

Analyse et Préviation des Economies Urbaines en Chine (1987-1991)

Huiwen Wang

Ecole de Gestion, BUAA, Beijing, Chine

Visiting Scholar, Groupe HEC, Jouy-en -Josas, France

Résumé

La Régression PLS (Partial Least Squares) a été utilisée pour étudier le développement économique urbain en Chine, et aussi pour établir des modèles de prévision. Les comportements économiques des différentes régions sont décrites d'une manière satisfaisante et les modèles de prévision économiques sont relativement bien adaptés.

Mots Clés: Régression PLS, Economie Urbaine en Chine,

I. Introduction

Jusqu'à une époque récente les villes, surtout les grandes villes, sont les endroits où les industries modernes, le commerce et les services se sont concentrés. Elles jouent un rôle très important pour réunir les éléments essentiels de l'économie et deviennent le centre du développement économique du pays.

La Chine est essentiellement rurale, la proportion de personnes habitant dans des villes est inférieure à la moyenne mondiale. En 1990, la population des villes en Chine représente 29,3% de la population totale du pays.

Pourtant en 1990, la production intérieure brute (PIB) fournie par les villes est de 1376,6 milliards Yuans, soit 57% du PIB total du pays. La production industrielle des villes se monte à 1582,5 milliards Yuans, soit 66,1% de celle du pays. A la fin de 1990 la Chine a 476 villes où se trouve donc concentrée une grande proportion de l'économie nationale. Le développement urbain détermine désormais la vitesse de la construction moderne du pays et il traduit aussi en grande partie la situation économique régionale. L'analyse et la prévision des économies urbaines en Chine est par conséquent un axe de recherche important pour contribuer aux choix politiques et stratégiques pour le développement économique nationale.

De 1949 à 1984, les choix économiques de la Chine ont accordé une priorité très importante au développement des villes industrielles. En 1949, il n'y a en Chine qu'un petit nombre de villes. La politique économique du gouvernement Chinois à cette époque était d'industrialiser préalablement les villes centrales. La stratégie du développement régional avait pour le but d'équilibrer les développements des régions. Et le gouvernement a donné une grande importance aux villes centrales. Durant cette période, sous le système de haute planification, l'état a concentré les investissements et les ressources du pays pour bien les développer, et effectivement les moyennes et grandes villes industrielles ont connu une forte croissance.

La réforme et l'ouverture des villes en Chine commence après 1984. La stratégie du développement régional a connu un changement très important. Le gouvernement a donné plus de libertés politiques aux villes au bord de la mer, favorisé les investissements locaux (impôts plus faibles, prix plus libres) et étrangers (contrôle des changes allégé). Et les administrations régionales ont obtenu plus d'autonomie économique. Depuis 1984, les villes au bord de la mer en Chine ont connu un très grand essor économique et social. La stratégie de réforme et d'ouverture a permis un progrès incontestable. La réforme et l'ouverture a conduit à une évolution évidente de la structure économique régionale en Chine.

2. Les données

Un échantillon de 80 villes typiques de Chine a été sélectionné à partir de l'Annuaire Statistique des Villes en Chine (Tableau 1). A l'aide d'une classification automatique (Wang, 1992), elles ont été partagées en quatre classes. La première classe est constituée des moyennes et grandes villes industrielles où l'état a concentré ses investissements depuis longtemps. La deuxième classe est formée des villes au bord de la mer qui ont progressé par bonds grâce à la réforme et l'ouverture. La troisième classe correspond aux villes des provinces de Jiangsu et Zhijiang, dont le commerce et les entreprises rurales ou privées sont beaucoup plus actifs que dans les autres régions. La dernière classe est celle des petites villes qui se sont développées lentement par rapport aux villes des autres classes ces dernières années; elles se trouvent plutôt à l'intérieur de Chine, principalement dans la région de l'ouest.

Nous allons analyser séparément ces quatre classes car elles ont des comportements économiques différents.

Afin de mieux comprendre les effets des éléments économiques, nous analysons ces 80 villes durant la période 1987-1991. Pendant cette période, l'économie chinoise connaissait successivement des phases de croissance, de dépression, de réajustement et de rétablissement.

Tableau 1 : Classification des 80 villes Chinoises

Classe 1 Villes moyennes et grandes	Classe 2 Villes au bord de la mer	Classe 3 Villes des provinces de Jiangsu et Zhijiang	Classe 4 Petite Villes
Shenyang	Dalian	Hangzhou	Xining
Nanjin	Qingdao	Suzhou	Yinchuan
Wuhan	Yantai	Wuxi	Yichang
Chongqing	Ningbo	Nantong	Jiamusi
X'an	Zhuhai	Changzhou	Chimeng
Chendu	Qinhuangdao	Wenzhou	Changzhi
Harbin	Zhanjiang		Kelamayi
Changcun	Shantou		Guilin
Zhenzhou	Lianyungang		Shashi
Taiyuan	Xiamen		Jindezhen
Kunming			Ganzhou
Lanzhou			Tonliao
Guiyan			Hanzhong
Wulumuqi			Xichang
Changsha			Yanji
Fuzhou			Xilinhaote
Jinan			Yiyayuguan
Luoyan			Meizhou
Shijiazhuang			Yan'an
Nanning			Heihe
Tatong			Puyang
Baotou			Chaoyang
Mudanjiang			Wulanhaote
Qiqihaer			Ku'erle
Liuzhou			Keshi
Nanchang			Baoshan
Hefei			Yinin
Dantong			Hetian
Huhehaote			Fujin
			Haila'er
			Aletai
			Lasa
			Tunhuang
			Akesu
			Nanyang

3. Les variables économiques

Nous étudions dans cet article les économies urbaines des différentes régions par la méthode de la régression PLS.

L'objectif de la régression PLS est de relier un ensemble de variables dépendantes $Y = \{Y_1, \dots, Y_p\}$ à un ensemble de variables indépendantes $X = \{X_1, \dots, X_M\}$ lorsque le nombre de variables indépendantes et/ou dépendantes est élevé, ou lorsque la multicollinéarité est forte. La régression PLS revient à effectuer des analyses en composantes principales des ensembles de variables X et Y , sous la contrainte que les composantes principales des X_j soient aussi corrélées que possible aux composantes principales des Y_k . Il est alors possible de prédire les Y_k à partir des X_j en séparant mieux l'information structurée du résidu aléatoire. Les livres de Martens et Naes (1989), de Höskuldsson (1996) et de Tenenhaus (1997) et l'article de Tenenhaus, Gauchi et Ménardo (1995) sont actuellement les références les plus complètes sur le sujet. Par ailleurs le logiciel SIMCA-P (version 1996) permet une utilisation facile de la régression PLS et fournit de nombreuses aides à l'interprétation.

Dans cette recherche, la production industrielle, la consommation des biens de détail, et le produit intérieur brut sont choisis comme les variables dépendantes. Elles représentent le développement de l'industrie, du commerce et le niveau de l'économie globale des villes.

Les variables indépendantes sont divisées en quatre parties.

La première partie est constituée de la variable Population. L'économie des villes est fortement liée à la taille de la population.

La deuxième partie est constituée des variables représentant les investissements, y compris l'investissement d'immobilisation des entreprises nationalisées et l'investissement d'immobilisation du secteur collectif. Ces dernières années, l'évolution économique de la Chine est étroitement liée au niveau de l'investissement total. Les investissements économiques sont encouragés ou freinés par des mesures étatiques. Le niveau de l'investissement est un mécanisme de réglage qui caractérise la situation économique en Chine.

Le troisième groupe de variables indépendantes est formé des variables investissement étranger, importation/exportation et acquisition d'exportations. Elles décrivent l'ouverture des villes en Chine.

La quatrième partie est la variable "nombre d'employés scientifiques". En Chine cet indice a deux aspects. D'une part, les villes bien développées ont besoin d'un grand nombre d'employés scientifiques. Par exemple ces dernières années le développement économique des villes au bord de la mer attire beaucoup d'employés scientifiques. D'autre part, beaucoup d'employés scientifiques sont concentrés dans les villes ayant une base industrielle planifiée par l'état. D'une certaine manière, le nombre d'employés scientifiques représente le niveau technique de ces villes.

Précisons les notations utilisées

Variables dépendantes :

- GDP = Produit Intérieur Brut (*Gross Domestic Product*)
- INDUPROD = Production Industrielle
- CONSUME = Consommation des biens de détail

Variables indépendantes :

- POPULAT = population
- INVSTATE = investissement d'immobilisation des entreprises nationalisées
- INVCOLLE = investissement d'immobilisation du secteur collectif
- INVFOREI = investissement étranger
- IMPOEXPO = importation/exportation
- EXPOACQU = acquisition d'exportation
- SCIENUMB = nombre des l'employés scientifiques

4. Evaluation des résultats de la Régression PLS

Les régressions PLS sont effectuées pour chacune des quatre classes de villes étudiées pendant la période 1987 à 1991.

