

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

PRÉVISION DE L'ACTIVITÉ ÉCONOMIQUE EN TEMPS DE CRISE AU
QUÉBEC ET AU CANADA

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN ÉCONOMIQUE

PAR
MANUEL PAQUETTE-DUPOIS

OCTOBRE 2022

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de mémoire, M. Dalibor Stevanovic, pour sa patience, sa disponibilité et son appui tout au long de mon cheminement. Je veux également remercier ma famille, mes amis et mes collègues qui m'ont grandement appuyé dans la réalisation de ce mémoire.

TABLE DES MATIÈRES

LISTE DES FIGURES	v
LISTE DES TABLEAUX	vii
RÉSUMÉ	viii
INTRODUCTION	1
CHAPITRE I	
APPROCHE ÉCONOMÉTRIQUE	6
1.1 Prédiction de la probabilité de récession	6
1.2 Prédiction de l'activité économique	7
1.2.1 Modèle autorégressif	7
1.2.2 Modèles ADL	8
1.2.3 Modèles AAR	8
1.2.4 Modèles des k plus proches voisins	10
CHAPITRE II	
DONNÉES	13
2.1 Données macroéconomiques	13
2.2 Données financières	15
2.3 Données de récessions	16
CHAPITRE III	
ANALYSE EN PSEUDO HORS ÉCHANTILLON	17
3.1 Méthodologie	17
3.1.1 Exercice de prédiction des modèles de récession	18
3.1.2 Exercice de prédiction des modèles de prédiction de l'activité économique	18
3.2 Résultats	21
3.2.1 Modèles de probabilité de récession	21

3.2.2 Modèles de prévision de l'activité économique	24
CONCLUSION	30
APPENDICE A	
DONNÉES FINANCIÈRES	32
A.1 Taux d'intérêt en niveau	32
A.2 Écart de taux d'intérêt	33
APPENDICE B	
RÉSULTATS DE L'EXERCICE EN PSEUDO HORS ÉCHANTILLON SUR LA PROBABILITÉ DE RÉCESSION	34
B.1 Probabilités de récession trimestrielles avec l'écart entre les taux 10 ans et les taux 3 mois	34
B.2 Probabilités de récession trimestrielles avec les écarts de taux sélec- tionnés selon les variables prévues avec le modèle AAR	35
APPENDICE C	
ÉCARTS DE TAUX SÉLECTIONNÉS DANS LES MODÈLES AAR ET ADL	36
C.1 Écarts de taux sélectionnés dans les modèles AAR mensuels	36
C.2 Écarts de taux sélectionnés dans les modèles AAR trimestriels	39
C.3 Écarts de taux sélectionnés dans les modèles ADL mensuels	40
C.4 Écarts de taux sélectionnés dans les modèles ADL trimestriels	43
APPENDICE D	
RÉSULTATS DE L'EXERCICE DE PRÉVISION EN PSEUDO HORS ÉCHANTILLON	45
D.1 Paramètres m et k du modèle AR-kNN mensuel	45
D.2 Paramètres m et k du modèle AAR-kNN mensuel	47
D.3 Paramètres m et k du modèle ADL-kNN mensuel	48
D.4 Paramètres m et k du modèle AR-kNN trimestriel	49
D.5 Paramètres m et k du modèle AAR-kNN trimestriel	50
D.6 Paramètres m et k du modèle ADL-kNN trimestriel	51
RÉFÉRENCES	52

LISTE DES FIGURES

Figure	Page
2.1 L'évolution du niveau de l'emploi	14
2.2 L'évolution du produit intérieur brut réel	14
3.1 Probabilité de récession en échantillon avec l'écart entre le taux 10 ans et le taux 3 mois	22
3.2 Probabilité de récession en pseudo hors échantillon avec l'écart entre le taux 10 ans et le taux 3 mois	23
3.3 Probabilités de récession mensuelles en pseudo hors échantillon pour le modèle AAR	25
B.1 Probabilités de récession en pseudo hors-échantillon avec l'écart entre les taux 10 ans et les taux 3 mois	34
B.2 Probabilités de récession trimestrielles en pseudo hors-échantillon pour le modèle AAR	35
C.1 Écarts de taux sélectionnés dans les modèles AAR mensuels pour les horizons 1 à 6	37
C.2 Écarts de taux sélectionnés dans les modèles AAR mensuels pour les horizons 7 à 12	38
C.3 Écarts de taux sélectionnés dans les modèles AAR trimestriels pour les horizons 1 à 6	39
C.4 Écarts de taux sélectionnés dans les modèles AAR trimestriels pour les horizons 7 à 8	40
C.5 Écarts de taux sélectionnés dans les modèles ADL mensuels pour les horizons 1 à 6	41
C.6 Écarts de taux sélectionnés dans les modèles ADL mensuels pour les horizons 7 à 12	42

C.7	Écart de taux sélectionnés dans les modèles ADL trimestriels pour les horizons 1 à 6	43
C.8	Écart de taux sélectionnés dans les modèles ADL trimestriels pour les horizons 7 à 8	44
D.1	Paramètres m et k du modèle AR-kNN mensuel pour les horizons 1 à 6	45
D.2	Paramètres m et k du modèle AR-kNN mensuel pour les horizons 7 à 12	46
D.3	Paramètres m et k du modèle AAR-kNN mensuel pour les horizons 1 à 12	47
D.4	Paramètres m et k du modèle ADL-kNN mensuel pour les horizons 1 à 12	48
D.5	Paramètres m et k du modèle AR-kNN trimestriel pour les horizons 1 à 8	49
D.6	Paramètres m et k du modèle AAR-kNN trimestriel pour les horizons 1 à 8	50
D.7	Paramètres m et k du modèle ADL-kNN trimestriel pour les horizons 1 à 8	51

LISTE DES TABLEAUX

Tableau		Page
3.1	Ratio d'erreur quadratique moyenne des modèles de l'emploi québécois sur le modèle autorégressif	26
3.2	Ratio d'erreur quadratique moyenne des modèles de l'emploi canadien sur le modèle autorégressif	27
3.3	Ratio d'erreur quadratique moyenne des modèles du PIB québécois sur le modèle autorégressif	28
3.4	Ratio d'erreur quadratique moyenne des modèles du PIB canadien sur le modèle autorégressif	29
A.1	Taux d'intérêt sélectionnés	32
A.2	Écarts de taux d'intérêt sélectionnés	33

RÉSUMÉ

L'objectif de ce mémoire est d'améliorer les modèles de prévisions en période de crise pour l'économie canadienne et québécoise. Deux techniques sont proposées pour capter les périodes de ralentissement, soit l'ajout de la probabilité de récession au modèle autorégressif tel que proposé par Kotchoni et Stevanovic (2020) et la correction des prévisions par la méthode des k plus proches voisins développée par Gueron-Quintana et Zhong (2017). Une des particularités de la recherche est de tester plusieurs écarts de taux d'intérêt canadiens pour comparer le pouvoir prédictif de différentes portions de la structure à terme sur l'activité économique. Les variables d'intérêts sont le taux de croissance de l'emploi et le taux de croissance du PIB. Un exercice de prévision hors échantillon est réalisé pour comparer les différents modèles proposés.

INTRODUCTION

La prévision de l'activité économique est un outil essentiel de la planification en société. Que ce soit pour la conduite de la politique monétaire ou fiscale ou l'allocation de ressources des institutions gouvernementales, une quantité massive de décisions et de ressources dépendent de prévisions macroéconomiques. En plus du secteur public, celles-ci jouent un rôle majeur dans la formation des anticipations des agents économiques, dans les décisions d'investissements et dans l'évaluation d'entreprises. Les erreurs de prévisions présentent donc des coûts importants pour la société. Autour des points de retournements du cycle, une information plus juste sur l'activité économique possède une valeur primordiale, car les anticipations doivent s'ajuster rapidement aux variations brusques de l'économie.

Les récessions représentent un défi de taille puisque l'erreur des prévisions de l'activité économique varie en fonction du cycle (McCracken, 2009). Durant ces périodes, les indicateurs économiques peuvent présenter des valeurs extrêmes et des changements brusques de direction. Ces caractéristiques augmentent les erreurs de prévisions des modèles macroéconométriques qui ont de la difficulté à s'adapter rapidement pour prédire correctement les variations futures. La crise économique mondiale de 2008 est un exemple où très peu de modèles ont réussi à prédire l'amplitude du ralentissement de l'activité.

Par contre, chaque récession peut aussi être perçue comme une nouvelle source d'information sur les interactions entre les différentes variables économiques et financières au cours d'un cycle économique. Qui plus est, avec le virage vers l'ère numérique, la quantité croissante d'indicateurs suivis par une multitude d'organi-

sations et d'agences statistiques détient un potentiel considérable. Le défi qu'apporte cette masse de données est cependant d'introduire cette information dans les modèles de façon efficace. Le choix du modèle pourrait ainsi permettre de réduire l'erreur de prévision en période de crise.

L'ajout de relations non linéaires peut améliorer la prévision des récessions en période de changements brusques, mais pas nécessairement en période d'expansion (Coulombe *et al.*, 2022). Il y a donc un avantage considérable à utiliser des modèles différents en période de crise que ceux utilisés en temps normal (Chauvet *et al.*, 2013; Kotchoni *et al.*, 2019). Par exemple, les prévisions réalisées par les acteurs privés majeurs de l'économie américaine présentent des erreurs beaucoup plus élevées autour des points de retournement du cycle. Or, comme le démontre Kundan (2010), les prévisions du *Greenbook* de la Réserve Fédérale américaine font mieux en période de crise. Par contre, ces prévisions ne peuvent être utilisées en temps réel, car elles sont rendues disponibles avec un retard de cinq ans. Si ce genre de modèle a été développé pour l'économie américaine, il existe peu de littérature pour l'économie canadienne.

L'objectif de cette recherche est d'améliorer les modèles de prévisions en période de crise pour l'économie canadienne et québécoise, notamment en réduisant l'erreur de prévision. Il existe de nombreuses façons d'ajuster un modèle de prévisions pour améliorer sa performance en temps de crise. Par exemple, Coulombe *et al.* (2021) explorent différentes méthodes d'apprentissage machine avec un grand ensemble de données afin de prévoir l'activité économique au Royaume-Uni durant la crise de la COVID-19. Ils concluent que les modèles non linéaires peuvent réduire considérablement l'erreur de prévision de l'activité durant cette période. Dueker (2005) propose plutôt un modèle Qual VAR, qui intègre une variable latente décrivant le cycle économique à un modèle VAR. Il trouve que l'ajout d'information sur la probabilité de récession permet de mieux prévoir le PIB durant la récession

de 2001 aux États-Unis.

Une autre approche est d'utiliser les modèles de prévision conditionnelle, tels qu'étudiés par Clark et McCracken (2014). Un exemple de ce type de modèle est l'approche développée par Kotchoni et Stevanovic (2020) où les auteurs ajoutent la probabilité de récession à une composante autorégressive pour améliorer le pouvoir prédictif du modèle. Ils utilisent un modèle probit incluant des variables financières pour modéliser le cycle. L'ajout non linéaire de ces variables dans le modèle permet de réduire l'erreur de prévision un an à l'avance en période de récession pour l'économie américaine. Ils améliorent ainsi la précision de la prévision de la croissance du PIB jusqu'à 30% par rapport au modèle de référence en période de récession.

Il existe plusieurs méthodes pour prévoir les probabilités de récession. Estrella et Mishkin (1998) utilisent la structure à terme des taux d'intérêt pour prévoir les récessions américaines. Les auteurs concluent qu'un modèle probit utilisant l'écart entre les taux à 10 ans et les taux à 3 mois sur les obligations du trésor américains parviennent à bien capter les points de retournements aux États-Unis. D'autres approches plus sophistiquées ont aussi été développées pour prévoir les récessions, par exemple Chauvet et Potter (2005) prennent en compte la présence de bris structurel dans la dynamique de la probabilité de récessions. Rudebusch et Williams (2009) montrent qu'un modèle utilisant la structure à terme en temps réel prédit mieux les récessions plus de deux trimestres à l'avance que les prévisionnistes professionnels.

Au Canada, l'écart entre les taux d'intérêt des obligations du gouvernement canadien de dix ans et plus et les taux de papiers commerciaux à 90 jours est un bon prédicteur du cycle (Clinton, 1995; Bernard et Gerlach, 1998). Par contre, l'utilisation de ce genre de modélisation du cycle nécessite d'observer les périodes

de récession passées. Or, au Canada, aucun organisme officiel n'établit les dates de récessions contrairement aux États-Unis où un comité du *National Bureau of Economic Research* est responsable de la tâche. Une façon de contourner ce problème est d'utiliser un modèle de Markov à changement de régime (Goodwin, 1993) ou d'utiliser les périodes de récession établies par l'Institut C.D. Howe (Cross et Bergevin, 2012).

Dans le but de prévoir d'améliorer la précision des prévisions en temps de crise, Guerron-Quintana et Zhong (2017) utilisent un modèle non paramétrique pour modéliser l'activité économique. Ils utilisent une méthode empruntée à l'apprentissage machine, soit la méthode des k plus proches voisins. La première étape est d'établir la séquence temporelle qui minimise la distance avec les données récentes. Ensuite, l'erreur de prévision du modèle sur cette séquence est ajoutée à la prévision du modèle dans h périodes. Les auteurs trouvent que leur modèle améliore la prévision d'un à douze mois à l'avance en période de crise par rapport aux modèles alternatifs paramétriques linéaires et non linéaires. En plus de bien capter les points de retournement rapides de plusieurs indicateurs économiques comme l'inflation, le taux de chômage et le revenu réel, l'approche produit aussi des prévisions précises en période d'expansion. Avec le même objectif, Foroni *et al.* (2022) montrent qu'une forme spécifique de l'approche des k plus proches voisins, la correction de l'ordonnée à l'origine, permet de mieux prédire la croissance aux États-Unis durant la crise de la Covid-19. Taieb et Hyndman (2012) utilise également la méthode des k plus proches voisins pour faire des prévisions sur des séries temporelles. Ils apportent une correction à la prévision itérative en ajoutant la modélisation de l'erreur de prévision réalisée de façon directe. La modélisation avec les voisins les plus proches permet de réduire la variance et le biais.

Cette recherche tente de produire les meilleures prévisions de l'activité économique en temps de crise en utilisant les techniques employées par Guerron-Quintana et

Zhong (2017) et Kotchoni et Stevanovic (2020). Le produit intérieur brut réel canadien et québécois ainsi que le niveau d'emploi sont les indicateurs choisis pour décrire l'activité économique. Le chapitre 1 présentera le modèle de prévision de la probabilité de récession et les modèles retenus prévoir l'activité économique. Ensuite, le chapitre 2 présentera les données utilisées. Finalement, le chapitre 3 exposera la méthodologie et les résultats de l'exercice de prévision hors échantillon pour comparer les différents modèles proposés.

CHAPITRE I

APPROCHE ÉCONOMÉTRIQUE

1.1 Prédiction de la probabilité de récession

La modélisation de la probabilité de récession est un instrument clé dans la prévision de l'activité économique en temps de crise. Un modèle probit, similaire à celui proposé par Estrella et Mishkin (1998), est utilisé pour modéliser les points de retournements du cycle. Ici, R_t est un indicateur de récession au temps t et X_t contient un ensemble de prédicteurs. L'indicateur de récession prend la valeur de 1 en cas de récession et 0 sinon. Le modèle prend donc la forme :

$$R_{t+h} = \Phi(X_t' \gamma_h) + u_{h,t} \quad \forall h, t, \quad R_{t+h} \in \{0, 1\}.$$

La prévision de la probabilité de récession est donnée par :

$$Pr(R_{t+h} = 1 | X_t) = \Phi(X_t \gamma_h). \quad (1.1)$$

Elle devient faisable lorsque le paramètre γ_h est estimé sur l'échantillon observé. La particularité de cette approche est la nécessité d'observer les dates de récession R_t , contrairement à la méthode de changement de régime markovien où la probabilité d'état de nature est modélisée de façon endogène. Nous privilégions ici la première approche à cause de sa simplicité, sa robustesse et l'inclusion facile des indicateurs avancés dans l'ensemble d'information X_t .

