Verification of Pre-Monsoon Temperature Forecasts over India during 2016 with focus on Heatwave Prediction

Harvir Singh, Kopal Arora, Raghavendra Ashrit, and EN Rajagopal

Ministry of Earth Sciences, National Centre for Medium Range Weather Forecasting, Noida, 201309, India

Correspondence to: Harvir Singh (harviriitkgp@gmail.com)

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 (EPS). This change has been recognized as an essentially useful tool for the medium range forecasting and is now finding its place in forecasting the extreme events. Here we investigate extreme events (Heatwaves) using a high-resolution numerical weather prediction and its ensemble forecast in union with the classical statistical scores to serve the verification purposes. With the advent of climate change related studies in the recent past, the rising extreme events and their plausible socio-economic effects have encouraged the need for forecasting and verification of extremes. Applying the traditional verification scores and the associated methods on both, deterministic and the ensemble forecast, we attempted to examine the performance of the ensemble based approach as compared to the traditional deterministic method. The results indicate towards an appreciable competence of the ensemble forecasting detecting extreme events as compared to 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 soaring signals, which thereby reduce the false alarms and thus prove to be more reliable than the deterministic forecast.

1. Introduction

Reliable weather forecasting plays a pivotal role in our everyday activities. Over the years NWP systems have been employed to serve the purpose. While the NWP models have demonstrated an improved 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 coldwaves, etc.) usually involve strong non-linear interactions ,often between small scale features in the atmosphere (Legg and Mylne, 2004). For example, 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 EPS (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 the practical way of estimating the forecast uncertainty and making probabilistic forecasts. It is based on multiple perturbed initial conditions, ensemble approach samples the errors in the initial conditions to estimate the forecast uncertainty (spread in member

forecasts). The skill of the ensemble forecast shows marked improvement over the deterministic forecast when comparing the ensemble mean to deterministic forecast after a short lead time

The new EPS at the NCMRWF is now running for operational purposes. This global medium-range weather forecasting system has been adopted from the UK Met Office (Sarkar et al., 2016). Generally, the model 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 from this model output for the extreme events over India to study the heatwave events. The heatwave is considered if maximum temperature of a station reaches at least 40°C or more for Plains and at least 30°C or more for Hilly regions. Based on departure from normal, a station is declared to have heatwave conditions if departure from normal is 4.5°C to 6.4°C and severe heatwave if the departure from normal is >6.4°C. In terms of the actual maximum temperature, a station is under heatwave when actual maximum temperature Ó45°C and severe heatwave when the maximum temperature is >47°C. There has been increasing interest in predicting such extremes, 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., 2005, Kothawale et al., 2010, Hartmann et al., 2013 and Rohini et al., 2016)- Mehdi and Dehkale (2016) in a climatological study of heat/cold waves show that over the Indian sub-continent between 1969 and 2013 there were more frequent cold and heatwave events over the Indo-Gangetic plains of India. In another study carried out for entire South Asia, Sheik et al., (2015) have 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 (IPCC, 2013, Rohini et al., 2016). Similar trends were also observed in India with the annual air surface temperature rise during 20th century. This is evident from the detailed study presented in Kothawale et al (2010) based on the data from 1901-2007. 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, an accelerated warming was observed during 1971-2007, mainly due to the last 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 IPCC report (2013) notified an ñunequivocalò proof of the increasing warming trend, globally which could be associated with the variations in the climate system. This indicates a need to comprehend the heatwave events on weather and climatic scales. This paper attempts to demonstrate the capability and strength of predicting such events using both ensemble and deterministic forecast. This research investigates the most recent heatwave events during the summer months March, April & May (MAM) 2016 in India. This investigation considers two case studies to demonstrate the strength and weaknesses of the EPS approach in predicting such extreme events.

With these factors in mind, we can say that temperature (Minimum and Maximum both), 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 adaption and mitigation planning by the agencies involved. The knowledge would

undoubtedly be useful for a layman and the society. Testing for the reliability of the NWP model results is efficiently done by the forecast verification methods. 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 method adopted in this study. However, not all score lead to the same conclusion. This is the challenging situation 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. Not all scores are equally efficient in describing a variable. This fact offers a choice and challenge to adopt 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 recent heatwave events over India which was verified using the deterministic and the ensemble forecasts ensemble forecasts. This paper talks about what an EPS can and canôt do. This also provides some important insights into the use of ensemble forecast over the deterministic forecast in predicting extreme events like a heatwave. However, this study is unable to encompass an entire discussion on the efficiency of the EPS in general as the work examines a narrow range of phenomena over a not so wider region.

The paper begins with a brief explanation of the observed temperature (*Tmax* & *Tmin*) data sets, model description and the methodology used. It will then go on to the results' section which encompasses two case studies from the recent 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)

Recently, IMD has developed a high resolution daily gridded temperature dataset at 0.5° x 0.5° resolution. Data processing procedure has been well documented (Srivastava et al., 2009). IMD has compiled, digitized, 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 exceeding known extreme values, minimum temperature greater than maximum temperature, same temperature values for many consecutive 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 Shepardôs 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 and Srivastava et al., 2009,) temperature (maximum and minimum) data to verify the realtime forecasts based on NCMRWF Unified Model (NCUM; deterministic) and NCMRWF Ensemble Predisction System (NEPS) ensemble mean forecast temperatures. The verification is carried out for the entire period from March 2016 to May 2016 at 0.5^{o}x0.5^{o} resolution over Indian land area.

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 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 (*Tmax* & *Tmin*) data have been interpolated to the 0.5^{o}x0.5^{o} resolution using 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 18Z and 10-day forecasts are made using the 00Z initial condition. The N400L70 forecast model consists of 800x600 grid points on the horizontal surface and has 70 vertical levels. Horizontal resolution of the model is approximately 33 km in the mid-latitudes. The 10-day control forecast run starts with the operational NCUM (N768L70) analysis and 44 ensemble members start from different perturbed initial conditions consistent with the uncertainty in initial conditions. The initial perturbations are generated using 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 ensemble forecasts. However, in the current study, we have used the POD, FAR, ETS, HK, and SEDI. A brief description of these scores is presented here.

POD Score or the Hit Rate (H): POD tries to answer the question, ñ*What fraction of the observed "yes" events were correctly forecast?"* It is very much sensitive to hits, but ignores false alarms and very sensitive to the climatologically 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.

$$POD = \frac{hits}{hits + misses} \qquad \textbf{Eq. 1}$$

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.

$$FAR = \frac{hits}{hits - false\ alarms} \qquad \textbf{Eq. 2}$$

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, does not depend on climatological event frequency. The expression is identical to ***HK = POD - POFD***, but the Hanssen and Kuipers score can also be interpreted as *(accuracy for events) + (accuracy for non-events) - 1*. The score ranges between -1 to 1, both inclusive along with 0, which indicates no skill and 1 denotes a perfect skill.

$$HK = \left[\frac{hits}{hits + misses}\right] - \left[\frac{false\ alarms}{false\ alarms + correct\ negatives}\right] \qquad \textbf{Eq. 3}$$

This score is efficient at verifying the most frequent events. Temperature possesses continuous values just like precipitation amount and a few other NWP variables. In such cases mean error, MSE, RMSE, correlation and anomaly correlation are best suitable (4^{th} international verification methods workshop, Helsinki, June 2009). Categorical values for instance precipitation occurrences are well suited for the verification analysis using POD, FAR, Heidke skill score, equitable threat score and H-K Statistics. However, in order to take advantage of these scores, for our continuous variable, temperature (Maximum and Minimum), we categorize it using the temperature ranges, 30-32, 32-34, 34-36, 36-38, 38-40, and 40-42 °C.

ETS: It is also known as, the Gilbert skill score describe how well the forecasted ñyesò events agree with the observed ñYesò events and thus exploring the hits by chance. This score ranges between -1/3 to 1. '0' shows no skill and 1 denotes the perfect skill. The score express the fraction of observed or the forecasted events projected accurately.

$$ETS = \frac{hits - hits_{random}}{hits + false\,alarms - hits_{random}} \qquad \textbf{Eq.4}$$

Where $hits_{random} = \frac{(hits - misses)(hits + false\,alarms)}{total}$

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 advance towards 0, where "X" denotes the variable that specifies the hit rate's convergence to 0 for the rarer events. SEDI is not influenced by the base rate SEDI score approaches 1.

$$SEDI = \frac{\ln F - \ln H + \ln(1-H) - \ln(1-F)}{\ln F + \ln H + \ln(1-H) + \ln(1-F)} \qquad \textbf{Eq.5}$$

3 Results and Discussions:

Traditionally, the performance of a forecast model is determined by a variety of statistical measures and scores which offer an effective way to quantify a model's capability. Before moving over to such methods, we begin with looking at the ensemble based and deterministic forecasts (on a daily basis) over a period of three hot summer months in India, March, April and May, and also compare it with the observations. The models are running operationally and are providing the forecasts out to 10 days every day. The verification is confined to MAM 2016, over six different temperature thresholds. For *Tmax*, the temperature thresholds are 32, 34, 36, 38, 40 & 42°C and for the *Tmin*, however, it is 22, 24, 26, 28, 30 & 32°C. The panels in Figure 1a,b show the observed and forecast (Day-3) frequency distribution for *Tmax* and *Tmin*. For lower temperature thresholds, the forecast underestimate the frequency, while it can be seen in the Figs-1(a & b), both, deterministic and ensemble mean converge towards observed relative frequency, especially for the temperature exceeding 38°C. NEPS performs better than the NCUM forecast (Figure1a), indicating better performance of the ensemble forecast over the deterministic one.

