UNIVERSITY OF CAMBRIDGE
Department of Engineering

Large Deformation Testing for Strain-Softening Parameter Calibration

D-SAS229-1

Author Name: Jonathan Smith
Supervisor: Dr Sam Stanier

Date: 27th May 2020

I hereby declare that, except where specifically indicated, the work submitted herein is my own original work.

*Signed:*____________________ *Date: 27th May 2020*

Contents

1 Technical Abstract

The strain-softening behaviour of fine-grained soils subject to large deformations has implications for many offshore geotechnical problems. The rapid deterioration of undrained strength, as a result of strain-softening, can cause instability and failure of geotechnical structures. Conventional element test based calibration techniques for strain-softening constitutive parameters are problematic; mainly the deformation is insufficient to reach a residual state (Wood, 2012). Additionally, tests often ignore localised deformation and failure (Gylland et al., 2014). Consequently, calibration becomes subjective and unconvincing.

Recent developments in full-field displacement measurement techniques (Stanier et al., 2016) and energy-based calibration methods (Grédiac & Pierron, 1998; Gueguin et al., 2015) provide a promising alternative to conventional approaches. In this vein, Singh and Stanier (unpublished) have developed and numerically validated a calibration technique for strain-softening constitutive parameters. Parameter optimisation is achieved by minimising the difference between the incremental internal and external work.

In the first part of the project, a Bayesian inferential approach, Markov Chain Monte Carlo (MCMC) methodology (with Metropolis-Hastings sampling algorithm (Hastings, 1970)) was incorporated into the proposed calibration framework. Input datasets consisted of two numerically-generated biaxial compression tests. Through MCMC, sample chains were generated and subsequently processed to generate parameter-likelihood distributions (Figure 1). This process was repeated, varying the magnitude of additive noise, replicating degrees of experimental error, and varying the degree of observed "remoulding", allowing comparison of proposed and conventional testing (left to right and top to bottom in Figure 1, respectively). The results of the analysis clearly show that the greater the precision in testing and the greater the proportion of degradation observed, the greater the calibration certainty. However, biaxial testing remains sub-optimal due to the modest amount of deformation mobilised; full remoulding to a residual state would improve the result. Additionally, a limit on acceptable error for the proposed test was identified.

In the second part of the project, a laboratory-based calibration test was designed and prepared. Due to COVID-19 related restrictions, it was not possible to conduct the planned experiment. The test (configuration shown in Figure 2) would have involved the cyclic displacement of a T-bar full-flow penetrometer in undrained consolidated Kaolin clay for full "remoulding" of the soil. The horizontal bar of the penetrometer, perpendicular to the exposed plane, extends the breadth of the test box, ensuring plane-strain conditions throughout. Thus, the exposed plane displacement field is representative of the full mechanism allowing the use of Particle Image Velocimetry techniques (GeoPIV-RG; Stanier et al., 2016). Incremental external work is calculated from measurements of penetration force and T-bar displacement.

Despite the inability to complete the experiment, this project made progress in its preparation and provided support for the proposed framework through other means. The functionality of the proposed calibration framework has been enhanced through the inherent incorporation of uncertainty in parameter definitions. Characterisation of uncertainty has also enabled a comparison of calibration quality. The reduction in uncertainty proved that accurate and representative testing improves calibration. Further work includes the completion of the experiment and the objective comparison of constitutive models.

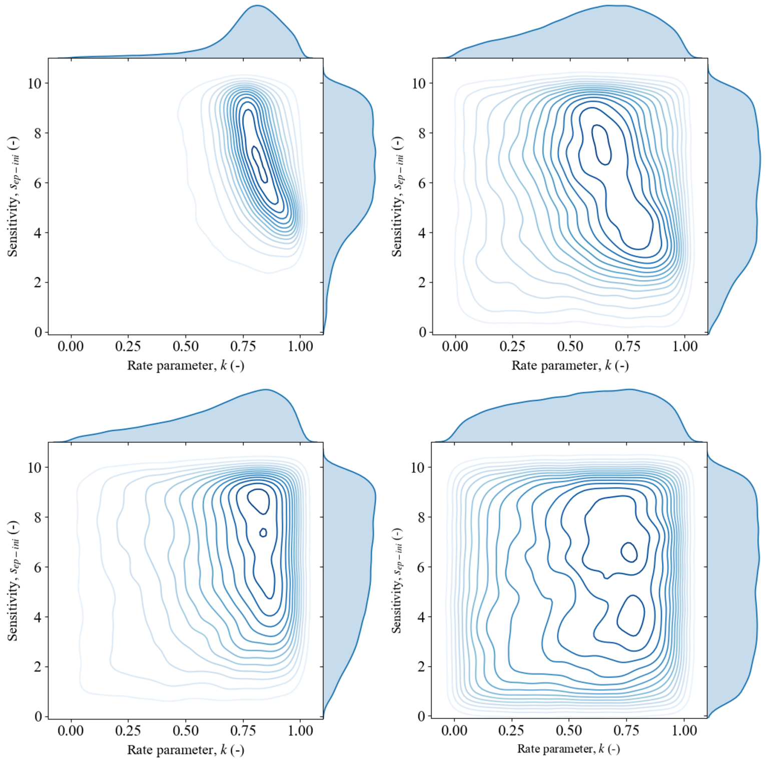

Figure 1: Parameter-likelihood distributions: full-frame dataset with negligible additive noise (top left), full-frame dataset with non-negligible additive noise (top right), half-frame dataset with negligible additive noise (bottom left) and half-frame dataset with non-neglible additive noise.

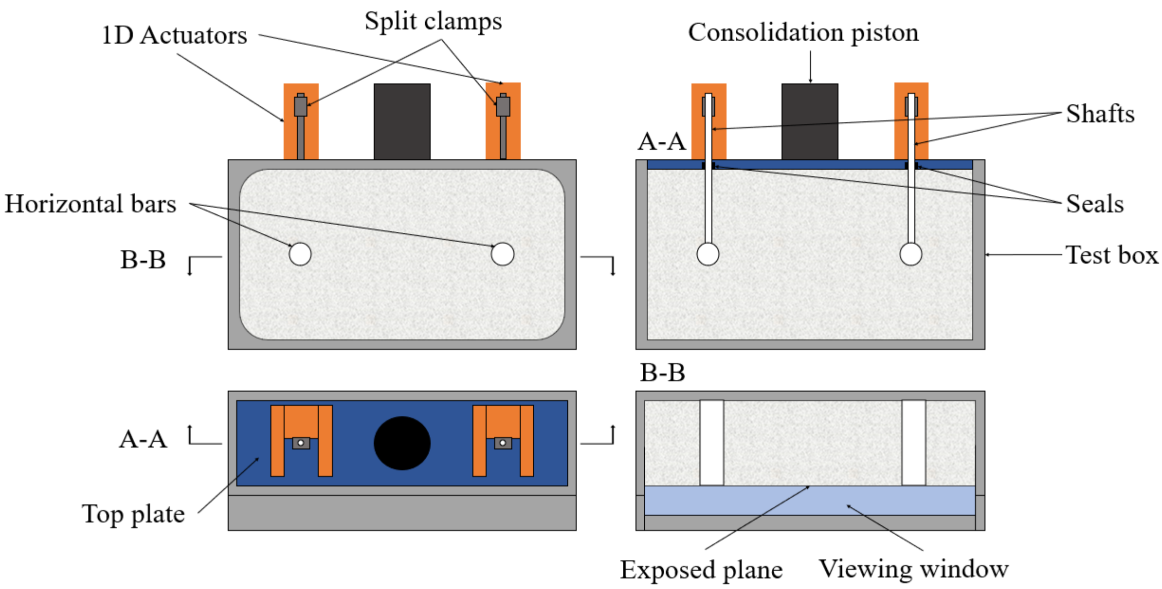

Figure 2: The test-rig configuration: front view (top left), plan view (bottom left), mid-box front view (top right) and mid-box plan view (bottom right).

2 Introduction

The strain-softening behaviour of fine-grained soils subject to large deformations has implications for many offshore geotechnical problems. Undrained shearing of normally consolidated fine-grained soils generates excess pore pressure, resulting in the rapid deterioration of undrained shear strength in a "remoulding" process. This process has been accounted for in a variety of constitutive models. However, the accuracy of numerical modelling is dependent upon appropriate calibration of the constitutive parameters.

Several issues exist with the suitability of conventional calibration methods (predominantly standard element tests) in deriving strain-softening parameters. Firstly, insufficient strains are achieved for a residual state to be observed (Wood, 2012). Secondly, tests record an average strain measured at the boundaries, implicitly assuming homogeneous element response. This is non-representative as it ignores localised deformation in shear band formation and the impact upon associated stress and strain paths (Gylland et al., 2014). Thirdly, subjectivity in determining model fit is introduced as a result of difficulty in calibration, as exemplified in the hypoplastic model simulated response (Figure 3; Ragni et al., 2015); a large range of strain-softening rate parameter values, k, could be justified for the current experimental data and constitutive model choice. None are particularly compelling.

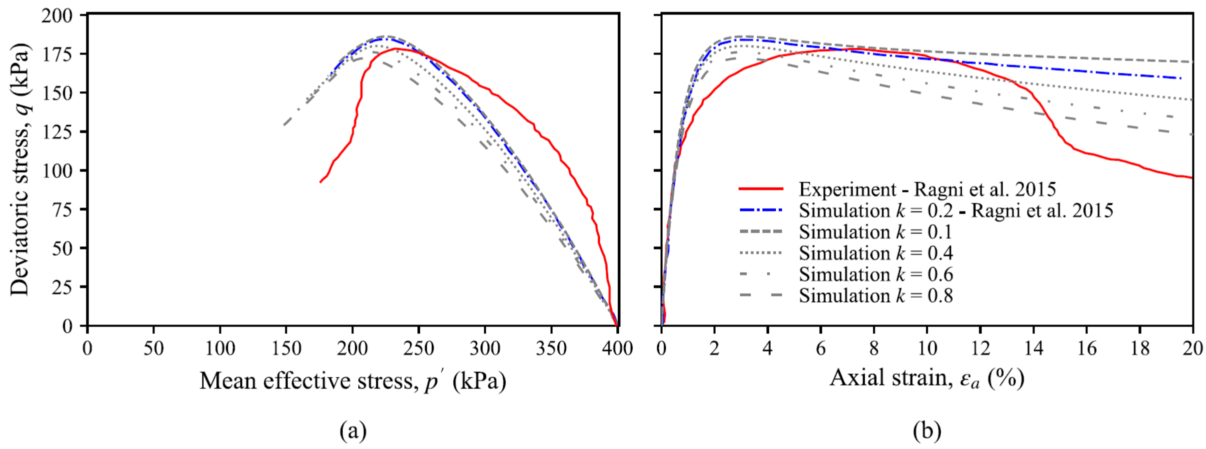

Figure 3: Triaxial test data for strain-softening parameter calibration illustrating fit difficulty from Ragni et al. (2015).

Recent developments in full-field displacement measurement techniques, including Direct Image Correlation (DIC) and Particle Image Velocimetry (PIV) (Stanier et al., 2016), provide alternative means of calibration. Full-field deformations are recorded in a sequence of images of an exposed plane of a sample during testing (thereby incorporating inhomogeneous strains). The principles of the conservation of energy and virtual work are applied within the Virtual Fields Method (Grédiac & Pierron, 1998). Recently, this method was employed for constitutive model parameter calibration in a simple technique proposed by Gueguin et al. (2015) and later applied by Charles et al. (2018).

Building on this foundation, Singh and Stanier (unpublished) have created a calibration technique targeting the strain-softening parameters. This enhancement implements full stress-strain integration algorithms (e.g. Sloan et al., 2001) to ensure a path-dependent response using more realistic constitutive models than Charles et al. (2018). In full form, the proposed framework uses full-field, image-based deformation measurements and PIV techniques before applying the Virtual Fields Method for energy-based parameter optimisation. The technique has proved successful in numerical validation (Singh and Stanier, pers. comm.).

Firstly, this project statistically justifies the potential improvement that could result from using full-field displacement measurement techniques for calibration. A Bayesian inferential methodology (see Bernardo and Smith, (1994) and Rogers and Girolami (2017) for comprehensive introductions) establishes strain-softening parameters as likelihood distributions based on numerically-generated datasets. In defining parameters as random variables, uncertainty is inherently incorporated. Comparison is made between calibration techniques employing uncertainty as a metric.

Secondly, this project develops the full-field, image-based calibration test, establishing the usefulness of the framework for practical application in industry. The experiment consists of cyclic T-bar tests performed in undrained Kaolin clay. The horizontal bar, perpendicular to the exposed plane, extends the breadth of the test box, allowing plane-strain conditions to be assumed. Therefore, the exposed plane is representative of full-field deformation. A Bayesian inferential methodology will be applied to the dataset produced to quantify uncertainty and select calibrated parameters for input into a numerical simulation replicating the experiment, as a means of validation.

The application of this project, in support of the larger framework established by Singh and Stanier (unpublished), is to facilitate accurate and representative parametric calibration from true boundary value problems. A Bayesian approach has enabled the characterisation of uncertainty in estimation, extending the framework functionality and usefulness. Further work could include the objective comparison of existing constitutive models, as well as the stochastic derivation of new models.

2.1 Constitutive Model

The strain-softening constitutive model to be calibrated consists of Structured Modified Cam Clay (SMCC) (Mašín, 2009) with the implementation of the Butterfield Compression Law (Butterfield, 1979). This elasto-plastic model requires eight material parameters; five basic parameters ($M, \lambda^*, \kappa^*, N$ and G or v) and three strain-softening parameters (k, A and s). The shape of initial yield surface, f, is defined by (1):

$$f = q^2 + M^2 p'(p' - s_{ep} p_c) \tag{1}$$

where q is the deviatoric stress, p' is the mean effective stress, p_c is the preconsolidation pressure and M is the critical state line (CSL) slope. The sensitivity state variable, s_{ep}, is measured along the elastic unload-reload line. Thus, the sensitivity state variable acts as a scalar multiplier on the size of the yield surface (1), representing the ratio of yield surface size between the intact and remoulded states. This is, therefore, defined in terms of the gradients of the normal consolidation line (NCL), λ^*, and unload-reload line (URL), κ^*, (2)(Figure 4):

$$s_{ep} = s_i^{\frac{\lambda^*}{\lambda^* - \kappa^*}} \tag{2}$$

where s_i is the initial sensitivity. Stress states on the yield surface strain plastically, degrading soil sensitivity (3):

$$\dot{s}_{ep} = -\frac{k}{(\lambda^* - \kappa^*)}(s_{ep} - s_f)\epsilon_s \tag{3}$$

where k controls the rate of sensitivity degradation, s_f is the reference sensitivity (typically unity) and ϵ_s is the equivalent plastic strain. Equivalent plastic strain is defined by (4) (modified from Mašín, 2009):

$$\epsilon_s = \sqrt{(1 - A)(\dot{\epsilon}_v)^2 + A(\dot{\epsilon}_q)^2} \tag{4}$$

where $\dot{\epsilon}_v$ and $\dot{\epsilon}_q$ are the plastic volumetric and deviatoric strain rates respectively. The parameter A controls the relative contribution of the two strain rates to sensitivity degradation. The remaining strain-softening parameters, k and s_{ep} are calibrated in the proposed framework.

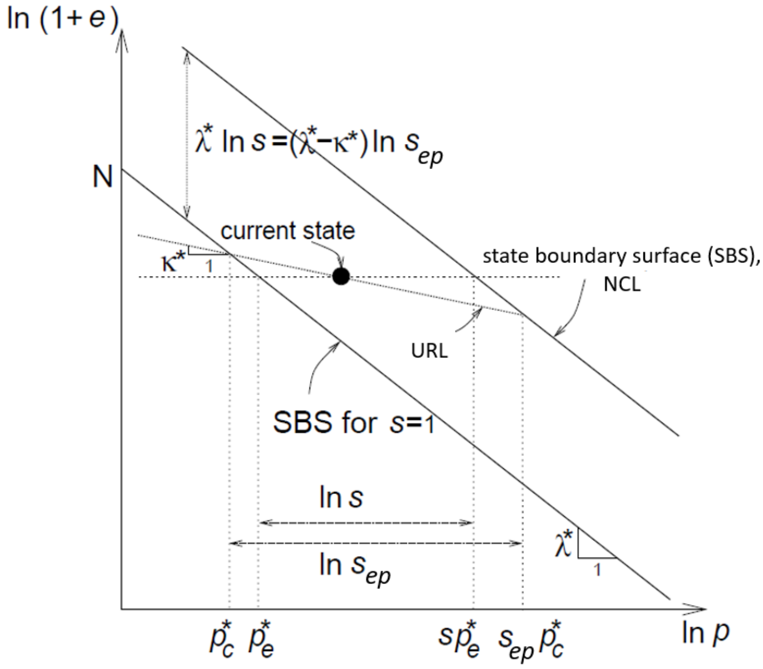

Figure 4: Definitions of sensitivities s and s_{ep}, quantities p_c^* and p_e^* and material parameters N, λ^* and κ^*. Diagram adapted from Mašín (2009).

2.2 Numerical Modelling Considerations

In large deformation strain-softening behaviour, strain fields tend to be inhomogeneous with deformations localising into thin shear bands (Georgiannou and Burland, 2006). In finite element (FE) modelling of this phenomenon, a non-local model must be implemented to avoid energy dissipation and localisation into a zone of vanishing volume (Bazant and Pijaudier-Cabot, 1988) and the development of mesh dependency (De Borst et al., 1993). The true width of a shear band in geomaterials is thought to be determined by an internal material length scale governed by microstructure (Desrues and Viggiani, 2004; Mánica et al., 2018).

Here, an integral-type, non-local model (Bažant et al., 1994) is employed. The strain-softening at any particular integration point is a function not only of its own plastic deformation, but also that of neighbouring points through a spatially-averaged weighting function (5) replacing (4):

$$\bar{\epsilon}_s(r_i) = \int_V \alpha(r_i, r_j)\epsilon_s(r_j)dV \tag{5}$$

where i and j index the current point and neighbouring point respectively, r is the point position, ϵ_s is the local equivalent plastic strain and dV is the volume represented at point i. α is a weighting function, applied in normalised form to preserve a uniform, non-local strain distribution (6). The weighting function, α_0, is dependent upon the distance between integration points i and j according to the function proposed by Galavi and Schweiger (2010) (7). Under this weighting function, the contribution of the current integration point is ignored.

$$\alpha(r_i, r_j) = \frac{\alpha_0(||r_i - r_j||)}{\int_V \alpha_0(||r_i - r_j||)dV} \tag{6}$$

$$\alpha_0 = \frac{||r_i - r_j||}{l_c} \exp\left[-\left(\frac{||r_i - r_j||}{l_c}\right)^2\right] \tag{7}$$

Mesh dependency is removed at the expense of an additional parameter for calibration, l_c, the characteristic length. This project observes the condition established by Galavi and Schweiger (2010) as expressed in (8), where L_{max} is the maximum element length.

$$l_c \geq L_{max} \tag{8}$$

2.3 Calibration Framework

Singh and Stanier's (unpublished) energy-based framework (Figure 5) has the potential to facilitate calibration from a single dataset, greatly simplifying the typical calibration process. The objective function (9), $f(x)$, is defined as the sum of the least squares of the difference between the incremental internal and external work (10), d^i, where i is the increment (image number or time step):

$$f(x) = \sqrt{\sum_{i=1}^{n}(d^i)^2} \tag{9}$$

$$W_I^i(\sigma, \epsilon) - W_E^i = d^i \tag{10}$$

$$W_E^i = F \cdot \delta \tag{11}$$

$$W_I^i = \int_V \sigma \Delta\epsilon dV \tag{12}$$

The incremental external work consists of work done on the soil by the displacing force (11), calculated from force-displacement ($F - \delta$) curves, measured experimentally using an Linear Variable Differential Transformer (LVDT) and a load cell, for example. In the evaluation of incremental internal work, input strain-softening parameters are varied in the constitutive model relating the observed strain field (calculated from full-field deformation measurements) to the derived stress field. The incremental internal work is calculated by integrating the stress, σ, and strain field over the volume (12). In perfect energy conservation, $d^i = 0$ for all i. Allowing for error, optimisation is achieved through minimising the objective function (9). The calibration framework has been scripted in Fortran by Singh (pers. comm.) and is used in this project.

