

Une étude comparative de logiciels de prévision automatique de séries chronologiques

Valentina STAN

Conservatoire National des Arts et Métiers,
292 Rue Saint Martin, F 75141 Paris Cedex 03, France
valentina_titu@yahoo.fr

Résumé : L'objectif de cette étude est de comparer 6 logiciels de prévision automatique de séries chronologiques. On analyse leurs performances sur 50 séries, selon 6 critères à l'aide de plusieurs analyses en composantes principales. Les séries utilisées proviennent de domaines différents : macroéconomie, industrie, finance, démographie, environnement.

Mots clés : prévision automatique, séries chronologiques, logiciels.

Abstract: The objective of this study is to compare 6 software, amongst the most used, which automatically estimate time series. For this software are analyzed the performances by using 50 time series, according to 6 criteria, with several principal components analyses. The used series result from different domains: macroeconomic, industry, finance, demography and environment.

Keywords: automatic forecasting, time series, software.

1. Brève présentation des logiciels testés

Cette étude, réalisée dans le cadre du CNAM-Paris sous la direction de G. Saporta, a pour objectif de comparer 6 logiciels de prévision automatique de séries chronologiques :

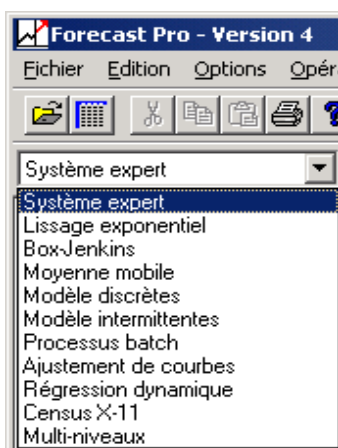
- quatre sont parmi les plus connus du marché : Forecast Pro 4.2, SAS 8.2, SPSS Decision Time 1.1, Statgraphics 5.1; les méthodes intégrées dans ces logiciels sont les méthodes classiques de prévision.
- le logiciel ForeScope+ 1.0, qui est encore en développement;
- le nouveau logiciel KTS 3.04, qui utilise des techniques proches des réseaux de neurones.

1.1. Forecast Pro 4.2

Forecast Pro est un «outil» conçu pour réaliser des prévisions dans le domaine commercial, marketing ou des études économiques.

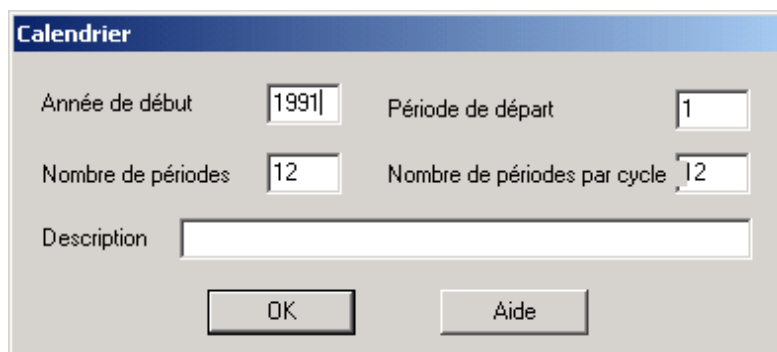
Le système expert du logiciel ajuste automatiquement les paramètres de chacun des modèles en compétition. Il choisit ensuite le modèle le mieux ajusté aux données.

Les méthodes intégrées sont:



Remarques :

- a. L'utilisateur doit spécifier si les données sont ou non saisonnières. «Nombre de périodes» indique le nombre de périodes par an : pour des données mensuelles par exemple, ce sera 12. Le nombre de périodes par cycle indique les périodes qu'il convient de comparer : pour des données mensuelles, le logiciel va comparer les mois de janvier, les mois de février, les mois de mars....



- b. Le logiciel réalise automatiquement la transformation de la série, lorsque cela est nécessaire.
- c. Le temps d'apprentissage est très court.
- d. On peut aussi obtenir des prévisions pour des séries très courtes.
- e. Le critère utilisé par défaut pour choisir la meilleure méthode de prévision est MAD (Mean Absolute Deviation).

1.2. ForeScope+ 1.0

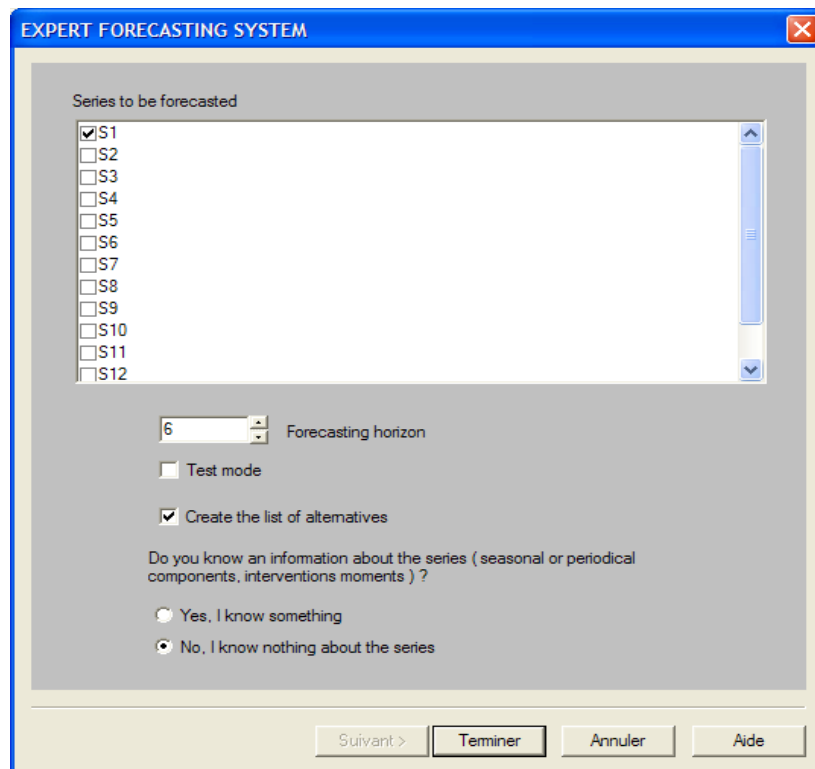
C'est un logiciel encore en cours de développement. Les méthodes de prévision proposées sont les suivantes :

- Prévision naïve;
- Marche aléatoire avec dérive;
- Moyenne mobile, de forme : $X(n+1) = A * MA(k) + B$;
- Méthodes de lissage exponentiel : Holt, Brown;
- Lissage exponentiel avec tendance et saisonnalité (Theil-Wage si additif et Winters si multiplicatif);
- Décomposition selon la tendance et saisonnalité, en modélisant la tendance par lissage exponentiel et par autorégression;
- Méthode de Box-Jenkins.

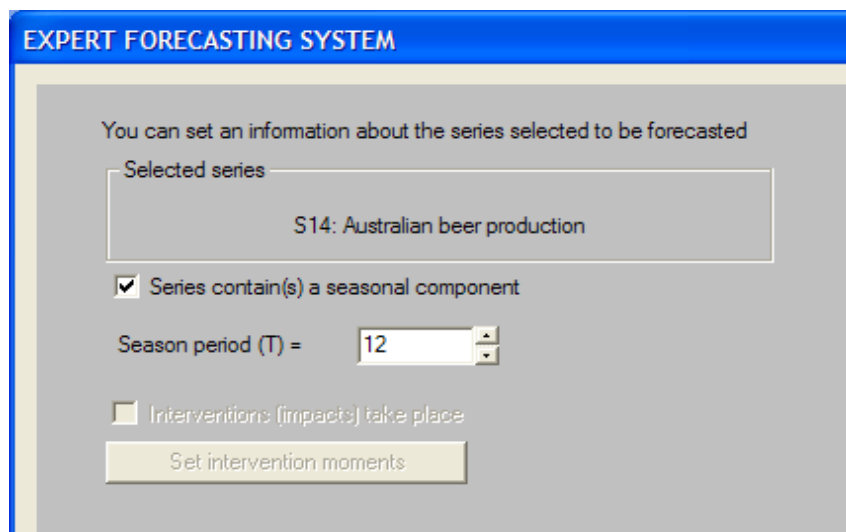
Remarques :

A l'aide du module «Expert», l'utilisateur peut obtenir des prévisions réalisées automatiquement par le logiciel.

- a. Dans la boîte de dialogue suivante, si «Test mode» est sélectionné, les dernières N observations de la série ne sont pas utilisées dans le processus de prévision, mais pour évaluer la qualité des prévisions :



- b. L'utilisateur a la possibilité d'introduire des informations sur la série, comme par exemple, la saisonnalité et les différents événements qui ont influencé cette série :



- c. Si nécessaire, Fore Scope applique automatiquement des transformations aux séries.
- d. Il réalise des prévisions pour les séries très courtes.
- e. Le temps d'apprentissage est très court.
- f. La procédure pour choisir la meilleure méthode est la suivante :
- Le logiciel réalise une analyse préliminaire de la série pour détecter les périodicités (les cycles et les saisonnalités) et la tendance, en utilisant l'analyse spectrale et la fonction d'autocorrélation ;

- chaque méthode est appliquée aux séries; pour les modèles saisonniers, toutes les périodicités trouvées sont utilisées comme facteur potentiel saisonnier. Dans la même catégorie de méthodes, la sélection du meilleur «prédicteur» est fondée sur la maximisation du coefficient de détermination ajusté. Pour choisir la meilleure méthode de prévision on utilise la formule suivante :

$$R = (RMSE + LAST_RMSE(k)) \times \sqrt{\frac{N}{(N-P)}}, \text{ où}$$

LAST_RMSE(k) est le RMSE (erreur quadratique moyenne) calculé pour les k dernières observations des séries,

N = longueur de la série;

P = nombre de paramètres du modèle.

