frontiers

Predicting climate change using an autoregressive long short-term memory model.

Chin Seokhyun[1], Victoria Lloyd[2*]

[1]Choate Rosemary Hall, Wallingford, CT, USA

[2]Department of Physics, Stony Brook University, Stony Brook NY, USA

Keywords: Climate forecasting, Machine learning, time series, Recurrent Neural Network, Model development, _.

메모 포함[GU1]: complete

Abstract

메모 포함[GU2]: Include as many keywords as you can in the first two sentences

Climate change is an issue that has and will affect humanity throughout the time that we are on Earth. Due to this fact, scientists have created numerous methods of modelling the climate, with mathematical models being used most widely. However, due to some limitations that these models hold, researchers have turned to utilize machine learning models to predict the future climate. This is because AI has a reputation for being able to handle complex data and recognize patterns that humans previously could not have. This study aims to create a baseline machine learning model that utilizes an Autoregressive Recurrent Neural network with a Long Short term memory implementation for the purpose of predicting climate. This type of model architecture has been utilized for other applications that relate to time series data, yet it has yet to be attempted to be used in the context of climate predictions. The data that was utilized was retrieved from the ensemble mean version of the ERA5 dataset. The model created from this study was able to predict the general trends of Earth(i.e, the poles are cold and the equator region is warm) for both when predicting the climate and when it was predicting weather. When predicting the climate, the model was able to have fair accuracy for a long period of time, with the ability to predict seasonal patterns. This feat is one that other researchers were not able to do with the complex reanalysis data that this study has utilized. This work demonstrated that this type of model can be utilized in a climate forecasting approach as a viable alternative to mathematical models and can be utilized to supplement current work that is mostly successful in short term predictions.

1 Introduction

From 1983 to 2012, the Northern Hemisphere experienced likely the warmest 30-year period in the last 1400 years. Furthermore, the globally average combined land and ocean surface temperature data predicts an approximate annual warming of 0.85 °C (Pachauri et al., 2015). Climate change itself is not new, every inhabited region on Earth is currently experiencing climate changes that have not been seen in a very long time. (UN REPORT, 2022; Lerner, 2023) The general scientific consensus is that human activity has significantly contributed to the acceleration of climate change beyond what would occur naturally. (Lerner, 2023)

메모 포함[GU3]: In text citation format

Indeed, extensive research has been conducted to predict the extent of climate change and estimate the magnitude of the challenges that humanity will ultimately face. Climate scientists relied on more traditional methods, such as divination, pattern recognition, and various other means. Developments have been made in statistical forecasting techniques, both multi and univariate, which fall short of numerical prediction models. (Stern and Easterling, 1999) Atmospheric general circulation models (AGCM), a type of numerical model, consist of a system of equations that describe the large-scale atmospheric balances of momentum, heat, and moisture, with schemes that approximate small-scale processes such as cloud formation, precipitation, and heat exchange with the sea surface and land. (Hurrell, 2003) AGCMs have the potential to make predictions multiple months and years into the future, and many different models emerged as time passed by, such as that of National Aeronautics and Space Administration and the University of California, Los Angeles. (Stern and Easterling, 1999; Edwards, 2010) However, there have been some issues due to decreasing the predicting ability of the model outside of the tropic region, (Kumar and Hoerling, 1995) In addition, added spatial complexity for more accurate predictions increased the computational complexity of GCMs. (Collins et al., 2012)

With the emergence of artificial intelligence (AI) and machine learning (ML), many attempts to explore the possibility of utilizing AI in climate and weather prediction have been done. Studies that utilize Convolutional Neural Networks to create weather predictions on specific regions were reported. (Scher and Messori, 2019; Weyn et al., 2019) Regression models and the random forest algorithm were made to create weather predictions as well. (Herman and Schumacher, 2018; Mansfield et al., 2020) To generate climate predictions, models that were utilized for weather predictions were regressively trained, and researchers were able to get stable climate predictions from relatively simple data that are pulled from GCMs. (Scher, 2018) However, when training on more complex data, models have struggled in climate predictions, not predicting seasonal patterns correctly and predicting unrealistic patterns. (Weyn et al., 2019)

Autoregressive long short-term memory (LSTM) networks have an ability to recognize behavioral patterns of time series data and utilize that to make predictions. They have been used in predicting air pollution (Kulkarni et al., 2018) and rainfall (Razak et al., 2016).

메모 포함[AC4]: expand on why this is new in an interesting way that we should care about

메모 포함[AC5R4]: Why is AI important in climate models? AI algorithms are known for being able to work with data in a much more flexible way that mathematical models because they can use so many different algorithms that can handle complex patterns that humans might not be able to know how to encorporate that into the mathematical models(be on the lookout for similar wordings)

This study aims to create a baseline model for the characterization of the long-term temperature of the Earth using LSTM networks.

