FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF MATHEMATICS,
STATISTICS AND ACTUARIAL SCIENCES

French Regional Wheat Prices : 1756-1872
Généralités versus Départements

E. Borghers

RESEARCH PAPER 2006-004
February 2006

University of Antwerp, Prinsstraat 13, B-2000 ANTWERP, Belgium
Research Administration - Room B.213
phone: (32) 3 220 40 32
e-mail: joeri.nys@ua.ac.be

D/2006/1169/004

French Regional Wheat Prices : 1756-1872
Généralités versus Départements

S'ils n'ont pas de pain, qu'ils mangent des brioches
Marie-Antoinette

Prof. dr. E. Borghers

University of Antwerp
Faculty of Applied Economics
Department of Mathematics, Statistics and Actuarial Sciences

Key Words : Multivariate Statistics, Wheat Prices, France

Acknowledgement

The helpful comments and suggestions made by dr. K. Van Rompay are greatly appreciated.

Corresponding Address

University of Antwerp
Faculty of Applied Economics
Room C 442
Prinsstraat 13
B 2000 Antwerp - Belgium

eddy.borghers@ua.ac.be
T +32 (0) 3 220 41 31 (Office)
T +32 (0) 3 220 41 54 (Secretary)
F +32 (0) 3 220 48 57 (Secretary)

Contents

Section 1 : Introduction

In a previous paper (See **Borghers [1]**) the French regional wheat prices for the period 1756-1872 were analyzed by using aggregated data. These data were called data set 1, 2 and 3 and are covering respectively the periods 1756-1790, 1797-1830 and 1831-1870. Precise information about these data sets can be found in Appendix 1 and Appendix 2. Each of these data sets consisted of nine price series, i.e. an aggregated series for each of the nine geographical regions called '*Grands Secteurs Teritoriaux*'.

In this paper almost exactly the same period will be analyzed by using disaggregated data. For reason of convenience these disaggregated data will be called 'larger' data sets as opposed to the aggregated and 'smaller' data sets used in the previous paper. The main purpose of this paper is to investigate to what extend the preliminary and promising results obtained by using the 'smaller' and aggregated series could be further refined by using the 'larger' and disaggregated data sets.

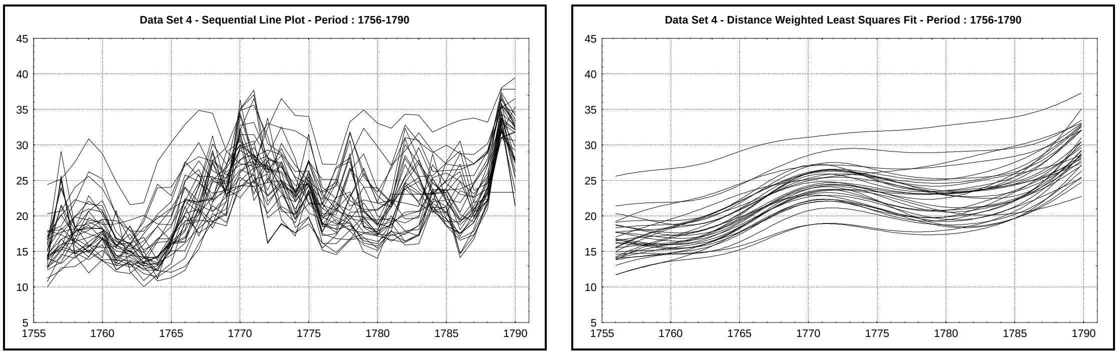
The statistical methodology and techniques used in this paper will be the same as those used in the previous paper. Therefore reference will rather be made to that previous paper than to the literature mentioned in the References.

Section 2 : Period 1756-1790 - Généralités

Introduction

For the period 1756-1790 disaggregated data about the price evolution in France of wheat are available on the level of the *généralités*. This data set, henceforth called data set 4, consists of the price of wheat for 32 regions. An additional 33rd series is available describing the price evolution for *Paris Ville*. Precise details about this data set can be found in Appendix 1. A list of these 33 regions is presented in Appendix 3.

Figure 1 : Wheat Price for 33 Généralités - Period : 1756-1790
Distance Weighted Least Squares Fit



The 33 regional wheat price series are graphically presented in Figure 1. The observed series are plotted on the left side while the distance weighted least squares fit for each of these time series is given on the right. A visual inspection of the graphical representations leads to the following main characteristics :

- a global and common upward long term trend
- a quasi cyclical behavior with a maximum around 1771-1772 and lowering prices for the periods 1761-1762 and 1781-1782
- the difference between lowest and highest prices remains fairly constant
- an asynchronous behavior of the price series resulting in a considerable variability

Cluster Analysis - Tree Clustering

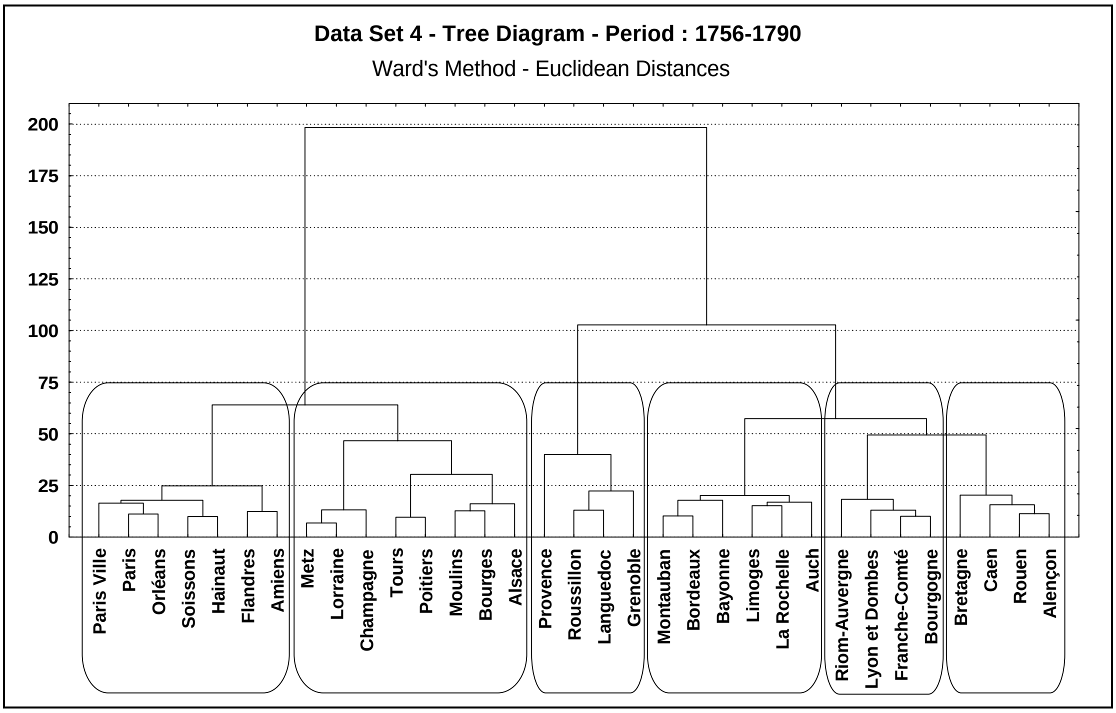
A first step in the analysis of data set 4 was the application of cluster analysis. In the exploratory phase of the analysis and in the absence of any a priori knowledge about structure and hierarchy, the most obvious choice was to start the analysis with a tree clustering of the 33 price series. The outcome of this tree clustering approach is summarized and graphically represented by the dendrogram given in Figure 2. From this dendrogram the following conclusions can be drawn :

- the hierarchical structure represented by the dendrogram is dominated by two main clusters composed by respectively the northern an southern part of France
- each of these main clusters is characterized by the distinction between *généralités* from the western and from the eastern part of France
- the only exception or contradicting behavior is shown by the *généralités* Alençon, Bretagne, Caen and Rouen. Although geographically situated in the north-west of France they are part, within the sub-cluster describing the *généralités* in the east, of the main cluster describing the southern part of France.

Cluster Analysis - k-Means Clustering

In order to get a better insight into the homogeneity of the several sub-clusters shown by the tree clustering the following experiment might be helpful. This experiment consists in applying the k-means clustering technique for a varying value for k, i.e. for a varying number of clusters ranging from two to six clusters. The composition for each of the five resulting clustering solutions from this experiment can be found in Table 1.

Figure 2 : Wheat Price for 33 Généralités - Period : 1756-1790
Tree Clustering Solution - Ward's Method - Dendrogram



The main results from this experiment can be summarized as follows :

- homogeneous geographical regions tend to stay together even when they are joining other regions
- by reducing the total number of clusters the original clusters are aggregating to meaningful and homogeneous geographical entities
- in the limit the further reduction of the number of clusters leads to a two cluster solution where a first cluster describes the northern part of France and a second cluster consists of the *généralités* in the south

One last remark that can be made is about the difference between the tree clustering solution and the k-means solution for k = 6. By comparing the results for the tree clustering, represented by the dendrogram in Figure 2, and the composition of the third cluster for the 6-means clustering, given in Table 1, it can be seen that for the tree clustering method this third cluster is split up into two clusters, i.e. an eastern sub-cluster consisting of the *généralités* Bourgogne, Franche-Comté, Lyon-Dombes and Riom-Auvergne (called cluster 7) and a western sub-cluster with the *généralités* Auch, Bayonne, Bordeaux, La Rochelle, Limoges and Montauban (former cluster 3). In other words, a slightly enhanced 6-means clustering solution is equivalent to the tree clustering solution.

As a final conclusion about the application of clustering techniques to analyze the wheat prices for the *généralités* in the period 1756-1790, a summarizing graphical representation can be constructed. Figure 3 consists of seven geographical maps. Each of the seven clusters is represented by its own map on which the *généralités* that are a member of that cluster are represented by a contrasting color. After a convenient rearrangement of these maps it is evident that for the period 1756-1790 France can be represented by three main regions, i.e.

- northern part of France, represented by the clusters 1, 6 and 4
- middle of France, represented by the clusters 5 and 7
- southern part of France, represented by the clusters 3 and 2

Table 1 : Wheat Price for 33 Généralités - Period : 1756-1790
k-Means Clustering - Varying Number of Clusters

	Number of Clusters																			
	Six						Five					Four				Three			Two	
Généralités	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**1**	**2**
Alençon	·						·					·				·				·
Bretagne	·								·				·			·			·	
Caen	·						·					·				·				·
Rouen	·						·					·				·				·
Grenoble		·								·				·			·		·	
Languedoc		·								·				·			·		·	
Provence		·								·				·			·		·	
Roussillon		·								·				·			·		·	
Auch			·						·				·			·			·	
Bayonne			·						·				·			·			·	
Bordeaux			·						·				·			·			·	
Bourgogne			·						·				·			·			·	
Franche-Comté			·						·				·			·			·	
La Rochelle			·						·				·			·			·	
Limoges			·						·				·			·			·	
Lyon - Dombes			·						·				·			·			·	
Montauban			·						·				·			·			·	
Riom-Auvergne			·						·				·			·				·
Champagne				·							·				·			·		·
Lorraine				·							·				·			·		·
Metz				·							·				·			·		·
Alsace					·			·							·			·		·
Bourges					·			·							·			·		·
Moulins					·			·							·			·		·
Poitiers					·			·							·			·		·
Tours					·			·							·			·		·
Amiens						·	·					·						·		·
Flandres						·	·					·						·		·
Hainaut						·	·					·						·		·
Orléans						·	·					·						·		·
Paris						·	·					·						·		·
Soissons						·	·					·						·		·
Paris Ville						·	·					·						·		·

Principal Component Analysis - Component Loadings

A first decision that has to be taken in order to conduct a principal component analysis is the decision about the number of components that will be used. As explained in Section 4 of **Borghers [1]**, all the main diagnostics, that can be used in deciding about the number of components that has to be extracted, are based on the eigenvalues of the correlation matrix. Numerical information about the ordered eigenvalues can be found in Table 2. Apart from the eigenvalues this table gives also the explained variance for each of the corresponding components. The scree plot, i.e. the graphical representation of the ordered and largest eigenvalues, is presented in Figure 4.

From this information and depending on the criterion used, it can be concluded that the number of components, i.e. the number of hypothetical independent explanatory variables, common to each of the 33 *généralités*, must be either three or four. Whereas the scree test is favoring a three-component solution the use of **Kaiser**'s criterion would result in extracting four principal components.

