JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

KRUMMEICH

Contribution à l'étude du mouvement de la population (suite)

Journal de la société statistique de Paris, tome 68 (1927), p. 191-199

<http://www.numdam.org/item?id=JSFS_1927__68__191_0>

© Société de statistique de Paris, 1927, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

III

CONTRIBUTION A L'ÉTUDE DU MOUVEMENT DE LA POPULATION *(Suite)*

15° *Comparaison de la formule de Pearl et de celle de Verhulst.* — Il est intéressant de voir si cette formule nous fournit un meilleur ajustement que

(1) L'indice des prix de détail pondéré s'établit à 915 en février 1927 (sur la base avril 1914), l'indice des prix de gros, à 854, l'indice des prix de détail non pondéré à 770. Ce dernier est un indice du genre monétaire, qui, en raison de sa constitution, ne semble pas traduire d'une façon correcte le niveau actuel des prix. Les prix ont légèrement baissé dans les mois suivants et semblent avoir atteint approximativement leur niveau d'équilibre.

la formule simple de Verhulst. Le tableau de comparaison plus haut nous montre une somme d'écarts positifs à peu près égale à celle des écarts négatifs, mais si on prend garde à la présence des deux écarts énormes de 1872 et 1876 parmi les positifs, on voit qu'en réalité les écarts négatifs sont plus importants. Le total des écarts est de 3,730, et l'écart moyen 0,196. Il est curieux de constater que cet écart monte à 0,204 si on ne tient pas compte de la région extrapolée 1896-1911, la zone extrapolée de la courbe est celle qui fournit le meilleur ajustement.

L'ajustement n'est donc pas meilleur qu'avec la formule de Verhulst qui nous donne un écart moyen de 0,196 par la méthode des ordonnées équidistantes, et de 0,185 par celle des sommes d'inverses. Du reste, on ne saurait invoquer en faveur de l'exactitude de la formule de Pearl (ni de celle de Verhulst qui en est un cas particulier), la perfection des ajustements qu'elle permet, sans quoi nous devrions accorder la même valeur à toute courbe qui nous fournira des ajustements aussi bons.

Le Dr Bowley a donné l'équation d'une simple parabole du second degré qui fournit pour la France un ajustement excellent. La formule, calculée par la méthode des moindres carrés est :

$$3{,}608 + 10{,}28\,t - 0{,}0602\,t^2$$

les chiffres étant exprimés en 10.000 habitants, l'origine étant 1864,7, et l'unité de temps 1 an. Nous donnons dans notre tableau comparatif l'ajustement par cette formule. La somme des écarts positifs est égale à celle des écarts négatifs, l'écart moyen est 0,212, c'est-à-dire presque égal à celui fourni par la formule de Pearl.

16° *Critique de la formule de Pearl.* — On peut invoquer en faveur de celle-ci son origine rationnelle et sa simplicité, mais la simplicité n'est pas une preuve, et aucune raison *a priori* ne nous permet de donner une forme linéaire à la fonction retardatrice. On a dit aussi que seule cette formule ne conduit pas à des résultats absurdes dès que l'on dépasse l'époque couverte par nos données, il n'est pas difficile d'imaginer des courbes répondant à cette condition et différentes des courbes de Pearl.

Du reste, les courbes de Verhulst qui constituent une courbe de Pearl ne sont pas particulièrement aptes à représenter l'accroissement de la population; elles conviennent pour représenter le développement des ferments, la pousse des dents des enfants de différents âges, la vitesse des réactions fermentées dans des liquides d'acidité ou d'alcalinité différente, etc..., en un mot, toutes les réactions dites auto-catalytiques, c'est-à-dire dans lesquelles un produit formé par la réaction accélère sa vitesse, leur interprétation mathématique est toujours la même.

L'argument que fournit l'observation d'un groupe de drosophiles vivant dans un espace fermé, placées dans des conditions bien définies de subsistance et de température, est douteux, car son évolution n'est pas assimilable à celle d'une population qui, comme les États-Unis, n'est homogène ni dans l'espace ni dans le temps, qui est ouverte à l'immigration, qui a vu sa mortalité par âge diminuer fortement au cours du siècle dernier, par suite des

progrès de l'hygiène, tandis que les drosophiles ne peuvent pas modifier leurs règles d'hygiène.