2 Method

This model was trained to input the temperature at 2m height(t2m) f every location on Earth for one year with an interval of 14 days (about 2 weeks) and output the state of Earth of the next half year with the same intervals

2.1 Model Architecture

This ML model is a deep autoregressive neural network that utilizes a Convolutional LSTM layer, which combines the properties of a convolutional layer alongside an LSTM layer. (https://arxiv.org/pdf/1506.04214v1.pdf) This allows for an increased ability to handle spatial-temporal data. Convolutional LSTM layers have often been utilized for predicting time-series data, which will be helpful in this specific application.[16] An Autoregressive model takes in the first few input labels as what is known as "warmup," and then utilizes the output of the previous LSTM iteration to train the next step of the model.

2.2 Model setup

The autoregressive model trains a single Convolutional LSTM layer regressively and runs the output of that through a series of dense layers that produces an output. Each dense layer has a LeakyRelU activation function, and the Convolutional LSTM layer has a dropout of 0.2. The training loop had maximum epochs of 50 with an EarlyStopping callback and a ReduceLRonPlateau callback. The

RMSprop optimizer with a learning rate of 0.001 and a decay rate of 0.9 was utilized, and the binary cross entropy loss function was used as well. The model was built using the open-source Keras Library for python (Collet, 2015) with Google's tensorflow backend. (Schneider and Xhafa, 2022)

2.3 Data acquisition and processing

The ensemble mean version of the ERA5 reanalysis dataset was used, which has an advantage that data are available for each grid point at each time step and that the data are consistent over the entire data window instead of observations for training. (Düben and Bauer, 2018) Temperature at two-meter height of the 1st and 15th of each month are considered during the time period 2002- 2022, leading to five hundred four-time steps overall. This temperature data is normalized by adding 80 to the temperature in Celsius and dividing that value by 160 so that the model's training can be done in a swifter manner. The latitude and longitude were recorded with a 0.5◦ resolution, resulting in a snapshot of 361*720 = 65160 grid points per each time step. Of the five hundred four-time steps, four hundred were used for training, forty-four as validation, and sixty as testing.

The data were converted from numerical tuples of (latitude, longitude, temperature, time) to an image for training with the two-meter height data being the color value at each latitude, longitude pair. Latitude-longitude of (90,-180) was set to (0,0) of the image, which is the top left corner.

2.4 Metrics

The model's performance is calculated using two values; the root mean squared error (RMSE) and the Mean absolute error (MAE), which are widely used metrics for calculating the accuracy of models.[20] These metrics will be calculated with the data un-normalized and back to its original scale.

The RMSE and MAE value are calculated as shown. (Chai and Draxler, 2014)

i. $RMSE = \square\{1\}\{n\}\sum\_\{i = 1\}\hat{}n\sqrt{\{e\_i\hat{}2\}}$ where n is the total number of samples and e_i denotes the error at the ith sample.

ii. $MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i|$ where n is the total number of samples and e_i denotes the error at the ith sample

2.5 Testing conditions

After the model is trained, the model's performance will be evaluated in three scenarios. The first is when images are generated on true images. Then, a scenario in which the model predicts future images with only the first input image being real and the model autoregressively feeding the output back as a part of the input will be considered. Finally, the scenario will be when climate predictions will be made twenty years into the future using the regressive methodology utilized in the second evaluation method will be conducted.

Mean temperature over time, the difference between the real data and generated data, and the RMSE and MAE values are considered. Afterward, results of model predictions in specific significant regions of the Earth-Arctic, Antarctic, East Asia, Europe, North America, North Africa, Northwest Asia, Oceania, South America, South Asia, and South Africa. These regions were set from Mansfield et al.'s study (2020), yet Antarctica was added due to its prominence in the images of earth.

3 Results

3.1 Analysis of the output generated solely on real images.

The model showed fair performance when using this method with RMSE and MAE values of 4.642 and 3.296, respectively. Figure 2 shows the RMSE and MAE values over each time step as as a line graph (a) and the average temperature change of the Earth over time for both the predicted values and the real values (b). It can be observed that there is a certain seasonal pattern, although the pattern's shape doesn't perfectly align with the real pattern. In addition, the predicted values have a generally higher prediction of temperature, as shown by the higher overall mean value.

Figure 3-a shows a model prediction of the temperature on December 15th, 2022, and the temperature observed on that date. The predicted image is much blurrier than the observed image. While the right image shows much more precise delineations in the color change, the left image has the colors more or less equal for regions where the color should turn out to be different. However, there are still differences that clearly divide the continents and the oceanic regions that can be observed in the left image, although the temperature difference is not well described in that image. In addition, the temperatures of regions closer to the equator are predicted to be hotter (more yellow), while the regions near the poles are predicted to be colder.

Figure 4-a shows a visualization of all the RMSE/MAE values over every predicted timestep in each major region of Earth. MAE values have less deviation than the RMSE values, and they are also much lower. The model is best at predicting the temperature in the South Africa region, with the lowest RMSE and MAE of 2.222 and 1.822, respectively. The model is the least competent in predicting temperature in the Northwest Asia region, with a high RMSE and MAE of 6.829 and 5.657. Figure 4-b shows the change in mean temperature of each region of the predicted versus the actual data. The Arctic and Antarctic region can be predicted with ease with the model, while regions such as Oceania/Southeast Asia, South America, North Africa and South Africa cannot be predicted so well using the model. The remaining regions – Northwest Asia, South Asia, East Asia, North America and Europe-seem to have relatively good accuracy, with very little deviation between the predicted mean temperature and the real mean temperature and little change between the patterns of increase and decrease.

