Self-supervised Classification of Weather Systems Based on Spatiotemporal Contrastive Learning

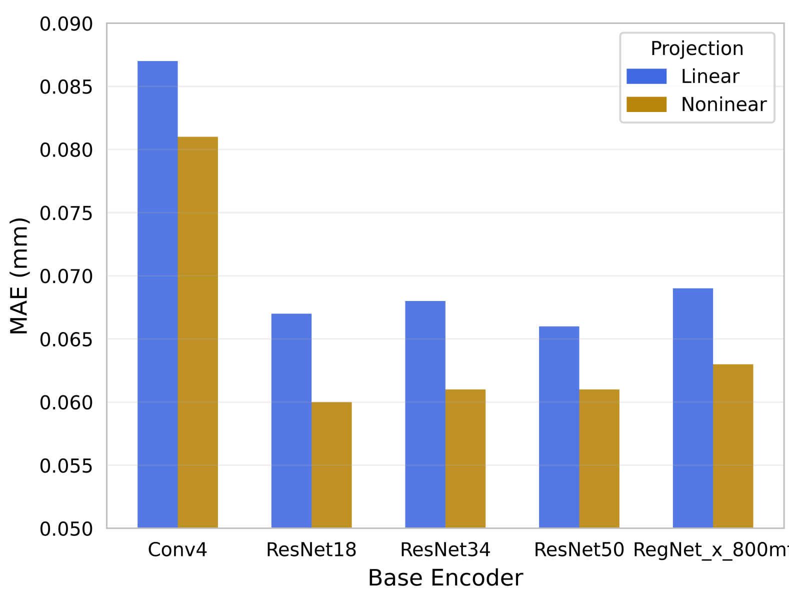
Liwen Wang[1], Qian Li[1], and Qi Lv[1]

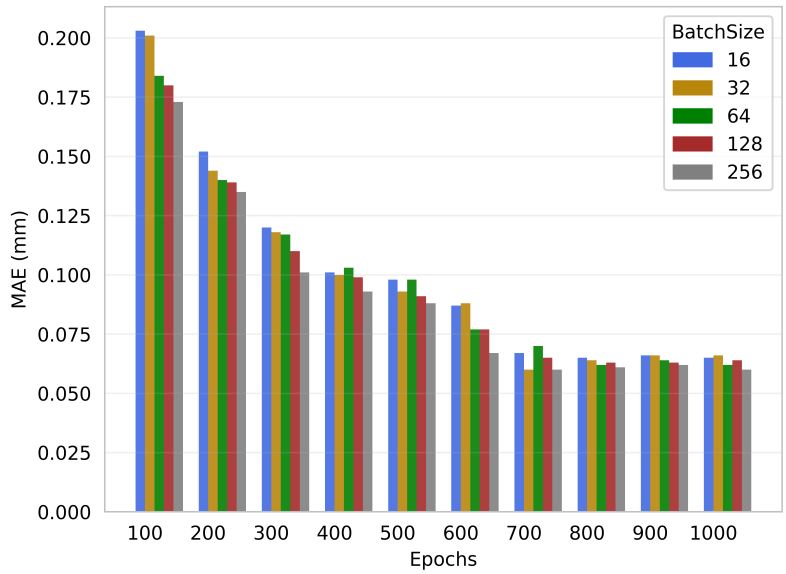
[1]National University of Defense Technology

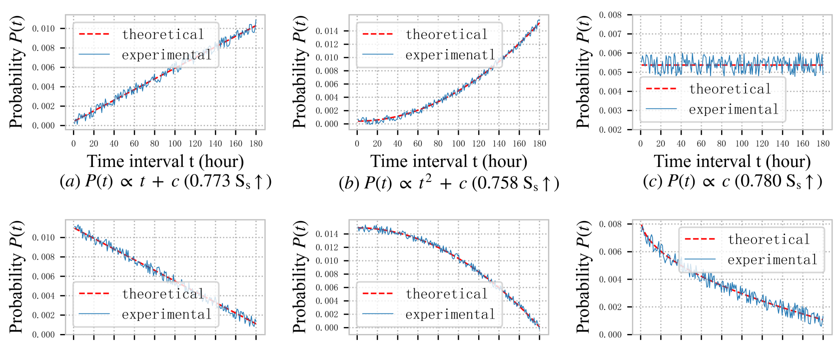
November 26, 2022

Abstract

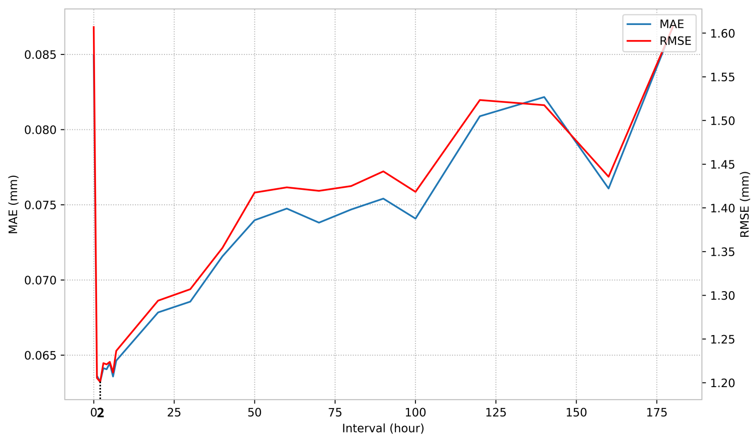
Classification of weather systems provides a simple description of atmospheric circulations and bridges the gap between large-scale atmospheric conditions and local-scale environmental variables. However, the existing classification methods are challenged due to lack of labels and inaccurate similarity measures between data samples. In this letter, we propose a self-supervised Spatiotemporal Contrastive Learning (SCL) framework for the classification of weather systems without manual labels. In particular, we operate both spatial and temporal augmentation on multivariate meteorological data to fully explore temporal context information and spatial stability in accordance with synoptic nature. With the classification results, we apply a statistical downscaling method based on analog forecasting for the assessment and comparison of classification results. The experimental results demonstrate that the proposed SCL model outperforms traditional classification methods.

