

**Forecast and control of dynamical systems
with data assimilation:**

Applications to COVID-19 epidemic and to Lorenz models

A dissertation submitted for the degree of
Doctor of Mathematical Science
September 2022

Qiwen SUN

Contents

Preface	iii
1 Overview: Background information of data assimilation	1
1.1 Dynamical systems	1
1.2 Probability measures generated by dynamical systems	2
1.3 Discrete-time data assimilation	3
1.3.1 The smoothing problem	4
1.3.2 The filtering problem	5
1.4 Kalman Filter	7
1.5 Ensemble Kalman filter (EnKF)	9
1.6 Several variants to EnKF	12
2 Analysis of COVID-19 in Japan with Extended SEIR model and ensemble Kalman filter	13
2.1 Introduction	13
2.2 Extended SEIR model	15
2.2.1 The SIR and SEIR models	15
2.2.2 The extended SEIR model	16
2.2.3 The reproduction number	18
2.2.4 The medical parameters	19
2.2.5 The observations	20
2.3 Ensemble Transform Kalman filter	21
2.4 Experiments	24
2.4.1 The experiments, using data from Tokyo	24
2.4.2 Technical discussion, using data from Tokyo	28
2.5 Discussion, comparisons, and future projects	31
3 Control simulation experiment with Lorenz’s butterfly attractor	39
3.1 Introduction	39
3.2 Experiments	41
3.3 Conclusions	44
3.A Appendix: The initial conditions of 40 CSEs	45
3.B Appendix: Additional sensitivity experiments	45

4	Control simulation experiments of extreme events with the Lorenz-96 model	49
4.1	Introduction	49
4.2	Background information	51
4.2.1	The Lorenz-96 model	51
4.2.2	Local Ensemble Transform Kalman Filter	52
4.2.3	Application of LETKF to the Lorenz-96 model	53
4.3	The Control Simulation Experiment	54
4.3.1	Extreme events	54
4.3.2	The Control Simulation Experiments	56
4.4	Local perturbations and partial observations	60
4.4.1	Local perturbations	60
4.4.2	Partial observations	63
4.5	Conclusion	65
	Epilogue	67
A	Bibliometric analysis on mathematics: 3 snapshots: 2005, 2010, 2015	71
A.1	Introduction	71
A.2	General pictures	73
A.3	The predictors and the response	76
A.4	Individual predictors	78
A.4.1	Number of authors	78
A.4.2	Number of countries	79
A.4.3	Number of institutes	80
A.4.4	Number of references	80
A.4.5	Number of pages	81
A.4.6	Number of keywords	82
A.4.7	Open access	83
A.4.8	Journal impact factor	84
A.4.9	Research areas	84
A.4.10	Categories	86
A.5	Investigations with decision trees	86
A.5.1	Tree classifier	87
A.5.2	The experiment	89
A.6	About countries	92
A.7	Conclusion	97
	Bibliography	99

Preface

Data assimilation is a widely used technique in numerical weather prediction and in other fields, as for example in earth science, in economics, and more recently in epidemiology. It is a mathematical tool that optimally combines theories and observations to find the precise signal and unknown parameters of a physical system, and to forecast its evolution. Ultimately, the question arises whether this technique can be used for controlling the evolution of the system in a prescribed direction?

In this thesis, we firstly apply data assimilation techniques to an infectious disease model to study the effectiveness of parameter estimations and system predictions. Secondly, we design the control simulation experiments (CSEs) of two chaotic dynamical systems and investigate the most effective perturbation signals.

The thesis is arranged as follows: In Chapter 1, the general framework of data assimilation is introduced. The theoretical developments of Kalman filter and ensemble Kalman filter (EnKF) are reviewed. Some special versions of EnKF are also discussed, as for example the ensemble transform Kalman filter (ETKF) and the ETKF with localization.

In Chapter 2, we introduce an extended SEIR infectious disease model together with a data assimilation scheme for the study of the spread of COVID-19. In this framework, undetected asymptomatic and pre-symptomatic cases are considered, and the impact of their uncertain proportion is fully investigated. The standard SEIR model does not consider these populations, while their role in the propagation of the disease is acknowledged. An ensemble Kalman filter is implemented to assimilate reliable observations of three compartments in the model. The system tracks the evolution of the effective reproduction number and estimates the unobservable subpopulations. The analysis is carried out for three main prefectures of Japan and for the entire population of Japan. We also perform sensitivity tests for different values of some uncertain medical parameters, like the relative infectivity of symptomatic / asymptomatic cases. The regional analysis results suggest the decreasing efficiency of the states of emergency.

In Chapter 3, we study the control of chaotic dynamical systems with data assimilation techniques. In numerical weather prediction (NWP), sensitivity to initial conditions often leads to an intrinsic limit to predictability, but it also implies an effective control in which a small control signal grows rapidly to make a substantial difference. In this chapter, we extend the well-known Observing Systems Simulation Experiment (OSSE) and design the CSE, in which the application of a small signal drags the systems in a prescribed direction. An idealized experiment with the Lorenz-63 three-variable system shows that we can control

the system to stay in a chosen wing of the Lorenz's butterfly attractor. Using longer lead time forecasts, we achieve more effective controls with a perturbation size of about 3% of the observation error. The idealized CSE is a starting point for CSEs applied to more realistic dynamical systems. A long-term aim would be for example to reduce weather disaster risks by adding small perturbations to the weather system.

The CSEs are further developed and applied to a more complicated scenario in Chapter 4. This CSE is aimed for reducing the number of extreme events in the Lorenz-96 model. The 40 variables of this model represent idealized meteorological quantities evenly distributed on a latitude circle. The reduction of occurrence of extreme events over a 100 years run of the model is discussed as a function of the parameters of the CSE: the ensemble forecast length for detecting extreme events in advance, the magnitude and localization of the perturbations, the quality and the coverage of the observations.

In addition to these studies, we also provide in an Appendix the work that drove us to the data assimilation framework, namely the bibliometric analysis of mathematical publications using tree-based methods. The framework for the bibliometric investigations (sketched during the master program) is added since the large-scale research has been performed at the beginning of the Ph.D. program. This Appendix contains several statistical results, and a detailed study of the factors that affect the citations of mathematical articles by using a tree-based classifier.

Let us finally mention that Chapter 3 and the Appendix correspond to published papers, Chapter 2 has been submitted for publication already in October 2021, while Chapter 4 is going to be submitted for publication in July 2022.

Acknowledgement

I wish to show my deepest appreciation to my supervisors, Prof. Serge Richard and Prof. Takemasa Miyoshi, for the constant support and guidance for each of my Ph.D. research projects.

I would also like to thank Dr. Hideyuki Sakamoto for his technical support during my stay at the Data assimilation research team at RIKEN Center for Computational Science. I wish to extend my thanks to all members of the team for their insightful comments and discussions, and to the team assistants, Yukie Komori, Saeko Imano, Sakiko Tashiro and Aki Mukunoki, for their support that ensures the smooth progress of my work.

I wish to thank students from Nagoya University who were involved in my research projects: Naohiro Tsuzu and Chang Sun for their contribution to the COVID-19 project.

Furthermore, I thank Prof. Toru Ohira, Prof. Yoshifumi Kimura and Prof. Takemasa Miyoshi for having accepted to be in the jury for my Ph.D. thesis. I also thank Masami Sato and Kichiyoshi Nagata for their help for the defense procedure.

I wish to thank my mother Kefen who encouraged me to pursue my Ph.D. and my father Fuyou who has always been kind and humorous. I am also grateful to my sister Qiqian and to my husband Yupeng who have supported me along the way. Finally, the research projects in this thesis were supported by the RIKEN Junior Research Associate Program and by FY2021 RIKEN President's Discretionary Funds for COVID-19.

Chapter 1

Overview: Background information of data assimilation

In this section, we introduce the necessary background information about data assimilation, and sketch the developments of the theory of the ensemble Kalman filter.

1.1 Dynamical systems

Let Ψ be a function from \mathbb{R}^ℓ to \mathbb{R}^ℓ and let x_0 be a point in \mathbb{R}^ℓ . The function Ψ defines an autonomous discrete-time dynamical system with initial condition x_0 by considering the sequence of points $\{x_j\}_{j \in \mathbb{N}}$ given by

$$x_{j+1} = \Psi(x_j). \quad (1.1)$$

For the continuous time scenario, let g be a C^1 function from \mathbb{R}^ℓ to \mathbb{R}^ℓ , and consider x a function from \mathbb{R} to \mathbb{R}^ℓ . Suppose x satisfies the following differential equation

$$\frac{dx(t)}{dt} = g(x(t)),$$

with initial condition $x(0) = x_0$. We assume that for any $x_0 \in \mathbb{R}^\ell$ and $t \in \mathbb{R}_+$, there exists such a solution. Then these assumptions guarantee the existence of the operator $\Psi(\cdot; t)$ with parameter t that satisfies

$$\begin{aligned} x(t) &= \Psi(x_0; t), & t \in \mathbb{R}_+, \\ \Psi(x_0; t+s) &= \Psi(\Psi(x_0; s), t), & s, t \in \mathbb{R}_+, \\ \Psi(x_0; 0) &= x_0. \end{aligned} \quad (1.2)$$

For $t > 0$, equation (1.2) has the same role as equation (1.1).

1.2 Probability measures generated by dynamical systems

For a given measurable space (Ω, \mathcal{F}) , a probability measure \mathbb{P} is a nonnegative countably additive set function from \mathcal{F} to $[0, 1]$ with $\mathbb{P}(\Omega) = 1$, namely it satisfies:

- (i) $\mathbb{P}(A) \geq \mathbb{P}(\emptyset) = 0$ for any $A \in \mathcal{F}$,
- (ii) for $\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{F}$ with $A_i \cap A_j = \emptyset$ whenever $i \neq j$, then $\mathbb{P}(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i)$.

Together with the measurable space, the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space [20, pp. 1].

Let $\mathcal{B}(\mathbb{R}^\ell)$ denote the σ -algebra of Borel sets on \mathbb{R}^ℓ generated by open sets [20, pp. 3]. We now consider a function $\kappa(\cdot, \cdot)$ from $\mathbb{R}^\ell \times \mathcal{B}(\mathbb{R}^\ell)$ to $[0, 1]$. Following [47, pp. 19], the function κ is a Markov kernel if it satisfies the following conditions:

- (i) For any fixed $x \in \mathbb{R}^\ell$, the map $A \rightarrow \kappa(x, A)$ is a probability measure on $(\mathbb{R}^\ell, \mathcal{B}(\mathbb{R}^\ell))$,
- (ii) For any fixed $A \in \mathcal{B}(\mathbb{R}^\ell)$, the map $x \mapsto \kappa(x, A)$ is $\mathcal{B}(\mathbb{R}^\ell)$ -measurable.

Let us come back to the discrete time dynamical system of the previous section, and consider a sequence of i.i.d. random variables $\{\xi_j\}_{j \in \mathbb{N}}$ on \mathbb{R}^ℓ with common probability density function (pdf) f_ξ . In the sequel, we shall only consider absolutely continuous random variables, and call them simply random variables. A stochastic dynamical system is formed by

$$X_{j+1} = \Psi(X_j) + \xi_j, \quad (1.3)$$

where X_0 is assumed to be a random variable on \mathbb{R}^ℓ , independent of ξ_j , $j \in \mathbb{N}$ [47, pp. 19]. Let us observe that $\{X_j\}_{j \in \mathbb{N}}$ defines a Markov chain, namely a sequence of random variables $\{X_j\}_{j \in \mathbb{N}}$ satisfying for any $j \in \mathbb{N}$, $j \neq 0$, and any $A_j \in \mathcal{B}(\mathbb{R}^\ell)$

$$\mathbb{P}(X_j \in A_j \mid X_{j-1} \in A_{j-1}, \dots, X_0 \in A_0) = \mathbb{P}(X_j \in A_j \mid X_{j-1} \in A_{j-1}). \quad (1.4)$$

In our setting, this condition is about the expression

$$\mathbb{P}(X_j \in A_j \mid X_{j-1} \in A_{j-1}, \dots, X_0 \in A_0)$$

for all $j \in \mathbb{N}$, $j \neq 0$, and $A_j \in \mathcal{B}(\mathbb{R}^\ell)$. Then, it is enough to replace X_j by $\Psi(X_{j-1}) + \xi_{j-1}$ and to observe that

$$\begin{aligned} & \mathbb{P}(\Psi(X_{j-1}) + \xi_{j-1} \in A_j \mid X_{j-1} \in A_{j-1}, \dots, X_0 \in A_0) \\ &= \mathbb{P}(\Psi(X_{j-1}) + \xi_{j-1} \in A_j \mid X_{j-1} \in A_{j-1}) \end{aligned}$$

because of the independence of ξ_{j-1} with X_0, \dots, X_{j-1} . Thus, one gets (1.4), as expected.

Starting again from (1.3), one observes that if X_j takes the value x_j , one has $X_{j+1} = \Psi(x_j) + \xi_j$, from which one directly infers that

$$f_{X_{j+1}|X_j}(x_{j+1} | x_j) = f_\xi(x_{j+1} - \Psi(x_j)), \quad (1.5)$$

where $f_{X_{j+1}|X_j}$ is the conditional pdf of X_{j+1} knowing X_j , and x_{j+1} is an arbitrary point in \mathbb{R}^ℓ . For any $A \in \mathcal{B}(\mathbb{R}^\ell)$, it then follows that

$$\mathbb{P}(X_{j+1} \in A | x_j) = \int_A f_\xi(x_{j+1} - \Psi(x_j)) dx_{j+1}.$$

Note that this expression defines a Markov kernel: With the notation introduced at the beginning of this section, the function κ defined by $\kappa(x_j, A) := \int_A f_\xi(x_{j+1} - \Psi(x_j)) dx_{j+1}$ is a Markov kernel.

Let us denote by f_{X_j} the pdf of the random variable X_j , for any $j \in \mathbb{N}$. It then follows from the previous observation that

$$\begin{aligned} \mathbb{P}(X_{j+1} \in A) &= \int_A f_{X_{j+1}}(x_{j+1}) dx_{j+1} \\ &= \int_A \int_{\mathbb{R}^\ell} f_{X_{j+1}|X_j}(x_{j+1} | x_j) f_{X_j}(x_j) dx_j dx_{j+1} \\ &= \int_A \int_{\mathbb{R}^\ell} f_\xi(x_{j+1} - \Psi(x_j)) f_{X_j}(x_j) dx_j dx_{j+1} \end{aligned}$$

At the level of densities, this equality reads:

$$f_{X_{j+1}}(x_{j+1}) = \int_{\mathbb{R}^\ell} f_\xi(x_{j+1} - \Psi(x_j)) f_{X_j}(x_j) dx_j.$$

for any $x_{j+1} \in \mathbb{R}^\ell$. Thus, following [47, pp. 20] we define the integral operator P by the formula

$$(Pf)(x) = \int_{\mathbb{R}^\ell} f_\xi(x - \Psi(x')) f(x') dx',$$

and the relation between $f_{X_{j+1}}$ and f_{X_j} is then provided by

$$f_{X_{j+1}} = Pf_{X_j}. \quad (1.6)$$

The pdf of the random variables $\{X_j\}_{j \in \mathbb{N}}$ are then obtained by iteration of (1.6).

1.3 Discrete-time data assimilation

We consider again the stochastic dynamical system (1.3). The sequence of random variables $\{X_j\}_{j \in \mathbb{N}}$ generated by (1.3) corresponds to the evolution of the model. Observations are usually obtained through the following process:

$$Y_j = h(X_j) + \eta_j, \quad (1.7)$$

where $h : \mathbb{R}^\ell \rightarrow \mathbb{R}^m$ is a function (also called observation operator) and $\{\eta_j\}_{j \in \mathbb{N}}$ is a sequence of i.i.d. random variables independent of X_0 and of $\{\xi_j\}_{j \in \mathbb{N}}$. We denote by η the random variable common to all η_j , and stress that the random variable η takes values in \mathbb{R}^m while the random variable ξ takes values in \mathbb{R}^ℓ . The random variables $\{Y_j\}_{j \in \mathbb{N}}$ are called the observations. The aim of data assimilation is to determine the unobservable signal $\{X_j\}_{j \in \mathbb{N}}$ by using the observations.

1.3.1 The smoothing problem

For $a, b \in \mathbb{N}$ with $a < b$ we introduce the notation $a : b$ for the set $\{j \in \mathbb{N} \mid a \leq j \leq b\}$. For example $y_{a:b}$ means $\{y_j\}_{a \leq j \leq b}$.

Let us fix $J \in \mathbb{N}$ with $J > 0$. In this section we shall consider the conditional pdf of $X_{0:J}$ given $y_{1:J}$. For that purpose, let $f_{X_{0:J}|y_{1:J}}$ represent this conditional pdf. This distribution is also called the smoothing distribution. To solve the smoothing problem, namely to find the smoothing distribution, we refer to Bayes' Theorem. Let $f_{X_{0:J}}$ be the pdf of $X_{0:J}$, also called prior pdf. Then, by the disintegration formula and the Markov property of $\{X_j\}_{j \in \mathbb{N}}$, one infers that [47, Sec. 2.3.2]

$$\begin{aligned} f_{X_{0:J}}(x_{0:J}) &= f_{X_J|X_{0:J-1}}(x_J \mid x_{0:J-1})f_{X_{0:J-1}}(x_{0:J-1}) \\ &= f_{X_J|X_{J-1}}(x_J \mid x_{J-1})f_{X_{0:J-1}}(x_{0:J-1}). \end{aligned}$$

The last factor $f_{X_{0:J-1}}(x_{0:J-1})$ can be further expanded in a similar manner, and one gets the equality [47, Sec. 2.3.2]

$$f_{X_{0:J}}(x_{0:J}) = f_{X_0}(x_0) \prod_{j=1}^J f_{X_j|X_{j-1}}(x_j \mid x_{j-1}).$$

Now, let us determine the likelihood function, namely the conditional pdf $f_{Y_{1:J}|X_{0:J}}(y_{1:J} \mid x_{0:J})$ of the observations $Y_{1:J}$ given $X_{0:J}$ [47, pp. 35]. By definition one has

$$f_{Y_{1:J}|X_{0:J}}(y_{1:J} \mid x_{0:J}) = \frac{f_{Y_{1:J}, X_{0:J}}(y_{1:J}, x_{0:J})}{f_{X_{0:J}}(x_{0:J})},$$

and by (1.7) one gets

$$\begin{aligned} f_{Y_{1:J}, X_{0:J}}(y_{1:J}, x_{0:J}) &= f_{Y_J|Y_{1:J-1}, X_{0:J}}(y_J \mid y_{1:J-1}, x_{0:J})f_{Y_{1:J-1}, X_{0:J}}(y_{1:J-1}, x_{0:J}) \\ &= f_{Y_J|X_J}(y_J \mid x_J)f_{Y_{1:J-1}, X_{0:J}}(y_{1:J-1}, x_{0:J}). \end{aligned}$$

By repeating this argument for $f_{Y_{1:J-1}, X_{0:J}}(y_{1:J-1}, x_{0:J})$, one finally obtains

$$f_{Y_{1:J}, X_{0:J}}(y_{1:J}, x_{0:J}) = f_{X_{0:J}}(x_{0:J}) \prod_{j=1}^J f_{Y_j|X_j}(y_j \mid x_j),$$

and thus

$$f_{Y_{1:J}|X_{0:J}}(y_{1:J} | x_{0:J}) = \prod_{j=1}^J f_{Y_j|X_j}(y_j | x_j).$$

By considering the relation between the η_j and y_j provided in (1.7), and by using the argument already explained for (1.5), one finds the likelihood of the observation Y_j given X_j has the same pdf as η_j , namely $f_{Y_j|X_j}(y_j | x_j) = f_\eta(y_j - h(x_j))$, where f_η denotes the pdf of the random variable η . Collecting these results, one obtains

$$f_{Y_{1:J}|X_{0:J}}(y_{1:J} | x_{0:J}) = \prod_{j=1}^J f_\eta(y_j - h(x_j)).$$

These different equalities together with Bayes' theorem lead finally to the posterior distribution $f_{X_{0:J}|Y_{1:J}}$, the pdf of $X_{0:J}$ after the observations $Y_{1:J}$ [47, pp. 35, Theorem 2.8]:

$$\begin{aligned} f_{X_{0:J}|Y_{1:J}}(x_{0:J} | y_{1:J}) &= \frac{f_{Y_{1:J}|X_{0:J}}(y_{1:J} | x_{0:J})f_{X_{0:J}}(x_{0:J})}{f_{Y_{1:J}}(y_{1:J})} \\ &= \frac{f_{X_{0:J}}(x_{0:J}) \prod_{j=1}^J f_\eta(y_j - h(x_j))}{f_{Y_{1:J}}(y_{1:J})}. \end{aligned}$$

Let us emphasize that so far, we have not used any specific form for the probability density functions of X_0 , of ξ , and of η . As shown in this section, the smoothing problem does not perform the analysis sequentially when new observations are available. In the next section, we introduce a sequential updating strategy described by the filtering problem. This approach is more interesting for solving problems in real time.

1.3.2 The filtering problem

For a time point j in \mathbb{N} with $j > 0$, the filtering distribution is defined as the conditional distribution of X_j given $Y_{1:j}$. Its pdf is denoted by $f_{X_j|Y_{1:j}}$. This function corresponds to the pdf of the unknown signal at time j given all observations from time 1 up to time j . The form of the filtering distribution inspires the potential of an update once a new observation is available. The online update can be carried out in two steps: the forecast step and the analysis step, see [47, pp. 38].

For some $j \in \mathbb{N}$, suppose that the pdf $f_{X_j|Y_{1:j}}$ is known. The forecast step consists in using this information for providing the forecast pdf $f_{X_{j+1}|Y_{1:j}}$ at time $j + 1$. The following

computation shows how it can be derived:

$$\begin{aligned}
f_{X_{j+1}|Y_{1:j}}(x_{j+1} | y_{1:j}) &= \int_{\mathbb{R}^\ell} f_{X_j, X_{j+1}|Y_{1:j}}(x_j, x_{j+1} | y_{1:j}) dx_j \\
&= \int_{\mathbb{R}^\ell} \frac{f_{X_j, X_{j+1}, Y_{1:j}}(x_j, x_{j+1}, y_{1:j})}{f_{Y_{1:j}}(y_{1:j})} dx_j \\
&= \int_{\mathbb{R}^\ell} \frac{f_{X_j, X_{j+1}, Y_{1:j}}(x_j, x_{j+1}, y_{1:j})}{f_{Y_{1:j}}(y_{1:j})} \frac{f_{X_j, Y_{1:j}}(x_j, y_{1:j})}{f_{X_j, Y_{1:j}}(x_j, y_{1:j})} dx_j \\
&= \int_{\mathbb{R}^\ell} \frac{f_{X_j, X_{j+1}, Y_{1:j}}(x_j, x_{j+1}, y_{1:j})}{f_{X_j, Y_{1:j}}(x_j, y_{1:j})} \frac{f_{X_j, Y_{1:j}}(x_j, y_{1:j})}{f_{Y_{1:j}}(y_{1:j})} dx_j \\
&= \int_{\mathbb{R}^\ell} f_{X_{j+1}|X_j, Y_{1:j}}(x_{j+1} | x_j, y_{1:j}) f_{X_j|Y_{1:j}}(x_j | y_{1:j}) dx_j \\
&= \int_{\mathbb{R}^\ell} f_{X_{j+1}|X_j}(x_{j+1} | x_j) f_{X_j|Y_{1:j}}(x_j | y_{1:j}) dx_j. \tag{1.8}
\end{aligned}$$

For the last step, we have used relation (1.3) which stipulates that X_j contains all necessary information for generating X_{j+1} , and that $Y_{1:j}$ are not necessary. In (1.8), $f_{X_{j+1}|X_j}$ can be computed by using (1.5) while $f_{X_j|Y_{1:j}}$ is supposed to be known.

Once the new observation y_{j+1} is available, one must updated the forecast pdf $f_{X_{j+1}|Y_{1:j}}$ to take this new observation into account, and get the analysis pdf $f_{X_{j+1}|Y_{1:j+1}}$. The computation takes the form:

$$\begin{aligned}
f_{X_{j+1}|Y_{1:j+1}}(x_{j+1} | y_{1:j+1}) &= f_{X_{j+1}|Y_{1:j}, Y_{j+1}}(x_{j+1} | y_{1:j}, y_{j+1}) \\
&= \frac{f_{X_{j+1}, Y_{1:j}, Y_{j+1}}(x_{j+1}, y_{1:j}, y_{j+1})}{f_{Y_{1:j}, Y_{j+1}}(y_{1:j}, y_{j+1})} \\
&= \frac{f_{Y_{j+1}|X_{j+1}, Y_{1:j}}(y_{j+1} | x_{j+1}, y_{1:j}) f_{X_{j+1}, Y_{1:j}}(x_{j+1}, y_{1:j})}{f_{Y_{1:j}, Y_{j+1}}(y_{1:j}, y_{j+1})} \\
&= \frac{f_{Y_{j+1}|X_{j+1}, Y_{1:j}}(y_{j+1} | x_{j+1}, y_{1:j}) f_{X_{j+1}|Y_{1:j}}(x_{j+1} | y_{1:j}) f_{Y_{1:j}}(y_{1:j})}{f_{Y_{1:j}, Y_{j+1}}(y_{1:j}, y_{j+1})} \\
&= \frac{f_{Y_{j+1}|X_{j+1}, Y_{1:j}}(y_{j+1} | x_{j+1}, y_{1:j}) f_{X_{j+1}|Y_{1:j}}(x_{j+1} | y_{1:j})}{\frac{f_{Y_{1:j}, Y_{j+1}}(y_{1:j}, y_{j+1})}{f_{Y_{1:j}}(y_{1:j})}} \\
&= \frac{f_{Y_{j+1}|X_{j+1}}(y_{j+1} | x_{j+1}) f_{X_{j+1}|Y_{1:j}}(x_{j+1} | y_{1:j})}{f_{Y_{j+1}|Y_{1:j}}(y_{j+1} | y_{1:j})}. \tag{1.9}
\end{aligned}$$

In (1.9), the denominator depends only on the observations, while the factor $f_{Y_{j+1}|X_{j+1}}$ can be found by relation (1.7). The formula provided in (1.9) gives the update of the forecast pdf to the analysis pdf.

By repeating the forecast-analysis cycle, one can update the filtering distribution in real time. So far, we have not taken any assumption of the form of distributions mentioned above, and it may result that there might be no closed solutions for (1.8) and for (1.9). In

this situation, numerical updates would be preferable. In the next section, we introduce a very special situation for which all solutions for the filtering distribution are closed. Later we will extend this to a more general situation with less assumptions and restrictions.

1.4 Kalman Filter

The material discussed in this section is based on reference [76]. We consider the system provided by the equations (1.3) and (1.7), but impose the following special form:

$$X_{j+1} = A_j X_j + \xi_j, \quad (1.10)$$

$$Y_j = H_j X_j + \eta_j. \quad (1.11)$$

Here, X_j are random variables taking values in \mathbb{R}^ℓ , ξ_j are independent random variables with values in \mathbb{R}^ℓ and following the normal distributions $N(0, \Sigma_j)$ with $\Sigma_j \in \mathbb{R}^{\ell \times \ell}$ and $\Sigma_j > 0$, and η_j are independent random variables with values in \mathbb{R}^m and following the distributions $N(0, \Gamma_j)$ with $\Gamma_j \in \mathbb{R}^{m \times m}$ and $\Gamma_j > 0$. The functions Ψ and h in (1.3) and (1.7) are replaced by the matrices $A_j \in \mathbb{R}^{\ell \times \ell}$ and $H_j \in \mathbb{R}^{m \times \ell}$ respectively. Note that these matrices can be different for different time j . Y_j correspond to the resulting observations, namely to some random variables with values in \mathbb{R}^m . The equation (1.10) holds for any $j \in \mathbb{N}$, with an initial condition X_0 following a distribution $N(m_0^a, P_0^a)$, while (1.11) is valid for $j \in \mathbb{N}^*$.

With these assumptions, the two equations (1.10) and (1.11) are linear, and this framework leads to closed solutions for the filtering problem mentioned in previous section, as we present now. First of all, relation (1.10) implies that $X_{j+1} - A_j X_j$ follows the distribution $N(0, \Sigma_j)$. Thus, if X_j takes the value x_j , then one has

$$(X_{j+1} | X_j = x_j) \sim N(A_j x_j, \Sigma_j), \quad (1.12)$$

where \sim precisely means that the l.h.s. follows the distribution provided by the r.h.s. Similarly, one infers from (1.11) that $(Y_j | X_j = x_j) \sim N(H_j x_j, \Gamma_j)$. Based on these relations, we derive the solutions of the forecast and of the analysis distributions, and show how the filtering problem can be solved. The proof is based on two lemmas which we recall:

Lemma 1 (Joint distribution of Gaussian variables, Lemma A.1 of [76]) *Let X be a random variable with values in \mathbb{R}^ℓ and Y be a random variable with values in \mathbb{R}^m which satisfy the following properties*

$$\begin{aligned} X &\sim N(v, P), \\ (Y | X) &\sim N(HX + u, Q), \end{aligned}$$

with $v \in \mathbb{R}^\ell$, $u \in \mathbb{R}^m$, $H \in \mathbb{R}^{m \times \ell}$, $P \in \mathbb{R}^{\ell \times \ell}$ and $P > 0$, $Q \in \mathbb{R}^{m \times m}$ and $Q > 0$. Then the joint distribution of X and Y satisfies

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left(\begin{pmatrix} v \\ H v + u \end{pmatrix}, \begin{pmatrix} P & P H^T \\ H P & H P H^T + Q \end{pmatrix} \right),$$

while the marginal distribution of Y is given by

$$Y \sim N(Hv + u, HPH^T + Q).$$

In the statement of this lemma, the notation H^T is used for the transpose of the matrix H . The same notation is used in the sequel.

Lemma 2 (Conditional distribution of Gaussian variables, Lemma A.2 of [76]) *If the random variables X and Y follow a distribution of the form:*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} A & C \\ C^T & B \end{pmatrix} \right),$$

for suitable vectors a and b , and conformable matrices A , B , and C , then the marginal distributions of X and Y , and conditional distribution $(X | Y)$ and $(Y | X)$ satisfy

$$\begin{aligned} X &\sim N(a, A), \\ Y &\sim N(b, B), \\ (X | Y) &\sim N(a + CB^{-1}(Y - b), A - CB^{-1}C^T), \\ (Y | X) &\sim N(b + C^T A^{-1}(X - a), B - C^T A^{-1}C). \end{aligned}$$

Let us provide the closed formula for the Kalman filter, following [76, p. 57–58]. We start with the computation of the forecast distribution, namely the pdf $f_{X_{j+1}|Y_{1:j}}$. Suppose at the time j , the analysis distribution is Gaussian and satisfies

$$X_j^a := (X_j | Y_{1:j}) \sim N(m_j^a, P_j^a). \quad (1.13)$$

By the rule of conditional probabilities, the pdf $f_{X_{j+1}, X_j | Y_{1:j}}$ can be rewritten as follows:

$$f_{X_{j+1}, X_j | Y_{1:j}} = f_{X_{j+1}, X_j^a} = f_{X_{j+1} | X_j^a} f_{X_j^a}. \quad (1.14)$$

The pdf represented by the first factor on the r.h.s. of (1.14) follows the distribution $N(A_j X_j^a, \Sigma_j)$ by (1.12). The pdf represented by the second factor on the r.h.s. of (1.14) follows also a normal distribution, by the assumption (1.13). Thus, by using Lemma 1, we infer that the joint distribution of X_j^a and X_{j+1} satisfies

$$\begin{pmatrix} X_j^a \\ X_{j+1} \end{pmatrix} \sim N \left(\begin{pmatrix} m_j^a \\ A_j m_j^a \end{pmatrix}, \begin{pmatrix} P_j^a & P_j^a A_j^T \\ A_j P_j^a & A_j P_j^a A_j^T + \Sigma_j \end{pmatrix} \right). \quad (1.15)$$

By using Lemma 2 and (1.15), we then obtain the forecast distribution, namely

$$X_{j+1}^f := (X_{j+1} | Y_{1:j}) \sim N(A_j m_j^a, A_j P_j^a A_j^T + \Sigma_j) =: N(m_{j+1}^f, P_{j+1}^f). \quad (1.16)$$

Now let us compute the analysis distribution $(X_{j+1} | Y_{1:j+1})$, once the observation Y_{j+1} is available. Consider the conditional joint distribution of X_{j+1} and Y_{j+1} given $Y_{1:j}$:

$$f_{X_{j+1}, Y_{j+1} | Y_{1:j}} = f_{Y_{j+1} | X_{j+1}} f_{X_{j+1} | Y_{1:j}} = f_{Y_{j+1} | X_{j+1}} f_{X_{j+1}^f}. \quad (1.17)$$

The first factor on the r.h.s of (1.17) follows a distribution $N(H_{j+1}X_{j+1}, \Gamma_{j+1})$, the second factor has the distribution given by (1.16). By Lemma 1, it follows that

$$\left(\begin{pmatrix} X_{j+1}^f \\ Y_{j+1} \end{pmatrix} \right) \sim N \left(\begin{pmatrix} m_{j+1}^f \\ H_{j+1}m_{j+1}^f \end{pmatrix}, \begin{pmatrix} P_{j+1}^f & P_{j+1}^f H_{j+1}^T \\ H_{j+1}P_{j+1}^f & H_{j+1}P_{j+1}^f H_{j+1}^T + \Gamma_{j+1} \end{pmatrix} \right). \quad (1.18)$$

By an application of Lemma 2, one gets

$$Y_{j+1} \sim N(H_{j+1}m_{j+1}^f, H_{j+1}P_{j+1}^f H_{j+1}^T + \Gamma_{j+1}).$$

Finally, by using Lemma 2 and (1.18), one obtains

$$(X_{j+1} | Y_{1:j+1}) \sim N(m_{j+1}^a, P_{j+1}^a), \quad (1.19)$$

where

$$m_{j+1}^a = m_{j+1}^f + P_{j+1}^f H_{j+1}^T (H_{j+1}P_{j+1}^f H_{j+1}^T + \Gamma_{j+1})^{-1} (Y_{j+1} - H_{j+1}m_{j+1}^f) \quad (1.20)$$

$$P_{j+1}^a = P_{j+1}^f - P_{j+1}^f H_{j+1}^T (H_{j+1}P_{j+1}^f H_{j+1}^T + \Gamma_{j+1})^{-1} H_{j+1}P_{j+1}^f \quad (1.21)$$

In (1.20), we observe that the mean of the analysis distribution is a sum of the mean of the forecast distribution and a weighted difference between the latest observation and a projection of the forecast mean to the observation space. The weight $P_{j+1}^f H_{j+1}^T (H_{j+1}P_{j+1}^f H_{j+1}^T + \Gamma_{j+1})^{-1}$ is called Kalman gain matrix and is usually denoted by K_{j+1} . By recursively computing the forecast distribution (1.16) and analysis distribution (1.19), one explicitly solves the data assimilation problem of the system given by (1.10) and (1.11).

1.5 Ensemble Kalman filter (EnKF)

As mentioned in the previous section, the Kalman filter provides a closed solution of the forecast distribution and the analysis distribution. In this framework, since both distributions are Gaussian, the mean can be used as an optimal estimator of the unknown signal, while the covariance matrix provides the uncertainty information on the estimates. However, Kalman filter requires strong restrictions on the dynamical model, on the observation model, and on the initial condition. Such requirements are difficult to meet in a general scenario. We shall thus relax the assumptions on A_j and on H_j , and re-write equations (1.10) and (1.11) in the more general form

$$X_{j+1} = \mathcal{A}_j(X_j) + \xi_j, \quad (1.22)$$

$$Y_j = \mathcal{H}_j(X_j) + \eta_j. \quad (1.23)$$

where $\mathcal{A}_j : \mathbb{R}^\ell \rightarrow \mathbb{R}^\ell$ and $\mathcal{H}_j : \mathbb{R}^\ell \rightarrow \mathbb{R}^m$ are functions. As a rule, all linear expressions available for the Kalman filter can no more be used in this context.

In order to develop a new machinery, let us consider again the solutions (1.20) and (1.21) of the Kalman filter. In these expressions, the forecast mean m_{j+1}^f and its covariance

P_{j+1}^f are transformed into the analysis mean m_{j+1}^a and its covariance P_{j+1}^a . In the ensemble Kalman filter, the analysis mean is computed from a set of analysis ensemble members $\{x_{j+1,1}^a, x_{j+1,2}^a, \dots, x_{j+1,N}^a\}$. These realizations of the analysis ensemble members are generated by a set of the forecast or background ensemble members $\{x_{j+1,1}^f, x_{j+1,2}^f, \dots, x_{j+1,N}^f\}$ through a suitable adaptation of (1.20), as described below.

Suppose at certain time point $j+1$, a set of background ensemble members is given. Then the analysis ensemble members can be updated by [41, eq. (4)]:

$$x_{j+1,i}^a = x_{j+1,i}^f + K_{j+1}(y_{j+1} - \mathcal{H}_{j+1}(x_{j+1,i}^f)) \quad (1.24)$$

where K_{j+1} is still called the Kalman gain matrix. However, the formula

$$K_{j+1} = P_{j+1}^f \mathcal{H}_{j+1}^T (\mathcal{H}_{j+1} P_{j+1}^f \mathcal{H}_{j+1}^T + \Gamma_{j+1})^{-1} \quad (1.25)$$

has now to be suitably interpreted, since it does not correspond to a product of matrices.

For that purpose, let us firstly recall the original idea of the ensemble Kalman filter, as introduced in [21]. By definition, the error covariance matrices P_{j+1}^f and P_{j+1}^a for the forecast and for the analysis are given by

$$P_{j+1}^f = \mathbb{E}((X_{j+1}^f - x_{j+1}^t)(X_{j+1}^f - x_{j+1}^t)^T), \quad (1.26)$$

$$P_{j+1}^a = \mathbb{E}((X_{j+1}^a - x_{j+1}^t)(X_{j+1}^a - x_{j+1}^t)^T), \quad (1.27)$$

where \mathbb{E} represents the expectation and x_{j+1}^t is the true state [23, eq. (11), (12)]. Since the true state is unknown, these two matrices can be estimated by replacing the true state by the forecast ensemble mean and the analysis ensemble mean respectively, namely

$$P_{j+1}^f = \mathbb{E}\left(\left(X_{j+1}^f - \overline{x_{j+1}^f}\right)\left(X_{j+1}^f - \overline{x_{j+1}^f}\right)^T\right) = \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^f - \overline{x_{j+1}^f}\right) \left(x_{j+1,i}^f - \overline{x_{j+1}^f}\right)^T, \quad (1.28)$$

$$P_{j+1}^a = \mathbb{E}\left(\left(X_{j+1}^a - \overline{x_{j+1}^a}\right)\left(X_{j+1}^a - \overline{x_{j+1}^a}\right)^T\right) = \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^a - \overline{x_{j+1}^a}\right) \left(x_{j+1,i}^a - \overline{x_{j+1}^a}\right)^T, \quad (1.29)$$

where the overline expressions represent the forecast ensemble mean and the analysis ensemble mean, respectively. Note that the last two expressions consist in the adjusted sample covariance, which is an unbiased estimator.

