Science Bulletin 66 (2021) 1358–1366

Contents lists available at ScienceDirect

Science Bulletin

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

Article

Unified deep learning model for El Niño/Southern Oscillation forecasts by incorporating seasonality in climate data

Yoo-Geun Ham [a,*], Jeong-Hwan Kim [a], Eun-Sol Kim [b], Kyoung-Woon On [b]

[a] Department of Oceanography, Chonnam National University, Gwangju 61186, South Korea
[b] Kakao Brain, Bundang-gu, Seongnam-si, Gyeonggi-do 13494, South Korea

ARTICLE INFO

Article history:
Received 1 September 2020
Received in revised form 5 February 2021
Accepted 8 February 2021
Available online 13 March 2021

Keywords:
Deep learning
ENSO forecasts
Seasonality of the ENSO

ABSTRACT

Although deep learning has achieved a milestone in forecasting the El Niño-Southern Oscillation (ENSO), the current models are insufficient to simulate diverse characteristics of the ENSO, which depends on the calendar season. Consequently, a model was generated for specific seasons which indicates these models did not consider physical constraints between different target seasons and forecast lead times, thereby leading to arbitrary fluctuations in the predicted time series. To overcome this problem and account for ENSO seasonality, we developed an all-season convolutional neural network (A_CNN) model. The correlation skill of the ENSO index was particularly improved for forecasts of the boreal spring, which is the most challenging season to predict. Moreover, activation map values indicated a clear time evolution with increasing forecast lead time. The study findings reveal the comprehensive role of various climate precursors of ENSO events that act differently over time, thus indicating the potential of the A_CNN model as a diagnostic tool.

1. Introduction

The behavior of El Niño-Southern Oscillation (ENSO) events can be distinguished based on calendar months [1–3]; an ENSO event typically begins to develop during boreal spring, rapidly grows during summer and autumn, and exhibits maximum amplitude in winter. This "seasonal phase-locking" of ENSO is caused by seasonal variations in the atmospheric response to a certain sea surface temperature (SST) [1,4–7] and oceans' basic states [2,8,9].

Seasonal ENSO phase-locking is a primary cause of seasonality in ENSO forecasting accuracy. The forecasting accuracy, a product of various dynamical and statistical models, tends to remain high during the boreal winter when ENSO events are in the mature phase. Contrastingly, it decreases rapidly during the boreal spring when the ENSO event starts to grow [10–12]. This unique feature of ENSO forecasts is often referred to as the "spring predictability barrier" [13,14].

In addition, ENSO forecasting accuracy involves seasonality because the significance of the ENSO precursor varies strongly with respect to the seasons. For example, the Indian Ocean Dipole (IOD), a well-studied precursor of the ENSO [15], peaks during boreal fall [16]; however, it is negligible in other seasons. Therefore, the IOD can be an optimal predictor only for forecasts initialized in a boreal fall [17]. Other well-known ENSO predictors (e.g., Atlantic Niño [18,19], North Tropical Atlantic SST [20], and Western Hemispheric Warm Pool [21]) additionally exhibit strong seasonality in their amplitudes; therefore, they are effective predictors only for a particular season. In other words, an ENSO precursor for a specific season would not be optimal for the remaining seasons.

As a reflection of predictor-specific seasonality, the deep learning-based ENSO forecast model proposed by Ham et al. [22] (hereafter referred to as H19) was formulated independently for each target season and forecast lead months. To produce 2-year ENSO forecasts for all target seasons, the number of H19 models required is 276 (that is, 23 (lead months) × 12 (target seasons)). Even though the follow-up studies developed deep learning models using the Convolutional Long Short-Term Memory Network (ConvLSTM) [23], LSTM [24], and the variant of the Convolutional Neural Network (CNN) [25], these models were also formulated separately according to the input season, which implies that the current deep learning models were not successful in simulating seasonally dependent diverse characteristics of the ENSO.

In addition, as the forecast results are independently generated through separate H19 models for each lead month, the predicted time series becomes less consistent over time. For instance, the predicted ENSO index of the H19 model contains an arbitrary month-to-month variation that is not perceived in the observation.

* Corresponding author.
E-mail address: ygham@jnu.ac.kr (Y.-G. Ham).

https://doi.org/10.1016/j.scib.2021.03.009
2095-9273/

The most intuitive solution to this problem is to formulate the single H19 model for all target seasons and forecast lead months. However, training all-month samples as a single set of data considerably degrades forecast accuracy of the ENSO events (Fig. S1a online).

The skill degradation in the H19 model trained by all–month samples might be due to the lack of proper consideration of the varying characteristics of different seasons in climate samples. The seasonally varying climatological states significantly modulate the strength of several air–sea coupled feedbacks (such as Wind-Evaporation-SST feedback, Cloud-Radiation-SST feedback, and Bjerknes feedback) [26,27]. Therefore, the time-evolution of the oceanic/atmospheric variables greatly differs from season to season even if the state vectors at the initial stage are almost similar. Hence, without accurate recognition of seasonality, the characteristics of the ENSO events in all-month climate samples are excessively diverse and cannot be expressed in a single set of model parameters. Thus, it is necessary to modify the H19 models to accurately recognize the differences in the ENSO characteristics for different seasons.

2. Methods

2.1. Convolutional neural network (CNN) model for all seasons

As shown in Fig. 1, the CNN for all seasons (denoted as All-Season CNN (A_CNN) hereafter) for ENSO forecasts is composed of an input layer, three convolutional layers, two pooling layers, two fully connected layers, and an output layer. The A_CNN takes gridded data of the SST, and the oceanic heat content (HC, defined as the vertically averaged oceanic temperature from the surface to 300 m) anomalies for three consecutive months with a dimension of $72 \times 24 \times 6$ ($5° \times 5°$ degree over 0°–360°E, 55°S–60°N) as input, and produces a total of 35 nodes, of which 23 nodes are Niño3.4 index (i.e., area-averaged SST anomaly over 170°–120°W, 5°S–5°N) from 1 to 23-lead months and 12 nodes are probability vectors that denote the observed calendar month of the input variable.

To obtain the probability, a softmax classifier is utilized to normalize the sum of the output values to one. Therefore, the output of the softmax classifier can be thought of as the likelihood for each class. The softmax function σ is defined as follows:

$$\sigma(x_i) = \frac{\exp(x_i)}{\sum_{j=1}^{12}\exp(x_j)}, \quad (1)$$

where x_i represents the input of the softmax classifier for output node i (of which there are 12 in this study). During labeling, each calendar month of the input variable is represented using a 12-dimensional one-hot vector [28], and a one-hot vector denotes a specific calendar month of the input variable. For example, the climate sample during the March-April-May (MAM) season is denoted as [0,0,1,0,0,0,0,0,0,0,0,0].

