymes	viii
Table des symboles	ix
Dédicaces	x
Remerciements	xii
Résumé	xiii
Abstract	xv
Introduction générale	xvii
Chapitre 1 :	
Cadre général du projet	20
1.1 Introduction	20
1.2 Présentation de l'organisme d'accueil	20
1.3 Cadre général du projet	23
1.3.1 Contexte du projet	23
1.3.2 Motivations et problématique	23
1.3.3 Objectif général	25
1.3.4 Objectifs spécifiques	25
1.4 Conduite de projet	26
1.4.1 Diagramme de Gantt pour le projet	26
1.4.2 Equipe de réalisation	27
1.5 Conclusion	27
Chapitre 2 :	
Données, outils et méthodes	28
2.1 Introduction	28
2.2 Outils et méthodes	28

2.2.1	Méthodes utilisées	28
2.2.1	Outils utilisés	30
2.3	Compréhension des données	32
2.3.1	Paramètres de la neige sur terre (SWE, SD, RSN, SC)	32
2.3.2	Tempratures de brillance micro-ondes	37
2.4	Récupération des données	40
2.4.1	Source des données de la neige	40
2.4.2	Source des données satellitaires	41
2.5	Conclusion	43

Chapitre 3 :

Généralités sur la neige et le rayonnement		44
3.1	Introduction	44
3.2	La neige	44
3.2.1	Formation de la neige	44
3.2.2	Les types de la couverture neigeuse	47
3.2.3	Répartition géographique de la neige sur la terre	48
3.2.4	Propriétés de la neige	52
3.2.5	Métamorphisme saisonnier des manteaux neigeux	53
3.2.6	Importance de la neige pour le contrôle du climat et son impact sur l'atmosphère	55
3.2.7	Importance de la neige et de son estimation pour les applications hydrologiques .	56
3.3	Interaction de la neige avec le rayonnement	57
3.3.1	Domaine de l'optique	57
3.3.2	Domaine des micro-ondes	57
3.4	Méthode d'inversion avec les micro-ondes passives issues du capteur AMSR2	60
3.5	Conclusion	62

Chapitre 4 :

Prétraitement des données et la sélection des variables potentielles		64
4.1	Introduction	64
4.2	Prétraitement des données	64
4.2.1	Pourquoi nettoyer les bases de données?	64
4.2.2	Prétraitement des données AMSR2	64
4.2.3	Prétraitement des données ERA5	65
4.2.4	Mise des données ERA5 et les observations MOP sur une même grille	65
4.3	Sélection des variables potentielles	66
4.4	Conclusion	68

Chapitre 5 :

Quelques techniques du machine learning (ML)		69
5.1	Introduction	69
5.2	Généralités sur le Machine Learning (ML)	69
5.3	Algorithmes de l'apprentissage supervisé utilisés dans la présente étude	72
5.3.1	Régression Linéaire Multiple (RLM)	72
5.3.2	Apprentissage profond : les réseau de neurones (RN)	73
5.4	Évaluation des modèles	85
5.4.1	Erreur absolue moyenne (MAE)	85
5.4.2	Èrreur quadratique moyenne (MSE)	85
5.4.3	Racine de l'erreur quadratique moyenne (RMSE)	86
5.4.4	Coefficient de détermination (R^2)	86
5.4.5	Coefficient de corrélation (r)	86
5.4.6	Analyse graphique de sensibilité du modèle	87
5.5	Généralisation, sur-ajustement et sous-ajustement	87
5.6	Conclusion	88

Chapitre 6 :

La modélisation de SWE sur terre à partir des températures de brillance	89
6.1 Introduction	89
6.2 Modélisation de SWE avec la RLM	89
6.2.1 Modélisation de SWE à partir des Tbs micro-ondes issues de AMSR2	89
6.2.2 Modélisation de SWE à partir des Tbs micro-ondes qui seront issues du radiomètre CIMR	94
6.3 Modélisation de SWE avec les réseaux de neurones	98
6.3.1 Modèles obtenus en utilisant les températures de brillance de tous les canaux (Tb6h/v, Tb10h/v, Tb18h/v, Tb23h/v, Tb36h/v et Tb89h/V.)	99
6.3.2 Modélisation de SWE en utilisant seulement les températures de brillance suivantes : Tb6h/v, Tb10h/v, Tb18h/v et Tb36h/v.	109
7 Conclusion	109
Conclusion et perspectives	110
Annexes	118
Les températures de brillance pour le premier novembre avec une projection polaire	118
Mesure de la qualité d'apprentissage avec les deux modèles DNN	120
Premier modèle	120
Deuxième modèle	122

TABLE DES FIGURES

1	Image de l'observatoire de paris	20
2	Organigramme de l'OP le 1er décembre 2020	22
3	Evolution de la température moyenne mondiale et la température moyenne en arctique entre 1950 et 2018	24
4	Diagramme de gantt pour la planification du projet de stage	27
5	Méthodologie et cycle de vie CRISP	30
6	SD pour la date 2019-02-15	32
7	RSN pour la date 2019-02-15	33
8	Histogramme de la densité de la neige sur terre pendant la période hivernale (11/2018, 12/2018, 01/2019, 02/201 et 03/2019) de l'année 2018-2019 dans l'hémisphère nord à partir de la latitude 30° avec et sans le Groenland	34
9	SWE sur la terre pour la date 2019-02-15	35
10	Histogramme de la densité de SWE sur terre pendant la période hivernale (11/2018, 12/2018, 01/2019, 02/201 et 03/2019) de l'année 2018-2019 dans le l'hémisphère nord à partir de la latitude 30°, pour l'histogramme à droite, tout le domaine géographique est inclus, pour celui à gauche, les régions délimitées dans la carte ci-dessous sont filtrées	36
11	Les régions filtrées pour produire l'histogramme à droite de la figure ci-dessus en utilisant l'image satellitaire issue de google map prise le 15 mai 2021	36
12	Histogramme de la distribution de SWE (en m) sur terre pendant la période hivernale (11/2018, 12/2018, 01/2019, 02/201 et 03/2019) de l'année 2018-2019 dans le l'hémisphère nord à partir de la latitude 30° en se limitant aux valeurs inférieures à 0.4 m	37
13	La moyenne de Tb sur les cinq premiers jours de décembre 2018 dans la bande de fréquence centrée sur 10 et 36 GHz polarisation horizontale et verticale dans cet ordre : tb10h, Tb10v, Tb36h, Tb36v	39
14	Les types de cristaux de neige dans les nuages selon la température	47
15	L'évolution de la couverture neigeuse sur la terre pour l'année 2018-2019	49
16	L'évolution de l'épaisseur des manteaux neigeux sur la terre pour l'année 2018-2019	51
17	Opacité de l'atmosphère en fonction de longueur d'ondes dans les hyperfréquences	59
18	La matrice de corrélation entre les variables élaborée sur l'ensemble de la base de données en filtrant le Groenland et les observations où SWE est inférieur à 0.001 m	67
19	Les trois types d'apprentissage automatique [47]	71
20	Illustration de la régression et la classification	71

21	Illustration simple d'un neurone biologique et artificiel	74
22	Descente de gradient avec taux d'apprentissage élevé (à droit) et taux d'apprentissage faible (à gauche)	79
23	La courbe de dispersion du modèle RLM développé pour la restitution de SWE à partir des Tbs micro-ondes AMSR2	90
24	NUAGE DE POINTS PRÉDITS-OBSERVÉS DU MODÈLE RLM DÉVELOPPÉ POUR LA RESTITUTION DE SWE À PARTIR DES TBS MICRO-ONDÉS AMSR2	91
25	DIFFÉRENCES ENTRE LES VALEURS PRÉDITES ET LES VALEURS ISSUES DES RÉANALYSES ERA5 EN FONCTION DES VALEURS DE SWE ISSUES DES RÉANALYSES ERA5	92
26	COMPARAISON ENTRE LE SWE DE ERA5 ET LE SWE PRÉDIT À L'AIDE DE LA RLM EN UTILISANT LES DONNÉES JOURNALIÈRES À TROIS JOURS DE LA PÉRIODE D'ÉTUDE	93
27	Courbe de dispersion du modèle RLM développé à partir des données AMSR pour la restitution de SWE à partir des Tbs micro-ondes CIMR	95
28	NUAGE DE POINTS PRÉDITS-OBSERVÉS DU MODÈLE RLM DÉVELOPPE À PARTIR DES TBS MICRO-ONDÉS CIMR POUR LA RESTITUTION DE SWE À PARTIR DES TBS MICRO-ONDÉS CIMR	96
29	DIFFÉRENCES ENTRE LES VALEURS PRÉDITES ET LES VALEURS ISSUES DES RÉANALYSES ERA5 EN FONCTION DES VALEURS SWE ISSUES DES RÉANALYSES ERA5	97
30	Courbe de dispersion du premier modèle de DNN (M1) en prenant en considération toutes les Tbs (les douze Tbs)	100
31	NUAGE DE POINTS PRÉDITS-OBSERVÉS AVEC LE MODÈLE M1 DU DNN EN UTILISANT TOUTES LES TBS	100
32	LA DENSITÉ DES NOYAUX DES DIFFÉRENCES ENTRE LES VALEURS PRÉDITES ET LES VALEURS ISSUES DE ERA5 DE SWE EN FONCTION DES VALEURS DE SWE ISSUES DE ERA5	101
33	Visualisation de la variabilité spatiale de l'erreur d'estimation de SWE par le premier modèle de DNN à trois dates de la période d'étude	102
34	NUAGE DE POINTS DE SWE DE ERA5 ET LE SWE PRÉDIT PAR LE MODÈLE M1 AVEC LA LIGNE LIMITANT LES VALEURS FILTRÉES DE LA BASE DE DONNÉES POUR ÉLABORER CE DEUXIÈME MODÈLE M2	103
35	La courbe de dispersion du deuxième modèle de DNN (M2) en prenant en considération toutes les Tbs	104
36	NUAGE DE POINTS PRÉDITS-OBSERVÉS AVEC LE MODÈLE M2 DU DNN EN PRENANT EN CONSIDÉRATION TOUTES LES TBS	104
37	LA DENSITÉ DES NOYAUX DES DIFFÉRENCES ENTRE LES VALEURS PRÉDITES ET LES VALEURS DE SWE DE ERA5 EN FONCTION DE SWE DE ERA5	105
38	Visualisation de la variabilité spatiale de l'erreur d'estimation de SWE par le deuxième modèle de DNN à trois dates de la période d'étude	106
39	Comparaison des performances des deux modèles DNN à travers des cartes d'erreurs élaborées pour trois dates de la période d'étude, les cartes à gauche représentent les erreurs du premier modèle DNN (M1), les cartes à droite représentent les erreurs du deuxième modèle (M2)	108
40	Les températures de brillance tb6h/v et tb10h/v au premier novembre 2018	118
41	Les températures de brillance tb18h/v, tb23h/v et tb36h/v au premier novembre 2018	119
42	Les températures de brillance tb89h/v au premier novembre 2018	120
43	Progression de la fonction perte aussi bien sur l'ensemble d'apprentissage (LOSS) que sur l'ensemble de validation (VAL LOSS) en fonction des itérations	120
44	Progression de l'erreur quadratique moyenne aussi bien sur l'ensemble d'apprentissage (MSE) que sur l'ensemble de validation (VAL MSE) en fonction des itérations	121
45	Progression de l'erreur absolue moyenne aussi bien sur l'ensemble d'apprentissage (MAE) que sur l'ensemble de validation (VAL MAE) en fonction des itérations	121
46	Progression de la fonction perte aussi bien sur l'ensemble d'apprentissage (LOSS) que sur l'ensemble de validation (VAL LOSS) en fonction des itérations pour le deuxième modèle	122
47	Progression de l'erreur quadratique moyenne aussi bien sur l'ensemble d'apprentissage (MSE) que sur l'ensemble de validation (VAL MSE) en fonction des itérations pour le deuxième modèle	122

48 Progression de l'erreur absolue moyenne aussi bien sur l'ensemble d'apprentissage (MAE) que sur l'ensemble de validation (VAL_MAE) en fonction des itérations pour le deuxième modèle	123
--	-----

LISTE DES TABLEAUX

1	Equipe de suivi et de réalisation	27
2	Les étapes de la méthode CRISP pour la présente étude	30
3	Les principales caractéristiques du radiomètre AMSR2 [7]	61
4	Description statistique brève de la base de données	66
5	Analogie entre un neurone artificiel et biologique	75
6	Les fonctions d'activation les plus utilisées pour les neurones cachés et les neurones de sortie [56]	76
7	Algorithme général de la descente de gradient	78
8	Algorithme de la descente de gradient pour l'apprentissage d'un réseau de neurones par lot	81
9	Les valeurs prises par les paramètres du premier modèle	99

TABLE DES ACRONYMES

<i>Abréviation</i>	<i>Signification</i>
Adam	Adaptive Moment Optimization
AMSR2	Advanced Microwave Scanning Radiometer 2 (radiomètre à balayage micro-ondes avancé)
CIMR	Copernicus Imager Microwave Radiometer
CNRS	Centre National de Recherche Scientifique
CRISP	CRoss-Industry Standard Process
DNN	Deep Neural Network (réseau de neurones profond)
ECMWF	European Centre for Medium-Range Weather Forecasts (centre européen pour les prévisions météorologiques à moyen terme)
ERA5	ECMWF Reanalysis 5th Generation
ESA	European Space Agency (agence Spatiale Européenne)
GCOM-W1	The Global Change Observation Mission 1st-Water
JAXA	Japan Aerospace eXploration Agency (agence japonaise d'exploration aérospatiale)
LERMA	Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères
ML	Machine Learning
MLR	Multiple Linear Regression (régression linéaire multiple)
MOP	Micro-Ondes Passives
MSR	Microwave Sensor Radiometer (radiomètre à balayage hyperfréquence)
NN	Neural Network (réseau de neurones)
OP	Observatoire de Paris
PSL	Université de Paris Sciences & Lettres
RL	Régression Linéaire
RMSProp	Root Mean Square Propagation
RSN	La densité de la neige
SC	Snow Cover (couverture neigeuse)
SD	Snow Depth (épaisseur de la neige)
SMMR	Scanning Multichannel Microwave Radiometer
SSM/I	Special Sensor Microwave/Imager
SWE	Snow Water Equivalent (équivalent en eau de la neige)
Tb(s)	Température de brillance (Températures de brillance)

TABLE DES SYMBOLES

<i>Symbol</i>	<i>Signification</i>
α	Albédo
$e^{-\tau}$	La transmissivité atmosphérique
Tb_{surf}	La température de brillance de l'émission de surface
R	La réflectivité de la surface
ρ_{eau}	La densité de l'eau
ν	La fréquence électromagnétique
p	La polarisation
H	La polarisation horizontale
V	La polarisation verticale
T	La température physique
e	L'émissivité
K	Kelvin
\hat{y}	La variable estimée
ϵ	Le résidu
∇	Le gradient
J	La fonction coût

DÉDICACES

*Au nom du Dieu clément et miséricordieux,
Je dédie cet humble travail :*

**À ma Chère Mère Mme. Fatima ALAOUI,
À mon Cher Père M. Mouhamed CHaabani,**

Qui n'ont jamais cessé de me soutenir, de m'encourager et de prier à mon égard. Mes Chers Parents qui ont toujours été là tout au long de mes études pour me consoler quand il fallait. Leurs prières et leur bénédiction m'ont été d'un grand secours pour mener à bien mes études. En ce jour mémorable pour nous, ce travail reçoit en signe de ma vive reconnaissance et ma profonde estime et je remercie dieu, le tout-puissant, de m'avoir donné la force et la patience d'accomplir cet humble travail dans lequel vous trouvez ma reconnaissance et tout mon amour. Que Dieu vous bénisse, vous préserve et vous accorde santé, longue vie et bonheur. Je vous aime énormément.

À mes Chères Soeurs Samira et Somia,

De partager les souvenirs d'une enfance pleine de moments agréables et pour le petit frère gâté que j'ai toujours été, le petit frère qui adorait, chamaillait, ignorait, jalouxait et imitait. Je n'oublierai jamais votre soutien, vos sacrifices et vos encouragements tout au long de mes études. Je vous dédie ce travail avec tous mes vœux de bonheur et je vous souhaite une vie pleine de bonheur et de succès et que Dieu, le tout-puissant, vous protège et vous garde. Je vous aime.

À mes Oncles Brahim, El Houssain, Ismail ...
*Pour leurs encouragements et les moments agréables et inoubliables
que nous avons partagés ensemble.*

**À mes Chers Amis Abdelaziz CHAQDID, Aki BOUAZDAY,
Abdelaali EL MAJDOUBI, Samir BOUTBOUT,**
*Ceux qui m'ont appris le vrai sens de l'amitié. Je vous dédie ce modeste
travail. Merci mes amis de toujours.*

À mes Chers Condisciples,
*Ceux qui m'ont appris la signifiance du partage et le goût d'un amour
sain. Ceux qui ont su me porter conseil quand j'en avais le plus grand
besoin. Ceux avec lesquels j'ai partagé des moments inoubliables.*

**À mes Chers Encadrants Mme. Catherine PRIGENT, M.
Noureddine SEMAN et M. Salem NAFIRI.**

**À mes Chers Professeurs M. Moustafa EL HAMLI, M. Driss
BARI, M. Rachid ILMEN ...**

À mes Chers Neveux Walid et Youssef.

Enfin, à tous les membres de ma famille.

CHAABANI Hamid

REMERCIEMENTS

Me voilà à l'étape d'écriture du remerciement, qui marque l'accomplissement de ce projet. Il est temps de remercier toutes les personnes qui m'ont aidée et je vais essayer de ne pas oublier des gens. Pour cela, je m'en excuse d'avance.

*Sans les personnes qui entouraient ce travail, il n'aurait aucun sens. Avant tout, je voudrais infiniment remercier mon encadrante externe **Mme Catherine PRIGENT** de m'avoir donné la chance d'entamer ce projet, pour la confiance qu'elle m'a accordée, pour son écoute, ses aiguillages judicieux, sa sympathie et ses sourires. C'est grâce à elle que j'ai appris à faire un travail de recherche scientifique. Elle m'a appris également qu'on ne finit jamais d'apprendre. Merci énormément **Mme. Catherine**.*

*Je tiens à remercier un monsieur que cet amas d'expressions ne me permet sincèrement pas d'y exprimer toute ma gratitude et ma reconnaissance, **M. Noureddine SEMANE** je vous remercie énormément pour l'aide que vous m'avez apporté dans ce stage pour surmonter toutes les difficultés. Je vous remercie non seulement pour votre encadrement lors du stage, mais encore pour votre accompagnement, vos conseils judicieux et votre exceptionnelle pédagogie d'enseignement tout au long de ces trois années, merci beaucoup. Je tiens également à remercier profondément mon encadrant interne **M. Salem NAFIRI** pour tous leurs précieux conseils durant les différentes réunions.*

*Ce travail n'a jamais été isolé. Il était toujours entouré de nombreuses personnes. Pour cela, je voudrais remercier l'équipe de recherche au département de LERMA de l'observatoire de paris **Catherine, Lise, Soriot, Carlos, Samuel, Filipe, Lan, Victor, Heberger, Karim** ... pour toutes les discussions scientifiques et techniques qui se sont déroulées dans les réunions hebdomadaires : Café Télédétection, et qui ont été très enrichissantes.*

Je tiens à exprimer encore ma gratitude à la direction et à l'ensemble du corps enseignant de notre chère école l'école Hassania des Travaux Publics (EHTP).

Enfin, je tiens à remercier toute personne qui, de près ou de loin, a contribué à l'aboutissement de ce projet et tous ceux et celles qui feuilleront ces pages.

RÉSUMÉ

La variation spatio-temporelle de la neige sur terre a un impact direct ou indirect sur plusieurs secteurs socio-économiques dans les moyennes et les hautes latitudes, notamment l'agriculture, l'approvisionnement en eau, la production de l'énergie hydraulique, le tourisme, etc. En plus, la couverture nivale affecte considérablement le climat à différentes échelles spatio-temporelles. Ainsi, une bonne estimation des propriétés de la neige sur terre est nécessaire, non seulement pour l'étude de la neige sur terre, mais encore en tant que l'une des entrées incontournables pour les modèles météorologiques de prévision du temps et les modèles hydrologiques en vue d'améliorer les prévisions de différents paramètres dépendants directement ou indirectement de la présence de la neige sur terre. Pour estimer les propriétés de la neige sur terre, les observations satellites dans le domaine des micro-ondes passives ont révélé un certain potentiel en complément des mesures *in situ*. Par conséquent, dans le but de répondre à la problématique concernant la modélisation des propriétés de la neige sur terre, notamment l'équivalent en eau de la neige et son épaisseur, à partir des observations micro-ondes passives issues du radiomètre à balayage micro-ondes avancé (en anglais : Advanced Microwave Scanning Radiometer 2), la présente étude propose une méthodologie basée sur l'approche Machine Learning, notamment la régression linéaire multiple et l'apprentissage profond qui servira à l'estimation de l'équivalent en eau de la neige sur terre et son épaisseur à partir d'observations micro-ondes passives entre 6 et 89 GHz.

Dans le but de déterminer la relation entre les températures de brillance micro-ondes passives et l'équivalent en eau de la neige, nous avons testé d'abord la régression linéaire multiple, en utilisant d'une part, les températures de brillance tb6h/v, tb10h/v, tb18h/v et tb36h/v comme variables d'entrée du modèle, et de l'autre part les températures de brillance tb6h/v, tb10h/v, tb18h/v, tb23h/V, tb36h/V et tb89h/v. Une validation croisée a été utilisée pour sélectionner le score le plus optimal et vérifier la pertinence des modèles. Toutefois, les scores obtenus avec les deux combinaisons des variables ont été très faibles faisant de la régression linéaire multiple une méthode trop limitée en terme de modélisation de l'équivalent en eau de la neige sur terre à l'échelle hémisphérique.

Les méthodes d'apprentissage profond ont eu un apport considérable en comparaison avec la régression linéaire multiple. En effet, en testant plusieurs combinaisons de paramètres d'un deep neural network (DNN), le coefficient de détermination en utilisant toutes les Tbs est passé de 0.13 à 0.46. En utilisant seulement les températures qui seront issues du futur radiomètre Copernicus Imager Microwave Radiometer (CIMR), le coefficient de détermination est passé de 0.05 à 0.3.

Malgré l'amélioration des performances des modèles en utilisant l'apprentissage profond pour estimer l'équivalent en eau de la neige à l'échelle globale, les résultats sont toujours insatisfaisants. Cela est peut-être à la dépendance du signal micro-ondes avec les conditions locales. En effet, l'analyse bibliographique et la visualisation des erreurs de l'estimation de l'équivalent en eau de la neige avec les modèles élaborés a montré bel et bien la sensibilité du signal micro-onde aux autres conditions, notamment la présence d'eau de surface, de couverture forestière, de terrains complexes, des accumulations importantes de neige, du givre en profondeur ou de l'humidité de la neige. Ces problèmes peuvent être traités soit en filtrant les zones dans lesquelles l'une ou plusieurs de ces conditions locales règnent, soit en introduisant dans l'entrée des modèles les paramètres qui affectent le signal micro-onde pour que le modèle puisse les prendre en compte. Dans la présente étude, la seule condition locale qui a été prise en considération pour l'instant est les accumulations importantes de la neige qui ont été filtrées en supprimant toutes les observations dont le SWE dépasse 0.5 m avant d'exploiter la base de données pour faire de la modélisation.

En ce qui concerne la modélisation de l'épaisseur de la neige (SD), il peut être simplement déduit à partir de SWE connaissant la densité de la neige qui peut être récupérée à partir des réanalyses ERA5 avec une grande précision.

Mots clés : La neige sur terre, l'équivalent en eau de la neige (SWE), AMSR2, les micro-ondes passives, les températures de brillance, la densité de la neige, l'épaisseur de la neige (SD), le réseau de neurones dense (DNN), le Machine Learning (ML), la régression linéaire multiple (RLM).

ABSTRACT

Snow variation over land has a direct or indirect impact on several socio-economic sectors in the middle and high latitudes, including agriculture, water supply, hydropower production, tourism, etc. In addition, snow cover significantly affects climate at different spatial and temporal scales. Thus, a good estimation of the properties of snow on land is necessary, not only for the study of snow on land, but also as one of the essential inputs for meteorological weather models and hydrological models to improve the prediction of different parameters dependent on the presence of snow over land. To estimate the properties of snow over land, several studies have conducted this problem using MOP observations instead of in situ observations or any other types of observations. Therefore, in order to address the problem of modeling snow properties over land, especially SWE and SD, from MOP observations from the AMSR2 radiometer, the present study proposes a methodology based on Machine Learning approach, especially MLR and deep learning that will be used to estimate SWE and SD over land from brightness temperatures corresponding to some microwave frequencies.

In order to determine the relationship between passive microwave brightness temperatures and SWE, we tested the MLR, using firstly tb6h/v, tb10h/v, tb18h/v, tb23h/V, tb36h/V and tb89h/v brightness temperatures as model input variables, then tb6h/v, tb10h/v, tb18h/v and tb36h/V brightness temperatures as input variables. A cross-validation was used to select the most optimal score and to check the relevance of the models. However, the scores obtained with the two combinations of variables were very low, making MLR a too limited method in terms of SWE modeling over land at the hemispheric scale.

Deep learning methods had a considerable contribution in comparison with MLR. Indeed, by testing several combinations of parameters of a DNN, the coefficient of determination using all Tbs increased from 0.13 to 0.46. Using only the Tbs that will be derived from the future CIMR radiometer, the coefficient of determination increased from 0.05 to 0.3.

Despite the improved performance of the models using deep learning to esti-

mate the SWE at the global scale, the results are still unsatisfactory. This is perhaps due to the dependence of the microwave signal on local conditions. Indeed, the literature review and the visualization of the errors of the SWE estimation with the developed models have shown the sensitivity of the microwave signal to other conditions, notably the presence of water bodies, forest cover, complex terrain, urban areas, large snow accumulations, deep frost and snow moisture. These problems can be addressed either by filtering out areas where one or more of these parameters prevail, or by introducing into the model input the parameters that affect the microwave signal so that the model can take them into account. In the present study, the only condition that was considered was significant snow accumulations which were filtered out by removing all observations with SWE above 0.5 m before mining the database for modeling.

Regarding the modeling of snow depth (SD), it can be simply deduced from SWE knowing the snow density that can be retrieved from ERA5 reanalysis.

Key words : Snow over land, snow water equivalent (SWE), AMSR2, passive microwaves, brightness temperatures, snow density, snow depth (SD), Dense Neural Network (DNN), Machine Learning (ML), multiple lineaire regression (MLR).

INTRODUCTION GÉNÉRALE

Les dernières années, la terre assiste à un réchauffement planétaire de son atmosphère avec des intensités différentes d'une région à une autre. Un tel réchauffement est accompagné d'une perturbation du climat mondiale. Ainsi, la variabilité spatio-temporelle de plusieurs paramètres (la température, les précipitations, le vent, les couvertures nuageuses, les écoulements, les propriétés de la neige sur terre, etc.) devient de plus en plus grande. L'observation de ces paramètres à l'échelle planétaire est désormais compliquée à cause de leur grande variabilité spatio-temporelle. Bien que les observations *in situ* des propriétés de la neige sur terre ont été améliorées afin de répondre aux applications météorologiques, climatiques, ou hydrologiques, ces observations restent insuffisantes à cause de leur rareté voir même leur absence dans certaines régions difficiles d'accès comme les pôles et les reliefs accidentés. Pour couvrir les lacunes des observations *in situ*, la télédétection spatiale de la terre a été adoptée et de nombreux satellites ont été lancés pour observer la terre à partir de l'espace. Les satellites d'observation de la terre peuvent être classifiés en deux grandes familles : les satellites polaires et les satellites géostationnaires. Les satellites polaires (nommés parfois satellites défilants) sont des satellites ayant une trajectoire polaire (ou légèrement oblique) autour de la terre leur permettant d'observer presque la totalité de la terre grâce à la combinaison d'un nombre de trajectoires successives en une période de temps qui dépend étroitement de la vitesse de rotation du satellite et la largeur spatiale de la fauchée. Les satellites polaires ont la particularité d'observer les régions des hautes latitudes (latitude supérieure à 60° ou les régions polaires) plusieurs fois en une période de 24 heures faisant des observations issues des satellites défilants les observations les plus adaptées et les plus adéquates aux études des différents paramètres dans ces régions, d'où vient le nom satellite polaire. Les satellites géostationnaires sont des satellites qui tournent avec la terre (leur vitesse par rapport à la terre est nulle.). Ils observent en permanence une région bien déterminée. Cette famille de satellite a comme avantage par rapport à la famille des satellites polaires la résolution temporelle importante. Toutefois, ils ne permettent pas une couverture globale de la terre. Sur chaque satellite, des instruments sont embarqués, ceux-ci permettent de capter les différents signaux électromagnétiques l'aide de différents canaux qui captent le signal dans différentes bandes spectrales centrées sur des

fréquences bien choisies en utilisant plusieurs critères. Ainsi, avec la télédétection spatiale, notamment la télédétection micro-onde passive, on va pouvoir restituer, à différentes échelles, de nombreux paramètres, à savoir les paramètres de la neige sur terre, les paramètres de végétation, les paramètres de la surface, etc.

Faisant partie des composantes de la cryosphère, la neige sur terre est au centre d'intérêt de plusieurs unités de recherche, notamment dans le continent eurasiatique et l'Amérique du Nord. En effet, plus de 80 % des terres en Eurasie et en Amérique du Nord sont recouvertes de neige au cœur de la période hivernale. Environ 46,5 millions de km² sont couverts de neige dans l'hémisphère nord en janvier, et 3,8 millions de km² même en août [27]. Son extension importante fait de la neige sur terre une composante influente du climat à l'échelle locale et globale et impacte même l'activité socio-économique de certains pays de moyennes et hautes latitudes. La neige sur terre a subi beaucoup de changements qui auraient pu résulter de changements dans les précipitations hivernales [15], ce qui rend l'estimation de ses paramètres une tâche délicate. En plus, la neige sur terre subit des mutations à cause de la rétroaction positive température-albédo. En effet, toute diminution de la couverture nivale qui résulte du réchauffement entraîne une augmentation de l'absorption de chaleur supplémentaire pour faire fondre encore la neige [3].

Malgré le développement des instruments d'observation de la neige sur terre, l'estimation de ses propriétés est encore compliquée en présence de montagnes et quand la neige est humide ou profonde. Ainsi, la sensibilité des observations satellitaires aux caractéristiques de la neige dépend des conditions locales (conditions météorologiques, végétation, topographie, etc.) [6].

Étant donné l'importance de la neige sur terre et de son estimation dans plusieurs applications, l'estimation de ses caractéristiques à partir des observations satellitaire à l'échelle globale reste encore une problématique importante, motivante et délicate en même temps. Par exemple, une estimation précise de l'équivalent en eau de la neige est cruciale pour l'approvisionnement en eau, la production d'électricité [20], etc. En plus, la neige sur terre est une composante essentielle de l'écosystème de la terre qu'il faut surveiller avec une grande précision.

Par conséquence, notre capacité à estimer les propriétés de la neige sur terre impacte notre capacité à surveiller le climat et le changement climatique [11] ainsi que notre capacité à gérer et à anticiper la production de l'énergie hydraulique et l'approvisionnement en eau douce. L'estimation des propriétés de la neige sur terre est également essentielle comme données d'entrée pour les modèles numé-

riques de la prévision opérationnelle et les modèles climatiques.

Dans ce sens, la présente étude a comme objectif principal l'étude de la sensibilité des micro-ondes passives issues du radiomètre AMSR2 embarqué au bord du satellite GCOM-W1 aux propriétés de la neige sur terre dans l'hémisphère nord pour l'année 2018-2019 en vue de développer un algorithme de restitution de l'équivalent en eau de la neige et de son épaisseur basé sur les méthodes du Machine Learning (ML) (notamment la régression linéaire multiple et l'apprentissage profond) à partir des résultats des modèles de la prévision numérique du temps (PNT) du centre européen pour les prévisions météorologiques à moyen terme (CEPMMT) et les données de températures de brillance micro-ondes passives issues du radiomètre AMSR2 pour préparer à une nouvelle mission (actuellement en étude) pour la prochaine génération de missions satellitaires CIMR (Copernicus Imager Microwaves radiometer).

Dans la perspective de comprendre au fond la problématique précitée, la présente mémoire a été subdivisée en six chapitres :

- **Chapitre 1** : ce premier chapitre parle en détail du cadre général du projet, à savoir le contexte, les motivations et la problématique, l'objectif général, etc.
- **Chapitre 2** : dans ce chapitre, nous listons les principaux outils et méthodes utilisés ainsi que les données exploitées pour mener à bien ce projet.
- **Chapitre 3** : ce chapitre est plus théorique que pratique puisque nous parlons en détails de la physique de la neige et sa télédétection spatiale en se référant à un amas d'articles scientifiques.
- **Chapitre 4** : ce chapitre présente les principaux prétraitements des données réalisés en vue d'obtenir une base de données exploitable. Il détaille également les principales étapes menées pour sélectionner les variables potentielles.
- **Chapitre 5** : un autre chapitre théorique dans lequel nous détaillons les principales techniques du Machine Learning servant à la présente étude.
- **Chapitre 6** : dans ce dernier chapitre, nous listons quelques résultats de notre étude portant sur la modélisation de l'équivalent en eau de la neige sur terre à partir des observations micro-ondes.

CHAPITRE 1 : CADRE GÉNÉRAL DU PROJET

1.1 Introduction

Cette partie est très importante pour un démarrage efficace du projet, partant de la contextualisation jusqu'à l'étape de formulation de l'objectif général. En effet, elle permet de se mettre dans le contexte du sujet et avoir une idée claire et concise du déroulement du projet avant d'entamer la réalisation de celui-ci. Ainsi, ce chapitre donnera premièrement un aperçu général sur l'organisme d'accueil. Puis, il parlera en détail du contexte, des motivations et des objectifs du projet. Enfin, nous évoquerons succinctement la conduite de ce projet.

1.2 Présentation de l'organisme d'accueil

Le début de l'observatoire de Paris remonte à l'année 1667, quand Colbert a acheté, au nom du roi Louis XIV, le terrain réservé à la construction des premières infrastructures de l'observatoire [2].



FIGURE 1 – Image de l'observatoire de paris

L'observatoire de Paris - PSL est un grand établissement appartenant au ministère de l'enseignement supérieur, de la recherche et de l'innovation. C'est le plus grand pôle national de recherche en astronomie [1].

L'Observatoire de Paris (OP) est composé de cinq départements scientifiques, deux services scientifiques, auxquels s'ajoutent des services communs et un institut. Il est déployé sur trois instituts, Paris, Meudon et Nançay. Les différentes entités scientifiques travaillent en coopération avec le CNRS et les grandes universités françaises. L'OP est récemment devenu un établissement composant de l'Université PSL [1].

L'OP a pour missions principales : la formation, la recherche et la diffusion de connaissance. Les activités de recherche au sein de l'OP sont liées à plusieurs champs, à savoir, l'astronomie et l'astrophysique contemporaines, la métrologie du temps et des fréquences. En plus, L'OP s'intéresse aussi aux sciences de la Terre, en coopération avec de grandes entités de recherche [1].

En plus de la recherche, la formation et la diffusion de connaissance, l'OP joue un rôle important dans la conception et la réalisation des grands instruments de l'observation de la terre et de l'espace et concourt aux grands relevés, aux grandes simulations, à l'observatoire virtuel et conduit des travaux en laboratoire [1].

Concernant la recherche, l'OP adopte une politique de recherche qui touche plusieurs thématiques scientifiques prioritaires en relation avec l'astronomie. À ces thématiques scientifiques s'ajoutent des thématiques méthodologiques dont les principaux axes de développement sont : l'exploration, l'observation spatiale et instrumentation sol et espace, la théorie, l'astrophysique de laboratoire et simulations numériques et l'observatoire virtuel. En ce qui concerne la formation, l'OP propose des formations en Astronomie, Astrophysique et Technologies spatiales à tous les niveaux de l'enseignement supérieur, du master sciences de l'univers et technologies spatiales jusqu'au doctorat. L'établissement dispense aussi des enseignements dans deux autres Masters et deux licences de l'université PSL. Les diplômes délivrés sont des diplômes de l'université PSL. L'OP peut aussi proposer des enseignements hors cursus. Il dispose aussi de nombreuses ressources pour la diffusion de l'information aux enseignants, aux élèves et au grand public [1].

Pour bien organiser le travail de recherche et de développement au sein de l'OP, ce dernier a été divisé en cinq départements scientifiques (GEPI, LESIA, LUTH, LERMA et SYRTE) et un institut (IMCCE) [1].

Mon PFE (Projet de fin d'études) a été réalisé au sein du département LERMA, un des départements de l'OP. Ce dernier se présente comme une unité mixte de recherche commune au CNRS et à 4 établissements d'enseignement supérieur :

OP, l'École normale supérieure (ENS), l'université Pierre et Marie Curie (UPMC - Paris 6) et l'université de Cergy-Pontoise (UCP) lesquelles hébergent les différentes composantes du LERMA.

Ce département lui-même est divisé en quatre pôles, lesquels s'intéressent à des thématiques de recherche différentes. Le premier pôle est nommé **Galaxies et cosmologie**, le deuxième s'intitule **Dynamique des milieux interstellaires et plasmas stellaires**, le troisième s'appelle **Molécules dans l'Univers** et le dernier duquel mon stage faisait partie est nommé **Instrumentation et télédétection**. Ce dernier touche les thématiques de recherche suivantes : composants et sous-systèmes THz, instruments hétérodynes THz, caractérisation des atmosphères claires, nuageuses et pluvieuses, caractérisation des surfaces de la Terre, des planètes et des comètes, et enfin le traitement, l'archivage et la valorisation des données [1].

