Title: **Verification of Pre-Monsoon Temperature Forecasts over India during 2016 with focus on Heat Wave Prediction**

NHESS-2016

We are thankful to the editor and the reviewer for their helpful suggestions which have helped us to improve the quality of the paper to a great extent. We have tried to incorporate as many of their suggestions as possible.

Reviewer #1 Specific Comments

1. *The verifications are based on extreme heat events for only one year, the authors could consider a few more years to support their results*

   Reply 1: The suggestion by the reviewer is very valid. 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 NCMRWF. For robust and conclusive results it is necessary that the study be based higher number of cases. This will be carried out in future.

2. *The authors have used gridded data. This would have definitely suppressed the extreme station temperature values*

   Reply 2: The temperature data from the stations distribution are discussed in the 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.

Reviewer #1 Technical corrections

3. *On Page 2, Kothawale (2005) and IPCC (2013) have been cited but not listed under references*

   Reply 3: Thanks for pointing out this error. The first reference Kothawale et al (2005) is removed and is replaced by more recent study by same author Kothawale et al (2010) on *page 2 line 20*. The IPCC(2013) is included in the references on page 12 line 4.

   Additionally, the other reference of Arora et al (2009) is eliminated in the text on page 2(*line 23*) and in the list of references.

4. *Page 6 : The y-axis in each of the figs (is this applicable to all figures ?)*

Reply 4: No. This refers to two panels in Figure 1. The text on page 6 near line 26 has been modified as-

The panels in Fig. 1a,b show the observed and forecast (Day-3) frequency distribution for *Tmax* and *Tmin*.

5. *Page 7: Line 23, mention is made of Table-1, but this table lists the abbreviations used*

   Reply 5: The text is corrected on page 8 line 9.

6. *Page 7 last 3 lines: Authors mention spatial distribution ----- but Fig. 8 and 9 show box plots*

   Reply 6: The discussion on two heatwave cases presented in sections 4.3.1 (and 4.3.2) are based on Figs 5,6 (and 7,8) respectively. The figure numbers are correctly represented in the revised manuscript on page 8 in line 31 and on page 9 line 12.

7. *Page 8: first Fig 11 is referred, then 10, then 9 and the 8 ??? Please follow sequence*

   Reply 7: The text on page 9 (line 12 and 3) is corrected to refer to Figs 7&8.

8. *Page 8 line 7 Mention of ETS plots (Fig.10) is made but this fig contains plots for FAR*

   Reply 8: In the revised manuscript on page 9 line 19, the ETS is discussed using Fig 10.

9. *Similarly Fig. 9 are ETS plots but in text something else is mentioned (page 8)*

   Reply 9: The change is reflected on the page 9 line 18, now it is HK scores.

10. *Page 8 line 23: Mention is made of SEDI score plot – fig number not mentioned*

    Reply 10: In the revised manuscript the Fig 13 showing SEDI is correctly referred on page 10 line 5.

11. *Several repetitions*
    Reply 12: Thanks for bringing this to our notice. We have taken extra care to avoid the repetitions in the manuscript.

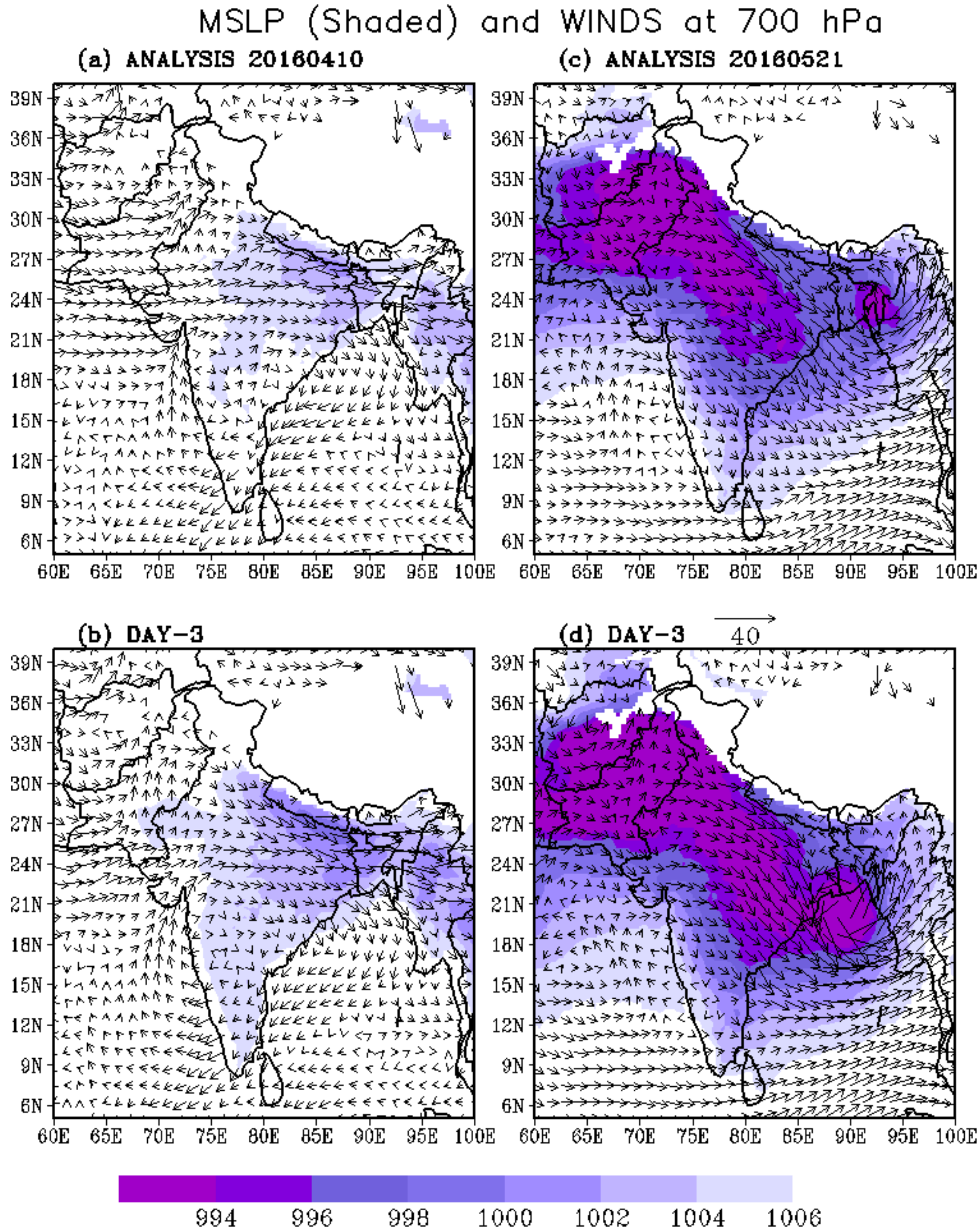
Reviewer #2 Specific Comments

1. *How much of the skill in predicting the heatwaves comes from persisting a heatwave already present in the initial conditions? How does the model perform when the heatwave evolves within the forecast range (e.g. Beyond days 2-3).*

Reply 1: 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.

2. *Synoptic evolution in heatwave case studies - It would have been good to also see the prevailing synoptic conditions and larger-scale flow conditions associated with these heatwaves (e.g. MSLP or low level winds) in both observations/analysis and deterministic and EPS (ensemble mean) forecasts. Perhaps also the time series of temperatures (deterministic and EPS members (at day 2, 5, 7), and Observations) over a specific region (e.g. Rajasthan) during one of the heatwave events would also give the reader a more physical feel for the predictability that is difficult to get just from the verification metrics alone. This is achieved to some extent by snapshots in Figs 4-7.*

Reply 2: Thank you for your insightful comment. As per your suggestion, we are adding a figure illustrating synoptic systems (both, MSLP & low-level winds) for the heat wave event considered in the present work (Dated: 20160521). We can see that the monsoon heat low shown by low MSLP values over NW Indian and adjoining Pakistan is an important semi-permanent system during the pre-monsoon season. The low MSLP values and high temperatures associated with that create strong land-sea temperature and pressure gradient in the lower troposphere which is crucial for onset and advance of monsoon. As can be seen in the figure below, during this pre-monsoon month, the low pressure is accompanied by the westerly and north-westerly winds and heat waves over the Indian and the neighboring countries. In the figure, we see it mainly occurring over the central India.



3. *Could the authors provide more detail on how the various categorical scores are calculated for the EPS. Are the scores based on the ensemble mean vs. observations or do they use all 44 individual ensemble members to construct a score?*

Reply 3: Computation of the scores is based on the ensemble mean (44 members). An ensemble mean is first computed from each member which is then treated as another model and is further used to obtain the scores. It is known that the ensemble mean has a higher skill than the deterministic forecast especially in the upper air fields (500 hPa) (cite: Ton Hamil et. al) and similar observation is justifiable for the low-level fields as well (fig: score plots).