From the spatial map Figure 2, the frequency of the observed maximum temperature *Tmax* Ó 40°C over the Maharashtra and adjoining regions show maximum (more than 70 counts) over the entire period of MAM 2016, which is picked up by both deterministic and ensemble forecasts. However, deterministic forecast is showing more frequency spread over MP, UP and Bihar, Odisha, 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 north and east India. Consequently, higher number of heatwave extremes was predicted by NCUM over east UP, Bihar, West-Bengal, Odisha, Jharkhand, Chhattisgarh and AP. On the other hand, NEPS (Figure3) prediction for the day -1 to day -9 is much subdued than in the NCUM forecasts. However, both models, NCUM and NEPS are, predicting more frequently the heatwaves over the above said regions. Comparatively, the ensemble based model NEPS is performing better (spatially) for the extremes of heatwave events than the NCUM over most of the Indian states up to day-9.

4. Case Studies for Extreme Heatwaves

4.1 Weather conditions during MAM-2016

Heatwave conditions prevailed at some places over the central and adjoining western parts of the country during last week of March-2016 (Climate Diagnostics Bulletin of India, March 2016) and over parts of central and northwest India (Climate Diagnostics Bulletin of India, April 2016) during the first week of April. These conditions prevailed over most parts of east India all through the second week. According to IMD official reports the severity and extent of heating increased during the next week resulting in the establishment of severe heatwave conditions over parts of north and eastern India. These

conditions continued to prevail over east India and also spread over parts of south India during the fourth week, however, its intensity and areal extent reduced towards the end of the week. During the last few days of the April month, heatwave conditions prevailed over parts of Odisha, Bihar, Gangetic West Bengal and Kerala.

During the month of May-2016 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). Severe heatwave / heatwave conditions developed and intensified over parts of northwest India from 15th May, spread and persisted over parts of central and north peninsular India till 22nd of the month. Jammu & Kashmir, west & east Rajasthan, west & east Madhya Pradesh and Vidarbha were especially affected during this period. Some stations of West Rajasthan viz. Barmer, Bikaner, Ganganagar, Jaisalmer, and Jodhpur observed severe heatwave conditions for 4 to 5 days in succession from 17 to 21 May and temperature observed Ó 50°C. Heatwave conditions gradually abated from most parts of the country after 23rd and prevailed only at isolated places over parts of Coastal AP and Telangana during last few days of the month.

4.2 Casualties reported during MAM-2016

Prevailing heatwave over India took a toll more than 500 loss of lives. Heatwave claimed one life each (Climate Diagnostics Bulletin of India, March 2016) in Maharashtra (Nanded, 13 March) & Kerala (Palakkad, 5March). A brief account of heatwave related deaths is listed in Table 2. It took a toll of over 200 lives (Climate Diagnostics Bulletin of India, April 2016) from central and peninsular India during the April month. Of these, 88 lives 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 heatwave claimed over 275 lives from central and peninsular parts of the country. Of these, over 200 lives (Climate Diagnostics Bulletin of India, April 2016) were reported from Telangana alone. 39 lives were reported from Gujarat and 34 from Maharashtra.

4.3 Synoptic features associated with Heatwaves during 2016

The panels in Figure 4 on the left show analysis (top) and Day-3 forecast (bottom) MSLP and winds at 700 hPa for 10th April 2016. Similarly, the panels on the right show analysis (top) and Day-3 forecast (bottom) MSLP and winds at 700 hPa for 21st May 2016. The typical synoptic features associated with the pre-monsoon season is depicted in the above figure, which shows the MSLP in hPa (shaded) and 700 hPa winds in m/s (vectors) over Indian sub-continent. The low pressure associated with continental heating (*heat low*) is prominent and an important semi-permanent system that drives the monsoon (Rao, Y. P. 1976). The heat low establishes over NW parts of India and adjoining 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 the central India. The pre-monsoon hot weather gets severe at times when the hot and dry northwesterlies penetrate deep into the peninsula and persist for several days. During May 2016, similar conditions caused severe heatwave conditions over parts of Maharastra, Telangana and Odisha.

4.3.1 Case-I Heatwaves on 11th April 2016

As per the IMD reports (Climate Diagnostics Bulletin of India, April 2016), the heatwave conditions prevailed over parts of central peninsular and east India during the second week of the April. It took a toll of over 200 lives (Table-2) from central and peninsular India during the April month. Observed and forecast *Tmax* valid for 11th April 2016 is shown for NCUM (Figure 5) and NEPS (Figure 6). The spatial distributions of *Tmax* shows prevailing heat-waves over Odisha, AP, Telangana, and some parts of Maharashtra on 11th April 2016. The observation shows more than 40°C spread over east UP, Bihar, West Bengal,east MP,Jharkhand, Chhattishgarh, Odisha, Maharashtra and Some parts of Karnataka and Tamil Nadu. In the NCUM forecast, on other hand showing marginally wider regions upto day-9 due to warm bias in the model and on the contrary NEPS forescasts also showing Ó40°C wider regions upto day-9 but marginally less than the NCUM frorecasts. Apart from the warm bias both the model forecast is showing cold bias in north-northeast of J&K. Hence theNEPS is better in predicting the extremes of heatwaves up to day -9 then the NCUM.

4.3.2 Case-II Heatwaves on 21st May 2016

The severe heatwave conditions developed and intensified over parts of northwest India entire third week of May-2016 and persisted over parts of central and north peninsular India some stations of West Rajasthan temperature observed Ó50°C viz. Barmer, Bikaner, Ganganagar, Jaisalmer & Jodhpur and observed severe heatwave conditions for 4 to 5 days in succession from 17th to 21st May-2016. The spatial distributions of NCUM & NEPS forecast *Tmax* with of observed IMD *Tmax* prevailing heat-waves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21st May 2016 is shown in Figure 7 & 8. Both the models deterministic and ensemble able to predict the extreme temperature (*Tmax* > 48°C) over west Rajasthan up day-3 only. However, the NCUM is predicting more wide-spreading *Tmax* > 46°C, over Rajasthan, MP, UP, Delhi, Haryana, Punjab and parts of Maharashtra all days forecast.

The H-K scores of the maximum temperature (*Tmax*) between the range 30-42 °C, constructed as box and whiskers for both NCUM and NEPS, indicate towards better performance of the ensemble based forecast as compared to the deterministic one. Interestingly, the forecast score does not fade away with the 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 (Figure 9).

Similar observations can be made from the ETS plots (Figure 10).The most obvious finding to emerge from the box and whiskers 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 *Tmax* thresholds (between 30 and 42°C), NEPS mean stands above the NCUM mean. The same observation holds during all the illustrated forecasts (Day1, 3, 5, 7, and 9). The scores falling under the 25% in the case of ensemble based forecast are either similar or lie little above the deterministic forecast unlike the values underlying 75% which in the NEPS case are markedly higher than that of the NCUM's.

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 low 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 forecast. Similarly, a comparatively higher score on Day-5, 9 and Day-7 respectively (Figure 11).

POD: 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 40-42°C where NCUM shows better performance (Figure 12.)

SEDI: At higher temperature ranges, representing rare events, the performance of NEPS and NCUM can be clearly seen from the SEDI score plot (Figure 13). We can notice a considerable difference between the performance of the two techniques for extreme events lying between 40 and 42 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 are listed in table 3.

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 model (NWP). One of the several predicted interactions could be climatologically extreme and are hence more difficult to predict. A small variation in the intensity, timing, and position of such anomalies could lead to a large difference in their prediction growth in time. Thus, despite the efficiency of 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 in predicting the extremes of heat-wave events than the deterministic forecast over all Indian states up to day-9.

1) The ensemble forecast provides appreciable forecasts on all the days and is most reliable after the Day-2 forecast. This characteristic is more pronounced for extreme events than for the less extreme events where 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 too. Thus, we need to understand the model performance at all the concerned ranges and based upon those verification results, employ the ensemble forecast accordingly for operational purposes.
2) Our forecasts were obtained for the current summer season in India, MAM and since, the severe events are rare in nature it limits the sample size for the ensemble forecast and thus pose a challenge for the efficient forecasting 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°C) can reliably be predicted for Day-2 onwards with less false alarms as 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 warm bias.