3.2 Analysis of the output generated by feeding in generated images.

The model still showed fair performance with this methodology, with RMSE and MAE values of 5.487 and 3.982, respectively. However, both the MAE and RMSE values increased when compared to the results of 3.1: an RMSE and MAE of 4.642 and 3.296. Figure 2 shows the RMSE and MAE values over each time step as a line graph (c) and the average temperature change of the earth over time for both the predicted values and the real values(d). The temperature pattern seems to show a slight increase in temperature over time that is larger than the actual data. In addition, the temperature predicted is also higher than the predictions made solely on real images.

Figure 3 (b) shows the model predictions of the temperature on December 15th, 2022, and the temperature observed on that date. The left image shows the image generated when data the output is generated on generated images, the middle image shows the image generated when the output is generated on real images, and the right image shows the observed image at the time. While the images do look quite similar, the region around the equator on the left image is the most yellow-meaning the hottest, while the arctic region is much less dark- meaning the coldest. The leftmost image is also the blurriest, although the major delineations (i.e., the continents) are still somewhat recognizable.

Figure 5 shows a visualization of all the RMSE/MAE values over every predicted timestep and the change in mean temperature of each region of the predicted versus the actual data in each major region of Earth. South Africa remains to be the region that the model predicts the best, with the lowest RMSE and MAE of 2.998 and 2.510, respectively. The least competent region of the model remained to be Northwest Asia, with the highest RMSE of 8.304 and 6.957. It can be shown from figure 5-b the model is not competent in predicting the mean temperature in Europe, South Asia, Oceania/Southeast Asia, South America, North Africa and South Africa

3.3 Model prediction of temperature

The model's output of mean temperature over the next twenty years was shown in Fig 6 with the projected mean temperature change by the model (a) and the predicted temperature on 1 year into the future, 5 years into the future, 10 years into the future and 20 years into the future (b). As shown in figure 6-b, the model progressively gets worse in its predictions, and more and more yellow as time goes on. The Arctic region first turns yellow, yet the Antarctic region doesn't follow that trend and remains cold. Figure 6-a shows that the model projects that the temperature will reach an average of roughly 18 degrees Celsius in the next twenty years, which is roughly an 11-degree raise in temperature. However, seasonal patterns of increase and decrease in temperature can still be observed in the predictions.

4 Discussion

Using an autoregressive LSTM network, we developed a predictive model for climate especially long-term temperature, with fair performance. Although future climate predictions are vital for addressing the challenges posed by global warming, there are only a few models built with the purpose of detecting long-term climate change. Future climate predictions can play a crucial role in decision-making and policy-planning, prevention and response to natural disasters, and resource management and environmental protection. These predictive model can be helpful in building to a sustainable future.

Global Climate Models (GCM), as a complex computer-based model, incorporate various components of the earth system, which provide insights into a wide range of climate variables, such as temperature, precipitation, wind patterns, and atmospheric circulation and risk of climate change. Due to their limitations and uncertainties, exploring more detailed relationships between emissions and multiregional climate responses still requires the application of GCMs that allow the behaviour of the climate to be simulated under various conditions on decadal to multi-centennial timescales. (Bitz and Polvani, 2012; Nowack et al., 2017; Hartmann et al., 2019) However, modeling climate at increasingly high spatial resolutions has significantly increased computational complexity. Therefore, numerous studies have been conducted to develop AI models to solve this issue and supplement these models.

A study predicting global patterns of long-term climate change from short-term simulations evaluated the performance using Ridge and Gaussian Process Regression (GPR) at a grid-cell level. Both predict broad features like enhanced warming over the Northern Hemisphere like pattern scaling. (Mansfield et al., 2020) The error of the temperature in this study remains in between one and two degrees, which is significantly lower than this study. However, their study is mainly focused on improving GCMs, while this study seeks to potentially replace GCMs. As a corollary to this fact, their study is pulled from the outputs of GCMs, while this study's outputs are from reanalysis data. In addition, the errors in that study were calculated on a grid-cell level, which can be misleading due to the fact that they penalize patterns that as broad features are predicted correctly but displaced marginally on the spatial grid.(Rougier, 2016) This crucial predicted difference in data could be the main cause of the difference in errors between this study and Mansfield et.al.'s study. The usage of different ML algorithms (Autoregressive LSTM vs Gaussian/ Ridge regression) could contribute to this difference as well.

A study about global mean surface temperature projections by employing advanced ensemble methods and using past information was reported. (Stobach and Bel, 2020) This study doesn't utilize RMSE as a metric that is displayed or discussed. The study does predict the change in global mean temperature, which is projected to be maximum of 4 degrees increase, although it is dependent on the type of algorithm utilized. This is a large difference from the projected increase of temperature of this study, which could be due to the difference in ML algorithms (Autoregressive LSTM vs Esemble models) and difference in data, like the reason that this study's results deviate from Mansfield et.al's study.

