

## UNIWIN VERSION 10.4.0

# ANALYSE EN COMPOSANTES PRINCIPALES FONCTIONNELLE

Révision : 15/09/2025

Définition.....	1
Entrée des données .....	2
Données manquantes .....	3
Exemple 1 : Fichier Berkeley .....	3
L'option Rapports .....	6
L'option Graphiques .....	9
Exemple 2 : Fichier CanadaTempératureMois .....	16
Exemple 3 : Fichier CanadaPrécipitationMois .....	21
Exemple 3 : Fichier Broyage (plan d'expériences) .....	25
ACPF versus ACP .....	35
Enregistrement des résultats .....	36
Références .....	36

### Définition

L'analyse en composantes principales fonctionnelle (ACPF) est une extension de l'analyse en composantes principales classique adaptée aux données fonctionnelles. Elle permet d'analyser des données qui représentent des fonctions ou des courbes plutôt que de simples vecteurs de valeurs discrètes.

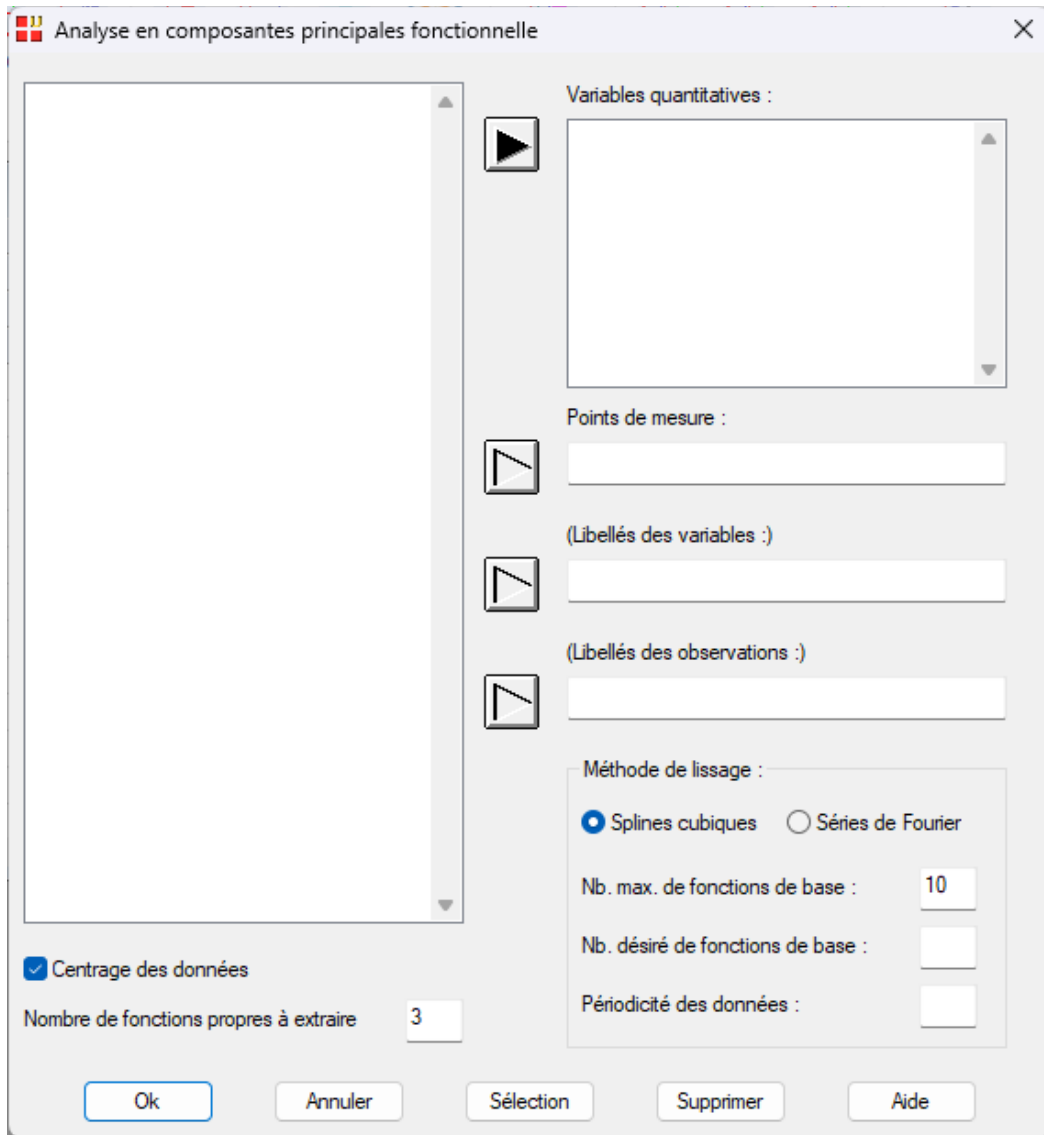
L'ACPF cherche à décomposer des fonctions aléatoires en une somme de fonctions déterministes (les composantes principales fonctionnelles) multipliées par des coefficients aléatoires.

L'ACPF est particulièrement utile pour l'analyse de données temporelles, spatiales ou spectrales où l'on souhaite capturer les variations continues dans les données tout en réduisant leur dimensionnalité.

Cette procédure est basée sur le package R 'fda'.

## Entrée des données

Cliquons sur l'icône ACPF dans le ruban Décrire pour afficher la boîte de dialogue montrée ci-dessous :



Cette boîte de dialogue permet de définir les variables quantitatives à utiliser ainsi que les points de mesure.

Elle permet également, en option, d'indiquer les noms des variables contenant les libellés des variables quantitatives et les libellés des observations.

Par défaut, les données sont centrées et le nombre de fonctions propres (harmoniques, composantes) à extraire est égal à 3.

Le choix de la méthode de d'interpolation et de lissage est proposé : base de splines cubiques (B-splines) pour des données non périodiques ou base de séries de Fourier pour des données périodiques.

Le nombre maximum de fonctions de base doit être précisé. Il est possible d'indiquer le nombre de fonctions de base désiré.

Dans le cas de séries de Fourier, la périodicité des données peut être indiquée. Si elle ne l'est pas, elle est définie comme égale à  $1 + \text{l'étendue des points de mesure}$ .

## **Données manquantes**

Les données manquantes ne sont pas autorisées.

Les observations ayant des données manquantes sont automatiquement éliminées des calculs.

## **Exemple 1 : Fichier Berkeley**

Nous utiliserons le fichier Berkeley pour illustrer ce premier exemple.

Dans cette étude de croissance (Physical growth of California boys and girls from birth to eighteen years, Tuddenham et Snyder, 1954), les tailles de 54 filles et 39 garçons âgés de 1 à 18 ans ont été enregistrées.

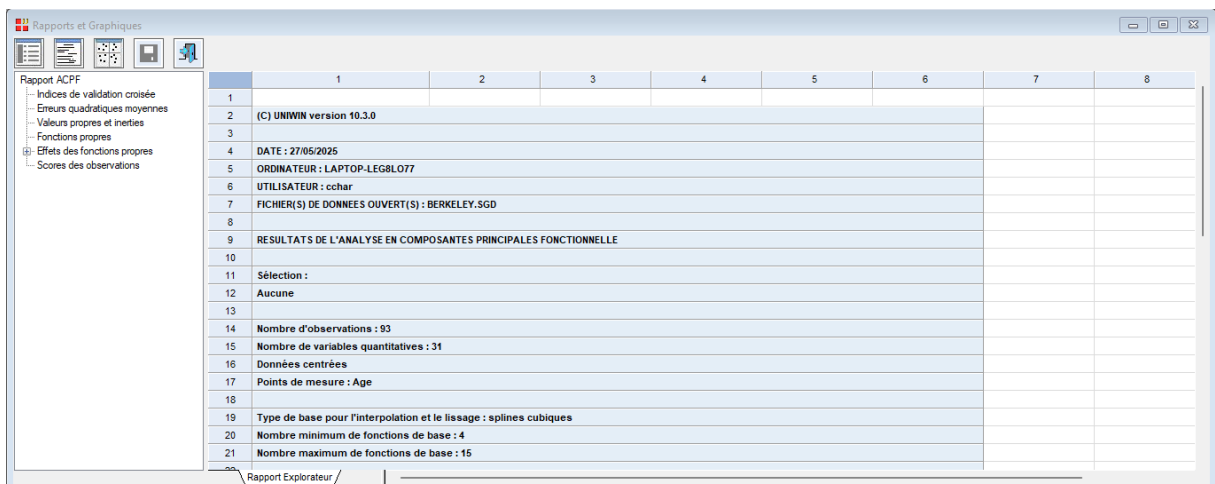
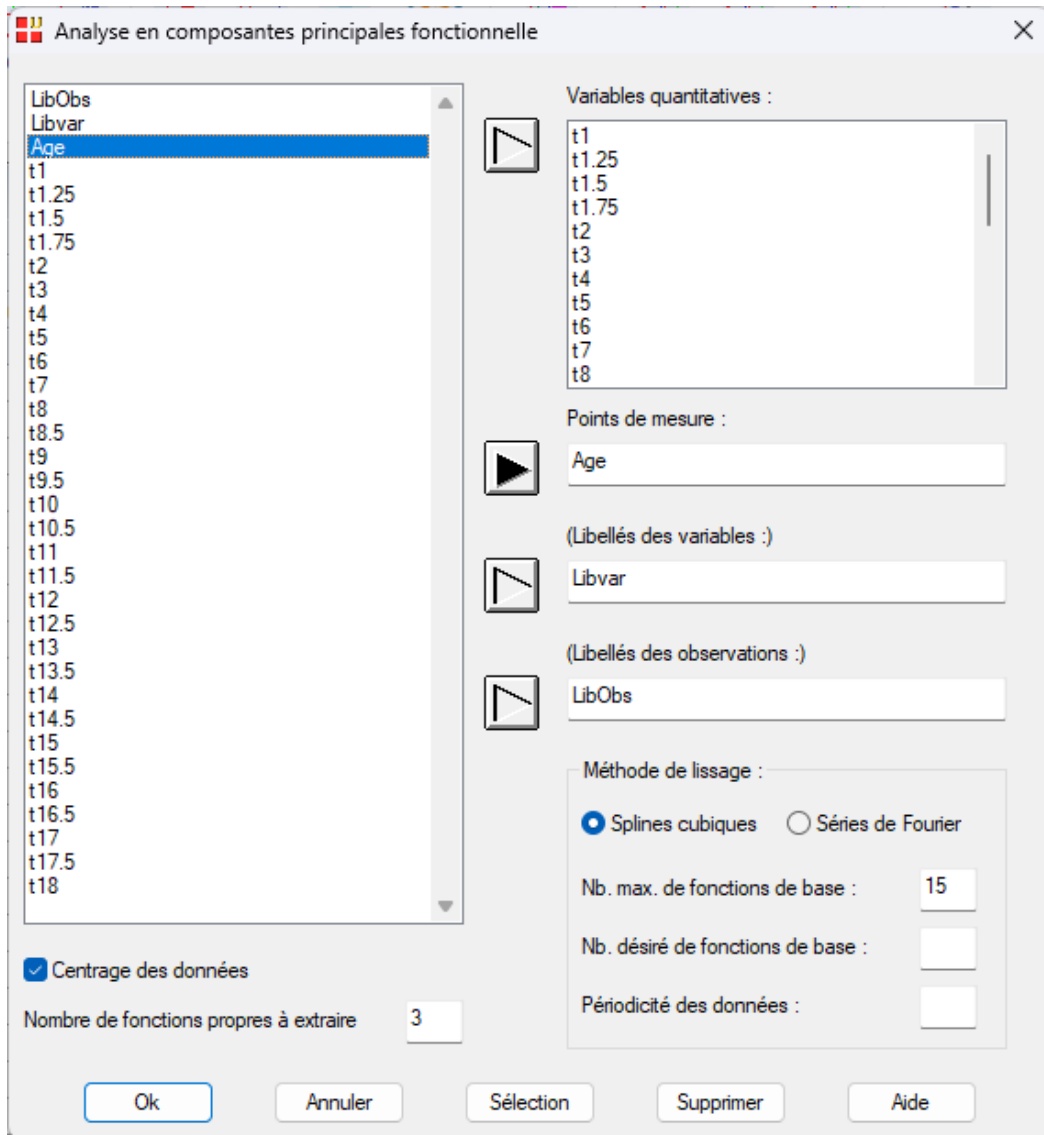
Ces tailles ont été mesurées à 31 instants pour chacun des 93 enfants.


Cliquons sur l'icône ACPF dans le ruban Décrire et renseignons la boîte de dialogue comme montré ci-après.


Les 31 variables quantitatives à utiliser sont  $t1$  à  $t18$ , les points de mesure sont *Age* (31 valeurs), les données sont centrées, une base d'au maximum 15 splines cubiques est demandée et 3 fonctions propres seront extraites.

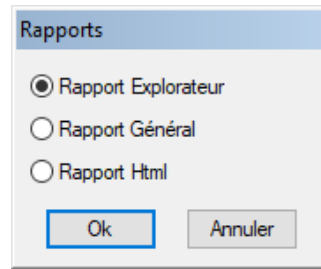
Les libellés des variables sont *Libvar* et ceux des observations *LibObs*.


Après quelques instants, l'écran montré ci-après s'affiche.

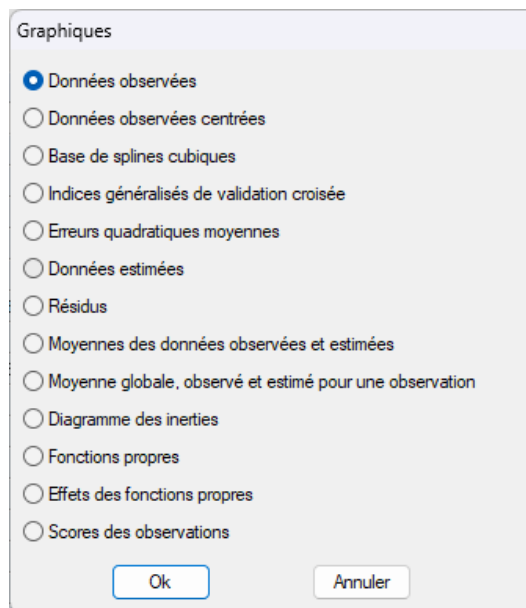



La barre d'outils 'Rapports et Graphiques' permet par l'icône 'Données'  de rappeler la boîte de dialogue d'entrée des données.

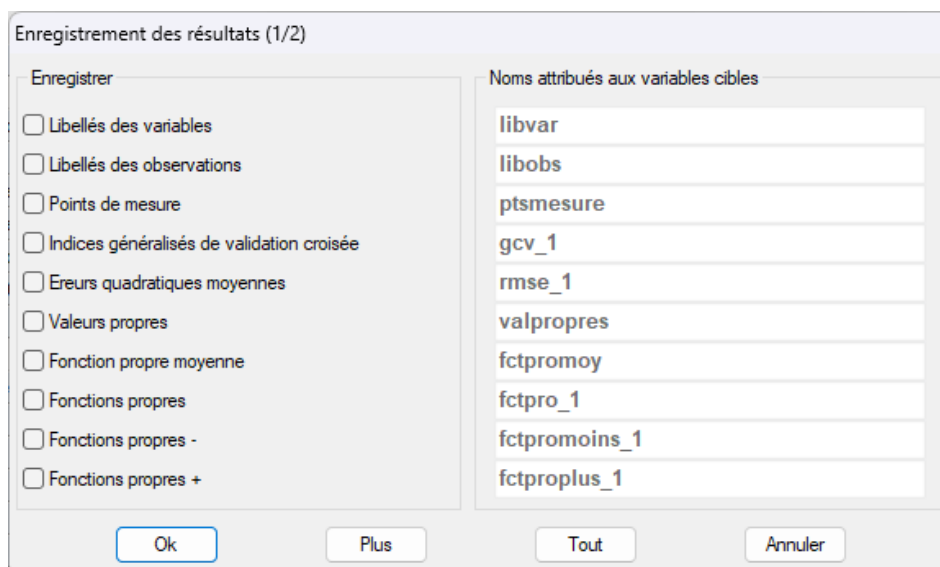
L'icône 'Rapports'  affiche la boîte de dialogue des options pour les rapports :



et l'icône 'Graphiques'  affiche la boîte de dialogue des options pour les graphiques.



L'icône 'Enregistrer'  permet de sélectionner les résultats de l'analyse à enregistrer dans un fichier.



## L'option Rapports

Cette option permet d'obtenir le rapport à l'écran sous la forme d'un explorateur, d'un tableau ou au format HTML.

Le premier tableau affiche un résumé des paramètres de l'analyse : nombre d'observations (93), nombre de variables (31), centrage des données, type de base pour l'interpolation et le lissage, nombres minimum (4) et maximum (15) de fonctions de base, nombre de fonctions propres à extraire (3).

	1	2	3	4	5	6	7	8
7	FICHIER(S) DE DONNEES OUVERT(S) : BERKELEY.SGD							
8								
9	RESULTATS DE L'ANALYSE EN COMPOSANTES PRINCIPALES FONCTIONNELLE							
10								
11	Sélection :							
12	Aucune							
13								
14	Nombre d'observations : 93							
15	Nombre de variables quantitatives : 31							
16	Données centrées							
17	Points de mesure : Age							
18								
19	Type de base pour l'interpolation et le lissage : splines cubiques							
20	Nombre minimum de fonctions de base : 4							
21	Nombre maximum de fonctions de base : 15							
22								
23	Nombre demandé de fonctions propres à extraire : 3							
24								
25								
26								
27								


Le deuxième tableau affiche l'évolution de l'indice généralisé de validation croisée (GCV) en fonction du nombre de fonctions de base.

Cet indice permet de déterminer la taille optimale de la base de splines cubiques (celle donnant la meilleure interpolation et le meilleur lissage des données).

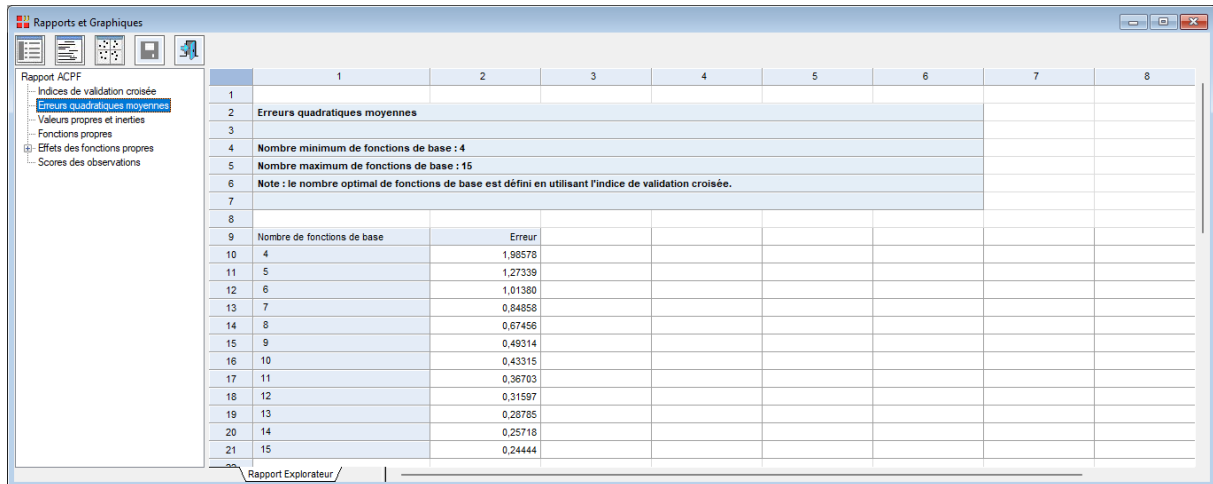
Elle est ici de 14 pour un indice GCV = 0,219930.

