



HAL
open science

Prévision de la propagation de la pandémie Covid-19 dans le monde

Jonas Kibala Kuma

► **To cite this version:**

Jonas Kibala Kuma. Prévision de la propagation de la pandémie Covid-19 dans le monde. 2020.
hal-02525455

HAL Id: hal-02525455

<https://hal.science/hal-02525455v1>

Preprint submitted on 31 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

« Préviation de la propagation de la pandémie Covid-19 dans le monde »

Jonas Kibala Kuma*

Université de Kinshasa
Centre de Recherches Economiques et Quantitatives (CREQ)

Working paper n°001/WP.CREQ/03-20

Mars 2020

Résumé

Nous avons cherché à déterminer la date approximative à partir de laquelle la pandémie ou maladie à Corona Virus (Covid-19) pourrait affecter ou toucher toute la population mondiale, entendue que celle-ci se situerait autour de 7.759.559.800 suivant nos projections sur base des estimations de l'ONU (en ligne sur : <https://coronavirus.politologue.com>), sous l'hypothèse que le rythme actuel de l'évolution de la pandémie ainsi que les mesures prises par différentes autorités demeurent inchangés (toutes choses restant égales par ailleurs). En se lançant dans un tel exercice de modélisation, notre souci est d'attirer l'attention des autorités dans le monde sur le temps qui nous est imparti pour éradiquer cette pandémie, à défaut de quoi il serait difficile ou même impossible d'arrêter l'hémorragie dans le sens où ladite pandémie pourrait frapper ou infecter toute la population mondiale à partir du 25 mai 2020 jusqu'au 25 juin 2020.

Pour réaliser cette analyse, c'est dire prévoir notre série d'étude ou le nombre de personnes infectées par la pandémie ou maladie à Corona Virus (Covid-19), nous avons fait recours à une forme fonctionnelle de type exponentiel (au regard de la structure de notre série) dont le paramètre essentiel, traduisant le taux de croissance moyen journalier du nombre d'infectés au covid-19 dans le monde, est estimé suivant une approche non linéaire qui est l'algorithme de Gauss-Newton. Des études ultérieures pourraient approfondir la question avec des spécifications plus abouties (par exemple, recourir à la cointégration non linéaire pour échapper aux régressions fallacieuses) ou des formes fonctionnelles bien plus meilleures que celle retenue dans cette analyse.

Abstract

We tried to determine the approximate date from which Corona Virus (Covid-19) pandemic could affect or touch all the world population, heard that this one would be evaluated around 7,759,559,800 according to our projections on the basis of estimate of UNO (see : <https://coronavirus.politologue.com>), under the assumption that the current rate/rhythm of the evolution of pandemic as well as the measurements taken by various authorities remain unchanged (all things remaining are equal). In such exercise of modelling, our concern is to draw the world authorities' attention about few time assigned to us for eradicating this pandemic, if not it would be difficult or even impossible to stop the haemorrhage because this pandemic could strike or infect all the world population as from May 25, 2020 to June 25, 2020.

To carry out this analysis, or to forecast our study time series or the number of people infected by Corona Virus (Covid-19) pandemic, we made recourse to a functional form of exponential type, according to our time series structure, of which the essential parameter, which is the average growth rate number of infected people by covid19 pandemic in the world by day, is estimated according to a nonlinear approach which is the Gauss-Newton algorithm. Future studies could look further into the question with more succeeded specifications (for example, to resort to the nonlinear cointegration to escape the fallacious regressions) or functional forms much better than that retained in this analysis.

Codes JEL : C13, C52, C53, C61.

Mots clés : Modélisation économétrique, analyse dynamique, différentiel et intégral.

* Economiste Chercheur à l'Université de Kinshasa (DEA en cours) et au Centre de Recherches Economiques et Quantitatives (CREQ). Nos propos n'engagent pas ces institutions. Mail : kibala.jonas@gmail.com. Nous remercions Oasis Kodila Tedika pour ses commentaires et orientations dans le choix du thème et Joël Kazadi pour les fructueux échanges.

I. INTRODUCTION

Connaitre le processus générateur d'une série, sa distribution (loi) statistique ou mieux sa forme fonctionnelle est une étape importante dans la démarche de modélisation aux fins de prévision. A côté des méthodes formelles (estimation, simulation, etc.), les techniques informelles peuvent aider à détecter le mode opératoire d'une série (elles ont l'avantage d'être moins complexes), notamment l'examen de son évolution graphique. Si l'on réussit à se faire une idée sur la forme fonctionnelle de sa série par une approche informelle, parlant d'une analyse uni-variée des séries temporelles, l'estimation des paramètres de sa fonction et la prévision des valeurs futures (ou dans l'échantillon) de ladite série exige (au modélisateur) de recourir à des outils statistiques rigoureux qui garantissent que les paramètres ont été estimés sans erreurs (le recours aux moindres carrés ordinaires) et que la prévision faite est statistiquement bonne. Dans le cadre de cet exercice qui consiste pour nous à estimer la fonction d'évolution du nombre de cas infectés par le Covid-19 et à prévoir le nombre de cas dans un avenir proche, nous tenons compte de toutes ces exigences. Nous cherchons plus précisément à déterminer la date approximative à partir de laquelle la pandémie pourrait affecter ou toucher toute la population mondiale, entendue que celle-ci se situerait autour de 7.759.559.800 suivant nos projections sur base des estimations de l'ONU ⁽¹⁾, sous l'hypothèse que le rythme actuel de l'évolution de la pandémie ainsi que les mesures prises par différentes autorités demeurent inchangés (toutes choses restant égales par ailleurs). En se lançant dans un tel exercice de modélisation, notre souci est d'attirer l'attention des autorités dans le monde sur le temps qui nous est imparti pour éradiquer cette pandémie, à défaut de quoi il serait difficile, voire même impossible, d'arrêter l'hémorragie dans le sens où ladite pandémie pourrait frapper ou infecter toute la population mondiale à partir du 25 mai 2020 jusqu'au 25 juin 2020.

