Process-informed Nonstationary Extreme Value Analysis (ProNEVA) User Manual

E. Ragno[1], A. AghaKouchak[1], L. Cheng[2], M. Sadegh[3]

[1] *Department of Civil and Environmental Engineering, University of California, Irvine, USA.*

[2] *University of Colorado, and Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA.*

[3] *Department of Civil Engineering, Boise State University, ID, USA.*

Process-informed **N**onstationary **E**xtreme **V**alue **A**nalysis (ProNEVA) is a Matlab software package designed to facilitate extreme value analysis (EVA) under both stationary and nonstationary assumptions. ProNEVA allows using time or a physically-based covariate to describe change in statistics of extremes. Examples of physically-based covariates include: change in runoff extremes in response to urbanization, or change in temperature extremes as a function of CO_2 in the atmosphere. ProNEVA estimates the parameters of the Generalized Extreme Value (GEV), the Generalized Pareto (GP), and the Log-Pearson Type III (LP3) distributions. The model includes a Bayesian approach and a hybrid Markov Chain Monte Carlo (MCMC) method for sampling from the posterior distribution. ProNEVA also provides diagnostic tests and return period-return level plots. The toolbox is released along with a Graphical User Interface (GUI) so that it can reach a broad audience. Moreover, the toolbox can be a valuable educational tool for advanced data analysis courses. By using ProNEVA, the users agree to the below disclaimer.[1]

Reference Publication

Ragno E, AghaKouchak A, Cheng L, Sadegh M, 2019, A Generalized Framework for Process-informed Nonstationary Extreme Value Analysis, *Advances in Water Resources*, 130, 270-282, doi: 10.1016/j.advwatres.2019.06.007.

[1] **Disclaimer:** The Process-informed Nonstationary Extreme Value Analysis (ProNEVA) software package is provided 'as is' without any endorsement made and without warranty of any kind, either express or implied. While we strive to ensure that ProNEVA is accurate, no guarantees for the accuracy of the codes, output information and figures are made. ProNEVA codes and outputs can only be used at your own discretion and risk and with agreement that you will be solely responsible for any damage and that the authors and their affiliate institutions accept no responsibility for errors or omissions in ProNEVA codes, outputs, figures, and documentation. In no event shall the authors, developers or their affiliate institutions be liable to you or any third parties for any special, direct, indirect

ProNEVA Folder Content

The software package ProNEVA contains the following components (Figure 1):

- **GUIpackage.** This folder includes the source codes of the ProNEVA GUI. Do not rename or move the folder or its content.
- **ProNEVApackage.** This is the folder that contains the source codes for stationary and nonstationary extreme value analysis, diagnostic tests, and plots. Do not rename or move the folder or its contents.
- **RUN_GUI_for_ProNEVA.** This is the code for running ProNEVA using a Graphical User Interface (GUI). The GUI will guide users in selecting inputs and run specifics.
- **RUN_ProNEVA.** This is the source code of ProNEVA and you can use this to run the code without the GUI. Using this code, the user should edit the portion of the code associated with the inputs before running it. This standalone code allows advanced users to incorporate ProNEVA within their own codes.
- **Disclaimer.** By using ProNEVA users agree to the disclaimer. Please read the document before using the software.
- **UserManual_ProNEVA**.

RUN ProNEVA

The ProNEVA software package can be executed via (1) the GUI or (2) the main source code based on the preference of the user. The outputs of ProNEVA do not depend on this choice.

ProNEVA via GUI

(1) **Open** Matlab and **select** the folder "ProNEVA" as the current folder.

---

or consequential damages and financial risks of any kind, or any damages whatsoever, resulting from, arising out of or in connection with the use of ProNEVA. The user of ProNEVA agrees that the codes and algorithms are subject to change without notice.

