Additional file 1 for “Parameter estimation in models of biological oscillators: an automated regularised estimation approach.”

Jake Alan Pitt\textsuperscript{1,2} and Julio R. Banga\textsuperscript{1*}

\textsuperscript{1} (Bio)Process Engineering Group, Spanish National Research Council, IIM-CSIC, C/Eduardo Cabello 6, 36208, Vigo, Spain

\textsuperscript{2} RWTH-Aachen University Hospital, Joint Research Centre for Computational Biomedicine (JRC-COMBINE), 52074, Aachen, Germany

\textsuperscript{*}Correspondence should be directed to Prof J. R. Banga at julio@iim.csic.es

S1.1 Analysis of logarithmic scaling in eSS

The original scatter search method used for parameter estimation in dynamic models already incorporated logarithmic sampling of the search space (Rodriguez-Fernandez et al., 2006). In that algorithm, a logarithm scaling was applied to the generation of the initial reference set of the parameter values, allowing for an initial diversification of the exploration of the parameter space. However, logarithm scaling was not used in other parts of the search. In recent years, several studies (Raue et al., 2013; Fröhlich et al., 2017; Kreutz, 2016; Villaverde et al., 2018) have proposed to perform the whole parameter estimation in logarithm space.

In this study we compared three different sets-ups: eSS with the entire search in log scale, eSS with all the search except for the local solver in log scale and the default formulation with only the initial reference set being in log scale. We tested each of the scaling set-ups for all four case studies, for both noisy and noiseless data (the first fitting data set is used as the noiseless data). To test the methods we run eSS 30 times for each case running eSS with the stopping criteria of reach the global optima’s (known) cost value or a hard cut off of 3 hours. We found that in every case running the local search in linear scale and rest of eSS in log scale was the most efficient method.

S1.1.1 Analysis with noisy data
**Figure S1.1:** FHN case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noisy data.

**Figure S1.2:** GO case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noisy data.
**Figure S1.3:** RP case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noisy data.

**Figure S1.4:** EO case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noisy data.
S1.1.2 Analysis with noiseless data

**Figure S1.5:** FHN case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noiseless data.

**Figure S1.6:** GO case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noiseless data.

In summary, we found that, at least for the problems considered here, the most efficient search scaling in eSS to perform the diversification search in logarithm scale but not the local search. We found that this was the most efficient method when fitting to both noisy
**Figure S1.7:** RP case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noiseless data.

**Figure S1.8:** EO case study: A comparison of how the using eSS with different sections of the solver in log scale effects eSS’ efficiency when fitting to noiseless data.

and noiseless models, for all the oscillators case studies considered.
S1.2 Computational reproducibility of eSS

Reproducibility is a major issue in computational research. The fact that eSS is a stochastic solver needs to be taken into account. Stochastic optimisers use some type of random number generator in their sampling of search space. Modern pseudo-random number generators use a seed which, if fixed, determines the sequence of pseudo-random values. Therefore, starting eSS runs with different seeds will result in different optimisation path followed, although most runs will converge to essentially the same final solution (depending on the stopping criteria chosen).

To illustrate this, in Figure S1.9 we plot the contours of the ENSO problem (considering the projection for 2 parameters) and then use eSS to solve the problem starting from the same initial guess multiple times, changing the seed randomly in each run. This results in different optimisation paths which ultimately arrive to essentially the same final solution.

In theory, fixing the seed should result in the same optimisation path. However, in GEARS we use all the information from the path taken in the form of parameter-cost distributions, which can lead to different solutions when the procedure is re-run. That is, even when starting from the same seed (so that the same pseudo-random sequence will be generated), slight differences caused by the stopping criteria can affect the result. This is because some types of stopping criteria are not checked during the local search phases. For example, if we have set eSS to stop after 2000 function evaluations in one run, it might however stop after 2003 evaluations in another. As all the information for every parameter point is used, these 3 extra points would cause differences that would be passed downstream and cause a lack of strict computational reproducibility.