Cette taille optimale de la base sera automatiquement utilisée pour la suite des calculs.

	1	2	3	4	5	6	7	8
1								
2	Indices généralisés de validation croisée (GCV)							
3								
4	Nombre minimum de fonctions de base : 4							
5	Nombre maximum de fonctions de base : 15							
6	Nombre optimal de fonctions de base pour l'indice GCV : 14							
7	GCV optimal = 0,219930							
8								
9								
10	Nombre de fonctions de base		Indice					
11	4	5,19828						
12	5	2,30516						
13	6	1,58033						
14	7	1,20139						
15	8	0,82662						
16	9	0,48285						
17	10	0,40884						
18	11	0,32365						
19	12	0,26576						
20	13	0,24577						
21	14	0,21993						
22								

Si une taille différente est souhaitée, il faut rappeler la boîte de dialogue d'entrée des données via l'icône  et préciser cette taille dans le champ 'Nombre désiré de fonctions de base'.

Le troisième tableau affiche l'évolution de l'erreur quadratique moyenne en fonction du nombre de fonctions de base.



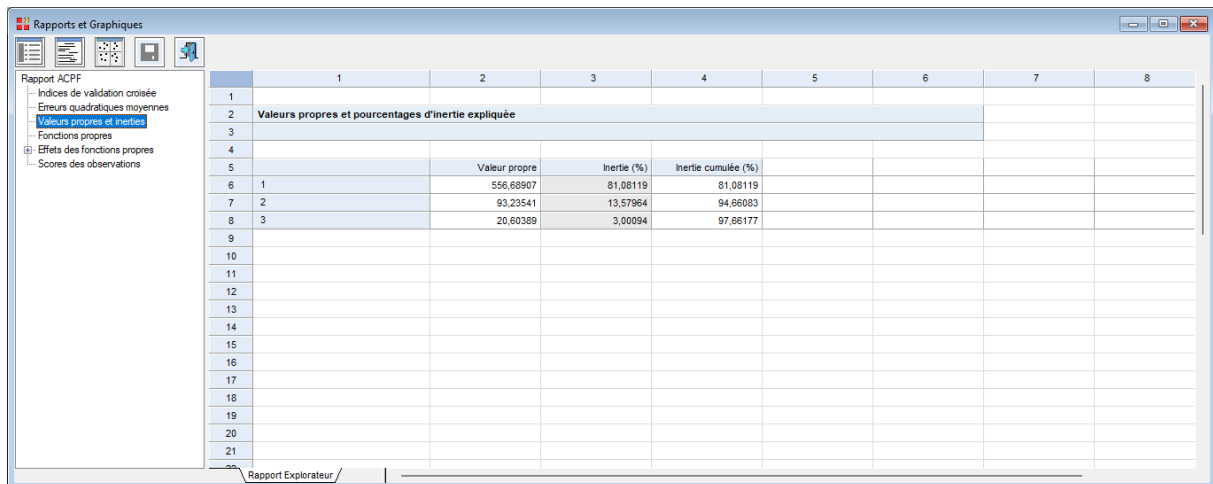
The screenshot shows a software window titled 'Rapports et Graphiques'. On the left is a navigation tree with 'Erreurs quadratiques moyennes' selected. The main area displays a table with the following data:

	1	2	3	4	5	6	7	8
1								
2	<b>Erreurs quadratiques moyennes</b>							
3								
4	<b>Nombre minimum de fonctions de base : 4</b>							
5	<b>Nombre maximum de fonctions de base : 15</b>							
6	<b>Note : le nombre optimal de fonctions de base est défini en utilisant l'indice de validation croisée.</b>							
7								
8								
9	Nombre de fonctions de base	Erreur						
10	4	1,98578						
11	5	1,27339						
12	6	1,01380						
13	7	0,84858						
14	8	0,67456						
15	9	0,49314						
16	10	0,43315						
17	11	0,36703						
18	12	0,31597						
19	13	0,28785						
20	14	0,25718						
21	15	0,24444						

Note : le nombre optimal de fonctions de base est défini en utilisant l'indice de validation croisée.

Le quatrième tableau affiche les valeurs propres et les pourcentages d'inertie expliquée par les 3 fonctions propres demandées.

Près de 95% de l'inertie totale est expliquée par les 2 premières fonctions propres.



The screenshot shows the same software window with 'Valeurs propres et pourcentages d'inertie expliquée' selected in the navigation tree. The main area displays a table with the following data:

	1	2	3	4	5	6	7	8
1								
2	<b>Valeurs propres et pourcentages d'inertie expliquée</b>							
3								
4								
5		Valeur propre	Inertie (%)	Inertie cumulée (%)				
6	1	556,68907	81,08119	81,08119				
7	2	93,23541	13,57964	94,66083				
8	3	20,60389	3,00094	97,66177				
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								

Le cinquième tableau affiche la fonction propre moyenne et chacune des 3 fonctions propres extraites (également appelées harmoniques) évaluées aux 31 points de mesure (Age).

	1	2	3	4	5	6	7	8
1								
2	<b>Fonctions propres (harmoniques)</b>							
3								
4								
5	Points de mesure	Fonction moyenne	Fonction 1	Fonction 2	Fonction 3			
6	1,000000	74,73793	0,08410	-0,01389	0,17332			
7	1,250000	78,71644	0,08982	0,00543	0,23055			
8	1,500000	82,16385	0,09510	0,02090	0,25838			
9	1,750000	85,16478	0,10009	0,03293	0,28471			
10	2,000000	87,80385	0,10491	0,04190	0,25743			
11	3,000000	96,36364	0,12502	0,05558	0,23931			
12	4,000000	103,74480	0,14350	0,06927	0,29769			
13	5,000000	110,84239	0,15268	0,12428	0,29134			
14	6,000000	117,59055	0,16489	0,16164	0,24671			
15	7,000000	123,86683	0,18588	0,15281	0,21183			
16	8,000000	129,84731	0,20414	0,15254	0,17941			
17	8,500000	132,75745	0,21216	0,15860	0,16372			
18	9,000000	135,62266	0,22050	0,16586	0,14892			
19	9,500000	138,47725	0,22895	0,17784	0,13106			
20	10,000000	141,38001	0,23840	0,20316	0,10246			
21	10,500000	144,38880	0,24177	0,25001	0,05540			

Le sixième tableau affiche les effets des fonctions propres sur la fonction propre moyenne.

Effet positif = fonction moyenne + 2 x  $\sqrt{\text{valeur propre}} \times \text{harmonique}$

Effet négatif = fonction moyenne - 2 x  $\sqrt{\text{valeur propre}} \times \text{harmonique}$

Sous l'hypothèse de normalité, cette plage définit une région qui contient approximativement 95% des données.

	1	2	3	4	5	6	7	8
1								
2	<b>Effets de la fonction propre n° 1 sur la fonction propre moyenne</b>							
3								
4	<b>Effets = fonction moyenne +/- 2 x racine carrée(valeur propre) x harmonique</b>							
5	<b>(sous hypothèse de normalité, région contenant approximativement 95% des données)</b>							
6								
7								
8	Points de mesure	Moyenne	Effet négatif	Effet positif				
9	1,000000	74,73793	70,76943	78,70643				
10	1,250000	78,71644	74,47792	82,95496				
11	1,500000	82,16385	77,67807	86,65164				
12	1,750000	85,16478	80,44192	89,88765				
13	2,000000	87,80385	82,85353	92,75416				
14	3,000000	96,36364	90,46412	102,26316				
15	4,000000	103,74480	96,97305	110,51656				
16	5,000000	110,84239	103,63784	118,04893				
17	6,000000	117,59055	109,80977	125,37133				
18	7,000000	123,86683	115,09550	132,63816				
19	8,000000	129,84731	120,21414	139,48048				
20	8,500000	132,75745	122,74612	142,76877				
21	9,000000	135,62266	125,21782	146,02750				

Le septième tableau affiche les scores des 93 observations sur les 3 composantes extraites.

	1	2	3	4	5	6	7	8
1								
2	Scores des observations							
3								
4								
5		Composante 1	Composante 2	Composante 3				
6	garçon01	52,91464	-12,73941	-6,17119				
7	garçon02	-5,96365	-7,32045	12,47372				
8	garçon03	-18,83691	-13,21234	0,86124				
9	garçon04	10,07882	-14,21759	1,24711				
10	garçon05	-18,02085	-16,10865	-6,68527				
11	garçon06	-8,89128	-9,71299	2,42183				
12	garçon07	-15,28779	-7,08121	6,90707				
13	garçon08	25,64638	-10,44873	-4,32763				
14	garçon09	-18,90611	-10,73514	9,44679				
15	garçon10	46,25064	-3,14971	-1,95030				
16	garçon11	-4,18844	-10,45029	8,21557				
17	garçon12	30,29390	-13,42646	-0,29307				
18	garçon13	10,93390	-12,94759	1,12032				
19	garçon14	30,95570	-7,73179	-2,28169				
20	garçon15	23,28189	-0,63737	2,65599				
21	garçon16	27,05462	-6,47879	-1,38250				

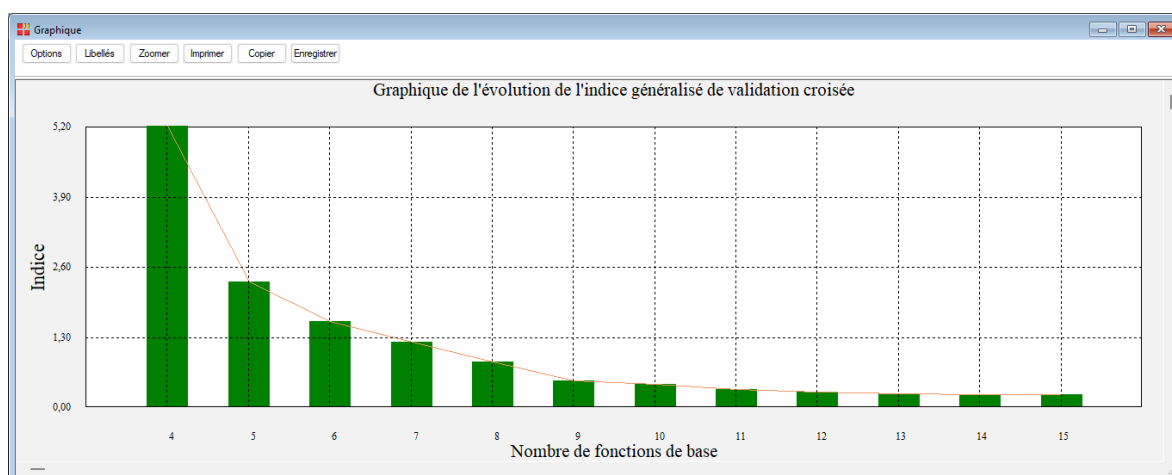
## L'option Graphiques

Cette option permet d'obtenir divers graphiques pour l'analyse ACPF.

- Graphique des indices généralisés de validation croisée

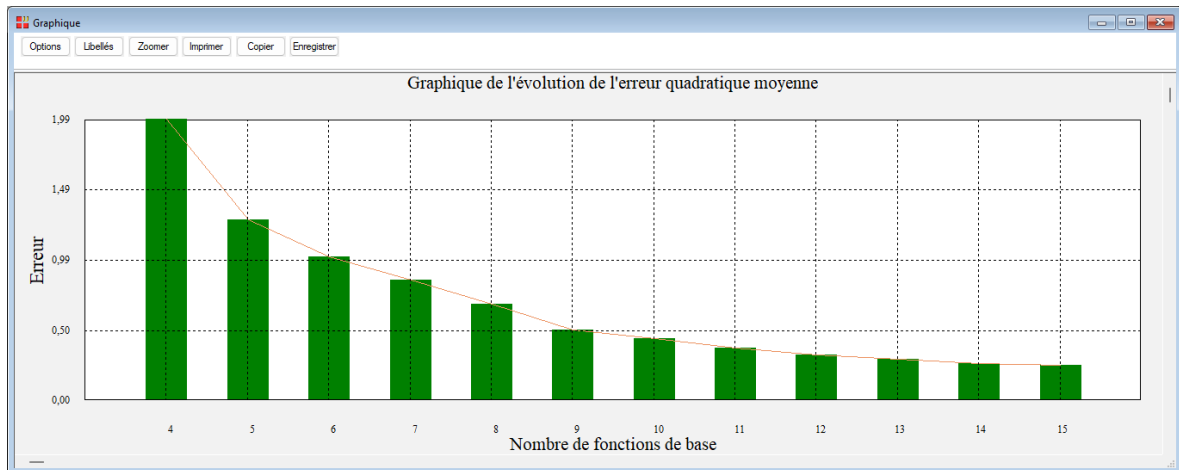
Ce graphique affiche l'évolution de l'indice généralisé de validation croisée en fonction du nombre de fonctions de base.

Il permet de visualiser le nombre optimal de fonctions de base (ici 14) et d'éventuellement de choisir un nombre différent si cela est souhaité. Pour la prise en compte de ce nombre différent du nombre optimal, revenir à la boîte de dialogue d'entrée des données et entrer ce nombre dans le champ 'Nombre désiré de fonctions de base'.



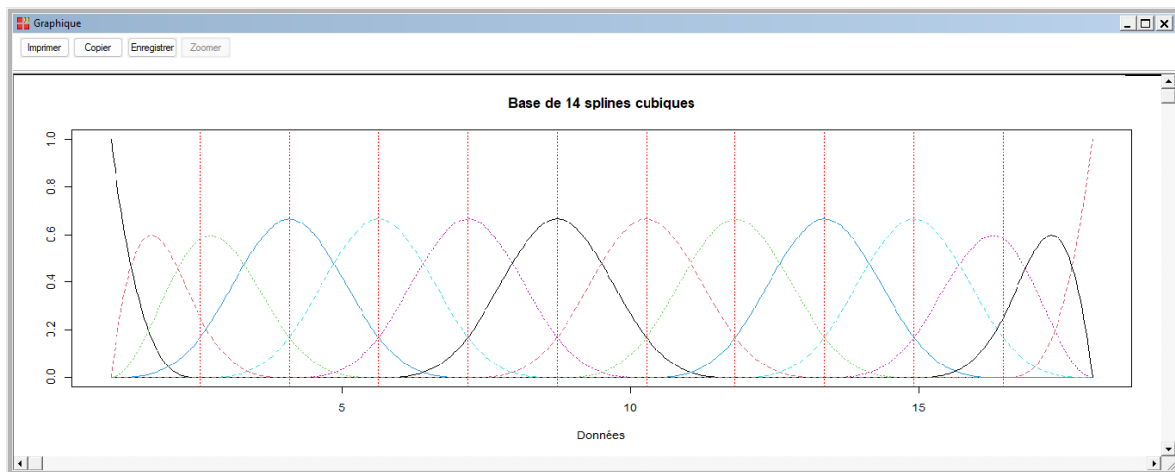
- Graphique des erreurs quadratiques moyennes

Ce graphique affiche l'évolution de l'erreur quadratique moyenne en fonction du nombre de fonctions de base.



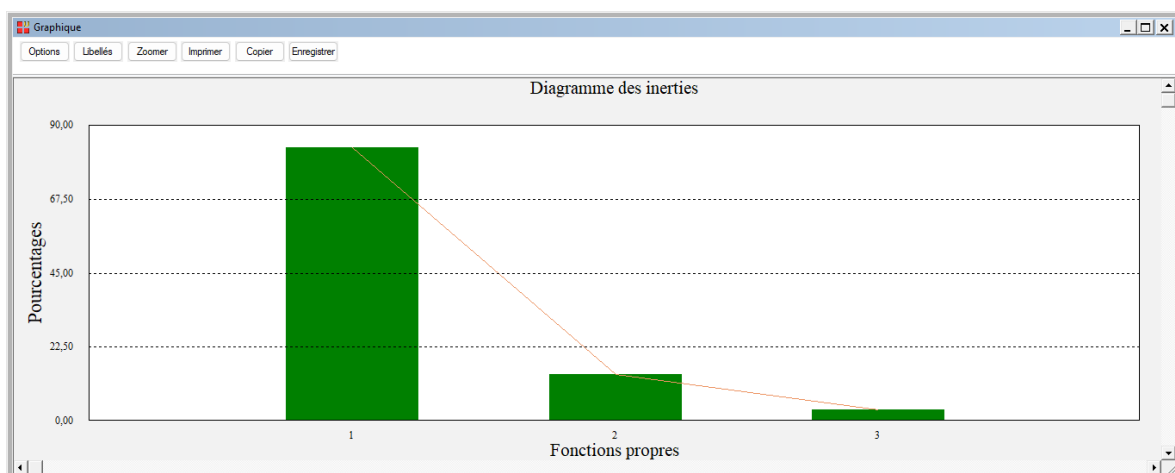
- Graphique de la base de splines cubiques

Ce graphique affiche les 14 splines cubiques utilisés pour l'interpolation et le lissage des données.



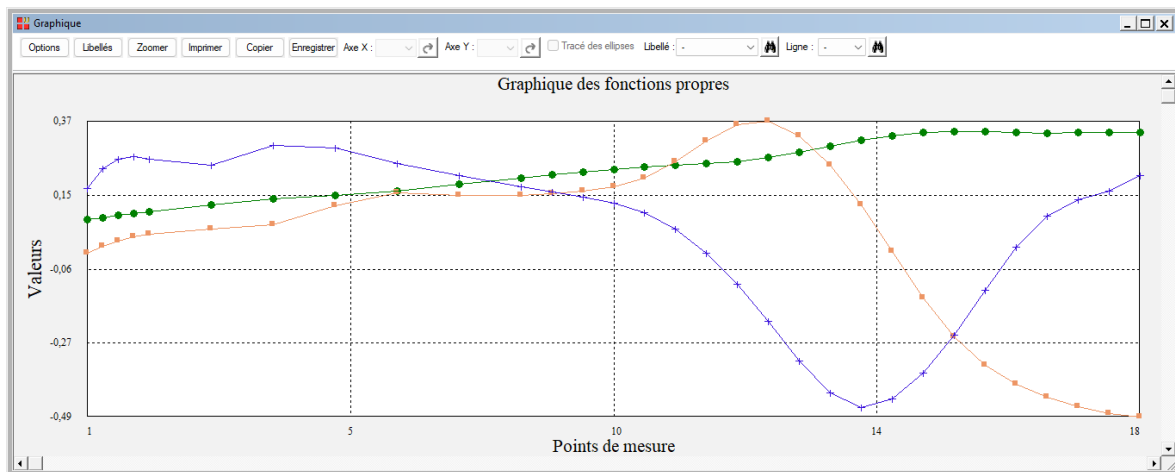
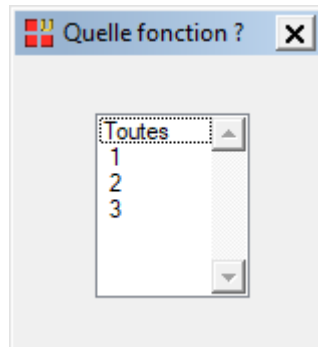
- Diagramme des inerties

Ce graphique affiche les pourcentages d'inertie expliqués par chacune des 3 fonctions propres extraites.