Advancing the proposal, this project implements a Bayesian inferential methodology for strain-softening parameter variation in the evaluation of incremental internal work. This allows for consideration of uncertainty in parameter estimation, as will be detailed in Section 3.

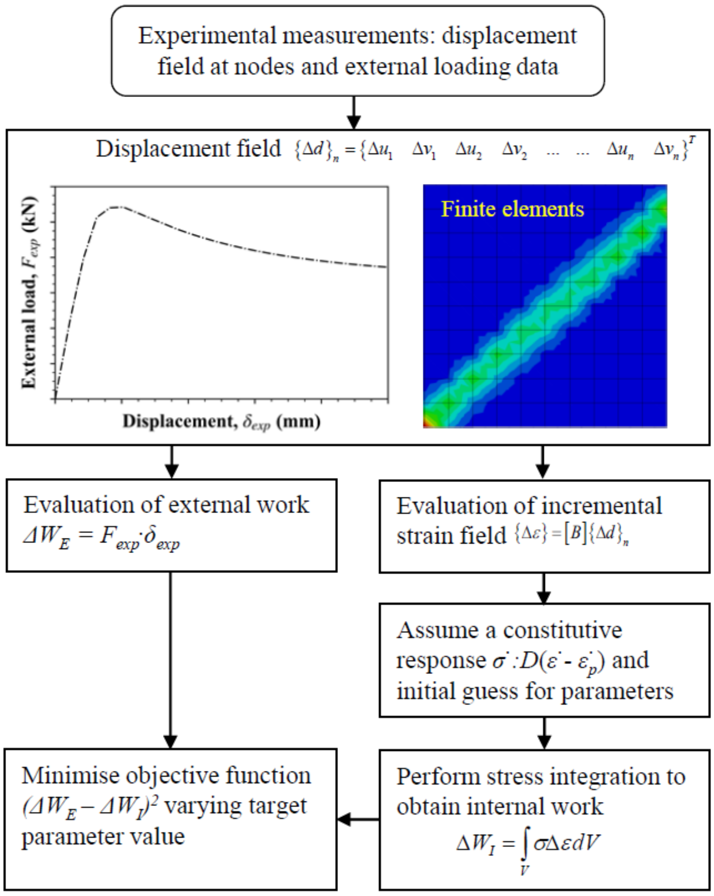

Figure 5: Optimisation strategy (shown for a numerically-generated dataset). Figure reproduced from Singh and Stanier (unpublished).

3 Bayesian Inference

3.1 Introduction

The application of Bayesian inference, as a coherent framework with inherent uncertainty characterisation, provides a statistical foundation to retrospectively justify the alternative calibration method proposed (Singh and Stanier, unpublished). Numerically-generated datasets were used to produce input parameter probability distributions. Subsequently, input parameter uncertainty was related to dataset error, validating an image-based approach; and to the proportion of observed strength degradation, validating large deformation testing.

3.2 Theory

3.2.1 General

Consider a general, simplified model (see Rogers and Girolami, 2017):

$$\mathbf{y} = m(\mathbf{x}; \boldsymbol{\theta}) + \boldsymbol{\epsilon} \tag{13}$$

where m is the functional form of the constitutive model, M, relating the input dataset, $\mathbf{x}$ with model parameter set, $\boldsymbol{\theta}$, to the output dataset, $\mathbf{y}$. In the proposed calibration framework, $\mathbf{x}$ describes the strain-field measurements and $\boldsymbol{\theta} = \{k, s_{ep}\}$, the strain-softening parameter set (with all other parameters defined via conventional means). The model functions, m, also derive the stress field and perform stress integration (12), according to the constitutive equations, producing the dataset of incremental internal work. In parallel, $\mathbf{y}$ is the dataset of incremental external work with values defined by (11). The observed data in testing is summarised by $\mathcal{D} = \{\mathbf{y}, \mathbf{x}\}$.

As a simplification, it is assumed that all error generated experimentally can be represented by an additive Gaussian noise, $\boldsymbol{\epsilon} \sim \mathcal{N}(0, v)$, on the model output. Therefore, the random variable y_i is described by the density function (14).

$$p(y_i|\mathbf{x}_i, \boldsymbol{\theta}, v, M) = \mathcal{N}(y_i; m_i(\mathbf{x}; \boldsymbol{\theta}), v) \tag{14}$$

Additionally, it is assumed that error is spatially and temporally independent. The validity of these assumptions is reliant on the selection of a representative variance, v, and the absence of growing systematic error in testing.

The posterior parameter distribution is given by Bayes' Rule:

$$P(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X}, v, M) = \frac{p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}, v, M)p(\boldsymbol{\theta}|M)}{p(\mathbf{y}|\mathbf{X}, v, M)} \tag{15}$$

As the strain-softening parameters are considered to be independent, the prior distribution (existing beliefs before data is seen), $p(\boldsymbol{\theta}|M)$, can be expressed as the product of individual parameter priors (16). A relatively uninformative prior is used; a uniform density over a finite space defined by fixed limits, k_{lim} and $s_{ep_{lim}}$ (17).

$$p(\boldsymbol{\theta}|M) = p(k)p(s_{ep}) \tag{16}$$

$$p(k) = \mathcal{U}_k(0, k_{lim}), \quad p(s_{ep}) = \mathcal{U}_{s_{ep}}(0, s_{ep_{lim}}) \tag{17}$$

The likelihood, $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}, v, M)$, is the joint conditional density of its constituent data points (18). For the purposes of calibration in the framework proposed, it is more precise numerically to compute the natural logarithm of the likelihood (19).

$$L = p(y_1, ..., y_N|\mathbf{x}_1, ..., \mathbf{x}_N, \boldsymbol{\theta}, v, M) = p(\mathbf{y}|\mathbf{X}, \boldsymbol{\theta}, v, M) = \prod_{i=1}^{N} p(y_i|\mathbf{x}_i, \boldsymbol{\theta}, v, M) \tag{18}$$

$$\begin{aligned} \ln L &= \sum_{i=1}^{N} ln\left(\frac{1}{\sqrt{2\pi v}} \exp\left[-\frac{1}{2v}(y_i - m(\mathbf{x}_i; \boldsymbol{\theta}))^2\right]\right) \\ &= -\frac{N}{2}\ln(2\pi v) - \frac{N}{2v^2}\sum_{i=1}^{N}(y_i - m(\mathbf{x}_i; \boldsymbol{\theta}))^2 \end{aligned} \tag{19}$$

The marginal likelihood, $p(\mathbf{y}|\mathbf{X}, v, M)$, is the likelihood of the data averaged across all parameter values, acting to normalise the product of likelihood and prior. Thus, it can be expanded:

$$p(\mathbf{y}|\mathbf{X}, v, M) = \int p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}, v, M)p(\boldsymbol{\theta}|M)d\boldsymbol{\theta} \tag{20}$$

The multi-dimensional integration involved in evaluating the terms of (15) has a general form equivalent to the expectation of the function, $f(\mathbf{z})$, with respect to a prior density, $\pi(\mathbf{z})$:

$$\int f(\mathbf{z})\pi(\mathbf{z})d\mathbf{z} \equiv E_{\pi}\{f(\mathbf{z})\} \tag{21}$$

$$E_{\pi}\{f(\mathbf{z})\} \approx \frac{1}{S}\sum_{s=1}^{S} f(\boldsymbol{\theta_s}) \tag{22}$$

where S is the number of samples taken.

As the constitutive model consists of complex non-linear ordinary differential equations (ODEs), solving ceases to be analytical and presents significant numerical computational difficulty. If it is necessary to calculate the posterior density, approximation methods of varying complexity exist. For example, a Monte Carlo approximation can be made with sufficient sampling of the function's parameter-space (22).

However, several problems arise concerning sampling from a density that lacks analytical form. Regions of high-likelihood in the parameter-space can be small, requiring a high magnitude of sampling for accurate estimation. Difficulty is exacerbated by increased dimensionality as the relative size of the region tends to decrease (Vyshemirsky and Girolami, 2007). For example, a simple Monte Carlo-based approach, the Prior Arithmetic Mean estimator (McCulloch and Rossi, 1991), required roughly 50 million draws from the prior distribution to achieve an acceptable error level in a study by Lewis and Raftery (1997). Furthermore, a high proportion of samples are drawn from low-likelihood regions, introducing large inefficiency and skewing the generated expectation, thus performing poorly as an estimator (Kass and Raftery, 1995).

3.2.2 Markov Chain Monte Carlo

The Markov Chain Monte Carlo (MCMC) methodology, implemented with the Metropolis-Hastings (MH) sampling algorithm (Hastings, 1970), is used as an efficient and effective approach for the calibration framework. An MCMC approach provides a means of sampling a continuous random variable probability density. Each proposed transition of state is dependent solely on the existing state, selected from a distribution that is centred on the current sample (Girolami, 2008). The sampling algorithm implemented sets acceptance criteria for each proposal in the generation of the chain.

Sampling from the unknown target distribution is achieved through a recognition of proportionality to the non-normalised known product of likelihood and prior (23):

$$P(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X}, v, M) \propto p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}, v, M)p(\boldsymbol{\theta}|M) \tag{23}$$

The transition density, $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$, where '*' denotes the proposed parameter set, defines the probability distribution within which local Markov transitions occur. Here, a Gaussian transition density (24), with a diagonal covariance matrix, $\mathbf{C}$ (25), is assumed for ease of sampling and for symmetry (26). This will prove a useful property in the implementation of the MH algorithm. The assumption will be validated by the similarity in distribution shape between the transition density and the posterior (Rogers and Girolami, 2017).

$$q(\boldsymbol{\theta}^*|\boldsymbol{\theta}) = \mathcal{N}_{\boldsymbol{\theta}^*}(\boldsymbol{\theta}, \mathbf{C}) \tag{24}$$

$$\mathbf{C} = \begin{bmatrix} v_k & 0 \\ 0 & v_{s_{ep}} \end{bmatrix} \tag{25}$$

$$q(\boldsymbol{\theta}^*|\boldsymbol{\theta}) = q(\boldsymbol{\theta}|\boldsymbol{\theta}^*) \tag{26}$$

The acceptance criterion (27) is dependent upon the ratio of transition densities and the ratio of posterior densities, H_r (28), thereby cancelling the normalising terms and allowing substitution by the proportional likelihood-prior products. If $\boldsymbol{\theta}^*$ has greater likelihood than $\boldsymbol{\theta}$, $H_r > 1$ and the proposal will be accepted. If $\boldsymbol{\theta}$ has greater likelihood than $\boldsymbol{\theta}^*$, $H_r < 1$ and the proposal will be accepted with probability H_r. In this case a random variable, u, is taken from the uniform distribution, $\mathcal{U}(0,1)$, and the proposal is accepted if $u \leq H_r$. Figure 6 displays steps of the MH algorithm diagrammatically.

$$\min\{1, H_r\} \tag{27}$$

$$H_r = \frac{p(\mathbf{y}|\boldsymbol{\theta}^*, \mathbf{X}, v, M)p(\boldsymbol{\theta}^*|M)}{p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}, v, M)p(\boldsymbol{\theta}|M)} \frac{q(\boldsymbol{\theta}^*|\boldsymbol{\theta})}{q(\boldsymbol{\theta}|\boldsymbol{\theta}^*)} \tag{28}$$

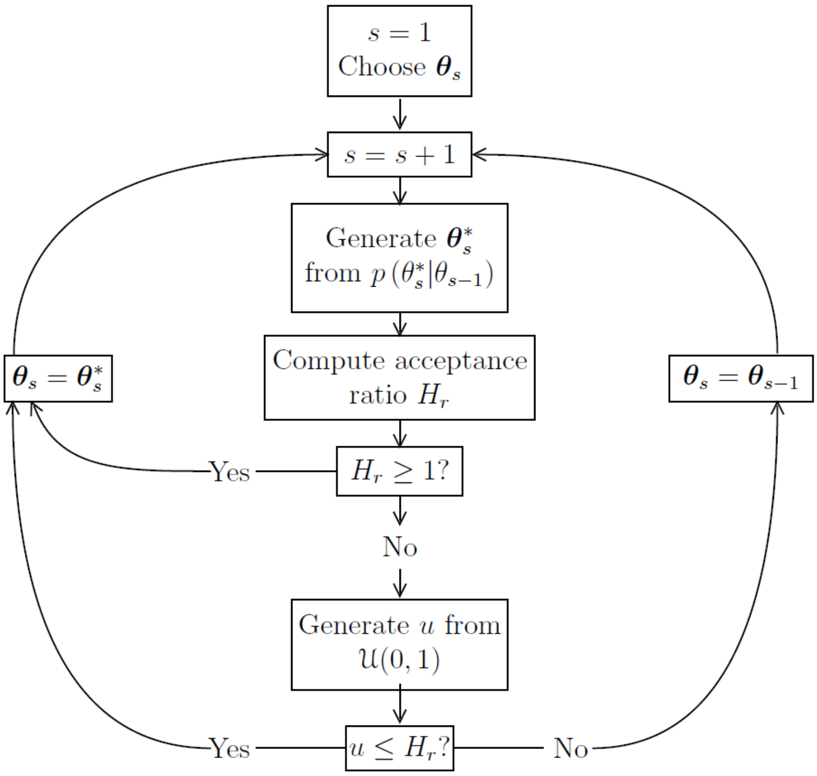

Figure 6: Metropolis-Hastings algorithm, after Rogers and Girolami (2017).

3.2.3 Convergence

Theoretically, any distribution would eventually converge but, when drawing a finite number of samples, distribution convergence within the chain length must be assessed (Rogers and Girolami, 2017). For computational efficiency, chain length should be the minimum that converges to the stationary posterior. Yet, decreasing the number of samples drawn increases the sensitivity. For example, if initial samples are taken from a region of low posterior-likelihood, the sampled distribution will be skewed, delaying convergence to the true distribution.

Without the ability to evaluate the true distribution, the convergence of multiple chains to the same distribution is assessed, establishing independence with regard to starting position. In practice, a measure of convergence, $\hat{R}$, is calculated for each parameter as a function of chain length (Gelman et al., 2014; Gelman and Rubin, 1992).

Given $s = 1, ..., S$ samples of $c = 1, ..., C$ chains, the chain mean, chain variance and global mean (μ_c, σ_c^2 and μ respectively), are calculated as a function of the total number of samples taken, S:

$$\mu_c = \frac{1}{S} \sum_{s=1}^{S} \theta_{sc} \tag{29}$$

$$\sigma_c^2 = \frac{1}{S-1} \sum_{s=1}^{S} (\theta sc - \mu_c)^2 \tag{30}$$

$$\mu = \frac{1}{C} \sum_{c=1}^{C} \mu_c \tag{31}$$

where θ_{sc} indicates the s^{th} sample from the c^{th} chain. Similarly, intra-chain and inter-chain variances, W (32) and B (33) respectively, are calculated as a function of S. Subsequently, V, an estimate of the marginal posterior variance of θ, is given by (34).

$$W = \frac{1}{C} \sum_{c=1}^{C} \sigma_c^2 \tag{32}$$

$$B = \frac{S}{C-1} \sum_{c=1}^{C} (\mu_c - \mu)^2 \tag{33}$$

$$V = \frac{S-1}{S} W + \frac{1}{S} B \tag{34}$$

Finally, $\hat{R}$ (35) is calculated as a function of S. As S tends to infinity, $\hat{R}$ tends to 1. The acceptance threshold selection process is a balance between ensuring the chains have truly converged and accepting realistic variation between converged distributions. Here, a typical threshold of $|\hat{R} - 1| < 0.01$ is used to define convergence (Rogers and Girolami, 2017). The samples taken before this point are discarded as the "burn-in" phase.

$$\hat{R} = \sqrt{\frac{V}{W}} \tag{35}$$

3.2.4 Autocorrelation

For approximating (21) using (22), the samples, s, must be independent. However, in implementing an MH algorithm which ensures that samples are taken from the region of interest, a high degree of consecutive sample dependency is introduced. Two correction methods are partnered to regain sample independence. Firstly, sampling is designed to maximise independence through suitable selection of transition density covariance. Matrix optimisation is governed by two relationships: independence increases with the magnitude of covariance factors; proposal rejection increases simultaneously, increasing dependence. Secondly, the remaining dependence can be removed through a thinning process, where all but every k^{th} sample are discarded.

Inter-sample dependence can be measured using autocorrelation, the correlation between samples being separated by a lag. The autocorrelation of each chain is calculated using (36):

$$a_k = \frac{1}{(S-k)\sigma^2} \sum_{s=1}^{S-k} (\theta_s - \mu)(\theta_{s+k} - \mu) \tag{36}$$

where k is the lag and σ^2 and μ are the variance and mean of the samples, S, respectively. a_k is plotted as a function of k to determine a suitable degree of thinning. Accepted samples can be used for likelihood integral evaluation, circumventing the aforementioned difficulties of sampling for Monte Carlo estimation (Girolami, 2008).

3.3 Method

3.3.1 Set-up

Data Generation Due to the required analysis mesh, the relative computational cost of the T-bar field integration is too high given the project timeline. Instead, two undrained biaxial compression datasets (Table 1) were generated in FE simulations performed by Singh (pers. comm.) for use in the Bayesian analysis.

Table 1: Structured Modified Cam Clay model parameters for biaxial datasets B1 and B2. The target variables are indicated in red.

Parameter	Description	B1	B2
M	CSL slope	0.8984	0.8984
λ^*	NCL slope	0.0811	0.0811
κ^*	URL slope	0.019	0.019
N	Intercept at 1kPa	1.195	1.195
v	Poisson's ratio	0.3	0.3
k	Softening rate	0.2	0.8
A	Volume/shear strain control	0.2	0.2
s_i	Initial soil sensitivity	2.8	6.8
s_f	Sensitivity of remoulded soil	1.0	1.0

Transition Density Covariance Appropriate magnitudes of transition covariance matrix factors were identified through an evaluation of the rejection ratio. A rejection ratio, r, of approximately 0.4 is desirable for matrix optimisation (Girolami pers. comm.) as described in Section 3.2.4. Using dataset B1, three chains, with covariance factors displayed in Table 2, were initiated at the B1 FE input parameters ($\boldsymbol{\theta} = \{0.2, 2.8\}$).

Table 2: Transition density covariance matrix input parameters for optimisation.

B1	v_k	$v_{s_{ep}}$
C_1	0.2	1.0
C_2	0.1	0.5
C_3	0.04	0.2

3.3.2 Markov Chain Monte Carlo

Chains of 100,000 accepted samples were produced with the inputs in Table 3 via a Python script (Appendix A). All chains were initiated at the dataset input parameters as a region of known high-likelihood, to ensure swift convergence. These chains were used to generate kernel density estimations of the posterior-likelihood distributions.

