

Sophie GOHIN
Christelle HERBIN
Marine JOSSET
Julien LEPAGE

ÉNERGIE MENSUELLE APPELÉE PAR LA CONSOMMATION ENTRE 1955 ET 1992

Groupe 8
Enseignant tuteur : M. Philippe CHABAULT
Année 2003 – 2004
IUT René Descartes (Paris 5)
Section STID (statistiques et traitement informatique des données)

TABLE DES MATIÈRES

RÉSUMÉ	- 1 -
INTRODUCTION.....	- 2 -
FICHE D'IDENTITÉ	3
INTRODUCTION DE LA PARTIE MATHÉMATIQUE DU PROJET	4
I. PARTIE MATHÉMATIQUE ET APPLICATION AU JEU DE DONNÉES.....	4
1. ÉLABORATION DES DIVERS MODELES	4
1. <i>Tendance</i>	4
2. <i>Modèle additif</i>	5
3. <i>Modèle additif logarithmique</i>	6
4. <i>Modèle multiplicatif</i>	7
5. <i>Modèle multiplicatif évolutif</i>	7
6. <i>Et aussi</i>	9
2. DETERMINATION DES ACCIDENTS.....	10
CONCLUSION.....	17
BIBLIOGRAPHIE.....	18
ANNEXES	19
GRAPHIQUES DES ACCIDENTS :	20

RÉSUMÉ

1. Fiche d'identité

- Synthèse de notre série de départ
- Représentation graphique de la moyenne et de l'écart-type par mois
- Représentation graphique de la moyenne et de l'écart-type par année
- Monotone décroissante de la série
- Auto corrélation de la série (25 décalages)

2. Partie mathématique et application au jeu de données

- Construction de modèles :
 - Moyennes mobiles
 - Moyennes mobiles centrées
 - Détermination de l'équation de la tendance
 - Méthode des MC
 - Calcul de \hat{a} et \hat{b} (respectivement 68.44 et 7293.69)
 - Méthode des 2 points moyens
 - Moyenne des 120 premières valeurs et moyenne des 120 autres.
 - Détermination de l'équation d'une droite de tendance de coefficients a et b (68.68 et 7264.33)
 - Construction d'un modèle additif (mais pas vraiment adapté)
 - CCL = 0.976
 - Construction d'un modèle additif logarithmique
 - CCL = 0.986
 - Construction d'un modèle multiplicatif
 - CCL = 0.989
 - Construction d'un modèle multiplicatif « évolutif »
 - CCL = 0.991
- Étude des accidents
 - Détails concernant la méthode utilisée pour faire ressortir les accidents à plus ou moins 1 écart-type et à plus ou moins 2 écarts-type)
 - Accidents remarquables :
 - Sous-consommation :
 - Mai 1968
 - Août 1975
 - Juillet, Septembre et Octobre 1984
 - Surconsommation :
 - Avril 1965
 - Novembre et Décembre 1980
 - Décembre 1981
 - Février et Mars 1984
 - Explications possibles
- Extension de l'étude

INTRODUCTION

Les étudiants de la section STID (statistiques et traitement informatique des données) de l'IUT René Descartes de Paris 5 ont eu un projet à réaliser au cours de l'année 2004. Les étudiants ont formé des groupes de trois ou quatre personnes, et un sujet pour environ trois groupes a été distribué.

Notre sujet porte sur l'énergie mensuelle appelée par la consommation entre janvier 1955 et décembre 1992. Cette série contient 456 données c'est pourquoi elle a été divisée en trois parties : 1955-1974, 1965-1984, 1975-1992. Notre groupe a été chargé d'étudier la période 1965-1984. Le but de notre projet est d'une part de pouvoir modéliser notre série de données afin de pouvoir prévoir la consommation future, et d'autre part de repérer les dates durant lesquelles une surconsommation, ou à l'inverse une sous-consommation d'électricité, a eu lieu.

Notre rapport a été découpé en deux grandes parties :

- La première est une analyse descriptive des données. Nous l'avons rendue dans un premier rapport : le rapport de Statistiques Descriptives. Nous l'avons résumée en ce que nous appelons une fiche d'identité.
- La seconde s'est articulée autour de calculs et de recherches afin de construire un modèle fiable et de faire ressortir les accidents. Elle représente la partie mathématique de notre projet ainsi que son application au jeu de données.

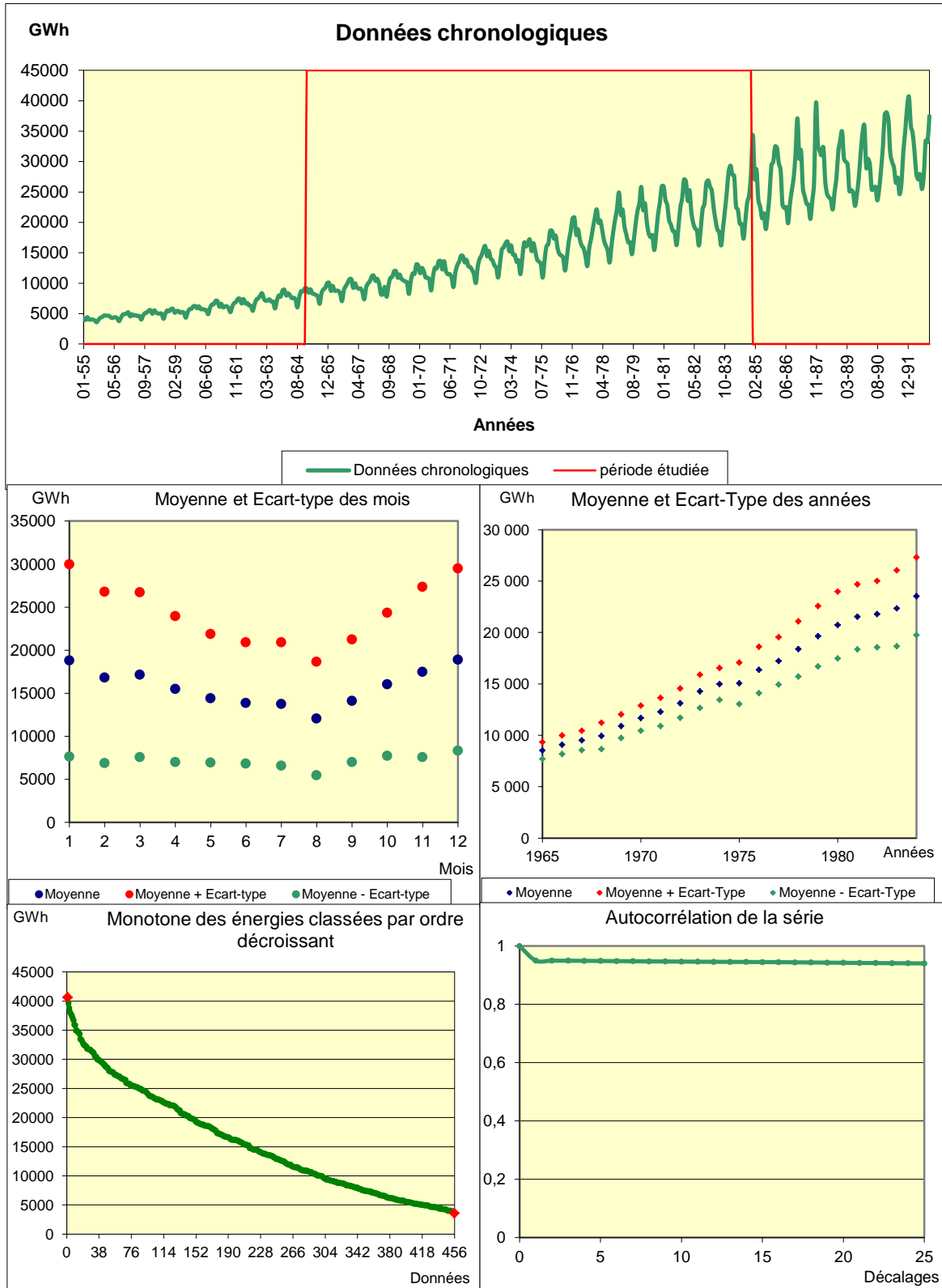
Le plan de notre seconde partie se compose de la manière suivante :

