Robust Calibration of Complex ThermoSysPro Models using Data Assimilation Techniques: Application on the Secondary Loop of a Pressurized Water Reactor

Luis Corona Mesa-Moles[1] Jean-Philippe Argaud[1] Audrey Jardin[1] Amine Benssy[1] Yulu Dong[1]

[1]EDF R&D, France, {luis.corona-mesa-moles, jean-philippe.argaud, audrey.jardin, amine.benssy, yulu.dong}@edf.fr

Abstract

ThermoSysPro (TSP) is a library for the modeling and simulation of power plants and energy systems. It has been developed by EDF and it is released under open source license. When developing models with TSP it is necessary to ensure that they match reality. In practice, this operation is performed by adjusting the value of the parameters appearing in the model. This major step corresponds to model calibration.

Calibration can be performed through various methods. A classical way to do so with Modelica models is by model inversion. The major inconvenience of this method, in addition of potential convergence problems for complex models, is that it is necessary to have exactly the same number of measurements as parameters to be calibrated, which is not often the case in practice.

This paper shows how data assimilation techniques can robustly be used for calibration of complex TSP models avoiding the inconveniences associated to calibration by model inversion while ensuring an optimal use of the available measurements. A complex TSP model of the secondary loop of a Pressurized Water Reactor (PWR) is considered for this purpose.

Keywords: Modelica, ThermoSysPro, ADAO, data assimilation, model calibration, thermal-hydraulics, pressurized water reactor.

1 Introduction and context

Physical models of energy systems such as power plants can be advantageously used for the engineering of these systems all along their lifecycle from the design phase till the operation phase. They can be employed to test different design or retrofit alternatives, to evaluate the impact of changes in safety or environmental rules, to validate the performance of new components during their commissioning, to train operators, or even to help diagnose component's failures or sensor's drifts during operation and predict the system evolution in these conditions.

Modelica (Modelica, 2018) is a language perfectly suited for this kind of modelling thanks to its equation-based and acausal features:

(1) The engineer can use physical equations to capture in the same model the different phenomena governing the system behavior from the mechanical, hydraulic, thermal, electrical, and so on points of view;

(2) The equations are expressed in an acausal (i.e. non-oriented) way such that the engineer can reuse the same model for different computation purposes. From the same equation, one may, for instance, deduce the perfect sizing of a component to match a given operating point or compute the resulting operating point given the characteristics of on-shelf component.

A generic Modelica library, called ThermoSysPro (TSP), has been developed by EDF to model and simulate power plants and other kinds of energy systems. It is released under open source license and freely distributed with the OpenModelica simulation tool (OpenModelica, 2018) downloadable here: https://openmodelica.org/download/download-windows#.

Numerous organizations and individuals worldwide now use TSP and a large spectrum of use-cases exist from nuclear, thermal, to combined-cycle through biomass or even concentrated solar plants (El Hefni B. and Bouskela D., 2017).

In the design phase, the engineer has no other choice than calibrating such models with design assumptions and theoretical performances of each component issued from manufacturer data.

In the operation phase, when measurements within the modelled system are available, it is possible to use them to calibrate the model. One way to perform calibration is by model inversion which consists in computing the values of n parameters that deterministically correspond to a given set of n measurements. Model inversion can be performed using the Modelica feature to express inverse problems. This method gives satisfying results but it can be difficult to implement in practice for complex models. The main drawbacks associated to this method are:

- the necessity to readapt some part of the model in order to express the inverse problem; in some cases it may require to develop new modules facilitating the convergence of the inverse model;
- the necessity to consider exactly the same number of measurements as of parameters to calibrate (which does not happen often);
- the fact that the different measurements are considered homogenously (even if they have not been obtained in the same conditions, or if the model is not intended to be representative of all the available measurements in the same way);
- no consideration of measurement uncertainties (including on the boundary conditions of the model).

Data assimilation framework (Asch M. et al., 2016; Bouttier B. et al., 1999; Kalnay E. et al., 2003) provides a number of alternative methods and techniques that can be used to overcome these difficulties during model calibration.

For illustration purposes and to better understand the main differences between the two approaches, hereafter is presented a calibration problem of a simple TSP model, for more details see *Modeling and simulation of a complex ThermoSysPro model with OpenModelica* (El Hefni B. and Bouskela D., 2017). The model is presented in Figure 1. It corresponds to a singular pressure loss module with given boundary conditions (in this case the inlet and outlet pressures).

