





Université de Montréal

Prédiction des éruptions solaires par assimilation de  
données avec les modèles d'avalanches

par

**Christian Thibeault**

Département de Physique  
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales  
en vue de l'obtention du grade de  
Maître ès sciences (M.Sc.)  
en Physique

26 août 2020



# Université de Montréal

Faculté des études supérieures et postdoctorales

Ce mémoire intitulé

## Prédiction des éruptions solaires par assimilation de données avec les modèles d'avalanches

présenté par

**Christian Thibeault**

a été évalué par un jury composé des personnes suivantes :

*Pierre Bergeron*

---

(président-rapporteur)

*Paul Charbonneau*

---

(directeur de recherche)

*Antoine Strugarek*

---

(co-directeur)

*Yashar Hezaveh*

---

(membre du jury)

Mémoire accepté le :

---



# Sommaire

---

Les éruptions solaires sont des tempêtes de rayonnement électromagnétique, de particules relativistes et parfois de masse coronale provoquées par la libération d'énergie magnétique provenant de la couronne solaire. Si ces tempêtes atteignent l'environnement terrestre, elles peuvent poser un danger à la santé des astronautes en hautes orbites et causer des perturbations importantes sur les systèmes GPS. Dans certains cas, elles peuvent même induire des dommages aux infrastructures technologiques, dont les réseaux électriques.

La prédiction des éruptions solaires est donc considérée comme un des plus importants défis de la météorologie spatiale. Par contre, à ce jour, aucune méthode présentée dans la littérature n'est capable de produire des prédictions fiables, ce qui met en évidence la nature complexe du déclenchement des éruptions solaires. Nous présentons donc dans ce mémoire une méthode alternative aux techniques statistiques habituelles, basée sur l'assimilation de données couplée avec des modèles rapides en automate cellulaire appelés *modèles d'avalanche*.

Les modèles d'avalanche sont une simplification drastique de la physique du déclenchement des éruptions solaires. Malgré leur simplicité, ils reproduisent assez bien les statistiques à long terme de la taille des éruptions. Nous présentons dans ce mémoire des analyses empiriques de la capacité prédictive de quatre modèles : le modèle de Lu et Hamilton (LH) (Lu & Hamilton, 1991, ApJ, 412, 841-852), deux modèles à forçage déterministes (D) (Strugarek & Charbonneau, 2014, SoPh, 289(8), 2993-3015) et finalement deux modèles maximisant l'énergie libérée, appelées modèles DMC, qui sont fortement inspirés du modèle présenté par Farhang et al. (2018, ApJ, 859(1), 41). Les modèles DMC ont été développés dans le cadre de cette maîtrise et donc un chapitre de ce mémoire est dédié à leur présentation et aux analyses plus détaillées de leurs caractéristiques. Nous montrons que pour les modèles D ainsi que les modèles DMC, une mémoire existe entre les évènements simulés de grandes tailles, malgré la forte stochasticité de chacun de ces modèles.

Nous présentons de plus dans ce mémoire un nouveau protocole de prédiction des éruptions solaires, utilisant l'assimilation de données couplée avec les modèles d'avalanches. Notre protocole se base sur une méthode de *recuit simulé* pour ajuster la condition initiale du modèle jusqu'à ce qu'elle reproduise de façon satisfaisante une série d'évènements observés. Une fois cette condition initiale optimisée produite, la simulation qui en résulte représente notre prédiction. Nous montrons dans ce mémoire le succès de notre protocole à bien assimiler une centaine d'observations synthétiques (produit par les modèles d'avalanche eux-mêmes).

**Mots-clés :** Éruptions solaires, Modèles d'avalanche, Assimilation de données, Prédiction, Systèmes complexes



# Summary

---

Solar flares are sudden releases of electromagnetic radiation, relativistic particles and occasionally coronal mass, caused by the release of magnetic energy from the solar corona. They pose a danger to astronauts in high orbits and directly impact the Earth, including significant disturbances on GPS systems, and can even cause damage to technological infrastructures, including electrical networks. Predicting solar flares is therefore considered to be one of the most critical challenges in space weather. However, no method presented in the literature can produce reliable predictions, highlighting the complex nature of the triggering of solar flares. We, therefore, present in this thesis an alternative method to the usual statistical forecasting techniques. Our method is based on data assimilation coupled with computationally inexpensive cellular automaton models called *avalanche models*.

Avalanche models are a drastic simplification of the physics underlying the triggering of solar flares. Despite their simplicity, they reproduce reasonably well the long-term statistics of solar flares sizes. In this thesis, we present empirical analyses of the predictive capabilities of four models: the Lu and Hamilton (LH) model (Lu & Hamilton, 1991, ApJ, 412, 841-852), two deterministic-driven (D) models (Strugarek & Charbonneau, 2014, SoPh, 289(8), 2993-3015) and finally two models using the principle of minimum energy during magnetic reconnection, called DMC models, which are strongly inspired by the models presented by Farhang et al. (2018, ApJ, 859(1), 41). The DMC models were developed during this project; therefore, a chapter of this thesis is dedicated to their presentation and more detailed analyses of their characteristics. We show that for D and DMC models, a memory exists between large simulated events, despite the high stochasticity present within each of these models.

We finally present in this thesis a new protocol for predicting solar flares, using data assimilation coupled with avalanche models. Our protocol is based on a *simulated annealing*

method to adjust the initial condition of the model until it satisfactorily reproduces a series of observed events. Once this optimal initial condition is found, the resulting simulation produces our prediction. In this thesis, we show our algorithm's success in assimilating hundreds of synthetic observations (produced by the avalanche models themselves).

**Keywords:** Solar flares, Avalanche models, Data assimilation, Forecasting, Complex systems

# Table des matières

---

<b>Sommaire</b> .....	v
<b>Summary</b> .....	vii
<b>Liste des tableaux</b> .....	xiii
<b>Table des figures</b> .....	xv
<b>Liste des sigles et des abréviations</b> .....	xix
<b>Remerciements</b> .....	xxi
<b>Chapitre 1. Introduction</b> .....	1
1.1. Éruptions solaires .....	1
1.1.1. Observations des éruptions solaires .....	2
1.1.2. Statistiques des éruptions solaires .....	5
1.1.3. Modèles physiques des éruptions solaires .....	6
1.1.3.1. Notions de base de la magnétohydrodynamique .....	6
1.1.3.2. Dissipation passive du champ magnétique coronal .....	9
1.1.3.3. La reconnexion magnétique .....	10
1.1.3.4. Modèle de Sweet-Parker .....	10
1.1.3.5. Modèle de nanoflares de Parker .....	13
1.1.4. Simulations numériques des éruptions solaires .....	14
1.1.5. Les éruptions solaires comme phénomène de criticalité auto-organisée .....	15
1.2. Modèles d'avalanches .....	16
1.2.1. Modèle de Lu et Hamilton 1993 (LH) .....	18
1.2.2. Modèles déterministes (D) .....	20

1.2.3.	Modèles déterministes avec minimisation Monte-Carlo (DMC) .....	21
1.3.	Assimilation de données .....	22
1.3.1.	Base théorique de l'assimilation de données .....	22
1.3.1.1.	3D-Var .....	24
1.3.1.2.	4D-Var .....	25
1.3.1.3.	Recuit simulé .....	26
1.4.	Prédiction des éruptions solaires .....	27
1.5.	Contributions .....	30
<b>Chapitre 2.</b>	<b>Modèles déterministes à minimisation Monte-Carlo (DMC) ..</b>	<b>33</b>
2.1.	Introduction .....	33
2.2.	Présentation du modèle .....	34
2.3.	Caractéristiques des avalanches .....	36
2.4.	Statistiques des avalanches .....	37
2.5.	Capacités prédictives .....	40
2.6.	Conclusion .....	44
<b>Chapter 3.</b>	<b>Article: Predicting Solar Flares by Data Assimilation in Avalanche Models. II. Efficient assimilation and predictive potential.....</b>	<b>47</b>
	Abstract .....	47
3.1.	Introduction .....	48
3.2.	Avalanche models for solar flares .....	50
3.2.1.	The Lu and Hamilton Model .....	51
3.2.2.	The Deterministically Driven Sandpile Model .....	52
3.2.3.	Avalanching properties of LH and D models .....	54

3.3. Predictive capabilities of avalanche models .....	55
3.3.1. Robustness of large avalanches .....	55
3.3.2. Short-term correlations between large events .....	57
3.4. Data assimilation .....	60
3.4.1. Data assimilation Procedure .....	61
3.4.1.1. Data to assimilate .....	61
3.4.1.2. Comparing two sets of discrete events: design of an efficient cost function	61
3.4.1.3. Minimization of the cost function .....	63
3.4.2. Results .....	64
3.5. Discussion and Conclusions .....	67
<b>Chapitre 4. Conclusion</b> .....	<b>71</b>
<b>Bibliographie</b> .....	<b>77</b>



## Liste des tableaux

---

1.1	Synthèse des résultats de prédiction 'all-clear' des éruptions de taille $\geq$ M1.0 pendant 12 h, tirée de l'article résumant l'atelier 'all-clear' à Boulder : Barnes et al. 2016. ....	29
-----	---	----





# Table des figures

---

1.1	Deux observations d'une éruption solaire de grande taille. ....	2
1.2	Trois images du Soleil prises le 24 octobre 2013 (minuit - Temps universel) par le <i>Solar Dynamics Observatory</i> (SDO). ....	4
1.3	Image dans la bande 195 Å d'une boucle coronale (TRACE/NASA). ....	4
1.4	Distributions des intensités maximales des éruptions solaires pendant trois cycles solaires, tirées de Aschwanden 2011. ....	6
1.5	Schéma d'un processus de reconnexion magnétique. ....	11
1.6	Représentation schématique du modèle de Sweet-Parker. ....	12
1.7	Représentation schématique du modèle de Parker des nanoflares. ....	14
1.8	Réseau cartésien à deux dimensions utilisé pour les modèles d'avalanche. Un nœud particulier $A_{i,j}$ (en rouge) interagit avec ses quatre voisins les plus près (en bleu). ....	17
1.9	Représentation schématique d'une boucle coronale avec la définition de nos axes de $z$ et $\phi$ . ....	17
1.10	Représentation schématique de l'analyse $x_a$ comme étant une distribution plus compacte et quelque part entre l'observation $y$ et l'état-modèle $x_b$ . ....	23
2.1	Exemple de l'énergie du réseau (Haut) et l'énergie libérée (bas) pour le modèle DMC1. ....	37
2.2	Série temporelle et des images de l'énergie libérée durant une avalanche de grande taille avec le modèle DMC1. ....	38
2.3	Statistiques des avalanches des modèles DMC1 et DMC2. ....	41

2.4	Distributions de l'énergie totale libérée ( $E/e_0$ ) du modèle DMC1 à différentes tailles de réseau, avec $N_{min} = 25$ et pour $10^7$ itération chacun. Les petites bosses aux parties droite sont des artefacts numériques causés par les effets de bords....	42
2.5	Distributions de l'énergie totale libérée ( $E/e_0$ ) du modèle DMC1 pour différent nombre d'itérations de minimisation $N_{min}$ , avec taille de réseau $48 \times 48$ et pour $10^7$ itération chacun.....	42
2.6	Distributions des itérations de déclenchement des avalanches de grandes tailles ( $E/e_0 > 10^5$ ), pour les modèles DMC1 (haut) et DMC2 (bas) à $48 \times 48$ noeuds et avec $N_{min} = 25$ .....	44
3.1	Top: Energy release time series for model LH (left) and D01 (right). Bottom: Compression of the same energy release time series, where each line represents the total energy released during the avalanche.....	55
3.2	Top: Waiting time distributions between events of all sizes. Bottom: Waiting time distribution between events over $10^4$ for Model LH and over $10^5$ for Model D01.....	56
3.3	Top: Probability density function (PDF) of energy released from 1000 avalanches produced from the same initial condition, but with varying random-number sequences. Bottom: Relative standard deviations of 1000 distributions of energy released ( $E$ ) such as the one presented in the top panel of this figure, with respect to the average energy released of each distribution.....	58
3.4	Distribution of the timing of large events (over $10^4$ for the LH model and over $10^5$ for both D models) from 10 000 models runs, each with identical initial conditions but with varying random-number sequences. ....	59
3.5	Illustration of our definition of a match, miss and false alarm used in our cost function between the reference (shown in black) and model data (shown in orange).	62
3.6	Evolution of the cost function during 100 simulated annealing algorithm runs on Model D01 on a single synthetic observation.....	65

3.7	Event energy time series of the best (top, blue) and worst (bottom, red) output of 100 data assimilation runs using Model D01 on a common synthetic data. ....	66
3.8	Top: Distribution of results of 10 000 data assimilation runs, consisting of 100 runs on each of 100 synthetic observations. Bottom: cumulative distribution of the best results ( $J_{min}$ ) on each of the 100 synthetic observations.....	68
3.9	Distribution of the timing of large avalanches ( $< 10^5$ ) produced by 100 assimilated (in red) and non-assimilated (in blue) initial conditions from model D01.....	70
4.1	Distributions du temps de déclenchement des avalanches produit par 100 conditions initiales assimilées (rouge) et non assimilées (bleu) avec le modèle D01.....	73



## Liste des sigles et des abréviations

---

GOES	Geostationary Operational Environmental Satellite
SOHO	Solar and Heliospheric Observatory
NOAA	National Oceanic and Atmospheric Administration
HMI	Helioseismic and Magnetic Imager
TRACE	Transition Region and Coronal Explorer
NASA	The National Aeronautics and Space Administration
MHD	Magnétohydrodynamique
SDO	Solar Dynamics Observatory
SOC	Criticalité auto-organisée
LH	Modèle de Lu et Hamilton
D	Modèle à forçage déterministe
DMC	Modèle à forçage déterministe avec minisation Monte-Carlo



# Remerciements

---

Je tiens premièrement à remercier mon directeur de recherche, Paul Charbonneau, de m'avoir guidé pendant tout ce projet de maîtrise. J'ai eu la chance de grandement bénéficier de son expertise, sa patience, sa curiosité et son professionnalisme scientifique. J'ai bien hâte de travailler avec lui dans de futurs projets.

Je tiens aussi à remercier mon co-directeur, Antoine Strugarek, pour toute son aide au niveau scientifique ainsi que numérique. Ce projet de recherche fut beaucoup plus intéressant par ses conseils et ses travaux impressionnants.

Merci à Benoit Tremblay pour son aide, particulièrement durant les premières parties de ma maîtrise où je devais préparer des présentations orales, affiches scientifiques, etc.

Merci à Carter Rhea, qui non seulement fut un très bon modèle à suivre d'un bon étudiant à la maîtrise, mais aussi un très bon ami. Notre amitié fut définitivement une aide importante dans ma motivation d'aller au bureau chaque jour.

Finalement, merci à mes parents, Julia et Raymond Thibeault, pour leur soutien continu durant toutes mes années d'études. Un merci aussi à mon frère, Stéphane Thibeault, pour m'avoir transmis dès l'adolescence son amour pour la science et les mathématiques.





# Chapitre 1

---

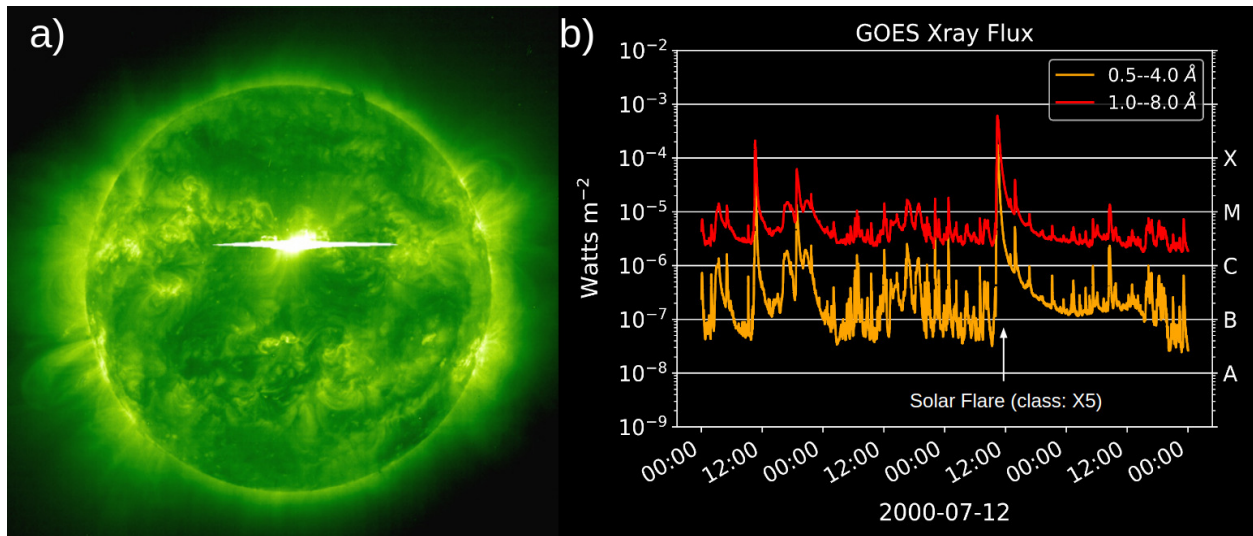
## Introduction

Le Soleil est une étoile de masse intermédiaire qui se distingue par sa grande proximité à la Terre. Cette proximité nous permet de bénéficier du Soleil comme source principale d'énergie sur Terre. Par contre, elle nous rend également vulnérables à des phénomènes éruptifs qui se déclenchent de manière intermittente dans son atmosphère étendue. Ces phénomènes éruptifs sont à l'origine de la *météorologie spatiale*, une branche de la physique spatiale dédiée à l'étude de l'impact des événements éruptifs et transitoires sur l'environnement géospatial.

Comme la plupart des étoiles, le soleil possède un champ magnétique maintenu par des écoulements plasma intérieurs. Ce champ magnétique est la source principale de *l'activité solaire* et des phénomènes éruptifs qui se produisent à la surface. Parmi ces phénomènes, les *éruptions solaires* sont d'une grande importance. Elles ont des impacts parfois violents sur l'environnement géospatial et sur la haute atmosphère terrestre. Leur mécanisme de déclenchement est encore aujourd'hui un phénomène mystérieux.

### 1.1. Éruptions solaires

Les éruptions solaires sont caractérisées par une libération soudaine d'énergie emmagasinée dans la couronne solaire. Cette libération cause un réchauffement très intense et spatialement localisé dans la couronne ainsi qu'une accélération importante des particules environnantes, y élevant brusquement les températures jusqu'à environ  $10^7$  K. Le tout émet une grande quantité de radiation en rayon X et dans l'ultraviolet, ainsi que des particules relativistes. Un exemple d'une éruption solaire de grande taille observée en extrême ultraviolet est présenté à la figure 1.1a. Elles posent un danger pour les astronautes et ont des impacts directs sur la Terre, dont des perturbations importantes sur les communications



**FIGURE 1.1.** a) Image en extrême ultraviolet (195 Å) d'une éruption solaire de classe X le 14 juillet 2000. Cette image en domaine public provient du "Solar and Heliospheric Observatory" (SOHO) (<https://sohowww.nascom.nasa.gov/gallery/top30/>). b) Séquence temporelle des émissions en rayons X pendant quatre jours (du 12 au 15 juillet 2000 inclusivement) mesurée par le réseau de satellites GOES (NOAA) (domaine public).

satellites (incluant les systèmes GPS), et peuvent induire des dommages aux infrastructures technologiques, dont les réseaux électriques. Pour cette raison, les éruptions solaires ont été d'un grand intérêt pour les physiciens solaires depuis leur découverte en 1859 par R. C. Carrington & R. Hodgson [1]. Leur prédiction est considérée comme un des plus grands défis de la météorologie spatiale.

### 1.1.1. Observations des éruptions solaires

La façon standard de détecter des éruptions solaires est par la mesure de pics d'intensité de radiation à courte longueur d'onde. La figure 1.1b présente une période particulièrement impressionnante d'éruptions solaires captées par les satellites *Geostationary Operational Environmental Satellite* (GOES) de la NOAA.

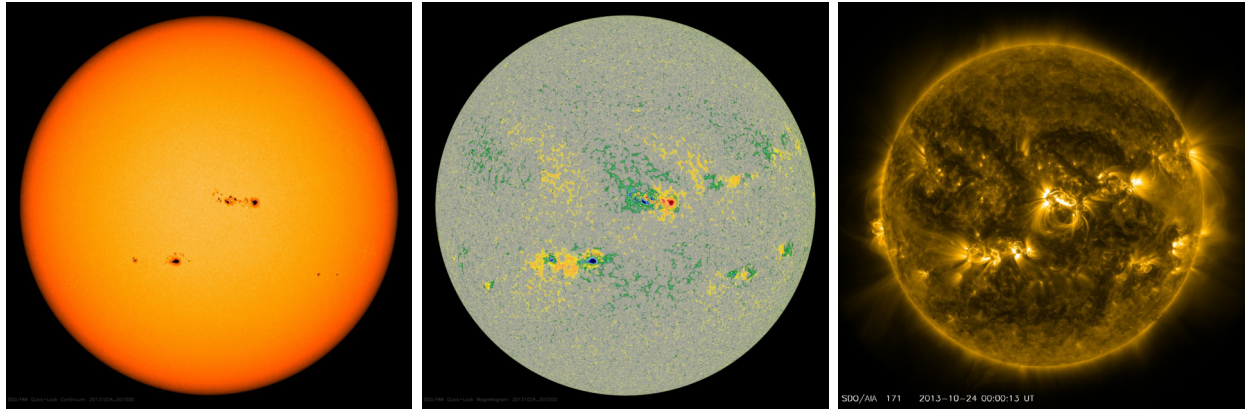
Vers midi le 14 juillet 2000, l'intensité mesurée dans la bande 1.0 – 8.0 Å augmente de 2 à 3 ordres de grandeur en à peine 10 à 100 secondes. Cette augmentation soudaine est la manifestation d'une éruption solaire de très grande taille. La figure 1.1a présente une image en extrême ultraviolet de cette même éruption. On remarque que cet évènement extrême se produit clairement dans une région spécifique de la couronne solaire. Les autres pics visibles sur la figure 1.1b représentent également des éruptions solaires. Un aspect impressionnant

des éruptions solaires est la grande différence d'échelle temporelle entre l'accumulation et la libération de l'énergie. En fait, l'accumulation se produit sur quelques jours, tandis que la libération se produit soudainement en quelques secondes. Cette caractéristique contraint particulièrement les mécanismes physiques responsables de leur déclenchement.

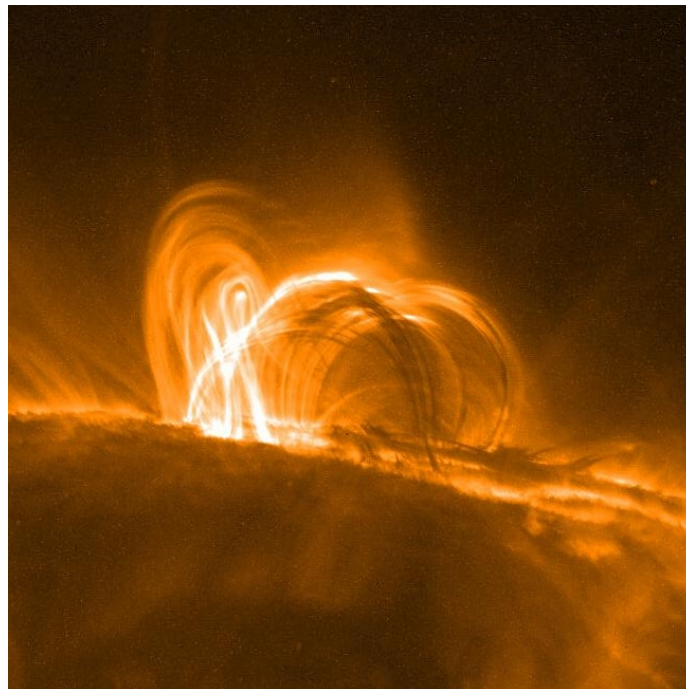
On divise les éruptions solaires en 5 types selon l'ordre de grandeur du flux maximal mesuré dans la bande 1-8 Å (en rouge dans la figure 1.1b). Donc A =  $10^{-8}$ – $10^{-7}$ , B =  $10^{-7}$ – $10^{-6}$ , C =  $10^{-6}$ – $10^{-5}$ , M =  $10^{-5}$ – $10^{-4}$  et X >  $10^{-4}$  W m<sup>-2</sup>. On ajoute à ces lettres un chiffre de 0 à 9 pour sous-diviser chacune de ces catégories en 10 parties linéairement égales. Par exemple, une éruption de classe M4 correspond à un flux entre 4 et  $5 \times 10^{-5}$  W m<sup>-2</sup>.

La grande majorité des éruptions solaires se produisent dans des *régions actives*. Les régions actives sont des régions fortement magnétisées de la surface du Soleil, facilement localisables par la présence de taches sombres à la photosphère, appelées *taches solaires*. L'image de gauche de la figure 1.2 montre une image du Soleil dans le visible, contenant quelques taches solaires. Les taches sont en paires horizontales avec des tailles variables. Au centre de la figure 1.2, on voit un magnétogramme du Soleil durant la même période. On remarque qu'il y a une coïncidence spatiale très évidente entre la position des taches solaires et les régions fortement magnétisées. De plus, les paires de taches solaires constituent des paires de pôles magnétiques opposés. À la droite de la figure 1.2, on voit une image à 171 Å du Soleil. Des structures en forme de filaments sont visibles, émanant des régions actives. Ces structures sont en fait du plasma très chaud confiné par les lignes de champ magnétique entrant et sortant de la photosphère. Ces filaments forment habituellement des boucles, appelées *boucles coronales*, ancrées sur la photosphère. Une image en ultraviolet de boucles coronales est présentée à la figure 1.3.

La grande majorité des éruptions solaires se produisent précisément sur les régions actives. Cette coïncidence spatiale suggère fortement que les éruptions solaires puisent leur énergie dans le champ magnétique coronal. Plus précisément, des études plus récentes montrent que le lieu de déclenchement des éruptions solaires se trouve au-dessus de la photosphère, à la base de la couronne [2]. Dans cette région, la conductivité électrique est tellement grande que le temps caractéristique de la dissipation ohmique est beaucoup plus grand que celui des éruptions solaires. Un mécanisme probable pour expliquer la libération soudaine d'énergie



**FIGURE 1.2.** Trois images du Soleil prises le 24 octobre 2013 (minuit) par le *Solar Dynamics Observatory* (SDO). (GAUCHE) Image en visible du Soleil. On peut y remarquer des taches sombres à la photosphère. (CENTRE) Magnétogramme du Soleil au même moment, provenant de l'instrument HMI (*Helioseismic and Magnetic Imager*). L'intensité du champ magnétique visant vers l'extérieur de la Terre est représentée en rouge, tandis que l'intensité du champ magnétique visant vers la Terre est représentée en bleu. (DROITE) Image à 171 Å du Soleil au même moment. Cette longueur d'onde correspond à une raie d'émission du FeIX à des températures d'environ 600 000 K. Elle nous montre donc des structures plus élevées de la couronne solaire incluant des boucles coronales. Ces images en domaine public proviennent du site internet de l'observatoire SDO (<https://sdo.gsfc.nasa.gov/data/aiahmi/>).



**FIGURE 1.3.** Image dans la bande 195 Å d'une boucle coronale (TRACE/NASA)([https://sdowww.lmsal.com/TRACE/POD/images/T195\\_020824\\_0050Xflare.jpg](https://sdowww.lmsal.com/TRACE/POD/images/T195_020824_0050Xflare.jpg)).

magnétique sur les boucles coronales est donc la *reconnexion magnétique* (discutée à la section 1.5).

### 1.1.2. Statistiques des éruptions solaires

Des séries temporelles du flux de rayon X, comme celles présentées à la figure 1.1b, nous montrent que les éruptions de grandes tailles semblent moins fréquentes que les plus petites. Pour visualiser les statistiques à plus long terme, la figure 1.4 présente des distributions des énergies de pics ( $P$ ) des éruptions solaires pendant trois cycles solaires.