- Dans un premier temps, nous vous expliquerons tous les calculs effectués avec les chiffres de notre série de départ pour déterminer un modèle fiable.
- Nous vous ferons part des dates accidentelles, repérées grâce à notre modèle.
- Nous tenterons d'expliquer certaines de ces dates en formulant des hypothèses cohérentes.
- Nous finirons par une extension de l'étude.

FICHE D'IDENTITÉ

Nombre de données : 38 ans * 12 mois = 456 données

Moyenne : 15 714,77 GWh — Écart-type : 8 938,70 GWh



Introduction de la partie mathématique du projet

I. Partie mathématique et application au jeu de données

1. Élaboration des divers modèles

L'objectif du projet est de pouvoir réaliser des prévisions de la consommation. Pour ce faire, il nous faut établir des modèles mathématiques ajustant au mieux la série de départ (SdD). Nous avons, dans un premier temps, représenté graphiquement cette série. Il apparaissait alors clairement qu'elle est constituée d'une tendance, d'une saisonnalité et d'accidents.

Ne pouvant modéliser les accidents, nous allons tenter de les réduire au maximum. Pour cela, deux indices nous permettront de mesurer la fiabilité de chacun des modèles que nous établirons :

- Le coefficient de corrélation linéaire (CCL).
- L'étude des résidus entre le modèle et la série de départ.

1. Tendance

Grâce au cours de séries chronologiques, nous avons pu tracer la série des moyennes mobiles centrées (MMC) de notre série. Celle-ci nous permet de mettre en évidence l'allure de la tendance en supprimant la saisonnalité et en atténuant les accidents. Nous avons calculé la série des moyennes mobiles simples d'ordre 12 (La saisonnalité étant évidemment d'ordre 12 (Un an)), notée MM(12).

Définition : La série des moyennes mobiles d'ordre 12 (MM(12)) est la série des moyennes de 12 observations consécutives et elle prend des valeurs aux dates moyennes \bar{t}_i .

$$\bar{t}_i = \frac{t_i + t_{i+1} + \dots + t_{i+12}}{12}$$
$$MM(12)_{\bar{t}_i} = \frac{y_{t_i} + y_{t_{i+1}} + \dots + y_{t_{i+12}}}{12}$$

Notre tendance ayant une allure linéaire, nous l'avons ajustée par la droite des moindres carrés (DMC). Son équation est du type :

$$\hat{y}_t = \hat{a}t + \hat{b} \quad \text{où} \quad \hat{a} = \frac{Cov(y_t, t)}{Var(t)}$$
$$\hat{b} = \bar{y}_t - \hat{a}\bar{t}$$

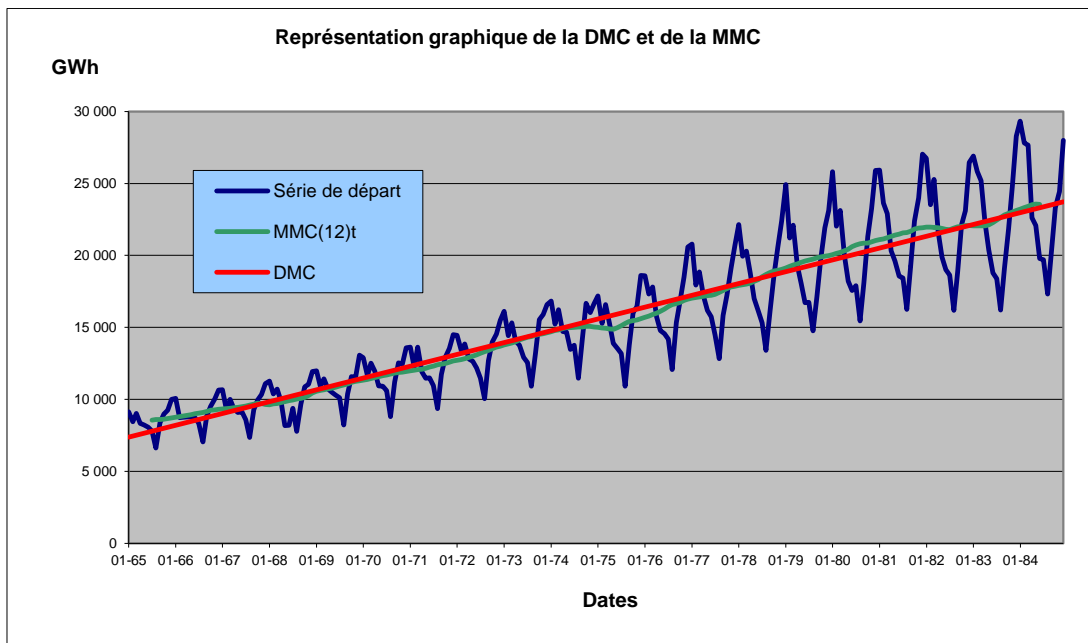
\bar{y}_t : Moyenne des données de la série de départ

\bar{t} : Moyenne des dates

La série des moyennes mobiles centrées est une série de moyennes mobiles d'ordre 2 faite sur la série des moyennes mobiles d'ordre 12. Elle sert simplement à faire concorder les dates.

(Voir documents en annexes sur \hat{a} et \hat{b}).

Graphique représentant la série de départ ainsi que la série des moyennes mobiles centrées et la droite des moindres carrés :



2. Modèle additif

Le modèle additif est composé de la tendance à laquelle nous ajoutons la saisonnalité :

$$y_t = f_t + s_t$$

t : Date

y_t : Série de départ

f_t : Tendance (ici DMC)

s_t : Saisonnalité

En faisant l'application numérique, nous trouvons pour la DMC : $f_t = 68,44 t + 7293,69$.

Les 12 coefficients saisonniers ont été calculés en faisant la moyenne des saisons de Janvier, celle des saisons de Février, etc.

Nous utilisons le coefficient de corrélation linéaire (CCL) afin de comparer notre modèle à celui de la série. Celui-ci est défini par :

$$CCL(y_t, t) = \frac{Cov(y_t, t)}{\sqrt{Var(t) \times Var(y_t)}}$$

Le calcul du CCL donne un nombre compris entre 0 et 1. Plus il est proche de 1 et plus notre modèle est fiable.

Nous notons la série suivant le modèle additif y_A .