Du reste, si des populations du XIX^e^ siècle suivent un cycle de Verhulst, il faut bien remarquer que ce siècle est remarquable par la régularité de l'essor industriel, les ressources par sol ont permis un rendement proportionné aux efforts d'exploitation : mais rien ne prouve qu'il sera toujours ainsi, et malgré le progrès du machinisme, lorsque les ressources facilement exploitables seront utilisées, à un doublement de l'effort d'extraction ne correspondra peut-être plus un doublement de la production. On pourrait en dire autant de la production agricole, le rendement d'un terrain n'augmentant pas indéfiniment avec l'augmentation d'engrais.

Il serait faux de croire que la formule de Pearl et celle de Verhulst donnant des valeurs presque identiques à la population française limite, celle-ci soit déterminée avec quelque certitude. Le trait de plume d'un dessinateur nous aurait donné les mêmes résultats. Cette concordance est due accidentellement à la faible courbure de la courbe française d'accroissement vers 1900. Ainsi, suivant qu'on introduit ou non la constante p, on obtient pour la population anglaise une valeur limite variant de 70 à 100 millions.

La formule de Pearl ne permettrait pas la plus légère extrapolation même si on pouvait démontrer son exactitude, car à chaque instant le cycle démographique en cours peut être interrompu pour faire place à un cycle nouveau ou à un changement de genre de vie. On en trouve un exemple très net dans la population allemande vers 1866 montrant une branche asymptotique analogue à celle de la France aujourd'hui, puis entrant dans une nouvelle et rapide phase de progression. On ne peut donc pas prévoir le sens de la courbure, puisqu'à chaque instant on peut se trouver en présence du point d'inflexion causé par tout changement de cycle.

Notons ici que ce qui rend la formule de Pearl plus souple que celle de Verhulst, la rend aussi plus critiquable, car, pourvu qu'on la partage en tronçons assez petits, toute fonction continue et toujours croissante pourra être représentée avec une bonne approximation par des arcs de courbes de Verhulst.

Il faudrait bien se garder de conclure de ce qui précède que les formules de Pearl et de Verhulst sont inutiles. Elles conviennent très bien pour ajuster les populations industrielles du XIX^e^ siècle, et probablement celles qui se trouveront plus tard dans des conditions économiques ou industrielles analogues, et cela avec un nombre restreint de constantes. Si leur justification rationnelle est critiquable, elle est cependant très plausible pourvu que l'on limite leur emploi à un intervalle de temps assez court, on ne saurait en dire autant des autres courbes interpolatrices proposées qui, elles, n'ont souvent aucune justification rationnelle.

Leur emploi sera un instrument précieux pour guider les recherches relatives aux différentes régions d'un pays, ou au développement des différentes subdivisions de sa population active.

17° *Forme linéaire de l'accroissement.* — Nous venons de voir que si y_t et y_{t+h} sont les valeurs de la population à deux époques t et $t+h$ séparées par un intervalle de temps h, on a, dans l'hypothèse de Verhulst :

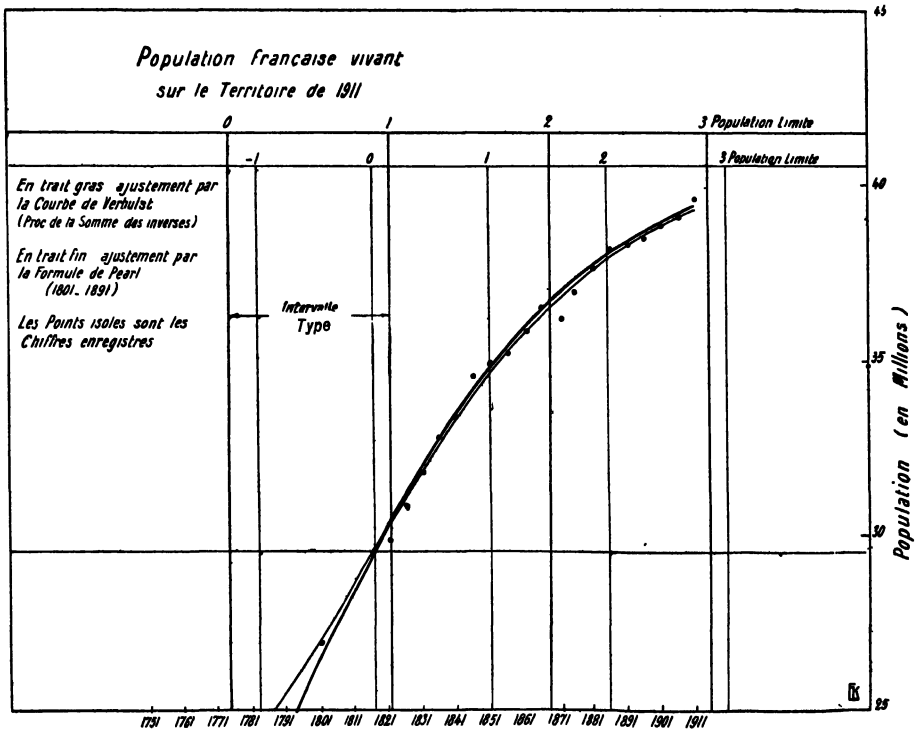
$$\frac{y_{t+h}-y_t}{y_t} = \left(1-\frac{y_{t+h}}{K}\right)\left(e^{\frac{h}{\alpha}}-1\right).$$