3) Comparatively, low efficiency of the ensemble based prediction on a shorter time scales (< Day-2) propose that the ensemble prediction may need a longer duration of time for the perturbation growth. This observation would prove to be an important aspect to consider for the future evolution of the ensemble based forecasting. If this hypothesis is true, for the short-range forecasts, ensemble based prediction could fall at the back of other methods like moist SV's optimization (Coutinho et al., 2004), the ETKF (12, 13). However, over a medium range forecast and for the 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 modes is available only from 2016. Ensemble based forecasts in realtime using the NEPS started in November 2015 at NCMRWF. For a robust and conclusive result it is necessary that the study be based on higher number of cases. This will be carried out in future.

The temperature data from the stations distribution are discussed in this paper which is used to obtain the gridded Tmax and Tmin 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 stations data network is sparse 395 and often there are missing values. Gridded data field provides a continuous and gap free data to work with.

Extreme events like heat waves are rare in nature and here we provided a general view of the two particular heat wave events (11 April & 21 May). From our experience as well as the forecast for the post heat wave 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 figure (5, 6, 7 & 8). The overall prediction of warm conditions is nicely predicted but at closer lead times, the events are better predicted. 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 a reasonable skill. So, overall the NEPS specifically, has a good skill in predicting the extreme event and is relatively robust.

Acknowledgments: The authors are grateful to India Meteorological Department (IMD) for providing the gridded observed *Tmax* and *Tmin* 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 manuscript. Thanks are due for anonymous reviewers for their comments and suggestions which have helped in revising the manuscript.

References:

Alexander, L.V. , X. Zhang, T. C. Peterson, J. Caesar, B. Gleason, A. Klein Tank, M. Haylock, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, P. Ambenje. Rupa Kumar, K., Revadekar, J.V., Griffiths, G., Vincent, L., Stephenson,D., Burn,

J., E. Aguilar, M. Brunet, M. Taylor, M. New, P. Zhai, M. Rusticucci, J. L. Vazquez-Aguirre, Global observed changes in daily climate extremes of temperature and precipitation, Journal of Geophysical Research - Atmospheres, Vol. 111 (D05109), doi:10.1029/2005JD006290 2006.

Bishop, C. H., Etherton, B. J. , and Majumdar, S. J. : Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. Mon. Wea. Rev, **129,** 420ï 436, 2001.

Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B. and Beare, S. E. (2008), The MOGREPS short-range ensemble prediction system. Q.J.R. Meteorol. Soc., 134: 703ï 722. doi:10.1002/qj.234

Buizza, R., Houtekamer, P.L., Toth, Z., Pellerin, G., Wei, M., Yuejian, Z.: A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems. Mon. Weather Rev. **133**: 1076ï 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, 2006.

Climate Change 2013: The Physical Science Basis, Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)

Climate Diagnostics Bulletin of India, April 2016, Near Real - Time Analyses National Climate Centre, Pune

Climate Diagnostics Bulletin of India, March 2016, Near Real - Time Analyses National Climate Centre, Pune.

Climate Diagnostics Bulletin of India, May 2016, Near Real - Time Analyses National Climate Centre, Pune

Coutinho, M. M., Hoskins B. J., and Buizza R.: The influence of physical processes on extratropical singular vectors. J. Atmos. Sci, **61,** 195ï 209, 2004.

Hartmann, D.L., A.M.G. Klein Tank, M. Rusticucci, L.V. Alexander, S. Brönnimann, Y. Charabi, F.J. Dentener, E.J. Dlugokencky, D.R. Easterling, A. Kaplan, B.J. Soden, P.W. Thorne, M. Wild and P.M. Zhai, 2013: Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 159ï 254, doi:10.1017/CBO9781107415324.008.

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. Climate, **16**, 3560ï 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.Climate Res., 42,89-102.,doi: 10.3354.cr00857, 2010

Legg, T. P. and Mylne, K. R:Early Warnings of Severe Weather from Ensemble Forecast Information Weather and Forecasting, 19, 891-906, 2004.

Lorenz, E. N.: The predictability of a flow which possesses many scales of motion, Tellus, 21**,** 289ï 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. Climate Dyn., 17, 753ï 775, 2001.

Mehdi, S, S., and B. S. Dehkale 2017: Long term climatology and trends of heat and cold waves over southern Bihar, India. J. Earth Syst. Sci., 125, (8), 1557ï 1567, DOI 10.1007/s12040-016-0762-2,

Molteni, F., R. Buizza, Palmer, T. N. and Petroliagis, T.:The ECMWF Ensemble Prediction System: Methodology and validation. Quart. J. Roy. Meteor. Soc, **122,** 73ï 119, 1996.

Mureau, R., Molteni, F. and Palmer, T. N.: Ensemble prediction using dynamically conditioned perturbations. Quart. J. Roy. Meteor. Soc, **119,** 299ï 323, 1993.

New, M., Hulme, M. and Jones, P.: Representing twentieth century spaceï time climate variability. Part II: Development of 1901ï 96 monthly grids of terrestrial surface climate. J. Climate, **13**, 2217ï 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. Cycles, **10**, 752ï 782, 1996.

Rajeevan, M., Jyoti, B., Kale, J.D., Lal, B.: Development of a high resolution daily gridded rainfall data for Indian region. *IMD Meterological Monograph. No: Climatology 22, Pune, India.* 27 pp, 2005.

Rao, Y.P.: Southwest Monsoon, Meteorological Monograph (Synoptic Meteorology), No. 1/1976, India Meteorological Department, New Delhi, 366pp

Rohini, P., M. Rajeevan and A. K. Srivastava 2016: On the variability and increasing trend of Heat Waves over India. Sci. Rep. 6, 26153; doi: 10.1038/srep26153 (2016).

Sarkar, A., Chakraborty, P., John, P. G., Rajagopal, E. N.: Implementation of Unified Model Based Ensemble Prediction System at NCMRWF (NEPS), ***NMRF/TR/02/2016***, 2016.

Sheikh MM, Manzoor N, Ashraf J, Adnan M, Collins D, Hameed S, Manton MJ, Ahmed AU, Baidya SK, Borgaonkar HP, Islam N, Jayasinghearachchi D, Kothawale DR, Premalal KHMS, Revadekar JV, and Shrestha ML. 2015. Trends in extreme daily rainfall and temperature indices over South Asia, June 2015, *International Journal of Climatology*, 35, DOI:10.1002/joc.4081, 1625-1637

Shepard, D.: A two-dimensional interpolation function for irregularly spaced data. Proc. 23rd National Conf., New York, NY, Association for Computing Machinery, 517ï 524, 1968.

Srivastava, A. K., Rajeevan, M. and Kshirsagar, S. R.: Development of a high resolution daily gridded temperature data set (1969ï 2005) for the Indian region. *Atmos Science Lett*. 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 E. Kalnay: Ensemble forecasting at NCEP and the breeding method. Mon. Wea. Rev, **125,** 3297ï 3319, 1997.