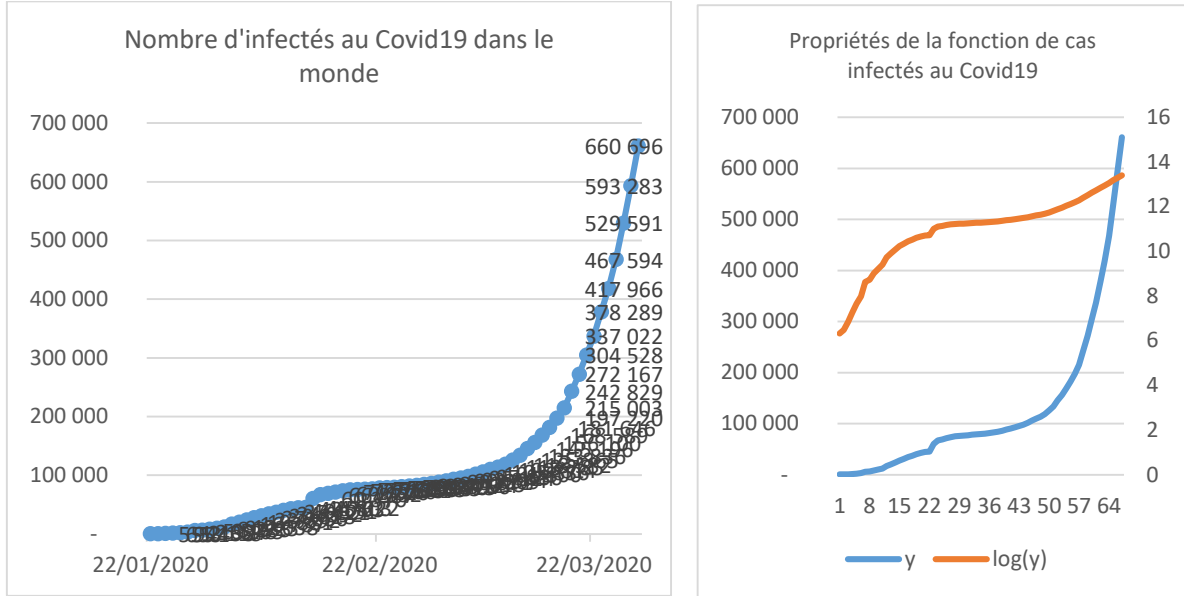
Rappelons que, suivant les derniers chiffres (considérant le moment où nous rédigeons cette note) du Coronavirus, issus du CSSE (Canadian Society of Safety Engineering, leur site : <https://www.csse.org/>) en date du samedi 28 mars 2020, le nombre total d'infections (nombre de personnes infectées par la pandémie Covid19) au niveau mondial est de 660 696 ; le nombre de guérisons est de 139 415 et le nombre de décès est de 30 652. Le taux de mortalité est de 4,64%, le taux de guérison est de 21,10% et le taux de personnes encore infecté est de 74,26% (source : <https://coronavirus.politologue.com/>).

II. MODELISATION

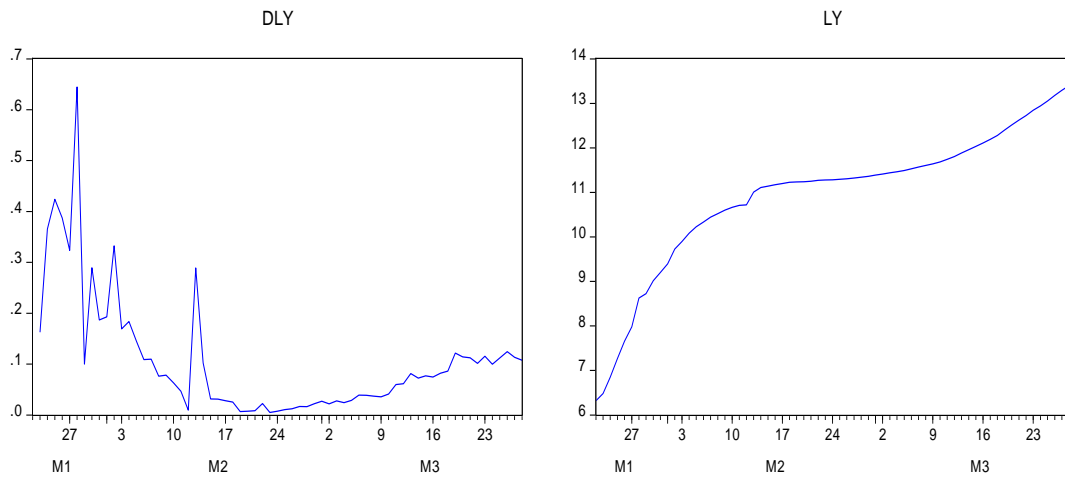
Au regard de la structure de notre série (Y), comme cela se dégage dans la figure 1 ci-dessous, il apparait que le nombre de personnes infectées par la pandémie ou maladie à Corona Virus (Covid-19), soit lié au temps de façon exponentielle.

¹ voir sur : <https://coronavirus.politologue.com>.

Fig.1 : Evolution de cas infectés au Covid19 dans le monde



Source : <https://coronavirus.politologue.com/> et nos calculs pour log(y)



Source : l'auteur

Cette intuition, fondée par les faits, peut se formaliser de deux manières en recourant à la technique de capitalisation en temps discret et celle en temps continu qui fait appel aux notions de différentiel et intégral. En effet :

(i) Modélisation en temps continu

Pour modéliser l'évolution de cas infectés par le Covid19 en temps continu, soit la série Y_t , l'on va supposer que cette série croît au taux « b » comme suit :

$$\frac{dY_t}{Y_t} = b \dots \dots (1)$$

Partant, pour trouver la fonction Y l'on doit résoudre l'équation différentielle. Pour ce faire, introduisons l'intégral dans les deux membres comme suit et résolvons après :

$$\int \frac{dY_t}{Y_t} = \int b dt \rightarrow \ln Y_t = bt + c$$

Introduisons la fonction exponentielle dans les deux membres, on a :

$$e^{\ln Y_t} = e^{bt} * e^c \rightarrow Y_t = c * e^{bt}$$

Si $t = 0$, alors : $Y_0 = ce^0 = c$. Ce qui amène à écrire :

$$Y_t = Y_0 * e^{bt} \dots \dots (2)$$

Avec : Y_t = série observée au temps t (journée en cours) ; Y_0 = série observée au temps t-1 (jour passé ou hier) ; b = le taux de croissance de Y_t (paramètre à estimer) et t = la variable temps qui traduit ici les jours (la tendance : t = 1 pour le 22/01/2020 ; 2 pour le 23/01/2020 ; 3 pour le 24/01/2020 ;... ; 67 pour le 28/03/2020).

