Chemometrics and Intelligent Laboratory Systems 120 (2013) 25–41

Contents lists available at SciVerse ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

# Hierarchical multivariate regression-based sensitivity analysis reveals complex parameter interaction patterns in dynamic models

Kristin Tøndel [a,*], Jon Olav Vik [a], Harald Martens [a], Ulf G. Indahl [a], Nicolas Smith [b], Stig W. Omholt [c]

[a] Centre for Integrative Genetics (CIGENE), Dept. of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway
[b] Department of Biomedical Engineering, Kings College London, The Rayne Institute, Lambeth Wing, St Thomas' Hospital, London SE1 7EH, United Kingdom
[c] CIGENE, Dept. of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P. O. Box 5003, N-1432 Ås, Norway

ARTICLE INFO

*Article history:*
Received 28 March 2012
Received in revised form 18 September 2012
Accepted 26 October 2012
Available online 7 November 2012

*Keywords:*
Sensitivity analysis
Parameter interactions
Hierarchical multivariate regression analysis
Metamodelling
Nonlinear dynamic models
Cluster analysis

ABSTRACT

Dynamic models of biological systems often possess complex and multivariate mappings between input parameters and output state variables, posing challenges for comprehensive sensitivity analysis across the biologically relevant parameter space. In particular, more efficient and robust ways to obtain a solid understanding of how the sensitivity to each parameter depends on the values of the other parameters are sorely needed.

We report a new methodology for global sensitivity analysis based on Hierarchical Cluster-based Partial Least Squares Regression (HC-PLSR)-based approximations (metamodelling) of the input–output mappings of dynamic models, which we expect to be generic, efficient and robust, even for systems with highly nonlinear input–output relationships. The two-step HC-PLSR metamodelling automatically separates the observations (here corresponding to different combinations of input parameter values) into groups based on the dynamic model behaviour, then analyses each group separately with Partial Least Squares Regression (PLSR). This produces one global regression model comprising all observations, as well as regional regression models within each group, where the regression coefficients can be used as sensitivity measures. Thereby a more accurate description of complex interactions between inputs to the dynamic model can be revealed through analysis of how a certain level of one input parameter affects the model sensitivity to other inputs. We illustrate the usefulness of the HC-PLSR approach on a dynamic model of a mouse heart muscle cell, and demonstrate how it reveals interaction patterns of probable biological significance not easily identifiable by a global regression-based sensitivity analysis alone.

Applied for sensitivity analysis of a complex, high-dimensional dynamic model of the mouse heart muscle cell, several interactions between input parameters were identified by the two-step HC-PLSR analysis that could not be detected in the single-step global analysis. Hence, our approach has the potential to reveal new biological insight through the identification of complex parameter interaction patterns. The HC-PLSR metamodel complexity can be adjusted according to the nonlinear complexity of the input–output mapping of the analysed dynamic model through adjustment of the number of regional regression models included. This facilitates sensitivity analysis of dynamic models of varying complexities.



## 1. Introduction

Dynamic models describing complex biological systems, processes or traits are normally rich in input parameters, i.e. quantities that are constant over the time-scale of the particular dynamic model being studied but can be varied between simulations to create variation in model output. In cases where a dynamic model is sensitive to changes in a set of parameters, and the effects of change in one parameter are not dependent on the values of the other parameters, the causal structure of the system is simple, although possibly nonlinear, and the associated sensitivity analysis is relatively trivial across the whole parameter range giving rise to biologically meaningful results. However, for the majority of nonlinear complex dynamic models, the effects of changes in a parameter are often highly dependent on the values of other parameters (the parameters interact), precluding a parameter-by-parameter approach. This situation is likely to become even more pronounced with the emergence of ever more high-resolution, multi-scale dynamic models characterised by high-dimensional input parameter- and output state variable spaces due to improved genomics and phenomics technologies [1].

Means to systematically elucidate the sensitivity features of such dynamic models, including ways to reveal complex interaction patterns between input parameters manifested in high-dimensional

* Corresponding author. Tel.: +47 64 96 52 83; fax: +47 64 96 51 01.
*E-mail address:* kristin.tondel@umb.no (K. Tøndel).

0169-7439 
http://dx.doi.org/10.1016/j.chemolab.2012.10.006

phenotypes, are instrumental for efficient model construction, validation and application. However, many traditional methods for sensitivity analysis are primarily suitable for systems with relatively few input- and output variables and for analysing the effects on only one output at a time [2–5]. A generic sensitivity analysis methodology must be able to handle even the most complex modelling situations, such as highly nonlinear high-dimensional systems [6–8]. Ideally, it should reveal the sensitivity of all dynamic model outputs to any parameter as a function of all other parameters, within the entire operative domain of the analysed model.

In statistical sensitivity analysis, a major branch of the sensitivity analysis field, a selection of data points is derived by experimental design or (semi-) random sampling, and the input–output relations are analysed by statistical methods such as e.g. regression methodology [3] (see Section 5 for more details). In such regression-based sensitivity analysis, the regression coefficients provide direct measures of the impact of the individual inputs on the output (model sensitivity). A major concern is that most regression-based sensitivity analyses published are based on relatively simple linear regression models fitted by ordinary least squares (OLS) regression. Since the input–output relations may be highly nonlinear, linear regression analysis may lead to suboptimal descriptions of dynamic model behaviour, and subsequent difficulties with revealing important interaction patterns. Simple curvature and interaction effects may be modelled successfully by polynomial regression with cross-terms, but when functionally distinct input parameter space regions with clearly different input–output relations and complex interaction patterns between inputs are present, more flexible multivariate analysis methods are needed for a detailed analysis of dynamic model behaviour. Furthermore, most regression-based sensitivity analysis methods are primarily focused on analysing the effects on a single output variable at a time. In many situations it might be advantageous to explore the effects of simultaneous input variation on the whole set of output variables to reveal intricate covariance patterns within both the input- and the output space, in addition to relationships between inputs and outputs. This motivates the development of multivariate *metamodels* [9]; statistical approximations to the input–output mappings of dynamic models that facilitate accurate analysis of their sensitivity features even if these vary substantially across parameter space. In the following, the term “model” refers to the analysed dynamic simulation model if not otherwise specified. Metamodels and regression models are specified as such when discussed.

We recently showed that metamodelling based on Hierarchical Cluster-based Partial Least Squares Regression (HC-PLSR) [10], which involves a combination of global and regional regression analysis, was more accurate than ordinary Partial Least Squares Regression (PLSR) [11] (see also [12,13]) and OLS regression for a range of nonlinear dynamic gene regulatory and physiological models. See Supplementary electronic material: Appendix A for a description of these data analysis methods. In general, PLSR is more effective than OLS for handling multiple output variables simultaneously, since it utilises inter-correlations between the response variables for regression model stabilisation, and is therefore used in both the global and the regional regression steps of HC-PLSR. Furthermore, in contrast to OLS, the PLSR does not require linear independency of the input parameters. In multivariate metamodelling of complex dynamic models this is an advantage, since in cases where the number of input variables is large, highly reduced experimental designs or random sampling must often be used to set up the parameter value combinations for the computational experiment, leading to potentially linearly dependent inputs. For some dynamic models the simulations may also fail to converge under certain conditions, leading to non-orthogonal inputs to the metamodelling. The usefulness of global PLSR for sensitivity analysis was recently demonstrated by Sobie [14] and Martens *et al.* [15]. As a nonlinear extension of PLSR, the HC-PLSR separates the input- or output space into local regions based on clustering in an initial global metamodel. Thereafter a regional metamodel is fitted for each cluster. This allows a simpler description of highly nonlinear effects of input parameters, e.g. causing output variations that may apply only in parts of the input space. The HC-PLSR provides a semi-parametric representation of complex interaction patterns that allows e.g. non-monotone parameter-to-phenotype maps to be modelled more accurately.

Here we introduce a flexible and generic methodology for global sensitivity analysis of complex dynamic models. It is based on the two-step HC-PLSR, and can reveal complex, regional interaction patterns between inputs in a multi-dimensional output setting. Both the global and regional regression modelling steps in HC-PLSR provide scores, loadings, regression coefficients and residual matrices that reflect the sensitivity of the dynamic model to variations in the different inputs. Whereas the initial, global regression model provides an overall summary, the subsequent regional regression models can detail input–output relations that pertain only to parts of the dataset. Hence, modifications of the effects of certain parameters on the dynamic model output dependent on the values of other parameters (reflecting complex parameter interactions) can be identified. Furthermore, the regional sensitivity analysis provides the opportunity to test whether a parameter showing little impact on the output in a global sensitivity analysis still has some impact in local regions of the biologically relevant parameter space.

We illustrate our approach using a detailed model of a mouse heart muscle cell (a ventricular myocyte) [16], primarily built to account for the action potential (the time-course of transmembrane voltage, i.e. the cell's electrical signal) and calcium transient (the time-course of the calcium concentration in the cell fluid, which is linked to muscle contraction) of the cell. These are modelled in terms of a large number of constituent ion currents and voltage- and calcium-sensitive ion channels in the cell, represented by a set of coupled ordinary differential equations (ODEs). Our hypothesis was that highly nonlinear dynamic models, like the model analysed here, will exhibit complex interaction effects that are not identifiable using a global regression-based sensitivity analysis alone. The present analysis includes a wide variety of phenotypic measures (outputs from ODE model simulations) related to the action potential (AP), the calcium transient (CT) as well as the dynamics of a range of other state variables (including ion concentrations in the cell fluid and various cellular compartments, and the state distributions of ion channels, whose transition rates between open, closed, and inactivated conformations may depend on transmembrane voltage and calcium concentration). Several auxiliary output variables, such as ion currents, whose magnitude is a function of system state, are also included. Complex interaction patterns between input parameters are revealed through analysis of the sensitivity of output variables to changes in the individual model inputs, conditional on the levels of the other input parameters.

Illustrating how our methodology can lead to new biological insights, we provide a detailed biological interpretation of how the model parameters interact with respect to four of the outputs; the AP time-to-peak, the AP duration to 25% repolarisation, the CT time-to-peak and the CT decay rate. We chose to focus on these four outputs since we know that the AP and the CT are key cell-level phenotypes of consequence for tissue and organ function. We compare our results to those obtained by a global PLSR-based sensitivity analysis, and show that additional parameter interactions can be identified by supplementing the global sensitivity analysis with a regional analysis.

A dynamic model from computational biology is used here to illustrate the two-step metamodelling methodology for sensitivity analysis. However, we believe that the method is generic, and that the HC-PLSR is a promising approach as part of a semi-automatic methodological framework. The reason for this is its possibility for automatic adjustment of the number of regional regression models according to the nonlinear complexity of the response surface of the analysed dynamic model.

## 2. Theory

### 2.1. The mouse ventricular myocyte model

The mouse ventricular myocyte (heart muscle cell) model [16] used here describes the flow of ions across the cell membrane, and the resulting difference in electrical potential between the intra- and extra-cellular space (the transmembrane potential). The major ions in the model are calcium, sodium, and potassium. Ion channels are specialised proteins that modulate their conductance to the passage of ions across the cell membrane, opening or closing in response to physiological state such as ionic concentration and/or transmembrane potential. Ion channels differ in their thresholds, as well as in how fast they switch between states. The cell also contains e.g. ion pumps, which expend energy to actively transport ions across membranes, and ion exchangers whose cycling is driven passively by concentration gradients.

Generically speaking, this dynamic model represents a nonlinear, high-dimensional mechanistic model of a complex system that is not yet fully understood, and therefore to be submitted to global sensitivity analysis. More specifically, the model was developed [16] as an extension of that of Bondarenko *et al.* [17], with more realistic calcium handling, better consistency checking by conservation of charge, and detailed re-parameterisation to new experimental data. The state variables of the model include concentrations of sodium, potassium and calcium in the cytosol, calcium concentration in the sarcoplasmic reticulum (SR) and the state distribution of ion channels. Formulated as a system of 36 coupled ODEs, this model provides a comprehensive representation of membrane-bound channels and transporter functions as well as fluxes between the cytosol and intracellular organelles.

The modelled mechanisms can be described briefly as follows: the resting transmembrane potential is negative, largely set by the potassium gradient which is maintained between the intracellular space (where it is high) and the extracellular space (where it is low). Given a sufficient electric impulse (inflow of positive ions, depolarising the membrane), the cell responds by opening more ion channels, increasing the flow of positive ions into the cell. This, in turn through calcium ion binding, triggers the muscle cell contraction proteins, followed by cellular repolarisation due to transport of positive (mainly potassium) ions out of the cell.

In more detail, the first step of the AP is the upstroke (depolarisation) phase, which consists of a large influx of Na-ions (the fast Na current, $i_{Na}$) through specific Na-channels [18]. Following the activation phase, the $Na^+$ permeability rapidly decreases to the resting value through inactivation of Na-channels, but the sudden change in the voltage resulting from the rising phase of the AP activates L-type calcium channels, leading in turn to a relatively slower influx of Ca ($i_{CaL}$). After depolarisation, the membrane potential falls rapidly due to an interplay between several currents (this phase is slower in human heart muscle cells). The major contributors in apical myocytes (the type of mouse ventricular myocytes modelled here) are three outward $K^+$ currents (the rapid transient outward $K^+$ current ($i_{Kto,f}$), the ultrarapidly activating delayed rectifier $K^+$ current ($i_{Kur}$) and the non-inactivating steady-state $K^+$ current ($i_{Kss}$), and $i_{CaL}$ [17]. The outward $K^+$ currents slow the rate of depolarisation and initiate repolarisation, and dominate the depolarising inward current, $i_{CaL}$. The final stage of repolarisation is relatively slow and is controlled by the slower $K^+$ currents (the slow transient outward $K^+$ current ($i_{Kto,s}$), a time-independent (inward rectifying) $K^+$ current ($i_{K1}$), $i_{Kur}$ and $i_{Kss}$), as well as the Na/Ca exchanger ($i_{NaCa}$).

