Chapter 1

Introduction to data assimilation and inverse problems

1.1 ▪ Introduction

What exactly is DA? The simplest view is that it is an approach/method for combining observations with model output with the objective of improving the latter. But do we really need DA? Why not just use the observations and average them or extrapolate them (as is done with regression techniques (McPherson, 2001), or just long-term averaging)? The answer is that we want to predict the state of a system, or its future, in the best possible way! For that we need to rely on models. But when models are not corrected periodically by reality, they can be of little value. Thus, we need to fit the model state in an optimal way to the observations, before an analysis or prediction is made. This fitting of a model to observations is a special case (but highly typical) of an *inverse problem*.

According to J. B. Keller [1966], two problems are inverse to each other if "the *formulation* of each involves all or part of the *solution* of the other." One of the two is named the direct problem, whereas the other is the inverse problem. The direct problem is usually the one that we can solve satisfactorily/easily. There is a back-and-forth transmission of information between the two. This is depicted in Figure 1.1, which represents a typical case: we replace the unknown (or partially known) medium by a model that depends on some unknown model parameters, $\boldsymbol{m}$. The inverse problem involves reversing the arrows—by comparing the simulations and the observations (at the array) to find the model parameters. In fact, the direct problem involves going from cause to effect, whereas the inverse problem attempts to go from the effects to the cause.

The comparison between model output and observations is performed by some form of optimization—recall that we seek an optimal match between simulations of the model and measurements taken of the system that we are trying to elucidate. This optimization takes two forms: classical and statistical. Let us explain. Classical optimization involves minimization of a positive, usually quadratic cost function that expresses the quantity that we seek to optimize. In most of the cases that we will deal with, this will be a function of the error between model and measurements—in this case we will speak of least-squares error minimization. The second form, statistical optimization, involves minimization of the variability or uncertainty of the model error and is based on statistical estimation theory.

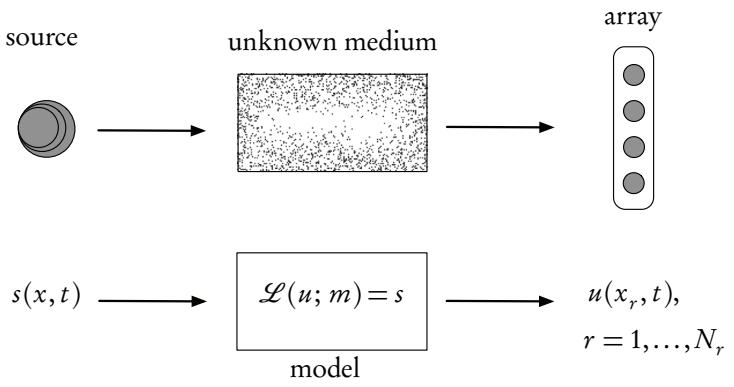


Figure 1.1. *Ingredients of an inverse problem: the physical reality (top) and the direct mathematical model (bottom). The inverse problem uses the difference between the model-predicted observations, u (calculated at the receiver array points, x_r), and the real observations measured on the array to find the unknown model parameters, m, or the source, s (or both).*

The main sources of inverse problems are science (social sciences included!) and engineering—in fact any process that we can model and measure satisfactorily. Often these problems concern the determination of the properties of some inaccessible region from observations on the boundary of the region, or at discrete instants over a given time interval. In other words, our information is *incomplete*. This incompleteness is the source of the major difficulties (and challenges…) that we will encounter in the solution of DA and inverse problems.

1.2 ▪ Uncertainty quantification and related concepts

Definition 1.1. *Uncertainty quantification (UQ) is the science of quantitative characterization and reduction of uncertainties in both computational and real-world applications. It tries to determine how likely certain outcomes are if some aspects of the system are not exactly known.*

The system-science paradigm, as expounded in Jordan [2015], exhibits the important place occupied by DA and inverse methods within the "deductive spiral"—see Figure 1.2. These methods furnish an essential link between the real world and the model of the system. They are intimately related to the concepts of *validation* and *verification*. Verification asks the question, "are we solving the equations correctly?"—this is an exercise in mathematics. Validation asks, "are we solving the correct equations?"—this is an exercise in physics. In geophysics, for example, the concept of validation is replaced by *evaluation*, since complete validation is not possible.

UQ is a basic component of model validation. In fact it is vital for characterizing our confidence in results coming out of modeling and simulation and provides a mathematically rigorous certification that is often needed in decision-making. In fact, it gives a precise notion of what constitutes a validated model by replacing the subjective concept of *confidence* by mathematically rigorous methods and measures.

There are two major categories of uncertainties: epistemic and aleatory. The first is considered to be reducible in that we can control it by improving our knowledge of

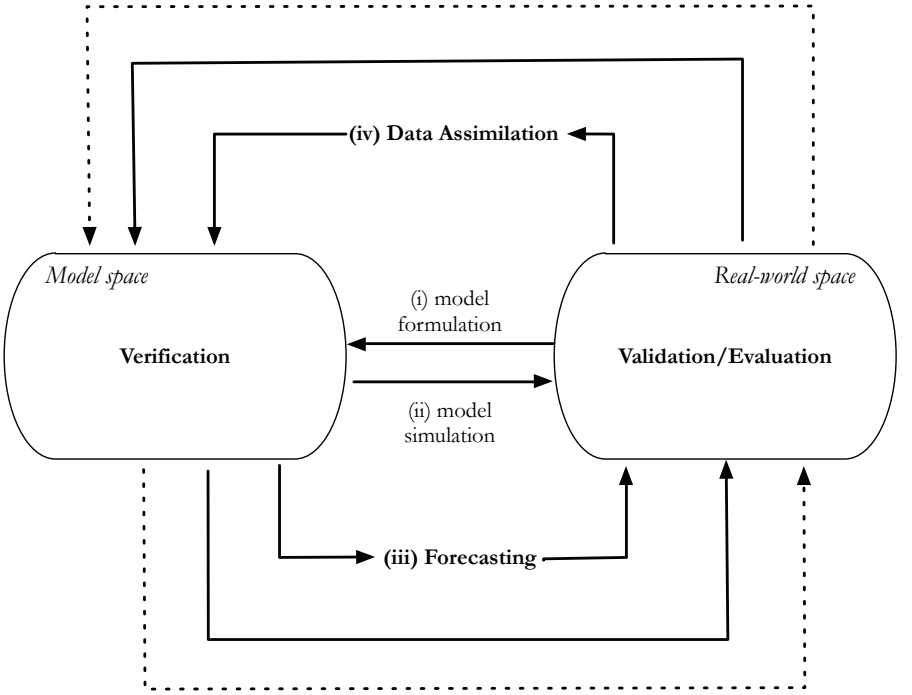


Figure 1.2. *The deductive spiral of system science (adapted from Jordan [*2015*]). The bottom half represents the direct problem (from model to reality); the top half represents the inverse problem (from reality to model). Starting from the center, with* (i) *model formulation and* (ii) *simulation, one works one's way around iteratively over* (iii) *forecasting and* (iv) *DA, while passing through the two phases of UQ (validation/evaluation and verification).*

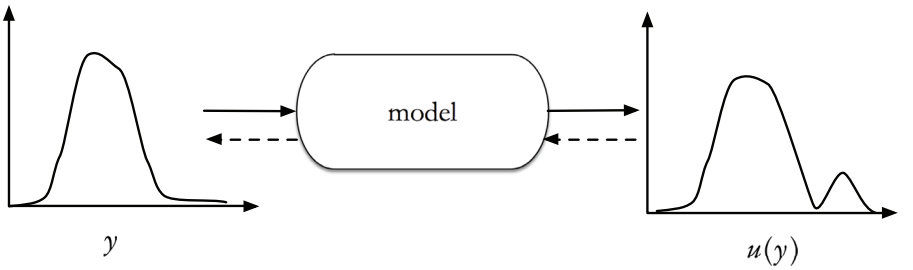


Figure 1.3. *UQ for a random quantity y: uncertainty propagation (left to right); uncertainty definition (right to left).*

the system. The second is assumed to be irreducible and has to do with the inherent noise in, or stochastic nature of, any natural system. Any computation performed under uncertainty will forcibly result in predictive simulations (see the introduction to Chapter 3 for more details on this point).