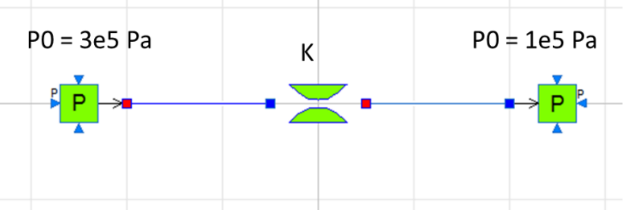

Figure 1. Model of singular pressure loss

The calibration of the model consists in determining the value of the pressure loss coefficient (K) of the pressure loss module. The measurement available to perform this calibration corresponds to the mass flow rate through the pressure loss module (Q).

For the calibration by model inversion, the observed mass flow rate is directly used to compute the exact value of the pressure loss coefficient since both appear in the same physical equation.

The corresponding physical equation is presented below:

$$P_i - P_o = K \cdot \frac{Q \cdot |Q|}{\rho}$$

P_i and P_o are the fluid pressure at the inlet and at the outlet of the singular pressure loss respectively, ρ is the average density of the fluid, Q is the mass flow rate and K is the friction pressure loss coefficient.

In the calibration using data assimilation techniques, the approach is different. From the physical knowledge of the system it is possible to give a guess value to the K coefficient (or use directly the default value of the TSP library), this corresponds to the *a priori* value of the parameter to be calibrated. This *a priori* value is used as a starting point and will be iteratively corrected to find the best value of the calibrated parameter, "best" in the sense that the results given by the model should be in the end as close as possible to the available measurements.

The objective of the article is to show how data assimilation techniques can be used in general to have a more robust approach of the calibration phase.

It illustrates on an industrial-size use-case, which is the model of the secondary loop of a pressurized water reactor, what are the concrete benefits of this approach compared to the traditional one in place using model inversion.

2 Model of the secondary loop of a PWR

2.1 Nuclear power plant performance monitoring

The secondary loop of a 1300 MW PWR nuclear power plant has been modelled with TSP modules in order to determine the best efficiency rate that can be expected from the thermo-hydraulic cycle, given various boundary conditions. This theoretical best efficiency operation setpoint gives an estimation of several physical quantities like pressures and temperatures across the cycle. They are the references against which the on-site measurements will be compared, allowing to identify any deviation causing energy losses. These symptoms will then be processed in order to identify their potential causes.

The more accurate the model is, the better the diagnosis will be.

2.2 Model description

Secondary loops of PWRs are classical Rankine cycles that convert thermal energy into electrical power.

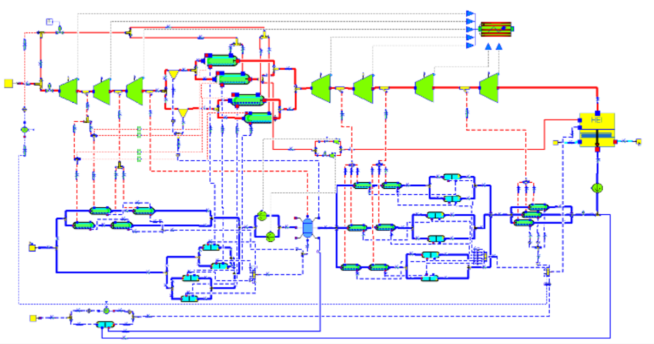

Figure 2. Model of 1300 MW PWR secondary loop with TSP

The TSP model developed to represent such PWR's secondary loop is static and composed of the following key systems (Figure 2):

- a turbogenerator set made of high-pressure (HP) and low-pressure (LP) turbines and one generator;
- two sets of Moisture Separator Reheaters;
- one condenser;
- one feedwater tank and gas stripper system;
- two turbine-driven feedwater pumps;
- low (LP) and high pressure (HP) feedwater heaters.

Once properly calibrated, the model calculates the nominal operation setpoint from thirteen boundary conditions. Among which the more important are: plant's cooling water temperature and pressure, Steam Generator's (SG) thermal power, SG's moisture carryover level, SG's pressure at the outlet, SG's feedwater flow.