In addition, there are two major ion exchange mechanisms, working to maintain the normal balance of ions inside the cell; The Na/Ca exchanger protein (NCX) and the $Na^+/K^+$-ATPase (Na-pump). NCX removes 1 $Ca^{2+}$ from the inside of the cell in exchange for 3 $Na^+$, while the Na-pump transports 3 $Na^+$ out of the cell for every 2 $K^+$ pumped into the cell. The Na-pump is important for maintaining the relatively high concentration of $K^+$ and the low concentration of $Na^+$ found inside normal cells.

### 2.2. Multivariate analysis methodology

PLSR produces a set of PLS components (PCs), which constitute a sequence of orthogonal linear combinations of the original regressor variables *X* that maximise the explained covariance between *X* and the response variables *Y* (see Supplementary electronic material: Appendix A for a more detailed description of the PLSR methodology). In this particular sense the PCs represent a subspace of the original *X*-variable space (here input parameters) that is most relevant for describing the relationship to the *Y*-variables (here dynamic model outputs). Each PC can be considered as an estimated latent variable (*score vector*), an abstract component defined as a weighted linear combination of the original *X*-variables where the associated coefficients are specified in a so-called *loading vector*. Correlation-loading vectors are scale invariant, and defined as the vectors of correlation coefficients between each PC and the original *X*- or *Y*- variables.

The multi-response PLSR (PLS2), the version of PLSR used in this study, provides the opportunity to calibrate a common PLSR model for many *Y*-variables, utilising the inter-correlations between these response variables for regression model stabilisation, and finally also for selection of a common rank for modelling of all the response variables. This is in contrast to single-response PLSR (PLS1), where one PLSR model is made for each response variable without consideration of the other response variables present. In some situations PLS1 may give better prediction results, but when a large number of related response variables is required for the data analysis, PLS2 is often a more efficient choice.

In Hierarchical Cluster-based Partial Least Squares Regression (HC-PLSR) [10], fuzzy *C*-means (FCM) clustering [19–22] is used to separate the observations into clusters according to a chosen similarity measure, and local PLSR models are calibrated within each cluster. The clustering is done on either the *X*-scores or the predicted *Y*-scores from a global PLSR model calibrated using all observations. HC-PLSR thus produces one global PLSR model based on all observations in the calibration set, and a number of local/regional regression models based on the observations in the different clusters. New observations can then be projected into the global PLSR model and classified into the various clusters (several different options for classification exist in the current HC-PLSR implementation). Predictions for new observations can be done either by choosing the most probable (according to the classification) regional regression model for prediction of the response for each new observation, or by using a weighted sum of the predictions obtained from each of the regional regression models, where the estimated cluster membership probabilities from the classification are used as weights. A flow-chart of the HC-PLSR algorithm can be found in Supplementary electronic material: Appendix A, Fig. A.1.

Different clustering methods may be used in the HC-PLSR. We have found the fuzzy *C*-means (FCM) clustering very useful. In FCM clustering, a membership $u_{ij}$ is defined for each object *i* and cluster *j*. The membership values are between 0 and 1, and must sum up to one for each object *i*. In FCM clustering the membership values are found by minimising Eq. (1).

$$J = \sum_{j=1}^{C}\sum_{i=1}^{N} u_{ij}^{m} d_{ij}^{2}, m \geq 1 \text{ subject to} \sum_{i=1}^{N} u_{ij} = 1 \tag{1}$$

Here $d_{ij}$ is the Euclidean distance between object *i* and cluster *j* ($i = 1,2,\ldots,N$, $j = 1,2,\ldots,C$), *m* is a fuzzifier parameter that usually is set to be equal to 2.0. With $m = 1$, FCM is the same as *K*-means clustering. The FCM algorithm can be described as follows: the number of clusters, *C*, is chosen by the user. The estimation procedure is

then initialised randomly for the cluster membership values $U=\{u_{ij}\}$. Then, an iterative procedure for minimising criterion $J$ is started, each iteration consisting of two steps: First, $J$ is minimised for the given memberships $U=\{u_{ij}\}$ by setting the cluster centres $v_j$ equal to the "fuzzy means" (i.e. the weighted averages, see Eq. (2)), and the object-to-cluster-mean distances $D=\{d_{ij}\}$ are computed. Secondly, the membership values $U=\{u_{ij}\}$ are calculated from the given distances $D=\{d_{ij}\}$ using Eq. (3). Thereby the memberships $U$, the set of cluster centres $v_j$ and the distances $D$ are updated in each iteration. The procedure continues until convergence.

$$v_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{2}$$

$$u_{ij} = \left( \sum_{k=1}^{C} \left( \frac{d_{ij}^2}{d_{ik}^2} \right)^{\frac{1}{m-1}} \right)^{-1} \tag{3}$$

This basic FCM algorithm seeks spherical clusters. To find clusters with other shapes, modifications of the FCM algorithm must be applied, see for instance [23,24].

## 3. Materials and Methods

### 3.1. Generation of the in silico data set

Data for the mouse heart muscle cell function was generated using the dynamic model recently published by Li *et al.* [16]. In the non-pacemaker cells of an intact heart, the action potential is initialised by an electrical stimulus from neighbouring cells. In isolated cells, the stimulus is mimicked by briefly applying an electrical current ( "pacing" the cell) at regular intervals. In the myocyte model, the stimulus control is represented as a term called $i_{Stim}$ in the differential equation for transmembrane voltage, $V$. Specifically, the stimulus current gives a positive contribution of 15 mV/ms to $dV/dt$ that lasts 3 ms and is applied every *stim.period* (a specified parameter) ms. In our simulations, *stim.period* was either 166.67, 333.33 or 500 ms (see below). The numerical simulations were carried out using the CVODE algorithm [25,26] with adaptive time-steps, scripted using in-house Python code, available on request. A newer version of the Python code is available at http://github.com/jonovik/cgptoolbox.

Ten different parameters (see Table 1) were varied in a full factorial design (FFD) with three levels of each parameter (baseline value ± 50%), resulting in 59,049 simulations. Hence, all possible combinations of parameter values within these three levels were included in the set of simulations. Because the parameter space of the mouse heart cell model has hitherto been explored to little extent, we designed for screening by using only three levels of each parameter. Screening designs are often used to identify which factors are most important. The range for each parameter is given in Table 1. The varied parameters and the cellular mechanisms they control are illustrated in Fig. 1.

For each set of parameter values, regular pacing was applied until the cellular dynamics converged, i.e. until the multivariate trajectory (time series) was virtually identical in successive stimulus intervals (induced "heart beats"). The convergence criterion for each state variable was based on its value at the beginning of each interval and the integral of its trajectory over that interval, both being constant to within a relative tolerance of 0.001. Details of alternans (alternating strong and weak beats) were not pursued, as this would make the format of the phenotypic data heterogeneous and complicate the application of our methodology and the interpretation of the results. Cell dynamics were categorised as "failed", and excluded from the statistical analyses described below, if the cell did not converge to stable dynamics within 10 min. of simulated time. This happened in 11,669 (19.8%) of the 59,049 simulations. The non-converging simulations did not cluster in any particular region of the parameter space. The data set resulting from the heart cell simulations consisted of values of 36 state variables and 83 auxiliary variables (including ion currents that can be monitored and manipulated in patch clamp experiments (see [16]) calculated over 100 time steps each, for the set of 47,380 combinations of values of the ten varied input parameters for which the cell dynamics converged.

The multivariate trajectories that made up the phenotypic outputs were summarised and represented by scalar characteristics as shown in Fig. 2. Action potential and calcium transient statistics included base and peak levels, time to peak, and time to 25%, 50%, 75% and 90% repolarisation/recovery (Fig. 2, left), as well as amplitude (peak minus base) and decay rate (estimated by fitting an exponential decay from 50% to 90% repolarisation/recovery). In addition, for all state variables, ion currents, and other auxiliary variables, we computed the statistics shown in Fig. 2 (right). Finally, we included each state variable's value at the end of the last stimulus interval. This resulted in a total of 1125 aggregated phenotypes.

### 3.2. Redundancy analysis of the aggregated phenotypes

Many of the 1125 aggregated phenotypes were highly correlated, and a subset of them was selected for the sensitivity analysis to reduce redundancy in the phenotypes and thereby simplify the graphical interpretation. Redundancy can be seen directly from the loading plots from PLSR. However, in order to get a reliable and automatic selection of representative phenotypes, the following procedure was used: Each phenotype was used as regressor to explain the other phenotypes using OLS regression, the phenotype explaining the largest portion of the total variance was picked, and the phenotype matrix was deflated with respect to this variable. This procedure was repeated until the cumulative sum of the percent explained variance from the selected variables reached 99.5%. This resulted in 104 selected

**Table 1**
Description and range of the parameters varied in the simulations with the mouse ventricular myocyte model.

| Parameter name | Unit | Description | Minimum value | Baseline value | Maximum value |
|---|---|---|---|---|---|
| *Ko* | uM | Extracellular potassium concentration | 2700 | 5400 | 8100 |
| *Nao* | uM | Extracellular sodium concentration | 67,000 | 134,000 | 201,000 |
| *Cao* | uM | Extracellular calcium concentration | 700 | 1400 | 2100 |
| *stim.period* | ms | Stimulus period | 166.67 | 333.33 | 500 |
| *vmupinit* | uM/ms | Scaling coefficient for calcium reuptake from cytosol to sarcoplasmic reticulum (SR) by SERCA | 0.2530 | 0.5059 | 0.7589 |
| *PCaL* | $ms^{-1}$ | Scaling coefficient for the L-type calcium current | 1.25 | 2.5 | 3.75 |
| *VmaxNCX* | pA/pF | Scaling coefficient for the sodium-calcium exchanger (NCX) current | 1.9695 | 3.9390 | 5.9085 |
| *gNa* | mS/uF | Scaling coefficient for the fast sodium current | 8 | 16 | 24 |
| *gK1* | mS/uF | Scaling coefficient for the time-independent (inward rectifying) potassium current | 0.1750 | 0.35 | 0.5250 |
| *gKr* | mS/uF | Scaling coefficient for the rapid delayed rectifier potassium current | 0.0083 | 0.0165 | 0.0248 |

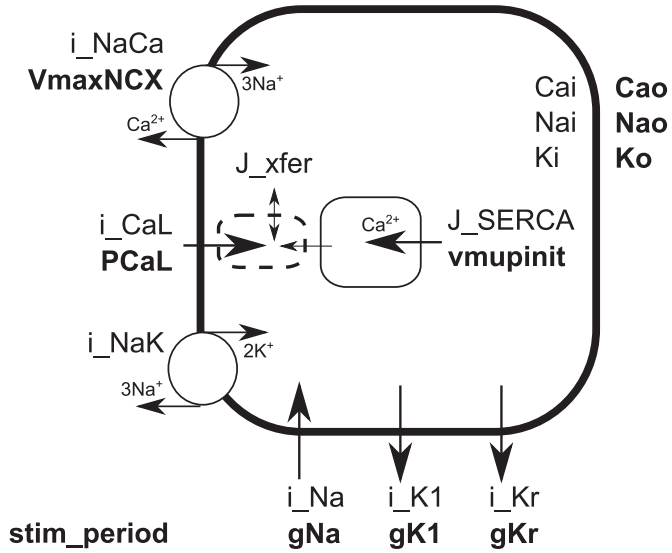


**Fig. 1.** Illustration of the varied input parameters and the cellular mechanisms they control. The ten varied input parameters (in bold face) are illustrated, together with the ion currents and ion channels that they control. *i_NaCa* is the Na/Ca exchange (NCX) current, *i_CaL* is the L-type calcium current, *i_NaK* is the $Na^+/K^+$-ATPase (Na-pump) current, *J_SERCA* is the calcium reuptake from cytosol to sarcoplasmic reticulum (SR) by SERCA (sarco/endoplasmic reticulum $Ca^{2+}$-ATPase), *J_xfer* is the $Ca^{2+}$ flux from the subspace volume to the bulk myoplasm, *i_Na* is the fast $Na^+$ current, *i_K1* is the time-independent (inward rectifying) $K^+$ current and *i_Kr* is the rapid delayed rectifier $K^+$ current, while *Cai*, *Nai* and *Ki* are the intracellular Ca, Na and K-concentrations, respectively.

aggregated phenotypes that were used in the subsequent sensitivity analysis (listed in Supplementary electronic material: Appendix B). How well the selected phenotypes covered the entire phenotype space was tested as described in Supplementary electronic material: Appendix B.

### 3.3. Regression-based sensitivity analysis

HC-PLSR was used here to generate a multivariate metamodel of the mouse ventricular myocyte model, and thereby study the regional differences in model sensitivity to the input parameters. A flow chart of the entire analysis is given in Fig. 3. In HC-PLSR, a global PLSR model is first made, and the observations are thereafter separated in groups based on fuzzy *C*-means (FCM) clustering [19–22] (see Section 2.2) on the PLS scores from the global PLSR model. Regional PLSR models are then made within each group of observations. For sensitivity analysis, the regression coefficients from PLSR are used as sensitivity measures, since when the input parameters are used as regressor variables and the model outputs are used as response variables in the PLSR, the regression coefficients are measures of the effects of variations in the input parameters on the response. Hence, HC-PLSR predicting outputs from inputs provide sensitivity analysis both based on a global PLSR model and regional sensitivity analyses using the regional regression models for comparison. Regional differences in model sensitivity can reveal complex parameter interaction patterns that are difficult to detect using only polynomial regression.

