Non-stationary Temperature Duration Frequency Curves for the North-West Homogeneous Region of India

Meera G. Mohan[✉] [iD] and S. Adarsh [iD]

TKM College of Engineering, APJ Abdul Kalam Technological University, Kollam, Kerala 691 005, India
meeragmohan90@gmail.com

Abstract. The long-term rise of global temperatures due to climate change has resulted in a larger increase in the probability of occurrence of extreme temperature events. The warming trend with an increase in the intensity, frequency and duration of heat waves is observed mainly in the North western regions of India. The current development of Temperature Intensity Duration Frequency (TDF) curves relies on the assumption of stationarity which does not hold valid due to the recent evidences in global warming. In this study, India Meteorological Department (IMD) 1° × 1° gridded temperature dataset is used to examine the frequency of occurrence of extreme temperatures over the North Western homogeneous temperature region of India covering 50 grid points during the period of 1951–2019. The maximum daily temperatures for six different durations of 1-day, 2-day, 4-day, 6-day, 8-day and 10-day were extracted. This study proposes a non-stationary approach to the development of TDF curves keeping time as a covariate. Five Non-Stationary TDF models were developed varying location and scale parameters linearly, exponentially and their combinations with respect to time. The goodness-of-fit is improved when using a non-stationary approach with time covariates. Model, NSGEV-1 for which location parameter was linearly varied with respect to time turned out to be the best fitting model for more than 84% of the grid points for different durations. A comparison of Stationary TDF curve and Non-Stationary TDF curves were also carried to illustrate the impacts of considering non-stationarity in extreme temperature analysis.

Keywords: Climate change · Non-stationarity · Trend · Extreme temperature · Covariate · TDF curves

1 Introduction

Urbanization and industrialization, the effect of natural climatic variability, and increased greenhouse gasses in the atmosphere have been suggested as the leading causes of the increase in duration, magnitude, and frequency of temperature extremes and heatwaves in the context of the current climate evolution [1]. Study of temperature extremes is important for understanding climate variability which can vary spatially and temporally at different local, regional and global scales. Rising global temperatures are causing increases

in the frequency and severity of extreme climatic events, such as floods, droughts, and heat waves [2]. Mean temperatures across India have risen by more than 0.5 °C over the period of 1960–2009, with statistically significant increases in heat waves.

The development of Temperature-Duration-Frequency (TDF) curves, which relate the intensity of heat events of different durations to their frequencies, can be useful tools for the analysis of heat extremes [3]. The TDF curves are developed based on historical maximum temperature time series data by fitting a theoretical probability distribution of annual maximum extreme temperature series. TDF curves are developed similar to Rainfall Intensity-Duration-Frequency (IDF) curves. Conventionally, frequency analysis is based on the concept of stationary extreme value theory i.e., the probability distribution parameters are invariant with time. This assumption of stationarity does not hold valid due to the recent evidences in global warming. Various studies analysing the patterns of change in temperature extremes in terms of heatwaves and cold waves [4–8], summer and winter temperature extremes [9] and trend analysis [10, 11] were carried out in the past years. Khaliq et al. [12] developed the concept of TDF curves assuming that the characteristics of the probability distribution of extreme heat events are invariant through time. This study was conducted for four stations in southern Quebec region, Canada. This work was further extended to the non-stationary concept of TDF curves by Ouarda [3] for six stations in the Province of Quebec, Canada. As far as India is concerned, the climate is interesting and complex due to diverse topographical features and large geographical area, hence in case of a local variability in climate change, it is important to know which area could be more affected by these changes. These results can aid in verification of climate model studies, especially for the future. High resolution daily gridded temperature data (1969–2005) was developed by Srivastava et al. [13] which has supported many researchers in their temperature analysis. Even though stationary and non-stationary frequency analysis in terms of rainfall intensity, flood, drought severity has been explored by various researchers in the past, the notion of establishing TDF curves for India still remains to be unexplored. The semi-arid north-western part of India will likely warm more rapidly than the remaining parts making it more prone to extreme temperatures.