Uncertainty, in models of physical systems, is almost always represented as a probability density function (PDF) through samples, parameters, or kernels. The central

objective of UQ is then to represent, propagate, and estimate this density—see Figure 1.3.

As a process, UQ can be decomposed into the following steps:

1. Define the system of interest, its response, and the desired performance measures.
2. Write a mathematical formulation of the system—governing equations, geometry, parameter values.
3. Formulate a discretized representation and the numerical methods and algorithms for its solution.
4. Perform the simulations and the analysis.
5. Loop back to step 1.

The numerical simulations themselves can be decomposed into three steps:

1. DA, whose objective is to compute the PDFs of the input quantities of interest. This is the major concern of this book and the methods described herein.
2. Uncertainty propagation, whose objective is to compute the PDFs of the output quantities of interest. This is usually the most complex and computationally intensive step and is generally based on Monte Carlo and stochastic Galerkin (finite element) methods—see Le Maitre and Knio [2010].
3. Certification, whose objective is to estimate the likelihood of specific outcomes and compare them with risk or operating margins.

For a complete, recent mathematical overview of UQ, the reader is referred to Owhadi et al. [2013]. There are a number of research groups dedicated to the subject—please consult the websites of UQ groups at Stanford University,[3] MIT,[4] and ETH Zurich[5] (for example).

1.3 ▪ Basic concepts for inverse problems: Well- and ill-posedness

There is a fundamental, mathematical distinction between the direct and the inverse problem: direct problems are (invariably) well-posed, whereas inverse problems are (notoriously) ill-posed. Hadamard [1923] defined the concept of a well-posed problem as opposed to an ill-posed one.

Definition 1.2. *A mathematical model for a physical problem is* well-posed *if it possesses the following three properties:*

WP1 *Existence of a solution.*

WP2 *Uniqueness of the solution.*

WP3 *Continuous dependence of the solution on the data.*

[3] `http://web.stanford.edu/group/uq/`.

[4] `http://uqgroup.mit.edu`.

[5] `http://www.sudret.ibk.ethz.ch`.

Note that existence and uniqueness together are also known as "identifiability," and the continuous dependence is related to the "stability" of the inverse problem. A more rigorous mathematical formulation is the following (see Kirsch [1996]).

Definition 1.3. *Let X and Y be two normed spaces, and let $K : X \to Y$ be a linear or nonlinear map between the two. The problem of finding x given y such that*

$$Kx = y$$

is well-posed if the following three properties hold:

WP1 *Existence—for every $y \in Y$ there is (at least) one solution $x \in X$ such that $Kx = y$.*

WP2 *Uniqueness—for every $y \in Y$ there is at most one $x \in X$ such that $Kx = y$.*

WP3 *Stability—the solution, x, depends continuously on the data, y, in that for every sequence $\{x_n\} \subset X$ with $Kx_n \to Kx$ as $n \to \infty$, we have that $x_n \to x$ as $n \to \infty$.*

This concept of ill-posedness will be the "red thread" running through the entire book. It will help us to understand and distinguish between direct and inverse models. It will provide us with basic comprehension of the methods and algorithms that will be used to solve inverse problems. Finally, it will assist us in the analysis of what went wrong in our attempt to solve the inverse problems.

1.4 ▪ Examples of direct and inverse problems

Take a parameter-dependent dynamical system,

$$\frac{dz}{dt} = g(t, z; \theta), \qquad z(t_0) = z_0,$$

with g known, z_0 an initial state, $\theta \in \Theta$ (a space, or set, of possible parameter values), and the state $z(t) \in \mathbb{R}^n$. We can now define the two classes of problems.

Direct Given parameters θ and initial state z_0, find $z(t)$ for $t \geq t_0$.

Inverse Given observations $z(t)$ for $t \geq t_0$, find $\theta \in \Theta$.

Since the observations are incompletely known (over space-time), they must be modeled by an observation equation,

$$f(t, \theta) = \mathcal{H} z(t, \theta),$$

where $\mathcal{H}$ is the observation operator (which could, in ideal situations, be the identity). Usually we have a finite number, p, of discrete (space-time) observations

$$\left\{\tilde{y}_j\right\}_{j=1}^{p},$$

where

$$\tilde{y}_j \approx f(t_j, \theta)$$

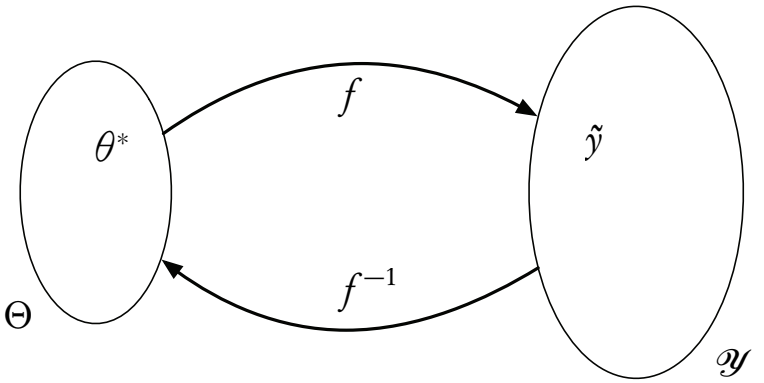
and the approximately equal sign denotes the possibility of measurement errors.

We now present a series of simple examples that clearly illustrate the three properties of well-posedness.

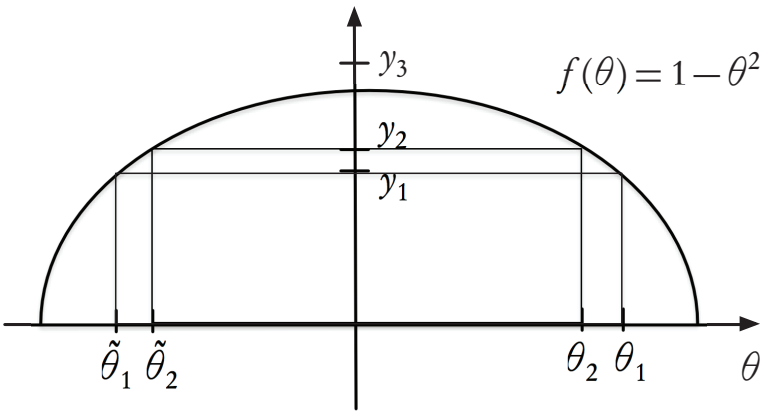
Example 1.4. (Simplest case—inspired by an oral presentation of H.T. Banks). Suppose that we have one observation, $\tilde{y}$, for $f(\theta)$ and we want to find the pre-image

$$\theta^* = f^{-1}(\tilde{y})$$

for a given $\tilde{y}$:



This problem can be severely ill-posed! Consider the following function:



Nonexistence There is no θ_3 such that $f(\theta_3) = y_3$.

Nonuniqueness $y_j = f(\theta_j) = f(\tilde{\theta}_j)$ for $j = 1, 2$.

Lack of continuity $|y_1 - y_2|$ small $\not\Rightarrow |f^{-1}(y_1) - f^{-1}(y_2)| = \left|\theta_1 - \tilde{\theta}_2\right|$ small.