Fore Scope choisit le modèle pour lequel la valeur de R est minimale.

1.3. KTS – 304

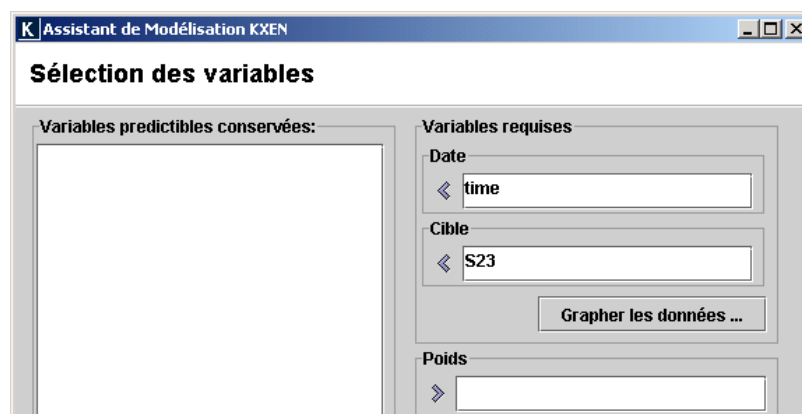
KXEN Time Series (KTS) est un des composants analytiques de la suite KXEN Analytic Framework. KTS construit ses modèles en extrayant les quatre composantes suivantes :

- La tendance - celle-ci est une fonction du temps ou une simple différentiation du signal.
- Les périodicités – décrivent les parties récurrentes dans le signal et peuvent être de 2 types : les cycles (parties récurrentes indépendantes de la date) ou les saisonnalités (parties récurrentes dépendant de la date).
- Les fluctuations - représentent les perturbations qui affectent une série chronologique et qui dépendent de son état passé : c'est le "phénomène de mémoire". Ces fluctuations sont expliquées par un modèle de type autorégressif.
- Le résidu - la partie non-expliquée de l'information contenue dans le signal.

La prévision donnée par KTS est donc la somme des trois premières composantes : la tendance, les périodicités et les fluctuations. Les modèles comparés par KTS sont une combinaison d'une ou plusieurs de ces composantes.

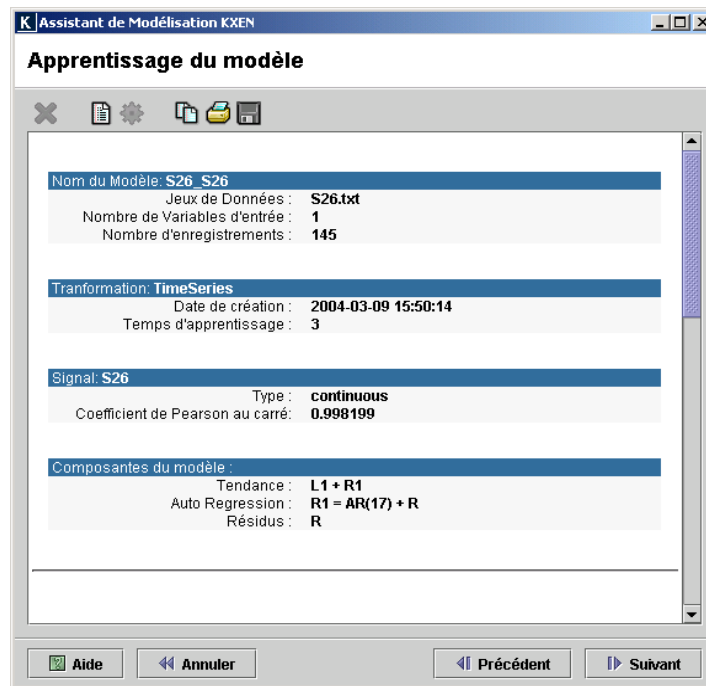
Remarques :

- a. KTS cherche automatiquement la ou les périodicités (cycles et/ou saisonnalités) contenues dans le signal. Cependant, une variable «Date» doit être spécifiée pour chacune des séries :



- b. La rubrique composante du modèle décrit les différentes parties extraites du signal :

Exemples :



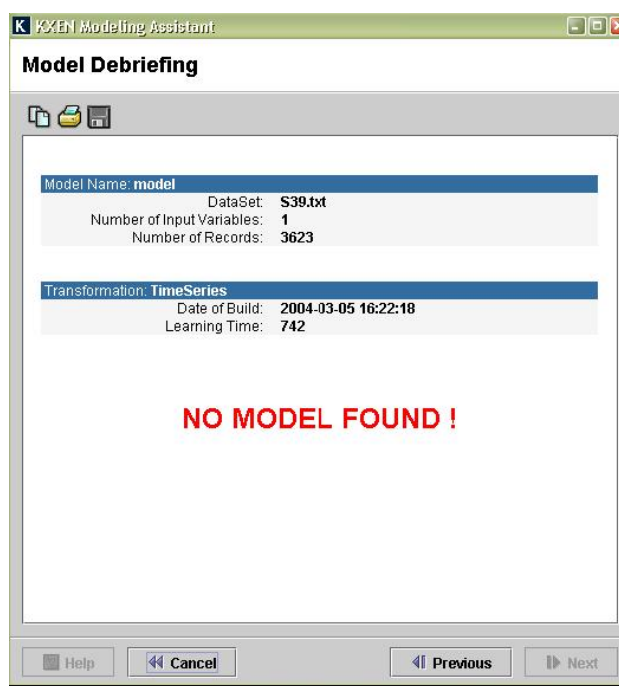
Nom du Modèle: S23_S23	
Jeux de Données :	S23.txt
Nombre de Variables d'entrée :	1
Nombre d'enregistrements :	168
Transformation: TimeSeries	
Date de création :	2004-03-09 15:42:18
Temps d'apprentissage :	6
Signal: S23	
Type :	continuous
Coefficient de Pearson au carré :	0.835616
Composantes du modèle :	
Tendance :	linear(time) + R1
Comp. Cyclique :	R1 = Cyclic(12(NotSmoothed)) + Cyclic(32(NotSmoothed)) + R
Résidus :	R

K2R est le moteur de «Robust Regression» de KXEN ; le chiffre entre parenthèses pour l'AR est son ordre (le nombre de données passées utilisées). Par exemple :

- AR(3) correspond à $X_t = a_1X_{t-1} + a_2X_{t-2} + a_3X_{t-3}$; (X_t est le signal auquel on a extrait les composantes tendances et périodes si cela était possible);
 - L1 correspond à une marche aléatoire; il correspond également au modèle dit «naïf» d'une série chronologique;
- c. KTS permet d'introduire des variables événements prédictibles qui ont pu et qui pourraient affecter la série chronologique.
- d. Le temps d'apprentissage pour les séries très longues (comme la série testée S34 : 3333 observations ; S35 : 1302 observations ; S37 : 5113 observations) est très grand par rapport aux autres logiciels.
- e. Pour les séries très courtes (12 observations, dont 10 ont été utilisées pour prévoir) le logiciel ne peut pas réaliser des prévisions - c'est une limite "technique". En effet, KTS utilise un ensemble d'apprentissage et un ensemble de validation (les autres logiciels utilisent généralement par défaut toutes les données en apprentissage). L'ensemble de validation permettra à KTS de déterminer quel est le meilleur modèle. Le problème est que

les séries courtes parmi les séries utilisées ne présentent plus que 2 ou 3 données dans l'ensemble de validation, ce qui est trop peu pour valider le choix d'un modèle.

f. Notons que pour la série S38, KTS n'a pas pu trouver de modèle :

