Assessment of Dynamic NMME Models for Meteorological Drought Forecasting in Western Regions of Iran

Mehdi Moghasemi
Islamic Azad University, Ahvaz Branch
Narges Zohrabi
nargeszohrabi@gmail.com

Islamic Azad University, Ahvaz Branch
Hossein Fathian
Islamic Azad University, Ahvaz Branch
Alireza Nikbakht Shahbazi
Islamic Azad University, Ahvaz Branch
Mohammadreza Yeganegi
Islamic Azad University Central Tehran Branch

Research Article

Keywords: NMME, SPEI, Iran, Karun, Karkheh

Posted Date: March 5th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3948603/v1

Additional Declarations: No competing interests reported.

Abstract

This study assessed the performance of North American multimodel ensemble (NMME) dynamic systems in forecasting meteorological drought within the western and southwestern watersheds of Iran. Without suitable observational data in this region, the global Precipitation Climatology Centre (GPCC) precipitation and Climatic Research Unit (CRU) temperature datasets served as the foundation for comparative analysis. The standardized precipitation evapotranspiration index (SPEI) was employed for drought evaluation. The findings indicated that longer forecast horizons significantly reduced model accuracy. Furthermore, the assessment of drought predictability based on SPEI revealed that both CanCM3 and CanCM4 models could predict seasonal drought variations, particularly in the northern regions, with a correlation coefficient (CC) exceeding 0.93 at a forecast horizon of 0.5 months. While both models performed similarly at the watershed level in terms of root mean square error (RMSE), the CanCM4 model displayed a higher characteristic stability index (CSI) correlation (above 0.08) than CanCM3 in diagnosing drought. Seasonal variations were evident, with better drought predictions in northern regions during spring and more noticeable model performance differences between northern and southern regions in summer. The evaluation of forecasting capability in both hindcast and forecast periods showed no significant disparity between the models, albeit the CanCM4 model exhibited superior performance in some instances. These results provide valuable insights for water resource planners, enabling more effective decision-making in drought adaptation.

1. Introduction

In recent decades, climate change and global warming have caused droughts to become more common worldwide (Jasim & Awchi, 2020; Kesravani et al., 2022). Drought is one of the most costly and unknown natural disasters, which imposes devastating consequences on agriculture, water supply, ecosystems, public health, watershed management, etc. In this regard, seasonal forecasting models have been involved in developing drought early warning systems (DEWS) and predicting hydroclimatic variables. Hence, the development of early DEWS is significant for managers and decision-makers of water resources.

A reliable forecast of extreme climate changes can lead to more effective management of watersheds in these areas, and many added socio-economic values. So far, various methods, comprising statistical and dynamic models, have been developed to predict drought. The North American multimodel ensemble (NMME) system is one of the most common dynamic models. In recent years, many studies have been conducted to evaluate NMME dynamic systems' ability to predict temperature and precipitation variables at the level of watersheds.

NMME temperature precipitation forecasts have been discussed in the northern hemisphere (Becker & van del Dole, 2016), the United States (Tian et al., 2014), East Asia (Zhou et al., 2015; Xu et al., 2018), Europe (Slater et al., 2016), East Africa (Shukla et al., 2014), South Asia (Cash et al., 2017) and Iran (Shirvani & Landman, 2016; Moghasemi et al., 2022; Yazdandoost et al., 2023).

Sheffield et al. (2014) investigated the online system for drought monitoring and seasonal forecasting of water resources and food security in Africa. In this system, microscale (downscaled) temperature and precipitation data of the CFS ver2 were applied to the predicting model for drought conditions. The results revealed the priority of this system compared to the standard statistical methods, for instance, the robustness of this system in predicting the Great drought in Africa in 2010–2011.

Shukla et al. (2019) demonstrated the limited skill of the NMME system in precipitation forecasting in East Africa, such that only in northern Ethiopia with a significant dataset, precipitation could be predicted from 2 months ago, whereas, in East Africa, these models are more reliable for temperature prediction of the precipitation.

Barbero et al. (2017) investigated the capability of the NMME system to precipitation forecast for seventeen hydroclimatological regions. The results confirmed that this system's performance depends on the geographical area of each watershed and seasonal variation.

Najafi et al. (2018) predicted the seasonal temperature in the watersheds of Iran. They applied the climatic research unit (CRU) dataset as the observational temperature into the NMME model for seasonal temperature prediction in Iran. The results indicated the priority of CSFv2 and GFDL-FLOR-BO5 models over others.

The NMME models were downscaled for precipitation and temperature datasets in the western region of the United States. The results revealed that the seasonal and geographical changes can be extracted in a downscaled form, not specified in the raw output, with a spatial resolution of 1 degree (Slater et al., 2019). Dehban et al. (2019) employed NMME models to forecast seasonal precipitation in the Sefidroud River's watershed in northwestern Iran. The results showed that due to the inadequate accuracy of the individual NMME models, the combinatorial form led to a higher accuracy by 20%. They also considered the uncertainty of the precipitation forecast and concluded that the combination of the models reduced the forecast uncertainty.

Moradian and Yazdandoost (2021) illustrated an experimental system to forecast seasonal drought based on meteorological predictions generated by NMME models. This experimental tool consists of (1) NMME and observational data, (2) post-processing techniques, especially GrandNMME and bias correction methods for statistical precipitation forecasts, (3) evaluation criteria of the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) model to select the most appropriate enhanced post-processed data, and (4) calculator of the standardized precipitation index (SPI), draws the distribution maps of the seasonal drought predictions. According to their results, in the arid regions of Iran facing a significant water crisis, including drought, the proposed NMME-based drought forecasting tool has significant skill in drought forecasting. It provides crucial information for early warnings, medium-term response planning, and preventive measures during critical times. Shirvani et al. (2023) evaluated NMME precipitation estimates using Taylor diagrams and ranked probability skill scores (RPSS) for a 29-year test period (1991–2019) over Iran. The results indicated that out of the Global Precipitation Climatology Center (GPCC) 2020 release dataset for model verification, among the

individual NMME models, generally, the Geophysical Fluid Dynamics Laboratory-Seamless System for Prediction and Earth System Research (GFDL-SPEAR) has the overall highest estimation skill.

Although so far, some studies have evaluated the capability of NMME dynamic models to forecast the values of temperature and precipitation variables at the level of the watershed in Iran (Najafi et al., 2018; Moradian & Yazdandoost, 2021; Shirvani et al., 2023), no research has been conducted to evaluate the these models' ability in predicting drought and estimate the accuracy of predictions via SPEI index.

In this regard, the performance of dynamic methods to forecast seasonal drought has been investigated in this study. For this purpose, the accuracy and precision of the NMME models to predict meteorological drought of the watersheds located in the west and southwest of Iran, comprising Great Karun, Karkheh, Zohre, Helleh, and Jarahi rivers, have been evaluated through the SPEI. The results help to determine the performance accuracy of zoned forms of the NMME models and evaluate the time series for different forecast horizons on a seasonal scale, leading to the development of high-resolution drought forecasting systems.

2. Materials and methods

2.1. The study area

In this research, the watersheds were located in the geographical range of 46° to 52° 30' east longitude and 28° to 34° north latitude, consisting of the second-degree watersheds in the west and southwest of Iran, namely Great Karun, Karkheh, Zohre, Helleh, and Jarahi rivers.

(Fig. 1). The overall area of these watersheds is 180,718 square kilometers, about 42% of the total area of the first-degree watershed of the Persian Gulf and Oman Sea.

2.2. Basic meteorological information

To determine the accuracy of drought prediction by NMME models and statistical methods, the gridded precipitation of the GPCC product (https://psl.noaa.gov/) and the minimum and maximum temperatures of the Climatic Research Unit gridded Time Series (CRU TS) product (https://www.ceda.ac.uk/) for the long-term period of 1982–2018 were used in this research.

The gridded nature of the NMME dynamic model caused to choose the gridded dataset instead of synoptic stations; meanwhile, the precipitation product of GPCC and CRU have more accurate spatial and temporal resolutions than the sparse observational stations. Figure 2 demonstrates the improper spatial distribution of the stations with complete statistics throughout the study area.

2.2.1. Global Precipitation Climatology Center (GPCC)

The Global Precipitation Climatology Center (GPCC) provides free access to the gridded precipitation data for users to analyze this gridded data accurately. This product includes monthly precipitation data on a regular grid with a spatial resolution of 2.5, 1, 0.5 arc degrees. The corresponding data were extracted from surveying systems under quality control (QC).

In this study, the latest version of GPCC (GPCC V2020) has been applied, covering the monthly period from 1901 to the end of 2020. Due to the acceptable performance of the observed precipitation, this dataset was selected and evaluated in detail in previous research (Navidi Nassaj et al., 2022).

2.2.2. Climatic research unit (CRU)

The Climatic Research Unit (CRU) is affiliated with the University of East Anglia, one of the leading institutions in studying natural and anthropogenic climate changes, which prepares and publishes several climate datasets within different resolutions. The CRU dataset has a spatial resolution of 0.5 degrees, which version 4 covers the period from 1901 to the end of 2021. A very suitable performance to estimate Iran's air temperature caused the selection of this dataset, discussed in previous studies (Navidi Nassaj et al., 2021; Jafarpour et al., 2022).

2.3. Forecast data of the NMME

In this research, the NMME outputs comprising precipitation and minimum and maximum temperatures have been involved in predicting drought. These models' datasets were provided monthly with a forecast horizon of 0.5 to 3.5 months and a spatial resolution of one degree for the statistical period of 1982 to 2018 via the portal iridl.ldeo.columbia.edu. The characteristics of these models are presented in Table 1.

Table 1
Characteristics of NMME models applied in this research

Model	Number of NMME model	Forecast horizon (month)	Reference
CMC1-CanCM3	10	11	Merryfield et al. (2013)
CMC2-CanCM4	10	11	Merryfield et al. (2013)

It should be noted that all datasets were reproduced into GPCC spatial specifications using the re-grid technique to establish the coincidence among the spatial accuracy of gridded datasets and NMME models.

2.4. Standardized precipitation evapotranspiration index (SPEI)

The standardized precipitation evapotranspiration index (SPEI) applies the simple water balance equation, the difference between precipitation and potential evapotranspiration, based on the Torrent White (TW) equation for different time scales (more details in Vicente-Serrano et al. 2010). This well-known index has been used in different time scales in this study. A 3-month scale of the SPEI denotes the medium-term droughts (Lotfirad et al., 2022; Navidi Nassaj et al., 2021; Hao & Aghakouchak, 2014).

2.5. Statistical criteria

In this study, a wide range of statistical indices, including correlation coefficient (CC), root mean square error (RMSE), percent bias error (PBIAS), as well as the consensus table indices comprising probability of detection (POD), false alarm ratio (FAR), and characteristic stability index (CSI) have been incorporated to evaluate the potential of precipitation and drought. The corresponding formula has been illustrated in Table 2. In these equations, M and RG denote calculated precipitation in the model and GPCC data, respectively; also, $\bar{M}$ and $\bar{RG}$ indicate the corresponding average.

Hits is the number of precipitation events that both statistical databases. Misses denote the number of precipitation events that occurred while the model did not announce. In addition, FalseAlarm represents the number of precipitation events announced by the model rainfall but not happened in reality.