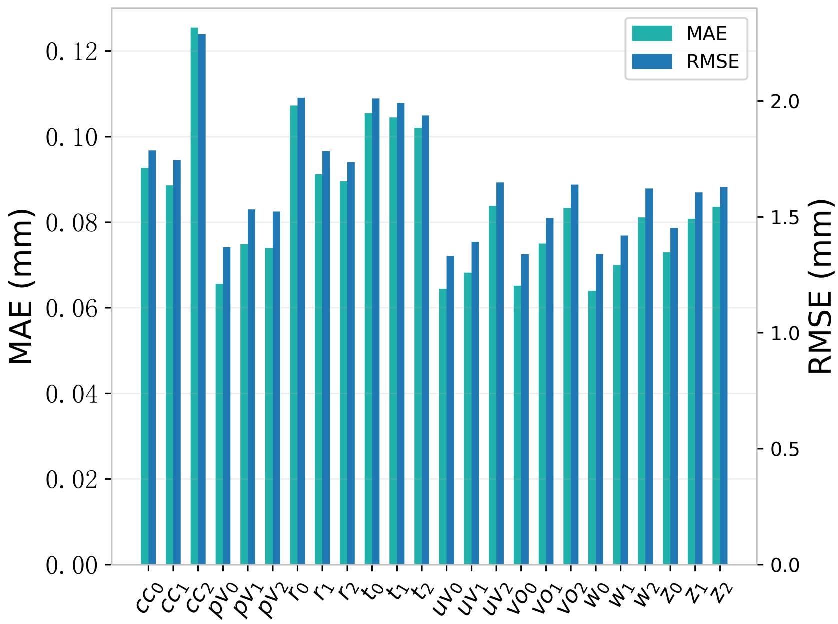






(a) $P(t) \propto t + c$ (0.773 $S_s\uparrow$)
(b) $P(t) \propto t^2 + c$ (0.758 $S_s\uparrow$)
(c) $P(t) \propto c$ (0.780 $S_s\uparrow$)
(d) $P(t) \propto -t + c$ (0.843 $S_s\uparrow$)
(e) $P(t) \propto -t^2 + c$ (0.824 $S_s\uparrow$)
(f) $P(t) \propto -t^{0.5} + c$ (0.803 $S_s\uparrow$)


MAE
RMSE
MAE (mm)
RMSE (mm)
Interval (hour)


MAE
RMSE
MAE (mm)
RMSE (mm)

Self-supervised Classification of Weather Systems Based on Spatiotemporal Contrastive Learning

Liwen Wang, Qian Li, Qi Lv

[1]College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China

Key Points:

- A Spatiotemporal Contrastive Learning (SCL) convolutional neural network is presented for the classification of weather systems without manual labels, fully utilizing temporal context information and spatial stability of element fields in accordance with synoptic nature.
- A statistical downscaling method is utilized to assess the model based on analog forecasting.

Corresponding author: Qian Li, `public_liqian@163.com`

Abstract

Classification of weather systems provides a simple description of atmospheric circulations and bridges the gap between large-scale atmospheric conditions and local-scale environmental variables. However, the existing classification methods are challenged due to lack of labels and inaccurate similarity measures between data samples. In this letter, we propose a self-supervised Spatiotemporal Contrastive Learning (SCL) framework for the classification of weather systems without manual labels. In particular, we operate both spatial and temporal augmentation on multivariate meteorological data to fully explore temporal context information and spatial stability in accordance with synoptic nature. With the classification results, we apply a statistical downscaling method based on analog forecasting for the assessment and comparison of classification results. The experimental results demonstrate that the proposed SCL model outperforms traditional classification methods.

Plain Language Summary

In recent years, deep learning has contributed greatly to the field of meteorology and we perform an investigation into the use of convolutional neural networks (CNNs) to classify typical weather systems without manual labels. Although CNNs have produced remarkable results in image classification, few works have evaluated their efficiency and accuracy in weather system classification. Highly accurate and automated weather system classification approaches, especially the technology of mining temporal context information, are essential for discovering the relationships between atmospheric circulation and local weather, climate and environmental variables. Moreover, explicit classification of typical weather systems would promote the study of climate change. Therefore, a discriminative and comprehensive weather system classification model, called SCL, is built for CNN training, which could achieve remarkable progress compared with conventional approaches. We also propose a method of generating pseudo labels for the training of a linear classifier. Moreover, statistical downscaling forecasting is utilized to assess the classification results of SCL and various conventional methods.

1 Introduction

Classification of weather systems (CWS) refers to the categorization of high-dimensional multivariate meteorological data into a reasonable and manageable number of typical weather systems that share similar meteorological fields, physical characteristics and evolutionary trends. Therefore, CWS has been applied broadly in weather forecasts and statistical climatology. For example, CWS has been utilized to facilitate weather forecasts, where each weather pattern of ensemble members is assigned to the closest matching type, hence reducing the ensemble forecasts to a sequence of circulation type probabilities (Chattopadhyay et al., 2020; Neal et al., 2016; Ohba et al., 2018). In climatology, classification helps to examine climate-scale changes in the frequency of circulation types (Luong et al., 2020; Lynch et al., 2006; Gibson et al., 2016).