L'organigramme suivant montre compendieusement les différentes composantes de l'OP :

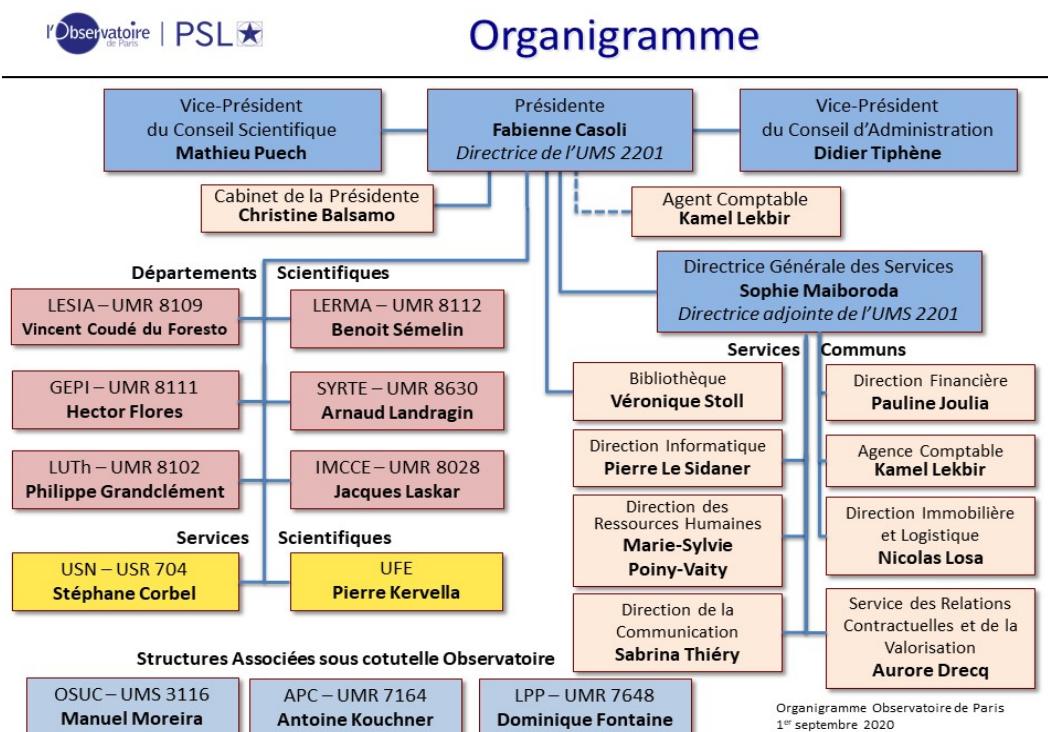


FIGURE 2 – Organigramme de l'OP le 1er décembre 2020

1.3 Cadre général du projet

1.3.1 Contexte du projet

La neige sur terre est une composante de la cryosphère caractérisée par une variabilité spatio-temporelle importante, surtout, dans les régions où la présence de la neige sur terre est restreinte à la période hivernale. Ainsi, l'estimation des principales propriétés de la neige avec une précision satisfaisante reste compliquée. Les observations in situ des propriétés de la neige sur terre souffre d'un manque de représentativité spatio-temporelle. En effet, le réseau d'observations in situ des propriétés de la neige n'est pas suffisamment dense pour avoir une bonne représentativité spatiale des paramètres de la neige sur terre. Les observations satellitaires peuvent combler les lacunes des observations in situ. En effet, elles ont une couverture globale et parfois des résolutions spatiales et temporelles convenables. Néanmoins, les données restituées des observations satellitaires nécessitent des algorithmes d'inversion qui sont souvent des méthodes empiriques ce qui engendre toujours une perte d'information. Malgré l'évolution dans les systèmes d'observation par télédétection spatiale, à savoir les capteurs à canaux micro-ondes comme AMSR2, SMM/I, SMMR, l'estimation des propriétés de la neige sur terre souffre de manque de précision surtout dans les reliefs accidentés.

Pour ce fait, plusieurs recherches ont été menées pour restituer les propriétés de la neige à partir des micro-ondes passifs par satellite en utilisant des méthodes statistiques classiques auxquelles s'ajoutent parfois des méthodes plus avancées de l'apprentissage profond. Les micro-ondes ont été utilisées, car elles sont sensibles aux propriétés de la neige et sont peu affectées par l'atmosphère.

Dans ce sens, la présente étude s'inscrit dans le cadre des projets réalisés en vue de l'amélioration de l'estimation des produits de la neige à partir des observations micro-ondes passifs du capteur AMSR2 embarqué au bord du satellite GCOM-W1, en utilisant quelques méthodes classiques du ML et l'apprentissage profond dans le cadre de la préparation à la nouvelle mission CIMR (Copernicus Imager Microwave Radiometer).

1.3.2 Motivations et problématique

L'augmentation anormale des gaz à effet de serre dans l'atmosphère à cause de la contribution anthropique, qu'appellent parfois les spécialistes des sciences du climat l'effet de serre naturel amplifié par les activités anthropiques a pro-

voqué une hausse de température moyenne mondiale, c'est le réchauffement climatique.

De ce fait, la machine climatique a été perturbée. Parmi les changements significatifs introduits dans le système terre-atmosphère, on peut citer : l'augmentation de la température de l'air dans l'arctique de plus du double de la moyenne mondiale au cours des deux dernières décennies. [3].

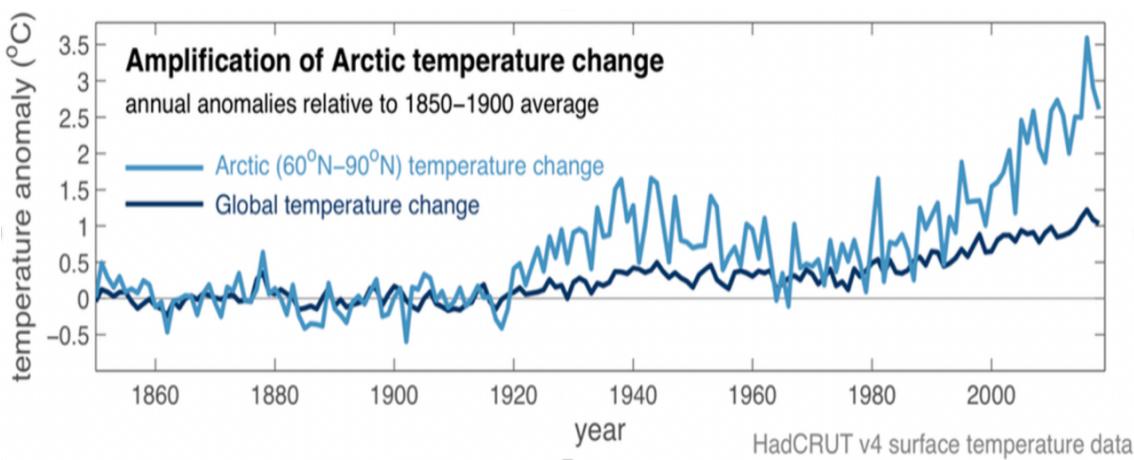


FIGURE 3 – Evolution de la température moyenne mondiale et la température moyenne en arctique entre 1950 et 2018

Ainsi, les régions polaires et les hautes latitudes subissent un réchauffement accéléré. Par conséquent, elles perdent des quantités importantes de la neige à cause de la fonte accélérée des manteaux neigeux avec tout ce qu'elle a de répercussions, non seulement sur le climat, mais aussi sur les secteurs de l'eau et de l'énergie, puisque, la présence de la neige sur terre dans les moyennes et les hautes latitudes est d'une grande importance pour le processus de potabilisation de l'eau et la production de l'énergie hydraulique, car sa fonte contribue largement aux apports en eaux de surface et les eaux souterraines pendant les périodes qui souffrent d'une rareté ou d'un manque de précipitations.

Une estimation suffisamment précise des propriétés de la neige, et surtout le SWE aura une grande importance dans plusieurs domaines, à savoir la météorologie, l'hydrologie, l'étude du cycle de l'eau et les changements globaux [42]. Elle est aussi très intéressante pour l'aménagement des ressources en eau et le contrôle des écoulements [42].

D'ailleurs, la couverture neigeuse peut affecter directement la dynamique de l'atmosphère. Ainsi, une bonne estimation de la couverture et de la masse neigeuse dans les manteaux neigeux saisonniers et permanents peut augmenter la précision de la surveillance du climat, le changement climatique et tester les si-

mulations des modèles climatiques [11]. Une bonne observation des propriétés de la neige permet aussi d'anticiper et de s'adapter aux nouvelles conditions climatiques et d'évaluer les risques comme les écoulements et les avalanches [8].

Puisque les mesures in situ de SWE sont coûteuses et non-accessibles en permanence et en temps réel, ce qui justifie le recours à quelques méthodes alternatives pour estimer le SWE à partir d'observations couvrant toute la surface terrestre et qui sont continues dans le temps [21]. En plus, à l'échelle globale, il est difficile de dévoiler une corrélation entre les mesures satellitaires et le SD (épaisseur de la neige). Autrement, il est compliqué de construire un algorithme de restitution de SD à partir des micro-ondes passifs par satellite [9].

Au mépris des efforts déployés par les chercheurs en vue de fournir une restitution précise de SWE en utilisant les micro-ondes passives par satellite, les produits SWE actuels ne répondent toujours pas aux exigences minimales des applications hydrologiques [28].

De ce fait, la problématique à relever par cette étude consiste à développer un algorithme de restitution de l'équivalent en eau de la neige à l'échelle globale basé sur les méthodes du Machine Learning à partir de données des réanalyses météorologiques ERA5 issues du modèle ECMWF et les observations satellites dans le domaine des micro-ondes.

1.3.3 Objectif général

L'objectif général de la présente étude est le développement d'un algorithme de restitution de produits de la neige, SD et SWE à partir des observations satellitaires micro-ondes passives issues du capteur AMSR2 embarqué au bord du satellite GCOM-W1 en utilisant les méthodes du Machine Learning.

1.3.4 Objectifs spécifiques

Pour mener à bien ce projet, il fallait mettre des objectifs spécifiques dans un ordre permettant d'atteindre l'objectif général dans les meilleurs délais. Ainsi, quatre objectifs spécifiques ont été fixés au début du stage dans l'ordre suivant :

1. L'étude théorique de la physique de la neige sur terre en lien avec la télé-détection satellitaire et l'étude des anciens algorithmes de restitution des produits de la neige à travers l'analyse d'une trentaine d'articles traitant la même problématique.

2. La visualisation des différents paramètres de la neige et des températures de brillance dans le domaine spectral MO pour vérifier premièrement l'intégrité des données récupérées, et étudier ensuite la répartition de la neige sur terre et la variabilité spatio-temporelle de ses différents paramètres et celle des températures de brillance.
3. L'étude de sensibilité de ces températures de brillance des différents canaux micro-ondes du radiomètre AMSR2 aux propriétés de la neige à travers des tests statistiques de corrélation en vue de sélectionner les variables potentielles à impliquer dans le modèle. Cette étape de la sélection des variables potentielles peut être effectuée également en faisant recours à l'étude des liens physiques entre la variable expliquée et les éventuelles variables explicatives ou bien la sélection des variables explicatives en se référant aux études présentes dans la littérature scientifique et qui traitent la même problématique. Dans la présente étude nous combinons ces trois méthodes pour sélectionner les variables potentielles.
4. La construction des modèles pour l'estimation de SWE et de SD dans l'hémisphère nord à partir des températures de brillances issues du radiomètre AMSR2.

1.4 Conduite de projet

1.4.1 Diagramme de Gantt pour le projet

Pour suivre l'avancement, réajuster et déplacer les étapes d'un projet, il faudra utiliser une méthode de gestion de projet, celle-ci permettra de visualiser l'état d'avancement des différentes activités et avoir une idée claire et concise sur la durée approximative que durera le projet. Ainsi, dans notre cas, nous avons utilisé le diagramme de gantt pour projeter les différentes étapes du projet sur la durée totale de celui-ci.

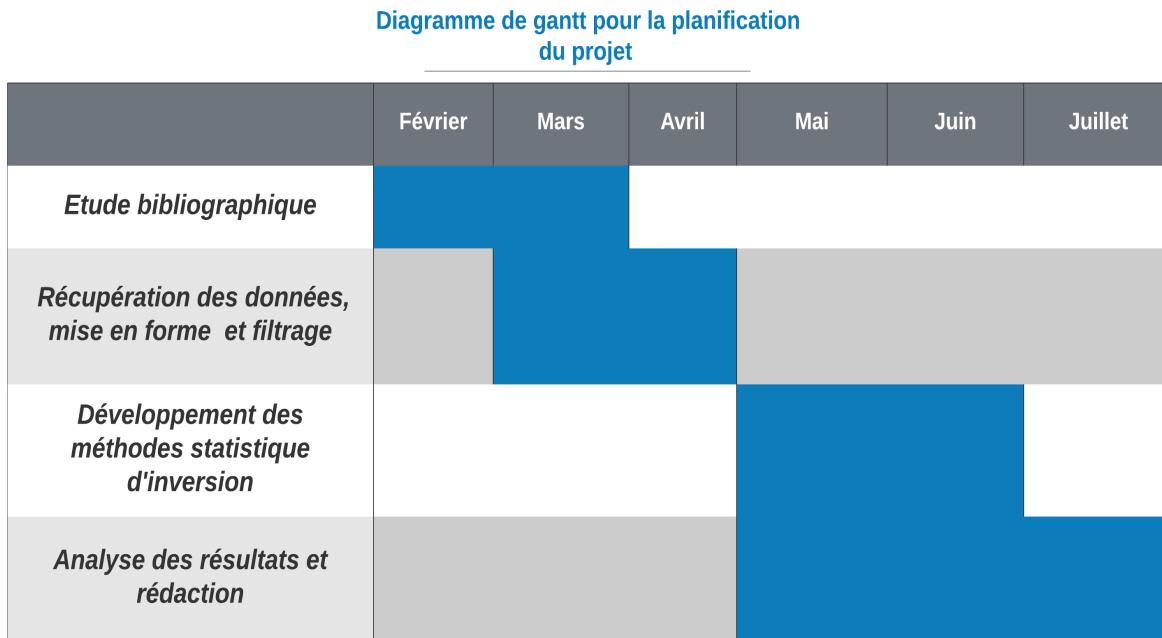


FIGURE 4 – Diagramme de gantt pour la planification du projet de stage

1.4.2 Equipe de réalisation

	Nome et prénom	Rôle
Suivi	Dr. PRIGENT Catherine (CNRS / LERMA Observatoire de Paris)	Encadrante externe et chef de projet
Suivi	Pr. SEMANE Nourredine (EHTP)	Encadrant interne
Suivi	Pr. NAFIRI Salem (EHTP)	Encadrant interne
Réalisation	CHAABANI Hamid	Stagiaire

TABLE 1 – Equipe de suivi et de réalisation

1.5 Conclusion

Dans ce chapitre, nous avons pu donner une idée globale sur les étapes de déroulement du projet en décrivant exhaustivement sa problématique, son objectif général et ses objectifs spécifiques. Dans le chapitre suivant, nous parlerons des données, des outils et des méthodes utilisées pour réaliser ce projet.

CHAPITRE 2 : **DONNÉES, OUTILS ET MÉTHODES**

2.1 Introduction

Un ensemble d'outils et de méthodes a été nécessaire pour réaliser ce projet. Ainsi, dans ce chapitre, nous détaillons les principaux outils informatiques utilisés ainsi que les méthodes statistiques adoptées. D'ailleurs, n'importe quel projet data science ne peut être réalisé sans un ensemble de données fiables servant à l'apprentissage automatique (entraînement des modèles statistiques). Pour ce faire, nous explicitons dans ce chapitre toutes les étapes de la préparation de données, depuis la récupération jusqu'à la mise en forme.

2.2 Outils et méthodes

2.2.1 Méthodes utilisées

Avant de commencer ce projet, il fallait mettre en place un ensemble de méthodes à utiliser, celles-ci dépendent étroitement de l'objectif général et des objectifs spécifiques du sujet. Ainsi, le projet a suivi une méthode facile, rapide, claire et adaptée aux projets data science, c'est la méthode CRISP (Cross-Industry Standard Process).

Elle a été développée en 1996 par IBM et utilisée dans le monde de l'entreprise, en premier temps dans les projets data mining. Elle a été généralisée à tous les projets data science grâce à son schéma d'application standard.

La méthode CRISP divise le déroulement d'un projet data science en six étapes principales :

1. **La compréhension du problème :** c'est une étape cruciale dans chaque projet data science, elle consiste à comprendre la problématique du sujet à travers une analyse bibliographique des travaux similaires en se référant aux divers articles scientifiques. Elle permet enfin de fixer l'objectif général et les objectifs spécifiques du projet.
2. **La compréhension des données :** pour une bonne exploitation des données à notre disposition, il fallait les comprendre d'abord. Ainsi, cette étape permet de comprendre les données à utiliser et leur relation avec le problème traité à travers une étude exploratoire consistant à faire des visualisations et le calcul de quelques indices statistiques.
3. **La préparation des données :** c'est une étape de collection, de filtrage et de réorganisation des données dans le but d'avoir une base de données facilement manipulable par les algorithmes et peut servir de base d'entraînement pour un problème de modélisation. En effet, les données brutes sont souvent complexes et nécessitent des pré-traitements.
4. **La modélisation :** En général, modéliser signifie la représentation du comportement d'un phénomène, dans le but de résoudre un problème concret. Dans notre cas, elle consiste à sélectionner les modèles statistiques ainsi que leurs paramètres et qui sont susceptibles de répondre à notre problématique (prédiction, décision, identification, etc.). Autrement dit, les modèles sont entraînés sur des données appelées données d'apprentissage pour pouvoir faire des prédictions sur d'autres jeux de données.
5. **L'évaluation et interprétation :** elle consiste à juger les performances des modèles élaborés dans la partie modélisation et comparer les résultats produits avec les objectifs à atteindre.
6. **Le déploiement :** À la fin du projet, c'est la phase pendant laquelle la solution est installée et de nouvelles pratiques sont mises en place suite aux conclusions tirées du projet.

En plus de la méthode susmentionnée, on peut lister d'autres méthodes, c'est le cas par exemple d'AgileKDD, ASUM-DM et AABA. Mais la méthode CRISP est la plus utilisée, car elle est agile et favorise les retours en arrière grâce à son approche itérative et cyclique.

La figure ci-dessous illustre l'approche CRISP :

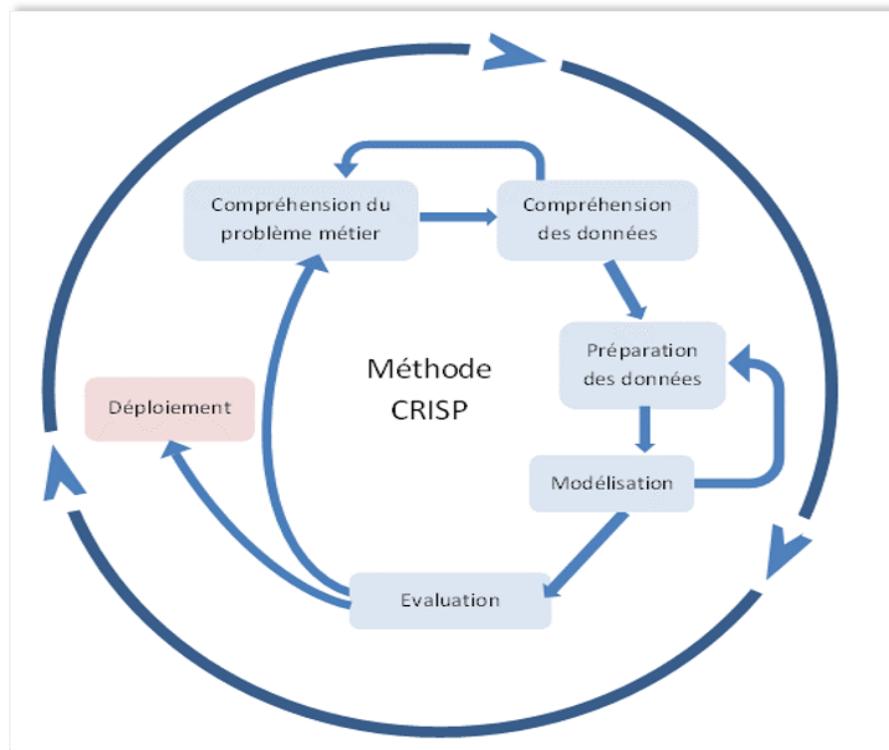


FIGURE 5 – Méthodologie et cycle de vie CRISP

Dans notre cas, les étapes de cette méthode ont été adoptées comme il est mentionné dans le tableau suivant :

La compréhension du problème	La compréhension des données	La préparation des données	La modélisation	L'évaluation et l'interprétation
Analyse des articles scientifiques. Analyse des données en relation avec le problème.	Visualisation. Calcul des indices statistiques	Filtrage, réorganisation et mise en forme des données.	Sélection des variables potentielles Régression linéaire multiple (RLM) Réseau de neurones (DNN)	Calcul des indices de performance et l'analyse graphique de sensibilité des modèles Validation croisée pour vérifier la pertinence des modèles.

TABLE 2 – Les étapes de la méthode CRISP pour la présente étude

2.2.1 Outils utilisés

Pour pouvoir exécuter les différentes tâches du projet, depuis la récupération des données jusqu'à l'évaluation et l'interprétation des modèles, un ensemble d'outils informatiques a été utilisés. Les principaux outils informatiques sont listés ci-dessous :

- **Le système d'exploitation linux Ubuntu 20.04** : Linux est un système d'exploitation très riche. On peut l'utiliser sous deux façons, via une interface graphique conviviale similaire à Windows ou via un terminal en utilisant des commandes shell. On peut exploiter de nombreuses commandes préétablies ou créer nos propres commandes en utilisant des scripts shell. Ainsi, ce système d'exploitation met à notre disposition un ensemble de commandes qui facilite largement le travail de la préparation des données.
- **Filezilla** : c'est un client FTP permettant le transfert de fichiers d'un serveur vers un ordinateur ou vice-versa. C'est un logiciel qui a une interface graphique simplifiant son utilisation. Il permet de se connecter à tous les serveurs FTP. Nous l'avons utilisé dans notre cas pour la récupération des données satellitaires. Il suffit d'avoir le hostname ou l'adresse IP du serveur, le nom d'utilisateur et le mot de passe du client FTP.
- **Panoply** : est un logiciel permettant la visualisation de certains jeux de données stockées dans des fichiers de formats spécifiques comme GRIB, netCDF, HDF, etc.
- **Python 3** : c'est un langage de programmation interprété. Il détient de nombreuses librairies facilitant l'analyse des données et la visualisation (numpy, matplotlib.pyplot, seaborn, basemap, etc.). En plus, il est doté d'outils très avancés pour l'exploration des données et la modélisation (scikit-learn, keras, TensorFlow, etc.).
- **Google Colab** : c'est un environnement python destiné principalement à l'analyse de données et le ML. Il permet d'écrire et d'exécuter les codes python en ligne par le biais du navigateur. En travaillant sous Google Colab, aucune configuration ou installation de librairies n'est exigée.

2.3 Compréhension des données

2.3.1 Paramètres de la neige sur terre (SWE, SD, RSN, SC)

SD (Snow Depth)

Le SD détermine la profondeur ou l'épaisseur du manteau neigeux. Toutefois, il ne reflète pas la quantité de la masse de la neige contenue dans un manteau neigeux. En effet, deux manteaux neigeux peuvent avoir le même SD, mais la masse neigeuse totale n'est pas la même si les densités RSN sont différentes.

Le SD sur terre varie souvent entre 0 m et 33 m d'après les réanalyses ERA5-Land. La figure ci-dessous est une cartographie de SD sur la terre le 15 février 2019. Nous nous sommes limités à 2 m pour que les petites variations de SD puissent être visibles à l'œil nu.

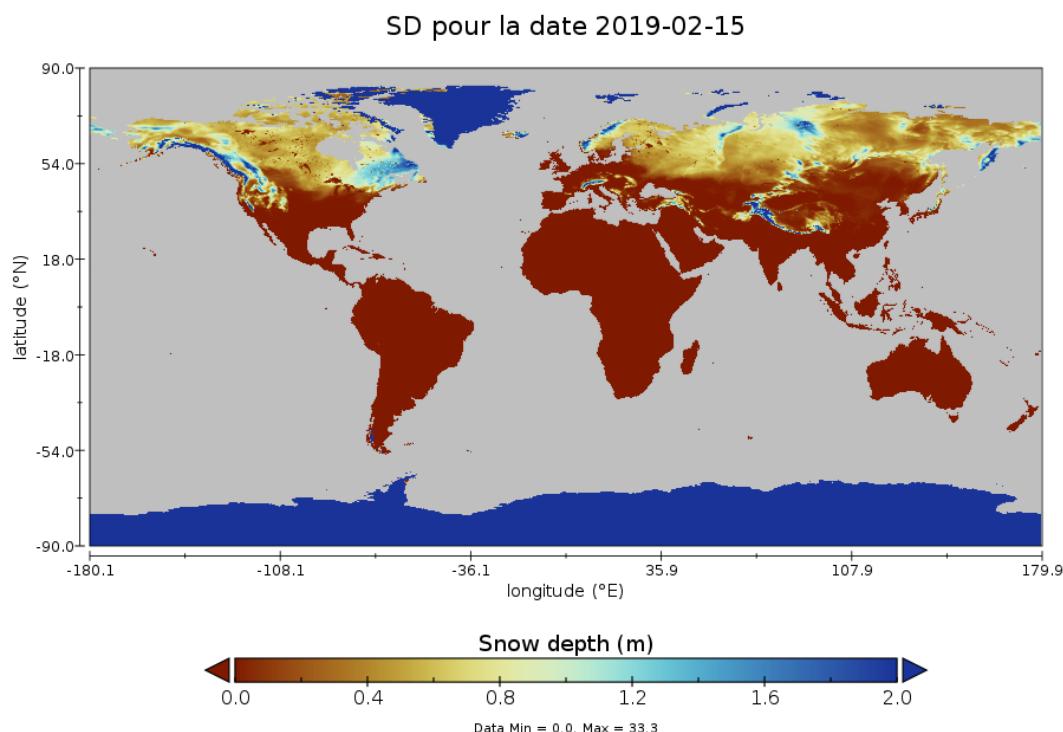


FIGURE 6 – SD pour la date 2019-02-15

RSN (densité de la neige)

RSN est la densité de la neige, elle représente la masse de la neige par mètre cube dans la couche de neige. Comme nous allons le voir dans le métamorphisme saisonnier de la neige, ce paramètre augmente avec le temps pour un manteau neigeux tant qu'il n'est pas totalement fondu.

La carte ci-dessous illustre les variations de RSN sur terre pour la date 2019-02-15 :

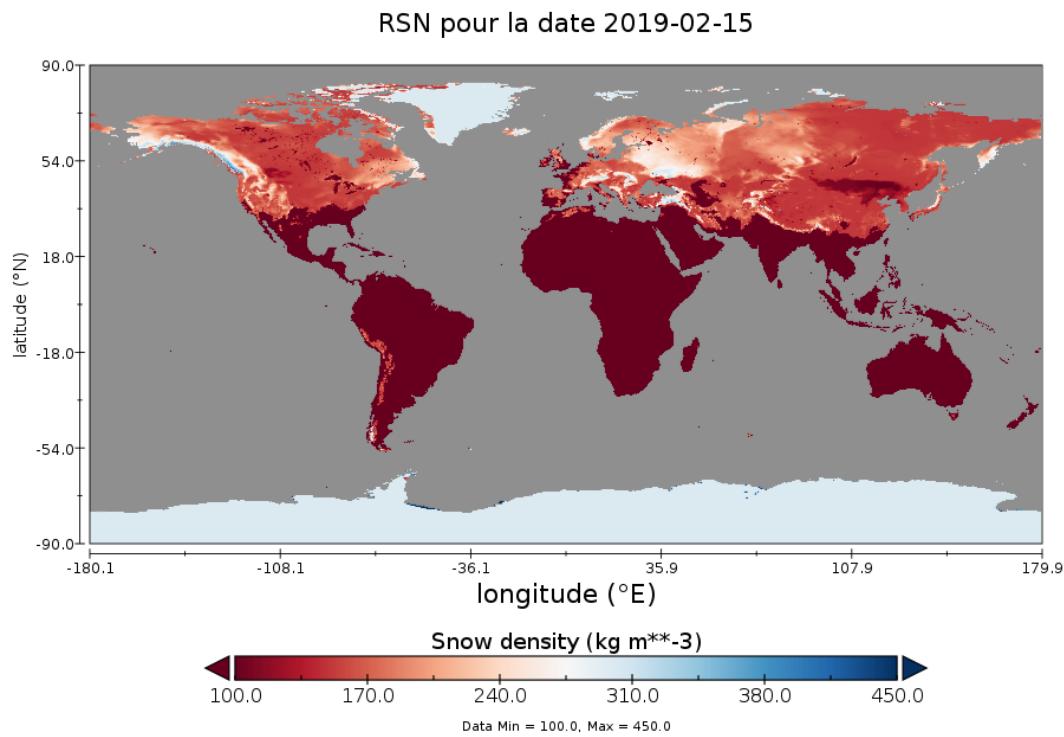


FIGURE 7 – RSN pour la date 2019-02-15

Les variations observées de RSN illustrent bel et bien le changement de densité que subissent les manteaux neigeux avec le temps. En effet, dans les régions qui détiennent des manteaux neigeux plus anciens, RSN prend des valeurs plus élevées qui sont approximativement de l'ordre de 300 kg m^{-3} , c'est le cas par exemple de l'Antarctique, le Groenland, les alpes, le sud-est d'Alaska, etc. Pour la neige saisonnière, au niveau des terres eurasiatiques et l'Amérique du Nord, les valeurs prises par la densité de la neige des manteaux neigeux formés pendant la période couvrant la saison hivernale de l'année 2018-2019 sont souvent inférieures à 180 kg m^{-3} .

La figure suivant montre deux histogrammes de la distribution de densité de la neige sur terre sur toute la période hivernale de l'année 2018-2019 dans l'hémisphère nord à partir de la latitude 30° en prenant en considération le Groenland (histogramme à gauche) et en filtrant le Groenland (histogramme à droite) :

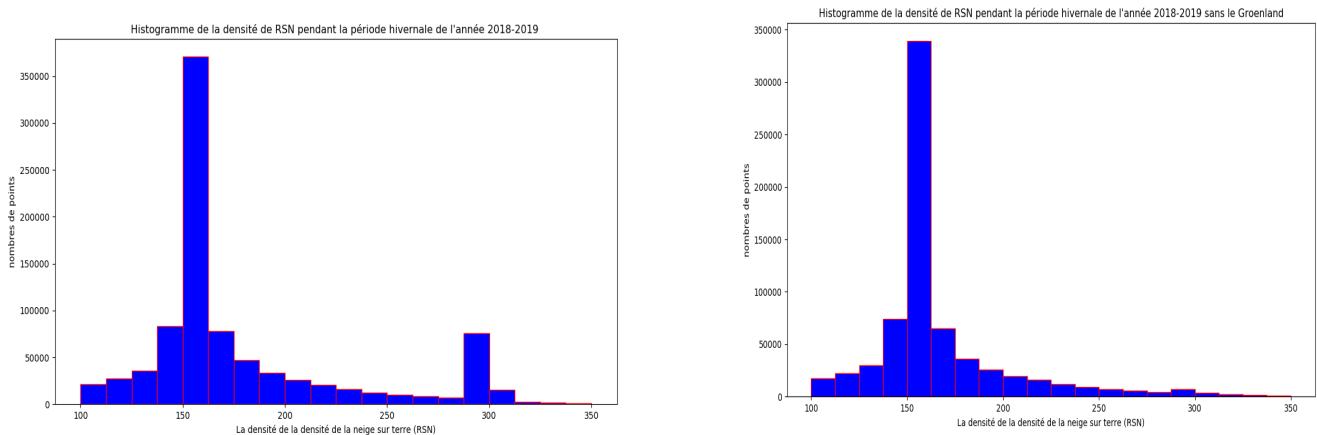


FIGURE 8 – Histogramme de la densité de la neige sur terre pendant la période hivernale (11/2018, 12/2018, 01/2019, 02/2019 et 03/2019) de l'année 2018-2019 dans l'hémisphère nord à partir de la latitude 30° avec et sans le Groenland

D'une part, il est très clair d'après les deux histogrammes ci-dessus que les densités comprises entre 140 et 160 $kg\ m^{-3}$ prédominent, d'autre part, nous observons une densité importante autour de la valeur $300\ kg\ m^{-3}$, ces valeurs sont rencontrées souvent dans les régions détenant des manteaux neigeux anciens comme le Groenland et l'Arctique.

Le deuxième pic à droit de l'histogramme à gauche a disparu dans l'histogramme à droit, cela montre bel et bien que le Groenland détient des masses neigeuses très entassées et très compactées, car elles sont plus anciennes. Elles auront indubitablement des propriétés radiatives différentes par rapport aux manteaux neigeux qui sont saisonniers. Il sera ainsi plus judicieux de filtrer ces régions dans l'étape de la modélisation.

SWE (Snow Water Equivalent)

Le SWE (l'équivalent en eau de la neige) donne l'épaisseur de la couche d'eau liquide qui correspondrait à la fonte de la neige d'une manière uniforme sur l'ensemble de la maille. Il s'exprime en mètre d'équivalent d'eau (m) [26].

Cette variable est très utilisée en hydrologie, car elle fournit la quantité totale d'eau (liquide et solide) stockée dans le manteau neigeux [21].

Connaissant l'épaisseur de la neige (SD) et sa densité (RSN), on peut déduire l'équivalent en eau en mètre par la simple relation :

$$SWE \ (m) = \frac{RSN \ (kg \ m^{-3})}{\rho_{eau} \ (kg \ m^{-3})} \ SD \ (m) \quad (1)$$

La relation susmentionnée peut être simplifiée et le SWE en millimètres de précipitations (mm) peut être obtenu en multipliant le SD en mètres par la densité RSN du manteau neigeux en $kg \ m^{-3}$ [21].

$$SWE \ (mm) = RSN \ (kg \ m^{-3}) \ SD \ (m) \quad (2)$$

Où le mm de précipitations représente le nombre des litres dans un mètre carré.

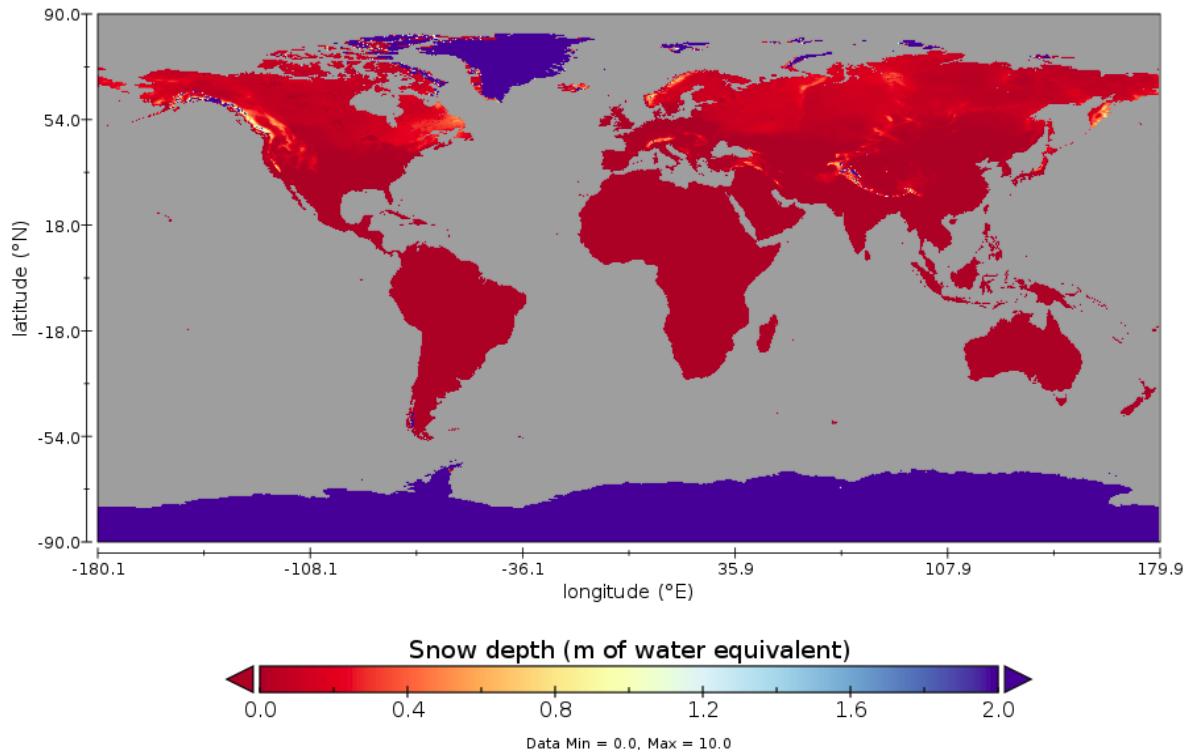


FIGURE 9 – SWE sur la terre pour la date 2019-02-15

D'après les cartes de SD et RSN, les régions où ces deux paramètres sont importants sont caractérisées par un SWE aussi important, c'est le cas de Victoria Island, Groenland, Svalbard, Terre du Nord, etc. Cela est conforme à la relation susmentionnée entre SD, RSN et SWE.