4. *Page 6, lines 11-12 - "Deterministic forecast hardly shows any variation in either of the considered days and illustrates quasi-stationary characteristics of the deterministic forecast from Day-1 through Day-10 forecast". I don't really understand this or know which figure/result it is referencing. Can the authors clarify.*

Reply 4: Suitably modified and change is reflected on page 7 line 3.

5. *Figure 1 suggests that the deterministic forecasts (and to a lesser extent the EPS) underpredicts the frequency of heatwaves compared to observations over Indian land points. This appears to be inconsistent with later discussions around figures 2 and 3 which suggest that the deterministic and EPS over predict the number of heatwave days (>40) compared to the Observations? Can the authors explain this inconsistency?*

Reply 5: The figure was prepared to choose ranges of the verification metrics and does not serve a purpose to indicate any sort of over or under prediction. This is because the figure represents the "fraction" of the total number of days and the grid points (i.e. counts/92X2686 (days X grids)). The denominator includes all the grid points with or without the Tmax > 40C.

6. *In Fig 6. the NCUM and to lesser extent the NEPS forecasts show a growing warm bias over NW India with FC range. Do the authors have any physical explanation for this bias (e.g. soil moisture initialisation, model systematic errors in circulation?)*

Reply 6: In 21 May case, warming is increasing drastically for both, NEPS and NCUM. This is not based on one initial condition and includes several different initial conditions. We have error growth and warm bias in the present study the impact of soil moisture feedback is not attempted. The land surface scheme involves soil moisture data assimilation using extended Kalman Filter technique. The soil moisture analysis prepared based on screen level humidity and temperature observations and ASCAT surface soil wetness observations from MetOP-A satellite (C- band, Level2 product). Systematic errors in circulation have been widely and extensively studied and documented for monsoon (JJAS) season. Typically for the pre-monsoon conditions such detailed analysis would be useful and will be taken up as a follow up of this study.

7. *Predictability of heatwaves - In the summary the authors state "Unless the atmosphere is in a highly predictable state, we should not expect an ensemble to forecast extreme events with a high probability". It would be good to see some discussion of whether these heatwave events are highly predictable (e.g. links to large scale flow anomalies), given they seemed to be predictable several days ahead? Was the ensemble spread of Tmax smaller or larger than normal in these heatwaves?*

Reply 7: As stated earlier in response to the second point, the extreme events are rare which offers' a small sample size, thereby making their predictability and verification difficult as such. However, signature of the events are noticeable in the synoptic systems, a few days ahead of the event (ex. Wind patterns and MSLP, fig).

8. *Are there plans to use these EPS predictions of heatwaves to give warnings to the public? Perhaps some discussion in summary?*

Reply 8: This preliminary study indicates potential skill in forecasting heat wave conditions. With the use of suitable calibration/downscaling and bias correction methods, these forecasts of heat waves could be useful for the forecasters at operational agency Indian Meteorological Department.

Technical corrections

This manuscript suffers from a lot of technical errors and inconsistencies that make it difficult to read. Some of these relate to English usage but many are just errors that are easily corrected. I have listed the main errors below

1. *A number of variations on the word "heatwave" appear in the manuscript (Heat wave, Heat Wave, heat wave and heatwave). Suggest authors provide a consistent spelling (e.g. heatwave).*
   Reply 1: Thanks for bringing this to our notice; we have incorporated a consistent word *"heatwave"* in the entire revised manuscripts.
2. *Authors also refer to "deterministic models" and "ensemble models". This should be replaced with "deterministic forecasts" and "ensemble forecasts" throughout the text as both actually use the same model (UM).*
   Reply 2: We have replaced *"deterministic models" and "ensemble models"* with *"deterministic forecasts" and "ensemble forecasts"* throughout the text.
3. *Page 1, Line 9 removed "the" in this sentence - here we investigate extreme events (heatwaves)*
   Reply 3: We have removed "the" from the sentence. This change can be seen on page 1 and line 9.
4. *Page 1, Line 22 - replace "...prediction the extreme events" with "...prediction of extreme events"* **Reply 4:** We have replaced "*…prediction the extreme events" with "...prediction of extreme events"* in page 1, line 22.
5. *Page 1, Line 22 - I don't understand the sentence "Extreme Weather events comprehend non-linear interactions..."*
   Reply 5: The said sentence is replaced as follows in the first para in Introduction on page 1 lines 23-25.

   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.

6. *Page 1, Line 30 - simplify this sentence "Based on multiple perturbed initial conditions, ensemble approach samples the errors in the initial..." to "It is based on*
   Reply 6: At page 1, Line 32 has been simplified and added "It is" in the beginning of the sentence.
7. *Page 2, line 1 - remove the first reference to Sarkar et. al., 2009, as it is repetitive.*
   Reply 7: The reference 'Sarkar et. Al., 2009' has removed from the text on page 2 line 4.
8. *Page 2 line 2 - Replace "Met office" with "Met Office" 2*
   Reply 8: "Met office" replaced with "Met Office" on page 2 line 5.
9. *Page 2 line 14 - replace "0.85 0°C" with "0.85 °C"*
   Reply 9: In the text we have replaced "0.85 0°C" with "0.85 °C" in line 18 on page 2.
10. *Page 2 line 15 - don't understand how Molteni et. al. (1996) could be used as refrence for a warming trend covering 1880-2012!*
    Reply 10: We have removed *Molteni et. al. (1996)* from the text. It was inadvertently typed on page 2 line 19.
11. *Page 2 line 17 - assume that the annual mean temperature of 0.42 C per 100 years refers to the globally averaged temperatures. This should be made clear.*
    Reply 11: We have elaborated in the text and updated with the new reference "*Kothawale et. al. 2010*" in line 20 on page 2.
12. *Page 2 line 21 - this paragraph begins with a sentence "Another study…" but the reference at the end of the sentence is Arora et. al. (2009) which was the same study discussed in the previous paragraph.*

Reply 12: Same as reply 11.

13. *Page 2 Line 24 - not sure what "recently reiterated" means?*
    Reply 13: Same as reply 11
14. *Page 2 line 28 - "sales" should read "scales"*
    Reply 14: We have replaced sales with scales now in line 3 on page 3.
15. *Page 2 line 29 - the sentence "...using ensemble forecast forecasts both, deterministic and ensemble forecasting." is very convoluted, can I suggest "...using both ensemble and deterministic forecasts"*
    Reply 15: The sentence is modified as per the suggestion in place of the *"...using ensemble forecast forecasts both, deterministic and ensemble forecasting."* to *"...using both ensemble and deterministic forecasts"* in the text in line 4 on page 3.
16. *Page 3 line 9 - delete "and" in the following "...adopt and the most compatible score type"*
    Reply 16: We have deleted "and" from the sentence "...adopt and the most compatible score type" from the text in line 18 on page 3.
17. *Page 3 line 11 - this sentence is very repetitive.*
    Reply 17: We have removed the repetitive sentence from the revised manuscripts.
18. *Page 3 line 23 - remove "... which was 1°x1° resolution a few years earlier over Indian land area." As it is irrelevant for this study.*
    Reply 18: We have removed the sentence on page 4 in line 1 "... which was 1°x1° resolution a few years earlier over Indian land area." from the revised manuscripts on page 4 line 1.
19. *Page 3 Line 32 - replace "operational NCMRWF" with "operational at NCMRWF*
    Reply 19: In the text we have replaced "operational NCMRWF" with "operational at NCMRWF"
20. *Page 4 line 8 - replace "...MET Office" with "...Met Office MOGREPS system (Bowler et. al. 2008)" where reference is 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: 703–722. doi:10.1002/qj.234*
    Reply 20: "...MET Office" replaced (on page 4 line 25) with "...Met Office MOGREPS system (Bowler et. al. 2008)" and reference is 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: 703–722. doi:10.1002/qj.234 added in the reference list on page 11 lines 27-28.
21. *Page 4 Line 14 - replace "Uncertainty in forecasting model..." with "Uncertainty in the forecasting model..."*
    Reply 21: The sentence "Uncertainty in forecasting model..." replaced with "Uncertainty in the forecasting model..." in the revised manuscripts the change can be seen at page 4, line 32.

22. *Page 4 line 16 - Remove this line as it is repetitive (see line 4-5 on this page which says the same thing)*
    Reply 22: We have removed the line, "In this study, the forecast data is interpolated to match the grid and resolution of observations i.e. 0.5^ox0.5^o .for verification.", from the text on page 5 lines 1-2.
23. *Page 5 line 8 - Heidke skill score mentioned but not defined or used. Remove this reference?*
    Reply 23: Heidke skill (HK) score is used at page 5, from line 16-20 and at page 9,line 15.
24. *Page 5 line 26 - replace "...efficiency" with "...capability"?*
    Reply 24: The word "...efficiency" replaced with "...capability" at page 6, line 16.
25. *Page 6 line 9 - replace "... the figures (Fig. 5) and (Fig. 4)." with "Fig. 5 and Fig. 4."*
    Reply 25: The figures (Fig. 5) and (Fig. 4).", replaced with "Fig. 5 and Fig. 4."
26. *Page 6 line 11 - use "The deterministic forecast..."*
    Reply 26: The sentence started with the *"The deterministic forecast..."* in the text.

