Efficient sampling of dependent data

Anne Eggels

October 11th, 2017

NWO Applied and Engineering Sciences

CWI

Problem setting (1/2)

We have a complex simulation with many inputs.

Input is given by data rather than distributions.

Uncertainty in the input leads to uncertainty in the output.

It is not feasible to evaluate the simulation output for all the input data. ($N = 10^5$, $K = 10^2$)

$\Longrightarrow$ well-known problem in uncertainty quantification (UQ)

CWI

Problem setting (2/2)

However:

The data is dependent; regular methods for UQ do not perform well.

$\Longrightarrow$ new methods are needed to select input samples.

(Small) example: sediment transport at the Noordwijk coast

CWI

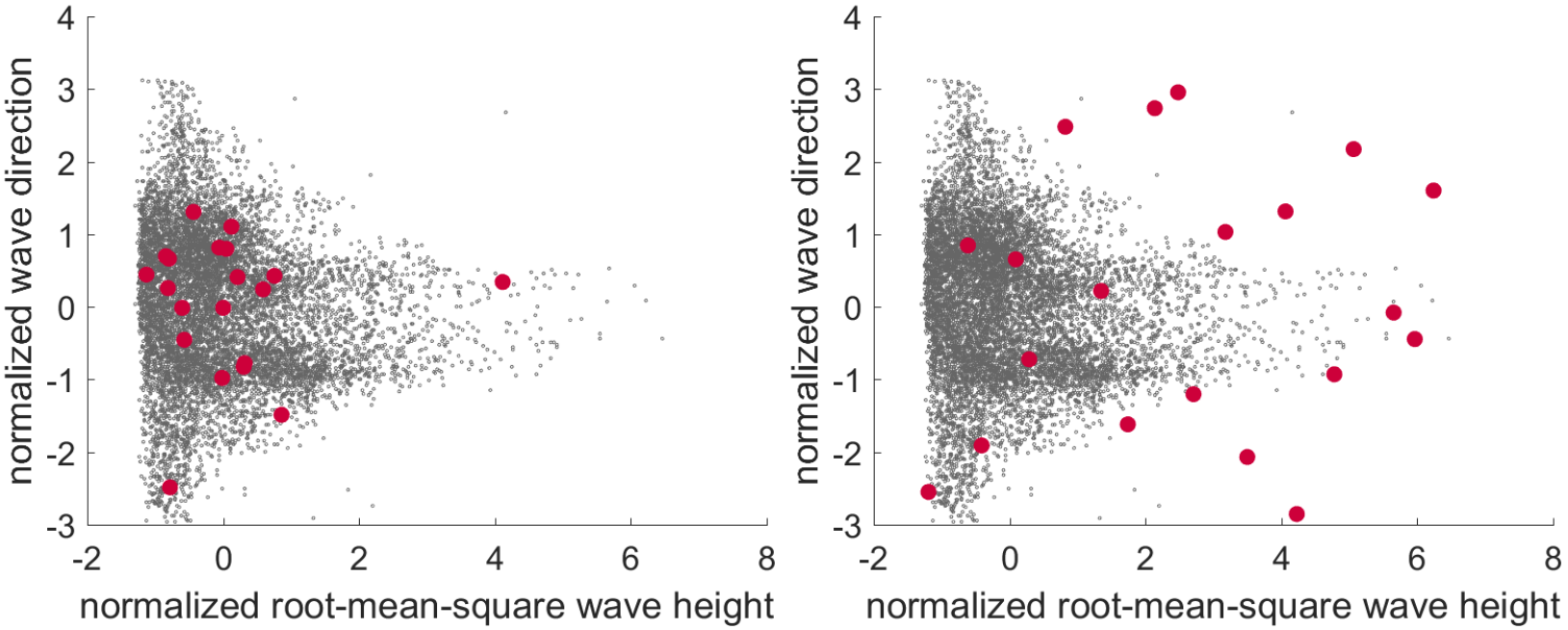
Sampling and dependent data

How to go from dataset to input samples with dependent data?