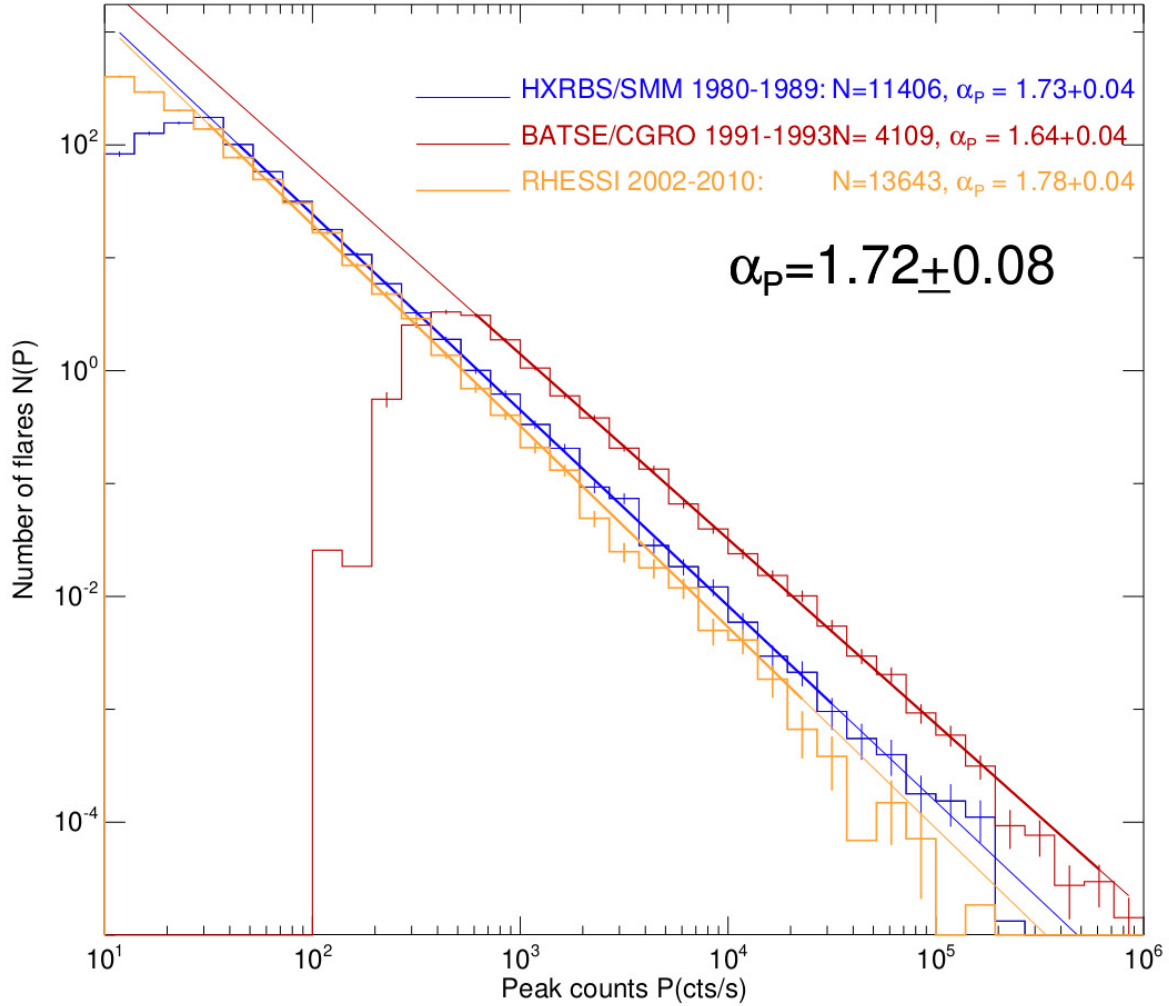
Cette figure met en évidence la statistique en loi de puissance des éruptions solaires, de la forme suivante :

$$f(P) = f_0 P^{-\alpha_p} . \quad (1.1.1)$$

où  $\alpha_p$  est l'indice d'exposant et  $f_0$  est un facteur de normalisation.

Ce qui est particulièrement surprenant de cette distribution observée est qu'elle couvre plusieurs ordres de grandeur (plus de quatre dans le cas de la figure 1.4). En fait, cette distribution préserve sa forme à différentes phases du cycle d'activité magnétique, avec seul changement le facteur  $f_0$  qui ne fait que décaler verticalement la courbe. Le tout suggère une invariance d'échelle du mécanisme de déclenchement des éruptions solaires, c'est-à-dire qu'il y a un mécanisme commun derrière le déclenchement des événements de toutes les tailles et rien ne distingue fondamentalement une très faible éruption d'une très intense. Les temps d'attente précédant une éruption mesurée dans la bande 1–8 Å suivent également une loi de puissance pour des temps d'attente plus grands que  $\sim 6$  heures [3].

Un modèle théorique du déclenchement des éruptions solaires doit donc satisfaire entre autres à toutes les contraintes présentées dans cette section. Il doit permettre une libération très rapide de l'énergie magnétique coronale ainsi que de permettre une grande variété d'échelles énergétiques. Dans la prochaine section, nous présenterons un cadre théorique habituellement utilisé pour modéliser le mécanisme de déclenchement des éruptions solaires.



**FIGURE 1.4.** Distributions des intensités maximales des éruptions solaires pendant trois cycles solaires, tirées de Aschwanden 2011 [3]. Les trois histogrammes représentent les résultats des mesures de trois instruments : le *Hard X-Ray Burst Spectrometer* sur le *Solar Maximum Mission* (HXRBS/SMM, en bleu), le *Burst And Source Transient Experiment* sur le *Compton Gamma Ray Observatory* (BATSE/CGRO, en rouge), et le *Ramaty High-Energy Solar Spectroscopic Imager* (RHESSI, en orange). Les lignes droites sont les meilleurs ajustements d’une loi de puissance (voir équation 1.1.1) sur les observations.

### 1.1.3. Modèles physiques des éruptions solaires

#### 1.1.3.1. Notions de base de la magnétohydrodynamique

Le lieu de déclenchement habituel des éruptions est la boucle coronale ancrée sur les régions actives (voir figures 1.3). Dans cet environnement, on retrouve un plasma très chaud et fortement magnétisé. Ce plasma est confiné par le champ magnétique de la couronne. De plus, les lignes de champ magnétique ne s’entrecroiseront pas, sauf dans le cas d’une instabilité

locale. Un cadre théorique approprié à ce genre d'environnement est la *magnétohydrodynamique* (MHD). Elle décrit à la fois le mouvement des fluides électriquement conducteurs ainsi que leur interaction avec les champs électromagnétiques.

La conservation de la masse d'un fluide de densité  $\rho$  s'écoulant à vitesse  $\mathbf{u}$  s'écrit de la forme suivante :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 . \quad (1.1.2)$$

La conservation du moment cinétique est exprimée par l'équation de *Navier Stokes* :

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \frac{1}{\mu_0 \rho} (\nabla \times \mathbf{B}) \times \mathbf{B} - \frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{g} , \quad (1.1.3)$$

où  $p$  est la pression du fluide,  $\nu$  la viscosité cinématique et  $\mathbf{g}$  est toutes autres forces massiques extérieures, incluant la gravité.

Les équations de Maxwell sont utilisées comme point de départ des aspects électromagnétiques :

$$\nabla \cdot \mathbf{E} = \frac{\rho_e}{\epsilon_0} , \quad (1.1.4)$$

$$\nabla \cdot \mathbf{B} = 0 , \quad (1.1.5)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} , \quad (1.1.6)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} , \quad (1.1.7)$$

où  $\mathbf{E}$  est le champ électrique,  $\rho_e$  la densité électrique,  $\epsilon_0$  la permittivité du vide,  $\mathbf{B}$  le champ magnétique,  $\mu_0$  la perméabilité magnétique du vide et  $\mathbf{J}$  la densité de courant électrique.

Il est parfois plus pratique de représenter le champ magnétique comme un potentiel vecteur  $\mathbf{A}$ , de sorte que  $\mathbf{B} = \nabla \times \mathbf{A}$ . Cette décomposition en potentiel vecteur est toujours possible grâce à l'équation 1.1.5. Il est important de noter que le potentiel vecteur  $\mathbf{A}$  est dégénéré au gradient d'une quantité scalaire près, demandant ainsi généralement le choix d'une jauge.

La version généralisée de la Loi d'Ohm pour un courant de plasma s'écrit de la façon suivante :

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) , \quad (1.1.8)$$

où  $\sigma$  est la conductivité électrique. Pour la suite, nous supposons que nos écoulements sont non-relativistes, ce qui implique qu'on néglige le deuxième terme du membre de droite de la loi d'Ampère (1.1.7). En combinant cette équation simplifiée avec la Loi de Faraday (1.1.6), et en définissant la *diffusivité magnétique*  $\eta$  comme étant  $\frac{1}{\mu_0\sigma}$ , on obtient l'équation d'induction magnétohydrodynamique :

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B} - \eta \nabla \times \mathbf{B}) . \quad (1.1.9)$$

Cette équation décrit l'évolution du champ magnétique dans un plasma. Le premier terme du membre de droite représente la contribution de l'écoulement d'un fluide électriquement conducteur comme source du champ magnétique. Le deuxième terme représente la dissipation ohmique du champ magnétique avec le temps.

Une estimation adimensionnelle de l'importance relative entre les deux termes du membre de droite de l'équation (1.1.9) est représentée par le *Nombre de Reynolds magnétique*  $R_m$  :

$$R_m = \frac{u_0 l}{\eta} , \quad (1.1.10)$$

où  $u_0$  représente la vitesse caractéristique du courant et  $l$  une échelle spatiale. Le nombre de Reynolds quantifie l'importance relative entre l'advection (terme source du champ magnétique) par rapport à la diffusion du champ magnétique. Plus ce terme est grand, plus l'écoulement  $\mathbf{u}$  contribue à générer le champ magnétique. Dans le cas où  $R_m \ll 1$ , le champ magnétique aura tendance à se dissiper selon un *temps de diffusion magnétique* :

$$\tau = \frac{l^2}{\eta} . \quad (1.1.11)$$

Dans l'environnement local d'une boucle coronale (et dans la plupart des systèmes astrophysiques), le nombre de Reynolds magnétique est très élevé ( $R_m \gg 1$ ). On est donc dans un régime appelé *MHD idéale*. Dans ce régime, le terme dissipatif de l'équation d'induction (1.1.9) disparaît et après une intégration autour d'une surface  $S$  quelconque, on obtient l'équation suivante :



$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \int_S \mathbf{B} \cdot \hat{\mathbf{n}} dS = 0 . \quad (1.1.12)$$

Cette équation indique que le flux magnétique, donné par l'intégrale au membre de droite de l'équation (1.1.12), suit l'écoulement du fluide en régime MHD idéal. En autre mot, dans ce régime les lignes de champ magnétique sont "gelées" dans le fluide et ne peuvent s'entrecroiser.

Finalement, un plasma magnétisé de volume  $V$ , délimité par une surface  $S$  et sujet à une force magnétique  $\mathbf{F}$  aura une évolution temporelle de son énergie magnétique exprimée de la façon suivante :

$$\frac{d\mathcal{E}}{dt} = - \oint_S \left( \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \cdot \hat{\mathbf{n}} \right) dS - \int_V (\mathbf{u} \cdot \mathbf{F}) dV - \int_V \sigma^{-1} \mathbf{J}^2 dV , \quad (1.1.13)$$

Où  $\mathcal{E}$  est l'énergie magnétique définie par :

$$\mathcal{E} = \int_V \frac{\mathbf{B}^2}{2\mu_0} dV . \quad (1.1.14)$$

L'équation (1.1.13) présente trois façons de convertir l'énergie magnétique vers une autre forme. Ces trois formes sont : l'énergie électromagnétique entrant et sortant du volume  $V$  (premier terme du membre de droite), le travail exercé par la force magnétique (deuxième terme) et la dissipation ohmique (dernier terme).

### 1.1.3.2. Dissipation passive du champ magnétique coronal

Nous avons déjà établi dans les sections précédentes qu'une théorie satisfaisante d'une éruption solaire doit permettre une libération rapide ( $\sim 100$ s) d'une grande quantité d'énergie ( $\sim 10^{24}$  J).

Dans une boucle coronale, la diffusivité magnétique  $\eta$  est de l'ordre de  $1 \text{ m}^2\text{s}^{-1}$ . Donc, pour atteindre un temps de diffusion de l'ordre de 100s, il faudrait une longueur caractéristique de  $\sim 10$ m, selon l'équation (1.1.11). Par contre, l'énergie magnétique contenue dans une région de  $10\text{m}^3$  n'est pas suffisante pour atteindre une libération de  $\sim 10^{24}$  J. En effet, dans une boucle coronale, la valeur typique du champ magnétique est  $\sim 10^{-2}\text{T}$ , ce qui mène approximativement à une énergie magnétique totale relativement faible de l'ordre de  $10^6$  J pour une région de  $10\text{m}^3$ , selon l'équation (1.1.14). Ceci suggère que la dissipation passive du champ magnétique dans une boucle coronale ne peut pas expliquer la conversion énergétique produisant des éruptions solaires.

Il faut donc un processus alternatif pour atteindre le régime de libération d'énergie nécessaire. Un mécanisme prometteur pour atteindre ce régime est la *reconnexion magnétique*. La section suivante présentera quelques modèles de bases de la reconnexion magnétique.

### 1.1.3.3. *La reconnexion magnétique*

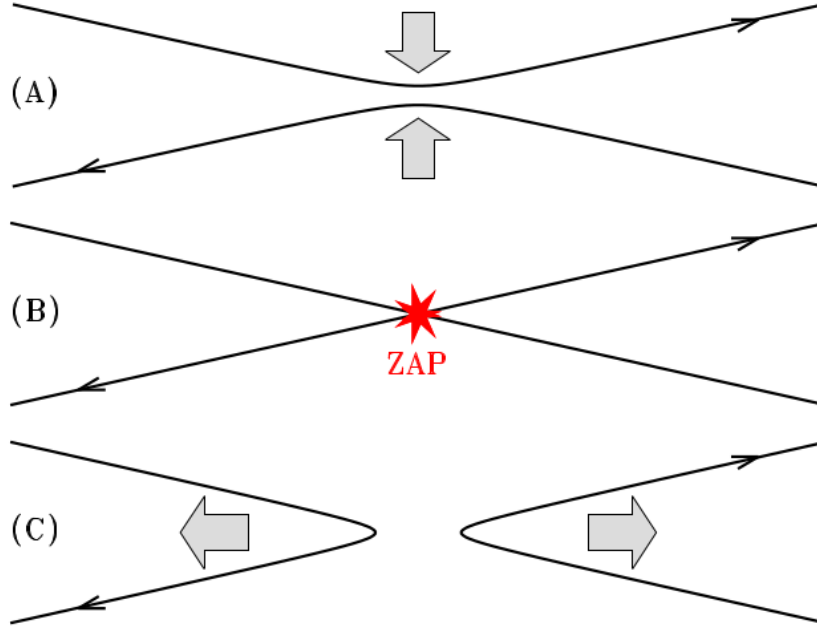
L'idée générale de la reconnexion magnétique est la suivante : un changement soudain de la topologie des lignes de champ magnétique engendre une conversion rapide de l'énergie magnétique en énergie cinétique et thermique. Cette idée est schématisée à la figure 1.5. Deux lignes de champ magnétique pointant dans des directions opposées se rapprochent. Ce rapprochement peut provenir entre autres de l'écoulement de plasma qui pousse les lignes de champ magnétique l'une vers l'autre. Si les lignes de champ magnétique se rapprochent assez pour que la longueur caractéristique  $l$  devienne assez petite, le nombre de Reynolds magnétique  $R_m$  deviendra assez petit pour que la diffusion magnétique se produise (voir équation (1.1.10)). Dans ce nouveau régime, les lignes de champ magnétique ne sont plus confinées à suivre l'écoulement local. Il y aura donc une reconnexion, et les deux nouvelles lignes de champ magnétique se redresseront étant donnée la nouvelle tension magnétique et entraîneront avec elles le plasma à l'extérieur du point de rencontre (où le  $R_m$  est encore très élevé).

Quelques modèles plus détaillés ont été développés pour quantifier l'effet de la reconnexion magnétique sur l'environnement local. La section suivante présentera brièvement le *modèle de Sweet-Parker*.

### 1.1.3.4. *Modèle de Sweet-Parker*

Le modèle de Sweet-Parker est basé sur le système présenté à la figure 1.5, mais avec des détails théoriques sur la dynamique du processus [4]. Elle suppose qu'il y a une nappe de courant entre les deux régions à champ magnétique de sens opposé. Le modèle est présenté à la figure 1.6.

Au centre, on a la zone de diffusion, où on retrouve une nappe de courant avec une densité de courant  $J$  qui pointe vers la direction sortant de la page. Cette nappe de courant aura comme effet de contribuer aux champs magnétiques opposés ( $\pm B_i$ ), une conséquence directe de la loi d'Ampère (voir équation (1.1.7)). Encore ici, les lignes de champ magnétique se reconnectent à la zone de diffusion et engendrent un écoulement horizontal ( $\pm v_o$ ). Par la

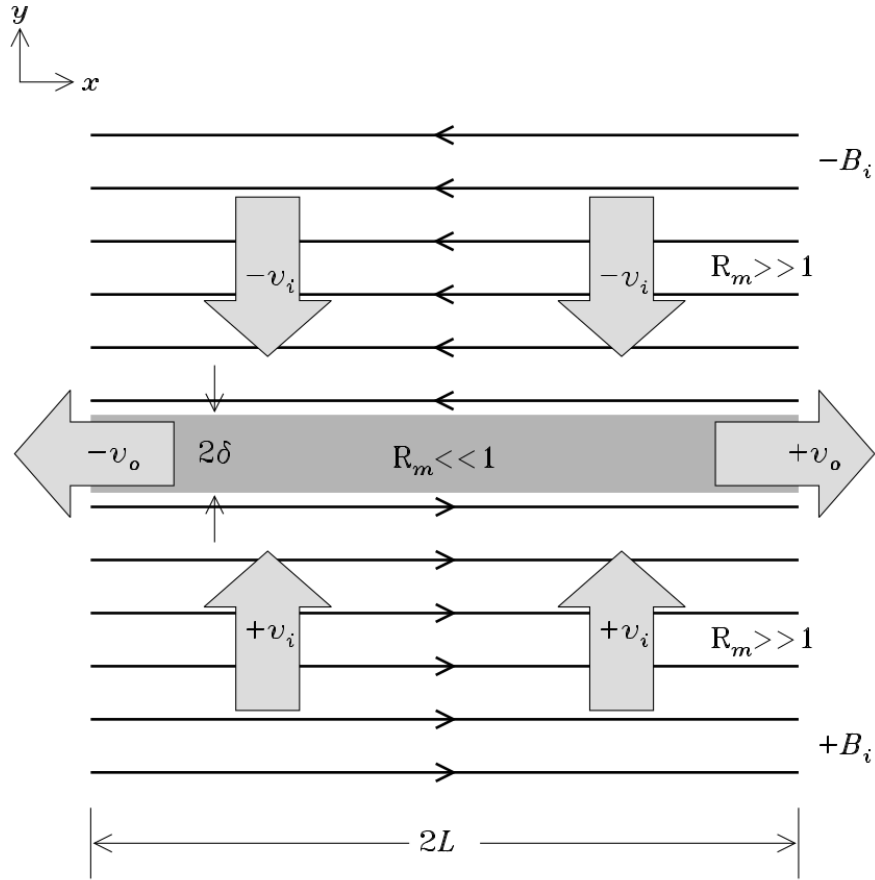


**FIGURE 1.5.** Schéma d'un processus de reconnexion magnétique en trois étapes. A) Deux lignes de champ magnétique pointant dans des directions opposées se rapprochent sous l'influence d'un écoulement extérieur. B) Les deux lignes se rapprochent suffisamment pour que le nombre de Reynolds magnétique ( $R_m$ ) devienne suffisamment petit pour que la dissipation magnétique se déclenche. Les lignes de champ magnétique vont donc se reconnecter. C) Les deux nouvelles paires de lignes de champ magnétique se redresseront sous l'effet de la très grande tension magnétique et le plasma local les suivra par un écoulement dans la même direction (flèches grises). Cette image est tirée des notes du cours de *Fluides Astrophysique* (PHY-6756) donné par Paul Charbonneau en 2018 à l'Université de Montréal.

conservation de la masse (équation (1.1.2)), le système compensera cette perte de masse à la zone de diffusion en forçant l'écoulement vertical ( $\pm v_i$ ) à se maintenir.

Dans ce modèle, l'énergie magnétique initiale est convertie en énergie thermique dans la zone de diffusion et en énergie cinétique par l'écoulement sortant. Si on suppose que la nappe de courant englobe toute la largeur d'une boucle coronale de  $\sim 10^8$ m, le temps caractéristique de reconnexion magnétique associé à ce modèle est de  $\tau_{SP} \sim 10^9$ s. Ce modèle a donc une vitesse de libération d'énergie magnétique plus grande que celle de la diffusion passive, mais n'est pas suffisant pour atteindre la vitesse mesurée des éruptions solaires.

Par contre, un modèle semblable à celle de Sweet-Parker a été développé avec l'ajout de l'effet des ondes de choc magnétohydrodynamique. Ce modèle est dit *de Petschek* [5]. Les ondes de choc magnétohydrodynamiques accélèrent d'avantage l'écoulement sortant de la zone de diffusion. Cet ajout mène à un temps caractéristique de reconnexion magnétique



**FIGURE 1.6.** Le modèle de Sweet-Parker. La région foncée représente la zone de diffusion. Les flèches grises représentent le courant entrant dans la zone de diffusion ( $v_i$ ) et celui sortant ( $v_o$ ). Cette image est tirée des notes du cours de *Fluides Astrophysiques* (PHY-6756) donné par Paul Charbonneau en 2018 à l'Université de Montréal.

de  $\tau_P \sim 10^4$ s. Cette valeur est calculée en supposant que la nappe de courant contenait toute la largeur de la boucle coronale ( $L \sim 10^8$ ). Cependant, il est raisonnable de supposer que l'échelle spatiale de la reconnexion magnétique est plus petite à l'intérieur de la boucle coronale, ce qui mène à des vitesses de reconnexion suffisamment grandes pour correspondre aux éruptions solaires mesurées.

Depuis le développement du modèle de Petschek, la reconnexion magnétique est considérée comme le mécanisme le plus prometteur pour la libération d'énergie par les éruptions solaires [6]. Elle permet une vitesse de libération suffisamment rapide. Par contre, des simulations numériques magnétohydrodynamiques de la reconnexion magnétique produisent encore à ce jour des résultats qui diffèrent des prédictions du modèle de Petschek (ainsi que celui de Sweet-Parker). Ce désaccord représente encore aujourd'hui un grand défi de la recherche sur

la reconnexion magnétique. Il y a donc plusieurs travaux de recherche en cours sur ce sujet au niveau analytique ainsi que numérique [7].

Peu importe le mécanisme de reconnexion magnétique proposé, il n'explique pas à lui seul la grande disparité d'échelles énergétiques des éruptions solaires. Les sections suivantes présentent quelques mécanismes nécessaires pour régler ce problème, dont le modèle de *nanoflare* de Eugene Parker.

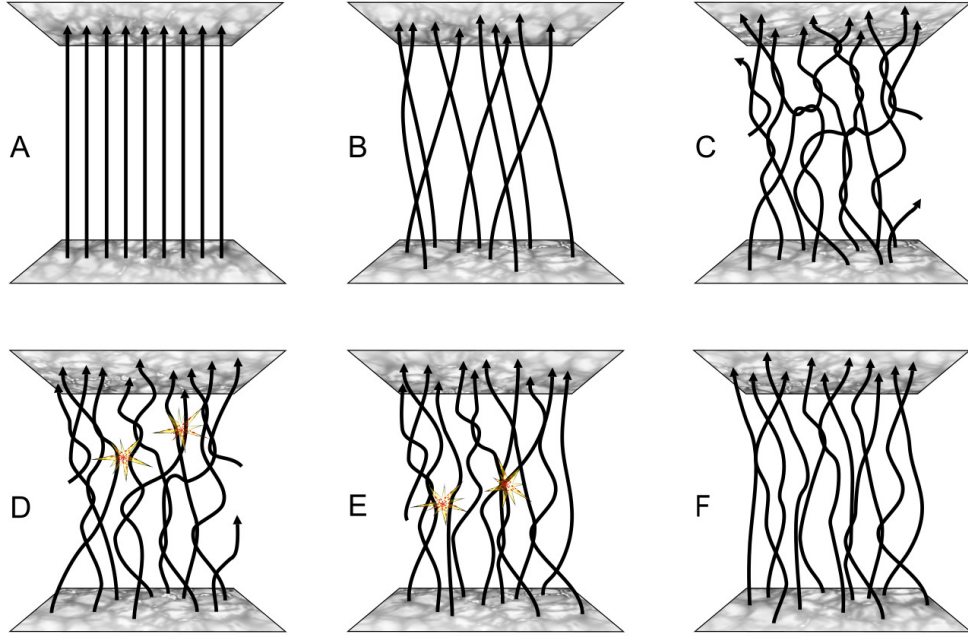
#### 1.1.3.5. *Modèle de nanoflares de Parker*

L'idée générale du modèle de *nanoflares* de Parker est présentée à la figure 1.7. On représente tout d'abord une boucle coronale comme étant une série de lignes de champ magnétique parallèles (A). Le haut et le bas de chacune de ces figures représentent la photosphère, donc les lignes droites verticales représentent les lignes de champ magnétique en forme de "U", mais redressées pour permettre la visualisation. Avec les mouvements aléatoires du plasma photosphérique, les lignes de champ magnétique se déplaceront et s'entortilleront, mais sans s'entrecroiser étant donnée la haute conductivité électrique de la couronne solaire (B-C). Ceci va ajouter de la pression et de la tension magnétique sur le système qui va lui permettre d'accumuler de l'énergie. Quand les lignes de champs magnétiques se rapprochent à un point où une instabilité magnétique se déclenche (dont celle décrite par le modèle de Sweet-Parker), il y a une reconnexion magnétique qui change la topologie locale des lignes de champ magnétique du système. Ces reconnexions vont localement libérer de l'énergie cinétique qui perturbera le champ magnétique dans les régions voisines. Ces perturbations peuvent ainsi causer des nouvelles instabilités, déclenchant à leur tour de nouvelles reconnexions magnétiques. On a donc une réaction en chaîne qui se propage jusqu'à ce que le système revienne à l'équilibre et que la topologie des lignes de champ magnétique redevienne stable. La taille de la réaction en chaîne dictera donc l'intensité de l'éruption solaire. Parker estime que chacun des événements individuels de reconnexion, appelés *nanoflares*, libère en moyenne  $\sim 10^{17}$  J d'énergie [8], ce qui représente une très petite fraction de l'énergie totale d'une très grosse éruption de  $\sim 10^{25} - 10^{26}$  J.

On a donc une théorie prometteuse d'un mécanisme responsable des éruptions de toute taille. Toutefois, à ce jour, les détections des nanoflares sont rares et très difficiles<sup>1</sup>. Les

---

1. En juillet 2020, l'agence spatiale européenne a présenté au public des vidéos récentes de la surface du Soleil captées par la sonde spatiale *Solar Orbiter* (voir [www.esa.int/Science\\_Exploration/Space\\_Science/Solar\\_Orbiter](http://www.esa.int/Science_Exploration/Space_Science/Solar_Orbiter)).



**FIGURE 1.7.** Représentation schématique du modèle de Parker des nanoflares. Les lignes verticales représentent une version “redressée” des lignes de champ magnétique d’une boucle coronale, de sorte que les deux extrémités des lignes représentent les points d’ancrage à la photosphère. Quand deux lignes de champ magnétique se rapprochent suffisamment, une instabilité se déclenche, causant un changement rapide de la topologie magnétique, appelée reconnexion magnétique. Ceci cause une réaction en chaîne de reconnexion magnétique, libérant chaque fois de l’énergie, jusqu’à ce que le système retourne à l’équilibre. Les lignes de champ en apparence coupées se reconnectent à l’extérieur de la portion de la boucle coronale représentée ici. Cette image est tirée du chapitre 12 de Aschwanden 2013 (écrit par Paul Charbonneau) [9]

simulations numériques sont donc un outil important pour la recherche en éruption solaire, pour tenter de simuler ces mécanismes difficiles à mesurer, mais potentiellement importants.

#### 1.1.4. Simulations numériques des éruptions solaires

Il y a eu plusieurs tentatives de produire des simulations magnétohydrodynamiques des éruptions solaires à partir de principes de base comme ceux présentés à la section (1.1.3.1). Par contre, la grande disparité des échelles spatiales (de nanoflares à éruptions complètes) et temporelles (accumulation et libération d’énergie) rend une simulation magnétohydrodynamique complète d’une éruption solaire numériquement très coûteuse. De plus, parmi les nombreuses instabilités plasmas proposées (comme celle du modèle de Sweet-Parker), il n’y

---

Ils ont montré pour la première fois des mini éclats à la couronne solaire qui pourraient être une version ‘grande échelle’ de ces fameuses nanoflares. Les futures observations de Solar Orbiter dans les années à venir sont ainsi particulièrement excitantes pour cette problématique.

a encore pas de consensus sur laquelle est véritablement derrière le déclenchement des éruptions. Une simulation MHD d'une éruption solaire devra donc inévitablement introduire une approximation représentant une instabilité générique.

Des modèles rapides dits de *criticalité auto-organisée* (ou SOC) ont justement été développés comme une simplification drastique de la physique des éruptions solaires tout en conservant leur disparité des échelles spatiales et temporelles. Ces modèles sont présentés à la section suivante.

### 1.1.5. Les éruptions solaires comme phénomène de criticalité auto-organisée

Historiquement, il y a eu plusieurs théories du déclenchement des éruptions solaires de type charge-décharge, où le système accumule de l'énergie jusqu'à un seuil et le libère instantanément. Ces théories furent une des premières tentatives d'explication des lois de puissance observées dans les distributions de taille des éruptions solaires [10]. En revanche, des études ultérieures ont montré qu'il n'y a pas de corrélations entre le temps d'attente avant une éruption et sa taille [11]. Ceci contredit les théories de type charge-décharge, car les plus grosses éruptions solaires ne semblent pas nécessiter un plus grand temps d'accumulation d'énergie.

