

Modèle prévisionnel touristique : la démarche d'élaboration

Report final

Ministère du Tourisme du Québec
et

Institut de la statistique du Québec

Mai 2008

Pour tout renseignement concernant l'Institut de la statistique du Québec et le ministère du Tourisme du Québec ainsi que les données statistiques qui y sont disponibles, s'adresser à :

| | |
|---|--|
| Ministère du Tourisme du Québec 900, boulevard René-Lévesque Est, bureau 400 Québec (Québec) G1R 2B5 Téléphone : (418) 643-9141 Sans frais : 1 800 482-2433 Ou Internet : www.bonjourquebec.com/mto/index.asp | Institut de la statistique du Québec 200, chemin Sainte-Foy Québec (Québec) G1R 5T4 Téléphone : (418) 691-2408 Sans frais : 1 800 463-4090 Ou Internet : www.stat.gouv.qc.ca |
|---|--|

© Ministère du tourisme

Dépôt légal :

Bibliothèque et Archives nationales du Québec, 2008

Bibliothèque et Archives Canada, 2008

ISBN : 978-2-550-53319-1

Toute reproduction est interdite sans l'autorisation expresse du ministère du Tourisme du Québec et de l'Institut de la statistique du Québec

Date : octobre 2008

Coordination et rédaction

Yves Lefrançois, Philippe DesRoches et Ghislain Dumas
 Agents de recherche
 Direction de la recherche et de la prospective
 Ministère du Tourisme du Québec

Danielle Bilodeau
 Agente de recherche
 Direction des statistiques économiques et sociales
 Institut de la statistique du Québec

Assistance technique

Denise Lemieux
 Technicienne
 Direction de la recherche et de la prospective
 Ministère du Tourisme du Québec

Sous la direction

Carmen Bellemare
 Directrice
 Direction de la recherche et de la prospective
 Ministère du Tourisme du Québec

Camille Courchesne
 Directeur
 Direction générale adjointe aux statistiques et à l'analyse
 Institut de la statistique du Québec

Révision linguistique

C. P. R. L.

Dans cette publication, la forme masculine désigne, lorsque le contexte s'y prête, aussi bien les femmes que les hommes.

Pour tout renseignement concernant le contenu de ce cahier, prière de contacter :

| | |
|---|---|
| Danielle Bilodeau Agente de recherche Institut de la statistique du Québec Direction des statistiques économiques et sociales 200, chemin Sainte-Foy, 3 ^e étage Québec (Québec) G1R 5T4 Téléphone : (418) 691-2411, poste 3191 Télécopieur : (418) 643-4129 Courriel : danielle.bilodeau@stat.gouv.qc.ca | Yves Lefrançois Agent de recherche Direction de la recherche et de la prospective Ministère du Tourisme 900, boulevard René-Lévesque Est, bureau 400 Québec (Québec) G1R 2B5 Téléphone : (418) 643-2658 Sans frais : 1 800 482-2433 Télécopieur : (418) 644-7514 Yves.lefrancois@tourisme.gouv.qc.ca |
|---|---|

Abréviations

AIC : Akaike information criterion
 DRP : Direction de la recherche et de la prospective
 EVC : Enquête sur les voyages des Canadiens
 EVI : Enquête sur les voyages internationaux
 ICRT : Institut canadien de recherche sur le tourisme
 ISQ : Institut de la statistique du Québec
 MAPE : Mean Average Percentage Error
 MTO : Ministère du Tourisme
 RMSE : Root Mean Square Error
 SAS : Statistical Analysis System
 TSFS : Time Series Forecasting System

Signes conventionnels

... N'ayant pas lieu de figurer
 e Donnée estimée
 n Nombre
 k En milliers
 M En millions
 G En milliards
 \$ En dollars

Avant-propos

La modélisation économétrique est un puissant outil d'analyse et son application à l'activité touristique du Québec constitue un pas vers une meilleure compréhension de toute la dynamique du comportement des agents économiques, à travers les données observées, qui constituent ce marché.

Cette étude a l'avantage d'apporter une aide appréciable aux agents de développement économique, afin d'établir des stratégies et des politiques spécifiques en matière d'intervention pour le développement économique du tourisme au Québec, grâce à l'anticipation et à la compréhension des phénomènes relatifs au secteur touristique québécois.

Cette étude fait suite à une collaboration de recherche entre le ministère du Tourisme du Québec (MTO) et l'Institut de la statistique du Québec (ISQ). Elle relate les différentes approches explorées et fait état des résultats obtenus avec les approches retenues. Le modèle prévisionnel touristique qui en résulte établit ainsi une base qui pourra être développée par la suite de façon à bien cerner les différentes problématiques de ce secteur.

Ce rapport s'adresse à tous ceux qui s'intéressent à l'activité touristique, plus particulièrement à celle du Québec, ainsi qu'à la modélisation économétrique. Toutefois, il est destiné davantage à un lecteur initié aux analyses de séries chronologiques étant donné le niveau technique de son contenu.

Le directeur général adjoint aux statistiques et à l'analyse.

Camille Courchesne

Avant-propos

Le ministère du Tourisme est fier de présenter à ses partenaires de l'industrie touristique un outil statistique permettant d'examiner l'évolution prochaine de l'activité touristique au Québec.

Des modèles prévisionnels ont été élaborés pour fournir aux acteurs de l'industrie des prévisions de l'évolution du nombre de touristes provenant des principaux marchés et des dépenses qu'ils effectueront au Québec. La démarche a été entreprise en 2005 pour combler les besoins d'information relativement aux statistiques permettant le suivi de la Politique touristique du Québec, *Vers un tourisme durable*.

L'exercice de modélisation, engagé par la Direction de la recherche et de la prospective, est le résultat d'une fructueuse collaboration avec l'Institut de la Statistique du Québec (ISQ) qui en a assuré la supervision méthodologique. Cette collaboration se poursuivra au cours de la prochaine année. Elle devrait permettre d'améliorer le modèle en approfondissant tout particulièrement l'analyse des principaux déterminants de la demande touristique.

Le lecteur constatera rapidement que le document s'adresse davantage aux personnes familiarisées avec l'analyse des séries chronologiques et les méthodes de prévision. Nous les invitons à nous acheminer tout commentaire susceptible d'améliorer l'outil et la justesse des raisonnements présentés.

La Directrice de la recherche et de la prospective

Carmen Bellemare

Table des matières

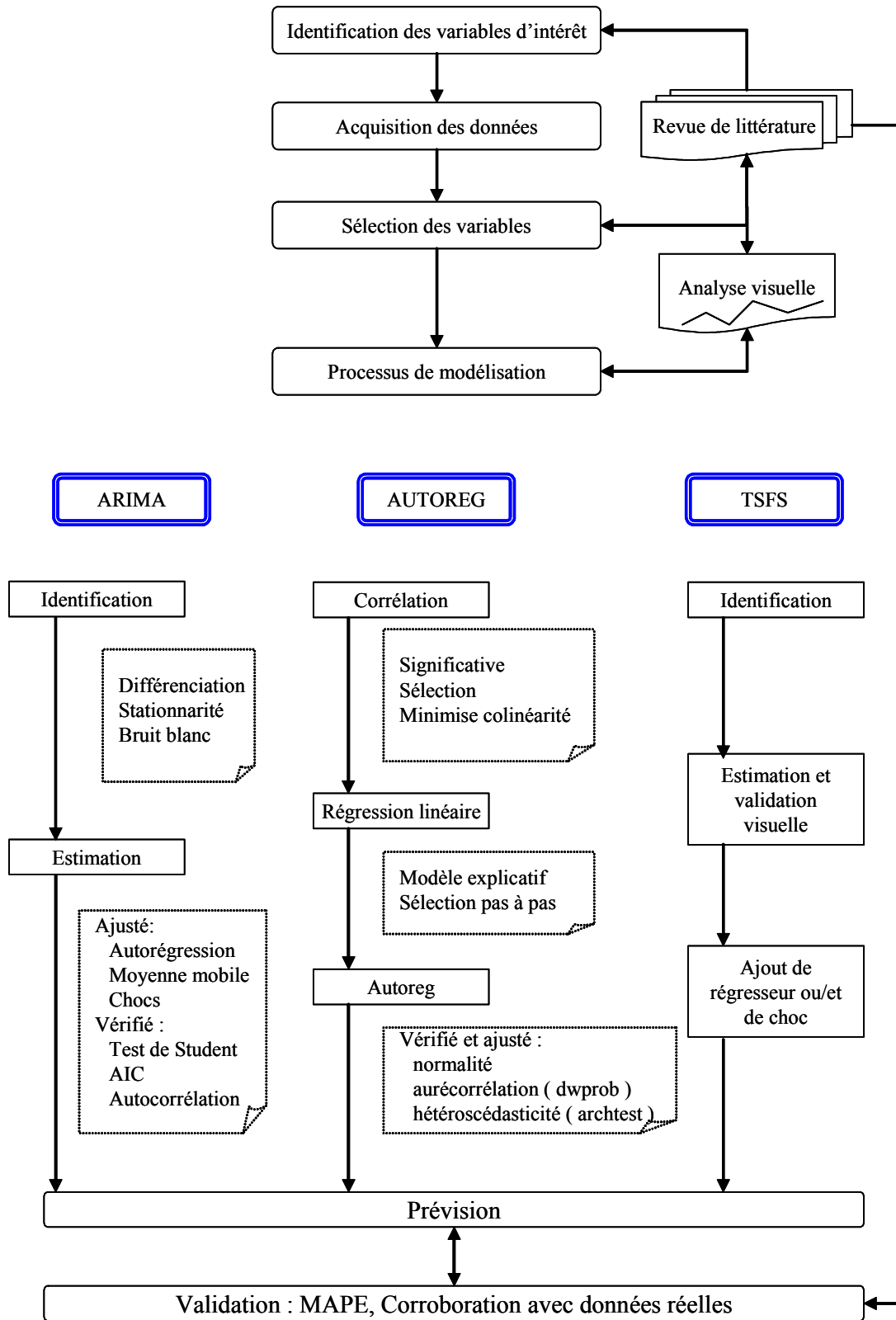
| | |
|---|----|
| Vue d'ensemble du processus de modélisation..... | 3 |
| Introduction..... | 4 |
| 1. Rappels méthodologiques..... | 5 |
| 1.1 Revue de littérature | 5 |
| 1.2 Méthodes | 5 |
| 1.2.1 Extrapolatives | 5 |
| Naïves..... | 6 |
| Décomposition classique..... | 6 |
| Lissage | 6 |
| Approche Box-Jenkins | 7 |
| 1.2.2 Causales..... | 8 |
| Modèle à équations uniques..... | 8 |
| Modèle à équations multiples | 9 |
| 1.3 Variables..... | 9 |
| 1.3.1 Dépendantes..... | 11 |
| 1.3.2 Indépendantes..... | 11 |
| Croissance économique et revenus..... | 11 |
| Prix de voyage au Québec et au Canada..... | 12 |
| Intervention marketing..... | 13 |
| Dichotomiques..... | 14 |
| Tendance..... | 14 |
| Autres types de variables | 15 |
| 1.4 Logiciels | 17 |
| 2. Modélisation | 18 |
| 2.1 Modèle ARIMA | 20 |
| 2.1.1 Étapes..... | 20 |
| A. Identification..... | 20 |
| B. Estimation..... | 21 |
| C. Prévision | 22 |
| 2.1.2 Option additionnelle | 23 |
| 2.1.3 Modèles retenus..... | 23 |
| 2.2 Modèles à équations uniques | 24 |
| 2.2.1 Étapes..... | 24 |
| A. Identification des variables explicatives..... | 24 |
| B. Détection des problèmes de multicolinéarité..... | 24 |
| C. Sélection des variables explicatives les plus significatives | 25 |
| D. Modélisation avec AUTOREG | 25 |
| E. Vérification de la justesse de la prévision | 28 |
| 2.2.2 Modèles retenus..... | 28 |

| | | |
|-------|---|----|
| 2.3 | Option Application de Prévisions de séries temporelles | 29 |
| 2.3.1 | Étapes..... | 29 |
| 2.3.2 | Options additionnelles..... | 30 |
| 2.3.3 | Modèles retenus..... | 30 |
| 3. | Analyses des résultats | 32 |
| 3.1 | Prévisions par modèle..... | 32 |
| 3.2 | Évaluation..... | 33 |
| 3.3 | Choix | 35 |
| 4. | Conclusion | 36 |
| 4.1 | Pousser plus loin la démarche de prévision | 36 |
| 4.2 | Recommandations..... | 38 |
| 5. | Annexes | 39 |
| 5.1 | Annexe 1 : Listes des variables utilisées pour le modèle prévisionnel touristique | 39 |
| 5.2 | Annexe 2 : Présentation visuelle des séries chronologiques assignées comme variables dépendantes..... | 43 |
| 5.3 | Annexe 3 : Liste des modèles préétablis du TSFS (ceux utilisés sont en caractère gras) | 46 |
| 6. | Bibliographie..... | 47 |

Liste des tableaux

| | | |
|-----------|--|----|
| TABLEAU 1 | Variations observées dans les séries des variables dépendantes (voir la représentation graphique en annexe 2)..... | 19 |
| TABLEAU 2 | Avantages et inconvénients des trois approches adoptées pour la modélisation de l'activité touristique..... | 20 |
| TABLEAU 3 | Modèles ARIMA retenus pour chaque variable dépendante..... | 23 |
| TABLEAU 4 | Modèles AUTOREG retenus pour chaque variable dépendante | 28 |
| TABLEAU 5 | Modèles associés à TSFS retenus pour chaque variable dépendante | 31 |
| TABLEAU 6 | Taux de variation de l'année prévue par rapport à l'année précédente (réelle) pour chaque variable dépendante..... | 32 |
| TABLEAU 7 | Évaluation des modèles retenus pour chaque variable dépendante..... | 34 |

Vue d'ensemble du processus de modélisation



Introduction

À la demande du ministère du Tourisme du Québec (MTO), l'Institut de la statistique du Québec (ISQ) a collaboré au projet en apportant l'expertise pertinente au développement d'un **modèle prévisionnel touristique** permettant d'estimer le nombre de visiteurs et les dépenses qu'ils ont effectuées au Québec pour les trimestres où Statistique Canada révisait l'Enquête sur les voyages des Canadiens (EVC).

Les résultats de cette démarche ont permis au MTO de fournir aux intervenants et à ses partenaires des estimations de l'évolution des volumes¹ et dépenses touristiques² effectués au Québec pour l'année 2005³ et ont rendu possible le suivi de la Politique touristique du Québec *Vers un tourisme durable*.

À moyen terme, la démarche de modélisation de l'activité touristique québécoise permettra au MTO, et tout particulièrement à la Direction de la recherche et de la prospective (DRP), de répondre aux demandes d'estimation et de prévision relatives à la performance touristique qui lui sont périodiquement adressées.

Sans pour autant diminuer la qualité des résultats, des limites ont toutefois restreint les possibilités de modélisation. Ainsi, certains modèles envisagés n'ont pu être expérimentés du fait que les séries chronologiques des variables à estimer étaient soit trop courtes, soit comportaient trop de données manquantes. Un autre facteur non négligeable est celui des perturbations auxquelles l'industrie touristique a eu à faire face au cours des dernières années et qui ont amené de l'instabilité ainsi que des cassures dans les séries, rendant la modélisation ardue.

Le présent document présente brièvement la démarche poursuivie pour atteindre les objectifs fixés et propose une série d'outils statistiques essentiels aux travaux d'estimation et de prévision du MTO. Le rapport fait état, dans un premier temps, de quelques rappels méthodologiques qui situent la démarche dans un cadre plus général basé sur une revue de littérature. On y traite des méthodes statistiques utilisées en prévision touristique et des éléments théoriques qui appuient la modélisation de la demande touristique. On y présente, également, les variables qui sont généralement considérées dans ce type de modélisation et les logiciels utilisés pour effectuer les travaux de modélisation. Dans un deuxième temps, les étapes à suivre lors de la modélisation sont commentées, et ce, pour chacune des approches retenues. Dans un troisième temps, la performance des modèles proposés est analysée en fonction de critères d'évaluation afin de retenir ceux qui proposent les prévisions les plus plausibles pour chaque variable analysée. Finalement, des pistes de recherche sont proposées afin de peaufiner le modèle prévisionnel touristique ainsi que des recommandations qui pourraient guider les travaux futurs du MTO.

-
1. Le volume touristique réfère au nombre de personnes qui ont visité le Québec.
 2. Les dépenses touristiques incluent les dépenses effectuées au Québec par tous les visiteurs et celles qui ont été effectuées au Québec par les Québécois pour des voyages réalisés à l'extérieur du Québec.
 3. L'exercice de prévision a été effectué sur l'année 2005 pour tous les marchés, même si les données des marchés internationaux étaient disponibles. Cependant, de façon générale, le modèle prévisionnel touristique sera utilisé deux fois par an, soit au printemps pour les estimations révisées de l'année précédente et les prévisions de l'année courante, et à l'automne pour les estimations de l'année courante et les prévisions de l'année suivante.

1. Rappels méthodologiques

Afin d'en arriver à développer un modèle prévisionnel, tout un travail de recherche documentaire, de collecte de données et de choix de logiciels a été nécessaire pour orienter le travail et fournir des bases solides pour le processus de modélisation.