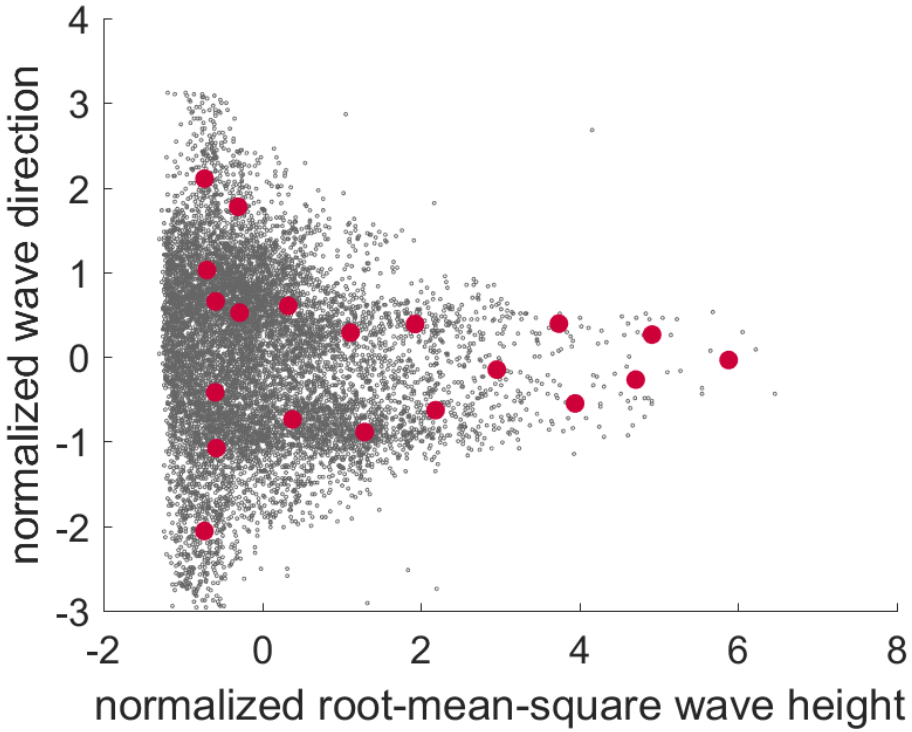
- Complete evaluation is not feasible
- Monte Carlo sampling is not desired
- Latin Hypercube sampling is not realistic

CWI

Examples


normalized wave direction
normalized root-mean-square wave height
4
3
2
1
0
-1
-2
-3
-2
0
2
4
6
8

Solution: clustering



A.W. Eggels, D.T. Crommelin & J.A.S. Witteveen, Clustering-based collocation for uncertainty propagation with multivariate dependent inputs,
International Journal for Uncertainty Quantification, to appear, 2017.