Par ailleurs, nous testerons plusieurs ensembles d'information X_t . Bien que l'écart

entre le taux 10 ans et le taux 3 mois soit communément utilisé dans la littérature pour mesurer la pente de la courbe de rendement, il n'a pas nécessairement le meilleur pouvoir prédictif pour prévoir l'activité économique. Ainsi différents écarts de taux seront examinés, créant plusieurs vecteurs de récession. Les détails seront expliqués plus tard dans l'application.

1.2 Préviation de l'activité économique

Plusieurs modèles sont considérés pour prévoir l'activité économique, y_t . L'objectif est de prévoir l'activité économique sur plusieurs horizons, h . Deux modèles de référence sont considérés, soit le modèle autorégressif, AR, et le modèle autorégressif à retards échelonnés, ADL. Ensuite, deux modèles ayant comme objectif de réduire l'erreur de prévision en temps de crise sont utilisés, soit le modèle autorégressif augmenté, AAR, et le modèle des k plus proches voisins, kNN. Dans tous les cas, la prévision directe sera considérée, i.e. que la cible y_{t+h} est projetée directement sur l'ensemble d'information X_t pour tout horizon de prévision h .

1.2.1 Modèle autorégressif

Le modèle de référence, dénoté AR, est un modèle autorégressif d'ordre 1. Il prend la forme :

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \nu_{t+h}.$$

Et la prévision h périodes à l'avance est donnée par :

$$y_{t+h|t} = \mathbf{E}(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t.$$

Afin de produire la prévision, les paramètres $\rho_{h,0}$ et $\rho_{h,1}$ sont estimés par les moindres carrés ordinaires.

1.2.2 Modèles ADL

La deuxième classe de modèle considérée est Autoregressive Distributed Lag (ADL). Ici, un ensemble de variables explicatives, X_t , est ajouté à la composante autorégressive. Le modèle s'écrit :

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + X_t'\beta_h + \tilde{v}_{t+h} \quad (1.2)$$

et la prévision est donnée par :

$$y_{t+h|t} = \mathbf{E}(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t + X_t\beta_h.$$

Comme dans le modèle AR, les paramètres sont estimés par les moindres carrés ordinaires. À noter que des spécifications plus riches des modèles AR et ADL peuvent être utilisées, en particulier en permettant une structure de retards plus compliquée. Nous optons quand même pour les modèles les plus parcimonieux et faisons varier considérablement les horizons de prévision.

1.2.3 Modèles AAR

Le troisième modèle fait partie de la classe des modèles autorégressifs augmentés tels que considérés par Kotchoni et Stevanovic (2020). En plus d'une composante autorégressive, il incorpore la prévision du point de retournement du modèle probit. Le modèle prédictif prend la forme suivante :

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \delta_h R_{t+h} + v_{t+h}, \quad (1.3)$$

pour $t = 1, \dots, T - h$, où $v_{t+h} \sim N(0, \sigma_h^2)$. Cette innovation peut être corrélée avec R_{t+h} mais non corrélée avec les valeurs retardées de y_t . Ce modèle n'est pas faisable en temps réel puisque R_{t+h} n'est pas connu en t .

En prenant l'espérance conditionnelle à l'information observée en temps t de l'équation 1.3, nous obtenons

$$E(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t + \delta_h \Pr(R_{t+h} = 1|y_t, X_t).$$

Nous assumons que la probabilité de récession en $t + h$ dépend seulement de X_t . Ceci est pour diminuer le risque de surajustement si les valeurs historiques de y_t ont été utilisées pour inférer sur les dates de récession R_t . La probabilité de récession est d'ailleurs modélisée par un probit comme dans la section 1.1. Dans ce cas, la valeur espérée de y_{t+h} devient

$$E(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t + \delta_h \Phi(X_t' \gamma_h).$$

Par conséquent, le modèle prédictif AAR peut s'écrire

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \delta_h \Phi(X_t' \gamma_h) + \tilde{v}_{t+h}, \quad (1.4)$$

où $\tilde{v}_{t+h} \equiv v_{t+h} + \delta_h (R_{t+h} - \Phi(X_t' \gamma_h))$ est un terme d'erreur. Notons que ce modèle exploite l'information X_t d'une façon non linéaire comparée au modèle ADL dans 1.2 qui fait intervenir les prédicteurs de manière linéaire. De plus, lorsque la probabilité de récession $\Phi(X_t' \gamma_h)$ est proche de zéro, ces prédicteurs n'ont pas d'effet sur la prévision de y_{t+h} . Dans le cas du modèle ADL, l'effet marginal de X_t est constant (donné par β_h). La comparaison entre ce modèle et le modèle AAR avec le même ensemble de variables explicatives permet de déterminer si le pouvoir explicatif de l'ensemble est amélioré par la non-linéarité.

La prévision directe du modèle s'écrit comme :

$$y_{t+h|t} = \mathbf{E}(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t + \delta_h \Phi(X_t' \gamma_h). \quad (1.5)$$

À noter qu'il s'agit d'une procédure à deux étapes : (1) modélisation prédictive de la probabilité de récession ; (2) modélisation prédictive de la variable cible y . Nous avons choisi de procéder avec un modèle simple et robuste dans la première

partie, mais n'importe quel autre modèle de détection de probabilité de récession peut s'implémenter.

1.2.4 Modèles des k plus proches voisins

Le modèle des k plus proches voisins (kNN) peut être utilisé pour améliorer la prévision de séries temporelles (Farmer et Sidorowich, 1987; Yakowitz, 1987; Diebold et Nason, 1990; Mulhern et Caprara, 1994). Ici, les deux algorithmes d'appariement proposés sont inspirés de la méthodologie de Guerron-Quintana et Zhong (2017). L'idée générale de cette méthode est de trouver la séquence dans une série observée dont les variations sont les plus semblables à celles de la dernière séquence observée. La série $Y_T = y_1, y_2, \dots, y_T$ est décomposée en t_n séquences de longueur k . L'ensemble des séquences prend la forme $C_j = \{[y_{t_1}, y_{t_1+1}, \dots, y_{t_1+k-1}], [y_{t_2}, y_{t_2+1}, \dots, y_{t_2+k-1}], \dots, [y_{t_n}, y_{t_n+1}, \dots, y_{t_n+k-1}]\}$.

Chaque algorithme d'appariement choisira une séquence de l'ensemble C_j et l'assignera à la séquence la plus récente $[y_{T-k+1}, y_{T-k+2}, \dots, y_T]$. Le premier algorithme d'appariement compare les différentes séquences en niveau directement. Par exemple, si les séquences $[y_{t+1-k}, \dots, y_t]$ et $[y_1, y_2, \dots, y_k]$ sont comparées, l'algorithme consiste à minimiser la fonction de distance W_{niv} :

$$W_{niv} = \sum_{i=1}^k \omega(i)(y_{t+1-i} - y_{k+1-i})^2. \quad (1.6)$$

Le second algorithme d'appariement compare les séquences en déviations par rapport à leur moyenne locale. La fonction de distance, W_{moy} , est donnée par :

$$W_{moy} = \sum_{i=1}^k \omega(i)((y_{t+1-i} - \bar{y}_t) - (y_{k+1-i} - \bar{y}_k))^2 \quad (1.7)$$

où \bar{y}_t représente la moyenne de la séquence :

$$\bar{y}_t = \frac{1}{k} \sum_{i=1}^k y_{t-k+i}.$$

Dans les deux cas, les poids accordés à chacune des observations décroissent linéairement avec i , de sorte à accorder le poids le plus élevé aux observations les plus récentes :

$$\omega(i) = \frac{1}{i}.$$

En minimisant la fonction de distance W_{niv} , l'algorithme d'appariement en niveau prend en compte à la fois le niveau des observations dans chacune des séquences et la direction de ses variations. Avec la seconde fonction de distance, W_{moy} , l'algorithme en déviation par rapport à moyenne considère seulement la direction des variations.

L'objectif est ensuite d'ajuster la prévision d'un modèle avec l'erreur du même modèle durant la séquence passée la plus similaire. Par exemple, avec un modèle AR, la prévision s'exprime comme :

$$\hat{y}_{t+h} = (y_{t_1+k+h} - \hat{y}_{t_1+k+h,AR}) + \hat{y}_{t+h,AR}$$

où $\hat{y}_{t_1+k+h,AR}$ est la prévision du modèle AR h périodes suivant la dernière observation de la séquence appariée et y_{t_1+k+h} est la valeur observée.¹ Il est également possible de sélectionner les m séquences qui minimisent la fonction de distance, la prévision devient alors :

$$\hat{y}_{t+h} = \frac{1}{m} \sum_{i=1}^m (y_{t(i)+k+h} - \hat{y}_{t(i)+k+h,AR}) + \hat{y}_{t+h,AR}$$

où $[t(i), t(i) + 1, \dots, t(i) + k]$ est la i ème séquence la plus rapprochée de la séquence la plus récente. Les paramètres k et m sont déterminés en minimisant l'erreur quadratique moyenne parmi les km combinaisons de paramètres possibles :

$$EQM = \frac{1}{t} \sum_{i=1}^t (y_t - \hat{y}_t)^2.$$

1. Le conditionnement à l'ensemble d'information est omis ici seulement pour faciliter l'écriture.

En exercice hors échantillon, comme les paramètres sont sélectionnés à chaque itération, il faut choisir une fenêtre sur laquelle sélectionner les paramètres. Cette fenêtre commence à t_1 et se termine à l'observation t . Une prévision est produite avec la méthode kNN pour les observations de t_1 à $t - h$ avec chacune des combinaisons de paramètres. Les paramètres sont ensuite sélectionnés comme :

$$\operatorname{argmin}_{k,m} EQM_{t,km} = \frac{1}{t - t_1} \sum_{\tau=t_1}^t (y_\tau - \hat{y}_{\tau,km})^2.$$

CHAPITRE II

DONNÉES

2.1 Données macroéconomiques

Plusieurs indicateurs sont disponibles pour mesurer l'activité économique au Québec et au Canada. Pour s'assurer de bien évaluer le pouvoir prédictif de chacun des modèles, il est important de tenir compte des échantillons disponibles pour chacun des indicateurs. Les indicateurs avec un délai de publication plus court sont aussi privilégiés, car leur information peut être utilisée plus promptement dans le processus décisionnel. Selon ces critères, deux ensembles de données ont été retenus au Québec et au Canada en fréquence mensuelle et trimestrielle.

Le premier ensemble est composé des données d'emplois provenant de l'enquête sur la population active en fréquence mensuelle au Québec et au Canada (Statistique Canada, Tableau 14-10-0287-01). Ces deux séries mesurent le nombre de travailleurs âgés de 15 ans et plus chaque mois et sont désaisonnalisées par Statistique Canada. La figure 2.1 montre les deux séries en niveau. Les résultats de l'enquête sur la population active sont les premières à être rendues publiques parmi les principaux indicateurs d'activité, les résultats étant publiés seulement 10 jours après la fin de la collecte de donnée. L'échantillon utilisé s'étale de janvier 1976 à décembre 2019.

Le deuxième ensemble de données est composé du produit intérieur brut en termes

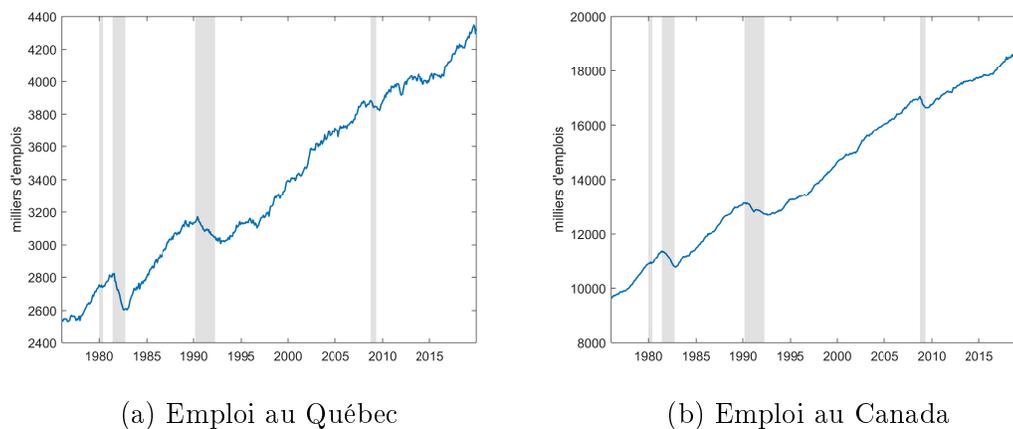


Figure 2.1: L'évolution du niveau de l'emploi

de dépenses en dollars enchaînés de 2012 pour le Québec et le Canada (Statistique Canada, Tableau 36-10-0104-01). Les deux séries proviennent des comptes nationaux canadiens et sont désaisonnalisées par l'agence statistique. La figure 2.2 montre le niveau du PIB au Québec et au Canada. Les comptes nationaux sont publiés dans les 60 à 75 jours suivant la période de référence. L'échantillon utilisé s'étale du premier trimestre 1981 au quatrième trimestre 2019.

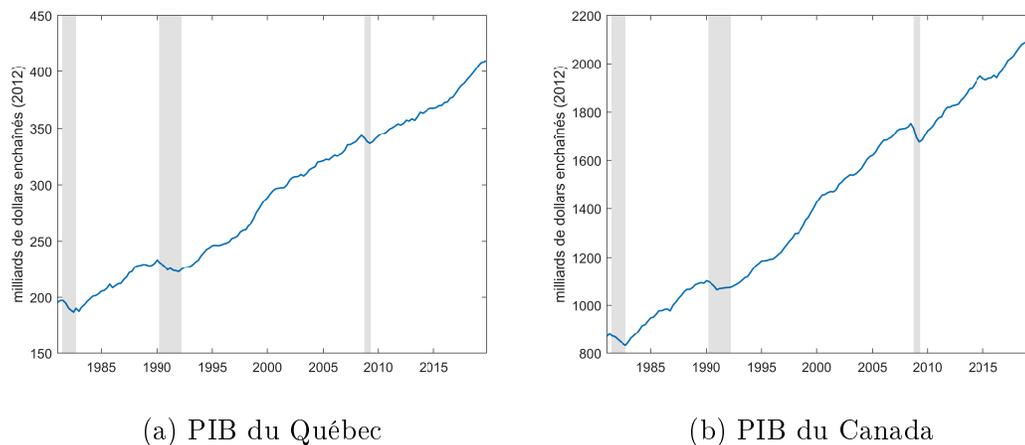


Figure 2.2: L'évolution du produit intérieur brut réel

Les données d'emplois et de PIB sont stationnarisées en utilisant une différence de

logarithme. Le test de Dickey-Fuller augmenté Dickey et Fuller (1981) est exécuté sur chacune des séries pour s'assurer de leur stationnarité.

2.2 Données financières

Les données financières utilisées sont un ensemble de taux de rendement compilés par la Banque du Canada (Statistique Canada, Tableau 10-10-0122-01). Cette base de données contient 33 séries de différents taux de rendement canadiens en fréquence mensuelle. Elle comprend, entre autres, le taux officiel d'escompte de la Banque du Canada, le taux de rendement des certificats de dépôts des entreprises et le taux des bons du Trésor et des obligations canadiennes à différentes maturités. Pour maximiser l'échantillon utilisé dans l'estimation des probabilités de récessions, seules huit séries de taux de rendement sont sélectionnées. L'échantillon s'étend de janvier 1973 à décembre 2019. Ensuite, 22 écarts de taux sont calculés pour mesurer différentes portions de la courbe de rendements. La liste complète des séries sélectionnées est disponible en appendice A. Cet ensemble de données permettra de déterminer quelle portion de la structure à terme contient le plus de pouvoir prédictif sur l'activité économique réelle. Deux autres méthodes sont aussi étudiées pour tenter d'extraire l'information des taux d'intérêt. La première consiste à diviser l'ensemble de variables financières selon leur maturité. Un groupe comprend les taux de rendement de court terme, avec une maturité de un an et moins, et l'autre groupe comprend les taux d'intérêt de long terme, avec une maturité de plus d'un an. Le premier facteur est ensuite extrait des deux groupes selon Stock et Watson (2002) et l'écart entre le facteur de long terme et celui de court terme est calculé. La deuxième méthode consiste à extraire le premier facteur de l'ensemble des 22 écarts de taux considérés. Pour l'exercice trimestriel, la moyenne des taux de rendement est calculée sur le trimestre.