Table 3: Markov Chain Monte Carlo input parameters.

Dataset	Prior bounds		Frames	Noise variance, v	Reference
	k_{lim}	$s_{ep_{lim}}$			
B1	1.0	5.0	54/108	1×10^{-7}	B1H1
				1×10^{-6}	B1H2
				1×10^{-5}	B1H3
			108/108	1×10^{-7}	B1F1
				1×10^{-6}	B1F2
				1×10^{-5}	B1F3
B2	1.0	10.0	54/108	1×10^{-7}	B2H1
				1×10^{-6}	B2H2
				1×10^{-5}	B2H3
			108/108	1×10^{-7}	B2F1
				1×10^{-6}	B2F2
				1×10^{-5}	B2F3

Multiple chains with varying initial locations were run for B1H2, B1F2, B2H2 and B2F2 to confirm convergence within the total chain length and to set a "burn-in" length. Initial chain points were selected to be well-distributed across the predicted posterior-space and are detailed in Table 4.

Table 4: Initial parameter sets, $\boldsymbol{\theta} = \{k, s_{ep}\}$ for convergence chains.

Chains	B1	B2
A	$\{0.2, 2.8\}$	$\{0.8, 6.8\}$
B	$\{0.8, 4.0\}$	$\{0.8, 2.0\}$
C	$\{0.5, 2.5\}$	$\{0.2, 6.8\}$

The relationship between strength degradation observation and parameter calibration, by means of uncertainty, was explored by varying the number of time steps used in integration. Numerical analysis generated 108 frames. The full-frame dataset, 'F', and half-frame dataset, 'H', were used to represent the data produced from large deformation testing (where a significant proportion of remoulding has occurred) and traditional calibration testing (where typically only peak strength is reached) respectively (Figure 7). As a consequence of the input parameters used, a further comparison can be drawn concerning the description of the remoulding process with B2 nearer to the complete remoulded state by the 108^{th} frame.

Similarly, the relationship between experimental error and uncertainty was ascertained by varying the order of magnitude of additive Gaussian noise. The variance levels, $v = 1 \times 10^{-7}$, $v = 1 \times 10^{-6}$ and $v = 1 \times 10^{-5}$, correspond with internal work error of the order of 0.01 %, 0.1 % and 1 % respectively. This was considered to represent calibration of raw numerical datasets (negligible error) and high-precision experimental datasets.

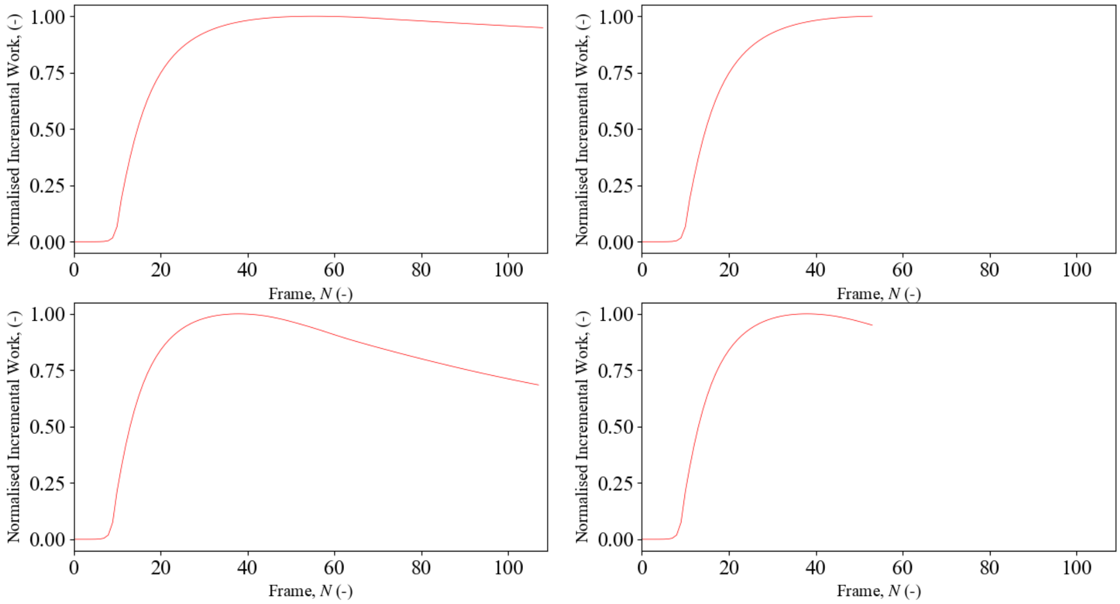

Figure 7: Normalised incremental external work for B1F (top left), B1H (top right), B2F (bottom left) and B2H (bottom right).

3.4 Results and Discussion

3.4.1 Transition Density Covariance

Transition covariance matrix, C_2, possessed the most appropriate rejection ratio (Table 5) and was thus adopted for the remaining analysis. The autocorrelation threshold lag values, $k_{t,k}$ and $k_{t,s_{ep}}$, were evaluated to exemplify the relationship recognised in section 3.2.4.

Table 5: Rejection ratios of transition density covariance matrix input parameters and autocorrelation thresholds.

	Rejection ratio, r	$k_{t,k}$	$k_{t,s_{ep}}$
C_1	0.585	12	10
C_2	0.391	49	47
C_3	0.200	>200	>200

3.4.2 Convergence

The convergence of the multiple chains, initiated as specified in Table 4 at the $v = 1 \times 10^{-6}$ error level, is exemplified in Figure 8 (see Appendix B; Figures 23 to 25 for the remaining convergence plots). Convergence to the true posterior distribution cannot be proven within a finite chain. However, in all cases, the chains failed to prove failure of convergence by converging to a stationary distribution within a finite number of samples (Table 6). Pre-convergence samples impose a bias upon the final distribution. For the remaining MCMC analysis, the maximum of the threshold values, S_t, for each dataset was set as a "burn-in" period, B_i, within which samples were discarded.

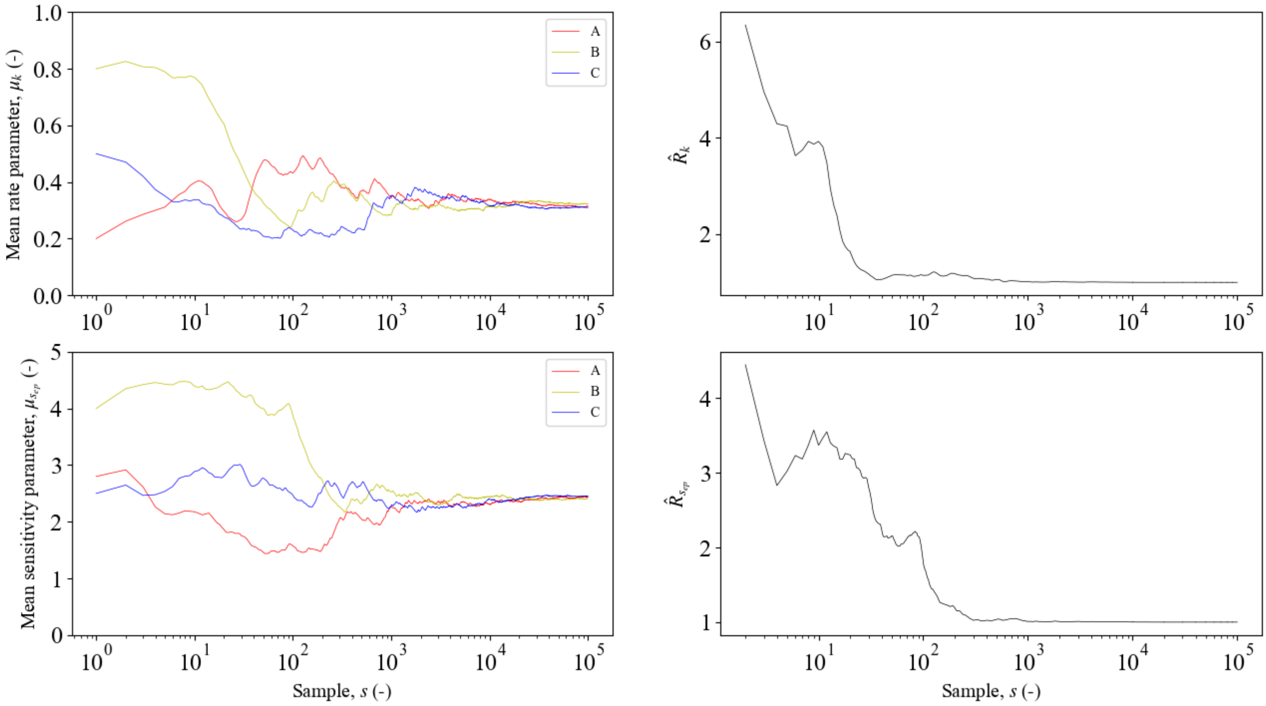

Figure 8: Convergence of multiple chains for dataset B1F2; posterior distribution means (left) and $\hat{R}$ (right) for k (top) and s_{ep} (bottom).

Table 6: Samples required for convergence, S_t, and "burn-in" period, B_i.

Dataset	$S_{t,k}$	$S_{t,s_{ep}}$	B_i
B1F2	1044	988	1044
B1H2	304	876	876
B2F2	73	7155	7155
B2H2	180	3374	3374

The B2 datasets converged after a greater number of samples than B1. Most significantly, B2F2 convergence was slower, by a factor of 7, compared to B1F2. Considering Figure 24 and Table 6, $\hat{R}_k$ for B2 converged quickly in 73 samples. However, $\hat{R}_{s_{ep}}$ appears to remain just above the specified convergence threshold for a large number of samples; initially crossing the thresholds $|\hat{R} - 1| < 0.1$ and $|\hat{R} - 1| < 0.02$ at samples 648 and 3484 respectively. This could be associated with initial parameter set position relative to the posterior. However, this was not a concern given that further analysis only used chains initiated at the input parameter set, guaranteeing convergence before the end of the "burn-in" period.

A discrepancy in converged chain mean exists for dataset B1H2, where chain A (initiated at $\boldsymbol{\theta} = \{0.2, 2.8\}$) settled to $\mu_k = 0.35$ compared to $\mu_k = 0.43$ in the other two chains (compare Figure 23 with 8, 24 or 25). While overall convergence was attained, this was an atypical result. If this stage were not purely for comparative purposes, this discrepancy should have been investigated further.

3.4.3 Autocorrelation

The "burnt" chains, initiated at the numerical input parameter sets, were thinned according to the results of autocorrelation (Figure 26 to 29 in Appendix B). Comparison between thinned and unthinned chains ensured that thinning made no change to the resulting distributions. Once established, the unthinned chains were used in subsequent analysis.

Table 7: Lag required to remove sample autocorrelation, k. k_T is the threshold lag used for distribution comparison.

	$k_{t,k}$	$k_{t,s_{ep}}$	k_T
B1F	49	47	50
B1H	42	33	50
B2F	40	99	75
B2H	44	108	75

The majority of lag values achieving the autocorrelation threshold, $a_k \leq 0.1$, were under 50 (Table 7). Autocorrelation, as a function of sample independence, is heavily dependent on the transition density matrix (as discussed in Section 3.2.4). In the interest of fair comparison between datasets, the transition density matrix, optimised for dataset B1, was unaltered for use with dataset B2. As a result, upon running the chains, a lower rejection ratio was seen for B2 than B1 (Table 8) which generated a greater dependency between samples. From the B2 distribution shape, the prior bound, $s_{ep} \leq 10.0$, appeared to restrict the extent of the distribution in the respective dimension, hence the threshold lag value, $k_{t,s_{ep}}$, was significantly higher. Meanwhile, for B1, there was a steep decline in likelihood at $s_{ep} \sim 1$ (transition from strain-softening to strain-hardening), before the prior bound, $0.0 \leq s_{ep}$, acting to curtail sampling density and reduce the severity of an otherwise binary distinction. For further analysis, a separate transition density matrix should be optimised for the B2 dataset.

Table 8: Stabilised rejection ratios.

Rejection Ratio, r	F1	F2	F3	H1	H2	H3
B1	0.377	0.391	0.336	0.336	0.313	0.308
B2	0.364	0.347	0.226	0.197	0.203	0.255

As a consequence of the large lag, the method was inefficient at producing independent sample sets as $97 - 99$ % of samples were discarded in the thinning process. Reduction in autocorrelation could be achieved by altering the sampling method employed. For example, Gibbs sampling (Geman and Geman, 1984; Gelfand and Smith, 1990) inherently mitigates consecutive sample correlation since elements of the parameter set are resampled, conditioned on the current values of the other parameters (Rogers and Girolami, 2017).

For the purpose of understanding the posterior distributions, chain-thinning was used only to affirm consistency of the result with independent sampling. Thinning minimises memory usage and post-processing time (Link and Eaton, 2011). However, for the greatest precision in posterior distribution understanding, the full chains should be used (Geyer, 1992; MacEachern and Berliner, 1994).

Comparing the thinned chains with their full counterparts (Figure 9), negligible change was observed in the mean and standard deviation of the marginalised distributions (Table 9). The consistency of the sample sets permitted further analysis with unthinned chains.

Table 9: Percentage difference in distribution properties between unthinned chains and chains thinned according to k_T.

Dataset	Mean, μ		Standard deviation, σ	
	$\Delta\mu_k$ %	$\Delta\mu_{s_{ep}}$ %	$\Delta\sigma_k$ %	$\Delta\sigma_{s_{ep}}$ %
B1F2	0.9	-0.6	2.7	1.0
B1H2	1.6	1.0	1.1	-0.5
B2F2	-0.3	0.2	-0.4	0.8
B2H2	-0.2	0.5	1.9	0.8

3.4.4 Posterior Distributions

The posterior distributions (properties in Table 10) may be seen in Figures 10 and 11. All posterior distributions use Gaussian kernel distributions and the Scott bandwidth method.

Shape Datasets B1 and B2 exhibited kernel density estimations of different shape, most clearly observed in B1F1 (Figure 10 (top left)) and B2F1 (Figure 11 (top left)). Both exhibited clear high-likelihood regions in the form of single ridges within the bounds of the respective prior distributions. Furthermore, both ridges have qualitatively taken a curved characteristic form, $y \propto \frac{1}{x^n}$ ($0 < x$ and $1 \leq n$). This behaviour was more prominent and distinctive in dataset B1 due to its proximity to the apparent asymptote, $s_{ep} = 1.0$ (i.e. fully remoulded). The test frames can be normalised such that they represent not time, nor progression, but proportion of observed strength degradation (justified by the similar form of incremental external work curves between datasets, as displayed in Figure 7). Considering a proportionality argument, simulation B2 mobilised a greater proportion of degradation within the same number of frames:

$$N \propto \frac{s_{ep}}{k} \tag{37}$$

$$\frac{N_{B1}}{N_{B2}} = \frac{s_{ep,B1}}{s_{ep,B2}} \frac{k_{B2}}{k_{B1}} \tag{38}$$

where N is the number of frames to observe full strength degradation. Using the numerical input for the other parameters, $N_{B1}/N_{B2} = 1.65$. Consequently, B1 was seen to have greater parametric uncertainty in k, the rate of strength degradation. This was particularly apparent for lower s_{ep}, representing the magnitude of strength degradation and thereby the magnitude of decrease in the incremental external work. With greater observation of the degradation slope, uncertainty of the slope rate decreases, reducing the ridge tail seen tending towards the asymptote, $s_{ep} = 1.0$. This causes the ridge shape to increase in similarity with that of dataset B2.

The ridge proportions and orientations, are indicative of the existing limitations with current calibration techniques. Namely, in the absence of the full observation of strength degradation to an asymptotic state, there is significant parametric uncertainty. This is exemplified further when comparing half-frame and full-frame dataset analyses.

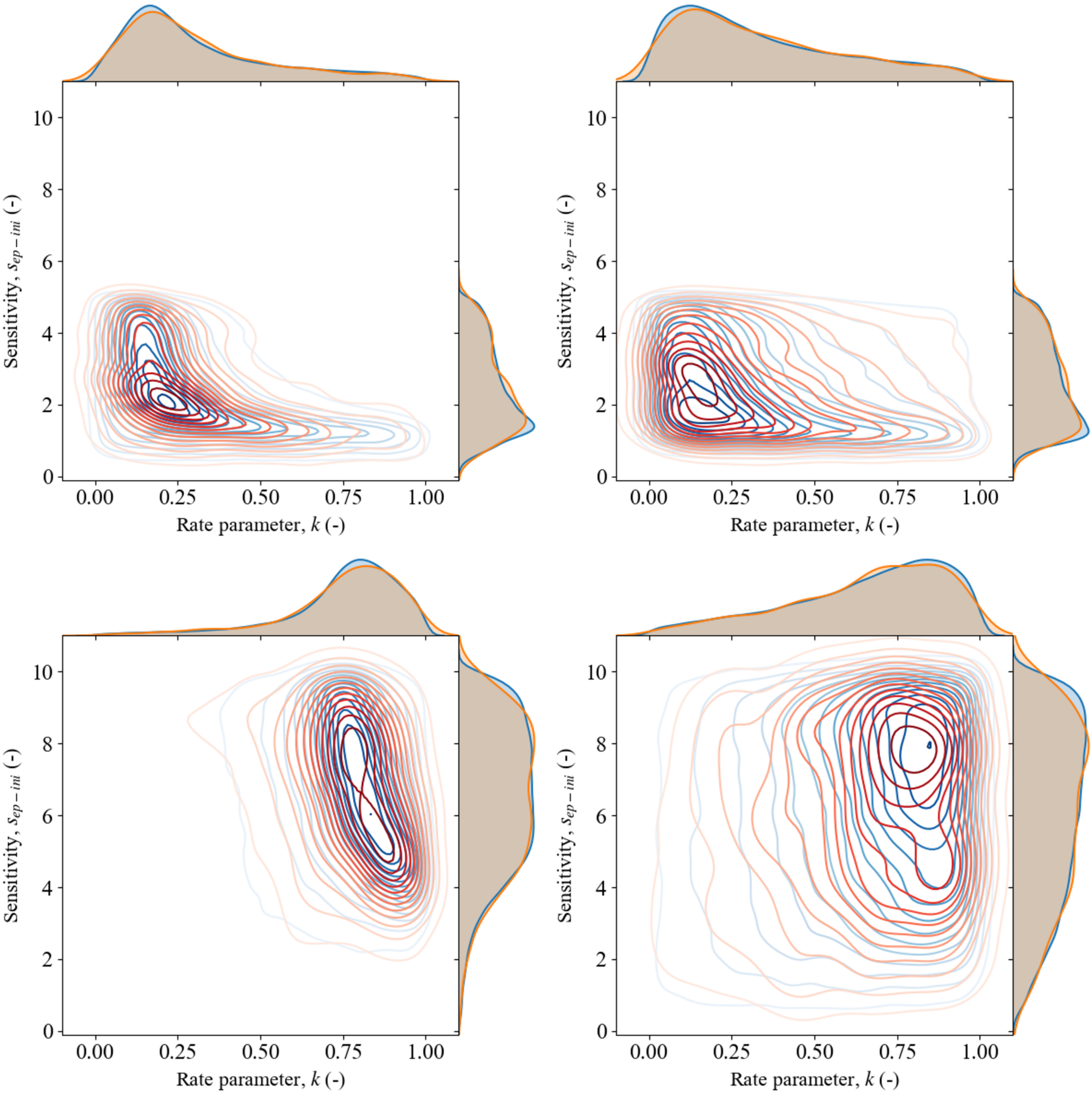

Figure 9: Comparative kernel density estimations of the thinned chain (reds) and unthinned chain (blues) posterior-likelihood distributions. All chains were initiated at the numerical input parameter set of the respective dataset. "Burn-in" and thinning were conducted according to Table 6 and 7 respectively.

