Applied Energy 228 (2018) 808–820

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

An optimized gene expression programming model for forecasting the national CO_2 emissions in 2030 using the metaheuristic algorithms

Taehoon Hong[a], Kwangbok Jeong[a,*], Choongwan Koo[b]

[a] Department of Architecture and Architectural Engineering, Yonsei University, Seoul 03722, Republic of Korea
[b] Department of Architectural Engineering, Kyonggi University, Suwon 16227, Republic of Korea

HIGHLIGHTS

- A model was developed to forecast the national CO_2 emissions in 2030.
- Model was developed using gene expression programming and harmony search algorithm.
- The mean absolute percentage error of developed model was estimated to be 2.06%.
- National CO_2 emissions in 2030 is increased by 10.0–21.91% than in 2016.
- Developed model can help policy makers forecast the national CO_2 emissions.

ARTICLE INFO

Keywords:
National CO_2 emissions
Gene expression programming
Harmony search optimization algorithm
Particle swarm optimization algorithm

ABSTRACT

To cope with the approaching POST-2020 scenario, the national CO_2 emission in the building sector, which accounts for 25.5% of the total CO_2 emissions, should be managed effectively and efficiently. To do this, it is essential to forecast the national CO_2 emissions in the building sector by region. As the South Korean government does not currently do this by region, regional characteristics are rarely taken into consideration when managing the national CO_2 emissions in the building sector. Towards this end, this study developed an optimized gene expression programming model for forecasting the national CO_2 emissions in 2030 using the metaheuristic algorithms. Compared to the forecasting performance of the gene expression programming model, the forecasting performance of the optimized gene expression programming – harmony search optimization model has improved by 7.11, 2.05, and 2.06% in terms of the mean absolute error, root mean square error, and mean absolute percentage error, respectively. Various national CO_2 emissions scenarios in the building sector were established in order to better analyze the variation range of the national CO_2 emissions in the building sector. Compared to the national CO_2 emissions in 2016 (i.e., scenario 1: 41,337 $ktCO_2$; scenario 2: 45,373 $ktCO_2$; scenario 3: 46,024 $ktCO_2$) in multi-family housing complexes, the national CO_2 emissions in 2030 (i.e., scenario 1: 37,579 $ktCO_2$; scenario 2: 37,736 $ktCO_2$; scenario 3: 37,754 $ktCO_2$) in multi-family housing complexes are forecasted to increase by 10.00–21.91%. The developed optimized gene expression programming – harmony search optimization model will potentially be able to assist policymakers in central and local governments forecast the national CO_2 emissions in 2030. Through this, national CO_2 emission management that more closely reflects the characteristics at the regional or national level can be supported.

1. Introduction

To reduce greenhouse gas (GHG) emissions, a key factor contributing to global warming, the Paris Agreement was adopted in December 2015, and the POST-2020 framework, participated in by all the developed and developing countries in the world, was launched. Each country established its nationally determined contribution to the global greenhouse gas (GHG) emission reduction efforts: the national CO_2 emission reduction target (CERT) in 2030 [1]. The South Korean government established a national CERT so as to reduce South Korea's GHG emissions by 37% (specifically, the building sector was targeted at 18.1%) below its business-as-usual level by 2030 [2,3]. To address this challenge, the South Korean government established various policies, such as the "GHG target management system" and the "GHG emissions trading scheme" [4,5]. As these policies, however, which are being implemented by the central government, were established by the

* Corresponding author at: Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea.
E-mail addresses: hong7@yonsei.ac.kr (T. Hong), kbjeong7@yonsei.ac.kr (K. Jeong), cwkoo@kyonggi.ac.kr (C. Koo).

https://doi.org/10.1016/j.apenergy.2018.06.106
Received 5 March 2018; Received in revised form 19 June 2018; Accepted 21 June 2018
0306-2619/