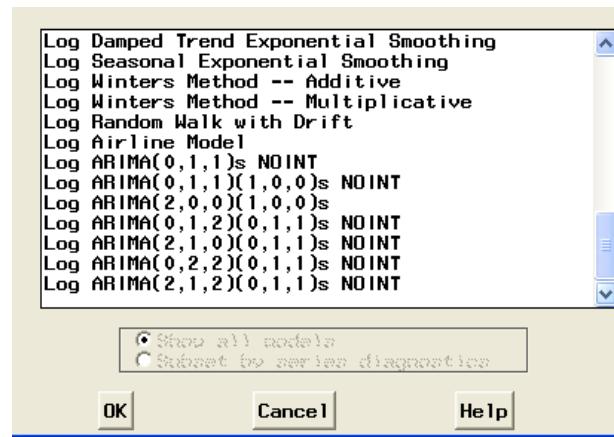
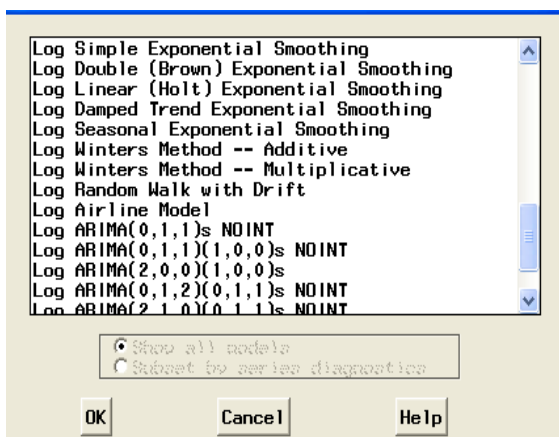
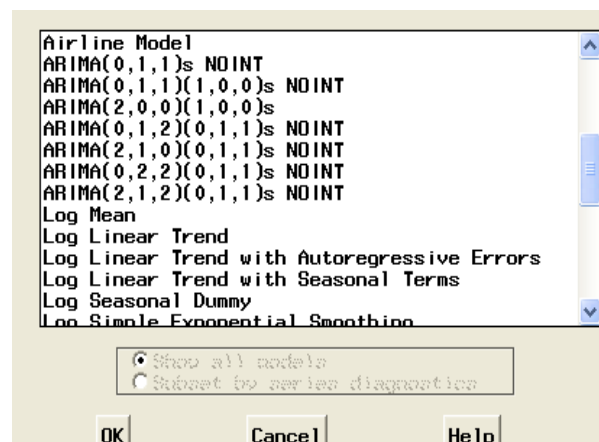
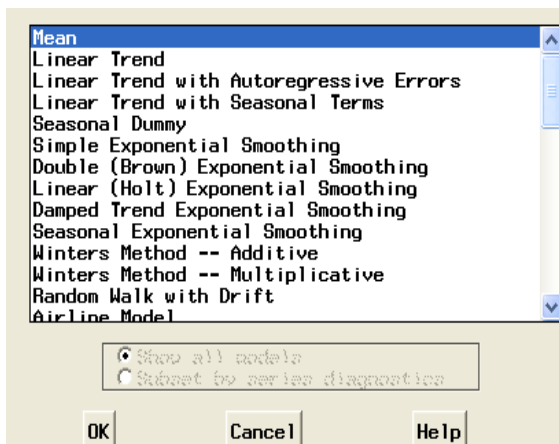


g. Pour choisir le meilleur modèle les critères utilisés par le logiciel sont des critères spécifiques :

- « la confiance sur l'horizon le plus grand possible par rapport à l'horizon de prévision choisi »;
- « l'erreur minimum cumulée sur l'horizon maximum de confiance trouvé précédemment ».

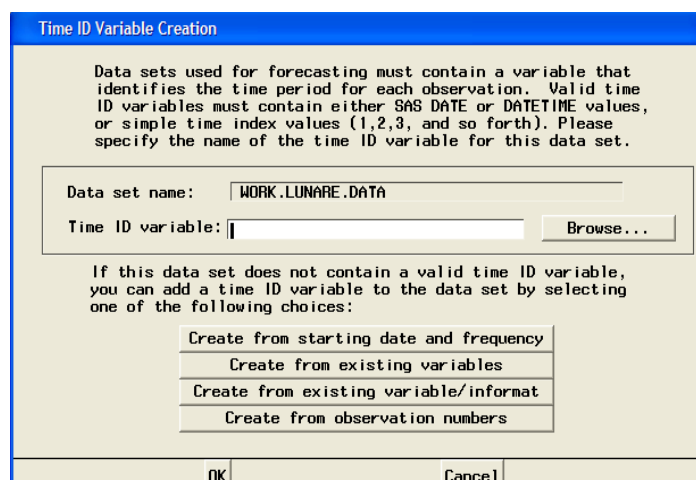
1.4. SAS 8.2

L'option «*Application de prévision de séries temporelles*» a été introduite à partir de la version 8.0, les méthodes disponibles étant les suivantes :



Remarques :

- Pour démarrer, l'utilisateur doit spécifier le type des données utilisées; le logiciel va créer automatiquement une variable «ID» :



Data set name:

Starting Date:

Interval:

New ID variable name:

(NOTE: The system automatically recognizes the time ID variable if you specify the name DATE.)

- b. Lorsque cela est nécessaire, SAS transforme la série initiale et précise le type de transformation utilisée.
- c. Le temps d'apprentissage est très court.
- d. Il réalise aussi des prévisions pour les séries très courtes.
- e. Pour choisir le meilleur modèle de prévision, le critère par défaut est le RMSE (Root Mean Square Error). Cependant, l'utilisateur a la possibilité de choisir lui-même un autre critère (dans cette étude nous avons considéré le critère par défaut) parmi la liste:

Model Selection Criterion

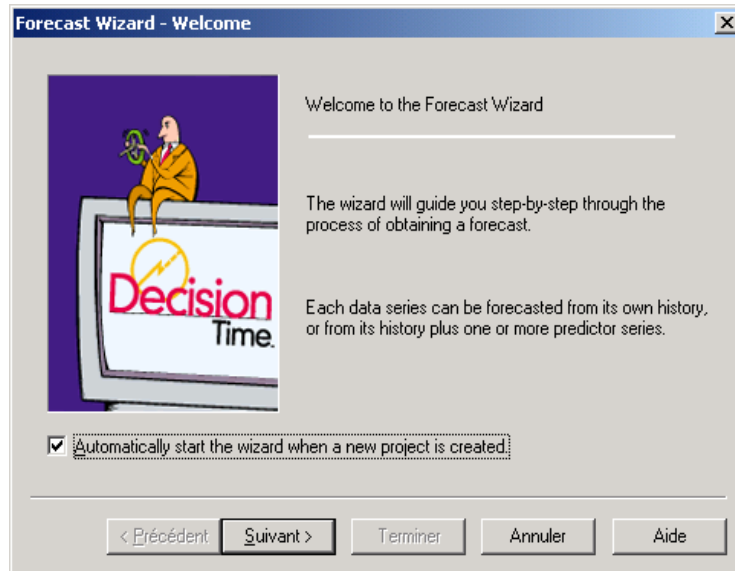
Sum of Square Error
Mean Square Error
Root Mean Square Error
Mean Absolute Error
Mean Absolute Percent Error
Akaike Information Criterion
Schwarz Bayesian Information Criterion
R-Square
Adjusted R-Square
Random Walk R-Square
Amemiya's Adjusted R-Square
Amemiya's Prediction Criterion

Show subset Show all

1.5. SPSS Decision Time 1.1

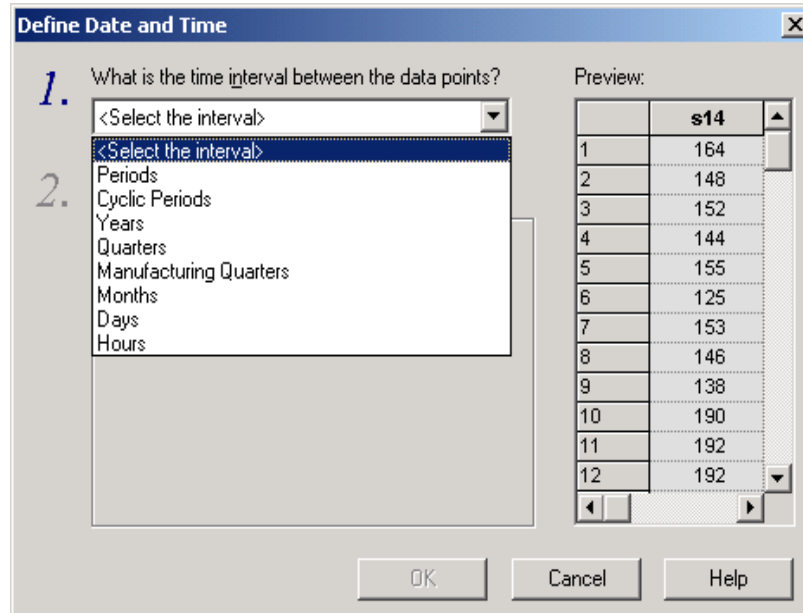
SPSS Decision Time 1.1 est le logiciel de prévision réalisé par SPSS Inc. Il est destiné plus spécifiquement à prévoir des séries microéconomiques. Les méthodes de prévision disponibles vont des classiques lissages exponentiels aux modèles ARIMA.