3 Calibration methodology

3.1 Data assimilation framework

Data assimilation is a general well established framework (Asch M. et al., 2016) for computing the optimal estimate of the true state of a system, over time if necessary. It combines knowledge between observations and *a priori* models, including information about their errors. The goal is to obtain the best possible estimate of the system real state and of its stochastic properties. Moreover, data assimilation provides deterministic techniques in order to perform very efficiently the estimation job. Because data assimilation looks for the best possible estimate, its underlying procedure always integrates optimization in order to find this estimate.

The calibration of a model consists in looking for the value (of part) of the parameters of a model, in such a way that the simulation obtained with these parameters is better adapted to real measurements on the same simulated system, in the sense that the distance between model predictions and measurements is smaller. The use of data assimilation for calibration requires the acquisition of measured information in the same conditions under which the simulated system is to be calibrated. The collection and prior analysis of these measurements also establishes elements of confidence and compared quality of the measurements, which will be interesting in the use of algorithms. In addition, the numerical model used must be functional over a validity domain that includes the range of variation of the parameters to be calibrated.

All quantities representing the description of physics in a model are likely to be calibrated in a data assimilation process, whether they are model parameters, initial conditions or boundary conditions. Their simultaneous consideration is greatly facilitated by the data assimilation framework, which makes it possible to objectively process a heterogeneous set of available information.

3.2 Data assimilation applied to 0D/1D models with ADAO

To perform data assimilation, a specialized LGPL free distributed tool ADAO (Salome, 2018) is used to simplify the application of data assimilation for the simulation of complex systems. Available in the Python environment that allows the simulation of Modelica models and hence of TSP models, it allows to easily automatize the calibration of 0D/1D models and the development of complex calibration scenarios according to the states of the analyzed physical system. ADAO was initially developed to perform data assimilation with 2D/3D models. Its adaptation to 0D/1D models has been coded during this work and now simplifies the specification of model parameters to be calibrated and simulated quantities to be compared to measurements, which are known in the Modelica description of the system. In addition, an advanced and simultaneous management of the various possible operating conditions enhances the physical representativeness of the overall calibration of the simulated system.

The use of ADAO in a Modelica/Python environment allows to simply describe the data assimilation problem, through a Modelica representation of the simulation and of the named data for measurements as well as for the *a priori* values of the parameters to be calibrated. Since the different measurements are not obtained with the same sensors, the confidence accorded to the different available measurements can be easily modified as well. Moreover, when calibrating a large number of parameters, a sensitivity analysis can be performed to reduce the set of parameters that should be calibrated to the only ones that have a real impact on the quantities observed through the measurements. The entire data assimilation process is then automated and depends only on the ability of the model to simulate the system for the required parameter values through optimization. The stability and convergence of the simulated system over its entire domain of validity are therefore essential to allow an efficient search for a set of calibration parameters. The availability of complete or aggregated outputs provided by ADAO for a simulation ensemble is also crucial to ensure that the optimal simulation can be analyzed in detail and that the calibrated parameters are relevant.

3.3 Calibration procedure

The specialized tool ADAO allows to easily define the different elements necessary to perform model calibration using data assimilation techniques. These elements are:

- Parameters (including initial conditions or boundary conditions if required) to be calibrated (with given *a priori* values);
- Available measurements (taking into account whether they have been obtained under the same conditions, i.e. with the same boundary conditions, or not);
- Modelica model (describing the physical connection existing between the parameters to be calibrated and the observations/measurements).

It is important to note that a variable confidence error can individually be associated to the different measurements available, under the form of a covariance matrix. This information is then used by ADAO to compute the optimal values of the parameters. This process is illustrated in Figure 3 (in blue the necessary information to be provided to ADAO).

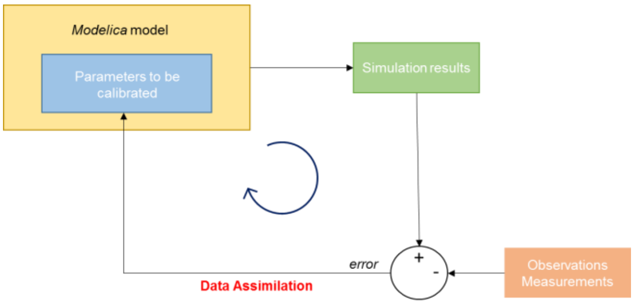

Figure 3. Illustration of the calibration procedure using data assimilation techniques

3.4 Analysis criteria

In order to evaluate how good a calibration of the model is, it is necessary to establish a certain number of criteria. In a model calibration procedure as here, the objective is that the variables computed by the model are as close as possible to the available measurements. Therefore, these indicators should be based on the differences between the available on-site measurements and the corresponding variables computed by the model.