Subjective CWS is heavily dependent on expert experience, which makes it labor intensive and time consuming. Furthermore, subjective classification results lack generalizability due to regional differences in underlying surfaces and atmospheric evolution laws. Therefore, a significant demand has arisen for objective classification with high accuracy in weather analysis. There are three main typical objective methods with computer-assisted analysis. The first is based on clustering, such as k-means clustering (Cuell & Bonsal, 2009; Esteban et al., 2006) or correlation-based methods (Brinkmann, 2000), in which element fields (temperature, humidity, etc.) are directly clustered into different groups in accordance with Euclidean distance between two samples. However, these methods incur high computational costs and have difficulty converging on nonconvex meteorological data. The second approach is based on the empirical orthogonal function (EOF)

(Dilinuer et al., 2021; Miró et al., 2018), which allows dominant spatial weather modes to be identified by means of matrix decomposition and samples to be classified through similarity search. However, this kind of methods assume a linear relationship between and within the elements for matrix decomposition and fail to utilize nonlinear relationship. For example, turbulence dominates the vertical exchange of elements such as momentum, heat and moisture in the atmospheric boundary layer (Monin, 1967), which is a highly nonlinear system. The last approach is mainly based on neural networks, for example, self-organizing maps (SOMs) (Bao & Wallace, 2015; Berkovic, 2017; Da-wei et al., 2018), which project samples onto a two-dimensional lattice and classify them into major weather patterns. At present, the SOM method is superior to the first two objective analysis approaches (Iseri et al., 2009), due to its ability to deal with nonlinear relationships.

Nevertheless, previous computer-assisted methods are heavily reliant on Euclidean distance between lattice points for similarity measures, which are applicable only to low-dimensional data. On the other hand, current classification methods deal with data matrixes as separate linear arrays, thereby discarding temporal context information and inner spatial stability; thus, the accuracy and stability of the classification results are still unsatisfactory for meteorological data, which is characterized by high dimensions and multiple elements.

In recent years, deep learning has achieved great progress in abstract representation learning for high-dimensional data by virtue of its inherent highly nonlinear transformation characteristics. Research has begun to focus on how to integrate deep learning into the meteorological field, fully utilize massive-scale observation data and build models that are more suitable for practical application needs to improve the refinement and accuracy of meteorological forecasts (Wu et al., 2021; Xing et al., 2021). To address the issues with the aforementioned methods, we propose a self-supervised Spatiotemporal Contrastive Learning (SCL) framework, in which a deep neural network for classification is trained with pseudo labels based on contrastive learning, in order to learn the invariance of key features in weather systems after transformations and then perform CWS on unlabeled data. The rest of this letter is structured as follows. We introduce the analyzed data in section 2. Section 3 describes the proposed SCL model, and section 4 presents the experimental details, results and analysis. Finally, conclusions are summarized in section 5.

2 Data

In this study, the selected area for model training and testing is in East China, in the range of 105-125°E and 25-35°N, with a resolution of 0.25°×0.25°. The hourly gridded atmospheric data is constructed from the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) data from 2014 to 2019. In the training procedure, seven variables are used: u and v components of the wind at 850 hPa; relative humidity at 850 hPa; temperature at 1000 hPa; geopotential height at 850 hPa; vertical velocity at 850 hPa; total precipitation. The first six variables are physically related to the total precipitation, which is the predictand. The u and v components of the wind can reveal low-level convergence and divergence, while the water vapor content of the lower atmosphere relates to relative humidity and surface temperature. The 850-hPa geopotential height and vertical velocity determine whether the initial conditions for precipitation are met. Then, an independent dataset in 2020 is utilized to test the model and perform synoptic analysis on classification results.

3 Methodology

3.1 The Framework for SCL

Inspired by Chen et al. (2020) and Qian et al. (2021), we propose a self-supervised framework based on contrastive learning for representation learning on multivariate meteorological data, in which spatiotemporal transformations are applied for data augmentation in order to utilize the invariance of key features after transformations. Contrastive learning (Jaiswal et al., 2020) is a discriminative approach that aims to group similar samples closer in the embedding space while spreading diverse dissimilar samples farther from each other.

As depicted in Figure 1, given an initial dataset of N samples with C elements, we randomly sample a minibatch of n examples and operate spatial augmentation ($F_i \sim \Gamma$) (described in section 3.2.1) and temporal augmentation ($t_i \sim \tau$) (described in section 3.2.2) on x_i and x_j, which are any two examples derived from the minibatch, to obtain 4 augmented data samples, among which (A_i, A_i') and (A_j, A_j') are positive pairs while (A_i, A_j), (A_i, A_j'), (A_i', A_j) and (A_i', A_j') are negative pairs. After feature extraction, a contrastive loss is utilized to group positive pairs closer and separate negative pairs farther from each other.