Les systèmes physiques qui produisent des distributions d'évènements en loi de puissance sur plusieurs ordres de grandeur ont souvent été liés à la théorie de la criticalité auto-organisée, ou "SOC" pour le terme anglais "Self-Organized Criticality". Un système dynamique SOC est caractérisé comme ayant un état critique comme attracteur. L'exemple standard d'un système SOC est le tas de sable sur une table circulaire, sur laquelle on laisse tomber des grains de sable. Le tas de sable va éventuellement former un cône inversé d'un diamètre égal à celui de la table. Il produira des avalanches de sable de tailles variables pour éventuellement créer un équilibre entre les grains de sable qui entrent dans le système, et ceux qui en sortent en tombant du bord de la table. Dans ce sens, le tas de sable a atteint son état critique sans qu'on ait besoin d'ajuster ses paramètres ; on parle alors d'un système à criticalité auto-organisée.

Le concept de criticalité auto-organisée a été appliqué au champ magnétique d'une boucle coronale qui émet de l'énergie selon le mécanisme de nanoflares de Parker. Cette idée fut publiée pour la première fois par Edward T. Lu et Russell J. Hamilton en 1991 [12]. Dans cette

optique, ils ont développé des modèles sur réseau très simples appelés *modèles d'avalanches*, en référence aux avalanches des tas de sable. Ces modèles tentent de représenter les réactions en chaîne de reconnections magnétiques durant une éruption solaire par un réseau de nœuds interconnectés, chacun représentant un aspect magnétique local d'une boucle coronale.

Dans la prochaine section, les modèles d'avalanches étudiés dans ce projet de maîtrise seront présentés.

## 1.2. Modèles d'avalanches

Dans le contexte de la modélisation des éruptions solaires, les modèles d'avalanches sont des automates cellulaires de nœuds interconnectés sur un réseau généralement cartésien. La figure 1.8 présente un schéma d'un réseau à deux dimensions. La valeur de chacun des nœuds est dénotée par un nombre réel  $A_{i,j}$ , où  $(i,j)$  sont les coordonnées de chacun nœud dans le réseau 2D. Dans cet exemple, chaque nœud interagit seulement avec ses quatre voisins les plus rapprochés.

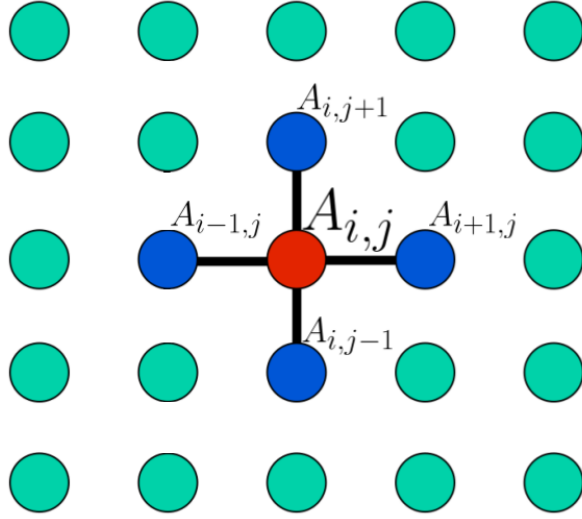
L'interprétation physique habituelle de cette variable  $A$  est la composante du potentiel vecteur le long d'une boucle coronale, qu'on définit comme la composante  $z$  (voir la figure 1.9). Comme  $\mathbf{B} = \nabla \times \mathbf{A}$ , le  $A_z$  est lié au degré de torsion toroïdale (en  $\phi$ ) du champ magnétique sur cette tranche. La simulation de ce réseau représente donc l'évolution temporelle de la torsion magnétique d'une tranche de la boucle coronale [13]. On suppose que  $B_z$  est essentiellement fixe, donc  $A_x$  et  $A_y$  le sont aussi. L'interaction entre les nœuds est liée à la dynamique de réaction en chaîne produite par les nanoflares, comme présentée à la figure 1.7.

Puisque  $A$  est associé à une mesure du champ magnétique local, on définit l'énergie du réseau comme étant :

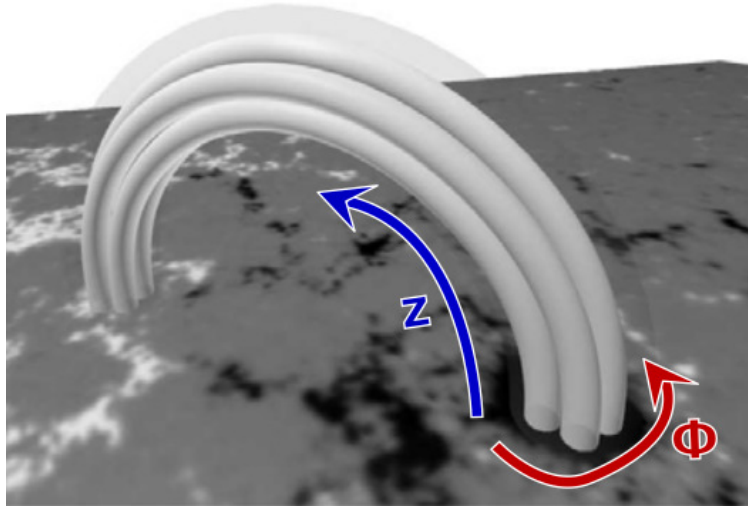
$$E_t = \sum_k A_k^2, \quad (1.2.1)$$

où on somme sur tous les nœuds du réseau. Ce calcul est un analogue adimensionnel au calcul d'énergie magnétique habituel présenté à l'équation (1.1.14). Bien sûr, l'équation de l'énergie magnétique (1.1.14) est une intégration volumique du champ magnétique, et non du potentiel vecteur. Toutefois, des études antérieures de ce genre de modèle ont montré qu'il existe une très forte corrélation entre la somme des valeurs nodales au carré et la somme





**FIGURE 1.8.** Réseau cartésien à deux dimensions utilisé pour les modèles d'avalanche. Un nœud particulier  $A_{i,j}$  (en rouge) interagit avec ses quatre voisins les plus près (en bleu).



**FIGURE 1.9.** Représentation schématique d'une boucle coronale avec la définition de nos axes de  $z$  et  $\phi$ . Cette figure est tirée de Strugarek et al. 2014 [14]

de son rotationnel au carré [9]. Cette corrélation nous permet de simplifier la définition de l'énergie de  $\sum(\nabla \times A)^2$  à  $\sum A^2$  de façon raisonnable, puisque les unités et l'échelle des valeurs de  $A$  sont arbitraires.

À toutes les itérations, on permet soit au système d'accumuler de l'énergie, soit de la libérer sous forme d'*avalanches* (analogue du modèle aux éruptions solaires), mais jamais les deux simultanément. Cette séparation est imposée pour respecter la séparation des échelles temporelles entre l'accumulation et la libération de l'énergie d'une boucle coronale.

La façon de simuler cette dynamique d'accumulation et libération sur ce réseau dépend du modèle d'avalanche utilisé. On présentera en premier lieu le modèle de Lu et Hamilton 1993 (LH).

### 1.2.1. Modèle de Lu et Hamilton 1993 (LH)

Un modèle surnommé LH91 fut le premier publié dans le contexte de la simulation des éruptions solaires par Edward T. Lu et Russell J. Hamilton [12] et est devenu un standard dans la littérature des modèles d'avalanches. En 1993 les mêmes auteurs ont proposé une variation sur leur modèle LH91, surnommée LH93 [13]. C'est ce modèle qui a été étudié dans le cadre de cette maîtrise.

Dans ce modèle, on permet au système d'accumuler de l'énergie en additionnant un petit incrément  $\delta A$  sur un nœud sélectionné aléatoirement :

$$A_{\mathbf{r}}^{n+1} = A_{\mathbf{r}}^n + \delta A , \quad (1.2.2)$$

où  $\mathbf{r}$  représente un vecteur d'indice  $(i,j)$  choisi de façon aléatoire (d'une distribution uniforme), et  $n$  est l'indice temporel. La valeur de l'incrément  $\delta A$  est choisie de façon aléatoire d'une distribution uniforme sur l'intervalle  $[\sigma_1, \sigma_2]$ . Les valeurs de  $\sigma_1$  et  $\sigma_2$  sont choisies de sorte que  $\frac{\sigma_1 + \sigma_2}{2} > 0$ , c'est-à-dire qu'en moyenne le système accumule de l'énergie par le forçage de l'équation (1.2.2). Cette étape simule les fluctuations du champ magnétique à des endroits aléatoires dans la boucle coronale.

Le réseau est forcé par l'équation (1.2.2) jusqu'à ce qu'un nœud soit instable. La stabilité du système dépend de la courbure de chaque nœud, calculé de la façon suivante :

$$\Delta A_{i,j}^n = A_{i,j}^n - \frac{1}{4} \sum_k A_k^n , \quad (1.2.3)$$

où  $k$  représente les indices de chacun des nœuds voisins du nœud  $(i,j)$ . Le système est considéré instable si la courbure  $\Delta A_{i,j}^n$  d'un des nœuds dépasse une valeur seuil de  $Z_c$ . L'équation (1.2.3) est une version discrète du laplacien  $\nabla^2(A_z \hat{\mathbf{z}})$ , une mesure du courant électrique le long de la composante  $z$ . Ce critère de stabilité est donc basé sur un seuil du courant  $J_z$ , au-dessus duquel une nappe de courant assez forte induira une instabilité magnétique produisant la reconnexion magnétique, comme celle présentée au modèle de Sweet-Parker (voir figure 1.6).

Quand il y a une instabilité dans le système, on coupe le forçage (équation 1.2.2) et on redistribue les variables nodales  $A$  des nœuds instables de la façon suivante :

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5}Z , \quad (1.2.4)$$

$$A_{i\pm 1,j\pm 1}^{n+1} = A_{i\pm 1,j\pm 1}^n + \frac{1}{5}Z . \quad (1.2.5)$$

Cette règle de redistribution est conservative, c'est-à-dire que la somme de  $A$  sur les 5 nœuds impliqués est conservée suite à la redistribution. De plus, chacun des quatre voisins du nœud instable reçoit une part égale de l'excès de  $A$  du nœud instable (une redistribution dite uniforme).

Cette redistribution peut faire qu'un ou plusieurs des noeuds voisins dépassent le seuil de stabilité, causant à leur tour une redistribution de la variable nodale à l'itération suivante. Cette cascade de redistributions simule la réaction en chaîne de reconnexion magnétique dans une boucle coronale (les nanoflares).

Les nœuds instables vont continuer à redistribuer leur variable nodale jusqu'à ce que le système retourne à l'équilibre global et recommence le forçage par l'équation (1.2.2). Les nœuds sur les frontières du système sont maintenus à  $A = 0$  et sont toujours considérés stables.

Selon notre définition de l'énergie du réseau, on peut montrer que chacune des redistributions définies par les équations (1.2.4) et (1.2.5) va mener à une diminution de l'énergie totale du réseau par la quantité suivante :

$$\Delta e_{i,j}^n = \frac{4}{5} \left( 2 \frac{|\Delta A_{i,j}^n|}{Z_c} - 1 \right) Z_c^2 . \quad (1.2.6)$$

La somme de tous les  $\Delta e_{i,j}^n$  pour une itération donnée représente l'énergie totale libérée par le réseau. Une unité énergétique habituellement utilisée est l'énergie libérée par un seul noeud qui dépasse infinitésimalement le seuil de stabilité, exprimée comme :

$$e_0 = \frac{4Z_c^2}{5} . \quad (1.2.7)$$

### 1.2.2. Modèles déterministes (D)

Les modèles dits *déterministes* dans le contexte des éruptions solaires ont été publiés pour la première fois en 2014 par Strugarek et al. [14]. Ils se basent sur les fondements du modèle de LH, mais avec quelques changements. Dans ce mémoire, nous présenterons une version spécifique de ces modèles déterministes : les modèles non conservatifs à seuil fixe et à redistribution uniforme (correspondant aux modèles D2 et D3 dans Strugarek et Charbonneau 2014 [15]). On surnomme cette version spécifique *modèle D*.

Dans les modèles D, le forçage de l'énergie dans le système se fait en multipliant tous les nœuds simultanément par une valeur très près de 1 :

$$A_{i,j}^{n+1} = A_{i,j}^n \cdot (1 + \epsilon), \quad \epsilon \ll 1. \quad (1.2.8)$$

De cette façon, la courbure de chacun des nœuds augmente de façon déterministe (d'où vient le nom). L'idée physique derrière ce forçage est que les points d'ancrage des boucles coronales subissent l'effet des écoulements photosphériques à des longueurs caractéristiques supérieures ou égales à la largeur d'une boucle coronale (voir figure 1.3). En réponse à ce mouvement de la base, la boucle accumulera une torsion globale qui affectera simultanément toutes les régions d'une tranche de la boucle (notre réseau).

Le critère de stabilité pour les modèles D est le même que pour celle du modèle LH93 (voir équation (1.2.3)).

Encore ici, une instabilité dans le système mène à une redistribution des variables nodales. Cette redistribution est la même que celle de LH93, mais avec une perte aléatoire de valeur nodale à chaque redistribution :

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5}Z, \quad (1.2.9)$$

$$A_{i\pm 1, j\pm 1}^{n+1} = A_{i\pm 1, j\pm 1}^n + \frac{r_{nc}}{5}Z, \quad (1.2.10)$$

où  $r_{nc}$  est un nombre aléatoire choisi d'une distribution uniforme dans l'intervalle  $[D_{nc}, 1]$ , avec  $D_{nc}$  défini comme le "degré de conservation". Une valeur de  $D_{nc} = 1$  nous ramène à une redistribution conservative, comme celle utilisée dans le modèle LH93 (équations (1.2.4) et (1.2.5)). Quand  $D_{nc} < 1$ , la redistribution est non conservative, et représente la seule

source de stochasticité dans les modèles déterministes. Notons que l'existence d'une forme de stochasticité est nécessaire pour atteindre l'état SOC dans ces modèles.

Dans ce mémoire, nous avons étudié deux modèles de type D. Le premier, surnommé D09, est caractérisé par un degré de conservation  $D_{nc}$  plutôt élevé de 0.9. Le second, surnommé D01, contient un  $D_{nc}$  de 0.1 et donc est considéré plutôt non-conservatif.

### 1.2.3. Modèles déterministes avec minimisation Monte-Carlo (DMC)

Les modèles déterministes avec minimisation Monte-Carlo, qu'on surnomme *modèle DMC*, ont été développés dans le cadre de ce projet de maîtrise. Une description et analyse plus détaillées de ces modèles sont présentées au chapitre 2 de ce mémoire. Ces modèles sont directement inspirés d'un modèle d'avalanche publié par Farhang et al. en 2019 [16], qui a été lui-même inspiré par les modèles déterministes présentés à la section 1.2.2.

Quand il y a au moins un noeud instable dans le système, les noeuds instables redistribuent leurs excès de façon inégale à tous leurs voisins, mais de façon conservative. La part que chaque voisin recevra durant la redistribution sera choisie de sorte à minimiser le plus possible l'énergie après la redistribution. Cette minimisation est justifiée du fait qu'un système dissipatif instable tend à se déplacer vers son état d'énergie minimale compatible avec les conditions frontières. L'équation de la redistribution est donc :

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5} Z_c, \quad (1.2.11)$$

$$\begin{aligned} A_{i+1,j+1}^{n+1} &= A_{i+1,j+1}^n + \frac{1}{5} \frac{r_1}{a} Z_c, \\ A_{i-1,j+1}^{n+1} &= A_{i-1,j+1}^n + \frac{1}{5} \frac{r_2}{a} Z_c, \\ A_{i+1,j-1}^{n+1} &= A_{i+1,j-1}^n + \frac{1}{5} \frac{r_3}{a} Z_c, \\ A_{i-1,j-1}^{n+1} &= A_{i-1,j-1}^n + \frac{1}{5} \frac{r_4}{a} Z_c, \end{aligned} \quad (1.2.12)$$

où  $a = r_1 + r_2 + r_3 + r_4$ . Pour minimiser l'énergie post-redistribution, on utilise une méthode approximative. On pose  $\mathbf{r} = (r_1, r_2, r_3, r_4)$ , où chacun des coefficients est choisi d'une distribution aléatoire entre 0 et 1. Pour un noeud instable donné, on génère  $N_{min}$  vecteurs  $\mathbf{r}$ . On choisit le vecteur  $\mathbf{r}$  qui permettra la plus grande libération d'énergie suite à la redistribution définie par (2.2.5). Ce vecteur  $\mathbf{r}$  est donc utilisé pour produire la redistribution du noeud instable. Cette minimisation imparfaite est inspirée du fait que la reconnexion magnétique

dans une boucle coronale aura tendance à minimiser son énergie magnétique locale, mais pas nécessairement jusqu'à un minimum absolu.

Dans ce mémoire, on sépare les modèles DMC en deux versions. Dans la première, appelée modèle DMC1, le seuil de stabilité  $Z_c$  sur la courbure  $\Delta A$  d'un nœud est tiré à chaque itération parmi une distribution normale autour de  $\bar{Z}_c$  et de largeur à mi-hauteur de  $\sigma$  (voir section 2.5 dans Strugarek et al. 2014 [14]). Dans la deuxième, appelée modèle DMC2, le seuil de stabilité est fixe, comme dans le cas des modèles D. La seule stochasticité du modèle DMC2 n'est donc que dans la redistribution et la condition initiale.

Nous avons présenté dans cette section tous les modèles d'avalanche étudiés dans le cadre de cette maîtrise. Des analyses de ces modèles seront présentées au chapitre 2 pour les modèles DMC et dans l'article pour les modèles D et LH. Dans la section suivante, nous présenterons les bases théoriques derrière *l'assimilation de données*, l'outil numérique de prédiction utilisé dans le cadre de cette maîtrise.

### 1.3. Assimilation de données

L'assimilation de données est une technique qui consiste à incorporer des données observées sur un modèle numérique pour produire des simulations plus rapprochées de la réalité. C'est une technique statistique qui prend en compte les incertitudes reliées à chacune de ses sources. Elle est depuis longtemps utilisée en météorologie, où les observations disponibles de l'état de l'atmosphère sont combinées avec des résultats de modèles d'évolution météorologique pour produire des estimations plus optimisées des conditions météorologiques futures<sup>2</sup>.

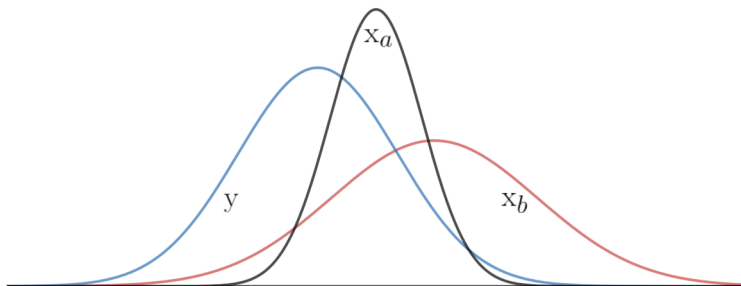
#### 1.3.1. Base théorique de l'assimilation de données

Supposons qu'on cherche à déterminer un état-modèle  $x$  qui soit le plus près possible de la valeur réelle  $x_t$  d'un système. Supposons également que nous avons accès à deux autres valeurs, un premier estimé de l'état-modèle généré par un modèle quelconque ( $x_b$ ) ainsi qu'une observation reliée à notre système ( $y$ ). Chacune de ces variables contient également une incertitude (reliée à la justesse du modèle ou à la précision de la mesure).

En combinant l'information de nos deux sources ( $x_b$  et  $y$ ), on peut calculer un état-modèle optimisé  $x_a$ , qu'on nomme *l'analyse*, qui sera quelque part entre l'observation  $y$  et

---

2. Pour des détails supplémentaires, voir *Atmospheric data analysis (No. 2)* de Daley, R (1994) [17] et *Atmospheric modeling, data assimilation and predictability* de Kalney, E (2002) [18]



**FIGURE 1.10.** Représentation schématique de l’analyse  $x_a$  comme étant une distribution plus compacte et quelque part entre l’observation  $y$  et l’état-modèle  $x_b$ .

l’état-modèle  $x_b$ . Le concept en une dimension est présenté à la figure 1.10. L’incertitude de  $x_a$  est la plus petite des trois distributions, résultat de l’ajout d’informations statistiques des deux sources  $y$  et  $x_b$ .

En version vectorielle (de dimension  $N$  pour les états-modèles et  $M$  pour les observations), l’analyse  $\mathbf{x}_a$  sera exprimé comme :

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - H(\mathbf{x}_b)) , \quad (1.3.1)$$

où  $\mathbf{K}$  est appelé la *matrice de gain* et  $H$  est l’*opérateur d’observation*.

L’opérateur d’observation  $H$  sert à effectuer un lien quantitatif entre un état-modèle  $\mathbf{x}$  et des observations  $\mathbf{y}$ . Par exemple, dans le cas d’une analyse météorologique, on cherche à créer un profil de température ( $\mathbf{x}$ ) sans avoir accès à des mesures directes de température. On utilise donc des mesures indirectes comme la radiance captée par des satellites ( $\mathbf{y}$ ). Pour comparer les observations avec notre modèle météorologique, on traduit le profil de température modélisé ( $\mathbf{x}_b$ ) en radiance avec l’opérateur  $H$  qui dépend de la théorie du transfert radiatif. En autres mots,  $H(\mathbf{x}_b)$  représente la mesure que notre instrument ferait si notre modèle trouve toujours la valeur réelle ( $\mathbf{x}_b = \mathbf{x}_t$ ) et que notre instrument était parfait.

La *matrice de gain*  $\mathbf{K}$  est déterminée de manière à optimiser  $\mathbf{x}_a$  (diminuer l’écart entre  $\mathbf{x}_a$  et la valeur réelle  $\mathbf{x}_t$ ). Elle prend en compte les statistiques de l’erreur de prévision du modèle ( $\epsilon_b = \mathbf{x}_b - \mathbf{x}_t$ ) ainsi que des observations ( $\epsilon_0$ ). Pour manipuler les erreurs de chaque source en version vectorielle, on introduit les trois matrices suivantes : matrice de covariance

d'erreur de prévision  $\mathbf{B} = \langle \epsilon_b \epsilon_b^T \rangle$ , matrice de covariance d'erreur d'observation  $\mathbf{R} = \langle \epsilon_o \epsilon_o^T \rangle$  et matrice de covariance d'erreur d'analyse  $\mathbf{P}_a = \langle \epsilon_a \epsilon_a^T \rangle = \langle (\mathbf{x}_a - \mathbf{x}_t)(\mathbf{x}_a - \mathbf{x}_t)^T \rangle$ <sup>3</sup>.

Dans tous les problèmes d'assimilation de données, on se définit une fonction coût, dénotée par  $J$ , une mesure de l'écart entre  $\mathbf{x}_a$  et  $\mathbf{x}_t$  qu'on cherche à minimiser. Par exemple, on peut montrer qu'une fonction coût définie par  $J = \langle \epsilon_a^T \epsilon_a \rangle$  sera minimisée pour une matrice de gain définie comme :

$$\mathbf{K} = \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} . \quad (1.3.2)$$

Cette expression est utile pour nous donner une intuition mathématique du fonctionnement de l'assimilation de données. Par contre, les matrices de covariance  $\mathbf{B}$  et  $\mathbf{R}$  sont en général de taille immense (de l'ordre du carré du nombre de variables dans le modèle et du nombre d'observations) et l'inversion présente dans l'équation (1.3.2) devient très coûteuse. On utilisera donc des méthodes approximatives et numériques pour minimiser la fonction coût  $J$ , dont les méthodes entre autres de 3D-Var, 4D-Var ou le Recuit simulé.

#### 1.3.1.1. 3D-Var

Le 3D-Var utilise une définition de la fonction coût  $J$  qui nous aide à trouver son minimum de façon plus efficace. On suppose premièrement que la distribution de probabilité d'une mesure suit une distribution normale autour de la valeur réelle, de même pour les paramètres d'un modèle (comme schématisé à la figure 1.10). La fonction coût est définie de la façon suivante :

$$J(\mathbf{x}) = -\ln(p(\mathbf{x}|\mathbf{y})) , \quad (1.3.3)$$

où  $p(\mathbf{x}|\mathbf{y})$  est la probabilité que l'état modèle  $\mathbf{x}$  soit vrai ( $\mathbf{x} = \mathbf{x}_t$ ) étant donné que notre instrument de mesure observe  $\mathbf{y}$  (avec une certaine incertitude). Par exemple, si on mesure systématiquement des valeurs exceptionnelles de  $\mathbf{y}$ , il est plus probable que la valeur réelle de l'état-modèle ( $\mathbf{x}_t$ ) soit aussi une valeur exceptionnelle, donc  $p(\mathbf{x}|\mathbf{y})$  aura un maximum sur cette valeur exceptionnelle.

---

3. En pratique, les moyennes de  $\mathbf{B}$  et  $\mathbf{P}_a$  peuvent être estimées à partir de plusieurs réalisations du même modèle.



En appliquant une approche bayésienne pour résoudre  $p(\mathbf{x}|\mathbf{y})$ , on arrive à la fonction coût suivante :

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}) + Const . \quad (1.3.4)$$

On a donc une fonction coût sous une forme quadratique, dépendant également de l'erreur du modèle  $\mathbf{B}$  que celle des observations  $\mathbf{R}$  [19].

Cette fonction est minimisée quand  $\mathbf{x}$  prend la valeur de  $\mathbf{x}_a$  de l'équation (1.3.1), mais comme mentionné précédemment, nous n'avons habituellement pas accès à la forme exacte de la matrice  $\mathbf{K}$ . Pour cette raison, la méthode de 3D-var utilise plutôt une descente de gradient pour estimer  $\mathbf{x}_a$ , ce qui est particulièrement pratique dans le cas d'une fonction de forme quadratique comme la fonction (1.3.4).

#### 1.3.1.2. 4D-Var

Dans certains cas, on ne cherche pas seulement à reproduire une observation  $\mathbf{y}$  spécifique, mais plutôt une série temporelle d'observation  $\mathbf{y}(t)$ . Le 4D-Var est une méthode qui produit un ajustement de paramètres d'une condition initiale  $\mathbf{x}$  pour bien reproduire une série d'observations. Cette méthode prend donc en compte l'évolution temporelle du modèle dans sa minimisation de la fonction coût  $J$ , ce qui n'est pas le cas dans la méthode 3D-Var.

On représente  $H_i$  comme étant l'opérateur qui avance l'état-modèle  $\mathbf{x}$  dans le temps par une simulation jusqu'à l'itération  $i$ , et le converti en équivalent observation. La fonction coût du 4D-Var pour une fenêtre temporelle de  $n$  vecteur d'observation s'écrit donc comme :

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^n (H_i(\mathbf{x}) - \mathbf{y}_i)^T \mathbf{R}^{-1}(H_i(\mathbf{x}) - \mathbf{y}_i) + Const . \quad (1.3.5)$$

On a donc une fonction coût qui prend en compte toutes les observations d'une certaine fenêtre temporelle (les  $\mathbf{y}_i$  avec  $i$  de 0 à  $n$ ) [20]. La difficulté du 4D-var est dans la minimisation de  $J(\mathbf{x})$ . Le calcul du gradient de  $J(\mathbf{x})$  nécessite transposer l'opérateur  $H_i$ , ce qui implique une version inversée du modèle d'évolution temporelle, appelée modèle adjoint (permettant de revenir dans un état passé étant donnée un état  $\mathbf{x}$  quelconque). Ce modèle adjoint peut parfois être difficile à implémenter, surtout dans le cas d'un modèle avec des composantes

stochastiques. Malgré cette difficulté, cette technique a déjà été implémentée avec les modèles d’avalanche par Bélanger et al. (2007) [21], mais les résultats se sont avérés peu probants. Ceci provient essentiellement du très grand nombre de minima locaux des fonctions coût utilisées dans le cadre de ce projet. Même si l’algorithme de minimisation est très efficace à calculer les gradients locaux, cette localité nous empêche généralement de trouver efficacement un minimum global.

Pour cette raison, dans le cadre de ce projet de maîtrise nous avons plutôt utilisé la méthode du recuit simulé pour chercher un minimum de  $J$ .

### 1.3.1.3. Recuit simulé

Le recuit simulé est une méthode stochastique de minimisation de fonction qui s’inspire des principes statistiques de la thermodynamique<sup>4</sup>. Elle est particulièrement utilisée pour trouver le minimum global pour des fonctions contenant plusieurs minima secondaires.