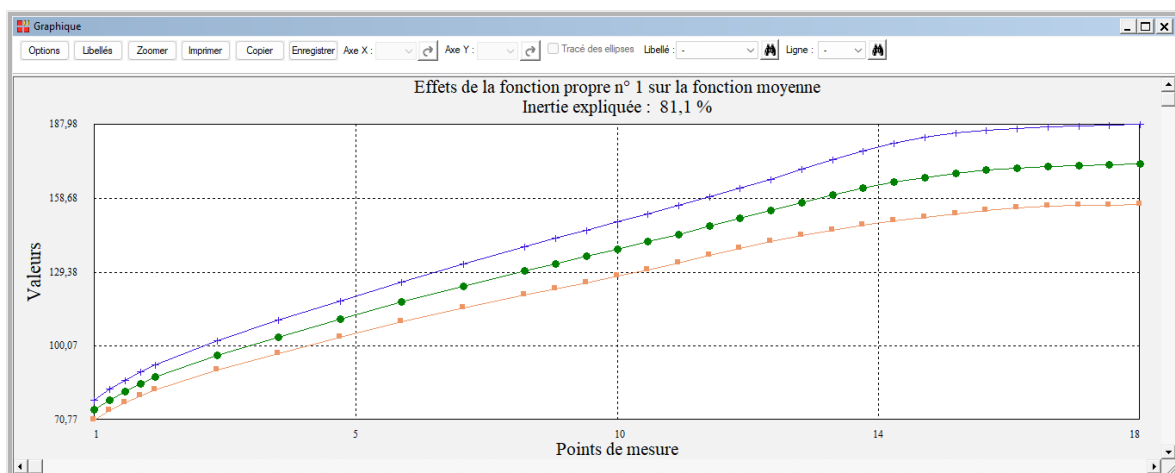
- Graphique des fonctions propres

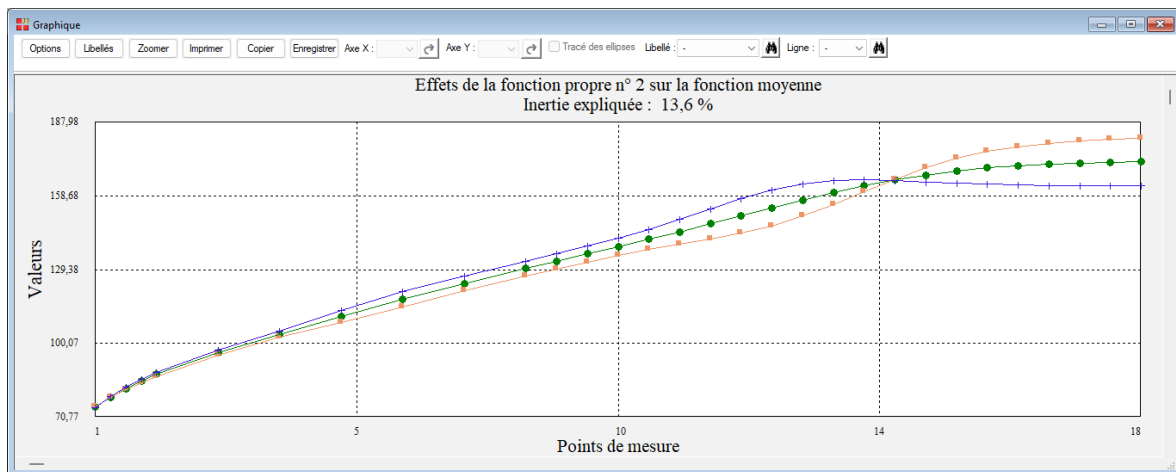
Après avoir choisi l'affichage de toutes les fonctions propres ou d'une unique fonction propre, ce graphique affiche la ou les fonctions propres par rapport aux points de mesure.



- Graphique des effets des fonctions propres

Après le choix de la fonction propre, ce graphique affiche les effets positif et négatif de chaque fonction propre sur la fonction propre moyenne.





Les effets sont définis par :

Effet positif = fonction moyenne + 2 x  $\sqrt{\text{valeur propre}} \times \text{harmonique}$

Effet négatif = fonction moyenne - 2 x  $\sqrt{\text{valeur propre}} \times \text{harmonique}$

Les fonctions propres sont ainsi considérées comme des perturbations de la fonction propre moyenne.

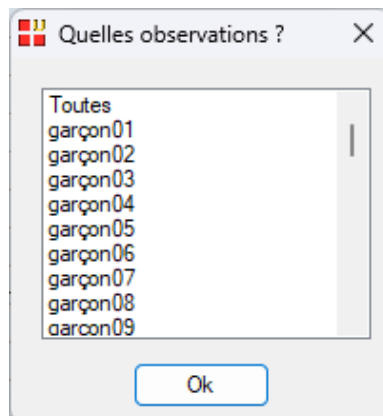
Sous l'hypothèse de normalité, cette plage définit une région qui contient approximativement 95% des données.

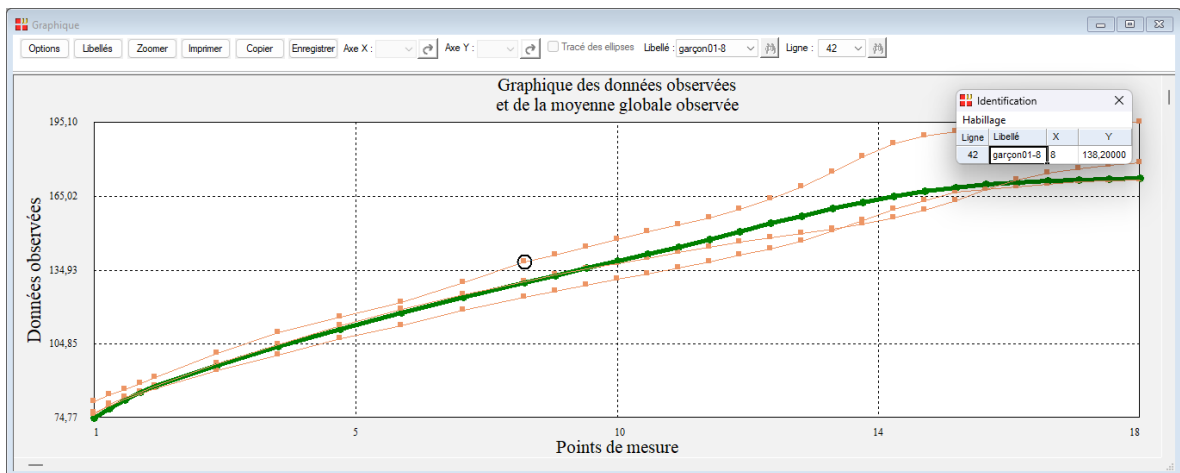
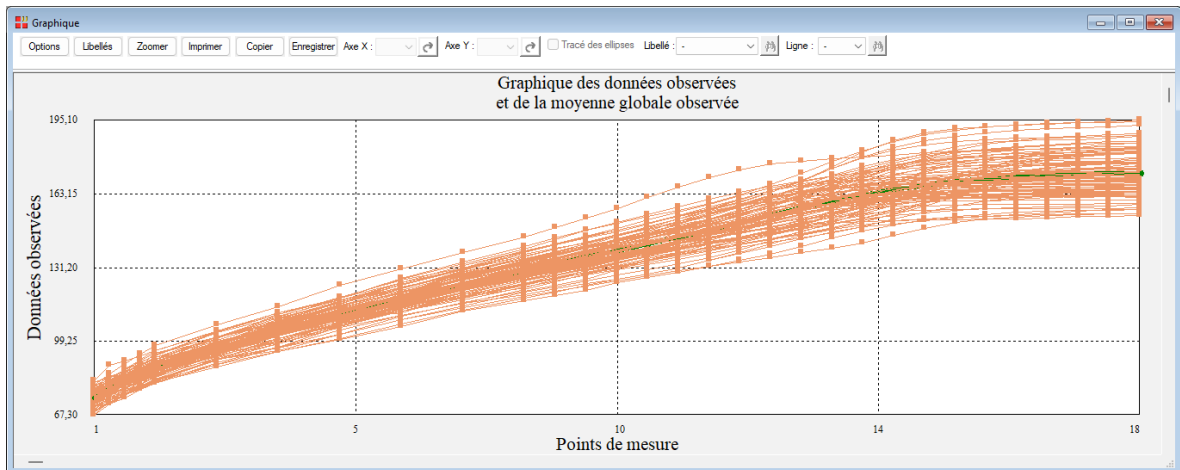
La première fonction propre capture les écarts globaux (positifs ou négatifs) par rapport à la moyenne.

La seconde révèle des schémas associés aux différences de vitesse de croissance. En particulier, elle présente un maximum suivi d'un changement de signe aux alentours de la puberté.

- Graphique des données observées

Ce graphique affiche les données observées ainsi que la moyenne globale observée par rapport aux points de mesure. Il est possible de choisir les observations à afficher :

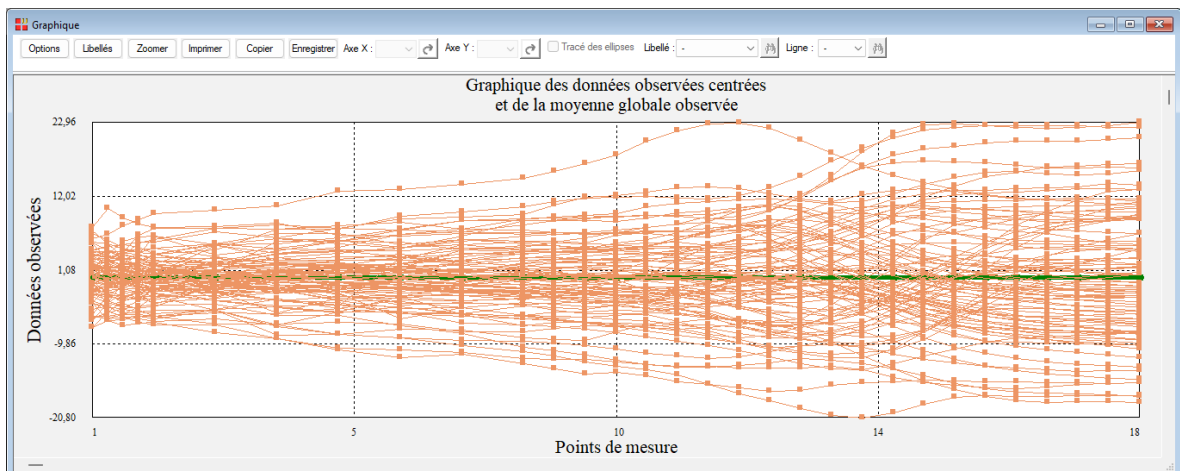




Cliquer sur un point dans le graphique permet de l'identifier.

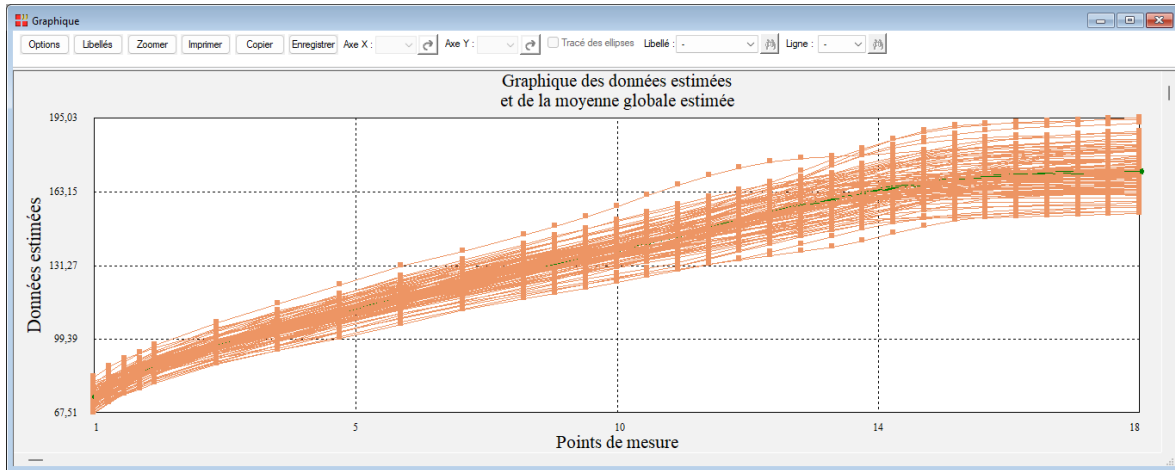
- Graphique des données observées centrées

Si le centrage des données a été demandé, ce graphique affiche les écarts à la moyenne des données observées par rapport aux points de mesure. Comme précédemment, il est possible de choisir les observations à afficher.



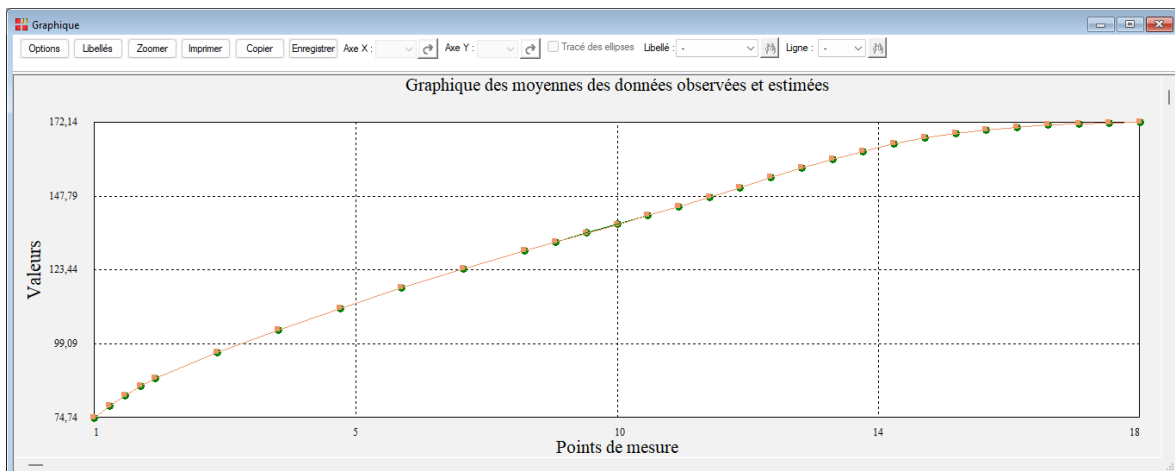
- Graphique des données estimées

Ce graphique affiche les données estimées ainsi que la moyenne globale estimée par rapport aux points de mesure. Comme précédemment, il est possible de choisir les observations à afficher.



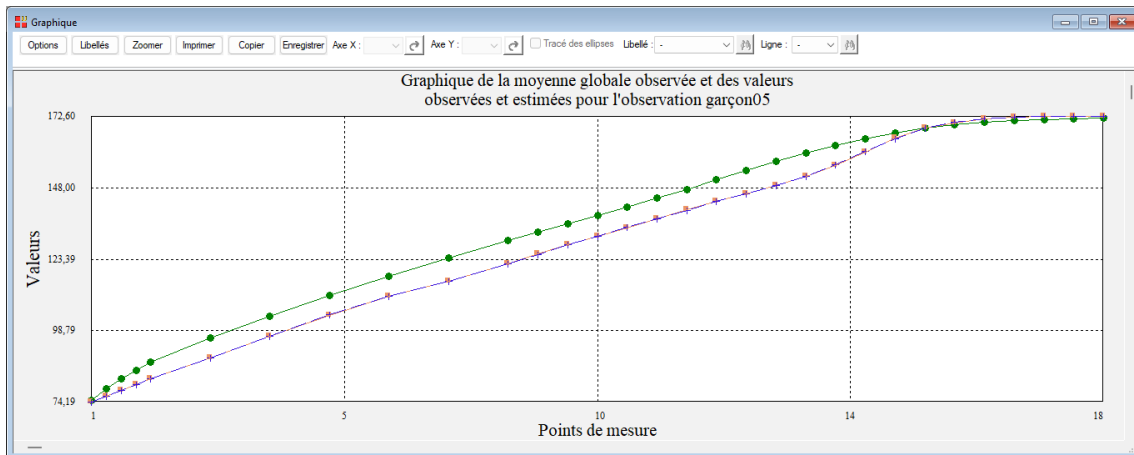
- Graphique des moyennes des données observées et estimées

Ce graphique affiche les moyennes des données observées et estimées par rapport aux points de mesure.



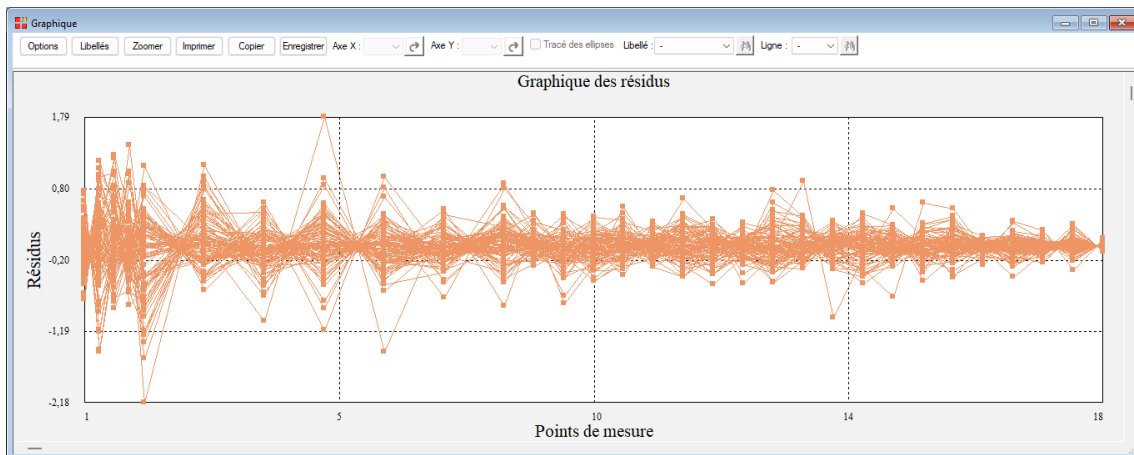
- Graphique moyenne globale, observé et estimé pour une observation

Après le choix de l'observation, ce graphique affiche la moyenne globale observée et les valeurs observées et estimées pour cette observation.



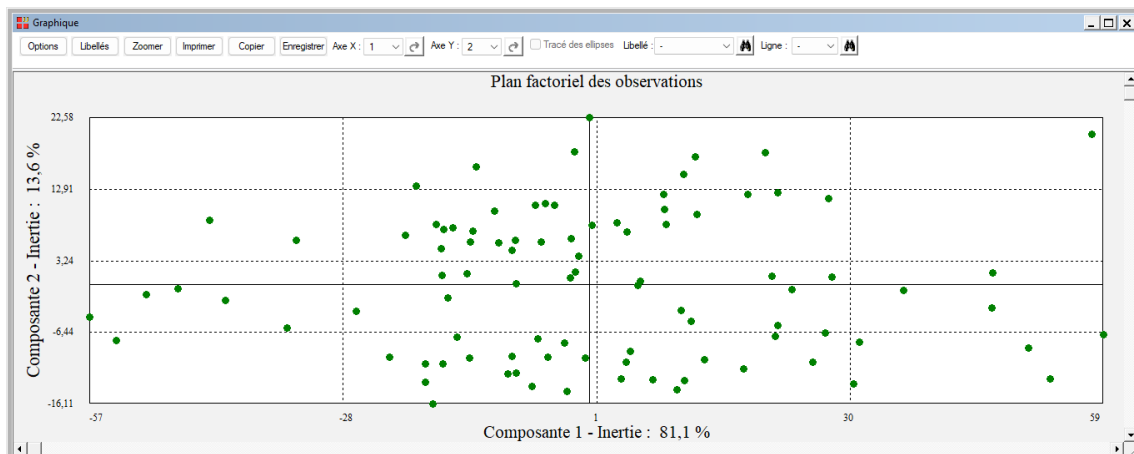
- Graphique des résidus

Ce graphique affiche les résidus (valeurs observées – valeurs estimées) par rapport aux points de mesure. Comme précédemment, il est possible de choisir les observations à afficher.