Table 2 : Wheat Price for 33 Généralités - Period : 1756-1790
Eigenvalues and Explained Variance

	Eigenvalue		Variance %	
#	**Absolute**	**Cumulative**	**Absolute**	**Cumulative**
1	24.86610	24.86610	75.35182	75.35182
2	2.98029	27.84639	9.03118	84.38300
3	1.31784	29.16423	3.99346	88.37645
4	1.00517	30.16940	3.04598	91.42243
5	0.73901	30.90841	2.23942	93.66185

Figure 3 : Wheat Price for 33 Généralités - Period : 1756-1790
Tree Clustering Solution - Geographical Representation

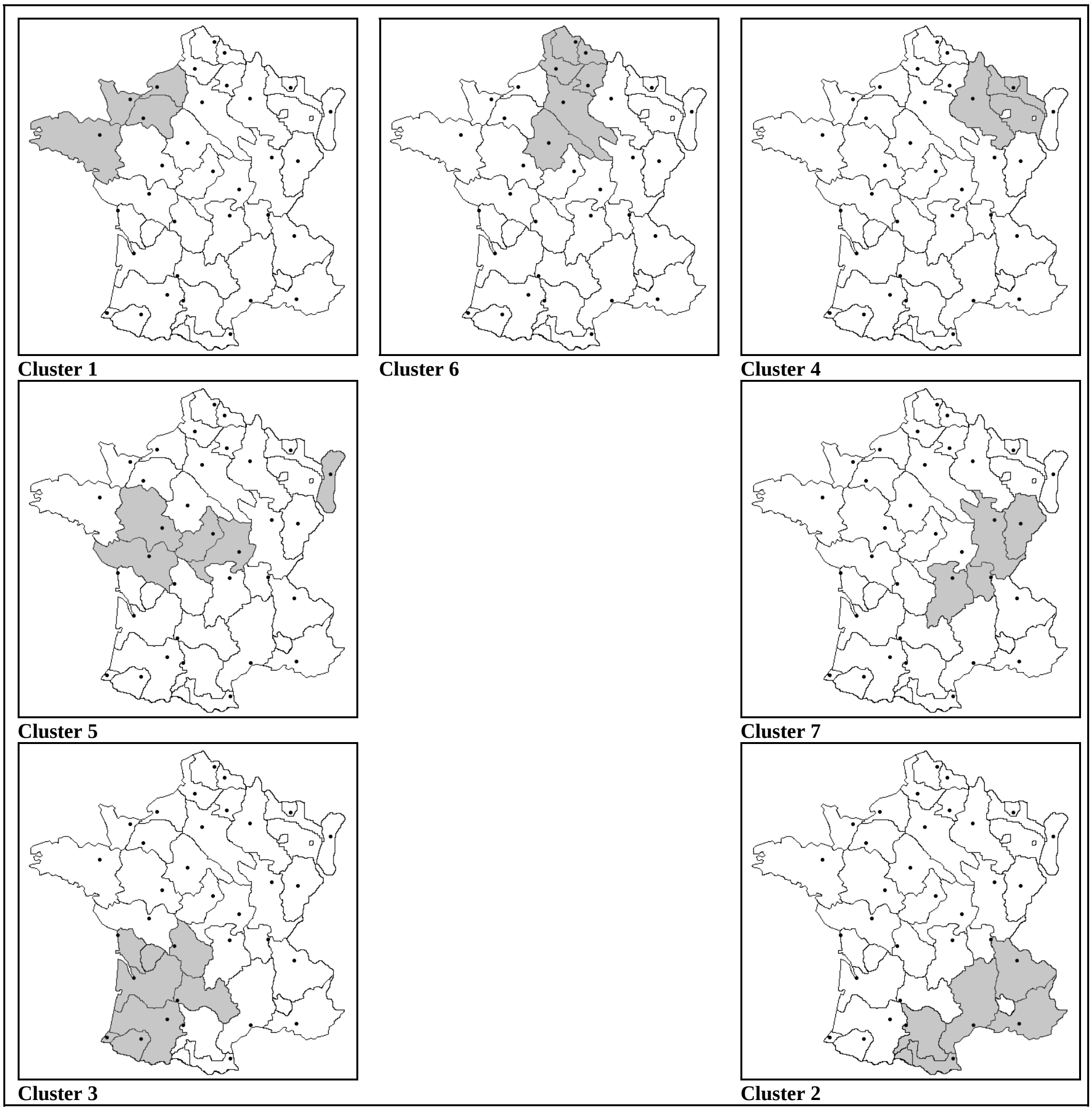


Figure 4 : Wheat Price for 33 Généralités - Period : 1756-1790
Largest Eigenvalues - Scree Plot

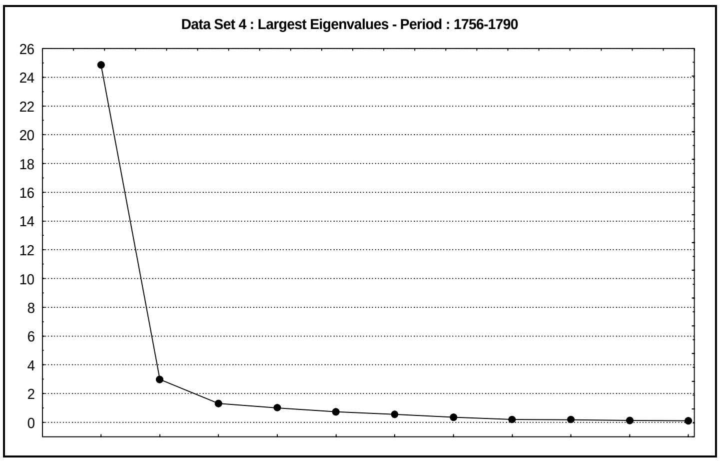


Table 3 : Wheat Price for 33 Généralités - Period : 1756-1790
Component Loadings - Varimax Rotation - 4- and 3-Component Solution

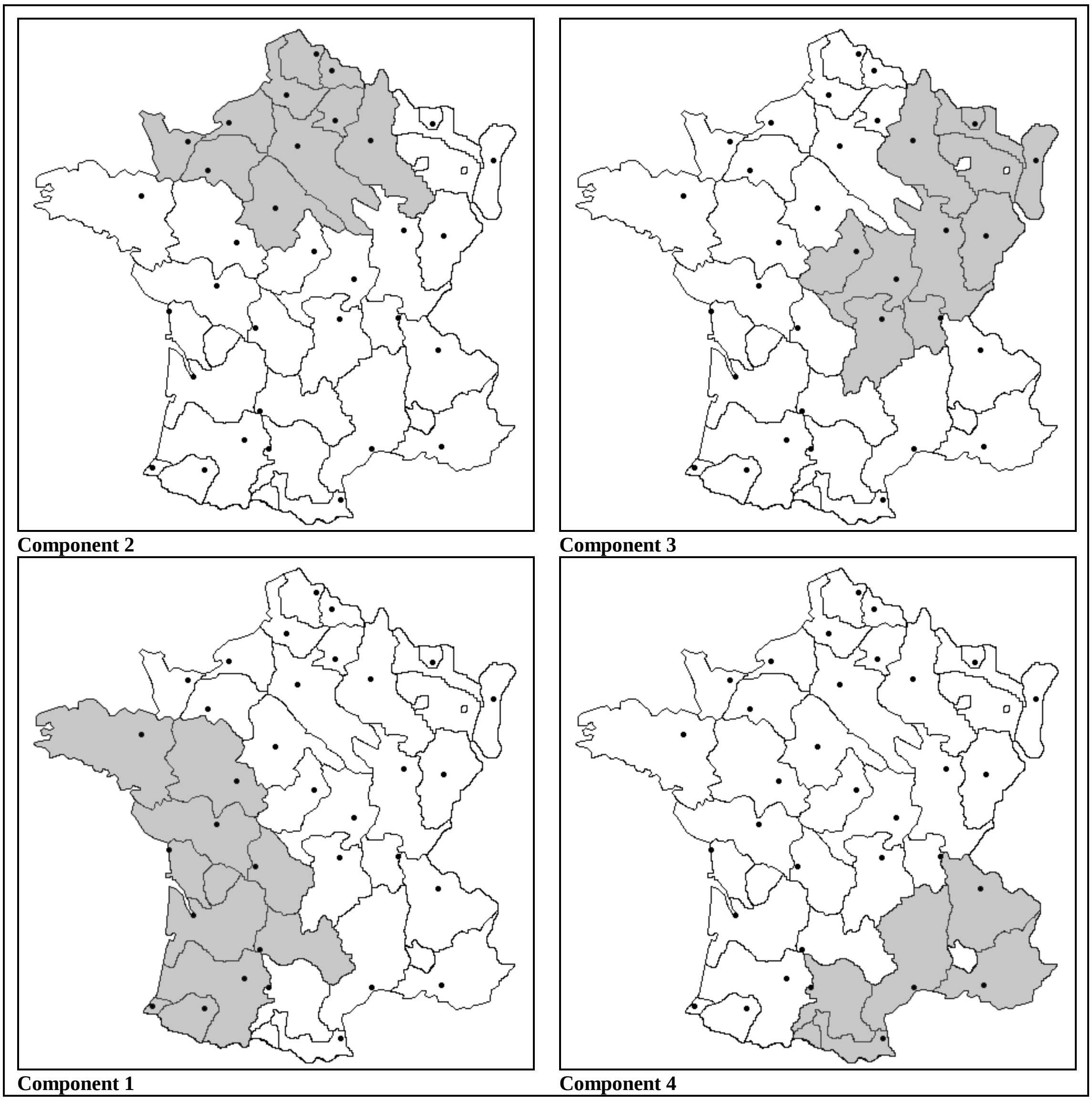
	Component Loadings						
	4-Component Solution				3-Component Solution		
Généralités	**Comp. 1**	**Comp. 2**	**Comp. 3**	**Comp. 4**	**Comp. 1**	**Comp. 2**	**Comp. 3**
Auch	**0.66**	0.19	0.42	0.51	**0.81**	0.22	0.44
Bayonne	**0.66**	0.36	0.40	0.32	**0.67**	0.40	0.44
Bordeaux	**0.78**	0.27	0.33	0.41	**0.82**	0.32	0.37
Bretagne	**0.62**	0.53	0.34	0.24	**0.58**	**0.57**	0.38
La Rochelle	**0.77**	0.35	0.30	0.37	**0.79**	0.40	0.35
Limoges	**0.67**	0.27	0.54	0.36	**0.71**	0.30	0.58
Montauban	**0.65**	0.24	0.49	0.46	**0.77**	0.28	0.52
Poitiers	**0.81**	0.36	0.34	0.25	**0.73**	0.42	0.40
Tours	**0.75**	0.51	0.31	0.19	**0.63**	0.57	0.36
Alençon	0.40	**0.78**	0.40	0.07	0.29	**0.80**	0.43
Amiens	0.20	**0.86**	0.22	0.15	0.20	**0.87**	0.23
Caen	0.43	**0.66**	0.45	0.10	0.34	**0.69**	0.49
Champagne	0.26	**0.65**	**0.64**	0.23	0.31	**0.65**	**0.65**
Flandres	0.24	**0.88**	0.12	0.17	0.25	**0.89**	0.14
Hainaut	0.24	**0.85**	0.30	0.25	0.30	**0.86**	0.31
Orléans	0.43	**0.72**	0.44	0.18	0.39	**0.74**	0.47
Paris	0.26	**0.81**	0.44	0.22	0.30	**0.81**	0.46
Rouen	0.20	**0.88**	0.31	0.15	0.20	**0.89**	0.33
Soissons	0.23	**0.86**	0.34	0.20	0.26	**0.87**	0.35
Paris Ville	0.22	**0.79**	0.33	0.19	0.25	**0.79**	0.35
Alsace	0.44	0.40	**0.73**	0.21	0.43	0.42	**0.75**
Bourges	0.57	0.37	**0.65**	0.14	0.48	0.40	**0.69**
Bourgogne	0.26	0.43	**0.78**	0.26	0.34	0.43	**0.78**
Franche-Comté	0.36	0.40	**0.77**	0.31	0.44	0.40	**0.78**
Lorraine	0.35	0.44	**0.70**	0.21	0.36	0.45	**0.72**
Lyon	0.33	0.35	**0.73**	0.43	0.51	0.35	**0.73**
Metz	0.32	0.52	**0.66**	0.16	0.30	0.53	**0.68**
Moulins	0.43	0.40	**0.75**	0.21	0.42	0.42	**0.78**
Riom-Auvergne	0.39	0.40	**0.71**	0.29	0.45	0.41	**0.72**
Grenoble	0.27	0.32	0.55	**0.65**	**0.63**	0.31	0.53
Languedoc	0.48	0.17	0.33	**0.75**	**0.85**	0.19	0.33
Provence	0.24	0.23	0.20	**0.87**	**0.77**	0.22	0.16
Roussillon	0.56	0.17	0.18	**0.74**	**0.90**	0.19	0.19
Expl. Var. %	**0.23**	**0.30**	**0.25**	**0.13**	**0.30**	**0.32**	**0.27**

In order to get a better insight into the conflicting situation in relation to the final decision about the number of components to extract it was decided to first compare both solutions. To facilitate this comparison and after rearranging the *généralités* according to their component loadings the three- and four-component solution were tabulated in one and the same table. These results can be found in Table 3. From these results the following conclusions can be drawn :

- the four-component model reveals four homogeneous geographical regions, i.e. the *généralités* in the north-west and the northern part of France, described by the second component, the *généralités* that are situated in the north-east and the east, explained by the third component, the *généralités* in the south-eastern part of France, bounded by the fourth component and the *généralités* that are situated in the south-west and west and for which the first component acts as the common explanatory variable
- by reducing the number of dimensions, and thus the number of components, from four to three results in a solution where the *généralités* in the south-east and south are linked with those in the south-west and west
- in both the four- and the three-component model the *généralité* Champagne is on the borderline between two regions, i.e. Champagne is almost equally shared by the component describing the *généralités* in the north and north-west and the component explaining the price behavior of the *généralités* situated in north-east and east
- also the *généralité* Bretagne is on the borderline, i.e. in the four-component model Bretagne is linked to the *généralités* situated in the south-west while in the three-component solution it is linked with the *généralités* in the north-west

An alternative and even more convenient way to visualize all these results is to use geographical maps. In Figure 5 the graphical representation consists of four maps, one for each of the components of the four-component model. On each of these four maps the *généralités*, explained by the corresponding component, are marked by using a contrasting foreground color. Taking into account the specific situation for the *généralités* Bretagne and Champagne, the transition from the four-component solution to the three-component solution can easily be made.