Notons u_1 la première composante des variables dépendantes Y_k et t_1 la première composante des variables indépendantes X_j . La liaison entre u_1 et t_1 est visualisée pour chaque classe de villes dans la Figure 1. On peut voir clairement que u_1 et t_1 ont des liaisons linéaires bien évidentes dans ces quatre graphiques. Cela signifie qu'il y a une forte liaison entre les variables indépendantes et les variables dépendantes pour chacune de ces quatre classes.

Le résumé de la régression PLS sur la classe des villes industrielles moyennes et grandes est dans le Tableau 2. Le modèle de régression PLS extrait deux composantes efficaces (critère $Q^2=0.130 > \text{limite}=0.097$). Ces deux composantes expliquent 88,6% de la variance des variables dépendantes Y et 66,6% des variances des variables indépendantes X . Pour des commentaires sur le critère de validation croisée Q^2 et la justification de la limite, on peut consulter Tenenhaus (1997).

Tableau 2: Régression PLS. Villes industrielles moyennes et grandes

Composante	1	2	Cumul
R2X	0.557	0.109	0.666
R2Y	0.856	0.029	0.886
Q2	0.850	0.130	
Limite	0.097	0.097	

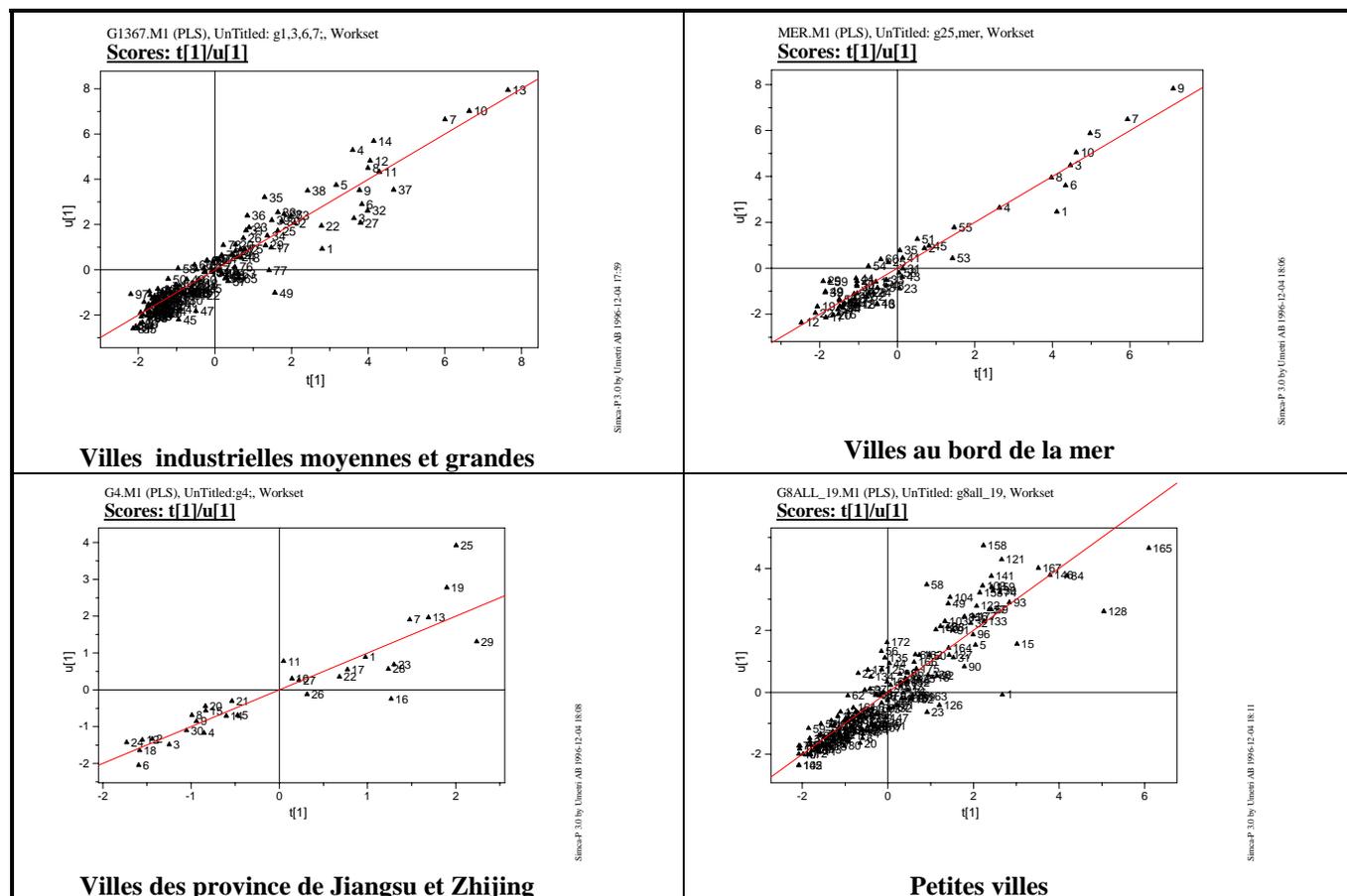


Figure 1 : Liaison entre t_1 et u_1

Dans le tableau 3, nous voyons que la régression PLS extrait une composante efficace pour la deuxième classe. Le modèle explique 91,2% de variance des variables dépendantes. L'analyse pour les villes au bord de la mer conduit à une plus grande précision par rapport aux autres classes, puisque la tendance du développement économique est plus forte et plus concentrée entre les villes de ces classes. Les trois variables GDP, INDUPROD et CONSUME sont bien prédites. Les modèles expliquent 69,1% des variances de ces variables.

Tableau 3: Régression PLS. Villes au bord de la mer

Composante	1	2(*)	Cumul
R2X	0.691	0.132	0.823
R2Y	0.912	0.008	0.920
Q2	0.904	0.044	
Limite	0.097	0.097	

(*) Bien que non significative, la deuxième composante est retenue pour permettre la construction de la carte des deux premières composantes PLS.

Dans la Figure 1, pour la classe des villes des provinces de Jiangsu et Zhijiang, on peut constater qu'il existe une corrélation assez forte entre les variables dépendantes et les variables indépendantes. Mais les éléments du Tableau 4 montrent que le modèle explique 66.7% de la variance de Y et 42.6% de la variance de X. A l'aide d'une analyse plus précise (voir section 7), on trouvera que les modèles de prévision pour GDP et CONSUME sont très satisfaisants. Pourtant le modèle pour INDUPROD a une variance de prévision plus grande et deux villes sont très mal prévues. La capacité d'explication de X n'est pas suffisante pour la variable dépendante INDUPROD. En effet, le développement de l'industrie de cette région compte plutôt sur le secteur rural et privé. Malheureusement, il n'y a pas de statistique de ce secteur pendant les années 1986-1991. Nous utilisons la variable des investissements du secteur collectif pour expliquer une partie de cet élément économique, mais le modèle de prévision de INDUPROD n'est pas très satisfaisant.

Tableau 4: Régression PLS. Villes des provinces de Jiangsu et Zhijiang

Composante	1	2(*)	Cumul
R2X	0.237	0.189	0.426
R2Y	0.556	0.111	0.667
Q2	0.486	0.063	
Limite	0.097	0.097	

(*) SIMCA-P a retenu deux composantes car les critères Q2 des variables GDP et CONSUME valent 0.354 et 0.250 respectivement pour la troisième composante.

Dans le tableau 5, le modèle de la régression PLS a trois composantes efficaces qui expliquent 77.1% de la variance des variables dépendantes et 67.7% de la variance des variables indépendantes. La tendance commune du développement de ces villes est bien évidente. Mais son degré de concentration est moins fort que les classes de villes au bord de la mer et les villes industrielles moyennes et grandes.

Tableau 5: Régression PLS. Petites villes

Composante	1	2	3(*)	Cumul
R2X	0.333	0.184	0.163	0.677
R2Y	0.701	0.051	0.018	0.771
Q2	0.687	0.134	0.015	
Limite	0.097	0.097	0.097	

(*) SIMCA-P a retenu 3 composantes car le critère Q2 de la variable GDP vaut 0.169 pour la troisième composante.

5. Comparaison des Comportements Economiques des Quatre Classes

Les figures des critères VIP et des poids peuvent nous aider à bien situer les différences de comportement économique entre les quatre classes. VIP signifie “*Variable Importance in the Projection*”. Il représente l'importance de chaque variable indépendante pour expliquer Y.

