

LES PROJETS COMME MOYEN DE CONTRÔLE DES CONNAISSANCES EN STATISTIQUE

Guy Mélard

ECARES CP114/4 Université libre de Bruxelles,
Avenue Franklin Roosevelt, 50, 1050 Bruxelles, Belgique
gmelard@ulb.ac.be , <http://homepages.ulb.ac.be/~gmelard>

RÉSUMÉ

L'enseignement par projets a des avantages indéniables. L'utilisation de projets pour le contrôle des connaissances dans le cadre d'un enseignement traditionnel est moins courante. L'article décrit une expérience qui a été menée depuis une quinzaine d'année auprès d'un public déterminé pour un enseignement de statistique et présente les avantages et les inconvénients de la démarche. L'article se termine par des conseils en vue d'une généralisation de la démarche.

ABSTRACT

Teaching through projects has undeniable advantages. Using projects for the control of knowledge in a traditional teaching is less common. In this paper an experiment that was conducted over the last 15 year to a specific audience for a course of statistics. Also advantages and disadvantages of the approach are presented. We provide some advice on a possible generalization of the approach.

MOTS-CLÉS

Enseignement de statistique, enseignement par projets, méthodes de prévision.

KEY WORDS

Teaching statistics, teaching by projects, forecasting methods.

INTRODUCTION

Depuis 1983 l'auteur enseigne un cours de 24 heures de techniques quantitatives de gestion I (GEST-D-615) au niveau post-master, le master complémentaire en gestion technologique et industrielle (anciennement DES en gestion) à la Solvay Brussels School of Economics and Management (anciennement Solvay Business School) de l'Université libre de Bruxelles. Ce programme s'adresse à un public très diversifié allant d'ingénieurs à des juristes en passant par tous les programmes de sciences exactes et humaines (à l'exception des sciences de la gestion), tous fraîchement diplômés et dont beaucoup ont une activité professionnelle. Il a pour but d'enseigner la gestion en deux années en horaire aménagé. Le nombre d'étudiants est de l'ordre de 80 en première année.

A l'origine, le cours était un cours de mise à niveau en statistique. Devant la difficulté pédagogique d'intéresser un public aussi varié, l'auteur l'a transformé en 1995 en un cours de méthodes de prévision. Il faut le préciser pour la bonne compréhension de la suite. Cette matière a l'avantage d'être directement utilisable en gestion et de ne pas être dans la formation antérieure du public cible. Le support principal du cours était celle des chapitres 1 à 10 de Mélard (1990), c'est-à-dire les méthodes de régression linéaire simple et multiple, courbes de croissance, décomposition saisonnière, lissage exponentiel, modèles ARIMA, tout en réduisant quelque peu les aspects mathématiques et de calcul. La matière est néanmoins difficile.

Vers 2001, le contenu a été enrichi et une version en auto-apprentissage a été mise à disposition qui a abouti à la 2e édition du livre, Mélard (2007). Cohen et Mélard (2009) ont décrit cette évolution qui n'est pas l'objet du présent article. Cet article se concentre sur le contrôle des connaissances. Pour motiver les étudiants et les sensibiliser à l'application des concepts théoriques, l'examen a été remplacé par la réalisation d'un projet de prévision par groupes homogènes de 2 étudiants. Chaque groupe choisit son sujet, et donc son jeu de données, dans le respect d'un cahier de charges que nous

préciserons dans le paragraphe suivant.

L'enseignement par projets a des avantages indéniables, par exemple Proulx (2005). L'utilisation de projets pour le contrôle des connaissances dans le cadre d'un enseignement traditionnel est moins courante. Pour des généralités sur le contrôle de connaissances, voir par exemple Kubiszyn et Borich (1996).

Plus de détails sont mis à disposition dans une version étendue de cet article (Mélard, 2011). Je remercie le lecteur anonyme et Mme Atika Cohen pour leurs remarques très utiles sur les premières versions de ce texte.

PROJET COMME SUBSTITUT A L'EXAMEN

Avant la réorientation du cours vers les méthodes de prévision, la mise en œuvre de l'examen de contrôle de connaissances posait déjà problème compte tenu de la diversité du public. La solution retenue consistait à présenter, d'une part, des questions relativement simples de calcul et de réflexion, et d'autre part, des études de cas. Par la suite, il s'est avéré préférable d'imposer seulement aux étudiants formés au quantitatif d'aborder les processus aléatoires et les modèles ARIMA. Le remplacement de l'examen par un projet à réaliser s'est aussi présenté naturellement.

Les principaux éléments du cahier des charges, voir l'annexe 1 de Mélard (2011) qui contient également quelques recommandations, sont les suivantes :

- le projet doit, d'une manière ou d'une autre, être lié à la gestion (au sens large, incluant le secteur non marchand et le secteur de l'énergie) ;
- le projet doit être pratique et comporter au moins une série temporelle réelle ;
- l'obligation de traiter au moins deux des méthodes de prévision vues au cours et de préférence plus de deux méthodes ;
- la production de prévisions ex post et comparaison aux données réservées et de préférence production de prévisions ex ante ;
- l'obtention d'intervalles de prévision et discussion de leur interprétation ;
- la rédaction d'un rapport écrit avec introduction, conclusion, annexes de calcul, références ;
- la fourniture des données et des moyens de vérifier les calculs (classeurs du tableur, sorties) ;
- tous les étudiants doivent aborder au moins une des méthodes plus avancées (voir ci-dessous) ; de plus les étudiants avec une formation mathématique forte (ingénieurs et assimilés) doivent traiter les modèles ARIMA ;
- le travail doit être déposé à une date précise, tout retard étant l'objet de pénalité ; selon les règles en vigueur, les étudiants ont droit à une deuxième session en août, où ils peuvent éventuellement présenter une version corrigée de leur travail.