A l'aide de «Forecast Wizard», l'utilisateur est guidé dans les étapes du processus de prévision :

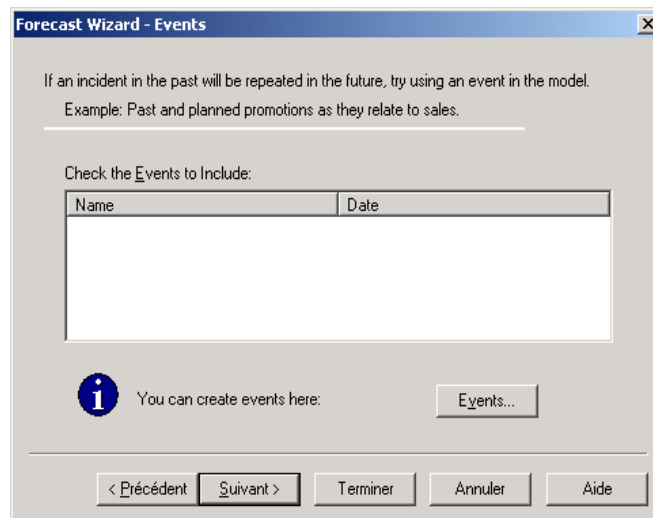
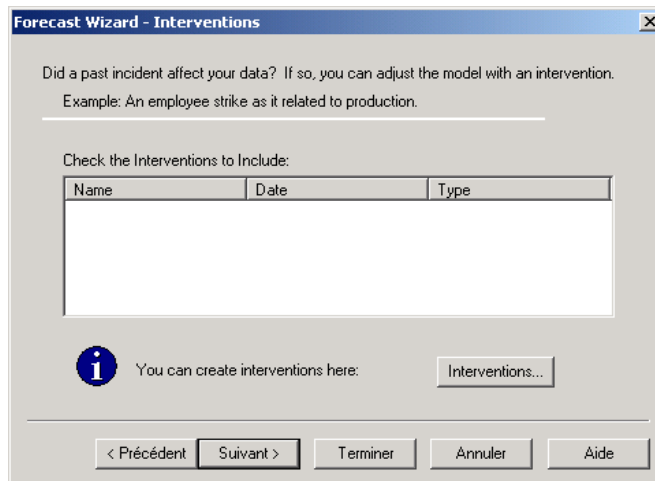


Remarques :

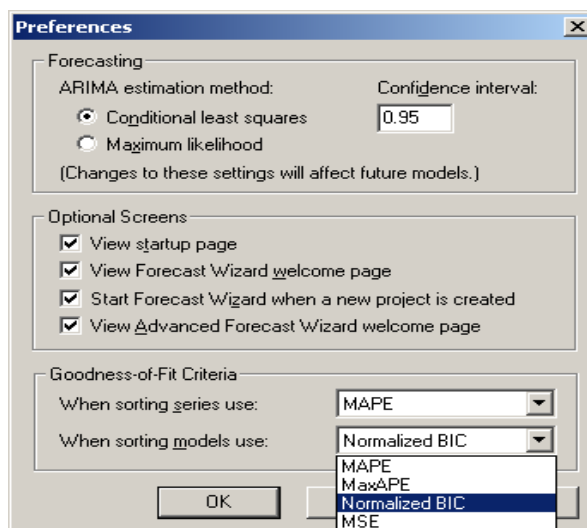
a. L'utilisateur doit indiquer le type de données utilisées :



b. SPSS Decision Time offre la possibilité d'introduire dans le processus de prévision des évènements qui ont affecté la série et qui sont considérés importants par l'utilisateur :

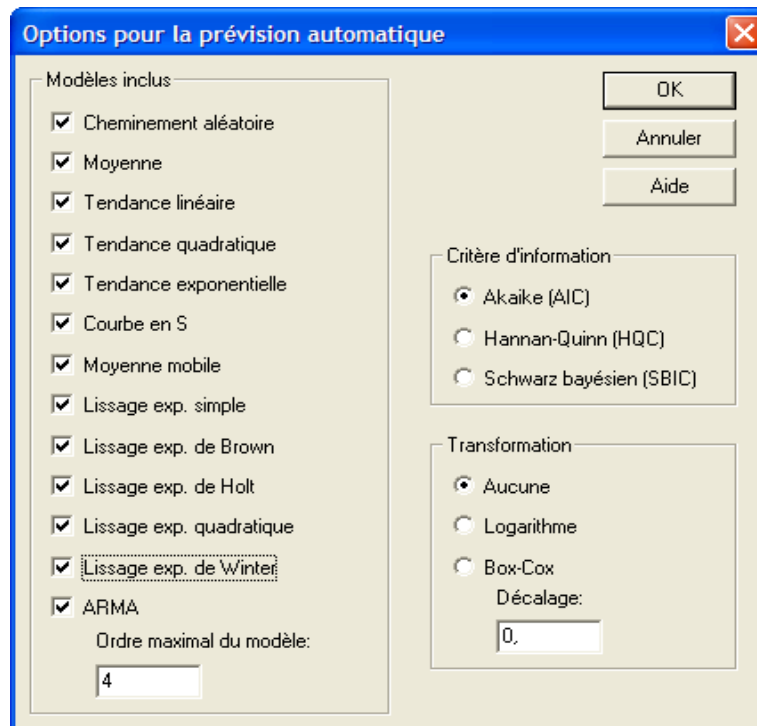


- c. Si nécessaire, SPSS applique automatiquement des transformations aux séries.
- d. Les séries très courtes sont acceptées.
- e. Le temps d'apprentissage est très court.
- f. Pour choisir la meilleure méthode, le critère par défaut est le MAPE (écart absolu moyen en pourcentage), mais l'utilisateur a la possibilité de choisir un autre critère (dans cette étude nous avons utilisé le critère proposé par défaut) :



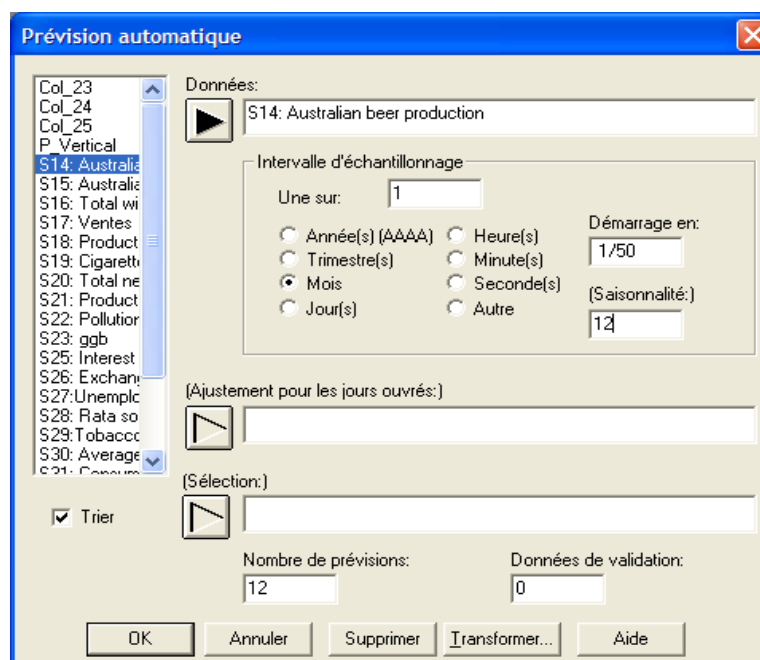
1.6. Statgraphics 5.1

Avec l'option «Prévision automatique» du menu «Analyse de séries temporelles» le logiciel sélectionne automatiquement le meilleur modèle pour la série chronologique à prévoir. Les méthodes disponibles pour réaliser cette opération sont les suivantes :



Remarques :

- L'utilisateur doit indiquer si la série est ou non saisonnière :



- Parmi les logiciels utilisés, Statgraphics est le seul qui présente une comparaison des modèles existants :

Exemple pour la série S14 :

Modèles

-
- (A) Cheminement aléatoire
 - (B) Moyenne constante = 151,203
 - (C) Tendence linéaire = $158,965 + -0,344997 t$
 - (D) Tendence quadratique = $160,97 + -0,606533 t + 0,00581191 t^2$
 - (E) Tendence exponentielle = $\exp(5,06792 + -0,0022725 t)$
 - (F) Courbe en S = $\exp(5,00207 + 0,148122 / t)$
 - (G) Moyenne mobile simple de 3 termes
 - (H) Lissage exponentiel simple avec $\alpha = 0,1185$
 - (I) Lissage exponentiel de Brown avec $\alpha = 0,066$
 - (J) Lissage exponentiel de Holt avec $\alpha = 0,1369$ et $\beta = 0,1541$
 - (K) Lissage exp. quadratique de Brown avec $\alpha = 0,0487$
 - (L) Lissage exponentiel de Winter avec $\alpha = 0,1115$, $\beta = 0,9904$, $\gamma = 0,2708$
 - (M) ARMA(0,0) SARMA(0,0)
 - (N) ARMA(1,0) SARMA(1,0)
 - (O) ARMA(2,1) SARMA(2,1)
 - (P) ARMA(3,2) SARMA(3,2)
 - (Q) ARMA(4,3) SARMA(4,3)

Période d'estimation

Modèle	RMSE	MAE	MAPE	ME	MPE	AIC
(A)	12,7677	10,1497	6,92642	-0,424332	-0,6068	5,09384
(B)	9,14034	6,66621	4,49649	-0,0135941	-0,368958	4,47085
(C)	8,06273	6,1717	4,20827	-0,0133376	-0,285197	4,26541
(D)	8,12847	6,06037	4,13947	0,0049368	-0,283499	4,32711
(E)	8,06092	6,17055	4,20201	0,199932	-0,143246	4,26496
(F)	8,4939	6,11346	4,12108	0,248932	-0,155617	4,36961
(G)	9,10787	6,83894	4,70417	-0,53257	-0,629646	4,41828
(H)	8,73201	6,78757	4,67177	-1,92211	-1,58701	4,37945
(I)	8,81841	6,89174	4,75603	-2,11021	-1,69528	4,39914
(J)	8,81992	6,61529	4,51483	0,00194716	-0,295216	4,44493
(K)	8,93648	7,0335	4,85039	-2,1815	-1,73102	4,42574
(L)	10,1517	7,63795	5,23003	0,968574	0,314189	4,77165
(M)	19,3721	15,1198	9,93289	1,2919E-14	-1,5274	5,97312
(N)	10,0758	7,31007	5,02042	-2,18181	-1,70985	4,75664
(O)	8,79284	6,04064	4,11601	0,17432	-0,0794479	4,66606
(P)	4,14392	1,74728	1,22185	-0,00165055	-0,0769192	3,34329
(Q)	2,64037	0,960803	0,690077	0,00248515	-0,0333923	2,62366

- c. Statgraphics ne réalise pas automatiquement la transformation des séries ; c'est l'utilisateur qui doit indiquer le type de transformation.
- d. Le logiciel réalise des prévisions même pour les séries très courtes.
- e. Le temps d'apprentissage est très court.
- f. Le critère de choix de la meilleure méthode est le minimum du critère d'information d'Akaike (c'est ce critère que nous avons utilisé dans cette étude).