Figure S1.9: ENSO case study: Contour plots with multiple starts of the eSS solver from the same initial point showing that the stochastic nature of the solver results in different paths being taken.
S1.3 Multimodality in the ENSO case study

In the case of the ENSO model by plotting contour plots we can clearly see the multimodality of the model, even when only considering two of the parameters (for visualisation purposes). In Figure S1.10 we can see the existence of many local optima in the search space. This many peak situations is essentially a worst case scenario in parameter estimation, showing extreme multimodality, making the parameter estimation problem extremely challenging to solve.

![Figure S1.10: ENSO case study: Contour plots of the b(4) and b(7) parameters plotted in three dimensions.](image)

S1.4 GO problem: detailed results

The results for running the analysis on the GO case study can be found here.
Table S1.1: GO case study: A summary of the results for the regularised fit to the first fitting data set.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Confidence (95%)</th>
<th>Coeff of variation (%)</th>
<th>Bounds status</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>2.4731</td>
<td>$\pm 0.034618$</td>
<td>0.714172</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.0917</td>
<td>$\pm 0.011088$</td>
<td>6.17051</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.9477</td>
<td>$\pm 0.13256$</td>
<td>7.13619</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$k_4$</td>
<td>0.1428</td>
<td>$\pm 0.019794$</td>
<td>7.07017</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$k_5$</td>
<td>1.2455</td>
<td>$\pm 0.24837$</td>
<td>10.1739</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$k_6$</td>
<td>0.0861</td>
<td>$\pm 0.015879$</td>
<td>9.40842</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$K_i$</td>
<td>0.9638</td>
<td>$\pm 0.18518$</td>
<td>9.80278</td>
<td>Bounds not active</td>
</tr>
<tr>
<td>$n$</td>
<td>11.7721</td>
<td>$\pm 2.1469$</td>
<td>9.30473</td>
<td>Bounds not active</td>
</tr>
</tbody>
</table>

Table S1.2: GO case study: NRMSE values for the fitting for each fitting data set, with and without regularisation.

<table>
<thead>
<tr>
<th></th>
<th>Regularised</th>
<th>Non-regularised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitting set 1</td>
<td>5.96798</td>
<td>4.77509</td>
</tr>
<tr>
<td>Fitting set 2</td>
<td>37.3928</td>
<td>4.28762</td>
</tr>
<tr>
<td>Fitting set 3</td>
<td>23.0941</td>
<td>5.07253</td>
</tr>
<tr>
<td>Fitting set 4</td>
<td>9.13458</td>
<td>7.46763</td>
</tr>
<tr>
<td>Fitting set 5</td>
<td>30.7735</td>
<td>30.7977</td>
</tr>
<tr>
<td>Fitting set 6</td>
<td>21.6336</td>
<td>5.21221</td>
</tr>
<tr>
<td>Fitting set 7</td>
<td>30.5141</td>
<td>30.3792</td>
</tr>
<tr>
<td>Fitting set 8</td>
<td>12.699</td>
<td>10.3537</td>
</tr>
<tr>
<td>Fitting set 9</td>
<td>29.6946</td>
<td>29.6743</td>
</tr>
<tr>
<td>Fitting set10</td>
<td>29.0661</td>
<td>8.60332</td>
</tr>
</tbody>
</table>
Table S1.3: GO case study: NRMSE values for the cross-validation for each regularised fit to the fitting data. Here, CV denotes cross-validation data set and F denotes fitting data set.

