Heliyon 8 (2022) e12138

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Forecasting petroleum products consumption in Cameroon's household sector using a sequential GMC(1,n) model optimized by genetic algorithms

Flavian Emmanuel Sapnken [a,*], Khazali Acyl Ahmat [b], Michel Boukar [b], Serge Luc Biobiongono Nyobe [a], Jean Gaston Tamba [a,b]

[a] *Laboratory of Technologies and Applied Science, University Institute of Technology, University of Douala, PO Box 8698, Douala, Cameroon*
[b] *Transports and Applied Logistics Laboratory, University Institute of Technology, University of Douala, PO Box 8698, Douala, Cameroon*

ARTICLE INFO

Keywords:
Forecast
Petroleum products
CO2 emissions
Household sector
Cameroon
Sequential-GMC(1,n)-GA hybrid model

ABSTRACT

Forecasting energy consumption is a major concern for policymakers, oil industry companies, and many other associated businesses. Though there exist many forecasting methodologies, selecting the most appropriate one is critical. GM(1,1) has proven to be one of the most successful forecasting tool. GM(1,1) does not require any specific information and can be adapted to predict energy consumption using a minimum of four observations. Unfortunately, GM(1,1) on its own will generate too large forecast errors because it performs well only when data follow an exponential trend and should be implemented in a political-socio-economic free environment. To reduce these errors, this study proposes a new GM(1,n) convolution model optimized by genetic algorithms integrating a sequential selection mechanism and arc consistency, abbreviated Sequential-GMC(1,n)-GA. Practically, the proposed approach, on one hand, highlights the forecast for petroleum products consumption in Cameroon's household sector. On the other hand, it estimates the amount of CO2 that would be reduced if petroleum products in this sector were switched to clean energy. The new model, like some recent hybrid versions, is robust and reliable, according to the results. Households petroleum products needs by 2025 are estimated to be 150,318 kilo tons of oil equivalent with MAPE of 1.44%, and RMSE of 0.833. Therefore, households GHG emissions would be reduced by 733.85 kilo tons of CO2 equivalent if clean energy was used instead of petroleum products. As a result, the new hybrid model is a valid forecasting tool that can be used to track the growth of Cameroon's household energy demand.

1. Introduction

In 2017, the household sector was Cameroon's largest energy consumer, accounting for 70% of overall energy consumption, thereby over distancing transports (15%) and industries (11%) [1]. It was also the fastest growing sector over the period 2001-2010, going from 1942.78 kilotons of oil equivalent (ktoe) to 3995.57 ktoe [1, 2, 3], although strongly dominated by biomass with a share of 94.74% (3785.6 ktoe), followed by petroleum products (PP) (125.13 ktoe, 3.13%) and electricity (84.84 ktoe, 2.12%) [1]. In the following years, this trend is projected to continue, due to major urbanization projects initiated by the State of Cameroon. Indeed, the State of Cameroon intends to launch colossal projects aimed at making large cities the centers of production and consumption necessary for the industrial sector's development. One of the objectives is to limit deforestation in the tropical regions and desertification in the Sahelian regions [4]. Another objective is to reduce GHGs by popularizing clean and modern energies such as electricity, so that the 40% of the current population having access to electricity become 60% by 2025 [2, 3, 5]. Finally, the State of Cameroon plans to build 17,000 social housing units [6] in the three main cities of Douala, Yaoundé and Bafoussam. Cameroon thus finds itself torn between two major challenges: ensuring the development of its household sector without ignoring climatic issues. To achieve these projects, the certainty on energy consumption prediction is quite important, and many studies have been carried out on this subject with various techniques [7, 8, 9].

1.1. Interest of study

The legal obligation imposed by regulatory authorities is one reason for conducting forecasting studies [10]. In Cameroon's oil market, the

* Corresponding author.
E-mail address: sapnken.emmanuel@gmail.com (F.E. Sapnken).

https://doi.org/10.1016/j.heliyon.2022.e12138
Received 23 February 2022; Received in revised form 26 April 2022; Accepted 26 November 2022

flow of information to predetermined market authorities is mandatory in accordance with regulations [3]. Besides, in cases where distributors cause an imbalance in the supply chain of PPs, related regulations and laws impose a high penalty payment on them [10]. This has the effect of putting pressure on distributors to accurately forecast end-users' consumption, in particular during this period of shutdown of the National Refining Company's (SONARA) production units which is the country's unique refinery [11].

Forecast studies are also important because PPs supply is provided through spot markets [10], where the bulk of the demand is in urban areas. Most households in Cameroon consume the bare minimum of their PP needs, as the latter are only used for lighting (kerosene) and/or cooking (kerosene and liquefied petroleum gas (LPG)) [11]. It is therefore incumbent on the State, and according to the market's structure, to manage its own strategic stocks in order to regulate PP's consumption.

Cameroon ratified the 2016 Paris Agreement [12], which established an international consensus on efforts to decrease GHGs [13]. Thus, although GHGs from Cameroon's energy sector are certainly not as high as those in developed countries, Cameroon has promised to help stabilize GHGs to a level that would limit threatening involvement of human activities with climate [14, 15]. To assist Cameroon's government to fulfill its commitment, this study also estimates CO2 emissions from households' PP consumption, and guides future policies to put in place mechanisms to reduce these emissions without harming energy needs.

1.2. Preceding studies

Energy demand forecasting has already been the subject of several studies in different geographical areas, ranging from a city scale, region, to country level [7, 8, 9]. One of the first models found in the literature is that of Hubbert [16]. In the studies which followed Hubbert's model, ordinary least squares [17, 18, 19], Kalman filters [20]; ARIMA models [21, 22], Box-Jenkins modeling [23, 24], nonlinear regression models [25, 26, 27], artificial neural networks (ANNs) [28, 29, 30], nonlinear mixed effects [31], fuzzy logic [32, 33, 34], genetic algorithms (GA-DEM) [35], and Grey modeling (GM) [36, 37, 38, 39, 40] have also been used in forecasting studies.

GA and GM are different from all the others by their simplicity. They do not need much data. With a minimum of four data points, GM can be applied as well as GA [37, 38, 41]. Unfortunately, GMs alone are not sufficient to forecast energy demand because, very often, forecast results are not sufficiently precise [39, 42, 43] and sometimes, too strong assumptions are made [37]. These assumptions could irremediably lead to unacceptable forecast errors. GA have recently gained the attention of researchers, as they are robust stochastic search algorithms that are used to solve various problems, including forecasting [35, 44, 45]. GA have several considerable advantages over conventional methods. They can provide practical solutions, by analyzing the search space from many starting points, without making any prior assumption about the model or the underlying function [46, 47, 48].

Among existing grey models, GM(1,1) has had success in energy [37], earthquakes [48], finance [49], food production [50], education [51], road safety [52] and transports [53]. However, GM(1,1) works well when data follow an exponential trend and this is only applicable in a political-socio-economic free environment. Yet, such a neutral environment does not reflect the reality of a generalized grey system. This is why data collected from such an environment cannot be modeled with GM(1,1) because there would be incomplete information. To complete this information, more variables must be inserted into the system. This tends to suggest that a multivariate grey model (GM(1,n)) is a way forward to whiten the system. Unfortunately, as demonstrated by Tien [54], GM(1,n) and GM(m,n) all lead to inaccurate modeling. Despite this, we still encounter some erroneous works based on the classical GM(1,n), but there are also many correct works on improved GM(1,n) such as the convolution integral model (GMC(1,n)) [54, 55, 56]. However, GMC(1,n) still has defects [55]. Thus, in order to improve the GMC(1,n), some authors recently proposed additional structural forms. For instance, in order to lower the morbidity and increase the structural stability of the current GMC(1,n) forecasting model, Zhu [57] presented a derived multivariable grey model based on the derivation approach. Ye et al.'s [58] accumulative time-delay multivariate grey prediction model discretizes the time response function's convolution integral using a Gaussian formula, and PSO algorithm determines the best weight coefficients. Ding and Li [59] followed the same process as [58] by effectively approximating the integration of a function of GMC(1,n) using Simpson's method. Still in the previous line of thought, Ding et al. [60] developed a GMC model from a collaborative optimization standpoint. The following are the main innovations of this model. The intrinsic flaws of GMC(1,n) grey models are first addressed by building a collaborative framework that integrates background value optimization, data preprocessing, and model structure enhancement. Second, to increase the adaptability and flexibility of the suggested model, the PSO method is then chosen to find the ideal values of the damping accumulation parameters.

The traditional GM(1,n) can be improved by adjusting its structure to make it more flexible [61], by directly optimizing and stabilizing its parameters [56, 62, 63], and by considering the hysteresis of the related variables [64, 65]. Other lines of thought include new AGO [66], hybridization [39, 67], or by integrating a functional mechanism into the GM's structure [68, 69]. It therefore emerges that no work to date has been carried out on the integration of specific mechanisms into the forecasting framework of GM. This could take into consideration key drivers of energy consumption, by connecting GM with GA.

1.3. This study: objective, contribution and novelty

Motivated by the importance of PP's forecasting in order to replace them with clean and modern energy, this study proposes a sequential grey multivariable convolution model optimized by GA (Sequential-GMC(1,n)-GA). The technical innovation here is to integrate the sequential selection mechanisms in GMC(1,n) and the arc-consistency (AC) in GA as a filter with the aim of reducing CPU execution time. In this perspective, GMC(1,n) is trained by GA in order to enhance all parameters and improve prediction accuracy. Baseline models are used for comparison. The respective precisions of these models are evaluated alongside with forecasts of PP demand by 2025, a decade before the country's anticipated emergence date of 2035 [6].

The fundamental objective of this study is therefore to contribute to accurate forecasting of kerosene and LPG demand for the coming years, and judiciously estimate the quantities of CO2 emissions that could be reduced if these PPs were substituted with clean energy. The contributions are twofold:

i. Sequential-GMC(1,n)-GA is put forward as a new forecasting tool, and is used to optimally determine all GMC's parameters. The proposed model is a modified GMC(1,n) model that incorporates a sequential mechanism algorithm, GA and AC in order to increase the prediction accuracy of the traditional GMC while reducing CPU execution time. This combination reinforces the modeling in three ways: firstly, the model's structure is more flexible and intelligently optimizes the parameters of the basic GMC(1,n) model. Second, the proposed model does not need to pretreat the modeling data. Finally, the sequential mechanism reduces the need for modeling knowledge in the viewpoint of intelligent systems.
ii. Estimate the amount of CO2 that could be reduced if households consumed clean energy, namely hydroelectricity, instead of PP. This is used to evaluate the model's accuracy and viability. Furthermore, we compared the Sequential-GMC(1,n)-GA model to the traditional GMC(1,n) and OGMC(1,n) multivariate grey model using the same dataset and numerical findings for prediction accuracy.

Fig. 1. Reference map of Cameroon.

As novelty, this study proposes a GMC that fully excavates the evolution law of multivariate time series without modifying the model's framework. Sequential-GMC(1,n)-GA proactively improves all model parameters in each projected time-frame, regardless of data patterns. Thus, Sequential-GMC(1,n)-GA eliminates the inconsistency issue between parameterization approach that is based on grey derivative and the model's minimal sum of squares of prediction errors, resulting in more accurate forecasts. Finally, the model is capable of using input data with different sizes and still competes with performant forecasting models.

Section 2 that follows provides an overview of Cameroon's energy sector. Section 3 presents the method. Section 4 presents the application using real data. Results and discussion are presented in Section 5 followed by policy suggestions. Section 6 concludes the study, with a synopsis of future works.

2. Overview of Cameroon's energy sector

Since independence, development planning in Cameroon (**Fig. 1**) has been implemented through reference documents. The Growth and Employment Strategy Paper (GESP) [5] was the last document implemented for the period 2009-2019. GESP expired in 2019. The National Development Strategy (SND30) [6] is the new reference document which takes over GESP for the period 2020-2030 in view of the achievement of emergence objectives by 2035. This periodicity has been chosen to better define the strategy with the world schedule of sustainable development objectives, but also to stabilize planning horizons and integrate delays in the implementation of strategies.

SND30 specific objectives include: (i) the development of the vast national hydroelectric potential; (ii) the development of alternative energies in order to better meet specific needs such as cooking food, lighting, etc.; (iii) strengthening the adaptation and mitigation measures of the effects of climate change and environmental management to ensure economic growth and sustainable social development; and (iv) make large cities economic growth poles and modern urban agglomerations by building 17,000 social housing units.

Primary energy production in 2017 was 10.12Mtoe including biomass (50%), hydroelectricity (7.3%), and oil (37%), of which 5.7% natural gas [11, 70, 71]. Electricity generation is largely dominated by hydroelectric plants[1] (73%). Thermal electricity is strongly dominated by self-production (79%) from diesel [1, 72]; and the remaining 21% in the form of public lighting [72], using diesel and gasoline. Overall installed electric capacity increased from 935MW in 2009 to 1545MW in 2017 [70, 72]. This additional 610MW is basically of thermal origin [70].

Cameroon is a modest oil producer with 72,000 barrels/day averagely in 2019 [72]. Oil production dropped by more than 36% between 1985[2] and 2019 [11]. Consequently, in order to meet the enormous PP demand, SONARA supplies 13% of the local market[3] [1] while 87% is imported from Nigeria. With the ongoing production rate, Cameroon's oil reserves will run out in less than 10 years [11].

Out of 4.5 Mt of crude produced in 2015, merely 409.225 kt were refined by SONARA[4] [11]. The refinery's poor technological profile (hydroskyming or top reforming) and the competitive market force SONARA to rely on other supplies, meanwhile demand for PP keeps rising. This encourages importation to fill the gap.

The oil industry in Cameroon is dominated by a few entities. The National Hydrocarbon Corporation (SNH) is responsible for the promotion, development, and monitoring of oil activities in the upstream sector [72]. SNH was Cameroon's sole crude supplier until the 1985 reforms,[5] but nowadays it competes with PERENCO, ADDAX and PECTEN. SONARA refines crude alone in the downstream sector. Storage is provided by the Cameroonian Company of Petroleum Deposits (CCPD); HYDRAC is in charge of regulating the quality and amount of PP sold in Cameroon [11]; while Marketers (Total, Bocom, Tradex, Greenoil, Ola, Neptune oil, MRS, SCTM, Camgaz, Delta oil and Gulfin) distribute PP under the supervision of the Hydrocarbon Prices Stabilization Fund (HPSF) [11].