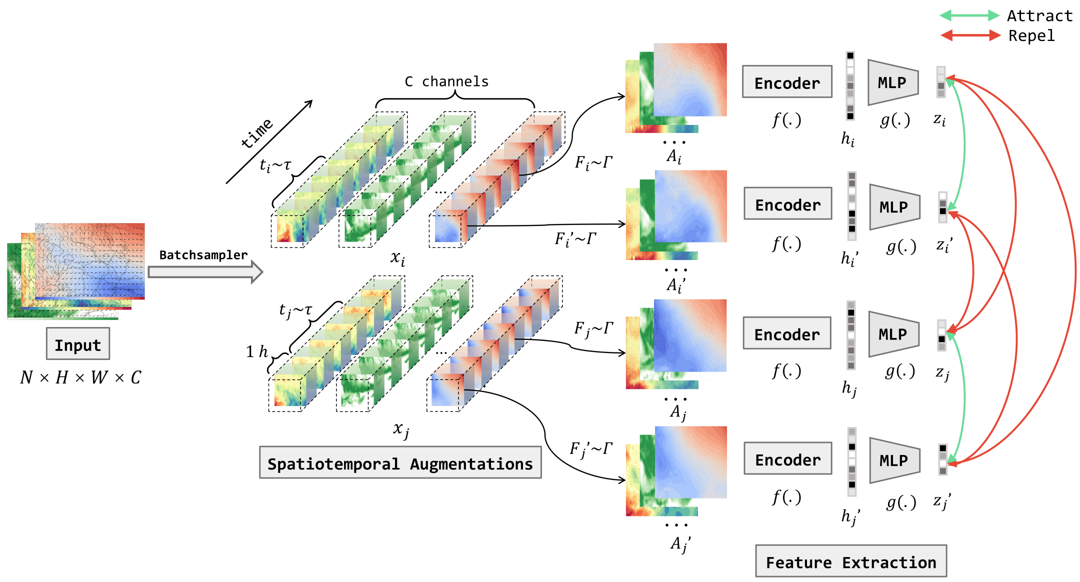
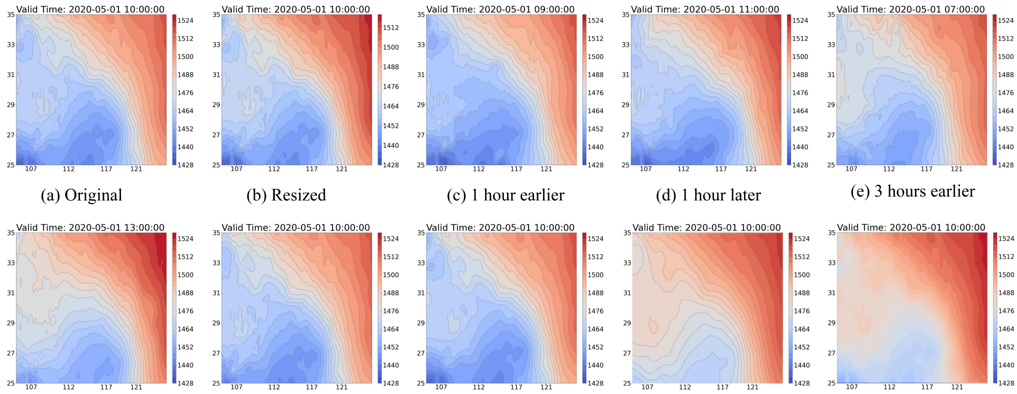


Figure 1. Overview of the proposed SCL framework.

3.2 Data Augmentation

In the field of machine learning, data augmentation methods mainly utilize the strong correlations between augmented and original samples to expand the training set. In this process, key features should still be retained. In addition, in contrastive learning, sample pairs are built based on the prior knowledge that the labels remain the same after various augmentation operations so as to reduce the workload of manual labeling, which is a key step in contrastive learning for teaching a model to differentiate positive samples from negative ones. In this letter, we augment the dataset from the perspectives of both space and time by utilizing spatial stability and temporal invariance in weather systems. Figure 2 shows augmented samples with spatiotemporal augmentation operators



(a) Original (b) Resized (c) 1 hour earlier (d) 1 hour later (e) 3 hours earlier

(f) 3 hours later (g) RandomCrop (h) GaussianBlur (i) kernel_5×5 (j) Gaussiannoise

Figure 2. Illustrations of spatiotemporal augmentation operators.

3.2.1 Spatial Transformations

As a general rule, a certain spatial displacement of a weather system does not change the overall attributes of the synoptic situation. For example, an area controlled by a trough of low pressure could consistently be characterized by cloudy and rainy conditions, lasting for several days, although the meteorological elements (wind, humidity, etc.) at individual grid points would not remain constant. Based on this assumption, three kinds of spatial transformations are introduced in SCL for dataset augmentation, namely, resizing, random cropping and mean filtering, to change grid point values while ensuring that the system type, structures and dominant characteristics of the original data remain constant.

The first spatial transformation is to resize the original data matrixes from (40, 80) to (100, 200) through bicubic interpolation; then, we randomly crop 90×180 patches as augmented samples. Furthermore, mean filtering is applied to filter out random fluctuations of the multivariate meteorological data by sliding a 5×5 mean filter across data, replacing the center value with the average of all pixel values in the window.

3.2.2 Temporal Transformations

Weather systems are relatively stable on a certain time scale; for example, samples from consecutive time instances generally belong to the same weather system, and the shorter the time interval is, the more similar the weather characteristics are. However, current classification methods assume independence between samples during analysis, thereby discarding temporal context information. Inspired by Qian et al. (2021), this paper takes as prior knowledge that different samples from adjacent time tend to belong to the same weather system; accordingly, data samples separated by short time intervals are treated as positive samples, while samples from different periods are treated as negative samples. In this way, pseudo labels are established for all training samples on the basis of temporal correlation, thereby reducing the workload of sample labeling.

In this letter, the following time series sampling strategy is adopted: given a sampling interval of t hours and an initial sample x_i at time H, the sample x_i' at time $H+t_i$ is selected to form a positive pair (x_i, x_i') with the initial sample. The sampling interval t_i is selected from a distribution $\tau(t)$ over $[0, T]$, and the probability distribution P decreases monotonically with increasing time (discretely from 0 to 180 hours). In this way, samples separated by shorter time intervals are pulled closer in the embedding space.

3.3 Representation Learning Based on Contrastive Learning

Abstract representations for multivariate meteorological data are extracted using an encoder $f(\cdot)$, as depicted in Figure 1. We adopt ResNet18 (He et al., 2016) as the base encoder, that is $h_i = f(A_i)) = ResNet(A_i)$. Then, we use a multilayer perceptron with one hidden layer (and rectified linear unit (ReLU) activation), denoted by $g(\cdot)$, to map the output h_i to the embedding space where the contrastive loss is applied, and thus obtain the feature, that is, $z_i = g(h_i)$.