Il est intéressant de voir comment la population française vérifie cette équation. On peut, à cet effet, utiliser les recensements qui sont effectués en général de cinq en cinq ans. Si par exemple nous prenons $h = 50$ ans, la valeur de l'accroissement pour cent pendant la période s'écoulant de 1861 à 1911 sera :

$$100 \times \frac{y_{1911} - y_{1861}}{y_{1861}} = 100 \times \frac{39{,}602 - 35{,}835}{35{,}835} = 10{,}51\,\%.$$

Nous pourrons donc construire un tableau indiquant dans la colonne (3) l'accroissement dans l'intervalle de 50 années finissant l'année indiquée dans la colonne (1). La colonne (2) indique les populations utilisées. La colonne (4) indique les accroissements théoriques dans l'hypothèse de Verhulst, la courbe utilisée étant déterminée par les recensements de 1801-1856-1911. Les colonnes (5) et (6) donnent les écarts absolus et relatifs.

On peut constater que dans l'ensemble, les nombres enregistrés sont assez voisins des nombres ajustés, l'écart relatif moyen étant 0,05. La vérification se fait cependant mal pour les périodes 1846-1896 et 1851-1901 où on voit le taux croître de 11,49 à 11,63, tandis que la théorie prévoit une décroissance de 1,10; le taux est encore de 11,59 pour la période 1856-1906.



TABLEAU

Taux d'accroissement pour cent de la population française pendant la période de 50 ans se terminant l'année indiquée dans la colonne (1).

(1) ANNÉE	(2) POPULATION	(3) TAUX enregistré	(4) TAUX théorique	ÉCARTS (5) absolus	ÉCARTS (6) relatifs
1801.	27,000	»	»	»	»
1831.	31,788	»	»	»	»
1836.	32,760	»	»	»	»
1841.	33,401	»	»	»	»
1846.	34,547	»	»	»	»
1851.	34,902	29,27	28,12	+ 1,15	+ 0,041
1856.	35,174	»	»	»	»
1861.	35,835	»	»	»	»
1881	37,672	18,51	17,75	+ 0,76	+ 0,043
1886.	38,219	16,66	16,39	+ 0,27	+ 0,016
1891.	38,343	14,80	15,11	— 0,31	— 0,021
1896.	38,518	11,49	13,93	— 2,44	— 0,161
1901.	38,962	11,63	12,83	— 1,20	— 0,094
1906.	39,252	11,59	11,81	— 0,22	— 0,019
1911.	39,602	10,51	10,86	— 0,35	— 0,032

$$\text{Écart relatif moyen} = \frac{0{,}427}{8} = 0{,}05.$$

La vérification est beaucoup moins bonne si au lieu de 50 ans nous donnons à h une valeur plus faible, par exemple 10 ans. Il fallait s'y attendre, car les écarts de la population au chiffre de population ajustée peuvent être de l'ordre de grandeur de l'accroissement entre deux recensements successifs.

On peut représenter le phénomène par un graphique en portant en abscisses les valeurs de la population, et en ordonnées les accroissements pour cent. Les taux théoriques viennent se placer sur une droite qui passe par le point d'ordonnée nulle et d'abscisse 42,754 (valeur de la population limite, pour laquelle, on le sait, l'accroissement devient nul). On pourrait achever de déterminer la droite par sa pente, mais il est plus simple de la faire passer par un quelconque des points théoriques. On constate que les points expérimentaux se groupent grossièrement autour de cette droite.