Table 2
Statistical criteria

Evaluation index	Equation	Maximum value
POD	$\frac{Hits}{Hits+\mathrm{Misses}}$	1
FAR	$\frac{False\ Alarms}{Hits+False\ Alarms}$	0
CC	$\frac{\sum_{i=1}^{n}(\mathrm{M}_i-\bar{\mathrm{M}})(\mathrm{RG}_i-\bar{\mathrm{RG}})}{\sqrt{\sum_{i=1}^{n}(\mathrm{M}_i-\bar{\mathrm{M}})^2\sum_{i=1}^{n}(\mathrm{RG}_i-\bar{\mathrm{RG}})^2}}$	1
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{M}_i-\mathrm{RG}_i)^2}$	0
PBIAS	$\frac{\sum_{i=1}^{n}(\mathrm{M}_i-\mathrm{RG}_i)}{\sum_{i=1}^{n}(\mathrm{RG}_i)}$	0
NSE	$1-\frac{\sum_{i=1}^{n}(M_i-RG_i)^2}{\sum_{i=1}^{n}(M_i-\bar{\mathrm{M}})^2}$	1
MIA	$1-\frac{\sum_{i=1}^{n}(M_i-RG_i)^2}{\sum_{i=1}^{n}(\lvert RG_i-\bar{\mathrm{M}}\rvert+\lvert M_i-\bar{\mathrm{M}}\rvert)^2}$	1

In the following, the GPCC gridded dataset has been evaluated to use as a basis for comparison (in terms of similarity with station values); in this regard, the ability of NMME models to forecast precipitation and then meteorological drought was investigated via the SPEI. Due to the conditional nature of the precipitation variable, an accurate estimation of this variable is very important. Meanwhile, the estimation

of minimum and maximum temperature variables by the CRU dataset confirmed that the minimum CC and maximum RMSE values were 0.96 and 3°C, respectively, indicating the high accuracy of this dataset for temperature estimation.

3. Results and Discussion

3.1. Evaluating the capability of gridded datasets to estimate the observational precipitation

Figure 3 demonstrates the performance of GPCC to estimate observational precipitation in the study area. A high correlation (CC > 0.6) between the ground-truth (station) data and the GPCC product was observed in most parts of the watershed, especially the Karkheh River. The lowest CC was observed in the southern parts of the Great Karun watershed. Meanwhile, in most parts of the study areas, in terms of RMSE, the estimation error of the GPCC dataset for precipitation is less than 60 mm per year.

Except for small parts of the south of the study area with underestimation of the observed precipitation, within other southern parts of the watershed, the PBIAS always varied between 0 and 20 mm per year, denoting the high accuracy of GPCC.

In the northern part, the GPCC overestimated the observed precipitation (always less than 50 mm per year). To more accurately evaluate the performance of GPCC, the results were obtained based on Nash-Sutcliffe efficiency (NSE) and modified index of agreement (MIA) (Fig. 3). The NSE of the GPCC product is positive (more than 0.6 for the northern, eastern, and southern parts) throughout the study area, demonstrating this model's acceptable performance. In addition, the MIA denotes suitable agreement between GPCC data and the station observations.

3.2. Drought prediction based on SPEI

To predict drought, SPEI has been employed in a period of 3 months and two short-term horizons, including 0.5 months and seasonal (3.5 months). The SPEI of NMME models versus the observed values has been depicted for the forecast horizon of 0.5 months (Fig. 4) and 3.5 months (Fig. 5).

Figures 4 and 5 illustrated that for both CanCm3 and CanCM4 models, the seasonal variation of SPEI was well identified by NMME models in the forecast horizon of 0.5 months. However, some errors were observed in forecast period ranges (such as the spring season of 2011). In the forecast horizon of 3.5 months, particularly in the forecast period, both models have predicted the occurrence of droughts a little earlier than the real.

Although the performance of CanCM3 and CanCM4 is very similar, both models have generally not performed well to detect severe droughts. In order to evaluate the spatial accuracy of NMME models for drought detection, zoning maps of the CC, RMSE, and CSI have been presented in Figs. 6 and 7.

Accordingly, in the forecast horizon of 0.5 months, both models have a robust correlation with SPEI observations, so for both models, the correlation coefficient in the northern parts of the watershed is more than 0.93, which indicates a satisfactory agreement with the observed values.

In terms of RMSE, the performance of the models is at an optimal level (a maximum value of 0.47 occurred in minimal parts of the watershed, while the best value happened within the north of the watershed (Karkheh River watershed). In general, the CanCM3 slightly outperformed the CanCM4 in terms of the RMSE model; in terms of the CSI, the priority of the CanCM4 model was observed (Fig. 6), so that almost within the entire watershed, except for small parts in the south, the CSI was above 80%, denoting a very high power of this model for drought detection, it should be mentioned that in terms of CSI, a relatively suitable performance of the CanCM3 was observed.

Figure 6. Spatial performance of NMME models to detect SPEI for a forecast horizon of 0.5 months
As indicated in Fig. 7, the accuracy of the models was significantly reduced by increasing the forecast horizon to 3.5 months. In this forecast horizon, unlike the horizon of 0.5 months, which showed the best correlation in the northern areas, the best performance was achieved in the southern parts of the watershed with a CC of 0.3 to 0.45, while the CC less than 0.1 for northern, and a weak CC of 0.2 to 0.3 for the central regions.

Regarding RMSE, the southern parts have an error of less than one, representing a desirable condition. Based on the CSI, it is revealed that in this forecast horizon, the highest value of CSI was obtained at 0.3 for the southern regions. In contrast, the northern parts had a CSI of less than 0.2.

Figure 7. Spatial performance of NMME models to detect SPEI for the forecast horizon of 3.5 months
Figures 8 to 15 present a schematic illustration of the models' accuracy on a seasonal scale in terms of the CC and RMSE for different seasons. In the winter season, the correlation of both models, especially in the forecast horizon of 0.5 months, has been very similar, so the southwestern parts of the watershed have the best situation with a CC of 0.6 to 0.7, while the eastern and northern parts with a CC of 0.4 to 0.5 had the worst situation.

Figure 8. Seasonal performance of NMME models to detect SPEI for a forecast horizon of 0.5 months in winter
Figure 9. Seasonal performance of NMME models to detect SPEI for the forecast horizon of 3.5 months (at a 3.5-month lead time) in winter
In spring, there is a powerful CC in both the CanCM3 and CanCM4 models (CC is always above 0.8). The best performance was assigned to the CanCM3 model for the north of the study area (CC is more than 92%). Regarding RMSE, for both models, the north of the watershed has outperformed the southern parts, so the RMSE in the northern parts varies from 0.13 to 1.5. Contrary to the winter, there is a desirable CC throughout the watershed at a 3.5-month lead time. For example, both models had a CC of more than 0.7 in the southwestern parts of the watershed. This result denoted the high accuracy of these models for predicting drought in the spring season. In addition, RMSE is desirable, so that it reached 1.8 in the worst case.

Figure 10. Seasonal performance of NMME models to detect SPEI at a 0.5-month lead time in spring
Figure 11. Seasonal performance of NMME models to detect SPEI at a 3.5-month lead time in spring
In summer, at a 0.5-month lead time, there is a significant difference between the CC for different parts of the watershed, so that for the northern and southern parts, the CC was above 0.5 and around 0.1 to 0.2, respectively. In the forecast horizon of 3.5 months, however, the CC is very weak, so it reached 0.25 in the best situation. Although RMSE was more than 0.5 months in the forecast horizon of 3.5 months, it did not exceed 1.8.

Figure 12. Seasonal performance of NMME models to detect SPEI for the forecast horizon of 0.5 months in summer

Figure 13. Seasonal performance of NMME models to detect SPEI for the forecast horizon of 3.5 months in summer
Similar to the spring season, in the autumn, a solid correlation was established in both models (CC of 0.93 to 0.96 in Most parts of the watershed); in addition, the RMSE was much lower compared to other seasons of the year, such that within almost all the watershed, it is in the range of 0.1 to 0.25, which confirmed the high accuracy of the models during this season. However, in the forecast horizon of 3.5 months, the RMSE has increased while the CC has decreased.

In this forecast horizon, the best performance of the CanCM3 was obtained for the northern, northeastern, and eastern parts of the study area (CC more than 0.5), whereas, for the CanCM4, the highest CC was acquired at 0.5 in the northern part.

Figure 14. Seasonal performance of NMME models to detect SPEI for the forecast horizon of 0.5 months in autumn

Figure 15. Seasonal performance of NMME models to detect SPEI for the forecast horizon of 3.5 months in autumn

For the horizon of 0.5 months, the results of the CSI were presented for two seasons, summer and autumn, to evaluate the ability of models for deterministic diagnosis of the drought. It should be noted that for the forecast horizon of 3.5 months in other seasons, it was impossible to calculate CSI due to its calculation relationship and the evaluated time series because, in some regions, an unlimited CSI was defined. Figure 16 denotes the priority of the CanCM4 model during the autumn so that in most parts of the watershed, CSI was more than 0.95. Meanwhile, the CanCM3 model had a significant performance this season.

In the summer, CanCM4 has outperformed (CSI more than 0.9) than CanCM3 (the maximum CSI of 0.65) within the entire watershed.

Figure 16. The CSI of NMME models in two seasons of autumn and summer for the forecast horizon of 0.5 months (Note: The range of the map legend is different)

To summarize the results, Table 3 shows the statistical criteria per regional average. It can be recognized that the performance of both models was very similar. Also, regarding CC, the best performance has always occurred in the spring.

Table 3
Seasonal performance indices of NMME models per regional average

Model	Seasonal period	CC	RMSE	CSI
0.5-month lead time				
Cancm3	Overall	0.92	0.3	0.81
	Winter	0.56	0.11	-
	Spring	0.91	0.19	0.25
	Summer	0.41	0.67	0.95
	Autumn	0.93	0.23	0.43
Cancm4	Overall	0.91	0.39	0.84
	Winter	0.54	0.1	-
	Spring	0.9	0.19	0.14
	Summer	0.34	0.7	0.94
	Autumn	0.93	0.23	0.58
3.5-month lead time				
Cancm3	Overall	0.13	1.3	0.16
	Winter	0.19	0.46	-
	Spring	0.61	1.7	0.03
	Summer	-0.3	1.5	0.39
	Autumn	0.36	0.9	0
Cancm4	Overall	0.18	1.2	0.2
	Winter	0.23	0.47	-
	Spring	0.67	1.6	0.04
	Summer	-0.2	1.6	0.36
	Autumn	0.05	0.96	0

In evaluating the NMME's performance to forecast drought, a central point is the corresponding performance during hindcast and forecast periods. In this regard, Table 4 shows the results of `performance indices for the periods of hindcast and forecast, denoting no considerable difference between these periods. Also, both models have very close performance, such that, in the prediction horizon of 0.5 months, the CC of both models was 0.89. This result revealed that further investigations only on the overall period could be sufficient, and there is no need to divide the evaluation period into two parts: hindcast and forecast.

Table 4
The performance of NMME models for the hindcast and forecast periods separately

0.5-Month lead time						
Indices	CC	RMSE	CSI	CC	RMSE	CSI
Model	Hindcast			Forecast		
CanCM3	0.93	0.34	0.84	0.89	0.46	0.72
CanCM4	0.92	0.38	0.85	0.89	0.43	0.8
3.5-Month lead time						
Indices	CC	RMSE	CSI	CC	RMSE	CSI
Model	Hindcast			Forecast		
CanCM3	0.15	1.3	0.17	0.08	1.3	0.13
CanCM4	0.19	1.26	0.26	0.11	1.27	0.15

To more accurately evaluate the performance of the models, mainly based on the seasonal performance, Taylor diagrams have been presented in Fig. 17.

Figure 17. Taylor diagram of the NMME models' performance to detect observational SPEI

Figure 18. Taylor diagram of the NMME models' seasonal performance to detect observational SPEI

4. Conclusion

Regarding the arid and semi-arid climate of Iran, reliable drought forecasting is necessary; meanwhile, limited studies have been conducted on drought index evaluation based on the NMME models' output for precipitation and temperature forecasting; hence, in this research, these models were investigated in the study area, including the watersheds of Great Karun, Karkheh, Helleh, and Handijan-Jarahi rivers for a long-term period of 1982–2018.