Figure 5 : Wheat Price for 33 Généralités - Period : 1756-1790
4-Component Solution - Geographical Representation

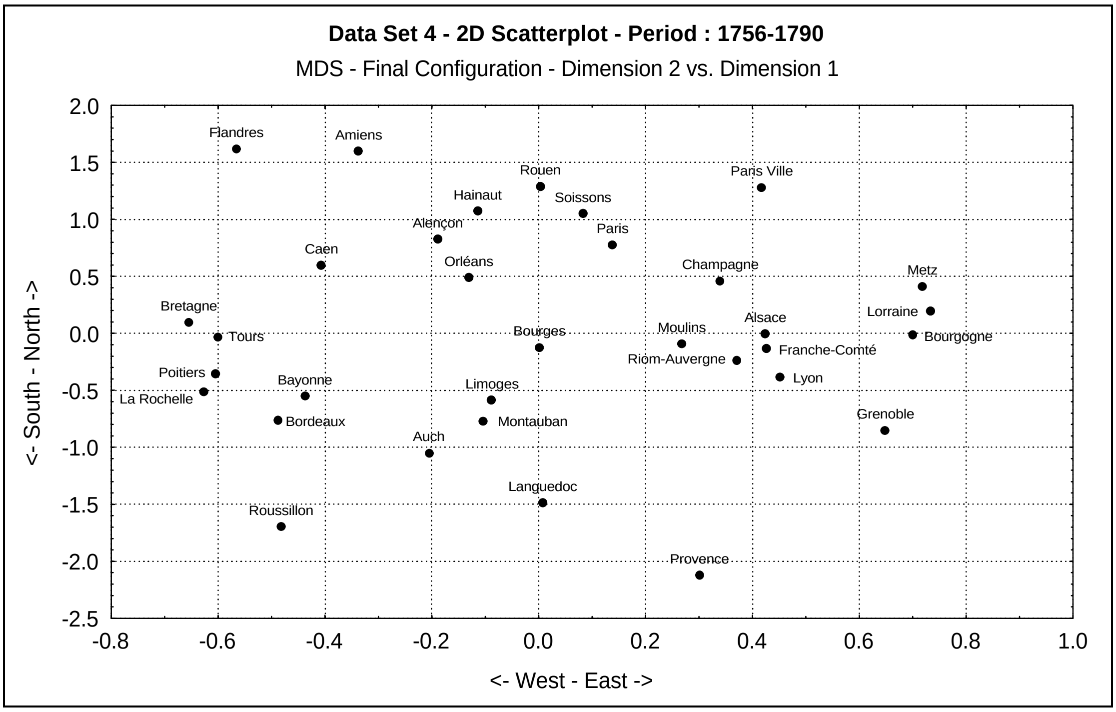


Multi Dimensional Scaling

Data set 4 was also used for a MDS analysis. Similar to the MDS analysis of the data sets 1, 2 and 3 in **Borghers [1]**, the **Pearson** correlation matrix for the 33 disaggregated price variables was used to start the analysis. In a first step a solution for three dimensions was derived. In a second step it was decided to use this three-dimensional solution as a starting configuration for a solution in two dimensions. The graphical representation of this final two-dimensional configuration can be found in Figure 6.

Apart from a relative small number of isolated cases, the two-dimensional scatterplot, representing the output of the MDS analysis, matches fairly well the geographical map of France. These isolated cases and most striking wrongly allocated *généralités* are : Paris Ville, Alsace, Languedoc, Provence, Roussillon and Bayonne.

Figure 6 : Wheat Price for 33 Généralités - Period : 1756-1790
Multi Dimensional Scaling - Final 2D Configuration



Not only the procedure followed in conducting a MDS analysis of the prices in the 33 *généralités* but also the final conclusion that follows from this analysis is the same as for the analysis of data set 1, 2 and 3 (See Section 7 in **Borghers [1]**). In the majority of the cases the physical geographical distance between the *généralités* seems to act as a good fitting explanatory variable in explaining the differences between the wheat prices in these *généralités* for the period 1756-1790.

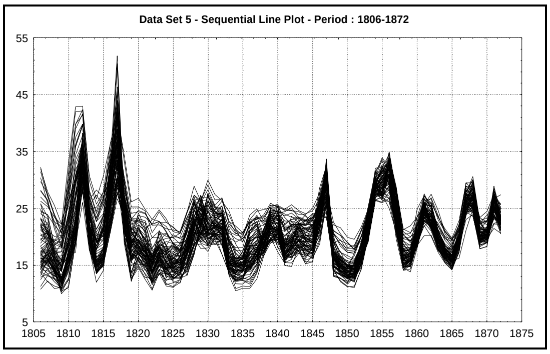
Section 3 : Period 1806-1872 - Départements

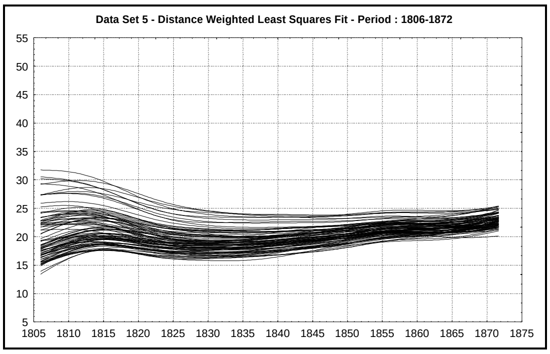
Introduction

The basic idea to divide France into *départements* goes back to the year 1778. It was the intention to create new geographical areas that should be comparable in size and wealth. Finally in 1790, 85 *départements* were replacing the provinces and *généralités* as the new administrative entities.

For the period 1806-1872 disaggregated data are available at the level of these new entities. From the original 90 *départements* for which the data are available only 85 of these regions will be used. The reason for excluding five of these data series from further analysis is that the missing values for these *départements* would have too large an influence on the final results. Further details about the 90 *départements* can be found in Appendix 1. A list of the *départements* is given in Appendix 4. In Figure 7 the 85 raw series, called data set 5, is graphed at the left and the corresponding weighted least squares fits are shown at the right.

Figure 7 : Wheat Price for 85 Départements - Period : 1806-1872
Distance Weighted Least Squares Fit





As could be expected the price evolution in the 85 *départements* can be compared with the evolution shown by the more aggregated series from data set 2 (1797-1830) and data set 3 (1831-1870) in **Borghers [1]**. The main characteristics of the price behavior of wheat during the period 1806-1872 can be summarized as follows :

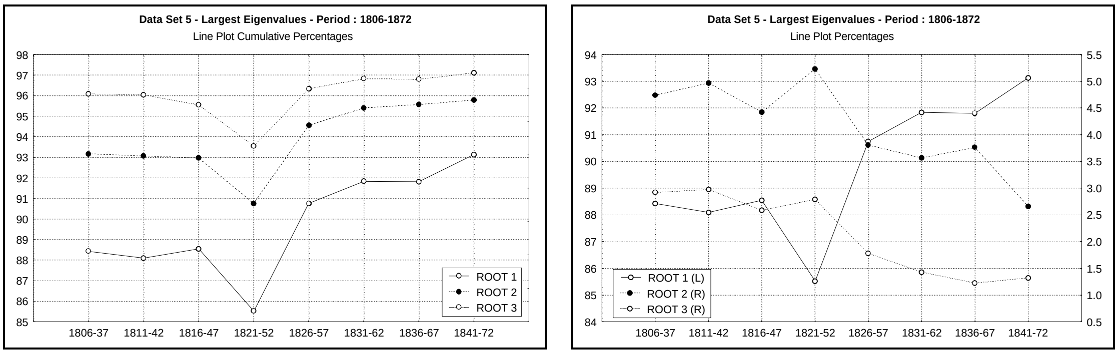
- fairly constant mean price level sustained during almost the whole period
- large peaks for the years 1812 and 1817
- appearance of quasi cyclical price behavior
- increased asynchronous evolution during the sub-period 1820-1845
- starting mainly around 1845 a gradually shrinking gap between highest and lowest prices

Principal Component Analysis - Component Loadings

Since data set 5 is covering a period of 67 years, it was decided to split the whole period 1806-1872 into smaller sub-periods of 32 observations each. A first advantage of working with these smaller subsets is that they are of a size comparable to the other data sets used so far. A second and major advantage, however, is that use can be made of the sliding window technique. With this technique the whole period can be analyzed by sliding a subset, consisting of a relative small number of observations, along the time axis. Starting in 1806 with a subset of 32 observations and shifting this subset repeatedly by five years will result in eight sub-periods, i.e. 1806-1837, 1811-1842, …, 1841-1872. It must be mentioned that from these eight sub-periods only the first and the last period will be non-overlapping and will lead to independent results.

For each of the eight sub-periods the eigenvalues of the respective correlation matrices were calculated. Based on the comparison of the eight scree tests, i.e. the graphical representation of the eigenvalues ordered in ascending order, it was decided that three components would be a reasonable compromise to describe the wheat prices for the 85 *départements* in each of the eight sub-periods. The relative importance of these three main components is given in Figure 8.

Figure 8 : Wheat Price for 85 Départements - Period : 1806-1872
Largest Eigenvalues - Scree Plot - Sliding Window



For each of the eight sub-periods the cumulated effect of these three components is given on the left of Figure 8. From this cumulated explanatory effect it can be seen that, apart from the period 1821-1852, the three principal components are responsible for at least 95% of the total variance of the price variables of wheat in the 85 *départements.* On the right side of Figure 8 the relative importance of the three principal components is represented in absolute percentages. This graph reveals that the explanatory power of the main and most dominant component ranges from about 88% for the period 1806-1837 to almost 94% for the period 1841-1872. In contrast to the increasing dominance of the main principal component, the explanatory power of the second and third component sharply decreased from about 5% and 3% for the period 1806-1837 to respectively 2.5% and 1.5% for the period 1841-1872.

Special attention should be paid to the sub-period 1821-1852. It is striking that for this period the explanatory contribution of the three components drops below the 94%. The main reason for this lower value is the relative sharp decrease in the contribution of the first and main component. This decrease is only partially compensated by the increase of the influence of the second and third component.

This last conclusion could be expected when the price behavior of the 85 variables during the sub-period 1821-1852 is taken into account. From Figure 7 it can be seen that, contrasting with both the previous and following periods, the behavior of the 85 series during the sub-period 1821-1852 is characterized by an asynchronous evolution. It is evident that in case of a much weaker common price behavior, the general and common explanatory principal component will be less dominant.

Starting around 1850 the price series for wheat in the 85 *départements* are following more or less the same pattern. As a consequence it can be expected that for those sub-periods, consisting of observations dating after 1850, the first principal component will gain importance. The more the sliding window will consist of observations dated after 1850, the more this first common explanatory component will be dominating. The graphical representation in Figure 8 clearly confirms this expectation. For the sub-period 1831-1872, i.e. the last of the eight sliding windows, the total cumulated effect of the first and second component is almost equal to the total cumulated explanatory power of the first three components for the period 1816-1847.

For the interpretation of the component loadings, representing the correlation between the 85 variables and the three components, one is confronted with a practical problem. For 85 variables, three components and eight sub-periods the total number of loadings is 2040! For this reason it was decided not to tabulate the component loadings but to represent the results graphically by using geographical maps. The resulting graphical representations can be found in Figure 9.

Figure 9 : Wheat Price for 85 Départements - Period : 1806-1872
Principal Components - 3-Component Model - Sliding Window

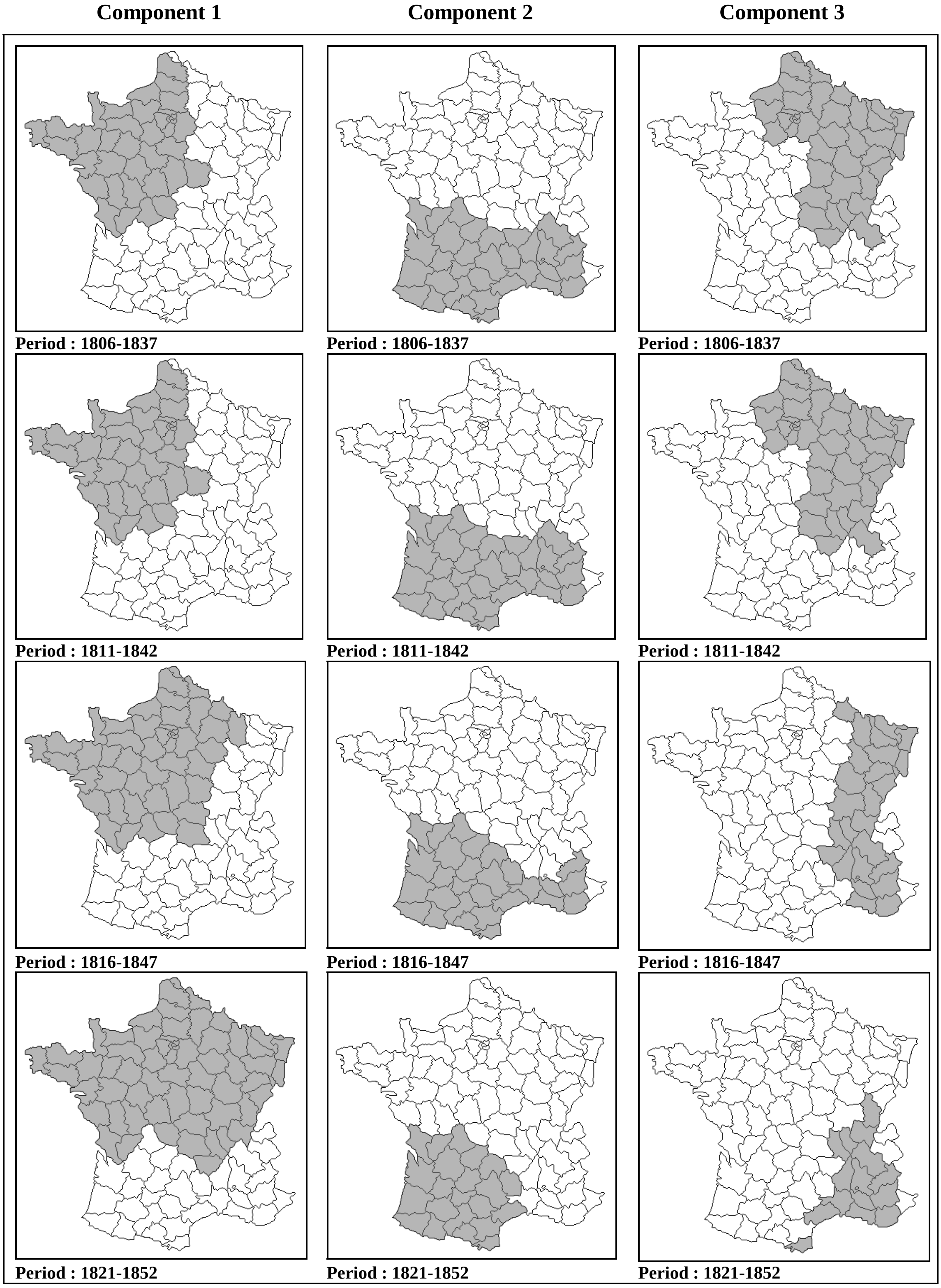


Figure 9 : Wheat Price for 85 Départements - Period : 1806-1872
Principal Components - 3-Component Model - Sliding Window (Continued)

