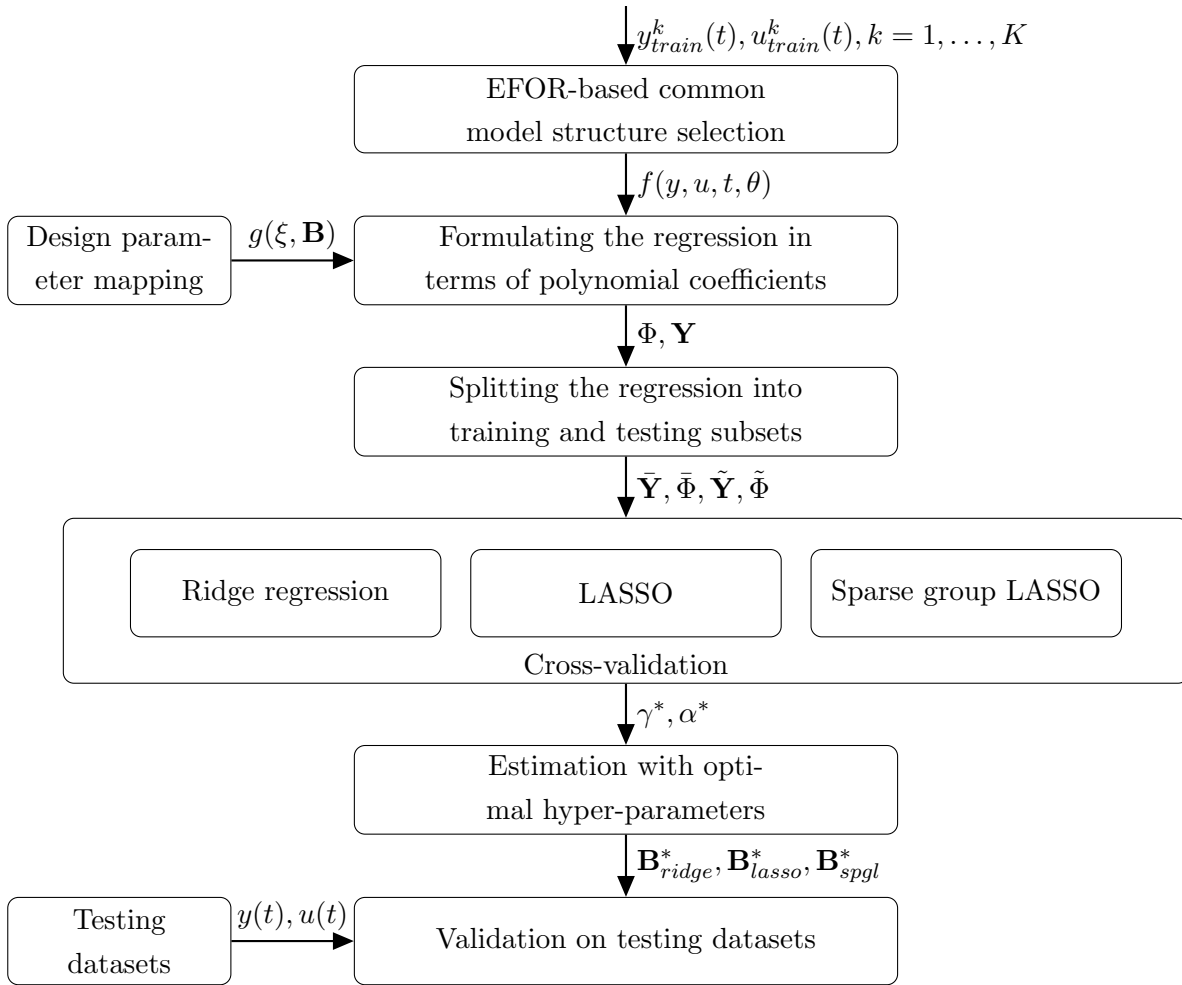# Identification framework overview

1 Framework structure

The schematic overview is of the implemented identification framework is presented in Figure 1, with further details on the constituent parts described in Table 1. The notation is consistent with the previous report.

Figure 1: A schematic overview of the identification framework.