Les principales différences entre les logiciels :

- Statgraphics est le seul qui ne réalise pas automatiquement la transformation de la série, c'est l'utilisateur qui doit l'indiquer. Les autres logiciels cherchent automatiquement la transformation à appliquer à la série parmi les transformations usuelles (logarithme, racine carrée ...).
- Pour Forecast Pro et Statgraphics l'utilisateur doit spécifier si la série est ou non saisonnière ; les autres déterminent automatiquement la saisonnalité.
- Parmi les 6 logiciels, Statgraphics est le seul qui présente une comparaison des modèles.

- ForeScope, KTS et SPSS Decision Time offrent la possibilité d'introduire dans le processus de prévision des événements qui ont affecté la série et qui sont considérés importants par l'utilisateur.

2. Séries utilisées

Pour réaliser cette étude nous avons utilisé 50 séries chronologiques de plusieurs types :

Périodicité des données	Type de séries						Total
	Macro	Industrie	Finance	Démographie	Environnement	Autres	
Annuelle	1	7	0	3	2	0	13
Trimestrielle	2	4	1	0	0	0	7
Mensuelle	4	11	2	2	1	0	20
Journalière	0	0	3	1	1	1	6
Autres	0	0	0	0	0	4	4
Total	7	22	6	6	4	5	50

Il y a des séries très courtes et aussi des séries qui dépassent 1000 observations, dont la plus longue a 5083 observations. La description complète des séries est faite dans l'**Annexe 1**.

La principale source pour sélectionner ces séries est la M3-Compétition, étude coordonnée par Makridakis et Hibon (2000) qui compare des méthodes de prévision pour 3003 séries.

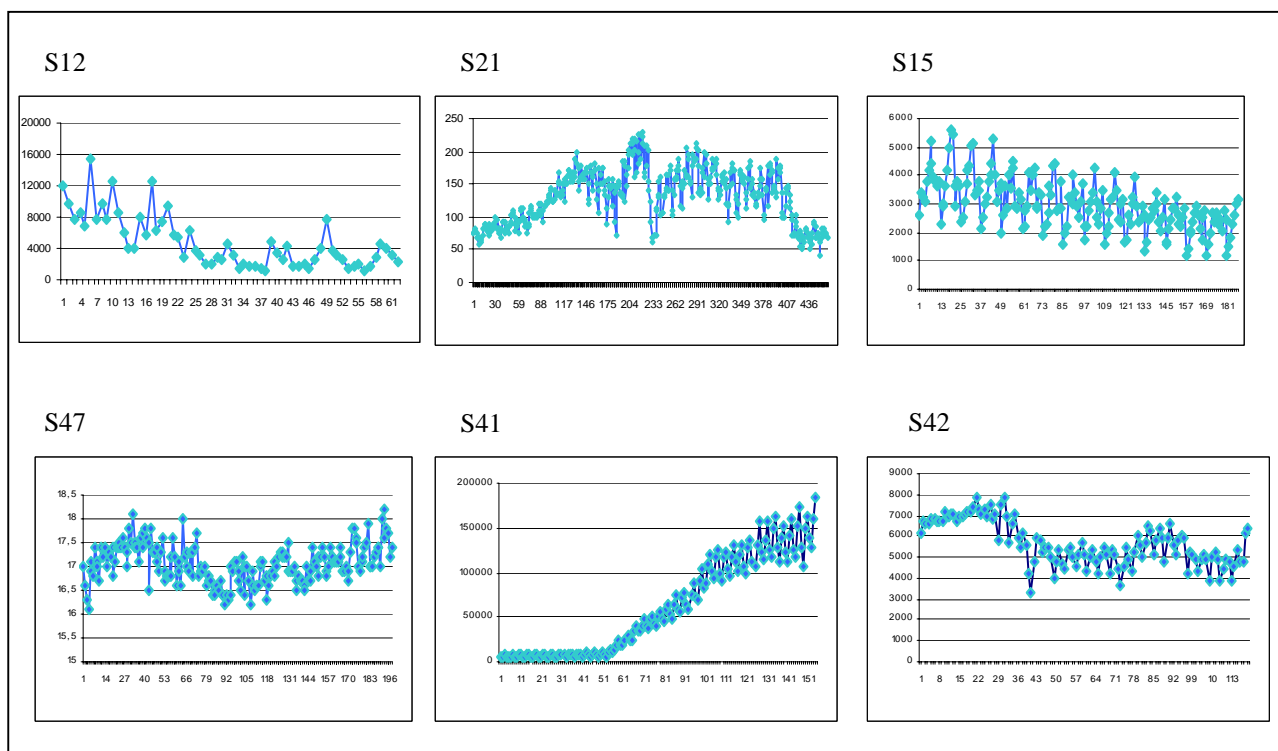


Figure n°1 : 6 exemples parmi les 50 séries utilisées

Pour comparer les résultats obtenus en utilisant chaque logiciel, on a procédé de la manière suivante:

- on supprime les dernières observations :
 - 6 pour les séries annuelles longues et 2 pour les séries courtes;
 - 8 pour les séries trimestrielles;

- 12 en général pour les séries mensuelles, mais 2 pour les séries très courtes;
 - 30 pour les séries journalières;
 - 18 pour les autres types de séries, sauf la série 50 pour laquelle 10 observations sont supprimées.
- chaque logiciel doit prévoir, pour chaque série, les observations supprimées;
 - les valeurs prévues sont comparées avec les valeurs réelles observées.

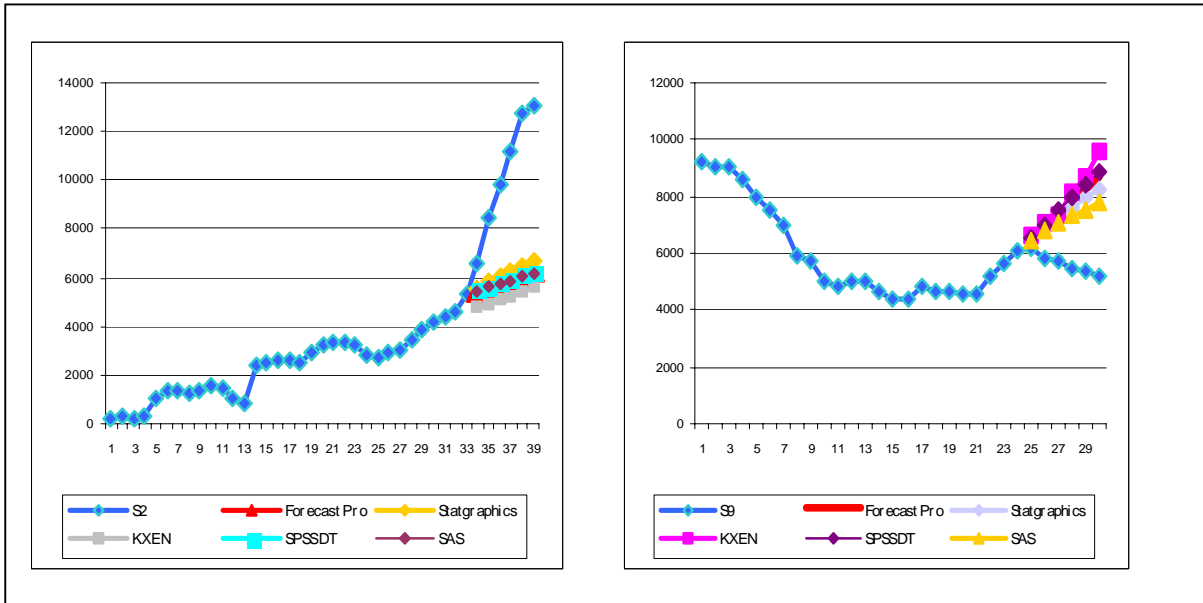


Figure n°2 : Deux exemples de prévisions

3. Critères et normalisations

En notant e_i les erreurs de prévision, les critères suivants ont été utilisés pour évaluer les performances de chaque logiciel :

1) l'écart type : $SD = \sqrt{\frac{1}{n} (e_i - \bar{e})^2}$;;

2) l'écart absolu moyen : $MAD = \frac{1}{n} \sum_{i=1}^n |e_i - \bar{e}|$

3) l'erreur absolue moyenne : $MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$;

4) l'écart absolu moyen en pourcentage : $MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{e_i}{y_i} \right|$;

5) l'erreur quadratique moyenne : $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$;

6) l'erreur maximale = $\max |e_i|$.