In order to account for nonlinearities both in terms of cross-terms and second order terms (polynomial terms) of the input parameters and in terms of more complex parameter interactions represented by regional differences in the input–output relationships, a second order polynomial HC-PLSR [10] metamodel was made using 67% of the converging observations (the remaining 33% was used for test set prediction). Here the input parameters in Table 1 and their cross-terms and second order terms (in total 65 variables) were used as regressors (*X*) and used to predict the dynamic model output (response, *Y*), represented by the 104 aggregated phenotypes selected in the redundancy analysis described above. Model sensitivity to the input parameters was evaluated using the PLSR-based regression coefficients and the PLS correlation loadings from both the global and the regional regression models produced by the HC-PLSR analysis. The amount of regional differences in model sensitivity to variation in the different input parameters was analysed by inspection of the variation in the PLS regression coefficient values between the four clusters used in the HC-PLSR.

The fuzzy *C*-means clustering of the observations in the HC-PLSR was based on the first three PCs of the *X*-scores (*PC1-PC3*). Using only the first three PCs ensures that only the information in *X* most relevant for the covariance between *X* and *Y* is being used for clustering. The number of clusters was chosen based on inspection of the *X*-scores from the global PLSR model and by comparison of the predictive ability (in terms of explained *Y*-variance) of HC-PLSR models (within the calibration set) using from 1 to 10 clusters. In the selection of the number of clusters to use, the observations in the calibration set were treated as if they were new observations in the prediction stage (i.e. the same procedure as for the test set was used, including a classification prior to the PLSR prediction). For large calibration sets, such as the one used here, this gives the same validation accuracy as using cross-validation (cross-validation would be too time consuming with such a large number of observations).

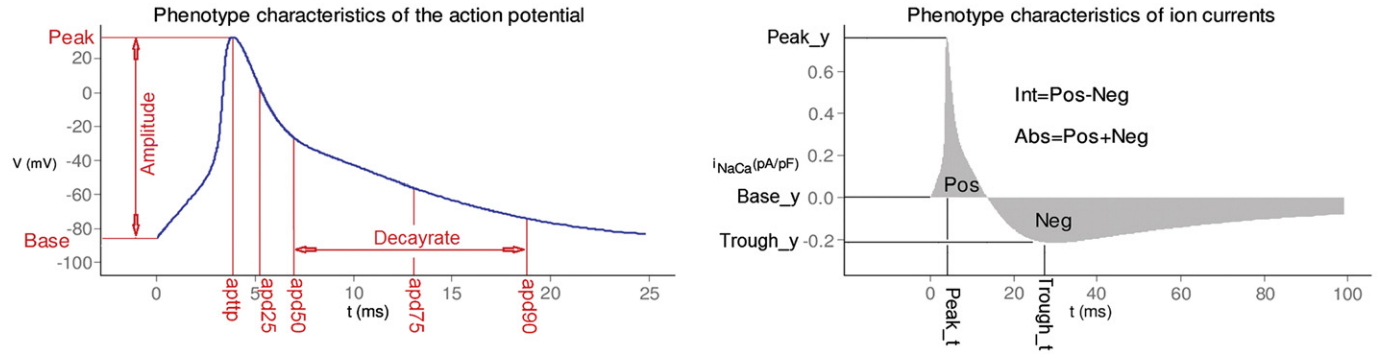


**Fig. 2.** Aggregated phenotypes calculated from the state trajectories and used as outputs in the sensitivity analysis. The figure illustrates the aggregated phenotypes calculated for state trajectories over time (t), exemplified by the action potential (left) and the integrated traffic across ion channels, exemplified by the Na/Ca exchange current $i_{NaCa}$ (right). The aggregated phenotypes illustrated in the right part of the figure were calculated from all state trajectories and auxiliary variable trajectories, and for the action potential and the calcium transient, the time to 25%, 50%, 75% and 90% repolarisation/recovery, the decay rate and the amplitude (peak minus base) were computed (left part of the figure). The decay rates were calculated by fitting an exponential decay from 50% to 90% repolarisation/recovery. *apttp* = action potential time-to-peak and *apd25*, *apd50*, *apd75*, *apd90* = action potential duration to 25, 50, 75 and 90% repolarisation, respectively. *Pos* and *Neg* denote the integrals of positive and negative values of the trajectories, respectively. In total, 1125 aggregated phenotypes were calculated (36 state variables∗10 (the 9 phenotypes in the right part of the figure + the value at the end of the last stimulus interval) + 83 auxiliary variables∗9 (phenotypes in the right part of the figure) + 18 (phenotypes in the left part of the figure calculated for AP and CT)).

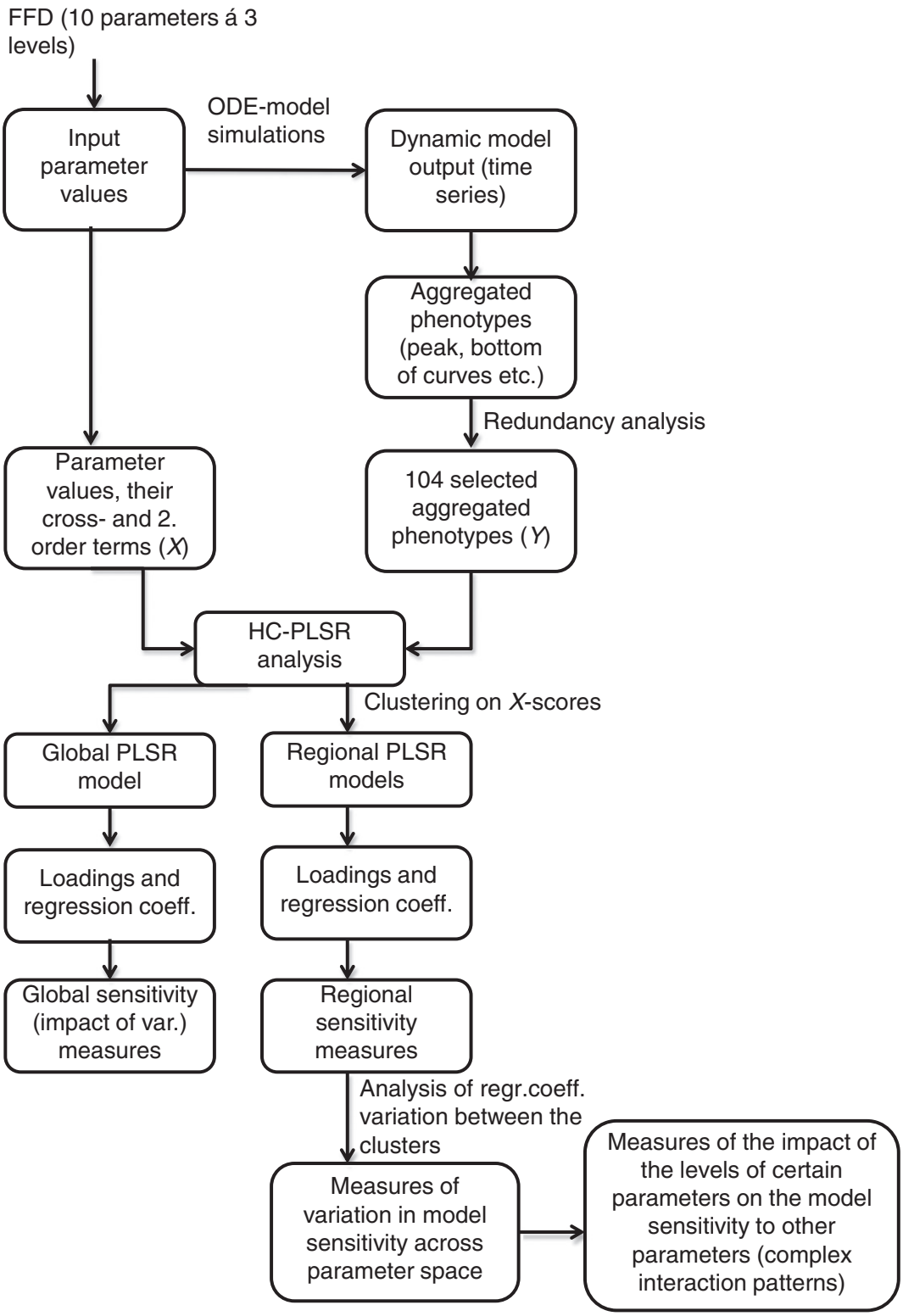


**Fig. 3.** Flow chart of the regression-based sensitivity analysis. First, a full factorial design (FFD) was made varying 10 different input parameters at 3 levels each. The ODE-based mouse ventricular myocyte model was then run for all the $3^{10}$ parameter value combinations, generating model output time series (trajectories). From these trajectories, the aggregated phenotypes illustrated in Fig. 2 were calculated. Based on a redundancy analysis, 104 of these aggregated phenotypes were chosen to be used in the subsequent sensitivity analysis. The parameter values, together with their cross-terms and second order terms were then used as input ($X$) to a HC-PLSR analysis with the 104 aggregated phenotypes as response variables ($Y$). This generated one global PLSR metamodel based on all observations, and four regional PLSR metamodels based on sets of the observations found by fuzzy $C$-means clustering. All these PLSR models generated loading vectors and regression coefficients that formed the basis for the sensitivity analysis. Interactions between the input parameters were represented both by the regression coefficients for the cross-terms between the input parameters in the regression, and by variations in the regression coefficients between the different regional PLSR models.

Using four clusters was considered optimal (the minimum number of clusters giving sufficient predictive ability was chosen in order for the metamodel to be as easily interpretable as possible and to avoid over-fitting). The partition coefficient [19] (a measure of the amount of overlap between the clusters) was also evaluated for fuzzy clustering using from 2 to 10 clusters, and showed decreasing values (high values means low degree of overlap) when more than two clusters was used (See Supplementary electronic material: Appendix C). However, to balance between predictive ability and overlap between the clusters, four clusters was chosen. Based on a classification using the Euclidian distances to all cluster centres for each observation (see [10]), the regional regression model corresponding to the most probable cluster was chosen for HC-PLSR prediction, since this gave better results than using a weighted average of the regional regression models.

The input parameters and phenotypes were auto-scaled (i.e. mean-centred and standardised (divided by their standard deviations)) globally prior to the regression analysis. The cross-terms and second order terms of the input parameters were calculated from globally mean-centred parameter values, prior to the final scaling. In the regional regression models, the matrices of cross-terms and second order terms were deflated with respect to the variation described by the first order terms (in an OLS regression) in order to better separate the effects of nonlinear terms and first order terms.

Both in the global and regional regression analyses the optimal number of PLS components to use was found by 10-fold cross-validation. In 10-fold cross-validation, 10 cross-validation segments are used, and each segment is kept out of the calibration once, and predicted using a PLSR model based on the other observations. This generates ten alternative PLSR models, and the reported regression model is the mean of these ten models. The number of PCs to use was chosen so that each included PC explains at least 1% of the total cross-validated mean squared error in $Y$ (when 0 PCs are used).

To validate the HC-PLSR model for internal consistency over the entire analysed input parameter space, 33% of the observations for which the heart cell simulations converged (randomly selected) were used as an independent test set. Test set validation was used in order to ensure that the predictive ability of the regression model was satisfactory in the entire analysed input space. Test set validation is faster than e.g. cross-validation, and it gives reliable results, provided that the test set is sufficiently large and representative compared to the complexity of the covariation structure of the system to be modelled [27] (as it is here due to random selection among a large number of observations). The Euclidian distance to the cluster centres found in the fuzzy clustering was used to classify the test set observations. The HC-PLSR was carried out in MATLAB® version 7.9.0.529 (R2009b) [28], using in-house code [10] that can be obtained from the authors upon request. The results were plotted in MATLAB® and R [29].

## 4. Results

### 4.1. Sensitivity patterns of the dynamic model revealed by the global metamodel

#### 4.1.1. Overview of input parameter- and interaction effects on the entire set of phenotypes

From extensive, statistically designed mechanistic computations with the mouse ventricular myocyte model, varying ten of the model's input parameters at three levels each (see Table 1) in a full factorial design, we generated two data matrices; the first containing the different parameter value combinations that were used as inputs to the simulations, together with their cross-terms and second order terms (the $X$-matrix), and the other containing the trajectories (time series) of the dynamic model's state- and auxiliary variables resulting from the simulations. To enhance the overview, we computed a set of 1125 aggregated phenotypes from the trajectory data. These are shown schematically in Fig. 2.

Fig. 3 outlines the sensitivity analysis procedure. Since there was a large degree of redundancy in the model outputs, we carried out a redundancy analysis to reduce the number of phenotypes. The sequential algorithm described in Section 3.2 was used for selecting the smallest possible subset of the original set of 1125 phenotypes that was capable of explaining 99.5% of the total variance in the original phenotype set. This subset was chosen in order to further increase the overview and ease the interpretation of the results, and consisted of 104 out of the 1125 aggregated phenotypes. A test of the redundancy analysis (described in Supplementary electronic material: Appendix B) showed that we succeeded in finding a set of aggregated phenotypes that represented most of the variance in the entire set of 1125 phenotypes (see Supplementary electronic material: Appendix B, Fig. B.1). These 104 aggregated phenotypes are listed in Supplementary electronic material: Appendix B, and the data for these phenotypes is in the following referred to as the $Y$-matrix.

The rationale behind utilising as many as 104 different phenotypes in the sensitivity analysis instead of choosing just a few phenotypes of known biological significance on the whole-organ level, was that PLSR utilises inter-correlations between the outputs for regression model stabilisation (for empirical data: against random noise, in this case: against over-fitting the many small nonlinear nuances in the input–output relations). Moreover, utilising a large number of outputs simultaneously can give a more complete overview of the dynamic model behaviour, and can reveal unexpected response patterns. Sobie *et al.* [14,30] has shown that analysing several outputs simultaneously reduces model sloppiness (i.e. many input parameter values leading to the same model output), and leads to more confident conclusions about the relationships between input parameters and model output. However, in order to illustrate the potential of our methodology, we interpret the *biological details* of the sensitivity patterns only for four phenotypes known to be relevant on the whole-organ level (see below).