PP sold in Cameroon include: diesel, gasoline, LPG, kerosene, jet-A1, and fuel oil [73, 74]. When PP are released from CCDP, they are promptly disseminated on the market through 625 filling stations, with over 65% of these outlets found in major cities [11]. Regarding natural gas, the Kribi power plant was commissioned in 2016 with an initial installed power of 67 Megawatt (MW) (extensible to 216 MW) and that of Logbaba (150 MW) in the Littoral region also works on natural gas [70, 71]. Cameroon's natural gas reserves are estimated at 580 bcm, of which 167 bcm have been proved by SNH [71].

In 2017, in terms of end-use energy consumption, PP ranked second with 19.69% of shares [71], the largest shares were attributed to biomass (72.6%), electricity (7.28%) and others (0.43%) (see Fig. 2a). The household sector outstands all other sectors with 70% of shares [70] (Fig. 2b), transports represent 15% and 9% for other sectors. End-use energy consumption in the industrial sector only represents 5% of the shares, reflecting the low level of industrialization in Cameroon [72].

According to FAO, Cameroon's average annual deforestation rate was 0.6%, 1.0% and 1.07% for the periods 1990–2000, 2000–2005 and 2005–2010 respectively [71]. Agriculture, fuelwood usage, and bushfires are the main causes forest degradation in Cameroon [71, 75, 76].

1 Cameroon's hydroelectric potential is ranked 3rd in Sub-Saharan Africa with about 23 GW [70].

2 The year 1985 marks Cameroon's oil production peak [72].

3 It should be noted that Cameroonian crude is heavy meanwhile SONARA can only refine light crudes.

4 Currently, the SONARA oil refinery has been shut down indefinitely after 4 of its 13 production units were destroyed in a massive fire in 2019. Rehabilitation works will last 4 years [11].

5 The reforms focused on: the removal of SNH monopolies (crude delivery), SONARA (PP supply on the local market), CCPD (PP storage) and the release of distribution margins.

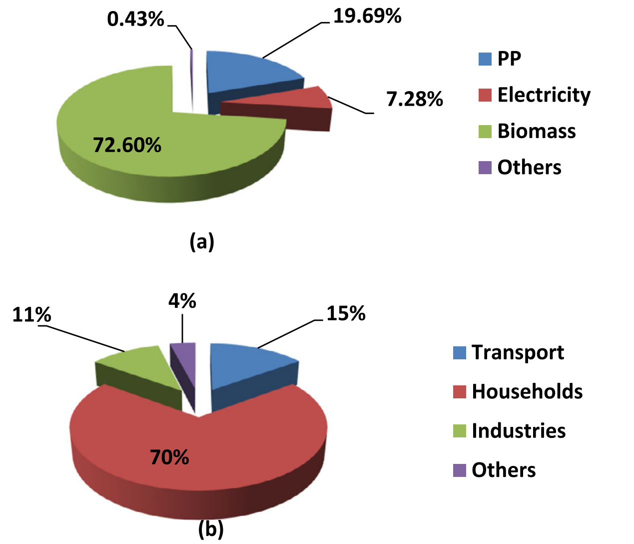


Fig. 2. (a) Total end-use energy consumption in 2017. **(b)** Shares of energy consumption by sector in 2017.

Deforestation in the tropical zone[6] and desertification in the Sahelian zone is very worrisome for the government [4]. Fuelwood usage leads to severe deforestation since reforestation is not practiced. Therefore, fuelwood that was abundant behind kitchens regardless of geographical zones is now scarce. Nowadays, villagers trek up to 35 km daily to fetch firewood [4]. This scarcity could increase in the next years given the number of households that has been growing since the 2000s.

Indeed, the number of households in Cameroon was estimated at 3,142,210 households in 2017 [6]. This number has increased by more than 33% compared to its size in 2005, which was 2,362,000 households [6]. The number of households has been growing at 4.8% on average yearly since 2005, which implies that its size will double in about 15 years from 2017 if it had to keep the same dynamic. Specifically, it could rise to more than 6,348,000 households by 2032. Currently, more than half (53.2%) of these households are located in urban areas. Regions with the most important demographic weights are the Center (19.6%), Far North (18%), Littoral (15.2%) and North (11%) [3, 6]. The relatively large size of the Center and Littoral regions is due to the fact that the political and economic capitals of Cameroon are found in these regions.

LPG and kerosene are the only PP used in Cameroonian households. Kerosene is widely used in rural areas for cooking, lighting and as a fire accelerant, while LPG founds prominent use in urban areas and is only used for cooking [4, 11, 71]. PP access rate in urban areas is 65% compared to only 14% in rural areas and the majority of rural communities are not served [71]. However, what is worrisome is the fact that in addition to being GHG emitters, PP demand (especially LPG) has exceeded supply. And yet, Cameroon is strongly dependent on LPG imports because its production only represents 1% of SONARA's annual production [71]. LPG dependency rate was 65% in 2017, reflecting a strong vulnerability of the LPG supply system [1, 11].

The World Bank's data [77] on GHG inventories show that Cameroon emissions were estimated at 465.99ktCO2-eq on average between 1994 and 2017. Over this period, the transport sector contributed 61% of these emissions compared to 17% for the household sector, 11% for manufacturing industries. The shares of industrial sectors were 9% and 2% for the other sectors. According to these data, the amount of GHGs emitted by Cameroonian households reached a record of 865.82ktCO2-eq in 2015. This represents 22% of total GHGs and is expected to increase in the coming years.

Fig. 3 illustrates the growth of GDP, PP consumption and CO2 emissions from Cameroonian households over the period 1994-2017. With an average growth rate of 6.94% annually, PP consumption in households increased from 21.66ktoe in 1994 to 33.56ktoe in 2000, of which LPG represented more than 70% of the shares [11]. The growth rate of PP consumption was 10.3% (52.2 ktoe) yearly in 2009. In 2015, households demand for PP had reached 87.58 ktoe. This surge is mainly due to the accelerated urbanization development, which is a result of the Cameroonian economy's ongoing expansion.

Faced with the challenges of climate change, which on one hand requires the reduction of GHGs, and on the other hand the adaptation of Cameroon households to modern and clean energies, massive investments will be needed in the coming decades to achieve these objectives [6]. In this context, it is therefore necessary to have methods and tools to estimate GHG emissions and provide more transparency to stakeholders through reporting.

3. Method

3.1. Convolution grey prediction model

GMC(1,n) model is based on six steps [54, 55]:

Step 1: Constructing input sequences

Let $X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}$ be variables of a grey system. $X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}$ are used to construct the input sequence. Each variable in the input sequence is defined by Eq. (1):

$$X_i^{(0)} = \left\{ x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(k) \right\}; \quad i = 1, 2, \dots, n; \ k \geq 4 \tag{1}$$

$X_i^{(0)}$ is a nonnegative sequence; k is the sample number of the ith input variable, and the superscript (0) denotes the original sequences. $X_1^{(0)}$ is the output and $X_i^{(0)}$, $i = 2, 3, \dots, n$ are the variables.

Step 2: Accumulated Generating Operation (AGO)

When $X_i^{(0)}$ is subject to AGO, the following Eq. (2) is obtained:

$$X_i^{(1)} = \left\{ x_i^{(1)}(1), x_i^{(1)}(2), \dots, x_i^{(1)}(k) \right\} \tag{2}$$

where:

$$x_i^{(1)}(t) = \sum_{m=1}^{t} x_i^{(0)}(m); \quad t = 2, 3, \dots, T \tag{3}$$

The superscript (1) in Eq. (3) represents the first order AGO of the original sequences. $x_i^{(1)}(k)$ increases continuously. The sequence is therefore monotonous [37].

Step 3: Designing the background value of the system

The system's background value is defined as in Eq. (4):

$$Z_1^{(1)}(t) = \theta X_1^{(1)}(t) + (1 - \theta) X_1^{(1)}(1 - t); \quad t = 2, 3, \dots, T \tag{4}$$

The horizontal adjustment coefficient θ is taken such that $0 < \theta < 1$ [37, 38]. The value of θ is generally taken as 0.5 but it should be chosen so as to reduce the forecast errors [37].

Step 4: Establishing Grey systems of equations

Grey system is trained by establishing a link between known and unknown sequences:

[6] The vegetation of the ten regions of Cameroon includes: the dense wet forest (found in the South, East, Littoral, Southwest, and Center regions). The semi-desert Sahelian savannah vegetation (found in the Adamawa, North and Far North regions). The grasslands prairies (found in the West and Northwest regions).

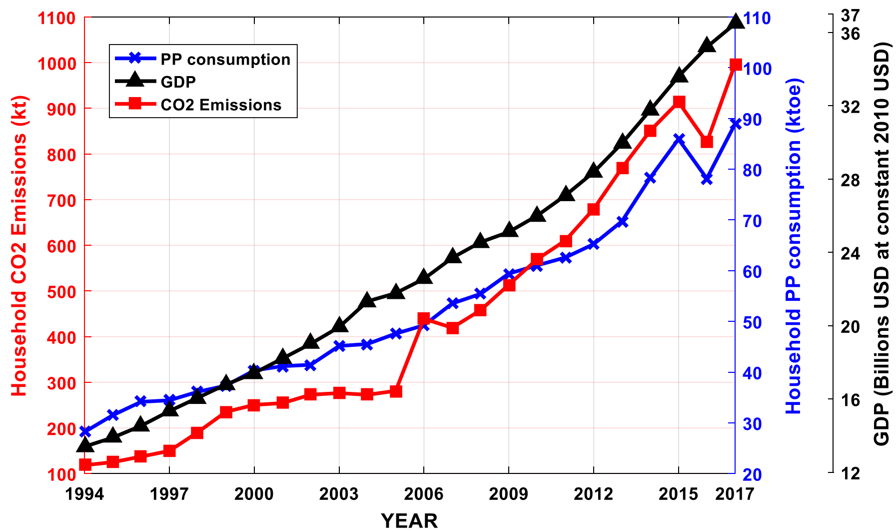


Fig. 3. GDP, PP consumption and household CO2 emissions in Cameroon from 1994 to 2017.

$$dx_1^{(1)}(t)/dt + ax_1^{(1)}(t) = b_2x_2^{(1)}(t) + b_3x_3^{(1)}(t) + \cdots + b_nx_n^{(1)}(t) + u \tag{5}$$

a is the development coefficient, u is GMC(1,n) parameter, while $b_{j=1,\ldots,n}$ are grey input coefficients [43, 44, 49]. Tien [54] starts by considering the right hand side of Eq. (5) as a function $f(t)$. Then, using the trapezoid formula, and integrating both sides of Eq. (5) from $t-1$ to t, Eq. (5) is approximated by the following difference equation [54]:

$$x_1^{(0)}(t) + az_1^{(1)}(t) = b_2z_2^{(1)}(t) + b_3z_3^{(1)}(t) + \cdots + b_nz_n^{(1)}(t) + u \tag{6}$$

As a result, Eq. (6) can be regarded as a linear equation system in relation to the coefficients $[a\ b_2\ b_3\ \ldots\ b_n\ u]^T$. These coefficients are calculated by least squares method [43, 44, 49]. Thus, by applying least squares, $[a\ b_2\ b_3\ \ldots\ b_n\ u]^T$ are determined as shown in Eq. (7):

$$A = \begin{bmatrix} a \\ b_2 \\ b_3 \\ \vdots \\ b_n \\ u \end{bmatrix} = \left(B^T B\right)^{-1} B^T Y \tag{7}$$

Where:

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & x_3^{(1)}(2) & \cdots & x_n^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & x_3^{(1)}(3) & \cdots & x_n^{(1)}(3) \\ \vdots & & & & \\ & \vdots & \vdots & \ddots & \vdots \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & x_3^{(1)}(n) & \cdots & x_n^{(1)}(m) \end{bmatrix}; \quad Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}$$

Step 5: Solving the system's differential equation

The solution to Eq. (5) is [54]:

$$\hat{x}_1^{(1)}(t) = x_1^{(0)}(1)e^{a(1-t)} + \int_1^t e^{a(\tau-t)} f(\tau)d\tau; \quad t \geq 2 \tag{8}$$

Whose value can only be approximated with numerical methods because of the presence of convolution integral. With trapezoid formula, Eq. (8) becomes:

$$\hat{x}_1^{(1)}(t) = x_1^{(0)}(1)e^{a(1-t)} + 0.5h(t)\sum_{i=2}^{n}\left[f(\tau)e^{a(\tau-t)} + f(\tau-1)e^{a(\tau-t-1)}\right]; \quad t \geq 2 \tag{9}$$

Recall that $\hat{x}_1^{(1)}(1) = x_1^{(0)}(1)$, and $h(t)$ in Eq. (9) is defined as: $h(t) = 0$ if $t < 2$; else $h(t) = 1$.

Step 6: Inverse AGO (IAGO)

Finally, forecasted values of $\hat{x}_1^{(0)}(t)$ are obtained by IAGO as shown in Eq. (10):

$$\hat{x}_1^{(0)}(t) = \hat{x}_1^{(1)}(t) - \hat{x}_1^{(1)}(t-1); \quad t \geq 2 \tag{10}$$

3.2. Optimizing GMC(1,n) with GA using arc consistency (AC) technique

In section 1.2, we highlighted that GMC(1,n) is a significant improvement over the basic GM(1,n) in terms of quality. It does, however, have several obvious flaws, as Shen has demonstrated [55]. Refs. [65, 78, 79, 80] demonstrated that k, θ, a and $b_{j=1,\ldots,n}$ affect the precision of GMs. In this study, the best values of these parameters are determined experimentally (instead of using arbitrary values reported in the literature). Thus, a new sequential GMC(1,n) optimized by GA is proposed to address these flaws in order to improve precision, stability, and improve performance in terms of CPU execution time.