Les résultats obtenus pour tous les critères sont dans l'Annexe 2.

Sauf MAPE, dont on verra plus loin qu'il se comporte différemment des autres, ces critères sont sensibles à l'échelle de mesure au niveau des séries et ne peuvent donc être comparés d'une série à l'autre.

Deux normalisations ont été utilisées : soit en divisant par la moyenne de la série, soit en divisant par la moyenne des valeurs du critère obtenues par les différents logiciels pour la même série. Dans cet article on présente seulement les résultats pour le premier type de normalisation.

Exemple: normalisation en fonction de la moyenne de la série

Série	Logiciel	SD	SD normalisé
S1	Forecast Pro 4.2	952,6863	0,3719 (=952,6863/2561,7805)
S1	ForeScope+ 1.0	536,4608	0,2094
S1	KTS - 304	481,1381	0,1878
S1	SAS 8.2	712,2782	0,2780
S1	SPSS - Decision Time 1.1	712,2256	0,2780
S1	Statgraphics 5.1	797,0265	0,3111
Moyenne de la série		2561,7805	

Ensuite, en utilisant les valeurs normalisées, on a considéré pour chaque normalisation deux types d'analyses en composantes principales.

4. Les deux types d'analyses en composantes principales

Comme KTS ne réalise pas de prévisions pour les séries très courtes ni pour la série qui contient beaucoup de valeurs nulles (au total 7 séries), nous avons considéré 2 cas :

- on utilise les 6 logiciels avec les 43 séries qui peuvent être prévues par tous les logiciels;
- on utilise 5 logiciels (en excluant KTS) et les 50 séries.

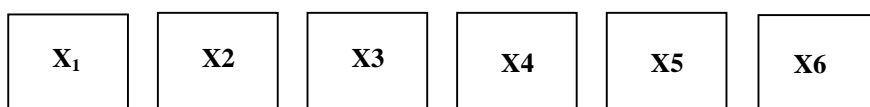
Nous ne présentons ici que les résultats correspondant au premier cas.

Pour chaque logiciel, on dispose donc d'un tableau de dimension : (43 x 6) contenant les valeurs des 6 critères statistiques normalisés. Les 6 logiciels étudiés conduisent donc à 6 tableaux analogues. Pour analyser simultanément ces tableaux on peut les assembler de deux manières différentes :

4.1 Juxtaposition horizontale

Objectif : étudier, dans le cadre de chaque logiciel, les corrélations entre critères.

Les 43 séries sont les individus et les 6 critères croisés avec les logiciels sont les variables continues actives.



4.2 Empilement vertical :

X1

Objectif : étudier les corrélations entre les 6 critères. Il y a donc 6 colonnes et 258 individus (43 séries x 6 logiciels).

:

A chaque série correspond donc 6 points.

X6

Les logiciels sont alors les modalités d'une variable nominale illustrative supplémentaire, ce qui permettra de réaliser un classement pour les logiciels.

5. Résultats

5.1 Juxtaposition horizontale

Les deux premiers axes factoriels extraient 95,41% de l'inertie du nuage de points :

HISTOGRAMME DES 36 PREMIERES VALEURS PROPRES

NUMERO	VALEUR PROPRE	POURCENTAGE	POURCENTAGE CUMULE	
1	29.2824	81.34	81.34	*****
2	5.0669	14.07	95.41	*****
3	0.6682	1.86	97.27	**
4	0.4295	1.19	98.46	**
5	0.1910	0.53	98.99	*

Le cercle de corrélation montre que le critère MAPE a un comportement indépendant des autres et que tous les autres critères sont très semblables entre eux quel que soit le logiciel.

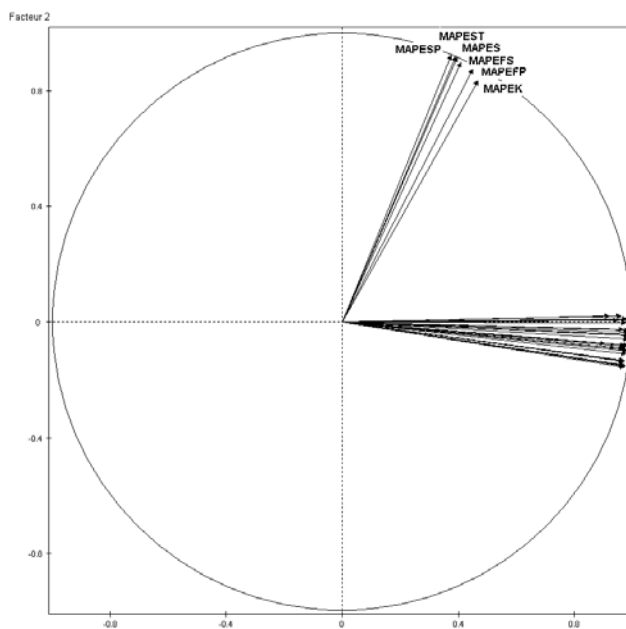


Figure n°3 : Cercle de corrélation

On peut observer qu'il existe des corrélations très fortes entre :

- le premier axe et tous les critères de tous les logiciels, sauf MAPE (les valeurs sont entre 0,9127 et 0,9934);
- le deuxième axe et MAPE (avec des valeurs entre 0,8338 et 0,9183).

Plan principal :

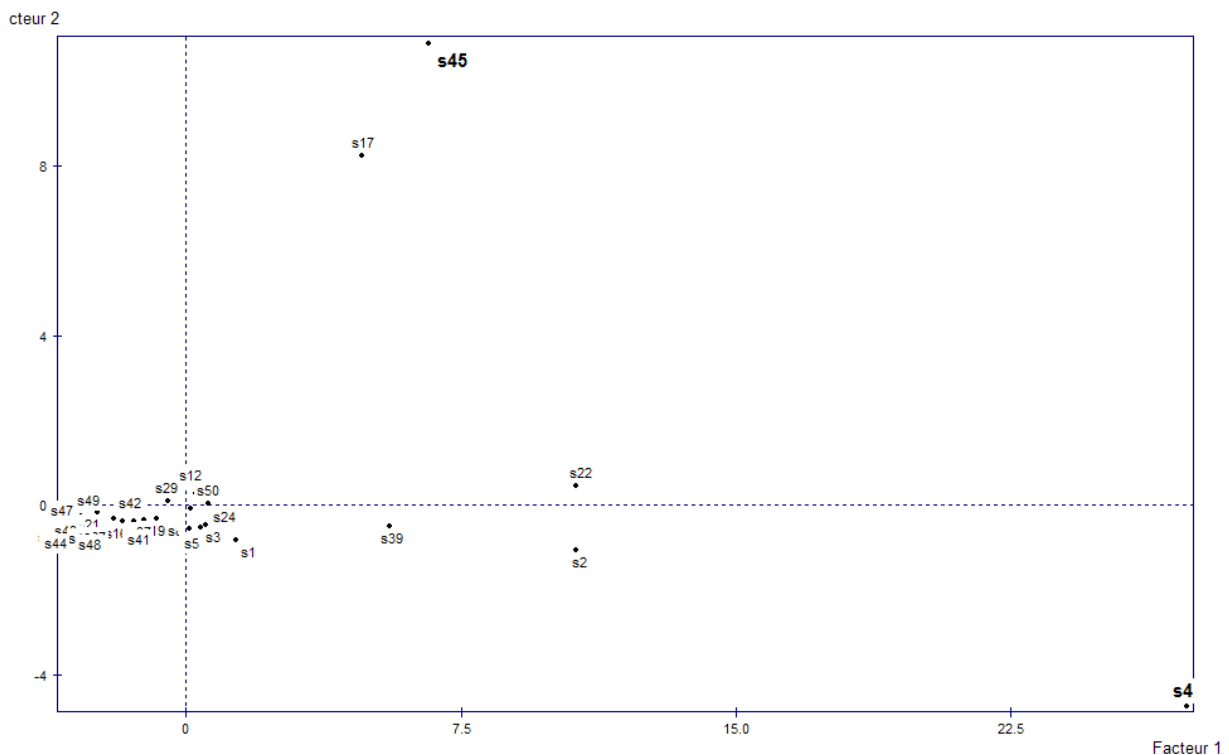


Figure n°4 : Graphique des 43 séries (points individus)