1.1 Revue de littérature

La démarche entreprise s'est tout d'abord appuyée sur une revue de littérature qui a révélé l'abondance des travaux scientifiques en matière d'estimation et de prévision de l'activité touristique⁴. Ceux qui ont retenu le plus d'attention ont été réalisés par Douglas C. Frechtling (2001), Peter J. Brockwell et Richard A. Davies (2002), Christine Lim et Michael McAleer (2001), Jean Stafford et Bruno Sarrasin (2005), Stephen F. Witt et Christine F. Witt (1992), Cooper, Fletcher, Fyall, Gilbert et Wanhill (2005). Ils ont, entre autres, permis de connaître ou de recenser les principales méthodes quantitatives de prévisions utilisées par les chercheurs en tourisme et d'identifier les types de variables retenues dans les modélisations de la demande touristique. Les travaux de Madrikakis, Wheelwright et Hyndman (1999), de Nash et Nash (2002), de Nau (2005), de Brockwell et Davis (2002) ont été particulièrement appréciés dans l'apprentissage des méthodes quantitatives prévisionnelles.

1.2 Méthodes

Les méthodes quantitatives de prévision retenues sont de deux types, soit les méthodes extrapolatives et les méthodes causales, dites de régression multiple. Chacune se subdivise en sous-groupes qui se distinguent selon le type et le niveau de complexité des techniques statistiques qu'ils impliquent.

1.2.1 Extrapolatives

Le premier type rassemble les méthodes extrapolatives, c'est-à-dire celles qui permettent de prévoir une variable touristique à partir de son comportement ou de ses mouvements historiques. Ces techniques présument ainsi que le passé conditionne l'avenir. On peut aussi les appeler méthodes d'analyse et de prévision des séries chronologiques univariées et non causales (Witt et Witt, 1992), car elles impliquent une seule variable et ne permettent pas d'identifier directement les phénomènes qui expliquent les changements ou les modifications de la variable prévue.

Les méthodes extrapolatives les mieux connues sont : les méthodes naïves, la décomposition classique, les méthodes de lissage et l'approche Box-Jenkins.

4. Voir la bibliographie

Naïves

Les méthodes naïves s'appuient essentiellement sur les raisonnements suivants qui disent que la valeur de la variable prévue sera égale :

- à celle de la plus récente période pour laquelle l'information est disponible (naïve 1);
- à la valeur actuelle multipliée par le taux de croissance qui a prévalu au cours de la période précédente (naïve 2).

Ces méthodes sont, la plupart du temps, utilisées comme référence lors de l'évaluation des prévisions basées sur les autres méthodes qui impliquent des techniques statistiques plus complexes. Elles peuvent également être adaptées à des séries qui présentent des variations saisonnières (*seasonal naïve*) en utilisant comme période de référence ou de projection la période (mois, trimestre, etc.) de l'année précédente.

Décomposition classique

La décomposition classique établit la prévision à partir de la désagrégation successive de la variable prévue selon chacune des quatre composantes d'une série chronologique. Elle met ainsi en évidence la tendance ou le mouvement à long terme de la série, le cycle qui réfère, dans le présent exercice, à l'évolution à moyen terme de la demande (Hadj-Hamou, 2004), la saisonnalité qui réfère aux variations périodiques et répétées qu'imprègne le passage du temps en cours d'année (jours, mois, trimestres, saisons) à la série (Frechtling, 2005) et les résidus aléatoires qui représentent l'ensemble des variations qui ne sont pas expliquées par les autres composantes (Cordeau, 2004). On reconnaît généralement que cette méthode demande peu de moyens et qu'elle permet de produire des prévisions relativement précises en prolongeant dans un premier temps la tendance avec les coefficients estimés.

Lissage

Hadj-Hamou (2004) identifie trois principales méthodes de lissage : les moyennes mobiles (simples et doubles), les lissages exponentiels et les méthodes dites Holt-Winters. Les premières bonifient les méthodes naïves et sont utilisées lorsque les valeurs récentes d'une variable sont atypiques. Toutefois, elles présentent un intérêt limité pour les prévisions touristiques puisqu'elles rendent difficilement compte des variations saisonnières.

Même si les méthodes de lissage exponentiel sont plus complexes que les méthodes naïves et les moyennes mobiles, leur utilisation ne requiert qu'un chiffrier électronique. Ces techniques se distinguent de celles basées sur le calcul de la moyenne mobile en permettant de varier le poids des valeurs récentes de la variable (pondération de la moyenne) et de prendre en compte toute l'information contenue dans la série. Kalekar (2004) reconnaît que cette procédure permet de réviser continuellement les prévisions à la lumière des plus récentes expériences. Choy (1984) considère que leur efficacité en matière de prévision touristique est supérieure à celle des moyennes mobiles et qu'elle se compare avantageusement à l'approche plus complexe développée par Box-Jenkins. Le

même auteur ajoute que ces méthodes sont particulièrement efficaces pour prévoir des résultats mensuels.

Le lissage exponentiel simple permet de prévoir les séries chronologiques stationnaires (moyenne et variance constantes) sans saisonnalité, alors que le lissage exponentiel double⁵ s'applique aux séries qui sont soumises à une tendance linéaire dont on a retiré la saisonnalité. Leur utilisation dans le cadre de prévisions touristiques implique donc une étape préalable de désaisonnalisation⁶ et de différenciation⁷ des données. Ces méthodes exigent aussi, pour atteindre des résultats satisfaisants, des séries chronologiques longues, et elles ne s'appliquent qu'à un horizon de prévision qui couvre au plus deux périodes (Frechtling, 2001).

Holt et Winters ont poussé plus loin la méthode de lissage en développant une approche qui permet de composer avec une série comportant une tendance linéaire et des variations saisonnières (Dufour, 2002). Cette technique consiste en trois lissages exponentiels simultanés et utilise trois paramètres ou coefficients, notés α , β et γ , qui doivent être faibles si l'on considère que la valeur de la variable à un moment précis dépend d'un grand nombre d'observations antérieures, et élevés si on prétend le contraire. Elle fournit donc pour chaque période une estimation de la tendance, du coefficient saisonnier et de la valeur observée. Cette méthode a démontré un potentiel intéressant dans quelques études (Chu, 1998, Lim et McAleer, 2001) et permet d'effectuer des prévisions trimestrielles.

Approche Box-Jenkins

La littérature consultée identifie enfin un quatrième type de méthodes extrapolatives digne d'intérêt : l'approche Box-Jenkins. Cette dernière semble plus porteuse que les précédentes puisqu'elle permet de composer avec un large éventail de séries touristiques, de saisir les tendances non linéaires telles que la saisonnalité et de simuler les cycles. Après lui avoir reconnu de solides assises théoriques et une efficacité certaine en matière de prévision touristique, Frechtling présente cette approche comme la combinaison de deux méthodes de prévision, l'autorégression et la moyenne mobile, et de leur adaptation aux séries non stationnaires. Cette approche, qui a d'abord été popularisée par Box-Jenkins, est également appelée modélisation ARIMA (AutoRegressive/Integrated/Moving Average). La méthode ARIMA prévoit une ou plusieurs valeurs courantes d'une série comme une combinaison linéaire de ses propres valeurs passées, de certains chocs ou innovations passés, et de valeurs courantes et passées d'autres séries chronologiques. Elle est bien adaptée à l'horizon des prévisions prévues dans le présent projet (4 à 12 trimestres), à l'amplitude et à la discontinuité des séries à traiter.

Lorsque la méthode ARIMA inclut d'autres séries chronologiques comme variables explicatives, elle est parfois appelée ARIMAX. Plus particulièrement, dans la modélisation du marché touristique du

5. Le lissage exponentiel double répète deux fois le lissage simple pour tenir compte d'une tendance probable.

6. La désaisonnalisation d'une série économique consiste essentiellement à utiliser certaines techniques mathématiques afin d'expurger la série de ces variations infra-annuelles périodiques. (ISQ, 1997)

7. Type le plus courant de transformation des données. On la réalise en remplaçant chaque valeur de la série par la différence entre cette valeur et une valeur adjacente. Le nombre de périodes qui séparent la valeur de celle par laquelle on la soustrait représente l'ordre de différenciation.

Québec, les séries se rapportant aux volumes seront d'abord prévues, puis elles serviront d'intrants pour prévoir les séries se rapportant aux dépenses, et ce, pour chacun des marchés (Québec, autres provinces canadiennes, États-Unis et autres pays que les États-Unis). Il faut noter toutefois que la méthode ARIMA permet d'incorporer seulement un nombre limité de variables explicatives pour effectuer les prévisions, comparativement à ce que permettent les méthodes de régression multiple.

1.2.2 Causales

Les méthodes causales utilisent les techniques de régression multiple pour estimer et prévoir la demande touristique. Elles permettent d'identifier les relations fonctionnelles entre la variable qui fait l'objet de la prévision, dite dépendante, et un certain nombre de variables explicatives, dites indépendantes, qui dans la perspective présente, s'imposeront comme autant de déterminants du volume et des dépenses touristiques. On considère aussi que ces méthodes offrent la possibilité de développer un modèle explicatif du comportement de la variable dépendante.

La littérature identifie plusieurs méthodes causales. Les plus connues sont : les modèles de régression multiple (les modèles à équations uniques et les modèles à équations multiples), les modèles de simulation et les modèles d'espace état. Faute de temps et de ressources financières, les présents travaux ont été concentrés sur le développement de modèles de régression multiple puisque ce type de technique est vite apparu comme le plus porteur pour l'atteinte de l'objectif de départ qui était l'estimation du nombre de visiteurs et de leurs dépenses touristiques pour l'année 2005.

Modèle à équations uniques

Les avantages du modèle de régression multiple constitué d'équations uniques sont nombreux (Frechtling, 2001). Il permet de mesurer l'effet de phénomènes qui ont des impacts incontestables sur l'activité touristique. On peut ainsi quantifier l'effet de la hausse de prix, de la croissance des revenus, de l'évolution de la compétition, etc. Il s'avère, en outre, un outil de planification important puisqu'il permet de simuler les effets prévisibles des phénomènes identifiés plus haut. On reconnaît aussi que cette technique s'appuie sur un grand nombre de mesures qui permettent d'en évaluer la précision et qu'elle offre la possibilité de nuancer la simultanéité des relations en permettant de décaler ou de retarder (« lag »⁸) l'effet d'une variable.

Le modèle autorégressif d'erreurs prend en compte les écarts récents de la tendance pour produire les prévisions. Il permet de traiter les valeurs manquantes des variables dépendantes et indépendantes. Les données doivent être ordonnées et également distantes, ce qui est le cas dans la base de données de la demande touristique.

8. Un « lag » correspond à l'intervalle de temps qui sépare deux observations successives d'une série chronologique.

Modèle à équations multiples

Un autre modèle de régression multiple considéré est celui constitué d'équations multiples que la procédure VARMAX de SAS permet de traiter. Cette méthode estime les paramètres d'un modèle et génère les prévisions associées à un processus de vecteur (V) autorégressif (AR) et de moyenne mobile (MA) avec variables explicatives exogènes (X) ou endogènes. Souvent, les variables économiques ou financières ne sont pas seulement corrélées entre elles, mais elles sont aussi corrélées aux valeurs passées des autres variables. La procédure VARMAX permet de modéliser les relations dynamiques entre les variables dépendantes et entre les variables dépendantes et indépendantes. Cette procédure est d'un intérêt extrême pour le cas du modèle de la demande touristique. Autre avantage non négligeable de la méthode, il n'est pas nécessaire d'inclure des valeurs futures de variables explicatives pour générer des prévisions pour les variables dépendantes. Toutefois, la procédure VARMAX ne supporte pas actuellement la présence de valeurs manquantes et tient compte seulement du groupe d'observations continues dans une série. Elle exige un nombre d'observations qui dépend du nombre de paramètres à estimer. Les séries des variables sur les volumes et les dépenses des touristes québécois et des autres provinces canadiennes au Québec, variables pour lesquelles l'effort de modélisation est prioritaire, ne débutent qu'en 1998 (on y retrouve des données manquantes avant cette année), ce qui donne un nombre insuffisant d'observations pour les paramètres à estimer.

1.3 Variables

Tel que précisé auparavant, la revue de littérature a également permis d'identifier les types de variables qui sont généralement retenues pour prévoir la demande touristique. Il a été ainsi considéré que cette dernière se mesure généralement en nombre de touristes, en nuitées et en dépenses qui se répartissent selon l'origine des touristes qui les effectuent (Crouch, 1994). Il apparaît toutefois que les statistiques relatives aux dépenses touristiques sont généralement moins fiables, moins disponibles et plus difficilement mesurables, c'est pourquoi la plupart des analyses ne s'y attardent pas (Witt et Witt, 1992).

Les variables utilisées pour la prévision de la demande touristique peuvent être regroupées selon diverses typologies qui s'appuient sur des critères aussi variés que : leur rôle dans la procédure de prévision (dépendantes <-> indépendantes), la nature des phénomènes qu'elles expriment (sociaux, technologiques, économiques et politiques)⁹, ou leur effet appréhendé sur la demande touristique (Freichtling, 2001).

Dans cette dernière perspective qui apparaît sans aucun doute la plus complète et intéressante, un premier type de variables réunit celles qui sont reconnues comme des incitatifs au voyage, également appelées « *push* » ou « *emissive factors* ». Leur évolution encouragerait la population d'un marché touristique à voyager à l'extérieur du pays. On associe à ce type de variables la taille de la population,

9. Cette typologie est également appelée STEP (in. Cooper, Fletcher, Fyall, Gilbert et Wanhill, 2005, p. 122).

la croissance économique, la progression des revenus, la scolarité et l'accès au temps libre. Un deuxième type de variables regroupe celles qui sont reconnues pour attirer (« *pull* ») les visiteurs dans une destination. On leur associe le climat, les relations commerciales, les échanges culturels, les programmes de marketing, la présence de réseau de distribution, les événements spéciaux, etc. La typologie identifie enfin un troisième groupe de variables qui réunit celles qui freinent le développement des voyages entre un pays d'origine et une destination. Les principales variables de ce type sont : les prix dans la destination, les prix dans les destinations concurrentes, les prix des biens essentiels, le taux de change, le niveau de taxation, les actions promotionnelles des compétiteurs, la distance, le temps de voyage, le contexte politique, les catastrophes naturelles.

Le temps qui avait été imparti au projet n'a pas permis de constituer une base de données qui rassemble l'ensemble des variables identifiées par Freichtling. En fait, il a été nécessaire de s'appuyer sur une classification plus limitée, approche adaptée des travaux de Witt et Witt (1992) et Crouch (1994 et 1995). Leurs classifications des variables s'appuient sur une théorie économique et sur une méta-analyse de la demande touristique. En plus des variables dépendantes touristiques (volume et dépenses selon le pays d'origine), elles retiennent la population, la croissance économique et le revenu, les prix des biens et services touristiques dans le pays récepteur, les prix dans les destinations concurrentes, les taux de change, les dépenses promotionnelles, les variables muettes ou binaires (« *dummy* ») et la tendance. Également, des variables relatives à l'emploi des secteurs touristiques, à la confiance des consommateurs, aux revenus des entreprises et toute une série de données touristiques qui sont généralement utilisées pour évaluer la performance touristique lorsque les résultats officiels ne sont pas encore publiés ont été ajoutées.

Les variables sélectionnées devaient respecter un certain nombre de critères. Elles devaient ainsi être non désaisonnalisées, trimestrielles (ou offrir la possibilité de les ramener sur une base trimestrielle¹⁰), disponibles à compter de 1990 au minimum et produites par des sources fiables et reconnues.

Le MTO a fait l'inventaire des variables pouvant servir à l'analyse et a recueilli la plus grande partie des séries. L'ISQ a fourni les données des comptes économiques et recueilli certaines autres séries de façon à compléter cette base de données.

Voici les variables qui ont été considérées et, brièvement exposées, les considérations théoriques et pratiques qui ont justifié leur choix ou leur exclusion dans la modélisation du marché touristique québécois. Pour consulter la liste complète des variables utilisées pour le modèle prévisionnel touristique, veuillez vous référer à l'annexe 1 à la fin du rapport.

10. Exemple : prendre la moyenne des données mensuelles ou quotidiennes du trimestre concerné.

1.3.1 Dépendantes

Puisque le MTO exprime la performance touristique du Québec en ce qui a trait aux volumes et aux dépenses touristiques réparties selon quatre grands marchés, la démarche de modélisation a porté sur les huit variables dépendantes suivantes :

- Volume de touristes québécois au Québec;
- Volume de touristes des autres provinces canadiennes au Québec;
- Volume de touristes américains au Québec;
- Volume de touristes des autres pays que les États-Unis au Québec;
- Dépenses des touristes québécois au Québec;
- Dépenses des touristes des autres provinces canadiennes au Québec;
- Dépenses des touristes américains au Québec;
- Dépenses des touristes des autres pays que les États-Unis au Québec.

Les séries associées à ces variables sont trimestrielles et elles ont été produites par Statistique Canada dans le cadre de l'Enquête sur les voyages des Canadiens (EVC) et de l'Enquête sur les voyages internationaux (EVI). Il faut, par ailleurs, préciser que les données de l'EVC s'étendent de 1990 à 2004, qu'elles ont été diffusées sur une base bisannuelle entre 1990 et 1998 et qu'elles ont fait l'objet de nombreux changements méthodologiques (1994, 1996, 2000 et 2002). Les données de l'EVI sont plus stables. Elles sont produites depuis le début des années 1980 et elles n'ont fait l'objet de changements méthodologiques mineurs qu'en 1990 et 2000.