À partir d’un état-modèle de départ  $\mathbf{x}$ , on choisit un déplacement aléatoire  $\delta\mathbf{x}$  et on calcule sa fonction coût  $J(\mathbf{x} + \delta\mathbf{x})$ . On calcule ensuite la différence de coût causée par ce déplacement par :  $\Delta J = J(\mathbf{x} + \delta\mathbf{x}) - J(\mathbf{x})$ . La probabilité que  $\mathbf{x} + \delta\mathbf{x}$  devienne le nouveau point de départ de la prochaine itération (qu’on appelle un *pas*) est donnée par :

$$p = \min [1, \exp(-\Delta J/kT)] , \quad (1.3.6)$$

où  $T$  représente la température qui peut varier d’une itération à l’autre, et  $k$  est une constante choisie au départ. On remarque que si  $\Delta J$  est négatif, le pas est effectué avec certitude. De cette façon, l’algorithme a tendance à privilégier les pas descendants. Si  $\Delta J$  est positif, il y a une probabilité non nulle que le pas se fasse de façon ascendante. Cette probabilité dépend de la grandeur de  $\Delta J$  (pour décourager les trop grandes augmentations de  $J$ ) et de la température. Ce pas ascendant permet à l’algorithme de se sortir des minima locaux pour trouver de nouveaux puits de fonction coût. Habituellement, la température  $T$  diminue graduellement au cours de la séquence d’optimisation, ce qui permet à l’algorithme de se ‘coincer’ dans un minimum. Des détails supplémentaires de cette méthode sont présentés dans l’article inclus dans ce mémoire.

---

4. La méthode du recuit simulé est expliquée en détail avec des exemples numériques dans le chapitre 10.9 du livre de *Numerical Recipes in Fortran 77* par Press et al. [22].

## 1.4. Prédiction des éruptions solaires

Dans certaines circonstances, les éruptions solaires peuvent causer des dommages importants sur les satellites et les infrastructures technologiques sur Terre, en plus de présenter un danger d'irradiation pour les astronautes en haute orbite. Il y a donc eu plusieurs tentatives de développer des méthodes de prédiction des éruptions solaires pour éventuellement pouvoir mitiger certains de ces effets néfastes. Par contre, il n'y a aucune méthode à ce jour capable de produire des prédictions significativement plus précises que des prévisions climatologiques, c'est-à-dire basées purement sur le comportement moyen à long terme des éruptions sans tenir compte des aspects à court terme. Pour mettre ce défi en évidence, un atelier a eu lieu en 2009 à Boulder au Colorado (É.-U.) pour comparer les performances de prédiction de 11 équipes de scientifiques [23]. Chacune de ces équipes avait déjà publié auparavant des algorithmes de prédiction qui semblaient prometteurs. Par contre, elles utilisaient toujours leurs propres métriques de qualité de prédiction, ce qui rendait très difficile la comparaison des algorithmes publiés dans la littérature. Pour cette raison, l'atelier de Boulder a fixé une métrique de qualité ainsi qu'une base de données commune pour tous les participants.

Chacune des équipes devait effectuer des prédictions "all clear", c'est-à-dire prédire s'il va y avoir oui ou non une éruption solaire d'une certaine intensité durant une fenêtre temporelle bien définie. Elles avaient accès à des données quotidiennes de magnétogrammes et de cartes d'intensité lumineuse (comme ceux présentés aux panneaux gauche et centre de la figure 1.2) durant les années 2000-2005 inclusivement, chacune provenant du Solar and Heliospheric Observatory/Michelson Doppler Imager (SOHO/MDI) [24]. Elles avaient aussi accès à une liste d'éruptions solaires selon leur jour de déclenchement ainsi que leurs tailles selon trois catégories :  $\geq C1.0$ ,  $\geq M1.0$  et  $\geq M5.0$ . Cette liste provient des séries temporelles du flux en rayon X mesuré par les satellites GOES (comme celui présenté à la figure 1.1b). Les organisateurs reconnaissent que ces données sont limitées et que plusieurs autres paramètres temporels et spatiaux pourraient certainement être utiles pour la prédiction, mais le but était une comparaison juste entre plusieurs algorithmes de prédiction déjà publiés. Chacun de ces algorithmes comprenait deux étapes principales : la première était de caractériser les paramètres physiques de la région active d'où proviendrait une éruption (le flux magnétique, la topologie inférée des lignes de champ magnétique, etc.), et la deuxième était d'utiliser

des méthodes statistiques à partir du passé pour prédire le comportement futur de la région active cible. Voici quelques exemples de méthodes utilisées dans cet atelier.

Schrijver (2007) [25] utilise un paramètre  $\mathcal{R}$  qui est une mesure du flux magnétique dans la région d'inversion de polarité entre deux taches solaires de pôle opposées (ces paires sont assez communes, voir panneau du centre de la figure 1.2). Ce flux magnétique serait lié au courant électrique qui, selon la théorie de la reconnexion magnétique (voir section 1.5), est une cause d'instabilité provoquant les éruptions solaires. Quelques études ont déjà montré une corrélation entre ce courant électrique et le déclenchement d'une éruption solaire (voir Schrijver (2007) [25] et Dhuri et al. 2019 [26]). Pour prédire une éruption sur une région active donnée, la méthode de Schrijver (surnommée  $\log(\mathcal{R})$ ) se base sur un certain seuil de  $\mathcal{R}$ , précalculé par rapport à des statistiques du comportement passé des régions actives.

Un autre exemple est une technique appelée *Big Bear Solar Observatory/Machine Learning* (BBSO), proposée par Yuan et al. 2010 [27]. La première étape est de caractériser la région active selon trois paramètres : (1) le flux magnétique absolu, (2) la longueur de la ligne neutre à fort gradient et (3) le total de la dissipation d'énergie magnétique. Avec ces trois paramètres, on fait des prédictions avec deux méthodes d'apprentissage machine supervisée : la régression logistique [28] et les machines à vecteurs de support (SVM) [29].

Malgré toutes ces idées prometteuses, le résultat de cet atelier a montré qu'aucune méthode, incluant les deux exemples présentés ci-dessus, ne performait mieux que les autres de façon significative, incluant les prévisions purement climatologiques. Une prévision climatologique se base sur les statistiques à long terme des éruptions solaires (dont les distributions en loi de puissance de la taille des éruptions présentée à la figure 1.4) pour calculer la probabilité qu'une éruption d'une taille quelconque ait lieu.

Dans le cas des prédictions probabilistes (par exemple : "42% de chance qu'il y aura une éruption  $\geq$  M1.0 dans l'espace de 12 heures"), une des métriques utilisées est le *Brier Skill Score* (BSS), défini comme :

$$\text{BSS} = \frac{\text{MSE}(p_f, o) - \text{MSE}(\langle o \rangle, o)}{0 - \text{MSE}(\langle o \rangle, o)}, \quad (1.4.1)$$

où MSE est l'erreur quadratique moyenne (défini comme  $\text{MSE}(\mathbf{a}, \mathbf{b}) = \langle (\mathbf{a} - \mathbf{b})^2 \rangle$ ),  $p_f$  est la prédiction probabiliste,  $o$  est l'observation ( $o = 0$  si aucune éruption,  $o = 1$  s'il y a une éruption) et  $\langle o \rangle$  est la moyenne des observations. Le score parfait du BSS est donc 1, et un

Method	BSS (M1.0+, 12hrs)
$B_{eff}$	0.03
ASAP	-0.01
BBSO	0.06
$WL_{SG2}$	0.06
NWRA MAG 2-VAR	0.13
$\log(\mathcal{R})$	0.10
GCD	0.03
NWRA MCT 2-VAR	0.14
SMART2	-4.31
Even Statistics, 10 prior	0.10
$WL_{SG2}$	-0.06

**TABLE 1.1.** Synthèse des résultats de prédiction 'all-clear' des éruptions de taille  $\geq$  M1.0 pendant 12 h, tirée de l'article résumant l'atelier 'all-clear' à Boulder : Barnes et al. 2016.

score de 0 indique que  $MSE(p_f, o) = MSE(\langle o \rangle, o)$ , ce qui signifie que la prédiction  $p_f$  fait aussi bien que la prévision climatologique  $\langle o \rangle$ . Un score BSS négatif indique que la performance est pire que la prévision climatologique.

Le tableau 1.1 présente tous les scores BSS pour toutes les équipes pour un des exercices de l'atelier (des prédictions 'all-clear' pour les éruptions de taille  $\geq$  M1.0 pendant les 12 heures suivant l'observation (magnétogramme et carte optique)). On peut donc voir qu'aucune méthode ne produit des résultats particulièrement satisfaisants, le meilleur score n'étant que 0.13 — 13% au-dessus de la prévision 'climatologique'. Cet atelier a donc montré la très grande difficulté de la prédiction des éruptions solaires ainsi que l'importance de l'utilisation d'une métrique commune pour évaluer toutes les méthodes.

Une des plus grandes difficultés dans la prédiction des éruptions solaires de grandes tailles est leur rareté (voir figure 1.4), ce qui diminue grandement l'efficacité des méthodes statistiques. Pour cette raison, ce projet de maîtrise vise justement à développer une méthode alternative aux procédures statistiques habituelles. Notre but est d'utiliser les modèles d'avalanches comme base à une procédure d'assimilation de données pour faire des prédictions sur le temps de déclenchement des éruptions solaires de très grandes tailles. La grande différence ici par rapport aux méthodes proposées dans l'atelier est que toutes nos prédictions passent par les simulations numériques, qui présentent de bonnes capacités prédictives pour les grands évènements. Les observations utilisées sont les séries temporelles de la radiation solaire, comme celles en rayon X du réseau GOES présentées à la figure 1.1b.

Dans ce projet de maîtrise, nous avons tout d’abord analysé les capacités prédictives des modèles d’avalanches et ensuite développé notre procédure d’assimilation de données sur des données synthétiques produites par les modèles d’avalanches eux-mêmes. Quelques résultats importants sont présentés dans l’article inclus dans ce mémoire. Nous avons également inclus une description plus détaillée ainsi qu’une analyse des modèles DMC (brièvement introduits à la section 1.2.3) au chapitre 2.

## 1.5. Contributions

Le projet de recherche présenté dans ce mémoire était déjà commencé avant le début de ma maîtrise. Le modèle LH était bien étudié depuis déjà quelques décennies, et les modèles D avaient été publiés et étudiés par Antoine Strugarek (mon codirecteur de maîtrise) et Paul Charbonneau (mon directeur de maîtrise) en 2014 [15]. De plus, le premier couplage du modèle d’avalanche avec l’assimilation de données par la méthode 4D-Var était déjà développé et publié par Bélanger et al. 2007 [21] (dans le cadre d’études graduées à l’Université de Montréal). Plus tard, Antoine Strugarek a développé la méthode d’assimilation de données par recuit simulé ainsi que l’optimisation de l’algorithme par décomposition en fonction propre [30]. Le point de départ de ma contribution fut donc ce code dû à Antoine Strugarek.

Mes contributions dans ce projet de recherche sont les suivantes (pas nécessairement dans cet ordre) :

- J’ai proposé, implémenté, testé et étudié les modèles DMC (voir chapitre 2), le tout en collaboration avec Paul Charbonneau et en nous inspirant directement du modèle publié par Farhang et al. 2018 [16].
- J’ai proposé et codé des versions automatisées des analyses de caractéristiques et capacités prédictives de chaque modèle. De cette façon, nous avons identifié leurs aspects plus globaux.
- J’ai testé et amélioré la performance de notre protocole d’assimilation de données sur des observations synthétiques. Cela impliquait un changement dans les détails du code ainsi qu’un ajustement des paramètres de l’algorithme de recuit simulé.
- J’ai développé de nombreux outils numériques d’analyses pour étudier la performance des résultats de l’assimilation de données, incluant les prédictions.

L'article inclus dans ce mémoire a été écrit par moi, Paul Charbonneau et Antoine Strugarek. Il sera soumis à la revue *Solar Physics* dans les prochains mois (automne 2020). Ma contribution a été :

- J'ai écrit toutes les sections sauf l'introduction (qui est écrite par Paul Charbonneau). La section 3.4.1 a été co-écrite par moi et Antoine Strugarek. Toutes les sections ont été corrigées à plusieurs reprises par les coauteurs.
- J'ai produit tous les calculs et figures (sauf la figure 3.5 produite par Antoine Strugarek) présentés dans l'article.

Durant mes deux ans à la maîtrise, j'ai présenté mes travaux à plusieurs reprises dans des congrès scientifiques. Dès le début de ma maîtrise, j'avais déjà assez de résultats pour faire une présentation orale à la conférence nationale *Atelier solaire canadien 2018*. J'ai ensuite présenté l'année suivante à la *Rencontre annuelle du Centre de recherche en astrophysique du Québec (CRAQ) 2019*. J'ai également présenté une affiche scientifique à la conférence internationale *Space Climate 7 2019* et au *Symposium Annuel de Physique pour un Avenir en Recherche et en Industrie (SAPHARI) 2019*. Cette affiche a remporté le prix du jury du SAPHARI 2019.





# Chapitre 2

---

## Modèles déterministes à minimisation Monte-Carlo (DMC)

### 2.1. Introduction

Les modèles déterministes à minimisation Monte-Carlo (DMC) sont inspirés d'un modèle d'avalanche publié par Farhang et al. 2018 [16]. Ce dernier est lui-même inspiré du modèle déterministe (Strugarek et al. 2014 [14], aussi présenté dans ce mémoire à la section 1.2.2), mais avec un ajout de minimisation énergétique durant les redistributions et un seuil de stabilité variable, tirée aléatoirement d'une gaussienne. La justification physique de la minimisation est basée sur le fait qu'un système dissipatif instable tend à se déplacer vers l'état d'énergie minimale compatible avec les conditions frontières. Les modèles DMC développés dans le cadre de cette maîtrise sont essentiellement une modification du modèle de Farhang et al. 2018, gardant le principe de minimisation énergétique, mais avec un temps de calcul plus court.

La motivation derrière sa création est d'avoir un modèle à forçage déterministe (comme présenté à la section 1.2.2) tout en ajoutant un nouvel aspect physique prometteur à un modèle déjà hautement simplifié. Nous cherchons également à développer un modèle capable de produire des statistiques plus rapprochées des observations solaires. Par exemple, la distribution des temps d'attentes entre les éruptions solaires suit en général une loi de puissance. Dans le cas des modèles d'avalanches habituels, cette distribution est plutôt exponentielle. Par contre, Farhang et al. 2018 [16] réussissent avec leur modèle à produire une distribution en loi de puissance, comme celle déduite des observations des éruptions, ce qui nous motive davantage à développer un modèle semblable au leur.

## 2.2. Présentation du modèle

Le modèle est défini par un réseau cartésien 2D. Chacun des noeuds contient une variable nodale dénotée par  $A_{i,j}^n$ , où  $n$  est un indice temporel et  $(i,j)$  sont les coordonnées spatiales. Ils sont chacun connectés à leurs 4 voisins les plus près  $(i+1, j)$ ,  $(i-1, j)$ ,  $(i, j+1)$  et  $(i, j-1)$ . Ce réseau est une représentation hautement simplifiée de la torsion magnétique d'une boucle coronale, donc la définition utilisée de l'énergie magnétique du réseau  $E$  est établie comme :

$$E = \sum_k A_k^2, \quad (2.2.1)$$

où on somme sur tous les noeuds. Le forçage de ce modèle est déterministe et se fait en multipliant simultanément tous les noeuds par une valeur légèrement supérieure à 1 :

$$A_{i,j}^{n+1} = A_{i,j}^n \cdot (1 + \epsilon), \quad \epsilon \ll 1. \quad (2.2.2)$$

Ce forçage continue jusqu'à ce qu'au moins un noeud devienne instable. Le critère de stabilité des noeuds dépend de sa courbure locale, une représentation simplifiée du courant électrique ( $J_e = \nabla^2 A$ ) exprimée par :

$$\Delta A_{i,j}^n = A_{i,j}^n - \frac{1}{4} \sum_k A_k^n, \quad (2.2.3)$$

où  $k$  représente les indices de chacun des noeuds voisins du noeud  $(i,j)$ . Le système est stable jusqu'à ce que la courbure  $\Delta A$  d'au moins un noeud dépasse une valeur seuil de  $Z_c$ .

Jusqu'ici, tous les aspects du modèle DMC présentés ci-dessus sont identiques au modèle D décrit au chapitre précédent (voir section 1.2.2). Les aspects suivants sont ceux qui diffèrent du modèle D.

Dans le cadre de cette maîtrise, nous présentons deux versions du modèle DMC. Dans la première, appelée modèle DMC1, la valeur de  $Z_c$  est recalculée à chaque itération et pour chaque noeud. Elle est choisie aléatoirement d'une distribution gaussienne centrée sur  $\bar{Z}$  avec déviation standard  $\sigma$ . Les valeurs utilisées sont  $\bar{Z} = 2$  et  $\sigma = 0.01$ . Ces valeurs sont choisies pour éviter que le déclenchement soit trop aléatoire (donc  $\sigma \ll Z_c$ ) et pour s'assurer que la taille de l'incrément soit beaucoup plus petite que le seuil de déclenchement (donc  $\bar{Z} \gg A_{i,j}^n \epsilon$ ). Dans le cas de la deuxième version, appelée modèle DMC2, le seuil est fixe comme dans le cas des modèles D. Ceci revient à poser  $\bar{Z} = 2$  et  $\sigma = 0$ . Ce modèle est

donc en principe plus déterministe que tous les autres modèles présentés dans ce mémoire. Ses seuls aspects stochastiques sont dans la condition initiale et la règle de redistribution. La règle de redistribution présentée aux paragraphes suivants est la même pour les modèles DMC1 et DMC2.

Quand un noeud est instable, il redistribuera une partie de sa variable nodale à ses quatre voisins de façon non uniforme, mais conservatrice (aucune perte de variable nodale durant la redistribution). La part que chacun des voisins recevra durant la redistribution sera choisie pour minimiser le plus possible l'énergie du réseau après la redistribution (et donc en maximisant l'énergie libérée par le réseau). Ceci est une façon de simuler le fait qu'un système dissipatif instable aura tendance à évoluer vers un état d'énergie minimale compatible avec les conditions frontières et/ou autres contraintes globales du système. Notons que dans un tel système, le minimum absolu d'énergie n'est pas nécessairement atteignable en raison des contraintes topologiques imposées par les conditions frontières. La règle de redistribution s'écrit donc comme :

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5} Z_c , \quad (2.2.4)$$

$$\begin{aligned} A_{i+1,j}^{n+1} &= A_{i+1,j}^n + \frac{1}{5} \frac{r_1}{a} Z_c , \\ A_{i-1,j}^{n+1} &= A_{i-1,j}^n + \frac{1}{5} \frac{r_2}{a} Z_c , \\ A_{i,j+1}^{n+1} &= A_{i,j+1}^n + \frac{1}{5} \frac{r_3}{a} Z_c , \\ A_{i,j-1}^{n+1} &= A_{i,j-1}^n + \frac{1}{5} \frac{r_4}{a} Z_c , \end{aligned} \quad (2.2.5)$$

où  $r_1, r_2, r_3$  et  $r_4$  sont des valeurs aléatoires entre 0 et 1 et  $a = r_1 + r_2 + r_3 + r_4$ . Pour un noeud instable donné, on génère  $N_{min}$  combinaisons de  $r_1, r_2, r_3$  et  $r_4$ . Chacune des valeurs de  $r$  est tirée aléatoirement d'une distribution uniforme entre 0 et 1 inclusivement. Pour chacune des combinaisons, on calcule l'énergie qui serait libérée par cette redistribution. On choisit finalement la combinaison qui produit la plus grande libération d'énergie pour faire la redistribution. Le résultat est une sorte de minimisation énergétique imparfaite à la Monte-Carlo. Cette méthode produit une minimisation statistiquement isotrope, c'est-à-dire qu'elle se comporte de la même façon pour tous les noeuds voisins impliqués. Ceci n'est pas le cas

pour la minimisation proposée par le modèle de Farhang et al. 2018 [16], qui ajuste un seul paramètre libre (correspondant à  $r_4$  dans l'équation 2.2.5) pour chaque redistribution.

Le calcul de l'énergie libérée par redistribution (ce qu'on veut maximiser) est exprimé par :

$$e_{i,j}^n = \frac{8}{5} Z_c \Delta A_{i,j}^n + \frac{2}{5} Z_c \sum_k \left( A_k^n \left( 1 - \frac{r_k}{a} \right) - \frac{16 + (r_k/a)^2}{5} Z_c \right), \quad (2.2.6)$$

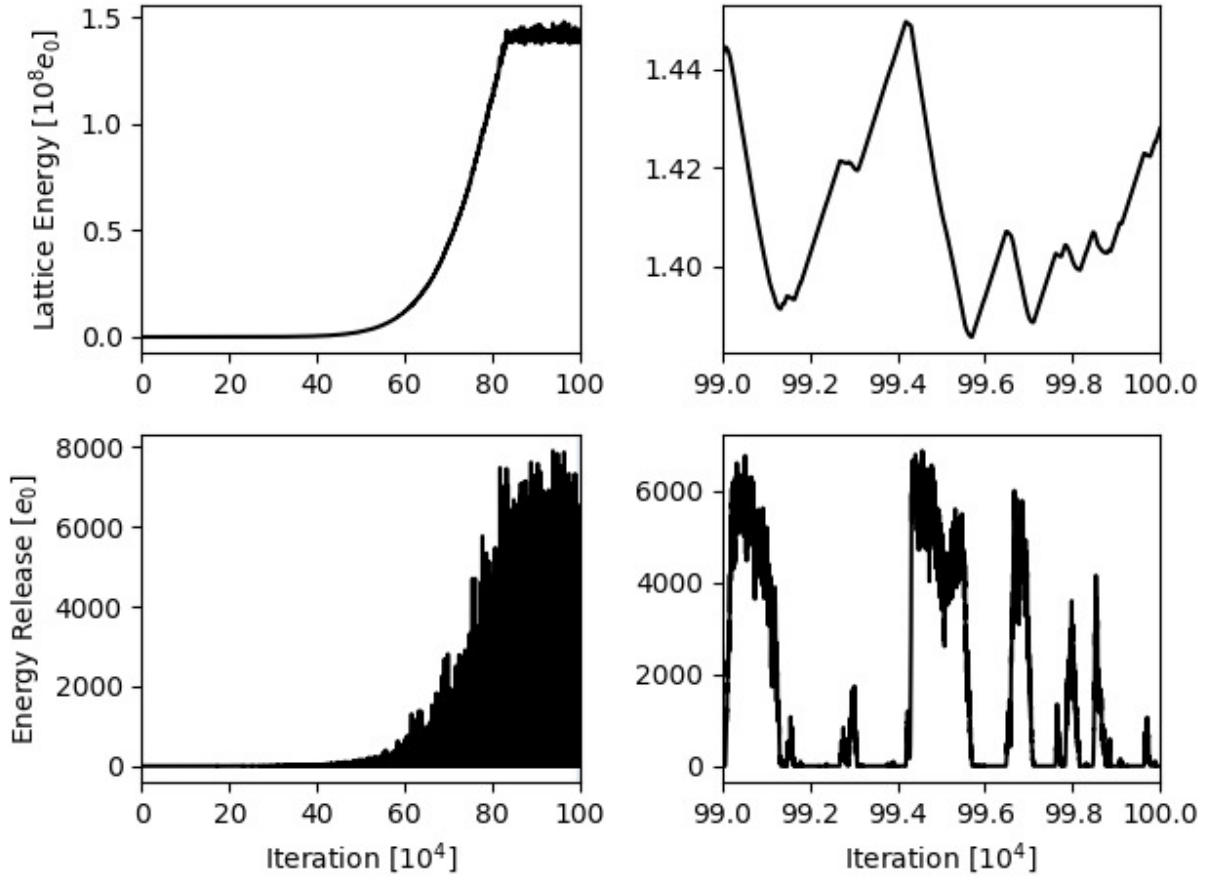
où la somme sur  $k$  est faite sur tous les quatre voisins.

Le nombre  $N_{min}$  de combinaisons d'essais pour la redistribution est un paramètre libre à choisir avant la simulation. Si  $N_{min} = 1$  pour DMC1, on retombe au modèle déterministe en version conservative avec seuil aléatoire, présenté dans Strugarek et al. 2014 comme le modèle 'C5' [14]. Plus  $N_{min}$  est élevé, plus le modèle DMC1 se distingue du modèle C5 par sa capacité à maximiser l'énergie libérée par redistribution.

### 2.3. Caractéristiques des avalanches

La figure 2.1 présente des séries temporelles de l'énergie du réseau (haut) et de l'énergie libérée (bas) du modèle DMC1 à partir d'une condition initiale où chaque variable nodale est tirée d'une distribution aléatoire entre 0 et 1 (sauf pour les noeuds frontières qui sont fixés à 0). Dans cet exemple, le système prend environ  $80 \times 10^4$  itérations avant d'atteindre un état statistiquement stationnaire, où le système en moyenne perd autant d'énergie par les avalanches qu'il en gagne par le forçage. Cet état d'équilibre est l'état de la criticalité auto-organisée, ou état SOC. Dans cet état, l'énergie du réseau fluctue autour d'une valeur moyenne (figure 2.1 haut droit) et les énergies totales libérées par les avalanches prennent des valeurs qui s'étendent sur plusieurs ordres de grandeur (figure 2.1 bas droit).

La figure 2.2 présente des détails spatiaux de l'évolution d'une avalanche de grande taille, produit encore ici par le modèle DMC1. Durant les premières itérations de cet exemple, l'avalanche se produit à petite échelle dans le coin supérieur gauche (A). Ensuite, l'étendue spatiale augmente rapidement (B) jusqu'à ce qu'elle couvre presque tout le réseau (C). Il diminue en intensité (D), remonte, et finalement termine avec une présence particulière de noeuds instables sur les bords du réseau (E). La figure du coin en bas à droite montre le nombre total de redistributions par noeud sur toute l'avalanche. On peut donc voir que le réseau a tendance à produire des noeuds instables sur les côtés, plutôt que sur les coins et au



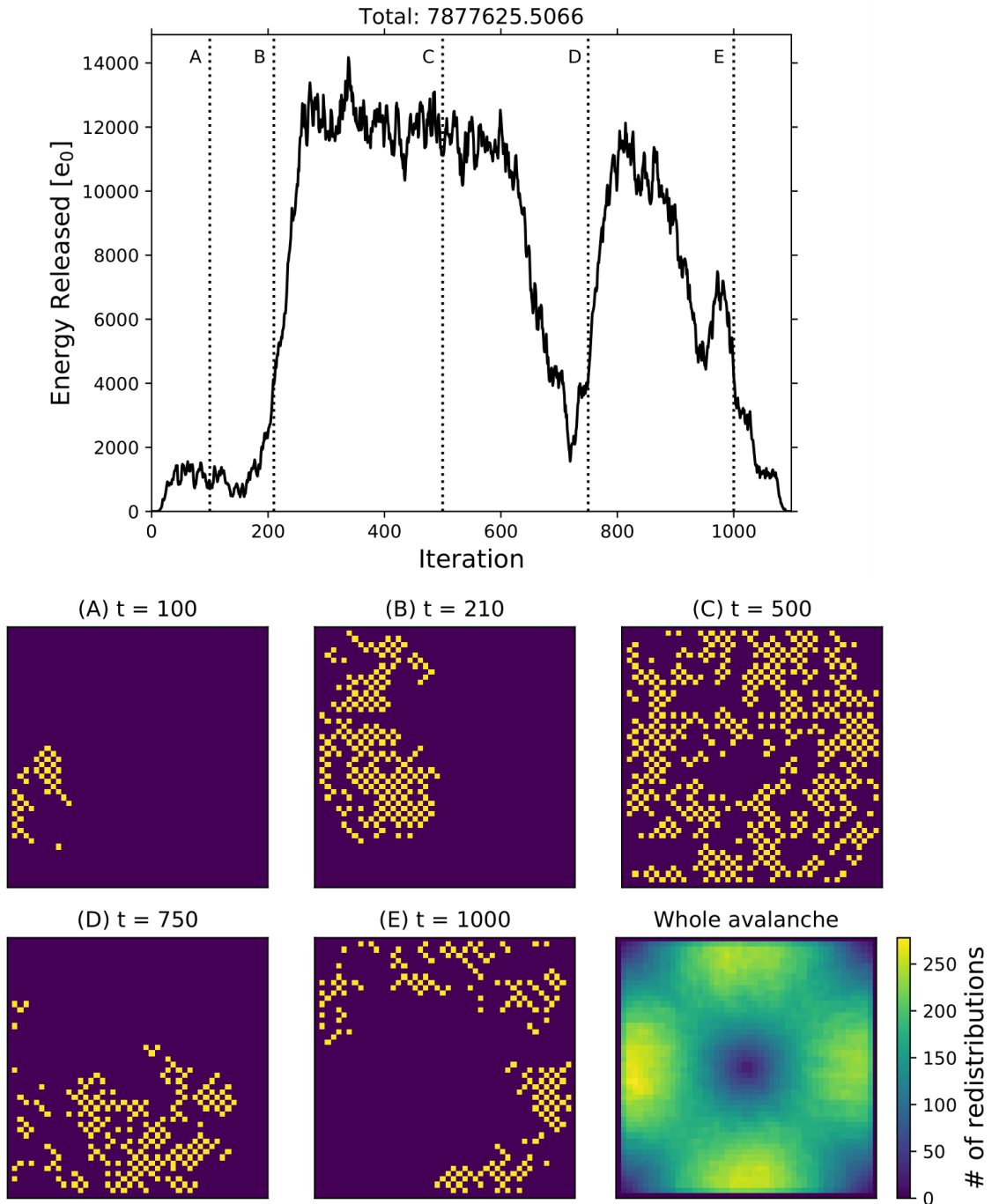
**FIGURE 2.1.** Énergie du réseau (haut) et énergie libérée (bas) à partir d’une condition initiale où chaque variable nodale est tirée d’une distribution aléatoire entre 0 et 1. Cette figure est produite avec le modèle DMC1 à  $N_{min} = 5$ ,  $\epsilon = 10^5$  et avec dimensions  $48 \times 48$ .

centre. Nous avons vérifié plusieurs autres exemples d’avalanches de très grande taille, soit comparable aux dimensions du réseau, et ces patrons sont toujours présents.