En ce qui concerne la distribution, les valeurs les plus rencontrées sont les valeurs inférieures à 0.4 m. Cela peut être confirmé et affiné en traçant l'histogramme de la distribution de la SWE suivant :

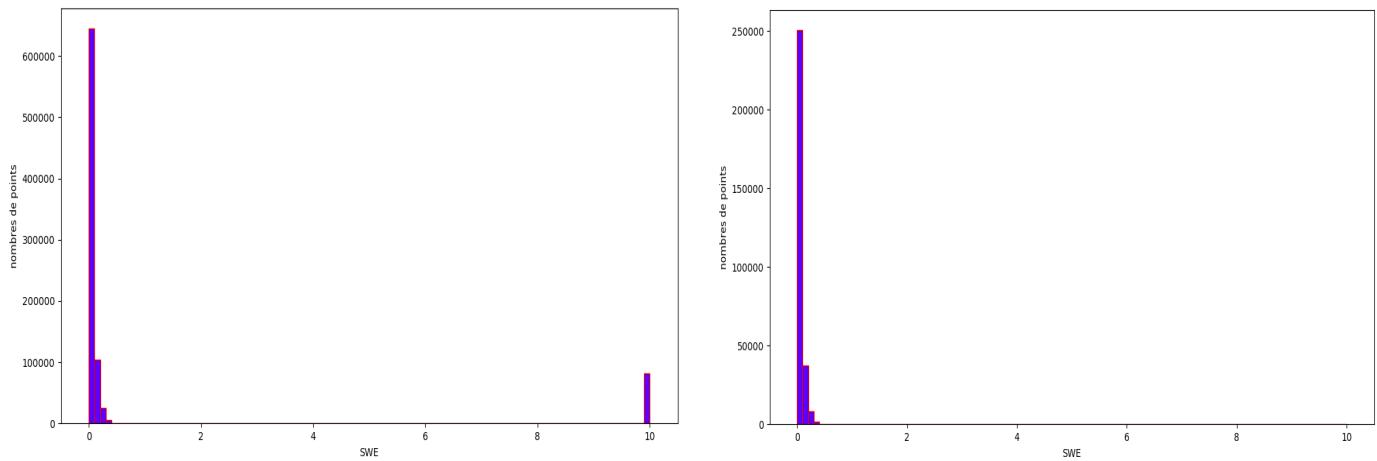


FIGURE 10 – Histogramme de la densité de SWE sur terre pendant la période hivernale (11/2018, 12/2018, 01/2019, 02/201 et 03/2019) de l’année 2018-2019 dans le l’hémisphère nord à partir de la latitude 30°, pour l’histogramme à droite, tout le domaine géographique est inclus, pour celui à gauche, les régions délimitées dans la carte ci-dessous sont filtrées

Sans compter les régions délimitées dans la carte ci-dessous, où la quasi-totalité des valeurs de SWE sont supérieures à 9 m, on remarque que presque toutes les valeurs de SWE sont inférieures à 0.4 m.



FIGURE 11 – Les régions filtrées pour produire l’histogramme à droite de la figure ci-dessus en utilisant l’image satellitaire issue de google map prise le 15 mai 2021

Pour bien affiner, nous allons nous limiter dans l’histogramme suivant aux valeurs qui sont inférieures à 0.4 m. On obtient l’histogramme suivant :

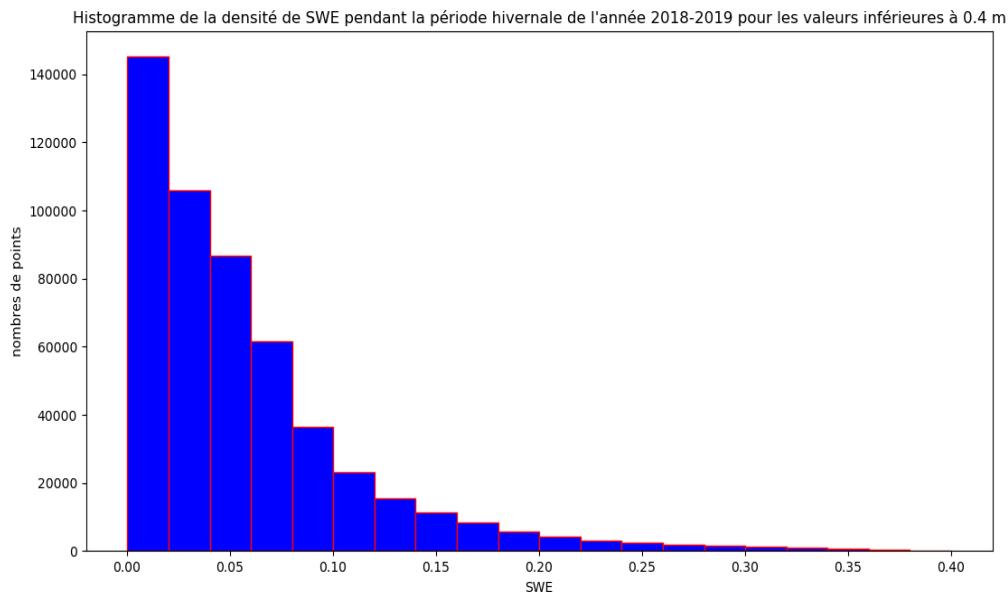


FIGURE 12 – Histogramme de la distribution de SWE (en m) sur terre pendant la période hivernale (11/2018, 12/2018, 01/2019, 02/2019 et 03/2019) de l'année 2018-2019 dans le l'hémisphère nord à partir de la latitude 30° en se limitant aux valeurs inférieures à 0.4 m

L'histogramme ci-dessus dévoile une décroissance exponentielle de la densité de SWE. Les valeurs proches de zéro sont les plus rencontrées.

SC (Snow Cover)

La couverture neigeuse (SC) est une variable prognostique du modèle que l'on calcule à partir du SD (l'épaisseur de la neige) et RSN (la densité de la neige) :

$$SC = \min\left(1, \frac{\left(\rho_{eau} \frac{SD}{RSN}\right)}{0.1}\right) \quad (3)$$

Le seuil est de 0.1 m au-dessus duquel on considère 100% de couverture neigeuse.

2.3.2 Temprératures de brillance micro-ondes

D'une manière générale, la température de brillance d'un corps est la température qu'aurait un corps noir qui émettrait la même quantité d'énergie [45]. Cette grandeur est exprimée en Kelvin (K).

Dans la nature, un corps noir n'existe pas, les émissivités de tous les corps natu-

rels sont inférieures à 1. Autrement dit, aux températures physiques égales, un corps naturel émet moins qu'un corps noir. Ainsi, la température physique d'un corps est toujours supérieure à sa température de brillance [45].

On peut ainsi définir la température de brillance avec la relation suivante :

$$Tb_{\nu,p} = e_{\nu,p} T \quad (4)$$

En ce qui concerne les surfaces terrestres enneigées, le rayonnement émis par un manteau neigeux est polarisé. Par conséquence, l'intensité du champ électromagnétique dépend de l'orientation considérée par rapport à la direction de propagation de l'onde [45].

Les données disponibles des températures de brillance des micro-ondes sont des mesures issues des radiomètres à différentes bandes de fréquences (6, 7, 10, 18, 23, 36, 89 GHz, etc.), selon les deux polarisations linéaires, horizontale et verticale [45].

L'écart entre les deux polarisations, verticale et horizontale est souvent important et provient principalement des réflexions à l'interface air-neige et à l'intérieur du manteau neigeux [46].

Par conséquence, les deux polarisations ne sont pas sensibles aux mêmes caractéristiques des manteaux neigeux.

Pour bien illustrer ces variations, nous allons visualiser quelques Tbs issues du capteur AMSR2 embarqué au bord du satellite GCOM-W1 en deux polarisations.

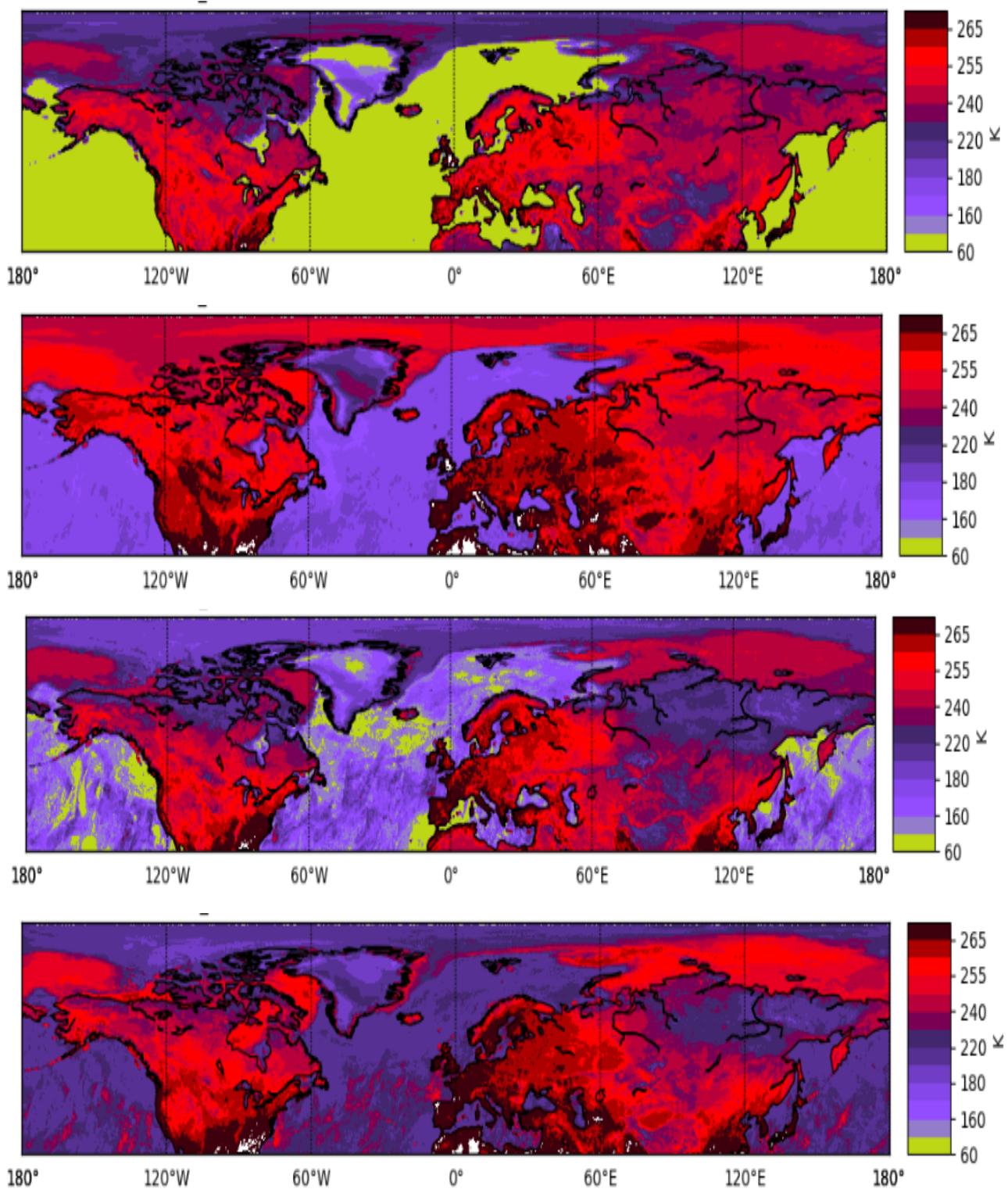


FIGURE 13 – La moyenne de Tb sur les cinq premiers jours de décembre 2018 dans la bande de fréquence centrée sur 10 et 36 GHz polarisation horizontale et verticale dans cet ordre : tb10h, Tb10v, Tb36h, Tb36v

D'après les visualisations des températures de brillance dans certaines bandes de fréquences pour les cinq premiers jours du mois de décembre, on remarque d'abord un grand contraste terre-mer. En effet, les océans ont en général des émissivités plus faibles que les continents. On peut constater aussi une diffé-

rence significative entre les deux polarisations pour le même canal, notamment au niveau des océans et les surfaces enneigées. Par exemple, la Tb dans la bande centrée sur la fréquence 10 GHz est souvent plus forte en polarisation verticale qu'en polarisation horizontale.

Par exemple, la température de brillance dans la bande de fréquence centrée sur 10 GHz polarisation horizontale est plus grande que celle dans la bande de fréquence centrée sur 36 GHz polarisation horizontale dans les terres russes. En effet, en décembre, la plupart des terres russes sont recouvertes par la neige saisonnière avec des épaisseurs faibles à modérées et que les signaux de hautes fréquences sont plus impactés par les manteaux neigeux que les signaux de basses fréquences.

En conséquence, la combinaison entre les différentes bandes spectrales avec les deux polarisations permet certainement de maximiser l'information.

2.4 Récupération des données

2.4.1 Source des données de la neige

Les données récupérées de la neige sont les données de SWE et celles de la densité de la neige RSN à partir des ré-analyses ERA5. Les données sur la couverture neigeuse et l'épaisseur de la neige sont récupérées à partir d'ERA-Land (résolution de 0.1°) et sont utilisées pour faire seulement de la visualisation.

Réanalyses : ERA5

ERA5 est la cinquième génération des réanalyses du climat et des conditions météorologiques à l'échelle mondiale menées par le ECMWF. Elle a remplacé la réanalyse d'ERA-Interim.

La réanalyse consiste à combiner les observations du monde entier avec les données du modèle en un ensemble de données cohérent et complet. Ce processus est appelé l'assimilation de données, celle-ci est basée sur la méthode 4D-Var (l'analyse variationnelle en 4 dimensions), les méthodes de corrections de biais variationnelles pour les données satellitaires et d'autres améliorations du traitement des données, dans le but de produire d'une manière optimale la meilleure estimation possible de l'état de l'atmosphère, appelée l'analyse, avec laquelle les

modèles de prévision du temps sont forcés pour émettre une prévision.

La réanalyse fonctionne de la même manière, mais à une résolution réduite pour permettre la fourniture d'un ensemble de données couvrant plusieurs décennies. Quant à la méthode d'assimilation de données pour produire l'analyse servant de forçage pour les modèles de prévision du temps, la réanalyse n'a pas la contrainte d'émettre des prévisions en temps voulu, ce qui permet d'avoir suffisamment de temps pour collecter des observations, injecter de versions améliorées des observations originales dans le passé, ce qui améliore la qualité du produit de réanalyse.

Les données ont été redressées sur une grille régulière latitude longitude de résolution $0,25^\circ \times 0,25^\circ$ pour la réanalyse.

Ces données sont disponibles et regroupées en quatre sous-ensembles principaux : des produits horaires et mensuels, à la fois sur les niveaux de pression (champs d'air supérieur) et au niveau unique (les paramètres atmosphériques dont ceux de la surface terrestre). Les estimations de l'incertitude sont également disponibles, elles sont étroitement liées au contenu d'informations du système d'observation disponible.

Les données ERA5 sont open source et téléchargeables en utilisant des scripts python qu'on obtient en sélectionnant les critères des données à télécharger sur le lien suivant : [ERA5 Data](#).

La base de données récoltée

La base de données récupérée des deux paramètres de la neige susmentionnées s'étale du début novembre 2018 jusqu'au mars 2019. Pour chaque mois, seulement quatre jours sont pris en compte ; le premier, le dixième, le vingtième et le dernier jour. Le domaine géographique considéré pour la base de données brute s'étale à partir de la latitude 30° puisqu'on s'intéresse dans cette étude plus précisément au continent eurasiatique et l'amérique du nord. La base de données finale est constituée d'un ensemble de fichiers NetCDF horaires.

2.4.2 Source des données satellitaires

Les données satellitaires exploitées dans ce projet sont les températures de brillance (Tb) micro-ondes passifs, plus précisément les fréquences suivantes : 6.925, 10.65, 18.7, 23.8, 36.5 et 89.0 GHz.

Les données de températures de brillance dans les micro-ondes sont issues des radiomètres micro-ondes : AMSR2, SMMR, SSM/I, etc. Aucun des deux derniers instruments n'a été conçu explicitement pour les applications sur la neige, mais il s'est avéré qu'ils sont efficaces pour les applications qui concernent la neige [10] [36].

Le radiomètre AMSR2 embarqué au bord du satellite GCOM-W1 appartient à la quatrième génération du radiomètres hyperfréquences (à micro-ondes) satellitaires de JAXA [7]. Il détecte l'énergie hyperfréquence polarisée horizontalement et verticalement prévenante de la surface terrestre et l'atmosphère à 6.925, 7.3, 10.65, 18.7, 23.8, 36.5 et 89.0 GHz faisant de ce radiomètre une source incontournable pour la fourniture des données de températures de brillance dans les hyperfréquences.

JAXA a publié deux produits de données pour les Tbs AMSR2 : le niveau 1R (L1R) et le niveau 1B (L1B) [7]. Les données du premier produit qui nous intéresse dans la présente étude sont stockées dans le serveur de fourniture de données GCOM-W1 depuis 2012 "standard/GCOM-W/GCOM-W.AMSR2/L1R/2". Depuis 2015, ces produits peuvent être téléchargés gratuitement à partir du service de fourniture de données GCOM-W1 Data Providing en utilisant des outils informatiques comme Fillezilla.

Base de données récoltée

Comme pour les données des paramètres de la neige récoltées à partir des réanalyses ERA5, le domaine géographique et la période restent les mêmes pour les données des Tbs AMSR2. La base de données finale est constituée d'un ensemble de fichiers hdf comportant seulement les données des orbites descendantes. En effet, chaque fichier hdf correspond à une orbite. L'orbite lui-même est deux types : orbite descendante et orbite ascendante. Les orbites descendantes sont toujours nocturnes en heure locale et vice-versa. Les Tbs MOP récoltées pendant la nuit ont l'avantage de ne pas être affectées par le rayonnement solaire.

2.5 Conclusion

Ainsi, ce projet a nécessité la mise en œuvre de plusieurs méthodes, divers outils informatiques et des données pour explorer le phénomène et la modéliser. Pour ce fait, la méthode CRISP a été amenée pour le bon déroulement de ce projet, divers logiciels et langages de programmation ont été utilisés ainsi qu'une base de données des paramètres de la neige sur terre issue des réanalyses ERA5 et des températures de brillances AMSR2.

CHAPITRE 3 :

GÉNÉRALITÉS SUR LA NEIGE ET LE RAYONNEMENT

3.1 Introduction

Une étude Data Science ne pourra guère aboutir sans une compréhension solide des différents processus physiques qui régissent le phénomène étudié. De ce fait, le présent chapitre aborde quelques notions de la physique de la neige sur terre, à savoir sa formation, son métamorphisme, sa variabilité spatio-temporelle, etc. Nous parlons aussi de l'interaction de la neige sur terre avec le rayonnement, et enfin le problème d'inversion. Ce travail bibliographique a été mené en s'appuyant sur divers articles scientifiques bien renommés dans la littérature scientifique.

3.2 La neige

3.2.1 Formation de la neige

La neige est un ensemble de cristaux de glace individuels qui se sont développés en suspension dans la troposphère, puis ils sont précipités et se sont accumulés sur le sol où ils subissent de nouvelles transformations [44].

Le processus de formation de la neige débute par le déclenchement des nuages de glace ou des nuages de glace plus les gouttelettes d'eau dans la troposphère. Suite au phénomène de l'évapotranspiration, des quantités importantes de la vapeur d'eau sont mises dans l'atmosphère. Cette vapeur d'eau fait augmenter l'humidité des masses d'air qui survolent les surfaces évapo-transpirées. Ces masses d'air enrichies en vapeur d'eau peuvent gagner de l'altitude dans la tro-

osphère suite à l'un des quatre facteurs de soulèvement. Le premier est nommé soulèvement horographique, où les basses couches d'air gagnent de l'altitude quand elles sont poussées par le vent en direction d'un relief suffisamment grand, la masse d'air n'arrive donc pas à contourner le relief, mais elle le surmonte. Puis, le soulèvement frontal, il se produit quand deux masses d'air de caractéristiques différentes entrent en contact, celles-ci ne se mélangeant pas, mais celle qui a une température potentielle relativement grande surmonte celle dont la température potentielle est relativement basse, elle est ainsi rejetée en altitude. Ensuite, le soulèvement par convection, où les couches inférieures se réchauffent à partir de la base, elles deviennent plus légères et prennent de l'altitude, elles sont remplacées par l'air froid et sec qui descend des couches supérieures de la troposphère. Enfin, le soulèvement par convergence de l'air, quand les flux d'air de différentes directions se rencontrent pour créer une accumulation de masses d'air. Comme l'air ne peut pas pénétrer la surface terrestre, il prend de l'altitude.

Les masses d'air humides rejetées en altitude transportent de l'énergie de la surface vers les couches supérieures de la troposphère sous forme de chaleur latente. Les nuages sont formés comme réponse au forçage énergétique par cette chaleur.

Plus précisément, un nuage n'est qu'un ensemble de gouttelettes d'eau ou des cristaux de glace formés via la condensation de la vapeur d'eau dans les couches supérieures de la troposphère. Cette condensation elle-même est due à la baisse de la température avec l'altitude, ce qui fait augmenter l'humidité des parcelles fluides transportées vers les couches supérieures via l'un des quatre facteurs de soulèvement susmentionnés. Ainsi, la vapeur d'eau se condense pour construire une goutte initiale ou cristal de glace selon la température. Dans l'air, les molécules d'eau entrent en collision de façon aléatoire et se forment des embryons de gouttelettes. On appellera ce processus la nucléation homogène dans le cas de l'eau pure et la nucléation hétérogène lorsqu'une gouttelette ou un cristal de glace se forme sur un noyau de condensation de nuages ou un noyau de glace. La nucléation Homogène se produit rarement dans l'atmosphère, car elle nécessite une valeur de sursaturation inexistante dans la nature. Ainsi, les noyaux de condensation jouent un rôle très important dans le passage de la vapeur d'eau en gouttelettes d'eau ou cristaux de glace juste à une humidité légèrement supérieure à 100 %. Les principaux noyaux de condensation sont : les particules solubles, les particules insolubles, les noyaux mixtes et les aérosols atmosphériques.

Les gouttelettes ou les cristaux de glace formés ne sont pas suffisamment lourds

pour qu'ils tombent sous l'effet de la gravitation sous forme de précipitations. Ainsi, dans l'atmosphère, plusieurs mécanismes entrent en jeu pour favoriser le grossissement des gouttelettes d'eau et des cristaux de glace. Les gouttelettes d'eau grossissent par deux principaux mécanismes : la croissance par condensation et la croissance par collection. Concernant le premier, la croissance de la goutte dépend de la densité de la vapeur d'eau dans son voisinage. La collection fait augmenter davantage le volume de la goutte, il se produit lorsqu'une goutte collecte des petites gouttes qui se sont trouvées dans son passage. Les équations qui régissent la croissance des gouttes d'eau sont : l'équation uniforme et l'équation stochastique. Le grossissement des cristaux de glace est un peu différent de celui des gouttelettes d'eau, il est régi par trois principaux mécanismes : la croissance par condensation solide, la croissance par accrétion et la croissance par agrégation. Concernant le premier mécanisme, les particules peuvent grossir par déposition de vapeur d'eau. Pour le deuxième, les particules de glace grossissent par collecte des gouttelettes d'eau surfondues qui se congèlent à leur contact. Le troisième mécanisme, concerne la croissance par collision des cristaux de glace, et il faut que les cristaux de glace qui entrent en collision restent soudés après leur collision.

Ces mécanismes peuvent amener à différents types de nuages, ainsi à différents types de précipitations. La forme des nuages dépend essentiellement de la stabilité de l'air. Dans une atmosphère stable, on aura souvent des nuages stratoformes, mais dans une atmosphère instable, les nuages formés sont souvent de type cumuliforme. Leur extension verticale dépend principalement du profil vertical de l'humidité et leur phase dépend de leur température de formation. Les nuages convectifs ont un dynamisme particulier, ils s'organisent souvent sous forme de cellules qui passent par trois stades d'évolution : stade de cumulus, stade de maturité et stade de dissipation. En plus, le cisaillement vertical du vent horizontal et les vents forts en altitude ont tendance à incliner les cellules et les mélanger avec l'air environnant plus froid et plus sec. Mais il y a plusieurs mécanismes compensateurs qui entrent en jeu pour le renforcement des cellules convectives et l'accroissement de leur durée de vie, à savoir : l'effet d'obstacle, l'inclinaison de l'ascendance, la création d'un pseudo-front froid et la convergence d'altitude. Ainsi, des cellules géantes et des systèmes multicellulaires ayant des durées de vie assez grandes et des intensités de précipitation importantes prennent naissance.

Par conséquent, ces nuages formés peuvent amener à plusieurs types de précipitations, comme la bruine, la pluie, la grêle, la neige, les averses, etc. Ainsi, quand les conditions météorologiques favorisent les précipitations de la neige, celle-ci peut prendre plusieurs formes en fonction de la température du nuage.

Les principales formes sont : les plaquettes, les aiguilles, les colonnes creuses, les étoiles et les dendrites.

L'image ci-contre illustre les différents types de cristaux de neige dans les nuages avec leurs plages de température :

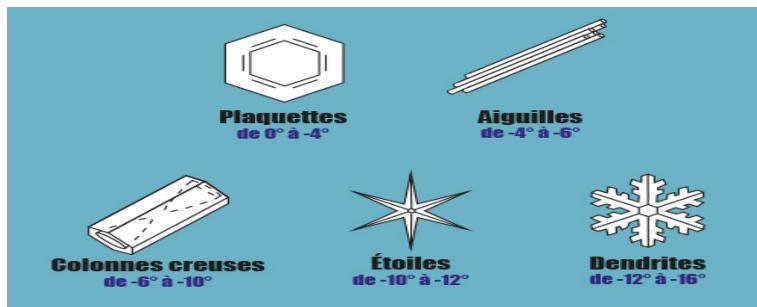


FIGURE 14 – Les types de cristaux de neige dans les nuages selon la température

Suite à des précipitations neigeuses, la neige se dépose sur la terre et prend de l'épaisseur si la température de surface est aussi favorable pour que la neige perdure. Après un certain temps, la neige tombée arrive donc à se tasser et à se congeler après un début de fusion. L'action du vent peut changer aussi la forme de la couverture neigeuse sur terre. En bref, plusieurs facteurs vont gouverner la forme de la couverture neigeuse.

Le dépôt de la neige sur terre forme ce qu'on appelle un manteau neigeux, celui-ci est constitué d'une accumulation stratifiée de la neige. Il se présente comme un mélange de cristaux de glace, de l'air et de l'eau liquide quand la température est proche de 0° ou suite à des précipitations liquides.

3.2.2 Les types de la couverture neigeuse

Selon les caractéristiques climatiques et géographiques, les conditions météorologiques, la nature des précipitations neigeuses et la durée de persistance de la neige sur terre, on distingue six types de couverture neigeuse [5] :

- **La neige fraîche (new snow)** : il s'agit d'un dépôt de neige plus récent, il le reste tant qu'on peut reconnaître la forme originale des cristaux de glace.
- **Le névé** : il s'agit d'une neige jeune, regelée, compactée et partiellement fondu. Les névés sont à l'origine de la formation des glaciers sur terre.

- **La vieille neige (old snow)** : elle indique une neige déposée et dont les cristaux se sont métamorphosés jusqu'au point où on n'arrive plus à reconnaître la forme originale des nouveaux cristaux.
- **La neige saisonnière (seasonal snow)** : elle désigne la neige qui perdure une seule saison ou la neige qui s'accumule pendant une saison.
- **La neige pérenne (perennial snow)** : elle indique la neige qui a persisté au sol année après année.
- **La neige poudreuse (powder snow)** : c'est la neige la plus fraîche. Elle est sèche et composée de cristaux de glace frais et détachés.

3.2.3 Répartition géographique de la neige sur la terre

La couverture nivale s'étend approximativement sur une superficie minimale de $4.10^6 km^2$ pendant le mois d'août dans l'hémisphère nord et sur $47.10^6 km^2$ comme couverture maximale en Janvier [17]. Ceci montre une variabilité saisonnière importante de la masse et de la couverture neigeuse dans l'hémisphère nord. Une telle variabilité sera bien illustrée par la suite en visualisant les différents paramètres de la neige pour quelques dates de l'année 2018-2019. Les variables que nous avons visualisées sont : la couverture neigeuse et l'épaisseur de la neige.

Couverture neigeuse sur la terre

La figure ci-dessous illustre la variabilité saisonnière de la couverture neigeuse sur la Terre pour l'année 2018-2019, chaque carte est une représentation de la couverture nivale à minuit pour un jour de la période d'étude. Les données utilisées sont issues de jeu de données ERA5-Land d'ECMWF, c'est un jeu de données fournissant une vue cohérente de l'évolution des variables terrestres sur plusieurs décennies à une résolution améliorée (0.1° latitude longitude) par rapport à ERA5, elles sont disponibles depuis 1981 jusqu'à nos jours. Il s'agit des réanalyses formées avec une combinaison de données du modèle avec les observations du monde entier pour produire un ensemble de données complet et cohérent à une résolution spatiale de 0.1° et une heure comme résolution temporelle en utilisant les lois de la physique.

La couverture neigeuse est donnée en pourcentage, et représente le pourcentage de la surface recouverte par la neige dans une maille du modèle.

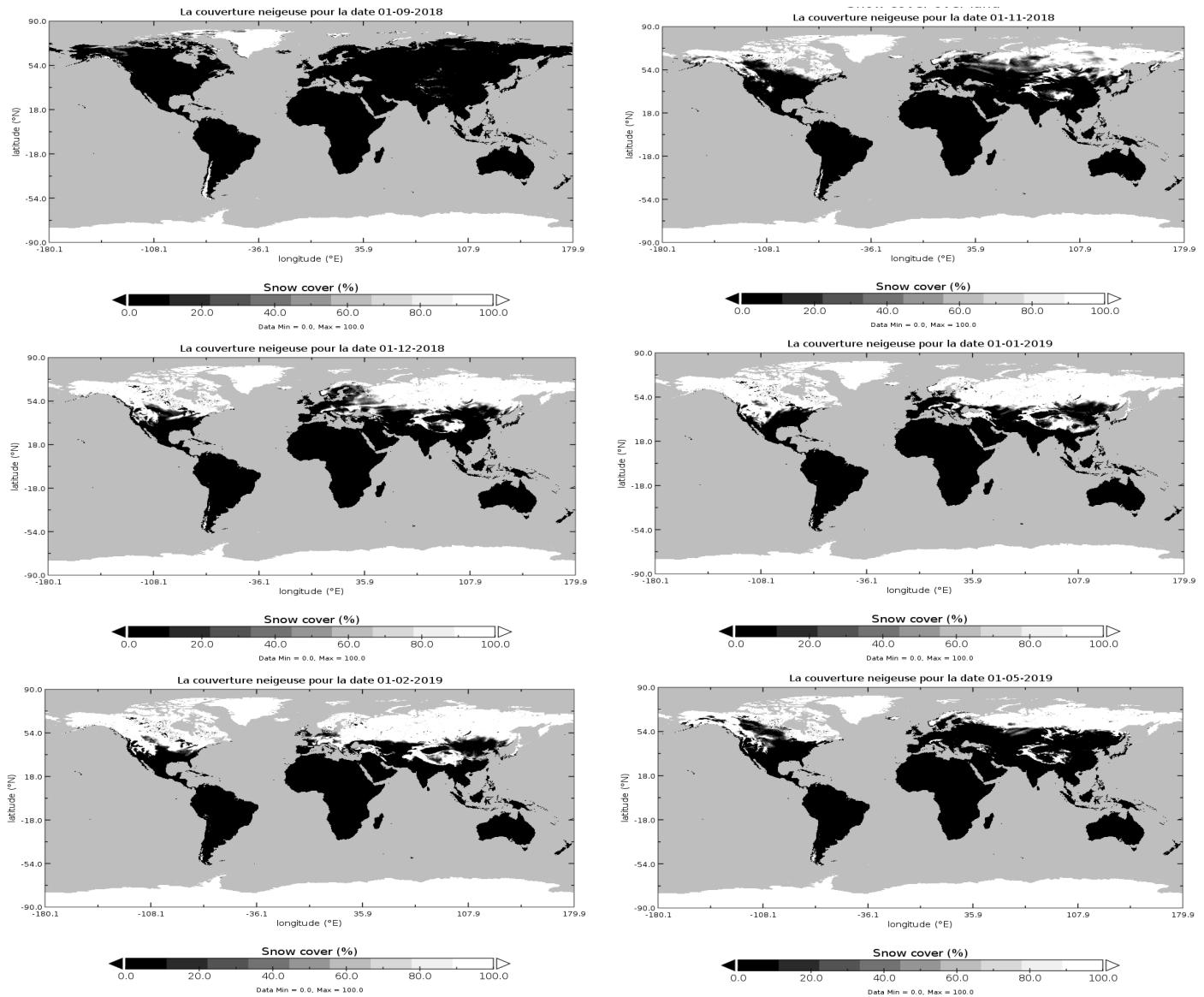


FIGURE 15 – L'évolution de la couverture neigeuse sur la terre pour l'année 2018-2019

Pour la date 2018-09-01, celle-ci correspond approximativement à la fin de l'été boréal, la présence significative de la neige sur terre est restreinte à l'Antarctique, au Groenland, à la chaîne de l'Himalaya, aux quelques régions d'Alaska, aux Cordillère des Andes, à l'extrême nord du Canada, aux Alpes et aux quelques îles à l'extrême nord de la Russie comme l'île Severny, l'île de la révolution d'octobre et l'île Komsomolets. Ces terres correspondent aux régions où les manteaux neigeux sont permanents. Dans ces terres, la neige n'est pas totalement fondue.

Au cœur de la saison d'automne, on remarque d'après la carte ci-dessus que la neige commence à s'étendre sur les terres eurasiatiques et sur l'Amérique du Nord. Par contre, elle disparaît de plus en plus dans l'hémisphère sud.

Vers la fin de la saison d'automne boréal, la neige couvre une grande partie dans l'hémisphère nord, surtout l'Asie, l'Europe Occidentale et l'Amérique du Nord.

La neige continue à couvrir la terre dans l'hémisphère nord vers les premiers jours de la saison d'hiver, tandis que la couverture neigeuse baisse considérablement dans l'hémisphère sud suite à l'arrivée de l'été austral.

La couverture neigeuse atteint presque son maximal vers le début février et la majorité des terres dans l'Asie, l'Amérique du Nord et l'Europe Occidentale sont couvertes par la neige saisonnière avec des épaisseurs, qui seront bien évidemment différentes selon les conditions climatiques, géographiques et météorologiques.

La dernière carte dévoile bien un abaissement de la couverture neigeuse suite à la fonte printanière de la neige à cause de l'augmentation de la température de l'air et la diminution des précipitations de neige. En contrepartie, des manteaux neigeux réapparaissent dans certaines régions de l'hémisphère sud comme les Cordillère des Andes, le Chili et la Nouvelle-Zélande.

L'épaisseur de la neige (SD) sur la terre

Le SD représente l'épaisseur totale de neige sur le sol. Ce paramètre est très important car il permet d'estimer le SWE en connaissant la densité de la neige quand on ne dispose que des observations *in situ*, comme il permet d'affiner l'estimation de l'albédo α d'une surface couverte par la neige [27]. Son estimation reste encore limitée à cause de sa grande variabilité spatio-temporelle ce que nous allons explorer dans les cartes qui suivront en utilisant les données ERA-Land.

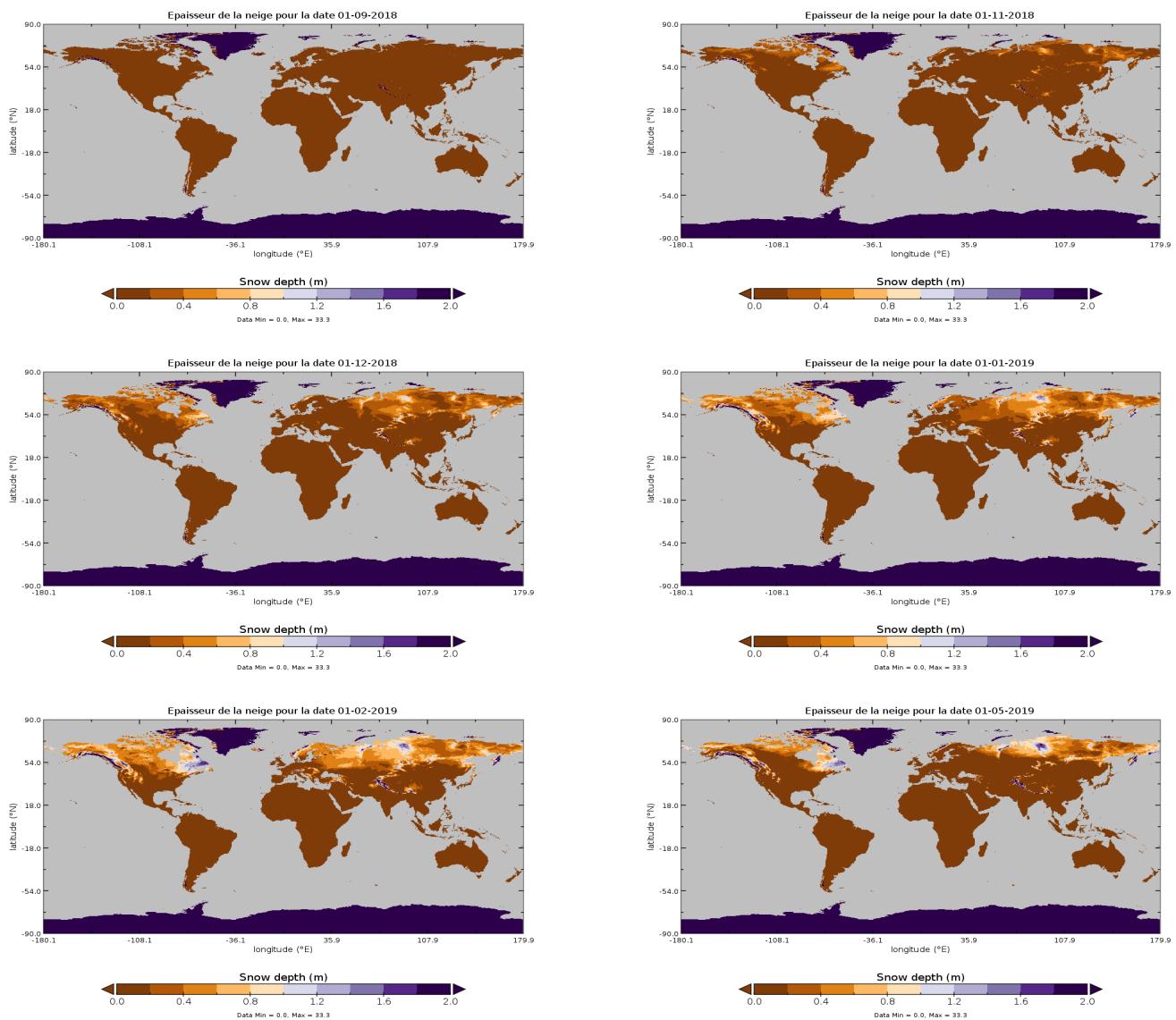


FIGURE 16 – L'évolution de l'épaisseur des manteaux neigeux sur la terre pour l'année 2018-2019

Pour visualiser ces cartes, nous nous sommes limités à 2 m pour le SD, cela ne signifie pas que le SD maximal est 2 m, mais cette façon nous a servi pour démasquer les petites variations du SD et pouvoir les visualiser à l'œil nu. En effet, sur plus de 80% de surfaces couvertes par la neige pendant la période hivernale, le SD prend des valeurs inférieures à 1 mètre, ce qu'on va montrer dans la partie destinée à l'exploration des données.