1.3.2 Indépendantes

Ces variables n'ont évidemment servi que dans l'exercice de régression multiple, sauf pour certains indicateurs touristiques dans ARIMAX. En outre, il faudra se rappeler que compte tenu du caractère exploratoire de la démarche, il a été convenu de retenir le plus grand nombre possible de variables explicatives recensées dans la revue de littérature. Plus d'une centaine de variables produites par différentes agences statistiques ont été retenues pour rendre compte de l'évolution de l'activité économique et de la performance touristique. Cette démarche était d'autant plus justifiée que les techniques statistiques de corrélation et de régression dites de pas à pas (*stepwise*), utilisées pour élaborer les modèles, permettaient d'utiliser un grand nombre de variables explicatives et d'éliminer celles qui n'expliquaient pas les variables dépendantes.

Croissance économique et revenus

Ce type de variables explicatives est le plus souvent utilisé dans les études de prévisions touristiques (Crouch, 1994). On reconnaît généralement qu'une bonne performance économique génère davantage de ventes, d'emplois et de revenus personnels qu'une économie stagnante. Ces conséquences entraînent à leur tour une hausse de la consommation des produits de luxe et, par

conséquent, davantage de voyages d'affaires et de loisirs. Les variables qui expriment le mieux ces phénomènes sont le produit intérieur brut (PIB), le revenu personnel disponible et le salaire.

Vingt et une variables de ce type ont été retenues dans la présente démarche dont sept sur le PIB des principaux marchés touristiques québécois (Québec, Canada, États-Unis, France, Royaume-Uni, Allemagne et Europe des 15) et six sur le revenu personnel disponible dans les mêmes marchés. Pour chacune de ces catégories de variables, les données relatives au Québec et au Canada ont été fournies par l'ISQ, alors que celles des autres marchés ont été tirées du portail Eurostat¹¹ et de l'Institut canadien de recherche sur le tourisme (ICRT).

Les modèles ont également été bonifiés par l'ajout de trois variables prévisionnelles de la conjoncture économique. Deux indicateurs avancés de Statistique Canada ont été retenus. Il s'agit d'outils de prévision à court terme (six à neuf mois) permettant de saisir les points de retournement des cycles de l'activité économique canadienne et américaine. Le troisième est l'indice précurseur Desjardins (IPD) qui se veut un outil d'analyse et de prévision permettant de mieux suivre l'évolution conjoncturelle de l'économie du Québec (trois à six mois).

Par ailleurs, comme ces trois variables permettent de prévoir l'activité économique, elles ont été retardées ou décalées (« *lag* ») en fonction de la période de prévision qu'elles sont théoriquement sensées couvrir (trois à neuf mois). Cet exercice permettait de synchroniser les variables dépendantes et la réalité économique « prévue » par les indices. Les indicateurs du Canada et des États-Unis ont ainsi été avancés de deux et trois trimestres, alors que l'indice précurseur Desjardins a été approché de un et deux trimestres.

Prix de voyage au Québec et au Canada

Les prix retiennent aussi l'attention des chercheurs. Ils s'imposent, en fait, comme la deuxième plus importante catégorie de variables explicatives utilisée dans les modèles de prévision depuis plus de trois décennies. Selon la typologie de Frechtling, les prix appartiennent à la catégorie de variables qui freinent le développement touristique puisque les clientèles potentielles ont à composer avec un revenu limité qui restreint leur propension à consommer des produits de luxe, comme les voyages, au fur et à mesure que les prix de ces derniers augmentent. L'inclusion de cette variable est d'autant plus justifiée que la sensibilité de la demande à la variation des prix tendrait à s'accroître avec le temps, notamment avec l'apparition de nouveaux modes de tarification. La littérature suggère aussi d'inclure le prix de plusieurs biens et services touristiques (Crouch, 1994).

Vingt-deux variables relatives aux prix des biens et des services touristiques du Canada et du Québec ont été retenues. Deux d'entre elles mesurent les variations de prix des biens et services consommés régulièrement par les Canadiens et les Québécois, à savoir l'indice des prix à la consommation (IPC) du Québec et celui du Canada. Neuf rapportent l'évolution du prix des produits et des services touristiques spécifiques comme la location d'unités dans les établissements d'hébergement du

11. Le portail Eurostat peut être consulté à l'adresse suivante : [http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_33076576&_dad=portal&_schema=PORTAL].

Québec, l'hébergement pour voyageurs (Québec et Canada), les services de voyage (Québec et Canada), le transport interurbain (Québec et Canada), l'utilisation de véhicules de loisirs (Canada) et les voyages organisés (Canada). Sept sont tirées des indicateurs nationaux du tourisme (INT) et montrent les changements de prix des services d'hébergement, des services de transport, des services de restauration, des autres biens et services touristiques, de l'ensemble des biens et services touristiques, des dépenses touristiques et des voyages. Une attention particulière a été apportée aux coûts de transports aériens puisqu'ils s'imposent comme un des premiers facteurs qu'un touriste doit considérer dans le choix d'une destination. Finalement, quatre autres variables reconnues pour leur effet sur les flux touristiques ont été retenues : le taux de change du dollar canadien par rapport à l'Euro, la livre Sterling, le dollar américain et le prix du baril de pétrole américain.

Les effets de la variation du taux de change sont très bien documentés (étude de Statistique Canada et rapport de l'ICRT). Il semble maintenant acquis qu'un taux de change défavorable diminue la compétitivité d'une destination et génère les effets suivants : une diminution du nombre de voyages des visiteurs étrangers, une augmentation des départs de la clientèle « domestique » vers l'étranger, une réduction des dépenses (notamment des dépenses consacrées aux voyages d'affaires) et des durées de séjour. À l'inverse, un taux de change favorable attirerait de nouvelles clientèles touristiques et susciterait plus de déplacements transfrontaliers.

Les conséquences de la hausse des prix du pétrole sont également bien documentées. Une étude menée par le Conference Board of Canada (CBC, 2004), pour le compte de la Commission canadienne du tourisme (CCT), montrait que le maintien des prix du pétrole à un niveau élevé pouvait aussi bien entraîner une baisse sensible de la demande touristique « domestique » qu'une diminution de la demande internationale.

Même s'il est fortement recommandé d'utiliser des variables qui expriment le rapport entre les prix en vigueur dans une destination par rapport à ceux du pays d'origine et de tenir compte du prix des biens et des services de voyages dans les destinations concurrentes, il n'a pas été possible de développer une série d'indicateurs qui répondent à ces exigences. Cette inclusion aurait nécessité la consultation d'autres instances du MTO, notamment de la Direction des marchés touristiques, ainsi qu'une fouille plus approfondie des sources de données et des traitements statistiques.

Parce qu'ils peuvent avoir un effet retardé sur la dynamique touristique, c'est-à-dire qu'une hausse des prix peut entraîner une baisse de volume quelques mois plus tard en raison des délais qui prévalent généralement entre la période de planification des voyages et leur réalisation, il a été nécessaire de décaler trente-deux variables relatives aux prix sur des périodes variant entre un et quatre trimestres. Les résultats permettent de voir dans quelle mesure les variations qui se produisent sur le plan des variables de prix ont un effet immédiat ou décalé dans le temps.

Intervention marketing

Il aurait été intéressant d'inclure les interventions marketing dans la démarche de modélisation puisqu'on reconnaît généralement que les dépenses et les actions qui y sont associées sont

susceptibles d'influencer sensiblement le niveau de la demande touristique internationale pour une destination. Certaines études ont estimé que les ratios coûts/bénéfices associés à ce type d'interventions pouvaient atteindre, selon les destinations, entre 1 :9 et 1 :220. Il serait d'autant plus justifié de les inclure dans la démarche prévisionnelle que les dépenses promotionnelles s'imposent comme une réponse à l'augmentation de la compétition sur le marché touristique international et comme une reconnaissance de l'importance économique de l'activité touristique dans la plupart des pays (Crouch, 1994).

Toutefois, l'utilisation de ce type de variables confronte la plupart du temps les chercheurs à des problèmes de colinéarité avec d'autres variables, notamment celles qui informent sur l'évolution du revenu. Par surcroît, les dépenses promotionnelles n'ont pas de signification intéressante lorsqu'on les considère sur une base trimestrielle et on peut croire que leur pouvoir prévisionnel aurait été limité dans le cadre de cet exercice qui vise, on se le rappellera, à estimer et à prévoir les flux et les dépenses touristiques sur une base trimestrielle. Ces limites imposaient, enfin, des transformations de la variable que le temps restreint n'a pas permis d'effectuer.

Dichotomiques

La plupart des études consultées utilisaient des variables dichotomiques¹² ou binaires (dummy) pour rendre compte de phénomènes et d'événements qui ont profondément marqué l'évolution de la demande touristique. Ce type de variable est également retenu pour mettre en évidence les discontinuités des enquêtes et les mouvements saisonniers. Quinze variables ont été identifiées, soit quatre pour rendre compte de l'effet de chacune des saisons touristiques et onze pour mesurer les impacts de certains événements et perturbations (Coupe du monde de soccer de 1998, attentats du 11 septembre 2001, Guerre en Irak, SRAS, etc.) ou pour considérer l'effet de changements inexplicables qui se sont produits dans la demande touristique.

Tendance

Il est d'usage d'inclure dans les variables explicatives l'effet de tendance ou le *momentum* d'une destination. Cet effet reflète les cycles de popularité, les modes et la persistance d'habitudes qui commandent l'évolution des destinations (Crouch, 1994, p.15). Le raisonnement qui soutient l'utilisation de ce type de variable se résume ainsi : le nombre de personnes qui choisissent une destination dépendrait du nombre de personnes qui y sont allées au cours des années précédentes. La tendance saisirait aussi l'effet de certaines autres variables qui ne sont pas explicitement comprises dans les équations, mais qui influencent incontestablement la performance touristique. Parmi elles, on compte les changements démographiques dans les pays d'origine, l'amélioration de la desserte aérienne, etc. Les variables tendance, ou *momentum*, sont relativement simples à construire. Il suffit de créer une nouvelle série en convertissant la première période analysée (mois, trimestre, année, etc.) en nombre et de l'incrémenter jusqu'à la fin de la période couverte ou encore de décaler

12. Une variable dichotomique, dite « dico », consiste à diviser une variable en deux classes ou catégories selon des critères arbitrairement fixés.

chronologiquement (*lag*) la variable dépendante. On considère toutefois que l'utilisation de ce type de variables doit être limitée puisqu'elle génère souvent un problème de multicolinéarité¹³ entre les variables explicatives et qu'elle en limite, par le fait même, l'influence dans le modèle explicatif.

Le caractère exploratoire de la démarche a permis de retenir, en dépit de cette dernière restriction, dix-sept variables de ce type. Ainsi, en plus de la tendance qui a été élaborée par incrémentation, il y a eu décalage des quatre variables dépendantes de volumes pour chaque trimestre qui compose une année.

Autres types de variables

D'autres variables ont été retenues pour vérifier l'impact du **niveau de consommation au Québec** et de **l'évolution du climat général de consommation** dans les principaux marchés touristiques. Les dépenses personnelles de biens et services de consommation au Québec et au Canada et celles qui sont associées aux biens et services touristiques apparaissent, dans cette perspective, susceptibles d'aider à prévoir les mouvements et les fluctuations de l'activité touristique. Les séries retenues ont été les dépenses en services de restauration et d'hébergement et celles de l'ensemble des dépenses en biens et services de consommation effectuées au Québec et au Canada.

Sept **indices de confiance des consommateurs** ont également été intégrés dans l'analyse. Ils expriment, pour leur part, les pronostics des consommateurs d'un pays ou d'une région à l'endroit des conditions économiques et leurs fluctuations reflètent les changements de comportements des consommateurs. On estime aussi que ces indices constituent un indicateur précurseur des ventes au détail, des dépenses de consommation et aussi de l'activité économique (Desjardins, 2004). Dans cette perspective, on estime donc que plus les consommateurs sont confiants en ce qui concerne leur situation financière et l'emploi, plus ils auront tendance à augmenter leurs dépenses de voyages.

Les indices de confiance des consommateurs des sept pays et régions suivants ont été inclus dans l'analyse : Allemagne, Canada, États-Unis, Europe, France, Québec et Royaume-Uni.

Le potentiel précurseur de ce type d'indices sous-tend qu'un écart pourrait exister entre le moment du changement de niveau de confiance et sa concrétisation en voyages. C'est pourquoi il y a eu création d'une variable décalée d'un trimestre pour chacun des indices précisés auparavant.

À l'instar du *Conference Board of Canada*, les **données économiques relatives à l'emploi** ont été utilisées dans l'exercice de modélisation. L'utilisation de ce type de variables apparaissait justifiée dans la mesure où les indicateurs d'emploi sont les premiers indicateurs mensuels qui permettent de connaître l'évolution de l'activité économique. On estime, en effet, qu'une augmentation du nombre d'emplois reflète généralement une hausse de l'activité des entreprises¹⁴. Ainsi, transposées au

13. La multicolinéarité apparaît lorsqu'il existe une relation linéaire parfaite ou presque parfaite entre deux variables explicatives d'un modèle.

14. Toutefois, il faut considérer que l'emploi est un indicateur retardé de l'activité économique, car l'augmentation de la demande entraîne tout d'abord une hausse des heures supplémentaires dans les entreprises et, par la suite, une hausse du nombre d'employés lorsque les effectifs ne suffisent plus à répondre à la demande

contexte touristique, les données relatives à l'emploi pouvaient s'avérer un reflet intéressant et relativement précis de la vitalité des entreprises touristiques. Quatre indicateurs d'emplois spécifiques au Québec ont été retenus : l'emploi dans l'ensemble des industries associées au tourisme, l'emploi dans les services d'hébergement et de restauration, l'emploi dans le secteur de la restauration et des débits de boissons, et l'emploi dans les services d'hébergement.

Les délais impartis pour réaliser le présent mandat n'ont pas permis d'inclure des indicateurs qui rapportent l'évolution du marché du travail dans les principaux marchés touristiques québécois. Il aurait fallu tenir compte des assertions qui avancent que l'évolution du marché du travail et les incertitudes qui l'entourent influencent non seulement le volume de la demande touristique, mais qu'elles en déterminent aussi la nature (Cooper, Fletcher, Fyall, Gilbert et Wanhill, 2005, p. 115)¹⁵.

Également, les **revenus des entreprises** de certains secteurs touristiques ont été retenus puisque ces derniers sont reconnus par la plupart des organismes de développement touristique (Travel industry association of America, Commission canadienne du tourisme, Observation, Développement et ingénierie touristiques, etc.) comme des indicateurs importants de la performance touristique d'un pays. Les seules séries statistiques de cette nature qui étaient disponibles sur une longue période (1991-2005) et sur une base suffisamment fiable étaient les recettes des restaurants et celles des établissements d'hébergement. Les premières ont été extraites de l'enquête sur les statistiques des restaurants, traiteurs et tavernes de Statistique Canada, qui présente les estimations mensuelles des recettes totales de ces établissements sur une base canadienne et provinciale. C'est l'Enquête sur les établissements d'hébergement, produite à chaque mois par l'Institut de la statistique du Québec pour le compte du MTO, qui a permis de colliger les deuxièmes.

Étant orientés sur l'estimation rétroactive (l'année précédente) et la prévision à court terme (un trimestre à deux années), il a été jugé important d'insérer dans la modélisation un ensemble d'indicateurs touristiques qui, même s'ils ne sont pas utilisés dans les exercices de modélisation de la demande touristique examinés, sont reconnus par les intervenants comme les premiers indicateurs de performance de l'activité touristique. Ils doivent être considérés comme des indicateurs prédéterminés dans un sens statistique, car ils sont recueillis avant les données des variables dépendantes.

Ces indicateurs touristiques sont produits dans trois cadres différents. Le premier informe sur les entrées aux frontières des touristes étrangers et sur les entrées des Canadiens revenant de l'étranger. Ceux-ci ont été désagrégés en dix-huit variables selon les deux grands marchés étrangers (États-Unis et autres pays que les États-Unis) et les trois portes d'entrées suivantes : Québec, Ontario et Canada. Les entrées aux frontières sont diffusées sur une base mensuelle, six semaines après le mois de référence, par Statistique Canada. Elles pourraient jouer un rôle de premier plan dans les estimations et les prévisions du volume et des dépenses des touristes puisqu'elles constituent le seuil minimal du volume de touristes en question.

15. Un ralentissement du marché du travail entraînerait ainsi des retards de réservations, davantage de tourisme domestique, des séjours plus courts, une diminution de la demande pour l'hébergement commercial et une augmentation des visites chez les parents et amis.

Le deuxième cadre concerne les indicateurs nationaux du tourisme canadien. Ceux-ci permettent de suivre l'évolution du tourisme au Canada en fonction des six thèmes suivants : la demande touristique, la demande touristique intérieure, les exportations touristiques, l'offre touristique, l'emploi touristique et le produit intérieur brut. Ils sont publiés huit semaines après le trimestre de référence. Ceux-ci permettraient d'évaluer l'évolution des volumes et des dépenses générés par les quatre grands marchés (Québec, autres provinces canadiennes, États-Unis et autres pays que les États-Unis). Les cinq indicateurs suivants ont été retenus dans cette perspective : la balance au compte de voyages internationaux, la demande touristique intérieure, la demande touristique totale, les dépenses au Canada par les résidents étrangers, les dépenses dans des pays étrangers par les Canadiens et les exportations touristiques.

