



# Verification of pre-monsoon temperature forecasts over India during 2016 with a focus on heatwave prediction

Harvir Singh<sup>1</sup>, Kopal Arora<sup>2</sup>, Raghavendra Ashrit<sup>1</sup>, and Ekkattil N. Rajagopal<sup>1</sup>

<sup>1</sup>National Centre for Medium Range Weather Forecasting, Ministry of Earth Sciences, Noida, 201309, India

<sup>2</sup>Swiss Re Global Business Solutions, Bangalore, 560027, India

Correspondence to: Harvir Singh ([harvirnitkgp@gmail.com](mailto:harvirnitkgp@gmail.com))

Received: 22 August 2016 – Discussion started: 4 October 2016

Revised: 11 May 2017 – Accepted: 26 July 2017 – Published: 6 September 2017

**Abstract.** The operational medium-range weather forecasting based on numerical weather prediction (NWP) models are complemented by the forecast products based on ensemble prediction systems (EPSs). This change has been recognised as an essentially useful tool for medium-range forecasting and is now finding its place in forecasting the extreme events. Here we investigate extreme events (heatwaves) using a high-resolution NWP model and its ensemble models in union with the classical statistical scores to serve verification purposes. With the advent of climate-change-related studies in the recent past, the rising number of extreme events and their plausible socio-economic effects have encouraged the need for forecasting and verification of extremes. Applying the traditional verification scores and associated methods to both the deterministic and the ensemble forecast, we attempted to examine the performance of the ensemble-based approach in comparison to the traditional deterministic method. The results indicate an appreciable competence of the ensemble forecast at detecting extreme events compared to the deterministic forecast. Locations of the events are also better captured by the ensemble forecast. Further, it is found that the EPS smoothes down the unexpectedly increasing signals, thereby reducing the false alarms and thus proving to be more reliable than the deterministic forecast.

forecasting capability in general, they still have a challenge in the accurate prediction of severe weather/extreme events. Severe weather events (thunderstorms, cloudburst, heatwaves and cold waves, etc.) usually involve strong non-linear interactions, often between small-scale features in the atmosphere (Legg and Mylne, 2004), for example the development of deep convection and thunderstorms in the tropics. These small-scale interactions are difficult to predict accurately (Meehl et al., 2001) and a small deviation in these could lead to completely different results, as a result of the forecast evolution process (Lorenz, 1969). The inherent uncertainty in the weather and climate forecasts can be well handled by employing ensemble-based forecasting (Buizza et al., 2005). The ensemble prediction systems (EPSs) (Mureau et al., 1993; Molteni et al., 1996; Toth and Kalnay, 1997) were first introduced in the 1990s in an effort to quantify the uncertainty caused by the synoptic-scale baroclinic instabilities in the medium-range weather forecasting (Legg and Mylne, 2004). Ensemble forecasting has emerged as a practical way of estimating the forecast uncertainty and making probabilistic forecasts (Buizza et al., 2005; Ashrit et al., 2013). It is based on multiple perturbed initial conditions, ensemble approach samples and the errors in the initial conditions used to estimate the forecast uncertainty (spread among member forecasts). The skill of the ensemble forecast shows a marked improvement over the deterministic forecast when comparing the ensemble mean to the deterministic forecast after a short lead time (Buizza et al., 2005).

The new EPS at the NCMRWF has been running for operational purposes since November 2015. This global medium-range weather forecasting system has been adopted from the UK Met Office (Sarkar et al., 2016). Generally, the model

## 1 Introduction

Reliable weather forecasting plays a pivotal role in our everyday activities. Over the years numerical weather prediction (NWP) systems have been employed to serve this purpose. While the NWP models have demonstrated an improved

and the ensemble forecast applications in addition to their verifications are used for prevalent events with a limited focus on the rare extreme weather events. It would be for the first time that the EPS technique has been employed to study the heatwave events from this model output for the extreme events over India. The heatwave is considered if the maximum temperature of a station reaches at least 40 °C or more in the plains and at least 30 °C or more for hilly regions. Based on the departure from the average, a station is declared to have heatwave conditions if the departure from the average is 4.5 to 6.4 °C and severe heatwave if the departure from the average is > 6.4 °C. In terms of the actual maximum temperature, a station is under a heatwave when the actual maximum temperature  $\geq 45$  °C and severe heatwave when the maximum temperature is > 47 °C. There has been increasing interest among the general public, media and local administration in predicting extremes such as the heatwave and cold wave events in India due to the associated loss of life. An increasing number of extreme temperature events over India were documented in several recent studies (Alexander et al., 2006; Kothawale et al., 2010; Hartmann et al., 2013; Rohini et al., 2016; Mehdi and Dehkale 2016). A climatological study of heatwaves and cold waves shows that over the Indian subcontinent between 1969 and 2013 there were more frequent cold waves and heatwave events over the Indo-Gangetic plains of India. In another study, carried out for the whole of southern Asia, Sheik et al. (2015) reported that warm extremes have become more common and cold extremes less common.

The global temperatures have exhibited a warming trend of about 0.85 °C due to anthropogenic activities between 1880 and 2012 (Hartmann et al., 2013; Rohini et al., 2016). Similar trends were also observed in India with the annual air surface temperature rising during the 20th century. This is evident from the detailed study presented in Kothawale et al. (2010) based on the data from 1901 to 2007. The study (Kothawale et al., 2010) shows that Indian mean, maximum and minimum annual temperatures have significantly increased by 0.51, 0.71 and 0.27 °C per 100 years respectively, during 1901–2007. However, accelerated warming was observed during 1971–2007, mainly due to the decade 1998–2007. The study (Kothawale et al., 2010) highlights that the mean temperature during the pre-monsoon season (March–May) shows an increasing trend of 0.42 °C per 100 years. On the other hand, a recently reiterated by Hartmann et al. (2013) noted “unequivocal” proof of the increasing warming trend globally, which could be associated with variations in the climate system. This indicates a need to comprehend the heatwave events on weather and climatic scales.

The current paper attempts to demonstrate the capability and strength of predicting such events using both the ensemble and deterministic models. The present study investigates the heatwave events during the summer months, March, April and May (MAM), 2016 in India. This study uses two cases to demonstrate the strengths and weaknesses of the EPS approach in predicting such extreme events.

With these factors in mind, we can say that temperature (minimum and maximum) forms a vital component of weather and climatic studies, which are becoming increasingly important and challenging. Reliable projections of such changes in our weather and climate are critical for adaptation and mitigation planning by the agencies involved. The knowledge would undoubtedly be useful for laymen and wider society. Forecast verification plays an important role in addressing two main questions: how good is a forecast and how much confidence can we have in it?

Verification by employing statistical scores is a well-established and recommended method by the World Meteorological Organization (WMO) that was adopted in this study (WMO, 2008). It is challenging when one needs to decide how much confidence can be placed in a model. Depending upon the statistical characteristics of the variable addressed, the score type is chosen and is employed for the verification. This offers the choice and challenge of adopting the most compatible score type. The set of verification scores used here are listed and briefly discussed in the next section.

In this paper, we investigate the utility of the ensemble prediction system over the deterministic forecast in studying extreme events like heatwaves. This forms the first documented study of the 2016 heatwave events over India which was verified using the deterministic and the ensemble forecasts. This also provides some important insights into the use of the ensemble forecast over the deterministic forecast in predicting extreme events like heatwaves. However, this study is unable to encompass an entire discussion on the efficiency of the EPS in general, as it examines a narrow range of phenomena over a not so wide region.

The paper begins with a brief explanation of the observed temperature ( $T_{\max}$  and  $T_{\min}$ ) data sets, model description and the methodology used. It will then go on to the results section which encompasses two case studies from the 2016 heatwave events in India, followed by the verification results and finally ending with the discussions and conclusions.

## 2 Observation, model description and verification methodology

### 2.1 Observed temperature (maximum and minimum)

IMD developed a high-resolution daily gridded temperature data set at  $0.5^\circ \times 0.5^\circ$  resolution, which is available to the users ([http://www.imdpune.gov.in/Seasons/Temperature/max/Max\\_Download.html](http://www.imdpune.gov.in/Seasons/Temperature/max/Max_Download.html)). The data processing procedure has been well documented (Srivastava et al., 2009). IMD has compiled, digitised, quality controlled and archived these data at the National Data Centre (NDC). Based on maximum data availability, some stations were subjected to quality-control checks like rejecting values greater than known extreme values, minimum temperatures greater than maximum temperatures, the same temperature values for many consec-

utive days, etc. After these quality checks, 395 stations were selected for further development of gridded data. IMD used measurements at these selected stations and interpolated the data into grids with the modified version of Shepard's angular distance weighting algorithm (Shepard, 1968). In this study, we have used IMD's real-time daily gridded (Shepard, 1968; Piper and Stewart, 1996; New et al., 2000; Kiktev et al., 2003; Rajeevan et al., 2005; Caesar et al., 2006; Srivastava et al., 2009) temperature (maximum and minimum) data to verify the real-time forecasts based on NCMRWF unified model (NCUM; deterministic) and NCMRWF ensemble prediction system (NEPS) ensemble mean forecast temperatures. The verification is carried out for the entire period from March 2016 to May 2016 at  $0.5^\circ \times 0.5^\circ$  resolution over India.

## 2.2 NCMRWF unified model (NCUM)

The unified model (John et al., 2016), operational at NCMRWF consists of an observation processing system (OPS 30.1), four-dimensional variational data assimilation (VAR 30.1) and unified model (UM 8.5). This analysis system makes use of various conventional and satellite observations. The analysis produced by this data assimilation system is being used as the initial condition for the daily operational high-resolution (N768L70) global NCUM 10-day forecast since January 2016. The horizontal resolution of NCUM system is 17 km and has 70 levels in the vertical extends from surface to 80 km height. The NCUM model forecast temperature ( $T_{\max}$  and  $T_{\min}$ ) data have been interpolated to the  $0.5^\circ \times 0.5^\circ$  resolution using a bilinear interpolation method to match the resolution and grids of the observed data.