La grille d'évaluation est basée sur ce cahier des charges. Détaillons-en certains points avant de discuter de l'évaluation.

Concernant le choix du sujet, si le titulaire du cours accepte de formuler des conseils sur les sujets proposés par les étudiants, il n'a jamais fourni de suggestions de sujets, se contentant de donner éventuellement des exemples de sujets déjà traités, voir l'annexe 2 de Mélard (2011).

Concernant les méthodes, les principales ont été mentionnées dans l'introduction mais il y a bien sûr des variantes. Par exemple, la décomposition saisonnière peut être réalisée dans le cadre d'un modèle additif ou multiplicatif, par plusieurs méthodes de comparaison des données à une estimation de la tendance-cycle et en employant plusieurs paramètres de synthèse. Parmi les méthodes de lissage exponentiel, outre le lissage exponentiel simple, plusieurs autres méthodes plus avancées sont mentionnées (lissage double, lissage de Holt), sans ou avec correction saisonnière (c'est-à-dire appliquée sur données corrigées des variations saisonnières) ou encore avec composante saisonnière incorporée (méthode de Winters). Les méthodes avancées sont les suivants : régression linéaire multiple, autocorrélation, méthodes avancées de lissage exponentiel. Les étudiants les moins bien préparés n'ont donc pas à traiter toutes les méthodes plus avancées et encore moins les modèles ARIMA qui nécessitent au moins le concept de processus aléatoire.

Quant à la production de prévisions ex post, elle requiert aussi une explication. Il s'agit de

l'obligation de réserver la partie la plus récente des données (par exemple la dernière année), de choisir les modèles, d'estimer leurs paramètres sur les données antérieures (l'historique), de calculer les prévisions (dites ex post) sur cette partie et de les comparer aux données à l'aide de critères à choisir (carré moyen des erreurs, moyenne des erreurs absolues en pourcentage).

Du point de vue des logiciels, un tableur peut suffire pour une partie des méthodes, mais un logiciel statistique est recommandé pour le chapitre sur les modèles ARIMA, voir Mélard *et al.* (2003). Une mise à jour de Time Series Expert (sur le CD-Rom de Mélard, 2007) a été fournie et une liste de logiciels a été communiquée. Certains des étudiants emploient un logiciel disponible sur le lieu de travail (par exemple Minitab), un logiciel libre (Demetra) ou une version d'évaluation d'un autre logiciel. De nombreux ingénieurs utilisent Matlab.

La correction des projets tient compte de la qualité du travail et du respect du cahier des charges et s'opère en quatre phases étalées sur quelques semaines.

1. Première vérification des livrables (rapport, annexes, ...), saisie des informations de base dans un formulaire, voir l'annexe 3 de Mélard (2011) qui avait été fournie aux étudiants, lequel servira de grille d'aide à l'évaluation, et communication aux étudiants d'un accusé de réception mentionnant éventuellement les éléments manquants à compléter (CD illisible, signature manquante, ...).

2. Dans une seconde phase, lecture des rapports et rédaction dans les formulaires d'une première version des rapports d'évaluation. Par exemple, les étudiants ont-ils fourni les intervalles de prévision demandés, au moins pour quelques méthodes et en précisant la probabilité de couverture ? Ont-ils interprétés ces résultats ? Les points qui méritent une vérification sont mentionnés.

3. Ensuite, vérification des annexes informatiques, par coup de sonde et notamment sur les points mentionnés comme étant à vérifier. Il arrive, par exemple, que les coefficients saisonniers soient calculés dans un tableur mais soient erronés. En cas de doute, le titulaire effectue des traitements complémentaires. Les rapports d'évaluation sont mis à jour en conséquence.

4. Enfin, les projets sont regroupés par niveau de qualité et une note leur est attribuée. Une dernière vérification des rapports d'évaluation complétés par les notes est alors effectuée. Le titulaire du cours fournit alors le rapport d'évaluation à chaque groupe par projet d'une page environ en signalant les problèmes rencontrés. Il arrive que la note soit revue à la suite de la réaction d'un étudiant.

L'annexe 4 de Mélard (2011) contient trois exemples de rapports rendus anonymes pour des raisons évidentes. Comme pour un travail de fin d'études, il ne s'est pas avéré possible d'employer une grille d'évaluation purement quantitative.

AVANTAGES DU CONTRÔLE PAR PROJET

Parmi les avantages du contrôle par projet, on peut citer

- l'aspect formatif pour les étudiants qui ont l'occasion de pratiquer ce qui a été vu au cours ;
- l'opportunité offerte de traiter un sujet proche de l'entreprise ; certains ont même la possibilité d'intéresser au projet l'entreprise qui les emploie ;
- la qualité des travaux est élevée grâce à la motivation liée au choix personnel du sujet ;
- pour le correcteur : lecture de travaux tous très différents donc absence de monotonie ;
- certains sujets intéressants lui ont donné des idées pour des projets de recherche ;
- en général les étudiants ne se plaignent pas des évaluations ni de l'ampleur du travail.

INCONVÉNIENTS DU CONTRÔLE PAR PROJET

Parmi les inconvénients du contrôle par projet, on peut citer

- la difficulté de trouver un sujet de la part des étudiants ;
- le temps de réalisation qui inclut la maîtrise d'une grande partie du cours, et s'avère supérieur à celui d'un examen, même si du temps de travail personnel est prévu ;
- le risque de fraude, cas où le travail est réalisé par autrui ;
- la plus grande difficulté d'évaluation de travaux de longueur et de profondeur très variable ;
- la plus grande durée de correction et de réalisation des rapports d'évaluation (2 à 3 heures par projet) ;
- la possibilité de biais lié au choix du sujet.