3.2.1. Formalism of constraint satisfaction problems (CSPs)

A CSP [81] $\mathcal{P}$ is a quadruplet $\mathcal{P} = (\chi, \mathcal{D}, \mathcal{C}, \mathcal{R})$, where:

- χ represents a finite set of N variables $\{k, \theta, b_{j=1,\ldots,n}\}$.
- $\mathcal{D}$ denotes a set of N domains $\{\mathcal{D}_1, \ldots, \mathcal{D}_N\}$. Each domain $\mathcal{D}_i$ is the finite set of values for the variable χ_i.
- $\mathcal{C}$ represents a finite set of m constraints $\{\mathcal{C}_1, \ldots, \mathcal{C}_m\}$. Each constraint $\mathcal{C}_p$ is defined by the set of variables $Var(\mathcal{C}_p) = \{\chi_{\mathcal{P}_1}, \ldots, \chi_{\mathcal{P}N_\mathcal{P}}\} \subseteq \chi$.
- $\mathcal{R}$ denotes a set of m relations $\{\mathcal{R}_1, \ldots, \mathcal{R}_m\}$. Each relation $\mathcal{R}_\mathcal{P}$ is a subset of the cartesian product $\mathcal{D}_{\mathcal{P}_1} \times \mathcal{D}_{\mathcal{P}_2} \times \ldots \times \mathcal{D}_{\mathcal{P}N_\mathcal{P}}$.

The arity of a constraint is the number of variables on which the constraint relates. A CSP with arity larger than 2 is an N-ary CSP. A solution is the assignment of a value to all variables of the problem such that each constraint is satisfied. CSP is consistent if and only if it admits a solution. A tuple t of the constraint $\mathcal{C}_p$ is said to be allowed if and only if $t \in \mathcal{R}_\mathcal{P}$ and said to be supported for a value $v \in \mathcal{D}_i, \chi_i \in Var(\mathcal{C}_p)$, if and only if t is allowed and t contains v in the position corresponding to χ_i in the constraint. Checking whether a given tuple is allowed by a constraint is called consistency testing. The possible combinations of variable instantiations in a CSP create a search space that can be seen as a search tree.

3.2.2. Arc-consistency and forward checking N-ary

AC remains one of the fundamental properties of CSPs. It guarantees that any value in the domain of a variable has at least one support in any constraint. This property can be established as a preprocessing step or during the search. The basic algorithms are AC1, AC2 and AC3, but AC3 remains the simplest AC known so far [82]. In order to preserve the simplicity of AC3 while improving its efficiency, Ref. [81] proposed AC2000 and AC2001 and asserted that these new ideas can be easily generalized to N-ary versions. AC3 requires the management of a set Q, that stores the remaining revisions to be performed, which in theory corresponds to a set of arcs [83]. However, it is also possible to consider it as a set of variables or constraints [81]. This paper uses AC algorithms to non-binary CSPs with nAC3 N-ary algorithm based on the principle of AC3 [81].

For constraint checking, we used the more popular nFC3 forward search algorithm. When instantiating a variable, nFC3 removes from the current domains of future variables all values incompatible with this instantiation. When a new variable is considered, we can then be sure that all the values of its current domain are consistent with the past variables. nFC3 is quite a well-developed algorithm to present all the contours of its process here. Nevertheless, implementation details are given in [81]. The computational steps of nAC3 N-ary algorithm based on the principle of AC3 is given below:

CSP Algorithm

Start

Step 1: Creat a variable set $\chi = \{\chi_1, \chi_2, \dots, \chi_N\}$

$\chi = \{k, \theta, b_{j=1,\dots,n}\}$

Where:

$$\begin{cases} \chi_1 = k; \chi_2 = \theta \\ \chi_3 = b_1; \dots; \chi_N = b_n \end{cases}$$

Step 2: Creat a domain set $\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_N\}$

Where:

$$\begin{cases} \mathcal{D}_1 = \mathcal{D}_{(k)}; \mathcal{D}_2 = \mathcal{D}_{(\theta)} \\ \mathcal{D}_3 = \mathcal{D}_{(b_1)}; \dots; \mathcal{D}_N = \mathcal{D}_{(b_n)} \end{cases}$$

Step 3: Creat a constraint set with χ and $\mathcal{D}$ after considering the constraints C

$C = \{C_1, \dots, C_m\}$

Where:

$$\begin{cases} C_1 : 4 \le k < \frac{n}{2} \\ C_2 : 0 < \theta < 1 \\ C_3 : b_1 \ne b_2 \ne \dots \ne b_n \end{cases}$$

Step 4: Do while (C_i : $true$, $\forall i$) and ($\mathcal{D}_i$: $true$, $\forall i$)

$$AC(C_{\chi_i,\chi_j}) = \forall \chi_i \in \mathcal{D}_{(\chi_i)}, \exists \chi_j \in \mathcal{D}_{(\chi_j)}, C_{\chi_i,\chi_j}(\chi_i, \chi_j)$$

End while

Output $\chi = \{k, \theta, b_{j=1,\dots,n}\}$

End

3.2.3. Genetic algorithms

We consider a population made of n individuals and we want to optimize $m = 3+n$ parameters, i.e. k, θ, a and $b_{j=1,\dots,n}$. So each chromosome will have m genes. The population $P(t)$ is given by Eq. (11):

$$P(t) = \begin{cases} I_1 = \left(k_1^{(1)}, \theta_2^{(1)}, a_3^{(1)}; b_4^{(1)}; b_5^{(1)}; \dots; b_m^{(1)}\right) \\ I_2 = \left(k_1^{(2)}, \theta_2^{(2)}, a_3^{(2)}; b_4^{(2)}; b_5^{(2)}; \dots; b_m^{(2)}\right) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ I_N = \left(k_1^{(N)}, \theta_2^{(N)}, a_3^{(N)}; b_4^{(N)}; b_5^{(N)}; \dots; b_m^{(N)}\right) \end{cases} \tag{11}$$

Binary forms of k, θ, a and $b_{j=1,\dots,n}$ are obtained from Eq. (12) [46, 84, 85].

$$k, \theta, a \quad \text{or} b_j = L + \frac{\alpha}{2^\beta - 1}(U - L) \tag{12}$$

where α is the number in decimal form which is represented in binary form, β is the number of bits, L and U are the lower and upper threshold values respectively [51]. Selection, crossover and mutation operators are then used (over a series of generations) by a standard GA to guide $P(t)$ towards convergence at the global optimum.

Once selection, reproduction and mutation operators have been applied, new phenotypes (population) are created (which theoretically has a better forecast accuracy). The generational counter is incremented by one. The whole process is reiterated until g_{max} is reached or until the desired forecast accuracy is achieved.

Unluckily, GA are well-known for being CPU-intensive (they sometimes take a long time when executing a given problem). To tackle this issue, we assumed that each individual in $P(t)$ is a CSP that we are attempting to solve using AC, with $P(t = 1)$ serving as the outcome.

3.2.4. Sequential GMC(1,n) prediction model optimized by GA

k, θ, a and $b_{j=1,\dots,n}$ are found in the literature to be 0.5 and 4 for θ and k respectively, while a and $b_{j=1,\dots,n}$ are calculated only once during the modeling stage [37, 43]. However, periodic values of these parameters can significantly increase accuracy of GMC(1,n) [37]. Hence, instead of using values reported in the literature, we rather calculate them experimentally (using AC-GA) for each forecast period h.

Sequential mechanism makes use of the most up-to-date simulations for prediction by depicting the latest characteristics of data. This allows the sequential mechanism to improve forecasting accuracy. This mechanism employs p sequences to model GMC(1,n), and q consumption data for forecasting. At each loop, new values of k, θ, a, $b_{j=1,\dots,N}$ are computed and optimized by GA. The process is described below:

i. Construct GMC(1,n) using $[X_i^{(0)}(1), X_i^{(0)}(2), \dots, X_i^{(0)}(p)]$ and compute $k, \theta, a, b_{j=1,\dots,n}$
ii. Optimize $k, \theta, a, b_{j=1,\dots,n}$ with AC-GA and forecast $[\hat{X}_1^{(0)}(p+1), \hat{X}_1^{(0)}(p+2), \dots, \hat{X}_i^{(0)}(p+q)]$.
iii. Remove $[X_i^{(0)}(1), X_i^{(0)}(2), \dots, X_i^{(0)}(p)]$ from the sequence, reconstruct GMC(1,n) with the most up-to-date p sequences, i.e. $[X_i^{(0)}(q+1), X_i^{(0)}(q+2), \dots, X_i^{(0)}(q+p)]$ and compute new values of $k, \theta, a, b_{j=1,\dots,n}$.
iv. Use AC-GA to optimize newly computed $k, \theta, a, b_{j=1,\dots,n}$, and forecast $[\hat{X}_1^{(0)}(p+q+1), \hat{X}_1^{(0)}(p+q+2), \dots, \hat{X}_i^{(0)}(p+2q)]$.
v. Repeat steps (ii) to (iv) till all required consumption data are forecasted.

This process is called sequential GMC(1,n) prediction optimized by GA (sequential-GMC(1,n)-GA) and is summarized in Fig. 4. The above-mentioned intructions have been written with the AC (steps ii and iv). It is therefore a sequential-GMC(1,n)-GA with AC. By editing the instruction "with AC" in steps ii and iv, it becomes a sequential-GMC(1,n)-GA without AC.

3.3. Energy conversion

Based on PP forecasts, equivalent electricity Q_{elec} (in Gigawatt hour (GWh)) necessary to substitute LPG and kerosene in households is estimated using Eq. (13):

$$Q_{elec} = 0.08598 \cdot \frac{\mu_{PP}}{\mu_{elec}} \cdot Q_{PP} \tag{13}$$

0.08598 is the conversion factor, Q_{PP} is the forecasted PP demand (in ktoe), while μ_{PP} and μ_{elec} represent efficiencies of PP and electric appliances respectively. We use energy efficiency values reported in Ref. [86, 87] who experimentally determined the energy efficiencies of several cooking and lighting household appliances.

3.4. Estimating CO2 emissions

The bottom-up Tier 2 approach is used in the IPCC Guidelines [88] to quantify GHGs from PP combustion based on country-specific emission factors and the amount of PP used in a particular sector [14, 15]. In general, for each sector, GHGs are calculated by Eq. (14).

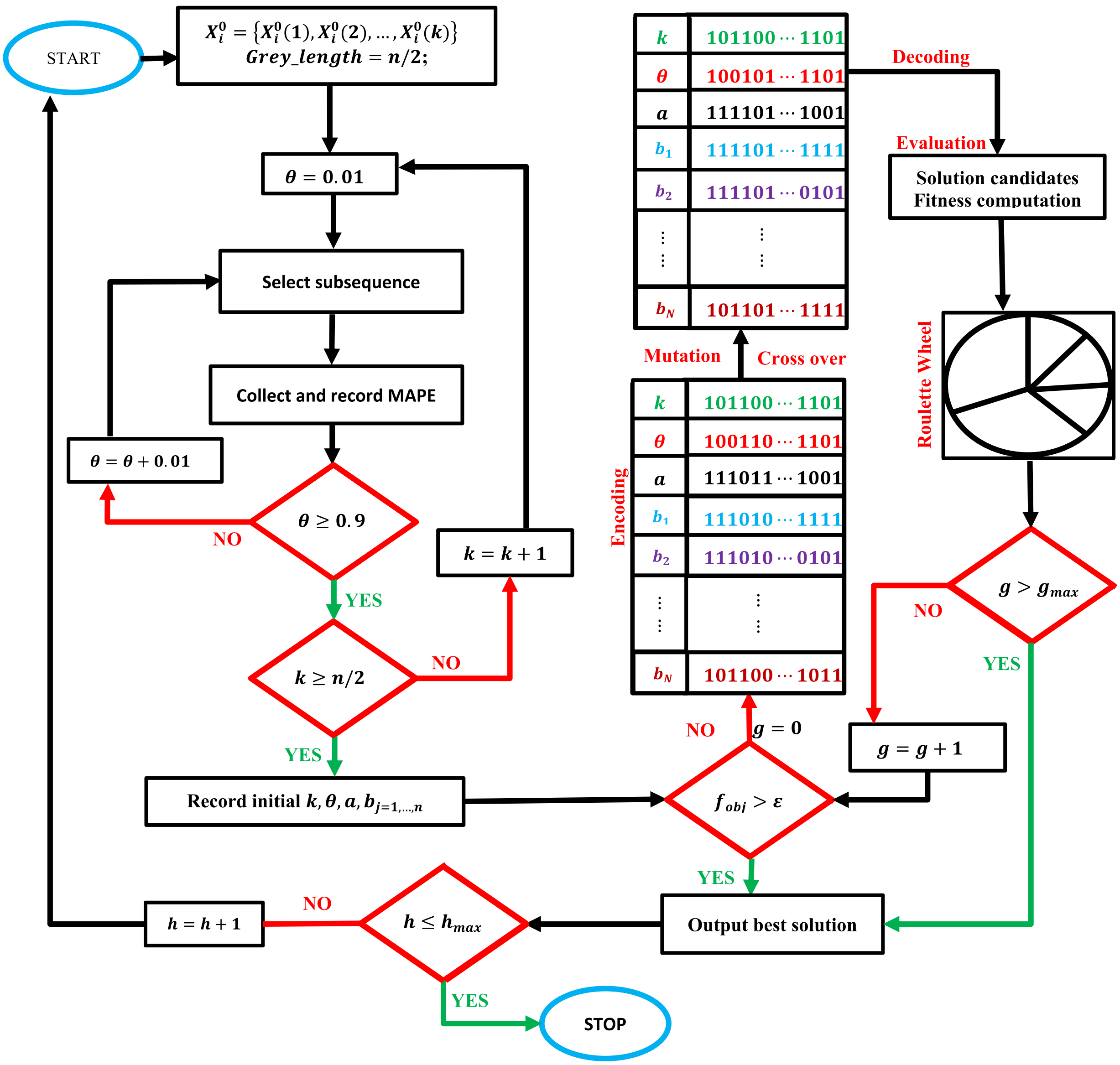


Fig. 4. Flow chart for the proposed Sequential-GMC(1,n)-GA.

$$E = FC_i \cdot EF_{i_{gas}} \tag{14}$$

where E represents GHG emissions in kilotons of CO2 equivalent (ktCO2-eq). FC_i is the quantity of PP consumed (in m^3), $EF_{i_{gas}}$ is the country-specific emission factor (in ktCO2-eq/TJ) per gas. The subscript i throughout this section represents the type of PP concerned.