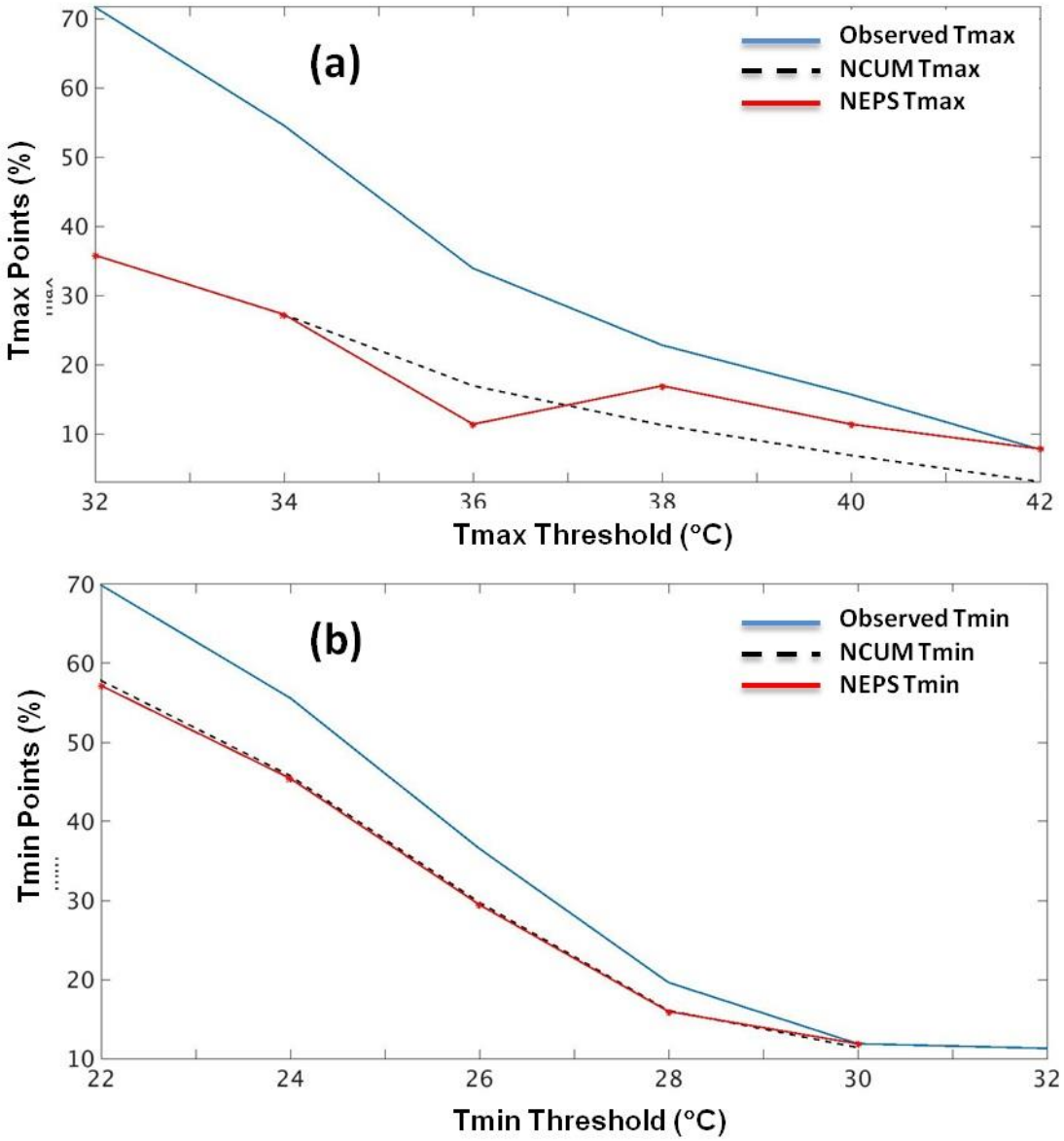

Figure 1. Frequency distribution of observed, and forecast (NCUM and NEPS) (a) *Tmax* (°C) and (b) *Tmin* (°C) over India during March-May 2016. (NCUM stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System) [Observed data from India Meteorological Department and forecast data from NCMRWF]

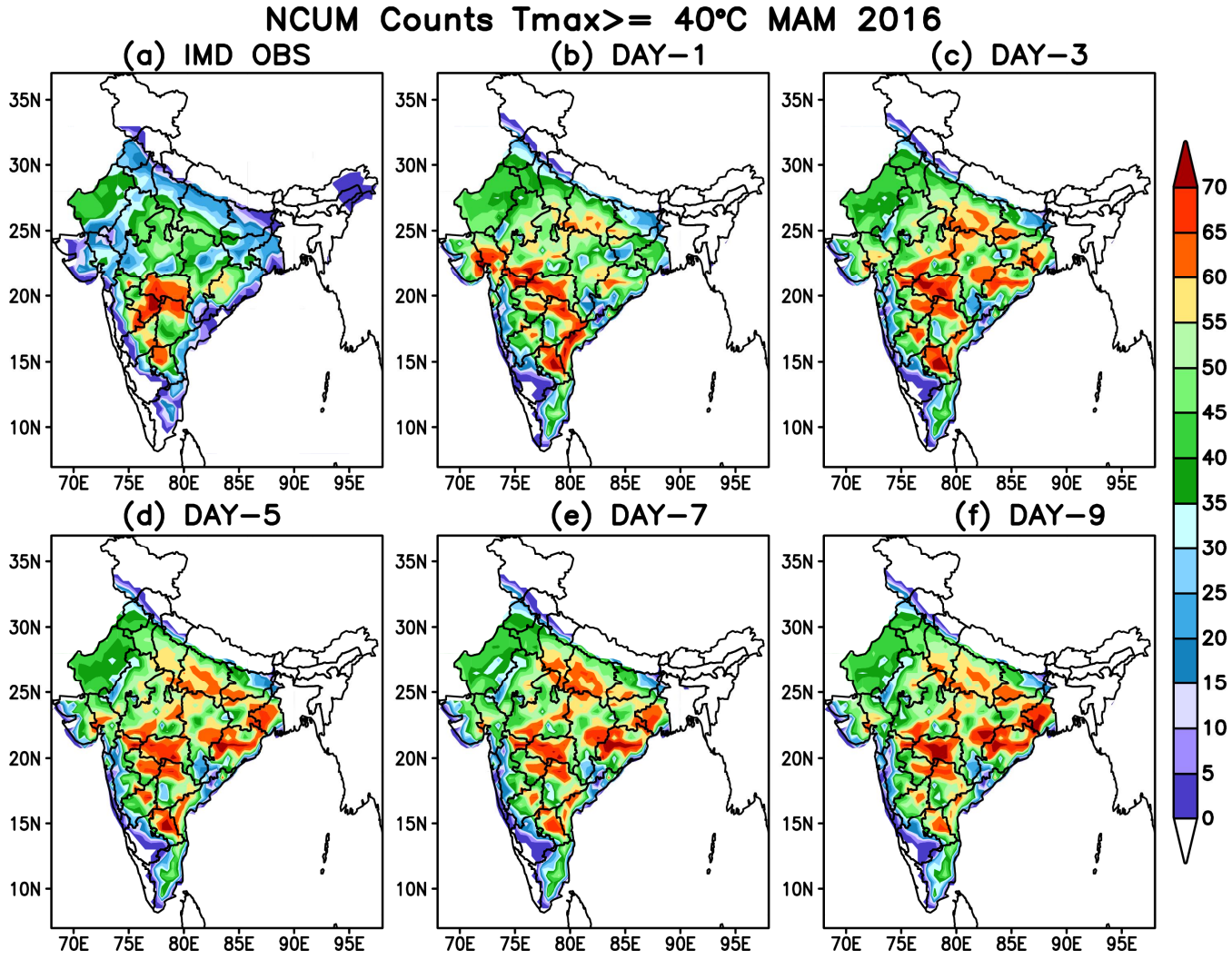

Figure 2. Spatial distribution of observed and NCUM forecasts number of days with *Tmax* Ó 40°C during the period of March to May 2016. (NCUM stand for NCMRWF Unified Model) [Observed data from India Meteorological Department and forecast data from NCMRWF]