Les avalanches produites par le modèle DMC2 ont essentiellement le même comportement que ceux du modèle DMC1, présentés aux figures 2.2 et 2.1. Pour cette raison, nous avons choisi de seulement présenter ceux du modèle DMC1.

## 2.4. Statistiques des avalanches

Dans cette section, nous étudions les statistiques de 4 mesures qui caractérisent les avalanches. La première est l’énergie totale libérée  $E/e_0$ , calculée par l’aire sous la courbe d’une série temporelle de l’énergie libérée durant une avalanche (comme celle présentée à la figure 2.2, panneau du haut). La deuxième est la valeur maximale de l’énergie libérée durant une



**FIGURE 2.2.** Haut : Énergie libérée durant une avalanche de grande taille avec le modèle DMC1 ( $48 \times 48$  noeuds), avec  $N_{min} = 25$ . Les lignes pointillées sont placées à des itérations de références utilisées dans la figure du bas. Bas : photos instantanées du réseau durant les itérations en pointillées de la figure du haut. En jaune sont les noeuds instables et en violet sont les noeuds stables. Le nombre de redistributions total de chaque noeud durant l'avalanche au complet est présenté à l'image en bas à droite.

avalanche, dénotée par  $P/e_0$  (environ 14000 dans le cas de la figure 2.2). La troisième est la durée en itération des avalanches, dénotée par  $T$  (environ 1100 dans la figure 2.2). Le dernier est le temps d’attente entre la fin d’une avalanche et le début du prochaine, dénoté par  $\Delta T$ .

Les statistiques à long terme des avalanches des deux modèles DMC, incluant l’énergie totale libérée ( $E/e_0$ ), le pic énergétique ( $P/e_0$ ), la durée ( $T$ ) et les temps d’attente ( $\Delta T$ ) des avalanches, sont présentées à la figure 2.3. Ces statistiques sont toutes calculées dans l’état SOC du modèle (voir figure 2.1), ce qui représente un état statistiquement stationnaire de la simulation. Cette figure met en évidence la distribution en loi de puissance pour les valeurs de  $E/e_0$ ,  $P/e_0$  et  $T$  (on remarquera que les échelles sont en log-log), exprimé sous la forme :

$$f(X) = f_0 X^{-\alpha_X} , \quad (2.4.1)$$

où  $f$  est la fonction de probabilité,  $f_0$  est un facteur de normalisation et  $X$  est la variable en question (ici  $E/e_0$ ,  $P/e_0$  ou  $T$ ). Les temps d’attente suivent plutôt une distribution exponentielle, sous la forme :

$$f(\Delta T) = f_0 e^{-\alpha_{\Delta T} \Delta T} . \quad (2.4.2)$$

Ces distributions sont similaires aux autres modèles d’avalanches présentés dans l’introduction de ce mémoire (modèle LH et les modèles D). On remarque également que les modèles DMC semblent s’éloigner d’une distribution en loi de puissance pour les très grosses avalanches ; on voit une petite bosse à la partie droite des distributions de  $E/e_0$  et  $P/e_0$ , un artefact numérique du modèle. Ces bosses sont également présentes pour les modèles déterministes, particulièrement en version conservative [14]. Les valeurs calculées des  $\alpha_X$  pour  $E/e_0$ ,  $P/e_0$  et  $T$  sont assez semblables aux modèles déterministes, particulièrement ceux du modèle C5 publiées dans Strugarek et al. 2014 [14]. Par contre, elles sont assez éloignées des valeurs publiées par Farhang et al. 2018 [16] avec leurs modèles. Par exemple, ils trouvent un  $\alpha_E$  d’environ 2 pour un modèle semblable au DMC1 mais avec une minimisation analytique et anisotrope ainsi qu’une définition de l’énergie du réseau défini par  $E = 2 \sum (A \cdot \Delta A)$ . De plus, leur distribution des temps d’attente suit plutôt une loi de puissance.

Les deux modèles DMC semblent se comporter de façon presque identique au niveau des statistiques de  $E/e_0$ ,  $P/e_0$  et  $T$ . Ils se distinguent par contre dans les statistiques des temps d’attente  $\Delta T$ , avec le modèle DMC1 (en bleu) ayant une pente visiblement plus raide que

celle du modèle DMC2 (en rouge). Le seuil de stabilité stochastique, présent dans le modèle DMC1, semble privilégier davantage les courts temps d'attente par rapport au plus long.

Dans les analyses statistiques qui suivent, nous allons étudier les distributions de l'énergie totale libérée ( $E/e_0$ ). Puisque les deux modèles DMC présentent des distributions quasi identiques de  $E/e_0$ , nous allons seulement présenter les résultats du modèle DMC1.

La figure 2.4 présente l'effet de la taille du réseau sur la distribution des énergies libérées par les avalanches ( $E/e_0$ ). On remarque que la pente des distributions ( $\alpha_E$ ) change très peu par rapport à la taille du réseau. Toutefois, plus le réseau est grand, plus la distribution atteint des événements de grande taille. Par exemple, le réseau à  $32^2$  noeuds s'arrête aux événements de taille  $\sim 10^6$ , tandis que le réseau à  $80^2$  noeuds s'arrête à  $\sim 10^8$ , une différence de deux ordres de grandeur. Ceci est tout à fait caractéristique pour un système dans l'état SOC, qui génère des événements qui sont invariants en échelle, limités que par la taille du réseau ainsi que le minimum possible d'énergie libérée par un noeud.

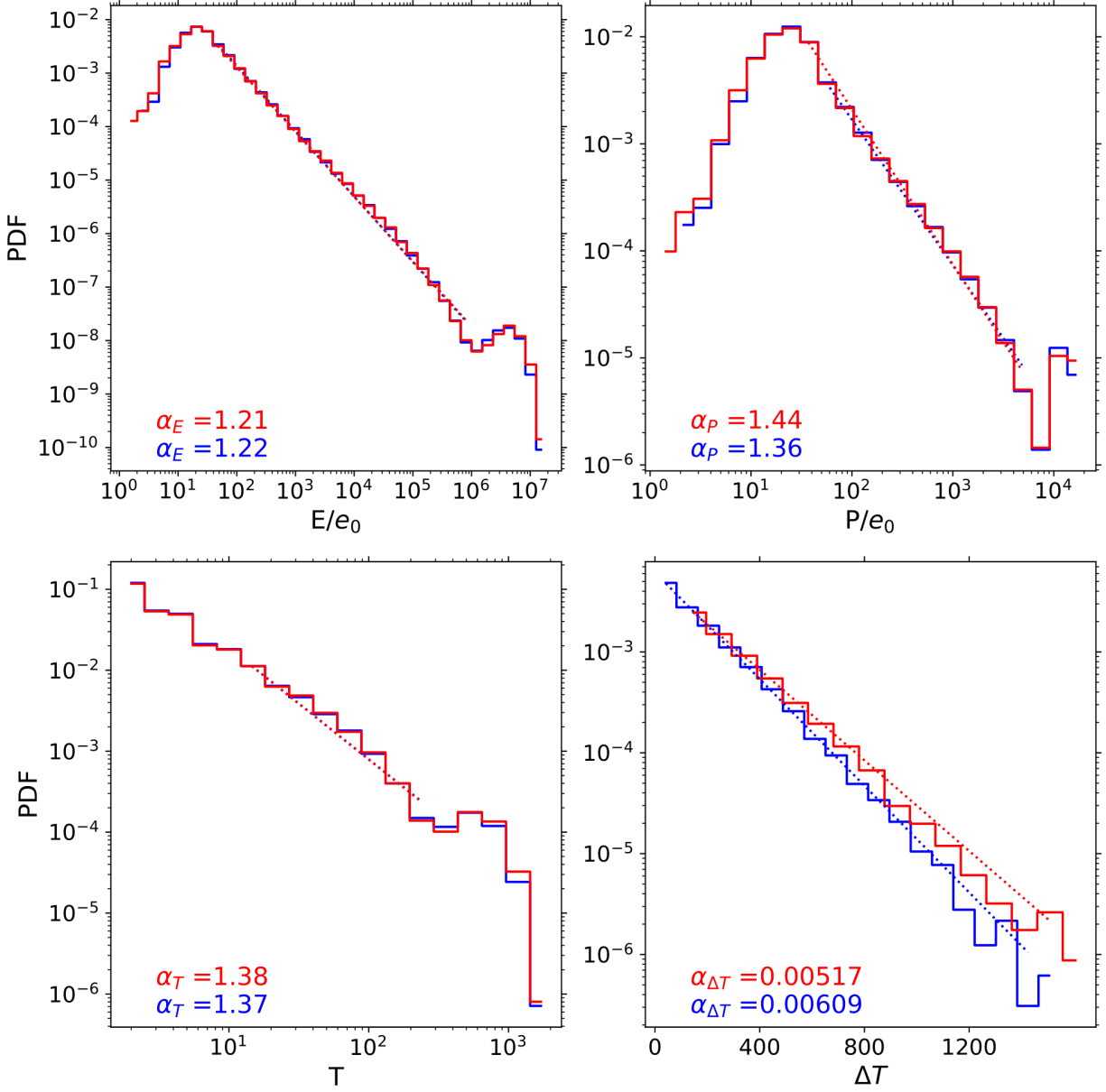
La petite bosse à la partie droite de la distribution se décale vers la droite plus la taille du réseau augmente. Ceci semble confirmer que, pour les règles de redistributions et stabilité utilisées ici, les avalanches couvrant des régions de surface  $\sim N^2/4$  (où  $N^2$  est la taille du réseau) sont artificiellement favorisées par les effets de bords.

La figure 2.5 présente l'effet de  $N_{min}$  (nombre d'itérations pour produire la minimisation de chaque redistribution) sur la distribution de la taille des avalanches ( $E/e_0$ ). On rappelle ici qu'à  $N_{min} = 1$  pour le modèle DMC1, on se retrouve avec le modèle déterministe version conservatrice (modèle C5 dans Strugarek et al. 2014 [14]). La présence des très petites avalanches ( $< 3 \times 10^1$ ) est inversement proportionnelle à  $N_{min}$ . Plus la minimisation est efficace, plus chaque redistribution tendra à libérer plus d'énergie et donc moins il y aura de très petites avalanches à une fraction de  $e_0$ . Encore ici, les statistiques des événements ne semblent pas particulièrement changer par rapport à  $N_{min}$ .

## 2.5. Capacités prédictives

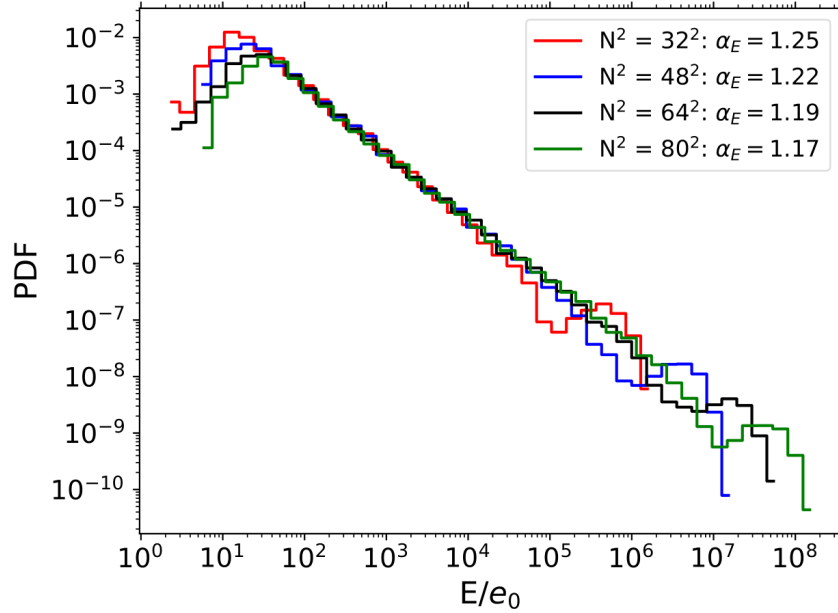
Un des buts principaux du projet de recherche présenté dans ce mémoire est d'utiliser les modèles d'avalanches pour faire des prédictions par assimilation de données. On cherche donc à étudier les capacités prédictives des modèles pour motiver son utilisation dans notre



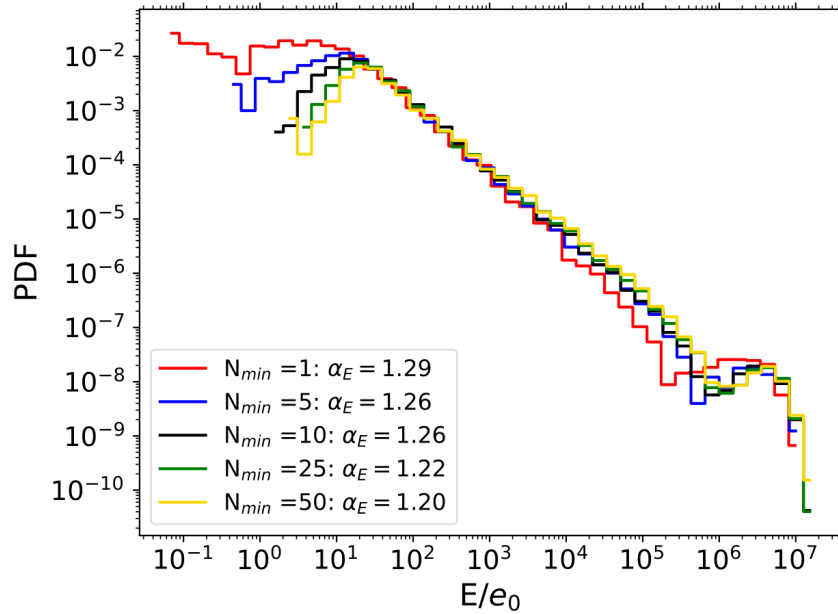


**FIGURE 2.3.** Lignes pleines : distributions de l'énergie totale libérée ( $E/e_0$ ), le pic énergétique ( $P/e_0$ ), la durée ( $T$ ) et les temps d'attente ( $\Delta T$ ) des avalanches produites pendant  $10^7$  itérations des modèles **DMC1** (en bleu) et **DMC2** (en rouge), chacun avec  $N_{min} = 25$  et un réseau de  $48 \times 48$ . Lignes pointillées : meilleur ajustement d'une loi de puissance pour les distributions de  $E$ ,  $P$  et  $T$  et d'une distribution exponentielle pour  $\Delta T$ . Les valeurs de  $\alpha$  associées à chaque meilleur ajustement sont incluses au bas de chaque figure.

éventuel protocole de prédiction. La figure 2.6 présente une distribution des temps de déclenchement des avalanches de grandes tailles ( $E/e_0 > 10^5$ ) pour 10000 simulations des deux modèles DMC ( $N_{min} = 25$  et de dimension  $48 \times 48$ ), toujours à partir de la même condition initiale, mais avec différentes séquences de nombres aléatoires.



**FIGURE 2.4.** Distributions de l'énergie totale libérée ( $E/e_0$ ) du modèle DMC1 à différentes tailles de réseau, avec  $N_{min} = 25$  et pour  $10^7$  itération chacun. Les petites bosses aux parties droite sont des artefacts numériques causés par les effets de bords.

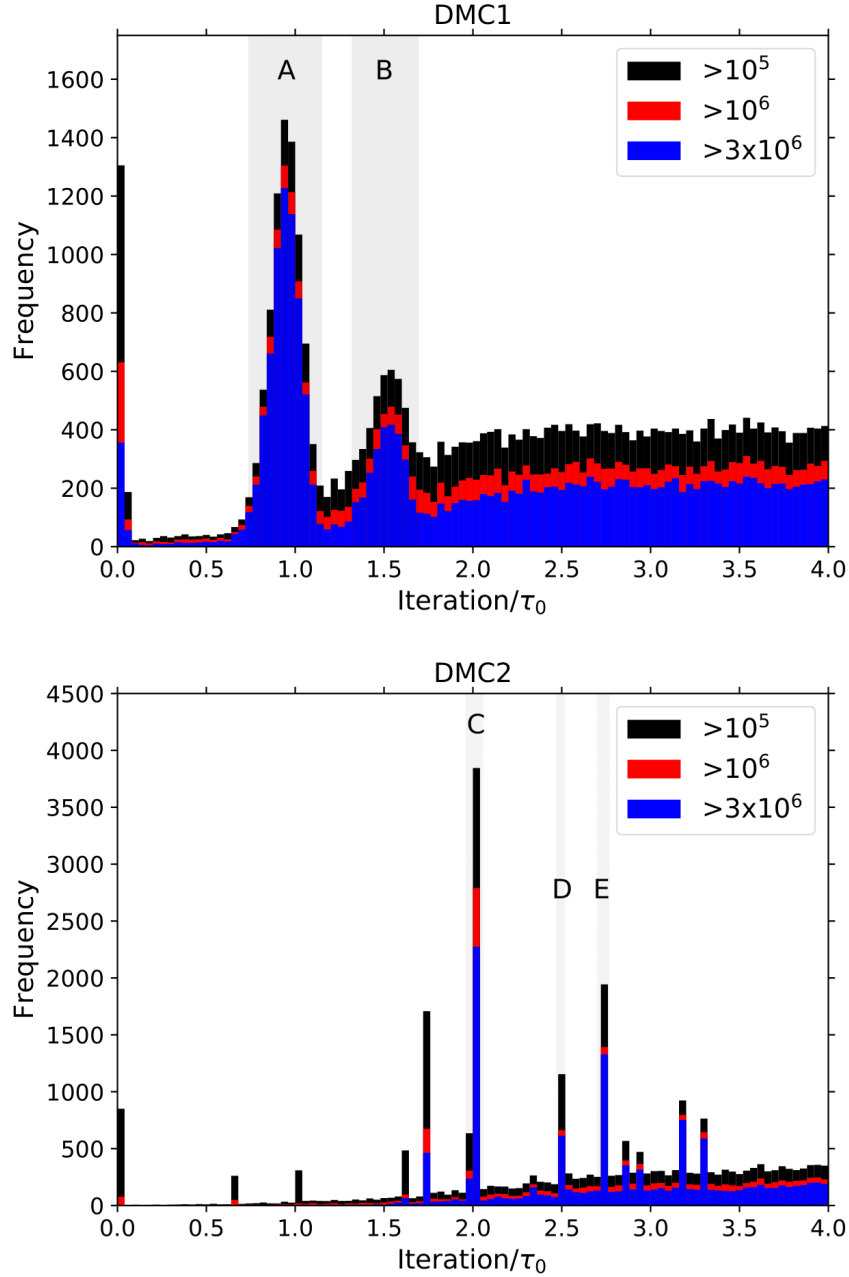


**FIGURE 2.5.** Distributions de l'énergie totale libérée ( $E/e_0$ ) du modèle DMC1 pour différent nombre d'itérations de minimisation  $N_{min}$ , avec taille de réseau  $48 \times 48$  et pour  $10^7$  itération chacun.

Dans le cas du modèle DMC1, on remarque un premier pic mince aux toutes premières itérations, ce qui est attendu pour des modèles à forçage déterministe qui augmentent les courbures des noeuds ( $\Delta A$ ) simultanément. Nous allons donc omettre ce premier pic dans l'analyse qui suit. Plus loin, on a deux régions (dénotées par A et B) avec des formes en fonction gaussienne, conséquence directe du seuil gaussien dans le seuil de stabilité. On a donc des évènements de grandes tailles qui tendent à se produire autour de deux temps distincts (A et B), malgré les changements de séquence de nombres aléatoires. Ceci est une indication que le temps de déclenchement des avalanches de grandes tailles dépend particulièrement de la condition initiale, qui contient des patrons de stress formés par des avalanches précédentes. De plus, 50.7% des simulations qui ont produit un évènement de grande taille dans la région B ont aussi produit un grand évènement dans la région A. Ces deux pics gaussiens sont donc raisonnablement bien corrélés. Après ces deux régions, on tombe dans un régime plutôt aléatoire ou il n'y a plus de temps de déclenchement préférentiel. Ceci indique que pour cette condition initiale, on a un temps de cohérence d'environ  $1.75\tau_0$  avant qu'on perde les capacités prédictives du modèle. Des expériences supplémentaires ont montré que ce temps de cohérence varie d'une condition initiale à l'autre, allant d'environ 0.5 à  $1.9\tau_0$ .

Dans le cas du modèle DMC2, le résultat diffère du modèle DMC1 sur plusieurs aspects. Le temps de cohérence est visiblement plus long, jusqu'à environ  $3.25\tau_0$ . Encore ici, ce temps de cohérence varie par rapport à la condition initiale, de environ 0.5 à  $3.75\tau_0$ . Les pics sont beaucoup plus minces, dû au seuil stabilité fixe. Il y a également beaucoup plus de pics qui ressortent du bruit. La corrélation entre les pics semble dépendre de leur proximité temporelle. Dans cet exemple, 46.7% des simulations qui ont déclenché une grosse avalanche à la fenêtre D en ont aussi déclenché une à la fenêtre C. Par contre, il y a seulement 23.5% pour la fenêtre E par rapport à la fenêtre C. Cet exemple suggère que cette corrélation temporelle entre les grosses avalanches diminue avec l'intervalle de temps.

Les évènements de taille moyenne ( $< 3 \times 10^6$ ) semblent être plus présents pour chaque pic par rapport au modèle DMC1 qui contient des pics plus dominés par les très grosses avalanches (en bleu). Le niveau de bruit des deux modèles est à peu près au même niveau à environ 400.



**FIGURE 2.6.** Distributions des itérations de déclenchement des avalanches de grandes tailles ( $E/e_0 > 10^5$ ), pour les modèles DMC1 (haut) et DMC2 (bas) à  $48 \times 48$  noeuds et avec  $N_{min} = 25$ . Les itérations sont normalisées par la médiane des temps d’attente entre les avalanches de grandes tailles  $\tau_0$ . Les sections A à E en gris sont ajoutées pour faciliter la discussion dans le texte.

## 2.6. Conclusion

Nous avons présenté un nouveau modèle d’avalanche développé dans le cadre de cette maîtrise et fortement inspiré du modèle publié par Farhang et al 2018 [16]. Ses statistiques

dans l'état SOC sont très semblables aux modèles D, avec quelques exceptions particulièrement aux avalanches de très petites tailles. L'ajout d'une minimisation à la Monte-Carlo ajoute un réalisme physique aux modèles DMC, ainsi qu'une diminution de la stochasticité dans la règle de distribution. Les capacités prédictives du modèle sont également prometteuses, avec les temps de déclenchement des avalanches de grandes tailles particulièrement dépendant sur les patrons de stress formés par les avalanches précédentes.

Par contre, nous n'avons pas réussi à produire des indices d'exposants  $\alpha_E$  et  $\alpha_P$  qui se rapprochent à ceux des distributions observées par le soleil, un des buts derrière le développement de ce modèle. Nous croyions à priori que ce modèle nous rapprocherait de quelques valeurs élevées publiées par Farhang et al. 2018 [16] ( $\alpha_E = 2.053 \pm 0.002$  et  $\alpha_P = 1.843 \pm 0.002$ ). Les trois différences principales entre le modèle DMC1 et le modèle 'Case A' de Farhang et al. 2018 sont (1) l'anisotropie de leur règle de redistribution, (2) leur minimisation globale et analytique (plutôt que par une méthode Monte-Carlo dans les modèles DMC) et (3) une définition différente de l'énergie du réseau :  $E = 2 \sum (A \cdot \Delta A)$  (où  $\Delta A$  est la courbure définie à l'équation 2.2.3). Il serait intéressant pour un projet futur d'explorer ces différents aspects et vérifier leurs effets sur les statistiques des avalanches.



# Chapter 3

---

## Article: Predicting Solar Flares by Data Assimilation in Avalanche Models. II. Efficient assimilation and predictive potential

C. Thibeault,<sup>1</sup> A. Strugarek,<sup>2</sup> P. Charbonneau,<sup>1</sup>

<sup>1</sup>*Département de physique, Université de Montréal, C.P. 6128 Succ. Centre-ville, Montréal H3C 3J7, Canada*

<sup>2</sup>*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France*

Cet article sera envoyé à *Solar Physics* (SoPh).

### Abstract

The prediction of solar flares is still a significant challenge in space weather research, with no techniques currently capable of producing reliable forecasts above climatological estimates. In this paper, we present an alternative forecasting technique using data assimilation coupled with computationally inexpensive cellular automaton models called avalanche models. Our data assimilation algorithm uses the simulated annealing method to find an optimal initial condition that reproduces well a set of observations. We present and empirically analyze the predictive capabilities of three avalanche models, namely the Lu and Hamilton model (LH)

and two deterministically-driven models (D). Despite their stochasticity, we have shown that deterministically-driven models contain some degree of memory between simulated events, a well-needed condition for data assimilation. We also present our new data assimilation algorithm and its success in assimilating synthetic observations produced by the avalanche models themselves.

**Keywords:** Flares, Forecasting, Avalanche models, Data assimilation, Self-organized criticality

### 3.1. Introduction

The accurate prediction of large solar flares is one of the desired milestones in ongoing space weather research efforts. A number of semi-empirical and/or data-driven techniques have been designed over the years, with various degrees of success. Nonetheless, none of them seem so far to be consistently doing very much better than so-called climatological forecasting, which consists in simply predicting according to empirically constructed statistical distributions. The flare forecasting exercise reported upon in Barnes et al. 2016 [23] (see also Leka et al. 2019a [31] and Leka et al. 2019b [32]) is particularly interesting in quantifying the relative merits of these various techniques.

That reliable flare forecasting should prove hard to achieve is not at all surprising. The first major challenge is related to the extremely wide range of spatial and temporal scales characterizing the flaring phenomenon. This makes brute force approaches such as full magnetohydrodynamical simulations extremely challenging, if not unpractical (see Cheung et al. 2019 [33] for the closest yet). Empirical/statistical techniques face another related challenge: the observed size distributions of flares takes the form of a steep power law (e.g. Denis 1985 [34]; Aschwanden 2011 [3]), meaning that large flares, which are the most important to accurately forecast, are rare. Consequently, the observational statistics are dominated by small events, implying in turn that methods such as machine learning have (relatively) few examples of large flares available for training purposes.

The power-law shape of flare properties (sizes, duration, energy) is an indication of scale invariance, which is a strong hint as to the inner workings of the flare phenomenon. Furthermore, similar power laws are constructed from stellar flare data (see Aschwanden et al. 2008 [35] and Namekata et al. 2017 [36]), pointing to some universality in the flaring



process. This observed scale invariance has led to the consideration of flares as possibly arising from an avalanche of small reconnection events cascading across coronal loops or other magnetic structures. Lu & Hamilton 1991 [12, 13] have designed the first lattice models capturing this avalanching process, with numerous variations on the theme subsequently proposed as explanatory frameworks for flares in general (for a review, see Charbonneau et al. 2001 [37] and Aschwanden et al. 2016 [38]). A key aspect of the Lu & Hamilton proposal is that their lattice, an idealized representation of a coronal structure loading magnetic energy, autonomously reaches a critical state in response to slow external forcing, from which results scale-invariant impulsive energy release. This represents an instance of self-organized-criticality (hereafter SOC; see Bak et al. 1987 [39] or Jensen 1998 [40]), now understood to be a robust generator of scale-invariant behavior, including power-law size distribution for energy release events in general. Indeed, the SOC framework has been applied to phenomena as diverse as flares, earthquakes, and geomagnetic substorms, to name but a few [3].

In the context of solar flare, the SOC hypothesis is buttressed by a physical scenario due to Parker [8], according to which photospheric fluid motions twist and braid the footpoints of magnetic fieldlines within coronal loops, leading to the inexorable buildup of magnetic tangential discontinuities (or current sheets) becoming unstable and releasing thermal energy in the form of what Parker dubbed nanoflares. Although originally designed as a model for coronal heating, Parker’s scenario contains all required elements for scale-invariant release of magnetic stress by cascades of localized reconnection events within a coronal loop. This is the physical picture captured by the lattice models of the general type introduced by Lu & Hamilton.

All avalanche models proposed in the flare context involve stochastic elements, either in lattice loading, avalanche triggering, or internal redistribution in the course of avalanches. This may lead one to expect that such models, even if they properly capture statistical flaring behavior, should be useless for prediction of individual events. For any of the aforementioned avalanche models, the triggering and unfolding of the numerous small avalanches, equivalent to the more frequent, smaller flares, are indeed strongly affected by the stochastic elements embedded in the model, and so are truly unpredictable.