The softmax classifier with a dimension of 12 to denote the observed calendar months in the output layer is a significant modification to A_CNN when compared to H19. With this modification, the weighting coefficient in the model is expected to become a function of the calendar month, enabling the model to predict ENSO events while considering seasonality. While the H19 models are formulated separately for each target season and forecast lead month, the A_CNN model can be used universally for all target seasons and all lead months. Note that the overall performance of the A_CNN is quite similar to that of the CNN which adds the calendar month information to the neuron in the 2nd FC layer (not shown here). However, by adding the calendar month information in the output layer, it is advantageous to analyze the degree of recognition of the calendar months in the samples can be quantified by comparing the predicted calendar month to the true observed month.

The kernel size of the convolutional layers is 8×4, 4×2, 4×2, respectively. Spatial pooling is carried out between each convolution layer by a 2×2 max-pooling layer with a stride of 2. The total number of convolutional kernels and nodes in the fully connected layers is determined using hyperparameter tuning [29], which is set to 35 and 50, respectively. We generated 40 ensemble members with different random initial weights, and the predicted Niño3.4 indices and the one-hot vector value were averaged to obtain the final forecast results.

We applied transfer learning [22]. During the initial training, the size of the batch for each epoch was set to 400, and we applied early stopping by using an independent validation dataset. For fine-tuning, the number of epochs is set to 40. The hyperbolic tangent function (tan$h$) is chosen as an activation function to consider the symmetric feature between El Niño and La Niña. The Xavier initialization is applied to define the initial weights [30]. A_CNN utilizes an Adam optimizer [31]. The learning rate was fixed at 0.005 for the 1st training and 0.0005 for the 2nd training. Regularization and dropout were not used. Our source code is available at https://github.com/jeonghwan723/A_CNN.

2.2. The loss function

The loss function (L) of A_CNN is defined as the sum of the mean square error for the Niño3.4 index and the cross-entropy for calendar months:

$$L = \alpha\left(\frac{1}{23}\sum_{i=1}^{23}\left(\widehat{N}_i - N_i\right)^2\right) - (1-\alpha)\left(\sum_{j=1}^{12} M_j \log \widehat{M}_j\right). \quad (2)$$

Here, $N(\widehat{N})$ and $M(\widehat{M})$ denote the observed (predicted) Niño3.4 index, and the class denoting the calendar months of the input variables, respectively. i and j denote the forecast lead and calendar

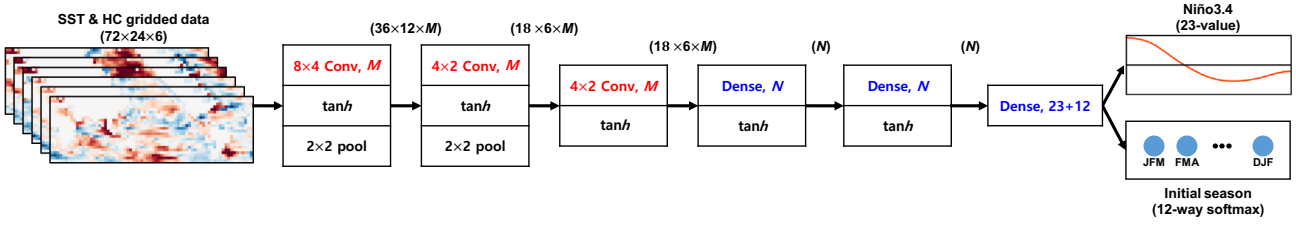

Fig. 1. Architecture of the All-season CNN (A_CNN) model for ENSO forecasts. The A_CNN model comprises one input layer (a.k.a., predictor), three convolutional layers, two max-pooling layers, and two fully connected layers. There are two types of output variables (a.k.a., predictands): one with 23 scalar values that denote the 3-month moving average Niño3.4 index from $\tau + 0$ to $\tau + 22$ and one 12-dimension probability vector that denotes the observed calendar months of the input variable. The variables of the input layer correspond to the SST and the oceanic heat content anomaly maps from time τ–3 to τ–1, between 0°–360°E and 55°S–60°N. M denotes the number of convolutional kernels (which is 35 in this study), while N denotes the number of nodes in the FC layer (which is 50 in this study).

months, respectively. α is a parameter that determines the weighting between the categorical forecast (for predicting calendar month) and the Niño3.4 forecast, which is determined as 0.8 through hyperparameter tuning [29].

2.3. Reanalysis data

A 103 years of the monthly mean SST and heat content data from the reanalysis data (between years 1871 and 1973) from the Simple Ocean Data Assimilation version 2.2.416 [32] were used for fine-tuning of the CNN model. The reanalysis product after 1974 was not used in any training process to ensure that the training and validation periods are independent of each other. To test the performance of the model by comparison with the observed values, the Global Ocean Data Assimilation System reanalysis during 1984–2017 was used [33]. For training the CNN model, the horizontal resolution is coarsened to 5° × 5° to reduce the number of weights [22], and for the analysis, the original horizontal resolution (i.e., 1° × 1°) is kept. For the CNN model training, the domain over 0°–360°E, 55°S–60°N is utilized. All the observed data are used as a form of the anomaly, which is defined as a deviation from the long-term time-mean values at the corresponding calendar month.

2.4. Model simulations

The historical simulations produced by the 21 Coupled Model Intercomparison Project phase 5 (CMIP5) [34] models were used to train the CNN model. It is used for 1st training and validation, and detailed information on the dataset is listed in Table S1 (online). The data are prescribed as a form of the anomaly for training the CNN models. The domain over 0°–360°E, 55°S–60°N is utilized, and the horizontal resolution of the CMIP5 samples is coarsened to 5° × 5° degree for the CNN model training. The forecast output from the SINTEX-F [12] during 1984–2017 was used to compare the prediction performance of the A_CNN model.

2.5. Activation map analysis

To quantify the contribution of each grid point to the output variable, the activation map is calculated by the following equation [22]: note that the activation map in this study is basically same as the Regression Activation Maps [35] to localize the discriminative region towards the regression output, even though the CNN system in this previous literature is different from ours in terms of usage of a fully connected layer and the type of pooling.

$$h^{x,y} = \sum_{n=1}^{N} \left\{ \tan h \left[\sum_{m=1}^{M_L} \left(W_{F,m,n}^{x,y} v_{L,m}^{x,y} \right) + \frac{b_{F,n}}{X_L Y_L} \right] W_{O,n} \right\} + \frac{b_O}{X_L Y_L}, \tag{3}$$

where $h^{x,y}$ denotes the activation at grid points (x, y), and X_L and Y_L denote the dimensions of the feature map in the last (i.e., 3rd) convolutional layer (i.e., X_L=18; Y_L=6). N denotes the number of neurons in the fully connected layer, $W_{F,m,n}^{x,y}$ denotes the weight at grid point (x, y) (used to link the mth feature map in the last convolutional layer L to the nth neuron in the fully connected layer F), $v_{L,m}^{x,y}$ denotes the value of the mth feature map of the last convolutional layer L at grid point (x, y), $b_{F,n}$ denotes the bias of the nth neuron in the fully connected layer F; $W_{O,n}$ denotes the weight (used to link the nth neuron in the fully connected layer to the output layer O), and b_O denotes the bias of the output layer O.