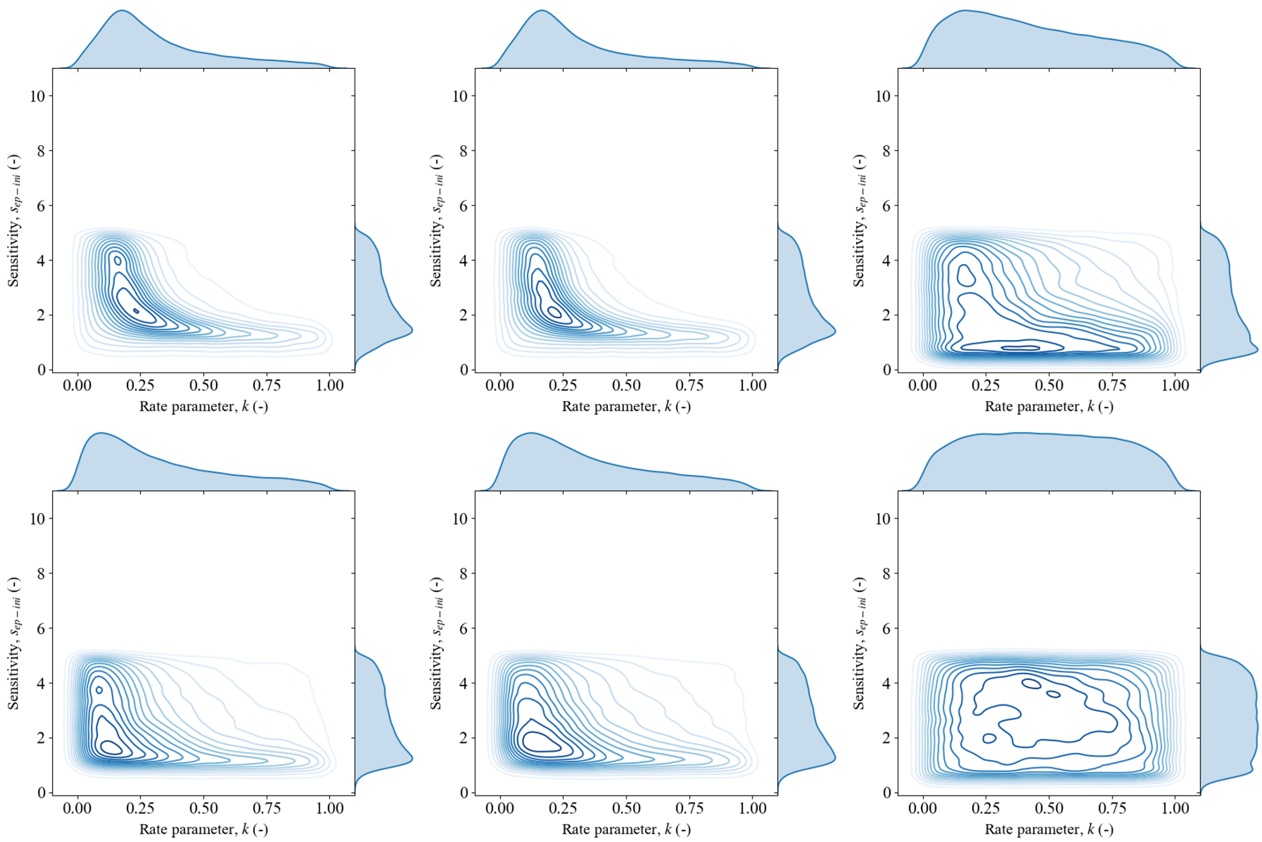

Figure 10: B1 kernel density estimation posterior distribution: F1-F3 (top, left to right) and H1-H3 (bottom, left to right).

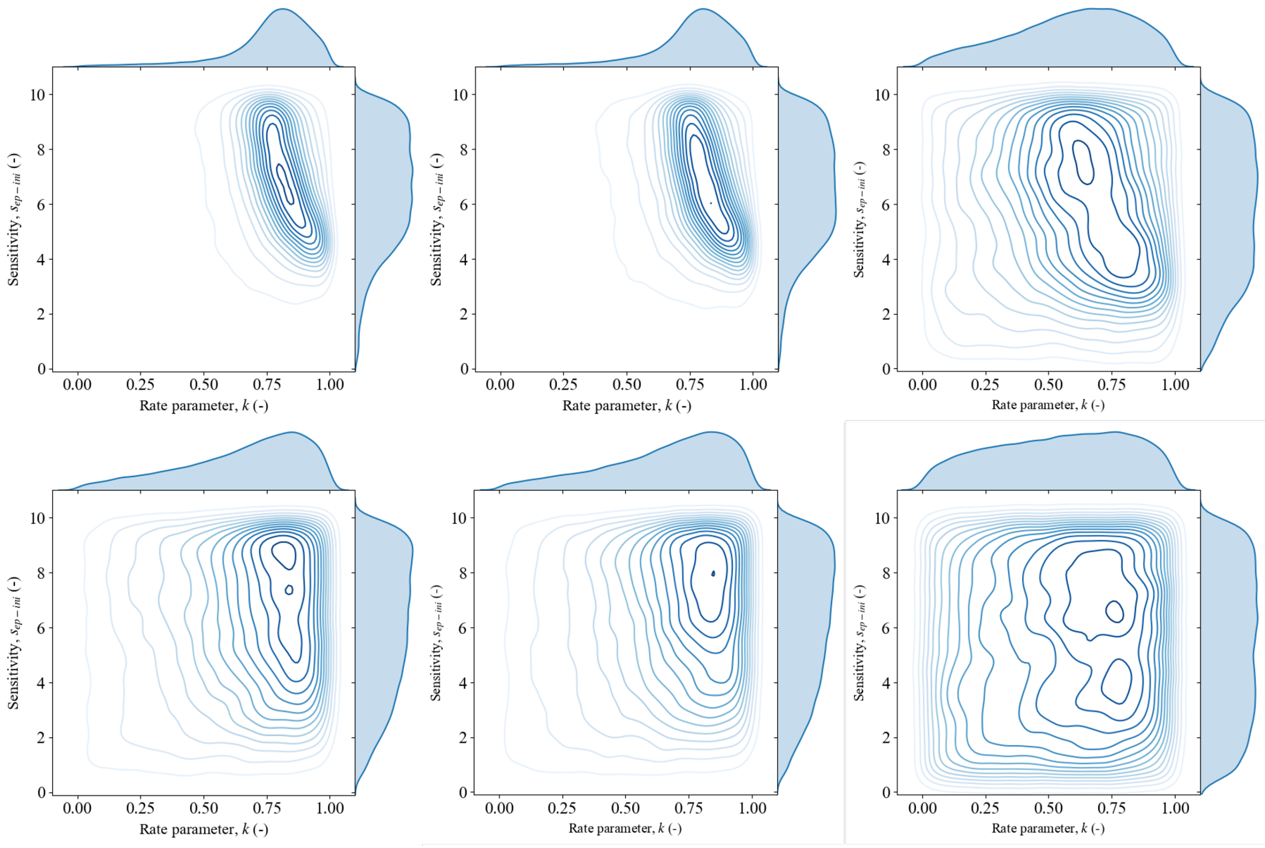

Figure 11: B2 kernel density estimation posterior distribution: F1-F3 (top, left to right) and H1-H3 (bottom, left to right).

Table 10: Posterior distribution properties. Univariate mean and standard deviation percentage change is measured relative to the lowest error level for each dataset group (e.g. B1F1 for all B1F). Z, the peak magnitude, is a relative property only, normalised by Z_{B1F1} for all B1 and by Z_{B2F1} for all B2.

Dataset	Mean, μ				Standard Deviation, σ				KDE Peak			KDE Covariance Matrix, Σ			
	μ_k	$\mu_{s_{ep}}$	$\Delta\mu_k$ %	$\Delta\mu_{s_{ep}}$ %	σ_k	$\sigma_{s_{ep}}$	$\Delta\sigma_k$ %	$\Delta\sigma_{s_{ep}}$ %	k	s_{ep}	Z	v_k	$v_{s_{ep},k}$	$v_{k,s_{ep}}$	$v_{s_{ep}}$
B1F1	0.3214	2.4660	0.0	0.0	0.2248	1.1575	0.0	0.0	0.24	2.12	1.00	0.0011	-0.0025	-0.0025	0.0290
B1F2	0.3142	2.4357	2.2	1.2	0.2289	1.1726	-1.8	-1.3	0.21	2.10	1.02	0.0011	-0.0025	-0.0025	0.0297
B1F3	0.4178	2.2622	-30.0	8.3	0.2627	1.3021	-16.8	-12.5	0.43	0.77	0.45	0.0015	-0.0015	-0.0015	0.0367
B1H1	0.3321	2.5263	0.0	0.0	0.2576	1.1756	0.0	0.0	0.14	1.57	0.74	0.0014	-0.0015	-0.0015	0.0299
B1H2	0.3471	2.5277	-4.5	-0.1	0.2552	1.1662	0.9	0.8	0.14	1.79	0.70	0.0014	-0.0012	-0.0012	0.0294
B1H3	0.4877	2.6678	-46.8	-5.6	0.2677	1.2536	-3.9	-6.6	0.43	3.98	0.34	0.0015	-0.0002	-0.0002	0.0340
B2F1	0.7600	6.5442	0.0	0.0	0.1749	2.0338	0.0	0.0	0.81	7.05	1.00	0.0007	0.0002	0.0002	0.0914
B2F2	0.7477	6.4959	1.6	0.7	0.1804	1.9975	-3.1	1.8	0.84	6.01	0.77	0.0007	-0.0005	-0.0005	0.0881
B2F3	0.5882	5.7132	22.6	12.7	0.2394	2.4313	-36.9	-19.5	0.62	7.91	0.33	0.0013	-0.0008	-0.0008	0.1305
B2H1	0.6531	5.9928	0.0	0.0	0.2400	2.4003	0.0	0.0	0.84	8.71	0.41	0.0013	0.0007	0.0007	0.1256
B2H2	0.6540	6.0392	-0.1	-0.8	0.2399	2.3960	0.1	0.2	0.84	7.87	0.42	0.0013	0.0012	0.0012	0.1251
B2H3	0.5420	5.3219	17.0	11.2	0.2636	2.6342	-9.8	-9.7	0.76	6.52	0.23	0.0015	0.0006	0.0006	0.1512

Error Level Increasing the magnitude of additive noise acts to diminish the distinctiveness (relative magnitude) of the posterior high-likelihood ridge in transition to a uniform plateau across the prior parameter-space (seen left to right in Figure 10 and 11). Conceptually, this is congruent with the order of magnitude of the noise approaching that of the decrease in incremental external work. As mean trends in the finite number of frames become less identifiable, given noise, parametric uncertainty increases, thereby eroding the evidence supporting a high-likelihood region and broadening the distribution. In dataset B1, parameter sets associated with no, or little, strength degradation increase in likelihood. However, all dataset variations eventually plateau with a mean shift towards the centre of the prior parameter-space and an increased distribution uniformity.

Considering both full-frame and half-frame datasets, little change in univariate distribution properties are observed between the first two error levels; at worst, a 4.5 % change in mean (μ_k B1F1-B1F2, Table 10) and a 3.1 % increase in standard deviation (σ_k B2F1-B2F2, Table 10). This is congruent with the bivariate kernel density estimations which observed little change in covariance matrix parameters and a peak of similar magnitude remaining within the initial high-likelihood region. The similarity in posteriors across these error levels suggests a capacity for absorbing error, proving the robustness of the proposed calibration framework.

Upon further increase in error level, the distinctiveness of the ridge is lost as the distribution plateaus across the prior-space. For example, in the full-frame datasets there are univariate mean changes of up to 30.0 % (μ_k B1F1-B1F3, Table 10) with a shift towards the centre of the prior-space and univariate standard deviation increases of up to 36.9 % (σ_k B2F1-B2F3, Table 10). There is a large drop in the relative magnitude of the bivariate peak, Z, with it typically halving in size, accompanied by a significant increase in variance. This demonstrates that the full-field strain measurements need to be of high precision for the proposed calibration framework to be robust. Furthermore, observing a greater proportion of the strength degradation process acts to reduce error sensitivity, as mean trends and curve characteristics become more identifiable.

Univariate changes in k are found to be typically greater than those in s_{ep} as a consequence of the initial distribution proportions (B1F1 and B2F1, Figure 10 and 11). s_{ep} spans a larger proportion of the prior-space as there is greater uncertainty, therefore smaller changes reach the prior boundaries. However, the relative magnitude decrease of the s_{ep} univariate peak is of similar order to k, as indicated by that in the bivariate kernel density estimation.

Degradation Observation Increasing the proportion of the degradation process observed in the calibration input dataset decreases parametric uncertainty. In this comparison, the half-frame dataset is analogous to current calibration practice in which testing ceases near the beginning of the "remoulding" process typically accompanied by strain-localisation. Since dataset B1 and B2 are biaxial tests, they are not directly analogous to the proposed calibration tests. Overall, the difference in MCMC outputs is generally representative of the impact of more comprehensive observation of the strength degradation process. Fundamentally, the justification for full observation of the "remoulding" process is that of interpolation being preferential to extrapolation from an uncertainty perspective.

Between frame proportions, the univariate means shift towards the centre of the prior-space, corresponding to a loss in posterior informativeness. Correspondingly, mean values increase for datasets B1 and decrease for dataset B2 (see Table 11). In the majority of cases, the standard deviation is significantly larger for half-frame datasets (up to 37.3 % for B2F1-B2H1) associated with greater parametric uncertainty. The greatest exception, $\sigma_{s_{ep}}$ for dataset B1, is a result of the s_{ep} distribution spanning the majority of the prior breadth (excluding values corresponding to strain-hardening) such that minimal change in $\sigma_{s_{ep}}$ will occur with the removal of frames.

Table 11: Univariate mean and standard deviation percentage changes between full-frame and half-frame datasets.

	$\Delta\mu_k$ %	$\Delta\mu_{s_{ep}}$ %	$\Delta\sigma_k$ %	$\Delta\sigma_{s_{ep}}$ %
B1F1-B1H1	3.3	2.4	14.6	1.6
B1F2-B1H2	10.5	3.8	11.5	-0.5
B1F3-B1H3	16.7	17.9	1.9	-3.7
B2F1-B2H1	-14.1	-8.4	37.3	18.0
B2F2-B2H2	-12.5	-7.0	33.0	20.0
B2F3-B2H3	-7.9	-6.8	10.1	8.3

The change in standard deviation between the half-frame and full-frame datasets acts as a metric for determining the informativeness of the section of the dataset including the majority of the strength degradation process. Since the degradation rate differs between the two datasets, a different relative proportion of the degradation process is represented. Considering $\Delta\sigma_k$ %, the increase in spread resulting from the full-frame to half-frame reduction is significantly greater for B2. This shows that, the greater the proportion of the degradation process input into the calibration, the greater the parametric certainty gained from the MCMC analysis. This would have been achieved to an even greater degree by the alternative experiment proposed in the latter part of this report.

Summary The certainty of the parameter distributions show high sensitivity to additive noise, an inevitable error in real-world experimentation. Since conventional calibration methods are known to be less accurate and unrepresentative, this study has provided strong support for implementing full-field deformation measurement techniques in the calibration process.

The analysis developed here also proves that the use of a biaxial experiment is sub-optimal due to the modest amount of deformation mobilised. In contrast, an experiment that fully remoulds the soil to its residual state would certainly reduce uncertainty in parameter assessment using the proposed MCMC approach.

4 Experiment

4.1 Introduction

Naturally, the experimental approach has been determined in response to the requirements of the proposed calibration framework. This consists of the accurate evaluation of internal and external work throughout the full "remoulding" process. Consequently, a test consisting of the cyclic displacement of a T-bar full-flow penetrometer (FFP) was selected.

Originally developed for profiling soft clay centrifuge samples (Stewart and Randolph, 1991), FFP tests are increasingly used for site investigation particularly for establishing the shear strength of intact and "remoulded" soils in offshore environments (see Randolph et al., 2018). FFP tests demonstrate numerous benefits over conventional methods (Randolph, 2004; Einav and Randolph, 2005; Yafrate et al., 2009) with specific advantages for constitutive model parameter calibration.

Considering the evaluation of internal work through full-field displacement measurement techniques, FFP testing is beneficial due to the form of failure being displayed. In conventional techniques (e.g. triaxial), failure is often discrete, consisting of the formation of thin, inhomogeneous shear bands, which prove difficult to analyse (e.g. Viggiani and Finno, 1994). Failure in FFP testing, however, consists of well-defined flow mechanisms with closed-form theoretical solutions (Einav and Randolph, 2005). Furthermore, in the case of a T-bar, plane-strain flow is established such that observation of a single plane can characterise the entire failure mechanism.

The absence of permanent, localised failure enables full "remoulding" via cyclic displacement. FFPs are also quicker and more accurate than similar alternatives, such as field vane tests (FVTs) (DeJong et al., 2011).

In the proposed experiment, a T-bar is cyclically displaced in a test box filled with consolidated Kaolin clay (Figure 12). In evaluation of incremental internal work, an array of RP cameras capture images of the exposed plane for processing via GeoPIV-RG to produce displacement fields, and following differentiation, strain fields. In evaluation of external work, a strain gauge in the T-bar shaft and the actuator motor measures the force and displacement respectively. Subsequently, calibration can be conducted using the virtual fields method as described previously.

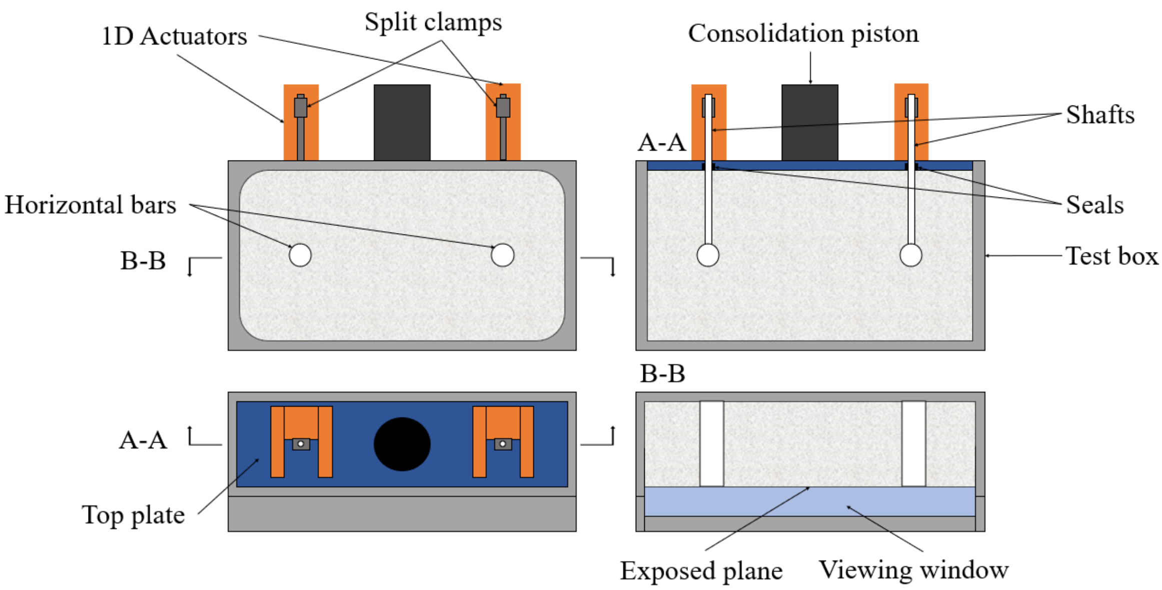

Figure 12: The test-rig configuration: front view (top left), plan view (bottom left), mid-box front view (top right) and mid-box plan view (bottom right).

