TRUE (Time-varying Return Levels of Univariate Extremes)

User Manual

Usman Mohseni[1], Mayank Gailakoti[2], Dr. Vinnarasi Rajendran[3]

[1]Research Scholar, Department of Civil Engineering, Indian Institute of Technology Roorkee, Uttarakhand, India. email: mohseni_ua@ce.iitr.ac.in

[2]M.Tech Student, Department of Civil Engineering, Indian Institute of Technology Roorkee, Uttarakhand, India. email: mayank_sgailakoti@ce.iitr.ac.in

[3]Assistant Professor, Department of Civil Engineering, Indian Institute of Technology Roorkee, Uttarakhand, India. email: vinnarasi@ce.iitr.ac.in

TRUE (Time-varying Return Levels of Univariate Extremes) is a MATLAB-based software package designed to facilitate non-stationary multi-covariate flood frequency analysis (NSMCFFA) under both stationary and non-stationary assumptions. TRUE allows users to incorporate multiple covariates such as Annual Precipitation (AP), Local Temperature Anomaly (LTA), Southern Oscillation Index (SOI), Indian Ocean Dipole (IOD), NDVI (Normalized Difference Vegetation Index), and Urbanization (URBAN) to describe changes in the statistical behavior of univariate extreme floods. TRUE estimates the parameters of candidate distributions (e.g., Generalized Extreme Value (GEV), Generalized Pareto (GP), Lognormal, and Gamma Type). TRUE utilizes Bayesian Inference (BI) in conjunction with the Differential Evolution Markov Chain Monte Carlo (DE-MCMC) technique for parameter estimation. The software supports both stationary models (constant parameters) and non-stationary models (parameters varying with covariates). The GUI provides diagnostic tools, return level plots, and exportable results. The toolbox is made freely available with a Graphical User Interface (GUI), enabling broad accessibility and serving as a valuable resource for advanced data analysis courses. Moreover, TRUE equips policymakers, researchers, and engineers with more precise tools to design hydraulic structures, optimize reservoir operations, and implement adaptive flood mitigation strategies.

TRUE Toolbox Contents

The TRUE toolbox consists of following components (Figure 1):

- TRUE_Code_Files: Contains all the codes and functions for the TRUE GUI.
- TRUE_RUN: Code to launch TRUE via the GUI.
- TRUE_GUI_NSMCFFA: Source codes for the TRUE GUI.
- TRUE_ NSMCFFA.m: Source code version for advanced users.
- Data: Consists of streamflow data for Handia station along with the six covariates.
- main.mlapp: The GUI for TRUE.
- TRUE_UserManual

Note: Do not rename, delete or move the folder or any contents of the TRUE Toolbox. The toolbox is compatible with MATLAB versions R2022a and later.

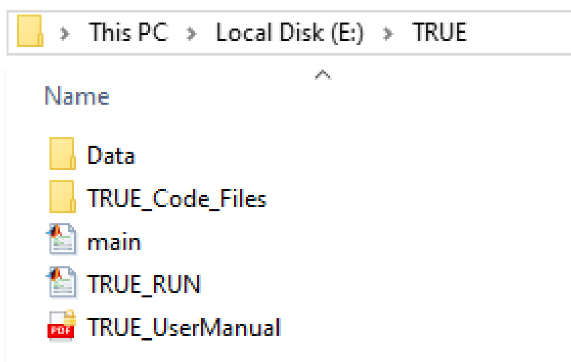

Figure 1. Contents of the TRUE Toolbox

Running TRUE

Note: Before running the GUI or source code, users are advised to ensure that the following MATLAB add-ons are installed on their system: (i) Curve Fitting Toolbox, (ii) Control System Toolbox, (iii) Parallel Computing Toolbox, and (iv) Statistics and Machine Learning Toolbox.

Running via GUI

1. Open MATLAB and set the folder 'TRUE' as the current directory.
2. Run the file TRUE_RUN.m'. The GUI window will open.
3. Select Data: Upload your data file (.csv) containing the time series. For covariates, each variable should be provided in a separate .csv file, and all files

must be stored in the same folder (see the ‘Covariates’ folder in the Data folder for reference).