Global Warming Potential (GWP) compares each GHG's potential to trap heat in the atmosphere relative to another gas. Specifically, GWP measures how much heat will be absorbed by one ton of gas emitted over a particular length of time compared to one ton of CO2 [15, 89]. The greater the GWP for a given gas, the more that gas warms the atmosphere over time compared to CO2.[7] GWP for CO2 is 1, 21 for CH4 and 310 for N2O [15, 89]. Therefore, Eq. (14) becomes:

$$GHG\ emissions = \left(FC_i(t) \cdot \rho_i \cdot LHV_i\right)\left(EF_{i_{CO2}} + 21 EF_{i_{CH4}} + 310 EF_{i_{N2O}}\right) \tag{15}$$

[7] The time period for GWPs is 100 years.

In Eq. (15), $FC_i(t)$ is the amount of PP consumed in year t, ρ_i is the density at 15 °C in Mg/m^3, LHV_i is the lower heating value in MJ/kg. $EF_{i_{CO2}}$, $EF_{i_{CH4}}$ and $EF_{i_{N2O}}$ are emission factors for each gas in ktCO2-eq/TJ. ρ_i, LHV_i and $EF_{i_{CO2}}$ are specific to Cameroon [14, 15]. Default emission factors for LPG and kerosene[8] are obtained from the works of Tamba et al. [14, 15].

4. Application using real data

The first step in building the models is to collect useful variables that are available for this study. Low impact variables should be eliminated and only the most influential ones are considered. Refs. [90, 91, 92] proved that stepwise regression analysis (SRA) [93] is more efficient for variable selection than alternative methods. So, we adopt SRA to select variables. SRA process is detailed in appendix A. All useful variables used for this work are in annual frequency and are presented in Table 1.

[8] These values are those recommended by the IPCC [88] for countries which do not have the necessary technologies and equipment to calculate these emission factors [15].

Table 1. Inputs used in this study.

	Variable	Start	End
X_1	Petroleum products consumption	1994	2017
X_2	Lag 1 of petroleum products consumption	1993	2016
X_3	Lag 2 of petroleum products consumption	1992	2015
X_4	Price of petroleum products	1992	2017
X_5	Real income	1994	2014
X_6	Number of households	2000	2014
X_7	Urbanization rate	1998	2014

4.1. Description of datasets

The modeling period was selected for 1992–2008, while the test period is 2009-2017. Prices and historical consumption were collected from HPSF (https://www.csph.cm) and confirmed by CCPD data (https://www.scdp.cm). Urbanization rate and number of households were provided by the National Institute of Statistics (https://www.minefop.gov.cm), while real income was collected from the World Bank statistics (https://data.worldbank.org).

4.2. Performance measures

Validation criteria were determined numerically (see Table 2). N is the sample size, p is the number of explanatory variables, $X_1(t)$ is actual consumption, $\overline{X}_1$ denotes the mean of X_1, and $sd(X_1)$ represents the standard deviation of X_1. Additionally, $\hat{X}_1(t)$ is the predicted consumption for year t, and $\overline{\hat{X}}_1$ is the average of the predicted consumption. MAPE, r, R^2_{adj} and RMSE are used as criteria to validate the new model. Threshold levels for MAPE, r and R^2_{adj} are given in Table 2. For RMSE and AAE, low values close to 0 indicate the best predictions [10]. However, the reference performance measure is MAPE.

The speed of convergence and stability, i.e. quality and reliability, are used to assess GA's performance. The best solution obtained in relation to the number of cycles (computational cost) is used to determine speed of convergence, while the standard deviation of the results from 10 optimization runs per algorithm is used to determine stability.

Dataset was divided into simulation (or training) set and validation (or test) set to ensure that the model is neither overfitting nor underfitting. Indeed, overfitting and underfitting are the main causes of poor performance of predictive models generated by Machine Learning algorithms. Thus, dividing the data into two as indicated above makes it possible to check the generalizability of the predictive model on the data that it has not yet seen during the learning stage. Training dataset spans from 1992 to 2008 while validation dataset is from 2009-2017.

5. Results and discussion

Simulations are done with MATLAB™R2016a software on a personal computer with AMD Ryzen™ 3 3200U@2.60GHz and 8.0 GB RAM. Sequential-GMC(1,n)-GA is first simulated without AC, then the model is boosted with this filtering technique.

5.1. Results from simulations and model validation

Sequential-GMC(1,n)-GA is initialized with parameters as indicated in Table 3. Once approximate values of $k, \theta, a, b_{j=1,\ldots,n}$ are obtained, they are encoded into binary digits and integrated into GA. Among these parameters, mutation probability must be very low, as low as 0.05 or even smaller [97, 98]. Because a higher value could destroy the solution. P_m is set as the inverse of chromosome length. Crossover probability P_c depends on the problem at hand. Its optimal value is estimated after different runs based on the targeted precision ϵ. If optimal values of P_m and P_c are settled, then number of generations and population size are simply set arbitrarily. However, a population size of 100 and 50 generations is advised [99].

The optimal values of $k, \theta, a, b_{j=1,\ldots,n}$ are presented in Table 4. Experimentally, $k = 4$ and 5 for kerosene and LPG models respectively, unlike the theoretical value of $k = 4$ reported in the literature. Likewise, instead of the theoretical value of $\theta = 0.5$, simulations reveal that θ is slightly higher/lower for LPG/kerosene. As for a and $b_{j=1,\ldots,n}$, their values given by each model are all different (taken in pairs) although not exceeding 16% deviation.

Without AC, execution takes 97 seconds and consistency is achieved at the 34$^{\text{th}}$ generation. Thereafter, each chromosome in $P(t)$ is assumed to be a CSP while each gene is considered as a variable. The global optimum is obtained at the 13$^{\text{th}}$ generation after generating the CSP at random and removing inconsistent values from the variable domains. The new execution takes 5 seconds. So, AC technique reduced execution time by 92 seconds. This gain also explains from 34 to 13 generations. As a result, GA needed fewer iterations to converge to a global optimum.

AAEs for each model are shown in Figs. 5(a-c) and Figs. 6(a-c). These figures show that prediction deviations given by Sequential-GMC(1,n)-GA are insignificant compared to GMC(1,n) and OGMC(1,n). Graphs of real and predicted consumption of all three models are shown in Figs. 10 and 11, with annual residual curves. The residual dynamics, which are not explicitly revealed in each model, can be analyzed using these graphics. Forecast curves produced by Sequential-GMC(1,n)-GA almost fit to perfection with real data. Moreover, annual residuals curves of Sequential-GMC(1,n)-GA are much more spread over time than GMC(1,n) and OGMC(1,n). Therefore, Figs. 10 and 11 confirm the results given by Figs. 5(a-c) and Figs. 6(a-c).

As a result, these graphs (Figs. 5(a-c) and Figs. 6(a-c)) enable us to study the residual dynamics–which the Sequential-GMC(1,n)-GA, GMC(1,n) and OGMC(1,n) models do not instantly reveal–and to contrast the suggested models in terms of the temporal distribution of the residuals. To further show that the dynamics of the time series have been modeled, autocorrelation (AC) and partial autocorrelation (PAC) plots of the residuals are added to the residual plots in Figs. B.12 to B.17 (in Appendix B). Performance statistics are presented in Tables 5 and 6. According to these tables, Sequential-GMC(1,n)-GA significantly outperforms OGMC(1,n) and GMC(1,n) regarding all indicators presented in section 4.2, especially for MAPE. Figs. 5(a-c) and Figs. 6(a-c) confirm the superiority of Sequential-GMC(1,n)-GA on GMC(1,n) and OGMC(1,n).

Table C.7 allows comparing between Sequential-GMC(1,n)-GA and other hybrid GM based models. On the basis of MAPE, RMSE and AAE, we can conclude that Sequential-GMC(1,n)-GA offers better precision and much more reliable forecasting capabilities. Sequential-GMC(1,n)-GA achieves these outstanding results by accounting for all stimuli of PP consumption, all of which are a priori defined by a significant link with demand.

Further tests and comparisons show that Sequential-GMC(1,n)-GA can perform as well as some new expert systems. Table C.8 shows that Sequential-GMC(1,n)-GA can compete with VMD-EELM (Variational Mode Decomposition-Evolutionary Extreme Learning Machine) [100], VMDSVM-PSO (VMD coupled to SVM and improved by PSO) [101], and VMD-SRSVRCBCS (VMD hybridized with Self recurring mechanism and SVR, and optimized by CBCS) [102].

Ultimately, we compare the above findings to those of other research that meet the same performance requirements. The standard GM(1,n) model, as shown in Table C.9 and validated by Shen et al. [55], performs well only when input data is monotonically increasing and is the worst for other data characteristics. The GMC(1,n) model improves the standard GM(1,n) model in terms of quality, but as demonstrated by Tien [54] and Shen et al. [55], the effect is not perceptible once the issue parameter estimation and parameter application mismatch arises. Both models proposed by Wu and Zhang [65] and Ding and Li [59] successfully tackle the issues with MAPEs of 1.91% and 2.48% respectively. The model structure, on the other hand, is rather basic, and the forecasting accuracy might be enhanced. The novel Sequential-GMC(1,n) model (with 1.44% MAPE) effectively solves the GMC(1,n) model's dif-

Table 2. Accuracy measures and threshold levels [10, 37, 94, 95, 96].

Criteria	Formula	Threshold levels			
		1st (Perfect)	2nd (Good)	3rd (Acceptable)	4th (Poor)
$MAPE^{\ddagger}$	$\frac{1}{N}\sum_{t=1}^{N}\lvert\frac{X_1(t)-\hat{X}_1(t)}{X_1(t)}\rvert\times 100$	≤ 0.01	≤ 0.05	≤ 0.1	> 0.1
r^{*}	$\frac{\frac{1}{N-1}\sum_{t=1}^{N}(X_1(t)-\overline{X}_1)(\hat{X}_1(t)-\overline{\hat{X}}_1)}{\sqrt{sd(X_1)^2 sd(\hat{X}_1)^2}}$	≥ 0.98	≥ 0.95	≥ 0.90	< 0.90
R^2_{adj} **	$1-\frac{N-1}{N-p-1}(1-\frac{\sum_{t=1}^{n}(X_1(t)-\hat{X}_1(t))^2}{\sum_{t=1}^{n}(X_1(t)-\overline{X}_1(t))^2})$	≥ 0.98	≥ 0.95	≥ 0.90	< 0.90
$RMSE^{\dagger}$	$\sqrt{\frac{1}{N}\sum_{t=1}^{N}(X_1(t)-\hat{X}_1(t))^2}$				
$AAE^{\dagger\dagger}$	$\frac{1}{N}\sum_{t=1}^{N}\lvert X_1(t)-\hat{X}_1(t)\rvert$				

‡ mean absolute percentage error, * coefficient of correlation, ** adjusted r, † root mean squared error, †† average absolute error.

Table 3. Initialization parameters.

Approach	Parameter	Value
GMC(1,n)	horizontal adjustment coefficient, θ	0.01
	k	4
	Number of variables, n	7
	Grey length	12
Approach	Parameter	Value
GA	Population size	100
	Maximum number of generations	50
	Chromosomes length	83
	Mutation probability, P_m	1/83
	Crossover probability, P_c	0.6
	Elitism	Activated
	Precision, ε	0.05

ficulties of parameter mismatching and relatively basic model structure, while GA helps it outperform all competing models, regardless of input data's characteristics.

Finally, a comparison was made with similar optimization algorithms, namely: Differential evolution (DE) [103] and Fruit Fly optimization Algorithm (FOA) [104] (see Table C.10). These algorithms were tested with the same dataset as that in this study in order to have a fair basis for comparison. The criteria for comparison are: convergence accuracy (given by MAPE), convergence speed (in seconds) and convergence stability (given by the standard deviation for 10 runs). From the results presented in Table C.10, it appears that FOA-F-SVM has the best convergence accuracy. As concluded by [104], this is due to the fact that FOA-F-SVM algorithm considers both the radius-margin error bound and the parameters optimization. This algorithm still comes first in terms of stability and speed of convergence and thus outperforms all others including that proposed in this study. However, we note that GAs coupled with AC as suggested in this study manage to outperform the other competing algorithms thereby ranking second (especially when considering the convergence speed and stability criteria). Therefore, this further convinces us that the proposed method can be very useful in optimization problems.

5.2. Results from convergence and stability tests

We ran 10 trials of Sequential-GMC(1,n)-GA with and without AC, and the average value and standard deviation was computed to compare the speed of convergence and stability of the method. The results are shown in Figs. 7–9.

Using LPG data, after 50 evaluations of Sequential-GMC(1,n), GA without AC finds mean solution at the 21st evaluation with MAPE of 4.622%. Using AC techniques combined with GA, mean solution is achieved at the 11th generation with MAPE of 1.178% (Fig. 7). Sequential-GMC(1,n)-GA without AC exhibit early convergence, and Sequential-GMC(1,n)-GA with AC catches up at around 7 evaluations. This better converging algorithm also displays high stability (see Fig. 9).

With kerosene data, after 50 evaluations, Sequential-GMC(1,n)-GA without AC finds a solution at the 30th evaluation with MAPE of 1.4% whereas Sequential-GMC(1,n)-GA with AC technique finds a solution at the 13th evaluation with 0.388% MAPE (see Fig. 8). As for stability, Fig. 9 reveals a large difference between Sequential-GMC(1,n)-GA with AC and Sequential-GMC(1,n)-GA without AC. It is already well known that GA slows down its convergence speed if it needs to fine-tune weights and biases. However, there are two possible explanations why Sequential-GMC(1,n)-GA with AC outperforms Sequential-GMC(1,n)-GA without AC. First, the threshold to stop GA without AC is too low for this task so that efficiency was lost when a plateau is encountered. The second reason is that Sequential-GMC(1,n)-GA with AC is faster even in the initial search of near optimal solutions.