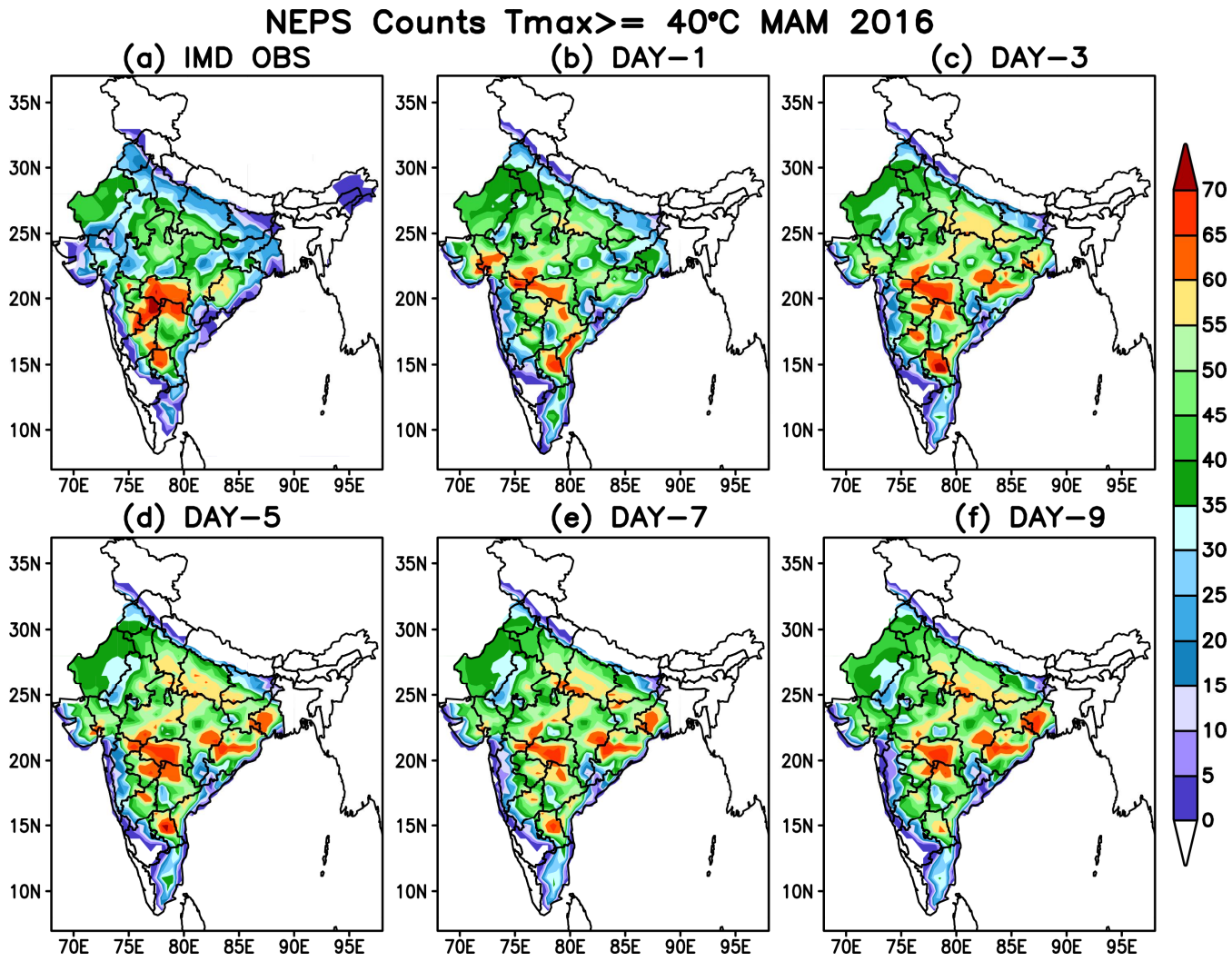

Figure 3. Spatial distribution of observed and NEPS forecasts number of days with *Tmax* Ó40°C during the period of March to May 2016. (NEPS stands for NCMRWF Ensemble Prediction System) [Observed data from India Meteorological Department and forecast data from NCMRWF]

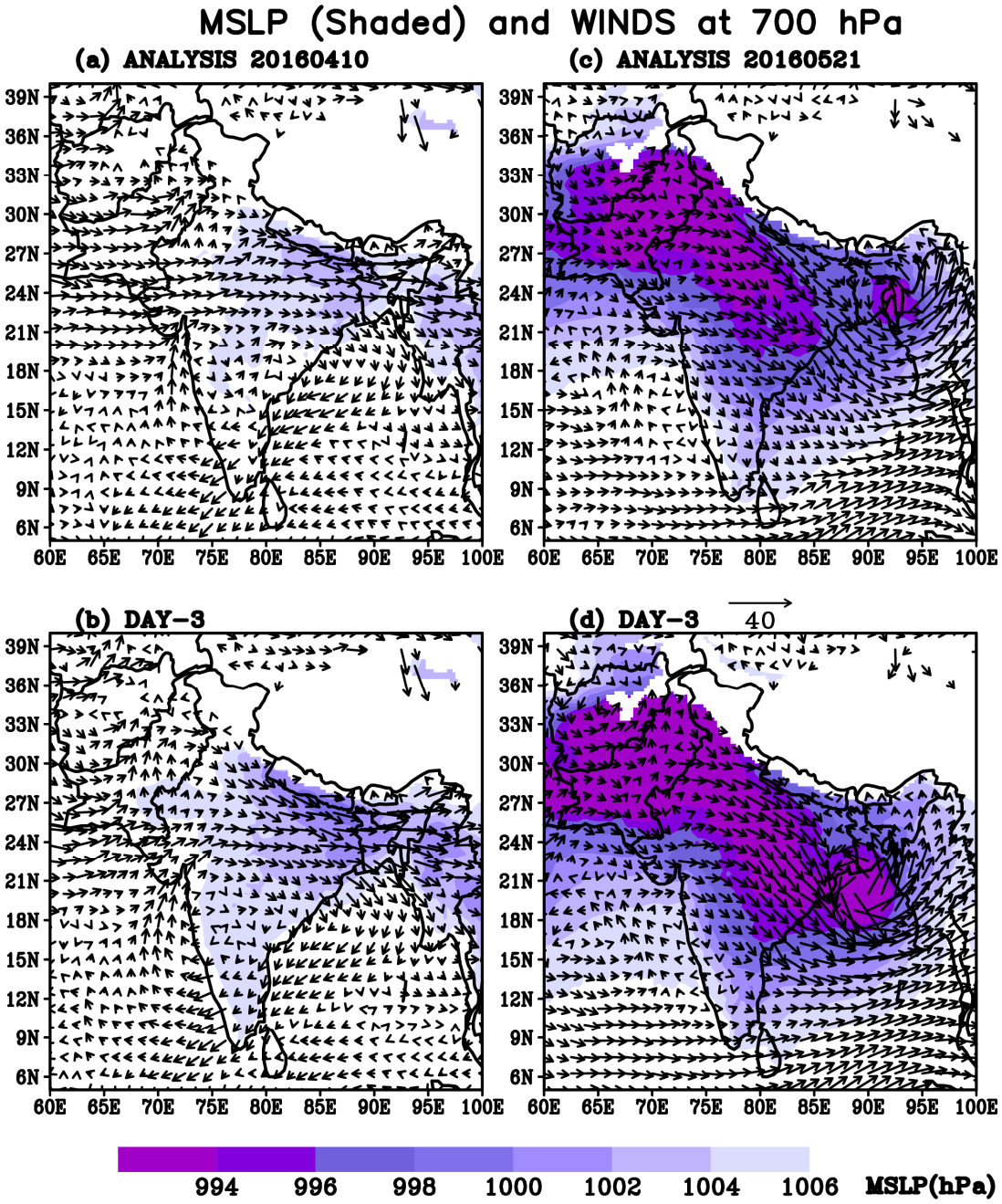

Figure 4. Mean Sea Level Pressure (MSLP) shaded and winds at 700 hPa showing heat low in the (a) Analysis of 20160410 (b) Day 3 forecast valid for 20160410 (c) Analysis of 20160521 (d) Day 3 forecast valid for 20160521. The analysis and the forecasts are based on the NCMRWF Unified Model.

