Deep Assessment Methodology Using Fractional Calculus on Mathematical Modeling and Prediction of Population of Countries

Ertuğrul Karacuha[1], Vasil Tabatadze[1], Nisa Ozge Onal[1], Kamil Karacuha[1], Derya Bodur[1]

[1] Informatics Institute, Istanbul Technical University, Istanbul 34467, Turkey.

The modelling of data and prediction for the upcoming years or events are one of the main concerns of not only all countries but also companies, investors, manufacturers, and institutions. The scientists investigate on a relation among telecommunication, economic growth, and financial development using technical, economic, social events and data. Besides, the population changes affect balances in any aspect. Therefore, the prediction of the population for each country is prominent and essential for the other prediction. In this study, the dataset consists of the populations of the USA, France, Britain, Italy, Spain, and Turkey were used from 1960 until 2018. In our study, populations of countries are tried to be modelled, and the predictions are made for the upcoming years, using fractional calculus. The newly developed approach uses the data for the modelling by employing Least Squares method and fractional calculus.

Keywords: deep assessment; fractional calculus; least squares; modelling; population; prediction

MSC Classification: 26A33, 93E24

1 | INTRODUCTION

In the last quarter of the century, the development of telecommunication technology and services in parallel with the information technologies has increased the information produced, transmitted and evaluated. It can be said that there is a relation among telecommunication, economic growth, and financial development [1-6]. The modeling of all technical, economic, social events and data has been the interest of scientists for many years [7-9] (Beaudry and Portier 2006; Simpson et al. 2005; Ho et al. 2018).

Fractional calculus is an approach that studies fractional-order differential and integral operators – this idea first offered at the end of the seventeenth century. Fractional calculus (FC) as a question is firstly arisen in 1695 by French Mathematician Marquis de L'Hopital (1661-1704) to Wilhelm Leibniz (1646-1716) [10-11]. The main object of interest was what would be if the order of differentiation were a real number instead of an integer number. The FC idea has been developed then, by many mathematicians throughout the eighteenth and nineteenth centuries. Now, there exist several definitions of the fractional-order derivative, including Grünwald-Letnikov, Riemann-Liouville, Weyl, Riesz, and the Caputo representation. In the last decades, using fractional operators provides more accurate models in many branches of science. The scope of the fractional calculus expands day by day and is used in many different applications like in biomedicine, biology, modeling world economies, plane wave diffraction problem and heat transfer process [12-21]. In this study, a new mathematical model is developed by using the fractional calculus methodology. Previously proposed models in our studies, such as the modelling of child growth, subscriber's numbers of operators, Gross Domestic Product (GDP) per capita were modelled and compared with other modelling approaches. In other words, the values of height, weight, and body mass index for the child growth is modelled and compared with Linear and Polynomial Models [22,23]. The number of subscribers of operators in Turkey are modelled and compared with Polynomial Model [24].

According to the results of those studies, our developed fractional models had better performance to the Linear and Polynomial Models [22-24]. Note that, the proposed fractional approach in the study is developed not only for modeling as it is done previously but also for prediction of next coming values.

In this paper, the modelling, testing, and prediction are done for the populations of countries. In this regard, using fractional calculus, USA, France, England, Italy, Spain, and Turkey's populations are tried to be modelled, and the predictions are made for the coming years.

The structure of the article is as follows. The formulation of the problem is expressed in Section 2. In the following section, namely, the solution of the problem, modelling and simulation, testing, and prediction parts are explained in detail. Then, in Section 4, the conclusion is drawn.

