

# Study of Climate Change Detection in North-East Africa Using Machine Learning and Satellite Data

Sara K. Ibrahim<sup>ID</sup>, Ibrahim E. Ziedan, and Ayman Ahmed<sup>ID</sup>

**Abstract**—The study of climate change has become an important topic because of its negative impact on human life. The North-East African part lacks the studies for climate change detection, despite it being one of the most affected parts worldwide. The relationship between the emission of greenhouse gases (GHGs) and climate change is an important factor to understand. To investigate this linkage, we used machine-learning (ML) models based on essential climate variables (ECVs) to investigate the relationship between the GHGs and the rhythm of climate variable change. The article investigates how ML techniques can be applied to climatic data to build an ML model that is able to predict the state of climate variables for the short and long term. By selecting a candidate model, it will help in climate adaptation and mitigation, also determine at what level GHGs should be kept and their corresponding concentrations in order to avoid climate events and crises. The used models are long short-term memory, autoencoders, and convolutional neural network (CNN). Alternatively, the dataset has been selected from U.K. National Centre for Earth Observation and Copernicus Climate Change Services. We compared the performance of these techniques and the best candidate was the Head—CNN; based on performance metrics such as root-mean-squared-error: 5.378, 2.395, and 15.923, mean-absolute-error: 4.157, 1.928, and 11.672, Pearson: 0.368, 0.649, and 0.291, and  $R^2$  coefficient: 0.607, 0.806, and 0.539 for the ECVs temperature, CO<sub>2</sub>, and CH<sub>4</sub>, respectively. We were able to link the GHG emission to ECVs with high accuracy based on the reading of this geographic area.

**Index Terms**—Climate change, greenhouse gases (GHGs), machine learning (ML), neural network, space systems.

## I. INTRODUCTION

AFRICA has been affected by extreme climate change recently; East Africa was exposed to unprecedented rainfall over the period of October 2019 to January 2020. Rainfall impacts floods and landslides across the region causing a natural disaster in many countries including Ethiopia, Kenya, Somalia, Uganda, Tanzania, and Djibouti, adversely affecting over 2.8 million people [1]. However, these regions face a poor rainy season from March to May, resulting in food shortages and famines [2]. Moreover, climate change makes these regions more

Manuscript received April 17, 2021; revised July 14, 2021 and September 10, 2021; accepted October 12, 2021. Date of publication October 19, 2021; date of current version November 10, 2021. This work was supported by the Space Plasma Nanosatellite Experiment Mission Alliance, funded by Academy of Scientific Research and Technology and the Egyptian Space Agency, Egypt under Grant SPNEX-2021. (*Corresponding author: Sara K. Ibrahim*)

Sara K. Ibrahim and Ibrahim E. Ziedan are with the Department of Computer & Systems, Faculty of Engineering, Zagazig University, Zagazig 44519, Egypt (e-mail: sara.khalil@zu.edu.eg; iezedan@gmail.com).

Ayman Ahmed is with the Egyptian Space Agency, Cairo 1564, Egypt (e-mail: ayman.ahmed@egsa.gov.eg).

Digital Object Identifier 10.1109/JSTARS.2021.3120987

precarious than ever, leading to additional attention to the study of climate change, especially in the North-Eastern part of Africa.

Atmospheric greenhouse gases (GHGs) affect climate change the most, leading to a need of monitoring the anthropogenic GHG concentrations from space using artificial satellites. The sources of more than 40% of global annual anthropogenic emissions of carbon dioxide (CO<sub>2</sub>) are power plants operated by coal burning. Additionally, methane (CH<sub>4</sub>) leakage can be attributed by anthropogenic sources, such as coal mines, oil-gas systems, livestock, wastewater management, rice cultivation, and landfills, in addition to natural sources, such as termites, inland waters, and wetlands.

Ever since the start of the industrial revolution, consumption of fossil fuels has modified the Earth's climate in unforeseen ways. A 1.5 °C increase in temperature can be observed because of global warming compared with the preindustrial levels and global GHG emissions as reported in the Intergovernmental Panel on Climate Change (IPCC) 2018 [3]. The change consequences to the climate are seemingly everywhere such as rising average temperatures and sea levels. Moreover, these climate changes affect the quality and availability of water resources, food production, and the ecosystem. The industrial sector contributes 65% of the global emissions of the world's GHGs emissions. Oil and gas extraction, heavy industry, power sector, and mining produce 55% of the world's CO<sub>2</sub> emissions. Additionally, agriculture gives a majority contribution to anthropogenic emissions of NO<sub>2</sub> and CH<sub>4</sub> by approximately 75% and 40% of global emissions, respectively [4]–[7].

The assessment of the satellite-derived column of XCO<sub>2</sub> measurements and the CO<sub>2</sub> emission estimates for each other has been found in different applications, for example, in determining the impact of regional fossil fuel emissions on global XCO<sub>2</sub> fields [8], in determining CO<sub>2</sub> emissions from individual power plants from space through plume model simulations [9], and for observing the emission of anthropogenic CO<sub>2</sub> by deriving CO<sub>2</sub> anomalies through deseasonalizing and detrending XCO<sub>2</sub> column measurements [10]. Moreover, the hyperspectral measurement of Earth-reflected radiation in the shortwave spectral range (300–2500 nm) improves the monitoring of CO<sub>2</sub>, as well as the classification of gases using artificial intelligence techniques [11]–[13].

Climate change is based on many variables that can be correlated in one way or another. The variables have many data acquired by sensor readings over a period of time. The linkage between those variables and each other as well as the effect of each phenomenon and climate change over a very large number

of readings cannot be perceived by the human mind, which is where the role of the machines comes in. A large amount of data, as well as the relation between variables, can be processed by a well-developed machine-learning (ML) approach [14].

In this article, we introduce ML models based on variables affecting climate change to relate GHGs to the rhythm of climate change. We compared different ML techniques to assess their performance. The dataset used in this article is acquired from different sensors, including U.K. National Centre for Earth Observation (NCEO) as part of the European Space Agency (ESA) GHG–climate change initiative (GHG–CCI) and coper-nicus climate change services (C3S) projects [15]–[17].

The rest of this article is organized as follows. First, we introduce the essential variables affecting climate change based on a literature review. Then, we describe the dataset and explain the preprocessing of data acquired from different sensors, as well as the time-series continuity of variables. The third section describes the proposed ML techniques, after which we introduce the evaluation methodology and justify its use in assessment. Finally, we discuss the assessment of the algorithms and conclude the article with a discussion on the results.

## II. FACTORS AFFECTING CLIMATE CHANGE AND GLOBAL WARMING

Climate models are usually used to forecast climate status in the future. It handles the interaction between the physical processes in the atmosphere, land, and water system, and the surrounding environment. Climate models generate projections of future temperature and precipitation based on GHGs concentration scenarios [18].

Anthropogenic (human-caused) climate change occurs due to GHGs trapping more heat in the environmental ecosystem. The distribution of global average temperatures is expected to shift to the warmer side; therefore, sometimes anthropogenic climate change is referred to as “global warming.”

### A. Climate Variables

The Global Climate Observing System has defined a set of parameters known as essential climate variables (ECV); these variables are related to land, ocean, and atmosphere. Studying these variables helps in understanding the past, current, and future climate variability and possible changes. The ECVs could be physical, chemical, or biological variables or a set of related parameters that potentially characterize the Earth’s climate. The ECV for the atmosphere are precipitation, temperature, water vapor, clouds, pressure, wind speed and direction, CO<sub>2</sub>, CH<sub>4</sub>, and other GHGs, aerosol, and ozone. Meanwhile, the land’s ECVs are river discharge, groundwater, lakes, soil moisture, snow and glaciers, albedo, fire, and anthropogenic GHG fluxes. Moreover, the ECVs for the ocean are ocean surface heat flux, sea ice, level, salinity, temperature, currents, color, biogeochemical, and ecosystem marine habitat and plankton.

In this article, we carried out the analysis for only “XCO<sub>2</sub>, XCH<sub>4</sub>, pressure, temperature, total column water vapor (TCWV), and total column rain water (TCRW).” The ECVs datasets contribute to understand and predict the variability of

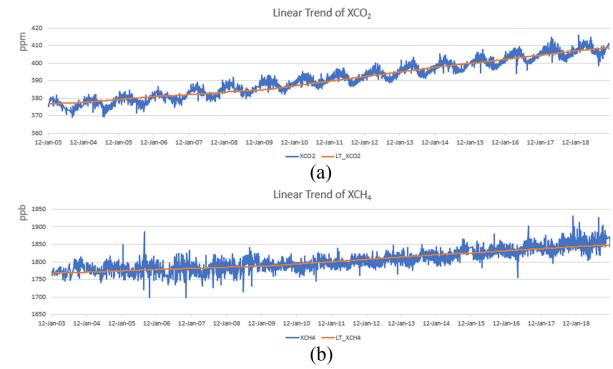


Fig. 1. Linear trend of column-average dry-air mixing ratios of CO<sub>2</sub> and CH<sub>4</sub> for North-East part of Africa for the period 2003–2018.

climate, tackling these changes by mitigation, and adaptation processes, estimate the impacts of climate change, and thus, enhance climatic services.

### B. Greenhouse Gases Effect

GHGs are important in understanding the climatic system and how climate might change [6], [7]. As the concentration of these gases gets higher in the atmosphere, the warming effect becomes greater. GHGs in the atmosphere absorb radiation emitted by the Earth’s surface and reradiate it in scattered directions. Since some of this reradiation is downward, the net effect is to inhibit heat loss from the surface to the upper reaches of the atmosphere.

Gases that have a greater effect on climate change and surface temperatures are CO<sub>2</sub> and CH<sub>4</sub>. CH<sub>4</sub> can last only for a decade so it is relatively shorter lived than CO<sub>2</sub> that persists for a century or more. However, during its lifespan, CH<sub>4</sub> is about 85 times more effective at trapping heat than CO<sub>2</sub> [6].

The impacts of CO<sub>2</sub> changes induce several important effects. The first effect is from additional water vapor that results from higher temperatures. Because water vapor is also a GHG, increasing CO<sub>2</sub> by a small percentage heats the atmosphere by trapping more heat. This permits more water vapor in the atmosphere, which also traps more heat (positive water vapor feedback) [19]. Chapter 5 of the fifth assessment report [20] discussed the impacts of anthropogenic GHGs emissions on climate change specifically CO<sub>2</sub> and CH<sub>4</sub> describing them as a warming impact. These gases are mostly derived from human daily life: industry, transport, buildings, energy, and waste. Fig. 1 shows a significant increase in CO<sub>2</sub> and CH<sub>4</sub> over the North-East part of Africa; however, this region is neither industrial nor energy—oil & gas—producer.

1) *Measuring of GHGs:* With the exponential growth in space-based remote sensing, algorithmic complexity, and computing power, end-to-end learning systems are becoming gradually accessible to academic researchers and the private sector. There are also feasible methods using remote sensing data that may be used to measure the emission levels of GHGs [21].

The first method is measuring emissions directly using sensors on airborne and spaceborne instruments. The other way is the indirect method, which use various identifiable parameters and characteristics to build an estimated model for GHG emissions.

Monitoring emissions by direct method demands the use of special instruments of airborne or satellites to accurately monitor GHGs with high precision from space. This is a challenging task because of their relatively small signal compared to other atmospheric constituents. Alternatively, the advances in both sensor technology and retrieval models are leading to more accurate detection and interpretation [10].

The indirect methodology should identify the main features associated with GHG emissions. For illustration, factory use rates are inherently correlated to the ratio of GHGs emissions. By using spaceborne tools combined with *in situ* factory-level production data gives the ability to model temporal variations in a factory use rate. Employing this projection of the use rate, an estimate of the emissions can then be obtained using a standardized model, such as those outlined in the IPCC reports [3], [5], [6].

