

# **Model Validation Guide**

**Release v3.5**

**June 2012**



---

# Model Validation Guide

Release v3.5

June 2012

Copyright © 1997-2012 Process Systems Enterprise Limited

Process Systems Enterprise Limited  
6th Floor East  
26-28 Hammersmith Grove  
London W6 7HA  
United Kingdom  
Tel: +44 20 85630888  
Fax: +44 20 85630999  
WWW: <http://www.psenterprise.com>

## Trademarks

gPROMS is a registered trademark of Process Systems Enterprise Limited ("PSE"). All other registered and pending trademarks mentioned in this material are considered the sole property of their respective owners. All rights reserved.

## Legal notice

No part of this material may be copied, distributed, published, retransmitted or modified in any way without the prior written consent of PSE. This document is the property of PSE, and must not be reproduced in any manner without prior written permission.

## Disclaimer

gPROMS provides an environment for modelling the behaviour of complex systems. While gPROMS provides valuable insights into the behaviour of the system being modelled, this is not a substitute for understanding the real system and any dangers that it may present. Except as otherwise provided, all warranties, representations, terms and conditions express and implied (including implied warranties of satisfactory quality and fitness for a particular purpose) are expressly excluded to the fullest extent permitted by law. gPROMS provides a framework for applications which may be used for supervising a process control system and initiating operations automatically. gPROMS is not intended for environments which require fail-safe characteristics from the supervisor system. PSE specifically disclaims any express or implied warranty of fitness for environments requiring a fail-safe supervisor. Nothing in this disclaimer shall limit PSE's liability for death or personal injury caused by its negligence.

## Acknowledgements

ModelBuilder uses the following third party free-software packages. The distribution and use of these libraries is governed by their respective licenses which can be found in full in the distribution. Where required, the source code will be made available upon request. Please contact [support.gPROMS@psenterprise.com](mailto:support.gPROMS@psenterprise.com) in such a case.

Many thanks to the developers of these great products!

---

**Table 1. Third party free-software packages**

| <b>Software/Copyright</b>                             | <b>Website</b>  | <b>License</b>                  |
|---|---|---------------------------------|
| <b>ANTLR</b>  | <a href="http://www.antlr2.org/">http://www.antlr2.org/</a>   | Public Domain                   |
| <b>Batik</b>  | <a href="http://xmlgraphics.apache.org/batik/">http://xmlgraphics.apache.org/batik/</a>                 | Apache v2.0                     |
| Copyright © 1999-2007 The Apache Software Foundation. |   |                                 |
| <b>BLAS</b>   | <a href="http://www.netlib.org/blas">http://www.netlib.org/blas</a>                                     | BSD Style                       |
| Copyright © 1992-2009 The University of Tennessee.    |   |                                 |
| <b>Boost</b>  | <a href="http://www.boost.org/">http://www.boost.org/</a>   | Boost                           |
| Copyright © 1999-2007 The Apache Software Foundation. |   |                                 |
| <b>Castor</b>   | <a href="http://www.castor.org/">http://www.castor.org/</a>   | Apache v2.0                     |
| Copyright © 2004-2005 Werner Guttman                  |   |                                 |
| <b>Commons CLI</b>                                    | <a href="http://commons.apache.org/cli/">http://commons.apache.org/cli/</a>                             | Apache v2.0                     |
| Copyright © 2002-2004 The Apache Software Foundation. |   |                                 |
| <b>Commons Collections</b>                            | <a href="http://commons.apache.org/collections/">http://commons.apache.org/collections/</a>             | Apache v2.0                     |
| Copyright © 2002-2004 The Apache Software Foundation. |   |                                 |
| <b>Commons Lang</b>                                   | <a href="http://commons.apache.org/lang/">http://commons.apache.org/lang/</a>                           | Apache v2.0                     |
| Copyright © 1999-2008 The Apache Software Foundation. |   |                                 |
| <b>Commons Logging</b>                                | <a href="http://commons.apache.org/logging/">http://commons.apache.org/logging/</a>                     | Apache v1.1                     |
| Copyright © 1999-2001 The Apache Software Foundation. |   |                                 |
| <b>Crypto++ (AES/Rijndael and SHA-256)</b>            | <a href="http://www.cryptopp.com/">http://www.cryptopp.com/</a>   | Public Domain                   |
| Copyright © 1995-2009 Wei Dai and contributors.       |   |                                 |
| <b>Fast MD5</b>                                       | <a href="http://www.twmacinta.com/myjava/fast_md5.php">http://www.twmacinta.com/myjava/fast_md5.php</a> | LGPL v2.1                       |
| Copyright © 2002-2005 Timothy W Macinta.              |   |                                 |
| <b>HQP</b>  | <a href="http://hqp.sourceforge.net/">http://hqp.sourceforge.net/</a>                                   | LGPL v2                         |
| Copyright © 1994-2002 Ruediger Franke.                |   |                                 |
| <b>Jakarta Regexp</b>                                 | <a href="http://jakarta.apache.org/regexp/">http://jakarta.apache.org/regexp/</a>                       | Apache v1.1                     |
| Copyright © 1999-2002 The Apache Software Foundation. |   |                                 |
| <b>JavaHelp</b>                                       | <a href="http://javahelp.java.net/">http://javahelp.java.net/</a>                                       | GPL v2 with classpath exception |
| Copyright © 2011, Oracle and/or its affiliates.       |   |                                 |
| <b>JXButtonPanel</b>                                  | <a href="http://swinghelper.dev.java.net/">http://swinghelper.dev.java.net/</a>                         | LGPL v2.1 (or later)            |
| Copyright © 2011, Oracle and/or its affiliates.       |   |                                 |
| <b>LAPACK</b>   | <a href="http://www.netlib.org/lapack/">http://www.netlib.org/lapack/</a>                               | BSD Style                       |
| <b>libodbc++</b>                                      | <a href="http://libodbcxx.sourceforge.net/">http://libodbcxx.sourceforge.net/</a>                       | LGPL v2                         |

| Software/Copyright  | Website   | License                         |
|---|---|---------------------------------|
| Copyright © 1999-2000 Manush Dodunekov <manush@stendahls.net> |   |                                 |
| Copyright © 1994-2008 Free Software Foundation, Inc.          |   |                                 |
| <b>lp_solve</b>   | <a href="http://lpsolve.sourceforge.net/">http://lpsolve.sourceforge.net/</a>             | LGPL v2.1                       |
| Copyright © 1998-2001 by the University of Florida.           |   |                                 |
| Copyright © 1991, 2009 Free Software Foundation, Inc.         |   |                                 |
| <b>MiGLayout</b>  | <a href="http://www.miglayout.com/">http://www.miglayout.com/</a>                         | BSD                             |
| Copyright © 2007 MiG InfoCom AB.                              |   |                                 |
| <b>Netbeans</b>   | <a href="http://www.netbeans.org/">http://www.netbeans.org/</a>                           | SPL                             |
| Copyright © 1997-2007 Sun Microsystems, Inc.                  |   |                                 |
| <b>omniORB</b>  | <a href="http://omniorb.sourceforge.net/">http://omniorb.sourceforge.net/</a>             | LGPL v2                         |
| Copyright © 1996-2001 AT&T Laboratories Cambridge.            |   |                                 |
| Copyright © 1997-2006 Free Software Foundation, Inc.          |   |                                 |
| <b>TimingFramework</b>  | <a href="http://timingframework.dev.java.net/">http://timingframework.dev.java.net/</a>   | BSD                             |
| Copyright © 1997-2008 Sun Microsystems, Inc.                  |   |                                 |
| <b>VecMath</b>  | <a href="http://vecmath.dev.java.net/">http://vecmath.dev.java.net/</a>                   | GPL v2 with classpath exception |
| Copyright © 1997-2008 Sun Microsystems, Inc.                  |   |                                 |
| <b>Wizard Framework</b>                                       | <a href="http://wizard-framework.dev.java.net/">http://wizard-framework.dev.java.net/</a> | LGPL                            |
| Copyright © 2004-2005 Andrew Pietsch.                         |   |                                 |
| <b>Xalan</b>  | <a href="http://xml.apache.org/xalan-j/">http://xml.apache.org/xalan-j/</a>               | Apache v2.0                     |
| Copyright © 1999-2006 The Apache Software Foundation.         |   |                                 |
| <b>Xerces-C</b>   | <a href="http://xerces.apache.org/xerces-c/">http://xerces.apache.org/xerces-c/</a>       | Apache v2.0                     |
| Copyright © 1994-2008 The Apache Software Foundation.         |   |                                 |
| <b>Xerces-J</b>   | <a href="http://xerces.apache.org/xerces2-j/">http://xerces.apache.org/xerces2-j/</a>     | Apache v2.0                     |
| Copyright © 1999-2005 The Apache Software Foundation.         |   |                                 |

This product includes software developed by the Apache Software Foundation, <http://www.apache.org/>.

gPROMS also uses the following third party commercial packages:

- **FLEXnet Publisher** software licensing management from Acreesso Software Inc., <http://www.acresso.com/>.
- **JClass DesktopViews** by Quest Software, Inc., <http://www.quest.com/jclass-desktopviews/>.
- **JGraph** by JGraph Ltd., <http://www.jgraph.com/>.

---

# Table of Contents

|   |    |
|---|----|
| 1. Introduction .....   | 1  |
| Experiments in gPROMS .....                                     | 1  |
| Experiment operating policy .....                               | 1  |
| Measurements, Sensors and Variance models .....                 | 2  |
| gPROMS Model requirements .....                                 | 3  |
| 2. Parameter Estimation in gPROMS .....                         | 5  |
| What is Parameter Estimation? .....                             | 5  |
| Performed Experiments .....                                     | 6  |
| General information .....                                       | 7  |
| Controls .....  | 9  |
| Measured Data .....   | 10 |
| Simulating a Performed Experiment .....                         | 13 |
| The Parameter Estimation entity .....                           | 14 |
| Parameters to be estimated .....                                | 15 |
| Experiments and measurements .....                              | 16 |
| Execution of a Parameter Estimation activity .....              | 18 |
| Running the Parameter Estimation activity .....                 | 19 |
| Monitoring the activity progress .....                          | 20 |
| Tuning the Parameter Estimation execution .....                 | 21 |
| Analysing the Parameter Estimation results .....                | 24 |
| Variable Trajectories .....                                     | 24 |
| The Comprehensive report for Parameter Estimation .....         | 26 |
| The summary report file .....                                   | 27 |
| Statistical analysis in Excel .....                             | 30 |
| A Parameter Estimation example .....                            | 37 |
| 3. Experiment Design in gPROMS .....                            | 43 |
| What is Experiment Design? .....                                | 43 |
| Underlying mathematics of Experiment Design .....               | 43 |
| Experiments to be designed .....                                | 45 |
| General information .....                                       | 46 |
| Controls .....  | 48 |
| Constraints .....   | 49 |
| Measurements and sensors .....                                  | 51 |
| The Experiment Design entity .....                              | 52 |
| General information .....                                       | 53 |
| Experiments and measurements .....                              | 54 |
| Execution of Experiment Design activities .....                 | 55 |
| Variable Trajectories .....                                     | 56 |
| The Comprehensive report for Experiment Design .....            | 57 |
| The Experiment Design text report and statistics files .....    | 59 |
| Experiment SCHEDULE files .....                                 | 63 |
| Designed Experiment templates .....                             | 65 |
| Controlling the execution of Experiment Design activities ..... | 65 |
| An Experiment Design example .....                              | 66 |
| 4. A Model Validation Example .....                             | 69 |
| Reaction kinetics .....   | 69 |
| Experiment description .....                                    | 70 |
| Parameter Estimation .....                                      | 71 |
| Experiment Design .....   | 71 |

---

## List of Figures

|  |    |
|--|----|
| 1.1. Iterative model development and validation procedure. ....  | 1  |
| 2.1. Experiments Performed entity editor. ....   | 7  |
| 2.2. Specifying the Process entity that will be used in conjunction with the experimental data. ....                                     | 8  |
| 2.3. Specifying the Saved Variable Sets that will be used to facilitate initialisation. ....   | 8  |
| 2.4. Specifying the initial conditions for the relevant differential variables in an experiment. ....                                    | 9  |
| 2.5. Specifying a piecewise linear profile for the behavior of a control variable during the experiment. ....                            | 10 |
| 2.6. Tabulation of the measured data for an experimental run. ....   | 11 |
| 2.7. Defining the measurement variance model and its associated parameters. ....   | 12 |
| 2.8. Simulating an experiment ....   | 13 |
| 2.9. Text comparison of an original Process and an automatically generated Process. ....   | 14 |
| 2.10. Parameter Estimations Entity editor ....   | 15 |
| 2.11. Parameters to be estimated tab ....  | 16 |
| 2.12. Specifying which experiments to include in a Parameter Estimation problem ....   | 17 |
| 2.13. Setting a Variance model for a Sensor group ....   | 17 |
| 2.14. Sensor groups and Variance models ....   | 18 |
| 2.15. Executing a Parameter Estimation activity from the Project Tree ....   | 19 |
| 2.16. The Parameter Estimation activity execution control dialog ....  | 20 |
| 2.17. Monitoring the progress of a Parameter Estimation activity. ....   | 21 |
| 2.18. Predicted trajectory of a measured variable in an experiment. ....   | 25 |
| 2.19. Parameter Estimation comprehensive report. ....  | 26 |
| 2.20. Parameter Estimation measurement plots ....  | 27 |
| 2.21. Browsing for the Batch Reactor Parameter Estimation example in the gPROMS distribution<br>directory. ....                          | 38 |
| 2.22. Batch reactor. ....  | 38 |
| 2.23. Comparison of experimental composition data with model predictions based on nominal values of<br>$k_0$ , $E$ and $\Delta h_r$ .... | 41 |
| 2.24. Comparison of experimental temperature data with model predictions based on nominal values of<br>$k_0$ , $E$ and $\Delta h_r$ .... | 41 |
| 3.1. Interpretation of the different design criteria. ....   | 45 |
| 3.2. Experiments to be designed entity editor ....   | 46 |
| 3.3. Specifying the Process entity that will be used in conjunction with the experimental data. ....                                     | 47 |
| 3.4. Specifying the initial conditions for the relevant differential variables in an experiment to be<br>designed. ....                  | 48 |
| 3.5. Specifying the bounds on a piecewise constant control profile that is to be optimally determined. ....                              | 49 |
| 3.6. Experiment Design - imposing design constraints ....  | 51 |
| 3.7. The measurements and sensors tab ....   | 51 |
| 3.8. Experiment Design entity editor. ....   | 53 |
| 3.9. Specifying the objective function for the Experiment Design problem. ....   | 53 |
| 3.10. Experiments Performed entities that can be included in the estimation problem. ....  | 55 |
| 3.11. Executing an Experiment Design activity. ....  | 56 |
| 3.12. Predicted trajectory of a measured variable in an experiment. ....   | 57 |
| 3.13. Comprehensive Experiment Design report. ....   | 58 |
| 3.14. Parameter Estimation measurement plots ....  | 59 |
| 3.15. Browsing for the Experiment Design example in the gPROMS distribution directory. ....  | 67 |
| 3.16. Fed-batch fermentation ....  | 67 |
| 4.1. Batch Reactor Flowsheet ....  | 69 |
| 4.2. Reaction Network ....   | 70 |

---

## List of Tables

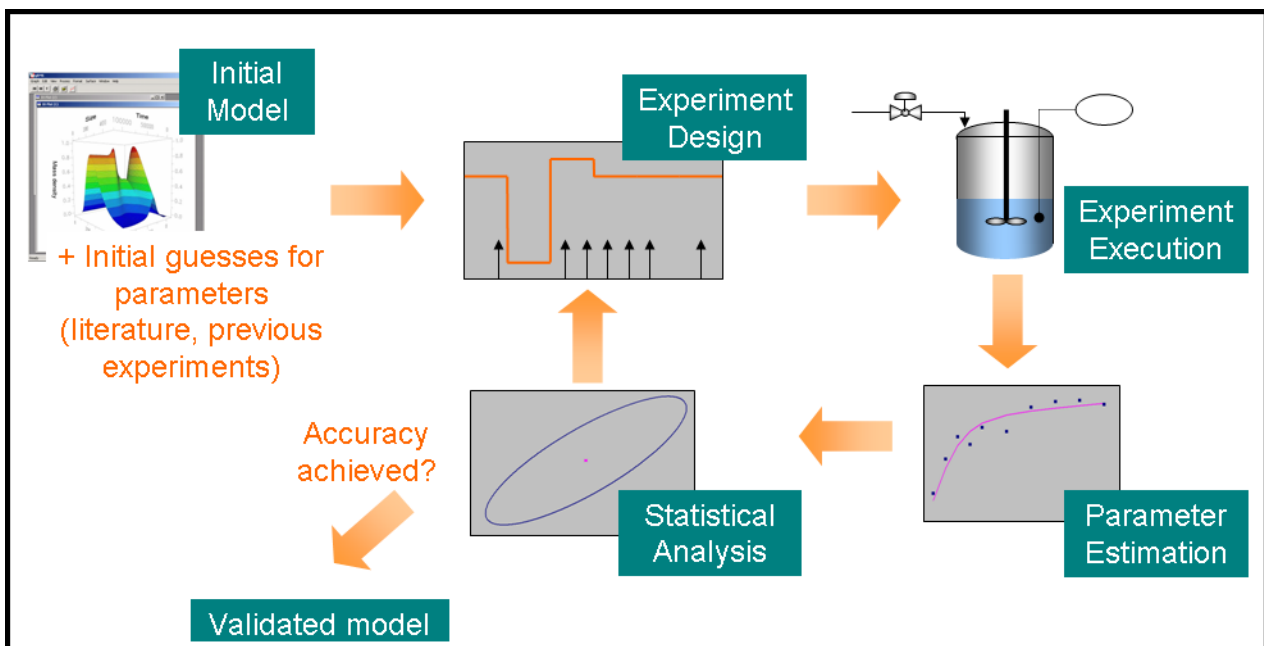
|  |    |
|--|----|
| 1. Third party free-software packages .....                                | 3  |
| 1.1. Statistical variance models. ....                                     | 2  |
| 2.1. Objective function symbol definitions. ....                           | 6  |
| 2.2. Parameters in batch reactor Model. ....                               | 39 |
| 2.3. Variables in batch reactor Model. ....                                | 39 |
| 2.4. Conditions for batch reactor experiments #1 to #4 .....               | 40 |
| 2.5. Data from the batch reactor experiments .....                         | 40 |
| 3.1. Symbol definitions. ....  | 44 |
| 3.2. Outputlevel .....   | 65 |
| 3.3. Parameters and design variables in fed-batch fermentation model. .... | 68 |
| 3.4. Conditions for the fed-batch fermentation experiment .....            | 68 |
| 4.1. Components of Phenol hydro-deoxidation .....                          | 69 |
| 4.2. Initial conditions of experiments .....                               | 70 |
| 4.3. Measurements for experiment 1 ( $T = 523.15 \text{ K}$ ) .....        | 71 |
| 4.4. Measurements for experiments 2 ( $T = 533.15 \text{ K}$ ) .....       | 71 |

# Chapter 1. Introduction

gPROMS employs a rigorous optimisation-based approach for Model Validation by offering *Parameter Estimation* capabilities (i.e. fitting model parameters to experimental data) together with *Experiment Design* capabilities (i.e. determining the best set of experiments to perform that will give the most precise parameter estimates).

Model Validation involves an iterative procedure as shown in the figure below. Once you perform the experiments that have been designed and collect the appropriate measurement data, these can be used to generate better estimates of the parameters. If necessary, the model with the updated parameters can then be used to design further experiment(s) and the cycle is repeated until you are satisfied that the model is validated.

**Figure 1.1. Iterative model development and validation procedure.**



## Experiments in gPROMS

Experiments are used to improve the understanding of processes and create accurate models. The quality of information generated by experiments depends strongly on the experimental conditions as well as what is measured and when it is measured. In gPROMS we can consider:

1. The processing of data from experiments to estimate the values of unknown model parameters - this is *Parameter Estimation*
2. The use of models to design experiments for the purposes of estimating the values of the unknown model parameters - this is *Experiment Design*

To use experiments and experimental data in gPROMS we need to be able to represent this information mathematically; there are two fundamental aspects to this

1. A description of the experiment operating policy
2. A description of the experimental measurements

## Experiment operating policy

In gPROMS an experimental recipe consists of the following information



1. The experiment duration
2. The initial conditions in the experimental equipment (such as the initial temperature and amount of material in the experimental equipment)
3. The time variation of the controlled inputs to the system (such the heat input to the system)
4. The values of time-invariant quantities (such as the diameter of the experimental equipment)

Given this information, a model of the experimental equipment can be used to predict the behaviour of the measured quantities. Note that for dynamic experiments all of the above information is required to define an experiment; whereas for steady-state experiments only the values of time-invariant quantities need to be provided. In Parameter Estimation the above information is known; whereas the purpose of Experiment Design is to determine this information.

In gPROMS in order to specify the time variation of controlled inputs we consider the time horizon of the experiment to be broken into a number of time intervals. A time-varying control is simply an ASSIGNED variable that is adjusted during an experiment. The user can specify the variation in a Variable value using one of the following mechanisms:

- Piecewise constant controls - the quantity holds a different constant value during each time interval during the experiment.
- Piecewise linear controls - the quantity varies linearly (from a start value to an end value) in a specified time interval during the experiment.

To simplify specification of the problem and the problem size the same control time intervals are used for all piecewise constant and piecewise linear controls within an experiment..

## Measurements, Sensors and Variance models

Measurements of key quantities (corresponding to Model Variables) are taken throughout the duration of an experiment, each measurement consists of three pieces of information:

- The value of the measured quantity
- The time the measurement was taken
- The uncertainty to which the measurement is known - this can be expressed as the statistical **standard deviation or variance** of the measurement.

Experimental measurements are taken using **sensors**: the uncertainty of the measurement is a property of the measurement technique associated with the sensor. Of course, it is often the case that the same sensor measures more than one variable. For example, a single composition sensor may measure the composition of multiple components in a given stream. Or, a given type of thermocouple may be used in a number of places on a plant to measure temperature.

When solving a Model Validation problem all measured Variables are associated with a sensor. The variance model of a given sensor comprises information associated with the variance of the error of the measurement produced by the sensor. The errors of the measurements are assumed to be statistically independent and normally distributed with zero mean. There are several types of sensor variance models. These can be considered to take the general form:

$$\sigma^2 = \sigma^2(z, B)$$

where  $z$  is the model prediction of the measured quantity and  $B$  is a set of parameters. The table below shows the variance models supported by gPROMS.

**Table 1.1. Statistical variance models.**

| Variance model (gPROMS Keyword)       | Mathematical Description |
|---------------------------------------|--------------------------|
| Constant variance (CONSTANT_VARIANCE) | $\sigma^2 = \omega^2$    |

| Variance model (gPROMS Keyword)                            | Mathematical Description                               |
|--|--|
| Constant relative variance<br>(CONSTANT_RELATIVE_VARIANCE) | $\sigma^2 = \omega^2 \cdot (z^2 + \varepsilon)$        |
| Heteroscedastic variance (HETEROSCEDASTIC)                 | $\sigma^2 = \omega^2 \cdot (z^2 + \varepsilon)^\gamma$ |
| Linear variance (LINEAR_VARIANCE)                          | $\sigma^2 = (\alpha z + \beta)^2 + \varepsilon$        |

The following points should be noted:

- The set of parameters, B , in the general statistical variance model comprises the parameters  $\omega$  and  $\gamma$ , as appropriate.
- $\varepsilon$  is a small but non-zero constant that ensures that the variance is still defined for predicted values that are equal to zero or very small.
- If  $\gamma = 0$  in the HETEROSCEDASTIC variance model, then this reduces to the CONSTANT\_VARIANCE model.
- If  $\gamma = 1$  in the HETEROSCEDASTIC variance model, then this reduces to the CONSTANT\_RELATIVE\_VARIANCE model.

As the estimation of the parameters and the statistical analysis of the results depend strongly on the given or estimated standard deviations of the measurement errors, attention should be paid to the formulation of the variance model and the values of the respective variance model parameters. You are encouraged to specify the variance model and the values of the according parameters as precisely as possible. gPROMS will determine optimal estimates for the values of the variance model parameters within the bounds that you specify.

## gPROMS Model requirements

Both Parameter Estimation and Experiment Design problems in gPROMS require a working gPROMS process model that runs robustly for steady-state and/or dynamic simulation.

Particular focus should be given to the following parts of the Process Entity:

- The **Assign** specifications mark certain system variables as fixed for the purposes of dynamic simulation. As far as Parameter Estimation and Experiment Design is concerned, some of these variables will belong to one of the following categories:
  - time-varying control variables;
  - time-invariant control parameters; or
  - unknown parameters to be estimated.

It is very important to note that, for the purposes of model validation, unknown parameters are declared as Variables in gPROMS Model entities, and **not** as Parameters<sup>1</sup>

- The **Initial** specifications in the Process will be used as the **default** initial conditions for the experiments. Some or all of the Initial specifications may then be overridden by specifying a value in an experiment.

In the Process Entity, Initial specifications can be given in the form of a general equation. However to specify Initial conditions in your experiments, corresponding Initial specifications must appear in the Process in the restricted form:

*VariableName* = *ConstantExpression*;

<sup>1</sup>This is consistent with the general gPROMS rule that Parameters can *never* be the result of any gPROMS calculation. Of course, once the model validation problem has been solved successfully and their values have been found, they can either be declared as Parameters (Set to the estimated value) or be left as Variables (Assigned to the estimated value)

where *VariableName* is the name of a differential or algebraic variable, and *ConstantExpression* is either a numerical value or an expression consisting of only parameters and numerical values.

- The gPROMS solvers for Parameter Estimation and Experiment Design are configured in the SOLUTIONPARAMETERS section of a Process entity - refer to Controlling the Execution of Parameter Estimation activities and Controlling the Execution of Experiment activities
- Any **Schedule** specification in the Process is ignored for the purposes of Parameter Estimation and Experiment Design. This also means that any Intrinsic Tasks<sup>2</sup> used by your Models will not be executed.

Experience indicates that most of the effort in defining Parameter Estimation and Experiment Design problems is, in fact, incurred in the construction of a robust model of your process. This will probably be exactly the same model as that used for dynamic simulations within gPROMS.

Therefore, it is worth investing some effort in ensuring that the model behaves properly over the *entire* range of possible parameter values. In particular, you should check that the differential and algebraic variables  $x$  and  $y$  remain within any specified bounds even for extreme values of the control variables and the parameters to be estimated.