Considering the added value that zoned drought forecasting provides, in this study, to validate the NMME models' predictions, GPCC gridded precipitation datasets and CRU temperature were incorporated as the benchmark data.

At the beginning of this research, the capability of these two datasets was evaluated as an alternative to station observations to ensure the accuracy of these surrogate data to observational ones. The results confirmed a very appropriate accuracy.

In terms of the SPEI, the results indicated that both CanCM3 and CanCM4 models have well reflected the seasonal changes of drought so that a robust correlation with observational data was obtained in the forecast horizon of 0.5 months (CC above 0.93 for northern parts of the watershed).

Relatively similar performance of both models was observed without much error in terms of RMSE. The CanCM4 model outperformed the CanCM4 (CSI always more than 0.80). According to the SPEI, the accuracy of the models sharply decreases with the increase of the forecast horizon. In the forecast horizon of 3.5 months, the best performance was assigned to the southern parts of the watershed.

Seasonal evaluations denoted a similar performance of CanCM3 and CanCM4 models for drought prediction in winter, such that the highest and lowest CC were acquired for the southwest and northeast of the watershed, respectively. In the spring season, a robust correlation (CC over 0.8) was obtained for both models, in which the north of the watershed showed the highest value. In general, in the spring, a more reliable drought prediction was conducted for the northern than the southern parts of the watershed. In summer, however, the difference between the performance of the models in the northern and southern parts is much more significant, so the correlation in the south and the north of the watershed was about 0.5 and 0.1, respectively.

In autumn, high accuracy was established by both models (CC ~ 0.96); meanwhile, the RMSE was the least among all seasons. Also, the models provided a desirable performance at a 3.5-month lead time. Regarding the ability to detect drought events, CanCM4 has outperformed CanCM3 (CSI above 0.9); generally, the results demonstrated that the models have the highest agreement with the observational data in spring and autumn. The evaluation of the predictive power in two hindcast and forecast time scales also indicated no significant difference between the performance of the models for these time scales.

The results obtained from this research provided a comprehensive outlook of the performance of NMME models' performance in predicting drought on a seasonal scale.

These results could be valuable for water resources planners to develop drought forecasting systems. In addition, in the forecast horizon of 0.5 months, NMME can provide a reliable forecast. Since the variable of the time is essential in drought management, this can lead to accurate management decisions to adapt to drought and develop appropriate response plans to drought conditions.

Declarations

Ethical Approval

The manuscript is an original work with its own merit, has not been previously published in whole or in part, and is not being considered for publication elsewhere.

Consent to Participate

Mehdi Moghasemi, Narges Zohrabi, Hossein Fathian, Alireza Nikbakht Shahbazi and Mohammadreza Yeganegi have read the final manuscript, have approved the submission to the journal, and have accepted full responsibilities pertaining to the manuscript's delivery and contents.

Consent to Publish

Mehdi Moghasemi, Narges Zohrabi, Hossein Fathian, Alireza Nikbakht Shahbazi and Mohammadreza Yeganegi agree to publish this manuscript upon acceptance.

Authors Contributions Mehdi Moghasemi, Narges Zohrabi, Hossein Fathian, Alireza Nikbakht Shahbazi and Mohammadreza Yeganegi declare that they have contributed to the preparation of this manuscript.

Funding

The authors did not receive support from any organization for the submitted work.

Competing Interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and materials

All data, models, and code are available from the corresponding author by request.

References

1. Barbero, R., Abatzoglou, J. T., & Hegewisch, K. C. (2017). Evaluation of statistical downscaling of North American multimodel ensemble forecasts over the western United States. *Weather and Forecasting*, *32*(1), 327-341.
2. Becker, E. J., & Van den Dool, H. M. (2018, December). Assessing Probabilistic Forecasts of S2S Climate Extremes Using the NMME. In AGU Fall Meeting Abstracts (Vol. 2018, pp. A33E-05B).
3. Cash, B. A., Manganello, J. V., & Kinter, J. L. (2019). Evaluation of NMME temperature and precipitation bias and forecast skill for South Asia. *Climate dynamics*, *53*, 7363-7380.
4. Dehban, H., Ebrahimi, K., Araghinejad, S., & Bazrafshan, J. (2019). Evaluation of NMME models in forecasting of monthly rainfall (Case study: Sefidrood Basin). Journal of Agricultural Meteorology,

7(1), 3-12.

5. Hao, Z., AghaKouchak, A., Nakhjiri, N., & Farahmand, A. (2014). Global integrated drought monitoring and prediction system. *Scientific data*, *1*(1), 1-10.
6. Jafarpour, M., Adib, A., & Lotfirad, M. (2022). Improving the accuracy of satellite and reanalysis precipitation data by their ensemble usage. Applied Water Science, 12(9), 232.
7. Jasim, Ansam I., and Taymoor A. Awchi. "Regional meteorological drought assessment in Iraq." Arabian Journal of Geosciences 13, no. 7 (2020). doi:10.1007/s12517-020-5234-y.
8. Lotfirad, M., Esmaeili-Gisavandani, H., & Adib, A. (2022). Drought monitoring and prediction using SPI, SPEI, and random forest model in various climates of Iran. Journal of Water and Climate Change, 13(2), 383-406.
9. Moghasemi, M., Zohrabi, N., Fathian, H., Nikbakht Shahbazi, A., & Yeganegi, M. (2022). Drought Prediction Using North American Multimodel Ensemble (NMME) Over Western Regions of Iran. Journal of Water and Soil Resources Conservation, 11(4), 25-40.
10. Moradian, S., & Yazdandoost, F. (2021). Seasonal meteorological drought projections over Iran using the NMME data. Natural Hazards, 108, 1089-1107.
11. Najafi, H., Massah Bavani, A. R., Irannejad, P., & Robertson, A. W. (2018). Evaluation of NMME seasonal temperature forecasts over Iran's river basins. Journal of Agricultural Meteorology, 6(1), 19-30.
12. Nassaj, B. N., Zohrabi, N., Shahbazi, A. N., & Fathian, H. (2022). Evaluating the performance of eight global gridded precipitation datasets across Iran. *Dynamics of Atmospheres and Oceans*, *98*, 101297.
13. Navidi Nassaj, Behzad, Narges Zohrabi, Alireza Nikbakht Shahbazi, and Hossein Fathian. "Evaluation of three gauge-based global gridded precipitation datasets for drought monitoring over Iran." Hydrological Sciences Journal 66, no. 14 (2021), 2033-2045. doi:10.1080/02626667.2021.1978444.
14. Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., ... & Ogallo, L. (2014). A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bulletin of the American Meteorological Society*, *95*(6), 861-882.
15. Shirvani, A., & Landman, W. A. (2016). Seasonal precipitation forecast skill over Iran. *International Journal of Climatology*, *36*(4), 1887-1900.
16. Shirvani, A., Landman, W. A., Barlow, M., & Hoell, A. (2023). Evaluation of the forecast skill of North American Multimodel Ensemble for monthly and seasonal precipitation forecasts over Iran. International Journal of Climatology, 43(2), 1141-1166.
17. Shukla, S., McNally, A., Husak, G., & Funk, C. (2014). A seasonal agricultural drought forecast system for food-insecure regions of East Africa. Hydrology and Earth System Sciences, 18(10), 3907-3921.
18. Shukla, S., Roberts, J., Hoell, A., Funk, C. C., Robertson, F., & Kirtman, B. (2019). Assessing North American multimodel ensemble (NMME) seasonal forecast skill to assist in the early warning of anomalous hydrometeorological events over East Africa. *Climate Dynamics*, *53*, 7411-7427.

19. Slater, L. J., Villarini, G., & Bradley, A. A. (2017). Weighting of NMME temperature and precipitation forecasts across Europe. Journal of Hydrology, 552, 646-659
20. Slater, L. J., Villarini, G., & Bradley, A. A. (2019). Evaluation of the skill of North-American Multimodel Ensemble (NMME) Global Climate Models in predicting average and extreme precipitation and temperature over the continental USA. Climate dynamics, 53, 7381-7396.
21. Tian, D., Martinez, C. J., Graham, W. D., & Hwang, S. (2014). Statistical downscaling multimodel forecasts for seasonal precipitation and surface temperature over the southeastern United States. Journal of Climate, 27(22), 8384-8411.
22. Xu, L., Chen, N., Zhang, X., & Chen, Z. (2018). An evaluation of statistical, NMME and hybrid models for drought prediction in China. Journal of hydrology, 566, 235-249.
23. Yazdandoost, F., Zalipour, M., & Izadi, A. (2023). Statistical refinement of the North American multimodel ensemble precipitation forecasts over Karoon basin, Iran. Journal of Water and Climate Change.

Figures

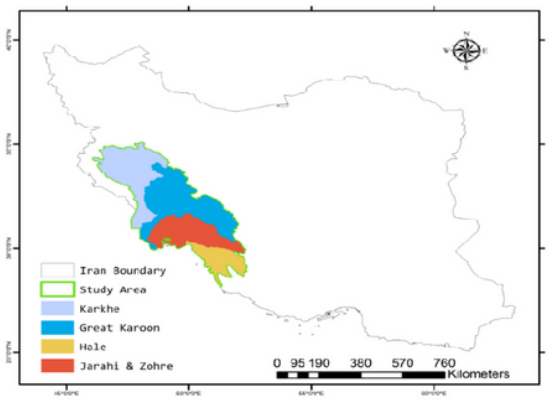


Figure 1. The study area

Figure 1

See image above for figure legend.