Le troisième et dernier cadre a été tiré de l'enquête sur les établissements d'hébergement que réalise l'ISQ pour le compte du MTO. Comme son nom l'indique, les variables qui en découlent informent sur l'évolution de l'industrie québécoise de l'hébergement et, plus spécifiquement, sur la performance des établissements hôteliers et des résidences de tourisme. On estime que ces variables sont performantes pour l'estimation des volumes et des dépenses touristiques internationales et canadiennes. Trois variables ont été retenues : le taux d'occupation dans les établissements d'hébergement du Québec, les unités disponibles et les unités louées dans les établissements d'hébergement du Québec.

Enfin, même s'il appert généralement que les pays peuplés génèrent davantage de voyages (Witt et Witt, 1991), aucune variable démographique n'a été retenue dans l'exercice puisqu'on reconnaît que ce type de variable engendre souvent des problèmes de multicolinéarité dans les modélisations par régression des marchés touristiques. Ce problème se produit généralement lorsque deux variables explicatives sont fortement corrélées entre elles. Il entraîne de l'instabilité dans le modèle et des difficultés à produire des résultats fiables.

L'annexe 1 identifie les séries de données trimestrielles utilisées et regroupées selon le type de variable (dépendante et indépendante). On y indique également le trimestre pour lequel la série débute et se termine ainsi que la nature de la variable qui réfère aux catégories théoriques identifiées auparavant (ex. : croissance et revenus, prix, variables dichotomiques, etc.).

1.4 Logiciels

Les logiciels retenus pour cet exercice de modélisation sont Excel et SAS. Excel, pour sa facilité d'utilisation, son accessibilité pour les autres ressources impliquées dans la démarche et ses fonctionnalités de transfert vers Word pour la production de rapports, a été retenu pour gérer la base de données, analyser et conserver les prévisions des modèles. Quant à SAS, un logiciel performant d'analyse statistique, il a servi à effectuer la démarche de modélisation et à produire les prévisions.

2. Modélisation

La démarche pour la modélisation doit respecter un certain nombre d'étapes qui diffèrent selon les modèles utilisés. La présente partie est donc structurée de façon à mettre en évidence les techniques statistiques utilisées pour chacune des trois approches testées afin de produire des prévisions.

Tel que précisé auparavant, l'effort de modélisation a porté sur les variables dépendantes suivantes : le volume de touristes québécois au Québec, les dépenses des touristes québécois au Québec, le volume de touristes américains au Québec, les dépenses des touristes américains au Québec, le volume de touristes des autres provinces canadiennes au Québec, les dépenses des touristes des autres provinces canadiennes au Québec, le volume de touristes des autres pays que les États-Unis au Québec et les dépenses des touristes des autres pays que les États-Unis au Québec.

Une analyse visuelle des séries relatives à l'ensemble des variables a d'abord été effectuée. Cette étape essentielle vise à bien connaître les caractéristiques (tendance, saisonnalité, cycle, données extrêmes, etc.) des séries. Elle permet même de suggérer quelquefois des explications aux variations de la série (Makridakis, Whellwright et Hyndmand, p.23). L'information ainsi obtenue a permis d'appuyer sur une base objective le choix du modèle d'estimation (SAS/ETS Users Guide, version 8, p. 1281).

Ce premier examen visuel a révélé que :

- les séries analysées présentent d'importantes variations saisonnières (saisonnalité);
- la présence d'une tendance dans la série n'est pas toujours observable;
- la dispersion des valeurs des variables autour de leur valeur moyenne est faible, sauf pour le volume de touristes des autres pays que les États-Unis au Québec;
- les fluctuations saisonnières ne s'accroissent pas avec le temps (modèle additif de décomposition de séries);
- certaines séries comportent de l'instabilité;
- la présence de chocs modifie profondément le patron dans quelques cas.

Le tableau 1 présente les résultats de l'analyse visuelle pour chacune des variables dépendantes.

TABLEAU 1 Variations observées dans les séries des variables dépendantes (voir la représentation graphique en annexe 2)

| Variable | Saisonnalité | Tendance | Dispersion | Chocs ¹⁶ |
|---|--------------|----------|------------|---|
| Volume de touristes au Québec | | | | |
| Québécois | Oui | Non | Faible | Choc négatif au 3 ^e trimestre 2001 à la suite des attentats du 11 septembre 2001 sur les deux tours du World Trade Center Changement structurel positif à partir du 1 ^{er} trimestre 2002 à la suite du changement méthodologique de l'EVC en 2002 Choc négatif au 3 ^e trimestre 2003 à la suite des cas d'infections au syndrome respiratoire aigu sévère (SRAS) et de la guerre en Irak |
| Autres provinces canadiennes | Oui | Non | Faible | Choc négatif inexplicable au 3 ^e trimestre 1998 Changement temporaire négatif inexplicable allant du 1 ^{er} au 4 ^e trimestre 2000 Changement structurel négatif à partir du 1 ^{er} trimestre 2003 à la suite des cas d'infections au syndrome respiratoire aigu sévère (SRAS) et de la guerre en Irak |
| Américains | Oui | Possible | Faible | Changement temporaire positif du 1 ^{er} au 4 ^e trimestre 2002 à la suite des attentats du 11 septembre 2001 sur les deux tours du World Trade Center Changement temporaire négatif du 2 ^e au 3 ^e trimestre 2001 à la suite des attentats du 11 septembre 2001 sur les deux tours du World Trade Center Changement structurel positif inexplicable à partir du 3 ^e trimestre 1998 |
| Autres pays que les États-Unis | Oui | Non | Moyenne | Choc négatif au 4 ^e trimestre 2001 à la suite des attentats du 11 septembre 2001 sur les deux tours du World Trade Center Choc négatif au 3 ^e trimestre 2003 à la suite des cas d'infections au syndrome respiratoire aigu sévère (SRAS) et de la guerre en Irak Choc positif inexplicable au 3 ^e trimestre 1997 |
| Dépenses des touristes au Québec | | | | |
| Québécois | Oui | Possible | Faible | Identiques à ceux du volume de touristes québécois au Québec |
| Autres provinces canadiennes | Oui | Oui | Faible | Série trop instable pour fins d'analyse |
| Américains | Oui | Oui | Faible | Identiques à ceux du volume de touristes américains au Québec |
| Autres pays que les États-Unis | Oui | Possible | Faible | Identiques à ceux du volume de touristes des autres pays que les États-Unis au Québec |

L'analyse visuelle devient un élément important faisant ressortir les caractéristiques des séries tout en suggérant un choix sur la méthode d'estimation. Parmi les méthodes décrites précédemment, deux ont été privilégiées afin de produire des prévisions d'estimations des variables dépendantes : ARIMA (approche Box-Jenkins) et AUTOREG (modèle à équation unique). La modélisation de ces deux méthodes (ou modèles) a été effectuée à partir des procédures décrites avec le progiciel SAS. Pour fin de comparaison et de mesure d'efficacité, la modélisation a également été réalisée avec l'outil interactif d'**application de prévision de séries temporelles**¹⁷ disponible dans le module ETS de la version 9.1.3. Le tableau 2 résume les avantages et les inconvénients de ces trois approches.

16. Voir la liste des variables dichotomiques en annexe 1.

17. Ce système est aussi appelé le Time Series Forecasting System (TSFS).

TABLEAU 2 *Avantages et inconvénients des trois approches adoptées pour la modélisation de l'activité touristique.*

| Méthode | Avantages | Inconvénients |
|---|--|---|
| Modèle ARIMA | Construction assez simple et rapide Permet de faire des estimations par la variable elle-même Permet de traiter des séries chronologiques comportant de la saisonnalité, des données manquantes ou des chocs | Peut inclure seulement un nombre limité de variables explicatives N'explique pas les causes de l'évolution des variables dépendantes Requiert un nombre minimal d'observations pour donner des prévisions acceptables Exige d'avoir des données pour les variables explicatives sur l'horizon de prévision si des variables prévisionnelles sont présentes |
| Modèle de régression linéaire multiple (AUTOREG) | Inclut des variables explicatives Peut expliquer les causes de l'évolution de la variable dépendante Permet de traiter des séries chronologiques comportant de la saisonnalité, des données manquantes ou des chocs | Construction plus complexe Exige d'avoir des données pour les variables explicatives sur l'horizon de prévision Requiert un nombre d'observations minimal pour donner des prévisions acceptables. De plus, ce nombre est d'autant plus élevé que le modèle est complexe. |
| Option automatique du module TSFS | Construction simple et rapide Choix de 42 modèles prédéfinis Permet de traiter des séries chronologiques comportant de la saisonnalité et des données manquantes Requiert, pour certains modèles, peu d'observations pour donner des prévisions acceptables | Programmes prédéfinis avec des limites d'options Ne peut expliquer les causes de l'évolution des variables dépendantes |

2.1 Modèle ARIMA

La procédure ARIMA de SAS analyse et prévoit des valeurs courantes ou futures de séries chronologiques univariées comportant des données également distantes (jour, semaine, mois, année, trimestre) en utilisant un modèle **AutoRegressive Integrated Moving-Average** (ARIMA ou approche Box-Jenkins), modèle autorégressif à moyenne mobile intégrée en français.

2.1.1 Étapes

La modélisation ARIMA s'effectue en trois étapes : l'identification, l'estimation et la prévision.

A. Identification

Cette première étape permet d'identifier la variable à analyser, de vérifier la stationnarité de la série chronologique (stabilité dans le temps pour les valeurs d'autocorrélation des résidus) et de déterminer la combinaison d'autorégression et de moyenne mobile la plus pertinente.

Vérification de la stationnarité¹⁸ de la série

On vérifie si le comportement de la série est animé par une tendance, des variations d'amplitude ou des mouvements saisonniers et on modifie les données au besoin en les transformant selon leur

18. La série est dite stationnaire lorsque la moyenne de la série ainsi que la variance sont constantes dans le temps.

valeur logarithmique ou en les différenciant¹⁹. L'analyse de la stationnarité s'effectue à partir des graphiques d'autocorrélation et de bruit blanc. Le bruit blanc est mesuré par une vérification de l'autocorrélation des résidus (les valeurs du Khi-2 doivent être non significatives). Une fois la série analysée transformée, s'il y a toujours présence d'une tendance ou de fluctuations importantes, cela révèle qu'elle n'est pas stationnaire. Dans ce cas, l'approche Box-Jenkins n'est pas recommandée.

Détermination des combinaisons d'autorégression et de moyenne mobile

Une fois stationnaire, il faut déterminer les modèles ARIMA potentiels selon des combinaisons d'autorégression et de moyenne mobile qui sont les plus pertinentes pour assurer une bonne prévision. Les modèles que cet exercice permet de déterminer sont suggérés par l'examen des graphiques d'autocorrélation et d'autocorrélation partielle. Les options ESACF, SCAN et MINIC de SAS génèrent des tableaux dont l'interprétation permet d'estimer les valeurs pour les corrections d'autorégression et de moyenne mobile.

B. Estimation

Cette deuxième étape sert à faire l'examen des modèles identifiés à l'étape 1, à les estimer par maximum de vraisemblance et à déterminer, selon divers tests, lequel serait le plus performant. Ce choix est effectué en fonction des statistiques de diagnostic suivantes²⁰ :

Test de Student

Une première statistique, le test de Student, donne le niveau de signification des termes d'autorégression (AR), de moyenne mobile (MA) et de la moyenne du modèle (MU ou constante du modèle) dans le modèle. Cette statistique indique si certains termes ne sont pas nécessaires dans le modèle vu leur faible degré de signification. On s'assure également que le signe du paramètre s'interprète de façon adéquate.

Akaike information criterion (AIC)

Une deuxième statistique, qui rend compte de la qualité de l'ajustement (Goodness-of-fit) du modèle, ici AIC, permet de comparer les divers modèles de la même série entre eux. Le modèle ayant le plus petit AIC devrait être celui qui est le plus en mesure de prévoir avec justesse les données.

Corrélation entre les paramètres

Une troisième statistique permet de voir la corrélation entre les paramètres d'autorégression, de moyenne mobile et la constante. Si deux paramètres corrélaient trop fortement entre eux (plus de 0,8

19. Réfère à la procédure de différenciation définie plus haut.

20. Si le diagnostic établi avec les diverses statistiques montre que le modèle testé rencontre des problèmes, il faut essayer les autres modèles en reprenant les mêmes étapes, soit les étapes d'identification et d'estimation.

ou de -0,8), cela signifie qu'ils estiment la même chose. Il vaut donc mieux dans ce cas n'en conserver qu'un seul, soit généralement le plus significatif des deux selon le test t de Student.

Autocorrélation des résidus

Une quatrième statistique de vérification d'autocorrélation des résidus, le Chi-Deux, montre si les résidus du modèle contiennent des informations additionnelles pouvant être prises en compte par un modèle plus complexe.

Détection des chocs

Il existe aussi une autre statistique, que l'on retrouve dans l'énoncé « Outlier », qui permet de voir si le modèle tient compte de toutes les variations de la série et de détecter des chocs qui viendraient expliquer ces variations. Il existe plusieurs types d'« outlier », mais dans le présent exercice, trois ont été retenus : l'« outlier » additif (« additive » [AO]), changement temporaire (« temporary change » [TC]) et changement de niveau (« level shift » [LS]). L'additif est un choc ayant un impact momentané, soit sur une seule période. Le changement temporaire est un choc qui a une influence sur un nombre plus ou moins élevé de périodes se situant généralement, pour les trimestres, entre deux et quatre périodes. Le changement de niveau se veut une cassure dans la série qui sépare celle-ci en deux niveaux distincts. Les types d'« outlier » ainsi découverts forment des variables explicatives que l'on peut ajouter au modèle, en créant une variable dichotomique, pour essayer de l'améliorer. Il est important d'éviter d'introduire trop de variables de type « outlier » dans un modèle, car il y a danger de « sur-ajustements » (« over-fitting »). Pour les fins de l'exercice, un maximum de trois « outlier » était attribué par variable dépendante. La façon de procéder pour introduire ces nouvelles variables dans le modèle sera discutée au point « Option additionnelle » un peu plus bas.

C. Prévision

À la troisième étape, on effectue les prévisions des données futures de la série et on génère un intervalle de confiance pour ces prévisions pour le modèle retenu. La validation se fait en comparant l'évolution (amplitude et sens de la variation) de la variable prévue à celle d'indicateurs reconnus et fortement corrélés.

De même, à cette étape, il faut vérifier la justesse de la prévision. Cela implique essentiellement de tester les modèles qui ont eu les meilleurs résultats dans les étapes précédentes pour prévoir la dernière année connue ou disponible de la série²¹. Le modèle retenu est celui qui présente le plus petit écart entre la valeur réelle et la valeur prévue. Un modèle est retenu seulement si cet écart est inférieur à 5 %²².

21. Si la série est assez longue, il peut être bon de faire le test sur deux années. Dans le présent exercice, nous avons testé l'année 2004.

22. La sélection du modèle pourrait également impliquer la consultation d'experts en tourisme.

2.1.2 Option additionnelle

Comme expliqué précédemment, le modèle ARIMA permet d'incorporer des variables explicatives pour effectuer les prévisions, mais en nombre plus limité que ce que permettent les modèles de régression multiple. Cet ajout s'effectue en trois étapes :

- déclarer à l'étape d'**identification** les variables explicatives (incluant les « outlier ») à l'aide de l'option « crosscorr » en attribuant les mêmes valeurs de différenciation que la variable;
- introduire ces nouvelles variables à l'étape **estimation** à l'aide de l'option « input » ;
- vérifier la contribution de chacune d'elles à l'aide des tests précédemment décrits.

2.1.3 Modèles retenus

Le tableau 3 présente les modèles ARIMA retenus pour chacune des variables dépendantes. La notation adoptée est la notation standard de ces modèles et celle du logiciel SAS, soit le modèle $ARIMA(p,d,q)(p,d,q)_s$, où S indique la partie saisonnière. Dans le marché touristique, on a $S=4$, car les données sont trimestrielles. Les autres lettres représentent les éléments suivants :

- p est la partie autorégressive;
- d est l'ordre de différenciation;
- q est l'ordre du processus de moyenne mobile.

Ainsi, $p=1$ dans la première partie, indique un processus autorégressif d'ordre un et $p=1$ dans la partie saisonnière indique un processus autorégressif d'ordre quatre puisque les données sont trimestrielles. Il arrive qu'un modèle comporte seulement des paramètres saisonniers et la notation s'inscrit alors $ARIMA(p,d,q)_s$.