27. *Page 6 lines 11-12 Replace "...any variation in either of the considered days and illustrates quasi-stationary characteristics of the deterministic forecast from Day-1 through Day-10 forecast" with "... any variation in either of the days and illustrates quasi-stationary characteristics from Day-1 through Day-10"*
    Reply 27: The text is modified as per the suggestion on.
28. *Page 6 line 13 - Remove "...and vary in not so distinctive fashion".*
    Reply 28: A part of sentence *"...and vary in not so distinctive fashion"* is removed from the text.
29. *Page 6 line 15 - "Fig. (2.)" should read "Fig.2"*
    Reply 29: "Fig. (2.)" replace with "Fig.2"
30. *Page 6 line 15 - Remove "..(Tmax).."*
    Reply 30: Tmax removed from the text.
31. *Page 8 line 7 - Fig 10. should read Fig 9.*
    Reply 31: The figure " Fig 10" is replace with "Fig 9" in the text.
32. *Page 8 line 18 - Fig. 9 should read Fig. 10.*
    Reply 32: The figure " Fig 9" is replace with "Fig 10" in the text.
33. *Page 8 line 21 - missing end parentheses ")"*
    Reply 33: The missing end parentheses ")" is inserted in the text.
34. *Page 8 line 23 - missing figure number.*
    Reply 34: The missing figure number is inserted in the text.

Figures and tables

1. *Figure 2 and 3 - the colour bar is labelled °C When the quantity is a count.*
   Reply 1: The colour bar label °C is removed from the figure 2.
2. *Table 2 title - "Causalities" should read "Casualties".*
   Reply 2: The word "Causalties" replaced by the "Casualties" in the table title.

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

Harvir Singh, Kopal Arora, Raghvendra 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 ~~the~~ extreme events (~~Heat waves~~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 ~~(Heat waves)~~ as compared to deterministic forecast. Locations of the events are also better captured by the ~~ensemble model~~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 ~~the~~ of severe weather/extreme events. ~~Extreme weather events comprehend non-linear interactions usually between the small scale natural processes~~ 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 ~~2014~~). 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, Toth and Kalnay, 1997, Molteni et al., 1996) 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 ~~2014~~). Ensemble forecasting has emerged as the practical way of estimating the forecast uncertainty and making probabilistic forecasts. ~~Based~~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~~ (Sarkar et al., 2016).~~. This global medium-range weather forecasting system has been adopted from the UK Met ~~o~~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 ~~heat wave~~heatwave events. The ~~heat wave~~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 ~~heat wave~~heatwave conditions if departure from normal is 4.5°C to 6.4°C and severe ~~heat wave~~heatwave if the departure from normal is >6.4°C. In terms of the actual maximum temperature, a station is under ~~heat wave~~heatwave when actual maximum temperature Ó45°C and severe ~~heat wave~~heatwave when the maximum temperature is >47°C. There has been increasing interest in predicting such extremes, the ~~heat wave~~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 by a few recent studies (Qin et al., 2013). A study conducted over the Indian sub-continent between 1969 and 1999 indicated more frequent cold and ~~heat wave~~heatwave events over the Indo-Gangetic plains of India. 5-6 ~~heat wave~~heatwave events and 2-3 cold wave events are reported to occur every year in the Northern parts of the country. The global temperatures have exhibited a warming trend of about 0.85~~ 0~~°C due to anthropogenic activities between 1880 and 2012~~ (Molteni et al., 1996).~~. Similar trends were also observed in India with the annual air surface temperature rise ~~between 1901 and 2014~~during 20th century. This is evident from the detailed study presented in Kothawale et al (2010) based on the data from 1901-2007.

~~The annual mean temperature has been shown to increase by 0.42°C per 100 years while the maximum land temperature over the India has shown an overall increase of 0.92°C per 100 years (Arora et al., 2009). Minimum temperature, on the other hand, does also show a warming trend over the Indian sub-continent but with a smaller magnitude of about 0.09°C (per 100 years) (Arora et al., 2009).Another study reported a significant rising trend of 0.05°C in the mean surface temperate of India between 1901 and 2003 has documented a warming by about 0.22°C per decade (Arora et al., 2009).~~ The 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 highlights that the mean temperature during the pre-monsoon season (March-May) shows an increasing trend of 0.42°C per 100 years. ~~This period also includes the monsoon months which additionally represent unprecedented warming, unusual according to the authorôs experience (Kothawale, 2005).~~ 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 ~~heat wave~~heatwave events on weather and climatic scales. While there is an extensive literature discussing the ~~heat wave~~heatwave events and their trends on the climatic scales, however, the literature is rather limited (especially over India) focusing such events on monthly scales. This paper thus tries to fill in the gap and attempt to demonstrate the capability and strength of predicting such events using both ensemble ~~forecast forecasts both,~~ and deterministic ~~and ensemble forecasting.~~forecast. This research investigates the most recent ~~heat wave~~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 ~~and~~ 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 ~~heat waves~~heatwaves. This forms the first documented study of the recent ~~heat wave~~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 can't do. This also provides some important insights into the use of ensemble forecast over the deterministic forecast in predicting extreme events like a ~~heat wave~~heatwave ~~and a cold wave~~. 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 ~~heat wave~~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, which was 1°x1° resolution a few years earlier over Indian land area.. 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 Shepardǒs angular distance weighting algorithm (Shepard,1968).In this study, we have used IMD's real-time daily gridded (Srivastava et al., 2009, Caesar et al., 2006, Kiktev et al., 2003, New et al., 2000, Piper and Stewart, 1996, Rajeevan et al., 2005 and Shepard, 1968) 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 from NCUM and NEPS respectively. 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 at 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 (Sarkar et al., 2016) adapted from the UK METMet 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). In this study, the forecast data is interpolated to match the grid and resolution of observations i.e. 0.5°x0.5° .for verification.

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 (4th 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.

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

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

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.

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

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 ~~efficiency~~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 ~~threshold~~ temperature thresholds~~s~~. 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.

~~*Tmax* & *Tmin* forecasts from deterministic and ensemble-based approach are averaged over 3 months (MAM) and are illustrated in the Figs.-1(a) and 1(b).~~

~~The y-axis in each of the Figs represents the number of forecasts points possessing the temperature value above a threshold. The values on the Y-axis are divided by 2686 to present it in an easy to read format, which is then multiplied with 100 to express the values as percentages.~~The panels in Figure 1a,b show the observed and forecast (Day-3) frequency distribution for *Tmax* and *Tmin*. ~~As~~ For lower temperature thresholds, the forecast underestimate the frequency, while it can be seen in the Figs-1(a & b), both, deterministic and ensemble~~-based approaches~~ mean ~~predict better temperature during the warm days~~converg towards observed relative frequency, especially for the temperature exceeding 38°C. NEPS performs better than the NCUM forecast (~~Fig~~ Figure1a), indicating better performance of the ensemble forecast over the deterministic one. ~~This~~

feature has also been highlighted in the figures (Fig. 5Figure 6) and (Fig.Figure 4).. Both, NCUM and NEPS forecasts converge to the observations at higher temperatures for both, *Tmax* and *Tmin*.

DeterministicThe deterministic forecast hardly shows any variation in either of the considered days and illustrates quasi-stationary characteristics of the deterministic forecast from Day-1 through Day-10 forecast. In the case of *Tmin* forecast, both the models underperform in terms of the temperature prediction and vary in not so distinctive fashion. Both the forecasts tend to converge to the observations at higher temperatures.

From the spatial map Fig.Figure (2.),, the frequency of the observed maximum temperature (*Tmax*) *Tmax* Ó40°C in over the Maharashtra and adjoining regions shown maximum (more than 70 counts) over the entire period of MAM 2016, which is picked up by both deterministic and ensemble modelensemble forecasts. However, deterministic modeldeterministic forecast is showing more frequency spread over MP, UP and Bihar, OrissaOdisha, AP and adjoining states from day-1 to day-9. As forecast lead time increases from day-1 to day-9 the heat waveheatwave frequency increases from central India to north and east India. Consequently, higher number of heat waveheatwave extremes was predicted by deterministic modeldeterministic forecast NCUM over east UP, Bihar, West-Bengal, OrissaOdisha, Jharkhand, Chhattisgarh and AP. On the other hand, NEPS (Fig.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 heat waveshеatwaves over the above said regions. Comparatively, the ensemble based model NEPS is performing better (spatially) for the extremes of heat-wave events than the deterministic modeldeterministic forecast NCUM over most of the Indian states up to day-9.

4. Case Studies for Extreme Heat wavesHeatwaves

4.1 Weather conditions during MAM-2016

Heat waveHeatwave 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 tThe severity and extent of heating increased during the next week resulting in the establishment of severe heat waveheatwave 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, heat waveheatwave 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 ~~heat wave~~heatwave conditions for 4 to 5 days in succession from 17 to 21 May and temperature observed Ó50°C. ~~Heat wave~~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 ~~heat wave~~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 ~~heat wave~~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 ~~Heat Waves~~Heatwaves on 11th April 2016

As per the IMD reports (Climate Diagnostics Bulletin of India, April 2016), the ~~heat wave~~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-~~1~~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 the~~The~~ NEPS is better in predicting the extremes of ~~Heat Waves~~heatwaves up to day -9 then the NCUM.