2) *Satellite Sensors Used to Measure GHGs:* Usually, satellite sensors orbiting the earth at low sun-synchronous orbit are used to measure GHGs at a fixed observing time each day for the same geographical area, which is known as the temporal resolution. CO<sub>2</sub> and CH<sub>4</sub> are measured as column-averaged dry air mole fractions (XCO<sub>2</sub> and XCH<sub>4</sub>) in the along-track nadir [22]. Some instruments are also used to measure cross-track in the off-nadir to enhance the spatial coverage. Shortwave infrared (SWIR) imaging is highly recommended for measuring GHGs to differentiate the spectral signature of each element. The CH<sub>4</sub> spectral characteristic is located in SWIR wavelengths around 1650 and 2300 nm, and at 8000 nm in the thermal infrared (TIR) spectrum. While CO<sub>2</sub> wavelength regions can be derived around 760, 1610, and 2060 nm, measurements in the SWIR are more sensitive to the near-surface atmospheric layers where the emitting sources are located, whereas TIR measurements provide a more regular sampling of the entire atmospheric column. This makes TIR retrievals less sensitive to CO<sub>2</sub> and CH<sub>4</sub> emissions on local and regional scales [12], [23]. Table I lists GHG sensors onboard satellite missions. Using such sensors onboard earth-orbiting satellite provides continuous measurements with wide coverage and near real-time observation. An example of such sensors onboard satellites is SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) on ESA's ENVIRONMENTAL SATellite (ENVISAT), and Thermal and Near infrared Sensor for carbon Observation–Fourier Transform Spectrometer (TANSO–FTS) on Greenhouse Gases Observing Satellite (GOSAT) [24]. SCIAMACHY/ENVISAT was a spectrometer installed on ESA's ENVISAT satellite (2002–2012). SCIAMACHY [25], [26] covers the spectral region from the ultra-violet to the SWIR spectral region (240–2380 nm) at moderate spectral resolution (0.2–1.5 nm) and observes the Earth's atmosphere in various viewing geometries (nadir, limb, and solar or lunar occultation). SCIAMACHY permits the retrieval of XCO<sub>2</sub> and XCH<sub>4</sub> [27], [28] from the appropriate spectral regions in the SWIR (around 1.6 μm).

The GOSAT is dedicated to measuring the concentrations of CO<sub>2</sub> and CH<sub>4</sub> from space. TANSO–FTS/GOSAT is an FTS onboard the Japanese GOSAT satellite [23], [29]. The GOSAT satellite was launched and operated successfully in 2009. GOSAT measures at 1650 nm with a high spectral resolution for CH<sub>4</sub> retrieval on a continental or regional scale. It

TABLE I  
ILLUSTRATES SUMMARY OF GHG MISSIONS THAT HAVE BEEN LAUNCHED

Satellite Name	Sensor Name	Date	Spectral Characteristics of The Sensor
NOAA-10	HIRS-2/TOVS	Sep 1986 - Aug 2001	Filter radiometer, 20 channels
ADEOS I	IMG	Aug 1996– Jun 1997	12 bands from 0.402 - 12.5 μm
Envisat	SCIAMACHY	Mar 2002– May 2012	214-2386 nm
Aqua (EOS PM-1)	AIRS	May 2002– ongoing	2,300 spectral channels, 0.4 - 15.4 μm
MetOp	IASI	19 Oct 2006	3 bands, 3.62 - 15.5 μm
GOSAT	TANSO-FTS	Jan 2009– ongoing	4 bands, 0.758- 0.775, 1.56- 1.72, 1.92-2.08, 5.5-14.3 μm
OCO	3 diffraction gratings spectrometers	Feb 2009 - fail	0.765 μm, 1.61 μm, and 2.06 μm.
OCO-2	3 diffraction gratings spectrometers	July 2014– ongoing	0.765 μm, 1.61 μm, and 2.06 μm.
GHGSat-D (Claire)	WFA-P / C&A	Jun 2016– ongoing	WFA-P, 1600- 1700nm, C&A, 400-1000nm
TanSat (CarbonSat)	CarbonSpec, CAPI	Dec 2016– ongoing	758-778, 1594- 1624, 2042- 2082 nm
FY-3D	GAS-FTS	Nov 2017– ongoing	0.75-0.77, 1.56- 1.72, 1.92-2.08, 2.20-2.38 μm
GaoFen-5	GMI	May 2018– ongoing	2.4 – 13.3μm 5 band, 0.75-
GOSAT-2	TANSO-FTS-2	Oct 2018– ongoing	0.77, 1.56-1.69, 1.92-2.33, 5.5- 8.4, 8.4-14.3μm
OCO-3	3 diffraction gratings spectrometers	May 2019– ongoing	0.765, 1.61, and 2.06 μm.
GHGSat-C1 (IRIS)	Fabry-Perot (FP) interferometer	Jun 2020	1600-1700nm

should measure at predefined pixel locations that are partially separated by more than 250 km, and this may lead to gaps in spatial observation. It is also designed to deliver concentrations of CO<sub>2</sub> using the 2060 nm fitting window. The data is delivered as needed for accurate XCO<sub>2</sub> and XCH<sub>4</sub> retrieval [22], [23], [30]. Moreover, a two-year comparison of airborne measurements of CO<sub>2</sub> and CH<sub>4</sub> with GOSAT at the railroad valley, Nevada has demonstrated a highly effective validation method for GOSAT by comparing with the Japan Aerospace Exploration Agency *in situ* data [31]. In 2018, GOSAT-2 [32] was launched, which has a high precision and capability for detecting CH<sub>4</sub> in the 2300 nm wavelength.

NASA's Orbiting Carbon Observatory-2 (OCO-2) mission [33] has been successfully launched in July 2014. The objective of OCO-2 is to collect space-based measurements of CO<sub>2</sub> with the required resolution, precision, and coverage to identify its sources and sinks after quantifying their changes over the seasonal cycle. OCO-2's single instrument is based on three high-resolution grating spectrometers designed to measure

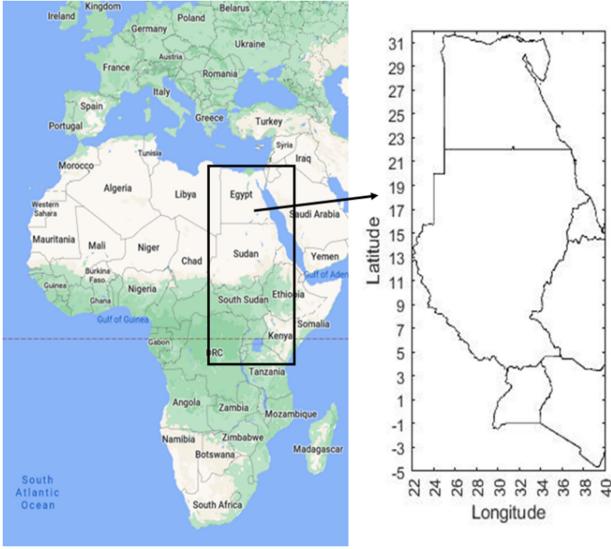


Fig. 2. Study is conducted in the North-East part of Africa, with a covered area from longitude 22 to 40 and latitude from  $-5$  to  $32$ .

the near-infrared absorption of reflected sunlight by  $\text{CO}_2$  and molecular oxygen [34]. OCO-2 covers similar spectral bands as SCIAMACHY and GOSAT; however, OCO-2 has much smaller ground pixels and the swath width is smaller than SCIAMACHY. OCO-2 only delivers  $\text{XCO}_2$  but not  $\text{XCH}_4$  [35].

The European Copernicus program [36] is, currently, preparing the ultra-violet/visible/near infrared/short wave infrared spectrometer instrument onboard Sentinel-5 to be launched in 2021. The mission's main target is the operational monitoring and tracing of gas concentrations for atmospheric chemistry and climate applications. Moreover, the Chinese  $\text{CO}_2$  observation satellite (TanSat) has a spatial resolution of  $1 \times 2 \text{ km}^2$  and a 16-day for temporal resolution [37].

Another small satellite mission that has been recently developed for the precise allocation of GHGs emissions is the commercial Canadian satellite GHGSat CLAIRE, IRIS, and Hugo. These missions were launched in 2016, 2020, and 2021, respectively. GHGSat has the capability to detect only the  $\text{CH}_4$  emissions at industrial facilities from space [38].

### III. CASE STUDY—NORTH-EAST AFRICA

Climate change puts Africa in a hotspot like other areas in the world. For example, Tunisia and Egypt have seen a sharp decrease in rainfall and higher temperatures over the last 50 years [39]–[41]. Dry spells and floods are the most widely recognized as hotspots of environmental condition events in Africa [42] that characterize direct threats to lives and socioeconomic matters. Furthermore, rising temperature related to climatic change is predicted to scale back the agricultural land cover in this part of Africa. As ocean levels rise, Alexandria city, on the Mediterranean coast of Egypt, is predicted to sink [43].

North-East Africa, Fig. 2, faces several problems involving geographical climate impacts in water inadequacy and soil degradation. Moreover, in Egypt, it is projected that a 1-m water level rise can affect millions of people “poor,” living within the Nile River delta basin. Thus, this weakening within the Nile

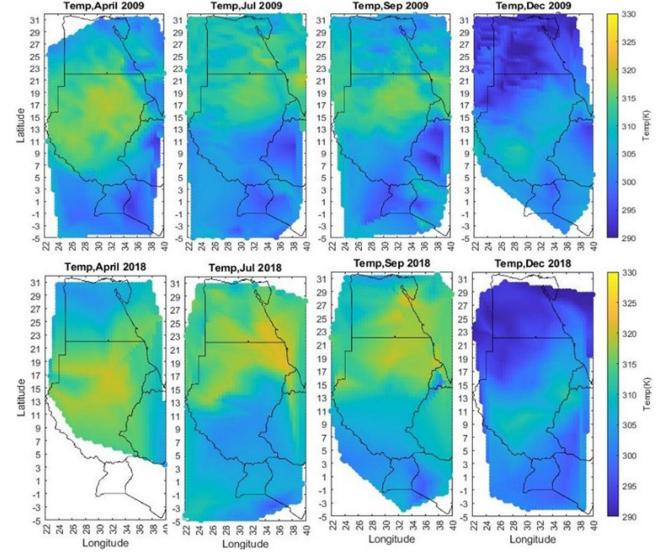


Fig. 3. Change of seasonal temperature ( $^{\circ}\text{K}$ ) by a decade for North-East part of Africa: TANSO-FTS L2 SWIR data.

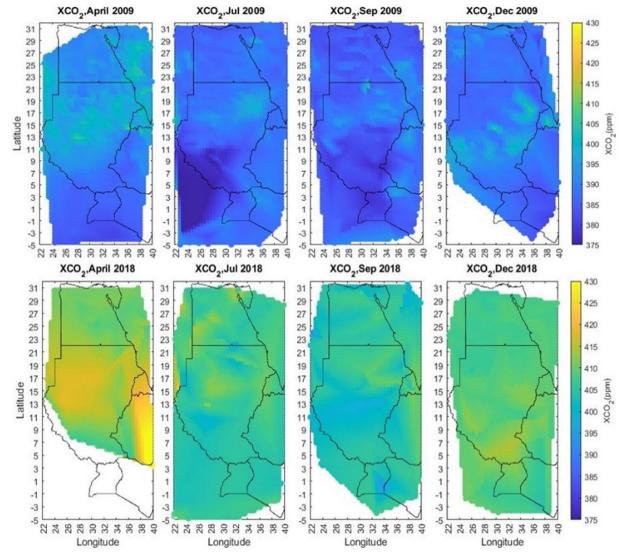


Fig. 4. Change of total column of  $\text{CO}_2$  (ppm) by a decade for North-East part of Africa: TANSO-FTS L2 SWIR data.

River Delta that is heavily populous and used for agriculture shall impact several people [44]. Thus, this weakening within the Nile River Delta that is heavily populous and used for agriculture shall impact several people [44].