Les zones colorées en bleu marine correspondent aux zones géographiques contenant en grande partie la neige pérenne (manteaux neigeux plus anciens) et dont l'épaisseur peut dépasser 33 m, elles sont principalement : le Groenland et l'Antarctique. Mais on peut trouver aussi de la neige pérenne à faible couverture spatiale dans d'autres régions du monde comme l'Alaska, l'extrême nord du Canada, la Sibérie, les Alpes, L'Himalaya et quelques îles russes.

En premier novembre, suite aux premières précipitations, des manteaux neigeux saisonniers s'installent au nord asiatique et en Amérique du Nord (Canada et Alaska) et prend de l'épaisseur sans dépasser le 1 m.

Vers la fin de la saison d'automne, la neige s'accumule encore et gagne en épaisseur. Cette accumulation est très remarquable dans les terres sibérienne, l'Himalaya et la province québécois.

Avec l'arrivée d'hiver, les précipitations neigeuses hivernales dans les hautes et les moyennes latitudes donnent suite à l'augmentation de l'épaisseur des manteaux neigeux saisonniers où il dépasse 1 m dans certaines régions de l'hémisphère nord, à savoir, la Russie centrale, la Norvège, le Suède, Finlande et Newfoundland Labrador.

Au début février, le SD atteint des valeurs importantes dans plusieurs régions de l'hémisphère nord.

Vers la fin de la saison printanière, le SD ralentit et les manteaux neigeux disparaissent dans plusieurs régions des moyennes latitudes à cause de la fonte et la décroissance des précipitations neigeuses.

Les cartes susmentionnées des deux paramètres de la neige ont montré une forte variabilité spatio-temporelle de cette composante de la cryosphère en terme de couverture et de l'épaisseur. Une telle variabilité est due principalement à des facteurs topographiques, climatiques, météorologiques et géographiques. Ceci rend plus délicate l'estimation des propriétés de la neige à l'échelle planétaire.

3.2.4 Propriétés de la neige

La neige est caractérisée par nombreuses propriétés qu'on peut rassembler en quatre principales catégories : les propriétés électromagnétiques de la neige, les propriétés microphysiques de la neige, les propriétés macrophysiques de la neige et la composition de la neige [8].

- **Les propriétés microphysiques de la neige :** elles expliquent les caractéristiques de la microstructure de la neige, à savoir : la surface spécifique,

la forme des grains, etc. Dans le but de développer des algorithmes dynamiques permettant de récupérer des variables importantes de la neige à partir des observations satellitaires et qui sont souvent basés sur des hypothèses sur la microphysique de la neige, il est important d'estimer précisément les propriétés microphysiques de la neige [16] [17].

- **Les propriétés macrophysiques de la neige :** elles s'intéressent aux caractéristiques plus générales des manteaux neigeux ou d'une couche de neige. L'estimation de ces propriétés est incontournable comme informations d'entrée pour les modèles hydrologiques et les modèles de prévision météorologique numérique. Les propriétés macrophysiques les plus utilisées sont : la présence de neige, l'épaisseur de neige SD et l'équivalent en eau SWE. Ces paramètres ont un réseau d'observation plus étendu par comparaison à d'autres paramètres. On peut également citer d'autres propriétés macrophysiques qui sont mesurées uniquement dans les régions montagneuses pour la prévision des avalanches comme : la force de la neige, la dureté de la neige et la pénétrabilité de la surface de la neige [16] [17].
- **Les propriétés électromagnétiques de la neige :** elles servent à caractériser l'interaction de la neige avec le rayonnement électromagnétique. L'albédo de la neige par exemple est un paramètre clé pour le bilan énergétique et massique de la neige, et surtout pour le ruissellement des eaux de fonte. En effet, dans le spectre du visible, l'albédo de la neige contrôle la quantité d'énergie solaire qui est réfléchie vers l'atmosphère et absorbée par le manteau neigeux. En plus, l'estimation de l'albédo de la neige dans l'infrarouge peut servir pour récupérer les propriétés microstructurales de la neige. En effet, dans la zone spectrale du proche infrarouge, l'albédo de la neige est plus faible et dépend fortement des caractéristiques microstructurelles de la couche proche de la surface [8].
- **La composition de la neige :** elle décrit le contenu des manteaux neigeux en d'autres matières comme les impuretés.

3.2.5 Métamorphisme saisonnier des manteaux neigeux

Le métamorphisme saisonnier des manteaux neigeux est l'ensemble des phénomènes physiques affectant le manteau neigeux, et produisant des change-

ments dans ses propriétés.

La neige fraîchement tombée ne garde pas ses propriétés d'origine, mais elle se compacte, se tasse et subit beaucoup de transformations avec le temps. Ainsi, les propriétés des manteaux neigeux saisonniers se voient largement changées notamment en période printanière. Les processus métamorphiques constructifs et destructifs modifient la métrique et la densité des grains dans un manteau neigeux lors de l'évolution saisonnière de la neige [28].

Généralement, les propriétés qui subissent de transformations significatives sont : la forme des grains, la taille des grains, les propriétés mécaniques de la neige, la densité de la neige, la composition chimique de la neige et les autres propriétés physiques de la neige, à savoir : la perméabilité, l'albédo, la conductivité thermique, etc.

Les principaux facteurs qui interviennent indirectement dans le métamorphisme de la neige sont : le refroidissement radiatif suite à l'émission par les manteaux neigeux, le réchauffement solaire diurne, la circulation du vent, la gravitation et le flux géothermique entre la couche inférieure du manteau neigeux et la surface et entre les différentes couches. Ces facteurs favorisent l'échange de la vapeur d'eau entre les différentes couches de la neige et avec l'atmosphère et établissent un gradient thermique dans le manteau neigeux. Le gradient thermique est dû principalement au flux de chaleur provenant du substrat vers la surface du manteau neigeux [28]. En plus, la faible conductivité de la neige facilite l'apparition de ce gradient qui contrôle la fonte des manteaux neigeux et leur évolution stratigraphique [30].

Plus précisément, pour la neige fraîche, la densité et le diamètre des grains sont généralement plus petits. Autrement, pour un manteau neigeux, la densité et le volume des grains sont minimaux lors de la formation du manteau neigeux suite aux précipitations solides. Après quelques jours, la densité et le diamètre des grains augmentent. Cette augmentation ne vient pas évidemment du néant, mais les grains qui commencent à se frotter naturellement, ainsi, les branches cristallines se brisent et les cristaux de neige commencent à devenir arrondis et plus cohésifs [39]. Ainsi, la taille des grains peut augmenter d'un facteur de 2 à 3 à la fin de l'hiver [9]. Des gros grains sphériques se forment également en printemps suite à des cycles de fonte et de recongélation subies par la neige [9].

Ainsi, la taille, la forme et la cohésion des grains de neige subissent des transformations radicales et rapides. Ce métamorphisme affecte largement les propriétés de base de la neige [29]. La teneur en eau liquide est aussi l'une des transfor-

mations que peut subir la neige sur terre où elle modifie considérablement les propriétés radiatives du manteau neigeux selon que la teneur en eau liquide est faible ou élevée [29].

De ce fait, on peut classer la neige en fonction des processus dominants de son métamorphisme. En effet, un manteau neigeux peut être constitué de plusieurs couches ayant des caractéristiques de grains et de densité différentes. De point de vue géographique et topographique, les propriétés de ces couches dépendent de la position géographique et l'altitude et reflètent l'histoire du manteau neigeux [9]. Les principales classes de la neige sont : toundra, taïga, alpine, maritime, prairie et éphémère. Chaque type a un ensemble unique de caractéristiques stratigraphiques et texturales, notamment l'épaisseur des couches de neige, leur densité, la taille des grains et la morphologie des cristaux [9].

Ce métamorphisme affecte considérablement le comportement des micro-ondes dans les manteaux neigeux. En effet, les émissions radiatives à travers les manteaux neigeux dépendent étroitement de la densité et de la taille des grains de neige [28].

3.2.6 Importance de la neige pour le contrôle du climat et son impact sur l'atmosphère

L'atmosphère est affectée directement par les conditions de surface. En d'autres termes, une masse d'air acquiert ses propriétés qui la caractérise (la température potentielle, l'humidité, etc.) à partir de la surface qu'elle surmonte.

Ainsi, la couverture neigeuse va jouer un rôle important dans la modification des caractéristiques des masses d'air et le contrôle du climat à différentes échelles spatiales. En effet, un sol couvert de neige refroidit les basses couches de l'atmosphère et garde des températures trop froides grâce à une forte énergie nécessaire à sa fonte et son fort albédo qui lui permet de réfléchir une grande partie du rayonnement solaire incident [4]. Ainsi, grâce à ces propriétés radiatives et thermiques, la neige modifie les échanges d'énergie entre la surface et l'atmosphère. Elle est considérée comme un élément important dans le contrôle de la machine climatique [4]. Plus précisément, le bilan énergétique de l'hémisphère nord est très contrôlé par le cycle saisonnier de la masse neigeuse et de la couverture neigeuse [12].

De nombreuses études ont montré l'importance de la couverture nivale dans le contrôle du climat et l'impact que pourrait avoir le réchauffement climatique

sur la répartition de la masse neigeuse à cause des fortes rétroactions positives caractérisant la neige et sa grande sensibilité aux changements climatiques [2]. Parmi les changements remarquables évoqués dans la littérature scientifiques, on peut citer la modification de la circulation atmosphérique dans l'hémisphère nord suite à la diminution de l'étendue de la couverture neigeuse [3] [8]. De plus, la couverture neigeuse est très sensible aux variations des précipitations et de la température de l'air [26]. C'est pour cela qu'il est très important de détecter les tendances dans la masse et la couverture neigeuse sur terre pour mesurer le degré d'impact des changements climatiques. La couverture neigeuse sur terre est notamment considérée comme un indicateur très utile pour mesurer le degré des changements climatiques [14].

L'état de l'atmosphère est ainsi fortement impacté par la couverture neigeuse, faisant des paramètres de la neige des entrées importantes pour les modèles de prévision météorologique. Les paramètres de la neige peuvent aussi être utilisés pour la validation des modèles climatiques en confrontant les résultats d'une simulation issue d'un modèle aux différents paramètres observés [43].

3.2.7 Importance de la neige et de son estimation pour les applications hydrologiques

Dans les moyennes et les hautes latitudes, la neige constitue un apport précieux pour les eaux de surface. En effet, les manteaux neigeux accumulent de grandes quantités d'eau qui sont libérées suite à la fonte des manteaux neigeux pendant la période printanière et estivale. En plus, la fonte lente de la neige sur terre, notamment dans les régions froides, permet aux eaux libérées de prendre le temps pour s'infiltrer et alimenter la nappe phréatique. Pour cela, la neige dans les moyennes et les hautes latitudes constitue une source importante pour les eaux souterraines et de nombreuses rivières [43] et impacte fortement le bilan hydrologique mondial [9]. Une diminution des manteaux neigeux pour une année peut entraîner une crise dans le domaine de l'agriculture, car les rivières et les cours d'eau sont secs à l'automne et les récoltes vont être évidemment mauvaises [6].

Néanmoins, pour bien exploiter cette source et pouvoir anticiper son comportement, une estimation précise de ses paramètres est cruciale, car l'absence d'une information précise sur les propriétés de la neige sur terre est la source d'une grande lacune dans la connaissance du bilan hydrique global [8]. Autrement dit, les observations des paramètres de la neige sont utilisées pour initialiser les modèles hydrologiques [43]. À titre d'exemple, pour la production de l'énergie

hydraulique, la prévision et la planification des inondations, les études climatologiques globales de la recharge des eaux souterraines et l'approvisionnement en eau et la gestion des ressources en eau, il est particulièrement crucial d'estimer l'équivalent en eau de la neige (SWE), la couverture neigeuse, la profondeur de la neige (SD) avec la plus grande précision possible [20] [43] [28].

3.3 Interaction de la neige avec le rayonnement

3.3.1 Domaine de l'optique

Les observations satellitaires dans l'infrarouge et le visible servaient largement à estimer la couverture neigeuse et en produire des cartes diurnes [8] [35] [34]. Au cours des quatre dernières décennies, des produits de la couverture nivale ont été fournis à partir du signal issu du domaine de l'optique notamment le visible [33]. C'est le cas de NOAA (National Oceanic and Atmospheric Administration), depuis 1966, elle a produit des cartes terrestres de l'hémisphère nord pour la couverture nivale en utilisant les images satellitaires issues du visible [33]. En outre, ce qui a encouragé l'utilisation du spectre de l'optique est la bonne résolution spatiale qui est inférieure au kilomètre [4].

Toutefois, le signal dans le domaine de l'optique est trop affecté par les conditions météorologiques, notamment les nuages [4] [8]. C'est pourquoi, il est difficile d'effectuer des acquisitions quotidiennes. Non seulement l'impact des conditions climatiques qui limitent l'utilisation de l'optique, mais encore les conditions d'ensoleillement où les observations ne peuvent pas être effectuées pendant la nuit et que le nombre des heures non ensoleillées dépasse de loin celui des heures ensoleillées durant l'hiver dans les hautes et les moyennes latitudes [4]. Sans oublier les problèmes de nuit polaire où certaines régions polaires peuvent demeurer sans ensoleillement pendant des mois.

3.3.2 Domaine des micro-ondes

Les micro-ondes sont des ondes électromagnétiques dont la longueur varie entre 1 mm (300 GHz) et 30 cm (1 GHz). En d'autres termes, la gamme des micro-ondes correspond approximativement aux longueurs d'onde intermédiaire entre les ondes radio et l'infrarouge bien que les limites entre ces gammes ne

sont pas exactes et dépendent du champ d'étude.

Les émissions et les diffusions issues des manteaux neigeux dans le domaine des micro-ondes sont principalement composées de deux parties : l'émission par le sol sous-jacent et l'émission par les couches de neige [10] [13]. La diffusion et l'absorption de la couche nivale et les propriétés de réflexion et de transmission des interfaces sol-neige et neige-air sont les principaux facteurs dont dépendent les deux contributions susmentionnées. Ces phénomènes dépendent étroitement de la longueur d'onde d'observation.

En d'autres termes, l'émission des micro-ondes à travers les manteaux neigeux dépend des propriétés de la neige. Néanmoins, le signal micro-onde issue d'un manteau neigeux ne dépend pas seulement des propriétés de la neige, mais encore de l'état du sol sous la neige [10].

En effet, la diffusion des micro-ondes à travers les couches nivales est contrôlée par l'épaisseur de ces couches (SD) [13] [9] [33]. La taille des grains de neige et leur forme jouent également un rôle important dans le contrôle de la diffusion [13] [9]. Car la diffusion devient importante pour le rayonnement ascendant quand la taille des grains de neige est proche de la longueur d'onde [13]. Un autre paramètre de la neige qui affecte fortement le signal issu des manteaux neigeux est l'humidité de la surface et la teneur en eau liquide [9] [40]. En effet, la neige humide est un fort émetteur dans le spectre des micro-ondes [33]. Une quantité minimale d'eau liquide dans la neige peut fortement augmenter la température de brillance [41]. D'autres paramètres peuvent également impacter le comportement des micro-ondes dans les couches nivales, à savoir les cycles de fonte et de congélation, la végétation qui la recouvre ou qui s'y trouve et le gradient vertical de température [9].

En conséquence, le signal micro-onde issu des terres recouverts par la neige dépend principalement des propriétés de celle-ci, faisant des micro-ondes un signal incontournable pour restituer les propriétés de la neige sur terre, notamment l'épaisseur et la densité de la neige sur terre. Comme le SWE est proportionnel au produit de SD avec la densité de la neige, il est possible d'inverser le SWE à partir du signal micro-ondes.

Non seulement la dépendance du signal micro-ondes avec les propriétés de la neige qui justifie le recours aux micro-ondes, mais encore la largeur des fenêtres atmosphérique que représentent quelques fréquences des MOP (micro-ondes passives) où seulement les nuages denses (souvent de type cumuliformes)

peuvent interférer avec les observations MOP [28]. En plus, les observations MOP ne dépendent pas de l'éclairement solaire et sont disponibles de jour comme de nuit [28].

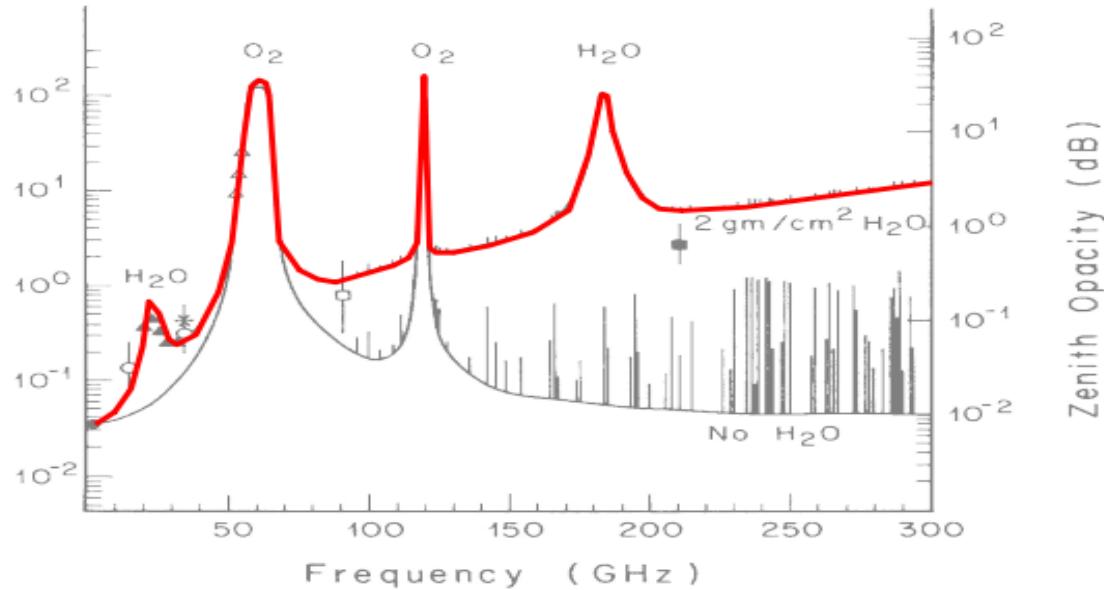


FIGURE 17 – Opacité de l’atmosphère en fonction de longueur d’ondes dans les hyperfréquences

On remarque que certaines bandes de fréquences sont trop atténuerées par l’oxygène et d’autre par la vapeur d’eau. Mais, dans la plage des fréquences inférieures à 100 GHz, on constate l’existence de quelques fenêtres atmosphériques, notamment les fréquences inférieures à 50 GHz à l’exception de la bande centrée sur 25 GHz.

C’est pourquoi, plusieurs études renforcent l’utilisation des MOP pour la récupération des paramètres de la surface avec des précisions importantes. Par exemple, plusieurs algorithmes ont été développés pour la récupération de SD et SWE à partir des observations MOP issues des capteurs comme SMMR, SSM/I, AMSR2, etc. Cependant, il est toujours difficile de déterminer le SWE à partir des observations MOP lorsque la neige est mouillée, mais plutôt il est possible de détecter sa présence en raison de sa différence de polarisation très élevée [33].

De point de vue pratique, seulement quelques canaux de la gamme 6–90 GHz sont utilisés avec les deux polarisations : verticale et horizontale, car elles n’ont pas le même comportement vis-à-vis des propriétés de la neige. Les basses fréquences de la gamme 6–18 GHz sont moins diffusées que les hautes fréquences (fréquences supérieures à 23 GHz), celles-ci sont souvent sensibles à la taille des grains, la densité de la neige et son épaisseur.

3.4 Méthode d'inversion avec les micro-ondes passives issues du capteur AMSR2

L'instrument AMSR2

Le 17 mai 2012, (JAXA a lancé le satellite GCOM-W1 (Global Change Observation Mission 1st Water) dans le but de surveiller le cycle de l'eau. Au bord de ce satellite, il est embarqué le radiomètre à balayage micro-ondes avancé-2 (AMSR2) [7].

Le satellite GCOM-W1 tourne autour de la Terre dans un orbite polaire 233 fois en une période récurrente de 16 jours. Lorsqu'il se trouve au-dessus de l'équateur, l'heure locale est $13 : 30 (1 : 30) \pm 15$ min pour une trajectoire ascendante (descendante), faisant des trajectoires ascendantes des trajectoires diurnes et des trajectoires descendantes des trajectoires nocturnes [7].

AMSR2 est équipé de récepteurs micro-ondes multifréquences à puissance totale ayant la capacité de détecter l'énergie micro-onde (température de brillance) polarisée horizontalement et verticalement provenant de l'atmosphère et de la surface de la Terre. Le capteur AMSR2 mesure les Tbs (températures de brillance) dans les bandes centrées sur les fréquences suivantes : 6.925, 7.3, 10.65, 18.7, 23.8, 36.5 et 89.0 GHz [7].

Les spécifications principales du radiomètre AMSR2 sont présentées dans le tableau suivant :

Frequency [GHz]	6.925	7.3	10.65	18.7	23.8	36.5	89(A)	89(B)
Polarization	Verticale et horizontale							
Band width [MHz]	350	350	100	200	400	1000	3000	3000
-3 dB Beam width [deg]	1.8	1.8	1.2	0.65	0.75	0.35	0.15	0.15
Footprint corresponding to -3 dB Beam width (Cross-track [km] x Along-track [km])	35x62	34x58	24x42	14x22	15x26	7x12	3x5	3x5
Sampling interval [km]	10	10	10	10	10	10	5	5
Swath width [km]	1.450 (effectively more than 1.600)							
Incidence angle	55	55	55	55	55	55	55	54.5
Integration time [msec]	2.6	2.6	2.6	2.6	2.6	2.6	1.3	1.3
Radiometric resolution ($NE\Delta T$ [K] for 1σ at 150k)	≤ 0.34	≤ 0.43	≤ 0.7	≤ 0.7	≤ 0.6	≤ 0.7	≤ 1.2	≤ 1.2

TABLE 3 – Les principales caractéristiques du radiomètre AMSR2 [7]

Le radiomètre à MOP mesure la température de brillance (Tb) qui correspond au signal ascendant brut avec contribution de l'atmosphère et de la surface. Une telle température peut être exprimée sous la forme suivante :

$$Tb = (R Tb_{atmosphere\ downwelling} + (1 - R) Tb_{surf}) e^{-\tau} + Tb_{atmosphere\ upwelling} \quad (5)$$

Pour les fréquences en dessous de 10 GHz, on néglige généralement $Tb_{atmosphere\ downwelling}$ et $Tb_{atmosphere\ upwelling}$, car elles sont très faibles. Pour limiter l'effet de l'atmosphère, notamment l'absorption par la vapeur d'eau et l'oxygène, des fenêtres atmosphériques micro-ondes peuvent être utilisées, ce qui maximise la transmissivité atmosphérique ($e^{-\tau} \sim 1$).

Le problème d'inversion des propriétés de la neige sur terre à partir des MOP

La méthode d'inversion des propriétés de la neige sur terre à partir des MOP est une méthode empirique dont le but est la recherche de corrélation entre l'une des propriétés de la neige et les observations multi-fréquences dans le domaine spectral micro-ondes afin de construire une fonction statistique permettant la restitution des propriétés de la neige connaissant les valeurs des observations MOP par satellite avec la moindre erreur possible, et ce, en utilisant des méthodes statistiques classiques ou des méthodes de ML très avancées selon la nature du problème.

On peut modéliser le problème d'inversion des paramètres de la neige avec l'équation suivant :

$$V_{variable\ restitue} = f(P_1, P_2, \dots, P_n) \quad (6)$$

Avec P_1, P_2, \dots, P_n sont les prédicteurs (les observations multi-fréquences dans le domaine des micro-ondes) et $V_{variable\ restitue}$ est la variable à restituer, elle peut être le SWE, SD, etc.

Les dernières décennies, diverses techniques de résolution des problèmes inverses en télédétection pour la reconstitution de plusieurs variables sur terre ont été utilisées, au niveau de l'océan et dans l'atmosphère [13]. Ce recours vers les observations satellitaires est justifié par plusieurs avantages qu'offrent les observations MOP par satellite contrairement aux observations in situ, notamment la grande couverture spatiale des observations MOP par satellite et la capacité d'observer même dans les zones difficiles d'accès.

3.5 Conclusion

Ce chapitre nous a permis d'explorer les processus de formation de la neige sur terre, sa répartition sur la terre, ses propriétés, les types de la couverture nivale, son impact sur le climat et son importance dans le secteur de l'eau et de l'énergie. En plus, nous avons détaillé les exigences des estimations fiables des produits de la neige sur terre pour les prévisions météorologiques, hydrologiques et les études du climat. Enfin, une étude bibliographique du problème d'inversion et l'interaction de la neige avec le rayonnement, notamment le spectre optique et MOP a été abordée.

On conclut donc que la restitution des produits de la neige sur terre est faisable à partir des observations MOP et plusieurs algorithmes ont été développés pour la restitution de SD et de SWE à partir des MOP par satellite, mais la grande majorité de ces méthodes sont valables pour des régions bien déterminées à des échelles souvent régionales ou au niveau des bassins hydrologiques quand les produits de la neige sont destinés seulement aux applications hydrologiques. Mais, à nos jours, aucune méthode de restitution à l'échelle planétaire n'a montré de satisfaction, surtout pour les besoins dans le domaine de l'hydrologie, et

la restitution des deux paramètres SD et SWE à partir des observations MOP par satellite à l'échelle planétaire ou hémisphérique reste encore un défi. C'est l'objectif de notre étude de tenter d'améliorer ces restitutions.

CHAPITRE 4 :

PRÉTRAITEMENT DES DONNÉES ET LA SÉLECTION DES VARIABLES POTENTIELLES

4.1 Introduction

Dans le présent chapitre, nous présenterons toutes les étapes qui ont été menées pour préparer la base de données servant à faire du ML ainsi que les diverses méthodes utilisées pour sélectionner les variables potentielles (températures de brillance) susceptibles d'expliquer le comportement de SWE et le SD sur terre.

4.2 Prétraitements des données

4.2.1 Pourquoi nettoyer les bases de données ?

Les bases de données brutes sont souvent sous des formats difficiles à manipuler directement, non seulement pour faire du ML, mais encore pour des simples visualisations. D'ailleurs, les données brutes sont souvent dotées des erreurs, des ruptures d'homogénéité, etc. Pour remédier à ce problème, nous procérons au filtrage et la préparation des données, qui est une démarche qui consiste à transformer les données brutes avant l'analyse et la modélisation.

4.2.2 Prétraitements des données AMSR2

Avant de les combiner avec les données ERA5, les données d'observations MOP ont subi trois principaux prétraitements :

1. Récupération des données à partir de la base de données brute.
2. Les valeurs de la latitude et la longitude ont été arrondies à deux chiffres après la virgule pour mettre les données des MOP (micro-ondes passives) et les données de la neige sur la même grille.
3. Elles sont mises dans un tableau sous format de base de donnée relationnelle : date, latitude, longitude, Tb1, Tb2, etc.

4.2.3 Prétraitemet des données ERA5

En ce qui concerne les données ERA5 sur les deux paramètres de la neige SWE et RSN, les valeurs de la latitude et la longitude ont été arrondies à deux chiffres après la virgule. Ensuite, les points dont la valeur de SWE est inférieure à 0.001 m ont été retirées de la base de données. En effet, à ce seuil, supposé, très faible, la neige peut être absente. La suppression de ces points de grilles permet de garder seulement les zones qui ont une forte chance d'être enneigées. Autrement dit, cela consiste à supprimer le bruit. Puis, les points dont la valeur de SWE est supérieure à 0.5 m ont été retirées de la base de données pour éviter les accumulations importantes de la neige sur terre.

4.2.4 Mise des données ERA5 et les observations MOP sur une même grille

Cette étape consiste à combiner les données des températures de brillance avec les données de la neige en les regroupant par les deux variables latitude et longitude. Pour ce faire, il suffisait d'arrondir les valeurs de la variable latitude et longitude à un chiffre après la virgule sans faire recours aux méthodes d'interpolation complexes.

Le tableau suivant résume les caractéristiques statistiques de la base de données qui compte 1 507 413 lignes après les opérations susmentionnées :

Variable	Minimum	Moyenne	Maximum	Ecart-type
RSN (kg/m^3)	100	179.90	450	53.4
SWE (m)	0.001	1.19	10	3.04
tb6v (K)	144.11	247.02	381.63	17.9
tb6h (K)	79.79	227.53	370.15	29.81
tb7v (K)	144.7	248.01	370.15	18.44
tb7h (K)	80.69	228.56	362.53	29.90
tb10v (K)	152.23	247.35	369.63	17.4
tb10h (K)	82.34	227.88	386.04	29.4
tb18v (K)	157.17	244.64	357.91	16.7
tb18h (K)	92.4	225.35	393.58	27.7
tb23v (K)	154.32	241.60	279.91	17.2
tb23h (K)	106.34	224.94	278.54	26.2
tb36v (K)	144.92	229.67	278.88	20.7
tb36h (K)	111.30	215.98	277.40	26.4
tb89v (K)	137.51	218.59	278.73	24.0
tb89h (K)	128.98	209.22	277.26	26.3

TABLE 4 – Description statistique brève de la base de données

4.3 Sélection des variables potentielles

Dans un problème data science, les variables potentielles sont des variables susceptibles d'expliquer le comportement du prédictand. On peut les sélectionner en suivant de nombreuses méthodes, les principales méthodes sont :

Étude des liens physiques entre le prédictand et les prédicteurs

Dans cette méthode, on fait le recours à notre connaissance des liens physiques existants entre le prédictand et la variable potentielle. Dans notre cas, la croissance des grains de neige via le processus du métamorphisme de la neige pendant l'hiver affecte largement les émissivités de la neige et la diffusion à travers celle-ci. Ce phénomène augmente avec la fréquence, puisqu'il est déjà très important à la bande spectrale centrée sur la fréquence 37 GHz, il serait indubitablement important pour les fréquences 36 GHz et 89 GHz [9].

Étude des travaux similaires

Cette méthode consiste à sélectionner des variables potentielles adoptées dans d'autres études et qui ont donné de résultats satisfaisants. Par exemple, dans cet article [6], pour tenir compte de la neige profonde et peu profonde, deux différences de températures de brillances ont été utilisées : tb10v - tb18v et tb18v - tb36v. Ces trois canaux de température de brillance se sont avérés prometteurs pour la restitution de SWE.

La sélection des variables potentielles à travers une étude de corrélation

L'étude de corrélation entre les variables sur une base de données suffisamment grande peut s'avérer encore importante pour la sélection des variables potentielles. Pour ce faire, nous produisons la matrice de corrélation sur toute la base de données et pour toutes les variables susmentionnées :

Matrice de corrélation entre les différentes variables

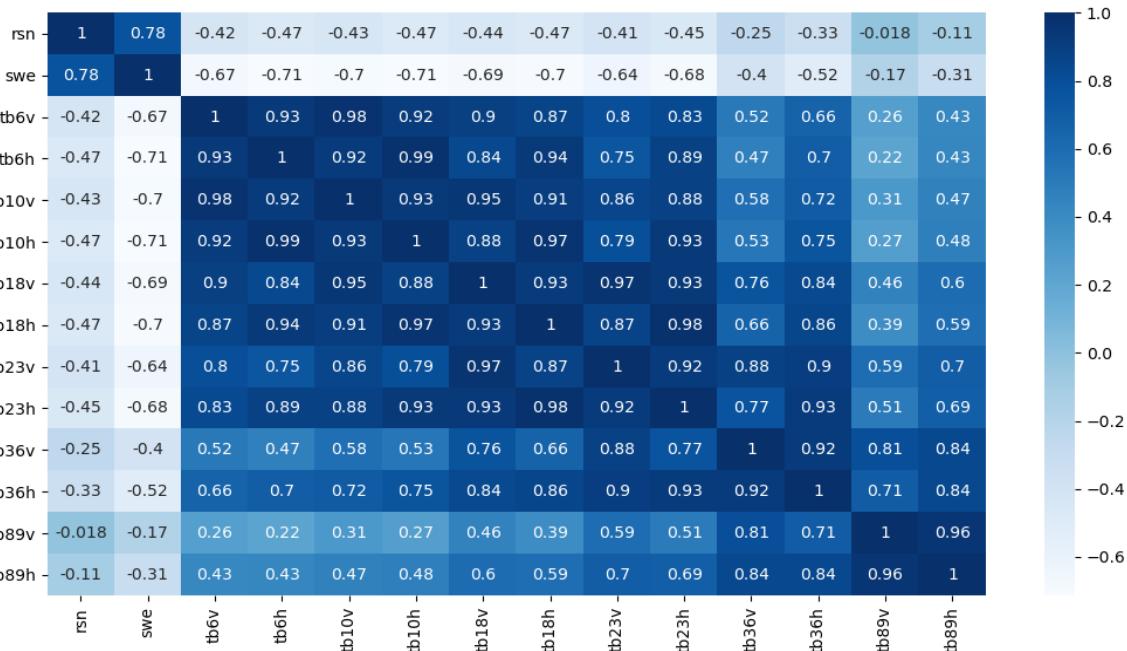


FIGURE 18 – La matrice de corrélation entre les variables élaborée sur l'ensemble de la base de données en filtrant le Groenland et les observations où SWE est inférieur à 0.001 m

On constate ainsi de fortes anticorrélations entre le SWE et les températures de brillance tb6h, tb6v, tb10h, tb10v, tb18h, tb18v, tb23h, tb23v et tb36h. Des anticorrélations moyennes sont observées entre le prédictand SWE et les températures de brillances tb36v, tb89h et tb89v est faiblement anticorrélée au SWE.

En combinant les variables sélectionnées en utilisant les trois méthodes susmentionnées et en prenant en considération les données de températures de brillance fournies par les deux radiomètres AMSR2 et CIMR, nous allons développer deux principaux algorithmes : le premier sera valable pour la restitution des propriétés de la neige à partir des températures de brillance issues des canaux 6v/h, 10v/h, 18v/h, 23v/h, 36v/h, 89v/h du radiomètre AMSR2. Le deuxième algorithme sera développé pour restituer les propriétés de la neige à partir des températures de brillance qui seront fournies par la nouvelle génération des radiomètres, le radiomètre CIMR en utilisant les canaux : 6v/h, 10v/h, 18v/h, 36v/h. Les données servant à l'apprentissage, l'évaluation des modèles et les tests sont toujours les observations fournies par le radiomètre AMSR2.

4.4 Conclusion

Nous avons vu dans ce chapitre qu'il a été important de procéder à un prétraitement de la base de données brute avant de l'exploiter pour explorer des données et faire du ML. Trois méthodes ont été amenées pour sélectionner les variables potentielles susceptibles d'expliquer le comportement des propriétés de la neige sur terre. Ainsi, les prédicteurs qu'il fallait garder pour développer les modèles sont : 6v/h, 10v/h, 18v/h, 23v/h, 36v/h, 89v/h.

CHAPITRE 5 :

QUELQUES TECHNIQUES DU MACHINE LEARNING (ML)

5.1 Introduction

Dans ce chapitre, nous expliquons succinctement quelques notions du Machine Learning et évoquer les diverses techniques qui seront utilisées pour élaborer les modèles de l'estimation des propriétés de la neige. Pour ce faire, nous évoquons quelques généralités sur le Machine Learning. Puis, nous parlons en détail de l'apprentissage supervisé. Ensuite, nous abordons les techniques utilisées pour évaluer et améliorer les modèles. Enfin, nous terminons avec la généralisation des modèles.

5.2 Généralités sur le Machine Learning (ML)

Le Machine Learning (ML) est un champ d'étude très large faisant partie de l'IA (Intelligence Artificielle) qui se base sur des modèles statistiques et des algorithmes dans le but de donner aux ordinateurs la capacité d'apprendre et de se développer pour effectuer automatiquement des tâches spécifiques sans être explicitement programmés et en s'appuyant sur une base de données (des exemples) qui leur permet de s'entraîner [47] [53]. Elle constitue une manière de modéliser des phénomènes, non seulement dans les sciences de la terre, mais encore dans le domaine industriel et commercial afin de prendre des décisions opérationnelles, tactiques ou stratégiques [54].

