Programs for
"Multiscale Inference for Nonparametric Time Trends"

Marina Khismatullina
University of Bonn

Michael Vogt
University of Bonn

April 30, 2018

In this document we describe the programs and functions that can be used to replicate all the empirical results reported in our paper. We first explain the model simulations and then the empirical results that were obtained with the temperature data from UK Met Office. The results are stored either as figures or as `.tex` files (for tables) in the folder `.\Plots`.

All programs are written in R with some functions in C. They are all quite self-explanatory and commented. The code is self-sufficient, the only thing that needs to be changed is the extensions of dynamic libraries as they are OS-specific. Dynamic libraries are defined in the beginning of each main file. At the moment everything is configured for running on Windows, that is, the `.dll` extension is used. In order to make the programs Linux or MacOS compatible, this extension should be changed to `.so`. For example,

```
dyn.load("Shape/C_code/estimating_sigma.dll")
```

should be replaced by

```
dyn.load("Shape/C_code/estimating_sigma.so").
```

All the dynamic libraries are already provided, additional compilation is not neccessary.

1 Model Simulations

1.1 Testing for the presence of the time trend

The folder `.\Shape` contains all the necessary functions, data and simulated distributions that were used for calculatig size and power of our method for testing for the presence of time trends in Section 7. In order to obtain Tables 1 and 2, one needs to run `.\main_shape_simulation.R`. As was already mentioned, the only possible adjustment that needs to be done is the change of extension of dynamic libraries in the beginning of `.\main_shape_simulation.R`.

The simulated distributions of Gaussian statistic Φ'_T for various lengths of time trends T can be found in the directory `.\Shape\distribution`. In order to be able to recalculate these distributions, the stored files need to be deleted or removed; the computation is then done automatically.

Size and power of the test are calculated by generating 1000 samples. This number can be set as `N_rep` in `.\main_shape_simulation.R`. We have run our simulations for 10000 samples as well, and the running time is quite small.

The dynamic libraries and R-wrappers of respective C functions are provided in the folder `.\Shape\C_code`.

All the auxiliary functions can be found in `.\Shape\functions.R`.

1.2 Testing for equality of the trends

The folder `.\Equality` contains all the necessary functions, data and simulated distributions that were used for testing the equality of time trends in Section 7. Specifically, in order to obtain Tables 3, 4 and 5, one needs to run `.\main_equality_simulation.R`. Do not forget to change the extensions of dynamic libraries in the very beginning of `.\main_equality_simulation.R` (if needed).

The directory `.\Equality\distribution` comprises the simulated distributions of gaussian statistic $\Phi_{n,T}$ for $n = 15$ time series (the number that was used for the simulation study) and different lengths of time trends T. If you want to recalculate the distribution, the stored files need to be deleted or removed. The recomputation is then done automatically. However, note that with increasing values of time series length T the running time increases significantly. For example, for $T = 1000$ and $n = 15$ calculation of $\Phi_{n,T}$ will probably take several hours.

The dynamic libraries and R-wrappers of respective C functions are provided in the folder `.\Equality\C_code`.

All the auxiliary functions can be found in `.\Equality\functions.R`.

Size and power of the test are calculated generating 1000 samples. This number can be set as `N_rep` in `.\main_equality_simulation.R`, though rising this parameter is not advisable due to computational tractability.

2 Application

2.1 Testing for the presence of the time trend

The file `.\Shape\data\cetml1659on.dat` contains the monthly and yearly mean central England temperature from the year 1659 up to year 2017. It was downloaded from

`https://www.metoffice.gov.uk/hadobs/hadcet/data/download.html`.

Full description of the data can be accessed on the official website of The UK Met Office `https://www.metoffice.gov.uk`.

As for the simulation study, the dynamic libraries and R-wrappers of respective C functions are provided in the directory `.\Shape\C_code`.

The distribution of the Gaussian statistic Φ'_T, that was used for calculating quantiles for the application, is stored in the file

`.\Shape\distribution\distr_T_359_testing_constant_type_ll.RData`.

As before, the recalculating of the Gaussian statistic distribution can be done by removing or deleting this file before running `N_rep` in `.\main_equality_simulation.R`.

All the auxiliary functions can be found in `.\Shape\functions.R`.

2.2   Testing for equality of the trends

The directory `.\Equality\data` contains the historic station data in `.txt` format downloaded from

`https://www.metoffice.gov.uk/public/weather/climate-historic/#?tab=climateHistoric`

and the description file `.\Equality\Description_file.xlsx`. The data consists of:

- Mean daily maximum temperature ($tmax$)

- Mean daily minimum temperature ($tmin$)

- Days of air frost (af)

- Total rainfall ($rain$)

- Total sunshine duration (sun)

The monthly mean temperature that was used in the paper is calculated from the average of the mean daily maximum and mean daily minimum temperature, i.e. ($tmax + tmin$)/2. Full description of the data can be accessed on the official website of The UK Met Office `https://www.metoffice.gov.uk`.

The distribution of the Gaussian statistic $\Phi'_{n,T}$, that was used for calculating quantiles for this application, is stored in the file

`.\Equality\distribution\`
`distr_for_application_T_386_and_N_ts_25_and_method_ll.RData`.

You can recalculate the distribution in order to check our results. In order to do so, delete or move the file mentioned above and the dstribution of the new Gaussian statistic will be computed automatically.

As for the simulation study, the dynamic libraries and R-wrappers of respective C functions are provided in the folder `.\Equality\C_code`.

All the auxiliary functions can be found in `.\Equality\functions.R`.