Pour estimer le paramètre « b » par la méthode des moindres carrés ordinaires/MCO, laquelle méthode exige de minimiser la somme des carrés des erreurs du modèle, l'on doit commencer par linéariser l'expression (2) pour se conformer à l'une des hypothèses fondamentales sous-tendant l'utilisation des MCO (la linéarité du modèle qui justifie même le recours au coefficient de détermination pour apprécier la qualité de l'ajustement sans quoi le R^2 va tendre vers 0 tendant à biaiser l'appréciation sur la bonté globale du modèle). Ainsi, la transformation logarithmique de l'expression (2) est :

$$\ln Y_t = \ln Y_0 + bt \ln e$$

$$(\ln Y_t - \ln Y_0) = bt + u_t \dots \dots (3), \quad u_t \sim i. i. d$$

Avec : $\ln Y_t - \ln Y_0$ = la différence logarithmique de la série Y_t ; u_t = le terme d'erreur ou terme aléatoire qui est identiquement (absence d'hétéroscédasticité) et indépendamment distribuée (absence d'autocorrélation des erreurs).

La relation estimée s'écrit :

$$E(\ln Y_t - \ln Y_0) = E(b)t + E(u_t) \rightarrow (\ln \widehat{Y_t} - \ln Y_0) = \widehat{b}t$$

Avec : $E(b) = \widehat{b}$ (les estimateurs des MCO sont des meilleurs estimateurs linéaires sans biais, ils sont BLUE en anglais) ; $E(u_t) = 0$ (les résidus de l'estimation par les MCO sont en moyenne nuls). Après estimation, le paramètre « \widehat{b} » sera dit statistiquement significatif, ce qui atteste la pertinence de la relation estimée ou de la variable explicative considérée (ici la variable temps t ou le nombre de jours), s'il est supérieur à deux fois son écart-type ($\widehat{b} > 2 * \delta_{\widehat{b}}$) suivant la règle de Puce ; cela va amener à rejeter l'hypothèse nulle selon laquelle ledit paramètre n'est pas statistiquement significatif ($H_0: b = 0$), avec une statistique t de Student calculée (t_c) supérieure à sa valeur lue sur la table (t_c) ou la probabilité associée audit t calculé de Student inférieure à 5%, au profit de l'alternative qui admet la significativité dudit paramètre estimé ($H_1: b \neq 0$).

Sur Eviews 10, pour estimer le paramètre « b », l'on pourra taper directement la commande (estimation du modèle (2)) : $NLS Y=Y(-1)*EXP(C(1)*T)$.

(ii) Modélisation en temps discret

Pour modéliser l'évolution de cas infectés par le Covid19 (la série Y_t) en temps discret, l'on va supposer que cette série croît comme suit :

$$Y_t = Y_0 * [(1 + b)^t] \dots \dots (4)$$

Si l'on passe par la transformation logarithmique de l'expression (4) ci-dessus, elle devient :

$$\ln Y_t = \ln Y_0 + t * \ln(1 + b) \dots \dots (5)$$

Si l'on pose : $\ln Y_t = X_t$; $\ln Y_0 = a_0$ et $\ln(1 + b) = a_1$, l'expression se r crit :

$$X_t = a_0 + a_1 * t + u_t \dots \dots (6), \quad u_t \sim i.i.d$$

Apr s estimation de la relation (6), on retrouve les param tres de d part comme suit (tous ces param tres sont estim s, bien que n'ayant pas de chapeau) :

$$\ln Y_0 = a_0 \rightarrow e^{\ln Y_0} = e^{a_0} \rightarrow Y_0 = e^{a_0}$$

$$\ln(1 + b) = a_1 \rightarrow e^{\ln(1+b)} = e^{a_1} \rightarrow 1 + b = e^{a_1} \rightarrow b = e^{a_1} - 1$$

III. ESTIMATION

Pour besoin de comparaison et de p dagogie, nous allons estimer les expressions (2) et (3), soit la capitalisation en temps continue. Les donn es sont observ es sur une fr quence journali re et couvrent 67 jours depuis le d but de la pand mie.

L'estimation de l'expression (3) par les MCO,   l'aide du logiciel Eviews 10, donne les r sultats suivants :

Dependent Variable: DLY				
Method: Least Squares				
Date: 03/29/20 Time: 20:32				
Sample (adjusted): 1/23/2020 3/28/2020				
Included observations: 66 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
T	0.001680	0.000459	3.659160	0.0005
R-squared	-0.503737	Mean dependent var		0.107304
Adjusted R-squared	-0.503737	S.D. dependent var		0.119882
S.E. of regression	0.147008	Akaike info criterion		-0.981624
Sum squared resid	1.404738	Schwarz criterion		-0.948448
Log likelihood	33.39360	Hannan-Quinn criter.		-0.968515
Durbin-Watson stat	0.478058			

Et, l'estimation directe de l'expression (2) par les moindres carr s non lin aires suivant l'algorithme de Gauss-Newton ⁽²⁾,   l'aide du logiciel Eviews 10, donne les r sultats suivants :

Dependent Variable: Y				
Method: Least Squares (Gauss-Newton / Marquardt steps)				
Date: 03/29/20 Time: 20:39				
Sample (adjusted): 1/23/2020 3/28/2020				
Included observations: 66 after adjustments				
Convergence achieved after 74 iterations				
Coefficient covariance computed using outer product of gradients				
Y=Y(-1)*EXP(C(1)*T)				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.001680	3.84E-05	43.71915	0.0000
R-squared	0.999363	Mean dependent var		125025.9

² Lire   ce sujet notre manuel consacr    la r gression non lin aire, disponible en ligne sur : <https://hal.archives-ouvertes.fr/cel-02168940>.

Adjusted R-squared	0.999363	S.D. dependent var	144417.4
S.E. of regression	3645.284	Akaike info criterion	19.25529
Sum squared resid	8.64E+08	Schwarz criterion	19.28847
Log likelihood	-634.4246	Hannan-Quinn criter.	19.26840
Durbin-Watson stat	0.745902		

Les résultats de l'estimation de l'expression (2) sont préférables à ceux de (3) au regard du coefficient de détermination élevé et du coefficient de Durbin-Watson supérieur, bien que les critères AIC et SIC sont plus minimal pour l'expression (3). Ainsi, on retient le modèle (2) estimé et on note ce qui suit :

- Le modèle est globalement bon, expliquant à 99% la dynamique du nombre de cas infectés par le Covid19 dans le monde. Toutefois, le coefficient d'autocorrélation de Durbin-Watson, qui atteste l'absence d'autocorrélation des erreurs d'ordre 1 s'il s'approche de 2, est largement inférieur à 2 tendant à remettre en cause la bonté de la spécification globale du modèle ou nécessitant d'intégrer d'autres variables pertinentes qui pourraient avoir été ignorées ici. Des études ultérieures pourraient approfondir la question avec des spécifications plus abouties (par exemple, recourir à la cointégration non linéaire pour échapper aux régressions fallacieuses) ou des formes fonctionnelles bien plus meilleures que celle retenue dans cette analyse. Toutefois, le test de Ljung-Box et Box-Pierce (bien qu'informel) nous renseigne une absence d'autocorrélation d'erreurs (voir la figure 2 ci-dessous).