When \mathcal{H}_{j+1} is linear, namely when \mathcal{H}_{j+1} is represented by a matrix $H_{j+1} \in \mathbb{R}^{m \times \ell}$, the

expression $P_{j+1}^f H_{j+1}^T$ appearing in (1.25), with P_{j+1}^f provided by (1.28), can be rewritten as

$$\begin{aligned} P_{j+1}^f H_{j+1}^T &= \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right) \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right)^T H_{j+1}^T \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right) \left(H_{j+1} (x_{j+1,i}^f - \overline{x_{j+1}^f}) \right)^T \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right) \left(H_{j+1} x_{j+1,i}^f - \overline{H_{j+1} x_{j+1}^f} \right)^T. \end{aligned}$$

Similarly, for another expression appearing in (1.25) the following equality hold:

$$\begin{aligned} H_{j+1} P_{j+1}^f H_{j+1}^T &= H_{j+1} \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right) \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right)^T H_{j+1}^T \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(H_{j+1} (x_{j+1,i}^f - \overline{x_{j+1}^f}) \right) \left(H_{j+1} (x_{j+1,i}^f - \overline{x_{j+1}^f}) \right)^T \\ &= \frac{1}{N-1} \sum_{i=1}^N \left(H_{j+1} x_{j+1,i}^f - \overline{H_{j+1} x_{j+1}^f} \right) \left(H_{j+1} x_{j+1,i}^f - \overline{H_{j+1} x_{j+1}^f} \right)^T. \end{aligned}$$

Inspired by the previous expressions, and following [40, eq. (2), (3)], we can now provide the meaning of the expressions appearing in (1.25):

$$\begin{aligned} P_{j+1}^f \mathcal{H}_{j+1}^T &:= \frac{1}{N-1} \sum_{i=1}^N \left(x_{j+1,i}^f - \overline{x_{j+1}^f} \right) \left(\mathcal{H}_{j+1}(x_{j+1,i}^f) - \overline{\mathcal{H}_{j+1}(x_{j+1}^f)} \right)^T \\ \mathcal{H}_{j+1} P_{j+1}^f \mathcal{H}_{j+1}^T &:= \frac{1}{N-1} \sum_{i=1}^N \left(\mathcal{H}_{j+1}(x_{j+1,i}^f) - \overline{\mathcal{H}_{j+1}(x_{j+1}^f)} \right) \left(\mathcal{H}_{j+1}(x_{j+1,i}^f) - \overline{\mathcal{H}_{j+1}(x_{j+1}^f)} \right)^T. \end{aligned}$$

As emphasized in [40], the first expression can be understood as the cross-covariance matrix between X_{j+1}^f and $\mathcal{H}(X_{j+1}^f)$, while the second expression corresponds to the covariance matrix of $\mathcal{H}(X_{j+1}^f)$.

The expression for the Kalman gain matrix provided in (1.25) is now fully explained. As a result, each analysis ensemble member can be computed through (1.24). The mean value of the analysis ensemble members becomes the optimal estimate of the signal, and the spread of this ensemble becomes a natural definition of the error to the mean [23, Sec. 3.4.1]. To generate the next forecast ensemble, one uses the equation (1.22) to evolve each analysis ensemble member. As a result, the filtering problem with the nonlinear equations (1.22) and (1.23) can be solved in real time.

1.6 Several variants to EnKF

The EnKF introduced in the previous section has been further developed in many directions, and this have led to several advanced algorithms. In this section, we sketch some of them, and in particular mention the ones that will be used in the following chapters. Additional information can be found in [41, Sec. 2], from which this section is inspired.

The computation of the analysis error covariance, as described in (1.21), has the troublesome feature of underestimating the error covariance of the analysis mean. This outcome is partially due to the fact that every ensemble members are updated by using the same observations [14, Sec. 3]. Based on this finding, a refined version of EnKF consists in treating the observations as random variables. For that purpose, i.i.d. random errors are added to the observation at each time point, and this creates a set of ensemble members [14, Sec. 3, (12)]. Each forecast ensemble member is updated by using the corresponding perturbed observation. The errors which are added to the observations are usually assumed to be Gaussian distributed with zero mean, especially when no additional information is available. This kind of algorithms are called stochastic filters [41, Sec. 2, b(1)].

Let us also mention other types of deterministic filters (in which the observations are not perturbed randomly) as for example the ensemble square root filter (EnSRF), the ensemble adjustment Kalman filter (EAKF), and the ensemble transform Kalman filter (ETKF) [41, Sec. 2, b(2)]. For these filters, the main idea is to assume the existence and the knowledge of the optimal Kalman gain, and to deduce the analysis error covariance from (1.21). From this formula, it follows that the analysis covariance matrix is smaller than the forecast covariance matrix, but unfortunately sampling errors might lead to analysis error bigger than the forecast error. To solve this problem, covariance inflation can be adopted, and this approach leads to results similar to the stochastic filters but with a smaller ensemble size. Detailed information can be found in [41, Sec. 4,a(1,2,3)] about different methods for increasing the ensemble spread. In our work, we have employed ETKF for the study related to COVID-19, and the theory for this filter is provided in Section 2.3.

In order to reduce the cost of numerical computations when the inverse of a matrix is involved, sequential and local filters have been developed [41, Sec. 2, c]. One of the famous local filters is called the local ensemble transform Kalman filter (LETKF). An outline of this filter is presented in Section 4.2.2, since this filter has been used for the control simulation experiment involving the Lorenz-96 model.

In summary, there exist numerous variants of data assimilation techniques, and we have presented only the main ones, and the ones explicitly used in our investigations. We refer to [41] for a recent review paper presenting many outcomes and challenges for these filters and for their applications to atmospheric problems. This reference contains also a very long list of additional references on the subject.

Chapter 2

Analysis of COVID-19 in Japan with Extended SEIR model and ensemble Kalman filter

2.1 Introduction

The outbreak of coronavirus 2019 (COVID-19) had a huge impact on human society, and its devastating effects are still present more than 18 months after its official acknowledgment. One of the specific characteristics of this disease increases the difficulties for any scientific research based on infectious disease models, namely the existence of asymptomatic cases. Their existence, but also their potential infectiousness, blur the full picture of the disease spreading inside the society [95]. The simultaneous existence of mild symptomatic cases, which often remain undocumented, also generates an additional challenge for any epidemiological investigations. For example, one study discussing the potential fraction of undocumented cases at the beginning of the outbreak ends up with a ratio of 86% undocumented infections [49].

There exists now a huge literature related to COVID-19, but only a small number of papers are using techniques of data assimilation, as we shall do. For that reason, we present only a few references closely related to our investigations. To get a better understanding of the spread of COVID-19, some studies start with a specially designed infectious disease model with appropriate structures. Those elaborated models may accurately describe the process in reality but they also bring more uncertainties and unknown parameters. For example in [3], a 22 variables model is adopted with each variable representing a subpopulation. The structure related to asymptomatic, pre-symptomatic, detected, and undetected cases are all properly described by the structure. However, the inference is done with only observations of 4 compartments among all variables. Another example is [25], where an extended SEIR model with 11 age-classes is used to estimate the posterior distribution of parameters and make short period predictions. The distribution of mild, severe and fatal cases for each age class is introduced as pre-defined parameters. Also, a matrix which describes the relative transmissions between each pair of age groups is predefined.

Despite these elaborated models, most of the investigations are still performed on the

common infectious disease models, namely the SIR and SEIR models. Such models are easier to implement and contain less parameters, but they also sacrifice some details of the process. For example the SEIR model is used to study the dynamical behavior of COVID-19 outbreaks at the regional level in [26], and the SIR model was used to track the effective reproduction number for 124 countries in [4]. A slightly more elaborated model is also introduced in [64], namely the SITR model with T standing for *under treatment*, for inferring infection numbers and parameters values.

For dealing with real data, one of the recent tools employed for the study of infectious diseases is data assimilation. The techniques developed for data assimilation have the ability to optimally meld dynamical systems with noisy and imperfect observations. They also provide forecasts and estimations of parameters and variables [74]. As examples of investigations studying the details of the implementation of these techniques to epidemic models, let us mention [61] and [73]. The previously cited work [3] also uses statistical data assimilation for identifying the measurements required for getting accurate states and parameters estimations. Finally, in [31] an ensemble Kalman filter is adopted to forecast the COVID-19 pandemic in Saudi Arabia with an extended SEIR model including vaccination.

Let us now present our research. First, we construct an extended SEIR model which takes into account one of the main specificities of COVID-19: the simultaneous existence of an asymptomatic/pre-symptomatic population and of a symptomatic population. These two infected populations have different characteristics which are encoded with different parameters. The model is also constructed such that it can be fed with the data of only three compartments, namely the Hospitalized (or treated) population, the Recovered population, and the Deceased population, naturally identified by the letters H , R , and D , respectively. Note that the population R or D are coming from the population H , and these three populations are considered as fully recorded by health services. Except for the first couple of weeks at the beginning of the epidemic, these data are also considered as the most reliable ones. With this model and these data, our main aim is to provide information about the effective reproduction number and about the population of asymptomatic/pre-symptomatic or undetected symptomatic, with uncertainty ranges. For these investigations, we shall also mainly concentrate on the population of Tokyo, but provide additional analysis with two additional prefectures (Osaka and Kanagawa) and with the entire population of Japan. Let us mention that targeting a specific population gives the opportunity to rely on local parameters and to use information provided by local medical research or health services. We also emphasize that for data assimilation, we use the ensemble transform Kalman filter [42] which is efficient compared with the standard extended Kalman filter.

As a result of our investigations we get an effective reproduction number for the four populations mentioned before, as well as a time dependent estimate for the number of asymptomatic/pre-symptomatic in these four populations. A comparison between our estimated effective reproduction number and the same statistic provided by Toyokeizai [87] shows the reliability of our strategy. As an asset of our approach, we test the sensitivity of our model and its outcomes to the values of uncertain parameters borrowed from the literature. These experiment's results show a robust performance of the strategy used in our investigations. Another clear result from our computation of the effective reproduction number is the decay

of the effectiveness of the states of emergency as their number increases. This effect is clearly visible independently for the three prefectures and for the entire country. Note that a similar effect is also visible in the use of public transportation, see for example [57]. Other outcomes of our investigation are gathered in Sections 2.4 and 2.5.

Let us finally describe the structure of this paper. In Section 2.2, we recall the standard SIR and SEIR models, and introduce the extended SEIR model. We also provide information for the computation of the effective reproduction number, and discuss the constant parameter values and the observations. In Section 2.3, we explain the technical details of ensemble transform Kalman filter. Readers familiar with data assimilation and Kalman filters can skip this section without any loss for the application to infectious diseases. In Section 2.4, background information about the experiments, and subsequently technical information, are provided. The main results of our investigations are also presented in this section. The discussion and the comparisons between the different populations are provided in Section 2.5. We finally prepare the ground for future projects.

2.2 Extended SEIR model

In this section, we recall a few information on the SIR and SEIR models, and introduce the extended SEIR model. We also discuss medical parameters, and provide some explanations about the observations.

2.2.1 The SIR and SEIR models

The SIR model is a deterministic epidemic model expressed by a system of differential equations. Its construction is based on Kermack–McKendrick theory, and it describes the transmission process of an infectious disease. A given population of size N is divided into three mutually exclusive sub-populations call compartments: the susceptible hosts S , the infectious hosts I , and the recovered hosts R . For any given time t , $S(t)$, $I(t)$, and $R(t)$ describe the number of individuals in each respective compartment, and they satisfy $S(t)+I(t)+R(t) = N$. Individuals in compartment S can be infected by individuals in compartment I through direct contact. Once a successful contact happens, the infected individual leaves compartment S and becomes a member of compartment I . The number of newly infected individuals per unit time is given by $\beta I(t)S(t)/N$, where the parameter β is called the transmission coefficient. Similarly, once a patient recovers, this person leaves compartment I and goes to compartment R immediately. The transfer between I and R is controlled by the recovery rate γ . The transfer rate, namely the number of transfer individuals per unit time, is given by $\gamma I(t)$.

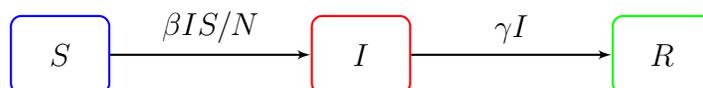


Figure 2.1: Transfer diagram for the SIR model

Figure 2.1 shows the process of the SIR model, and the following differential system describes its dynamic:

$$\begin{aligned}\frac{dS}{dt} &= -\beta IS/N, \\ \frac{dI}{dt} &= \beta IS/N - \gamma I, \\ \frac{dR}{dt} &= \gamma I.\end{aligned}\tag{2.1}$$

Note that the system (2.1) assumes a permanent immunity once recovered. In other words, no individual can be infected a second time. The model also assumes a constant total population: there is no inflow to the system, or outflow from the system.

The SEIR model is very similar to the previous model, but with an additional compartment E between the compartments S and I . The individuals in E are exposed, namely they have been contaminated, but they are not infectious yet. The time spent in E corresponds to the incubation period, before becoming infectious. The transfer rate from E to I is given by $\varepsilon E(t)$.

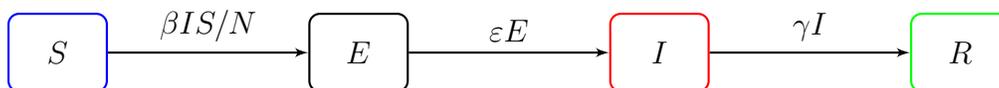


Figure 2.2: Transfer diagram for the SEIR model

Figure 2.2 shows the process of the SEIR model, and the following differential system describes its dynamic:

$$\begin{aligned}\frac{dS}{dt} &= -\beta IS/N, \\ \frac{dE}{dt} &= \beta IS/N - \varepsilon E, \\ \frac{dI}{dt} &= \varepsilon E - \gamma I, \\ \frac{dR}{dt} &= \gamma I.\end{aligned}$$

2.2.2 The extended SEIR model

The extended SEIR model has been developed based on some specific features of the COVID-19 epidemic, as described now. In early 2020, health services already noticed that some infected individuals, who did not show any symptom, could spread the disease. These persons correspond either to asymptomatic hosts or to pre-symptomatic hosts. In the first cohort, people will never show any symptom, while the second cohort corresponds to individuals who will show some symptoms subsequently. Clearly, asymptomatic hosts and pre-symptomatic hosts are difficult to be identified by health services, even though they play an important

role in the spread of the disease. Symptomatic individuals are also infectious, but they can be more easily identified precisely because of their symptoms. Thus, if symptomatic individuals correspond to the compartment I of the above SIR or SEIR models, then there is no compartment left for the asymptomatic or the pre-symptomatic individuals.

The existence of COVID-19 infectious spreaders who do not show (yet or at all) any symptom already brings a lot of uncertainties to health services. In addition, whenever the symptoms are relatively mild, some symptomatic individuals do hesitate to report the health service [66, p. 11]. As a consequence, daily new cases are under-reported. Combining this effect with some delayed information, some inaccurate tests (especially in the first phase of the epidemic), and other reasons that we are not aware of, it turns out that the reported data are not very accurate. In such a situation one needs to use proper strategies to analyse the observation data.

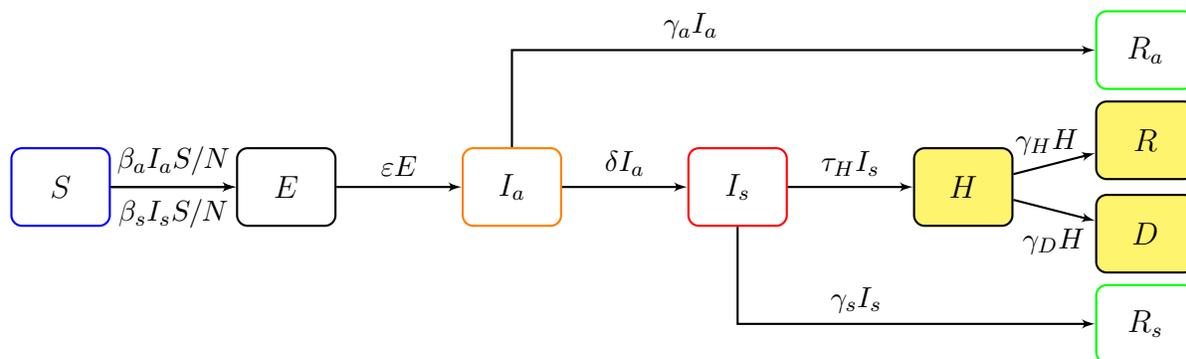


Figure 2.3: Transfer diagram for the extended SEIR model. Compartment I of the SEIR model is divided into two compartments I_a (asymptomatic/pre-symptomatic) and I_s (symptomatic)

In order to meet the special features of COVID-19, we separate the compartment I of the SEIR model into two compartments I_a and I_s . These compartments correspond to asymptomatic/pre-symptomatic and to symptomatic states, respectively. Both I_a and I_s can infect S . As shown in Figure 2.3, the transmission processes related to I_a and to I_s have transmission coefficients β_a and β_s respectively. Once an individual in S gets infected, this person becomes a member of E , and then moves to I_a when it becomes infectious. In I_a , part of these individuals (the asymptomatic) will never generate any symptoms. They will thus spend some time in I_a , and then recover. The corresponding compartment is denoted by R_a . In contrast, the other part of individuals in I_a will develop symptoms, and then move to I_s . In I_s , a majority of individuals will be identified by health services, but as mentioned above a minority will recover without being identified by any health services. Compartment R_s corresponds to those recovered individuals who have not been registered. The identified persons in I_s will then move to the compartment H , which corresponds to hospitalized or treated patients. The ones staying at home but under medical control, or the ones isolated in a hotel, are all included in the compartment H . Finally, the patients in H will recover, and thus move to the compartment R , or will decrease and end up in the compartment D . As

for the SIR or SEIR models, we assume a permanent immunity, which means that there is no flow from R_a , R_s , or R to S . Also, the total number N of individual is constant, namely at all time one has

$$N = S + E + I_a + I_s + H + R + D + R_a + R_s. \quad (2.2)$$

Based on the Figure 2.3 and on the explanations provided above, the differential system of the extended SEIR model reads:

$$\begin{aligned} \frac{dS}{dt} &= -\beta_a I_a S/N - \beta_s I_s S/N & \frac{dR}{dt} &= \gamma_H H \\ \frac{dE}{dt} &= \beta_a I_a S/N + \beta_s I_s S/N - \varepsilon E & \frac{dD}{dt} &= \gamma_D H \\ \frac{dI_a}{dt} &= \varepsilon E - \delta I_a - \gamma_a I_a & \frac{dR_a}{dt} &= \gamma_a I_a \\ \frac{dI_s}{dt} &= \delta I_a - \tau_H I_s - \gamma_s I_s & \frac{dR_s}{dt} &= \gamma_s I_s \\ \frac{dH}{dt} &= \tau_H I_s - \gamma_H H - \gamma_D H \end{aligned} \quad (2.3)$$

where $\beta_a, \beta_s, \varepsilon, \delta, \gamma_a, \gamma_s, \gamma_D$ and γ_H are some medical parameters. Note that some of them are time dependent.

2.2.3 The reproduction number

The basic reproduction number, denoted by R_0 , can be interpreted as the average number of secondary cases generated by on primary case in a susceptible population. It is commonly admitted that $R_0 = 1$ is a threshold value, as explained below. We also refer to [18, 35] for more explanations and more precise statements.

To study R_0 for different models, a general definition of the basic reproduction number is introduced based on the notion of disease free equilibrium (DFE). In a DFE the population remains in the absence of disease. For example, in the SIR or SEIR models, it means that $I = 0$ while in the extended SEIR model it means that $I_a = I_s = 0$. In this context, the basic reproduction number can be defined as the number of new infections produced by a typical infectious individual in a population at a DFE. The main feature of R_0 is that it corresponds to a threshold parameter, namely if $R_0 < 1$, the DFE is locally asymptotically stable, while if $R_0 > 1$, the DFE is unstable and an outbreak is possible.

The precise expression for R_0 is clearly model dependent, but numerous examples are available in the literature. For example, let us consider a general compartmental disease transmission model. Such models are represented by a system of ordinary differential equations, and under natural assumptions it can be shown that these models have a DFE, see [18]. In this reference, the precise expression for R_0 is then provided for some classes of models, and the extended SEIR model meets the assumptions of the staged progression model, as presented in [18, Sec. 4.3]. For the extended SEIR model, the expression for R_0 then reads:

$$R_0 = \frac{\beta_a}{\delta + \gamma_a} + \frac{\beta_s \delta}{(\delta + \gamma_a)(\gamma_s + \tau_H)}, \quad (2.4)$$

where β_a and β_s are the initial transmission coefficients at time 0.

To understand the above expression, observe that $\delta/(\delta + \gamma_a)$ corresponds to the fraction of individuals of I_a going to the compartment I_s , while $1/(\delta + \gamma_a)$ and $1/(\gamma_s + \tau_H)$ define the average times an infected individual spends in compartments I_a and I_s respectively. Thus, each term on the R.H.S. of (2.4) represents the number of new infections generated by an infectious individual during the time spent in the compartments I_a and I_s .

In contrast, the definition of the effective reproduction number R_t at time t is much more delicate. The various challenges and possible errors have recently been discussed in [30], to which we refer for a thorough discussion. In our setting, we shall keep the interpretation of R_t as the average number of secondary cases generated by one primary case. This approach is possible because the transmission coefficients at time t are available in our approach, and therefore one can compute R_t with (2.4) and the corresponding β_a and β_s at time t . Additional information on R_t will be provided in Section 2.5.

2.2.4 The medical parameters

It clearly appears in Figure 2.3 and in the corresponding system (2.3) that several parameters control the flows between the compartments. The values of these parameters may result in very different behaviors of the model. We refer for example to [48, Sec. 1.4] and to [10, Chap. 2] for general discussions of model behaviors and the role of parameters. In our setting, some parameters are easy to evaluate, as for example the recovery rate γ_H or the death rate γ_D , but others are difficult to estimate, as for example the transmission coefficients β_a and β_s . Let us also stress that some parameters depend on location. Since our experiment is based on data from Japan, see Section 2.4, the medical parameters are chosen accordingly. For that reason, we use research or survey results specific to Japan, or even more precisely to specific prefectures in Japan.

In Table 2.1 we list the estimated values of some parameters for the extended SEIR model, and provide the sources of information. Several parameters in the table have the form of the product of a percentage and the inverse of a time length. A similar setting for the construction of the parameters can be found for example in [25, 46]. For δ and γ_a , the percentage parts should sum up to 1. For the parameters τ_H and γ_s , some information can be found in the surveys [66, 13] and the health services website [85]. For parameter γ_H and γ_D , instead of using constant value estimated by health services, we shall use the observation data to estimate their values on a daily basis. The details will be discussed in Section 2.4.

Let us stress that the value assigned to some of these parameters has evolved during the last 12 months. For example, the ratio of asymptomatic individuals was thought to be very high at the beginning of the epidemic (up to 80%), but some recent research [15, 12, 72] have revised this ratio to 17% to 20%. Our knowledge about the length of infectious periods has also evolved, and the possible values are spread over a rather long interval. In Table 2.1 we list some mean values, but in the simulations additional uncertainties are implemented. Since pre-symptomatic patients become infected before the appearance of symptoms, the incubation period (encoded in ε) has been shorten a little bit, and the last 2 days of this incubation period have been moved to I_a .

One very delicate question is the relation between β_a and β_s , namely between the transmission coefficient for asymptomatic/pre-symptomatic and the transmission coefficient for symptomatic. For our investigations, we shall rely on the result of the systematic review [15] which asserts that the relative risk of asymptomatic transmission is 42% lower than that for symptomatic transmission. As a consequence, we shall fix

$$\beta_a = 0.58\beta_s \quad \text{or equivalently} \quad \beta_s = 1.72\beta_a. \quad (2.5)$$

This factor 0.58 is slightly smaller but of a comparable scale compared to earlier investigations, see for example [12].

parameter	estimation	source	remark
ε	$(3 \text{ days})^{-1}$	[59]	incubation period, not contagious
δ	$83\% \times (2 \text{ days})^{-1}$ (95% CI 80% to 86%)	[15, 59]	proportion of pre-symptomatic \times (duration of pre-symptomatic) $^{-1}$
τ_H	$78\% \times (8.3 \text{ days})^{-1}$ ($\sim 2020/5/31$) $78\% \times (5.2 \text{ days})^{-1}$ ($2020/6/1 \sim$)	[66, 56]	proportion of detected symptomatic \times (days of symptoms onset) $^{-1}$
γ_a	$17\% \times (9 \text{ days})^{-1}$ (95% CI 14% to 20%)	[15, 72]	proportion of asymptomatic \times (duration of asymptomatic) $^{-1}$
γ_s	$22\% \times (7 \text{ days})^{-1}$	[66, 13]	proportion of undetected symptomatic \times (duration of symptomatic) $^{-1}$
γ_H	computed by observations	[87]	discussed in Section 2.4
γ_D	computed by observations	[87]	discussed in Section 2.4

Table 2.1: Medical parameters

2.2.5 The observations

As introduced in Section 4.1, one of the aims of this study is to estimate the unknown parameters β_a and β_s , and the unobservable compartments I_a and I_s . Because of relation (2.5), it is clear that for the parameters only β_s has to be evaluated. The method that we are going to introduce in Section 2.3 requires observations of some compartments in the extended SEIR model. In Figure 2.3, we highlight the three compartments with observations in yellow, namely H , R and D . The data corresponding to these compartments may not be perfect, and for the analysis based on these data we shall take some uncertainties into consideration. More precisely, some random noise will be added to the observations.

For our experiment, we shall firstly and mainly use the data about Tokyo, with a total population $N = 13'955'000$ (approximate mean of the population of 2020 and 2021). Subsequently, other prefectures in Japan are also considered. As already mentioned, the compartment H corresponds to the identified individuals either hospitalized or treated at home or in a hotel. On the other hand, R and D describe the accumulated number of recovered and deceased individuals. The information about these three compartments are very easy to find for Tokyo, but also for any region in Japan. One can check for example the website of Ministry of Health, Labour and Welfare, prefecture's websites or some mass communication companies' websites to get more details. The information is provided on a daily basis. Note however, that these records were not very accurate at the beginning of the outbreak. This was caused by the delay of reports, but also by some changes in the policy for the collect of information. Our analysis will take this source of uncertainties into consideration.

2.3 Ensemble Transform Kalman filter

In this section, we introduce the main tool employed in the study, namely the ensemble Kalman filter. The analysis of parameter values and unobservable compartments is performed based on it.

The method Ensemble Transform Kalman Filter (ETKF) used in this paper has been introduced and developed in [8, 42]. It is based on the Ensemble Kalman Filter [21, 23, 24], but it provides analyses in a more efficient way.

Let us consider a discrete time state space model given for any time $t \in \mathbb{N} := \{1, 2, \dots\}$ by

$$\mathbf{x}_t = M_t(\mathbf{x}_{t-1}) \quad (2.6)$$

$$\mathbf{y}_t = H_t(\mathbf{x}_t) + \epsilon_t, \quad (2.7)$$

where \mathbf{x}_t is an l -dimensional state vector and \mathbf{y}_t an m -dimensional observational vector, M_t is the operator which defines the dynamics of the state, H_t is an operator corresponding to the observation model, and ϵ_t provides the observation error. All these vectors or operators are explicitly t dependent. The vector \mathbf{x}_t represents the state of the dynamical system, while \mathbf{y}_t is called the noisy observation, both at time t . The observation error ϵ_t follows a normal distribution $N(0, \mathbf{R}_t)$, where \mathbf{R}_t is an $m \times m$ observation error covariance matrix. In this framework, the general question is: given a list of noisy, unbiased observations $(\mathbf{y}_t)_t$, how can one find the best estimates for $(\mathbf{x}_t)_t$?

Under the assumption of unbiased Gaussian observation error and of Gaussian initial distributions, one can consider the maximum likelihood approach. For time $t = 1, \dots, N$, the likelihood of $(\mathbf{x}_t)_t$ is proportional to

$$\prod_{t=1}^N \exp \left(-\frac{1}{2} [\mathbf{y}_t - H_t(\mathbf{x}_t)]^T \mathbf{R}_t^{-1} [\mathbf{y}_t - H_t(\mathbf{x}_t)] \right), \quad (2.8)$$

where the superscript T denotes the transpose of a vector or of a matrix. Clearly, maximizing (2.8) can be transformed into a minimization problem. More precisely, by using equation (2.6), one first defines the cost function

$$\sum_{t=1}^N [\mathbf{y}_t - H_t(M_t(\mathbf{x}_{t-1}))]^T \mathbf{R}_t^{-1} [\mathbf{y}_t - H_t(M_t(\mathbf{x}_{t-1}))]. \quad (2.9)$$

Then, one ends up in looking for a list of estimates $(\mathbf{x}_t^a)_t$ of $(\mathbf{x}_t)_t$ which minimize (2.9).

When the operators M_t and H_t are linear, they can be represented by matrices, denoted also by M_t and H_t . In such a situation, and with our assumption of Gaussian observation error, the linear model leads to the propagation of Gaussian distributions for the states. More precisely, assume that at some time t , there is a prior estimate \mathbf{x}_t^b of the state \mathbf{x}_t with its covariance estimate \mathbf{P}_t^b (also called the background error covariance). Then the Kalman filter [43, 44] provides an optimal solution to minimize the variance of its uncertainty. For this reason, it is known as the minimum variance estimator. Formally, the prior estimate is updated by using the information of the observation \mathbf{y}_t at time t as follows [41, Eq. (1), (2), (8) & (10)]:

$$\mathbf{x}_t^a = \mathbf{x}_t^b + \mathbf{K}_t[\mathbf{y}_t - H_t\mathbf{x}_t^b], \quad (2.10)$$

$$\mathbf{K}_t = \mathbf{P}_t^b H_t^T (H_t \mathbf{P}_t^b H_t^T + \mathbf{R}_t)^{-1}, \quad (2.11)$$

$$\mathbf{P}_t^a = (I - \mathbf{K}_t H_t) \mathbf{P}_t^b, \quad (2.12)$$

$$\mathbf{P}_{t+1}^b = M_{t+1} \mathbf{P}_t^a M_{t+1}^T, \quad (2.13)$$

where \mathbf{K}_t is called the Kalman gain, \mathbf{x}_t^a the analysis, and \mathbf{P}_t^a its estimated covariance. With the analysis, one can use the forecast model (2.6) to evolve the analysis and get the prior estimate and the covariance at the next time step.

When the operators M_t and H_t are nonlinear, the Kalman filter is not applicable in its original form, but a simple extension with linear approximation is known to be effective. The straightforward nonlinear extension of the Kalman filter is known as the extended Kalman filter (EKF). In this case, M_{t+1} and H_t in equation (2.10) to (2.13) are replaced by their linear approximations respectively. For low dimensional models, the cost of EKF is low. However, when the model complexity and the state dimension increases, one may consider the ensemble Kalman filter. In this case, an ensemble of model states is used to estimate the covariance matrices. Namely, let $\{\mathbf{x}_{t-1}^{a(i)}, i = 1, \dots, k\}$ be an ensemble of estimates of the model states at time $t - 1$. By letting this ensemble evolve according to the model (2.6), one gets the forecast or background ensemble $\{\mathbf{x}_t^{b(i)}, i = 1, \dots, k\}$ at time t , where

$$\mathbf{x}_t^{b(i)} = M_t(\mathbf{x}_{t-1}^{a(i)}).$$

The mean of the ensemble can be used as the most probable estimate of the state, with its covariance estimated by the sample covariance of the ensemble:

$$\bar{\mathbf{x}}_t^b = \frac{1}{k} \sum_{i=1}^k \mathbf{x}_t^{b(i)} \quad \text{and} \quad \mathbf{P}_t^b = \frac{1}{k-1} (\mathbf{X}_t^b)(\mathbf{X}_t^b)^T,$$

where \mathbf{X}_t^b is an $l \times k$ matrix with the i^{th} column $\mathbf{x}_t^{b(i)} - \bar{\mathbf{x}}_t^b$.

The next step is to update the forecast ensemble mean and its covariance by using the information of the observation \mathbf{y}_t . The updated ensemble mean is denoted by $\bar{\mathbf{x}}_t^a$ and its covariance by \mathbf{P}_t^a . To get them, one first needs to determine an analysis ensemble $\{\mathbf{x}_t^{a(i)}, i = 1, \dots, k\}$. Similar to the background ensemble, the mean and covariance of the analysis ensemble are given by

$$\bar{\mathbf{x}}_t^a = \frac{1}{k} \sum_{i=1}^k \mathbf{x}_t^{a(i)}, \quad (2.14)$$

$$\mathbf{P}_t^a = \frac{1}{k-1} (\mathbf{X}_t^a)(\mathbf{X}_t^a)^T. \quad (2.15)$$

In our application the operator H_t is linear, so that H_t is given by a matrix. In this case and for ETKF, the computations can be written as follows [41, Eq. (15)~(17)]:

$$\bar{\mathbf{x}}_t^a = \bar{\mathbf{x}}_t^b + \mathbf{X}_t^b \tilde{\mathbf{P}}_t^a (H_t \mathbf{X}_t^b)^T \mathbf{R}_t^{-1} (\mathbf{y}_t - \overline{H_t \mathbf{x}_t^b}), \quad (2.16)$$

$$\tilde{\mathbf{P}}_t^a = [(k-1)I + (H_t \mathbf{X}_t^b)^T \mathbf{R}_t^{-1} H_t \mathbf{X}_t^b]^{-1}, \quad (2.17)$$

$$\mathbf{X}_t^a = \mathbf{X}_t^b [(k-1)\tilde{\mathbf{P}}_t^a]^{\frac{1}{2}}, \quad (2.18)$$

where the $k \times k$ matrix $\tilde{\mathbf{P}}_t^a$ is called the analysis error covariance in ensemble space. For (2.18), the background perturbation \mathbf{X}_t^b is transformed into the analysis perturbation \mathbf{X}_t^a with the weight $[(k-1)\tilde{\mathbf{P}}_t^a]^{\frac{1}{2}}$. By adding $\bar{\mathbf{x}}_t^a$ to each column of \mathbf{X}_t^a , one gets the analysis ensemble $\{\mathbf{x}_t^{a(i)}, i = 1, \dots, k\}$ that satisfies (2.14). Note that the ETKF is a deterministic filter since no randomly perturbed observations are used in the computation [94], and it is also a square-root filter because it takes the power $\frac{1}{2}$ of the matrix $\tilde{\mathbf{P}}_t^a$ [84].

To avoid a variance underestimation, caused for example by the limited ensemble size, an artificial inflation of the ensemble spread is usually applied. In [41], several ways to perform the variance inflation are discussed. In our study, we apply multiplicative inflation and additive inflation. For multiplicative inflation, the background error covariance \mathbf{P}_t^b is multiplied by a tunable factor $\rho > 1$ before the analysis. The multiplicative inflation can be thought as a procedure to increase the influence of the current observations on the analysis. For additive inflation, since the extended SEIR model is a conserved closed system, it is then applied by adding a random vector to the background ensemble. The random vector can be sampled from the normal distribution with mean $\mathbf{0}$ and some covariance matrix. This covariance matrix is assumed to be proportional to the background error covariance \mathbf{P}_t^b by a tunable parameter $\alpha \in (0, 1)$ [41].

For the current study with a low dimensional model, we could use EKF or EnKF. However, by considering the possibility of increasing the model complexity in future, we opted for ETKF with 50 ensemble members.

2.4 Experiments

In this section, we explain the design of the experiments and illustrate the outcomes by using observations from Tokyo Metropolis. These observations (H , R , and D) are obtained from the website [87]. Discussions and comparisons with other prefectures are provided in the subsequent section. As mentioned in Section 2.2.5, health officials started providing data from February 2020 but they included some uncertainties caused by delay or by policy changes. Our experiments start with the observations from March 6th, 2020, when the record of data became more systematic.

2.4.1 The experiments, using data from Tokyo

For the state space model mentioned in Section 2.3, the dynamical system (2.6) is described by the differential system (2.3). To estimate the parameter β_s , we use the augmented state by adding one more equation $\frac{d\beta_s}{dt} = 0$ to the system (2.3), assuming persistence for β_s . Namely, β_s will stay at the same value during the time integration process, and will be updated during the analysis step of the data assimilation. In other words, if β_s has a correlation with the observations, β_s will be updated together with the other states. In system (2.3), the unit of time $t = 1$ represents one day. Thus, the one day forecast from day n is obtained by integrating the system (2.3) on $[n, n + 1]$ with the initial values equal to the analysis on day n . To avoid negative values in the analysis step, all the compartments are transformed to log scale with base e . The lower case will be used for these new variables, as for example $s(t) := \log S(t)$, and the corresponding equation becomes $\frac{ds(t)}{dt} = \frac{1}{S(t)} \frac{dS(t)}{dt}$. The DA analysis update directly applies to the log-transformed values.

Now the forecast value of \mathbf{x}_t is a 9-dimensional vector

$$(e(t), i_a(t), i_s(t), h(t), r(t), d(t), r_a(t), r_s(t), \log \beta_s)^T. \quad (2.19)$$

Note that the compartment $s(t)$ is not explicitly included since it can be deduced from the conservation relation (2.2). Before performing the DA analysis update, the above vector has to be projected to the observation space by the observation system (2.7). In our setting, the operator H_t introduced in (2.7) is defined by the projection which sends the forecast vector (2.19) on $(h(t), r(t), d(t))^T$, namely on the three compartments with observations.