This study aims to conduct the frequency analysis of temperature extremes with time-varying distribution functions to characterize temperature extremes in a Non-Stationary (NS) context for the North-West region of India. The TDF analysis is carried out to explore the following objectives: (i) to develop Stationary (S) TDF model assuming location, scale and shape parameter to be a constant and five NS-TDF models by varying location and scale parameter linearly and exponentially with respect to time and the shape parameter as a constant; (ii) to identify the best fit model among the S and NS models for each duration extreme temperature series; (iii) to build and compare S-TDF curve and NS-TDF curves based on the best fit NS model for quantifying the percentage variation; and (iv) to compute the temperature return levels for various return periods with reference to the NS-TDF curves.

2   Study Area and Data

India is a vast country stretching between latitudes 8°4′N and 37°6′N and longitudes 68°7′E and 97°25′E. It is fractionated into seven homogeneous temperature regions

namely, North-West (NW), Western Himalayas (WH), West Coast (WC), East Coast (EC), North-East (NE), North Central (NC) and Interior Peninsula (IP) by the Indian Institute of Tropical Meteorology, Pune (online at www.tropmet.res.in). In this paper, the NW homogenous region which comprises of the Indian desert is considered as the study area due to its prominent temperature. The daily maximum temperature gridded data (1°latitude × 1°longitude) observed by the India Meteorological Department (IMD) is procured for NW region for a period of 1951–2019 (69 years). A total of 50 grid points spread out over Gujarat, Rajasthan, Haryana, Punjab, Uttar Pradesh and Himachal Pradesh belongs to the NW region (see Fig. 1).
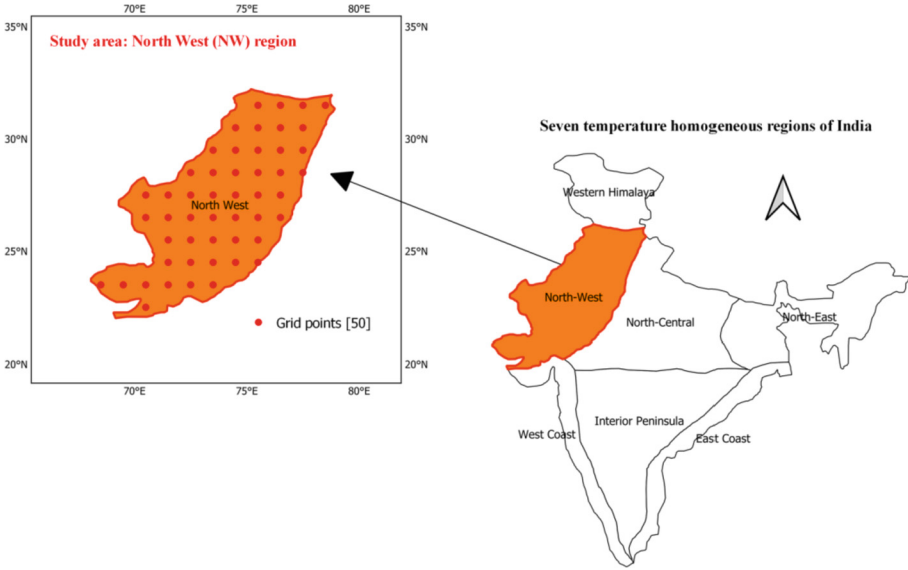

Fig. 1. Study area – North West temperature homogeneous region of India

3   Methodology

3.1   Significance of Trend

Firstly, Mann Kendell (MK) Trend analysis test [14, 15] was carried out to ascertain the presence of statistically significant trend. For a time series, X(t), the Mann–Kendall statistic is defined as:

$$S = \sum_{t=1}^{N-1} \sum_{i=t+1}^{N} sign\big(X(i) - X(t)\big) \tag{1}$$

where N is number of data and sign (•) represents a signum function given by:

$$sign(a) = \begin{cases} \frac{a}{|a|}, & if\ x \neq 0 \\ 0, & if\ x = 0 \end{cases} = \begin{cases} 1, & if\ a > 0 \\ 0, & if\ a = 0 \\ -1, & if\ a < 0 \end{cases} \tag{2}$$

Sign and value of the S statistic show the direction and intensity of the trend. The increasing or decreasing trends are indicated by positive or negative values of S respectively, while zero represents no trend. A normalized test statistic z can be used to statistically quantify the significance of the trend.

$$z = \begin{cases} \frac{S-1}{\sqrt{var(S)}} & if\ S > 0 \\ 0 & if\ S = 0 \\ \frac{S+1}{\sqrt{var(S)}} & if\ S < 0 \end{cases} \tag{3}$$

The positive values of z represent increasing trend while negative values indicate decreasing trend. At 95% confidence interval, if the value of z is more than 1.96 it represents significant non-stationary increasing trend. Similarly, if z is less than -1.96, it represents significant non-stationary decreasing trend.