## 2.3 NCMRWF ensemble prediction system (NEPS)

NEPS is a global medium-range ensemble forecasting system adapted from the UK Met Office MOGREPS system (Bowler et al., 2008). The configuration consists of four cycles of assimilation corresponding to 00Z, 06Z, 12Z and 18Z, and 10-day forecasts are made using the 00Z initial condition. The N400L70 forecast model consists of  $800 \times 600$  grid points on the horizontal surface and has 70 vertical levels. The horizontal resolution of the model is approximately 33 km in the midlatitudes. The 10-day control forecast run starts with the operational NCUM (N768L70) analysis and 44 ensemble members start from different perturbed initial conditions consistently with the uncertainty in the initial conditions. The initial perturbations are generated using the ensemble transform Kalman filter (ETKF) method (Bishop et al., 2001). Uncertainty in the forecasting model is taken into account by making small random variations to the model and using a stochastic kinetic energy backscatter scheme (Tennant et al., 2010).

## 2.4 Verification metrics

There are several scores available for the categorical verification of forecasts (Wilks, 2011). However, in the current study, we have used the probability of detection (POD), false alarm ratio (FAR), equitable threat score (ETS), Hanssen and Kuipers score (HK score) and symmetrical extremal dependence index (SEDI). A brief description of these scores is presented here.

### 2.4.1 Probability of detection (POD), also known as hit rate (H)

The POD tries to answer the question, "What fraction of the observed yes events were correctly forecasted?" It is very sensitive to hits, but ignores false alarms, and very sensitive to the climatological frequency of the event. It is good for rare events and can be artificially improved by issuing more yes forecasts to increase the number of hits. Its value varies from 0 to 1 for perfectly forecasted events  $POD = 1$  and it is computed by Eq. (1):

$$POD = \frac{\text{hits}}{\text{hits} + \text{misses}}. \quad (1)$$

### 2.4.2 False alarm ratio (FAR) (F)

What fraction of the predicted yes events actually did not occur? FAR is sensitive to false alarms, but ignores misses, very sensitive to the climatological frequency of the event and should be used in conjunction with the probability of detection. FAR is computed by Eq. (2):

$$FAR = \frac{\text{hits}}{\text{hits} - \text{false alarms}}. \quad (2)$$

### 2.4.3 Hanssen and Kuipers score (HK)

It reveals the true skill statistic and focuses on how well the forecast separates the "yes" events from the "no" events. HK uses all elements in the contingency table and does not depend on climatological event frequency. The score ranges between  $-1$  to  $1$ . This includes  $0$ , which indicates no skill and  $1$  denotes a perfect skill. It is computed by Eq. (3):

$$HK = \left[ \frac{\text{hits}}{\text{hits} + \text{misses}} \right] - \left[ \frac{\text{false alarms}}{\text{false alarms} + \text{correct negatives}} \right], \quad (3)$$

This score is efficient at verifying the most frequent events. Temperature possesses continuous values like precipitation amount and a few other NWP variables. In such cases mean error, mean square error (MSE), root mean square error (RMSE), correlation and anomaly correlation are most suitable (4th international verification methods workshop, Helsinki, June 2009). Categorical values, for instance precipitation occurrences, are well suited to verification analysis using POD, FAR, Heidke skill score, equitable threat

score and HK score statistics. However, in order to take advantage of these scores for our continuous variable, temperature (maximum and minimum), we categorise it using the temperature ranges 30–32, 32–34, 34–36, 36–38, 38–40 and 40–42 °C.

#### 2.4.4 Equitable threat score (ETS)

It is also known as the Gilbert skill score and describes how well the forecasted yes events agree with the observed yes events, thus exploring the hits by chance. This score ranges between  $-1/3$  and 1; 0 shows no skill and 1 denotes the perfect skill. The score expresses the fraction of observed or the forecasted events projected accurately and ETS is computed by the following Eq. (4):

$$\text{ETS} = \frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{false alarms} - \text{hits}_{\text{random}}}, \quad (4)$$

$$\text{where } \text{hits}_{\text{random}} = \frac{(\text{hits-misses})(\text{hits} + \text{false alarms})}{\text{total}}.$$

#### 2.4.5 Symmetric extremal dependence index (SEDI)

It expresses the association between a forecast and the observed rare events. It ranges between  $-1$  and 1, where the perfect score is 1. This score converges to  $(2X - 1)$  as the event frequency advances towards 0, where  $X$  denotes the variable that specifies the hit rate convergence to 0 for the rarer events. SEDI is not influenced by the base rate SEDI score approaching 1 and is computed by the following Eq. (5):

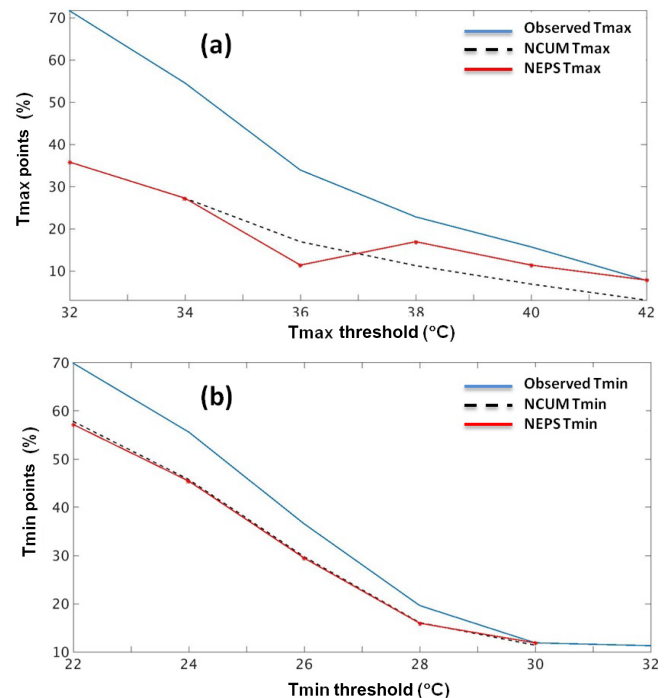
$$\text{SEDI} = \frac{\ln F - \ln H + \ln(1 - H) - \ln(1 - F)}{\ln F + \ln H + \ln(1 - H) + \ln(1 - F)}. \quad (5)$$

Here  $H$  is hit rate (POD) and  $F$  is FAR.

### 3 Results and discussions

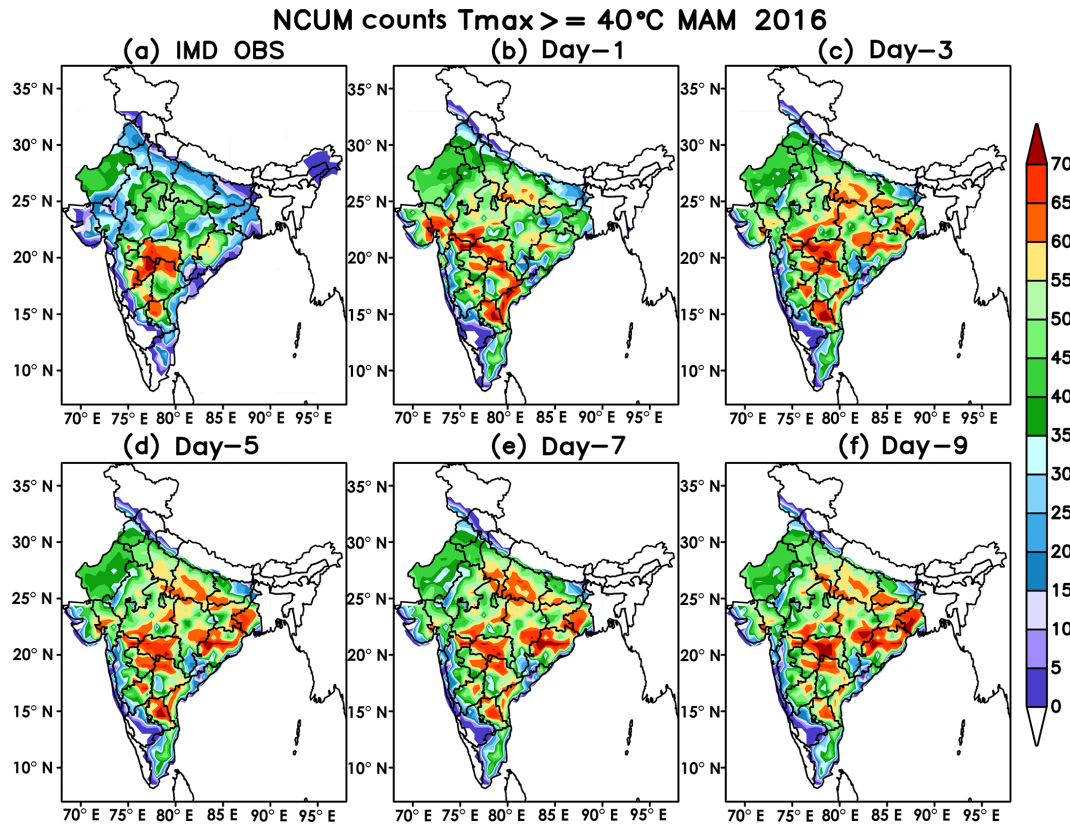
The verification of temperature forecasts is presented in this section. The models run operationally and provide the forecasts every day for 10 days. The verification is confined to MAM 2016 over six different temperature thresholds. For  $T_{\text{max}}$ , the temperature thresholds are 32, 34, 36, 38, 40 and 42 °C, and for the  $T_{\text{min}}$ , however, it is 22, 24, 26, 28, 30 and 32 °C. The panels in Fig. 1a, b show the observed and forecast (day 3) frequency distribution for  $T_{\text{max}}$  and  $T_{\text{min}}$ . For lower temperature thresholds, the forecast underestimates the frequency, while both deterministic and ensemble mean converge towards observed relative frequency, especially for the temperature exceeding 38 °C. The NEPS performs better than the NCUM forecast (Fig. 1a), indicating better performance of the ensemble model.