3. Results

3.1. Minimizing the "spring predictability barrier" problem by incorporating seasonality in climate data

The all-season correlation skill of the Niño3.4 index in A_CNN exhibited a similar degree to that of H19, even though the former's correlation skill was slightly higher than those of the H19 models for the forecasts earlier than a 7-month lead (Fig. S2a online). This demonstrates that the recognition of the seasonality in the climate data in A_CNN allows the successful prediction of the Niño3.4 index for all target seasons and forecast lead months with a significantly lower computational cost. The computational cost for formulating the A_CNN is reduced to 1/30 of that for the H19 models.

While the superiority of the A_CNN is only shown for early forecast lead months for all-season correlation, the increase in the forecast skill in A_CNN is particularly robust for most of the lead months for the forecasts initiated in the boreal spring, which is a season that is especially difficult to predict. For example, accuracy improvements were prominent for the forecasts initiated using anomalies during the MAM season (or forecasts initiated on 1st June) (Fig. 2a). The correlation skill of the Niño3.4 index in the H19 models is only 0.65 between the 4- and 10-lead month period (i.e., between the August-September-October (ASO)(0) and MAM(1) seasons), whereas it remained above 0.75 in the A_CNN model. The higher forecast skill in A_CNN is confirmed by the predicted time-series of the Niño3.4 time-series for individual El Niño events (Fig. S3 online) and La Niña events (Fig. S4 online). The predicted Niño3.4 index in A_CNN is systematically closer to the observed, than that in H19 in a larger number of forecast cases.

To confirm the skill improvement of A_CNN, we performed statistical significance testing by leveraging the bootstrap method. First, we randomly selected the same number of ensemble members among all ensemble members (i.e., 40) by allowing the overlapping. Then, we calculated the forecast skill of the ensemble mean of the randomly selected members. This procedure is repeated 1000 times; then, the upper and lower 2.5% values are denoted. Therefore, the error bar denotes the 95% confidence range. The bootstrapping test confirms that the forecast skill in A_CNN is outside of the correlation skill range of H19, confirming that the forecast skill in A_CNN is significantly higher than that in H19.

This result was mainly achieved because the decrease in correlation skill for the 1- to 3-month lead in A_CNN was not as severe as that in H19, which indicated that A_CNN was less affected by the spring predictability barrier problem, which is a common problem in most ENSO prediction models. For example, the correlation skill drops to 0.7 within 4-lead months in H19, contrastingly, while it maintains around 0.8 in the A_CNN. This fast drop in the forecast skill is evident for most of the state-of-the-art prediction systems [11], where the forecasts initiated on June 1st exhibit correlation skill around 0.6 at 6-months lead; contrastingly, it reaches up to 0.75 in the A_CNN.

The successful forecasts of Niño3.4 for the forecasts initiated in the boreal spring season are likely due to A_CNN being trained simultaneously for all forecast lead months. As the dynamic relationship for the time evolution of the Niño3.4 for 2-year forecast periods was implicit in the training samples for A_CNN, the predicted Niño3.4 values at the later forecast stages constrained those at the early forecast stages. As Niño3.4 during boreal winter was relatively easy to predict, these predictions could potentially help determine whether an ENSO event was initiated during the preceding boreal spring.

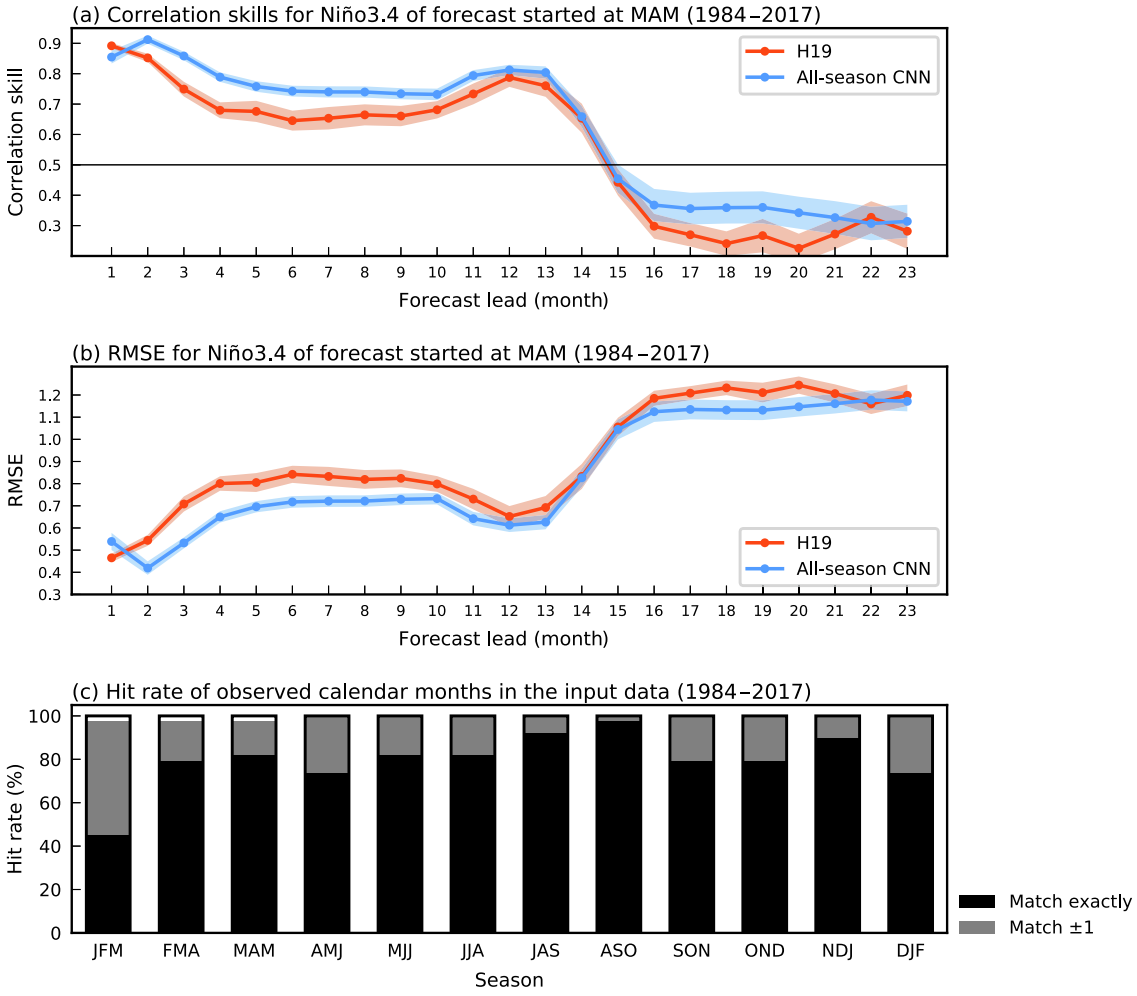

Fig. 2. Forecast accuracy of the output variables in A_CNN. (a) Correlation skill, and (b) root-mean-square-error (RMSE) of the 3-month moving average Niño3.4 index obtained using the H19 (red) and A_CNN (blue) models for forecasts employing input data from the MAM season. Red, and blue areas denote 95% confidence range of the H19, and A_CNN using the bootstrapping method, respectively. (c) Hit rate (ratio of the number of correctly predicted cases to the total cases, unit: %) of observed calendar months in the input data. The black bar denotes the observed calendar months that are accurately recognized by A_CNN, whereas the gray bar denotes the observed calendar months that are predicted with a 1-month error.