To initiate the experiments, initial values for all compartments and parameters have to be provided. When these initial values are far away from the observations, the system goes through an unreasonable transition period. To avoid or reduce the unreasonable transition, preliminary experiments were performed for determining suitable initial values. Namely, we run the system for a few days, starting with a few individuals in some compartments and a presumed value for β_s . Different values of β_s were tested, until values comparable to the data of the compartments H, D, R on day zero (March 6th, 2020) were obtained. The corresponding values for the different compartments were then chosen as initial conditions. However, independent normal distributed random errors are also added to create 50 ensemble members $\{\mathbf{x}_0^{a(i)}, i = 1, \dots, 50\}$. Note that each member $\mathbf{x}_0^{a(i)}$ is a 9-dimensional vector containing the initial conditions for the compartments and for the parameter β_s .

The ensemble on day zero are then integrated by the model (2.3) for one day and the ETKF will update the one day forecast by using the observations. The procedure of the experiments is shown in Figure 2.4.

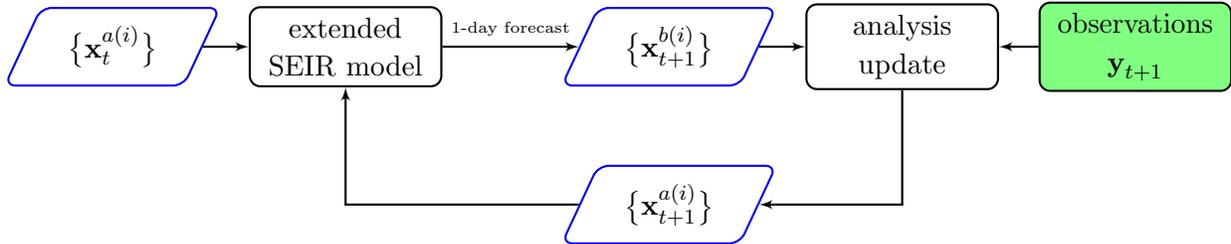


Figure 2.4: Data assimilation flow-chart

As mentioned in Section 2.2.4, parameters γ_H and γ_D are estimated by assimilating the observations. Consider equation $\frac{dR}{dt} = \gamma_H H$ in model (2.3). Since the time is discrete, we can rewrite this relation as $R(t+1) - R(t) = \gamma_H(t)H(t)$ which leads to

$$\gamma_H(t) = \frac{R(t+1) - R(t)}{H(t)}. \quad (2.20)$$

Similar computation can be generated for γ_D also, namely,

$$\gamma_D(t) = \frac{D(t+1) - D(t)}{H(t)}. \quad (2.21)$$

The daily values of both parameters are shown in Figure 2.5. For the implementation of these parameters for the computation of the 1-day forecast, we have used a slightly smoothed version obtained by a 7-day convolution with the symmetric weights $\frac{1}{64}(1, 6, 15, 20, 15, 6, 1)$. The effect is to decrease the amplitude of the weekly oscillations.

For the integration process, we have also included some uncertainties to all parameters: to the ones presented in Table 2.1, but also to the values of the parameters γ_H, γ_D . Thus, the 1-day forecast is obtained with the parameters of the previous day, each of them perturbed by a normal distribution $N(0, (M/10)^2)$, where M corresponds to the value of this parameter. These perturbations are independent and randomly generated for each 1-day forecast.

The last initial setting for data assimilation is the observation error covariance matrix \mathbf{R}_t . We assume that the covariance of the errors between different observations is zero, namely, the matrix \mathbf{R}_t is diagonal. Since we observe three states, \mathbf{R}_t is a 3×3 diagonal matrix. Clearly, the bigger the variance, the weaker contribution the corresponding observation will make to the update.

For simplicity, we assume that the observation error covariance is independent of time. The diagonal elements are chosen as $(\log(1.3))^2$ for any time t , representing the observation error variance: Namely, for $x \in \{h, r, d\}$ one has $[x - \log(1.3), x + \log(1.3)]$ as 68% CI and $[x - 2\log(1.3), x + 2\log(1.3)]$ as 95% CI. Considering that all the observations have been transformed into log scale, that is, the observations in original scale are distributed in $[X/(1.3), (1.3)X]$ as 68% CI and $[X/(1.69), (1.69)X]$ as 95% CI.

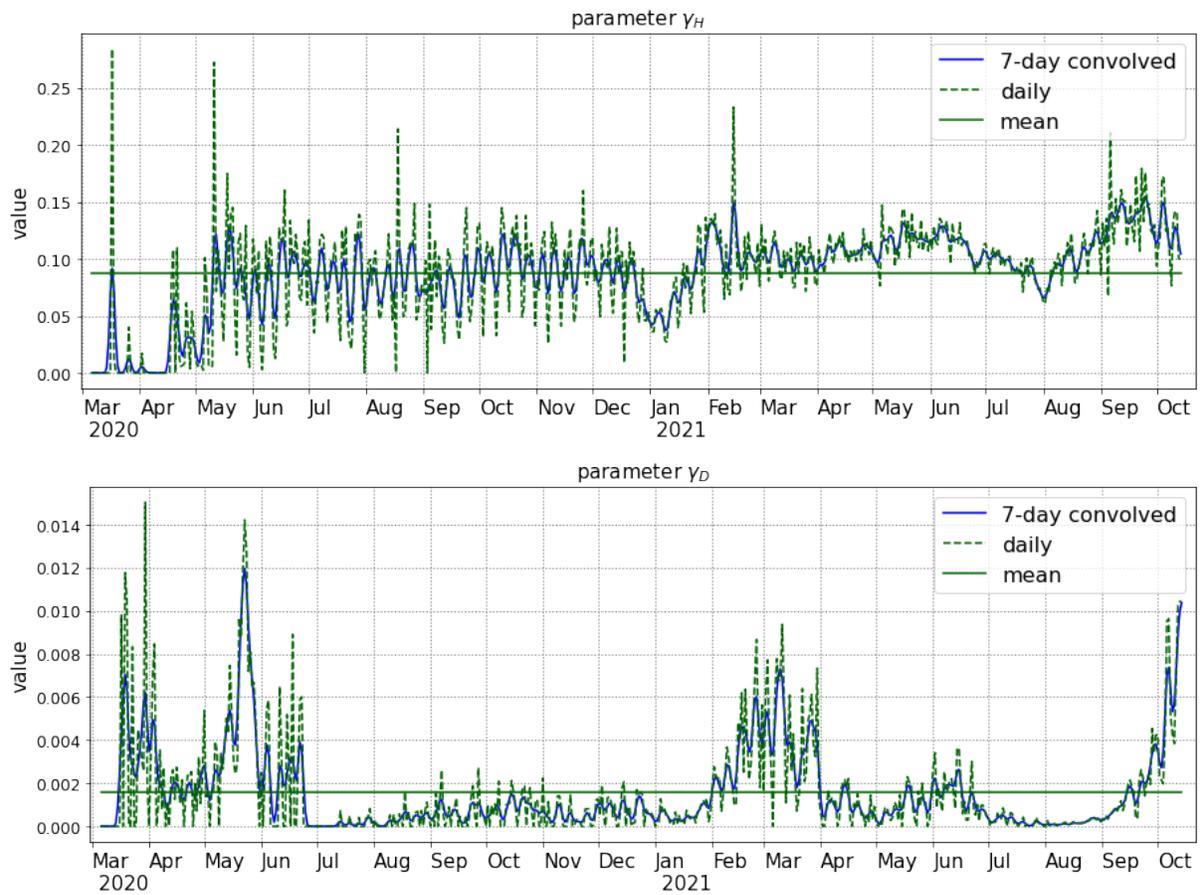


Figure 2.5: Daily estimation of the parameters γ_H and γ_D

Before showing the analysis result, a scatter plot of the values of h and of $\log(\beta_s)$ for the different ensemble members on a certain day is provided in Figure 2.6. Indeed, for the parameter estimation, one assumption is the linear relation between the ensemble of parameters and the ensemble of compartments with observations. As shown in Figure 2.6, a positive correlation is observed between h and $\log(\beta_s)$. Thus we can use the observation of H for the estimation of β_s .

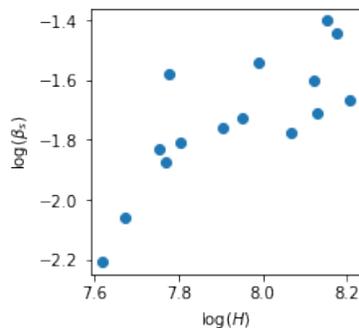


Figure 2.6: Scatter plot of h and $\log(\beta_s)$ on November 17, 2020

In Figure 2.7 we finally provide the analysis value for R_t from March 6th 2020 to the middle of October 2021. The black curve represents the mean value of the ensemble. The two shaded regions represent the 68% CI (dark orange) and 95% CI (light orange). The analysis value of R_t is computed by using β_s together with the information contained in (2.4) and in (2.5). The red curve shows the R_t provided by Toyokeizai.net for reference. The grey shaded regions represent the periods of state of emergency in Tokyo. Note that the jump at the end of May 2020 is due to an abrupt change in the value of τ_H as indicated in Table 2.1. Additional discussions about this graph will be provided in the subsequent section.

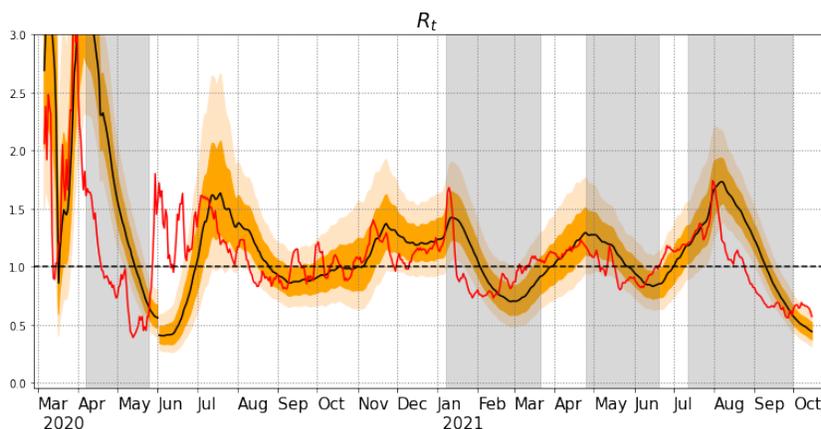


Figure 2.7: Analysis value for R_t , and comparison with the value of Toyokeizai.net

The analysis results for the compartments E , I_a , I_s , H , R , D with confidence intervals

are also presented in Figure 2.8. For compartments H , R , D , the red dots represent the observations of Tokyo. For all pictures, the y -axis is in \log_{10} scale.

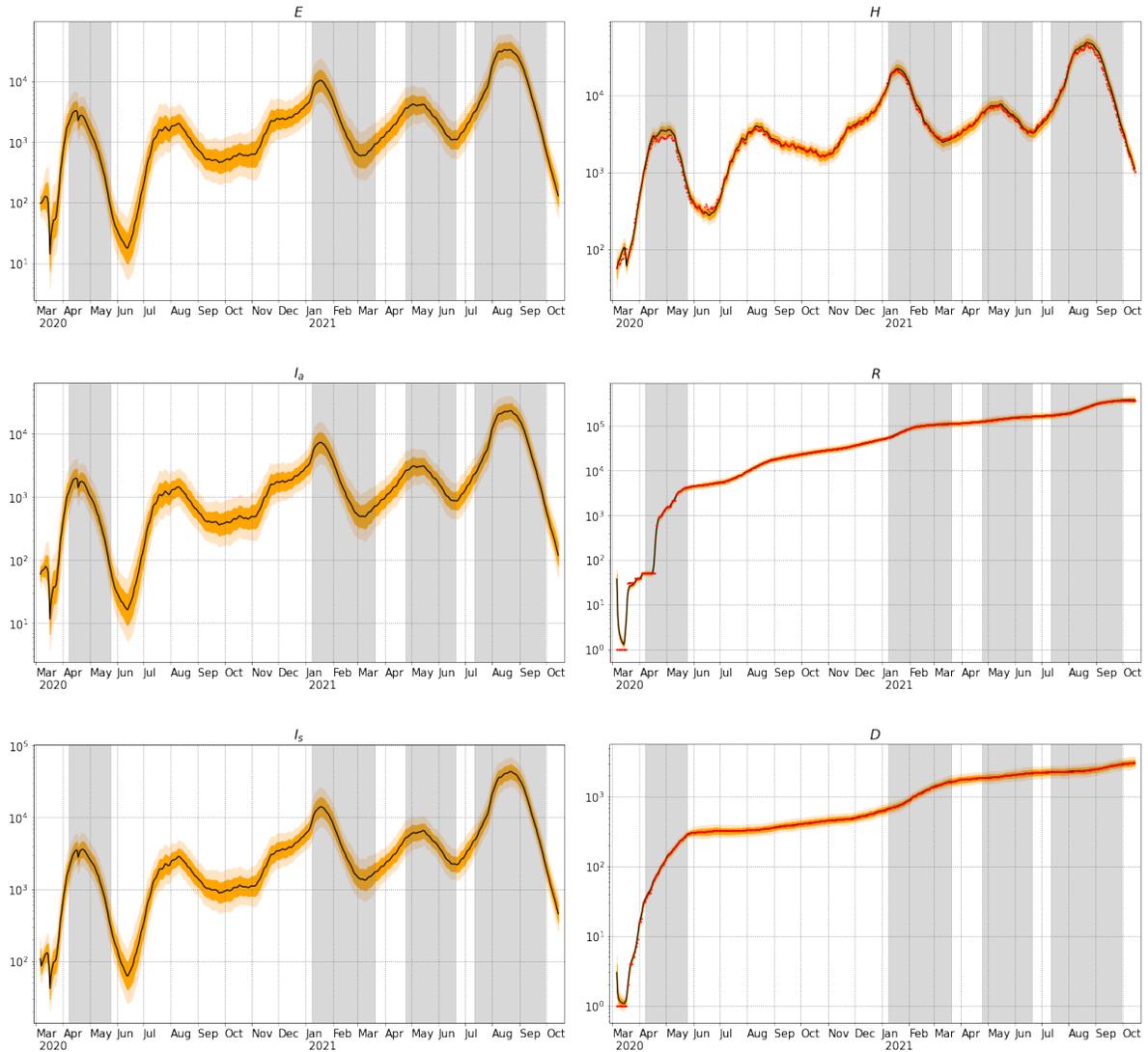


Figure 2.8: The 6 compartments for Tokyo

2.4.2 Technical discussion, using data from Tokyo

In Figure 2.7, one first notices that the similarity between the analysis R_t and the reference R_t provided by Toyokeizai.net. The Toyokeizai R_t on day t is computed by the following

formula:

$$\left(\frac{\sum_{i=t-6}^t \text{new confirmed cases on day } i}{\sum_{i=t-13}^{t-7} \text{new confirmed cases on day } i} \right) \frac{\text{average generation time}}{\text{reporting interval}},$$

where the average generation time is assumed to be 5 days and the reporting interval is assumed to be 7 days. The formula is a simplified version of a maximum likelihood estimation for the effective reproduction number provided by H. Nishiura [65]. Note that this formula is based on the daily confirmed cases, which experiences quick changes everyday. On the other hand, the analysis R_t and its confidence interval is computed by formula (2.4) which involve the analysis ensemble of β_s and β_a . The average over the ensemble member provides a smoother evolution of R_t , which is certainly closer to a true statistic's evolution. Note that a third approach for the computation of R_t is available in [83]: it involves an agent-based model together with a particle filter.

In Figure 2.8, we show the analysis mean of compartments E , I_a , I_s with the 68% CI (dark orange) and 95% CI (light orange), and the same results for compartments H , R and D with the daily observations represented by red dots. One notes that the analyses of H , R , D fit the observations appropriately except at the beginning of the observation period. This may be related to the initial guess and the model imperfections.

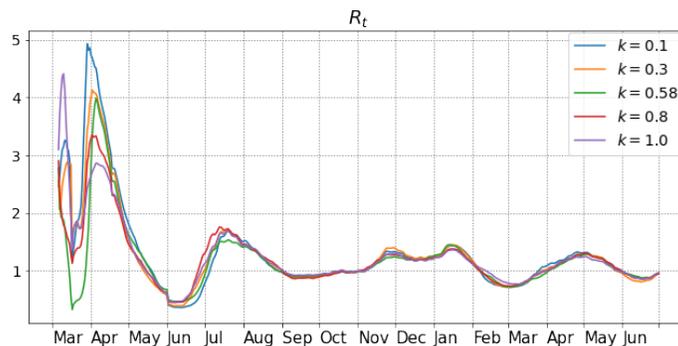


Figure 2.9: Comparison of analysis results with different ratio k between β_a and β_s

As mentioned in Section 2.2.4, the ratio between β_a and β_s is fixed at 0.58 based on the information provided by the literature. Since this ratio is quite uncertain, we performed similar investigations with different ratios varying from 0.1 to 1.0. For $k = \frac{\beta_a}{\beta_s}$ with $k = 0.1, 0.3, 0.58, 0.8$ and 1.0, we run the experiments independently (always with the data from Tokyo) and the analysis mean of R_t are shown in Figure 2.9. The patterns of each R_t are similar, especially after May 31st, 2020. However, at the beginning of the outbreak and during the first state of emergency, all the analysis value of R_t experience quick changes but with different speed. In Figure 2.10 we also provide the analysis mean of the different compartments, since they also depend on the ratio between β_a and β_s . The same conclusion is obtained: except for the first three months, the possible ratio do not generate any noticeable difference.

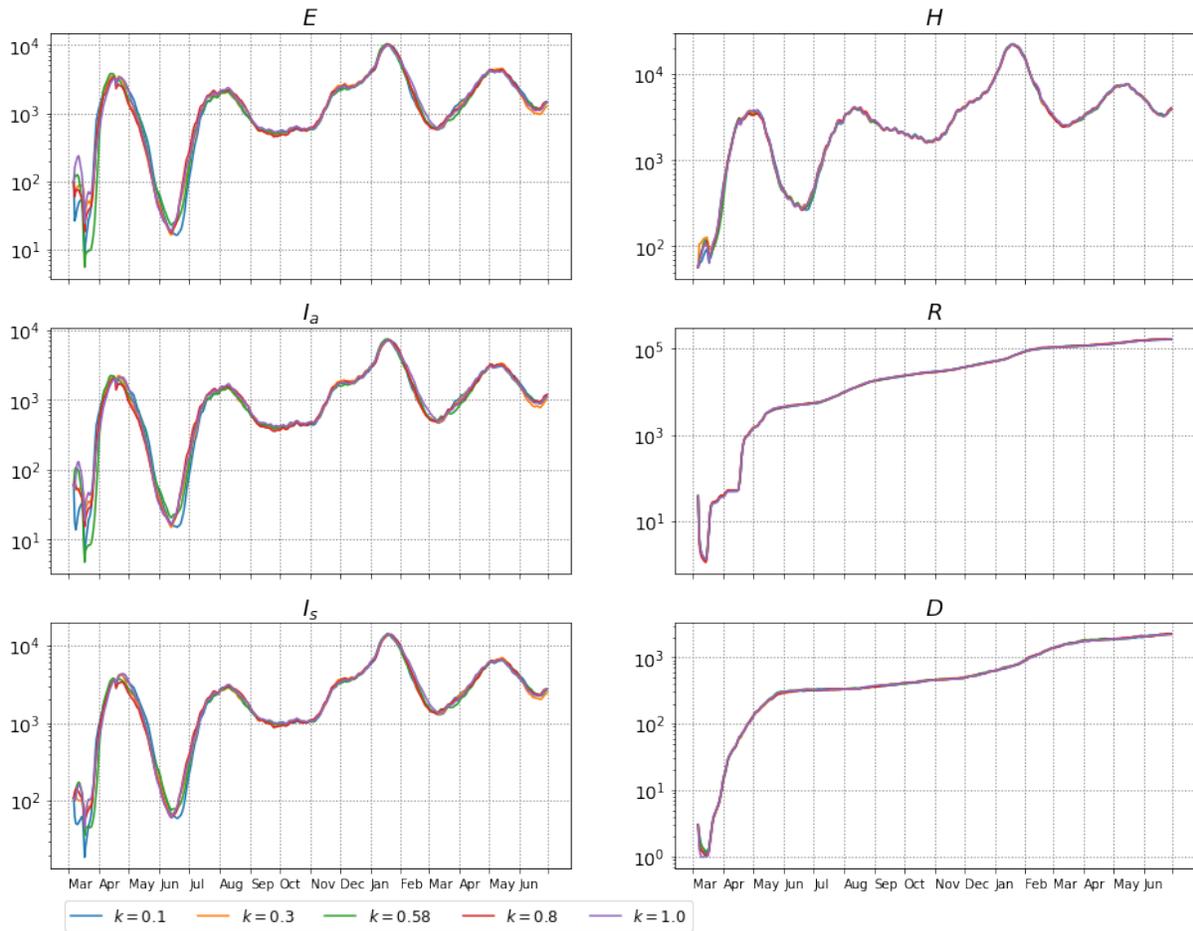


Figure 2.10: Analysis mean of compartments with different ratio k

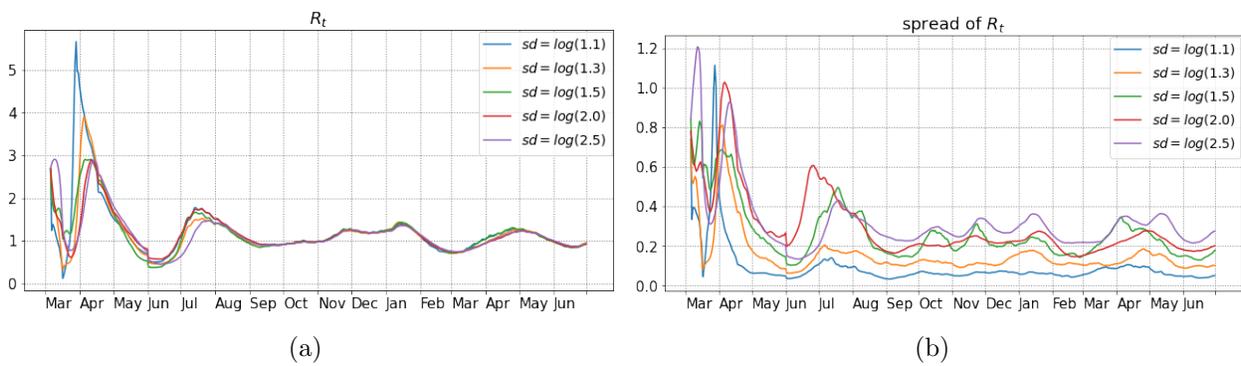


Figure 2.11: Analysis mean and spread of R_t with different observation covariance

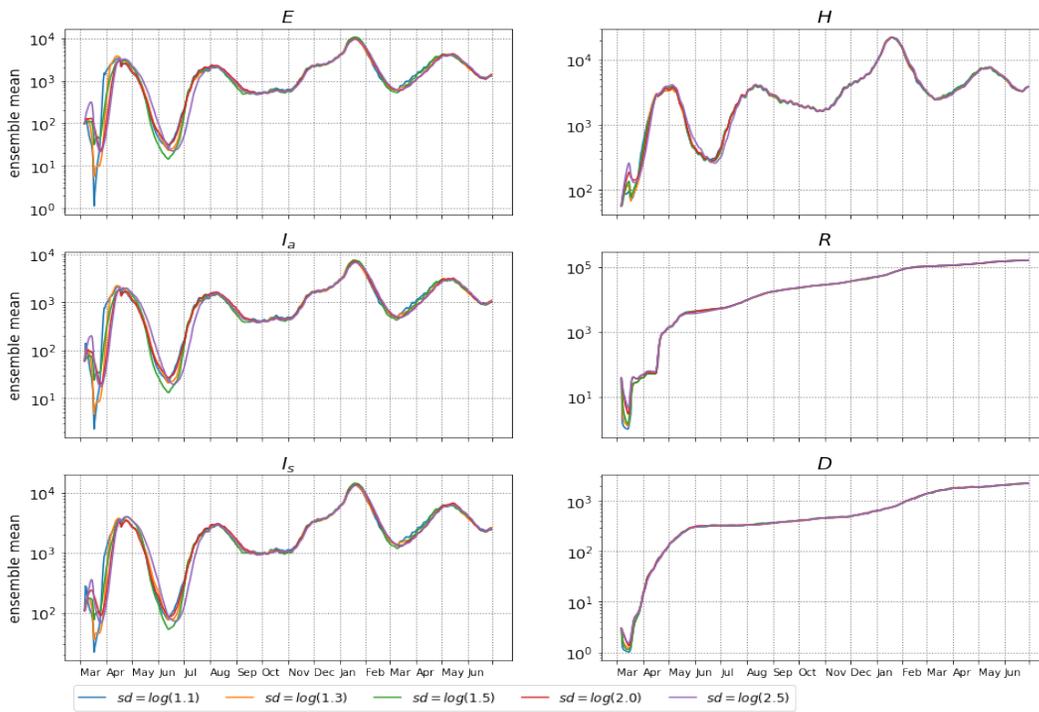
In Section 2.4.1, the error covariance matrix \mathbf{R}_t is defined as a diagonal matrix $((\log(1.3))^2 I$ where I is the identity 3×3 matrix. We also run a sensitivity test for different choices of the covariance matrix, and checked if it has an effect on the analysis results. We choose \mathbf{R}_t equals to $\text{sd}^2 I$ with $\text{sd} = \log(1.1), \log(1.3), \log(1.5), \log(2.0),$ and $\log(2.5)$. For these investigations the ratio between β_a and β_s is fixed as 0.58. In Figure 2.11(a), the analysis means are provided, and they are clearly close to each other except at the beginning stage of the outbreak. The different spreads are shown in Figure 2.11(b), and one notices that the spread increases with the sd , but for sd larger than $\log(1.5)$, the increase is less clear. We also show the analysis mean of compartments E, I_a, I_s, H, R and D in Figure 2.12(a). Again, the analysis results of these compartments with different observation error covariance are quite close to each other. In Figure 2.12(b), the spread is computed by the standard deviation of ensemble members of the \log -transformed compartments. The relations of spread are quite clear for the three compartments with observations, and there are some overlaps for spread of e, i_a and i_s , but in general the spread increases with sd .

As shown in Table 2.1, the proportion of symptomatic and asymptomatic used in our investigation is 83% and 17%, respectively. These values were determined based on sources mentioned in Section 2.2.4. However, since asymptomatic cases are very difficult to detect, and since we can not be fully confident in the above ratio, a sensitivity test is necessary. To do this, we choose two different combinations which are 70% and 30%, and 50% and 50%, and keep the observation error covariance matrix at $((\log(1.3))^2 I$ and $k = 0.58$. Compared to the previous settings, in these two scenarios more people become asymptomatic and recover without showing any symptoms, as illustrated in Figure 2.13(a). On the other hand, the three curves for R_t are close to each others, but the one corresponding to only 50% of symptomatic is slightly higher than the other two, see Figure 2.13(b). In order to understand this, let us recall that the relation between the transmission coefficients β_a and β_s is set to be 0.58, it means that the asymptomatic cases are less infectious than the symptomatic and symptomatic cases. Thus, when more people do not show any symptoms, the transmission coefficient has to be bigger to create enough cases to fit the observations. As a consequence, R_t will also be bigger.

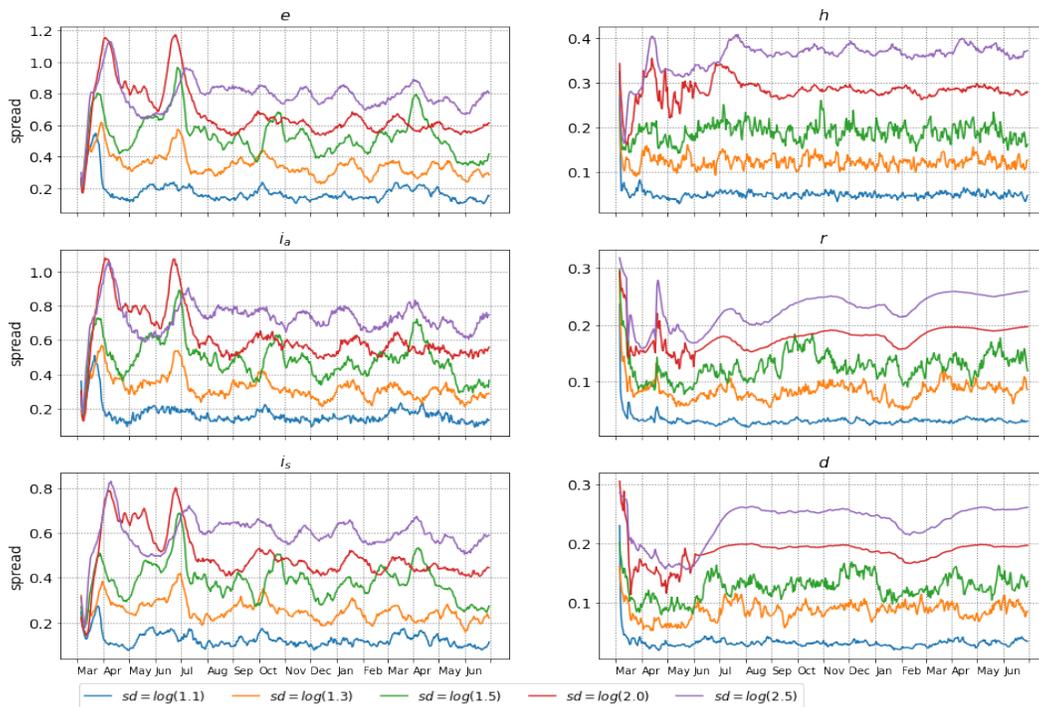
2.5 Discussion, comparisons, and future projects

Let us start by a few easy observations about the results obtained in the previous section.

The analysis values of R_t quickly decrease during the first state of emergency, and this parameter is lower than 1 around the middle of May 2020. This indicates that the disease has started to die out. During the second state of emergency, analysis values of R_t had a much slower decay, but nevertheless in early February 2021, R_t successfully drops below 1. However it starts to increase from March while the state of emergency had not been released. In the third state of emergency, the analysis R_t had the slowest decay, compared to the previous two states of emergency. During the first several weeks of the fourth state of emergency, the effective reproduction number continued increasing, even at a faster speed. It is only after the end of the first week of August that a decay finally took place.



(a)



(b)

Figure 2.12: Analysis mean (a) and spread (b) of the compartments with different observation covariance

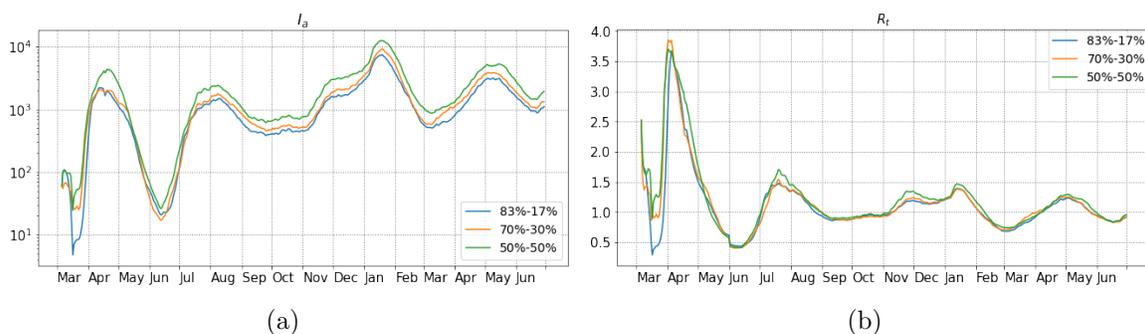


Figure 2.13: Analysis mean of state I_a (a) and R_t (b) with different ratios between symptomatic and asymptomatic

For E , I_a and I_s , one finds that the three values increase at the beginning of the first state of emergency, then they slow down and change the direction to decrease. Similar behaviors can be observed in the second state of emergency, and the values start to increase around the middle of March 2021 while the R_t starts to increase at the beginning of March. There is clearly a delay between the changes of R_t and their effect on the three compartments E , I_a and I_s .

Let us now provide a comparison with two different regions from Japan, namely Osaka and Kanagawa, and also provide the analysis for Japan as a whole. The experiments for Osaka, Kanagawa, and Japan start from March 26th 2020, March 18th 2020, March 1st 2020 respectively. The observed compartments of these regions are the same as those for Tokyo, namely H , D and R . The way to determine the initial setting is the same as that for Tokyo and we use the same setting for the parameters which have been presented in Table 2.1. The daily values for the parameter γ_H and γ_D for each region are also computed by using (2.20) and (2.21). The ratio between β_a and β_s is 0.58, and the observation error covariance matrix is assumed to be $(\log(1.3))^2 I$.

In Figure 2.14 we provide the analysis values of R_t with 68% CI and 95% CI for each region. The additional red curves are again R_t given by Toyokeizai.net. The grey shaded regions mark the respective periods of state of emergency. Though each region declared some states of emergency at different periods of time, they all have the common feature that the recent declarations are less efficient than the first ones. In Figure 2.15, 2.16, and 2.17, we provide the analysis results for compartments E , I_s , I_a , H , R and D for each region.

In Figure 2.18, the three analysis R_t curves for Tokyo, Osaka and Kanagawa are drawn simultaneously. The analysis R_t for Japan is represented by red dots. One observes that the general trends for each region are similar, but some delays are also visible. For example in early April 2020, R_t in Tokyo reached the first peak then started to decline, while R_t in Kanagawa reached the first peak in the middle of April, and then start to decline. Subsequently, these two regions are quite correlated, but Osaka has a slightly independent behavior, both in the summer 2020 and in April 2021.

Based on these observations, one can suspect that the movement of populations might be

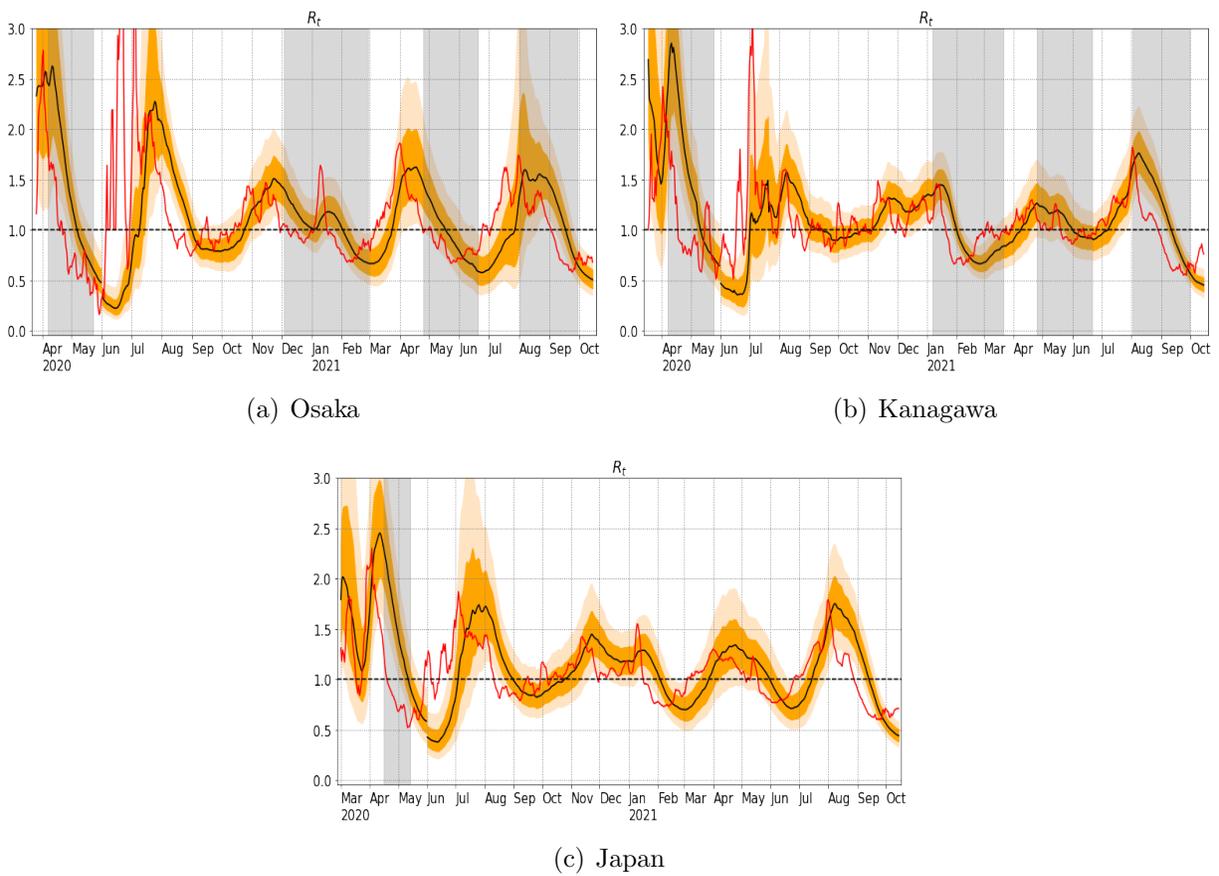


Figure 2.14: Analysis value of R_t for Osaka, Kanagawa and Japan

related to the evolution of R_t for regions which are close to each other. The study of the effect of movements can be carried out by constructing a proper model with multi-populations connected to each others. Some parameters may be tuned, for example for simulating movement restriction's policies. Such simulations could be useful for public health officials or for local governments to understand and estimate different disease-control strategies. We plan to work in this direction in the future.