In order to analyse the relationships between the model input parameters and the generated dynamic model outputs, we first constructed a global metamodel based on a second order global polynomial regression model $Y=f(X)$ using PLSR, with the 104 selected aggregated phenotypes as response variables ($Y$) and the ten input parameters in Table 1 with cross-terms and second order terms as regressors ($X$). The regression coefficients from PLSR are measures of the effect of variations in the input parameters on the various model outputs/aggregated phenotypes, and are therefore useful for sensitivity analysis. Moreover, the regression coefficients for the cross-terms and second order terms indicate interactions between input parameters and nonlinearities predicted by the dynamic model, respectively.

The statistics of the global PLSR model are summarised in Fig. 4 and in Supplementary electronic material: Appendix C. In Fig. 4A the cross-validated (CV) mean squared error (MSE) for the predicted $Y$ is plotted against the number of PCs used in the global PLSR modelling. Since the present data come from noise-free simulations in a deterministic, but highly nonlinear dynamic model, we expected the first few components to be of most relevance, and the minor predictive improvements from the last components to represent nonlinear adjustments. Hence, we chose to use 11 PCs (explaining 57% of the total cross-validated $Y$-variance) as an optimal model rank, in order to balance the ease of interpreting results and the amount of explained variance. The scores for the first three PCs of the global PLSR model are plotted in Fig. 4B, where four clusters of observations (later used in HC-PLSR) are shown. The percentage of the cross-validated $Y$-variance explained by each PC is also indicated. Fig. 5 gives the regression coefficients for the 11-dimensional PLSR model, in terms of the main effects and the second-order interactions and squared effects.

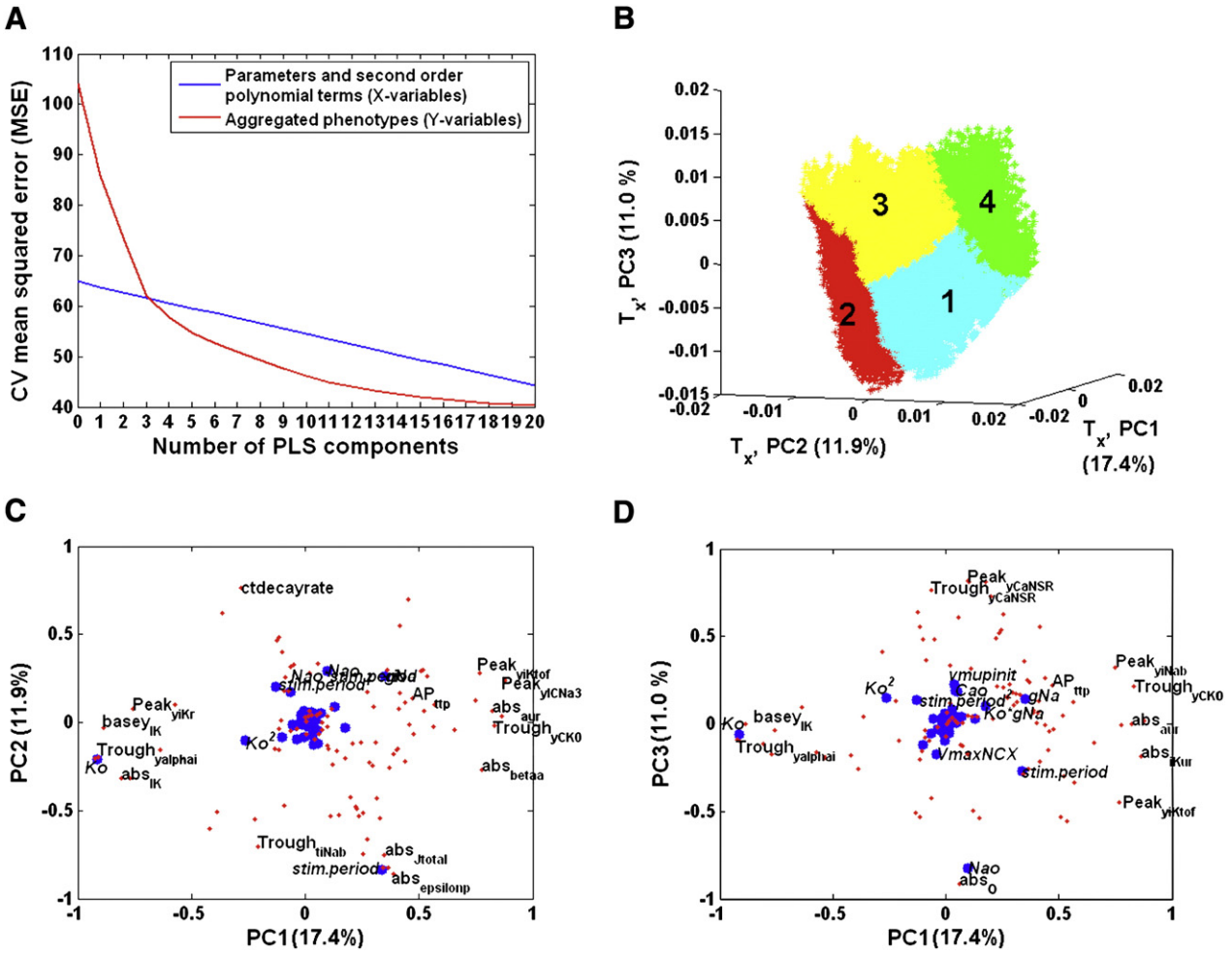


**Fig. 4.** Statistics for the global PLSR model. A) Mean squared error (MSE) of prediction from the 10-fold cross-validation (CV) for *X* and *Y* versus the number of PLS components (PCs) included in the global PLSR of the aggregated phenotypes. *X* is the matrix containing the input parameters and their cross-terms and second order terms, while *Y* is the matrix of the 104 aggregated phenotypes selected in the redundancy analysis. Each PC can be considered as a latent variable (*score vector*) defined as a linear combination of the original *X*-variables where the associated coefficients are specified in a *loading vector*. The minimal number of PCs giving approximately minimal MSE is usually included in the PLSR model. Here, using 11 PCs was considered sufficient in order to balance between predictive ability and metamodel complexity. B) 3D-plot of the first three *X*-score vectors (*PC1-PC3*) from the global PLSR, showing four clusters found by fuzzy clustering of the observations (simulated cells) based on these first three PCs of the *X*-scores. The cells are coloured according to cluster memberships: cluster 1 = cyan, cluster 2 = red, cluster 3 = yellow, cluster 4 = green. This clustering result was later used in the HC-PLSR. The explained CV *Y*-variance is shown in parenthesis for each PC. C–D) Plot of the *X*- and *Y* correlation loadings for *PC1* to *PC3* from the global PLSR. *X*-loadings (representing the input parameters) are shown as blue dots while *Y*-loadings (representing the aggregated phenotypes) are shown in red. Correlation-loading vectors are defined as the vectors of correlations between each PC and the original *X*- or *Y*- variables. Variables placed close to each other in the correlation loading plots are positively correlated, while variables placed opposite each other are negatively correlated. To increase overview, the *X*-variables having low loading-values as well as many of the *Y*-variables were not named in the plot.

Both the correlation loading plots in Fig. 4C and D and the regression coefficients in Fig. 5 showed that extracellular potassium concentration (*Ko*), the scaling coefficient for the fast sodium current (*gNa*), the stimulus period (*stim.period*) and extracellular sodium concentration (*Nao*) were the input parameters having the largest covariance with the 104 aggregated phenotypes used here, and had therefore the largest overall impact on the model output phenotypes (with different signs for different phenotypes). This makes biological sense, as sodium influx is the first stage of the action potential that eventually leads to cardiac muscle cell contraction, and potassium efflux is important for restoring the cell membrane potential to the value at rest [18] (see the description of the dynamic model system in Section 2.1). Increasing the length of the stimulus period gives the cell more time to recover between the action potentials, and will therefore also affect many of the state variables in the myocyte model.

Normal probability plots of the regression coefficients from the global PLSR model (shown in Supplementary electronic material: Appendix C, Fig. C.1) indicated that an approximate significance limit for the global regression coefficients was $\pm 0.2$ for the main effects and $\pm 0.1$ for the interaction effects. We detected significant quadratic effects and global pair-wise interactions between the four parameters *Ko*, *Nao*, *gNa* and *stim.period*, indicating a nonlinear parameter-to-phenotype map. However, cross-terms represent very simplified measures of input parameter interactions, and a corresponding regional sensitivity analysis may therefore reveal more detailed aspects of the interaction patterns.

#### *4.1.2. Input parameter- and interaction effects on key cell-level phenotypes revealed by the global metamodel*

The action potential (AP) and the calcium transient (CT) are key phenotypes of importance for signal propagation and muscle contraction, respectively, and are represented by the trajectories of the state variables *V* and *Cai* in the mouse ventricular myocyte model. These two state variables are known to be especially relevant for the translation from cell to whole-organ models. They were represented by

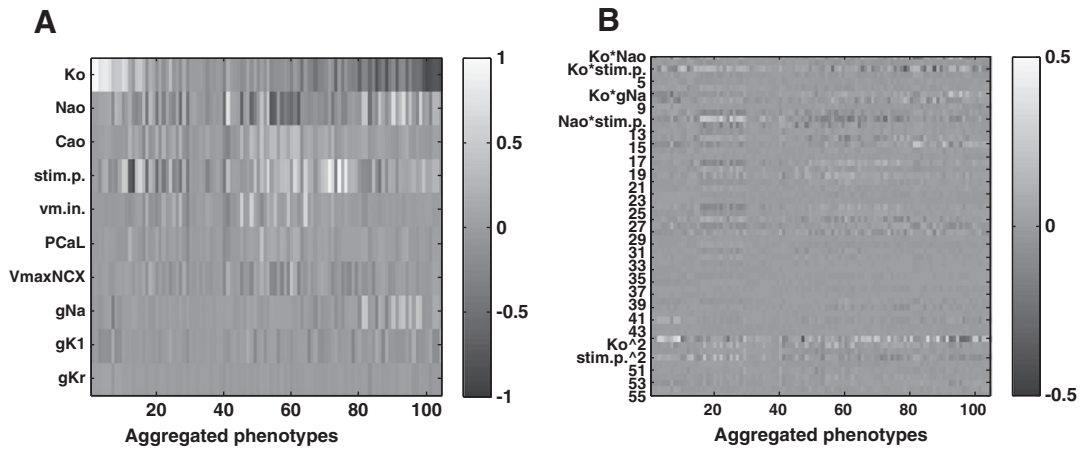


**Fig. 5.** Results from the regression-based sensitivity analysis with global PLSR. A) Regression coefficients of the model input parameters *Ko*, *Nao*, *Cao*, *stim.period*, *vmupinit*, *PCaL*, *VmaxNCX*, *gNa*, *gK1* and *gKr* (stim.p. = *stim.period*, vm.in. = *vmupinit*), for all aggregated phenotypes in the global PLSR model of the aggregated phenotypes as functions of the parameters (using 11 PCs). Each cell in the matrix represents the regression coefficient for a particular input parameter on a given output phenotype, and constitutes a measure of the sensitivity of the given dynamic model output to variations in that input parameter. The phenotypes are sorted according to decreasing loading values of *PC1*. The sorted list of the 104 phenotypes is given in Supplementary electronic material: Appendix B. B) Regression coefficients of the cross-terms and second order terms of the model input parameters for all aggregated phenotypes in the global PLSR model of the aggregated phenotypes (using 11 PCs). The variables 1–9 correspond to the cross-terms *Ko*Nao*, *Ko*Cao*, *Ko*stim.period*, *Ko*vmupinit*, *Ko*PCaL*, *Ko*VmaxNCX*, *Ko*gNa*, *Ko*gK1*, *Ko*gKr*, 10–17 correspond to *Nao*Cao*, *Nao*stim.period*, *Nao*vmupinit*, *Nao*PCaL*, *Nao*VmaxNCX*, *Nao*gNa*, *Nao*gK1*, *Nao*gKr*, 18–24 correspond to *Cao*stim.period*, *Cao*vmupinit*, *Cao*PCaL*, *Cao*VmaxNCX*, *Cao*gNa*, *Cao*gK1*, *Cao*gKr*, 25–30 correspond to *stim.period*vmupinit*, *stim.period*PCaL*, *stim.period*VmaxNCX*, *stim.period*gNa*, *stim.period*gK1*, *stim.period*gKr*, 31–35 correspond to *vmupinit*PCaL*, *vmupinit*VmaxNCX*, *vmupinit*gNa*, *vmupinit*gK1*, *vmupinit*gKr*, 36–39 correspond to *PCaL*VmaxNCX*, *PCaL*gNa*, *PCaL*gK1*, *PCaL*gKr*, 40–42 correspond to *VmaxNCX*gNa*, *VmaxNCX*gK1*, *VmaxNCX*gKr*, 43–44 correspond to *gNa*gK1*, *gNa*gKr*, 45 correspond to *gK1*gKr* and 46–55 correspond to the second order effects $Ko^2$, $Nao^2$, $Cao^2$, $stim.period^2$, $vmupinit^2$, $PCaL^2$, $VmaxNCX^2$, $gNa^2$, $gK1^2$ and $gKr^2$, respectively. The phenotypes are sorted according to the loading values of *PC1* (see Supplementary electronic material: Appendix B).

four out of the 104 aggregated phenotypes selected in the redundancy analysis; the action potential time-to-peak (*apttp*), the action potential duration to 25% repolarisation (*apd25*), the calcium transient time-to-peak (*ctttp*) and the calcium transient decay rate (*ctdecayrate*). The original data set contained also other phenotypes related to the AP and the CT, such as the amplitudes and the integrals under the state variable trajectory curves, but these were not selected in the redundancy analysis (i.e. most of their variation was represented by the selected phenotypes). The global regression coefficients of the parameters and their cross-terms and second order terms for these four aggregated phenotypes are shown in Fig. 6.

