

Perspective on Predictive Modeling: Current Status, New High-Order Methodology and Outlook for Energy Systems

Dan Gabriel Cacuci

Center for Nuclear Science and Energy, University of South Carolina, Columbia, SC 29208, USA;
cacuci@cec.sc.edu

Abstract: This work presents a perspective on deterministic predictive modeling methodologies, which aim at extracting best-estimate values for model responses and parameters along with reduced predicted uncertainties for these best-estimate values. The two oldest such methodologies are the data-adjustment method, which stems from the nuclear energy field, and the data-assimilation method, which is implemented in the geophysical sciences. Both of these methodologies attempt to minimize, in the least-square sense, a user-defined functional that represents the discrepancies between computed and measured model responses. These two methodologies were briefly reviewed and shown to be inconsistent even to first-order in the sensitivities of the response to the model parameters. In contrast to these methodologies, it was shown that the “maximum entropy”-based predictive modeling methodology (called BERRU-PM) that was developed by the author not only dispenses with the subjective “user-chosen functional to be minimized” but is also inherently amenable to high-order formulations. This inherent potential was illustrated by presenting a novel, higher-order, MaxEnt-based predictive modeling methodology, labelled BERRU-PM-2+, which is complete and exact to second-order sensitivities and moments of both the a priori and posterior distributions of responses and parameters, while explicitly including third- and fourth-order sensitivities and correlations, thus indicating the mechanism for incorporating information of orders higher than second in predictive modeling. The presentation of this new predictive modeling methodology also aims at motivating a widespread application of predictive modeling principles and methodologies in the energy sciences for obtaining best-estimate results with reduced uncertainties.

Citation: Cacuci, D.G. Perspective on Predictive Modeling: Current Status, New High-Order Methodology and Outlook for Energy Systems. *Energies* **2023**, *16*, 933.

<https://doi.org/10.3390/en16020933>

Academic Editors: Paul Stewart and George Halkos

Received: 30 October 2022

Revised: 5 December 2022

Accepted: 10 January 2023

Published: 13 January 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: predictive modeling; data adjustment; data assimilation; maximum entropy principle; least squares; response sensitivities to model parameters; model calibration; model validation; high-order sensitivities; high-order correlations

1. Introduction

The modeling of a physical system includes the following elements:

- (a) a mathematical model comprising equations that relate the system’s independent variables and parameters to the system’s state (i.e., dependent) variables;
- (b) inequality and/or equality constraints that delimit the ranges of the system’s parameters;
- (c) one or several computational results, customarily referred to as “system responses” (or objective functions, or indices of performance), which are computed using the mathematical model; and
- (d) experimentally measured responses, with their respective nominal (mean) values and uncertainties (variances, covariances, skewness, kurtosis, etc.).

The results of measurements and computations are never perfectly accurate but are subject to unavoidable errors. Thus, the results of measurements are inevitably influenced by experimental errors stemming from imperfect instruments or imperfectly known calibration standards. Around any reported experimental value, therefore, there always

exists a range of values that may also be plausibly representative of the true but unknown value of the measured quantity. Similarly, computations are subject to errors stemming from uncertain model parameters, initial and boundary conditions, imperfectly known physical processes, imperfectly known physical geometry caused by imperfect material boundaries and, finally, inexact numerical computations. Therefore, knowledge of just the nominal values of experimentally measured or computed quantities is insufficient for determining the reliability of results in practical applications. The quantitative uncertainties accompanying measurements and computations are also needed, along with the respective nominal values.

The uncertainties inherent to experimental and computational results provide the basic motivation for performing quantitative model verification, validation, qualification, and predictive estimation. The activity of “code/model verification” addresses the question “are the equations underlying the mathematical model solved correctly?” The activity of “code/model validation” addresses the question “does the respective model represent reality?” Model validation requires benchmarking the respective model against independently obtained experimental results, including all accompanying uncertainties (computational, experimental, etc.) to quantify the respective model’s accuracy by comparing the computational results produced by the model, including their accompanying statistical characteristics (standard deviations, skewness, kurtosis, and corresponding correlations) with the corresponding experimental results. The activity “code/model qualification” addresses the certification that a proposed simulation and/or design methodology satisfies all performance and safety specifications regarding the system under consideration.

In the author’s view, “predictive modeling” comprises three key elements, namely: model calibration, model extrapolation, and estimation of the validation domain. Model calibration addresses the integration of experimental data for the purpose of updating the parameters underlying the computational model. Important components include the estimation of inconsistencies in the experimental data, and quantification of the biases between model predictions and experimental data. The state-of-the-art of model calibration is fairly well developed, but current methods are still hampered in practice by the significant computational effort required for performing exhaustive (i.e., including all of the sensitivities of model responses to the model’s parameters) model calibration for large-scale models. Reducing the computational effort is paramount, and methods based on adjoint models show great promise in this regard. “Model extrapolation” addresses the quantification of the uncertainties in predictions under new conditions, including both untested regions of the parameter space and higher levels of system complexity in the validation hierarchy. Extrapolation of models and the resulting increase of uncertainty are poorly understood, particularly the estimation of uncertainty that results from nonlinear coupling of two or more physical phenomena that were not coupled in the existing validation database. The quantification of the validation domain underlying the models of interest requires estimation of contours of constant uncertainty in the high-dimensional space that characterizes the application of interest. In practice, this involves the identification of areas where the predictive estimation of uncertainty meets specified requirements for the performance, reliability, or safety of the system of interest. The conceptual and mathematical development of methods for quantifying the validation domain is in an incipient stage.

The earliest activities aimed at extracting best-estimate values for model parameters were initiated in the 1960s [1–3] in the nuclear energy field. These activities aimed at evaluating neutron cross sections by using time-independent reactor physics experiments for measuring “integral quantities” (i.e., model responses) such as reaction rates and multiplication factors. These activities reached conceptual maturity under the name of “cross section adjustment” [4–8], which used a least-square procedure, weighted with first-order response sensitivities, for combining uncertainties in the model responses with uncertainties in the corresponding experimental data. The resulting “adjusted” neutron cross

sections (model parameters) and their “adjusted” uncertainties were subsequently employed in the respective reactor core model to predict improved model responses (e.g., improved reaction rates and reaction rate ratios, reactor multiplication factors, Doppler coefficients) in an extended application domain (e.g., a new or improved reactor core design). Since the neutron transport (or diffusion) equation is linear in the neutron flux (dependent variable), it admits a corresponding adjoint transport (or diffusion) equation for the adjoint neutron flux, which was in turn used to compute efficiently the first-order response sensitivities, which appeared as weighting functions in the least squares adjustment procedure. The “data-adjustment” methodology can thus be considered to have been the earliest systematic methodology that embodies principles of “predictive modeling”. The principles underlying this data-adjustment methodology for the linear, time-independent nuclear energy systems modeled by the neutron transport equation are briefly summarized in Section 2.1.

A pioneering formalism for performing data adjustment in the context of time-dependent nonlinear system was conceived by Cacuci’s group at Oak Ridge National Laboratory [9], using the adjoint method of sensitivity analysis for nonlinear systems conceived and developed by Cacuci [10], who had formally introduced the concepts of sensitivity analysis by using concepts of nonlinear functional analysis. The data adjustment formalism for time-dependent nonlinear system [9] also relies on the minimization, in the least-square sense, of a user-defined functional; its underlying principles are presented in Section 2.2.

In the late 1980s and during the 1990s, the fundamental concepts underlying “data adjustment” seem to have been rediscovered in the atmospheric and geophysical sciences while developing the so-called “data assimilation” procedure, in that the concepts underlying data assimilation are the same as those underlying the previously developed “data adjustment” procedure, relying on the minimization, in the least-square sense, of a user-defined “cost functional. The principles underlying data assimilation procedures are amply described in the books by Lewis et al. [11], Lahoz et al. [12], Farago et al. [13] and Cacuci et al. [14]. These principles are briefly summarized in Section 2.3, in order to facilitate comparisons with the other procedures that aim at the same goals.

In contradistinction to the methods used for data adjustment and/or assimilation, the BERRU predictive modeling methodology developed by Cacuci [15,16] dispenses with the need to minimize an a priori chosen “cost functional” (usually a quadratic functional that represents the weighted errors between measured and computed responses), by employing the “maximum entropy” (MaxEnt) principle [17] to combine computational and experimental information for obtaining best-estimate predicted mean values for model responses and parameters, together with reduced predicted uncertainties for these best-estimate values. The BERRU predictive modeling methodology [15,16] incorporates first-order sensitivities of model responses with respect to the model parameters, computed using the adjoint method [10].

The conception of the second- and higher-order comprehensive adjoint sensitivity analysis methodology by Cacuci [18–21] has enabled the efficient computation of exact expression for arbitrarily-high order sensitivities, thus opening the path to further advances of all of the activities which could use such higher-order sensitivities. In particular, the availability of the second- and higher-order model response sensitivities to model parameters has also enabled the extension of the BERRU predictive modeling methodology from first-order to higher-orders. This extended, novel MaxEnt-based predictive modeling formalism, is labeled the “BERRU-PM-2+” methodology because it incorporates not only all of the possible second-order sensitivities and correlations, but also provides the mechanism for incorporating arbitrarily-high-order sensitivities of model responses with respect to model parameters. The novel “BERRU-PM-2+” methodology is presented in Section 3 of this work.

Section 4 concludes this work by presenting a comparative discussion of the deterministic methodologies currently available in the context of their abilities to predict best-

estimate mean values and uncertainties (variances/covariances) for calibrating and validating large-scale models of energy-related physical systems.

2. Traditional Least-Square Based Deterministic Predictive Modeling Methodologies

This Section briefly reviews the traditional deterministic methodologies which comprise elements of predictive modeling, as follows:

- (i) The so-called “data-adjustment methodology”, which is perhaps the oldest such methodology in use and was developed in the nuclear energy field. The initial “data-adjustment methodology”, which was developed in the 1960s for the large-scale time-independent linear systems modeled by the linear neutral particle transport equation, is briefly reviewed in Section 2.1. The data-adjustment methodology for time-dependent nonlinear systems, which could be considered to be the predecessor for the so-called “4D-Var” methodology mentioned below, is reviewed in Section 2.2.
- (ii) The so-called “data assimilation” methodology, which is implemented for the large-scale time-dependent systems encountered in the atmospheric and geophysical field, is briefly reviewed in Section 2.3.

2.1. Time-Independent Least-Squares Based Data Adjustment

The latest versions of the “data-adjustment” methodology for the time-independent systems are described in the leading nuclear engineering code-packages SCALE [22] and CRISTAL [23]. This data-adjustment methodology is intended for assimilating experimental information in order to “adjust” (calibrate) model parameters (namely nuclear cross sections) describing linear reactor physics problems modeled by the linear Boltzmann equation. The data-adjustment methodology relies on a “Generalized Linear Least Squares Methodology (GLLSM)” which involves minimizing the following user-defined Lagrangian functional, denoted below as $L(\mathbf{y}, \mathbf{z})$:

$$\min_{\mathbf{y}, \mathbf{z}} L(\mathbf{y}, \mathbf{z}); L(\mathbf{y}, \mathbf{z}) \triangleq (\mathbf{y}, \mathbf{z})^{\dagger} \begin{pmatrix} \mathbf{C}_{mm} & \mathbf{C}_{m\alpha} \\ \mathbf{C}_{\alpha m} & \mathbf{C}_{\alpha\alpha} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} + 2\lambda^{\dagger} (\mathbf{S}_k \mathbf{z} - \mathbf{y}). \quad (1)$$

The quantities appearing in functional $L(\mathbf{y}, \mathbf{z})$ defined in Equation (1) have the following meanings/definitions:

- (i) The (symmetric) relative covariance matrix for the model parameters, $\mathbf{C}_{\alpha\alpha}$, is defined as follows:

$$\mathbf{C}_{\alpha\alpha} \triangleq \left[\frac{\text{cov}(\alpha_n, \alpha_p)}{\alpha_n \alpha_p} \right]; n, p = 1, \dots, M, \quad (2)$$

where α_n denotes the n th model parameter (cross sections); $\text{cov}(\alpha_n, \alpha_p)$ denotes the covariance between the two respective model parameters; M denotes the total number of model parameters.

- (ii) The (symmetric) relative covariance matrix for the measured model responses (effective multiplication factors), $\mathbf{C}_{mm}$, is defined as follows:

$$\mathbf{C}_{mm} \triangleq \left(\frac{m_i}{k_i} \right) \left[\frac{\text{cov}(m_i, m_j)}{m_i m_j} \right] \left(\frac{m_j}{k_j} \right); i, j = 1, \dots, I, \quad (3)$$

where k_i denotes the nominal values of the computed effective multiplication factor, using the nominal values of the model parameters (cross sections); m_i denotes the nominal values of the measured effective multiplication factor; $\text{cov}(m_i, m_j)$ denotes the

covariance between two measured effective multiplication factors; I denotes the total number of model responses (measured or computed effective multiplication factors).

(iii) The rectangular matrix $\mathbf{C}_{am}$ contains as elements the relative covariances between the measured responses and model parameters, and is defined as follows:

$$\mathbf{C}_{am} \triangleq \left[\frac{\text{cov}(\alpha_n, m_i)}{\alpha_n m_i} \right] \left(\frac{m_i}{k_i} \right); n = 1, \dots, M; i = 1, \dots, I; \mathbf{C}_{m\alpha} \triangleq \mathbf{C}_{am}^\dagger. \quad (4)$$

(iv) Each model response, denoted as $k_i(\mathbf{a})$, is considered to be a linear function of the model parameters, denoted as $\mathbf{a} \triangleq (\alpha_1, \dots, \alpha_M)^\dagger$, having the following form:

$$k_i(\mathbf{a} + \delta\mathbf{a}) = k_i(\mathbf{a}) \left[1 + \sum_{n=1}^M S_n^i \frac{\delta\alpha_n}{\alpha_n} \right]; i = 1, \dots, I; \delta\mathbf{a} \triangleq (\delta\alpha_1, \dots, \delta\alpha_M)^\dagger, \quad (5)$$

where $\delta\alpha_n$ denotes an arbitrary perturbation in the n -th model parameter, and where S_n^i represents the relative sensitivity of the i -th response to the n -th parameter, computed at the nominal parameter values, and is defined as follows:

$$S_n^i \triangleq \frac{\partial k_i}{\partial \alpha_n} \frac{\alpha_n}{k_i}; i = 1, \dots, I; n = 1, \dots, M. \quad (6)$$

The linear relationship representing the model response as a function of the model parameters is re-written in the following form:

$$\mathbf{y} = \mathbf{d} + \mathbf{S}_k \mathbf{z}, \quad (7)$$

where

$$\mathbf{S}_k \triangleq [S_n^i] \triangleq \left[\frac{\partial k_i}{\partial \alpha_n} \frac{\alpha_n}{k_i} \right]; i = 1, \dots, I; n = 1, \dots, M; \quad (8)$$

$$\mathbf{z} \triangleq (z_1, \dots, z_M)^\dagger; z_n \triangleq \frac{\delta\alpha_n}{\alpha_n}; \quad (9)$$

$$\mathbf{d} \triangleq (d_1, \dots, d_I)^\dagger; d_i \triangleq \frac{k_i(\mathbf{a}) - m_i}{k_i(\mathbf{a})}; \quad (10)$$

$$\mathbf{y} \triangleq (y_1, \dots, y_I)^\dagger; y_i \triangleq \frac{k_i(\mathbf{a} + \delta\mathbf{a}) - m_i}{k_i(\mathbf{a})}. \quad (11)$$

(v) The (column) vector of Lagrange multipliers $\boldsymbol{\lambda} \triangleq (\lambda_1, \dots, \lambda_I)^\dagger$, which appears in the functional $R(\mathbf{y}, \mathbf{z})$ defined in Equation (1), enforces the linear model defined by Equation (7) as a “hard constraint.”

The “adjusted” parameter and response values are obtained by solving the following equations, which are to be satisfied at the minimum of the functional $R(\mathbf{y}, \mathbf{z})$:

$$\frac{\partial R(\mathbf{y}, \mathbf{z})}{\partial \mathbf{z}} = \frac{\partial R(\mathbf{y}, \mathbf{z})}{\partial \mathbf{y}} = 0. \quad (12)$$

The solution of Equation (12) yields the following relations for the “adjusted” parameter and response values:

$$\mathbf{z} = (\mathbf{C}_{am} - \mathbf{C}_{a\alpha} \mathbf{S}_k^\dagger) \mathbf{C}_{dd}^{-1} \mathbf{d}, \quad (13)$$

$$\mathbf{y} = (\mathbf{C}_{mm} - \mathbf{C}_{m\alpha} \mathbf{S}_k^\dagger) \mathbf{C}_{dd}^{-1} \mathbf{d}, \quad (14)$$

$$\mathbf{C}_{dd} \triangleq \mathbf{S}_k \mathbf{C}_{\alpha\alpha} \mathbf{S}_k^\dagger + \mathbf{C}_{mm} - \mathbf{S}_k \mathbf{C}_{\alpha m} - \mathbf{C}_{m\alpha} \mathbf{S}_k^\dagger. \quad (15)$$

The “adjusted covariance matrix”, $\mathbf{C}_{mm}^{adj}$, for the “adjusted” responses (measured or computed) and “the adjusted covariance matrix”, $\mathbf{C}_{\alpha\alpha}^{adj}$, for the “adjusted” parameters, respectively, are obtained by using Equations (13) and (14), to obtain the following expressions:

$$\mathbf{C}_{mm}^{adj} = \mathbf{C}_{mm} - (\mathbf{C}_{mm} - \mathbf{C}_{m\alpha} \mathbf{S}_k^\dagger) \mathbf{C}_{dd}^{-1} (\mathbf{C}_{mm} - \mathbf{S}_k \mathbf{C}_{\alpha m}), \quad (16)$$

$$\mathbf{C}_{\alpha\alpha}^{adj} = \mathbf{C}_{\alpha\alpha} - (\mathbf{C}_{\alpha m} - \mathbf{C}_{\alpha\alpha} \mathbf{S}_k^\dagger) \mathbf{C}_{dd}^{-1} (\mathbf{C}_{m\alpha} - \mathbf{S}_k \mathbf{C}_{\alpha\alpha}). \quad (17)$$

Care must be exercised, however, since the indiscriminate incorporation of all seemingly relevant experimental-response data could produce a set of calibrated (adjusted) parameter values that might differ unreasonably much from the corresponding original nominal values and might even fail to improve the agreement between the calculated and measured values of some of the very responses that were used to calibrate the model parameters. When calibrating (adjusting) a library of model parameters, it is tacitly assumed that the given parameters are basically “correct”, in that their true values lie within some reasonable uncertainty of their reported value. However, calibration can be used to increase the accuracy of a model beyond that which is constrained by its reliance on the reported value. The calibration procedure uses additional data (e.g., experimentally measured responses) for improving the parameter values while reducing their uncertainties. Although such additional information induces modifications of the original parameter values, the adjusted parameters are still generally expected to remain consistent with their original nominal values, within the range of their original uncertainties. However, calibration of a library of model parameters by experimental responses which significantly deviate from their respective computed values would significantly modify the resulting adjusted parameters, perhaps even violating the restriction of linearity expressed by Equation (7). Such unlikely adjustments would most probably lead to failure of even reproducing the original experimental responses.

On the other hand, calibrating model parameter by using measured responses that are very close to their respective computed values would cause minimal parameter modifications and a nearly perfect reproduction of the given responses by the calibrated parameters, as would be expected. In such a case, the given responses would be considered as being consistent with the parameter library, in contradistinction to adjustment by inconsistent experimental information, in which case the adjustment (calibration) could fail because of inconsistencies. These considerations clearly underscore the need for using a quantitative indicator to measure the mutual and joint consistency of the information available for model calibration.

2.2. Time-Dependent Least-Squares Based Data Adjustment

The time-dependent least-squares based data-adjustment methodology [9] was developed in the early 1980's to analyze time-dependent problems of fundamental interest in nuclear engineering (e.g., thermal-hydraulics transient computations and experiments). The underlying concepts are as above and the transient phenomena can be analyzed either globally (“off-line”), when the entire information during the time interval of interest is being analyzed at once or locally, when the data assimilation/adjustment is performed one time-step at a time, as nowadays done in the atmospheric sciences.

