Supplementary Material to "Directional Outlyingness for Multivariate Functional Data"

Wenlin Dai and Marc G. Genton[1]

This document serves as a supplement to the main text. It contains extra simulation studies, in which the covariance functions of the random processes have a short effective spatial range. Thus, the scaled RMD^2 is better approximated by the F-distribution with parameters calculated by the procedure described in Section 4. Besides, we also include one figure to show that the plot of $(\mathrm{MO}, \mathrm{VO})^T$ is robust to different smoothing methods applied to the raw data.

S1 Additional Numerical Studies

We consider four models with different types of outliers and different contamination levels $\epsilon = 0$ (uncontaminated), 0.1 and 0.2. These models all have a relatively short effective spatial range. For the new Models (S1-S4), we simply changed the covariance function of Models 1-4 as $r(s,t) = \exp\{-1000|t-s|\}$. The contamination styles remain the same. The results are shown in Table S1. Overall, the comparison results are similar with that drawn from the studies in the main paper: the proposed method using directional outlyingness outperforms

[1]CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia. E-mail: wenlin.dai@kaust.edu.sa, marc.genton@kaust.edu.sa
This research was supported by the King Abdullah University of Science and Technology (KAUST).

Table S1: Mean and standard deviation (in parentheses) of the percentage of correctly and falsely detected outliers over 500 simulation runs with four univariate models. Bold font is used to highlight the worst performance on both measurements for each setting. Int.Sqe: integrated squared error method; Rob.Mah: robust Mahalanobis distance; Out.Grm: adjusted outliergram; Dir.Out: method based-on directional outlyingness.

Method	Model S1		Model S2		Model S3		Model S4	
	p_c	p_f	p_c	p_f	p_c	p_f	p_c	p_f
$\epsilon = 0$								
Int.Sqe	-	**12.6** (3.1)	-	**12.6** (3.1)	-	**12.6** (3.1)	-	2.5 (1.5)
Rob.Mah	-	1.7 (1.6)	-	1.7 (1.6)	-	1.7 (1.6)	-	**2.7** (1.3)
Out.Grm	-	0.6 (0.9)	-	0.6 (0.9)	-	0.6 (0.9)	-	0.1 (0.2)
Dir.Out	-	0.3 (0.7)	-	0.3 (0.7)	-	0.3 (0.7)	-	0.0 (0.1)
$\epsilon = 0.1$								
Int.Sqe	**59.5** (24.0)	**11.5** (3.1)	100 (0.0)	**10.8** (2.9)	**24.4** (26.0)	0.1 (0.3)	93.0 (6.1)	**1.9** (1.5)
Rob.Mah	100 (0.0)	0.5 (0.9)	**41.7** (17.9)	0.8 (1.0)	100 (0.0)	**0.5** (0.8)	**74.7** (16.9)	1.4 (1.0)
Out.Grm	100 (0.0)	0.6 (1.0)	66.5 (17.2)	0.8 (1.2)	100 (0.0)	**0.5** (0.8)	90.4 (10.1)	1.4 (1.2)
Dir.Out	100 (0.0)	0.3 (0.6)	100 (0.4)	0.3 (0.6)	100 (0.0)	0.0 (0.1)	100 (0.0)	0.0 (0.0)
$\epsilon = 0.2$								
Int.Sqe	**45.3** (24.8)	**10.2** (3.3)	100 (0.0)	**8.1** (2.8)	**14.9** (21.5)	0.0 (0.2)	92.9 (2.9)	1.2 (1.3)
Rob.Mah	100 (0.0)	0.0 (0.2)	**25.1** (11.4)	0.2 (0.5)	100 (0.0)	0.1 (0.3)	**72.1** (11.8)	0.3 (0.6)
Out.Grm	100 (0.0)	1.7 (2.4)	52.6 (15.7)	0.8 (1.3)	98.1 (3.6)	**0.4** (0.7)	96.1 (5.4)	**4.0** (2.2)
Dir.Out	100 (0.0)	0.1 (0.4)	100 (0.2)	0.2 (0.4)	100 (0.0)	0.0 (0.1)	100 (0.0)	0.0 (0.1)

the other methods. Addtionally, the false detection rates are always lower due to the better approximation of the tail distribution of the scaled RMD^2.

S2 Robustness to Different Smoothing Methods

The plot of $(\text{MO}, \text{VO})^{\text{T}}$ is robust not only to the amount of smoothing but also to the selected smoothing methods. We provide one example in Figure S1, which shows that the plot remains similar with the Figure 5 in the main paper, which is based on B-spline basis functions, when we fit the curves with Fourier basis functions.

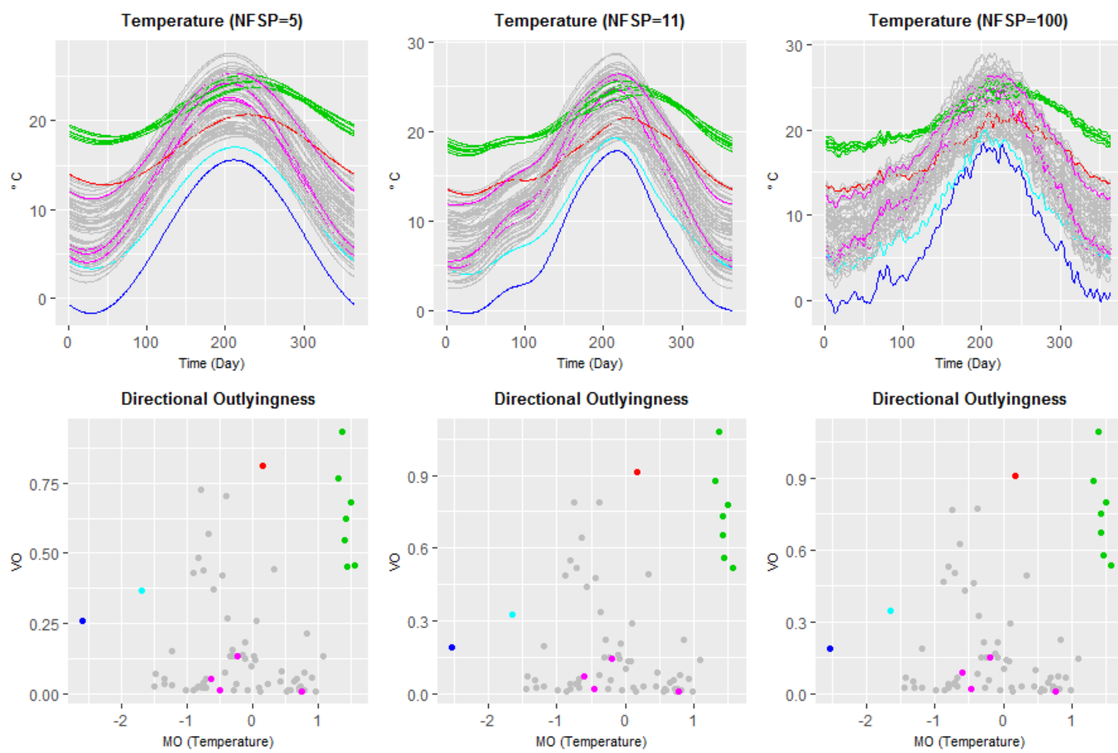

Figure S1: Plots of $(\mathrm{MO}, \mathrm{VO})^{\mathrm{T}}$ for curves fitted with different numbers of Fourier basis functions.