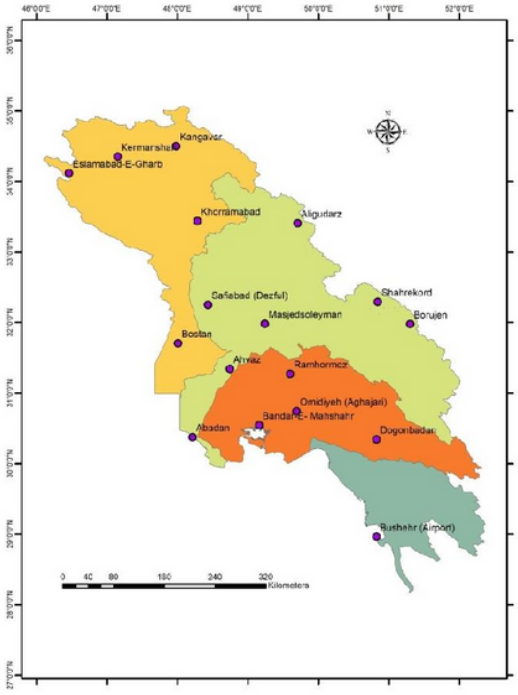


Figure 2

Synoptic meteorological stations with long-term statistics in the study area

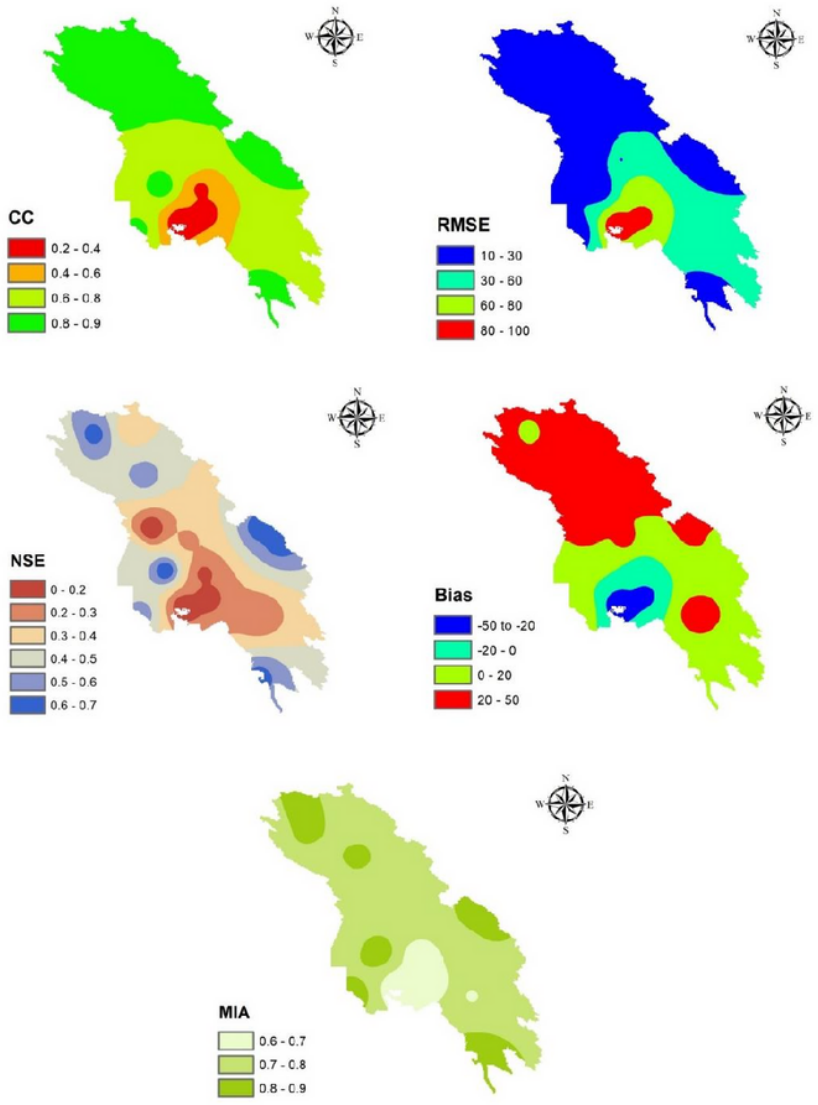


Figure 3

The spatial accuracy of GPCC precipitation versus synoptic station data is based on different indices comprising CC, RMSE, PBIAS, NSE, and MIA.

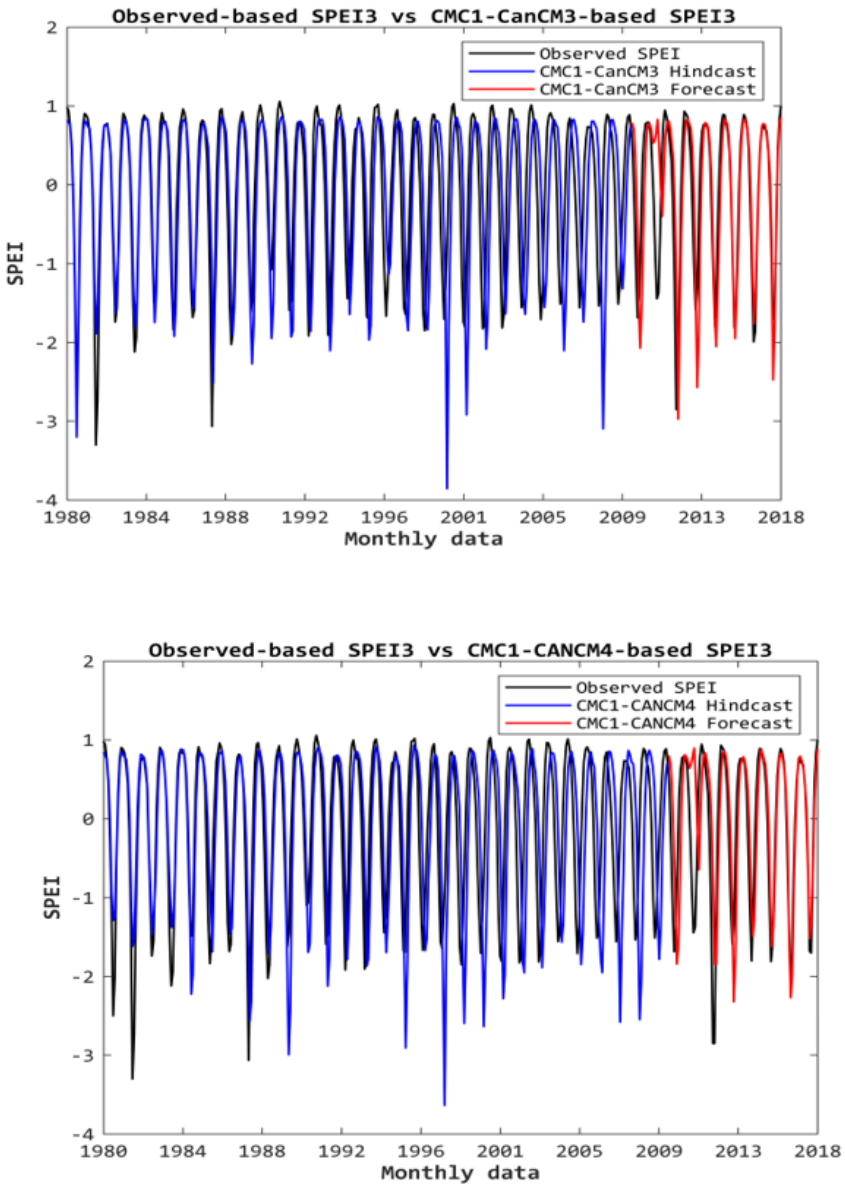
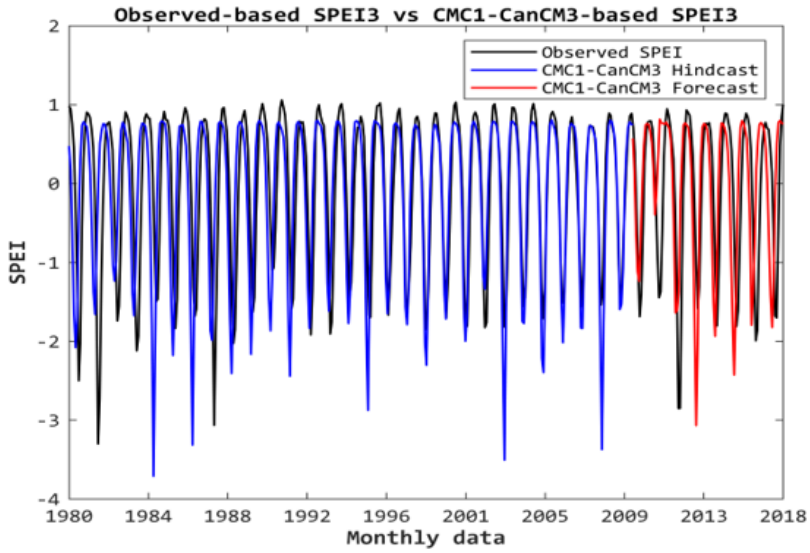


Figure 4

Seasonal variation of SPEI for two periods of hindcast and forecast in the horizon of 0.5 months



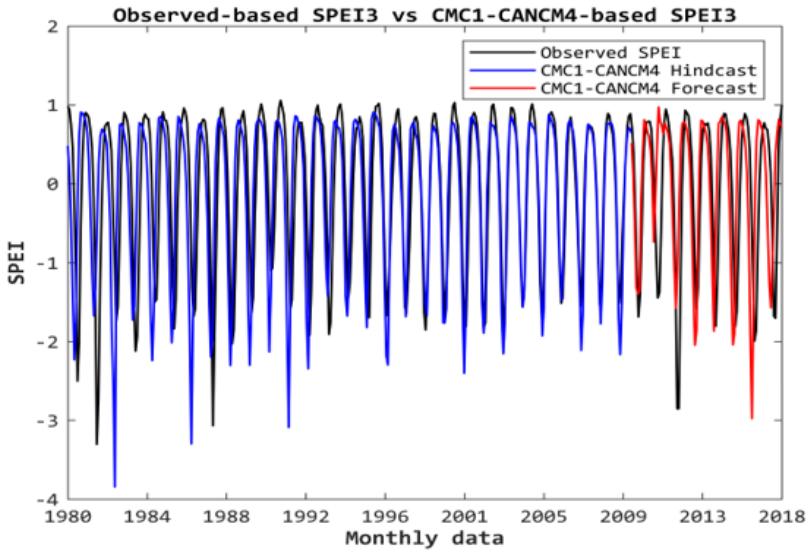


Figure 5

Seasonal variation of SPEI for two periods of hindcast and forecast in the horizon of 3.5 months

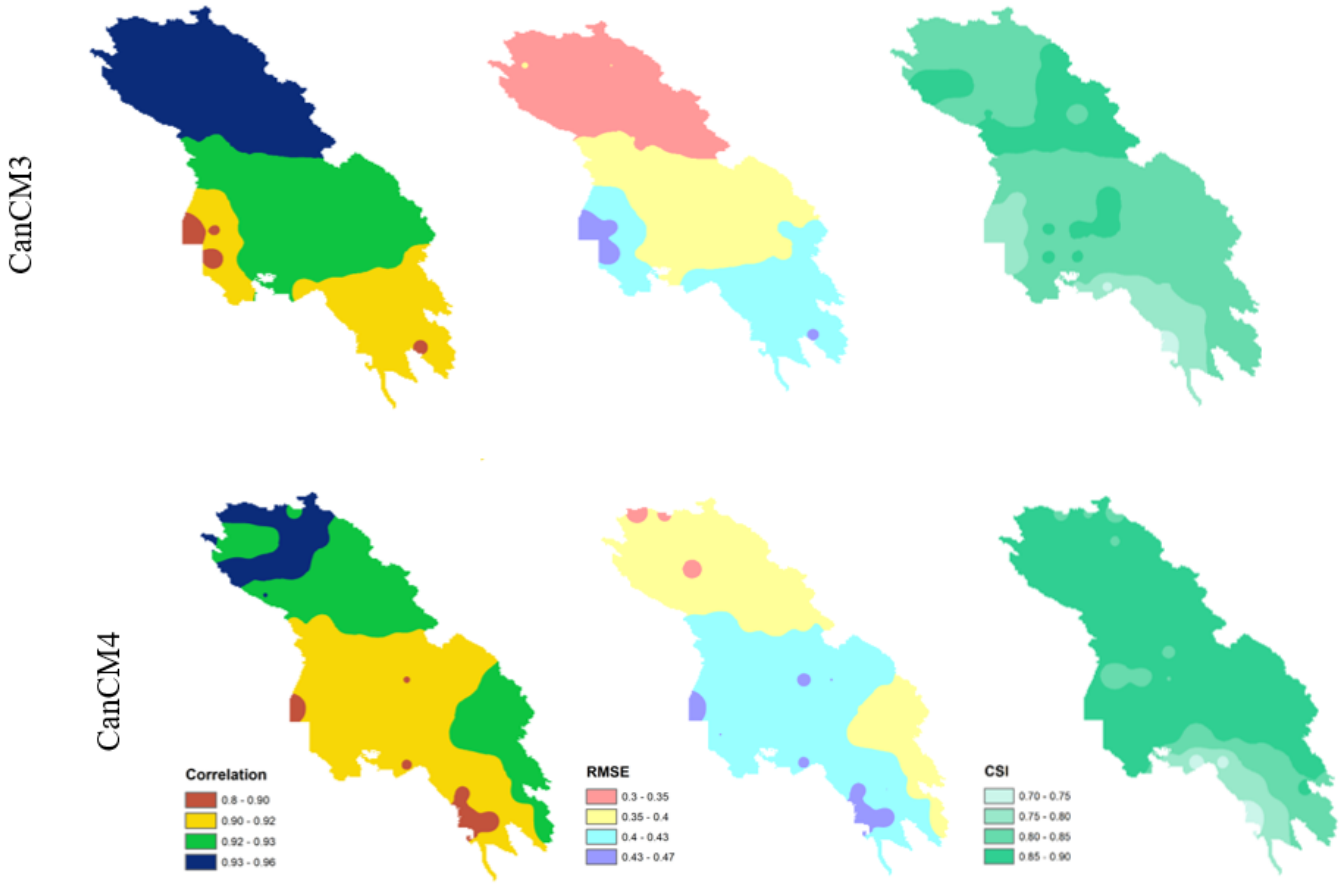


Figure 6

Spatial performance of NMME models to detect SPEI for a forecast horizon of 0.5 months

