EPJ Nuclear Sciences & Technologies

REGULAR ARTICLE

OPEN ⌑ ACCESS

Evaluation of relevant information for optimal reflector modeling through data assimilation procedures

Jean-Philippe Argaud[*], Bertrand Bouriquet, Thomas Clerc, Flora Lucet-Sanchez, and Angélique Ponçot

EDF Recherche et développement, 1 avenue du Général de Gaulle, 92141 Clamart cedex, France

Abstract. The goal of this study is to look after the amount of information that is mandatory to get a relevant parameters optimisation by data assimilation for physical models in neutronic diffusion calculations, and to determine what is the best information to reach the optimum of accuracy at the cheapest cost. To evaluate the quality of the optimisation, we study the covariance matrix that represents the accuracy of the optimised parameter. This matrix is a classical output of the data assimilation procedure, and it is the main information about accuracy and sensitivity of the parameter optimal determination. From these studies, we present some results collected from the neutronic simulation of nuclear power plants. On the basis of the configuration studies, it has been shown that with data assimilation we can determine a global strategy to optimise the quality of the result with respect to the amount of information provided. The consequence of this is a cost reduction in terms of measurement and/or computing time with respect to the basic approach.

1 Introduction

The modeling of the reflector part of a nuclear PWR core is crucial to model the physical behaviour of the neutron fluxes inside the core. However, this element is represented by a parametrical model in the diffusion calculation code we use. Thus, the determination of the reflector parameters is a key point to obtain a good agreement with respect to reference calculation such as transport one, used as pseudo-observations. This can be done by optimisation of reflector parameters with respect to reference values. This optimisation needs to be done with care, avoiding in particular the production of aberrant results by forcing the model to match data that are not accurate enough or irrelevant. A good way is to use data assimilation, to optimise by taking into account the respective accuracy of core model and reference values. This method allows to find a good compromise between the information provided by the model and the ones provided by a reference calculation.

Data assimilation techniques have already proven to be efficient in such an exercise, as well as in field reconstruction problems [1–5]. In particular, it has been shown that there is a logarithmic-like progression of the quality of the reconstruction as a function of the number of instruments available.

Thus, there is an optimal amount of information that provides suitable results without too many measurements.

The purpose of this work is to generalise and extend the results, obtained previously on field reconstruction, for the case of parameters optimisation. It is interesting to look for the amount of information that is mandatory to get a relevant parameters optimisation, and to determine what is the best information to reach the optimum of accuracy at the cheapest cost. This question is very important in an industrial environment, as such knowledge helps to select the most relevant reference values and then to reduce the overall cost (measurement and/or computing cost) for parameters determination.

In Section 2, we present a short review of data assimilation concepts, giving the mathematical framework of the method. Then we develop the specific equations that are related to the purpose of information qualification. Those developments highlight the opportunity given by data assimilation to quantify the quality of the results. We study the evolution of the trace of the so-called analysis matrix **A** that represents the accuracy of the optimised parameter. This covariance matrix is a classical output of the data assimilation procedure, and this is the main information about accuracy and sensitivity of the optimal parameter determination.

In Section 3, we present some results collected in the field of neutronic simulation for nuclear power plants. Using the neutronic diffusion code COCAGNE [6], we seek to

* e-mail: jean-philippe.argaud@edf.fr

optimise the reflector parameters that characterise the neutronic reflector surrounding the whole reactive core in the nuclear reactor. Those studies are done on several cases which are required to be similar to the realistic configuration of Pressurised Water Reactors (PWR) of type 900 MWe or 1300 MWe, respectively named PWR900 and PWR1300 hereafter.

Finally, we conclude with the best strategy, in order to get the best result at the cheapest cost. This cost can be either evaluated in terms of computing time or number of reference measurements to provide.

2 Data assimilation and evaluation of the quality

2.1 Data assimilation

Here we briefly introduce the key points of data assimilation. More complete references on data assimilation can be found in various publications [7–16]. However, data assimilation is a wider domain and these techniques are, for example, the keys of the meteorological operational forecast. It is through advanced data assimilation methods that the weather forecast has been drastically improved during the last 30 years. Those techniques use all of the available data, such as satellite measurements, as well as sophisticated numerical models.

The ultimate goal of data assimilation methods is to estimate the inaccessible true value of the system state, $\mathbf{x^t}$, where the $\mathbf{t}$ index stands for "true state" in the so-called "control space". The basic idea of most of data assimilation methods is to combine information from an a priori estimate on the state of the system (usually called $\mathbf{x^b}$, with $\mathbf{b}$ for "background"), and measurements (referenced as $\mathbf{y^o}$ with $\mathbf{o}$ for "observation"). The background is usually the result of numerical simulations, but can also be derived from any a priori knowledge. The result of the data assimilation is called the analysis, denoted by $\mathbf{x^a}$, and it is an estimation of the researched true state $\mathbf{x^t}$.