---

<sup>2</sup>See the section "Defining Tasks" in the Model Developer Guide*i*

---

# Chapter 2. Parameter Estimation in gPROMS

gPROMS can be used to perform *Parameter Estimation* for complex models using both dynamic and steady-state experimental data. This chapter covers the following issues:

- a description of what a Parameter Estimation problem is,
- the Performed Experiment and Parameter Estimation entities that are used to specify Parameter Estimation problems in gPROMS,
- running and controlling the execution of a Parameter Estimation activity, and
- how to analyse the results of a Parameter Estimation run.

Explanatory screen shots are taken from a batch reactor example included in the gPROMS distribution. Some basic familiarity with the gPROMS language and concepts is assumed in this guide.

## What is Parameter Estimation?

A detailed gPROMS process model is constructed from equations describing the physical and chemical phenomena that take place in the system. These equations usually involve parameters that can be adjusted to make the model predictions match observed reality. Examples of model parameters include reaction kinetic constants, heat transfer coefficients, distillation stage efficiencies, constants within physical property correlations, and so on. The more accurate these parameters are, the closer the model response is to reality.

The process of fitting these parameters to laboratory or plant data is called Parameter Estimation. gPROMS contains powerful, state-of-the-art Parameter Estimation capabilities that have been applied successfully to a wide range of problems. Key features are:

- Multiple parameters occurring in dynamic or steady-state models may be estimated simultaneously. Nonlinear models of arbitrary size and complexity – including multi-unit flowsheets – may be used.
- Data from both dynamic and steady-state experiments may be used.
- The results of the estimation are subjected to extensive statistical analysis.

Parameter Estimation in gPROMS is based on the Maximum Likelihood formulation which provides simultaneous estimation of parameters in both:

- the physical model of the process
- the variance model of the measuring instruments - the **Sensor**, which can be:
  - constant variance (e.g. a thermocouple with an accuracy of +/- 1K)
  - constant relative variance (e.g. a composition analyser with an error of +/- 2%)
  - heteroscedastic variance, combining both of the above.

When solving a Maximum Likelihood Parameter Estimation problem, gPROMS attempts to determine values for the uncertain physical and variance model parameters,  $\theta$ , that maximise the probability that the mathematical model will predict the measurement values obtained from the experiments. Assuming independent, normally distributed measurement errors,  $\epsilon_{ijk}$ , with zero means and standard deviations,  $\sigma_{ijk}$ , this maximum likelihood goal can be captured through the following objective function:

$$\Phi = \frac{N}{2} \ln(2\pi) + \frac{1}{2} \min_{\theta} \left\{ \sum_{i=1}^{NE} \sum_{j=1}^{NV_i} \sum_{k=1}^{NM_{ij}} \left[ \ln(\sigma_{ijk}^2) + \frac{(\tilde{z}_{ijk} - z_{ijk})^2}{\sigma_{ijk}^2} \right] \right\},$$

where the symbols have the following definitions:

**Table 2.1. Objective function symbol definitions.**

|                   |  |
|-------------------|--|
| $N$               | Total number of measurements taken during all the experiments.   |
| $\theta$          | Set of model parameters to be estimated. The acceptable values may be subject to given lower and upper bounds, i.e. $\theta^l \leq \theta \leq \theta^u$ . |
| $NE$              | Number of experiments performed.   |
| $NV_i$            | Number of variables measured in the $i$ th experiment.   |
| $NM_{ij}$         | Number of measurements of the $j$ th variable in the $i$ th experiment.  |
| $\sigma_{ijk}^2$  | Variance of the $k$ th measurement of variable $j$ in experiment $i$ . This is determined by the measured variable's variance model.                       |
| $\tilde{z}_{ijk}$ | $k$ th measured value of variable $j$ in experiment $i$ .  |
| $z_{ijk}$         | $k$ th (model-)predicted value of variable $j$ in experiment $i$ .   |

## Performed Experiments

A Parameter Estimation problem makes use of data gathered from a set of experiments. Each experiment is characterised by a set of conditions under which it is performed, namely:

- the overall duration;
- the initial conditions,  $v$ ;
- the variation of the control variables,  $u(t)$ ; and
- the values of the time-invariant parameters,  $p$  (these could be things like vessel volume and so on).

The Experiments Performed entity is used to specify the full details of an experiment. To create a new Performed Experiments entity, do the following:

- Pull-down the Entity menu from the top pane in gPROMS ModelBuilder.
- Click on New Entity. A dialog box will appear.
- Choose Performed Experiment for the Entity type and fill in the Name field.

An Experiments Performed entity editor window will open up in the right hand pane, as shown in the figure below.

**Figure 2.1. Experiments Performed entity editor.**

PERFORMED EXPERIMENT C13T3 (ReactorEst)

Process:

Saved variable set(s):

Initial conditions

☒ Dynamic ☐ Steady state

| Variable name  | Initial value |
|----------------|---------------|
| R101.HOLDUP(1) | 100.0         |
| R101.HOLDUP(2) | 100.0         |
| R101.HOLDUP(3) | 1.0           |
| R101.HOLDUP(4) | 1.0           |
| R101.T         | 293.0         |
| <new>          |               |

Delete row

General Controls Measured data gPROMS language Properties

The Performed experiment entity editor has the following tabs for entering information:

- The General tab - a form for defining the Process entity to use, the Saved variable sets to help with initialisation and the initial conditions for the experiment.
- The Controls tab - a form for specifying the values of the time-invariant, piecewise constant and piecewise linear variables that are adjusted during the course of the experiment.
- The Measured data tab - a form for entering the measured data and sensor used. There is also the option of specifying a measurement variance model if you wish to have a specific model or set of parameters for this particular experiment.
- The gPROMS language tab - this provides a convenient means of reviewing all the information regarding the performed experiment in a text-based format. You are not advised to use this tab to enter data.
- Properties (a record of when the entity was created, last edited as well as the default tab to open on)

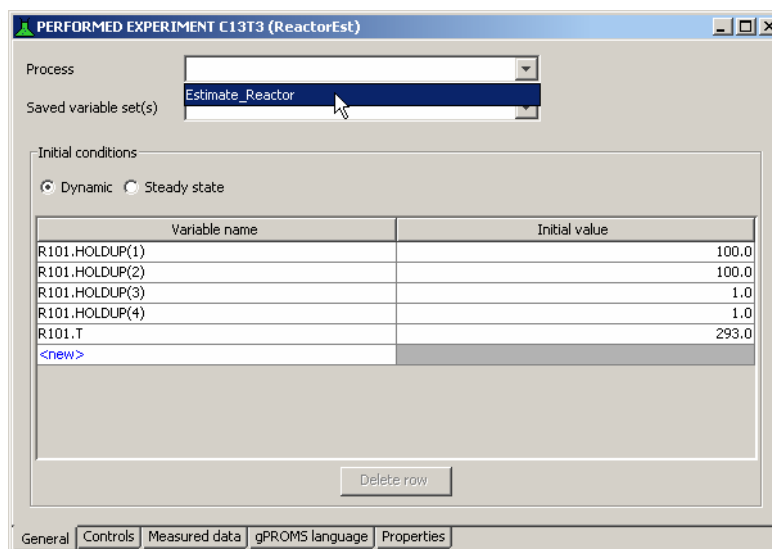
In addition, a Performed experiment can be automatically converted into a Process and simulated.

## General information

In order to specify the Process entity to use for the Parameter Estimation, click on the General tab in the Experiments Performed entity editor and then choose the desired Process from the drop down menu at the top of the window (see figure below) - you can select a Process Entity from the same project as that containing this Experiment entity or any projects cross-referenced by the project.<sup>1</sup>

<sup>1</sup>For backwards compatibility with v2.3, it is valid for the Process field to remain blank. In this case, the Parameter Estimation and/or Experiment Design activities will look for a Process with the same name as the activity being run. However, if you do not set the Process field, assisted pathname completion will not work within this editor.

**Figure 2.2. Specifying the Process entity that will be used in conjunction with the experimental data.**

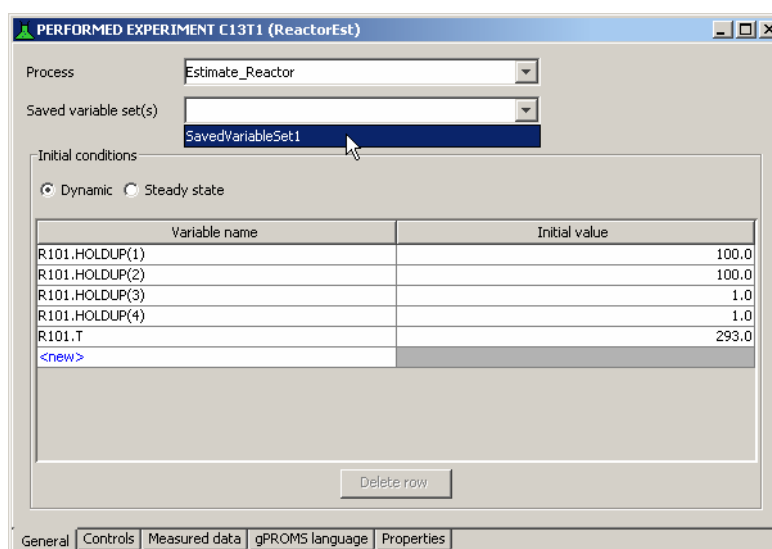


In some situations, initialisation is helped by specifying a Saved variable set. This can be done by choosing the desired Saved file from the drop down menu shown below, or it may be typed in manually.

**Advanced notes:**

1. A Saved variable set specified in the Performed Experiment will be applied **in addition** to specifications given in the PRESET section of the Process. Each entry of the Saved variable set file(s) specified in the Performed Experiment supersede any specification given in the Process.
2. gPROMS will also accept a path to a .gSTORE file located in the file system (as opposed to being embedded in the project).
3. Multiple files may be listed, separated using a comma.

**Figure 2.3. Specifying the Saved Variable Sets that will be used to facilitate initialisation.**



You have the option to provide different initial conditions to those contained in the Process. An experiment can have steady-state or dynamic initial conditions. If you to choose steady-state initial conditions, then simply check

the steady-state button in the Initial conditions section of the window. For dynamic initial conditions, check the dynamic button and enter information in the table underneath. Start by clicking the <new> cell in the table and type the full pathname of the gPROMS differential variable that you wish to specify an initial condition for. You **MUST** only provide initial conditions for variables that are given initial values directly in the Process.

When entering a variable name in the left hand column, *pathname completion* can be activated by pressing **Ctrl+Space**. Thereafter simply type in its value in the right hand column of the table as shown below.

**Figure 2.4. Specifying the initial conditions for the relevant differential variables in an experiment.**

PERFORMED EXPERIMENT C13T1 (ReactorEst)

Process: Estimate\_Reactor

Saved variable set(s):

Initial conditions:

☒ Dynamic ☐ Steady state

| Variable name  | Initial value |
|----------------|---------------|
| R101.HOLDUP(1) | 100.0         |
| R101.HOLDUP(2) | 100.0         |
| R101.HOLDUP(3) | 1.0           |
| R101.HOLDUP(4) | 1.0           |
| R101.T         | 293.0         |
| <new>          |               |

Delete row

General Controls Measured data gPROMS language Properties

## Controls

An experiment control is a variable that is adjusted during an experiment (such as a heating policy in a reactor). Experiment controls therefore form part of the "recipe" for running the experiment. The user can specify the variation in a Variable value using one of three different mechanisms:

- Time-invariant controls: provide a single variable value, this quantity will be the same for the duration of the experiment. Appropriate for steady-state and dynamic experiments.
- Piecewise constant controls: provide multiple variable values, the quantity holds a different constant value for specified time intervals during the experiment. Appropriate for dynamic experiments only.
- Piecewise linear controls: provide multiple variable values, the quantity varies linearly (from a start value to an end value) in a specified time interval during the experiment. Appropriate for dynamic experiments only.

Note that any Variables that are Assigned values in a Process but are not specified as controls in the experiment will retain their Assigned values, whether these are constants or functions of TIME.

The Control *interval durations* are synchronised for all piecewise constant and piecewise linear controls and adding or deleting control intervals therefore affects *all* piecewise control variables.

In order to define the controls, click on the Controls tab in the Experiments Performed entity editor and do the following:

- Click the <new> cell in the Control variables table at the top of the window and type the full pathname of the gPROMS control variable. This variable should have been ASSIGNED in the Process entity

### Tip

*Pathname completion* can be activated by pressing **Ctrl+Space**.



- Choose from the drop down list in the right-hand column to specify whether the control variable is time-invariant, piecewise constant or piecewise linear.
- Repeat the above for each new control variable you wish to use.

Having chosen the format for the time-varying behavior of the control variable, select the relevant control by clicking on it in the top pane, and go to the lower part of the window to enter the details of its behavior during the experiment. Time-invariant controls simply need a value. Piece-wise constant controls need a duration for each interval<sup>2</sup>, together with the corresponding value of the control. Piecewise linear controls need the interval duration, together with starting and final values for the control variable during each interval (as illustrated in the following figure).

To duplicate or delete an entire interval, select it and click on the Duplicate or Delete buttons at the bottom of the window.

**Figure 2.5. Specifying a piecewise linear profile for the behavior of a control variable during the experiment.**

| Interval duration | Start time | End time | Start value | End value |
|-------------------|------------|----------|-------------|-----------|
| 100.0             | 0.0        | 100.0    | 505000.0    | 505000.0  |
| 300.0             | 100.0      | 400.0    | 0.0         | 0.0       |
| <new>             |            |          |             |           |

## Measured Data

During an experiment, measurements of key variables are taken using sensors. The measured data for an experimental run should be entered in the table found on the Measured data tab of the Experiments Performed entity (shown below).

<sup>2</sup>When you enter the interval duration gPROMS automatically determines the start and end times of that interval, but the relevant cells are greyed out and you cannot edit them.

**Figure 2.6. Tabulation of the measured data for an experimental run.**

| Time  | Variable Name | Sensor        | Variance model |          |
|-------|---------------|---------------|----------------|----------|
|       | R101.C(1)     | <unspecified> | <unspecified>  | <new>    |
|       | R101.C(3)     | <unspecified> | <unspecified>  |          |
|       | R101.T        | <unspecified> | <unspecified>  |          |
| 10.0  |               |               |                | 3746.911 |
| 20.0  |               |               |                | 2518.6   |
| 30.0  |               |               |                | 1504.12  |
| 40.0  |               |               |                | 883.7833 |
| 50.0  |               |               |                | 548.0505 |
| 60.0  |               |               |                | 348.6161 |
| 70.0  |               |               |                | 239.5115 |
| 80.0  |               |               |                | 175.4503 |
| 90.0  |               |               |                | 135.0508 |
| 100.0 |               |               |                | 100.9699 |
| 110.0 |               |               |                | 87.2978  |

The Measured data table accepts the following information: the gPROMS name of the measured variable, a name for the sensor used (*optional*), a description of the measurement variance model for this sensor (*optional*) and the measured values at discrete points in time.

To add the experimental results for a particular variable, do the following:

- Click the <new> cell in the Variable name part of the table at the top of the window and type the full pathname of the gPROMS variable.

## Tip

*Pathname completion* can be activated by pressing **Ctrl+Space**.

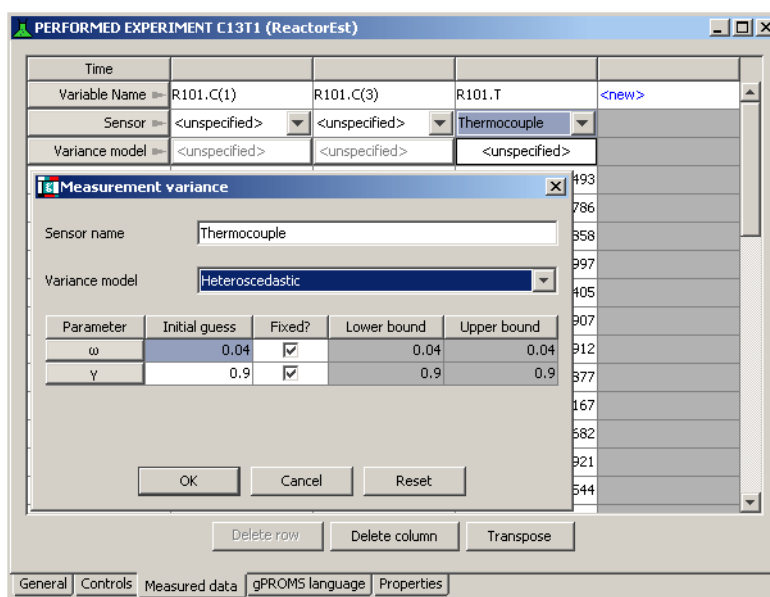
- If you wish to give a name to the sensor that was used for the measurement, then do so in the corresponding part of the table, and press **Enter**.
- If you wish to specify the variance model information for this particular sensor then click on the Variance model cell in the table. When you do so a dialog box will open. You need to then do the following:
  - Select the desired variance model from a drop down list and enter the initial guess and the lower and upper bounds for the relevant statistical parameters that will be used in the solution of the maximum likelihood parameter estimation problem [6].
  - If you wish to estimate the value of a given variance model parameter, then *uncheck* the Fixed value? box and enter appropriate lower and upper bounds.
  - Click OK to close the dialog box and apply the changes.
- To add the measurement data points, click on the <new time> cell in the table on the Measured data tab. Enter the time and the corresponding value for the measured variable. It is possible to copy and paste data into the table from an Microsoft Excel™ spreadsheet. To do so, first make sure that your Excel data has the columns in the same order as in the gPROMS table. Then do the following:
  - select the numerical data in Excel and copy it (Ctrl+C). Do not select the column headings in Excel
  - go to the gPROMS table and click on the cell containing <new time>. Paste the numerical Excel data (Ctrl+V).

## Tip

Measured variables do not need to have a value recorded at every time point. It is valid to leave cells in the table blank.

- To clear a single cell quickly, select the cell and press the **space** key.
- To delete an entire row or column, select a cell in the desired row or column, then click the Delete row or Delete column button below the table (as appropriate).
- It is possible to transpose the table using the transpose button below the table.

**Figure 2.7. Defining the measurement variance model and its associated parameters.**



## When to assign a Sensor name

As mentioned previously, assigning a Sensor name to a measured variable is optional. If you are putting together your first experiments, you may wish to simply leave these <unspecified>. In this case, the measured variables will appear "by name" in the Parameter Estimation and/or Experiment Design editors (and reports).

Once you assign a sensor name to more than one variable in an experiment, those variables share the sensor. Changing the sensor's variance model will affect all variables with this sensor, but **only within the experiment**. Sensor names are local to an experiment, so a sensor specified in one experiment has no effect on any other experiment.

Once a sensor name has been given, that name will appear in the Parameter Estimation and Experiment Design editors, and in some places in the reports it will be used instead of the variable name.

Specifying a sensor name is useful:

- as a record that the variables were measured by the same physical measuring device, in the same batched measurement or taken under the same calibration of the measuring equipment,
- to help manage a large number of measured variables. When the Performed Experiment is added into a Parameter Estimation or Experiment Design entity, the variables are automatically grouped by their sensor name, reducing the number of items listed in the sensor group editor and helping to sort them.

## When to set a Variance model

One approach is to leave the variance model <unspecified>, which allows all the variance model information to be given in a central location. It reduces the duplication of the variance model information. For example, you may intend to use the same variance specifications across a large number of experiments.

Even if a sensor name has been given, the variance model can be left <unspecified>. This effectively "links" the measured variables together, but omits the specification of the variance model (which must be given later in the a Parameter Estimation or Experiment Design entity).

The alternative approach is to record the variance model information within each experiment. This provides a more complete record of the conditions under which the experiment was performed. This is most useful if the variance information differs between experimental runs.

Note that a variance model specification given in the Parameter Estimation or Experiment Design editor will **override** a variance model specified in the Performed Experiments. This is an important ordering of precedence. This allows, for example, different Parameter Estimation activities to be defined, each having a different variance model specifications, but *without* requiring any changes to be made to the Performed Experiments themselves. This helps to maintain the integrity of the Performed Experiment entity as a record of experimental conditions.

## Simulating a Performed Experiment

To help with debugging and diagnosis it is also possible to simulate the control strategy defined in a single experiment.

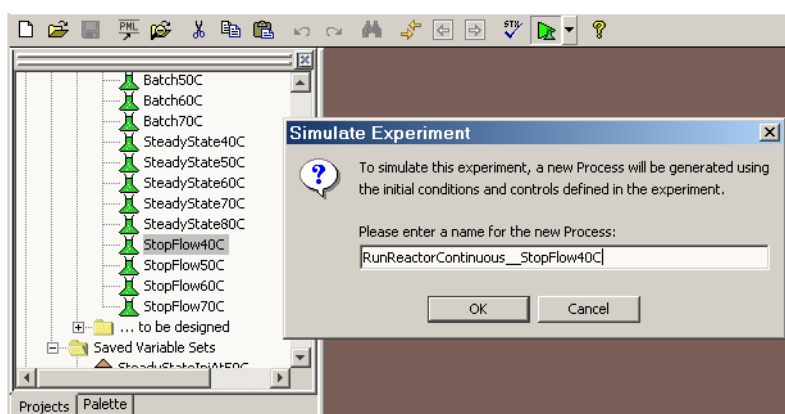
To do this:

1. Ensure a Process has been specified on the General tab of the Performed Experiment.
2. Select the Experiment entity in the Project Tree.
3. From the Activities menu, select Simulate selected experiment... (or click the short-cut button on the tool bar)

gPROMS will prompt you for the name of a new Process entity. This will be created within the same project as the original Process and, if there are no problems reported, will be automatically run as a Simulation activity.

Once the Simulation has completed, you can either delete the newly created Process or keep it.

**Figure 2.8. Simulating an experiment**



## The Process generated to Simulate a Performed Experiment

To simulate an experiment, gPROMS creates a copy of the existing Process and then automatically modifies the text to match the conditions specified in the Performed Experiment.

This approach has a number of advantages. Firstly, the structure of the Process remains intact, including any comments. Secondly, it provides a convenient way to view the modifications that gPROMS has made. To do so, simply select the two versions of the Process and press **F12** to activate the *text comparison* tool.

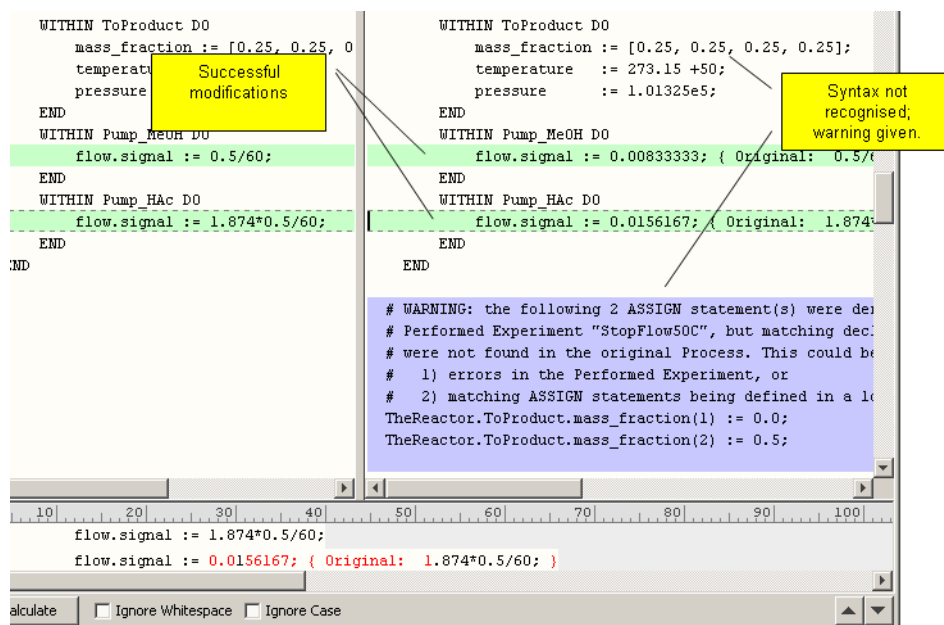
Specifically, the following modifications are made:

- Each Initial conditions is located in the Initial section of the Process, and modified with the value given in the Performed Experiment. Alternatively, if the Performed Experiment's initial condition is steady state, then this is set accordingly.
- Any Controls given in the Performed Experiment are located in the Assign section of the Process, and their values are modified.
- If any of the Performed Experiment's Controls are time-varying, a new Schedule section will be generated that mimics the control strategy.
- Any Saved Variable Set(s) specified in the Performed Experiment are added to the end of the Preset section of the Process (creating this section if necessary).

**Limitations of automatic Process generation:** The text modification approach employed to Simulate a Performed Experiment is not fool-proof. If a problem occurs during the generation of the Process, a message dialog will inform you of this fact. You will need to open the newly created Process and manually make some fixes, which are typically quite straightforward. Search for the text "Warning", and you will find the problems listed.

One problem, for example, occurs when elements of an array are specified using the square bracket notation `[value, value, ...]`, and that array variable is referenced in the Performed Experiment. In this scenario gPROMS is unable to locate the correct location in the Process for automatic modification. See the figure below, which shows how this problem might look when viewed using the *text comparison* viewer.

**Figure 2.9. Text comparison of an original Process and an automatically generated Process.**



After the required corrections have been made, select the newly created Process and Simulate it as you would any other Process (e.g. by pressing **F5**).

## The Parameter Estimation entity

The complete specification of a Parameter Estimation problem requires some additional information which is not provided in the gPROMS Process or Experiments Performed entities. This includes information on the

unknown parameters to be estimated, the number of experiments that were performed, and sensor specifications (for example, if they have not already been defined in the Experiments Performed entity).

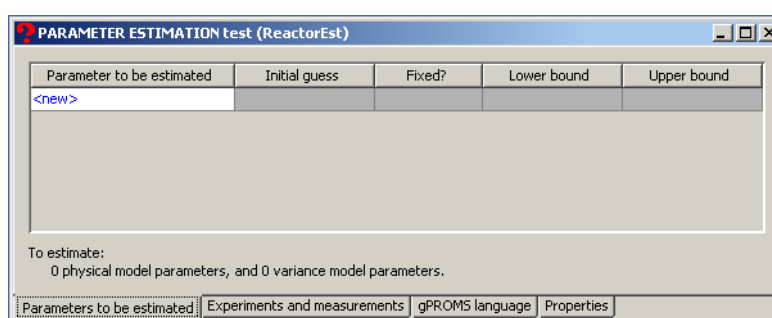
All of the above information has to be specified in a separate entity which appears under the Parameter Estimations entry in the gPROMS project tree.

To create a Parameter Estimation entity:

- From the Entity menu, select New Entity.
- Choose Parameter Estimation for the Entity type and fill in the Name field.

A Parameter Estimation entity editor window will open, as shown in the figure below.