Note that all three well-posedness properties, WP1, WP2, and WP3, are violated by this very basic case. Why is this so important? Couldn't we just apply a good least-squares algorithm (for example) to find the best possible solution? Let's try this. We define a mismatch-type cost function,

$$J(\theta) = |y_1 - f(\theta)|^2,$$

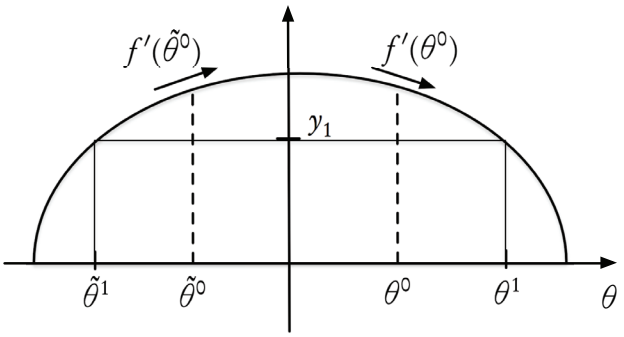
for a given y_1 and try to minimize this square error by applying a standard iterative scheme, such as direct search or gradient-based minimization [Quarteroni et al., 2007], to obtain a solution. For example, if we apply Newton's method, we obtain the following iteration:

$$\theta^{k+1} = \theta^k - \left[J'(\theta^k)\right]^{-1} J(\theta^k),$$

where

$$J'(\theta) = 2(y_1 - f(\theta))\big(-f'(\theta)\big).$$

Let us graphically perform a few iterations on the above function:



- $J'(\theta^0) = 2(\overbrace{y_1 - f(\theta^0)}^{(-)})(\overbrace{-f'(\theta^0)}^{(+)}) < 0 \quad \Rightarrow \quad \theta^1 > \theta^0$, etc.,
- $J'(\tilde{\theta}^0) = 2(\overbrace{y_1 - f(\tilde{\theta}^0)}^{(-)})(\overbrace{-f'(\tilde{\theta}^0)}^{(-)}) > 0 \quad \Rightarrow \quad \tilde{\theta}^1 < \tilde{\theta}^0$, etc.,

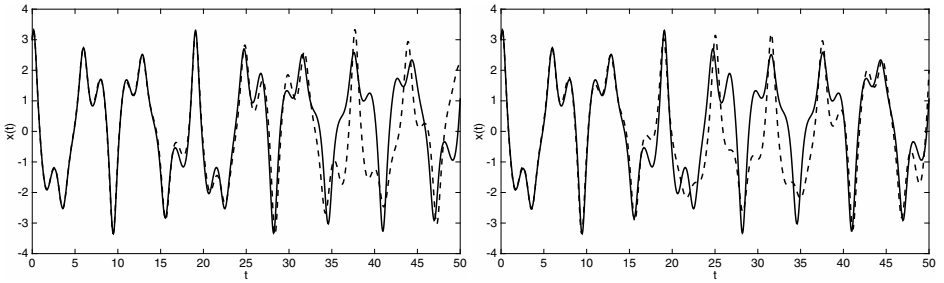
where in the last two formulas for J' we have indicated the sign $(+/-)$ above each of the terms. We observe that in this simple case we have a highly unstable, oscillating behavior: at each step we move from positive to negative increments due to the changing sign of the gradient, and convergence is not possible. So what went wrong here? This behavior is not the fault of descent algorithms. It is a manifestation of the *inherent ill-posedness* of the problem. How to fix this problem has been the subject of much research over the past 50 years! Many remedies (fortunately) exist, such as explicit and implicit constrained optimizations, regularization, and penalization—these will be referred to, when necessary, in what follows. ∎

To further appreciate the complexity, let us briefly consider one of the remedies: Tykhonov regularization (TR)—see, for example, Engl et al. [1996] and Vogel [2002]. The idea here is to replace the ill-posed problem for $J(\theta) = |y_1 - f(\theta)|^2$ by a "nearby" problem for

$$J_\beta(\theta) = |y_1 - f(\theta)|^2 + \beta\,|\theta - \theta_0|^2,$$

where β is a suitably chosen regularization/penalization parameter. When it is done correctly, TR provides convexity and compactness,[6] thus ensuring the existence of a

[6]Convexity means that the function resembles a quadratic function, $f(x) = x^2$, with positive second derivative; compactness means that any infinite sequence of functions must get arbitrarily close to some function of the space.



(a) Initial error of 0.03% (b) Initial error of 0.06%

Figure 1.4. *Duffing's equation with small initial perturbations. Unperturbed (solid line) and perturbed (dashed line) trajectories.*

unique solution. However, even when done correctly, it *modifies the problem*, and new solutions may be far from the original ones. In addition, it is not trivial to regularize correctly or even to know if we have succeeded in finding a solution.

Example 1.5. The highly nonlinear Duffing's equation [Guckenheimer and Holmes, 1983],

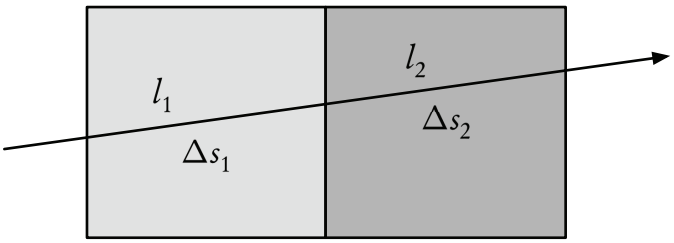
$$\ddot{x} + 0.05\dot{x} + x^3 = 7.5\cos t,$$

exhibits great *sensitivity to the initial conditions* (WP3). We will observe that two very closely spaced initial states can lead to a large discrepancy in the trajectories.

- Let $x(0) = 3$ and $\dot{x}(0) = 4$ be the true initial state.
- Introduce an error of 0.03% in the initial state—here we have an accurate forecast until $t = 35$ (see Figure 1.4(a)).
- Introduce an error of 0.06% in the initial state—here we only have an accurate forecast until $t = 20$ (see Figure 1.4(b)).

The initial perturbations are scarcely visible (and could result from measurement error), but the terminal states can differ considerably. ∎

Example 1.6. Seismic travel-time tomography provides an excellent example of non-uniqueness (WP2):



A signal seismic ray (or any other ray used in medical or other imaging) passes through a two-parameter block model.

- The *unknowns* are the two block slownesses (inverse of seismic velocity), $(\Delta s_1, \Delta s_2)$.
- The *data* consist of the observed travel time of the ray, Δt_1.
- The *model* is the linearized travel time equation

$$\Delta t_1 = l_1 \Delta s_1 + l_2 \Delta s_2,$$

where l_j is the length of the ray in the jth block.

Clearly we have one equation for two unknowns, and hence there is *no unique solution*. In fact, for a given value of Δt_1, each time we fix Δs_1 we obtain a different Δs_2 (and vice versa). ∎

We hope the reader is convinced, based on these disarmingly simple examples, that inverse problems present a large number of potential pathologies. We can now proceed to examine the therapeutic tools that are at our disposal for attempting to "heal the patient."

1.5 ▪ DA methods

Definition 1.7. *DA is the approximation of the true state of some physical system at a given time by combining time-distributed observations with a dynamic model in an optimal way.*

DA can be classically approached in two ways: as variational DA and as statistical[7] DA—see Figure 1.5. They will be briefly presented here and then in far more detail in Chapters 2 and 3, respectively. Newer approaches are also becoming available: nudging methods, reduced methods, ensemble methods, and hybrid methods that combine variational and statistical[8] approaches. These are the subject of the chapters on advanced methods—see Part II.