The global (“off-line”) methodology can be conveniently summarized by using the notation provided in [16], which considers that the time span of interest for the nonlinear transient system being analyzed is partitioned into $(N_t - 1)$ intervals. At every time instance ν , the quantities J_α^ν , and J_r^ν are used to denote the ordered set of integers $J_\alpha^\nu \triangleq \{1, \dots, N_\alpha^\nu\}$ and, respectively, $J_r^\nu \triangleq \{1, \dots, N_r^\nu\}$, where N_α^ν and N_r^ν denote the

numbers of distinct system parameters and distinct responses, respectively, at time instance ν . The quantity J_t denotes the set $J_t \triangleq \{1, \dots, N_t\}$.

At every time instance ν , the imprecisely known system parameters are considered to be variates with mean values denoted as $(\alpha_i^\nu)^0$. The covariances between two parameters α_i^ν and α_j^μ , at two time instances μ and ν , is denoted as $\text{cov}(\alpha_i^\nu, \alpha_j^\mu)$. These local covariances are used to construct the following symmetric global parameter covariance matrix:

$$\mathbf{C}_\alpha \triangleq \begin{pmatrix} \mathbf{C}_\alpha^{11} & \mathbf{C}_\alpha^{12} & \dots & \dots \\ \mathbf{C}_\alpha^{21} & \mathbf{C}_\alpha^{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \mathbf{C}_\alpha^{N_t N_t} \end{pmatrix}, \mathbf{C}_\alpha^{\mu\nu} \triangleq [\text{cov}(\alpha_i^\nu, \alpha_j^\mu)]; i, j = 1, \dots, J_\alpha^\nu; \nu, \mu = 1, \dots, J_t. \quad (18)$$

Similarly, the imprecisely known measured responses are characterized by local mean values denoted as r_i^ν , at a time instance ν . The covariances between two responses measured at time instances μ and ν are denoted as $\text{cov}(r_i^\nu, r_j^\mu)$. These covariances are used to construct the symmetric global matrix, $\mathbf{C}_m$, for measured responses, defined below:

$$\mathbf{C}_m \triangleq \begin{pmatrix} \mathbf{C}_m^{11} & \mathbf{C}_m^{12} & \dots & \dots \\ \mathbf{C}_m^{21} & \mathbf{C}_m^{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \mathbf{C}_m^{N_t N_t} \end{pmatrix}; \mathbf{C}_m^{\mu\nu} \triangleq [\text{cov}(r_i^\nu, r_j^\mu)]; i, j = 1, \dots, J_r^\nu; \nu, \mu = 1, \dots, J_t. \quad (19)$$

In the most general case, the measured responses may be correlated to the parameters through local response-parameter covariances, denoted as $\text{cov}(\alpha_i^\nu, r_j^\mu)$, between a parameter α_i^ν and a response r_j^μ , which are used to construct the global rectangular parameter-response covariance matrix, $\mathbf{C}_{ar}$, having the following form:

$$\mathbf{C}_{ar} \triangleq \begin{pmatrix} \mathbf{C}_{ar}^{11} & \mathbf{C}_{ar}^{12} & \dots & \dots \\ \mathbf{C}_{ar}^{21} & \mathbf{C}_{ar}^{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \mathbf{C}_{ar}^{N_t N_t} \end{pmatrix}; \mathbf{C}_{ar}^{\mu\nu} \triangleq [\text{cov}(\alpha_i^\nu, r_j^\mu)]; i = 1, \dots, J_\alpha^\nu; j = 1, \dots, J_r^\nu; \nu, \mu = 1, \dots, J_t. \quad (20)$$

At any given time instance ν , a response r_i^ν can be a function of not only the system parameters at time instance ν , but also of the system parameters at previous time instances μ , $1 \leq \mu \leq \nu$; this means that $\mathbf{r}^\nu = \mathbf{R}^\nu(\mathbf{p}^\nu)$, where $\mathbf{p}^\nu \triangleq (\mathbf{a}^1, \dots, \mathbf{a}^\mu, \dots, \mathbf{a}^\nu)$. As in the previous Subsection, the computed response is considered to depend linearly on the model parameters, i.e., the computed response is linearized via a functional Taylor-series expansion around the nominal values, $\mathbf{p}_0^\nu \triangleq ((\mathbf{a}^0)^1, \dots, (\mathbf{a}^0)^\mu, \dots, (\mathbf{a}^0)^\nu)$, of the parameters $\mathbf{p}^\nu$, as follows:

$$\mathbf{r}^\nu \triangleq \mathbf{R}^\nu(\mathbf{p}^\nu) = \mathbf{R}^\nu(\mathbf{p}_0^\nu) + \sum_{\mu=1}^{\nu} \mathbf{S}^{\nu\mu}(\mathbf{p}_0^\mu) [\mathbf{a}^\mu - (\mathbf{a}^0)^\mu] + \text{higher order terms}, \nu \in J_t, \quad (21)$$

where $\mathbf{R}^\nu(\mathbf{p}_0^\nu)$ denotes the vector of computed responses at a time instance ν , at the nominal parameter values $\mathbf{p}_0^\nu$, while $\mathbf{S}^{\nu\mu}(\mathbf{p}_0^\mu)$, $1 \leq \mu \leq \nu$, represents the $(J_r^\nu \times J_\alpha^\mu)$ -dimensional matrix containing the first-order Gateaux-derivatives of the computed responses with respect to the parameters, defined as follows:

$$\mathbf{S}^{\nu\mu}(\mathbf{p}_0^\mu) \triangleq \begin{pmatrix} S_{II}^{\nu\mu} & \cdots & S_{IN}^{\nu\mu} \\ \vdots & s_{in}^{\nu\mu} & \vdots \\ S_{II}^{\nu\mu} & \cdots & S_{IN}^{\nu\mu} \end{pmatrix} \triangleq \begin{pmatrix} \frac{\partial R_l^n(\mathbf{p}_0^\mu)}{\partial \alpha_l^\mu} & \cdots & \frac{\partial R_l^n(\mathbf{p}_0^\mu)}{\partial \alpha_N^\mu} \\ \vdots & \frac{\partial R_l^\nu}{\partial \alpha_n^\mu} & \vdots \\ \frac{\partial R_l^\nu(\mathbf{p}_0^\mu)}{\partial \alpha_l^\mu} & \cdots & \frac{\partial R_l^\nu(\mathbf{p}_0^\mu)}{\partial \alpha_N^\mu} \end{pmatrix}, \quad l \leq \mu \leq \nu. \quad (22)$$

Since the response $\mathbf{R}^\nu(\mathbf{p}_0^\nu)$ at time instance ν can depend only on parameters $(\mathbf{a}^0)^\mu$ which appear up to the current time instance ν , it follows that $\mathbf{S}^{\nu\mu} = \mathbf{0}$ when $\mu > \nu$, implying that non-zero terms in the expansion shown in Equation (21) can only occur in the range $l \leq \mu \leq \nu$. The linear model expressed by Equation (21) can be written in the form

$$\mathbf{r} = \mathbf{R}(\mathbf{a}^0) + \mathbf{S}(\mathbf{a} - \mathbf{a}^0) + \text{higher order terms}; \quad (23)$$

where

$$\begin{aligned} \mathbf{a} &\triangleq (\mathbf{a}^I, \dots, \mathbf{a}^\mu, \dots, \mathbf{a}^{N_t}); \quad \mathbf{a}^\nu = \{\alpha_n^\nu \mid n \in J_\alpha^\nu\}, \quad \nu \in J_t; \\ \mathbf{r} &\triangleq (\mathbf{r}^I, \dots, \mathbf{r}^\mu, \dots, \mathbf{r}^{N_t}); \quad \mathbf{r}^\nu = \{r_i^\nu \mid i \in J_r^\nu\}, \quad \nu \in J_t; \\ \mathbf{R}(\mathbf{a}^0) &\triangleq (\mathbf{R}^I, \dots, \mathbf{R}^\mu, \dots, \mathbf{R}^{N_t}), \quad \nu \in J_t; \end{aligned} \quad (24)$$

$$\mathbf{S} \triangleq \begin{pmatrix} \mathbf{S}^{II} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{S}^{N_t I} & \cdots & \mathbf{S}^{N_t N_t} \end{pmatrix}. \quad (25)$$

The information provided in Equations (18)–(25) is now used to construct the following Lagrangian functional $P(\mathbf{z})$ to be minimized:

$$P(\mathbf{z}) \triangleq Q(\mathbf{z}) + 2\lambda^\dagger [\mathbf{Z}(\mathbf{a}^0)\mathbf{z} + \mathbf{d}] = \min, \quad \text{at } \mathbf{z} = \mathbf{z}^{be}, \quad (26)$$

where $\lambda = (\lambda^I, \dots, \lambda^\nu, \dots, \lambda^{N_t})$ denotes the corresponding vector of Lagrange multipliers and where

$$Q(\mathbf{z}) \triangleq \mathbf{z}^\dagger \mathbf{C}^{-1} \mathbf{z}; \quad \mathbf{z} \triangleq \begin{pmatrix} \mathbf{a} - \mathbf{a}^0 \\ \mathbf{r} - \mathbf{r}_m \end{pmatrix}; \quad \mathbf{C} \triangleq \begin{pmatrix} \mathbf{C}_\alpha & \mathbf{C}_{\alpha r} \\ \mathbf{C}_{r\alpha} & \mathbf{C}_m \end{pmatrix}; \quad \mathbf{a}^0 \triangleq [(\mathbf{a}^0)^I, \dots, (\mathbf{a}^0)^\mu, \dots, (\mathbf{a}^0)^{N_t}]^\dagger; \quad (27)$$

$$\mathbf{d} \triangleq \mathbf{R}(\mathbf{a}^0) - \mathbf{r}_m; \quad \mathbf{r}_m \triangleq (\mathbf{r}_m^I, \dots, \mathbf{r}_m^\mu, \dots, \mathbf{r}_m^{N_t}); \quad \mathbf{Z} \triangleq (\mathbf{S}, \mathbf{U}); \quad \mathbf{U} \triangleq \begin{pmatrix} -\mathbf{I}^{II} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & -\mathbf{I}^{N_t N_t} \end{pmatrix}. \quad (28)$$

In Equation (28), the quantities $\mathbf{I}^{jj}$, $j = 1, \dots, N_t$ denote the identity matrices of corresponding dimensions.

Setting $\partial P(\mathbf{z})/\partial \mathbf{z} = 0$ and solving the resulting equations yields the following results, which have the same formal structure as shown in Equations (13)–(17):

(i) calibrated best-estimate parameter values:

$$\begin{pmatrix} (\mathbf{a}^{be})^I \\ \vdots \\ (\mathbf{a}^{be})^{N_t} \end{pmatrix} = \begin{pmatrix} (\mathbf{a}^0)^I \\ \vdots \\ (\mathbf{a}^0)^{N_t} \end{pmatrix} + \begin{pmatrix} \mathbf{C}_{ar}^{II} - \mathbf{C}_\alpha^{II} (\mathbf{s}^\dagger)^{II} & \dots & \mathbf{C}_{ar}^{IN_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_\alpha^{I\rho} (\mathbf{s}^\dagger)^{N_t\rho} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{ar}^{N_t I} - \mathbf{C}_\alpha^{N_t I} (\mathbf{s}^\dagger)^{II} & \dots & \mathbf{C}_{ar}^{N_t N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_\alpha^{N_t \rho} (\mathbf{s}^\dagger)^{N_t \rho} \end{pmatrix} \times \begin{pmatrix} \sum_{\eta=1}^{N_t} \mathbf{K}_d^{I\eta} \mathbf{d}^\eta \\ \vdots \\ \sum_{\eta=1}^{N_t} \mathbf{K}_d^{N_t \eta} \mathbf{d}^\eta \end{pmatrix}, \quad (29)$$

where $\mathbf{K}_d^{v\eta}$ denotes the corresponding (v, η) -element of the block-matrix $\mathbf{C}_d^{-1}$, where:

$$\begin{aligned} \mathbf{C}_d &\triangleq \begin{pmatrix} \mathbf{C}_d^{II} & \dots & \mathbf{C}_d^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_d^{N_t I} & \dots & \mathbf{C}_d^{N_t N_t} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{rc}^{II} + \mathbf{C}_m^{II} & \dots & \mathbf{C}_{rc}^{IN_t} + \mathbf{C}_m^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{rc}^{N_t I} + \mathbf{C}_m^{N_t I} & \dots & \mathbf{C}_{rc}^{N_t N_t} + \mathbf{C}_m^{N_t N_t} \end{pmatrix} \\ &- \begin{pmatrix} \mathbf{C}_{ra}^{II} (\mathbf{s}^\dagger)^{II} + \mathbf{s}^{II} \mathbf{C}_{ar}^{II} & \dots & \mathbf{s}^{II} \mathbf{C}_{ar}^{IN_t} + \sum_{\rho=1}^{N_t} \mathbf{C}_{ra}^{I\rho} (\mathbf{s}^\dagger)^{N_t \rho} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{ra}^{N_t I} (\mathbf{s}^\dagger)^{II} + \sum_{\rho=1}^{N_t} \mathbf{s}^{N_t \rho} \mathbf{C}_{ar}^{I\rho} & \dots & \sum_{\rho=1}^{N_t} [\mathbf{C}_{ra}^{N_t \rho} (\mathbf{s}^\dagger)^{N_t \rho} + \mathbf{s}^{N_t \rho} \mathbf{C}_{ar}^{N_t \rho}] \end{pmatrix}, \end{aligned} \quad (30)$$

and where the covariance matrix of the computed responses, $\mathbf{C}_{rc}$, is defined as follows:

$$\mathbf{C}_{rc} \triangleq \begin{pmatrix} \mathbf{C}_{rc}^{II} & \dots & \mathbf{C}_{rc}^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{rc}^{N_t I} & \dots & \mathbf{C}_{rc}^{N_t N_t} \end{pmatrix} = \mathbf{S} \mathbf{C}_\alpha \mathbf{S}^\dagger; \mathbf{C}_{rc}^{v\mu} = \sum_{\eta=1}^v \sum_{\rho=1}^\mu \mathbf{s}^{v\eta} \mathbf{C}_\alpha^{\eta\rho} (\mathbf{s}^{\mu\rho})^\dagger = (\mathbf{C}_{rc}^{\mu\nu})^\dagger; v, \mu \in J_t. \quad (31)$$

Written in component form, Equation (29) indicates that the vector $(\mathbf{a}^{be})^v$, representing the calibrated best-estimates for the system parameters at a specific time instance v , takes on the following expression:

$$(\mathbf{a}^{be})^v = (\mathbf{a}^0)^v + \sum_{\mu=1}^{N_t} \left\{ \left[\mathbf{C}_{ar}^{v\mu} - \sum_{\rho=1}^\mu \mathbf{C}_\alpha^{v\rho} (\mathbf{s}^\dagger)^{\mu\rho} \right] \left[\sum_{\eta=1}^{N_t} \mathbf{K}_d^{\mu\eta} \mathbf{d}^\eta \right] \right\}, \quad v \in J_t \quad (32)$$

(ii) The calibrated best-estimate covariance matrix, $\mathbf{C}_\alpha^{be}$, corresponding to the calibrated best-estimates system parameters:

$$\begin{aligned} \mathbf{C}_\alpha^{be} &\triangleq \begin{pmatrix} (\mathbf{C}_\alpha^{be})^{II} & \dots & (\mathbf{C}_\alpha^{be})^{IN_t} \\ \vdots & \ddots & \vdots \\ (\mathbf{C}_\alpha^{be})^{N_t I} & \dots & (\mathbf{C}_\alpha^{be})^{N_t N_t} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_\alpha^{II} & \dots & \mathbf{C}_\alpha^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_\alpha^{N_t I} & \dots & \mathbf{C}_\alpha^{N_t N_t} \end{pmatrix} \\ &- \begin{pmatrix} \mathbf{C}_{ad}^{II} & \dots & \mathbf{C}_{ad}^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{ad}^{N_t I} & \dots & \mathbf{C}_{ad}^{N_t N_t} \end{pmatrix} \begin{pmatrix} \mathbf{K}_d^{II} & \dots & \mathbf{K}_d^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_d^{N_t I} & \dots & \mathbf{K}_d^{N_t N_t} \end{pmatrix} \begin{pmatrix} \mathbf{C}_{ad}^{II} & \dots & \mathbf{C}_{ad}^{IN_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{ad}^{N_t I} & \dots & \mathbf{C}_{ad}^{N_t N_t} \end{pmatrix}^\dagger \end{aligned} \quad (33)$$

where

$$\begin{pmatrix} \mathbf{C}_{ad}^{1l} & \cdots & \mathbf{C}_{ad}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{ad}^{N_t l} & \cdots & \mathbf{C}_{ad}^{N_t N_t} \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{C}_{ar}^{1l} - \mathbf{C}_{\alpha}^{1l} (\mathbf{s}^\dagger)^{1l} & \cdots & \mathbf{C}_{ar}^{1N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_{\alpha}^{1\rho} (\mathbf{s}^\dagger)^{N_t \rho} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{ar}^{N_t l} - \mathbf{C}_{\alpha}^{N_t l} (\mathbf{s}^\dagger)^{1l} & \cdots & \mathbf{C}_{ar}^{N_t N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_{\alpha}^{N_t \rho} (\mathbf{s}^\dagger)^{N_t \rho} \end{pmatrix}. \quad (34)$$

The block-matrix expression in Equation (33) can be written in component form, for the calibrated best-estimate parameter covariance matrix $(\mathbf{C}_{\alpha}^{be})^{\nu\mu}$ between two (distinct or not) time instances $\nu, \mu \in J_t$, as follows

$$(\mathbf{C}_{\alpha}^{be})^{\nu\mu} = \mathbf{C}_{\alpha}^{\nu\mu} - \sum_{\eta=1}^{N_t} \sum_{\rho=1}^{N_t} \left[\mathbf{C}_{ar}^{\nu\rho} - \sum_{\pi=1}^{\rho} \mathbf{C}_{\alpha}^{\nu\pi} (\mathbf{s}^\dagger)^{\rho\pi} \right] \mathbf{K}_d^{\rho\eta} \left[\mathbf{C}_{ra}^{\eta\mu} - \sum_{\pi=1}^{\eta} \mathbf{s}^{\eta\pi} \mathbf{C}_{\alpha}^{\pi\mu} \right]. \quad (35)$$

(iii) The vector of calibrated best-estimate system responses at all time instances $\nu \in J_t$:

$$\begin{pmatrix} (\mathbf{r}^{be})^l \\ \vdots \\ (\mathbf{r}^{be})^{N_t} \end{pmatrix} = \begin{pmatrix} (\mathbf{r}_m)^l \\ \vdots \\ (\mathbf{r}_m)^{N_t} \end{pmatrix} + \begin{pmatrix} \mathbf{C}_m^{1l} - \mathbf{C}_{ra}^{1l} (\mathbf{s}^\dagger)^{1l} & \cdots & \mathbf{C}_m^{1N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_{ra}^{1\rho} (\mathbf{s}^\dagger)^{N_t \rho} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_m^{N_t l} - \mathbf{C}_{ra}^{N_t l} (\mathbf{s}^\dagger)^{1l} & \cdots & \mathbf{C}_m^{N_t N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_{ra}^{N_t \rho} (\mathbf{s}^\dagger)^{N_t \rho} \end{pmatrix} \begin{pmatrix} \sum_{\eta=1}^{N_t} \mathbf{K}_d^{1\eta} \mathbf{d}^\eta \\ \vdots \\ \sum_{\eta=1}^{N_t} \mathbf{K}_d^{N_t \eta} \mathbf{d}^\eta \end{pmatrix}. \quad (36)$$

Written in component form, Equation (36) provides the following expression for the vector $(\mathbf{r}^{be})^\nu$, of calibrated best-estimates for the responses at a specific time instance ν :

$$(\mathbf{r}^{be})^\nu = (\mathbf{r}_m)^\nu + \sum_{\mu=1}^{N_t} \left\{ \left[\mathbf{C}_x^{\nu\mu} - \sum_{\rho=1}^{\mu} \mathbf{C}_{ra}^{\nu\rho} (\mathbf{s}^\dagger)^{\mu\rho} \right] \left[\sum_{\eta=1}^{N_t} \mathbf{K}_d^{\mu\eta} \mathbf{d}^\eta \right] \right\}, \quad \nu \in J_t. \quad (37)$$