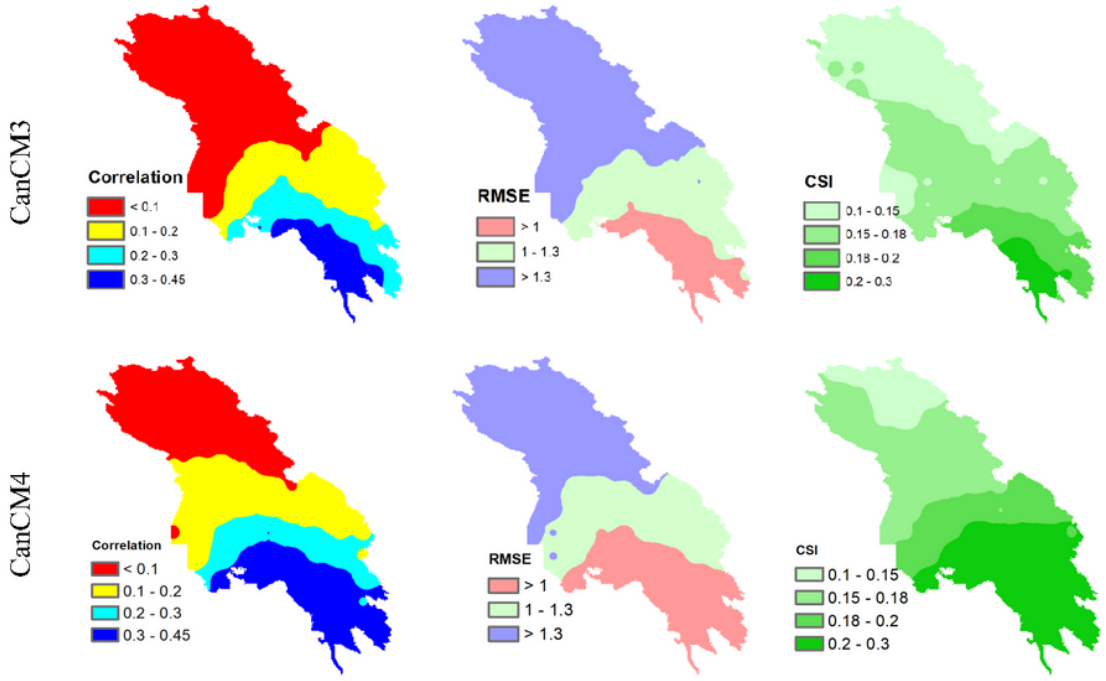


Figure 7

Spatial performance of NMME models to detect SPEI for the forecast horizon of 3.5 months

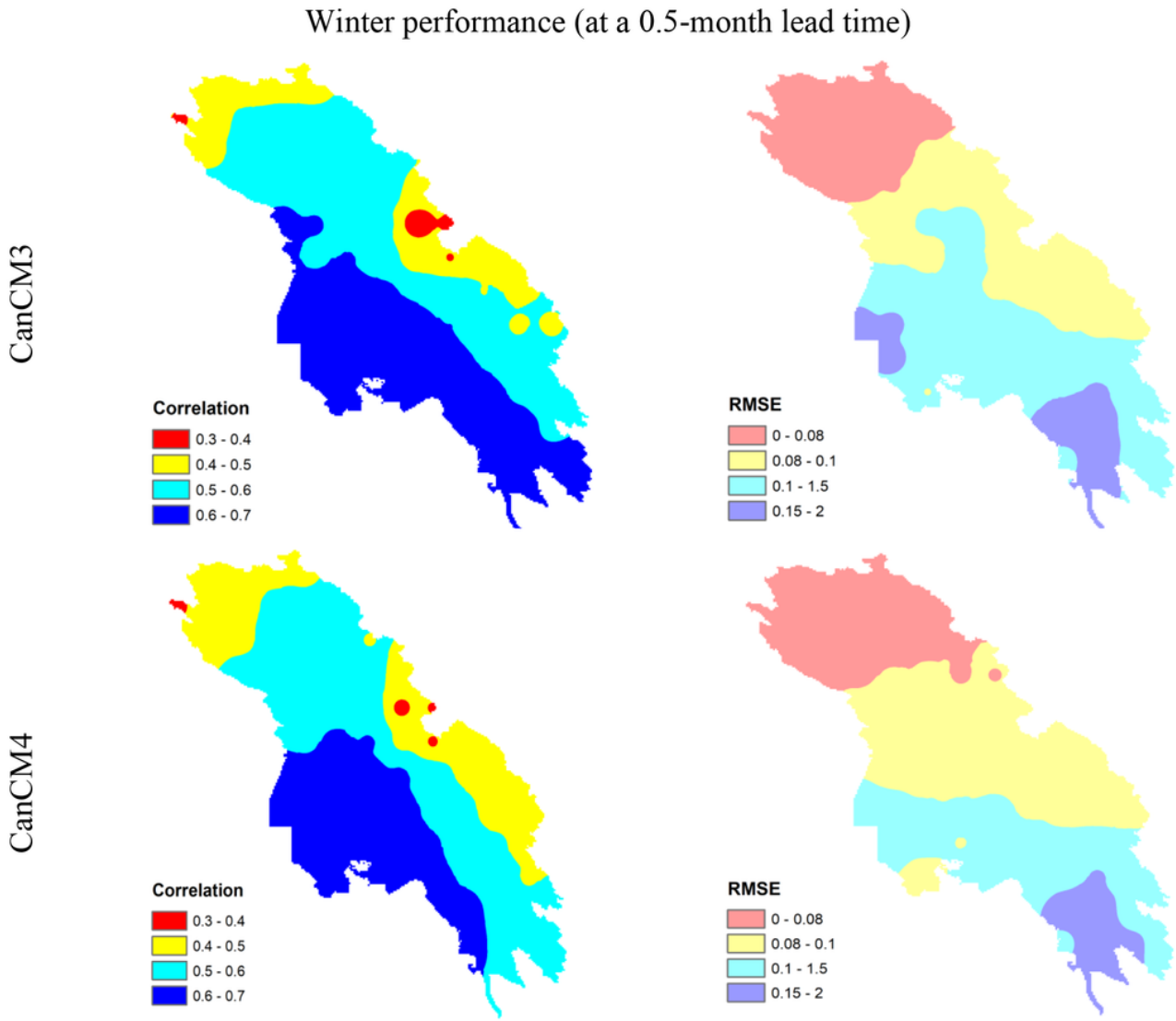


Figure 8

Seasonal performance of NMME models to detect SPEI for a forecast horizon of 0.5 months in winter

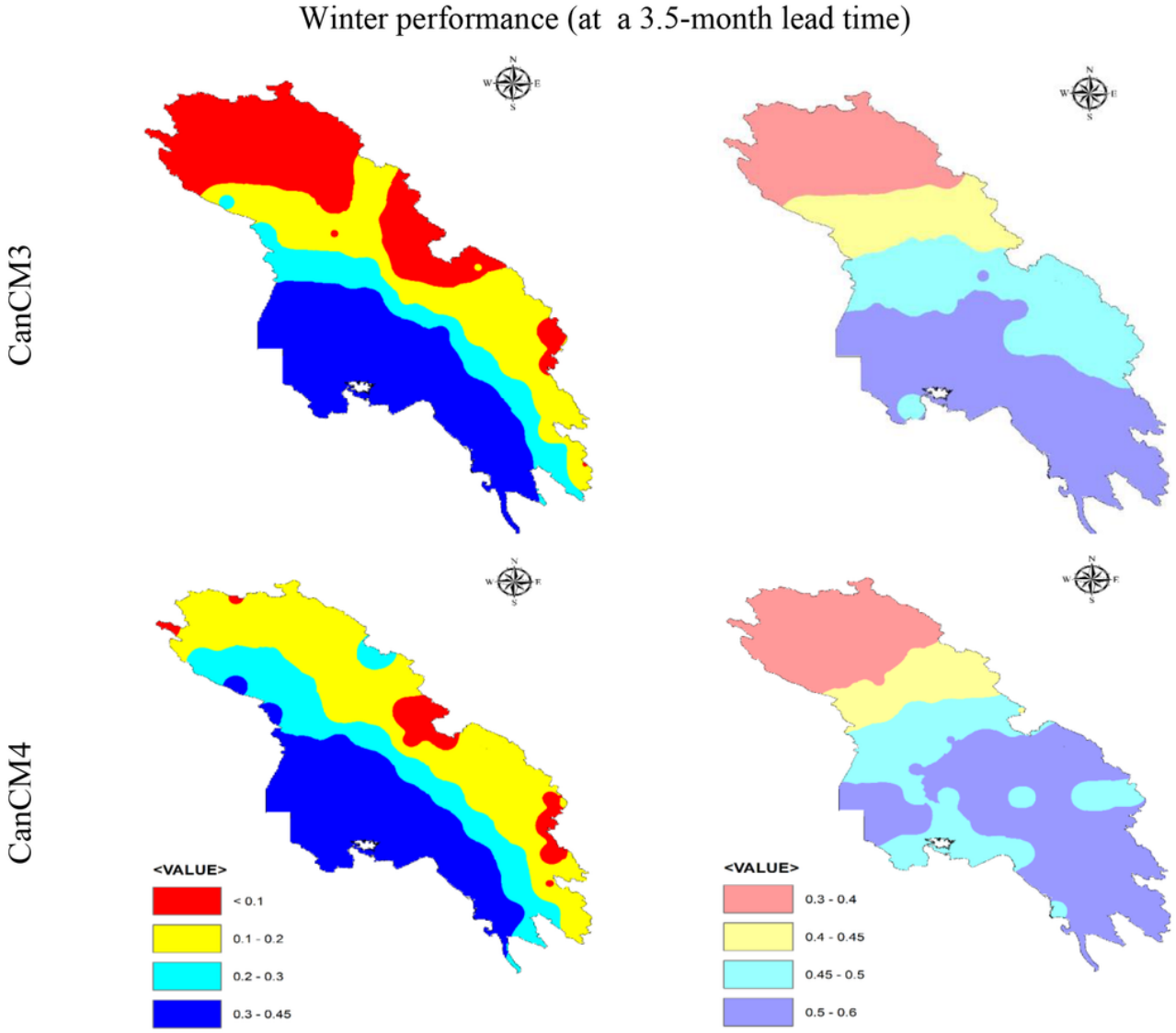


Figure 9

Seasonal performance of NMME models to detect SPEI for the forecast horizon of 3.5 months (at a 3.5-month lead time) in winter