From the spatial map Fig. 2, the frequency of the observed maximum temperature  $T_{\text{max}} \geq 40$  °C over Maharashtra and adjoining regions shows a maximum (more than 70 counts)



**Figure 1.** Frequency distribution of observed and forecast (NCUM and NEPS) (a)  $T_{\text{max}}$  (°C) and (b)  $T_{\text{min}}$  (°C) over India during March to May 2016 (NCUM stands for NCMRWF unified model and NEPS stands for NCMRWF ensemble prediction system; observed data from India Meteorological Department and forecast data from NCMRWF). Both panels indicate that the forecasts have a lower frequency compared to observations at lower thresholds. At higher temperature thresholds ( $> 40$  °C for  $T_{\text{max}}$  and  $> 28$  °C for  $T_{\text{min}}$ ), there is better agreement.

over the entire period of MAM 2016, which is picked up by both deterministic and ensemble forecasts. However, the deterministic forecast shows more frequency spread over Madhya Pradesh (MP), Uttar Pradesh (UP) and Bihar, Odisha, Andhra Pradesh (AP) and adjoining states from day 1 to day 9. As forecast lead time increases from day 1 to day 9 the heatwave frequency increases from central India to the northern and eastern India. Consequently, a higher number of heatwave extremes were predicted by NCUM over eastern UP, Bihar, West Bengal, Odisha, Jharkhand, Chhattisgarh and AP. On the other hand, NEPS (Fig. 3) prediction for the day 1 to day 9 is much more subdued than in the NCUM forecasts. However, both models, NCUM and NEPS, predict the heatwaves more frequently over the above-mentioned regions. Comparatively, the ensemble-based model NEPS is performing better (spatially) for the extremes of heatwave events than the deterministic-based model NCUM over most of the Indian states up to day 9.



**Figure 2.** Spatial distribution of observed (a) and NCUM forecasts (b–f) number of days with  $T_{\max} \geq 40^{\circ}\text{C}$  during the period of March to May 2016 (NCUM stand for NCMRWF unified model; observed data from India Meteorological Department and forecast data from NCMRWF). Regions with over 60 days of  $T_{\max} \geq 40^{\circ}\text{C}$  are indicated in orange and red. In the observations (a) this is confined to a small part over the peninsula. The NCUM forecasts (b–f) show that  $T_{\max} \geq 40^{\circ}\text{C}$  is widespread over northern and eastern parts of India.

## 4 Case studies for extreme heatwaves

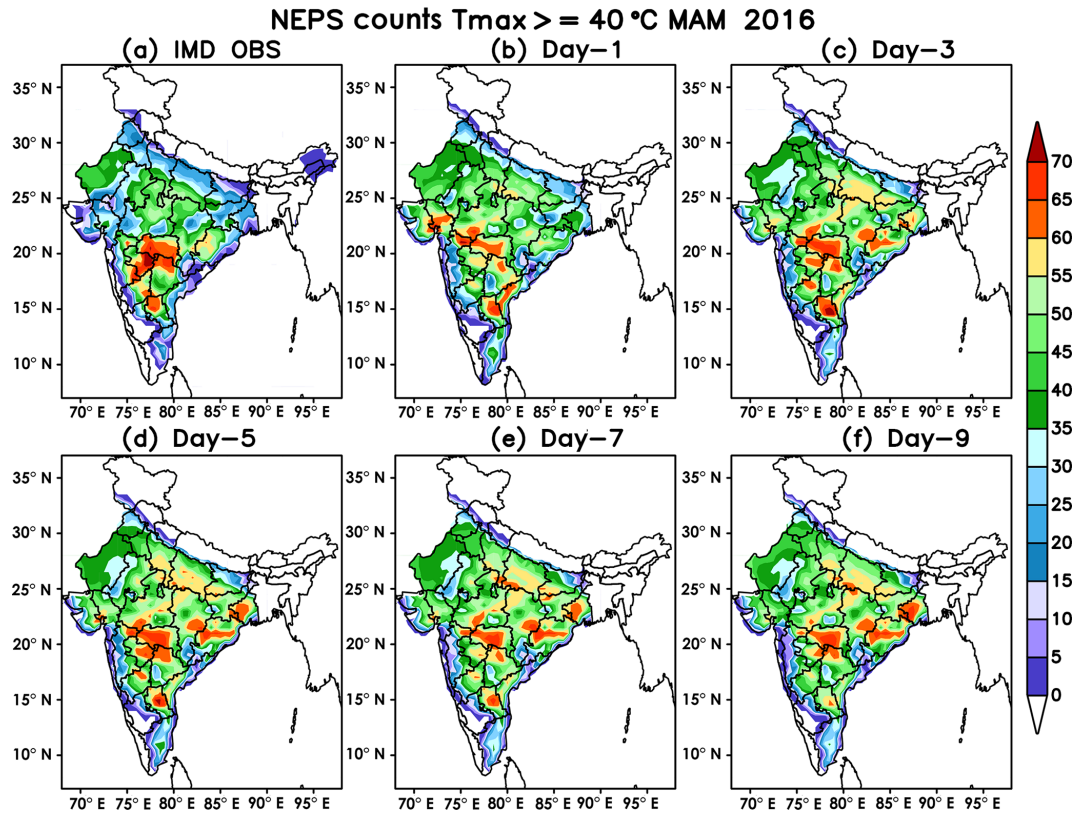
### 4.1 Weather conditions during March–May 2016

The heatwave conditions prevailed at some places over the central and adjoining western parts of the country during the last week of March 2016 (Climate Diagnostics Bulletin of India, March 2016) and over parts of central and north-western India during the first week of April (Climate Diagnostics Bulletin of India, April 2016). These conditions prevailed over most parts of eastern India throughout the second week. According to IMD official reports the severity and extent of heating increased during the third week, resulting in the establishment of severe heatwave conditions over parts of northern and eastern India. These conditions continued to prevail over eastern India and also spread over parts of southern India during the fourth week. However, its intensity and areal extent reduced towards the end of the week. During the last few days of April, heatwave conditions prevailed over parts of Odisha, Bihar, Gangetic West Bengal and Kerala. During the month of May 2016, heatwave conditions were reported at isolated places on some occasions over parts of Rajasthan, Punjab, Odisha, Gangetic West Bengal and Kerala

during the first fortnight of the month (Climate Diagnostics Bulletin of India, May 2016). As per the Climate Diagnostics Bulletin of India for May, severe heatwave conditions developed and intensified over parts of north-western India. From 15 May the heatwave conditions spread and persisted over parts of central and northern peninsular India until 22 of the month. Jammu and Kashmir, western and eastern Rajasthan, western and eastern MP and Vidarbha were especially affected during this period. Some stations of western Rajasthan viz. Barmer, Bikaner, Sri Ganganagar, Jaisalmer and Jodhpur observed severe heatwave conditions for 4 to 5 days in succession from 17 to 21 May and temperature observed  $\geq 50^{\circ}\text{C}$ . The heatwave conditions gradually abated from most parts of the country after 23 March and prevailed only at isolated places over parts of coastal AP and Telangana during the last few days of the month.

### 4.2 Casualties reported during March–May 2016

As per the official IMD reports (Climate Diagnostics Bulletin of India, March, April and May 2016) the prevailing heatwave over India caused more than 500 loss of lives. The heatwave claimed one life each (Climate Diagnostics Bulletin of



**Figure 3.** Spatial distribution of observed (a) and NEPS forecasts (b–f) number of days with  $T_{max} \geq 40\text{ }^{\circ}\text{C}$  during the period of March to May 2016 (NEPS stands for the NCMRWF ensemble prediction system; observed data from India Meteorological Department and forecast data from NCMRWF). Regions with over 60 days of  $T_{max} \geq 40\text{ }^{\circ}\text{C}$  are indicated in orange and red. In the observations (a) this is confined to a small part over peninsula. The NEPS forecasts (b–f) show  $T_{max} \geq 40\text{ }^{\circ}\text{C}$  widespread over northern and eastern parts of India. NEPS forecasts (b–f) have a better agreement with the observations (a) compared to NCUM forecasts (Fig. 2b–f).