4.2 Theory

4.2.1 Full-flow Penetrometer Test

General Flow During penetration, the FFP induces non-localised turbulent shearing (flow) as the soil is forced around in a viscous, fluid-like manner (DeJong et al., 2011). The axisymmetric T-bar axis extends perpendicularly to the plane of motion, constraining flow to plane-strain conditions. The flow mechanism generated from displacement of the FFP consists of shearing conditions at the base and top, similar to triaxial compression and extension, respectively, as well as simple shear at mid-height (Randolph and Andersen, 2006).

Remoulding Using an upper bound solution, the average cumulative shear strain rate per pass of the T-bar penetrometer at a given depth is estimated to be 400 % (Einav and Randolph, 2005). This is supported further by a more recent FE modelling estimation of 370 % (Zhou and Randolph, 2009). A common conservative assumption is that remoulded soil is formed by the end of 10 complete cycles (Yafrate et al., 2009).

Geometry The permanent displacement induced by the T-bar shaft alters the deformation pattern from that of an infinitely long cylinder (DeJong et al., 2011). To ensure a full-flow mechanism is established, the relative size and influence of the shaft to the overall mechanism must be small. This is considered in terms of the area ratio:

$$A_R = \frac{A_p}{A_s} \tag{39}$$

where A_R is the area ratio, A_p is the projected area of the bar (perpendicular to motion) and A_s is the cross-sectional area of the shaft. Penetration resistance stabilises at $A_R \approx 10$ (Yafrate et al., 2007) and this limit is therefore used as a design guide.

In absolute terms, the size of T-bar is proportional to the volume of soil mobilised in the failure mechanism. As T-bar size increases, deployability decreases, penetration resistance increases and the effective length of heterogeneity (the length below which, the influence of local property changes are cancelled along the bar length) decreases. These factors must be balanced in the overall design.

Rate Since real soil behaviour is time-dependent, the experimental response is dependent upon the rate of penetration. In undrained conditions, penetration resistance increases with rate due to viscous effects (inherently velocity dependent) (DeJong et al., 2011). However, if the rate is too slow, partial drainage and consolidation can occur (see Figure 13). In variable rate FFP testing the minimum penetration resistance was identified to occur at the transition point between undrained and partially drained behaviour (Bemben and Myers, 1974; Chung et al., 2006).

4.2.2 PIV

General In geotechnics, PIV was developed to describe planar deformation fields precisely without requiring target markers, in the observed medium, which could influence the overall behaviour (White et al., 2003). The method involves recording a sequence of digital images of an exposed textured plane during deformation. Incremental displacement fields are then generated by cross-correlating discretised patches (or subsets) between consecutive images, creating a normalised plane with peaks indicating highest-degree match vectors (Figure 14).

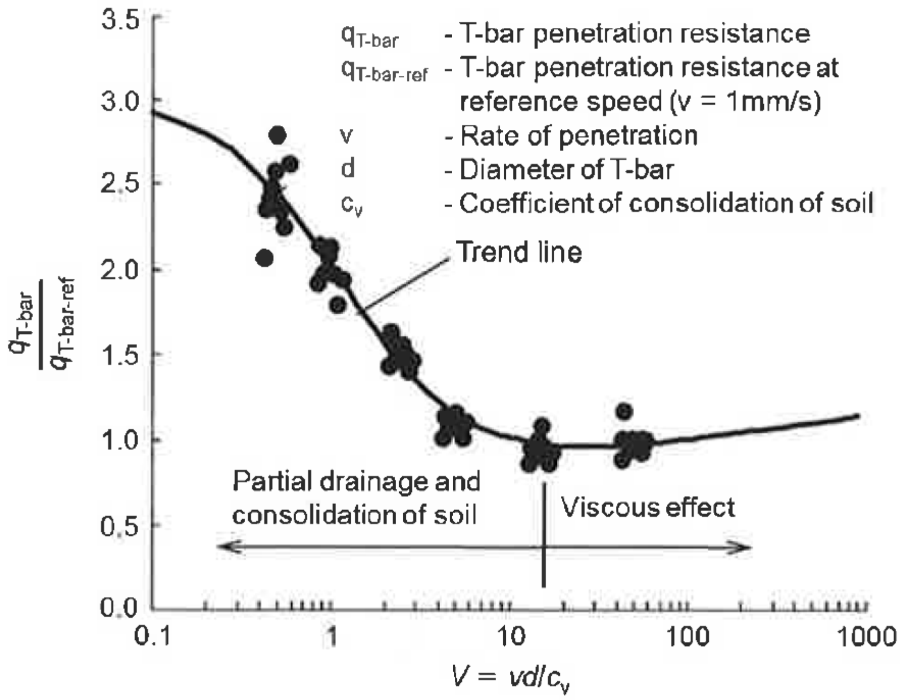

Figure 13: The relationship between penetration rate and resistance (Ganesan and Bolton, 2013; Randolph and Hope, 2004)

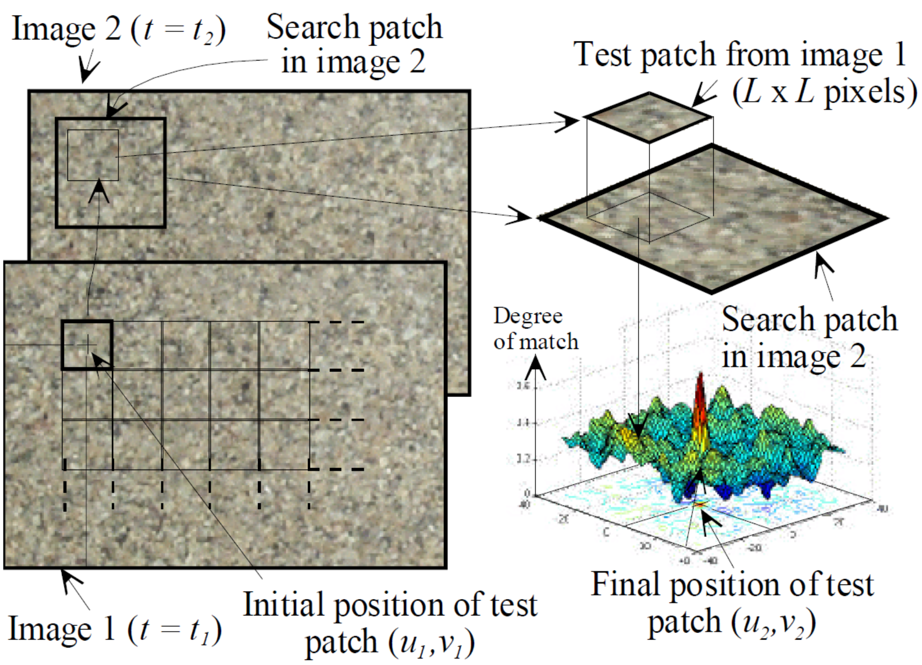

Figure 14: Basic principle of PIV exemplified in a GeoPIV framework (White and Take, 2003).

The ability to employ image correlation to generate displacement fields depends on four implicit assumptions (Stanier et al., 2016):

1. The observed particles are homogeneously distributed across the image.
2. The image texture is perfectly representative of the actual soil displacement.
3. There is sufficient similarity between the "reference" and "target" images such that correlation is computable for all subsets.
4. The image correlation shape function is consistent with the measured deformation.

Experimental success within the proposed calibration framework rests heavily on the second assumption. However, the observed displacements may be unrepresentative of displacement through the thickness due to the effect of wall friction. Given that the internal friction angle of the clay is significantly greater than that at the clay-perspex interface when greased, it is reasonable to assume that the influence of wall friction is likely to be negligible.

GeoPIV-RG (Stanier et al., 2016) is implemented in this project. This uses a first order subset shape function, bi-quintic B-spline image intensity interpolation and an Inverse Compositional Gauss-Newton (IC-GN) solver (Pan et al., 2013) to obtain the optimal subset shape function deformation parameters.

Seeding Artificial seeding increases the precision of PIV by producing "texture" on the exposed plane, improving target identification. The particles, differing in intensity to the matrix, create distinctive patterning. The success of the method is dependent on the seeding proportion as well as the size and colour of the particles.

Since under-seeding or over-seeding could result in insufficient contrast, optimisation must be conducted. The artificial seeding ratio (ASR) is optimised to maximise the potential contrast within PIV:

$$ASR = \frac{\overline{I_{px}} - \overline{I_{px,m}}}{\overline{I_{px,p}} - \overline{I_{px,m}}} \tag{40}$$

where $\overline{I_{px}}$ is the mean pixel intensity of a seeded plane, $\overline{I_{px,m}}$ is the mean pixel intensity of an unseeded plane (matrix only) and $\overline{I_{px,p}}$ is the mean pixel intensity of a fully seeded plane (seeding particle only). The seeding optimisation process was exemplified by Stanier and White (2013). Assuming good particle distribution, the higher the standard deviation of pixel intensity within an image, the lower the error, and thus, the optimum ASR is that with highest standard deviation (Figure 15).

Despite the plane-strain conditions imposed in PIV experimentation, seeding particles tend to lose contact with the exposed plane in large deformation regions (Stanier, pers. comm.). Optimisation through energy conservation is sensitive to contact loss as differences in test and search patches creates correlation error and work done in an unseeded region is not recorded. Therefore, the clay matrix was uniformly seeded throughout, not just on the surface. Assuming minimal contact between seeding particles, bulk mechanical properties are largely unaffected. Regardless, the project objective, as a proof of concept, is to identify the strain-softening properties of the seeded matrix, using those parameters in a simulation to see if the experimental response was accurately reproduced with the parameters resulting from the calibration exercise.

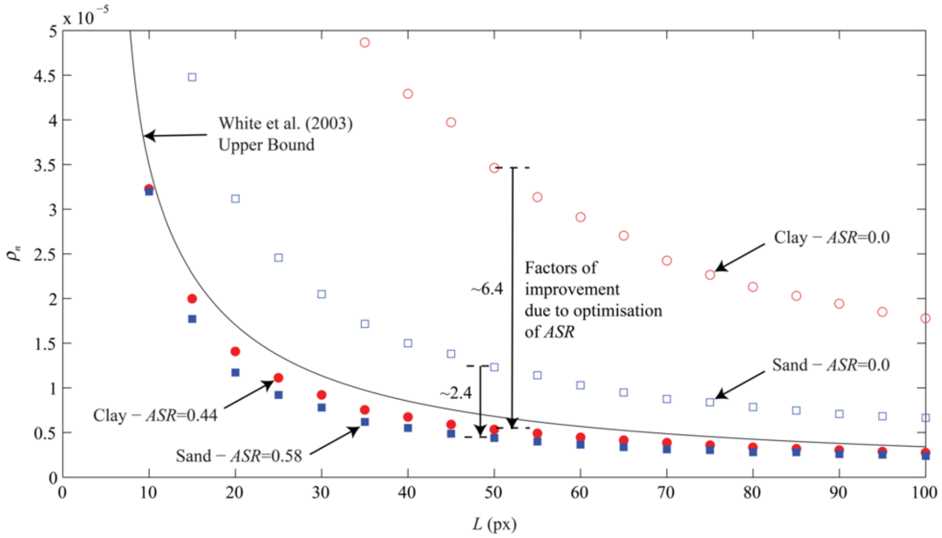

Figure 15: "Standard error", ρ_n (see reference), plotted against subset size, $L(px)$, for different ASRs (Stanier and White, 2013).

Peak Locking Initially, cross-correlation is determined at integer pixel values. At small particle size ratios, d/p (particle to pixel size), the correlation of neighbouring pixels is low relative to that of the sharp peak. In the subsequent sub-pixel refinement, bias is generated towards integer pixel displacements, degrading the process in a phenomenon known as "peak locking" (Westerweel, 1997). As d/p increases, the correlation of neighbouring pixels increase, improving the efficacy of the sub-pixel refinement process. Through successful interpolation, inferred displacements more closely correlate to those imposed (Figure 16). Stanier et al., (2016) conclude that it is critical that $d/p \geq 2$ and that ideally $d/p \approx 4$.

Peak locking can be qualitatively observed in abnormal artefacts including displacement field banding and strain field "vermiculation" (Stanier et al., 2016). In particular, evidence of peak locking can be identified in displacement histograms and sub-pixel component histograms. In the case of peak locking, the displacement histogram will present regular peaks at integer pixel values, rather than a smooth distribution, and the sub-pixel component histograms will exhibit bias towards zero rather than a uniform distribution.

In observance of the peak locking resolution requirements, the field of view (FoV) should be checked to ensure there is a suitable particle size ratio in the current camera positioning. The marginal (41) and ideal (42) cases are expressed as:

$$d/p = \min\left(\frac{di}{X}, \frac{dj}{Y}\right) \geq 2 \tag{41}$$

$$d/p = \min\left(\frac{di}{X}, \frac{dj}{Y}\right) \approx 4 \tag{42}$$

where d is the particle size (diameter, mm), i is the horizontal image size (pixels), j is the vertical image size (pixels), X is the horizontal model dimension (mm) and Y is the vertical model dimension (mm).

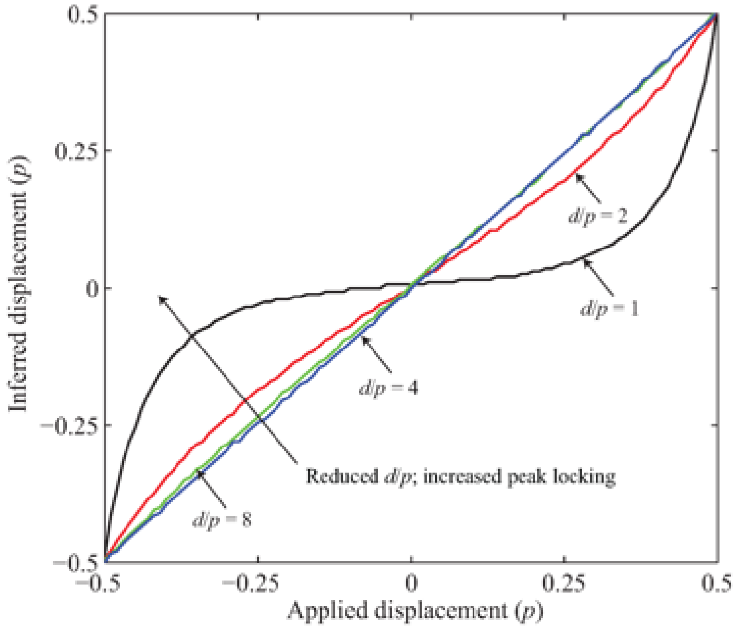

Figure 16: The relationship between the inferred and applied displacement at several d/p values. The generated image subject to sub-pixel displacement consisted of a Gaussian intensity distribution. Figure reproduced from Stanier et al., (2016).

Image Quality Several metrics can be used to evaluate image properties enabling quantitative consideration of suitability. A crude measure of contrast is the standard deviation of subset pixel intensities, σ_{I_s}:

$$\sigma_{I_s} = \sqrt{\frac{1}{n_s} \sum_{i,j \in n_s} \left(I_{i,j} - \bar{I}\right)^2} \tag{43}$$

where n_s is the total number of pixels in the subset, $I_{i,j}$ is the intensity of the i^{th} pixel on the j^{th} row and $\bar{I}$ is the mean subset pixel intensity. For good contrast, $\sigma_{I_s} \geq 15$ is desirable (Stanier and White, 2013). However, consideration of the standard deviation of subset pixel intensities does not take into account the spatial distribution of pixels, crucial for generating the necessary texture. A better approach, which considers the spatial arrangement of intensities, is the sum of the square of the subset intensity gradients (SSSIG), a measure of contrast frequency:

$$SSSIG_x = \sum_{i,j \in n_s} (I'_x)^2 \tag{44}$$

$$SSSIG_y = \sum_{i,j \in n_s} (I'_y)^2 \tag{45}$$

$$\overline{SSSIG} = \sum_{i,j \in n_s} \frac{1}{2} \left[(I'_x)^2 + (I'_y)^2\right] \tag{46}$$

where I'_x and I'_y are the intensity gradients in the x and y direction respectively and n_s is the total number of pixels. Assuming a suitable d/p, $\overline{SSSIG} \geq 1 \times 10^5$ should be acquired for sufficient quality (Pan et al., 2008).

Distortion The image-space and the object-space must be accurately related by a transformation function. Distortion consists of deviation from rectilinear projection. In practice, removing the sources of distortion is impractical, so, instead, images must be calibrated with reference to known positions.

The experimental set-up provides several sources of distortion. Minor misalignment of the camera, such that the viewing axis is not perfectly perpendicular to the object plane, can create a large loss in accuracy (White et al., 2003). Furthermore, the refraction of light through the viewing window causes image scale variation.

Limitations exist with the pinhole camera model approximation; a real lens has radial (Slama, 1980) and tangential distortion resulting from optical effects such as non-collinear centres of curvature, generating "barrelling" or "pin-cushion" behaviour in the final image. The pixels will not be perfectly square, requiring scaling according to the pixel aspect ratio (Heikkila and Silven, 1997).

4.3 Method

4.3.1 Seeding

Several seeding options were trialled. In proportion to larger scale production, the Kaolin clay matrix ($30 - 50$ μm grain size) was mixed to 120 % water content, approximately twice the liquid limit.

A seeded clay sample (with visibly uniform distribution) was applied evenly across a perspex plate ~ 80 mm above a Raspberry Pi (RP) camera. Once recorded, sample photos were processed by creating a histogram of pixel intensity ($0 - 255$) to characterise image texture and identify optimum seeding proportions for contrast.

Seeding proportions were altered by mixing in additional unseeded clay to the sample. Residual loss (considered in proportion to the clay seeding) was minimised by thorough scraping. The maximum proportion of mass lost, which decreased with increasing amount of sample, was 2.5 %.

In summary, of the seeding particles that did not proceed beyond initial trials:

Titanium coated mica Contrast with the matrix was provided by the high reflectivity of the flakes ($10 - 60$ μm diameter), creating a sparkling or glittering effect in ambient lighting. To provide sufficient contrast, it was considered necessary to add black ink to the clay. However, no visible contrast was found despite the large range in proportions tested. It is believed that the dyed clay coated the flakes, rendering them ineffective for seeding.

Strontium aluminate particles Contrast with the matrix was provided by the particles (35 μm diameter) glowing under UV backlight in an otherwise dark room. Standard RP cameras failed to capture the contrast as they contain an infrared filter which, acting as a band filter, also removed UV light. Further tests were conducted using a RP noIR camera (filter removed) with limited improvement in contrast capture. The filter fine-tuning required to capture contrast with an array of RP cameras would have potential but was considered out of project scope.

Grade E sand (250 μm diameter) was dyed black with indelible ink, dried and sieved before being added as a seeding particle. Photographs were taken under ambient lighting over a wide range of proportions ($0.5 - 23\%$ by mass to matrix).