This article is conducted in the North-East part of Africa with a covered area from longitude 22 to 40 and latitude from  $-5$  to  $32$ . A sample representation of temperature and total column of  $\text{CO}_2$  for our study area is shown in Figs. 3 and 4. The temperature reached over  $320^{\circ}\text{K}$  for some areas, and  $\text{CO}_2$  reached over 415 parts per million (ppm). The samples are a representation of data for five consecutive files from day 20 to 24 from the months April, July, September, and December for 2009 and 2018; the data are acquired from GOSAT satellite data retrieved by ESA GHG-CCI projects [24]. There is a remarkable change in temperature and  $\text{XCO}_2$  for the same period over

different years, as we can see from Fig. 3 summer gets more increase in temperature, and the winter gets warmer which is a demonstration of excess CO<sub>2</sub> emissions.

### A. Dataset Description

In this article, we used level 2 processed datasets produced by the U.K. NCEO, which is a part of the ESA GHG–CCI and C3S projects [15], [17], [27]. NCEO produced algorithms for retrieving near-surface sensitive column-averaged mixing ratios of CO<sub>2</sub> and CH<sub>4</sub>, denoted XCO<sub>2</sub> and XCH<sub>4</sub>, from SCIAMACHY/ENVISAT (nadir mode) and TANSO/GOSAT [24].

There are two types of retrieval algorithms full-physics (FP) and proxy (PR); FP product is preferred as this product is independent of any (model) assumptions on CO<sub>2</sub>. Where the PR products suffer from scattering related biases and have much more data points [45], we used the FP algorithm results to reduce the uncertainty due to assumptions.

The dataset used in this article has been gathered from the SCIAMACHY which is analyzed based on optimal estimation theory introduced to the remote sensing community by Rodgers (1976) [46]. The Bremen Optimal Estimation DOAS (BESD) algorithm is used to retrieve the column-average dry-air mole fraction of atmospheric CO<sub>2</sub>. Meanwhile, the iterative maximum *a posteriori*-DOAS (IMAP-DOAS) algorithm is used to retrieve CH<sub>4</sub>.

The OCO FP retrieval (OCFP) algorithm is applied to retrieve CO<sub>2</sub> and CH<sub>4</sub> from GOSAT SWIR spectra at the University of Leicester. The OCFP retrieval algorithm was developed to retrieve XCO<sub>2</sub> and XCH<sub>4</sub> from a simultaneous fit of the near-infrared O<sub>2</sub>-A Band spectrum at 0.76 μm and the CO<sub>2</sub> bands at 1.61 and 2.06 μm as measured by the OCO-2 instrument. While the algorithm was developed to retrieve XCO<sub>2</sub> from OCO and OCO-2 observations, it was designed to be adaptable to analyze data from other instruments for algorithm testing and validation. The OCFP algorithm has successfully been used to analyze observations from GOSAT, and ground-based FTS.

We conducted our study based on CO<sub>2</sub>-GOS-OCFP and CH<sub>4</sub>-GOS-OCFP data produced by the University of Leicester. In addition, we used algorithms retrieved for SCIAMACHY: CO<sub>2</sub>-SCI-BESD and CH<sub>4</sub>-SCI-IMAP. The selection of these approaches is based on the studies [16], [27], [47] according to retrieval accuracy compared with other approaches and with the total column carbon observing network (TCCON).

The dataset is presented in the form of column-average dry-air mixing ratios (mole fractions) [24] of CO<sub>2</sub> and CH<sub>4</sub>, denoted XCO<sub>2</sub> (in ppm) and XCH<sub>4</sub> (in parts per billion, ppb). The first sensor data received from the GOSAT satellite TANSO-FTS sensor represent the period from April 2009 to December 2018. Parameters extracted are XCO<sub>2</sub>, XCH<sub>4</sub>, surface air pressure (hPa), and air temperature (°K) of the first level based average kernel. The second sensor data is E SA's ENVISAT satellite–SCIAMACHY sensor, we used data for the period January 2003 to April 2009. Parameters extracted are XCO<sub>2</sub>, XCH<sub>4</sub>, and surface air pressure (hPa). Because of SCIAMACHY temperature parameter is not presented in the dataset provided by Copernicus Climate Data Store, we completed our dataset using ERA5 sensor [48]. We used the ERA5 “2-m temperature” parameter

TABLE II  
COMPLETE STRUCTURE OF THE DATASET INCLUDING PARAMETERS, SOURCE OF DATASET OVERTIME

Parameter	Sensor Data Source	Period of Time
XCO <sub>2</sub>	SCIAMACHY/ENVISAT CO2-SCI-BESD	Jan. 2003 – Apr. 2009
	TANSO-FTS/GOSAT CO2-GOS-OCFP	Apr. 2009 – Dec. 2018
XCH <sub>4</sub>	SCIAMACHY/ENVISAT CH4-SCI-IMAP	Jan. 2003 – Apr. 2009
	TANSO-FTS/GOSAT CH4-GOS-OCFP	Apr. 2009 – Dec. 2018
Pressure	SCIAMACHY/ENVISAT	Jan. 2003 – Apr. 2009
	TANSO-FTS/GOSAT	Apr. 2009 – Dec. 2018
Temp.	ERA5	Jan. 2003 – Apr. 2009
TCWV	ERA5	Apr. 2009 – Dec. 2018
TCRW	ERA5	Jan. 2003 – Dec. 2018

to fill up our temperature parameter for the period from 2003 to April 2009. Furthermore, we used parameters from ERA5 for TCWV (kg m<sup>-2</sup>) and TCRW (kg m<sup>-2</sup>) that are measured as a column density. The temporal description of the dataset is presented in Table II.

The GOSAT temperature is measured over 20 levels of pressure; we select the first level measurements which are the nearest to the earth’s surface to be more compatible with the ERA5 2-m temperature parameter.

Similar studies combined XCO<sub>2</sub> measurements derived from SCIAMACHY and GOSAT for generating global CO<sub>2</sub> maps with high spatiotemporal resolution [49], [50]. Buchwitz *et al.* [47] used datasets of XCH<sub>4</sub> retrievals from SCIAMACHY and GOSAT—generated by different research teams of the GHG–CCI project of the ESA CCI—to identify CH<sub>4</sub> hotspot emissions. The GHG–CCI products used from SCIAMACHY are CH<sub>4</sub>\_SCL\_WFMD and CH<sub>4</sub>\_SCI\_IMAP. In addition, they used the two GOSAT products CH<sub>4</sub>-GOS-OCPR and CH<sub>4</sub>-GOS-SRFP.

## IV. THEORETICAL ANALYSIS AND METHODOLOGY

In this article, we collected atmospheric GHGs datasets from multiple satellite-observations to produce a long time series of spatiotemporal continuous XCO<sub>2</sub> and XCH<sub>4</sub> from satellite observations. Additionally, we collected a set of climatic variables such as temperature, pressure, TCWV, and TCRW, from ERA5 reanalysis dataset to assess the impacts of GHGs emissions on climate variables especially the temperature as described in section “Dataset Description” in (III-A). The flowchart of the work flow in this article is shown in Fig. 5.

The following section describes the flow of the methodology used in the article including intercalibration, spatiotemporal scaling, interpolation/data processing, dataset splitting, and ML approaches.

### A. Intercalibration

The retrieval algorithm uses an iterative retrieval scheme based on Bayesian optimal estimation to estimate a set of atmospheric/surface/instrument parameters, referred to as the state vector, from measured, calibrated spectral radiances.

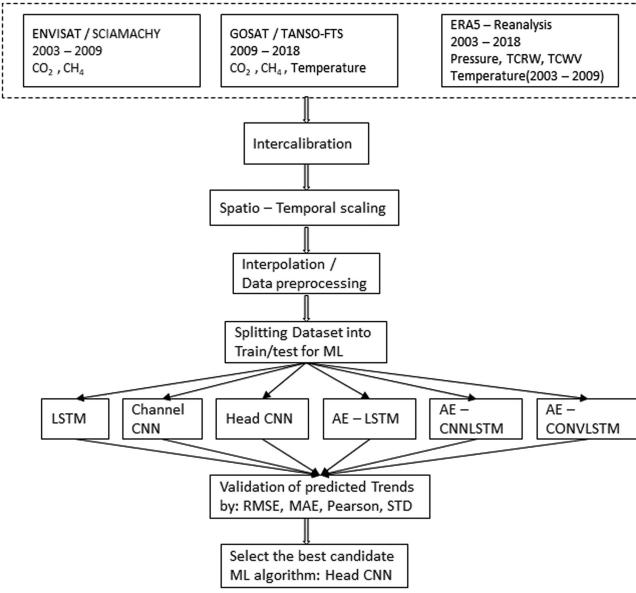


Fig. 5. Flowchart for climate change detection and prediction using machine learning approaches and satellite data.

The inverse method employs the Levenberg–Marquardt modification of the Gauss–Newton method to find the estimate of the state vector with the maximum *a posteriori* probability, given the measurement  $y$  [51]. The state vector will typically include a CO<sub>2</sub> (or CH<sub>4</sub>) profile together with a non-CO<sub>2</sub> state vector. After this iterative retrieval process has converged to a solution, the error covariance matrix  $\hat{S}$  and the averaging kernel (AK) matrix  $A$  are calculated using the *a priori* covariance matrix  $S_a$  and the measurement covariance matrix  $S_\varepsilon$ . XCO<sub>2</sub> is inferred by averaging the retrieved CO<sub>2</sub> profile weighted by the pressure weighting function,  $h$ . The associated column AK for a level  $i$  is then given by  $(a_{\text{CO}2})_i$ , and the variance of XCO<sub>2</sub> by  $\sigma_{\text{XCO}2}$ . The iterative retrieval process equations are represented in (1)–(5):

$$\hat{S} = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} \quad (1)$$

$$A = \frac{\partial \hat{x}}{\partial x} = \hat{S} K^T S_\varepsilon^{-1} K \quad (2)$$

$$\text{XCO}_2 = h^T \hat{x} \quad (3)$$

$$(a_{\text{CO}2})_i = \frac{\partial X\text{CO}_2}{\partial u_i} \frac{1}{h_i} = (h^T A)_i \frac{1}{h_i} \quad (4)$$

$$\sigma_{\text{XCO}2} = h^T \hat{S} h. \quad (5)$$

The variance  $\sigma$  is used to measure the retrieved data uncertainty; the retrieved X<sub>GHG</sub> and its uncertainties data after applying the retrieval process for the research study area, North-East part of Africa are represented in Figs. 6 and 7. Comparing the output of the retrieval algorithms with ground-based TCCON observations exposed that the retrieved data are very comparable with respect to biases. The XCO<sub>2</sub> bias range of the retrieval algorithms BESD/SCIAMACHY and OCFP/GOSAT should not exceed the 8 ppm as shown in Fig. 6(b). Meanwhile, the XCH<sub>4</sub> bias range of the OCFP/GOSAT retrieval algorithm should be below 34 ppb as shown for the period 2009–2018 in Fig. 7(b).

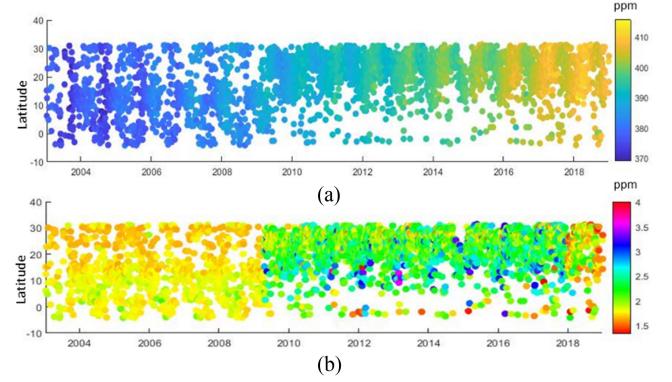


Fig. 6. Retrieved XCO<sub>2</sub> after performing the retrieval process for the research study area North-East part of Africa: (a) CO<sub>2</sub> and (b) CO<sub>2</sub> bias.