La vérification est meilleure si, conformément à l'hypothèse de Pearl, on admet l'existence d'une constante correspondant aux cycles précédents. Les quantités y_t et y_{t+h} que nous aurons à considérer ne seront plus les chiffres fournis par les recensements mais ces mêmes chiffres diminués de la constante p. La vérification ne saurait donc être complètement expérimentale car la valeur de p fournie par le calcul implique l'hypothèse que la population est partie d'une valeur initiale donnée pour décrire un cycle nouveau (mais non l'hypothèse que ce cycle nouveau est un cycle de Verhulst, puisqu'en effet on introduit une seule constante calculée, et non quatre). Nous adopterons pour p la valeur 18,462 millions trouvée plus haut. Comme précédemment, nous obtiendrons les valeurs de :

TABLEAU

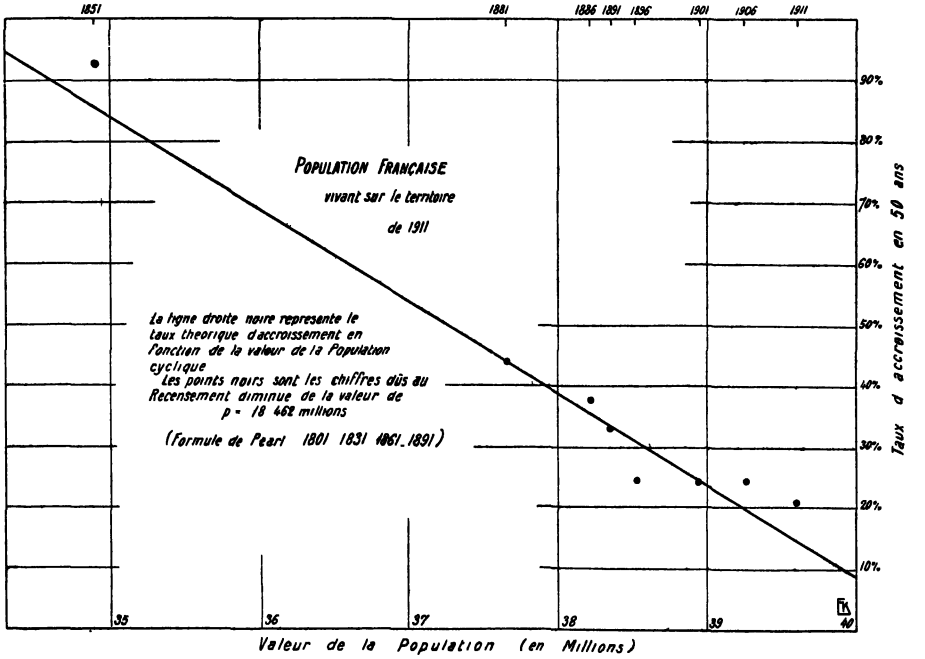
Taux d'accroissement pour cent de la population française (diminuée du chiffre atteint pendant les cycles précédents) pendant la période de 50 ans se terminant l'année indiquée dans la colonne (1).

(1) ANNÉE	(2) POPULATION diminuée de 18.462 millions	(3) TAUX enregistrés	(4) TAUX théorique	ÉCARTS (5) absolus	ÉCARTS (6) relatifs
1801	8,538	»	»	»	»
1831	13,326	»	»	»	»
1836	14,298	»	»	»	»
1841	14,939	»	»	»	»
1846	16,085	»	»	»	»
1851	16,440	92,55	89,55	+ 3,00	+ 0,034
1856	16,712	»		»	»
1861	17,373	»	»	»	»
1881	19,210	44,15	44,15	»	»
1886	19,757	38,32	38,84	— 0,52	— 0,013)
1891	19,881	33.08	34,10	— 1,02	— 0,030)
1896	20,056	24,69	(29,87)	(— 5,18)	(— 0,173)
1901	20,500	24,69	(26,12)	(— 1,43)	(— 0,055)
1906	20,790	24,40	(22,81)	(+ 1,59)	(+ 0,070)
1911	21,140	21,68	(19,89)	(+ 1,79)	(+ 0,900)