2.3 Données de récessions

Les dates de récessions établies par le conseil des cycles économiques de l'institut canadien C.D. Howe¹ sont utilisées pour calibrer le modèle de probabilité de récession. Depuis la publication du premier rapport sur les dates de récessions canadiennes depuis 1926 (Cross et Bergevin, 2012), le conseil des cycles économiques se réunit annuellement pour déterminer les points de retournement du cycle. Pour l'exercice sur les données du Québec, les récessions québécoises sont considérées comme étant les mêmes que les récessions canadiennes.

1. Les dates de récession peuvent être retrouvées sur <https://www.cdhowe.org/council/business-cycle-council>.

CHAPITRE III

ANALYSE EN PSEUDO HORS ÉCHANTILLON

Ce chapitre aborde la méthodologie et les résultats de l'analyse des modèles présentés au chapitre 1. L'objectif est de déterminer quel modèle offre le meilleur pouvoir prédictif. L'analyse peut être séparée en deux étapes distinctes. D'abord, les probabilités de récession historiques sont estimées avec différents ensembles de données. Dans un second temps, les modèles de prévisions de l'activité économique sont estimés et on compare leur performance historique.

3.1 Méthodologie

Une course de modèle sera étudiée pour chacun des quatre indicateurs d'activité économique présentés au chapitre 2. L'objectif est de répliquer l'utilisation de chaque modèle à travers le temps avec les données qui auraient été disponibles à chaque étape. La performance de chaque modèle sera évaluée de janvier 1990 à décembre 2019 pour les données d'emplois et du premier trimestre de 2000 au quatrième trimestre de 2019 pour les données de PIB. L'exercice trimestriel est plus court, car il faut s'assurer d'avoir un échantillon suffisamment long pour faire la première prévision. Pour chaque période de temps à l'étude, les paramètres de chaque modèle seront estimés. Plusieurs horizons de prévisions sont aussi testés pour chaque exercice. La performance des modèles mensuels est testée avec des horizons de prévision de un à douze mois, et les modèles trimestriels avec des hori-

zons de un à huit trimestres. Par exemple, pour la prévision de l'emploi canadien de janvier 1990 à l'horizon 12 mois, chaque modèle est estimé avec les données allant jusqu'à janvier 1989.

Il est important de noter qu'il ne s'agit pas d'un exercice en temps réel, les révisions dans les données et les délais de parution ne sont pas pris en compte. Il ne s'agit donc pas d'une réplique exacte de l'utilisation du modèle à travers le temps, mais plutôt d'une approximation. Même si les données d'emplois et de PIB sont parfois révisées par Statistique Canada, les intrants des modèles, soit les variables financières, ne sont pas révisés. L'impact sur les résultats présentés plus loin est donc moins important que dans un exercice utilisant des données macroéconomiques comme intrants.

3.1.1 Exercice de prévision des modèles de récession

L'estimation des probabilités de récession est la première étape des modèles AAR. Pour chaque fréquence et chaque horizon, les probabilités de récession sont estimées avec les données disponibles à chaque période de prévision. On cherche à connaître quelle portion de la courbe de rendement donne le plus d'information sur la probabilité de récession, un vecteur de probabilité est donc estimé pour chaque écart de taux de rendement. Le vecteur avec le meilleur pouvoir prédictif sera sélectionné à la seconde étape des modèles AAR.

3.1.2 Exercice de prévision des modèles de prévision de l'activité économique

Le calcul de l'erreur moyenne de chaque modèle permettra de déterminer quels modèles ont le meilleur pouvoir prédictif. Le modèle AR servira de modèle de référence, il sera comparé à l'erreur moyenne des autres modèles. Comme on ne connaît pas la meilleure spécification des modèles ADL et AAR, et qu'elle peut évoluer dans le temps, on doit utiliser une méthode de sélection de modèles. Le

critère d'information bayésien (BIC) est choisi. À chaque période prévue, l'écart de taux sélectionné pour le modèle ADL et AAR est le résultat de :

$$\underset{i}{\operatorname{argmin}} BIC_{t,n,i} = t \ln(\widehat{\sigma}_{t,n,i}^2) + k \ln(t)$$

où

$$\widehat{\sigma}_{t,n,i}^2 = \frac{1}{t} \sum_{j=1}^t (y_{j,n,i} - \widehat{y}_{j,n,i})^2$$

et i représente un écart de taux de rendement parmi les 24 étudiés et j représente le modèle, ADL ou AAR. L'intrant choisi dans chacun des modèles peut donc évoluer dans le temps, et différer d'un modèle et d'un horizon de prévision à l'autre.

La méthode kNN est ensuite utilisée sur les modèles AR, ADL et AAR. L'ensemble d'information utilisé dans les modèles AAR-kNN et ADL-kNN pour chaque variable et chaque horizon est le même que celui des modèles AAR et ADL. Les hyperparamètres k et m sont sélectionnés à chaque période de prévision avec la méthode présentée à la section 1.2.4. Le paramètre k est testé sur l'intervalle de 2 à 18 pour l'exercice mensuel et de 2 à 10 pour l'exercice trimestriel. Les séquences choisies peuvent donc s'étaler jusqu'à 18 mois pour les variables d'emploi jusqu'à 10 trimestres pour les variables de PIB. Le paramètre m est testé sur l'intervalle 2 à 20 pour les deux exercices.

La comparaison des modèles de prévisions présentés dans la prochaine section est réalisée avec l'erreur quadratique moyenne. Après avoir estimé chacun des modèles en pseudo hors échantillon, la valeur prévue, \hat{y}_t est comparée à la vraie valeur observée, y_t avec :

$$EQM = \frac{1}{T} \sum_{t=1}^t (y_t - \hat{y}_t)^2.$$

Ensuite, des tests de Diebold et Mariano (1995) sont effectués pour vérifier si la performance des modèles étudiés est significativement supérieure au modèle de référence AR. Pour réaliser le test, la première étape consiste à calculer la différence entre les erreurs quadratiques de chacun des modèles. Ainsi, pour un chacun des modèles, cette différence est définie comme :

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$$

où

$$d_t = e_{t+h}^2 - e_{t+h,AR}^2$$

et $e_{t+h} = y_{t+h} - \hat{y}_{t+h}$. L'hypothèse nulle et l'hypothèse alternative sont données par :

$$H_0 : \bar{d} = 0$$

$$H_1 : \bar{d} \neq 0.$$

Sous l'hypothèse nulle, la statistique de test est :

$$DM = \frac{\bar{d}}{\hat{\sigma}_{\bar{d}}} \sim \mathcal{N}(0, 1)$$

où

$$\hat{\sigma}_{\bar{d}} = \sqrt{\frac{1}{T} \left(\hat{\gamma}_0 + 2 \sum_{\tau=1}^{h-1} \hat{\gamma}_{\tau} \right)}$$

et

$$\hat{\gamma}_{\tau} = \frac{1}{T} \sum_{t=\tau+1}^T (d_t - \bar{d})(d_{t-\tau} - \bar{d}).$$

Un rejet l'hypothèse nulle du test de Diebold Mariano suggère donc qu'un modèle a une performance significativement différente du modèle de référence.

3.2 Résultats

3.2.1 Modèles de probabilité de récession

La première étape de la modélisation AAR est d'estimer un modèle de probabilité de récession. Comme illustré dans la section 2, un simple modèle probit statique a été choisi afin de modéliser et prévoir les points de retournement du cycle de l'économie canadienne et québécoise. Une stratégie commune pour modéliser la probabilité de récession est d'utiliser l'écart entre le taux moyen des obligations du gouvernement du Canada à 10 ans et plus et le taux des bons du trésor canadien à 3 mois comme utilisé par Clinton (1995). La figure 3.1 montre les prévisions mensuelles du modèle probit aux horizons trois, six, neuf et douze mois sur l'échantillon complet. Les bandes grises représentent les périodes de récession de l'Institut C.D. Howe.

Sur l'échantillon complet, l'écart entre le taux 10 ans et le taux 3 mois réussi à bien prédire les récessions de 1980, 1981 et 1990 pour tous les horizons. Plusieurs mois avant chacune de ces récessions, la courbe de rendement canadienne s'est inversée, c'est-à-dire que les taux longs avaient un rendement inférieur à celui des taux courts. Une inversion de courbe peut se produire lorsque les investisseurs s'attendent à un ralentissement de la croissance dans le futur, ce qui fait pression à la hausse sur le prix des obligations à long terme et à la baisse sur le rendement. Pendant ce temps, si la banque centrale maintient un taux directeur élevé, les obligations de court terme maintiennent leur prix et leur rendement, créant un écart négatif entre les rendements de court terme et ceux de long terme. La récession de 2008 est moins bien captée par le modèle, puisque dans les mois précédant cette crise, l'écart entre les taux 10 ans et ceux de 3 mois est devenu négatif, mais pas avec la même amplitude que durant les trois récessions précédentes.

La figure 3.2 montre les prévisions du même modèle probit en pseudo hors échan-

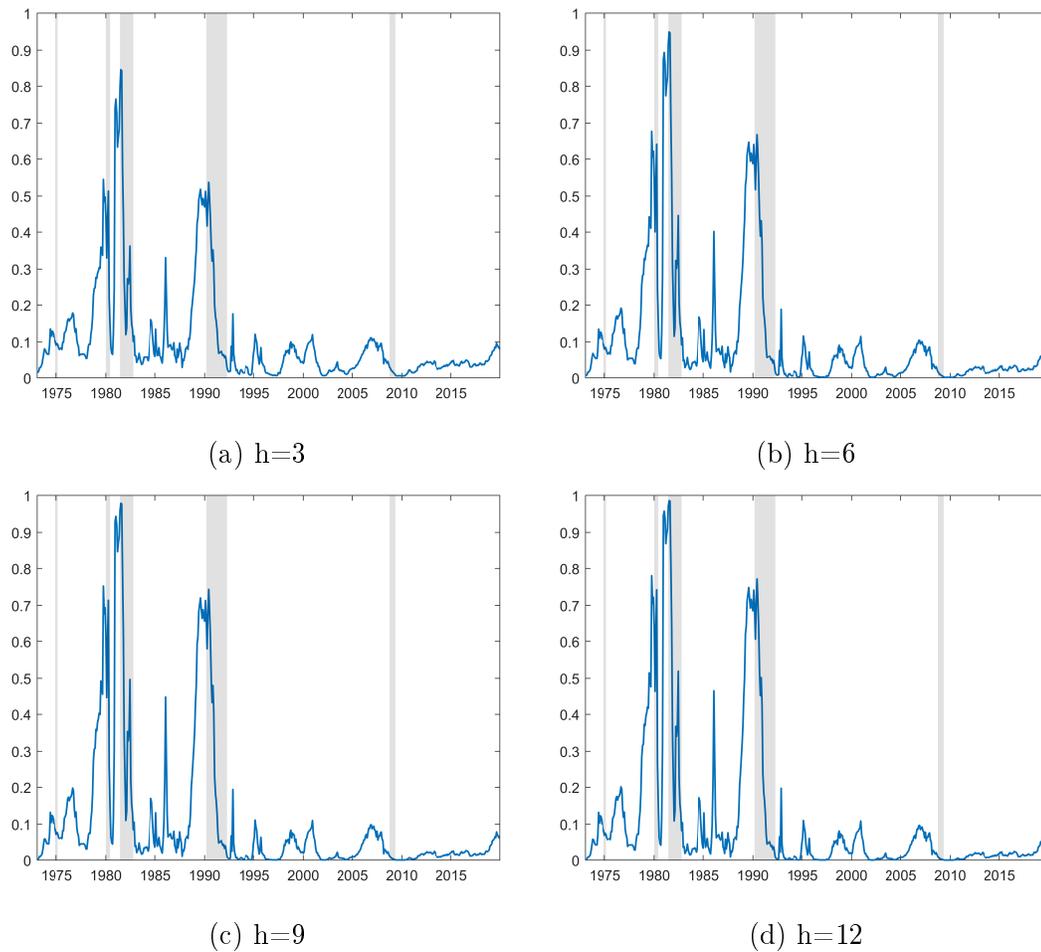


Figure 3.1: Probabilité de récession en échantillon avec l'écart entre le taux 10 ans et le taux 3 mois

tillon. Ici, la structure à terme des prévisions est présentée, au lieu des prévisions pour un horizon donné. Chaque observation représente la prévision de la probabilité de récession d'un mois donné, réalisée h mois à l'avance. Ceci est plus informatif puisque la forme (convexité) de cette structure à terme peut indiquer l'arrivée imminente d'un point de retournement comme proposé dans Kotchoni et Stevanovic (2020). Les bandes rouges représentent les récessions déterminées par l'Institut C.D. Howe. Encore une fois, le signal envoyé pour la récession de 1990

est plus fort que celui de la récession de 2008. Notons aussi la présence de certains faux signaux prévus par le modèle probit, lorsque la probabilité augmente sans qu'une récession soit déclarée. Par exemple, en 2001, la probabilité de récession a augmenté, mais l'Institut C.D. Howe n'a pas déclaré de récession comparative-ment au NBER qui a déclaré une récession aux États-Unis. Il est probable que l'économie canadienne ait subi un ralentissement, sans qu'il soit assez important pour être considéré comme une récession. Cette information peut tout de même avoir une valeur prédictive pour l'activité économique réelle.