3.2   Stationary and Non-stationary Temperature Models

Extreme Value Theory (EVT) is a framework used for analysing the risks associated with climate extremes and their return levels [1]. The Generalized Extreme Value (GEV) distribution is a continuous probability distribution that combines Gumbel, Fréchet and Weibull extreme value distributions. The frequency model using this family of distributions is often referred to as the block maxima approach [1]. The GEV distribution depends on three parameters: location (μ), scale (σ), and shape (ξ) with a cumulative distribution function expressed as:

$$F(x, \mu, \sigma, \xi) = \begin{cases} exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \sigma > 0, 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0, \xi \neq 0 \\ exp\left\{-exp\left[-\frac{(x-\mu)}{\sigma}\right]\right\}, \sigma > 0, \xi = 0 \end{cases} \tag{4}$$

In this study Stationary (S) model and five Non-Stationary (NS) models were developed as shown in Table 1. The NS setting for the location and scale parameters of GEV distribution was introduced as a function of time. The shape parameter was assumed to be a constant for all models.

Table 1. Development of stationary and non-stationary temperature models

Model name	Model parameters		
	Location	Scale	Shape
SGEV	$\mu_t = \mu$	$\sigma_t = \sigma$	$\xi_t = \xi$
NSGEV-1	$\mu_t = \mu_0 + \mu_1 Time$	$\sigma_t = \sigma$	$\xi_t = \xi$
NSGEV-2	$\mu_t = \mu$	$\sigma_t = \sigma_0 + \sigma_1 Time$	$\xi_t = \xi$
NSGEV-3	$\mu_t = \mu_0 + \mu_1 Time$	$\sigma_t = \sigma_0 + \sigma_1 Time$	$\xi_t = \xi$
NSGEV-4	$\mu_t = \mu$	$\sigma_t = exp(\sigma_0 + \sigma_1 Time)$	$\xi_t = \xi$
NSGEV-5	$\mu_t = \mu_0 + \mu_1 Time$	$\sigma_t = exp(\sigma_0 + \sigma_1 Time)$	$\xi_t = \xi$

The distribution parameters are estimated using Maximum Likelihood Estimation (MLE) method. MLE involves defining a likelihood function for calculating the conditional probability of observing the data sample given a probability distribution and distribution parameters. The log likelihood is derived from Eqs. (5) and (6). The location and scale parameters in the above equations are replaced with the corresponding equations in Table 1 based on the model application. MLE is done by maximizing the likelihood estimates obtained from Eqs. (5) and (6).

For $\xi \neq 0$,

$$\log L(\mu, \sigma, \xi | X) = -n \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n} \log \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]$$

$$- \sum_{i=1}^{n} \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{-1/\xi}, 1 + \xi \left(\frac{x_i - \mu}{\sigma}\right) > 0 \qquad (5)$$

For $\xi = 0$,

$$\log L(\mu, \sigma | X) = -n \log \sigma - \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^{n} \exp \left[-\left(\frac{x_i - \mu}{\sigma}\right)\right] \qquad (6)$$

The Akaike Information Criterion (AIC), which penalizes the minimized negative log likelihood for the number of parameters estimated, is used to select the best model [16]. AIC is used to estimate the best fit GEV model among the SGEV model and 5 NSGEV models. It is computed using Eq. (7)

$$AIC = 2K - 2\ln(L) \qquad (7)$$

where K is the number of independent variables used and L is the log-likelihood estimate.

3.3 Quantification of Temperature Intensities for Various Return Periods

Once the best NS model is identified, the NS temperature intensity is estimated using the best NS model parameters. Unlike S model, the NS models' location and/or scale parameter value are bound to be time dependant. In this study for NS framework models, 95 percentiles of the location and scale parameter values from the historical observation obtained from Eqs. (8) and (9) is substituted into Eq. (10) for estimating the non-stationary temperature intensities for various return periods (T in years).