In addition, the improvement of the ENSO forecast skill was achieved because the A_CNN model was optimally trained with an increased number of climate samples from all months. Compared to the H19 model initialized during the MAM season, the number of climate samples used to train the A_CNN model was 12 times larger, and it was advantageous to train the model optimally. This implies that the recognition of seasonality in ENSO events contributes to improving the ENSO forecast skill by training with samples for all target seasons and forecast lead months simultaneously.

To examine how thoroughly the seasonality in climate data is recognized by the A_CNN model, Fig. 2c shows the hit rate of the initial season in the model. The hit rate was calculated as the percentage of cases that correctly predicted the observed calendar months of the input data divided by the total forecast cases. It should be noted that we considered the probability vector output with a 1-month error as the correct forecast because the primary goal was the accurate recognition of the seasonality. For example, for the case of MAM (i.e., the arguments of the maxima (argmax) = 3), the probability vector output denoting February-March-April (i.e., argmax = 2) or AMJ (i.e., argmax = 4) was also treated as the correct forecast. The hit rate was higher than 95% for all calendar months. This demonstrated that A_CNN thoroughly recognized the observed calendar season of input variables.

To examine the source of seasonality in the input data, Fig. 3a presents an activation map for the output of the softmax classifier denoting the MAM season with given SST and HC anomalies for MAM 1996 (see Section 2.5 for the activation map calculation). The activation is positive for most of the globe, and the amplitude of positive values is stronger than that of the activations for other seasons (Fig. S5 online). This indicates that A_CNN accurately predicted the observed calendar months of the input variables for this case.

Specifically, the positive activation is robust over the northeastern Pacific, where the dipole SST anomalies are located (i.e., positive SST anomalies over the far northeastern Pacific and negative SST anomalies over the northern Central Pacific; contours in Fig. 3a). This dipole SST pattern over the North Pacific is similar to the typical pattern of the Pacific Meridional Mode (PMM), which exhibits the strongest variability during boreal spring [36,37]. To demonstrate this point, we defined the index as the SST difference between the northeastern Pacific box (red box in Fig. 3a) and the northern Central Pacific box (blue box in Fig. 3a), where the SST anomalies during MAM 1996 were the strongest. The regressed SST fields exhibited a dipole pattern over the North Pacific (shadings in Fig. 3b), which was similar to the typical spatial distribution of the positive phase of the PMM event (contours in Fig. 3b). The pattern correlation between them over 120°E–120°W, 5°–60°N is

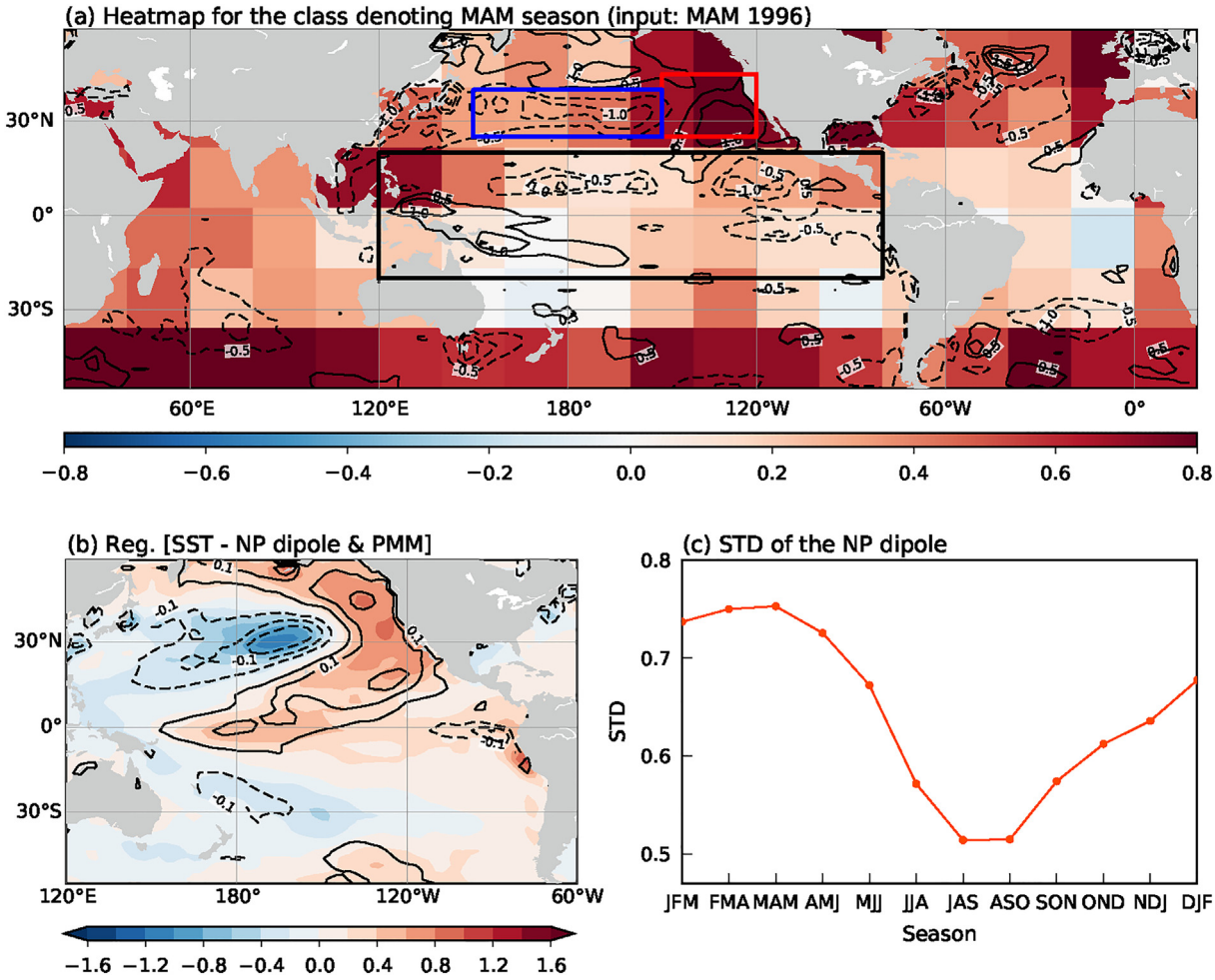

Fig. 3. The imprinted seasonality in the climate data recognized by A_CNN. (a) Heat map values for the class denoting the MAM season (i.e., argmax = 3) with input variables from MAM 1996. The contour denotes the heat content (inside the black box in the panel) or SST (outside the black box in the panel) anomalies from MAM 1996. (b) Regression coefficients of SST anomalies during the MAM season with respect to the SST dipole index over the north Pacific (i.e., [SST averaged over 150°–120°W, 20°–45°N] - [SST averaged over 150°E–150°W, 20°–40°N], shading) and the spatial distribution of the Pacific Meridional Mode (PMM, contour). (c) Standard deviation of the SST dipole index over the north Pacific with respect to the calendar months.