4. Select Simulations: Any real number of simulations is acceptable. For preliminary analysis, a value of 1 may be used, while for more precise results a minimum of 20 simulations is recommended. Higher values (e.g., 100) can further improve accuracy but may increase computation time. Then, click the “Continue” button on the toolbar. The window 2 will pop up which is the second step.
5. Select Distribution: Select the appropriate probability distribution(s) for the given dataset. Multiple distributions can be chosen from the following options: GEV, GP, Lognormal, Gamma. Then, click the “RUN” button on the toolbar to obtain the best-fit distribution and the best covariate. Please be patient after clicking “RUN,” as the model may take some time to complete depending on the number of simulations and the specifications of the system on which TRUE is running.
6. Click “CONTINUE” to continue. One of the following windows will pop up (step 3), based on the result of the best fit in step 2:

   For the best distribution (GEV, GP, Gamma, or Lognormal), specify the prior distribution and associated parameters for each parameter: lower and upper bounds for uniform distribution; mean and standard deviation for normal distribution; shape and scale for gamma distribution. For GP, additionally select the type of threshold and its quantile value, and insert the number of observations in one year for plotting return level curves. In the case of a nonstationary analysis, the trend sections will be active. Select the type of trend: “none” refers to a constant parameter. If “none” is selected for all parameters, a stationary analysis will be performed. For more details about the types of distributions, refer to MATLAB help.
7. Click “CONTINUE” to continue. The last window will pop up (step 4). Specify the number of chains and iterations, and the burn-in period for MCMC. Select “YES” to plot and to save the results.
8. Click “RUN” to start the analysis. When the option to save the results is selected, all output files will be saved in the folder specified by fileLocation = "E:\TRUE" in the following formats: .fig, .tiff, .png, and .emf.

Running via Source Code

Advanced users can modify 'TRUE_NSMCFFA.m' located in the TRUE_Code_Files folder to directly load data, select distributions, and define covariates without using the GUI.

<u>Outputs</u>

Numerical Outputs

- Best distribution based on Akaike Information Criterion (AIC) and Kolmogorov–Smirnov (KS) test. The result is stored in the variable “best_fit_PF”.
- Best covariate based on minimum Deviance Information Criteria (DIC) (stored in “DIC_PF1_average”). The result of best covariate is stored in the variable “Best_Covariate_PF”. Refer to the file “Handia_Acceptable_Model.xlsx” in the Data folder as a reference for selecting the best covariate from the list of acceptable covariates.
- Stationary and non-stationary parameter estimates (Location, Scale, and Shape) using Bayesian Inference in conjunction with Differential Evolution Markov Chain Monte Carlo (DE-MCMC). For stationary analysis, the estimates are stored in the variables “sta_para_PF”. For non-stationary analysis, the estimates are stored in “para_PF”.

Graphical Outputs

- Figure 1: Peak flow series obtained using Block Maxima approach.
- Figure 2: Trend in parameters of the best fitted distribution.
- Figure 3: Trace (left panel) and Histogram (right panel) plots for the posterior samples of the non-stationary model of peak flow.
- Figures 4 & 5: Stationary return levels (Figure 4) and Non-stationary return levels (Figure 5). With the shaded region representing the ensemble spread, the solid blue line denotes the mean return level, while the dashed red lines indicate the 95% confidence interval (upper bound) and the 5% confidence interval (lower bound).

Example Case Study: Narmada River Basin

TRUE was to annual maxima streamflow series for five gauging stations in the Narmada River Basin (NRB). The figure 2 illustrates the application of TRUE for Handia station using the TRUE (GUI).