$$\widehat{\mu_{95}} = Q_{95}(\widehat{\mu_{t1}}, \widehat{\mu_{t2}}, \ldots \ldots, \widehat{\mu_{tn}}) \qquad (8)$$

$$\widehat{\sigma_{95}} = Q_{95}(\widehat{\sigma_{t1}}, \widehat{\sigma_{t2}}, \ldots \ldots, \widehat{\sigma_{tn}}) \qquad (9)$$

$$I = \begin{cases} \widehat{\mu} + \frac{\widehat{\sigma}}{\widehat{\xi}} \left[\left(-\log\left(1 - \frac{1}{T}\right)\right)^{-\widehat{\xi}} - 1\right], \widehat{\xi} \neq 0 \\ \widehat{\mu} + \widehat{\sigma} \left[-\log\left(-\log\left(1 - \frac{1}{T}\right)\right)\right], \widehat{\xi} = 0 \end{cases} \qquad (10)$$

4   Results and Discussion

The daily maximum temperature data obtained from IMD for 50 grid points belonging to NW region was used to extract the annual maximum temperature over 6 durations (1-day, 2-day, 4-day, 6-day, 8-day and 10-day). The trend analysis was carried over the NW region using MK test and it was found that number of grid points following non-stationarity trend increases as the duration increases. The same is represented in a bar graph in Fig. 2.
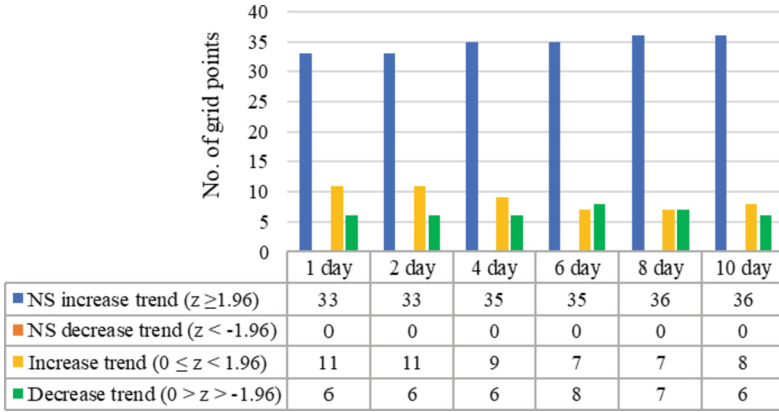

	1 day	2 day	4 day	6 day	8 day	10 day
■ NS increase trend (z ≥ 1.96)	33	33	35	35	36	36
■ NS decrease trend (z < -1.96)	0	0	0	0	0	0
■ Increase trend (0 ≤ z < 1.96)	11	11	9	7	7	8
■ Decrease trend (0 > z > -1.96)	6	6	6	8	7	6

Fig. 2. Significance of trend in grid points of NW region

The S TDF and NS TDF analysis was carried out by developing GEV models as mentioned in Sect. 3.2 and the best fit model was identified with the help of AIC values. From Fig. 3 it can be seen that the best fit models vary for the same grid point as the duration changes in some cases while in others the best fit model remains constant throughout for all durations.
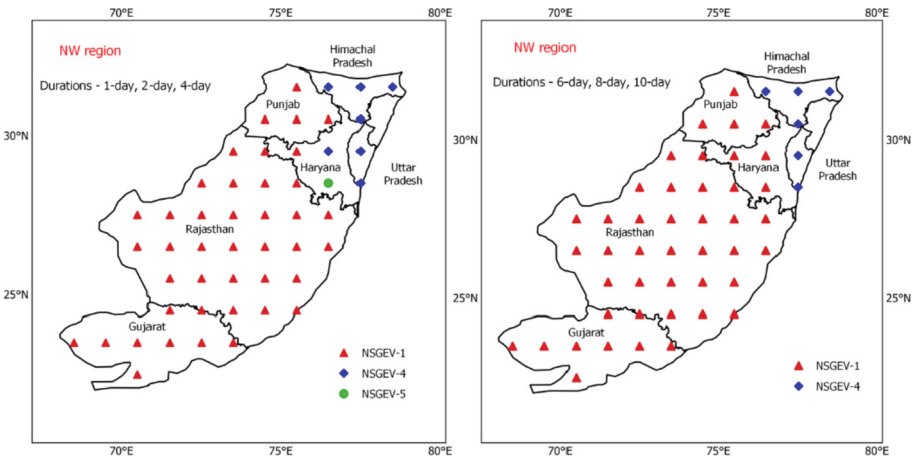