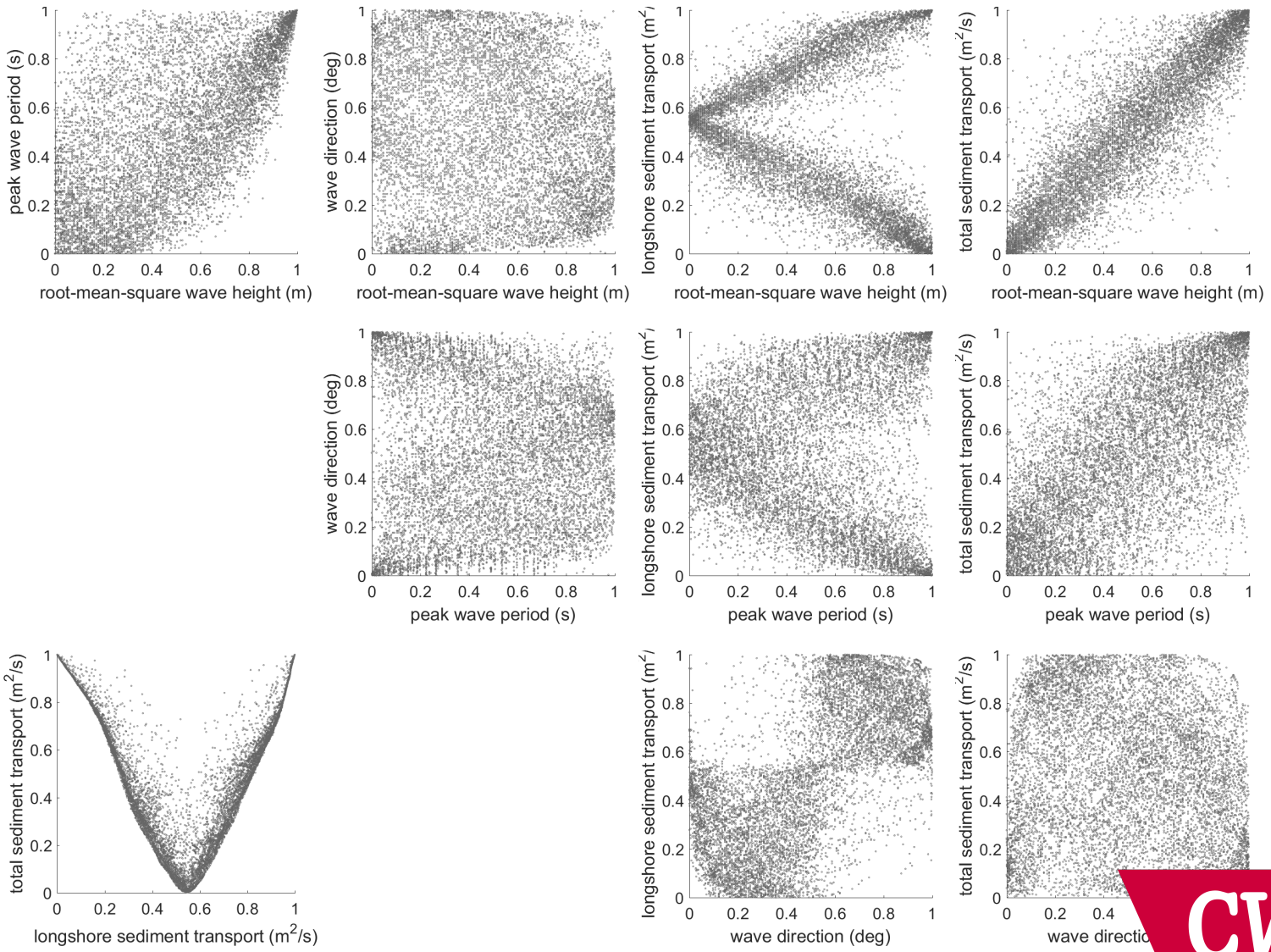
CWI

Benefits

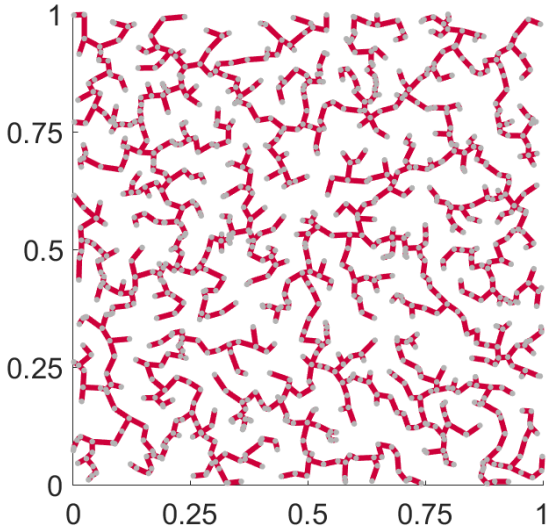
- Samples are well-distributed across the data
- No unrealistic samples
- Freedom in choosing the clustering method

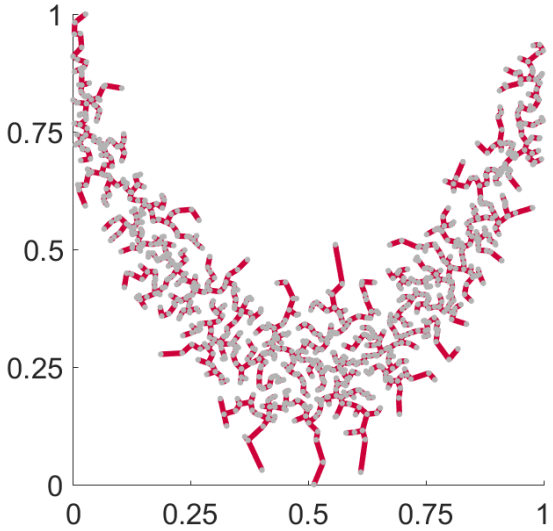
CWI

Can we measure the dependencies?



A new measure derived from entropy





	T_p	θ	S_y	S
H_{rms}	0.5598	0.6260	0.5024	0.4919
T_p		0.6113	0.5893	0.5883
θ			0.5486	0.6333
S_y				0.2847

CWI

A.W. Eggels & D.T. Crommelin, Quantifying dependencies in large datasets, *to be submitted, 2017.*

Conclusions

We developed a new method to reduce the number of samples needed in the case of dependent inputs.

We developed a new method to quantify the dependencies in a dataset.

Hence,

- dependencies do not need to be a problem;
- but they cannot be neglected.

CWI