2 | FORMULATION OF THE PROBLEM

In this section, the mathematical background is presented. First, it is better to assume an arbitrary function $g(x)$ as the finite summation of previous values of the same function weighted with unknowns coefficients α_k.

$$g(x) \cong \sum_{k=1}^{l} \alpha_k \, g(x-k) \tag{2.1}$$

In the study, x and $g(x)$ would be time and the population, respectively. After assuming (2.1), the function $g(x)$ can be expanded as the summation of polynomials with unknown constant coefficients, a_n as given in (2.2.a)

$$g(x) = \sum_{n=0}^{\infty} a_n \, x^n \tag{2.2.a}$$

Then, $g(x-k)$becomes

$$g(x-k) = \sum_{n=0}^{\infty} a_n \, (x-k)^n \tag{2.2.b}$$

The final form of $g(x)$ is given as (2.3).

$$g(x) \cong \sum_{k=1}^{l} \sum_{n=0}^{\infty} \alpha_k a_n \, (x-k)^n \tag{2.3}$$

Here, after defining $\alpha_k a_n$ as a_{kn} and approximating (2.3), (2.4) is obtained.

$$g(x) \cong \sum_{k=1}^{l} \sum_{n=1}^{M} a_{kn} \, (x-k)^n \tag{2.4}$$

In the following expression given in (2.5), the definition of Caputo's fractional derivative is presented [25]. Throughout the study, Caputo's description of the fractional derivative is employed.

$$\mathfrak{D}_x^{\gamma} g(x) = \frac{dg(x)}{dx^{\gamma}} = \frac{1}{\Gamma(1-\gamma)} \int_0^x \frac{g^{(1)}(k)dk}{(x-k)^{\gamma}} \tag{2.5}$$

$$(0 < \gamma < 1)$$

Here, the derivative is taken concerning x in the order of γ and $g^{(1)}$ stands for the first derivative with respect to x. After giving the expression of expanding an arbitrary function $g(x)$ with its previous values and Caputo's fractional derivative, it is better to express Deep Assessment Methodology by using fractional calculus for modelling and prediction. Here, the fractional derivative of $f(x)$ in the order of γ is assumed to be equal to (2.6).

$$\frac{d^{\gamma} f(x)}{dx^{\gamma}} \cong \sum_{k=1}^{l} \sum_{n=1}^{\infty} a_{kn} . n . (x-k)^{n-1} \tag{2.6}$$

where, $f(x)$ is the population of the countries and x stands for the time.

The main purpose is to find a_{kn} given in (2.6). Therefore, first, the Laplace transform is taken for (2.6) in order to solve the differential equation, then, by using inverse Laplace transform techniques, the final formula for $f(x)$ is given in (2.7a) [25].

$$f(x) \cong f(0) + \sum_{k=1}^{l} \sum_{n=1}^{\infty} a_{kn} \, c_{kn} \tag{2.7a}$$

where,

$$c_{kn} \triangleq \frac{\Gamma(n+1)}{\Gamma(n+\gamma)} (x-k)^{n+\gamma-1}$$

For the numerical calculation, the infinite summation is converted into a finite summation, as given in (2.7b).

$$f(x) \cong f(0) + \sum_{k=1}^{l} \sum_{n=1}^{M} a_{kn} \, c_{kn} \tag{2.7b}$$

Here $f(0)$ and a_{kn} are unknown coefficients to be determined.

3 | THE SOLUTION OF THE PROBLEM

Modeling and Simulation

In this part, the approach for the solution of the problem is given in detail. In order to predict the upcoming years, the problem is divided into four regions as given in Figure 1. Note that, up to end of Region 3 (P_m), the modelling of discrete data is done, in region 4, the prediction is obtained. Therefore, the actual discrete data is known for Region 1, 2, and 3. There are parameters given in the Formulation of the Problem Section such as M, l, γ needed to be determined before the prediction. Here, in Region 1, called as the historical data, the data is known. Due to having a solution approach where a function is represented as the weighted summation of the same function with its previous values as given in (2.1), in Region 1, the actual data is employed. In Region 2, the modelling is done to find the coefficients a_{kn} in (2.7b). For the modeling, Least Squares Method is utilized, which is discussed later in this section. In order to find optimum M, l, γ values for the prediction, Region 3 is needed. In the region, there is an iterative approach where the actual discrete data is known. For example, in Region 3, we want to find $f(m_1 + 1)$. By using the proposed method and Least Squares method, the required $f(m_1 + 1)$ is found with a minimum error by optimizing M, l, γ values. Then, $f(m_1 + 1)$ is used for $f(m_1 + 2)$ and up to $f(m)$. Then, with optimized M, l, γ, $f(m_x)$ is predicted in Region 4.