When adjusting parameters, the state $\mathbf{x}$ is used to simulate the system using an operator named H, in order to get output that can be compared to observations. This operator embeds not only a physical equations simulator, but also sampling or post-processing of the simulation in order to get values that have the same nature as the observation. For our data assimilation purpose, we use a linearised form $\mathbf{H}$ of the full non-linear operator H. The inverse operation, to go from the observations space to the background's space, is formally given by the transpose $\mathbf{H^T}$ of the linear operator $\mathbf{H}$.

Two other ingredients are necessary. The first is the covariance matrix of the observation errors, defined as $\mathbf{R} = \mathrm{E}[(\mathbf{y^o}{-}\mathbf{H}(\mathbf{x^t})).(\mathbf{y^o}{-}\mathbf{H}(\mathbf{x^t}))^{\mathbf{T}}]$, where E[.] denotes the mathematical expectation. It can be obtained from the known errors on unbiased measurements, which means E$[\mathbf{y^o}{-}\mathbf{H}(\mathbf{x^t})] = 0$. The second ingredient is the covariance matrix of background errors, defined as $\mathbf{B} = \mathrm{E}[(\mathbf{x^b}{-}\mathbf{x^t}).(\mathbf{x^b}{-}\mathbf{x^t})^{\mathbf{T}}]$. It represents the error on the a priori state, assuming it to be unbiased following the E$[(\mathbf{x^b}{-}\mathbf{x^t})] = 0$ no

bias property. There are many ways to define this a priori state and background error matrices. However, those matrices are commonly the output of a model, the evaluation of accuracy, or the result of expert knowledge.

It can be proved [7,8], within this formalism, that the Best Unbiased Linear Estimator (BLUE, denoted as $\mathbf{x^a}$), under the linear and static assumptions, is given by the following equation:

$$x^a = x^b + K\big(y^o - H\big(x^b\big)\big), \tag{1}$$

where $\mathbf{K}$ is the gain matrix:

$$K = BH^T\big(HBH^T + R\big)^{-1}. \tag{2}$$

Equivalently, using the Sherman-Morrison-Woodbury formula, we can write the $\mathbf{K}$ matrix in the following form that we will exploit later:

$$K = \big(B^{-1} + HR^{-1}H^T\big)^{-1}H^TR^{-1}. \tag{3}$$

Additionally, we can get the analysis error covariance matrix $\mathbf{A}$, characterising the $\mathbf{x^a}$ analysis errors. This matrix can be expressed from $\mathbf{K}$ as:

$$A = (I - KH)B, \tag{4}$$

where $\mathbf{I}$ is the identity matrix.

If the probability distribution is Gaussian, solving equation (1) is equivalent to minimize the following function J($\mathbf{x}$), $\mathbf{x^a}$ being the optimal solution:

$$\begin{aligned} J(x) = &\big(x - x^b\big)^T B^{-1}\big(x - x^b\big) \\ &+ (y^o - H(x))^T R^{-1}(y^o - H(x)). \end{aligned} \tag{5}$$

The minimization of this error (or cost function) is known in data assimilation as the 3D-Var methodology [10].

In the present case, we adjust the reflector parameter D_1 setting the diffusion characteristic of fast group neutrons in diffusion equations, as described in details in reference [17].

2.2 Quality of data assimilation

One of the key points of data assimilation is that, as shown in equation (4), we gain access to the $\mathbf{A}$ covariance matrix. This matrix characterises the quality of the obtained analysis. The smaller the values in the matrix, the more "accurate" is the final result. However, understanding the full matrix content is tedious, so we choose to summarize it. For this purpose, we look, classically, at the trace of the $\mathbf{A}$ matrix, denoted as tr ($\mathbf{A}$), that is in fact the sum of all the analysis variances. The smaller this value is overall, the better is the agreement. According to our choice of the matrix trace as a quality indicator, we take the $\mathbf{R}$ and $\mathbf{B}$ matrices as diagonal in order to get a posteriori information without adding too much a priori information through terms outside the diagonal part of the matrices. Moreover, this choice is the best if no extra information is available on variances and covariances of

errors. Thus, we consider diagonal matrices, with only one parameter σ_B or σ_R, respectively, representing the overall variance for each type of errors:

$$\begin{cases} B = \sigma_B I_n \\ R = \sigma_R I_p \end{cases}. \tag{6}$$

In these equations, the indices n and p stand for the size of the space that is involved in control space ($\mathbf{x^b}$) and observation space ($\mathbf{y}$), respectively. With such a formulation, we can write the gain matrix $\mathbf{K}$ under the following form:

$$K = \sigma^2 \left(I_n + \sigma^2 HH^T \right)^{-1} H^T, \tag{7}$$

with $\sigma^2 = \frac{\sigma_B^2}{\sigma_R^2}$

Then the matrix $\mathbf{A}$ is:

$$A = \sigma_B^2 \left(I_n - \sigma^2 \left(I_n + \sigma^2 HH^T \right)^{-1} H^T H \right). \tag{8}$$

Assuming that we only want to get the optimal value of one parameter, as is often the case in optimal parameter simple determination, we get $\mathbf{HH^T} = h^2$, i.e. a scalar. Thus we can write:

$$tr(A) = \frac{\sigma_B^2}{1 + \sigma^2 h^2}. \tag{9}$$

This formulation has a very interesting asymptotic behaviour. When the measurements are very inaccurate, the result is only a function of σ_B, and when the background is accurate the result depends on both the matrix product of the observation operator $\mathbf{HH^T}$ and on the accuracy of the measurement.