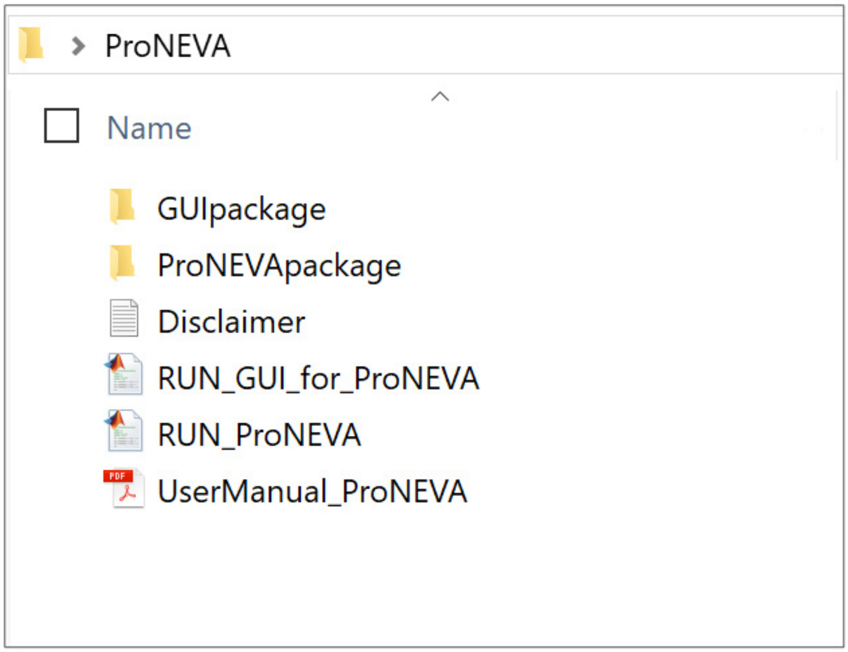

Figure S 1: Content of the folder ProNEVA.

(2) **Run** RUN_GUI_for_ProNEVA.m code by either (a) typing "RUN_GUI_for_ProNEVA" on the command window or (b) opening "RUN_GUI_for_ProNEVA.m" in the editor subsection and clicking the "RUN" button on the toolbar. The window in Figure 2 will pop up.

(3) In the section "Select Data", click on **"BROWSE"** and navigate to the file containing the data. The data must be formatted as **one vector** stored in a **text file** (*.txt* extension). Before uploading the file, make sure that the data are processed as follows:

GEV: Block maxima.

GP: Complete time series. The threshold will be selected in the next step.

LP3: Annual maxima in real or original space. ProNEVA will automatically transform the data in log-space.

(4) In the section "Select Distribution", **select** the desired distribution – GEV, GP, or LP3 – in accordance with uploaded data.

(5) In the section "Select Model", **select** between stationary and nonstationary analysis

(6) In the case of a nonstationary analysis, the section "Covariate" will be active. **Select** the type