In order to be more precise, it is better to give details not only for general expressions but also for our specific problem. Let us assume that each of the countries that we will make population modeling has data for each year. In our model, we will have four main regions, as shown in Figure 1. In the figure, P_i corresponds to the discrete known data that we build modelling for it. The first region contains historical data. Each of the coefficients $(x - k)^{n+\gamma-1}$ given in (2.7a) will include the contribution coming from previous values multiplied by coefficients of different weights. That is why we use the concept of deep assessment. The second region is the modeling region in which the population would be modelled, and the unknown coefficients is determined as mentioned above. The third region contains data used to test for upcoming predictions. Finally, the fourth region is the region where we will find the values that we cannot predict yet.

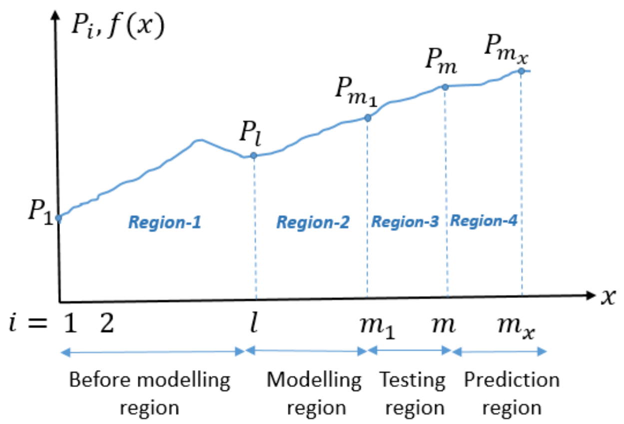

Figure 1: The regions of the dataset.

The $f(i)$ can be expressed as (3.1). By doing so, from the discrete data, continuous curve is obtained. In order to find the unknowns a_{kn} and $f(0)$, the Least Squares Method is utilized.

$$f(i) = f(0) + \sum_{k=1}^{l} \sum_{n=1}^{M} a_{kn}\, c_{kn} \tag{3.1}$$

In (3.2), the squares of total error ${\epsilon_T}^2$ is defined. The main purpose of the modelling region is to minimize ${\epsilon_T}^2$.

$$ {\epsilon_T}^2 = \sum_{i=l}^{m_1} \big(P_i - f(i)\big)^2 \tag{3.2}$$

In order to minimize the square of the total error, the followings are required.

$$\frac{\partial {\epsilon_T}^2}{\partial f(0)} = 0$$

and

$$\frac{\partial {\epsilon_T}^2}{\partial a_{rt}} = 0$$

where, $r = 1, 2, 3, \ldots l$ and $t = 1, 2, 3, \ldots M$

Then, a system of linear algebraic equations (SLAE) is obtained as (3.3).

$$[A].[B] = [C] \tag{3.3}$$

where, [A], [B], and [C] is showed in (3.4), (3.5), and (3.6), respectively.

In order to find the curve with minimum error, optimum γ is searched between (0,1). Then, with optimum γ, the unknown coefficients are determined.

In our study, the populations of the USA, France, Britain, Italy, Spain, and Turkey were used from 1960 until 2018 [26]. Among these, 2013-2018 years were evaluated for the test of our prediction.

Here,

t (years): [1960, 1961, ..., 2018]

i (points): [1, 2, ..., 59]

P_i (value of i): $[P_1, P_2, \ldots, P_{59}]$

P_i: It shows the population of each country in each i^{th} year. For example, P_1 is the population of the country in 1960.

i: It is the corresponding number for each year. For example, $i = 1$ for 1960, $i = 2$ for 1961 and $i = 59$ for 2018.