After feature extraction, a contrastive loss is constructed to ensure that positive samples are close in the embedding space while negative samples are far away to drive the learning of abstract representations for samples of similar weather systems. We do not apply this contrastive loss directly to h_i because Chen et al. (2020) has verified that it is beneficial to define the contrastive loss based on z rather than h. We sample N raw weather clips and obtain $2N$ augmented views, consisting of one positive pair and $2(N-1)$ negative pairs. Let $sim(z_i, z_i') = z_i^\top z_i'/\|z_i\|\|z_i'\|$ denote the dot product between the l_2-normalized z_i and z_i', the loss function for a positive pair of examples (z_i, z_i') is defined as

$$\ell_{i,i'} = -\log \frac{\exp\left(\operatorname{sim}\left(\boldsymbol{z}_i, \boldsymbol{z}_i'\right)/\sigma\right)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp\left(\operatorname{sim}\left(\boldsymbol{z}_i, \boldsymbol{z}_k\right)/\sigma\right)} \tag{1}$$

where $\mathbb{1}_{[k \neq i]} \in \{0, 1\}$ is an indicator function evaluating to 1 if $k \neq i$ and σ denotes a temperature parameter. The final loss is computed across all positive pairs, both (z_i, z_i') and (z_i', z_i), in a minibatch.

4 Experimental Results

To verify our model, we conduct experiments on Ubuntu 20.04 with a single RTX 3090 GPU. The Adam optimizer is adopted for parameter optimization. We apply a minibatch size of 256 and an initial learning rate of 3e-4 to train ResNet18. First, we compare our model with traditional methods (k-means and SOM). Then, we conduct ablation studies to verify the selection of augmentation operators and parameters of contrastive learning. Finally, the hourly precipitation is used to assess the classification results, for which the skill score ($S_s\uparrow$) (Perkins et al., 2007a), brier score ($S_b\downarrow$) and mean absolute error (MAE$\downarrow$) are used as evaluation metrics.

4.1 Evaluation Metrics

Considering the difficulty of obtaining manual labels for CWS, we focus on analog forecasting (prediction with similar patterns) to assess the classification results, based on the premise that the better classification results are, the more accurate the analog forecasting results will be; this approach has also been applied to assess CWS models by Nishiyama et al. (2004), Pu and Zhihong (2016) and Xianghua et al. (2018).

Specifically, we utilize precipitation probability density functions (PDFs) (Nishiyama et al., 2004; Perkins et al., 2007b; Yin et al., 2011) for a more complete assessment of a climate model's capacity to forecast the complete range of observations at daily time scales. These PDFs are less likely to be influenced by observation errors than the mean or standard deviation. To measure the common area between two PDFs, the metric S_s is used, which is calculated as the cumulative minimum value of each binned value between the two distributions, expressed as follows:

$$S_s = \sum_{1}^{n} \min(P_m, P_o) \tag{2}$$

where n is the number of bins used to calculate the PDF for a given region (16 by default), P_m is the frequency of values in a given bin from the model, and P_o is the frequency of values in a given bin from the observed data. If a model inaccurately classifies weather systems and forecasts the observed PDF poorly, it will have a skill score close to zero with negligible overlap between the observed and modeled PDFs. This provides a robust and comparable measure of the relative similarity between modeled PDFs and observed PDFs. To measure the accuracy of probabilistic forecasts, we adopt the brier score (S_b), which quantitatively represents the region where the predicted and observed PDFs do not coincide, expressed as follows:

$$S_b = \frac{1}{n}\sum_{i=1}^{n}(P_m - P_o)^2 \tag{3}$$

The closer S_b is to zero, the smaller the PDF difference between forecasting and observation is. Therefore, a lower value of S_b and a higher value of S_s indicate better classification performance.

To further measure the difference between forecasts and observations, we also use the MAE of hourly precipitation to comprehensively assess our model.

4.2 Evaluation of Data Augmentation

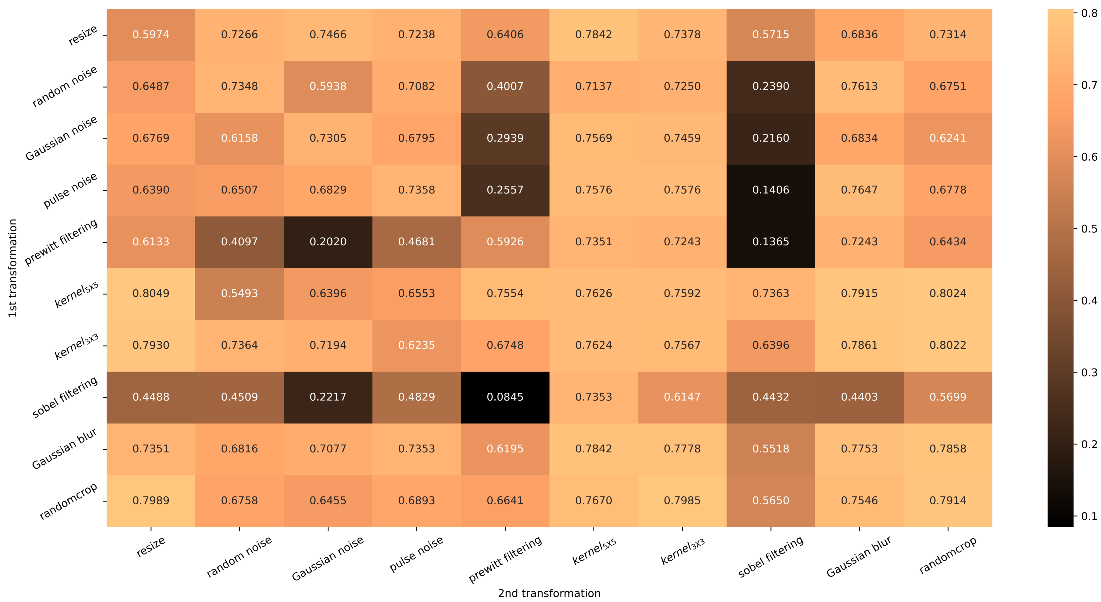


Figure 3. Skill scores (↑) under individual or composition of data augmentation, applied only to one branch. For all columns but the last, diagonal entries correspond to single transformation, and off-diagonals correspond to composition of two transformations (applied sequentially). The last column reflects the average over the row. If a model classifies weather systems perfectly, the skill score will equal to one.