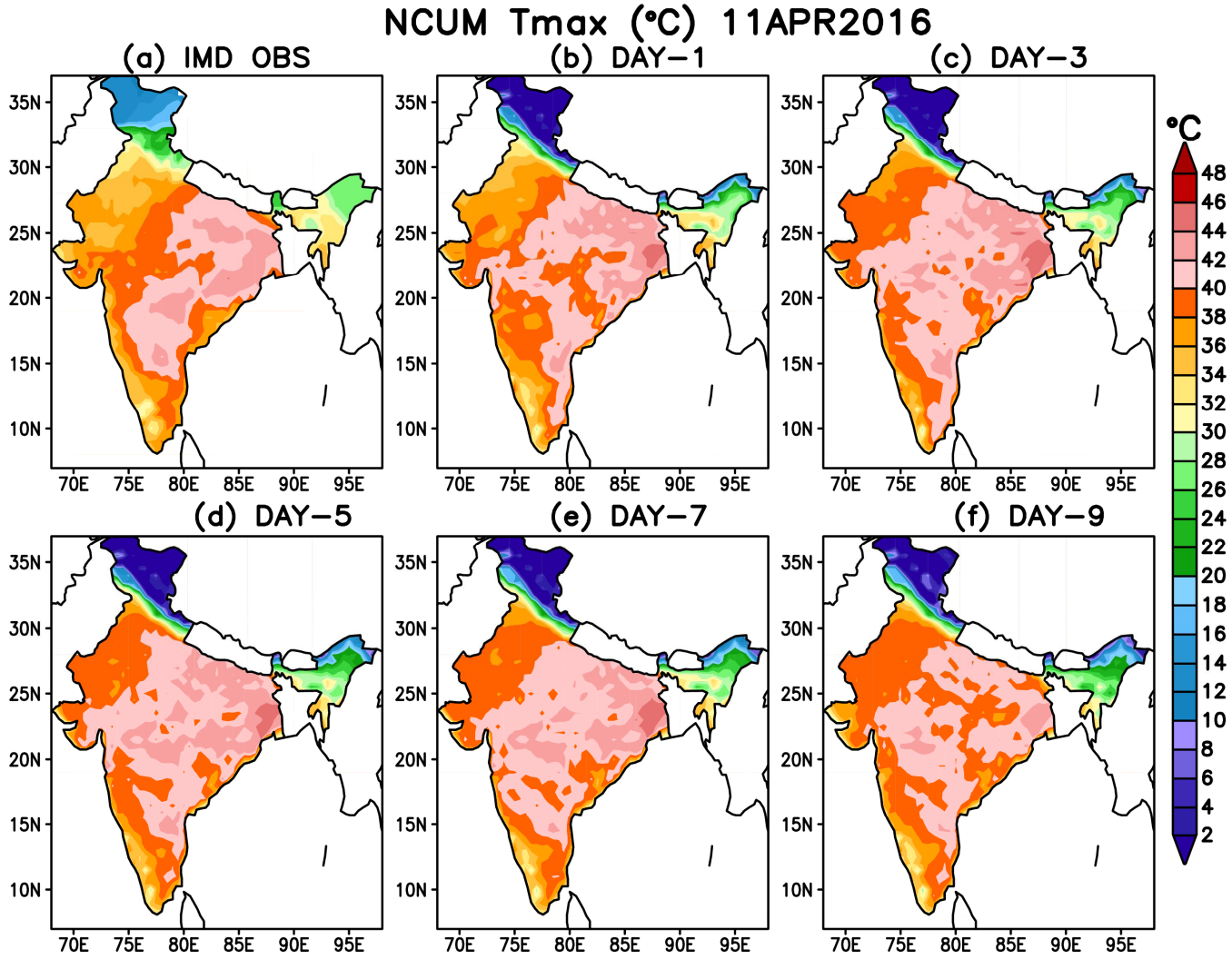

Figure 5. Spatial distributions of Observed *Tmax* and NCUM forecast *Tmax* prevailing heat-waves over, MP, Odisha, AP, Telangana and some parts of Maharashtra on 11th April 2016. (NCUM stand for NCMRWF Unified Model). [Observed data from India Meteorological Department and forecast data from NCMRWF]

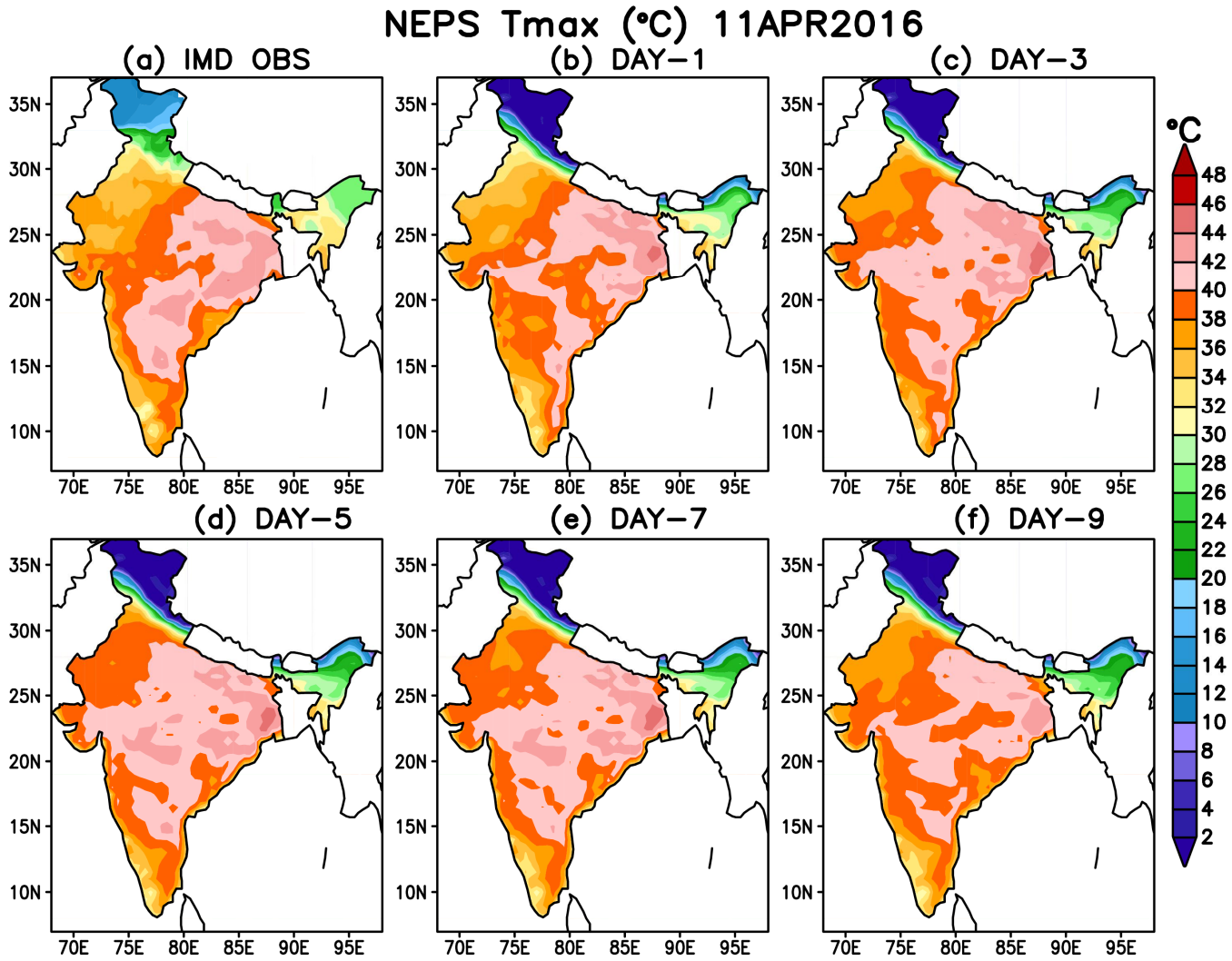

Figure 6. Spatial distributions of Observed *Tmax* and NEPS forecast *Tmax* prevailing heat-waves over, MP, OdiZZsha, AP, Telangana and some parts of Maharashtra on 11th April 2016. (NEPS stands for NCMRWF Ensemble Prediction System). [Observed data from India Meteorological Department and forecast data from NCMRWF]

5

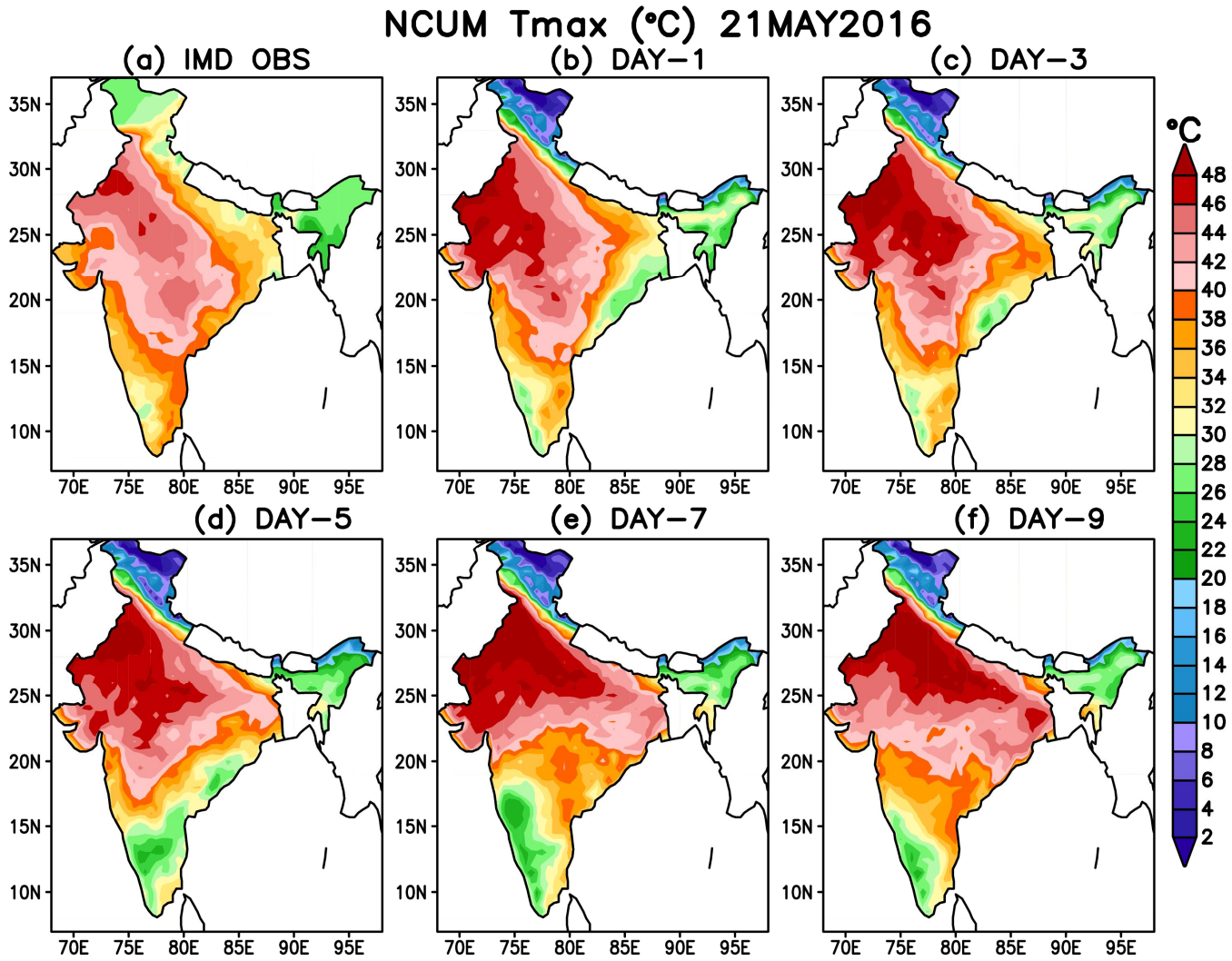

Figure 7. Spatial distributions of Observed *Tmax* and NCUM forecast *Tmax* prevailing heat-waves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21st May 2016. (NCUM stand for NCMRWF Unified Model). [Observed data from India Meteorological Department and forecast data from NCMRWF]