Bilan du modèle :

- $CCL(y_A, t) : 0,918$
- $CCL(y_A, y_t) : 0,976$

- Moyenne des résidus à la série de départ : 0,000
- Écart-type des résidus à la série de départ : 1160

En conclusion, bien qu'étant très bon, ce modèle ne correspond pas vraiment à la courbe puisque l'écart entre les maximums et les minimums reste constant au fil du temps alors que sur la série de départ il augmente. Aussi avons-nous pensé qu'un modèle additif logarithmique serait plus adapté.

3. Modèle additif logarithmique

Le modèle logarithmique se construit comme le modèle additif, mais non pas sur la série de départ mais sur son logarithme.

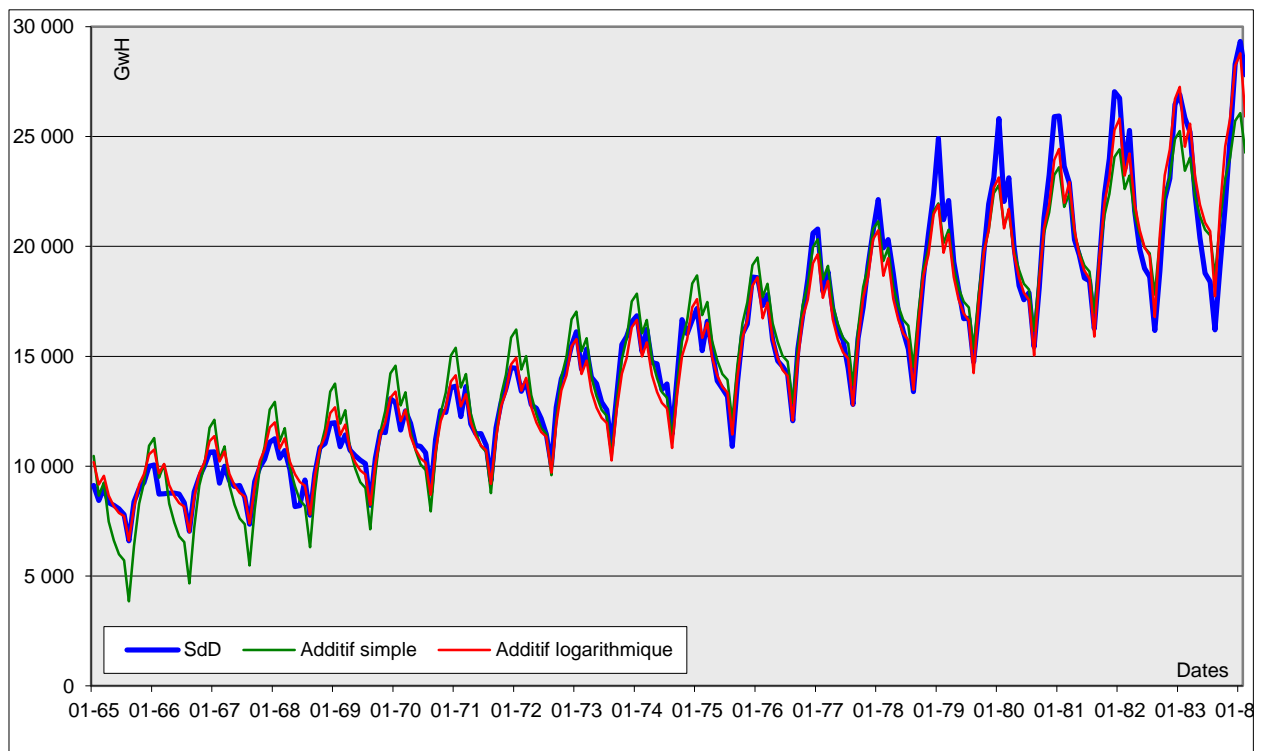
Pour revenir aux mêmes unités que la série, il suffit donc de prendre l'exponentielle des valeurs données par ce modèle.

Nous notons la série suivant le modèle additif logarithmique y_{AL} .

Bilan du modèle :

- $CCL(y_{AL}, t)$: 0,914
- $CCL(y_{AL}, y_t)$: 0,986
- Moyenne des résidus à la série de départ : -20,336
- Écart-type des résidus à la série de départ : 881

Graphique représentant la série de départ ainsi que les modèles additif simple et additif logarithmique :



Commentaire du graphique :

Le modèle additif conserve les mêmes amplitudes tandis que celles de notre série ne cesse d'accroître. Ainsi pour le début de la série, les variations sont trop importantes et pour la fin de la série, elles ne le sont pas assez.

Le modèle additif logarithmique se rapproche déjà plus de la série initiale de par l'augmentation progressive de ses variations. Néanmoins, il ne correspond pas tout à fait à notre série car les mois de janvier et février ne sont pas correctement modélisés.

Conclusion :

Ce modèle est encore meilleur que le précédent et réduit les résidus. De plus, l'allure de ce modèle correspond bien plus à l'allure de la série. Mais nous avons voulu également essayer un modèle multiplicatif qui pourrait convenir à l'allure de la série.

4. Modèle multiplicatif

Le modèle multiplicatif est très similaire au modèle additif à ceci près que la saisonnalité n'est pas ajoutée mais multipliée à la tendance : $y_t = f_t \times s_t$

Ceci implique bien entendu une technique de calcul différente pour les coefficients saisonniers (ils ne sont plus calculés à partir des valeurs de départ auxquelles on retranche la tendance, mais à partir des valeurs de départ que l'on divise par la tendance).

Nous notons la série suivant un modèle multiplicatif y_M .

Bilan du modèle :

- $CCL(y_M, t) : 0,919$
- $CCL(y_M, y_t) : 0,989$
- Moyenne des résidus à la série de départ : 25,241
- Écart-type des résidus à la série de départ : 772

Ce modèle est légèrement meilleur que le précédent et réduit encore plus l'écart-type des résidus. Et là encore, l'allure du modèle correspond particulièrement bien à celle de la série.

Cependant, nous avons eu une idée afin d'aller plus loin. Dans le modèle multiplicatif, nous utilisons comme tendance la droite des moindres carrés de formule $\hat{y}_t = \hat{a}t + \hat{b}$.

Seulement, \hat{a} et \hat{b} sont ici des constantes. L'idée que nous avons eue est de calculer \hat{a} et \hat{b} non pas sur les 240 valeurs, mais sur la 1^{ère}, les 2 premières, les 3 premières, et ce, jusqu'aux 240 valeurs. Nous obtenons ainsi une série de \hat{a} et \hat{b} « évolutifs ».

5. Modèle multiplicatif évolutif

En traçant la courbe des \hat{a} et celle des \hat{b} , nous constatons que les premières valeurs sont assez « chaotiques ». Aussi nous avons décidé de considérer la série à partir de $t = 61$ (janvier 1970) pour \hat{a} et à partir de $t = 30$ (juin 1967) pour \hat{b} .

Nous avons pu alors établir un modèle pour \hat{a} (modèle multiplicatif) et un pour \hat{b} (modèle exponentiel).