In table 1, the regressor matrix Φ reflects both the common structure of the dynamical model identified using EFOR algorithm, f, and the arbitrary structure of the polynomial parameter mapping, g. The joint regression matrix and output vector are partitioned for the cross validation into the training data, $\bar{\Phi}, \bar{\mathbf{Y}}$, and the testing data, $\tilde{\Phi}, \tilde{\mathbf{Y}}$. The estimation algorithms are performed simultaneously with the same regularisation coefficient. Optimal regularisation coefficients are found using

Table 1: Constituent parts of the framework

Algorithm	Input	Output	Criterion
EFOR	Dictionary of regressors	Model structure	AERR - for determining significant terms BIC - for truncating number of terms
Tikhnonov	Training data, $\{\bar{\mathbf{Y}}, \bar{\Phi}\}$; Lagrangian, γ	Polynomial coefficients, $\hat{\mathbf{B}}^*(\gamma)$	$\left\{ \left\| \left(\bar{\mathbf{Y}} - \bar{\Phi}\mathbf{K}\bar{\mathbf{B}} \right) \right\|_2^2 + \gamma \left\| \bar{\mathbf{B}} \right\|_2^2 \right\}$
LASSO	Training data, $\{\bar{\mathbf{Y}}, \bar{\Phi}\}$; Lagrangian, γ	Polynomial coefficients, $\hat{\mathbf{B}}^*(\gamma)$	$\left\{ \left\| \left(\bar{\mathbf{Y}} - \bar{\Phi}\mathbf{K}\bar{\mathbf{B}} \right) \right\|_2^2 + \gamma \left\| \bar{\mathbf{B}} \right\|_1 \right\}$
Sparse group LASSO	Training data, $\{\bar{\mathbf{Y}}, \bar{\Phi}\}$; Lagrangians, γ, α	Polynomial coefficients, $\hat{\mathbf{B}}^*(\gamma, \alpha)$	$\left\{ \left\| \left(\bar{\mathbf{Y}} - \bar{\Phi}\mathbf{K}\bar{\mathbf{B}} \right) \right\|_2^2 + \right.$ $\left. \gamma \left((1-\alpha) \sum_i^N \sqrt{L_i} \|B_i\|_2^2 + \alpha \sum_i^N \|B_i\|_1 \right) \right\}$
CV	Testing data, $\{\tilde{\mathbf{Y}}, \tilde{\Phi}\}$; Parameter estimates, $\hat{\mathbf{B}}^*(\gamma, \alpha)$	Optimal regularisation parameters, γ^*, α^*	PRESS $= \sum_{l=1}^L \sum_t (\tilde{y}^l(t) - \hat{y}^l(t))$, where L is number of folds

random search algorithm with 200 iterations. The testing datasets for the final stage are not utilised in the model identification and differ from the testing subsets in the cross-validation algorithm.

Previously demonstrated results were obtained using K-folds validation that implicitly relies on OSA-validation as the data partition does not preserve the time-series structure in $\tilde{\Phi}, \tilde{\mathbf{Y}}$. The identified model structure thus corresponds to the minimal OSA error, while model-predicted output differs from the measured one(see Figure 2).

The framework has been amended so that the PRESS is obtained using model-predicted output

$$\hat{y}(t+1) = f(\hat{y}(t), \ldots, \hat{y}(t-n_y), u(t), \ldots, u(t-n_u), \theta(\mathbf{B})). \tag{1}$$

In order to accommodate MPO validation into the CV algorithm, the regression was partitioned so that the testing is performed against a dataset from one experiment at a time. See the blocks of the regression matrix in Figure 3.

2 Results file

The matlab file with results contains the following variables:

- Betas_nonreg_opt – estimate of polynomial coefficients obtained using OLS

- Betas_tikh_opt – estimate of polynomial coefficients obtained using ridge regression

- Betas_lasso_opt – estimate of polynomial coefficients obtained using LASSO algorithm

- Betas_spl_opt – estimate of polynomial coefficients obtained using sparse group LASSO regression

- Files – indices of files used for model training

- testFiles – indices of files used for model testing

- A – matrix of polynomial model terms for training files so that the following relationship holds

$$\Theta = BA^\top.$$

- A_valid – matrix of polynomial model terms for testing files

- A_symb – symbolic array of polynomial model terms that (structure of $g(\cdot)$)

- f_model – inline functions for lagged terms in the dynamical model $f(\cdot)$

- g_model – inline functions for polynomial terms in the parameter model $g(\cdot)$

- Terms – symbolic array with identified significant model terms (structure of $f(\cdot)$)