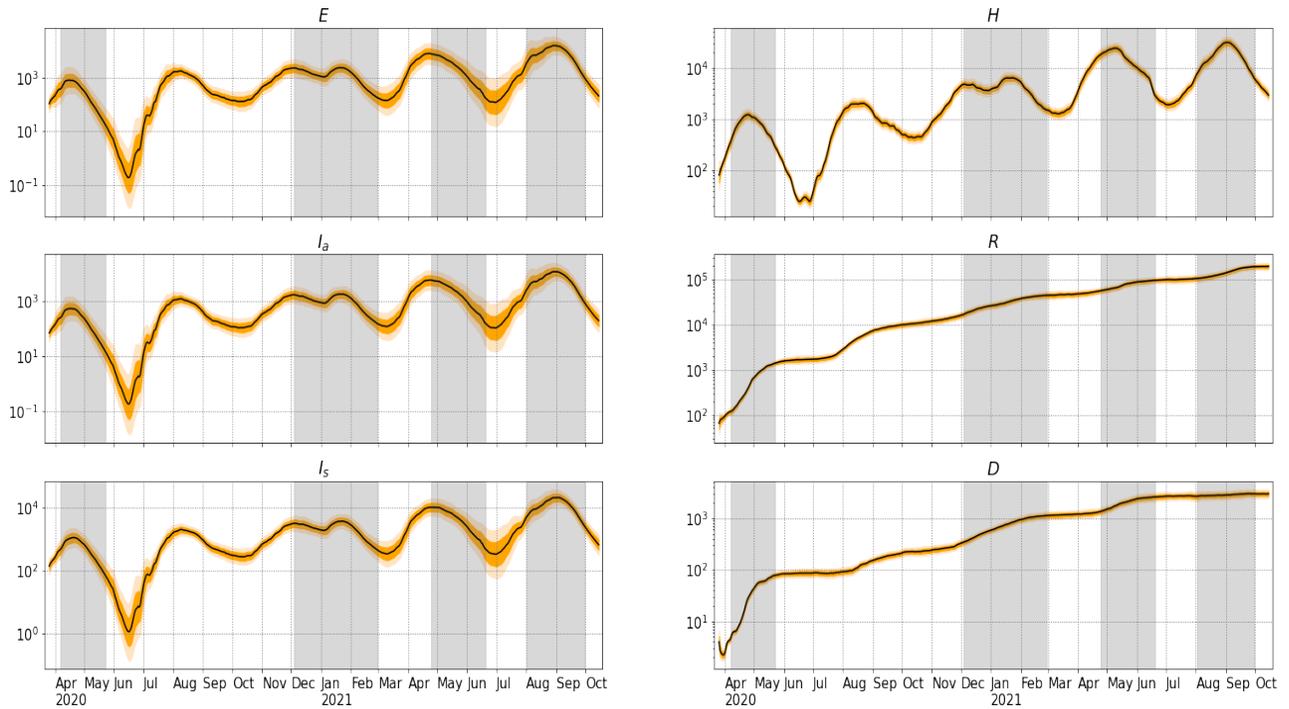


Figure 2.15: Analysis results of Osaka

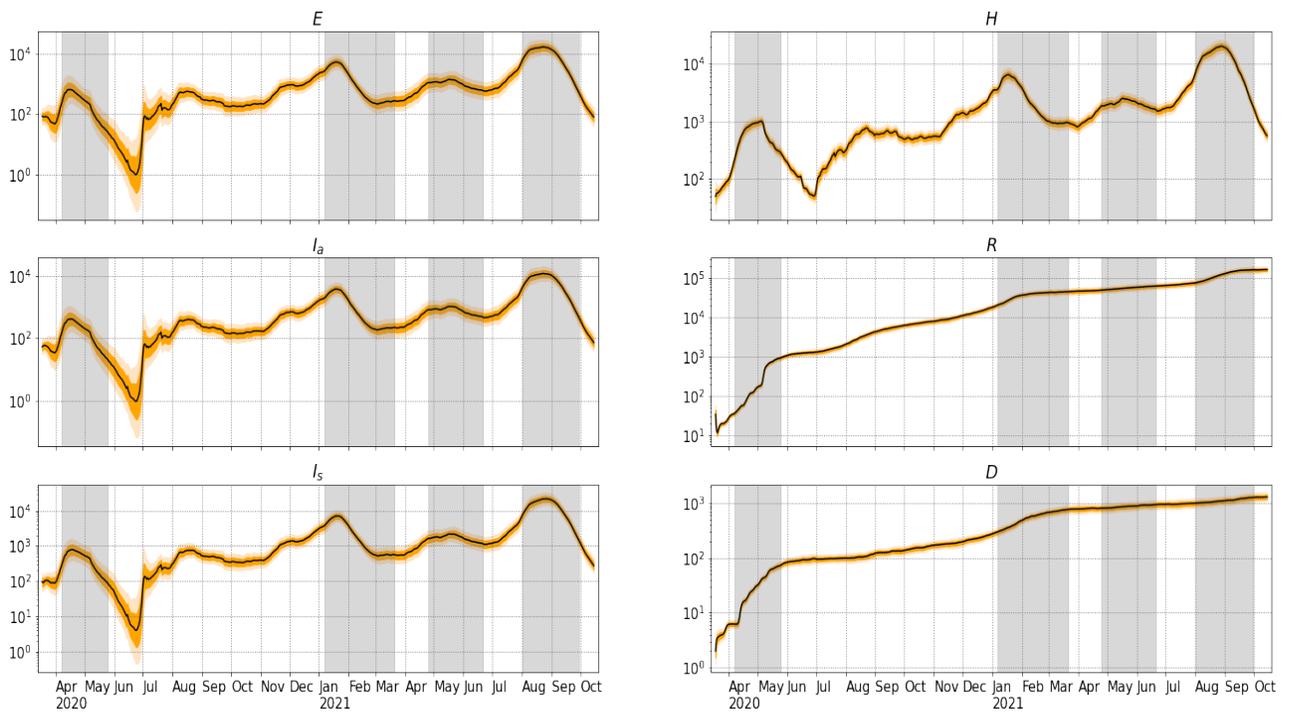


Figure 2.16: Analysis results of Kanagawa

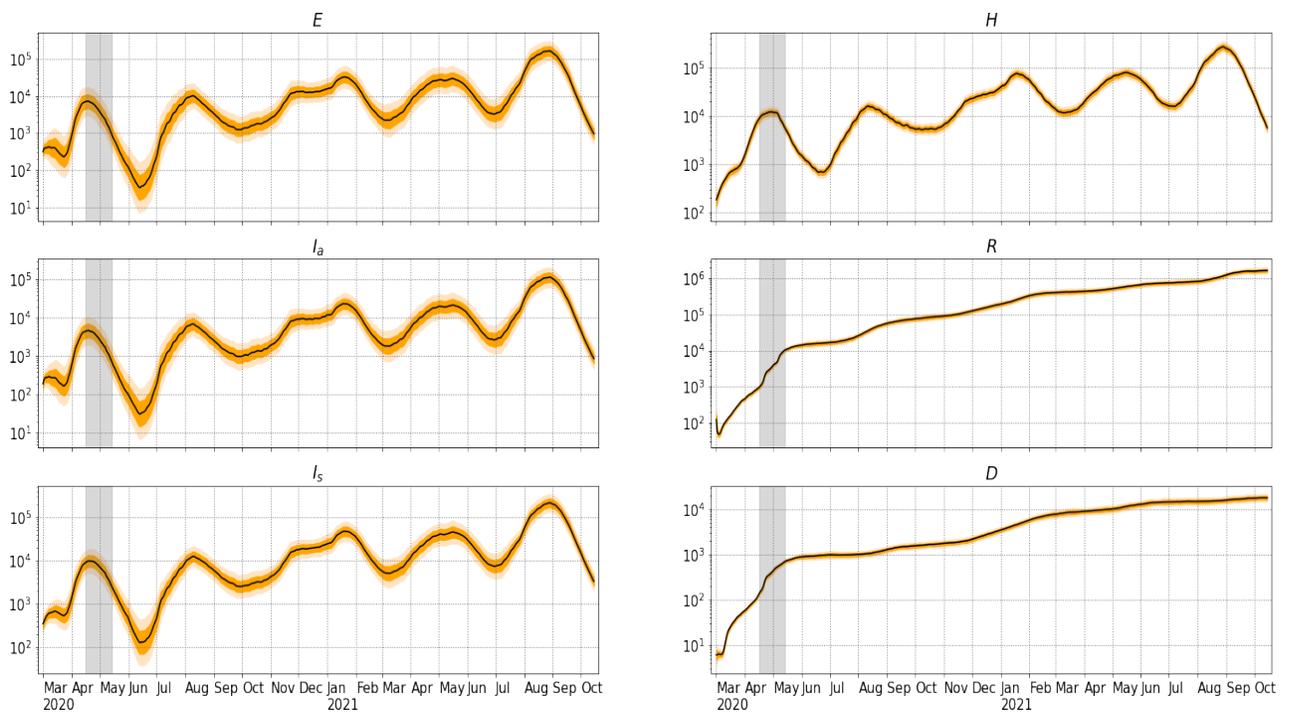
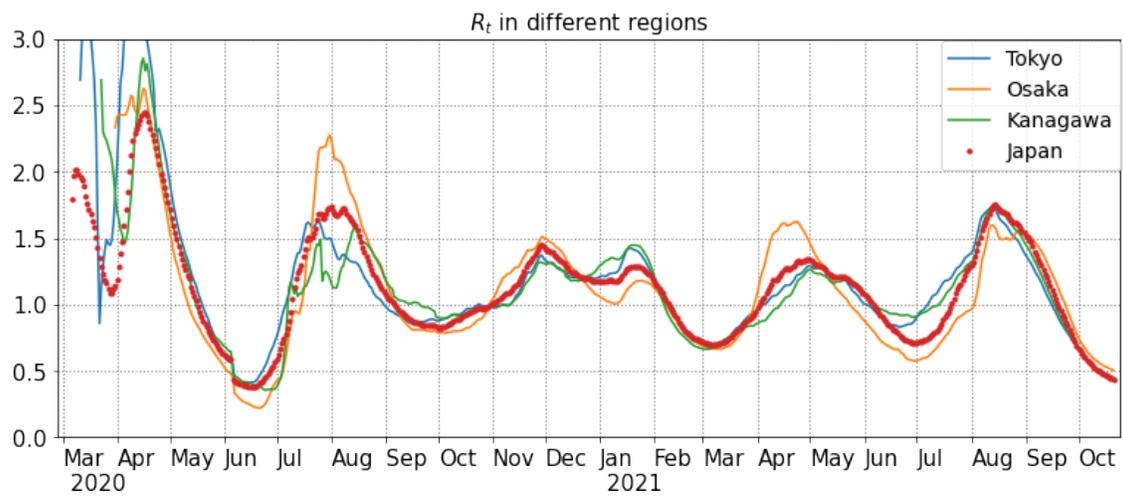


Figure 2.17: Analysis results of Japan

Figure 2.18: R_t for each region

Chapter 3

Control simulation experiment with Lorenz’s butterfly attractor

3.1 Introduction

The “butterfly effect”, discovered by Lorenz in the 1960s [51, 52], is a phenomenon that an infinitesimal perturbation like “a butterfly flapping its wings in Brazil” causes a big consequence like “a tornado in Texas”. This extreme sensitivity brings chaotic behaviors and an intrinsic limit to predictability, but it also allows us to design an effective control which was explored as “the control of chaos” in the 1990s (e.g., a review by [9]). That is, we could take advantage of the “butterfly effect” and design an effective control with a series of infinitesimal interventions leading to a desired future. The control of weather is humans’ long-time desire, and if we know when and where to put a “butterfly”, we could lead a better life by, for example, reducing the risks of tornadoes.

Predictability has been studied extensively, and we enjoy current high-quality weather prediction that is consistently being improved. However, studies on controllability are limited because we had to first improve the prediction accuracy and because our engineering power may be insufficient to enforce large enough perturbations to the atmosphere. Based on recent high-quality numerical weather prediction (NWP), this study attempts to explore a computational simulation approach to weather controllability. The simulation studies reveal what perturbations are needed to modify and control the weather. Mutual interactions between the simulation studies and the intervention techniques would be essential for future developments toward real-world applications.

Previous efforts in weather modification include rain enhancement studies (e.g., a review by [28]) by cloud seeding with ground-based facilities and aircraft injecting smokes and dry ices into moist air, so that the aerosols act as cloud condensation nuclei and enhance cloud formation. These studies greatly helped advance our knowledge about physical processes of clouds and precipitation, but in terms of controlling the weather, we had only limited success with unclear implications for high-impact weather events, mainly because this method works only with supersaturated air. On the climate scale, geoengineering is a widely discussed concept, such as launching mirror satellites to reflect the sunlight and injecting dusts

into the stratosphere to block the sunlight to cool the air. [50] performed computational simulations and explored potential rain enhancements in the Sahel region by implementing large-scale wind and solar farms over the Sahara and modulating the global atmospheric circulation. However, actual geoengineering operations are controversial because they may cause irreversible unexpected side-effects due to our limited knowledge of the Earth system. The accepted and currently ongoing operations to counteract the current climate change may be limited in reducing the greenhouse gas emissions and enhancing renewables and recycles.

Our focus here is different. We aim to apply “the control of chaos” to the weather. We do not aim to cause a permanent irreversible change to nature, but we would like to control the weather within its natural variability and to aid human activities, for example, by shifting the location of an extreme rain region to avoid disasters without causing a side-effect on the global climate. For extreme weather that occurs in a chaotic manner under natural variations, the control of chaos suggests that proper infinitesimal perturbations to the natural atmosphere alter the orbit of the atmospheric dynamics to a desired direction. If the proper infinitesimal perturbations are within our engineering capability, we could apply the control in the real world. However, we cannot be too cautious about potential side-effects and must consider and address every possible consequence. We will come back to this issue later in conclusion.

Here we develop a method of the control simulation experiment (CSE). It would be straightforward to extend the method to broader fields with chaotic dynamics beyond NWP. Weather prediction has been improved consistently by studying predictability and better initial conditions for NWP. Data assimilation (DA) combines the NWP model and observation data for optimal prediction. The method of DA shares that of optimal control, such as the Kalman filter [43], where prediction and control are the two sides of a coin. DA has been studied extensively to improve the prediction, and this study illuminates the control.

The Observing Systems Simulation Experiment (OSSE) is a powerful method to simulate an NWP system (e.g., [5], [38]). The OSSE can be designed to assess the impact of certain observing systems and is useful, for example, for evaluating the potential value of a new satellite sensor before launch. The OSSE can also be designed to evaluate DA methods. In the OSSE, an independent model run acts as a synthetic “nature run” (NR), and we simulate observations by sampling the NR. The NWP system is blind to the NR, takes the simulated observations, and estimates the NR. We compare the estimation accuracy among different OSSEs with different observations and different DA methods.

Here we extend the OSSE and apply small perturbations to the NR to alter the orbit to a desired direction. Investigating effective perturbations would address the controllability. As a proof of concept, we focus on the essence of the problem and use Lorenz’s three-variable model (L63, [51]) instead of using a complex large-scale NWP model. In predictability studies, OSSEs are often performed with such simple idealized models like L63 to explore new DA methods before application to real NWP models (e.g., [45], [97]). L63 is often used to focus on the essence of the problem since L63 shows typical chaotic behaviors, with the solution manifold being a well-known “butterfly attractor” (Fig. 3.1(a)), which has two regimes or wings corresponding to the positive and negative values for variable x . The regime shifts randomly, and the predictability is limited due to chaos. [27] revealed predictability of the regime shift from rapidly growing uncertainties given by the growth rate of specific

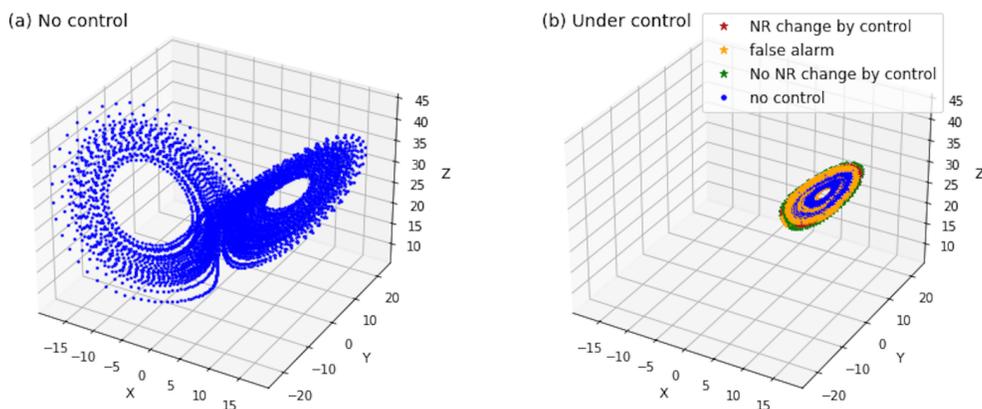


Figure 3.1: Phase space of the three-variable Lorenz model. **(a)** Lorenz’s butterfly attractor from the NR without control; **(b)** the NR under control ($D = 0.05$, $T = \lceil 4T_0 \rceil$). Each dot shows every time step for 8000 steps. See also a movie at <https://doi.org/10.5446/54893> [62].

growing perturbations known as the bred vectors [86].

3.2 Experiments

We first perform a regular OSSE following the previous studies ([45], [97]). The L63 system with the standard choice of the parameters ([51]) is discretized in time by the Runge–Kutta fourth-order scheme with a time step of 0.01 units. We define one step as 0.01 units throughout the paper. We assimilate observations every $T_a = 8$ steps. A round of the orbit, i.e., from a maximum to the next maximum for variable x , corresponds to $T_0 = 75.1$ steps on average. We use the ensemble Kalman filter (EnKF, e.g., [21], [41]) with three ensemble members, which represent equally probable state estimates. For simplicity, we observe all three variables in this study but any subset of observations except for observing only z variable results in the same conclusion, as suggested by the previous study on chaos synchronization [96]. The observation noise is generated from the normal distribution for each variable independently with the variance of 2.0. The EnKF results in an accurate state estimation of the root mean square error (RMSE) of 0.32, consistent with the previous studies.

Next, we extend the OSSE and design a CSE. The goal of the control is to stay in a wing of the butterfly attractor without shifting to the other. It is essential that our prediction and control system is blind to the NR and takes only the imperfect observations. The control system finds when and what perturbations to add to the NR as follows (cf. Fig. 3.2).

1. Perform a DA update using the observations at time t ($t = 0$ in Fig. 3.2).
2. Run an ensemble forecast for T steps from time t to $t + T$ ($T = \lceil 4T_0 \rceil$ in Fig. 3.2, where $\lceil \cdot \rceil$ indicates rounding up to the closest integer since the model integration is discretized).

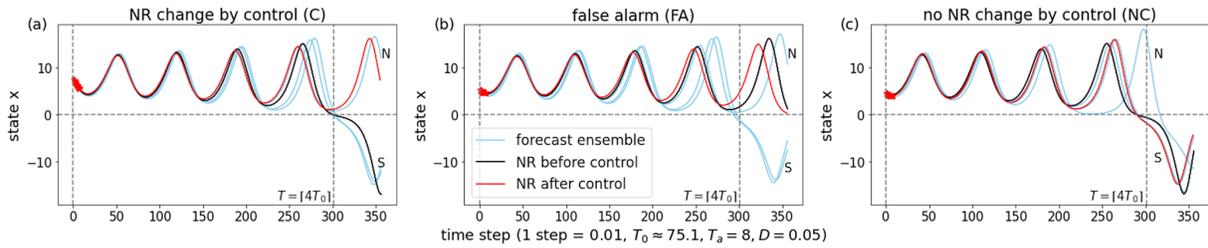


Figure 3.2: Control cases with $T = \lceil 4T_0 \rceil$ and $D = 0.05$ for (a) NR changed (C), (b) false alarm (FA), and (c) NR unchanged (NC). Red ticks at the beginning ($t = 1, \dots, 7$) show addition of perturbations to the NR.

3. If at least one ensemble member shows the regime shift, activate the control (step 4); otherwise, go to step 1 for the next DA at time $t + T_a$.
4. Add perturbations with Euclidean norm D to the NR at every step from $t + 1$ to $t + T_a - 1$. More precisely, at time $t + i$ ($i = 1, \dots, T_a - 1$), the NR state is evolved from the previous NR state at time $t + i - 1$ and is perturbed by adding (dx, dy, dz) , where $\sqrt{dx^2 + dy^2 + dz^2} = D$ (Fig. 3.2 red ticks, indicating perturbations added to the NR with $D = 0.05$).
5. At time $t + T_a$, the new NR is used to simulate the observations; go to step 1 for the next DA at time $t + T_a$.

Step 4 requires perturbations added to the NR. Investigating different strategies to generate the perturbations addresses controllability. Randomly chosen perturbations are found to be ineffective, but instead we find the following strategy effective. We choose an ensemble member “S” showing the regime shift and another ensemble member “N” not showing the regime shift. If all three ensemble members show the regime shift, we use the ensemble members from the former initial times for an extended forecasting period and identify an ensemble member “N” not showing the regime shift during the period from t to $t + T$. Take the differences of the two ensemble members $S - N$ for every step from $t + 1$ to $t + T_a - 1$ (1 to 7 in Fig. 3.2) before the next observations are available at $t + T_a$ (8 in Fig. 3.2). The differences are used as perturbations added to the NR at appropriate time steps. Here, we consider the limitation of our intervention and include only a subset of the three variables (x, y, z) with a limited perturbation size. The choice of the variables and norm D are the parameters for intervention.

Figure 3.2 illustrates three different cases with perturbations added to all three variables (x, y, z) with $D = 0.05$ and $T = \lceil 4T_0 \rceil$. With these settings the control is successful, as shown in Fig. 3.1(b) for 8000 steps. Figure 3.2(a) shows the case in which the NR is changed by the control and stays in the positive- x regime successfully. Figure 3.2(b) shows the case of a false alarm (FA), in which the NR does not show the regime shift but the ensemble prediction does. Therefore, the perturbations are added unnecessarily but do not hurt. Figure 3.2(c) shows the case in which the NR is not changed by control and still shows the regime shift (simply “NC” for no change).

(a) control (x, y, z)									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0	0	0	0.025
[2T ₀]	0	0	0	0	0	0.025	0	0	0
[2.5T ₀]	0	0	0	0.025	0.250	0.275	0.250	0.200	0.175
[3T ₀]	0	0.050	0.425	0.550	0.800	0.675	0.400	0.100	0.200
[4T ₀]	0.050	0.975	0.925	0.975	0.825	0.825	0.725	0.500	0.525

	≥0.8	≥0.6	≥0.4	≥0.2	>0	0

(b) control (x, y)									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0	0.025	0.075	0.050
[2T ₀]	0	0	0	0	0	0	0	0	0
[2.5T ₀]	0	0	0	0	0.275	0.350	0.400	0.375	0.375
[3T ₀]	0	0	0.050	0.200	0.500	0.375	0.275	0.100	0
[4T ₀]	0	0.275	0.900	0.975	0.900	0.825	0.700	0.550	0.575

(c) control (x, z)									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0.025	0	0	0
[2T ₀]	0	0	0	0	0	0	0	0	0
[2.5T ₀]	0	0	0	0	0.100	0.200	0.050	0.075	0.100
[3T ₀]	0	0	0	0.200	0.600	0.800	0.650	0.350	0.225
[4T ₀]	0	0.175	0.900	0.950	0.775	0.775	0.600	0.625	0.550

(d) control (y, z)									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0	0	0	0.025
[2T ₀]	0	0	0	0	0.025	0.050	0	0	0.025
[2.5T ₀]	0	0	0	0.025	0.250	0.125	0.175	0.225	0.075
[3T ₀]	0	0.025	0.250	0.650	0.800	0.675	0.375	0.175	0.175
[4T ₀]	0	0.850	0.950	0.950	0.850	0.750	0.625	0.550	0.550

(e) control x									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0	0	0	0
[2T ₀]	0	0	0	0	0	0	0	0	0
[2.5T ₀]	0	0	0	0	0	0.450	0.725	0.525	0.225
[3T ₀]	0	0	0	0	0	0.425	0.425	0.350	0.175
[4T ₀]	0	0	0	0	0.825	0.975	0.875	0.850	0.525

(f) control y									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0	0.025	0.025	0
[2T ₀]	0	0	0	0	0	0	0	0	0
[2.5T ₀]	0	0	0	0	0.050	0.225	0.375	0.175	0.250
[3T ₀]	0	0	0	0.100	0.350	0.225	0.150	0.050	0.025
[4T ₀]	0	0.050	0.350	0.875	0.900	0.775	0.600	0.500	0.400

(g) control z									
T \ D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5
[1.5T ₀]	0	0	0	0	0	0	0	0	0
[2T ₀]	0	0	0	0	0	0	0	0	0
[2.5T ₀]	0	0	0	0	0	0.025	0	0	0
[3T ₀]	0	0	0	0	0.45	0.45	0.275	0.35	0.275
[4T ₀]	0	0	0.225	0.825	0.875	0.6	0.55	0.325	0.425

Figure 3.3: Rates of successful control out of 40 CSEs for perturbations added to variables (a) x, y, z , (b) x, y , (c) x, z , (d) y, z , (e) x , (f) y , and (g) z .

To investigate the sensitivity to the parameters T and D and the choice of the perturbed variables, we perform 40 independent experiments for each setting for 8000 steps (1000 DA cycles; cf. Appendix 3.A for the exact choices of the initial conditions) and count the number of successful experiments in which the NR stays in a single regime under control. Higher success rates correspond to better controllability. With longer forecasts (larger T), control is generally more effective (Fig. 3.3). With small $T < [2.5T_0]$, the success rates are very low. The mean transition time for the regime shift is approximately $2.3T_0$, which may be the minimum forecast length for effective control. With very small perturbations ($D = 0.02$), the control is difficult, but a larger D does not necessarily improve the success rate. The perturbations are added every step, and the state evolves by approximately 0.5 (Euclidean norm) in one step on average (Table 3.1). This is about half of the evolution without control, suggesting that the perturbations effectively drag the NR states toward more stable regions of the attractor (cf. Fig. 3.1 and a movie at <https://doi.org/10.5446/54893> [62]). Adding larger perturbations with a similar size to the one-step model evolution tends to reduce the effect of control. Although observing only z is not sufficient for DA, it is good for control. Perturbing only one variable y or z is effective with $T = [4T_0]$ and $D > 0.04$, only an eighth of the analysis error of 0.32 or only 3% of the observation error standard deviation of $\sqrt{2}$. In short, the L63 regime change is considerably controllable.

We further investigate the rates of FAs and NR changed (C) and unchanged (NC) by perturbations (Fig. 3.4(a)). With larger D , we find generally fewer interventions. With smaller D , we have more interventions mostly by FA. With smaller D , higher rates of NC suggest that longer-term small interventions are needed. Additional experiments by not applying FA and/or NC perturbations reveal the relative importance of these perturbations (Fig. 3.4(b)). These experiments require knowing the NR T steps in advance and therefore are not practical but are useful for understanding the roles of these perturbations. For $D = 0.2$ and smaller, not applying FA perturbations does not significantly contribute to the control (Fig. 3.4(b), yellow), whereas not applying NC has a significant impact on reducing the effect

D	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4	0.5	No control
OME	0.694	0.608	0.594	0.577	0.536	0.488	0.461	0.422	0.403	0.956
D/OME	0.029	0.049	0.067	0.087	0.186	0.410	0.651	0.947	1.239	NA

Table 3.1: Averaged one-step model evolution in the Euclidean norm (OME) and the relative size of perturbations (D/OME). Only successful control cases are considered for CSEs with $T = \lceil 4T_0 \rceil$ and perturbations added to variables x , y , and z . “NA” indicates not available.

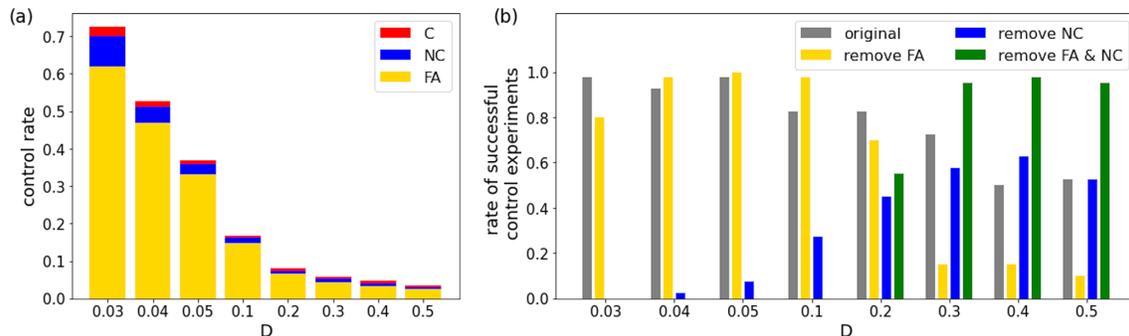


Figure 3.4: **(a)** Rates of the cases of C (red), FA (yellow), and NC (blue) for the successful control experiments with $T = \lceil 4T_0 \rceil$ and perturbations added to variables x , y , and z . The rates indicate the number of cases out of a total of 1000 DA cycles. **(b)** Rates of successful control experiments with $T = \lceil 4T_0 \rceil$ and perturbations added to variables x , y , and z for the original CSE (grey; cf. Fig. 3.3), the CSE without applying FA perturbations (yellow), the CSE without applying NC perturbations (blue), and the CSE without applying FA and NC perturbations (green).

of control (Fig. 3.4(b), blue, green). That is, the accumulation of NC perturbations would be essential for effective control. With large $D > 0.2$, not applying FA and NC perturbations significantly enhances the effect of control (Fig. 3.4(b), green). With the perturbation size similar to or even larger than the one-step model evolution (Table 3.1), a single instance of C perturbations is quite significant. In these cases, FA and NC perturbations are found to be harmful.

Finally, we perform additional sensitivity experiments with a longer DA interval of $T_a = 25$ steps and with partial observations; i.e., only one or two variables are observed. The results generally agree with what has been shown so far (cf. Appendix 3.B).

3.3 Conclusions

In this study, we proposed the CSE with numerical demonstration using the L63 three-variable model. The OSSE is a well-known, powerful approach to study predictability and to evaluate DA methods and observing systems without having real-world observation data. The CSE is an extension to the OSSE to study controllability and can be applied to various dynamical systems including full-scale NWP models. Our future studies apply the CSE

to more complex models and investigate different control scenarios such as controlling the occurrences of extreme events. Such studies will address critical issues like how manageable interventions in terms of cost and energy can make differences to extreme events. This study is only a small step toward broad investigations that may lead to effective control of weather events.

As we described in the introduction, any real-world application requires extensive caution. For the case of the L63 model, one side of the attractor may not be desirable for all aspects. We must consider and assess every potential impact caused by the control and have proper protocols for social, ethical, and legal agreement about real-world operations.

3.A Appendix: The initial conditions of 40 CSEs

The OSSE with the L63 model follows that of the previous studies [45, 97, 60, 22]. Here we describe the additional details that were not provided in the previous papers but that are necessary to repeat the experiment in this study. The initial condition for the NR was chosen to be $(x, y, z) = (8.20747939, 10.0860429, 23.86324441)$ after running the L63 model for 1000 steps initialized by the three state variables taken from independent random draws from a normal distribution with mean 0 and variance 2.0. The NR was 8 million steps long, and the OSSE was performed for the same period as the NR.

The CSEs were performed for a total of 378 combinations of T , D , and the choice of intervention. There were nine, five, and seven choices of T , D , and intervention, as shown in Fig. 3.3. For each combination, 40 independent CSEs were performed for 8000 steps. The initial conditions for the 40 CSEs were chosen from the analyzed states of the OSSE at different time points as shown in Table 3.2. Figure 3.1(b) shows CSE no. 1 and Fig. 3.1(a) the corresponding period of the NR.

3.B Appendix: Additional sensitivity experiments

CSEs are performed with a longer DA interval of $T_a = 25$ steps, which results in a RMSE of 0.76, consistent with the previous studies. The results are generally consistent (Fig. 3.5 compared with Fig. 3.3).

CSEs are performed with different observing coverages, and the results are summarized in Table 3.3. Multiplicative inflation is manually tuned for each observing coverage.

CSE index	Time point of the NR	CSE index	Time point of the NR
1	106 069	21	150 056
2	107 043	22	150 796
3	109 371	23	152 308
4	111 261	24	155 048
5	112 987	25	155 666
6	114 146	26	162 753
7	122 065	27	164 411
8	124 720	28	168 461
9	125 339	29	172 109
10	125 854	30	173 399
11	126 902	31	173 894
12	128 058	32	175 011
13	130 718	33	179 671
14	132 342	34	184 480
15	133 311	35	197 270
16	138 699	36	199 278
17	140 562	37	200 712
18	144 953	38	201 304
19	147 614	39	208 511
20	149 418	40	209 397

Initial time point = time point of the NR $- T - [(time\ point\ of\ the\ NR - T) \bmod T_a]$.

Table 3.2: Time points of the NR providing the initial conditions of the 40 independent CSEs. The initial time point coincides with the time when observations are available, i.e., only every $T_a = 8$ steps, and the formula underneath the table provides the exact initial time point from the value in the table for a given parameter of T .

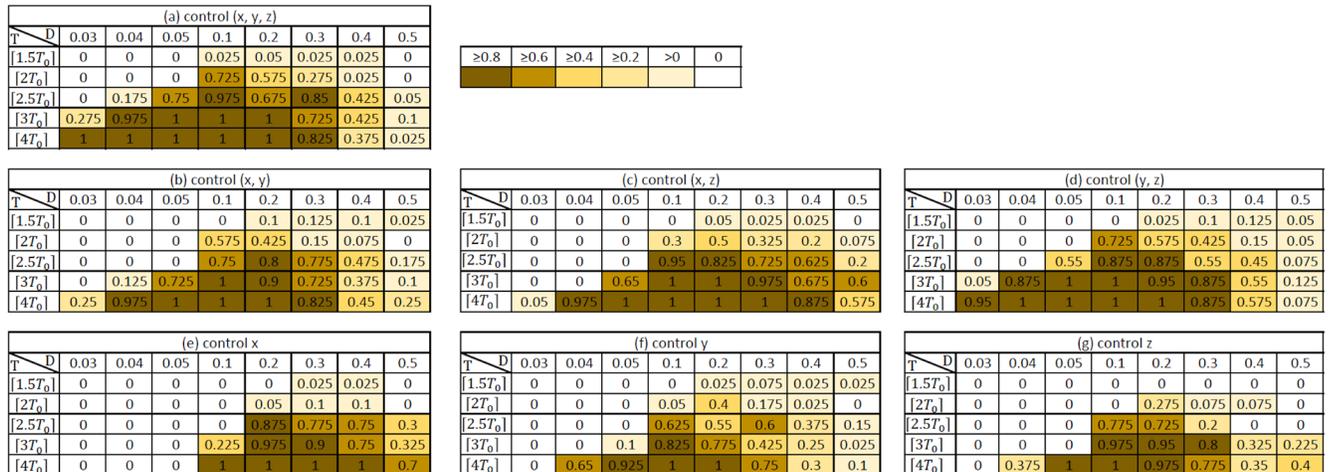


Figure 3.5: Similar to Fig. 3.3 but for the case with a longer DA interval of $T_a = 25$.

D	Obs x	Obs y	Obs x, y	Obs x, z	Obs y, z	Obs x, y, z
0.02	0	0.025	0.05	0.125	0	0.05
0.03	1	0.95	0.95	0.975	0.975	0.975
0.04	1	0.975	0.95	1	1	0.925
0.05	1	1	0.975	1	1	0.975
0.1	1	1	1	1	1	0.825
0.2	0.975	0.925	0.85	0.975	0.975	0.825
0.3	0.95	0.925	0.675	0.975	0.95	0.725
0.4	0.95	0.8	0.78	0.975	0.875	0.5
0.5	0.9	0.75	0.65	0.95	0.85	0.525
Ensemble spread	0.807	0.469	0.376	0.477	0.323	0.27
RMSE	0.908	0.507	0.412	0.564	0.356	0.32
Multiplicative inflation	1.065	1.05	1.045	1.09	1.06	1.04

Table 3.3: Rates of successful control out of 40 CSEs with different observing coverage. $T = \lceil 4T_0 \rceil$ and perturbations are added to variables x , y , and z .

Chapter 4

Control simulation experiments of extreme events with the Lorenz-96 model

4.1 Introduction

In the recent study [63] of the first two authors, it has been demonstrated that a Control Simulation Experiment (CSE) leads to a successful control of the chaotic behavior of the famous Lorenz-63 model [51]. Namely, by applying small perturbations to its nature run, the evolution of the system can be confined to one of the two wings of the attractor. The method is based on an extension of the Observing Systems Simulation Experiment (OSSE) from predictability to controllability. This work can also be seen as an application of the control of chaos [9] to the framework of Numerical Weather Prediction (NWP).

The control of weather is certainly one of the oldest dreams of human being, and several earlier works have already investigated different approaches of weather control based on numerical simulations. As for some of the pioneering works, let us mention [36] and [37] in which R. Hoffman discusses a general scheme and some applications for the control of hurricanes. A case study is then presented in [34], where the temperature increments needed to limit the wind damage caused by Hurricane Andrew in 1992 is studied with a four-dimensional variational data assimilation technique (4D-Var). Such investigations have recently been further extended in [91] for the control of cyclones, with some numerical simulation experiments based on the Typhoon Mitag of 2019. Similarly, in [58] the effect of small control inputs are evaluated by using the Cane-Zebiak 33 000-state model of the El-Niño/Southern Oscillation (ENSO), and the outcome is a (theoretical) significant reduction of the ENSO amplitude with small control inputs. As a final example, let us mention that weather and climate modifications can also be explored as a problem of optimal control, see for example [79].

In the present paper, the CSE proposed in [63] is further developed and applied to the Lorenz-96 model with 40 variables [53]. In this setting, a notion of extreme events is defined first, based on the observation of 100 years of the nature run. Next, the aim of the CSE is to

drive any new nature run away from the extreme events, by a suitable application of small perturbations to the nature run whenever such an event is forecasted. By defining a success rate for such actions, it is possible to assess the performance of the CSE as a function of the different parameters involved in the process.

The underlying motivation for such investigations is quite clear and can be summarized in a single question: Can one control to avoid weather extremes within its intrinsic variability? Indeed, the sensitivity to initial conditions of the weather system advocates the implementation of suitable small perturbations for dragging the system away from any catastrophic event. With such an approach, heavy geoengineering actions become automatically obsolete. However, ahead of any such implementations, numerical demonstrations have to demonstrate the applicability and the effectiveness of such actions, and ethical, legal, and social issues have to be thoroughly discussed. With this ambitious program in mind, this paper provides a precise description of the CSE applied to the Lorenz-96 model, and some sensitivity tests with the CSE parameter settings. Clearly, this corresponds to a modest but necessary step before dealing with more realistic NWP models.

Let us now describe more concretely the CSE and the content of this paper. In Section 4.2 we briefly introduce the Lorenz-96 model (L96) with 40 variables, recall the main ideas of the Local Ensemble Transform Kalman Filter (LETKF), and describe the implementation of the LETKF with 10 ensemble members for our investigations on L96. By creating synthetic observations with the addition of uncorrelated random Gaussian noise to the nature run, we then test the performances of our system, and get the analysis and forecast RMSEs for different forecast lengths. Multiplicative inflation and observation localization are also briefly introduced in this section as essential techniques to stabilize the LETKF.

