Man-made flows from a fish's perspective: *in situ* flow classification of fishways using an artificial lateral line

Jeffrey A. Tuhtan[1*], Juan Francisco Fuentes-Perez[1], Gert Toming[1], Matthias Schneider[2], Richard Schwarzenberger[3], Martin Schletterer[4] and Maarja Kruusmaa[1]

[1]Centre for Biorobotics, Department of Computer Systems, Tallinn University of Technology

[2]SJE Ecohydraulic Engineering, GmbH

[3]ARGE Limnologie GesmbH

[4]Tiroler Wasserkraft AG

*Corresponding author: Email: jeffrey.tuhtan@ttu.ee, Phone: +49 176 724 16432

(Dated: March 18, 2017)

1. Introduction

The MATLAB script **HydroSigKMeans** ("m-file") and raw measurement data are provided from three vertical slot fishway (VSF) *in situ* measurements. The code and data generate hydrodynamic signatures and perform the k-means clustering of pool signatures as presented in our paper. The code requires that the user download freely available dependencies, and are discussed section 3 of this document.

This document supplies an outline of the m-files and data, and an example application. We use **bold text** for file names, and *italic text* for variables and commands.

2. Compilation

No compilation of the m-files is required for this release.

3. Script user variables

filePath – location of the locally saved raw data files for sequential processing and clustering.

tInt – sets the sample length in no. observations, default is 12000 (60 seconds).

pMax – the maximum value in kPa (0 – pMax) for the hydrodynamic signatures, default is 6 kPa.
pStep – number of increments used to visualize the hydrodynamic signatures, default is 6.

4. Dependencies

HydroSigKMeans requires the installation of two freely available m-files: **sort_nat** and **WindRose**, as well as the use of **kmeans** function which is part of the MATLAB Statistics and Machine Learning Toolbox.

- **sort_nat** – provides a natural order sort (1, 2, 3, … , . vs. 1, 11, 2, 22, 3, 33, … ,) which is needed for processing raw data files which are consecutively numbered. The m-file is provided by a third party and can be downloaded from the MATLAB File Exchange: https://www.mathworks.com/matlabcentral/fileexchange/10959-sort-nat--natural-order-sort
- **WindRose** – is a m-file which is used for plotting the hydrodynamic signatures of the 11 lateral line probe pressure sensors. It is also provided by a third party and can be downloaded from the MATLAB File Exchange: https://www.mathworks.com/matlabcentral/fileexchange/47248-wind-rose

5. Raw data

The raw data files are located in sub folders sorted first according to the VSF (Hirnbach, Runserau or Wenns), then by the type (Pool Cross-Sections or Slots). The numbering of the pool folders and slot measurements is based on the basin IDs which are shown in Figure 6 of the manuscript. The pool measurements are further organized based on their location, as they were measured from left to right when facing upstream into the flow. A diagram of all pool measurement locations and their orientation is also provided in the manuscript.

All measurement data has been directly converted from the ASCII text file as recorded in the field and converted to binary .MAT format to reduce storage requirements. The data have not been modified in any other manner, and are thus the raw lateral line probe sensor readings from the field measurements. The data format uses the structure type with the sensor name first, “ps1” for

pressure sensor number 1, and the data for each sensor are as follows: *a0.x, a0.y a0.z* and *a1.x*, *a1.y*, and *a1.z* are the linear accelerometers. Note that the accelerometers were not used in this manuscript, see [1] for a full technical overview of the lateral line probe. The pressure sensors outputs are *ps0.b, ps1.b, …, ps11.b* and are converted to pressure in kPa by the script before generating the hydrodynamic signatures.

6. Example application

To run **HydroSigKMeans**, it is first necessary to install all dependencies. This can be done by saving them to the same location as the **HydroSigKMeans** m-file. Alternately, the user can specify the location of the dependencies by pointing to them using the **addpath** function in MATLAB.

After the dependencies have been installed, the user needs to download a set of raw measurement data, for example the folder *…/Runserau/Pool_Cross-Sections/B02* which includes the full set of pool measurements in basin 2 of the Runserau VSF. In **HydroSigKMeans**, the user will then enter their local path for the variable *filePath* which is the location where the example data has been saved by the user. Finally, the m-file can be run either by typing in the command line *HydroSigKMeans* or alternately by selecting *Run* in the MATLAB Editor.

The user should see the consecutive list of file names displayed in the command line:
Importing file: 01-2016-04-22--10-58-49-401709749.mat ...
Importing file: 02-2016-04-22--11-01-18-093169352.mat ...
Importing file: 03-2016-04-22--11-07-56-466207000.mat ...

Upon successful completion of the m-file, the signatures are saved as a series of PNG image files to the source folder, and an elbow plot of the number of clusters using the max, min, mean and sum distances from each cluster centroid is displayed graphically:

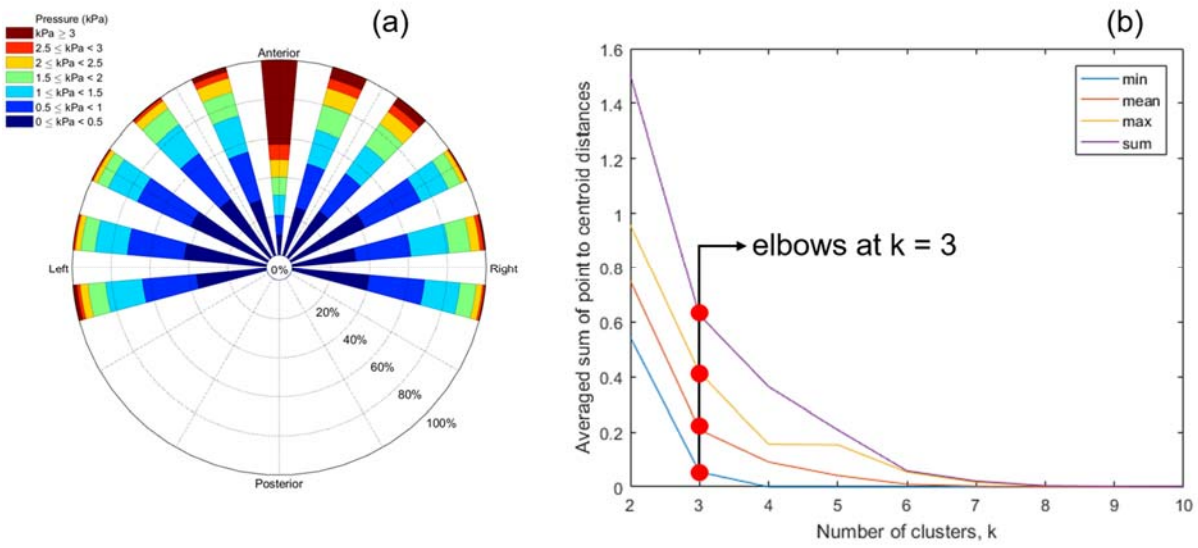

Figure 1: (a) Example of the exported PNG image showing the hydrodynamic signature for a single measurement location. (b) Plotted results of k-mean clustering using the elbow method, the example clearly shows the optimal value of k = 3.

7. References

1. Tuhtan J, Fuentes-Pérez J, Strokina N, Toming G, Musall M, Noack M, Kämäräinen J, Kruusmaa M. 2016 Design and application of a fish-shaped lateral line probe for flow measurement. *Rev. Sci. Instrum.* **87**, 045110.