<table>
<thead>
<tr>
<th></th>
<th>F 1</th>
<th>F 2</th>
<th>F 3</th>
<th>F 4</th>
<th>F 5</th>
<th>F 6</th>
<th>F 7</th>
<th>F 8</th>
<th>F 9</th>
<th>F 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>All CV</td>
<td>49.1531</td>
<td>45.6818</td>
<td>92.87566</td>
<td>82.02807</td>
<td>124.931</td>
<td>43.5472</td>
<td>45.5031</td>
<td>64.29963</td>
<td>132.8635</td>
<td>237.2202</td>
</tr>
<tr>
<td>CV 1</td>
<td>90.2766</td>
<td>44.5872</td>
<td>91.60485</td>
<td>98.67496</td>
<td>160.9028</td>
<td>42.7661</td>
<td>44.6539</td>
<td>54.58628</td>
<td>139.1305</td>
<td>81.22797</td>
</tr>
<tr>
<td>CV 2</td>
<td>17.4337</td>
<td>55.6168</td>
<td>108.2136</td>
<td>102.4116</td>
<td>137.4336</td>
<td>51.4413</td>
<td>55.3724</td>
<td>75.44783</td>
<td>181.2703</td>
<td>154.689</td>
</tr>
<tr>
<td>CV 3</td>
<td>59.2786</td>
<td>41.4897</td>
<td>132.7566</td>
<td>81.3378</td>
<td>190.7445</td>
<td>41.0993</td>
<td>41.7198</td>
<td>100.1812</td>
<td>151.4091</td>
<td>55.42433</td>
</tr>
<tr>
<td>CV 4</td>
<td>49.6364</td>
<td>41.3489</td>
<td>95.58716</td>
<td>57.79468</td>
<td>77.52544</td>
<td>37.2885</td>
<td>41.3348</td>
<td>65.96083</td>
<td>50.11638</td>
<td>364.066</td>
</tr>
<tr>
<td>CV 5</td>
<td>15.5789</td>
<td>48.1839</td>
<td>103.2592</td>
<td>87.23705</td>
<td>107.1258</td>
<td>48.9272</td>
<td>47.8465</td>
<td>37.84106</td>
<td>135.5939</td>
<td>46.12598</td>
</tr>
<tr>
<td>CV 6</td>
<td>33.8966</td>
<td>45.9419</td>
<td>66.3954</td>
<td>64.4661</td>
<td>96.05532</td>
<td>46.6294</td>
<td>45.4043</td>
<td>82.39392</td>
<td>75.67943</td>
<td>85.2841</td>
</tr>
<tr>
<td>CV 7</td>
<td>45.3068</td>
<td>40.5613</td>
<td>52.82711</td>
<td>76.48417</td>
<td>70.68241</td>
<td>36.401</td>
<td>40.1927</td>
<td>33.8625</td>
<td>44.96142</td>
<td>396.7472</td>
</tr>
<tr>
<td>CV 8</td>
<td>77.0161</td>
<td>45.5434</td>
<td>81.13562</td>
<td>90.45528</td>
<td>127.3902</td>
<td>43.8462</td>
<td>45.1954</td>
<td>55.16384</td>
<td>139.7722</td>
<td>258.7255</td>
</tr>
<tr>
<td>CV 9</td>
<td>14.8661</td>
<td>42.0491</td>
<td>111.9855</td>
<td>83.81761</td>
<td>126.1179</td>
<td>38.2899</td>
<td>42.1025</td>
<td>52.71671</td>
<td>186.9935</td>
<td>62.97651</td>
</tr>
<tr>
<td>CV 10</td>
<td>11.4834</td>
<td>49.3559</td>
<td>91.35434</td>
<td>65.43839</td>
<td>105.2309</td>
<td>46.1197</td>
<td>49.1384</td>
<td>55.64005</td>
<td>135.0444</td>
<td>398.5941</td>
</tr>
</tbody>
</table>
Table S1.4: GO case study: NRMSE values for the cross-validation for each non-regularised fit to the fitting data. Here, CV denotes cross-validation data set and F denotes fitting data set.