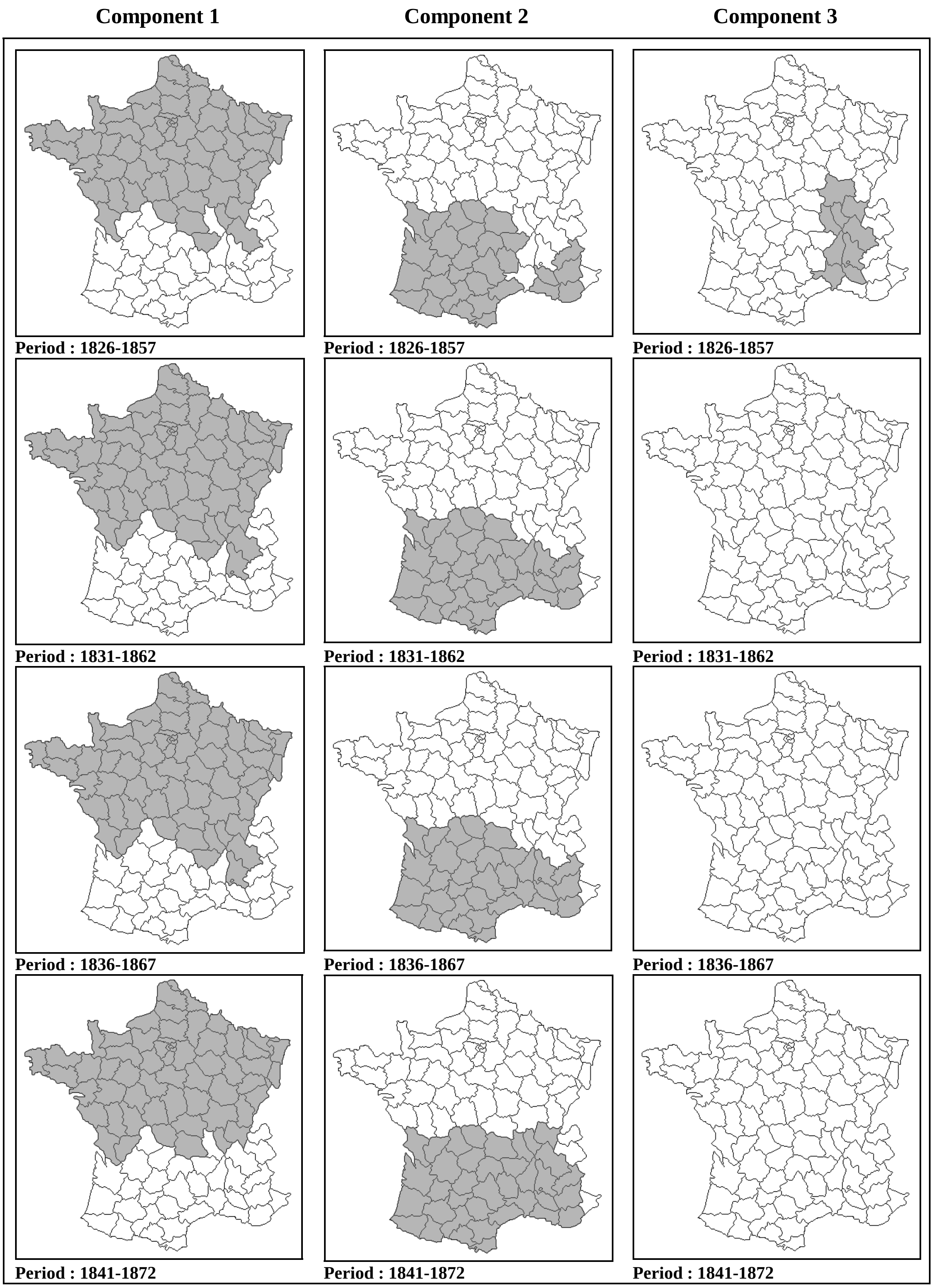


Figure 9 consists of three columns, i.e. one column per component. The criterion for coloring a particular *département* is that the correlation between the price series for that *département* with the principal component must be at least equal to 0.6. In other words, this means that in order to be called a dominant component the latter must explain at least one third, i.e. to be precise 36% or more, of the total variance of the corresponding price variable.

Conclusions

A first conclusion that follows from these results is about the interpretation of the three common components. From the maps in Figure 9 it can be seen that the first component is highly correlated with the price variables of the *départments* in the west, north-west and north. The third component can be seen as the explanatory variable for the *départments* in the north-east and the east, while the second principal component acts as the explanatory variable for the *départments* situated in the southern part of France.

This first conclusion is an illustration that the PCA is a technique that not only can be used to represent the original observed data by means of the smallest possible number of orthogonal explanatory components or variables, but that this technique can also be used as a classification method.

A second and by far the most important conclusion that can be drawn from these results is about the dynamical changing pattern that can be observed during almost the whole period 1806-1872. Using the results for the eight sub-periods two simultaneous occurring shifts can be distinguished. A first shift is related to the gradually linking of the *départements* from the eastern part of France with those from the northern part of France. The second shift is related to the *départements* of the south-east that are linked temporally with the *départements* in the eastern part. The final and net effect of these two simultaneous movements is the partition of France into a northern and a southern region.

Section 4 : Disaggregation versus Aggregation

Introduction

In this and the previous paper **Borghers [1]** use was made of the principal component analysis. The applications were entirely focused on the component loadings, i.e. the correlations between the principal components and each of the regional price series. In this section special attention will be paid to the behavior of these components as explanatory variables. To be more precise, the actual values of the individual observations of these principal components, called component scores, will be analyzed in more detail.

Principal components are not just ordinary explanatory variables. One of the main characteristics of each component is that it acts as an explanatory variable to the whole set of variables by explaining the maximum of the remaining variance left in the data after the extraction of the previous components. In other words the first component will contain a maximum of information scattered among the variables in the set, the second component, being orthogonal and thus not correlated with the first, will explain a maximum of the remaining information, etc. It follows that the first few components can be seen as a compact summary of the behavior described by the whole set of variables.

Intercorrelations of Principal Components

A first aspect that will be analyzed is the comparison of the component scores of the smaller data sets with those of the larger data sets. For this purpose the first and second component of the data sets 1, 2 and 3 will be compared with the first and second component of the larger data sets 4 and 5. For practical reasons data set 5, i.e. the data set for the period 1806-1872 and based on the 85 *départements*, was split up into two subsets. The first subset, called data set 5.1, will cover the period 1806-1830, while the second subset, called data set 5.2, will describe the regional price behavior during the period 1831-1870.

The six available data sets thus derived enable the calculation of the correlation between the first and the second component extracted from a smaller data set and those derived from a larger and more disaggregated data set. The numerical results can be found in Table 4, 5 and 6.

Table 4 : Component Scores - Intercorrelations - Period : 1756-1790

		Data Set 1		Data Set 4	
Data	**Components**	**Comp. 1**	**Comp. 2**	**Comp. 1**	**Comp. 2**
Set 1	**Comp. 1**	1.0000	0.0000	0.9504	0.1664
	Comp. 2	0.0000	1.0000	-0.0846	0.9500
Set 4	**Comp. 1**	0.9504	-0.0846	1.0000	0.0000
	Comp. 2	0.1664	0.9500	0.0000	1.0000

Table 5 : Component Scores - Intercorrelations - Period : 1806-1830

		Data Set 2		Data Set 5.1	
Data	**Components**	**Comp. 1**	**Comp. 2**	**Comp. 1**	**Comp. 2**
Set 2	**Comp. 1**	1.0000	0.1663	0.9878	0.0671
	Comp. 2	0.1663	1.0000	0.1249	0.9788
Set 5.1	**Comp. 1**	0.9878	0.1249	1.0000	0.0000
	Comp. 2	0.0671	0.9788	0.0000	1.0000

Table 6 : Component Scores - Intercorrelations - Period : 1831-1870

		Data Set 3		Data Set 5.2	
Data	**Components**	**Comp. 1**	**Comp. 2**	**Comp. 1**	**Comp. 2**
Set 3	**Comp. 1**	1.0000	0.0000	0.9970	-0.0004
	Comp. 2	0.0000	1.0000	0.0060	0.9973
Set 5.2	**Comp. 1**	0.9970	0.0060	1.0000	0.0000
	Comp. 2	-0.0004	0.9973	0.0000	1.0000

Prior to the formulation of any conclusion from these results, a remark must be made about the results for the period 1806-1830. It can be seen from Table 5 that the correlation between the first and second component for data set 2 is not equal to zero whereas the orthogonality of the components implies a zero correlation. The reason for this is that in order to correlate these components with those of data set 5.1 the observations of the components for data set 2 for the period 1797-1805 had to be left out from the calculation of the correlation coefficients.

The main and dominant conclusion that can be drawn from these results is that both the first and second component from the smaller data sets are highly correlated with the corresponding components for the larger data sets.

Component Scores and Aggregated Price Series

Another interesting aspect about the components is the way they are correlated with the general aggregated price series, i.e. data set 0. Detailed information about the six data sets is tabulated in Table 7. For each of these six data sets also the coefficients of correlation between the first two components and the general aggregated price series are given.

The general conclusion about these results is that the components based on the larger data sets do show almost exactly the same correlations as those obtained for the smaller data sets. The results for the period 1756-1790 seem to make a an exception on this rule. But even these results do not support the hypothesis that the larger the data set, the stronger the first two components will be correlated with the general aggregated price series.

Table 7 : Component Scores and General Aggregated Price Series - Correlations

		Data		Correlation	
Data	**Period**	**Variables**	**Observations**	**Comp. 1**	**Comp. 2**
Set 1	**1756-1790**	9	35	0.7792	0.6252
Set 4	**1756-1790**	33	35	0.6947	0.7190
Set 2	**1797-1830**	9	34	0.7468	0.6613
Set 5.1	**1806-1830**	84	25	0.7412	0.6711
Set 3	**1831-1870**	9	40	0.7462	0.6656
Set 5.2	**1831-1870**	85	40	0.7487	0.6629

In order to get an even better insight into the behavior of these first two components a graphical comparison can be made. To facilitate this comparison, use can be made of the graphs presented by Figure 10. The first row of Figure 10 consists of the general aggregated price series for each of the three sub-periods. In the second and third row the first and second component are given for the smaller data sets while these components for the larger data sets can be found on the fourth and fifth row. All graphs are overlaid by the distance weighted least squares fit.

Visual comparison of these plots do confirm the previous formulated conclusions. By far the most important characteristic specifically shown by these graphs is the behavior of the second component. In all six cases, i.e. for the smaller as well as for the larger data sets, the second component shows an almost cyclical behavior with a mean periodicity which might be very close to what could be called a business cycle, i.e. a periodicity of roughly four to nine years. This is a good example of one of these characteristics that needs further investigation by using more specific time series analysis techniques.

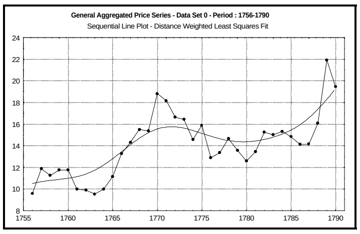
Component Scores and Cluster Means

In this last subsection the first two components will be confronted with the cluster means for the 2-means clustering solution. The comparison will be applied to each of the three smaller data sets. The idea behind this comparison is that, apart from the first two components, also the cluster means can be seen as summarizing the information of the data sets. The numerical results for this comparison are tabulated in Table 8, 9 and 10.

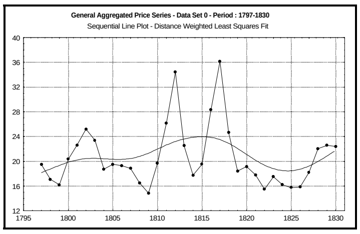
Before to proceed with the interpretation of the results it has to be taken into account that the first cluster consists of the southern regions while the second cluster relates to the northern part of France. The reverse is true for the principal components. Using the pattern of the component loadings as a criterion and after rotation by the varimax procedure, the first component can be associated with the northern regions, while the second component can be seen as an explanatory variable for the price behavior in the southern regions.

Figure 10 : Component Scores and General Aggregated Price Series
Sequential Line Plot - Distance Weighted Least Squares Fit

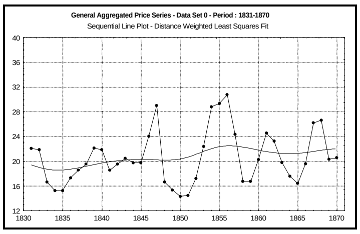
Period : 1756-1790 | **Period : 1797(1806)-1830** | **Period : 1831-1870**