**Figure 2.10. Parameter Estimations Entity editor**



The Parameter Estimations entity editor has the following tabs:

- The Parameters to be estimated tab ( a form for defining the parameters that are to be estimated)
- The Experiments and measurements tab (a form for defining the set of experiments and measurement variance descriptions that will be used in the estimation)
- The gPROMS language tab - this provides a convenient means of reviewing all the information regarding the performed experiment in a text-based format. You are not advised to use this tab to enter data<sup>3</sup>.
- Properties (a record of when the entity was created, last edited as well as the default tab to open on)

## Parameters to be estimated

The Parameters to be estimated are a subset of the ASSIGNED variables in the Process entity, that will change the model predictions if they are adjusted. One could manually try and change them and run repeated simulations until the model predictions match the experiment data, or you can use the Parameter Estimation capabilities in gPROMS to do this in a rigorous and automated fashion.

In order to define these parameters<sup>4</sup> you need to click on the Parameters to be estimated tab in the Parameter Estimations entity editor (as shown in the figure below).

<sup>3</sup>The Parameter Estimation editor is *destructively round-trip*: editing can be performed in the gPROMS language tab, and valid changes will be accepted when you perform a syntax check (**F4**) or change to a form-based view. However, any subsequent edits using the form-based editor will regenerate the language, and all comments and/or formatting will be lost.

<sup>4</sup>Note: even though we call it a *parameter* to be estimated, it should be defined in the gPROMS model as a **variable** and ASSIGNED in the Process entity.

**Figure 2.11. Parameters to be estimated tab**

| Parameter to be estimated | Initial guess | Fixed?                   | Lower bound | Upper bound |
|---------------------------|---------------|--------------------------|-------------|-------------|
| R101.ACTIVATIONENERGY     | 15000.0       | <input type="checkbox"/> | 10000.0     | 20000.0     |
| R101.ARRHENIUSCONSTANT    | 5.0E-4        | <input type="checkbox"/> | 1.0E-4      | 0.0010      |
| R101.REACTIONENTHALPY     | 55000.0       | <input type="checkbox"/> | 35000.0     | 75000.0     |
| <new>                     |               |                          |             |             |

To estimate:  
3 physical model parameters, and 5 variance model parameters.

Parameters to be estimated | Experiments and measurements | gPROMS language | Properties

Then simply click on the <new> cell in the table and type the full pathname to the variable.

## Tip

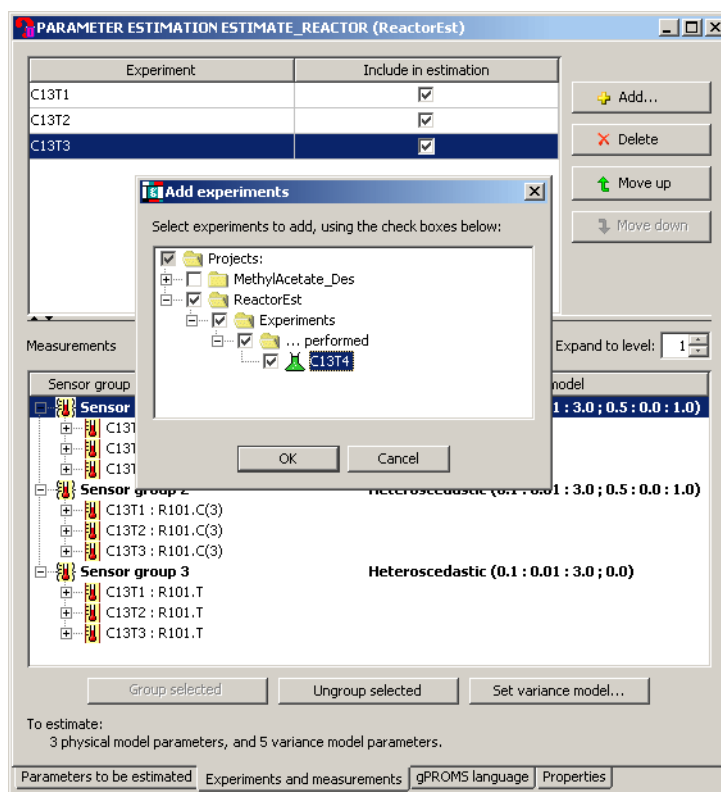
*Pathname completion* can be activated by pressing **Ctrl+Space**.

Having entered the *parameter* you want to estimate, enter the initial guess and the lower and upper bounds in the column alongside. These will be used by the optimizer when solving the Parameter Estimation problem. If you wish to fix the *parameter* at the value defined by its initial guess, then check the cell in the Fixed? column of the table - this means the optimiser will not adjust this value when trying to minimise the deviation of the model from the experimental data.

## Experiments and measurements

To define the experiments that will be used in the Parameter Estimation problem, click on the Experiments and measurements tab and then click the Add button at the top right of the window. gPROMS will then open a new window that includes a project tree with all the experiment entities that have been defined in this project or a cross-referenced project (as shown below).

**Figure 2.12. Specifying which experiments to include in a Parameter Estimation problem**



In the Add experiments dialog, check the boxes of the experiments that you wish to include and select OK.

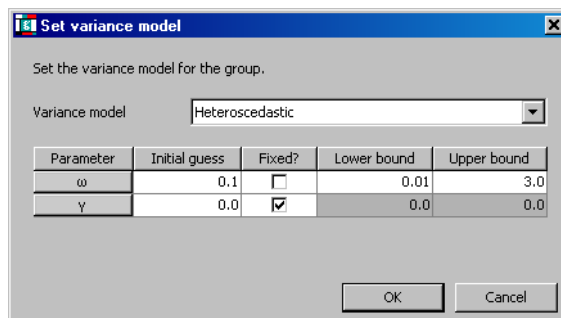
After all the experiments have been added you may want to consider the order of the experiments in the table which can be adjusted using the Move up and Move down buttons. The order of the experiments in the table corresponds to the order in which gPROMS solves them.

gPROMS will automatically create the table of all the sensors and the sensor groups in the lower half of the window. This is both a summary view and an editor. Here you may specify the statistical variance models for all the sensors and sensor groups. Any information provided here will **override** the sensor information provided in the individual experiments and if you have not provided a definition of a sensor for any of the measured variables in an experiment then it **must** be provided here. To change or add a Variance model for a sensor or sensor group, select the item in the tree and click the Set Variance model... button; this will open a window (as shown below) which allows you to set the Variance model. Enter the initial guess and the lower and upper bounds for the relevant statistical parameters that will be used in the solution of the maximum likelihood Parameter Estimation problem [6].

## Note

Setting a variance model will always create a Sensor group, if one does not exist.

**Figure 2.13. Setting a Variance model for a Sensor group**



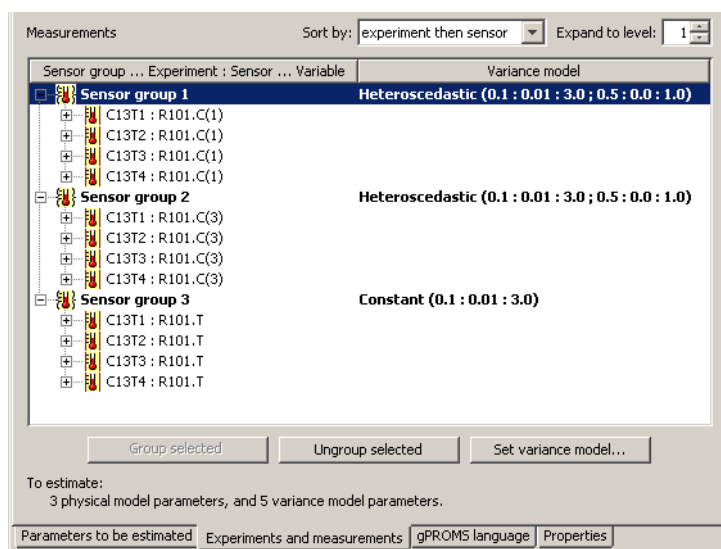


The information at the bottom of the window shows the total number of parameters to be estimated. If you have provided explicit values for all the sensors (or sensor groups) then 0 model variance model parameters will need to be estimated. On the other hand, if you have asked gPROMS to estimate the uncertainty of the measured variables then the number of the variance parameters will be displayed here.

## Grouping sensors

When gPROMS solves a Parameter Estimation problem, the total number of parameters to be estimated is the sum of the unknown physical parameters together with the unknown variance model parameters. It is possible to reduce this number by grouping variance parameters together. For example, two different experiments may be performed where the concentration of a substance is measured using the same device. If it is expected that the measurement variance is the same for both experiments, then one would group them together and have a common variance model for both.

**Figure 2.14. Sensor groups and Variance models**



If you wish to **group measurements** and have a common variance model across all the experiments for those measurements, select the desired sensors and sensor groups<sup>5</sup>, then click the Group selected... button at the bottom of the window. There are now three possibilities:

- if you have not provided a Variance model for any of the sensors to be grouped then you will be asked to select a Variance model and give the appropriate variance parameter values for the new sensor group
- if one or more of the sensors to be grouped has a Variance model then this will be the variance model for the new sensor group assuming that there are no conflicts
- if two or more sensors to be grouped have conflicting Variance models then you will be asked to resolve the conflict and select a single variance model for the new sensor group

If you wish to **ungroup sensors**, select the appropriate sensor group, then click the Ungroup selected... button.

### Note

When you ungroup a sensor, it will revert to using the variance model specification provided in the experiment entity, or to <unspecified> if it was not provided there.

## Execution of a Parameter Estimation activity

For a fully defined Parameter Estimation problem the following entities are required:

<sup>5</sup>There are a number of ways to select multiple items: (1) click and drag, (2) click the first item in a range, then **Shift**+Click the last item, and (3) you can add or remove individual items from the selection using **Ctrl**+Click.

- one or more *Model* entities;
- a *Process* entity;
- an *Estimation* entity
- one or more *Performed Experiments* entities

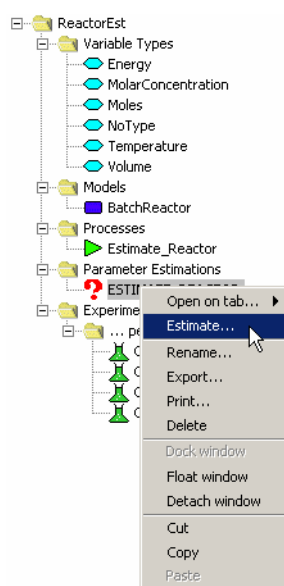
Note that the Experiment entities must all refer to the same Process.

## Running the Parameter Estimation activity

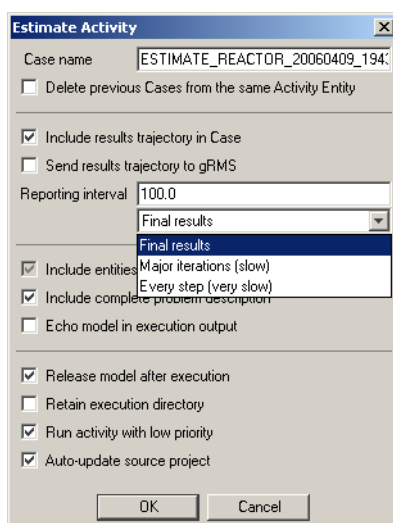
You can execute a Parameter Estimation activity in a number of ways. First, select the Estimation entity from the project tree, and then either;

1. drop down the Activities menu and select Estimate... , or
2. click the red question mark in the tool bar, or
3. press **F5**.

**Figure 2.15. Executing a Parameter Estimation activity from the Project Tree**



If there are any syntactical or cross-referencing mistakes etc., these will be reported. Otherwise, the execution options dialog (below) will be displayed. Upon clicking OK, an execution Case is created and the activity begins.

**Figure 2.16. The Parameter Estimation activity execution control dialog**

**Parameter Estimation reporting intervals:** Most of the execution options for Parameter Estimation are identical to those for Simulation. The only extra option is in the reporting interval. In addition to the reporting interval period, you may also select at which stages of the optimisation the values are recorded into the results trajectory. The three options are:

- Final results - no values are recorded until an optimal solution has been found.
- Major iterations - values are recorded at every major iteration; that is, after a set of parameter values is accepted and a new search direction is being computed.
- Every step - as above, and in addition, values are recorded during the line-search steps of the optimisation. This may include experiment evaluations that do not complete - and so may be useful for diagnosing problems, such as problematic combinations of parameters.

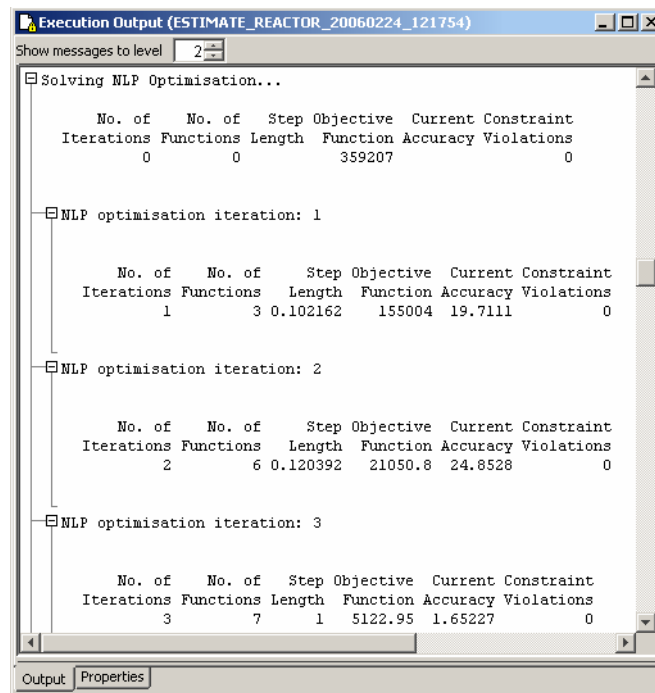
## Note

In addition to the reporting interval, the values of monitored variables are also recorded at each measurement time of the Performed Experiment. Therefore, if you wish to view the values of monitored variables at *just* the measurement times, simply enter a large reporting interval.

See also the Variable Trajectories section.

## Monitoring the activity progress

You can monitor the progress that is being made during any Parameter Estimation activity from the Execution Output window. From the Execution Output window you can expand or collapse portions of the output by clicking on the + or - signs in the tree associated with the executing activity.

**Figure 2.17. Monitoring the progress of a Parameter Estimation activity.**

Adjusting the output level in the SOLUTIONPARAMETERS section of the Process will enable you to change the amount of information sent to the execution window during the execution of an Estimation activity.

## Tuning the Parameter Estimation execution

It is possible to fine tune the numerical parameters associated with the underlying solvers used by gPROMS in the solution of a Parameter Estimation problem. This is done in the SOLUTIONPARAMETERS section of the Process entity. There is currently only one standard mathematical solver for Parameter Estimation in gPROMS, namely MAXLKHD.

The PESolver solution parameter may be used to configure the Parameter Estimation solver. The algorithmic parameters used by MAXLKHD along with their default values are shown below. Since the MINLP optimisation solver is integral to the Parameter Estimation activity, the default MINLP solver (SRQPD) is also shown, with selected relevant settings and their default values. Please see the Optimisation Guide for full details of the SRQPD solver.

### SOLUTIONPARAMETERS

```

PESolver := "MAXLKHD" [
  "OutputLevel" := 0;
  "Statistics" := 0;
  "DASolver" := "DASOLV";
  "MINLPSolver" := "SRQPD" [
    "OutputLevel" := 0;
    "MaximumLineSearchSteps" := 20,
    "MinimumLineSearchStepLength" := 1e-005,
    "OptimisationTolerance" := 0.001,
    "Scaling" := 0
  ]
]

```

### Solution Parameters for the MAXLKHD solver:

- OutputLevel: An integer in the range [-1, 2].

The amount of information generated by the solver. The following table indicates the lowest level at which different types of information are produced:

|    |  |
|----|--|
| -1 | (None)   |
| 0  | Failed integrations and initialisations, optimisation failure, summary information from the SQP nonlinear programming code, final solution point and constraint values, best available point after failure |
| 1  | Name of each experiment being evaluated  |
| 2  | Parsing of experiment-specific Saved Variable Sets.  |

- **Statistics:** An integer in the range [0, 2].

Controls the information produced by the solver in the machine-readable estimation statistics **.stat-mr** file at the end of the computation:

- **0 :** The parameter vector used for the calculation of the variance/covariance matrices includes both the model parameters and the variance model parameters.
- **1 :** The parameter vector used for the calculation of the variance/covariance matrices includes only the model parameters.
- **2 :** The variance/covariance matrices are not calculated.
- **DASolver:** A quoted string specifying a differential-algebraic equation solver.

The solver to be used for integrations of the model equations and their sensitivity equations at each iteration of the parameter estimation. This can be either the standard DASOLV solver or a third-party differential-algebraic equation solver (see the gPROMS System Programmer Guide). The default is DASOLV.

This parameter can be followed by further specifications aimed at configuring the particular solver by setting values to its own algorithmic parameters.

- **MINLPSolver:** A quoted string specifying a nonlinear programming optimisation solver.

#### **Solution Parameters for the SRQPD sub-solver:**

*The following are only a subset of the parameters available in the default MINLP, SRQPD.*

- **OutputLevel:** An integer in the range [-1, 4].

The amount of information generated by the solver. The following table indicates the lowest level at which different types of information are produced:

|    |  |
|----|--|
| -1 | (None)   |
| 0  | Failed integrations and initialisations, optimisation failure, summary information from the SRQPD nonlinear programming code, final solution point and constraint values, best available point after failure |
| 1  | Values of optimisation decision variables, objective function and constraints in each major optimisation iteration   |
| 2  | Start and end times of each interval of integration, optimisation decision variables and objective function at each line search trial  |
| 3  | Derivatives of objective function and constraints  |

- **MaximumLineSearchSteps**: An integer in the range [1, 100].

The maximum number of line search steps in one optimisation iteration.

- **MinimumLineSearchStepLength**: A real number in the range  $[10^{-10}, 1.0]$ .

The minimum length of a line search step.

- **OptimisationTolerance**: A real number in the range [0.0, 1.0].

The solution tolerance for the parameter estimation. Convergence is deemed to occur when the following convergence criterion is satisfied:

$$\frac{1}{|\Phi^*|+1.0} \left( \left| \sum_j \frac{\partial \Phi^*}{\partial \theta_j} \delta \theta_j \right| + \sum_j |\mu_j| \max(0, \theta_j^L - \theta_j^*, \theta_j^* - \theta_j^U) \right)$$

$$\sum_j \max(0, \theta_j^L - \theta_j^*, \theta_j^* - \theta_j^U) \leq \text{OptTol}$$

where:

- $\theta_j$  is the  $j^{\text{th}}$  parameter to be estimated (including both model parameters and variance model parameters);
- $\theta_j^*$  is the final value of parameter  $\theta_j$ ;
- $\theta_j^L$  is the lower bound imposed on parameter  $\theta_j$ ;
- $\theta_j^U$  is the upper bound imposed on parameter  $\theta_j$ ;
- $\Phi^*$  is the final value of the maximum likelihood objective function;
- $\delta \theta_j$  is the step taken in parameter  $\theta_j$  at the last iteration of the parameter estimation calculation;
- $\mu_j$  is the Lagrange multiplier that corresponds to the bound constraints imposed on parameter  $\theta_j$ ;
- **Scaling**: An integer in the range [0, 3].

The parameters to be determined in the context of a single parameter estimation problem may vary significantly in magnitude, which may adversely affect the performance of the optimisation algorithms. Consequently, appropriate scaling of these parameters is strongly recommended.<sup>6</sup>

The scaling performed is of the general mathematical form:

$$\tilde{\theta}_j = \frac{\theta_j - c_j}{d_j}$$

where  $\theta_j$  is the  $j^{\text{th}}$  original parameter to be estimated and  $\tilde{\theta}_j$  the corresponding scaled parameter. The constants  $c_j$  and  $d_j$  are determined automatically depending on the value of **Scaling**, as described below:

- **Scaling** = 0 : No scaling (default).

$$d_j = 1, \quad (1a)$$

$$c_j = \frac{e_j}{d_j} = 0. \quad (1b)$$

---

<sup>6</sup>A useful indication as to whether scaling is necessary is the condition number estimate that is printed out at each iteration of the optimisation calculation. It is recommended that scaling be undertaken for problems with condition numbers exceeding  $10^{10}$ .

- **Scaling = 1** : Scaling according to the ranges of the parameters so that the scaled parameters vary between -1 and 1.

$$d_j = \frac{1}{2} \left( \theta_j^{\max} - \theta_j^{\min} \right), \quad (2a)$$

$$c_j = \frac{1}{2} \left( \theta_j^{\max} + \theta_j^{\min} \right) \quad (2b)$$

- **Scaling = 2** : Scaling according to the initial guesses of the parameters.

$$d_j = \begin{cases} \theta_j^0 & \text{if } |\theta_j^0| > \varepsilon, \\ \frac{1}{2} \left( \theta_j^{\max} - \theta_j^{\min} \right) & \text{otherwise} \end{cases} \quad (3a)$$

$$c_j = 0 \quad (3b)$$

where  $\theta_j^0$  is the initial guess for the  $j^{\text{th}}$  parameter and  $\varepsilon$  is a small constant (currently set at  $10^{-8}$ ).

- **Scaling = 3** : Scaling according to the value and the gradients of the objective function at the initial guess.

$$d_j = \begin{cases} \frac{1 + |\Phi(\theta^0)|}{2 \left| \frac{\partial \Phi}{\partial \theta_j} \right|_{\theta^0}} & \text{if } \left| \frac{\partial \Phi}{\partial \theta_j} \right|_{\theta^0} > \varepsilon, \\ \frac{1}{2} \left( \theta_j^{\max} - \theta_j^{\min} \right) & \text{otherwise} \end{cases} \quad (4a)$$

$$c_j = 0 \quad (4b)$$

where  $\theta^0$  is the vector of initial guesses of the parameters and  $\varepsilon$  is a small constant (currently set at  $10^{-8}$ ).

## Analysing the Parameter Estimation results

You can analyse the results of a Parameter Estimation problem in a number of ways:

1. Variable values, stored in the Trajectories folder of the Case, allow you to view the model predictions of the measured variables - and optionally other variables - while the activity is running.
2. Information on the values of the estimation problems; the status of the optimisation problem and the statistical confidence of the estimated results is provided in three key results files accessed from the Results folder of the Case:
  - Comprehensive Parameter Estimation report file (in HTML format): `PPP`
  - Parameter Estimation report file (in plain text format): `PPP.out`
  - Statistical results for analysis in MS Excel : `PPP.stat-mr`

(where `PPP` is the name of the Parameter Estimation entity that has been executed to produce these results.)

## Variable Trajectories

In the Execution Case results Trajectories are stored for all of the control and measured variables in each of the experiments. Note that from the execution dialog you have the option to choose whether to store only the results from the final solution or whether to record them for all of the major steps or even all of the steps during the solution of the Parameter Estimation problem.

To view the results for a particular variable:-

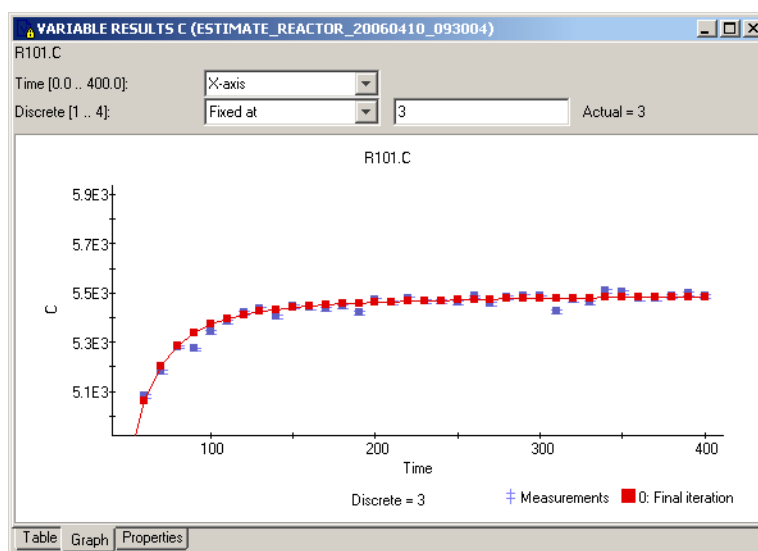
- open the Trajectories folder in the Case
- select the desired Experiment and navigate to the desired Unit by expanding the results tree.
- double click on the desired variable. This will open up a results window for that variable. There are three tabs, corresponding to tabulated results, graphical results and a properties tab that summarises when the execution was carried out.

The figure below shows an example of a 2D plot of a measured variable. The model predictions for all variables are shown in red. In the 2D plots of a measured variable, the measured values (taken from data in the performed experiment) are also shown in blue, with error bars indicating the standard deviation of the measurements.

## Note

The standard deviations here are indicative only, based on the *initial guess* of the variance model parameters. If the variance model parameters themselves are being estimated (and therefore may change during the parameter estimation activity), then see the comprehensive report for the calculated standard deviations.

**Figure 2.18. Predicted trajectory of a measured variable in an experiment.**



**Selecting variables to monitor.** If no MONITOR section has been specified in the Process, then a default subset of the variables are recorded in the Trajectory. This subset consists of: the measured variables, the control variables, and the parameters being estimated.

To customise which variables are recorded, you can create a MONITOR section in the Process, as shown in the example below.

```
UNIT
  TheReactor AS Reactor

MONITOR
  TheReactor.Reactor.*;
  TheReactor.Pump_*.mass_flowrate;

...
```

The MONITOR section will replace the default subset of variables being recorded.

## Caution

Be aware that the amount of data being logged will increase if you monitor extra variables. The storage requirements for the variables will be multiplied by both the number of experiments and the number of



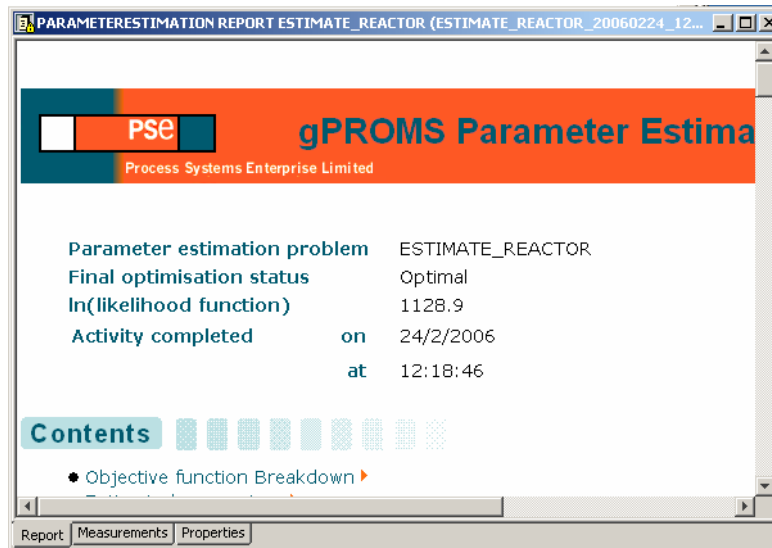
iterations reported, which it could result in a large amount of data being stored in the Case project, or sent to gRMS.