- extParams – design parameter vector indexed according to the tables 1 and 2 in the previous report
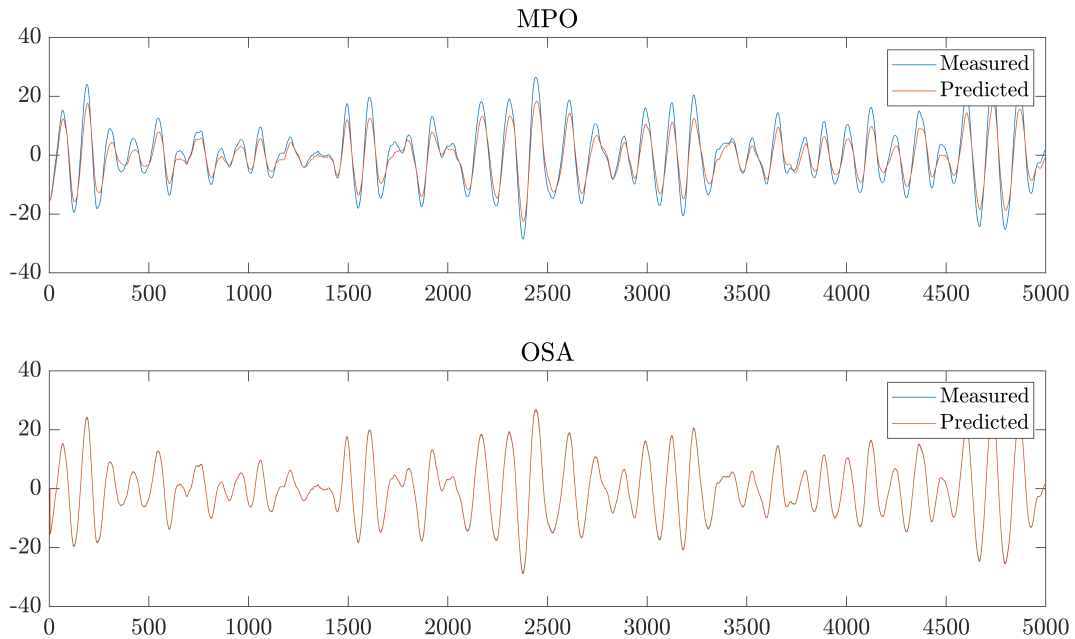
Figure 2: Model output from the model identified with OSA-PRESS optimal regularisation parameters for dataset HS4.