However, and perhaps counter-intuitively, this may not be the case for the larger avalanches. In lattice models, large avalanches release large-scale stress patterns having

built up in the lattice in response to slow external forcing, but also via the unfolding of earlier avalanches, especially the larger ones. Once they reach the self-organized critical state, lattice models exhibit long range spatiotemporal correlations [40], so that even in the presence of truly stochastic elements, the system does not behave completely stochastically. What this implies is that information useful for predicting future avalanching behavior is contained, in principle, in past avalanching behavior. How to extract that information then becomes the key challenge.

Data assimilation is one attractive possibility. In short, the idea is to use past observations to adjust the state of an underlying physical model to best reproduce these observations, yielding an “optimal” initial condition for model-based forecasting. Such procedures have been extensively used for (earthly) weather forecasts. Belanger et al. 2007 [21] and Strugarek & Charbonneau 2014 [15] have presented proof-of-concept and exploratory results on the use of various data assimilation methods in the context of avalanche models for flares. These latter authors, in particular, have shown example of successful assimilation of both synthetic data as well as GOES X-Ray flux time series. This prior work is the motivation and starting point of the work reported upon in this paper.

In this article, we first introduce in section 3.2 the three avalanche models used in our analysis, namely the Lu and Hamilton model (LH) and two deterministically-driven models (D). In section 3.3, we present empirical analyses and comparison of the predictive capabilities of these three models. Finally, in section 3.4, we present our data assimilation protocol and our results in assimilating synthetic observations produced by the avalanche models.

### **3.2. Avalanche models for solar flares**

Avalanche models are cellular automata that can be used to represent the flaring solar corona in a highly simplified view. Despite their simplicity, they reproduce reasonably well the power-law size distributions of solar flares and draw physical support from the nanoflare picture proposed by Parker [8]. In comparison to MHD models of coronal magnetic fields, avalanche models are computationally inexpensive and thus, in principle, can be more easily coupled with data assimilation methods. For detailed reviews on avalanche models for solar flares, we refer the reader to Charbonneau et al. 2001 [37] and Aschwanden et al. 2016 [38].

### 3.2.1. The Lu and Hamilton Model

The first sandpile model of solar flares was published by Lu and Hamilton in 1991 [12]. As such, they are often considered as the canonical avalanche models for solar flares, which we denote as "LH" hereafter. Here, we use a 2D Cartesian grid version of this model, with  $48 \times 48$  nodes each containing a real number denoted  $A_{i,j}^n$ , where  $n$  is a discrete time index and  $(i,j)$  are its spatial coordinates. The driving of the lattice is done by adding a small increment value of  $\delta A$  on a randomly selected node. The value of  $\delta A$  is randomly selected through a uniform distribution between  $-0.2$  and  $0.8$ . The boundary of the lattice is kept at  $A = 0$ .

The driving of the lattice is done once every iteration until one node is considered unstable. The stability criterion of a node is defined by its local curvature, expressed as:

$$\Delta A_{i,j}^n \equiv A_{i,j}^n - \frac{1}{4} \sum_{k=1}^4 A_k^n, \quad (3.2.1)$$

where  $k$  denotes each nearest neighbour of the node  $(i,j)$ . Once  $\Delta A_{i,j}^n$  exceeds a predetermined threshold  $Z_c$ , the driving stops and redistribution occurs to regain stability. We therefore explicitly separate the timescales of the driving and release of lattice energy. When redistribution occurs, an amount of nodal value of the unstable node  $(i,j)$  is redistributed evenly to each neighbour according to the following rules:

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5}Z, \quad (3.2.2)$$

$$A_{i\pm 1,j\pm 1}^{n+1} = A_{i\pm 1,j\pm 1}^n + \frac{1}{5}Z, \quad (3.2.3)$$

where  $Z \equiv Z_c \Delta A_{i,j}^n / |\Delta A_{i,j}^n|$ . This process is deterministic and conservative, meaning that the total amount of nodal value is conserved during the redistribution. The redistribution of an unstable node will occasionally cause neighbouring nodes to become unstable. This leads to chain reactions of redistribution events, which we call *avalanches*, the model equivalent of a solar flare. The lattice eventually recovers a global stability, and the forcing starts over until the next avalanche.

Since the nodal value  $A_{i,j}^n$  is generally associated with a measure of the magnetic twist in a coronal loop, the usual measure of the magnetic energy of the lattice is defined by:

$$E^n = \sum_{i,j} A_{i,j}^2, \quad (3.2.4)$$

where we sum on all nodes of the lattice. Each redistribution will therefore cause a decrease in lattice energy by the amount:

$$\Delta e_{i,j}^n = \frac{4}{5e_0} \left( 2 \frac{|\Delta A_{i,j}^n|}{Z_c} - 1 \right) Z_c^2, \quad (3.2.5)$$

where  $e_0$  is the energy liberated by a single node exceeding the stability threshold by an infinitesimal amount, expressed as:

$$e_0 = \frac{4Z_c^2}{5}. \quad (3.2.6)$$

The total energy released by the lattice per iteration ( $\Delta E_r^n$ ) is equal to the total energy loss of all redistributing nodes. Note that other physical interpretations of the lattice nodal variable can lead to other measures of energy release in avalanche models (for a more in-depth discussion, see chapter 12 in Aschwanden's book *Self-organized criticality systems* [9]).

Sandpile models describe the evolution of a system on two characteristic timescales, the forcing timescale corresponding to the typical evolution time of an active region, of the order of hours to days, and the eruption timescale corresponding to the impulsive flare of the order of seconds. In that sense, the total energy of an avalanche in a sandpile model can be compared to the peak energy of a solar flare, as the decrease of the X-ray flux after the flare associated with the cooling of the flaring region is not accounted for in the sandpile model.

### 3.2.2. The Deterministically Driven Sandpile Model

A deterministically driven model in the context of solar flare simulations was first proposed by Strugarek et al. 2014 [14], notably inspired by deterministic driving schemes used for seismic faults [41]. The idea was to shift the stochastic aspects towards the redistribution rules, rather than in the driving mechanism as in the LH model. In this way, the trigger of an avalanche is deterministic, whereas the unfolding of the avalanche picks up a stochastic component. We will denote these models as "D" in what follows.

The lattice is driven by simultaneously multiplying each node by a value slightly above unity:

$$A_{i,j}^{n+1} = A_{i,j}^n \times (1 + \epsilon), \quad \epsilon \ll 1, \quad \forall(i,j), \quad (3.2.7)$$

where  $\epsilon$  is the driving rate. The stability criterion remains the same as in the LH model (equation (3.2.1)). A physical analogy between the deterministically-driven model and twisted coronal loops nicely applies [14], where the granular flow coherently displaces the footpoints. The avalanche criterion is directly mapped to a critical electrical current threshold.

The stochastic component in the deterministic models resides in the redistribution rules. Once the curvature of a node exceeds a threshold  $Z_c$ , redistribution occurs evenly among all neighbours as in the LH models, but with random loss of nodal value induced per redistribution:

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5}Z, \quad (3.2.8)$$

$$A_{i\pm 1,j\pm 1}^{n+1} = A_{i\pm 1,j\pm 1}^n + \frac{r_0}{5}Z, \quad (3.2.9)$$

where  $r_0$  is extracted from a uniform distribution between  $D_{nc}$  and 1. The predetermined parameter  $D_{nc} \in ]0,1[$  represents the conservation parameter. This redistribution rule is non-conservative, meaning each redistribution involves a random loss of nodal value in the lattice, which brings a stochastic aspect to the evolution of each avalanche.

In this article, we consider two types of deterministically-driven models with different levels of stochasticity (both corresponding to model D2 and D3 in Strugarek & Charbonneau 2014 [15]). The first one, which we call here Model D09, uses a conservation parameter  $D_{nc} = 0.9$  and is considered a fairly conservative model. The second, called Model D01, is a less conservative model with  $D_{nc} = 0.1$ , and therefore incorporates a larger degree of stochasticity in the evolution of its avalanches.

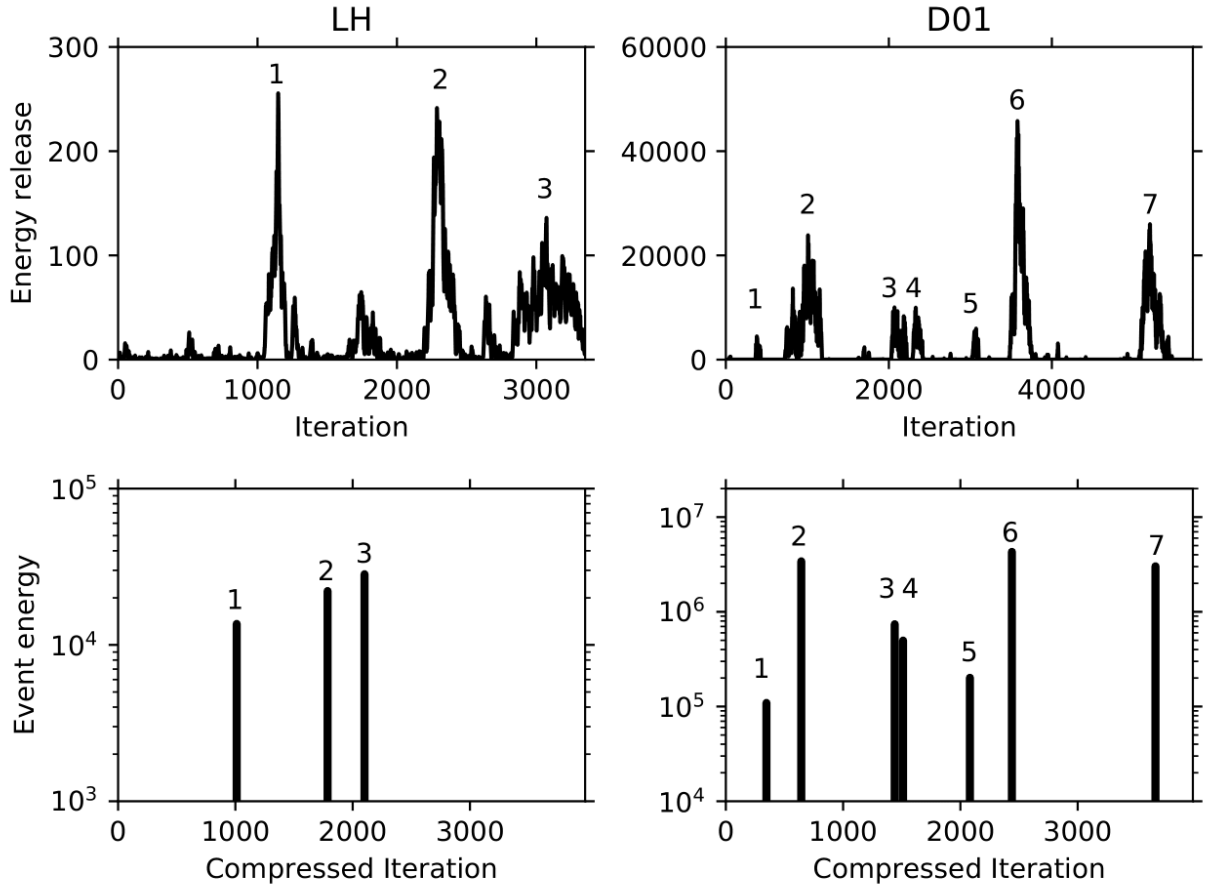
Another class of deterministically-driven models for solar flares has been recently proposed by Farhang et al. 2018 [16], where the liberated energy of each redistribution is maximized to mimic energy minimization during magnetic reconnection. Here, we only explore the initial model of Strugarek & Charbonneau 2014 [15] and defer the detailed analysis of this new class of deterministic models for future works.

### 3.2.3. Avalanching properties of LH and D models

We illustrate typical energy release time series for models LH and D01 in the top panels of Figure 3.1. Model D01 clearly tends to produce much larger avalanches than model LH. To explicitly mimic the large separation of timescales between the release of energy in solar flares and their waiting times, we compress avalanches into discrete events (shown in the bottom panel of figure 3.1), each representing the sum of all energy released during an avalanche [13]. Strugarek & Charbonneau 2014 [15] have shown that D models hold better predictive capabilities for large avalanches. Therefore, we remove all small events in the compression process, applying an event energy threshold of  $10^5$  for models D01 and D09, and  $10^4$  for the LH model. This focus on large events is also motivated by the difficulty and importance of predicting large solar flares.

Figure 3.2 shows histograms of waiting time between events of all sizes (blue, top) and between larger events (red, bottom) for models LH and D01. In both models, a clear exponential distribution is seen for waiting times of all events, reminiscent of a random Poisson process [42]. This result is expected for model LH, which triggers avalanches using a random driving process, although more surprising for model D01 with a deterministic triggering of avalanches. For larger events, model D01 departs from this exponential distribution, reminiscent of a load-unload cycle [10]. This is not the case with model LH, which maintains its exponential distribution even for higher events.

We recognize that our exponential distributions do not match the power-law waiting times distributions observed from solar flares (see Wheatland 2000 [42]). Other published avalanche models have addressed this problem by applying a non-stationary driving mechanism [43], or more recently, an energy minimization method during the redistribution of unstable nodes (Farhang et al. 2018 [16]). Since this work concentrates on the predictive capabilities of avalanche models and our data assimilation algorithm, we chose to address this problem in future work.

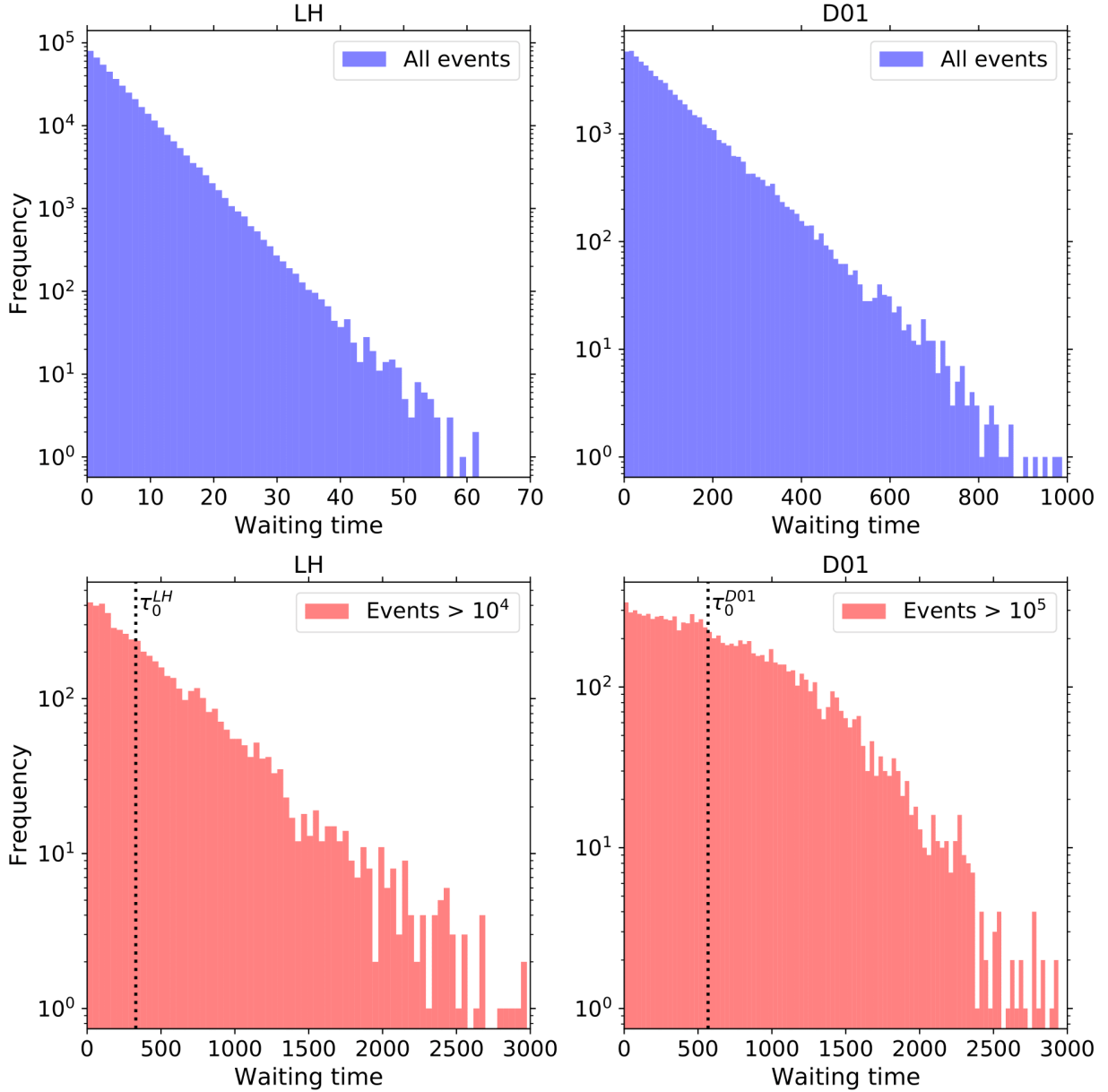


**Figure 3.1.** Top: Energy release time series for model LH (left) and D01 (right). Bottom: Compression of the same energy release time series, where each line represents the total energy released during a large avalanche (over  $10^5$  for model D01 and  $10^4$  for the LH model). Individual avalanches are labelled in each panel (1 to 3 for left panels, 1 to 7 for right panels).

### 3.3. Predictive capabilities of avalanche models

#### 3.3.1. Robustness of large avalanches

Strugarek & Charbonneau (2014) showed with a limited exploration that deterministically-driven models have a strong potential to robustly predict large events compared to LH models [15]. We further expand their analysis to systematically characterize this ability. From one fixed initial condition preceding a large avalanche, we used 1000 different random-number sequences to drive the model until an avalanche is triggered. The distribution of energy released in avalanches in all of the 1000 runs is presented in the top of Figure 3.3. In this example, models D01 and D09 both have a very high tendency to produce large avalanches, regardless of their random-number sequences. This suggests



**Figure 3.2.** Top: Waiting time distributions between events of all sizes. Bottom: Waiting time distribution between events over  $10^4$  for Model LH and over  $10^5$  for Model D01. The dotted lines shows the median waiting time for each model, denoted as  $\tau_0$ .

that a pre-existing stress pattern in the lattice is an important factor in producing a large avalanche. In the case of the LH model, the distribution is much wider, indicating a larger sensitivity to the random-number sequence in triggering a large avalanche.

We then investigated how these distributions of events changes with respect to different initial conditions preceding avalanches of different sizes. To do so, we produced 1000 distributions of energy released from initial conditions preceding large avalanches, such as

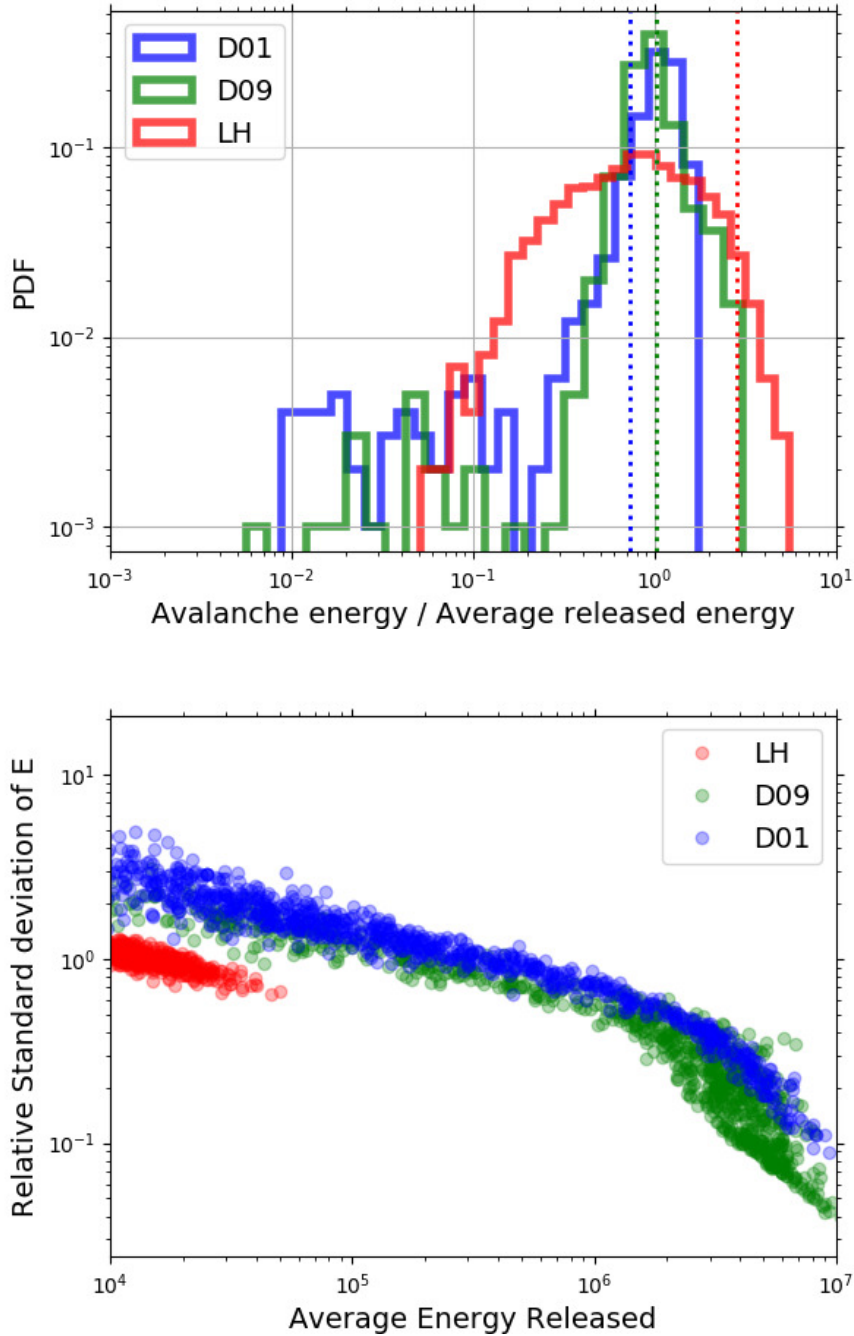


the ones presented in top panel of Figure 3.3. With each of these distributions, we extract the average and relative standard deviation of the energy release. The bottom panel of Figure 3.3 shows the relationship between these two extracted parameters. In general, the distribution of events tends to be more concentrated as the average energy released increases for all models. For events lower than  $10^5$ , the LH model tends to have more concentrated distributions than the D models. The cutoff of the LH model to events higher than  $10^5$  is a characteristic of the model itself, which tends to produce smaller events. For the D models, the standard deviation of the distributions drops for events higher than  $10^6$ , especially in the case of model D09. This confirms previous findings that D models tends to favor large avalanches in its predictive capabilities.

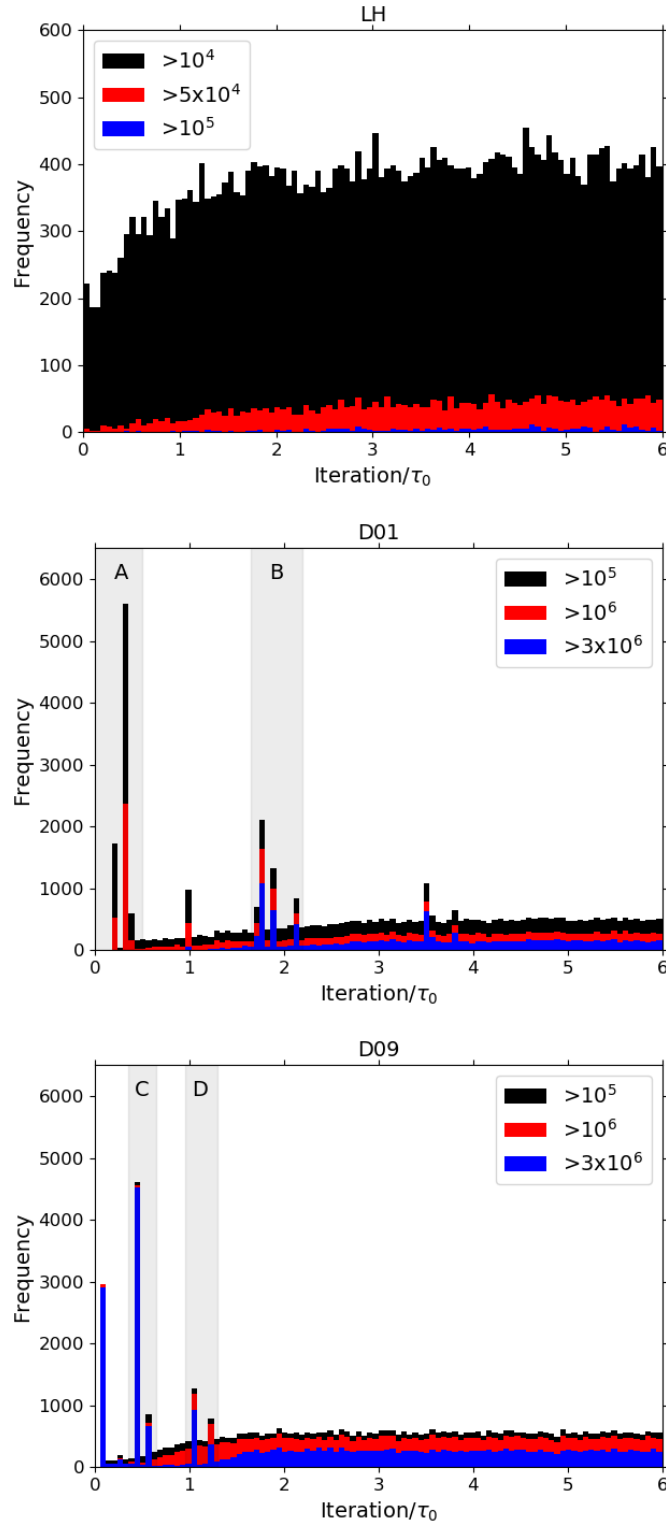
### 3.3.2. Short-term correlations between large events

The results shown in the previous section already suggests that the timing of a large avalanche is dictated by the stress patterns established across the lattice by previous avalanches. To assess whether data assimilation can be successfully coupled to avalanche models, we verified that these models also contain short-term correlations between large events. We also characterized the memory timescale associated with stress patterns. To do so, we extracted the timing of large events within 10 000 model runs, each with a common initial condition but with varying random-number sequences. The resulting histograms of these timing for each model are shown in Figure 3.4.

The time axis is normalized by the median waiting time ( $\tau_0$ ) between events larger than  $10^5$  for D models and  $10^4$  for the LH model (these characteristic times are also indicated in Figure 3.2). In the case of the LH model (top pannel of Figure 3.4), we see no preferential timing of large events. These results fit with previous findings of the low predictive capabilities of the LH model [15]. For the deterministically driven model D01 (center pannel), the results show clear peaks in the timing of large events. These peaks seem to cluster on time windows smaller than  $\tau_0/2$  and persist until up to  $4\tau_0$ . It suggests that the lattice contains stress patterns that tend to produce large avalanches at specific times. The presence of the first peak is not surprising, since the timing of the first avalanche, regardless of its size, is completely determined by the initial conditions. Further analysis shows that 58% of D01 simulations that produced large events ( $> 10^5$ ) in the B window (from  $1.75\tau_0$  to  $2.15\tau_0$ ) also



**Figure 3.3.** Top: Probability density function (PDF) of energy released from 1000 avalanches produced from the same initial condition, but with varying random-number sequences. The initial condition is chosen to precede a large avalanche. Bottom: Relative standard deviations of 1000 distributions of energy released ( $E$ ) such as the one presented in the top panel of this figure, with respect to the average energy released of each distribution.



**Figure 3.4.** Distribution of the timing of large events (over  $10^4$  for the LH model and over  $10^5$  for both D models) from 10 000 models runs, each with identical initial conditions but with varying random-number sequences. The grey highlighted areas, labeled as A, B, C and D are time windows used to analyse correlations between the timing of large events.

produced large events in the A window, the initial phases of the simulation. Model D01 therefore shows short-term correlations between events, meaning that large events produced in one time window will coherently affect the timing of future events in the short-term. In the case of model D09, the peaks are composed mainly of very large events ( $> 3 \times 10^6$ , represented in blue). They persist on a shorter time window of around  $1.2\tau_0$ . This model also shows a lower correlation between groups of large events. For example, only 33% of runs that produced large events in the D time window also produced large events in the C window. This shows a lower and less consistent correlation between the timing of large events, in comparison with model D01.

In summary, we have shown significant differences between the short term correlations in our three avalanche models, despite their similar long-term statistics. We have shown that short-term correlations exist between large avalanches for Model D01. Conversely, the D09 and LH models show little to no short-term correlations between large-events. The timing distributions shown in Figure 3.4 are representative of tens of such analyses from different initial conditions, therefore the analysis presented in the previous paragraphs reasonably represents the general features of each model. Since the LH model does not show promising predictive capabilities, we chose to omit this model in its coupling with data assimilation in what follows.