(iv) The expression of the calibrated best-estimate covariance block-matrix, $\mathbf{C}_r^{be}$, for the best-estimate responses:

$$\begin{aligned} \mathbf{C}_r^{be} &\triangleq \begin{pmatrix} (\mathbf{C}_r^{be})^{1l} & \cdots & (\mathbf{C}_r^{be})^{1N_t} \\ \vdots & \ddots & \vdots \\ (\mathbf{C}_r^{be})^{N_t l} & \cdots & (\mathbf{C}_r^{be})^{N_t N_t} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_m^{1l} & \cdots & \mathbf{C}_m^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_m^{N_t l} & \cdots & \mathbf{C}_m^{N_t N_t} \end{pmatrix} \\ &- \begin{pmatrix} \mathbf{C}_{rd}^{1l} & \cdots & \mathbf{C}_{rd}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{rd}^{N_t l} & \cdots & \mathbf{C}_{rd}^{N_t N_t} \end{pmatrix} \begin{pmatrix} \mathbf{K}_d^{1l} & \cdots & \mathbf{K}_d^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_d^{N_t l} & \cdots & \mathbf{K}_d^{N_t N_t} \end{pmatrix} \begin{pmatrix} \mathbf{C}_{rd}^{1l} & \cdots & \mathbf{C}_{rd}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{rd}^{N_t l} & \cdots & \mathbf{C}_{rd}^{N_t N_t} \end{pmatrix}^\dagger, \end{aligned} \quad (38)$$

where

$$\begin{pmatrix} \mathbf{C}_{rd}^{1l} & \cdots & \mathbf{C}_{rd}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{rd}^{N_t l} & \cdots & \mathbf{C}_{rd}^{N_t N_t} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_m^{1l} - \mathbf{C}_{ra}^{1l} (\mathbf{s}^\dagger)^{1l} & \cdots & \mathbf{C}_m^{1N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_{ra}^{1\rho} (\mathbf{s}^\dagger)^{N_t \rho} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_m^{N_t l} - \mathbf{C}_{ra}^{N_t l} (\mathbf{s}^\dagger)^{1l} & \cdots & \mathbf{C}_m^{N_t N_t} - \sum_{\rho=1}^{N_t} \mathbf{C}_{ra}^{N_t \rho} (\mathbf{s}^\dagger)^{N_t \rho} \end{pmatrix}. \quad (39)$$

The block-matrix expression given in Equation (38) can be written in component form, where each of the calibrated best-estimate parameter covariance matrix $(\mathbf{C}_r^{be})^{\nu\mu}$ between two (distinct or not) time instances $\nu, \mu \in J_t$ has the following form:

$$(\mathbf{C}_{r\alpha}^{be})^{\nu\mu} = \mathbf{C}_m^{\nu\mu} - \sum_{\eta=1}^{N_t} \sum_{\rho=1}^{N_t} \left[\mathbf{C}_m^{\nu\rho} - \sum_{\pi=1}^{\rho} \mathbf{C}_{r\alpha}^{\nu\pi} (\mathbf{S}^\dagger)^{\rho\pi} \right] \mathbf{K}_d^{\rho\eta} \left[\mathbf{C}_m^{\eta\mu} - \sum_{\pi=1}^{\eta} \mathbf{S}^{\eta\pi} \mathbf{C}_{\alpha r}^{\pi\mu} \right]. \quad (40)$$

(v) The best-estimate response-parameter covariance block-matrix $\mathbf{C}_{\alpha r}^{be}$:

$$\begin{aligned} \mathbf{C}_{r\alpha}^{be} &\triangleq \begin{pmatrix} (\mathbf{C}_{r\alpha}^{be})^{11} & \cdots & (\mathbf{C}_{r\alpha}^{be})^{1N_t} \\ \vdots & \ddots & \vdots \\ (\mathbf{C}_{r\alpha}^{be})^{N_t 1} & \cdots & (\mathbf{C}_{r\alpha}^{be})^{N_t N_t} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{r\alpha}^{11} & \cdots & \mathbf{C}_{r\alpha}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{r\alpha}^{N_t 1} & \cdots & \mathbf{C}_{r\alpha}^{N_t N_t} \end{pmatrix} \\ &- \begin{pmatrix} \mathbf{C}_{rd}^{11} & \cdots & \mathbf{C}_{rd}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{rd}^{N_t 1} & \cdots & \mathbf{C}_{rd}^{N_t N_t} \end{pmatrix} \begin{pmatrix} \mathbf{K}_d^{11} & \cdots & \mathbf{K}_d^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_d^{N_t 1} & \cdots & \mathbf{K}_d^{N_t N_t} \end{pmatrix} \begin{pmatrix} \mathbf{C}_{\alpha d}^{11} & \cdots & \mathbf{C}_{\alpha d}^{1N_t} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{\alpha d}^{N_t 1} & \cdots & \mathbf{C}_{\alpha d}^{N_t N_t} \end{pmatrix}^\dagger. \end{aligned} \quad (41)$$

Each of the calibrated best-estimate parameter-response covariance matrices $(\mathbf{C}_{r\alpha}^{be})^{\nu\mu}$ for $\nu, \mu \in J_t$, which appear as block-matrix components of the matrix $\mathbf{C}_{r\alpha}^{be}$ in Equation (41) has the following expression:

$$(\mathbf{C}_{r\alpha}^{be})^{\nu\mu} = \mathbf{C}_{r\alpha}^{\nu\mu} - \sum_{\eta=1}^{N_t} \sum_{\rho=1}^{N_t} \left[\mathbf{C}_x^{\nu\rho} - \sum_{\pi=1}^{\rho} \mathbf{C}_{r\alpha}^{\nu\pi} (\mathbf{S}^\dagger)^{\rho\pi} \right] \mathbf{K}_d^{\rho\eta} \left[\mathbf{C}_{\alpha r}^{\eta\mu} - \sum_{\pi=1}^{\eta} \mathbf{S}^{\eta\pi} \mathbf{C}_{\alpha}^{\pi\mu} \right]. \quad (42)$$

Computing the calibrated best-estimate quantities by using Equations (32), (35), (37), (40) and (42) is definitely more advantageous in terms of storage requirements than the direct computations of the corresponding full block-matrices. The largest requirement of computational resources arises when inverting the matrix $\mathbf{C}_d$. In view of Equation (30), it is important to note that the inverse matrix, $\mathbf{C}_d^{-1}$, incorporates simultaneously all of the available information about the system parameters and responses at all time instances (i.e., $\nu = 1, 2, \dots, N_t$). In other words, at any time instance ν , $\mathbf{C}_d^{-1}$ incorporates information not only from time instances prior to, and at, ν (i.e., information regarding the “past” and “present” states of the system) but also from time instances posterior to ν (i.e., information about the “future” states of the system). Therefore, at any specified time instance ν , the calibrated best-estimates parameters $(\mathbf{a}^{be})^\nu$ and responses $\mathbf{r}(\mathbf{a}^{be}) \triangleq \mathbf{r}^{be}$ together with the corresponding calibrated best-estimate covariance matrices $(\mathbf{C}_{\alpha}^{be})^{\nu\mu}$, $(\mathbf{C}_r^{be})^{\nu\mu}$, and $(\mathbf{C}_{\alpha r}^{be})^{\nu\mu}$ will also incorporate automatically, through the matrix $\mathbf{C}_d^{-1}$, all of the available information about the system parameters and responses at all time instances, i.e., $(\nu = 1, 2, \dots, N_t)$.

Note that the mathematical formalism underlying the time-dependent data-adjustment methodology is written in terms of absolute sensitivities and covariances, in contrast to the data adjustment formalism for time-independent systems, which is written in terms of relative sensitivities and covariances. Nevertheless, the two formalisms produce end-results which look formally quite similar.

The actual application of the predictive modeling results presented in Equations (32), (35), (37), (40) and (42) to a physical system characterized by nominal values and uncertainties for model parameters together with the computed and measured responses is straightforward, in principle, although it can become computationally very demanding regarding both data handling and computational speed.

The minimum value, $\mathcal{Q}_{min}$, of the functional $\mathcal{Q}(\mathbf{z})$ has the following expression:

$$\mathcal{Q}_{min} \triangleq \mathcal{Q}(\mathbf{z}^{be}) = \mathbf{d}^\dagger \left[\mathbf{C}_d(\mathbf{a}^0) \right]^{-1} \mathbf{d} = \chi^2. \quad (43)$$

As the above expression indicates, the minimal value, $Q_{min} \triangleq Q(\mathbf{z}^{be})$, represents the square of the length of the vector $\mathbf{d}$, measuring (in the corresponding metric) the deviations between the experimental and nominally computed responses. The quantity $Q_{min} \triangleq Q(\mathbf{z}^{be})$ can be evaluated directly from the given data (i.e., given parameters and responses, together with their original uncertainties) after having inverted the deviation-vector covariance matrix $\mathbf{C}_d(\mathbf{a}^0)$. It is also very important to note that $Q_{min} \triangleq Q(\mathbf{z}^{be})$ is independent of calibrating (or adjusting) the original data. As the dimension of $\mathbf{d}$ indicates, the number of degrees of freedom characteristic of the calibration under consideration is equal to the number of experimental responses. In the extreme case of absence of experimental responses, no actual calibration takes place since $\mathbf{d} = \mathbf{R}(\mathbf{a}^0)$, so that the best-estimate parameter values are just the original nominal values, i.e., $(\mathbf{a}^{be})^k = (\mathbf{a}^0)^k$. An actual calibration (adjustment) occurs only when including at least one experimental response.

In turn, the variate Q_{min} follows a χ^2 -distribution with n degrees of freedom, where n denotes the total number of experimental responses considered in the calibration (adjustment) procedure. The quantity Q_{min} is the “ χ^2 of the calibration (adjustment) at hand” and can be used as an indicator of the agreement between the computed and measured responses, measuring essentially the consistency of the measured responses with the model parameters. A similar quantity exists for the time-independent data adjustment procedure presented previously, in Section 2.1.

In order to facilitate a comparison with the time-dependent data-assimilation methods used in the atmospheric sciences (which are summarized in Section 2.3, below), which customarily involve just two consecutive time-steps, Cacuci [16] has obtained explicit formulas by simplifying the expressions presented in Equations (32), (35), (37), (40) and (42) by including only two consecutive time-steps, namely $\nu = k - l, k$; $k = (l, 2, \dots, N_t)$, and by presenting the explicit form of the inverse matrix $\mathbf{K}_d^{\nu\eta}$. The simplified expressions thus obtained are as follows:

- (i) The expressions for the calibrated best-estimate parameter values take on the following particular form of Equation (32) at time node k :

$$(\mathbf{a}^{be})^k = (\mathbf{a}^0)^k + \sum_{\mu=k-1}^k \left\{ \left[\mathbf{C}_{ar}^{k\mu} - \sum_{\rho=1}^{\mu} \mathbf{C}_{ar}^{k\rho} (\mathbf{s}^\dagger)^{\mu\rho} \right] \left[\sum_{\eta=k-1}^k \mathbf{K}_d^{\mu\eta} \mathbf{d}^\eta \right] \right\}, \quad (44)$$

where

$$\begin{aligned} \mathbf{K}_d^{k-l,k-l} &= \left[\mathbf{C}_d^{k-l,k-l} - \mathbf{C}_d^{k-l,k} (\mathbf{C}_d^{k,k})^{-1} \mathbf{C}_d^{k,k-l} \right]^{-1} \\ &= (\mathbf{C}_d^{k-l,k-l})^{-1} + (\mathbf{C}_d^{k-l,k-l})^{-1} \mathbf{C}_d^{k-l,k} \mathbf{K}_d^{k,k} \mathbf{C}_d^{k,k-l} (\mathbf{C}_d^{k-l,k-l})^{-1}, \end{aligned} \quad (45)$$

$$\begin{aligned} \mathbf{K}_d^{k-l,k} &= -(\mathbf{C}_d^{k-l,k-l})^{-1} \mathbf{C}_d^{k-l,k} \left[\mathbf{C}_d^{k,k} - \mathbf{C}_d^{k,k-l} (\mathbf{C}_d^{k-l,k-l})^{-1} \mathbf{C}_d^{k-l,k} \right]^{-1} \\ &= -(\mathbf{C}_d^{k-l,k-l})^{-1} \mathbf{C}_d^{k-l,k} \mathbf{K}_d^{k,k}, \end{aligned} \quad (46)$$

$$\begin{aligned} \mathbf{K}_d^{k,k} &= \left[\mathbf{C}_d^{k,k} - \mathbf{C}_d^{k,k-l} (\mathbf{C}_d^{k-l,k-l})^{-1} \mathbf{C}_d^{k-l,k} \right]^{-1} \\ &= (\mathbf{C}_d^{k,k})^{-1} + (\mathbf{C}_d^{k,k})^{-1} \mathbf{C}_d^{k,k-l} \mathbf{K}_d^{k-l,k-l} \mathbf{C}_d^{k-l,k} (\mathbf{C}_d^{k,k})^{-1}, \end{aligned} \quad (47)$$

$$\begin{aligned}\mathbf{K}_d^{k,k-l} &= -(\mathbf{C}_d^{k,k})^{-1} \mathbf{C}_d^{k,k-l} \left[\mathbf{C}_d^{k-l,k-l} - \mathbf{C}_d^{k-l,k} (\mathbf{C}_d^{k,k})^{-1} \mathbf{C}_d^{k,k-l} \right]^{-1} \\ &= -(\mathbf{C}_d^{k,k})^{-1} \mathbf{C}_d^{k,k-l} \mathbf{K}_d^{k-l,k-l}.\end{aligned}\quad (48)$$

- (ii) The components $(\mathbf{C}_\alpha^{be})^{\nu\mu}$, $\nu, \mu = k-l, k$, of the calibrated best-estimate covariance matrix, $\mathbf{C}_\alpha^{be}$, have the following particular form of Equation (35):

$$\begin{aligned}(\mathbf{C}_\alpha^{be})^{\nu\mu} &= \mathbf{C}_\alpha^{\nu\mu} - \sum_{\eta=k-l}^k \sum_{\rho=k-l}^k \left[\mathbf{C}_{\alpha r}^{\nu\rho} - \sum_{\pi=l}^{\rho} \mathbf{C}_{\alpha}^{\nu\pi} (\mathbf{S}^\dagger)^{\rho\pi} \right] \mathbf{K}_d^{\rho\eta} \left[\mathbf{C}_{r\alpha}^{\eta\mu} - \sum_{\pi=l}^{\eta} \mathbf{S}^{\eta\pi} \mathbf{C}_\alpha^{\pi\mu} \right] \\ \text{for } \nu &= k-l, k; \text{ and } \mu = k-l, k.\end{aligned}\quad (49)$$

- (iii) The vector $(\mathbf{r}^{be})^k$, representing the calibrated best-estimates for the system parameters at a time instance k , takes on the following particular form of Equation (37):

$$(\mathbf{r}^{be})^k = (\mathbf{r}_m)^k + \sum_{\mu=k-l}^k \left\{ \left[\mathbf{C}_m^{k\mu} - \sum_{\rho=l}^{\mu} \mathbf{C}_{r\alpha}^{k\rho} (\mathbf{S}^\dagger)^{\mu\rho} \right] \left[\sum_{\eta=k-l}^k \mathbf{K}_d^{\mu\eta} \mathbf{d}^\eta \right] \right\}.\quad (50)$$

- (iv) The components $(\mathbf{C}_r^{be})^{\nu\mu}$ of the calibrated best-estimate covariance block-matrix $\mathbf{C}_r^{be}$ for the best-estimate responses takes on the following particular form of Equation (38) for $\nu, \mu = k-l, k$:

$$(\mathbf{C}_r^{be})^{\nu\mu} = \mathbf{C}_m^{\nu\mu} - \sum_{\eta=k-l}^k \sum_{\rho=k-l}^k \left[\mathbf{C}_x^{\nu\rho} - \sum_{\pi=k-l}^{\rho} \mathbf{C}_{r\alpha}^{\nu\pi} (\mathbf{S}^\dagger)^{\rho\pi} \right] \mathbf{K}_d^{\rho\eta} \left[\mathbf{C}_m^{\eta\mu} - \sum_{\pi=k-l}^{\eta} \mathbf{S}^{\eta\pi} \mathbf{C}_{r\alpha}^{\pi\mu} \right].\quad (51)$$

- (v) The matrix-valued components $(\mathbf{C}_{\alpha r}^{be})^{\nu\mu}$, $\nu, \mu = k-l, k$, of the best-estimate response-parameter covariance matrix $\mathbf{C}_{\alpha r}^{be}$ reduce to the following particular forms of Equation (42):

$$(\mathbf{C}_{\alpha r}^{be})^{\nu\mu} = \mathbf{C}_{r\alpha}^{\nu\mu} - \sum_{\eta=k-l}^k \sum_{\rho=k-l}^k \left[\mathbf{C}_m^{\nu\rho} - \sum_{\pi=k-l}^{\rho} \mathbf{C}_{r\alpha}^{\nu\pi} (\mathbf{S}^\dagger)^{\rho\pi} \right] \mathbf{K}_d^{\rho\eta} \left[\mathbf{C}_{r\alpha}^{\eta\mu} - \sum_{\pi=k-l}^{\eta} \mathbf{S}^{\eta\pi} \mathbf{C}_\alpha^{\pi\mu} \right].\quad (52)$$

2.3. Least-Squares Based Variational Data Assimilation

The predecessors to data-assimilation methods were called “objective analyses” [24], to distinguish them from the “subjective analyses” in which data was manipulated according to the opinion of experts. Subsequently, the so-called “nudging” or “Newtonian relaxation” methods were introduced [25] to drive the model variables towards the observations by adding empirical forcing terms. The variational principle introduced by Sasaki [26,27] provided the basis for the development of “optimal interpolation,” “three-dimensional variational data assimilation,” and “physical space statistical analysis” methods (see, e.g., the books [11–14]), which were developed independently but were shown by Lorenc [28] to be equivalent.

The introduction of Cacuci’s adjoint method [10] to the atmospheric sciences [29–31] has enabled the subsequent development of the modern techniques nowadays called “variational data assimilation” methodology (also called “4D-VAR”; “4D” refers to a framework involving the four independent variables “time plus three space dimensions”). Variational data assimilation aims at determining the model’s evolution “trajectory” which fits “best”—in a “least-square” sense—the observational data over an assimilation time interval by adjusting the initial conditions supplied for advancing (integrating) the forecast model forward in time. The earliest work on developing the 4D-VAR methodology using adjoint functions was by Derber [32,33] and Hoffmann [34], while Navon [35–37] has introduced augmented Lagrangian methods for enforcing the conservation of integral invariants. Most of the meteorological centers around the world have since

implemented 4D-VAR methods in conjunction with comprehensive numerical weather prediction models [38–40].

The so-called strong constraint “variational data assimilation (VDA)” or classical VDA framework assumes that the forecast model perfectly represents the evolution of the actual physical system (atmosphere), and the “best fit” model trajectory is obtained by adjusting only the initial conditions via minimization of a user-defined cost functional, subject to the model equations as strong constraints. However, like any other models of large-scale processes, numerical weather prediction (NWP) models are affected by errors arising from the omission of sub-grid processes; also, the discretization of continuous processes introduces dissipative and dispersion errors. Furthermore, the mathematical modeling of the boundary conditions and forcing terms is imperfect, and most of the physical processes and their interactions in the atmosphere are parameterized. In general, these modeling approximations/errors comprise both systematic (so-called “bias”) and stochastic components which vary in space and time. Customarily, the model error (ME) term is formally introduced as a correction to the time derivatives of model variables and the resulting methodology is called the weak constraint VDA.