**Table 1.** Casualties reported during March to May 2016 due to pre-  
vailing heatwaves over India.

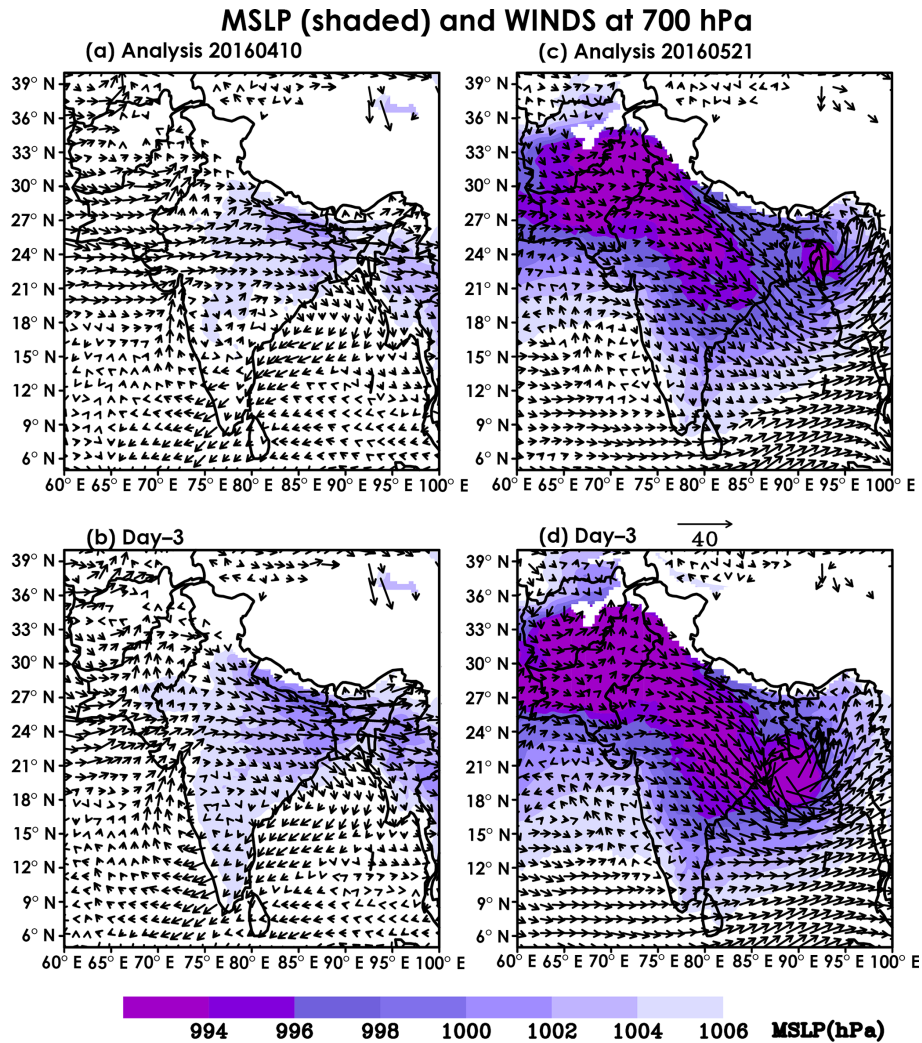
Month	State/region	No. of loss of lives	Total
March	Maharashtra	1	2
	Kerala	1	
April	Odisha	88	220
	Telangana	79	
	AP	40	
	Maharashtra	9	
	Karnataka	1	
	Tamil Nadu	1	
May	Telangana	200	273
	Gujarat	39	
	Maharashtra	34	

Data from the Climate Diagnostic Bulletin of India, March 2016, April 2016 and May 2016, India Meteorological Department.

India, March 2016) in Maharashtra (Nanded, 13 March) and Kerala (Palakkad, 5 March). A brief account of heatwave-related deaths is listed in Table 1. It took a toll of over 200 deaths (Climate Diagnostics Bulletin of India, April 2016) from central and peninsular India during the month of April. Of these, 88 deaths were reported from Odisha, 79 from Telangana, 40 from AP, 9 from Maharashtra and one each from Karnataka and Tamil Nadu. In the month of May the heatwave caused over 275 deaths from central and peninsular parts of the country. Of these, over 200 deaths (Climate Diagnostics Bulletin of India, April 2016) were reported from Telangana alone, 39 deaths were reported from Gujarat and 34 from Maharashtra.

### 4.3 Synoptic features associated with heatwaves during 2016

The panels in Fig. 4 on the left show the analysis (top) and day 3 forecast (bottom) MSLP and winds at 700 hPa for 10 April 2016. Similarly, the panels on the right show the analysis (top) and day 3 forecast (bottom) of MSLP and winds at 700 hPa for 21 May 2016. The typical synoptic fea-

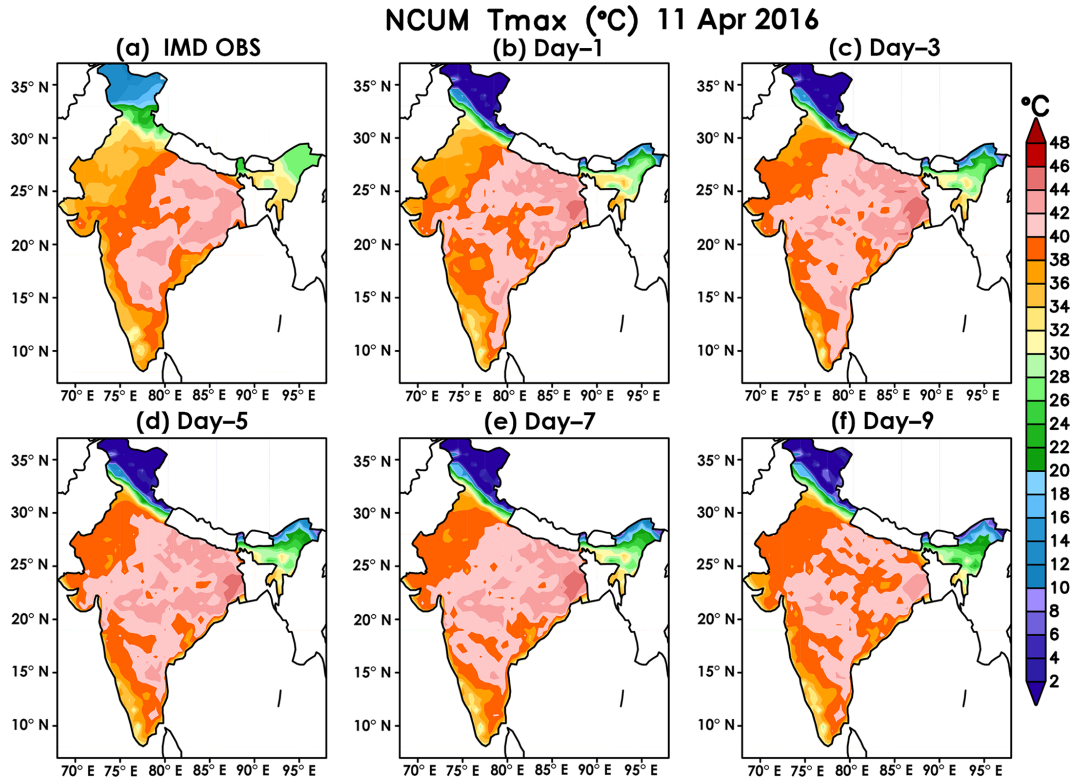


**Figure 4.** Shaded area shows mean sea level pressure (MSLP) and winds at 700 hPa, showing heat low (a). Analysis of 20160410 (b) day 3 forecast valid for 20160410 (c). Analysis of 20160521 (d) day 3 forecast valid for 20160521. The analysis and the forecasts are based on the NCMRWF unified model. Both analyses (a, b) and forecasts (c, d) feature dry north-westerlies.

tures associated with the pre-monsoon season is depicted in the above figure, which shows the MSLP in hPa (shaded) and 700 hPa winds in  $\text{ms}^{-1}$  (vectors) over the Indian subcontinent. The low pressure associated with continental heating (heat low) is prominent and is an important semi-permanent system that drives the monsoon (Rao, 1976). The heat low establishes over NW parts of India and adjoining parts of Pakistan and is seen to extend over India. The day 1 and day 3 forecasts successfully capture this broad-scale feature of the heat low. The 700 hPa winds over central India are predominantly north-westerlies driving the hot and dry air from over the Thar desert towards central India. The pre-monsoon hot weather becomes severe at times when the hot and dry north-westerlies penetrate deep into the peninsula and persist for several days. During May 2016, similar conditions caused a severe heatwave over parts of Maharashtra, Telangana and Odisha.

### 4.3.1 Case I heatwaves on 11 April 2016

As per the IMD reports (Climate Diagnostics Bulletin of India, April 2016), the heatwave conditions prevailed over parts of central peninsular and eastern India during the second week of April. It took a toll of over 200 deaths (Table 1) from central and peninsular India during April. Observed and forecast  $T_{\text{max}}$  valid for 11 April 2016 are shown for NCUM (Fig. 5) and NEPS (Fig. 6). The spatial distributions of  $T_{\text{max}}$  show prevailing heatwaves over Odisha, AP, Telangana and some parts of Maharashtra on 11 April 2016. The observation shows more than  $40^\circ\text{C}$  spread over eastern UP, Bihar, West Bengal, eastern MP, Jharkhand, Chhattisgarh, Odisha, Maharashtra, some parts of Karnataka and Tamil Nadu. The NCUM forecast, in contrast, shows marginally wider regions up to day 9 due to a warm bias in the model, and NEPS forecasts also show  $\geq 40^\circ\text{C}$  wider regions up to day 9 but



**Figure 5.** Spatial distributions of observed  $T_{max}$  (a) and NCUM forecast  $T_{max}$  (b–f) prevailing heatwaves over MP, Odisha, AP, Telangana and some parts of Maharashtra on 11 April 2016 (NCUM stands for NCMRWF unified model; observed data from India Meteorological Department and forecast data from NCMRWF). Observed (a)  $T_{max} \geq 40^\circ\text{C}$  is widespread over eastern India and a large part of the peninsula. NCUM forecasts (b–f) show a similar distribution.

marginally less than the NCUM forecasts. Apart from the warm bias, both the model forecasts show a cold bias in the north–north-east of J&K. Hence the NEPS is better at predicting the extremes of heatwaves up to day 9 than the NCUM.