- Graphique des scores des observations

Ce graphique affiche les scores des observations par rapport à deux des composantes

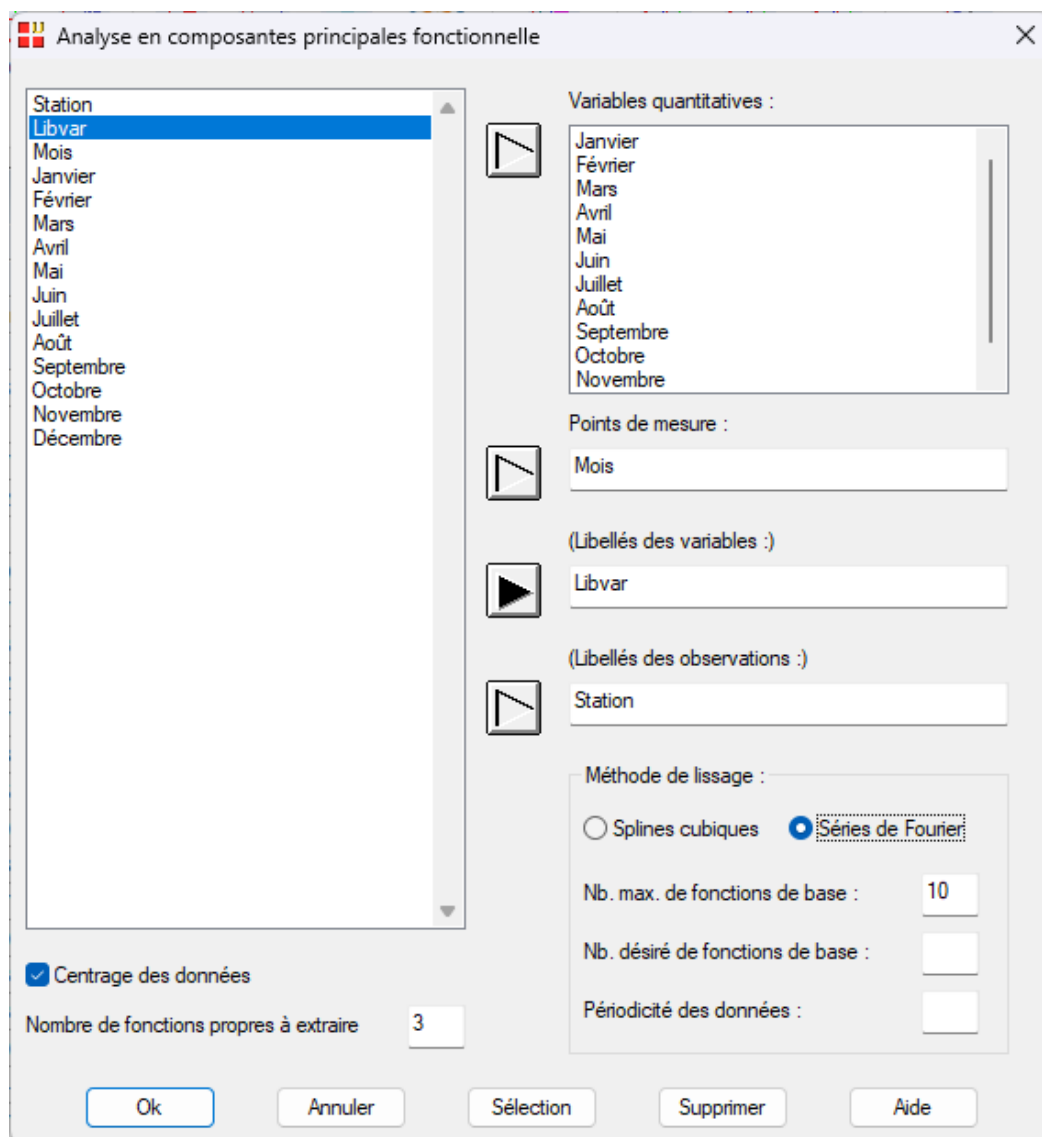


La barre d'outils permet notamment de choisir les composantes à afficher, de gérer les options graphiques, d'afficher ou non les libellés des observations. Il est possible d'interroger le graphique en utilisant les options Libellés et Lignes.

## Exemple 2 : Fichier CanadaTempératureMois

Ce fichier contient les températures (Celsius) mensuelles relevées dans 35 stations météorologiques au Canada moyennées sur la période 1960 à 1994. Cliquons sur l'icône ACPF dans le ruban Décrire et renseignons la boîte de dialogue comme montré ci-après.

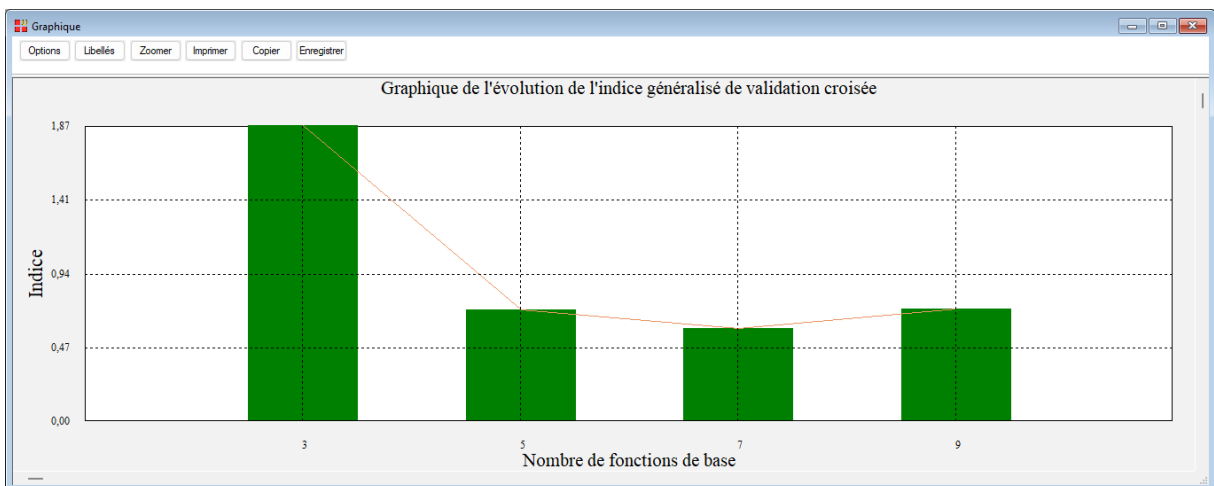
Les 12 variables quantitatives à utiliser sont *Janvier* à *Décembre*, les points de mesure sont *Mois* (12 valeurs), les données sont centrées, une base d'au maximum 10 séries de Fourier est demandée et 3 fonctions propres seront extraites. La périodicité n'étant pas indiquée, elle sera par défaut de 12. Les libellés des variables sont *Libvar* et ceux des observations *Station*.



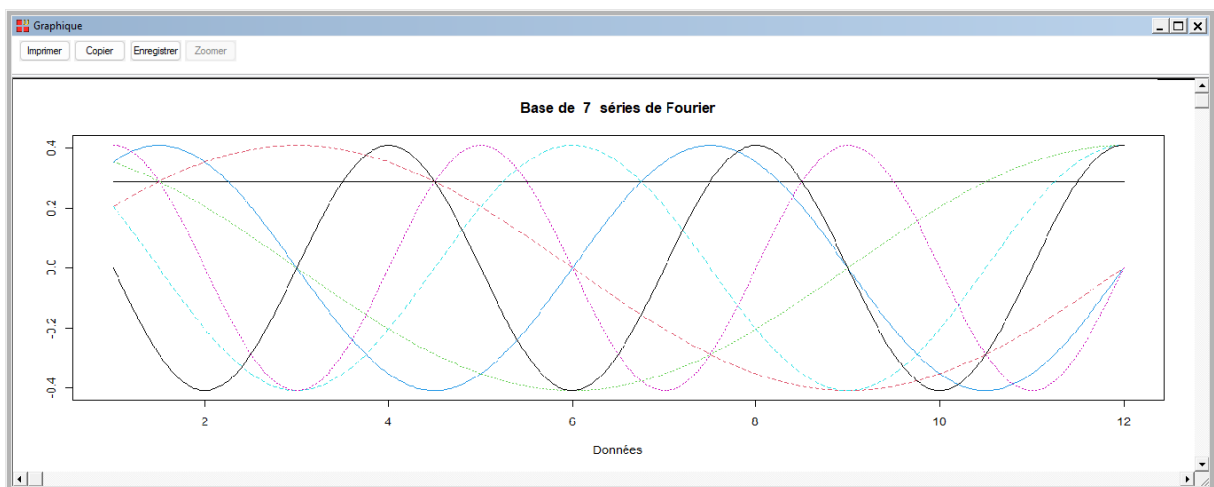
Après quelques instants, l'écran montré ci-après s'affiche.

	1	2	3	4	5	6	7	8
7	FICHER(S) DE DONNEES OUVERT(S) : CANADATEMPERATUREMOIS.SGD							
8								
9	RESULTATS DE L'ANALYSE EN COMPOSANTES PRINCIPALES FONCTIONNELLE							
10								
11	Sélection :							
12	Aucune							
13								
14	Nombre d'observations : 35							
15	Nombre de variables quantitatives : 12							
16	Données centrées							
17	Points de mesure : Mois							
18								
19	Type de base pour l'interpolation et le lissage : séries de Fourier							
20	Périodicité de la série : 12							
21	Nombre minimum de fonctions de base : 3							
22	Nombre maximum de fonctions de base : 9							
23								
24	Nombre demandé de fonctions propres à extraire : 3							
25								
26								
27								

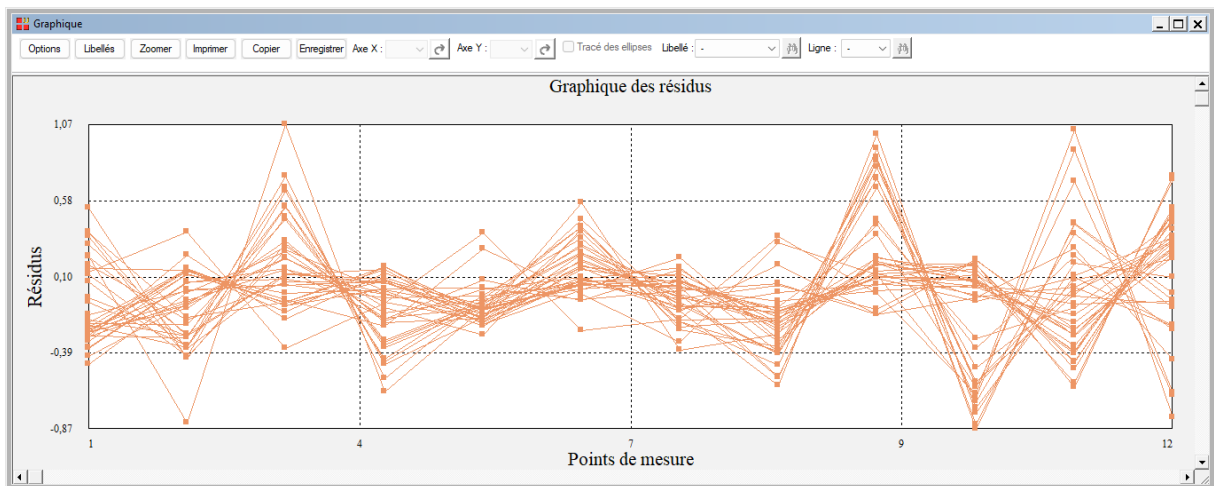
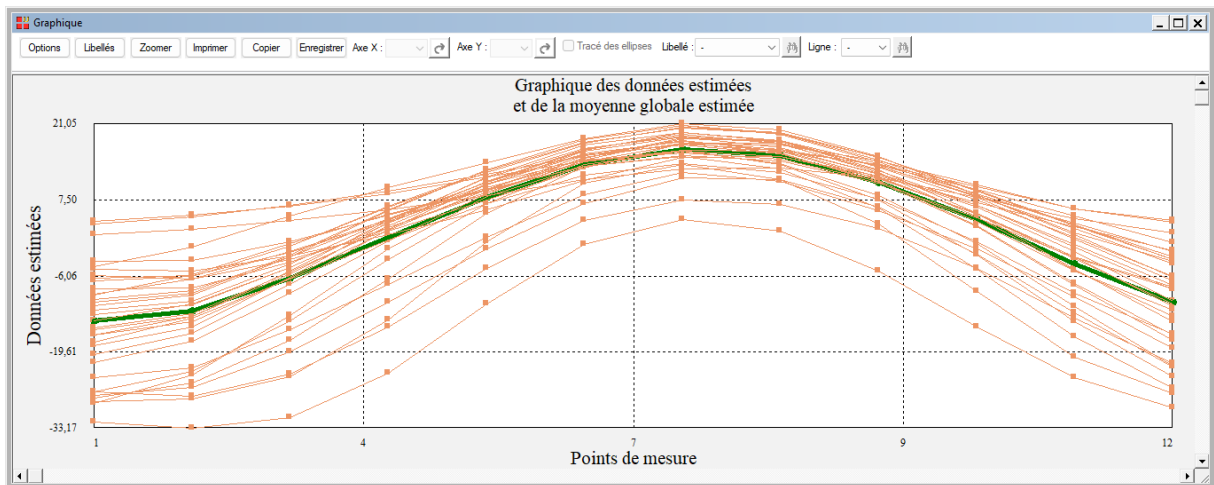
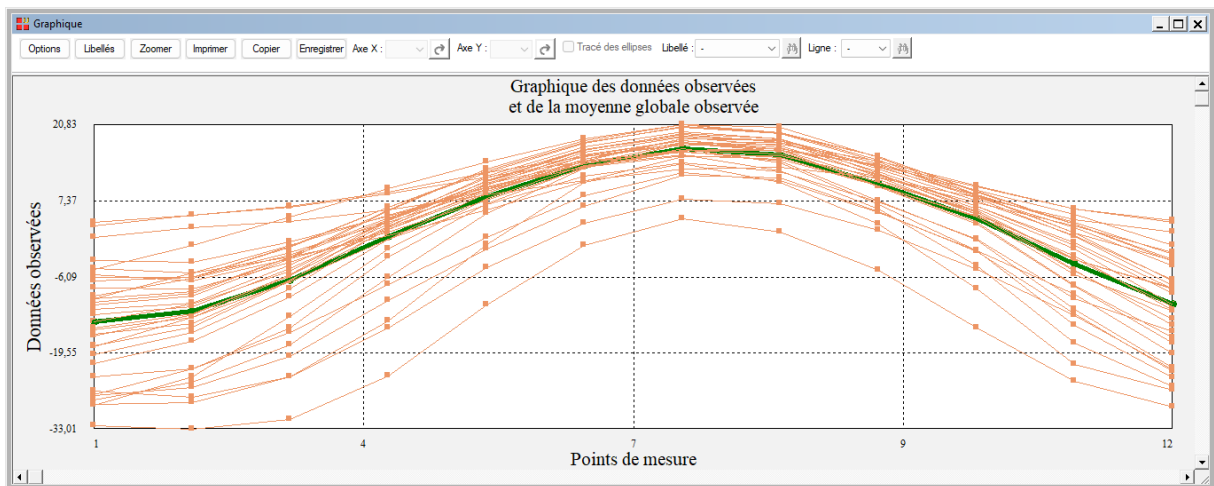
Visualisons les indices généralisés de validation croisée.



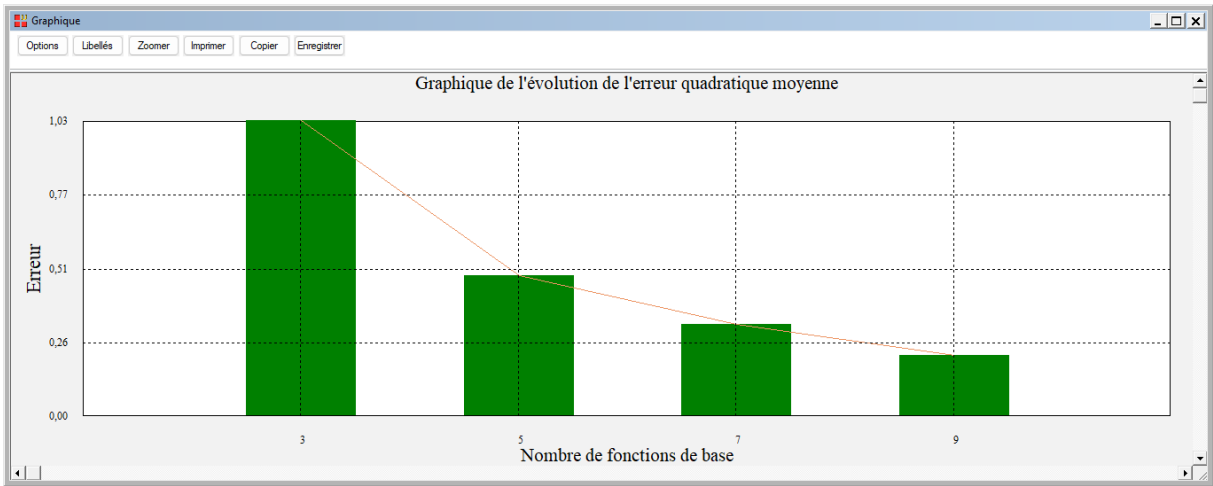
L'indice optimal est obtenu pour 7 fonctions de base : constante, sinus1, cosinus1, sinus2, cosinus2, sinus3, cosinus3. Visualisons cette base de séries de Fourier.