Since seeding particles can be added to the exposed face after consolidation, under-seeding is preferable to over-seeding. For a 100×100 pixel subset, $SSSIG$ was 1.2×10^5 for 12.0 % seeding by mass, exceeding the 1.0x10^5 threshold (Pan et al., 2008). All available black sand was used (making ASR optimisation redundant), a proportion of 9.0 % (to mass of unseeded clay matrix), providing a proportion of 12.4 % once consolidated.

4.3.2 Test Box and Clay Preparation

The box (internal dimensions 790x450x180 mm), with consolidation extension attached (300 mm high), was cleaned and greased (Figure 17). An even base layer of sand was poured to connect drainage ports and a boundary layer (filter paper with a wet tissue paper perimeter) was then placed to stop clay seepage into the drainage layer. The clay, seeded in hand-mixed batches (to avoid seal damage to the clay mixer), was poured into the prepared box evenly and topped with a second boundary layer and the top plate (11.8 kg) before being loaded incrementally up to 50 kPa with a piston. A thin water layer was maintained on top of the clay during consolidation to prevent drying out. Additionally, the drainage ports at the base were connected to a tank of water held at the top clay level, maintaining appropriate base pore pressure during consolidation.

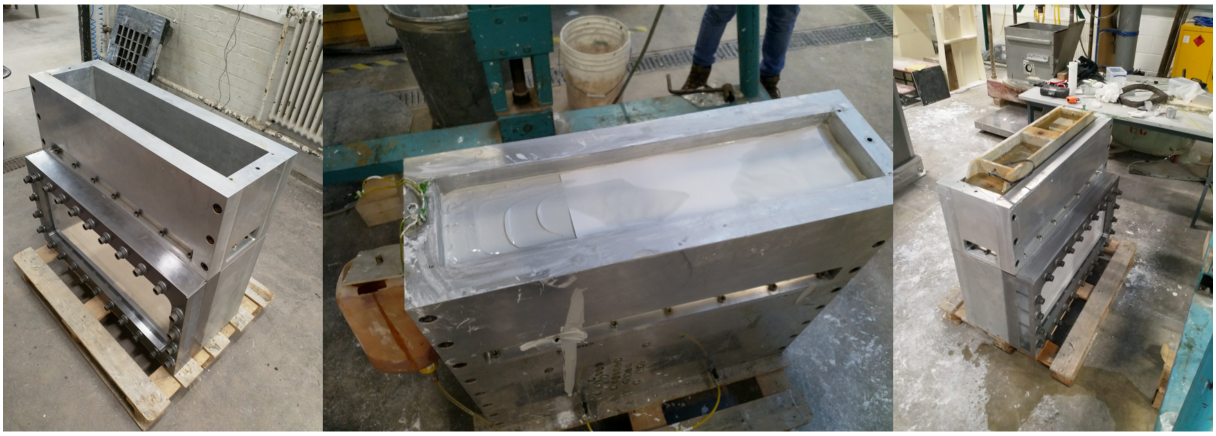

Figure 17: Test box preparation: extension attachment, sand base and greasing (left), clay pouring (centre) and pre-loading (right).

Consolidation (double drainage) under 50 kPa for a duration of 10.5 days was sufficient to achieve 0.23 m settlement (90 % relative settlement); excess clay extending out of the test box would have been removed before testing. The effective stress was applied uniformly to the clay by a piston acting on the top plate.

4.3.3 T-bar Preparation

The T-bar (Figure 18) consisted of a 180 mm long (spanning the box depth) horizontal aluminium cylinder (40 mm external diameter). This ensured that the exposed plane was bound by the plane-strain condition of the T-bar full-flow mechanism (see section 4.2.1). The T-bar connected to the actuator by means of a vertical shaft (20 mm external diameter). The relative shaft and T-bar dimensions achieved an area ratio of 23 ensuring full-flow mechanism stabilisation. This was significantly higher than the requirement ($A_R \approx 10$) as a result of other geometric requirements.

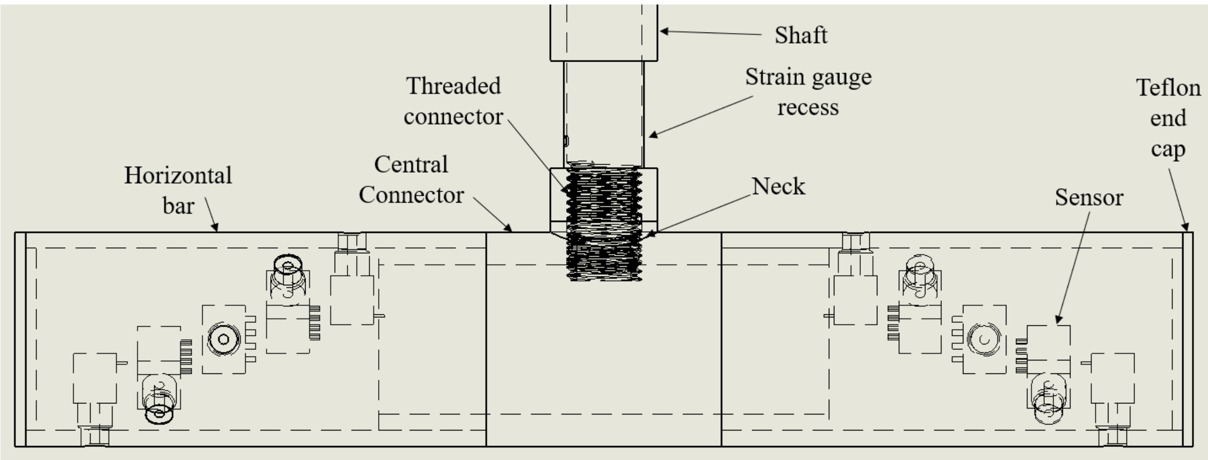

Figure 18: Technical drawing side view layout of T-bar (not to scale).

Figure 19: Wiring of the shaft load cell (left), the horizontal bar with sensors and connecting components (centre) and the horizontal bar end view (right).

The horizontal bar was hollow (34 mm internal diameter, 3 mm wall thickness), minimising excess material and allowing space for pore-pressure sensors. Along the two arm components of the horizontal bar, the sensors were positioned at 12 mm intervals in a 45^o spiral (restricted by spatial compatibility) with four per arm, achieving 360^o rotation across the total length of the bar (Figure 19 (centre)). Honeywell 26PCCFA6G sensors (miniature low pressure sensor series, ~ 100 kPa pressure capacity, fluorosilicone sealed, straight port) were used, primarily for their size. To minimise spatial requirements, each sensor was wired with a length of Mogami four-core (one for each of the sensor terminals) 2 mm diameter cable with heat-shrink sealed terminals. Once the sensors were positioned, the drill holes were sealed internally using epoxy resin for water tightness (Figure 19 (right)) and the sensor shafts were cut flush to the T-bar surface. To minimise friction with the test box surfaces, the relevant ends of the arm components were capped with Teflon discs.

A 20 mm deep lap joint with an epoxy resin adhesive secured the arms (female) to a central connector (male). The wall thickness of the central connector increased to 6 mm providing greater joint capacity. A threaded, hollow cylinder screwed the end of the shaft (internally) to the wall of the connector. This provided sufficient space for the sensor wires to pass up through the shaft and a passage for air to maintain atmospheric pressure internally for the sensors.

A load cell was placed in an external recess near the shaft base, above the connector thread, to measure penetration resistance (Figure 19 (left)). This positioning had been selected with the aim of minimising the contribution of shaft friction measured as penetration resistance by the load cell. The cell consisted of four 350 Ω strain gauges evenly spaced in alternating passive (perpendicular to shaft axis) and active (parallel to shaft axis) directions in a Wheatstone Bridge. To avoid wire inlet-hole stress concentrations influencing the measurement, a passive gauge was aligned here. An epoxy resin was used to seal the load cell from the clay, supporting the wire inlet-hole and filling the recess up to the shaft outer diameter, creating a smooth finish. The cell was wired up using the same type of cable as the pressure sensors for spatial functionality. A split clamp secured the top of the shaft onto the table of the $1D$ actuator. The sensor and load cell wires passed out through the top of the shaft.

In response to initial poor signal to noise ratio (SNR) of 4 : 1 in load cell calibration, the recess depth was increased from 1.5 mm to 2.5 mm, reducing wall thickness to 0.5 mm. Theoretically, this provided a threefold increase in signal strength. Conservative estimations (neglecting the contribution of the epoxy resin fill and treating the section as a pin-ended Euler column) confirmed that sufficient structural integrity was maintained for the experiment.

Control Primarily, experimental control is yet to be enacted centrally using a "LabJack T7", a measurement and automation device. Actuator motor control and RP image capture would have been operated from a Python script via the T7, simplifying the experimental process. Additionally, measured data (sensors and load cell) and images would have been transferred to the central operating computer via the T7.

Actuator The T7-controlled 1D actuators, attached to the loading plate, would have driven the T-bars. The loading plate (already constructed) consisted of 20 mm thick aluminium (7.1 kg) with appropriate fixtures for actuator connection and shaft intersection. Ultimately, the loading force will be carried by the piston applying surcharge via the experiment top plate. The shaft holes were lined with double-lipped, rotary shaft seals to minimise risk of binding and prevent clay loss during testing. The T-bar centres were positioned 4 diameters away from the end walls (to ensure a fully visible full-flow mechanism) and 11.5 diameters away from each other (to ensure test separation and space for top plate piston loading).

4.3.4 Image Capture

Trial images were captured by an array of standard RP cameras attached to a configurable frame (seen during use in Figure 21) by 1/4" bolts, as would be intended for the test. All cameras were approximately perpendicular and equidistant to the exposed clay face.

The image areas must have minimal overlap, ensuring continuity while minimising error in the observed internal work. RP cameras (V2) have a sensor resolution of 3280×2464 pixels with a vertical and horizontal FoV of 62.2^o and 48.8^o respectively. Taking into consideration a single camera and fixing the pixel size according to the critical and ideal d/p ratio, the captured area and distance from the surface was estimated (Table 12 and Figure 20) to guide array positioning. The refractive index ($RI = \sin i / \sin r$) of perspex was taken as 1.495.

Table 12: Camera positioning estimation. "H" signifies horizontal and "V", vertical.

d/p	Pixel length, μm	Dimension	Clay length, l_c, m	Perspex length, l_p, m	Distance, z, m
4	62.5	H	0.205	0.120	0.100
		V	0.154	0.069	0.076
2	125	H	0.410	0.344	0.285
		V	0.308	0.242	0.267

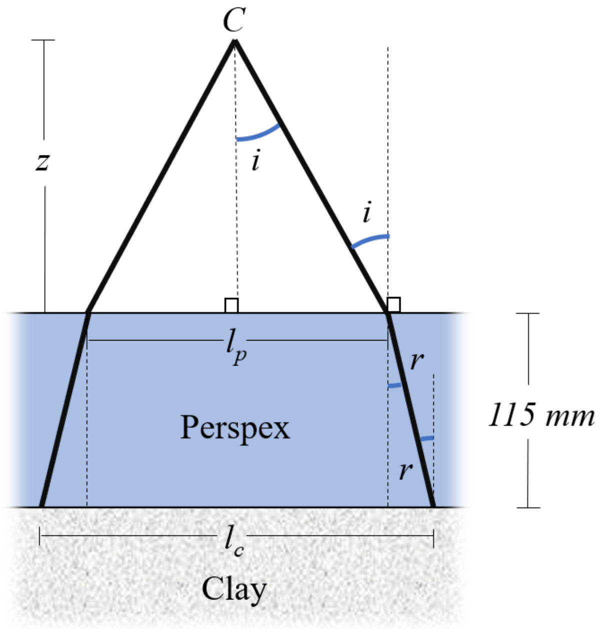

Figure 20: Field of view top-view schematic. "C" indicates the camera position and the thick black lines represent the bounds of the FoV.

The precision of PIV is dependent upon the record of representative pixel intensities. The exposed face must be evenly illuminated to avoid the influence of lighting upon measured intensity, typically achieved through using LED strips or panels. Since reflections create image blind spots, diffusers should be used and the lighting, distributed around the edge of the exposed plane and oriented appropriately.

Several different lighting arrangements were trialled and compared qualitatively (Figure 21). The LED strip produced an uneven illumination and reflection and the LED bar produced a strobing effect in images due to the AC power supply. The best arrangement consisted of two 192 LED panels with DC power supply, one above and one below the window, angled to avoid shadow or reflection in the region of interest.

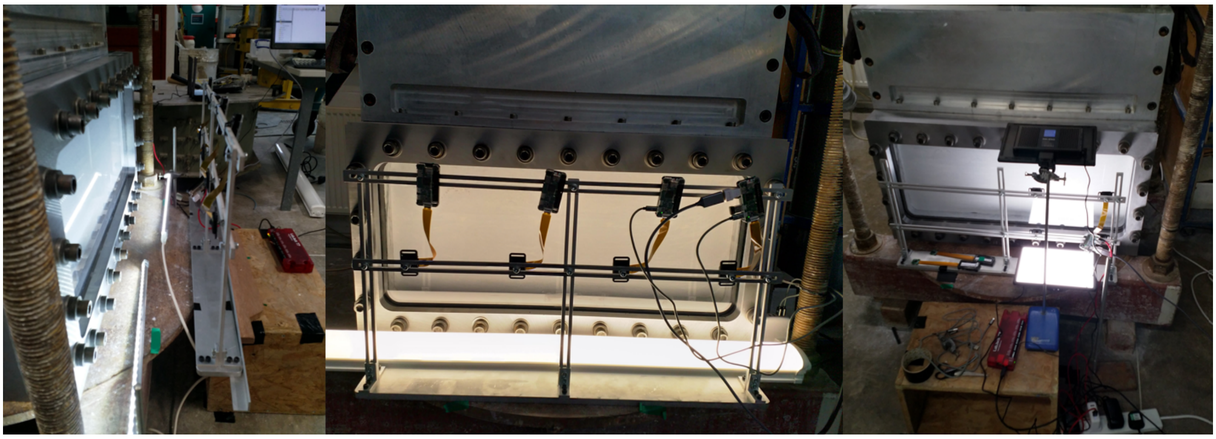

Figure 21: Pre-experiment lighting trials using the camera mounting frame: LED strip (left), LED bar (centre) and LED panel (right).

4.3.5 Image Quality Assessment

Once consolidation was completed, an assessment of image quality was made under experimental lighting conditions using an RP camera positioned across a range of distances (Figure 22). The SSSIG was evaluated within the GeoPIV-RG framework using subset size of 50 and 100 pixel diameters (25 and 30 pixel spacing respectively) in the configuration shown in Figure 21 (right). The results of this analysis in Table 13 indicate a sufficient SSSIG was attained to meet the 1×10^5 threshold with 100 pixel diameter subsets. Additionally, the furthest distance assessed suggested a drop in SSSIG for the 100 pixel diameter subsets, providing a second limit on the required RP set-up.

Figure 22: Exposed plane texture from a SSSIG trial image ($z = 72.5\ mm$).

Distortion would have been accounted for by imaging a Charuco board at the exposed plane under the same experimental set-up. This would have provided a grid of known reference points with which to calibrate images.

Table 13: SSSIG values calculated from images taken of the final exposed plane under lighting at a variety of distances. The SSSIG was calculated for two subset sizes and spacings: 50 pixel size with 25 pixel spacing and 100 pixel size with 30 pixel spacing. Parameter definition can be found in Figure 20.

	SSSIG	
Distance, z, mm	50, 25	100, 30
72.5	3.04×10^4	1.26×10^5
87.5	3.63×10^4	1.27×10^5
107.5	3.57×10^4	1.32×10^5
127.5	3.19×10^4	1.29×10^5
147.5	3.28×10^4	1.30×10^5
167.5	3.12×10^4	1.07×10^5

4.3.6 Full-flow Penetrometer Experimental Process

Upon the completion of sample preparation, the experiment would have been conducted in the following steps:

1. The consolidation surcharge (piston loading) and top plate would have been removed. Once positioned, the T-bars would have initially penetrated the clay through the incremental loading of the experiment top plate to the original consolidation pressure (stress state on the NCL). A foam layer would have been positioned between the horizontal bars and the top plate to ensure minimal uneven bulging of the surface. After initial penetration, the actuators would have driven the T-bars to mid-depth at a rate of 1.5 mm/s ($vd/c_v \sim 6$ to achieve the minima indicated in Figure 13).

2. Final RP camera array configuration, including specification of aperture and exposure, would have been conducted with the T-bar at mid-depth immediately before the experiment. This would have ensured optimal accounting for ambient lighting conditions.

3. Only one T-bar would have been operational at any one time. In the experimental process, a T-bar would have been oscillated 10 times through a total amplitude of 120 mm (3 diameters centred at mid-depth), at a rate of 1.5 mm/s. During this process, measurements (load cell and pore pressure sensors) would have been monitored and photos taken at a rate of 4 frames per second. Since the actuator motor speed, sensor measurements and RP image capture are all controlled by a single script through the T7 system, fully-integrated synchronisation would have been achieved.

4. The excess pore pressure generated in testing would have been allowed to dissipate before the subsequent test began.

5. Image processing would have been conducted according to the standard GeoPIV-RG framework. Calibration would have then been conducted, employing the Bayesian approach, to quantify the experimental uncertainty.

Unfortunately, due to the COVID-19 pandemic, it was not possible to conduct the experiment (Appendix D). The experiment preparation was near completion when COVID-19 restrictions were imposed.

5 Conclusion

This project has explored testing methods for the strain-softening parameter calibration technique proposed by Singh and Stanier (unpublished). In the proposed approach, the magnitude and rate of sensitivity degradation within a SMCC constitutive model are calibrated through an equating of internal and external work.

The enhancement of the approach, namely the inclusion of Bayesian inference, proved fruitful, enabling a richer appreciation of the uncertainty associated with the calibrated parameters. Instead of producing singular "best-fit" parameters, the method described a high-likelihood region. This provides a basis for more robust and informed parameter choice, allowing the consideration of a range of potential strain-softening behaviours. Additionally, comparison of the shape and size of the region produced evidence for improvements to the calibration process by identifying factors that reduce parameter uncertainty.

The cyclic displacement of a T-bar penetrometer, fully remoulding the soil to its residual state, partnered with PIV techniques, enabling high-precision full-field displacement measurements, was to be trialled as an alternative parameter calibration method. While restrictions imposed during the COVID-19 pandemic (Appendix D) prevented the experiment from being conducted, the test plan and the necessary equipment have been designed and produced, ready for immediate further work when restrictions allow [1]. If successful, the proposed technique would gain experimental support, establishing it as a viable, practicable alternative.

The implementation of the proposed framework would remove dependence on the problematic assumptions that uphold conventional calibration methods. If the calibration can be conducted successfully, the limiting factor to design for strain-softening would become the fit of the constitutive model itself.

[1] This project will be continued in the commencement of a PhD.

6 Bibliography

Bažant, Z.P. Belytschko, T.B. and Chang, T.P. (1994). Continuum model for strain-softening. Journal of Engineering Mechanics, 110(12):1666–1692.

Bažant ZP, Pijaudier-Cabot G. (1988) Nonlocal continuum damage, localization instability and convergence. J Appl Mech;55(2):287–93.

Bemben, S. M., and Myers, H. J. (1974). "The influence of rate of penetration on static cone resistance in Connecticut River Valley varved clay." Proc., European Symp. on Penetration Testing (ESOPT), National Swedish Building Research, Stockholm, Sweden, 33–34.

Bernardo, J.M., Smith, A.F.M. (1994) Bayesian Theory, Wiley.