Le promoteur principal de cette science moderne est la base de données. En effet, cette science se base sur des statistiques des données pour dévoiler des liai-

sons entre deux variables, une variable et plusieurs variables ou bien dévoiler des comportements qui ne sont pas connus auparavant à partir des statistiques sur un ensemble de variables. La multiplication des appareils de mesure et les systèmes informatiques de traitement et de stockage des données qui ont été développées les dernières décennies ont facilité largement la mise en œuvre des bases de données géantes et exploitable servent à développer des algorithmes statistiques pour la modélisation des divers phénomènes.

Le Machine Learning peut être divisé en deux grandes familles : l'apprentissage automatique et la modélisation statistique (régression simple ou multiple). Dans cette dernière, les modèles formés par les statistiques classiques sont des équations mathématiques modélisées en se basant sur des hypothèses posées à priori. Quant à l'apprentissage automatique, on élaborer un modèle sur une base de données indépendante de toute hypothèse, par conséquent la puissance prédictive est plus grande.

Le Machine Learning n'est donc qu'un ensemble d'algorithmes d'apprentissage automatique. Ces derniers peuvent être classés en trois grandes classes :

1. **Les algorithmes d'apprentissage supervisé** : ils sont les plus couramment utilisés. Ils servent à entraîner des modèles pour qu'ils reproduisent un comportement prédéterminé.
2. **Les algorithmes d'apprentissage non supervisé** : ils sont utilisés dans le cas où le résultat attendu n'est pas connu au préalable. Ainsi, le modèle n'a pas une sortie précise et définie au préalable.
3. **Les algorithmes d'apprentissage par renforcement** : ils permettent d'apprendre un comportement à partir d'un ensemble d'observations.

La figure ci-contre illustre les 3 types d'apprentissage automatique : supervisé, non supervisé et par renforcement.

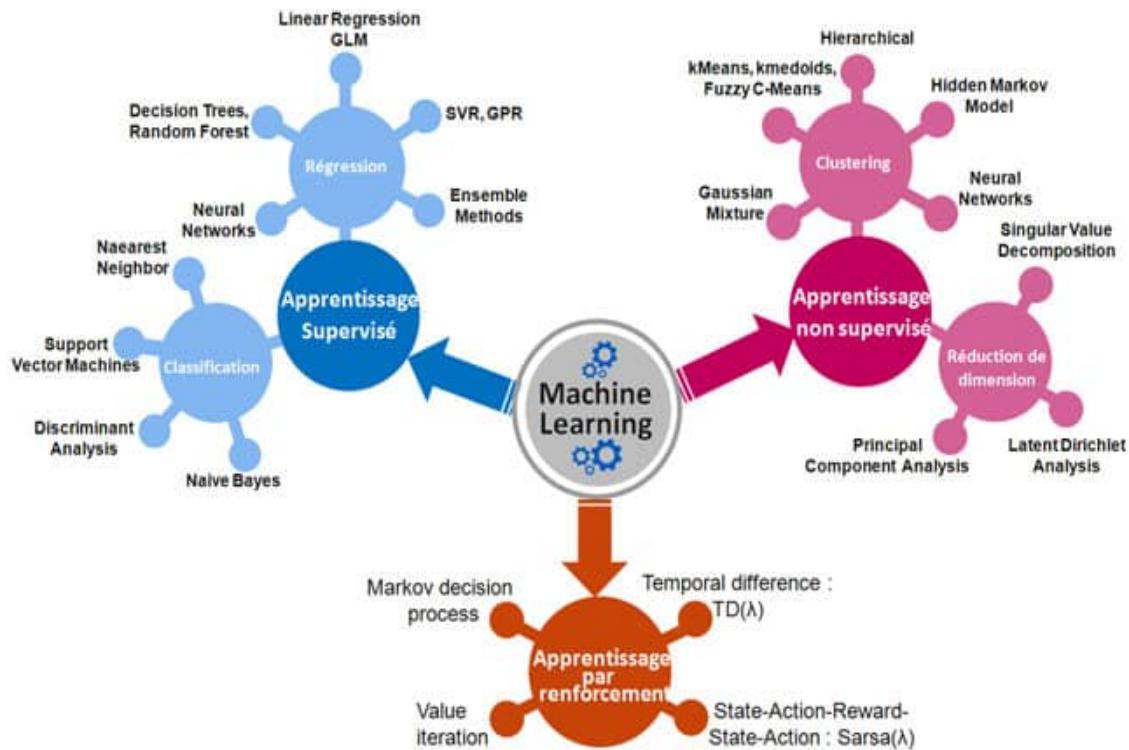


FIGURE 19 – Les trois types d'apprentissage automatique [47]

Le caractère de la variable de sortie que l'on attend du programme nous aidera dans le choix de l'algorithme du ML à utiliser. Quand la variable de sortie est continue (elle peut prendre une infinité de valeurs dans un intervalle.), on est alors confronté à un problème de régression. Quand la variable est discrète (elle ne peut prendre qu'un nombre limité de valeurs dans l'intervalle.), on est confronté à un problème de classification.

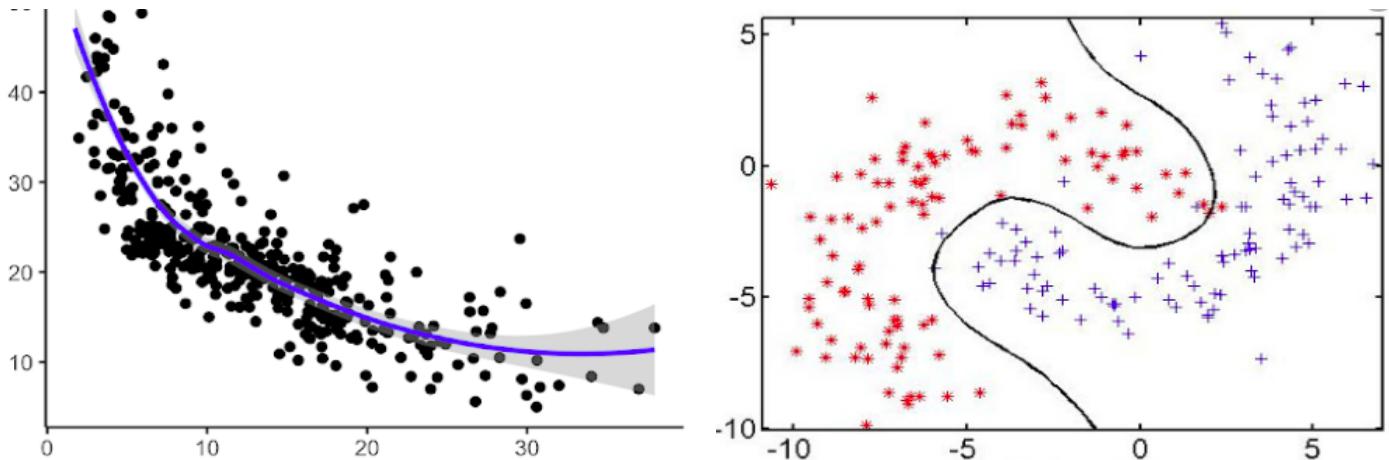


FIGURE 20 – Illustration de la régression et la classification

5.3 Algorithmes de l'apprentissage supervisé utilisés dans la présente étude

5.3.1 Régression Linéaire Multiple (RLM)

La régression linéaire est une méthode servant à modéliser une liaison linéaire entre une variable quantitative (ou prédictand) y , et un ou plusieurs variables explicatives (ou prédicteurs) x_1, x_2, \dots, x_p par ajustement d'une équation linéaire des données observées [49].

On peut résumer le processus de la régression linéaire multiple comme suit :

$$y_j = \sum_{i=1}^p b_i x_{ij} + b_0 + \epsilon(x)_j \quad (7)$$

Où y_j est la j^{eme} observation de la variable expliquée, p est le nombre des variables explicatives, x_{ij} est la j^{eme} observation de la i^{eme} variable explicative, $\epsilon(x)_j$ est le résidu de la j^{eme} observation qui détermine l'information manquante dans l'explication linéaire des valeurs de y_j à partir des $x_{i1}, x_{i2}, \dots, x_{ip}$. Les b_i sont les paramètres qu'on veut estimer de la régression et b_0 est la coordonnée à l'origine.

Afin d'estimer les coefficients (b_0 et b_i) de la régression qui donnent la meilleure approximation de y_j , Saporta suggère l'utilisation d'une méthode dite des moindres carrés qui permet de minimiser la somme des carrés des résidus, de telle sorte que : $\hat{y}_j = b_0 + \sum_{i=1}^p b_i x_{ij}$, où \hat{y}_j est la j^{eme} observation estimée [50].

Une fois les coefficients de régression sont ajustés sur les observations à notre disposition, en connaissant une nouvelle observation des prédicteurs, on peut prédire la valeur de y correspondant à cette nouvelle observation.

Géométriquement, le problème de la régression linéaire multiple consiste à trouver le meilleur hyperplan de dimension $p-1$ qui permet de minimiser la somme des distances entre les points (les enregistrements) et cet hyperplan dans un espace de dimension p .

Toutefois, la régression linéaire a plusieurs limitations liées aux hypothèses sur lesquelles se base la RL (régression linéaire). Certaines hypothèses concernent les variables explicatives et la taille de la base de données et sont appelées les hypothèses structurelles, d'autres concernent les résidus et les observations et sont appelées les hypothèses stochastiques.

Les hypothèses structurelles

- H_1 : absence de colinéarité entre les variables explicatives.
- H_2 : le nombre d'observations est strictement supérieur au nombre de variables.

Les hypothèses stochastiques

- H_3 : les observations de la base de données de l'entraînement sont déterminées sans erreurs. Cette hypothèse n'est jamais vérifiée à 100%.
- H_4 : les erreurs sont linéairement indépendantes des variables explicatives.
- H_5 : les résidus suivent une loi normale multidimensionnelle [49].

Néanmoins, les hypothèses susmentionnées ne sont pas toujours vérifiées dans les bases de données utilisées pour faire de la modélisation. En plus, la régression linéaire permet de modéliser seulement les relations linéaires qui peuvent exister entre les variables alors que les variables peuvent être liées par des relations non-linéaires et qui sont souvent très complexes. D'où le besoin de modélisation avec d'autres méthodes qui peuvent prendre en considération la complexité des données et fusionner l'information. C'est pour cela que d'autres méthodes de modélisation ont été mises en place, à savoir l'apprentissage profond que nous allons évoquer dans le prochain paragraphe.

5.3.2 Apprentissage profond : les réseaux de neurones (RN)

L'apprentissage profond (en anglais Deep Learning) joue un rôle de plus en plus important dans différentes sciences. Il a déjà eu un impact considérable dans des domaines tels que le transport (la conduite autonome des voitures), la médecine (diagnostic du cancer) et la reconnaissance vocale [48]. Son utilisation s'est développée ces dernières années dans plusieurs disciplines, à savoir la météorologie, l'écologie et l'environnement, la biologie et la médecine, etc. D'une manière générale, l'apprentissage profond sert à la résolution des problèmes de classification, de prédiction, de catégorisation ou de reconnaissance des formes et de mémoire associative [55].

5.3.2.1 Définition et fonctionnement de l'apprentissage profond

Le deep learning est une technique d'apprentissage automatique qui permet de modéliser les relations entre les variables avec un très haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non-linéaires. Le Deep Learning repose sur des algorithmes des réseaux de neurones artificiels (profonds), ceux-ci inspirent leur fonctionnement de la structure du système nerveux humain. Il se compose d'un ensemble d'unités de traitement organisées en couches d'entrée (prédicteurs), cachées et de sortie (prédictand). Chaque couche est composée d'un nombre d'unités ou nœuds appelés les neurones, ceux-ci sont connectés aux nœuds des couches adjacentes par des connexions. Chaque connexion a un poids (valeur numérique). Pour chaque unité, les entrées sont multipliées par des poids respectifs et additionnées. Ensuite, la somme subit une transformation par une fonction d'activation [48].

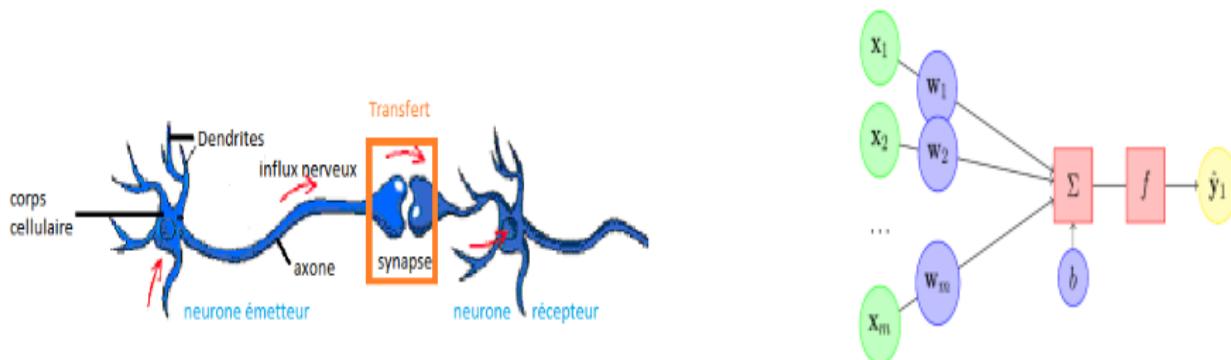


FIGURE 21 – Illustration simple d'un neurone biologique et artificiel

Le neurone biologique est une cellule qui se compose des dendrites, des synapses, d'axones et de noyaux qui activent les sorties en fonction des stimulations en entrée.

Dans la figure ci-dessus, le rectangle orange indique la région où les deux neurones communiquent. Un neurone reçoit des signaux d'autres neurones à travers les dendrites. Ces dernières sont caractérisées chacune par un poids, appelé poids synaptique, celui-ci est multiplié par le signal entrant. Les signaux des dendrites sont accumulés dans le corps cellulaire, et si la force du signal résultant dépasse un certain seuil, le neurone s'active (transmet le message à l'axone). Sinon, le neurone reste désactivé, le signal ne se propage pas à travers ce neurone. Ainsi, on peut définir le potentiel d'activation comme une variation de la force du signal qui détermine si la communication doit se faire ou non.

L'analogie entre un neurone artificiel et biologique

Neurone biologique	Neurone artificiel
Les dendrites	Les entrées du neurone
Le cône d'émergence (point de départ de l'axone)	La sortie du neurone
Les actions excitatrices et inhibitrices des synapses	Les coefficients numériques (les poids synaptiques) associés aux entrées

TABLE 5 – Analogie entre un neurone artificiel et biologique

Les coefficients d'un réseau de neurones sont ajustés lors d'une phase d'apprentissage. Dans la version basique d'un neurone, celui-ci calcule la somme des entrées reçues pondérées par les poids appropriés, à cette somme s'ajoute le biais du neurone. Ainsi, la sortie du neurone est le résultat de la fonction d'activation du neurone appliquée à la somme susmentionnée à laquelle le biais a été ajouté.

La fonction d'activation f est une fonction mathématique, souvent non-linéaire, appliquée à un signal pour produire la sortie d'un neurone. Les fonctions d'activation les plus utilisées sont illustrées dans le tableau ci-dessous [51]. Les fonctions illustrées dans le tableau ci-contre sont les plus utilisées parce qu'elles ont une dérivée mathématiquement favorable, ce qui facilite le calcul des dérivées partielles de l'erreur δ par rapport aux poids individuels [48].

Fonction	Définition	Description	Représentation graphique	Intervalle de définition
Tangente hyperbolique	$\frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}}$	Elle est convenable et idéale pour les perceptrons multicouches, notamment pour les couches cachées		[-1, 1]
Sigmoïde	$\frac{1}{1+e^{-x}}$	Une courbe en "S", permet d'exprimer la valeur d'entrée sous forme de probabilité		[0, 1]
ReLU	$\max(0, \alpha)$	Elle permet de pallier le problème de saturation des deux fonctions susmentionnées, comme elle permet d'augmenter la convergence		[0, +∞]

TABLE 6 – Les fonctions d’activation les plus utilisées pour les neurones cachés et les neurones de sortie [56]

Deux principaux algorithmes sont nécessaires pour entraîner un modèle :

La propagation en avant (Forward-Propagation) :

La couche d’entrée dont le nombre des neurones est égal au nombre des variables explicatives reçoit les données d’apprentissage, puis renvoie les résultats aux couches cachées, puis la dernière couche cachée renvoie le résultat à la couche de sortie qui donne la première estimation de la variable ou les variables expliquées. Le résultat de sortie est ensuite comparé aux observations réelles de (des) variable(s) expliquée(s). Une telle comparaison consiste à calculer l’erreur que nous cherchons à minimiser. Si l’erreur obtenue n’est pas suffisamment faible, une propagation en arrière est nécessaire pour ajuster les poids du réseau afin de minimiser davantage l’erreur d’estimation.

La propagation en arrière (Backward-Propagation) :

La propagation en arrière sert à modifier les valeurs des poids afin de minimiser l'erreur obtenue avec la propagation en avant. C'est là où nous introduisons la notion de descente de gradient. Cette alternance de propagation en avant et en arrière continue tant que la condition d'arrêt n'est pas encore vérifiée. La condition d'arrêt peut être le nombre d'itérations que l'on fixe ou bien la condition sur le seuil de l'erreur que l'on veut atteindre.

5.3.2.2 Algorithmes d'optimisation

La phase d'apprentissage du réseau de neurones est une phase dans laquelle les poids des connexions sont adaptés selon les erreurs de prédiction obtenues à chaque itération.

La méthode la plus utilisée pour l'adaptation des poids est la rétro-propagation du gradient (backpropagation) [57]. La rétro-propagation détermine le gradient de l'erreur pour chaque neurone du réseau en partant de la dernière couche et en arrivant jusqu'à la première couche cachée [52]. Il s'agit ainsi d'un problème d'optimisation qui utilise la méthode de descente de gradient qui a montré de grandes performances d'optimisation d'un réseau neuronal. Nous évoquerons dans les prochains paragraphes quelques notions de cette méthode d'optimisation.

La descente du gradient est un algorithme d'optimisation très général et efficace qui permet de trouver des solutions optimales à un grand nombre de problèmes. Plus précisément, cet algorithme consiste à chercher la valeur de x qui minimise la fonction coût $E(\theta)$, avec : $\frac{\partial E}{\partial \theta} = 0$, vu qu'il n'y a pas de solution analytique au problème de minimisation, l'optimisation se fait de manière itérative en utilisant des méthodes numériques. La plupart de ces méthodes reposent sur l'algorithme de descente du gradient suivant [58] :

Algorithm 1 : Algorithmes de la descente de gradient.

Données : une base de données B où B_t est le t^{eme} couple (x, c) et dont la taille est N.
Entrées : une fonction CritereArret() qui renvoie vraie quand l'apprentissage est terminé.
 Un réseau de neurones muni d'une collection de paramètres θ .
 Un hyper-paramètre α donnant la vitesse d'apprentissage de l'algorithme.

```

tant que CritereArret( $\theta$ ) faire
    pour chaque couple ( $x, c$ ) de B faire
         $y \leftarrow Propagation(x);$                                 /* La prdiction du rseau */
         $J_t \leftarrow Retropagation(y - c);$                       /* Le teme gradient */
    finprch
     $J \leftarrow \frac{1}{N} \sum_{t=1}^N J_t;$                       /* Le gradient moyen */
     $\theta \leftarrow \theta - \alpha J;$                             /* Mise à jour des paramètres */
     $E \leftarrow \frac{1}{N} \sum_{t=1}^N \|y_t - c_t\|_2^2;$           /* Erreur moyenne de prédiction */
fintq
```

Sorties : Un réseau de neurones ayant des paramètres θ optimisés.

TABLE 7 – Algorithme général de la descente de gradient

L'algorithme initialise les valeurs des θ_k aléatoirement, puis les améliore progressivement en direction du gradient descendant, en tentant à chaque étape de faire décroître la fonction coût (MSE), jusqu'à ce que l'algorithme converge vers un minimum local ou global.

Un autre paramètre très important dans l'optimisation est le taux d'apprentissage qu'on doit ajuster dans l'algorithme de descente du gradient. En effet, à chaque itération, l'algorithme de descente de gradient multiplie le taux d'apprentissage (α_k) par le gradient. Le produit ainsi généré est appelé le pas de gradient. Selon le taux d'apprentissage choisi, on peut se retrouver avec deux situations illustrées dans la figure ci-dessous :

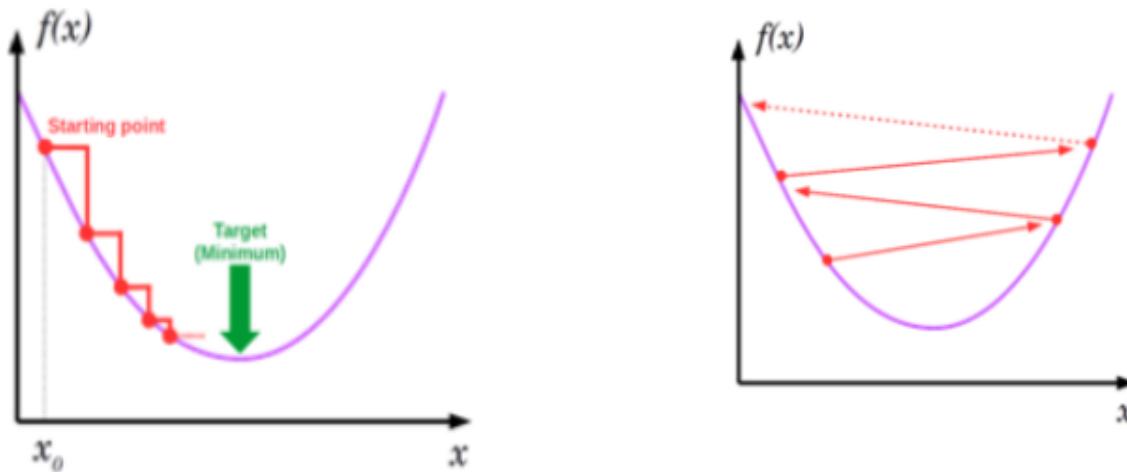


FIGURE 22 – Descente de gradient avec taux d'apprentissage élevé (à droit) et taux d'apprentissage faible (à gauche)

D'après la figure ci-dessus, on constate que si le taux d'apprentissage est trop élevé, on risque de sauter le point le plus bas et de se retrouver de l'autre côté. Autrement dit, on risque de faire diverger l'algorithme, avec des valeurs de plus en plus grandes, ce qui va ralentir, voire même empêcher la découverte d'une solution optimale.

Néanmoins, si le taux d'apprentissage est trop petit, l'algorithme devra tourner un grand nombre de fois avant de converger, ce qui augmente le temps de calcul.

À ces deux contraintes s'ajoute une troisième contrainte. En fait, la fonction coût n'a pas la forme d'une cuvette régulière, comme dans la figure susmentionnée, mais en revanche, il peut y avoir des crêtes, des trous, des plateaux et toutes sortes de terrains irréguliers, ce qui ralentit et complique encore la convergence vers le minimum global voire même vers un minimum local. Pour pallier ce problème, on a souvent recours à l'algorithme de **descente de gradient par lot**. En effet, la descente de gradient par lot consiste en une minimisation d'une fonction qui peut être écrite sous la forme de la somme de fonctions différentiables. Ce processus de minimisation est réalisé de manière itérative sur des lots de données tirés aléatoirement. L'utilisation de cette méthode de minimisation de la fonction coût, donne comme résultat une approximation de la fonction coût global.

La descente de gradient par lot

L'entraînement des modèles de RN par lot améliore considérablement la vitesse d'apprentissage en utilisant un échantillon de données d'entraînement (x, c) au

lieu de la base de données entière où x représente un vecteur d'entrée du lot et c représente la cible.

Cette méthode peut être utile et efficace pour avancer plus rapidement. Toutefois, l'estimation de la moyenne du gradient reste sans biais alors que la variance de l'estimation de la moyenne augmente, car les calculs se font sur un échantillon [58]. Une méthode d'accélération du calcul surtout en utilisant l'algèbre linéaire optimisée, consiste à passer à la propagation, tous les vecteurs d'entrée d'un lot à la fois, sous forme de matrice, ce qui permet un calcul direct et une seule fois pendant une seule propagation et rétro-propagation. C'est-à-dire que dans ce cas, un lot est un couple (x, c). C'est une matrice de dimension $n \times 1$ composée de l vecteurs et d'un vecteur c respectivement.

La méthode de descente du gradient par lot est décrite par l'équation suivante :

$$\theta_{t+1} = \theta_t - \alpha \frac{1}{B_s} \sum_{i=1}^{B_s} \nabla_{\theta} J_t^i(\theta_t, x_t^i, y_t^i) \quad (8)$$

Avec :

B_s la taille du lot.

$B_t = (x_t^i, y_t^i)_i \in [1, B_s]$ le lot de données tiré à l'étape t .

x_t^i et y_t^i sont respectivement la donnée de la vérité terrain et la donnée estimée dans ce lot.

$J_t = \frac{1}{B_s} \sum_{i=1}^{B_s} J_t^i$ est l'approximation stochastique de la fonction coût globale à l'étape t sur le lot B_s décomposée en somme de fonctions différentiables J_t^i liées à chaque paire (x_t^i, y_t^i) .

L'algorithme suivant décrit le fonctionnement de cette méthode [58] :

Algorithm 2 : Algorithmes de la descente de gradient par lot pour l'apprentissage d'un réseau de neurones.

Données : une base de données B où B_t est le t^{eme} couple (x, c) et dont la taille est N .
Entrées : une fonction CritereArret() qui renvoie vraie quand l'apprentissage est terminé.
 Un réseau de neurones muni d'une collection de paramètres θ .
 Un hyper-paramètre α donnant la vitesse d'apprentissage de l'algorithme.

```

tant que CritereArret( $\theta$ ) faire
     $B_{lot} \leftarrow$  couple  $(x, c)$  tirés au hasard et sans remise de  $B$ ;
    pour chaque couple  $(x, c)$  de  $B_{lot}$  faire
         $y \leftarrow Propagation(x);$  /* La prdition du rseau */
         $J_t \leftarrow Retropagation(y - c);$  /* Le teme gradient */
    finprch
     $J \leftarrow \frac{1}{N} \sum_{t=1}^N J_t;$  /* Le gradient moyen */
     $\theta \leftarrow \theta - \alpha J;$  /* Mise à jour des paramètres */
     $E \leftarrow \frac{1}{N} \sum_{t=1}^N \|y_t - c_t\|_2^2;$  /* Erreur moyenne de prédiction */
fintq
Sorties : Un réseau de neurones ayant des paramètres  $\theta_k$  optimisés.

```

TABLE 8 – Algorithme de la descente de gradient pour l'apprentissage d'un réseau de neurones par lot

La descente de gradient stochastique

On parle d'algorithmes de descente de gradient stochastique ou en ligne lorsque la taille du lot dans l'algorithme de descente de gradient par lot, égale à 1. Ce qui justifie la différence entre les deux méthodes est la méthode du choix du couple individuel (x, c) utilisé dans le calcul du gradient. En effet, si le couple (x, c) est choisi dans l'ordre des données d'apprentissage, on parle de la descente de gradient en ligne, alors qu'il s'agit d'une descente de gradient stochastique lorsque un tel choix est fait de manière aléatoire.

Extensions de la descente de gradient

La descente de gradient moyennée est une méthode proposée et étudiée par Polyak [59]. Elle consiste à remplacer le calcul du paramètre θ_t , par le calcul de la moyenne temporelle de ses valeurs, et ce, à partir des mises à jour obtenues par la descente du gradient :

$$\bar{\theta}_t = \frac{1}{t} \sum_{t=0}^T \theta_t \quad (9)$$

Adagrad : cette méthode a été proposée en 2011 [60], elle consiste à ajuster le taux d'apprentissage en fonction de "l'éparsité" des paramètres du réseau. Adagrad abaisse le taux d'apprentissage, selon la pente, en effet, dans le cas d'une

pente plus prononcée, le taux d'apprentissage s'affaiblit plus rapidement que dans le cas d'une pente douce, mathématiquement le pas est décrit par :

$$\forall i \quad (\theta_{t+1})_i = (\theta_t)_i - \alpha \frac{(\nabla J(\theta_t))_i}{\sqrt{\sum_{k=1}^t (\nabla J(\theta_k))_i^2}} \quad , \alpha \geq 0 \quad (10)$$

L'un des principaux avantages d'Adagrad est qu'il évite d'ajuster manuellement le taux d'apprentissage, que la plupart des algorithmes prennent par défaut égale à 0.01. Néanmoins, sa principale faiblesse est son accumulation des gradients carrés dans le dénominateur : la somme cumulée ne cesse de croître pendant l'entraînement, ce qui réduit considérablement le taux d'apprentissage jusqu'à devenir infiniment petit, par conséquent, l'algorithme devient incapable d'acquérir des connaissances supplémentaires. Les algorithmes suivants permettent de résoudre ce problème.

RMSProp : cette méthode est proposée par Geoff Hinton, elle fonctionne de la même manière que la méthode d'Adagrad, sauf qu'elle ne cumule pas les gradients des itérations récentes, elle utilise donc une moyenne glissante [61] :

$$(\nabla_{t+1})_i = \delta(\nabla_t)_i + (1 - \delta)(\nabla J(\theta_t))_i^2 \quad (11)$$

$$\forall i \quad (\theta_{t+1})_i = (\theta_t)_i + \alpha \frac{\nabla J(\theta_t)}{\sqrt{(\nabla_{t+1})_i}} \quad , \alpha \geq 0 \quad (12)$$

avec $\delta(\nabla_t)_i$ est la moyenne quadratique glissante du gradient. La division du gradient de la fonction objective par la racine de la moyenne quadratique glissante (c'est-à-dire l'amplitude.) améliore la convergence.

Hinton suggère que δ soit défini par 0.9, alors qu'une bonne valeur par défaut pour le taux d'apprentissage α est 0.001 [61].

Adam : cet algorithme a été développé dernièrement, sa particularité est qu'il a une grande capacité d'optimisation par descente de gradient. Son principe de fonctionnement est le même que celui de RMSProp et Adagrad. Autrement dit, il adapte automatiquement son taux d'apprentissage pour chaque paramètre. Cependant, sa particularité est qu'il calcule de plus (m_t, v_t) : des estimations adaptatives du moment. C'est une généralisation de l'algorithme d'Adagrad [62].

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla J(\theta_t) \quad (13)$$

$$\nu_t = \beta_2 \nu_{t-1} + (1 - \beta_2) \nabla^2 J(\theta_t) \quad (14)$$

m_t et ν_t sont des estimations du premier moment (la moyenne) et du deuxième moment (variance non centrée) des gradients, d'où le nom de la méthode. Comme m_t et ν_t sont initialisés en tant que vecteurs de 0, les auteurs d'Adam observent qu'ils sont biaisés vers zéro, en particulier pendant les pas de temps initiaux, et en particulier lorsque les taux de décroissance β_1 et β_2 sont faibles, c'est-à-dire sont proches de 1. Ils corrigent ces biais en calculant des estimations de premier et second moment :

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (15)$$

$$\hat{\nu}_t = \frac{\nu_t}{1 - \beta_2^t} \quad (16)$$

Ces moments sont utilisés ensuite pour mettre à jour les paramètres, ce qui donne la règle de mise à jour Adam suivant :

$$\theta_{t+1} = \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{\nu}_t}} \quad (17)$$

β_1 et β_2 sont des nouveaux hyperparamètres introduits dans l'algorithme. Ils ont de très bonnes valeurs par défaut de 0.9 et 0.999 respectivement. Ainsi, les expériences montrent qu'Adam est meilleur par rapport aux méthodes d'apprentissage adaptatif ainsi présentées [61].

La méthode Quasi-Newton

Cet algorithme nécessite la connaissance des valeurs de la fonction objectif à minimiser et du vecteur des dérivées premières au point solution courante. Le vecteur des dérivées est évalué analytiquement. Avec les méthodes Quasi-Newton, on cherche à généraliser la formule itérative de Newton :

$$x^{k+1} = x^k - \alpha^k [\nabla^2 f(x^k)]^{-1} \nabla f(x^k) \quad (18)$$

Le principe des méthodes de résolution de type Quasi-Newton est de générer une séquence de matrices symétriques définies positives qui soient des approximations améliorées de la matrice Hessienne réelle ou de son inverse. Sois H_k la suite d'approximation de l'inverse de la matrice Hessienne $[\nabla^2 f(x^k)]^{-1}$ et B_k la suite des estimations de la matrice Hessienne $\nabla^2 f(x^k)$. La méthode consiste à chercher une méthode, telle que, dans le cas d'un problème quadratique, la matrice B_k converge vers la valeur exacte des dérivées secondes (constantes dans ce cas), et qu'enfin, une convergence de type newton prend lieu. Si la méthode est appliquée à une fonction quelconque, B_k peut être considérée à chaque instant, comme une approximation (définie, positive.) du Hessien.

Algorithme de BFGS

L'algorithme de BFGS (du nom de ses inventeurs : Broyden, Fletcher, Goldfarb et Shanno) est une règle d'ajustement des paramètres qui a l'expression suivante [63] :

$$\theta^k = \theta^{k-1} \pm \mu_k M_k \nabla J(\theta^{k-1}) \quad (19)$$

où M_k est une approximation calculée itérativement, de l'inverse de la matrice Hessienne. L'approximation de l'inverse du Hessien est modifiée à chaque itération selon la règle suivante :

$$M_k = M_{k+1} + \left[1 + \left(\frac{\gamma_{k-1}^T M_k \gamma_{k-1}}{\gamma_{k-1}^T \delta_{k-1}^T} \right) \right] \frac{\delta_{k-1}^T \delta_{k-1}}{\delta_{k-1}^T \gamma_{k-1}} \pm \frac{\delta_{k-1} \gamma_{k-1}^T M_{k-1} + M_{k-1} \gamma_{k-1} \delta_{k-1}^T}{\delta_{k-1}^T \gamma_{k-1}} \quad (20)$$

Avec $\gamma_{k-1} = \nabla J(\theta^k) \pm \nabla J(\theta^{k-1})$ et $\delta_{k-1} = \theta^k \pm \theta^{k-1}$ La valeur initiale de M est la matrice identité. Si à une itération, la matrice calculée n'est pas définie positive, elle est réinitialisée à la matrice identité.

5.4 Évaluation des modèles

Pour juger les performances des modèles et leur capacité à être généralisés, on utilise plusieurs indices numériques ou graphiques calculés sur un ensemble de données qui ne faisait pas partie de l'ensemble ayant servi à l'apprentissage, cet ensemble est dite : l'ensemble de tests. Cette partie évoquera un ensemble des indices utilisés dans des études similaires.

5.4.1 Èrreur absolue moyenne (MAE)

L'erreur absolue moyenne mesure l'ampleur moyenne des erreurs dans un ensemble de prédiction, sans tenir compte de leur direction. C'est la moyenne sur l'échantillon de test des différences absolues entre la prédiction et l'observation réelle lorsque toutes les différences individuelles ont le même poids.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (21)$$

5.4.2 Èrreur quadratique moyenne (MSE)

C'est un critère quantitatif très utilisé pour comparer les valeurs prédictives \hat{y}_i avec les observations y_i . Pour chaque observation, la différence entre la valeur donnée par le modèle et la valeur observée est calculée, ensuite le carré de la différence est moyenné sur toutes les observations.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (22)$$

5.4.3 Racine de l'erreur quadratique moyenne (RMSE)

Le RMSE est la racine carrée du MSE. Elle est introduite pour que l'échelle des erreurs soit la même que celle de la variable cible.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} = \sqrt{MSE} \quad (23)$$

Le RMSE, le MAE et le MSE varient entre 0 et $+\infty$, ils sont indifférents à la direction de l'erreur et plus l'erreur est faible plus le modèle est performant.

Le RMSE donne plus de poids aux grandes erreurs (le carré), plus celles-ci sont importantes plus la courbe du RMSE s'éloigne du MAE.

5.4.4 Coefficient de détermination (R^2)

Le coefficient de détermination ou l'efficience est un autre critère utilisé pour mesurer l'accord entre les valeurs réelles et les valeurs prédictes.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} = 1 - \frac{MSE(\text{modèle})}{MSE(\text{moyenne})} \quad (24)$$

Ce coefficient permet de conclure à quel point le modèle s'approche du modèle simpliste qui associe à chaque valeur réelle y_i sa valeur moyenne \hat{y}_i .

Ce coefficient est compris entre 0 et 1, et croît avec le potentiel prédictif du modèle.

5.4.5 Coefficient de corrélation (r)

Ce coefficient permet de détecter la présence ou l'absence d'une relation linéaire entre deux caractères quantitatifs continus.

Le coefficient de corrélation de deux caractères X et Y est égal à la covariance

de X et Y divisée par le produit des écarts-types de X et Y.

$$r = \frac{COV(X, Y)}{\sigma_x \sigma_y} \quad (25)$$

5.4.6 Analyse graphique de sensibilité du modèle

L'analyse des performances d'un modèle peut se faire également grâce à plusieurs types de graphes, à savoir les nuages de points des valeurs prédictives/observées, les graphes de la densité des noyaux en 2D (en anglais 2-D kernel density plot), les histogrammes, les boîtes à moustaches (en anglais box plots), les cartes de chaleur (en anglais heat maps), etc.

5.5 Généralisation, sur-ajustement et sous-ajustement

La généralisation dans les modèles est très importante. En effet, il détermine à quel point le modèle pourrait s'appliquer aux nouvelles données qu'il n'a pas rencontrées pendant l'entraînement. Ainsi, le but d'un modèle n'est pas seulement un bon ajustement, mais encore une bonne généralisation ce qui permet au modèle de faire des prédictions sur de nouvelles données, et c'est l'objectif principal de la modélisation statistique.