**4.3.2 Case II heatwaves on 21 May 2016**

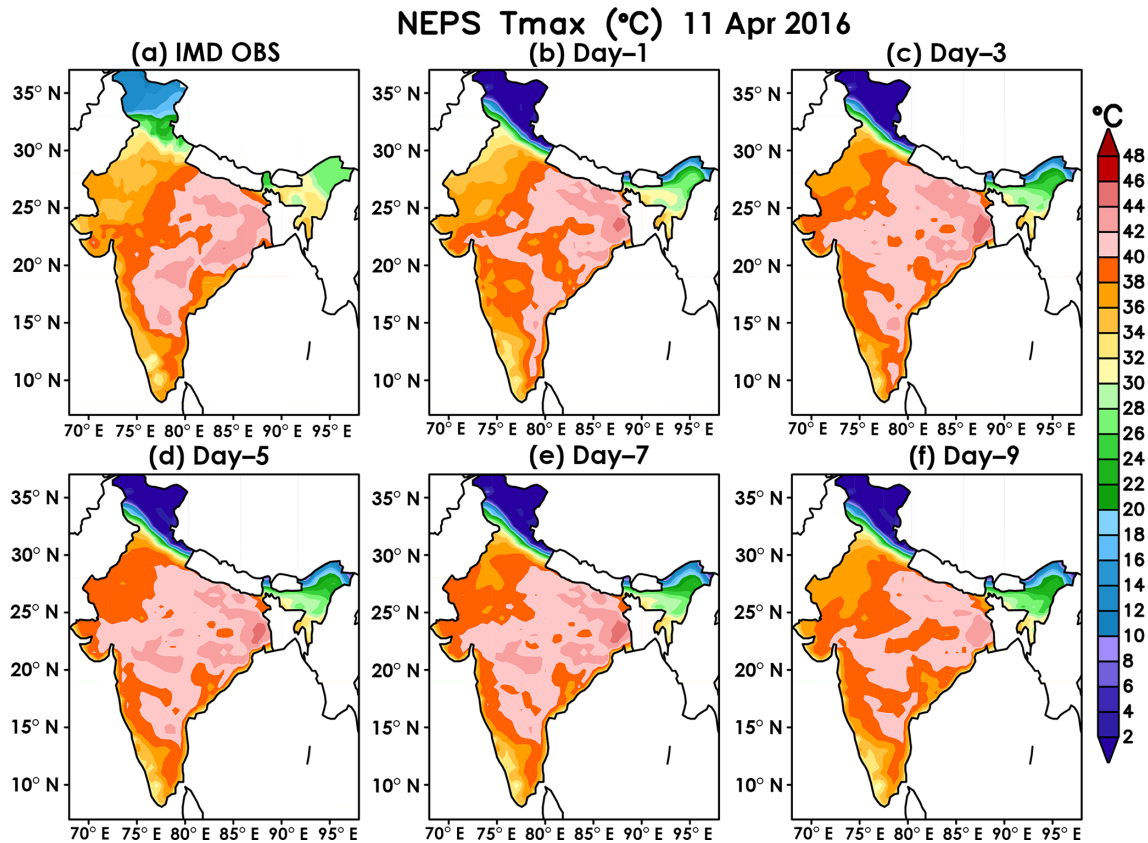
The severe heatwave conditions developed and intensified over parts of north-western India during the third week of May 2016 and persisted over parts of central and northern peninsular India. At some stations in western Rajasthan temperatures of  $\geq 50^\circ\text{C}$ , viz. at Barmer, Bikaner, Ganganagar, Jaisalmer and Jodhpur, were observed for the severe heatwave conditions for 4 to 5 days in succession from 17 May to 21 May in 2016. The spatial distributions of NCUM and NEPS forecast  $T_{max}$  with observed IMD  $T_{max}$  prevailing heatwaves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21 May 2016 is shown in Figs. 7 and 8. Both the models, deterministic and ensemble, are able to predict the extreme temperature ( $T_{max} > 48^\circ\text{C}$ ) over western Rajasthan up to day 3 only. However, the NCUM predicts more widespread  $T_{max} > 46^\circ\text{C}$  over

Rajasthan, MP, UP, Delhi, Haryana, Punjab and parts of Maharashtra for all days forecast.

The HK scores of the maximum temperature ( $T_{max}$ ) between the range  $30\text{--}42^\circ\text{C}$ , constructed as box-and-whisker plots for both NCUM and NEPS, indicate better performance of the ensemble-based forecast compared to the deterministic one. Interestingly, the forecast score does not fade away with lead time, contrary to the expectation. This depicts that the NEPS performs better and its prediction skill remains quasi-constant throughout the lead time of 10 days (Fig. 9).

Similar observations can be made from the ETS plots (Fig. 10). The most obvious finding to emerge from the box-and-whisker plots of the ETS scores is the better performance of the ensemble-based forecast (NEPS) than that of the deterministic forecast (NCUM). This result is consistent with the earlier documented findings. At all the  $T_{max}$  thresholds (between  $30$  and  $42^\circ\text{C}$ ), the NEPS mean stands above the NCUM mean. The same observation holds during all the illustrated forecasts (days 1, 3, 5, 7 and 9). The scores falling under the 25% in the case of the ensemble-based forecast are either similar or lie slightly above the deterministic forecast unlike the values underlying 75%, which in the case of NEPS are markedly higher than that of the NCUM.





**Figure 6.** Spatial distributions of observed  $T_{\max}$  (a) and NEPS forecast  $T_{\max}$  (b–f) prevailing heatwaves over MP, Odisha, AP, Telangana and some parts of Maharashtra on 11 April 2016 (NEPS stands for NCMRWF ensemble prediction system; observed data from India Meteorological Department and forecast data from NCMRWF). Observed (a)  $T_{\max} \geq 40^{\circ}\text{C}$  is widespread over eastern India and a large part of the peninsula. NEPS forecasts (b–f) show a similar distribution.

This finding raises an intriguing question regarding the difference in the characteristic distribution of both NEPS and NCUM forecasts. This result also advocates better performance of the ensemble-based forecast over the deterministic forecast.

Importantly, the ensemble-based forecast predicts a lower false alarm than its counterpart, NCUM, especially in the high temperature range. In the low temperature range, between 30 and 32, NEPS has low FAR score (where 0 denotes the perfect score) for day 1 and day 3 forecasts. Similarly, it has a comparatively higher score on days 5, 7 and 9 (Fig. 11).

The probability of detection of ensemble-based forecast is higher than the deterministic forecast during all the lead times and at all the temperature thresholds except for the day 1 forecast score for the NEPS in the range between 40 and 42  $^{\circ}\text{C}$ , where NCUM shows better performance (Fig. 12).

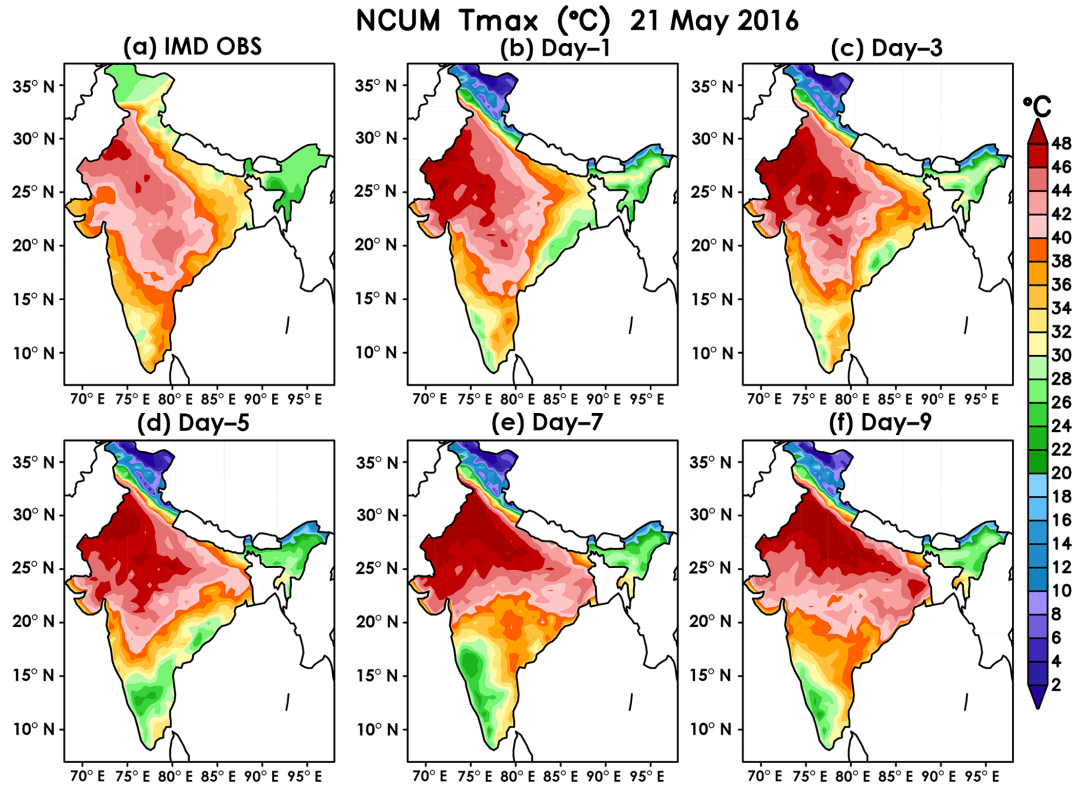
At higher temperature ranges, representing rare events, the performance of NEPS and NCUM can be clearly seen from the SEDI score plot (Fig. 13). We can notice a considerable difference between the performance of the two techniques for extreme events lying between 40 and 42  $^{\circ}\text{C}$  on all the days.

Apparently, NEPS demonstrates higher skill than that of NCUM in predicting the heatwave events. The heatwave event prediction skill is best seen on the day 5 forecast with NEPS's SEDI score encompassing the score value of 0.7. Monthly scores also are listed in Table 2.

A consistent result attained from the NEPS and NCUM verification demonstrates the better skill of the ensemble forecasts compared to the deterministic forecast for the considered cases.