In the case of one parameter to optimise p measurements, the linear operator $\mathbf{H}$ is a vector with p components h_i that are respectively the linearisation of the ith component of the operator H for the parameter λ around the value λ that are such that there is then one value of h_i by measure with:

$$h_i = \frac{dH_i}{d\lambda}(\lambda). \tag{10}$$

Under such conditions we can then write:

$$h^2 = HH^T = \sum_{i=1}^p \frac{dH_i^2}{d\lambda}(\lambda). \tag{11}$$

And then using equation (9) we can write:

$$tr(A) = \frac{\sigma_B^2}{1 + \sigma^2 \sum_{i=1}^p \frac{dH_i^2}{d\lambda}(\lambda)}. \tag{12}$$

This formulation is very interesting as we see that the more data we have, the higher is the value of h^2, and the higher is this value, the lower is the value of the trace of the analysis.

There are two main ways to increase the value of h^2. The first one is to add more measurements in order to get a higher value of p, assuming that we had a non-null value of the derivative. This leads to the second option, which is the increase of the sensitivity of the observation H_i with respect to the parameter λ.

Moreover, the global structure of the function $tr(\mathbf{A})$ as $(1 + x)^{-1}$ shows that, as already expected, the improvement of the quality is non-linear and saturating. Thus the derivative is such that the first amount of information provided has a huge impact in terms of decreasing the error function. Then when the value of x increases as we add information, the function decreases more slowly (lower value of the derivative), so a lot more information is required to keep enhancing the quality.

In the present case, it is also worth noting that the formulation in equation (10) is completely independent of the experimental values themselves. It is only the setup and the sensitivity to the parameter of the experiment that is meaningful and that is given in the operator $\mathbf{H}$.

3 Setup of the test and results

3.1 Setup of the problem for neutronic case

All the concepts and quality indicators that have been presented in the previous section are now used in the case of the neutronic problem, solved with the two energy groups diffusion theory. The problem we plan to address is the determination of an optimal parameter of the reflector D_1 in a Lefebvre-Lebigot approach, where D_1 is the diffusion coefficient of the fast group in the reflector. The aim is then to know how many campaigns are necessary to determine the best value of this parameter. To perform this task, we use the core calculation code COCAGNE developed by EDF, with the default parameters of several campaign that are representative of the nuclear fleet, either of type PWR900 or type PWR1300. The campaigns have various lengths, frequency of measurement and loading pattern setup to get a rather realistic situation. We base our study on a set of 3 PWR900 campaigns, and on a set of 6 PWR1300 campaigns.

As already mentioned, there is no use of experimental or measurement data in this study. We use equation (12) that is dependent on H but not $\mathbf{y}$ (as in Eq. (1) and Eq. (5)), so it is only the design of the campaign that matters through the modeling of the H observation operator and its linearisation $\mathbf{H}$. In the considered campaign, for several values of burnup (irregular to get closer to a real case), we consider the map of activity in the core obtained through MFC (Mobile Fission Chamber), as in the regular operation of operating plants. The combination of the COCAGNE code and of the mandatory post-processing to obtain the response of MFC is the H observation operator. Its linearised approximation $\mathbf{H}$ is then obtained with finite differences around the reference point.

3.2 Parameter determination on one campaign

We first study the impact of removing one or two instruments on the final result through the value of

tr(**A**), the trace of A. We make this study both on cases similar to REP 900 and cases similar to PWR1300.

Figure 1 represents the evolution of the trace of A as we remove, respectively, one or two maps from the collection of available maps, for each studied campaign. The trace of **A** is given as a percentage of the one calculated using all of the available campaigns. When removing one measurement, we can only decrease the quality of the analysis, leading to an increase of the variance or the trace of **A**, and so all the curves are above 100%. With such normalisation to the limiting value, the curves can be compared between all of the campaigns. In the top panel, we notice several points. The first one, which is mathematically obvious but needs to be recalled, is that the quality of the optimization decreases when some maps are removed. This degradation is stronger for the maps that are located at the beginning of the campaign, and rather small for the maps located at the end. Globally speaking, the decrease is steady between the beginning and the end of the campaigns. For the bottom panel of Figure 1, the conclusions are the same. The further into the campaign a map is located, the smaller the effect of removing this map.

In order to obtain a more global overview of the result, we study the case of the PWR1300. In Figure 2, we plotted the same information as in Figure 1, but for 3 cases among the 6 of the PWR1300 set.