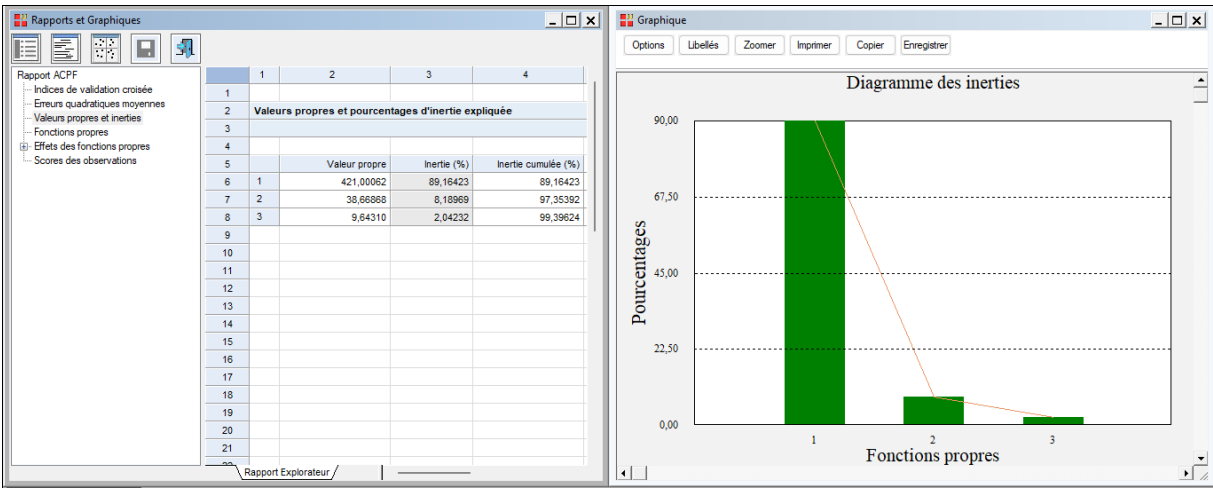
Visualisons les graphiques des données observées, de données estimées et des résidus.



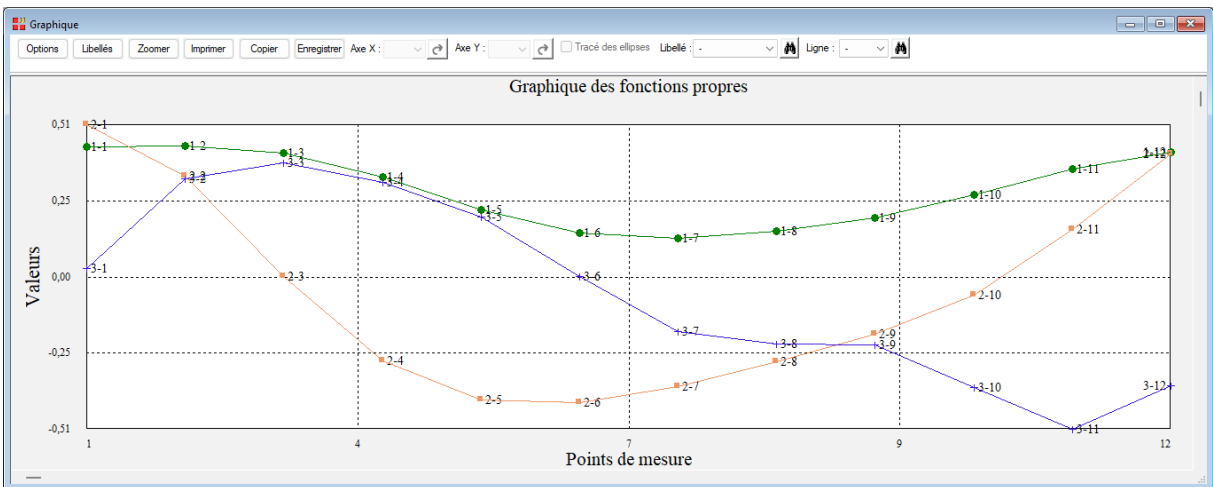
Augmenter la base à 9 permettrait de réduire les résidus. Conservons toutefois, la base définie par l'indice généralisé de validation croisée.



Le diagramme des inerties nous indique qu'environ 97% de l'inertie est expliqué par les deux premières fonctions propres.



Visualisons les fonctions propres.



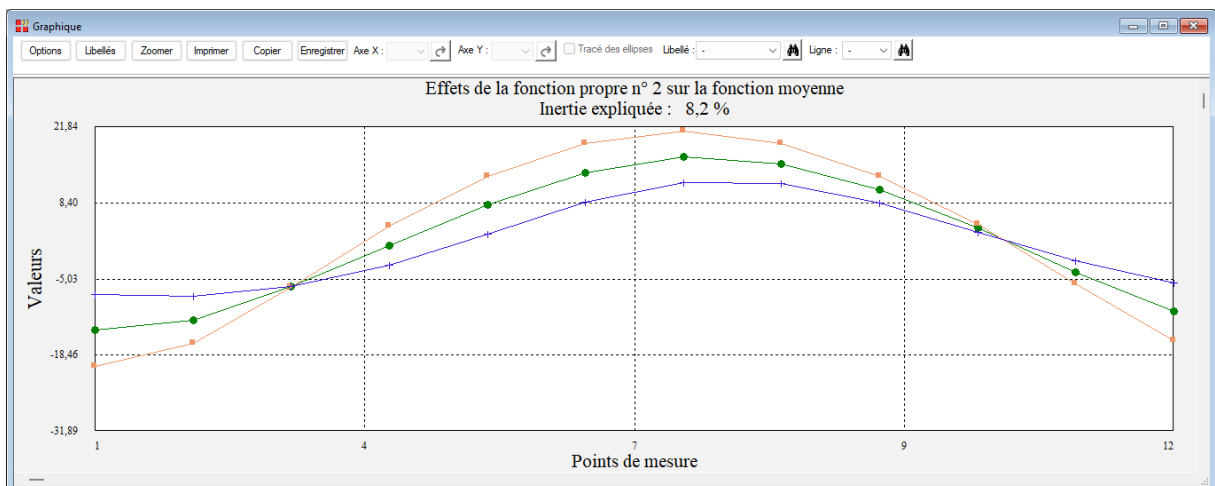
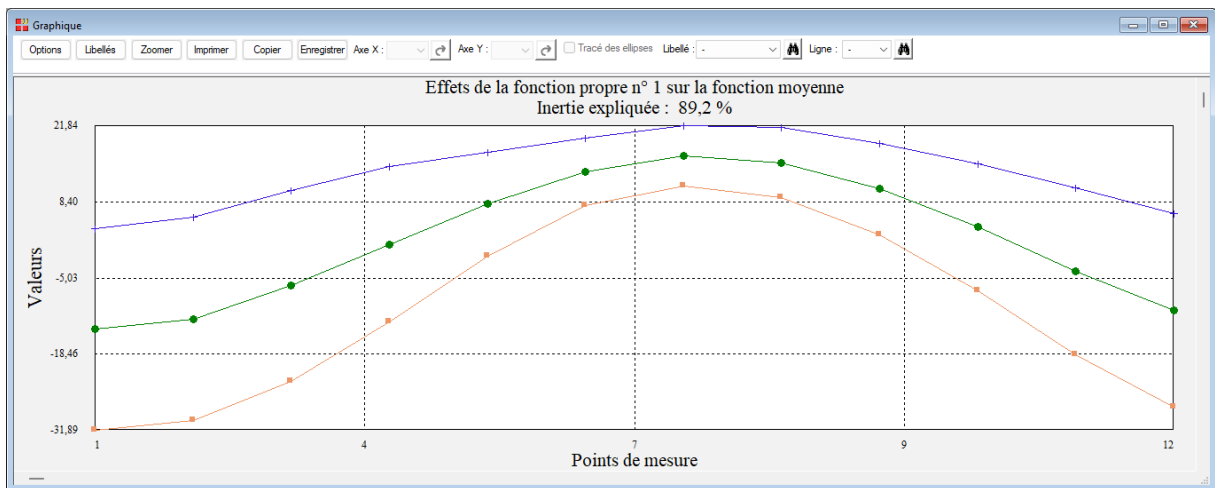
La première fonction propre montre les différences entre les mois d'hiver et ceux d'été.

Les stations ayant des valeurs positives élevées pour cette fonction ont des hivers beaucoup plus chauds que la moyenne et des étés plus chauds que la moyenne. Les stations ayant des valeurs négatives élevées pour cette fonction ont des hivers beaucoup plus froids que la moyenne et des étés plus froids que la moyenne.

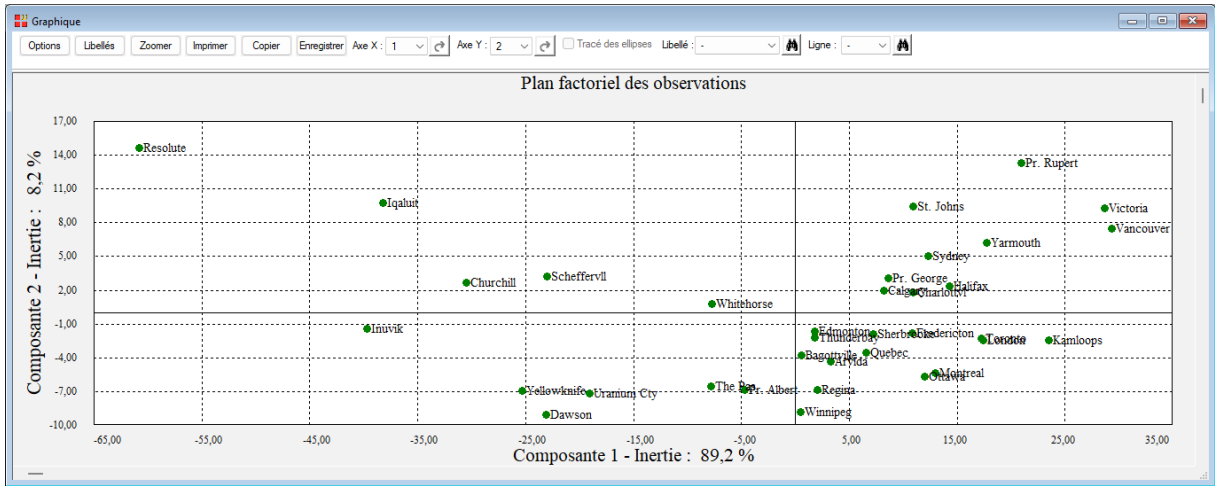
La deuxième fonction propre montre les différences entre les hivers plus chauds qu'en moyenne et les étés plus frais qu'en moyenne.

Les stations ayant des valeurs positives élevées pour cette fonction ont des hivers plus chauds que la moyenne et des étés plus froids que la moyenne. Les stations ayant des valeurs négatives élevées pour cette fonction ont des hivers plus froids que la moyenne et des étés plus chauds que la moyenne.

Visualisons les effets des fonctions propres sur la fonction propre moyenne (courbe verte).



Enfin, visualisons les scores pour les 35 stations.



**Note :** Le fichier CanadaTempératureJour contient les températures journalières pour les 35 stations.

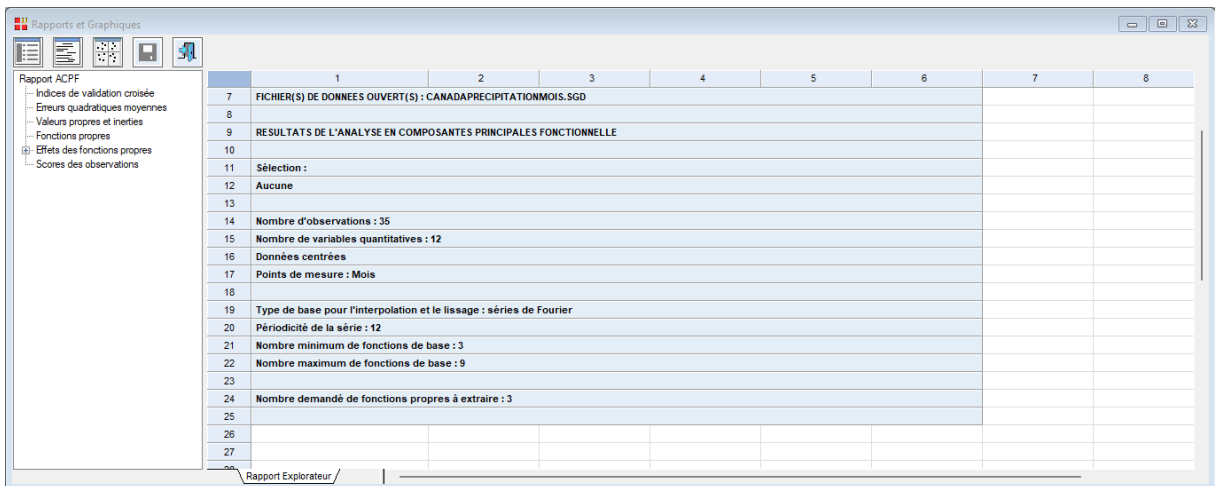
### Exemple 3 : Fichier CanadaPrécipitationMois

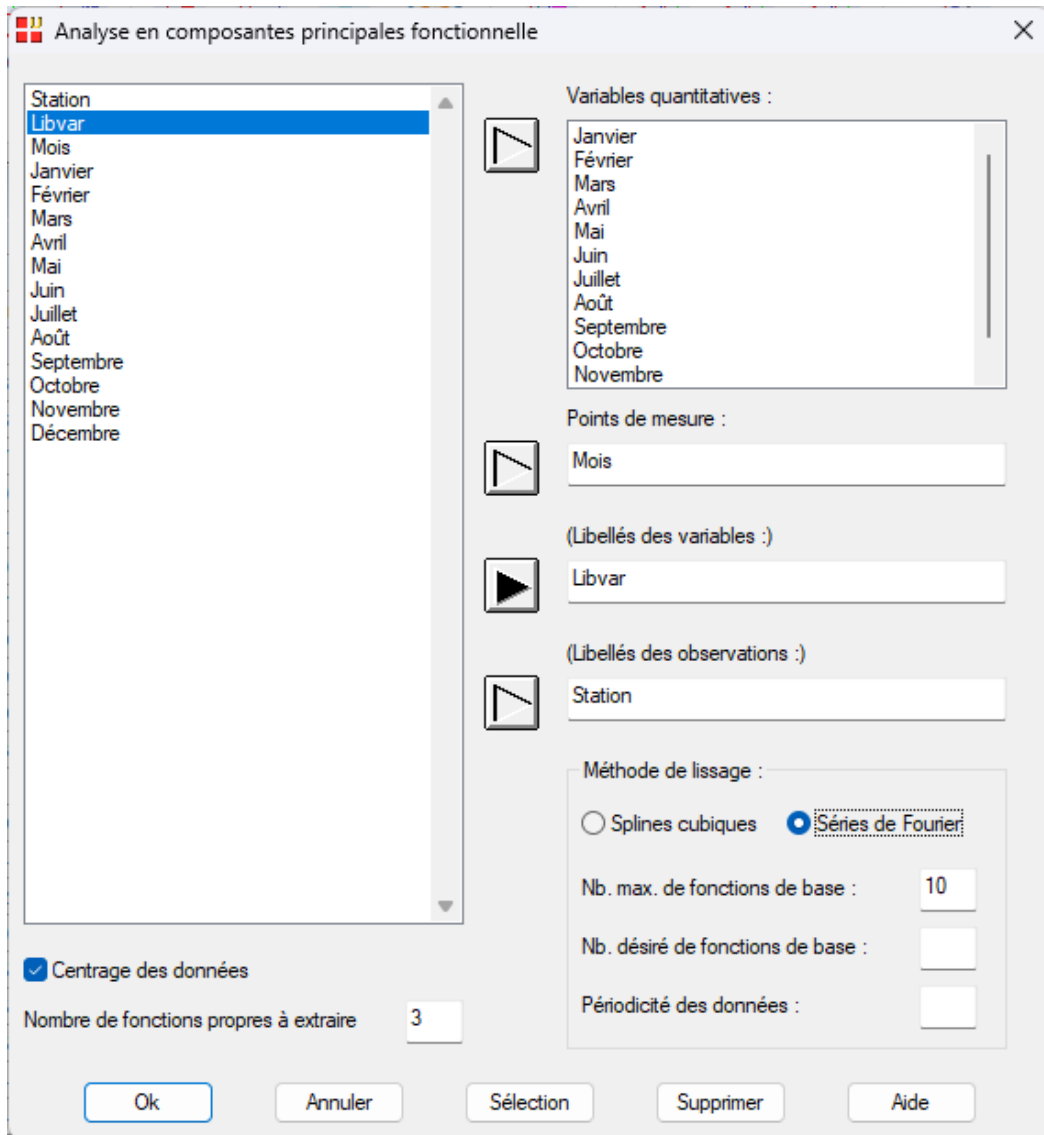
Ce fichier contient les précipitations mensuelles (logarithmes décimaux des millimètres) relevées dans 35 stations météorologiques au Canada moyennées sur la période 1960 à 1994.

Cliquons sur l'icône ACPF dans le ruban Décrire et renseignons la boîte de dialogue comme montré ci-après.

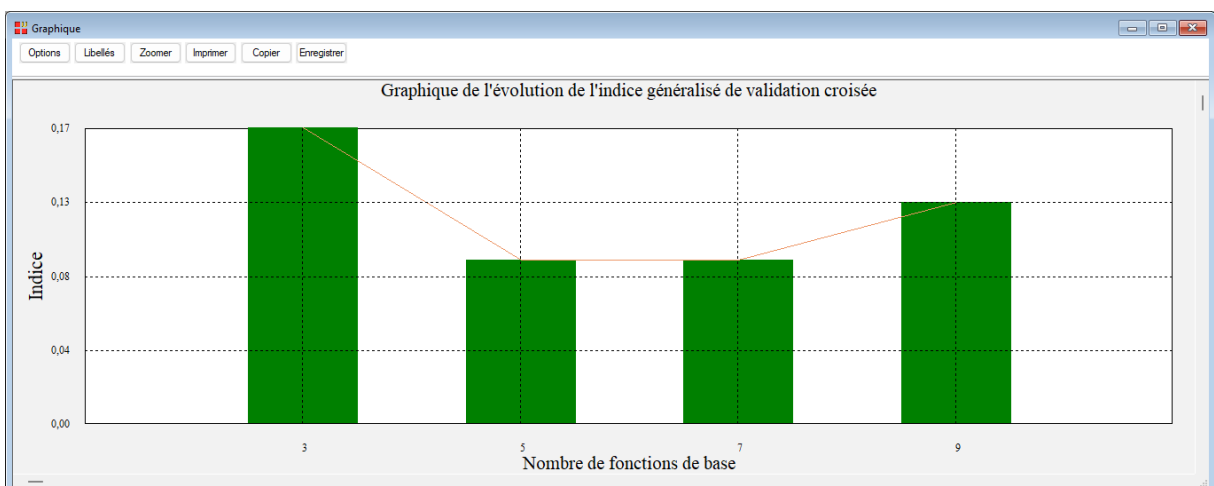
Les 12 variables quantitatives à utiliser sont *Janvier* à *Décembre*, les points de mesure sont *Jour* (12 valeurs), les données sont centrées, une base d'au maximum 10 séries de Fourier est demandée et 3 fonctions propres seront extraites. La périodicité n'étant pas indiquée, elle sera par défaut de 12. Les libellés des variables sont *Libvar* et ceux des observations *Station*.