For forecasting large-scale spatial patterns of precipitation across the western United States, training on thousands of seasons of climate model simulations and testing on the historical observational period (1980-2020) are done. It could compete with or outperforming existing dynamical models from the North American Multi Model Ensemble. Gibson et al., 2021) This study utilizes numerous ML models such as LSTM, NN, XGBoost and Random Forest algorithm to predict the precipitation of North America. Their LSTM model was a consistent performer, with an accuracy score of roughly 0.5 for all categories. This prediction is higher than the other models tested in the study. Our model have quite modest errors as this one does. To circumvent this issue, here we explore the feasibility of training various machine learning approaches on a large climate model ensemble, providing a long training set with physically consistent model realizations.

Previous researchers also utilized Convolutional Neural Networks for predicting weather. (Scher and Messori, 2019, Weyn et al, 2019). In Weyn's study, the error of the LSTM model begins at 10 and increases as time moves on, which is similar to the results in this study. The error in their study, increases at a more rapid rate, increasing every 12 hours opposed to the longer rate of ours. This is mainly due to the fact that Weyn's study takes data every 6 hours, and we took it every 14 days, yet our prediction length with relatively low RMSE- 60-time steps- is much longer than that of their study. Scher's study had the limitation of poor seasonal cycle and unrealistic predictions when trained for long periods of time. In this study, although still producing unrealistic predictions, succeed in predicting seasonal cycles successfully for up to thirty years- an improvement that is made from Scher's study. Our study's RMSE values remain equal or lower than Scher's results, which is also an improvement that is made. This is because of the different model architecture that was utilized (I.e autoregressive LSTM, Convolutional Neural Network), since the data that was used are both the ERA5 reanalysis.

This study has shown that the autoregressive LSTM architecture succeed in improving the long-term climate predictions of ML models. The largest issue of the model currently is that it seems to be underfitting. When observing the data generated by the model, it is learning some broad patterns like how the poles are cold, the rest is warm, and mountain ranges and seasons. The issue is that the model is not learning specific enough information to be useful, so the model isn't fitting to the data well enough. This is mainly due to the limitation of computing power. To resolve this issue, increasing the complexity of the model such as denser layers, more filters and increasing the amount of training data should be considered.

This paper's model should serve solely as a baseline model. With many improvements, the model should reach its full potential. The ability of this architecture that makes it successful compared to other architectures (I.e., possibility of predicting global patterns over a long period of time) is still shown in this study, which makes it valuable to pursue further research in this area.

5 Conclusion

In this study, the potential of utilizing Autoregressive LSTM models for climate predictions were considered. The model outputted fair predictions of the global temperature at 2 meters for a long period of time, with it succeeding in predicting the global mean temperature progression of multiple regions. This feat is one that has not been achieved in previous work that utilized Neural Networks, making this study useful in that sense. The model developed here should be used as a baseline for further development with increases in computing power and data availability.

- List results of my work(don't give numbers) and why my work is important.,

메모 포함[AC6]: dont need- do want to wrap up in discussion!

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

Chin SH and Llyod V conceptualized the study. Chin SH performed the simulations, model analysis, and wrote the first draft of the text. Lloyd V contributed to the text and figure revisions and the final manuscript.

8 Acknowledgments

I would like to acknowledge the Inspirit X AI mentorship program for facilitating this research project.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Zheng, X., et al. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.

Bitz, C. M. and Polvani, L. M. (2012). Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model. Geophysical Research Letters, 39(20), L20705. doi:10.1029/2012GL053393

Chai, T., and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscientific Model Development, 7(3), 1247–1250. doi:10.5194/gmd-7-1247-2014

Chollet, F. (2015). Keras: Deep learning library for theano and tensorflow. Keras.io, 7(8), T1, Retrieved from https://keras. io/

Collins, M. J., Chandler, R. B., Cox, P. T., Huthnance, J. M., Rougier, J., and Stephenson, D. B. (2012). Quantifying future climate change. Nature Climate Change, 2(6), 403–409. doi:10.1038/nclimate1414

Düben, P., and Bauer, P. (2018). Challenges and design choices for global weather and climate models based on machine learning. Geoscientific Model Development, 11(10), 3999–4009. doi:10.5194/gmd-11-3999-2018

Edwards, P. N. (2010). History of climate modeling. Wiley Interdisciplinary Reviews: Climate Change, 2(1), 128–139. doi:10.1002/wcc.95

Gibson, P. B., Chapman, W. E., Altinok, A., Monache, L. D., DeFlorio, M. J., and Waliser, D. E. (2021). Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. Communications Earth and Environment, 2, 159. doi:10.1038/s43247-021-00225-4

Hartmann, D. L., Blossey, P. N. and Dygert, B. D. (2019). Convection and climate: what have we learned from simple models and simplified settings? Current Climate Change Reports, 5, 196–206. doi:10.1007/s40641-019-00136-9

Herman, G. S., and Schumacher, R. S. (2018). Money Doesn't Grow on Trees, but Forecasts Do: Forecasting Extreme Precipitation with Random Forests. Monthly Weather Review, 146(5), 1571–1600. doi:10.1175/mwr-d-17-0250.1

Hewamalage, H., Bergmeir, C., and Bandara, K. (2021). Recurrent Neural Networks for Time Series Forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388–427. doi:10.1016/j.ijforecast.2020.06.008

Hurrell, J. W. (2003). CLIMATE VARIABILITY | North Atlantic and Arctic Oscillation. In Encyclopedia of Atmospheric Sciences.

