

SORBONNE UNIVERSITÉ

ÉCOLE DOCTORALE DES SCIENCES DE L'ENVIRONNEMENT

LSCE/ESTIMR

Transport optimal pour quantifier l'évolution d'un attracteur climatique et corriger ses biais

Par Yoann ROBIN

Thèse de doctorat de climatologie

Dirigée par Pascal YIOU et Philippe NAVEAU

Présentée et soutenue publiquement le 4 avril 2018

Devant un jury composé de :

M. Francis CODRON	Professeur	<i>Examineur</i>
M. Valerio LUCARINI	Professeur	<i>Rapporteur</i>
M. Claudio PIANI	Professeur	<i>Rapporteur</i>
Mme. Bérengère DUBRULLE	Directrice de Recherche	<i>Examinatrice</i>
Mme. Anne-Catherine FAVRE	Professeure	<i>Examinatrice</i>
M. Pascal YIOU	Directeur de recherche	<i>Directeur de thèse</i>
M. Philippe NAVEAU	Directeur de recherche	<i>Directeur de thèse</i>

Remerciements

Il m'apparaît comme étant difficile de pouvoir décrire l'exercice que représente une thèse. Plusieurs années sur un sujet, que l'on commence sans le comprendre en pensant le comprendre, que l'on termine en comprenant qu'on ne commence qu'à le comprendre. Fort heureusement, j'ai pu compter sur de nombreux soutiens : intellectuels, émotionnels, et personnels.

Les premières personnes sont sans aucun doute mes directeurs, Pascal Yiou et Philippe Naveau. Guider un aspirant chercheur en maintenant un équilibre entre autonomie et accompagnement revient à mon sens à franchir le Grand Canyon sur un fil de nylon, les yeux bandés après avoir fait trois tours sur soi-même. C'est pourtant un exercice qu'ils ont réalisé à cloche pied, et je ne connais aucun mot assez fort pour exprimer ma reconnaissance. Je dirai donc simplement merci.

Valério Lucarini et Claudio Piani ont toute ma gratitude pour avoir accepté de rapporter mon manuscrit, ainsi que Bérengère Dubrulle, Anne-Catherine Favre et Francis Codron, pour leur participation à mon jury. Leurs critiques étaient constructives, leurs questions stimulantes. Qu'aurais-je pu espérer de plus ?

Je suis également gré à Aurélien Ribes et Pierre Ailliot pour avoir fait partie de mon comité de thèse. Leur enthousiasme lors de nos réunions et la pertinence de leurs interrogations m'ont bien aidé.

Un nombre important de collègues méritent aussi ma reconnaissance. À Mathieu Vac, pour m'avoir initié à la correction de biais, un immense merci. À Ara Arakelian, pour nos discussions sur le climat, les sciences, la politique, la psychologie et l'informatique (et un peu de tout), un grand merci. À ma co-thésarde, Aglaé Jézéquel, avec qui je partage le sujet de thèse, un grand merci pour nos discussions et notre co-soutien. Un grand merci s'impose aussi pour Davide Faranda, pour m'avoir aidé lors des premiers temps de ma thèse, et m'avoir permis d'exercer mon esprit critique. Un peu plus pêle-mêle, un grand merci à John Bazire (pour m'avoir caché mes affaires), à Boutheina Oueslati-Khalfaoui,

Soulivanh Thao (ton aide précieuse en R), Jean-Yves (pour ton panier), Françoise (pour ta bonne humeur), François Marabelle et Julien Bruna (pour avoir accepté d'installer tout ce qu'il me fallait sur le cluster). Et enfin très pêle-mêle, un grand merci à Gilles Ramstein, Didier Paillard, Didier Roche, Jeremy Pantou, Thomas Noel, et tous ceux que j'ai oubliés.

Je tiens aussi à remercier l'équipe et le laboratoire qui m'ont accueilli : ESTIMR et le LSCE, ainsi que le CEA et le CNRS. Ma thèse a été réalisée dans le cadre du projet européen d'ERC A2C2.

Finalement, comment aurais-je pu résister à ces années de thèse sans l'aide de mes proches ? Mille mercis pour ma femme, Camille Marceau, pour m'avoir soutenu, encouragé, supporté, relu mon manuscrit, et avoir même eu le courage de me demander de l'épouser ! Un immense merci pour ma mère, Florence Gerry, pour m'avoir enseigné la combativité, et à mon père, Éric Robin, pour m'avoir donné le goût des sciences. Un grand merci pour mon frère David Robin et ma belle soeur, Aurélie Lalbie, pour leur soutien. De larges remerciements aussi pour ma belle famille. À Patrick et Véronique Marceau, pour leur franc-parler, éternellement rafraichissant. À Amandine Marceau et Loïc Lecorvaisier, et nos soirées jeux. À Aurélien et Annabelle Marceau, pour les bons moments. Un autre grand merci à toute ma famille pour être venu me soutenir le jour de ma soutenance. Un grand merci final à mes amis. À Solène Colmet-Daâge, quoi de plus relaxant que l'étude de l'évolution de la place de la femme dans les James Bond ? À Cyril Fraioli pour nos soirées comics.

Une dernière mention à mes boules de poils : Maya, Schrodinger et Tigrou (une superbe Golden Retriever et deux monstrueux chats). Leur effet relaxant ne peut qu'aider un thésard à garder le moral !