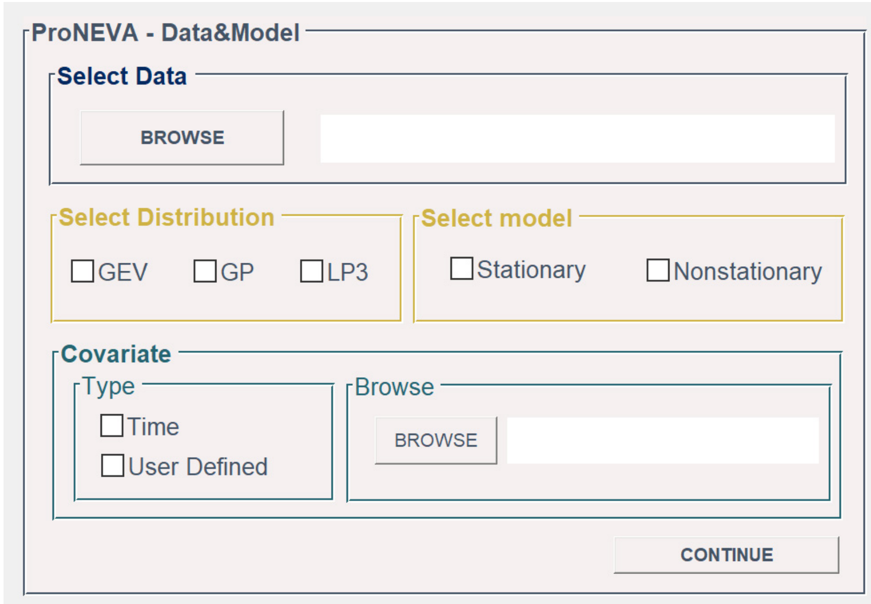

Figure S 2: ProNEVA GUI - Data&Model for selecting data, distribution, and model type.

of covariate between “Time” and “User Defined.” In the case of “User Defined” covariate, click on “BROWSE” and navigate to the text file containing the data. The data must be formatted as **one vector** and the **same length** as the main variable selected in Step (3).

(7) Click **“CONTINUE”** to continue. One of the following windows will pop up, based on the distribution type selected in Step (4):

Figure 3 when a **GEV** distribution is chosen. For each parameter of the GEV distribution, specify the prior distribution and associated parameters. Specifically, lower and upper bound for uniform distribution; mean and standard deviation for normal distribution; shape and scale for gamma distribution. For more details about the types of distributions refer to Matlab help. 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 the parameters, a stationary analysis will be performed.

Figure 4 when a **GP** distribution is chosen. Select the type of threshold. Select the quantile to determine the value of the threshold parameter. Insert the number of observation in one year; this variable will be used for plotting return level curves. For each parameter of the GP distribution, specify the prior distribution and associated parameters. Specif-