As in the case of the PWR900, it seems for PWR1300 that the maps at the beginning of the campaign are more influential on the reflector result than those located at the end. This result seems to be even clearer, as the slope of the initial decrease looks to be sharper in Figure 2 than in Figure 1.
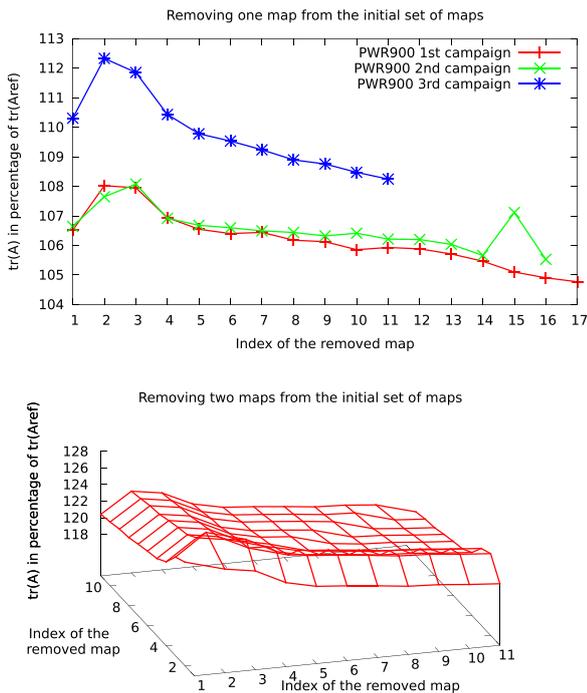

Fig. 1. Impact of removing one (top subplot) or two (bottom subplot) maps on the quality of the parameter determination, as a function of the index of the chosen map in 3 cases of PWR900 campaigns.
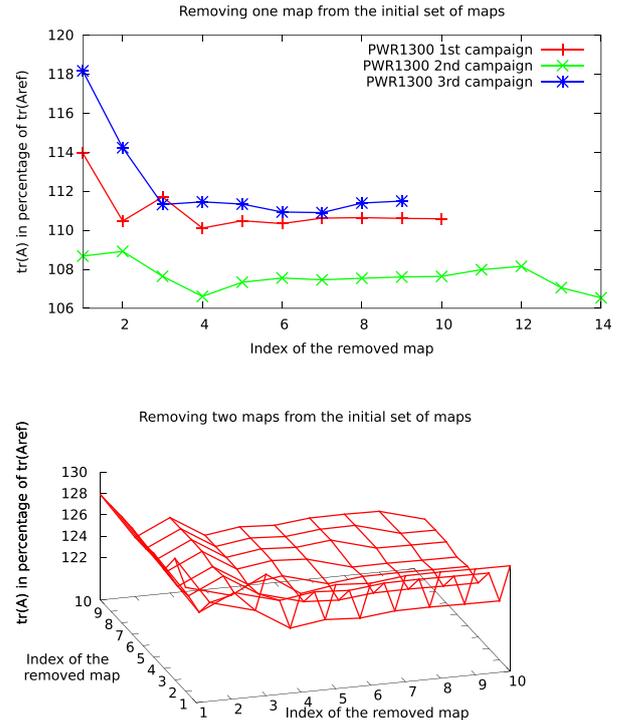

Fig. 2. Impact of removing one (top subplot) or two (bottom subplot) maps on the quality of the parameter determination, as a function of the index of the chosen map in 3 cases of PWR900 campaigns.

For the PWR900 campaigns, it is not straightforward to conclude that the maps at the beginning of the campaign are more interesting for data assimilation than the ones at the end. To show more clearly that the maps at the beginning are more important, we redo the work described in Figure 3 in another way. For this purpose, we include in the data assimilation more and more experimental flux maps in two ways: the first consists of taking the map into account following the increase of the burnup characteristic (red curve) and the second consists of the opposite, that is using the maps in decreasing order of burnup (green curve). A third way is to look at the map that gives the best result, and adding maps one by one in the same way (blue curve).
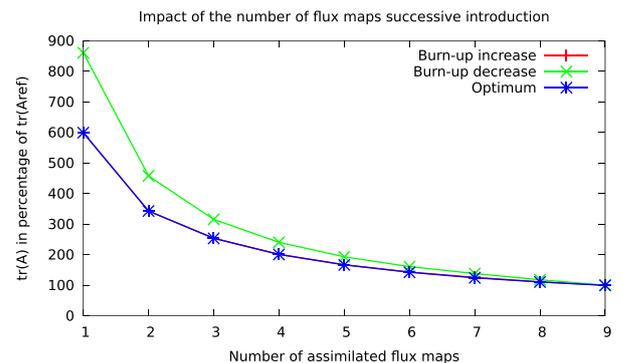

Fig. 3. Variation of the evolution of the trace of **A** as a function of the order of introduction of the flux maps in the data assimilation process.