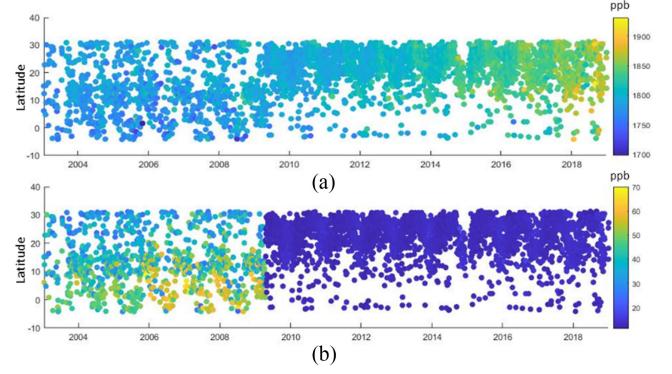


Fig. 7. Retrieved XCH<sub>4</sub> after performing the retrieval process for the research study area North-East part of Africa: (a) CH<sub>4</sub> and (b) CH<sub>4</sub> bias.

This is also true for the estimated measurement precisions for IMAP/SCIAMACHY for the time period 2003–2005; however, the SCIAMACHY detector suffered from a major degradation in its spectral region needed for CH<sub>4</sub> retrieval, which explains the major bias values after 2005.

#### B. Spatiotemporal Scaling

Since the data are gathered from multiple satellite sensors as mentioned above in Section III-A, we had to unify the measurement scale and calibrate for the time of different data acquisition, as represented in Table II. GOSAT revisit time (time of regular data acquisition) over North-East Africa is between 10:00 and 11:35 AM, while the other part of data acquired by ERA5 is an hourly tagged dataset, so we used data with the same time tag. In contrast for SCIAMACHY/ENVISAT, its revisit time was from 7:00 to 8:35 AM; therefore, we used the ERA5 parameters at 11 AM to be compatible with GOSAT variables, especially for the temperature.

#### C. Interpolation/Data Processing

AKs are used in case of using different satellite products because of measurement sensitivity at different altitudes through the atmospheric column. The corrections are based on vertical priori profiles which should be adjusted to a common profile

when comparing XGHGs from different instruments. Moreover, the retrieval smoothing effect must be considered by applying the AKs [44] to decrease the effects from different instruments on XGHGs retrievals.

The AK and related parameters are provided for “layer-based AKs” and “level-based AKs” for layer-based AK;  $m$  is the number of layers which are defined by  $k = m+1$  pressure levels. For level-based AK, only levels are used; all parameters have the same number of elements, namely  $m$  levels. Here, the number of pressure levels is also  $m$  (i.e.,  $k = m$ ). The layer-based AK approach is applied on CO<sub>2</sub>-SCI-BESD and CH<sub>4</sub>-SCI-IMAP, while level-based AK approach is applied on the University of Leicester products CO<sub>2</sub>-GOS-OCFP and CH<sub>4</sub>-GOS-OCFP.

The column AK vector is defined as sensitivity of the retrieved XCO<sub>2</sub> to the true layered CO<sub>2</sub> mixing ratios. In the ideal case, the column AK vector would be identical unity. This would mean that a XCO<sub>2</sub> change introduced by a change of the  $i$ 'th layer is one-to-one reproduced by the retrieved XCO<sub>2</sub>.

The following equation is used for layer and level based average kernels:

$$\text{XGHG}^{\text{mod}} = \sum_{i=1}^m [\text{VMR}_i^{\text{apri}} + (\text{VMR}_i^{\text{mod}} - \text{VMR}_i^{\text{apri}}) \text{AK}_i] \text{pw}_i \quad (6)$$

where XGHG<sup>mod</sup> is the desired modeled XCO<sub>2</sub> or XCH<sub>4</sub> value, which corresponds to the satellite retrievals. The sum is over the  $m$  atmospheric layers (located between pressure levels  $p_i$  and  $p_{i+1}$  with  $i = 1 \dots m$ ). Here,  $i = 1$  corresponds to the bottom of the atmosphere and  $i = k = m + 1$  corresponds to the top of the atmosphere.

$\text{pw}_i$  is a layer-dependent pressure weight vector.  $\text{VMR}_i^{\text{apri}}$  is the satellite *a priori* layer-averaged CO<sub>2</sub> or CH<sub>4</sub> volume mixing ratio (VMR) or, more precisely, dry mole fraction (DMF), between pressure levels  $P_i$  and  $P_{i+1}$  ( $P_i > P_{i+1}$ ).

$\text{VMR}_i^{\text{mod}}$  is the corresponding value of the model (CO<sub>2</sub> or CH<sub>4</sub>) VMR (DMF) between pressure levels  $P_i$  and  $P_{i+1}$ .  $\text{AK}_i$  is the satellite XCO<sub>2</sub> or XCH<sub>4</sub> AK for layer  $i$ .

#### D. Dataset Splitting

We used data for the period from 2003 to 2015 as training, while data from 2016 to 2018 for validation. We split the data into standard weeks, by organizing the data into weeks gives 677 full standard weeks for training a predictive model, whereas the test dataset has 156 weeks. We used neural networks and as such, they are slow to train but fast to evaluate. The preferred usage of the models is to build them on historical data and then use them to predict each step of the walk-forward validation. The models are static (not updated) during their evaluation. Table II shows the structure of the dataset, which includes parameters of climate change and its corresponding data sources; the parameters are XCO<sub>2</sub>, XCH<sub>4</sub>, pressure, temperature, TCWV, and TCRW.

For temperature, CO<sub>2</sub>, and CH<sub>4</sub>, we used ML algorithms to predict their values over time of validation (677 weeks for training and 156 weeks for validation). We used all parameters

for neural network training; that means we included the effect of pressure, water vapor, and rainwater on the temperature, CO<sub>2</sub>, and CH<sub>4</sub>, in addition to the effect of these parameters on each other.

We implemented the ML algorithms using Python3 programming language. The neural network models are implemented using Keras deep learning library on top of Google TensorFlow [52].

We used the following performance evaluation metrics to assess the models prediction outcome: root-mean-squared-error (RMSE) in (7), mean-absolute-error (MAE) in (8), Pearson in (9), standard deviation in (10), and  $R^2$  coefficient to evaluate the performance of each model [53]. The results of Pearson and  $R^2$  vary between  $\pm 1$  that indicates how the variables are related to a positive or negative relationship. For the naïve methods, we used only RMSE as it is more punishing of prediction errors, unlike MAE:

$$\text{RMSE}(\bar{X}, X) = \sqrt{\frac{1}{N} \cdot \sum_{n=1}^N (\bar{X}_n - X_n)^2} \quad (7)$$

$$\text{MAE}(\bar{X}, X) = \frac{1}{N} \cdot \sum_{n=1}^N |\bar{X}_n - X_n| \quad (8)$$

$$r_{\bar{X}X} = \frac{N \sum_{n=1}^N X_n \bar{X}_n - \sum_{n=1}^N X_n \sum_{n=1}^N \bar{X}_n}{\sqrt{N \sum_{n=1}^N X_n^2 - \left( \sum_{n=1}^N X_n \right)^2} \sqrt{N \sum_{n=1}^N \bar{X}_n^2 - \left( \sum_{n=1}^N \bar{X}_n \right)^2}} \quad (9)$$

$$\text{STD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (10)$$

where  $\bar{X}_n$  and  $x_i$ ,  $X_n$ ,  $N$ ,  $\mu$ , and  $r_{\bar{X}X}$  represent predicted value, observed value, the mean of predicted values  $x_i$ , the number of data, and Pearson correlation coefficient, respectively; STD is the standard deviation.

The neural network models are evaluated using a scheme called walk-forward validation [54]. This is where the model is required to make a one-week prediction, and then the actual data for that whole week are made available to the model so it can be used as a basis for predicting the subsequent week.

## V. MACHINE-LEARNING APPROACHES

A group of ML experts [4] described climate change as one of the greatest challenges facing humanity. Furthermore, they invite the ML community to join the global effort against climate change. Recent trends have created opportunities for ML to advance the state-of-the-art in climate prediction [55]. Many challenges have been the motivation for introducing the ML technique as a reliable tool for addressing climate change phenomena. First, GHGs monitoring satellites are creating petabytes of climate observation data. Second, enormous computational

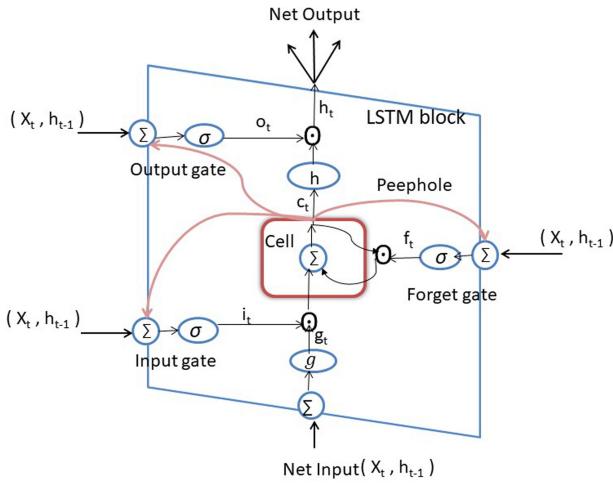


Fig. 8. Long short-term memory block.

climate modeling projects generate petabytes of simulated climate data. Third, climate forecasting is computationally expensive [56], [57]; meanwhile, ML models are becoming gradually fast to train and run, particularly on next-generation computing hardware [57]–[59].

ML may be able to help with many aspects of CO<sub>2</sub> removal, and also identify and characterize potential storage locations. A deep learning approach in [60] is used to improve the retrieval of temperature and humidity profiles from a ground-based microwave radiometer.

We examined different types of neural network models for multistep multivariate time series forecasting: long short-term memory (LSTM), convolutional neural network (CNN) (multiple input channels and multiple input heads), and autoencoders (LSTM, CNNLSTM, and CONVLSTM). Moreover, we applied different naïve linear regression (LR) approaches such as LR, ridge regression, huber regression (Hber), least-angle regression (LARS), passive-aggressive regression, random sample consensus (RANSAC) regression, and stochastic gradient descent (SGD) regression [61].

#### A. Long Short-Term Memory

LSTM is developed from recurrent neural networks (RNN) by replacing the RNN hidden layer neurons with LSTM blocks, as shown in Fig. 8. Each block has a memory cell that helps overcome the RNN vanishing gradient problem [62], illustrated by (11)–(16). This technique has been recently used in many fields including commercial, medical, and space fields [14]. However, replacing matrix multiplication with the convolution operation at each gate in the LSTM block results in the ConvLSTM approach, which is effective in multidimensional time-series data [63] as illustrated by the equations in (17)–(22). In [64], ConvLSTM is used as a Hurricane-Tracker for its capability of tracking and predicting extreme climate events.

LSTM can be used in our article to handle a large amount of data received from satellites; since the data are huge and arranged in a time-series manner, we expected to have good results.

A recent study on LSTM applied to satellite data, suggested that, the algorithm performs well whenever larger data are arranged in time series, compared to gated recurrent units (GRU) and multilayer perceptron [65], [66].