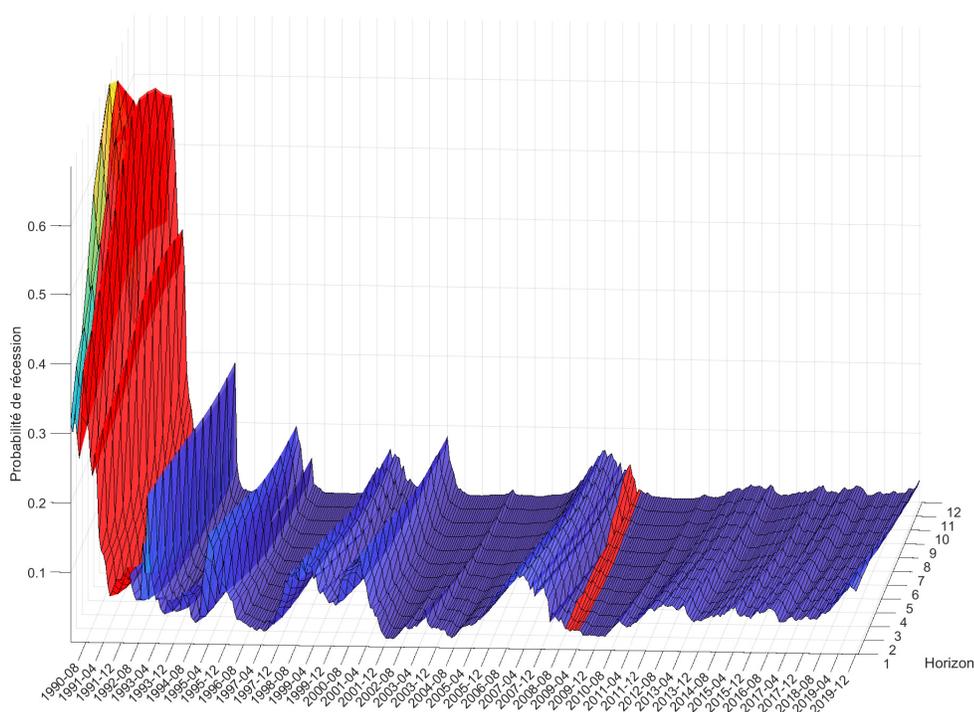


Figure 3.2: Probabilité de récession en pseudo hors échantillon avec l'écart entre le taux 10 ans et le taux 3 mois

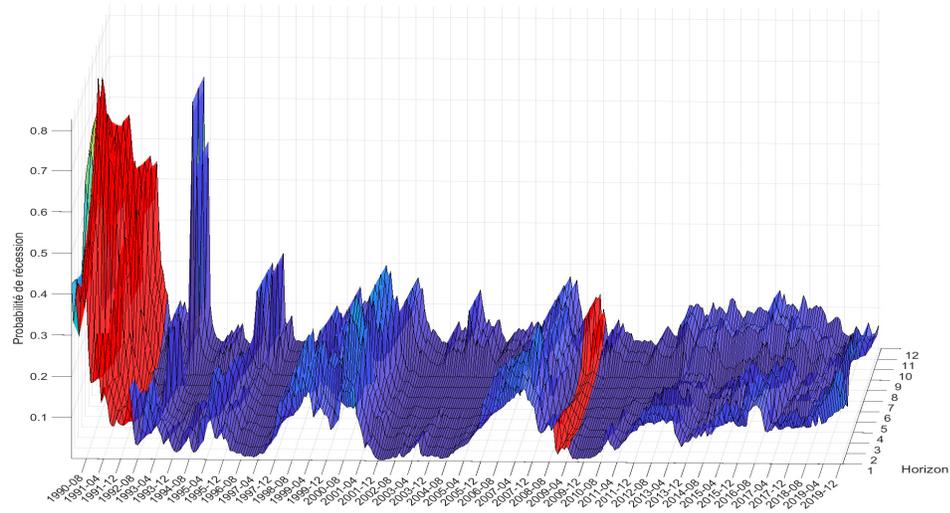
La figure 3.3 présente les probabilités de récession utilisées dans le modèle autorégressif augmenté pour l'emploi du Québec et du Canada. Comme illustré à

la section 3.1.2, pour chaque période de prévision et chaque horizon, la probabilité de récession avec chaque écart de taux est testée et celle qui minimise le critère d'information est choisie. On peut voir que ces probabilités envoient plus de faux signaux, mais le pouvoir prédictif pour chacune des variables devrait être supérieur. Les résultats de l'exercice en pseudo hors échantillon sur les variables d'activité économique seront présentés à la prochaine section. Les probabilités pour l'exercice trimestriel sur le PIB du Québec et du Canada sont montrées en appendice B et les écarts de taux sélectionnés en appendice C.

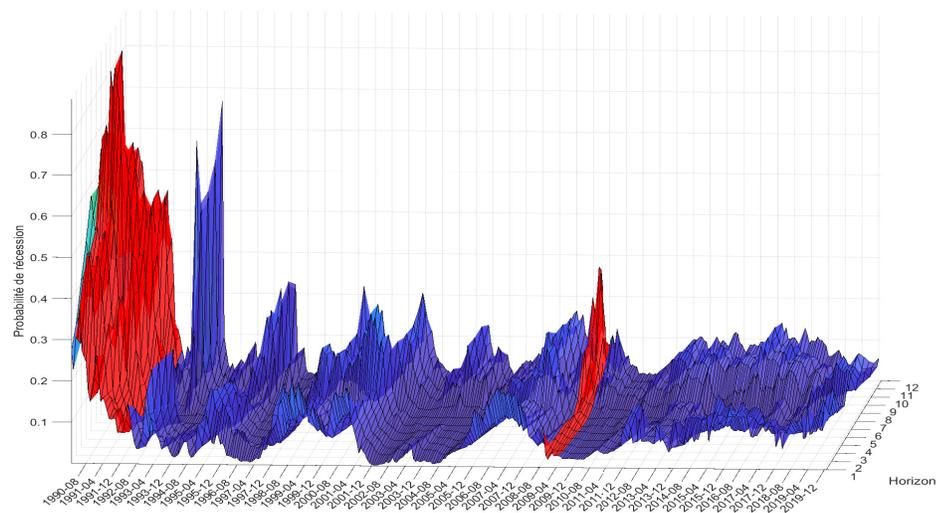
3.2.2 Modèles de prévision de l'activité économique

Le tableau 3.1 présente les ratios d'erreur quadratique moyenne des prévisions de l'emploi au Québec de 1 à 12 mois à l'avance. Une valeur supérieure à un signifie que l'erreur quadratique du modèle est supérieure à celle du modèle AR et vice-versa. Le suffixe « kNN » indique les modèles corrigés avec l'approche des k plus proches voisins en niveau, utilisant la fonction de distance (1.6). Le suffixe « kNN-dev » indique la correction utilisant la fonction de distance (1.7), en déviation par rapport à la moyenne. On peut voir que les modèles ADL et AAR produisent une erreur similaire à celle réalisée par le modèle AR pour l'ensemble des horizons de prévisions. Cette similarité peut s'expliquer par le fait que l'emploi met plus de temps à répondre aux périodes de ralentissement ou aux récessions. Ensuite, si on compare les modèles AR, ADL et AAR à leur version corrigée par l'approche des k plus proches voisins, on peut voir que la correction apporte peu de gain de performance pour la plupart des horizons et augmente l'erreur moyenne sur certains horizons. Les écarts de taux sélectionnés dans les modèles AAR et ADL sont présentés en appendice C. L'appendice D montre les paramètres m et k sélectionnés pour les modèles AR-kNN, AAR-kNN et ADL-kNN.

Le prochain tableau, 3.2, montre les ratios d'erreur quadratique pour les mêmes



(a) Avec les écarts de taux sélectionnés pour l'emploi québécois



(b) Avec les écarts de taux sélectionnés pour l'emploi canadien

Figure 3.3: Probabilités de récession mensuelles en pseudo hors échantillon pour le modèle AAR

modèles et les mêmes horizons pour l'emploi canadien. Les résultats sont similaires à ceux de l'emploi du Québec. Encore une fois, l'erreur moyenne des modèles ADL

Tableau 3.1: Ratio d'erreur quadratique moyenne des modèles de l'emploi québécois sur le modèle autorégressif

Modèle	Horizons (mois)											
	1	2	3	4	5	6	7	8	9	10	11	12
ADL	1.006	1.003**	1.002	1.002	1.006**	1.006	1.003	1.003	1.000	1.006	1.004	1.000
AAR	1.001	0.999	0.999	1.001	0.997	0.997	1.002	1.002	1.008*	1.000	0.999	1.006
AR-kNN	0.998	1.034**	1.054**	1.055	1.002	1.008	0.991	1.001	0.999	0.974	0.991	1.023*
ADL-kNN	1.004	1.036**	1.055**	1.063*	1.009	1.015	0.990	1.004	1.002	0.980	0.997	1.018*
AAR-kNN	0.999	1.026*	1.039*	1.059*	1.001	1.008	1.006	1.007	1.009	0.968	0.985	1.030*
AR-kNN-dev	1.000	0.998	1.054**	1.006	1.005	1.001	0.999	0.994	1.000	0.979	1.001	1.016
ADL-kNN-dev	1.006	1.019	1.056**	1.011	1.012	1.000	0.999	0.994	0.999	0.983	1.001	1.016
AAR-kNN-dev	1.000	0.997	1.048**	1.007	0.996	0.999	1.004	0.999	1.007	0.974	0.994	1.016

Note : Les valeurs en gras indique l'erreur quadratique moyenne la plus faible par horizon de prévision. L'absence de valeur en gras pour un horizon indique que le modèle AR a l'erreur la plus faible. Les étoiles représentent le niveau de significativité du test de Diebold-Mariano. Ainsi, *, ** et *** correspondent à des niveaux de significativité de 10%, 5% et 1%, respectivement.

et AAR est près de celle produite par le modèle AR pour tous les horizons de prévisions, la valeur des ratios s'approche de un. L'information contenue dans la structure à terme a donc un pouvoir prédictif faible sur l'emploi, peu importe si elle est intégrée de façon linéaire, dans les modèles ADL, ou de façon non linéaire, dans les modèles AAR.

Ensuite, le tableau 3.3 présente les ratios d'erreur quadratique pour le PIB québécois, en fréquence trimestrielle. Pour la prévision un trimestre à l'avance, le modèle AAR produit le meilleur résultat, mais le gain n'est pas significatif. Ici, la modélisation non linéaire de l'écart de taux est pertinente puisque le modèle AAR fait mieux que le modèle ADL. Ensuite, pour les horizons deux à trois, le modèle AAR est amélioré par la méthode des k plus proches voisins en niveau. L'erreur quadratique moyenne du modèle est inférieure aux modèles AR et ADL et les

Tableau 3.2: Ratio d'erreur quadratique moyenne des modèles de l'emploi canadien sur le modèle autorégressif

Modèle	Horizons (mois)											
	1	2	3	4	5	6	7	8	9	10	11	12
ADL	1.008	1.007	1.008	1.005	1.012	1.008*	1.006***	1.004***	1.005	1.010	1.005	0.992
AAR	1.004	0.995	1.001	0.998	0.996	0.996	0.997	0.999	1.001	0.998	0.999	1.003
AR-kNN	0.981	1.060***	1.030	1.031*	1.018***	1.015	1.019*	1.008	1.005	1.001	1.014*	1.017
ADL-kNN	0.991	1.046*	1.039	1.038*	1.032**	1.024*	1.025**	1.013	1.006	1.012	1.017**	1.009*
AAR-kNN	0.986	1.048**	1.031**	1.018	1.007	0.999	0.999	1.002	0.995	0.988	1.002	1.021
AR-kNN-dev	0.996	1.013***	1.024	1.041*	1.045**	1.017	1.009	1.003	1.035*	0.994	1.032**	1.037**
ADL-kNN-dev	1.004	1.026***	1.031	1.046*	1.052*	1.026*	1.014	1.012	1.042**	1.019*	1.033*	1.031
AAR-kNN-dev	0.999	1.019*	1.021	1.048	1.035*	1.004	0.993	1.001	1.031	1.006	1.008	1.035*

Note : Les valeurs en gras indique l'erreur quadratique moyenne la plus faible par horizon de prévision. L'absence de valeur en gras pour un horizon indique que le modèle AR a l'erreur la plus faible. Les étoiles représentent le niveau de significativité du test de Diebold-Mariano. Ainsi, *, ** et *** correspondent à des niveaux de significativité de 10%, 5% et 1%, respectivement.

deux méthodes kNN contribuent à réduire l'erreur de prévision, mais le gain est significatif seulement pour AAR-kNN. Par contre, pour les horizons plus longs, de quatre à huit, l'ajout des écarts de taux d'intérêt ne semble pas améliorer la prévision, peu importe le type de modèle. Aux horizons six et huit, la méthode kNN en niveau et en déviation par rapport à la moyenne améliorent marginalement la prévision.

Les ratios d'erreur quadratique moyenne en pseudo hors-échantillon du PIB canadien sont présentés dans le tableau 3.4. De façon similaire à l'exercice sur le PIB québécois, le modèle AAR produit une erreur plus faible que les modèles AR et ADL aux horizons un à trois. Encore une fois, ce résultat démontre la pertinence d'intégrer la prévision du modèle probit au lieu d'intégrer les écarts de taux de façon linéaire. La méthode des k plus proches voisins permet aussi d'améliorer la

Tableau 3.3: Ratio d'erreur quadratique moyenne des modèles du PIB québécois sur le modèle autorégressif

Modèle	Horizons (trimestre)							
	1	2	3	4	5	6	7	8
ADL	1.011*	0.992	0.986*	1.009	1.015	1.015	1.019**	1.031*
AAR	0.988	0.986	0.973	1.023	1.021**	1.021**	1.014**	1.008***
AR-kNN	1.009	0.992	1.005	1.008	1.010	0.993	1.019*	1.019
ADL-kNN	1.030	0.998	0.991	1.019	1.028	1.009	1.037***	1.054***
AAR-kNN	0.996	0.945*	0.949**	1.023	1.023***	1.017***	1.024	1.024
AR-kNN-dev	1.037	0.994	1.004	1.013	1.016	1.008	1.014***	0.995
ADL-kNN-dev	1.057	0.993	0.997	1.025	1.035	1.023**	1.037***	1.027*
AAR-kNN-dev	0.998	0.959	0.967	1.027	1.028***	1.012	1.031**	1.002

Note : Les valeurs en gras indique l'erreur quadratique moyenne la plus faible par horizon de prévision. L'absence de valeur en gras pour un horizon indique que le modèle AR a l'erreur la plus faible. Les étoiles représentent le niveau de significativité du test de Diebold-Mariano. Ainsi, *, ** et *** correspondent à des niveaux de significativité de 10%, 5% et 1%, respectivement.

prévision aux horizons un à trois. Pour les horizons un et trois, la méthode en niveau surpasse le modèle de référence alors que c'est la méthode en déviation par rapport à la moyenne qui fait le mieux pour l'horizon deux. Or, le gain de performance est seulement significatif pour la prévision réalisée un trimestre à l'avance. Comme pour le PIB québécois, le pouvoir prédictif des écarts de taux est faible pour la prévision quatre trimestres à l'avance et plus. En effet, les modèles ADL et AAR ont une erreur plus grande que le modèle AR sur ces horizons. Toutefois, la méthode kNN permet de réduire l'erreur des trois modèles à partir de l'horizon cinq. La méthode en niveau fait le mieux, et le modèle AR-kNN produit l'erreur la plus faible pour ces horizons. La différence de performance n'est toutefois pas significative.

Les résultats des exercices de prévision en pseudo hors échantillon démontrent que

Tableau 3.4: Ratio d'erreur quadratique moyenne des modèles du PIB canadien sur le modèle autorégressif

Modèle	Horizons (trimestre)							
	1	2	3	4	5	6	7	8
ADL	1.012	1.006	1.002	1.010	1.025*	1.032**	1.039**	1.040*
AAR	0.969	0.946	0.943	1.041	1.027	1.027***	1.006	1.006
AR-kNN	0.905	1.037**	1.012	1.016	0.954	0.950	0.959	0.969
ADL-kNN	1.012	1.036***	1.014	1.007	0.975	0.975	0.999	1.008
AAR-kNN	0.887*	0.955	0.913	1.044	0.957	0.973	0.967	0.978
AR-kNN-dev	1.004	1.025	1.031	1.010	0.962	0.984	1.015	1.052*
ADL-kNN-dev	1.019	1.030	1.029	1.018	0.982	1.014	1.025*	1.017
AAR-kNN-dev	0.993	0.931	0.915	1.077	0.976	1.000	0.997	0.979

Note : Les valeurs en gras indique l'erreur quadratique moyenne la plus faible par horizon de prévision. L'absence de valeur en gras pour un horizon indique que le modèle AR a l'erreur la plus faible. Les étoiles représentent le niveau de significativité du test de Diebold-Mariano. Ainsi, *, ** et *** correspondent à des niveaux de significativité de 10%, 5% et 1%, respectivement.

les modèles présentés semblent avoir une meilleure performance sur les variables en fréquence trimestrielle. Dans l'exercice présenté, l'information contenue dans la structure à terme possède un meilleur pouvoir prédictif sur les variations du produit intérieur brut que sur les variations de l'emploi au Québec et au Canada. Cette information semble avoir le plus de pertinence sur la prévision des horizons inférieurs à un an. De plus, incorporer les erreurs passées d'un modèle à l'aide de la méthode des k plus proches voisins semble améliorer la prévision pour plusieurs modèles et plusieurs horizons.