Finally, when all of the information is added, all of the curves reach the same point. With respect to what is plotted in Figure 3, we noticed that a data assimilation procedure taking into account the maps following increasing burnups gives results that are very close to the optimal. On the contrary, the assimilation procedure taking into account the maps following decreasing burnups is not a good choice when only a few maps are taken into account. Thus, even if the curves that give the impact of removing one map are not uniformly decreasing, we can still conclude that the maps from the beginning of the campaign are more influential than the maps at the end.

Thus, for data assimilation methods, some data are more influential than others. The more the core is burnt, the less the map seems to be influential in decreasing the trace of **A**. The burning of the core tends to make the flux map spatially more homogeneous ("flat"). As the D_1 reflector parameter governs mainly the global curvature of the flux inside the reactive core, if it is determined on a quasi-fully-burnt core with very flat flux map, the resulting D_1 will be very insensitive to the burnup.

It is then possible to determine a strategy in using this framework to get optimal results at the lowest cost.

3.3 Parameter determination on several campaigns

As a global multi-campaigns strategy of optimisation can be determined, we go to the next step. We compare the trace of the matrix calculated from multiple campaigns to that coming only from one campaign. We build multi-campaigns cases in the following way. Assuming that we dispose of the flux maps coming from 3 campaigns, we solve a data assimilation problem with the first map of each campaign. Thus we obtain a first analysis and value of $\mathrm{tr}(\mathbf{A})$. Then, we use in addition the second map of each of the three campaigns, and we obtain a new analysis for a total of six maps, and so on.

To make a good comparison, we also show the evolution of $\mathrm{tr}(\mathbf{A})$ for data assimilation with only one campaign. For those curves, the first point is obtained using the first available map then the second using the two first maps, and so on.

In order to make a comparison between the multi-campaigns data assimilation, we calculate statistical indicators (mean and standard deviation) for each kind of multi-campaigns data assimilation (use of 2 campaigns, 3 campaigns . . .) on all the possible combinations. For example, for the data assimilation of 3 campaigns over a set of 6, like in PWR1300, there are 20 possible combinations. Those statistical indicators are given in all the coming figures. It is worth pointing out that the standard deviation does not correspond to the standard deviation on the analysis (which is here directly $\mathrm{tr}(\mathbf{A})$ but to the standard deviation of $\mathrm{tr}(\mathbf{A})$. Here, we are studying the variability of the indicator $\mathrm{tr}(\mathbf{A})$ as a function of the chosen set of campaigns.

For the PWR900, we have 3 campaigns with associated calculations of flux maps. In Figures 4a–4c, we present the evaluation of the trace of A as a function of the total number of maps assimilated, respectively for 3 scenarios of

data assimilation: mono-campaign data assimilation (3 cases), 2-campaign data assimilation (3 cases), and 3-campaign data assimilation (1 case). The data for mono-campaign and 2-campaign are a mean value on the different results for each scenario. Figure 4d presents the compared evolution of the standard deviation for the mono-campaign and 2-campaigns cases.
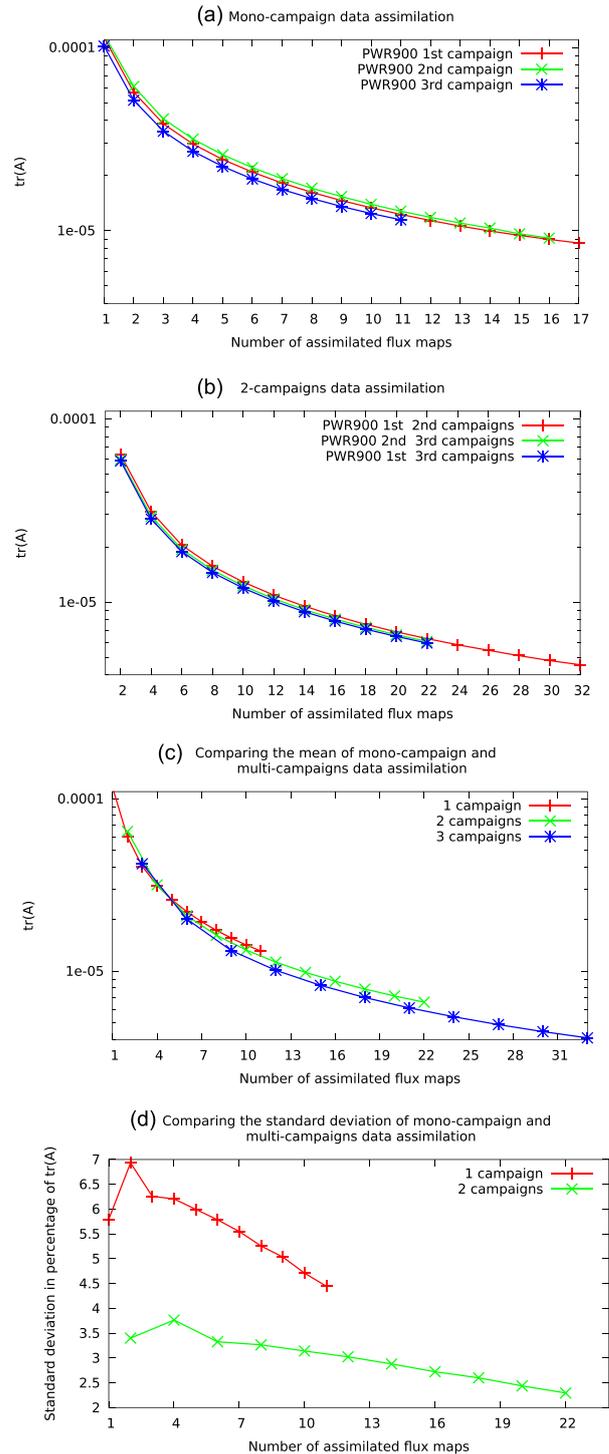