The LSTM block equations are shown in (11)–(16):

$$i_t = \sigma(w_{xi}x_t + w_{hi}h_{t-1} + w_{ci}p_{t-1} + b_i) \quad (11)$$

$$f_t = \sigma(w_{xf}x_t + w_{hf}h_{t-1} + w_{cf}p_{t-1} + b_f) \quad (12)$$

$$o_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + w_{co}p_t + b_o) \quad (13)$$

$$g_t = g(w_{xg}x_t + w_{hg}h_{t-1} + b_g) \quad (14)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (15)$$

$$h_t = o_t \odot h(c_t). \quad (16)$$

While the ConvLSTM block equations are in (17)–(22):

$$i_t = \sigma(w_{xi} * x_t + w_{hi} * h_{t-1} + w_{ci} * p_{t-1} + b_i) \quad (17)$$

$$f_t = \sigma(w_{xf} * x_t + w_{hf} * h_{t-1} + w_{cf} * p_{t-1} + b_f) \quad (18)$$

$$o_t = \sigma(w_{xo} * x_t + w_{ho} * h_{t-1} + w_{co} * p_t + b_o) \quad (19)$$

$$g_t = g(w_{xg} * x_t + w_{hg} * h_{t-1} + b_g) \quad (20)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (21)$$

$$h_t = o_t \odot h(c_t) \quad (22)$$

where  $\sigma$  is a sigmoid activation function with range [0, 1],  $g$  and  $h$  are tanh activation function with range [-1, 1],  $*$  is the convolutional operation,  $\odot$  is a pointwise vector product,  $w_{xi}$ ,  $w_{hi}$  ... are the weight matrices, and  $b_i$ ,  $b_f$  ... are the bias.

The input, forget, and output gates are represented by  $i_t$ ,  $f_t$ , and  $o_t$  in (11)–(13), respectively. The new memory content  $g_t$  is represented in (14), (15) is the current memory content  $c_t$ , while (16) is the LSTM block output  $h_t$  [67].

#### B. Convolutional Neural Network

A CNN is the basis of most computer vision technologies. It uses “convolution” and “pooling” operations to reduce input data into its essential features and uses those features to understand and obtain the desired output [68]. The convolutional layers can read large input data as a sequence and automatically extract features. The pooling layers can refine the extracted features and focus attention on the most salient elements. Then, the fully connected layers can interpret the internal representation and output a vector representing multiple time steps. The key benefits of the approach are the automatic feature learning and the ability of the model to output a multistep vector directly [56], [69]. The configuration of the CNN architectures is divided into two groups multihead (H-CNN) and multichannel (CH-CNN). The H-CNN uses independent single-channel convolutional heads to process each time series sensor data separately, and the output of all heads is concatenated before reaching the fully connected part, as shown in Fig. 9. The latter, CH-CNN, uses a single convolutional head with multiple channels to process all the time-series data [70].

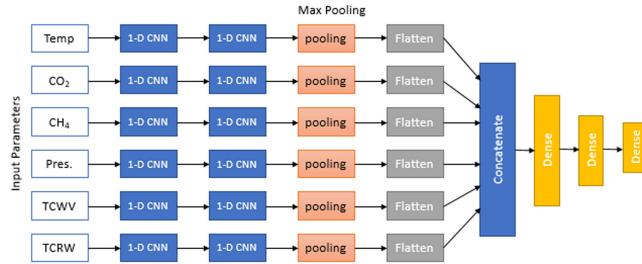


Fig. 9. Multiheaded multistep CNN applied on essential climate variables.

CNN and RNN approaches are not mutually exclusive, combining CNN and RNN (LSTM) results in CNNLSTM architecture, which is particularly promising in various domains such as weather recognition [71], gesture recognition [72], estimation of the TCWV [13], and speech recognition [73]. Furthermore, they facilitate the analysis of inputs over longer periods than could be achieved with lower level RNN architecture types [74].

### C. Autoencoders

An unsupervised artificial neural network technique called autoencoder that learns in what way to efficiently compress and encode data after that knows how to reconstruct the data back from the reduced encoded representation to a representation that is as adjacent to the original input as possible [62]. The design of autoencoders works on decreasing data dimensions by learning how to ignore the amount of noise in the data. The network architecture for autoencoders can vary between a simple Feedforward, LSTM, CNN, ConvLSTM, and CNNLSTM networks or a mixture of them depending on the use case [69], [75]. Furthermore, an unsupervised analysis for change detection and clustering approach is applied on satellite image time-series data by combining neural network GRU with autoencoder-based model [76].

In this article, the autoencoder is used to extract the characteristics of original climate variables data. Then, we use LSTM, CNN, ConvLSTM, or CNNLSTM models to predict the climate variables based on data acquired from satellite sensors.

### D. Linear Methods for Regression

LR is a linear approach to modeling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables). The ridge and lasso approaches shrink the regression coefficients by imposing a penalty on their size. The ridge coefficients minimize a penalized residual sum of squares. LARS is an algorithm for fitting LR models to high-dimensional data.

Hber is an alternative approach that is based on a slightly modified loss function, called Huber loss. It defines a threshold that makes the loss function switch from a squared error to an absolute one [77].

RANSAC is an iterative method to estimate parameters of a mathematical model from a set of observed data that contains outliers. It is a nondeterministic algorithm in the sense that it

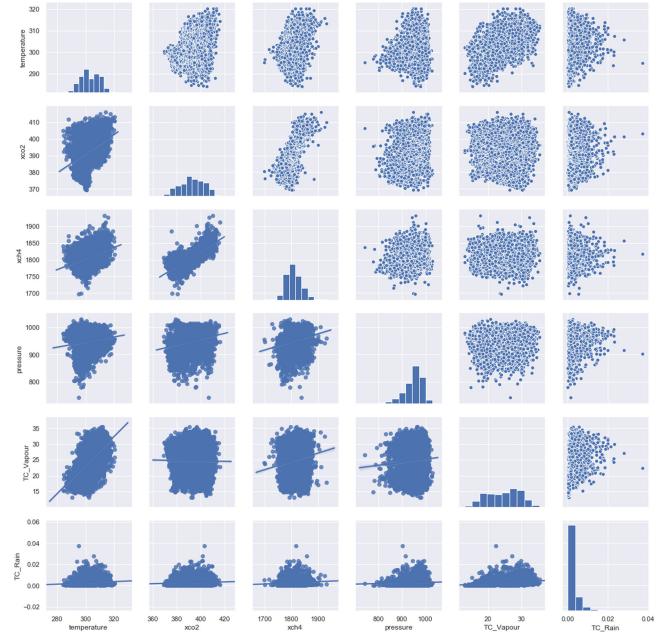


Fig. 10. Relation between the essential climate variables for the North-East part of Africa dataset.

produces a reasonable result only with a certain probability with this probability increasing as more iterations are allowed.

Passive-aggressive behavior is characterized by a pattern of passive hostility and an avoidance of direct communication. Inaction, where some action is socially customary, is a typical passive-aggressive strategy.

SGD is an iterative method for optimizing an objective function with suitable smoothness properties (e.g., differentiable or subdifferentiable). It can be regarded as a stochastic approximation of gradient descent optimization, since it replaces the actual gradient (calculated from the entire dataset) by an estimate thereof (calculated from a randomly selected subset of the data). SGD is a simple and very efficient approach to fitting linear classifiers and regressors under convex loss functions such as (linear) support vector machines and logistic regression [61].

## VI. RESULTS AND ANALYSIS

The relationship between the climate variables ( $XCO_2$ ,  $XCH_4$ , pressure, temperature, TCWV, and TCRW) used in this article is shown in Fig. 10. The graphs show the trends and dependency between variables and each other; the trend line is shown crossing the data graphs. Each trend indicates the dependency between two variables. From the relationship between the climate variables presented in Fig. 10, we found that temperature,  $XCO_2$ , and  $XCH_4$  are mostly dependent. This is clear from the regression line between the temperature and  $XCO_2$  which shows a correlation of 39% as well as the same percent for  $XCH_4$ . This result indicates that these three parameters (ECVs) are the most significant factor in studying the climate change. For each of these variables, we applied the ML models for prediction, and then we calculated the evaluation metrics for assessment. In Tables III and IV, the assessment of prediction accuracy

TABLE III  
COMPARISON OF THE PERFORMANCE OF DIFFERENT METHODS

	Method	RMSE	MAE	Pearson	$R^2$	STD
<b>Temp</b>	LSTM	5.421	4.182	0.329	0.574	3.883
	CH-CNN	6.118	4.819	0.417	0.646	4.498
	H-CNN	5.378	4.157	0.368	0.607	5.252
	AE-LSTM	7.681	6.365	0.435	0.660	3.919
	AE-CNNLSTM	5.606	4.304	0.291	0.539	3.170
	AE-CONVLSTM	5.287	4.095	0.320	0.566	3.648
<b>CO<sub>2</sub></b>	LSTM	6.107	5.649	0.389	0.624	2.427
	CH-CNN	1.836	1.347	0.627	0.792	2.582
	H-CNN	2.395	1.928	0.649	0.806	2.751
	AE-LSTM	6.599	6.316	0.553	0.743	2.283
	AE-CNNLSTM	7.661	7.015	0.074	0.272	2.107
	AE-CONVLSTM	6.979	6.411	0.210	0.459	2.275
<b>CH<sub>4</sub></b>	LSTM	17.137	11.766	0.227	0.476	14.155
	CH-CNN	17.846	13.051	0.219	0.468	13.997
	H-CNN	15.923	11.672	0.291	0.539	13.211
	AE-LSTM	16.305	12.044	0.323	0.568	13.063
	AE-CNNLSTM	17.850	13.294	0.251	0.501	13.851
	AE-CONVLSTM	18.435	14.032	0.139	0.373	10.687

TABLE IV  
RMSE COMPARISON OF THE PERFORMANCE OF DIFFERENT NAÏVE METHODS

Method	LR	Ridge	Hber	LARS	PA	Rsac	SGD
TEMP.	5.178	5.168	5.248	5.178	6.723	5.755	5.12
CO <sub>2</sub>	1.69	1.686	1.706	1.69	4.856	1.696	1.682
CH <sub>4</sub>	19.472	19.513	16.154	19.472	17.32	17.473	18.838

of each model applied for the essential climatic variables was summarized.

#### A. Prediction Methods Accuracy Assessment

1) *ML Prediction Accuracy Assessment:* In the results of Table III, a comparison between different architectures of LSTM and CNN is presented. Multihead CNN and the multichannel CNN give a comparable result; however, multihead CNN is promising better accuracy compared to multichannel architectures as the number of features extracted from concatenated heads presents more information that helps in the prediction process. Furthermore, H-CNN, CH-CNN, and LSTM give the best prediction accuracy for RMSE and MAE, where the H-CNN gives 5.378, and 4.157 ppm in RMSE and MAE, respectively, comparing these values w.r.t the uncertainty from the XCO<sub>2</sub> retrieval algorithm discussed in Section IV-A that should be less than 8 ppm. Additionally, the H-CNN gives 15.923 and 11.672 ppb in RMSE and MAE, respectively, for XCH<sub>4</sub> which is still less than the uncertainty of the XCH<sub>4</sub> retrieval algorithms

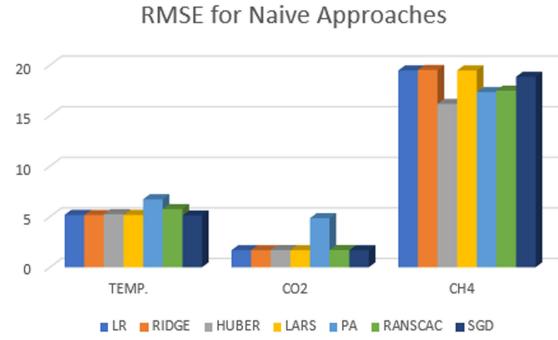


Fig. 11. RMSE for different Naïve approaches for the essential climate variables of temperature, CO<sub>2</sub>, and CH<sub>4</sub>.

that equal to 34 ppb; these results can overcome the degradation that occurred to CH<sub>4</sub> detectors of SCIAMACHY after 2005.

Alternatively, autoencoder LSTM and multichannel CNN (CH-CNN) achieved the highest performance accuracy regarding the Pearson coefficient and  $r^2$  correlation coefficient.

The autoencoders (AE-ConvLSTM and AE-CNNLSTM) show a comparable accuracy especially for STD since the convolutional part is to extract the most related characteristics from the data. Despite that, the autoencoder part performs well on encoding and decoding, this might decrease extracted features. One of the main differences between the proposed approaches is the number of features extracted from each one. Consequently, more features the architecture can obtain means more information to get better results.

This means that multihead CNN is the most reliable model for prediction of ECVs; however, multihead CNN should be used with careful to avoid overfitting.