Écart relatif moyen pour la période 1801-1891 $= \frac{0,077}{4} = 0,01.9$

$$\frac{(y_{t+h} - p) - (y_t - p)}{y_t - p} = \frac{y_{t+h} - y_t}{y_t - p}.$$

A côté de ces nombres fournis par la colonne (3) du tableau nous avons disposé dans la colonne (4) les taux théoriques d'accroissement, la courbe utilisée étant déterminée par les recensements de 1801-1831-1861-1891. Les colonnes (5) et (6) indiquent respectivement les écarts absolus et les écarts relatifs. On voit que l'écart relatif moyen n'est que de 0,019 si l'on ne considère pas les écarts postérieurs à 1891, et correspondant à des valeurs théoriques de la population obtenues par extrapolation et dont on ne doit pas tenir compte.



On peut, comme précédemment, traduire les résultats par un graphique. Le groupement des points autour de la droite est sensiblement amélioré. Malheureusement, l'absence de certains recensements ne nous permet de déterminer que huit de ces points. D'autre part, la grandeur de leur écart à la droite théorique ne nous permet pas de voir si leur disposition suggère l'idée d'une droite ou celle d'une autre courbe, c'est-à-dire si la fonction retardatrice introduite par Verhulst est linéaire ou non. Par contre, on peut en tirer un argument en faveur de la constante supplémentaire introduite par Pearl.

Rappelons qu'inversement, on pourrait par la méthode des moindres carrés, déterminer les constantes de la formule de M. Pearl, ainsi que M. Yule l'a fait pour la formule simple de Verhulst, mais la façon médiocre dont les points expérimentaux vérifient le principe rendent cette détermination illusoire.

Du reste, l'application de la méthode des moindres carrés n'est légitime que si on possède un grand nombre d'équations, ce n'est pas le cas ici.

18° *Ajustement de la population française active.* — En exposant l'hypothèse d'additivité des cycles de Pearl et Reed, nous avons donné comme type de phase cyclique l'accroissement de population dû à l'essor industriel du XIX^e^ siècle. Cette hypothèse n'exclut nullement la coexistence dans un même pays de plusieurs types de population suivant chacun leur courbe particulière qui peut ne pas être une courbe de Verhulst (si toutes ces courbes étaient des courbes de Verhulst, leurs constantes devraient vérifier certaines identités, car la somme de deux telles fonctions n'est généralement pas une fonction de Verhulst).

Il est évidemment d'un grand intérêt de voir si pour la France il est possible de faire un tel ajustement de différentes portions de la population. Un premier moyen d'y parvenir serait de distinguer les départements industriels des départements agricoles, puisque pour chaque département nous possédons des statistiques distinctes, mais il serait entaché de trop de causes d'erreurs; ainsi vers 1800, on ne trouverait trace de grande industrie dans aucun département, alors que même dans les régions réputées agricoles, on voit surgir de nos jours de puissantes usines. La classification, même si on pouvait la faire à chaque recensement, serait inexacte. Il est préférable d'utiliser les statistiques qui nous fournissent le total pour la France entière, des individus exerçant chaque profession prévue dans la nomenclature. Malheureusement nous ne disposons à cet effet que de statistiques récentes.

Le premier recensement des industries et des professions eut lieu en 1861 au moyen de deux enquêtes séparées, mais c'est seulement celui de 1866 qui donna de bons résultats, les suivants furent plus sommaires. En 1896 on perfectionna beaucoup l'enquête par la création du bulletin individuel double permettant le double classement par industries et par professions. Nous utiliserons les recensements des années 1866, 1881, 1896, 1901, 1906, 1911.

Pour rendre les chiffres du recensement de 1866 comparables à ceux des recensements suivants, qui se rapportent au territoire français en 1911, nous admettrons que la répartition des professions en Alsace-Lorraine était la même que sur le reste du territoire, ce qui revient à multiplier les chiffres bruts de 1866 par le rapport :

$$\frac{\text{France diminuée de l'Alsace-Lorraine}}{\text{Population totale}} = \frac{36{,}485}{38{,}067} = 0{,}958442.$$

Dans l'étude de la population industrielle, il ne semble pas justifié d'ajouter à la population active. les personnes appartenant à la famille, car elles peuvent exercer une profession différente de celle du chef de famille. Par contre, le développement industriel entraîne l'augmentation du nombre d'employés, de fonctionnaires, etc... Nous avons pris comme chiffre de population industrielle le total de la population active de la France, diminué seulement de la population agricole et des domestiques.