Fig. 4. Evolution of the trace of **A** as a function of the number of assimilated flux maps for the PWR900.

For all of the scenarios, we notice that the trace of **A** decreases very steadily as a function of the total number of assimilated flux maps, and that we roughly lose one order of magnitude of the trace of **A** between the value obtained with only one map and the trace calculated with all the available ones (between 11 and 32 maps according to the scenario). For mono-campaign data assimilation, where Figure 4a presents the evolution of the trace, we notice that all the curves are very similar. For this case, the standard deviations calculated on those 3 assimilations vary between 4.5% and 7% as we can see on Figure 4d. Apart from one point, there is a constant decrease of the value of the standard deviation as a function of the number of maps used, as expected. For the data assimilation on the 2-campaign case, shown on Figure 4b, the 3 curves are very close. We also notice on Figure 4d that there is a regular decrease of the standard deviation, but the amplitude of variation is lower (around 3%). If we make a comparison between the 2-campaign and 3-campaign cases, shown in Figures 4b and 4c, we notice that all scenarios of data assimilation have the same behaviour as a function of the number of assimilated maps. However, when the number of maps increases, we notice that the 3-campaign data assimilation is better than the 2-campaign data assimilation, that is itself better than the mono-campaign data assimilation.

For the PWR1300, we take into account 6 different campaigns and we do the same process as for PWR900. As for the PWR900 cases, we compute the evolution of the trace of **A** as a function of the total number of assimilated maps, for different sets of n campaigns (n-tuples). In this case also, the curves in Figures 5a–5e present the same behaviours: the trace of **A** decreases steadily as a function of the number of assimilated maps. However, in comparison to the result obtained in the PWR900 case, we notice a higher variability of the results as a function of the chosen n-tuple (8%, at most, of standard deviation in Fig. 5d with respect to 25% in Fig. 4b).

Figure 6 allows to compare the different scenarios of data assimilation through the mean value and the standard deviation over all of the possible combinations, for each of the 6 scenarios. For example, for the data assimilation with 4 campaigns, there are $C_6^4 = 15$ ways to choose 4 campaigns among the 6 taken into account. The different scenarios of data assimilation behave in a similar way. The first result is that, for an equal number of assimilated maps, the trace of **A** is smaller if more campaigns are used.

With the same number of assimilated maps, it is more valuable to choose them in the largest amount of campaigns, as the trace of **A** is smaller if we choose the maps in different campaigns rather than in the same campaign. Moreover, the study of the variation of the standard deviation of the trace of **A**, as a function of the n-tuple of campaigns used, shows that, for any given number of campaigns used, the more maps we assimilate, the more we reduce the standard deviation of the result. This variability also decreases as more campaigns are taken into account.