## The Comprehensive report for Parameter Estimation

Found in the Results folder, the Parameter Estimation report file has the same name as the estimation activity that was executed. To view the report simply open it by double-clicking.

The figure below shows an example of a comprehensive report.

**Figure 2.19. Parameter Estimation comprehensive report.**



The report has three tabs:

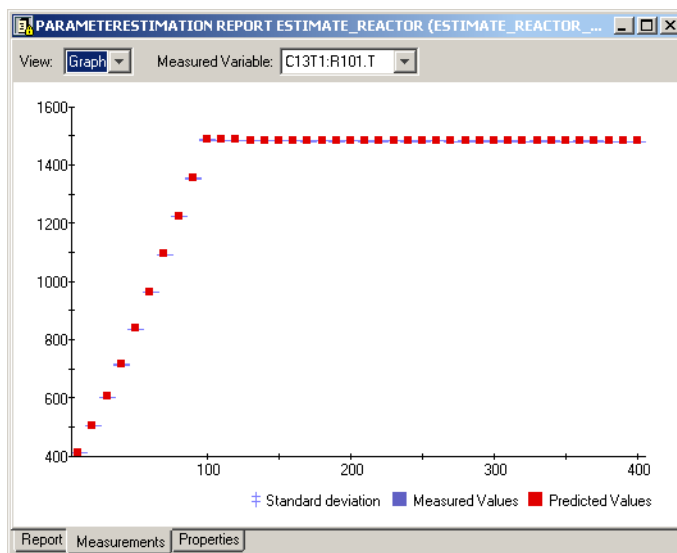
- the Report tab
- the Measurements tab
- the Properties tab - this provides details regarding the time of creation.

The Report tab includes the following information:

- a table of contents that allows quick access to the information listed below via "hyperlinks";
- general information such as the date and time of the execution of the activity, and its final status;
- a detailed breakdown of the value of the likelihood function in terms of the constant, variance and residual terms of each measured variable in each experiment;
- the initial guesses and final (optimal) values of all model parameters being estimated, together with the corresponding lower and upper bounds, the 90%, 95% and 99% confidence intervals, the 95% t-value and the standard deviation;
- the same information as above for any parameters describing the variance behaviour of the various sensors used for the measurements; these parameters will also have been estimated;
- a detailed analysis of each experiment, including the experimental controls and initial conditions, and the measured and estimated values of each measured point, as well as the corresponding difference ("residual");
- the variance/covariance matrix [31];
- the correlation matrix [31];
- a goodness-of-fit analysis based on the 95%  $\chi^2$  criterion.

From the Measurements tab you can compare the predicted and experimental values for the measured variables as well as seeing the statistical confidence in the results. You can select the variable to plot and the confidence interval from a drop down menus. You can also choose to view the data for the measured variable in tabular format..

**Figure 2.20. Parameter Estimation measurement plots**



## The summary report file

The text report (.out) file contains a summary report on the Parameter Estimation activity (giving the same information as the Comprehensive report), including:

- the outcome of the estimation run
- the final value of the objective function
- the final values of the estimated model parameters
- the final values of the sensor's variance model parameters

The file also contains computational statistics for the estimation calculation.

## Example Parameter Estimation text output file

gPROMS Parameter Estimation

```

Estimation Name           : ESTIMATE_REACTOR
Final Optimisation Status : Optimal Solution Found
Objective function        : 1128.91
    Constant term         : 441.09
    Variance term         : 462.61
    Least squares term    : 225.207
  
```

Statistical Analysis:  
=====

Estimated Parameters:

# Parameter Estimation in gPROMS

| Parameter              | Final Value | Lower Bound | Upper Bound | 95% Conf. Interval | 95% t-value | S |
|------------------------|-------------|-------------|-------------|--------------------|-------------|---|
| R101.ACTIVATIONENERGY  | 1.503E+004  | 1E+004      | 2E+004      | 55                 | 2.7E+002    | 2 |
| R101.ARRHENIUSCONSTANT | 0.0007556   | 0.0001      | 0.001       | 1.1E-005           | 70          | 5 |
| R101.REACTIONENTHALPY  | 5.502E+004  | 3.5E+004    | 7.5E+004    | 32                 | 1.7E+003    | 1 |

## Variance Models:

|                                 |        |      |      |       |     |   |
|---------------------------------|--------|------|------|-------|-----|---|
| Sensor Group 1: HETEROSCEDASTIC |        |      |      |       |     |   |
| omega                           | 0.1536 | 0.01 | 3    | 0.31  | 0.5 | 0 |
| gamma                           | 0.6371 | 0    | 1    | 0.51  | 1.3 | 0 |
| Sensor Group 2: HETEROSCEDASTIC |        |      |      |       |     |   |
| omega                           | 0.1305 | 0.01 | 3    | 0.031 | 4.2 | 0 |
| gamma                           | 0.585  | 0    | 1    | 0.034 | 17  | 0 |
| Sensor Group 3: HETEROSCEDASTIC |        |      |      |       |     |   |
| omega                           | 0.4858 | 0.01 | 3    | 0.072 | 6.7 | 0 |
| gamma                           | 0      | 0[*] | 0[*] |       |     |   |

## Variance-covariance matrix of estimated parameters:

```

792
0.000144 3E-011
-107 -2.73E-005 266
2.3 5.87E-007 -0.717 0.0244
-3.77 -9.69E-007 1.19 -0.0403 0.0671
-0.0101 -1.29E-009 0.000115 -7.02E-006 1.7E-005 0.000252
-0.00546 -9.55E-010 0.000612 -3.94E-006 1.02E-006 -0.000226 0.000294
-0.044 -4.24E-009 0.000694 3.09E-005 -5.54E-005 -4.59E-006 5.16E-006 0.00135

```

## Correlation matrix of estimated parameters:

```

1
0.937 1
-0.234 -0.306 1
0.523 0.687 -0.282 1
-0.517 -0.683 0.283 -0.996 1
-0.0225 -0.0148 0.000445 -0.00284 0.00413 1
-0.0113 -0.0102 0.00219 -0.00147 0.000229 -0.833 1
-0.0426 -0.0211 0.00116 0.00539 -0.00583 -0.00789 0.0082 1

```

## Experimental Settings:

Experiment 1: C13T1

## Initial States:

```

R101.HOLDUP(1) 100
R101.HOLDUP(2) 100
R101.HOLDUP(3) 1
R101.HOLDUP(4) 1
R101.T 293

```

## Intervals:

```
-----
Interval # 1    100
Interval # 2    300
```

Time-Varying Controls:

```
-----
1.  R101.Q
    (piecewise constant)
```

```
Interval    Value
1           505000
2              0
```

Measurements:

```
-----
Measured Variable  R101.C(1)      Sensor Group Sensor Group 1
```

| Time | Meas.Value | Pred.Value | Deviation | %Deviation | StdDev  | Dev/StdDev |
|------|------------|------------|-----------|------------|---------|------------|
| 10   | 3746.911   | 3720.9496  | 25.9614   | 0.692875   | 29.0673 | 0.893149   |
| 20   | 2518.6     | 2513.4938  | 5.10624   | 0.202741   | 22.5678 | 0.226263   |
| 30   | 1504.12    | 1509.8021  | -5.6821   | -0.377769  | 16.2498 | -0.349672  |
| 40   | 883.7833   | 888.95174  | -5.16844  | -0.584809  | 11.58   | -0.446325  |
| 50   | 548.0505   | 543.84439  | 4.20611   | 0.767467   | 8.54054 | 0.492487   |
| 60   | 348.6161   | 352.88585  | -4.26975  | -1.22477   | 6.40183 | -0.666958  |
| 70   | 239.5115   | 242.7614   | -3.2499   | -1.35689   | 5.04007 | -0.644812  |
| 80   | 175.4503   | 175.71184  | -0.261537 | -0.149066  | 4.13345 | -0.0632732 |
| 90   | 135.0508   | 132.66329  | 2.38751   | 1.76786    | 3.49862 | 0.682415   |
| 100  | 100.9699   | 103.68624  | -2.71634  | -2.69024   | 2.90685 | -0.934459  |
| 110  | 87.2978    | 84.205283  | 3.09252   | 3.54249    | 2.64949 | 1.16721    |
| 120  | 74.13863   | 70.888503  | 3.25013   | 4.38385    | 2.38755 | 1.36128    |
| 130  | 62.0976    | 61.209633  | 0.887967  | 1.42995    | 2.13261 | 0.416376   |
| 140  | 54.54854   | 53.85582   | 0.69272   | 1.26991    | 1.96357 | 0.352786   |
| 150  | 49.05154   | 48.079397  | 0.972143  | 1.98188    | 1.83508 | 0.529756   |

<... snip ...>

|     |          |           |           |          |          |           |
|-----|----------|-----------|-----------|----------|----------|-----------|
| 350 | 14.34733 | 15.287814 | -0.940484 | -6.55511 | 0.838487 | -1.12164  |
| 360 | 15.20852 | 14.783675 | 0.424845  | 2.79347  | 0.870214 | 0.488207  |
| 370 | 14.26656 | 14.311725 | -0.045165 | -0.31658 | 0.835477 | -0.054059 |
| 380 | 14.57341 | 13.868976 | 0.704434  | 4.83369  | 0.846882 | 0.831797  |
| 390 | 11.8482  | 13.452799 | -1.6046   | -13.543  | 0.742229 | -2.16186  |
| 400 | 14.05076 | 13.060872 | 0.989888  | 7.04509  | 0.827403 | 1.19638   |

<...and for all experiment measurement....>

| Experiment | Variable  | Contribution |
|------------|-----------|--------------|
| C13T1      | R101.C(1) | 38.160398    |
| C13T1      | R101.C(3) | 139.93596    |
| C13T1      | R101.T    | -14.315856   |
| C13T2      | R101.C(1) | 38.383615    |
| C13T2      | R101.C(3) | 128.82448    |
| C13T2      | R101.T    | -0.31047035  |

|                       |           |            |
|-----------------------|-----------|------------|
| C13T3                 | R101.C(1) | 36.729493  |
| C13T3                 | R101.C(3) | 129.8304   |
| C13T3                 | R101.T    | -4.881258  |
| C13T4                 | R101.C(1) | 58.162784  |
| C13T4                 | R101.C(3) | 149.39776  |
| C13T4                 | R101.T    | -12.100016 |
| N/2*ln(2*PI)          |           | 441.0905   |
| Sum of variance terms |           | 462.61015  |
| Sum of residual terms |           | 225.20714  |

Computational Statistics

|  |   |          |                                 |
|--|---|----------|---------------------------------|
| Total CPU Time                                 | : | 22.063   | seconds                         |
| ESTIMATE_REACTOR Optimiser Statistics          |   |          |                                 |
| CPU Time                                       | : | 0.016    | seconds (0.072520 % of total)   |
| Number of NLP Iterations                       | : | 34       |                                 |
| Number of NLP Line Search Steps                | : | 38       |                                 |
| DASOLV Integrator Statistics                   |   |          |                                 |
| CPU Time                                       | : | 8.75902  | seconds (39.700020 % of total)  |
| CPU Time Spent on Integration Only             | : | 3.01301  | seconds (34.398934 % of DASOLV) |
| 40104 steps, 60972 residuals                   | : | 0.781008 | seconds                         |
| 11848 Jacobians                                | : | 0.200001 | seconds                         |
| CPU Time Spent on Sensitivity Integration Only | : | 5.74601  | seconds (65.601066 % of DASOLV) |
| 38060 steps, 95136 residuals                   | : | 0.562005 | seconds                         |
| 49220 Jacobians                                | : | 0.372004 | seconds                         |
| Mean (Sensitivity+State)/(State) CPU Ratio     | : | 1.52437  |                                 |

## Statistical analysis in Excel

gPROMS provides an advanced statistical analysis tool in Microsoft Excel™ that allows the user to further analyse the results and create several statistical plots. The data that are necessary for the operation of this facility are stored in the *.stat-mr file*. It contains all the information that appears in the text based results file but in machine-readable format for import into Excel.

To use the Excel based tool on a Windows platform - double click the *.stat-mr* file; this will open Excel and will automatically import the statistics file<sup>7</sup>. If you running gPROMS on a non-Windows platform then you can export the *.stat-mr* data file for transfer to a MS Windows-based system.

The Excel tool makes it easy to

1. View and manipulate measurement and statistical data tables
2. Create overlay, residual and confidence ellipsoid plots

## Data spreadsheets

A number of spreadsheets are created automatically when the statistics data file is loaded in Excel.

- **Summary.** This sheet summarises the contributions to the objective function of the constant, variance and residual terms, as well as the types of variance models used for each measured variable in each experiment and their respective contributions to the objective function. It also shows the estimated parameters. Parameters in red colour indicate either that the estimated value is at one of its bounds or that the confidence intervals are very large (see also the "Statistical significance" sheet). A sample summary sheet for the batch reactor example can be found in the next chapter.

<sup>7</sup>The security settings within Excel must allow the execution of macros.

- **Parameter estimates.** This sheet summarises the estimated values of the model parameters and the parameters of the different statistical variance models used, together with their initial guesses, and lower and upper bounds (see next chapter).
- **Variance matrix.** This sheet prints the variance-covariance matrix,  $V$ . This matrix contains the variances and covariances of the estimated process model and variance model parameters. The square root of each diagonal element,  $\sqrt{v_{ii}}$ , is the approximated standard deviation of the respective estimated parameter.

The following approximation to the variance-covariance matrix is used:

$$\mathbf{V} = \mathbf{H}^{*-1} \Gamma \mathbf{V}_z \Gamma^T \mathbf{H}^{*-1}$$

where

$$\mathbf{H}^* = \begin{bmatrix} \frac{\partial^2}{\partial \theta^2} \Phi & \frac{\partial^2}{\partial \beta_{i,j} \partial \theta} \Phi \\ \frac{\partial^2}{\partial \beta_{i,j} \partial \theta} \Phi & \frac{\partial^2}{\partial \beta_{i,j}^2} \Phi \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \frac{\partial^2}{\partial \tilde{z}_{i,j,k} \partial \theta} \Phi \\ \frac{\partial^2}{\partial \tilde{z}_{i,j,k} \partial \beta_{i,j}} \Phi \end{bmatrix}$$

and  $\mathbf{v}_z = \text{diag}(\sigma_{ijk}^2)$  denotes the variance-covariance matrix of the measurement errors. The variance-covariance matrix of the estimated parameters,  $V$ , is of size  $\mathbf{r}^{n_p \times n_p}$ , where  $N_p$  is the number of all estimated parameters (process model parameters  $\theta$  and variance model parameters  $\beta$ ) whose values do not lie at one of their respective lower or upper bounds.

- **Correlation matrix.** This sheet prints the correlation matrix,  $R$ , which is calculated from the variance-covariance matrix,

$$R_{ij} = \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}}, \quad i \neq j,$$

where  $R_{ij} = 1, \quad i = j.$

Values with absolute value close to one in the off-diagonals indicate a high correlation of the corresponding parameters  $i$  and  $j$ , and vice versa. The high correlation of these parameters can also be seen in the corresponding confidence ellipsoid plots (see below).

- **Information matrix.** This sheet prints the Fischer information matrix,  $M$ , which is equal to the inverse of the variance-covariance matrix  $V$ . It also prints the 90%, 95% and 99% F-values for this matrix (calculated using internal statistical functions).

An approximate  $(1 - \alpha)$  highest posterior density region for the parameters  $\mathbf{p}^t = (\theta^t, \beta^t)$  is given by

$$(\mathbf{p} - \hat{\mathbf{p}})^T \cdot \mathbf{M} \cdot (\mathbf{p} - \hat{\mathbf{p}}) \leq N_p s^2 F(\alpha, N_p, N - N_p)$$

with

$$s^2 = \frac{1}{N - N_p} \sum_{i=1}^{NE} \sum_{j=1}^{NV_i} \sum_{k=1}^{NM_{ij}} \frac{(\tilde{z}_{ijk} - z_{ijk})^2}{\sigma_{ijk}^2}$$

which forms an ellipsoid in the  $N_p$ -dimensional parameter space.<sup>8</sup> An  $\alpha\%$  confidence region means that if we repeat the experiments (which produces nearly the same measurements, but with slightly different observation values and therefore a different distribution of the measurement errors), and estimate the parameters out of the repeated experimental data, the values of the estimated parameters will lie in this confidence region with  $\alpha\%$  probability.<sup>9</sup> Confidence ellipsoids for any chosen pair of parameters can be plotted by selecting the Plot option (see below).

<sup>8</sup>Again only those parameters whose values are not at one of their bounds are considered.

<sup>9</sup>Note that the confidence ellipsoid is only a linear approximation of the non-linear confidence region and may not be very accurate for models which are highly non-linear in the parameters.

- **Statistical significance.** This sheet prints:
  - the estimated values of the process model and variance model parameters;
  - the 90%, 95% and 99% confidence intervals for the estimated model and variance model parameters. The confidence ellipsoid is bounded by the box
 
$$\otimes_{i=1}^{N_p} [\hat{p}_i - X_i(\alpha); \hat{p}_i + X_i(\alpha)]$$

The sides of the box,  $[\hat{p}_i - X_i(\alpha); \hat{p}_i + X_i(\alpha)]$ , are called the two-sided joint  $\alpha\%$  confidence intervals.

These are calculated from  $X_i(\alpha) = t\left(\frac{1+\alpha}{2}, N - N_p\right) \cdot \sqrt{V_{ii}}$
  - 95% t-values for the estimated parameters. These are calculated from:
 
$$t_i = \frac{\hat{p}_i}{X_i(0.95)}$$

The t-values show the percentage accuracy of the estimated parameters with respect to the 95% confidence intervals.

The associated t-values,  $t_i$ , are compared with the reference 95% t-value,  $t(0.95, N - N_p)$ , which is again calculated using internal statistical functions. A t-value larger than the reference t-value indicates that the corresponding parameter has been accurately estimated (the standard deviation and the confidence interval are small compared to the value of the estimated parameter); a smaller value indicates a poor estimate of the corresponding parameter.
  - the standard deviations  $\sqrt{v_{ii}}$  of the estimated parameters.
- **Measured variable information.** For each measured variable in each experiment, a sheet is produced with the measured and predicted values, the standard deviations and the absolute, relative and weighted deviations. One such sheet for the variable  $R101.C(1)$  in the batch reactor experiment C13T1 is shown in the next chapter.

## Sample summary sheet.

|     | A                                 | B                                | C                                 | D   | E | F | G | H | I | J |
|-----|-----------------------------------|----------------------------------|-----------------------------------|---|---|---|---|---|---|---|
| 1   | Parameter estimation problem:     |                                  | <b>ESTIMATE_REACTOR (Optimal)</b> |   |   |   |   |   |   |   |
| 2   |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 3   | Completed at: <b>12:18:46 PM</b>  |                                  | on <b>24/02/2006</b>              |   |   |   |   |   |   |   |
| 4   |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 5   | Objective Function                |                                  |                                   |   |   |   |   |   |   |   |
| 6   |                                   | <b>Value</b>                     | <b>1128.908</b>                   |   |   |   |   |   |   |   |
| 7   |                                   | <b>Constant Term</b>             | 441.090                           |   |   |   |   |   |   |   |
| 8   |                                   | <b>Variance Sum</b>              | 462.610                           |   |   |   |   |   |   |   |
| 9   |                                   | <b>Residual Sum</b>              | 225.207                           |   |   |   |   |   |   |   |
| 10  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 11  | Model parameters to be estimated: |                                  |                                   |   |   |   |   |   |   |   |
| 12  | R101.ACTIVATIONENERGY             |                                  |                                   |   |   |   |   |   |   |   |
| 13  | R101.ARRHENIUSCONSTANT            |                                  |                                   |   |   |   |   |   |   |   |
| 14  | R101.REACTIONENTHALPY             |                                  |                                   |   |   |   |   |   |   |   |
| 15  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 16  |                                   | <b><u>Variables Measured</u></b> | <b><u>Variance Model</u></b>      | <b><u>Objective Function Contribution</u></b> |   |   |   |   |   |   |
| 17  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 18  | Experiment #1:                    | <b>C13T1</b>                     |                                   |   |   |   |   |   |   |   |
| 19  |                                   | R101.C(1)                        | HETEROSCEDASTICITY                | 38.160  |   |   |   |   |   |   |
| 20  |                                   | R101.C(3)                        | HETEROSCEDASTICITY                | 129.936                                       |   |   |   |   |   |   |
| 21  |                                   | R101.T                           | HETEROSCEDASTICITY                | -14.316                                       |   |   |   |   |   |   |
| 22  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 23  | Experiment #2:                    | <b>C13T2</b>                     |                                   |   |   |   |   |   |   |   |
| 24  |                                   | R101.C(1)                        | HETEROSCEDASTICITY                | 38.384  |   |   |   |   |   |   |
| 25  |                                   | R101.C(3)                        | HETEROSCEDASTICITY                | 128.824                                       |   |   |   |   |   |   |
| 26  |                                   | R101.T                           | HETEROSCEDASTICITY                | -0.310  |   |   |   |   |   |   |
| 27  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 28  | Experiment #3:                    | <b>C13T3</b>                     |                                   |   |   |   |   |   |   |   |
| 29  |                                   | R101.C(1)                        | HETEROSCEDASTICITY                | 36.729  |   |   |   |   |   |   |
| 30  |                                   | R101.C(3)                        | HETEROSCEDASTICITY                | 129.830                                       |   |   |   |   |   |   |
| 31  |                                   | R101.T                           | HETEROSCEDASTICITY                | -4.881  |   |   |   |   |   |   |
| 32  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 33  | Experiment #4:                    | <b>C13T4</b>                     |                                   |   |   |   |   |   |   |   |
| 34  |                                   | R101.C(1)                        | HETEROSCEDASTICITY                | 58.163  |   |   |   |   |   |   |
| 35  |                                   | R101.C(3)                        | HETEROSCEDASTICITY                | 149.398                                       |   |   |   |   |   |   |
| 36  |                                   | R101.T                           | HETEROSCEDASTICITY                | -12.100                                       |   |   |   |   |   |   |
| 37  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 38  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 39  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 40  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 41  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 42  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 43  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 44  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 45  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 46  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 47  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 48  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 49  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 50  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 51  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 52  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 53  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 54  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 55  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 56  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 57  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 58  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 59  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 60  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 61  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 62  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 63  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 64  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 65  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 66  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 67  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 68  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 69  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 70  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 71  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 72  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 73  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 74  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 75  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 76  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 77  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 78  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 79  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 80  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 81  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 82  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 83  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 84  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 85  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 86  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 87  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 88  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 89  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 90  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 91  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 92  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 93  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 94  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 95  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 96  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 97  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 98  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 99  |                                   |                                  |                                   |   |   |   |   |   |   |   |
| 100 |                                   |                                  |                                   |   |   |   |   |   |   |   |

## Sample parameter estimates sheet

|    | A                                  | B                     | C                  | D                          | E                | F             |
|----|------------------------------------|-----------------------|--------------------|----------------------------|------------------|---------------|
| 1  | Parameter estimates for problem:   |                       |                    | ESTIMATE_REACTOR (Optimal) |                  |               |
| 2  |                                    |                       |                    |                            |                  |               |
| 3  | Model parameter estimates          |                       |                    |                            |                  |               |
| 4  |                                    |                       |                    |                            |                  |               |
| 5  |                                    | Parameter             |                    | Optimal Estimate           | Initial Guess    |               |
| 6  |                                    | R101ACTIVATIONENERGY  |                    | 1.503103E+04               | 1.500000E+04     |               |
| 7  |                                    | R101ARRHENIUSCONSTANT |                    | 7.556302E-04               | 5.000000E-04     |               |
| 8  |                                    | R101REACTIONENTHALPY  |                    | 5.502167E+04               | 5.500000E+04     |               |
| 9  |                                    |                       |                    |                            |                  |               |
| 10 |                                    |                       |                    |                            |                  |               |
| 11 |                                    |                       |                    |                            |                  |               |
| 12 | Variance model parameter estimates |                       |                    |                            |                  |               |
| 13 |                                    |                       |                    |                            | OMEGA            |               |
| 14 | Experiment                         | Measured Variable     | Variance Model     | Measured / Predicted       | Optimal Estimate | Initial Guess |
| 15 | C13T1                              | R101.C(1)             | HETEROSCEDASTICITY | MEASURED                   | 1.536279E-01     | 1.000000E-01  |
| 16 | C13T2                              | R101.C(1)             | "                  | "                          | "                | "             |
| 17 | C13T3                              | R101.C(1)             | "                  | "                          | "                | "             |
| 18 | C13T4                              | R101.C(1)             | "                  | "                          | "                | "             |
| 19 | C13T1                              | R101.C(3)             | HETEROSCEDASTICITY | MEASURED                   | 1.304752E-01     | 1.000000E-01  |
| 20 | C13T2                              | R101.C(3)             | "                  | "                          | "                | "             |
| 21 | C13T3                              | R101.C(3)             | "                  | "                          | "                | "             |
| 22 | C13T4                              | R101.C(3)             | "                  | "                          | "                | "             |
| 23 | C13T1                              | R101.T                | HETEROSCEDASTICITY | MEASURED                   | 4.858160E-01     | 1.000000E-01  |
| 24 | C13T2                              | R101.T                | "                  | "                          | "                | "             |
| 25 | C13T3                              | R101.T                | "                  | "                          | "                | "             |
| 26 | C13T4                              | R101.T                | "                  | "                          | "                | "             |
| 27 |                                    |                       |                    |                            |                  |               |
| 28 |                                    |                       |                    |                            |                  |               |
| 29 |                                    |                       |                    |                            |                  |               |
| 30 |                                    |                       |                    |                            |                  |               |
| 31 |                                    |                       |                    |                            |                  |               |
| 32 |                                    |                       |                    |                            |                  |               |
| 33 |                                    |                       |                    |                            |                  |               |
| 34 |                                    |                       |                    |                            |                  |               |
| 35 |                                    |                       |                    |                            |                  |               |
| 36 |                                    |                       |                    |                            |                  |               |

Summary

Parameter Estimates

Statistical Significance

Variance Matrix

Correlation Matrix

Summary Parameter Estimates Statistical Significance Variance Matrix Correlation Matrix