Détaillons certains de ces points. La difficulté de trouver un sujet est différente selon le profil des étudiants. Ceux qui travaillent doivent obtenir l'accord de leur direction et éventuellement maquiller les données et/ou demander que le titulaire signe un accord de non-divulgaration. Ceux qui ne travaillent pas ou ceux, assez nombreux, qui sont en même temps chercheurs à l'université doivent contacter des proches ou des connaissances.

Concernant les risques de fraude, comme pour les travaux de fin d'étude, ils ne sont pas nuls. Il est évidemment difficile de l'estimer. Heureusement le risque de réutilisation d'un travail de l'extérieur est limité par plusieurs des contraintes du cahier des charges, quoiqu'il soit arrivé d'observer un travail contenant des tests de racine unité, très fréquents chez nos collègues. Nous avons observé à plusieurs reprises, soit la réutilisation naïve du travail d'une année précédente avec simple substitution de nom, soit l'utilisation d'un sujet neuf mais avec des données fabriquées à partir de celle d'une année précédente et récupération de quelques éléments de quelques classeurs Excel et de bribes de l'ancien rapport. Ces fraudes ont été repérées grâce à la base de données des travaux antérieurs.

Pour ce qui est du biais lié au choix du travail, il est clairement annoncé qu'un sujet intéressant sera avantagé encore que gain en nombre de points n'est probablement pas proportionnel au temps additionnel perdu pour sa recherche, par rapport à un sujet pris sur internet. En revanche, la qualité des prévisions obtenues n'est pas favorisée.

GÉNÉRALISATION DE L'APPROCHE

Nous discutons maintenant dans quelle mesure cette procédure de contrôle de connaissances par projets est généralisable à d'autres contextes d'enseignement de la statistique. Pour des raisons pratiques (durée de correction, risque de fraude), nous ne la recommandons pas pour des cours qui s'adressent à de grands nombres d'étudiants. Cela écarte donc les cours de base de statistique. Cette approche ne s'applique pas à des cours de statistique mathématique mais plutôt à des cours de statistique appliquée. Notre avantage de disposer d'étudiants qui sont souvent dans un environnement professionnel n'est pas une nécessité. Des étudiants exclusivement aux études peuvent rencontrer des conditions similaires à l'occasion de stage en entreprises. Des projets similaires peuvent être développés pour des données scientifiques ou de livres. On exigera néanmoins des références très précises des sources. Nous avons mentionné ici le traitement de données temporelles. D'autres types de données peuvent être traités mais on remarquera que les données d'enquête sont souvent plus difficiles à obtenir pour des raisons de confidentialité. La réalisation complète d'une enquête par questionnaire est également formative à condition de bien l'encadrer en demandant la lecture de recommandations générales.

Bibliographie

- [1] Cohen, A. et Mélard, G. (2009) Retombées d'une expérience d'enseignement en analyse de données temporelles, 41èmes Journées de Statistique, Bordeaux, 25-29 mai 2009, <http://hal.inria.fr/inria-00386777/fr/>.
- [2] Kubiszyn, T. et Borich, G. (1996) *Educational testing and measurement : classroom application and practice*, 5e éd., HarperCollins, New York.
- [3] Mélard G. (1990) *Méthodes de prévision à court terme*, Editions de l'Université de Bruxelles, Bruxelles, et Editions Ellipses, Paris.
- [4] Mélard G. (2007) *Méthodes de prévision à court terme*, 2e édition (avec CD-Rom), Editions de l'Université de Bruxelles, Bruxelles, et Editions Ellipses, Paris.
- [5] Mélard G. (2011) Contrôle des connaissances en statistique par projets : avantages et inconvénients, <http://homepages.ulb.ac.be/~gmelard/submTunis.pdf>.
- [6] Mélard G., Cohen A., Lotfi, S. and Njimi, H. (2003) Software for a time series analysis course, International Symposium on Forecasting, ISF2003, Mérida, México, June 15-18, 88.
- [7] Proulx, J. (2005) *L'apprentissage par projet*, Presses de l'Université du Québec, Québec.

Cahier des charges du travail d'examen

1. Introduction

La note est attribuée sur base d'un travail. Le travail doit être relatif au cours, être réalisé par un groupe homogène d'élèves (2, éventuellement 3, si le volume le justifie et après accord du titulaire), représenter en temps au moins le temps d'étude d'un cours de 24 heures et respecter pour le fond comme pour la forme les instructions générales de la SBS-EM :

- effectuer le travail personnellement sans aide sauf éventuellement pour la compréhension du cours,
- citer les références utilisées,
- éviter les copies textuelles sauf à mentionner la source (avec mention de la page),
- ne pas employer de données confidentielles (sauf avec accord),
- éventuellement maquiller les données si cela peut satisfaire le fournisseur.

2. Le sujet et les données

Les meilleurs travaux sont ceux dont on se sent le plus proche: plutôt que de traiter des données officielles ou des données trouvées sur l'internet, *il est plus intéressant d'offrir ses services à une entreprise ou à une collectivité et de dialoguer avec un partenaire intéressé par le projet.*

Autres recommandations:

- Pour certaines méthodes (celles des deux derniers chapitres en particulier), les séries chronologiques doivent être mensuelles ou trimestrielles et comporter au moins une soixantaine de données.
- De façon générale, il est conseillé d'employer des séries aussi longues que possible sous réserve qu'elles soient homogènes.
- Réfléchir où placer les données dans le temps: en fin de mois (variable de niveau) ou en milieu de mois (variable de flux).
- Essayer d'établir des liens avec les autres cours sans provoquer de double emploi.
- Introduire le problème traité (intérêt de la prévision, terminologie, qualité des données, ...); en revanche, il n'est pas nécessaire de reprendre des éléments du cours (sauf à la demande d'un partenaire extérieur, ...), le titulaire le connaissant suffisamment.
- Joindre les données sous forme de tableau ou sur disquette afin de permettre la reproductibilité des résultats. Présenter le graphique des données.