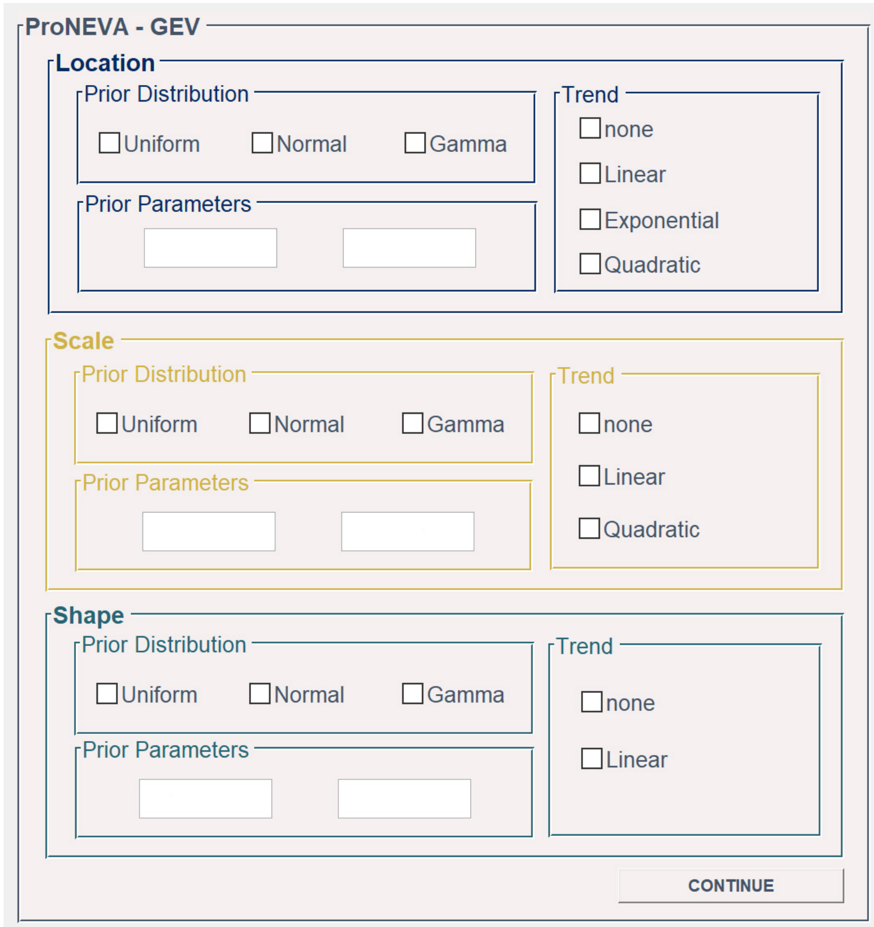

Figure S 3: ProNEVA GUI. Select priors and trends for the GEV parameters.

ically, lower and upper bound for uniform distribution; mean and standard deviation for normal distribution; shape and scale for gamma distribution. For more details about the types of distributions, refer to the Matlab help. 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 the parameters, a stationary analysis will be performed.

Figure 5 when a **LP3** distribution is chosen. For each parameter of the LP3 distribution, specify the prior distribution and associated parameters. Specifically, lower and upper bound for uniform distribution; mean and standard deviation for normal distribution; shape and scale for gamma distribution. For more details about the types of distributions, refer to the Matlab help. 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

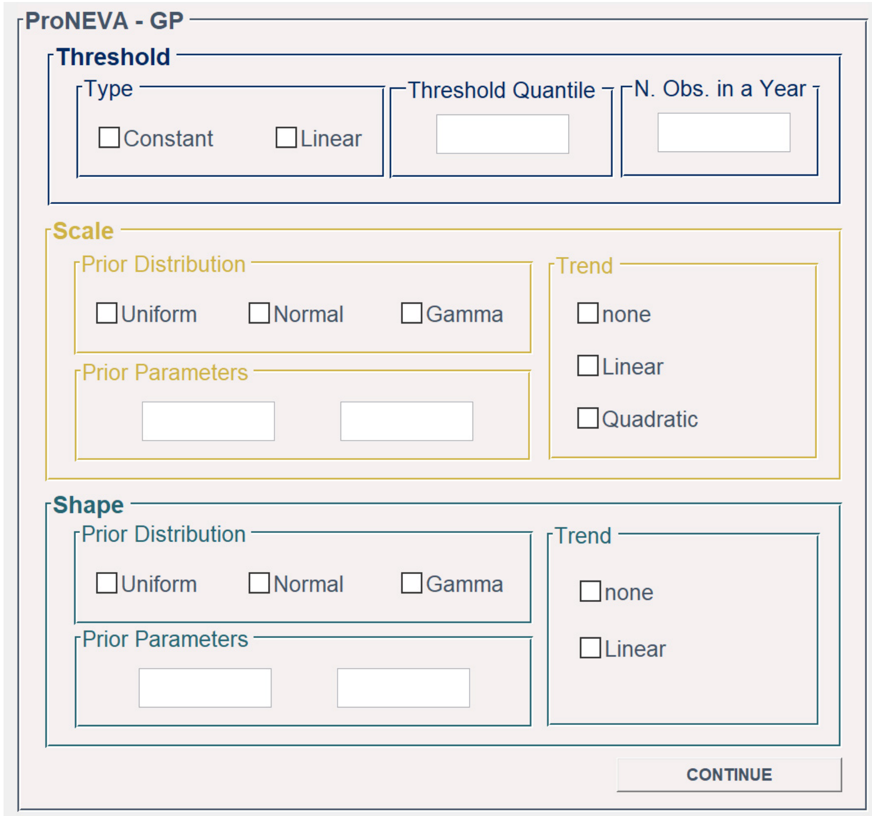

Figure S 4: ProNEVA GUI. Select threshold type along with priors and trends for the GP parameters.

selected for all the parameters, a stationary analysis will be performed.

(8) Click **"CONTINUE"** to continue. The last window will pop up, Figure 6. Specify the number of chains and iterations, and the burn-in period for MCMC. Specify the maximum return period of interest for return level curves. Select "YES" to the three questions to perform Man-Kendall trend test and White test, to plot return level curves, and to save the results.