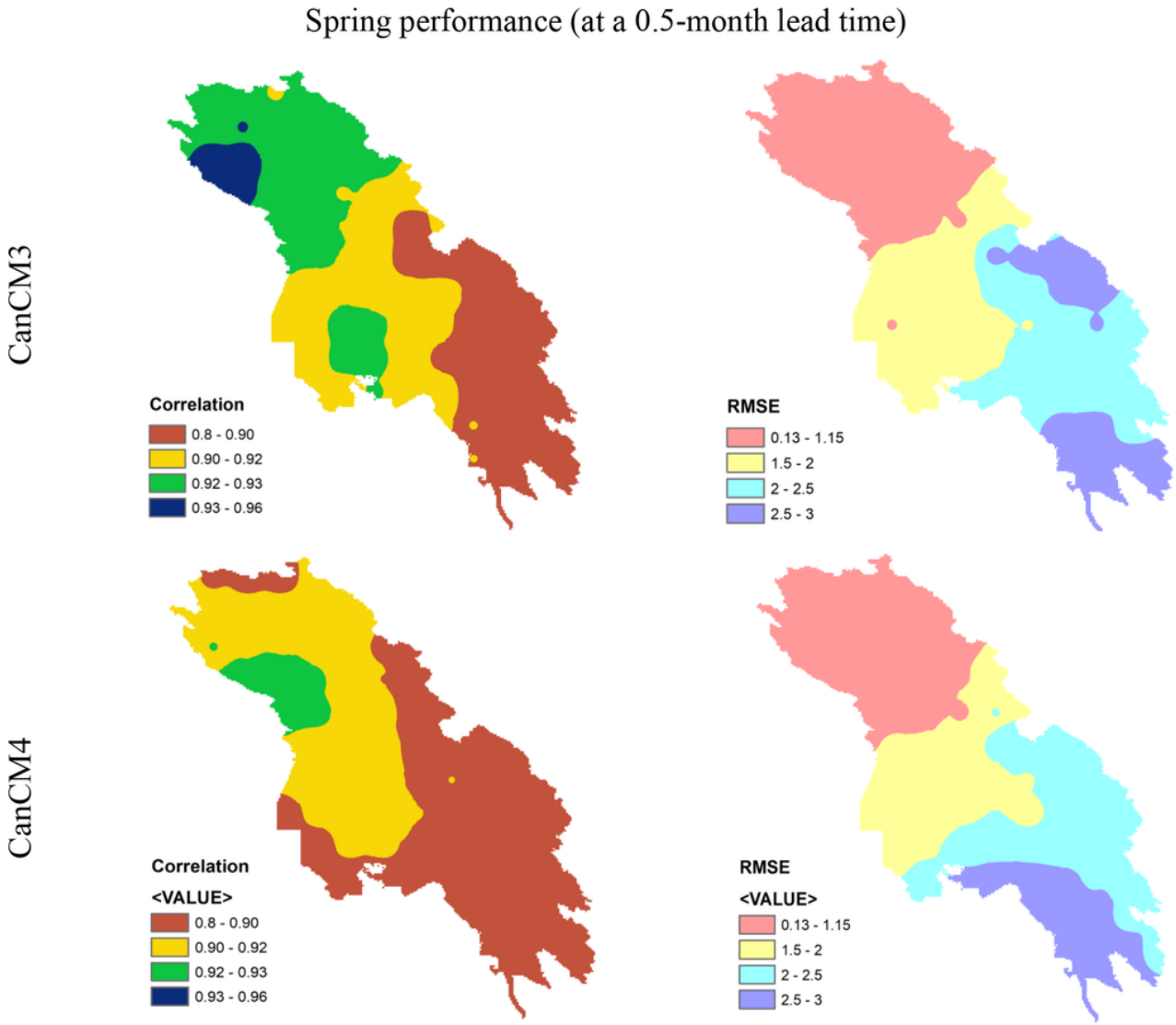


Figure 10

Seasonal performance of NMME models to detect SPEI at a 0.5-month lead time in spring

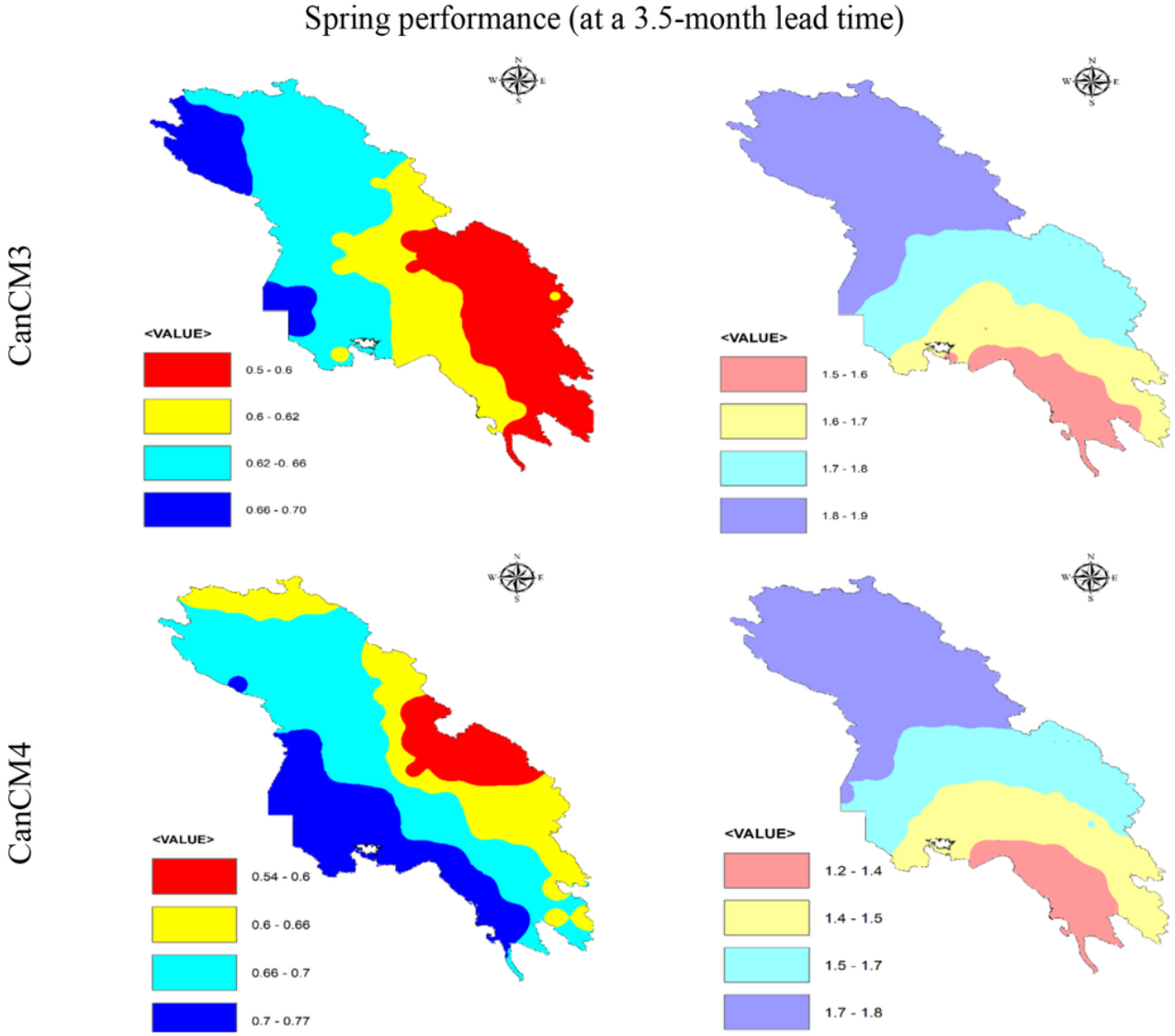


Figure 11

Seasonal performance of NMME models to detect SPEI at a 3.5-month lead time in spring

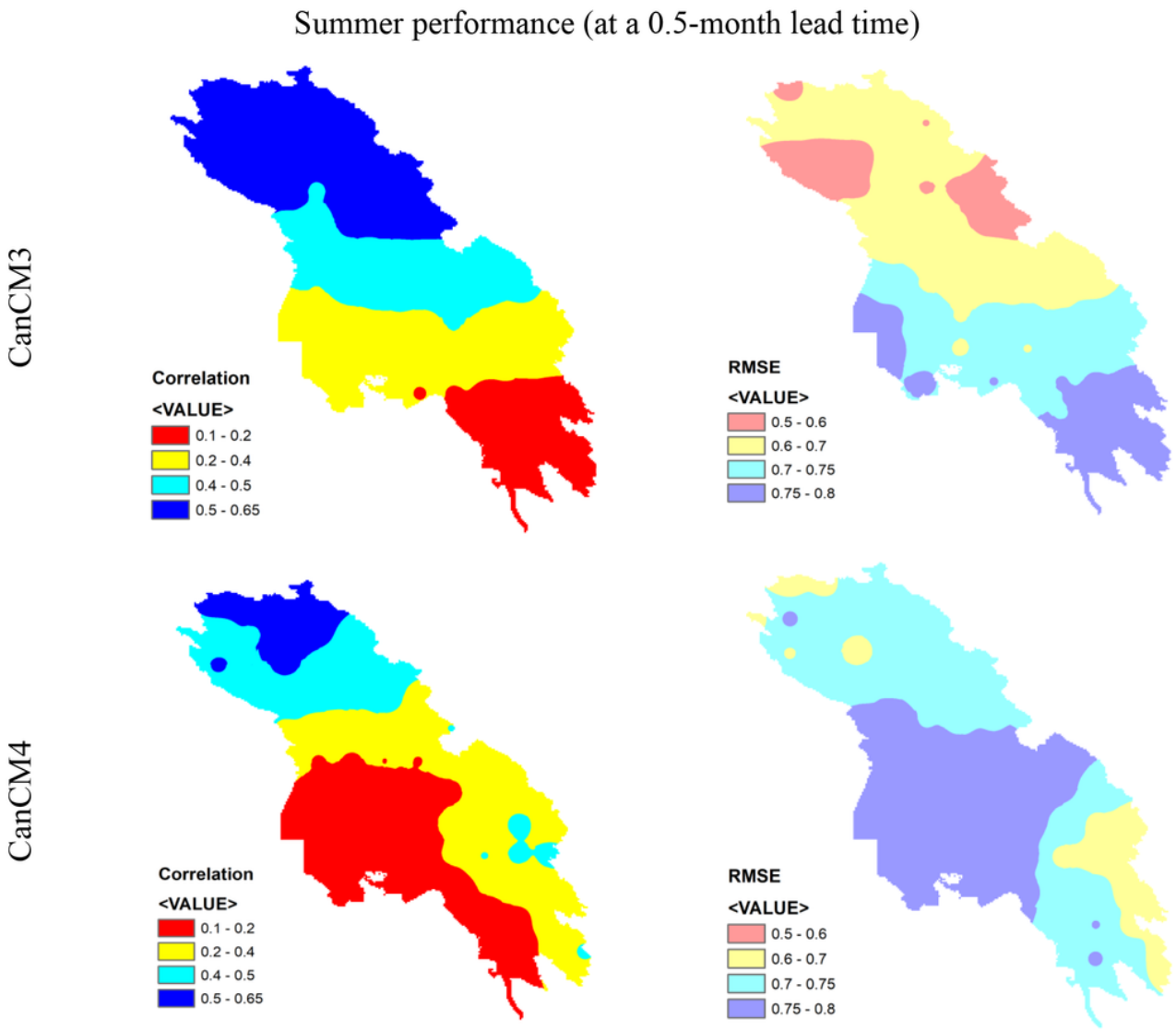


Figure 12

Seasonal performance of NMME models to detect SPEI for the forecast horizon of 0.5 months in summer