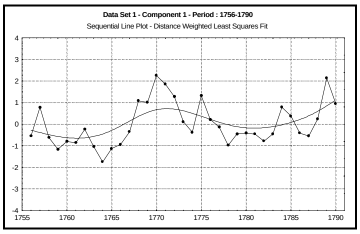
Data Set 0



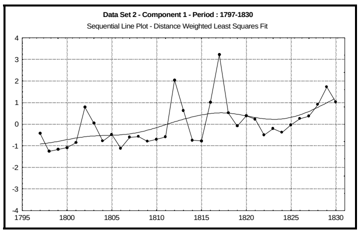
Data Set 0



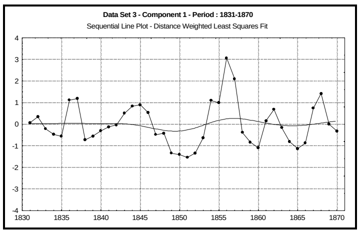
Data Set 0



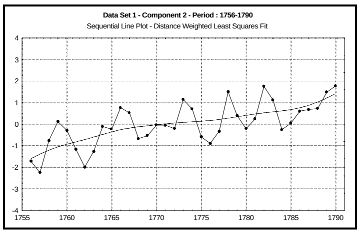
Data Set 1 - Component 1



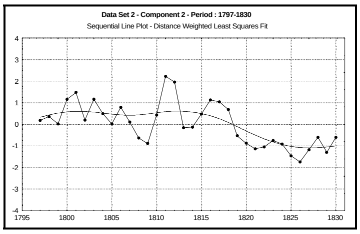
Data Set 2 - Component 1



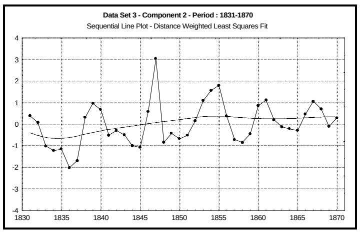
Data Set 3 - Component 1



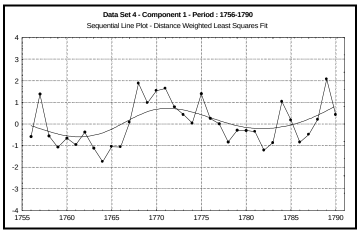
Data Set 1 - Component 2



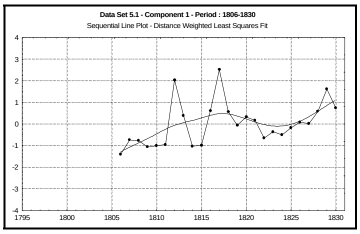
Data Set 2 - Component 2



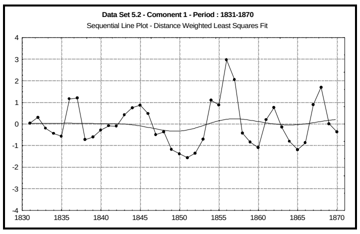
Data Set 3 - Component 2



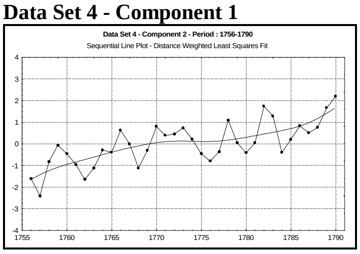
Data Set 4 - Component 1



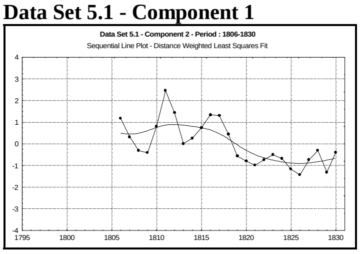
Data Set 5.1 - Component 1



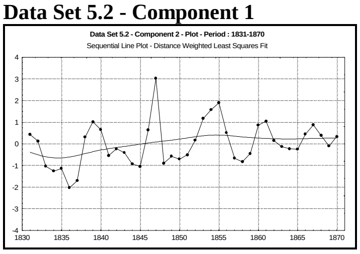
Data Set 5.2 - Component 1



Data Set 4 - Component 2



Data Set 5.1 - Component 2



Data Set 5.2 - Component 2

The conclusions can be briefly summarized as follows :

- as could be expected the cluster means are (highly) correlated while the principal components, being orthogonal, are not correlated at all
- the correlations between cluster mean and component for the same geographical area of France, i.e. north-north or south-south, are substantial and are ranging from 0.78 to 0.86
- the correlations between the northern and the southern regions are systematically lower and are situated between 0.51 and 0.58

Table 8 : Component Scores and Cluster Means - Correlations
Data Set 1 - Period : 1756-1790

	Clusters		Components	
	Mean 1	**Mean 2**	**Comp. 1**	**Comp. 2**
Mean 1	1.0000	0.8791	0.5126	0.8557
Mean 2	0.8791	1.0000	0.8518	0.5214
Comp. 1	0.5126	0.8518	1.0000	0.0000
Comp. 2	0.8557	0.5214	0.0000	1.0000

Table 9 : Component Scores and Cluster Means - Correlations
Data Set 2 - Period : 1797-1830

	Clusters		Components	
	Mean 1	**Mean 2**	**Comp. 1**	**Comp. 2**
Mean 1	1.0000	0.8886	0.5508	0.8290
Mean 2	0.8886	1.0000	0.8615	0.5052
Comp. 1	0.5508	0.8615	1.0000	0.0000
Comp. 2	0.8290	0.5052	0.0000	1.0000

Table 10 : Component Scores and Cluster Means - Correlations
Data Set 3 - Period : 1831-1870

	Clusters		Components	
	Mean 1	**Mean 2**	**Comp. 1**	**Comp. 2**
Mean 1	1.0000	0.9547	0.6197	0.7828
Mean 2	0.9547	1.0000	0.8153	0.5774
Comp. 1	0.6197	0.8153	1.0000	0.0000
Comp. 2	0.7828	0.5774	0.0000	1.0000

Section 5. Including External Data

One point of criticism about the given analysis might be that it was entirely based on the data from one single source, i.e. the data published by **Labrousse**. To what extend these data can be considered as representative remains an open question. Although using one and the same data source may give the (subjective) appearance of credibility it still ought to be verified.

In order to check the performance of external collected data an additional data set was constructed. This new data set, henceforth called data set 6, consists of six price series covering about exactly the same period, i.e. the period 1756-1790. More precise information about these six series can be found in Appendix 1.

At least two series from data set 6 can be compared with the corresponding series from data set 4, i.e. both data sets contain price series for Montauban and Poitiers. In order to prevent any confusion the series will be numbered with 'one' for the series from data set 4 and with 'two' for the series originating from data set 6.

Since for both Montauban as well as for Poitiers two series are available they can easily be compared by plotting both series along the time axis. As a first example the two series describing the price of wheat in Montauban are given in Figure 11. Simple visual inspection of both graphical representations reveals some interesting characteristics.

From Figure 11 it can be seen that there is almost no difference in level between the two price series for Montauban. Apart from the decade between 1765 and 1775 both time series are remarkably close. No wonder that the **Pearson** correlation coefficient between the two series for Montauban is 0.96. This last result can be found in the correlation matrix tabulated in Table 11. The enlarged data set from Table 11 consists of the six series from data set 6 completed with the data for Montauban and Poitiers from data set 4.

To enable the comparison of the two price series for Poitiers Figure 12 was constructed. From the graphical representation of the two series it is evident that, in contrast to the situation for the series for Montauban, the two series for Poitiers clearly behave on a different level. The main reason for the difference in level is the different unit of measurement that was used for the standard weight, i.e. Setier de Paris (240 livres) for Poitiers 1 and hectoliter for Poitiers 2. Since further analysis will be based on the correlation concept it was decided not to correct one or other series.

Figure 11 : Price Series for Montauban 1 & 2 Sequential Line Plot

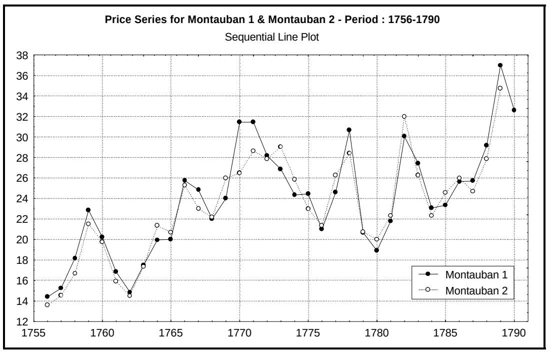
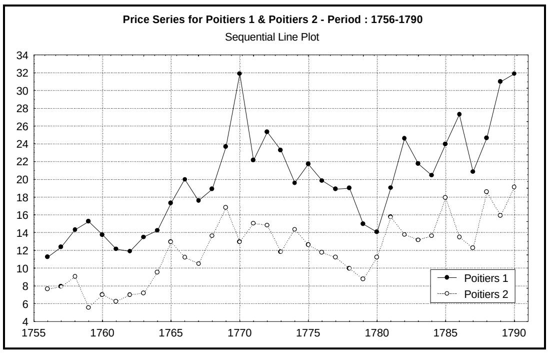


Figure 12 : Price Series for Poitiers 1 & 2 Sequential Line Plot



An even more interesting characteristic resulting from the comparison is the difference of the turning points of the series. A close inspection of the peaks and the lowest values leads to the conclusion that the price series for Poitiers from data set 6 is leading the corresponding series of data set 4. This last finding can be quantified by the cross-correlation function between these two series. The visualization of the cross-correlation function can be found in Figure 13. In a static context, and thus leaving the time shifts out of consideration, the two series for Poitiers are showing a correlation coefficient of almost 0.80. In other words each of these variables can be accounted for explaining more than 60% of the variance of the other.

Table 11 : Enlarged Data Set 6 - Correlation Matrix

	Aix	Arles	Drag.	Mont. 2	Mont. 1	Poit. 2	Poit. 1	Toul.
Aix	**1.00**	**0.70**	**0.86**	0.71	0.66	0.57	0.60	0.79
Arles	**0.70**	**1.00**	**0.87**	0.71	0.66	0.54	0.61	0.74
Draguignan	**0.86**	**0.87**	**1.00**	0.75	0.65	0.58	0.58	0.81
Montauban 2	0.71	0.71	0.75	**1.00**	**0.96**	0.69	0.85	0.97
Montauban 1	0.66	0.66	0.65	**0.96**	**1.00**	0.63	0.86	0.94
Poitiers 2	0.57	0.54	0.58	0.69	0.63	**1.00**	**0.78**	0.64
Poitiers 1	0.60	0.61	0.58	0.85	0.86	**0.78**	**1.00**	0.78
Toulouse	0.79	0.74	0.81	0.97	0.94	0.64	0.78	1.00

Figure 13 : Price Series for Poitiers 1 and Poitiers 2 Cross-Correlation Function

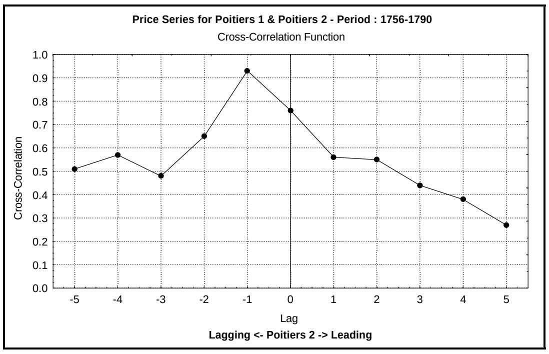
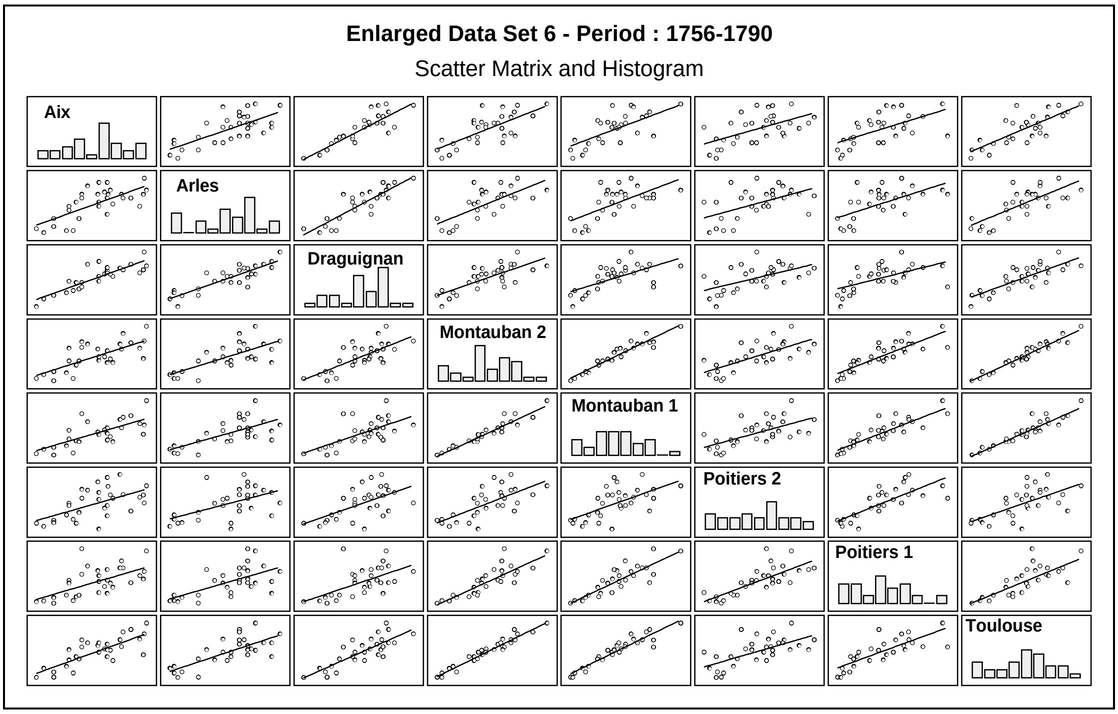
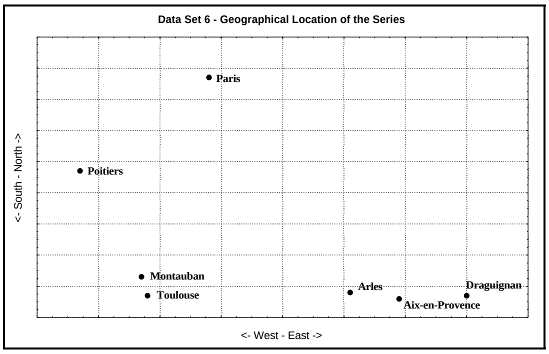


Figure 14 : Enlarged Data Set 6 - Scatter Matrix and Histogram



A more detailed analysis of the characteristics of the static interactions between the price series from the enlarged data set 6 can be based on the correlation matrix given in Table 11. The scatter matrix presented in Figure 14 and the actual distances between the geographical locations involved are a valuable help for the interpretation of the correlation coefficients. The graphical representation of these distances can be found in Figure 15. On this map the relative position of Paris can be used as a point of reference.

Figure 15 : Data Set 6 - Geographical Location of the Series



The conclusions about the analysis of the correlation coefficients can be formulated as :

Arles, Aix-en-Provence and Draguignan

- the correlation between the three price series for Arles, Aix and Draguignan are high
- the correlation coefficients between these three locations and both the series for Poitiers are smaller than the correlations with Toulouse and the two series for Montauban

Montauban

- the two series for Montauban are highly correlated, i.e. a correlation coefficient of 0.96
- both the series for Montauban are also highly correlated with the series for Toulouse

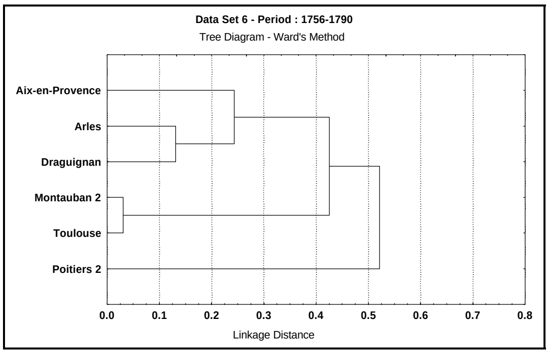
Poitiers

- the highest correlations are those between the two series for Poitiers and the correlations between Poitiers and Montauban and Toulouse
- the smallest correlation coefficients are those between the two price series for Poitiers and the three locations in the south-east

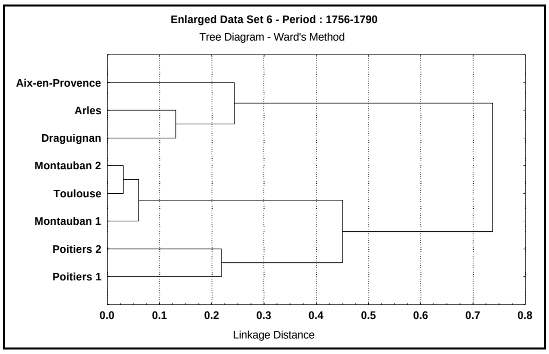
One can conclude that the closer the locations are the higher the correlation coefficients for the price variables will be.