Our results showed that the extracellular sodium and potassium concentration and the sodium conductance had the largest effects on *apttp* and *apd25*. The effects of all three parameters were negative (Fig. 6), meaning that the action potential reaches both the peak and 25% repolarisation earlier with increasing values of these parameters. Extracellular Na, extracellular K, extracellular calcium concentration (*Cao*), *PCaL* (scaling coefficient for the L-type calcium current) and *VmaxNCX* (scaling coefficient for the Na/Ca exchanger (NCX) current) had significant negative effects on CT time-to-peak. The following significant interaction effects were identified by the global regression model through analysis of the regression coefficients of the cross-terms between the input parameters in Fig. 6B (see Supplementary electronic material: Appendix C, Fig. C.1 for normal probability plots of all regression coefficients): *Ko*Nao*, *Ko*stim.period*, *Ko*gK1*, *Nao*stim.period* and *stim.period*gK1*, where *gK1* is the scaling coefficient for the inward rectifying potassium current. The following second order terms had large regression coefficients for the analysed phenotypes: $Ko^2$ and $stim.period^2$. These results were all expected since these parameters are important for the key ion currents making up the action potential (see Supplementary electronic material: Appendix C for a more detailed explanation of these results).

Increasing the stimulus period shortened CT time-to-peak and reduced CT decay rate. Long stimulus period gives the cell more time to recover, prolonging the decay phase. The predicted negative effect of long stimulus period on CT time-to-peak was not as obvious, and is likely to be caused by a complex nonlinear interaction, calling for a more detailed analysis of the dynamic model and the restitution curve [31].

Our results showed that the global regression model revealed the main sensitivity patterns of the mouse ventricular myocyte model. However, as shown in the following sections, an accompanying regional sensitivity analysis can reveal more subtle predicted interactions, generating new testable hypotheses.

### 4.2. Additional sensitivity patterns of the dynamic model revealed by the HC-PLSR-based metamodel

#### 4.2.1. Characteristics of the HC-PLSR clusters

Using four clusters was found optimal by evaluation of the predictive ability (see Supplementary electronic material: Appendix C, Fig. C.2). As described in Section 3.3, the four clusters were identified by fuzzy *C*-means clustering [19–22] based on the *X*-scores from the global PLSR. The cluster-wise parameter ranges in Table 2 showed that extracellular potassium concentration and the stimulus period were the only parameters for which the ranges varied between the clusters. However, the four clusters still represented different distinct regions in the parameter space due to a different *distribution* of combinations of parameter values. The cluster-characteristics given in Table 3 were found by analysis of the bi-plots in Fig. 7 (indicating where in the score space each cluster was placed and which parameters and phenotypes that dominated each PLS component) and by inspection of the mosaic-plots [32] in Supplementary electronic material: Appendix C, Fig. C.4. Fig. 8 gives an example of more detailed interpretation of these clusters in terms of two of the phenotypes, colour coded according to stimulus period (Fig. 8A) and cluster identification (Fig. 8B). The *X*-loading plots for the regional regression models made in each cluster are shown in Supplementary electronic material: Appendix C, Fig. C.5.

The action potential (AP) and the calcium transient (CT) were plotted for the three levels of each of the parameters, for the four clusters (Supplementary electronic material: Appendix C, Figs. C.6 and C.7), indicating that the clusters represented different types of

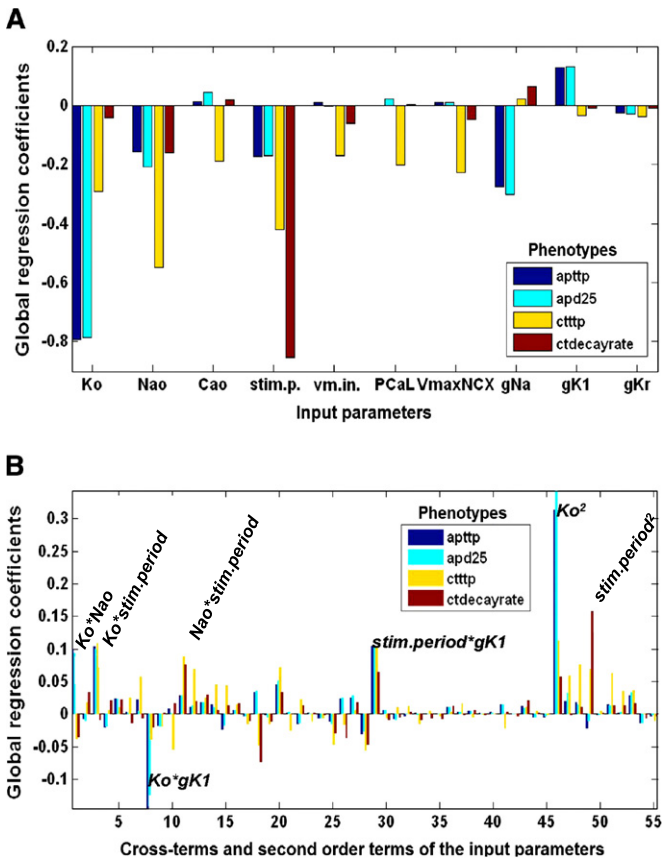


**Fig. 6.** Global regression coefficients for the aggregated phenotypes representing the action potential and the calcium transient. Regression coefficients from the global regression model (at optimal rank) of A) the input parameters *Ko*, *Nao*, *Cao*, *stim.period*, *vmupinit*, *PCaL*, *VmaxNCX*, *gNa*, *gK1* and *gKr* (stim.p. = *stim.period*, vm.in. = *vmupinit*) and B) their cross-terms and second order terms for the analysed key cell-level phenotypes: action potential time-to-peak (*apttp*), action potential duration to 25% repolarisation (*apd25*), calcium transient time-to-peak (*ctttp*) and calcium transient decay rate (*ctdecayrate*) are shown. The following cross-terms and second order terms had regression coefficient values above the approximate significance limit of ±0.1 (see Supplementary electronic material: Appendix C, Fig. C.1 for normal probability plots of all regression coefficients): *Ko*Nao*, *Ko*stim.period*, *Ko*gK1*, *Nao*stim.period*, *stim.period*gK1*, $Ko^2$ and $stim.period^2$. The regression coefficients are measures of the sensitivity of these phenotypes to variations in the input parameters.

temporal behaviour for the mouse ventricular myocyte model. Fig. 8 illustrates that the clustering was partly based on how fast the CT rises and decays. There were two distinct groupings, one with high CT decay rate and fast pacing, and one containing simulated cells with medium and long stimulus period. Clusters 3 and 4 were clearly distinguished by the stimulus period, while clusters 1 and 2 were more mixed. Both Figs. 7 and 8 showed that cluster 4 contained myocytes having high CT decay rates, cluster 3 contained cells with low CT decay rates, while clusters 1 and 2 contained a mix of high and low CT decay rates. According to Fig. 7, the AP time-to-peak separated to a large degree cluster 1 from cluster 2 (since AP time-to-peak pointed in the direction in which the two clusters were most clearly separated), where cluster 1 contained cells with longer AP time-to-peak while cluster 2 contained cells with short time-to-peak. As shown in Supplementary electronic material: Appendix C, Fig. C.6, cluster 2 contained only cells with short time-to-peak while cluster 1 actually contained a mix of short and long AP time-to-peak, so the AP time-to-peak did not completely separate cluster 1 and cluster 2. However, as seen from Supplementary electronic material: Appendix C, Fig. C.7, the CT amplitude was higher, on average, in cluster 2 than in cluster 1. The reason why the clusters could not be completely distinguished based on one phenotype was that the clustering was based on

**Table 2**
Parameter ranges (with mean ± standard deviations within clusters in parenthesis) for the groups of cells belonging to each of the four clusters in the HC-PLSR of the aggregated phenotypes for the mouse ventricular myocyte.

| Cluster | *Ko* | *Nao* | *Cao* | *stim.period* | *vmupinit* | *PCaL* | *VmaxNCX* | *gNa* | *gK1* | *gKr* |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2700–8100 | 67,000–201,000 | 700–2100 | 166.7–500 | 0.2530–0.7589 | 1.25–3.75 | 1.9695 – 5.9085 | 8–24 | 0.1750–0.5250 | 0.0083– 0.0248 |
| | (4188 ± 1349) | (175629 ± 34398) | (1344 ± 558) | (386 ± 105) | (0.4712 ± 0.2050) | (2.56 ± 0.98) | (4.0805 ± 1.5879) | (16.06 ± 6.45) | (0.3548 ± 0.1429) | (0.0167 ± 0.0067) |
| 2 | 5400–8100 | 67,000–201,000 | 700–2100 | 166.7–500 | 0.2530–0.7589 | 1.25–3.75 | 1.9695 – 5.9085 | 8–24 | 0.1750–0.5250 | 0.0083– 0.0248 |
| | (8074 ± 263) | (135578 ± 54288) | (1348 ± 561) | (355 ± 125) | (0.4895 ± 0.2068) | (2.53 ± 1.01) | (4.0228 ± 1.6035) | (11.94 ± 4.85) | (0.3766 ± 0.1382) | (0.0167 ± 0.0068) |
| 3 | 2700–8100 | 67,000–201,000 | 700–2100 | 333–500 | 0.2530–0.7589 | 1.25–3.75 | 1.9695 – 5.9085 | 8–24 | 0.1750–0.5250 | 0.0083– 0.0248 |
| | (4332 ± 1746) | (83827 ± 29759) | (1516 ± 563) | (430 ± 82.3) | (0.5550 ± 0.2008) | (2.50 ± 1.03) | (3.7568 ± 1.6170) | (16.26 ± 6.53) | (0.3479 ± 0.1416) | (0.0164 ± 0.0067) |
| 4 | 2700–8100 | 67,000–201,000 | 700–2100 | 166.7 | 0.2530–0.7589 | 1.25–3.75 | 1.9695 – 5.9085 | 8–24 | 0.1750–0.5250 | 0.0083– 0.0248 |
| | (5221 ± 2024) | (116120 ± 50703) | (1491 ± 557) | | (0.5211 ± 0.2049) | (2.58 ± 1.01) | (3.7382 ± 1.5982) | (17.39 ± 6.30) | (0.3280 ± 0.1411) | (0.0164 ± 0.0067) |

**Table 3**
Distributions of combinations of parameter levels characterising the four clusters in the HC-PLSR.

| Cluster | *Ko* | *Nao* | *gNa* | *stim.period* |
|---|---|---|---|---|
| 1 | Low/medium | Medium/high | – | Medium/high |
| 2 | High | – | Low/medium | – |
| 3 | Low/medium | Low/medium | – | Medium/high |
| 4 | –[a] | Low/medium[b] | Medium/high[b] | Low |

[a] Blank cells indicate an even distribution of values of the parameter.
[b] Cluster 4 had only a tendency towards low and medium *Nao* and medium and high *gNa*, the distribution was not as skewed as for the other combinations of clusters and parameter levels.

estimated latent variables from PLSR, which are linear combinations of the original variables in the analysis.

New cells can be classified into one of the identified clusters based on the parameter values, since the clustering was based on *X*-scores from the global PLSR. The fuzzy clustering algorithm estimates the probability of each combination of parameter values (i.e. cell) to belong to each of the clusters. The reason why we chose to cluster based on the parameter values and not on physiological states or phenotypes was that we wanted to search for complex interaction patterns between parameters described by differences in model sensitivities across the parameter space. Since dynamic models are generally sloppy, several different parameter combinations can generate the same output, and hence, clustering on the output would most likely produce clusters that would correspond to non-continuous or scattered regions in the parameter space. Hence, the effects of certain levels of one parameter on the sensitivity of the dynamic model to a second parameter would be difficult to identify.

*4.2.2. Gain in mapping approximation accuracy using HC-PLSR compared to global PLSR*

HC-PLSR separating the observations (here: different simulated cells with different combinations of parameter values) into four clusters representing different parameter space regions provided a clear gain in the ability to describe the parameter-to-output mapping for the mouse ventricular myocyte model (Fig. 9 and Supplementary electronic material: Appendix C). The number of well-predicted phenotypes (having squared correlation coefficient ($R^2$) values higher than 0.9) increased from 19 in the global PLSR metamodelling to 38 in the HC-PLSR metamodelling. An analysis of the polynomial HC-PLSR prediction residuals showed that the HC-PLSR tended to pick up the interpretable variation while relegating artefacts to the residuals in this study (see Supplementary electronic material: Appendix C, Fig. C.3). Including polynomial terms in the HC-PLSR (Fig. 9D) gave the opportunity to model more complex interaction patterns, through the ability to detect regional differences in the effects of cross-terms and second order terms. The HC-PLSR-based metamodel generates both the global PLSR model described in Section 4.1 and several regional regression models, and has therefore an implicit possibility to compare the sensitivity patterns revealed by the global and the regional regression models.

*4.2.3. Comparison of model sensitivity patterns in different regions of the parameter space*

From the regression coefficient plots (each regression model using the optimal number of PCs) shown in Fig. 10, regional differences between the effects of the various input parameters, their cross-terms and second order terms were identified for most of the 104 aggregated phenotypes used as responses in the metamodelling. This illustrates complex high-order interactions between parameters, since modifications of the model sensitivity to variations in a particular input parameter by increasing or decreasing the values of other parameters indicate that the parameter effects are dependent on each other. Regional differences between the effects of cross-terms and second order terms therefore represent 3rd and 4th order parameter interactions, and can be used as a supplement to cross-terms in order to describe more complex parameter interaction patterns.