Le sur-ajustement se produit quand le modèle se rapproche beaucoup de la fonction jusqu'au point où il donne trop d'attention au bruit. Le modèle apprend la relation avec trop de détails, et même le bruit dans les données est modélisé. Par conséquence, lorsque le modèle est confronté aux nouvelles données, il ne pourra pas généraliser les prédictions, car il dépend trop des données d'apprentissage. On peut ainsi se retrouver avec des modèles qui sont irréalistes.

Le sous-ajustement est le contraire du sur-ajustement. Il se produit lorsque le modèle ne s'ajuste pas assez à la fonction. Il est donc incapable de capter la tendance des données. Autrement dit, il a tendance à ignorer une grande partie

des informations. Lorsqu'il reçoit de nouvelles données, il ne pourra pas prédire avec une bonne précision, car il est appliqué aux données comportant des situations auxquelles il n'a pas été confronté.

Pour remédier à ces problèmes, une des solutions les plus connue pour éviter le sur-ajustement est la **validation croisée**. Elle consiste en une division des données d'entraînement en différents sous-ensembles qu'on utilise pour entraîner et tester le modèle plusieurs fois. Ensuite, on calcule une estimation de la performance sur les différents entraînements. Pour éviter le sous-ajustement, il suffit d'avoir une base de données très large permettant de capturer toutes les fluctuations possibles du phénomène étudié.

5.6 Conclusion

Dans ce chapitre, nous avons donné un aperçu général sur le ML et ses applications. La modélisation linéaire, l'apprentissage profond et les étapes de modélisation ont été abordés d'un point de vue théorique. Néanmoins, le ML est un domaine large et nous ne pouvions pas évoquer toutes les techniques, notamment celles de l'apprentissage non supervisé et l'apprentissage par renforcement, mais seulement quelques techniques d'apprentissage supervisée qui seront utilisées dans notre problématique, notamment la RLM et l'apprentissage profond (DNN). Dans le prochain chapitre, nous détaillerons les principaux résultats de la modélisation des propriétés de la neige sur terre à partir des températures de brillance micro-ondes en utilisant les deux techniques précitées.

CHAPITRE 6 :

LA MODÉLISATION DE SWE SUR TERRE À PARTIR DES TEMPÉRATURES DE BRILLANCE

6.1 Introduction

Une fois terminées les étapes précédentes (l'étude bibliographique, la préparation des données et la sélection des variables potentielles), nous entamons l'étape la plus cruciale et la plus passionnante, c'est la modélisation de SWE sur terre. Pour ce faire, les principaux résultats de la modélisation de SWE à partir des températures de brillance micro-ondes en utilisant la régression linéaire multiple et les réseaux de neurones sont détaillés dans cette partie.

6.2 Modélisation de SWE avec la RLM

En pratique, nous visons à développer deux algorithmes de restitution de SWE à partir des Tbs micro-ondes passives à l'échelle hémisphérique, l'un serait destiné à la restitution de SWE à partir des Tbs micro-ondes passives issues du radiomètre AMSR2, l'autre sera développé pour la restitution de SWE à partir des Tbs issues du future radiomètre CIMR qui sera bientôt lancé par l'agence spatiale européenne.

6.2.1 Modélisation de SWE à partir des Tbs micro-ondes issues de AMSR2

Dans ce cas, nous avons à notre disposition 12 températures de brillance comme variables d'entrée : Tb6v, Tb6h, Tb10v, Tb10h, Tb18v, Tb18h, Tb23h, Tb23v, Tb36v, Tb36h, Tb89h et Tb89v.

Avant de procéder à la modélisation, nous avons divisé aléatoirement la base de données en deux parties, une partie pour l'entraînement avec 80% de la totalité des données et le reste pour l'évaluation du modèle. En plus, nous avons choisi les données journalières pour trois jours : 01-11-2018, 01-01-2019 et 20-03-2019 pour envisager des tests.

La restitution de SWE à partir des Tbs micro-ondes passives précitées en utilisant la régression linéaire multiple a donné de mauvais résultats (voir la figure 23). On constate une forte sous-estimation des valeurs de SWE qui sont supérieures à 0.05 m. En plus, un coefficient de détermination trop faible.

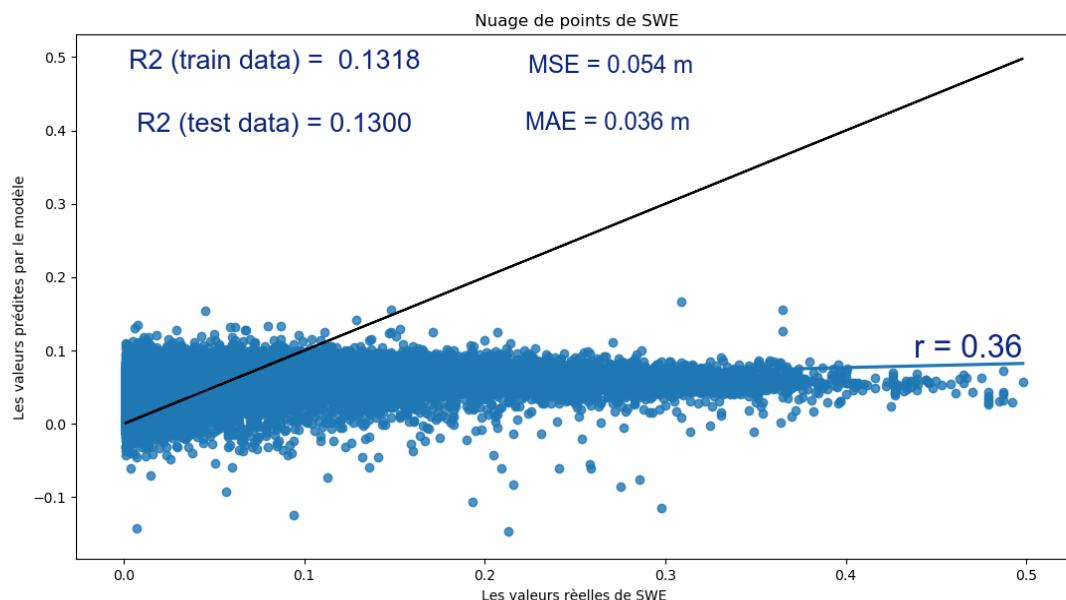


FIGURE 23 – La courbe de dispersion du modèle RLM développé pour la restitution de SWE à partir des Tbs micro-ondes AMSR2

Ces résultats sont également confirmés par le nuage de points observés et prédits par le modèle (voir la figure 24) :

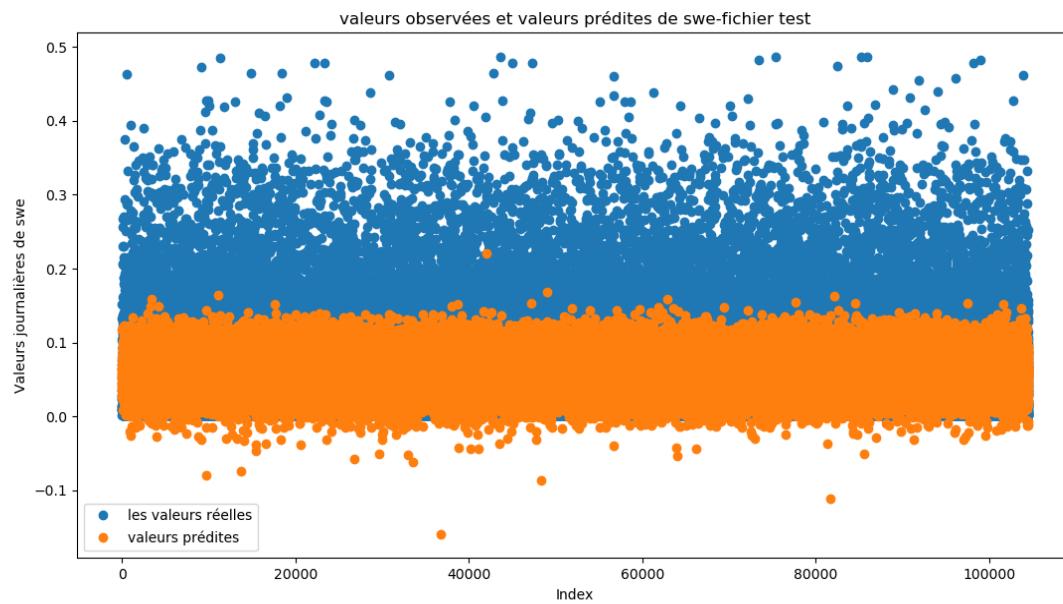


FIGURE 24 – Nuage de points prédits-observés du modèle RLM développé pour la restitution de SWE à partir des Tbs micro-ondes AMSR2

On remarque que presque la totalité des valeurs de SWE qui sont supérieures à 0.13 m sont mal prédites. D'après la figure ci-dessus, on peut suggérer que les valeurs élevées de SWE affectent la qualité du modèle.

Pour bien affiner encore l'étude de la sensibilité du modèle, nous allons visualiser la densité des noyaux des différences entre les valeurs prédites et les valeurs issues des réanalyses ERA5 en fonction des valeurs de SWE issues des réanalyses ERA5 (voir la figure 25).

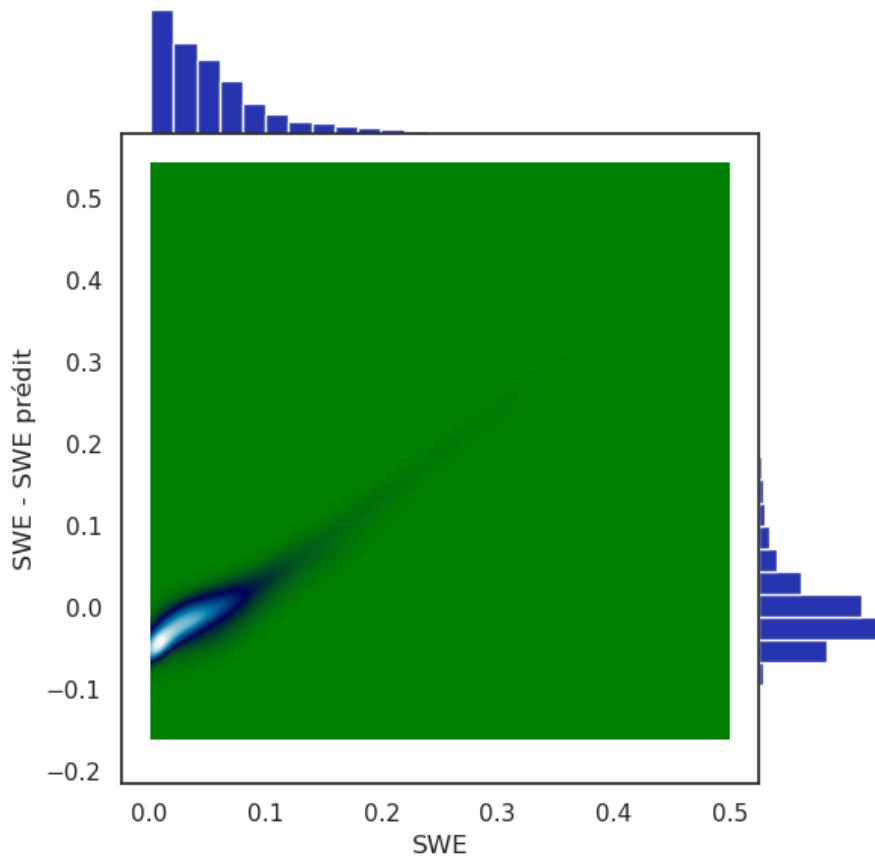


FIGURE 25 – Différences entre les valeurs prédites et les valeurs issues des réanalyses ERA5 en fonction des valeurs de SWE issues des réanalyses ERA5

La figure ci-dessus confirme encore que les valeurs élevées de SWE sont sous-estimées. Cette figure montre en plus qu'une grande partie des valeurs faibles de SWE sont sur-estimées.

Ainsi, on peut affirmer que la RLM n'est pas adaptée à ce problème puisqu'elle n'a pas pu modéliser une partie non-négligeable de la variabilité contenue dans les données de notre base de données. Pourtant, les coefficients de détermination sur la base de données d'apprentissage et celle de test ont été proches l'un de l'autre (voir la figure 23), cela montre que la RLM ne modélise que les relations réelles qui peuvent exister entre les variables. Autrement dit, la RLM ne modélise pas le bruit.

Un tel modèle a été appliqué à trois jeux de données journalières pour produire des cartes de SWE prédites pour trois jours dans la période d'étude afin de comparer le SWE prédict avec le SWE issu des données ERA5.

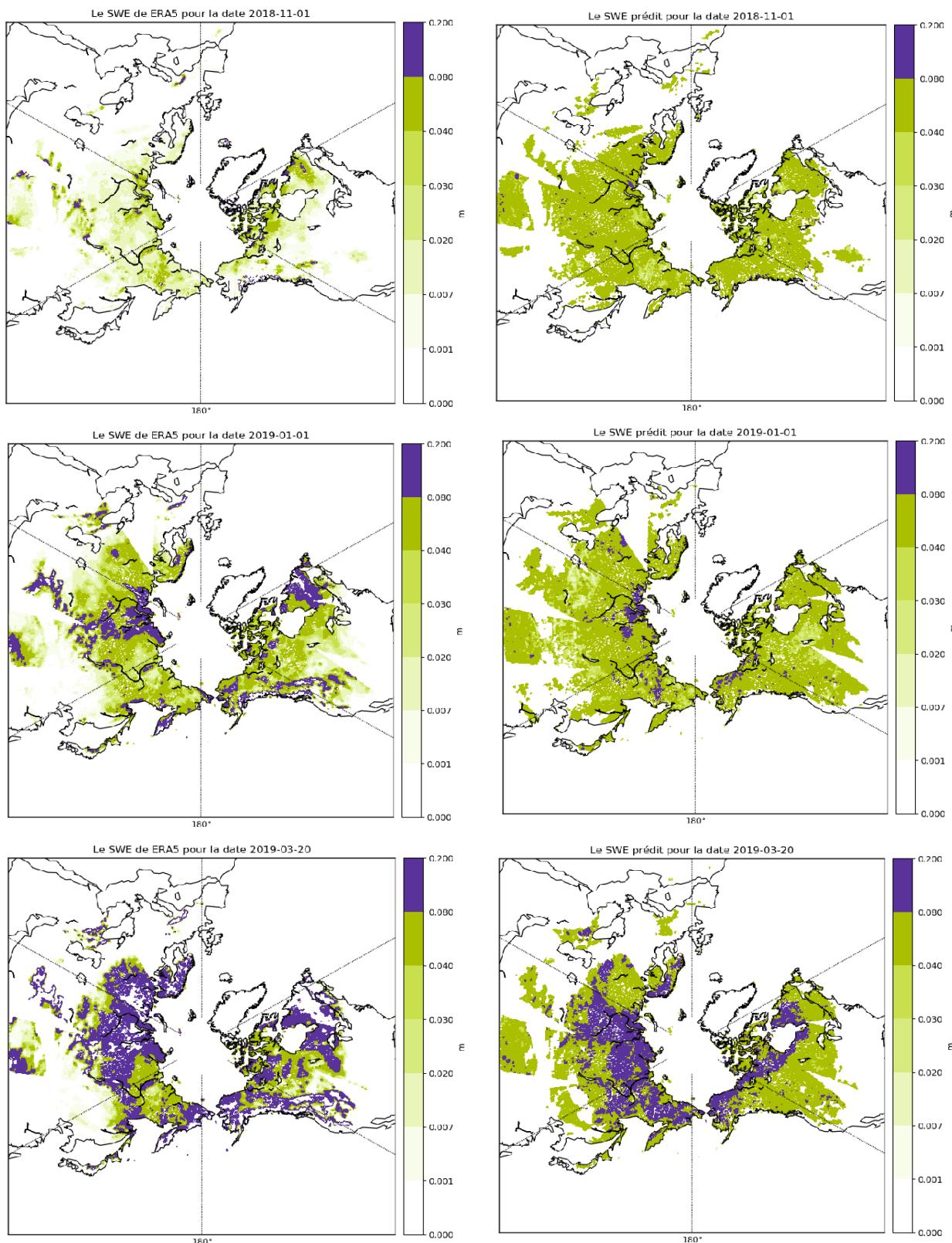


FIGURE 26 – Comparaison entre le SWE de ERA5 et le SWE prédit à l'aide de la RLM en utilisant les données journalières à trois jours de la période d'étude

Les résultats obtenus avec ces trois tests sont conformes aux résultats lors de l'évaluation du modèle. En effet, le 1er novembre et à une échelle hémisphérique, les valeurs de SWE sur terre sont trop faibles et le modèle sur-estime

le SWE. Par contre, vers la fin de la période hivernale, les manteaux neigeux prennent de l'épaisseur suite aux précipitations neigeuses hivernales et les valeurs de SWE sont ainsi importantes. Dans ce cas, le modèle a tendance à sous-estimer les valeurs de SWE.

6.2.2 Modélisation de SWE à partir des Tbs micro-ondes qui seront issues du radiomètre CIMR

Nous visons à développer dans ce cas un modèle de restitution de SWE à partir des Tbs micro-ondes qui proviendront du radiomètre CIMR, mais l'entraînement et l'évaluation du modèle se fait toujours avec les températures de brillances issues du radiomètre AMSR2 sans utilisation de canaux Tb23h/v et Tb89h/v.

Dans ce cas, nous avons à notre disposition 8 températures de brillance comme variables d'entrée : Tb6v, Tb6h, Tb10v, Tb10h, Tb18v, Tb18h, Tb36v et Tb36h.

La restitution de SWE à partir des Tbs micro-ondes passives précitées en utilisant la régression linéaire multiple a donné aussi de mauvais résultats (voir la figure 27).

On constate une forte sous-estimation des valeurs de SWE qui sont supérieures à 0.05 m et un coefficient de détermination trop faible.

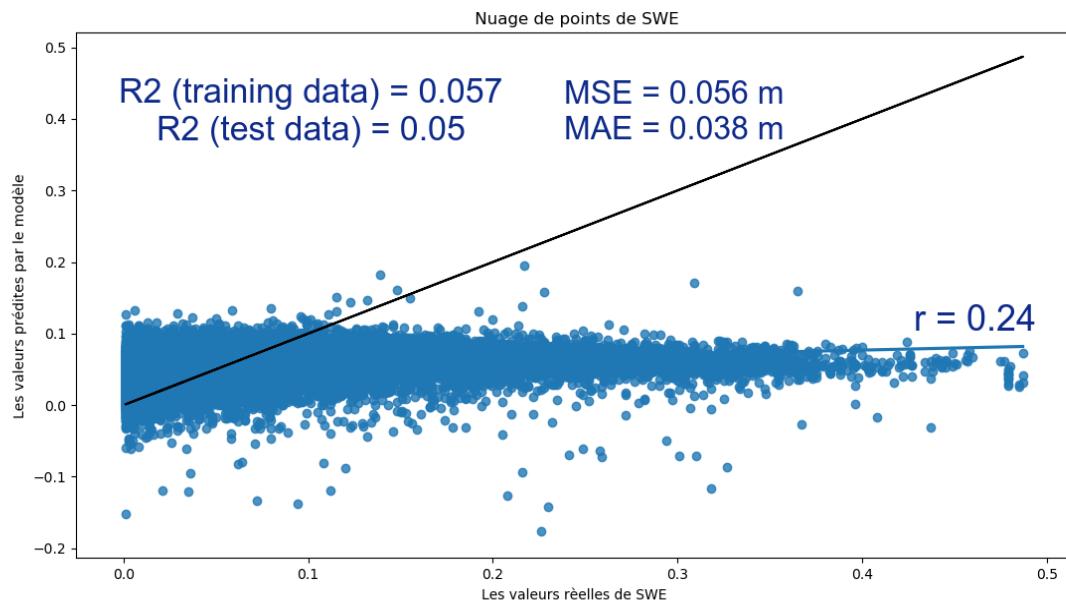


FIGURE 27 – Courbe de dispersion du modèle RLM développé à partir des données AMSR pour la restitution de SWE à partir des Tbs micro-ondes CIMR

En terme de comparaison de ce modèle avec le précédent, on ne peut pas dire qu'un modèle est meilleur qu'un autre en se basant seulement sur le coefficient de détermination. En effet, en augmentant le nombre de variables explicatives, le R^2 augmente aussi d'une manière mécanique, même dans le cas où ces variables ne sont pas pertinentes. Pour comparer ainsi deux modèles dont le nombre de variables est différent, on adopte le critère du R^2 corrigé. Ce dernier, noté \bar{R}^2 , prend en considération le nombre de degrés de liberté. Autrement dit, il tient compte du nombre de variables introduites dans le modèle. Son expression mathématique est la suivante :

$$\bar{R}^2 = 1 - \frac{n-1}{n-q-1} \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (26)$$

Où n est la taille de la base de données de test, q est le nombre de variables explicatives introduites dans le modèle, y_i est la valeur observée du prédictand, \hat{y}_i est la valeur prédictée par le modèle et \bar{y} est la moyenne des y_i .

Il est possible aussi d'exprimer le \bar{R}^2 à partir de R^2 comme suit :

$$\bar{R}^2 = 1 - \frac{n-1}{n-q-1} (1 - R^2) \quad (27)$$

On constate qu'on ajoutant une variable supplémentaire dans le modèle, deux effets antagonistes s'opposent : \bar{R}^2 augmente avec l'amélioration de R^2 et diminue quand le nombre des variables explicatives augmente.

Avec $n = 110060$, $R^2 = 0.13$ et $q = 12$, le calcul de \bar{R}^2 pour le modèle précédent a donné une valeur de 0.129.

Avec $n = 110060$, $R^2 = 0.13$ et $q = 8$, le calcul de \bar{R}^2 pour le présent modèle a donné une valeur de 0.049.

Ainsi, le modèle développé avec les douze températures de brillance est plus performant que celui développé avec seulement les huit températures de brillance.

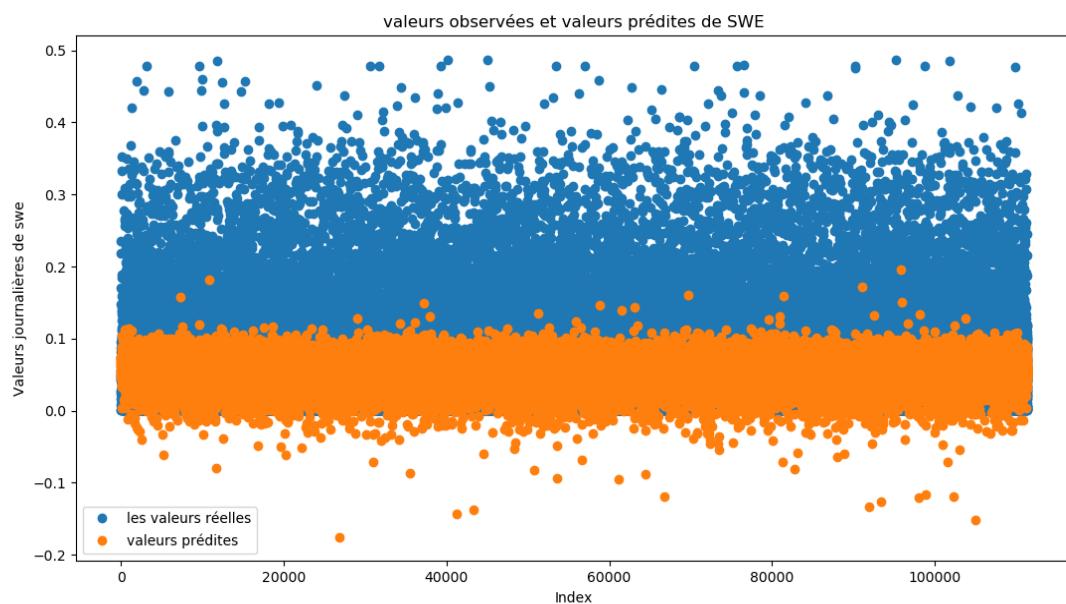


FIGURE 28 – Nuage de points prédits-observés du modèle RLM développé avec les données Tbs de AMSR2 pour la restitution de SWE à partir des Tbs micro-ondes CIMR

En ce qui concerne l'analyse de sensibilité de ce modèle, on peut faire les mêmes remarques que celles faites sur le modèle précédent. On remarque que presque la totalité des valeurs de SWE qui sont supérieures à 0.1 m sont mal prédites.

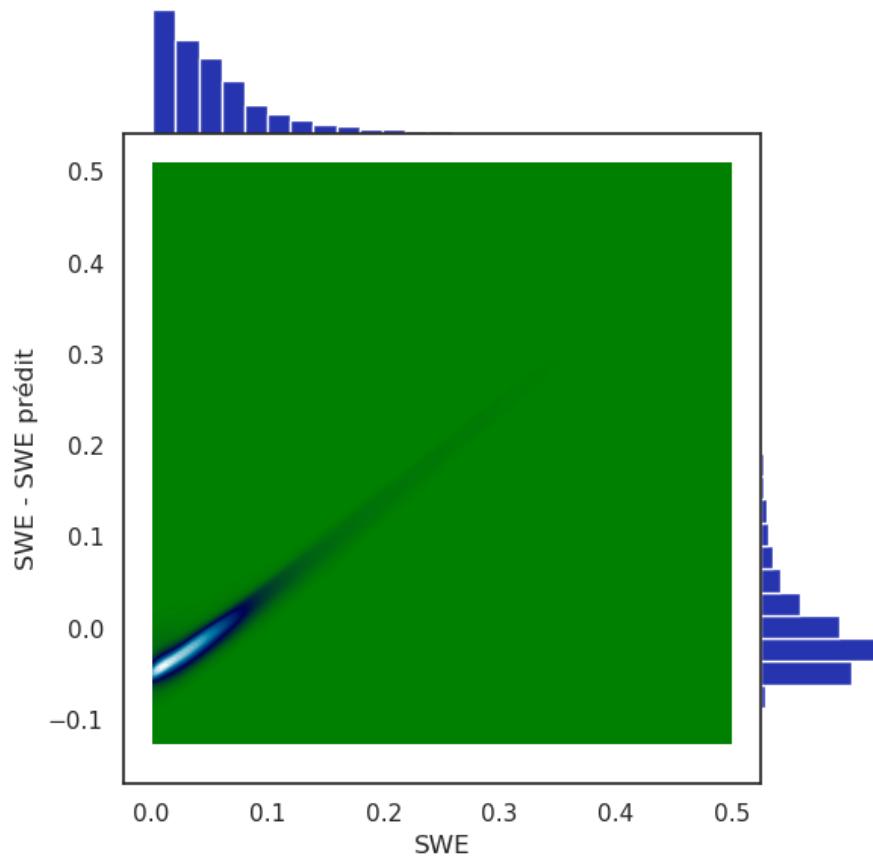


FIGURE 29 – Différences entre les valeurs prédites et les valeurs issues des réanalyses ERA5 en fonction des valeurs SWE issues des réanalyses ERA5

Les différences entre les valeurs prédites et les valeurs issues des réanalyses ERA5 en fonction des valeurs SWE issues des réanalyses ERA5 montrent une sous-estimation dominante des valeurs supérieures à 0.08 m et une sur-estimation des faibles valeurs (les valeurs proches de zéro).

De ce fait, avec la RLM, nous n'avons pas pu modéliser le SWE à partir des températures de brillances micro-ondes passives à l'échelle hémisphérique. Cela dévoile la complexité de la restitution des propriétés de la neige sur terre à l'échelle planétaire à partir des observations micro-ondes passives, ce qui justifie encore notre recours à d'autres méthodes de modélisation plus complexes qui peuvent modéliser les moindre détails de l'information contenue dans la donnée, à savoir la modélisation à partir des réseaux de neurones que nous allons voir dans les prochains paragraphes.

6.3 Modélisation de SWE avec les réseaux de neurones

Les architectures des réseaux de neurones sont assez nombreuses. Dans notre cas, nous adoptons l'architecture la plus simple et la plus utilisée dans les problèmes de régression, il s'agit d'une architecture dans laquelle le réseau est complètement connecté. Autrement dit, chaque neurone d'une couche donnée est connecté à tous les neurones de la couche ou des deux couches adjacentes.

Comme première étape de l'application du réseau de neurones, il fallait normaliser (centrer et réduire) les variables d'entrée pour éviter d'exploser le gradient.

Le modèle comporte plusieurs paramètres qu'on peut changer afin de chercher le meilleur modèle possible. Premièrement, on peut faire varier le nombre de couches cachées, le nombre de neurones dans chaque couche et la fonction d'activation associée à chaque couche. On peut faire varier également l'algorithme d'optimisation, le nombre d'itérations et la taille du lot (batch-size). Ces deux derniers paramètres sont très importants, car ils déterminent le nombre de fois où les poids du réseau sont mis à jour. Si par exemple un algorithme DNN (deep neural network) est appliqué à une base de données d'apprentissage de taille 1000 avec un nombre d'itérations de 100 et un batch-size de 50. Cela veut dire que l'algorithme va diviser l'ensemble d'apprentissage en plusieurs morceaux de taille 50 observations (étant donné la taille de la base de données qui est de 1000, on aura 20 morceaux) et à chaque itération les poids du réseau vont être mis à jour sur chaque lot d'apprentissage. Autrement dit, les poids du réseau vont être mis à jour $20 \times 100 = 2000$ fois pendant la période d'apprentissage. Selon le volume du lot (batch-size), on distingue trois types d'algorithmes d'apprentissage :

- **Batch Gradient Descent** : volume du lot = volume de l'ensemble d'apprentissage
- **Stochastic Gradient Descent** : volume du lot = une seule observation (=1)
- **Mini-Batch Gradient Descent** : $1 < \text{volume du lot} < \text{volume de l'ensemble d'apprentissage}$

En général, il n'existe pas de méthode magique pour le choix des hyperparamètres d'un réseau de neurones. On teste plusieurs combinaisons de hyperparamètres de l'algorithme. Une méthode de recherche par grille peut être menée en initialisant chaque paramètre du modèle à une ou plusieurs valeurs et faire tourner le modèle avec toutes les combinaisons possibles de ces paramètres, puis sélectionner enfin la combinaison qui a donné le meilleur score. Pourtant, si la

taille de la base données est importante, la recherche par grille ralentit l'exécution du modèle et devient irréalisable avec les moyens de calcul basiques.

Dans notre cas, nous avons testé une trentaine de modèles appliqués aux observations dont le SWE est compris entre 0.001 m et 0.5 m d'équivalent d'eau. Puis, nous avons sélectionné le modèle avec le meilleur score après une validation croisée pour vérifier sa pertinence.

6.3.1 Modèles obtenus en utilisant les températures de brillance de tous les canaux (Tb6h/v, Tb10h/v, Tb18h/v, Tb23h/v, Tb36h/v et Tb89h/V.)

6.3.1.1 Premier modèle DNN (M1)

Le tableau suivant illustre les valeurs prises par les hyperparamètres de ce modèle :

Algorithme d'optimisation :	Adam	
Nombre d'itérations :	500	
Batch size :	64	
Couches cachées	Nombre de neurones	Fonction d'activation
Couche C1	12	ReLU
Couche C2	24	ReLU
Couche C3	12	ReLU

TABLE 9 – Les valeurs prises par les paramètres du premier modèle

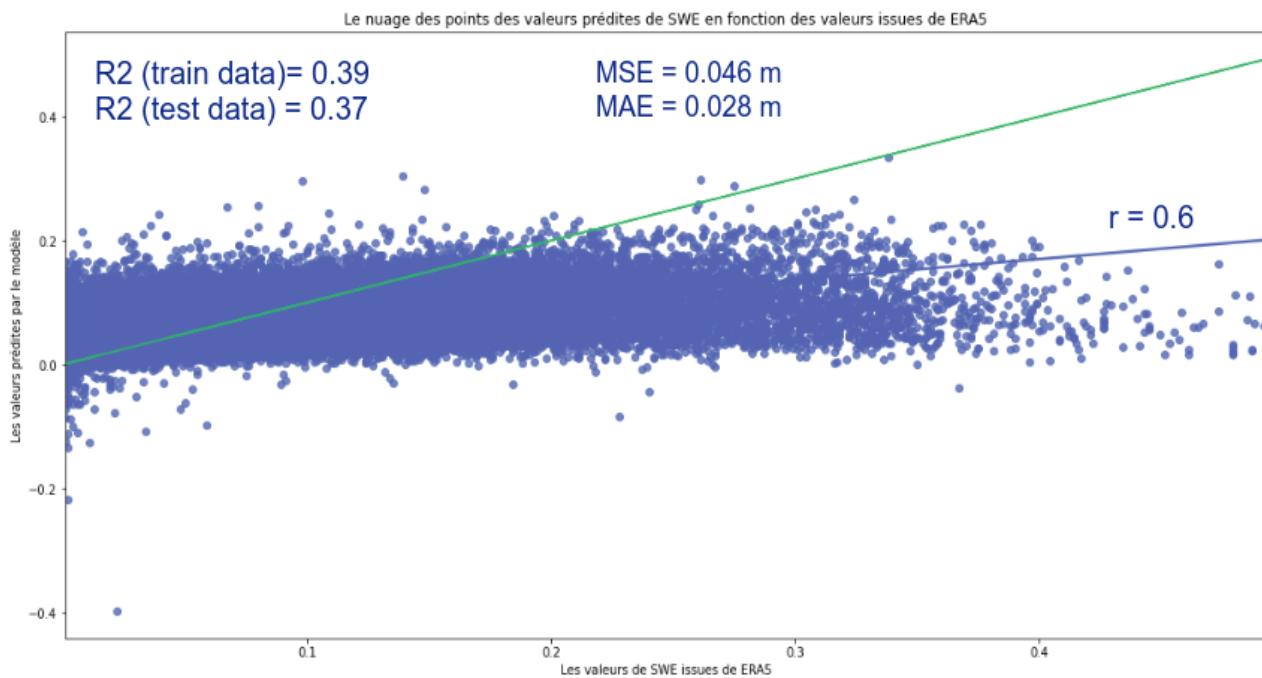


FIGURE 30 – Courbe de dispersion du premier modèle de DNN (M1) en prenant en considération toutes les Tbs (les douze Tbs)

En comparaison avec les résultats de la RLM, on remarque une amélioration considérable des scores. Autrement dit, ce premier modèle de DNN modélise le SWE à partir des Tbs beaucoup mieux que la RLM. Toutefois, on constate encore une sous-estimation dominante des valeurs élevées de SWE et une sur-estimation des valeurs qui sont très faibles.

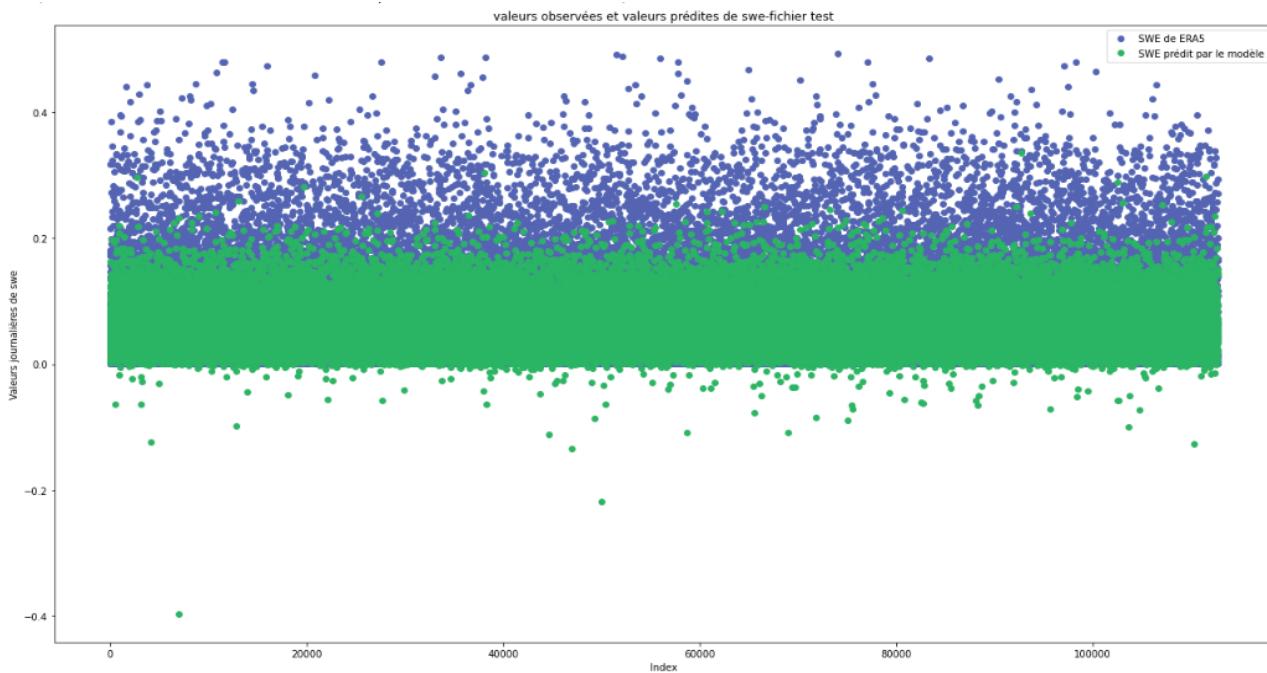


FIGURE 31 – Nuage de points prédits-observés avec le modèle M1 du DNN en utilisant toutes les Tbs

Le résultat est toujours meilleur que la RLM, mais une forte sous-estimation des valeurs élevées de SWE est toujours présente (plus précisément les valeurs de SWE qui sont supérieures à 0.2 m).