## 5 Summary and conclusions

Unless the atmosphere is in a highly predictable state, we should not expect an ensemble to forecast extreme events with a high probability (Legg and Mylne, 2004). This is due to the small-scale non-linear interactions involved in a NWP model. One of the several predicted interactions could be climatologically extreme; hence they are more difficult to predict. A small variation in the intensity, timing and position of such anomalies could lead to a large difference in their predicted growth with time. Thus, despite the efficiency of

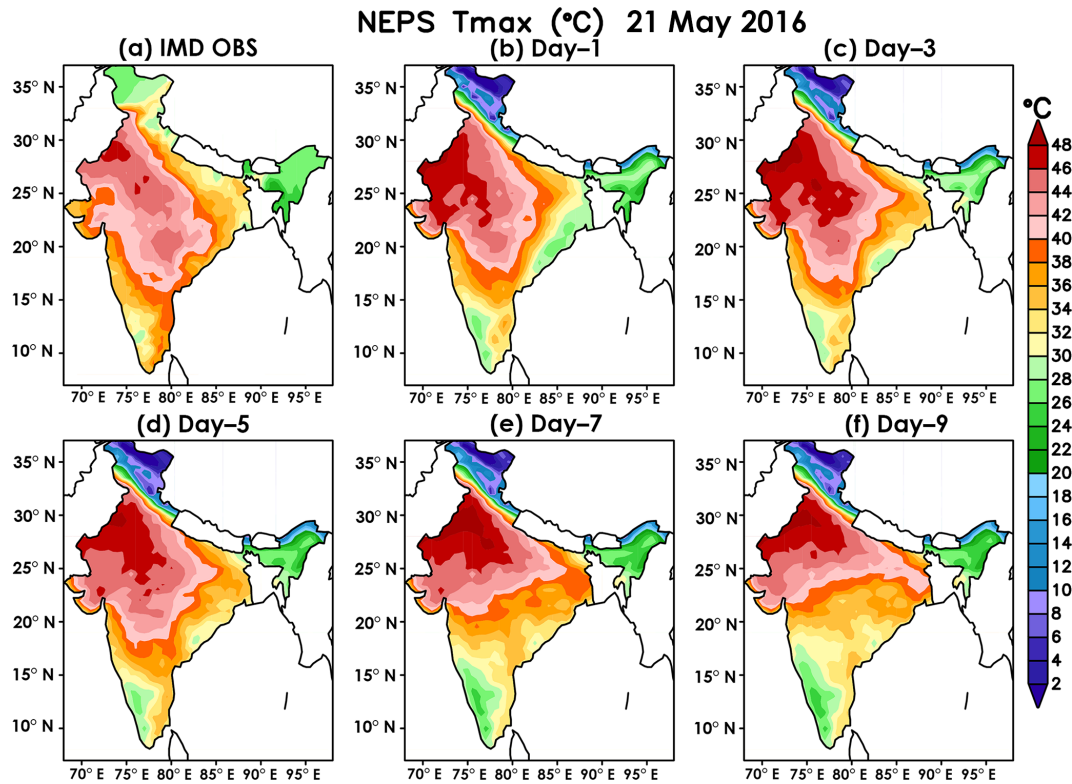


**Figure 7.** Spatial distributions of observed  $T_{max}$  (a) and NCUM forecast  $T_{max}$  (b–f) prevailing heatwaves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21 May 2016 (NCUM stands for NCMRWF unified model; observed data from India Meteorological Department and forecast data from NCMRWF). Observed (a)  $T_{max} \geq 40^\circ\text{C}$  is widespread over north-western and central India. NCUM forecasts (b–f) show a similar distribution. However, the forecasts indicate much stronger warming compared to the observations. The  $T_{max} \geq 46^\circ\text{C}$  is confined to a small region in the observations, while in the NCUM forecasts, it is seen as widespread over north-western India.

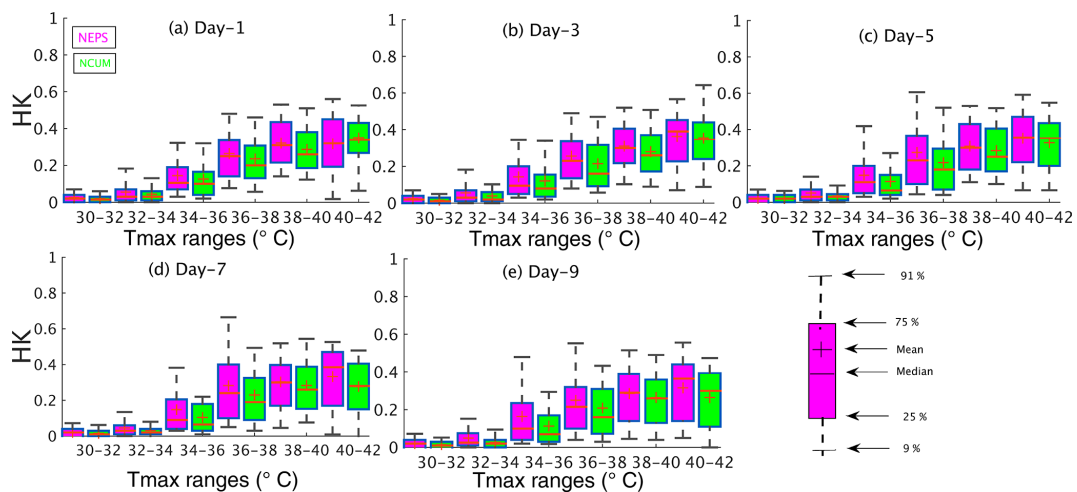
**Table 2.** Monthly verification scores for  $T_{max} > 40^\circ\text{C}$  for NCUM and NEPS forecast for day 1 to day 9 lead times with Indian Meteorological Department (IMD) observed temperature.

Month	Score	NCUM					NEPS				
		Day 1	Day 3	Day 5	Day 7	Day 9	Day 1	Day 3	Day 5	Day 7	Day 9
March	POD	0.25	0.23	0.27	0.30	0.28	0.23	0.20	0.22	0.24	0.22
	FAR	0.81	0.71	0.75	0.75	0.79	0.49	0.54	0.53	0.53	0.43
	ETS	0.09	0.09	0.09	0.08	0.08	0.10	0.09	0.10	0.11	0.11
	HK	0.22	0.21	0.24	0.27	0.25	0.21	0.18	0.21	0.23	0.21
	SEDI	0.33	0.32	0.36	0.38	0.36	0.31	0.30	0.34	0.34	0.33
April	POD	0.39	0.39	0.38	0.36	0.36	0.43	0.43	0.41	0.42	–
	FAR	0.66	0.65	0.66	0.66	0.66	0.62	0.61	0.62	0.61	0.62
	ETS	0.16	0.16	0.15	0.15	0.15	0.19	0.19	0.19	0.19	0.19
	HK	0.30	0.29	0.28	0.27	0.26	0.34	0.34	0.34	0.33	0.33
	SEDI	0.46	0.45	0.45	0.43	0.42	0.51	0.51	0.52	0.51	0.50
May	POD	0.30	0.30	0.28	0.26	0.24	0.32	0.34	0.31	0.31	0.27
	FAR	0.70	0.71	0.72	0.74	0.75	0.67	0.69	0.70	0.71	0.75
	ETS	0.12	0.11	0.11	0.10	0.09	0.14	0.14	0.13	0.12	0.10
	HK	0.22	0.22	0.21	0.19	0.17	0.25	0.26	0.24	0.23	0.19
	SEDI	0.39	0.38	0.36	0.33	0.30	0.43	0.43	0.40	0.39	0.33

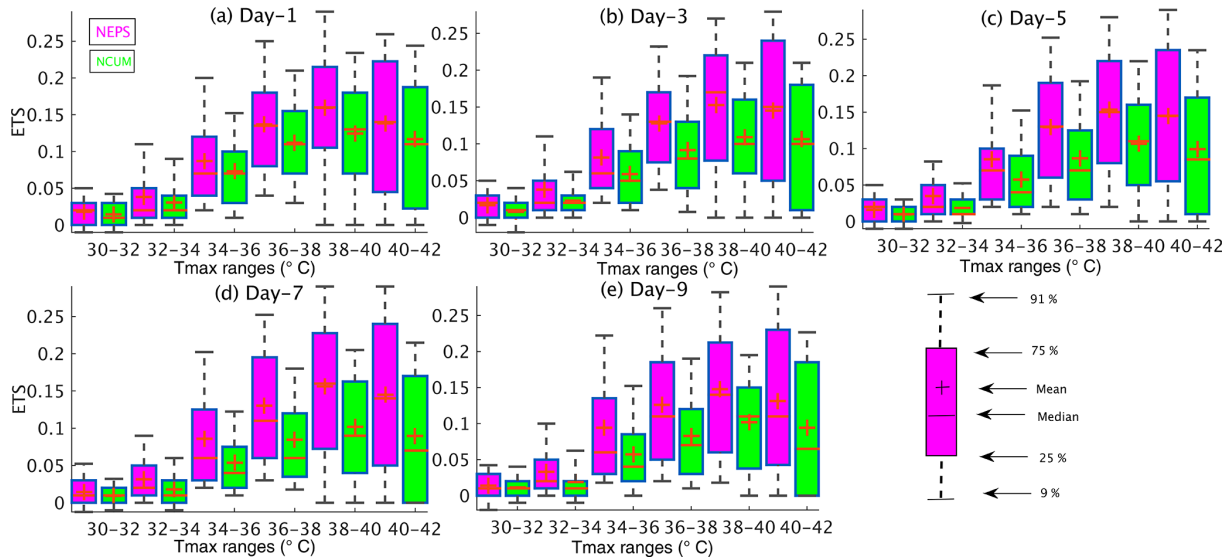
NCUM stands for NCMRWF unified model, and NEPS stands for NCMRWF ensemble prediction system.