In order to obtain the optimum modelling, l value is investigated. Then, the value of l is found 6, 15, 6, 16, 16, and 10 for the USA, France, Britain, Italy, Spain, and Turkey, respectively. In other words, for the modeling of the populations of each country, the required previous data of past years used in the algorithm differs. When the test region is considered, the optimal M values are found to be 2, 2, 2, 5, 3, and 4, respectively.

The data related to the modeling were determined with high accuracy, as shown clearly in Table 1. These errors of countries are 0.26%, 0.18%, 0.19%, 0.12%, 0.52%, and 0.15% for USA, France, Britain, Italy, Spain, and Turkey, respectively.

Testing

The predictions obtained in the test region (3rd region) $(m_1 < i < m)$ are also given in Table 1. The data up to $m_1 = 54$ have been taken into consideration in our operations. The $f(m_1 + 1)$ value was found from the obtained modeling. Then, the value is kept and the next step was done again for $(f(m_1 + 2))$.These operations are continued until the final value of the test zone $m = 59$.

Projections for this region, the error of the USA, France, Britain, Italy, Spain, and Turkey values were found as 0.26%, 0.18%, 0.19%, 0.12%, 0.52%, and 0.15%, respectively.

Prediction

The predictions obtained for Region 4 in Figure 1 are given in Table 2. After having a testing region, the unknown coefficients in (3.1) are found with minimum error. Then, the prediction region is started. In the region, the first prediction $f(m + 1)$ is found by using the coefficients calculated by the testing region. Then, the first predicted value $(f(m + 1))$ is also included in the test region for the upcoming prediction $f(m + 2)$. This procedure is repeated up to $f(m_x)$. For example, as of the end of 2020 $(f(m + 2))$, the US, France, Britain, Italy, Spain, and Turkey's populations are expected to be 331601950, 67464467, 67928961, 59442390, 44109433, and 85661569, respectively.

In Figure 2, there is given the flowchart of the deep assessment algorithm. First, the data $(l, M, x_1, x_2, \ldots x_m$ and $P_1, P_2, \ldots P_m)$ is initialized. After that, the variable N is introduced, which counts the number of prediction steps. The total number of predicted steps is n_0. After that, the fractional-order ν is assigned 0, and the increment step is 0.01. For each value of ν between 0 and 1, matrix A given as (3.4) is created according to the part, namely Formulation of the Problem given in the study, and the unknown coefficients are evaluated given in (2.7a). Then, the modelling of data between P_l and P_m is done. After that, the least squares error is calculated. The value of the error is compared to previously obtained values. If it is less than the previous one, we memorized the corresponding fractional-order value. At the end of Loop II, we determine the optimal value of the fractional-order, which corresponds to the best modelling and we memorize corresponding coefficients given in (2.7a). After that, the prediction of the next upcoming value of the initial data with (2.7a) is made. Then, we repeat all the procedures starting from the increment of N so that the predicted value is added to initial data for the next step prediction. This process continues up to the termination of Loop I. As a result, n_0 values are predicted.