TABLEAU 3 *Modèles ARIMA retenus pour chaque variable dépendante*

| Variable dépendante | Modèle retenu | |
|---|---------------------------------|--|
| | Notation | Variables explicatives |
| Volume de touristes au Québec | | |
| Québécois | $ARIMA(0,1,0)_4$ | Unités louées mensuellement dans les établissements d'hébergement du Québec |
| Autres provinces canadiennes | $ARIMA(1,1,0)_4$ | LS2001 et AO200303 |
| Américains | $ARIMA(0,1,1)_4$ | Entrées aux frontières des touristes américains au Québec |
| Autres pays que les États-Unis | $ARIMA(0,1,0)_4$ | Entrées aux frontières des touristes des autres pays que les États-Unis + AO200303 |
| Dépenses des touristes au Québec | | |
| Québécois | $ARIMA(0,1,0)_4$ | Unités louées mensuellement dans les établissements d'hébergement + prix de location dans les établissements d'hébergement |
| Autres provinces canadiennes | Aucun | |
| Américains | $ARIMA(0,1,1)_4$ avec constante | Volume de touristes américains au Québec + AO2003Q3 |
| Autres pays que les États-Unis | $ARIMA(0,1,0)_4$ avec constante | Volume de touristes des autres pays que les États-Unis au Québec |

Il ressort clairement que pour prévoir la demande touristique du Québec, la saisonnalité joue un grand rôle. Il appert aussi que les variables touristiques (unités louées mensuellement dans les établissements d'hébergement du Québec, entrées aux frontières des touristes américains au Québec, entrées aux frontières des touristes des autres pays que les États-Unis) sont de bonnes variables prévisionnelles pour les volumes et que la prévision des volumes, dans le cas des marchés internationaux, sert à prévoir les dépenses.

Les variables touristiques utilisées comme variables explicatives sont de type endogène, ce qui peut amener des problèmes de simultanéité. Cependant, on peut considérer ces variables comme des variables prédéterminées, dans un sens statistique, par rapport aux variables dépendantes puisqu'elles sont recueillies avant. De plus, l'ajout de variables explicatives touristiques aux modèles ARIMA a rendu ces derniers moins complexes tout en améliorant leurs performances de prévision. Le but premier étant de produire des prévisions justes et fiables, il était approprié de les retenir.

2.2 Modèles à équations uniques

La construction des modèles à équation unique s'est effectuée en cinq étapes.

2.2.1 Étapes

A. Identification des variables explicatives

Une revue de littérature a d'abord permis d'identifier les variables qui sont généralement considérées dans les modèles de prévision de la demande touristique. Ces dernières ont été, pour l'essentiel, présentées dans la section 1.3.

Les variables pouvant avoir l'effet le plus significatif sur la variable dépendante ont été identifiées et retenues sur la base du niveau de corrélation qu'elles entretenaient avec les variables dépendantes. Dans cette perspective, seules les variables présentant une corrélation suffisamment élevée étaient retenues (R de Pearson $> 0,3$ et $< -0,3$).

La procédure CORR de SAS a été utilisée à cette étape. De plus, pour qu'une variable explicative soit retenue, le lien entre cette variable et la variable dépendante devait avoir un sens reconnu dans la littérature et la théorie.

B. Détection des problèmes de multicollinéarité

Ensuite, certaines variables explicatives, qui étaient fortement corrélées ou qui variaient de la même manière en fonction du temps, ont été supprimées. Ce problème de multicollinéarité, fréquemment rencontré en prévision touristique (Crouch, 1994), peut générer de l'instabilité dans l'appréciation des erreurs et la production des estimations. Une forte corrélation nuit au calcul des coefficients de corrélation associés à chaque variable. Frechtling rapporte même que des logiciels de traitement de

séries chronologiques peuvent fournir des estimations différentes en utilisant les mêmes modèles et les mêmes données lorsqu'ils ont à composer avec la multicollinéarité.

Afin d'éviter ce type de problème, des procédures de validation ont été mises en place selon la méthodologie proposée par Frechtling. Il faut calculer les coefficients de corrélation de Pearson entre les différentes variables indépendantes retenues à l'étape précédente. Si la valeur est supérieure à 0,80, il y a risque de colinéarité, car les deux variables indépendantes peuvent avoir plus d'influence entre elles qu'avec la variable dépendante, d'où un problème important pour l'interprétation du modèle. À ce stade, il faut éviter de conserver des variables dont le coefficient de Pearson est plus élevé que chacun pris indépendamment avec la variable dépendante. Si la situation se présente, il faut éliminer l'une des deux variables, ou, si la littérature confirme l'importance des deux, transformer l'une des deux variables, soit en forçant un retard d'une période dans le temps (*lag*), soit en les différenciant d'une période.

C. Sélection des variables explicatives les plus significatives

La procédure REG de SAS et deux de ses options ont été utilisées pour sélectionner automatiquement les variables les plus significatives et les plus pertinentes pour la modélisation de la demande touristique. L'option « pas à pas » (*stepwise*) a permis d'isoler, à l'aide de la statistique F, les variables qui expliquent le mieux la variance des variables dépendantes. Il a été aussi jugé intéressant de mesurer le niveau d'incidence des variables sélectionnées avec l'option INFLUENCE²³. Cette démarche a permis d'éliminer les variables qui n'étaient pas statistiquement pertinentes et dont la présence accroissait dans le futur la possibilité d'erreurs de prévision.

La procédure REG se veut une régression par moindres carrés ordinaires qui constitue déjà un choix de modèle pour AUTOREG, car elle signifie implicitement qu'il n'est pas nécessaire d'effectuer une correction due au fait que les résidus sont autocorrélés dans le temps.

D. Modélisation avec AUTOREG

La quatrième étape est celle de la modélisation proprement dite. Elle est effectuée à partir de la procédure AUTOREG de SAS à laquelle on applique les variables indépendantes sélectionnées à l'étape 3.

Le modèle retenu à l'étape précédente a été soumis à la procédure AUTOREG, en y ajoutant au besoin certaines variables indépendantes, dichotomiques ou encore une variable de tendance, si l'on juge qu'elles ont leur place dans le modèle sur une base théorique ou autre. À l'occasion, il a été nécessaire d'appliquer un opérateur de délai d'ordre n à une variable dépendante ou indépendante introduite comme variable explicative dans le modèle, de façon à mieux faire ressortir une relation qui serait retardée (*lag*). À ce moment-là, le suffixe ln (n correspond à l'ordre du délai) est ajouté au nom de la variable.

23. Cette option détecte si une observation a particulièrement plus d'influence (quand la statistique Dffits en valeur absolue est ≥ 2) que les autres sur l'estimation des paramètres.

La méthode d'estimation privilégiée est l'estimation par maximum de vraisemblance (*maximum likelihood*). Elle a été retenue parce qu'elle est plus générale et qu'elle est en mesure d'apporter les modifications qui permettent de répondre aux exigences de la distribution aléatoire et de l'égalité de la variance des erreurs ou des résidus (correction de l'hétéroscédasticité des résidus). Elle offre, en effet, la possibilité d'utiliser des modèles de type GARCH ou des modèles mixtes de type AR-GARCH.

La procédure AUTOREG produit une série de tests statistiques de diagnostic qui permettent de vérifier la validité des conditions de base, soit l'indépendance des erreurs, la normalité et l'homogénéité de la variance (à l'opposé, l'hétéroscédasticité) des résidus. Voici les tests effectués :

Durbin-Watson

Le premier test, celui de Durbin-Watson, permet de vérifier si le modèle répond à la condition d'indépendance des erreurs et d'apporter au besoin les corrections d'usage. On reconnaît aussi qu'il permet d'évaluer si une ou plusieurs variables indépendantes importantes n'ont pas été considérées dans l'exercice de modélisation.

En effet, il arrive fréquemment que les résultats d'une analyse de régression linéaire de données chronologiques présentent des erreurs ou des résidus qui ne sont pas indépendants dans le temps. On dit alors de ces erreurs qu'elles sont sériellement corrélées ou autocorrélées. Cette caractéristique entraîne des biais et des erreurs d'estimation des paramètres du modèle de régression qui deviennent par le fait inefficaces, l'estimation de leur écart-type étant biaisée. Ainsi, les tests associés aux paramètres indiquant si ces derniers sont significatifs ou non ne sont plus fiables. On pourrait donc rejeter une variable explicative, alors qu'elle est significative ou l'inverse, en se basant sur les tests de Student utilisés pour vérifier la signification des paramètres de la régression.

Pour tester la présence ou non d'autocorrélation, les statistiques généralisées de Durbin-Watson (DW), leurs probabilités marginales (option DWPROB) et les valeurs p de probabilité exacte sont données pour tout ordre d'autocorrélation spécifié. Dans le modèle prévisionnel touristique, on pourrait s'attendre à des autocorrélations d'ordre un, corrélation sérielle en regard du trimestre précédent, et à des autocorrélations d'ordre quatre, corrélation saisonnière en regard du même trimestre de l'année précédente. La probabilité d'autocorrélation positive ($Pr < DW$) et négative ($Pr > DW$) doit être supérieure à 0,05, sinon il faut corriger chaque délai pour lequel un ordre a été identifié comme significatif. La procédure AUTOREG peut sélectionner de façon automatique le modèle autorégressif pertinent des erreurs à l'aide de l'option « backstep ».

Test du portemanteau

Pour tester la présence d'hétéroscédasticité, ce qui contreviendrait à l'hypothèse d'homoscédasticité à laquelle doit répondre un modèle de régression linéaire, il faut utiliser le test du portemanteau et les tests statistiques du multiplicateur d'Engle Lagrange. En effet, l'analyse de régression ordinaire suppose que la variance des erreurs est la même pour chaque observation. Lorsque la variance des

erreurs n'est pas constante, on dit que les données sont hétéroscédastiques et les estimations par moindres carrés ordinaires ne sont plus efficaces. L'hétéroscédasticité affecte aussi la précision des intervalles de confiance des prévisions.

Les tests statistiques et les valeurs de probabilité p indiquant le niveau de signification sont donnés pour l'hétéroscédasticité conditionnelle pour les délais de 1 à 12. La famille des modèles GARCH, au moyen de la procédure de SAS, est également en mesure d'estimer et de corriger la variabilité changeante des données. Pour les modèles de type GARCH, la procédure AUTOREG produit l'erreur de prévision conditionnelle des variances des erreurs prévues ainsi que les estimations et les covariances des paramètres estimés.

L'hétéroscédasticité peut ainsi être éliminée à l'aide de l'option ARCHTEST qui assure que les variances des erreurs du modèle sont homogènes. Les probabilités des tests Q ($Pr > Q$) et LM ($Pr > LM$) doivent alors être supérieures à 0,05, sinon il faut y apporter des corrections en spécifiant l'option GARCH= (q=1, p=1) qui est la correction la plus courante et la plus facilement interprétable. Une inspection visuelle du graphique des résidus permet aussi de dépister la présence d'hétéroscédasticité.

Normalité

L'option NORMAL de la procédure réalise le test de normalité de Bera-Jarque. Ce test permet de tester la non-normalité conditionnelle des résidus, ou en d'autres mots, de savoir si la distribution des résidus est normale, ce qui est une autre condition de l'analyse de régression linéaire multiple. La probabilité du Chi-Deux ($pr > ChiSq$) doit être supérieure à 0,05 et une analyse visuelle des résidus est également effectuée.

Statistique R-Square

La statistique du R-Carré (R-Square) permet de juger du potentiel de prévision du modèle. Plus le R-Carré se rapproche de 1,00, plus le modèle explique les variations de la variable dépendante et devrait ainsi être en mesure de la prévoir avec justesse.

AIC et test de Student

Les statistiques AIC et du test de Student sont aussi des statistiques utilisées dans les modèles AUTOREG. Ils ont les mêmes fonctions que dans les modèles ARIMA puisqu'ils permettent de comparer les modèles de même série entre eux et de mesurer la signification des paramètres, ici les variables explicatives, dans le modèle. À moins qu'il y ait une justification théorique importante, on garde seulement les variables explicatives significatives et dont le signe s'interprète de façon adéquate.

E. Vérification de la justesse de la prévision

On évalue ultimement la performance de la spécification du modèle grâce à sa capacité de prévision. Tout comme dans l'exercice de modélisation avec ARIMA, le modèle retenu est celui qui présente le plus petit écart entre la valeur réelle et la valeur prévue, et un modèle est retenu seulement si son écart est inférieur à $\pm 5\%$.

2.2.2 Modèles retenus

Le tableau 4 présente les modèles retenus pour chacun des marchés touristiques selon les volumes et les dépenses touristiques.

TABLEAU 4 *Modèles AUTOREG retenus pour chaque variable dépendante*

| Variable dépendante | Modèle retenu | |
|---|--|--|
| | Variables explicatives et « dico » | Correction des résidus selon l'ordre d'autocorrélation |
| Volume des touristes au Québec | | |
| Québécois | <ul style="list-style-type: none"> – Unités louées mensuellement dans les établissements d'hébergement du Québec – AO200103 – LS200201 | Aucun |
| Autres provinces canadiennes | <ul style="list-style-type: none"> – Unités louées mensuellement dans les établissements d'hébergement du Québec | Nlag = 1 |
| Américains | <ul style="list-style-type: none"> – Entrées aux frontières des touristes américains au Québec – Prix de location dans les établissements d'hébergement du Québec – AO200303 | Nlag = (1) |
| Autres pays que les États-Unis | <ul style="list-style-type: none"> – entrées aux frontières des touristes des autres pays que les États-Unis au Québec | Nlag = (4) |
| Dépenses des touristes au Québec | | |
| Québécois | <ul style="list-style-type: none"> – constante – unités louées mensuellement dans les établissements d'hébergement du Québec – prix de location dans les établissements d'hébergement du Québec | Nlag = (4) |
| Autres provinces canadiennes | <ul style="list-style-type: none"> – constante – volume des touristes des autres provinces canadiennes – prix de location dans les établissements d'hébergement du Québec | Aucun |
| Américains | Aucun | Aucun |
| Autres pays que les États-Unis | Aucun | Aucun |

Trois constats majeurs ressortent de ces modèles : les indicateurs touristiques, qui ont une forte corrélation avec les variables dépendantes de volume, auront grandement servi dans les modèles AUTOREG; les variables dépendantes de dépenses sont difficilement modélisables; et il existe une forte corrélation saisonnière des résidus.

2.3 Option Application de Prévisions de séries temporelles

L'interface *Time Series Forecasting System* (TSFS) de SAS, disponible dans la version 9.1 de SAS, offre un ensemble de fonctionnalités qui donnent accès aux principaux outils de prévisions du module SAS/ETS. Elle permet d'explorer et de prévoir avec un support interactif les séries chronologiques en utilisant 42 méthodes de prévisions, au nombre desquelles on compte les méthodes basées sur la moyenne mobile, celles qui recourent au lissage exponentiel et plusieurs modèles ARIMA préétablis. La liste de ces derniers apparaît à l'annexe 2.

Cette interface a permis de tester les techniques qui ne proposent pas la construction d'un modèle complexe de prévision et qui sont reconnues comme efficaces dans la littérature pour la prévision de séries touristiques. Les modèles ARIMA suggérés ont été mis de côté puisqu'ils ont été traités précédemment dans la partie Modèles ARIMA. Les techniques les plus porteuses dans cette perspective reposent sur l'examen de la tendance linéaire et saisonnière et sur les techniques de lissage exponentiel.

Une attention particulière a été portée aux méthodes de lissage pour données saisonnières qui permettent de nuancer le poids des données selon la place qu'elles occupent dans la série chronologique en accordant plus d'importance aux données récentes. Ces méthodes ont également pour « principale qualité d'exiger peu de données sur l'évolution passée de la série temporelle » (Stafford et Sarrasin, 2005). On leur reconnaît aussi une précision certaine pour produire des prévisions sur un grand nombre de séries économiques (Brockwell et Davies, 2002, p. 317).

Ces méthodes posent comme principal défi le choix judicieux et optimal des coefficients ou constantes de lissage α , β et γ . Ces constantes doivent varier entre 0 et 1 et il est, dans cette perspective, recommandé de :

- diminuer le poids des observations récentes lorsque les bruits jouent un rôle important dans le profilage de la série;
- donner plus de poids aux observations récentes lorsque la moyenne de la série change rapidement;
- changer les poids des constantes pour chaque série qui fait l'objet d'une prévision (SAS/ETS User's guide, p. 609).

TSFS a été utilisé sur toutes les variables dépendantes afin de comparer les résultats à ceux des deux modèles utilisés pour produire les prévisions décrites précédemment.

2.3.1 Étapes

La démarche associée au TSFS nécessite les étapes suivantes :

- À partir du menu Application de prévision de séries temporelles, identifier la variable temporelle de référence (id), préciser la date de début de la série en format SAS et déterminer le type d'intervalle (mensuel, trimestriel, etc.) de la série.

- Examiner sa configuration graphique, les graphiques d'autocorrélation et les tests de stationnarité et de bruit blanc afin de déceler la présence d'une tendance, de mouvements saisonniers et de cycles.
- Valider le diagnostic proposé par l'étape 2 avec l'outil de diagnostic de séries du TSFS et transformer au besoin les données de base (LOG, Première différence, Deuxième différence).
- Identifier les modèles et préciser le nombre ou le type de modèles devant être testés en utilisant les options « Model selection list » et « Automatic Fit ». Comme précisé auparavant, les modèles ARIMA ont été exclus de la sélection parce que les deux séries analysées couvraient une période trop courte (ce qui créait de l'instabilité) et que ce type de modèles avait déjà fait l'objet d'une analyse plus fine précédemment.
- Sélectionner les statistiques de mesure de la qualité d'ajustement (*goodness-of-fit statistic*) qui permettront d'évaluer la performance des différents modèles. TSFS utilise par défaut le Root Mean Square Error (RMSE), mais la valeur du Mean Average Percentage Error (MAPE) a été privilégiée, car utilisée dans un grand nombre d'études consultées tout au long de la démarche.
- Utiliser la procédure « Automatic Fit » proposée par l'option « Fit model » de l'onglet Edit.
- La fenêtre Automatic Model Fitting Result ordonne les modèles sélectionnés préalablement selon leur score RMSE et MAPE. Elle fournit également pour chacun d'eux une série de tableaux et de graphiques qui permettent d'analyser plus finement les modèles performants et de les modifier. Ces tableaux et graphiques fournissent des représentations graphiques du modèle de prévisions et d'erreurs, les autocorrélations des erreurs, les tests de stationnarité et de bruits blancs, les paramètres d'estimations, les statistiques de précision, et bien entendu, les prévisions.
- Vérifier la justesse de prévision des modèles qui présentent une MAPE inférieure à 10 %, ce qui révèle un niveau élevé de précision.