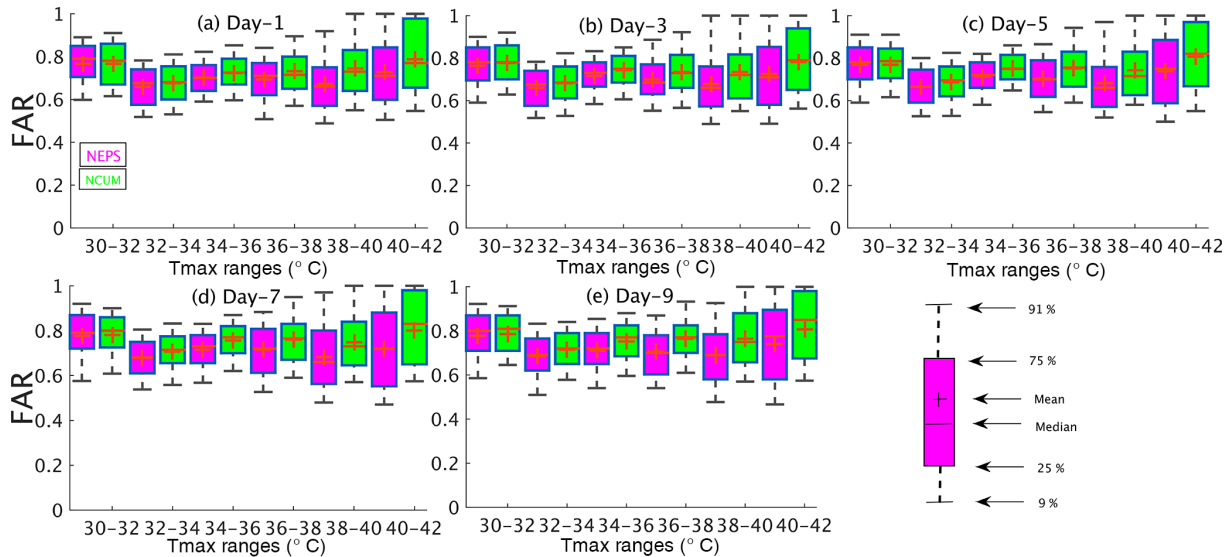
**Figure 8.** Spatial distributions of observed  $T_{max}$  (a) and NEPS forecast  $T_{max}$  (b–f) prevailing heatwaves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21 May 2016 (NEPS stands for NCMRWF ensemble prediction system; observed data from India Meteorological Department and forecast data from NCMRWF). Observed (a)  $T_{max} \geq 40^{\circ}\text{C}$  is widespread over north-western and central India. NEPS forecasts (b–f) show a similar distribution. However, the forecasts indicate much stronger warming compared to the observations. The  $T_{max} \geq 46^{\circ}\text{C}$  is confined to a small region in the observations, while in the NEPS forecasts it is widespread over north-western India. Additionally, in NEPS day 7 and day 9 forecasts (e, f)  $T_{max} \geq 40^{\circ}\text{C}$  does not extend to over the peninsula and is seen stretching eastwards compared to the observations (a).



**Figure 9.** Box plots (a–e) for Hanssen and Kuipers (HK) scores for different temperature ranges ( $T_{max}$ ) NCUM and NEPS from March to May 2016 (NCUM stands for NCMRWF unified model and NEPS stands for NCMRWF ensemble prediction system). HK score values are higher in NEPS forecasts, particularly at high temperature ranges ( $> 36^{\circ}\text{C}$ ), and for all lead times compared to NCUM forecasts, suggesting improved performance of NEPS at higher temperature ranges.



**Figure 10.** Box plots (a–e) for equitable threat score (ETS) for NCUM and NEPS from March to May 2016 (NCUM stands for NCMRWF unified model and NEPS stands for NCMRWF ensemble prediction system). ETS values are higher in NEPS forecasts, particularly at high temperature ranges, (>36 °C) and for all lead times compared to NCUM forecasts, suggesting improved performance of NEPS at higher temperature ranges.

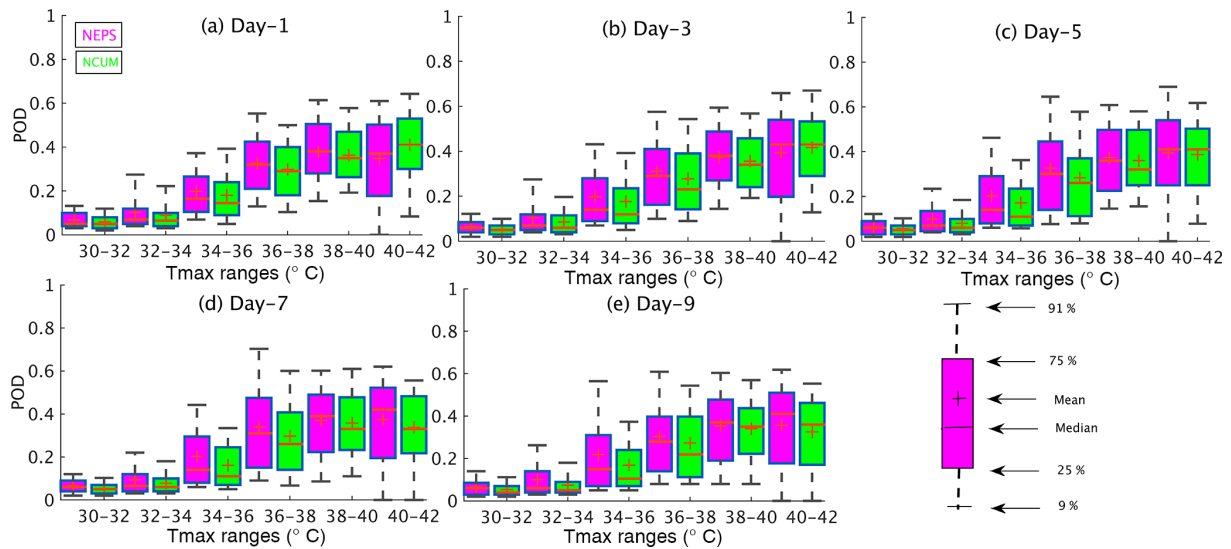


**Figure 11.** Box plots (a–e) for false alarm ratio (FAR) for NCUM and NEPS from March to May 2016 (NCUM stands for NCMRWF unified model and NEPS stands for NCMRWF ensemble prediction system). FAR values are lower in NEPS forecasts, particularly at high temperature ranges (>36 °C), and for all lead times compared to NCUM forecasts, suggesting improved performance of NEPS at higher temperature ranges.

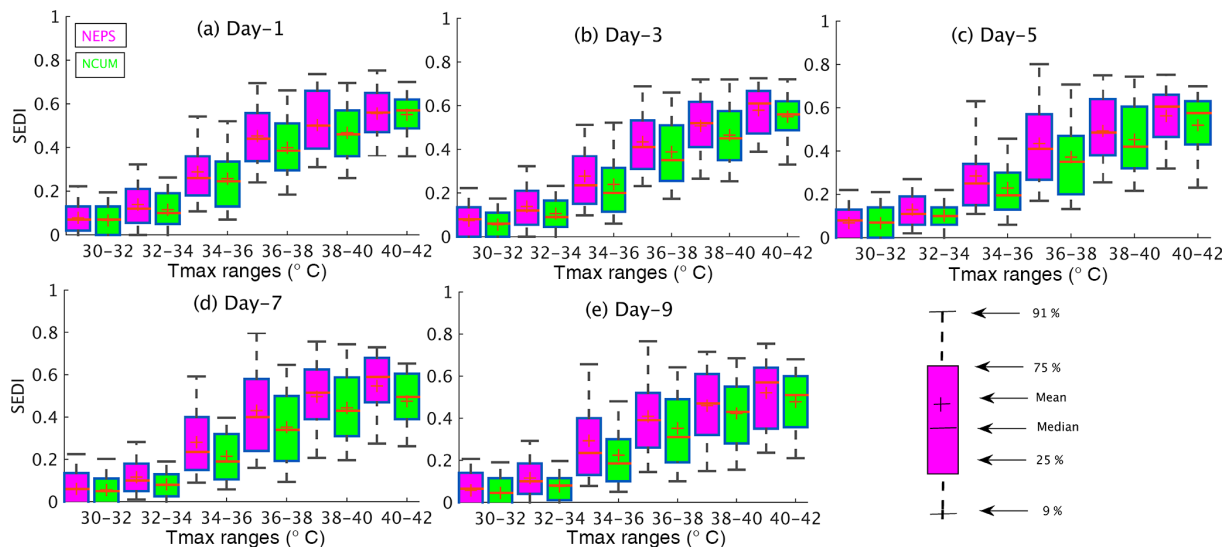
the EPS over the deterministic forecast in detecting extreme events, we should be extremely careful in declaring it locally as explained above. The ensemble mean is relatively better at predicting the extremes of heatwave events than the deterministic forecast over all Indian states up to day 9.

The ensemble forecast provides appreciable forecasts on all days and is most reliable after the day 2 forecast. This characteristic is more pronounced for extreme events than for

the less extreme events, whereas the ensemble forecast after day 2 is less reliable, as can be seen from the FAR and POD scores at the lower thresholds. This suggests that the performance of EPS on different thresholds is different; that is, if it performs well at higher thresholds, it does not necessarily mean that it would perform equally well at the lower thresholds. Thus, we need to understand the model performance at all the concerned ranges and, based upon those verification



**Figure 12.** Box plots (a–e) for probability of detection (POD) for NCUM and NEPS from March to May 2016 (NCUM stands for NCMRWF unified model and NEPS stands for NCMRWF ensemble prediction system). POD values are higher in NEPS forecasts, particularly at high temperature ranges ( $>36^{\circ}\text{C}$ ), and for all lead times compared to NCUM forecasts, suggesting improved performance of NEPS at higher temperature ranges.



**Figure 13.** Box plots (a–e) for symmetric extremal dependence index (SEDI) for NCUM and NEPS from March to May 2016 (NCUM stands for NCMRWF unified model and NEPS stands for NCMRWF ensemble prediction system). SEDI values are higher in NEPS forecasts for all temperature ranges and for all lead times compared to NCUM forecasts, suggesting improved performance of NEPS at all temperature ranges.

results, employ the ensemble forecast accordingly for operational purposes.

Our forecasts were obtained for the 2016 pre-monsoon season in India, MAM and since the severe events are rare in nature it limits the sample size for the ensemble forecast and thus poses a challenge for the efficient forecast verification. Despite the caveats involved, the ensemble forecast has shown to predict the heatwaves several days ahead of

the event, as discussed in the results. The severe heatwaves ( $>40^{\circ}\text{C}$ ) can reliably be predicted for day 2 onwards with fewer false alarms compared to the deterministic forecast as observed here. This could be explained by the inherent smoothing characteristic of the ensemble-based prediction contrary to the deterministic one, which in our case shows a warm bias.

Comparatively, low efficiency of the ensemble-based prediction on shorter timescales (< day 2) proposes that the ensemble prediction may need a longer duration of time for the perturbation growth. However, over a medium-range forecast and for extreme events like heatwaves, the ensemble-based approach proves to be one of the most economic and effective tools.