$$
A=\begin{bmatrix}
m_1-l+1 & \sum_{i=l}^{m_1} c_{11} & \sum_{i=1}^{m_1} c_{12} & \dots & \sum_{i=1}^{m_1} c_{1M} & \sum_{i=1}^{m_1} c_{21} & \sum_{i=1}^{m_1} c_{22} & \dots & \sum_{i=1}^{m_1} c_{2M} & \dots & \sum_{i=1}^{m_1} c_{l1} & \sum_{i=1}^{m_1} c_{l2} & \dots & \sum_{i=1}^{m_1} c_{lM} \\
\sum c_{11} & \sum_{i=1}^{m_1} c_{11}c_{11} & \sum_{i=1}^{m_1} c_{12}c_{11} & \dots & \sum_{i=1}^{m_1} c_{1M}c_{11} & \sum_{i=1}^{m_1} c_{21}c_{11} & \sum_{i=1}^{m_1} c_{22}c_{11} & \dots & \sum_{i=1}^{m_1} c_{2M}c_{11} & \dots & \sum_{i=1}^{m_1} c_{l1}c_{11} & \sum_{i=1}^{m_1} c_{l2}c_{11} & \dots & \sum_{i=1}^{m_1} c_{lM}c_{11} \\
\sum c_{12} & \sum_{i=1}^{m_1} c_{11}c_{12} & \sum_{i=1}^{m_1} c_{12}c_{12} & \dots & \sum_{i=1}^{m_1} c_{1M}c_{12} & \sum_{i=1}^{m_1} c_{21}c_{12} & \sum_{i=1}^{m_1} c_{22}c_{12} & \dots & \sum_{i=1}^{m_1} c_{2M}c_{12} & x\dots & \sum_{i=1}^{m_1} c_{l1}c_{12} & \sum_{i=1}^{m_1} c_{l2}c_{12} & \dots & \sum_{i=1}^{m_1} c_{lM}c_{12} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\sum c_{lm} & \sum_{i=1}^{m_1} c_{11}c_{lM} & \sum_{i=1}^{m_1} c_{12}c_{lM} & \dots & \sum_{i=1}^{m_1} c_{1M}c_{lM} & \sum_{i=1}^{m_1} c_{21}c_{lM} & \sum_{i=1}^{m_1} c_{22}c_{lM} & \dots & \sum_{i=1}^{m_1} c_{2M}c_{lM} & \dots & \sum_{i=1}^{m_1} c_{l1}c_{lM} & \sum_{i=1}^{m_1} c_{l2}c_{lM} & \dots & \sum_{i=1}^{m_1} c_{lM}c_{lM}
\end{bmatrix}
\tag{3.4}
$$

$$
[B]=[f(0)\quad a_{11}\quad a_{12}\quad \dots\quad a_{1M}\, a_{21}\quad a_{22}\quad \dots a_{2M}\quad a_{31}\quad \dots a_{3M}\quad a_{l1}\dots\quad a_{lM}]^T \tag{3.5}
$$

$$
[C]=\left[\sum_{i=l}^{m_1} P_i\quad \sum_{i=l}^{m_1} P_i c_{11}\quad \sum_{i=l}^{m_1} P_i c_{12}\quad \dots\quad \sum_{i=l}^{m_1} P_i c_{lM}\right]^T \tag{3.6}
$$