0.96. This SST dipole index exhibited the strongest variability during the MAM season (Fig. 3c), thus allowing A_CNN to accurately recognize seasonality in climate samples. The overall positive activations near the Southern Ocean might be attributed to the delay in the remote SST response to the ENSO by a few months to exhibit peak phase during boreal spring [38].

3.2. Physical interpretation using activation map

To compare the performance of the ENSO forecasts obtained using the A_CNN and H19 models in greater detail, Fig. 4a shows the predicted Niño3.4 time-series initialized during MAM 1996. When using the H19 model, as Niño3.4 was predicted independently using multiple models for each forecast lead month, artificial noise was observed in the predicted Niño3.4 indices between adjacent seasons. In other words, the Niño3.4 time-series predicted using the H19 model was not as smooth as that in the observation. However, the predicted Niño3.4 index of A_CNN exhibited a smooth time evolution with a similar degree to the observed index. In addition, the Niño3.4 predicted using A_CNN was positive from the boreal summer 1997 to boreal spring 1998, and exhibited a peak phase during DJF 1997; by comparison, this index could not be predicted successfully using the H19 model. We note that although A_CNN better predicts the time evolution of the 1997/98 El Niño event, there is still significant underprediction of the total magnitude.

In addition, A_CNN exhibited clear advantages with regard to the physical interpretations of the forecasts. The activation map extracted from A_CNN was more spatially organized than that from the H19 model (Fig. 4b–g). For the latter, the activation map for 6-month lead forecasts with initial conditions during MAM 1996 exhibits scattered positive and negative values; therefore, the large-scale signal is not elucidated (Fig. 4b). This prevents a clear physical interpretation of the H19 model results. Conversely, the activation map from A_CNN shows large-scale signals over the equatorial Pacific (Fig. 4e). In other words, the large-scale negative activations, which denote the contribution to La Niña, are prominent over the eastern Pacific, and weak positive values, which denote the contribution to El Niño, are shown across the western and off-equatorial central Pacific.

These activations can be explained physically. Specifically, the negative HC anomaly over the eastern Pacific (contours in Fig. 4e) is associated with the upwelling of the colder subsurface temperature leading to La Niña events [39]. In contrast, the positive HC anomaly over the equatorial western Pacific and the negative HC anomaly over the off-equatorial Pacific denote the clear recharged HC over the equatorial Pacific, which facilitates El Niño events [40,41]. However, as the time lag of 6 months is, to a certain

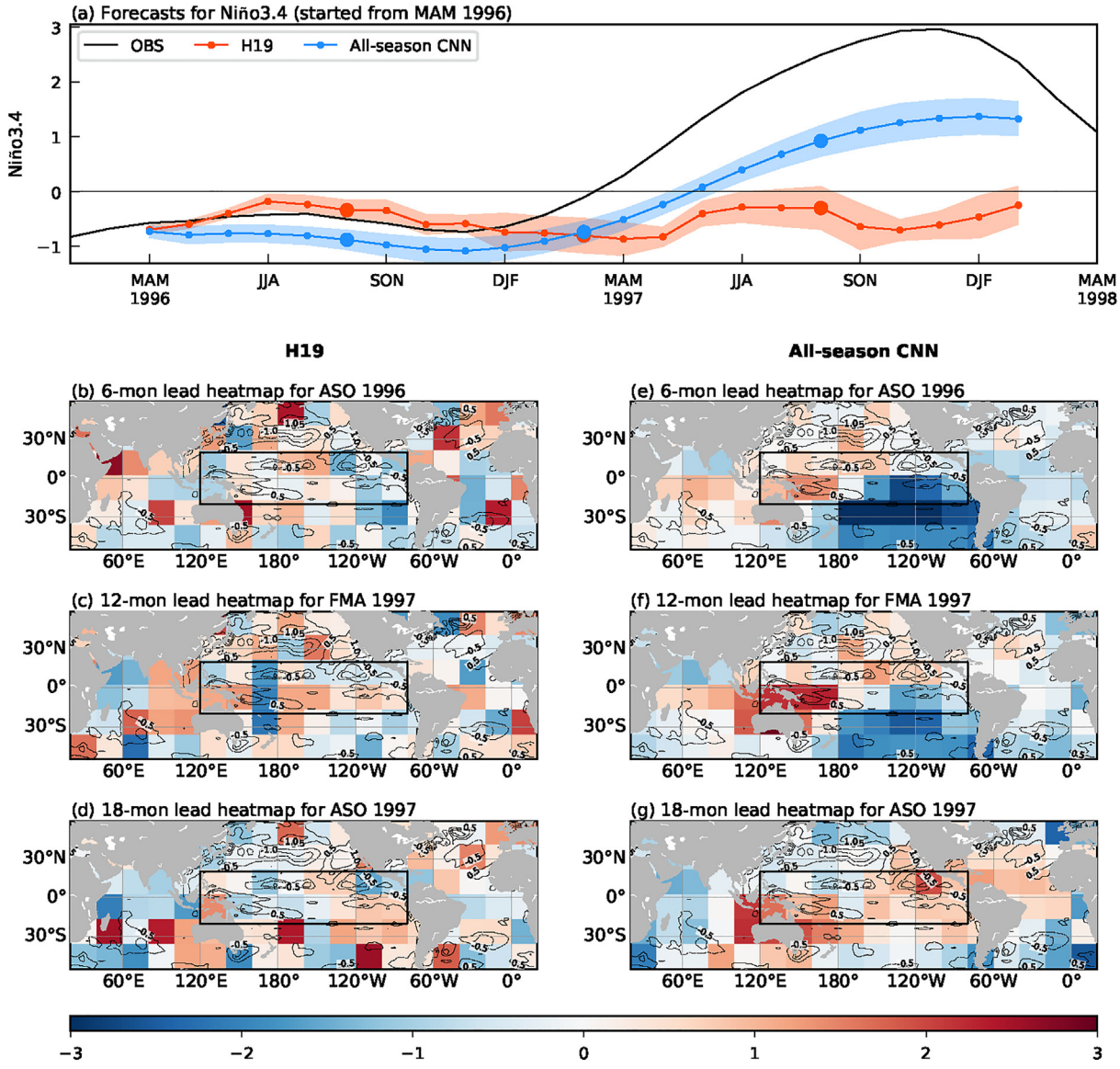

Fig. 4. Comparison of heat map values for the 1997/1998 El Niño event. (a) Predicted 3-month moving average Niño3.4 time-series with input variables at MAM 1996 obtained using the H19 model (red) and A_CNN (blue). The reference time-series from 1996 to MAM 1998, determined via the reanalysis, is also shown (black). Red, and blue areas denote 95% confidence range of the H19, and A_CNN using the bootstrapping method, respectively. Heat map values (shading) for 6- (b), 12- (c), and 18-month (d) lead forecasts initialized with anomalies at MAM 1996 for the H19 model. Heat maps for the 6- (e), 12- (f), and 18-month (g) lead forecasts for A_CNN. The contours in panels (b)–(g) denote the heat content (inside the black box in the panel) or SST (outside the black box in the panel) anomalies at MAM 1996.