In both we seek an optimal solution—statistically we will, for example, seek a solution with minimum variance, whereas variationally we will seek a solution that minimizes a suitable cost (or error) function. In fact, in certain special cases the two approaches are identical and provide exactly the same solution. However, the statistical approach, though often more complex and time-consuming, can provide a richer information structure: an average solution and some characteristics of its variability (probability distribution). Clearly a temperature forecast of "15°C for tomorrow" is much less informative than a forecast of "an *average* of 15°C with a *standard deviation* of 1.5°C," or "the *probability* of a temperature below 10°C is 0.125 for tomorrow," or, as is now quite common (on our smartphones), "there is a 60% chance of rain at 09h00 in New York." Ideally (and this will be our recommendation), one should attempt to combine the two into a single, "hybrid" approach. We can then take advantage of the relative rapidity and robustness of the variational approach, and at the same time obtain an information-rich solution thanks to the statistical/probabilistic approach. This is easily said but (as we will see) is not trivial to implement and will be highly problem dependent and probably computationally expensive. However, we can and

[7] Alternatives are "filtering" or "probabilistic."
[8] *ibid.*

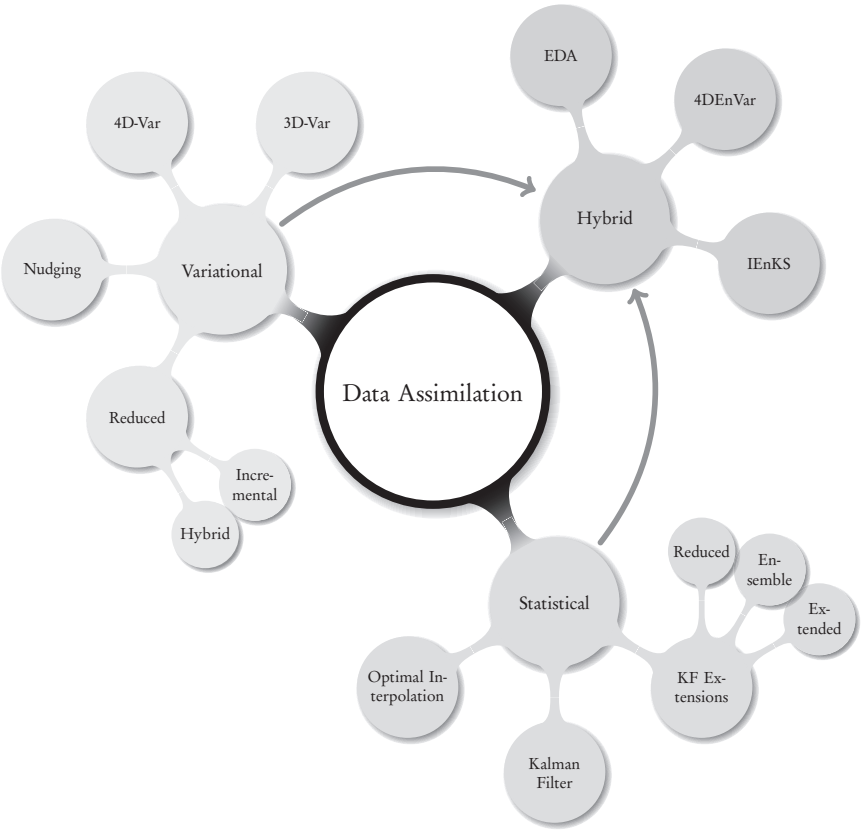


Figure 1.5. *DA methods: variational (Chapters* 2*,* 4*,* 5*); statistical (Chapters* 3*,* 5*,* 6*), hybrid (Chapter* 7*).*

will provide all the necessary tools (theoretical, algorithmic, numerical) and the indications for their implementation. It is worthwhile to point out that recently (since 2010) a number of the major weather-forecasting services across the world (Canada, France, United Kingdom, etc.) have started basing their operational forecasting systems on a new, hybrid approach, 4DEnVar (presented in Chapter 7), which combines 4D-Var (variational DA) with an ensemble, statistical approach.

To complete this introductory chapter, we will now briefly introduce and compare the two approaches of variational and statistical. Each one is subsequently treated, in far greater detail, in its own chapter—see Chapters 2 and 3, respectively.

1.5.1 ▪ Notation for DA and inverse problems

We begin by introducing the standard notation for DA problems as formalized by Ide et al. [1997]. We first consider a discrete model for the evolution of a physical (atmospheric, oceanic, mechanical, biological, etc.) system from time t_k to time t_{k+1}, described by a dynamic state equation

$$\mathbf{x}^{\mathrm{f}}(t_{k+1}) = \mathbf{M}_{k+1}\left[\mathbf{x}^{\mathrm{f}}(t_k)\right], \tag{1.1}$$

where $\mathbf{x}$ is the model's state vector of dimension n (see below for the definition of the superscripts) and $\mathbf{M}$ is the corresponding dynamics operator that can be time dependent. This operator usually results from a finite difference [Strikwerda, 2004] or finite element [Hughes, 1987] discretization of a (partial) differential equation (PDE). We associate an error covariance matrix $\mathbf{P}$ with the state $\mathbf{x}$ since the true state will differ from the simulated state (1.1) by random or systematic errors.

Observations, or measurements, at time t_k are defined by

$$\mathbf{y}^{\mathrm{o}} = \mathbf{H}_k\left[\mathbf{x}^t(t_k)\right] + \boldsymbol{\epsilon}_k^{\mathrm{o}}, \tag{1.2}$$

where $\mathbf{H}$ is an *observation operator* that can be time dependent and $\boldsymbol{\epsilon}^{\mathrm{o}}$ is a *white noise process* with zero mean and associated covariance matrix $\mathbf{R}$ that describes instrument errors and representation errors due to the discretization. The observation vector, $\mathbf{y}_k^{\mathrm{o}} = \mathbf{y}^{\mathrm{o}}(t_k)$, has dimension p_k, which is usually much smaller than the state dimension, $p_k \ll n$.

Subscripts are used to denote the discrete time index, the corresponding spatial indices, or the vector with respect to which an error covariance matrix is defined; superscripts refer to the nature of the vectors/matrices in the DA process:

- "a" for *analysis*,
- "b" for *background* (or initial/first guess),
- "f" for *forecast*,
- "o" for *observation*, and
- "t" for the (unknown) *true* state.

Analysis is the process of approximating the true state of a physical system at a given time. Analysis is based on

- observational data,
- a model of the physical system, and
- background information on initial and boundary conditions.

An analysis that combines time-distributed observations and a dynamic model is called *data assimilation*.

Now let us introduce the continuous system. In fact, continuous time simplifies both the notation and the theoretical analysis of the problem. For a finite-dimensional system of ODEs, the equations (1.1)–(1.2) become

$$\dot{\mathbf{x}}^{\mathrm{f}} = \mathscr{M}(\mathbf{x}^{\mathrm{f}}, t),$$

and

$$\mathbf{y}^{\mathrm{o}}(t) = \mathscr{H}(\mathbf{x}^{\mathrm{t}}, t) + \boldsymbol{\epsilon},$$

where $\dot{(\,)} = \mathrm{d}/\mathrm{d}t$ and $\mathscr{M}$ and $\mathscr{H}$ are nonlinear operators in continuous time for the model and the observation, respectively. This implies that $\mathbf{x}$, $\mathbf{y}$, and $\boldsymbol{\epsilon}$ are also continuous-in-time functions. For PDEs, where there is in addition a dependence on space, attention must be paid to the function spaces, especially when performing variational analysis. Details will be provided in the next chapter. With a PDE model, the field (state) variable is commonly denoted by $\boldsymbol{u}(\mathbf{x}, t)$, where $\mathbf{x}$ represents the space

variables (no longer the state variable as above!), and the model dynamics is now a nonlinear partial differential operator,

$$\mathcal{M} = \mathcal{M}\left[\partial_{\mathbf{x}}^{\alpha}, \boldsymbol{u}(\mathbf{x}, t), \mathbf{x}, t\right],$$

with $\partial_{\mathbf{x}}^{\alpha}$ denoting the partial derivatives with respect to the space variables of order up to $|\alpha| \leq m$, where m is usually equal to two and in general varies between one and four.