Kaur, J., Parmar, K. S., and Singh, S. (2023). Autoregressive models in environmental forecasting time series: a theoretical and application review. Environmental Science and Pollution Research, 30(8), 19617–19641. doi:10.1007/s11356-023-25148-9

Kulkarni, G., Muley, A., Deshmukh, N., and Bhalchandra, P. (2018). Autoregressive integrated moving average time series model for forecasting air pollution in Nanded city, Maharashtra, India. Modeling Earth Systems and Environment, 4(4), 1435–1444. doi: 10.1007/s40808-018-0493-2

Kumar, A., and Hoerling, M. P. (1995). Prospects and Limitations of Seasonal Atmospheric GCM Predictions. Bulletin of the American Meteorological Society, 76(3), 335–345. Retrieved from http://www.jstor.org/stable/26232354

Lerner, K. L. (2023). Climate change. In Gale Science Online Collection. Gale. Retrieved from https://link.gale.com/apps/doc/KODPIQ908201214/SCIC?u=wall96493&sid=bookmark-SCIC&xid=97316fbe

Mansfield, L., Nowack, P., Kasoar, M., Everitt, R. G., Collins, W. J., and Voulgarakis, A. (2020). Predicting global patterns of long-term climate change from short-term simulations using machine learning. Npj Climate and Atmospheric Science, 3(1). doi:10.1038/s41612-020-00148-5

Nowack, P. J., Braesicke, P., Luke Abraham, N. and Pyle, J. A. (2017). On the role of ozone feedback in the ENSO amplitude response under global warming. Geophysical Research Letters, 44(8), 3858–3866. doi:10.1002/2016GL072418

Pachauri, R. K., Allen, M. R., Barros, V. R., Broom, J., Cramer, W., Christ, R., et al. (2015). Climate Change 2014: Synthesis Report. The Intergovernmental Panel on Climate Change. Retrieved from https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf

Razak, N. A., Aris, A. Z., Ramli, M., Looi, L. J., and Juahir, H. (2016). Temporal flood incidence forecasting for Segamat River (Malaysia) using autoregressive integrated moving average modelling. Journal of Flood Risk Management, 11, S794–S804. doi:10.1111/jfr3.12258

Rougier, J. (2016). Ensemble averaging and mean squared error. The Journal of Climate, 29, 8865–8870.

Scher, S. (2018). Toward Data-Driven Weather and Climate Forecasting: Approximating a Simple General Circulation Model with Deep Learning. Geophysical Research Letters. doi:10.1029/2018GL080704

Scher, S., and Messori, G. (2019). Weather and climate forecasting with neural networks: using general circulation models (GCMs) with different complexity as a study ground. Geoscientific Model Development, 12(7), 2797–2809. doi:10.5194/gmd-12-2797-2019

Schneider, P., and Xhafa, F. (2022). Chapter 3–Anomaly detection: Concepts and methods. In Elsevier eBooks, 49–66. doi:10.1016/b978-0-12-823818-9.00013-4

Stern, P. C., and Easterling, W. E. (1999). Making Climate Forecasts Matter. National Academies Press eBooks. doi:10.17226/6370

Strobach, E., and Bel, G. (2020). Learning algorithms allow for improved reliability and accuracy of global mean surface temperature projections. National Communications, 11, 451. doi:10.1038/s41467-020-14342-9

UN report : Climate Change is irrevocalble. (2022, January-February). Discover, 43(1), 38. Retrieved from https://link.gale.com/apps/doc/A693733918/SCIC?u=wall96493&sid=bookmark-SCIC&xid=f30c64a2

Weyn, J. A., Durran, D. R., and Caruana, R. (2019). Can Machines Learn to Predict Weather? Using Deep Learning to Predict Gridded 500-hPa Geopotential Height From Historical Weather Data. Journal of Advances in Modeling Earth Systems, 11(8), 2680–2693. doi:10.1029/2019ms001705

9 Data Availability Statement

Pulled all data from the ERA5 reanalysis ensemble mean data. Refer to this link: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

Code availability

Code generating figures and processed data will be available upon request

10 Figure legend

Figure 1: Architecture of auto-regressive neural network models. The first 24 inputs are utilized as “warmup,” and the output of the final input is continuously fed through the model to generate the predictions.