2) *Naïve Methods Prediction Accuracy Assessment:* The performance of naïve methods is shown in Table IV and Fig. 11; the RMSE of each approach illustrates that the SGD has less RMSE in the prediction of CO<sub>2</sub>, CH<sub>4</sub>, and temperature as well. SGD's nature always tries to converge to the global minimum which reduces the variance of the parameter updates, leading to be more stable convergence.

The SGD regressor performing the best with an overall RMSE for the three essential climate parameters (temperature, CO<sub>2</sub>, and CH<sub>4</sub>) compared to other regression methods, shown in Table IV, followed by HuberRegressor, LinearRegression, and LARS. As a result, for naïve methods, we can conclude the SGD as the best candidate, which makes it a best choice as an optimization algorithm for state-of-the-art deep learning libraries.

#### B. ML Prediction Efficiency

The evaluation of predicted values of the different techniques compared to the real data is shown in Figs. 12 and 13; as we can see from the chart, the multihead CNN gives a comparable result to the real values as it has the highest features extracted; other techniques such as channel CNN and LSTM models have the second better prediction performance followed by AE-ConvLSTM and AE-CNNLSTM.

Fig. 13 shows the monthly mean temperature from 2016 to 2018, the graph shows the real data (red) versus the predicted

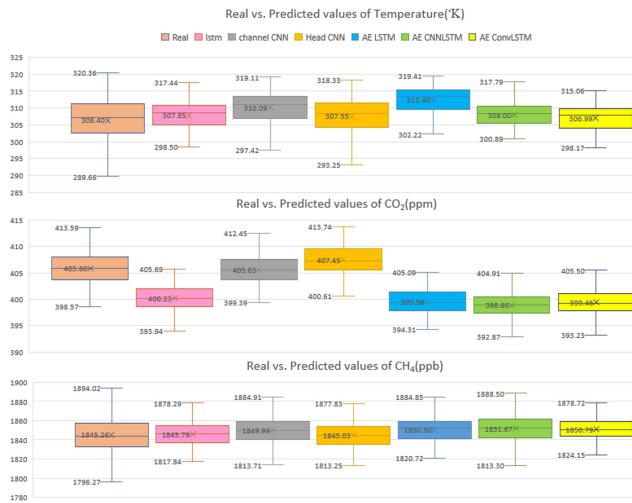


Fig. 12. Real vs. predicted values of different applied ML approaches for the essential climate variables of temperature, CO<sub>2</sub>, and CH<sub>4</sub>.

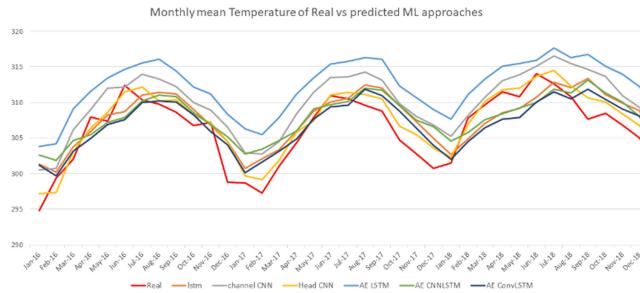


Fig. 13. Monthly mean temperature (°K) of real data (red) vs. the different ML approaches used for prediction.

values of different ML models. The H-CNN (yellow) gives similar behavior to the real trend. The model can successfully mimic the real data promising to be used as a simulation tool to forecast future trends of the temperature.

The head-CNN model vs. the real data for ECVs are shown in Fig. 14, where the black line represents a perfectly steady slope, while the blue line is the regression line representing the data distribution trend, for the ECVs (temperature, CO<sub>2</sub>, and CH<sub>4</sub>).

### C. Prediction of Essential Climatic Variables

We can rely on these data to conclude that the multiheaded CNN and autoencoder models are best candidates in climate change prediction based on the given ECVs. The predicted maps of ECVs using multihead CNN are presented in Figs. 15–17 for the years 2016, 2017, 2018, and predicting the trend in 2030. The subfigures show how the temperature, CO<sub>2</sub>, and CH<sub>4</sub> are changed over the years. By 2030, the increase in atmospheric CO<sub>2</sub> and CH<sub>4</sub> concentrations may reach 440 ppm for CO<sub>2</sub>, and more than 2000 ppb for CH<sub>4</sub>; and therefore, there would be a significant increase in the temperature values that may reach 2 °C by 2030. Further studies have been performed to forecast the temperature and GHGs emissions in 2030 and obtained a similar result. Lazar and Williams [78] evaluated the climate change resulted from increased GHG emissions. They used the

general circulation models to predict the change in temperature and precipitation for the years 2030 and 2100 using GHGs emissions scenario for the years 2030 and 2100, relative to 1990. They found that the temperature will increase 2 °C by 2030 and 6 °C by 2100. Amin *et al.* [79] used a climate model based on ensemble learning for evaluation and analysis of temperature for historical (1996–2015) and projected (2030–2060) climates in Pakistan. They concluded that maximum temperature projected for 2030 and 2060 was worth noting where negative trends were increased as compared to past temperature trends on annual and seasonal (summer) basis; furthermore, minimum temperature projected for future 2030 and 2060 revealed that negative trends were increased by almost three times for the number of weather stations (larger for 2060 than 2030) then past weather trends (1996–2015) for summer and annual.

## VII. DISCUSSION

Building an adaptation model is a process which is used to tackle climate change; a carbon—CO<sub>2</sub> and CH<sub>4</sub>—emission model using ML approaches and remote sensing data can improve the accuracy of carbon monitoring and estimation. The developed model helps to unify the world's sustainability data into a single scheme. The ML models have the ability to recalculate the carbon emissions and identify which features influence the emissions the most. Moreover, the ML approaches are used to identify the impacts of these emissions on temperature and global warming.

In this article, we studied the general climate parameters, and justified our selection for the ECVs; then we applied different ML models to these ECVs. Best model has been identified and applied for these climate variables for prediction.

The prediction results have been discussed in some details and we concluded a reliable model for climate trend prediction. The multihead CNN used for climate change detection and prediction revealed that by 2030, there will be a significant increase in temperature that may reach 2 °C above the normal temperature range; the algorithm also showed that this excess of temperature is due to, mainly, the predicted growth rate of atmospheric CO<sub>2</sub> which is expected to reach around 440 ppm, in addition to the concentration of CH<sub>4</sub> that is predicted to reach more than 2000 ppb.

These results give an early warning alert to countries across the world about the industrial emissions of GHGs. Despite that ML techniques are still predictive methods that use available data for learning, it has become a reliable tool for decision making and supports futuristic studies. The general limitation of ML models can be summarized in next section.

### A. Prediction Limitations of ML Models vs. Climate Models

Many climate prediction models depend on data from weather stations; these models suffer from limited data. Existing climate models deal with this limitation by relying heavily on physical laws that is computationally expensive to be included in global climate models; however, ML models are able to efficiently solve such systems. Gentine *et al.* [80] trained a deep learning model to emulate the behavior of a high-resolution cloud simulation,

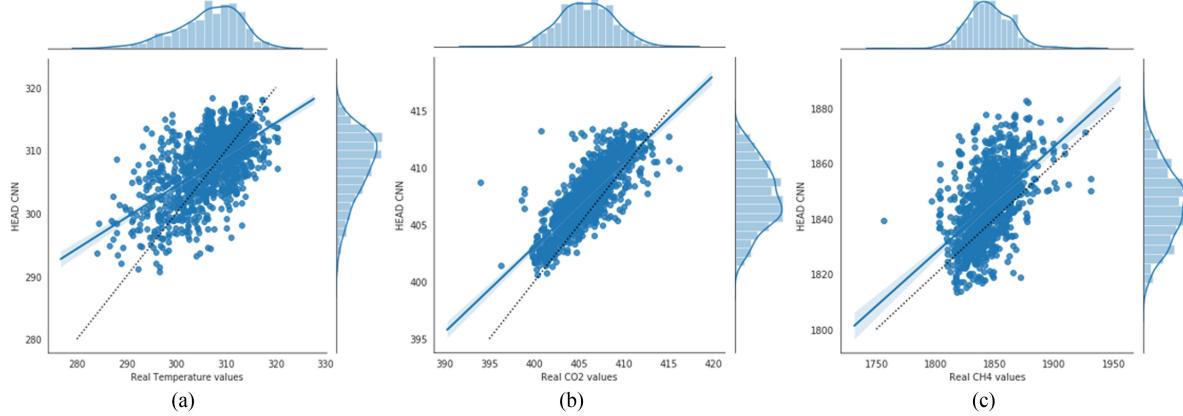


Fig. 14. Head-CNN model vs. the real data for essential climate variables where the black line represents a perfectly steady slope, while the blue line is the regression line representing the data distribution trend; for A-Temperature, B-CO<sub>2</sub>, and C-CH<sub>4</sub>.

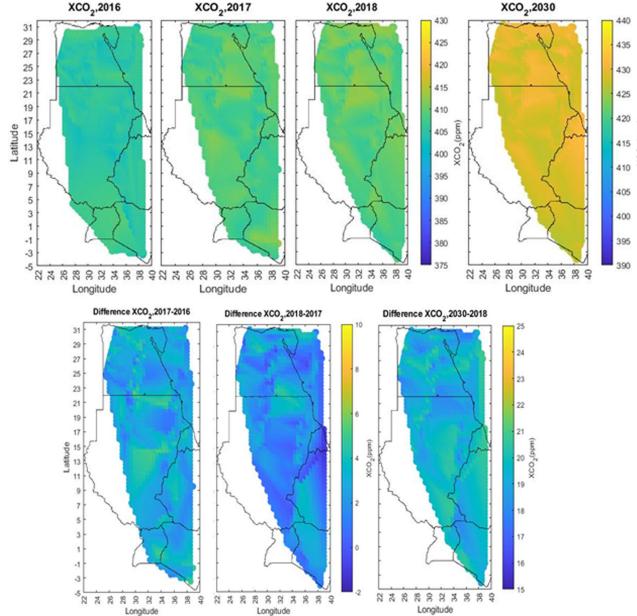


Fig. 15. Predicted maps of CO<sub>2</sub> using multihead CNN for the years 2016–2018 and the prediction for 2030. The subfigures are the CO<sub>2</sub> difference for years 2017–2016, 2018–2017, and the predicted 2030–2018.

and their model gave comparable results for a fraction of the cost and was stable in a simplified global model.

Climate models are applied at the scale of years to decades; their long-term trends are driven by slow, predictable changes of ocean, land, and ice, which make them chaotic on short time scales. ML models, in contrast, could bridge that gap by making good predictions at intermediate time scales of weeks to months. Running such models on weather data gives an early caution by forecasting extreme events [4].

From an ML perspective, learning from an existing model has many advantages: modelers can generate new training and test data on-demand, and the new ML model inherits some community trust from the old one. Despite all of the ML advantages, there are some limitations of ML models especially when applied in the geosciences [81] as follows:

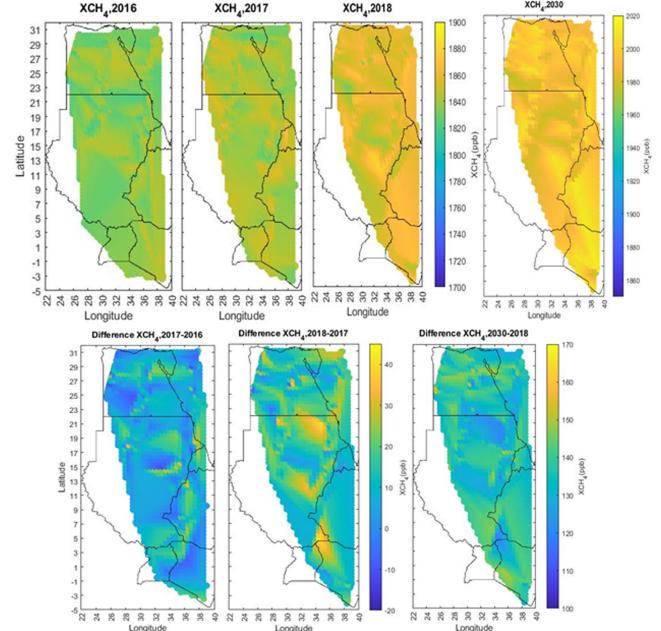


Fig. 16. Predicted maps of CH<sub>4</sub> using multihead CNN for the years 2016–2018 and the prediction for 2030. The subfigures are the CH<sub>4</sub> difference for years 2017–2016, 2018–2017, and the predicted 2030–2018.