3. Les méthodes

Parmi les méthodes étudiées dans le cours, la régression multiple et les modèles ARIMA sont les plus aptes à alimenter une discussion intéressante. Il ne faut pas négliger pour autant les moyennes mobiles, la décomposition saisonnière, les différentes formes de lissage exponentiel. On essayera toujours d'avoir au moins deux modèles de façon à pouvoir les comparer.

Pour tenir compte des différences de formation préalable des étudiants, il est demandé de constituer des binômes homogènes et il est recommandé de limiter les méthodes utilisées comme suit, en utilisant la notation en étoiles de l'annexe B :

- Pour ceux dont la formation préalable a comporté des cours de mathématiques et de statistique approfondis (ingénieurs, masters en sciences, en sciences économiques, ...) : jusqu'à **** ;
- Pour ceux sans aucune formation préalable quantitative (masters en droit, en philologie, ...) : jusqu'à ** plus un sujet de niveau ***, au choix ;
- Pour les autres étudiants : jusqu'à ***.

Afin que la comparaison de méthodes de prévision soit justifiée, on estimera les modèles en laissant de côté quelques données (entre 6 mois et 2 ans, en général) qui ne seront utilisées que pour juger de la validité des méthodes. Utiliser à cette fin les critères vus dans le chapitre 1, notamment les critères RMSE et MAPE.

Certaines méthodes (régression linéaire, modèles ARIMA) permettent de réaliser des intervalles de prévision. Une méthode élémentaire décrite dans le chapitre 1 permet de représenter la fonction de distribution de la valeur future, de façon approchée. On peut en déduire un intervalle de prévision tout autant approché mais ceci quelle que soit la méthode de prévision utilisée.

Privilégier des modèles qui peuvent être formulés a priori, sans connaître les données et qui sont donc de ce fait susceptibles d'une explication. D'autre part, les données étudiées sont chronologiques. Les dates auxquelles arrivent des résidus importants sont donc intéressantes et peuvent correspondre à des faits historiques répertoriés. Outre la littérature spécialisée, des encyclopédies ou des ouvrages comme le "Quid" (Editions Robert Laffont) peuvent être consultés. L'accès aux numéros anciens de journaux demande plus de temps. Penser aux ressources de l'Internet.

La modélisation peut être un jeu dangereux. A plusieurs endroits dans le cours on met en garde contre le fait d'employer plus de paramètres qu'il n'est nécessaire (surparamétrisation) et contre le danger des tests statistiques multiples (si 100 tests sont réalisés au niveau de 5%, on doit s'attendre à 5 rejets de l'hypothèse dans le cas où celle-ci est vraie). C'est surtout dangereux avec les modèles ARIMA, où on a parfois tendance à employer des modèles trop complexes.

Autres recommandations:

- Privilégier des modèles qui peuvent être formulés a priori, sans connaître les données et qui sont donc de ce fait susceptibles d'une explication.
- D'autre part, les données étudiées sont chronologiques. Les dates auxquelles arrivent des résidus importants sont donc intéressantes et peuvent correspondre à des faits historiques répertoriés. Outre la littérature spécialisée, des encyclopédies ou des ouvrages comme le "Quid" (Editions

Robert Laffont) peuvent être consultés. L'accès aux numéros anciens de journaux demande plus de temps. Penser éventuellement aux ressources de l'Internet.

- La modélisation peut être un jeu dangereux. A plusieurs endroits dans le cours on met en garde contre le fait d'employer plus de paramètres qu'il n'est nécessaire (surparamétrisation) et contre le danger des tests statistiques multiples (si 100 tests sont réalisés au niveau de 5%, on doit s'attendre à 5 rejets de l'hypothèse dans le cas où celle-ci est vraie). C'est surtout dangereux avec les modèles ARIMA, où on a parfois tendance à employer des modèles trop complexes.

Voici quelques remarques au sujet des différentes méthodes.

- Quelle que soit la méthode envisagée, commencer par une étape de familiarisation avec les données (paragraphe 10.2 dans l'ouvrage de référence) et une analyse préliminaire (paragraphe 10.3) au moins sous forme sommaire.
- Certaines méthodes nécessitent certaines conditions pour être employées: par exemple, le lissage exponentiel simple n'est pas applicable s'il y a une tendance (voir alors le lissage double de Brown ou le lissage de Holt) ou s'il y a une saisonnalité (voir alors le lissage de Winters ou appliquer le lissage simple sur la série corrigée des variations saisonnières, voir ci-dessous).
- Réfléchir avant d'agir. Ce n'est pas gênant qu'une méthode soit appliquée alors qu'il ne faudrait pas à condition que ceci soit remarqué et commenté dans le rapport.
- Pour la méthode de prévision par moyenne mobile sur des données mensuelles, le choix d'un ordre 12 est le plus mauvais qu'on puisse faire pour la prévision puisque la saisonnalité est rabotée; de plus, le centrage est justifié pour du lissage mais pas pour de la prévision.
- Il y a fréquemment choix entre un modèle additif et un modèle multiplicatif (ou un modèle additif sur la série en logarithmes). Justifier ce choix par l'examen graphique (voir chapitre 5).
- Pour que la décomposition saisonnière soit bien réalisée il convient que la tendance soit déterminée non pas à partir des moyennes mobiles sur un an, mais à partir des moyennes annuelles (voir l'ouvrage de référence). S'il y a une grande instabilité dans la comparaison données - tendance-cycle, on peut obtenir les coefficients saisonniers autrement que par une moyenne (moyenne tronquée, voire médiane).
- L'analyse des résidus (moyenne, étude de l'homoscédasticité, détection des valeurs aberrantes, autocorrélation) fait partie intégrante de la régression multiple et de la modélisation ARIMA mais il n'y a pas de raison pour ne pas l'utiliser sur les erreurs de prévision des autres méthodes.
- On insiste dans le cours sur les liens entre le lissage exponentiel et les modèles ARIMA. Il est conseillé d'exploiter ces liens.
- Certaines méthodes ne sont pas adaptées à la présence d'une saisonnalité, comme les lissages exponentiels simple et double. Il faut alors les appliquer sur les séries corrigées des variations saisonnières, et restituer la saisonnalité aux prévisions (c'est très facile à faire dans TSE).
- La régression multiple comme les modèles ARIMA permettent d'inclure de l'information extérieure. De l'information qualitative peut être introduite à l'aide de variables binaires, notamment.