De Borst R, Sluys LJ, Mühlhaus H-B, Pamin J. (1993) Fundamental issues in finite element analyses of localization of deformation. Eng Comput;10(2):99–121.

Butterfield, R. (1979). A natural compression law for soils (an advance in e-log p') Géotechnique 1979 29:4, 469-480.

Charles, J.A., Smith C.C. and Black J.A. (2018). Identification of soil stress-strain response from full field displacement measurements in plane strain model tests. Physical Modelling in Geotechnics, London, 835–840.

Chung, S. F., Randolph, M. F., and Schneider, J. A. (2006). "Effect of penetration rate on penetrometer resistance in clay." J. Geotech. Geoenviron. Eng., 132(9), 1188–1196.

DeJong, J. T., Yafrate, N. J., and DeGroot, D. J.(2011) Evaluation of Undrained Shear Strength Using Full-Flow Penetrometers Journal of Geotechnical and Geoenvironmental Engineering, Vol. 137, No. 1, January 1, 2011.

Desrues J, Viggiani G. (2004) Strain localization in sand: an overview of the experimental results obtained in Grenoble using stereophotogrammetry. International Journal Numerical Analysis Methods Geomechanics;28(4):279–321.

Einav, I., and Randolph, M. F. (2005). "Combining upper bound and strain path method for evaluating penetration resistance." Int. J. Numer. Methods Eng., 63(14), 1991–2016.

Galavi, V. and Schweiger, H.F. (2010) Nonlocal multilaminate model for strain softening analysis. International Journal of Geomechanics, 10(1):30–44.

Ganesan, S.A. and Bolton, M.D. (2013). Characterisation of a high plasticity marine clay using a T-bar penetrometer. Underwater Technology, Vol.31, No.4, pp.179-185.

Gelfand, A.E. and Smith, A.F.M. (1990), "Sampling Based Approaches to Calculating Marginal Densities." Journal of the American Statistical Association, 85, 398-409.

Geman, S. and Geman, D. (1984), "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, 6, 721-741.

Georgiannou VN., Burland, JB., (2006) A laboratory study of slip surface formation in an intact natural stiff clay. Géotechnique 2006;56(8):551-9.

Geyer, C.J. (1992) Practical Markov Chain Monte Carlo. Statistical Science, 7,473–483.

Girolami, M.A. (2008) Bayesian inference for differential equations. Theoretical Computer Science 408;4-16.

Grédiac, M. & Pierron, F. (1998). A T-shaped specimen for the direct characterization of orthotropic materials. International Journal for Numerical Methods in Engineering, 41(September 1996), 293–309.

Gueguin, M., Smith C. and Gilbert, M. (2015). Use of digital image correlation to directly derive soil stress-strain response from physical model test data, Proceedings of the XVI ECSMGE Geotechnical Engineering for Infrastructure and Development, 3881–3886.

Gylland, A.S., Jostad, H.P. & Nordal, S. (2014). Experimental study of strain localization in sensitive clays. Acta Geotechnica, 9: 227-240.

Hastings, W.K. (1970) Monte Carlo sampling methods using Markov chains, and their applications, Biometrika 57:97-109.

Heikkila, J. and Silven, O. (1997) A four-step camera calibration procedure with implicit image correction. Proc. IEEE Computer Soc. Conf. Computer Vision and Pattern Recognition (CVPR'97), San Juan, 1106–1112.

Kass, R.E., Raftery, A.E. (1995) Bayes factors. Journal of the American Statistical Association 90(430); 773-795.

Lewis,S.M. and Raftery,A.E. (1997) Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator. J. Am. Stat. Assoc., 92, 648–655.

Link, W. A. and Eaton, M. J. (2011) On thinning of chains in MCMC. Methods in Ecology and Evolution. First published online 17 June 2011. doi: 10.1111/j.2041-210X.2011.00131.x

MacEachern, S.N. and Berliner, L.M. (1994) Subsampling the Gibbs sampler.The American Statistician, 48, 188–190.

Mánica, M.A., Gens, A., Vaunat, J., Ruiz, D.F. (2018) Nonlocal plasticity modelling of strain localisation in stiff clay. Computers and Geotechnics 103;138-150.

Mašín, D. (2009). Comparison of predictive capabilities of selected elasto-plastic and hypoplastic models for structured clays. Soils and Foundations, 49(3):381–390.

McCulloch,R.E. and Rossi,P.E. (1991) Bayes factors for nonlinear hypotheses and likelihood distributions. Technical Report 101. Statistics Research Center, University of Chicago, Graduate School of Business.

Pan, B., Xie H., Wang Z., Qian, K, and Wang, Z. (2008) Study on subset size selection in digital image correlation for speckle patterns. Opt Express 2008;16(10):7037–48.

Pan, B., Li, K., and Tong, W. (2013) Fast, robust and accurate digital image correlation calculation without redundant computations. Experimental Mechanics, 53: 1277–1289.

Ragni, R., Wang, D., Cassidy, M.J., Bienen, B. and Mašín, D. (2015). Modelling the effects of pauses during spudcan penetration on the further installation behaviour, The 15th International Conference: The Jack-Up Platform: Design, Construction and Operation.

Randolph, M. F. (2004). "Characterisation of soft sediments for offshore applications." Proc., ISC-2: 2nd Int. Conf. on Site Characterization, V. da Fonseca and P. Mayne, eds., Millpress, Rotterdam, The Netherlands, 209–232.

Randolph, M. F., and Andersen, K. H. (2006). "Numerical analysis T-bar penetration in soft clay." Int. J. Geomech., 6(6), 411–420.

Randolph, M., & Hope, S. (2004). Effect of cone velocity on cone resistance and excess pore pressures. In T. Matsui, Y. Tanaka, & M. Mimura (Eds.), Proceedings of the IS Osaka - Engineering Practice and Performance of Soft Deposits (Osaka, Japan ed., Vol. n/a, pp. 147-152). Osaka, Japan: Yodogawa Kogisha Co. Ltd.

Randolph, M.F., Stanier, S.A., O'Loughlin, C.D., Chow, S.H., Bienen, B., Doherty, J.P., Mohr, H., Ragni, R., Schneider, M.A., White, D.J. and Schneider, J.A. (2018). Penetrometer equipment and testing techniques for offshore design of foundations, anchors and pipelines. Cone Penetration Testing, Taylor and Francis Group.

Rogers, S., Girolami, M. (2017) First Course in Machine Learning, 2nd Edition, Taylor & Francis Group, LLC.

Slama, C. C. (ed.) (1980). Manual of photogrammetry, 4th edn. American Society of Photogrammetry, VA, USA.

Sloan, S.W., Abbo, A.J. and Sheng, D. (2001). Refined explicit integration of elastoplastic models with automatic error control. Engineering Computations, 18(1): 121-194.

Stanier, S.A., and White, D.J. (2013) Improved image-based deformation measurement for the centrifuge experiment. Geotechnical Testing Journal. 36(6): 915-928.

Stanier, S.A., Blaber, J., Take, W.A. and White, D.J. (2016). Improved image-based deformation measurement for geotechnical applications. Canadian Geotechnical Journal, 53(5): 727–739.

Stanier, S.A., Dijkstra, J., Leśniewska, D., Hambleton, J., White, D. and Wood, D.M. (2016). Vermiculate artefacts in image analysis of granular materials. Computers and Geotechnics, 72: 100-113.

Stewart, D. P., and Randolph, M. F. (1991). "A new site investigation tool for the centrifuge." Proc. Centrifuge, 91, 531–538.

Viggiani, G. and Finno, R.J. (1994) Experimental observations of strain localisation in plane strain compression of a stiff clay. Localisation and Bifurcation Theory for Soils and Rocks, Chambon, Desrues, Vardoulakis (eds) 1994. Balkerna, Rotterdam.

Vyshemirsky, V. and Girolami, M.A. (2008) Bayesian ranking of biochemical system models. Bioinformatics, Vol. 24 no. 6;833-839.

Westerweel, J. (1997) Fundamentals of digital particle image velocimetry. Meas Sci Technol 1997;8(12):1379.

White, D. J., Take, W. A. and Bolton, M. D. (2003). Soil deformation measurement using particle image velocimetry (PIV) and photogrammetry Géotechnique 53, No. 7, 619–631.

Wood, D.M. (2012). Heterogeneity and soil element testing. Géotechnique Letters, 2(4):217–219.

Yafrate, N. J., DeJong, J. T., and DeGroot, D. J. (2007). "The influence of full-flow penetrometer area ratio on undrained strength measurements." Proc., 6th Int. Conf. on Offshore Site Investigation and Geotechnics, Society for Underwater Technology (SUT), London, 461– 468.

Yafrate, N. J., DeJong, J. T., DeGroot, D. J., and Randolph, M. F. (2009). "Evaluation of remolded shear strength and sensitivity of soft clay using full-flow penetrometers." J. Geotech. Geoenviron. Eng., 135(9), 1179–1189.

Zhou, H., and Randolph, M. F. (2009). "Resistance of full-flow penetrometers in rate-dependent and strain-softening clay." Geotechnique, 59(2), 79–86.

Appendices

A MCMC Code

```
"""cg.py
Chain Generator: MCMC parameter optimisation implementing Metropolis-Hastings Algorithm.
Date:27/02/2020, Author: Jonathan Smith
Based on the work of Vikram Singh (numerical models) and Sam Stanier (initial code).
For operation ensure the necessary files are present - contact author for details."""

# Import modules
import numpy as np
import os

# Modeldata.txt input parameters
flag = 2
beginFrame = 1
endstep = 108
soil_width = 0.4
soil_height = 0.6
embedment = 0.26
dim_str = 0.04
thick_str = 0.01
clay_MM = 0.8984
clay_lambda =  0.0811
clay_kappa = 0.019
clay_N_rmld = 1.195
clay_poi_ratio = 0.3
clay_kk_max = 1.0
AA = 0.2
s_f = 1.0
OCR = 1.0
bulk_w = 0.0
void_ini_clay = 1.441
Sep_ini_max = 3.0
chrlnt = 0.01
gamma_c = 0.0
clay_perm_k = 0.0
state_var_no = 15
p_surcharge = 100
youngs_mod_steel = 200000000
poi_ratio_steel = 0.25
W_p = 0.0
disp_single_step = 0.01
step_inc_max = 0.001
step_inc_min = 0.0000001
```

```
step_inc_ini = 0.0001
steptime = 0.1
meshdensity = 0.04
alpha = 0
Vel_d = 1
itransit = 200.0
drainage = 400
gamma_w = 10.0

# Variable input
k_min = 0.0 # k uniform prior lower bound
k_max = 1.0 # k uniform prior upper bound
s_ep_min = 0 # s_ep uniform prior lower bound
s_ep_max = 10 # s_ep uniform prior upper bound
n = 100000 # Final chain length
v_s_ep = 0.5 # Transition distribution variance s_ep
v_k = 0.1 # Transition distribution variance k
end_inc = 54 # Number of frames to consider in analysis (H->54, F->108)
I_IWorkAr = np.zeros([end_inc,n+1]) # Empty internal work array
mu = 0.0 # Additive noise mean
v = 0.0000001 # Additive noise variance (1->1x10^-7, 2->1x10^-6, 3->1x10^-5)
p = '0868_0868' # Prefix for generated chain files ([Dataset]_[initial parameters])
dtog = 1 # Toggle (0/1) high to load existing chains
if dtog == 1: # Load existing chains and set final values to current
        k_c = np.loadtxt(p+'_k_chain.txt',dtype=float,delimiter=',')
        s_ep_c = np.loadtxt(p+'_s_ep_chain.txt',dtype=float,delimiter=',')
        LnL_c = np.loadtxt(p+'_LnL_chain.txt',dtype=float,delimiter=',')
        ra_c = np.loadtxt(p+'_ra_chain.txt',dtype=float,delimiter=',')
        k_n = k_c[-1]
        s_ep_n = s_ep_c[-1]
        ra_n = ra_c[-1]
        a = len(k_c)
        r = int(ra_p*a/(1-ra_p))
else: # Chain initialisation
        k_n = 0.8
        s_ep_n = 6.8
        ra_n = 0
        s_ep_c = np.asarray(s_ep_n)
        k_c = np.asarray(k_n)
        ra_c = np.asarray(ra_n)
        a = 0
        r = 0

# Load External Work Files
EWork=np.loadtxt('External_work.txt',dtype=float,delimiter=',') # Load external work
I_EWork= EWork[0:end_inc,1] # Number of considered frames limited
MaxV_EWork=np.loadtxt('MaxV_External_work.txt',dtype=float) # Load maximum external work

```

```
# Functions
def run_integration(s_ep, k):
        """InternalWork.exe operation: sets up modeldata.txt with variable input,
        runs full integration for the internal work and calculates the log likelihood
        of the input parameter set."""
        parameters = np.asarray([flag, beginFrame, endstep, soil_width, soil_height,
        embedment, dim_str, thick_str, clay_MM, clay_lambda, clay_kappa, clay_N_rmld,
        clay_poi_ratio, k, AA, s_f, OCR, bulk_w, void_ini_clay, s_ep, chrlnt, gamma_c,
        clay_perm_k, state_var_no, p_surcharge, youngs_mod_steel, poi_ratio_steel, W_p,
        disp_single_step, step_inc_max, step_inc_min, step_inc_ini, steptime, meshdensity,
        alpha, Vel_d, itransit, drainage, gamma_w])
        np.savetxt('modeldata.txt', np.transpose(parameters)) # Update 'modeldata.txt' file
        os.system('InternalWork.exe') # Run internal work executable
        IWork = np.loadtxt('Internal_work.txt',dtype=float,delimiter=',') # Access output
        er = np.random.normal(mu, v, len(IWork)) # Generate additive noise
        I_IWork = IWork[:,1] + er # Apply additive noise
        I_IWork = I_IWork[0:end_inc] # Number of considered frames limited
        N = len(I_IWork)
        I_IWork = I_IWork/np.max(I_IWork)
        LnL = -0.5*N*np.log(2*np.pi*v)-0.5*N*v**-2*np.sum((I_EWork-I_IWork)**2) # Log likelihood
        return LnL, I_IWork

def file_record():
        """Opens/clears chain txt files and saves current chains to them."""
        file = open(p+'_k_chain.txt', 'w+')
        file.close()
        file = open(p+'_s_ep_chain.txt', 'w+')
        file.close()
        file = open(p+'_LnL_chain.txt', 'w+')
        file.close()
        file = open(p+'_ra_chain.txt', 'w+')
        file.close()
        np.savetxt(p+'_k_chain.txt', np.transpose(k_c))
        np.savetxt(p+'_s_ep_chain.txt', np.transpose(s_ep_c))
        np.savetxt(p+'_LnL_chain.txt', np.transpose(LnL_c))
        np.savetxt(p+'_ra_chain.txt', np.transpose(ra_c))

#MCMC
i = a + r # Iteration counter
LnL_n, I_IWork_n = run_integration(s_ep_n, k_n) # Initial internal work evaluation
I_IWorkAr[:,0] = I_IWork_n
LnL_c = np.asarray(-1/LnL_n)
while a < n: # Initiate chain loop
        tog = 0 # Reset toggle for new iteration
        i += 1 # Update iteration counter
        k_star = np.random.normal(loc=k_n, scale=v_k) # Generate proposed k value
        s_ep_star = np.random.normal(loc=s_ep_n, scale=v_s_ep) # Generate proposed s_ep value
        if k_star < k_min or k_star > k_max or s_ep_star < s_ep_min or s_ep_star > s_ep_max:
```

```
            r += 1 # Update rejection counter if outside uniform prior
            continue # Continue to next iteration
        LnL_star, I_IWork_star = run_integration(s_ep_star, k_star) # Proposed parameter set
        Hr = LnL_n/LnL_star # Acceptance probability - transition ratio = 1, log inversion
        if 1 < Hr: # Proposal accepted
            tog = 1 # Update toggle
            a += 1 # Update acceptance counter
        elif np.random.uniform(low = 0.0, high = 1.0) <= Hr: # Random variable comparison
            tog = 1 # Update toggle
            a += 1 # Update acceptance counter
        else: # Proposal rejected
            r += 1 # Update rejection counter
        ra = r/(r+a) # Compute rejection ratio
        if tog == 1: # Record iteration values if accepted
            s_ep_n = s_ep_star
            k_n = k_star
            LnL_n = LnL_star
            ra_n = ra
            s_ep_c = np.append(s_ep_c, s_ep_n)
            k_c = np.append(k_c, k_n)
            LnL_c = np.append(LnL_c, -1/LnL_n)
            ra_c = np.append(ra_c, ra_n)
            I_IWorkAr[:, a-1] = I_IWork_star
            print("i: {}, r: {}, a:{}, ra:{}, Hr: {}, k_n: {}, s_ep_n: {}"
                .format(round(i, 3), r, a, round(ra, 3), round(Hr, 3),
                round(k_n, 3), round(s_ep_n, 3)))
        if i%1000==0: # Save the chains at regular intervals
            file_record()
file_record()
```

B Bayesian Inference

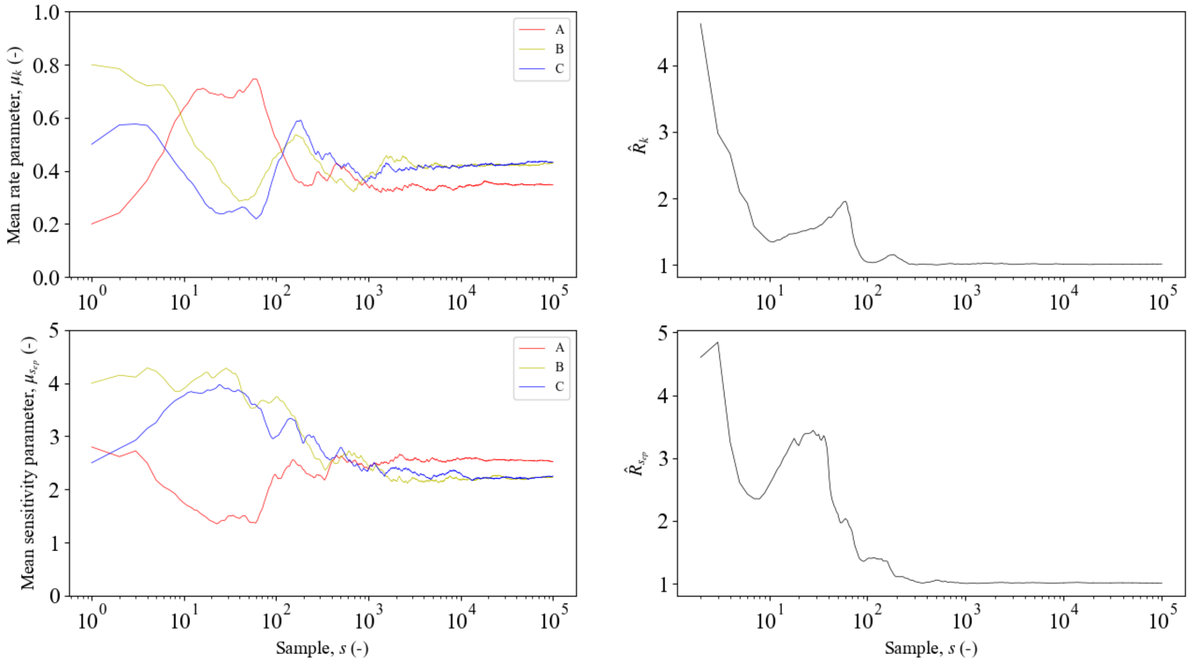

Figure 23: Convergence of multiple chains for dataset B1H2; posterior distribution means (left) and $\hat{R}$ (right) for k (top) and s_{ep} (bottom).