Using the notation of [14], the state of the physical system (e.g., atmosphere) can be described by the following evolution equation:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{M}[\mathbf{x}(t)] + \mathbf{T}[\boldsymbol{\eta}(t)], \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad (53)$$

where: (i) the vector $\mathbf{x}(t)$ represents the state of the atmosphere, with an initial condition $\mathbf{x}_0$ at the initial time t_0 ; (ii) the operator $\mathbf{M}[\cdot]$ denotes all the mathematical operations involved in the model; (iii) $\boldsymbol{\eta}(t)$ represents the model errors; (iv) $\mathbf{T}[\cdot]$ represents an operator that maps the space of ME to the space of the model state $\mathbf{x}(t)$. If the model state has an associated model error at every grid point, then the operator $\mathbf{T}[\cdot]$ is identically equal to the unit matrix and the dimension of $\boldsymbol{\eta}(t)$ is the same as the dimension of the model state $\mathbf{x}(t)$. On the other hand, if only some components of the state vector (e.g., the atmosphere in the polar regions) are affected by modeling errors, then the operator $\mathbf{T}[\cdot]$ can be specified in such a way that only those model grid points (in the polar regions) have modeling errors, while (by comparison) the rest of the model states are free of ME's. Some authors formally/artificially write $\boldsymbol{\eta}(t)$ as the sum of a stochastic error (“noise”) component and a “systematic error” component but such a separation is superfluous in practice since both components are anyway estimated based on “expert opinion” rather than based on exact modeling.

The model error is considered to evolve according to the following evolution model:

$$\frac{d\boldsymbol{\eta}}{dt} = \boldsymbol{\Phi}[\boldsymbol{\eta}(t), \mathbf{x}(t)], \quad \boldsymbol{\eta}(t_0) = \boldsymbol{\eta}_0, \quad (54)$$

where the rate of growth or decay of $\boldsymbol{\eta}(t)$ in time is governed by the particular form of the mapping $\boldsymbol{\Phi}$.

Data-assimilation schemes determine the analyzed atmospheric state as an optimal combination of a-priori background information and observational information, by minimizing the following user-defined functional:

$$J[\mathbf{x}(t_0), \boldsymbol{\eta}(t_0)] \triangleq J_o + J_b + J_\eta, \quad (55)$$

where

$$J_b \triangleq \frac{1}{2} [\mathbf{x}_0 - \mathbf{x}^b]^T \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^b], \quad (56)$$

$$J_o \triangleq \frac{1}{2} \sum_{i=0}^n [\mathbf{H}(\mathbf{x}_i) - \mathbf{y}^o(t_i)]^T \mathbf{R}^{-1} [\mathbf{H}(\mathbf{x}_i) - \mathbf{y}^o(t_i)], \quad (57)$$

$$J_\eta \triangleq \frac{1}{2} [\eta_0 - \eta^b]^T \mathbf{Q}^{-1} [\eta_0 - \eta^b]. \quad (58)$$

The components of the functionals J_b , J_o and J_η are defined as follows:

1. The functional J_b quantifies, in a least-square sense, the squared differences between the initial state $\mathbf{x}_0$ and the background state $\mathbf{x}^b$; $\mathbf{B}$ denotes the estimated background covariance matrix. The background state $\mathbf{x}^b$ provides an initial guess for minimization of J .
2. The functional J_o quantifies, in a least-square sense, the squared differences between the observed state and initial state at the time instance t_i ; $\mathbf{R}$ denotes the estimated observation covariance matrix and $\mathbf{H}(\mathbf{x}_i)$ denotes the observation operator at the time instance t_i .
3. The functional J_η quantifies, in a least-square sense, the model errors; $\mathbf{Q}$ denotes the covariance matrix of model errors.

The constrained minimization of the functional $J(\mathbf{x}_0, \boldsymbol{\eta}_0)$ subject to the “hard constraints” represented by Equations (53) and (54) is performed by minimizing the unconstrained augmented Lagrangian functional defined below:

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\eta}, \mathbf{x}^*, \boldsymbol{\eta}^*) = & J(\mathbf{x}_0, \boldsymbol{\eta}_0) + \int_{t_0}^{t_n} \left\langle \mathbf{x}^*(t), \left\{ \frac{d\mathbf{x}(t)}{dt} - \mathbf{M}[\mathbf{x}(t)] - \boldsymbol{\eta}(t) \right\} \right\rangle dt \\ & + \int_{t_0}^{t_n} \left\langle \boldsymbol{\eta}^*(t), \left\{ \frac{d\boldsymbol{\eta}}{dt} - \boldsymbol{\Phi}[\boldsymbol{\eta}(t), \mathbf{x}(t)] \right\} \right\rangle dt, \end{aligned} \quad (59)$$

where $\mathbf{x}^*(t)$ and $\boldsymbol{\eta}^*(t)$ are the Lagrange multiplier vectors corresponding to $\mathbf{x}(t)$ and $\boldsymbol{\eta}(t)$, respectively, while $\langle \cdot, \cdot \rangle$ denotes Euclidean inner product in the three-dimensional physical space. The extrema of $L(\mathbf{x}, \boldsymbol{\eta}, \mathbf{x}^*, \boldsymbol{\eta}^*)$ are the solutions of the following Euler–Lagrange equations:

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{0}, \quad \frac{\partial L}{\partial \boldsymbol{\eta}} = \mathbf{0}, \quad (60)$$

$$\frac{\partial L}{\partial \mathbf{x}^*} = \mathbf{0}, \quad \frac{\partial L}{\partial \boldsymbol{\eta}^*} = \mathbf{0}. \quad (61)$$

Imposing Equation (61) ensures that the equations describing the evolution of model state and ME shown in Equations (53) and (54) are satisfied. Solving Equation (60) yields the following adjoint equations which describe the evolution of the adjoint variables $\mathbf{x}^*(t)$ and $\boldsymbol{\eta}^*(t)$:

$$-\frac{d\mathbf{x}^*(t)}{dt} = \left[\frac{\partial \mathbf{M}}{\partial \mathbf{x}} \right]^T \mathbf{x}^*(t) + \left[\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}} \right]^T \boldsymbol{\eta}^*(t) + \delta(t - t_i) \sum_{i=0}^n \left[\frac{\partial H}{\partial \mathbf{x}} \right]^T \mathbf{R}^{-1} [H(\mathbf{x}_i) - \mathbf{y}^o(t_i)]; \quad \mathbf{x}^*(t_n) = \mathbf{0}, \quad (62)$$

$$-\frac{d\boldsymbol{\eta}^*(t)}{dt} = \left[\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{\eta}} \right]^T \boldsymbol{\eta}^*(t) + \mathbf{x}^*(t); \quad \boldsymbol{\eta}^*(t_n) = \mathbf{0}. \quad (63)$$

As indicated by Equations (62) and (63), the evolution of $\mathbf{x}^*(t)$ is coupled to the evolution of $\boldsymbol{\eta}^*(t)$ via the operator $\boldsymbol{\Phi}$. Backward integration of the adjoint models represented by Equations (62) and (63) from time $t_n \rightarrow t_0$ provides the values of initial adjoint states: $\mathbf{x}^*(t_0)$ and $\boldsymbol{\eta}^*(t_0)$. Furthermore, the optimality conditions expressed by Equations (60) and (61) also provide the following expressions for the partial gradients of the

functional $L(\mathbf{x}, \boldsymbol{\eta}, \mathbf{x}^*, \boldsymbol{\eta}^*)$ or, equivalently, the functional $J(\mathbf{x}_0, \boldsymbol{\eta}_0)$, with respect to the initial model state, $\mathbf{x}_0$, and the initial ME state, $\boldsymbol{\eta}_0$:

$$\nabla_{\mathbf{x}_0} J = \nabla_{\mathbf{x}_0} J_b + \nabla_{\mathbf{x}_0} J_o = \mathbf{B}^{-1}[\mathbf{x}_0 - \mathbf{x}^b] + \mathbf{x}^*(t_0); \quad (64)$$

$$\nabla_{\boldsymbol{\eta}_0} J = \nabla_{\boldsymbol{\eta}_0} J_{\boldsymbol{\eta}} + \nabla_{\boldsymbol{\eta}_0} J_o = \mathbf{Q}^{-1}[\boldsymbol{\eta}_0 - \boldsymbol{\eta}^b] + \boldsymbol{\eta}^*(t_0). \quad (65)$$

If the model's errors are neglected, Equation (65) would not exist, so the amount of computations would be just half of the computations needed when modeling errors are explicitly taken into account.

3. BERRU-PM-2+: Second-Order-Plus MaxEnt Forward and Inverse Predictive Modeling Methodology

Cacuci [15,16] conceived a predictive modeling methodology which uses the maximum entropy (MaxEnt) principle [17] to combine experimental and computational information in the joint phase-space of responses and model parameters in order to provide optimal mean values for model parameters, as well as reduced uncertainties for the optimally predicted responses and parameters. This methodology has been called the “BERRU-PM” methodology, which is an acronym for “Best-Estimate Results with Reduced Uncertainties: Predictive Modeling”. Using the MaxEnt principle [17], the BERRU-PM methodology eliminates the need for introducing and “minimizing” a user-chosen “cost functional quantifying the discrepancies between measurements and computations”, thus yielding results that are free of subjective user-interferences while generalizing and significantly extending the current dynamic data assimilation procedures. The use of the maximum entropy principle enables the BERRU-PM to consider the model as an “imprecisely known” entity, as opposed to being appended by Lagrange multipliers as a “hard constraint” devoid of model errors. Incorporating correlations, including those between the imprecisely known model parameters and computed model responses, the BERRU-PM also provides a quantitative metric, constructed from sensitivity and covariance matrices, for determining the degree of agreement among the various computational and experimental data while eliminating discrepant information. This Section will introduce the BERRU-PM-2+ predictive modeling methodology, which further generalizes the BERRU-PM methodology provided in [15,16] by including all of the second-order terms in response sensitivities and providing explicitly the mechanism for incorporating not only second-order but arbitrarily high-order (hence the “2+” in the acronym BERRU-PM-2+) sensitivities of responses with respect to parameters, overcoming the “curse of dimensionality” [41].

3.1. Construction of the Second-Order-Accurate MaxEnt Probability Distribution $p_c(\mathbf{z}|\mathbf{z}_c, \mathbf{C}_c)$ of Computational Model Responses and Parameters

The model parameters are denoted as α_i , $i = 1, \dots, TP$, where TP denotes the “total number of parameters” and are considered, without loss of generality, to be real-valued scalars. For subsequent mathematical operations, it is convenient to consider that these model parameters are the components of the vector of parameters denoted as $\boldsymbol{\alpha} \triangleq (\alpha_1, \dots, \alpha_{TP})^\top \in \mathbb{R}^{TP}$. In practice, the model parameters are obtained from experimental measurements, which are external to the model and are performed independently of the manner in which the parameters are used in a computational model. Thus, the model parameters are not known exactly; they are subject to uncertainties. Formally, the model parameters can be considered to obey an unknown multivariate probability distribution, denoted as $p_\alpha(\boldsymbol{\alpha})$, which is defined on a domain (of definition of parameters) denoted as $D_\alpha \in \mathbb{R}^{TP}$ and is normalized such that $\int_{D_\alpha} p_\alpha(\boldsymbol{\alpha}) d\boldsymbol{\alpha} = 1$. Formally, the expected (or mean) value of a model parameter α_j , denoted as $\alpha_j^0 \triangleq E_c(\alpha_j)$, is defined as follows:

$$\alpha_j^0 \triangleq \int_{D_a} \alpha_j p_a(\mathbf{a}) d\mathbf{a}, \quad j = 1, \dots, TP; \quad \mathbf{a}^0 \triangleq (\alpha_1^0, \dots, \alpha_{TP}^0)^\dagger. \quad (66)$$

The covariance, $\text{cov}(\alpha_i, \alpha_j)$, of two parameters, α_i and α_j , is formally defined as follows:

$$\text{cov}(\alpha_i, \alpha_j) \triangleq \int_{D_a} \delta\alpha_i \delta\alpha_j p_a(\mathbf{a}) d\mathbf{a}; \quad i, j = 1, \dots, TP; \quad \delta\alpha_k \triangleq \alpha_k - \alpha_k^0; \quad k = 1, \dots, TP. \quad (67)$$

The variance, $\text{var}(\alpha_i)$, of a parameter α_i , is formally defined as the following particular case of Equation (67):

$$\text{var}(\alpha_i) \triangleq \int_{D_a} (\delta\alpha_i)^2 p_a(\mathbf{a}) d\mathbf{a}; \quad i = 1, \dots, TP. \quad (68)$$

The standard deviation of α_i is denoted as σ_i and is defined as follows: $\sigma_i \triangleq \sqrt{\text{var}(\alpha_i)}$. The definitions of the higher-order correlations among parameters, which are used together with those shown in Equations (66)–(68) to construct the moments of the distribution of computed responses (i.e., results of interest computed using the model), are presented in Appendix A.

A model response is a function (implicit and, occasionally, explicit) of the model parameters and is denoted as $r_k(\mathbf{a})$. Evidently, the model response is also an imprecisely known quantity, subject to uncertainties stemming from the model parameters, in addition to uncertainties stemming from the numerical computational procedures. A model response thus comprises by both systematic and random errors. For subsequent mathematical operations, it is convenient to consider that the responses $r_k(\mathbf{a})$, $k = 1, \dots, TR$, where TR denotes the “total number of responses” computed using the model, are the components of the vector of responses denoted as $\mathbf{r}(\mathbf{a}) \triangleq [r_1(\mathbf{a}), \dots, r_{TR}(\mathbf{a})]^\dagger$.

Conceptually, the computational model’s responses can be considered to obey an unknown multivariate probability distribution function, denoted as $p_c(\mathbf{r})$, which is defined on a domain denoted as $D_r \in \mathbb{R}^{TR}$ and is normalized such that $\int_{D_r} p_c(\mathbf{r}) d\mathbf{r} = 1$. Further-

more, the parameters and computed responses can be considered to follow a joint probability distribution function denoted as $p_c(\mathbf{a}, \mathbf{r}) \triangleq p_a(\mathbf{a}) p_c(\mathbf{r})$, which is not known, but is formally defined on a computational domain denoted as $D_c = D_a \cup D_r$, which comprises the reunion of the domains of definitions of all of the model’s parameters and responses, and is properly normalized, i.e., $\int_{D_c} p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = 1$. The subscript “c” is used to indicate

quantities pertaining to the “computational model”.

In addition to results produced by computational models, in practice there also exist independent measurements of responses and parameters, which are subject to uncertainties which stem from the respective measurement devices and procedures, independently of the computational model. Since such measurements are scalar quantities which are compared to the computational results, it is convenient to consider that the computational responses are also real-valued scalar quantities, to enable direct comparisons with the corresponding measured responses.

Since the system’s response is a function of the system’s parameters, it is assumed that each computed response can be formally expanded in a multivariate Taylor-series around the parameters’ mean values. In particular, the fourth-order Taylor-series of a system response, denoted as $r_k(\mathbf{a})$, around the expected (or nominal) parameter values $\mathbf{a}^0$ has the following formal expression:

$$\begin{aligned}
r_k(\mathbf{a}) = & r_k(\mathbf{a}^0) + \sum_{j_1=1}^{TP} \left\{ \frac{\partial r_k(\mathbf{a})}{\partial \alpha_{j_1}} \right\}_{\mathbf{a}^0} \delta \alpha_{j_1} + \frac{1}{2} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \left\{ \frac{\partial^2 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2}} \right\}_{\mathbf{a}^0} \delta \alpha_{j_1} \delta \alpha_{j_2} \\
& + \frac{1}{3!} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \left\{ \frac{\partial^3 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3}} \right\}_{\mathbf{a}^0} \delta \alpha_{j_1} \delta \alpha_{j_2} \delta \alpha_{j_3} \\
& + \frac{1}{4!} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \sum_{j_4=1}^{TP} \left\{ \frac{\partial^4 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3} \partial \alpha_{j_4}} \right\}_{\mathbf{a}^0} \delta \alpha_{j_1} \delta \alpha_{j_2} \delta \alpha_{j_3} \delta \alpha_{j_4} + \varepsilon_k.
\end{aligned} \quad (69)$$

In Equation (69), the quantity $r_k(\mathbf{a}^0)$ indicates the computed value of the response using the expected/nominal parameter values $\mathbf{a}^0 \triangleq (\alpha_1^0, \dots, \alpha_{TP}^0)^\top$, while the notation $\{ \}_{\mathbf{a}^0}$ indicates that the quantities within the braces (i.e., first- and second-order sensitivities of the response with respect to the parameters) are also computed using the expected/nominal parameter values. The quantity ε_k in Equation (69) comprises all quantifiable errors in the representation of the computed response as a function of the model parameters, including the truncation errors $O(\delta \alpha_j)^5$ of the Taylor-series expansion, possible bias-errors due to incompletely modeled physical phenomena, and possible random errors due to numerical approximations. The radius/domain of convergence of the series in Equation (69) determines the largest values of the parameter variations $\delta \alpha_j$ which are admissible before the respective series becomes divergent. In turn, these maximum admissible parameter variations limit, through Equation (67), the largest parameter covariances/standard deviations which can be considered for using the Taylor-expansion for the subsequent purposes of computing moments of the distribution of computed responses.

As is well known, and as indicated by Equation (69), the Taylor-series of a function of TP -variables [e.g., $r_k(\mathbf{a})$] comprises TP 1st-order derivatives, $TP(TP+1)/2$ distinct second-order derivatives, and so on. The computation by conventional methods of the n^{th} -order functional derivatives (called “sensitivities” in the field of sensitivity analysis) of a response with respect to the TP -parameters (on which it depends) would require at least $O(TP^n)$ large-scale computations. The exponential increase—with the order of response sensitivities—of the number of large-scale computations needed to determine higher-order sensitivities is the manifestation of the “curse of dimensionality in sensitivity analysis”, by analogy to the expression coined by Belmann [41] to express the difficulty of using “brute-force” grid search when optimizing a function with many input variables. The “ n^{th} -order Comprehensive Adjoint Sensitivity Analysis Methodology for Nonlinear Systems” (n^{th} -CASAM-N) conceived by Cacuci [20] and the “ n^{th} -order Comprehensive Adjoint Sensitivity Analysis Methodology for Response-Coupled Forward/Adjoint Linear Systems” (n^{th} -CASAM-L) conceived by Cacuci [21] are currently the only methodologies that enable the exact and efficient computation of arbitrarily high-order sensitivities while overcoming the curse of dimensionality.

Uncertainties in the model’s parameters will evidently give rise to uncertainties in the model responses $r_k(\mathbf{a})$. The approximate moments of the unknown distribution of $r_k(\mathbf{a})$ are obtained by using the so-called “propagation of errors” methodology, integrating formally, over the unknown distribution, various expressions involving the truncated Taylor-series expansion of the response provided in Equation (69), as first performed by Tukey [42]. Tukey’s results were generalized to sixth-order by Cacuci [21].

The expectation value, $E(r_k)$, of a response $r_k(\mathbf{a})$ is obtained by integrating formally Equation (69) over $p_c(\mathbf{a}, \mathbf{r})$, i.e.,:

$$E_c(r_k) \triangleq \int_{D_c} r_k(\mathbf{a}) p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}. \quad (70)$$

The approximations of $E(r_k)$ up to fourth-order in the parameters' standard deviations are presented in Appendix A.

The expression of the correlation between a computed responses and a parameter variance, which is denoted as $\text{cov}(\alpha_i, r_k)$, is defined below

$$\text{cov}(\alpha_i, r_k) \triangleq \int_{D_c} \delta\alpha_i[r_k(\mathbf{a}) - E_c(r_k)] p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}. \quad (71)$$

The approximations of $\text{cov}(\alpha_i, r_k)$ up to fourth-order in the parameters' standard deviations are presented in Appendix A.

The covariance between two responses r_k and r_ℓ is denoted as $\text{cov}(r_k, r_\ell)$ and is defined below:

$$\text{cov}(r_k, r_\ell) \triangleq \int_{D_c} [r_k(\mathbf{a}) - E_c(r_k)][r_\ell(\mathbf{a}) - E_c(r_\ell)] p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}. \quad (72)$$

The approximations of $\text{cov}(r_k, r_\ell)$ up to fourth-order in in the parameters' standard deviations are presented in Appendix A.