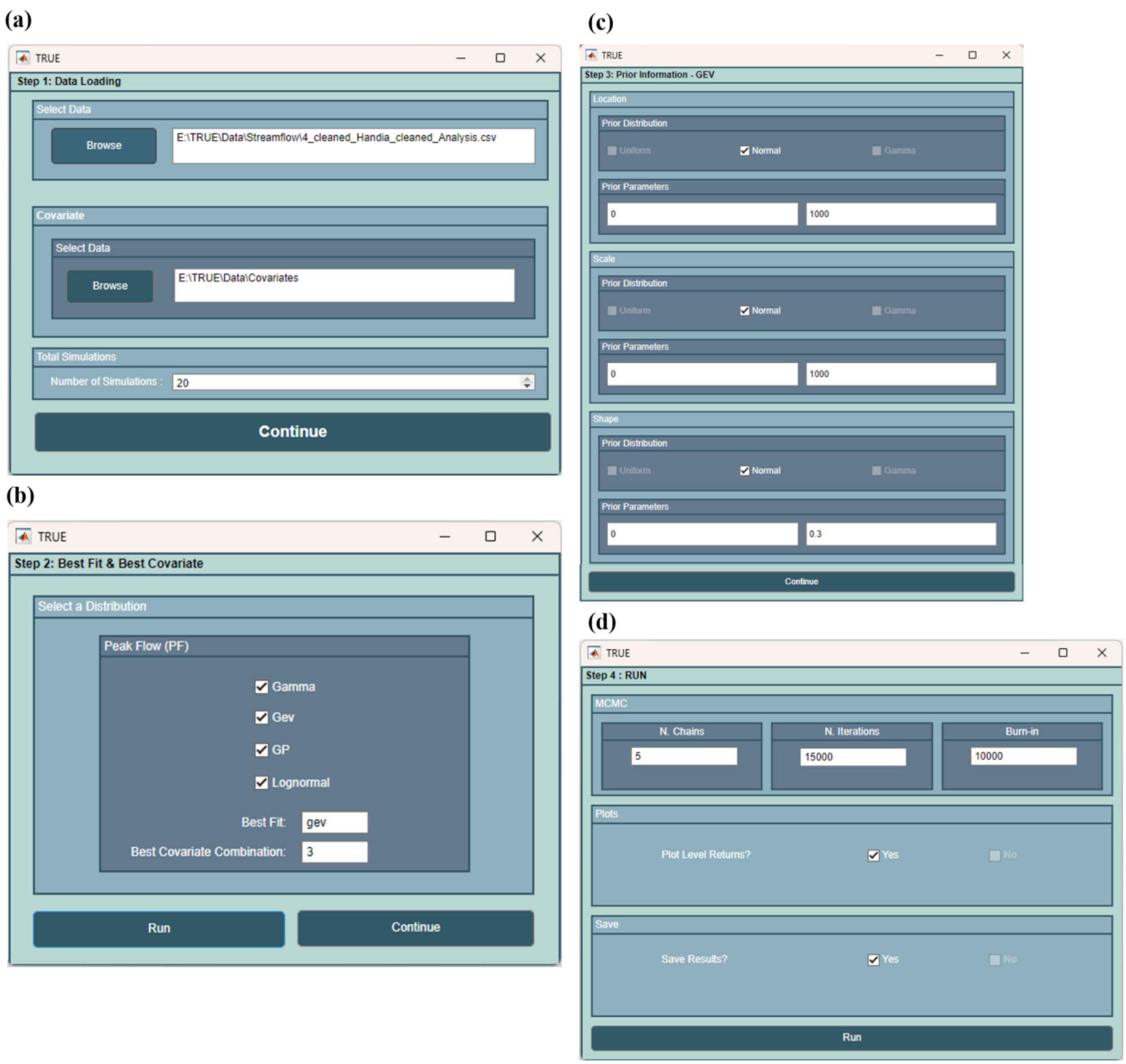

Figure 2. GUI version TRUE (Time-varying Return level for Univariate Extremes) for Handia station. (a) Interface for uploading data and choice of simulations; (b) Interface for selecting the choice of distribution (GEV/GP/Lognormal/Gamma); (c) Interface specific to the choice of distribution for selecting priors and nonstationary model; (d) Interface for selecting MCMC information, plotting options, and saving results.

Disclaimer

TRUE is provided 'as is' without warranty of any kind. The developers are not responsible for errors, misuse, or misinterpretation of results. Users should validate the outputs before applying them in design or decision-making. For inquiries regarding collaboration or professional training on the use of this toolbox, please contact the first author (**Usman Mohseni: mohseni_ua@ce.iitr.ac.in**) or third author (**Dr. Vinnarasi Rajendran: vinnarasi@ce.iitr.ac.in**).

References

Coles, S. (2001). An Introduction to Statistical Modeling of Extreme. In Journal of Chemical Information and Modeling (Vol. 53, Issue 9).

Cheng, L., AghaKouchak, A., Gilleland, E., & Katz, R. W. (2014). Non-stationary extreme value analysis in a changing climate. Climatic Change, 127(2), 353–369. https://doi.org/10.1007/s10584-014-1254-5

Luke, A., Vrugt, J. A., & Sanders, B. F. (2017). Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the United States. Water Resources Research, 5(3), 2–2. https://doi.org/10.1111/j.1752-1688.1969.tb04897.x

Ragno, E., AghaKouchak, A., Cheng, L., & Sadegh, M. (2019). A generalized framework for process-informed nonstationary extreme value analysis. Advances in Water Resources, 130(June), 270–282. https://doi.org/10.1016/j.advwatres.2019.06.007