CONCLUSION

Le but de cette recherche est de déterminer quels modèles permettent d'améliorer les prévisions de l'activité économique en temps de crise. Les méthodes proposées sont le modèle autorégressif augmenté de la probabilité récession de Kotchoni et Stevanovic (2020) et la correction par les k plus proches voisins de Guerron-Quintana et Zhong (2017). Une des particularités de la recherche est de modéliser la probabilité de récession avec différents écarts de taux d'intérêt canadien. Cette approche permet de sélectionner l'information sur la courbe de rendement qui a le meilleur pouvoir prédictif sur les variables d'activité économique.

Deux variables, canadiennes et québécoises, sont sélectionnées pour décrire l'activité économique, soit l'emploi et le PIB réel. Dans le cas du taux de croissance de l'emploi, le modèle autorégressif augmenté produit la meilleure prévision sur certains horizons dans l'exercice en pseudo hors échantillon. Par contre, les gains de pouvoir prédictif sont mineurs par rapport au modèle autorégressif. Généralement, dans le cas du taux de croissance du PIB réel, les modèles proposés font mieux que pour l'emploi par rapport au modèle de référence. Pour certains horizons de moins d'un an, le modèle autorégressif augmenté incorporant la correction des k plus proches voisins améliore significativement la prévision du PIB. Pour le PIB canadien, les gains vont jusqu'à 11,3% en termes d'erreur quadratique moyenne par rapport au modèle autorégressif.

Alors que cet exercice a été réalisé sur les données historiques qui ont subi des révisions, il serait intéressant d'évaluer les prévisions dans un exercice en temps réel. Ceci pourrait avantager les méthodes AAR et kNN puisque les révisions sont

souvent majeures durant les récessions. Ces révisions ont pu surévaluer la performance des modèles standards dans notre étude, qui aurait eu plus de difficulté à capter des changements brusques dans les données. Par contre, les données en temps réel ne sont pas disponibles sur de longs historiques au Canada. Il faudra donc attendre que des séries temporelles plus longues soient rendues disponibles avant de pouvoir effectuer un tel exercice.

Il serait aussi intéressant de tester différentes méthodes d'agrégation des prévisions, avec objectif de réduire l'incertitude sur le choix de modèle et l'erreur de prévision (Moral-Benito, 2015). Également, tester différents modèles avec la méthode kNN permettrait possiblement d'obtenir des prévisions plus fiables. Une autre option pour améliorer les résultats serait de tester le pouvoir prédictif de différents modèles de prévision tel qu'un modèle probit dynamique comme celui développé par Hao et Ng (2011) pour l'économie canadienne, ou encore par un modèle logit lasso utilisé par Fortin-Gagnon *et al.* (2018).

APPENDICE A

DONNÉES FINANCIÈRES

A.1 Taux d'intérêt en niveau

Tableau A.1: Taux d'intérêt sélectionnés

Code	Nom
brate	Taux officiel d'escompte
bill2	Bons du Trésor : à 2 mois
bill3	Bons du Trésor : à 3 mois
bill6	Bons du Trésor : à 6 mois
bond1	Rendements moyens des obligations du gouv. canadien : de 1 à 3 ans
bond3	Rendements moyens des obligations du gouv. canadien : de 3 à 5 ans
bond5	Rendements moyens des obligations du gouv. canadien : de 5 à 10 ans
bond10	Rendements moyens des obligations du gouv. canadien : de plus de 10 ans

A.2 Écart de taux d'intérêt

Tableau A.2: Écarts de taux d'intérêt sélectionnés

Identifiant	Nom
1	bond1 - brate
2	bond1 - bill2
3	bond1 - bill3
4	bond1 - bill6
5	bond3 - brate
6	bond3 - bill2
7	bond3 - bill3
8	bond3 - bill6
9	bond3 - bond1
10	bond5 - brate
11	bond5 - bill2
12	bond5 - bill3
13	bond5 - bill6
14	bond5 - bond1
15	bond5 - bond3
16	bond10 - brate
17	bond10 - bill2
18	bond10 - bill3
19	bond10 - bill6
20	bond10 - bond1
21	bond10 - bond3
22	bond10 - bond5
23	Écart entre la composante principale de court terme et celle de long terme
24	Composante principale des écarts 1 à 28

APPENDICE B

RÉSULTATS DE L'EXERCICE EN PSEUDO HORS ÉCHANTILLON SUR LA PROBABILITÉ DE RÉCESSION

B.1 Probabilités de récession trimestrielles avec l'écart entre les taux 10 ans et les taux 3 mois

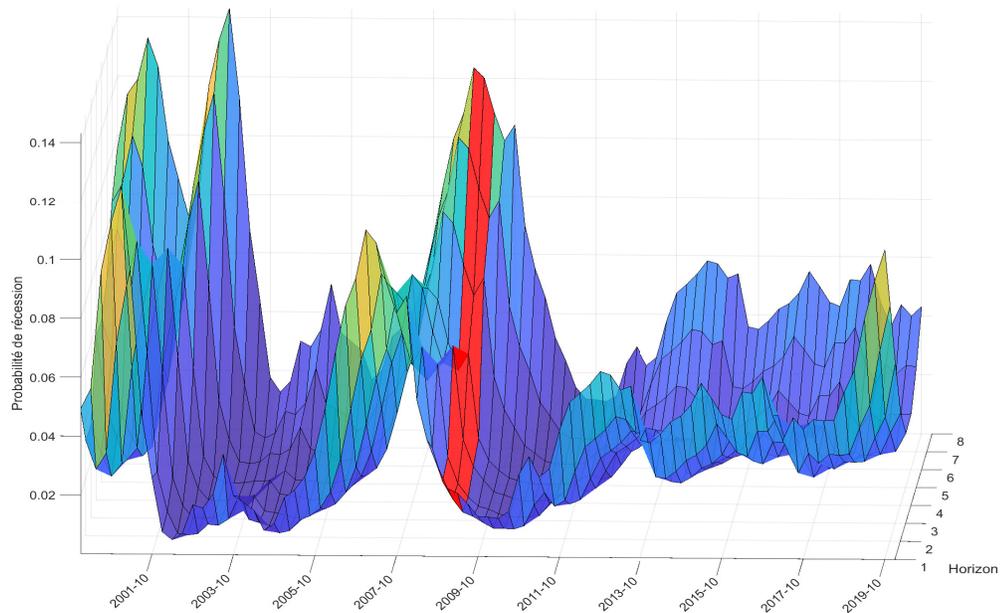
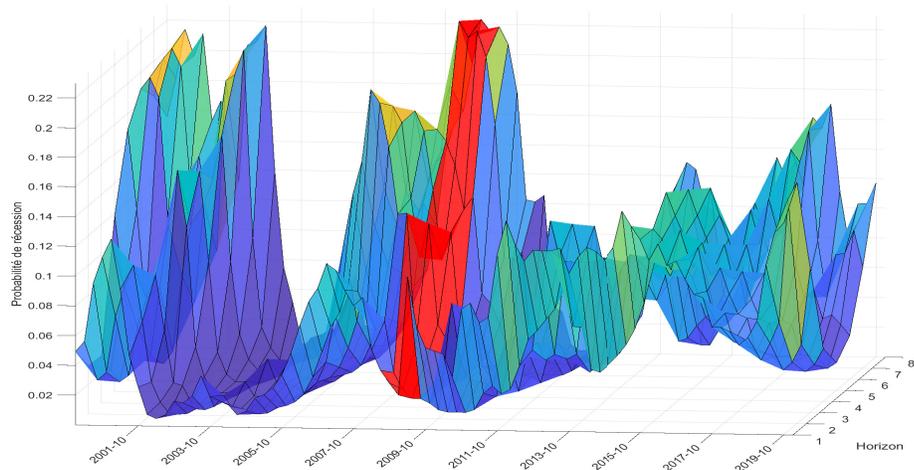
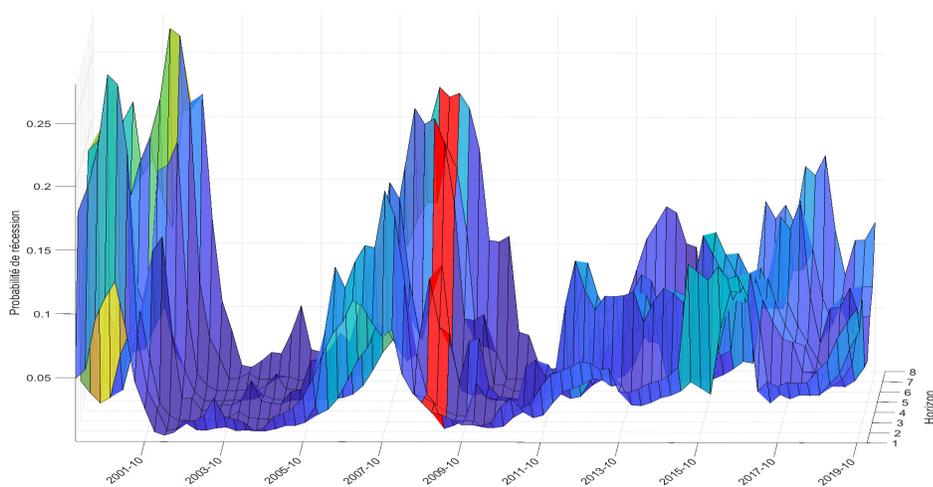


Figure B.1: Probabilités de récession en pseudo hors-échantillon avec l'écart entre les taux 10 ans et les taux 3 mois

B.2 Probabilités de récession trimestrielles avec les écarts de taux sélectionnés selon les variables prévues avec le modèle AAR



(a) Avec les écarts de taux sélectionnés pour le PIB du Québec



(b) Avec les écarts de taux sélectionnés pour le PIB du Canada

Figure B.2: Probabilités de récession trimestrielles en pseudo hors-échantillon pour le modèle AAR

APPENDICE C

ÉCARTS DE TAUX SÉLECTIONNÉS DANS LES MODÈLES AAR ET ADL

C.1 Écarts de taux sélectionnés dans les modèles AAR mensuels

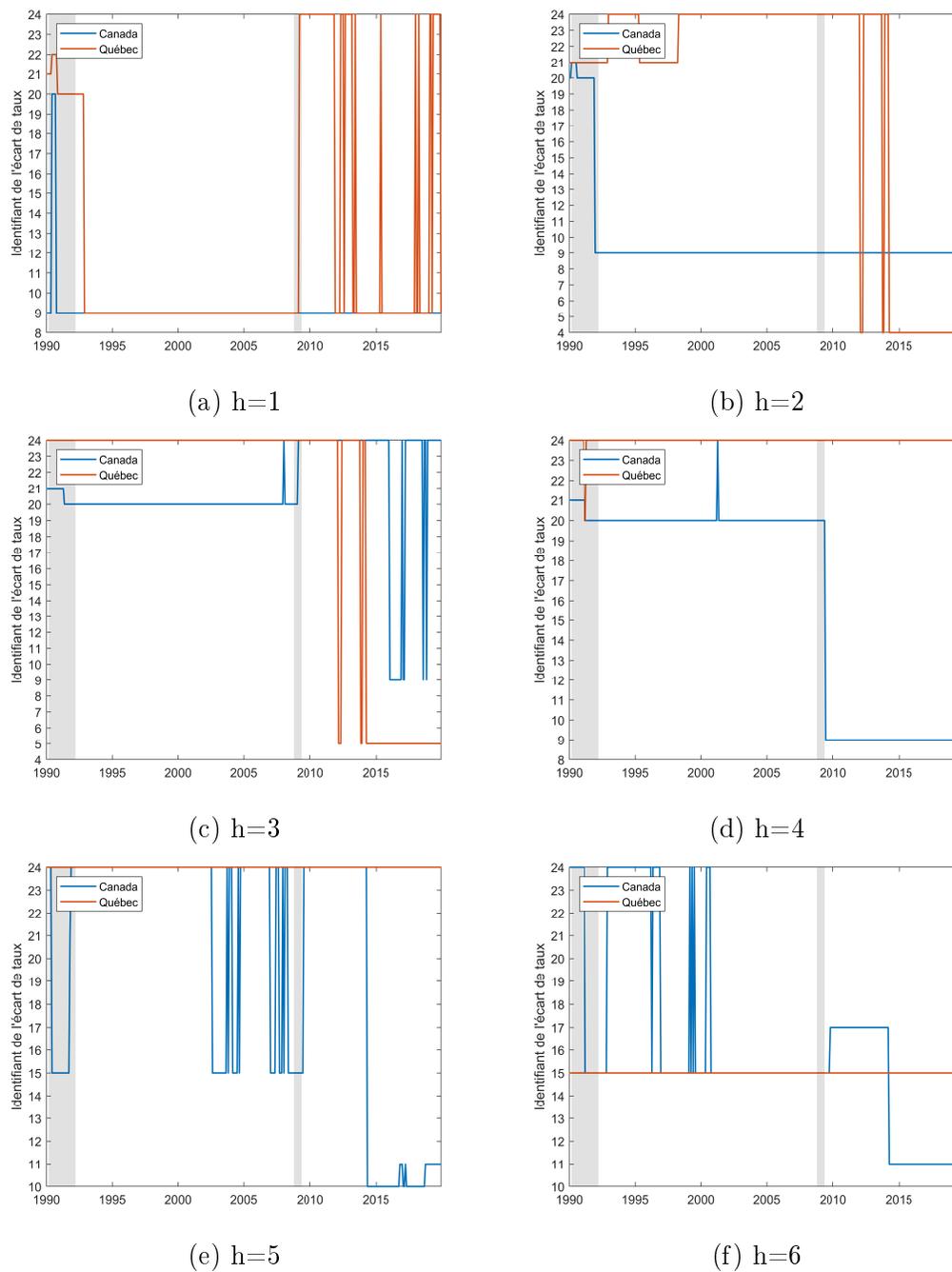


Figure C.1: Écarts de taux sélectionnés dans les modèles AAR mensuels pour les horizons 1 à 6

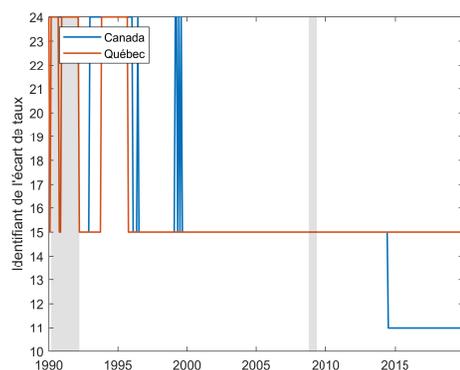
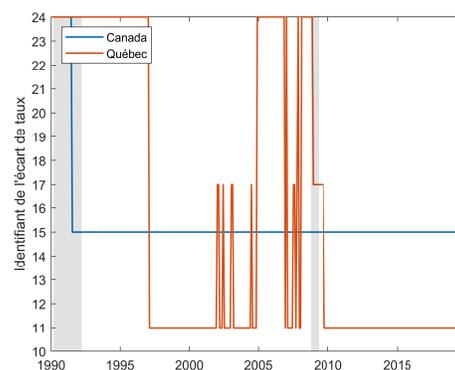
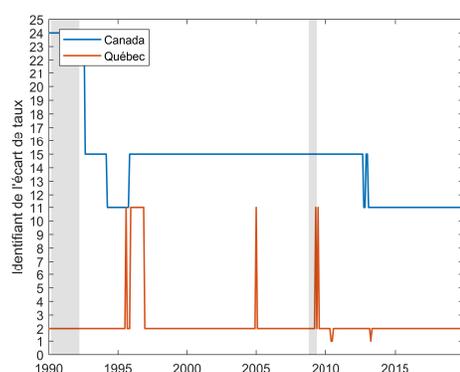
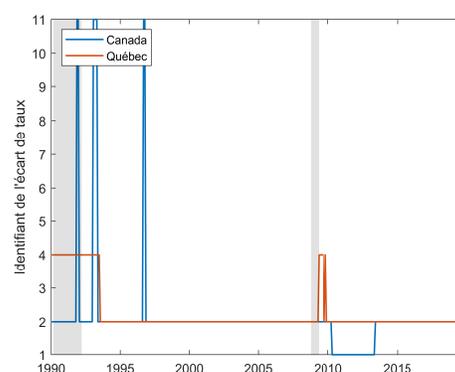
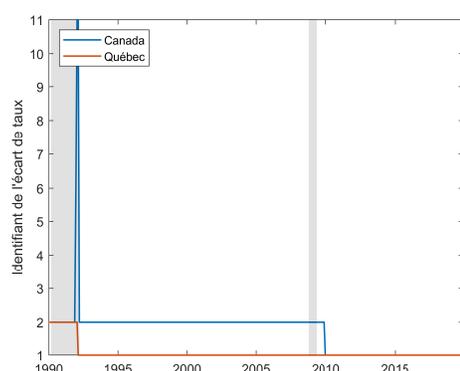
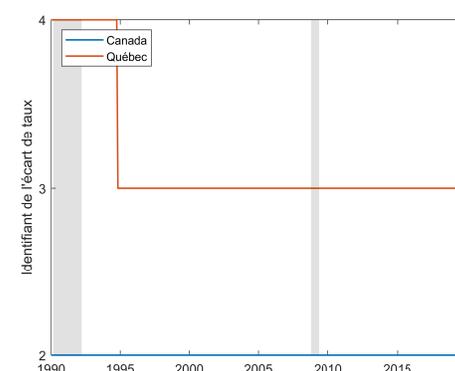
(a) $h=7$ (b) $h=8$ (c) $h=9$ (d) $h=10$ (e) $h=11$ (f) $h=12$