1.5.2 ▪ Statistical DA

Practical inverse problems and DA problems involve measured data. These data are inexact and are mixed with random noise. Only *statistical models* can provide rigorous, effective means for dealing with this measurement error. Let us begin with the following simple example.

1.5.2.1 ▪ A simple example

We want to estimate a *scalar* quantity, say the temperature or the ozone concentration at a fixed point in space. Suppose we have

- a model forecast, x^{b} (background, or a priori value), and
- a measured value, x^{o} (observation).

The simplest possible approach is to try a linear combination of the two,

$$x^{\mathrm{a}} = x^{\mathrm{b}} + w(x^{\mathrm{o}} - x^{\mathrm{b}}),$$

where x^{a} denotes the analysis that we seek and $0 \leq w \leq 1$ is a weight factor. We subtract the (always unknown) true state x^{t} from both sides,

$$x^{\mathrm{a}} - x^{\mathrm{t}} = x^{\mathrm{b}} - x^{\mathrm{t}} + w(x^{\mathrm{o}} - x^{\mathrm{t}} - x^{\mathrm{b}} + x^{\mathrm{t}}),$$

and, defining the three errors (analysis, background, observation) as

$$e^{\mathrm{a}} = x^{\mathrm{a}} - x^{\mathrm{t}}, \quad e^{\mathrm{b}} = x^{\mathrm{b}} - x^{\mathrm{t}}, \quad e^{\mathrm{o}} = x^{\mathrm{o}} - x^{\mathrm{t}},$$

we obtain

$$e^{\mathrm{a}} = e^{\mathrm{b}} + w(e^{\mathrm{o}} - e^{\mathrm{b}}) = we^{\mathrm{o}} + (1-w)e^{\mathrm{b}}.$$

If we have many realizations, we can take an ensemble average,[9] denoted by $\langle\cdot\rangle$:

$$\langle e^{\mathrm{a}}\rangle = \left\langle e^{\mathrm{b}}\right\rangle + w\left(\langle e^{\mathrm{o}}\rangle - \left\langle e^{\mathrm{b}}\right\rangle\right).$$

Now if these errors are centered (have zero mean, or the estimates of the true state are *unbiased*), then

$$\langle e^{\mathrm{a}}\rangle = 0$$

also. So we are logically led to look at the *variance* and demand that it be as small as possible. The variance is defined, using the above notation, as

$$\sigma^2 = \left\langle (e - \langle e\rangle)^2\right\rangle.$$

[9]Please refer to a good textbook on probability and statistics for all the relevant definitions—e.g., DeGroot and Schervisch [2012].

So by taking variances of the error equation, and using the zero-mean property, we obtain

$$\sigma_{\mathrm{a}}^2 = \sigma_{\mathrm{b}}^2 + w^2 \left\langle \left(e^{\mathrm{o}} - e^{\mathrm{b}}\right)^2 \right\rangle + 2w \left\langle e^{\mathrm{b}} \left(e^{\mathrm{o}} - e^{\mathrm{b}}\right) \right\rangle.$$

This reduces to

$$\sigma_{\mathrm{a}}^2 = \sigma_{\mathrm{b}}^2 + w^2 \left(\sigma_{\mathrm{o}}^2 + \sigma_{\mathrm{b}}^2\right) - 2w\sigma_{\mathrm{b}}^2$$

if e^{o} and e^{b} are uncorrelated. Now, to compute a minimum, take the derivative of this last equation with respect to w and equate to zero,

$$0 = 2w \left(\sigma_{\mathrm{o}}^2 + \sigma_{\mathrm{b}}^2\right) - 2\sigma_{\mathrm{b}}^2,$$

where we have ignored all cross terms since the errors have been assumed to be independent. Finally, solving this last equation, we can write the optimal weight,

$$w_* = \frac{\sigma_{\mathrm{b}}^2}{\sigma_{\mathrm{o}}^2 + \sigma_{\mathrm{b}}^2} = \frac{1}{1 + \sigma_{\mathrm{o}}^2/\sigma_{\mathrm{b}}^2},$$

which, we notice, depends on the ratio of the observation and the background errors. Clearly $0 \leq w_* \leq 1$ and

- if the observation is perfect, $\sigma_{\mathrm{o}}^2 = 0$ and thus $w_* = 1$, the maximum weight;
- if the background is perfect, $\sigma_{\mathrm{b}}^2 = 0$ and $w_* = 0$, so the observation will not be taken into account.

We can now rewrite the analysis error variance as

$$\begin{aligned} \sigma_{\mathrm{a}}^2 &= w_*^2 \sigma_{\mathrm{o}}^2 + (1 - w_*)^2 \sigma_{\mathrm{b}}^2 \\ &= \frac{\sigma_{\mathrm{b}}^2 \sigma_{\mathrm{o}}^2}{\sigma_{\mathrm{o}}^2 + \sigma_{\mathrm{b}}^2} \\ &= (1 - w_*)\sigma_{\mathrm{b}}^2 \\ &= \frac{1}{\sigma_{\mathrm{o}}^{-2} + \sigma_{\mathrm{b}}^{-2}}, \end{aligned}$$

where we suppose that $\sigma_{\mathrm{b}}^2,\ \sigma_{\mathrm{o}}^2 > 0$. In other words,

$$\frac{1}{\sigma_{\mathrm{a}}^2} = \frac{1}{\sigma_{\mathrm{o}}^2} + \frac{1}{\sigma_{\mathrm{b}}^2}.$$

Finally, the analysis equation becomes

$$x^{\mathrm{a}} = x^{\mathrm{b}} + \frac{1}{1 + \alpha}(x^{\mathrm{o}} - x^{\mathrm{b}}),$$

where $\alpha = \sigma_{\mathrm{o}}^2/\sigma_{\mathrm{b}}^2$. This is called the BLUE—best linear unbiased estimator—because it gives an unbiased, optimal weighting for a linear combination of two independent measurements.

We can isolate three special cases:

- If the observation is very accurate, $\sigma_{\mathrm{o}}^2 \ll \sigma_{\mathrm{b}}^2$, $\alpha \ll 1$, and thus $x^{\mathrm{a}} \approx x^{\mathrm{o}}$.