Fig.2 : corrélogramme des résidus de l'estimation du modèle (2)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.009	0.009	0.0057	0.940
		2	-0.072	-0.072	0.3708	0.831
		3	-0.057	-0.056	0.6039	0.896
		4	-0.076	-0.081	1.0251	0.906
		5	-0.057	-0.066	1.2678	0.938
		6	-0.008	-0.024	1.2727	0.973
		7	0.022	0.003	1.3102	0.988
		8	-0.021	-0.038	1.3440	0.995
		9	-0.018	-0.028	1.3682	0.998
		10	0.022	0.012	1.4062	0.999
		11	0.075	0.070	1.8704	0.999
		12	-0.013	-0.016	1.8847	1.000
		13	-0.038	-0.032	2.0078	1.000
		14	-0.016	-0.011	2.0301	1.000
		15	-0.034	-0.028	2.1313	1.000
		16	0.002	0.002	2.1316	1.000
		17	-0.031	-0.045	2.2180	1.000
		18	-0.017	-0.031	2.2460	1.000
		19	-0.037	-0.047	2.3742	1.000
		20	-0.021	-0.032	2.4175	1.000
		21	-0.020	-0.043	2.4572	1.000
		22	0.015	-0.011	2.4793	1.000
		23	0.024	0.005	2.5381	1.000
		24	0.063	0.056	2.9579	1.000
		25	0.075	0.073	3.5763	1.000
		26	0.051	0.065	3.8725	1.000
		27	-0.022	-0.001	3.9262	1.000
		28	-0.023	0.011	3.9871	1.000

- Les résidus de l'estimation sont homoscedastiques (voir tableau ci-dessous), bien que non normalement distribués. La forme fonctionnelle estimée est non linéaire (avec une approche d'estimation non linéaire de Gauss-Newton), ce qui peut justifier la violation de quelques hypothèses des MCO.

Heteroskedasticity Test: White			
F-statistic	0.143025	Prob. F(1,64)	0.7065
Obs*R-squared	0.147166	Prob. Chi-Square(1)	0.7013
Scaled explained SS	0.219637	Prob. Chi-Square(1)	0.6393

- Le nombre de personnes infectées par la pandémie ou maladie à Corona Virus (Covid-19) croit au taux $b = 0,001680$, soit 0,16% en moyenne journalière. Le paramètre « b » ainsi estimé est statistiquement significatif au regard de la statistique t de Student calculée (soit $t_c = 43,71$)

qui est largement supérieure à sa valeur lue sur la table ($t_c = 1,96$) qui avoisine 2 pour tout échantillon dont la taille dépasse 30 observations. Cette significativité statistique de « \hat{b} » est attestée également par la probabilité associée au t calculé de Student qui est de 0,00, ce qui amène à dire que le paramètre estimé est statistiquement significatif au seuil de 1%.

Par ailleurs, dans le souci de tester la robustesse de nos estimations et prévisions, nous nous sommes exercés à trouver une variante du modèle estimé « $Y=Y(-1)*EXP(C(1)*T)$ » (Cfr expression (2)), en remplaçant $y(t-1)$ par la valeur de départ de l'échantillon ou le nombre de personnes infectées par le Covid-19 dans le monde en date du 22/01/2020, soit 555 personnes. Ainsi, nous avons également estimé le modèle suivant ($Y=555*EXP(C(1)*T)$) :

$$Y_t = 555 * e^{bt} \dots \dots (2a)$$

Les principaux résultats d'estimation de l'expression (2a) sont les suivants (les résultats d'estimation sont donnés après les commentaires) :

- Le modèle est globalement bon, expliquant à 93% la dynamique du nombre de cas infectés par le Covid19 dans le monde, avec une forte autocorrélation des erreurs d'ordre 1 (suivant la statistique de Durbin-Watson) et d'ordre supérieur à 1 selon les résultats de Breusch-Godfrey Serial Correlation LM Test, soit un problème typique aux régressions fallacieuses.
- Le nombre de personnes infectées par la pandémie ou maladie à Corona Virus (Covid-19) croit au taux $b = 0,105628$, soit 10,56% en moyenne journalière sur base de l'effectif de départ. Ce paramètre apparait aussi statistiquement significatif à 1%.
- Les résidus de l'estimation de l'expression (2a), bien que normalement distribués au regard de la probabilité associée à la statistique calculée de Jarque-Bera qui est supérieure à 5% (on accepte l'hypothèse nulle de normalité des résidus), sont hétéroscédastiques (la variance n'est pas minimale) en plus d'être autocorrélés. Aussi, comparés aux résultats d'estimation de l'expression (2), l'expression (2a) affiche des coefficients AIC et SIC (respectivement Aikaike et Schwarz information criterion) plus grands, ce qui la rend moins préférable à (2).

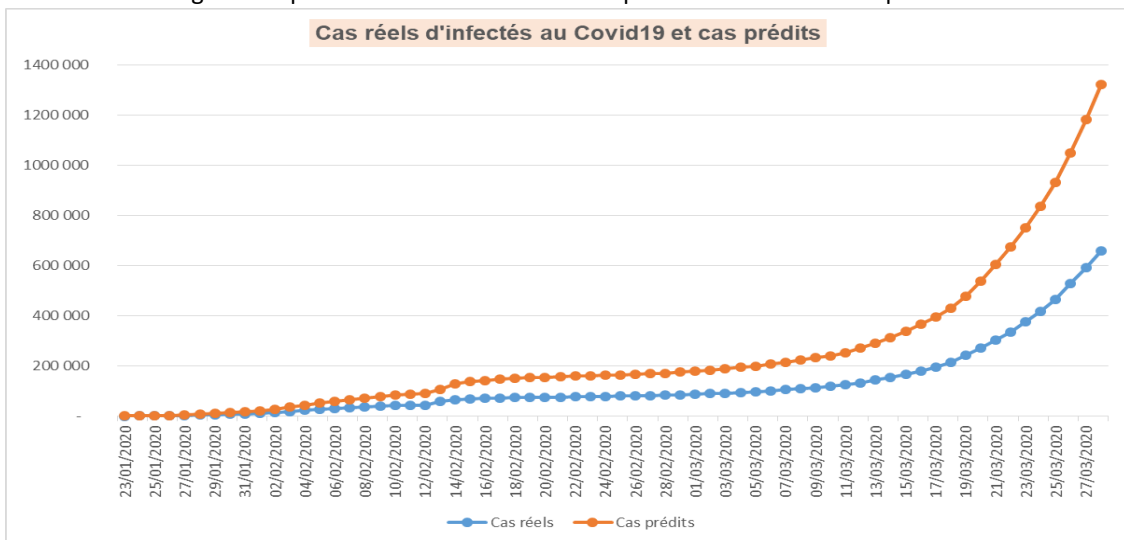
Dans la partie qui suit, nous essayons tout de même de comparer aussi les résultats de prévision issus de ces deux procédés ou spécifications (2 et 2a).