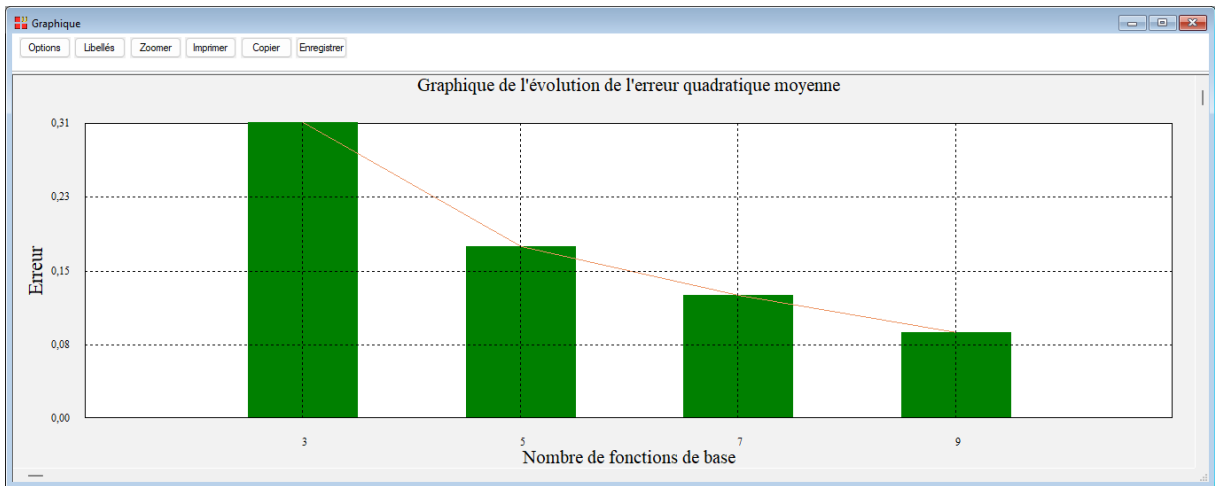
Après quelques instants, l'écran montré ci-après s'affiche.






Visualisons les graphiques des indices généralisés de validation croisée et les erreurs quadratiques moyennes.





A la vue de ces deux graphiques, il semble envisageable de choisir une base de 7 séries de Fourier. Pour cela, il suffit de rappeler la boîte de dialogue d'entrée des données via l'icône  et de préciser cette taille dans le champ 'Nombre désiré de fonctions de base'.

Méthode de lissage :

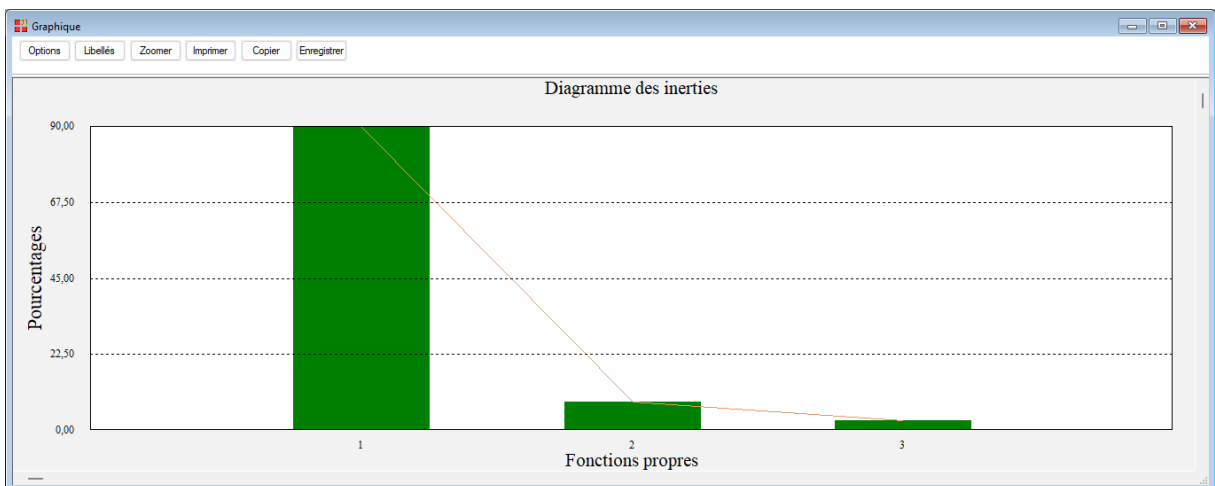
Splines cubiques     Séries de Fourier

Nb. max. de fonctions de base :

Nb. désiré de fonctions de base :

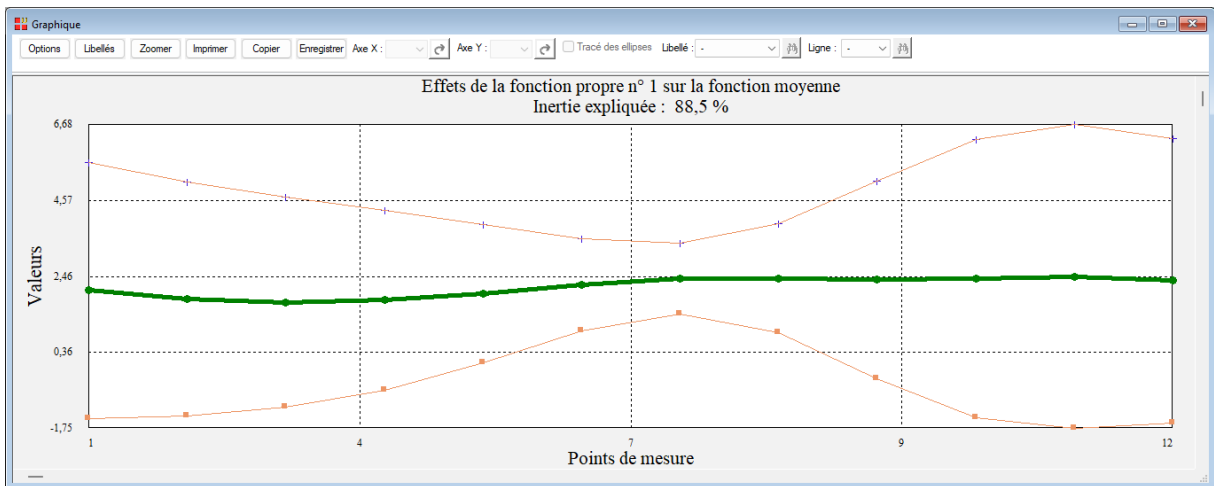
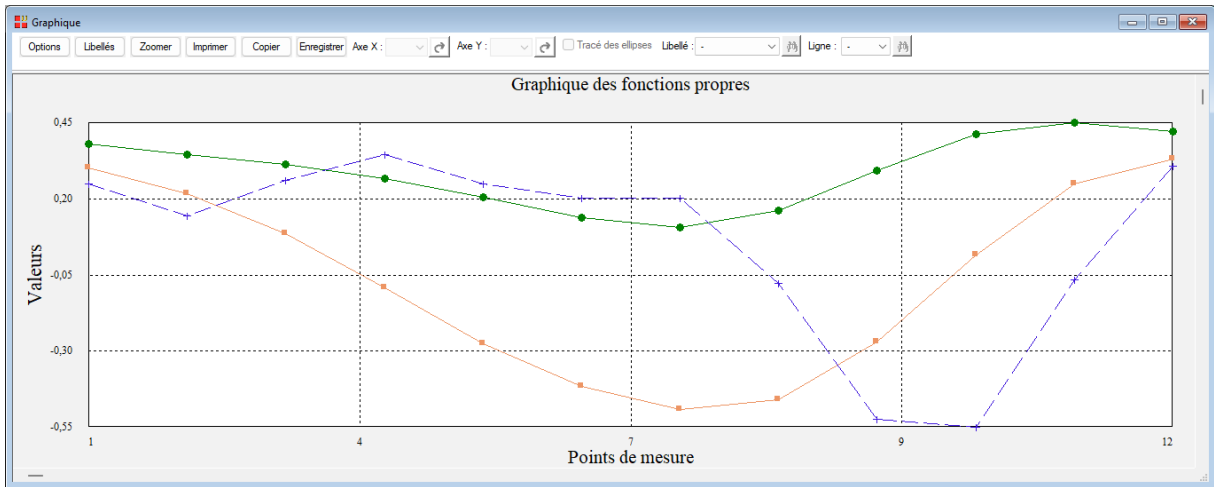
Périodicité des données :

Visualisons le diagramme des inerties.

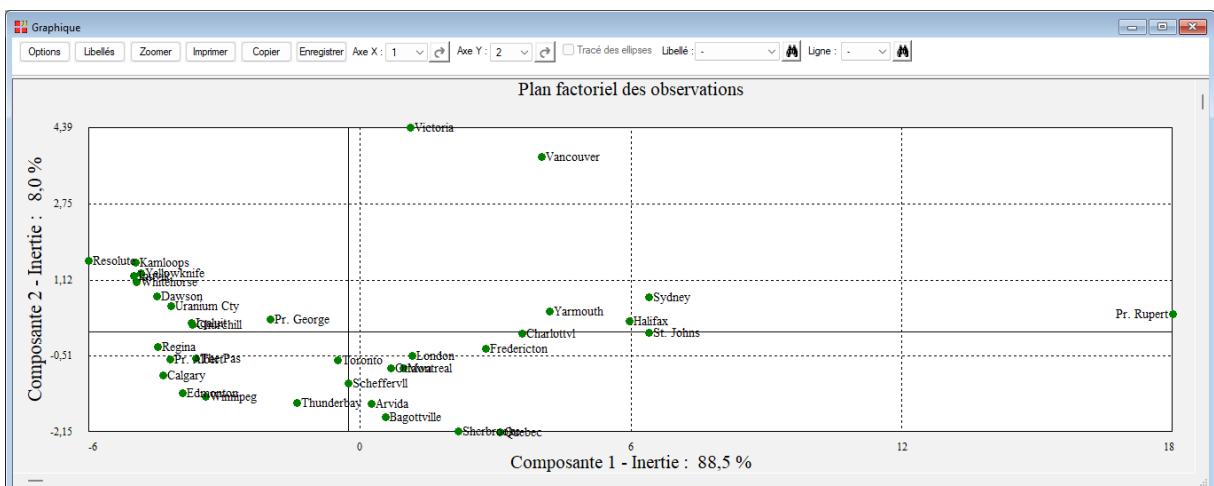


Le premier plan factoriel représente environ 96 % de l'inertie totale.

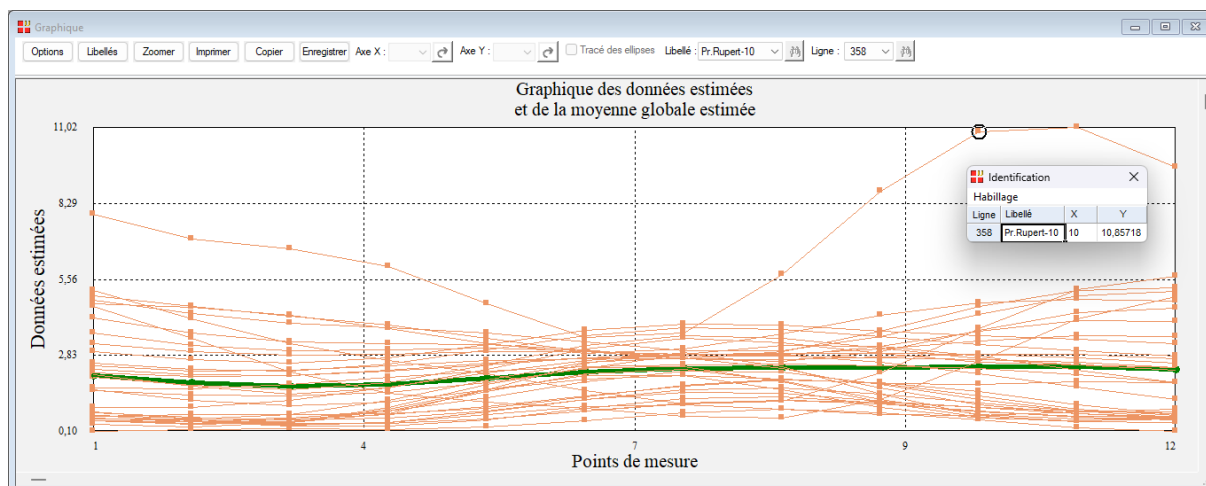
Visualisons les fonctions propres et les effets de la première fonction propre sur la fonction propre moyenne.



Affichons le graphique des scores.



Ce graphique montre que la station Pr. Rupert a un profil particulier comme l'indique le graphique des données estimées ci-dessous.



**Note :** Le fichier CanadaPrécipitationJour contient les précipitations journalières (mm) pour les 35 stations.

### Exemple 3 : Fichier Broyage (plan d'expériences)

Cet exemple illustre l'usage de la méthode ACPF dans le cadre d'un plan d'expériences élaboré pour étudier le broyage de particules de pigment.

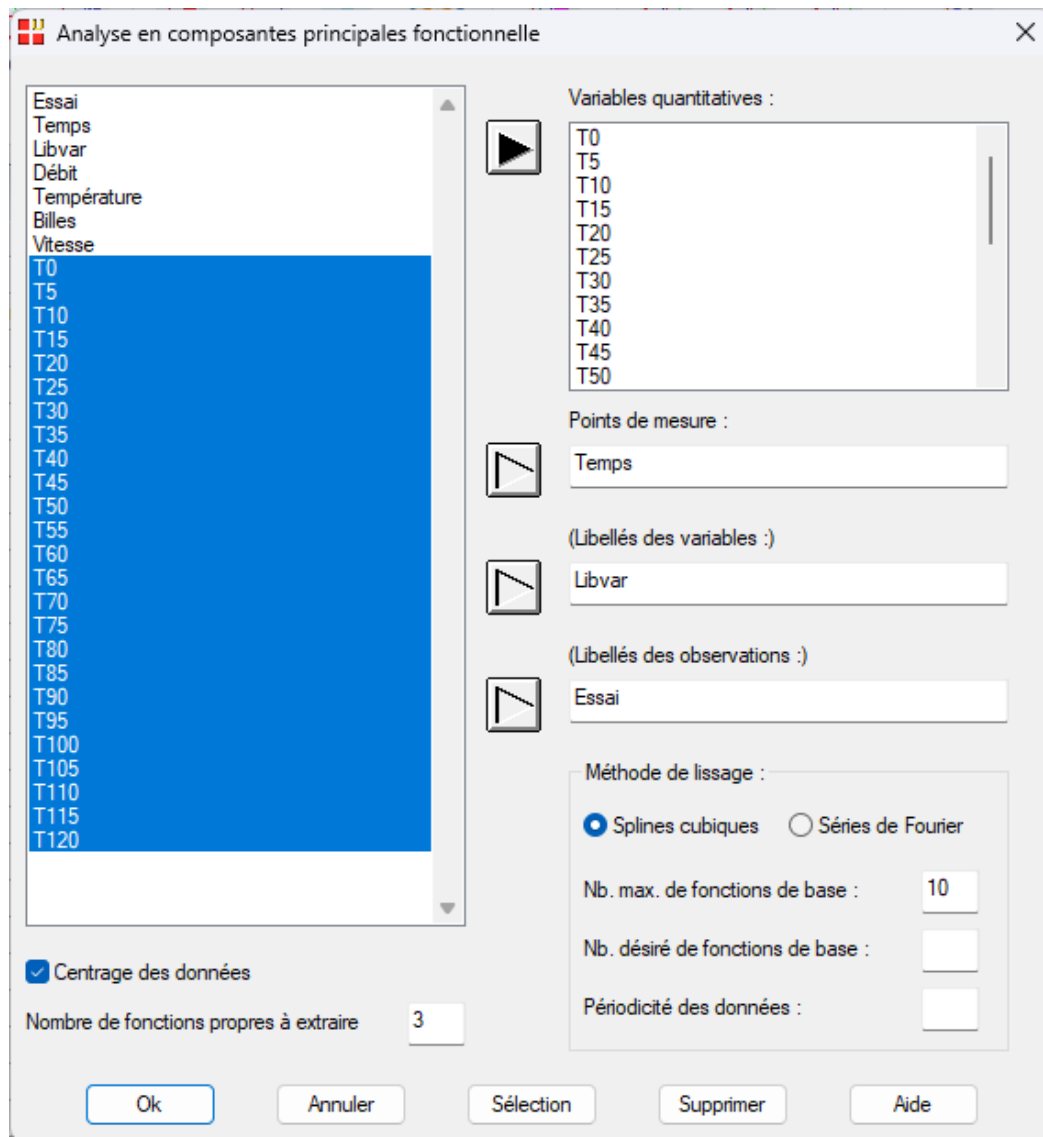
**Note :** cet exemple est inspiré d'une présentation faite par Chris Gotwalt et Tom Donnelly à DATAWorks 2019.

Le plan d'expériences utilisé est un plan factoriel complet comportant 4 facteurs ayant chacun 2 niveaux. Trois expériences ont été réalisées au centre du domaine expérimental. Le nombre total d'essais est ainsi de  $2^4 + 3 = 19$ .

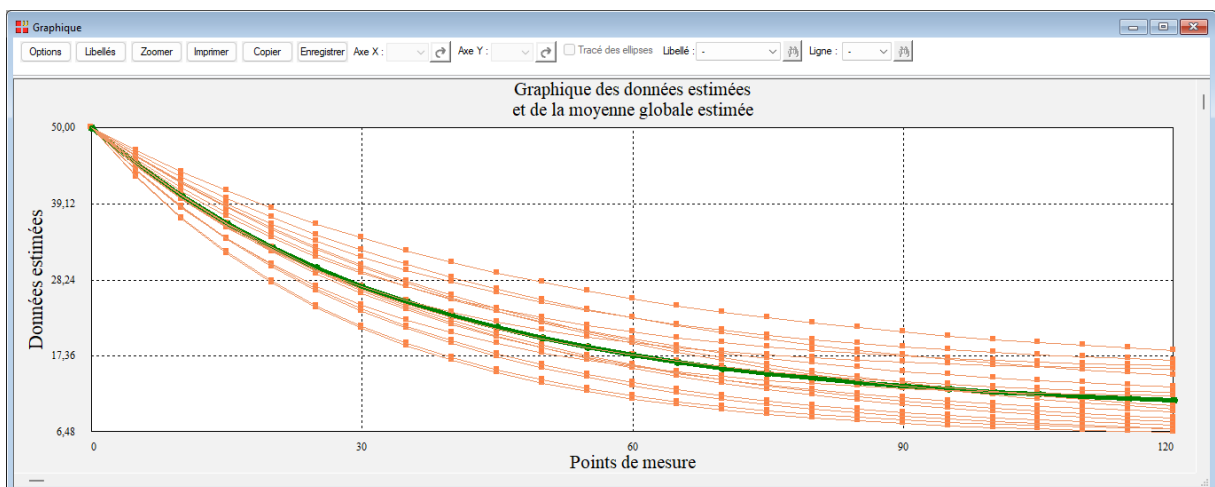
Facteurs	Domaine expérimental
Débit (mL/min)	de 50 à 150
Température (°C)	de 20 à 40
Billes (%)	de 60 à 80
Vitesse (rpm)	de 1000 à 2000

La réponse est la taille des particules ( $\mu\text{m}$ ) collectée pendant 120 minutes chaque 5 minutes. Il y a donc 25 mesures temporelles de la réponse pour chaque essai du plan d'expériences. Il est souhaité obtenir une taille comprise entre 8  $\mu\text{m}$  et 10  $\mu\text{m}$  (spécifications) le plus rapidement possible et y rester le plus longtemps possible.