Notons t_h la $h^{\text{ème}}$ composante construite sur les variables X , w_{jh}^* représente le poids de la variable X_j dans la construction de t_h , c_{kh} est le coefficient de régression de la variable Y_k sur la composante t_h . A l'aide du graphique des coefficients $(w_{j1}^*, c_{k1}) / (w_{j2}^*, c_{k2})$, les liaisons entre les variables Y et X peuvent être bien visualisées.

1. Villes Industrielles Moyennes et Grandes

Le graphique des VIP de la Figure 2 montre que la variable la plus importante pour expliquer le niveau économique des villes est la population. A part la population, les variables importantes sont les investissement d'état et le nombre d'employés scientifiques.

Les poids w^*c de la Figure 2 indiquent que l'économie globale est très corrélée avec le développement de l'industrie et du commerce. Les trois points GDP, INDUPRO et CONSUME sont presque superposés. Les variables les plus proches de ces trois variables dépendantes sont l'investissement d'état et le nombre d'employés scientifiques. L'investissement étranger a aussi influencé le développement des villes industrielles moyennes et grandes, mais la liaison est moins forte. Enfin l'effet de la variable IMPOEXPO est le plus faible.

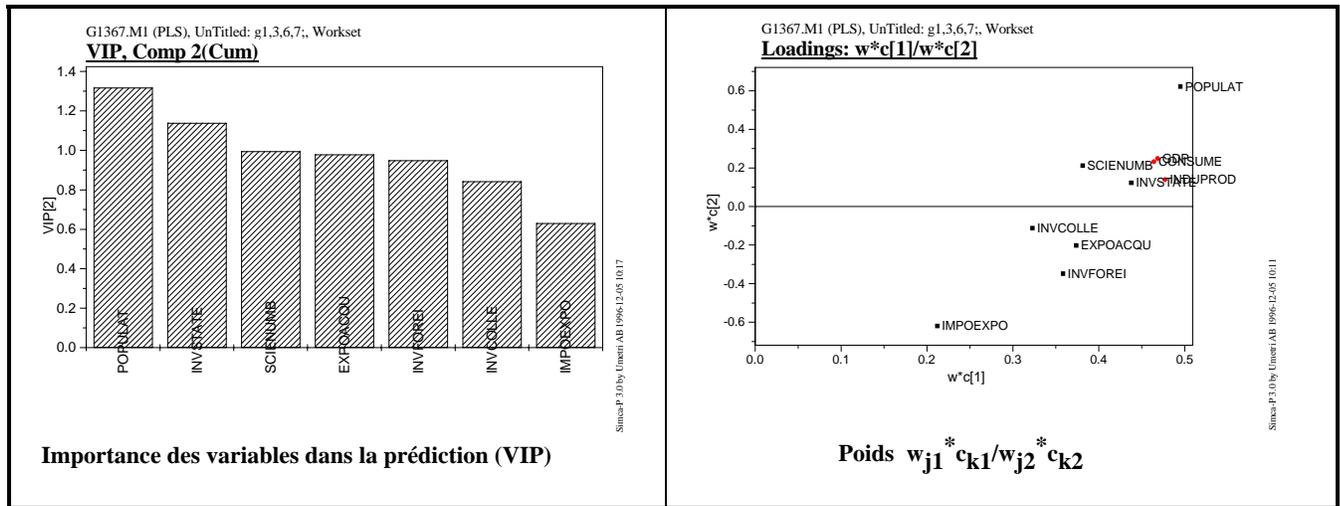


Figure 2: VIP et poids w^*c pour les villes industrielles moyennes et grandes

Durant une longue période, le système économique en Chine était extrêmement centralisé et planifié. Le développement des villes dépendait entièrement du développement de l'industrie, avec priorité à l'industrie lourde. Les villes de cette classe, où l'état a beaucoup investi, se sont développés plus rapidement. Elles ont établi leurs bases industrielle et économique et possèdent une puissance scientifique et technologique qui constitue la force motrice du développement de ces villes. Cependant depuis 1984, la réforme et l'ouverture urbaine ont commencé par les villes au bord de la mer. Le gouvernement Chinois a donné plus de priorité et de liberté à ces régions. Par conséquent l'ouverture de ces grandes et

moyennes villes industrielles qui se trouvent à l'intérieur de la Chine est plus faible que les villes au bord de la mer. Les variables qui expriment l'ouverture n'ont pas des effets décisifs dans cette classe.

2. Les Villes au bord de la Mer

Dans la figure 3, on trouve, qu'à part la population, les variables économiques les plus importantes sont Importation/Exportation et Investissement d'état. De plus toutes les autres variables explicatives influent sur le développement de villes au bord de la mer.

Dans la carte des variables de la figure 3, le développement industriel et commercial sont prédictifs de l'économie globale, GDP essentiellement. Les trois variables dépendantes ont presque les mêmes liaisons avec tous les éléments économiques. Parmi toutes les variables explicatives, IMPOEXPO, INVSTATE, SCIENUMB et EXPOACQU ont les corrélations les plus élevées avec les variables dépendantes. En comparant avec les autres classes les liaisons entre les trois variables dépendantes et les variables indépendantes sont beaucoup plus corrélés (toutes les variables sont à droite dans la carte de la figure 3). Encouragé par le stratégie de réforme et d'ouverture, les villes au bord de la mer ont connu un développement très rapide et leurs économie et société sont beaucoup plus prospères.

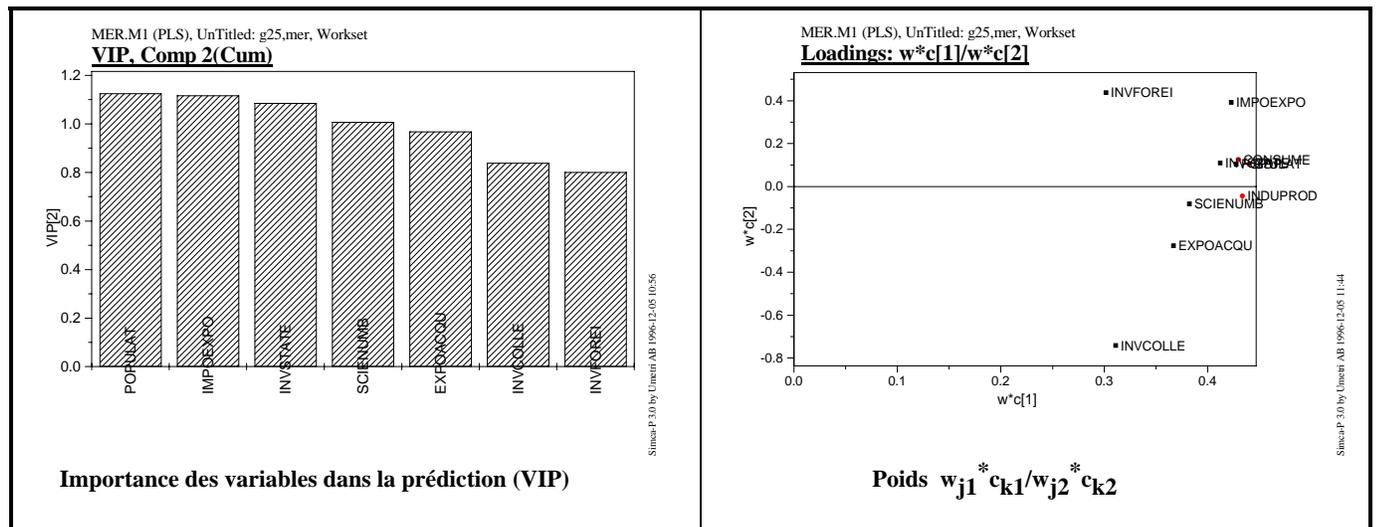


Figure 3 : VIP et poids $w_{j1} * c_{k1} / w_{j2} * c_{k2}$ pour les villes au bord de mer

3. Les Villes des Provinces Jiangsu et Zhejiang

Les comportements économiques des villes des provinces Jiangsu et Zhejiang sont très particuliers. Les liaisons entre les variables indépendantes et le développement industriel, le commerce et l'économie globale de cette classe sont très différents de celles observées dans les autres régions.

Premièrement le graphique des VIP (figure 4) montre qu'à part la population, la variable la plus importante dans cette région est l'investissement d'état. Le phénomène est ici différent car l'effet de l'investissement du secteur collectif semble être plus important que dans les autres régions.