## Sample correlation matrix sheet

|    | A   | B                      | C             | D                          | E             |
|----|---|------------------------|---------------|----------------------------|---------------|
| 1  | Correlation matrix for problem  |                        |               | ESTIMATE_REACTOR (Optimal) |               |
| 2  |   |                        |               |                            |               |
| 3  |   |                        |               |                            |               |
| 4  |   | Parameter              | Optimal Value | Parameter Number           | 1             |
| 5  |   | R101.ACTIVATIONENERGY  | 1.503103E+04  | 1                          | 1.000000E+00  |
| 6  |   | R101.ARRHENIUSCONSTANT | 7.556302E-04  | 2                          | 9.372804E-01  |
| 7  |   | R101.REACTIONENTHALPY  | 5.502167E+04  | 3                          | -2.337364E-01 |
| 8  |   | Sensor_Group_1.Omega   | 1.536279E-01  | 4                          | 5.230276E-01  |
| 9  |   | Sensor_Group_1.Gamma   | 6.371410E-01  | 5                          | -5.170103E-01 |
| 10 |   | Sensor_Group_2.Omega   | 1.304752E-01  | 6                          | -2.254801E-02 |
| 11 |   | Sensor_Group_2.Gamma   | 5.849756E-01  | 7                          | -1.132908E-02 |
| 12 |   | Sensor_Group_3.Omega   | 4.858160E-01  | 8                          | -4.257078E-02 |
| 13 |   |                        |               |                            |               |
| 14 |   |                        |               |                            |               |
| 15 | The above matrix omits the following parameters, the optimal estimates of which lie on either |                        |               |                            |               |
| 16 |   |                        |               |                            |               |
| 17 |   | Parameter              |               |                            |               |
| 18 | C13T1   | R101.C(1).OMEGA        |               |                            |               |
| 19 | C13T1   | R101.C(1).GAMMA        |               |                            |               |
| 20 | C13T1   | R101.C(3).OMEGA        |               |                            |               |
| 21 | C13T1   | R101.C(3).GAMMA        |               |                            |               |
| 22 | C13T1   | R101.T.OMEGA           |               |                            |               |
| 23 | C13T1   | R101.T.GAMMA           |               |                            |               |
| 24 |   |                        |               |                            |               |
| 25 |   |                        |               |                            |               |
| 26 |   |                        |               |                            |               |
| 27 |   |                        |               |                            |               |
| 28 |   |                        |               |                            |               |

Statistical Significance Variance Matrix Correlation Matrix Information Matrix C13T1..R10



## Sample information matrix sheet

|    | A                        | B                          | C                  | D                  | E          | F | G | H | I |
|----|--------------------------|----------------------------|--------------------|--------------------|------------|---|---|---|---|
| 1  | Information matrix       | ESTIMATE_REACTOR (Optimal) |                    |                    |            |   |   |   |   |
| 2  |                          |                            |                    |                    |            |   |   |   |   |
| 3  |                          |                            |                    |                    |            |   |   |   |   |
| 4  |                          |                            |                    |                    |            |   |   |   |   |
| 5  |                          |                            |                    |                    |            |   |   |   |   |
| 6  |                          |                            |                    |                    |            |   |   |   |   |
| 7  | F values for this        |                            | 90%                | 95%                | 99%        |   |   |   |   |
| 8  |                          |                            | 1.683255           | 1.958015           | 2.54903    |   |   |   |   |
| 9  |                          |                            |                    |                    |            |   |   |   |   |
| 10 |                          |                            |                    |                    |            |   |   |   |   |
| 11 |                          |                            |                    |                    |            |   |   |   |   |
| 12 |                          |                            |                    |                    |            |   |   |   |   |
| 13 |                          |                            |                    |                    |            |   |   |   |   |
| 14 |                          |                            |                    |                    |            |   |   |   |   |
| 15 |                          |                            |                    |                    |            |   |   |   |   |
| 16 |                          |                            |                    |                    |            |   |   |   |   |
| 17 |                          |                            |                    |                    |            |   |   |   |   |
| 18 |                          |                            |                    |                    |            |   |   |   |   |
| 19 |                          |                            |                    |                    |            |   |   |   |   |
| 20 |                          |                            |                    |                    |            |   |   |   |   |
| 21 |                          |                            |                    |                    |            |   |   |   |   |
| 22 |                          |                            |                    |                    |            |   |   |   |   |
| 23 |                          |                            |                    |                    |            |   |   |   |   |
| 24 |                          |                            |                    |                    |            |   |   |   |   |
| 25 |                          |                            |                    |                    |            |   |   |   |   |
| 26 |                          |                            |                    |                    |            |   |   |   |   |
| 27 |                          |                            |                    |                    |            |   |   |   |   |
| 28 |                          |                            |                    |                    |            |   |   |   |   |
| 29 |                          |                            |                    |                    |            |   |   |   |   |
|    | Statistical Significance | Variance Matrix            | Correlation Matrix | Information Matrix | C13T1..R10 |   |   |   |   |

## Sample statistical significance sheet

|    | A   | B                          | C                  | D                    | E                | F                   | G            |
|----|---|----------------------------|--------------------|----------------------|------------------|---------------------|--------------|
| 1  | Statistical information for problem:  | ESTIMATE_REACTOR (Optimal) |                    |                      |                  |                     |              |
| 2  |   |                            |                    |                      |                  |                     |              |
| 3  | A 95% t-value for a parameter component smaller than the reference t-value indicates that the data is not sufficient to estimate this parameter p |                            |                    |                      |                  |                     |              |
| 4  |   |                            |                    |                      |                  |                     |              |
| 5  | Model parameters  |                            |                    |                      |                  |                     |              |
| 6  |   |                            |                    |                      |                  | Confidence Interval |              |
| 7  |   | Parameter                  |                    | Optimal Estimate     | 90%              | 95%                 |              |
| 8  |   | R101ACTIVATIONENERGY       |                    | 1.503103E+04         | 4.638211E+01     | 5.531497E+01        |              |
| 9  |   | R101ARRHENIUSCONSTANT      |                    | 7.556302E-04         | 3.019370E-06     | 1.075643E-05        |              |
| 10 |   | R101REACTIONENTHALPY       |                    | 5.502167E+04         | 2.686523E+01     | 3.203927E+01        |              |
| 11 |   |                            |                    |                      |                  |                     |              |
| 12 |   |                            |                    |                      |                  |                     | Referen      |
| 13 |   |                            |                    |                      |                  |                     |              |
| 14 |   |                            |                    |                      |                  |                     |              |
| 15 | Variance model parameters   |                            |                    |                      |                  |                     |              |
| 16 |   |                            |                    |                      |                  | OMEGA               |              |
| 17 |   |                            |                    |                      |                  | Confidence Interval |              |
| 18 | Experiment  | Measured Variable          | Variance Model     | Measured / Predicted | Optimal Estimate | 90%                 | 95%          |
| 19 | C13T1   | R101C(1)                   | HETEROSCEDASTICITY | MEASURED             | 1.536279E-01     | 2.572101E-01        | 3.067469E-01 |
| 20 | C13T2   | R101C(1)                   | "                  | "                    | "                | "                   | "            |
| 21 | C13T3   | R101C(1)                   | "                  | "                    | "                | "                   | "            |
| 22 | C13T4   | R101C(1)                   | "                  | "                    | "                | "                   | "            |
| 23 | C13T1   | R101C(3)                   | HETEROSCEDASTICITY | MEASURED             | 1.304752E-01     | 2.613760E-02        | 3.117151E-02 |
| 24 | C13T2   | R101C(3)                   | "                  | "                    | "                | "                   | "            |
| 25 | C13T3   | R101C(3)                   | "                  | "                    | "                | "                   | "            |
| 26 | C13T4   | R101C(3)                   | "                  | "                    | "                | "                   | "            |
| 27 | C13T1   | R101T                      | HETEROSCEDASTICITY | MEASURED             | 4.858160E-01     | 6.051082E-02        | 7.216476E-02 |
| 28 | C13T2   | R101T                      | "                  | "                    | "                | "                   | "            |
| 29 | C13T3   | R101T                      | "                  | "                    | "                | "                   | "            |
| 30 | C13T4   | R101T                      | "                  | "                    | "                | "                   | "            |
| 31 |   |                            |                    |                      |                  |                     |              |
| 32 |   |                            |                    |                      |                  |                     |              |
| 33 |   |                            |                    |                      |                  |                     |              |
| 34 |   |                            |                    |                      |                  |                     |              |
| 35 |   |                            |                    |                      |                  |                     |              |
| 36 |   |                            |                    |                      |                  |                     |              |
|    | Statistical Significance  | Variance Matrix            | Correlation Matrix | Information Matrix   | C13T1..R10       |                     |              |

## Sample measured variable sheet

|  | A                              | B                        | C                | D  | E             | F            | G             |
|--|--------------------------------|--------------------------|------------------|--|---------------|--------------|---------------|
| 1  | Parameter estimation problem:  |                          |                  | ESTIMATE REACTOR (Optimal)                     |               |              |               |
| 2  |                                |                          |                  |  |               |              |               |
| 3  | Measured variable information: |                          |                  |  |               |              |               |
| 4  | C13T1.R101.C(1)                |                          |                  |  |               |              |               |
| 5  |                                |                          |                  |  |               |              |               |
| 6  | Variable Values                |                          |                  | Deviation                                      |               |              |               |
| 7  | Measurement Time               | Experimental Measurement | Model Prediction | Standard Deviation of Experimental Measurement | Absolute      | Percentage   | Weighted      |
| 8  | 10                             | 3.74691E+03              | 3.720950E+03     | 2.906728E+01                                   | 2.596140E+01  | 0.692874844  | 8.931432E-01  |
| 9  | 20                             | 2.518600E+03             | 2.513494E+03     | 2.256775E+01                                   | 5.106240E+00  | 0.2027412    | 2.262627E-01  |
| 10   | 30                             | 1.504120E+03             | 1.508002E+03     | 1.624977E+01                                   | -5.682098E+00 | -0.377768929 | -3.496725E-01 |
| 11   | 40                             | 8.837833E+02             | 8.889517E+02     | 1.157939E+01                                   | -5.168441E+00 | -0.584808606 | -4.463253E-01 |
| 12   | 50                             | 5.480505E+02             | 5.438444E+02     | 8.540537E+00                                   | 4.206106E+00  | 0.76746678   | 4.924872E-01  |
| 13   | 60                             | 3.486161E+02             | 3.528859E+02     | 6.401829E+00                                   | -4.269753E+00 | -1.22477221  | -6.669583E-01 |
| 14   | 70                             | 2.39515E+02              | 2.427614E+02     | 5.040071E+00                                   | -3.249896E+00 | -1.356885245 | -6.448116E-01 |
| 15   | 80                             | 1.754503E+02             | 1.757188E+02     | 4.134552E+00                                   | -2.615368E+00 | -0.149068025 | -6.327322E-02 |
| 16   | 90                             | 1.350508E+02             | 1.326633E+02     | 3.498625E+00                                   | 2.387515E+00  | 1.767864162  | 6.824152E-01  |
| 17   | 100                            | 1.009639E+02             | 1.036862E+02     | 2.906893E+00                                   | -2.716335E+00 | -2.690242781 | -9.344591E-01 |
| 18   | 110                            | 8.729780E+01             | 8.420528E+01     | 2.649495E+00                                   | 3.092517E+00  | 3.542491369  | 1.167210E+00  |
| 19   | 120                            | 7.413863E+01             | 7.088850E+01     | 2.387549E+00                                   | 3.250127E+00  | 4.383850925  | 1.361282E+00  |
| 20   | 130                            | 6.209760E+01             | 6.120963E+01     | 2.132612E+00                                   | 8.879673E-01  | 1.429954363  | 3.163756E-01  |
| 21   | 140                            | 5.454854E+01             | 5.385582E+01     | 1.963568E+00                                   | 6.927200E-01  | 1.269914911  | 3.527863E-01  |
| 22   | 150                            | 4.905154E+01             | 4.807940E+01     | 1.835077E+00                                   | 9.721427E-01  | 1.981880042  | 5.297557E-01  |
| 23   | 160                            | 4.298377E+01             | 4.342218E+01     | 1.687002E+00                                   | -4.384078E-01 | -1.019338008 | -2.598740E-01 |
| 24   | 170                            | 3.961377E+01             | 3.958755E+01     | 1.601488E+00                                   | 2.621978E-02  | 0.066188557  | 1.637214E-02  |
| 25   | 180                            | 3.571447E+01             | 3.637551E+01     | 1.439170E+00                                   | -6.610380E-01 | -1.850896872 | -4.409359E-01 |
| 26   | 190                            | 3.239367E+01             | 3.364579E+01     | 1.425523E+00                                   | -6.462240E-01 | -1.958280108 | -4.533243E-01 |
| 27   | 200                            | 3.212896E+01             | 3.129694E+01     | 1.401445E+00                                   | 8.320247E-01  | 2.589641027  | 5.936908E-01  |
| 28   | 210                            | 2.969873E+01             | 2.925456E+01     | 1.332944E+00                                   | 4.441718E-01  | 1.495592011  | 3.332262E-01  |
| 29   | 220                            | 2.643707E+01             | 2.746243E+01     | 1.237715E+00                                   | -1.025360E+00 | -3.878432916 | -8.284297E-01 |
| 30   | 230                            | 2.695787E+01             | 2.587719E+01     | 1.253195E+00                                   | 1.080677E+00  | 4.008760926  | 8.623370E-01  |
| 31   | 240                            | 2.475480E+01             | 2.448500E+01     | 1.186937E+00                                   | 2.898039E-01  | 1.170637818  | 2.441611E-01  |
| 32   | 250                            | 2.254412E+01             | 2.318996E+01     | 1.118261E+00                                   | -6.548383E-01 | -2.904696495 | -5.855862E-01 |
| 33   | 260                            | 2.262180E+01             | 2.205751E+01     | 1.120715E+00                                   | 5.642871E-01  | 2.494439578  | 5.035065E-01  |
| 34   | 270                            | 2.140928E+01             | 2.102314E+01     | 1.082060E+00                                   | 3.861447E-01  | 1.803632479  | 3.568607E-01  |
| 35   | 280                            | 2.096655E+01             | 2.008143E+01     | 1.067749E+00                                   | 8.851248E-01  | 4.221604417  | 8.289632E-01  |
| ◀ ▶ 🔍   Correlation Matrix   Information Matrix   Sheet20   C13T1.R101.C(1)   C13T1.R101.C(3)   C13T |                                |                          |                  |  |               |              |               |

## Graphical analysis

In addition to the spreadsheets, the user can also create three types of statistical plots by choosing the Plot option from the Parameter Estimation menu on the MS Excel tool bar. Each of these is described below:

- **Overlay chart.** This overlays the measured and predicted values for whichever variable is chosen from a specified experiment. The user has the option of showing error bars by checking the Show standard deviations dialog box (see appropriate figure in the next chapter). These error bars correspond to  $\tilde{z}_{ijk} \pm \sigma_{ijk}$ . The user selects one or more variables (choose Flat list, if you want to select more than one variable). The charts can be plotted on the respective variable data worksheets, on a new sheet for each chart or all charts on one new sheet. A figure showing a plot of the batch reactor variable R101.C(1) in experiment C13T1 can be found in the following chapter. It can be seen that the agreement between the predicted values using the estimated parameters and the measured values is very close.
- **Residual chart.** This allows the user to plot the absolute deviation, relative deviation and/or weighted (scaled) residual for each measured variable in each experiment. As for the "Overlay chart", the user may select one or more variables and plot the charts on the respective data worksheets, on a new sheet for each chart or all the charts on one new sheet. Figures in the following chapter show the dialog window and weighted residual plot, respectively, for the batch reactor variable R101.C(1) in experiment C13T1.
- **Confidence ellipsoids.**

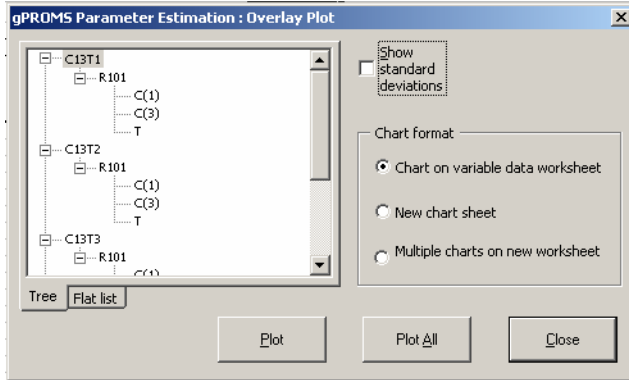
This option gives the confidence ellipsoids in two-dimensional space for any pair of parameters ( $p_i, p_j$ ) according to the joint confidence region:

$$\begin{pmatrix} p_i - \hat{p}_i \\ p_j - \hat{p}_j \end{pmatrix}^T \begin{pmatrix} V_{ii} & V_{ij} \\ V_{ij} & V_{jj} \end{pmatrix}^{-1} \begin{pmatrix} p_i - \hat{p}_i \\ p_j - \hat{p}_j \end{pmatrix} \leq N_p s^2 F(\alpha, N_p, N - N_p).$$

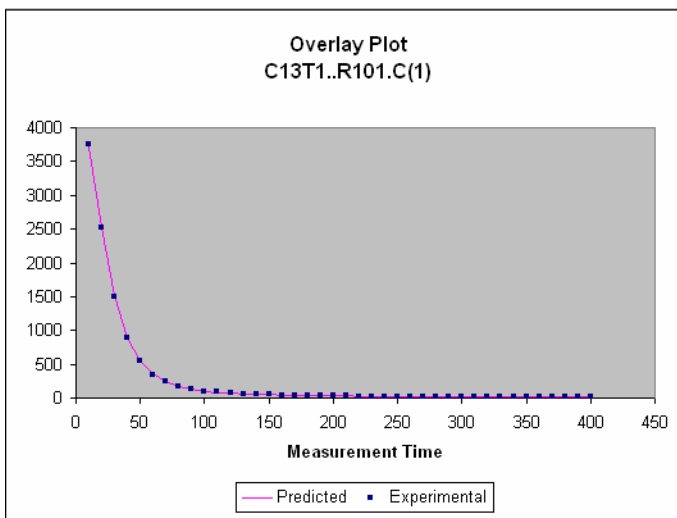
The parameters can be plotted against each other using different confidence levels (90%, 95% or 99%). The user may select any pair of parameters or plot the confidence ellipsoids for all pairs. An example of a confidence ellipsoid is shown in the next chapter.

**Note:** Although the variance model parameters have uncertainties, the user may wish to exclude these uncertainties from the calculations. In this case the vector  $p^t$  does not contain the parameters  $\beta^t$  and is defined as  $p^t = \theta^t$ .

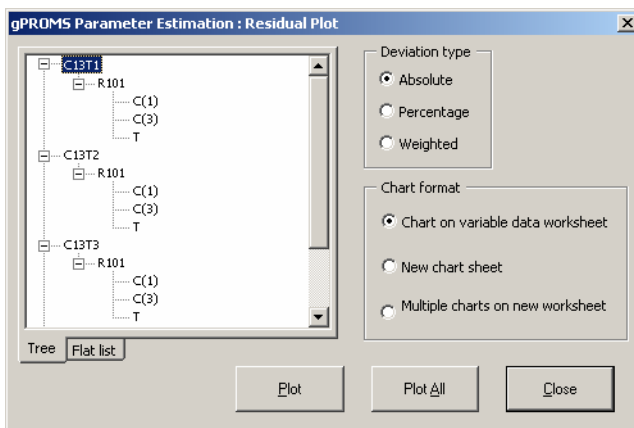
## Overlay Plot Dialog Window



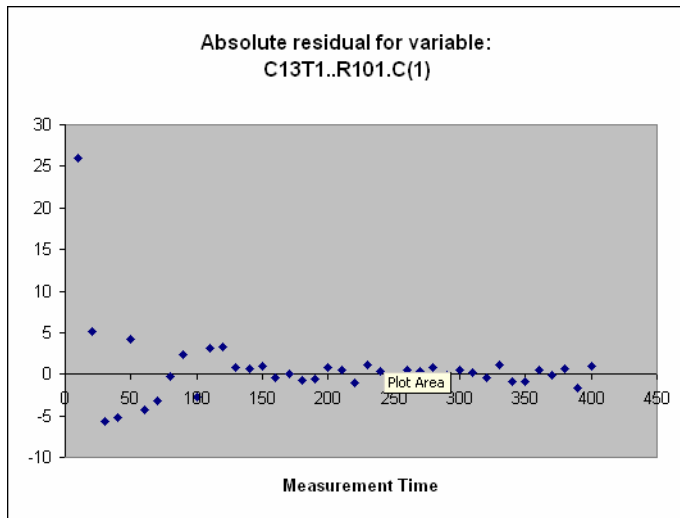
## Sample Overlay Plot



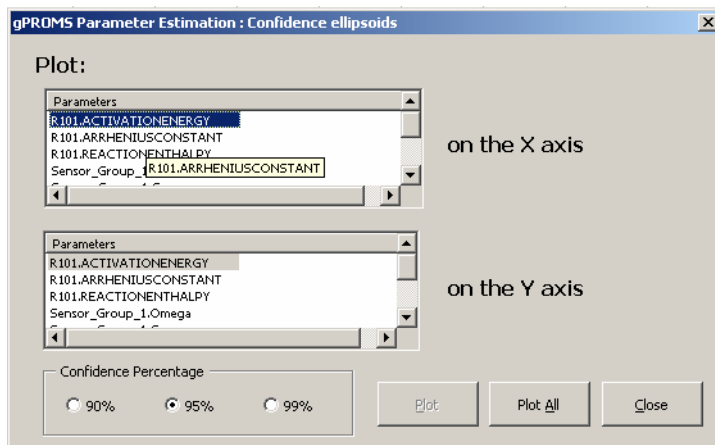
## Residual Plot Dialog Window



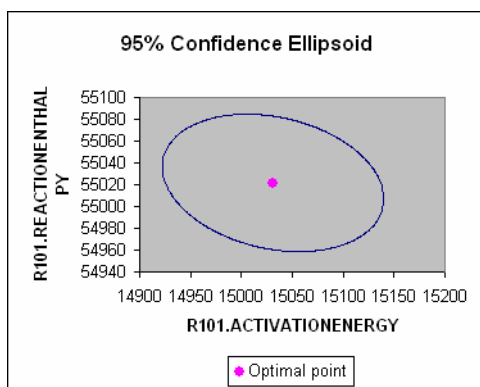
## Sample Weighted Residual Plot



## Confidence Ellipsoid Dialog Window



## Sample Confidence Ellipsoid Plot.

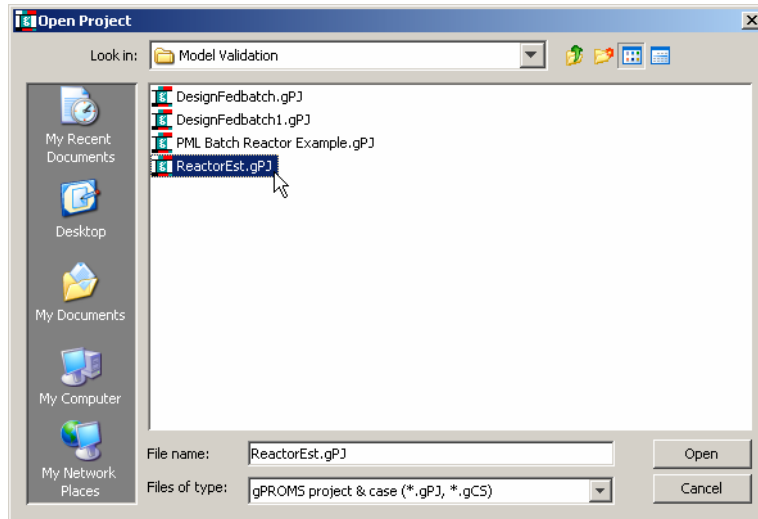


## A Parameter Estimation example

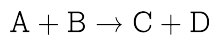
The gPROMS installation includes a simple Parameter Estimation example - it is highly recommended that you take a look at this. You can access this by clicking on the browse examples button on the gPROMS Tool bar

and then navigating to General capabilities\Model Validation\ReactorEst.gPJ (as shown below). This section explains the features of the model.

**Figure 2.21. Browsing for the Batch Reactor Parameter Estimation example in the gPROMS distribution directory.**

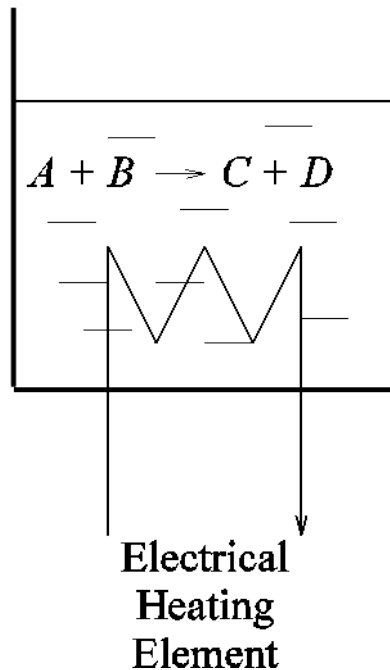


The example considers the batch reactor shown in the figure below. The reactor is used to carry out the following liquid-phase endothermic reaction:



An electrical heating element is used to provide the necessary heat to the reactor. The duty of this heating element is the control variable in the experiment.

**Figure 2.22. Batch reactor.**



Experimental and theoretical evidence indicates that the order of the reaction with respect to both  $A$  and  $B$  is one. We also know that the kinetic constant of the reaction follows an Arrhenius-type temperature dependence relationship. Based on this information, and assuming perfect mixing and ideal liquid mixture behaviour, we can construct the following mathematical model for the process:

Component mass balance

$$\frac{dM_i}{dt} = \nu_i r V, \quad i = A, B, C, D.$$

Energy balance

$$\frac{dH}{dt} = r V (-\Delta H_R) + Q$$

$$r = k C_A C_B$$

Arrhenius temperature dependence

$$k = k_0 e^{-E/RT}$$

Component concentrations

$$C_i = \frac{M_i}{V}, \quad i = A, B, C, D, V = \sum_{i=A,B,C,D} \frac{M_i}{\rho_i}$$

Energy content

$$H = \sum_{i=A,B,C,D} M_i \hat{H}_i, \hat{H}_i = \alpha_i (T - T_{ref}) + \frac{\beta_i}{2} (T^2 - T_{ref}^2), \quad i = A, B, C, D$$

The tables below summarise the parameters and variables that appear in this model.