5.3. Prediction of Cameroon's households PP demand

Household LPG and kerosene demand forecast for the period 2018–2025 is reported in Table C.11. The results show that LPG demand will continue to increase at a rate of 4.9% yearly for the next few years. LPG demand will reach 150,212 MT by 2025, exceeding the current national production of 48,000 MT/year. With this overwhelming gap, domestic production alone will not meet household needs.

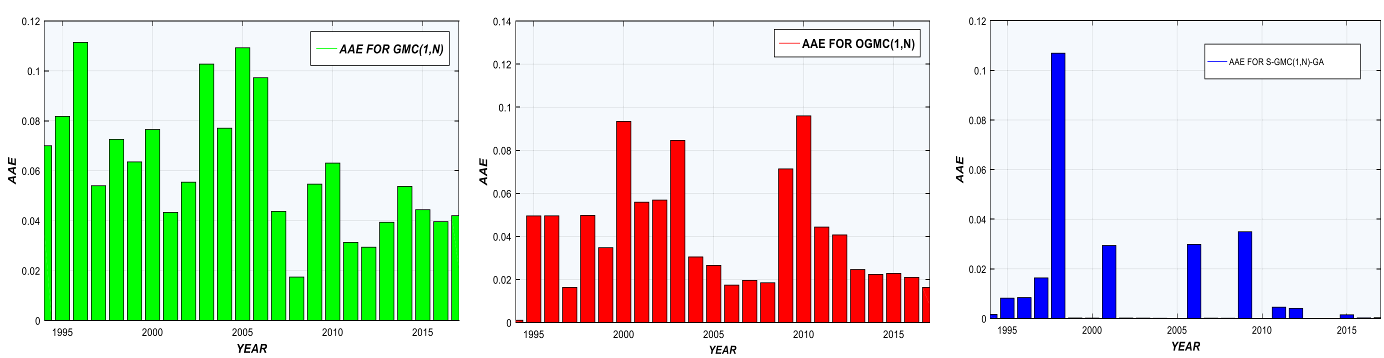
Rather, the dynamics of kerosene demand will follow the opposite trend. Kerosene demand will decrease very slowly (about −0.61% yearly) to reach 106,159 m^3 by 2025. However, consumption will stabilize after 2022, and wave around 100,000 m^3 beyond 2025 which corresponds to the amount of kerosene consumption in rural areas [3]. Given that LPG demand is more prominent in cities than in the backcountry, it is therefore undeniable that in the coming years, households LPG demand will largely take precedence over kerosene demand, especially in urban areas.

5.4. Forecast of equivalent electricity demand

If LPG and kerosene consumption in Cameroonian households were substituted by hydroelectricity, the corresponding electricity needs for the period 2018–2025 will be as reported in Table C.12. This table shows that equivalent electricity will increase by 12.8% on average yearly to reach 1426.14 GWh in 2025, which is more than twice compared to 2018. Guefano et al. [69] predicted that households electricity needs will be 1466.04 GWh in 2023. Therefore, in case of energy transition, a supplementary 990.72 GWh should be provided for households as from 2022 added to Guefano's et al. predictions. Overall, households' electricity needs could reach 2456.76 GWh in 2023. According to Table E.10 and Guefano's et al. predictions, electricity demand in Cameroon households could amount to 3301.7 GWh by 2025. These findings are critical for energy planning and climate issues. As a result, if LPG and kerosene in Cameroon households are substituted by electricity, the government must take sufficient preparations to fulfill future massive energy demands.

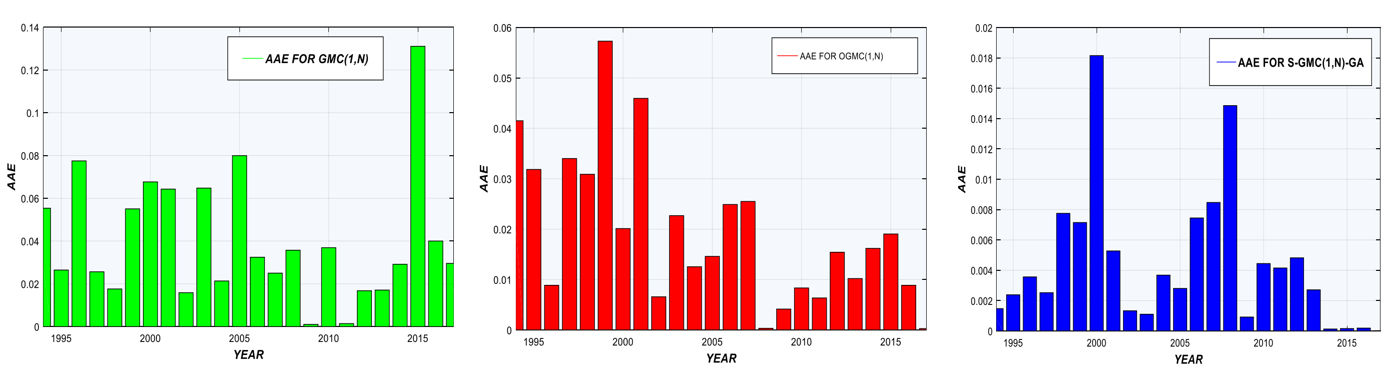
Table 4. Chromosome creation and parameter optimization.

Parameter	Modeled value		*L-U*	Genetic structure	Number of genes	Optimal value given by AC-GA
	GMC(1,n)	OGMC(1,n)				
Kerosene model						
θ	0.51	0.48	0.01-0.99	111	3	0.49
k	4	4	4-10	1111	4	4
$-a$	0.05	0.08	0.01-0.99	111111	6	0.05
b_1	34.1	29.6	0-40	1111111111	10	31.8
b_2	38.7	30.3	0-45	1111111111	10	35.8
b_3	27.4	23.4	0-36	1111111111	10	25.1
b_4	51.3	48.3	0-59	1111111111	10	49.1
b_5	53.7	48.7	0-55	1111111111	10	50.1
b_6	18.7	14.9	0-20	1111111111	10	16.9
b_7	23.4	22.2	0-25	1111111111	10	23.7
			Length of chromosome:		83	
LPG model						
θ	0.47	0.45	0.01-0.99	111	3	0.51
k	4	5	4-12	1111	4	5
$-a$	0.03	0.02	0.01-0.99	111111	6	0.08
b_1	24.2	30.1	0-31	1111111111	10	27.4
b_2	20.7	33.5	0-35	1111111111	10	25.5
b_3	39.3	31.7	0-41	1111111111	10	36.2
b_4	43.1	39.2	0-45	1111111111	10	39.1
b_5	41.5	39.3	0-42	1111111111	10	35.2
b_6	30.8	28.2	0-32	1111111111	10	25.1
b_7	49.6	41.2	0-51	1111111111	10	49.3
			Length of chromosome:		83	



(a) **(b)** **(c)**

Fig. 5. AAEs of **(a)** GMC, **(b)** OGMC, **(c)** S-GMC(1,n)-GA model for LPG training and test data.



(a) **(b)** **(c)**

Fig. 6. AAEs of **(a)** GMC, **(b)** OGMC, **(c)** S-GMC(1,n)-GA model for kerosene training and test data.

Table 5. Performance of GMC(1,n), OGMC(1,n) and S-GMC(1,n)-GA for annual LPG demand prediction.

Measure	GMC(1,n)		OGMC(1,n)		S-GMC(1,n)-GA	
	Train	Test	Train	Test	Train	Test
MAPE (%)	7.1689	4.5205	6.0276	4.9891	1.3445	0.727
r	0.9947	0.9979	0.998	0.9975	0.9993	0.9999
R^2_{adj}	0.9772	0.9834	0.9761	0.972	0.9828	0.9916
RMSE	2.588	3.309	0.157	3.5821	0.8103	0.8325
AAE	0.2368	0.127	0.1336	0.1321	0.422	0.172

Table 6. Performance of GMC(1,n), OGMC(1,n) and S-GMC(1,n)-GA for annual kerosene demand prediction.

Measure	GMC(1,n)		OGMC(1,n)		S-GMC(1,n)-GA	
	Train	Test	Train	Test	Train	Test
MAPE (%)	6.9332	4.0314	4.9014	1.9429	1.6147	0.4396
r	0.9861	0.9975	0.9875	0.9994	0.9988	1.0000
R^2_{adj}	0.9354	0.9712	0.9365	0.9725	0.9610	0.9997
RMSE	4.3026	6.4998	1.9364	3.2911	8.273	8.308
AAE	0.2792	0.1534	0.1628	0.1204	0.6583	0.135

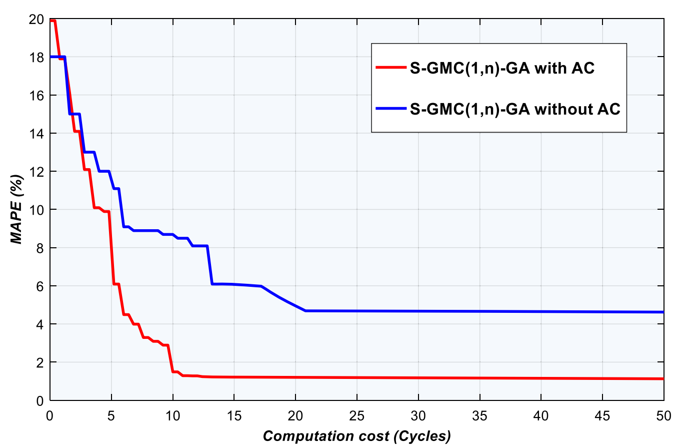


Fig. 7. Convergence of S-GMC(1,n)-GA model with and without AC using LPG data.

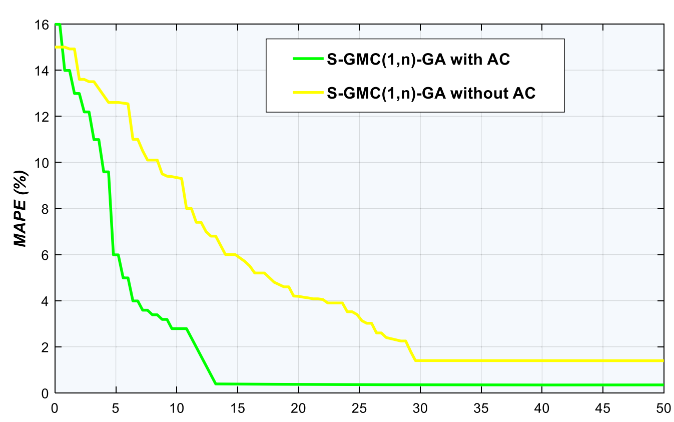


Fig. 8. Convergence of S-GMC(1,n)-GA model with and without AC using kerosene data.

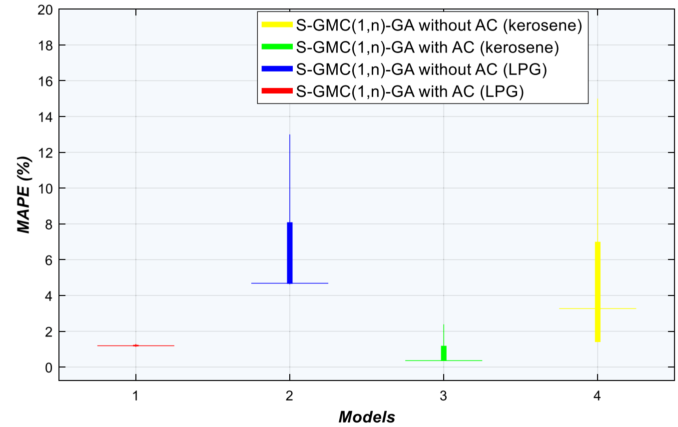


Fig. 9. Boxplot of minimum MAPE after 50 evaluations from 10 runs.

5.5. Estimation of reduced CO2 emissions

CO2 emission estimates in Cameroon households between 2018 and 2025 are presented in Table C.13. If no clean energy transition is undertaken, total GHGs from Cameroonian households would increase from 605.04ktCO2-eq in 2018 to 733.86ktCO2-eq by 2025. At the same time, GHGs from LPG and kerosene consumption would rise from 90.79 and 514.46ktCO2-eq in 2018 to 205.49 and 528.37ktCO2-eq in 2025 respectively.

Although kerosene demand is expected to decrease considerably during the period 2018–2025, its consumption alone will contribute on average to 80% of total GHGs from Cameroon's household, while LPG consumption will contribute to the remaining proportion. Although LPG's emissions shares are quite low over this period, it is important to note that LPG shares in households PP demand will be 60% by 2025 [4, 6, 11]. Consequently, LPG's emissions shares in total households GHGs will be much higher in the coming years. In view of the above, if LPG and kerosene demand in Cameroonian households were replaced by a clean energy source, the amount of GHG reduced could be 733.86ktCO2-eq by 2025.

Under these circumstances, what strategies must be taken to meet households' energy needs and reduce CO2 in the same move?

5.6. Policy implications

Given the severe PP shortage, the cessation of SONARA's refining activities, and policies currently being pursued by the Cameroonian government aimed at limiting kerosene consumption in households, it is proper time for policy makers to start developing the electricity sector. Electricity needs in case of energy transition before 2025 will be as shown in Table C.12. Unfortunately, access to electricity in Cameroon is still insufficient and unaffordable for many localities. Increasing electricity production and accessibility is therefore necessary to ensure economic and social development.

To get there, the Cameroonian government will need to devise an efficient strategy for the development of the electricity sector in order to address the country's deficits. For this, hydropower has many advantages [105, 106, 107] and there are many hydro sites throughout the country [70, 108, 109]. Ultimately, the success of development plans in the electricity sector will make it possible to meet household energy needs, limit deforestation and desertification in tropical and sahelian regions respectively, and above all considerably reduce household CO2 emissions.