For the standard deviation, the curves are decreasing more sharply if only a few campaigns are involved in the data assimilation processes. Thus, assimilation with one campaign is very dependent on the number of assimilated
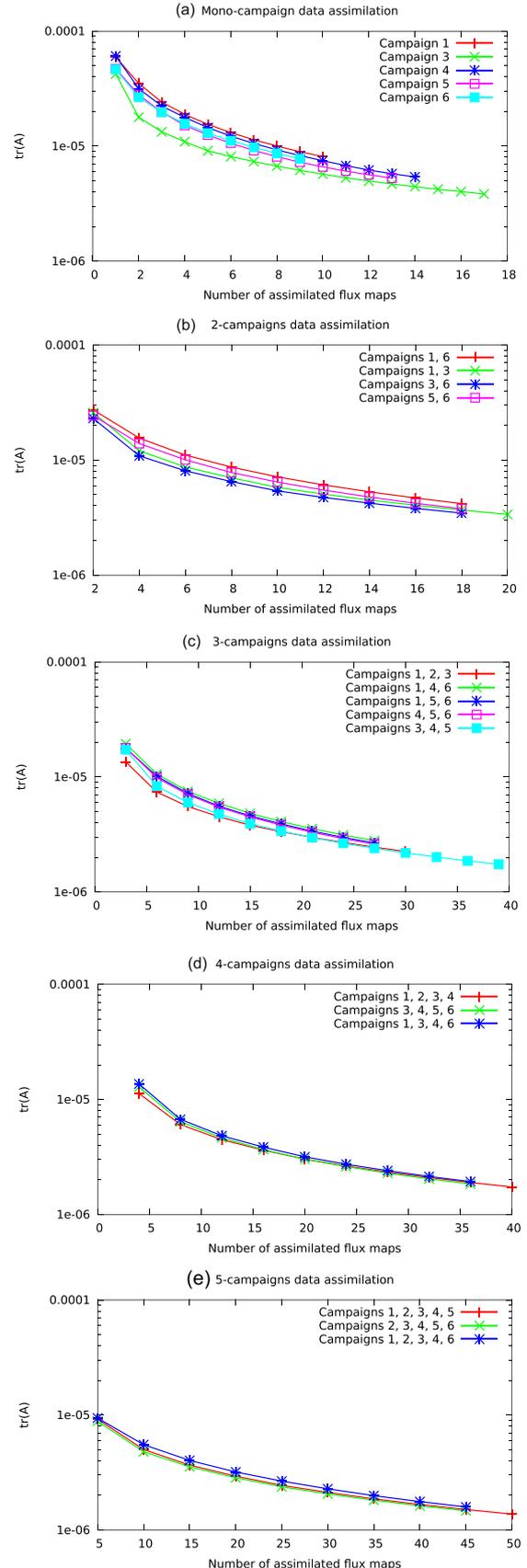

Fig. 5. Evolution of the trace of **A** as a function of the number of assimilated flux maps for the PWR1300.

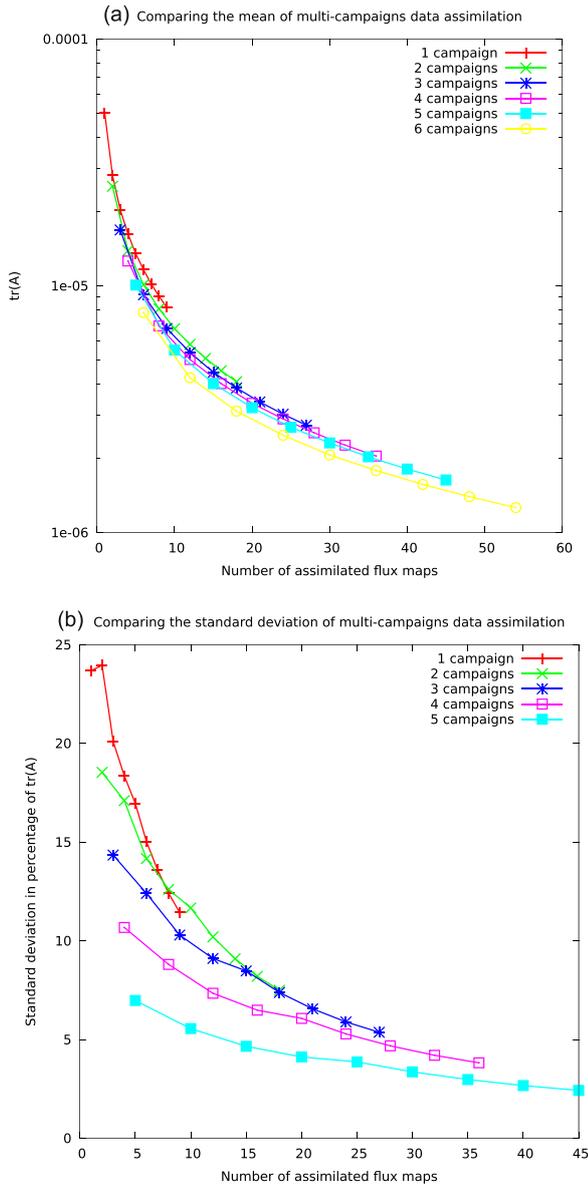**Fig. 6.** Comparison of the evolution of the trace of **A** (mean on top (a), standard deviation on bottom (b) as a function of the number of flux maps used for various n-tuplets of campaigns in the case of PWR1300.

maps, and on the choice of the used campaign (standard deviation around 20% of the 3 first assimilated maps). By contrast, there is very small variability in the result for the assimilation involving 5 campaigns (variation of the standard deviation of only a few percent). These results are still true whatever the number of assimilated maps. This means choosing maps in several campaigns makes the adjustment of the parameter more robust.

We can also guess that, using maps coming from several campaigns, we contribute to removing the systematic error that can be specific to a given campaign.

To summarise, data assimilation on several campaigns shows that it is more valuable to use p maps coming from n campaigns than $n \times p$ maps coming from the same campaign, for three reasons: it reduces the trace of **A**, it reduces the variability of the trace, and it reduces the systematic error.

We therefore conclude that, in both the PWR900 and PWR1300 cases, multi-campaigns data assimilation is more beneficial that mono-campaign data assimilation.