For the present study, the data from the two models is available only from 2016. Ensemble-based forecasts in real-time using the NEPS started in November 2015 at NCM-RWF. For a robust and conclusive result, it is necessary that the study is based on the higher number of cases. This will be carried out in future.

The temperature data from the station's distribution are discussed in this paper and are used to obtain the gridded  $T_{\max}$  and  $T_{\min}$  data. It is indeed likely that some of the station extremes are smoothed out in the gridded data. It should also be noted that the station's data network is sparse with 395 stations, and often there are missing values. The gridded data field provides continuous and gap-free data to work with.

Extreme events like heatwaves are rare in nature and here we provided a general view of the two particular heatwave events (11 April and 21 May). From our experience, as well as from the forecast for the post-heatwave event days, we can state that the skill of predicting an event with the initial conditions of no indication of severity is comparatively lower than when the signature is present in the initial conditions. Even before the event, there is some signature of it as can be seen in the Figs. 5, 6, 7 and 8. The overall prediction of warm conditions is fairly accurate, but at closer lead times the events are better predicted. The same can be seen in the box-and-whisker plots for ETS (and rest of the score plots as well). For instance, the skill of NEPS does not fall drastically from day 2 to day 7 and thus depicts reasonable skill. So, overall the NEPS is good at predicting the extreme event and is relatively robust.

*Data availability.* The observed temperature (maximum and minimum) data over India can be downloaded from [http://www.imdpune.gov.in/Seasons/Temperature/max/Max\\_Download.html](http://www.imdpune.gov.in/Seasons/Temperature/max/Max_Download.html) and [http://www.imdpune.gov.in/Seasons/Temperature/min/Min\\_Download.html](http://www.imdpune.gov.in/Seasons/Temperature/min/Min_Download.html). NCUM and NEPS temperature forecast data over India are made available upon request to [director@ncmrwf.gov.in](mailto:director@ncmrwf.gov.in).

**Appendix A**

## List of abbreviations

---

AP	Andhra Pradesh
EPS	Ensemble prediction system
ETKF	Ensemble transform Kalman filter
ETS	Equitable threat score
FAR	False alarm ratio
HK	Hanssen and Kuipers
IMD	Indian Meteorological Department
J&K	Jammu & Kashmir
MAM	March, April and May
MOGREPS	Met Office Global and Regional Ensemble Prediction System
MP	Madhya Pradesh
MSE	Mean square error
MSLP	Mean sea level pressure
NCMRWF	National Centre for Medium Range Weather Forecasting
NCUM	NCMRWF unified model
NDC	National Data Centre
NEPS	NCMRWF ensemble prediction system
NWP	Numerical weather prediction
OPS	Observation processing system
POD	Probability of detection
RMSE	Root mean square error
SEDI	Symmetric extremal dependence index
$T_{\max}$	Maximum temperature
$T_{\min}$	Minimum temperature
UK	United Kingdom
UM	Unified model
UP	Uttar Pradesh
WMO	World Meteorological Organization

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The authors are grateful to India Meteorological Department (IMD) for providing the gridded observed  $T_{\max}$  and  $T_{\min}$  data in real-time, which is a very useful product for verification of NWP forecasts. The authors are thankful for discussions and feedback provided by colleagues at NCMRWF, which has helped in improving the quality of the manuscript. Thanks are due for anonymous reviewers for their comments and suggestions which have helped in revising the manuscript.

Edited by: Bruce D. Malamud

Reviewed by: two anonymous referees

## References

- Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Ambenje, P., Rupa Kumar, Revadekar, K., Griffiths, J. V., Vincent, G., Stephenson, L., Burn, D., Aguilar, J. E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L.: Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.-Atmos.*, 111, D05109, <https://doi.org/10.1029/2005JD006290>, 2006.
- Ashrit, R., Iyengar, G. R., Sankar, S., Ashish, A., Dube, A., Dutta, S. K., Prasad, V. S., Rajagopal, E. N., and Basu S.: Performance of Global Ensemble Forecast System (GEFS) During Monsoon 2012, Res. Rep., NMRF/RR/01/2013, 1–25, 2013.
- Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive sampling with the ensemble transform Kalman filter, Part I: Theoretical aspects, *Mon. Weather Rev.*, 129, 420–436, 2001.
- Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., and Beare, S. E.: The MOGREPS short-range ensemble prediction system, *Q. J. R. Meteorol. Soc.*, 134, 703–722, <https://doi.org/10.1002/qj.234>, 2008.
- Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M., and Yuejian, Z.: A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems, *Mon. Weather Rev.*, 133, 1076–1097, 2005.
- Caesar, J., Alexandar, L., and Vose, R.: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set, *J. Geophys. Res.*, 111, D05101, <https://doi.org/10.1029/2005JD006280>, 2006.
- Climate Diagnostics Bulletin of India, Near Real – Time Analyses National Climate Centre, Pune, April 2016.
- Climate Diagnostics Bulletin of India, Near Real – Time Analyses National Climate Centre, Pune, March 2016.
- Climate Diagnostics Bulletin of India, Near Real – Time Analyses National Climate Centre, Pune, May 2016.
- Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M., and Zhai, P. M.: Observations: Atmosphere and Surface, in: *Climate Change 2013: The Physical Science Basis, The contribution of Working Group I to the Fifth Assessment Report* of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, <https://doi.org/10.1017/CBO9781107415324.008>, 159–254, 2013.
- John, P. G., Rani, S. R., Jayakumar, A., Mohandas, S., Mallick, S., Lodh, A., Rakhi, R., Sreevathsa, M. N. R., and Rajagopal, E. N.: NCUM Data Assimilation System, NMRF/TR/01/2016, 2016.
- Kiktev, D., Sexton, D. M. H., Alexander, L., and Folland, C. K.: Comparison of modeled and observed trends in indices of daily climate extremes, *J. Clim.*, 16, 3560–3571, 2003.
- Kothawale, D. R., Munot, A. A., and Krishna Kumar, K.: Surface air temperature variability over India during 1901–2007, and its association with ENSO, *Clim. Res.*, 42, 89–102, 2010.
- Legg, T. P. and Mylne, K. R.: Early Warnings of Severe Weather from Ensemble Forecast Information, *Weather Forecast.*, 19, 891–906, 2004.
- Lorenz, E. N.: The predictability of a flow which possesses many scales of motion, *Tellus*, 21, 289–307, 1969.
- Meehl, G. A., Lukas, R., Kiladis, G. N., Weickmann, K. M., Matthews, A., and Wheeler, M.: A conceptual framework for time and space scale interactions in the climate system, *Clim. Dynam.*, 17, 753–775, 2001.
- Mehdi, S. S. and Dehkale, B. S.: Long-term climatology and trends of heat and cold waves over southern Bihar, India, *J. Earth Syst. Sci.*, 125, 1557–1567, <https://doi.org/10.1007/s12040-016-0762-2>, 2017.
- Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T.: The ECMWF Ensemble Prediction System: Methodology and validation, *Q. J. Roy. Meteor. Soc.*, 122, 73–119, 1996.
- Mureau, R., Molteni, F., and Palmer, T. N.: Ensemble prediction using dynamically conditioned perturbations, *Q. J. Roy. Meteor. Soc.*, 119, 299–323, 1993.
- New, M., Hulme, M., and Jones, P.: Representing twentieth-century space–time climate variability, Part II: Development of 1901–96 monthly grids of terrestrial surface climate, *J. Clim.*, 13, 2217–2238, 2000.
- Piper, S. C. and Stewart, E. F.: A gridded global data set of daily temperature and precipitation for terrestrial biospheric modeling, *Global Biogeochem. Cy.*, 10, 752–782, 1996.
- Rajeevan, M., Jyoti, B., Kale, J. D., and Lal, B.: Development of a high resolution daily gridded rainfall data for Indian region, IMD Meteorological Monograph, Climatology 22, Pune, India, 27 pp., 2005.
- Rao, Y. P.: Southwest Monsoon, Meteorological Monograph (Synoptic Meteorology), No. 1/1976, India Meteorological Department, New Delhi, 366 pp., 1976.
- Rohini, P., Rajeevan, M., and Srivastava, A. K.: On the variability and increasing trend of Heat waves over India, *Sci. Rep.*, 6, 26153, <https://doi.org/10.1038/srep26153>, 2016.
- Sarkar, A., Chakraborty, P., John, P. G., and Rajagopal, E. N.: Implementation of Unified Model Based Ensemble Prediction System at NCMRWF (NEPS), NMRF/TR/02/2016, 2016.
- Sheikh, M. M., Manzoor, N., Ashraf, J., Adnan, M., Collins, D., Hameed, S., Manton, M. J., Ahmed, A. U., Baidya, S. K., Borgaonkar, H. P., Islam, N., Jayasinghearachchi, D., Kothawale, D. R., Premalal, K. H. M. S., Revadekar, J. V., and Shrestha, M. L.: Trends in extreme daily rainfall and tempera-



- ture indices over South Asia, *Int. J. Climatol.*, 35, 1625–1637, <https://doi.org/10.1002/joc.4081>, 2015.
- Shepard, D.: A two-dimensional interpolation function for irregularly spaced data, *Proc. 23rd National Conf.*, New York, NY, Association for Computing Machinery, 517–524, 1968.
- Srivastava, A. K., Rajeevan, M., and Kshirsagar, S. R.: Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region, *Atmos. Sci. Lett.*, 10, 249–254, <https://doi.org/10.1002/asl.232>, 2009.
- Tennant, W. J., Shutts, G. J., Arribas, A., and Thompson, S. A.: Using a Stochastic Kinetic Energy Backscatter Scheme to Improve MOGREPS Ensemble forecast Skill, *Mon. Weather Rev.*, 139, 1190–1206, 2010.
- Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, *Mon. Weather Rev.*, 125, 3297–3319, 1997.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, International Geophysical Series Vol. 100., Academic Press, 2011.
- WMO: Recommendations for the verification and intercomparison of QPFS and PPFPS from operational NWP models, WMO TD No. 1485, 2008.