Figure 2: RMSE, MAE, and mean temperature change over the entire earth. the RMSE and MAE in each time step solely on real images (a) the mean temperature change over the entire earth that the model predicted solely on real images (b). RMSE and MAE in each time step by feeding in generated images (c) the mean temperature change over the entire earth that the model predicted by feeding in generated images (d)

Figure 3: Image of predicted temperature on December 15th, 2022. (a) Left image is the image generated by the model, and the right image is the image that is derived from real data solely on real images. (b) Left image generated by the model when fed autoregressively; middle image generated by model when trained on all real images when fed autoregressively; right image derived from data when fed autoregressively

Figure 4. RMSE, MAE (a) and change in mean temperature (b) over each region trained solely on real images.

Figure 5. RMSE, MAE(a), and change in mean temperature (b) over each region trained solely by feeding in generated images.

Figure 6. Changes in average temperature over time (a) and the predicted temperature on 1 year into the future, 5 years into the future, 10 years into the future and 20 years into the future (b).

Figures

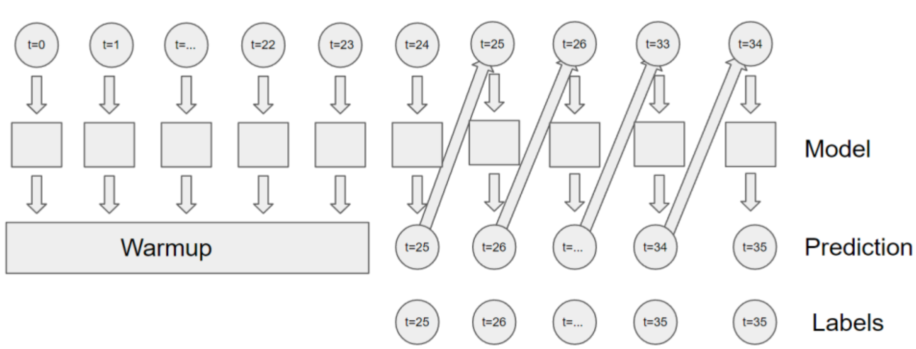


Figure 1: Architecture of auto-regressive neural network models. The first 24 inputs are utilized as “warmup,” and the output of the final input is continuously fed through the model to generate the predictions.

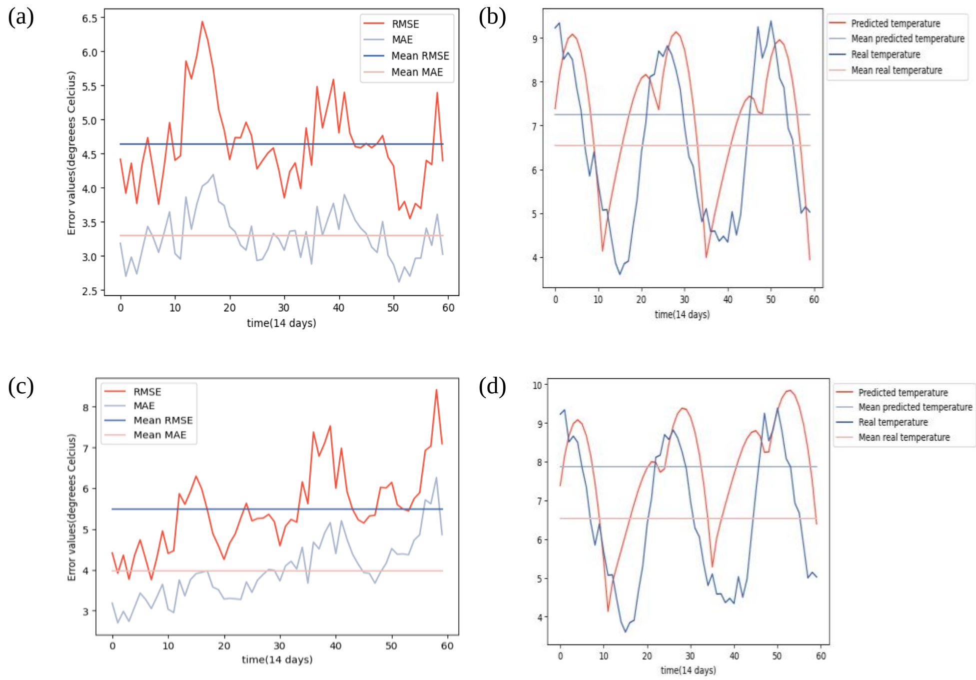


Figure 2: RMSE, MAE, and mean temperature change over the entire earth. the RMSE and MAE in each time step solely on real images (a) the mean temperature change over the entire earth that the model predicted solely on real images (b). RMSE and MAE in each time step by feeding in generated images (c) the mean temperature change over the entire earth that the model predicted by feeding in generated images (d)