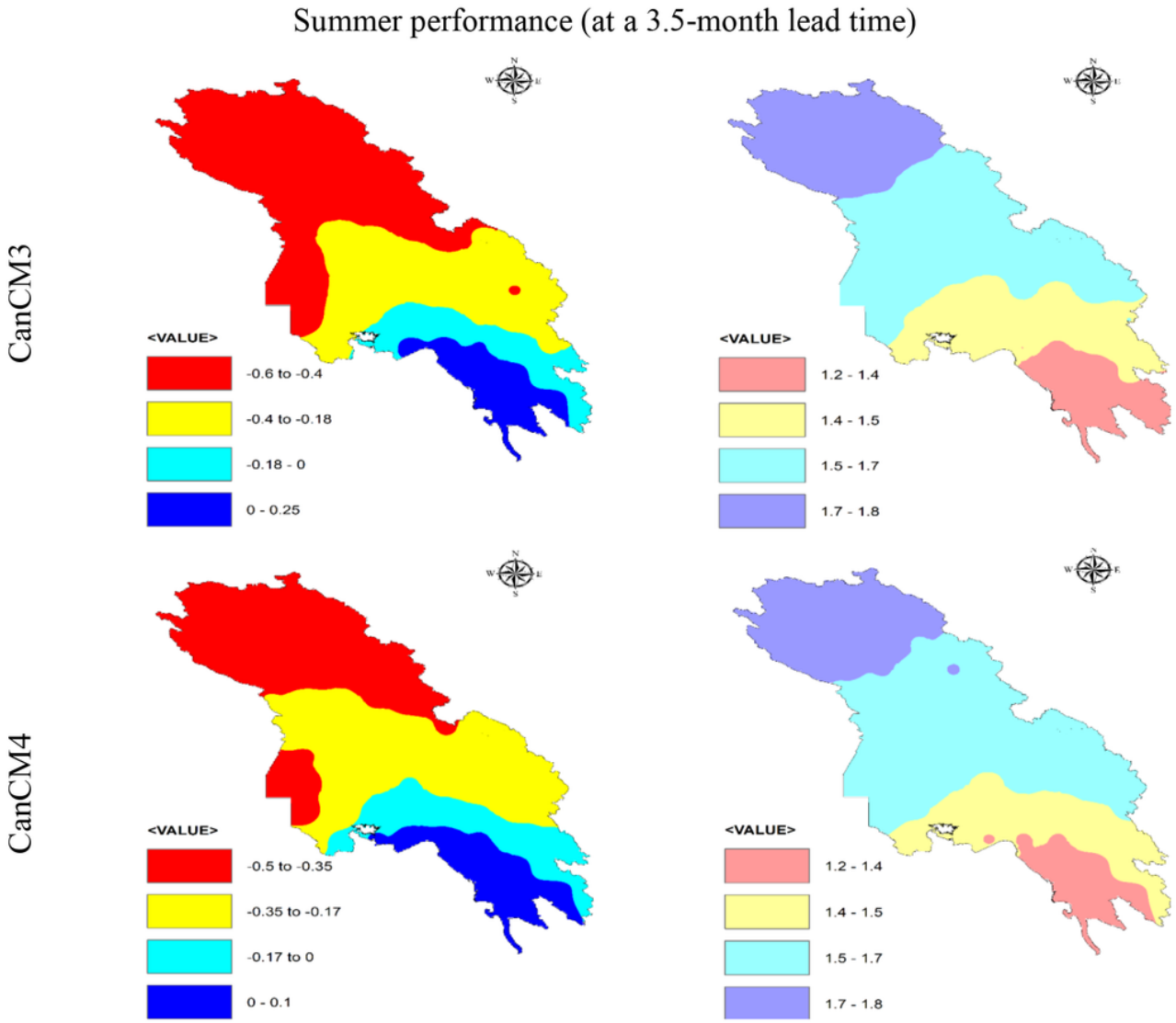


Figure 13

Seasonal performance of NMME models to detect SPEI for the forecast horizon of 3.5 months in summer

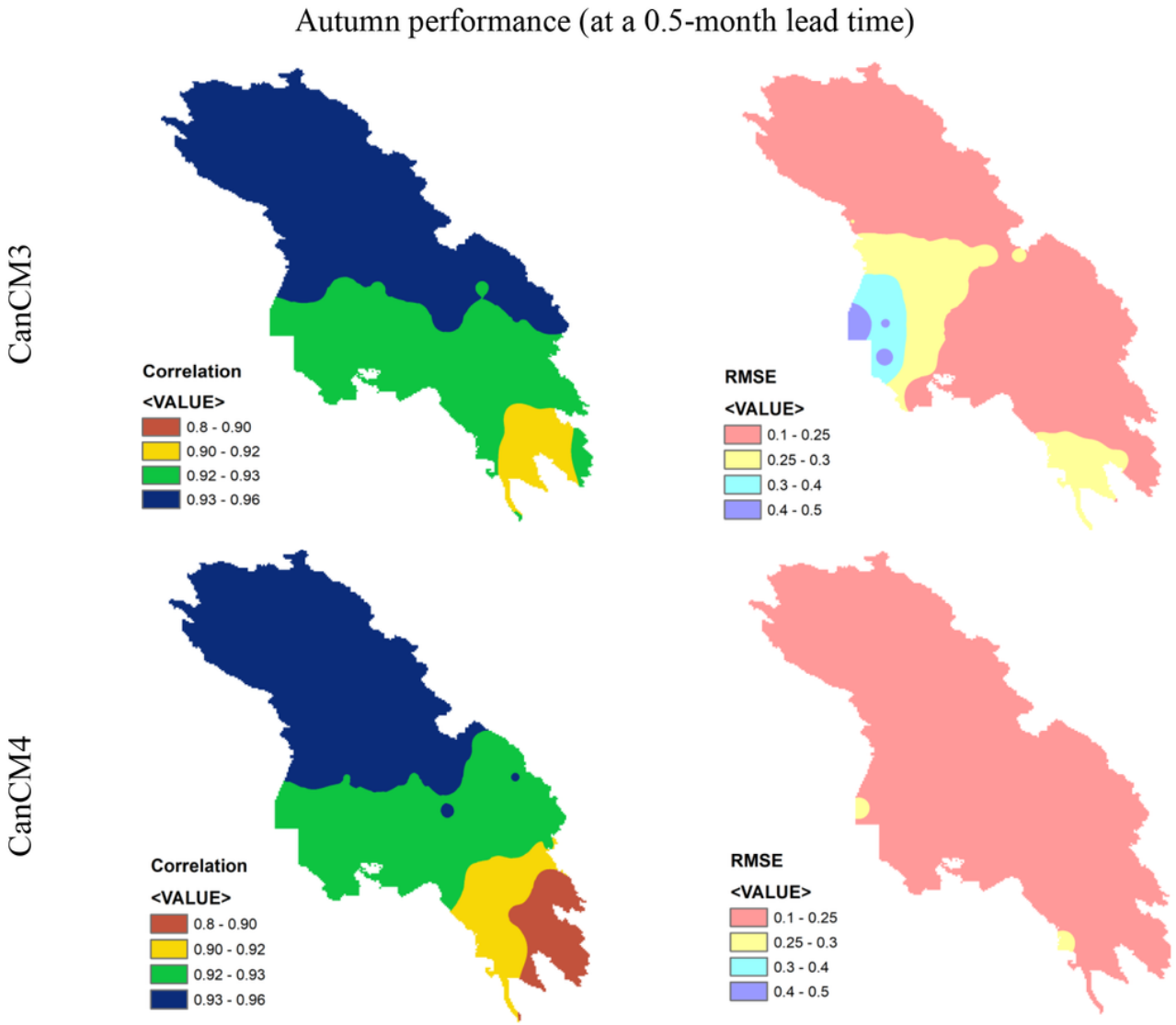


Figure 14

Seasonal performance of NMME models to detect SPEI for the forecast horizon of 0.5 months in autumn

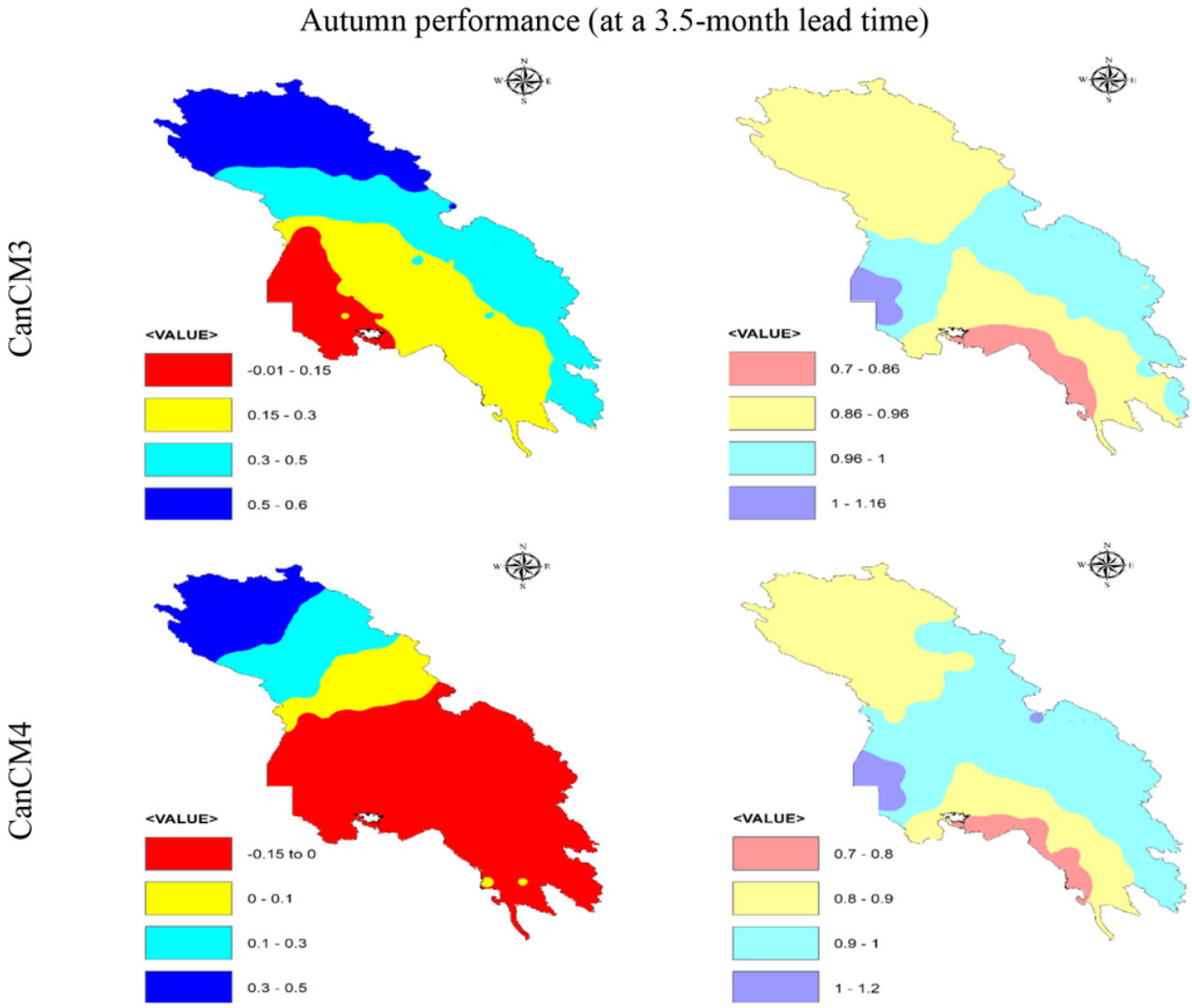


Figure 15

Seasonal performance of NMME models to detect SPEI for the forecast horizon of 3.5 months in autumn