(9) Click **"RUN"** and the analysis will start. When the option to save the results is selected, the folder "Results" containing the outputs (.mat and .fig files) will be created in the folder "ProNEVA".

ProNEVA via Source Code

Users can perform stationary and nonstationary extreme value analysis using ProNEVA avoiding the GUI. To do so:

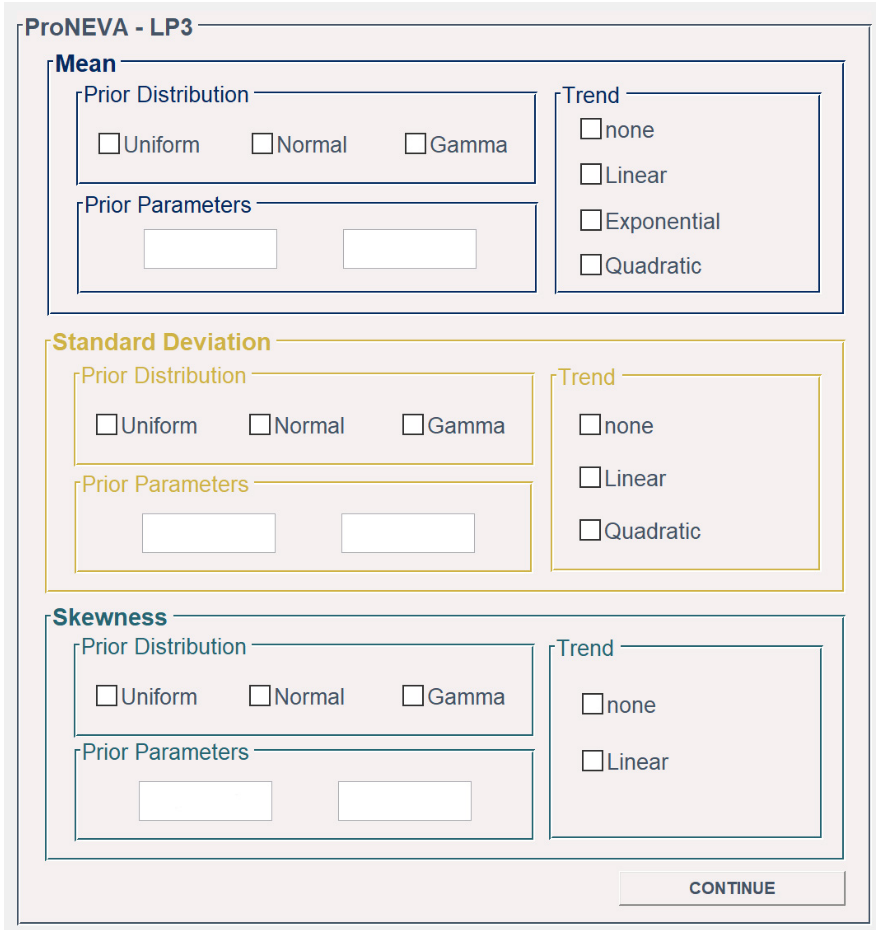

Figure S 5: ProNEVA GUI. Select priors and trends for the LP3 parameters.

(1) **Open** Matlab and **select** the folder “ProNEVA” as the current folder.

(2) **Open** “RUN_ProNEVA.m” in Matlab Editor.