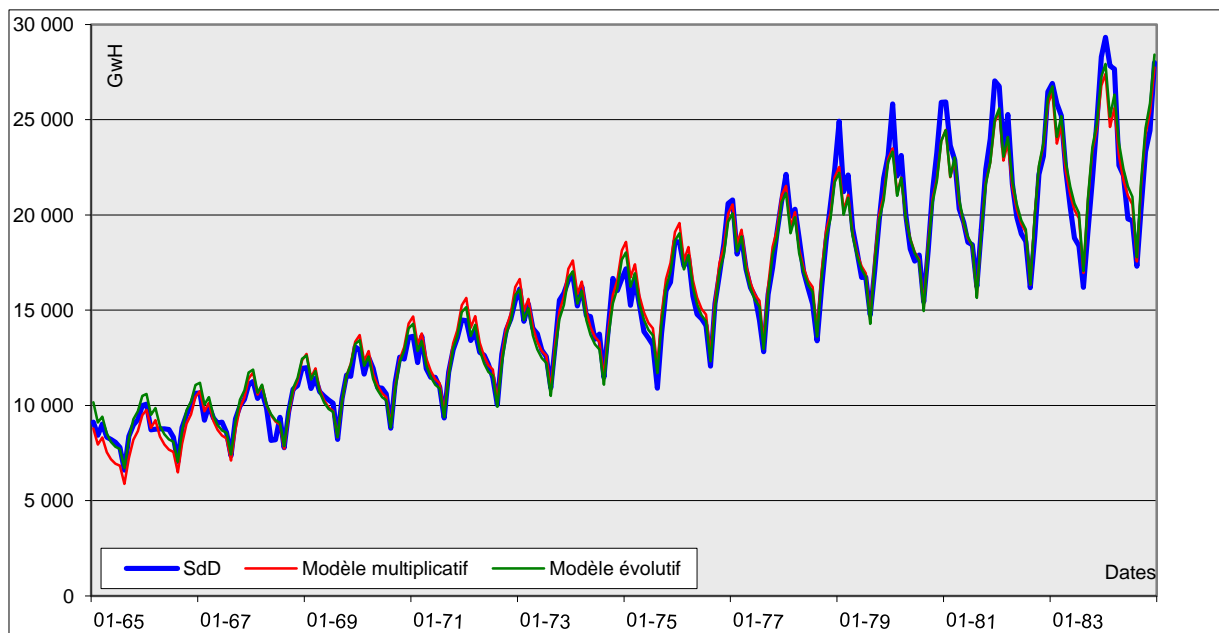
Nous notons la série suivant le modèle multiplicatif évolutif y_{ME} .

Bilan du modèle :

- $CCL(y_{ME}, t) : 0,914$
- $CCL(y_{ME}, y_t) : 0,991$
- Moyenne des résidus à la série de départ : 0,000
- Écart-type des résidus à la série de départ : 724

Le modèle obtenu est donc, encore une fois, meilleur que le précédent, réduisant une fois de plus l'écart-type des résidus.

Graphique représentant la série de départ ainsi que les modèles multiplicatif simple et multiplicatif évolutif :



Commentaire du graphique :

On constate que les deux modèles se ressemblent beaucoup mais c'est le début de la série qui fait la différence. En effet, juste avant les années 1970, la courbe du modèle multiplicatif est en dessous de la courbe du modèle modifié, ce qui permet au nouveau modèle d'être plus proche de la série ; mais très vite, les deux courbes se rejoignent pour finalement être presque confondues.

6. Et aussi ...

Une autre idée que nous avons eue consistait à modéliser de la même manière les coefficients saisonniers, mais le modèle ainsi obtenu n'obtenait pas de meilleurs résultats.

En superposant chacun des 4 modèles obtenus (additif, logarithmique, multiplicatif et multiplicatif « évolué ») sur un même graphique. On constate alors clairement que certains modèles ajustent mieux la courbe que d'autres par moment.

Aussi fûmes-nous tentés de combiner ces quatre modèles de façon à en obtenir un cinquième encore plus performant.

Nous avons donc créé une série constituée de 1, 2, 3 et 4 selon quel modèle marchait mieux que les autres. Mais le coefficient de corrélation entre cette nouvelle série et le temps est seulement de 0,065...

Ce qui interdit du fait tout modèle en fonction du temps.

En revanche, nous avons supposés que le fait qu'un modèle marche mieux qu'un autre pouvait dépendre du mois. Nous avons donc déterminé 12 coefficients saisonniers.

Mais même avec cette méthode le 5^e modèle restait moins bon que le 4^e.

Nous avons alors essayé de faire une moyenne pondérée des 4 modèles en fonction des effectifs sur la série globale puis des effectifs sur chaque mois. Mais aucune de ces méthodes ne nous permirent de dépasser le 4^e modèle.

2. Détermination des accidents

Dans cette partie, nous allons expliquer comment nous avons pu déterminer les « accidents », c'est-à-dire les dates auxquelles on a relevé une hausse ou une baisse notable de la consommation d'électricité.

Nous avons utilisé le modèle multiplicatif.

- On a fait la différence entre les données de notre série de départ et les données de la série composée par un modèle multiplicatif : On a ainsi obtenu les résidus :

$$e_t = y_t - y_M$$

- Nous avons retranscrit ces données dans le tableau ci-dessous :

Année	J	F	M	A	M	J	J	A	S	O	N	D
1965	341	501	704	788	1040	1124	970	736	1146	780	607	509
1966	290	-91	-471	399	828	1056	785	545	851	466	487	187
1967	-89	-467	-138	145	359	705	306	252	554	73	-85	-316
1968	-472	-205	-330	-153	-1322	-959	382	50	177	136	-257	-431
1969	-736	-569	-518	-110	224	379	403	-117	127	18	-650	-267
1970	-792	-686	-347	301	-88	266	166	-148	215	116	-617	-716
1971	-1036	-956	-145	-546	-334	99	-203	-218	-1	-340	-400	-773
1972	-1194	-672	-837	-496	66	50	-424	-148	191	-130	-292	-732
1973	-518	-543	-267	-56	413	50	-39	123	-32	590	198	-604
1974	-771	-622	-273	-210	547	-121	422	64	107	889	-558	-1488
1975	-1416	-1454	-821	-456	-1004	-802	-871	-1121	-963	-604	-1009	-487
1976	-995	-278	-515	-816	-871	-531	-578	-576	-159	-594	155	537
1977	240	-522	-372	-228	-254	-89	-1030	-443	-427	-1081	-124	-313
1978	602	584	165	562	-189	-410	-890	-479	-812	-548	453	402
1979	2406	991	1044	263	73	-584	-206	264	-663	-281	923	187
1980	2334	938	1167	169	-533	-469	240	334	-377	456	1376	2002
1981	1462	1658	21	-348	85	-213	47	536	-6	693	1228	2173
1982	1301	665	1496	142	-420	-516	-474	-157	-1114	-405	-551	635
1983	474	2104	508	41	-678	-1480	-1440	-745	-1653	-1525	345	1507
1984	1918	3203	2065	-520	237	-1206	-853	-257	-1215	-902	-980	257

- Nous avons calculé la moyenne et l'écart type pour chaque mois :

Mois	J	F	M	A	M	J	J	A	S	O	N	D
Moy	167	179	107	-56	-91	-183	-164	-75	-203	-110	12	113
σ	1 157	1 124	781	400	591	671	629	450	689	638	670	920

Ce qui nous intéresse, c'est de regarder quelles ont été les dates qui ont connu une forte ou faible consommation, du moins, par rapport à ce à quoi on pouvait s'attendre.