An alternative approach for the analysis of the interaction of the price series of data set 6 is to use multivariate techniques. The first technique that was used was a cluster analysis of the six additional price series. The final result of the analysis can be represented by the hierarchical structure for the six series. The resulting tree diagram is given at the left of Figure 16.

Figure 16 : Data Set 6 and Enlarged Data Set 6 - Tree Diagram



Data Set 6



Enlarged Data Set 6

From the tree diagram three clusters can be identified. A first cluster with the three series for Arles, Aix and Draguignan, a second cluster with the series for Montauban and Toulouse and a third cluster consisting of just one single series, i.e. the price series for Poitiers.

An interesting experiment is to repeat the analysis but to replace the six series of data set 6 by the enlarged data set. The idea behind this experiment is to see how the two series for Potiers and the two series for Montauban will behave within the hierarchy. The resulting tree diagram from this experiment is represented by the right graph of Figure 16.

Not totally unexpected it turns out that also this second tree diagram consists of three clusters. The first cluster remains unchanged with the three regions in the south-east of France, i.e. Arles, Aix-en-Provence and Draguignan. The second cluster is jointed by the second series for Montauban and the two price series for Poitiers are building up the third cluster.

In order to get an even better insight into the behavior of the six additional price series of data set 6 within the 33 series of data set 4 the previous experiment can be generalized. Therefore it was decided to perform a cluster analysis for the 39 series resulting from combining the data from data set 4 with the six additional series from data set 6. The purpose of this experiment is to get information on the precise effect of the introduction of the additional price series where the effect refers to both the number and the exact composition of the clusters.

The results of this comparison can be represented in at least three different ways. The first and most obvious representation is the tree diagram of Figure 17. The analysis of the resulting tree diagrams for the combined data set 4 and 6 becomes very appealing when it is compared with the tree diagram for data set 4. As can be seen from Figure 18 the tree diagram for the price series of data set 4 reveals four main clusters. These four clusters can roughly be situated respectively in the north, the north-east and east, the south-east and the south-west and west of France. Taking the hierarchy into consideration the cluster for the north-east and east is much closer to the southern part of France than the cluster describing the northern part of France.

The tree diagram for data set 4 shows both an important resemblance but also interesting differences with the tree diagram obtained after the introduction of the six additional price series from data set 6. Concentrating on the two clusters in the northern part of France the resemblance between the two tree diagrams is striking. For both solutions two clusters are describing the price behavior in the north. The only difference between the two clusters is the allocation of the price series for the Champagne region. This last result is not totally surprising. From this and previous results one can conclude that the price series for Champagne can be situated just between the north-west and the north-east of France.

Even more interesting are the results obtained for the south. The introduction in the analysis of the six additional series from data set 6 resulted in a solution with three instead of just two clusters. To better understand the important difference it is preferable to make use of two alternative representations of the results of the cluster analysis.

A first alternative is to represent each cluster and its composition on a separate map of France. These maps can be found respectively as Figure 19 for the combined data set and as Figure 20 for the clusters obtained for the series of data set 4.

The third alternative representation, though not a real graphical representation, is at least as informative. Almost exactly the same information as represented by the tree diagram and the geographical map can also be represented in a tabulated form. A good example for this third alternative can be found in Table 12.