**Table 2.2. Parameters in batch reactor Model.**

| Model Parameters   | Description   |
|--------------------|---|
| $E$                | Reaction activation energy                            |
| $k_0$              | Arrhenius constant                                    |
| $T_{ref}$          | Reference temperature                                 |
| $\alpha_i \beta_i$ | Specific molar enthalpy coefficients of component $i$ |
| $\delta h_r$       | Heat of reaction                                      |
| $\nu_i$            | Stoichiometric coefficient of component $i$           |
| $\rho_i$           | Molar density of component $i$                        |

**Table 2.3. Variables in batch reactor Model.**

| Model Variables | Description                              |
|-----------------|--|
| $C_i$           | Molar concentration of component $i$     |
| $H$             | Total enthalpy content of the reactor    |
| $\hat{H}_i$     | Specific molar enthalpy of component $i$ |
| $k$             | Reaction constant                        |
| $M_i$           | Molar holdup of component $i$            |
| $Q$             | Heating load                             |
| $r$             | Reaction rate                            |
| $T$             | Reactor temperature                      |

|     |                                       |
|-----|---------------------------------------|
| $V$ | Total volume of liquid in the reactor |
|-----|---------------------------------------|

Before this model can be used to simulate or optimise the operation of the reactor, all parameters that appear in it must be given fixed values. Imagine, however, a situation where the kinetic characteristics of the reaction are unknown. That is, we do not know the values of  $k_0$ ,  $E$  and  $\delta h_r$ . In order to determine these, we can perform a number of experiments and measure the values of some or all process variables.

Four experiments have been performed, all under identical conditions. The duration of each experiment is 400s, the reactor is initially loaded with 100 mol of  $A$ , 100 mol of  $B$ , 1 mol of  $C$  and 1 mol of  $D$ , and the initial temperature is 293K. The heating element power supply is set to 505kW for the first 100s of operation and to zero thereafter. The experimental conditions are summarised below:

**Table 2.4. Conditions for batch reactor experiments #1 to #4**

| Initial conditions      | Heating policy                             |
|-------------------------|--|
| $M_A = 100 \text{ mol}$ | $Q = 505 \text{ kW}, 0 \leq t(s) \leq 100$ |
| $M_B = 100 \text{ mol}$ | $Q = 0, 100 < t(s) \leq 400$               |
| $M_C = 1 \text{ mol}$   |  |
| $M_D = 1 \text{ mol}$   |  |
| $T = 293 \text{ K}$     |  |

The data from four typical experiments are shown in the table below. The reactor temperature and the concentrations of reactant  $A$  and product  $C$  are automatically measured every 10 seconds.

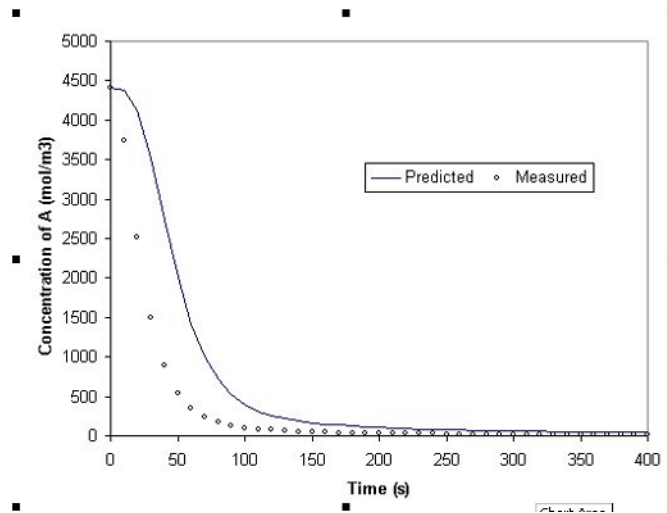
**Table 2.5. Data from the batch reactor experiments**

| Experiment # | Time $t$ (s) | Temperature $T$ (K) | $C_A$ (mol m <sup>-3</sup> ) | $C_C$ (mol m <sup>-3</sup> ) |
|--------------|--------------|---------------------|------------------------------|------------------------------|
| 1            | 10           | 412.549             | 3746.911                     | 898.552                      |
|              | 20           | 504.179             | 2518.600                     | 2358.859                     |
|              | 30           | 603.486             | 1504.120                     | 3634.895                     |
|              | 40           | 715.300             | 883.783                      | 4407.922                     |
|              | 50           | 837.841             | 548.051                      | 4849.769                     |
|              | ...          | ...                 | ...                          | ...                          |
| 2            | 10           | 411.804             | 3700.752                     | 898.216                      |
|              | 20           | 504.530             | 2520.929                     | 2380.372                     |
|              | 30           | 604.027             | 1502.396                     | 3626.625                     |
|              | 40           | 714.709             | 894.412                      | 4388.874                     |
|              | 50           | 837.154             | 545.142                      | 4828.031                     |
|              | ...          | ...                 | ...                          | ...                          |
| 3            | 10           | 412.823             | 3705.683                     | 896.799                      |
|              | 20           | 503.740             | 2513.977                     | 2386.384                     |
|              | 30           | 604.359             | 1516.819                     | 3633.759                     |
|              | 40           | 715.539             | 895.847                      | 4419.044                     |

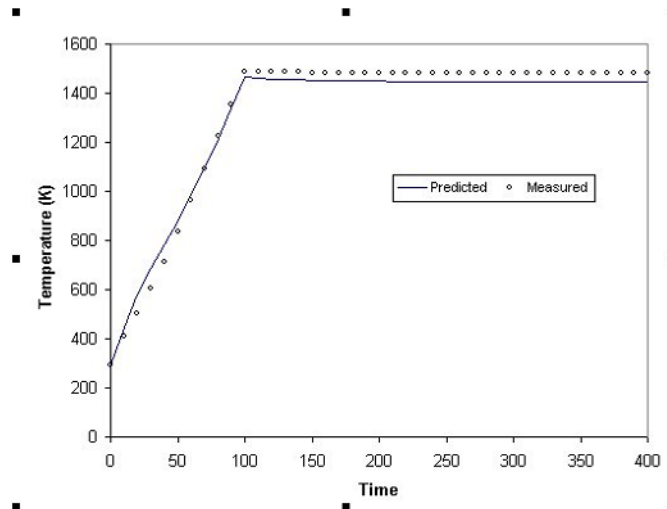
| Experiment # | Time $t$ (s) | Temperature $T$ (K) | $C_A$ (mol m <sup>-3</sup> ) | $C_c$ (mol m <sup>-3</sup> ) |
|--------------|--------------|---------------------|------------------------------|------------------------------|
| 4            | 50           | 836.518             | 551.290                      | 4796.965                     |
|              | ...          | ...                 | ...                          | ...                          |
|              | 10           | 411.235             | 3711.353                     | 902.858                      |
|              | 20           | 503.830             | 2526.062                     | 2409.363                     |
|              | 30           | 603.157             | 1527.653                     | 3653.779                     |
|              | 40           | 716.325             | 895.004                      | 4400.964                     |
|              | 50           | 837.703             | 543.662                      | 4824.272                     |
|              | ...          | ...                 | ...                          | ...                          |

We note that, for any given set of values of the unknown parameters, the model equations can be solved to predict the reactor behaviour at the experimental conditions. For instance, the plots below compares the predicted variation of  $C_A(t)$  and  $T(t)$  for  $k_0 = 6.5 \times 10^{-4} \text{ m}^3 \text{ mol}^{-1} \text{ s}$ ,  $E = 20000 \text{ J/mol}$  and  $\Delta H_R = 55000 \text{ J/mol}$ , against the data obtained from the first experiment. It can be seen that the model predictions are rather poor.

**Figure 2.23. Comparison of experimental composition data with model predictions based on nominal values of  $k_0$ ,  $E$  and  $\delta h_r$**



**Figure 2.24. Comparison of experimental temperature data with model predictions based on nominal values of  $k_0$ ,  $E$  and  $\delta h_r$**





In practice, we obviously need to determine the values of the unknown parameters,  $k_0$ ,  $E$  and  $\delta h_r$ , in order to maximise the probability that the model will predict the values obtained from the experiments. This is the objective of the Parameter Estimation problem.

In our example we will assume that the following initial guess and bounds hold for the 3 different parameters of interest:

- The Arrhenius coefficient,  $k_0$  (gPROMS pathname R101.ARRHENIUSCONSTANT), has an initial guess 5E-4 and the lower and upper bounds on its value are 1E-4 and 1E-3 respectively.
- The reaction activation energy,  $E$  (gPROMS pathname R101.ACTIVATIONENERGY), has an initial guess of 15000 and the lower and upper bounds on its value are 10000 and 20000 respectively.
- The heat of reaction,  $\delta h_r$  (gPROMS pathname R101.REACTIONENTHALPY), has an initial guess of 55000 and the lower and upper bounds on its value are 35000 and 75000 respectively.

Here the measured variable is assumed to be described by the same variance model for each of the experiments that were performed. Hence:

- The concentration of reactant A,  $C_A$  (gPROMS pathname R101.C(1)), and the concentration of product C,  $C_C$  (gPROMS pathname R101.C(3)), are both described by HETEROSCEDASTIC models on the PREDICTED\_VALUES. The initial guess for  $\omega$  is 0.1, with lower and upper bounds of 0.01 and 3 respectively, while the initial guess for  $\gamma$  is 0.5 with bounds of 0 and 1.
- The reactor temperature,  $T$  (gPROMS pathname R101.T), is also described by a HETEROSCEDASTIC model. The initial guess for  $\omega$  is 0.1, with bounds of 0.01 and 3. However,  $\gamma$  is fixed at zero (the measurement error for the temperature is a constant value and does not depend on the size of  $T$ ).

---

# Chapter 3. Experiment Design in gPROMS

gPROMS can be used to perform *Experiment Design for parameter precision* for complex steady-state and dynamic models. This chapter covers the following issues:

- a description of what an Experiment Design problem is,
- the Experiments to be designed and Experiment Design entities that are used to specify Parameter Estimation problems in gPROMS,
- how to analyse the results from an Experiment Design activity,
- how to control the execution of an Experiment Design activity.

Explanatory screen shots are taken from a fed batch fermentation problem included in the gPROMS distribution. Some basic familiarity with the gPROMS language and concepts is assumed.

## What is Experiment Design?

Experiments are used to improve the understanding of processes and create accurate models. The quality of information generated by experiments depends strongly on the experimental conditions as well as what is measured and when it is measured. The aim is to design experiments in order to maximise the information content of the measurements in the context of their utilisation for estimating the model parameters. This is equivalent to minimising the variances of the parameters to be estimated. The variances are a measure for the uncertainty of the parameters, also represented by individual confidence interval approximations. Experiment Design therefore aims at minimising the variances (the uncertainty) of the parameters to be estimated.

Experiment Design in gPROMS supports the design of one or more new experiments. Designing more than one experiment simultaneously is of interest in cases where multiple sets of experimental apparatus are available, allowing several experiments to be carried out in parallel, or when the experiment duration is relatively short, in which case one may wish to design simultaneously all experiments to be performed within a single day or week. The Experiment Design calculations can take into account any experiments that have already been performed. This may, for example, result in different regions of the operating space being explored from the point of view of yielding maximum information.

Experiment Design for parameter precision aims at determining optimal experimental settings and measurement times in order to maximise the information content from the measured data generated by these experiments. This is equivalent to minimising the confidence ellipsoid of the parameters to be estimated.

## Underlying mathematics of Experiment Design

In mathematical terms, we want to minimise some measure  $\psi$  of the variance-covariance matrix,  $\mathbf{V}_\theta$ , of the parameters to be estimated:

$$\min_{\xi} \Psi(\mathbf{V}_\theta)$$

The experiment decision variables  $\xi$  may be subject to equality or inequality constraints:

$$c(\xi) \geq 0$$

$$\xi^{\min} \leq \xi \leq \xi^{\max}$$

The variance-covariance matrix is of the form:

$$\mathbf{V}_\theta = (H_\theta^*)^{-1}$$

The information matrix  $h_\theta^*$  is a  $n_\theta \times n_\theta$  matrix, where  $n_\theta$  is the number of parameters to be estimated,  $\theta$ . It is given by

$$(H_\theta^*)_{\mu,\nu} = \sum_{l=1}^{N_{exp}} \sum_{i \in \mathcal{MV}_l} \sum_{m=1}^{NM_{il}} \left( \frac{\left( \frac{\partial}{\partial \theta_\mu} z_{il}(\rho_{iml}) \right) \left( \frac{\partial}{\partial \theta_\nu} z_{il}(\rho_{iml}) \right)}{\sigma_{il}^2(z_{il}(\rho_{iml}), \beta_{il})} \right), \mu, \nu = 1, \dots, N_\theta$$

The symbols in these equations have the following definitions:

**Table 3.1. Symbol definitions.**

|   |   |
|---|---|
| $\mathbf{V}_\theta$                             | The variance-covariance matrix of the parameters to be estimated.   |
| $H_\theta^*$                                    | & The information matrix of the parameters to be estimated.   |
| $\xi$   | The set of experiment decision variables in all experiments.  |
| $n_{exp}$                                       | The number of experiments.  |
| $\mathcal{MV}_l$                                | The set of measured variables in experiment $l$ , i.e. $\mathcal{MV}_l \subseteq \{1, \dots, N_{x_l} + N_{y_l}\}$ |
| $nm_{il}$                                       | The number of sampling points for measured variable $i$ in experiment $l$ , $i \in \mathcal{MV}_l$ .              |
| $\rho_{iml}$                                    | The $m$ -th measurement time for variable $i$ in experiment $l$ .   |
| $z_{il}(\rho_{iml})$                            | The model-predicted value of variable $i$ at time point $\rho_{iml}$ in experiment $l$ .                          |
| $\sigma_{il}^2(z_{il}(\rho_{iml}), \beta_{il})$ | The variance of the measurement error of variable $i$ at time point $\rho_{iml}$ in experiment $l$ .              |

Equipment design and resource availability may impose certain operational and safety limits within which our control manipulations should be maintained. Experiment Design seeks to provide answers to the following questions

- What should be the initial conditions for the experiment?
- How long should we run the experiment for?
- How should we vary the controls (e.g. the time profiles of feed flowrates and external heating/cooling loads)?
- When should we take the measurement samples?

The overall aim is to generate the maximum amount of information for a subsequent estimation of the parameters ( $\theta_1 \dots \theta_4$ ) using gPROMS' Parameter Estimation facilities, while trying to maintain the process within the required operating envelop.

In order to compare the magnitude of different variance-covariance matrices, various real-valued functions have been suggested as a measure of "smallness". gPROMS supports three well-known criteria:

- **A-optimality:** minimise the trace of the variance-covariance matrix:

$$\Psi_A(\mathbf{V}_\theta) = \frac{1}{N_\theta} \sum_{\mu=1}^{N_\theta} (\mathbf{V}_\theta)_{\mu,\mu}$$

This minimises the sum of the variances of the individual parameter estimates. It corresponds to minimising the dimensions of the smallest hyper rectangle within which the confidence ellipsoid can be inscribed.

- **D-optimality:** minimise the determinant of the variance-covariance matrix:

$$\Psi_D(\mathbf{V}_\theta) = \det(\mathbf{V}_\theta)^{\frac{1}{N_\theta}}$$

This is also known as the minimum volume criterion since it minimises the volume of the confidence ellipsoid.

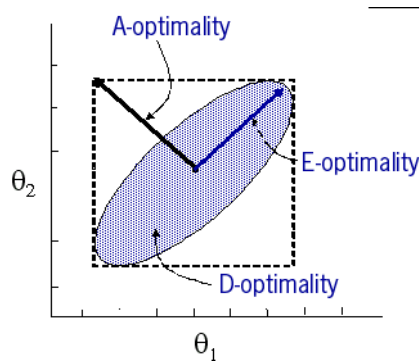
- **E-optimality:** minimise the largest eigenvalue of the variance-covariance matrix:

$$\Psi_E(\mathbf{V}_\theta) = \lambda_{\max}(\mathbf{V}_\theta)$$

The eigenvalues of the variance-covariance matrix correspond to the lengths of the minor and major axes of the confidence ellipsoid. By minimising the largest eigenvalue, the design renders the confidence ellipsoid as spherical as possible.

The figure below shows a graphical interpretation of the different design criteria for a two-dimensional confidence ellipsoid.

**Figure 3.1. Interpretation of the different design criteria.**



## Experiments to be designed

The Experiments to be designed entity is used to specify the full details of the experiments to be designed.

Note experiments that have already been performed and which you want to take into account when designing an experiment are specified in exactly the same way as for Parameter Estimation.

In order to create a new Experiments Designed entity, do the following:

- Pull-down the Entity menu from the top pane in gPROMS ModelBuilder.
- Click on New Entity. A dialog box will appear.
- Choose Experiment to be designed for the Entity type and fill in the Name field.

When you do the above, an Experiments to be designed entity editor window will open up in the right hand pane, as shown in the figure below.

**Figure 3.2. Experiments to be designed entity editor**

The Experiments to be designed entity editor has the following tabs:

- The General tab - a form for defining the Process entity to use; the type of experiment to determine (steady-state or dynamic) and which initial conditions should be determined .
- The Controls tab - a form for specifying the time-invariant, piecewise constant and piecewise linear variables that can be adjusted during the course of the experiment and that are to be optimally determined by the Experiment Design problem.
- The Constraints tab - a form for defining any endpoint or interior-point constraints that need to be placed on the experiment that will be designed.
- The Measured data tab - a for entering the quantities that can be measured and the sensor used to do this.
- The gPROMS language tab - this provides a convenient means of reviewing all the information regarding the experiment to be designed in a text-based format. You are not advised to use this tab to enter data.
- Properties (a record of when the entity was created, last edited as well as the default tab to open on)

## General information

In order to define the Process entity to use for a given experiment, click on the General tab in the Experiments to be designed entity editor and then choose the desired Process from the drop down menu at the top of the window as shown in the figure below.

**Figure 3.3. Specifying the Process entity that will be used in conjunction with the experimental data.**

EXPERIMENT TO BE DESIGNED FEDBATCH\_DES (DesignFedbatch)

Process: FedBatch\_Des

Type of experiment: ☒ Dynamic ☐ Steady state

Experiment duration:

| Initial guess | Fixed? | Lower bound | Upper bound |
|---------------|--------|-------------|-------------|
|               |        |             |             |

Initial conditions:

| Variable name | Initial guess | Fixed? | Lower bound | Upper bound |
|---------------|---------------|--------|-------------|-------------|
| <new>         |               |        |             |             |

Delete row

General Controls Constraints Measurements and sensors gPROMS language Properties

gPROMS can be used to design both steady-state and dynamic experiments. If designing a dynamic experiment then you need to decide whether to optimally determine the duration of the experiment and the initial conditions for the experiment:

- Determine experiment duration?
  - Fixed the experiment duration; give the value in the Initial guess field
  - Give the initial guess and lower and upper bounds for the duration of the experiment - gPROMS will determine an experiment duration between these bounds
- What initial conditions?
  - Type the full pathname of the gPROMS differential variable that you wish to specify an initial condition for in the <new> cell in the initial conditions table.. You MUST only provide initial conditions for variables that are given initial values directly in the Process.
  - If you do not wish to type the pathname you can select the appropriate variable from a drop down list by holding the CTRL key down and hitting the SPACE button on your keyboard. The first time you do this, gPROMS will give you a list of relevant Units and Variables. Select the desired one and then repeat as necessary to drill down to the desired Variable.
  - Give the initial guess and lower and upper bounds for the initial condition - gPROMS will determine the value of the initial condition between these bounds. If you want to fix the initial condition at a constant value then *check* the fixed box.

**Figure 3.4. Specifying the initial conditions for the relevant differential variables in an experiment to be designed.**

EXPERIMENT TO BE DESIGNED FEDBATCH\_DES (DesignFedbatch)

Process: FedBatch\_Des

Type of experiment: ☒ Dynamic ☐ Steady state

Experiment duration:

| Initial guess | Fixed?                   | Lower bound | Upper bound |
|---------------|--------------------------|-------------|-------------|
| 10.0          | <input type="checkbox"/> | 5.0         | 10.0        |

Initial conditions:

| Variable name | Initial guess | Fixed?                   | Lower bound | Upper bound |
|---------------|---------------|--------------------------|-------------|-------------|
| FERMENT.y1    | 7.0           | <input type="checkbox"/> | 1.0         | 10.0        |
|               |               |                          |             |             |

Delete row

General Controls Constraints Measurements and sensors gPROMS language Properties

## Controls

An experiment control is a variable that can be adjusted during an experiment (such as a heating policy in a reactor). Experiment controls therefore form part of the "recipe" for running the experiment. The user can specify the variation in a Variable value using one of three different mechanisms:

- *Time-invariant controls* - provide a single variable value, this quantity will be the same for the duration of the experiment. Appropriate for steady-state and dynamic experiments.
- *Piecewise constant controls* - provide multiple variable values, the quantity holds a different constant value for specified time intervals during the experiment. Appropriate for dynamic experiments only.
- *Piecewise linear controls* - provide multiple variable values, the quantity varies linearly (from a start value to an end value) in a calculated time interval during the experiment. Appropriate for dynamic experiments only.

So when designing an experiment; gPROMS determines the optimal values and profiles for time-invariant, piecewise constant and piecewise linear control variables.

Note that any Variables that are Assigned values in a Process but are not specified as controls in the experiment will retain their Assigned values, whether these are constants or functions of TIME.

In order to define the experiment controls, click on the *Controls* tab in the *Experiments to be designed* entity editor and do the following:

- Click the <new> cell in the *Control variables* table at the top of the window and type the full pathname of the gPROMS control variable<sup>1</sup>. If you do not wish to type the pathname you can select the appropriate variable from a drop down list by holding the CTRL key down and hitting the SPACE button on your keyboard. The first time you do this, gPROMS will give you a list of relevant Units and Variables. Select the desired one and then repeat as necessary to drill down to the desired Variable.
- Choose from the drop down list in the right-hand column whether the control variable is time-invariant, piecewise constant or piecewise linear. Note that if a steady-state experiment has been specified on the general tab then you only have the option to select time-invariant controls.
- Repeat the above for each new control variable you wish to use.

<sup>1</sup>NOTE: This variable should be one of those you ASSIGNED in the associated Process entity.

Having chosen the format for the time-varying behavior of the control variable, select the relevant control by clicking on it, and then go to the lower part of the window to enter the details of its allowable behavior during the experiment. If one or more piece-wise constant or piece-wise linear controls have been selected then first define the time intervals.: **the Control time intervals are synchronised for all piecewise constant and piecewise linear controls and adding or deleting control intervals therefore affects all control variables**

- Provide an initial guess and bounds on the allowable time duration for each interval
  - To add another interval enter the value in the <new> row<sup>2</sup>
1. *Time-invariant controls* - provide a value together with lower and upper bounds. gPROMS will determine a value between these bounds.
  2. *Piece-wise constant controls* - for each interval specify an initial guess for the control along with lower and upper bounds; gPROMS will determine a value between these bounds.
  3. *Piece-wise linear controls* - for each interval gPROMS will determine a starting value for the control and a final value for the control variable - provide an initial guess for the starting and final value of the control along with lower and upper bounds; gPROMS will determine values between these bounds.

Should you wish to duplicate or delete an entire interval, simply select it and click on the Duplicate or Delete buttons at the bottom of the window.

**Figure 3.5. Specifying the bounds on a piecewise constant control profile that is to be optimally determined.**

| Interval guess | Fixed interval?          | Interval lower bound | Interval upper bound | Start time | End time | Initial guess | Fixed value?             | Lower bound | Upper bound |
|----------------|--------------------------|----------------------|----------------------|------------|----------|---------------|--------------------------|-------------|-------------|
| 2.0            | <input type="checkbox"/> | 0.5                  | 10.0                 | 0.0        | 2.0      | 0.2           | <input type="checkbox"/> | 0.05        | 0.2         |
| 2.0            | <input type="checkbox"/> | 0.5                  | 10.0                 | 2.0        | 4.0      | 0.2           | <input type="checkbox"/> | 0.05        | 0.2         |
| 2.0            | <input type="checkbox"/> | 0.5                  | 10.0                 | 4.0        | 6.0      | 0.2           | <input type="checkbox"/> | 0.05        | 0.2         |
| 2.0            | <input type="checkbox"/> | 0.5                  | 10.0                 | 6.0        | 8.0      | 0.2           | <input type="checkbox"/> | 0.05        | 0.2         |
| 2.0            | <input type="checkbox"/> | 0.5                  | 10.0                 | 8.0        | 10.0     | 0.2           | <input type="checkbox"/> | 0.05        | 0.2         |
| <new>          |                          |                      |                      |            |          |               |                          |             |             |

## Constraints

The purpose of the Constraints tab is to allow the user to impose safety and operational limits on the designed experiment to ensure. This ensures that the result of the Experiment Design problem is an experiment that is allowable and feasible to run in practice. To capture these limits, gPROMS allows for the specification of *endpoint* and *interior-point* constraints on the predicted behavior of variables in the experiment that is to be designed. There are no interior point constraints for steady-state models.

<sup>2</sup>When you enter the interval duration gPROMS automatically determines the start and end times of that interval, but the relevant cells are greyed out and you cannot edit them.



In order to define an **endpoint constraint**, click on the Constraints tab in the Experiments to be designed entity editor and do the following:

- Click the <new> cell in the Variable name column of the Endpoint constraints table at the top of the window and type the full pathname of the gPROMS variable. Note that *pathname completion* can be activated by pressing **Ctrl+Space**.
- Specify the lower and upper bounds on the Variable. When gPROMS solves the Experiment Design problem it will ensure that the Variable indicated is restricted to within the specified bounds. Alternatively, check the Fixed? cell and specify the constraint value if you desire an endpoint equality constraint. Note: endpoint equality constraints are very challenging for numerical reasons and it is usually better to specify an endpoint inequality constraint with a tight lower and upper bound.
- Repeat the above for each new endpoint constraint you wish to add. If you wish to delete an endpoint constraint, click the *Delete* button at the bottom of the window.

Interior-point constraints are *inequality constraints* defined at the *start* of every time interval. Such constraints are slightly more complicated to specify because of the fact that it is possible to have the bounds varying from one interval to another.

In order to define an **interior-point constraint**, click on the Constraints tab in the Experiments to be designed entity editor and do the following:

- Click the <new> cell in the Variable name column of the Endpoint constraints table at the top of the window and type the full pathname of the gPROMS variable. If you do not wish to type the pathname you can select the appropriate variable from a drop down list by holding the CTRL key down and hitting the SPACE button on your keyboard. The first time you do this, gPROMS will give you a list of relevant Units and Variables. Select the desired one and then repeat as necessary to drill down to the desired Variable.
- Specify the lower and upper bounds on the Variable. When gPROMS solves the Experiment Design problem it will ensure that the Variable indicated is restricted to within the specified bounds at the start of each time interval. Alternatively, check the Varying? cell and then specify the lower and upper bounds on the interior point constraint for the start of each interval in the experiment to be designed.
- Repeat the above for each new interior-point constraint you wish to add. If you wish to delete an interior-point constraint, click the *Delete* button at the bottom of the window.