Dependent Variable: Y				
Method: Least Squares (Gauss-Newton / Marquardt steps)				
Date: 03/30/20 Time: 17:51				
Sample: 1/22/2020 3/28/2020				
Included observations: 67				
Convergence achieved after 17 iterations				
Coefficient covariance computed using outer product of gradients				
Y=555*EXP(C(1)*T)				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.105628	0.000385	274.4290	0.0000
R-squared	0.935904	Mean dependent var		123168.1
Adjusted R-squared	0.935904	S.D. dependent var		144123.7
S.E. of regression	36488.04	Akaike info criterion		23.86217
Sum squared resid	8.79E+10	Schwarz criterion		23.89508
Log likelihood	-798.3827	Hannan-Quinn criter.		23.87519
Durbin-Watson stat	0.008792			

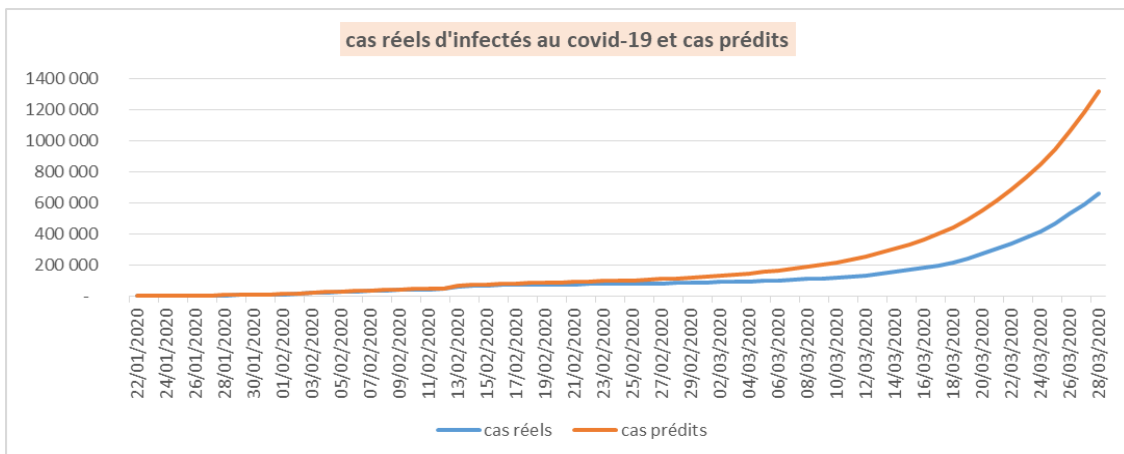
IV. PREVISION

L'une des conditions que doit remplir un modèle estimé pour une prévision de bonne qualité, statistiquement parlant, c'est la stabilité de ses paramètres dans l'échantillon considéré. Ce test (le Cusum test de stabilité) n'était pas adapté à l'approche d'estimation non linéaire que nous avons adoptée, autant pour le test de Ramsey concernant la bonté de la spécification choisie. Le coefficient de Theil affiche une valeur proche de 0 pour l'expression 2a et proche de 1 pour l'expression 2 (voir les figures 3c et 3b, respectivement), pourtant ledit coefficient avoisine 0 ou est faible pour un modèle meilleur comparé à un autre. Comparée à 2, l'expression 2a paraît fournir ainsi des bonnes prédictions au regard de son faible coefficient de Theil. Toutefois, la figure 3 ci-dessous témoigne de la bonté relative de nos spécifications 2 et 2a en ce qu'elles ont permis de reproduire la tendance exponentielle de la série de départ, ainsi que ses pics. Aussi, pour l'expression 2, on a constaté que les écarts de prévision s'amplifient au fil du temps, caractéristique des fonctions exponentielles, ce qui nous empêche de prévoir le nombre d'infectés au Covid19 sur une période plus longue, afin de limiter le biais (la figure 3a rend compte de l'évolution desdits écarts). Pour ce qui est de l'expression 2a, le constat est que les écarts sont plutôt plus importants au départ et tendent à se réduire dans le temps, avec le risque de s'amplifier sur le long terme, ce qui implique de limiter l'intervalle ou la plage de prévision pour échapper à d'éventuels biais.

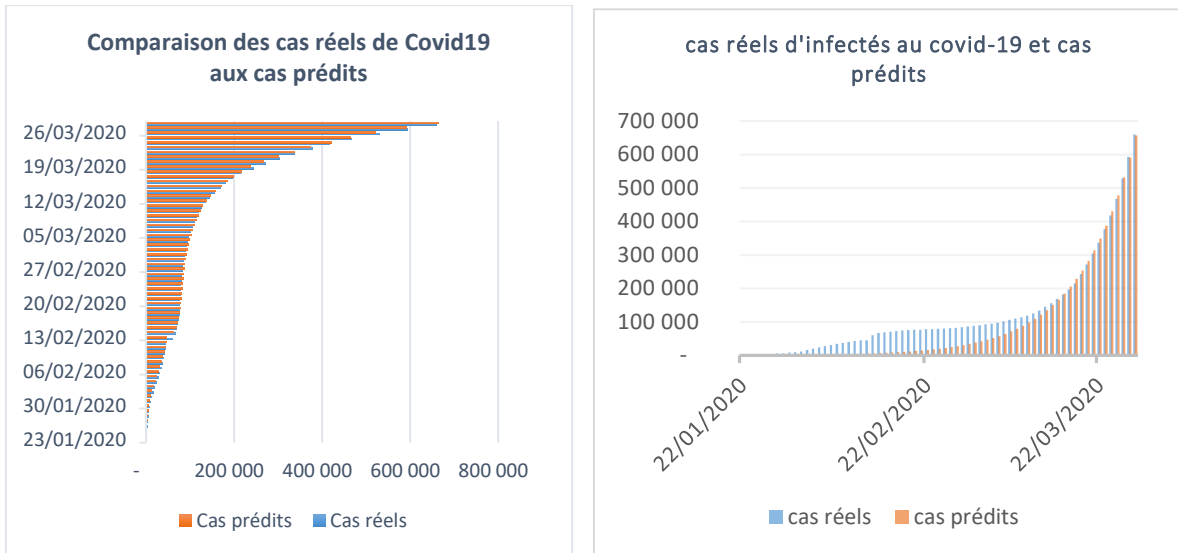
Fig.3 : comparaison des cas réels infectés par Covid19 et des cas prédits