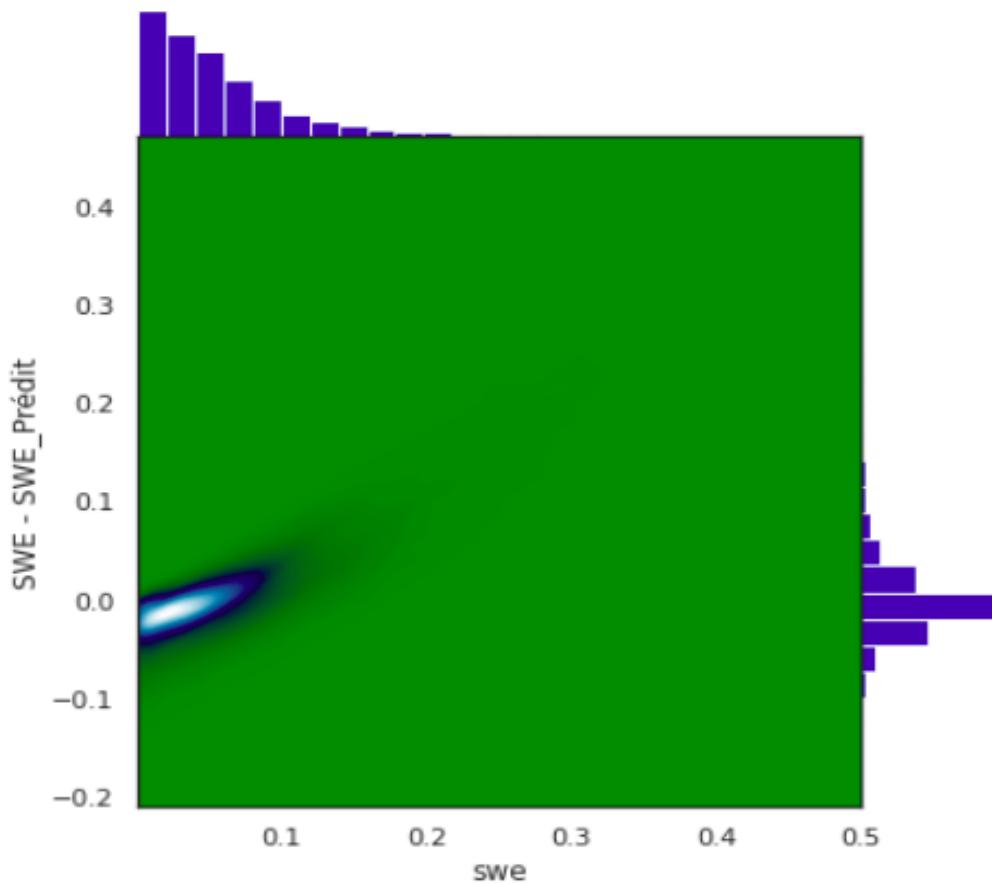


FIGURE 32 – La densité des noyaux des différences entre les valeurs prédites et les valeurs issues de ERA5 de SWE en fonction des valeurs de SWE issues de ERA5

La figure ci-dessus confirme également que les valeurs élevées de SWE sont sous-estimées et qu'une grande partie des valeurs les plus faibles de SWE sont sur-estimées. À partir de cette densité de noyaux, on peut affirmer que les valeurs appartenant aux intervalles où la densité des points est la plus importante sont mieux estimées que d'autres valeurs. En effet, quand le nombre d'observations est plus important dans un intervalle de la plage de variations des valeurs de la variable expliquée, le modèle aura plus de chance d'apprendre un maximum d'informations dans cette plage de valeurs.

Pour bien visualiser la variabilité spatiale des erreurs, nous avons appliqué ce modèle à trois jeux de données journalières pour produire des cartes des valeurs absolues des erreurs d'estimation.

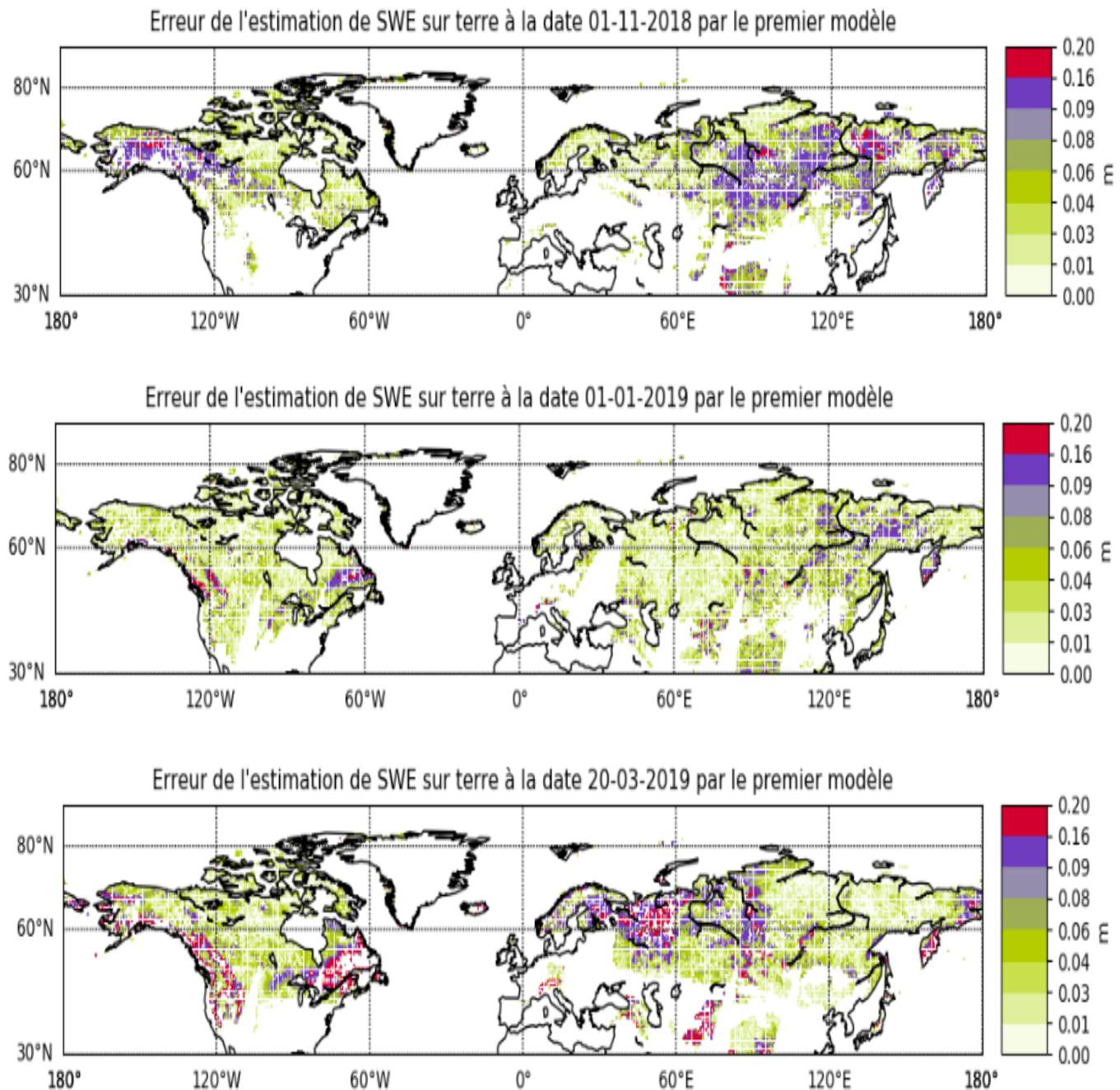


FIGURE 33 – Visualisation de la variabilité spatiale de l'erreur d'estimation de SWE par le premier modèle de DNN à trois dates de la période d'étude

D'après les cartes ci-dessus, on remarque bel et bien que l'erreur de l'estimation de SWE par ce modèle est moins importante au cœur de la période hivernale qu'à son début et sa fin. Pour expliquer cette différence, on peut supposer que les valeurs de SWE sur terre au premier novembre sont trop faibles et le modèle a tendance à les sur-estimer. Par contre, vers la fin de la période hivernale, les valeurs de SWE sur terre sont très importantes suite aux précipitations neigeuses et le modèle a tendance à les sous-estimer. Néanmoins, cette différence n'est pas expliquée seulement par l'hypothèse précitée relative au modèle, mais

plusieurs facteurs peuvent affecter la modélisation de SWE sur terre à partir des Tbs micro-ondes, comme la topographie, la végétation, la température de l'air, la couverture nuageuse, la teneur en eau liquide dans la neige sur terre, etc. C'est pour cela qu'on constate une variation longitudinale et zonale de l'erreur pour la même date.

6.3.1.1 Deuxième modèle DNN (M2)

En vue d'améliorer le résultat obtenu avec le premier modèle, nous allons filtrer les observations dont la valeur de SWE est supérieure à 0.2 m (voir la figure ci-contre) car le SWE prenant ces valeurs est mal estimé, ces observations risquent ainsi d'être influentes. Ensuite, nous allons diminuer le Batch-size de 64 à 32.

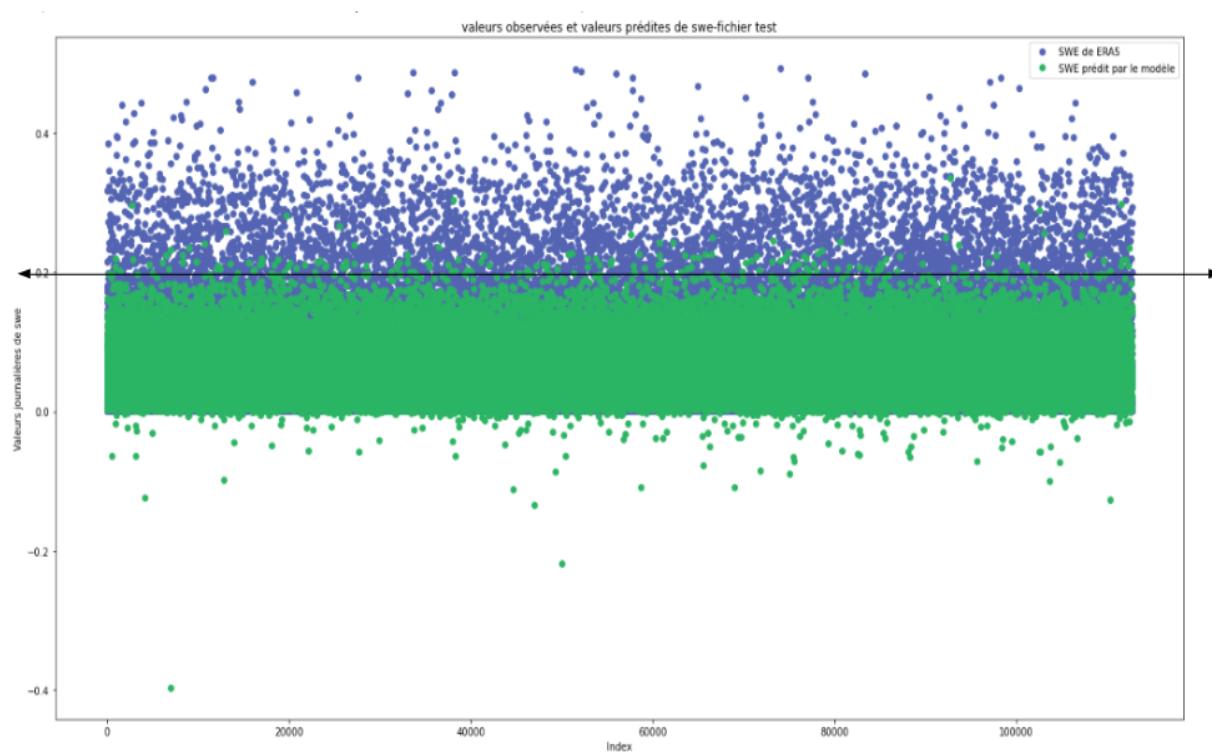


FIGURE 34 – Nuage de points de SWE de ERA5 et le SWE prédit par le modèle M1 avec la ligne limitant les valeurs filtrées de la base de données pour élaborer ce deuxième modèle M2

Une amélioration dans les scores du deuxième modèle par rapport au premier a été observée (voir la figure ci-contre)

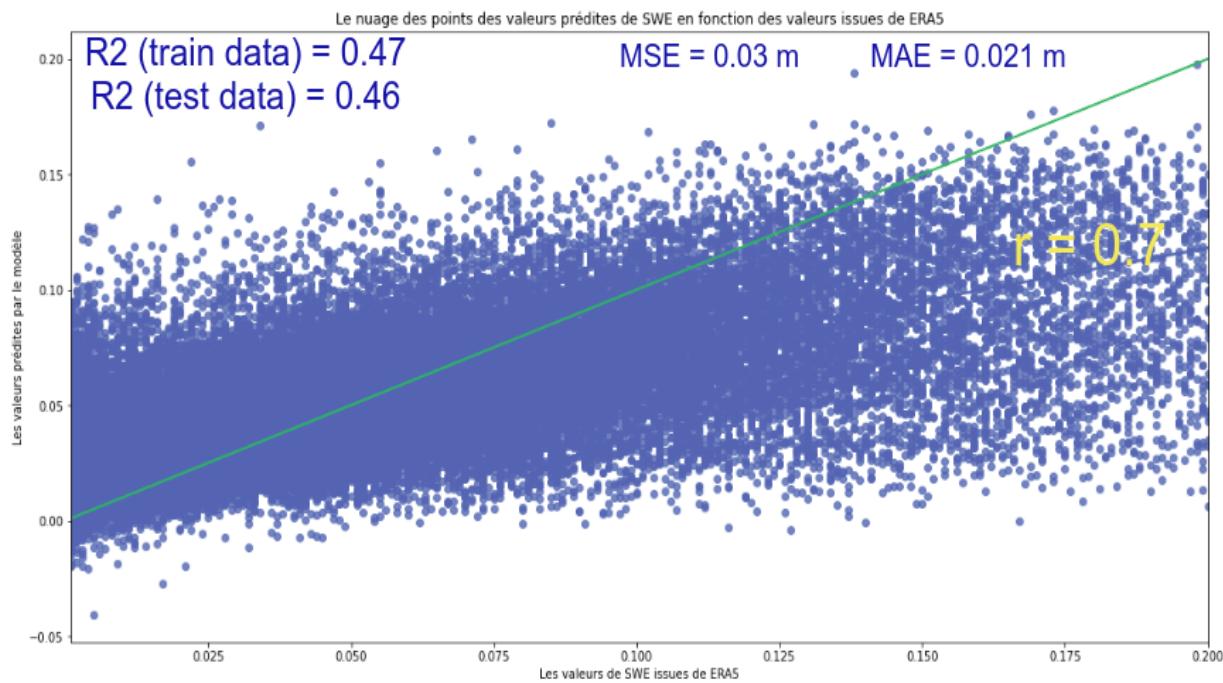


FIGURE 35 – La courbe de dispersion du deuxième modèle de DNN (M2) en prenant en considération toutes les Tbs

En filtrant les points dont le SWE est supérieur à 0.2 et en faisant varier le Batch-size de 64 à 32, le coefficient de détermination est passé de 0.37 à 0.46.

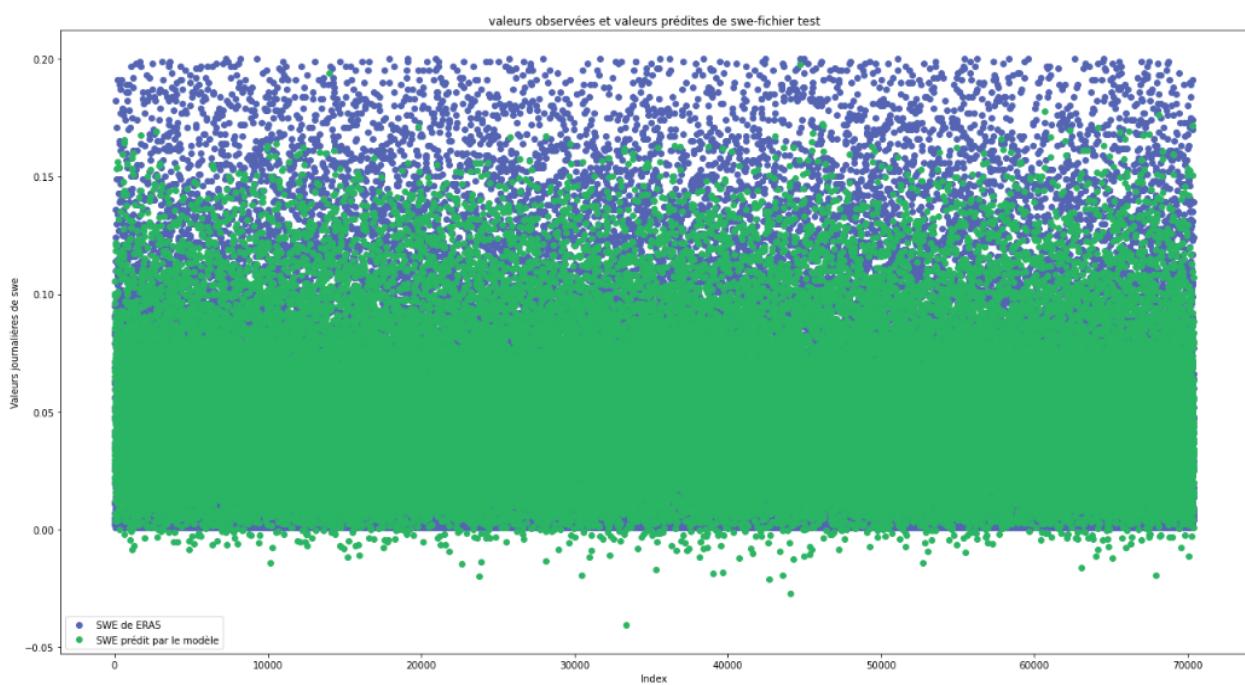


FIGURE 36 – Nuage de points prédicts-observés avec le modèle M2 du DNN en prenant en considération toutes les Tbs

Les deux nuages de points des valeurs de SWE issu de ERA5 et les valeurs de SWE prédict par le modèle commencent à s'ajuster. Néanmoins, les valeurs de

SWE de ERA5 qui sont supérieures à 0.15 m sont toujours trop mal estimées.

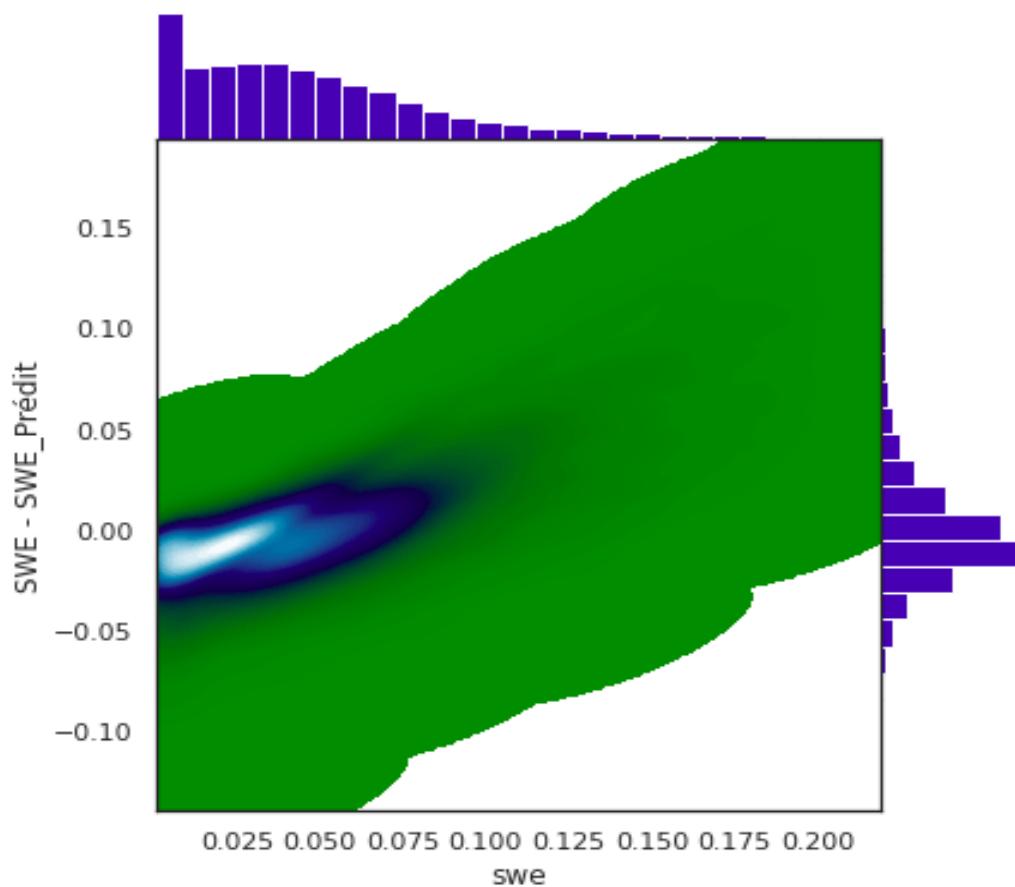


FIGURE 37 – La densité des noyaux des différences entre les valeurs prédites et les valeurs de SWE de ERA5 en fonction de SWE de ERA5

La figure de densité de noyaux montre ensuite qu'une grande partie des valeurs de SWE qui sont inférieures à 0.05 m sont légèrement sous-estimées.

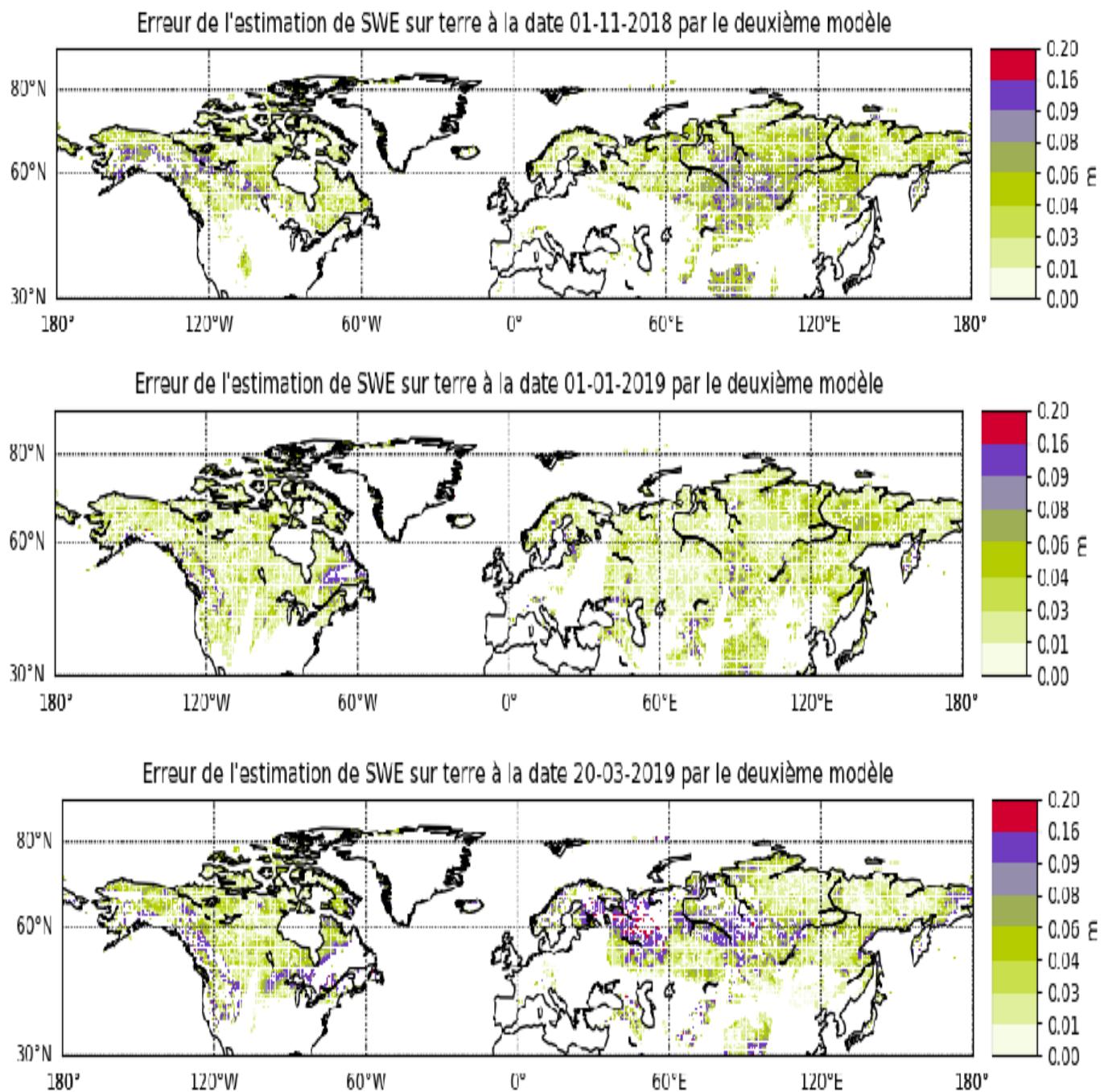


FIGURE 38 – Visualisation de la variabilité spatiale de l'erreur d'estimation de SWE par le deuxième modèle de DNN à trois dates de la période d'étude

On remarque toujours que les erreurs et leur variabilité spatiale sont moins importantes au cœur de la période hivernale qu'à sa fin et la période automnale.

IL faut noter aussi que ce deuxième modèle est confronté à un problème de sur-ajustement (voir la figure 46 dans les annexes).

6.3.1.1 Comparaison des performances des deux modèles de DNN à travers des cartes d'erreurs

Dans cette section, nous visons à comparer les performances des deux modèles à travers des cartes d'erreurs élaborées pour trois jours de la période d'étude, le premier novembre, le premier janvier et le vingt mars. Les données de ces trois dates n'ont pas été incluses ni dans l'ensemble de l'apprentissage ni dans celui de validation.

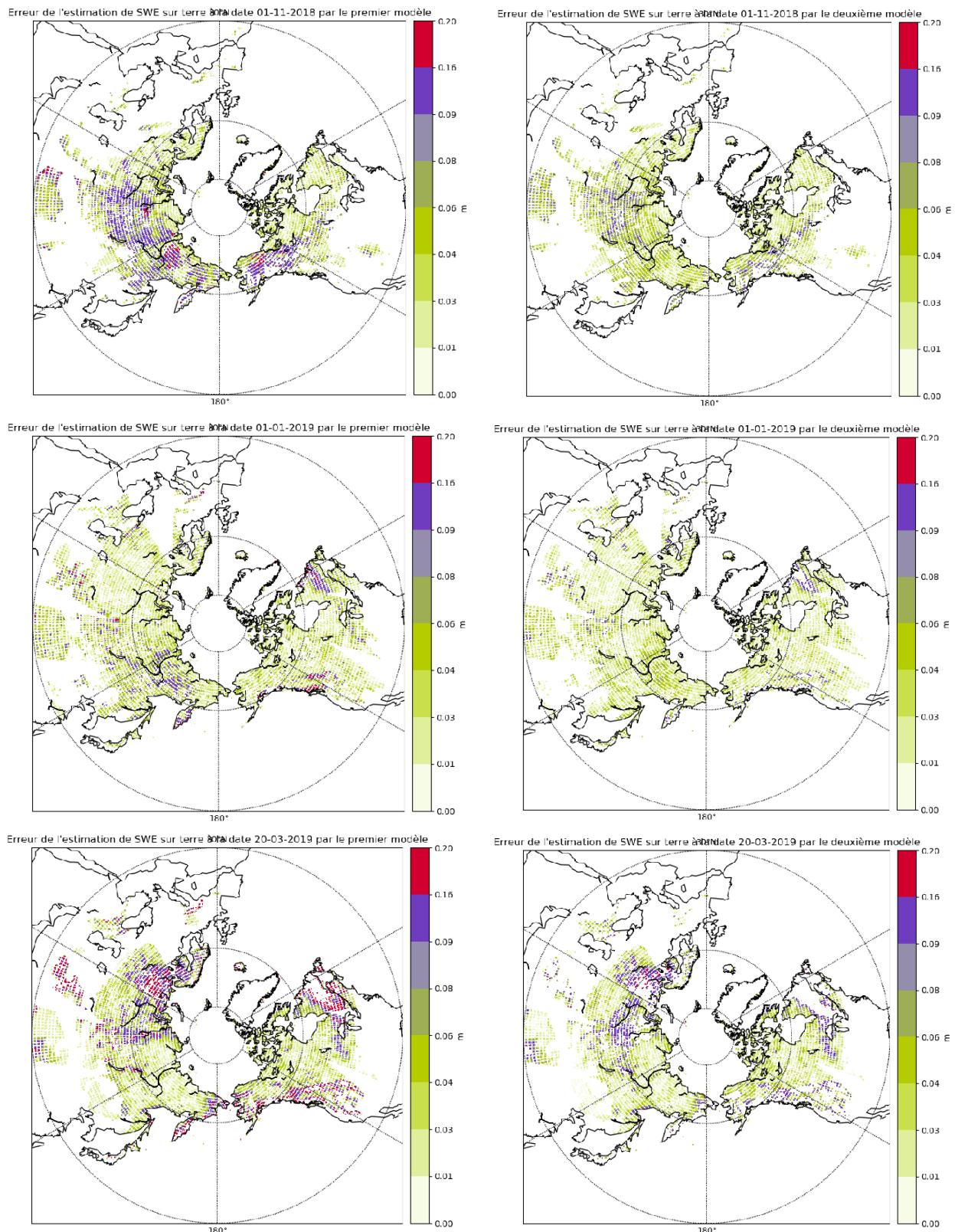


FIGURE 39 – Comparaison des performances des deux modèles DNN à travers des cartes d'erreurs élaborées pour trois dates de la période d'étude, les cartes à gauche représentent les erreurs du premier modèle DNN (M1), les cartes à droite représentent les erreurs du deuxième modèle (M2)

On constate bel et bien que les erreurs du deuxième modèle sont plus faibles que celles du premier modèle. Pourtant, les erreurs les plus importantes sont

localisées approximativement dans les mêmes régions, ce qui justifie l'impact des conditions locales sur la modélisation de SWE à partir des Tbs micro-ondes passives. Les conditions locales sont nombreuses et peuvent être la température de l'air ou de la surface, la végétation, la topographie, la couverture nuageuse, etc.

6.3.2 Modélisation de SWE en utilisant seulement les températures de brillance suivantes : Tb6h/v, Tb10h/v, Tb18h/v et Tb36h/v.

En utilisant seulement les canaux du futur radiomètre CIMR et l'algorithme du deuxième modèle précité, le coefficient de détermination a été diminué de 0.46 à 0.3 sur l'ensemble de validation. Cela signifie que les canaux 23 et 89 avec les deux polarisations ont une contribution non-négligeable dans l'amélioration de la variance expliquée de SWE par les températures de brillance micro-ondes passifs.

7 Conclusion

En peu de mots, la RLM est très limitée quand-t-il s'agit de modéliser des propriétés de la neige sur terre à partir des Tbs micro-ondes passifs à l'échelle globale. En testant un DNN plusieurs fois en faisant varier ses paramètres en vue d'améliorer les scores, le meilleur modèle obtenu est un modèle permettant d'expliquer 46 % de la variance de SWE à partir de toutes les températures micro-ondes de AMSR2 à l'exception de celle du canal à 7 GHz. En utilisant seulement les Tbs qui seront issues des canaux du futur radiomètre CIMR, le coefficient de détermination est passé de 0.46 à 0.3.

CONCLUSION ET PERSPECTIVES

Ce travail, réalisé dans le cadre de projet de fin d'études, a porté sur quatre grands axes, le premier a eu comme objectif l'étude de l'interaction du rayonnement avec la neige en vue de dévoiler physiquement la sensibilité des MOP aux propriétés de la neige sur terre pour déterminer à quel point les propriétés de la neige sur terre impactent le rayonnement micro-onde. Le deuxième axe a consisté en une étude exploratoire des données relatives aux propriétés de la neige sur terre et des températures de brillance afin de vérifier l'intégrité des données récoltées d'une part et d'illustrer la variabilité spatio-temporelle des propriétés de la neige et des Tbs micro-ondes de AMSR2 ainsi que les distributions statistiques des données de l'autre part. Le troisième axe a porté sur les principaux prétraitements des données dans le but d'avoir une base de données exploitable et la sélection des variables potentielles. Le dernier axe a eu comme objectif le développement des algorithmes de restitution de SWE en utilisant la RLM et les méthodes de l'apprentissage profond (deep learning).

S'agissant du premier objectif spécifique, nous avons découvert à travers l'analyse d'une trentaine d'articles que les variations des propriétés de la neige sur terre, notamment le SD (l'épaisseur de la neige sur terre) et le RSN (la densité de la neige sur terre) affectent considérablement l'émission à travers les manteaux neigeux dans le domaine spectral des micro-ondes. Puisque le SWE est fonction de SD et de RSN, on peut affirmer bel et bien que les variations de SWE sur terre impactent l'émission à travers les manteaux neigeux dans le spectre des micro-ondes. C'est pour cela que les radiomètres à MOP se sont avérés efficaces pour la télédétection passive de la neige sur terre.

Une première exploration des données a été également menée. Elle nous a permis en premier temps de nous assurer de la fiabilité des données récoltées, et en deuxième temps de comprendre la variabilité spatio-temporelle de la neige sur terre dans l'hémisphère nord et les Tbs micro-ondes passives, et ce, en visualisant la couverture nivale et l'épaisseur de la neige sur terre pour quelques

jours de la période d'étude qui s'étale du premier novembre 2018 jusqu'au 30 mars 2019 et les Tbs micro-ondes de AMSR2. Ces visualisations ont dévoilé une grande variabilité spatio-temporelle de l'épaisseur de la neige sur terre dans l'hémisphère nord et une grande couverture des moyennes et les hautes latitudes pendant la période hivernale. Pour les températures de brillances MOP, nous avons remarqué un contraste terre-mer important et une grande variabilité spatiale à l'intérieur des continents.

En ce qui concerne le prétraitement des données et la sélection des variables potentielles, quelques filtrages et mises en forme de la base de données brute ont été réalisées afin d'obtenir une base de données exploitable par les différents programmes python des algorithmes d'apprentissage. Ensuite, les variables potentielles ont été sélectionnées en se référant aux liens physiques, aux études scientifiques traitant la même problématique et à l'étude de corrélations entre les variables en prenant en considération seulement les canaux micro-ondes que détiennent les radiomètres à notre disposition. Ainsi, les variables potentielles sélectionnées ont été 6v/h, 10v/h, 18v/h, 23v/h, 36v/h, 89v/h.

La dernière partie vient par la suite pour développer des algorithmes de restitution des propriétés de la neige en utilisant la RLM et l'apprentissage profond. Elle consiste précisément à choisir le modèle du Machine Learning le plus performant pour l'estimation de SWE à partir des températures de brillance, et ce, par le biais de la comparaison des résultats de celle-ci en termes des scores statistiques à savoir le RMSE, MSE, MAE, R^2 ou bien \bar{R}^2 quand il s'agit de comparer des modèles dont les nombres des variables d'entrée sont différents. En utilisant la RLM, les résultats obtenus ont été insatisfaisants et le meilleur score était de 13 % avec une sur-estimation des faibles valeurs de SWE et une sous-estimation dominante de la quasi-totalité des valeurs qui sont supérieures à 0.05 m d'équivalent en eau. Concernant les méthodes de l'apprentissage profond, les résultats ont été considérablement améliorés par rapport à la RLM, et le meilleur modèle réalisé a eu un coefficient de détermination de 0.46 sans problème de surajustement (en effet, le R2 sur l'ensemble d'apprentissage été de 0.47, celui-ci est très proche de R2 sur l'ensemble de validation) en utilisant toutes les températures micro-ondes d'AMSR2 à l'exception de celle du canal à 7 GHz.

En visualisant les cartes des erreurs de l'estimation de SWE par les méthodes de l'apprentissage profond, nous avons remarqué que les erreurs sont plus importantes à la fin de la période hivernale et en période automnale qu'au cœur de l'hiver. En plus, une variabilité spatiale des erreurs a été observée, ce qui a prouvé la dépendance du signal micro-onde prévenant de la surface avec les conditions locales. Ceci nous pousse à mettre des hypothèses que nous allons

prendre en considération pour améliorer la modélisation de SWE sur terre à partir des Tbs micro-ondes à l'échelle globale. Ces hypothèses sont formées en se basant sur la distribution des erreurs des modèles que nous avons construits et l'étude bibliographique. Les plans d'eau, la couverture forestière, les terrains complexes et les zones urbaines affectent largement les estimations globales de SD et SWE [32]. En termes d'étapes que nous pouvons aborder pour limiter l'effet de ces facteurs sur l'estimation de SWE ou de SD à l'échelle globale, nous pouvons masquer les terrains complexes, les zones urbaines et les plans d'eau, puis introduire l'indice de végétation comme une variable d'entrée parmi les autres dans les modèles. Les estimations de SWE sur terre sont particulièrement difficiles dans les zones de forte densité d'accumulation de neige [20]. L'étude [25] a montré également l'importance d'isoler les autres phénomènes qui peuvent influencer le signal micro-onde, comme le givre en profondeur, l'humidité de la neige, les terrains accidentés et la couverture forestière dense. Ainsi, la solution qui persiste est de masquer les zones présentant l'une ou plusieurs de ces conditions locales ou bien introduire certaines variables en entrée des modèles, c'est la problématique à relever dans le dernier mois du stage (Juillet 2021).

En vue d'améliorer encore les résultats, nous pourrons adopter une autre méthode de modélisation, c'est la méthode XGBoost qui est aussi une méthode souvent plus performante qu'un réseau neuronal.