4. Les logiciels

Du point de vue des logiciels, les salles informatiques disposent notamment de EViews (version 3.1), d'Excel (version 2003 ou 2007) et de SPSS. D'autres logiciels (gratuits, en version d'évaluation limitée dans le temps ou éventuellement disponibles sur le lieu de travail comme SAS, SPSS, Statistica, Systat, XLStat, Statgraphics, Minitab, ...) peuvent être employés. Sur le CD-Rom de Mélard (2007) sont disponibles TSE pour DOS (version 2.4) et TSE pour Windows (version 3.2). TSE pour Windows a été fortement amélioré entre-temps. Une version TSE 3.3 est disponible sur le site du cours dans l'Université virtuelle, voir annexe H. Remarquons ce qui suit:

- Les assistants des salles informatiques ne sont pas engagés pour aider à l'emploi des logiciels (un cours a été donné sur ce sujet par le titulaire).
- Excel est très bien adapté pour la présentation de tableaux et de graphiques, pour les moyennes mobiles, la décomposition saisonnière et le lissage exponentiel.
- Dans EViews, il faut spécifier explicitement la constante dans un modèle.
- EViews et TSE permettent de traiter les modèles ARIMA, pas Excel. La notation de EViews pour les coefficients d'un polynôme moyenne mobile n'est pas la même que dans le cours (les coefficients sont changés de signe).
- Time Series Expert ou TSE en version 2.2 (version téléchargeable sur l'Internet), 2.3 (version installée dans les salles Renaissance) ou 2.4 (version sur l'Université virtuelle ou sur le CD-Rom joint à la 2^e édition du livre) n'étant pas un logiciel conçu pour Windows, il peut s'avérer difficile (voire impossible sous Windows XP et suivants) de copier/coller les graphiques. Pour les systèmes où cela marche, on peut ouvrir une fenêtre de commande et employer l'option "Edit" de la case système pour marquer et copier et ensuite coller dans WordPad ou un traitement de texte. Pour les textes et tableaux, le mieux est de sauver les fichiers et de les ouvrir dans le traitement de texte, comme fichiers texte MS-DOS. En configurant TSE, on peut aussi sauver les graphiques en mode PostScript (avec une extension EPS) et les insérer dans Word ou les convertir dans un programme approprié (Adobe Illustrator, par exemple). La version 2.4 de TSE dispose d'un programme de réalisation de graphiques sous Windows qui facilite les récupérations de graphiques. Cette version DOS ne fonctionne pas du tout dans un système Microsoft à 64 bits (XP, Vista ou Seven) et pas non plus sous MacOS ou Linux à moins d'employer la DosBox (<http://www.dosbox.com/>) qui permet de l'employer virtuellement partout.
- Time Series Expert ou TSE en version 3.2 est un logiciel dont l'interface est conçue pour Windows. Parce qu'il emploie encore des programmes 16 bits, il ne fonctionne pas sous les systèmes 64 bits. Cette version était fournie sur le CD-Rom de Mélard (2007).
- Time Series Expert ou TSE en version 3.3 est un logiciel dont l'interface est conçue pour Windows. Il existe(ra) en 4 éditions. A l'exception de l'édition de démonstration, il fonctionne sous les systèmes Windows 64 bits. L'édition professionnelle est disponible sur l'Université virtuelle. Les deux autres éditions seront commercialisées (<http://www.itse.be/>) et permettent de traiter plus de 400 observations. L'édition professionnelle étendue contiendra le module TSE-

AX de modélisation automatique partiellement automatique, c'est-à-dire une version améliorée du module optionnel TSE-AX de la version 2.3.

- Le passage entre Eviews ou TSE 2.4, d'une part, et Excel, d'autre part, peut se faire par l'intermédiaire du format WK1 (feuille de calcul de Lotus 1-2-3 version 2). Le tableur incorporé de TSE 3.2 ou 3.3 permet les opérations de copier/coller des données avec les tableurs.
- Les fichiers de données de EViews, d'extension .DB, peuvent être lus et écrits par TSE. En revanche les fichiers de type "workfile", d'extension .WF1 ne peuvent pas être récupérés.

5. Le rapport

Quelques conseils

- Fournir un rapport écrit imprimé et relié (une version électronique **ne suffit pas**).
- Donner les noms, les diplômes principaux et les adresses de courrier électronique de chacun des membres du groupe pour faciliter la communication.
- Pour chaque étudiant, faire figurer (en page 2) la mention "*J'affirme sur l'honneur que j'ai effectué ce travail personnellement*" et signer.
- **Signaler la répartition du travail entre les participants, qui devrait être équilibrée.**
- Commencer par une introduction au problème mentionnant les objectifs poursuivis et justifiant les méthodes utilisées.
- Ne pas nécessairement reprendre tous les tableaux et tous les graphiques de résultats. Se limiter aux éléments essentiels, en particulier à ceux qui servent à prendre une décision fondamentale. Il est fortement recommandé de joindre les détails dans une version électronique (sur disquette, CD ou par courrier électronique (à condition de n'envoyer les fichiers que sous la forme d'un seul fichier compressé, dans ce dernier cas).
- Si les tableaux ne sont pas récupérés d'un logiciel mais sont saisis à nouveau, on peut se contenter des chiffres les plus significatifs (2 à 4, le plus souvent). Des résultats statistiques à 10 décimales sont rarement plus corrects que ceux à 4 décimales.
- Eviter autant que possible le jargon propre au domaine étudié comme le jargon statistique. Donner les équations des modèles utilisés. Choisir le nom des variables (plutôt que de prendre X, Y ou VAR). Si les données ont été fournies par un tiers, rédiger le texte de manière à ce que l'essentiel lui soit compréhensible.
- Ne pas oublier les conclusions, y compris sur l'utilité des méthodes utilisées.
- Prendre l'habitude de soigner la forme. Un gestionnaire du 21e siècle doit maîtriser les outils mis à sa disposition (traitement de texte, tableur, logiciel de dessin) afin de réaliser la *communication* de sa connaissance.