Source : nos estimations et prédictions à l'aide du logiciel Eviews 10 (expression 2)



Source : nos estimations et prédictions à l'aide du logiciel Eviews 10 (expression 2a)



Source : nos estimations et prédictions à l'aide du logiciel Eviews 10 (à gauche pour l'expression 2 et à droite pour l'expression 2a)

Fig.3a : comparaison des cas réels infectés par Codvid19, des cas prédits et des résidus (expression 2)

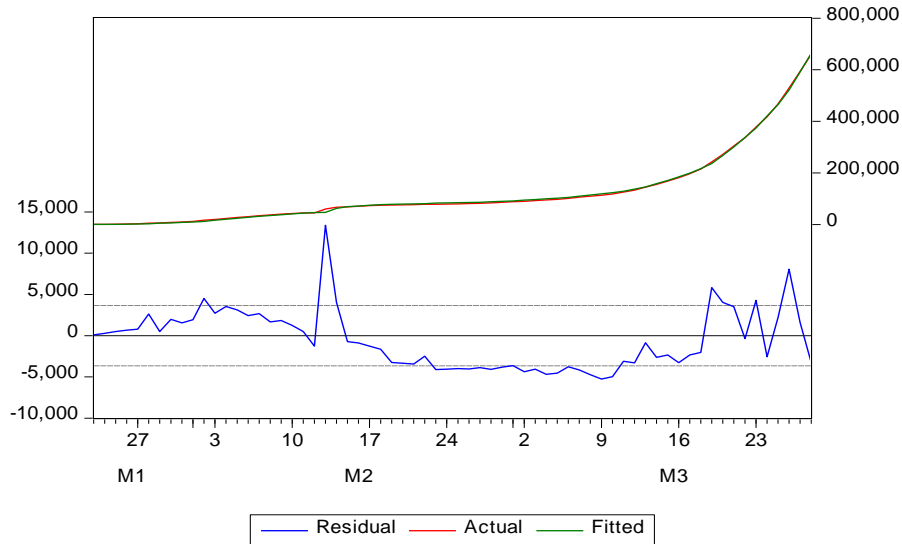


Fig.3b : le coefficient de Theil estimé (expression 2)