To verify the performance of different types of spatial augmentations, ten spatial transformations are tested in this letter. We evaluate their performance both separately and in combination. Figure 3 shows that when augmentations are combined, the contrastive prediction task becomes more difficult, but the quality of representation could be improved because more redundant information is eliminated while only key features are retained, which helps the SCL model learn more robust features. The combination of augmentations that yields the best result is resizing and $kernel_{5\times5}$. In addition, random cropping outperforms other single spatial augmentations. We observe that blurring also achieves good performance while sobel filters and prewitt filters have a detrimental impact on data augmentation because such excessive transformations can destroy key features in weather systems. Therefore, it is beneficial to combine random cropping with resizing and $kernel_{5\times5}$ to help our model learn more generalizable features.

4.3 Synoptic Interpretation of SCL clustering

The classification results are shown in Figure 4 in 2020. The frequencies of each type, denoted by h, evenly distribute in $[4\%, 11\%]$, showing that SCL achieves good performance in separating dissimilar systems and grouping similar systems and could fairly reveal characteristics of local climate. The 850 hPa geopotential height field, lying between 1200 m and 1900 m, is selected for analysis, which represents a strong correlation with the near-surface boundary layer, and will not be significantly affected by small local topographic features.

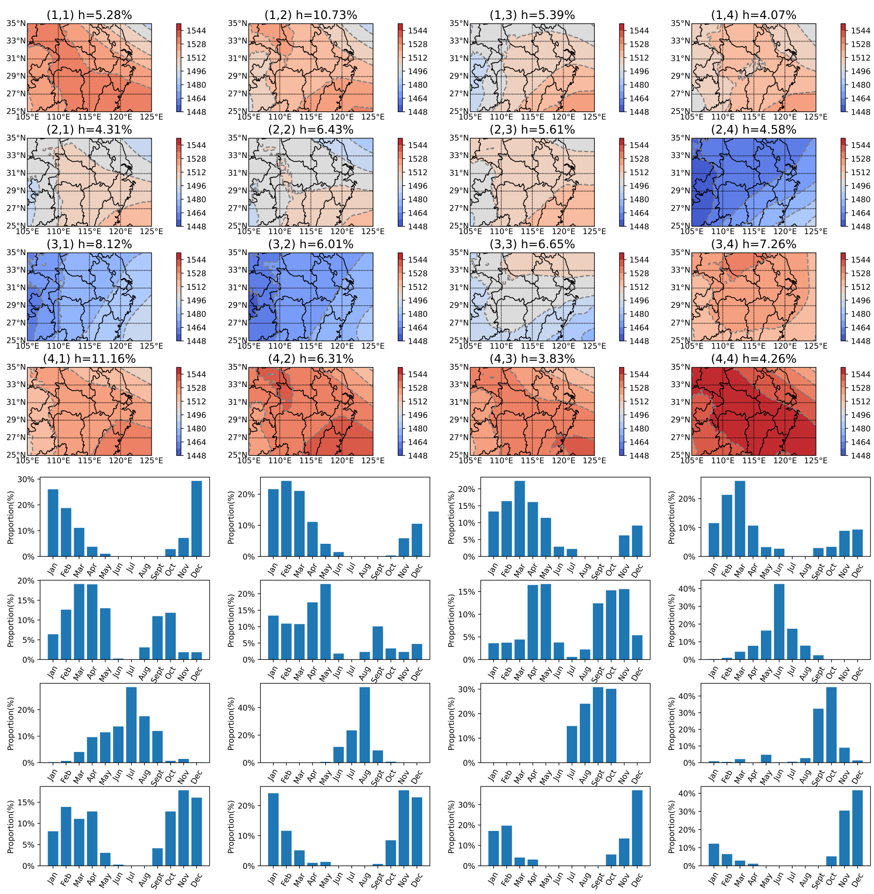


Figure 4. SCL classification results for the 850 hPa geopotential height fields (gpm), obtained by averaging data samples in the same category. Results are arranged in chronological order. The corresponding occurrence frequencies are presented in the lower plots, arranged in chronological order, from January to December.

For simplicity, we set the number of weather types to 16. The study area is located on the eastern coast of the subtropical continent, where tropical marine air mass and polar continental air mass alternately dominate and compete for each other (Ke-yi et al., 2020). As a result, the four seasons are distinct with different circulation characteristics, as depicted in Figure 4.

In the upper left and bottom right corners of Figure 4, the weather types (1,1), (1,2), (4,2), (4,3) and (4,4) are in winter, mainly dominated by the Siberia high. The western Pacific shows a lower pressure field due to the land-sea thermal contrast. Weather types (2,4), (3,1) and (3,2) indicate that the study area is influenced by the Qinghai-Tibet high pressure system in summer. Autumn ((3,4) and (4,1)) is the transitional season from summer to winter, when the angle of solar radiation decreases. In early September ((3,3) and (3,4)), there is frequent southward movement of cold air into the middle and lower reaches of the Yangtze River, prompting a rapid southward movement of warm and humid air that remains there in summer (Lei et al., 2021; Ming et al., 2019). Therefore, in September and October, the surface of the middle and lower reaches of the Yangtze River is often controlled by a cold high-pressure zone (fang Sang, 2012). However, the Pacific subtropical high pressure in summer has not returned south, so this area remains under the control of high pressure (Qin et al., 2022). After October ((4,2) and (4,3)), the upper subtropical high moves southward, and the middle and lower reaches of the Yangtze River come under the control of the western wind belt, resulting in more rainfall than in autumn (Shengnan & Zhihong, 2019; Ke-yi et al., 2020).