2.3.2 Options additionnelles

Il est possible de développer des modèles dans TSFS, ce qui offre la possibilité d'augmenter les possibilités et d'aller plus loin dans les options que ce que permet l'option automatique. Ainsi, on peut ajouter des variables explicatives et « outlier » aux modèles existants. De plus, les programmes produits pour les modèles associés à TSFS peuvent être modifiés afin d'introduire des statistiques de diagnostic ou toute autre option jugée pertinente.

2.3.3 Modèles retenus

L'option TSFS a permis d'identifier plusieurs modèles de prévision acceptables sur la base du MAPE et du RMSE pour chaque variable dépendante. Le tableau 5 illustre les modèles retenus pour chaque variable dépendante.

TABLEAU 5 *Modèles associés à TSFS retenus pour chaque variable dépendante*

| Variable dépendante | Modèle retenu | MAPE |
|---|----------------------------------|-------|
| Volume des touristes au Québec | | |
| Québécois | Winters Method -- Additive | 8,88 |
| Autres provinces canadiennes | Seasonal Exponential Smoothing | 11,74 |
| Américains | Winters Method -- Additive | 4,85 |
| Autres pays que les États-Unis | Winters Method -- Multiplicative | 9,56 |
| Dépenses des touristes au Québec | | |
| Québécois | Seasonal Exponential Smoothing | |
| Autres provinces canadiennes | Aucun | Aucun |
| Américains | Aucun | Aucun |
| Autres pays que les États-Unis | Aucun | Aucun |

On retient que les modèles de lissage exponentiel saisonnier semblent efficaces pour prévoir la demande touristique et que tous les modèles retenus tiennent compte de la saisonnalité, qui est fortement présente dans les variables dépendantes analysées. On remarquera aussi la difficulté de modéliser les variables de dépenses comme dans le cas des modèles ARIMA et AUTOREG.

3. Analyses des résultats

Cette section présente les résultats obtenus pour chaque variable dépendante, selon les modèles ARIMA, AUTOREG et de TSFS, et l'évaluation qui a permis de faire un choix de modèle prévisionnel donnant les prévisions qui s'avèrent les plus justes pour chacune de ces variables.

Quoique les modèles soient construits sur des séries trimestrielles, l'analyse des résultats se fait sur des prévisions annuelles étant donné que les prévisions trimestrielles sont moins précises et que le besoin du MTO était d'obtenir des prévisions annuelles. Ces prévisions annuelles sont tout simplement l'addition des quatre trimestres prévus pour une variable donnée.

3.1 Prévisions par modèle

Les prévisions associées aux modèles sont présentées dans le tableau 6. Les résultats des variables de dépenses y sont présentés, même si en bout de ligne, elles n'ont pas été retenues étant donné que les modèles n'étaient pas concluants.

TABLEAU 6 *Taux de variation de l'année prévue par rapport à l'année précédente (réelle) pour chaque variable dépendante*

| Variable | Modèles retenus | 2004 réel | 2005 prévue | Variation (%) |
|---------------------------------------|---|------------|-------------|---------------|
| Volume des touristes au Québec | | | | |
| Québécois | ARIMA(0,1,0) ₄ avec unités louées mensuellement dans les établissements d'hébergement du Québec | 21 359 000 | 21 367 314 | 0,0 |
| | AUTOREG (unités louées mensuellement dans les établissements d'hébergement du Québec + AO200103 + LS200201) | | 20 933 705 | -2,0 |
| | TSFS - Winters Method -- Additive | | 21 555 063 | 0,9 |
| Autres provinces canadiennes | ARIMA(1,1,0) ₄ LS200101 + AO200303 | 3 643 000 | 3 619 576 | -0,6 |
| | AUTOREG (unités louées mensuellement dans les établissements d'hébergement du Québec) avec nlag = (1) | | 3 552 692 | -2,5 |
| | TSFS - Seasonal Exponential Smoothing | | 3 480 573 | -4,5 |
| Américains | ARIMA(0,1,1) ₄ + entrées aux frontières des touristes américains au Québec + AO200303 | 2 362 800 | 2 231 875 | -5,5 |
| | AUTOREG (entrées aux frontières des touristes américains au Québec + prix de location dans les établissements d'hébergement du Québec + AO200303) avec nlag = (1) | | 2 127 892 | -9,9 |
| | TSFS - Winters Method -- Additive | | 2 422 422 | 2,5 |
| Autres pays que les États-Unis | ARIMA(0,1,0) ₄ + entrées aux frontières des touristes des autres pays que les États-Unis + AO200103 + AO200303 | 978 500 | 1 066 636 | 9,0 |
| | AUTOREG (entrées aux frontières des touristes des autres pays que les États-Unis au Québec) avec nlag = (4) | | 1 133 166 | 15,8 |
| | TSFS - Winters Method -- Multiplicative | | 975 770 | -0,3 |

| Variable | Modèles retenus | 2004 réel | 2005 prévue | Variation (%) |
|---|--|---------------|---------------|---------------|
| Dépenses des touristes au Québec | | | | |
| Québécois | ARIMA(0,1,0) ₄ + unités louées mensuellement dans les établissements d'hébergement du Québec + prix de location dans les établissements d'hébergement du Québec) | 3 666 321 000 | 3 869 336 770 | 5,5 |
| | AUTOREG (constante + unités louées mensuellement dans les établissements d'hébergement du Québec + prix de location dans les établissements d'hébergement du Québec) avec nlag = (4) | | 3 819 741 983 | 4,2 |
| | TSFS - Seasonal Exponential Smoothing | | 3 637 114 929 | 4,6 |
| Autres provinces canadiennes | AUTOREG (constante + volume des touristes des autres provinces canadiennes + prix de location dans les établissements d'hébergement du Québec) | 1 218 771 000 | 1 253 493 500 | 2,8 |
| Américains | ARIMA(0,1,1) ₄ constante + volume de touristes américains au Québec + ao200303 | 1 386 304 500 | 1 290 120 198 | -6,9 |
| Autres pays que les États-Unis | ARIMA(0,1,0) ₄ constante + volume de touristes des autres pays que les États-Unis au Québec | 958 629 200 | 1 103 362 512 | 15,1 |

Les modèles ont donné des prévisions allant la majorité du temps dans le même sens, et avec des taux de variations passablement similaires, ce qui permet d'une certaine manière de valider la démarche de modélisation. Afin de déterminer quels modèles sont les plus performants, il a fallu faire appel à certains tests qui sont décrits au point suivant.

3.2 Évaluation

Afin de déterminer lequel des modèles sera finalement choisi pour chaque variable dépendante, trois critères de comparaison ont été utilisés :

- un test de mesure d'amplitude des erreurs;
- l'écart entre la donnée réelle et prévue de la dernière année disponible;
- la vraisemblance des prévisions.

Test de mesure d'amplitude des erreurs

Le test utilisé est l'écart moyen en valeur absolue en pourcentage ou MAPE (*Mean Absolute Percent Error*), un test que recommande Douglas C Frechtling dans son livre *Forecasting tourism demand - methods and strategies*. Il calcule la somme des erreurs absolues pour chaque période de temps divisée par la valeur actuelle de la période et cette somme est divisée par le nombre de périodes pour obtenir la valeur moyenne (p. 26). Plus le MAPE est petit, plus le modèle produit des petits pourcentages d'erreur. L'interprétation du MAPE se fait, selon Frechtling, de la façon suivante :

- <10 excellent
- 10-20 bon
- 20-50 raisonnable
- >50 mauvais

Écart absolu entre la donnée réelle et prévue de la dernière année disponible

L'écart absolu entre la donnée réelle et prévue de la dernière année disponible est l'élément clé du choix d'un modèle puisqu'il montre avec quelle précision il est capable de prévoir les valeurs futures d'une série. Dans la modélisation, un écart inférieur à 5 % est considéré comme excellent.

Vraisemblance des prévisions

La vraisemblance des prévisions consiste à regarder si celles-ci ont du sens par rapport à l'information qualitative et quantitative disponible sur la période de référence. Il s'agit en quelque sorte de mesurer le degré de confiance qu'on attribue à la prévision. Un modèle peut effectivement être statistiquement très performant, mais ses prévisions peuvent ne pas être représentatives de la réalité, comme un modèle peut être statistiquement moins performant, mais très bien prévoir ce qui s'est vécu sur le terrain.

Le tableau 7 permet de comparer les modèles retenus pour les variables dépendantes à partir de ces trois éléments.

TABEAU 7 *Évaluation des modèles retenus pour chaque variable dépendante*

| Variable | Modèles retenus | MAPE | Écart absolu 2004 (%) | Vraisemblance |
|---------------------------------------|--|-------|-----------------------|---------------|
| Volume des touristes au Québec | | | | |
| Québécois | ARIMA (0,1,0) ₄ avec unités louées mensuellement dans les établissements d'hébergement du Québec | 9,84 | 3,6 | Oui |
| | AUTOREG (unités louées mensuellement dans les établissements d'hébergement du Québec + AO200103 +LS200201) | 8,79 | 2,0 | Oui |
| | TSFS - Winters Method -- Additive | 8,88 | 3,0 | Non |
| Autres provinces canadiennes | ARIMA (1,1,0) ₄ LS200101 + AO200303 | 11,05 | 4,8 | Oui |
| | AUTOREG (unités louées mensuellement dans les établissements d'hébergement du Québec) avec nlag = (1) | 11,15 | 0,1 | Oui |
| | TSFS - Seasonal Exponential Smoothing | 11,74 | 0,9 | Oui |
| Américains | ARIMA (0,1,1) ₄ + entrées aux frontières des touristes américains au Québec + AO200303 | 2,42 | 0,1 | Oui |
| | AUTOREG (entrées aux frontières des touristes américains au Québec + prix de location dans les établissements d'hébergement du Québec + AO200303) avec nlag = (1) | 2,43 | 1,4 | Oui |
| | TSFS - Winters Method -- Additive | 4,85 | 1,7 | Non |
| Autres pays que les États-Unis | ARIMA (0,1,0) ₄ + entrées aux frontières des touristes des autres pays que les États-Unis + AO200103 + AO200303 | 8,77 | 0,8 | Oui |
| | AUTOREG (entrées aux frontières des touristes des autres pays que les États-Unis au Québec) avec nlag = (4) | 8,35 | 1,8 | Oui |
| | TSFS - Winters Method -- Multiplicative | 9,56 | 7,8 | Non |

| Variable | Modèles retenus | MAPE | Écart absolu 2004 (%) | Vraisemblance |
|---|---|-------|-----------------------|---------------|
| Dépenses des touristes au Québec | | | | |
| Québécois | ARIMA (0,1,0) ₄ + unités louées mensuellement dans les établissements d'hébergement du Québec + prix de location dans les établissements d'hébergement du Québec) | 10,78 | 3,6 | Oui |
| | AUTOREG (constante + unités louées mensuellement dans les établissements d'hébergement du Québec + prix de location dans les établissements d'hébergement du Québec) avec nlag = (4) | 10,84 | 1,5 | Oui |
| | TSFS - Seasonal Exponential Smoothing | 10,12 | 1,2 | Oui |
| Autres provinces canadiennes | AUTOREG (constante + volume des touristes des autres provinces canadiennes + prix de location dans les établissements d'hébergement du Québec) | 13,90 | 0,32 | Non |
| Américains | ARIMA (0,1,1) ₄ constante + volume de touristes américains au Québec + ao200303 | 6,32 | 0,3 | Oui |
| Autres pays que les États-Unis | ARIMA (0,1,0) ₄ constante + volume de touristes des autres pays que les États-Unis au Québec | 6,31 | 5,2 | Oui |

3.3 Choix

Les choix de modèle pour les variables dépendantes, basés sur les éléments de comparaison retrouvés dans le tableau précédent, sont donc :

| Volume des touristes au Québec | |
|---|---|
| Québécois | AUTOREG (avec unités louées mensuellement dans les établissements d'hébergement du Québec + AO200103 + LS200201) |
| Autres provinces canadiennes | ARIMA (1,1,0) ₄ LS200101 + AO200303 |
| Américains | AUTOREG (entrées aux frontières des touristes américains au Québec + prix de location dans les établissements d'hébergement du Québec) avec nlag = (4) |
| Autres pays que les États-Unis | AUTOREG (entrées aux frontières des touristes des autres pays que les États-Unis au Québec + AO200303) avec nlag = (4) |
| Dépenses des touristes au Québec | |
| Québécois | TSFS - Seasonal Exponential Smoothing |
| Autres provinces canadiennes | AUTOREG (constante + volume des touristes des autres provinces canadiennes + prix de location dans les établissements d'hébergement du Québec) |
| Américains | ARIMA (0,1,1) ₄ avec constante + volume de touristes américains au Québec |
| Autres pays que les États-Unis | ARIMA (0,1,0) ₄ avec constante + volume de touristes des autres pays que les États-Unis au Québec |

Les trois approches semblent offrir de bonnes performances prévisionnelles pour les volumes et les dépenses des touristes, sauf dans le cas des dépenses des touristes des autres provinces canadiennes au Québec dont la série est instable.

4. Conclusion

Le présent rapport a servi à exposer la démarche de modélisation entreprise par le MTO afin de répondre à la demande qui était de fournir des estimations et des prévisions pour le volume de touristes et leurs dépenses pour les trimestres où Statistique Canada a révisé l'Enquête sur les voyages des Canadiens (EVC). Le modèle prévisionnel touristique développé à cette fin permet aussi de répondre aux différentes demandes *ad hoc* sur la performance touristique que la Direction de la recherche et de la prospective du MTO reçoit de façon régulière.

Trois approches de modélisation ont été priorisées afin d'y parvenir et ont fait l'objet d'une analyse détaillée, soit les modèles ARIMA, les modèles AUTOREG et l'Option « Application de prévisions de séries temporelles ». Ils fournissent chacun de bons outils de base pour prévoir et modéliser les différents marchés touristiques et donnent de bons résultats, sauf pour certaines variables dépendantes de dépenses. Les choix de modèles pour les variables dépendantes se sont principalement arrêtés sur le modèle AUTOREG.

4.1 Pousser plus loin la démarche de prévision

Dans ce rapport, nous avons mentionné quelques pistes de développement. Cependant, en fonction des questions à éclaircir, des résultats à venir et des développements scientifiques relatifs à l'estimation économétrique et à la modélisation de marchés touristiques, il pourrait y avoir d'autres avenues à examiner ou à tester. Le domaine de la prévision et de la modélisation est extrêmement vaste. Voici quelques pistes de recherche qui pourraient être privilégiées pour bonifier et améliorer le modèle prévisionnel touristique.

- Compléter la sélection des variables

Certaines variables explicatives pourraient être ajoutées pour mieux cerner la réalité sous-jacente aux différents marchés. Les plus intéressantes, dans cette perspective, sont celles qui expriment l'évolution des prix dans les destinations compétitives du Québec et celles qui quantifient les interventions en marketing du MTO et de ses principaux partenaires. Il serait approprié d'utiliser la typologie proposée par Frechtling (2001) pour regrouper les variables et s'assurer de bien couvrir tous les effets appréhendés sur la demande touristique. Par la suite, une épuration devrait être effectuée afin de conserver seulement les variables qui ont ou auront un rôle certain à jouer dans la modélisation de l'activité touristique.

- Utiliser d'autres options dans les procédures

Pour les procédures ARIMA, AUTOREG et du TSFS, d'autres options pourraient être utilisées afin de compléter les diagnostics des séries. Par exemple, le Test de Chow dans la procédure AUTOREG pourrait être ajouté pour vérifier s'il y a un changement structurel dans une série. Pour le moment, le manque d'observations nous empêche de le faire.

- Recourir à des méthodes plus complexes pour établir les liens entre les variables

Les différents marchés ont été estimés de façon séparée par moindres carrés ordinaires, jusqu'à maintenant. Du point de vue de la modélisation, on peut subodorer qu'il existe en fait des liens entre ces différents marchés et que certaines variables dépendantes peuvent devenir les variables explicatives d'une autre variable dépendante. Dans ce cas, il faut recourir à des méthodes plus complexes, tels les moindres carrés à deux étapes ou des méthodes par maximum de vraisemblance, qui tiennent compte de la simultanéité et de la cointégration dans les équations.