Figure C.2: Écarts de taux sélectionnés dans les modèles AAR mensuels pour les horizons 7 à 12

C.2 Écarts de taux sélectionnés dans les modèles AAR trimestriels

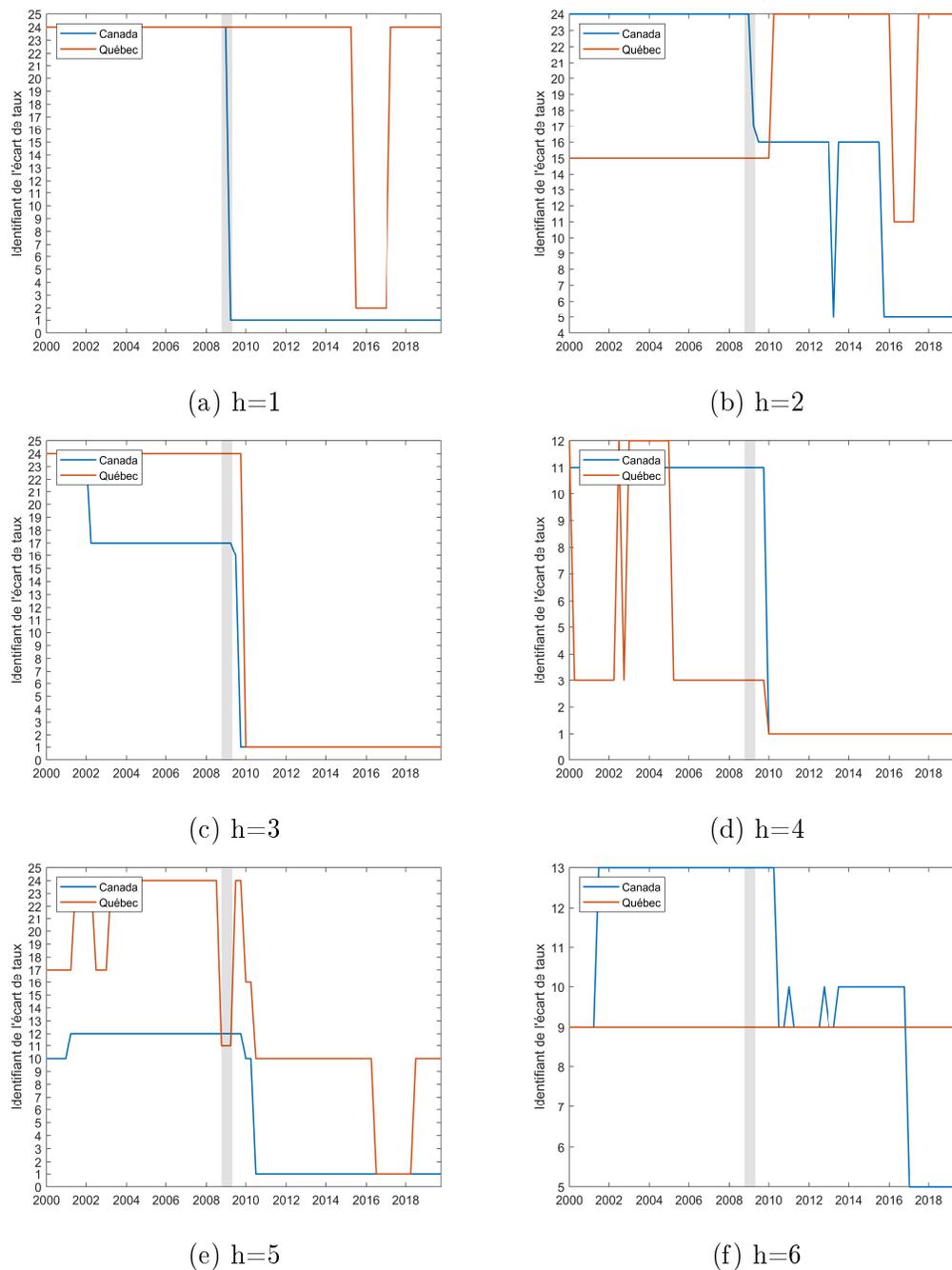


Figure C.3: Écarts de taux sélectionnés dans les modèles AAR trimestriels pour les horizons 1 à 6

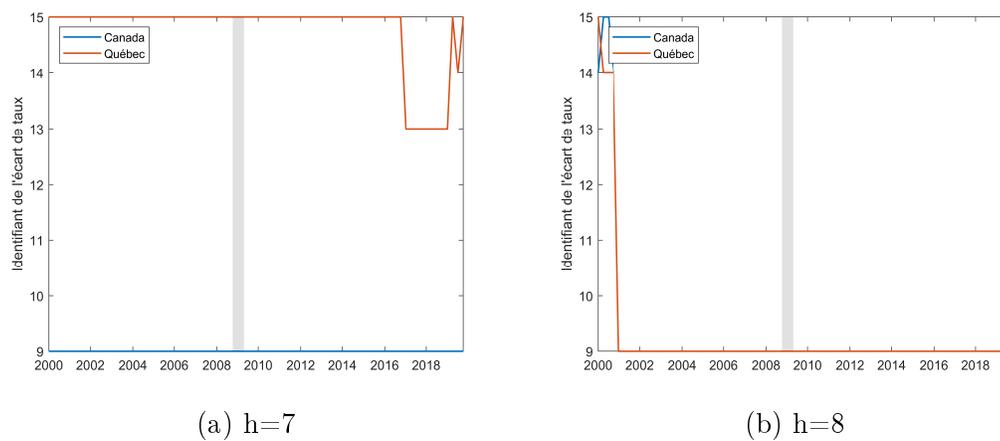


Figure C.4: Écarts de taux sélectionnés dans les modèles AAR trimestriels pour les horizons 7 à 8

C.3 Écarts de taux sélectionnés dans les modèles ADL mensuels

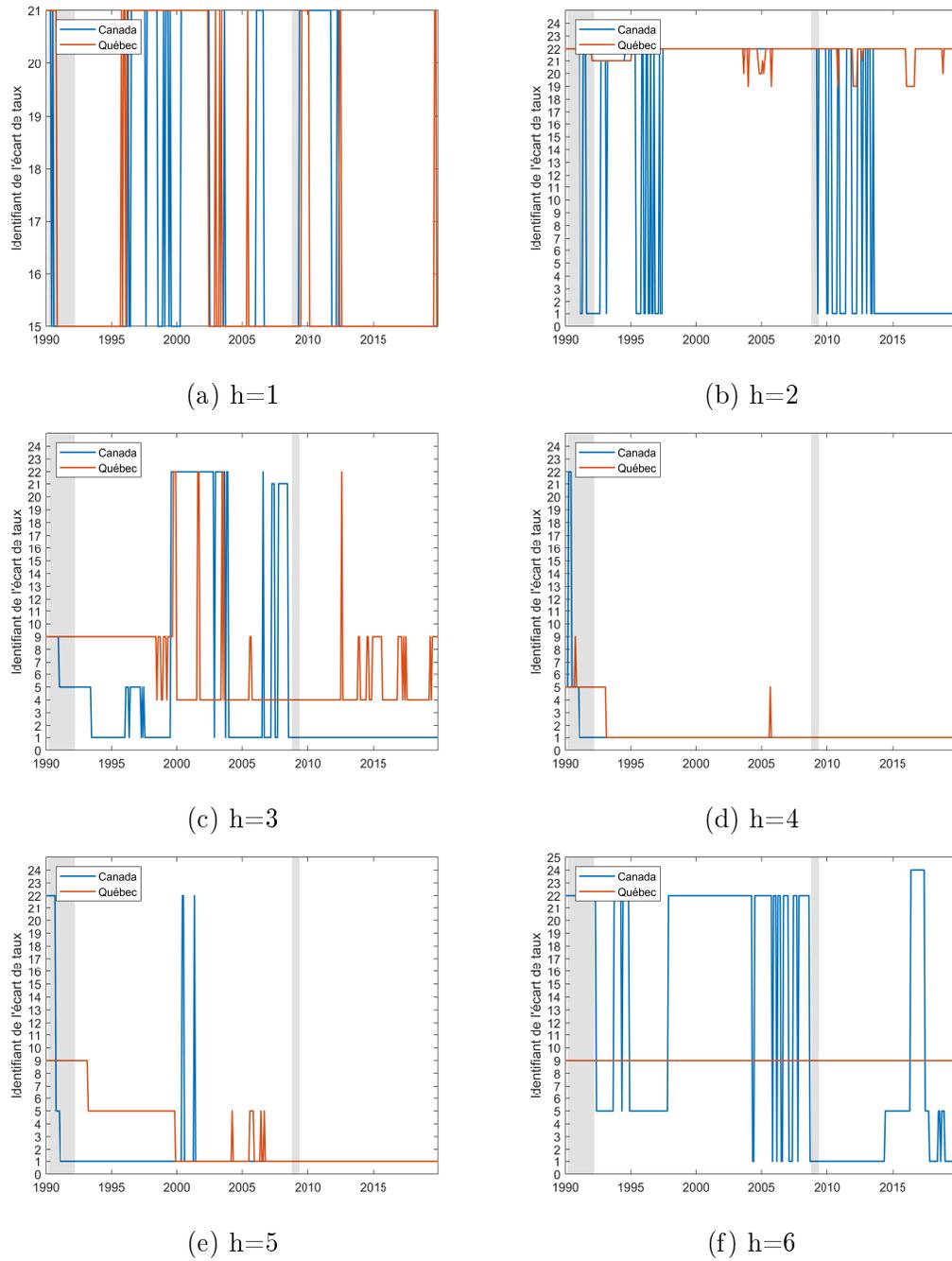


Figure C.5: Écarts de taux sélectionnés dans les modèles ADL mensuels pour les horizons 1 à 6

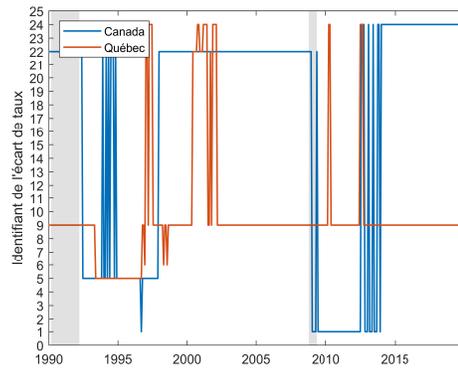
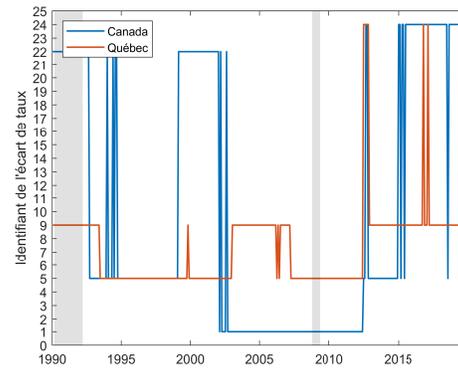
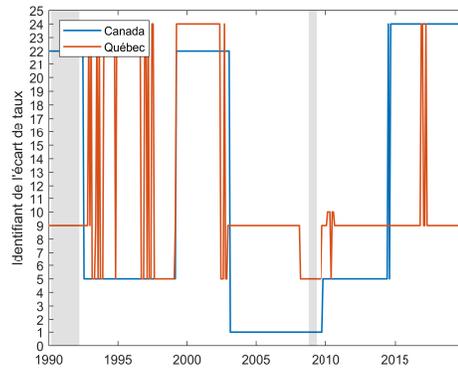
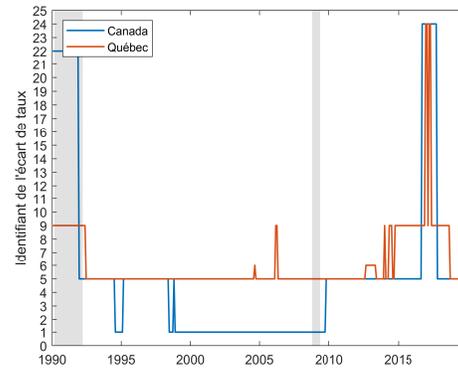
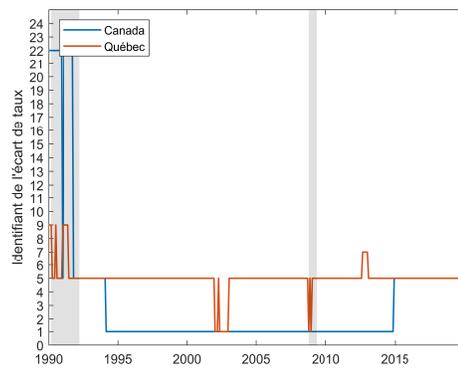
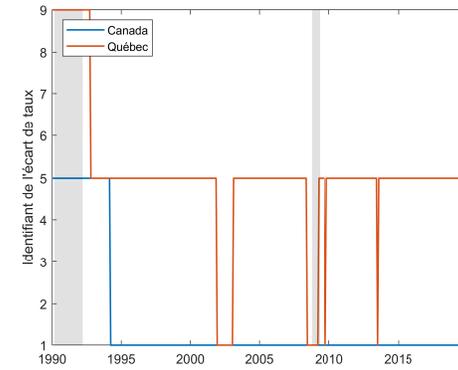
(a) $h=7$ (b) $h=8$ (c) $h=9$ (d) $h=10$ (e) $h=11$ (f) $h=12$

Figure C.6: Écarts de taux sélectionnés dans les modèles ADL mensuels pour les horizons 7 à 12