In order to illustrate the potential of this methodology we chose to focus on the effects on the key cell-level phenotypes AP and CT. Although a similar analysis could have been done for all 104 included phenotypes, we considered this to be more instrumental as part of separate studies, combined with physiological measurements related to the particular dynamic model outputs. The main focus here was to illustrate the methodology. Box-plots [24–33] of the regression coefficients for all combinations of parameters and the AP and CT-related phenotypes in the four regional regression models (Fig. 11) and analysis of the standard deviations of the regression coefficients over the clusters (Supplementary electronic material: Appendix D, Table D.1 and D.2) showed that the effects of the following input parameters on the AP and the CT varied the most between the four clusters (standard deviations of the regression coefficients above 0.1 in Table D.1): *Nao* (extracellular sodium), *Cao* (extracellular calcium), *gK1* (scaling coefficient for the inward rectifying potassium current), *vmupinit* (scaling coefficient for calcium reuptake

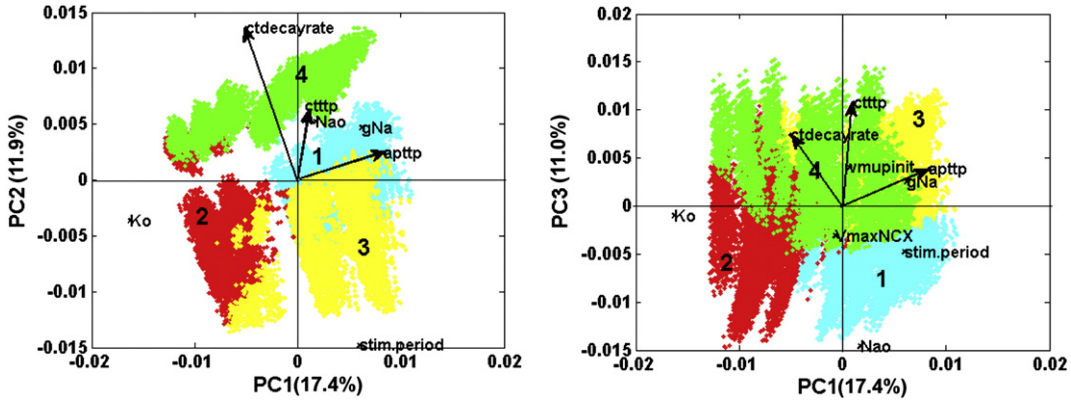


**Fig. 7.** The PLS scores and loadings showing the clustering results used in the HC-PLSR-based sensitivity analysis. The bi-plots show the *X*-scores (*PC1-PC3*) from the global PLS regression of the aggregated phenotypes with the clustering results (cluster 1 = cyan, cluster 2 = red, cluster 3 = yellow, cluster 4 = green) and the global *X*-loadings for the parameters spanning the first three PCs (as crosses), as well as the global *Y*-loadings (as points) for the action potential time-to-peak (*apttp*), the Ca-transient time-to-peak (*ctttp*) and the Ca-transient decay rate (*ctdecayrate*). The arrows point in the directions of the *Y*-loadings. The *X*- and *Y*-loadings were scaled by dividing by 10,000 in order to increase interpretability. The clustering was based on the first three PCs of the global *X*-scores. The plot illustrates which input parameters that are most related to the different phenotypes (lie close to/opposite each other in the plot), as well as which parametric and phenotypic characteristics the clusters possess.

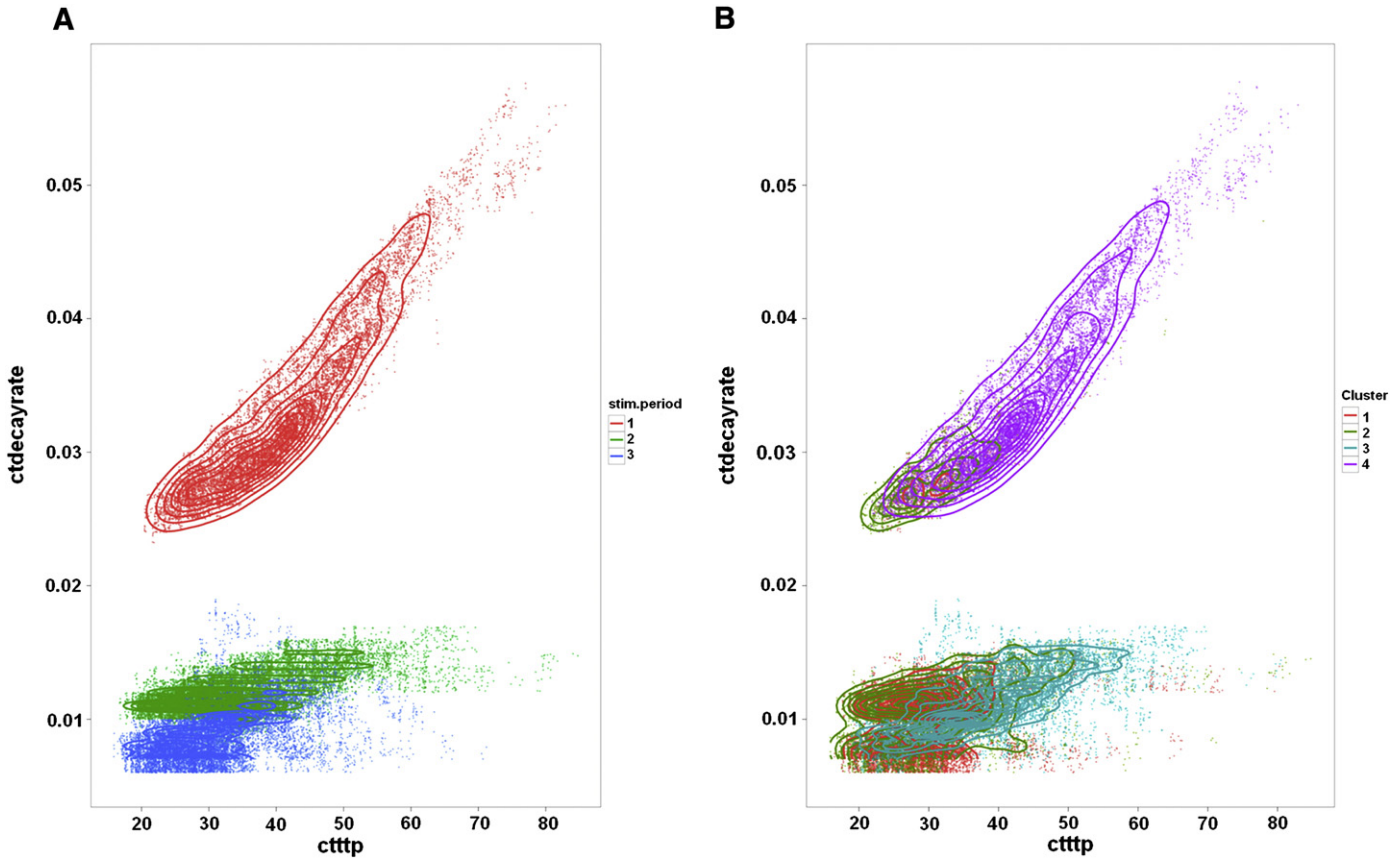


**Fig. 8.** Calcium transient decay rate (*ctdecayrate*) plotted against calcium transient time-to-peak (*ctttp*). Calcium transient (CT) decay rate is plotted against CT time-to-peak for A) the three different levels of the stimulus period and B) the four different clusters. The contours indicate the density of data points. The striated appearance was caused by a round-off of the plotted values. The plot was made using the "ggplot2" package for R [29].

from cytosol to sarcoplasmic reticulum (SR) by SERCA (sarco/endoplasmic reticulum $Ca^{2+}$-ATPase)) and *gNa* (scaling coefficient for the fast sodium current). A large standard deviation of the regression coefficient values for an input parameter indicates that the effect of variations in that parameter on the dynamic model output differs according to the values of other parameters (the clusters contain different parameter value combinations), meaning that the parameter interacts with other parameters. As shown in Table D.2, the effects of the extracellular potassium concentration and the stimulus period also varied a lot between the clusters, but the values of these parameters were almost constant in clusters 2 and 4, respectively (see Table 2). The effects of some of the cross-terms and second order terms between parameters also varied among the clusters, indicating complex, high-order interaction patterns, especially between *Ko*, *gNa* and *gK1*.

The parameter *gKr* (scaling coefficient for the rapid delayed rectifier potassium current) had a low impact on the phenotypes in all regional regression models, as well as in the global analysis, indicating that the mouse ventricular myocyte model is quite insensitive to this parameter, and might be simplified by setting this parameter to a nominal value in the model space we have studied here. The low sensitivity to this input parameter has also been identified by others [34,35]. Identification of such known patterns in dynamic model behaviour gives extra confidence to other results produced by this methodology.

In summary, based on the results in Fig. 11 and Table 3, the following parameter interaction effects on the AP and the CT characteristics were revealed in the HC-PLSR-based sensitivity analysis, that were not detected by the global regression model: The Na conductance (*gNa*) interacted with the inward rectifying $K^+$ channel conductance (*gK1*), the extracellular Na and K concentrations and the stimulus period, respectively; *gK1* and the Ca reuptake from cytosol to sarcoplasmic reticulum both interacted with the extracellular Na concentration; the extracellular Ca concentration interacted with the extracellular Na and K concentrations, respectively; there seemed to be a threshold value above which the effects of the extracellular Na concentration on the AP and the CT characteristics were not detectable, and complex, high-order interactions between *gNa*, *gK1*, the extracellular Na and K concentrations and the stimulus period were detected. More detailed explanations and biological interpretations of these results and in which parameter space regions they were manifested are given below (all interpretations are based on the results presented in Fig. 11, in the context of Table 3).

#### *4.2.4. Modification of the impact of increasing gNa and gK1 by the extracellular Na and K concentrations*

The effects of the sodium conductance (represented by *gNa*) on the AP time-to-peak and time to 25% AP repolarisation (*apttp* and *apd25*) were negative and stronger in cluster 3 than in the other clusters. However, no significant interaction effects involving the Na conductance on the AP and CT characteristics were detected by the global regression model. The effects on CT time-to-peak and CT decay rate (*ctttp* and *ctdecayrate*) were small for all clusters.

The low extracellular potassium and sodium concentration and long stimulus period in cluster 3 strengthened the effect of increasing the Na conductance on the AP time-to-peak and duration. A possible explanation is that this is a mechanism for the cell to compensate for the low extracellular sodium by a more effective sodium conduction in the initial phase of the AP. These interactions and possible compensatory mechanisms predicted by the myocyte model were not easily detectable from the model differential equations. If contradicted by

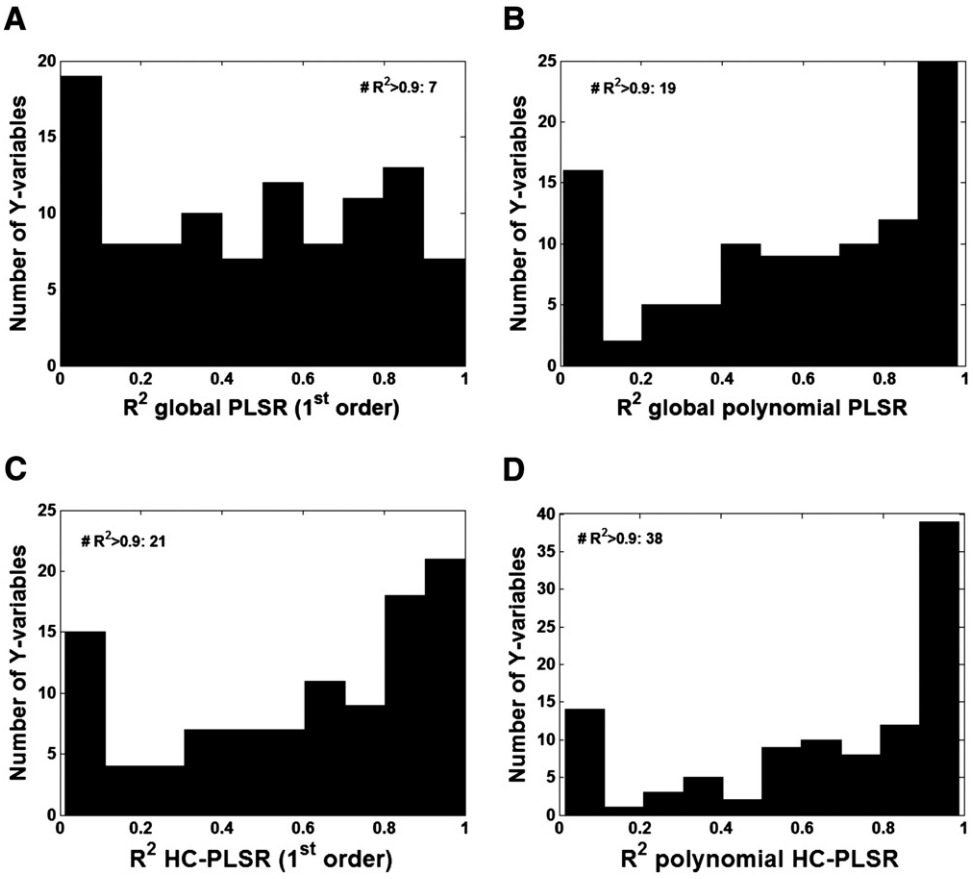


**Fig. 9.** Gain in prediction accuracy using HC-PLSR compared to global PLSR. In addition to the cross-validation carried out both in calibration of the global PLSR metamodel and in the calibration of the regional PLSR metamodels, a test set validation was carried out to further validate the HC-PLSR metamodelling, and compare it to global PLSR. This figure shows histograms over the correlation coefficient ($R^2$) values between test set predicted and reference values for the aggregated phenotypes using A) ordinary global first-order PLS regression (with 7 PCs), B) ordinary global polynomial PLS regression (with 11 PCs), C) first-order HC-PLSR (no cross-terms or second order terms of the input parameters were included in the regressor matrix), using from 8 to 10 PCs in the regional regression models and D) polynomial HC-PLSR (including cross-terms and second order terms of the input parameters in the regressor matrix), using from 9 to 11 PCs in the regional regression models. The number of well-predicted aggregated phenotypes is shown in each panel.

future laboratory experiments, they represent potential for refinement of the myocyte model. It is important to note, however, that these interactions predicted for the model parameterisations applied in this study may not apply in general or for a different set of model parameter scenarios.