extent, shorter than that of the recharged equatorial HC anomalies over the western Pacific that catalyze the El Niño signal [42], the positive activations over the equatorial western Pacific are weaker than the negative activations over the equatorial eastern Pacific. As the large-scale negative activations are overwhelmed by the positive values, the final Niño3.4 prediction is weakly negative for the 6-month lead forecast (denoted by a blue circle in ASO 1996).

For the 12- and 18-month lead forecasts, the activations from the H19 models still do not exhibit a large-scale feature. Moreover, the spatial distribution of the activation map is significantly different among different forecast lead months. This supports the notion that the H19 model lacks consistency in the time evolution of ENSO events. However, in A_CNN, a large-scale cancelation between the negative values over the eastern Pacific and positive values over the western Pacific is evident for the 12-month lead forecasts (Fig. 4c, f respectively). Compared to the activation map for the 6-month lead forecast, the 12-month lead forecast map shows that the area with negative activations is diminished and that the area with positive activations is expanding. This indicates that the recharged equatorial HC anomalies during the boreal spring over the western Pacific can significantly intensify an El Niño event after 12 months [42]. This is also consistent with the gradual increase of the predicted Niño3.4 index with increasing lead time, as indicated by A_CNN.

In addition, for the 18-month lead forecast, the positive activations are overwhelmed over the entire equatorial and northeastern Pacific. This means that the negative HC anomalies over the equatorial eastern Pacific can contribute to an El Niño event after 18 months. This is associated with an oscillatory feature of the ENSO, whereby a developing La Niña signal from the boreal spring season exhibits a peak phase in boreal winter in the same year, then turns into an El Niño in boreal fall after 1 year (i.e., biennial tendency of the ENSO) [43].

Another interesting feature is the positive activations over the northeast Pacific for the 12- and 18-month lead forecasts. This indicated that A_CNN captured the contribution of the positive

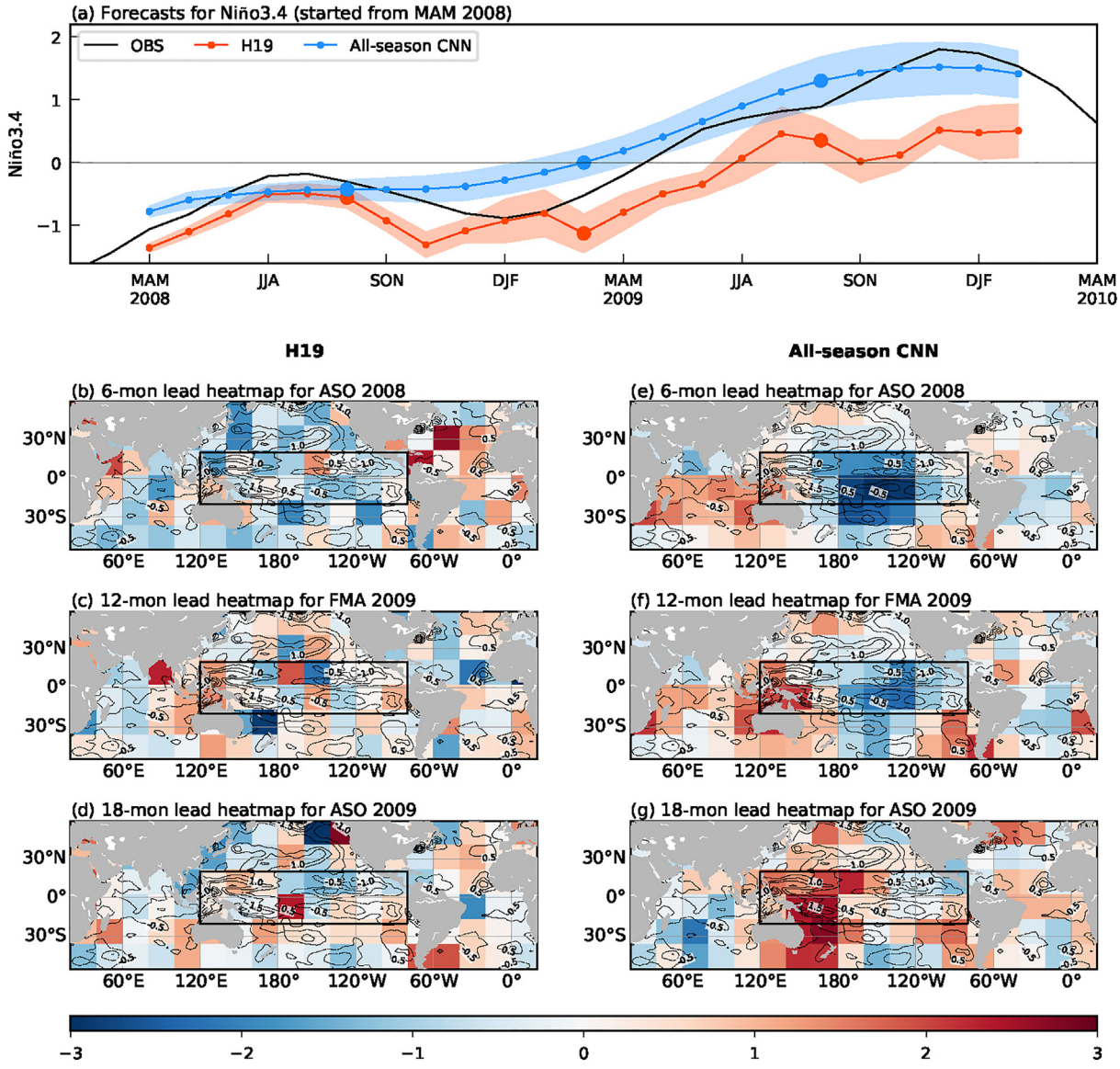

Fig. 5. Comparison of heat map values for the 2009/2010 El Niño. (a) Predicted 3-month moving average Niño3.4 time-series with input variables at MAM 2008 obtained using the H19 model (red) and A_CNN (blue). The reference time-series from 2008 to MAM 2010, determined via the reanalysis, is also shown (black). Red, and blue areas denote 95% confidence range of the H19, and A_CNN using the bootstrapping method, respectively. Heat map values (shading) for 6- (b), 12- (c), and 18-month (d) lead forecasts initialized with anomalies at MAM 2008 for the H19 model. Heat maps for the 6- (e), 12- (f), and 18-month (g) lead forecasts for A_CNN. The contours in panels (b)–(g) denote the heat content (inside the black box in the panel) or SST (outside the black box in the panel) anomalies at MAM 2008.

phase of the PMM that sustained up to the next boreal spring (i.e., MAM 1997) (Fig. S6 online) [44], leading to the El Niño event during the subsequent winter (i.e., DJF 1997/98) [45]. It should be noted that the positive activation over the northeastern Pacific was not observed for the 6-month lead forecast activation map, as the time lags from PMM that affect the ENSO events are longer than 6 months [21,38].