- 1) *Interpretability*: Improving predictive accuracy is important but insufficient. Certainly, interpretability and understanding are crucial, including visualization of the results for analysis by humans. This presents a challenge in this domain, which usually requires a domain expert to assure proper interpretation of data and results. Achieving a self-explanatory model is still far away.
- 2) *Physical consistency*: ML models can fit observations very well, but predictions may be physically inconsistent due to observational biases. This can be solved by teaching the model physical rules of the Earth system.
- 3) *Complex and uncertain data*: ML methods are needed to deal with complex statistics, multiple inputs and outputs, different noise sources, and high-dimensional spaces.

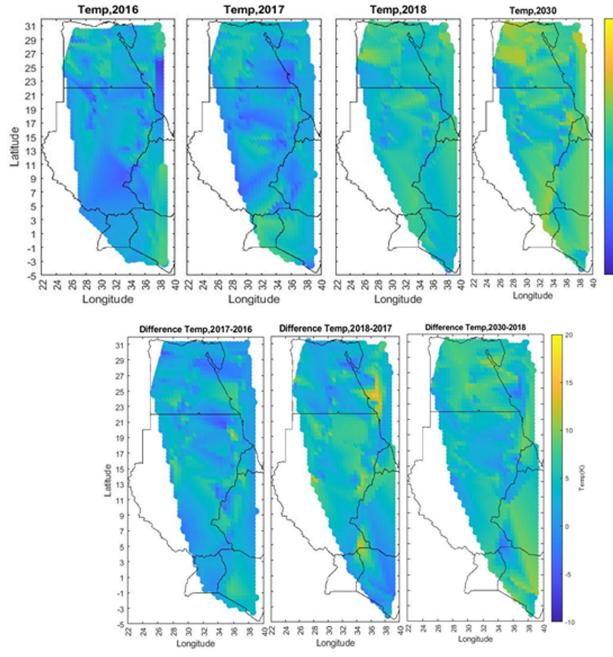


Fig. 17. Predicted maps of temperature using multihead CNN for the years 2016–2018 and the prediction for 2030. The subfigures are the temp. difference for years 2017–2016, 2018–2017, and the predicted 2030–2018.

4) *Computational demand:* There is a huge technical challenge regarding the high computational cost of current geoscience problems—an example is Google’s Earth Engine, which allowed the solution of real problems like deforestation and lake monitoring.

## VIII. CONCLUSION

In this article, we addressed the challenge of predicting the status of climate change based on a set of variables acquired from remote sensing data. We collected a dataset from space system sensors including SCIAMACHY/ENVISAT, CH<sub>4</sub>-SCI-IMAP, and CO<sub>2</sub>-SCI-BESD; TANSO-FTS/GOSAT, CH<sub>4</sub>-GOS-OCFP, and CO<sub>2</sub>-GOS-OCFP and ERA5. The data were preprocessed to unify the rhythm. We applied a set of ML approaches; starting from NAÏVE Methods to CNN and LSTM, based on parameters affecting climate change. The parameters are XCO<sub>2</sub>, XCH<sub>4</sub>, pressure, temperature, TCWV (kg m<sup>-2</sup>), and TCRW (kg m<sup>-2</sup>). The applied models are LSTM, CNN (multiple input channels and multiple input heads), and autoencoders (LSTM, CNNLSTM, and CONVLSTM). Moreover, we applied different naïve LR approaches such as LR, Ridge, huber regressor, LARS, passive-aggressive regressor, RANSAC regressor, and SGD regressor.

We compared the performance of ML techniques to determine the best candidate algorithm that can be used in forecasting climate variables based on data acquired from satellite sensors. The RMSE, MAE, Pearson coefficient,  $R^2$  coefficient, and standard deviation have been used to assess the algorithms.

We concluded that multihead CNN and autoencoder could be used effectively to predict climate variables compared with other neural networks such as traditional CNN, LSTM; while for

naïve methods, we concluded that SGD can be a good candidate as well. We can rely on these data to select the multiheaded CNN to generate predicted maps of ECVs for the years 2016, 2017, 2018, and predicting the trend in 2030.

By 2030, the increase in atmospheric CO<sub>2</sub> and CH<sub>4</sub> concentrations may reach 440 ppm for CO<sub>2</sub>, and more than 2000 ppb for CH<sub>4</sub>; and therefore, there would be a significant increase in the temperature values that may reach 2 °C by 2030. This obvious change in concentrations causes warming and is affecting various aspects of climate, including air and ocean temperatures, precipitation, and sea levels. Human health, agriculture, water resources, forests, wildlife, and coastal areas are all vulnerable to climate change.

The multihead CNN achieved the highest accuracy in linking between the temperature as the main climate variable from one side and CH<sub>4</sub> and CO<sub>2</sub> as GHG factors from the other side. With this approach, we can rely on the GHG measurements acquired from the neighborhood regions to forecast for the climate change variable in North-East Africa. The forecasting can help reduce the effect of climate changes on the life ecosystem in this region.

The future trend in this article will be to effectively combine the climate models, which are based on physics and the equation of climate environment, with ML models, which have the ability to forecast and deal with massive data. However, this trend is quite complex, it will unveil the full power of both models.

## REFERENCES

- [1] C. M. Wainwright, D. L. Finney, M. Kilavi, E. Black, and J. H. Marsham, “Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change,” *Weather*, vol. 76, pp. 26–31, 2021.
- [2] C. Funk, “Ethiopia, Somalia and Kenya face devastating drought,” *Nature*, vol. 586, pp. 645–645, 2020.
- [3] “Global warming of 1.5 °C. An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty,” IPCC-World Meteorol. Organ Geneva, Switzerland, 2018.
- [4] D. Rolnick *et al.*, “Tackling climate change with machine learning,” 2019. *arXiv:1906.05433*.
- [5] A. Ollhoff and J. M. Christensen, *Emissions Gap Report 2018*, United Nations Environment Programme, 2018.
- [6] T. F. Stocker *et al.*, *Climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of IPCC*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
- [7] R. T. Sutton, B. Dong, and J. M. Gregory, “Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations,” *Geophys. Res. Lett.*, vol. 34, 2007.
- [8] G. Keppel-Aleks, P. Wennberg, C. O’Dell, and D. Wunch, “Towards constraints on fossil fuel emissions from total column carbon dioxide,” *Atmos. Chem. Phys.*, vol. 13, pp. 4349–4357, 2013.
- [9] R. Nassar, T. G. Hill, C. A. McLinden, D. Wunch, D. B. Jones, and D. Crisp, “Quantifying CO<sub>2</sub> emissions from individual power plants from space,” *Geophys. Res. Lett.*, vol. 44, pp. 10,045–10,053, 2017.
- [10] J. Hakkarainen, I. Ialongo, and J. Tamminen, “Direct space-based observations of anthropogenic CO<sub>2</sub> emission areas from OCO-2,” *Geophys. Res. Lett.*, vol. 43, pp. 11,400–11,406, 2016.
- [11] A. M. Ahmed, S. K. Ibrahim, and S. Yacout, “Hyperspectral image classification based on logical analysis of data,” in *Proc. IEEE Aerosp. Conf.*, 2019, pp. 1–9.
- [12] M. Zhang, X.-Y. Zhang, R.-X. Liu, and L.-Q. Hu, “A study of the validation of atmospheric CO<sub>2</sub> from satellite hyper spectral remote sensing,” *Adv. Climate Change Res.*, vol. 5, pp. 131–135, 2014.