Analogous to the above, the effects of increasing the potassium conductance in the inward rectifying $K^+$ channels (represented by *gK1*) on AP time-to-peak and duration were positive for both phenotypes and strongest in cluster 3. The inward rectifying $K^+$ current plays an important role in stabilising the ventricular cells at rest. The concentration gradient between the inside and the outside of the cell stimulates the outward current of $K^+$ in the repolarisation phase. Low extracellular potassium will therefore shorten the AP duration, and the positive effect of increasing *gK1* on the *apttp* and *apd25* might be strengthened in order to compensate for this. Low extracellular potassium concentration also leads to less $K^+$ being pumped in by the $Na^+/K^+$-ATPase (Na-pump).

There were also significant negative effects of *gNa* and positive effects of *gK1* in cluster 1, although smaller in magnitude than for cluster 3. This showed that a combination of relatively low extracellular potassium and relatively slow pacing strengthened the effects of *gNa* and *gK1* on AP time-to-peak and duration, but having in addition low extracellular Na strengthened the effects even further. In cluster 4, the *X*-loading plot of *PC3* vs. *PC1* showed an interaction between *gNa* and *gK1* (Supplementary electronic material: Appendix C, Fig. C.5).

#### 4.2.5. Modification of the impact of increasing Ca reuptake through SERCA by the extracellular Na concentration

The parameter *vmupinit* scales the calcium reuptake from cytosol to the SR by SERCA, and had a negative effect on CT time-to-peak that was strongest in clusters 3 and 4. Increasing *vmupinit* increases the calcium uptake into the SR, causing the CT to reach the peak earlier. Increasing *vmupinit* might also be expected to lower the calcium concentration at rest.

The relatively low extracellular sodium concentration in clusters 3 and 4 strengthened the negative effect of *vmupinit* on CT time-to-peak. A possible explanation of this result is that low extracellular sodium concentration leads to less calcium-ions leaving through the NCX, leading to more $Ca^{2+}$ being taken up by SERCA due to a higher concentration of calcium in the cytosol. The total effect will be an *increase* in the CT time-to-peak, but SERCA itself will then play a more important role; hence the strengthened effect of *vmupinit* on CT time-to-peak. This interaction was also identified in the *X*-loading plot of *PC3* vs. *PC1* for cluster 4 (Supplementary electronic material: Appendix C, Fig. C.5).

#### 4.2.6. Modification of the impact of increasing extracellular Ca concentration by the extracellular Na and K concentrations

The effect of extracellular calcium concentration, *Cao*, on CT time-to-peak was negative (due to a faster inward current of calcium), and much larger in clusters 1 and 2 than in the other two clusters. Hence, the effect was strengthened by high extracellular sodium or

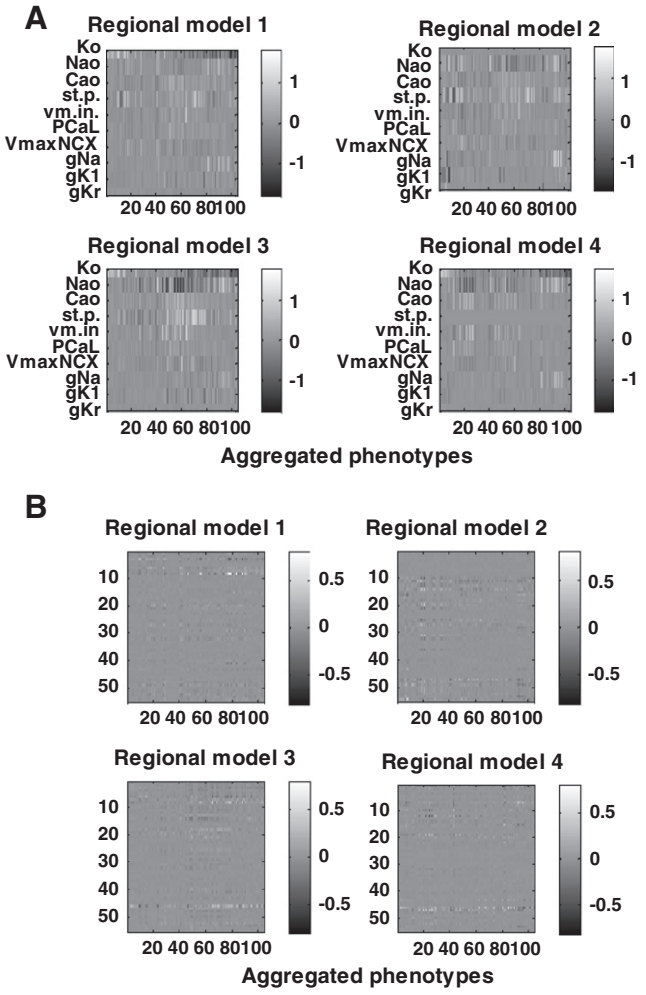


**Fig. 10.** Results from the regression-based sensitivity analysis with HC-PLSR. A) Regression coefficients of the model input parameters *Ko*, *Nao*, *Cao*, *stim.period*, *vmupinit*, *PCaL*, *VmaxNCX*, *gNa*, *gK1* and *gKr* (stim.p. = *stim.period*, vm.in. = *vmupinit*) for all aggregated phenotypes from the four regional regression models in the HC-PLSR of the aggregated phenotypes for the mouse ventricular myocytes. The four regional regression models use 11, 9, 11 and 10 PCs, respectively. The phenotypes are sorted according to decreasing loading values of *PC1* from the global regression model (see Supplementary electronic material: Appendix B), in the same way as in Fig. 5. B) Regression coefficients of the cross-terms and second order terms of the model input parameters for all the aggregated phenotypes from the four regional regression models in the HC-PLSR (using 11, 9, 11 and 10 PCs, respectively). The numbering of the variables and the sorting of the phenotypes correspond to that used in Fig. 5.

potassium concentration. This is probably due to a complex coupling between the NCX and the Na-pump, related to sodium-calcium overload [31]. In the global regression analysis, however, no significant interaction between *Cao* and other parameters was detected.

*4.2.7. Modification of the impact of increasing extracellular Na concentration by the extracellular Na and K concentrations*

The extracellular Na concentration had a negative effect on the AP time-to-peak and duration in clusters 3 and 4 (having low concentrations of extracellular Na), but the effects were negligible in clusters 1 and 2 (having high extracellular Na). Hence, there seemed to be a *threshold value* above which the effect of further increasing extracellular Na concentration had no effect on these phenotypes.

Analogous to the above, a threshold value was indicated for the effect of increasing extracellular Na concentration on CT time-to-peak. This might be caused by the fact that there are two mechanisms modulating Ca-ion concentration; NCX and SERCA. Increasing extracellular Na stimulates NCX to transport $Na^+$ in and $Ca^{2+}$ out, while SERCA continues to transport $Ca^{2+}$ into the SR.

The effect of extracellular Na on CT decay rate was negative and only significant with fast pacing, low extracellular Na concentration and high Na conductance (cluster 4). This result indicated a very complex coupling between the NCX and the Na-pump. Fast pacing and high Na conductance would cause more sodium to enter the cell, both through the Na-channels and the NCX. This would normally cause large amounts of Ca-ions to leave, but our results indicated that the low extracellular Na causing the Na-pump to transport more Na-ions out might offset the Na–Ca effect. This calls for further analysis of the dynamic model in order to fully understand these possible coupled effects.

## 5. Discussion

The literature on statistical sensitivity analysis contains several alternatives to regression-based sensitivity analysis, such as rank transformation, first- and second order reliability algorithms (FORM and SORM) and variance-based methods [3]. Rank transformation is an alternative to conventional regression-based sensitivity analysis in cases where the input–output relations are monotonically nonlinear, while reliability algorithms are used in cases where the primary focus is on a particular mode of failure of the system rather than the entire spectrum of possible outcomes. Variance-based methods, such as Sobol's method, use analysis of variance (ANOVA)-type decomposition of the output function into a polynomial expression including cross-terms between the input parameters. Partial variances are computed from each of the terms in the decomposition, and the sensitivity of each term is defined as the partial variance divided by the total output variance. However, these methods concentrate on the effects on one output variable at a time. Artificial Neural Network-based methods [36], on the other hand, can fit input–output relations including several outputs successfully, but the results produced by these methods are often not straight-forward to interpret.

Within regression-based sensitivity analysis, an alternative to regional regression modelling may be to expand the regression with high-order polynomial terms. However, this approach may be difficult to stabilise against over-fitting. Furthermore, representing input parameter interactions only through their cross-terms or higher-order product terms does not give good descriptions of the interactions between parameters if one of them equals zero (which is a parameter setting that is likely to affect the impact of other parts of the dynamic model on the output), and will not separate e.g. between the effects of two parameters both being positive contra both having negative values. Hence, combining polynomial regression with regional analysis gives insight into more aspects of parameter interaction.

Most of the regression-based sensitivity analyses published are based on OLS and extensions thereof, such as ANOVA or second-degree polynomial response surface methodology [2–5], that with few exceptions require the input parameters (regressors, $X$) to be linearly independent. Even in cases where the input parameter space of the computer experiment was originally spanned at full rank suitable for OLS, a subsequent regional sensitivity analysis may require multivariate regression in local subsets of the data that may not have full rank. Under such circumstances reduced-rank regression methods are required.

Bi-linear (BLM) regression methods [13], like PLSR and principal component regression (PCR) [37,38], identify subspaces of particular relevance for the relationship between regressors and the response ($Y$), providing considerable data compression possibilities. We consider PLSR to be the most informative of the two methods in a sensitivity analysis context due to its more $Y$-relevant optimisation criterion [12,39–42]. PLSR is recognised for its merits to capture covariance structures in large and complex data sets, and thereby reveal hidden correlations both within and between partially redundant

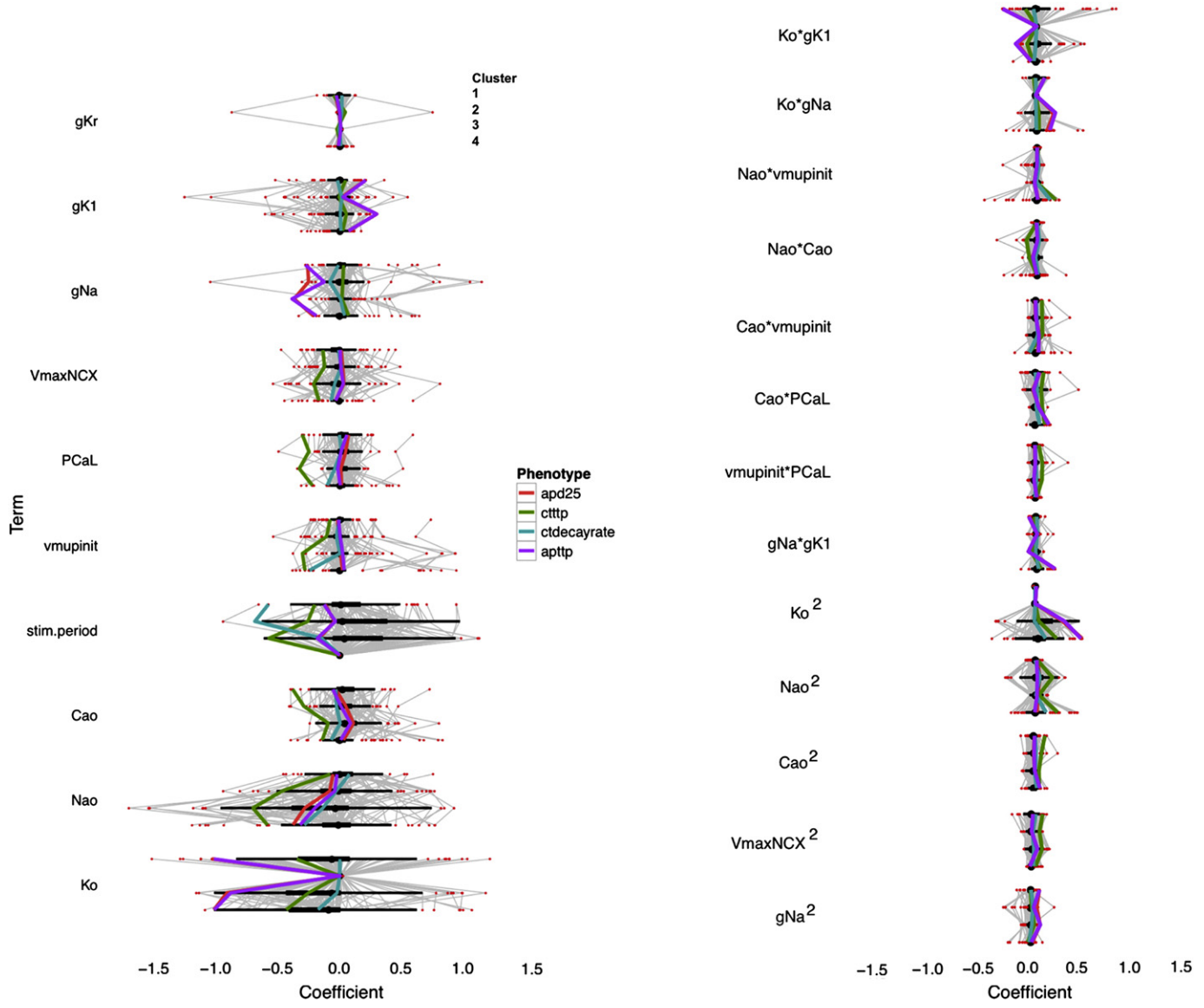


**Fig. 11.** Variability in the regression coefficients for the input parameters among the clusters in the HC-PLSR. The contrast between clusters in the distribution of effect sizes of the input parameters ($X$) on the 104 selected phenotypes ($Y$) that span most of the total phenotypic variance is illustrated. Variation in the values of the regression coefficients for the various input parameters on the output phenotypes between the HC-PLSR clusters reflects that certain levels of some input parameters (defining the clusters) can modify the effects of other parameters on the dynamic model output (the model sensitivity varies across the parameter space), i.e. there are interactions between the input parameters. Here, within each term and cluster, the thick line shows the interquartile range ($IQR = Q3 - Q1$) of the regression coefficients for each phenotype, the middle dot shows the median (second quartile), and the thin line shows "whiskers" extending to the smallest and the largest data point within 1.5 times $IQR$ of the $IQR$. Data points beyond that are plotted separately in red. Lines connect effects on the same phenotype of each term, colour highlights cell-level phenotypes, only four of which (*apttp*, *apd25*, *ctttp* and *ctdecayrate*) were included among the 104 phenotypes selected in the redundancy analysis. For the cross-terms and second order terms only those that had varying effects on these four phenotypes are shown. The plot was made using the function "boxplot.stats" in R [29]. The standard deviations of the regression coefficients among the clusters for each pair of input parameter and aggregated phenotype are listed in Supplementary electronic material: Appendix D.