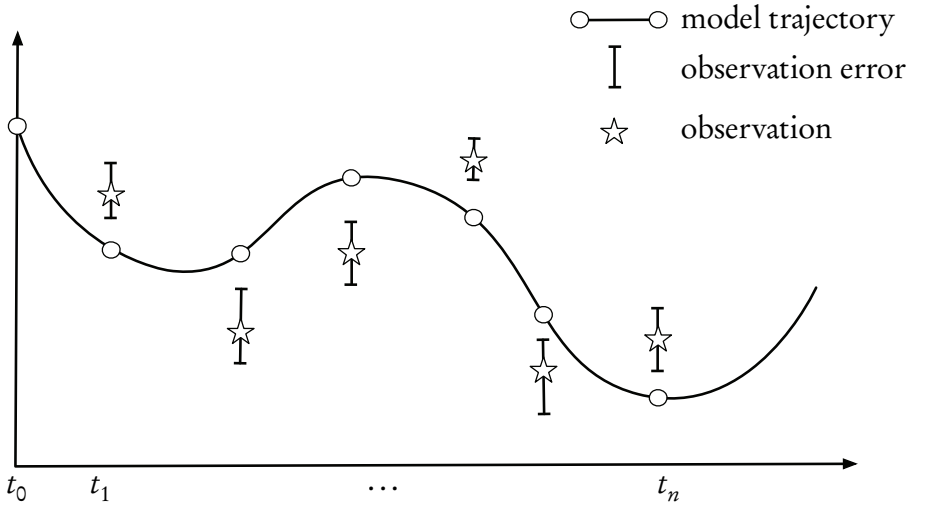


Figure 1.6. *Sequential assimilation. The x-axis denotes time; the y-axis is the assimilated variable.*

- If the background is accurate, $\alpha \gg 1$ and $x^{\mathrm{a}} \approx x^{\mathrm{b}}$.
- And, finally, if the observation and background variances are approximately equal, then $\alpha \approx 1$ and x^{a} is just the arithmetic average of x^{b} and x^{o}.

We can conclude that this simple, linear model does indeed capture the full range of possible solutions in a statistically rigorous manner, thus providing us with an "enriched" solution when compared with a nonprobabilistic, scalar response such as the arithmetic average of observation and background, which would correspond to only the last of the above three special cases.

1.5.2.2 ▪ The more general case: Introducing the Kalman filter

The above analysis of the temperature was based on a spatially dependent model. However, in general, the underlying process that we want to model will be time dependent. Within the significant toolbox of mathematical tools that can be used for statistical estimation from noisy sensor measurements, one of the most well-known and often-used tools is the Kalman filter (KF). The KF is named after Rudolph E. Kalman, who in 1960 published his famous paper describing a recursive solution to the *time-dependent* discrete-data linear filtering problem [Kalman, 1960].

We consider a dynamical system that evolves in time, and we seek to estimate a series of true states, $\mathbf{x}_k^{\mathrm{t}}$ (a sequence of random vectors), where discrete time is indexed by the letter k. These times are those when the observations or measurements are taken—see Figure 1.6. The assimilation starts with an unconstrained model trajectory from $t_0, t_1, \ldots, t_{k-1}, t_k, \ldots, t_n$ and aims to provide an optimal fit to the available observations/measurements given their uncertainties (error bars), depicted in the figure.

This situation is modeled by a stochastic system. We seek to estimate the state, $\mathbf{x} \in \mathbb{R}^n$, of a discrete-time dynamic process that is governed by the linear stochastic difference equation

$$\mathbf{x}_{k+1} = \mathbf{M}_{k+1}[\mathbf{x}_k] + \mathbf{w}_k,$$

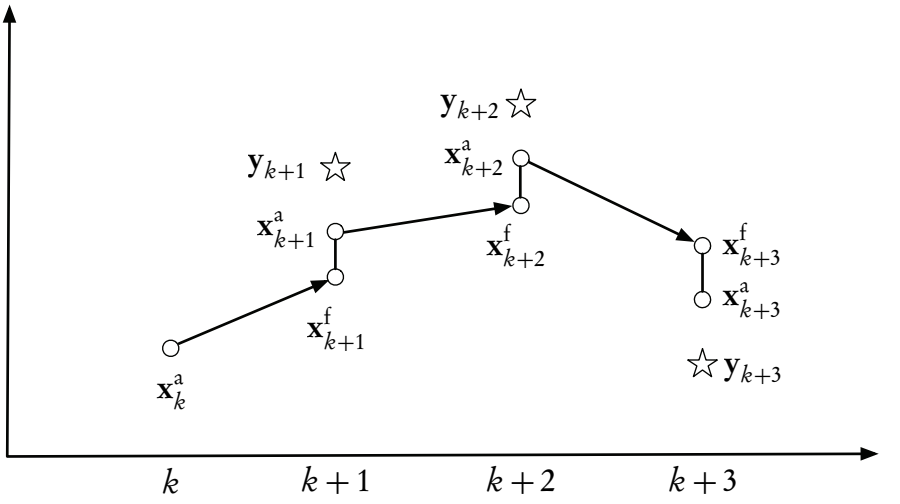


Figure 1.7. *Sequential assimilation scheme for the KF. The x-axis denotes time; the y-axis is the assimilated variable. We assume scalar variables.*

with a measurement/observation $\mathbf{y} \in \mathbb{R}^m$ defined by

$$\mathbf{y}_k = \mathbf{H}_k[\mathbf{x}_k] + \mathbf{v}_k.$$

The random vectors $\mathbf{w}_k$ and $\mathbf{v}_k$ represent the process/modeling and measurement/observation errors, respectively. They are assumed[10] to be independent, white, and with Gaussian/normal probability distributions,

$$\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k), \tag{1.3}$$
$$\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k), \tag{1.4}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are the covariance matrices (assumed known) and can in general be time dependent.

We can now set up a sequential DA scheme. The typical assimilation scheme is made up of two major steps: a *prediction/forecast* step and a *correction/analysis* step. At time t_k we have the result of a previous forecast, $\mathbf{x}^{\mathrm{f}}_k$ (the analogue of the background state $\mathbf{x}^{\mathrm{b}}_k$), and the result of an ensemble of observations in $\mathbf{y}_k$. Based on these two vectors, we perform an analysis that produces $\mathbf{x}^{\mathrm{a}}_k$. We then use the evolution model, which is usually (partial) differential equation–based, to obtain a prediction of the state at time t_{k+1}. The result of the forecast is denoted $\mathbf{x}^{\mathrm{f}}_{k+1}$ and becomes the background (or initial guess) for the next time step. This process is summarized in Figure 1.7.

We can now define forecast (a priori) and analysis (a posteriori) estimate errors in the same way as above for the scalar case, with their respective error covariance matrices, which generalize the variances used before, since we are now dealing with vector quantities. The goal of the KF is to compute an optimal a posteriori estimate, $\mathbf{x}^{\mathrm{a}}_k$, that is a linear combination of an a priori estimate, $\mathbf{x}^{\mathrm{f}}_k$, and a weighted difference between the actual measurement, $\mathbf{y}_k$, and the measurement prediction, $\mathbf{H}_k\left[\mathbf{x}^{\mathrm{f}}_k\right]$. This

[10] These assumptions are necessary in the KF framework. For real problems, they must often be relaxed.

is none other than the BLUE that we saw in the example above. The filter must be of the form

$$\mathbf{x}_k^{\mathrm{a}} = \mathbf{x}_k^{\mathrm{f}} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^{\mathrm{f}} \right), \tag{1.5}$$

where $\mathbf{K}$ is the Kalman gain. The difference $\left(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^{\mathrm{f}} \right)$ is called the *innovation* and reflects the discrepancy between the actual and the predicted measurements at time t_k. Note that for generality, the matrices are shown with a time dependence. Often this is not the case, and the subscripts k can then be dropped. The *Kalman gain* matrix, $\mathbf{K}$, is chosen to minimize the a posteriori error covariance equation. This is straightforward to compute: substitute (1.5) into the definition of the analysis error, then substitute in the error covariance equation, take the derivative of the trace of the result with respect to $\mathbf{K}$, set the result equal to zero, and, finally, solve for the optimal gain $\mathbf{K}$. The resulting optimal gain matrix is

$$\mathbf{K}_k = \mathbf{P}_k^{\mathrm{f}} \mathbf{H}^{\mathrm{T}} \left(\mathbf{H} \mathbf{P}_k^{\mathrm{f}} \mathbf{H}^{\mathrm{T}} + \mathbf{R} \right)^{-1},$$

where $\mathbf{P}_k^{\mathrm{f}}$ is the forecast error covariance matrix. Full details of this computation, as well as numerous examples, are provided in Chapter 3, where we will also generalize the approach to more realistic cases.