<table>
<thead>
<tr>
<th></th>
<th>F 1</th>
<th>F 2</th>
<th>F 3</th>
<th>F 4</th>
<th>F 5</th>
<th>F 6</th>
<th>F 7</th>
<th>F 8</th>
<th>F 9</th>
<th>F 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>All CV</td>
<td>52.0968</td>
<td>446.13711</td>
<td>188.8501</td>
<td>157.9609</td>
<td>183.2273</td>
<td>101.2102</td>
<td>45.2463</td>
<td>77.25336</td>
<td>179.9307</td>
<td>379.0375</td>
</tr>
<tr>
<td>CV 1</td>
<td>87.4309</td>
<td>70.334685</td>
<td>230.7914</td>
<td>97.35506</td>
<td>234.9919</td>
<td>105.9857</td>
<td>44.8763</td>
<td>62.66137</td>
<td>191.5425</td>
<td>49.92867</td>
</tr>
<tr>
<td>CV 2</td>
<td>19.8086</td>
<td>21.002245</td>
<td>206.6961</td>
<td>22.64062</td>
<td>209.0168</td>
<td>108.3317</td>
<td>54.5939</td>
<td>61.06428</td>
<td>245.0328</td>
<td>44.90806</td>
</tr>
<tr>
<td>CV 3</td>
<td>85.5105</td>
<td>81.308119</td>
<td>230.8422</td>
<td>86.60813</td>
<td>277.54</td>
<td>104.0653</td>
<td>41.9434</td>
<td>53.90163</td>
<td>207.3404</td>
<td>18.88773</td>
</tr>
<tr>
<td>CV 4</td>
<td>45.554</td>
<td>695.14654</td>
<td>120.9753</td>
<td>329.22</td>
<td>85.05097</td>
<td>98.38078</td>
<td>41.7637</td>
<td>138.5818</td>
<td>56.63287</td>
<td>692.9271</td>
</tr>
<tr>
<td>CV 6</td>
<td>58.6734</td>
<td>72.38579</td>
<td>145.8185</td>
<td>17.15099</td>
<td>126.1488</td>
<td>106.6444</td>
<td>44.6755</td>
<td>81.10734</td>
<td>95.45583</td>
<td>39.55832</td>
</tr>
<tr>
<td>CV 7</td>
<td>60.3253</td>
<td>1218.8307</td>
<td>103.5804</td>
<td>347.3703</td>
<td>83.11907</td>
<td>90.66787</td>
<td>40.1697</td>
<td>115.7737</td>
<td>48.90935</td>
<td>972.3396</td>
</tr>
<tr>
<td>CV 8</td>
<td>39.6467</td>
<td>58.421456</td>
<td>196.2636</td>
<td>36.97415</td>
<td>195.5502</td>
<td>98.56097</td>
<td>44.6943</td>
<td>82.0453</td>
<td>193.4304</td>
<td>50.32223</td>
</tr>
<tr>
<td>CV 9</td>
<td>16.1404</td>
<td>15.545342</td>
<td>205.9824</td>
<td>16.43928</td>
<td>197.5051</td>
<td>64.82991</td>
<td>41.7203</td>
<td>34.1444</td>
<td>254.2339</td>
<td>31.86496</td>
</tr>
</tbody>
</table>
Figure S1.11: GO case study: reduction in the parameter bounds with the estimated values and their 95% confidence intervals for the first fitting data set.
**Figure S1.12:** GO case study: convergence curve of the final regularised estimation for the first fitting data set.

**Figure S1.13:** GO case study: distribution of local solutions found using the nl2sol local solver fitting to noiseless data.
**Figure S1.14:** GO case study: distribution of local solutions found using the nl2sol local solver, with examples of local solutions and overfitting for the first fitting data set.

**Figure S1.15:** GO case study: violin plots showing the distribution of the NRMSE for the fit and cross-validation for all the data sets considered, both with and without regularisation.
Figure S1.16: GO case study: final regularised fit with uncertainty intervals for the first fitting data set.
Figure S1.17: GO case study: parameter correlation matrix for the final estimated regularised solution for the first fitting data set.
Figure S1.18: GO case study: predicted versus measured values for the first fitting data set.
Figure S1.19: GO case study: normalised residuals for the regularised fit for the first fitting data set.
Figure S1.20: GO case study: sampling from the initial estimation with the new parameter bound box, where the height of said box is the cost cut off for the first fitting data set.
Figure S1.21: GO case study: comparison of the fits with and without regularisation for the first fitting data set.
Figure S1.22: GO case study: comparison of the cross-validation with and without regularisation for the fit to the first fitting data set.

Figure S1.23: GO case study: results of the VisId analysis performed at the regularised solution for the first fitting data set.
References