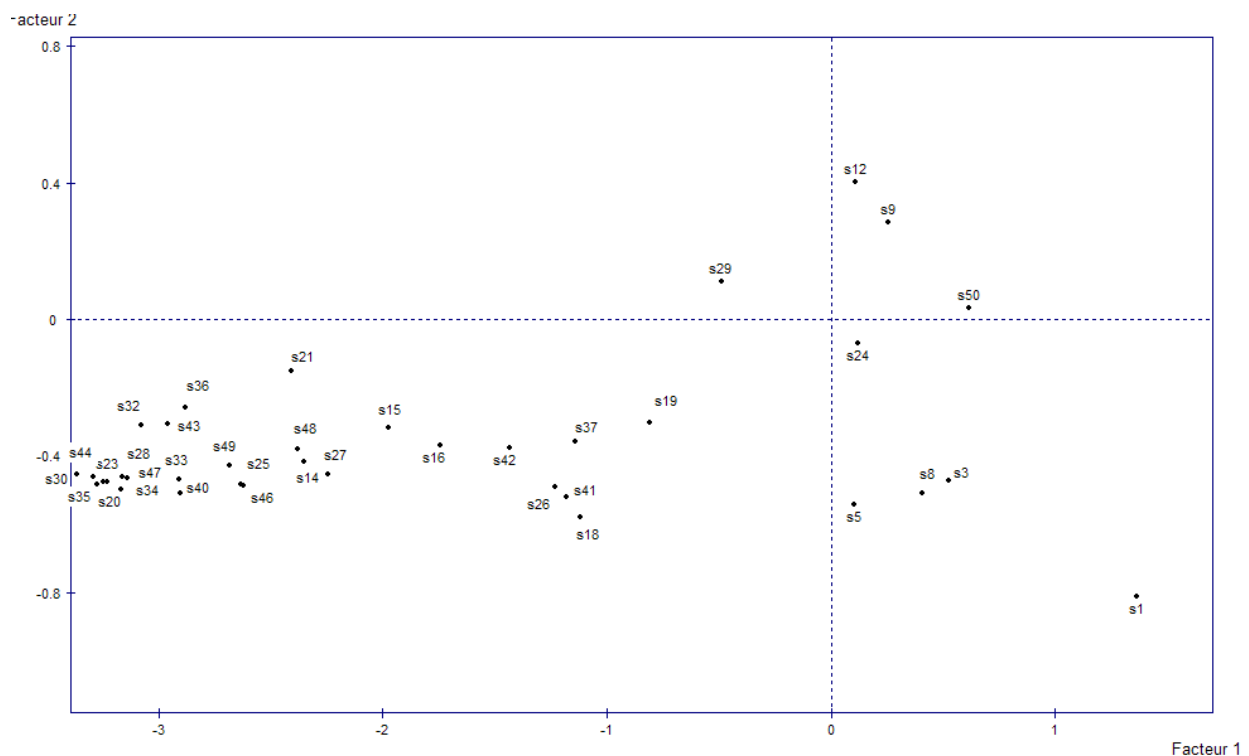


Figure n°5 : Graphique des séries (individus) – zoom au centre

En général, les séries sont proches du centre du graphique; donc les valeurs des critères sont globalement proches des valeurs moyennes.

On constate (fig. 4) que les séries 4 et 45 ont un comportement très différent des autres et contribuent fortement à l'inertie portée par chaque axe. La série qui contribue le plus au premier axe est S4 (59,3% de l'inertie de cet axe est due à cette série); S45 contribue le plus au deuxième axe

(avec 54,5%). Donc, ces deux séries ont une grande influence pour la détermination de ces axes. Cependant, si dans une nouvelle analyse, on considère les deux séries comme individus illustratifs, le cercle de corrélation est le même.

La série 4 («La valeur de marché du stock total de vente sur les échanges enregistrés entre 1947-1987») est une série annuelle sans saisonnalité et qui a une évolution constante jusqu'à une croissance brusque.

La série 45 («Le PIB réel aux Etats Unis - exprimé en \$ 1982 - entre le premier trimestre 1947 et le troisième trimestre 1991») est une série trimestrielle, sans saisonnalité mais qui présente de grandes fluctuations.

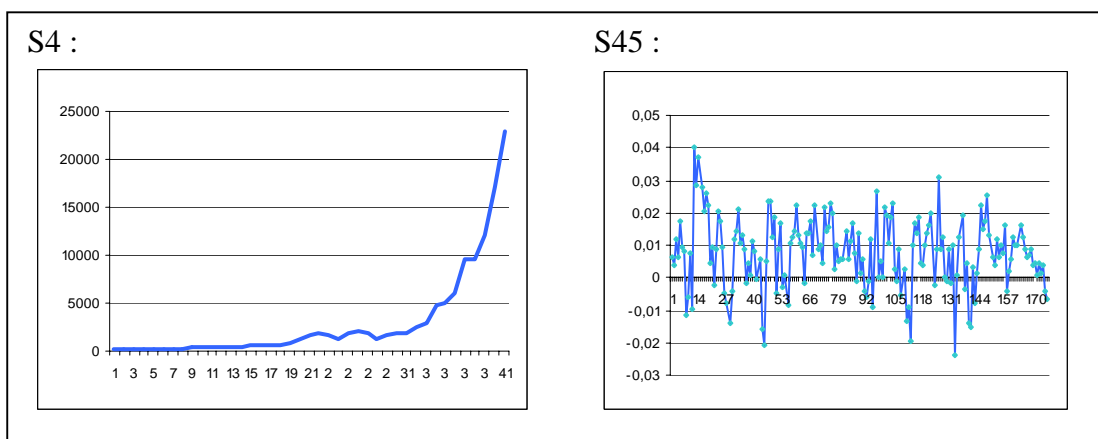


Figure n°6 : Les 2 séries atypiques

5.2 Empilement vertical

MATRICE DES CORRELATIONS						
	SD	MAD	MAE	MAPE	RMSE	EMAX
SD	1.00					
MAD	1.00	1.00				
MAE	0.92	0.92	1.00			
MAPE	0.33	0.31	0.36	1.00		
RMSE	0.95	0.95	1.00	0.36	1.00	
EMAX	0.98	0.98	0.97	0.37	0.99	1.00

HISTOGRAMME DES 6 PREMIERES VALEURS PROPRES

NUMERO	VALEUR PROPRE	POURCENTAGE	POURCENTAGE CUMULE	
1	5.0111	83.52	83.52	*****
2	0.8531	14.22	97.74	*****
3	0.1256	2.09	99.83	***
4	0.0094	0.16	99.99	*
5	0.0005	0.01	100.00	*
6	0.0002	0.00	100.00	*

La première composante principale extrait 83,52% de l'inertie totale.

On constate à nouveau que MAPE a un comportement différent des autres critères.

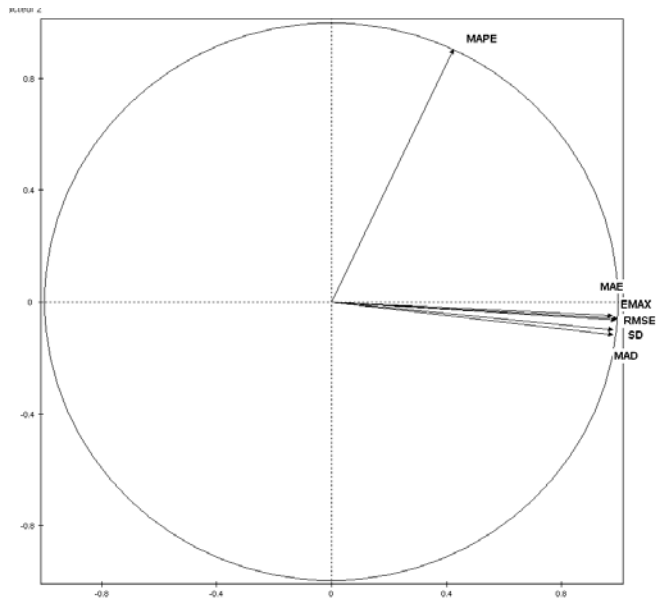


Figure n°7 : Cercle de corrélation

COORDONNEES DES VARIABLES SUR LES AXES 1 A 5
VARIABLES ACTIVES

VARIABLES	COORDONNEES					CORRELATIONS VARIABLE-FACTEUR					ANCIENS AXES UNITAIRES				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
SD	0.98	-0.10	0.18	-0.01	-0.02	0.98	-0.10	0.18	-0.01	-0.02	0.44	-0.11	0.51	-0.10	-0.73
MAD	0.98	-0.12	0.17	-0.03	0.01	0.98	-0.12	0.17	-0.03	0.01	0.44	-0.13	0.48	-0.35	0.65
MAE	0.97	-0.05	-0.22	-0.03	0.00	0.97	-0.05	-0.22	-0.03	0.00	0.44	-0.05	-0.61	-0.33	-0.14
MAPE	0.42	0.91	0.02	0.00	0.00	0.42	0.91	0.02	0.00	0.00	0.19	0.98	0.06	-0.02	0.01
RMSE	0.99	-0.07	-0.13	-0.01	0.00	0.99	-0.07	-0.13	-0.01	0.00	0.44	-0.07	-0.38	-0.09	0.08
EMAX	0.99	-0.06	0.00	0.08	0.00	0.99	-0.06	0.00	0.08	0.00	0.44	-0.06	-0.01	0.87	0.13