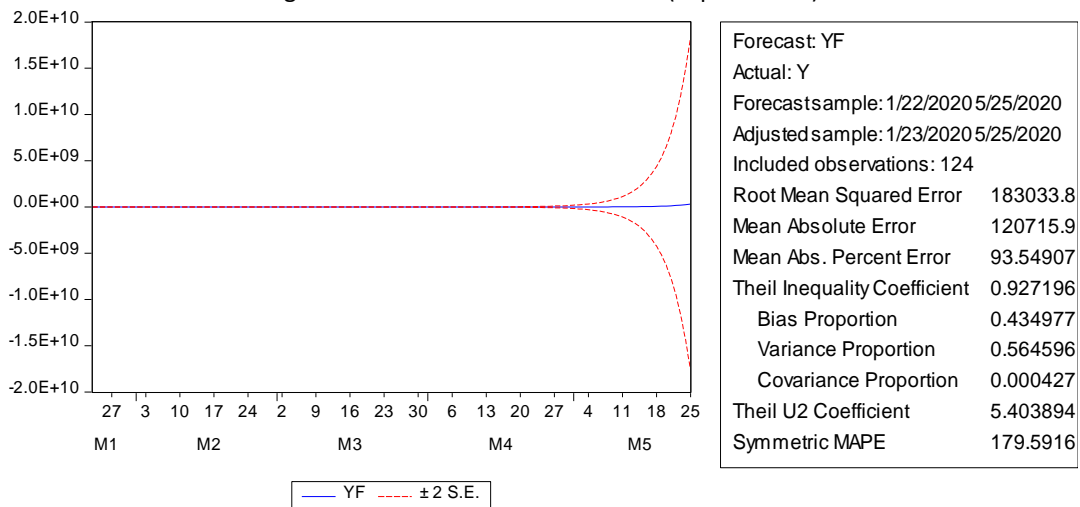
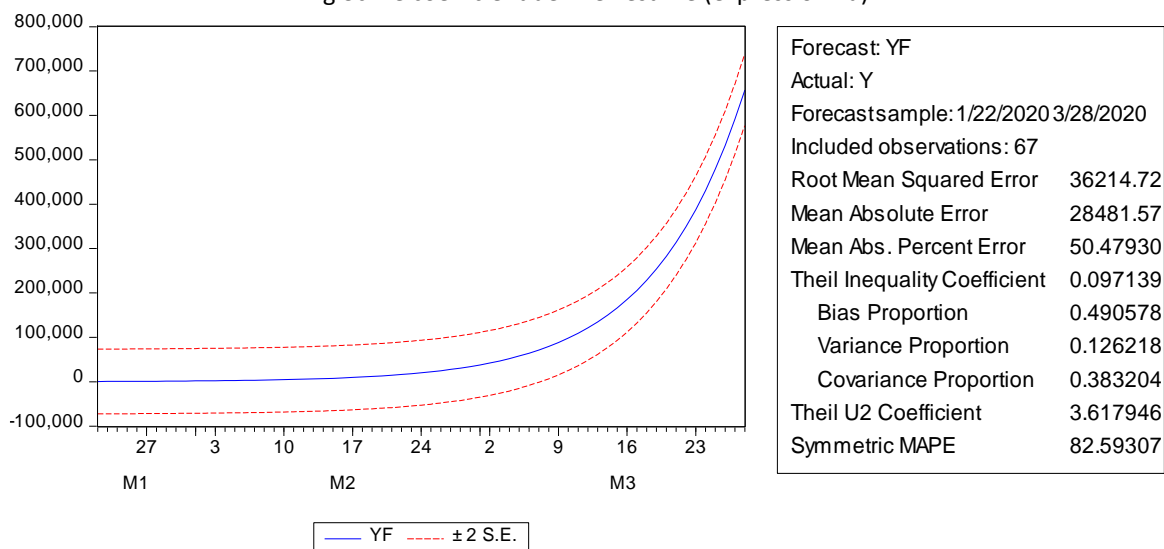
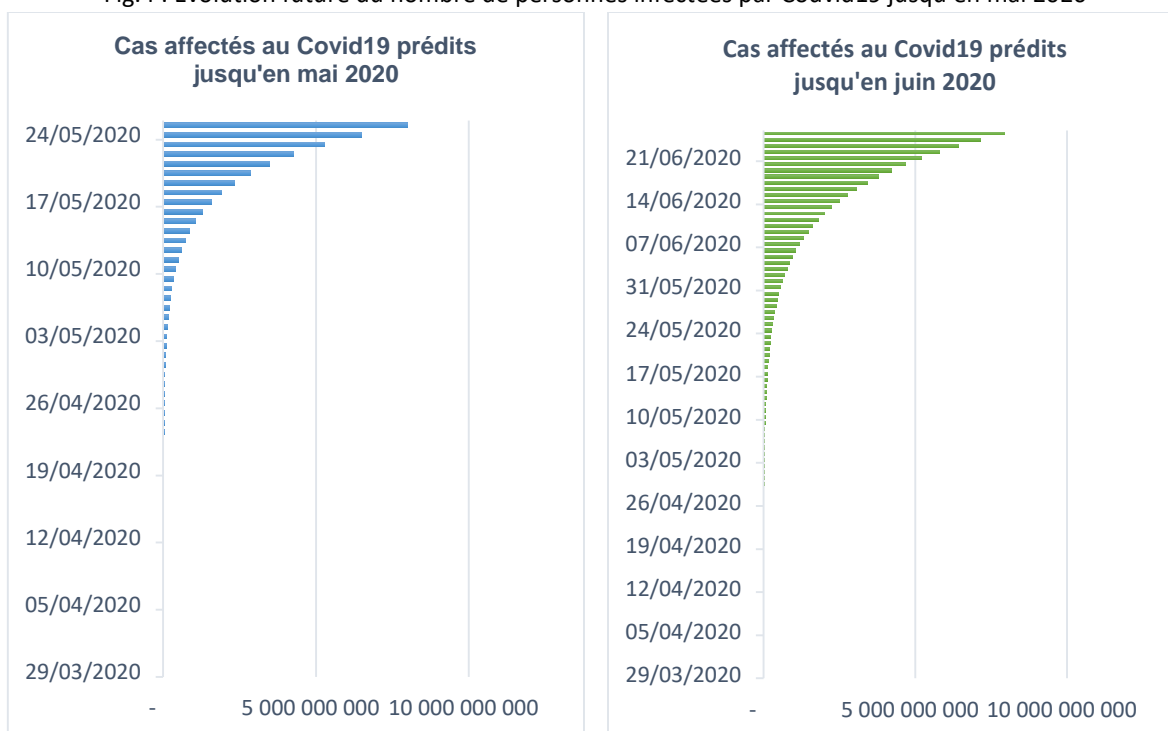


Fig.3c : le coefficient de Theil estimé (expression 2a)



A l'issue de nos estimations, les prévisions sont telles que la pandémie Covid-19 pourrait frapper ou infecter toute la population mondiale à partir du 25 mai 2020 jusqu'au 25 juin 2020 (la figure 4 illustre nos propos), si le rythme actuel de l'évolution de la pandémie ainsi que les mesures prises par différentes autorités demeurent inchangés (toutes choses restant égales par ailleurs).

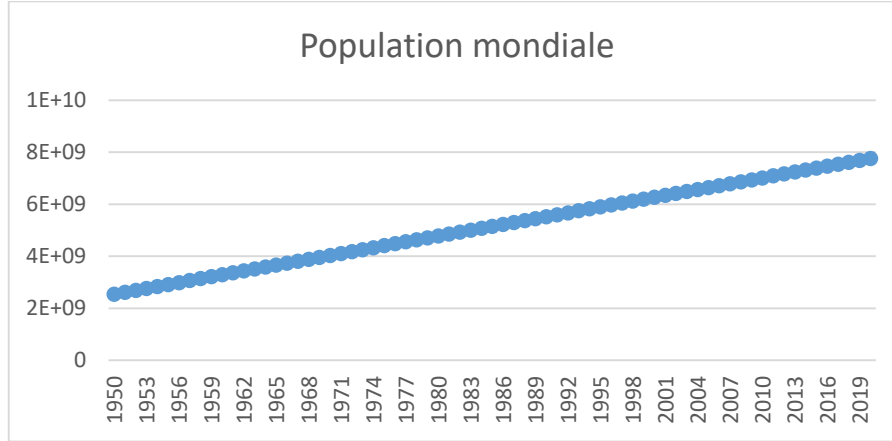
Fig.4 : Evolution future du nombre de personnes infectées par Covid19 jusqu'en mai 2020



Source : nos estimations (expression 2 pour le graphe à gauche et expression 2a pour le graphe à droite)

En effet, suivant les estimations de l'ONU (les derniers chiffres communiqués sont pour l'année 2015), la population mondiale est passée de 2,54 milliards d'habitants en 1950 à 7,38 milliards d'habitants en 2015 (la figure 5 est éloquent). Entre 1950 et 2015, soit pratiquement 65 ans, la population mondiale a progressé de 291,1% avec une moyenne de +74.565.140 d'habitants par an (source : <https://www.politologue.com/population-mondiale/>).

Fig.5 : Evolution de la population mondiale de 1950 à 2015



Source : <https://www.politologue.com/population-mondiale/>

Dans l’hypothèse où la population croit avec une moyenne de 74.565.140 d’habitants par an, on a fait une projection linéaire pour dégager la population mondiale estimée autour de 7.759.559.800 habitants en 2020. Suivant les estimations de modèles ou expressions 2 et 2a, nos projections tablent sur une pandémie (Covid-19) qui pourrait atteindre, à l’échelle mondiale, respectivement 8.008.369.197 personnes le 25 mai 2020 ou 7.954.420.767 personnes le 25 juin 2020.