In this paper, two different criteria are considered. Firstly, a global indicator that it is equal to the sum of the quadratic difference between the measurement and the corresponding variable in the model, for all the available measurements. This indicator is not very different from the cost function value minimized by ADAO in the calibration procedure. It provides a general overview of a given calibration. Secondly, an indicator for each measurement, considered individually, is necessary in order to detect a calibration that is unacceptable (i.e. outside of its *a priori* confidence interval) for a given measurement while being correct globally. For each measurement, the relative difference between this measurement and the corresponding model output is computed. The relative difference is defined as the absolute value of the ratio between the difference between the measurement and the model output and the value of the measurement. It provides therefore a homogenous indicator for all the available measurements. This indicator can as well be used to check precisely if the calibrated model is more representative for certain measurements, judged to be more important than the others.

4 Calibration of the secondary loop of a PWR

4.1 Scenario description

The calibration of the TSP model of PWR's secondary loop (Figure 2) is performed over 116 parameters. For comparison purposes, the same observations as for the calibration by model inversion are considered. They correspond to pressure, mass flowrate and temperature measurements. These observations correspond to 116 measures obtained in ten campaigns of measurements. Therefore, in total 10x116=1160 observations are used to perform the model calibration. However, contrary to the calibration by inversion, the observations are considered simultaneously for the calibration presented in this study. The use of ADAO and preprocessing facilities allows to adapt the boundary conditions for each set of 116 observations.

The calibration of the model is performed for two different configurations, which mainly differ from the *a priori* values given to the parameters as an initial guess. In the first configuration the *a priori* values for the parameters to be calibrated correspond to the values obtained by model inversion. In the second one, typical *a priori* values are considered, corresponding to what can be found in technical data sheets of the modeled components. This second case would correspond to a typical calibration procedure while the first one shows how data assimilation methods can improve the calibration obtained by the classical model inversion method which is now in current use in our engineering divisions.

For the first configuration, the following sub-scenarios are studied:

- High confidence on observations (scenario 1);
- High confidence on observations but according more confidence to some of them that are considered as more meaningful (scenario 2).

For the second configuration, the following sub-scenarios are considered:

- High confidence on observations (scenario 3);
- High confidence on observations but considering a reduced number of parameters to calibrate (the selection of these parameters is performed through a sensitivity analysis, 62 parameters are kept) (scenario 4).

For each scenario, the domain in which the optimal value of the parameters is searched is adjusted in order to ensure the convergence of the simulated model (see

3.2 for more details). These research domains are indicated in Table 1.

Table 1. Research domain for the optimal value of calibrated parameters.

Scenario	*Research domain*
Scenario 1	5% around *a priori* values
Scenario 2	5% around *a priori* values
Scenario 3	10% around *a priori* values
Scenario 4	60% around *a priori* values

It appears clearly that reducing the number of parameters to calibrate enables to enlarge the research domain for the optimal value of the parameters: the convergence of the model is facilitated compared to the situation in which all the parameters have to be calibrated and may vary.

4.2 Results and discussion

First of all, it is important to examine the optimal value of the parameters given by the data assimilation procedure. A key point is to check if the optimal value of the calibrated parameter reaches the bounds of the research domain. In such case, it is probable that right optimal value of the parameters is not reached. If there were no limitations, or if the non-convergence situations could be avoided, the calibration of the model would be more trustful. Table 2 summarizes this aspect for the scenarios described in the previous section: it indicates the number of times that the bounds of the research domain for a given parameter are reached.

Table 2. Number of times the bounds of the parameters research domain are reached.

Scenario	*Number of times the bounds of the research domain are reached*
Scenario 1	3
Scenario 2	1
Scenario 3	73
Scenario 4	9

For the scenarios in which the starting values of the parameters is the one obtained by model inversion (Scenarios 1 and 2), it can be checked that, even with a small research domain, the bounds are rarely reached. This seems logical as the value of the parameters obtained by model inversion is supposed to be close to an optimal value. For the scenarios in which typical *a priori* values are considered, the bounds of the domain research are often reached when all the parameters are kept. If the number of parameters is reduced and the research domain is enlarged (as in scenario 4 with respect to scenario 3), reaching the bounds is largely reduced: 63% of calibrated parameters reach the bounds in scenario 3 compared to only 15% in scenario 4.