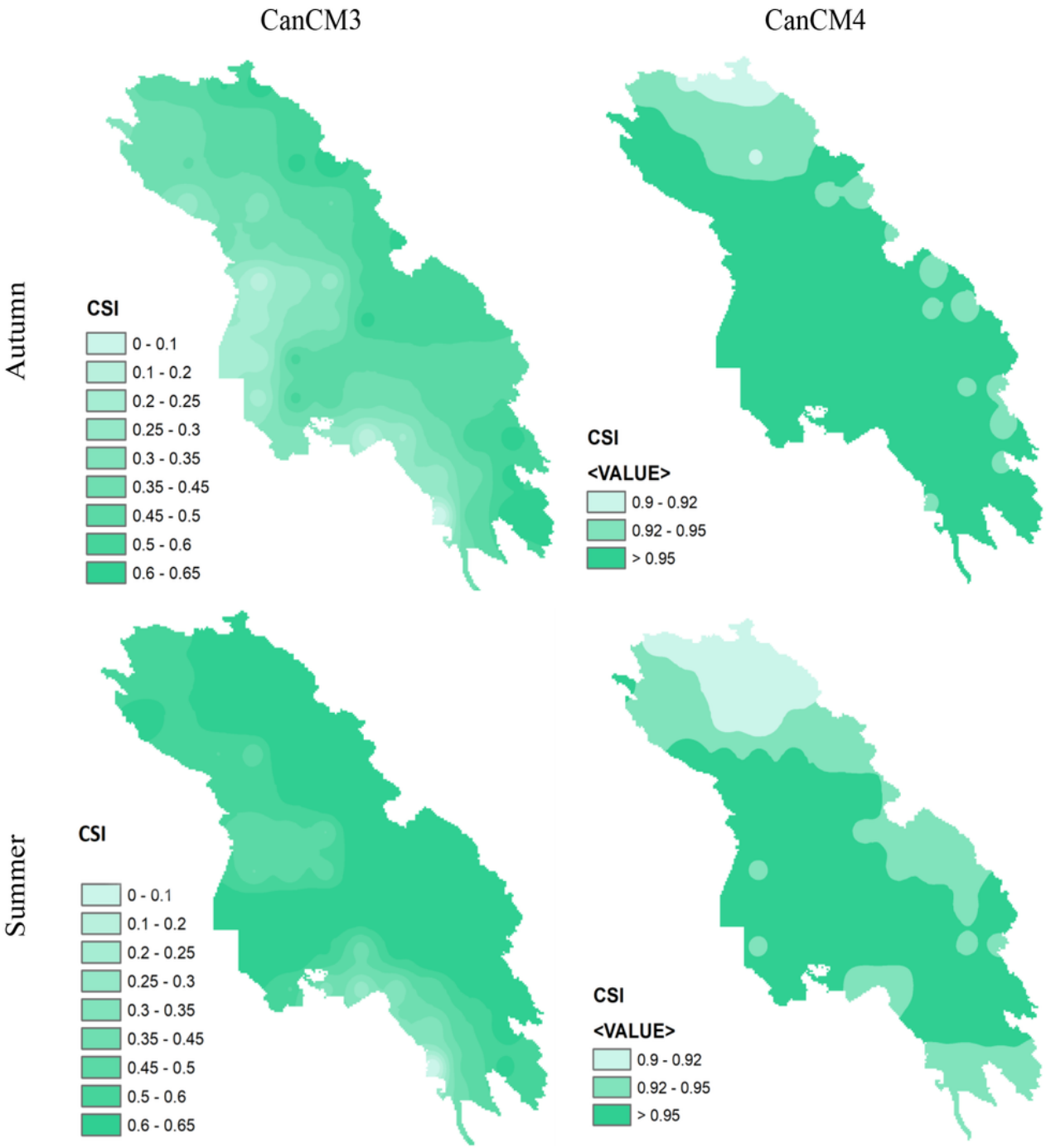


Figure 16

The CSI of NMME models in two seasons of autumn and summer for the forecast horizon of 0.5 months (Note: The range of the map legend is different)

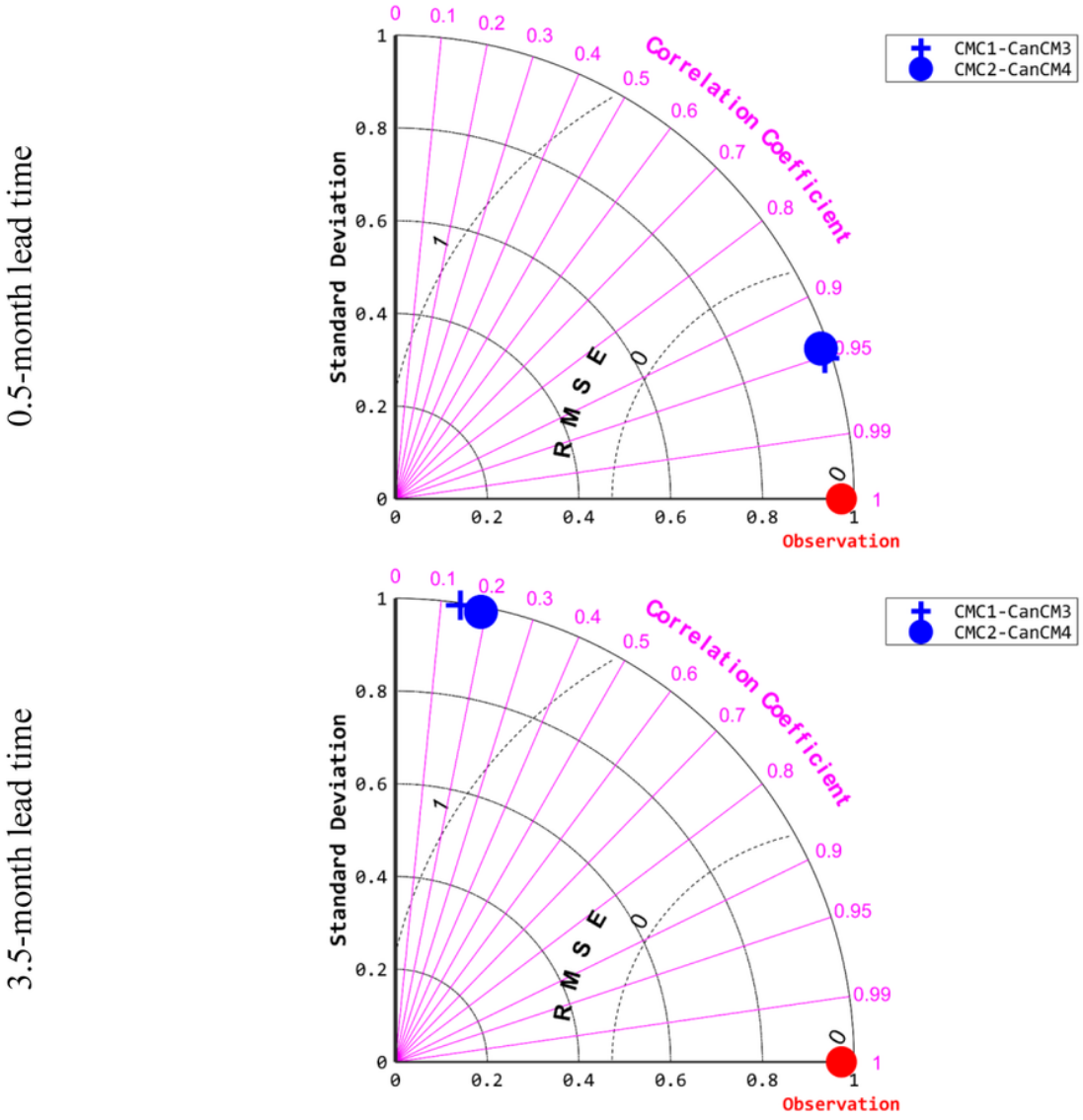


Figure 17

Taylor diagram of the NMME models' performance to detect observational SPEI

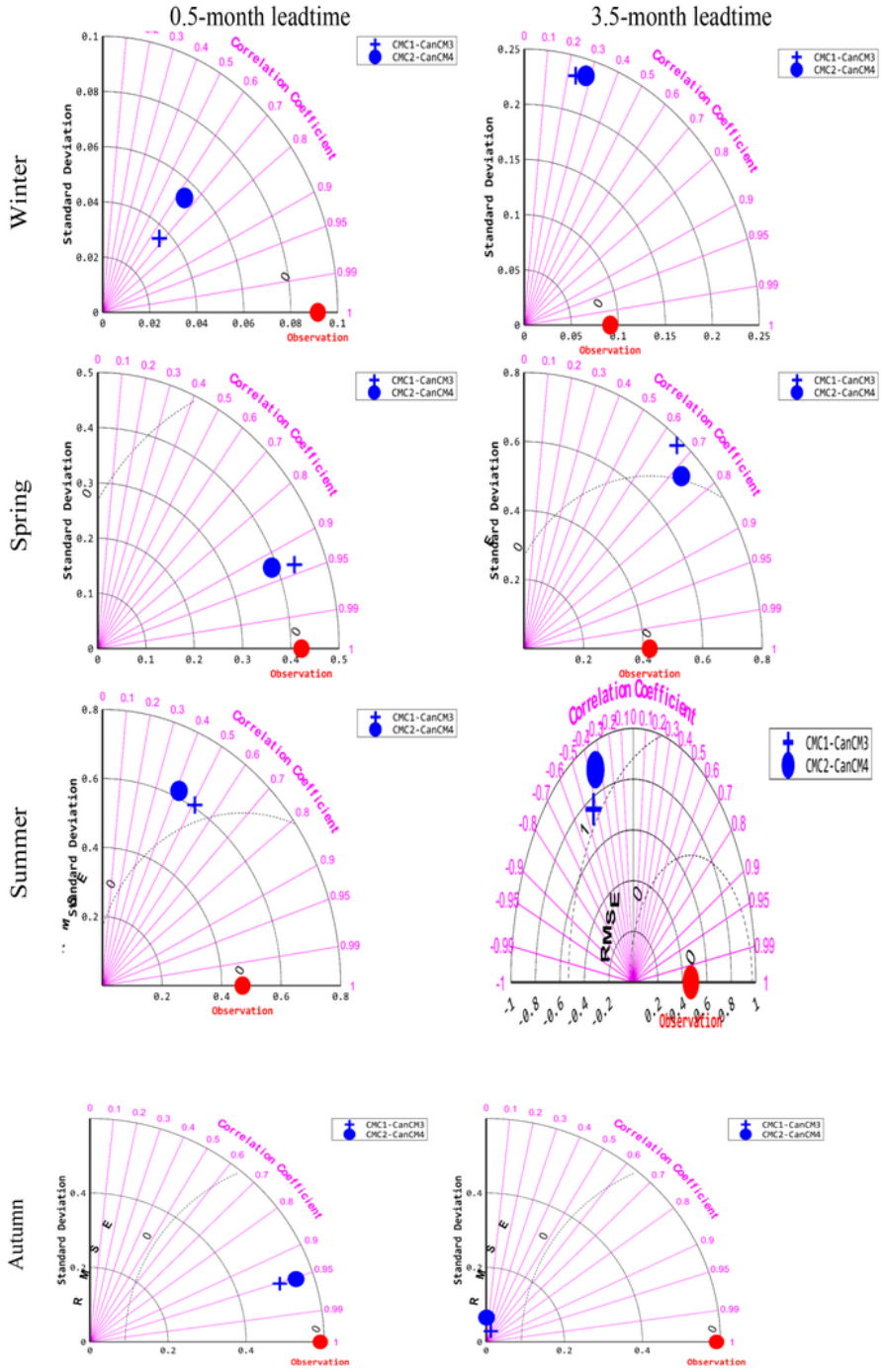


Figure 18

Taylor diagram of the NMME models' seasonal performance to detect observational SPEI