(3) **Edit** section “Load Data”. **Replace** “DataPath\MyData.txt‘ with the name of the desired file. Include the file path when the file .txt is outside the folder “ProNEVA”.

```
1 %% (1) EDIT - LOAD DATA
2 fileOBS = fopen('DataPath\MyData.txt');
```

(4) **Edit** section “Distribution Type”. Specify the distribution of interest in Line 7 based on the legend.

```
1 %% (2) EDIT - DISTRIBUTION TYPE
2 % RUNspec.DISTR.Type
```

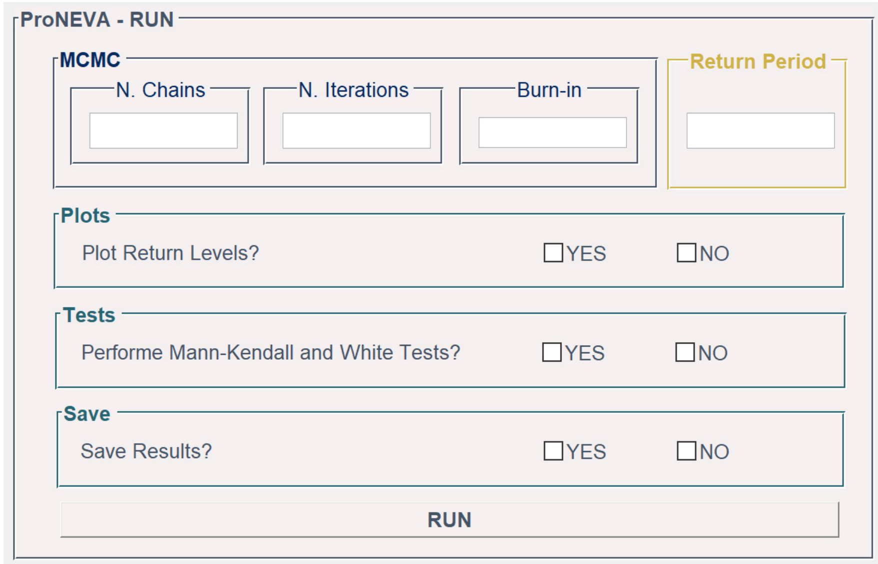

Figure S 6: ProNEVA GUI for selecting the parameters for MCMC along with complementary options such as tests, plots, and save.

```
% (i)   RUNspec.DISTR.Type = 'GEV'  Generalized Extreme Value
    Distribution
% (ii)  RUNspec.DISTR.Type = 'GP'   Generalized Pareto Distribution
% (iii) RUNspec.DISTR.Type = 'P3'   Pearson Typer III
RUNspec.DISTR.Type = 'GEV';
```

(5) **Edit** section "Model Type". Specify the type of model in line 5, stationary or nonstationary. In the case of nonstationary analysis, RUNspec.DISTR.Model = 'NonStat', specify the type of covariate. If RUNspec.COV.type = 'User', replace 'CovPath\MyCovariate.txt' with the name of the text file containing the vector of covariate. Include the file path if the file is outside the ProNEVA folder.

```
1 %% (3) EDIT - MODEL TYPE
2 % 'Stat'    : Stationary Analysis
3 % 'NonStat' : Nonstationary Analysis
4
5 RUNspec.DISTR.Model = 'Stat';
6
7 if strcmp(RUNspec.DISTR.Model, 'NonStat';
8
```

```
    %% EDIT - COVARIATE TYPE
    % RUNspec.COVtype:
    % (i)  RUNspec.COV.type = 'Time'
    % (ii) RUNspec.COV.type = 'User'
    RUNspec.COV.Type = 'Time';

    if strcmp(RUNspec.COV.Type, 'User')

        %% EDIT - SELECT FILE COVARIATE
        fileCOV = fopen('CovPath\MyCovariate.txt');
        % DO NOT EDIT
        textCOV  = textscan(fileCOV, '%f');
        fclose(fileCOVE);
        RUNspec.COV.X = textCOV{1}(:);
    end
end
```