In Section 4.3, we firstly introduce the definition of extreme events for L96, and then provide the precise description of the CSE. It essentially consists in simulating the evolution of the system for a certain forecast length, by integrating its evolution with a forth-order Runge-Kutta scheme, and by applying a suitable perturbation if an extreme event is forecast. The LETKF is then used for assimilating the noisy observation to the forecast value in order to get the analysis value for the next integration step. Definitions for a success rate and for a perturbation energy (energy delivered to the system through the perturbations) are also introduced in this section, and a discussion of these quantities as functions of the forecast length and of the magnitude of the perturbation vectors is provided. Based on a time scale commonly accepted for this model, our CSEs are taking place for a period of 100 years, the noisy observations are available every 6 hours, and the forecast length is chosen between 2 and 14 days.

Sensitivity tests are finally performed in Section 4.4. Namely, the performance of the method is evaluated when perturbation vectors are applied at fewer places, or when the observations are collected less frequently. In these settings, the success rate and the perturbation energy are again discussed as a function of the parameters. The reason for performing these tests is quite clear: In the framework of a more realistic numerical weather prediction model, this would correspond to the restricted ability of applying a global perturbation (even of a small magnitude), or to lack of regular measurements of the weather system.

These investigations are only the second step towards the development of a applicable

CSE, but the outcomes are already promising: For various settings a success rate close to 1 is obtained. Such a score depends on the forecast length, on the perturbation energy, on the number of sites on which the perturbations are applied, and on the regularity of the observations. The next step will be to consider an improved model, and to get a more quantitative information about the exact amount of energy introduced in the system, and a precise evaluation of the false alarms. Indeed, the ultimate goal of such investigations is to keep a high success rate, but to minimize the perturbation actions, in magnitude and in frequency. We plan to work on these issues in the future.

4.2 Background information

4.2.1 The Lorenz-96 model

Lorenz-96 model was first introduced in [53], and is expressed by a system of K differential equations of the form:

$$\frac{dX_k}{dt} = (X_{k+1} - X_{k-2})X_{k-1} - X_k + F \quad (4.1)$$

for $k \in \{1, 2, \dots, K\}$ and $K \geq 4$. Note that $F > 0$ is independent of k . It is also assumed that $X_0 = X_K$, $X_{-1} = X_{K-1}$, and $X_{K+1} = X_1$, which ensure that the system (4.1) is well-defined for each k . These conditions endow the system with a cyclic structure of the real *variables* X_k . Originally, these variables represented some meteorological quantities evenly-distributed on a latitude circle. In this framework, the constant term F simulates an external forcing, while the linear terms and quadratic terms simulate internal dissipation and advection respectively. A simple computation performed in [54, Sec. 2] shows that the average value over all variables and over a long enough time lies in $[0, F]$, and that the standard deviation lies in $[0, F/2]$.

As already observed in [53, Sec. 2], when F is small enough, all solutions decay to the steady solution defined by $X_k = F$ for $k = 1, 2, \dots, K$. Then, as F increases, the solutions show a periodic behavior, and end up with a chaotic behavior for larger F . For the following discussion, we choose $K = 40$ and $F = 8$, which is identical to the setting used in [54]. This setting ensures the chaotic behavior of the model. By studying the error growth, namely the difference between two solutions with slightly different initial values, it has been found that the leading Lyapunov exponent corresponds to a doubling time of 0.42 unit of time [54, Sec. 3]. By comparing this model with other up-to-date global circulation models, it is commonly accepted that 1 unit of time in this model is equal to 5 days in reality. As a consequence, the doubling time is about 2.1 days. Note that the doubling time may decrease when F increases further.

In [54, Sec. 3], some numerical investigations are performed on this model. In particular, the evolution of one variable is reported in Figure 4 of this reference over a long period of time, and this evolution is compared with the evolution of the same variable when a small perturbation is added at the initial time. For one month, the two trajectories are not distinguishable, while 2 months after the perturbation the two trajectories look completely independent. In the sequel, we shall consider similar perturbations of the system, but applied

to more than one variable. More precisely, since the evolution is taking place in \mathbb{R}^{40} , we shall consider perturbation vectors in this space, of different directions and of different sizes (magnitudes). By applying suitable perturbation vectors, our aim is to drive the evolution of the system in a prescribed direction within a limited time interval.

4.2.2 Local Ensemble Transform Kalman Filter

The *Local Ensemble Transform Kalman Filter* (LETKF) is a type of ensemble Kalman filter which was first introduced in [42]. It corresponds to a data assimilation method suitable for large, spatiotemporally chaotic systems. For $i \in \{1, \dots, N\}$, let $\mathbf{x}^{b(i)}$ denote a m -dimensional state vector, at a given fixed time. The exponent b stands for *background*, or *forecast*. One aim of LETKF is to transform the ensemble $\{\mathbf{x}^{b(i)}\}_{i=1}^N$ into an *analysis* ensemble $\{\mathbf{x}^{a(i)}\}_{i=1}^N$ such that its analysis mean $\bar{\mathbf{x}}^a$ minimizes the cost function:

$$J(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \bar{\mathbf{x}}^b) + [\mathbf{y} - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x})]. \quad (4.2)$$

In this expression, \mathbf{P}^b denotes the background covariance matrix defined by $(N-1)^{-1} \mathbf{X}^b (\mathbf{X}^b)^T$, where \mathbf{X}^b is the $m \times N$ background ensemble perturbation matrix whose i^{th} column is given by $\mathbf{x}^{b(i)} - \bar{\mathbf{x}}^b$. \mathbf{R} stands for the observation error covariance matrix and H corresponds to the observation operator. For LETKF, it is assumed that the number of ensemble members N is smaller than the dimension m of the state vector, and also smaller than the number of observations (which corresponds to the dimension of \mathbf{y}). In [42], it is shown that the minimization can be performed in the subspace S spanned by the vectors $\mathbf{x}^{b(i)}$, which is efficient in terms of reduced dimensionality.

To perform the analysis, the matrix \mathbf{X}^b is used as a linear transformation from some N -dimensional space S' onto S . Thus, assume that $\mathbf{w} \in S'$, then $\mathbf{X}^b \mathbf{w} \in S$ and $\bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w} \in S$ are vectors in the model variables. Similar to (4.2), a cost function J' for \mathbf{w} is defined by

$$J'(\mathbf{w}) = (k-1) \mathbf{w}^T \mathbf{w} + [\mathbf{y} - H(\bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w})]^T \mathbf{R}^{-1} [\mathbf{y} - H(\bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w})]. \quad (4.3)$$

In [42, Sec. 2.3], it is shown that if $\bar{\mathbf{w}}^a \in S'$ minimizes J' , then the vector $\bar{\mathbf{x}}^a := \bar{\mathbf{x}}^b + \mathbf{X}^b \bar{\mathbf{w}}^a$ minimizes J .

Let us now define the vectors $\mathbf{y}^{b(i)} := H(\mathbf{x}^{b(i)})$ and the corresponding matrix \mathbf{Y}^b whose i^{th} column is given by $\mathbf{y}^{b(i)} - \bar{\mathbf{y}}^b$. The linear approximation of $H(\bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w})$ is then provided by the expression $\bar{\mathbf{y}}^b + \mathbf{Y}^b \mathbf{w}$. By inserting this approximation into (4.3), one obtains an expression having the form of the Kalman filter cost function, and by standard computations, the solution $\bar{\mathbf{w}}^a$ and its analysis covariance matrix $\tilde{\mathbf{P}}^a$ can be inferred. By transforming this solution back to the model space, one gets

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b \bar{\mathbf{w}}^a, \quad (4.4)$$

$$\mathbf{P}^a = \mathbf{X}^b \tilde{\mathbf{P}}^a (\mathbf{X}^b)^T. \quad (4.5)$$

Finally, a suitable analysis ensemble perturbation matrix \mathbf{X}^a can be defined by $\mathbf{X}^a := \mathbf{X}^b \mathbf{W}^a$, where $\mathbf{W}^a := [(\tilde{\mathbf{P}}^a)^{\frac{1}{2}}]^T$. Here, the power $\frac{1}{2}$ represents the symmetric square root of the

matrix. One then easily checks that the sum of columns of \mathbf{X}^a is zero, and that the analysis ensemble $\mathbf{x}^{a(i)}$ (obtained by the sum of the i^{th} column of \mathbf{X}^a and $\bar{\mathbf{x}}^a$) agrees with the mean and covariance matrix given by (4.4) and (4.5).

4.2.3 Application of LETKF to the Lorenz-96 model

From now on, we consider the Lorenz-96 model with 40 variables and $F = 8$. The vector generated by the 40 variables will be used as the state vector \mathbf{x} for the LETKF, and will be referred to as the state variables. To study the predictability and analysis accuracy of LETKF, we run the model by using a forth-order Runge-Kutta scheme. The initial value is provided in the Appendix, and the integration time step is 0.01 unit of time. As mentioned in Section 4.2.1, 1 unit of time is assumed to equal to 5 days in reality. We let the model evolve for 110 years (8030 units of time) and disregard the first 10 years' results to avoid any transient effect. Note that for all subsequent experiments, the time 0 is set after these disregarded 10 years. For 100 years of nature run, we generate a synthetic observation every 6 hours (0.05 unit of time) by adding independently a $Normal(0, 1)$ distributed random number to the value of each variable. For the choice of the amplitude of these perturbations, we recall that for the Lorenz-96 model with $F = 8$, the average value over all variables and over a long enough time lies in $[0, 8]$, and the standard deviation lies in $[0, 4]$.

To proceed with data assimilation, we consider 10 ensemble members. For the initial data, we add to the nature run at time 0 a $Normal(0, 1)$ distributed random number independently to each variable. We then run the forecast-observe-analyze cycle every 6 hours over 100 years with LETKF and with the synthetic observations. To avoid underestimating the error variance, we implement a *multiplicative inflation* with $\rho = 1.06$, namely we consider an inflated background error covariance matrix $\rho\mathbf{P}^b$ instead of \mathbf{P}^b , see [41, Sec. 4]. Also, with the current parameter's setting of the Lorenz-96 model, the attractor has 13 positive Lyapunov exponents [54, Sec. 3] which is greater than the number of ensemble members. In this case, the forecast errors will not be corrected by the analysis, because the errors will grow in directions which are not accounted for by the ensemble [42, Sec. 2.2.3]. This problem can be solved by implementing a *localization*, namely by performing the analysis independently for each state variable and by taking the distance between variables into account.

In this study, we implement the \mathbf{R} localization described in [42, Sec. 2.3]. More precisely, every 6 hours we update each state variable independently, which means that we perform the analysis 40 times, and each time we update one state by using its forecast ensemble and some truncated observations. For the space dependence, recall that the states are located on a latitude circle, and therefore the analysis result of a given state should be affected more by its neighboring states rather than by states which are located farther from it. To implement this idea, let us consider how we perform of the analysis of the variable i . We first fix a suitable diagonal observation error covariance matrix \mathbf{R} which takes into account the distance between the site i and any site j : the entry (j, j) of this matrix is given by $e^{d(i,j)^2/2L^2}$, where $d(i, j) = \min\{|i - j \bmod 40|, |j - i \bmod 40|\}$ if $d(i, j) \leq 19$, and $L := 5.45$. Then the analysis of state i is performed with the formulas of Section 4.2.2 on a space of dimension 39, which means that the state and the observations associated with $j = i + 20 \bmod 40$

are disregarded. Note that the value of the tuning parameter L (and the related choice of a truncation at the maximum distance of 19) have been fixed by a minimization process of the analysis RMSE (obtained by the difference between the analysis values and the nature run). In this setting, which will be kept throughout the investigations, the analysis RMSE is 0.1989.

We also test the forecast RMSE for different forecast lengths. To do this, we run the forecast with the analysis value as initial value, and compute the RMSE by comparing the forecast with the nature run at different time points. Note that the previous experiment supplies a new initial analysis value every 6 hours. Figure 4.1 summarizes the forecast ability by providing the values of the forecast RMSE for different forecast lengths.

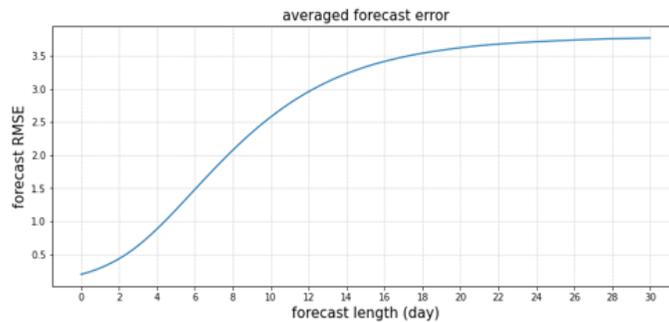


Figure 4.1: The forecast RMSE averaged over every 6 hours in 100 years run

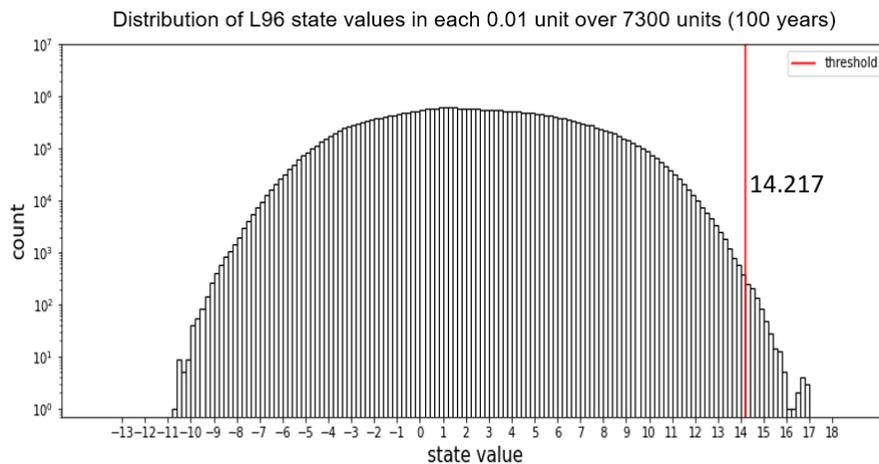
4.3 The Control Simulation Experiment

In this section, we explain the design and the settings of the Control Simulation Experiment (CSE). Several results are already presented in this section, while additional discussion are gathered in the following section. The aim of this CSE is to avoid extreme events, and this will be achieved by adding small perturbations to the dynamical system. These perturbations should drive the system into a preferable direction. Let us mention that some preliminary experiments consisting in replacing the perturbation vectors by randomly generated vectors have shown the same distribution of state values as the nature run without control. Namely, by applying random perturbations one can not reduce the number of extreme events.

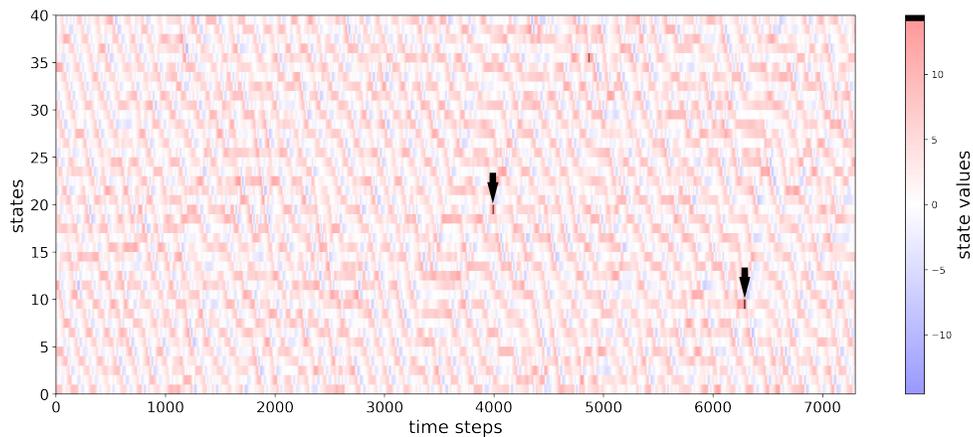
4.3.1 Extreme events

Let us first provide a definition of the extreme events for the Lorenz-96 model. For that purpose, we use the nature run mentioned in Section 4.2.3, namely the trajectories of the 40 state variables recorded every 0.01 unit of time and over 7300 units (100 years). In each period of 0.05 unit of time (6 hours), we look at the maximum value over all states, and keep this value as the local maximum for this period. The first 200 greatest local maxima over the 100 years are treated as extreme values. Note that the interest of considering intervals

of 6 hours is to minimize the representation of events involving several variables taking large values simultaneously or within a very short period of time. As a result of this definition, extreme values appear 2 times a year on average. This construction leads also to a lower threshold value for an extreme event at 14.217. Figure 4.2(a) shows the histogram of the state values (nature run) recorded every 0.01 unit of time for 7300 units, and Figure 4.2(b) shows the value of all states over one randomly chosen year. Values that exceed the threshold are colored in black.



(a)



(b)

Figure 4.2: (a) The distribution of the state values of Lorenz-96 recorded every 0.01 unit of time for 7300 units, (b) The values of all states over a randomly chosen year. The extreme values are colored in black and indicated by arrows.

4.3.2 The Control Simulation Experiments

Recall that synthetic observations are produced every 0.05 unit of time (6 hours). For the forecasts described below, a forecast length T equal to 2, 4, 6, 10, or 14 days is chosen.

1. At a certain time point t_m , we run a T days forecast by integrating the Lorenz-96 model with initial values provided by the analysis ensemble value at time t_m . The integration is performed by using a forth-order Runge-Kutta scheme with integration step of 0.01 unit of time.
2. If no extreme event is forecast during these T days, then we just move to Step 3. If the maximum among all states exceeds the threshold at any time point during these T days, then we apply some perturbations to the nature run from time $t_m + 0.01$ to time $t_m + 0.04$. After each perturbation, we use the perturbed system as initial value for the next integration over 0.01 unit of time. After the fourth perturbation, the system evolves to $t_m + 0.05$.
3. At time $t_m + 0.05$, we have the forecast ensemble and the observation generated by the independent addition of a $Normal(0, 1)$ distributed random number to each variable of the nature run obtained from Step 2. The LETKF is used to assimilate the observation, and an analysis ensemble at time $t_m + 0.05$ is produced. We then move back to Step 1 with t_m replaced by $t_m + 0.05$.

Figure 4.3 shows the steps which are described above. In the upper part of the figure, the green curves represent the T days forecast of one ensemble member. When a green curve crosses or touches the gray line representing the threshold value, it means that one (or more) state has a forecast value which is greater than or equal to the threshold. In the lower part of the figure, the black curves represent the nature run: the original nature run for the dashed curve, and the perturbed nature run for the solid curve. We shall call *the controlled nature run* the nature run obtained by the above process, and emphasize that it corresponds to a nature run including the perturbations, when these ones are applied.

Let us now describe how the perturbation vectors are chosen. For their definition we use two ensemble members A and B among the 10. The ensemble A corresponds to the ensemble showing the maximal extreme value during the forecast T days. The exact place (one variable k) where this maximum is reached, and its precise time t are recorded. For the ensemble member B , we first eliminate the candidates which show an extreme value at any location and at any time within the T days. With the remaining members, we choose the member who has the minimal value for the variable k and at the time t , see Figure 4.4. Then, a rescaled difference of the ensemble members B and A at times $t_m + 0.01$ to $t_m + 0.04$ is applied to the nature run at times $t_m + 0.01$ to $t_m + 0.04$, as indicated in the above Step 2. Let us emphasize that the perturbation vectors are 40 dimensional vectors, since the difference is computed on all variables. For the (Euclidean) norm of the perturbation vectors we choose them proportional to the value of the analysis RMSE $D_0 := 0.1989$. Thus, we fix the norm of this vector as $D := \alpha D_0$ with the magnitude coefficient $\alpha = 0.1, 0.2, \dots, 1.0$. Since the norm

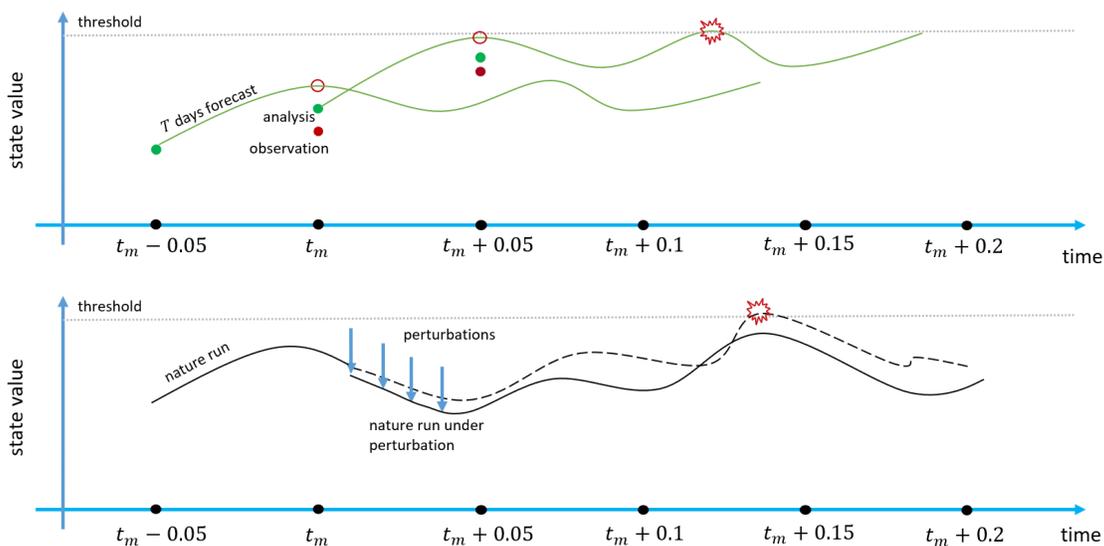


Figure 4.3: CSE description: When a T days forecast from t_m shows extreme values, perturbations are applied to the nature run between t_m and $t_m + 0.05$. At time $t_m + 0.05$, previous forecast and observation based on the controlled nature run are available, and a new analysis ensemble is created.

of the average displacement in each 0.01 unit of time is 1.1720, the norm of the perturbations correspond to less than 2% of the average displacement when $\alpha = 0.1$ and to 17% of the average displacement when $\alpha = 1.0$.

If ever the above process does not allow us to choose an ensemble member B , then we come back 6 hours before, use the analysis ensemble at time $t_m - 0.05$, and run a T days plus 6 hours forecast to find a suitable candidate.

In Figure 4.5, we superpose the distribution of the state values recorded every 0.01 unit of time for 7300 units already obtained in Figure 4.2(a) with the similar distribution for the controlled nature run. The parameters chosen are $T = 10$ days and $\alpha = 0.2$. It is clear that the number of values that exceed the threshold decreases significantly. The successive perturbations that were applied to the nature run effectively control the system and reduce the number of extreme events.

When the norm of the perturbation vector is too small, namely when α is too small, the control is less effective for avoiding extreme events. Similarly, when the forecast length is too short, it is often too late for applying successfully the perturbations. In order to understand the effectiveness of the choice of α and T , we perform the same experiment for different combinations of these parameters. For each combination, we run 10 independent experiments. To be consistent with the definition of extreme events, we define the effectiveness as the success rate given by the formula

$$\text{success rate} := 1 - \frac{\#\{6 \text{ hours intervals in 100 years with extreme events}\}}{200}.$$

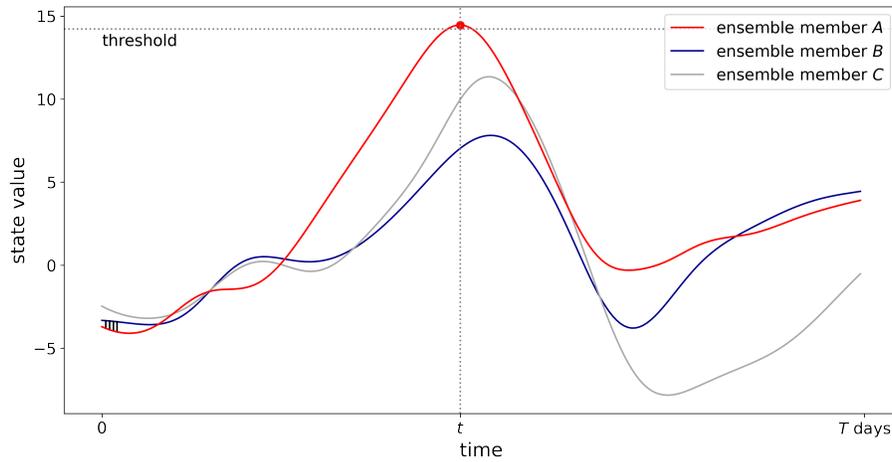


Figure 4.4: The forecast of a state with three ensemble members. Ensemble B is chosen with the farthest distance to ensemble A when A exhibits an extreme event. The four perturbation vectors (before rescaling) are indicated on the lower left part of the figure.

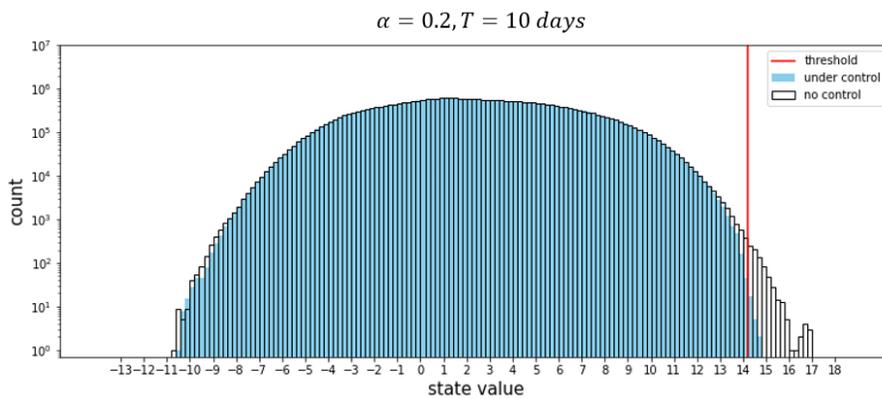


Figure 4.5: Distribution of state values for the nature run and for the controlled nature run.

The outcomes of these investigations are gathered in Figure 4.6.

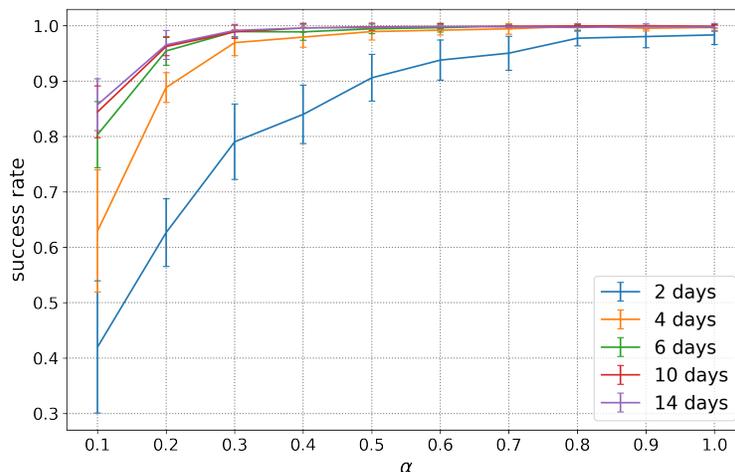


Figure 4.6: Success rates, as a function of the forecast length T and of the magnitude coefficient α .

In addition to the success rate, we also would like to know how many times a control of the system has been implemented, or equivalently how many times perturbation vectors have been applied. For that purpose, we call a *perturbation action* each group of four perturbations, as described in the above Step 2. We provide in Figure 4.7(a) the averaged values of perturbation actions over a period of 100 years and obtained with 10 independent simulations. The intervals corresponding to one standard deviation are also indicated. However, since the magnitude of the perturbations vectors are proportional to α , another quantity of interest is the *the perturbation energy* delivered to the system and defined by

$$\text{perturbation energy} := 4\alpha D_0 \times \#(\text{perturbation actions}).$$

Figure 4.7(b) provides the information about the perturbation energy over a period of 100 years. In other words, this value describes the total energy put into the dynamical system for controlling it during the 100 years run.

By looking at Figure 4.6, we observe that the success rate is quite low for the combination of short forecast length T and small coefficient α . Nevertheless, the lowest success rate is between 0.4 and 0.5, which means that more than 40% of the extreme events are successfully removed. With T longer than or equal to 4 days and α greater than 0.5, all performances are quite similar: a success rate close to 1. On the other hand, it is clearly visible on Figure 4.7(b) that the perturbation energy quickly increases with longer forecasts and larger coefficients α . In fact, the relation between the perturbation energy and the parameters T and α is quite subtle. On the one hand, one observes that for constant T , the number of perturbation actions is decreasing as a function of α , see Figure 4.7(a). Nevertheless, this decay is counterbalanced by the multiplication by the factor α itself, when computing the perturbation energy for any fixed T , as it appears in Figure 4.7(b). On the other hand, the

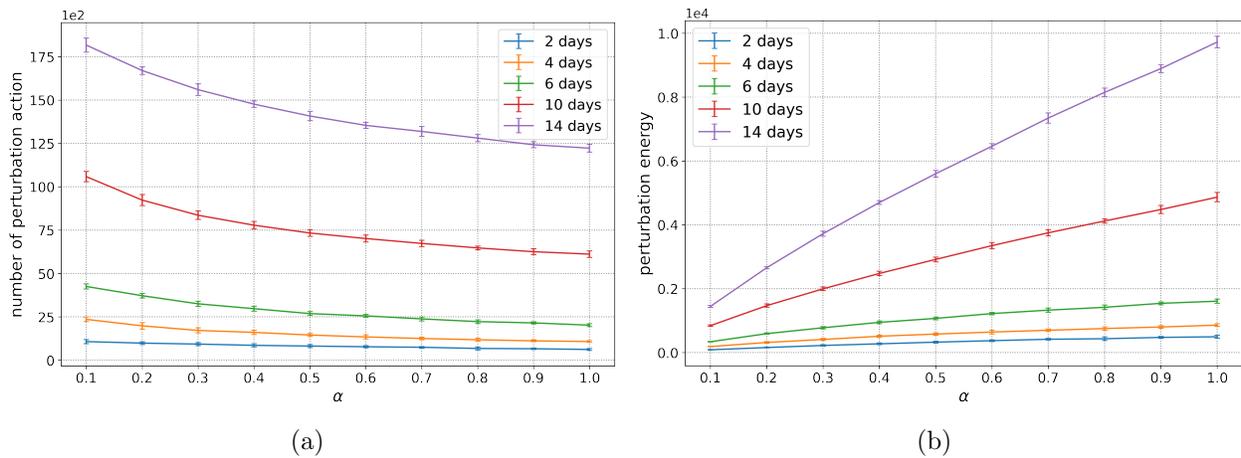


Figure 4.7: Number of perturbation action and perturbation energy over a period of 100 years, as functions of the forecast length T and of the magnitude coefficient α .

increase of T , for any fixed α , has a simple impact in both Figures 4.7(a) and 4.7(b): the corresponding values increase.

The increases with respect to T are quite clear: A longer forecast increases the probability of having one particle exhibiting an extreme value, and thus increases the number of perturbation actions triggered. This also result in an increase of the perturbation energy. For the dependence on α : a small value of this parameter will not be sufficient for dragging the system far away from an extreme event, and therefore a particle has a bigger chance to forecast an extreme event even after a perturbation action. On the other hand, a large value of α corresponds to larger perturbation vectors moving the system momentarily away from an extreme event. In summary and according to these observations, for a pretty good success rate with a small energy cost, the choice $T = 4$ days and $\alpha = 0.5$ looks like an efficient compromise.

4.4 Local perturbations and partial observations

In the previous section, information was available for the 40 variables, and the perturbation vectors were applied at the 40 locations. We shall now discuss the situation when only partial information is available, and when the perturbation vectors are applied only at a reduced number of sites.

4.4.1 Local perturbations

The reason of considering local perturbations is quite straightforward: Since the 40 variables of the Lorenz-96 model represent meteorological quantities evenly distributed on a latitude circle, it could be that perturbations can only be generated at a limited number of locations. Also, one may be interested in using only a smaller number of sites, in order to reduce the

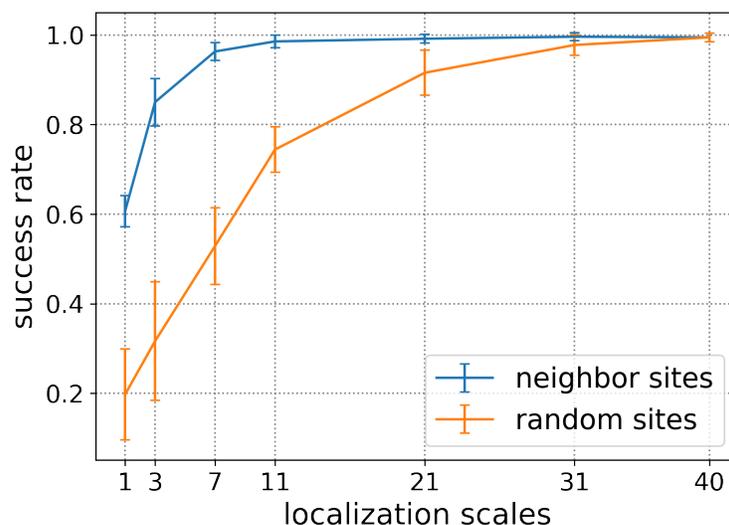


Figure 4.8: Success rates for the two experiments and for different localization scales. The forecast time T is set to 4 days and the magnitude coefficient α to 0.7.

perturbation energy delivered to the system. The purpose of this section is to test if such an approach is possible and efficient.

To study the interest and feasibility of local perturbations, we design two experiments: In the first one, we assume that a randomly selected set of locations can generate perturbations. During one realization (a 100-year run), these locations are fixed. In the second one, we assume that each location can generate perturbations, but only sites which are close to the targeted site, namely the site corresponding to the state variable at which an extreme event is forecast, will be turned on and generate perturbations. Quite naturally, we shall refer to the first experiment as the *random sites* experiment, while the second one will be referred to as the *neighbor sites* experiment.

Recall now that the perturbation vector was previously obtained by the rescaled difference of two ensemble members. For local perturbations, we shall additionally set some entries of the perturbation vector to 0. In other words, some locations among the 40 states will not be perturbed, even if a perturbation is applied to other states. These latter sites will be called *the perturbation sites*, but observe that the perturbation can accidentally be also equal to 0 at some of these sites. For the random sites experiments, we choose successively the numbers of perturbation sites as 1, 3, 7, 11, 21, and 31. These numbers are referred to as *the localization scales*. For the neighbor sites experiments, we shall successively consider perturbation sites which are located at a distance at most 0, 1, 3, 5, 10, and 15 from the targeted site. According to this maximal distance, the localization scale will also coincide with 1, 3, 7, 11, 21, and 31. For both experiments and when an action is suggested by the forecast, we first compute the difference of two the ensemble members are mentioned above, we rescale the resulting vector with the coefficient $\alpha = 0.7$, and set to 0 its entries which are not at a perturbation site. This final vector is used for the perturbation, and its Euclidean

norm provides the energy of this perturbation. For the perturbation energy, we consider the sum of these norms.

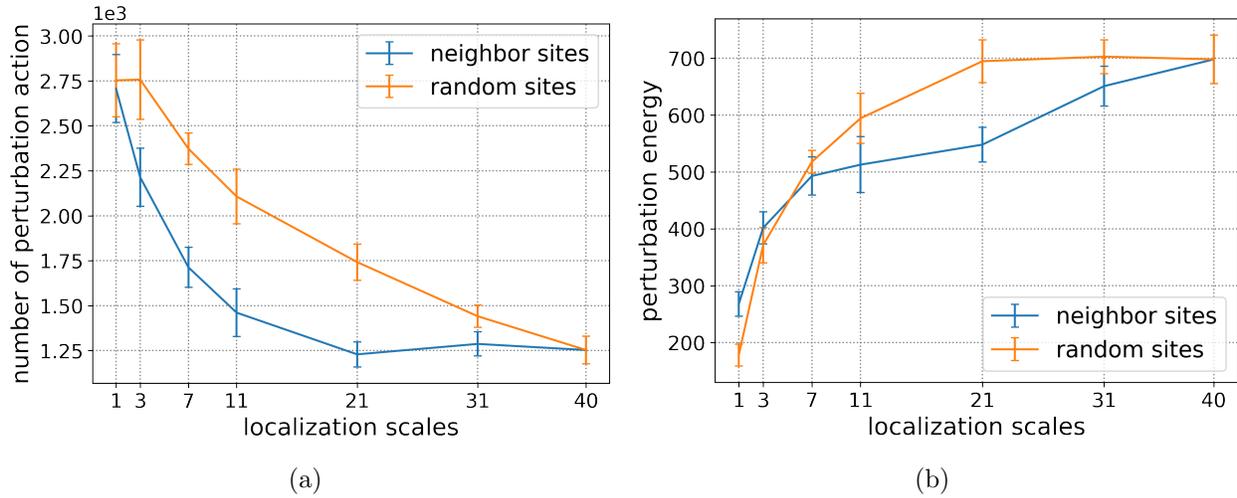


Figure 4.9: Number of perturbation action and perturbation energy for the two experiments and for different localization scales. The forecast time T is set to 4 days and the magnitude coefficient α to 0.7.

In Figure 4.8, we provide the success rate for the different localization scales: 1, 3, 7, 11, 21, 31, and 40. The coefficient α is fixed to 0.7, and the forecast length T to 4 days. The error bars mark the range within one standard deviation of 10 independent realizations. Clearly, using random sites is less effective than using neighbor sites. For localization scale equal to 1, which means only one site is perturbed, localizing the perturbation on the targeted site eliminates already more than 60% of the extreme events, while choosing a random position for the perturbation only eliminates about 20% of the extreme events. When the localization scale increases, the success rates of both experiments increases. However, when the localization scale is greater than 7, the success rates of the neighbor sites experiment does not increase anymore, while there is a constant increase for the random sites experiment, with a maximum at the localization scale equal to 40. For the uncertainty range, the neighbor sites experiment also shows a more stable performance compared to the random sites experiment.