4.3.2 Case-II ~~Heat Waves~~Heatwaves on 21st May 2016

The severe ~~heat wave~~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 ~~S~~some stations of West Rajasthan temperature observed Ó 50°C viz. Barmer, Bikaner, Ganganagar, Jaisalmer & Jodhpur and observed severe ~~heat wave~~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 ~~heat wave~~heatwave events. The ~~heat wave~~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, 2014~~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 model~~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 ~~heat waves~~heatwaves several days ahead of the event, as discussed in the results. The severe ~~heat waves~~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 ~~heat waves~~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 workwith.

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.

References:

~~Arora, M., Goel, N. K. and Singh, P.: Evaluation of temperature trends over India / Evaluation de tendances de température en Inde, Hydrological Sciences Journal, **50**:1, -93, 2005.~~

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.

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., and K. Rupa Kumar: On the recent changes in surface temperature trends over India, Geophysical Research Letters, Vol. 32, L18714, doi:10.1029/2005GL023528, 2005~~

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.

Molteni, F., R. Buizza, Palmer, T. N. and Petroliagis, T.:The ECMWF Ensemble Prediction System: Methodology and validation. Quart. J. Roy. Meteor. Soc, **122,** 73ï119, 1996.

Mureau, R., Molteni, F. and Palmer, T. N.: Ensemble prediction using dynamically conditioned perturbations. Quart. J. Roy. Meteor. Soc, **119,** 299ï323, 1993.

New, M., Hulme, M. and Jones, P.: Representing twentieth century spaceïtime climate variability. Part II: Development of 1901ï96 monthly grids of terrestrial surface climate. J. Climate, **13**, 2217ï2238, 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**, 752ï782, 1996.

Qin, D. *et al. Climate change 2013: The physical science basis (p. 1535)*. T. Stocker (Ed.). Cambridge, UK, and New York: Cambridge University Press (2014).

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

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.

Shepard, D.: A two-dimensional interpolation function for irregularly spaced data. Proc. 23rd National Conf., New York, NY, Association for Computing Machinery, 517ï524, 1968.

Srivastava, A. K., Rajeevan, M. and Kshirsagar, S. R.: Development of a high resolution daily gridded temperature data set (1969ï2005) 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.

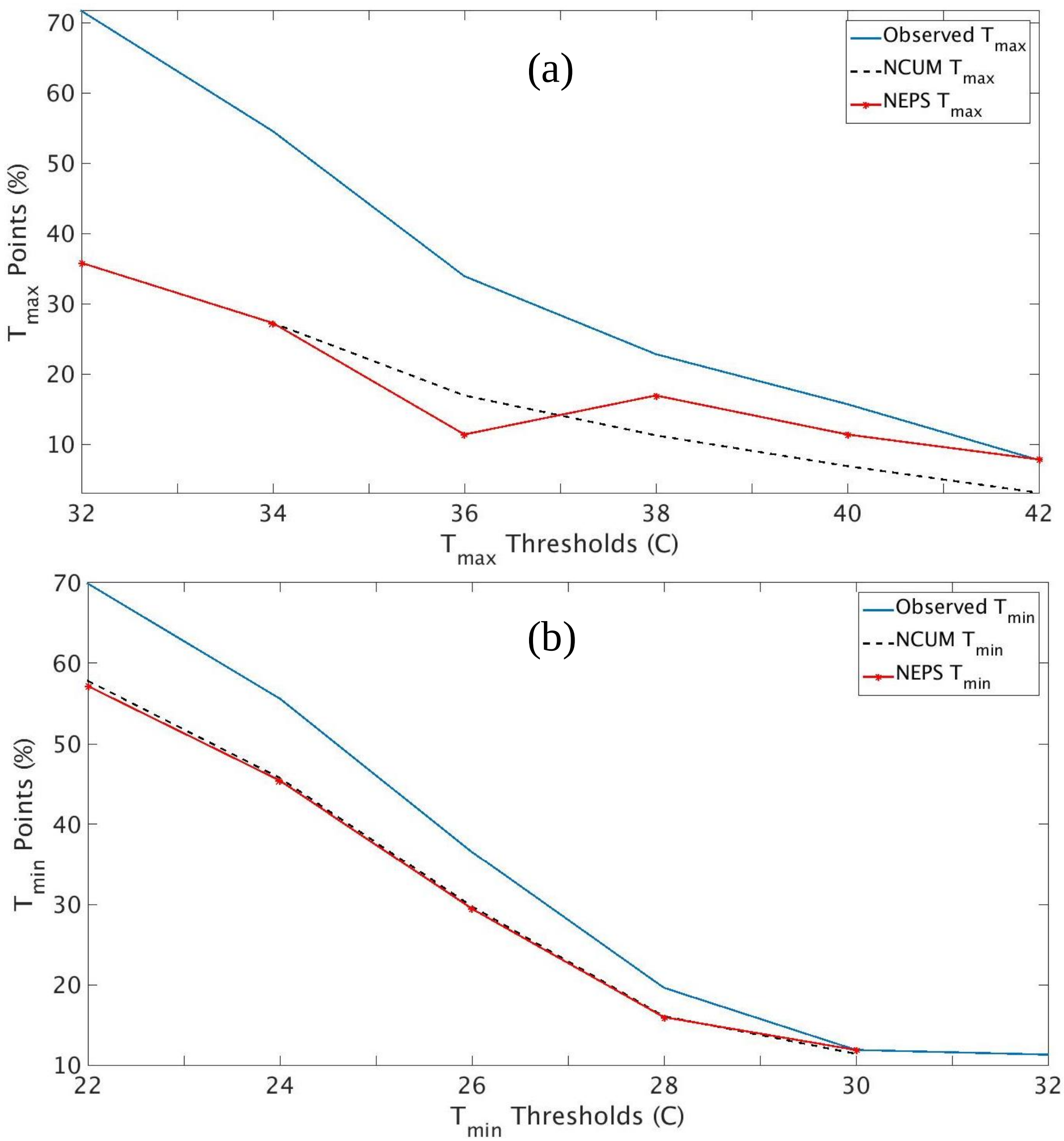


Figure 1. Frequency distribution of observed, and forecast (NCUM and NEPS) (a) *Tmax* (°C) and (b) *Tmin* (°C) over India during March-May 2016.

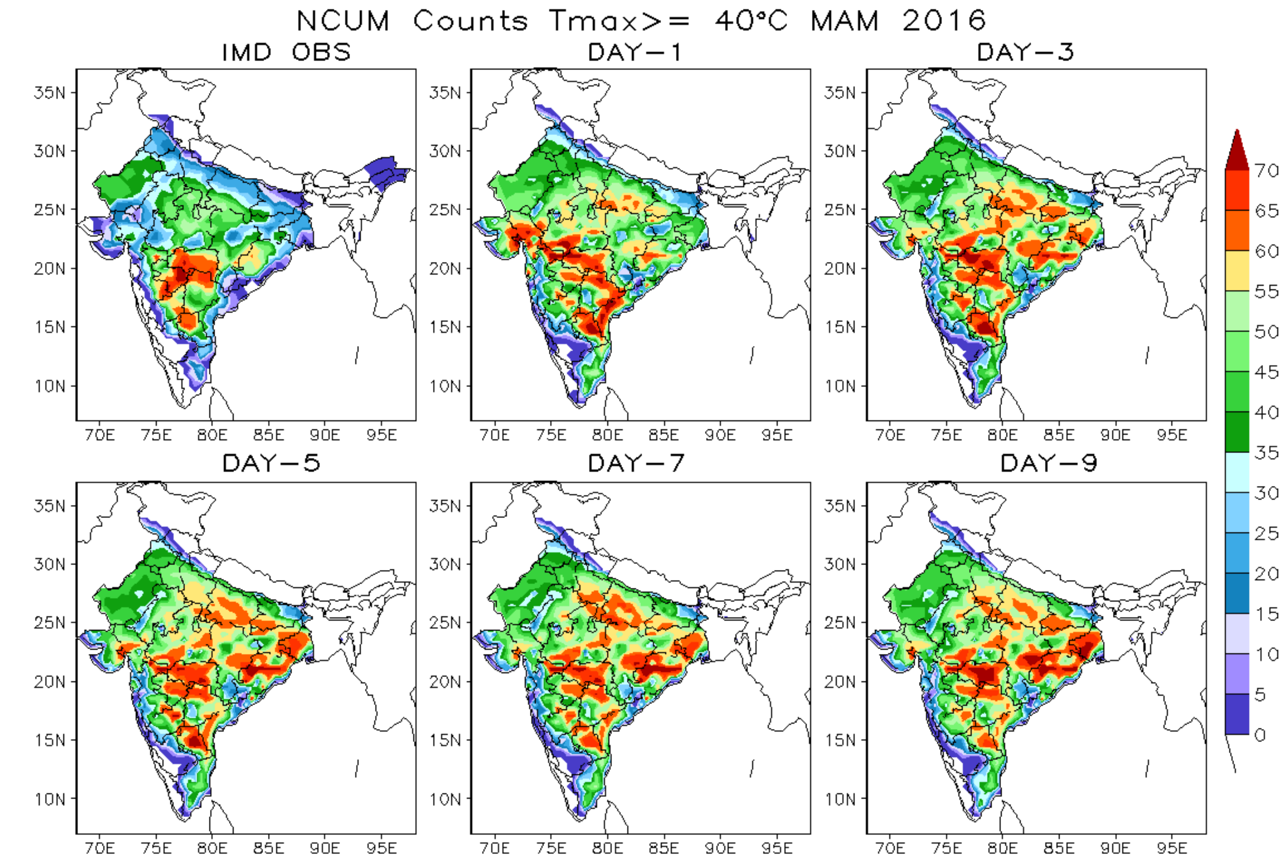


Figure 2. Spatial distribution of observed and NCUM forecasts number of days with *Tmax* Ó 40°C during the period of March to May 2016