La carte des variables donne une explication bien adaptée au comportement économique de cette région. Dans la figure 4, la variable GDP est corrélée fortement à la variable CONSUME. Cela signifie que le développement économique total de cette région dépend fortement de la prospérité de son commerce. En réalité dans ces deux provinces la densité de population est très grande et il y a beaucoup de gens qui travaillent dans le secteur du commerce. Le commerce de cette région est beaucoup plus actif. On peut voir aussi dans la figure.10 que le développement du commerce est bien corrélé avec la population. Cependant le développement industriel est largement éloigné de l'économie globale. L'industrie de cette région dépend plutôt de l'investissement du secteur collectif et de l'acquisition d'exportation.

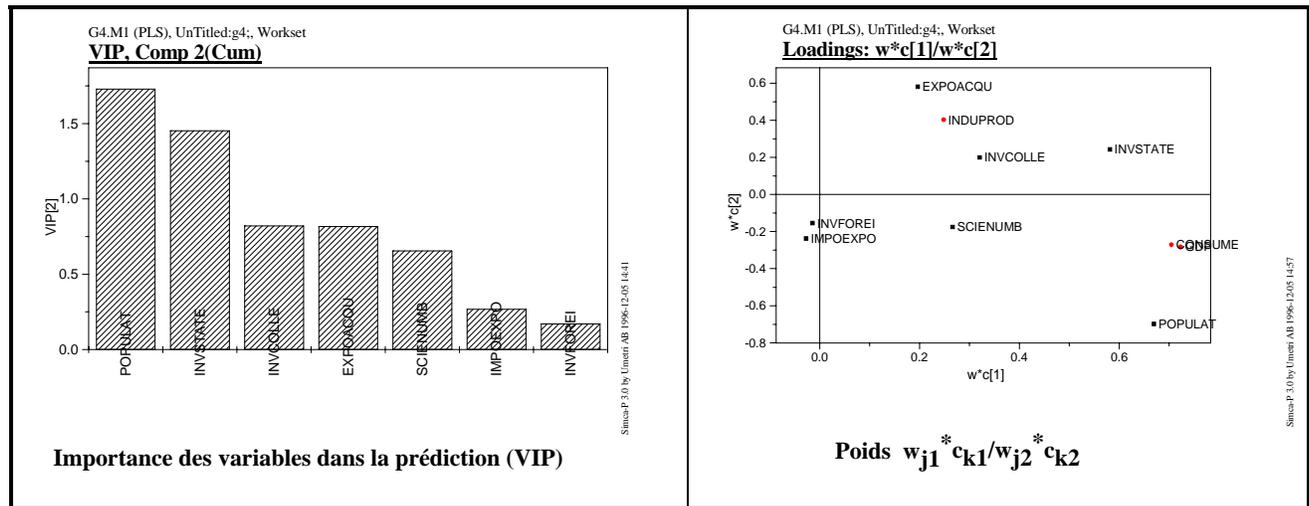


Figure 4: VIP et poids $w_{j1} * c_{k1} / w_{j2} * c_{k2}$ pour les villes de Jiangsu et Zhejiang

4. Les Petites Villes à l'Intérieur de la Chine

La figure 5 (VIP et carte des variables) pour les petites villes à l'intérieur de la Chine nous racontent beaucoup de faits intéressants. Le graphique des VIP montre que les trois éléments les plus importants dans la vie économique de cette région sont le nombre des employés scientifiques, l'investissement d'état et la population. En réalité parmi ces petites villes à l'intérieur du pays que nous avons choisies, il y en a une grande partie qui ont des

ressources industrielles spécifiques. Elles bénéficient de ressources particulières de l'état où les employés scientifiques sont assez nombreux. De plus l'état a organisé une politique de haut salaire pour attirer des employés efficaces. Donc plus le nombre des employés scientifiques est élevé, meilleure est la situation du développement industriel et du commerce. Pourtant leurs industries sont plutôt limitées à quelques domaines et l'économie globale n'est pas bien développée dans de nombreux secteurs.

Dans la carte des variables, nous constatons pour la deuxième fois le phénomène de la séparation entre les variables GDP, INDUPROD et CONSUME. Mais la situation est différente de celle de la région de Jiangsu et Zhejiang, ici l'économie globale est reliée étroitement avec l'industrie qui dépend principalement de l'investissement d'état et le nombre d'employés scientifiques. Dans cette région le commerce reste peu développé. La consommation des biens dépend essentiellement de la population. En outre par rapport aux habitants régionaux, les employés scientifiques ont un niveau de consommation plus élevé. Importation et exportation sont très faibles dans cette région. Durant la réforme urbaine décennale, les villes de cette classe se développent lentement. Son écart avec les autres classes s'élargit.

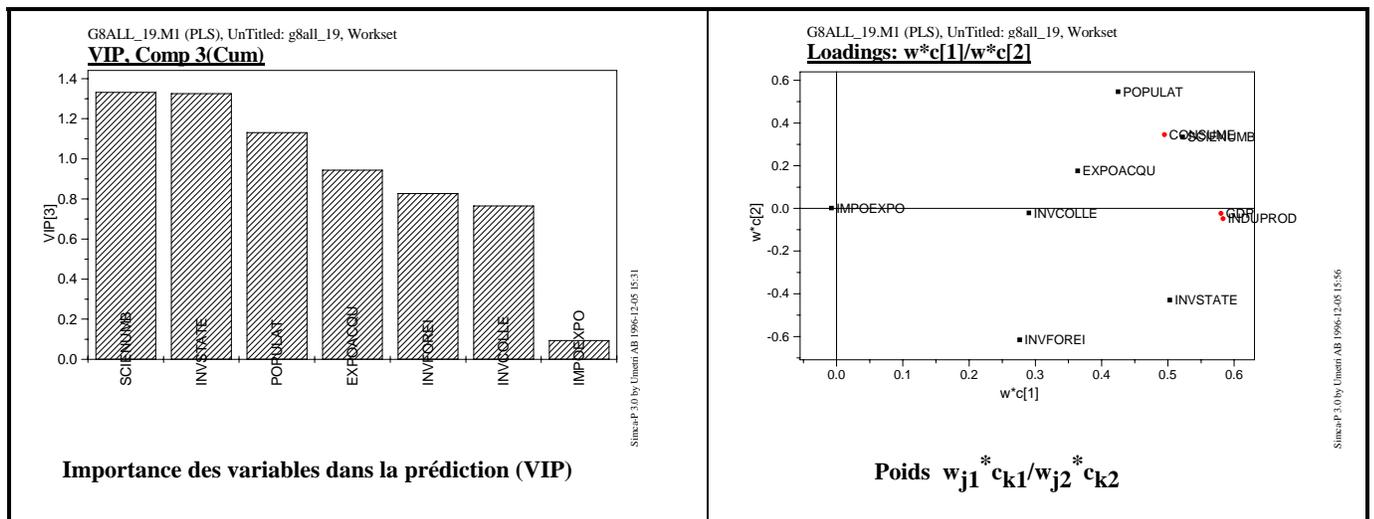


Figure 5 : VIP et poids $w_{j1} * c_{k1} / w_{j2} * c_{k2}$ pour les petite villes

6. Interprétation des Composants t_1 / t_2

Dans cette section nous allons interpréter les cartes des deux premières composantes t_1 / t_2 pour les différentes classes de villes. Dans SIMCA-P, on construit dans le plan (t_1, t_2) l'ellipse de Hotelling pour observer si les villes peuvent être considérées comme homogènes et détecter les points atypiques.

1. Les Grandes et Moyennes Villes Industrielles

Dans les tableaux 6 et 7, on peut voir que la composante t_1 est principalement construite avec la population, l'investissement d'état et le nombre d'employés scientifiques. Les autres variables économiques contribuent également de manière positive à sa construction. Donc il représente le facteur de niveau économique. La population a une contribution positive à la construction de la composante t_2 et toutes les variables du commerce international ont une contribution négative. Donc, dans la carte des villes, celles qui sont situées à droite sont les villes (de différentes années) qui ont les niveaux économiques les plus élevés. Et les villes vers le bas sont les villes plus ouvertes.