- Essayer le modèle à équations multiples

Il faut aussi analyser de façon plus complète la structure des délais qui permet d'établir des liens entre les variables. SAS supporte une procédure, soit la procédure VARMAX, qui a été utilisée récemment pour la modélisation des marchés touristiques²⁴ et qui semble offrir plusieurs avantages pour solutionner le type de problèmes auxquels nous avons fait face. De plus, elle est construite de façon à générer facilement des prévisions. Elle est en quelque sorte la combinaison d'ARIMA et d'AUTOREG et offre de plus la possibilité de générer des prévisions pour les variables dépendantes, sans qu'il ne soit nécessaire d'inclure des valeurs futures pour les variables explicatives. Elle permet aussi d'effectuer des analyses plus poussées des relations entre les variables et la structure des délais pour modéliser les interactions. Malheureusement, le nombre d'observations dont nous disposons à l'heure actuelle pour certaines variables dépendantes n'est pas suffisant pour s'en servir. On pourrait toutefois l'utiliser en modélisant le volume et les dépenses des touristes américains au Québec, le volume et les dépenses des touristes des autres pays que les États-Unis au Québec ou des variables touristiques fortement corrélées aux volumes et aux dépenses touristiques, puisque nous avons plus d'observations pour ces variables. Les entrées aux frontières, les unités louées dans les établissements d'hébergement et certaines séries tirées des indicateurs nationaux du tourisme sont, à l'évidence, les variables qui offrent le plus de potentiel.

D'autres méthodes causales pourraient aussi faire l'objet d'une lecture plus approfondie afin de voir s'il serait justifié de les essayer et si elles s'avèrent plus efficaces que les modèles à équation unique (AUTOREG).

- Faire appel à un panel d'experts de l'industrie

Dans les modèles prévisionnels, on fait souvent appel à un panel d'experts afin de compléter l'analyse et la validation des prévisions. Cette approche, de type qualitative, fait référence à la méthode Delphi dans la littérature. Elle vient apporter des éléments de précision que les méthodes statistiques n'arrivent pas à considérer dans leurs calculs.

- Utiliser les données annuelles au lieu des données trimestrielles

Étant donné que les prévisions trimestrielles sont moins précises que les prévisions annuelles, il pourrait s'avérer efficace d'utiliser des données annuelles dans le modèle prévisionnel touristique,

24. SONG, H., and S. F. WITT. *Tourism Demand Modeling and Forecasting: Modern Econometric Approaches*, Oxford, U.K.: Elsevier Science Ltd., 2000 et en ligne : <http://support.sas.com/rnd/app/examples/ets/tourism/index.htm>.

ce qui simplifierait d'autant plus la modélisation. Dans cette éventualité, si l'analyse ne se fait plus à partir des prévisions trimestrielles, il est possible de développer un indice saisonnier, à partir de la procédure X12 dans SAS, qui servirait à « trimestrialiser » les prévisions annuelles. Cet indice se calculerait à partir des séries trimestrielles dans lesquelles on retrouve, comme nous avons pu le constater dans ce rapport, une forte saisonnalité.

- Modéliser les équations d'offre et de demande

Par ailleurs, pour chacun des marchés, l'équation estimée est une équation mixte entre l'équation d'offre et l'équation de demande. La quantité observée est le fait de l'équilibre réalisé entre la quantité offerte et la quantité demandée. En principe, d'un point de vue théorique, on doit essayer de modéliser une équation d'offre et une équation de demande. Une fonction qui équilibre la quantité offerte et la quantité demandée joint les deux équations. Il n'est toutefois pas facile de trouver les variables explicatives adéquates afin d'arriver à identifier chacune des équations. C'est pourquoi, en pratique, cet effort souhaitable de modélisation est difficile à réaliser.

- Tester d'autres approches pour les variables dépendantes de dépenses touristiques

La modélisation des dépenses touristiques s'avérant plus ardue, du fait que ces variables sont généralement moins fiables et plus difficilement mesurables, il pourrait être intéressant d'essayer la méthode utilisée par le ministère du Tourisme de la Nouvelle-Zélande. Pour mesurer les dépenses totales, ils utilisent une combinaison des prévisions des nuitées avec celles des dépenses moyennes par nuit.

4.2 Recommandations

À la lumière de cet exercice de modélisation, des résultats et des démarches additionnelles suggérées précédemment, voici les recommandations qu'il serait important de réaliser dans les travaux futurs :

- **Compléter la sélection des variables** : en ajoutant, si elles sont disponibles, les variables de prix des destinations concurrentes et l'investissement promotionnel du MTO; en regroupant les variables selon la typologie de Frechtling; et en faisant une épuration des variables explicatives en conservant seulement celles qui ont un rôle certain à jouer dans la modélisation des marchés touristiques;
- **Développer d'autres modèles pour les variables dépendantes de dépenses touristiques** en essayant la méthode du ministère du Tourisme de la Nouvelle-Zélande qui utilise une combinaison des prévisions des nuitées avec celles des dépenses moyennes par nuit;
- **Essayer les modèles VARMAX** pour les variables dépendantes de volume et de dépenses des touristes américains et des autres pays que les États-Unis au Québec;
- **Faire appel à un panel d'experts de l'industrie** pour valider les prévisions.
- Nous croyons toutefois que les modèles actuels obtenus à l'aide des trois approches adoptées peuvent être utilisés pour répondre aux besoins à court terme d'estimation du volume et des dépenses touristiques par marché.

5. Annexes

5.1 Annexe 1 : Listes des variables utilisées pour le modèle prévisionnel touristique

| Variable dépendante | Nom | Début | Fin | Transformation |
|--|-----------------|--------|---------|----------------|
| Dépenses des touristes américains au Québec (\$) | deptourusaque | 1990-I | 2005-IV | - |
| Dépenses des touristes des autres pays que les États-Unis au Québec (\$) | deptourapaysque | 1990-I | 2005-IV | - |
| Dépenses des touristes des autres provinces canadiennes au Québec (\$) | deptouraprovque | 1990-I | 2004-IV | - |
| Dépenses des touristes québécois au Québec (\$) | deptourqueque | 1990-I | 2004-IV | - |
| Volume de touristes américains au Québec | voltourusaque | 1990-I | 2005-IV | Lag 1 à 4 |
| Volume de touristes des autres pays que les États-Unis au Québec | voltourapaysque | 1990-I | 2005-IV | Lag 1 à 4 |
| Volume de touristes des autres provinces canadiennes au Québec | voltouraprovque | 1990-I | 2004-IV | Lag 1 à 4 |
| Volume de touristes québécois au Québec | voltourqueque | 1990-I | 2004-IV | Lag 1 à 4 |

| Variable indépendante | Nom | Début | Fin | Transformation |
|--|-----------------|--------|---------|----------------|
| Balance au compte des voyages internationaux (recettes moins paiements), Canada (M \$) | balancecan | 1990-I | 2005-IV | - |
| Demande touristique intérieure, Canada (M \$) | demtourdomcan | 1990-I | 2005-IV | - |
| Demande touristique totale, Canada (M \$) | demtourtotcan | 1990-I | 2005-IV | - |
| Dépenses au Canada par les résidents étrangers (recettes) (M \$) | deptourintcan | 1990-I | 2005-IV | - |
| Dépenses au point d'origine des Québécois au Québec pour des voyages à l'extérieur du Québec (\$) | deporigque | 1998-I | 2004-IV | - |
| Dépenses dans des pays étrangers par les Canadiens (paiements) (M \$) | deptourcanint | 1990-I | 2005-IV | - |
| Dépenses des excursionnistes totales au Québec (\$) | depexctotque | 2000-I | 2004-IV | - |
| Dépenses des touristes québécois revenant de l'international (\$) | deptourqueint | 1995-I | 2004-IV | - |
| Dépenses des touristes québécois revenant des États-Unis (\$) | deptourqueusa | 1995-I | 2004-IV | - |
| Dépenses des touristes québécois revenant des autres pays que les États-Unis (\$) | deptourqueapays | 1995-I | 2004-IV | - |
| Dépenses personnelles en biens et services de consommation (\$ courants) - loisirs, divertissement, formation et culture - Québec (M \$) | deppersldfcque | 1990-I | 2005-IV | - |
| Dépenses personnelles en biens et services de consommation (\$ courants) - loisirs, divertissement, formation et culture - Canada (M \$) | deppersldfccan | 1990-I | 2005-IV | - |
| Dépenses personnelles en biens et services de consommation (\$ courants) - restaurants et hébergement - Canada (M \$) | deppersrhcan | 1990-I | 2005-IV | - |
| Dépenses personnelles en biens et services de consommation (\$ courants) - total - Canada (M \$) | depperstotcan | 1990-I | 2005-IV | - |
| Dépenses personnelles en biens et services de consommation (\$ courants) - total - Québec (M \$) | depperstotque | 1990-I | 2005-IV | - |
| Emploi - ensemble des industries - Québec (k) | emploitotque | 1990-I | 2005-IV | - |
| Emploi - hébergement et restauration - Québec (k) | emploihrique | 1990-I | 2005-IV | - |

| Variable indépendante | Nom | Début | Fin | Transformation |
|--|-------------------|--------|---------|----------------|
| Emploi - services de restauration et débits de boissons - Québec (k) | emploisrdbque | 1990-I | 2005-IV | - |
| Emploi - services d'hébergement - Québec (k) | emploishque | 1990-I | 2005-IV | - |
| Entrées aux frontières des touristes américains au Canada | eftourusacan | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes américains au Québec | eftourusaque | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes américains en Ontario | eftourusaont | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes canadiens revenant des autres pays que les États-Unis | eftourcanrevapays | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes canadiens revenant des États-Unis | eftourcanrevusa | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes canadiens revenant des pays internationaux | eftourcanrevint | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes des autres pays que les États-Unis au Canada | eftourapayscan | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes des autres pays que les États-Unis au Québec | eftourapaysque | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes des autres pays que les États-Unis en Ontario | eftourapaysont | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes internationaux au Canada | eftourintcan | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes internationaux au Québec | eftourintque | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes internationaux en Ontario | eftourintont | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes ontariens revenant des autres pays que les États-Unis | eftourontrevapays | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes ontariens revenant des États-Unis | eftourontrevusa | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes ontariens revenant des pays internationaux | eftourontrevint | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes québécois revenant des autres pays que les États-Unis | eftourquerevapays | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes québécois revenant des États-Unis | eftourquerevusa | 1991-I | 2005-IV | - |
| Entrées aux frontières des touristes québécois revenant des pays internationaux | eftourquerevint | 1991-I | 2005-IV | - |
| Exportations touristiques, Canada (M \$) | exptourtotcan | 1990-I | 2005-IV | - |
| Indicateurs économiques avancés, Canada (1992=100) | indavcan | 1990-I | 2005-IV | Lag 2 et 3 |
| Indicateurs économiques avancés, États-Unis (1992=100) | indavusa | 1990-I | 2005-IV | Lag 2 et 3 |
| Indice de confiance des consommateurs, Allemagne (2000?=0) | indconfconsall | 1990-I | 2005-IV | Lag 1 |
| Indice de confiance des consommateurs, Canada (1991=100) | indconfconscan | 1990-I | 2005-IV | Lag 1 |
| Indice de confiance des consommateurs, États-Unis (1991?=100) | indconfconsusa | 1990-I | 2005-IV | Lag 1 |
| Indice de confiance des consommateurs, Europe (2000?=0) | indconfconseuro | 1990-I | 2005-IV | Lag 1 |
| Indice de confiance des consommateurs, France (2000?=0) | indconfconsfr | 1990-I | 2005-IV | Lag 1 |
| Indice de confiance des consommateurs, Québec (1991=100) | indconfconsque | 1990-I | 2005-IV | Lag 1 |
| Indice de confiance des consommateurs, Royaume-Uni (2000?=0) | indconfconsuk | 1990-I | 2005-IV | Lag 1 |
| Indice des prix à la consommation - Ensemble - Canada (1992=100) | indprixconstotcan | 1990-I | 2005-IV | Lag 1 à 4 |
| Indice des prix à la consommation - Ensemble - Québec (1992=100) | indprixconstotque | 1990-I | 2005-IV | Lag 1 à 4 |

| Variable indépendante | Nom | Début | Fin | Transformation |
|---|-------------------|--------|---------|----------------|
| Indice des prix à la consommation - Hébergement pour voyageurs - Canada (1992=100) | indprixconshvcan | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Hébergement pour voyageurs - Québec (1992=100) | indprixconshvque | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Services de voyage - Canada (1992=100) | indprixconssvcan | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Services de voyage - Québec (1992=100) | indprixconssvque | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Transport interurbain - Canada (1992=100) | indprixconstican | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Transport interurbain - Québec (1992=100) | indprixconstique | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Utilisation de véhicules de loisirs - Canada (1992=100) | indprixconsuvlcan | 1990-I | 2005-IV | - |
| Indice des prix à la consommation - Voyages organisés - Canada (1992=100) | indprixconsvocan | 1990-I | 2005-IV | - |
| Indice des prix de voyage, Canada (1992=100) | indprixvoycan | 1990-I | 2005-IV | Lag 1 à 4 |
| Indice précurseur Desjardins, Québec (2003=100) | indpredesjqque | 1990-I | 2005-IV | Lag 1 et 2 |
| Indices implicites de prix - autres biens et services touristiques - Canada (1997=100) | indimpprixabstcan | 1990-I | 2005-IV | - |
| Indices implicites de prix - biens et services touristiques - Canada (1997=100) | indimpprixbstcan | 1990-I | 2005-IV | - |
| Indices implicites de prix - dépenses touristiques - Canada (1997=100) | indimpprixdtcan | 1990-I | 2005-IV | Lag 1 à 4 |
| Indices implicites de prix - services de restauration - Canada (1997=100) | indimpprixsrcan | 1990-I | 2005-IV | - |
| Indices implicites de prix - services de transport - Canada (1997=100) | indimpprixstcan | 1990-I | 2005-IV | - |
| Indices implicites de prix - services d'hébergement - Canada (1997=100) | indimpprixshcan | 1990-I | 2005-IV | - |
| Nuitées des touristes américains au Québec | nuitusaque | 1990-I | 2005-IV | - |
| Nuitées des touristes des autres pays que les États-Unis au Québec | nuitapaysque | 1990-I | 2005-IV | - |
| Nuitées des touristes des autres provinces canadiennes au Québec | nuitaprovque | 1990-I | 2004-IV | - |
| Nuitées des touristes québécois au Québec | nuitqueque | 1990-I | 2004-IV | - |
| Prix de location dans les établissements d'hébergement du Québec (\$) | ehprixlocque | 1995-I | 2005-IV | - |
| Prix du pétrole brut - WTI-Cushing, Oklahoma - USA (\$US/baril) | prixpetroleusa | 1990-I | | Lag 1 à 4 |
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), Allemagne (M euros) | piball | 1991-I | 2005-IV | - |
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), Canada (M \$) | pibcan | 1990-I | 2005-IV | - |
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), États-Unis (G \$) | pibusa | 1990-I | 2004-IV | - |
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), Euro15 (M euros) | pibeuro15 | 1995-I | 2005-IV | - |
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), France (M euros) | pibfr | 1990-I | 2005-IV | - |
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), Québec (M \$) | pibque | 1990-I | 2005-IV | - |

| Variable indépendante | Nom | Début | Fin | Transformation |
|---|----------------|--------|---------|----------------|
| Produit intérieur brut aux prix du marché, selon les revenus (\$ courants), Royaume-Uni (M euros) | pibuk | 1990-I | 2005-IV | - |
| Recettes des restaurants, traiteurs et tavernes, Québec (k \$) | recrttque | 1990-I | 2005-IV | - |
| Recettes touristiques du Québec (\$) | rectourque | 2000-I | 2004-IV | - |
| Revenu des établissements d'hébergement, Québec (\$) | ehrevenueque | 1995-I | 2005-IV | - |
| Revenu personnel disponible, Allemagne (M \$) | revpersdispall | 1991-I | 2005-IV | - |
| Revenu personnel disponible, Canada (M \$) | revpersdispcan | 1990-I | 2005-IV | - |
| Salaire, États-Unis (\$US) | salaireusa | 1990-I | 2005-IV | - |
| Revenu personnel disponible, France (M \$) | revpersdispfr | 1990-I | 2005-IV | - |
| Revenu personnel disponible, Québec (M \$) | revpersdispque | 1990-I | 2005-IV | - |
| Revenu personnel disponible, Royaume-Uni (M \$) | revpersdispuk | 1990-I | 2005-IV | - |
| Taux de change de la livre sterling (taux à midi), Royaume-Uni (\$) | tauxchangeuk | 1990-I | 2005-IV | Lag 1 à 4 |
| Taux de change de l'Euro (taux à midi), Euro15 (\$) | tauxchangeeuro | | 2005-IV | Lag 1 à 4 |
| Taux de change du dollar américain (taux à midi), États-Unis (\$) | tauxchangeusa | 1990-I | 2005-IV | Lag 1 à 4 |
| Taux d'occupation dans les établissements d'hébergement du Québec (%) | ehtauxocque | 1990-I | 2005-IV | - |
| Unités disponibles mensuellement dans les établissements d'hébergement du Québec | ehunitdispque | 1993-I | 2005-IV | - |
| Unités louées mensuellement dans les établissements d'hébergement du Québec | ehunitlouque | 1993-I | 2005-IV | - |
| Volume d'excursionnistes total au Québec | volexctotque | 2000-I | 2004-IV | - |

| Variable dichotomique | Nom | Commentaire |
|---|----------|--|
| Choc au 3 ^e trimestre 2001 sur la série des variables « voltourqueque » et « deptourqueque » | AO200103 | Attentats du 11 septembre 2001 sur les deux tours du World Trade Center |
| Changement temporaire du 1 ^{er} au 4 ^e trimestre 2002 sur la série des variables « voltourusaque » et « deptourusaque » | TC200201 | |
| Choc au 4 ^e trimestre 2001 sur la série des variables « voltourapaysque » et « deptourapaysque » | AO200104 | |
| Changement structurel à partir du 1 ^{er} trimestre 2002 sur la série des variables « voltourqueque » et « deptourqueque » | LS200201 | Changement méthodologique de l'EVC en 2002 |
| Choc au 3 ^e trimestre 2003 sur la série des variables « voltourqueque », « voltourapaysque », « deptourqueque », « deptourapaysque » | AO200303 | Infections au syndrome respiratoire aigu sévère (SRAS) et guerre en Irak |
| Changement structurel à partir du 1 ^{er} trimestre 2003 sur la série de la variable « voltouraprovque » | LS200301 | |
| Changement temporaire du 2 ^e au 3 ^e trimestre 2003 sur la série des variables « voltourusaque » et « deptourusaque » | TC200302 | |
| Changement temporaire du 1 ^{er} au 4 ^e trimestre 2000 sur la série de la variable « voltouraprovque » | TC200001 | Changements ou chocs inexplicables |
| Choc au 3 ^e trimestre 1998 sur la série de la variable « voltouraprovque » | AO199803 | |
| Changement structurel à partir du 3 ^e trimestre 1998 sur la série des variables « voltourusaque » et « deptourusaque » | LS199803 | |