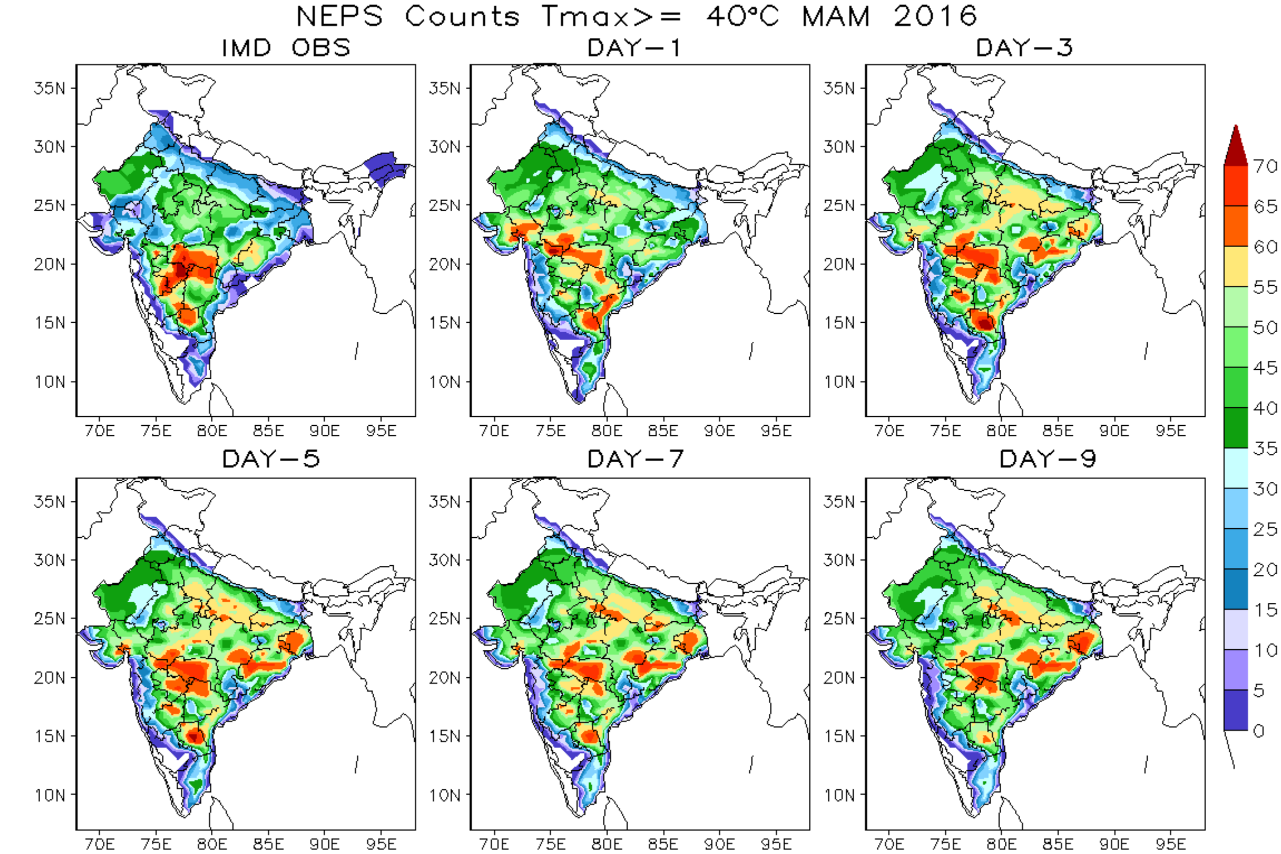


Figure 3. Spatial distribution of observed and NEPS forecasts number of days with *Tmax* Ó40°C during the period of March to May 2016

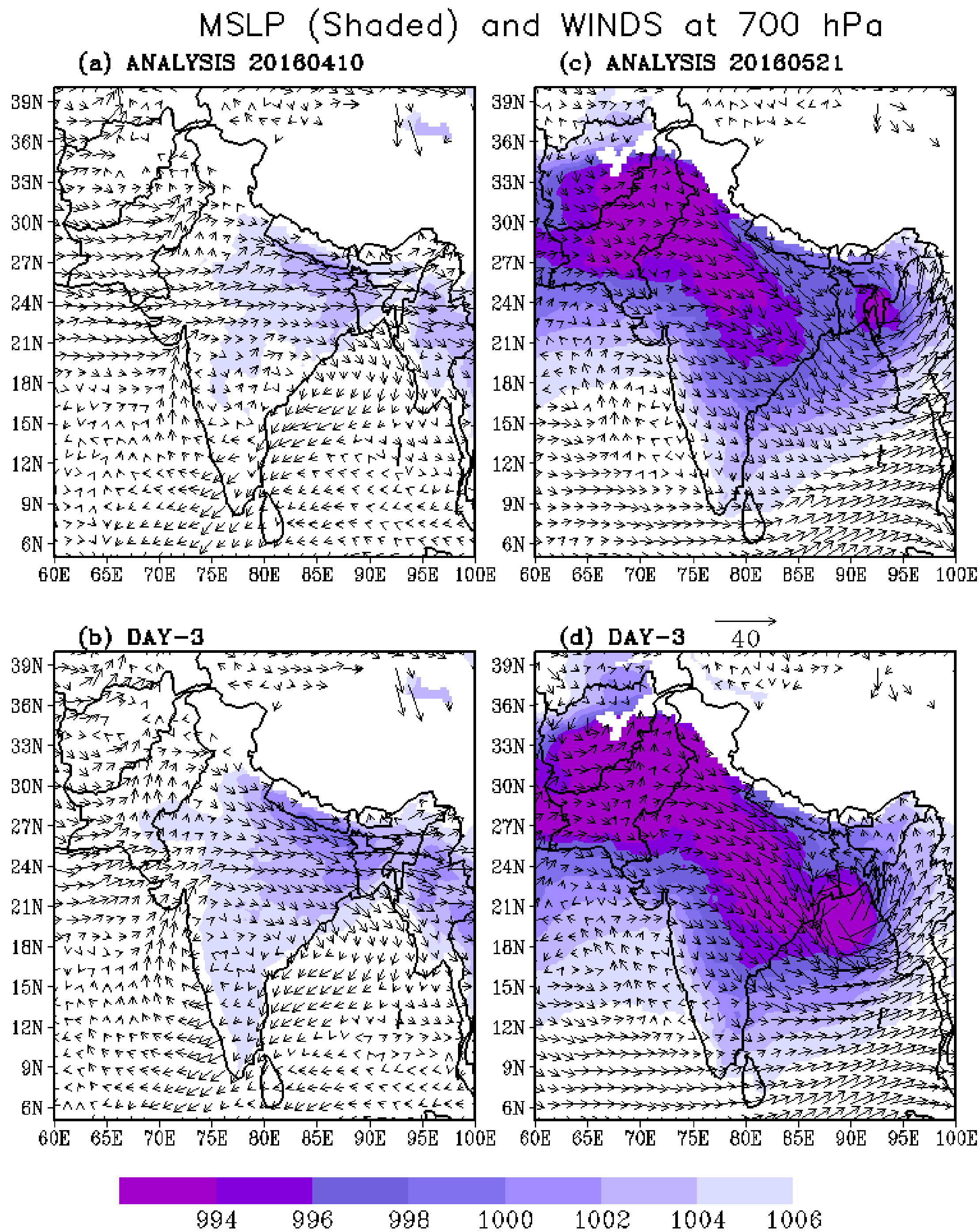


Figure 4. Mean Sea Level Pressure (MSLP) shaded and winds at 700 hPa showing heat low (a) Analysis of 20160410 (b) Day 3 forecast valid for 20160410 (c) Analysis of 20160521 (d) Day 3 forecast valid for 20160521