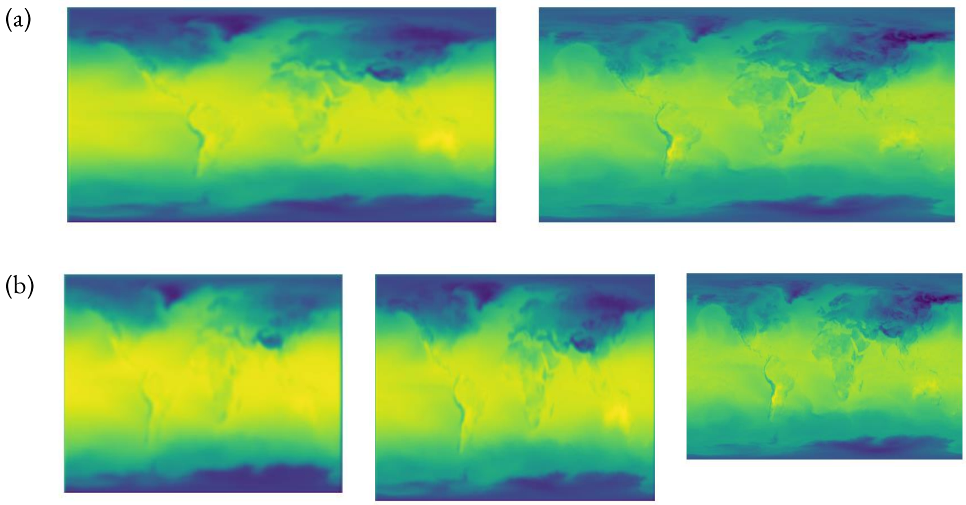


Figure 3. Image of predicted temperature on December 15th, 2022. (a) Left image is the image generated by the model, and the right image is the image that is derived from real data solely on real images. (b) Left image generated by the model when fed autoregressively; middle image generated by model when trained on all real images when fed autoregressively; right image derived from data when fed autoregressively

(a)

(b)

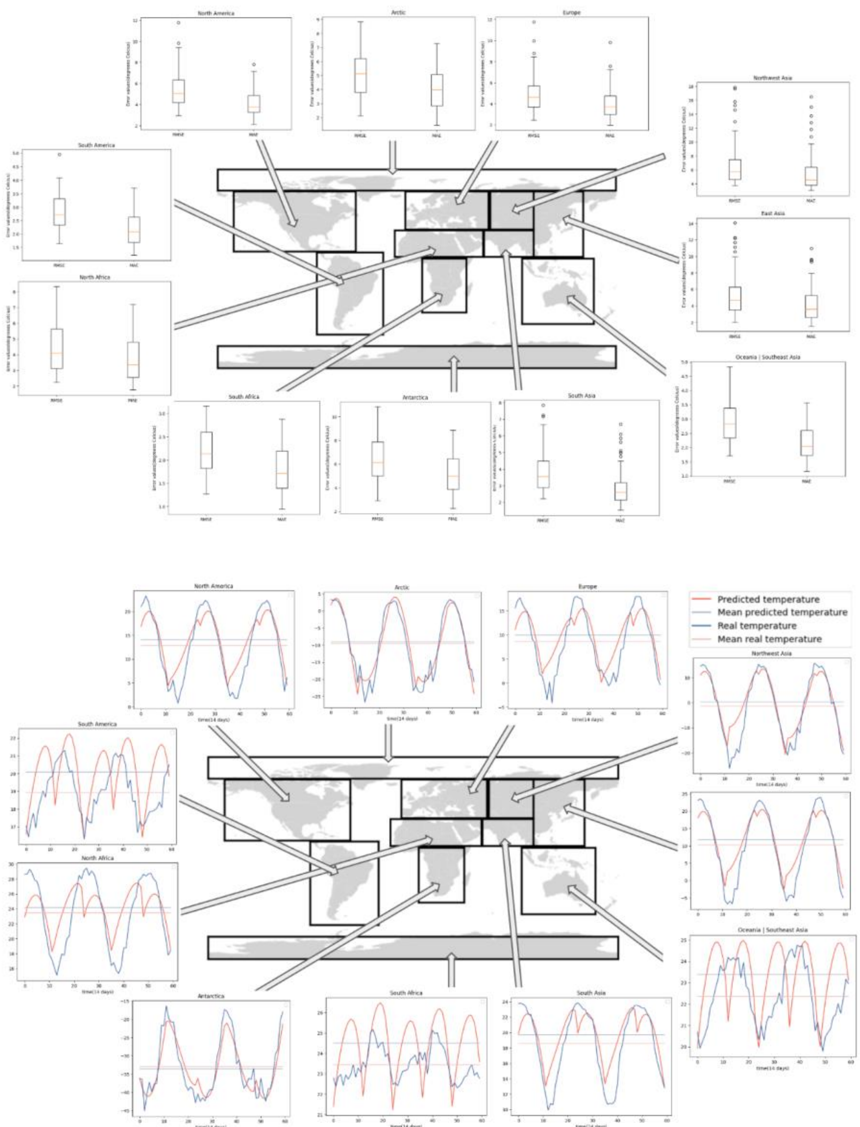
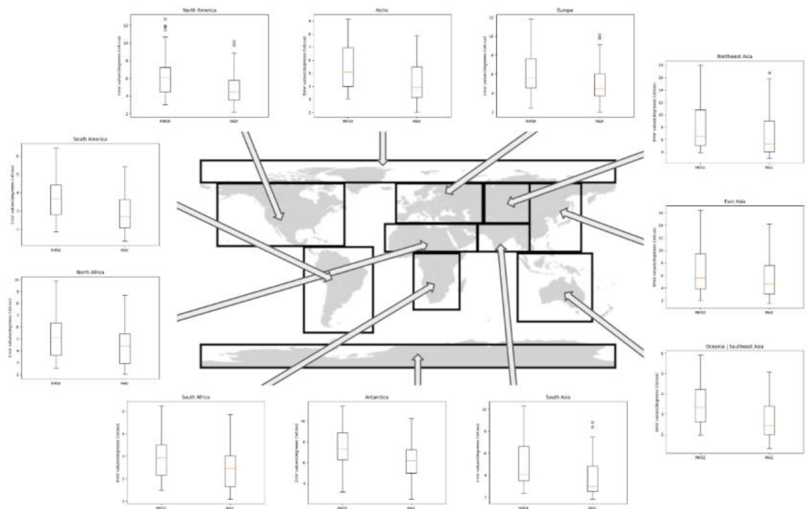