It is assumed in this work that the only information about the distribution $p_c(\mathbf{a}, \mathbf{r})$ is provided by the known first- and second-order moments of the joint distribution of model parameters and responses defined by Equations (66), (67), (70), (71) and (72). It is well known that when the underlying distribution $p_c(\mathbf{a}, \mathbf{r})$ needs to be determined from such incomplete information, the principle of maximum entropy (MaxEnt) originally formulated by Jaynes [17] provides the optimal compatibility with the available information, while simultaneously ensuring minimal spurious information content. According to the MaxEnt principle, the probability density $p_c(\mathbf{a}, \mathbf{r})$ would satisfy the “available information” without implying any spurious information or hidden assumptions if:

- (i) $p_c(\mathbf{a}, \mathbf{r})$ maximizes the Shannon [43] information entropy for the computational model, S_c , defined below:

$$S_c = - \int_{D_c} p_c(\mathbf{a}, \mathbf{r}) \ln[p_c(\mathbf{a}, \mathbf{r})] d\mathbf{a} d\mathbf{r}, \quad (73)$$

- (ii) $p_c(\mathbf{a}, \mathbf{r})$ satisfies the “moments constraints” defined by Equations (66), (67), (70), (71) and (72);
- (iii) $p_c(\mathbf{a}, \mathbf{r})$ satisfies the normalization condition:

$$\int_{D_c} p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = 1. \quad (74)$$

The MaxEnt distribution $p_c(\mathbf{a}, \mathbf{r})$ is obtained as the solution of the constrained variational problem $\partial H_c(p_c)/\partial p_c = 0$, where the entropy (Lagrangian functional) $H_c(p_c)$ is defined as follows:

$$\begin{aligned}
H_c(p_c) = & - \int_{D_c} p_c(\mathbf{a}, \mathbf{r}) \ln[p_c(\mathbf{a}, \mathbf{r})] d\mathbf{a} d\mathbf{r} - \lambda_0 \left[\int_{D_c} p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} - 1 \right] \\
& - \sum_{k=1}^{TR} \lambda_k^{(1)} \left[\int_{D_c} r_k p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} - E_c(r_k) \right] - \sum_{i=1}^{TP} \lambda_i^{(2)} \left[\int_{D_c} \alpha_i p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} - \alpha_i^0 \right] \\
& - \frac{1}{2} \sum_{k=1}^{TR} \sum_{\ell=1}^{TR} \lambda_{k\ell}^{(11)} \left[\int_{D_c} r_k r_\ell p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} - \text{cov}(r_k, r_\ell) - E_c(r_k) E_c(r_\ell) \right] \\
& - \sum_{k=1}^{TR} \sum_{i=1}^{TP} \lambda_{ki}^{(12)} \left[\int_{D_c} r_k \alpha_i p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} - \text{cov}(r_k, \alpha_i) - E_c(r_k) \alpha_i^0 \right] \\
& - \frac{1}{2} \sum_{i=1}^{TP} \sum_{j=1}^{TP} \lambda_{ij}^{(22)} \left[\int_{D_c} \alpha_i \alpha_j p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} - \text{cov}(\alpha_i, \alpha_j) - \alpha_i^0 \alpha_j^0 \right].
\end{aligned} \quad (75)$$

In Equation (75), the quantities $\lambda_k^{(1)}$, $\lambda_k^{(2)}$, $\lambda_{k\ell}^{(11)}$, $\lambda_{k\ell}^{(12)}$, and $\lambda_{ij}^{(22)}$ denote the respective Lagrange multipliers, and the factors $1/2$ are introduced for subsequent computational convenience.

Solving the equation $\partial H(p_c)/\partial p_c = 0$ yields the following expression for the resulting MaxEnt distribution $p_c(\mathbf{z})$:

$$p_c(\mathbf{z}) = \frac{1}{Z_c(\mathbf{b}, \Lambda_c)} \exp\left(-\mathbf{b}^\top \mathbf{z} - \frac{1}{2} \mathbf{z}^\top \Lambda_c \mathbf{z}\right), \quad (76)$$

where the subscript “c” indicates quantities related to the computational model and where the various vectors and matrices are defined as follows:

$$\mathbf{z} = \begin{pmatrix} \mathbf{r} \\ \boldsymbol{\alpha} \end{pmatrix}; \quad \mathbf{b} \triangleq \begin{pmatrix} \boldsymbol{\lambda}^{(1)} \\ \boldsymbol{\lambda}^{(2)} \end{pmatrix}; \quad \boldsymbol{\lambda}^{(1)} \triangleq \begin{pmatrix} \lambda_1^{(1)} \\ \vdots \\ \lambda_{TR}^{(1)} \end{pmatrix}; \quad \boldsymbol{\lambda}^{(2)} \triangleq \begin{pmatrix} \lambda_1^{(2)} \\ \vdots \\ \lambda_{TP}^{(2)} \end{pmatrix}; \quad (77)$$

$$\begin{aligned}
\Lambda_c & \triangleq \begin{pmatrix} \Lambda_c^{(11)} & \Lambda_c^{(12)} \\ \left[\Lambda_c^{(12)}\right]^\top & \Lambda_c^{(22)} \end{pmatrix}; \quad \Lambda_c^{(11)} \triangleq \begin{pmatrix} \lambda_{11}^{(11)} & \cdot & \lambda_{1,TR}^{(11)} \\ \cdot & \lambda_{k\ell}^{(11)} & \cdot \\ \lambda_{TR,1}^{(11)} & \cdot & \lambda_{TR,TR}^{(11)} \end{pmatrix}; \\
\Lambda_c^{(12)} & \triangleq \begin{pmatrix} \lambda_{11}^{(12)} & \cdot & \lambda_{1,TP}^{(12)} \\ \cdot & \lambda_{k\ell}^{(12)} & \cdot \\ \lambda_{TR,1}^{(12)} & \cdot & \lambda_{TR,TP}^{(12)} \end{pmatrix}; \quad \Lambda_c^{(22)} \triangleq \begin{pmatrix} \lambda_{11}^{(22)} & \cdot & \lambda_{1,TP}^{(22)} \\ \cdot & \lambda_{k\ell}^{(22)} & \cdot \\ \lambda_{TP,1}^{(22)} & \cdot & \lambda_{TP,TP}^{(22)} \end{pmatrix}.
\end{aligned} \quad (78)$$

The normalization constant $Z_c(\mathbf{b}, \Lambda_c)$ in Equation (76) is defined as follows:

$$Z_c(\mathbf{b}, \Lambda_c) = \int_{D_c} \exp\left(-\mathbf{b}^\top \mathbf{z} - \frac{1}{2} \mathbf{z}^\top \Lambda_c \mathbf{z}\right) d\mathbf{z}; \quad d\mathbf{z} \equiv d\mathbf{a} d\mathbf{r}. \quad (79)$$

In statistical mechanics, the normalization constant Z_c is called the partition function (or sum over states) and carries all of the information available about the possible states of the system, while the MaxEnt distribution $p_c(\mathbf{z})$ is called the canonical Boltzmann–Gibbs distribution. The integral in Equation (79) can be evaluated explicitly by conservatively extending the computational domain D_c to the entire multidimensional real vector space $\mathbb{R}^N$, where $N \triangleq TR + TP$, to obtain the following expression:

$$Z_c(\mathbf{b}, \Lambda_c) = \int_{\mathbb{R}^N} \exp\left(-\mathbf{b}^\top \mathbf{z} - \frac{1}{2} \mathbf{z}^\top \Lambda_c \mathbf{z}\right) d\mathbf{z} = \frac{(2\pi)^{N/2}}{\sqrt{\text{Det}(\Lambda_c)}} e^{\left(\frac{1}{2} \mathbf{b}^\top \Lambda_c^{-1} \mathbf{b}\right)}. \quad (80)$$

The Lagrange multipliers are determined in terms of the known information (means and covariances of parameters and responses) by differentiating the “free energy” $F(\mathbf{b}, \mathbf{C}_c) \triangleq -\ln Z_c(\mathbf{b}, \mathbf{C}_c)$ with respect to the components of the vector $\mathbf{b} \triangleq (\lambda^{(1)}, \lambda^{(2)})^\top$ to obtain the following expressions:

$$\frac{\partial F(\mathbf{b}, \mathbf{C}_c)}{\partial \lambda_k^{(1)}} = \frac{1}{Z_c} \int_{D_c} r_k e^{\left(-\mathbf{b}^\top \mathbf{z} - \frac{1}{2} \mathbf{z}^\top \Lambda_c \mathbf{z}\right)} d\mathbf{z} = \int_{D_c} r_k p_c(\mathbf{a}, \mathbf{r}) d\mathbf{z} = E_c(r_k); \quad k = 1, \dots, TR; \quad (81)$$

$$\frac{\partial F(\mathbf{b}, \mathbf{C}_c)}{\partial \lambda_i^{(2)}} = \frac{1}{Z_c} \int_{D_c} \alpha_i e^{\left(-\mathbf{b}^\top \mathbf{z} - \frac{1}{2} \mathbf{z}^\top \Lambda_c \mathbf{z}\right)} d\mathbf{z} = \int_{D_c} \alpha_i p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = \alpha_i^0; \quad i = 1, \dots, TP. \quad (82)$$

The results obtained in Equations (81) and (82) can be collectively written in vector-matrix form as follows:

$$\frac{\partial F(\mathbf{b}, \mathbf{C}_c)}{\partial \mathbf{b}} = \mathbf{z}_c; \quad \mathbf{z}_c \triangleq [\mathbf{E}_c(\mathbf{r}), \alpha^0]^\top; \quad \mathbf{E}_c(\mathbf{r}) \triangleq [E_c(r_1), \dots, E_c(r_{TR})]^\top. \quad (83)$$

On the other hand, it follows from Equation (80) that

$$\frac{\partial F(\mathbf{b}, \mathbf{C}_c)}{\partial \mathbf{b}} = -\Lambda_c^{-1} \mathbf{b}. \quad (84)$$

The relations obtained in Equations (83) and (84) imply the following relation:

$$\mathbf{b} = -\Lambda_c \mathbf{z}_c. \quad (85)$$

Differentiating a second time the relation provided in Equation (81) or (82) yields the following relations:

$$\begin{aligned} \frac{\partial^2 F(\mathbf{b}, \mathbf{C}_c)}{\partial \lambda_j^{(1)} \partial \lambda_k^{(1)}} &= -\frac{1}{Z_c^2} \frac{\partial Z_c}{\partial \lambda_j^{(1)}} E_c(r_k) + \frac{1}{Z_c} \int_{D_c} (-r_j) r_k e^{\left(-\mathbf{b}^\top \mathbf{z} - \frac{1}{2} \mathbf{z}^\top \Lambda_c \mathbf{z}\right)} d\mathbf{z} \\ &= E_c(r_j) E_c(r_k) - E_c(r_j r_k) \triangleq -\text{cov}(r_j, r_k); \quad j, k = 1, \dots, TR. \end{aligned} \quad (86)$$

$$\frac{\partial^2 F(\mathbf{b}, \mathbf{C}_c)}{\partial \lambda_i^{(2)} \partial \lambda_k^{(1)}} = E_c(r_k) \alpha_i^0 - E_c(\alpha_i r_k) \triangleq -\text{cov}(\alpha_i, r_k); \quad i = 1, \dots, TP; \quad k = 1, \dots, TR; \quad (87)$$

$$\frac{\partial^2 F(\mathbf{b}, \mathbf{C}_c)}{\partial \lambda_i^{(2)} \partial \lambda_j^{(2)}} = \alpha_i^0 \alpha_j^0 - E_c(\alpha_i \alpha_j) \triangleq -\text{cov}(\alpha_i, \alpha_j); \quad i, j = 1, \dots, TP. \quad (88)$$

The results obtained in Equations (86)–(88) can be collectively written in vector-matrix form as follows:

$$\begin{aligned} \frac{\partial^2 F(\mathbf{b}, \mathbf{C}_c)}{\partial \mathbf{b} \partial \mathbf{b}} &= -\mathbf{C}_c; \quad \mathbf{C}_c \triangleq \begin{pmatrix} \mathbf{C}_c^{rr} & \mathbf{C}_c^{r\alpha} \\ \mathbf{C}_c^{\alpha r} & \mathbf{C}_c^{\alpha\alpha} \end{pmatrix}; \quad \mathbf{C}_c^{rr} \triangleq [\text{cov}(r_j, r_k)]_{TR \times TR}; \\ \mathbf{C}_c^{r\alpha} &\triangleq [\text{cov}(r_k, \alpha_i)]_{TR \times TP} = (\mathbf{C}_c^{\alpha r})^\top; \quad \mathbf{C}_c^{\alpha\alpha} \triangleq [\text{cov}(\alpha_i, \alpha_j)]_{TP \times TP}. \end{aligned} \quad (89)$$

On the other hand, it follows from Equation (84) that

$$\frac{\partial^2 F(\mathbf{b}, \mathbf{C}_c)}{\partial \mathbf{b} \partial \mathbf{b}} = -\Lambda_c^{-1}. \quad (90)$$

The relations obtained in Equations (89) and (90) imply the following relation:

$$\Lambda_c^{-1} = \mathbf{C}_c. \quad (91)$$

Introducing the results obtained in Equations (84) and (91) into Equations (76) and (80) yields the following expression for the MaxEnt distribution $p_c(\mathbf{z}|\mathbf{z}_c, \mathbf{C}_c)$:

$$p_c(\mathbf{z}|\mathbf{z}_c, \mathbf{C}_c) = \frac{(2\pi)^{-N/2}}{\sqrt{\text{Det}(\mathbf{C}_c)}} \exp\left[-\frac{1}{2}(\mathbf{z}-\mathbf{z}_c)^\dagger \mathbf{C}_c^{-1}(\mathbf{z}-\mathbf{z}_c)\right]. \quad (92)$$

3.2. Construction of the Second-Order-Accurate MaxEnt Probability Distribution $p_e(\mathbf{z}|\mathbf{z}_e, \mathbf{C}_e)$ of Experimentally Measured Responses

In addition to information regarding the computational model's responses and parameters, in practice there also exists experimental information about the responses and model parameters, which are obtained independently of computations. In the most comprehensive setting, there would be independent measurements for the computed responses, for the model parameters and for measured correlations between the model parameters and the measured responses. In this idealized setting, each one of the computed responses would also be measured experimentally, so that information about the mean values and covariances for a total number TR of experimentally measured responses would also be available. The mean values of the experimentally measured responses are denoted as r_i^e , $i = 1, \dots, TR$, and the covariances of two measured responses are denoted as $\text{cov}(r_i, r_j)_e$, $i, j = 1, \dots, TR$. The letter "e" is used either as a superscript or a subscript to indicate experimentally measured quantities. Similarly, experimentally-determined values for the expected values α_i^e , $i = 1, \dots, TP$, and newly measured covariances $\text{cov}(\alpha_i, \alpha_j)_e$ for the model parameters could also be available. In principle, it is also possible to obtain correlations between some measured responses and some model parameters. In the nuclear energy field, for example, an important measured detector response is the fission rate in a material (e.g., reactor fuel) which contains (fissionable) uranium-235. However, the fission cross section of uranium-235 would also be a model parameter in the computation of the fission reaction rate, which would thus be correlated to the response.

Formally, the measured responses and parameters are considered to follow a joint unknown probability distribution function, which can be denoted as $p_e(\mathbf{a}, \mathbf{r})$ and which is formally defined on a domain denoted as $D_e \in \mathbb{R}^N$; $N \triangleq TP + TR$. The means and covariances of the experimentally measured responses can be defined formally as follows:

$$r_i^e \triangleq \int_{D_e} r_i p_e(\mathbf{a}, \mathbf{r}) d\mathbf{r} d\mathbf{a}, \quad i = 1, \dots, TR. \quad (93)$$

$$\text{cov}(r_i, r_j)_e \triangleq \int_{D_e} (r_i - r_i^e)(r_j - r_j^e) p_e(\mathbf{a}, \mathbf{r}) d\mathbf{r} d\mathbf{a}; \quad i, j = 1, \dots, TR. \quad (94)$$

The expected values of the measured responses are considered to constitute the components of a vector denoted as $\mathbf{r}_e \triangleq (r_1^e, \dots, r_{TR}^e)^\dagger$. The covariances (i.e., standard deviations and correlations) of the measured responses are considered to be components of the $TR \times TR$ -dimensional covariance matrix of measured responses, which are denoted as $\mathbf{C}_e^{rr} \triangleq [\text{cov}(r_i, r_j)_e]_{TR \times TR}$. Similarly, the mean values and covariances for the additional measurements of model parameters can be formally represented as follows:

$$\alpha_j^e \triangleq \int_{D_e} \alpha_j p_e(\mathbf{a}, \mathbf{r}) d\mathbf{r} d\mathbf{a}, \quad j = 1, \dots, TP. \quad (95)$$

The expected values of the measured parameters are considered to constitute the components of a vector denoted as $\mathbf{a}_e \triangleq (\alpha_1^e, \dots, \alpha_{TR}^e)^\top$. The covariance, $\text{cov}(\alpha_i, \alpha_j)_e$, of two independently/newly measured parameters, α_i and α_j , is defined as follows:

$$\text{cov}(\alpha_i, \alpha_j)_e \triangleq \int_{D_e} \delta\alpha_i^e \delta\alpha_j^e p_e(\mathbf{a}, \mathbf{r}) d\mathbf{r} d\mathbf{a}; i, j = 1, \dots, TP; \delta\alpha_i^e \triangleq \alpha_i - \alpha_i^e. \quad (96)$$

The covariances $\text{cov}(\alpha_i, \alpha_j)_e$ would constitute the components of the covariance matrix, denoted as $\mathbf{C}_e^{\alpha\alpha}$, of newly measured parameters. When correlations, denoted as $\text{cov}(\alpha_i, r_j)_e$, between the measured responses and the model parameters are available, they can formally be considered to be elements of a (very sparse, in practice) rectangular correlation matrix which is denoted as $\mathbf{C}_e^{\alpha r} \triangleq [\text{cov}(\alpha_i, r_j)_e]_{TP \times TR}$, where

$$\text{cov}(\alpha_i, r_j)_e \triangleq \int_{D_e} (\alpha_i - \alpha_i^e)(r_j - r_j^e) p_e(\mathbf{a}, \mathbf{r}) d\mathbf{r} d\mathbf{a}; i = 1, \dots, TP; j = 1, \dots, TR. \quad (97)$$

The MaxEnt principle [17] can now be applied to construct the least informative (and hence, most conservative) distribution using the information provided by the vector of mean experimental values $\mathbf{r}_e$ and covariance matrices $\mathbf{C}_e^{rr} \triangleq [\text{cov}(r_i, r_j)_e]_{TR \times TR}$ and $\mathbf{C}_e^{r\alpha} \triangleq [\text{cov}(\alpha_i, r_j)_e]_{TP \times TR}$. This MaxEnt distribution for the experimentally measured responses is denoted as $p_e(\mathbf{z}|\mathbf{z}_e, \mathbf{C}_e)$ and is constructed by applying the same steps as those leading to Equation (92), above, to obtain the following expression:

$$p_e(\mathbf{z}|\mathbf{z}_e, \mathbf{C}_e) = \frac{(2\pi)^{-N/2}}{\sqrt{\text{Det}(\mathbf{C}_e)}} \exp\left[-\frac{1}{2}(\mathbf{z} - \mathbf{z}_e)^\top \mathbf{C}_e^{-1}(\mathbf{z} - \mathbf{z}_e)\right], \quad (98)$$

where

$$\mathbf{C}_e \triangleq \begin{pmatrix} \mathbf{C}_e^{rr} & \mathbf{C}_e^{r\alpha} \\ \mathbf{C}_e^{\alpha r} & \mathbf{C}_e^{\alpha\alpha} \end{pmatrix}; \mathbf{z} = \begin{pmatrix} \mathbf{r} \\ \mathbf{a} \end{pmatrix}; \mathbf{z}_e = \begin{pmatrix} \mathbf{r}_e \\ \mathbf{a}_e \end{pmatrix}. \quad (99)$$

3.3. Construction of the Complete Second-Order-Accurate Joint Posterior MaxEnt Probability Distribution of Computed and Measured Responses and Model Parameters

The joint posterior probability distribution of all computed and experimentally measured quantities, which is denoted as $p_p(\mathbf{z}|\mathbf{z}_p, \mathbf{C}_p)$ where the subscript “p” indicates “posterior,” is obtained as the properly normalized product of the distributions $p_e(\mathbf{z}|\mathbf{z}_e, \mathbf{C}_e)$ and $p_c(\mathbf{z}|\mathbf{z}_c, \mathbf{C}_c)$. The posterior probability $p_p(\mathbf{z}|\mathbf{z}_p, \mathbf{C}_p)$ is defined on the domain D_p obtained as the union of the computational and experimental domains, i.e., $D_p \triangleq D_c \cup D_e$. Straightforward, albeit tedious, algebraic computations lead to the following expression for $p_p(\mathbf{z}|\mathbf{z}_p, \mathbf{C}_p)$:

$$p_p(\mathbf{z}|\mathbf{z}_p, \mathbf{C}_p) \triangleq p_c(\mathbf{z}|\mathbf{z}_c, \mathbf{C}_c) p_e(\mathbf{z}|\mathbf{z}_e, \mathbf{C}_e) = K \exp\left[-\frac{1}{2}Q(\mathbf{z}|\mathbf{z}_p, \mathbf{C}_p)\right], \quad (100)$$

where the normalization constant K and the quadratic form $Q(\mathbf{z}_p, \mathbf{C}_p)$ have the following expressions, respectively:

$$K \triangleq (2\pi)^{-N/2} [\text{Det}(\mathbf{C}_c + \mathbf{C}_e)]^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{z}_c - \mathbf{z}_e)^\top (\mathbf{C}_c + \mathbf{C}_e)^{-1}(\mathbf{z}_c - \mathbf{z}_e)\right]; \quad (101)$$

$$Q(\mathbf{z}|\mathbf{z}_p, \mathbf{C}_p) \triangleq (\mathbf{z} - \mathbf{z}_p)^\dagger \mathbf{C}_p^{-1} (\mathbf{z} - \mathbf{z}_p); \quad (102)$$

$$\mathbf{C}_p \triangleq (\mathbf{C}_c^{-1} + \mathbf{C}_e^{-1})^{-1} = \mathbf{C}_c - \mathbf{C}_c (\mathbf{C}_c + \mathbf{C}_e)^{-1} \mathbf{C}_c = \mathbf{C}_e - \mathbf{C}_e (\mathbf{C}_c + \mathbf{C}_e)^{-1} \mathbf{C}_e; \quad (103)$$

$$\mathbf{z}_p \triangleq \mathbf{C}_p (\mathbf{C}_c^{-1} \mathbf{z}_c + \mathbf{C}_e^{-1} \mathbf{z}_e). \quad (104)$$

The expression obtained in Equation (100) provides the exact first-order (mean values) and second-order (variances and co-variances) of the most comprehensive combined distribution of computations and measurements of responses and parameters. In practice, such a comprehensive amount of experimental information is highly unlikely to be available. Furthermore, even if such massive amount of experimental information were available, the inversion of the matrix $(\mathbf{C}_c + \mathbf{C}_e)^{-1}$ would require massive computational resources for large systems involving many parameters.