C.4 Écarts de taux sélectionnés dans les modèles ADL trimestriels

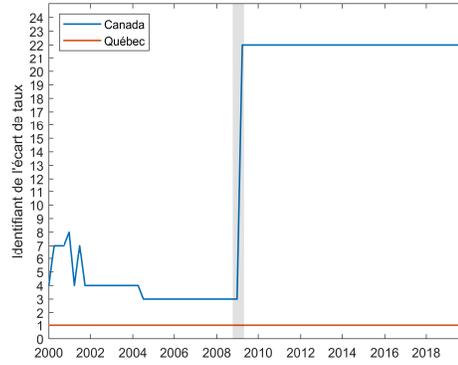
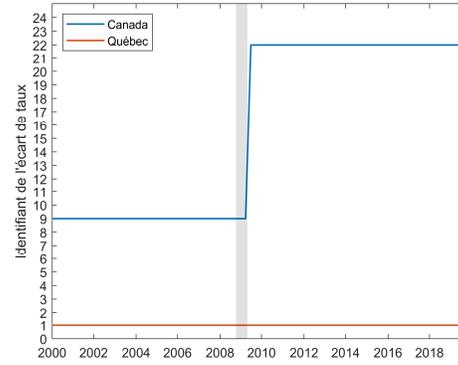
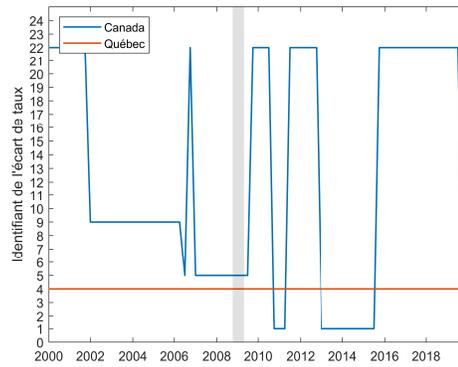
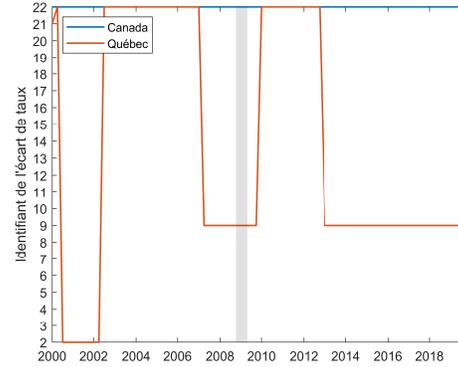
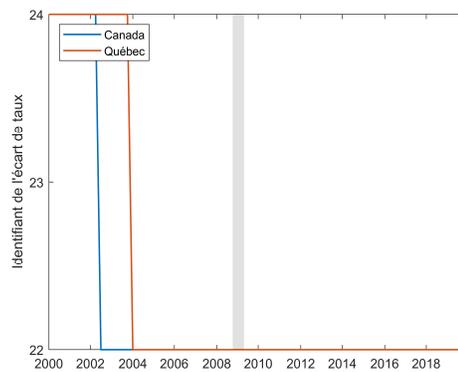
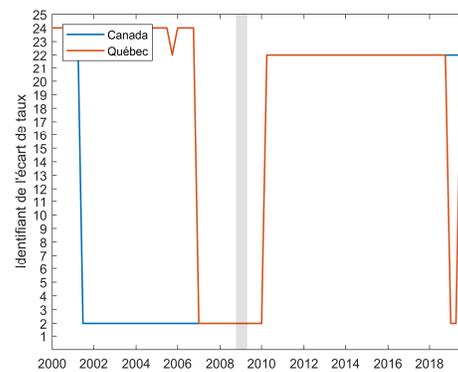
(a) $h=1$ (b) $h=2$ (c) $h=3$ (d) $h=4$ (e) $h=5$ (f) $h=6$

Figure C.7: Écarts de taux sélectionnés dans les modèles ADL trimestriels pour les horizons 1 à 6

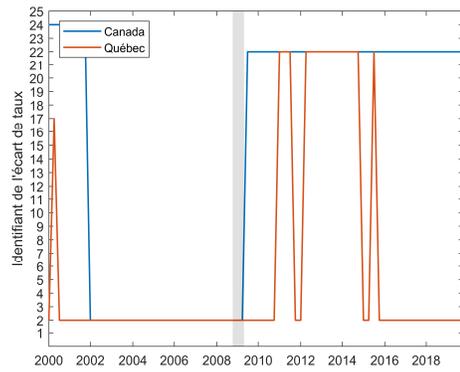
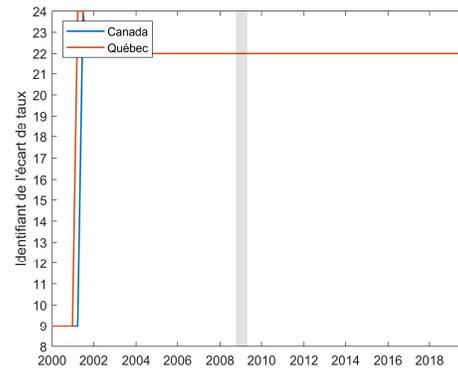
(a) $h=7$ (b) $h=8$

Figure C.8: Écarts de taux sélectionnés dans les modèles ADL trimestriels pour les horizons 7 à 8