# 3 Validation results for modified framework

Figures below present validation results for experimental data obtained for soft auxetic foams. The test file indexing corresponds to the tables 2 in the previous report.
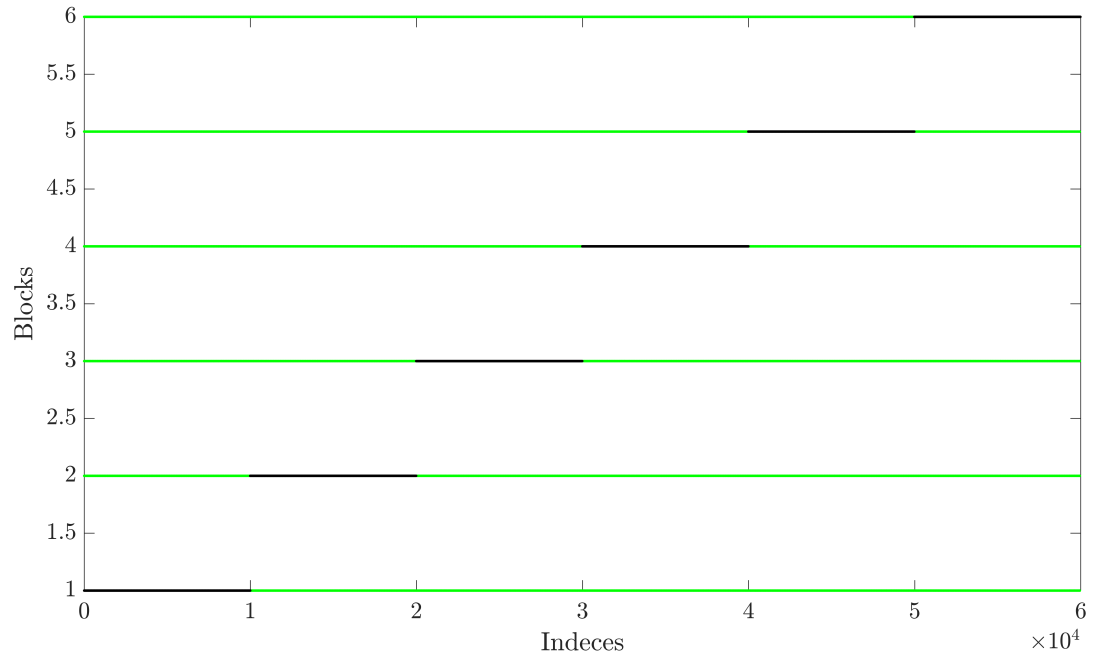
Figure 3: Block partitions of the regression data. Each