4 Conclusion

The goal of this paper was to provide an analysis of the observation impact, in order to get the best result at the cheapest cost for parameter optimisation. As we are in the optimal estimation framework of data assimilation, we are able to propose an estimate of the final quality based on the trace of the covariance matrix **A** of analysis error, where **A** is a natural sub-product of the data assimilation calculation. With this specific tool, we handle the adjustment of the reflector parameter D_1 for the simulation of various representative cores performed with the neutronic code COCAGNE. It is worth noting that, in this paper, no measurement data (or even pseudo-measurement data) has been used, as it is not necessary. It is only the study and the modeling of the core configurations that lead to the results obtained. To this purpose, we work on 3 or 6 campaign sets for PWR900 and PWR1300, respectively.

The studies demonstrate that, as expected, the more flux maps we assimilate, the more the trace of the matrix **A** of analysis error is reduced. However, if we want to choose the fewest number of maps possible, it has been shown that it is more favourable to take maps from the beginning of the campaign rather than from the end. Nevertheless, we emphasise the need to be careful not to choose maps only at the beginning of the campaign, in which case the parameter adjustment can be more sensitive or unstable. This result can be interpreted as a consequence of the core burnup. Actually, the flux maps become flatter during the depletion of the core, and then less information on the curvature of the flux map can be collected, which leads to more difficulties in adjusting the D_1 reflector parameter.

Another result is that using multi-campaign data significantly improves the efficiency of parameters optimisation. Indeed, for the same number of maps used, it is better to use p flux maps from n campaigns than to use $n \times p$ flux maps coming from only one campaign. We also notice that taking flux maps from various campaigns reduces the variability of the adjustment result. On the basis of the configuration studies, data assimilation on several campaigns allows one to obtain an analysis variance that is far more stable, and therefore far more predictive, than that obtained using only one campaign.

For both reasons, multi-campaign data assimilation gives better results than mono-campaign data assimilation.

References

1. B. Bouriquet et al., Differential influence of instruments in nuclear core activity evaluation by data assimilation, Nucl. Instrum. Methods Phys. Res. A **626-627**, 97 (2011)

2. B. Bouriquet et al., Robustness of nuclear core activity reconstruction by data assimilation, Nucl. Instrum. Methods Phys. Res. A **629**, 282 (2011)

3. D.G. Cacuci, M. Ionescu-Bujor, Best-estimate model calibration and prediction through experimental data assimilation - I: Mathematical framework, Nucl. Sci. Eng. **165**, 18 (2010)

4. D. Petruzzi, G. Cacuci, F. d'Auria, Best-estimate model calibration and prediction through experimental data assimilation - II: Application to a blowdown benchmark experiment, Nucl. Sci. Eng. **165**, 45 (2010)

5. J.P. Bouriquet, A.R. Cugnart, Optimal design of measurement network for neutronic activity field reconstruction by data assimilation, Nucl. Instrum. Methods Phys. Res. A **664**, 117 (2012)

6. T. Clerc et al., An advanced computational scheme for the optimization of 2D radial reflectors in pressurized water reactors, Nucl. Eng. Des. **273**, 560 (2014)

7. F. Bouttier, P. Courtier, Data assimilation concepts and methods, Meteorological training course lecture series, ECMWF, 1999

8. E. Kalnay, *Atmospheric modeling, data assimilation and predictability* (Cambridge University Press, 2003)

9. F. Rabier et al., The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics, Q. J. R. Meteorol. Soc. **126**, 1143 (2000)

10. O. Talagrand, Assimilation of observations, an introduction, J. Meteorol. Soc. Japan **75**, 191 (1997)

11. K. Ide, P. Courtier, M. Ghil, A.C. Lorenc, United notation for data assimilation: operational, sequential and variational, J. Meteorol. Soc. Japan **75**, 181 (1997)

12. D.F. Parrish, J.C. Derber, The National Meteorological Center's spectral statistical interpolation analysis system, Mon. Weather Rev. **120**, 1747 (1992)

13. R. Todling, S.E. Cohn, Suboptimal schemes for atmospheric data assimilation based on the Kalman filter, Mon. Weather Rev. **122**, 530 (1994)

14. S.M. Uppala et al., The ERA-40 re-analysis, Q. J. R. Meteorol. Soc. **131**, 2961 (2005)

15. E. Kalnay et al., The NCEP/NCAR 40-year reanalysis project, Bull. Am. Meteorol. Soc. **77**, 437 (1996)

16. G.J. Huffman et al., The Global Precipitation Climatology Project (GPCP) combined precipitation dataset, Bull. Am. Meteorol. Soc. **78**, 5 (1997)

17. S. Massart, S. Buis, P. Erhard, G. Gacon, Use of 3DVAR and Kalman filter approaches for neutronic state and parameter estimation in nuclear reactors, Nucl. Sci. Eng. **155**, 409 (2007)