1.5.3 ▪ Variational DA

Unlike sequential/statistical assimilation (which emanates from estimation theory), variational assimilation is based on optimal control theory [Kwakernaak and Sivan, 1972; Friedland, 1986; Gelb, 1974; Tröltzsch, 2010], itself derived from the *calculus of variations*. The analyzed state is not defined as the one that maximizes a certain PDF, but as the one that *minimizes a cost function.* The minimization requires numerical optimization techniques. These techniques can rely on the *gradient* of the cost function, and this gradient will be obtained here with the aid of *adjoint methods*.

1.5.3.1 ▪ Adjoint methods: An introduction

All descent-based optimization methods require the computation of the gradient, ∇J, of a cost function, J. If the dependence of J on the control variables is complex or indirect, this computation can be very difficult. Numerically, we can always manage by computing finite increments, but this would have to be done in all possible perturbation directions. We thus need to find a less expensive way to compute the gradient. This will be provided by the *calculus of variations* and the *adjoint approach*.

A basic example: Let us consider a classical inverse problem known as a parameter identification problem, based on the ODE (of convection-diffusion type)

$$\begin{cases} -b\,u''(x) + c\,u'(x) = f(x), \quad 0 < x < 1, \\ u(0) = 0,\ u(1) = 0, \end{cases} \tag{1.6}$$

where $'$ depicts the derivative with respect to x, f is a given function, and b and c are *unknown (constant) parameters* that we seek to identify using observations of $u(x)$ on the interval $[0,1]$. The mismatch (or least-squares error) cost function is then

$$J(b,c) = \frac{1}{2} \int_0^1 (u(x) - u^{\circ}(x))^2 \, \mathrm{d}x,$$

where u^o is the observational data. The gradient of J can be calculated by introducing the *tangent linear model* (TLM). Perturbing the cost function by a small perturbation[11] in the direction α, with respect to the two parameters, b and c, gives

$$J(b+\alpha\delta b, c+\alpha\delta c) - J(b,c) = \frac{1}{2}\int_0^1 (\tilde{u}-u^o)^2 - (u-u^o)^2\,\mathrm{d}x,$$

where $\tilde{u} = u_{b+\alpha\delta b, c+\alpha\delta c}$ is the perturbed solution and $u = u_{b,c}$ is the unperturbed one. Now we divide by α and pass to the limit $\alpha \to 0$ to obtain the directional derivative (with respect to the parameters, in the direction of the perturbations),

$$\hat{J}[b,c](\delta b, \delta c) = \int_0^1 (u-u^o)\,\hat{u}\,\mathrm{d}x, \tag{1.7}$$

where we have defined

$$\hat{u} = \lim_{\alpha\to 0} \frac{\tilde{u}-u}{\alpha}.$$

Then, passing to the limit in equation (1.6), we can define the TLM

$$\begin{cases} -b\hat{u}'' + c\hat{u}' = (\delta b)u'' - (\delta c)u', \\ \hat{u}(0) = 0,\ \hat{u}(1) = 0. \end{cases} \tag{1.8}$$

We would like to reformulate the directional derivative (1.7) to obtain a calculable expression for the gradient. For this we introduce the adjoint variable, p, satisfying *the adjoint model*

$$\begin{cases} -bp'' - cp' = (u-u^o), \\ p(0) = 0,\ p(1) = 0. \end{cases} \tag{1.9}$$

Multiplying the TLM by this new variable, p, and integrating by parts enables us to finally write an explicit expression (see Chapter 2 for the complete derivation) for the gradient based on (1.7),

$$\nabla J(b,c) = \left(\int_0^1 pu''\,\mathrm{d}x\,, -\int_0^1 pu'\,\mathrm{d}x\right)^{\mathrm{T}},$$

or, separating the two components,

$$\nabla_b J(b,c) = \int_0^1 pu''\,\mathrm{d}x,$$

$$\nabla_c J(b,c) = -\int_0^1 pu'\,\mathrm{d}x.$$

Thus, for the additional cost of solving the adjoint model (1.9), we can compute the gradient of the cost function with respect to either one, or both, of the unknown parameters. It is now a relatively easy task to find (numerically) the optimal values of b and c that minimize J by using a suitable descent algorithm. This important example is fully developed in Section 2.3.2, where all the steps are explicitly justified.

Note that this method generalizes to (linear and nonlinear) time-dependent PDEs and to inverse problems where we seek to identify the *initial conditions*. This latter problem is exactly the 4D-Var problem of DA. All of this will be amply described in Chapter 2.

[11] The exact properties of the perturbation will be fully explained in Chapter 2.

Algorithm 1.1 Iterative 3D-Var (in its simplest form).

$k = 0$, $x = x_0$
while $||\nabla J|| > \epsilon$ or $j \leq j_{\max}$
  compute J
  compute ∇J
  gradient descent and update of x_{j+1}
  $j = j + 1$
end

1.5.3.2 ▪ 3D-Var

We have seen above that the BLUE requires the computation of an optimal gain matrix. We will show (in Chapters 2 and 3) that the optimal gain takes the form

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

to obtain an analyzed state,

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}(\mathbf{y} - \mathbf{H}(\mathbf{x}^{\mathrm{b}})),$$

that minimizes what is known as the 3D-Var cost function,

$$J(x) = \frac{1}{2}\left(\mathbf{x} - \mathbf{x}^{\mathrm{b}}\right)^{\mathrm{T}} \mathbf{B}^{-1}\left(\mathbf{x} - \mathbf{x}^{\mathrm{b}}\right) + \frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{y})^{\mathrm{T}} \mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y}), \tag{1.10}$$

where $\mathbf{R}$ and $\mathbf{B}$ (also denoted $\mathbf{P}^{\mathrm{f}}$) are the observation and background error covariance matrices, respectively. But the matrices involved in this calculation are often neither storable in memory nor manipulable because of their very large dimensions. The basic idea of variational methods is to overcome these difficulties by attempting to directly minimize the cost function, J. This minimization can be achieved, for inverse problems in general (and for DA in particular), by a combination of (1) an adjoint approach for the computation of the gradient of the cost function with (2) a descent algorithm in the direction of the gradient. For DA problems where there is no time dependence, the adjoint is not necessary and the approach is named 3D-Var, whereas for time-dependent problems we use the 4D-Var approach.

We recall that $\mathbf{R}$ and $\mathbf{B}$ are the observation and background error covariance matrices, respectively. When the observation operator $\mathbf{H}$ is linear, the gradient of J in (1.10) is given by

$$\nabla J = \mathbf{B}^{-1}\left(\mathbf{x} - \mathbf{x}^{\mathrm{b}}\right) - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}).$$

In the iterative 3D-Var Algorithm 1.1 we use as a stopping criterion the fact that ∇J is small or that the maximum number of iterations, $j_{\max}$, is reached.