Figure 17 : Data Set 4 and Data Set 6 - Cluster Analysis - Tree Diagram

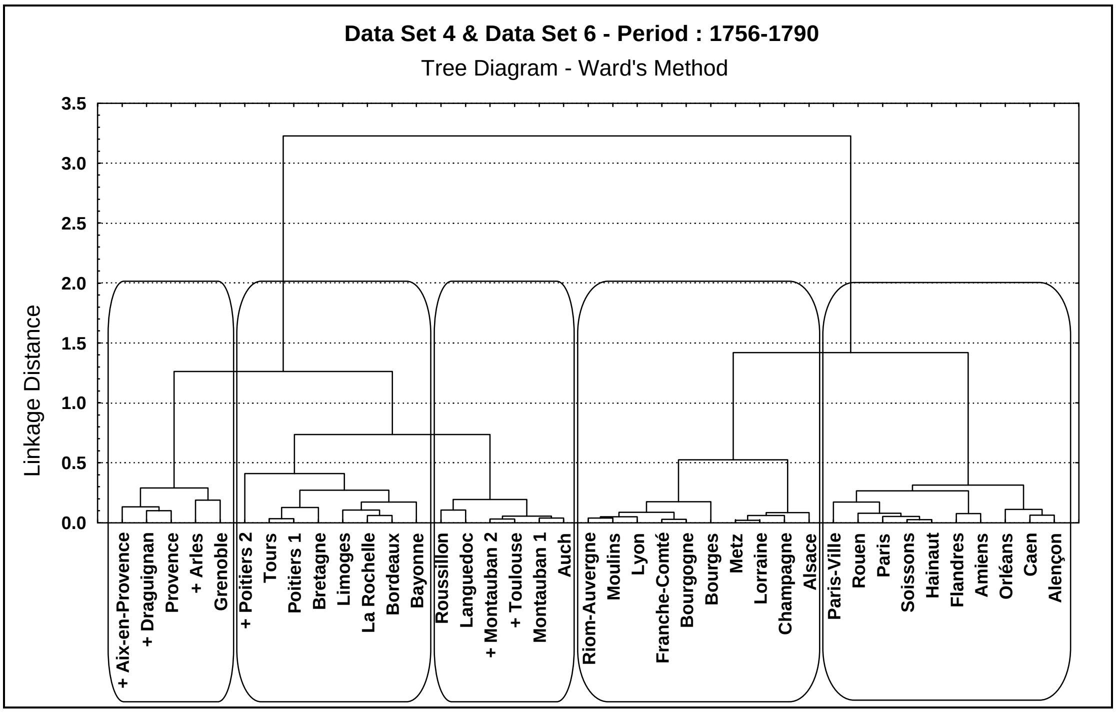


Figure 18 : Data Set 4 - Cluster Analysis - Tree Diagram

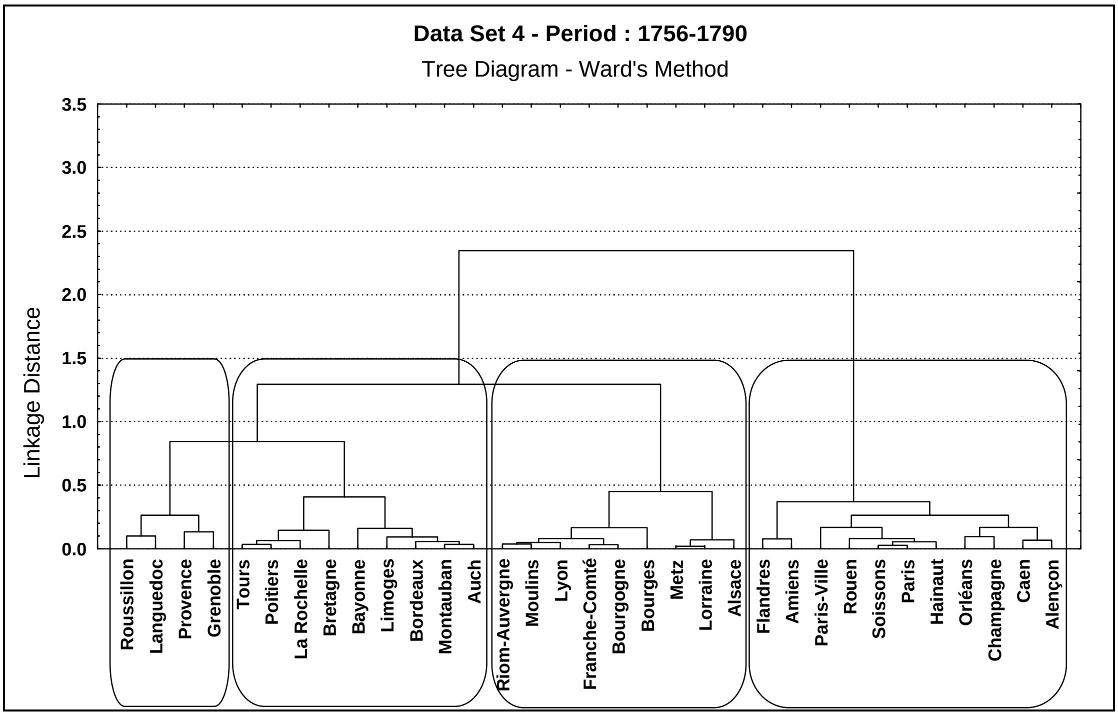


Figure 19 : Data Set 4 and Data Set 6 - Cluster Analysis - Map of Clusters

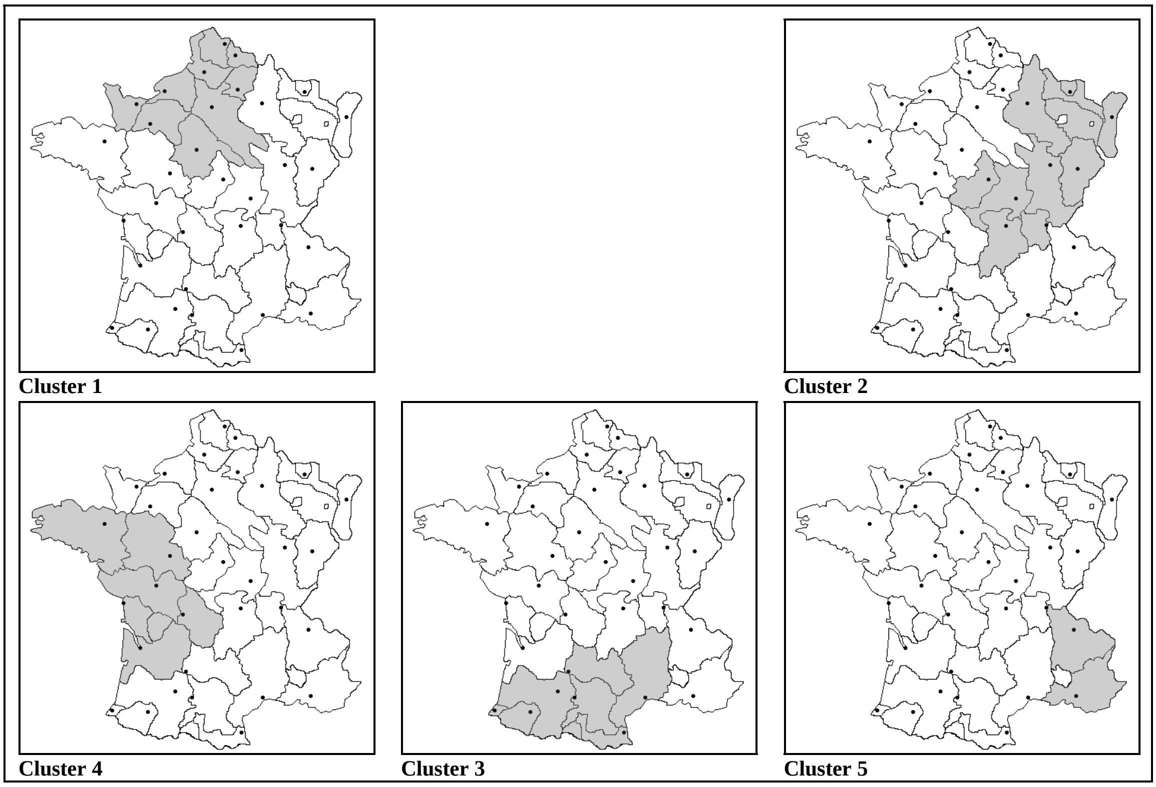


Figure 20 : Data Set 4 - Cluster Analysis - Map of Clusters

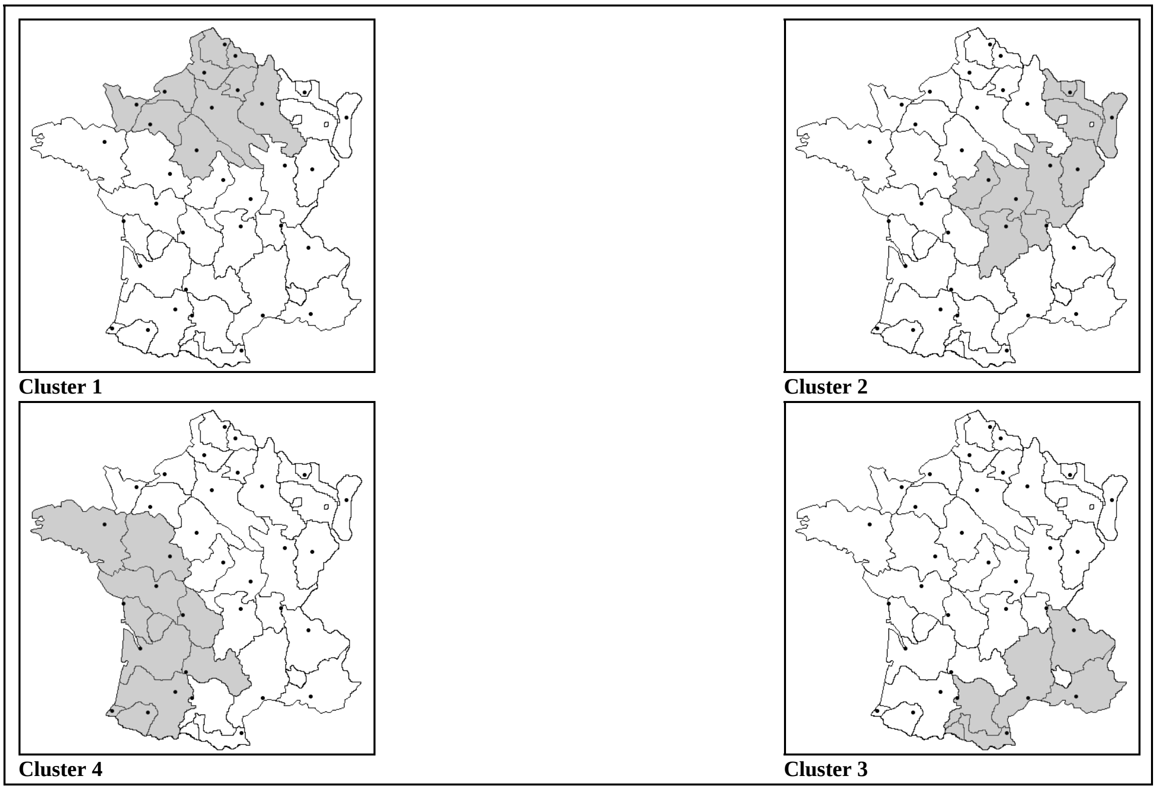


Table 12 : Data Set 4 and Combined Data Set 4 & 6 - Cluster Members

	Data Set 4 Clusters				Data Set 4 & 6 Clusters				
	1	2	3	4	1	2	3	4	5
Alençon	·				·				
Caen	·				·				
Champagne	·					·			
Orléans	·				·				
Hainaut	·				·				
Paris	·				·				
Soissons	·				·				
Rouen	·				·				
Paris-Ville	·				·				
Amiens	·				·				
Flandres	·				·				
Alsace		·				·			
Lorraine		·				·			
Metz		·				·			
Bourges		·				·			
Bourgogne		·				·			
Franche-Comté		·				·			
Lyon		·				·			
Moulins		·				·			
Riom-Auvergne		·				·			
Auch			·				·		
Montauban 1			·				·		
Bordeaux			·					·	
Limoges			·					·	
Bayonne			·					·	
Bretagne			·					·	
La Rochelle			·					·	
Poitiers 1			·					·	
Tours			·					·	
Grenoble				·					·
Provence				·					·
Languedoc				·			·		
Roussillon				·			·		
Toulouse							·		
Montauban 2							·		
Poitiers 2								·	
Arles									·
Draguignan									·
Aix-en-Provence									·

This table is an excellent help in explaining the switch from two to three clusters to describe the clustering of the price series of the combined data set for the southern part of France. The introduction of the series for Arles, Aix-en-Provence and Draguignan resulted in a correct allocation within the cluster covering the south-east. These three series are putting much more weight on the close proximity of the point of gravitation for the region. As a consequence both the Provence, including the new series, and Grenoble are incorporated within a separate cluster and are no longer associated with the neighboring Languedoc and Roussillon regions.

The same reasoning can be used to better understand the splitting up of the larger cluster for the west and south-west. The effect of the introduction of an additional price series for Toulouse and a second series for Montauban is twofold. The first effect is the incorporation of the Languedoc and Roussillon. The second effect is the rejection of the western regions. The final result of the combined effect is the appearance of a separate cluster describing the whole southern part of France.

Also the new and separate cluster for the west of France is the result of two effects. The effect of rejection of the regions in the west, due to the introduction of two new series in the south, is enforced by the introduction of a second series for Poitiers.

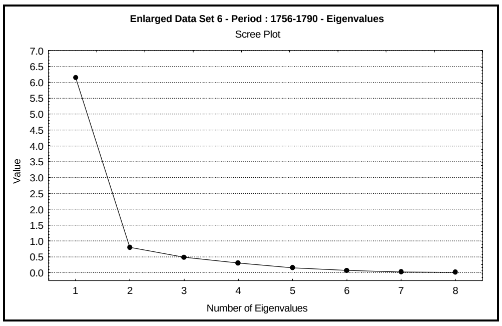
The results about the behavior of the price series of data set 6 in applying a cluster analysis can be easily summarized. None of the newly introduced price series is violating the pattern and order that was found after the analysis of the 33 series of data set 4.

The second multivariate technique that was used for the analysis of the price series from the enlarged data set 6 is the principal component analysis. A first step in the analysis is the calculation of the **Pearson**-correlation coefficients for the eight series. This matrix can be found in Table 11. The second step is to calculate the eigenvalues of the correlation matrix. The eigenvalues are then used as an important diagnostic to decide upon the number of components to retain. The basic information can be found in respectively Table 13 for the tabulated eigenvalues and in Figure 21 for the graphical representation.

Table 13 : Enlarged Data Set 6
Eigenvalues

	Eigenvalues		Cumul. Eigenvalues	
#	**Raw**	**%**	**Raw**	**%**
1	6.1546	76.9326	6.1546	76.9326
2	0.7978	9.9719	6.9524	86.9045
3	0.4839	6.0487	7.4362	92.9531
4	0.3010	3.7626	7.7373	96.7157
5	0.1545	1.9316	7.8918	98.6473
6	0.0701	0.8762	7.9619	99.5235
7	0.0242	0.3028	7.9861	99.8263
8	0.0139	0.1737	8.0000	100.0000

Figure 21 : Enlarged Data Set 6
Scree Plot



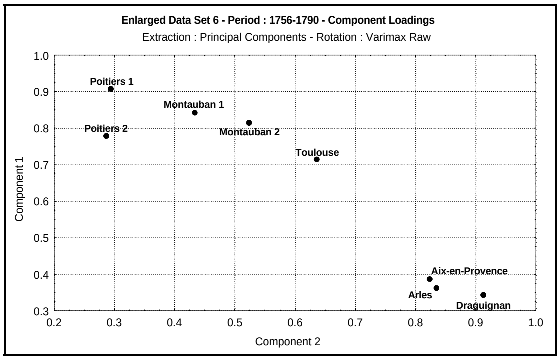
Using the previous information it was decided to opt for just two dimensions leading to the extraction of two principal components as a third step in the analysis. In other words two dimensions seem to be sufficient in explaining almost 87% of the total variance of the eight price series.

The fourth and last step is to calculate the component loadings, i.e. the correlations between the price variables and each of the two orthogonal components or hypothetical explanatory variables. To facilitate the interpretation of the component loading the solution was rotated using the varimax criterion. A tabulated version of the final results can be found in Table 14 while the graphical and geographical representation is given by Figure 22.

Table 14 : Enlarged Data Set 6
Component Loadings

	Loadings	
	Comp. 1	**Comp. 2**
Aix-en-Provence	0.3861	0.8242
Arles	0.3611	0.8349
Draguignan	0.3422	0.9134
Montauban 2	0.8138	0.5242
Montauban 1	0.8422	0.4335
Poitiers 2	0.7778	0.2871
Poitiers 1	0.9066	0.2945
Toulouse	0.7143	0.6362
Explained Variance	3.7052	3.2471
%	0.4632	0.4059

Figure 22 : Enlarged Data Set 6
Component Loadings



The final result is a good example of what can be called a simple structure. It demonstrates the potential capabilities of the principal component technique as a tool for data reduction. Even with only eight variables it turned out that the interpretation of the correlation matrix was not that simple. The interpretation of the rotated final solution however is rather straightforward.

Each of the two components is responsible for a clear and distinct group of locations. A first component is highly correlated with the price series for Arles, Aix and Draguignan while the correlations with the remaining series is rather weak. The behavior of the second component is totally the opposite, i.e. weak correlations with the regions in the south-east but highly correlated with the other regions.

Apart from this general and interesting characteristic Figure 22 reveals some further peculiarities about the precise allocation of the eight series. The three locations in the south-east are not exactly in the correct order but they are at least in close proximity of each other. However questions can be raised about the location of the other five series. Although the sequence Poitiers-Montauban-Toulouse seems not to be the major problem the orientation west-east instead of north-south may raise some questions.

Section 6 : Final Conclusions and Remarks

The main purpose of this paper was to investigate to what extend the preliminary results presented in a previous paper (See **Borghers [1]**) could be further refined by replacing the aggregated regional series by using disaggregated price series. The main conclusions resulting from the analysis of these 'larger' data sets can easily be summarized.

Period 1756-1790 - Généralités

Irrespective of the exploratory statistical technique used, the price behavior of wheat in France can be split up in both a northern and a southern region as well as a western and an eastern region. Furthermore it was found that the physical geographical distance between the *généralités* seems to act as a good fitting explanatory variable in explaining the differences between the wheat prices in these *généralités* for the period 1756-1790.

Period 1796-1872 - Départements

By using a sliding window technique it was shown that the dynamical changing pattern among the wheat prices in the 85 *départements* is characterized by two simultaneous acting shifts. The final and net effect of these two simultaneous movements is that at the end of the period, i.e. the period 1831-1872, the French territory can be partitioned into a northern and a southern region.

Disaggregation versus Aggregation

The main and dominant conclusion that can be drawn from the comparison of periods and sub-periods and of disaggregated and aggregated data is that in all these cases the price series are characterized by an almost cyclical behavior with a mean periodicity of roughly four to nine years.

Including External Data

In order to check whether the results derived from the analysis of the data given by **Labrousse [13]** are representative, external collected data (data set 6) were introduced in the analysis. The final conclusion is that none of the newly introduced price series is violating the general pattern that was found in the analysis of the series describing the price behavior in the 33 *généralités*.

A last remark must be made about the comparison of the derived results in this paper and those given by **Weir [9, p. 210]**. Both the analysis by **Weir** and the analysis followed in this paper are based on exactly the same data set 4, i.e. the price series for the 33 *généralités* for the period 1756-1790. However whereas **Weir** uses these series after detrending, i.e. by taking deviations from the moving average values, the analysis in this paper is based on the raw time series.

A second point on which both approaches differ is the basic idea behind the derived clusters. The clusters presented by **Weir** are based on three a priori chosen reference regions, i.e. Paris, Moulins and Languedoc. From the graphical representation of these clusters, given in Appendix 5, it can be seen that the price correlations of the regions surrounding these reference points are declining in accordance to their distance from the chosen center. The clusters derived in this paper are not a priori defined but are the outcome of descriptive exploratory statistical techniques. At least in one respect they confirm the findings by **Weir**, i.e. the association between geographical distance and the correlation between the wheat prices.

Appendix 1 : Data

1756<- Set 0 ->1870

1756<- Set 1 ->1790 | **1797<- Set 2 ->1830** | **1831<- Set 3 ->1870**

1756<- Set 4 ->1790

1806<- Set 5 ->1872

1806<- Set 5.1 ->1830 | **1831<- Set 5.2 ->1872**

1756<- Set 6 ->1790

Data Set 0 : Prix Moyens Nationaux Annuels de Froment - Par Année Civile

Source	**:** See **Labrousse et al. [14, pp. 9-11]**
Period	**:** 1756-1870 (1726-1913) Missing data for the period 1793-1796
Regions	**:** France
Series	**:** 1
Frequency	**:** Yearly
Observations	**:** 111 (184) per series
Price / Unit	**:** Livres et centièmes de livre / Hectolitre

Data Set 1 : Prix Moyens Interrégionaux Annuels de Froment - Grands Secteurs Territoriaux

Source	**:** See **Labrousse et al. [14, p. 23]**
Period	**:** 1756-1790
Regions	**:** Grands Secteurs Territoriaux - See Appendix 2
Series	**:** 9
Frequency	**:** Yearly
Observations	**:** 35 per series
Price / Unit	**:** Livres et centièmes de livre / Hectolitre

Data Set 2 : Prix Moyens Interrégionaux Annuels de Froment - Grands Secteurs Territoriaux

Source	**:** See **Labrousse et al. [14, pp. 23-24]**
Period	**:** 1797-1830
Regions	**:** Grands Secteurs Territoriaux - See Appendix 2
Series	**:** 9
Frequency	**:** Yearly
Observations	**:** 34 per series
Price / Unit	**:** Livres et centièmes de livre / Hectolitre

Data Set 3 : Prix Moyens Interrégionaux Annuels de Froment - Grands Secteurs Territoriaux

Source	**:** See **Labrousse et al. [14, pp. 27-35]**
Period	**:** 1831-1870
Regions	**:** Grands Secteurs Territoriaux - See Appendix 2
Series	**:** 9
Frequency	**:** Yearly
Observations	**:** 40 per series
Price / Unit	**:** Livres et centièmes de livre / Hectolitre

Data Set 4 : Les Prix du Blé en France dans 32 Généralités du Royaume et la Ville de Paris

Source	**:** See **Labrousse [13, pp. 106-113]**
Period	**:** 1756-1790
Regions	**:** Généralités du Royaume et Paris Ville - See Appendix 3
Series	**:** 33
Frequency	**:** Yearly
Observations	**:** 35 per series
Price / Unit	**:** Livres et centièmes de livre / Setier de Paris pesant 240 livres

Data Set 5 : Prix Moyens Annuels de l'Hectolitre de Froment par Département Intérieur

Source	**:** See **Labrousse et al. [14, pp. 45-219]**
Period	**:** 1806-1872 (IX-XIV and 1806-1872) Data for the period IX-XIV are excluded
Regions	**:** Départements Intérieurs - See Appendix 4
Series	**:** 85 - See Appendix 4
Frequency	**:** Yearly
Observations	**:** 67 per series
Price / Unit	**:** Livres et centièmes de livre / Hectolitre

Data Set 6.1 : Le Prix du Blé de Pays à Aix

Source	**:** See **Baehrel [11, p. 535]**
Period	**:** 1756-1789 (1570-1789)
Regions	**:** Aix-en-Provence
Series	**:** 1
Frequency	**:** Yearly
Observations	**:** 34
Price / Unit	**:** Livres et centièmes de livre / Livres-Tournois à charge (1 hl 6316)

Data Set 6.2 : Le Prix du Blé à Arles

Source	**:** See **Baehrel [11, p. 554]**
Period	**:** 1756-1789 (1570-1789)
Regions	**:** Arles
Series	**:** 1
Frequency	**:** Yearly
Observations	**:** 34 Missing observations for 1780 and 1787
Price / Unit	**:** Livres et centièmes de livre / Livres-Tournois au setier

Data Set 6.3 : Le Prix du Blé de Pays à Draguignan

Source	**:** See **Baehrel [11, p. 555]**
Period	**:** 1756-1790 (1616-1790)
Regions	**:** Draguignan
Series	**:** 1
Frequency	**:** Yearly
Observations	**:** 35
Price / Unit	**:** Livres et centièmes de livre / Livres-Tournois au setier until 1678 Livres et centièmes de livre / Livres-Tournois à la charge since 1681

Data Set 6.4 : Prix Annuels des Céréales au Marché de Toulouse

Source **:** See **Frêche & Frêche [12, pp. 85-91]**
Period **:** 1756-1790 (1486-1849)
Regions **:** Toulouse
Series **:** 1
Frequency **:** Yearly
Observations **:** 35
Price / Unit **:** Livres et centièmes de livre / Livres au setier (Mesure de Toulouse - 93,2 l)

Data Set 6.5 : Prix Annuel du Blé à Montauban

Source **:** See **Frêche & Frêche [12, pp. 101-103]**
Period **:** 1756-1789 (1691-1789)
Regions **:** Montauban
Series **:** 1
Frequency **:** Yearly
Observations **:** 34
Price / Unit **:** Livres et centièmes de livre / Mesure de Paris de 240 livres poids de marc

Data Set 6.6 : Le Prix du Froment sur le Marché de Poitiers

Source **:** See **Raveau [15, pp. 74-81]**
Period **:** 1756-1790 (1548-1860)
Regions **:** Poitiers
Series **:** 1
Frequency **:** Yearly
Observations **:** 35
Price / Unit **:** Livres et centièmes de livre / hectolitre

Appendix 2 : Grands Secteurs Territoriaux

#	Region
1.	North
2.	North-East
3.	East
4.	South-East
5.	South
6.	South-West
7.	West
8.	North West
9.	Center

Remark

In **Labrousse et al. [14, p. 21]** the following information about these regions is given :

'…*Grands Secteurs Teritoriaux…Secteurs constitués au XIX^{e} siècle par les services nationaux de la statistique agricole. On s'est efforcé de grouper les généralités de l'ancien régime dans le cadre de ces secteurs…*'.