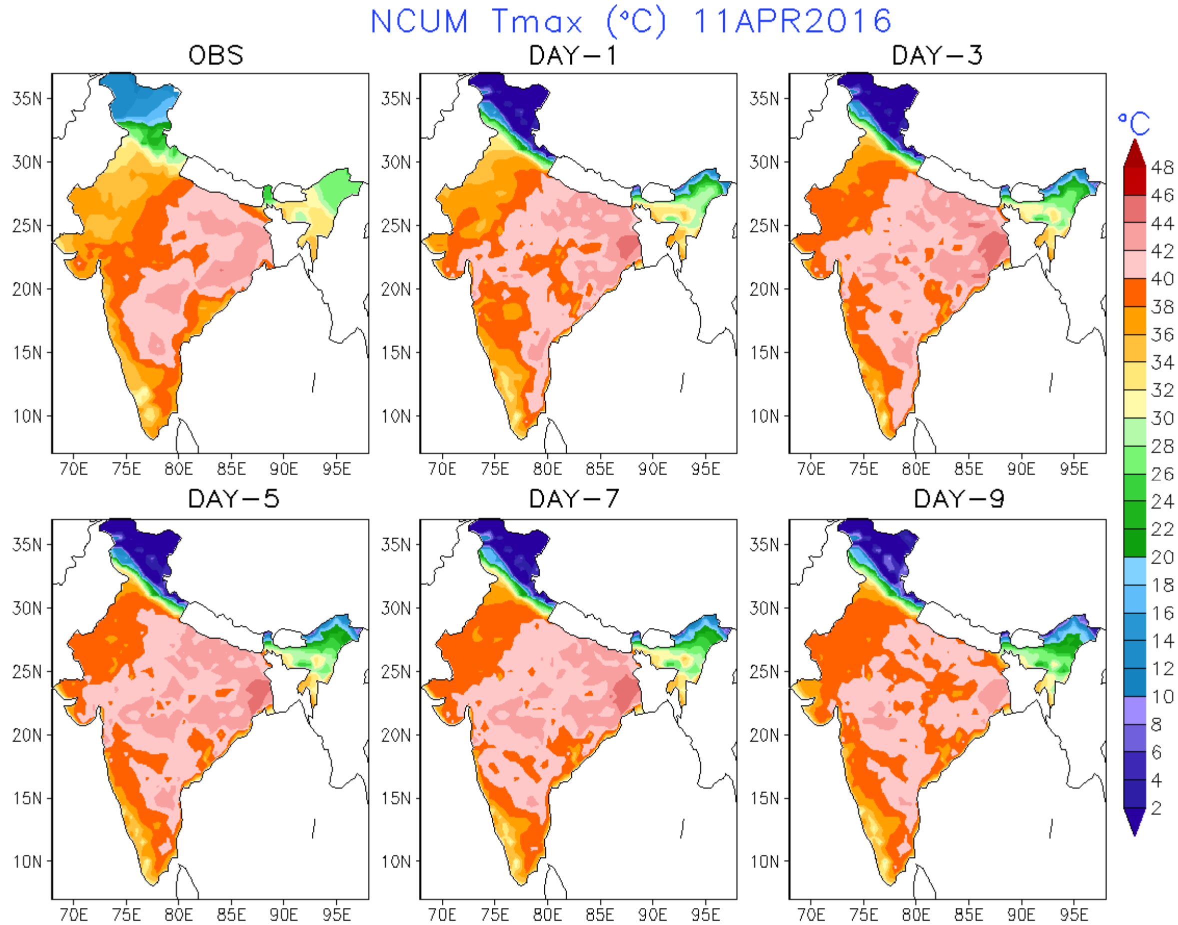


Figure 5. Spatial distributions of Observed *Tmax* and NCUM forecast *Tmax* prevailing heat-waves over, MP, Odisha, AP, Telangana and some parts of Maharashtra on 11th April 2016

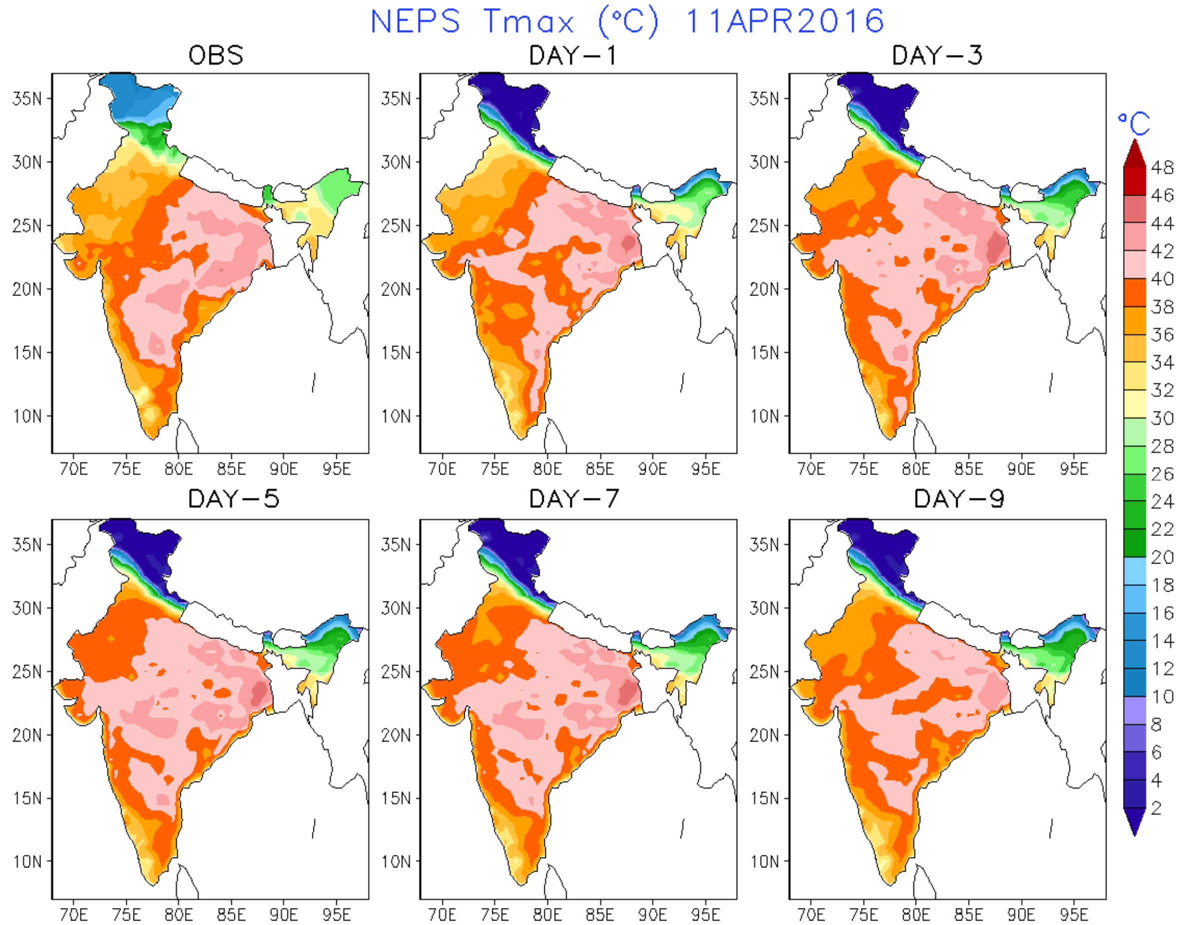


Figure 6. Spatial distributions of Observed *Tmax* and NEPS forecast *Tmax* prevailing heat-waves over, MP, Odisha, AP, Telangana and some parts of Maharashtra on 11th April 2016