APPENDICE D

RÉSULTATS DE L'EXERCICE DE PRÉVISION EN PSEUDO HORS ÉCHANTILLON

D.1 Paramètres m et k du modèle AR-kNN mensuel

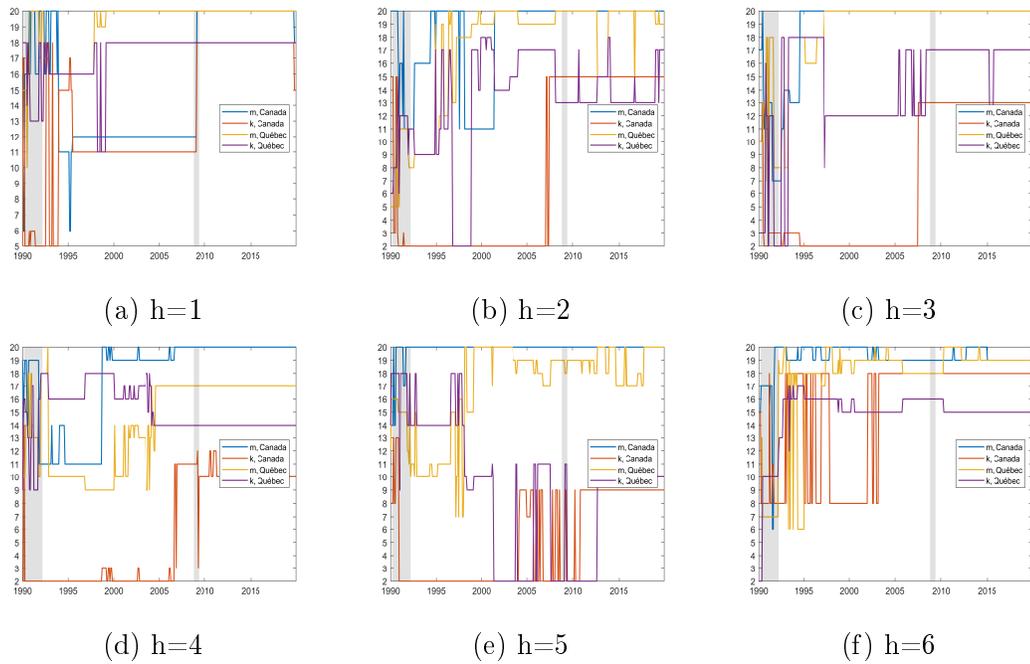


Figure D.1: Paramètres m et k du modèle AR-kNN mensuel pour les horizons 1 à

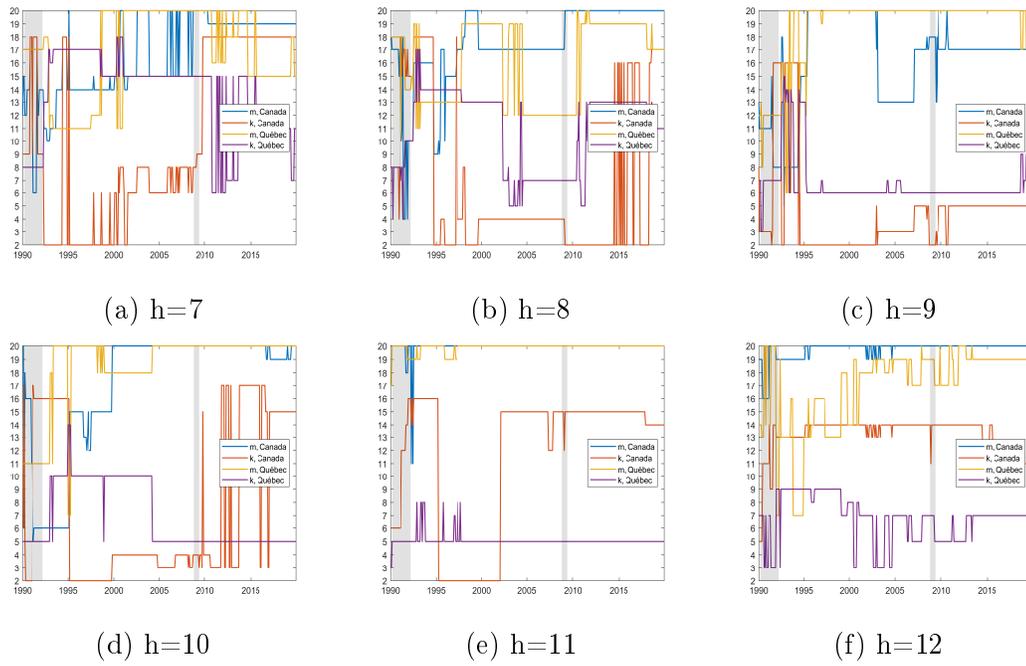


Figure D.2: Paramètres m et k du modèle AR-kNN mensuel pour les horizons 7 à 12

D.2 Paramètres m et k du modèle AAR-kNN mensuel

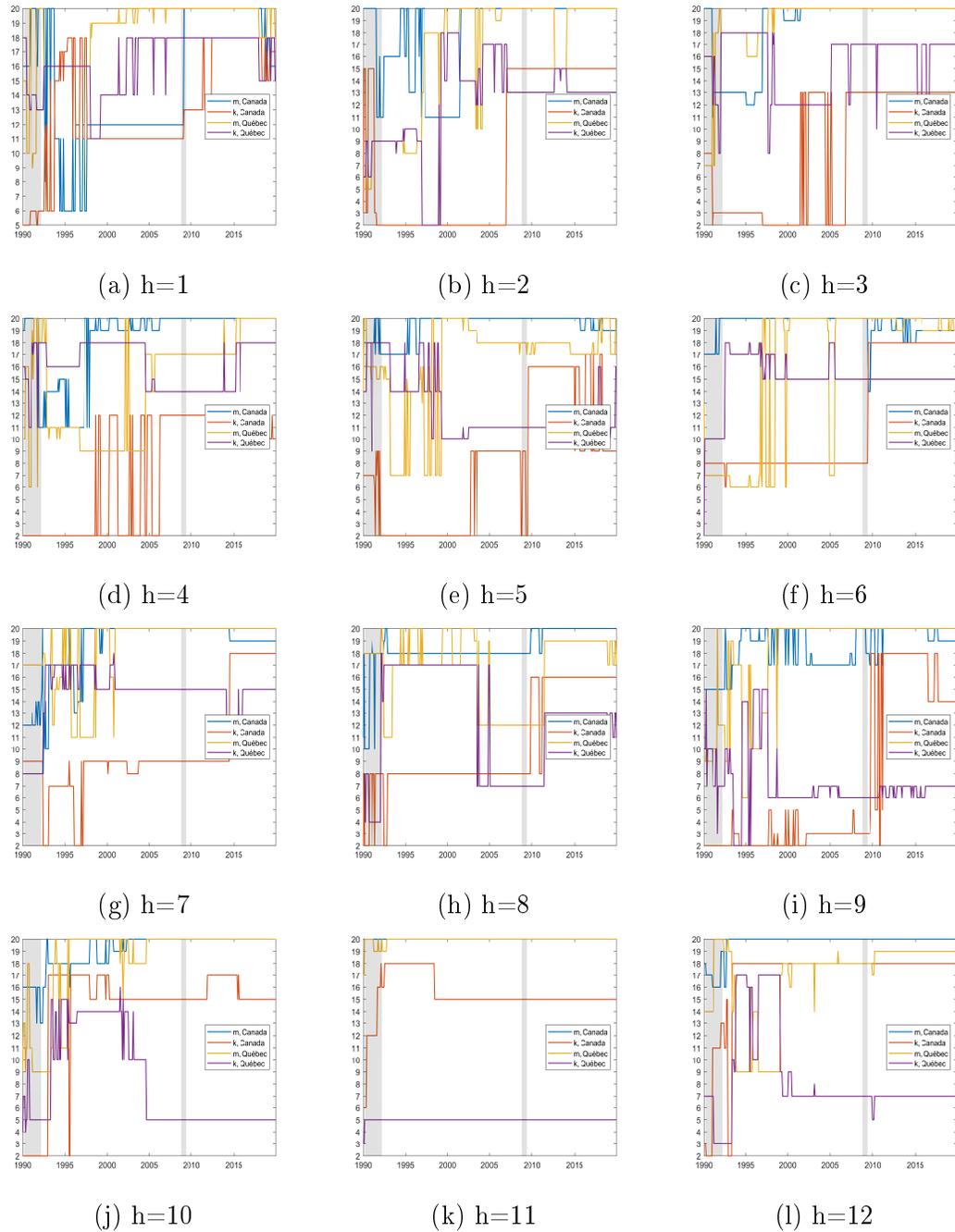


Figure D.3: Paramètres m et k du modèle AAR-kNN mensuel pour les horizons 1 à 12

D.3 Paramètres m et k du modèle ADL-kNN mensuel

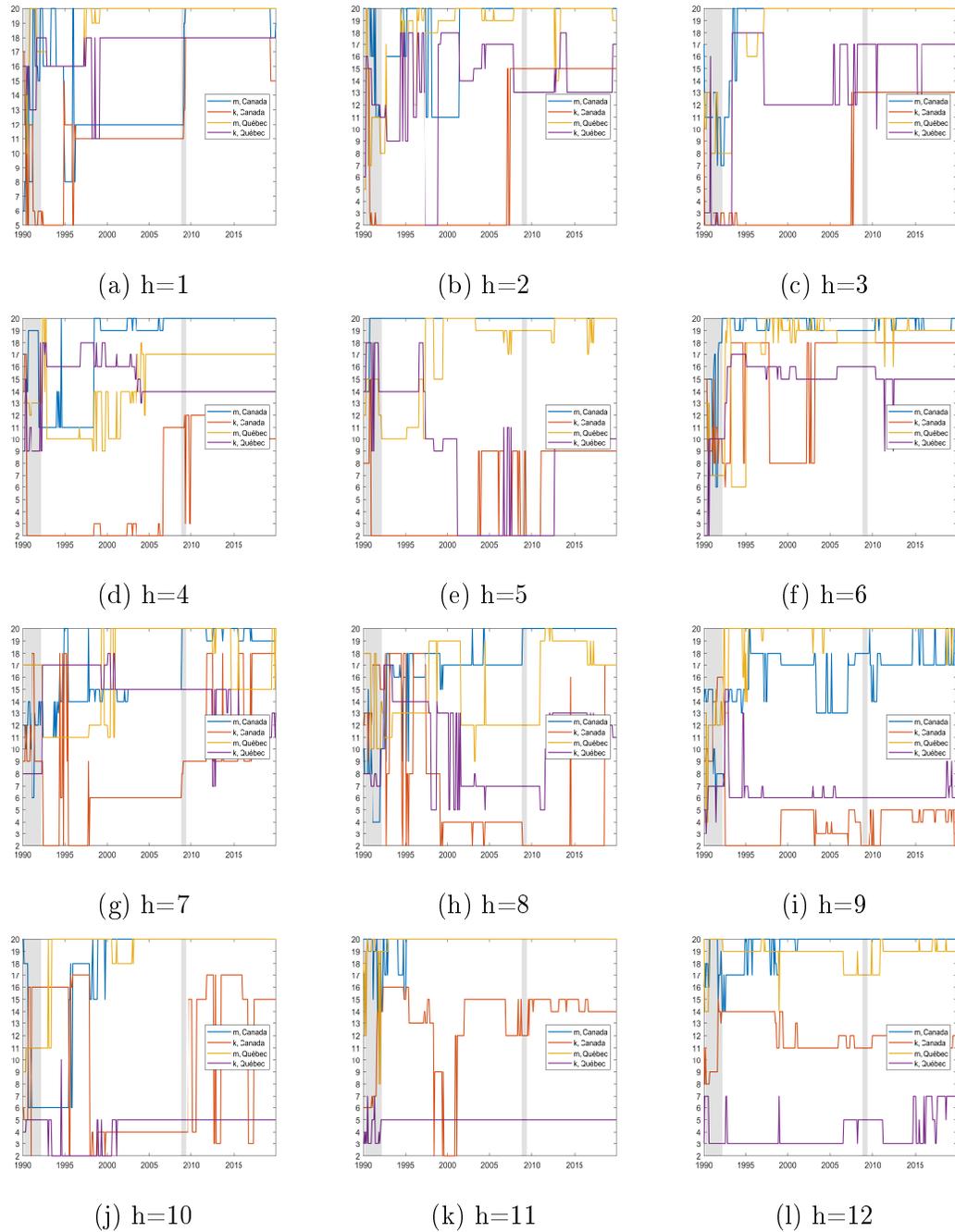


Figure D.4: Paramètres m et k du modèle ADL-kNN mensuel pour les horizons 1 à 12

D.4 Paramètres m et k du modèle AR-kNN trimestriel

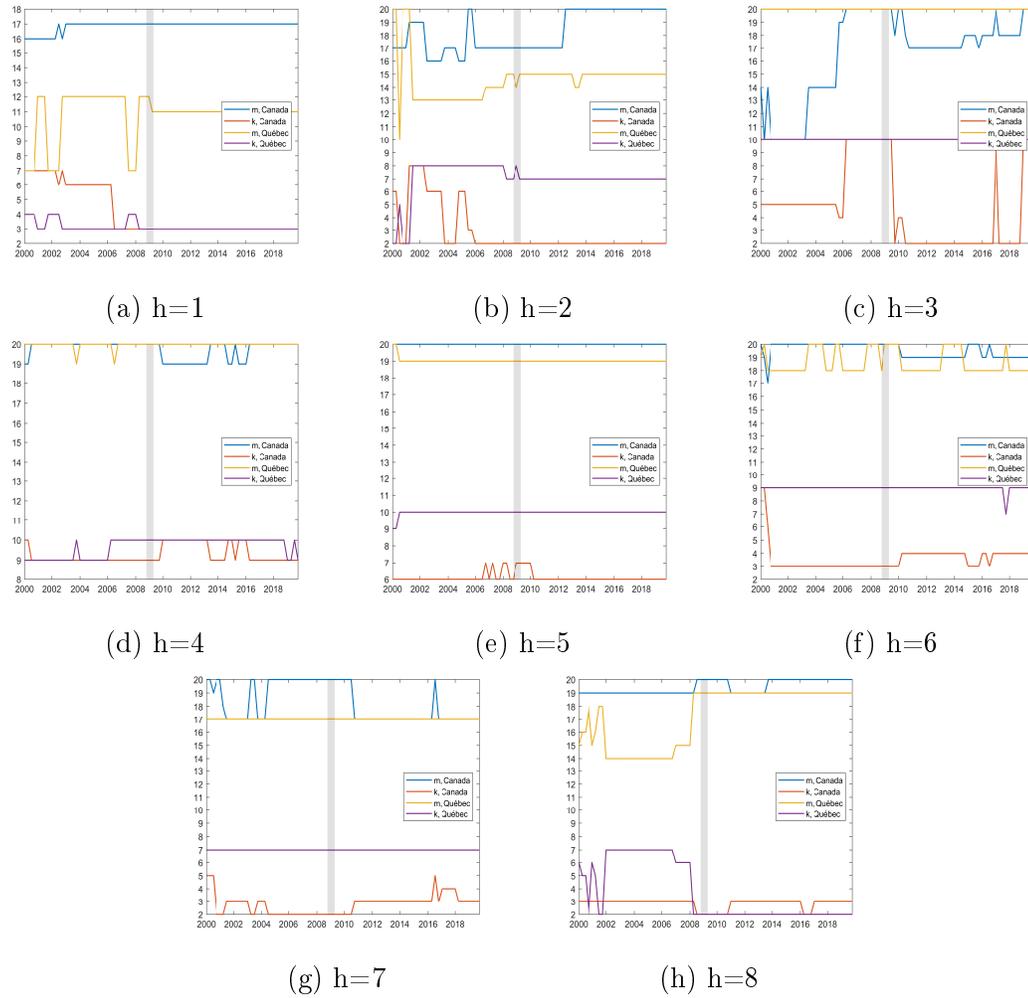


Figure D.5: Paramètres m et k du modèle AR-kNN trimestriel pour les horizons 1 à 8

D.5 Paramètres m et k du modèle AAR-kNN trimestriel

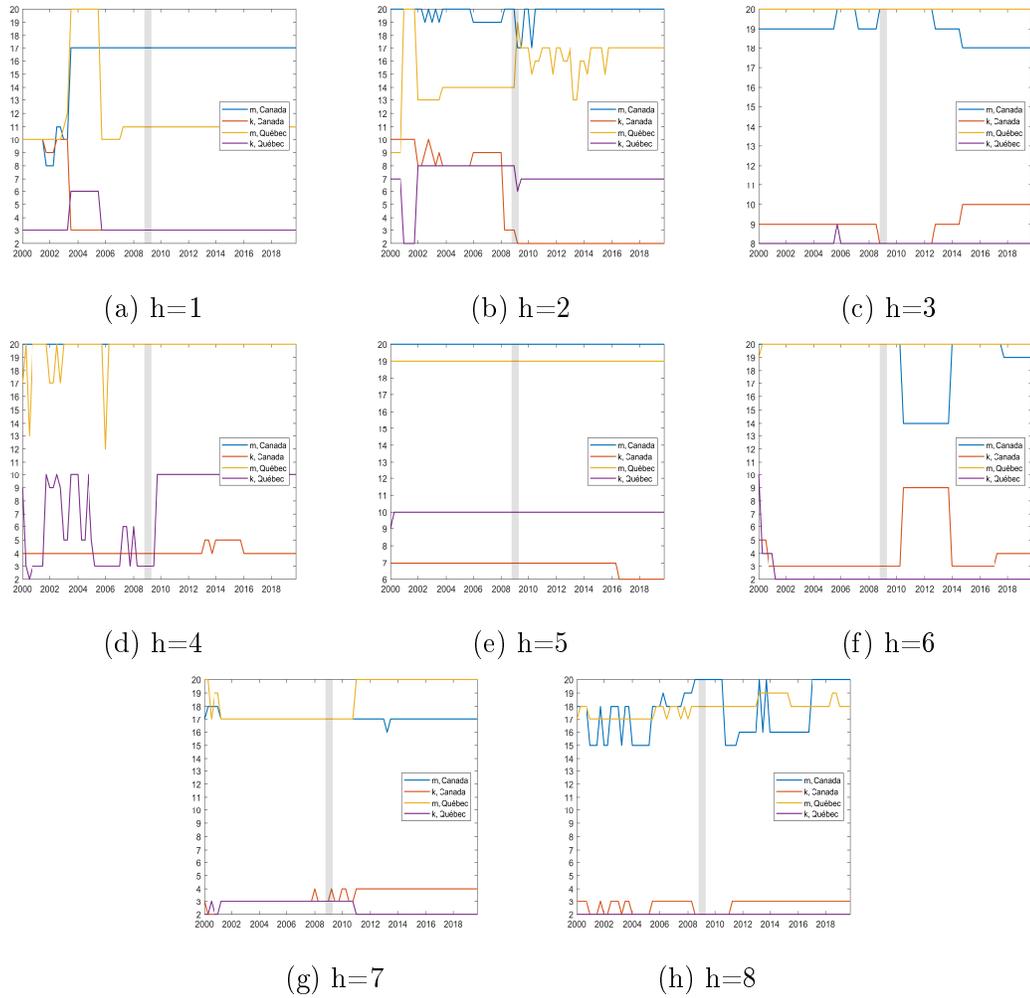


Figure D.6: Paramètres m et k du modèle AAR-kNN trimestriel pour les horizons 1 à 8

D.6 Paramètres m et k du modèle ADL-kNN trimestriel

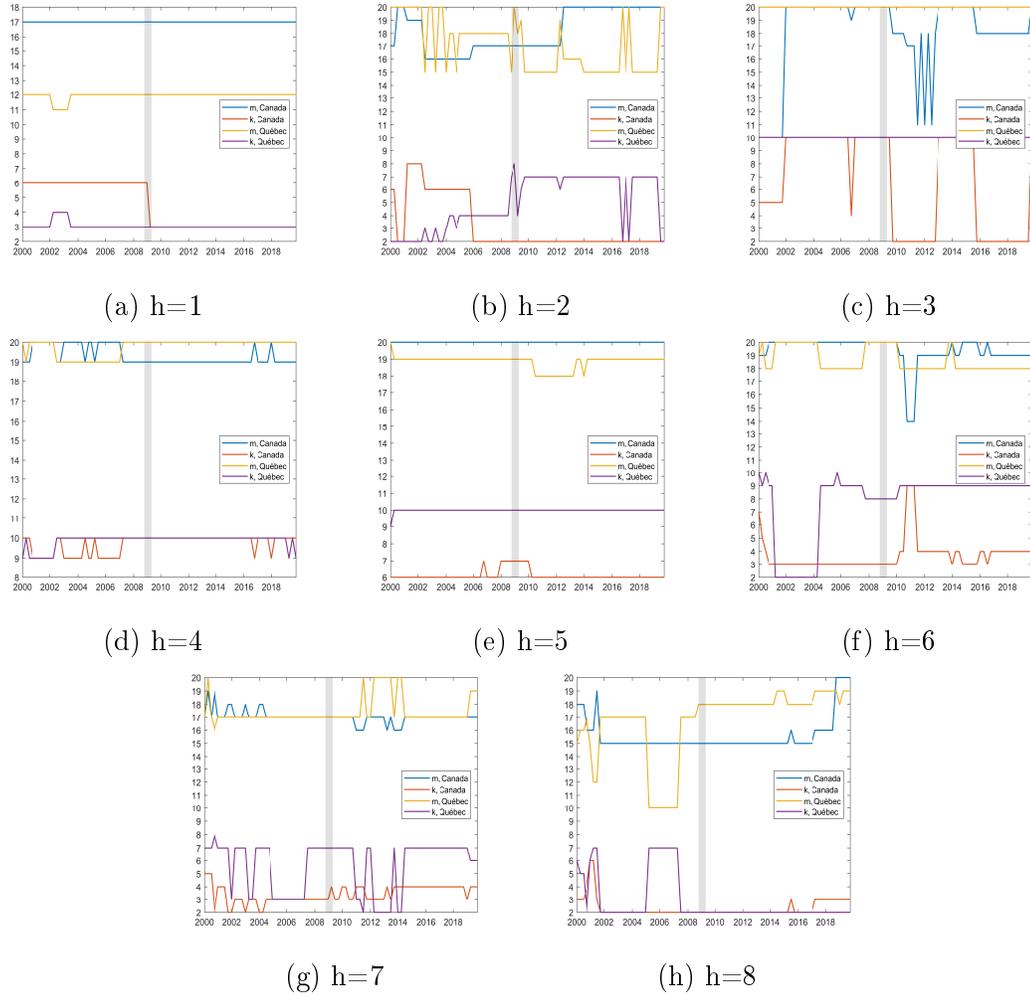


Figure D.7: Paramètres m et k du modèle ADL-kNN trimestriel pour les horizons 1 à 8

RÉFÉRENCES

- Bernard, H. et Gerlach, S. (1998). Does the term structure predict recessions? The international evidence. *International Journal of Finance & Economics*, 3(3), 195–215.
- Chauvet, M. et Potter, S. (2005). Forecasting recessions using the yield curve. *Journal of Forecasting*, 24, 77–103.
<http://dx.doi.org/10.2139/ssrn.274202>
- Chauvet, M., Potter, S. *et al.* (2013). Forecasting output. *Handbook of Economic Forecasting*, 2(Part A), 141–194.
- Clark, T. E. et McCracken, M. W. (2014). *Evaluating Conditional Forecasts from Vector Autoregressions*. Working papers (old series), Federal Reserve Bank of Cleveland.
- Clinton, K. (1995). The term structure of interest rates as a leading indicator of economic activity : A technical note. *Bank of Canada Review*, 1994(Winter), 23–40.
- Coulombe, P. G., Leroux, M., Stevanovic, D. et Surprenant, S. (à venir en 2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*.
- Coulombe, P. G., Marcellino, M. et Stevanović, D. (2021). Can machine learning catch the covid-19 recession? *National Institute Economic Review*, 256, 71–109.
- Cross, P. et Bergevin, P. (2012). Turning points : Business cycles in canada since 1926. *CD Howe Institute*, 366.
- Dickey, D. A. et Fuller, W. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49(4), 1057–1072.
- Diebold, F. X. et Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Diebold, F. X. et Nason, J. A. (1990). Nonparametric exchange rate prediction? *Journal of international Economics*, 28(3-4), 315–332.

- Dueker, M. (2005). Dynamic forecasts of qualitative variables. *Journal of Business & Economic Statistics*, 23(1), 96–104.
- Estrella, A. et Mishkin, F. S. (1998). Predicting U.S. Recessions : Financial Variables As Leading Indicators. *The Review of Economics and Statistics*, 80(1), 45–61.
- Farmer, J. D. et Sidorowich, J. J. (1987). Predicting chaotic time series. *Physical review letters*, 59(8), 845.
- Foroni, C., Marcellino, M. et Stevanovic, D. (2022). Forecasting the Covid-19 recession and recovery : Lessons from the financial crisis. *International Journal of Forecasting*, 38(2), 596–612.
<http://dx.doi.org/10.1016/j.ijforecast.2020>. Récupéré de
<https://ideas.repec.org/a/eee/intfor/v38y2022i2p596-612.html>
- Fortin-Gagnon, O., Leroux, M., Stevanovic, D. et Surprenant, S. (2018). *A large canadian database for macroeconomic analysis*. Rapport technique, Document de travail.
- Goodwin, T. H. (1993). Business-Cycle Analysis with a Markov-Switching Model. *Journal of Business & Economic Statistics*, 11(3), 331–339.
- Guerron-Quintana, P. et Zhong, M. (2017). *Macroeconomic forecasting in times of crises*. Finance and Economics Discussion Series 2017-018, Board of Governors of the Federal Reserve System (U.S.)
- Hao, L. et Ng, E. C. (2011). Predicting canadian recessions using dynamic probit modelling approaches. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(4), 1297–1330.
- Kotchoni, R., Leroux, M. et Stevanovic, D. (2019). Macroeconomic forecast accuracy in a data-rich environment. *Journal of Applied Econometrics*, 34(7), 1050–1072.
- Kotchoni, R. et Stevanovic, D. (2020). *GDP Forecast Accuracy During Recessions*. Rapport technique.
- Kundan, K. N. (2010). *The Superiority of Greenbook Forecasts and the Role of Recessions*. Rapport technique.
- McCracken, M. W. (2009). How accurate are forecasts in a recession ? *National Economic Trends*, (Feb).
- Moral-Benito, E. (2015). Model averaging in economics : An overview. *Journal of Economic Surveys*, 29(1), 46–75.

- Mulhern, F. J. et Caprara, R. J. (1994). A nearest neighbor model for forecasting market response. *International journal of forecasting*, 10(2), 191–207.
- Rudebusch, G. D. et Williams, J. C. (2009). Forecasting Recessions : The Puzzle of the Enduring Power of the Yield Curve. *Journal of Business & Economic Statistics*, 27(4), 492–503.
- Stock, J. H. et Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147–162.
- Taieb, S. B. et Hyndman, R. J. (2012). *Recursive and direct multi-step forecasting : the best of both worlds*. Rapport technique.
- Yakowitz, S. (1987). Nearest-neighbour methods for time series analysis. *Journal of Time Series Analysis*, 8(2), 235–247.