To achieve this goal, the government has to make huge investments. However, financing is a major obstacle to achieving low carbon transition. Although several strategies are feasible such as the sale of carbon credits [13, 110], promotion of renewable energy via local communities [111], microfinancing etc., there is still a financing gap that must be

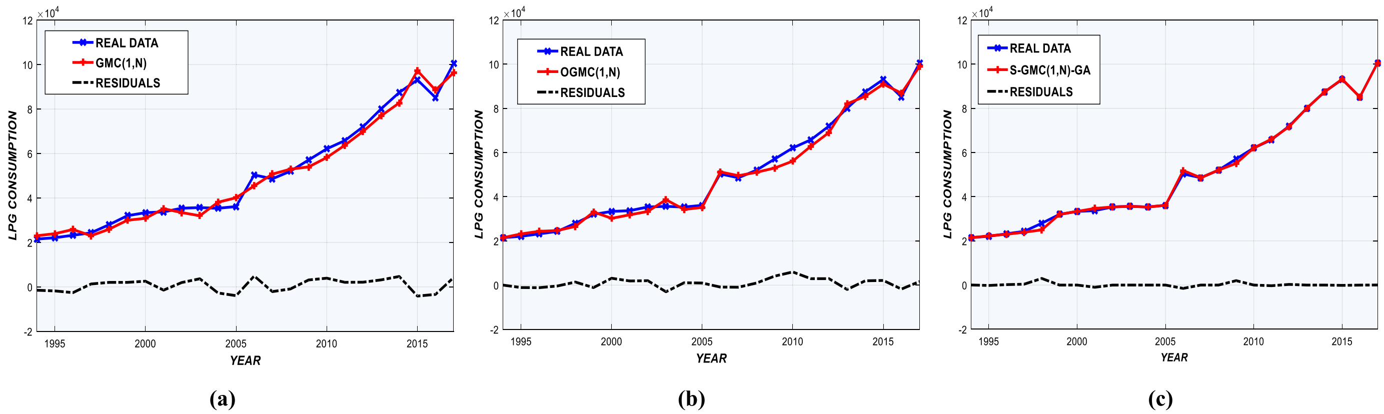


Fig. 10. Observed-predicted LPG consumption and residuals for **(a)** GMC, **(b)** OGMC, **(c)** S-GMC(1,n)-GA.

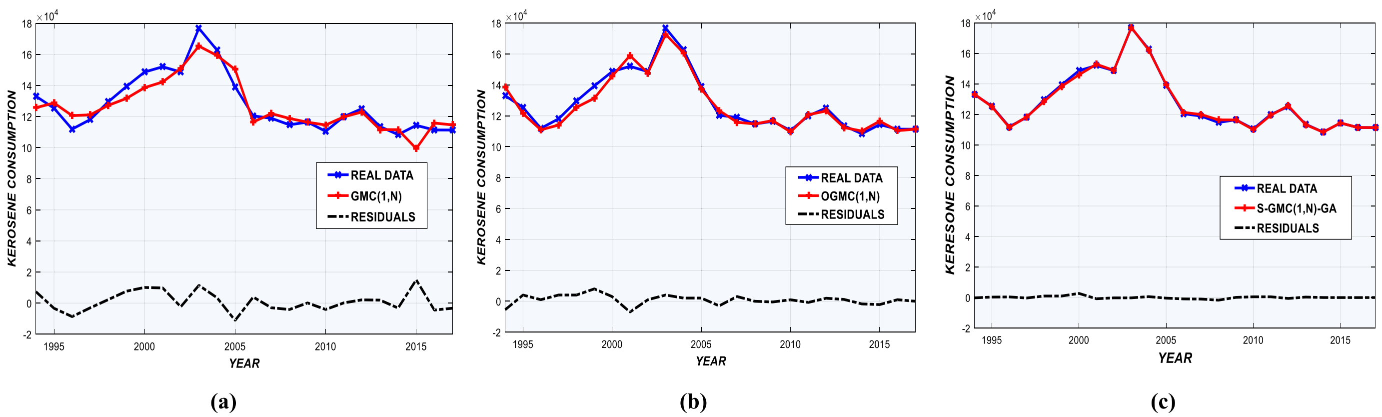


Fig. 11. Observed-predicted kerosene consumption and residuals for **(a)** GMC, **(b)** OGMC, **(c)** S-GMC(1,n)-GA.

filled urgently [111]. Financial institutions could play a role by offering soft loans to people who want to invest in hydropower or by creating and facilitating capital market programs. [111]. In this way, leading technology companies (such as Siemens, Vestas and Philips lighting), and oil firms will influence the pace of this transition [111].

6. Conclusion

In order to curb global warming and climate change without harming energy needs, it is more realistic to bring up solutions from each sector rather than proposing idealistic solutions to all sectors [39]. Thus, the main contribution of this article is to model PP demand in a Grey system to obtain reliable forecasts using limited data. This study takes the Cameroonian household sector as a case study because it is one of the most energy intensive end-users. Also, guaranteeing the correctness of CO2 emissions estimates is a critical aspect in the fight against climate change and global warming. Efficient demand prediction for PP demand paves the way for effective CO2 mitigation strategies and regulations. Modeling however becomes problematic due to a lack of data. To overcome this difficulty, this study develops a new hybrid forecasting model with a minimum of four data.

PP consumption models were produced with training datasets from 1992 to 2008, while validation datasets include data for the period 2009–2017. The results show that Sequential-GMC(1,n)-GA hybrid model is suit to predict energy demand than competing models. This superiority results from two facts: first, instead of using theoretical values of θ and k, all parameters $(\theta, k, a, b_{j=1,\dots,n})$ are determined experimentally and optimized before each forecasting period. Second, the use of a sequential method and AC gathers the most recent data characteristics, reducing forecasting errors and CPU execution time. As a result, this study demonstrates that GM's performance can still be improved.

Using Sequential-GMC(1,n)-GA, total households PP demand is estimated for the period 2018–2025 and CO2 emissions associated with this demand are estimated. The essence of estimating household CO2 emissions is to offer a baseline that policy makers may use to meet household energy needs and achieve CO2 emission reduction targets. On the basis of results presented in this study, the Cameroonian government's attention is drawn on the tremendous rise in PP demand and CO2 emissions for the coming years. It is therefore urgent to start considering the development of Cameroon's electricity sector. This is possible given the enormous hydro potential added to the fact that it is much cleaner than PP. There is, however, the obstacle of funding.

The findings of this study suggest a number of possibilities for using Sequential-GMC(1,n)-GA to other types of energy forecasting and even beyond energy. To begin with, if the series have a high degree of periodicity, volatility or if there are price shocks, Sequential-GMC(1,n)-GA will be unable to completely extract the evolution law due to the presence of disturbing characteristics. This problem can be solved by including a term that accounts for nonlinearities in a time-varying GMC(1,n). Second, CSPs in extension are memory intensive, it would be interesting to represent them in the best possible way, by moving on to intention representation. It may be necessary to develop new N-ary heuristics that are both performant and sufficiently diverse to improve efficiency. Finally, for a multivariable approach, OLS method for calculating constrains working with colossal matrices. This problem can be solved by using numerous data transformations. Thus, Sequential-GMC(1,n)-GA has many aspects where it can be improved. Overall, the model has good adaptability and feasibility, and can extract a grey system's evolution law.

Declarations

Author contribution statement

Flavian Emmanuel Sapnken: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; contributed reagents, materials, analysis tools or data; Wrote the paper. **Khazali Acyl Ahmat, Michel Boukar, Serge Luc Biobiongono Nyobe:** Conceived and designed the experiments; Analyzed and interpreted the data. **Jean Gaston Tamba:** Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

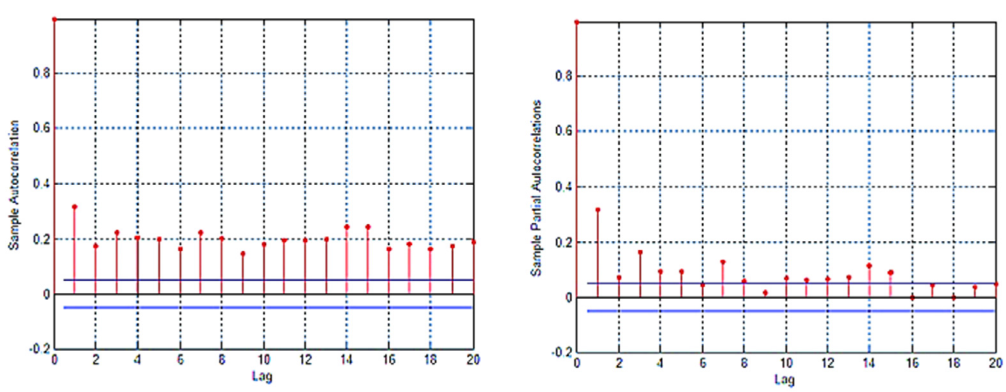
Additional information

No additional information is available for this paper.

Appendix A. Stepwise regression analysis (SRA)

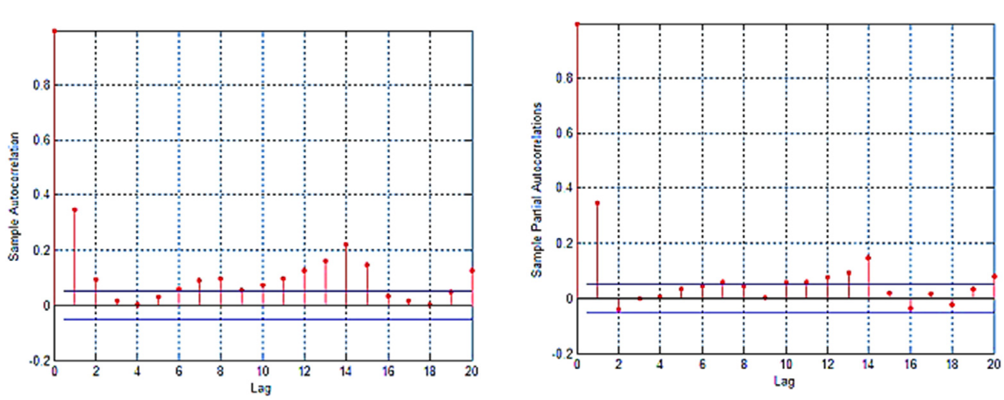
For a given modeling problem, useful variables must be separated from the unnecessary ones. SRA [93] defines predictor variables that most accurately characterize the variable to be predicted [112]. For this, we start by regressing the dependent variable on a single independent variable and set the significance level and values of F_r (F-to-remove) and F_e (F-to-enter) [112]; F being the Fisher statistic. The statistical significance of the F-test and the decrease in the sum of the squared error constitute the benchmark for adding or eliminating a variable [93, 112]. After inserting a new variable in the regression model, the partial F-value is calculated and compared to F_e and F_r; if $F > F_e$, then we include this variable; otherwise, if $F < F_r$, it is eliminated and will never return to the model again [93].

Appendix B



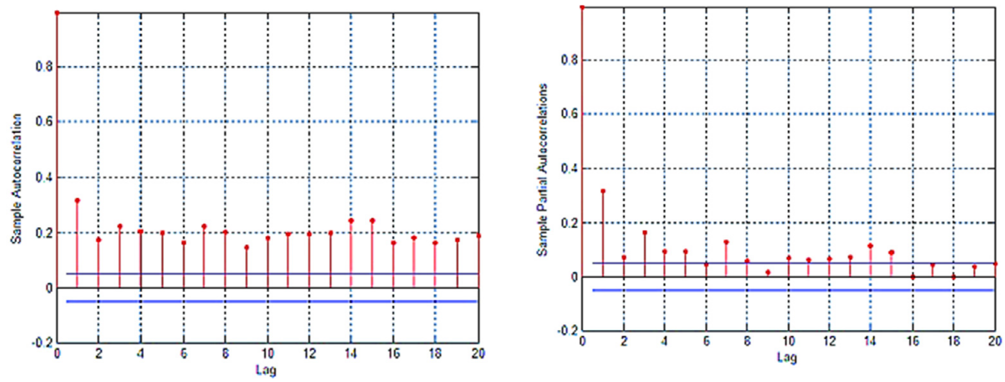
(a) Residual autocorrelations (b) Residual PAC

Fig. B.12. LPG consumption residuals for Sequential-GMC(1,n)-GA model.



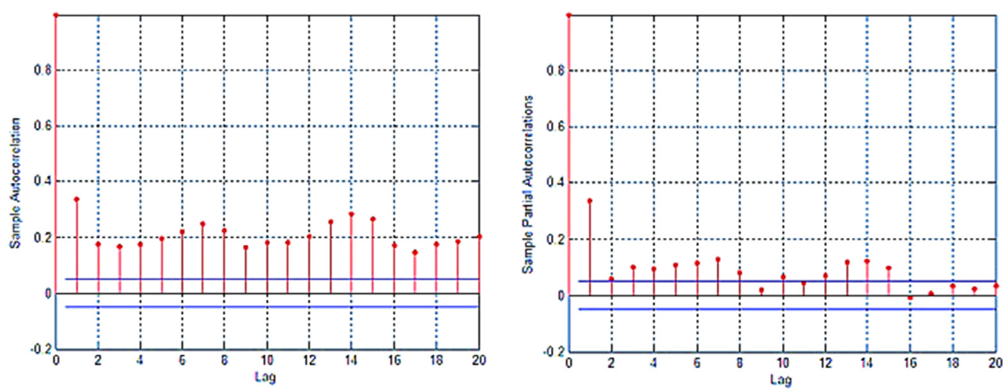
(a) Residual autocorrelations (b) Residual PAC

Fig. B.13. Kerosene consumption residuals for Sequential -GMC(1,n)-GA model.