At the level of the number of perturbation action, see Figure 4.9(a), the neighbor sites experiments require less perturbation actions than the random sites experiments. In addition, when the localization scale is greater than 21, the number of perturbation action for the neighbor sites experiments does not decrease anymore. This suggests the existence of an optimal setting for the localization scale around 21. On the other hand, in Figure 4.9(b), the relation between the two experiments is a little bit more confused, but for a localization scale bigger than or equal to 7, it also clearly appears that the neighbor sites experiment is more effective than the random sites experiment. For these localization scales, the perturbation energy is higher for the random sites experiment, even though the corresponding success rates are always smaller. In fact, more perturbation actions are taking place in the random setting, since the system is not carefully driven away from a forecast extreme event. For

the localization scale smaller than 7, the small perturbation energy for the random site experiment is due to the random choice of the components of the difference between the two ensemble members: in general these components are not related to the variable of the extreme event, and therefore to the biggest difference between the components.

By looking simultaneously at Figures 4.8, 4.9(a), and 4.9(b), let us still mention an interesting outcome: For a localization scale equal to 11, the success rate is equal to the success rate obtained for the localization scale equal to 40, but the corresponding perturbation energy in the former case is about $\frac{6}{7}$ of the perturbation energy used for controlling the 40 sites. This ratio can even decrease to about $\frac{5}{7}$ for a localization scale of 7, if we accept a small decay of the success rate. Thus, with a fine tuning of the localization scale, a reduction of the total perturbation energy can be obtained, without any decay of the success rate.

4.4.2 Partial observations

Our next aim is to study the performance of the CSE when only partial observations are available. The lack of information can either be spatial (with the observation of only 20 state variables) or temporal (with observations available only every 24 hours, which means every 0.2 unit of time). We provide the description of both experiments simultaneously.

When observations are sparse in space or in time, the available information from observations decreases, and the analysis becomes less accurate, accordingly. As a consequence, the parameter ρ of the multiplicative inflation and the \mathbf{R} localization parameter L , introduced in Section 4.2.3, need to be adapted. In particular, suitable values for these parameters can lead to a smaller analysis RMSE. Note that the average spread of ensemble members should be close to the RMSE, namely, the value of average should spread basically explain the potential accuracy of the analysis results. Recall that the spread is computed as the RMSE but by replacing the truth by the mean value of the ensemble members.

For the observations sparse in space, the 20 observed state variables are still chosen evenly distributed on a latitude circle. The average spread and RMSE are computed on a 100 years data assimilation experiment with different combinations of parameters ρ and L , and the results are provided in Figures 4.10(a) and 4.10(b). The smallest RMSE together with a spread of a similar value are found for the combination $(\rho, L) = (1.08, 4.85)$, and this common value is about 0.31 (to be compared with the value 0.1989 of the original experiment).

For the observations sparse in time, we have only tuned the parameter ρ and kept the \mathbf{R} localization parameter L at its initial value, since the space structure has not changed. Table 4.1 provides the RMSE and the spread for different value of ρ computed also on a 100 years data assimilation experiment, and the smallest RMSE is reached for $\rho = 1.50$, with a value almost two times bigger than the original RMSE.

With the parameters mentioned above, we run 10 independent CSE for $T = 4$ days and $\alpha = 0.1, 0.5$ and 1.0 . For the observations sparse in time, namely when an observation is taken place every 24 hours only, the system is still perturbed between two observation time and every every 0.01 unit of time. More precisely, if the forecast computed from a time t_m anticipate an extreme event, then the perturbation vectors are applied from time $t_m + 0.01$ to time $t_m + 0.19$, every 0.01 unit of time.

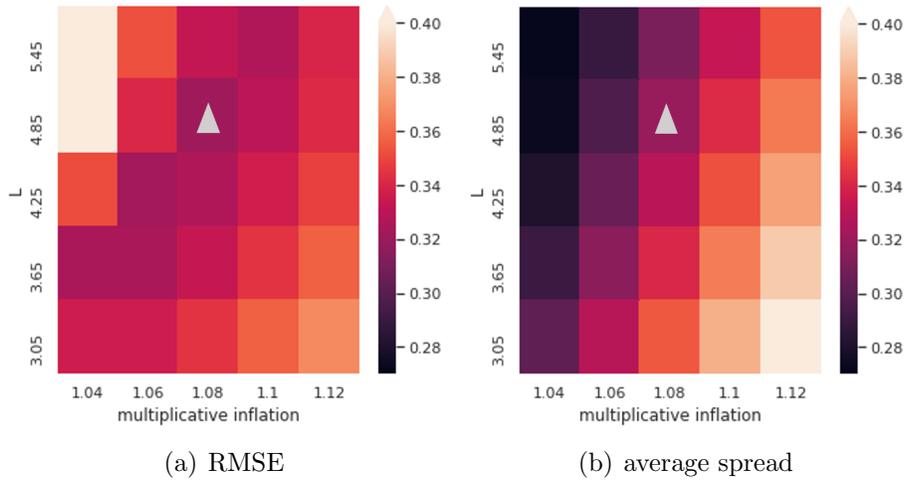


Figure 4.10: The RMSE (a) and the average spread (b) over 100 years run for different combination of multiplicative inflation parameter ρ and \mathbf{R} localization parameter L .

ρ	1.40	1.45	1.50	1.60	1.70
RMSE	0.443	0.429	0.422	0.427	0.434
spread	0.372	0.380	0.389	0.403	0.417

Table 4.1: The RMSE and spread over 100 years run with different values of ρ .

The various results obtained for partial observations are gathered in Figure 4.11. For a comparison, we also recall the results of the original experiment which assumes that all states are observed every 6 hours. Quite surprisingly, when half of the states are masked, and despite an increase of the analysis error by about 50%, the decay of the success rate is pretty small for $\alpha = 0.1$ and 0.5, and negligible when $\alpha = 1.0$. On the other hand, the corresponding perturbation energy is higher than for the original setting, meaning that a more frequent application of the perturbation vectors is taking place. This might be caused by less accurate forecasts and by less efficient perturbations.

In contrast, when the observations are taking place every 24 hours only, the decay of the success rate is more clear, and the increase of the perturbation energy is obvious. For the success rate, the bigger RMSE can certainly explain why the analysis ensemble members can not make accurate enough forecasts even on a very short range (2 days). On the other hand, the increase of the perturbation energy is linked to the repeated application of the perturbation vectors, which took place 19 times instead of 4 times at every alerts. In summary, the partial forecasts, both in space and in time, lead to a CSE which is less efficient and necessitates more perturbation energy.

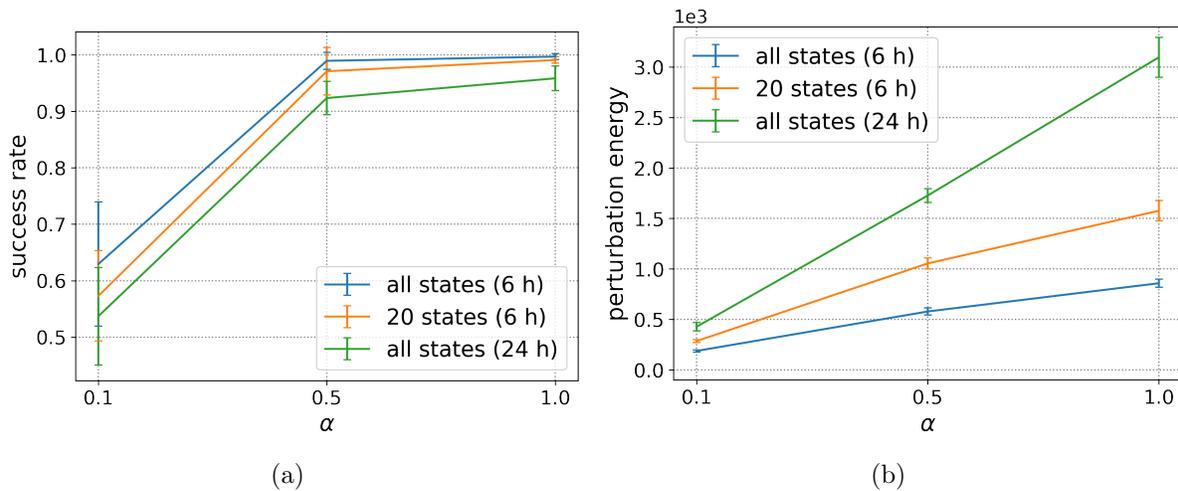


Figure 4.11: Success rate (a) and perturbation energy (b) for sparse observations experiments. The forecast time T is set to 4 days.

4.5 Conclusion

In this study, we developed a CSE for reducing the occurrences of extreme events in the Lorenz-96 model with 40 variables. As background information, the main properties of the model are recalled, and an introduction to LETKF is provided. The use of these tools for investigating the Lorenz-96 model has then been explained, and some parameters have been evaluated with preliminary experiments. Subsequently, we defined the notion of extreme events, and designed the experiment for reducing their occurrence during periods of 100 years. The notions of success rate and of perturbation energy are introduced for discussing the outcomes of the experiments. Sensitivity tests on localized perturbations (vs global perturbations) show the potential of refining the experiments for a decay of the perturbation energy without a concomitant decay of the success rate. Additional sensitivity tests with partial observations show the robustness of the strategy. This study is an extension of previous investigations of controllability of chaotic dynamical systems based on the Lorenz-63 model [63]. Further investigations on more complicated real world models or on numerical weather prediction models are planned. In this context, the potential side effects of the method should be carefully studied, and an global assessment of the CSE would be challenging but very interesting.

Epilogue

In summer 2019, I was accepted to work with the researchers of the Data Assimilation Research Team of RIKEN, for an internship of three weeks. During this stay, I was involved in the project “Analysis of 2-dimensional cellular automaton model using particle filters”. The project was mainly developed by Dr. Sakamoto, who was also my advisor during the internship. Through the research work on this project, I was impressed by the power of particle filters which is one important tool of data assimilation. I was also introduced to the methodology of data assimilation and to some elegant theories and tools in this field.

After the internship, I decided to further extend my knowledge on data assimilation. I deepened my understanding by studying the theories and by presenting them at the seminar arranged by my supervisor, Prof. Richard. During the first semester of the doctoral program, I also submitted my application and a research proposal to the Junior Research Associate Program at RIKEN. This program supports Ph.D. students for achieving their research at RIKEN during the period of their Ph.D. This successful application allowed me to be jointly supervised by Prof. Miyoshi, who is the team leader of Data Assimilation Research Team (DA team), and to continue my research with the members of this team. As a positive outcome, this program also triggered a collaboration between Prof. Richard of the Graduate School of Mathematics of Nagoya University and the DA team.

At the beginning of 2020, the COVID-19 pandemic started affecting people’s live worldwide. The special pattern of the spread of this disease inspires me to construct a model which fits this propagation pattern. Also, the daily reported statistics of the disease by health officials provided very good observation resources. With these regularly reported observations and with our newly designed infectious disease model, it quickly became possible to study the spread of the disease by using data assimilation techniques. One major outcome of our investigations has been the evaluation of the effective reproduction number R_t , one key factor for the evolution of an epidemic. With this new knowledge available, and with our growing understanding of data assimilation techniques, we started conceiving that large-scale observations and simulations could help for preventing further spread of the disease in the future.

Simultaneously, this idea resonated with the newly developed research projects of Control Simulation Experiments (CSE) for Lorenz-63 and for Lorenz-96 models, as presented respectively in Chapter 3 and in Chapter 4. For these models, the main idea was precisely to control the evolution of the underlying dynamical systems by exploiting the forecast power

of DA techniques and the sensitivity to initial conditions. However, in our investigations on COVID-19, as presented in Chapter 2, the amount of information and medical parameters involved in the extended SEIR model is much bigger than the corresponding ones involved in the two toy models. In other words, dealing with the realistic COVID-19 situation was much more difficult than with the two models from Lorenz. In addition, the performance of data assimilation tools on L63 or L96 had already been well studied by many researchers, which means that it was more straightforward to start the control simulation experiments for these toy models, rather than with the model involving several medical constants and human factors. As a result, it was quickly decided to carry out simultaneously both investigations: the CSE research on the Lorenz models, and the more basic but essential research of COVID-19.

In the Analysis of COVID-19, our aim was to understand how data assimilation techniques work on a model with real-world observations. Our main contributions have been the design of the extended SEIR model, the estimation of the unobservable states and of unknown parameters with data from Japan, and the study of the sensitivity of the method with respect to the parameters. To make the extended SEIR model informative but nevertheless simple, we firstly studied the medical research results and the policies related to the identification, the treatment, and the discharge of patients. We adopted the ETKF as the tool for the analysis and we tried different setting of parameters, for example the multiplicative inflation, the observation error covariance matrix, and the number of ensemble members to stabilize the performance of ETKF and to increase the analysis accuracy. Moreover, for the recovery rate and the death rate, we did not use the values suggested by health officials. Instead, we approximated these values by using the daily observations and this led to a more stable performance for the computations.

The analysis results for several cities in Japan and for the whole country confirmed our ability to estimate the effective reproduction number and to evaluate the unobservable states with ETKF. The effective reproduction number R_t computed with ETKF essentially agreed with the values provided by other sources (as for example from ToyoKeizai), but our analysis values for R_t had a smoother evolution. This fact looks rather natural if we consider the propagation property of an infectious disease. On the other hand, we also noticed that the reaction of R_t to some rapid external modifications (like the change of the variant, or rapid non pharmaceutical interventions) was not quick enough. This slow reaction has been corrected and improved by using extra correction procedures. Namely, we have tested several tools for improving our model, as for example the use of ARIMA model, an additional smoother process, and other stochastic procedures, and some of them have been implemented for the scenario simulations.

As a preliminary step for studying the control of a dynamical system through data assimilation, it is important to study the sensitivity to model parameters. As presented in the technical discussion of Chapter 2, we tested our model with different values of the external parameters. For example, since the ratio between symptomatic and asymptomatic patients has been subjected to many controversies (especially at the beginning of the epidemic), we have tested several values for this ratio. The results show that most of these parameters do not affect the analysis value of R_t , which means a great stability of our main outcome.

Once our model has been sufficiently understood and tested, it has been decided to

provide our analysis and some future scenarios on a dedicated COVID-19 website of the DA team, with a daily update. To make the procedure works efficiently, we refined the algorithms and implemented an automatic download of the observations, and an automatic upload of the results. My colleagues Sakamoto-san and Ishimuzu-san from the DA team helped me designing and constructing the website, and Tsuzu-san (a graduate student of Nagoya University) helped me to refine and test part of the code. The final product was available here

http://www.data-assimilation.riken.jp/covid-19/realtime/index_en.html

but this project has been completed on March 31, 2022, because of the end of the RIKEN presidential grant. It also turns out that with the change of variants, the observations became less reliable, and future predictions became less essential. However, our approach and the knowledge obtained have been summarized in a paper, and remain available for future epidemics.

Let us come back to the two Lorenz models. There already existed several applications of DA techniques to the partially observed Lorenz-63 and Lorenz-96 models, as for example the OSSE (Observing System Simulation Experiment). These research results allowed us to directly start the control simulation experiments with these two toy models, without having to deal with preliminary stability tests or the estimation of some parameters' settings. The predictability and the evaluation of data assimilation methods provided useful information for the design of the CSEs. Indeed, it allowed us to design the experiments with the given optimal setting of the data assimilation methods, for example the number of ensemble members or the size of the multiplicative inflation. The data assimilation methods adopted in the research of L63 and L96 are ETKF and LETKF respectively.

Our main contribution in the CSE for Lorenz-63 has been to implement the control simulation experiment and to study the performances of the CSEs by tuning the experimental parameters. It is commonly understood that long-term forecasts of chaotic dynamical systems are difficult because of the sensitivity to the initial conditions. The tricky part of in the design of the CSE is precisely to use this sensitivity to generate appropriate perturbation signals of small amplitudes. Our aim for this model has been to avoid a change of regime, or in other words to keep the evolution of the dynamical system on one wing of the butterfly attractor. If a change of regime is forecast, then a suitable perturbation has to be implemented. The ensemble members and their different trajectories provide natural candidates for the definition of the perturbation. Among these candidates, it is possible to find a good one which does not perform a change of regime during a prescribed time window. Then the signals are computed by using the selected candidate, and it is expected that the addition of those signals to the system at proper time points can keep its evolution in one regime.

The results have shown that the implementation of these well-designed perturbation signals can effectively control the Lorenz-63 model and avoid a regime change. Note that the effective perturbation signals are rather small compared to the analysis RMSE and to the one-step model evolution in the Euclidean norm (OME).

The successful application of the CSE to the Lorenz-63 model inspired the follow-up research related to the control of extreme events in more complex models. The Lorenz-96

model was chosen and the design of the CSE was also updated to deal with the spatial structure of this model. In this experiment, the definition of extreme events was not as trivial as the notion of regime change for the butterfly attractor. Consequently, we carefully defined the notion of extreme events to ensure that these events take place only two times a year on average, and are evenly distributed in the 100-year trajectories. Also, we introduced the idea of perturbation energy for a better understanding of the efficiency of the CSE.

Considering the spatial structure and the advection of the Lorenz-96 model, we also tested the efficiency of local perturbation, and the results have shown that the CSE is efficient enough when the perturbation is applied at a few pre-designed locations only. This observation indicates that further refinements of the CSE are possible, and this partial application might be favoured when the CSE would be applied to a more realistic dynamical model. In the long run, we keep in mind possible applications of the CSE for the control heavy rainfalls or of typhoons, or for the control the spread of a pandemic!

As a conclusion, to control a partially observed dynamical system by using data assimilation techniques, one needs a good knowledge of the predictability (forecast) of the model, and a good understanding of how data assimilation tools can be applied to the system. In the research on COVID-19, we carried out the preliminary research for the forecast of the model. For the research on Lorenz-63 and Lorenz-96, we were able to implement the CSE and to study its effects. The code for these investigations will be opened for public use. We expect that these results can be further extended to other models and it may benefit both academic researchers and any disasters prevention center.

Appendix A

Bibliometric analysis on mathematics: 3 snapshots: 2005, 2010, 2015

A.1 Introduction

This research has been triggered by a simple question: Do international collaborations increase the number of citations in mathematics? By looking at the existing studies about this question in various fields of research [55, 78, 90, 92, 93], the easy and naive answer would be positive. However, some investigations show that a more precise answer depends on the field of research, and that additional information should be taken into account, see for example [29, 80, 81]. Thus, from the original narrow question, our interest has shifted to the more general question: What are the important predictors for the publications in mathematics, if the response is the number of citations?

Similar bibliometric investigations have already been performed, as for example in [17], but mathematics was not considered in this reference, and the analysis is partially model dependent. Specific to mathematics, let us mention the early studies [32, 33] based on MathSciNet, followed by [7] which discusses the different citation indices for mathematics journals, [6] which performs a bibliometric analysis on the period 1868–2008, and [69] which focuses on mathematics education. For very recent investigations, let us also mention [82] which also studies citations but from an individual perspective, [89] which provides a detailed bibliometric analysis over 40 years but on publications of a single journal, and [70] which studies US mathematics faculties with some bibliometric tools.

In order to provide a broad picture about recent publications in mathematics, and therefore complement some of the publications introduced above, our initial hope was to use MathSciNet, which is familiar to all mathematicians and which really focus on mathematics publications. Unfortunately, MathSciNet does not allow any automated searching or downloading, and collecting enough information for any serious analysis turns out to be impossible (despite several requests). On the other hand, Web of Science, which is not specific to mathematics but contains information about mathematics among other fields, allows the collect of large amount of data, and its supporting team answered all our inquiries. For these reasons, after a comparison of the two databases provided in Section A.2, we concentrate in the

subsequent sections on data provided by Web of Science only.

Let us now be more specific about the content of this paper. Our investigations are focusing on publications in mathematics for three years: 2005, 2010, and 2015. This choice allows us to see an evolution in the publication records over a period of ten years, without overwhelming us with too much data. For each of these three years, we collected between 45'000 and nearly 80'000 items related to mathematics, and for each item we kept the record of 10 features as predictors together with the response, namely the number of citations up to November 2020. These predictors are introduced and discussed in Section A.3.

The preliminary analysis consists in looking at the response as a function of a single predictor. Let us immediately stress that since the response depends on time (the number of citations increases as years pass), all investigations are performed on the three years independently. Section A.4 contains these results, presented either with graphs or with tables. More precisely, the citations are provided successively as a function of

- (i) the number of authors,
- (ii) the number of countries associated with the authors,
- (iii) the number of institutes associated with the authors,
- (iv) the number of references provided by the authors,
- (v) the number of pages of the publication,
- (vi) the number of keywords provided by the authors
- (vii) the open access (or not) of the publication,
- (viii) the journal impact factor JIF (if the publication has appeared in a journal with a JIF),
- (ix) the research area of the publication,
- (x) the categories associated with the publication.

More explanations and comments are provided in Section A.4.

It clearly appears in these individual investigations that the response is related to some predictors, but how much information can be extracted from them, and what is their relative importance? These questions, and others, are discussed in Section A.5. Because of the diversity of the predictors we opted for an approach based on tree-based methods, as introduced in [11]. Indeed, unlike the approach provided in [78] we do not want to consider some linear relations between the predictors and the response, but prefer an approach which divides the predictor space into several regions and associates to each region a local response. Alternatively, we could have borrowed some *bibliometric* tools developed in [2] if our investigations were performed on R, but we opted for the tools available on the platform [77].

Several experiments are performed with trees, with some parameters chosen according to the year of publications and to the existence (or non-existence) of a JIF associated with the

publications. Based on these experiments, the predictors can be ranked according to their importance. Another outcome of tree classifiers is the ability of predicting the citations (at least within some predefined classes) based on the predictors. Clearly, the result is not very good, but the converse would have been even more surprising. However, the predictions are better than a random guess, as explained in Section A.5.

In Section A.6 we turn our attention to countries: What information can be deduced from the individual publications about the research in the countries of the authors? Can one measure a kind of performance for each country? And what about collaboration between countries, which is related to our very initial question, can one measure these collaborations, and say something about them? Data for answering these questions are presented in Section A.6 for the main countries, which means for the country producing the majority of publications in mathematics. In fact, data covering about 130 countries were available, but for some of them, the annual number of publications is too limited to support any analysis.

With this paper we provide three snapshots (2005, 2010, 2015) about the publications in mathematics, and extract as much bibliometric information as possible. As already mentioned, we would have preferred working on a database MathSciNet because some information would have been more accurate. It is quite unfortunate that the policy and the tools provided by this website do not allow such investigations, as implicitly acknowledged in [19]. On the other hand, by using Web of Science database, our investigations about mathematics have covered a slightly broader range of publications.

A.2 General pictures

In this section we provide general information about publications in mathematics for the last 20 years. A few comparisons between the two databases MathSciNet (MSN) and Web of Science (WoS) are also presented. Finally, the data we shall use in the following sections are introduced, and some statistics are exhibited.

Since this research is based on data about mathematics, let us first have a look at two important sources of information. MathSciNet is an electronic database operated by the American Mathematical Society focusing exclusively on publications in mathematics. In November 2020 it contains about 3.9 millions items. Web of Science is a much more general database operated by the private company Clarivate. It is possible to select publications in mathematics by choosing the *research area* mathematics (SU=mathematics). In November 2020, the

outcome for this general request is about 2.0 millions items. Note that WoS contains also *categories*, and one of them corresponds to mathematics. However, by choosing this request

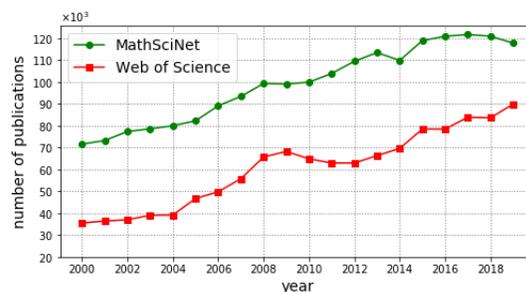


Figure A.1: Yearly new indexed math publications

(WC=mathematics) the number of items is 1.7 millions, and these items are strictly contained in the previous request about research area.

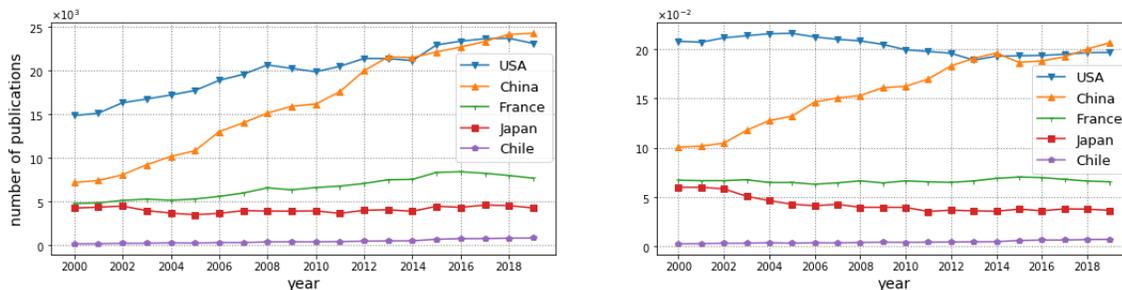


Figure A.2: Publications with at least one author from a given country: absolute and relative numbers from MSN

A comparison about the number of works published in the last 20 years is provided in Figure A.1. For MSN, all works are reported, while for WoS it is again the works corresponding to the research area mathematics. Let us provide one more comparison between the two databases, based on one information that will be used in the analysis. The information is related to the country in which research institutions (universities, research institutes, etc) are located. For simplicity, we shall call this the *country* of the research institution, and by extension the country of the author working in this research institution. Figures A.2 and A.3 show the yearly publications and their relative numbers with at least one author from a research institution in one of the following countries: USA, China, France, Japan, Chile. The relative numbers are with respect to the total number of publications in mathematics index by MSN and WoS (shown in Figure A.1). Thus, even if the difference between the total numbers of items in the two databases is not negligible, we expect that their shapes and trends are similar.

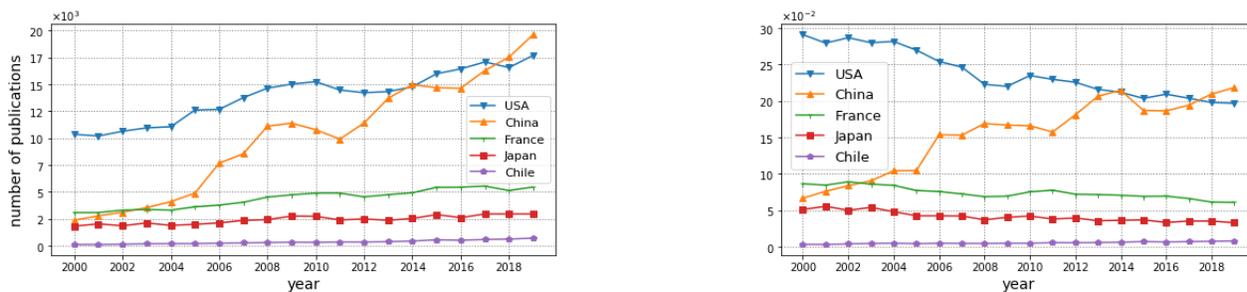


Figure A.3: Publications with at least one author from a given country: absolute and relative numbers from WoS

On the other hand, a unique feature of MSN is the Mathematics Subject Classification (MSC). This classification contains more than 60 subjects, and each one can be divided into numerous sub-subjects. Each publication is indexed by one primary subject or one primary

subject with several secondary subjects. The subjects are usually chosen by the author(s) of a publication, or carefully assigned by the editors of MSN. The MSC provides a rather precise information about the content of each publication. With this information, a refined plot of Figure A.1 for MSN is shown in Figure A.4. The eight combined subjects are elaborated in [75]. Unfortunately, WoS does not contain the MSC, and therefore we shall not be able to use this information in our investigations.

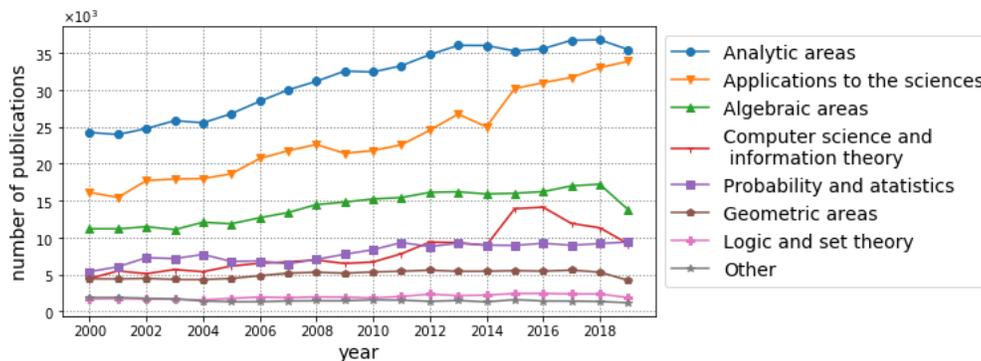


Figure A.4: Yearly new indexed math publications (8 fields) based on the MSC

As mentioned in the Introduction, our analysis is based on the data in three years. With the request *research area = mathematics*, and once the items with no clear author or with no clear affiliation for the author(s) have been removed, the number of items collected are:

$$\begin{aligned}
 &\text{research area = mathematics, clear author(s) and affiliation(s)} \\
 &2005 : 45'035 \text{ items} \quad 2010 : 62'945 \text{ items} \quad 2015 : 76'788 \text{ items.}
 \end{aligned}
 \tag{A.1}$$

The following statistics are computed on these numerous items. It has been observed in earlier publications that the average number of authors for each paper has been increasing over time, see for example [6, Figure 10]. Since our investigations are based on three distinct years, let us observe this effect on a period of 10 years, see Figure A.5.

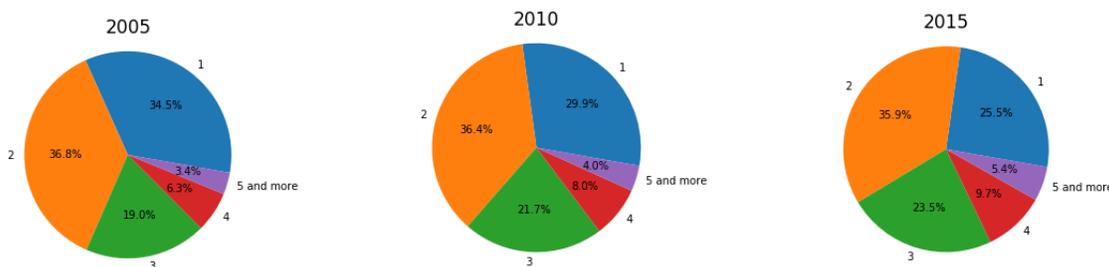


Figure A.5: Distribution of the number of authors per publication

It clearly appears that the proportions of publications with 1 and 2 authors are decreasing, while the ones with 3, 4, or 5 and more authors are increasing. The average number of authors

for these three years and based on the data mentioned in (A.1) are respectively:

$$2005 : 2.10 \text{ authors} \quad 2010 : 2.23 \text{ authors} \quad 2015 : 2.39 \text{ authors.} \quad (\text{A.2})$$

Note that other numbers confirm this increase in the collaborations for each publication. Indeed, if one looks at the average number of research institutes involved for each publication one gets

$$2005 : 1.57 \text{ institutes} \quad 2010 : 1.74 \text{ institutes} \quad 2015 : 1.90 \text{ institutes.} \quad (\text{A.3})$$

These numbers have been computed by counting the number of different addresses provided by the publications.

Since one of our interests is to study international collaborations, let us provide similar results for the average number of countries involved for each publication:

$$2005 : 1.23 \text{ countries} \quad 2010 : 1.28 \text{ countries} \quad 2015 : 1.32 \text{ countries.} \quad (\text{A.4})$$

Again, these numbers have been computed by counting the number of different countries mentioned in the list of addresses of the authors. If we look at the details, one obtains the distributions provided in Figure A.6.

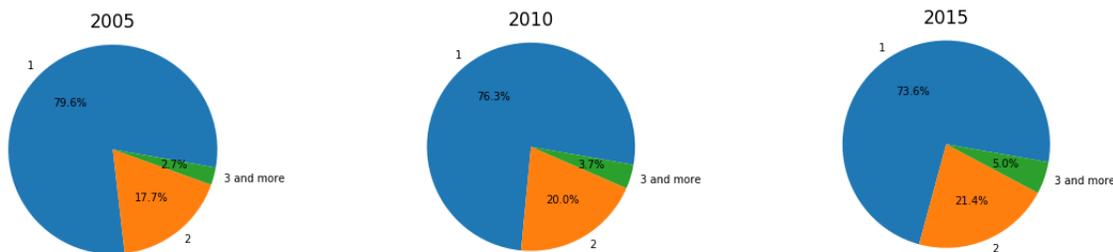


Figure A.6: Distribution of the number of countries per publication

A.3 The predictors and the response

In this section we introduce the predictors and the response that we have employed for our investigations, and make a few comments about them.

The predictors can be roughly divided into three categories, namely those related to the author(s) of a publication, those related to the publication itself, and those related to the journal or the physical support in which the publication has appeared. All of them have been extracted from the WoS database for the items mentioned in (A.1). Let us immediately stress that the author's names have been completely disregarded in our investigations.

- Author(s)

authors: the number of authors

countries: the number of countries

institutes: the number of research institutes

Note that in addition to the number of countries involved for each publication, the exact list of countries will also be investigated in Section A.6.

- Publication

references: the number of references

pages: the number of pages

keywords: the number of keywords provided by the authors

open access: open access

WoS provides some information about various types of open access. More precisely, publications which are partially or fully open access are identifiable in the database. This information is obtained in collaboration with *EndNote Click*, formerly called *Kopernio*. Since there exist various levels of open access, this predictor will be used cautiously.

- Journal

jif: journal impact factor

research areas: research areas

categories: categories

The *journal impact factor* is computed by WoS and assigned to several journals for each year. More information about its computation and its weaknesses can be found here¹. As already mentioned, *research areas* and *categories* are classification indices provided by WoS. Several research areas and several categories have been assigned to each journal in the WoS database, based on several criteria. Compared with the rather precise definitions of the categories, the research areas are less precisely defined (this has been confirmed by the technical support from Clarivate who answered our inquiries). Moreover, compared with Mathematics Subject Classification (MSC) from MathSciNet, these two predictors are hugely less precise. Not only these indices are not chosen by the authors, but they are common to all publications in one journal, and their assignment is not so clear. Nevertheless, they provide a vague information which deserves to be collected, and which will be further discussed later on.

For the response, the number of citations for each publication has been recorded. These numbers were collected in October/November 2020. It should be emphasized that these numbers range between 0 and some very large numbers. For example, one work published

¹https://en.wikipedia.org/wiki/Impact_factor

in 2005 has been cited up to 11'106 times (12'015 and 18'439 times for the most cited works published in 2010 and 2015).

A few other publications are also cited numerous times, which is quite unlikely in the field of mathematics. Since such publications have an enormous impact in the computations of means, we have decided not to consider them. More precisely, we have decided to keep only the publications with a number of citations strictly below 64. This number corresponds to the lower 95th percentile of the collected publications from 2005. For simplicity, we kept this upper limit of 64 also for the data from 2010 and 2015 (which corresponds respectively to the lower 97th and 99th percentiles). As a consequence, the data with less than 64 citations are

$$\begin{aligned} \text{research area} &= \text{mathematics, clear author(s) and affiliation(s), citation} < 64 \\ 2005 : & 42'792 \text{ items} \quad 2010 : 61'084 \text{ items} \quad 2015 : 76'168 \text{ items.} \end{aligned} \quad (\text{A.5})$$

In the subsequent computations, and in particular for computations of means, it is these items which are considered.

Let us finally mention that we could have considered the 95th percentiles for the three years, which means that the upper limit for 2010 and 2015 would have been lower than 64. As a result, the means related to these years would have been slightly smaller. However, since we can not compare directly the citations between publications produced respectively in 2005, 2010 and 2015, the simpler choice of keeping the upper bound 64 instead of keeping 95th percentiles does not affect our investigations.

A.4 Individual predictors

In this section we study the relations of the individual predictors with the number of citations. A comparison between the predictors, based on a tree classifier, will be presented in the next section.

A.4.1 Number of authors

The items mentioned in (A.5) have been divided according to the number of authors (1, 2, 3 and more) and the respective distributions of citations have been reported in the first column of Figure A.7.

For each group, the mean has been indicated with a vertical line, and the median has been reported with a vertical dashed line. These precise values are indicated in Table A.1, as well as the proportions of publications with 1, 2, or 3 and more authors.

On Table A.1, it clearly appears that the fractions of publications with 1 or 2 authors decrease over time, while the one with 3 and more authors increases over time. This observation goes in line with the content of Figure A.5, when no upper limit for the number of citations was imposed. On the other hand, it clearly appears that means and medians of the three years' record increase along with the number of authors.

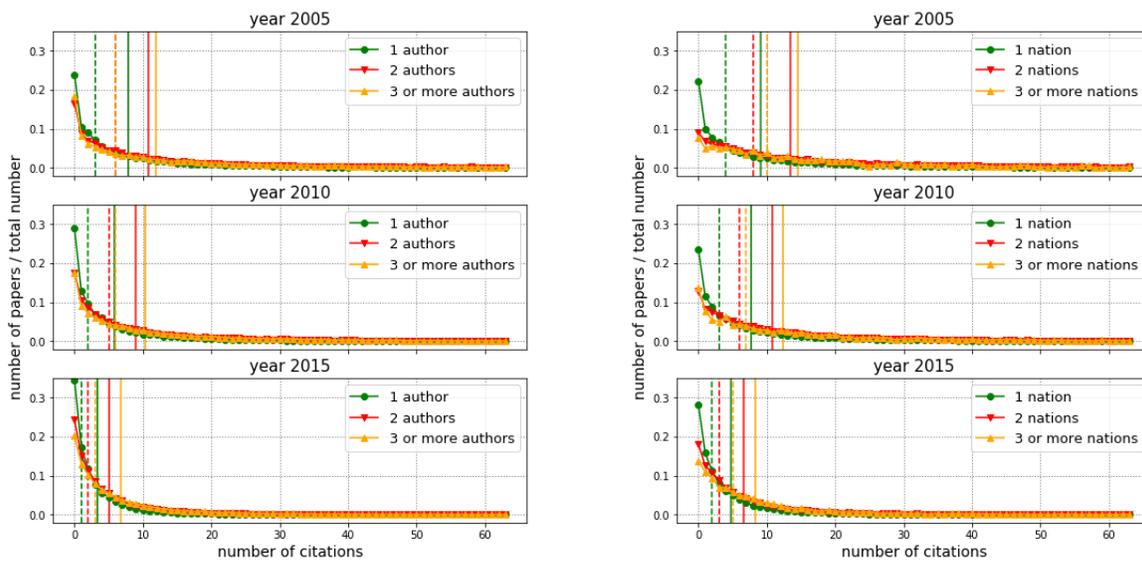


Figure A.7: Distributions depending on the number of authors or on the number of countries involved

	2005			2010			2015		
	%	mean	median	%	mean	median	%	mean	median
1 author	35.3	7.8	3	30.4	5.9	2	25.6	3.4	1
2 authors	36.6	10.8	6	36.5	9.0	5	35.9	5.1	2
≥ 3 authors	28.1	11.8	6	33.1	10.3	6	38.5	6.7	3

Table A.1: Citations depending on the number of authors

A.4.2 Number of countries

Let us now divide the items according to the number of countries appearing in the list of addresses of the authors. A division into 1, 2, 3 and more countries has also been performed, and the distribution of citations is reported in the second column of Figure A.7. The values of the means, the medians, and the proportions are provided in Table A.2.