Based on the relationship between the geopotential field and wind, in the middle and lower reaches of the Yangtze River, the northerly wind in winter ((4,3),(4,4) and (1,1)) comes from the interior of the continent, and the southerly wind in summer comes from the ocean, forming a monsoon climate pattern with simultaneous rain and heat. The summer monsoon wind is southeasterly, and the winter monsoon wind is northwesterly (Shengnan & Zhihong, 2019). In early July ((3,1)), the middle and lower reaches of the Yangtze River are controlled by the subtropical high-pressure zone. It is obvious that the subtropical high gradually moves north from January to July ((1,1) $\rightarrow$ (2,4)), while the Siberia High moves north too, backwards from August to December ((3,2) $\rightarrow$ (4,4)). In general, weather systems are separate from each other and could reveal the key features of local climate.

4.4 Comparison with Traditional Methods

Table 1. Performance assessment of SCL and traditional methods with different similarity metrics. SCL* refers to SCL ending with a clustering module (k-means by default); SCL† and SCL‡ indicate that only apply spatial and temporal augmentations are applied respectively.

Method	MAE(mm)(↓)	S_s(↑)	S_b(↓)
k-means-l_1	0.128	0.541	1.369×10^{-3}
k-means-l_2	0.121	0.510	1.715×10^{-3}
k-means-*cosine*	0.177	0.478	1.762×10^{-3}
SOM-l_1	0.164	0.388	2.001×10^{-3}
SOM-l_2	0.133	0.530	1.779×10^{-3}
SOM-*cosine*	0.191	0.461	1.945×10^{-3}
SCL*	0.095	0.538	1.37×10^{-3}
SCL†	0.085	0.790	8.60×10^{-4}
SCL‡	0.069	0.801	7.32×10^{-4}
SCL	**0.060**	**0.843**	$\mathbf{3.94 \times 10^{-4}}$

Our model is compared with commonly used traditional methods, k-means and SOM, in terms of the accuracy and stability of the results. Similarity research in traditional methods is necessary, and we choose l_1 distance, l_2 distance and cosine distance to measure to what extent two samples are similar.

After feature extraction, the labels calculated from hourly precipitation are utilized for the training of a linear classifier. We also replace the linear classifier with a clustering module in the embedding space (the default clustering algorithm is k-means), denoted by SCL*, to verify the effectiveness of representations learned in SCL.

As illustrated in table 1, SCL shows competitive performance compared with traditional models. Both spatial (SCL†) and temporal (SCL‡) augmentation can separately improve the classification performance separately, and their combination yields the best performance. SCL* (ending with a clustering module) still outperforms the traditional methods, which shows that our framework can extract effective representations from the raw data.

5 Conclusion

Based on the establishment of relationships between meteorological elements and weather systems, CWS has been widely used in the fields of weather forecasts and climate research. However, a lack of manual labels and inaccurate similarity measures limit the accuracy and stability of current methods. Inspired by contrastive learning proposed in recent years, we constructed a spatiotemporal contrastive learning model to address this issue. Temporal context information and spatial stability have been fully explored in accordance with synoptic nature to enhance the capacity of a model to learn key features of weather systems. In addition, a statistical downscaling rainfall prediction method based on analog forecasting is used to assess the model. As illustrated in experiments, SCL with spatiotemporal augmentation outperforms traditional classification methods in terms of accuracy. In future work, we will focus on representation learning and classification for weather processes instead of individual samples, taking advantage of RNN-based networks.

6 Open Research

All the data necessary to reproduce the results of this work could be downloaded at `https://doi.org/10.24381/cds.bd0915c6` and `https://doi.org/10.24381/cds.adbb2d47`. The scripts used for classification are freely available at `https://github.com/JOJOeth/SCL`.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant No. 42075139 and 41305138), the China Postdoctoral Science Foundation (Grant No. 2017M621700) and the Hunan Province Natural Science Foundation (Grant No. 2021JJ30773).

References

Bao, M., & Wallace, J. M. (2015). Cluster analysis of northern hemisphere wintertime 500-hpa flow regimes during 1920-2014. *Journal of the Atmospheric Sciences*, *72*, 3597-3608.

Berkovic, S. (2017). Winter wind regimes over israel using self-organizing maps. *Journal of Applied Meteorology and Climatology*, *56*, 2671-2691.

Brinkmann, W. A. R. (2000). Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation. *International Journal of Climatology*, *20*, 839-852.

Chattopadhyay, A., Nabizadeh, E., & Hassanzadeh, P. (2020). Analog forecasting of extreme-causing weather patterns using deep learning. *Journal of Advances in Modeling Earth Systems*, *12*.

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. E. (2020). A simple framework for contrastive learning of visual representations. *ArXiv*, *abs/2002.05709*.

Cuell, C., & Bonsal, B. R. (2009). An assessment of climatological synoptic typing by principal component analysis and kmeans clustering. *Theoretical and Applied Climatology*, *98*, 361-373.

Da-wei, L., Cholaw, B., & Zuo-wei, X. (2018). Classification of wintertime large-scale tilted ridges over the eurasian continent and their influences on surface air temperature. *Atmospheric and Oceanic Science Letters*, *11*, 404 - 411.