Pour ce faire, nous avons déterminé des limites, à savoir Moyenne + Ecart type pour la supérieure, et Moyenne – Ecart type pour l'inférieure.

De plus, afin de distinguer les « gros » accidents, nous avons fait la même chose à plus ou moins 1,5 écart type et 2 écarts type. Cependant, nous n'avons gardé que les données à 1 et à 2 écarts type.

- Nous avons voulu faire ressortir les dates qui dépassaient les limites de la moyenne plus ou moins un écart type ainsi que celles qui dépassaient celles de la moyenne plus ou moins deux écart type.
 1. Nous avons donc calculé, pour chaque mois :
 - Moyenne + 1 écart type
 - Moyenne – 1 écart type
 - Moyenne + 2 écarts type
 - Moyenne – 2 écarts type
 2. Nous avons ensuite créé quatre tableaux qui faisaient correspondre, à chaque date :
 - 1^{er} tableau : Résidu à la date t – (Moyenne + 1 écart type)
 - 2^e tableau : Résidu à la date t – (Moyenne – 1 écart type)
 - 3^e tableau : Résidu à la date t – (Moyenne + 2 écarts-types)
 - 4^e tableau : Résidu à la date t – (Moyenne – 2 écart type)

+ 1	J	F	M	A	M	J	J	A	S	O	N	D
1965	-984	-802	-183	444	541	636	505	361	659	251	-76	-524
1966	-1035	-1394	-1358	55	328	568	320	171	365	-63	-195	-846
1967	-1414	-1770	-1025	-199	-141	217	-159	-122	68	-455	-767	-1349
1968	-1796	-1508	-1218	-497	-1822	-1447	-83	-325	-309	-393	-939	-1464
1969	-2061	-1872	-1406	-454	-276	-109	-62	-491	-360	-510	-1332	-1300
1970	-2117	-1989	-1235	-43	-587	-222	-299	-522	-271	-413	-1300	-1749
1971	-2360	-2259	-1033	-890	-834	-388	-668	-593	-487	-869	-1082	-1806
1972	-2519	-1975	-1724	-839	-434	-438	-889	-522	-295	-658	-974	-1765
1973	-1843	-1846	-1154	-400	-86	-438	-504	-251	-518	61	-485	-1637
1974	-2096	-1925	-1160	-554	47	-609	-43	-311	-379	361	-1240	-2521
1975	-2740	-2757	-1708	-800	-1504	-1290	-1336	-1495	-1449	-1133	-1691	-1520
1976	-2320	-1581	-1403	-1160	-1371	-1019	-1043	-950	-645	-1123	-527	-496
1977	-1085	-1825	-1260	-572	-754	-577	-1495	-818	-913	-1609	-806	-1346
1978	-722	-719	-723	218	-688	-898	-1355	-853	-1299	-1077	-230	-631
1979	1081	-312	156	-81	-427	-1072	-671	-110	-1149	-810	241	-846
1980	1009	-365	280	-175	-1033	-957	-225	-41	-863	-73	694	969
1981	137	355	-866	-692	-414	-701	-418	162	-492	164	545	1140
1982	-23	-638	609	-202	-920	-1004	-939	-531	-1601	-933	-1233	-398
1983	-851	801	-380	-303	-1178	-1968	-1905	-1120	-2139	-2054	-337	474
1984	593	1900	1177	-864	-262	-1694	-1318	-631	-1701	-1431	-1662	-776

- 1	J	F	M	A	M	J	J	A	S	O	N	D
1965	1331	1446	1378	1245	1722	1977	1764	1261	2037	1528	1264	1315
1966	1280	854	203	856	1510	1909	1579	1070	1743	1214	1145	993
1967	900	478	536	602	1041	1558	1100	777	1446	821	572	490
1968	518	740	344	304	-641	-106	1175	575	1068	884	401	375
1969	254	376	156	347	906	1232	1197	408	1018	766	7	539
1970	198	259	327	757	594	1119	960	377	1107	864	40	91
1971	-46	-11	528	-89	347	953	591	307	891	408	258	33
1972	-204	273	-163	-39	748	903	369	377	1083	618	365	74
1973	472	402	407	401	1095	903	755	648	859	1338	855	202
1974	218	323	401	247	1229	732	1216	589	999	1637	99	-682
1975	-426	-509	-147	1	-322	51	-77	-596	-71	144	-352	319
1976	-5	667	159	-359	-190	322	215	-51	733	154	813	1343
1977	1230	423	302	229	428	764	-236	82	464	-333	533	493
1978	1592	1529	839	1019	493	443	-96	46	79	200	1110	1208
1979	3396	1936	1717	720	754	269	588	789	229	467	1581	993
1980	3324	1883	1841	626	149	384	1033	859	515	1204	2033	2809
1981	2451	2603	695	109	767	640	841	1061	885	1441	1885	2979
1982	2291	1610	2170	599	262	337	320	368	-223	343	106	1441
1983	1464	3049	1182	497	4	-627	-646	-220	-761	-777	1002	2313
1984	2908	4148	2739	-63	919	-353	-60	268	-323	-154	-322	1063

+ 2	J	F	M	A	M	J	J	A	S	O	N	D
1965	-2141	-1926	-964	44	-50	-35	-124	-88	-29	-387	-745	-1444
1966	-2192	-2518	-2139	-346	-263	-103	-309	-279	-324	-701	-865	-1765
1967	-2571	-2894	-1806	-599	-731	-453	-788	-572	-621	-1094	-1437	-2269
1968	-2954	-2632	-1998	-898	-2413	-2117	-713	-774	-998	-1031	-1609	-2384
1969	-3218	-2996	-2186	-854	-867	-779	-691	-941	-1048	-1149	-2002	-2219
1970	-3274	-3113	-2016	-444	-1178	-892	-928	-972	-960	-1051	-1969	-2668
1971	-3518	-3383	-1814	-1290	-1425	-1059	-1297	-1042	-1176	-1507	-1752	-2726
1972	-3676	-3099	-2505	-1240	-1024	-1109	-1519	-972	-984	-1297	-1644	-2684
1973	-3000	-2970	-1935	-800	-677	-1109	-1133	-701	-1207	-577	-1154	-2556
1974	-3253	-3049	-1941	-954	-544	-1280	-672	-760	-1068	-278	-1910	-3441
1975	-3898	-3881	-2489	-1201	-2094	-1961	-1965	-1945	-2138	-1771	-2361	-2440
1976	-3477	-2705	-2183	-1560	-1962	-1689	-1673	-1400	-1334	-1761	-1197	-1415
1977	-2242	-2949	-2040	-973	-1345	-1247	-2124	-1267	-1602	-2248	-1476	-2266
1978	-1879	-1843	-1504	-182	-1279	-1568	-1984	-1303	-1988	-1715	-899	-1551
1979	-76	-1436	-625	-482	-1018	-1742	-1300	-560	-1838	-1448	-429	-1765
1980	-148	-1489	-501	-575	-1624	-1627	-855	-490	-1552	-712	24	50
1981	-1020	-769	-1647	-1093	-1005	-1372	-1047	-288	-1181	-474	-124	220
1982	-1181	-1762	-172	-602	-1511	-1675	-1568	-981	-2289	-1572	-1903	-1318
1983	-2008	-323	-1160	-704	-1768	-2639	-2534	-1569	-2828	-2692	-1007	-445
1984	-564	776	397	-1264	-853	-2365	-1948	-1081	-2390	-2069	-2332	-1696