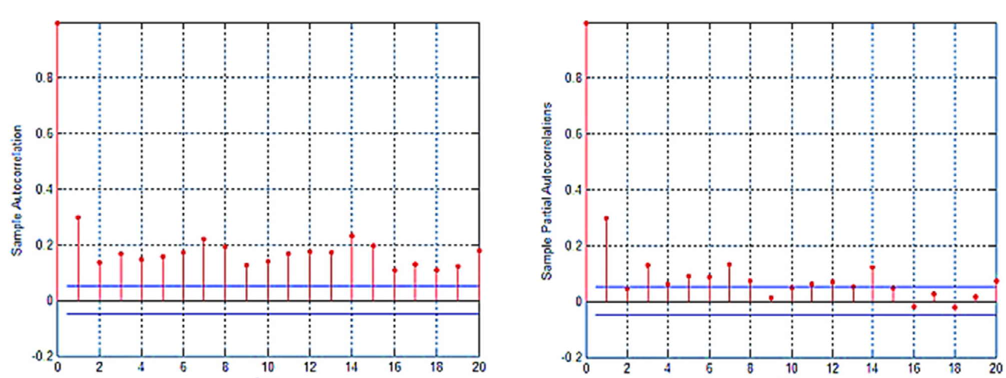
(a) Residual autocorrelations (b) Residual PAC

Fig. B.14. LPG consumption residuals for GMC(1,n) model.



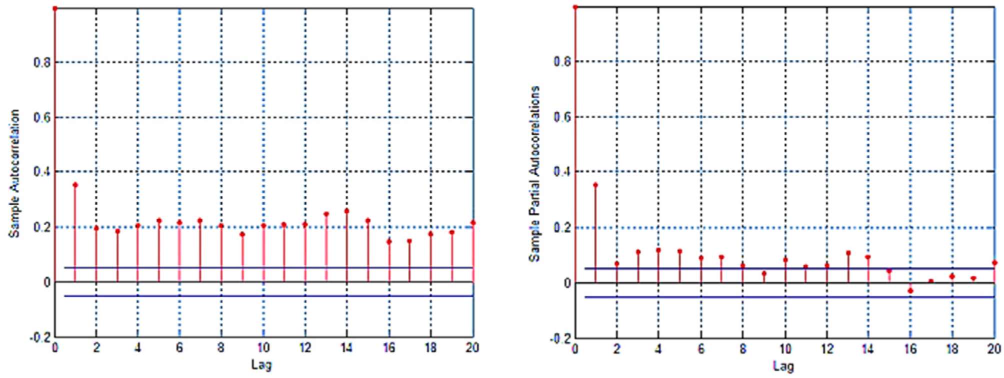
(a) Residual autocorrelations (b) Residual PAC

Fig. B.15. Kerosene consumption residuals for GMC(1,n) model.



(a) Residual autocorrelations (b) Residual PAC

Fig. B.16. LPG consumption residuals for OGMC(1,n) model.



(a) Residual autocorrelations (b) Residual PAC

Fig. B.17. Kerosene consumption residuals for OGMC(1,n) model.

Appendix C

Table C.7. Accuracy of hybrid GM-based models.

Author	Model	MAPE (%)	RMSE	AAE
Ofosu-Adarkwa et al. [39]	Verhulst-GM(1,n)	2.71	34	53
Lee and Tong [66]	GMGP(1,1)	4.20	–	–
Bahrami et al. [43]	Wavelet-GMPSO(1,5)	1.82	–	–
Zhao and Guo [112]	Rolling-ALO-GM(1,1)	4.04	–	–
Guefano et al. [69]	GM(1,1)-VAR(1)	1.63	15.42	1.5
This study	Sequential-GMC(1,n)-GA	1.44	0.8325	0.1204

Table C.8. Comparison between Sequential-GMC(1,n)-GA and other hybrid models, including artificial intelligence models.

Author	Model	MAPE (%)	RMSE	AAE
Yu et al. [113]	PSO-GA	0.54	–	–
Yu et al. [114]	EMD-RLSTM-ELM	3.56	–	–
	ARIMA-LR-ANN	2.45	–	–
Niu et al. [99]	VMD-EELM	2.96	471.781	–
Feng et al. [100]	VMD-SVM-PSO	–	3648.830	–
Zhang et al. [101]	VMD-SRSVRCBCS	0.9	–	62.3
This study	Sequential-GMC(1,n)-GA	1.44	0.8325	0.1204

Table C.9. Comparison between Sequential-GMC(1,n)-GA with similar studies in line with the performance criteria.

Author	Model and specifics	MAPE (%)	RMSE	AAE
Wu and Zhang [65]	GMC(1,n) with new information priorities	1.91	–	–
Tien [54]	Basic GMC(1,n) with Unit impulse response function	4.62	27.88	–
Shen et al. [55]	Optimized discrete GMC(1,n)	5.03	–	–
Ding and Li [59]	GMC(1,n) based on Simpson's rule	2.48	19.17	–
This study	Sequential-GMC(1,n)-GA	1.44	0.8325	0.1204

Table C.10. Comparison between GA-AC with similar optimization algorithms using the same dataset.

Reference paper	Optimization algorithm	Convergence accuracy/MAPE	Speed/s	Stability/Standard deviation
Wan et al. [102]	DE	3.61*	7*	0.089*
Gu et al. [103]	FOA-F-SVM	0.98***	1.2***	0.00014***
This study	GA	4.9	97	253.8
	GA-AC	1.44*	5**	0.02**

***, **, and * represents the performance rank of each algorithm. *** (1st rank), ** (2nd rank) and * (3rd rank).

Table C.11. Prediction values of Cameroon's household LPG and kerosene demand (2018-2025).

Year	LPG (MT)			Kerosene (m^3)			Total PP (ktoe)		
	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA
2018	108247.4	106190.6	101229.3	121444.1	119136.5	113570.1	108.368	106.309	101.342
2019	109408.3	109409.9	108340.2	114426.3	114427.9	113309.3	109.523	109.524	108.453
2020	117034.6	117035.9	115950.3	114417.3	114418.6	113357.4	117.149	117.150	116.063
2021	129434.3	126498.1	124095.7	116942.1	114289.1	112118.4	129.551	126.612	124.207
2022	133182.8	134123.5	132812.6	112509.3	113303.9	112196.2	133.295	134.237	132.924
2023	141451.9	141453.4	140141.4	112435.1	112436.2	111393.3	141.564	141.566	140.252
2024	148241.1	144878.1	142126.1	114433.5	111837.5	109713.5	148.356	144.990	142.236
2025	160626.5	151695.3	150212.5	113519.2	107207.3	106159.2	160.740	151.802	150.318

Table C.12. Equivalent electricity consumption forecasting in Cameroon's households.

Year	Electricity (GWh)		
	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA
2018	582.72	571.65	544.94
2019	672.26	672.27	665.70
2020	780.31	780.32	773.08
2021	899.61	879.20	862.50
2022	993.49	1000.50	990.72
2023	1115.59	1115.60	1105.25
2024	1319.95	1290.01	1265.50
2025	1525.02	1440.22	1426.14

Table C.13. Amount of GHGs (ktCO2-eq) that could be reduced in Cameroon's household sector (2018-2025).

Year	From LPG consumption			From kerosene consumption			Total household emissions		
	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA	GMC(1,n)	OGMC(1,n)	S-GMC(1,n)-GA
2018	97.08	107.19	90.79	550.13	514.01	514.46	647.21	621.2	605.04
2019	107.48	112.80	106.43	524.75	538.94	519.63	632.23	651.74	626.06
2020	131.05	126.11	129.84	524.21	524.43	519.36	655.26	650.54	649.20
2021	148.98	152.90	142.83	550.44	564.58	527.74	699.42	717.48	670.57
2022	160.79	179.64	160.34	538.28	601.41	536.79	699.07	781.05	697.13
2023	173.73	183.31	172.12	550.15	555.51	545.06	723.88	738.82	717.18
2024	194.92	178.75	186.88	554.76	561.81	531.88	749.68	740.56	718.76
2025	219.74	218.94	205.49	565.01	614.38	528.37	784.75	833.32	733.86

References

[1] J.G. Tamba, An analysis of Cameroon's energetic system, 2001–2010, Energy Sources, Part B, Econ. Plan. Policy 12 (2017) 216–222.

[2] SIE-Cameroon, Energetic situation of Cameroon, Report 2010, vol. 2, 2nd ed., Ministry of Energy and Water Resources, Yaoundé, 2011.

[3] SIE-Cameroon, Energetic situation of Cameroon, Report 2014, vol. 3, 1st ed., Ministry of Energy and Water Resources, Yaoundé, 2015.

[4] J.G. Tamba, LPG consumption and economic growth, 1975-2016: evidence from Cameroon, Int. J. Energy Sect. Manag. (2020).

[5] GESP, Government Reference Framework Action for 2010-2020, 2009th ed., Ministry of Economy, Planning and Regional Development, Washington, DC, 2009.

[6] MINEPAT, Stratégie Nationale de Développement 2020-2030: Pour la transformation structurelle et le développement inclusif, 2020th ed., Cameroun Emergent, Yaoundé, 2020.

[7] B. Soldo, Forecasting natural gas consumption, Appl. Energy 92 (2012) 26–37.

[8] J.G. Tamba, S.N. Essiane, E.F. Sapnken, F.D. Koffi, J.L. Nsouandélé, B. Soldo, et al., Forecasting natural gas: a literature survey, Int J Energy Econ Policy 8 (2018) 216–249.

[9] L. Suganthi, A.A. Samuel, Energy models for demand forecasting—a review, Renew. Sustain. Energy Rev. 16 (2012) 1223–1240.

[10] A. Özmen, Yı lmaz Y, G-W. Weber, Natural gas consumption forecast with MARS and CMARS models for residential users, Energy Econ. 70 (2018) 357–381.

[11] F.E. Sapnken, J.G. Tamba, S.E. Ndjakomo, F.D. Koffi, Oil products consumption and economic growth in Cameroon households: an assessment using ARDL cointegration and Granger causality analysis, Int. J. Energy Econ. Policy 10 (2020) 510–523.

[12] K. Peterman, M. Cordes, Paris: The Paris Agreement and the Path Beyond, Context. Clim. Change Link. Sci. Cult., ACS Publications, 2021, pp. 167–182.

[13] A. Rahil, R. Gammon, N. Brown, J. Udie, M.U. Mazhar, Potential economic benefits of carbon dioxide (CO2) reduction due to renewable energy and electrolytic hydrogen fuel deployment under current and long term forecasting of the Social Carbon Cost (SCC), Energy Rep. 5 (2019) 602–618.

[14] J.G. Tamba, D. Njomo, E.T. Mbog, Assessment of CO2 emissions in the petroleum refining in Cameroon, Univers. J. Environ. Res. Technol. 1 (2011).

[15] J.G. Tamba, D. Njomo, J.L. Nsouandele, B. Bonoma, S.B. Dongue, Assessment of greenhouse gas emissions in Cameroon's road transport sector (2012), Univers. J. Environ. Res. Technol. 2.

[16] M.K. Hubbert, Energy from fossil fuels, Science 109 (1949) 103–109.

[17] A.D. Papalexopoulos, T.C. Hesterberg, A regression-based approach to short-term system load forecasting, IEEE Trans. Power Syst. 5 (1990) 1535–1547.

[18] R.P. Timmer, P.J. Lamb, Relations between temperature and residential natural gas consumption in the Central and Eastern United States, J. Appl. Meteorol. Climatol. 46 (2007) 1993–2013.

[19] M. Ali, R. Prasad, Y. Xiang, R.C. Deo, Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms, Renew. Sustain. Energy Rev. 132 (2020) 110003.

[20] H.M. Al-Hamadi, S.A. Soliman, Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model, Electr. Power Syst. Res. 68 (2004) 47–59.

[21] M. Thaler, I. Grabec, A. Poredoš, Prediction of energy consumption and risk of excess demand in a distribution system, Phys. Stat. Mech. Appl. 355 (2005) 46–53.