Appendix 3 : Généralités du Royaume et Paris Ville

#	Généralité	#	Généralité
1.	Alençon	**17.**	La Rochelle
2.	Alsace	**18.**	Limoges
3.	Amiens	**19.**	Lorraine
4.	Auch	**20.**	Lyon et Dombes
5.	Bayonne	**21.**	Metz
6.	Bordeaux	**22.**	Montauban
7.	Bourges	**23.**	Moulins
8.	Bourgogne	**24.**	Orléans
9.	Bretagne	**25.**	Paris
10.	Caen	**26.**	Poitiers
11.	Champagne	**27.**	Provence
12.	Flandres	**28.**	Riom-Auvergne
13.	Franche-Comté	**29.**	Rouen
14.	Grenoble	**30.**	Roussillon
15.	Hainaut	**31.**	Soissons
16.	Languedoc	**32.**	Tours
		33.	Paris Ville

Appendix 4 : Départements Intérieurs

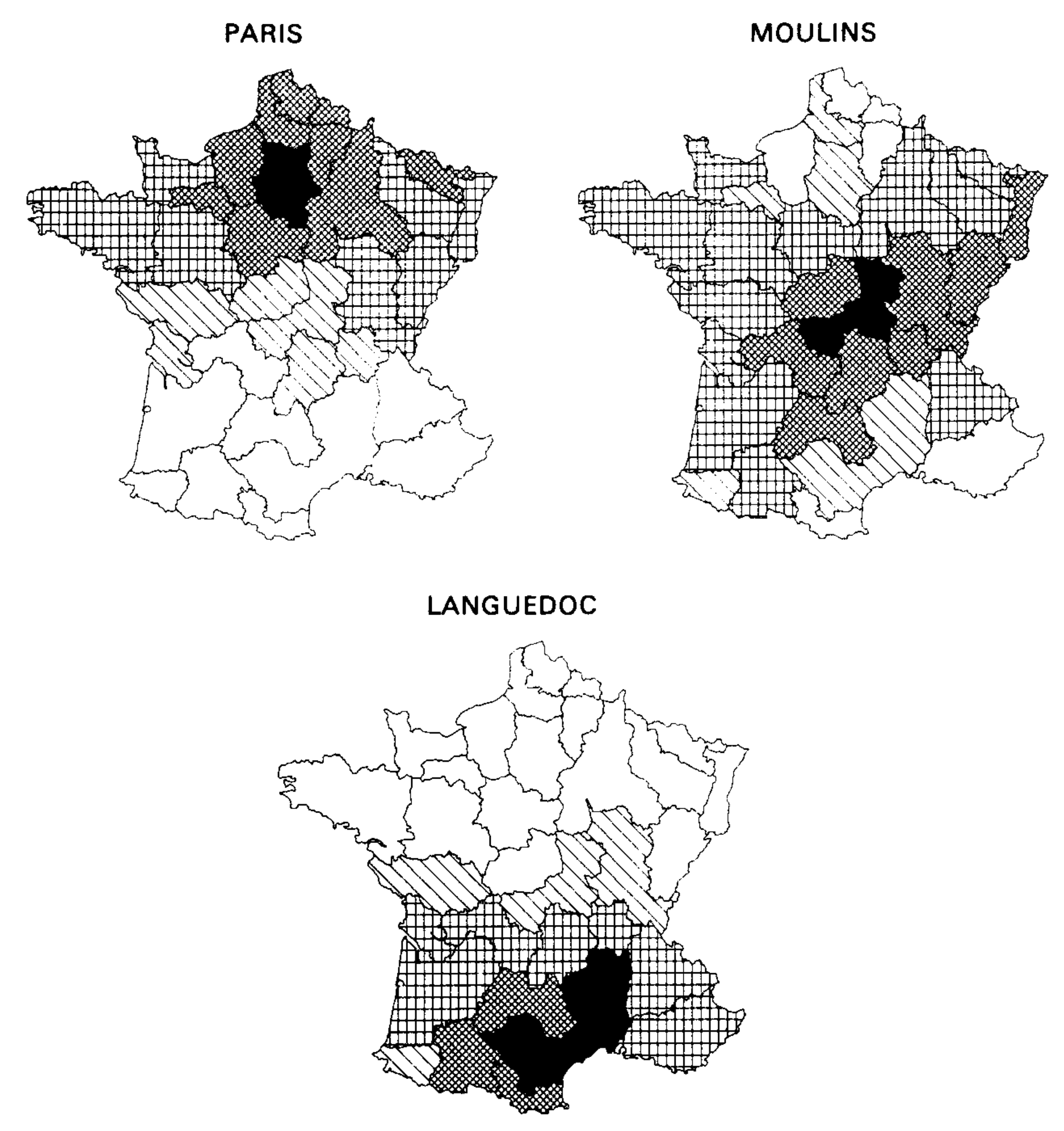
#	Département	#	Département	#	Département
1.	Ain	**31.**	Haute-Garonne	**61**	Orne
2.	Aisne	**32.**	Gers	**62.**	Pas-de-Calais
3.	Allier	**33.**	Gironde	**63.**	Puy-de-Dôme
4.	Basses-Alpes	**34.**	Hérault	**64.**	Basses-Pyrénées
5.	Hautes-Alpes	**35.**	Ille-et-Vilaine	**65.**	Hautes-Pyrénées
6.	Alpes-Maritimes (°)	**36.**	Indre	**66.**	Pyrénées-Orientales
7.	Ardèche	**37.**	Indre-et-Loire	**67.**	Bas-Rhin (°)
8.	Ardennes	**38.**	Isère	**68.**	Haut-Rhin (°)
9.	Ariège	**39.**	Jura	**69.**	Rhône
10.	Aube	**40.**	Landes	**70.**	Haute-Saône
11.	Aude	**41.**	Loire-et-Cher	**71.**	Saône-et-Loire
12.	Aveyron	**42.**	Loire	**72.**	Sarthe
13.	Bouches-du-Rhône	**43.**	Haute-Loire	**73.**	Savoie (°)
14.	Calvados	**44.**	Loire-Inférieure	**74.**	Haute-Savoie (°)
15.	Cantal	**45.**	Loiret	**75.**	Seine (°)
16.	Charente	**46.**	Lot	**76.**	Seine-Inférieure
17.	Charente-Inférieure	**47.**	Lot-et-Garonne	**77.**	Seine-et-Marne
18.	Cher	**48.**	Lozère	**78.**	Seine-et-Oise
19.	Corrèze	**49.**	Maine-et-Loire	**79.**	Deux-Sèvres
2A.	Corse-du-Sud (°)	**50.**	Manche	**80.**	Somme
2B.	Haute-Corse (°)	**51.**	Marne	**81.**	Tarn
21.	Côte-d'Or	**52.**	Haute-Marne	**82.**	Tarn-et-Garonne
22.	Côtes-du-Nord	**53.**	Mayenne	**83.**	Var
23.	Creuse	**54.**	Meurthe	**84.**	Vaucluse
24.	Dordogne	**55.**	Meuse	**85.**	Vendée
25.	Doubs	**56.**	Morbihan	**86.**	Vienne
26.	Drôme	**57.**	Moselle (°)	**87.**	Haute-Vienne
27.	Eure	**58.**	Nièvre	**88.**	Vosges
28.	Eure-et-Loire	**59.**	Nord	**89.**	Yonne
29.	Finistère	**60.**	Oise		
30.	Gard				

(°) **Remarks**

#	Département	
6.	Alpes-Maritimes	Excluded - Missing data for the period 1814-1860
2A	Corse-du-Sud	Excluded - Département Extérieur
2B	Haute-Corse	Excluded - Département Extérieur
57.	Moselle	Included - Missing data for the years 1871 and 1872
67.	Bas-Rhin	Included - Missing data for the years 1871 and 1872
68.	Haut-Rhin	Included - Missing data for the years 1871 and 1872
73.	Savoie	Excluded - Data available as from 1861
74.	Haute-Savoie	Excluded - Data available as from 1861
75.	Seine	Included - Missing data for the years 1871 and 1872

Appendix 5 : Graphical Results by Weir

Price Correlations with Three Regions - Reference Regions : Paris, Moulins and Languedoc (°)



(°) Remarks

- See **Weir [9, p. 210]**
- Data used are those from **Labrousse [13, pp. 106-113]**.
- Shading indicates the statistical significance of the correlation of prices (deviations from moving averages) between the reference region and the shaded region. The solid shade is the reference *généralité* to which the correlation refer. Cross-hatched regions are correlated at $p < 0.001$. Box-striped regions are correlated at $p < 0.01$ and diagonal regions are correlated at $p < 0.1$.
- After detrending, the years in observation are 1761-1784.

References

General

[1] Borghers, E.
French Regional Wheat Prices : 1756-1870 - An Illustrated Statistical Analysis
Research Paper 2006-003
Department of Mathematics, Statistics and Actuarial Sciences
Faculty of Applied Economics, University of Antwerp, Antwerp, February 2006.

[2] Cox, T.F. & Cox, M.A.A.
Multidimensional Scaling
Chapman & Hall, London, 1994.

[3] Everitt, B.S.
Cluster Analysis
Edward Arnold, London, 3rd Ed., 1995.

[4] Everitt, B.S. , Landau, S. & Leese, M.
Cluster Analysis
Edward Arnold, London, 4th Ed., 2001.

[5] Hakstian, A.R. , Rogers, W.D. & Cattell, R.B.
The Behavior of Numbers of Factor Rules with Simulated Data
Multivariate Behavioral Research, vol. 17, 1982, pp. 193-219.

[6] Jackson, J.E.
A User's Guide to Principal Components
John Wiley & Sons Inc., New York, 1991.

[7] Jolliffe, I.T.
Principal Component Analysis
Springer, New York, 2nd Ed., 2002.

[8] Ward, J.H.
Hierarchical Grouping to Optimize an Objective Function
Journal of the American Statistical Association, vol. 58, March, 1963, pp. 236-244.

[9] Weir, D.R.
Markets and Mortality in France, 1600-1789
In : Walter, J. & Schofield, R. (Eds.)
Famine, Disease and the Social Order in Early Modern Society
Cambridge University Press, Cambridge, 1989, pp. 201-234.

[10] Zwick, W.R. & Velicer, W.F.
Comparison of Five Rules for Determining the Number of Components to Retain
Psychological Bulletin, vol. 99, 1986, pp. 432-442.

Price Data

[11] Baehrel, R.
Une Croissance : La Basse-Provence Rurale - Fin du XVIe Siècle - 1789
Essai d’Economie Historique Statistique
Démographie et Sociétés VI
Ecole Pratique des Hautes Etudes - VIe Section
Centre de Recherches Historique
SEVPEN., Paris, 1961, p. 535, p. 554 & p. 555.

［12］**Frêche, G & Frêche, G.**
Les Prix des Grains, des Vins et des Légumes à Toulouse, 1486-1868
Tavaux et Recherches de la Faculté de Droit et des Sciences Economiques de Paris
Séries Sciences Historique, N° 10
Presse Universitaires de France, Paris, 1967, pp.85-91 & pp.101-103.

［13］**Labrousse, C.-E.**
Esquisse du Mouvement des Prix et des Revenus en France au XVIIIe Siècle
Collection Scientifique d'Economie Politique III
Tome 1 : Les Prix
Tome 2 : Les Prix (Fin), Les Revenus
Dalloz, Paris, 1933, pp. 106-113.

［14］**Labrousse, C.-E., Romano, R. & Dreyfus, F.-G.**
Le Prix du Froment en France
Au temps de la Monaie Stable, 1726-1913
Ecole Pratique des Hautes Etudes - VIe Section
Centre de Recherches Historiques
Monnaie - Prix - Conjoncture IX
SEVPEN, Paris, 1970, pp. 9-11, pp. 23-24 & pp. 27-35.

［15］**Raveau, P.**
Essai sur la Situation Economique et l'Etat Social en Poitou au XVIe Siècle
Bibliothèque d'Histoire Economique
Librairie des Sciences Politique et Sociales, Marcel Rivière, Paris, 1931, pp. 74-81.

Software

［16］**Statistica**
Kernel Release 5.5, Edition '99
Statsoft, Tulsa, 1999.