Fig. 3. Best fit models for NW temperature homogenous region for all durations

It can be concluded that more than 84% of the grid points have NSGEV-1 as the best fit model where location parameter is linearly time varying with a constant scale and shape parameter for all durations (see Fig. 4). S TDF curves and best fit NS TDF curves were developed for all the grid points for various return periods (RP) of 2 years, 5 years, 10 years, 25 years, 50 years and 100 years. Due to space constraints, a sample grid point (23.5°N, 68.5°E) of Gujarat is chosen for a detailed representation of S and best fit NS TDF curves for all return periods in Fig. 5. The dashed curves represent S-TDF curves and the thick line represent NS-TDF curves for various return periods. It was found that the temperature intensities decrease with increase in the daily maximum durations and increase when return period increases.
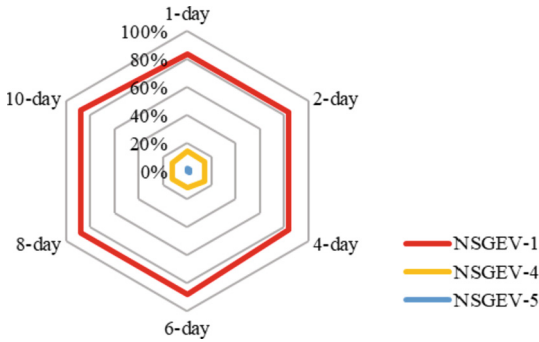

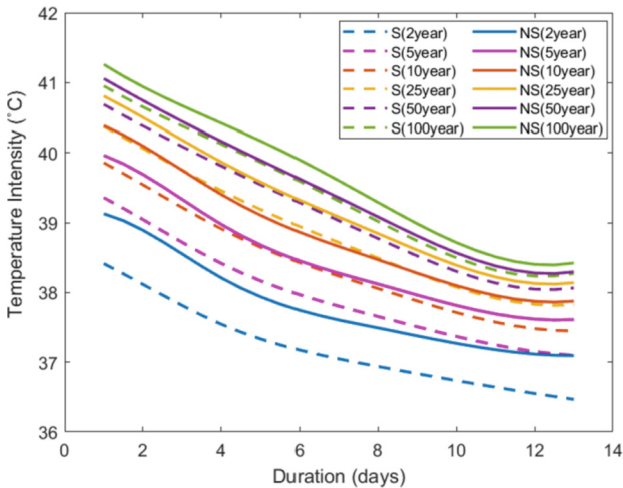**Fig. 4.** Percentage of grid points for best fit models in NW region

Fig. 5. S and best fit NS (NSGEV-1) TDF curves for grid point (23.5°N, 68.5°E) Gujarat

The percentage variations (PV) of return levels between S & NS cases were calculated for all the grid points. A spatial representation of the temperature intensities for both S

and NS case along with the PV for 1 day duration for all return periods for the entire NW region is shown in Fig. 6.
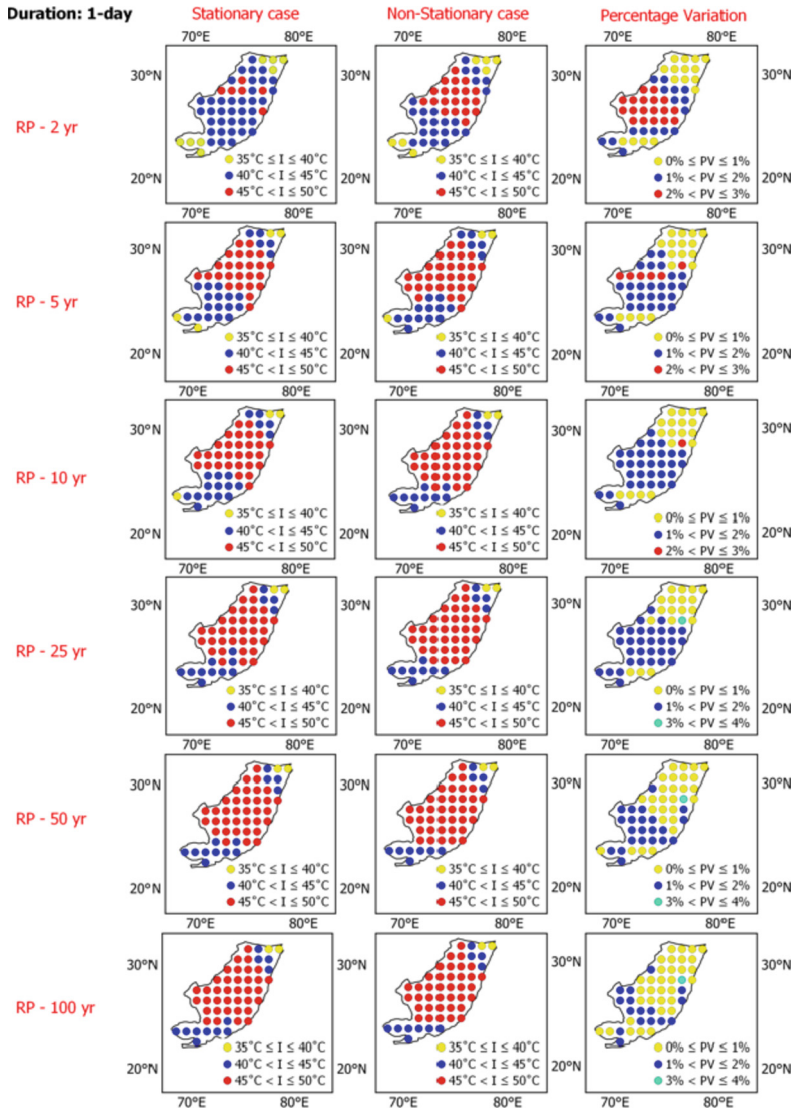