In order to evaluate how good the calibration is, an overall indicator is the quadratic difference between the observations and the model output (for the 1160 observations), see paragraph 3.4. The smaller this quadratic difference is, the better the calibration is. Table 3 summarizes this result for the four scenarios studied in the present work and for the calibration performed by model inversion, the so-called *Inverse calibration*. The results are presented based on the result obtained for the *Inverse calibration* method (a value lower than 1 indicates that the result is better than the result obtained by model inversion and a value higher than 1 indicates that it is worse).

Table 3. Quadratic difference between the observations and the model output – Inverse calibration as a reference

Scenario	*Quadratic difference*
Inverse calibration	1 – Reference result
Scenario 1	0.166
Scenario 2	0.245
Scenario 3	4.483
Scenario 4	0.167

In Table 3, the most important point is the value of the quadratic difference compared to the one obtained by model inversion that is set to 1 for comparison purposes. For scenarios 1 and 2, results show that this difference is largely reduced (almost by a factor from 5 to 10, especially for scenario 1). This show how data assimilation can improve an existing calibration.

For scenarios starting from typical *a priori* values, the results are very encouraging as well. When all the parameters are considered (scenario 3), the quadratic difference is only a few times higher than the one obtained by model inversion. However, when a fewer number of parameters are kept but with a larger variation range as indicated in Table 1 (scenario 4), Table 3 shows that the quadratic difference is much smaller than the one obtained by model inversion: similar results as for scenario 1 are obtained. This shows how important it is to ensure the model convergence in a domain as large as possible (scenario 3 should give better results than scenario 4, however as indicated in Table 2 for scenario 3 a large number of parameters reach the bounds of their research domain).

In addition of the overall overview of the calibration, it is important to ensure that the calibration provides good results for each observation separately, avoiding for example to reduce the error obtained for one single observation and increasing it for a large amount of them. As presented in section 3.4, a good indicator can be for example the relative difference between a given observation and the corresponding model output. Table 4 shows how many times this relative difference (averaged over the ten campaigns of measurements) is minimal, with a certain tolerance, for the 116 observations.

These results show that this indicator is improved (or at least not worsened) when starting from the values of the parameters obtained by model inversion, especially for scenario 2. This shows that even observation by observation, considered separately, data assimilation techniques can improve the model calibration. For scenarios 3 and 4 it is shown that good results are obtained for a significant number of observations as well.

Table 4. Results with respect to the relative difference between a given observation and the corresponding model output.

Scenario	*Number of times that the relative difference between a given observation and the corresponding model output is minimal (with a tolerance of 10%)*
Inverse calibration	48
Scenario 1	48
Scenario 2	67
Scenario 3	27
Scenario 4	25

Finally, in order to illustrate the effect of modifying the confidence on certain observations, a focus is performed on the observations for which a higher confidence has been considered (in scenario 2, compared to scenario 1 in which all the observations were considered in the same manner). These results are not provided for scenarios 3 and 4 since no specific focus on these observations was performed. Table 5 summarizes these results. The observations for which a higher confidence has been given, i.e. considered as more meaningful, are numbered from 1 to 16. For these observations, the relative difference between the observations and the corresponding model output (averaged over the ten campaigns of measurements) is indicated for the calibration by model inversion and for scenarios 1 and 2. For each observation, the minimal relative difference is put in bold. Moreover, the last line of Table 5 indicates the quadratic difference obtained over this subset of observations (as in Table 3, the results are given with respect to the results obtained with the calibration method by model inversion).

Table 5. Comparison between observations and model output for the observations on which a higher confidence is given.

Observation number	*Inverse calibration*	*Scenario 1*	*Scenario 2*
1	1.66%	1.54%	**1.33%**
2	**1.51%**	**1.51%**	1.54%
3	1.50%	1.39%	**1.31%**
4	0.74%	1.09%	**0.44%**
5	**4.28%**	4.30%	4.33%
6	**0.85%**	0.99%	0.89%
7	0.36%	0.40%	**0.34%**
8	0.41%	0.32%	**0.29%**
9	0.33%	0.23%	**0.21**%
10	0.37%	0.25%	**0.23**%
11	**0.11%**	0.20%	0.18%
12	0.14%	**0.10%**	**0.10%**
13	0.67%	0.72%	**0.59%**
14	1.05%	0.87%	**0.83%**
15	1.07%	0.88%	**0.82%**
16	1.42%	**1.16%**	1.31%
Overall quadratic difference	1 - Reference	0.822	0.676

Table 5 shows that scenario 2 provides better results for a large part of these observations considered individually (and when this is not the case, the relative difference is still very close to the one obtained by inverse calibration or in scenario1). Moreover, the overall indicator, giving the quadratic difference for this subset of observations, shows clearly that in both cases (scenario 1 and 2) the overall results obtained by data assimilation techniques are better than those obtained by model inversion. Therefore, it is possible, using data assimilation techniques, to easily obtain different calibrations of the model according to what the model is intended for or according to the quality of the observations.