inputs and outputs [14,43]. This can be used to prioritise measurements of both parameters and phenotypes in validation of dynamic models. Martens and Martens [44] demonstrated the use of PLSR as an alternative to ANOVA to facilitate the interpretation of multi-response data from designed experiments. Campbell *et al.* [45] recently showed that when compared to Legendre polynomials and Principal Component Analysis (PCA), the subspace of the PLSR gave the simplest and most predictive basis for sensitivity analysis of computational dynamic models. The suitability of PLSR for interpretation of complex biological systems and use of PLSR in sensitivity analysis were also demonstrated in [14,15]. This motivated our preference for PLSR both in the global and regional sensitivity analyses.

In HC-PLSR, differences in model sensitivity to input parameters conditional on the span of the other parameters can be revealed by an automatic combination of global and regional sensitivity analysis. HC-PLSR accounts for nonlinearities both through the regional regression analysis and through including polynomial terms in the regression, and represents an improved methodology to detect complex interaction patterns predicted by the analysed dynamic model that can form the basis for further validation and refinement of the analysed model. This can also reveal dynamic model behaviour and interaction patterns between inputs that differ in a diseased state compared to in a healthy state, and may therefore be effective for identification of clinically relevant characteristics of model behaviour.

Here, parameter interaction patterns predicted by the dynamic model of the mouse ventricular myocyte, that could not be detected by the global regression model, were revealed using HC-PLSR. This improvement in mapping approximation accuracy demonstrated the advantage of using regional regression analysis to analyse dynamic model behaviour. However, some identified patterns were considered to be artefacts; the much lower effect of the extracellular potassium concentration in regional regression model 2 compared to the other regional regression models was probably an artefact of the low variation in the values of this parameter in cluster 2; analogous for the effects of the stimulus period in cluster 4, which was constant at the lowest level in this cluster. According to Fig. 11, calcium reuptake from cytosol to SR had a negative effect on CT decay rate, which was only evident in cluster 4. This was probably also an artefact, since cluster 4 consisted of only outliers in the plot of *vmupinit* versus CT decay rate (Supplementary electronic material: Appendix C, Fig. C.8).

Using HC-PLSR, the number of aggregated phenotypes well predicted by the metamodelling was doubled compared to the global polynomial PLSR model. Hence, even for the heart muscle cell model, which has a moderately nonlinear parameter-to-phenotype mapping for the parameter sets and aggregated phenotypes used in this paper (compared to other model settings of potential interest [6–8]), regional regression modelling provided a clear gain in both prediction accuracy and analytical insight. Several new parameter interactions were identified by the HC-PLSR-based sensitivity analysis, such as that low extracellular potassium and sodium concentration strengthened the effect of increasing the sodium conductance on the AP time-to-peak and duration. This was here explained as a possible compensatory mechanism for the cell. Our results also indicated a threshold value for the effects of increasing extracellular Na concentration on the AP and CT-related phenotypes. Hence, our approach to sensitivity analysis has the potential to reveal new biological insight through the identification of complex parameter interaction patterns. However, the results from these simulations only tell us about the behaviour of the myocyte model under the specific model settings used in this particular study. Simulations using other settings, such as different values for model parameters that were not varied here, may produce other results due to the demonstrated variability in model sensitivity across the parameter space.

Larger gains in prediction accuracy have been observed when using the state trajectories themselves as response variables [10] due to a more non-monotone dependency on the model input parameters. Thus in cases with an even higher degree of nonlinear behaviour one would expect even lager gains by using a regional regression approach. In order to allow for a semi-automatic sensitivity analysis of complex dynamic models, robust methods that can handle the entire spectrum of dynamic models from simple and relatively monotonous systems to highly nonlinear and non-monotone parameter-to-phenotype mappings should be considered with particular interest, since ones insight into the complexity of the parameter-to-phenotype map is often limited prior to the analysis. We therefore anticipate that HC-PLSR, which is designed to semi-automatically adjust the number of separately analysed regions to suit the complexity of the analysed dynamic model, will be highly instrumental for this endeavour.

## Acknowledgements

This study was supported by the National Program for Research in Functional Genomics in Norway (FUGE) (RCN grant no. NFR151924/S10) and by the Norwegian eScience program (eVITA) (RCN grant no. NFR178901/V30).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.chemolab.2012.10.006.

## References

[1] D. Houle, D.R. Govindaraju, S. Omholt, Phenomics: the next challenge, Nature Reviews Genetics 11 (2010) 855–866.
[2] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, *et al.*, Global Sensitivity Analysis: the Primer, Wiley-Interscience, 2008.
[3] D.G. Cacuci, M. Ionescu-Bujor, I.M. Navon, 1st ed., Sensitivity and Uncertainty Analysis: Applications to Large-scale Systems, vol. 2, CRC Press, 2005.
[4] S. Marino, I.B. Hogue, C.J. Ray, D.E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, Journal of Theoretical Biology 254 (2008) 178–196.
[5] A. Saltelli, K. Chan, E.M. Scott, Sensitivity Analysis, 1st ed. Wiley, 2000.
[6] S. Kaplan, A. Bren, E. Dekel, U. Alon, The incoherent feed-forward loop can generate non-monotonic input functions for genes, Molecular Systems Biology 4 (2008) 203.
[7] K.L. Wang, J.R. Warner, Positive and negative autoregulation of REB1 transcription in *Saccharomyces cerevisiae*, Molecular and Cellular Biology 18 (1998) 4368–4376.
[8] A.B. Gjuvsland, J.O. Vik, J.A. Woolliams, S.W. Omholt, Order-preserving principles underlying genotype–phenotype maps ensure high additive proportions of genetic variance, Journal of Evolutionary Biology 24 (2011) 2269–2279.
[9] J.P.C. Kleijnen, Design and Analysis of Simulation Experiments, 1st ed. Springer, New York, USA, 2007.
[10] K. Tøndel, U.G. Indahl, A.B. Gjuvsland, J.O. Vik, P. Hunter, S.W. Omholt, *et al.*, Hierarchical Cluster-based Partial Least Squares Regression is an efficient tool for metamodelling of nonlinear dynamic models, BMC Systems Biology 5 (2011) 90.
[11] S. Wold, H. Martens, H. Wold, The multivariate calibration method in chemistry solved by the PLS method, in: Lecture Notes in Mathematics, Matrix Pencils, Springer-Verlag, Heidelberg, 1983, pp. 286–293.
[12] H. Martens, T. Næs, Multivariate Calibration, John Wiley and Sons, Chichester, UK, 1989.
[13] H. Martens, M. Martens, Multivariate Analysis of Quality: an Introduction, 1st ed. Wiley, 2001.
[14] E.A. Sobie, Parameter sensitivity analysis in electrophysiological models using multivariable regression, Biophysical Journal 96 (2009) 1264–1274.
[15] H. Martens, S. Veflingstad, E. Plahte, M. Martens, D. Bertrand, S. Omholt, The genotype–phenotype relationship in multicellular pattern-generating models — the neglected role of pattern descriptors, BMC Systems Biology 3 (2009) 87.
[16] L. Li, S.A. Niederer, W. Idigo, Y.H. Zhang, P. Swietach, B. Casadei, *et al.*, A mathematical model of the murine ventricular myocyte: a data-driven biophysically based approach applied to mice overexpressing the canine NCX isoform, American Journal of Physiology — Heart and Circulatory Physiology 299 (2010) H1045–H1063.
[17] V.E. Bondarenko, G.P. Szigeti, G.C.L. Bett, S.-J. Kim, R.L. Rasmusson, Computer model of action potential of mouse ventricular myocytes, American Journal of Physiology — Heart and Circulatory Physiology 287 (2004) H1378–H1403.
[18] R.E. Klabunde, Cardiovascular Physiology Concepts, Lippincott Williams & Wilkins, 2004.
[19] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Kluwer Academic Publishers, 1981.
[20] I. Berget, B.-H. Mevik, T. Næs, New modifications and applications of fuzzy C-means methodology, Computational Statistics & Data Analysis 52 (2008) 2403–2418.
[21] T. Næs, T. Isaksson, Splitting of calibration data by cluster analysis, Journal of Chemometrics 5 (1991) 49–65.
[22] T. Næs, E. Kubberød, H. Sivertsen, Identifying and interpreting market segments using conjoint analysis, Food Quality and Preference 12 (2001) 133–143.
[23] I. Gath, A.B. Geva, Unsupervised optimal fuzzy clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (1989) 773–780.
[24] H. Frigui, R. Krishnapuram, A robust competitive clustering algorithm with applications in computer vision, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (1999) 450–465.
[25] S.D. Cohen, A.C. Hindmarsh, CVODE, a stiff/nonstiff ODE solver in C, Computers in Physics 10 (1996) 138–143.
[26] A.C. Hindmarsh, P.N. Brown, K.E. Grant, S.L. Lee, R. Serban, D.E. Shumaker, *et al.*, SUNDIALS: suite of nonlinear and differential/algebraic equation solvers, ACM Transactions on Mathematical Software 31 (2005) 363–396.
[27] H.A. Martens, P. Dardenne, Validation and verification of regression in small data sets, Chemometrics and Intelligent Laboratory Systems 44 (1998) 99–121.
[28] MATLAB®, v. 7.13, The MathWorks™, 2011.
[29] R, v. 2.13.1, The R Foundation for Statistical Computing, 2011.
[30] A.X. Sarkar, E.A. Sobie, Regression analysis for constraining free parameters in electrophysiological models of cardiac cells, PLoS Computational Biology 6 (2010).
[31] T. Simor, T. Lóránd, B. Gaszner, G.A. Elgavish, The modulation of pacing-induced changes in intracellular sodium levels by extracellular $Ca^{2+}$ in isolated perfused rat hearts, Journal of Molecular and Cellular Cardiology 29 (1997) 1225–1235.
[32] M. Friendly, Mosaic displays for multi-way contingency tables, Journal of the American Statistical Association 89 (1994) 190–200.
[33] J.W. Tukey, Exploratory Data Analysis, University Microfilms International, 1988.
[34] L.J. Wang, E.A. Sobie, Mathematical model of the neonatal mouse ventricular action potential, American Journal of Physiology — Heart and Circulatory Physiology 294 (2008) H2565–H2575.
[35] G.X. Liu, J. Zhou, S. Nattel, G. Koren, Single-channel recordings of a rapid delayed rectifier current in adult mouse ventricular myocytes: basic properties and effects of divalent cations, Journal of Physiology 556 (2004) 401–413.
[36] E. El Tabach, L. Lancelot, I. Shahrour, Y. Najjar, Use of artificial neural network simulation metamodelling to assess groundwater contamination in a road project, Mathematical and Computer Modelling 45 (2007) 766–776.

[37] I.T. Jolliffe, A note on the use of principal components in regression, Journal of the Royal Statistical Society: Series C: Applied Statistics 31 (1982) 300–303.
[38] R. Kramer, Chemometric Techniques for Quantitative Analysis, 1st ed. CRC Press, 1998.
[39] A. Höskuldsson, PLS regression methods, Journal of Chemometrics 2 (1988) 211–228.
[40] I.S. Helland, Partial least squares regression and statistical models, Scandinavian Journal of Statistics 17 (1990) 97–114.
[41] I.S. Helland, Steps Towards a Unified Basis for Scientific Models and Methods, World Scientific, 2009.
[42] I.S. Helland, T. Almøy, Comparison of prediction methods when only a few components are relevant, Journal of the American Statistical Association 89 (1994) 583–591.
[43] K.A. Janes, J.G. Albeck, S. Gaudet, P.K. Sorger, D.A. Lauffenburger, M.B. Yaffe, A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis, Science 310 (2005) 1646–1653.
[44] M. Martens, H. Martens, Partial least squares regression, in: J.R. Piggott (Ed.), Statistical Procedures in Food Research, Elsevier Applied Sciences, London, 1986, pp. 293–360.
[45] K. Campbell, M.D. McKay, B.J. Williams, Sensitivity analysis when model outputs are functions, Reliability Engineering and System Safety 91 (2006) 1468–1472.