Fig. 6. S and NS temperature intensities and percentage variation for 1-day duration

When the stationary case is compared with non-stationary case it was found that the temperature intensities increased beyond 45 °C for many grid points when non-stationarity was considered. This points out that the stationary temperature intensities are lower than all of the corresponding non-stationary values for all grid points. The same was the case for all the other durations as well. Majority of the percentage variations

range between 0% and 2% i.e., about 0 °C to 0.9 °C. In the case of return period of 2 years around 36% of the grid points showed a percentage variation between 2% to 3% which reduced to the category of 0% to 2% as the return period increased. One grid point (28.5°N,76.5°E) showed a consistent increase in the percentage variation reaching a value of 3% to 4% (up to 1.6 °C) from 25 years to 100 years of return period. The pattern of percentage variations of return levels was observed closely by plotting a stem plot as shown in Fig. 7 for 1-day duration. It was seen that the graph follows two patterns i.e., PV decreases as RP increases and vice-versa. While comparing Fig. 6 with Fig. 3 it was observed that the grid points having NSGEV-1 as the best fit model followed the first pattern of PV decreases as RP increases while the grid points having NSGEV-4 and NSGEV-5 as the best fit model followed the second pattern of PV increases as RP increases.
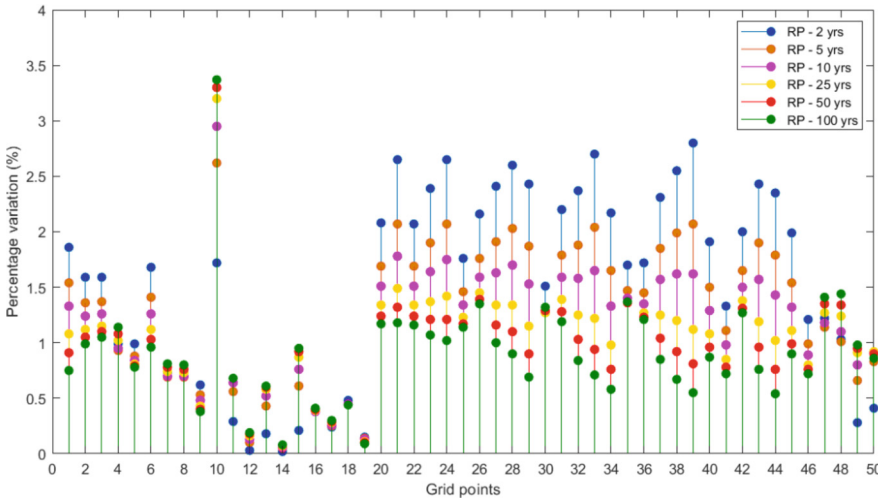

Fig. 7. Percentage variation of all grid points for 1-day duration

5   Conclusions

In this study, TDF analysis was carried out for NW temperature homogeneous region of India by developing S model and five NS models with time as covariate. It was found that more than 84% of the grid points approve NSGEV-1 model with location parameter linearly varying with respect to time as the best fit model with the help of AIC values. The best fit model tends to change for each duration for certain grid points. Temperature intensity versus duration plots were established for both S and NS case with each curve representing return period of 2 years, 5 years, 10 years, 25 years, 50 years and 100 years. A comparison of S-TDF curves and NS-TDF curves were carried for all the 50 grid points and it was found that the NS temperature intensities were more than the S estimates. The temperature intensity increased beyond 45 °C when non-stationarity was considered.