Le travail doit être rendu le jour convenu c'est-à-dire le jour spécifié à l'annexe B. **Une pénalité d'un point par jour de retard sera appliquée.**

Le titulaire du cours (ou son suppléant désigné) se réserve le droit de convoquer un étudiant pour discuter du travail et s'assurer ainsi que ce travail a bien été réalisé par l'étudiant.

ANNEXE 2. Les sujets traités en 2007

M/ref	Subject	Dates	Note	DiscData
TQG07.01	Ventes de champagne et produits similaires d'une grande surfa	03.01-01	14	Données originales, très bien décrites
TQG07.02	Nombre de passagers à Bruxelles-National (embarquements/k	00.01-01	19	Données intéressantes, assez bien discutées
TQG07.03	Consommation mensuelle nette d'électricité en Belgique	93.01-01	15	Données loin d'être originales mais très bien dis
TQG07.04	Ventes de vin rouge en Australie	80.01-91	16	Données de l'Internet mais effort méritoire pour
TQG07.05	Nombre de nuitées dans les hôtels en Belgique	90.01-01	16	Données intéressantes, assez bien discutées
TQG07.06	Evolution du trafic de XXX sur le marché des valeurs financière	93.01-01	18	Données originales, excellentement décrites
TQG07.07	Nombre de demandeurs d'emploi de moins de 25 ans en Belgi	82.01-01	15	Données disponibles mais effort de présentatio
TQG07.08	Courbes de températures	83.01.01	19	Données originales, excellentement décrites. Prob
TQG07.09	Consommation d'électricité en Belgique	94.01-01	14	Données loin d'être originales mais assez bien d
TQG07.10	Chiffre d'affaires de XXX	02.01-01	12	Données originales, mais très mal décrites
TQG07.11	Les XXX 4 XXX	04.10-01	16	Données originales (3 séries), très bien décrites
TQG07.12	Indice des exportations de la Belgique selon le concept nation.	98.01-01	15	Données de l'Internet sans description (seuleme
TQG07.13	Produit intérieur brut et chiffre d'affaire des entreprises belges	95.1-06.	16	Données de l'Internet. Double série pour varier le
TQG07.14	Naissances en Belgique	01.01-05	15	Données de l'Internet avec description très brèv
TQG07.15	Population des demandeurs d'emploi inoccupés	00.01-01	12	Données disponibles avec une bonne descriptio
TQG07.16	CONFIDENTIEL Ventes de polypropylène par une société pétr	98.01-01	17	Données originales, très bien décrites
TQG07.17	Consommation de gaz naturel en Belgique	91.01-05	18	Données disponibles avec une très bonne descri
TQG07.18	Nombre de joueurs de Lotto	99.01-01	15	Données intéressantes, très bien discutées
TQG07.19	Chiffre d'affaire de XXX	95.01-01	15	Données originales, très bien décrites
TQG07.20	Ventes de maisons aux Etats-Unis	94.01-01	14	Données de l'Internet avec une description très l
TQG07.21	Prêts de CD de la Médiathèque du Passage 44	00.07-11	16	Données originales, excellentement bien décrites
TQG07.22	Immatriculation de véhicules neufs et usagés	03.01-01	14	Données disponibles avec description moyenn
TQG07.23	Nombre de passagers de l'aéroport de Nice-Côte d'Azur	97.01-01	15	Données disponibles avec description moyenn
TQG07.24	Ventes de blocs de béton cellulaires	02.01-01	17	Données originales, très bien décrites
TQG07.25	Nombre de touristes à Bruxelles	98.01-01	14	Données loin d'être originales mais très bien dis
TQG07.26	Permis de bâtir de la région de Bruxelles-Capitale	96.01-01	13	Données de l'Internet mais très bien décrites
TQG07.27	XXX Belgique	02.01-01	15	Données originales, très bien décrites
TQG07.28	Nombre de passagers au départ et à l'arrivée de l'aéroport de C	99.01-01	14	Données intéressantes, très bien discutées
TQG07.29	Consommation de tabac en Espagne, 4e trimestre de 1986	non chi	10	Données intéressantes, bien discutées
TQG07.30	Cimenterie de Lukala	98.01-01	17	Données originales, excellentement décrites. Situa
TQG07.31	Consommation de gaz naturel en Belgique	85.01-01	17	Données loin d'être originales mais effort de pré
TQG07.32	Produit intérieur brut de l'Australie, sauf la production agricole	59.1-95.	13	Données de l'Internet sans description
TQG07.33	L'usine XXX	1998.01.	16	Données originales, excellentement décrites. Situa
TQG07.34	Cours Euronext de ING Belgique	2002.01	13	Données disponibles, avec intérêt particulier ma
TQG07.35	Visiteurs au musée royal de l'Afrique centrale de Tervuren	2001.01.	14	Données originales, excellentement décrites mais s
TQG07.36	Chiffre d'affaires de X	2002.01	13	Données originales, mais assez mal décrites
TQG07.37	Nombre d'accidents de la route	1999.01.	12	Données de l'Internet, moyennement bien décrit

ANNEXE 3. Document de base du rapport de retour (complété par publipostage à partir d'un classeur Excel)

Solvay Brussels School of Economics and Management, MCC1 en Gestion 2010-2011

GEST-D-615. Techniques quantitatives de gestion I, Prof. G. Mélard

Dans le contexte de la discussion de votre travail, voici mes appréciations. La qualité des travaux est toujours très bonne (moyenne 13,3) et j'ai apprécié les nombreux sujets de gestion. Je vous en félicite. Je reste à votre disposition pour plus de renseignements (gmelard@ulb.ac.be ou ☎ 026504604).