- 2	J	F	M	A	M	J	J	A	S	O	N	D
1965	2488	2570	2159	1646	2313	2647	2393	1710	2726	2166	1934	2235
1966	2437	1978	984	1256	2100	2580	2208	1520	2432	1852	1814	1913
1967	2058	1602	1317	1003	1632	2229	1729	1227	2135	1460	1242	1409
1968	1675	1864	1125	704	-50	565	1805	1024	1757	1522	1070	1295
1969	1411	1500	936	747	1497	1903	1826	858	1707	1405	677	1459
1970	1355	1383	1107	1158	1185	1790	1589	827	1796	1502	710	1010
1971	1111	1113	1309	311	938	1623	1220	756	1580	1046	927	952
1972	953	1397	618	362	1339	1573	999	827	1771	1257	1035	994
1973	1629	1526	1188	801	1686	1573	1384	1098	1548	1976	1525	1122
1974	1376	1447	1182	648	1819	1402	1845	1038	1688	2276	769	237
1975	731	615	634	401	269	721	552	-146	618	782	318	1239
1976	1152	1791	940	42	401	993	845	399	1421	792	1482	2263
1977	2387	1547	1082	629	1018	1435	393	531	1153	306	1203	1412
1978	2750	2653	1619	1420	1084	1114	533	496	768	838	1780	2127
1979	4553	3060	2498	1120	1345	940	1217	1239	918	1106	2250	1913
1980	4481	3007	2622	1027	740	1055	1663	1308	1203	1842	2703	3728
1981	3609	3727	1476	509	1358	1310	1470	1511	1574	2079	2555	3898
1982	3448	2734	2951	999	852	1007	949	818	466	982	776	2361
1983	2621	4173	1963	898	595	43	-17	229	-72	-139	1672	3233
1984	4065	5272	3520	337	1510	317	570	718	366	484	347	1982

Remarques : Pour les premier et troisième tableaux, les données qui apparaissent positives ont été surlignées en bleu. En effet, ces données représentent les surconsommations d'électricité.

Pour les deuxième et quatrième tableaux, les données qui apparaissent négatives ont été surlignées en bleu. Elles représentent les sous-consommations d'électricité.

Commentaire :

On remarque qu'à moins 1 écart-type, ce sont surtout des périodes qui se distinguent. En effet on observe que les données observées pour -1 écart-type marquent surtout des années alors que les données à plus 1 écart-type sont plutôt regroupées en « tas », ce sont plutôt des mois qui ressortent. Les deux premières années, c'est entre Avril et Octobre (plutôt l'été), alors que les six dernières, c'est entre Novembre et Mars (plutôt l'hiver).

3. Afin de faciliter la lecture des accidents, nous avons regroupé les accidents au-dessus et en dessous d'un écart type dans un même tableau. Nous avons fait de même pour ceux au-dessus et en dessous de deux écarts type. Nous allons d'ailleurs retenir les 11 accidents de ce tableau car ce sont les plus importants vu qu'ils subsistent après 2 écarts type (Tableau ci-dessous).

± 2	J	F	M	A	M	J	J	A	S	O	N	D
1965	5327	4950	2081	-72	116	92	298	151	80	839	1441	3226
1966	5342	4981	2104	434	552	264	682	424	787	1299	1569	3377
1967	5291	4636	2378	601	1193	1011	1362	702	1325	1596	1785	3197
1968	4948	4906	2247	632	-120	1196	1286	793	1754	1570	1722	3086
1969	4541	4494	2047	639	1297	1483	1262	807	1790	1614	1356	3238
1970	4436	4305	2232	514	1396	1597	1475	803	1723	1579	1398	2695
1971	3910	3765	2374	402	1337	1719	1583	788	1857	1577	1624	2596
1972	3504	4330	1548	449	1371	1745	1516	804	1743	1630	1702	2668
1973	4887	4532	2299	641	1142	1744	1568	769	1869	1141	1760	2868
1974	4475	4412	2294	618	989	1795	1240	789	1802	632	1469	816
1975	2851	2388	1577	482	563	1415	1085	-284	1320	1385	751	3022
1976	4006	4844	2051	65	787	1677	1413	558	1896	1395	1774	3203
1977	5352	4562	2209	612	1369	1790	834	673	1848	687	1776	3200
1978	5168	4889	2435	259	1387	1747	1058	646	1526	1438	1600	3299
1979	345	4394	1561	540	1369	1637	1582	694	1687	1601	965	3377
1980	663	4477	1313	591	1201	1717	1421	642	1868	1311	-65	-186
1981	3682	2865	2431	556	1365	1797	1539	435	1860	986	318	-858
1982	4071	4817	508	602	1288	1687	1488	802	1067	1543	1477	3110
1983	5263	1348	2277	632	1052	114	-42	360	-204	-374	1684	1439
1984	2293	-4093	-1396	427	1288	751	1109	776	874	1002	810	3362

Conclusion :

Les différentes dates d'accidents retenus sont :

- avril 1965 (consommation élevée)
- mai 1968 (consommation faible)
- août 1975 (faible)
- novembre et décembre 1980 (élevée)
- décembre 1981 (élevée)
- juillet 1983 (faible)
- septembre et octobre 1983 (faible)
- février et mars 1984 (élevée)

Malheureusement pour nous les données concernant ces périodes sont rares et les possibilités d'explication de ses données ont donc été un peu difficiles. C'est pourquoi nous ne tenterons d'expliquer que trois d'entre elles.

Explications possibles :

L'accident de consommation de **mai 1968** est le plus facile à tenter d'expliquer. Nul n'ignore en effets les événements de cette année. Après de nombreuses contestations étudiantes ce fut alors l'amorce de la deuxième phase, sociale (13-27 mai), du mouvement. Divers syndicats trouvèrent un terrain d'entente avec les étudiants en organisant, pour la journée du 13 mai 1968, une grande manifestation contre la répression policière, assortie d'une grève générale. Durant plus de cinq heures, quelque 200 000 ouvriers, étudiants, dirigeants syndicaux et politiques, (François Mitterrand, Guy Mollet, Pierre Mendès France) défilèrent à travers les rues de Paris. Le 14 mai, les grèves avec occupation d'usines

s'étendirent en province et paralysèrent la vie économique du pays. La France comptait 9 millions de grévistes.

Or environ un tiers de la série de donnée concerne les grandes entreprises ouvrières, ces grèves et surtout l'immobilisation totale des entreprises expliqueraient la chute importante de la consommation d'électricité. Mais cela n'est qu'une hypothèse, en effet des statisticiens professionnels étudient encore cette période.

Le quadruplement des prix du pétrole décidé par l'Organisation des pays producteurs de pétrole (OPEP), a provoqué de nombreux déséquilibres dans les pays occidentaux, qui se sont manifestés notamment par une hausse des taux d'inflation, une aggravation des déficits commerciaux liés à la facture pétrolière, un ralentissement de la croissance imposé par la hausse des coûts de production et des déficits extérieurs, et par la montée du chômage. Après 1973, on a enregistré une réduction des taux de croissance (récession), et non une baisse de la production (dépression). Ces explications peuvent montrer que la première crise pétrolière de 1975 peut être une des causes de la baisse de la consommation d'électricité en **1975**.