1.5.3.3 ▪ A simple example of 3D-Var

We seek two temperatures, x_1 and x_2, in London and Paris. The climatologist gives us an initial guess (based on climate records) of $x^b = (10\ 5)^{\mathrm{T}}$, with background error covariance matrix

$$\mathbf{B} = \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}.$$

Algorithm 1.2 4D-Var in its basic form

$j = 0$, $\mathbf{x} = \mathbf{x}_0$
while $||\nabla J|| > \epsilon$ **or** $j \leq j_{\max}$
  (1) compute J with the direct model M and H
  (2) compute ∇J with adjoint model $\mathbf{M}^{\mathrm{T}}$ and $\mathbf{H}^{\mathrm{T}}$ (reverse mode)
  gradient descent and update of $\mathbf{x}_{j+1}$
  $j = j + 1$
end

We observe $\mathbf{y}^{\mathrm{o}} = 4$ in Paris, which implies that $\mathbf{H} = (0\ 1)$, with an observation error variance $\mathbf{R} = (0.25)$. We can now write the cost function (1.10) as

$$\begin{aligned} J(\mathbf{x}) &= \begin{pmatrix} x_1 - 10 & x_2 - 5 \end{pmatrix} \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - 10 \\ x_2 - 5 \end{pmatrix} + R^{-1}(x_2 - 4)^2 \\ &= \begin{pmatrix} x_1 - 10 & x_2 - 5 \end{pmatrix} \frac{16}{15} \begin{pmatrix} 1 & -0.25 \\ -0.25 & 1 \end{pmatrix} \begin{pmatrix} x_1 - 10 \\ x_2 - 5 \end{pmatrix} + 4(x_2 - 4)^2 \\ &= \frac{16}{15} \big((x_1 - 10)^2 + (x_2 - 5)^2 - 0.5(x_1 - 10)(x_2 - 5) \big) + 4(x_2 - 4)^2 \\ &= \frac{16}{15} \big(x_1^2 - 17.5 x_1 + 100 + x_2^2 - 5x_2 - 0.5 x_1 x_2 \big) + 4(x_2^2 - 8x + 16), \end{aligned}$$

and its gradient can be easily seen to be

$$\nabla J(\mathbf{x}) = \frac{16}{15} \begin{pmatrix} 2x_1 - 0.5x_2 - 17.5 \\ 2x_2 - 5 - 0.5x_1 + \frac{15}{4}(2x_2 - 8) \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 32x_1 - 8x_2 - 280 \\ -8x_1 + 152x_2 - 560 \end{pmatrix}.$$

The minimum is obtained for $\nabla J(\mathbf{x}) = 0$, which yields

$$x_1 = 9.8, \quad x_2 = 4.2.$$

This is an optimal estimate of the two temperatures, given the background and observation errors.

1.5.3.4 ▪ 4D-Var

In 4D-Var,[12] the cost function is still expressed in terms of the initial state, $\mathbf{x}_0$, but it will include the model because the observation $\mathbf{y}_i^{\mathrm{o}}$ at time i is compared to $\mathbf{H}_i(\mathbf{x}_i)$, where $\mathbf{x}_i$ is the state at time i initialized by $\mathbf{x}_0$ and the adjoint is not simply the transpose of a matrix but also the "transpose" of the model/operator dynamics. To compute this will require the use of a more general adjoint theory, which is introduced just after the following example and fully explained in Chapter 2.

In step (1) of Algorithm 1.2, we use the equations

$$\mathbf{d}_k = \mathbf{y}_k^{\mathrm{o}} - \mathbf{H}_k \mathbf{M}_k \mathbf{M}_{k-1} \ldots \mathbf{M}_2 \mathbf{M}_1 \mathbf{x}$$

and

$$J(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x} - \mathbf{x}^{\mathrm{b}} \right)^{\mathrm{T}} \mathbf{B}^{-1} \left(\mathbf{x} - \mathbf{x}^{\mathrm{b}} \right) + \sum_{i=0}^{j} \mathbf{d}_i^{\mathrm{T}} \mathbf{R}_i^{-1} \mathbf{d}_i.$$

[12] The 4 refers to the additional time dimension.

In step (2), we use

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^{\mathrm{b}}\right) - \left[\mathbf{H}_0^{\mathrm{T}}\mathbf{R}_0^{-1}\mathbf{d}_0 + \mathbf{M}_1^{\mathrm{T}}\left[\mathbf{H}_1^{\mathrm{T}}\mathbf{R}_1^{-1}\mathbf{d}_1 + \mathbf{M}_2^{\mathrm{T}}\left[\mathbf{H}_2^{\mathrm{T}}\mathbf{R}_2^{-1}\mathbf{d}_2 + \cdots + \mathbf{M}_j^{\mathrm{T}}\mathbf{H}_j^{\mathrm{T}}\mathbf{R}_j^{-1}\mathbf{d}_j\right]\right]\right],$$

where we have assumed that $\mathbf{H}$ and $\mathbf{M}$ are *linear*.

1.6 ▪ Some practical aspects of DA and inverse problems

In this brief section we point out some important practical considerations. It should now be clear that there are four basic ingredients in any inverse or DA problem:

1. observation or measured data;
2. a forward or direct model of the real-world context;
3. a backward or adjoint model in the variational case and a probabilistic framework in the statistical case; and
4. an optimization cycle.

But where does one start? The traditional approach, often employed in mathematical and numerical modeling, is to begin with some simplified, or at least well-known, situation. Once the above four items have been successfully implemented and tested on this instance, we then proceed to take into account more and more reality in the form of real data, more realistic models, more robust optimization procedures, etc. In other words, we introduce uncertainty, but into a system where we at least control some of the aspects.

1.6.1 ▪ Twin experiments

Twin experiments, or synthetic runs, are a basic and indispensable tool for all inverse problems. To evaluate the performance of a DA system we invariably begin with the following methodology:

1. Fix all parameters and unknowns and define a reference trajectory, obtained from a run of the direct model—call this the "truth."
2. Derive a set of (synthetic) measurements, or background data, from this "true" run.
3. Optionally, perturb these observations to generate a more realistic observed state.
4. Run the DA or inverse problem algorithm, starting from an initial guess (different from the "true" initial state used above), using the synthetic observations.
5. Evaluate the performance, modify the model/algorithm/observations, and cycle back to step 1.

Twin experiments thus provide a well-structured methodological framework. Within this framework we can perform different "stress tests" of our system. We can modify the observation network, increase or decrease (even switch off) the uncertainty, test the robustness of the optimization method, and even modify the model. In fact, these experiments can be performed on the full physical model or on some simpler (or reduced-order) model.

1.6.2 ▪ Toy models and other simplifications

Toy models are, by definition, simplified models that we can play with, yes, but these are of course "serious games." In certain complex physical contexts, of which meteorology is a famous example, we have well-established toy models, often of increasing complexity. These can be substituted for the real model, whose computational complexity is often too large, and provide a cheaper test-bed.

Some well-known examples of toy models are

- Lorenz models—see Lorenz [1963]—which are used as an avatar for weather simulations;
- various harmonic oscillators that are used to simulate dynamic systems; and
- famous examples such as the Ising model in physics, the Lotka–Volterra model in life sciences, and the Schelling model in social sciences;

See Marzuoli [2008] for a more general discussion.

1.7 ▪ To go further: Additional comments and references

- Examples of inverse problems: 11 examples can be found in Keller [1966] and 16 in Kirsch [1996].
- As the reader may have observed, the formulation and solution of DA and inverse problems require a wide range of tools and competencies in functional analysis, probability and statistics, variational calculus, numerical optimization, numerical approximation of (partial) differential equations, and stochastic simulation. This monograph will not provide all of this, so the reader must resort to other sources for the necessary background "tools." A few bibliographic recommendations are
  - DeGroot and Schervisch [2012] for probability and statistics;
  - Courant and Hilbert [1989a] for variational calculus;
  - Nocedal and Wright [2006] for numerical optimization;
  - Kreyszig [1978] and Reed and Simon [1980] for functional analysis;
  - Strikwerda [2004] for finite difference methods;
  - Hughes [1987] and Zienkiewicz and Taylor [2000] for finite element methods;
  - Press et al. [2007] for stochastic simulation;
  - Quarteroni et al. [2007] for basic numerical analysis (integration, solution of ODEs, etc.); and
  - Golub and van Loan [2013] for numerical linear algebra.