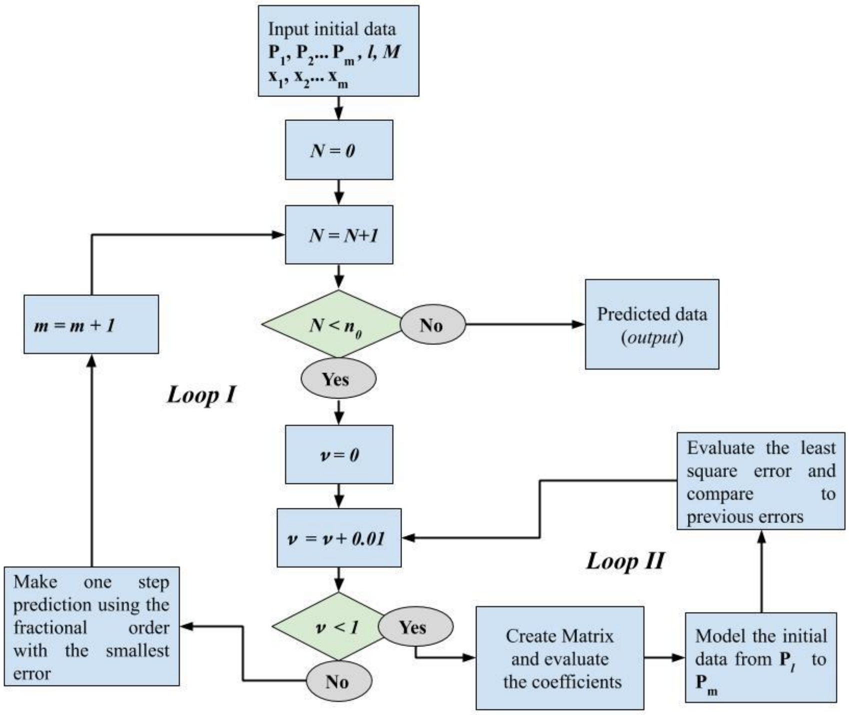

Figure 2: The algorithm for the prediction.

Table 1 Modelling and test results of countries.

Country	USA	France	Britain	Italy	Spain	Turkey
γ **test region**	0.56	0.58	0.99	0.22	0.86	0.25
γ **interpolation**	0.66	0.71	0.96	0.46	0.27	0.98
Deep Assessment $(l < i < m)$	0.26%	0.18%	0.19%	0.12%	0.52%	0.15%
Test Prediction $(m_1 < i < m)$	0.59%	0.45%	0.24%	2.1%	6.66%	1.41%

Table 2 Population Prediction of Countries [2019-23].

Country / Years	USA	France	Britain	Italy	Spain	Turkey
2019	329611042	67354537	67314116	60045141	45135583	84033932
2020	331601950	67464467	67928961	59442390	44109433	85661569
2021	333524876	67535658	68562216	58582961	42826902	87386115
2022	335378595	67566017	69213927	57424815	41266460	89215397
2023	337161950	67553519	69884140	55922601	39406271	91157524

4 | CONCLUSION

Our study showed that the mathematical model of the populations of the countries USA, France, Britain, Italy, Spain, and Turkey had been formed with our deep assessment method using fractional calculus. The approach is also used for the prediction. Considering to express the population as the summation of the previous finite values of the population with weighted unknown coefficients, employing with fractional differential equation and modeling with the Least Squares Method provide to the prediction of the upcoming values of the population with given method and algorithm. It is considered that better results can be obtained if it is modeled by joining other factors that may affect the modelling. In this paper, the modelling, testing, and prediction are done for the populations. It is an initial and unprecedented approach combining the fractional calculus, modeling and prediction. Therefore, the proposed method in the paper is called as the deep assessment method.

ACKNOWLEDGMENTS

This work is supported in part by Istanbul Technical University (ITU) Vodafone Future Lab under Project ITUVF20180901P11.

References

1. Forestier, E., Grace, J., & Kenny, C. (2002). Can information and communication technologies be pro-poor?. Telecommunications Policy, 26(11), 623-646.
2. Kretschmer, T. (2012). Information and communication technologies and productivity growth.
3. Han, E. J., & Sohn, S. Y. (2016). Technological convergence in standards for information and communication technologies. Technological forecasting and social change, 106, 1-10.
4. Grover, V., Purvis, R. L., & Segars, A. H. (2007). Exploring ambidextrous innovation tendencies in the adoption of telecommunications technologies. IEEE Transactions on Engineering Management, 54(2), 268-285.
5. Torero, M., & Von Braun, J. (Eds.). (2006). Information and communication technologies for development and poverty reduction: The potential of telecommunications. Intl Food Policy Res Inst.
6. Pradhan, R. P., Arvin, M. B., & Norman, N. R. (2015). The dynamics of information and communications technologies infrastructure, economic growth, and financial development: Evidence from Asian countries. Technology in Society, 42, 135-149.
7. Beaudry, P., & Portier, F. (2006). Stock prices, news, and economic fluctuations. American Economic Review, 96(4), 1293-1307.
8. Simpson, D. M., Rockaway, T. D., Weigel, T. A., Coomes, P. A., & Holloman, C. O. (2005). Framing a new approach to critical infrastructure modelling and extreme events. International journal of critical infrastructures, 1(2-3), 125-143.
9. Ho, T. Q., Strong, N., & Walker, M. (2018). Modelling analysts' target price revisions following good and bad news?. Accounting and Business Research, 48(1), 37-61.
10. Loverro, A. (2004). Fractional calculus: history, definitions and applications for the engineer. Rapport technique, University of Notre Dame: Department of Aerospace and Mechanical Engineering, 1-28.
11. Machado, J. T., Kiryakova, V., & Mainardi, F. (2011). Recent history of fractional calculus. Communications in Nonlinear Science and Numerical Simulation, 16(3), 1140-1153.
12. Ionescu, C., Lopes, A., Copot, D., Machado, J. T., & Bates, J. H. T. (2017). The role of fractional calculus in modeling biological phenomena: A review. Communications in Nonlinear Science and Numerical Simulation, 51, 141-159.
13. Gómez, F., Bernal, J., Rosales, J., & Cordova, T. (2019). Modeling and simulation of equivalent circuits in description of biological systems-a fractional calculus approach. Journal of Electrical Bioimpedance, 3(1), 2-11.