Ma référence: «Mref»
Noms: «Names»
Sujet: «Subject»
Dates: «Dates»
Note sur 20: «Note20»

Remarques générales:

«Paragr1»

«Paragr2»

«Parag3»

Discussion des données	«DiscData»
Tableau ou fichier des données fourni	«Tab_or_file»
Graphiques	«Graphs»
2 méthodes employées au moins	«Methods1»
Utilisation des critères	«Criteria»
Emploi d'une méthode simple	«SimpleMeth»
Emploi d'une méthode plus avancée	«MoreSophist»
Valeurs extrêmes et interventions si nécessaire	«InterventionOutliers»
Modèles correctement paramétrés	«Parametrisation»
Choix du modèle additif/multiplicatif, transformation	«ModelTransform»
Analyse des résidus	«ResidAnal»
Prévisions ex post bien réalisées	«ExPostForecast»
Intervalles de prévision.	«ForcIntervals»
Ecriture correcte des équations des modèles	«Equations»
Rapport bien structuré	«StructureReprt»
Présentation matérielle	«Presentation»
Présence de conclusions	«Conclusions»
La démarche vue au cours est bien comprise.	«CourseUnderstanding»

Solvay Brussels School of Economics and Management, MCC1 en Gestion 2010-2011

GEST-D-615. Techniques quantitatives de gestion I, Prof. G. Mélard

Dans le contexte de la discussion de votre travail, voici mes appréciations. La qualité des travaux est toujours très bonne et j'ai apprécié les nombreux sujets de gestion. Je vous en félicite. Je reste à votre disposition pour plus de renseignements (gmelard@ulb.ac.be ou ☎ 026504604).

Ma référence: TQG11.**1
 Noms: ***/***
 Sujet: Nombre d'admissions à l'hôpital X
 Dates: 2000.01-2005.12
 Note sur 20: 15

Remarques générales:

Notons d'abord (p. 4) que les données sont disponibles jusque 2005, pas 2006. La méthode de prévision naïve est en fait la méthode de prévision naïve 2 (PRECYR du chap. 4).

La décomposition saisonnière est mise en oeuvre dans Excel après une estimation de la régression linéaire simple sur les moyennes et les médianes annuelles mais par la méthode de comparaison aux moyennes mobiles (et non "Ce modèle sera appliqué aux moyennes mobiles centrées d'ordre 12"), en employant moyennes et médianes pour la synthèse (ceci est clair).

La méthode de Winters a été employée avec TSE. Les sorties sont fournies. Beaucoup (trop) d'efforts sont consacrés au choix d'une transformation, même l'inverse des racines carrées est considérée, en contradiction avec le choix du modèle additif pour la décomposition saisonnière. J'aurais éventuellement pu accepter la transformation racine carrée qui est finalement employée parce qu'elle est parfois recommandée pour des données de comptage (argument basé sur des propriétés de la loi de Poisson). Notons un diagramme de corrélation avec retard. La conclusion n'est pas très claire en ce qui concerne le choix d'une méthode de prévision.

Discussion des données	Sujet d'entreprise très bien documenté 3 pp.
Tableau ou fichier des données fourni	Graphique, fichier
Graphiques	Assez nombreux, bien réalisés mais parfois abscisses pas détaillées
2 méthodes employées au moins	Chap. 2, 4, 5, 6, 8
Utilisation des critères	Oui (5)
Emploi d'une méthode simple	Décomposition saisonnière
Emploi d'une méthode plus avancée	Lissage exponentiel de Winters
Valeurs extrêmes et interventions si nécessaire	Oui, plusieurs événements mentionnés
Modèles correctement paramétrés	Très bien
Choix du modèle additif/multiplicatif, transformation	Modèle additif assez bien justifié
Analyse des résidus	Distribution des erreurs et autocorrélations pour Winters
Prévisions ex post bien réalisées	Oui (12 mois)
Intervalles de prévision.	Oui à 80%, décomposition saisonnière et méthode de Winters
Ecriture correcte des équations des modèles	Pas pour la méthode de Winters
Rapport bien structuré	Pas toujours très logique
Présentation matérielle	Très correcte. 22 pp.
Présence de conclusions	Bonnes (1 p.) avec comparaison des méthodes (3)
La démarche vue au cours est bien comprise.	Presque très bien mais léger

GEST-D-615. Techniques quantitatives de gestion I, Prof. G. Mélard

Dans le contexte de la discussion de votre travail, voici mes appréciations. La qualité des travaux est toujours très bonne et j'ai apprécié les nombreux sujets de gestion. Je vous en félicite. Je reste à votre disposition pour plus de renseignements (gmelard@ulb.ac.be ou ☎ 026504604).

Ma référence: TQG11.**2
 Noms: ***/***
 Sujet: Ventes de tabac en France
 Dates: 2005.01-2009.12
 Note sur 20: 12

Remarques générales:

Problèmes pour trouver des données mais finalement sujet assez riche. Décomposition saisonnière : tout est fait par TSE donc c'est correct. Je cherchais des détails sur la méthode employée mais je les ai trouvés dans le fichier problème : comparaison à la moyenne mobile et synthèse par moyennes élaguées. Vous auriez pu essayer aussi autre chose, compte tenu de la facilité. Gros problème pour tout le travail : toutes les données sont employées donc pas de prévisions ex post comme exigé et aussi pas de calculs des critères. Il aurait fallu définir une date de fin d'estimation, plus petite, par exemple 2008.12, ce qui aurait laissé 12 mois pour la comparaison.