Le tableau suivant résume les données statistiques que nous utiliserons.

Il n'est guère possible de faire l'étude de la population agricole dont l'accroissement actuel est beaucoup trop lent pour permettre de déterminer avec quelque précision les constantes de la formule de Verhulst, si toutefois elle suit un tel cycle.

Population française active vivant sur le territoire de 1911.

	1866	1881	1896	1901	1906	1911
Population active.	14,514	16,881	18,994	19,715	20,721	20,931
Professions agricoles	6,899	7,856	8,429	8,177	8,777	8,517
Service domestique	1,220	1,156	0,903	0,956	0,946	0,929
Total des agricult et des domest	8,119	9,012	9,332	9,133	9,723	9,446
Population active diminuée des agriculteurs et des domestiques.	6,395	7,869	9,662	10,582	10,998	11,485

L'ajustement effectué par la formule de Pearl, par le procédé des quatre ordonnées équidistantes, nous donne, en utilisant les recensements de 1866, 1881, 1896, 1911.

$$y_t = 3{,}803 + \frac{12{,}051}{1 + e^{\frac{1897{,}338 - t}{24{,}209}}}.$$

La constante $p = 3{,}803$ peut s'interpréter par le fait qu'avant le grand essor industriel du XIX^e siècle existait ce qu'on appelle aujourd'hui la petite industrie, indispensable à toute civilisation agraire.

La valeur de $p + K = 3{,}803 + 12{,}051 = 15{,}854$.

La fixation vers 1897 du point d'inflexion indique une diminution dans l'accroissement de la population industrielle depuis cette date.

Population industrielle ajustée par la formule de Pearl.

ANNÉE	POPULATION ACTIVE INDUSTRIELLE			DIFFÉRENCE d'accroissement depuis 1866 entre la population industrielle et la population française totale (1) ajustée par la formule de Pearl.
	cyclique	ajustée	enregistrée	
1866.	2,592	6,395	6,395	
1872.	3,132	6,935	»	— 31
1876.	3,530	7,333	»	21
1881.	4,066	7,869	7,869	163
1886.	4,640	8,443	»	383
1891.	5,241	9,044	»	667
1896.	5,859	9,662	9,662	1,004
1901.	6,481	10,283	10,582	1,376
1906.	7,092	10,895	10,998	1,766
1911.	7,682	11,485	11,485	2,161

(1) 1801-1831-1861-1891.

Les données sont trop peu nombreuses pour décider de la valeur de l'ajustement. Les deux recensements de 1901 et 1906 nous donnent respectivement des écarts de 0,299 et 0,103 au-dessus de la valeur calculée.

Nous avons vu dans l'ajustement de la population comment l'extension des données modifiait complètement l'allure de la courbe. Il est cependant intéressant de rapprocher les résultats obtenus de ceux que fournit la formule de Pearl pour la population totale établie à l'aide des recensements 1836, 1861, 1886, 1911 :

	Population totale	Population active industrielle
K	14,243	12,051
α	28,252	24,209
β	1843	1897.

Les valeurs trouvées pour K sont assez voisines, elles sont presque égales si on considère que l'accroissement de la population agricole et domestique n'est pas terminée et qu'il a été de 1,327 million de 1866 à 1911.

L'intervalle type, mesure du rythme d'accroissement, est à peu près le même dans les deux courbes. Par contre, les points d'inflexion diffèrent de plus d'un demi-siècle, la courbe de population ayant précédé la courbe industrielle, ce qu'il faut attribuer sans doute à la présence de la population agricole dont l'évaluation est bien plus ancienne.

Si au lieu de l'ajustement de la population 1836-1861-1886-1911 nous utilisons l'ajustement 1801-1831-1861-1891, nous obtenons des constantes absolument différentes, rendant toute comparaison impossible. Il est permis de supposer que la connaissance d'un recensement industriel en 1801 nous fournirait un ajustement de la population industrielle différent de celui que fournissent les données existantes, peut-être parallèle à celui de la population totale; mais en l'absence de recensements professionnels antérieurs à 1866, rien ne nous permet de voir si le rapprochement entre le cycle industriel 1866-1911 et le cycle démographique 1836-1911 est accidentel ou s'il constitue une vérification de l'hypothèse de Pearl. (*A suivre.*)