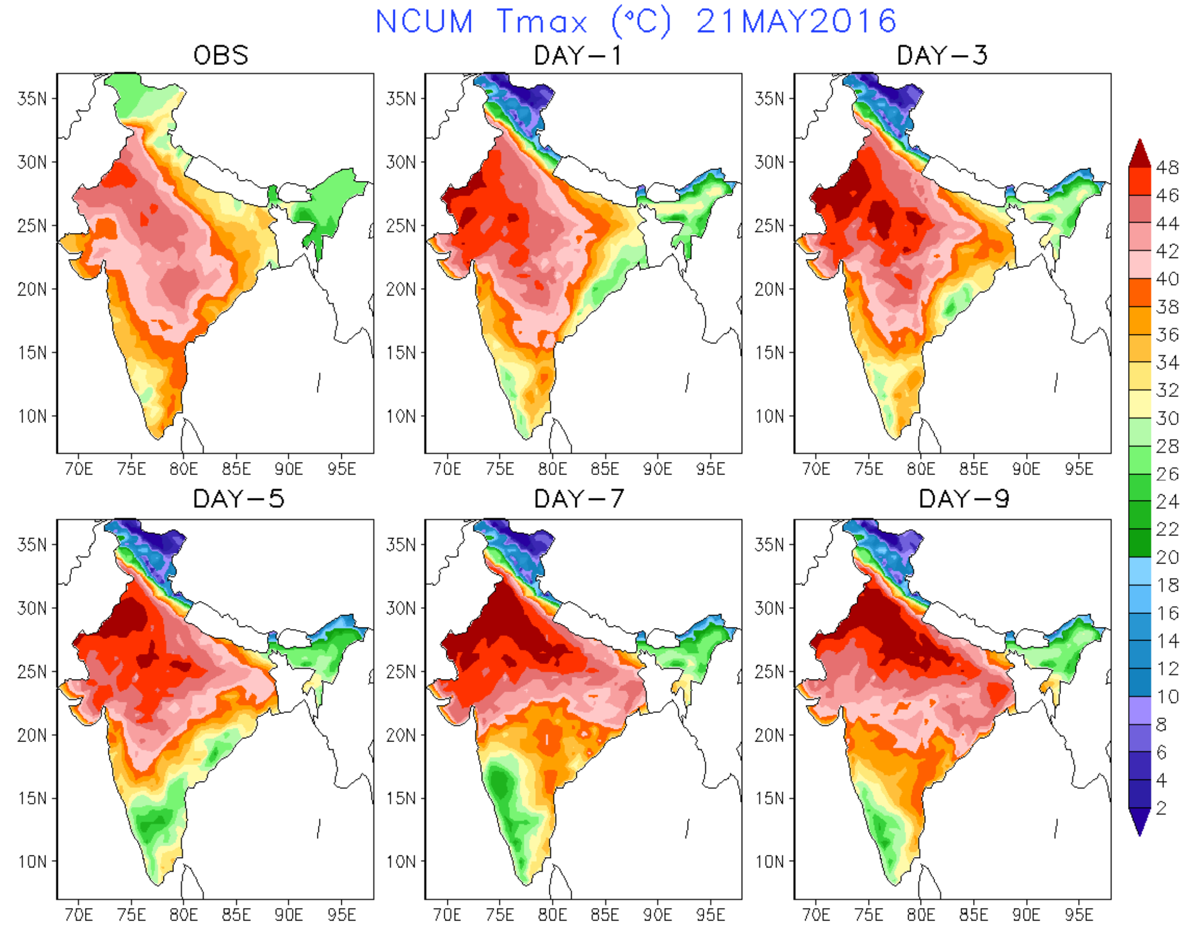


Figure 7. Spatial distributions of Observed *Tmax* and NCUM forecast *Tmax* prevailing heat-waves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21st May 2016

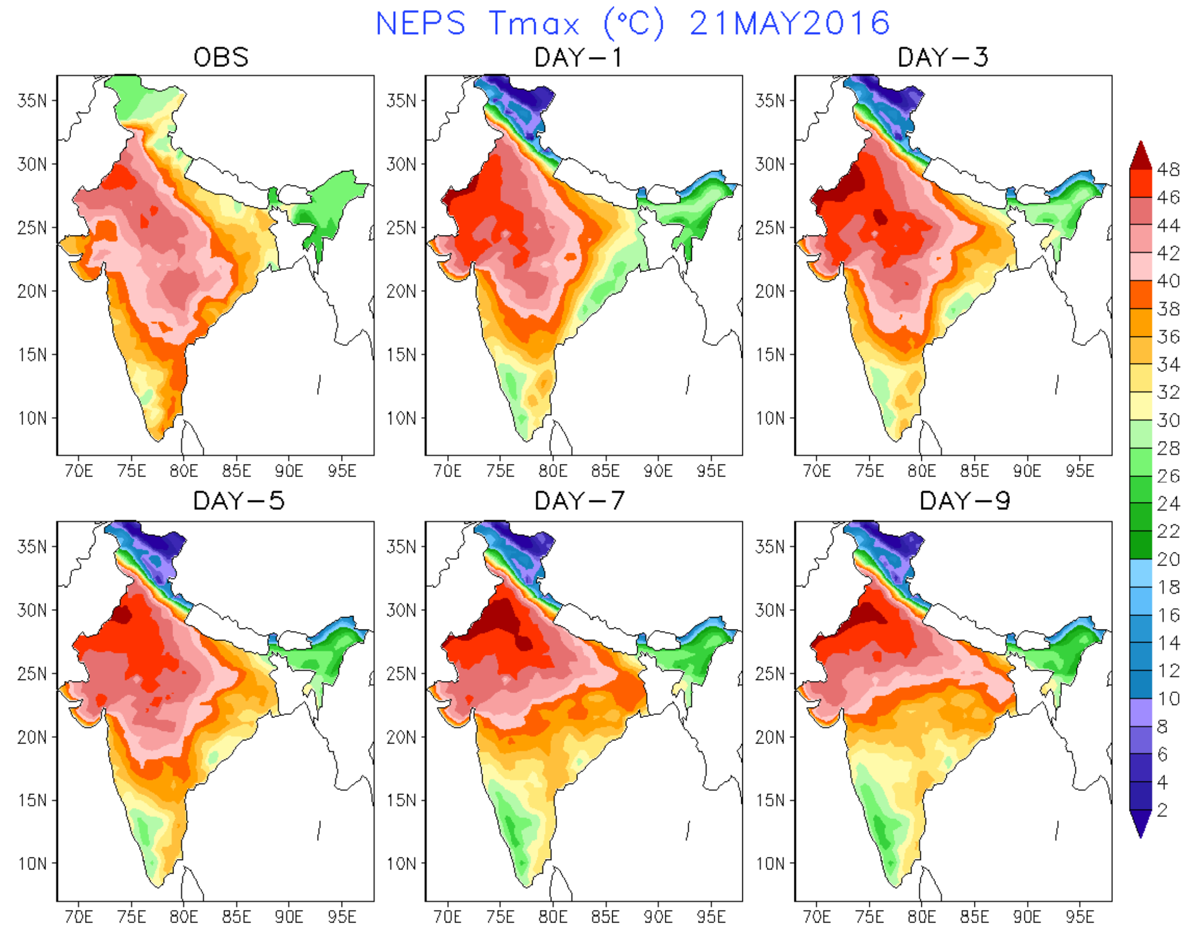


Figure 8. Spatial distributions of Observed *Tmax* and NEPS forecast *Tmax* prevailing heat-waves over Rajasthan, MP, UP, Delhi, Haryana, Punjab and some parts of Maharashtra on 21st May 2016

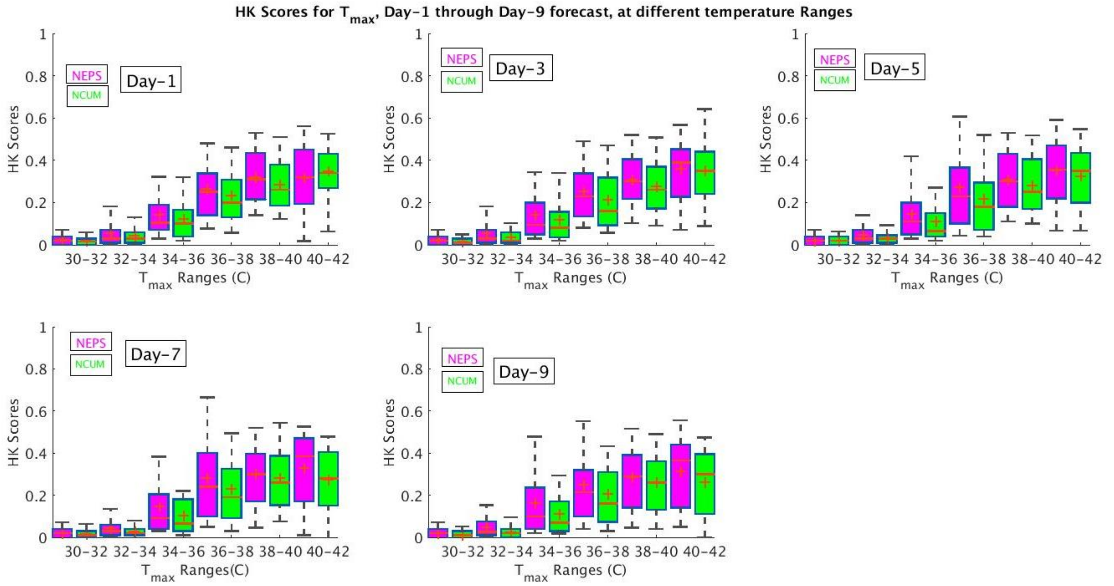


Figure 9. Box plots for HK scores for different temperature ranges (*Tmax*) NCUM and NEPS form March to May 2016