### 3.4. Data assimilation

Data assimilation consists of using a set of external data to alter a set of control parameters of a model until it reproduces the input data satisfactorily. It can be used to derive the most realistic parameters of a reduced model, to test if a given model can accommodate a given set of data, or to develop a physical model compatible with a given series of observations that can then be used to carry prediction [44]. We are interested in this paper in the latter application of data assimilation for large and rare solar flares.

The use of data assimilation to fit a model to observational data has been extensively used in many fields, notably in meteorological forecasting [18]. The general idea is to fit a simulated time series to a time window of observational data by adjusting the model's initial conditions. To achieve this, gradient descent methods, such as 4D-var, are typically used to minimize a cost function  $\mathcal{J}$  that measures the distance between the observed and modelled

time series. In the case of avalanche models for solar flares, the stochastic and discrete nature of the model output makes data assimilation a distinct challenge. We present in the following section an overview of our data assimilation procedure.

### 3.4.1. Data assimilation Procedure

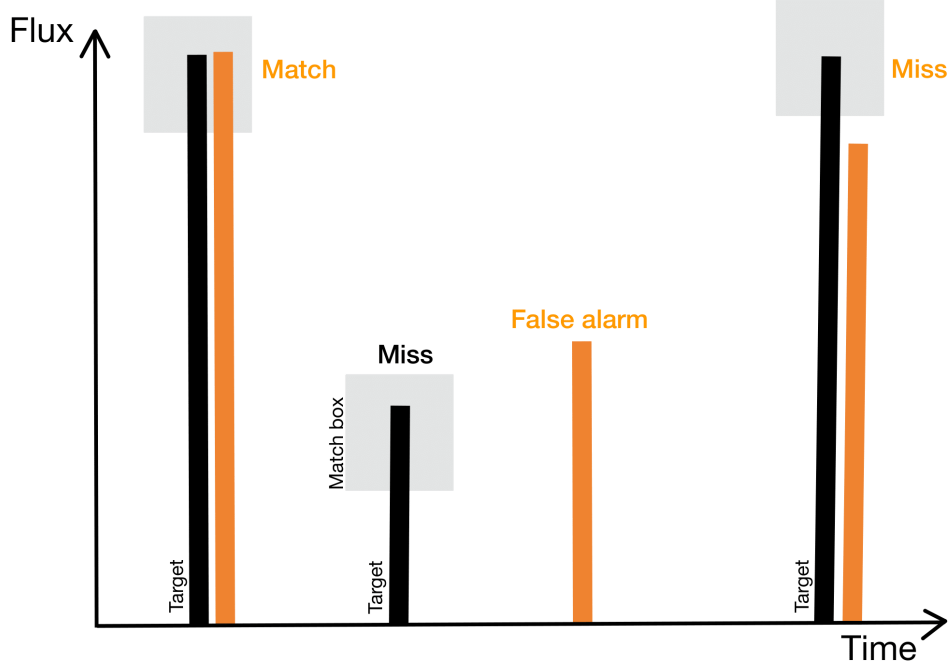
#### 3.4.1.1. Data to assimilate

The data to be assimilated consists of a series of discrete events characterized by an amplitude and an occurrence time (see e.g. bottom panels in Figure 3.1). The first type of data we will assimilate is *synthetic data*, produced independently by the sandpile model. Synthetic data can be used to validate the data assimilation procedure with data that we know can be produced by the model itself. More specifically, we use synthetic data to determine optimal numerical values for the various parameters defining our cost function and controlling the minimization algorithm.

#### 3.4.1.2. Comparing two sets of discrete events: design of an efficient cost function

The fundamental quantity we want to assimilate is thus a series of discrete events, which call here *reference data*. The output of the model is also a series of discrete events, which we refer to as *model data*. Simple least-squares minimization is inefficient due to the discrete nature of the data we are dealing with. We thus developed a dedicated cost function  $\mathcal{J}$  designed to compare the distance between two series of discrete events [21].

Our primary goal is to accurately reproduce the largest events in a given time series. We also do not want to miss any large events, as these set the long-range spatial correlations across the lattice that confer predictability to the model, despite its stochastic elements. Finally, false alarms also need to be avoided during the assimilation process, as they will likely decrease the predictive potential of the assimilated model. For each event in the reference data and in the model data, we check whether it is a *match*, a *miss* or a *false alarm*. Two events are considered to match when the model event occurs within a small time-window (here 400 model iterations) around the reference event, and with an energy in  $[E_0/2, 2E_0]$  (where  $E_0$  is the energy of the reference event). We illustrate the distinction between matched, missed and false alarm events in Figure 3.5. Note that our match criterion



**Figure 3.5.** Illustration of our definition of a match, miss and false alarm used in our cost function between the reference (shown in black) and model data (shown in orange). The grey squares represent areas in which the model data is considered a match.

is quite demanding, any event outside the gray boxes is considered a miss, including the rightmost pair of observation and assimilated events despite their similar timing.

The cost function  $\mathcal{J}$  is then defined as

$$\mathcal{J} = 1 - \left( \alpha \sum_{\text{match}} \frac{E_0}{E_{\text{tot}}} - \beta \sum_{\text{miss}} \frac{E_0}{E_{\text{tot}}} - \gamma \sum_{\text{False Alarm}} \frac{E_i}{E_{\text{tot}}} \right), \quad (3.4.1)$$

where  $E_{\text{tot}} = \sum E_0$  is the total energy in the reference data,  $E_i$  is the energy of events in the model data, and  $\alpha$ ,  $\beta$  and  $\gamma$  are adjustable coefficients that we set here to 1, 0.5 and 0.25. This cost function gives more weight to large events that we want to predict most reliably. It also gives more weight to matching events ( $\alpha = 1$ ), than missed events ( $\beta = 0.5$ ), and finally less weight to false alarms ( $\gamma = 0.25$ ). If each of the reference data is matched with a model data within the match box,  $\mathcal{J} = 0$ . Conversely, if none of the reference events are matched by the model events,  $\mathcal{J} \sim 1.75$  (no match, all data events are considered as false alarms and all reference events as misses).

The cost function (3.4.1) is only one example of how to design a cost function to compare two discrete series of events. We tested various flavours of  $\mathcal{J}$ , *e.g.* by varying the coefficients

$\alpha$ ,  $\beta$  and  $\gamma$  and by using the logarithm of the fractional energy of events  $E_0/E_{\text{tot}}$  and did not find any significant change in the results presented here.

#### 3.4.1.3. *Minimization of the cost function*

In order to minimize the cost function (3.4.1), several optimization techniques can be used. Before selecting a minimization algorithm, one must beforehand define the set of control parameters that will be allowed to evolve during the process. In this work, we do not want to alter the ingredients of the sandpile model to maintain the toy model’s solar-like statistics. The only parameter left is then the initial condition of the sandpile, which is typically of size  $N = N_x \times N_y \sim 10^3$ . Such a large set of control parameters slows down the assimilation process. In this work, we make use of the eigenvalue decomposition of the sandpile model itself to decrease the dimension of the set of control parameters to 50. The details of this technique can be found in Hung et al. 2017 [44] in the context of the prediction of the solar cycle, and in Strugarek et al. 2017 [30] in the context of flare forecasting with sandpile models.

The minimization of the cost function can be realized through standard down-gradient algorithms; advanced 4DVar method using an adjoint code to estimate the gradient of the cost function [21] (and hence significantly reducing the convergence time of the minimization); or more generic and costly algorithm such as simulated annealing to avoid local minima traps by allowing up-gradient exploration. We tested each of these methods. The adjoint of the sandpile model was automatically generated using the Tapenade software [45]. We used the simulated annealing algorithm implemented in Numerical Recipes [22]. We found by employing these different approaches that only the simulated annealing algorithm was able to robustly succeed. Indeed, the cost function (3.4.1) presents numerous local minima, and most down-gradient minimization methods are not able to find a sufficiently good minimum. An alternative approach could make use of neural networks to minimize  $\mathcal{J}$ , which we leave for future work.

The simulated annealing method allows occasional up-gradient explorations with a probability that depends on a parameter  $T$  that represents the temperature in the analogy of annealing in metallurgy. In this work,  $T$  is initialized at 10, then decreases linearly for 20 iterations until it reaches  $10^{-5}$ . As proposed in Numerical Recipes [22], each iteration of the

simulated annealing uses a downhill simplex method (also known as the amoeba method). Here we use 20 iterations of amoeba descent for each simulated annealing iteration. The choice of the cooling schedule of  $T$  and the number of iterations for both algorithms have been found empirically to achieve satisfying results while maintaining its relatively high calculation speed.

### 3.4.2. Results

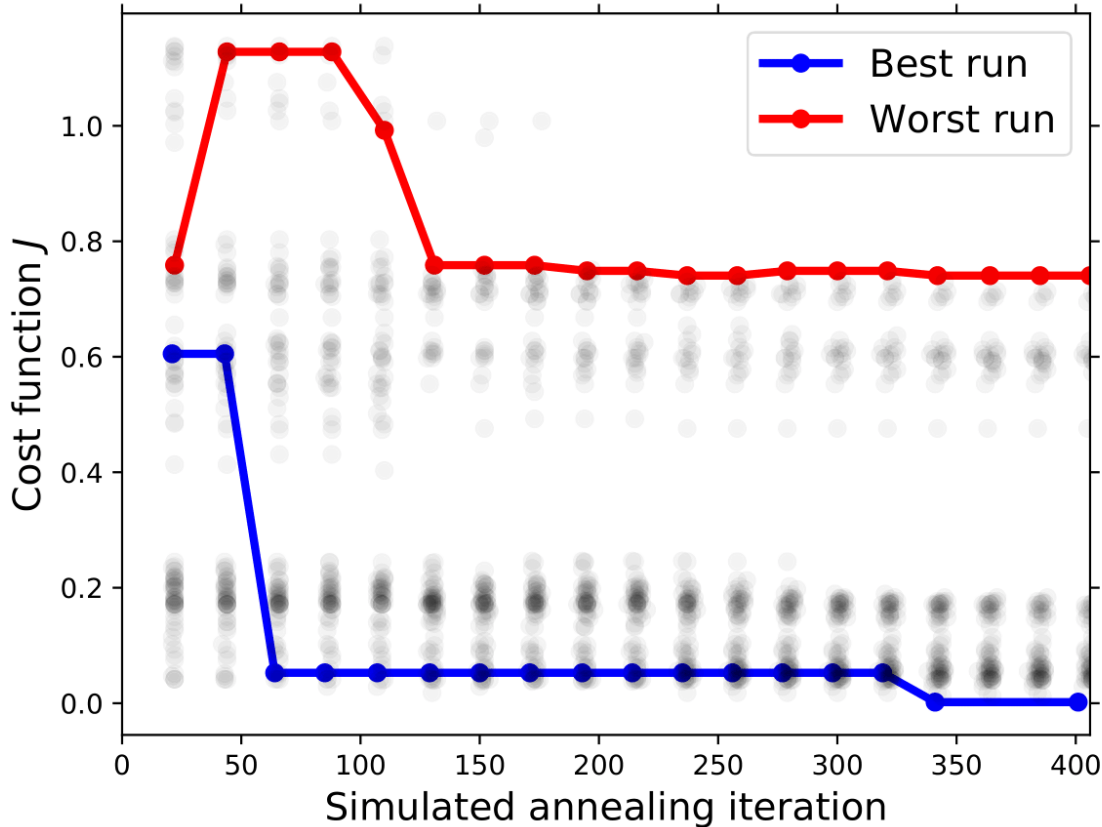
Figure 3.6 highlights two examples of the evolution of the cost function  $\mathcal{J}$  for Model D01 undergoing our minimization algorithm presented above. In the initial phase of the minimization, the temperature  $T$  is high enough for up-gradient exploration to occur fairly frequently, as shown in the initial phase of the worst run (in red). As the temperature diminishes with the iterations, the cost function stabilizes to a local minimum. The end result of these 100 minimization results contains two groups, one with higher cost functions of  $\mathcal{J} \sim 0.5 - 0.7$  and the other with more successful runs of  $\mathcal{J} \sim 0 - 0.2$ . These groups of results are quite frequent in our minimization outputs. This is due to the discrete nature of the penalties produced by misses and false alarms.

Figure 3.7 shows the event time series of the best (Top) and worst (Bottom) results of the same 100 data assimilation runs on a common synthetic observation from Figure 3.6.

In the first example, the assimilated run (blue peaks) succeeds in capturing all five synthetic reference events (black peaks). This leads to a perfect cost function of  $J = 0$ . In the second example (red peaks), only two reference events were captured, and three were missed. A particular big increase in our cost function of  $J = 0.7565$  is mostly due to the three missed events (recall that our cost function is designed to have a bigger penalty on larger missed events). The four false alarms also contribute, to a lesser extent, to this higher cost function (misses have twice the weight as false alarms). These two examples show variations of assimilation results on a single synthetic observation, but the results also fairly depends on the synthetic data being assimilated.

The top panel of Figure 3.8 shows a more general trend of our data assimilation results, produced by a set of 100 synthetic data, each being assimilated 100 times, for a total of 10 000 cost functions for each model D01 and D09 (labelled as 'assimilated'). In addition to the assimilated results, we included the cost function distributions of 10 000 non-assimilated

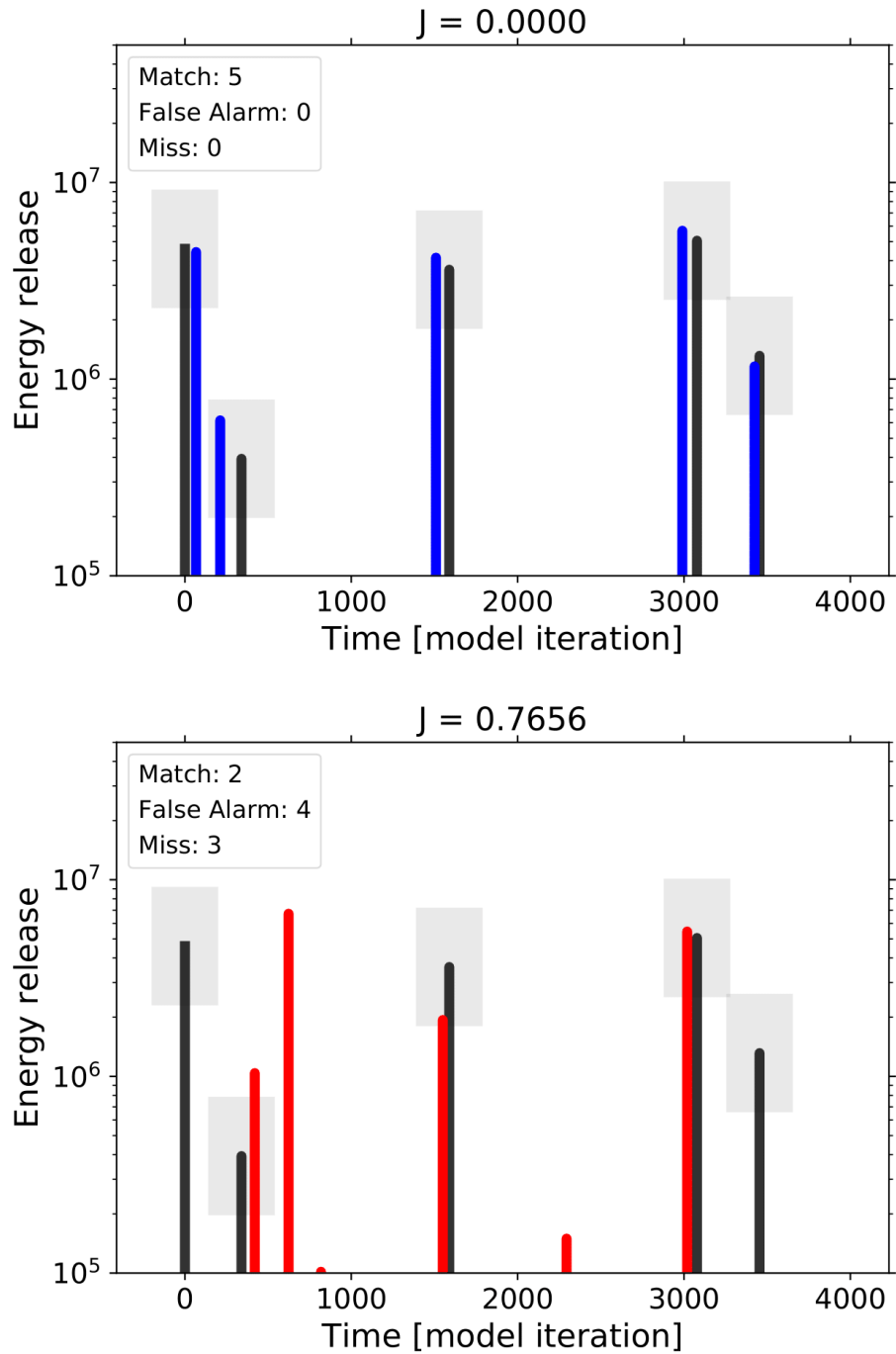




**Figure 3.6.** Evolution of the cost function during 100 simulated annealing algorithm runs on Model D01 on a single synthetic observation. The grey dots represent the results of all 100 simulated annealing runs, and the blue and red lines highlight the best and worst run respectively.

initial conditions, again with 100 runs for each of the 100 synthetic data (labelled as 'random'). Each of these non-assimilated initial conditions was selected at random iterations of a simulation run of  $\sim 10^8$  iterations.

Our assimilated runs clearly show significantly better results than the randomly generated initial conditions. They both exhibit similar results, a rising phase for low cost function, a peak at  $J \sim 0.5$  with finally a diminishing tail until  $J \sim 1.1 - 1.2$ . The low consistency of our assimilation results is mainly due to the stochastic nature of our simulated annealing method and the variety of synthetic observations to fit. That being said, the low computational cost of assimilating these sandpile models leads no problem to repeat the assimilation process many times, only to keep the best runs for one reference data.



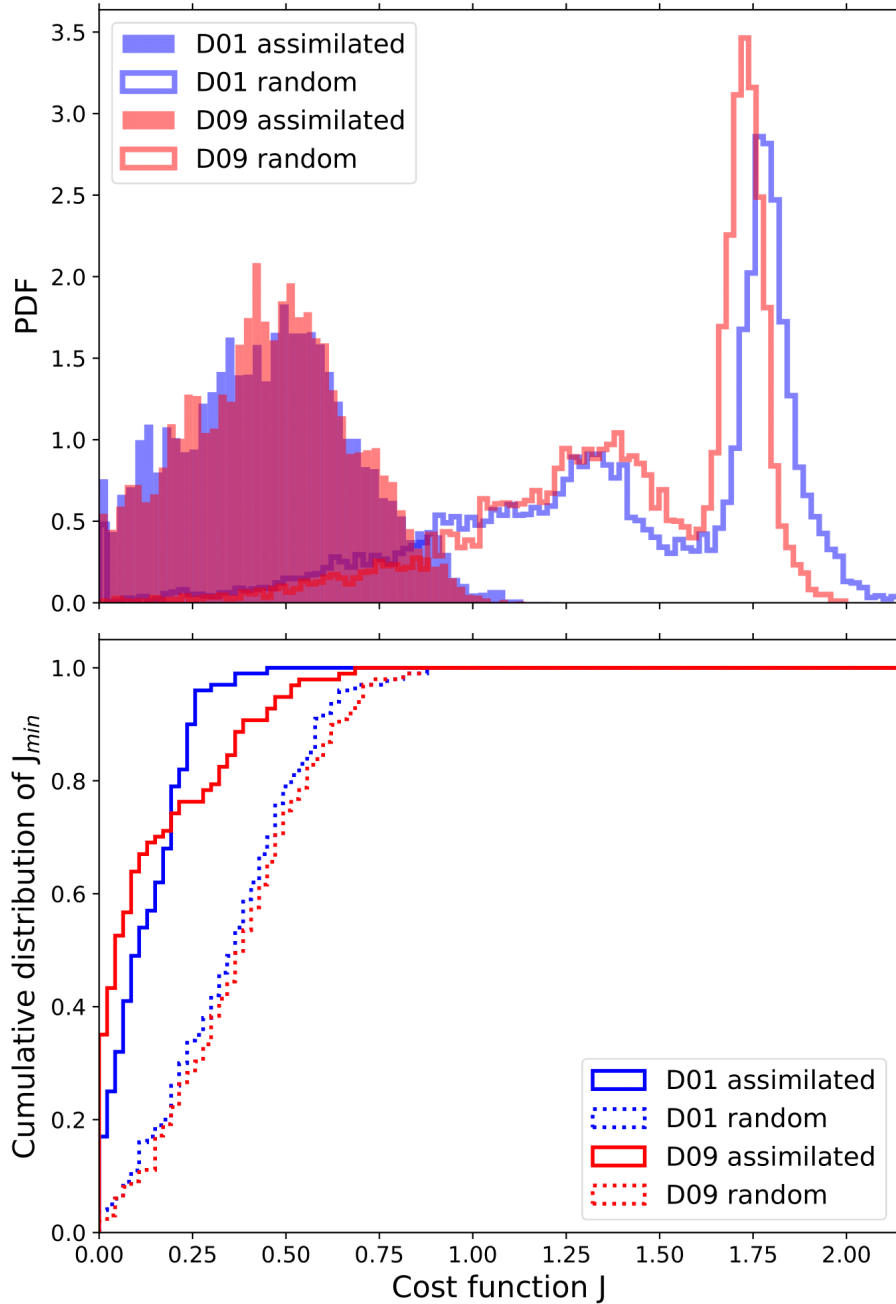
**Figure 3.7.** Event energy time series of the best (top, blue) and worst (bottom, red) output of 100 data assimilation runs using Model D01 on a common synthetic data. These two examples corresponds to the output of the same color-highlighted examples shown in Figure 3.6.

The bottom panel of Figure 3.8 shows the cumulative distribution of the best cost function achieved for each observation, denoted as  $J_{min}$ . Both these models clearly outperform predictions from a random initial condition (shown as dotted lines). Only around 17% of observations got a  $J_{min}$  of around 0 for model D01, and 36% for model D09. A cost function of 0 is difficult to achieve for most observations with our current data assimilation algorithm, since some have more events to assimilate. It is probable that more allowed iterations during our simulated annealing method would help get better results. Model D09 outperforms Model D01 until around  $J_{min} = 0.21$ . On the other hand, model D01 seems to be more consistent with its best results, with all of them being at least under  $J_{min} = 0.5$ . In summary, Model D09 seems better suited to achieve at least some great results ( $J_{min} < 0.2$ ), but Model D01 produces more consistent results.

### 3.5. Discussion and Conclusions

In this work, we explicitly tested the predictive capabilities of three avalanche models for solar flares (the Lu & Hamilton model (LH) and two deterministically-driven models (D01 and D09)). Although these models generally show very similar statistical properties (i.e. power-law size distribution), they differ in their short-term behaviour. We have shown that pre-existing stress patterns embedded in a given initial condition of the lattice are a strong factor in producing a large avalanche in the short-term, despite the important stochastic elements of these models (see figure 3.4). This correlation between a lattice configuration and the timing of its future large events is much more present for model D01 and D09, and quite absent for model LH. Model D01 also shows good temporal correlations between large avalanches themselves. Therefore, at least some information for predicting future avalanches is contained in past events, which is a well-needed ingredient for forecasting with data assimilation.

We presented our data assimilation protocols, which consists of using a simulated annealing method to fit our model output to a given set of observations, giving an optimal initial condition for forecasting. We use an eigenfunction decomposition of the lattice to reduce the control parameters to 50, rather than adjusting each of the  $48 \times 48 = 2304$  nodes individually. This greatly speeds up the fitting process. For example, for model D01, the algorithm takes



**Figure 3.8.** Top: Distribution of results of 10 000 data assimilation runs, consisting of 100 runs on each of 100 synthetic observations. The empty bars, labelled as 'random', represents 100 runs from random initial conditions on each of the same 100 synthetic observations, for a total of 10 000 results. Bottom: cumulative distribution of the best results ( $J_{min}$ ) on each of the 100 synthetic observations.

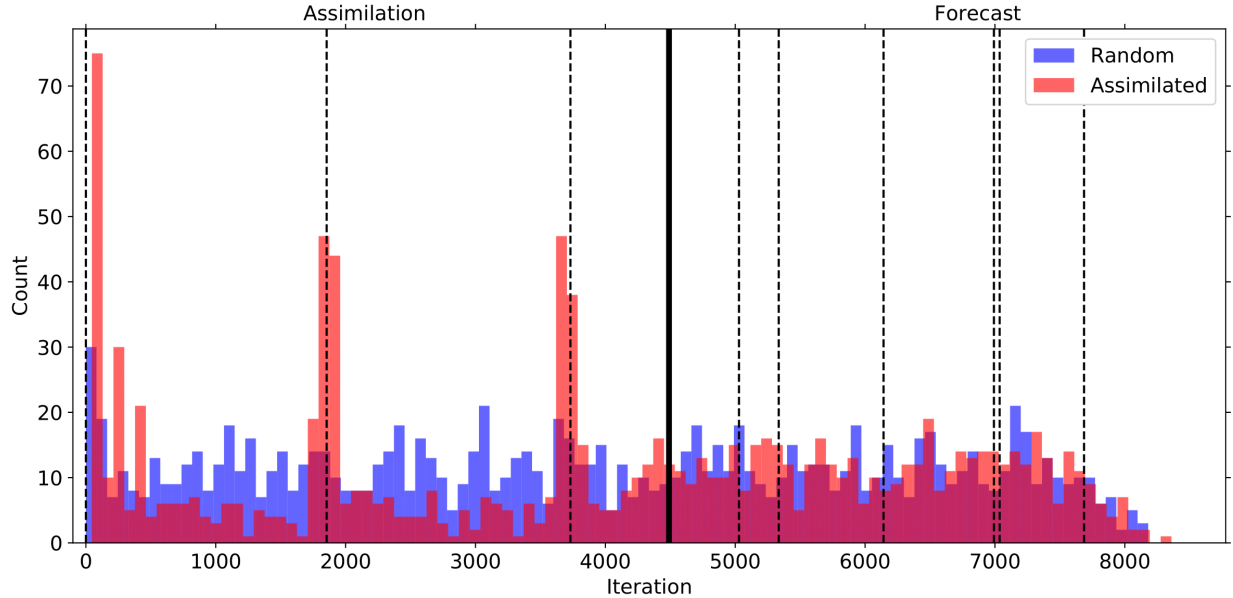
around 10 seconds to fit an observation set of 5 events. The observations used to test our data assimilation algorithm are synthetic observations created from the models themselves.

We show that our data assimilation protocols, coupled with model D01 and D09, successfully reproduces our set of 100 synthetic observations for each model (see figure 3.8). We can therefore conclude that our algorithm can fit reasonably well any set of synthetic observations produced by the avalanche models themselves, and that our fitting procedure is rapid enough to quickly repeat the process many times for a statistically relevant forecast.

The fact that we can successfully assimilate synthetic observation is a good first step. The next logical step is to assure that many assimilation runs on a single observation produces similar forecasts. The wanted behaviour would be that each initial condition that reproduces well our assimilation time window would all produce similar forecasts.

Unfortunately, we have not yet succeeded in producing these coherent forecasts. To show this explicitly, Figure 3.9 presents the timing of large avalanches ( $> 10^5$ ) produced by 100 assimilated initial conditions, all fitting the same set of synthetic observations. The events shown occur in both the assimilation time-window (on the left, where lies the observations used for data assimilation) and the forecast time-window (on the right). The dotted black lines are the timing of large events from the synthetic observations. For the assimilated initial conditions (shown in red), we see clear peaks of timing around the observations in the assimilation window. This is simply an indication that our data assimilation algorithm successfully fitted the observations. Once the assimilation time-window ends, the distribution of assimilated runs falls into a noise-like behaviour, just as the non-assimilated distribution (shown in blue). This suggests that our set of assimilated initial conditions does not have sufficient aspects in common to produce coherent forecasts. The number of states with acceptable results for assimilation seems to far outnumber those who also produce coherent forecasts. Further work is thus needed to fix this problem, in the details of our data assimilation procedure and/or in the models used themselves.

A promising alternative to the LH and D models used in this article is an avalanche model that uses the principle of minimum energy in magnetic reconnection, first proposed by Farhang et al. 2018 [16]. One of the main features of these models is that the unstable nodes will redistribute their excess nodal value unevenly to each neighbour, and this partition is chosen to minimize as much as possible the lattice energy post-redistribution. It seems reasonable to believe that since this minimization per redistribution is dependent on the landscape of the lattice at a given iteration, it will likely decrease the stochasticity of the



**Figure 3.9.** Distribution of the timing of large avalanches ( $< 10^5$ ) produced by 100 assimilated (in red) and non-assimilated (in blue) initial conditions from model D01. The black dotted lines show the timing of the synthetic observations, and the thicker black line shows the separation between the assimilation (left) and forecast (right) time-windows.

unfolding of an avalanche. Inspired by this novel approach to redistributions, we have recently developed another class of deterministically-driven models that also uses this minimization scheme. For each redistribution event, we use a Monte-Carlo method to find the partition that maximizes the energy release. We therefore named these models "deterministically-driven avalanche models with Monte-Carlo minimization", or DMC for short. Preliminary results of these models already show promising predictive capabilities, which could potentially help in our forecasting goal.