testing dataset has time-series structure.
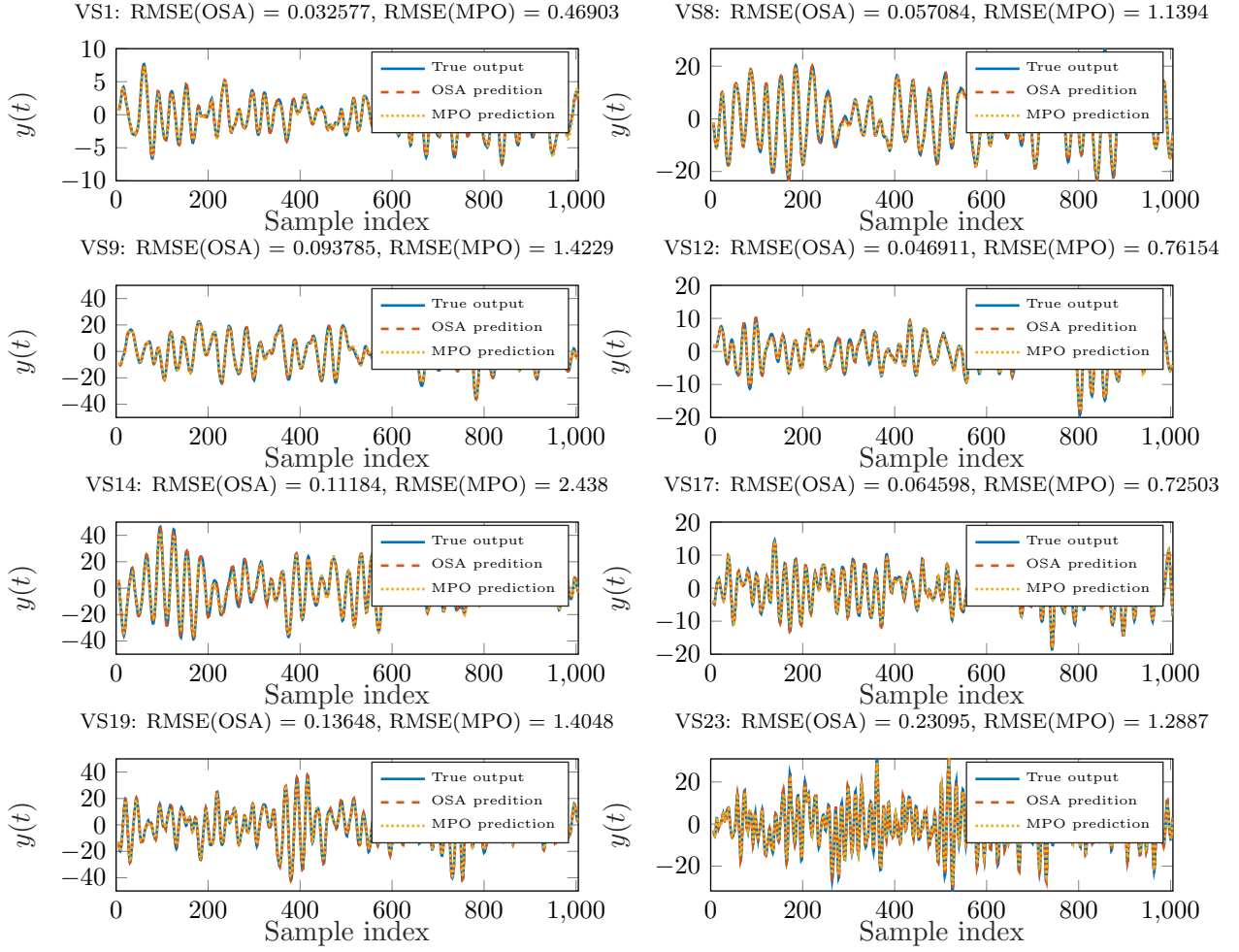
Figure 4: Validation results for the model estimated using OLS. Compression direction d3.
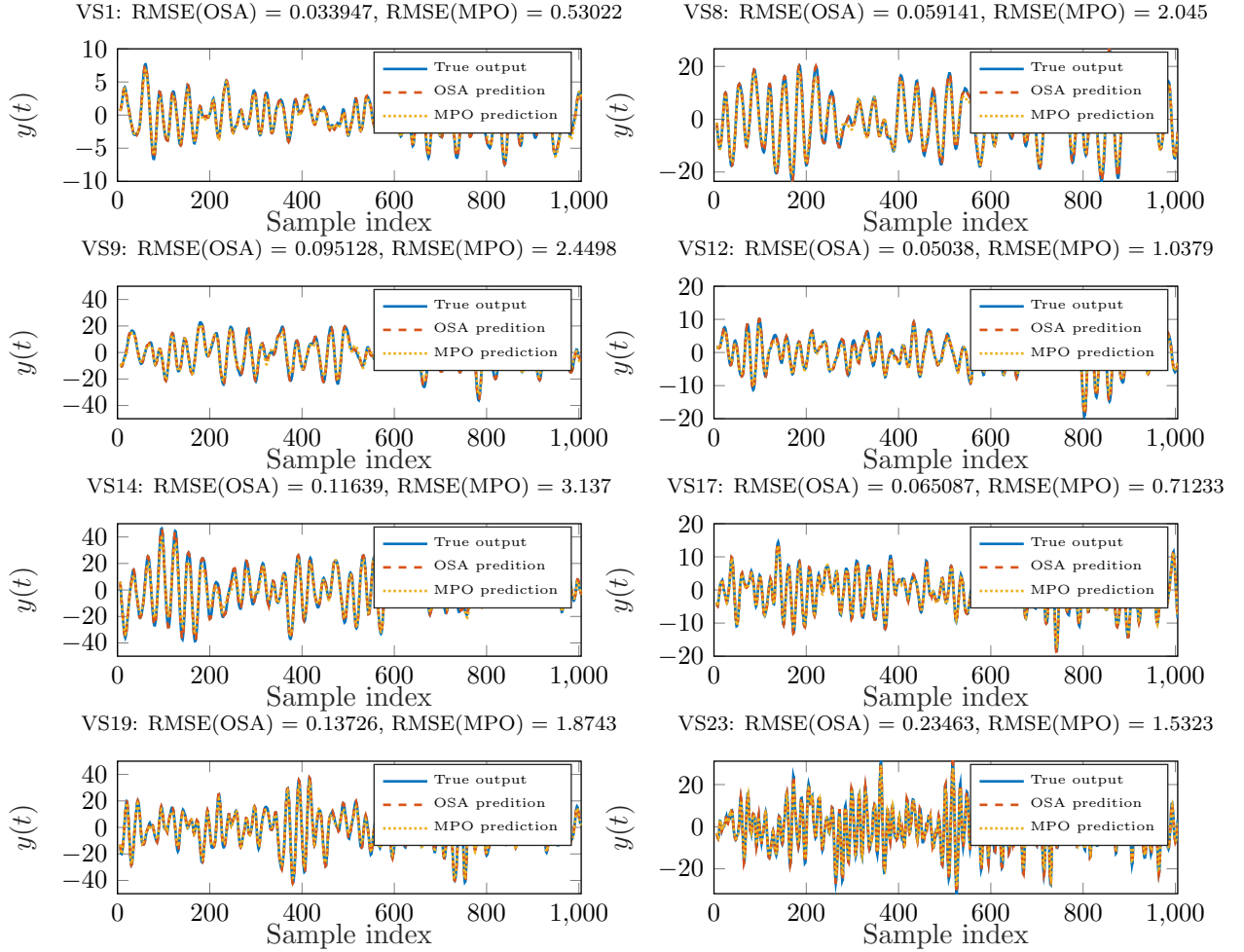
Figure 5: Validation results for the model estimated using LASSO. Compression direction d3.