It is worth emphasizing that the activations under the same initial conditions exhibited changes in the forecast lead time for the A_CNN model. For instance, the negative SST anomalies over the equatorial eastern Pacific contributed to the La Niña signal in the 6-month lead forecast, whereas they contributed to the El Niño signal in the 18-month lead forecast. This indicates that A_CNN successfully simulated different effects of the prescribed initial climate signals on ENSO events for different lead times. This also implied that A_CNN allowed a diagnosis of how the given climate signals in the input layer influenced the time evolution of ENSO events.

The activation map analysis for the forecasts initialized at MAM 2008 yielded similar features (Fig. 5). In A_CNN, (1) the predicted Niño3.4 showed a smoothed time evolution, as in the observation, with a higher accuracy (Fig. 5a), and (2) the activations exhibited a large-scale spatial distribution with time evolution. This confirms the notion that (1) A_CNN performs better in predicting the initiation of El Niño events, which are the most difficult to predict, and (2) physical interpretations of the forecasts are easier when using A_CNN. This implies that the role of a deep learning model as a diagnostic tool can be further enhanced by implementing physical constraints for the time evolution of ENSO events.

For the 6-month lead forecast, the negative activations in A_CNN were prominent throughout the equatorial Pacific, with the peak amplitude over the equatorial central-eastern Pacific. The negative HC anomalies over the central-eastern Pacific were responsible for the negative activation through the vertical advection of cold subsurface temperature anomalies to the surface layer. In addition, positive activations over the Indian Ocean were also

weakly indicated for the 6- and 12-month lead forecasts. This was, to a certain extent, consistent with previous literature, which indicates that the negative Indian Ocean SST during a previous boreal winter can increase the Niño3.4 index after 1 year [46,47]. The activation map for the 18-month lead forecast using A_CNN exhibited strong positive values over the equatorial western Pacific, which indicated that the recharged HC anomalies over the western Pacific contributed to El Niño events after 18 months (Fig. 5g). Conversely, the positive activations were less evident in the H19 models (Fig. 5d), which was consistent with the weaker amplitude of the predicted Niño3.4 index in these models.

4. Discussion and conclusion

Although H19 first demonstrated that deep learning is a promising algorithm to increase forecasting accuracy for the most dominant climate variability, that is, ENSO, the deep learning model in H19 was premature in terms of considering the seasonality of ENSO events. Specifically, the characteristics in the climate data are excessively diverse between seasons; therefore, the performance of the H19 model is significantly degraded once it is trained by simultaneously using climate samples from all months. A model should be able to recognize the seasonality in climate data to achieve accurate forecasts of the complete time evolution of an ENSO event.

To achieve this goal, a deep learning model for the ENSO was developed to recognize the seasonality in the climate data. We added the softmax classifier as an output variable that denoted the observed calendar months of the input climate data. With this modification, the weighting coefficients that constituted the deep learning model became a function of the calendar months, and the different time evolution features of the ENSO (i.e., seasonal ENSO phase-locking) were successfully considered.

For this proposed A_CNN model, the all-season correlation skill was slightly higher than that in H19, which demonstrated that the unique characteristics between climate samples in different seasons were accurately simulated. Specifically, the A_CNN model performed better for forecasts starting during the boreal spring season, as it minimized the spring predictability barrier problem.

The computational costs to formulate the deep learning model were also significantly reduced when using A_CNN. In H19, the total number of deep learning models for all target seasons and all forecast periods numbered in the hundreds (i.e., 23 (lead months) $\times$ 12 (target seasons) = 276). By comparison, a single A_CNN model could produce ENSO forecast results for all target seasons and all lead months. More importantly, as the physical constraints to predict 2-year Niño3.4 time-series were implemented in A_CNN, the activations exhibited a clear time evolution with increasing lead times. This allowed for an understanding of the comprehensive role of various climate precursors of ENSO events that act differently over time.

This study shows that the incorporation of seasonality imprinted in climate data benefited the deep learning model for ENSO forecasts by appropriately considering the season-dependent characteristics of these events (e.g., amplitude growth rate, phase transition speed, duration, interaction with other variabilities, and predictability). This modification is directly applicable to other climate phenomena (such as the Madden–Julian Oscillation and Atlantic Niño), whose characteristics are related to their phase [48,49].

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF)(NRF-2020R1A2C2101025).

Author contributions

Yoo-Geun Ham and Jeong-Hwan Kim designed the study. Yoo-Geun Ham wrote the majority of the manuscript. Jeong-Hwan Kim formulated a statistical model and conducted a hindcast experiments. All authors contributed to the interpretation of the results and the fine tuning of the model.

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2021.03.009.