En dernier lieu, il tient de rappeler que des études ultérieures pourraient approfondir l’analyse et améliorer nos résultats avec des spécifications plus abouties (par exemple, recourir à la cointégration non linéaire pour échapper aux régressions fallacieuses) ou des formes fonctionnelles bien plus meilleures que celle retenue dans cette analyse.

Bibliographie

- Bosonga B.L. (2019), « *Manuel d’Econométrie* », PUK, Editions Terabytes, RDC, 326 p.
- _____ (2011), « *Mathématiques appliquées à l’économie* », RDC, 312 p.
- Bourbonnais R. (2015), « *Econométrie : cours et exercices corrigés* », 9^e édition, éd. DUNOD, Paris, 32 p.
- Bourbonnais R. et Terraza M. (2016), « *Analyse des séries temporelles : Applications à l’économie et à la gestion – cours et exercices corrigés* », éd. Dunod, 4^e édition, 354.
- Dameron P. (2001), « *Mathématiques des modèles économiques : Analyse dynamique* », éd. Economica, Paris, 321 p.
- Dowling E.T. (1995), « *Mathématiques pour l’Economiste : cours et problèmes* », Série Schaum, éd. McGraw-Hill, 2^e édition, Paris, 482 p.
- Greene W. (2011), « *Econométrie* », éd. Pearson, 7^e édition, France, 988 p.
- Hubbard W. (1999), « *Equations différentielles et systèmes dynamiques* », Cassini.
- Kibala Kuma J. (2018), « *Application sur la régression linéaire simple (Spécifications, Estimation, Inférence et Interprétation des Résultats) : Examen de la convergence économique au sein de l’Union Européenne* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-01771444>), Kinshasa, RDC, 29 p.

- Kibala Kuma J. (2018), « *Econométrie Appliquée : Manuel des cas pratiques sur EViews et Stata* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-01771756>), Kinshasa, RDC, 88 p.
- Kibala Kuma J. (2018), « *Econométrie Appliquée : Recueil des cas pratiques sur Eviews et Stata* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-01771070>), Kinshasa, RDC, 204 p.
- Kibala Kuma J. (2018), « *Econométrie Appliquée : Recueil des cas pratiques sur EViews (Régression linéaire simple et multiple)* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-01771168>), Kinshasa, RDC, 35 p.
- Kibala Kuma J. (2018), « *Modèles de régression non linéaires : éléments de théorie et pratiques sur Logiciel* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-02168940>), Kinshasa, RDC, 26 p.
- Kibala Kuma J. (2018), « *Prév́ision par l'approche méthodologique de Box et Jenkins : Cas d'une série saisonnière et non stationnaire du type DS* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-01771475>), Kinshasa, RDC, 10 p.
- Kibala Kuma J. (2018), « *Prév́ision par l'approche méthodologique de Box et Jenkins : Cas d'une Série Non Saisonnière et Non Stationnaire du type TS (Pratique sur EViews et Stata)* », publié dans HAL (<https://hal.archives-ouvertes.fr/cel-01771600>), Kinshasa, RDC, 26 p.
- Kibala Kuma J. (2019), « *L'Econométrie avec Eviews : Répertoire de quelques commandes de base* », Editions Universitaires Européennes, 117 p.
- Kintambu Mafuku E.G. (2004), « *Principes d'Econométrie* », Presses de l'Université Kongo, 4è édition, 285 p.
- Marti R. (2002), « *Optimisation intertemporelle : application aux modèles macroéconométriques* », éd. Economica.
- Mignon V. (2008), « *Econométrie : Théorie et Applications* », éd. Economica, 236 p.
- Wade A. (2007), « *Les mathématiques de l'analyse économique moderne* », éd. Economica, Paris.

Sites internet consultés

- <https://experience.arcgis.com/experience/685d0ace521648f8a5beeeee1b9125cd>
- <https://coronavirus.politologue.com/>
- <https://www.csse.org>
- <https://data.europa.eu/euodp/fr/data/dataset?tags=corona+virus>
- <https://www.worldometers.info/coronavirus/>.

Annexe

• Estimation

```
Estimation Command:
=====
NLS Y=Y(-1)*EXP(C(1)*T)

Estimation Equation:
=====
Y=Y(-1)*EXP(C(1)*T)

Substituted Coefficients:
=====
Y=Y(-1)*EXP(0.00167995423935*T)
```

• **Extrait de l'évolution de la population mondiale**

Année	Population estimée	Variation
2015	7 383 008 820 hab.	+84 555 787 hab. (+1,16%)
2014	7 298 453 033 hab.	+85 026 581 hab. (+1,18%)
2013	7 213 426 452 hab.	+85 249 517 hab. (+1,20%)
2012	7 128 176 935 hab.	+85 168 349 hab. (+1,21%)
2011	7 043 008 586 hab.	+84 839 427 hab. (+1,22%)
2010	6 958 169 159 hab.	+84 428 105 hab. (+1,23%)
2009	6 873 741 054 hab.	+83 969 801 hab. (+1,24%)
2008	6 789 771 253 hab.	+83 352 660 hab. (+1,24%)
2007	6 706 418 593 hab.	+82 570 680 hab. (+1,25%)
2006	6 623 847 913 hab.	+81 688 530 hab. (+1,25%)
2005	6 542 159 383 hab.	+80 788 518 hab. (+1,25%)
2004	6 461 370 865 hab.	+79 961 878 hab. (+1,25%)
.....		
1963	3 210 271 352 hab.	+61 027 107 hab. (+1,94%)
1962	3 149 244 245 hab.	+58 938 966 hab. (+1,91%)
1961	3 090 305 279 hab.	+57 092 752 hab. (+1,88%)
1960	3 033 212 527 hab.	+55 387 841 hab. (+1,86%)
1959	2 977 824 686 hab.	+53 743 443 hab. (+1,84%)
1958	2 924 081 243 hab.	+52 128 965 hab. (+1,82%)
1957	2 871 952 278 hab.	+50 568 834 hab. (+1,79%)
1956	2 821 383 444 hab.	+49 140 909 hab. (+1,77%)
1955	2 772 242 535 hab.	+47 940 067 hab. (+1,76%)
1954	2 724 302 468 hab.	+47 072 110 hab. (+1,76%)
1953	2 677 230 358 hab.	+46 645 974 hab. (+1,77%)
1952	2 630 584 384 hab.	+46 767 598 hab. (+1,81%)
1951	2 583 816 786 hab.	+47 542 065 hab. (+1,87%)