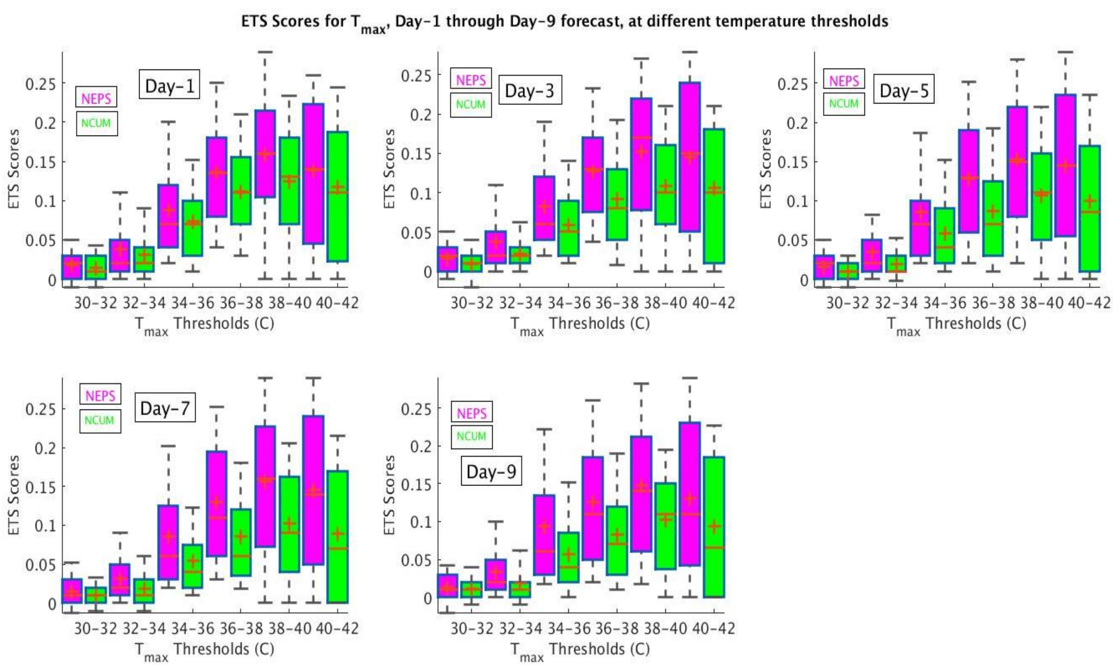


Figure 10. Box plots for Equitable Threat Score (ETS) for NCUM and NEPS form March to May 2016

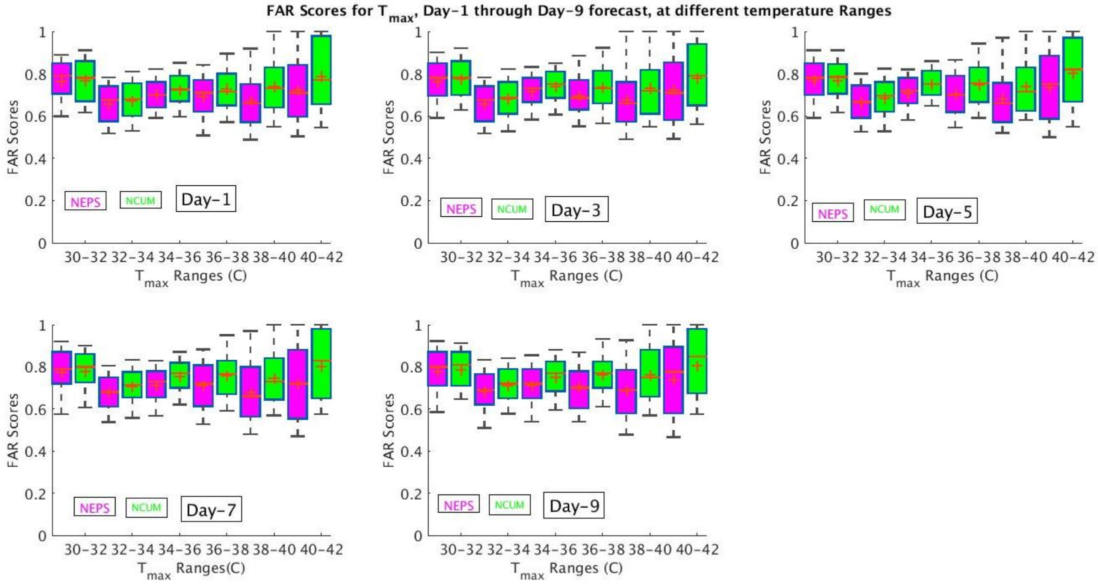


Figure 11. Box plots for False Alarm Ratio (FAR) for NCUM and NEPS form March to May 2016

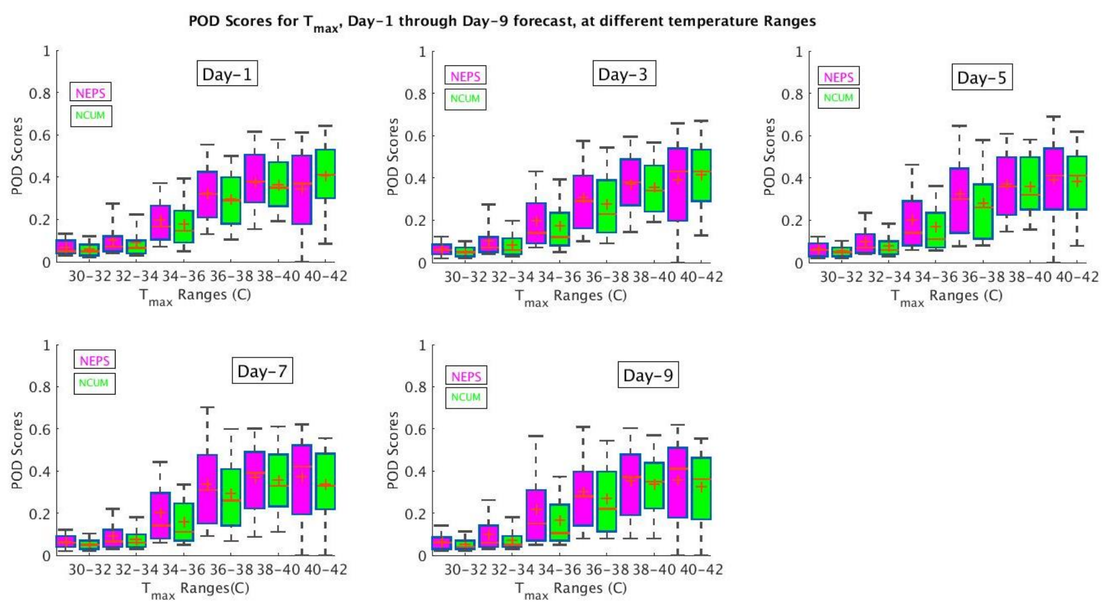


Figure 12. Box plots for Probability of Detection (POD) for NCUM and NEPS form March to May 2016

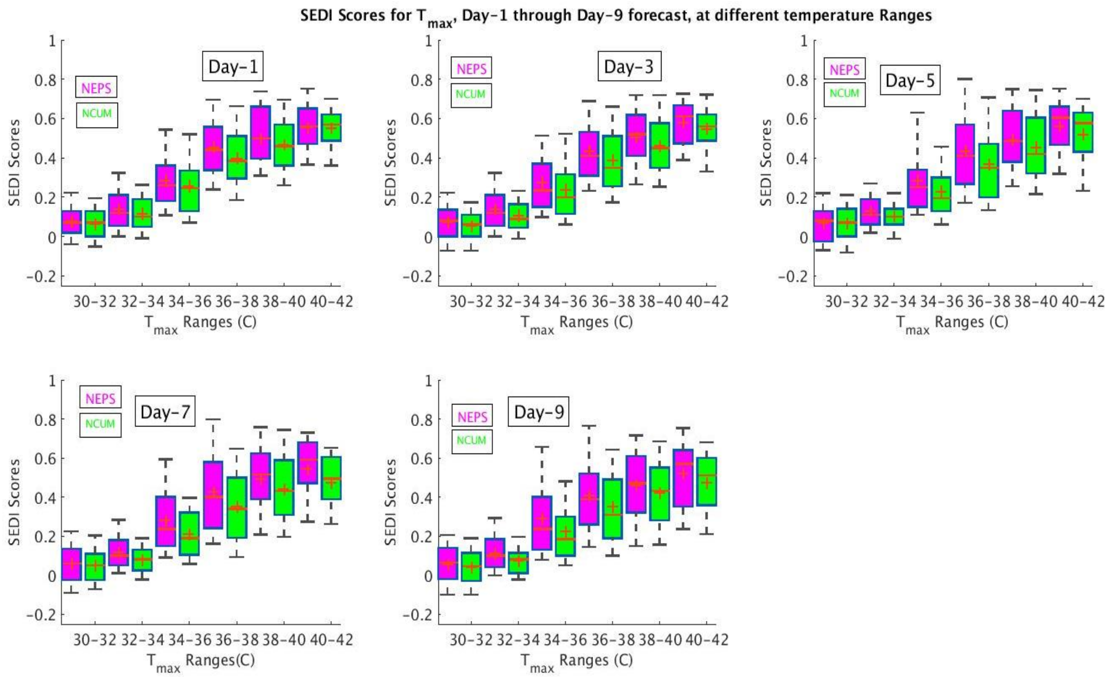


Figure 13. Box plots for Symmetric Extrernal Dependence Index (SEDI)) for NCUM and NEPS form March to May 2016

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
SV	**S**ingular **V**ector

Table 2. Ca~~u~~sualities reported during MAM-2016 due to prevailing ~~heat waves~~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	

Table 3. Monthly *Tmax* > 40°C scores for NCUM and NEPS forecast with 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