References

[1] Tziperman E, Cane MA, Zebiak SE, et al. Locking of El Niño's peak time to the end of the calendar year in the delayed oscillator picture of ENSO. J Clim 1998;11:2191–9.
[2] An SI, Wang B. Mechanisms of locking of the El Niño and La Niña mature phases to boreal winter. J Clim 2001;14:2164–76.
[3] Timmermann A, An SI, Jin FF, et al. El Niño-Southern Oscillation complexity. Nature 2018;559:535–45.
[4] Tziperman E, Cane MA, Zebiak SE. Irregularity and locking to the seasonal cycle in an ENSO prediction model as explained by the quasi-periodicity route to chaos. J Atmos Sci 1995;52:293–306.
[5] Tziperman E, Zebiak SE, Cane MA. Mechanisms of seasonal–ENSO interaction. J Atmos Sci 1997;54:61–71.
[6] Harrison DE, Vecchi GA. On the termination of El Niño. Geophys Res Lett 1999;26:1593–6.
[7] Vecchi GA, Harrison DE. On the termination of the 2002–03 El Niño event. Geophys Res Lett 2003;30.
[8] Xiao H, Mechoso CR. Seasonal cycle–El Niño relationship: validation of hypotheses. J Atmos Sci 2009;66:1633–53.
[9] Ham YG, Kug JS, Kim D, et al. What controls phase-locking of ENSO to boreal winter in coupled GCMs? Clim Dyn 2013;40:1551–68.
[10] Webster PJ, Yang S. Monsoon and ENSO: selectively interactive systems. Q J Roy Meteorol Soc 1992;118:877–926.
[11] Barnston AG, Tippett MK, Ranganathan M, et al. Deterministic skill of ENSO predictions from the North American Multimodel Ensemble. Clim Dyn 2019;53:7215–34.
[12] Luo JJ, Masson S, Behera SK, et al. Extended ENSO predictions using a fully coupled ocean–atmosphere model. J Clim 2008;21:84–93.
[13] Balmaseda MA, Davey MK, Anderson DL. Decadal and seasonal dependence of ENSO prediction skill. J Clim 1995;8:2705–15.
[14] Levine AF, McPhaden MJ. The annual cycle in ENSO growth rate as a cause of the spring predictability barrier. Geophys Res Lett 2015;42:5034–41.
[15] Izumo T, Vialard J, Lengaigne M, et al. Influence of the state of the Indian Ocean Dipole on the following year's El Niño. Nat Geosci 2010;3:168–72.
[16] Saji NH, Goswami BN, Vinayachandran PN, et al. A dipole mode in the tropical Indian Ocean. Nature 1999;401:360–3.
[17] Dayan H, Vialard J, Izumo T, et al. Does sea surface temperature outside the tropical Pacific contribute to enhanced ENSO predictability? Clim Dyn 2014;43:1311–25.
[18] Keenlyside NS, Latif M. Understanding equatorial Atlantic interannual variability. J Clim 2007;20:131–42.
[19] Ding H, Keenlyside NS, Latif M. Impact of the equatorial Atlantic on the El Niño Southern Oscillation. Clim Dyn 2012;38:1965–72.
[20] Ham YG, Kug JS, Park JY. Two distinct roles of Atlantic SSTs in ENSO variability: north tropical Atlantic SST and Atlantic Niño. Geophys Res Lett 2013;40:4012–7.
[21] Park JH, Kug JS, Li T, et al. Predicting El Niño beyond 1-year lead: effect of the Western Hemisphere warm pool. Sci Rep 2018;8:1–8.
[22] Ham YG, Kim JH, Luo JJ. Deep learning for multi-year ENSO forecasts. Nature 2019;573:568–72.
[23] Mahesh A, Evans M, Jain G, et al. Forecasting El Niño with convolutional and recurrent neural networks. Abstract in 33rd conference on neural information processing systems, 2019.
[24] Broni-Bedaiko C, Katsriku FA, Unemi T, et al. El Niño-Southern Oscillation forecasting using complex networks analysis of lstm neural networks. Artif Life Rob 2019;24:445–51.
[25] Yan J, Mu L, Wang L, et al. temporal convolutional networks for the advance prediction of ENSO. Sci Rep 2020;10:1–15.
[26] Li T, Wang B, Chang CP, et al. A theory for the Indian Ocean dipole–zonal mode. J Atmos Sci 2003;60:2119–35.

[27] Bjerknes J. Atmospheric teleconnections from the equatorial Pacific. Mon Weather Rev 1969;97:163–72.
[28] Harris D, Harris S. Digital design and computer architecture. 2nd ed. San Francisco: Morgan Kaufmann; 2012. p. 129.
[29] Snoek J, Larochelle H, Adams RP. Practical bayesian optimization of machine learning algorithms. Adv Neural Inf Process Syst 2020;25:2951–9.
[30] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. p. 249–56.
[31] Kingma P, Diederik B, Adam LJ. A method for stochastic optimization. arXiv:1412.6980, 2014.
[32] Giese BS, Ray S. El Niño variability in simple ocean data assimilation (SODA), 1871–2008. J Geophys Res Oceans 2011;116:C02024.
[33] Behringer DW, Xue Y. Evaluation of the global ocean data assimilation system at NCEP: the Pacific Ocean. Proceedings of eighth symposium on integrated observing and assimilation systems for atmosphere, oceans, and land surface (AMS 84th Annual Meeting), 2004. and Land Surface (AMS 84th Annual Meeting).
[34] Taylor KE, Stouffer RJ, Meehl GA. An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 2012;93:485–98.
[35] Wang Z, Yang J. Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation. arXiv:1703.10757, 2017.
[36] Vimont DJ, Wallace JM, Battisti DS. The seasonal footprinting mechanism in the Pacific: implications for ENSO. J Clim 2003;16:2668–75.
[37] Chang P, Zhang L, Saravanan R, et al. Pacific meridional mode and El Niño–Southern oscillation. Geophys Res Lett 2007;34:L16608.
[38] Wilson AB, Bromwich DH, Hines KM, et al. El Niño flavors and their simulated impacts on atmospheric circulation in the high southern latitudes. J Clim 2014;27:8934–55.
[39] An SI, Kang IS. A further investigation of the recharge oscillator paradigm for ENSO using a simple coupled model with the zonal mean and eddy separated. J Clim 2000;13:1987–93.
[40] Jin FF. An equatorial ocean recharge paradigm for ENSO. Part I: conceptual model. J Atmos Sci 1997;54:811–29.
[41] Ren HL, Jin FF. Recharge oscillator mechanisms in two types of ENSO. J Clim 2013;26:6506–23.
[42] Anderson BT. On the joint role of subtropical atmospheric variability and equatorial subsurface heat content anomalies in initiating the onset of ENSO events. J Clim 2007;20:1593–9.
[43] Rasmusson EM, Wang X, Ropelewski CF. The biennial component of ENSO variability. J Mar Syst 1990;1:71–96.
[44] Joh Y, Di Lorenzo E. Interactions between Kuroshio Extension and Central Tropical Pacific lead to preferred decadal-timescale oscillations in Pacific climate. Sci Rep 2019;9:1–12.
[45] Larson SM, Kirtman BP. The Pacific meridional mode as an ENSO precursor and predictor in the North American multimodel ensemble. J Clim 2014;27:7018–32.
[46] Kug JS, Kang IS. Interactive feedback between ENSO and the Indian Ocean. J Clim 2006;19:1784–801.
[47] Zhang F, Zhang G, Liu L, et al. The negative feedback effects of sea surface temperatures on El Niño Events in the West Indian Ocean. Atmos Sci Lett 2019;20:e924.
[48] Kim H, Vitart F, Waliser DE. Prediction of the Madden–Julian oscillation: a review. J Clim 2018;31:9425–43.
[49] Xie SP, Carton JA. Tropical Atlantic variability: patterns, mechanisms, and impacts. Earth's climate: the ocean-atmosphere interaction. Geophys Monogr 2004;147:121–42.

Yoo-Geun Ham got his Ph.D. degree from Seoul National University in 2009. He then joined NASA Goddard Space Flight Center in USA as a research scientist. In 2013, he joined Department of Oceanography in Chonnam National University as a professor. His research interest includes subseasonal-seasonal-decadal climate forecasts using deep learning and dynamical global climate models.