Tableau 6 : Villes industrielles moyennes et grandes

R^2 entre les composantes t_1 , t_2 et les variables X et Y

Variable	R2(Variable,t)	R2(variable,t2)
GDP	0.85	0.04
INDUPROD	0.88	0.01
CONSUME	0.83	0.03
POPULAT	0.79	0.12
INVSTATE	0.71	0.00
INVCOLLE	0.42	0.01
INVFOREI	0.58	0.05
IMPOEXPO	0.26	0.45
EXPOACQ	0.58	0.04
SCIENUMB	0.52	0.07

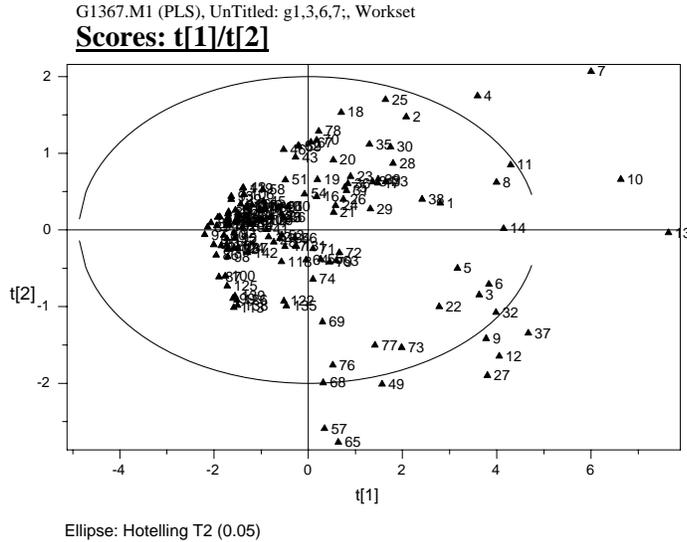
Tableau 7 : Villes industrielles moyennes et grandes

Coefficients $c1,w1^*$ et $c2,w2^*$

Variable	$c1,w1^*$	$c2,w2^*$
GDP	0.46	0.24
INDUPROD	0.47	0.13
CONSUME	0.46	0.23
POPULAT	0.49	0.62
INVSTATE	0.43	0.12
INVCOLLE	0.32	-0.11
INVFOREI	0.35	-0.34
IMPOEXPO	0.21	-0.61
EXPOACQ	0.37	-0.20
SCIENUMB	0.38	0.21

Dans la figure 6, tous les points vers le haut à droite correspondent à la ville Shenyang à différentes années. Shenyang est une grande base industrielle importante en Chine. Sa population est très nombreuse avec une puissance économique beaucoup plus grande que

le niveau moyen de cette classe. Les villes vers le bas à droite sont Fuzhou et Nanjing. L'ouverture de ces villes est plus forte qu'à Shenyang.



Simca-P 3.0 by Umetri AB 1996-12-08 09:15

Figure 6: Carte des deux première composantes PLS pour les villes industrielles moyennes et grandes

2. Les Villes au Bord de la Mer

Dans les tableaux 8 et 9, on constate que les variables ont toutes une contribution importante à la composante t_1 et positive. Elle représente donc un facteur de taille. Les poids des variables du commerce international sont plus importants sur t_2 , De l'autre côté le poids de la variables Investissement du secteur collectif est plus négatif.

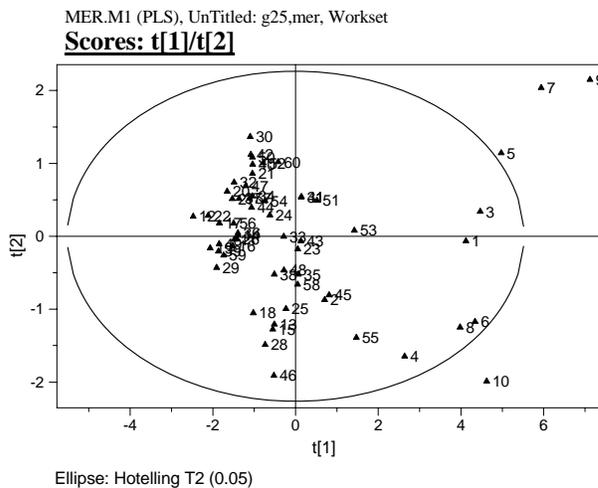
Tableau 8 : Villes au bord de la mer
 R^2 entre les composantes t_1 , t_2 et les variables X et Y

Variable	R2 (Variable,t1)
GDP	0.93
INDUPROD	0.90
CONSUME	0.89
POPULAT	0.85
INVSTATE	0.81
INVCOLLE	0.54
INVFOREI	0.44
IMPOEXPO	0.80
EXPOACQ	0.65
SCIENUMB	0.71

Tableau 9 : Villes au bord de la merCoefficients $c1, w1^*$ et $c2, w2^*$

Variable	$c1, w1^*$	$c2, w2^*$
GDP	0.43	0.10
INDUPROD	0.43	-0.04
CONSUME	0.42	0.12
POPULAT	0.42	0.10
INVSTATE	0.41	0.10
INVCOLLE	0.31	-0.74
INVFOREI	0.30	0.43
IMPOEXPO	0.42	0.39
EXPOACQ	0.36	-0.27
SCIENUMB	0.38	-0.08

Dans la carte des villes (Figure 7), les villes Dalian et Qingdao, dont les populations sont plus grandes (vers 1~2 millions), ont leur niveau économique beaucoup plus haut que les autres villes de cette classe. Les points vers le haut à droite sont Dalian au cours des différentes années. Cela signifie que la base industrielle et le commerce international sont beaucoup plus avancés à Dalian. L'autre grande ville Qingdao est moins forte que Dalian dans l'investissement étranger et dans l'importation/exportation. Mais son secteur non d'état se développe plus rapidement. Les autres villes au bord de la mer ont une population comprise entre 200 000 à 500 000 habitants. Cependant, encouragé par la stratégie de réforme et d'ouverture sur l'extérieur, leur rythme de croissance économique est explosif.



Simca-P 3.0 by Unscrt AB 1996:12:08 09:28

Figure 7 : Carte des deux premières composantes PLS pour les villes au bord de la mer

3. Les Villes des Provinces de Jiangsu et Zhejiang

L'examen des tableaux 10 et 11 montre que la composante t_1 est essentiellement construite à partir des investissements d'état et du secteur collectif. Elle représente donc la puissance des investissements. La population a une contribution négative à la construction de la composante t_2 , tandis que l'acquisition d'exportation a une contribution positive.

Tableau 10 : Villes des provinces de Jiangsu et Zhejiang

R^2 entre les composantes t_1 , t_2 et les variables X et Y

Variable	R^2 (Variable, t_1)	R^2 (variable, t_2)
GDP	0.80	0.08
INDUPROD	0.09	0.17
CONSUME	0.76	0.07
POPULAT	0.44	0.43
INVSTATE	0.50	0.08
INVCOLLE	0.25	0.01
INVFOREI	0.00	0.10
IMPOEXPO	0.01	0.00
EXPOACQ	0.19	0.61
SCIENUMB	0.24	0.07

Tableau 11 : Villes des provinces de Jiangsu et Zhejiang

Coefficients c_1, w_1^* et c_2, w_2^*

Variable	c_1, w_1^*	c_2, w_2^*
GDP	0.72	-0.28
INDUPROD	0.24	0.40
CONSUME	0.70	-0.27
POPULAT	0.66	-0.69
INVSTATE	0.58	0.24
INVCOLLE	0.32	0.19
INVFOREI	-0.01	-0.15
IMPOEXPO	-0.02	-0.23
EXPOAXQ	0.19	0.58
SCIENUMB	0.26	-0.17

7. Préviation

La prévision est l'objectif essentiel de la Régression PLS. On peut obtenir simultanément des modèles de régression pour les variables dépendantes GDP, INDUPROD et CONSUME. Dans SIMCA-P, en plus du tableau des coefficients de régression on trouve des visualisations de ces coefficients qui enrichissent l'explication économique d'une façon graphique. On peut aussi construire les graphiques (valeurs observées et valeurs prédites) pour les différentes variables dépendantes.

1. Les Villes Industrielles Moyennes et Grandes

Les coefficients de régression PLS sur les données centrées-réduites pour X et réduites, mais non centrées pour Y, apparaissent dans le tableau 14.

Tableau 14: Les coefficients de régression PLS pour les villes industrielles moyennes et grandes

Variable	GDP	INDUPROD	CONSUME
Constante	1.54	1.49	1.51
POPULAT	0.38	0.32	0.37
INVSTATE	0.23	0.22	0.23
INVCOLLE	0.12	0.13	0.12
INVFOREI	0.08	0.12	0.08
IMPOEXPO	-0.05	0.01	-0.04
EXPOACQ	0.12	0.15	0.12
SCIENUMB	0.23	0.21	0.22

Le graphique des coefficients de régression pour les villes industrielles moyennes et grandes est visualisé dans la Figure 10. La première impression évidente d'un coup d'oeil est l'équivalence des structures des coefficients dans les trois modèles. Ce phénomène confirme encore une fois que l'économie globale se développe simultanément avec l'industrie et le commerce dans cette région. Les coefficients de régression montrent que la population, l'investissement d'état et la puissance scientifique et technologie sont les éléments les plus importants dans la vie économique. Importation et exportation sont faibles dans ces villes . Il existe des villes qui ont une très bonne base industrielle et économique, mais un commerce international restant faible, d'où la corrélation négative entre leur économie et la variable IMPOEXPO.