	2005			2010			2015		
	%	mean	median	%	mean	median	%	mean	median
1 country	80.0	9.2	4	76.5	7.7	3	73.7	4.7	2
2 countries	17.4	13.5	8	19.8	10.8	6	21.4	6.6	3
≥ 3 countries	2.6	14.5	10	3.7	12.3	7	4.9	8.3	5

Table A.2: Citations depending on the number of countries involved

Similar to the content of Table A.1, the number of countries involved for each publication increases over time, and the means and medians grow along with the number of countries. The first observation confirms the trend observed in Figure A.6. These observations reflect the

internationalization of research and publication processes. However, let me mention a small effect inside this global picture. For most items, the appearance of n countries correspond to at least n authors, one (or more) in each country. However, there also exist tens of items with one author having two main addresses in two different countries. Since these situations confirm an additional face of internationalization, we haven't tried to separate these two effects.

A.4.3 Number of institutes

Another division according to the number of different research institutes appearing in the list of addresses is performed. We divided the items into 1, 2, 3, 4 and more institutes, and computed the proportion, the mean and the median for each of these groups. These numbers are reported in Table A.3.

	2005			2010			2015		
	%	mean	median	%	mean	median	%	mean	median
1 institute	60.7	8.0	3	51.7	6.8	3	44.7	4.2	2
2 institutes	27.6	12.5	8	31.0	9.6	5	33.1	5.5	3
3 institutes	8.8	14.5	9	12.4	10.8	6	14.6	6.6	3
≥ 4 institutes	2.9	16.5	12	4.9	12.8	8	7.6	8.1	4

Table A.3: Citations depending on the number of institutes

As for the previous two predictors, the number of publications involving only one institute decrease over time, while the number of publications involving two and more institutes increase over time. By looking at these three predictors, it appears quite clearly that an increase in the number of authors, countries, or institutes, corresponds to an increase in the number of citations. In fact, both the mean and the median are increasing with these predictors.

A.4.4 Number of references

For most items provided by WoS, the number of references mentioned by the authors of a publication is provided. A classification depending on the number of references has been realized, and Figure A.8 provides the information about the number of publications (y -axis) with a given number of references (x -axis), together with the citation mean (color). Note that in the three graphs, the grey color correspond to the mean citation over all data of the respective year. These means appear in Table A.6 (first row), but for the record let us already mention them:

$$2005 : 10.0 \text{ citations} \quad 2010 : 8.5 \text{ citations} \quad 2015 : 5.3 \text{ citations.} \quad (\text{A.6})$$

By computing the average number of references for the 3 years, the expected number of references are:

$$2005 : 18.3 \text{ references} \quad 2010 : 20.7 \text{ references} \quad 2015 : 23.8 \text{ references.} \quad (\text{A.7})$$

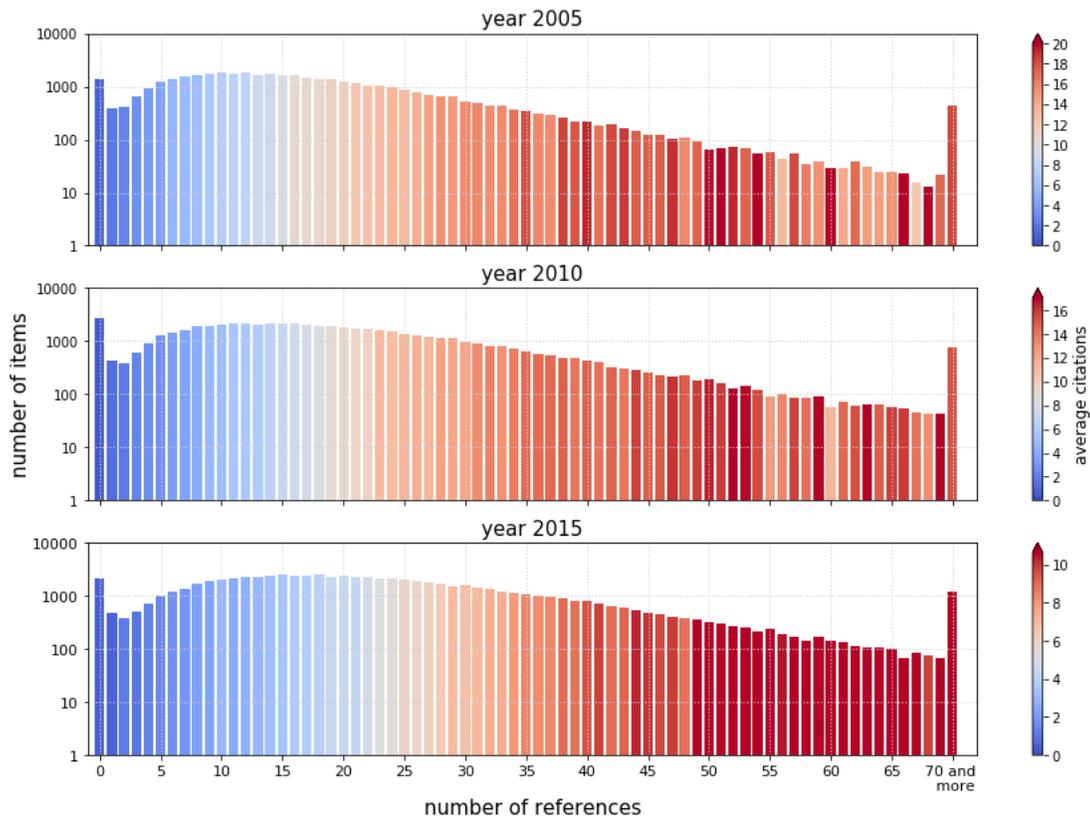


Figure A.8: Citations depending on the number of references

A similar computation can be realized by eliminating the items with 0 references. Indeed, it is rather doubtful that a publication does not mention any reference, but it seems more likely that WoS has not recorded the information for these publications. In order to eliminate this bias, we compute the expected number of references, once the items with 0 references are disregarded, and one gets:

$$2005 : 19.0 \text{ references} \quad 2010 : 21.7 \text{ references} \quad 2015 : 24.5 \text{ references.} \quad (\text{A.8})$$

The numbers in (A.7) and in (A.8) show a rather steep increase of the average number of references over a period of 10 years. Also, publications with the mean number of references received about the mean number of citations for each year. It would certainly be valuable to look at this predictor with data from additional years and over a longer period of time.

A.4.5 Number of pages

Similar investigations can be performed with the number of pages. Figure A.9 contains the outcomes of this investigation, with the number of items for a given number of pages together with an information about the citation mean.

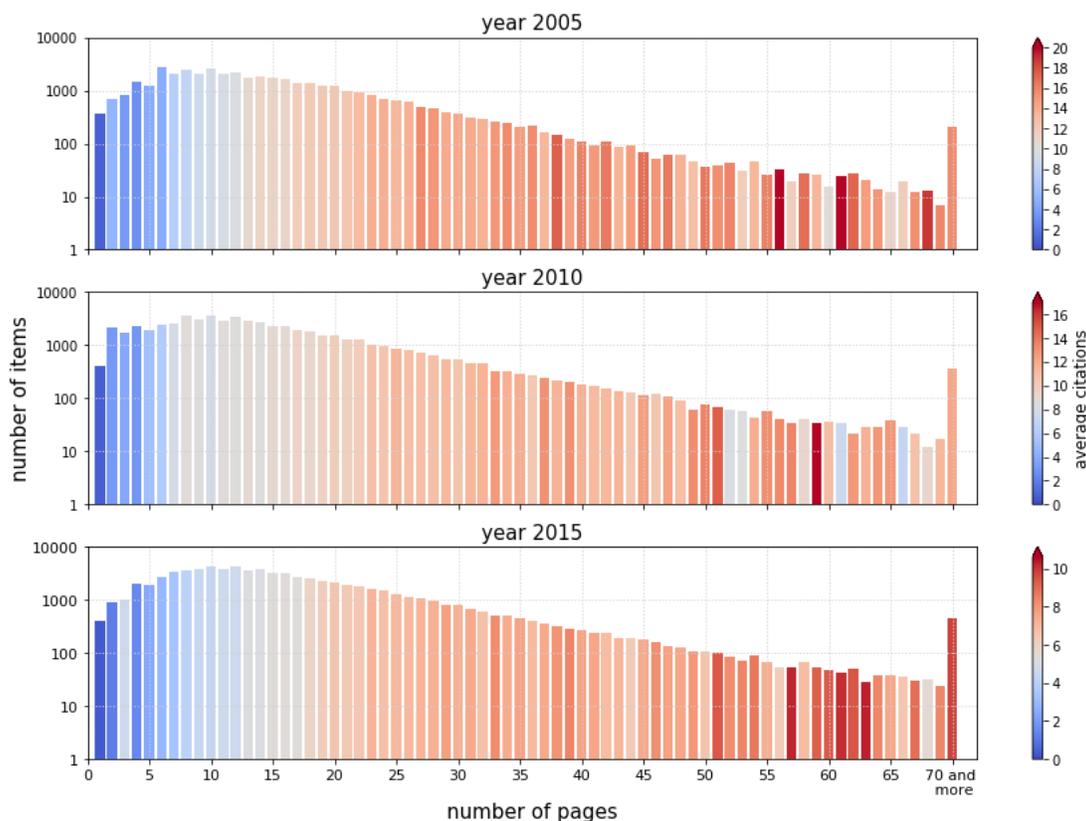


Figure A.9: Citations depending on the number of pages

The conclusion is similar to the one obtained for Figure A.8, namely a positive correlation between the number of pages and the citations. Once again, the average number of pages can be computed. The expected numbers of pages are

$$2005 : 15.5 \text{ pages} \quad 2010 : 15.5 \text{ pages} \quad 2015 : 16.9 \text{ pages.} \quad (\text{A.9})$$

Compared to the previous result with the number of references, the increasing trend of number of pages is less clear. Also, one observes that the contrast of colors is more important in Figure A.8 than in Figure A.9. This difference can be understood as the predominance of the number of references over the number of pages as a predictor for the number for citations. This fact will be confirmed in the next section.

A.4.6 Number of keywords

Let us now have a look at keywords. WoS lists the keywords provided by the authors, and a second list of keywords introduced by WoS. Let us immediately say that we were interested only in the first list. In Table A.4, we report the proportions of publications as a function of the number of keywords, and provide also the means and the medians. Note that there is no information about keywords for about 1/3 of the data for each year. This absence can be

due either to the lack of keywords provided in the original support of the publication, or in a failure in the collect of this information by WoS.

nb keywords	2005			2010			2015		
	%	mean	median	%	mean	median	%	mean	median
0	36.6	8.3	3	28.7	6.7	2	23.7	4.4	1
1	0.3	7.7	3	0.3	4.4	2	0.2	3.1	1
2	4.3	8.1	4	3.9	6.8	3	3.0	3.4	2
3	15.9	9.8	5	16.7	7.6	4	16.2	4.3	2
4	17.6	10.9	6	20.1	8.7	4	21.9	5.4	3
5	13.1	11.5	7	15.8	10.3	6	18.6	6.2	3
6	6.4	13.3	8	7.9	11.2	6	9.5	7.0	4
7	2.7	12.9	7	3.2	10.7	6	3.3	6.3	3
8	1.3	14.8	10	1.7	11.3	7	1.7	6.2	3
9	0.7	15.3	11	0.8	11.5	7	0.9	6.9	4
10	0.4	14.3	9	0.4	12.0	6	0.5	6.0	3
≥ 11	0.5	15.3	11	0.5	11.2	7	0.6	6.6	4

Table A.4: Citations depending on the number of keywords

What appears in this table is a positive relation between the number of keywords and the citation mean as long as the number of keywords is between 1 and 5. For the items with more than 5 keywords, any relation between the number of keywords and the citation mean is not really visible.

Let us still provide the average number of keywords. Based on WoS, the expected number of keywords are:

$$2005 : 2.78 \text{ keywords} \quad 2010 : 3.19 \text{ keywords} \quad 2015 : 3.49 \text{ keywords.} \quad (\text{A.10})$$

In order to eliminate the bias related to 0 keywords, the statistics with 0 keywords removed have also been computed, and one gets

$$2005 : 4.38 \text{ keywords} \quad 2010 : 4.48 \text{ keywords} \quad 2015 : 4.58 \text{ keywords.} \quad (\text{A.11})$$

As for the number of references and the number of pages, it is visible in (A.10) and in (A.11) that the number of keywords provided by the authors is also following an increasing trend.

A.4.7 Open access

As already mentioned, WoS lists the items which have a partial or a full open access. Since this information is available, let us just provide the proportion of items having any kind of open access, together with the respective mean and median. The outcomes are summarized in Table A.5.

The table contains one information which is not surprising: the proportion of items with an open access increases over the years. Also, both the means and the medians are larger for

	2005			2010			2015		
	%	mean	median	%	mean	median	%	mean	median
OA	16.4	12.2	7	21.2	9.7	5	32.3	5.7	3
no OA	83.6	9.6	5	78.8	8.1	4	67.7	5.1	2

Table A.5: Citations with or without open access

the items with an open access compared to the items without this access. However, even if these results look quite natural, we think that further and more precise information would be necessary to get a better picture about the impact of open access.

A.4.8 Journal impact factor

By definition, the journal impact factor (JIF) is available only for journals, and not for all items. Note that WoS does not link automatically the publications appearing in journals with the corresponding JIF. However, we were able to track this information for about 2/3 of our items. More precisely, a JIF has been associated to the following numbers of items from the list described in (A.5):

$$\begin{aligned}
 &\text{research area} = \text{mathematics, clear author(s) and affiliation(s), citation} < 64, \\
 &\text{JIF available} \tag{A.12} \\
 &2005 : 31'556 \text{ items} \quad 2010 : 44'639 \text{ items} \quad 2015 : 57'756 \text{ items.}
 \end{aligned}$$

It is not surprising that this list of items is biased. Indeed, if we look at the citation means over these items, the numbers are not exactly the ones appearing in (A.6) but

$$2005 : 12.0 \text{ citations} \quad 2010 : 10.2 \text{ citations} \quad 2015 : 6.3 \text{ citations.} \tag{A.13}$$

There are certainly many reasons why these subgroups are not fully representative of their original groups. For example, all publications appearing in proceedings or in books are not linked to a JIF. Thus, for the next graph only the items of the list (A.12) have been considered, but elsewhere let us stress that the more general list (A.5) was preferred.

In Figure A.10 we report the citation means and the citation medians as a function of the JIF. For the graphs, we have divided the possible values of the JIF into subintervals of length 0.1, collected all papers linked to a JIF in each subinterval and computed their mean and median. The information about the number of items corresponding to each subinterval is also mentioned. For a JIF below 2, a nearly linear relation between JIF and citation mean or median is quite visible. On the other hand, for values of the JIF above 2, the relation is less clear. But this part of the statistics is performed on much fewer items, and makes it less reliable.

A.4.9 Research areas

To each item, WoS associates one or several research area(s). As mentioned in the previous section, our selection was based on *research area = mathematics*, which means that all

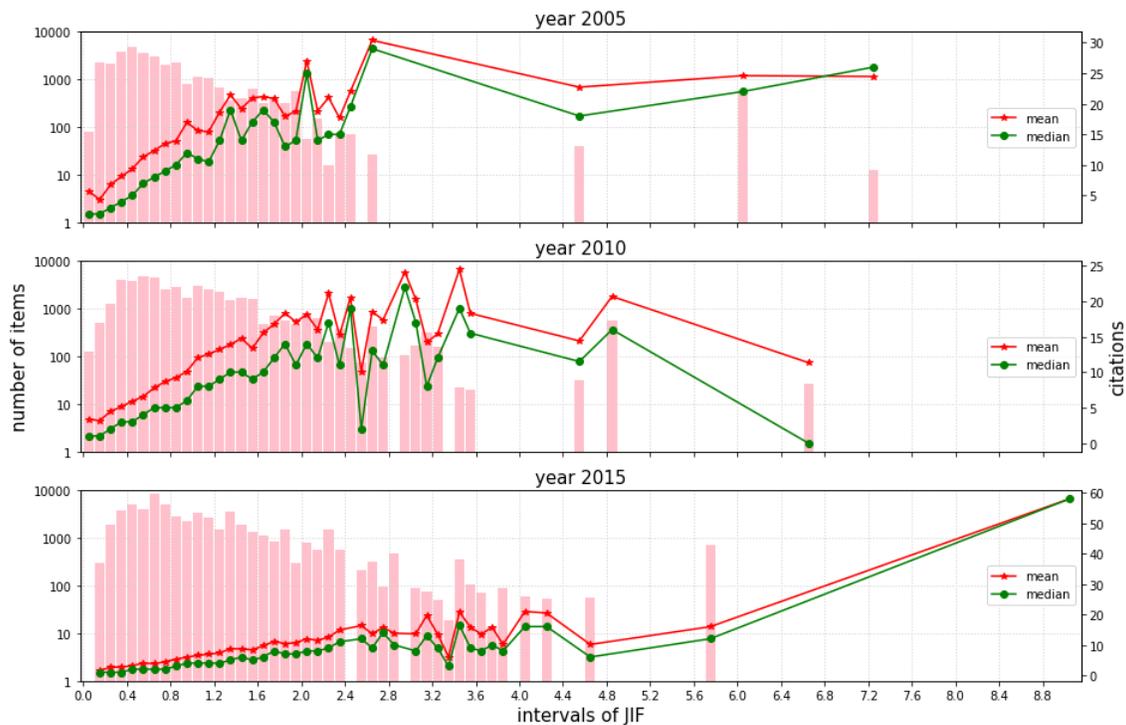


Figure A.10: Citations depending on the Journal Impact Factor

data have at least *mathematics* in the list of their research areas. However, most of them possess more than just one research area. For the data from 2005, it turns out that only 12 additional research areas were contained in at least 1% of the considered items. The list of these research areas is: *Mathematics, Computer Science, Engineering, Physics, Mechanics, Operations Research & Management Science, Mathematical & Computational Biology, Mathematical Methods in Social Sciences, Business & Economics, Biochemistry & Molecular Biology, Automation & Control Systems, Science & Technology, Biotechnology & Applied Microbiology*.

Based on this list, we have computed the fraction, the mean and the median of all the items introduced in (A.5) which possess one of these entries as a research area. The result is provided in Table A.6.

In this table, we have also reported the *relative citation* which corresponds to the citation mean of a particular research area divided by the citation mean of the same year for all the data. It turns out that these computations give some interesting results: the range of this relative citation is between 0.7 and 2.8, with the lowest value shared in 2005 and 2010 by *Engineering* and by *Operations Research & Management Science*. On the other hand, most of the highest values are reached by research areas related to biology or to biotechnology. Note that the median follows a similar pattern, but since the computation of a relative median does not look so natural, we have refrained from providing such a relative information. Thus, this table confirm that the citation mean or median really depend on the research areas, and that

	2005				2010				2015			
	%	mean	rel. c.	med.	%	mean	rel. c.	med.	%	mean	rel. c.	med.
Mathematics	100	10.0	1.0	5	100	8.5	1.0	4	100	5.3	1.0	2
Comp. Science	15.2	9.9	1.0	4	15.0	8.4	1.0	3	9.9	6.7	1.3	3
Engineering	8.7	7.1	0.7	1	11.3	6.1	0.7	1	8.7	6.8	1.3	3
Physics	8.7	11.3	1.1	6	7.9	7.7	0.9	3	7.3	5.1	1.0	2
Mechanics	4.5	9.4	0.9	4	5.3	10.6	1.2	5	3.7	10.9	2.1	7
Op. Research	4.3	6.7	0.7	2	4.4	6.0	0.7	2	2.7	5.6	1.1	3
Math. & C. Bio.	3.2	19.9	2.0	15	3.5	13.1	1.5	8	2.4	10.4	2.0	6
Math. M. Soc. S.	2.5	14.0	1.4	9	2.8	11.3	1.3	6	2.5	8.4	1.6	5
Bus. & Eco.	2.8	8.7	0.9	1	2.3	9.1	1.1	4	2.2	6.4	1.2	3
Bio. & M. Bio.	1.8	23.1	2.3	19	1.3	16.9	2.0	12	1.3	13.4	2.5	9
Auto. & C. Syst.	1.7	11.5	1.1	6	1.8	11.9	1.4	6	1.8	12.0	2.3	8
Science & Tech.	1.7	9.8	1.0	6	1.9	6.5	0.8	3	2.2	4.1	0.8	2
Bio. & A. Mic.	1.6	23.9	2.4	20	1.1	17.4	2.1	13	1.0	14.8	2.8	11

Table A.6: Citations for the main research areas

some striking differences exist. This fact is well documented, and has led to the developments of several relative indices, see for example [1, 16].

A.4.10 Categories

Categories correspond to another indexation of the items chosen by WoS. They correspond to rather broad research fields, but some of them coincide also with research areas. The main list of categories (different from any research area) appearing in our items are the following: *Applied Mathematics*, *Mathematics*, *Statistics & Probability*, *Mathematics (Interdisciplinary Applications)*, *Computer Science (Interdisciplinary Applications)*, *Mathematical Physics*, *Computer Science (Theory & Methods)*, *Engineering (Multidisciplinary)*, *Mechanics*.

For each of these categories, the proportion, the citation mean, relative citation, and the median have been reported in Table A.7. The variations already observed in Table A.6 are also visible here. Note that even in the three main categories, namely Mathematics, Applied mathematics, and Statistics and Probability, the means and the medians are clearly not equal, even if the variations are less pronounced than with categories further apart.

A.5 Investigations with decision trees

In the previous section, we studied the relationships between the number of citations and some individual predictors extracted from the WoS database. In several figures and tables, it clearly appears that the citations are related to these predictors. The problems left are how much information on the citations can be explained by the predictors and what is the

	2005				2010				2015			
	%	mean	rel. c.	med.	%	mean	rel. c.	med.	%	mean	rel. c.	med.
Math. app.	45.1	9.5	0.9	4	50.5	8.2	1.0	4	41.8	5.6	1.1	3
Mathematics	41.5	8.6	0.9	5	40.7	7.2	0.8	4	46.3	3.9	0.7	2
Sta. & Prob.	16.2	12.8	1.3	7	15.4	10.5	1.2	6	15.8	6.5	1.2	3
Math., Int. App.	14.9	11.8	1.2	6	12.3	11.5	1.4	6	16.4	7.2	1.4	3
C. S., Int. App.	7.1	10.6	1.1	4	6.9	8.4	1.0	3	3.7	7.9	1.5	4
Math. Phys.	6.3	10.9	1.1	5	5.2	8.5	1.0	3	3.7	7.7	1.4	4
C. S., T. & M.	5.0	9.6	1.0	4	5.0	6.8	0.8	3	3.8	4.7	0.9	2
Eng., Mult.	5.1	9.2	0.9	3	6.3	7.6	0.9	2	6.8	7.0	1.3	3
Mechanics	4.5	9.4	0.9	4	5.3	10.6	1.2	5	3.7	10.9	2.1	7

Table A.7: Citations for the main categories

relative importance among the predictors. To answer these questions, we use decision trees, as thoroughly introduced in [11].

A.5.1 Tree classifier

A tree classifier is a procedure that divides a data set into two or more subsets based on some predetermined criteria. Let us explain this on a simple example. Suppose that the response contains two classes: Yes and No, which are represented by red and blue dots in Figure A.11. Suppose also that there are two predictors X_1 and X_2 attached to each item. The tree classifier can identify the best split value of one of the predictors such that the purity in each subset is enhanced. After the splitting into two subsets, further splits will be carried out independently for each subset. The classifier goes through all the possible values of the predictors at each split. It stops splitting when a stopping rule is met. During the process, each subset is called a node, the nodes without further subsets are called leaves. When a node contains more than one class (which means that the node is not pure), the class with the majority of items is selected as the label class of the node. The *misclassification rate* of the node corresponds to the ratio of items in the node belonging to a class which is not the label class.

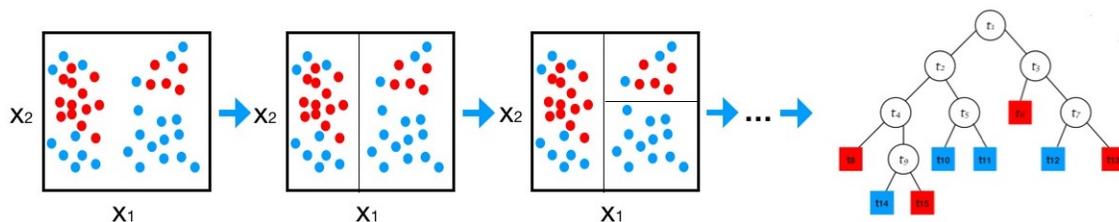


Figure A.11: Tree-structured classifier

There exist several criteria for the choice of the split value, but all of them are based on an *impurity function* which has to be minimized. More precisely, let H denote such an

impurity function, and let us consider a split of the content of one node t into two subsets, called t_{left} and t_{right} . Then one sets

$$Q_t = \frac{N_{t_{left}}}{N_t} H(t_{left}) + \frac{N_{t_{right}}}{N_t} H(t_{right}), \quad (\text{A.14})$$

where N_t , $N_{t_{left}}$, and $N_{t_{right}}$ denote the number of items in the note t and in the two subsets. The function H is also evaluated on the items of t_{left} and t_{right} . By considering all possible splits, we choose the one which generates the minimal value for Q_t .

For the impurity function, a few canonical choices are possible. In order to define them, consider now that the response contains J classes, and assume that in a node t the items are distributed following a distribution $\{p_j\}_{j=1}^J$, with p_j the proportion of items in the class j . Then, some canonical impurity functions are

$$\begin{aligned} \text{Gini index : } H(t) &= \sum_j p_j(1 - p_j) \\ \text{Entropy : } H(t) &= - \sum_j p_j \ln(p_j) \\ \text{Misclassification : } H(t) &= 1 - \max_j(p_j) \end{aligned}$$

In the subsequent investigations, we shall use the impurity function provided by the Gini index only.

By building a tree with the above process, we often end up with a very big tree: many leaves and a big height, which corresponds to the maximal distance between the first node (root) and the farthest leaf. Such a big tree leads often to an overfitting phenomenon. Thus, a pruning procedure need to be done to reserve the effective tree structure and to remove the risk of overfitting. One can perform the pruning procedure by removing successively the leaves with the least contribution of decreasing the impurity, namely the weakest leaves. There exists several ways for implementing such a process, let us therefore only sketch the main ideas of the cost complexity pruning. To each node t , one associates a real coefficient $\alpha_{eff}(t)$ which takes into account the misclassification of the node, the misclassification of the subtree having t as a root, and the number of leaves of the subtree, see for example [77]. Then, starting from 0 and by slowly increasing a parameter α , one prunes successively the nodes with α_{eff} smaller than α . Obviously, an additional stopping rule has to be fixed, otherwise the process would end up in keeping only the root, which means the original set of items without any subdivision. Again, several options exist. Before presenting one of these stopping rules, let us mention some outcomes of the pruning operation.

Consider that α is slowly increasing from 0, and that the pruning is taking place. It is clear that the number of leaves and the height of the successive trees are decreasing. On the other hand, the total misclassification of the tree tends to increase. Equivalently, the *train accuracy* (ratio of correctly classified items in the leaves) tends to decrease as α increases. It is thus natural to look for a suitable value of α , leading neither to a too small tree which is not able to do effective classification, nor to a too large tree with high risk of overfitting.

Such a suitable α can be obtained by testing the tree on new items. Indeed, consider a new item having all necessary predictors and labeled by one class. According to the value of its predictors, the item can be placed in a unique leaf with a labeled class. It may be correctly or incorrectly classified. By repeating this operation on several new items, one get a *test accuracy*, the ratio of correctly classified new items. Again, if one considers the family of trees obtained by pruning according to the parameter α , one observes that the test accuracy starts by increasing with α , before decreasing again. Since we are interested in the highest test accuracy, we then select the optimal value α_{opt} of the parameter α corresponding to the maximum test accuracy. A typical example of the train accuracy and the test accuracy as a function of α is provided in Figure A.12. It is then this α_{opt} which is used for stopping the pruning process.

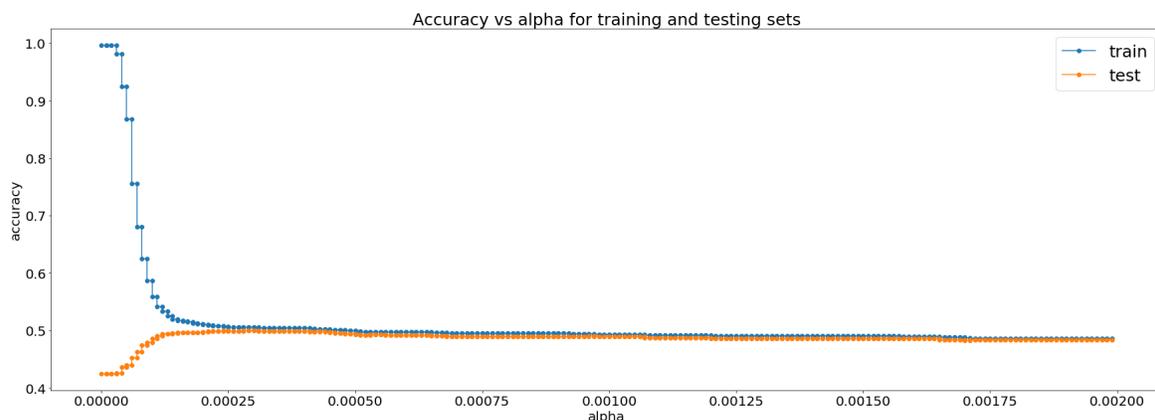


Figure A.12: Train accuracy and test accuracy, as a function of α

A.5.2 The experiment

In this section, we implement these processes and describe the precise experiment we have performed. The analyze tool package is provided by Scikit-learn (Supervised Learning / Decision Trees), see [71]. The experiments will be done independently on the three datasets of 2005, 2010, and 2015. For the predictors, we shall use those introduced and discussed in Section A.3. However, for the predictor *jif*, some items miss this information because they do not have an associated JIF. We have then decided to perform the experiment independently on two lists of items. The first list contains all items, as shown in (A.5), and the predictor *jif* will not be used. The second list contains items with a JIF, as shown in (A.12), and the predictor *jif* is included in the list of all possible predictors.

For the six lists of items mentioned above, we define a family of classes, namely a partition of the set $\{0, 1, 2, \dots, 63\}$, corresponding to the J classes mentioned in the previous section. The partitions should be relevant and understandable and they should also be chosen accordingly to the specificity of the different lists of items. For example, the following partition

will be used for the first list of items of 2005:

$$\text{weak: } [0, 5], \quad \text{normal: } [6, 12], \quad \text{good: } [13, 20], \quad \text{very good: } [21, 40], \quad \text{excellent: } [41, 63]. \quad (\text{A.15})$$

Let us now describe the precise construction of the tree classifier, and the pruning process.

- (i) Fix one dataset among the six presented in (A.5) and in (A.12),
- (ii) Fix one relevant partition for the dataset, as for example the one presented in (A.15). The number of classes defined by this partition is denoted by J and corresponds to the number of intervals,
- (iii) Label the items in the dataset with one of the J classes, according to their citations, and select randomly X items for each class. For our experiment we have chosen X equal to 80% of the items of the smallest class (which has always been the one corresponding to the highest citations class). One ends up with JX items equally distributed among the J classes,
- (iv) Divide randomly these JX items into K folds $\{\Lambda_k\}_{k=1}^K$ of equal size, for $K \in \mathbb{N}$, and fix $k = 1$,
- (v) The fold Λ_k is called the *test set*, and the others $K - 1$ folds are combined in a *training set*. A classification tree is build with the training set,
- (vi) On the classification tree, the pruning process introduced in the previous section is performed: the computation of $\alpha_{eff}(t)$ for each node t , the increase of α , the pruning of the weakest node (*i.e.* the ones with the smallest α_{eff}), the computation of the train accuracy and of the test accuracy. These accuracies, denoted respectively by $a_k(\alpha)$ and $b_k(\alpha)$, are computed for each α (but are piecewise constant) and are stored. The value k is updated by setting $k := k + 1$,
- (vii) The steps (v) and (vi) are repeated as long as $k \leq K$,
- (viii) For each α , the average $a(\alpha)$ and average $b(\alpha)$ are computed by averaging the K values and α_{opt} is deduced by generating a graph like Figure A.12 with the average values (K -fold cross-validation),
- (ix) The optimal tree is built based on the JX items and pruned with the α_{opt} found at the previous step.

Note that in our experiments we have used the parameter $K = 5$. Once this optimal tree is built, the relative importance of the predictors can be obtained. Indeed, the relative importance of the predictors can be computed by looking at the *weighted impurity decrease* at each node. By using the notation already introduced in (A.14), this quantity is provided for each node t by the formula

$$\frac{N_t}{N} \left(H(t) - \frac{N_{t_{left}}}{N_t} H(t_{left}) - \frac{N_{t_{right}}}{N_t} H(t_{right}) \right),$$

where N is the total number of items in the tree. Clearly, if a node is a leaf, there is no contribution to be subtracted. The above quantity provides an information about the decay in the impurity provided by a subdivision. Then, since each subdivision is associated with a single predictor, this decay in the impurity is gained by the corresponding predictor. By summing up the decay in impurity due to all subdivisions associated with one predictor, one obtains the total decay in impurity due to this predictor. The predictors are finally ordered by the decreasing order of their total decay in impurity.

In Table A.8, we provide various information obtained with the tree classifier, namely the relative importance of the predictors, as mentioned above, but also the number of leaves and the height of each tree. The average train accuracy and the test accuracy are also reported. These two values correspond to the two accuracies obtained at α_{opt} after cross-validation.

	2005		2010		2015	
JIF	No	Yes	No	Yes	No	Yes
classes	[0,5] [6,12] [13,20] [21,40] [41,63]	[0,6] [7,14] [15,23] [24,45] [46,63]	[0,4] [5,10] [11,20] [21,63]	[0,5] [6,12] [13,24] [25,63]	[0,3] [4,9] [10,63]	[0,4] [5,11] [12,63]
# leaves	24	15	27	29	49	25
height	7	6	8	6	10	6
av. train acc. (%)	32.9	34.5	39.3	40	51.3	51.5
av. test acc. (%)	31.3	30	38.4	37.8	50.4	50.7
references	1	2	1	2	1	2
authors		4	3	3	3	4
countries						3
institutes	2					
pages	4	3	4	4	2	
jif		1		1		1
research area	3 (ENG.)		2 (ENG.)		4 (Bio. & A. M.)	
Experiment (%)	39	40	44	43	53	53

Table A.8: Experiments with tree classifier

Let us now make several comments about Table A.8.

- (i) First of all, the number of classes and their precise values is partially arbitrary, and was determined after several preliminary tests. Note that we considered less classes for the recent data since the citations accumulate with time.
- (ii) The size of the resulting tree is determined by the computation of α_{opt} as explained above. We were surprised that these trees are relatively small, with a number of leaves between 15 and 49, and a height of maximum 10. On the other hand, the 6 trees contained between 4'800 and 20'000 items.
- (iii) About the accuracies: since the J classes are of equal size, a random guess for an item

of the test set would give a correct prediction with a probability of 20% for the data of 2005, 25% for the data of 2010, and 33.3% for the data of 2015. Thus, the difference between these values, and the test accuracy can be understood as the gain in prediction due to the tree. However, since the train set and the test set do not follow the initial distributions of items, a different computation has to be performed for an arbitrary set of items in the initial dataset (see below).

(iv) The ranking of the predictors for each of the 6 trees is reported in the table, but only for the first 4. It turns out that only 7 predictors among 10 appear in the ranking. For publications which appear in a *jifted* journal, this predictor is always chosen first, and the number of references used by the authors is chosen as the second predictor. On the other hand, for the larger set of publications, with no use of the predictor *jif*, it appears that the number of references is always chosen as the first predictor, while the second one is different in the three experiments. Note that the importance of the predictor *references* was already anticipated in Section A.4 just by looking at the sharp contrasts of Figure A.8.

(v) An additional experiment has been performed with the items not used for the construction of the trees. Indeed, the 20% of the smallest class was still available, and 20% of the untouched items in the other classes could be selected randomly. We apply the classification tree on the new test set which is made up of these items. The last row of Table A.8 provides the percentage of the correctly assigned classes. Clearly, these accuracies do not coincide with the train accuracy or the test accuracy, since the corresponding items do not share the same distributions. The interpretation is the following: given an arbitrary item from one of the three years, and based on the constructed trees, our ability to predict correctly the citation class corresponds to the last row of the table. Not surprisingly, this accuracy increases as the number of classes decreases, but the knowledge of the JIF for the item does not improve our prediction. Even if these numbers show the limitations of our approach, it also provides a heart-warming message: the content of a publication still matters for the citations, and any bibliometric analysis won't be able to predict this.

A.6 About countries

In this last section we provide some statistics related to countries. Indeed, by selecting the items according to the location of the corresponding research institutes, some additional information can be extracted.

In Table A.9 we provide a comparison between the main 25 countries (according to their number of publications). These statistics can be thought as a kind of relative performance. More precisely, Table A.9 contains the following entries:

% : Percentage of the publications having at least one author from the given country,

h.c.: (highly cited) Percentage of publications having more than 63 citations with at least one author from the given country,

mn: (mean) Citation mean for the publications having less than 64 citations and at least one authors from the given country,

md: (median) Median for the publications having less than 64 citations and at least one authors from the given country,

r.c.: (relative citation) Ratio of mn by the citation mean over all publications with less than 64 citations.