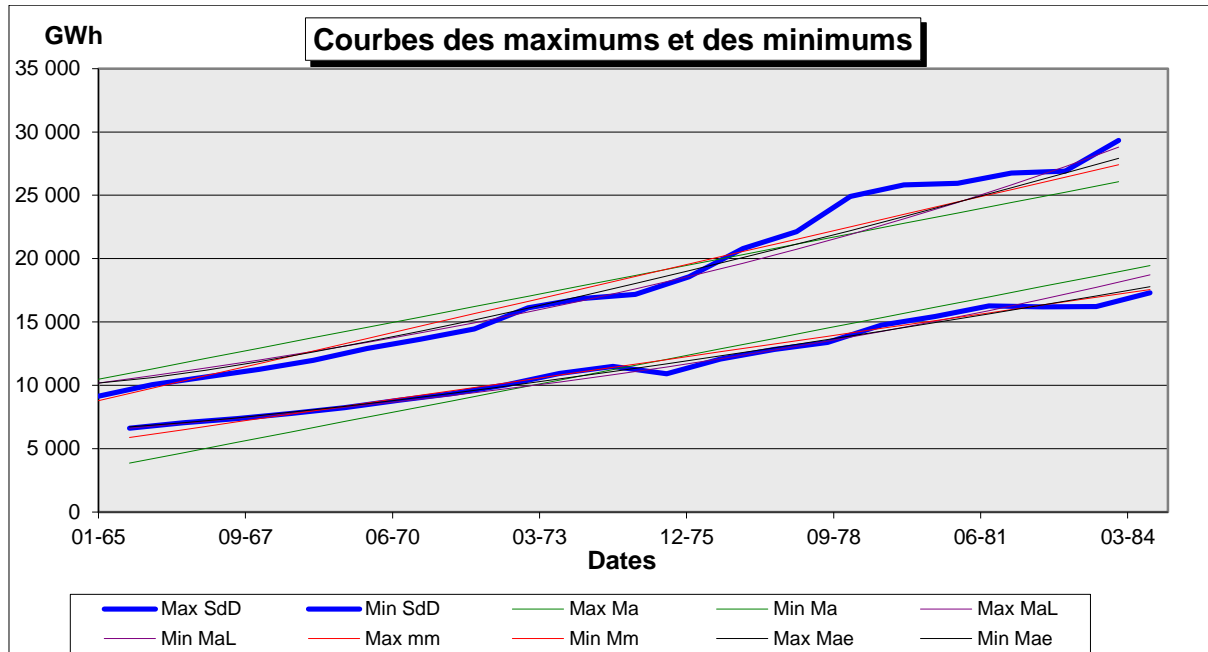
A la fin de l'année **1980** la hausse des prix peut peut-être s'expliquer par la recrudescence de la consommation d'électricité face à la hausse du prix du pétrole. La révolution iranienne puis la guerre entre l'Iran et l'Iraq furent des événements influençant fortement la hausse vertigineuse des prix du pétrole du fait de la réduction considérable des exportations de ces pays : de 14 dollars américains/baril en 1978 à 35 dollars américains/baril en 1981. Ces prix élevés permirent à la production hors moyen orient de progresser ainsi que de réduire la consommation du fait de politiques de lutte contre le gaspillage d'énergie (meilleure isolation des maisons, procédés industriels moins coûteux en énergie et parc automobile se augmentant plus lentement). Mais cette lutte contre le gaspillage d'énergie mis un certain temps à se mettre en place ce qui explique éventuellement la hausse de la consommation d'énergie à ces dates.

II. Extension de l'étude

Nous avons également fait le graphique représentant les maximums et minimums de :

- La série de départ (SdD)
- Le modèle additif (Ma)
- Le modèle additif logarithmique (MaL)
- Le modèle multiplicatif (Mm)
- Le modèle multiplicatif évolutif (Mae)

Nous avons considéré que les minimums correspondaient au mois d'août de chaque année et que les maximums étaient au mois de janvier de chaque année.



Commentaire :

Sur ce graphique sont représentés les maximums et les minimums de la série initiale, du modèle additif, du modèle additif logarithmique, du modèle multiplicatif, et du modèle multiplicatif « modifié ».

On constate que le minimum et le maximum du modèle additif ne correspondent pas du tout à ceux de la série étudiée. Aucune courbe ne modélise correctement celle des maximums de la série car la consommation d'énergie était beaucoup plus élevée que les années précédentes pour les mois de janvier de 1978 à 1981.

La courbe des minimums de la série subit moins de variations que les maximums. Elle est donc plus simple à modéliser. À part le modèle additif qui ne convient, les trois autres modélisent relativement bien la courbe de la série. Il n'y a que sur la fin du jeu de données que l'on remarque que le modèle additif logarithmique n'est pas si bien adapté.

Suggestion :

Une idée, à la vue de ce graphique, a fait surface. En effet, en observant l'évolution des minimums et des maximums de la série de départ, nous distinguons deux « périodes ». Jusqu'en 1978 à peu près, la série évolue à peu près linéairement, dans un intervalle relativement constant. Cependant, à partir de cette date, la montée rapide de la courbe des maximums tend plutôt à faire évoluer la série en forme de cône. On aurait donc pu essayer de modéliser la série en combinant un modèle additif avec un modèle multiplicatif.

CONCLUSION

BIBLIOGRAPHIE

Notre bibliographie est très modeste dans la mesure où nous n'avons pas eu recours à beaucoup de supports.

Livres, Encyclopédies :

- Deux livres que M. CHABAULT nous a prêtés concernant l'énergie mensuelle appelée par la consommation sous toutes ses formes. Nous avons complètement oublié de prendre leurs références.

Supports multimédias :

- Microsoft Encarta 1998
- Logiciels Excel et Word

Sites Internet :

