Supplementary material to the paper:
"Kernel-based Hidden Markov Conditional Densities"

by Jan G. De Gooijer, Gustav Eje Henter, and Ao Yuan

Published in: *Computational Statistics and Data Analysis*, 2022, 169.
https://doi.org/10.1016/j.csda.2022.107431.

The files/codes in the package **"kde-hmm-material-2016.zip"** can be used to replicate some of the results in the paper by Henter, Leijon, and Kleijn (2016). These files/codes are also used by De Gooijer, Henter, and Yuan (2022). The package **"kcde-hmm-material-2022.zip"** contains some additional files.

Package: kde-hmm-material-2016.zip

1. **"ar_synthetic":** This folder contains results and figures pertaining to Sect. 5 of the paper by Henter, Leijon, and Kleijn (2016).

2. **"ar_kde":** This folder contains materials pertaining to the experiments in the paper by Henter, Leijon, and Kleijn (2016) in general. This includes data, scripts, text logs, and subfolders with figures and Matlab (ML) code. The folder also contains data and results for a number of preliminary experiments not reported in the paper.

    - **"ar_kde/copied_code/ar_kde":** This folder contains ML codes and classes that implement the AR-HMM, KDE-MM, and KDE-HMM, as well as run experiments on selected datasets.
    - **"ar_kde/copied_code/PattRecClasses":** This folder contains an internal toolbox implementing standard models such as HMMs and Markov chains.
    - **"ar_kde/copied_code/my_util":** This folder contains assorted ML utility functions used, for instance, for plotting.
    - **"ar_kde/laser_3k_phaseinit_data.mat":** This file contains the time-series data used for the experiment on the laser data reported in the paper.
    - **"ar_kde/run_laser_3k_phaseinit.m":** This file launches the full laser data experiment reported in the paper. Very time-consuming.
    - **"ar_kde/laser_3k_phaseinit":** This folder contains figures pertaining to the experiment on the laser data reported in the paper.
    - **"ar_kde/laser_3k_phaseinit/laser_3k_phaseinit_results.mat":** This file contains the results (including trained models) for the experiment on the laser data reported in the paper.
    - **"ar_kde/ecg_3k_phaseinit_data.mat":** This file contains the time-series data used for the experiment on the ECG data reported in the paper.
    - **"ar_kde/ecg_3k_phaseinit":** This folder contains figures pertaining to the experiment on the ECG data reported in the paper.
    - **"ar_kde/run_ecg_3k_phaseinit.m":** This file launches the full ECG data experiment reported in the paper. Very time-consuming.
    - **"ar_kde/ecg_3k_phaseinit/ecg_3k_phaseinit_results.mat":** This file contains the results (including trained models) for the experiment on the ECG data reported in the paper.

Package: kcde-hmm-material-2022.zip

- **"CPAtest2.m":** This function performs the asymptotic conditional predictive ability test of Giacomini and White (2006).
- **"run_HMR-STAR-50windows-n1000-H2.m":** This function computes some of the test results reported in Section 5.1, Table 1 (HMRS-STAR model) of the CSDA paper. Forecast horizon $H = 2$, and $R = 1,000$.
  Supporting files: **"default_opts3.m"** and **"kdehmm_experiment5.m"**.
- **"run_AR_HMM_Gamma.m":** This function computes some of the test results reported in Section 5.2, Table 2 (AR-HMM model) of the CSDA paper. 4 distribution functions and Newey-West correction for autocorrelation.
  Supporting files: **"default_opts3.m"** and **"kdehmm_experiment6_M3.m"**.
- **"run_empirical.m":** This function computes various results plotted in Figures 3 and 4 (Section 4) of the CSDA paper.
  Data: **"SP500_T2012.dat"** and **"Rate_ALL_1214.dat"**.
  Supporting files: **"default_opts3.m"** and **"kdehmm_experiment6_M3.m"**.

The following files contain data and results for a number of experiments not reported in the paper by De Gooijer, Henter, and Yuan (2022).

- **"KCDE_HMM_SP500.pdf":** This file shows a plot of the conditional density of the S&P 500 series $\{Y_t\}$ conditional on the realized value Y_{t-1} at three quantile levels τ using the KCDE-HMM method: $\tau = 0.9$ (yellow curve), $\tau = 0.5$ (red curve; median), and $\tau = 0.1$ (blue curve). Similarly, the file **"KCDE_HMM_Ind_Prod.pdf"** shows a plot of the conditional density of the U.S. Industrial Production series using the KCDE-HMM method.
- **"KCDE_MM_SP500.pdf":** This file shows a plot of the conditional density of the S&P 500 series $\{Y_t\}$ conditional on the realized value Y_{t-1} at three quantile levels τ using the KCDE-MM method (Markov model order $p = 1$): $\tau = 0.9$ (yellow curve), $\tau = 0.5$ (red curve; median), and $\tau = 0.1$ (blue curve). Similarly, the file **"KCDE_MM_Ind_Prod.pdf"** shows a plot of the conditional density of the U.S. Industrial Production series using the KCDE-MM method.
- **"condlogprob.m"** gives a quick demonstration of how to compute and graph conditional probability densities for trained models. The code was used to construct the pdf-plots given above.
  Supporting files: **"SP500_2012_all.mat"**, **"Ind_Prod.mat"** and **"condlogprob.m"**.

References

De Gooijer, J.G, Henter, G.E. and Yuan, A. (2022). Kernel-based Hidden Markov Conditional Densities. *Computational Statistics and Data Analysis*, 169. https://doi.org/10.1016/j.csda.2022.107431.

Henter, G.E., Leijon, A. and Kleijn, W.B. (2016). Kernel density estimation-based Markov models with hidden state. Available at https://arxiv.org/pdf/1807.11320.pdf.