[22] J-F. Chen, W-M. Wang, C-M. Huang, Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting, Electr. Power Syst. Res. 34 (1995) 187–196.
[23] K. Gairaa, A. Khellaf, Y. Messlem, F. Chellali, Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: a combined approach, Renew. Sustain. Energy Rev. 57 (2016) 238–249.
[24] J.L. Piggott, Use of Box–Jenkins modelling for the forecasting of daily and weekly gas demand, IEE Colloq. Dig. 4 (1983).
[25] S. Fathi, R. Srinivasan, A. Fenner, S. Fathi, Machine learning applications in urban building energy performance forecasting: a systematic review, Renew. Sustain. Energy Rev. 133 (2020) 110287.
[26] J. Vondráček, E. Pelikán, O. Konár, J. Čermáková, K. Eben, M. Malý, et al., A statistical model for the estimation of natural gas consumption, Appl. Energy 85 (2008) 362–370.
[27] A. Gascón, E.F. Sánchez-Úbeda, Automatic specification of piecewise linear additive models: application to forecasting natural gas demand, Stat. Comput. 28 (2018) 201–217.
[28] C. Rui, W. Jian, W. Li, Y. Ningjie, Z. Pengyan, The forecasting of China natural gas consumption based on genetic algorithm, in: 2009 Fifth Int. Jt. Conf. INC IMS IDC, IEEE, 2009, pp. 1436–1439.
[29] S. Aslam, H. Herodotou, S.M. Mohsin, N. Javaid, N. Ashraf, S. Aslam, A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids, Renew. Sustain. Energy Rev. 144 (2021) 110992.
[30] Z. Tonković, M. Zekić-Sušac, M. Somolanji, Predicting natural gas consumption by neural networks, Teh. Vjesn. 16 (2009) 51–61.
[31] M. Brabec, O. Konár, E. Pelikán, M. Malý, A nonlinear mixed effects model for the prediction of natural gas consumption by individual customers, Int. J. Forecast. 24 (2008) 659–678.
[32] A. Azadeh, R. Arab, S. Behfard, An adaptive intelligent algorithm for forecasting long term gasoline demand estimation: the cases of USA, Canada, Japan, Kuwait and Iran, Expert Syst. Appl. 37 (2010) 7427–7437.
[33] J-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. Syst. Man Cybern. 23 (1993) 665–685.
[34] S. Kyung, S. Young, H. Dug, J. Gilsoo, Short-term load forecasting for the holidays using fuzzy linear regression method, IEEE Trans. Power Syst. 20 (2005) 96–101.
[35] D.E. Goldberg, Genetic Algorithms in Machine Learning, Search and Optimization, 1989.
[36] G-D. Li, C-H. Wang, S. Masuda, M. Nagai, A research on short term load forecasting problem applying improved grey dynamic model, Int. J. Electr. Power Energy Syst. 33 (2011) 809–816.
[37] C. Hamzacebi, H.A. Es, Forecasting the annual electricity consumption of Turkey using an optimized grey model, Energy 70 (2014) 165–171.
[38] S. Ding, K.W. Hipel, Y. Dang, Forecasting China's electricity consumption using a new grey prediction model, Energy 149 (2018) 314–328.
[39] J. Ofosu-Adarkwa, N. Xie, S.A. Javed, Forecasting CO2 emissions of China's cement industry using a hybrid Verhulst-GM(1,N) model and emissions' technical conversion, Renew. Sustain. Energy Rev. 130 (2020) 109945.
[40] S. Liu, Y. Lin, Grey Information: Theory and Practical Applications, Springer Science & Business Media, 2006.
[41] C. Hatice, A. Alper, Determination of monthly wind speed of Kayseri region with gray estimation method, in: 2019 IEEE Jordan Int. Jt. Conf. Electr. Eng. Inf. Technol. 2019, 2019, pp. 415–420.
[42] Y. Xie, M. Li, Research on prediction model of natural gas consumption based on grey modeling optimized by genetic algorithm, in: 2009 IITA Int. Conf. Control Autom. Syst. Eng. Case 2009, IEEE, 2009, pp. 335–337.
[43] S. Bahrami, R-A. Hooshmand, M. Parastegari, Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm, Energy 72 (2014) 434–442.
[44] N. Aras, Forecasting residential consumption of natural gas using genetic algorithms, Energy Explor. Exploit. 26 (2008) 241–266.
[45] J. Holland, Adaptation in Natural and Artificial Systems, second edition, University of Michigan Press, 1975.
[46] D.A. Coley, Introduction to Genetic Algorithms for Scientists and Engineers, World Scientific Publishing Co, Singapore 912805, 1999.
[47] M. Mitchell, An Introduction to Genetic Algorithms, 5th ed., A Bradford Book the MIT Press, Massachusetts, 1999.
[48] P. Li, Grey forecasting of earthquake (M greater than or equal to 6.0) frequency in China mainland, J. Grey Syst. 2 (1990) 133–151.
[49] Y-F. Wang, Predicting stock price using fuzzy grey prediction system, Expert Syst. Appl. 22 (2002) 33–38.
[50] G. Sun, Prediction of vegetable yields by grey model GM (1, 1), J. Grey Syst. 2 (1991) 187–197.
[51] H.-W.V. Tang, M-S. Yin, Forecasting performance of grey prediction for education expenditure and school enrollment, Econ. Educ. Rev. 31 (2012) 452–462.
[52] M. Mao, E.C. Chirwa, Application of grey model GM (1, 1) to vehicle fatality risk estimation, Technol. Forecast. Soc. Change 73 (2006) 588–605.
[53] C-I. Hsu, Y-H. Wen, Improved grey prediction models for the trans-pacific air passenger market, Transp. Plann. Technol. 22 (1998) 87–107.
[54] T-L. Tien, A research on the grey prediction model GM (1, n), Appl. Math. Comput. 218 (2012) 4903–4916.
[55] Q-Q. Shen, Y. Cao, L-Q. Yao, Z-K. Zhu, An optimized discrete grey multi-variable convolution model and its applications, Comput. Appl. Math. 40 (2021) 58.
[56] Z. He, Y. Shen, J. Li, Y. Wang, Regularized multivariable grey model for stable grey coefficients estimation, Expert Syst. Appl. 42 (2015) 1806–1815.
[57] H. Zhu, Multi-parameter grey prediction model based on the derivation method, Appl. Math. Model. 97 (2021) 588–601.
[58] L. Ye, N. Xie, A. Hu, A novel time-delay multivariate grey model for impact analysis of CO2 emissions from China's transportation sectors, Appl. Math. Model. 91 (2021) 493–507.
[59] S. Ding, R. Li, A new multivariable grey convolution model based on Simpson's rule and its applications, Complexity (2020) (2020) 1–14.
[60] S. Ding, Z. Tao, J. Hu, Forecasting the economic indices of the high-tech industries in China using the grey multivariable convolution model, Appl. Soft Comput. 126 (2022) 109301.
[61] B. Zeng, H. Duan, Y. Zhou, A new multivariable grey prediction model with structure compatibility, Appl. Math. Model. 75 (2019) 385–397.
[62] S. Ding, A novel discrete grey multivariable model and its application in forecasting the output value of China's high-tech industries, Comput. Ind. Eng. 127 (2019) 749–760.
[63] Z-X. Wang, Multivariable time-delayed GM (1, N) model and its application, Control Decis. 30 (2015) 2298–2304.
[64] S. Ding, Y.G. Dang, N. Xu, L. Wei, J. Ye, Multi-variable time-delayed discrete grey model, Control Decis. 32 (2017) 1997–2004.
[65] L. Wu, Z. Zhang, Grey multivariable convolution model with new information priority accumulation, Appl. Math. Model. 62 (2018) 595–604.
[66] Y-S. Lee, L-I. Tong, Forecasting energy consumption using a grey model improved by incorporating genetic programming, Energy Convers. Manag. 52 (2011) 147–152.
[67] D. Akay, Atak M. Grey, Prediction with rolling mechanism for electricity demand forecasting of Turkey, Energy 32 (2007) 1670–1675.
[68] B. Zeng, C. Li, Forecasting the natural gas demand in China using a self-adapting intelligent grey model, Energy 112 (2016) 810–825.
[69] S. Guefano, J.G. Tamba, T.E.W. Azong, L. Monkam, Forecast of electricity consumption in the Cameroonian residential sector by Grey and vector autoregressive models, Energy 214 (2021) 118791.
[70] L.L. Marius, N.A. Joel, Energy sector of Cameroon, Afr. Rev. 11 (2019) 34–45.
[71] J.G. Tamba, D. Njomo, T. Limanond, B. Ntsafack, Causality analysis of diesel consumption and economic growth in Cameroon, Energy Policy 45 (2012) 567–575.
[72] E.F. Sapnken, J.G. Tamba, F.L. Djanna Koffi, The variability of gasoline demand in Cameroon. Energy sources part B, Econ. Plan. Policy 12 (2017) 772–778.
[73] J.G. Tamba, J.L. Nsouandélé, E.F. Sapnken, F.L. Djanna Koffi, The variability of diesel demand in Cameroon. Energy sources part B, Econ. Plan. Policy 12 (2017) 868–875.
[74] G.P. Dkamela, The Context of REDD+ in Cameroon: Drivers, Agents and Institutions, vol. 57, CIFOR, 2011.
[75] O.A. Somorin, I.J. Visseren-Hamakers, B. Arts, D.J. Sonwa, A-M. Tiani, REDD+ policy strategy in Cameroon: actors, institutions and governance, Environ. Sci. Policy 35 (2014) 87–97.
[76] World Bank, WDI - environment. Environment 2019, https://datatopics.worldbank.org/world-development-indicators/themes/environment.html. (Accessed 17 May 2021).
[77] J.-L. Deng, Control problems of grey systems, Syst. Control Lett. 1 (1982) 288–294.
[78] C. Hamzacebi, H.A. Es, Forecasting the annual electricity consumption of Turkey using an optimized grey model, Energy 70 (2014) 165–171.
[79] F-M. Tseng, H-C. Yu, G-H. Tzeng, Applied hybrid grey model to forecast seasonal time series, Technol. Forecast. Soc. Change 67 (2001) 291–302.
[80] C. Bessiere, J-C. Régin, Refining the basic constraint propagation algorithm, in: IJCAI, vol. 1, 2001, pp. 309–315.
[81] M. Zouita, S. Bouamama, K. Barkaoui, Improving genetic algorithm using arc consistency technic, Proc. Comput. Sci. 159 (2019) 1387–1396.
[82] A.K. Mackworth, Consistency in networks of relations, Artif. Intell. 8 (1977) 99–118.
[83] P. Gray, W. Hart, L. Painton, C. Phillips, M. Trahan, J. Wagner, A Survey of Global Optimization Methods, 1997.
[84] G. Ozdemir, E. Aydemir, M.O. Olgun, Z. Mulbay, Forecasting of Turkey natural gas demand using a hybrid algorithm, Energy Sources, Part B, Econ. Plan. Policy 11 (2016) 295–302.
[85] M. Gen, R. Cheng, Genetic Algorithms and Engineering Optimization, vol. 7, John Wiley & Sons, 1999.
[86] W.O. Siyanbola, M.S. Adaramola, O.O. Omotoyin, O.A. Fadare, Determination of energy efficiency of non-bio-fuel household cooking stoves, Niger. J. Phys. 16 (2004) 171–179.
[87] H.S. Eggleston, L. Buendia, K. Miwa, T. Ngara, K. Tanabe, 2006 IPCC guidelines for national greenhouse gas inventories, 2006,.
[88] IPCC, Task force on National GHG inventories, Good Pract Guid Uncertain Manag Natl Greenh Gas Invent n.d., https://www.ipcc-nggip.iges.or.jp/public/gp/english/. (Accessed 30 August 2021).
[89] P-C. Chang, C-H. Liu, A TSK type fuzzy rule based system for stock price prediction, Expert Syst. Appl. 34 (2008) 135–144.

[90] E. Hadavandi, H. Shavandi, A. Ghanbari, Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting, Knowl.-Based Syst. 23 (2010) 800–808.
[91] A. Esfahanipour, W. Aghamiri, Adapted neuro-fuzzy inference system on indirect approach TSK fuzzy rule base for stock market analysis, Expert Syst. Appl. 37 (2010) 4742–4748.
[92] T.J. Burkholder, R.L. Lieber, Stepwise regression is an alternative to splines for fitting noisy data, J. Biomech. 29 (1996) 235–238.
[93] H. Verbois, P. Blanc, R. Huva, Y-M. Saint-Drenan, A. Rusydi, A. Thiery, Beyond quadratic error: case-study of a multiple criteria approach to the performance assessment of numerical forecasts of solar irradiance in the tropics, Renew. Sustain. Energy Rev. 117 (2020) 109471.
[94] J.M. González-Sopeña, V. Pakrashi, B. Ghosh, An overview of performance evaluation metrics for short-term statistical wind power forecasting, Renew. Sustain. Energy Rev. 138 (2021) 110515.
[95] S. Makridakis, S.C. W, V.E. McGee, Forecast Methods and Applications, second ed., John Wiley Sons, 1983.
[96] J.H. Holland, Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, MIT Press, 1992.
[97] P.J. Mitchell, F. Waldner, H. Horan, J.N. Brown, Z. Hochman, Data fusion using climatology and seasonal climate forecasts improves estimates of Australian national wheat yields, Agric. For. Meteorol. 320 (2022) 108932.
[98] I. Kucukkoc, A.D. Karaoglan, R. Yaman, Using response surface design to determine the optimal parameters of genetic algorithm and a case study, Int. J. Prod. Res. 51 (2013) 5039–5054.
[99] W. Niu, Z. Feng, Y. Chen, H. Zhang, C. Cheng, Annual streamflow time series prediction using extreme learning machine based on gravitational search algorithm and variational mode decomposition, J. Hydrol. Eng. 25 (2020) 04020008.
[100] Z. Feng, W. Niu, Z. Tang, Z. Jiang, Y. Xu, Y. Liu, et al., Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization, J. Hydrol. 583 (2020) 124627.
[101] Z. Zhang, W-C. Hong, J. Li, Electric load forecasting by hybrid self-recurrent support vector regression model with variational mode decomposition and improved Cuckoo search algorithm, IEEE Access 8 (2020) 14642–14658.
[102] X. Wan, X. Zuo, X. Zhao, A differential evolution algorithm combined with linear programming for solving a closed loop facility layout problem, Appl. Soft Comput. 121 (2022) 108725.
[103] Q. Gu, Y. Chang, X. Li, Z. Chang, Z. Feng, A novel F-SVM based on FOA for improving SVM performance, Expert Syst. Appl. 165 (2021) 113713.
[104] A. Kuriqi, A.N. Pinheiro, A. Sordo-Ward, M.D. Bejarano, L. Garrote, Ecological impacts of run-of-river hydropower plants—current status and future prospects on the brink of energy transition, Renew. Sustain. Energy Rev. 142 (2021) 110833.
[105] C. Llamosas, B.K. Sovacool, The future of hydropower? A systematic review of the drivers, benefits and governance dynamics of transboundary dams, Renew. Sustain. Energy Rev. 137 (2021) 110495.
[106] A.M. Mayeda, A.D. Boyd, Factors influencing public perceptions of hydropower projects: a systematic literature review, Renew. Sustain. Energy Rev. 121 (2020) 109713.
[107] F.H. Abanda, Renewable energy sources in Cameroon: potentials, benefits and enabling environment, Renew. Sustain. Energy Rev. 16 (2012) 4557–4562.
[108] E. Muh, S. Amara, F. Tabet, Sustainable energy policies in Cameroon: a holistic overview, Renew. Sustain. Energy Rev. 82 (2018) 3420–3429.
[109] F. Polzin, M. Sanders, How to finance the transition to low-carbon energy in Europe?, Energy Policy 147 (2020) 111863.
[110] S.A. Qadir, H. Al-Motairi, F. Tahir, L. Al-Fagih, Incentives and strategies for financing the renewable energy transition: a review, Energy Rep. 7 (2021) 3590–3606.
[111] A. Ghanbari, S.M.R. Kazemi, F. Mehmanpazir, M.M. Nakhostin, A cooperative ant colony optimization-genetic algorithm approach for construction of energy demand forecasting knowledge-based expert systems, Knowl.-Based Syst. 39 (2013) 194–206.
[112] H. Zhao, S. Guo, An optimized grey model for annual power load forecasting, Energy 107 (2016) 272–286.
[113] S. Yu, K. Zhu, X. Zhang, Energy demand projection of China using a path-coefficient analysis and PSO–GA approach, Energy Convers. Manag. 53 (2012) 142–153.
[114] L. Yu, Y. Zhao, L. Tang, Z. Yang, Online big data-driven oil consumption forecasting with Google trends, Int. J. Forecast. 35 (2019) 213–223.