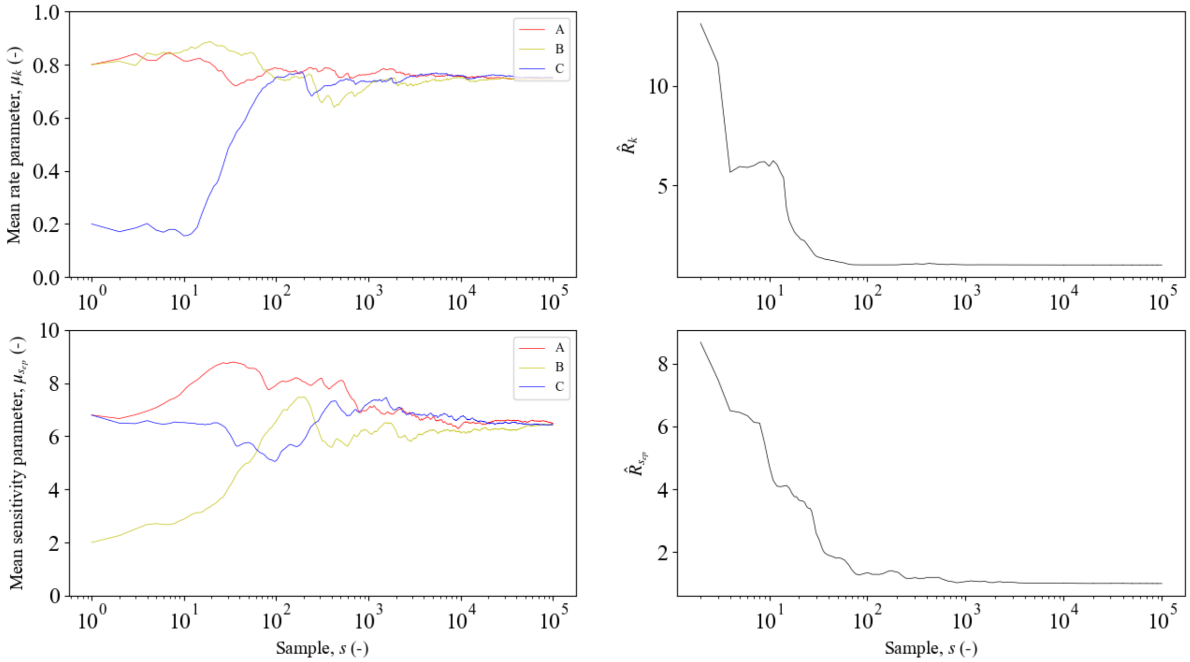

Figure 24: Convergence of multiple chains for dataset B2F2; posterior distribution means (left) and $\hat{R}$ (right) for k (top) and s_{ep} (bottom).

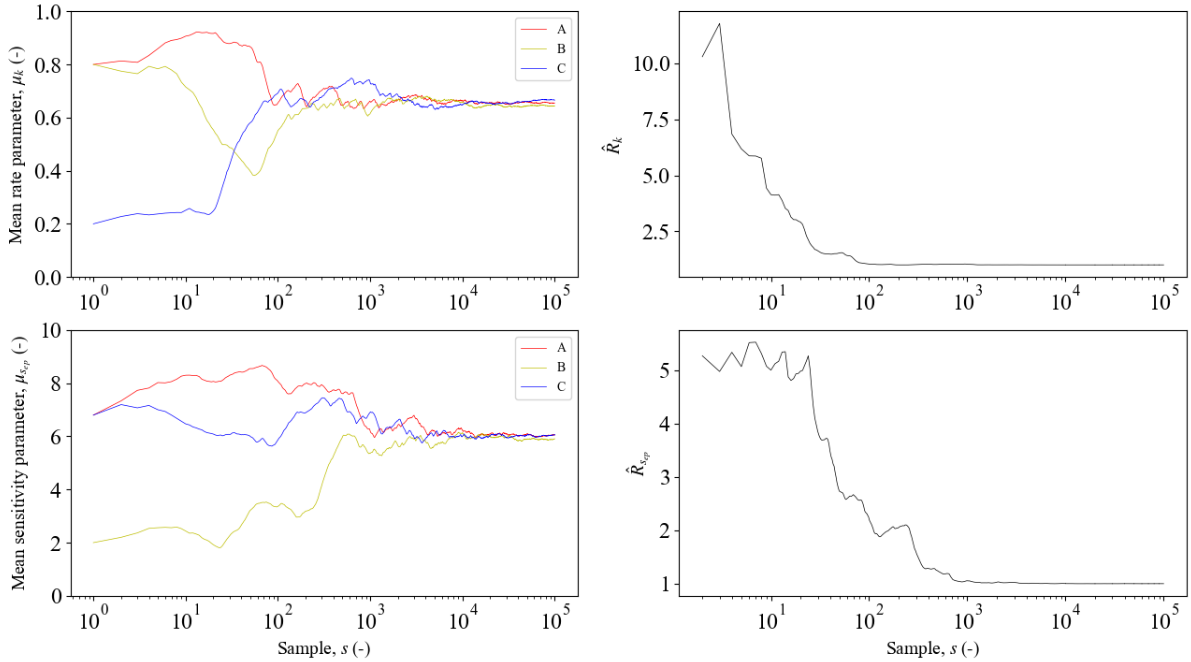

Figure 25: Convergence of multiple chains for dataset B2H2; posterior distribution means (left) and $\hat{R}$ (right) for k (top) and s_{ep} (bottom).

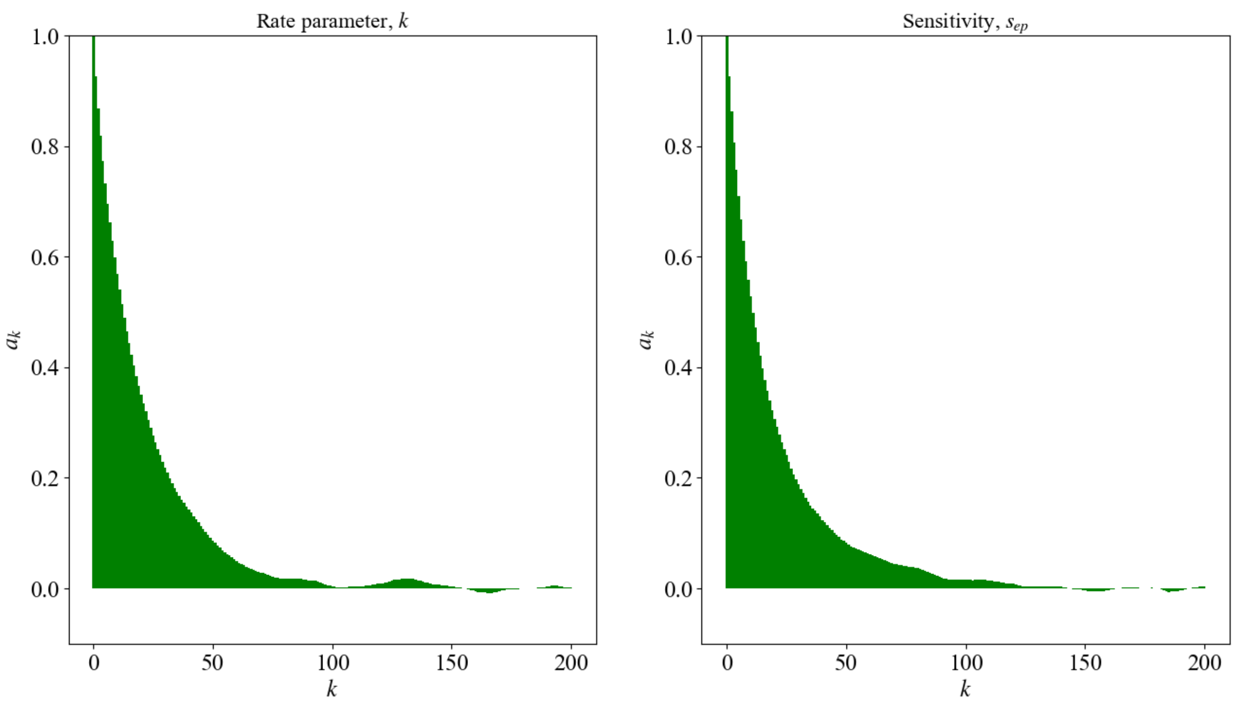

Figure 26: Autocorrelation of B1F2 initiated at $\boldsymbol{\theta} = \{0.2, 2.8\}$ with 1044 burn-in.

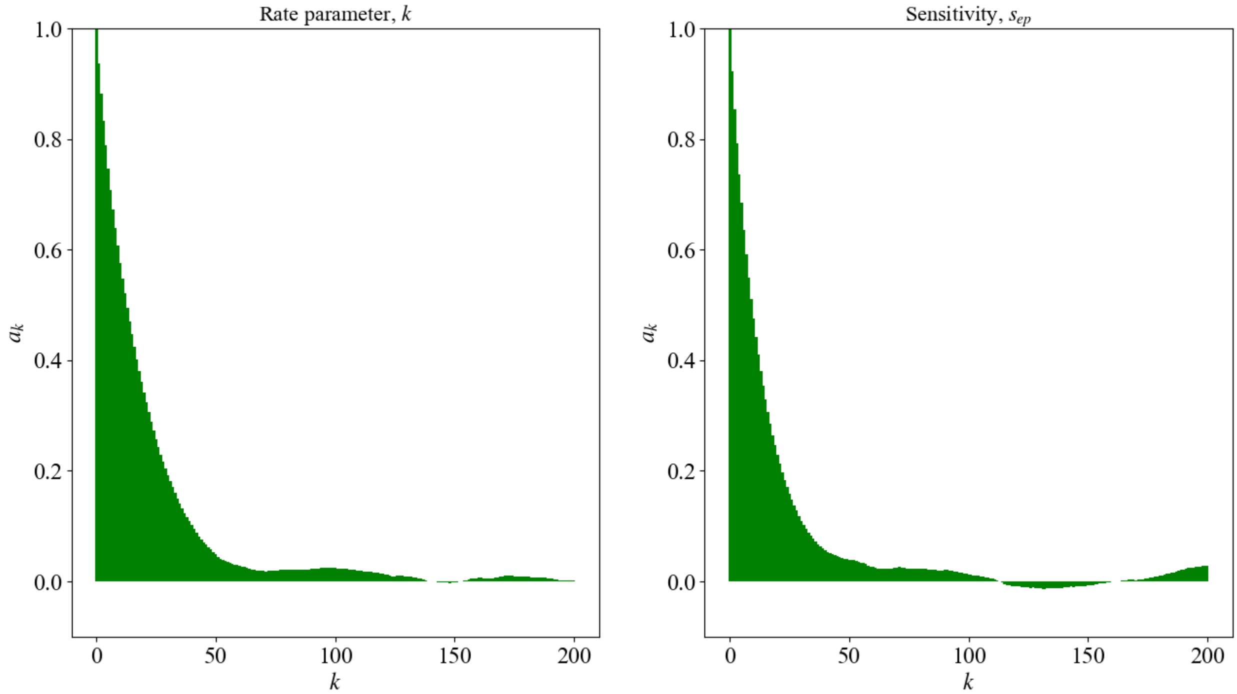

Figure 27: Autocorrelation of B1H2 initiated at $\boldsymbol{\theta} = \{0.2, 2.8\}$ with 876 burn-in.

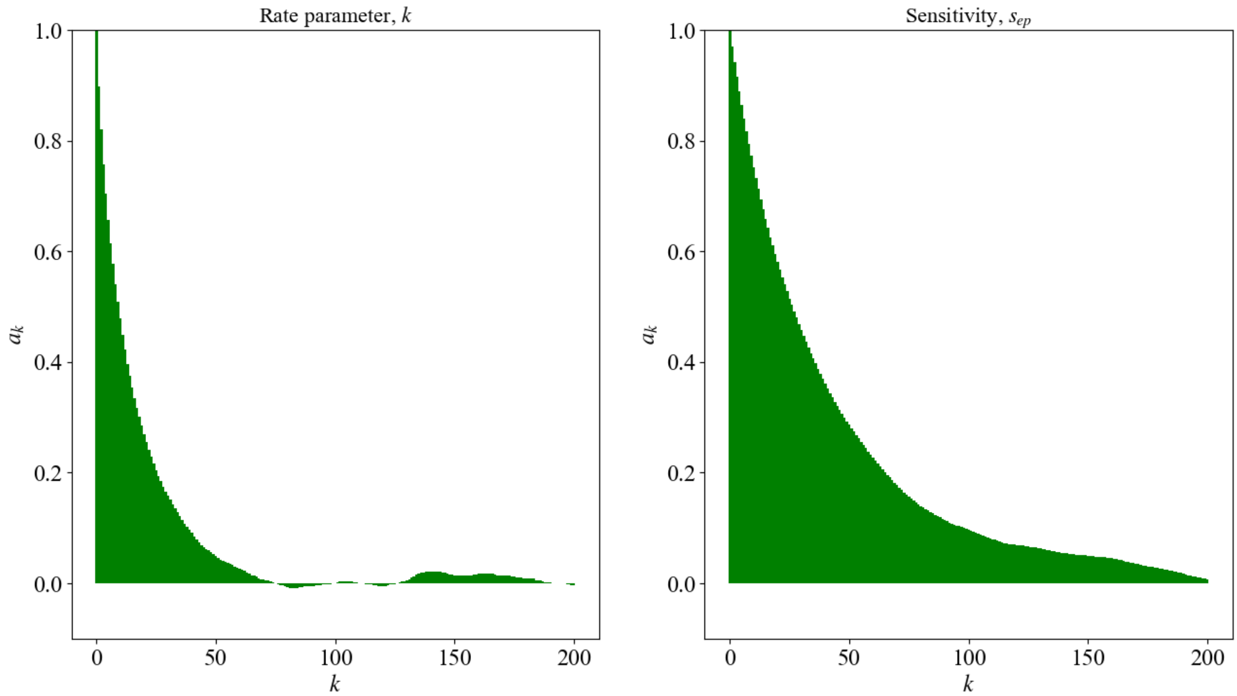

Figure 28: Autocorrelation of B2F2 initiated at $\boldsymbol{\theta} = \{0.8, 6.8\}$ with 7155 burn-in.

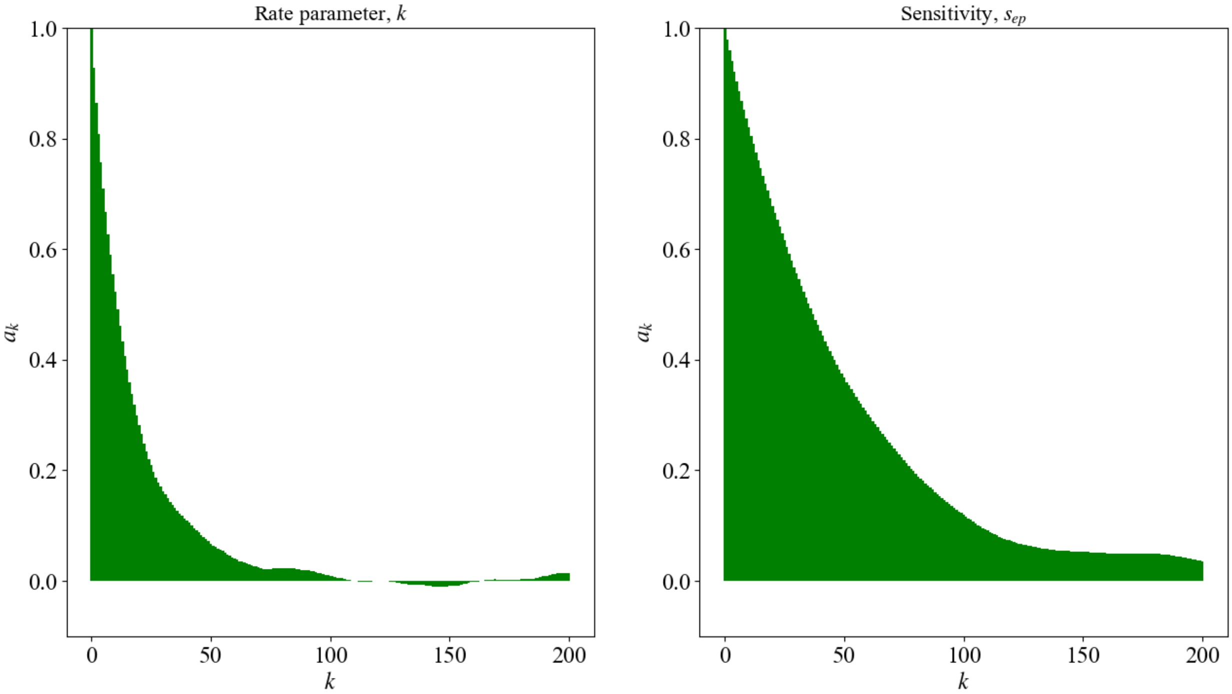

Figure 29: Autocorrelation of B2H2 initiated at $\boldsymbol{\theta} = \{0.8, 6.8\}$ with 3374 burn-in.

C Risk Assessment Retrospective

A COSHH form was submitted along with the risk assessment since the primary project-specific risks were identified as those associated with the handling and safe usage of dry, powdered material (kaolin powder and titanium coated mica). To minimise risk, the use of appropriate personal protective equipment (PPE) was stipulated. During the project, strontium aluminate particles and sand (grade E) were also used. In discussion with my supervisor, continuing with the current level of PPE was deemed acceptable. Were the risk assessment to be completed again, the use of overhead respiratory protection would be substituted with a fitted face mask for user comfort. Additionally, whilst the original assessment was sufficient and suitable for the project, earlier consideration of potential changes could provide greater flexibility.

Other risks were environmental, concerned with laboratory working, or were associated with heavy lifting. The necessary precautions primarily consisted of observance of the general health and safety protocols of the Schofield Centre. These measures were implemented with relative ease.

As a result of the initial risk assessment, the project was conducted safely.

D COVID-19

Due to the restrictions imposed during the COVID-19 pandemic, it was not possible to complete certain aspects of the project. Disruptions included:

Final experiment preparation (19/03/2020-24/03/2020): Final preparation for the planned experiment was not possible since the equipment could not be accessed in the days preceding and during lockdown. This would have required access to the Schofield centre which was closing over these dates. Additionally, the college required all students to leave college accommodation if they were able to by the 19/03/2020.

Experiment (25/03/2020-26/03/2020): It was not possible to conduct the project experiment as the equipment was not accessible in the days preceding and during lockdown. Since no experiment was conducted, no results were produced.

Experiment Analysis (26/03/2020-): In the absence of experimental data, the planned analysis, a large component of the project, could not be conducted.

In light of the disruption, the project scope was adapted. As a result of COVID-19 restrictions, a Bayesian inferential methodology was incorporated into the calibration process to characterise uncertainty and explore the parameter space. This concept had been discussed previously (late Michaelmas term and Lent term) but was considered to be out of project scope based on the available time. The approach was applied to biaxial compression test data numerically produced by Singh (my supervisor's PhD student). Ideally, the approach would have been applied to T-bar test data to replicate the planned experiment numerically, however, this was not possible in the given time frame. Within the available time, the project adaptation involved:

Coding the relevant scripts for the MCMC process in python.

Processing the MCMC processes for several dataset variations.

Analysing the results to draw conclusions for the report.

In this report, the Bayesian inferential approach provided evidence justifying the chosen experiment.