3.4. Practical Situation: Only Response Measurements Are Available to Be Assimilated

In practice, the information (mean values and covariances) about the model parameters indicated in Equations (66)–(68) is obtained prior to using this information in a model and, hence, for computing responses using the respective model. Additional experimental information would usually be available only about the responses of interest. Any experimental information that might be available about the model parameters and/or about correlations between some model parameters and some responses would need to be assessed against the similar information already available in the covariance/correlation matrices $\mathbf{C}^{\alpha\alpha}$ and/or $\mathbf{C}_c^{r\alpha}$, so the respective experimental and computational information could be combined into the appropriate components of these matrices. In practice, therefore, only information that becomes additionally available about measured responses would need to be assimilated explicitly, i.e., only the components of the vector $\mathbf{r}_e \triangleq (r_1^e, \dots, r_{TR}^e)^\dagger$ and of the matrix $\mathbf{C}_e^{rr} \triangleq [\text{cov}(r_i, r_j)]_{e, TR \times TR}$ would become available for assimilation and predictive modeling. These measured responses would be considered to follow an unknown probability distribution function, denoted as $p_e(\mathbf{r})$ and defined on a domain denoted as $D_e \in \mathbb{R}^{TR}$, so that the expressions in Equations (93) and (94) would take on the following forms, respectively:

$$r_i^e \triangleq \int_{D_e} r_i p_e(\mathbf{r}) d\mathbf{r}, \quad i = 1, \dots, TR. \quad (105)$$

$$\text{cov}(r_i, r_j)_e \triangleq \int_{D_e} (r_i - r_i^e)(r_j - r_j^e) p_e(\mathbf{r}) d\mathbf{r}; \quad i, j = 1, \dots, TR. \quad (106)$$

The MaxEnt distribution corresponding to the information provided in Equations (105) and (106) has the following expression:

$$p_e(\mathbf{r}|\mathbf{r}_e, \mathbf{C}_e^{rr}) = (2\pi)^{-TR/2} [\text{Det}(\mathbf{C}_e^{rr})]^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{r} - \mathbf{r}_e)^\dagger (\mathbf{C}_e^{rr})^{-1} (\mathbf{r} - \mathbf{r}_e)\right]. \quad (107)$$

Furthermore, when only the experimental information represented by the distribution $p_e(\mathbf{r}|\mathbf{r}_e, \mathbf{C}_e^{rr})$ is available, the posterior joint probability distribution of the computed and measured quantities, which is denoted as $p_p(\mathbf{r}, \boldsymbol{\alpha}) = p_c(\mathbf{z}|\mathbf{z}_c, \mathbf{C}_c) p_e(\mathbf{r}|\mathbf{r}_e, \mathbf{C}_e^{rr})$ and which is considered to be defined on the corresponding domain D_p , takes on the following form:

$$p_p(\mathbf{r}, \mathbf{a}) = N \exp \left[-\frac{1}{2} Q(\mathbf{r}, \mathbf{a}) \right]; \quad N \triangleq \left[\int_{D_p} \exp \left[-\frac{1}{2} Q(\mathbf{r}, \mathbf{a}) \right] d\mathbf{a} d\mathbf{r} \right]^{-1}; \quad (108)$$

$$Q(\mathbf{r}, \mathbf{a}) \triangleq (\mathbf{z} - \mathbf{z}_c)^{\dagger} \mathbf{C}_c^{-1} (\mathbf{z} - \mathbf{z}_c) + (\mathbf{r} - \mathbf{r}_e)^{\dagger} (\mathbf{C}_e^{rr})^{-1} (\mathbf{r} - \mathbf{r}_e).$$

The posterior distribution represented by Equation (108) is evidently not Gaussian. The posterior mean values for the best-estimate predicted responses and parameters are defined as before, namely:

$$\mathbf{r}_{be} \triangleq \left[\int_{D_p} p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} \right]^{-1} \int_{D_p} \mathbf{r} p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}; \quad (109)$$

$$\mathbf{a}_{be} \triangleq \left[\int_{D_p} p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} \right]^{-1} \int_{D_p} \mathbf{a} p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}; \quad (110)$$

The evaluation of the integrals appearing in Equations (109) and (110) can be performed to a high degree of accuracy, with a controlled error, by employing the saddle-point (Laplace) method [44–46]. For a ratio of integrals of the form

$$I = \left[\int \exp[-g(\mathbf{z})] d\mathbf{z} \right]^{-1} \int f(\mathbf{z}) \exp[-g(\mathbf{z})] d\mathbf{z}, \quad (111)$$

the saddle-point (Laplace) method yields the following result [44–46]:

$$I = \hat{f} - \frac{1}{2} \hat{f}_{,i} \hat{g}_{,i_2 i_3 i_4} \hat{g}^{i_2 i_3 i_4} + \frac{1}{2} \hat{f}_{,j_1 j_2} \hat{g}^{j_1 j_2} + O(\hat{f}_{,j_1 j_2 j_3}) \quad (112)$$

where:

- (i) the derivative of a function with respect to a component of $\mathbf{z}$ is denoted using a subscript, e.g., $f_i \triangleq \partial f / \partial z_i$, $f_{ij} \triangleq \partial^2 f / \partial z_i \partial z_j$, $i, j = 1, \dots, TI$, where TI denotes the total number of independent variables;
- (ii) the superscripts denote the respective component of the inverse Hessian of the respective function, e.g., f^{ij} denotes the (i, j) -element of the inverse Hessian matrix $[f_{ij}^{-1}]$;
- (iii) an index that appears as a subscript and a superscript implies a summation over all possible values of that index;
- (iv) the “hat” denotes that the respective quantity is evaluated at the saddle point of $\exp[-g(\mathbf{z})]$, which is defined as the point at which the gradient $g_z(\mathbf{z}) \triangleq (\partial g / \partial z_1, \dots, \partial g / \partial z_{TI})$ of $g(\mathbf{z})$ vanishes, i.e., $g_z(\mathbf{z}) = \mathbf{0}$.

The saddle-point of $p_p(\mathbf{r}, \mathbf{a})$ is denoted as $(\mathbf{r}_s, \mathbf{a}_s)$ and is defined by the following relations:

$$\frac{\partial Q(\mathbf{r}, \mathbf{a})}{\partial \mathbf{r}} = \mathbf{0}, \quad \frac{\partial Q(\mathbf{r}, \mathbf{a})}{\partial \mathbf{a}} = \mathbf{0}, \quad \text{at } (\mathbf{r}, \mathbf{a}) = (\mathbf{r}_s, \mathbf{a}_s). \quad (113)$$

To obtain the partial gradients (differentials) shown in Equation (113), it is convenient to write the matrix $\mathbf{C}_c^{-1}$ in the form $\mathbf{C}_c^{-1} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix}$, and use this form together with Equations (99) and (83) in Equation (108) to expand the functional $Q(\mathbf{r}, \mathbf{a})$ in the following form:

$$\begin{aligned} Q(\mathbf{r}, \mathbf{a}) = & [\mathbf{r} - \mathbf{E}_c(\mathbf{r})]^\dagger \mathbf{C}_{11} [\mathbf{r} - \mathbf{E}_c(\mathbf{r})] + [\mathbf{r} - \mathbf{E}_c(\mathbf{r})]^\dagger \mathbf{C}_{12} (\mathbf{a} - \mathbf{a}^0) \\ & + (\mathbf{a} - \mathbf{a}^0)^\dagger \mathbf{C}_{21} [\mathbf{r} - \mathbf{E}_c(\mathbf{r})] + (\mathbf{a} - \mathbf{a}^0)^\dagger \mathbf{C}_{22} (\mathbf{a} - \mathbf{a}^0) + (\mathbf{r} - \mathbf{r}_e)^\dagger (\mathbf{C}_e^r)^{-1} (\mathbf{r} - \mathbf{r}_e). \end{aligned} \quad (114)$$

Taking the partial differentials of the expression in Equation (114) yields the following equation at $(\mathbf{r}, \mathbf{a}) = (\mathbf{r}_s, \mathbf{a}_s)$:

$$\begin{pmatrix} \mathbf{r}_s - \mathbf{E}_c(\mathbf{r}) \\ \mathbf{a}_s - \mathbf{a}^0 \end{pmatrix} = - \begin{pmatrix} \mathbf{C}_c^{rr} & \mathbf{C}_c^{ra} \\ \mathbf{C}_c^{ar} & \mathbf{C}_c^{aa} \end{pmatrix} \begin{pmatrix} (\mathbf{C}_e^r)^{-1} (\mathbf{r}_s - \mathbf{r}_e) \\ \mathbf{0} \end{pmatrix}. \quad (115)$$

Solving Equation (115) leads to the following expressions for the saddle point coordinates $\mathbf{r}_s$ and $\mathbf{a}_s$:

$$\mathbf{r}_s = \mathbf{r}_e + \mathbf{C}_e^{rr} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} [\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e], \quad (116)$$

$$\mathbf{a}_s = \mathbf{a}^0 - \mathbf{C}_c^{ar} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} [\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]. \quad (117)$$

Using the results obtained in Equations (112), (116) and (117) in Equations (109) and (110), respectively, yields the following results for the best-estimate predicted values for the responses and, respectively, calibrated parameters:

$$\mathbf{r}_{be} = \mathbf{r}_s = \mathbf{r}_e + \mathbf{C}_e^{rr} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} [\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]; \quad (118)$$

$$\mathbf{a}_{be} = \mathbf{a}_s = \mathbf{a}^0 - \mathbf{C}_c^{ar} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} [\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]. \quad (119)$$

Since the components of the vector $\mathbf{E}_c(\mathbf{r})$, and the components of the matrices $\mathbf{C}_c^{rr}$ and $\mathbf{C}_c^{ar}$ can contain arbitrarily high-order response sensitivities to model parameters, as shown in Appendix A, the formulas presented in (118) and (119) generalize all of the previous formulas of this type found in data adjustment/assimilation procedures published to date (which contain at most second-order sensitivities). The best-estimate parameter values are the “calibrated model parameters” which can be used for subsequent computations with the “calibrated model”.

The second-order moments of the posterior distribution $p_p(\mathbf{r}, \mathbf{a})$ comprise the covariances between the best-estimated response, which are denoted as $\mathbf{C}_{be}^{rr}$, the covariances between the best-estimate parameters, which are denoted as $\mathbf{C}_{aa}^{be}$, and the correlations between the best-estimate parameters and responses, which are denoted as $\mathbf{C}_{ar}^{be}$. The expression of the “best-estimate” posterior parameter covariance matrix $\mathbf{C}_r^{be}$ for the best-estimate responses $\mathbf{r}_{be}$ is derived by using the results given in (118) and (119) to obtain the following expressions:

$$\mathbf{C}_{be}^{rr} \triangleq \int_{D_p} (\mathbf{r} - \mathbf{r}_{be})(\mathbf{r} - \mathbf{r}_{be})^\dagger p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = \mathbf{C}_e^{rr} - \mathbf{C}_e^{rr} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} \mathbf{C}_e^{rr}. \quad (120)$$

The following important result has been used to obtain the final expression provided in Equation (120):

$$[\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e][\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]^\dagger = [\mathbf{r} - \mathbf{r}_e - \mathbf{r} + \mathbf{E}_c(\mathbf{r})][\mathbf{r} - \mathbf{r}_e - \mathbf{r} + \mathbf{E}_c(\mathbf{r})]^\dagger = \mathbf{C}_e^{rr} + \mathbf{C}_c^{rr}. \quad (121)$$

As indicated in Equation (120), the initial covariance matrix $\mathbf{C}_e^{rr}$ for the experimentally measured responses is multiplied by the matrix $[\mathbf{I} - (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} \mathbf{C}_e^{rr}]$, which means that the variances contained on the diagonal of the best-estimate matrix $\mathbf{C}_{be}^{rr}$ are smaller than the experimentally measured variances contained in $\mathbf{C}_e^{rr}$. Hence, the incorporation

of experimental information reduces the predicted best-estimate response variances in $\mathbf{C}_{be}^{rr}$ by comparison to the measured variances contained a priori in $\mathbf{C}_e^{rr}$. Since the components of the matrix $\mathbf{C}_{be}^{rr}$ contain high-order sensitivities, the formula presented in Equation (120) generalizes all of the previous formulas of this type found in data adjustment/assimilation procedures published to date (which contain at most first-order sensitivities).

The expression of the “best-estimate” posterior parameter covariance matrix $\mathbf{C}_{be}^{\alpha\alpha}$ for the best-estimate parameters $\mathbf{a}_{be}$ is derived by using the result given in Equation (119) to obtain:

$$\mathbf{C}_{be}^{\alpha\alpha} \triangleq \int_{D_p} (\mathbf{a} - \mathbf{a}_{be})(\mathbf{a} - \mathbf{a}_{be})^\dagger p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = \mathbf{C}^{\alpha\alpha} - \mathbf{C}_c^{ar} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} \mathbf{C}_c^{ra}. \quad (122)$$

The matrices $\mathbf{C}^{\alpha\alpha}$ and $\mathbf{C}_c^{ar} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} \mathbf{C}_c^{ra}$ are symmetric and positive definite. Therefore, the subtraction indicated in Equation (122) implies that the components of the main diagonal of $\mathbf{C}_{be}^{\alpha\alpha}$ must have smaller values than the corresponding elements of the main diagonal of $\mathbf{C}^{\alpha\alpha}$. In this sense, the combination of computational and experimental information has reduced the best-estimate parameter variances on the diagonal of $\mathbf{C}_{be}^{\alpha\alpha}$. Since the components of the matrices $\mathbf{C}^{\alpha\alpha}$, $\mathbf{C}_c^{ar}$, and $\mathbf{C}_c^{ra}$ contain high-order response sensitivities, the formula presented in Equation (122) generalizes all of the previous formulas of this type found in data adjustment/assimilation procedures published to date (which contain at most first-order sensitivities).

The expression of the “best-estimate” posterior parameter correlation matrix $\mathbf{C}_{be}^{ra}$ and/or its transpose $\mathbf{C}_{be}^{ar}$, for the best-estimate parameters $\mathbf{a}_{be}$ and best-estimate responses $\mathbf{r}_{be}$, are derived by using the results given in Equations (118) and (119) to obtain the following expressions:

$$\mathbf{C}_{be}^{ar} \triangleq \int_{D_p} (\mathbf{a} - \mathbf{a}_{be})(\mathbf{r} - \mathbf{r}_{be})^\dagger p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = \mathbf{C}_c^{ar} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} \mathbf{C}_e^{rr}; \quad (123)$$

$$\mathbf{C}_{be}^{ra} \triangleq \int_{D_p} (\mathbf{r} - \mathbf{r}_{be})(\mathbf{a} - \mathbf{a}_{be})^\dagger p_p(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} = \mathbf{C}_e^{rr} (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} \mathbf{C}_c^{ra} = (\mathbf{C}_{be}^{ar})^\dagger. \quad (124)$$

Since the components of the matrices $\mathbf{C}_c^{ar}$ and $\mathbf{C}_c^{ra}$ contain high-order sensitivities, the formulas presented in Equations (123) and (124) generalize all of the previous formulas of this type found in data adjustment/assimilation procedures published to date (which contain at most first-order sensitivities).

It is important to note from the results shown in (118)–(124) that the computation of the best estimate parameter and response values, together with their corresponding best-estimate covariance matrices, only requires the computation of $(\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1}$, which entails the inversion of a matrix of size $TR \times TR$. This is computationally very advantageous, since $TR \ll TP$, i.e., the number of responses is much less than the number of model parameters in the overwhelming majority of practical situations.

The minimum value, $Q_{\min} = Q(\mathbf{a}_{be}, \mathbf{r}_{be})$, of the quadratic form $Q(\mathbf{a}, \mathbf{r})$, takes on the following expression, which is obtained by using Equation (115):

$$Q_{\min} = \chi^2 = \begin{pmatrix} \mathbf{r}_{be} - \mathbf{E}_c(\mathbf{r}) \\ \mathbf{a}_{be} - \mathbf{a}^0 \end{pmatrix}^\dagger \begin{pmatrix} \mathbf{C}_c^{rr} & \mathbf{C}_c^{ra} \\ \mathbf{C}_c^{ar} & \mathbf{C}^{\alpha\alpha} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{r}_{be} - \mathbf{E}_c(\mathbf{r}) \\ \mathbf{a}_{be} - \mathbf{a}^0 \end{pmatrix} + (\mathbf{r}_{be} - \mathbf{r}_e)^\dagger (\mathbf{C}_e^{rr})^{-1} (\mathbf{r}_{be} - \mathbf{r}_e) = [\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]^\dagger (\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1} [\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]. \quad (125)$$

As the expression obtained in Equation (125) indicates, the quantity $Q_{\min}$ represents the square of the length of the vector $[\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]$, measuring (in the corresponding metric) the deviations between the experimental and nominally computed responses. The quantity $Q_{\min}$ can be evaluated directly from the given data (i.e., model parameters and computed and measured responses, together with their original uncertainties) after having computed the matrix $(\mathbf{C}_e^{rr} + \mathbf{C}_c^{rr})^{-1}$. As the dimension of the vector $[\mathbf{E}_c(\mathbf{r}) - \mathbf{r}_e]$ indicates, the number of degrees of freedom characteristic of the calibration under consideration is equal to the number TR of experimental responses. It is important to note that $Q_{\min}$ is independent of calibrating (or adjusting) the original data. In the extreme case of absence of experimental responses, no actual calibration takes place. An actual calibration (adjustment) occurs only when including at least one experimental response.