(6) **Uncomment** and **edit** the sections in Listings 1, 2, and 3 based on the type of distribution chosen at Step (4). For each distribution's parameter, specify the type of prior and associated parameters, following the legend at the top. In the case of a GP distribution, specify the threshold quantile and the type, along with the number of observations in a year, used for return level plots. In the case of nonstationary analysis, define the type of trend for the different parameters.

```
%% (4) UNCOMMENT and EDIT if RUNspec.DISTR.Type = 'GEV'

%% Edit PRIOR
% (i)   'Uniform': parm1 = min   | parm2 = max
% (ii)  'Normal' : parm1 = mean  | parm2 = std
% (iii) 'Gamma'  : parm1 = shape | parm2 = scale

% Location - MU:
RUNspec.PRIOR.MUdistr = 'Normal';
RUNspec.PRIOR.MUparm1 = 0;
```

```
RUNspec.PRIOR.MUparm2 = 100;

% Scale - SI:
RUNspec.PRIOR.SIdistr = 'Normal';
RUNspec.PRIOR.SIparm1 = 0;
RUNspec.PRIOR.SIparm2 = 10;

% Shape - XI:
RUNspec.PRIOR.XIdistr = 'Normal';
RUNspec.PRIOR.XIparm1 = 0;
RUNspec.PRIOR.XIparm2 = 0.2;

% DO NOT EDIT
if strcmp(RUNspec.DISTR.Model, 'Stat')

    RUNspec.NS.MU = 'none';
    RUNspec.NS.SI = 'none';
    RUNspec.NS.XI = 'none';
else
    %% EDIT TREND  'NonStat' case
    % TREND LOCATION
    % 'none' | 'Linear' | 'Quadratic' | 'Exponential'
    RUNspec.NS.MU = 'Linear';
    % TREND SCALE
    % 'none' | 'Linear' | 'Quadratic'
    RUNspec.NS.SI = 'Linear';
    % TREND SHAPE
    % 'none' | 'Linear'
    RUNspec.NS.XI = 'none';
end
```

Listing 1: Section for GEV.

```
%% (4) UNCOMMENT AND EDIT if RUNspec.DISTR.Type = 'GP'

```

```
%% EDIT GP THRESHOLD
% RUNspec.THtype: (i) 'Const' | (ii) 'QR' - Quantile Regression
RUNspec.THtype = 'Const';
% RUNspec.THp: p-quantile for threshold definition [0 1]
RUNspec.THp    = 0.98;
% RUNspec.NobsY: Observations in a year
RUNspec.NobsY  = 365;

%% EDIT PRIOR
% (i)   'Uniform': parm1 = min   | parm2 = max
% (ii)  'Normal' : parm1 = mean  | parm2 = std
% (iii) 'Gamma'  : parm1 = shape | parm2 = scale

% Scale
RUNspec.PRIOR.SIdistr = 'Normal';
RUNspec.PRIOR.SIparm1 = 0;
RUNspec.PRIOR.SIparm2 = 10;

% Shape
RUNspec.PRIOR.XIdistr = 'Normal';
RUNspec.PRIOR.XIparm1 = 0;
RUNspec.PRIOR.XIparm2 = 0.2;

% DO NOT EDIT
if strcmp(RUNspec.DISTR.Model, 'Stat')

    RUNspec.NS.MU = 'none';
    RUNspec.NS.SI = 'none';
    RUNspec.NS.XI = 'none';
else
    %% EDIT TREND  'NonStat' case
    % TREND SCALE
    % 'none' | 'Linear' | 'Quadratic'
```

```
    RUNspec.NS.SI = 'Linear';
    % TREND SHAPE
    % 'none' | 'Linear'
    RUNspec.NS.XI = 'none';
end
```