**Acknowledgements** We thank A.S. Brun, L. Jouve, B. Tremblay and A. Vincent for multiple discussions on data assimilation procedures. We acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC). A. Strugarek thanks N. Vilmer for discussions on solar flares and self-organized criticality models. A. Strugarek acknowledges funding from ERC WHOLESUN 810218 grant, INSU/PNST, and CNES Solar Orbiter.

# Chapitre 4

---

## Conclusion

Dans ce mémoire, nous avons utilisé les modèles d’avalanches comme une version hautement simplifiée des processus de déclenchement des éruptions solaires. Ils reproduisent assez bien les statistiques en loi de puissance de la taille des éruptions solaires, et leurs règles d’évolution contiennent quelques analogues physiques hautement simplifiés des réactions en chaîne de reconnexion magnétique dans les régions actives.

Il a déjà été montré que certains de ces modèles, notamment les modèles à forçage déterministe, démontrent des capacités prédictives prometteuses malgré leurs aspects stochastiques assez importants [15]. Cette capacité prédictive est mesurée par la tendance au modèle à produire un résultat donné, peu importe la séquence de nombres aléatoires. L’insensibilité de ces modèles aux séquences de nombres aléatoires est une caractéristique importante pour ultimement produire des prédictions cohérentes. L’article inclus dans ce mémoire présente des analyses détaillées sur la capacité prédictive des modèles d’avalanche, concluant que le modèle D01 contient des fenêtres de prédiction de plus longue durée que les modèles D09 et LH ainsi qu’une meilleure corrélation entre les temps de déclenchement des avalanches de grandes tailles. Au chapitre 2 de ce mémoire, nous avons également introduit un nouveau modèle, appelé DMC, qui est aussi caractérisé par une bonne capacité prédictive. Bien sûr, ces aspects prometteurs peuvent sûrement être améliorés avec une optimisation des paramètres des modèles, ce qui n’a pas été fait ici. Nous avons tout de même jugé que ces modèles ont un minimum raisonnable de capacités prédictives pour le coupler à notre protocole d’assimilation de données.

Nous nous sommes ensuite penchés sur l’assimilation de données, qui est une technique de prédiction consistant à incorporer des données observées pour modifier notre modèle

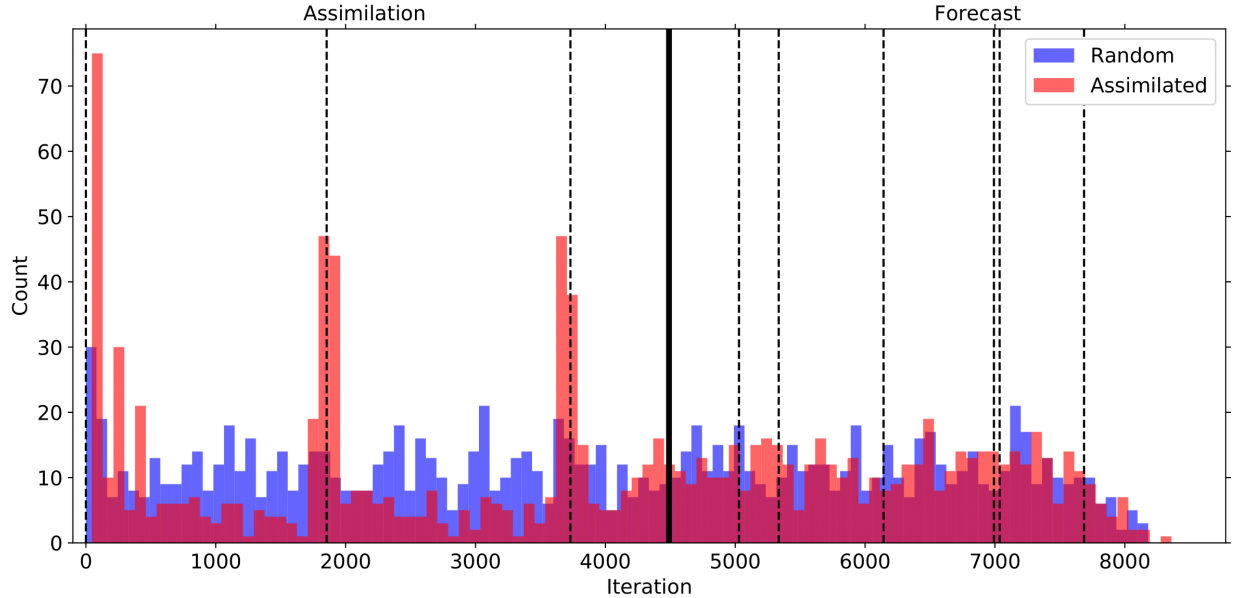
afin qu'il reproduise celles-ci le plus près que possible. Dans le cadre de ce projet, nous avons appliqué une méthode de recuit simulé pour ajuster la condition initiale de notre réseau de sorte à reproduire le mieux possible une série d'évènements observés. Une fois que les observations sont reproduites, on poursuit la simulation plus loin pour produire les prédictions. Les observations que nous avons utilisées ici sont des observations 'synthétiques', produites par les modèles d'avalanches eux-mêmes. De cette façon, nous avons pu ajuster notre procédure d'assimilation de données sur des observations qui sont cohérentes a priori avec le modèle utilisé.

Pour atteindre un bon protocole de prédiction, nous devons satisfaire aux trois conditions suivantes : 1) Notre procédure d'assimilation de données doit produire des conditions initiales qui reproduisent assez bien une série observations (réelle ou synthétique). 2) Un grand nombre de conditions initiales assimilées (qui reproduisent assez bien les observations) doivent produire des prédictions assez semblables. 3) Ces prédictions semblables doivent correspondre le mieux possible aux observations à prédire. Nous allons maintenant discuter de l'état actuel de notre protocole de prédiction par rapport à chacune de ces conditions.

La première condition est assez bien satisfaite, particulièrement dans le cas des observations synthétiques. Nous avons montré dans l'article inclus dans ce mémoire que notre algorithme d'assimilation de données reproduit très bien une centaine d'observations synthétiques avec des fonctions coût raisonnablement faibles. De plus, le temps de calcul est rapide, environ 10 secondes pour chaque minimisation dans le cas du modèle D01 pour une fenêtre temporelle contenant environ 5 évènements. Une telle rapidité dans l'assimilation est essentielle à tout système opérationnel visant la prédiction des éruptions sur une fenêtre de l'ordre de 6 heures. On peut donc recommencer l'assimilation de données une centaine de fois pour finalement choisir les meilleurs résultats, également présenté dans l'article de ce mémoire. Les séries temporelles du flux en rayon X des satellites GOES ont également été assimilées avec le même genre de succès par notre algorithme d'assimilation de données. Un exemple de ce résultat est présenté à la conclusion de l'article de Strugarek et al. 2014 [15].

La deuxième condition, celle des prédictions cohérentes pour un grand nombre d'assimilations sur une observation donnée, n'est pas encore satisfaite. Même si nous produisons 100 conditions initiales qui reproduisent chacun très bien une série d'observations, leurs prédictions subséquentes ne sont pas semblables les unes avec les autres. La figure 4.1 présente





**FIGURE 4.1.** Distributions du temps de déclenchement des avalanches produit par 100 conditions initiales assimilées (rouge) et non assimilées (bleu) avec le modèle D01. Les lignes noires pointillées représentent le temps de déclenchement des observations synthétiques, et la ligne noire au centre sépare la fenêtre d’assimilation (à gauche) et la fenêtre de prédiction (à droite).

justement ce résultat problématique par une distribution du temps de déclenchement des avalanches de grandes tailles produites par 100 conditions initiales assimilées. On voit bien que dans la fenêtre d’assimilation, qui contient les observations assimilées (représentées avec des lignes noires pointillées), les conditions initiales assimilées (en rouge) produisent des pics autour des observations. Ceci n’est qu’un rappel du succès de l’assimilation des observations. En revanche, dans la fenêtre de prédiction à droite, on retombe dans le bruit de façon comparable aux conditions initiales non assimilées (en bleu). Le comportement voulu serait la présence de pics bien définis également dans la fenêtre de prédiction, comme ceux présentés à la figure 3.4 dans l’article de ce mémoire. Cette analyse a été refaite pour une centaine d’observations synthétiques et avec le modèle D09, toujours avec des résultats semblables. Dans l’article inclus dans ce mémoire, on montre qu’une condition initiale donnée tend à produire de grosses avalanches dans des temps similaires pour les modèles D, indépendamment de la stochasticité. Par contre, nos ensembles de 100 configurations de réseau assimilés ne semblent pas avoir assez d’aspects en commun pour produire des prédictions similaires au-delà de la fenêtre d’assimilation.

Finalement, la troisième condition, celle des prédictions qui concordent aux observations à prédire, est également non réussie. Il est fort probable que cette condition puisse seulement être atteinte que si la deuxième l'est également.

Puisque cet algorithme est nouveau, il y a plusieurs analyses intéressantes qui n'ont pas encore été faites. Elles pourraient nous indiquer quelques pistes dans la bonne direction.

Il semble évident que l'ensemble de conditions initiales qui reproduisent bien la fenêtre d'assimilation est beaucoup plus grand que celles qui produisent des bonnes prédictions. On pourrait donc assimiler plutôt les deux fenêtres temporelles en même temps. Cela revient essentiellement à doubler la longueur de la fenêtre d'assimilation. On pourrait donc avoir une intuition de la difficulté du modèle à reproduire les deux fenêtres en vérifiant entre autres la distribution des résultats, et ça nous permettrait peut-être d'avoir accès à des échantillons de conditions initiales qui produisent de bonnes prédictions. Nous pourrions donc analyser leurs aspects communs et voir comment ils diffèrent des conditions initiales qui reproduisent bien seulement la fenêtre assimilée.

Nous avons également démontré dans ce projet de maîtrise que le temps de cohérence à partir d'une condition initiale donnée est assez réduit. Le temps entre le dernier événement assimilé et le premier à prédire est potentiellement un aspect important à prendre en considération. Par exemple, si notre condition initiale assimilée reproduit bien tous les événements observés sauf un, il est sûrement préférable que cet événement manqué soit dans la première partie de la fenêtre d'assimilation, plutôt que dans la dernière. On pourrait donc ajouter une dimension temporelle à notre fonction coût, mettant plus de pénalité pour les événements manqués plus près de la fenêtre de prédiction. De cette façon, l'algorithme tendra à mieux reproduire les événements près de la fenêtre de prédiction.

Il se peut aussi que l'assimilation de seulement les événements de grandes tailles ( $> 10^5$  dans ce mémoire) soit insuffisant pour reproduire les bons patrons de stress dans le réseau. On pourrait donc modifier notre protocole pour assimiler également les avalanches de petite taille. Cela nécessiterait sûrement plus d'itération de minimisation, une plus petite fenêtre d'assimilation et/ou une fonction coût plus souple.

Finalement, les détails du modèle couplé avec notre protocole de prédiction sont potentiellement très importants. Il y a plusieurs modèles d'avalanche publiés dans le contexte des éruptions solaires, et une infinité de combinaisons de paramètres pour chacun. Ces modèles

produisent des résultats parfois difficiles à saisir intuitivement, et il est possible que notre intuition de nos analyses de la capacité prédictive des modèles nous amène parfois dans la mauvaise direction. Il serait donc intéressant d'essayer quelques autres modèles d'avalanche avec notre protocole de prédiction, dont les modèles DMC présentés au chapitre 2. On pourrait aussi imaginer une maximisation de la performance de notre protocole de prédiction en modifiant à la fois les paramètres du modèle ainsi que le type de modèle, le tout avec des techniques numériques d'intelligence artificielle ou même des algorithmes de maximisation plus traditionnels. Il faudrait donc avoir en entrée un vecteur de paramètres du modèle, et l'algorithme maximise une quantification de sa performance prédictive (par exemple, la moyenne de la fonction coût de la prédiction).



## Bibliographie

---

- [1] R. C. Carrington. Description of a Singular Appearance seen in the Sun on September 1, 1859. *Monthly Notices of the Royal Astronomical Society*, 20 :13–15, November 1859. doi : 10.1093/mnras/20.1.13. URL <http://adsabs.harvard.edu/abs/1859MNRAS..20...13C>.
- [2] S. Masuda, T. Kosugi, H. Hara, S. Tsuneta, and Y. Ogawara. A loop-top hard X-ray source in a compact solar flare as evidence for magnetic reconnection. *Nature*, 371 : 495–497, October 1994. doi : 10.1038/371495a0. URL <http://adsabs.harvard.edu/abs/1994Natur.371..495M>.
- [3] Markus Aschwanden. *Self-Organized Criticality in Astrophysics : The Statistics of Non-linear Processes in the Universe*. Astronomy and Planetary Sciences. Springer-Verlag, Berlin Heidelberg, 2011. ISBN 978-3-642-15000-5. doi : 10.1007/978-3-642-15001-2. URL <https://www.springer.com/gp/book/9783642150005>.
- [4] E. N. Parker. Sweet’s mechanism for merging magnetic fields in conducting fluids. *Journal of Geophysical Research (1896-1977)*, 62(4) :509–520, 1957. ISSN 2156-2202. doi : 10.1029/JZ062i004p00509. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ062i004p00509>.  
\_eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/JZ062i004p00509>.
- [5] H. E. Petschek. Magnetic Field Annihilation. *NASA Special Publication*, 50 :425, 1964. URL <http://adsabs.harvard.edu/abs/1964NASSP..50..425P>.
- [6] Kazunari Shibata and Tetsuya Magara. Solar Flares : Magnetohydrodynamic Processes. *Living Reviews in Solar Physics*, 8(1) :6, December 2011. ISSN 1614-4961. doi : 10.12942/lrsp-2011-6. URL <https://doi.org/10.12942/lrsp-2011-6>.
- [7] Masaaki Yamada, Russell Kulsrud, and Hantao Ji. Magnetic reconnection. *Reviews of Modern Physics*, 82(1) :603–664, March 2010. doi : 10.1103/RevModPhys.82.603.

- URL <https://link.aps.org/doi/10.1103/RevModPhys.82.603>. Publisher : American Physical Society.
- [8] E. N. Parker. Nanoflares and the solar X-ray corona. *The Astrophysical Journal*, 330 : 474–479, July 1988. doi : 10.1086/166485. URL <http://adsabs.harvard.edu/abs/1988ApJ...330..474P>.
- [9] M. J. Aschwanden. Self-Organized Criticality Systems. *Self-Organized Criticality Systems, Edited by M.J. Aschwanden. e-book published by Open Academic Press, Berlin, Warsaw, 2013, 483pp.*, July 2013. URL <http://adsabs.harvard.edu/abs/2013socs.book.....A>.
- [10] R. Rosner and G. S. Vaiana. Cosmic flare transients - Constraints upon models for energy storage and release derived from the event frequency distribution. *The Astrophysical Journal*, 222 :1104–1108, June 1978. doi : 10.1086/156227. URL <http://adsabs.harvard.edu/abs/1978ApJ...222.1104R>.
- [11] Edward T. Lu. Constraints on Energy Storage and Release Models for Astrophysical Transients and Solar Flares. *The Astrophysical Journal*, 447 :416, July 1995. doi : 10.1086/175885. URL <http://adsabs.harvard.edu/abs/1995ApJ...447..416L>.
- [12] Edward T. Lu and Russell J. Hamilton. Avalanches and the distribution of solar flares. *The Astrophysical Journal Letters*, 380 :L89–L92, October 1991. doi : 10.1086/186180. URL <http://adsabs.harvard.edu/abs/1991ApJ...380L..89L>.
- [13] Edward T. Lu, Russell J. Hamilton, J. M. McTiernan, and Kenneth R. Bromund. Solar flares and avalanches in driven dissipative systems. *The Astrophysical Journal*, 412 : 841–852, August 1993. doi : 10.1086/172966. URL <http://adsabs.harvard.edu/abs/1993ApJ...412..841L>.
- [14] Antoine Strugarek, Paul Charbonneau, Richard Joseph, and Dorian Pirot. Deterministically Driven Avalanche Models of Solar Flares. *Solar Physics*, 289(8) :2993–3015, August 2014. ISSN 1573-093X. doi : 10.1007/s11207-014-0509-7. URL <https://doi.org/10.1007/s11207-014-0509-7>.
- [15] A. Strugarek and P. Charbonneau. Predictive Capabilities of Avalanche Models for Solar Flares. *Solar Physics*, 289(11) :4137–4150, November 2014. ISSN 1573-093X. doi : 10.1007/s11207-014-0570-2. URL <https://doi.org/10.1007/s11207-014-0570-2>.

- [16] Nastaran Farhang, Hossein Safari, and Michael S. Wheatland. Principle of Minimum Energy in Magnetic Reconnection in a Self-organized Critical Model for Solar Flares. *The Astrophysical Journal*, 859(1) :41, May 2018. ISSN 0004-637X. doi : 10.3847/1538-4357/aac01b. URL <https://doi.org/10.3847%2F1538-4357%2Faac01b>. Publisher : American Astronomical Society.
- [17] Roger Daley. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, February 1994. ISBN 978-0-521-45825-2.
- [18] Eugenia Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, New York, 1 edition edition, November 2002. ISBN 978-0-521-79629-3.
- [19] Yoshikazu Sasaki. SOME BASIC FORMALISMS IN NUMERICAL VARIATIONAL ANALYSIS. *Monthly Weather Review*, 98(12) :875–883, December 1970. ISSN 0027-0644. doi : 10.1175/1520-0493(1970)098<0875:SBFINV>2.3.CO;2. URL <https://journals.ametsoc.org/mwr/article/98/12/875/99064/SOME-BASIC-FORMALISMS-IN-NUMERICAL-VARIATIONAL>. Publisher : American Meteorological Society.
- [20] John M. Lewis and John C. Derber. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A : Dynamic Meteorology and Oceanography*, 37(4) :309–322, January 1985. ISSN null. doi : 10.3402/tellusa.v37i4.11675. URL <https://doi.org/10.3402/tellusa.v37i4.11675>. Publisher : Taylor & Francis \_eprint : <https://doi.org/10.3402/tellusa.v37i4.11675>.
- [21] Eric Bélanger, Alain Vincent, and Paul Charbonneau. Predicting Solar Flares by Data Assimilation in Avalanche Models. *Solar Physics*, 245(1) :141–165, September 2007. ISSN 1573-093X. doi : 10.1007/s11207-007-9009-3. URL <https://doi.org/10.1007/s11207-007-9009-3>.
- [22] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in FORTRAN 77 : Volume 1, Volume 1 of Fortran Numerical Recipes : The Art of Scientific Computing*. 1992.
- [23] G. Barnes, K. D. Leka, C. J. Schrijver, T. Colak, R. Qahwaji, O. W. Ashamari, Y. Yuan, J. Zhang, R. T. J. McAteer, D. S. Bloomfield, P. A. Higgins, P. T. Gallagher, D. A. Falconer, M. K. Georgoulis, M. S. Wheatland, C. Balch, T. Dunn, and E. L. Wagner. A

- Comparison of Flare Forecasting Methods, I : Results from the "All-Clear" Workshop. *The Astrophysical Journal*, 829(2) :89, September 2016. ISSN 1538-4357. doi : 10.3847/0004-637X/829/2/89. URL <http://arxiv.org/abs/1608.06319>. arXiv : 1608.06319.
- [24] P. H. Scherrer, R. S. Bogart, R. I. Bush, J. T. Hoeksema, A. G. Kosovichev, J. Schou, W. Rosenberg, L. Springer, T. D. Tarbell, A. Title, C. J. Wolfson, I. Zayer, and MDI Engineering Team. The Solar Oscillations Investigation - Michelson Doppler Imager. *Solar Physics*, 162 :129–188, December 1995. doi : 10.1007/BF00733429. URL <http://adsabs.harvard.edu/abs/1995SoPh..162..129S>.
- [25] Carolus J. Schrijver. A Characteristic Magnetic Field Pattern Associated with All Major Solar Flares and Its Use in Flare Forecasting. *The Astrophysical Journal*, 655 (2) :L117–L120, January 2007. ISSN 1538-4357. doi : 10.1086/511857. URL <https://doi.org/10.1086%2F511857>. Publisher : IOP Publishing.
- [26] Dattaraj B. Dhuri, Shraavan M. Hanasoge, and Mark C. M. Cheung. Machine learning reveals systematic accumulation of electric current in lead-up to solar flares. *Proceedings of the National Academy of Sciences*, 116(23) :11141–11146, June 2019. ISSN 0027-8424, 1091-6490. doi : 10.1073/pnas.1820244116. URL <http://arxiv.org/abs/1905.10167>. arXiv : 1905.10167.
- [27] Yuan Yuan, Frank Y. Shih, Ju Jing, and Hai-Min Wang. Automated flare forecasting using a statistical learning technique. *Research in Astronomy and Astrophysics*, 10 : 785–796, August 2010. ISSN 1674-4527. doi : 10.1088/1674-4527/10/8/008. URL <http://adsabs.harvard.edu/abs/2010RAA....10..785Y>.
- [28] J. S. Cramer. The Origins of Logistic Regression. Technical Report 02-119/4, Tinbergen Institute, December 2002. URL <https://ideas.repec.org/p/tin/wpaper/20020119.html>. Publication Title : Tinbergen Institute Discussion Papers.
- [29] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. In *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152. ACM Press, 1992.
- [30] Antoine Strugarek, Allan S. Brun, Paul Charbonneau, and Nicole Vilmer. Sandpile Models and Solar Flares : Eigenfunction Decomposition for Data Assimilation. *Proceedings of the International Astronomical Union*, 13(S335) :250–253, July 2017. ISSN 1743-9213, 1743-9221. doi : 10.1017/S1743921317007244. URL <https://www.cambridge>.



org/core/journals/proceedings-of-the-international-astronomical-union/  
article/sandpile-models-and-solar-flares-eigenfunction-decomposition-for-data-assim  
6217C836E893573DEE272CB39A5CCEE5. Publisher : Cambridge University Press.

- [31] K. D. Leka, Sung-Hong Park, Kanya Kusano, Jesse Andries, Graham Barnes, Suzy Bingham, D. Shaun Bloomfield, Aoife E. McCloskey, Veronique Delouille, David Falconer, Peter T. Gallagher, Manolis K. Georgoulis, Yuki Kubo, Kangjin Lee, Sangwoo Lee, Vasily Lobzin, JunChul Mun, Sophie A. Murray, Tarek A. M. Hamad Nageem, Rami Qahwaji, Michael Sharpe, Robert A. Steenburgh, Graham Steward, and Michael Terkildsen. A Comparison of Flare Forecasting Methods. II. Benchmarks, Metrics, and Performance Results for Operational Solar Flare Forecasting Systems. *The Astrophysical Journal Supplement Series*, 243(2) :36, August 2019. ISSN 0067-0049. doi : 10.3847/1538-4365/ab2e12. URL <https://doi.org/10.3847/2F1538-4365/2Fab2e12>. Publisher : American Astronomical Society.
- [32] K. D. Leka, Sung-Hong Park, Kanya Kusano, Jesse Andries, Graham Barnes, Suzy Bingham, D. Shaun Bloomfield, Aoife E. McCloskey, Veronique Delouille, David Falconer, Peter T. Gallagher, Manolis K. Georgoulis, Yuki Kubo, Kangjin Lee, Sangwoo Lee, Vasily Lobzin, JunChul Mun, Sophie A. Murray, Tarek A. M. Hamad Nageem, Rami Qahwaji, Michael Sharpe, Rob Steenburgh, Graham Steward, and Michael Terkildsen. A Comparison of Flare Forecasting Methods. III. Systematic Behaviors of Operational Solar Flare Forecasting Systems. *The Astrophysical Journal*, 881(2) :101, August 2019. ISSN 1538-4357. doi : 10.3847/1538-4357/ab2e11. URL <http://arxiv.org/abs/1907.02909>. arXiv : 1907.02909.
- [33] M. C. M. Cheung, M. Rempel, G. Chintzoglou, F. Chen, P. Testa, J. Martínez-Sykora, A. Sainz Dalda, M. L. DeRosa, A. Malanushenko, V. Hansteen, B. De Pontieu, M. Carlsson, B. Gudiksen, and S. W. McIntosh. A comprehensive three-dimensional radiative magnetohydrodynamic simulation of a solar flare. *Nature Astronomy*, 3 : 160–166, November 2019. ISSN 2397-3366. doi : 10.1038/s41550-018-0629-3. URL <http://adsabs.harvard.edu/abs/2019NatAs...3..160C>.
- [34] B. R. Dennis. Solar hard X-ray bursts. *Solar Physics*, 100 :465–490, October 1985. doi : 10.1007/BF00158441. URL <http://adsabs.harvard.edu/abs/1985SoPh...100..465D>.

- [35] Markus J. Aschwanden, Robert A. Stern, and Manuel Güdel. Scaling laws of solar and stellar flares. *The Astrophysical Journal*, 672(1) :659–673, January 2008. ISSN 0004-637X, 1538-4357. doi : 10.1086/523926. URL <http://arxiv.org/abs/0710.2563>. arXiv : 0710.2563.
- [36] Kosuke Namekata, Takahito Sakaue, Kyoko Watanabe, Ayumi Asai, Hiroyuki Maehara, Yuta Notsu, Shota Notsu, Satoshi Honda, Takako T. Ishii, Kai Ikuta, Daisaku Nogami, and Kazunari Shibata. Statistical Studies of Solar White-light Flares and Comparisons with Superflares on Solar-type Stars. *The Astrophysical Journal*, 851 :91, December 2017. doi : 10.3847/1538-4357/aa9b34. URL <http://adsabs.harvard.edu/abs/2017ApJ...851...91N>.
- [37] Paul Charbonneau, Scott McIntosh, Hanli Liu, and J. Bogdan. Avalanche models for solar flares (Invited Review). *Solar Physics*, pages 321–353, 2001. doi : 10.1023/A:1013301521745. URL <https://opensky.ucar.edu/islandora/object/articles%3A8921/>.
- [38] Markus J. Aschwanden, Norma B. Crosby, Michaila Dimitropoulou, Manolis K. Georgoulis, Stefan Hergarten, James McAteer, Alexander V. Milovanov, Shin Mineshige, Laura Morales, Naoto Nishizuka, Gunnar Pruessner, Raul Sanchez, A. Surja Sharma, Antoine Strugarek, and Vadim Uritsky. 25 Years of Self-Organized Criticality : Solar and Astrophysics. *Space Science Reviews*, 198(1-4) :47–166, January 2016. ISSN 0038-6308, 1572-9672. doi : 10.1007/s11214-014-0054-6. URL <http://link.springer.com/10.1007/s11214-014-0054-6>.
- [39] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality : An explanation of the 1/f noise. *Physical Review Letters*, 59(4) :381–384, July 1987. doi : 10.1103/PhysRevLett.59.381. URL <https://link.aps.org/doi/10.1103/PhysRevLett.59.381>. Publisher : American Physical Society.
- [40] Henrik Jeldtoft Jensen. *Self-Organized Criticality : Emergent Complex Behavior in Physical and Biological Systems*. Cambridge University Press, Cambridge ; New York, 1 edition edition, April 1998. ISBN 978-0-521-48371-1.
- [41] Zeev Olami, Hans Jacob S. Feder, and Kim Christensen. Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical Review Letters*, 68(8) :1244–1247, February 1992. doi : 10.1103/PhysRevLett.68.1244. URL

- <https://link.aps.org/doi/10.1103/PhysRevLett.68.1244>. Publisher : American Physical Society.
- [42] M. S. Wheatland. The Origin of the Solar Flare Waiting-Time Distribution. *The Astrophysical Journal Letters*, 536 :L109–L112, June 2000. doi : 10.1086/312739. URL <http://adsabs.harvard.edu/abs/2000ApJ...536L.109W>.
- [43] James P. Norman, Paul Charbonneau, Scott W. McIntosh, and Han-Li Liu. Waiting-Time Distributions in Lattice Models of Solar Flares. *The Astrophysical Journal*, 557 : 891–896, August 2001. doi : 10.1086/321678. URL <http://adsabs.harvard.edu/abs/2001ApJ...557..891N>.
- [44] Ching Pui Hung, Allan Sacha Brun, Alexandre Fournier, Laurène Jouve, Olivier Talagrand, and Mustapha Zakari. Variational estimation of the large scale time dependent meridional circulation in the Sun : proofs of concept with a solar mean field dynamo model. *The Astrophysical Journal*, 849(2) :160, November 2017. ISSN 1538-4357. doi : 10.3847/1538-4357/aa91d1. URL <http://arxiv.org/abs/1710.02114>. arXiv : 1710.02114.
- [45] Valérie Pascual and Laurent Hascoët. Extension of TAPENADE toward Fortran 95. In Martin Bücker, George Corliss, Uwe Naumann, Paul Hovland, and Boyana Norris, editors, *Automatic Differentiation : Applications, Theory, and Implementations*, Lecture Notes in Computational Science and Engineering, pages 171–179, Berlin, Heidelberg, 2006. Springer. ISBN 978-3-540-28438-3. doi : 10.1007/3-540-28438-9\_15.