Toutefois, la modélisation des propriétés de la neige sur terre à l'échelle globale ou hémisphérique à partir des observations MOP reste une tâche difficile, passionnante et sur la bonne voie.

BIBLIOGRAPHIE

- [1] Observatoire, «<https://www.observatoiredeparis.psl.eu/?lang=fr>». [En ligne].
- [2] R. Taton. 1976, Les origines et les débuts de l'observatoire de Paris.
- [3] H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegria, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.), *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*. IPCC, 2019.
- [4] Fanny Larue, *Développement d'un système d'assimilation de mesures satellites micro-ondes passives dans un modèle de neige pour la prévision hydrologique au Québec*. Décembre 2017.
- [5] nsidc, «<https://nsidc.org/cryosphere/snow> ». [en ligne].
- [6] Edward H. Bair, Andre Abreu Calfa, Karl Rittger, and Jeff Dozier, *Using machine learning for real-time estimates of snow water equivalent in the watersheds of Afghanistan*. 3 May 2018.
- [7] Takashi Maeda, Member, IEEE, Yuji Taniguchi, and Keiji Imaoka, *GCOM-W1 AMSR2 Level 1R Product : Dataset of Brightness Temperature Modified Using the Antenna Pattern Matching Technique* . FEBRUARY 2016.
- [8] Roberta Pirazzini, Leena Leppänen, Ghislain Picard, Juan Ignacio Lopez-Moreno, Christoph Marty, Giovanni Macelloni, Anna Kontu, Annakaisa von Lerber, Cemal Melih Tanis, Martin Schneebeli, Patricia de Rosnay, Ali Nadir Arslan, *European In-Situ Snow Measurements : Practices and Purposes* . 22 June 2018.
- [9] E. Cordisco, C. Prigent, F. Aires, *Snow characterization at a global scale with passive microwave satellite observations* . 4 October 2006.

- [10] A.T.C. Chang, J.L. Foster and D.K. Hall, *NIMBUS- 7 SMMR DERIVED GLOBAL SNOW COVER PARAMETERS*. 1987.
- [11] Richard E. Kelly, Alfred T. Chang, Fellow, IEEE, Leung Tsang, Fellow, IEEE, and James L. Foster, *A Prototype AMSR-E Global Snow Area and Snow Depth Algorithm* . FEBRUARY 2003.
- [12] L. R. M UDRYK, C. D ERKSEN, P. J. K USHNER, R. B ROWN, *Characterization of Northern Hemisphere Snow Water Equivalent Datasets, 1981–2010* . July 2015.
- [13] A. T. C. Chang, L. Tsang, *A Neural Network Approach to Inversion of Snow Water Equivalent from Passive Microwave Measurement* .3 March, 1992.
- [14] Jamie L. Dyer and Thomas L. Mote, *Spatial variability and trends in observed snow depth over North America*.31 August 2006.
- [15] Melinda A. Webster, Ignatius G. Rigor, Son V. Nghiem, Nathan T. Kurtz, Sinead L. Farrell, Donald K. Perovich, and Matthew Sturm , *Interdecadal changes in snow depth on Arctic sea ice*.22 AUG 2014.
- [16] Tao CHE, Xin LI, Rui JIN, Richard ARMSTRONG, Tingjun ZHANG, *Snow depth derived from passive microwave remote-sensing data in China*. 2008.
- [17] MATTHEW STURM, BRIAN TARAS, GLEN E. LISTON , CHRIS DERKSEN, TOBIAS JONAS, JON LEA, *Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes* . 10 August 2010.
- [18] BRUCE BRASNETT, *A Global Analysis of Snow Depth for Numerical Weather Prediction* . 3 August 1998.
- [19] A. S. BAMZAI, J. SHUKLA, *Relation between Eurasian Snow Cover, Snow Depth, and the Indian Summer Monsoon : An Observational Study* . 23 October 1998.
- [20] Andrew M.Snauffer, William W.Hsieh, Alex J.Cannon, Markus A.Schnorbus, *Improving gridded snow water equivalent products in British Columbia, Canada : multi-source data fusion by neural network models* . 12 Mar 2018.
- [21] Konstantin Franz Fotios Ntokas, Jean Odry, Marie-Amélie Boucher, and Camille Garnaud, *Using an ensemble of artificial neural networks to convert snow depth to snow water equivalent over Canada* . 18 Nov 2020.
- [22] Thian Yew Gan, Oscar Kalina, Purushottam Singh, *Comparison of snow water equivalent retrieved from SSM/I passive microwave data using artificial neural network, projection pursuit and nonlinear regressions* . 15 May 2009.

- [23] Hossein Tabari, S.Marofi, H.Zare Abyaneh, M.R.Sharifi, *Comparison of artificial neural network and combined models in estimating spatial distribution of snow depth and snow water equivalent in Samsami basin of Iran* .05 November 2009.
- [24] Safar Marofi, Hossein Tabari, Hamid Zare Abyaneh, *Predicting Spatial Distribution of Snow Water Equivalent Using Multivariate Non-linear Regression and Computational Intelligence Methods* . 12 January 2011.
- [25] A.B.Tait, *Estimation of Snow Water Equivalent Using Passive Microwave Radiation Data* . June 1998.
- [26] Yinsheng Zhang, Ning Ma, *Spatiotemporal variability of snow cover and snow water equivalent in the last three decades over Eurasia* . April 2018.
- [27] Robinson, D. A., F. T. Keiming, and R. R. Heim, *Global Snow Cover Monitoring : An Update*. 1993.
- [28] Nastaran Saberi, Richard Kelly, Margot Flemming and Qinghuan Li, *Review of snow water equivalent retrieval methods using spaceborne passive microwave radiometry*. 07 Sep 2019.
- [29] Colbeck, S, *An Overview of Seasonal Snow Metamorphism* . 1982.
- [30] Marshall, S. J, *The Cryosphere* . 2011.
- [31] Schott, J. R, *Remote Sensing : the Image Chain Approach* . 2007.
- [32] Kelly, R, *The Amsr-e Snow Depth Algorithm : Description and Initial Results* . 2009.
- [33] R. L. Armstrong and M. J. Brodzik, *Recent Northern Hemisphere snow extent : A comparison of data derived from visible and microwave satellite sensors* . 01 October 2001.
- [34] D. K. Hall, R. E. J. Kelly, G. A. Riggs, A. T. C. Chang, and J. L. Foster, *Assessment of the relative accuracy of hemispheric-scale snow-cover maps*. 2002.
- [35] D. K. Hall, J. L. Foster, V. V. Salomonson, A. G. Klein, and J. Y. L. Chien, *Development of a technique to assess snow-cover mapping errors from space*. Feb 2001.
- [36] C. Chen, B. Nijssen, J. Guo, L. Tsang, A. W. Wood, J. Hwang, and D.P. Lettenmaier, *Passive microwave remote sensing of snow constrained by hydrological simulations*. Aug 2001.
- [37] C. Mätzler, *Passive microwave signatures of landscapes in winter*. 1994.

- [38] A. Rango, A. T. C. Chang, and J. L. Foster, *The utilization of spaceborne microwave radiometers for monitoring snowpack properties.* 1979.
- [39] S. C. Colbeck, *An overview of seasonal snow metamorphism.* 1982.
- [40] Chang and Shiue, *Microwaves remote sensing of snowpack properties.* 1980.
- [41] T. C. Chang and P. Gloersen, Goddard Space Flight Center, Greenbelt, Maryland, *Microwaves emission from dry and wet snow.* 1975.
- [42] JianCheng SHI, Chuan XIONG and LingMei JIANG *Review of snow water equivalent microwave remote sensing.* January 6, 2016.
- [43] Clissold, *Candidate earth explorer core missions : Reports for assessment : CoReH2O-Cold regions hydrology high-resolution observatory.* 2008.
- [44] Hobbs, Peter V, *Ice Physics.* 2010.
- [45] Florent Dupont, *Télédétection micro-onde de surfaces enneigées en milieu arctique : étude des processus de surface de la calotte glaciaire Barnes, Nunavut, Canada.* 2014.
- [46] Sylviane Surdyk, *Using microwave brightness temperature to detect short-term surface air temperature changes in Antarctica : An analytical approach.* 2002.
- [47] Groupe HLI, *Le machine learning ou la mutation numérique de l'industrie - HLI Lab*
- [48] AJAY SHRESTHA, AUSIF MAHMOOD, *Review of Deep Learning Algorithms and Architectures.* April 22, 2019.
- [49] J. Zhang and W. Ding, *Prediction of air pollutants concentration based on an extreme learning machine : The case of Hong Kong* 2017.
- [50] Gilbert Saporta, *Probabilités Analyse des Donnes et Statistique.*
- [51] Jeff Howbert, *Classification / Regression Neural Networks.*
- [52] Schmitt, *Les réseaux de neurones artificiels.* apr 2010.
- [53] E. Scornet, *Promenade en forêts aléatoires.*
- [54] G. Louppe, *Understanding Random Forests : From Theory to Practice.* July, 2014.
- [55] A. Schmitt, *Les réseaux de neurones artificiels.* apr 2010.
- [56] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, , *Activation Functions : Comparison of trends in Practice and Research for Deep Learning.* 2018.

- [57] L. Rouvière, *Introduction aux méthodes d'agrégation : boosting , bagging et forêts aléatoires. Illustrations avec R*
- [58] A. Salvail-bérard,*Réseaux de neurones.*
- [59] N. Boland, J. Christiansen, B. Dandurand, A. Eberhard, J. Linderoth, and J. Luedtke,*Combining Progressive Hedging with a Frank-Wolfe Method to Compute Lagrangian Dual Bounds in Stochastic Mixed-Integer Programming*, vol.30.2017.
- [60] J. Duchi, E. Hazan, and Y. Singer, *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization*, vol. 12. 2011.
- [61] S. Ruder,*An overview of gradient descent optimization algorithms.* 2016.
- [62] D. P. Kingma and J. Ba, *Adam : A Method for Stochastic Optimization*.2014.
- [63] E. Lionel, G. Slimane, and C. Jérôme, *Mise en oeuvre de réseau de neurones pour la modélisation de la cinétiques reactinnelles en vus de la transposition batch/continu.* 2007.

ANNEXES

Les températures de brillance pour le premier novembre avec une projection polaire

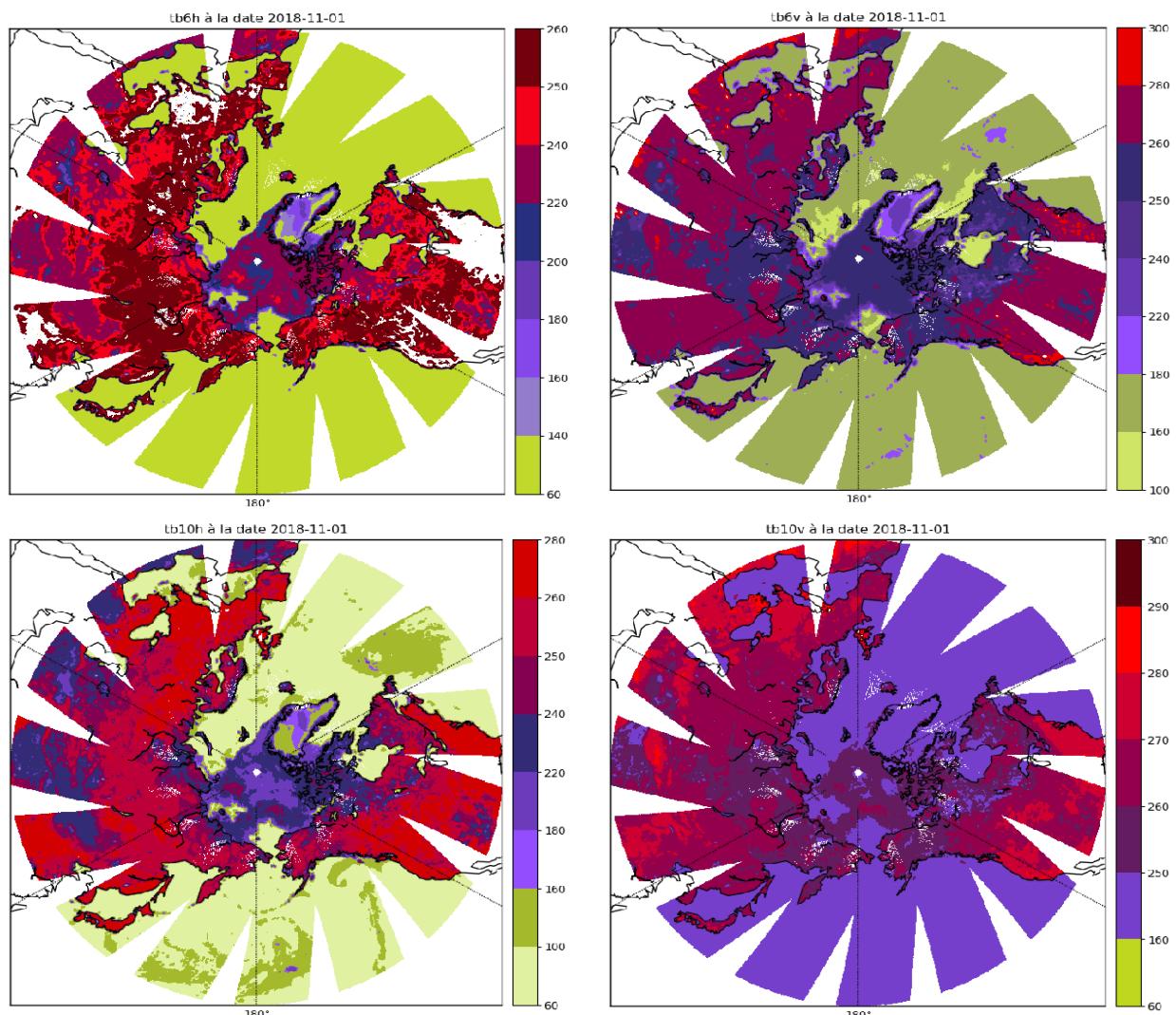


FIGURE 40 – Les températures de brillance tb6h/v et tb10h/v au premier novembre 2018

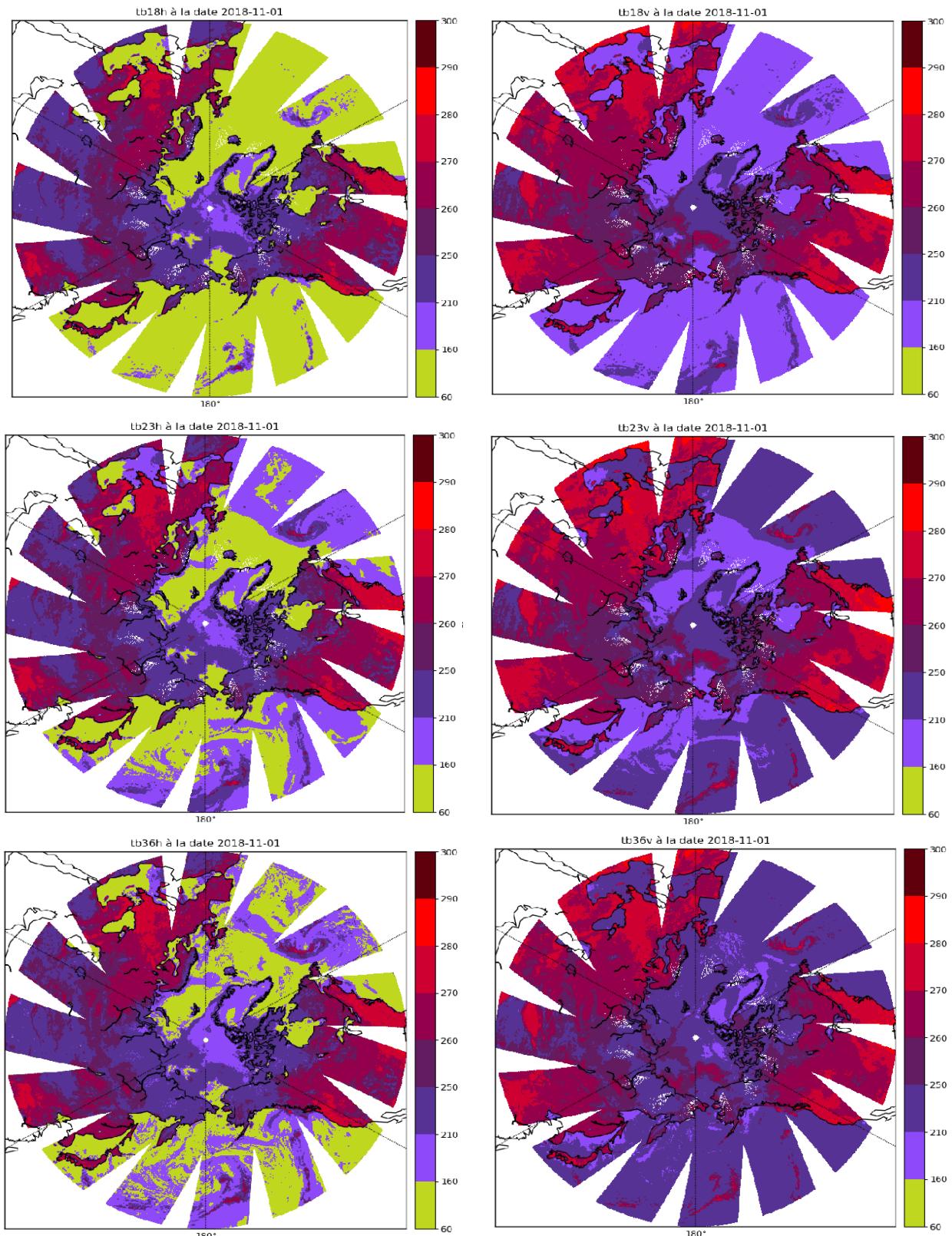


FIGURE 41 – Les températures de brillance tb18h/v, tb23h/v et tb36h/v au premier novembre 2018

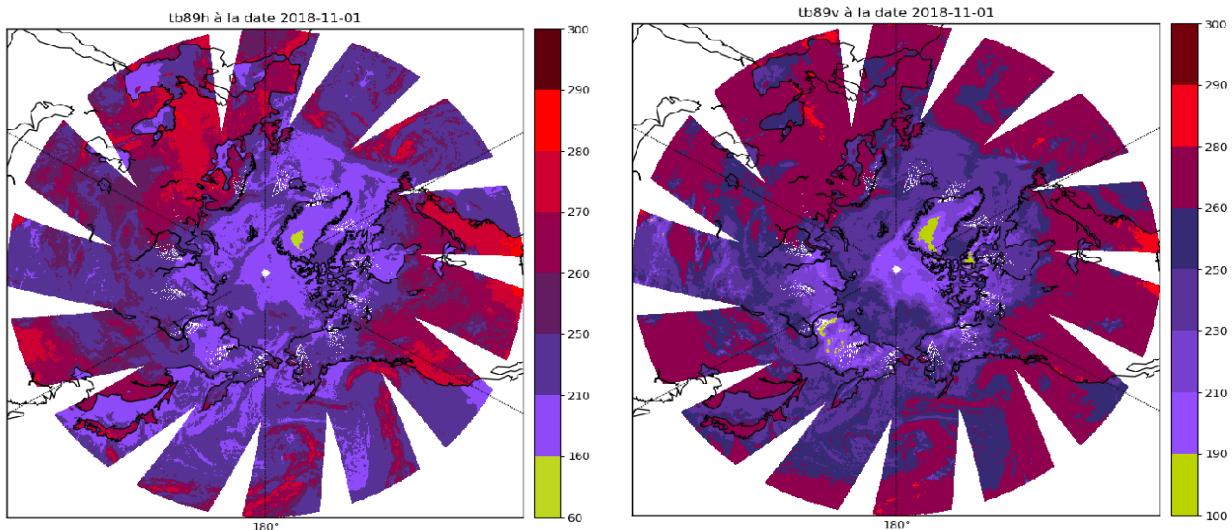


FIGURE 42 – Les températures de brillance tb89h/v au premier novembre 2018

Mesure de la qualité d'apprentissage avec les deux modèles DNN

Premier modèle

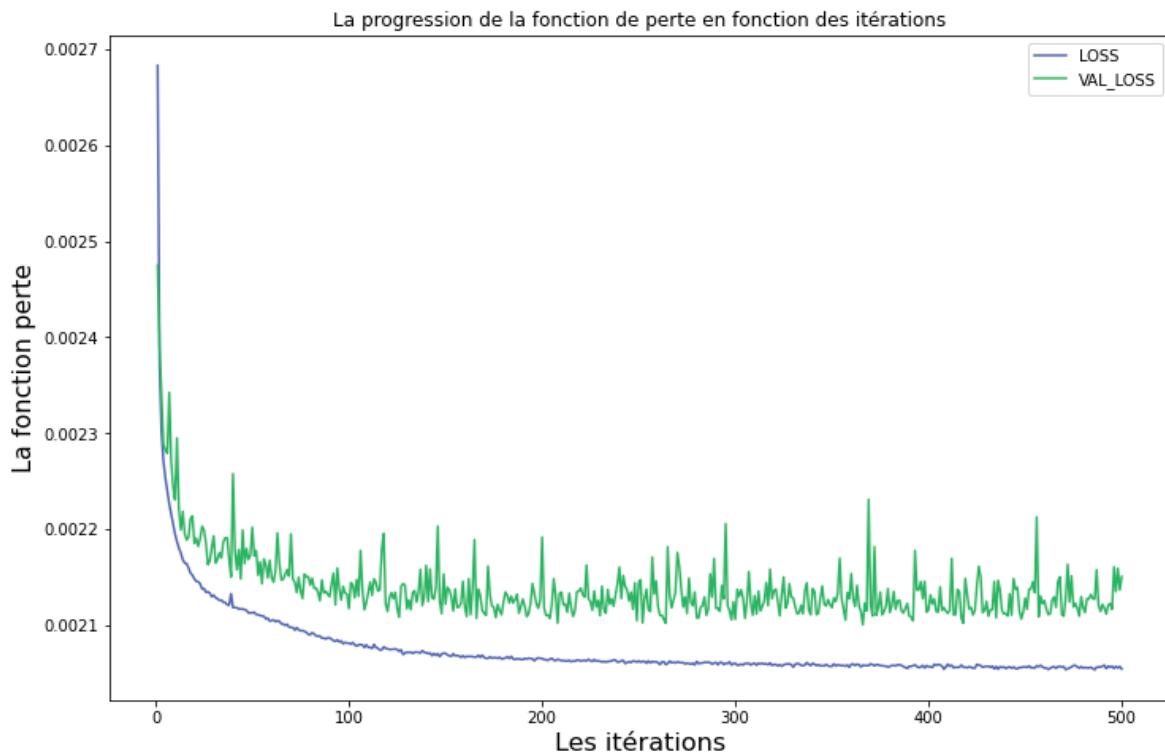


FIGURE 43 – Progression de la fonction perte aussi bien sur l'ensemble d'apprentissage (LOSS) que sur l'ensemble de validation (VAL_LOSS) en fonction des itérations

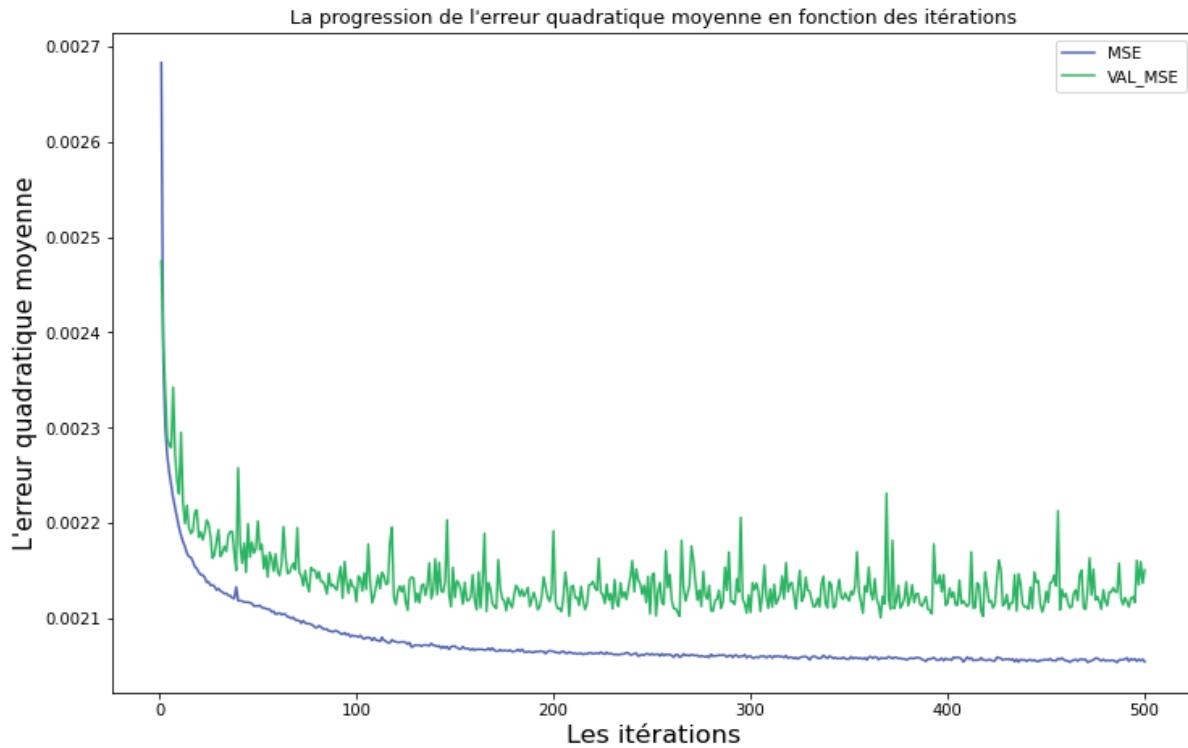


FIGURE 44 – Progression de l'erreur quadratique moyenne aussi bien sur l'ensemble d'apprentissage (MSE) que sur l'ensemble de validation (VAL_MSE) en fonction des itérations

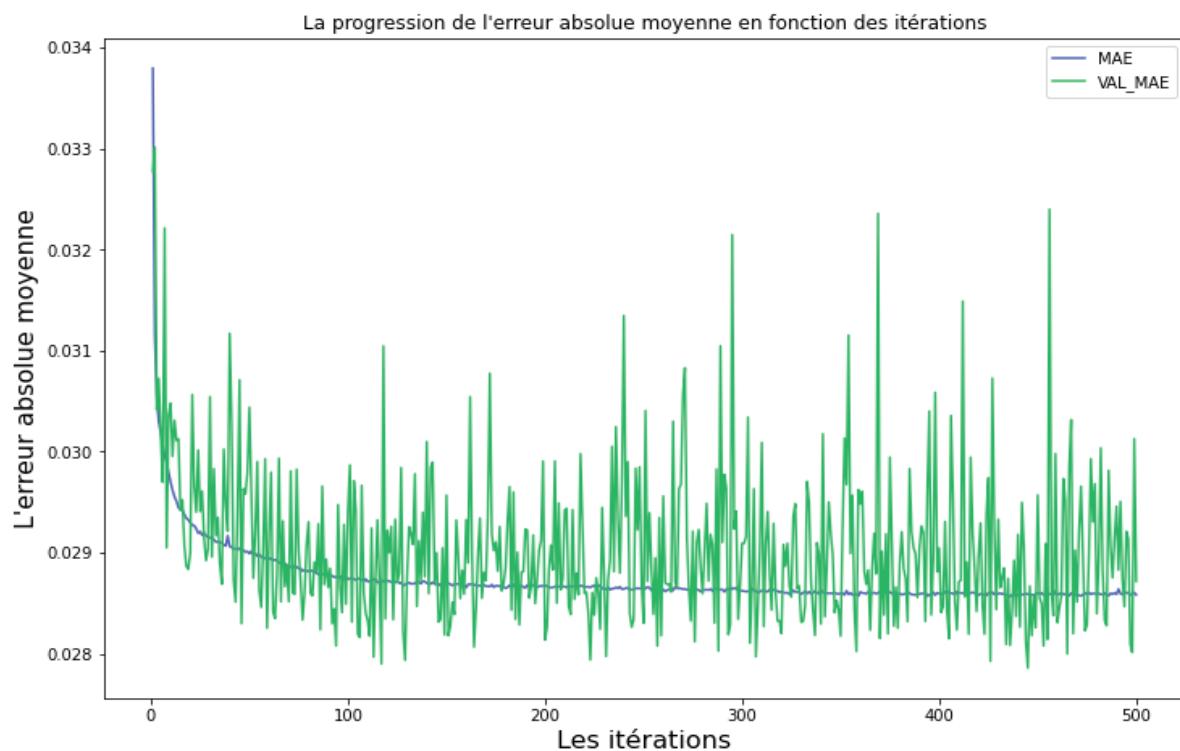


FIGURE 45 – Progression de l'erreur absolue moyenne aussi bien sur l'ensemble d'apprentissage (MAE) que sur l'ensemble de validation (VAL_MAE) en fonction des itérations

Deuxième modèle

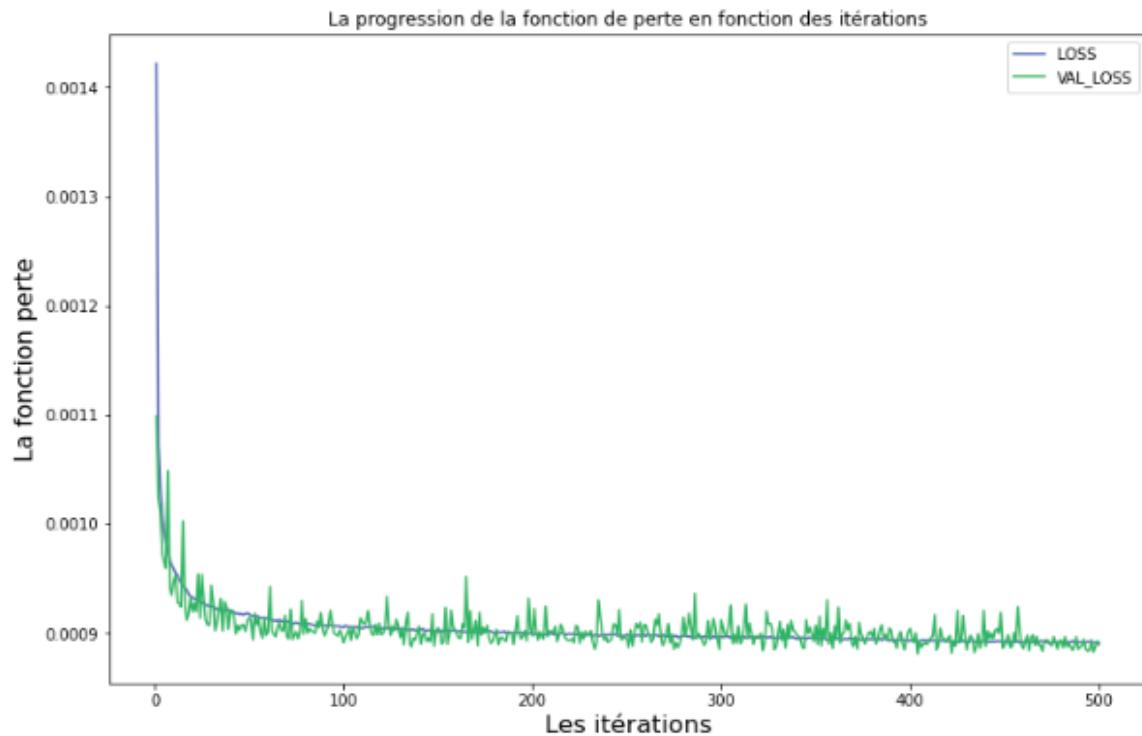


FIGURE 46 – Progression de la fonction perte aussi bien sur l'ensemble d'apprentissage (LOSS) que sur l'ensemble de validation (VAL_LOSS) en fonction des itérations pour le deuxième modèle

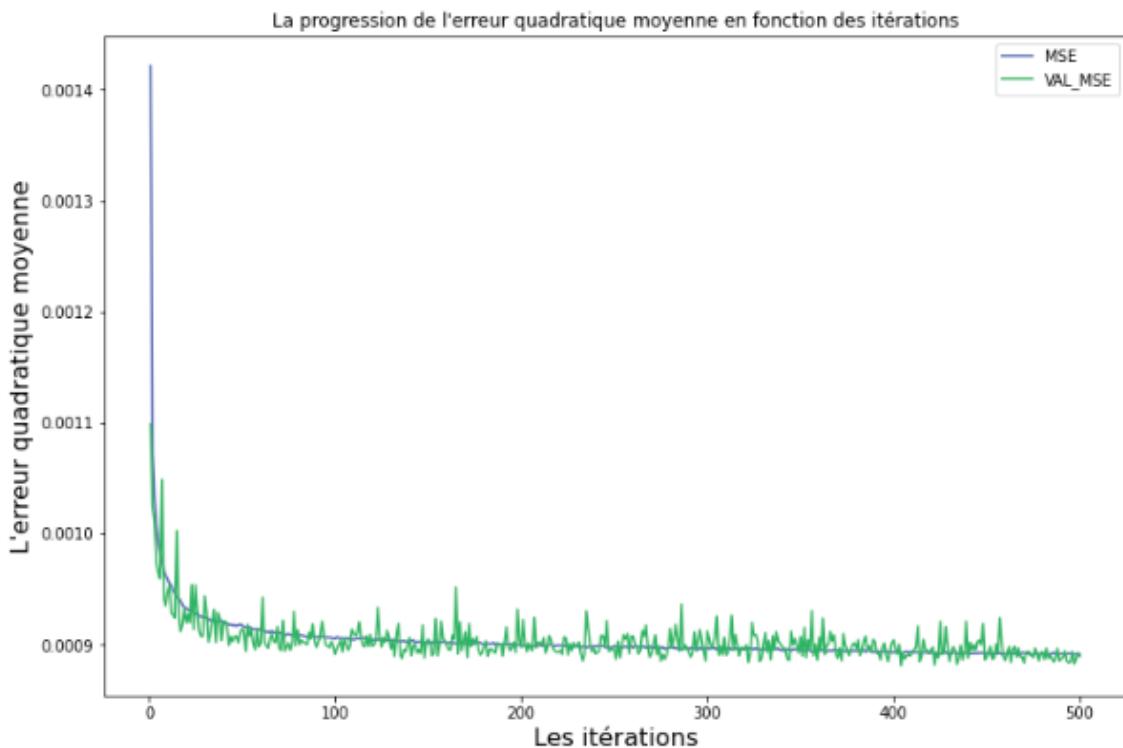


FIGURE 47 – Progression de l'erreur quadratique moyenne aussi bien sur l'ensemble d'apprentissage (MSE) que sur l'ensemble de validation (VAL_MSE) en fonction des itérations pour le deuxième modèle

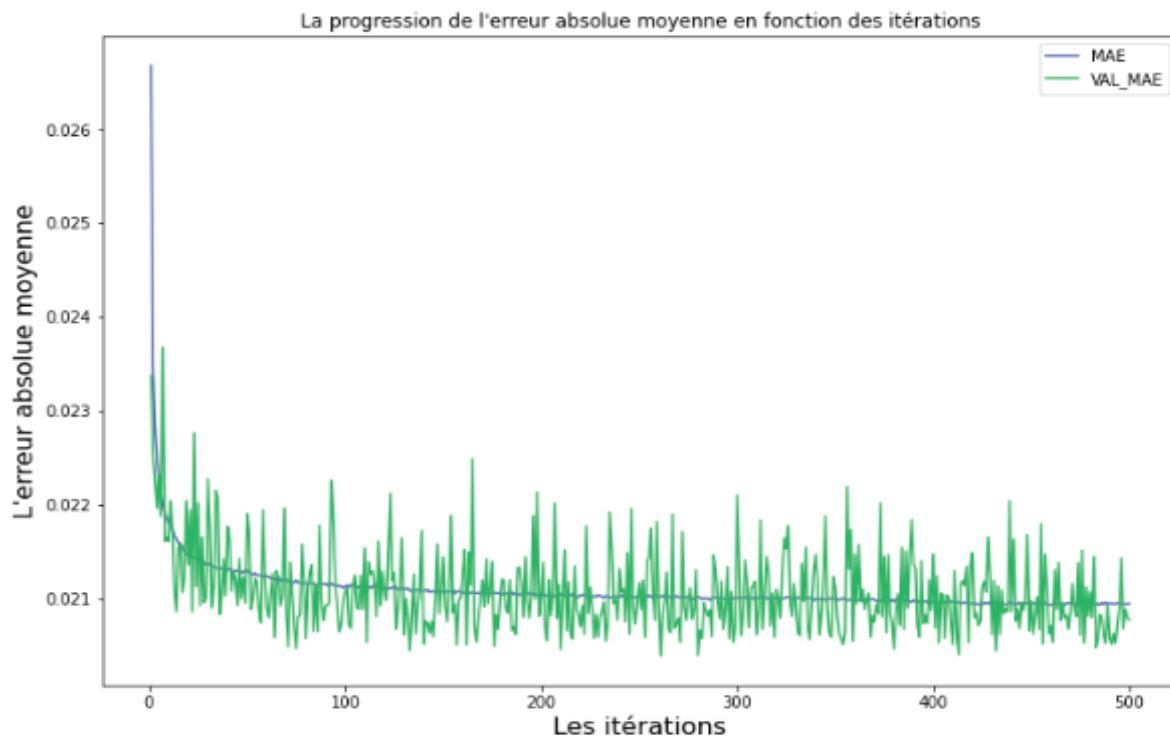


FIGURE 48 – Progression de l'erreur absolue moyenne aussi bien sur l'ensemble d'apprentissage (MAE) que sur l'ensemble de validation (VAL_MAE) en fonction des itérations pour le deuxième modèle