Listing 2: Section for GP.

```
%% (4) UNCOMMENT and EDIT if RUNspec.DISTR.Type = 'P3'

%% Edit PRIOR
% (i)   'Uniform': parm1 = min   | parm2 = max
% (ii)  'Normal' : parm1 = mean  | parm2 = std
% (iii) 'Gamma'  : parm1 = shape | parm2 = scale

% Location - MEAN:
RUNspec.PRIOR.MUdistr = 'Normal';
RUNspec.PRIOR.MUparm1 = 0;
RUNspec.PRIOR.MUparm2 = 100;

% Scale - STANDARD DEVIATION:
RUNspec.PRIOR.SIdistr = 'Normal';
RUNspec.PRIOR.SIparm1 = 0;
RUNspec.PRIOR.SIparm2 = 10;

% Shape - SKWENESS:
RUNspec.PRIOR.XIdistr = 'Normal';
RUNspec.PRIOR.XIparm1 = 0;
RUNspec.PRIOR.XIparm2 = 0.2;

% DO NOT EDIT
if strcmp(RUNspec.DISTR.Model, 'Stat')

    RUNspec.NS.MU = 'none';
    RUNspec.NS.SI = 'none';
```

```
    RUNspec.NS.XI = 'none';
else
    %% EDIT TREND  'NonStat' case
    % TREND LOCATION
    % 'none' | 'Linear' | 'Quadratic' | 'Exponential'
    RUNspec.NS.MU = 'Linear';
    % TREND SCALE
    % 'none' | 'Linear' | 'Quadratic'
    RUNspec.NS.SI = 'Linear';
    % TREND SHAPE
    % 'none' | 'Linear'
    RUNspec.NS.XI = 'none';
end
```

Listing 3: Section for LP3.

(7) **Edit** MCMC information and optional results in Listing 4. Specify the desired number of chains and iterations, and the burn-in period for MCMC approach. Specify the maximum return period for return level curves. Finally, specify whether ProNEVA will perform the Mann-Kendall trend test and White test, plot return level curves, and save the results. When the option to save the results is selected, a folder "Results" will be created in the folder "ProNEVA" containing the analysis outputs (.mat and .fig files).

```
%% (5) EDIT - MCMC AND EXTRA OPTIONS
% MCMC
% Number of Chains
RUNspec.Nchain  = 3;
% Number of Iterations
RUNspec.maxIT   = 10000;
% Burn-in period
RUNspec.brn     = 9000;
% Return Period
RUNspec.RP      = 100;

```

```
% Extra Options
% 'Y': Yes - 'N': No
% Save Results? 'Y' /'N'
EXTRAS.saveRES  = 'Y';
% Run Mann-Kendall and White Tests? 'Y'/'N'
EXTRAS.RunTests = 'Y';
% Plot Return Level? 'Y'/'N'
EXTRAS.PlotRL   = 'Y';
```

Listing 4: Section for LP3.

(8) **Run** the code.

ProNEVA Results

When the option to save the data is selected, a folder "Results" will be created and it will contain the outputs from the run. Table 1 summarizes the expected outputs. However, some outputs may not be available because of the type of the input previously selected.

OUTPUT.	
CH	Parameters of the selected distribution
RhatCH	Gelman $\hat{R}$ for convergency check
Z	Standardized Observations for GOF tests
RES	Residuals
EWT	Expected Waiting Time
Z_{q0}	Quantile associated with the EWT
RLplot.VC	Covariate values for the return level plots
PARvc	Distribution parameters for covariate equal to RLplot.VC
RLplot.RL95	Upper bound of return level curves (95-percentile)
RLplot.RL05	Lower bound of return level curves (5-percentile)
RLplot.RL50	Expected return level curves (median)
RLplot.RLm	Maximum likelihood return level curves
ERP.TT	Return period for effective return level plot
RLeff	Effective return level curves: each row corresponds to ERP.TT
PDFhat	Predictive PDF
DGN.	
KS.HH	Kolmogorov Smirnov (KS) Test Results
KS.RJrate	KS test rejection rate
AIC	Akaike Information Content
BIC	Bayesian Information Criterion
RMSE	Root Mean Square Error
NSE	NashSutcliffe model Efficiency Coefficient
TST.	
MK.H	Statistical significance of Mann-Kendall trend test
MK.p_value	p_value for the Mann-Kendall trend test
WT.H	Statistical significance of White test
WT.p_value	p_value for the White test

Table S 1: List of ProNEVA outputs