Nomenclature

Abbreviations

ANN	Artificial neural network
CAGR	Compounded annual growth rate
CERT	CO_2 emissions reduction target
GA	Genetic algorithm
GDP	Gross domestic product
GEP	Gene expression programming
GEP-HSO	Gene expression programming – harmony search optimization
GEP-PSO	Gene expression programming – particle swarm optimization
GHG	Greenhouse gas
GP	Genetic programming
GRDP	Gross regional domestic production
HM	Harmony memory
HMCR	Harmony memory considering rate
HMS	Harmony memory size
HSO	Harmony search optimization
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MFHC	Multi-family housing complex
PAR	Pitch adjusting rate
PSO	Particle swarm optimization
RMSE	Root mean square error
RNC	Random numerical constant
SVM	Support vector machine

industry and energy sectors, they suffer from limitations in being able to perform GHG management successfully in the building sector. Moreover, with the gradual increase in the importance at the local government level for GHG management in the building sector, it is necessary to forecast and manage building data, socio-economic data, and GHG emission data, which largely affect local-level GHG emissions [6]. For example, an increase in the gross domestic product (GDP) results in increased energy consumption and GHG emissions based on the increase in related economic activities [7,8]. In order to better manage GHG emissions in the building sector, first of all, it is necessary to forecast GHG emissions by region. In other words, GHG emissions by region can be integrated into the national-level GHG emission data, and ultimately, can be made part of the national GHG emission management. The current forecasting methods, however, face several limitations in forecasting the CO_2 emissions by region. First, such methods estimate the CO_2 emissions by country and cannot easily forecast detailed CO_2 emissions by reflecting the regional characteristics. In addition, while some local governments are working in some ways towards forecasting and managing their respective regions' CO_2 emissions, there is no standard for unified forecasting and management. Second, the current forecasting methods consider CO_2 emissions from the gas consumption of buildings and categorize the CO_2 emissions according to electricity consumption as in the transformation sector. Thus, it is difficult to calculate the total CO_2 emissions of a building in the operation phase. Third, the previous methods were the bottom-up type and were limited in that they were cost- and time-consuming as they collected historical data by building. Compared to the current forecasting method, to better solve the aforementioned three limitations, the objective of this study were as follows: (i) forecasting of CO_2 emissions by reflecting the regional characteristics; (ii) forecasting of CO_2 emissions by taking into consideration both electricity and gas consumption in building; and (iii) forecasting of CO_2 emissions in a manner more efficient in terms of both cost and time than previous methods. To address this challenge, this study aimed to develop an optimized gene expression programming (GEP) model for forecasting the national CO_2 emissions in 2030 using the metaheuristic optimization algorithms.

To develop the model, the research scopes are as following. First, it is observed that the energy consumption in the residential sector accounts for about 58% of the total energy consumption. Moreover, South Korea has a high rate of multi-family housing complexes (MFHCs), and the electricity consumption of MFHCs accounts for 68% of the residential electricity consumption [9,10]. Thus, this study developed an optimized GEP model targeting the MFHCs. Second, South Korea can be largely divided into 16 administrative divisions, and these regions have different characteristics in terms of socio-economic and building factors. Accordingly, the characteristics of both energy consumption and CO_2 emissions differ according to regional characteristic. As a result,

Table 1
Previous studies on a forecasting model for energy consumption and CO_2 emissions.

Authors	Soft computing approach			Target variable	Independent variables	Historical data	Forecast period
	MLA[a]	EA[b]	MA[c]				
Kandal et al. [15]	ANN[d]	–	–	Net energy consumption	GDP[j], population, import, export	1980–2007	2008–2014
Sun and Liu [16]	SVM[e]	–	–	Energy consumption and CO_2 emissions	Per capita GDP, consumer price index, etc.	1978–2012	2008–2012
Karabulut et al. [17]	–	GP[f]	–	Electricity consumption	Electricity demand, climate	1994–2005	2006–2010
Mousavi et al. [18]	–	GEP[g]	–	Electricity demand	GDP, population, stock index	1986–2003	2004–2009
Ardakani and Ardehali [21]	–	ANN[h]	PSO[h]	Electricity demand	GDP, population, import, export	1967–2009	2010–2030
Mostafavi et al. [22]	–	GP	SA[i]	Electricity demand	GDP, population, stock index	1986–2003	2004–2009

Note.
[a] MLA is the machine learning algorithm.
[b] EA is the evolutionary algorithm.
[c] MA is the metaheuristic algorithm.
[d] ANN is the artificial neural network.
[e] SVM is the support vector machine.
[f] GP is the genetic programming.
[g] GEP is the gene expressions programming.
[h] PSO is the particle swarm optimization.
[i] SA is the simulated annealing.
[j] GDP is the gross domestic production.