It should be noted that it is possible to supply constraints that make solution of the Experiment Design problem impossible - an *infeasible* problem (e.g. specification of too low a temperature may lead to an infeasible problem if there is only limited cooling available).

**Figure 3.6. Experiment Design - imposing design constraints**

EXPERIMENT TO BE DESIGNED FEDBATCH\_DES (DesignFedbatchb)

End-point constraints

| Variable name | Equality ? | Constraint value | Lower bound | Upper bound |
|---------------|------------|------------------|-------------|-------------|
| <new>         |            |                  |             |             |

Interior-point constraints

| Variable name | Varying ? | Lower bound | Upper bound |
|---------------|-----------|-------------|-------------|
| <new>         |           |             |             |

Delete row

General Controls **Constraints** Measurements and sensors gPROMS language Properties

## Measurements and sensors

During an experiment, measurements of key variables are taken using sensors. The purpose of the Measurement and Sensors tab is to define the variables that will be measured and to specify how often or when the measurements can be taken.

**Figure 3.7. The measurements and sensors tab**

EXPERIMENT TO BE DESIGNED FEDBATCH\_DES (DesignFedbatchb)

Measurements and sensors used

| Measured variable | Sensor | Variance model | Measurement times |
|-------------------|--------|----------------|-------------------|
| FERMENT.y1        | Hplc   | <unspecified>  | Optimised         |
| FERMENT.y2        | Hplc   | <unspecified>  | Optimised         |
| <new>             |        |                |                   |

Optimised measurement times for sensor: Hplc

Earliest measurement time: 0.0

Minimum time between measurements: 0.25

Initial guess: 1.0, 2.0, 3.0, 4.0, 5.0

Delete row

General Controls **Measurements and sensors** Constraints gPROMS language Properties

Each measured variable must be associated with a sensor. In Experiment Design a sensor has a Variance model and a set of measurement times: the measurement times are either fixed or determined optimally.

To introduce a Variable that can be measured:

- Click the <new> cell in the Measured Variable part of the table at the top of the window and type the full pathname of the gPROMS variable. If you do not wish to type the pathname you can select the appropriate

variable from a drop down list by holding the CTRL key down and hitting the SPACE button on your keyboard. The first time you do this, gPROMS will give you a list of relevant Units and Variables. Select the desired one and then repeat as necessary to drill down to the desired Variable.

- You should provide a name for the sensor in the next column or if appropriate associate this Variable with an existing sensor.

For each sensor, you should provide the Variance model and information on when or how the measurement times are to be determined.

- Click on the Variance model cell next to the sensor. When you do so a dialog box will open.
  - Select the desired variance model from a drop down list and
  - enter the appropriate values for the variance model you have selected, click OK to close the dialog box. Note that as gPROMS is not estimating parameters (as in Parameter Estimation) you must provide a fixed value for the uncertainty of the sensor's measurements.
- You must specify how frequently the sensor takes a measurement by selecting from the drop down menu in the adjacent cell.
  - for a fixed number of measurements, select *equidistant* - in the bottom half of the window; enter the number of measurements to be taken and the earliest possible measurement time.
  - to take measurements at fixed times, select *fixed* - in the bottom half of the window; enter the measurement times.
  - to let the optimiser choose the measurement times - in the bottom half of the window; enter initial guesses for the measurement times along with the minimum allowable time between measurements and the earliest possible measurement time .
  - for a steady state experiment you should choose *unspecified*.

*Should you wish to delete a measurement you can do so using the button at the bottom of the window.*

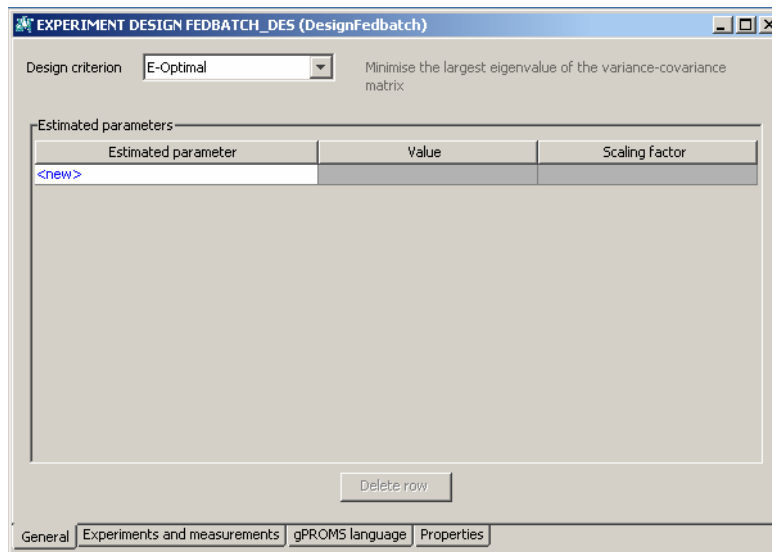
## The Experiment Design entity

This comprises information on the design criterion, the unknown parameters to be estimated and the measurements and sensors (together with their statistical variance models) used in one or more of the experiments.

In order to create a new Experiment Design entity, do the following:

- Pull-down the Entity menu from the top pane in gPROMS ModelBuilder.
- Click on New Entity. A dialog box will appear.
- Choose Experiment Design for the Entity type and fill in the Name field.

When you do the above, an Experiment Design entity editor window will open up in the right hand pane, as shown in the figure below.

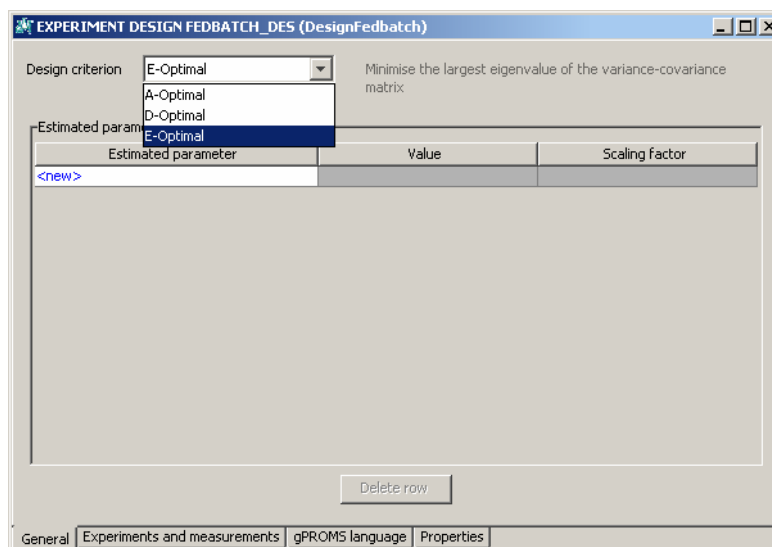
**Figure 3.8. Experiment Design entity editor.**

The Experiment Design entity editor has the following tabs:

- The General tab - a form for defining the Experiment Design objective function as well as the parameters to be estimated by the experiment.
- The Experiments and measurements tab - a form for defining the sets of experiments to be designed and those that have already been performed, together with the measurement variance descriptions that will be used in the estimation.
- The gPROMS language tab - this provides a convenient means of reviewing all the information regarding the experiment to be designed in a text-based format. You are not advised to use this tab to enter data.
- Properties (a record of when the entity was created, last edited as well as the default tab to open on)

## General information

In order to define the form of objective function for the Experiment Design problem, click on the General tab in the Experiment Design entity editor and then choose the desired objective from the drop down menu at the top of the window as shown in the figure below.

**Figure 3.9. Specifying the objective function for the Experiment Design problem.**

In addition to the objective, it is necessary to define the parameters that you are trying to get precise estimates for by designing optimal experiments: these should be added to the Estimated Parameters table along with the current estimate for the Parameter value.

3

- In order to define a new parameter to be estimated, simply click on the <new> cell in the table and type the full pathname to the variable. If you do not wish to type the pathname you can select the appropriate variable from a drop down list by holding the CTRL key down and hitting the SPACE button on your keyboard. The first time you do this, gPROMS will give you a list of relevant Units and Variables. Select the desired one and then repeat as necessary to drill down to the desired Variable.

For certain design criterions, scaling of the parameters can affect the solution obtained: so scaling of the parameter must be considered. **If you are interested equally in all the values of parameters to be estimated then these numbers should be the same.**

A- and E-optimal designs depend strongly on the actual value of the parameters to be estimated. If one parameter is much larger than the rest, the design will most likely try to reduce the variance of this specific parameter. As a side-effect, the variances of the other parameters may be reduced as well, but these are not taken into account directly by the optimisation. In order to make the confidence ellipsoid, given by the variance-covariance matrix  $V_\theta$ , as spherical as possible, it is advisable to scale all parameters to the same value, e.g. unity. On the other hand, if the you want some parameters to be statistically more reliable than others, it is advisable to scale these parameters to a higher value than the rest.

In mathematical terms, we divide each parameter  $\theta_\mu$  by its scaling factor  $scal_\mu$ :

$$\tilde{\theta}_{scal,\mu} = \frac{\theta_\mu}{scal_\mu}, \mu = 1, \dots, N_\theta$$

This is the same as multiplying the variance-covariance matrix with the diagonal matrix  $\text{diag}(scal_1^2, \dots, scal_{N_\theta}^2)$ ,

i.e.

$$V_{\theta_{scal}} = V_\theta \cdot \text{diag}(scal_1^2, \dots, scal_{N_\theta}^2)$$

The D-optimal design is invariant to scaling of the parameters to be estimated.

## Experiments and measurements

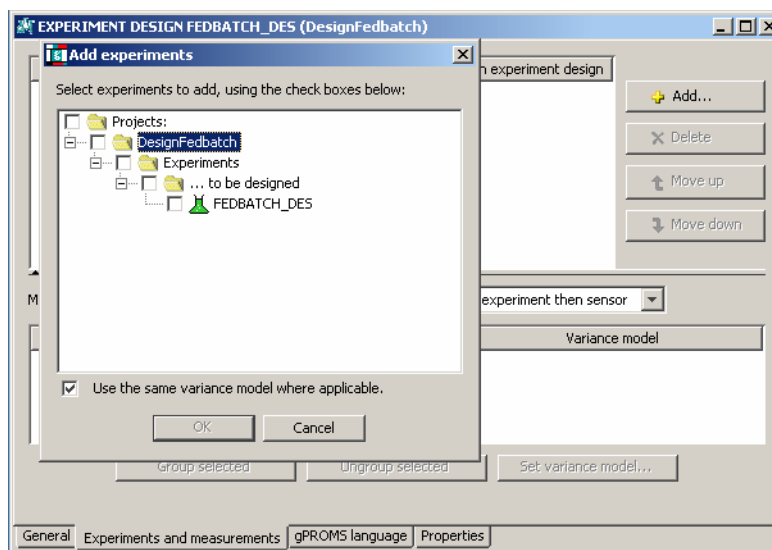
On this tab you select the experiments (both ....*performed* and ...*to be designed*) to be included in the Experiment Design problem. In addition, you have the opportunity to check that the variance models for all the sensors have been fully defined.

To add experiments click on the *Experiments and measurements* tab and then click the *Add* button at the top right of the window. gPROMS will then open a new window that includes a project tree with all the experiment entities that have been defined in all open projects:

---

<sup>3</sup>Note: even though we call it a "parameter" to be estimated, it should be define in the gPROMS model as a variable and ASSIGNED in the Process entity.

**Figure 3.10. Experiments Performed entities that can be included in the estimation problem.**



Check the boxes of the experiments that you wish to include, and click OK.

For a description of Sensors and Sensor groups tables refer to the information given for Performed Experiments.

For the purposes of Experiment Design, gPROMS needs to have a value for the Variance of each sensor:

- *Experiments... to be designed*: when defining the measurement you had the option to specify the variance model and its parameters for each sensor. If you chose not to provide the variance model parameters then you must specify these here.
- *Experiments... performed*: when defining the measurement you had the option to specify the variance model and its parameters for each sensor. You could also choose to estimate the parameters for the sensors; for the purposes of Experiment Design the initial guess for the Variance parameter is used as a fixed value. If you chose not to provide the variance model parameters then you must specify these here.

If any sensors have unspecified variance parameters then you will be warned at the bottom of the *Experiments and measurements* tab. To make it easier to specify variance parameters then you can group sensors and measurements together - note the grouping sensors only links their variance models and NOT the measurement time information provided in the experiment.

## Execution of Experiment Design activities

For a fully defined Experiment Design problem the following entities are required:

- one or more *Model* entities;
- a *Process* entity;
- an *Experiment Design* entity
- one or more *to be designed Experiment* entities

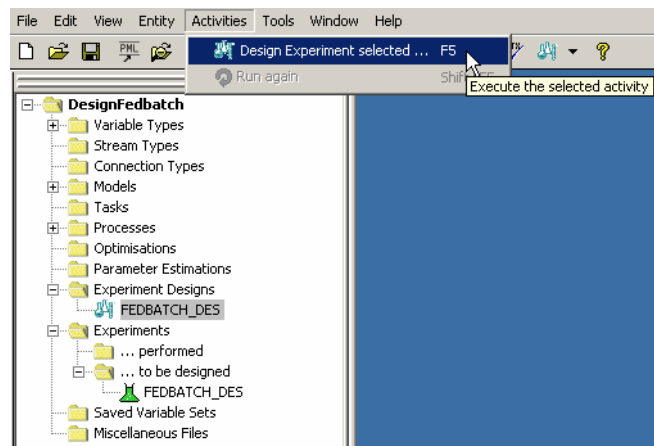
Note that the Experiment entities must all refer to the same Process.

To execute an Experiment Design problem, select the Experiment Design entity from the project tree and then either

1. from the Activities menu and select Design Experiment *name...* , or
2. click on the Design Experiment button at the top right of the tool bar, or
3. press F5.

If there are any syntactical, cross-referencing mistakes etc., these will be detected. Otherwise, an Execution Case is created.

**Figure 3.11. Executing an Experiment Design activity.**



You can analyse the results of an Experiment Design problem in a number of ways:

1. Variable values - stored in the Trajectories folder of the Case
2. Information on the values of the design problem; the status of the optimisation problem and the statistical analysis of the results is provided in three key results files - accessed from the Results folder of the Case:
  - Comprehensive Experiment Design report file (in HTML format): PPP
  - Experiment Design report and statistics files (in plain text format): PPP.out & PPP.stat
  - Experiment schedule files: *EXPERIMENTNAME1.SCHEDULE*, *EXPERIMENTNAME2.SCHEDULE* ...
3. Experiment templates - stored in the Experiment Templates folder of the Case

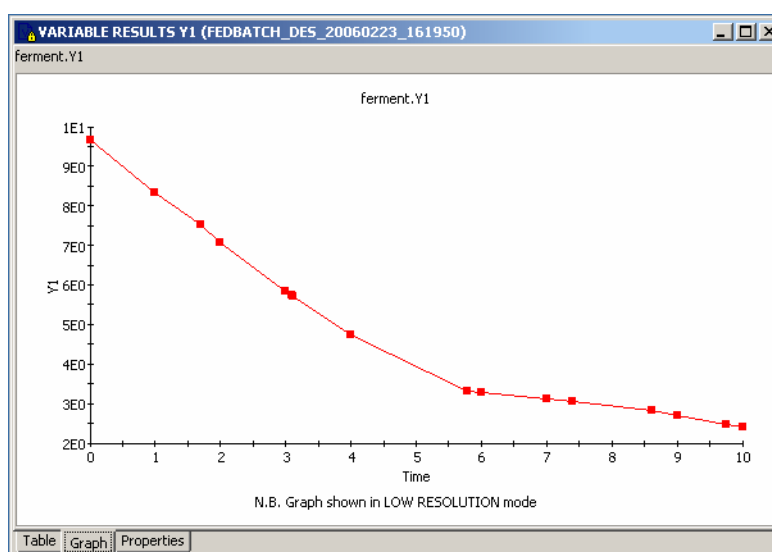
where *PPP* is the name (in capitals) of the Experiment Design entity that has been executed to produce these results.

## Variable Trajectories

In the Execution Case results Trajectories are stored for all of the control and measured variables in each of the experiments - these show the recommended behaviour for all the control variables and the anticipated behaviour (from the gPROMS model) of the measured variables. Note that from the execution dialog you have the option to choose whether to store only the results from the final solution or whether to record them for all of the major steps or even all of the steps during the solution of the Experiment Design problem.

To view the results for a particular variable:-

- open the Trajectories folder in the Case
- select the desired Experiment and navigate to the desired Unit by expanding the results tree.
- double click on the desired variable. This will open up a results window for that variable. There are three tabs, corresponding to tabulated results, graphical results and a properties tab that summarises when the execution was carried out. The figure below shows an example of graphical results for a measured variable.

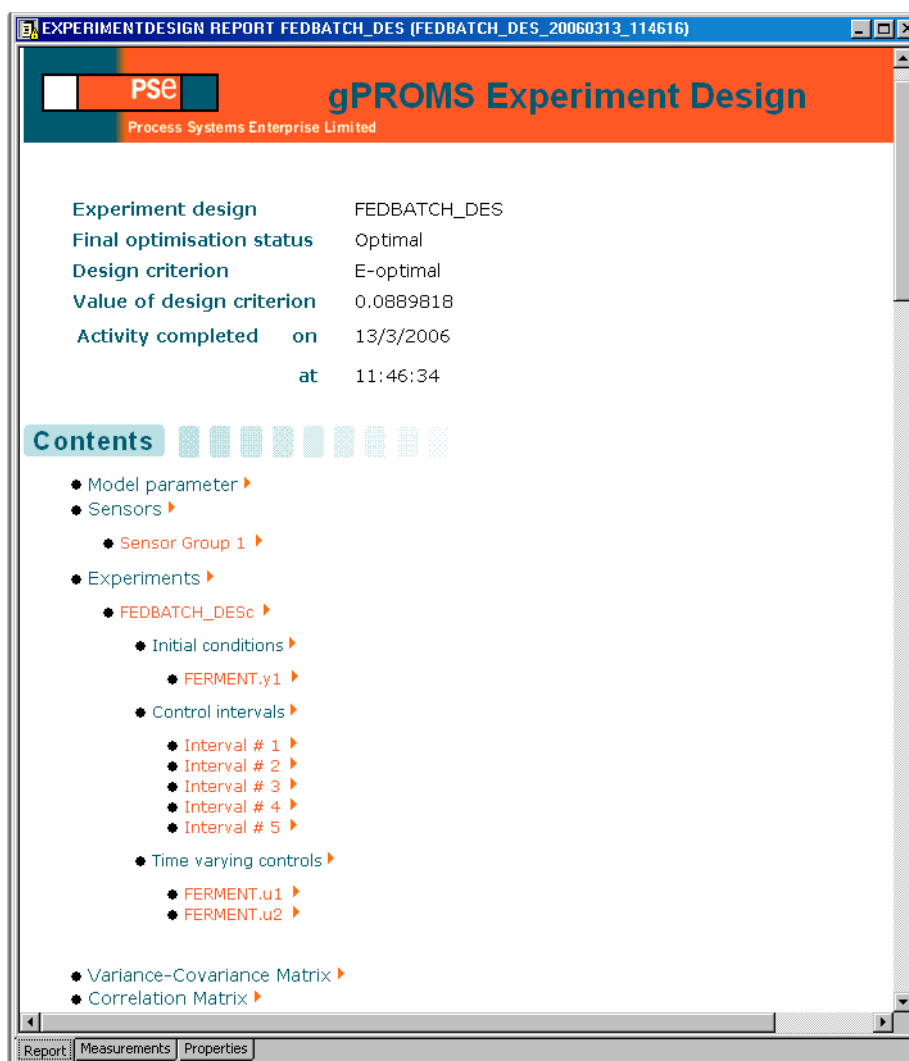
**Figure 3.12. Predicted trajectory of a measured variable in an experiment.**

## The Comprehensive report for Experiment Design

Found in the Results folder, the Experiment Design report file has the same name as the Experiment design activity that was executed. To view the report simply open it by double-clicking.

The figure below shows an example of a comprehensive report.



**Figure 3.13. Comprehensive Experiment Design report.**

The report has three tabs:

- the Report tab
- the Measurements tab
- the Properties tab - this provides details regarding the time of creation.

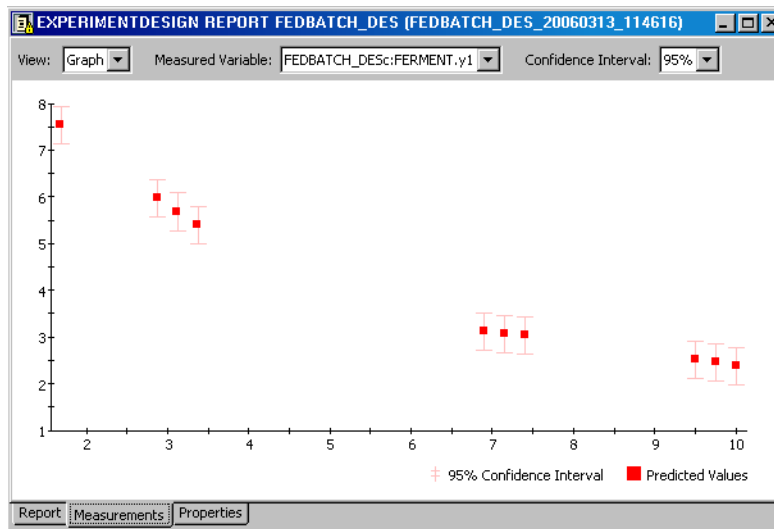
The Report tab includes the following information:

- a table of contents that allows quick access to the information listed below via "hyperlinks";
- general information such as the date and time of the execution of the activity, its final status and the type and value of the design criterion;
- information on the various experiments and experimental decision variables (duration of the experiment, control interval durations, time-invariant and time-varying controls), including the values of:
  - the initial guess used,
  - the final value obtained,
  - the lower and upper bounds,
  - the Lagrange multipliers corresponding to the above bounds,

- information on the interior point and end-point constraints imposed (all active bounds are automatically highlighted).
- the variance/covariance matrix [31];
- the correlation matrix [31];

The *Measurements* tab allows you to view the predicted values for the measured variables from each designed experiment as well as seeing the statistical confidence that is anticipated for these results. You can select the variable to plot and the confidence interval from a drop down menus. You can also choose to view the data for the measured variable in tabular format.

**Figure 3.14. Parameter Estimation measurement plots**



## The Experiment Design text report and statistics files

The .out and .stat plain text files contain a summary report of the Experiment Design run in a simple text format; but of course the information is the same as that presented in the Comprehensive Report file.

The text report (.out) file provides the following information:

- the outcome of the Experiment Design run;
- the design criterion used and the final value of it;
- for each experiment:
  - the final value of the experiment duration and the lengths of the time intervals;
  - the final values of the time-invariant control parameters, and the control-variable profiles; the latter are specified in terms of a single value per interval for piecewise-constant controls, and a pair of values for piecewise-linear controls (as usual, corresponding to the value of the control at the start and end of each interval);
  - the values of variables on which end-point and/or interior point constraints were specified, at the corresponding final and/or interior points;
  - the measured variables and the final values of sampling times and predicted values for the measured data and standard deviation of the measurement error at these points.
- Computational statistics on the performance of the numerical method.

The text statistics (.stat) file contains the results of the statistical analysis of the Experiment Design problem; in particular the following information is available

- the confidence intervals, individual t-values and standard deviations,
- the variance-covariance matrix,
- the correlation matrix of the parameters to be estimated.