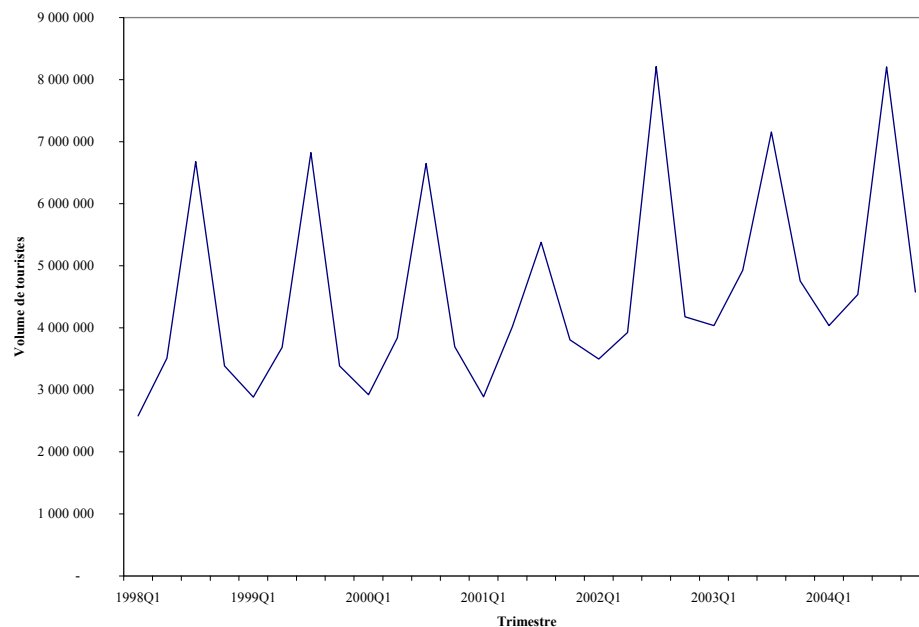
| Variable dichotomique | Nom | Commentaire |
|---|-----------|---------------------------------|
| Choc au 3 ^e trimestre 1997 sur la série des variables « voltourapaysque » et « deptourapaysque » | AO199703 | |
| Printemps | Printemps | Tenir compte de la saisonnalité |
| Été | Ete | |
| Automne | Automne | |
| Hiver | Hiver | |

| Autres variables | Nom | Commentaire |
|-------------------|----------|--|
| Tendance linéaire | Tendance | Tenir compte de la tendance dans la modélisation des variables dépendantes |

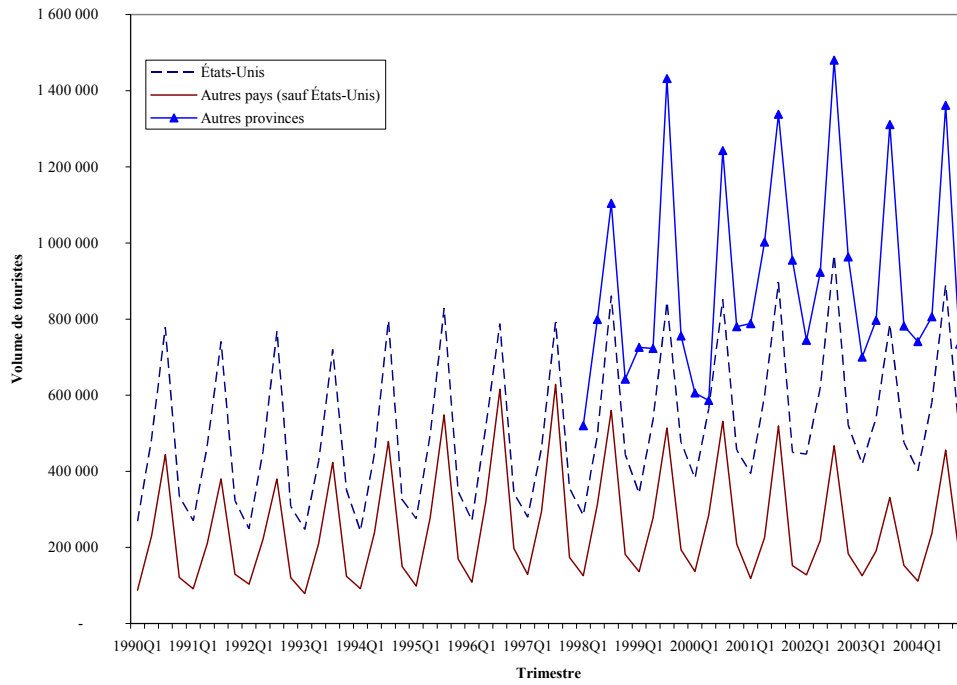
Sources : Statistique Canada, Conference Board of Canada, Institut de la statistique du Québec, ministère du Tourisme, SAS 9.1, Banque du Canada, Desjardins, U.S. Energy Information Administration, Bureau of Economic Analysis.

5.2 Annexe 2 : Présentation visuelle des séries chronologiques assignées comme variables dépendantes.

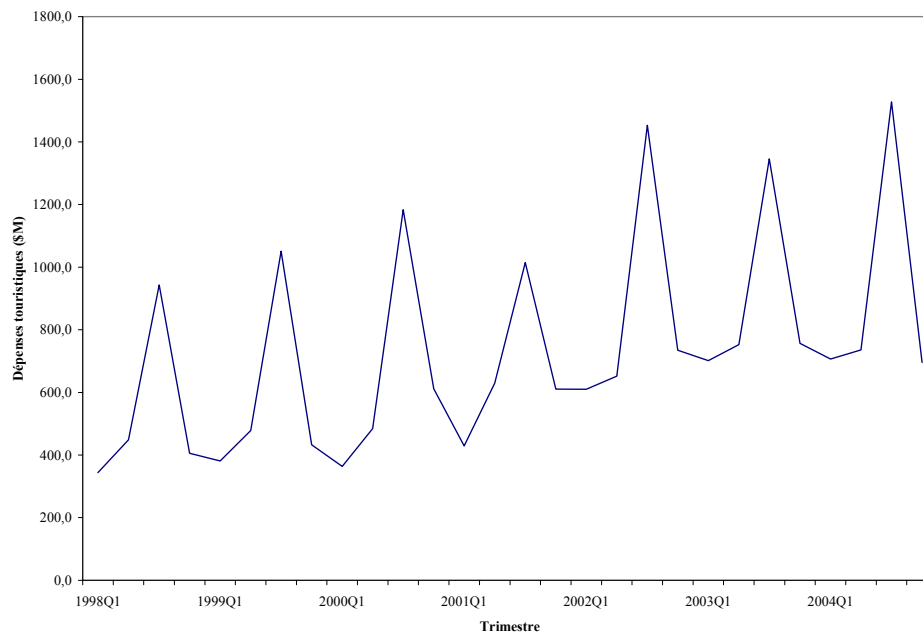
A) Volumes touristiques des Québécois au Québec



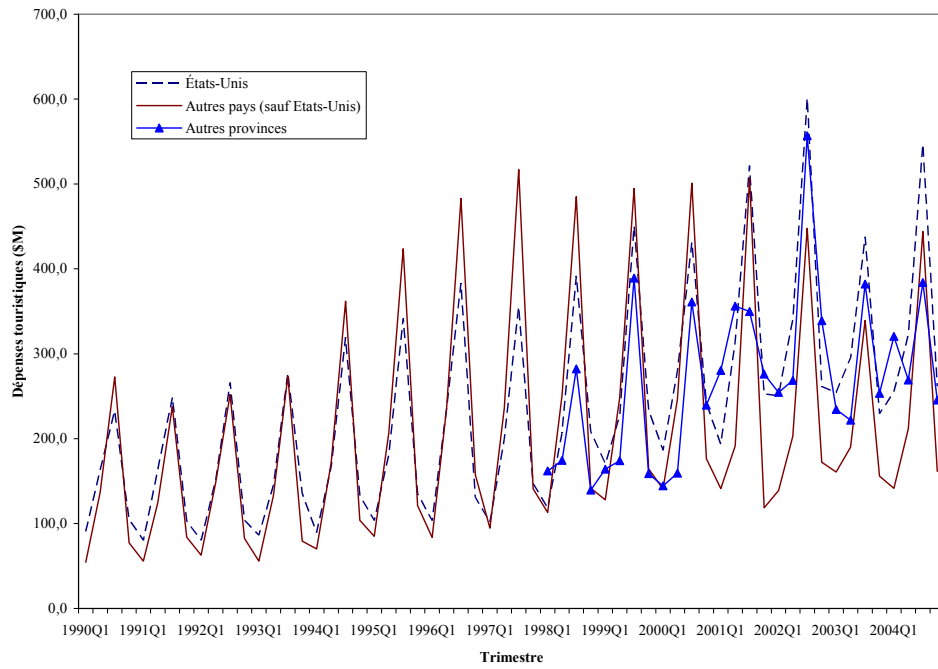
b) Volumes touristiques provenant des États-Unis et des autres pays au Québec



c) Dépenses touristiques des Québécois au Québec



d) Dépenses touristiques provenant des États-Unis et autres pays au Québec



5.3 Annexe 3 : Liste des modèles préétablis du TSFS (ceux utilisés sont en caractère gras)

1. Mean
2. Linear Trend
3. Linear Trend with Autoregressive Errors
4. Linear Trend with Seasonal Terms
5. Seasonal Dummy
6. Simple Exponential Smoothing
7. Double (Brown) Exponential Smoothing
8. Linear (Holt) Exponential Smoothing
9. Damped Trend Exponential Smoothing
10. Seasonal Exponential Smoothing
11. Winters Method –Additive
12. Winters Method --Multiplicative
13. Random Walk with Drift
14. Airline Model
15. ARIMA(0,1,1)_s NOINT
16. ARIMA(0,1,1)(1,0,0)_s NOINT
17. ARIMA(2,0,0)(1,0,0)_s
18. ARIMA(0,1,2)(0,1,1)_s NOINT
19. ARIMA(2,1,0)(0,1,1)_s NOINT
20. ARIMA(0,2,2)(0,1,1)_s NOINT
21. ARIMA(2,1,2)(0,1,1)_s NOINT
22. Log Mean
23. Log Linear Trend
24. Log Linear Trend with Autoregressive Errors
25. Log Linear Trend with Seasonal Terms
26. Log Seasonal Dummy
27. Log Simple Exponential Smoothing
28. Log Double (Brown) Exponential Smoothing
29. Log Linear (Holt) Exponential Smoothing
30. Log Damped Trend Exponential Smoothing
31. Log Seasonal Exponential Smoothing
32. Log Winters Method –Additive
33. Log Winters Method --Multiplicative
34. Log Random Walk with Drift
35. Log Airline Model
36. Log ARIMA(0,1,1)_s NOINT
37. Log ARIMA(0,1,1)(1,0,0)_s NOINT
38. Log ARIMA(2,0,0)(1,0,0)_s
39. Log ARIMA(0,1,2)(0,1,1)_s NOINT
40. Log ARIMA(2,1,0)(0,1,1)_s NOINT
41. Log ARIMA(0,2,2)(0,1,1)_s NOINT
42. Log ARIMA(2,1,2)(0,1,1)_s NOINT

6. Bibliographie

- CHARPENTIER, Arthur. *Séries temporelles: Théorie et applications*, Université Paris Dauphine, 2002, 182 p.
- BILODEAU Danielle. *La désaisonnalisation : pourquoi, quand et comment ?*, L'Écostat, Institut de la Statistique du Québec, décembre 1997, p. 9.
- CANADIAN TOURISM RESEARCH INSTITUTE. *The impact of high gas prices on Canada's tourism industry*, Members Briefing, may 2004, 2 p.
- TIDSWELL, Carmen, Trevor MULES, and Bill FAULKNER. *An Integrative Approach to Tourism Forecasting: A Glance in the Rearview Mirror*, Journal of Travel Research, vol. 40, november 2001, 162-171 p.
- Chris Cooper, and others. *Tourism: principles and practice*, Pearson Education, third edition, 2005.
- MARTIN, Christine A., and Stephen F. WITT. *Accuracy of econometric forecasts of tourism*, *Annals of Tourism Research*, vol. 16, Nb. 3, 1989, 407-428 p.
- LIM, Christine, and Michael MCALEER. *Forecasting Tourist Arrivals*, *Annals of Tourism Research*, vol. 28, n° 4, 2001, 965-977 p.
- CREASCIENCE. *Analyse des données socio-économiques*, Creascience Inc, 2005.
- CHOY, Dexter J.L. *Forecasting tourism revisited*, *Tourism Management*, september 1984, 71-176 p.
- FRECHTLING Douglas C. *An Assessment of Visitor Expenditure Methods and Models*, *Journal of Travel Research*, vol. 45, august 2006, 26-35 p.
- FRECHTLING, Douglas C. *Forecasting Tourism Demand: Methods and Strategies*, Butterworth Heinemann, 2001.
- FÉDÉRATION DES CAISSES DESJARDINS DU QUÉBEC. *Guide des indicateurs économiques des principaux pays industrialisés*, Desjardins Études économiques, première édition, avril 2004, 168 p.
- CHU, Fong-Lin. *Forecasting Tourist Arrivals: Nonlinear Sine Wave or ARIMA*, *Journal of Travel Research*, vol. 36, winter 1998, 79-84 p.
- LI, Gang, and others. *Tourism Demand Forecasting: A Time Varying Parameter Error Correction Model*, *Journal of Travel Research*, vol. 45, november 2006, 175-185 p.
- CROUCH, Geoffrey I. *The Study of International Tourism Demand: A Review of Findings*, *Journal of Travel Research*, summer 1994, 12-23 p.
- CROUCH, Geoffrey I. *The Study of International Tourism Demand: A Survey of Practice*, *Journal of Travel Research*, spring 1994, 41-55 p.
- CROUCH, Geoffrey I. *A meta-analysis of tourism demand*, *Annals of Tourism Research*, vol. 22, n° 1, 1995, 103-118 p.

- YU, Gongmei, and Zvi SCHWARTZ. *Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models*, Journal of Travel Research, vol. 45, november 2006, 194-203 p.
- SONG, Haiyan, Stephen F. WITT, Thoma C. JENSEN. *Tourism forecasting: accuracy of alternative econometric models*, International Journal of Forecasting 19, 2003, 123-141 p.
- INSTITUT CANADIEN DE RECHERCHE SUR LE TOURISME. *Taux de change : Les conséquences possibles d'un dollar canadien plus fort sur les voyages au Canada en 2004 et 2005*, Commission canadienne du tourisme, avril 2004, 27 p.
- VU, Jo Chau, and Lindsay W. TURNER, *Regional Data Forecasting Accuracy: The Case of Thailand*, Journal of Travel Research, vol. 45, november, 186-193 p.
- RABOTEUR, Joël. *Introduction à l'économie du tourisme*, Sociétés et Économies insulaires, janvier, 2000, 152 p.
- COSHALL, John. *Time series Analyses of UK Outbound Travel by Air*, Journal of Travel Research, vol. 44, february 2006, 335-347 p.
- MINISTRY OF TOURISM. *New Zealand tourism forecasting methodology*, september 2006.
- KULENDRAN, Nada, and Stephen F.WITT. *Cointegration versus least squares regression*, *Annals of Tourism Research*, vol. 28, n° 2, 2001, 291-311 p.
- KALEKAR, Prajakta S. *Time series Forecasting using Holt-Winters Exponential Smoothing*, Kanwal Rekhi School of Information Technology, december 2004, 13 p.
- GELLER, S. *Abrégé de statistique à l'usage des étudiants en médecine et en biologie*, Masson & Cie, deuxième édition, 1974.
- LEBLOND, Simon, et Isabelle BELLEY-FERRIS. *Guide d'économétrie appliquée à l'intention des étudiants du cours ECN 3950*, Département de sciences économiques, Université de Montréal, octobre 2004, 14 p.
- SHAW, Simon. *Forecasting*, 2005, 13 p.
- MAKRIDAKIS, Spyros, Steven C. WHEELWRIGHT, and Rob J. HYNDMAN. *Forecasting: Methods and Applications*, John Wiley & Sons, inc., third edition, 1998.
- STATISTIQUE CANADA. *La montée du buard et les voyages internationaux*, L'Observateur économique canadien, février 2005, 7 p.