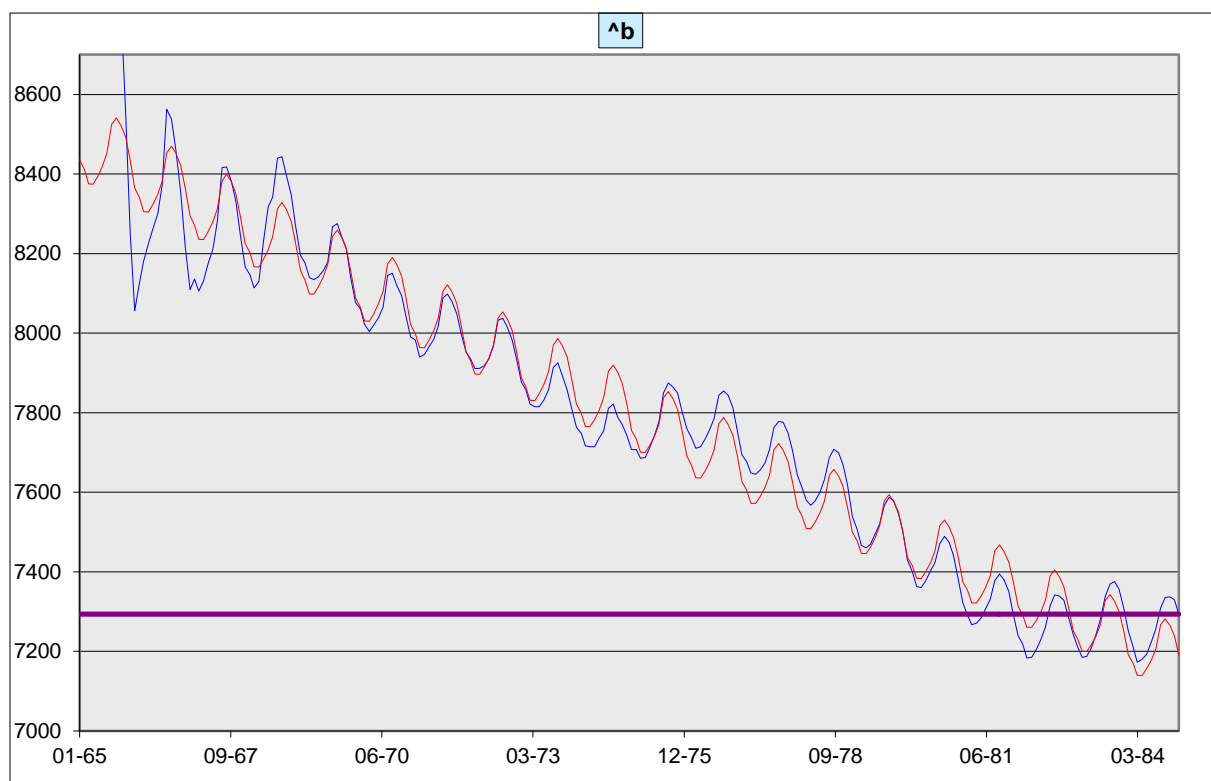
- <http://r0.unctad.org/infocomm/francais/indexfr.htm>
- <http://www.insee.fr>

Cours de STID :

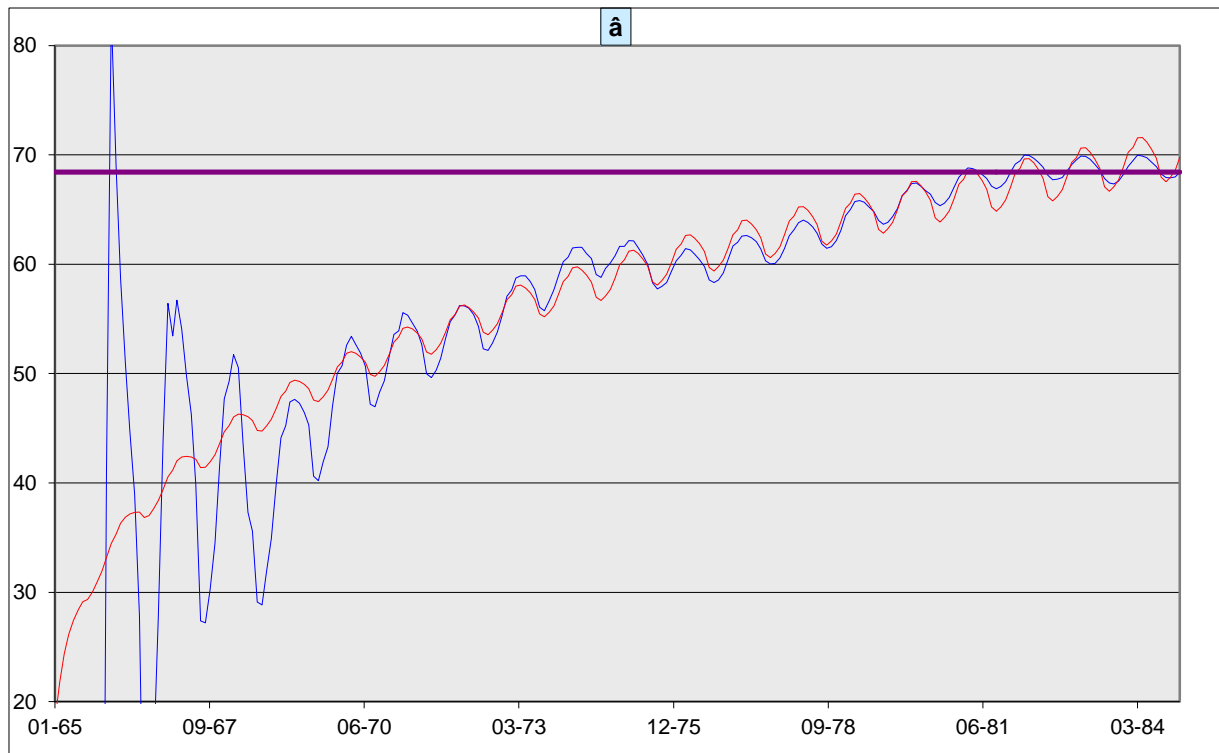
- Séries Chronologiques de Mme BRUNEL
- Statistiques descriptives de M. LEVY

Nous avons regardé plusieurs sites Internet, notamment pour tenter d'expliquer certaines hausses ou baisses de la consommation d'électricité mais n'avons pas relevé toutes les sources.

ANNEXES



La courbe **mauve** représente \hat{b} , calculé pour la droite des moindres carrés de la série. Il est bien sûr constant. Il a également servi à déterminer un modèle multiplicatif relativement proche de notre jeu de données. Mais le coefficient de corrélation linéaire n'étant que de 0.989, nous avons décidé de recalculer \hat{b} avec la première valeur, puis avec les deux premières valeurs, puis les trois premières, ... et ce jusqu'aux 240 valeurs que contient notre jeu de données. Ce nouveau \hat{b} est représenté par la courbe **bleue**. Nous avons ensuite tenté de le modéliser par un modèle multiplicatif. Nous avons obtenu la courbe **rouge**. Les courbes rouge et bleue sont relativement proches l'une de l'autre. Cependant, la bleue est une amplitude plus élevée que la rouge au début de la série, ce qui est expliqué par le fait que le \hat{b} recalculé ne débute qu'avec une seule valeur.



Ce graphique a été conçu selon le même principe que précédemment.

Ici encore les deux courbes se ressemblent mais les amplitudes diffèrent également, toujours au début de la série essentiellement. Au cours de cette période, la courbe **bleue**, qui représente le nouveau \hat{a} , montre de très grandes variations de celui-ci mais plus on avance dans le temps, plus nous distinguons la tendance de \hat{a} et ainsi se rapproche de la courbe **rouge**, qui est modélisée à partir d'une tendance exponentielle.

Graphiques des accidents :

Les accidents à 2 ET :

Les accidents sont calculés à plus ou moins un écart-type et plus ou moins deux écart-types, ce qui nous donne les résultats suivants avec :

- en rouge prononcé : +2 écart-types
- en vert prononcé : -2 écart-types
- en mauve : moyenne
- en rouge (trait fin) : +1 écart-type
- en vert (trait fin) : -1 écart-type
- en bleu : la courbe représentant l'année

Pour repérer les accidents les plus importants, seuls les points de la courbe qui se situent au-dessus de la courbe des +2 écart-types et en dessous de celle de -2 écart-types.

Voici les mois qui sont concernés : Avril 1965, Mai 1968, Août 1975, Novembre 1980, Décembre 1980, Décembre 1981, Juillet 1983, Septembre 1983, Octobre 1983, Février 1984, Mars 1984.