Les graphiques "valeurs observées et valeurs prédites" pour les variables GDP, INDUPROD et CONSUMED sont également donnés dans la Figure 10. Les qualités des prévisions de ces trois modèles sont proches et satisfaisantes.

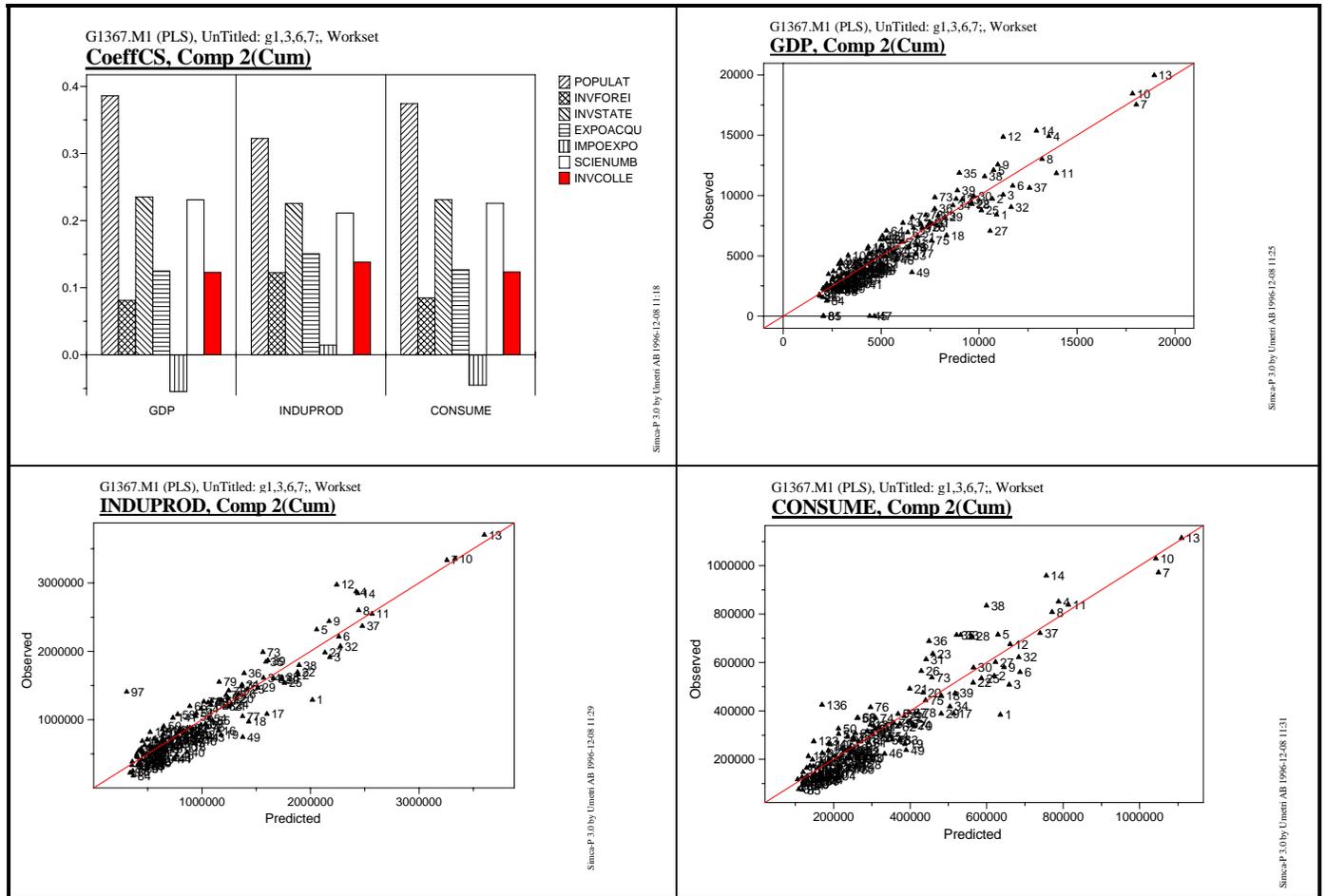


Figure 10 : Les coefficients de régression PLS pour les villes industrielles moyennes et grandes. Cartes des valeurs observées et prédites pour les villes industrielles moyennes et grandes

2. Les Villes au Bord de la Mer

Les coefficients de régression PLS pour les villes au bord de la mer sont dans le tableau 15.

Tableau 15: Les coefficients de régression PLS pour les villes au bord de la mer

Variable	GDP	INDUPROD	CONSUME
Constante	1.23	1.05	1.49
POPULAT	0.18	0.15	0.18
INVSTATE	0.18	0.17	0.17
INVCOLLE	0.13	0.13	0.13
INVFOREI	0.13	0.13	0.12
IMPOEXPO	0.18	0.18	0.18
EXPOACQ	0.16	0.15	0.15
SCIENUMB	0.16	0.16	0.16

Les structures des coefficients de régression pour les trois variables dépendantes sont très voisines pour les villes au bord de la mer. Dans le graphique des coefficients de régression (Figure 11), les trois variables dépendantes sont corrélées très fortement avec tous les éléments économique. Cela signifie que tous les éléments économiques ont eu leur importance pour le développement global. On peut remarquer que les variables Importation/Exportation et Investissement d'état sont les plus significatives.

Les cartes des valeurs observées et prédites des variables dépendantes de la Figure 11 montrent les qualités prédictives des modèles. Par ailleurs les grandes villes Dalian et Qindao ont un niveau de développement supérieur aux autre villes.

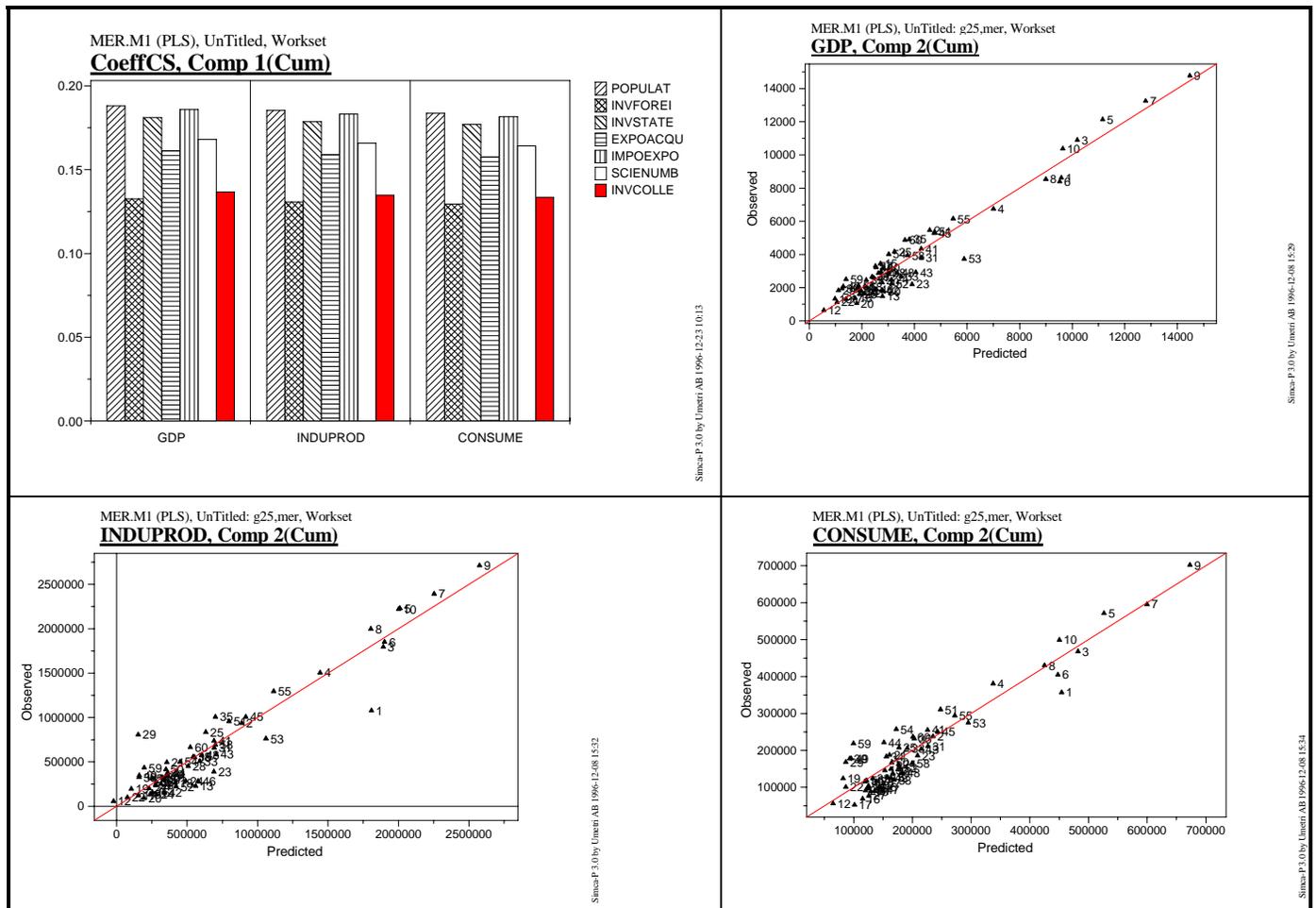


Figure 11 : Les coefficients de régression PLS pour villes au bord de la mer. Cartes des valeurs observées et prédites pour les villes au bord de la mer

3. Les Villes des Provinces de Jiangsu et Zhejiang

Les coefficients de régression PLS pour les villes des provinces de Jiangsu et Zhejiang sont dans le tableau 16.