Dilinuer, T., Yao, J., Chen, J., Mao, W., Yang, L., Yeernaer, H., & Chen, Y. (2021). Regional drying and wetting trends over central asia based on köppen climate classification in 1961–2015. *Advances in Climate Change Research*.

Esteban, P., Martín-Vide, J., & Mases, M. (2006). Daily atmospheric circulation catalogue for western europe using multivariate techniques. *International Journal of Climatology*, *26*, 1501-1515.

fang Sang, Y. (2012). Spatial and temporal variability of daily temperature in the yangtze river delta, china. *Atmospheric Research*, *112*, 12-24.

Gibson, P. B., Perkins-Kirkpatrick, S. E., & Renwick, J. A. (2016). Projected changes in synoptic weather patterns over new zealand examined through self-organizing maps. *International Journal of Climatology*, *36*, 3934-3948.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In (p. 770-778).

Iseri, Y., Matsuura, T., Iizuka, S., Nishiyama, K., & Jinno, K. (2009). Comparison of pattern extraction capability between self-organizing maps and principal component analysis. *Memoirs of the Faculty of Engineering. Kyushu University*, *69*, 37-47.

Jaiswal, A., Babu, A. R., Zadeh, M. Z., Banerjee, D., & Makedon, F. (2020). A survey on contrastive self-supervised learning. *ArXiv*, *abs/2011.00362*.

Ke-yi, W., Yuan, L., Bu-chun, L., Mei-juan, Q., Xiao-juan, Y., Yue-ying, Z., … Jing-yi, P. (2020). Spatio-temporal characteristics of cold wave in the middle and lower reaches of the yangtze river between 1958 and 2015. *Journal of Natural Resources*, *35*(12), 10.

Lei, Z., Jun, X., & Yin, Z. (2021). Spatial-temporal characteristics of extreme precipitation in the middle and lower reaches of the yangtze river. *Resources and Environment in the Yangtze Basin*.

Luong, T., Dasari, H. P., & Hoteit, I. (2020). Extreme precipitation events are becoming less frequent but more intense over jeddah, saudi arabia. are shifting weather regimes the cause? *Atmospheric Science Letters*, *21*.

Lynch, A. H., Uotila, P., & Cassano, J. J. (2006). Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 2: Antarctic. *International Journal of Climatology*, *26*, 1181-1199.

Ming, L., Xu-rong, C., Gui-wen, W., Wei-xia, H., & Lian-zhi, Z. (2019). Research on meteorological drought in the middle and lower reaches of the yangtze river. *Journal of Natural Resources*(2), 11.

Miró, J. R., Peña, J. C., Pepin, N. C., Sairouní, A., & Aran, M. (2018). Key features of cold-air pool episodes in the northeast of the iberian peninsula (cerdanya, eastern pyrenees). *International Journal of Climatology*, *38*, 1105-1115.

Monin, A. (1967). Turbulence in the atmospheric boundary layer. *Physics of Fluids*, *10*.

Neal, R., Fereday, D. R., Crocker, R., & Comer, R. E. (2016). A flexible approach to defining weather patterns and their application in weather forecasting over europe. *Meteorological Applications*, *23*, 389-400.

Nishiyama, K., Endo, S., Jinno, K., Uvo, C. B., Olsson, J., & Berndtsson, R. (2004). Identification of typical synoptic patterns causing heavy rainfall in the rainy season in japan by a self-organizing map. *Atmospheric Research*, *83*, 185-200.

Ohba, M., Kadokura, S., & Nohara, D. (2018). Medium-range probabilistic forecasts of wind power generation and ramps in japan based on a hybrid ensemble. *Atmosphere*.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., & McAneney, J. (2007a). Evaluation of the ar4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over australia using probability density functions. *Journal of Climate*, *20*, 4356-4376.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., & McAneney, J. (2007b). Evaluation of the ar4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over australia using probability density functions. *Journal of Climate*, *20*, 4356-4376.

Pu, Z., & Zhihong, J. (2016). Simulation and evaluation of statistical downscaling of regional daily precipitation over yangtze–huaihe river basin based on self-organizing maps. *Climatic and Environmental Research*(5), 13.

Qian, R., Meng, T., Gong, B., Yang, M.-H., Wang, H., Belongie, S. J., & Cui, Y. (2021). Spatiotemporal contrastive video representation learning. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 6960-6970.

Qin, P., Xu, H., Liu, M., Liu, L., Xiao, C., Mallakpour, I., … Sorooshian, S. (2022). Projected impacts of climate change on major dams in the upper yangtze river basin. *Climatic Change*, *170*, 1-24.

Shengnan, W., & Zhihong, J. (2019). Synoptic classification and precipitation characteristics in summer over the yangtze river basin based on self-organizing map. *Journal of the Meteorological Sciences*, *39*(5), 11.

Wu, Y., Teufel, B., Sushama, L., Belair, S., & Sun, L. (2021). Deep learning-based super-resolution climate simulator-emulator framework for urban heat studies. *Geophysical Research Letters*.

Xianghua, W. U., Meng, F., Xiong, P., Huaying, Y. U., Yan, N., & Liu, W. (2018). A statistical simulation study on spring-summer precipitation in jilin province using self-organizing maps. *Transactions of Atmospheric Sciences*.

Xing, J., Zheng, S., Li, S., Huang, L., Wang, X., Kelly, J. T., … Hao, J. (2021). Mimicking atmospheric photochemical modeling with a deep neural network. *Atmospheric Research*.

Yin, C., Li, Y., Ye, W., Bornman, J. F., & Yan, X. (2011). Statistical downscaling of regional daily precipitation over southeast australia based on self-organizing maps. *Theoretical and Applied Climatology*, *105*, 11-26.