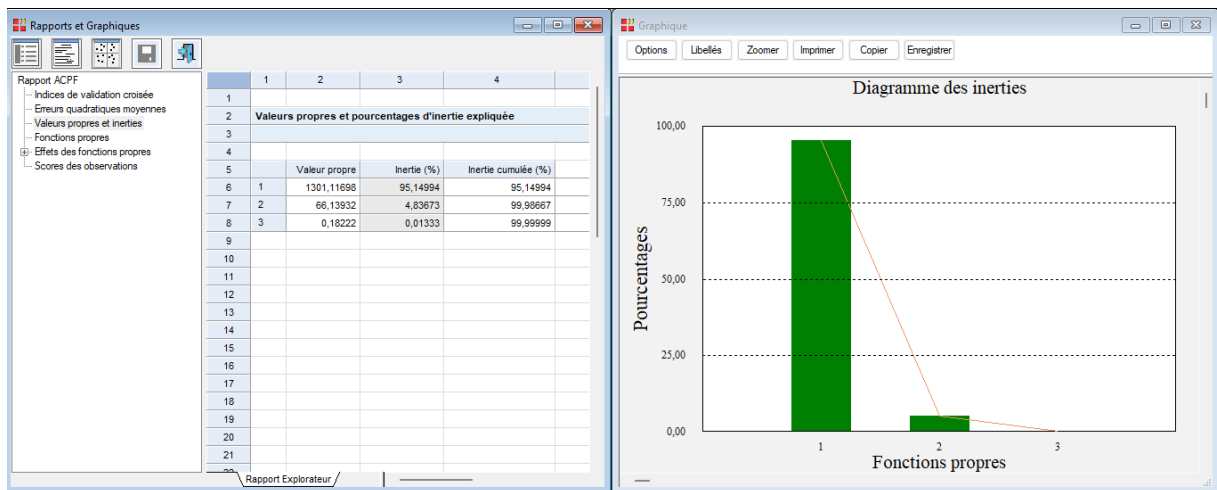
Renseignons la boîte de dialogue de l'ACPF comme montré ci-dessous. La variable Temps contient les points de mesure et les 25 variables T0 à T120 les tailles des particules pour chacun de ces points de mesure. Un lissage par une base de splines cubiques est demandé.



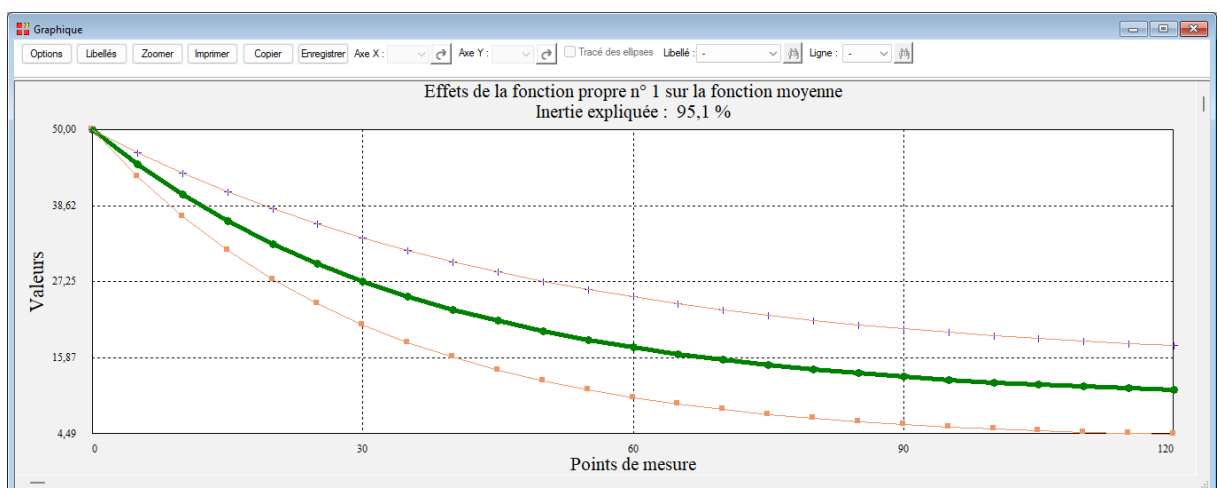
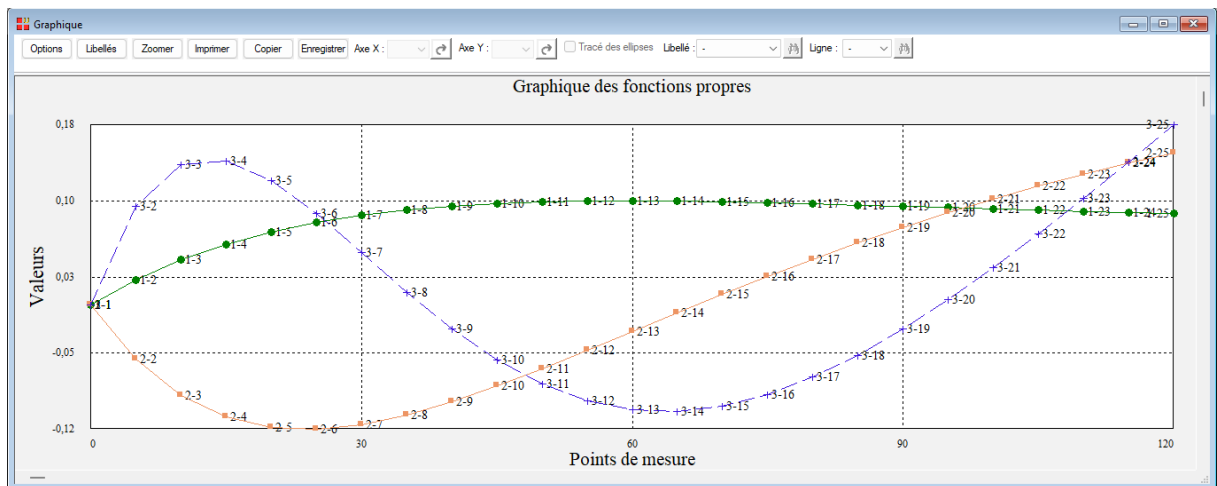
Visualisons le graphique des données estimées et de la moyenne globale estimée.

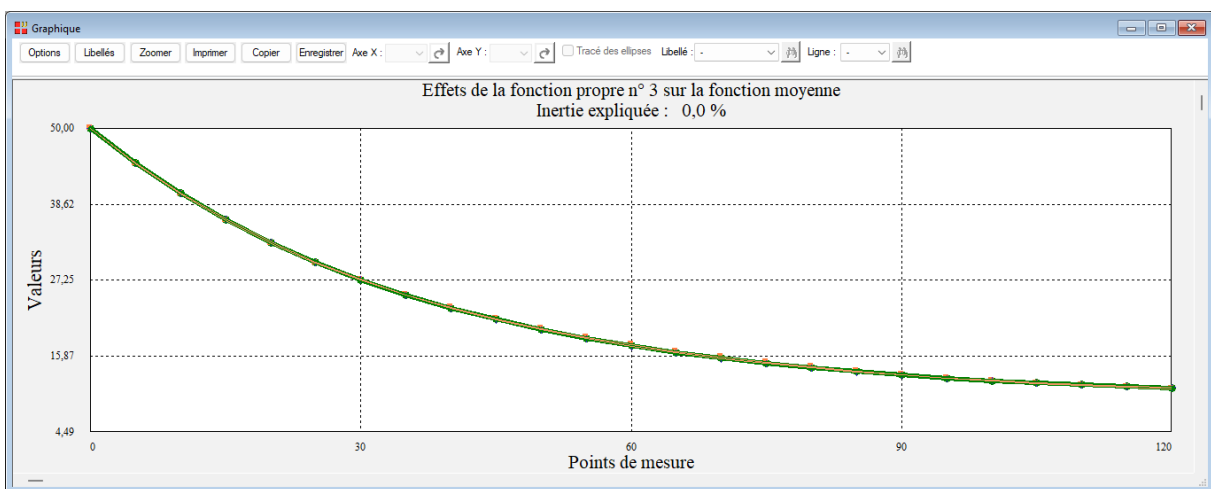
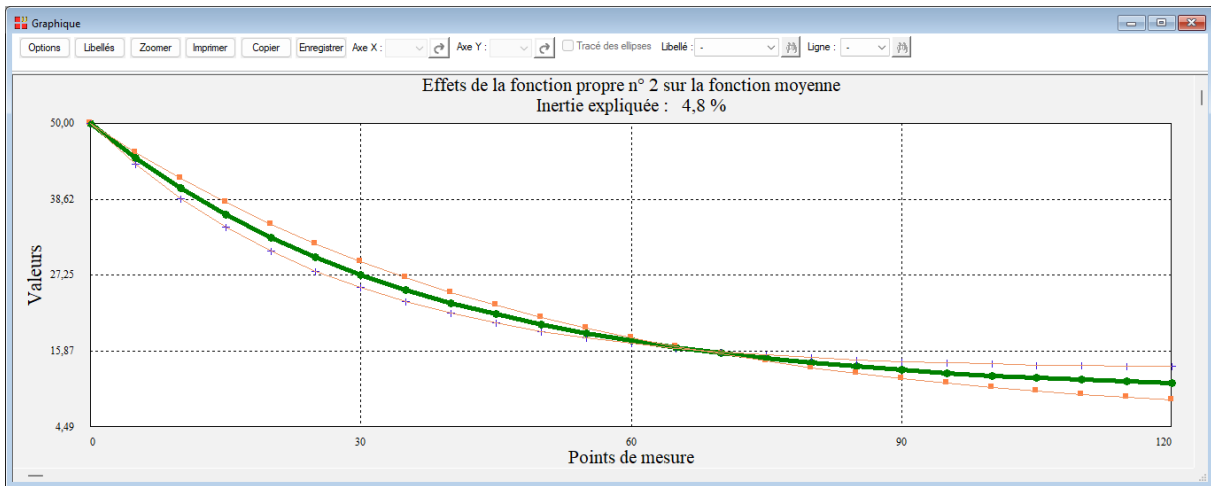


Trois fonctions propres expliquent presque toute l'inertie des données.

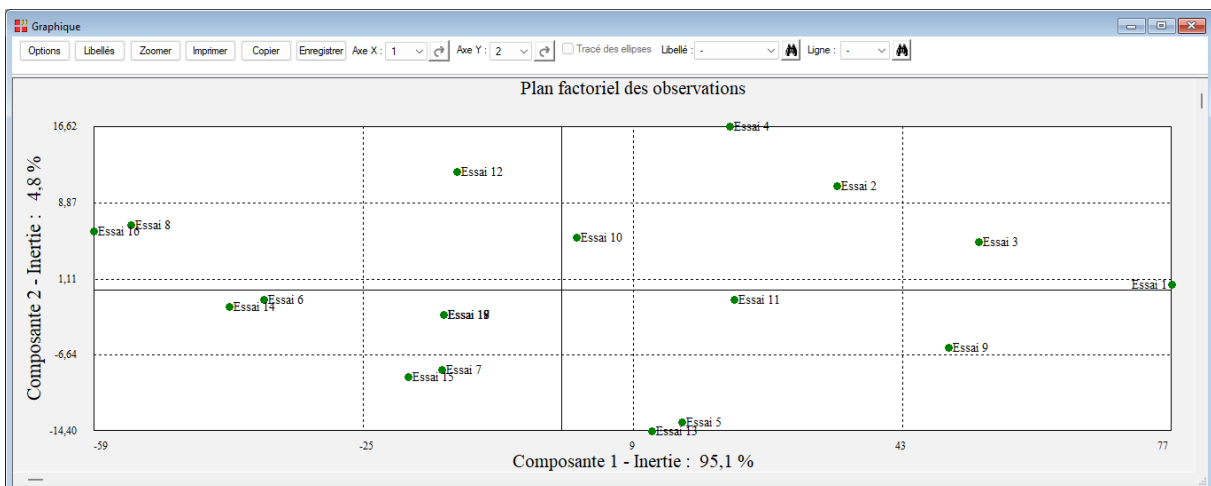


Visualisons les effets de ces trois fonctions propres sur la fonction propre moyenne.





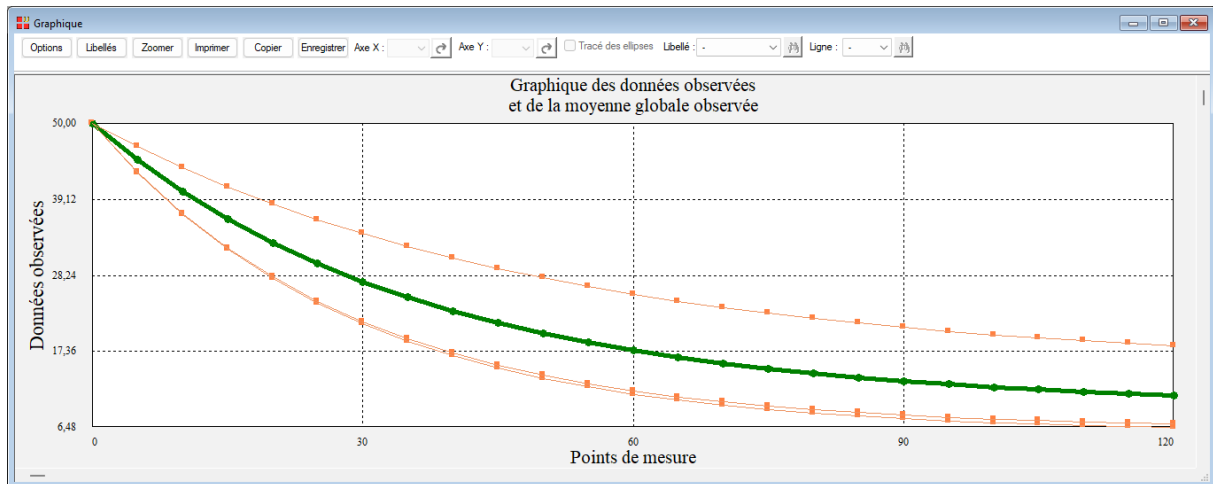
Visualisons le graphique des scores des observations (essais).



L'axe 1 (95,1 % de l'inertie) oppose principalement l'essai 1 aux essais 8 et 16.

Visualisons les courbes pour ces essais.

La courbe pour l'essai 1 est au-dessus de la courbe moyenne (en vert) et les courbes pour les essais 8 et 16 sont au-dessous, ce que montre bien le graphique des scores.



Les niveaux des 4 facteurs pour ces 3 essais sont :

Essai	Débit	Température	Billes	Vitesse
1	50	20	60	1000
8	150	40	80	1000
16	150	40	80	2000

Des niveaux hauts pour Débit, Température et Billes semblent préférables pour obtenir une décroissance rapide de la taille des particules.



Cliquons sur l'icône et enregistrons les informations suivantes dans le fichier ACPF\_Broyage : libellés des observations, fonction propre moyenne, fonctions propres, scores des observations.

Ouvrons le logiciel STATGRAPHICS Centurion et créons le fichier des données pour le plan d'expériences en combinant les informations :

- du fichier Broyage : Essai, Débit, Température, Billes, Vitesse
- du fichier ACPF\_Broyage : scores\_1, scores\_2, scores\_3

Enregistrons ce fichier sous le nom PEX\_Broyage.

Dans le ruban PEX, sélectionnons 'Convertir un fichier de données – Fichier pour l'assistant' et renseignons la boîte de dialogue comme montré ci-après.

Essai	Débit	Température	Billes	Vitesse	scores_1	scores_2	scores_3
	mL/min	°C	%	rpm	Scores des observations	Scores des observations	Scores des observations
Caractère	Numérique	Numérique	Numérique	Numérique	Numérique	Numérique	Numérique
Essai 1	50,0	20,0	60,0	1000,0	77,1535463389	0,51962126743	-0,2998693721
Essai 2	150,0	20,0	60,0	1000,0	34,7866749592	10,5404018294	0,38897690644
Essai 3	50,0	40,0	60,0	1000,0	52,7743961647	4,85820770898	0,41945654996
Essai 4	150,0	40,0	60,0	1000,0	21,2516243129	16,6200413238	-0,3164328486
Essai 5	50,0	20,0	80,0	1000,0	15,2688354523	-13,5318905550	-0,5874931842
Essai 6	150,0	20,0	80,0	1000,0	-37,564647025	-1,03550559240	0,27153054353
Essai 7	50,0	40,0	80,0	1000,0	-15,133112144	-8,12146907590	0,30954009342
Essai 8	150,0	40,0	80,0	1000,0	-54,443492166	6,54609097638	-0,6081486189
Essai 9	50,0	20,0	60,0	2000,0	48,9926466349	-5,8745786785	-0,4307538057
Essai 10	150,0	20,0	60,0	2000,0	1,86289956038	5,27273614569	0,33553244174
Essai 11	50,0	40,0	60,0	2000,0	21,8727969932	-1,0482491669	0,36943860109
Essai 12	150,0	40,0	60,0	2000,0	-13,193758928	12,0358475658	-0,4491793449
Essai 13	50,0	20,0	80,0	2000,0	11,4576721147	-14,397251411	-0,6052064662
Essai 14	150,0	20,0	80,0	2000,0	-42,020395493	-1,74840661980	0,26429762177
Essai 15	50,0	40,0	80,0	2000,0	-19,315188821	-8,92082109010	0,30277090023
Essai 16	150,0	40,0	80,0	2000,0	-59,105167863	5,92568781585	-0,6261139038
Essai 17	100,0	30,0	70,0	1500,0	-14,881776695	-2,54682081570	0,42055129552
Essai 18	100,0	30,0	70,0	1500,0	-14,881776695	-2,54682081570	0,42055129552
Essai 19	100,0	30,0	70,0	1500,0	-14,881776695	-2,54682081570	0,42055129552

Importer un plan d'expériences

Essai  
Débit  
Température  
Billes  
Vitesse  
scores\_1  
scores\_2  
scores\_3

Variables de réponse:

▶ scores\_1  
scores\_2  
scores\_3

Facteurs procédé:

▶ Débit  
Température  
Billes  
Vitesse

Composants de mélange:

▶

Facteurs de bruit:

▶

Colonnes supplémentaires:

▶

Trier les noms des colonnes

OK Annuler Supprimer Aide

Assistant pour les plans d'expériences

1) Définir les réponses à mesurer

Nom	Unités	Analyser	But	Cible	Impact	Sensibilité	Bas	Haut
scores_1	Scores des observations	Moyenne	Maximiser		3.0	Moyenne		
scores_2	Scores des observations	Moyenne	Maximiser		3.0	Moyenne		
scores_3	Scores des observations	Moyenne	Maximiser		3.0	Moyenne		

2) Définir les facteurs expérimentaux à faire varier

Nom	Unités	Type	Rôle	Bas	Haut	Niveaux
A:Débit	mL/min	Continu	Contrôlable	50.0	150.0	
B:Température	°C	Continu	Contrôlable	20.0	40.0	
C:Billes	%	Continu	Contrôlable	60.0	80.0	
D:Vitesse	rpm	Continu	Contrôlable	1000.0	2000.0	