The following abbreviations for the countries is used: US = USA, CN = People's R. China, FR = France, DE = Germany, IT = Italy, UK = England, CA = Canada, JP = Japan, ES = Spain, RU = Russia, AU = Australia, KR = South Korea, PL = Poland, IL = Israel, NL = Netherlands, IN = India, BR = Brazil, TW = Taiwan, BE = Belgium, CH = Switzerland, SE = Sweden, GR = Greece, CZ = Czech Republic, AT = Austria, TR = Turkey.

	2005					2010					2015				
	%	h.c.	mn	md	r.c.	%	h.c.	mn	md	r.c.	%	h.c.	mn	md	r.c.
US	27.7	40.6	12.0	7	1.2	24.0	37.3	10.1	5	1.2	20.7	32.9	6.5	3	1.2
CN	10.8	12.8	11.0	6	1.1	17.1	17.2	8.6	4	1.0	19.1	29.5	6.5	3	1.2
FR	8.0	9.3	12.0	7	1.2	7.7	7.1	10.2	6	1.2	7.0	7.4	6.0	3	1.1
DE	7.3	8.7	11.2	6	1.1	7.3	7.7	8.9	5	1.0	6.9	10.6	6.1	3	1.2
IT	5.5	4.8	9.9	5	1.0	5.1	4.2	9.4	5	1.1	5.1	7.4	6.6	4	1.3
UK	5.3	7.8	12.2	7	1.2	5.0	8.1	10.7	6	1.3	4.7	11.1	6.6	4	1.2
CA	4.6	5.0	11.2	6	1.1	4.2	5.2	9.8	5	1.2	3.5	3.7	5.7	3	1.1
JP	4.4	2.5	8.5	4	0.8	4.3	2.1	6.4	3	0.8	3.7	2.7	4.2	2	0.8
ES	4.1	3.4	11.0	6	1.1	4.3	4.2	8.9	5	1.1	3.7	3.2	5.8	3	1.1
RU	4.1	1.1	6.2	2	0.6	3.8	0.9	5.6	2	0.7	4.8	1.4	3.8	2	0.7
AU	2.9	3.3	9.7	4	1.0	2.1	3.1	9.7	5	1.1	2.5	4.0	6.1	3	1.2
KR	2.2	1.5	9.1	4	0.9	2.0	1.9	8.3	4	1.0	2.5	2.6	4.9	2	0.9
PL	2.1	0.7	8.7	4	0.9	2.2	0.8	6.9	3	0.8	2.6	1.8	4.7	2	0.9
IL	1.9	1.9	11.2	6	1.1	1.7	1.6	8.7	5	1.0	1.4	1.4	5.5	3	1.0
NL	1.8	2.1	11.6	6	1.2	1.7	2.3	9.8	6	1.2	1.3	1.6	6.6	3	1.3
IN	1.7	1.3	10.4	6	1.0	2.3	2.3	9.1	4	1.1	4.1	2.7	4.5	2	0.8
BR	1.6	1.3	10.3	6	1.0	1.8	1.6	8.7	5	1.0	2.1	0.8	5.0	3	1.0
TW	1.6	1.6	12.4	7	1.2	1.8	1.5	9.3	5	1.1	1.4	0.6	4.8	2	0.9
BE	1.4	1.9	11.5	6	1.1	1.3	1.3	9.8	5	1.2	1.1	1.9	6.9	4	1.3
CH	1.3	2.1	12.9	7	1.3	1.2	2.4	10.8	6	1.3	1.4	2.7	7.9	5	1.5
SE	1.2	1.2	10.8	6	1.1	1.1	0.8	7.9	4	0.9	1.1	1.6	5.4	2	1.0
GR	1.2	1.0	9.1	4	0.9	0.9	0.6	8.5	5	1.0	0.7	1.3	5.6	2	1.1
CZ	1.2	0.6	7.5	2	0.7	1.3	0.6	6.9	3	0.8	1.5	0.0	3.8	1	0.7
AT	1.1	1.5	10.4	5	1.0	1.2	1.8	9.7	5	1.1	1.2	2.3	6.1	3	1.1
TR	1.1	0.9	10.5	6	1.0	1.7	2.0	9.9	5	1.2	2.0	1.4	5.0	2	0.9

Table A.9: Statistics for the main countries

Let us make some comments about this table. First of all, columns % and h.c. sum up to more than 100 (when all countries are considered) because publications involving authors from different countries are counted more than once. Note also that it is the first (and last) time that outliers (namely publications with more than 63 citations) are used: Columns h.c. provides the information on how much the countries are involved in the highly cited publications. The columns with r.c. can be thought as a comparison between the performance of these countries: it is rather striking that the relative citations take values between 0.6 and

1.5. Some countries are clearly performing better than others. On the other hand, these data do not present a clear pattern over the 3 years considered, and the relative citations are quite stable. Only a few countries have a small variation of their relative citations over the 3 years, but the evolution is not so noticeable. Most probably, a period of 10 years is not long enough to really assert a real change of performance for a given country.

The next table, Table A.10, is more related to the importance of international collaborations for each country. By international collaboration we mean a publication with one author from the given country, and at least one author from another country. More precisely, Table A.10 contains the following entries, again for the main 25 countries:

% : Percentage of publications of a given country which are international collaborations,

mn: (mean) Citation mean for the publications of a given country which are international collaborations and which have less than 64 citations,

md: (median) Median for the publications of a given country which are international collaborations and which have less than 64 citations

rcc: (relative citation for international collaborations) Ratio of mn by the citation mean over all publications of this country with less than 64 citations,

rc2: (relative citations for international collaborations / 2 authors) Ratio of mn by the citation mean over all publications of this country with less than 64 citations and at least 2 authors.

Note that we have added the column rc2 in order to eliminate a bias. Indeed, an international collaboration involves at least 2 authors (with a few exceptions already mentioned in Section A.4.2) while arbitrary publications from any country also include many single author productions. Since these publications are usually less cited (see Table A.1), we eliminate this bias by considering only publications with at least 2 authors.

By looking at this table, the interest of international collaborations is quite clear. All the relative citations appearing in rcc or rc2 are bigger or equal to 1. In fact, some countries benefit a lot from international collaborations, having a factor rc2 taking a maximum value of 1.5. On the other hand, for some countries, publications which are international collaborations do not receive substantially more citations than publications involving only researchers in this country. As a rule and not surprisingly, authors in a country with a low relative citation in Table A.9 benefit more from international collaborations than authors from a country with a high relative citation. Fortunately, no researcher from any country is penalized by establishing international collaboration: such an unfortunate situation would end up with a rcc or rc2 smaller than 1 in Table A.10.

In Table A.11, we provide more specific information about bi-national collaboration. More precisely, since our dataset is large enough, collaboration between two countries can be extracted. It corresponds to publications involving at least one author in a country X , and one author in a country Y . Additional authors and/or countries can also be involved. This number can then be divided either by the total number of international collaborations

	2005					2010					2015				
	%	mn	md	rcc	rc2	%	mn	md	rcc	rc2	%	mn	md	rcc	rc2
US	30.1	14.4	9	1.2	1.0	34.9	12.3	7	1.2	1.0	42.4	7.5	4	1.2	1.0
CN	23.7	14.1	9	1.3	1.2	21.2	12.4	7	1.4	1.4	25.5	8.6	5	1.3	1.3
FR	38.7	14.4	9	1.2	1.1	48.1	11.5	7	1.1	1.0	54.7	6.6	4	1.1	1.0
DE	40.1	13.6	9	1.2	1.1	47.6	10.8	7	1.2	1.1	51.8	7.1	4	1.2	1.0
IT	33.7	13.6	8	1.4	1.2	42.4	11.5	7	1.2	1.1	49.5	7.8	5	1.2	1.1
UK	46.8	14.3	9	1.2	1.0	54.0	12.3	7	1.1	1.0	61.5	7.3	4	1.1	1.0
CA	51.2	13.6	8	1.2	1.1	53.7	11.4	7	1.2	1.1	60.8	6.4	3	1.1	1.0
JP	23.3	12.3	8	1.4	1.2	30.3	8.7	5	1.4	1.2	33.3	6.3	3	1.5	1.3
ES	36.0	14.2	10	1.3	1.2	46.2	10.5	6	1.2	1.1	54.4	6.4	4	1.1	1.1
RU	29.3	11.5	7	1.8	1.5	30.3	9.0	5	1.6	1.2	25.0	5.9	3	1.5	1.3
AU	40.9	14.2	9	1.5	1.4	56.8	11.2	6	1.2	1.1	57.7	8.2	4	1.3	1.2
KR	31.5	13.0	7	1.4	1.3	41.9	10.7	6	1.3	1.2	40.7	6.8	3	1.4	1.2
PL	32.0	11.9	8	1.4	1.2	34.1	9.3	5	1.3	1.2	38.1	6.1	3	1.3	1.1
IL	46.9	13.2	8	1.2	1.0	54.0	10.4	6.5	1.2	1.1	58.1	6.7	4	1.2	1.1
NL	42.6	13.9	9	1.2	1.1	54.9	11.5	7	1.2	1.1	63.7	6.8	3	1.0	1.0
IN	31.0	12.6	7	1.2	1.2	31.3	10.7	6	1.2	1.1	26.4	6.3	3	1.4	1.3
BR	39.8	13.4	9	1.3	1.2	44.1	10.2	6	1.2	1.1	43.8	6.6	4	1.3	1.3
TW	29.5	13.6	9	1.1	1.1	35.4	10.8	7	1.2	1.1	40.8	6.4	4	1.3	1.3
BE	45.1	14.5	10	1.3	1.1	57.4	11.1	6	1.1	1.1	66.9	7.1	4	1.0	1.0
CH	49.2	17.4	13	1.3	1.2	61.7	11.8	8	1.1	1.0	68.6	8.6	5	1.1	1.0
SE	44.6	12.8	7	1.2	1.1	50.9	9.8	5	1.2	1.0	62.2	6.3	3	1.2	1.1
GR	27.5	14.7	10	1.6	1.5	45.3	9.9	5	1.2	1.1	54.9	7.4	3	1.3	1.2
CZ	37.1	12.5	7	1.7	1.4	44.8	9.8	6	1.4	1.2	39.4	6.2	3	1.6	1.5
AT	40.9	14.4	10	1.4	1.2	59.5	10.7	6	1.1	1.0	62.2	6.7	4	1.1	1.0
TR	21.9	11.8	8	1.1	1.1	25.6	11.7	6	1.2	1.1	33.5	6.9	4	1.4	1.3

Table A.10: Statistics about international collaborations for the main countries

of the country X , or by the total number of international collaborations of the country Y . In the first case, it gives for the country X the relative importance of collaborations with the country Y , while in the second case it gives for the country Y the relative importance of collaborations with the country X . Table A.11 contains this information for 16 countries and for the three years. These three information are provided in each cell, with 2005 on the top, 2010 in the middle, and 2015 on the bottom.

A general trend is visible in this table: For all countries except for China, the ratios of collaborations with the USA slightly decrease. On the other hand, for all countries, USA included, the ratios of collaborations with China increase, even if for most of them, collaborations with the USA are still in much higher numbers compared to the collaborations with China. In that respect, Australia is quite an exception, with a higher percentage of collaborations with China than with the USA. Note that for the USA, the ratio of collaborations with China has doubled during a period of 10 years (from 11.3% to 22.2 %). For other bi-national collaborations, involving a few % of all international collaborations for both countries, a general trend is not clearly visible, and fluctuations are more important (also because the numbers of publications involved are smaller).

	US	CN	FR	DE	IT	UK	CA	JP	ES	RU	AU	KR	PL	IL	NL	IN
US		36.7 37.5 40.1	23.9 21.5 18.8	24.2 22.8 22.7	21.6 22.2 19.9	27.9 22.3 25	37 34.6 32.7	27.2 18.5 22.3	25.1 16.6 15.6	20.7 19.3 18.8	26.9 19.7 19.5	46.9 38.6 33.6	29 20.5 15	54.5 46.3 49.2	28.3 23.6 24.2	36.1 23 22.2
CN	11.3 16.1 22.2		3.7 5.6 6.7	5.3 5.3 6.8	3.3 3.5 3.1	6.4 7.7 10.3	13.5 13 18.7	13.1 18 16.7	2.2 2.9 4.2	0.4 3.1 4.1	15.2 22 26.8	15.3 21.6 19.6	2.3 2.3 5.4	2.4 2.1 6	1.7 3.6 5.6	4.1 6.5 8.3
FR	8.8 9.5 8.2	4.5 5.7 5.3		8.2 10 9.9	16.8 17.9 15.2	7.2 8.8 9.6	8 9.3 10	6.1 6.7 9.5	12.5 12.4 11.7	13.8 13.8 11.4	3.2 7.7 4.8	3.6 5.1 4.2	12.7 10.2 10.8	8 10.1 8.7	7.9 7.1 9.9	7 5.1 4.9
DE	8.4 9.4 9.3	6.1 5.1 5	7.8 9.3 9.2		11.5 11.7 12.1	10.1 11.7 11.6	5.4 5.5 6.2	13.1 8.5 8.8	5.8 8.2 7.5	14.6 12.6 12.2	6.1 8.8 6.1	3.9 4 5.9	13.7 14.7 12.9	7.3 10.1 10	13.6 19.7 19	11.9 12.5 5.3
IT	4.8 5.7 5.7	2.4 2.1 1.6	10.2 10.4 10	7.3 7.3 8.5		5.4 5.8 8.8	2.6 3.7 6.7	4.8 4.4 5	6.4 10.1 8.9	10.1 6.7 8.5	2.7 4.8 3.5	1.6 1.5 2.8	6 5.4 6.9	3.2 4 3.9	5.7 5.7 7.2	4.1 2.9 3
UK	8.2 7.1 8.2	6.1 5.7 6.1	5.8 6.4 7.3	8.5 9.1 9.3	7.2 7.3 10.1		5.5 6.6 7.6	3.7 7.3 6.1	5.5 6.2 6.7	10.7 8.4 8.3	15 11.8 12.8	5.5 4.5 3.7	4.7 6 6.8	4.4 6.4 7.4	5.7 11.2 10.2	7.8 4.5 4.1
CA	10.4 9.3 8	12.4 8.1 8.3	6.1 5.7 5.6	4.3 3.6 3.7	3.2 3.8 5.7	5.2 5.5 5.6		6.1 5.5 4.4	4.6 4.3 3.2	5.7 3.8 2.6	7.8 8.2 5.8	6.2 4.5 3.6	3.3 3.9 5	7.3 9.9 6	7.1 3.3 4.8	7.8 5.6 5.3
JP	3.3 2.9 3.2	5.2 6.5 4.3	2 2.4 3.1	4.6 3.2 3.1	2.6 2.7 2.5	1.5 3.5 2.6	2.6 3.2 2.5		1.6 1.9 1.9	2 4.1 2.6	2.5 2 2.7	11.1 9.8 8.8	5 4.4 4.2	2.2 0.5 1	3.1 3.3 1.4	4.1 4.3 2.1
ES	4.5 3.9 3.6	1.3 1.6 1.8	6.1 6.6 6.2	3 4.7 4.3	5.1 9.2 7.2	3.3 4.5 4.7	2.9 3.7 3	2.4 2.8 3		4.4 3.2 4	3.4 3.5 3.4	2 2.1 3.5	7 7.1 6	2.4 3.5 1.9	3.7 5.5 6.1	2.9 4.7 3.8
RU	3 2.6 2.6	0.2 1 1	5.4 4.2 3.6	6 4.1 4.1	6.6 3.5 4.1	5.2 3.5 3.5	2.9 1.9 1.5	2.4 3.5 2.5	3.6 1.9 2.4		2.8 2 2.4	1.3 0.6 2.6	3.3 5 3	5.1 3.3 4.8	4.2 2.6 2.2	1.2 0.4 1
AU	3.8 2.7 3.1	6.9 7.1 7.8	1.2 2.4 1.8	2.4 3 2.4	1.7 2.6 2	7.1 5.1 6.3	3.9 4.2 3.8	2.8 1.8 3	2.7 2.1 2.4	2.8 2.1 2.8		2 1.5 2.1	1.7 1.7 2	2.2 1.9 3.7	2.3 1.4 1.6	1.6 1.6 2.7
KR	3.8 3.9 3.9	4.1 5 4.1	0.8 1.2 1.1	0.9 1 1.7	0.6 0.6 1.1	1.5 1.4 1.3	1.8 1.7 1.7	7.4 6.3 7.1	0.9 0.9 1.7	0.7 0.4 2.2	1.1 1.1 1.5		0 1 0.9	0.5 0.3 2.1	0.8 0.5 0.5	5.3 5.4 6.2
PL	2.3 1.9 1.7	0.6 0.5 1.1	2.7 2.1 2.8	3.1 3.2 3.6	2.1 1.9 2.7	1.3 1.7 2.3	0.9 1.3 2.3	3.3 2.6 3.3	3.1 2.7 2.9	1.8 3.4 2.5	0.9 1.1 1.4	0	0.9 0.9		1 1.7 2.7	1.1 0.9 2.9
IL	6 5 4.5	0.9 0.5 1	2.4 2.5 1.8	2.3 2.7 2.3	1.6 1.7 1.2	1.6 2.2 2.1	2.8 4 2.2	2 0.4 0.6	1.5 1.6 0.8	3.9 2.7 3.2	1.7 1.5 0.6	0.7 0.4 1.7	1.3 2.1 2.2		4.2 3.6 1.6	1.6 0.7 1.7
NL	2.7 2.6 2.3	0.5 0.9 0.9	2 1.8 2.1	3.7 5.2 4.3	2.4 2.4 2.3	1.8 3.8 2.9	2.4 1.3 1.8	2.4 2.3 0.9	1.9 2.6 2.4	2.8 2.1 1.5	1.5 1.1 2.1	1 0.6 0.4	1.3 2.3 1.4	3.6 3.7 1.6		1.2 1.6 0.7
IN	2.3 1.9 2.7	0.9 1.3 1.8	1.2 1 1.4	2.2 2.6 1.6	1.2 1 1.3	1.7 1.2 1.5	1.8 1.8 2.7	2.2 2.3 1.8	1 1.7 2	0.6 0.3 0.9	0.8 1 2	4.2 4.5 6.6	0 0.8 3.1	1 0.5 2.3	0.8 1.2 1	

Table A.11: Bilateral collaborations: For a country in x -coordinate, the numbers correspond to the % of its international collaborations with a country of the y -coordinate, for 2005, 2010, and 2015

A.7 Conclusion

One of the unexpected outcomes of these investigations is the rapid increase of the number of publications, but also of the significant change of many predictors over a period of 10 years only. Indeed, if we gather some results obtained in (A.1), (A.2), (A.3), (A.4), (A.8), (A.9), (A.11), and in Table A.5, we obtain Table A.12. If we summarize in one sentence the content of this table, it would be that the publications in mathematics are becoming more collaborative, more international, longer, with more references and keywords, more freely accessible, and especially more numerous.

	2005	2010	2015
# publications	45'035	62'945	76'788
# authors	2.10	2.23	2.39
# institutes	1.57	1.74	1.90
# countries	1.23	1.28	1.32
# references	19	21.7	24.5
# pages	15.5	15.5	16.9
# keywords	4.38	4.48	4.58
open access (%)	16.4	21.2	32.3

Table A.12: The increase of many predictors

For the citations, it is quite natural that the Journal Impact Factor plays an important role. In that sense, its appearance as the most important predictor (whenever available) is not surprising. For more general publications, the importance of the number of references is also not so surprising: quite often, a paper with numerous references corresponds to a paper which is well nested in the research landscape, and as a consequence it can be cited by several authors. The importance of other predictors is less clear, and no conclusion can be established for them. At this level, the real content of a paper certainly matters more than any bibliometric predictor.

The content of Section A.6 seems new. From the point of view of the authors, these statistics are maybe not so relevant, since it is difficult to move to another country, just to get more citations. On the other hand, at the level of the scientific policy of countries, knowing the relative ranking of countries with respect to several macro statistics is certainly important. In this direction, our results are only partial and preliminary, but we hope that it can trigger additional investigations in the future.

Bibliography

- [1] Amodio, P. and Brugnano, L., *Recent advances in bibliometric indexes and the PaperRank problem*, Journal of Computational and Applied Mathematics, 267, 182–194, 2014.
- [2] Aria, M. and Cuccurullo, C., *bibliometrix: An R-tool for comprehensive science mapping analysis*, Journal of Informetrics, 11, 959–975, 2017.
- [3] Armstrong, E., Runge, M., and Gerardin, J., *Identifying the measurements required to estimate rates of COVID-19 transmission, infection, and detection, using variational data assimilation*, Infectious Disease Modelling 6, 133–147, 2021.
- [4] Arroyo-Marioli, F., Bullano, F., Kucinskis, S., and C. Rondón-Moreno, *Tracking R of COVID-19: A new real-time estimation using the Kalman filter*, PLoS ONE 16(1): e0244474, 2021.
- [5] Atlas, R., Kalnay, E., Baker, W. E., Susskind, J., Reuter, D., and Halem, M., *Simulation studies of the impact of future observing systems on weather prediction*, Preprints, Seventh Conf. on Numerical Weather Prediction, Montreal, QC, Canada, Amer. Meteor. Soc., 145–151, 1985.
- [6] Behrens, H. and Luksch, P., *Mathematics 1868–2008: a bibliometric analysis* Scientometrics, 86, 179–194, 2011.
- [7] Bensman, S., Smolinsky, L., and Pudovkin, A., *Mean Citation Rate per Article in Mathematics Journals: Differences From the Scientific Model*, Journal of the American Society for information science and technology, 61(7), 1440–1463, 2010.
- [8] Bishop, C. H., Etherton, B. J., and Majumdar, S. J., *Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects*, Mon. Wea. Rev. 129, 420–436, 2001.
- [9] Boccaletti, S., Grebogi, C., Lai, Y.-C., Mancini, H., and Maza, D., *The control of chaos: theory and applications*, Phys. Rep., 329, 103–197, [https://doi.org/10.1016/S0370-1573\(99\)00096-4](https://doi.org/10.1016/S0370-1573(99)00096-4), 2000.
- [10] Brauer, F., Driessche, P.v.d. and Wu, J., *Mathematical epidemiology*, Lecture Notes in Mathematics 1945, Springer, 2008.

- [11] Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J., *Classification and Regression Trees*, CHAPMAN & HALL/CRC, 1984.
- [12] Buitrago-Garcia, D., Egli-Gany, D., Counotte, M. J., Hossmann, S., Imeri, H., Ipekci, A.M., et al., *Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis*, PLoS Med 17(9): e1003346, 2020.
- [13] Bureau of social welfare and public health, *About death cases due to COVID-19 in Tokyo*, <https://www.fukushihoken.metro.tokyo.lg.jp>
- [14] Burgers, G., van Leeuwen P. J., and Evensen, G., *Analysis Scheme in the Ensemble Kalman Filter*, Monthly Weather Review 126.6 (1998): 1719-1724.
- [15] Byambasuren, O., Cardona, M., Bell, K., Clark, J., McLaws, M.-L., and Glasziou, P., *Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis*, J. Association of Medical Microbiology and Infectious Disease Canada 5 Issue 4, 223–234, 2020.
- [16] De Battisti, F. and Salini, S., *Robust analysis of bibliometric data*, Stat. Methods Appl., 22, 269–283, 2013.
- [17] Didegah, F. and Thelwall, M., *Which factors help authors produce the highest impact research? Collaboration, journal and document properties*, Journal of Informetrics, 7, 861–873, 2013.
- [18] Driessche P.v.d. and Watmough, J., *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission*, Mathematical Biosciences 180, 29–48, 2002.
- [19] Dunne, E., *Don't count on it*, Notices Amer. Math. Soc., 68 no. 1, 114–118, 2021.
- [20] Durrett, R., *Probability: Theory and Examples*, Fifth edition, Cambridge Series in Statistical and Probabilistic Mathematics 49, Cambridge University Press, Cambridge, 2019.
- [21] Evensen, G., *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, J. Geophys. Res., 99, 10143-10162, 1994.
- [22] Evensen, G., *Advanced data assimilation for strongly nonlinear dynamics*, Mon. Weather Rev., 125, 1342–1354, 1997.
- [23] Evensen, G., *The ensemble Kalman filter: Theoretical formulation and practical implementation*, Ocean Dyn., 53, 343-367, 2003.
- [24] Evensen, G., *Data Assimilation: The Ensemble Kalman Filter*, Springer, 2006.

- [25] Evensen, G., Amezcuca, J., Bocquet, M., Carrassi, A., Farchi, A., Fowler, A., Houtekamer, P. L., Jones, C. K., de Moraes, R. J., Pulido, M., Sampson, C., and Vossepoel, F. C., *An international initiative of predicting the SARS-CoV-2 pandemic using ensemble data assimilation*, Foundations of Data Science, American Institute of Mathematical Sciences, 2020.
- [26] Engbert, R., Rabe, M. M., Kliegl, R., and Reich, S., *Sequential Data Assimilation of the Stochastic SEIR Epidemic Model for Regional COVID-19 Dynamics*, Bulletin of Mathematical Biology 83:1, 2021.
- [27] Evans, E., Bhatti, N., Kinney, J., Pann, L., Peña, M., Yang, S. C., Kalnay, E., and Hansen, J., *RISE undergraduates find that regime changes in Lorenz’s model are predictable*, B. Am. Meteorol. Soc., 85, 521–524, 2004.
- [28] Flossmann, A. I., Manton, M., Abshaev, A., Bruintjes, R., Murakami, M., Prabhakaran, T., and Yao, Z., *Review of advances in precipitation enhancement research*, Bull. Am. Meteorol. Soc., 100, 1465–1480, <https://doi.org/10.1175/BAMS-D-18-0160.1>, 2019.
- [29] Goldfinch, S., Dale, T., and DeRouen, K., *Science from the periphery: Collaboration, networks and ‘Periphery Effects’ in the citation of New Zealand Crown Research Institutes articles, 1995–2000*, Scientometrics, 57(3), 321–337, 2003.
- [30] Gostic, K.M., McGough, L., Baskerville, E. B., Abbott, S., Joshi, K., Tedijanto, C., et al., *Practical considerations for measuring the effective reproductive number, R_t* , PLoS Comput Biol 16 No. 12, e1008409, 21 pages, 2020.
- [31] Ghostine, R., Gharamti, M., Hassrouny, S., and Hoteit, I., *An Extended SEIR Model with Vaccination for Forecasting the COVID-19 Pandemic in Saudi Arabia Using an Ensemble Kalman Filter*, Mathematics 9, 636. 2021.
- [32] Grossman, J. W., *The evolution of the mathematical research collaboration graph in Proceedings of the Thirty-third Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 2002)*, Congr. Numer., 158, 201–212., 2002.
- [33] Grossman, J. W., *Patterns of research in mathematics*, Notices Amer. Math. Soc., 52 no. 1, 35–41, 2005.
- [34] Henderson, J. M., Hoffman, R. N., Leidner, S. M., Nehrkorn, T., and Grassotti, C., *A 4D-VAR study on the potential of weather control and exigent weather forecasting*, Q.J R. Meteorol. Soc. 131, 3037–3052, Oct. (Part B) 2005.
- [35] Hethcote, H. W., *The mathematics of infectious diseases*, SIAM Rev. 42 No. 4, 599–653, 2000.

- [36] Hoffman, R. N., *Controlling the global weather*, Bulletin of the American Meteorological Society 83(2), 241–248, 2002.
- [37] Hoffman, R. N., *Controlling hurricanes. Can hurricanes and other severe tropical storms be moderated or deflected?* Scientific American, 291(4), 68–75, Oct. 2004.
- [38] Hoffmann, R. S. and Atlas, R., *Future observing system simulation experiments*, B. Am. Meteorol. Soc., 97, 1601–1616, <https://doi.org/10.1175/BAMS-D-15-00200.1>, 2016.
- [39] Houtekamer, P. L. and Mitchell, H. L., *Data Assimilation Using an Ensemble Kalman Filter Technique*, Monthly Weather Review, 126(3), 796–811, 1998.
- [40] Houtekamer, P. L. and Mitchell, H. L., *A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation*, Monthly Weather Review 129(1), 123–137, 2001.
- [41] Houtekamer, P. L. and Zhang, F., *Review of the ensemble Kalman filter for atmospheric data assimilation*, Monthly weather review, vol 144, 4489–4532, 2016.
- [42] Hunt, B. R., Kostelich, E. J., and Szunyogh, I., *Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter*, Physica D. 230, 112–126, 2007.
- [43] Kalman, R. E. *A new approach to linear filtering and prediction problems*, Trans. ASME Ser. D: J. Basic Eng. 82, 35–45, 1960.
- [44] Kalman, R. E. and Bucy, R. S., *New results in linear filtering and prediction theory*, Trans. ASME Ser. D: J. Basic Eng. 83, 95–108, 1961.
- [45] Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C., and Ballabrera-Poy, J., *4D-Var or Ensemble Kalman Filter?*, Tellus, 59A, 758–773, 2007.
- [46] Kuniya, T., Inaba, H., *Possible effects of mixed prevention strategy for COVID-19 epidemic: massive testing, quarantine and social distancing*, AIMS Public Health 7 No 3, 490–503, 2020.
- [47] Law, K., Stuart, A., and Zygalakis, K., *Data Assimilation: A Mathematical Introduction*. A mathematical introduction. Texts in Applied Mathematics 62, Springer, Cham, 2015.
- [48] Li, M., *An Introduction to Mathematical Modeling of Infectious Diseases*, Mathematics of Planet Earth 2, Springer 2018.
- [49] Li, R. et al., *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)*, Science 368, 489–493, 2020.

- [50] Li Y., Kalnay, E., Matesharrei, S., Rivas, J., Kucharski, F., Kirk-Davidoff, D., Bach, E., and Zeng, N., *Climate model shows large-scale wind and solar farms in the Sahara increase rain and vegetation*, Science, 361, 1019–1022, 2018.
- [51] Lorenz, E. N., *Deterministic nonperiodic flow*, J. Atmos. Sci., 20, 130–141, 1963.
- [52] Lorenz, E. N., *The Essence of Chaos*, University of Washington Press, Seattle, Washington, United States, 227 pp., ISBN 9780295975146, 1993.
- [53] Lorenz, E. N., *Predictability - A problem partly solved*, Seminar on Predictability, ECMWF, 1995.
- [54] Lorenz, E. N. and Emmanuel, K., *Optimal sites for supplementary weather observations: simulation with a small model*, Journal of the Atmospheric Science 55, 399–414, 1998.
- [55] Luo, J., Flynn, J. M., Solnick, R. E., Ecklund, E. H., and Matthews, K. R. W., *International Stem Cell Collaboration: How Disparate Policies between the United States and the United Kingdom Impact Research*, PLoS ONE, 6(3), e17684, 2011.
- [56] m3: <https://www.m3.com/open/iryoIshin/article/849820/>
- [57] MLIT: https://www.mlit.go.jp/tetudo/tetudo_fr1_000062.html
- [58] MacMynowski, D. G., *Controlling chaos in El Niño*, Proceedings of the 2010 American Control Conference, pp. 4090-4094, 2010.
- [59] McAloon, C., Collins, A., Hunt, K., et al., *Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research*, BMJ Open 10:e039652, 2020.
- [60] Miller, R., Ghil, M., and Gauthiez, F., *Advanced data assimilation in strongly nonlinear dynamical systems*, J. Atmos. Sci., 51, 1037–1056, 1994.
- [61] Mitchell, L. and Arnold, A., *Analyzing the effects of observation function selection in ensemble Kalman filtering for epidemic models*, Mathematical Biosciences 339, 2021.
- [62] Miyoshi, T. and Sun, Q., *Control Simulation Experiment (CSE) with the Lorenz-63 model*, TIB AV-Portal [video], <https://doi.org/10.5446/54893>, 2021.
- [63] Miyoshi, T. and Sun, Q., *Control simulation experiment with Lorenz’s butterfly attractor*, Nonlin. Processes Geophys. 29, 133–139, 2022.
- [64] Nadler, P., Wang, S., Arcucci, R., Yang, X., and Guo, Y., *An epidemiological modelling approach for COVID-19 via data assimilation*, European Journal of Epidemiology 35, 749–761, 2020.
- [65] Nishiura, H.: <https://github.com/contactmodel/COVID19-Japan-Reff>

- [66] Osaka prefecture government, *Citizens awareness and behavior change of measures against COVID-19*, http://www.pref.osaka.lg.jp/hodo/attach/hodo-40479_4.pdf
- [67] Ott, E., Hunt, B. R., Szunyogh, I., Corazza, M., Kalnay, E., Patil, D. J., Yorke, J. A., Zimin, A. V., and Kostelich, E. J., *Exploiting local low dimensionality of the atmospheric dynamics for efficient ensemble Kalman filtering*, Preprint: <https://arxiv.org/abs/physics/0203058v3>.
- [68] Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A., *A local ensemble Kalman filter for atmospheric data assimilation*, *Tellus A* 56, 415–428, 2004.
- [69] Özkaya, A., *Bibliometric analysis of the studies in the field of mathematics education*, *Educational Research and Reviews*, 13(22), 723–734, 2018.
- [70] Paik, J. and Rivin, I., *Bibliometric Analysis of Senior US Mathematics Faculty*, preprint arXiv:2008.11196, 2020.
- [71] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., *Scikit-learn: Machine Learning in Python*, *Journal of Machine Learning Research*, 12, 2825–2830, 2011.
- [72] Pollock, A. M. and Lancaster, J., *Asymptomatic transmission of covid-19*, *BMJ* 371:m4851, 2020.
- [73] Rebollo, T. C. and Coronil, D., *Predictive data assimilation through Reduced Order Modeling for epidemics with data uncertainty*, Preprint: <https://arxiv.org/abs/2004.12341>.
- [74] Rhodes, C. J. and Hollingsworth, T. D., *Variational data assimilation with epidemic models*, *Journal of Theoretical Biology* 258, 591–602, 2009.
- [75] Rusin, D., *A Gentle Introduction to the Mathematics Subject Classification Scheme*, link provided by https://en.wikipedia.org/wiki/Mathematics_Subject_Classification, 2015.
- [76] Särkkä, S., *Bayesian Filtering and Smoothing*, Institute of Mathematical Statistics Textbooks 3, Cambridge University Press, Cambridge, 2013.
- [77] *Scikit-learn.org*, <https://scikit-learn.org/stable/modules/tree.html>
- [78] Smith, M. J., Weinberger, C., Bruna, E. M., and Allesina, S., *The Scientific Impact of Nations: Journal Placement and Citation Performance*, *PLoS ONE*, 9(10), e109195, 2014.

- [79] Soldatenko, S. and Yusupov, R., *An Optimal Control Perspective on Weather and Climate Modification*, Mathematics 9, 305, 2021.
- [80] Sooryamoorthy, R., *Do types of collaboration change citation? Collaboration and citation patterns of South African science publications*, Scientometrics, 81(1), 177–193, 2009.
- [81] Sooryamoorthy, R., *Do types of collaboration change citation? A scientometric analysis of social science publications in South Africa*, Scientometrics, 111, 379–400, 2017.
- [82] Szomszor, M., Pendlebury, D. A., and Adams, J., *How much is too much? The difference between research influence and self-citation excess*, Scientometrics, 123, 1119–1147, 2020.
- [83] Sun C., Richard S., Miyoshi T., and Tsuzu N., *Analysis of COVID-19 Spread in Tokyo through an Agent-Based Model with Data Assimilation*, Journal of Clinical Medicine, 11(9):2401, 2022.
- [84] Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., Whitaker, J. S., *Ensemble square-root filters*, Mon. Wea. Rev. 131, 1485–1490, 2003.
- [85] Tokyo metropolitan government, *COVID-19 The information website*, <https://stopcovid19.metro.tokyo.lg.jp>.
- [86] Toth, Z. and Kalnay, E., *Ensemble Forecasting at NMC: The Generation of Perturbations*, B. Am. Meteorol. Soc., 74, 2317–2330, [https://doi.org/10.1175/1520-0477\(1993\)074<2317:EFANTG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2), 1993.
- [87] Toyokeizai, *Coronavirus disease (COVID-19) situation report in Japan*, <https://toyokeizai.net/sp/visual/tko/covid19/en.html>.
- [88] Van der Hoek, J. and Elliott, R. J., *Introduction to Hidden Semi-Markov Models*, London Mathematical Society Lecture Note Series 445, Cambridge University Press, Cambridge, 2018.
- [89] Verma, R., Lobos-Ossandón, V., Merigó, J.M., Cancino, C., and Sienz, J., *Forty years of applied mathematical modelling: A bibliometric study*, Applied Mathematical Modelling, 89, 1177–1197, 2021.
- [90] Wagner, C., Whetsell, T., and Leydesdorff, L., *Growth of international collaboration in science: revisiting six specialties*, Scientometrics, 110, 1633–1652, 2017.
- [91] Wang, T., Peng, Y., Zhang, B., et al., *Move a Tropical Cyclone with 4D-Var and Vortex Dynamical Initialization in WRF Model*, Journal of Tropical Meteorology 27(3), 191–200, 2021.

- [92] Wang, L., Thijs, B., and Glänzel, W., *Characteristics of international collaboration in sport sciences publications and its influence on citation impact*, *Scientometrics*, 105, 843–862, 2015
- [93] Wang, M., Zhang, J., Jiao, S., and Zhang, T., *Evaluating the impact of citations of articles based on knowledge flow patterns hidden in the citations*, *PLoS ONE*, 14(11), e0225276, 2019.
- [94] Whitaker, J. S. and Hamill, T. M., *Ensemble data assimilation without perturbed observations*, *Mon. Wea. Rev.* 130, 1913–1924, 2002.
- [95] World Health Organization: <https://www.who.int/>.
- [96] Yang, S.-C., Baker, D., Li, H., Cordes, K., Huff, M., Nagpal, G., Okereke, E., Villafañe, J., Kalnay, E., and Duane, G. S., *Data assimilation as synchronization of truth and model: experiments with the three-variable Lorenz system*, *J. Atmos. Sci.*, 63, 2340–2354, <https://doi.org/10.1175/JAS3739.1>, 2006.
- [97] Yang, S.-C., Kalnay, E., and Hunt, B., *Handling nonlinearity in an ensemble Kalman filter: experiments with the three-variable Lorenz model*, *Mon. Weather Rev.*, 140, 2628–2646, 2012.