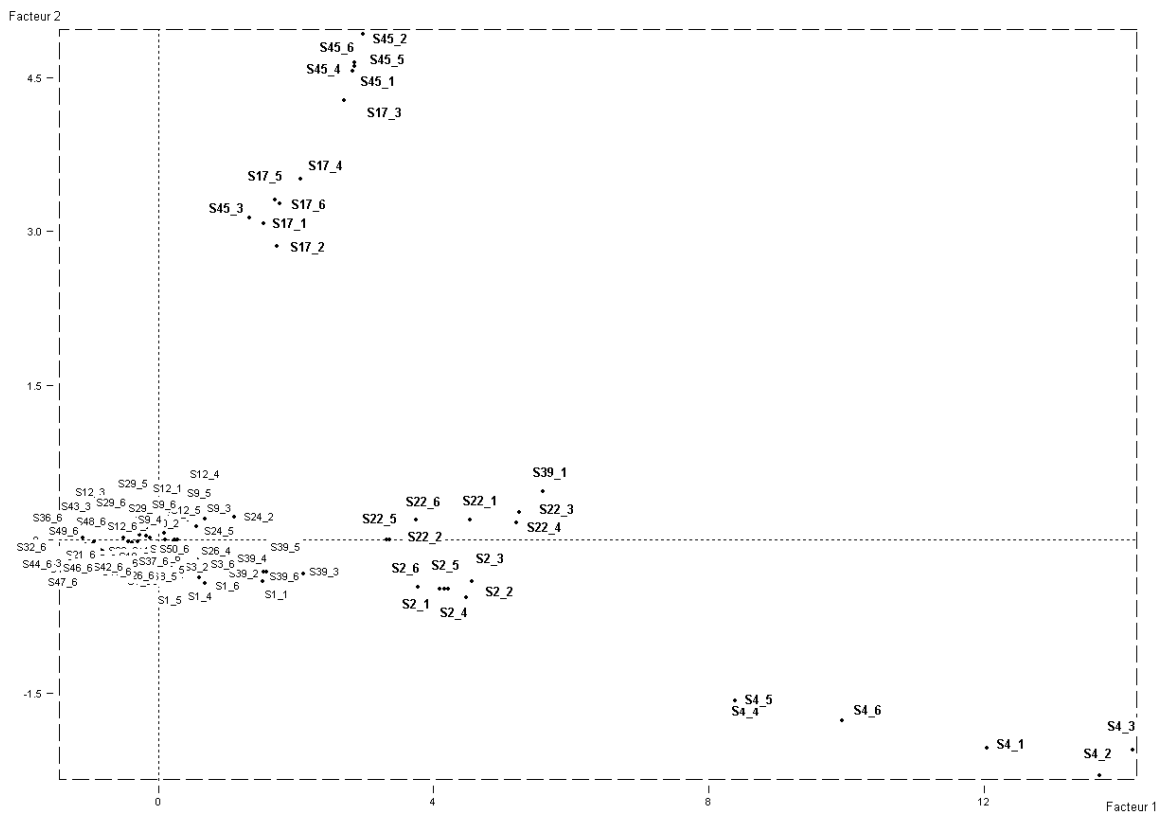


Figure n°8 : Graphique des 43 séries (points individus)

On peut constater que :

- les séries différentes S4 et S45 sont atypiques pour tous les logiciels;
- les sous-nuages associés aux 5 logiciels sont, en général, concentrés.

En représentant pour chaque logiciel le centre de gravité des 43 séries (fig. 9), le classement sur le premier axe donnerait de gauche à droite l'ordre des logiciels par performance moyenne : SPSS, Statgraphics, SAS, Fore Scope, KTS, Forecast Pro.

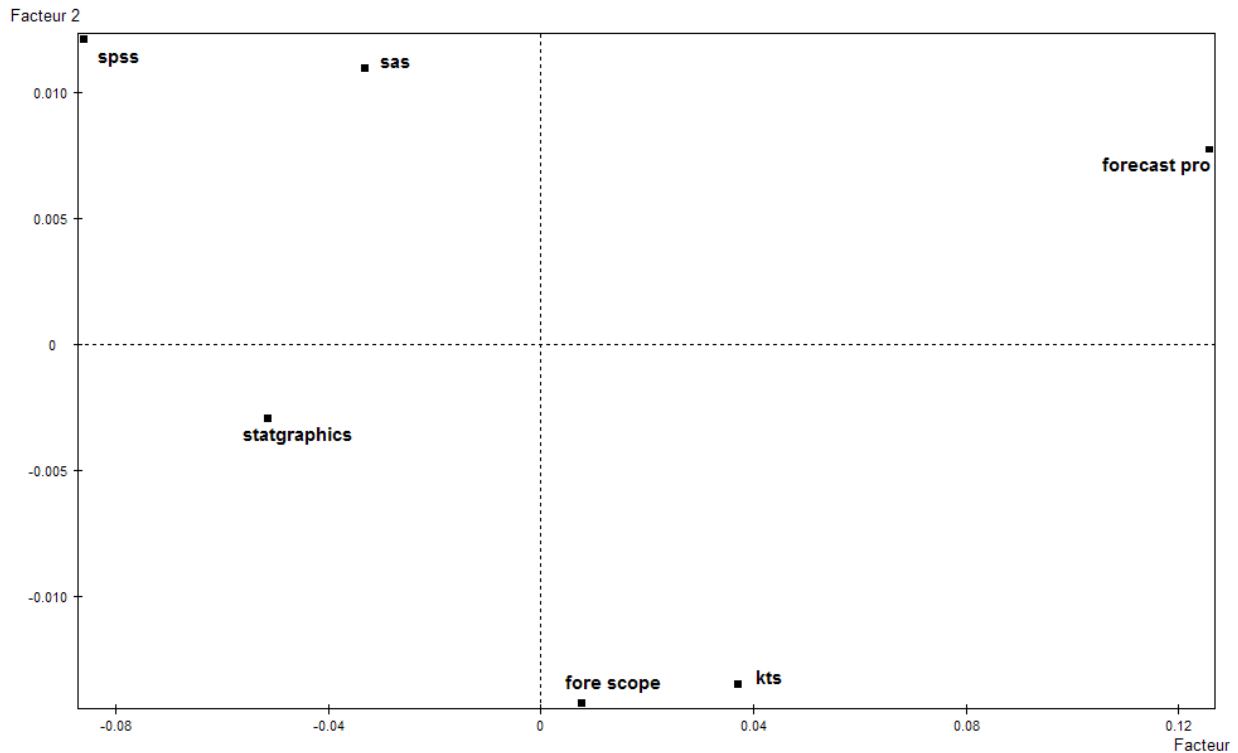


Figure n°9 : Graphique des logiciels analysés - zoom

Mais ce classement est illusoire car les zones de confiance se recouvrent largement (comme le montre la figure 10): on peut donc difficilement conclure à la supériorité d'un logiciel sur les autres.

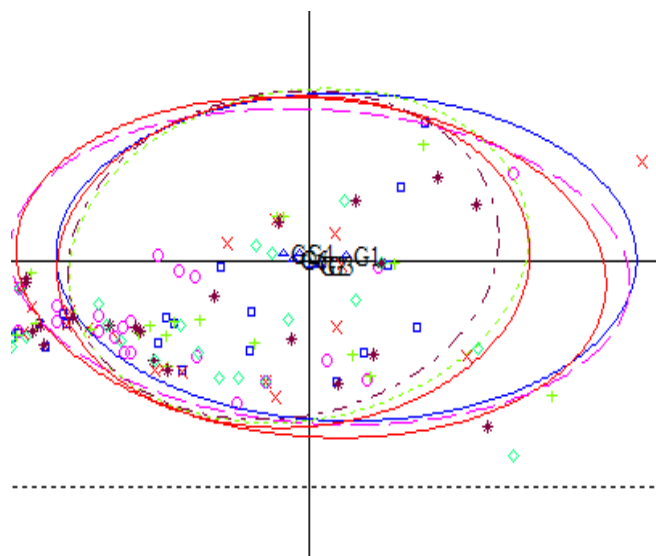


Figure n°10 : Zones de confiance

Remarques en guise de conclusion

- La principale conclusion (qui tient compte des 8 analyses en composantes principales réalisées pour les 2 normalisations) est qu'on ne peut pas fournir un classement global des 6 logiciels testés. En effet leurs performances ne sont pas significativement différentes du point de vue des 6 critères considérés, quelle que soit la normalisation utilisée. Cette remarque pourra aussi rassurer l'utilisateur embarrassé pour faire le choix d'un logiciel.
- Le problème de la confiance dans les prévisions réalisées automatiquement reste entier.
- Les situations sont fréquentes dans lesquelles des modèles complexes ne donnent pas de meilleurs résultats que des modèles simples. Par exemple, pour la série 37, les méthodes simples de lissage exponentiel utilisées par SPSS et SAS donnent des résultats meilleurs que les modèles ARIMA utilisés par Forecast Pro et Statgraphics.

Naturellement ces conclusions ne sont valables, en toute rigueur, que pour l'ensemble des 50 séries utilisées et pour les critères de performance considérés ici.

Bibliographie

Gourieroux, C., Monfort, A., *Séries temporelles et modèles dynamiques*, Economica, Paris, 1995

Makridakis, S., Wheelwright, S. C., Hyndman, R. J., *Forecasting: methods and applications*, Wiley, New York, 1998

Makridakis, S. et Hibon, M. (2000) *The M3-Competition: results, conclusion and implication*, International Journal of Forecasting, vol. 16

Mélard, G., *Méthodes des prévision à court terme*, Ellipses, Bruxelles, 1990

On trouvera des précisions sur les logiciels testés à l'adresse des sites suivant :

<http://www.forecastpro.com>

<http://www.kxen.com>

<http://www.sas.com>

<http://www.spss.com.fr>

<http://www.statgraphics.com>

Les analyses ont été effectuées à l'aide du logiciel SPAD 5.5.