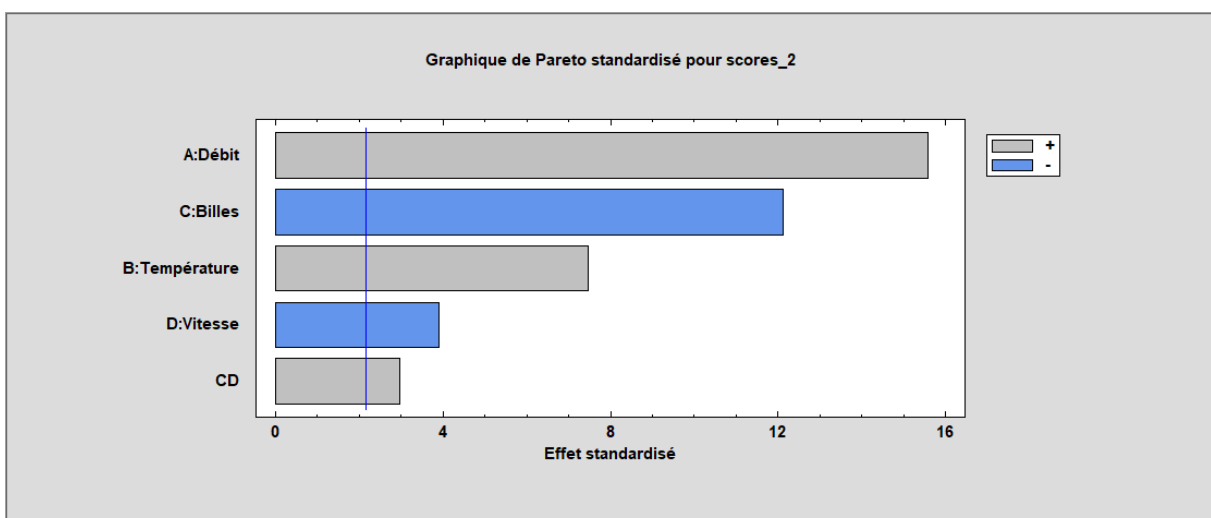
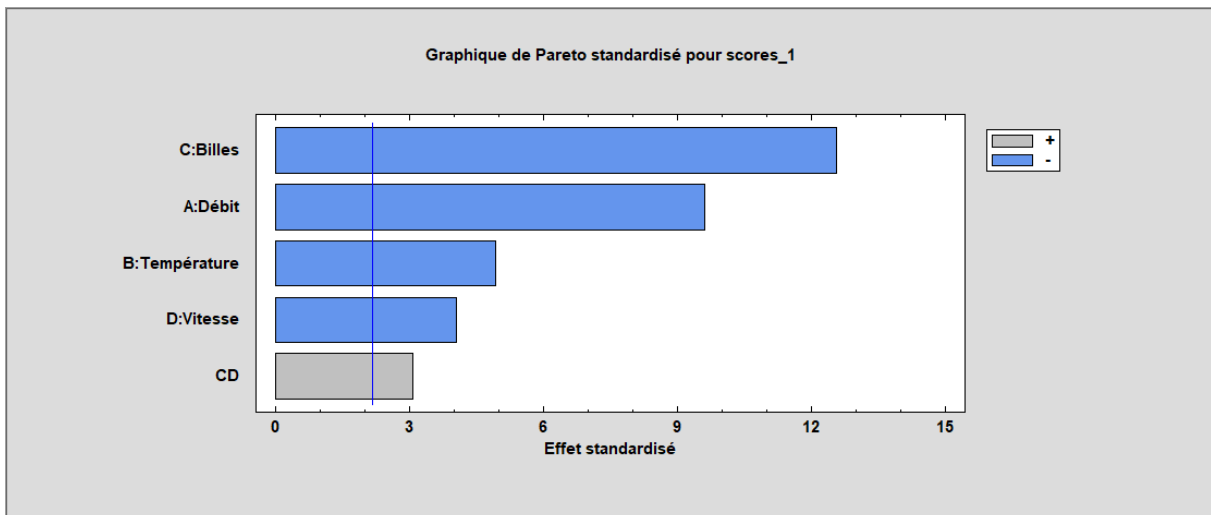
3) Sélectionner le plan d'expériences

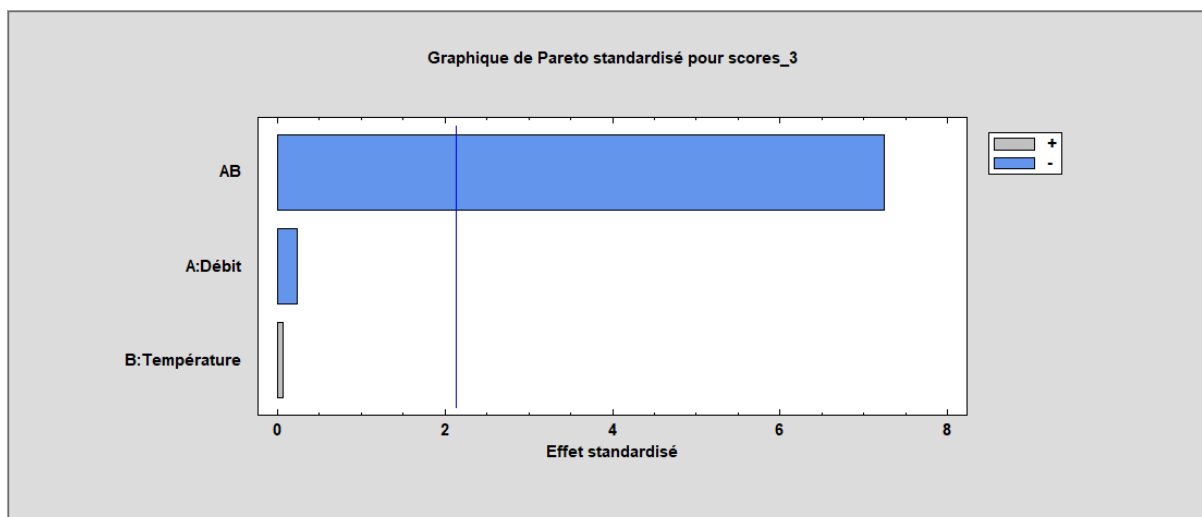
Type de facteurs	Type de plan	Points au centre par bloc	Emplacement des points au centre	Plan randomisé	Nombre de répétitions	Nombre total d'essais
Procédé	Défini par l'utilisateur					

Type de facteurs	Nombre total de blocs	D.D.L. pour l'erreur
Procédé		

Nombre d'échantillons par essai: 1

Analysons chacune des trois réponses scores\_1, scores\_2 et scores\_3.





Nous obtenons ainsi les valeurs des effets pour les modélisations des trois réponses :

EFFETS1	ESTIMATIONS1	EFFETS2	ESTIMATIONS2	EFFETS3	ESTIMATIONS3
Effets	Estimations	Effets	Estimations	Effets	Estimations
Numérique	Numérique	Numérique	Numérique	Numérique	Numérique
moyenne	-0,0000105263	moyenne	-0,00000784211	moyenne	5,26316E-8
A:Débit	-42,6872	A:Débit	12,5842	A:Débit	-0,0271777
B:Température	-21,9037	B:Température	6,01877	B:Température	0,0080395
C:Billes	-55,7945	C:Billes	-9,77595	AB	-0,823093
D:Vitesse	-17,9428	D:Vitesse	-3,14382		
CD	13,6651	CD	2,39431		

Enregistrons ces résultats dans le fichier PEX\_Broyage.sgp.

Ainsi avec ces informations et pour un ensemble quelconque d'essais, nous pouvons calculer les scores (sur facteurs en unités codés) et en utilisant la formule de reconstitution fonctionnelle obtenir pour chaque nouvel essai la courbe prévue correspondante :

Courbe = Fonction propre moyenne + (Valeur prévue de score\_1 \* Fonction propre 1) + (Valeur prévue de score\_2 \* Fonction propre 2) + (Valeur prévue de score\_3 \* Fonction propre 3)

Voici ci-dessous un exemple des calculs dans une feuille Microsoft Excel pour trois essais (en unités codées).

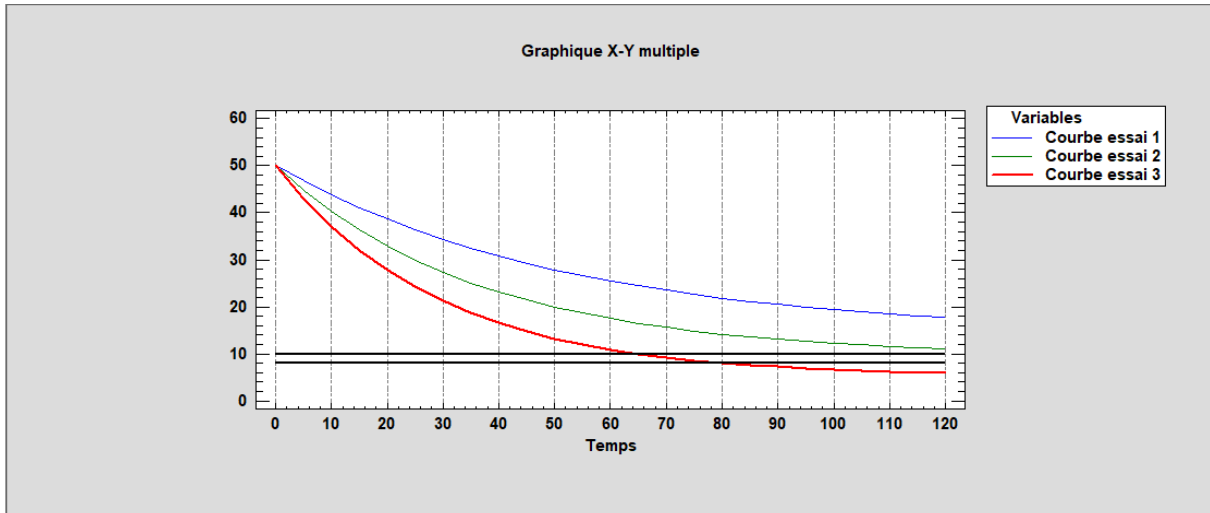
Essais	A	B	C	D	CD	AB	PEX score_1	PEX score_2	PEX score_3
Essai 1	-1	-1	-1	-1	1	1	75,9966447	-1,64444892	-0,40197737
Essai 2	0	0	0	0	0	0	-5,2632E-06	-3,9211E-06	1,64001E-06
Essai 3	1	1	1	1	1	1	-62,3315553	4,038751079	-0,42111557

Temps	fctpromoy	fctpro_1	fctpro_2	fctpro_3	Effets	ACPF scores_1	ACPF scores_2	ACPF scores_3
0	49,9999371	1,4008E-05	-3,0357E-05	8,01E-05	Moyenne	-1,05263E-05	-7,84211E-06	5,26E-08
5	44,7824542	0,02520504	-0,05449177	0,09862637	A:Débit	-42,6872	12,5842	-0,0271777
10	40,2520187	0,04527404	-0,09076722	0,14130591	B:Température	-21,9037	6,01877	0,0080395
15	36,3209026	0,06105747	-0,11239675	0,14457292	C:Billes	-55,7945	-9,77595	0
20	32,9024289	0,07338028	-0,12286642	0,12455219	D:Vitesse	-17,9428	-3,14382	0
25	29,9275708	0,08287492	-0,12475535	0,09183616	CD	13,6651	2,39431	0
30	27,3419473	0,09001406	-0,1198901	0,05242664	AB	0	0	-0,823093
35	25,0916314	0,09526544	-0,11007398	0,01218419				
40	23,1280436	0,09903937	-0,0968628	-0,02420186	Effets/2	ACPF scores_1	ACPF scores_2	ACPF scores_3
45	21,4166367	0,10159554	-0,08116296	-0,05511543	Moyenne	-5,26315E-06	-3,92106E-06	2,63158E-08
50	19,9251372	0,10316924	-0,06377562	-0,07943848	A:Débit	-21,3436	6,2921	-0,01358885
55	18,6220299	0,10398876	-0,04547449	-0,09614548	B:Température	-10,95185	3,009385	0,00401975
60	17,4832464	0,10421401	-0,0267632	-0,10512011	C:Billes	-27,89725	-4,887975	0
65	16,4889455	0,10396603	-0,00799213	-0,10676208	D:Vitesse	-8,9714	-1,57191	0
70	15,6193641	0,10336522	0,01049107	-0,10147856	CD	6,83255	1,197155	0
75	14,8573585	0,10251176	0,0284233	-0,08980733	AB	0	0	-0,4115465
80	14,1903988	0,10147028	0,04569046	-0,07251637				
85	13,6064248	0,10030177	0,06219363	-0,0503971				
90	13,093857	0,09906285	0,07784247	-0,02422798				
95	12,6440826	0,09778306	0,09259974	0,00529246				
100	12,2496196	0,09648163	0,10644844	0,03749617				
105	11,9030247	0,09517787	0,11937257	0,07171763				
110	11,5980749	0,09389464	0,131389	0,10737394				
115	11,3300168	0,09265903	0,14255408	0,14398168				
120	11,0941849	0,09149842	0,15292657	0,18106341				

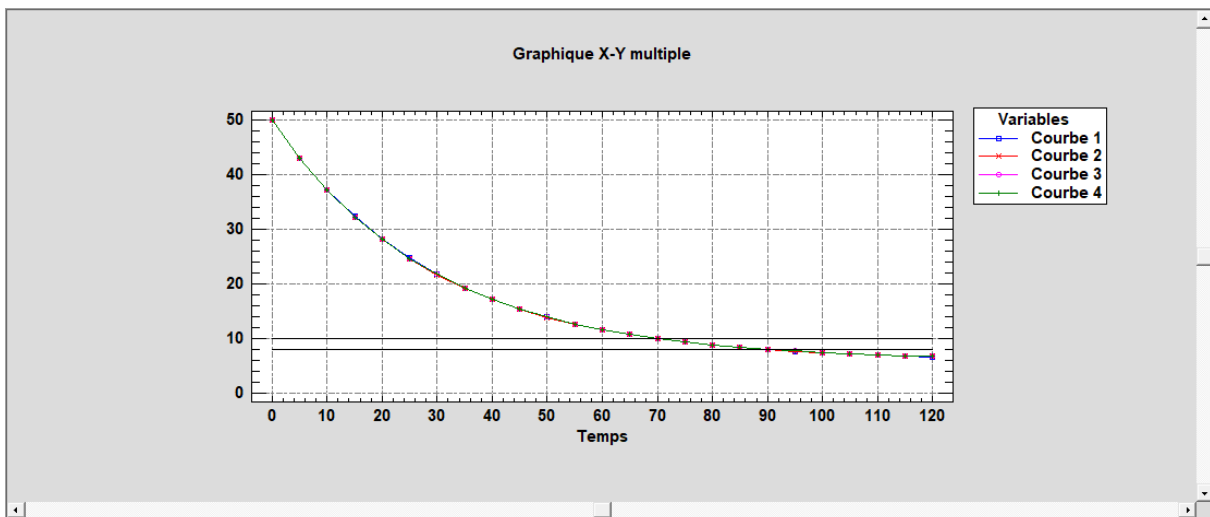
Temps	Courbe essai 1	Courbe essai 2	Courbe essai 3
0	50,0010194	49,99993711	49,99890763
5	46,74791571	44,78245445	42,94977334
10	43,78515442	40,25201905	37,0039248
15	41,08778106	36,32090299	32,00027127
20	38,63106445	32,9024292	27,7798441
25	36,39402435	29,92757099	24,21931872
30	34,35879285	27,34194742	21,22494691
35	32,50759834	25,09163139	18,70389619
40	30,82371794	23,12804344	16,57375273
45	29,29318062	21,41663638	14,77944143
50	27,90246125	19,9251368	13,27031718
55	26,63825551	18,62202938	11,99707701
60	25,48942785	17,48324583	10,9236032
65	24,44607376	16,48894477	10,02126186
70	23,49831428	15,61936338	9,26155406
75	22,63726852	14,85735774	8,620254724
80	21,85581395	14,19039799	8,080668641
85	21,14700704	13,60642395	7,626867072
90	20,50403225	13,09385613	7,243704713
95	19,92086407	12,64408177	6,92087105
100	19,39177803	12,2496187	6,649898087
105	18,91109267	11,90302388	6,422354665
110	18,47452781	11,59807407	6,230906826
115	18,07949196	11,33001604	6,069542911
120	17,7234942	11,09418408	5,932329984

Visualisons graphiquement les trois courbes obtenues et affichons les limites des spécifications (entre 8  $\mu\text{m}$  et 10  $\mu\text{m}$ ).

La courbe pour l'essai 3 atteint la spécification supérieure au temps 70 mais ne reste que jusqu'au temps 80 au-dessus de la spécification inférieure.



Une analyse plus fine peut aisément être mise en œuvre en maillant de façon régulière le domaine expérimental. A l'issue de cette analyse quatre courbes permettent d'atteindre le niveau de la spécification supérieure au temps 70 et de rester au-dessus de la limite de spécification inférieure jusqu'au temps 90.



Les réglages des facteurs pour ces courbes sont :

	Courbe 1	Courbe 2	Courbe 3	Courbe 4
Débit	1	1	1	1
Température	0,9	1	1	1
Billes	0,8	0,7	0,8	0,9
Vitesse	0,5	0,9	0,2	-0,7

soit dans les unités d'origine :

	Courbe 1	Courbe 2	Courbe 3	Courbe 4
Débit	150	150	150	150
Température	39	40	40	40
Billes	78	77	78	79
Vitesse	1750	1950	1600	1150

Note : Les fichiers ACPF\_Broyage.sgd, PEX\_Broyage.sgd, PEX\_Broyage.sgp et PEX\_Broyage.xlsx sont disponibles sur simple demande.

## **ACPF versus ACP**

L'analyse en composantes principales fonctionnelle (ACPF) présente plusieurs avantages substantiels par rapport à l'analyse en composantes principales (ACP) classique, même quand les résultats peuvent sembler similaires à première vue :

### Préservation de la structure temporelle/spatiale

L'ACPF conserve la continuité inhérente des données fonctionnelles. Même si les composantes principales semblent proches, l'ACPF capture les variations lisses et les dépendances locales que l'ACP discrète peut fragmenter ou lisser artificiellement.

### Robustesse aux choix de discrétisation

L'ACP classique dépend fortement du nombre et de la position des points d'échantillonnage. L'ACPF, travaillant directement sur l'espace fonctionnel, évite ces biais de discrétisation qui peuvent affecter l'interprétation même quand les résultats globaux paraissent similaires.

### Interprétation plus naturelle

Les composantes fonctionnelles ont une interprétation directe en termes de modes de variation des courbes (tendances, oscillations, pics), alors que les composantes discrètes peuvent mélanger ces aspects de façon moins intuitive.

### Capacité de prédiction et reconstruction

L'ACPF permet de reconstruire ou prédire des valeurs en tout point du domaine, pas seulement aux points observés. Cette propriété d'interpolation/extrapolation est cruciale pour de nombreuses applications.

### Gestion naturelle des données irrégulières

Quand les courbes sont observées à des instants différents ou avec des fréquences variables, l'ACPF gère naturellement cette hétérogénéité, contrairement à l'ACP qui nécessite un préalable d'interpolation.

## Enregistrement des résultats

Voici la liste des variables créées par la procédure.

<i>Variable</i>	<i>Contenu</i>
libvar	Libellés des variables
libobs	Libellés des observations
ptsmeasure	Points de mesure
gcv	Indices généralisés de validation croisée
rmse	Erreurs quadratiques moyennes
valpropres	Valeurs propres
fctpromoy	Fonction propre moyenne
fctpro	Fonctions propres
fctpromoins	Fonctions propres –
fctproplus	Fonctions propres +
estimes	Valeurs estimées
moyobservees	Valeurs observées moyennes
moyestimes	Valeurs estimées moyennes
résidus	Résidus
scores	Scores des observations

## Références

Documentation du package R 'fda'  
<https://cran.r-project.org/web/packages/fda/fda.pdf>

Ramsay J.O. and B.W. Silverman (2006) – « Functional data analysis (2<sup>ème</sup> édition). Springer »

Tuddenham et Snyder (1954) – « Physical growth of California boys and girls from birth to eighteen years »