Tableau 16: Les coefficients de régression PLS pour les villes de Jiansu et Zhejiang

Variable	GDP	INDUPROD	CONSUME
Constante	2.38	1.42	1.94
POPULAT	0.68	-0.11	0.66
INVSTATE	0.35	0.24	0.34
INVCOLLE	0.17	0.15	0.17
INVFOREI	0.03	-0.06	0.03
IMPOEXPO	0.04	-0.10	0.04
EXPOACQ	-0.02	0.28	-0.01
SCIENUMB	0.24	-0.00	0.23

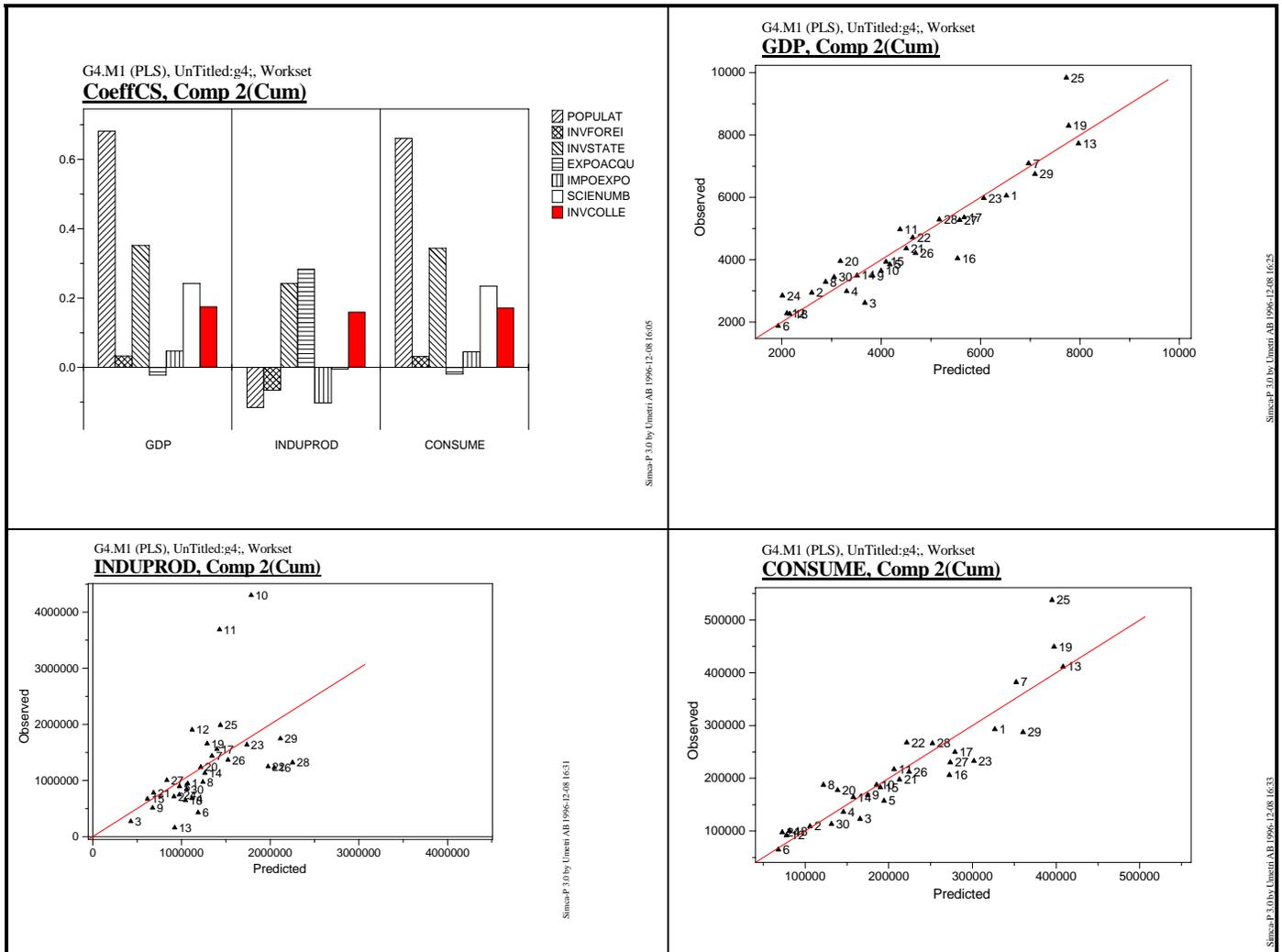


Figure 12 : Les coefficients de régression PLS pour les villes de Jiansu et Zhejiang. Cartes des valeurs observées et prédites pour les villes de Jiansu et Zhejiang

Dans la région de Jiansu et Zhejiang, les structures des coefficients de régression pour les variables GDP et Commerce sont voisines. Les éléments économiques les plus importants sont la Population et les Investissements d'état. De plus la Puissance Scientifique et le Secteur Collectif sont également importants. Cependant les coefficients de régression pour

IMPOEXPO, INVETLAN et EXPOACQU sont négligeables. Le Développement Industriel est plutôt relié avec l'Acquisition d'Exportation, l'Investissement d'État et l'Investissement du Secteur Collectif. En fait, il manque des variables explicatives importantes dans ce modèle, et les variables explicatives actuellement disponibles sont insuffisantes pour expliquer le développement industriel de cette région.

Les cartes des valeurs observées et prédites sont données dans la Figure 12. Les prévisions pour GDP et CONSUME sont satisfaisantes. On voit aussi que la grande ville de Hangzhou a un niveau de développement beaucoup plus élevé que les autres villes. Les prévisions pour la variable INDUPROD sont moins précises. Les productions industrielles de Suzhou (1988) et Wuxi (1988) sont très mal prévues. En effet, les secteurs rural et privé sont très actifs dans la région Jiangsu et Zhejiang. Malheureusement il n'y a pas de statistique pour ce secteur durant 1987-1991.

4. Les Petites Villes à l'Intérieur de la Chine

Les coefficients de régression PLS pour les petites villes à l'intérieur de la Chine sont dans le tableau 17. Les coefficients de régression PLS montrent que la structure économique globale de cette région dépend plutôt de son développement industriel. Les trois premiers éléments importants dans la vie économique sont l'investissement d'état, le nombre d'employés scientifiques et la population. Les effets de l'importation et de l'exportation sont négligeables. De plus le commerce n'est pas actif dans cette région. La consommation dépend essentiellement de la population et du nombre d'employés scientifiques. Le fait remarquable dans cette région est la liaison extrêmement faible entre d'une part l'économie globale et d'autre part le commerce international, l'introduction de l'investissement étranger et l'extension du secteur non d'état.