Figure 4. RMSE, MAE (a) and change in mean temperature (b) over each region trained solely on real images.

(a)



(b)

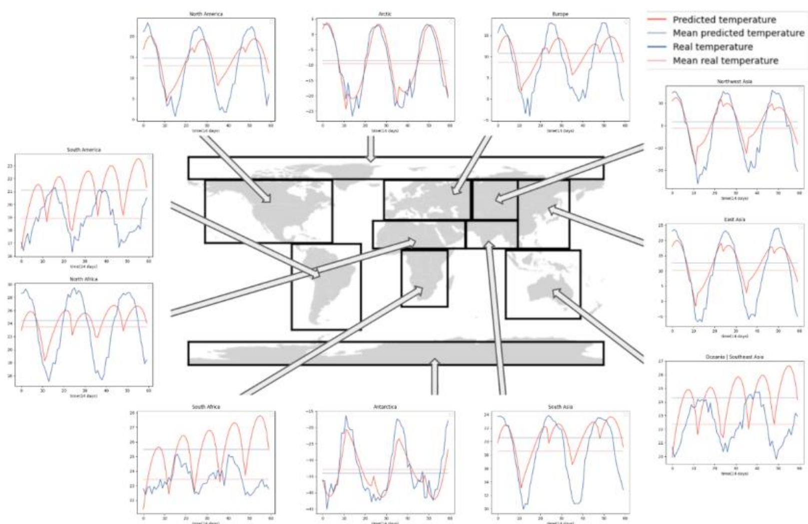
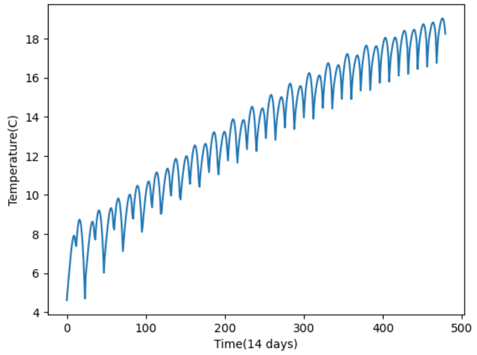
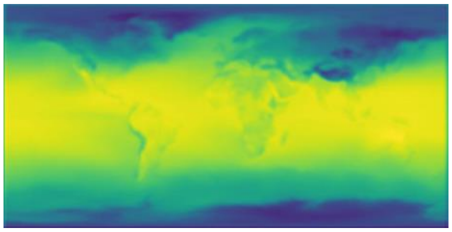


Figure 5. RMSE, MAE(a), and change in mean temperature (b) over each region trained solely by feeding in generated images.

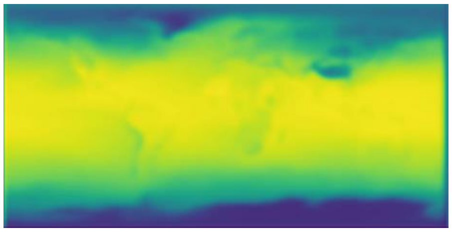
(a)



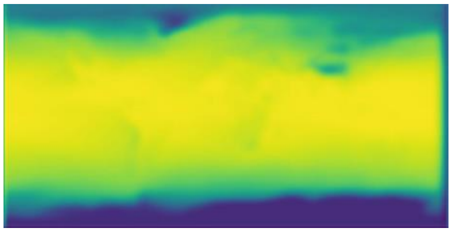
(b)



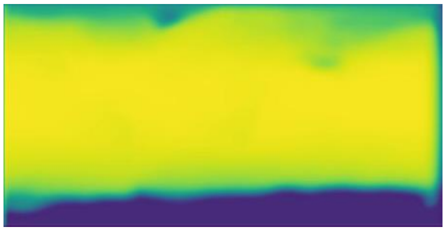
1 year into the future



5 years into the future



10 years into the future



20 years into the future

Figure 6. Changes in average temperature over time (a) and the predicted temperature on 1 year into the future, 5 years into the future, 10 years into the future and 20 years into the future (b).

Permission to reuse and Copyright

Permission must be obtained for use of copyrighted material from other sources (including the web). Please note that it is compulsory to follow figure instructions.