14. Machado, J. T., & Mata, M. E. (2015). Pseudo phase plane and fractional calculus modeling of western global economic downturn. Communications in Nonlinear Science and Numerical Simulation, 22(1-3), 396-406.
15. Moreles, M. A., & Lainez, R. (2017). Mathematical modelling of fractional order circuit elements and bioimpedance applications. Communications in Nonlinear Science and Numerical Simulation, 46, 81-88.
16. Tabatadze, V., Karaçuha, K., & Veliev, E. I. (2019). The Fractional Derivative Approach for the Diffraction Problems: Plane Wave Diffraction by Two Strips with the Fractional Boundary Conditions. Progress In Electromagnetics Research, 95, 251-264.
17. Sierociuk, D., Dzieliński, A., Sarwas, G., Petras, I., Podlubny, I., & Skovranek, T. (2013). Modelling heat transfer in heterogeneous media using fractional calculus. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371(1990), 20120146.
18. Lefaida, S., Soltani, F., & Mezache, A. (2018). Radar sea-clutter modelling using fractional generalised Pareto distribution. Electronics Letters, 54(16), 999-1001.
19. Tarasova, Valentina V., and Vasily E. Tarasov. "Elasticity for economic processes with memory: Fractional differential calculus approach." Fractional Differential Calculus 6.2 (2016): 219-232.
20. Tarasov, Vasily E., and Valentina V. Tarasova. "Macroeconomic models with long dynamic memory: Fractional calculus approach." Applied Mathematics and Computation 338 (2018): 466-486.
21. Tejado, I., Pérez, E., & Valério, D. (2019). Fractional calculus in economic growth modelling of the group of seven. Fractional Calculus and Applied Analysis, 22(1), 139-157.
22. Önal, N. Ö., Karaçuha, K., Erdinç, G. H., Karaçuha, B. B., & Karaçuha, E. (2019). A Mathematical Approach with Fractional Calculus for the Modelling of Children's Physical Development. Computational and mathematical methods in medicine, 2019.
23. Önal, N. Ö., Karaçuha, K., Erdinç, G. H., Karaçuha, B. B., & Karaçuha, E. (2019). A mathematical model approach regarding the children's height development with fractional calculus. International Journal of Biomedical and Biological Engineering, 13(5), 252-260.
24. Onal, Nisa Ozge, Kamil Karacuha, and Ertugrul Karacuha. "A Comparison of Fractional and Polynomial Models: Modelling on Number of Subscribers in the Turkish Mobile Telecommunications Market." International Journal on Applied Physics and Mathematics, Vol:10, pg:41, 2019.
25. Podlubny, I. (1999). Fractional differential equations, vol. 198 of Mathematics in Science and Engineering, pp.79.
26. World Bank Databank World Development Indicators. Available: https://databank.worldbank.org/source/world-development-indicators# (accessed: December 2019)