As the return period increases, the temperature intensities tend to increase for S and NS case. A PV of 0% to 2% was observed for majority of the grid points. In general, it can be concluded that for more than 84% of the grid points PV decreases as RP increases while for the remaining 16% PV increases as RP increases based on the best fit model for each grid point.

References

1. Hamdi, Y., Duluc, C.M., Rebour, V.: Temperature extremes estimation of non-stationary return levels and associated uncertainties. Atmosphere **9**(4), 129 (2018). https://doi.org/10.3390/atmos9040129.hal-02871800

2. Mazdiyasni, O., et al.: Increasing probability of mortality during Indian heat waves. Sci. Adv. **3**, e1700066 (2017)

3. Ouarda, T.B.M.J., Charron, C.: Nonstationary temperature-duration-frequency curves. Sci. Rep. **8**(1), 1–8 (2018). https://doi.org/10.1038/s41598-018-33974-y

4. Dash, S.K., Mamgain, A.: Changes in the frequency of different categories of temperature extremes in India. J. Appl. Meteorol. Climatol. **50**(9), 1842–1858 (2011). https://doi.org/10.1175/2011JAMC2687.1

5. Perkins, S.E., Alexander, L.V.: On the measurement of heat waves. J. Clim. **26**(13), 4500–4517 (2013). https://doi.org/10.1175/JCLI-D-12-00383.1

6. Rohini, P., Rajeevan, M., Srivastava, A.K.: On the variability and increasing trends of heat waves over India. Sci. Rep. **6**, 1–9 (2016). https://doi.org/10.1038/srep26153

7. Mishra, V., Mukherjee, S., Kumar, R., Stone, D.A.: Heat wave exposure in India in current, 1.5 °C, and 2.0 °C worlds. Environ. Res. Lett. **12**(12) (2017). https://doi.org/10.1088/1748-9326/aa9388

8. Khan, N., Shahid, S., Ismail, T., Ahmed, K., Nawaz, N.: Trends in heat wave related indices in Pakistan. Stoch. Env. Res. Risk Assess. **33**(1), 287–302 (2018). https://doi.org/10.1007/s00477-018-1605-2

9. Loikith, P.C., Broccoli, A.J.: The influence of recurrent modes of climate variability on the occurrence of winter and summer extreme temperatures over North America. J. Clim. **27**(4), 1600–1618 (2014)

10. Kothawale, D.R., Revadekar, J.V., Kumar, K.R.: Recent trends in pre-monsoon daily temperature extremes over India. J. Earth Syst. Sci. **119**(1), 51–65 (2010). https://doi.org/10.1007/s12040-010-0008-7

11. Ghasemi, A.R.: Changes and trends in maximum, minimum and mean temperature series in Iran. Atmos. Sci. Lett. **16**(3), 366–372 (2015). https://doi.org/10.1002/asl2.569

12. Khaliq, M.N., St-Hilaire, A., Ouarda, T.B.M.J, Bob'ee, B.: Frequency analysis and temporal pattern of occurences of Southern Quebec heatwaves. Int. J. Climatol. **25**, 485–504 (2005). https://doi.org/10.1002/joc.1141

13. Srivastava, A.K., Rajeevan, M., Kshirsagar, S.R.: Development of high resolution daily gridded temperature data set (1969–2005) for the Indian region. Atmos. Sci. Let. (2009). https://doi.org/10.1002/asl.232

14. Mann, H.B.: Non-parametric tests against trend. Econometrica **13**, 163–171 (1945)

15. Kendall, M.G.: Rank Correlation Techniques.Charles Griffen. London ISBN: 195205723 (1975)

16. Katz R.W.: Statistical methods for nonstationary extremes. In: AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., Sorooshian, S. (eds.) Extremes in a Changing Climate. Water Science and Technology Library, vol 65. Springer, Dordrecht (2013). https://doi.org/10.1007/978-94-007-4479-0_2