The variate $Q_{\min}$ follows a χ^2 -distribution with TR degrees of freedom, where TR denotes the total number of experimental responses considered in the assimilation/calibration (adjustment) procedure. Thus, the quantity $Q_{\min}$ is the “ χ^2 of the calibration at hand” and can be used as an indicator of the agreement between the computed and measured responses (which depend on the model parameters). Recall that the χ^2 -distribution is a measure of the deviation of a “true distribution” (in this case, the distribution of experimental responses) from the hypothetic one (in this case, a MaxEnt distribution constructed from the available information regarding mean values and correlations/covariances). For predictive modeling it is important to assess if: (i) the response and data measurements are free of gross errors (blunders such as wrong settings, mistaken readings, etc.), and (ii) the measurements are consistent with the assumptions regarding the respective means, variances, and covariances.

3.5. Characteristics of the BERRU-PM-2+ Methodology: Summary

The end-results of BERRU-PM-2+ methodology introduced in Section 3.4, above, are the expressions presented in Equations (118)–(125). An examination of these expressions highlights the fact that they can be utilized both for “forward/direct predictive modeling” and for “inverse predictive modeling”. The “forward” or “direct problem” solves the “parameter-to-output” mapping that describes the “cause-to-effect” relationship in the physical process being modeled. The “inverse problem” attempts to solve the “output-to-parameters” mapping. In particular, “measurement problems” are “inverse” to the direct problem, in that they seek to determine (from measurements) the properties of the host medium (e.g., composition, geometry, including internal interfaces) or the properties of the source (e.g., strength, location, direction) and/or the size of the medium on its boundaries. Such inverse problems are encountered in fields as diverse as astrophysics (in which one measures the intensity and spectral distribution of light in order to infer properties of stars), nuclear medicine (where radioisotopes are injected into patients and the radiation emitted is used in diagnostics to reconstruct body properties, e.g., tumors), non-destructive fault detection in materials, underground (oil, water) logging, and detection of sensitive materials. Some authors further group such inverse problems into “invasive”, when the interior particle distribution is accessible for measurements, as opposed to “non-invasive” ones, in which only particle distributions on the boundaries of (or exterior to) the medium can be measured.

Inverse problems are fundamentally ill-posed and/or ill-conditioned, unstable to uncertainties in the model parameters and/or the experimental measurements. In particular, inverse problems involving differential operators are ill-posed because the differentiation operator is not continuous with respect to any physically meaningful observation topology. The existence of a solution for an inverse problem is in most cases secured by defining the data space to be the set of solutions to the direct problem. This approach may fail if the data is incomplete, perturbed, or noisy. If the uniqueness of a solution cannot be secured from the given data, additional data and/or a priori knowledge about the solution

need to be used to restrict the set of admissible solutions. In particular, stability of the solution is the most difficult to ensure and verify. If an inverse problem fails to be stable, then small round-off errors or noise in the data will amplify to a degree that renders a computed solution useless. The procedures used to solve approximately an ill-posed problem are called “regularization” procedures and are customarily categorized as “explicit” or “implicit”; see, e.g., [47–49]. The historically older explicit methods attempt to manipulate the forward model equations in conjunction with measurements in order to estimate explicitly the unknown source and/or other unknown characteristics of the medium. On the other hand, implicit methods combine measurements with repeated solutions of the direct problem obtained with different values of the unknowns, iterating until an a priori selected functional, usually representing the user-defined “goodness of fit” between measurements and direct computations, is reduced to a value deemed to be “acceptable” by the user.

Since the framework of BERRU-PM methodology (including its present extension, BERRU-PM-2+) comprises the combined phase-space of parameters and responses, it can be used for solving both forward/direct and inverse problems. The solution of forward/direct problems is provided by the expression for the predicted best-estimate response given in Equation (116) together with the corresponding reduced predicted uncertainties provided by Equation (120). The remaining expressions for the best-estimate predicted parameters and their reduced predicted uncertainties provided in Equations (117), (122), along with the parameter-response correlations provided in Equation (123) and the “goodness of fit indicator” provided by Equation (125) are additional features which the BERRU-PM methodology provides. Conversely, the solution of “inverse problems” is provided by the expressions for the best-estimate predicted parameters and their reduced predicted uncertainties provided in Equations (117) and (122). The superior accuracy provided by the BERRU-PM methodology (by comparison to the traditional methods based on the minimization of chi-square/least-squares-type functionals for solving inverse problems) has been demonstrated in [16] for the inverse prediction, from detector responses in the presence of counting uncertainties, of the thickness of a homogeneous slab of material containing uniformly distributed gamma-emitting sources. In particular, it has been shown in [16] that for optically very thin slabs, both the traditional chi-square-minimization methods and the BERRU-PM methodology predict the slab’s thickness accurately. For optically thick slabs, however, the traditional inverse-problem methods based on the minimization of chi-square-type functionals fail to predict the slab’s thickness, while the BERRU-PM methodology correctly predicts the slab’s actual physical thickness when precise experimental results are assimilated, while also predicting the physically correct response within the selected precision criterion.

As shown in Appendix A, the expressions of the moments of the joint distribution of computed responses and model parameters defined in Equations (70)–(72), in terms of the moments of the distribution of model parameters, explicitly include the third- and higher-order sensitivities of responses with respect to the model parameters, as well as the systematic and random errors that would stem from incomplete modeling and numerical truncations. The expressions provided in Appendix A also indicate the path for incorporating explicitly higher-order sensitivities, if available. Since BERRU-PM-2+ mathematical framework introduced in Section 3 of this work enables the use of higher-order sensitivities, it is expected that the BERRU-PM-2+ methodology will improve considerably the already unparalleled efficiency and accuracy of the extant BERRU-PM methodology.

4. Discussion, Conclusions and Outlook

Although they are not identical, the data adjustment and data-assimilation methodologies share several fundamental common features which stem from their common root in least-square minimization procedures, so the expressions of the end-results produced by these methodologies look similar. Neither of these methodologies use the complete first-order information, however. The data-adjustment methodology is incomplete in

first-order because it has no provisions for incorporating the first-order correlations between the model parameters and the computed responses, which are defined by Equation (71) and which can be readily computed, as shown in Appendix A. Even though the provision for incorporating the correlations defined in Equation (4) between measured responses and model parameters is extant, this provision has not been used in applications thus far because of lack of actual data.

The applicability of the data-adjustment methodology is also restricted by the domain of applicability of the “generalized linear least-squares methodology” (GLLSM), which requires that: (i) the uncertainty of the calculated response (e.g., neutron multiplication factor) be dominated by uncertainties in the data (e.g., cross sections, neutron multiplicities, fission spectra), having negligible systematic errors; and (ii) the data uncertainties be sufficiently small for approximating the response (e.g., neutron multiplication factor) by a first-order Taylor-series expansion about the values of the given data. The analysis performed in [50,51] indicated that neither condition is fulfilled in general, since (i) the linear perturbation assumption underlying the GLLSM predictions breaks down for relative standard deviations (in the data) larger than ca. 1%; and (ii) systematic errors larger than ca. 1% also cause the break-down of the GLLSM predictions, since such systematic errors are not explicitly considered within the GLLSM formalism. Consequently, for real application cases, which are always characterized by systematic computational errors, the work reported in [50,51] indicates that true coverage probabilities of GLLSM confidence intervals for the computational bias may strongly differ from their expected confidence levels, particularly if a chi-square filter is used. The user-defined “cost functional” represented by Equation (1), which is minimized in the least-square sense, cannot be extended in an obvious manner to incorporate higher-order correlations and sensitivities. A fundamentally different type of “cost functional” would need to be introduced for incorporating higher-order correlations and sensitivities, the form of which has not been produced yet in the literature.

The data-assimilation methodology is also incomplete in first-order because it has no provisions for incorporating the first-order correlations between the model parameters and either the computed or the measured responses; specifically, neither the information represented by Equation (71) nor the information represented by Equation (97) can be incorporated in the user-defined cost functional underlying the extant data-assimilation methodologies. On the other hand, the currently most advanced 4D-VAR data-assimilation methodologies (including the extended/ensemble Kalman filter, second-order filter and variants thereof; see, e.g., references [11–14]) allow the incorporation of some partial second-order information through the so-called “Hessian-vector product” which appears as an additional term in the model dynamics/forecast step and the data assimilation step. Thus, although the most advanced data-assimilation methods can be considered to be “more inclusive” —in the sense of incorporating some partial second-order information— than the most advanced data-adjustment methods, neither of these methodologies are complete regarding the incorporation of all extant first- and/or second-order information, and neither methodology provides a mechanism (even a theoretical one) for incorporating response sensitivities (to model parameters) higher than first-order and/or incorporating correlations among responses and/or parameters higher than second-order. Of course, the predictions of the data adjustment or assimilation methodologies are also limited to first-order mean values and second-order covariances/correlations for parameters or responses.

The first-order (in response sensitivities) BERRU-PM methodology [16] is based on the MaxEnt principle rather than least-squares minimization, thereby eliminating the user-introduced preference (bias) in the definition of the functional “describing the discrepancies between computations and measurements” which is to be minimized. Although the end-products of the BERRU-PM procedure look superficially similar to those produced by the data adjustment and/or data assimilation procedure, the BERRU-PM procedure differs fundamentally from the data adjustment/assimilation procedures, both

regarding the actual contents of its end-products and also the potential for further generalization to higher-order, which are inherent to the BERRU-PM methodology [16] but are lacking within the frameworks of the other aforementioned methodologies.

Extending the BERRU-PM methodology [16], the BERRU-PM-2+ methodology presented in Section 3 of this work also employs the MaxEnt principle instead of a user-defined “cost functional to be minimized”. Employing the MaxEnt principle not only eliminates the impact of the user-defined functional which predetermines the outcome of the respective methodology’s predictions, but also makes it possible to incorporate the complete first- and second-order information (sensitivities and correlations among parameters and responses, computed and measured) by considering the complete joint-phase-space of parameters, computed and measured responses. By setting its mathematical framework in this most-inclusive phase-space, the BERRU-PM-2+ methodology can be used for solving both forward/direct problems and inverse problems, not only for applications related to energy-systems but for applications in all fields that combine computational and experimental information to produce best-estimate predictions for results of interest (responses) while also calibrating/improving the models of the underlying physical systems. Furthermore, the novel MaxEnt-based BERRU-PM-2+ methodology introduced in Section 3 explicitly includes the third- and fourth-order response sensitivities to model parameters, and clearly indicates the path for incorporating explicitly higher-order sensitivities, if available. The BERRU-PM-2+ expressions also comprise quantities representing the systematic and random errors that would stem from incomplete modeling and numerical truncations. It is therefore expected that the BERRU-PM-2+ methodology will further improve the already unparalleled efficiency and accuracy of the extant first-order BERRU-PM methodology.

The data-assimilation methodology is widely used in the atmospheric and geophysical sciences. On the other hand, predictive modeling applications to energy-related fields are very few, with the highly notable exception of applications in the nuclear energy field, where the original data-adjustment methodology was conceived. The perspective on the available predictive modeling methodologies presented in this work and, in particular, the introduction in this work of the BERRU-PM-2+ methodology, also aim to motivate a much wider use of predictive modeling in the energy sciences for obtaining best-estimate results with reduced uncertainties. In practice, predictive modeling methodologies should be an integral part of model validation methodologies to be applied prior to model qualification and licensing considerations.

Ongoing work aims at extending the BERRU-PM-2+ methodology to provide information about the joint parameters/responses posterior distribution’s third-order (skewness) and fourth-order (kurtosis) moments. The incorporation of third- and fourth-order information (i.e., high-order response sensitivities along with skewness and kurtosis of parameters) will greatly increase the range of validity of predicted results, as well as reduce the uncertainties in these predictions. In particular, the envisaged new fourth-order (BERRU-PM-4) predictive modeling methodology will significantly impact the activities of model calibration and model validation against experiments. Further work will aim at quantifying the impact of ignorance (lack of knowledge) caused by non-modeled/missed phenomena in computational models of physical/biological systems. It is envisaged to demonstrate the potential impact of this new, high-order, comprehensive predictive modeling methodology by applications to reduce uncertainties in modeling predictions for energy and related systems.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

This appendix presents the expressions, up to and including fourth-order in the standard deviations of parameters, of the moments of the computed responses in the combined phase-space of parameters and computed responses.

Recall from Equations (67) and (68) that the covariance, $\text{cov}(\alpha_i, \alpha_j)$, of two parameters, α_i and α_j , is formally defined as follows:

$$\text{cov}(\alpha_i, \alpha_j) \triangleq \int_{D_\alpha} \delta\alpha_i \delta\alpha_j p_\alpha(\mathbf{a}) d\mathbf{a} \triangleq \rho_{ij} \sigma_i \sigma_j; i, j = 1, \dots, TP.$$

The quantity ρ_{ij} denotes the correlation between the parameters α_i and α_j , while σ_i denotes the standard deviation of parameter α_i and σ_j denotes the standard deviation of parameter α_j , respectively. The variance, $\text{var}(\alpha_i)$, of a parameter α_i , is formally defined as the following particular case of Equation (67):

$$\text{var}(\alpha_i) \triangleq \int_{D_\alpha} (\delta\alpha_i)^2 p_\alpha(\mathbf{a}) d\mathbf{a} = \sigma_i^2; i = 1, \dots, TP.$$

The third-order correlation, t_{ijk} , of three parameters ($\alpha_i, \alpha_j, \alpha_k$) is formally defined as follows:

$$t_{ijk} \sigma_i \sigma_j \sigma_k \triangleq \int_{D_\alpha} \delta\alpha_i \delta\alpha_j \delta\alpha_k p_\alpha(\mathbf{a}) d\mathbf{a}; i, j, k = 1, \dots, TP. \quad (\text{A1})$$

The fourth-order parameter moment, $\mu_4^{j_1 j_2 j_3 j_4}$, and the associated fourth-order correlation $q_{j_1 j_2 j_3 j_4}$ among four parameters, are defined as follows:

$$q_{ijk\ell} \sigma_i \sigma_j \sigma_k \sigma_\ell \triangleq \int_{D_\alpha} \delta\alpha_i \delta\alpha_j \delta\alpha_k \delta\alpha_\ell p_\alpha(\mathbf{a}) d\mathbf{a}; i, j, k, \ell = 1, \dots, TP. \quad (\text{A2})$$

The expression of the fourth-order Taylor-series expansion of the computed response, which was provided in Equation (69), is re-written below in order to highlight the expressions of the contributions to various order in parameter deviations $\delta\alpha_j$:

$$r_k^{(1)}(\mathbf{a}) = r_k(\mathbf{a}^0) + \sum_{j_1=1}^{TP} \left\{ \frac{\partial r_k(\mathbf{a})}{\partial \alpha_{j_1}} \right\}_{\mathbf{a}^0} \delta\alpha_{j_1} + \varepsilon_k^{(1)}; \quad (\text{A3})$$

$$r_k^{(2)}(\mathbf{a}) = r_k^{(1)}(\mathbf{a}) + \frac{1}{2} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \left\{ \frac{\partial^2 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2}} \right\}_{\mathbf{a}^0} \delta\alpha_{j_1} \delta\alpha_{j_2} + \varepsilon_k^{(2)}; \quad (\text{A4})$$

$$r_k^{(3)}(\mathbf{a}) = r_k^{(2)}(\mathbf{a}) + \frac{1}{3!} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \left\{ \frac{\partial^3 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3}} \right\}_{\mathbf{a}^0} \delta\alpha_{j_1} \delta\alpha_{j_2} \delta\alpha_{j_3} + \varepsilon_k^{(3)} \quad (126)$$

$$r_k^{(4)}(\mathbf{a}) = r_k^{(3)}(\mathbf{a}) + \frac{1}{4!} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \sum_{j_4=1}^{TP} \left\{ \frac{\partial^4 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3} \partial \alpha_{j_4}} \right\}_{\mathbf{a}^0} \delta\alpha_{j_1} \delta\alpha_{j_2} \delta\alpha_{j_3} \delta\alpha_{j_4} + \varepsilon_k^{(4)} \quad (\text{A6})$$

Appendix A.1. Fourth-Order Approximate Expressions for the Expected Value of the Computed Response

The approximations to the expected value, $E_c(r_k)$, of the computed response $r_k(\mathbf{a})$ are obtained by integrating formally the expression provided in Equation (69) over the unknown distribution $p_\alpha(\mathbf{a})$ to obtain the following expressions for the first-order through fourth-order approximate expectation, $E_c^{(n)}(r_k)$, $n = 1, \dots, 4$:

$$E_c^{(1)}(r_k) = E_c^{(0)}(r_k) \triangleq r_k(\mathbf{a}^0) + \hat{\varepsilon}_k^{(2)}; \quad (\text{A7})$$

$$E_c^{(2)}(r_k) = E_c^{(1)}(r_k) + \frac{1}{2} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \left\{ \frac{\partial^2 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2}} \right\}_{\mathbf{a}^0} \rho_{j_1 j_2} \sigma_{j_1} \sigma_{j_2} + \hat{\varepsilon}_k^{(3)}; \quad (\text{A8})$$

$$E_c^{(3)}(r_k) = E_c^{(2)}(r_k) + \frac{1}{6} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \left\{ \frac{\partial^3 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3}} \right\}_{\mathbf{a}^0} t_{j_1 j_2 j_3} \sigma_{j_1} \sigma_{j_2} \sigma_{j_3} + \hat{\varepsilon}_k^{(4)}, \quad (\text{A9})$$

$$E_c^{(4)}(r_k) = E_c^{(3)}(r_k) + \frac{1}{4!} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \sum_{j_4=1}^{TP} \left\{ \frac{\partial^4 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3} \partial \alpha_{j_4}} \right\}_{\mathbf{a}^0} q_{j_1 j_2 j_3 j_4} \sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4} + \hat{\varepsilon}_k^{(5)}. \quad (\text{A10})$$

As indicated in Equation (A8), the second-order sensitivities contribute the leading correction terms to the response's expected value, causing it to differ from the response's computed value, $r_k(\mathbf{a}^0)$. The quantities $\hat{\varepsilon}_k^{(n)} = O(\sigma^n)$, $n = 2, 3, 4, 5$, which appear in Equations (A7)–(A10), represent the approximate mean values of the respective-order error term ε_k , which appeared in Equation (69). The systematic error in each error term $\hat{\varepsilon}_k^{(n)}$, $n = 2, 3, 4, 5$, is of order $O(\sigma^n)$; the actual numerical value for each error term $\hat{\varepsilon}_k^{(n)}$, $n = 2, 3, 4, 5$, is assigned by the user.

Appendix A.2. Fourth-Order Approximate Expressions for the Correlations Between Computed Responses and Model Parameters

The correlation between a model parameter and a computed response was defined in Equation (71). The approximations to the correlation between a model parameter α_i and a computed response $r_k(\mathbf{a})$ are obtained by multiplying the expression provided in Equation (69) by $\delta \alpha_i$ and by the corresponding expression of $[r_k^{(n)}(\mathbf{a}) - E_c^{(n)}(r_k)]$, and then integrating the resulting expression formally over the unknown distribution $p_c(\mathbf{a}, \mathbf{r})$. This sequence of mathematical operations yields the following expressions for the respective approximate correlations,

$$\text{cov}(\alpha_i, r_k^{(n)}) \triangleq \int_{D_a} \delta \alpha_i [r_k^{(n)}(\mathbf{a}) - E_c^{(n)}(r_k)] p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}, \quad n = 1, \dots, 4:$$

$$\text{cov}(\alpha_i, r_k^{(1)}) = \sum_{j=1}^{TP} \left\{ \frac{\partial r_k(\mathbf{a})}{\partial \alpha_j} \right\}_{\mathbf{a}^0} \text{cov}(\alpha_i, \alpha_j) + \varepsilon_{ik}^{(3)}; \quad (\text{A11})$$

$$\text{cov}(\alpha_i, r_k^{(2)}) = \text{cov}(\alpha_i, r_k^{(1)}) + \frac{1}{2} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \left\{ \frac{\partial^2 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2}} \right\}_{\mathbf{a}^0} t_{i, j_1 j_2} \sigma_{j_1} \sigma_{j_2} + \varepsilon_{ik}^{(4)}; \quad (\text{A12})$$

$$\text{cov}(\alpha_i, r_k^{(3)}) = \text{cov}(\alpha_i, r_k^{(2)}) + \frac{1}{6} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \left\{ \frac{\partial^3 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3}} \right\}_{\mathbf{a}^0} q_{i, j_1 j_2 j_3} \sigma_{j_1} \sigma_{j_2} \sigma_{j_3} + \varepsilon_{ik}^{(5)}; \quad (\text{A13})$$

The error terms $\varepsilon_{ik}^{(n)} = O(\sigma^n)$, $n = 3, 4, 5$, which appear in Equations (A11)–(A13), respectively, contain random errors, as well as systematic errors of order $O(\sigma^n)$, $n = 3, 4, 5$; their respective values are assigned by the user.