## Example Experiment Design text report file

-----  
gPROMS Experiment Design for Parameter Precision  
-----

gPROMS Process : FEDBATCH\_DES1  
Final Optimisation Status : Optimal Solution Found  
Design Criterion : E-optimal  
Value of Design Criterion : 0.089641

| Parameters to Be Estimated: | Nominal Values (fixed) | Scaling factor |
|-----------------------------|------------------------|----------------|
| -----                       | =====                  | =====          |
| FERMENT.theta(1)            | 0.1                    | 0.1            |
| FERMENT.theta(2)            | 0.1                    | 0.1            |
| FERMENT.theta(3)            | 0.1                    | 0.1            |
| FERMENT.theta(4)            | 0.1                    | 0.1            |

Sensors:

-----  
Hplc CONSTANT\_VARIANCE ( 0.2)

Experiment Decision Variables

-----  
([\*] denotes an active bound)

Experiment 1: FEDBATCH\_DES

-----

| Decision Variable | Value | Lower Bound | Upper Bound |
|-------------------|-------|-------------|-------------|
| =====             | ===== | =====       | =====       |
| Duration          | 10    | 5           | 10[*]       |

Initial States:

-----  
FERMENT.y1 9.5985 1 10

Control Intervals:

-----

|                     |        |     |    |
|---------------------|--------|-----|----|
| Control Interval #1 | 1.7312 | 0.5 | 10 |
| Control Interval #2 | 1.444  | 0.5 | 10 |
| Control Interval #3 | 2.6112 | 0.5 | 10 |
| Control Interval #4 | 2.8723 | 0.5 | 10 |
| Control Interval #5 | 1.3412 | 0.5 | 10 |

Time-Varying Controls:

-----

1. FERMENT.u1  
(piecewise constant)  
Control

| Interval<br>===== | Value<br>===== | Lower Bound<br>===== | Upper Bound<br>===== |
|-------------------|----------------|----------------------|----------------------|
| # 1               | 0.05           | 0.05[*]              | 0.2                  |
| # 2               | 0.17201        | 0.05                 | 0.2                  |
| # 3               | 0.2            | 0.05                 | 0.2[*]               |
| # 4               | 0.05           | 0.05[*]              | 0.2                  |
| # 5               | 0.097177       | 0.05                 | 0.2                  |

2. FERMENT.u2  
(piecewise constant)  
Control

| Interval<br>===== | Value<br>===== | Lower Bound<br>===== | Upper Bound<br>===== |
|-------------------|----------------|----------------------|----------------------|
| # 1               | 5.04           | 5                    | 35                   |
| # 2               | 33.694         | 5                    | 35                   |
| # 3               | 33.837         | 5                    | 35                   |
| # 4               | 19.863         | 5                    | 35                   |
| # 5               | 25.529         | 5                    | 35                   |

Measurements:

1. FERMENT.y1 , FERMENT.y2 (OPTIMISED)

Sensor: Hplc

| Measured Variable<br>===== | Time<br>===== | Predicted Value<br>===== | Standard Deviation<br>===== |
|----------------------------|---------------|--------------------------|-----------------------------|
| FERMENT.y1                 | 1.6971        | 7.4787                   | 0.2                         |
| FERMENT.y2                 | 1.6971        | 0.0034771                | 0.2                         |
| FERMENT.y1                 | 3.0063        | 5.8171                   | 0.2                         |
| FERMENT.y2                 | 3.0063        | 0.73037                  | 0.2                         |
| FERMENT.y1                 | 3.2563        | 5.5444                   | 0.2                         |
| FERMENT.y2                 | 3.2563        | 0.95598                  | 0.2                         |
| FERMENT.y1                 | 3.5063        | 5.2634                   | 0.2                         |
| FERMENT.y2                 | 3.5063        | 1.3486                   | 0.2                         |
| FERMENT.y1                 | 7.0885        | 3.093                    | 0.2                         |
| FERMENT.y2                 | 7.0885        | 2.8612                   | 0.2                         |
| FERMENT.y1                 | 7.3385        | 3.0517                   | 0.2                         |
| FERMENT.y2                 | 7.3385        | 2.3375                   | 0.2                         |
| FERMENT.y1                 | 7.5885        | 3.0103                   | 0.2                         |
| FERMENT.y2                 | 7.5885        | 1.8368                   | 0.2                         |
| FERMENT.y1                 | 9.5           | 2.5404                   | 0.2                         |
| FERMENT.y2                 | 9.5           | 0.49742                  | 0.2                         |
| FERMENT.y1                 | 9.75          | 2.4697                   | 0.2                         |
| FERMENT.y2                 | 9.75          | 0.57685                  | 0.2                         |
| FERMENT.y1                 | 10            | 2.402                    | 0.2                         |
| FERMENT.y2                 | 10            | 0.65817                  | 0.2                         |

Computational Statistics

|                                    |          |         |
|------------------------------------|----------|---------|
| Total CPU Time                     | : 55.102 | seconds |
| FEDBATCH_DES1 Optimiser Statistics |          |         |
| CPU Time                           | : 0.02   | seconds |
| Number of NLP Iterations           | : 11     |         |
| Number of NLP Line Search Steps    | : 22     |         |

## DASOLV Integrator Statistics

```

CPU Time : 39.829 seconds
CPU Time Spent on (First Order) Sensitivity Integration Only : 6.65001 seconds
    6013 steps, 9054 residuals : 0.410007 seconds
    7728 Jacobians : 1.69301 seconds
CPU Time Spent on Second Order Sensitivity Integration Only : 33.179 seconds
    2689 steps, 3980 residuals : 0.230001 seconds
    46507 Jacobians : 4.68701 seconds
Mean (1st+2nd Order Sensitivity)/(1st Order Sensitivity) CPU Ratio : 5.98931

```

**Example Experiment Design text statistics file**

```

Parameters to Be Estimated:      Nominal Values (fixed)      Scaling factor
-----
FERMENT.theta(1)                0.1                        0.1
FERMENT.theta(2)                0.1                        0.1
FERMENT.theta(3)                0.1                        0.1
FERMENT.theta(4)                0.1                        0.1
Variance Models:
-----

```

```

Sensor Group 1  CONSTANT_VARIANCE      (      0.2)

```

```

Experiment 1: FEDBATCH_DESc
-----

```

## Measurements:

```

FERMENT.y1      Sensor Group 1
FERMENT.y2      Sensor Group 1
(OPTIMISED)

```

| Measured Variable | Time   | Predicted Value | Standard Deviation |
|-------------------|--------|-----------------|--------------------|
| =====             | ====   | =====           | =====              |
| FERMENT.y1        | 1.6804 | 7.5428          | 0.2                |
| FERMENT.y2        | 1.6804 | 0.0034492       | 0.2                |
| FERMENT.y1        | 2.8671 | 5.9712          | 0.2                |
| FERMENT.y2        | 2.8671 | 0.72544         | 0.2                |
| FERMENT.y1        | 3.1171 | 5.6937          | 0.2                |
| FERMENT.y2        | 3.1171 | 0.9059          | 0.2                |
| FERMENT.y1        | 3.3671 | 5.4046          | 0.2                |
| FERMENT.y2        | 3.3671 | 1.2807          | 0.2                |
| FERMENT.y1        | 6.8971 | 3.1218          | 0.2                |
| FERMENT.y2        | 6.8971 | 3.3295          | 0.2                |
| FERMENT.y1        | 7.1471 | 3.0805          | 0.2                |
| FERMENT.y2        | 7.1471 | 2.7756          | 0.2                |
| FERMENT.y1        | 7.3971 | 3.0393          | 0.2                |
| FERMENT.y2        | 7.3971 | 2.2436          | 0.2                |
| FERMENT.y1        | 9.5    | 2.5315          | 0.2                |
| FERMENT.y2        | 9.5    | 0.46123         | 0.2                |
| FERMENT.y1        | 9.75   | 2.4603          | 0.2                |
| FERMENT.y2        | 9.75   | 0.52776         | 0.2                |
| FERMENT.y1        | 10     | 2.3921          | 0.2                |
| FERMENT.y2        | 10     | 0.59668         | 0.2                |

```

Variance-covariance matrix:
-----

```

```

0.0002385
-1.675E-005    0.0008739
0.0002338    -6.32E-005    0.0002345
0.0001552    -4.015E-005    0.0001496    0.0001105

```

Correlation matrix:

-----

```

      1
-0.03669      1
0.9886    -0.1396      1
0.9558    -0.1292    0.9292      1

```

Information matrix:

-----

```

1.985E+006
-1.07E+005    6937
-1.503E+006    8.133E+004    1.169E+006
-7.926E+005    4.272E+004    5.576E+005    3.827E+005

```

Parameters to Be Estimated:

-----

Nominal Values (fixed)

=====

Standard Deviation

=====

% Stan

=====

|                  |     |          |  |
|------------------|-----|----------|--|
| FERMENT.theta(1) | 0.1 | 0.015444 |  |
| FERMENT.theta(2) | 0.1 | 0.029561 |  |
| FERMENT.theta(3) | 0.1 | 0.015313 |  |
| FERMENT.theta(4) | 0.1 | 0.010513 |  |

## Experiment SCHEDULE files

For each experiment named *EXPERIMENTNAME*, a file *EXPERIMENTNAME.SCHEDULE* is created. This presents the *most recent* Experiment Design solution point in the form of a gPROMS Schedule

The Schedule file can be used to reproduce the detailed results of the Experiment Design optimisation for each experiment by carrying out a simulation activity within gPROMS. In order to do this:

1. Paste the Schedule into the Schedule section of the Process entity.
2. Paste the lines that appear inside the first Reset statement in the Schedule into the Assign section of the Process entity.
3. Paste the lines that appear inside the Initial statement in the Schedule into the Initial section of the Process entity.

Note that it is also possible to *Simulate* a designed Experiment using the Experiment templates.

*Remember:* the contents of the Schedule file do not necessarily represent an optimal or even a feasible solution to the problem: if the optimisation run is interrupted by the user, or ends without finding a satisfactory solution, the file will simply show the point last considered by gPROMS. Only if a comment at the top of the file states the following:

```
# Final Optimisation Status      : Optimal Solution Found
```

should the results be relied upon as a (locally) optimal solution.

## Example Experiment Schedule

```
# Schedule generated from Experiment Design for parameter precision
# for experiment FEDBATCH_DES
#
# Final Optimisation Status      : Optimal Solution Found
# Design Criterion               : E-optimal
# Value of Design Criterion      : 0.0896408
#
INITIAL
  FERMENT.Y1 = 9.59846;
  FERMENT.Y2 = 0;

SCHEDULE

SEQUENCE

  RESET
    FERMENT.u1 := 0.05;
    FERMENT.u2 := 5.04004;
  END

  CONTINUE FOR 1.73124

  RESET
    FERMENT.u1 := 0.172015;
    FERMENT.u2 := 33.6937;
  END

  CONTINUE FOR 1.44401

  RESET
    FERMENT.u1 := 0.2;
    FERMENT.u2 := 33.837;
  END

  CONTINUE FOR 2.61125

  RESET
    FERMENT.u1 := 0.05;
    FERMENT.u2 := 19.8626;
  END

  CONTINUE FOR 2.87231

  RESET
    FERMENT.u1 := 0.0971769;
    FERMENT.u2 := 25.5291;
  END

  CONTINUE FOR 1.34119

END
```

## Designed Experiment templates

After a successful Experiment Design activity; for each designed experiment an Experiment template is generated. This presents the results of the Experiment Design in a Performed Experiment entity: the calculated values for the controlled variables and the initial conditions are entered on the appropriate tabs and the sensor information (variance models and measurement times) is also preserved. Note that the measured data is not available because, of course, the experiment has yet to be carried out.

To use the Experiment template you should copy them into your working project.

The experiment templates satisfy two purposes:

- To facilitate the iterative Model Validation cycle
- To enable easy simulation of the designed experiment using the Simulate Experiment capability

## Controlling the execution of Experiment Design activities

It is possible to fine tune the numerical parameters associated with the underlying solvers used by gPROMS in the solution of an Experiment Design problem. This is done in the *SolutionParameters* section of the Process entity.

The specification for the default solver is shown below:

```
SOLUTIONPARAMETERS
  EDSolver := "EXPDES" ;
```

Obviously, a specification in the simple form shown above is redundant since EXPDES is already the default solver. However, the solver has a number of parameters that can be used for configuring its precise behaviour for any particular problem being solved, as shown in the example below:

```
SOLUTIONPARAMETERS
  EDSolver := "EXPDES" [ "OutputLevel"           := 0 ;
                        "MINLPSolver"           := "SRQPD" ;
                        "DASolver"              := "DASOLV" ] ;
```

- `Outputlevel`: An integer in the range [0, 3].

The amount of information generated by the solver. The following table indicates the lowest level at which different types of information are produced:

**Table 3.2. Outputlevel**

|   |   |
|---|---|
| 0 | Warning and error messages,<br>best available point after failure;                                    |
| 2 | Message "Evaluating contribution<br>of experiment <i>name</i> ";                                      |
| 3 | Variance-covariance matrix of the parameters to<br>be estimated (unscaled) in each optimisation step. |

- `MINLPSolver`: A quoted string specifying a nonlinear programming optimisation solver.

The optimisation solver to be used for solving the Experiment Design optimisation problem. This can be either the standard SRQPD solver or a third-party nonlinear programming solver. The default is SRQPD. This parameter can be followed by further specifications aimed at configuring the particular solver by setting values to



its own algorithmic parameters. Note that the solution parameter Scaling of the SRQPD solver defines different scaling methods of the Hessian of the SQP method (which can be seen as a scaling of the experiment decision variables  $\xi$ ) whereas the SCALING\_FACTOR of the parameters to be estimated in the Experiment Design entity defines the scaling of the parameters to be estimated.

- **DASolver**: A quoted string specifying a differential-algebraic equation solver.

The solver to be used for integrating the model equations and calculating their sensitivity equations at each iteration of the optimisation. This can be either the standard DASOLV solver (see Model Developer Guide) or a third-party differential-algebraic equation solver. The default is DASOLV. This parameter can be followed by further specifications aimed at configuring the particular solver by setting values to its own algorithmic parameters.

Experiment Design for parameter precision requires the calculation of first and second order sensitivities of the underlying differential-algebraic equation system. Second order sensitivities are calculated via so-called varied trajectories (or finite differences). This is similar to a finite difference calculation. The difference is that the algorithm uses the same discretisation for both the nominal solution and first and second order sensitivities, exploiting structure in the linear algebra calculations. Moreover, the accuracy for the first and second order sensitivities is retained at the same order of magnitude as the accuracy for the nominal solution.

The following algorithmic parameters used by DASOLV to control the calculation of second order sensitivities are shown below. This is followed by a detailed description of each parameter.

```
"DASOLV" [ "AbsolutePerturbationFactor" := 1e-07;  
           "RelativePerturbationFactor" := 1e-04;  
           "FiniteDifferences"          := FALSE];
```

- **AbsolutePerturbationFactor**: A real number in the range  $[10^{-20}, 10^{10}]$ .

The absolute perturbation factor of the optimisation decision variables for calculating varied trajectories for second order sensitivity calculation.

- **RelativePerturbationFactor**: A real number in the range  $[10^{-20}, 10^{10}]$ .

The relative perturbation factor of the optimisation decision variables for calculating varied trajectories for second order sensitivity calculation.

- **FiniteDifferences**: A boolean value.

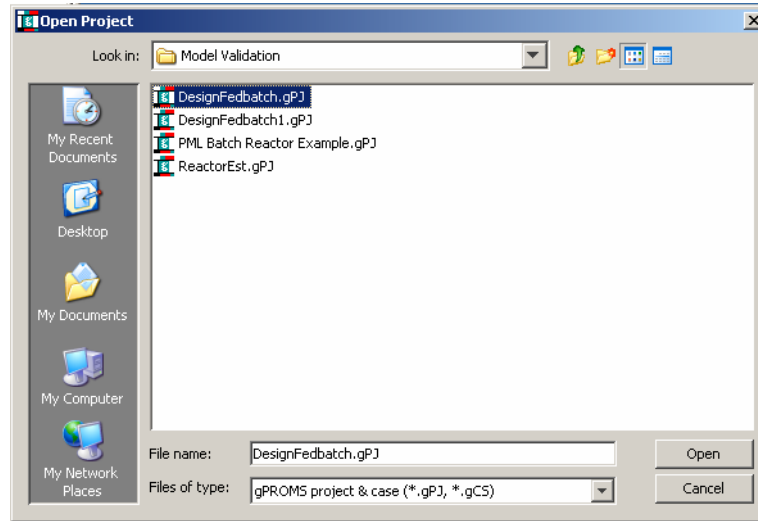
Second order sensitivities are by default calculated via so-called varied trajectories (see above). If this parameter is set to TRUE, second order sensitivities are calculated via ordinary finite differences on the first order sensitivities.

These solution parameters can be set in addition to those stated for the DASOLV solver (see Model Developer Guide).

## An Experiment Design example

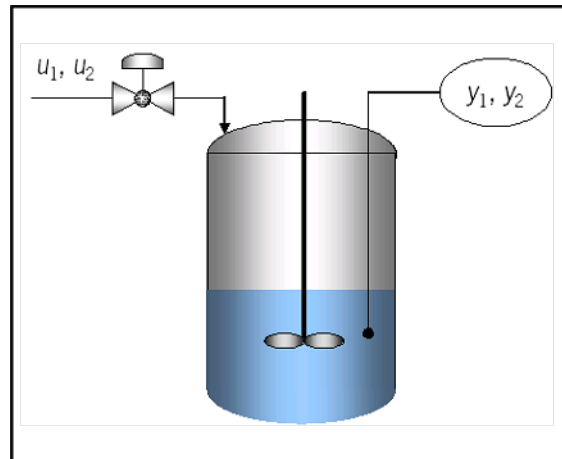
The gPROMS installation includes a simple Experiment Design example - it is highly recommended that you take a look at this. You can access this by clicking on the browse examples button on the gPROMS Tool bar and then navigating to General capabilities\Model Validation\DesignFedbatch.gPJ (as shown below). This section explains the features of the model.

**Figure 3.15. Browsing for the Experiment Design example in the gPROMS distribution directory.**



The example considers the fed-batch fermentation process shown in the figure below.

**Figure 3.16. Fed-batch fermentation**



The mathematical model for the process is as follows:

- *Biomass concentration*

$$\frac{dy_1}{dt} = (r_m - u_1) y_1 - \theta_4 y_1$$

- *Substrate concentration*

$$\frac{dy_2}{dt} = -\frac{r_m y_1}{\theta_3} + u_1 (u_2 - y_2)$$

- *Reaction rate*

$$r_m = \frac{\theta_1 y_2}{\theta_2 + y_2}$$

The table below summarises the parameters to be estimated and the design variables that appear in this model.

**Table 3.3. Parameters and design variables in fed-batch fermentation model.**

| Model Variables           |                              |
|---------------------------|------------------------------|
| $\theta_i, i = 1 \dots 4$ | Parameters to be estimated   |
| $u_2$                     | Dilution factor              |
| $u_2$                     | Feed substrate concentration |
| $y_1$                     | Biomass concentration        |
| $y_2$                     | Substrate concentration      |
| $r_m$                     | Reaction rate                |

Before this model can be used to simulate or optimise the fermentation process, all parameters that appear in it must be given values. Imagine, however, a situation where the kinetic parameters are unknown. That is, we do not know the values of  $\theta_i, i = 1 \dots 4$ . To determine these, we can perform a number of experiments, measure the values of some or all process variables, and estimate the kinetic parameters that best fit the measured data (see Parameter Estimation in gPROMS for a more detailed description).

In the example, it is assumed that we design one new experiment, take up to 10 samples at different times and measure the biomass and substrate concentration,  $y_1$  and  $y_2$ , in each sample using a High Pressure Liquid Chromatograph (HPLC). The duration of the experiment should be between 5 and 10 hours. The dilution factor  $u_1$  and the feed substrate concentration  $u_2$  may be varied over time. We may also change the initial value of the biomass concentration  $y_1(0)$  in the reactor.

Resource availability may impose an upper limit on the feed substrate concentration.

The Experiment Design problem seeks to provide answers to the following questions

- What initial biomass concentration  $y_1(0)$  should we use?
- How long should we run the experiment for?
- How should we vary the feed flowrate for the dilution factor  $u_1(t)$  and the feed substrate concentration  $u_2(t)$  over the duration of the experiment?
- When should we take the measurement samples?

The unknown parameters to be estimated are  $\theta_1$  (gPROMS pathname FERMENT.theta(1)) to  $\theta_4$  (gPROMS pathname FERMENT. $\theta(4)$ ). We use the estimated value of each of the parameters as the SCALING\_FACTOR so that the scaled values of all parameters are one.

An HPLC is used to measure both  $y_1$  and  $y_2$ . The variance model is described as CONSTANT\_VARIANCE with  $\omega$  equal to 0.5.

The table below contains the above information for the fed-batch fermentation experiment.

**Table 3.4. Conditions for the fed-batch fermentation experiment**

| Type of Decision Variable                 | Variable | Initial Guess | Lower Bound | Upper Bound |
|---|----------|---------------|-------------|-------------|
| Experiment Duration                       | $t_f$    | 10.0          | 5.0         | 10.0        |
| Initial Condition                         | $y_1(0)$ | 7.0           | 1.0         | 10.0        |
| Time-varying control (piecewise-constant) | $u_1(t)$ | 0.2           | 0.05        | 0.2         |
| Time-varying control (piecewise-constant) | $u_2(t)$ | 15.0          | 5.0         | 35.0        |

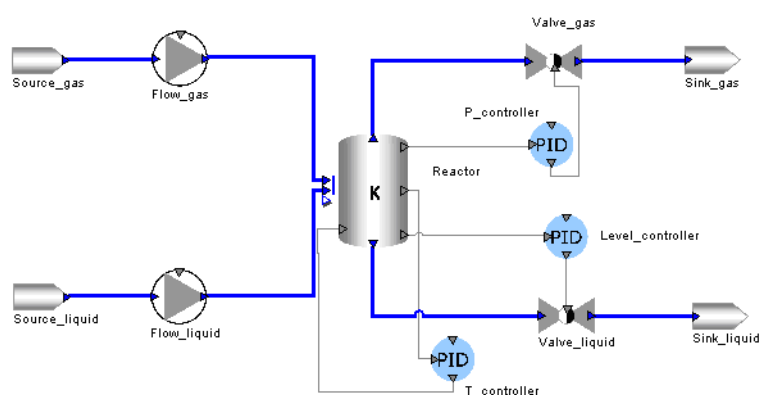
# Chapter 4. A Model Validation Example

In this example, the estimation of the kinetic parameters of a phenol hydro-deoxidation reaction is considered.

**Note:** You should load the Process Model Libraries and the PML Batch Reactor Example project file (from the examples/PML Flowsheets directory of your gPROMS installation)

The following figure shows the flowsheet developed using Process Model Library (PML) to represent a batch reactor. The reactor unit is an instance of the Reactor\_drum\_kinetic PML model. The influence of the vapour-liquid equilibrium is taken into account. The chemical reaction is assumed to happen in the liquid phase only.

**Figure 4.1. Batch Reactor Flowsheet**



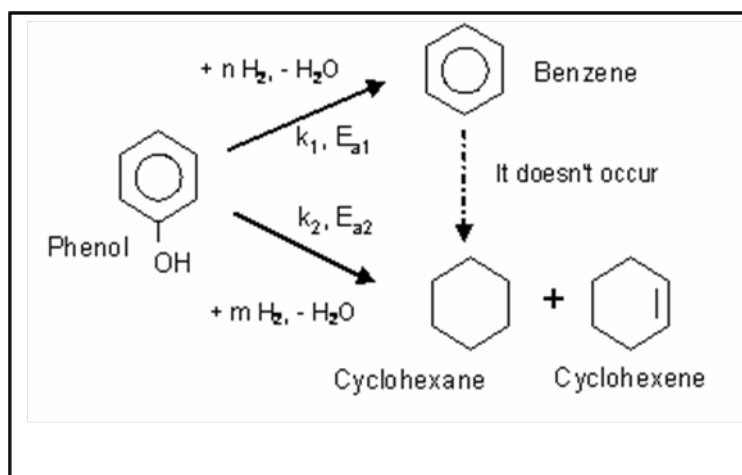
The table below shows the components used in the simulation and their corresponding array numbers in the gPROMS model:

**Table 4.1. Components of Phenol hydro-deoxidation**

| Array Number | Component   |
|--------------|-------------|
| 1            | hydrogen    |
| 2            | phenol      |
| 3            | benzene     |
| 4            | cyclohexane |
| 5            | water       |
| 6            | hexadecane  |

## Reaction kinetics

Phenol hydro-deoxidation over a sulphided CoMo/Al<sub>2</sub>O<sub>3</sub> catalyst follows the following reaction network:

**Figure 4.2. Reaction Network**

The kinetics of these reactions has been described as pseudo-first order in the liquid phase phenol concentration:

$$R_i = K_{0,i} \exp \{-E_a/RT\} C_i^{liquid} \quad i = 1, \dots, N_C$$

## Experiment description

### Reactor

A 50 ml stainless steel autoclave is used for the experiments. The autoclave is equipped with a spinning catalyst basket to avoid heat and mass transfer limitations between the catalyst particles and the liquid with the reactants.

### Experiment procedure

Enough quantity of pre-sulphide catalyst is packed into the catalyst basket. The reactor is initially filled with 15 ml of hexadecane ( $770.676 \text{ kg/m}^3$ ) and 1 g of phenol. The reactor is then sealed and, once the reaction temperature is reached, pure hydrogen is added until a pressure of 70.6 bar is reached. The degree of the reaction achieved during the temperature and pressure stabilisation is considered negligible with respect to the total degree of reaction at the end of the experiments. Five samples are withdrawn from the reactor liquid phase at minimum intervals of 6 min. The first sample can be taken after 3 min of the experiment start. The amount of liquid extracted during the sampling is negligible related to the total amount of liquid in the reactor. No hydrogen is added during the runs to compensate the pressure drop caused by the reaction and sampling.

The quantitative analysis of the liquid sample is performed by means of an on-line gas chromatography system. The concentration measurements have a relative error with a standard deviation of 0.02.

### Experiments

Two experiments were performed with similar initial conditions (see table below):

**Table 4.2. Initial conditions of experiments**

|                           |                 |
|---------------------------|-----------------|
| Phenol initial mass       | 1 g             |
| Hexadecane initial volume | 15 ml (11.56 g) |
| Pressure                  | 70.6 bar        |
| Product initial masses    | 0               |

The experiments were performed at different temperatures, 523.15 and 533.15 K, respectively. The measurements of hydrogen, phenol, and benzene are specified in the tables below:

**Table 4.3. Measurements for experiment 1 (T = 523.15 K)**

| Time (s) | C <sub>1</sub> (mol/m <sup>3</sup> ) | C <sub>2</sub> (mol/m <sup>3</sup> ) | C <sub>3</sub> (mol/m <sup>3</sup> ) |
|----------|--------------------------------------|--------------------------------------|--------------------------------------|
| 360.0    | 570.6                                | 2.4                                  | 1.7                                  |
| 1190.0   | 560.1                                | 8.0                                  | 5.5                                  |
| 2410.0   | 561.6                                | 15.9                                 | 11.6                                 |
| 3600.0   | 529.9                                | 22.4                                 | 16.7                                 |
| 4790.0   | 534.6                                | 30.9                                 | 22.3                                 |

**Table 4.4. Measurements for experiments 2 (T = 533.15 K)**

| Time (s) | C <sub>1</sub> (mol/m <sup>3</sup> ) | C <sub>2</sub> (mol/m <sup>3</sup> ) | C <sub>3</sub> (mol/m <sup>3</sup> ) |
|----------|--------------------------------------|--------------------------------------|--------------------------------------|
| 360.0    | 564.1                                | 4.1                                  | 2.4                                  |
| 1190.0   | 553.1                                | 13.1                                 | 7.8                                  |
| 2410.0   | 523.9                                | 25.4                                 | 15.0                                 |
| 3600.0   | 501.6                                | 36.8                                 | 22.7                                 |
| 4790.0   | 481.4                                | 49.4                                 | 29.3                                 |

## Parameter Estimation

In order to perform the Parameter Estimation, you need to do the following:

1. Load the Process Model Library and Batch reactor project.
2. Inspect the Batch *Estimation* entity; observe where the estimated parameters and sensor groups are identified.
3. Inspect the EXP\_1 and EXP\_2 *Performed Experiment* entities; observe how the initial conditions, control variables and measure data are defined.
4. Select the Batch *Estimation* Entity and execute the Estimation activity
5. After the activity has executed: open the Comprehensive report - this is found in the Results directory of the Execution Case. Have a look at the statistical analysis of the estimated parameters (e.g. confidence intervals, individual t-value, correlation matrix) and the lack-of-fit test.

## Experiment Design

In order to increase the precision/certainty of the parameters to be estimated, another experiment is to be designed. To do so, do the following:

1. Inspect the Batch *Experiment Design* Entity. Note that:
  - The E-optimal design criterion is used to design the new experiment.
  - The values determined from the Parameter Estimation runs are used
  - The variance model for the Gas Chromatography sensor is specified.
  - All the information from the already performed experiments is taken into account.
2. Inspect the EXP\_TBD *to be designed Experiment* entity; compare the specification of the initial conditions, control variables and measure data with the information provided for the *Performed Experiments*.
3. Select the Batch *Experiment Design* entity and execute the Experiment Design activity

4. After the activity has executed: open the Comprehensive report - this is found in the Results directory of the Execution Case. Have a look at the values of the experiment decision variables. Compare the statistical analysis with the one from the Parameter Estimation activity. Determine if the precision/certainty of the parameters to be estimated has increased.