Tableau 17: Les coefficients de régression PLS pour les petites villes

Variable	GDP	INDUPROD	CONSUME
Constante	1.33	1.07	1.55
POPULAT	0.29	0.22	0.40
INVSTATE	0.43	0.31	0.11
INVCOLLE	0.03	0.16	0.12
INVFOREI	0.11	0.18	-0.08
IMPOEXPO	-0.06	-0.00	-0.00
EXPOACQU	0.09	0.20	0.22
SCIENUMB	0.32	0.28	0.37

Les cartes des valeurs observées et prédites sont données dans la Figure 13. La prévision du PIB est satisfaisante dans cette classe. Les villes de Xining, Yinchuan, Jiamusi et Kelamayi sont plus développées que le niveau moyen des villes de cette classe. Dans les cartes de l'industrie et du commerce, les prévisions sont aussi satisfaisantes, mais avec des variances résiduelles plus importantes. En effet les facteurs de développement des villes de cette

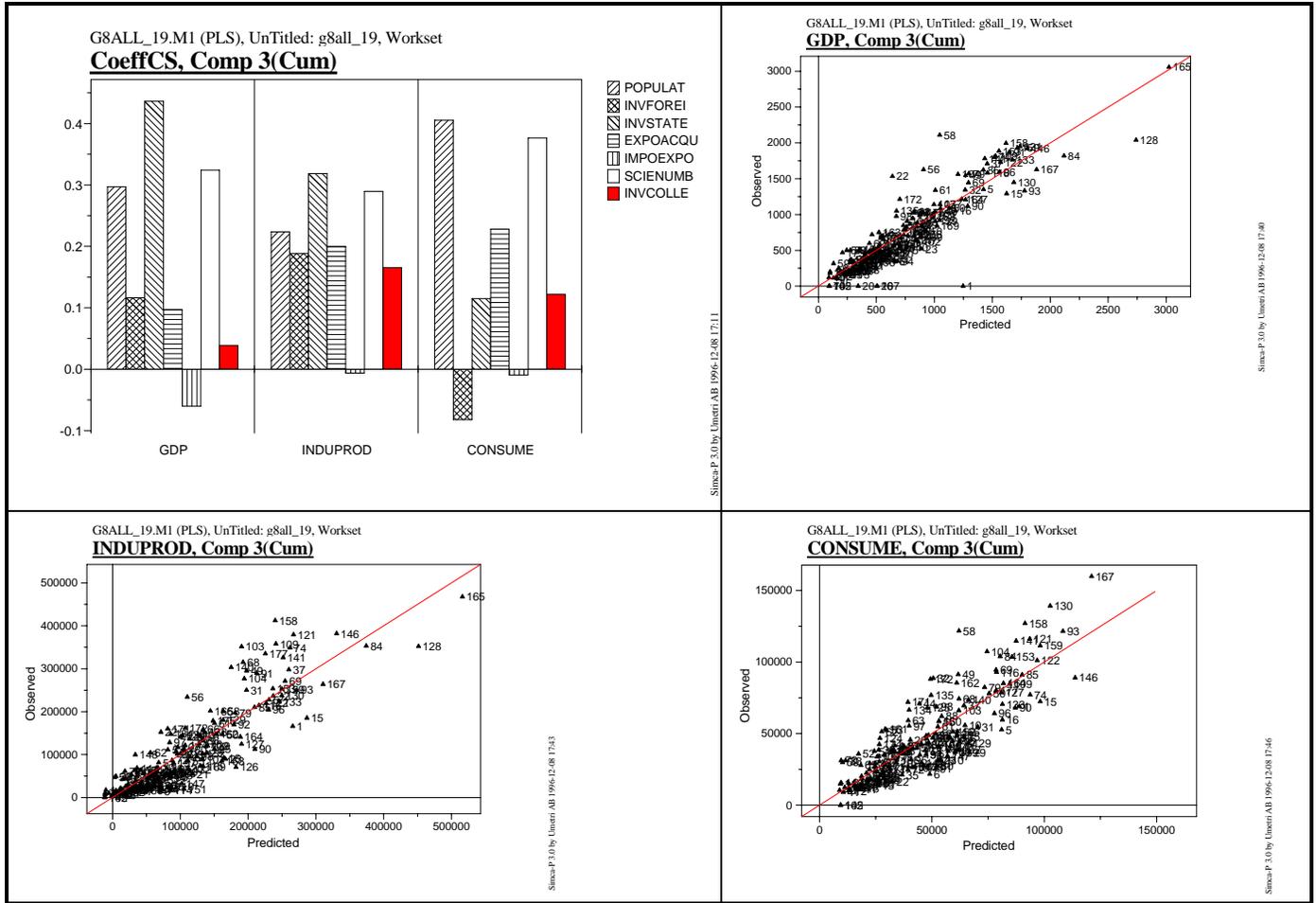


Figure 13: Les coefficients de régression PLS pour les petite villes. Cartes des valeurs observées et prédites pour les petite villes

classe sont plus hétérogènes que ceux des villes au bord de la mer qui ont fortement profité de la politique de réforme et d'ouverture. Donc pour obtenir un modèle plus précis, il faudrait faire une classification supplémentaire des villes de cette classe. Néanmoins la situation et le schéma de développement de cette région comparés avec les autre classes sont bien mis en évidence.

8. Bilan

Dans cette article le comportement économique et les tendances du développement économique urbain chinois sont étudiés par la régression PLS à l'aide du logiciel SIMCA-P. Nous avons choisi 80 villes typiques et une classification en quatre classes: les villes industrielles moyennes et grandes, les villes au bord de la mer, les villes des provinces de Jiangsu et Zhejiang, et les petites villes à l'intérieur surtout à l'ouest de la Chine. La modélisation par régression PLS des économies de ces quatre classes montre que les comportements économiques et les éléments importants au cours du développement sont très différents entre les classes.

Depuis la réforme et l'ouverture sur l'extérieur commencées en 1984, l'économie globale urbaine en Chine se développe. Les villes au bord de la mer ont particulièrement profité de la stratégie de réforme et se développent à un rythme de croissance très rapide. La réussite de ces villes est déjà un modèle qui influence largement les autres régions de la Chine.

Les villes industrielles moyennes et grandes possèdent déjà une bonne base industrielle. L'économie de ces villes s'est également développée pendant ces dernières années. Cependant la lenteur de l'ouverture sur l'extérieur ralentit leur développement. La prépondérance absolue de l'industrie nationalisée a déjà beaucoup diminué. Les villes qui dépendent fortement de l'investissement d'état ont rencontré des difficultés dans le système économique de marché appliqué durant la réforme.

Dans la région Jiangsu et Zhejiang le commerce est très prospère et les secteurs rural et privé sont plus étendus. La réussite économique de cette région au cours de ces dernières années est remarquable. Mais l'effet du commerce international sur l'économie de cette classe n'est pas très important.

Les petites villes à l'intérieur, surtout à l'ouest de la Chine, se développent plus lentement que les autres villes. Les bases économique et technologique sont faibles dans ces villes, dont l'infrastructure est mal adaptée aux besoins du développement. Durant la période de réforme, elles sont moins touchées par la stratégie de réforme et d'ouverture. Par exemple en 1990 l'acquisition d'exportation, l'exportation et l'importation des villes dans la région de l'est (au bord de la mer) occupent respectivement les proportions de 69,2%, 89,3% et 92,3% par rapport des villes de Chine. En opposition à l'ouest elles sont seulement de 5,7%, 2,0% et 2,7%. Le commerce est aussi sous-développé. Les écarts de développement avec les autres classes s'élargissent constamment. D'après la statistique de 1990, la surface urbaine de l'ouest occupe 41,5% du pays, mais son PIB urbain occupe seulement 12,6% du PIB urbain total chinois, et sa production industrielle 11,9%. Pourtant la région de l'est, avec sa surface urbaine 20,1% du pays, fournit 59,7% du PIB urbain et 61,3% de la production industrielle. Cependant les ressources naturelles, surtout en minerais sont plus importantes dans cette région que dans le reste de la Chine, ce qui donne des possibilités de développement futur à cette région.

Remerciements :

Ces recherches et l'écriture de cet article ont été menées en collaboration avec le Professeur Michel TENENHAUS. L'achèvement de ce travail a été facilité par ses conseils dans les théories mathématiques et l'utilisation de SIMCA-P. Nous voudrions lui exprimer ici notre remerciement très profond ainsi qu'à la Direction du Groupe HEC qui a soutenu financièrement cette recherche et à Jean STORA et Daniel BARRAULT de LAGERIE pour leurs aides dans tous les domaines. Enfin nous tenons à remercier très chaleureusement le Professeur Bernard BURTSCHY pour sa correction approfondie de cet article et les nombreuses remarques qui nous ont aidé à l'améliorer.

Références :

1. Annuaire Statistique des Villes en Chine (1988~1992), Maison d'Edition Statistique de Chine, Beijing, Vol. 1987~1992.
2. Höskuldsson, A. (1996) . Prediction Methods in Science and Technology. Vol 1. Basic Theory. Thor Publishing, Denmark.
3. Martens, H. et Næs, T. (1989). Multivariate Calibration. John Wiley & Sons, Chichester.
4. SIMCA-P for Windows (1996) :Multivariate Modeling, Analysis and SPC of Process Data. UMETRI AB, Box 7960, S-907 19 Umeå, Suède.
- 5 Tenenhaus, M. (1997) : La Régression PLS : Théorie et Pratique. Document de Recherche de la Fondation HEC, Jouy-en-Josas.
6. Tenenhaus, M., Gauchi, J.-P., Ménardo, C. (1995) : Régression PLS et applications. Revue de Statistique Appliquée, vol. XLIII, n° 1, p. 7-63.
7. Wang, Huiwen (1992) : Data Analysis of Cubic Time Series Data Table : Theory and Application. Thèse Docdorat, BUAA, Beijing.