Appendix A.3. Fourth-Order Approximate Expressions for the Correlations between Computed Responses

The covariance, $\text{cov}(r_k, r_\ell) \triangleq \int_{D_c} [r_k(\mathbf{a}) - E_c(r_k)][r_\ell(\mathbf{a}) - E_c(r_\ell)] p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r}$, between two computed responses r_k and r_ℓ has been defined in Equation (72). It is shown below that the n th-order approximation of the covariance $\text{cov}(r_k, r_\ell)$ can be defined in a either a “consistent” way, in which case the respective approximation is denoted as

$\left\{cov\left(r_k^{(n)}, r_\ell^{(n)}\right)\right\}_{inc}$, or an “inconsistent” was, in which case the respective approximation is denoted as $\left\{cov\left(r_k^{(n)}, r_\ell^{(n)}\right)\right\}_{con}$, for $n = 1, 2, \dots$, as follows:

$$\left\{cov\left(r_k^{(n)}, r_\ell^{(n)}\right)\right\}_{inc} \triangleq \int_{D_c} \left[r_k^{(n)}(\mathbf{a}) - E_c^{(n)}(r_k)\right] \left[r_\ell^{(n)}(\mathbf{a}) - E_c^{(n)}(r_\ell)\right] p_c(\mathbf{a}, \mathbf{r}) d\mathbf{a} d\mathbf{r} + O(\sigma^{2n}) \quad (\text{A14})$$

$$\left\{cov\left(r_k^{(n)}, r_\ell^{(n)}\right)\right\}_{con} \triangleq cov\left(r_k^{(n+1)}, r_\ell^{(n+1)}\right) - O(\sigma^{n+1}), \quad n = 1, 2, \dots \quad (\text{A15})$$

The standard deviation, $SD\left[r_k^{(n)}(\mathbf{a})\right]$, of $r_k^{(n)}(\mathbf{a})$ is obtained by setting $k = \ell$ in either (A14) or (A15) and taking the square root of the resulting expression, respectively, to obtain:

$$\left\{SD\left(r_k^{(n)}\right)\right\}_{inc} \triangleq \sqrt{\left\{var\left(r_k^{(n)}, r_k^{(n)}\right)\right\}_{inc}} + O(\sigma^n), \quad n = 1, 2, \dots \quad (\text{A16})$$

$$\left\{SD\left(r_k^{(n)}\right)\right\}_{con} \triangleq \sqrt{\left\{var\left(r_k^{(n)}, r_k^{(n)}\right)\right\}_{con}} + O(\sigma^{n+1/2}), \quad n = 1, 2, \dots \quad (\text{A17})$$

Thus, as indicated by the right-side of Equation (A16), the standard deviation $\left\{SD\left(r_k^{(n)}\right)\right\}_{inc}$ corresponding to the inconsistent variance $\left\{var\left(r_k^{(n)}, r_k^{(n)}\right)\right\}_{inc}$ will not comprise the complete number of terms that contain the n th-order parameter standard deviations. In contradistinction to $\left\{SD\left(r_k^{(n)}\right)\right\}_{inc}$, the consistent standard deviation, $\left\{SD\left(r_k^{(n)}\right)\right\}_{con}$, which corresponds to the consistent variance $\left\{var\left(r_k^{(n)}, r_k^{(n)}\right)\right\}_{con}$, will comprise all of the terms that contain the n th-order parameter standard deviations.

The first-order approximation of the covariance $cov(r_k, r_\ell)$ has the following expression:

$$cov\left(r_k^{(1)}, r_\ell^{(1)}\right) = \sum_{i=1}^{TP} \sum_{j=1}^{TP} \left\{ \frac{\partial r_k(\mathbf{a})}{\partial \alpha_i} \frac{\partial r_\ell(\mathbf{a})}{\partial \alpha_j} \right\}_{\mathbf{a}^0} cov(\alpha_i, \alpha_j) + O(\sigma^3). \quad (\text{A18})$$

Notably, the expression of $cov\left(r_k^{(1)}, r_\ell^{(1)}\right)$ provided in (A18) is consistent both in the highest-order (in this case: first-order) of sensitivities and also in the highest-order of parameter standard deviation (in this case, second-order) since all of the terms involving the product $\sigma_i \sigma_j$ are consistently included (i.e., none are missing) in the expression provided. Therefore, the standard deviation of the computed response $r_k(\mathbf{a})$ will also be correct to first-order in the standard deviations of the parameters, i.e.,

$$SD\left(r_k^{(1)}\right) = \left\{ \sum_{i=1}^{TP} \sum_{j=1}^{TP} \left\{ \frac{\partial r_k(\mathbf{a})}{\partial \alpha_i} \frac{\partial r_k(\mathbf{a})}{\partial \alpha_j} \right\}_{\mathbf{a}^0} cov(\alpha_i, \alpha_j) \right\}^{\frac{1}{2}} + O(\sigma^{3/2}). \quad (\text{A19})$$

The second-order approximation of the covariance $cov(r_k, r_\ell)$ has the following expression:

$$\begin{aligned}
\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{inc} &= cov\left(r_k^{(1)}, r_\ell^{(1)}\right) \\
&+ \frac{1}{2} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \left\{ \frac{\partial^2 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2}} \frac{\partial r_\ell(\mathbf{a})}{\partial \alpha_{j_3}} + \frac{\partial r_k(\mathbf{a})}{\partial \alpha_{j_1}} \frac{\partial^2 r_\ell(\mathbf{a})}{\partial \alpha_{j_2} \partial \alpha_{j_3}} \right\}_{\mathbf{a}^0} t_{j_1 j_2 j_3} \sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \\
&+ \frac{1}{4} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \sum_{j_4=1}^{TP} \left\{ \frac{\partial^2 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2}} \frac{\partial^2 r_\ell(\mathbf{a})}{\partial \alpha_{j_3} \partial \alpha_{j_4}} \right\}_{\mathbf{a}^0} (q_{j_1 j_2 j_3 j_4} - \rho_{j_1 j_2} \rho_{j_3 j_4}) \sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4} \\
&+ O\left(\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4}\right).
\end{aligned} \tag{A20}$$

Notably, the expression of $\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{inc}$ provided in Equation (A20) is consistent in the second-order of sensitivities but is inconsistent in the fourth-order of standard deviations of parameters since it has errors of order $O\left(\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4}\right)$. The missing terms of order $O\left(\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4}\right)$, which involve third-order response sensitivities to parameters, are provided in the expression below for the “consistent second-order parameter correlations” $\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{con}$:

$$\begin{aligned}
\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{con} &= \left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{inc} + \frac{1}{6} \sum_{j_1=1}^{TP} \sum_{j_2=1}^{TP} \sum_{j_3=1}^{TP} \sum_{j_4=1}^{TP} \left\{ \frac{\partial^3 r_k(\mathbf{a})}{\partial \alpha_{j_1} \partial \alpha_{j_2} \partial \alpha_{j_3}} \frac{\partial r_\ell(\mathbf{a})}{\partial \alpha_{j_4}} \right. \\
&\left. + \frac{\partial r_k(\mathbf{a})}{\partial \alpha_{j_1}} \frac{\partial^3 r_\ell(\mathbf{a})}{\partial \alpha_{j_2} \partial \alpha_{j_3} \partial \alpha_{j_4}} \right\}_{\mathbf{a}^0} q_{j_1 j_2 j_3 j_4} \sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4} + O\left(\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4} \sigma_{j_5}\right).
\end{aligned} \tag{A21}$$

Therefore, if the third-order sensitivities are available, then the expression provided in Equation (A21) should be used, since it includes all of the fourth-order terms containing the products $\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4}$. Consequently, the consistent second-order standard deviation $\left\{SD\left(r_k^{(2)}\right)\right\}_{con} \triangleq \sqrt{\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{con}}$ will not have any second-order errors in the parameter standard deviations, i.e.,

$$\left\{SD\left(r_k^{(2)}\right)\right\}_{con} \triangleq \sqrt{\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{con}} + O\left(\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4} \sigma_{j_5}\right)^{1/2}. \tag{A22}$$

On the other hand, the second-order inconsistent standard deviation $\left\{SD\left(r_k^{(2)}\right)\right\}_{inc} \triangleq \sqrt{\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{inc}}$ will have second-order errors in the parameter standard deviations, i.e.,

$$\left\{SD\left(r_k^{(2)}\right)\right\}_{inc} \triangleq \sqrt{\left\{cov\left(r_k^{(2)}, r_\ell^{(2)}\right)\right\}_{inc}} + O\left(\sigma_{j_1} \sigma_{j_2} \sigma_{j_3} \sigma_{j_4}\right)^{1/2}. \tag{A23}$$

References

1. Humi, I.; Wagschal, J.J.; Yeivin, Y. Multi-group constants from integral data. In Proceedings of the 3rd International Conference on the Peaceful Uses of Atomic Energy, Geneva, Switzerland, 31 August–9 September 1964; Volume 2, p. 398.
2. Cecchini, G.; Farinelli, U.; Gandini, A.; Salvatores, A. Analysis of integral data for few-group parameter evaluation of fast reactors. In Proceedings of the 3rd International Conference on the Peaceful Uses of Atomic Energy, Geneva, Switzerland, 31 August–9 September 1964; Volume 2.
3. Usachhev, L.N. Perturbation theory for the breeding ratio and for other number ratios pertaining to various reactor processes. *J. Nucl. Energy. Parts A/B React. Sci. Technol.* **1964**, *18*, 571. [https://doi.org/10.1016/0368-3230\(64\)90142-9](https://doi.org/10.1016/0368-3230(64)90142-9).
4. Rowlands, J. The production and performance of the adjusted cross-section set FGL5. In Proceedings of the International Symposium on Physics of Fast Reactors, Tokyo, Japan, 16–19 October 1973.
5. Gandini, A.; Petilli, M. *AMARA: A Code Using the Lagrange Multipliers Method for Nuclear Data Adjustment*; CNEN-RI/FI(73)39; Comitato Nazionale Energia Nucleare: Casaccia/Rome, Italy, 1973.

6. Kuroi, H.; Mitani, H. Adjustment to cross-section data to fit integral experiments by least squares method. *J. Nucl. Sci. Technol.* **1975**, *12*, 663. <https://doi.org/10.3327/jnst.12.663>.
7. Dragt, J.B.; Dekker, J.W.M.; Gruppelaar, H.; Janssen, A.J. Methods of adjustment and error evaluation of neutron capture cross sections. *Nucl. Sci. Eng.* **1977**, *62*, 11. <https://doi.org/10.13182/NSE77-3>.
8. Weisbin, C.R.; Oblo, E.M.; Marable, J.H.; Peelle, R.W.; Lucius, J.L. Application of sensitivity and uncertainty methodology to fast reactor integral experiment analysis. *Nucl. Sci. Eng.* **1978**, *66*, 307. <https://doi.org/10.13182/NSE78-3>.
9. Barhen, J.; Cacuci, D.G.; Wagschal, J.J.; Bjerke, M.A.; and Mullins, C.B.; Uncertainty analysis of time-dependent nonlinear systems: Theory and application to transient thermal hydraulics. *Nucl. Sci. Eng.* **1982**, *81*, 23–44.
10. Cacuci, D.G. Sensitivity theory for nonlinear systems: I. Nonlinear functional analysis approach. *J. Math. Phys.* **1981**, *22*, 2794–2812.
11. Lewis, J.M.; Lakshmivarahan, S.; Dhall, S.K. *Dynamic Data Assimilation: A Least Square Approach*; Cambridge University Press: Cambridge, UK, 2006.
12. Lahoz, W.; Khattatov, B.; Ménard, R. (Eds.) *Data Assimilation: Making Sense of Observations*; Springer: Heidelberg, Germany, 2010.
13. Faragó, I.; Havasi, Á.; Zlatev, Z. (Eds.) *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*; Bentham Science Publishers: Bussum, The Netherlands, 2013.
14. Cacuci, D.G.; Navon, M.I.; Ionescu-Bujor, M. *Computational Methods for Data Evaluation and Assimilation*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2014.
15. Cacuci, D.G. Predictive modeling of coupled multi-physics systems: I. Theory. *Ann. Nucl. Energy* **2014**, *70*, 266–278.
16. Cacuci, D.G. *BERRU Predictive Modeling: Best-Estimate Results with Reduced Uncertainties*; Springer: Berlin, Germany, 2019.
17. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
18. Cacuci, D.G. Second-order adjoint sensitivity analysis methodology (2nd-ASAM) for large-scale nonlinear systems: I. Theory. *Nucl. Sci. Eng.* **2016**, *184*, 16–30.
19. Cacuci, D.G. *The Second-Order Adjoint Sensitivity Analysis Methodology*; CRC Press: Boca Raton, FL, USA, 2018.
20. Cacuci, D.G. The nth-order comprehensive adjoint sensitivity analysis methodology for nonlinear systems (nth-CASAM-N): Mathematical framework. *J. Nucl. Eng.* **2022**, *3*, 163–190. <https://doi.org/10.3390/jne3030010>.
21. Cacuci, D.G. *The nth-Order Comprehensive Adjoint Sensitivity Analysis Methodology: Overcoming the Curse of Dimensionality. Volume I: Linear Systems*. Springer Nature Switzerland; Cham, Switzerland, 2022.
22. SCALE: A Modular Code System for Performing Standardized Computer Analyses for Licensing Evaluation; ORNL/TM 2005/39, Version 6; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2009.
23. Venard, C.; Santamarina, A.; Leclainche, A.; Mournier, C. The R.I.B. Tool for the determination of computational bias and associated uncertainty in the CRISTAL criticality safety package. In Proceedings of ANS Nuclear Criticality Safety Division Topical Meeting (NCSD 2009), Richland, Washington, DC, USA, 13–17 September 2009.
24. Gandin, L.S. *Objective Analysis of Meteorological Fields*; Gridromet: Leningrad, Russia, 1963; (In Russian).
25. Anthes, R.A. Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.* **1974**, *31*, 702. [https://doi.org/10.1175/1520-0469\(1974\)031<0702:DAAIOH>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<0702:DAAIOH>2.0.CO;2).
26. Sasaki, Y.K. A fundamental study of the numerical prediction based on the variational principle. *J. Meteor. Soc. Japan* **1955**, *33*, 262. https://doi.org/10.2151/jmsj1923.33.6_262.
27. Sasaki, Y.K. An objective analysis based on the variational method. *J. Meteor. Soc. Japan* **1958**, *36*, 77. https://doi.org/10.2151/jmsj1923.36.3_77.
28. Lorenc, A.C. Analysis-methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **1986**, *112*, 1177. <https://doi.org/10.1002/qj.49711247414>.
29. Hall, M.C.G.; Cacuci, D.G.; Schlesinger, M.E. Sensitivity analysis of a radiative-convective model by the adjoint method. *J. Atmos. Sci.* **1982**, *39*, 2038–2050.
30. Hall, M.C.G.; Cacuci, D.G. Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci.* **1983**, *40*, 2537–2546.
31. Práger, T.; Kelemen, F.D. Adjoint methods and their application in earth sciences. In *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*; Faragó, I., Havasi, Á., Zlatev, Z., Eds.; Bentham Science Publishers: Bussum, The Netherlands, 2013; Chapter 4, Part A, pp. 203–275.
32. Derber, J.C. The variational four-dimensional assimilation of analysis using filtered models as constraints, Ph.D. Thesis, University of Wisconsin–Madison, Madison, WI, USA, 1985.
33. Derber, J.C. Variational four-dimensional analysis using the quasigeostrophic constraint. *Mon. Weather. Rev.* **1987**, *115*, 998. [https://doi.org/10.1175/1520-0493\(1987\)115<0998:vfdauq>2.0.co;2](https://doi.org/10.1175/1520-0493(1987)115<0998:vfdauq>2.0.co;2).
34. Hoffmann, R.N. A four-dimensional analysis exactly satisfying equations of motion. *Mon. Weather. Rev.* **1986**, *114*, 388. [https://doi.org/10.1175/1520-0493\(1986\)114<0388:AFDAES>2.0.CO;2](https://doi.org/10.1175/1520-0493(1986)114<0388:AFDAES>2.0.CO;2).
35. Navon, I.M.; de Villiers, R. Combined penalty multiplier optimization methods to enforce integral invariants conservation. *Mon. Weather. Rev.* **1983**, *111*, 1228. [https://doi.org/10.1175/1520-0493\(1983\)111<1228:cpmomt>2.0.co;2](https://doi.org/10.1175/1520-0493(1983)111<1228:cpmomt>2.0.co;2).
36. Navon, I.M. A review of variational and optimization methods in meteorology. *Dev. Geomath.* **1986**, *5*, 29–34.
37. Navon, I.M. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn. Atmos. Ocean.* **1998**, *27*, 55–79.

38. Derber, J.C.; Parrish, D.F.; Lord, S.J. The new global operational analysis system at the National Meteorological Center. *Weather. and Forecast* **1991**, *6*, 538. [https://doi.org/10.1175/1520-0434\(1991\)006<0538:TNGOAS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1991)006<0538:TNGOAS>2.0.CO;2).
39. Mahfouf, J.F.; Buizza, R.; Errico, R.M. Strategy for including physical processes in the ECMWF variational data assimilation system. In Proceedings of the ECMWF workshop on non-linear aspects of data assimilation, Reading, UK, 9–11 September 1996.
40. Rabier, F. Overview of global data assimilation developments in numerical weather-prediction centers. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3215. <https://doi.org/10.1256/qj.05.129>.
41. Bellman, R.E. *Dynamic Programming*. Rand Corporation; Princeton University Press: Princeton, NJ, USA, 1957. ISBN 978-0-691-07951-6.
42. Tukey, J.W. *The Propagation of Errors, Fluctuations and Tolerances*; Technical Reports No. 10–12; Princeton University: Princeton, NJ, USA, 1957.
43. Shannon, C.E. A Mathematical theory of communication. *Bell System Tech. J.* **1948**, *27*, 379–423.
44. Evangelou, E.; Zhu, Z.; Smith, R.L. Estimation and prediction for spatial generalized linear mixed models using high order Laplace approximation. *J. Stat. Plan. Inference* **2011**, *141*, 3564–3577. <https://doi.org/10.1016/j.jspi.2011.05.008>
45. Shun, Z. and P. McCullagh. Laplace approximation of high dimensional integrals. *J. R. Stat. Society. Ser. B* **1995**, *57*, 749–760.
46. Tang, Y.; Reid, N. Laplace and saddle point approximations in high dimensions. *arXiv* **2021**, arXiv:2107.10885. <https://doi.org/10.48550/arXiv.2107.10885>.
47. Tichonov, A.N. Regularization of non-linear ill-posed problems. *Dokl. Akad. Nauk.* **1963**, *49*.
48. Tichonov, A.N. Solution of incorrectly formulated problems and the regularization method. *Sov. Math. Dokl.* **1963**, *4*, 1035. <https://doi.org/10.4236/jamp.2016.41017>.
49. Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2005.
50. Hoefer, A.; Buss, O.; Neuber, J.C. How confident can we be in confidence intervals for the computational bias obtained with the generalized linear least squares methodology?—A toy model analysis. In Proceedings of the International Conference on Nuclear Criticality, ICNC 2011, Edinburgh, Scotland, UK, 19–22 September 2011.
51. Hoefer, A.; Buss, O.; Neuber, J.C. Limitations of the generalized linear least squares methodology for bias estimation in criticality safety analysis. In Proceedings of the IAEA International Workshop on Burnup Credit Criticality Calculation Methods and Applications, Beijing, China, 25–28 October 2011. Available online: <https://www.researchgate.net/publication/276271639> (accessed on 5 December 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.