C'est très bien d'avoir établi des liens avec l'économie p. 8. Le problème de votre approche de régression est qu'elle nécessite de prévoir les variables explicatives, en fait la vente des médicaments. En effet, on peut considérer que le prix et les actions de communication sont choisies par l'autorité publique. Vous auriez dû essayer le modèle sans les ventes de médicaments. Aussi le fait de considérer les données du mois précédent surtout pour la communication où il doit y avoir un effet de retard. J'ai beaucoup regretté l'absence d'interprétation des coefficients et de leur signe et des statistiques de Student qui peuvent justifier l'approche.

Méthode de Winters. Vous ne donnez pas les coefficients estimés. Je ne suis pas certain que l'estimation des paramètres se soit bien déroulée. Cela vaudrait la peine de réessayer avec d'autres valeurs initiales. Absence de conclusion sur la comparaison des méthodes. Travail très améliorable.

Discussion des données	Sujet basé sur des données publiques, assez bien décrit 1,5 pp.
Tableau ou fichier des données fourni	Tableau, graphique, fichier
Graphiques	Assez nombreux, bien réalisés mais parfois abscisses pas assez détaillées
2 méthodes employées au moins	Chap. 5, 6, 7
Utilisation des critères	Non, pas sur les prévisions
Emploi d'une méthode simple	Décomposition saisonnière
Emploi d'une méthode plus avancée	Lissage exponentiel de Winters et régression linéaire multiple
Valeurs extrêmes et interventions si nécessaire	Oui, discussion des variations
Modèles correctement paramétrés	Méthodes sur série complète.
Choix du modèle additif/multiplicatif, transformation	Modèle additif avec justification
Analyse des résidus	Oui, graphes
Prévisions ex post bien réalisées	Non, seulement des prévisions ex ante pour 2 méthodes
Intervalles de prévision.	Oui à 80%, pour toutes les méthodes mais pas commentés
Ecriture correcte des équations des modèles	Non, pas d'équations
Rapport bien structuré	Prometteur mais un peu décevant
Présentation matérielle	Très correcte. 17 pp.
Présence de conclusions	Décevantes, pas de comparaison effective des méthodes
La démarche vue au cours est bien comprise.	Presque suffisant mais pas de prévisions ex post

GEST-D-615. Techniques quantitatives de gestion I, Prof. G. Mélard

Dans le contexte de la discussion de votre travail, voici mes appréciations. La qualité des travaux est toujours très bonne et j'ai apprécié les nombreux sujets de gestion. Je vous en félicite. Je reste à votre disposition pour plus de renseignements (gmelard@ulb.ac.be ou ☎ 026504604).

Ma référence: TQG11.**3
 Noms: ***/***
 Sujet: Volume du fret aérien aux aéroports de Paris
 Dates: 1998.07-2010.06
 Note sur 20: 17

Remarques générales:

Retard 1 jour. La plus grande partie effectuée avec TSE donc pas de problème de calcul. Décomposition saisonnière avec Excel : je ne trouve pas les résultats annoncés pour les coefficients saisonniers et le classeur est très confus. Je ne retrouve pas le contenu de la figure 4 dans les sorties de TSE mais c'est très proche. C'est bien d'avoir employé la fonction centile.

Lissage exponentiel de Holt et de Winters : c'est bien mis en oeuvre avec un modèle additif apparemment pour le second.

La modélisation ARIMA a été effectuée assez correctement mais bizarrement en logarithmes alors que la décomposition saisonnière employait un modèle additif. Le modèle a été estimé avec une intervention en janvier 2009, sans montrer le modèle "airline" intermédiaire (fichier ANSECH4.STA). Je n'ai pas compris pourquoi cette intervention-là a été effectuée, alors que décembre 2006 est une valeur extrême plus importante. Néanmoins, c'est très bien, la seule réserve étant que l'équation donnée p. 16 ne contient pas l'intervention (il aurait fallu écrire $y_t - I_t$ au lieu de y_t , avec $I_t = 0$ pour tout t sauf en $t = \text{janvier } 2009$ où elle vaut -6.13). Notez que l'erreur de prévision la plus importante en mai 2010 peut être due à la reprise après l'éruption du volcan islandais Eyjafjöll.

Discussion des données	Données publiques assez bien décrites 2 pp.
Tableau ou fichier des données fourni	Graphique, fichier
Graphiques	Assez nombreux, très bien réalisés avec annotations sauf parfois abscisses irrégulières
2 méthodes employées au moins	Chap. 1, 5, 6, 8-11
Utilisation des critères	Oui (3)
Emploi d'une méthode simple	Décomposition saisonnière
Emploi d'une méthode plus avancée	Modèles ARIMA avec intervention
Valeurs extrêmes et interventions si nécessaire	Intervention
Modèles correctement paramétrés	Oui (à une méthode près)
Choix du modèle additif/multiplicatif, transformation	Modèle additif sans justification; modèle ARIMA en logarithmes
Analyse des résidus	Oui, corrélogrammes prévision naïve et Winters
Prévisions ex post bien réalisées	Oui (12 mois)
Intervalles de prévision.	Oui à 95%, décomposition saisonnière et méthodes suivantes. Précisé une seule fois.
Ecriture correcte des équations des modèles	Oui
Rapport bien structuré	Très bonne qualité
Présentation matérielle	Très bonne 17 pp.
Présence de conclusions	Bonnes (0,5 p.) mais pas de comparaison synthétique des méthodes
La démarche vue au cours est bien comprise.	Très bien à excellent