These results show how the application of data assimilation techniques for the calibration of complex TSP models can give good calibration results, both in providing or in improving the optimal value of the calibrated parameters. Moreover, these calibration results can be obtained in about one day of calculations, compared to several weeks for the calibration by model inversion currently required (including the development of an inverse model, the pre-treatment of the measurements initially available and the different post-treatment techniques required to determine the optimal value of the parameters).

5 Conclusion and perspectives

A new method for robust and reliable model calibration, based on data assimilation techniques, for complex TSP models is currently under development. It already shows its important benefits compared to the traditional method using model inversion.

The results presented in this paper show how the application of data assimilation techniques to calibrate a complex TSP model of the secondary loop of a PWR is able to improve the calibration obtained by model inversion. In addition, it shows how a usual calibration procedure using these new techniques, coupled with a sensitivity analysis of the model, can as well provide better results than the traditional calibration method.

Therefore, compared to calibration by model inversion, this new method enables to handle conveniently situations that could not be treated before, or that would have required an important number of pre-treatments. For example, when more measurements than parameters to be calibrated are available, with the calibration by model inversion method, it was necessary to make a choice and loose some information, whereas with the new method presented in this paper it is not necessary. On the contrary, if not enough measurements are available it is not possible to calibrate the whole set of parameters using the traditional inversion method, whereas calibration approach based on data assimilation techniques is able to provide an optimal value for the whole set of parameters using efficiently all the available information. In addition, the consideration of measurements obtained in different operating conditions is greatly facilitated by data assimilation since they can all be considered simultaneously (it is therefore not necessary to post-treat independently the results obtained individually by one model inversion per each operating condition or campaign of measurements). Moreover, for some complex models, calibration by model inversion requires to develop new inverse modules when the convergence for inverse calculation is difficult, which may be very time-consuming.

In other word data assimilation method allows to automatize the model calibration procedure and hence to considerably reduce the time necessary to its calibration. Furthermore, it paves the way to improve the calibration accuracy, by enabling the use of additional information (e.g. more measurements than those strictly necessary for calibration by model inversion), or the use of available information in a specific way (e.g. according more confidence to some measurements). However, it is important to keep in mind that a good knowledge of the modelled system and of the model itself is very important in order to ensure that the results obtained applying data assimilation techniques are physically correct.

In the future, the current improvements under development should facilitate the application of this new calibration method. A major aspect is to ensure the convergence of the model over a large domain so that data assimilation techniques can provide even better results. Work on model initialization will in particular be done within the ongoing FUI ModeliScale project in partnership with Dassault Systèmes, INRIA and Phiméca. Other important point is from the methodological point of view to study: (1) how complementary studies such as sensitivity analysis of the model can be used more efficiently to properly formulate the calibration problem (e.g. by considering for calibration only the parameter that have a real impact on the variables of interest considered); (2) how data assimilation could be used for other purposes such as state estimation or prognosis.

References

Asch M., Bocquet M., Nodet M., Data Assimilation - Methods, Algorithms and Applications, SIAM, 2016.

Bouttier B., Courtier P., Data assimilation concepts and methods, Meteorological Training Course Lecture Series, ECMWF, 1999.

El Hefni B., Bouskela D., *Modeling and simulation of a complex ThermoSysPro model with OpenModelica – Dynamic Modeling of a combined power plant*, 12th International Modelica Conference, May 15-17, 2017, Prague, Czech Republic.

Kalnay E., Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Modelica, open-source modelling language, information available at: https://www.modelica.org/

OpenModelica, open-source Modelica-based modeling and simulation environment, information available at: https://openmodelica.org/

SALOME The Open Source Integration Platform for Numerical Simulation, information available at: http://www.salome-platform.org/

DOI
10.3384/ecp19157553