5

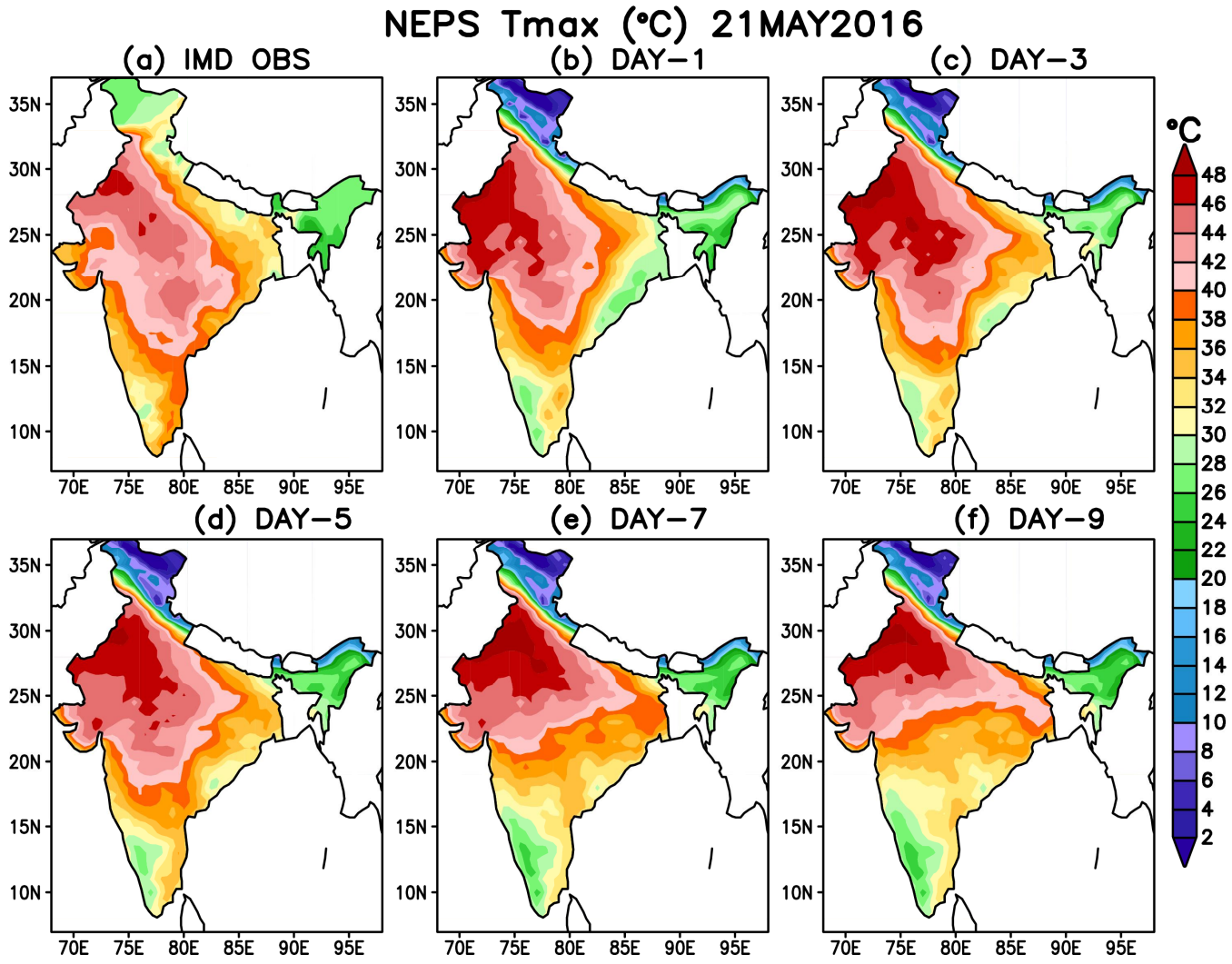

Figure 8. Spatial distributions of Observed *Tmax* and NEPS forecast *Tmax* prevailing heat-waves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21st May 2016. (NEPS stands for NCMRWF Ensemble Prediction System). [Observed data from India Meteorological Department and forecast data from NCMRWF]

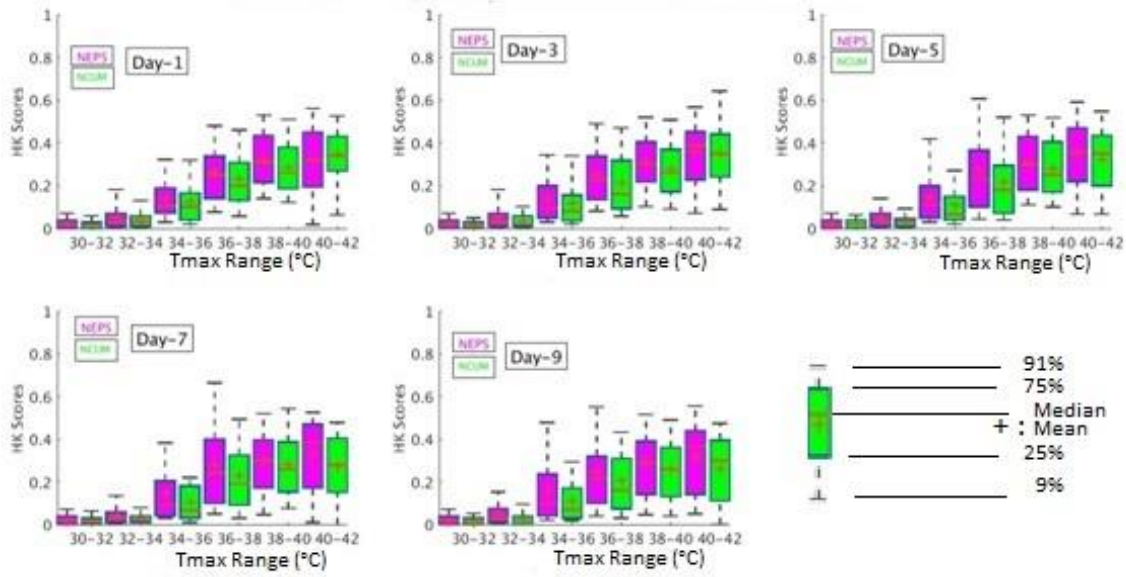

Figure 9. Box plots for HK scores for different temperature ranges (*Tmax*) NCUM and NEPS form March to May 2016. (NCUM stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System)

5

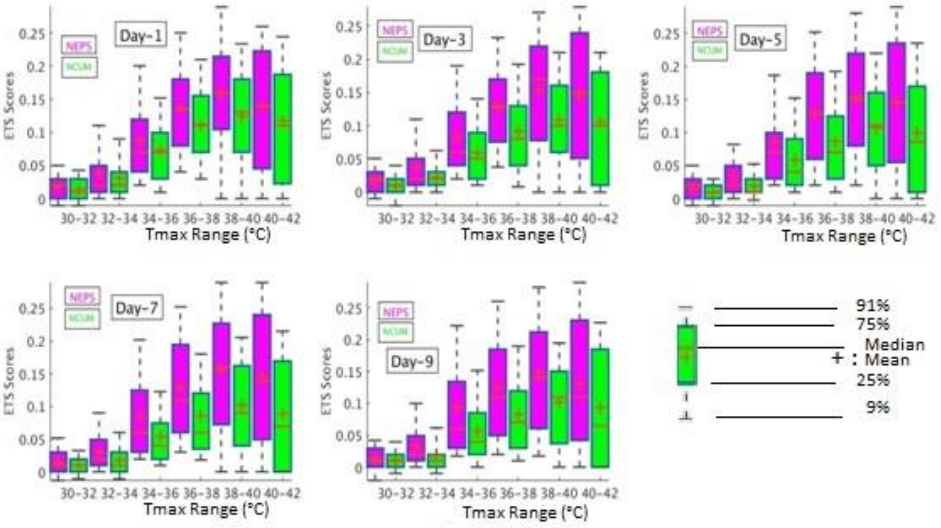

Figure 10. Box plots for Equitable Threat Score (ETS) for NCUM and NEPS form March to May 2016. (NCUM stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System)