- [13] N. Acito, M. Diani, and G. Corsini, "CWV-Net: A deep neural network for atmospheric column water vapor retrieval from hyperspectral VNIR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8163–8175, Nov. 2020.
- [14] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, pp. 197–387, 2014.
- [15] M. Buchwitz *et al.*, "The greenhouse gas project of ESAs climate change initiative (GHG-CCI): Overview, achievements and future plans," in *Proc. 36th Int. Symp. Remote Sens. Environ.*, 2015.
- [16] M. Buchwitz *et al.*, "The greenhouse gas climate change initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO<sub>2</sub> and CH<sub>4</sub> global data sets," *Remote Sens. Environ.*, vol. 162, pp. 344–362, 2015.
- [17] S. Plummer, P. Lecomte, and M. Doherty, "The ESA climate change initiative (CCI): A European contribution to the generation of the global climate observing system," *Remote Sens. Environ.*, vol. 203, pp. 2–8, 2017.
- [18] A. Gettelman and R. B. Rood, *Demystifying Climate Models: A Users Guide to Earth System Models*. Basingstoke, U.K.: Springer, 2016.
- [19] D. Rapp, *Assessing Climate Change: Temperatures, Solar Radiation and Heat Balance*. Berlin, Germany: Springer, 2014.
- [20] G. Blanco *et al.*, "Drivers, trends and mitigation," in *Climate Change 2014: Mitigation of Climate Change*. IPCC Working Group III Contribution to AR5, Cambridge Univ. Press, 2014, ch. 5.
- [21] B. Caldecott *et al.*, "Climate risk analysis from space: Remote sensing, machine learning, and the future of measuring climate-related risk," PTI Global Processes, GFZPOF, p. 3, 2018.
- [22] Y. Yoshida *et al.*, "Improvement of the retrieval algorithm for GOSAT SWIR XCO<sub>2</sub> and XCH<sub>4</sub> and their validation using TCCON data," *Atmos. Meas. Techn.*, vol. 6, pp. 1533–1547, 2013, doi: [10.5194/amt-6-1533-2013](https://doi.org/10.5194/amt-6-1533-2013).
- [23] A. Butz *et al.*, "Toward accurate CO<sub>2</sub> and CH<sub>4</sub> observations from GOSAT," *Geophys. Res. Lett.*, vol. 38, 2011, Paper L14812, doi: [10.1029/2011GL047888](https://doi.org/10.1029/2011GL047888).
- [24] B. Dils *et al.*, "The greenhouse gas climate change initiative (GHG-CCI): Comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO<sub>2</sub> and CH<sub>4</sub> retrieval algorithm products with measurements from the TCCON," *Atmos. Meas. Techn.*, vol. 7, no. 6, pp. 1723–1744, 2014.
- [25] P. Bergamaschi *et al.*, "Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2 Evaluation based on inverse model simulations," *J. Geophys. Res., Atmos.*, vol. 112, 2007, Paper D02304, doi: [10.1029/2006JD007268](https://doi.org/10.1029/2006JD007268).
- [26] D. R. Doelling *et al.*, "The intercalibration of geostationary visible imagers using operational hyperspectral SCIAMACHY radiances," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1245–1254, Mar. 2013.
- [27] M. Buchwitz *et al.*, "Copernicus climate change service (C<sub>3</sub>S) global satellite observations of atmospheric carbon dioxide and methane," *Adv. Astronaut. Sci. Technol.*, vol. 1, pp. 57–60, 2018.
- [28] M. Reuter *et al.*, "Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003–2018) for carbon and climate applications," *Atmos. Meas. Techn.*, vol. 13, no. 2, pp. 789–819, 2020.
- [29] F. Kataoka *et al.*, "Calibration, level 1 processing, and radiometric validation for TANSO-FTS TIR on GOSAT," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3490–3500, Jun. 2019.
- [30] L. Guo, L. Lei, Z.-C. Zeng, P. Zou, D. Liu, and B. Zhang, "Evaluation of spatio-temporal variogram models for mapping XCO<sub>2</sub> using satellite observations: A case study in China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 1, pp. 376–385, Jan. 2015.
- [31] T. Tanaka *et al.*, "Two-year comparison of airborne measurements of CO<sub>2</sub> and CH<sub>4</sub> with GOSAT at railroad valley, Nevada," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4367–4375, Aug. 2016.
- [32] T. Matsunaga *et al.*, "Early results of GOSAT-2 level 2 products," *AGUFM*, vol. 2019, pp. A52H-A502, 2019.
- [33] D. Crisp, "Measuring atmospheric carbon dioxide from space with the orbiting carbon observatory-2 (OCO-2)," in *Proc. Earth Observ. Syst. XX*, 2015, Art. no. 960702.
- [34] D. Crisp *et al.*, "The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products," *Atmos. Meas. Techn.*, vol. 10, pp. 59–81, 2017.
- [35] D. Wunch *et al.*, "Comparisons of the orbiting carbon observatory-2 (OCO-2) x CO<sub>2</sub> measurements with TCCON," *Atmos. Meas. Tech.*, vol. 10, pp. 2209–2238, 2017, doi: [10.5194/amt-10-2209-2017](https://doi.org/10.5194/amt-10-2209-2017).
- [36] A. P. Albiñana *et al.*, "Sentinel-5: The new generation European operational atmospheric chemistry mission in polar orbit," in *Proc. Infrared Remote Sens. Instrum. XXV*, 2017, Art. no. 104030P.
- [37] Y. Liu *et al.*, "The tansat mission: Preliminary global observations," *Sci. Bull.*, vol. 63, pp. 1200–1207, 2018.
- [38] A. Poghosyan and A. Golkar, "CubeSat evolution: Analyzing cubesat capabilities for conducting science missions," *Prog. Aerosp. Sci.*, vol. 88, pp. 59–83, 2017.
- [39] R. C. Ssenyunzi *et al.*, "Performance of ERA5 data in retrieving precipitable water vapour over East African tropical region," *Adv. Space Res.*, vol. 65, pp. 1877–1893, 2020.
- [40] P. Collier, G. Conway, and T. Venables, "Climate change and Africa," *Oxford Rev. Econ. Policy*, vol. 24, pp. 337–353, 2008.
- [41] F. Muthoni, "Spatial-temporal trends of rainfall, maximum and minimum temperatures over West Africa," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2960–2973, 2020, doi: [10.1109/JSTARS.2020.2997075](https://doi.org/10.1109/JSTARS.2020.2997075).
- [42] M. Hagenlocher, S. Lang, D. Hölbling, D. Tiede, and S. Kienberger, "Modeling hotspots of climate change in the sahel using object-based regionalization of multidimensional gridded datasets," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 229–234, Jan. 2014.
- [43] M. Boko *et al.*, *Africa Climate Change 2007: Impacts, Adaptation and Vulnerability: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, U.K.: Cambridge Univ. Press, 2018.
- [44] S. Nicholson, "A review of climate dynamics and climate variability in Eastern Africa," in *The Limnology, Climatology and Paleoclimatology of the East African Lakes Book*, 2002, pp. 25–56. [Online]. Available: <http://hdl.handle.net/20.500.12364/1095>
- [45] M. Buchwitz *et al.*, "The GHG-CCI project of ESA's climate change initiative: Data products and application," in *Proc. Living Planet Symp.*, European Space Agency, University of Leicester (Special Publication), ESA SP, SP-740, 2016. [Online]. Available: <https://hdl.handle.net/2381/42555>
- [46] C. D. Rodgers, "Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation," *Rev. Geophys.*, vol. 14, pp. 609–624, 1976.
- [47] M. Buchwitz *et al.*, "Satellite-derived methane hotspot emission estimates using a fast data-driven method," *Atmos. Chem. Phys.*, vol. 17, pp. 5751–5774, 2017.
- [48] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Quart. J. Roy. Meteorol. Soc.*, vol. 146, no. 730, pp. 1999–2049, 2020.
- [49] Z. He *et al.*, "Spatio-temporal mapping of multi-satellite observed column atmospheric CO<sub>2</sub> using precision-weighted Kriging method," *Remote Sens.*, vol. 12, 2020, Art. no. 576.
- [50] T. Wang, J. Shi, Y. Jing, T. Zhao, D. Ji, and C. Xiong, "Combining XCO<sub>2</sub> measurements derived from SCIAMACHY and GOSAT for potentially generating global CO<sub>2</sub> maps with high spatiotemporal resolution," *PloS One*, vol. 9, 2014, Art. no. e105050.
- [51] C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, vol. 2. Singapore: World Scientific, 2000.
- [52] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Oper. Syst. Des. Implementation*, 2016, pp. 265–283.
- [53] A. Botchkarev, "Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology," 2018, [arXiv:1809.03006](https://arxiv.org/abs/1809.03006).
- [54] M. Y. Hu, G. Zhang, C. X. Jiang, and B. E. Patuwo, "A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting," *Decis. Sci.*, vol. 30, pp. 197–216, 1999.
- [55] C. Huntingford, E. S. Jeffers, M. B. Bonsall, H. M. Christensen, T. Lees, and H. Yang, "Machine learning and artificial intelligence to aid climate change research and preparedness," *Environ. Res. Lett.*, vol. 14, 2019, Art. no. 124007.
- [56] Y. Liu *et al.*, "Application of deep convolutional neural networks for detecting extreme weather in climate datasets," 2016, [arXiv:1605.01156](https://arxiv.org/abs/1605.01156).
- [57] B. Knüsel *et al.*, "Applying big data beyond small problems in climate research," *Nature Climate Change*, vol. 9, pp. 196–202, 2019.
- [58] C. Monteleoni, G. A. Schmidt, and S. McQuade, "Climate informatics: Accelerating discovering in climate science with machine learning," *Comput. Sci. Eng.*, vol. 15, pp. 32–40, 2013.
- [59] Z. Han, J. Zhao, H. Leung, K. F. Ma, and W. Wang, "A review of deep learning models for time series prediction," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7833–7848, Mar. 2021.

- [60] X. Yan, C. Liang, Y. Jiang, N. Luo, Z. Zang, and Z. Li, "A deep learning approach to improve the retrieval of temperature and humidity profiles from a ground-based microwave radiometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8427–8437, Dec. 2020.
- [61] J. Friedman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning*, vol. 1 (series in statistics). New York, NY, USA: Springer, 2001.
- [62] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2016.
- [63] J. Feng, X. Wu, J. Chen, X. Zhang, X. Tang, and D. Li, "Joint multilayer spatial-spectral classification of hyperspectral images based on CNN and ConvLSTM," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 588–591.
- [64] S. Kim *et al.*, "Deep-hurricane-tracker: Tracking and forecasting extreme climate events," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2019, pp. 1761–1769.
- [65] S. K. Ibrahim, A. Ahmed, M. A. E. Zeidan, and I. E. Ziedan, "Machine learning methods for spacecraft telemetry mining," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 4, pp. 1816–1827, Aug. 2019.
- [66] S. K. Ibrahim, A. Ahmed, M. A. E. Zeidan, and I. E. Ziedan, "Machine learning techniques for satellite fault diagnosis," *AIN Shams Eng. J.*, vol. 11, pp. 45–56, 2020.
- [67] A. Graves, "Supervised sequence labelling," in *Supervised Sequence Labelling With Recurrent Neural Networks*. Berlin, Germany: Springer, 2012, pp. 5–13.
- [68] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [69] C. Yin, S. Zhang, J. Wang, and N. N. Xiong, "Anomaly detection based on convolutional recurrent autoencoder for IoT time series," *IEEE Trans. Syst., Man, Cybern., Syst.*, Feb. 2020, doi: [10.1109/TSMC.2020.2968516](https://doi.org/10.1109/TSMC.2020.2968516).
- [70] M. Canizo, I. Triguero, A. Conde, and E. Onieva, "Multi-head CNN-RNN for multi-time series anomaly detection: An industrial case study," *Neurocomputing*, vol. 363, pp. 246–260, 2019.
- [71] B. Zhao, X. Li, X. Lu, and Z. Wang, "A CNN-RNN architecture for multi-label weather recognition," *Neurocomputing*, vol. 322, pp. 47–57, 2018.
- [72] E. Tsironi, P. Barros, C. Weber, and S. Wermter, "An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition," *Neurocomputing*, vol. 268, pp. 76–86, 2017.
- [73] Y. Qian, M. Bi, T. Tan, and K. Yu, "Very deep convolutional neural networks for noise robust speech recognition," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 12, pp. 2263–2276, Dec. 2016.
- [74] J. Zhang, Y. Li, T. Li, L. Xun, and C. Shan, "License plate localization in unconstrained scenes using a two-stage CNN-RNN," *IEEE Sensors J.*, vol. 19, no. 13, pp. 5256–5265, Jul. 2019.
- [75] Z. Zhang, Z. Tian, Y. Zhang, M. Zhou, and B. Wang, "u-DeepHand: FMCW radar-based unsupervised hand gesture feature learning using deep convolutional auto-encoder network," *IEEE Sensors J.*, vol. 19, no. 16, pp. 6811–6821, Aug. 2019.
- [76] E. Kalinicheva, D. Ienco, J. Sublime, and M. Trocan, "Unsupervised change detection analysis in satellite image time series using deep learning combined with graph-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1450–1466, 2020, doi: [10.1109/JSTARS.2020.2982631](https://doi.org/10.1109/JSTARS.2020.2982631).
- [77] G. Bonacorso, *Mastering Machine Learning Algorithms: Expert Techniques to Implement Popular Machine Learning Algorithms and Fine-Tune Your Models*. Birmingham, U.K.: Packt Publishing, 2018.
- [78] B. Lazar and M. Williams, "Climate change in western ski areas: Potential changes in the timing of wet avalanches and snow quality for the aspen ski area in the years 2030 and 2100," *Cold Regions Sci. Technol.*, vol. 51, pp. 219–228, 2008.
- [79] A. Amin *et al.*, "Evaluation and analysis of temperature for historical (1996–2015) and projected (2030–2060) climates in Pakistan using SimCLIM climate model: Ensemble application," *Atmos. Res.*, vol. 213, pp. 422–436, 2018.
- [80] P. Gentile, M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis, "Could machine learning break the convection parameterization deadlock?," *Geophys. Res. Lett.*, vol. 45, pp. 5742–5751, 2018.
- [81] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, and N. Carvalhais, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, pp. 195–204, 2019.



**Sara K. Ibrahim** received the B.S. degree in computer and systems electrical engineering and the M.Sc. degree in spacecraft performance analysis and fault diagnosis using telemetry-mining applied on the EGYPSAT-1 telemetry data from Zagazig University, Zagazig, Egypt, in 2014 and 2018, respectively.

She is currently working toward the Ph.D. degree in machine learning applied to climate change as a joint research of Zagazig University and Egyptian Space Agency.

She has worked as a Teaching Assistant with the C.S. Department, Faculty of Engineering, Zagazig University. Her research interests include machine learning, hyperspectral imaging, and satellite telemetry data.



**Ibrahim E. Ziedan** received the M.Sc. and Ph.D. degrees in systems engineering from Aberdeen University, Aberdeen, U.K., in 1964 and 1967, respectively.

He is acting as a Professor with the C.S. Department of Zagazig University, Zagazig, Egypt. His research interests include machine learning, computer architecture, control engineering, and image processing.



**Ayman Ahmed** received the MBA degree in management from L'École Supérieure Libre des Sciences Commerciales Appliquées (ESLSCA), France/Egypt, in 2011, and the Ph.D. degree in engineering from Ain Shams University, Faculty of Engineering, Cairo, in 2013.

He has worked as a System Image Team Leader with the Egyptian Space Agency, Cairo, Egypt. He is experienced in satellite technology, payload imaging camera, satellite testing and verification, satellite system engineering, and artificial intelligence. He is currently a Principal Investigator for NARSSCube satellite series.