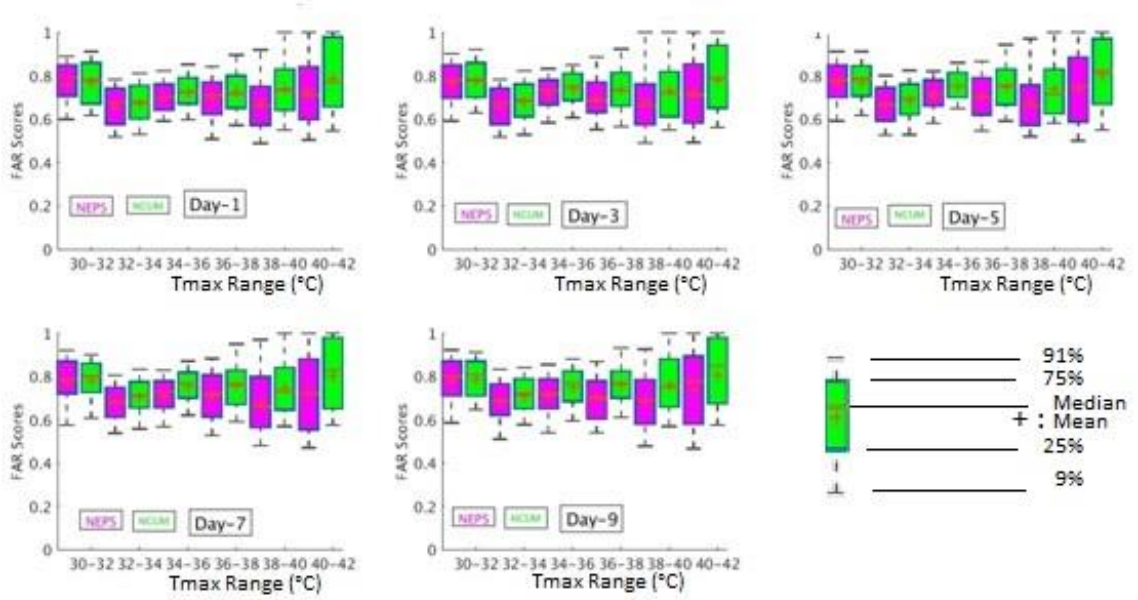

Figure 11. Box plots for False Alarm Ratio (FAR) for NCUM and NEPS form March to May 2016. (NCUM stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System)

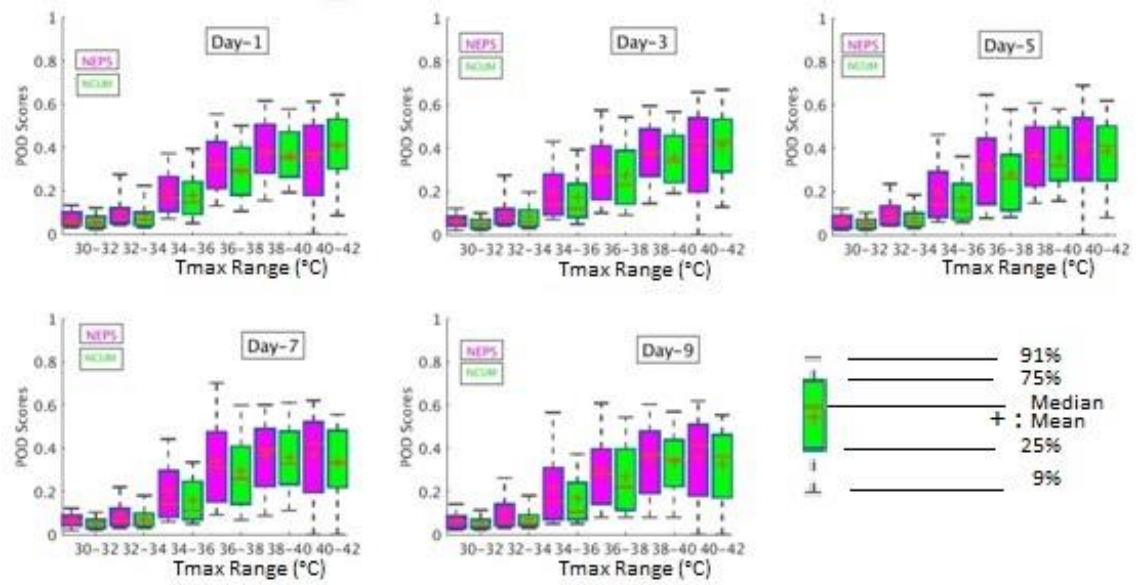

Figure 12. Box plots for Probability of Detection (POD) for NCUM and NEPS form March to May 2016. (NCUM stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System)

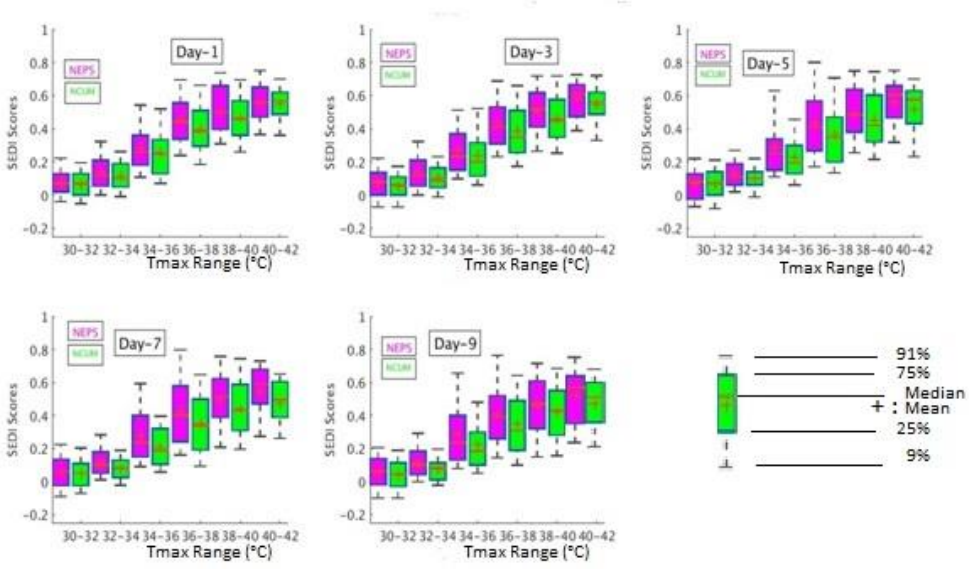

Figure 13. Box plots for Symmetric Extrernal Dependence Index (SEDI)) for NCUM and NEPS form March to May 2016. (NCUM stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System)

Table 1. List of Abbreviations

EPS	**E**nsemble **P**rediction **S**ystems
NCMRWF	**N**ational **C**entre **for M**edium **R**ange **W**eather **F**orecasting
NEPS	**NCMRWF E**nsemble **P**rediction **S**ystem
NCUM	**NCMRWF U**nified **M**odel
NWP	**N**umerical **W**eather **P**rediction
MAM	**M**arch, **A**pril and **M**ay
Tmax	**M**aximum **T**emperature
Tmin	**M**inimum **T**emperature
IMD	**I**ndian **M**eteorological **D**epartment
NDC	**N**ational **D**ata **C**entre
ETKF	**E**nsemble **T**ransform **K**alman **F**ilter
POD	**P**robability **O**f **D**etection
FAR	**F**alse **a**larm **r**atio
HK	**H**anssen and **K**uipers
ETS	**E**quitable **T**hreat **S**core
SEDI	**S**ymmetric **E**xtrernal **D**ependence **I**ndex
MP	**M**adhya **P**radesh
UP	**U**ttar **P**radesh
AP	**A**ndhra **P**radesh

Table 2. Casualities reported during March to May 2016 due to prevailing 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
	Gujrat	39	
	Maharashtra	34	

(Data: Climate Diagnostic Bulletin of India, March 2016, April 2016 and May 2016,India Meteorological Department)

Table 3. Monthwise verification scores for *Tmax* > 40°C for NCUM and NEPS forecast with India Meteorological Department (IMD) observed temperature.

		NCUM					**NEPS**				
Month	Score	Day 1	Day 3	Day 5	Day 7	Day 9	Day 1	Day 3	Day 5	Day 7	Day 9
MAR	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
APR	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 stand for NCMRWF Unified Model and NEPS stands for NCMRWF Ensemble Prediction System)