Introduction

Cette thèse a été réalisée dans le domaine de l'étude du climat. La première chose à comprendre, pour quelqu'un (comme moi) qui n'avait jamais étudié le climat, est sa définition. Pour cela, j'ai besoin de définir le *système climatique*. Je reproduis ici celle qui est donnée par le glossaire du GIEC (IPCC, 2014).

Système extrêmement complexe comprenant cinq grands éléments : l'atmosphère, l'hydrosphère, la cryosphère, la lithosphère et la biosphère, et qui résulte de leurs interactions. Ce système évolue avec le temps sous l'effet de sa propre dynamique interne et en raison de forçages externes tels que les éruptions volcaniques, les variations de l'activité solaire ou les forçages anthropiques (notamment les variations de la composition de l'atmosphère ou les changements d'affectation des terres).

Il est intéressant de voir que dans cette définition, nous avons deux grands objets distincts. D'une part, les différents éléments de notre planète que sont l'atmosphère, l'hydrosphère, etc, qui existent en tant qu'objets propres; et d'autre part, les interactions entre ces objets, et les éléments externes qui peuvent les altérer. Les éléments du système climatique sont autant de variables du système (des milliers de variables peuvent décrire un seul élément), tandis que les interactions sont les équations qui vont décrire les relations de ces éléments entre eux. Dans l'état actuel des connaissances, il n'est pas possible de décrire exactement le fonctionnement de ce système. Afin de comprendre ce qui fait la complexité du climat, je propose de décrire brièvement le fonctionnement de la dynamique atmosphérique.

L'élément qui engendre la dynamique atmosphérique est assez simple : le soleil *chauffe* notre planète, mais l'équateur et les pôles sont en *déséquilibre*. Ce phénomène est représenté sur la Fig. 1a). Les lignes rouges représentent les rayons du Soleil. La Terre étant (presque) sphérique, la quantité d'énergie reçue par la zone (A) est la même que celle de la zone (B), qui est deux fois plus grande. Autrement dit, la quantité d'énergie reçue

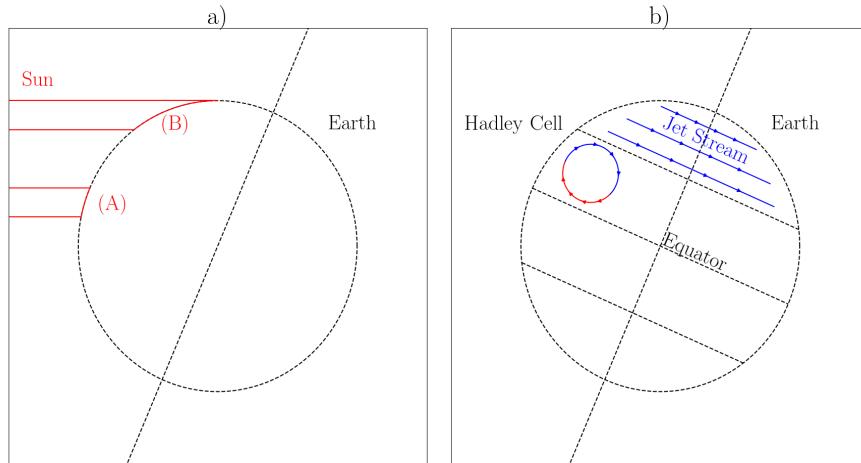


FIGURE 1 – **a)** Représentation schématique du forçage solaire. **b)** Représentation schématique de la dynamique atmosphérique.

par m^2 est plus faible aux pôles qu'à l'équateur. Le bilan énergétique est donc en surplus à l'équateur, et en déficit aux pôles. Ce déséquilibre énergétique engendre un transfert de chaleur de l'équateur vers les pôles par un mouvement de l'atmosphère. Comme la Terre est en rotation, ce transfert est dévié vers l'Est au delà de 30° de latitude par la force de Coriolis. Ceci engendre deux « cellules » dans les bandes $0 - 30^\circ \text{ N/S}$ de latitude, nommées *cellules de HADLEY (1735)*, comme représenté sur la Fig. 1b). Au delà de 30° , la différence entre l'air « chaud » venant des tropiques, et l'air « froid » venant des pôles entraîne un fort courant atmosphérique zonal appelé *jet-stream*. Ce courant est particulièrement instable à cause des perturbations de l'équilibre géostrophique à l'interface entre la cellule de HADLEY (1735) et les moyennes latitudes (i.e. les vents horizontaux ne sont plus proportionnels au gradient de pression). Cette instabilité est chaotique (dans un sens qui sera précisé dans le premier chapitre), ce qui génère la variabilité climatique.

Ces instabilités dans l'atmosphère nous amènent à une définition du climat, comme les *statistiques* (dans un sens que je préciserai après) du système climatique. Pour pouvoir faire ces statistiques, il nous faut des données du système climatique.

Nous pouvons tout d'abord utiliser les données a posteriori, que sont les observations. Ces observations peuvent être directes pour l'état récent (i.e. les 40 dernières années) grâce aux satellites et aux stations réparties sur le globe, ou indirectes pour l'état lointain dans le passé, comme les carottes glacières qui permettent d'estimer la température et la composition de l'atmosphère par l'analyse des teneurs en isotopes de l'oxygène dans des bulles d'air qui sont emprisonnées. Ces données soulèvent de nombreux problèmes, comme leurs répartitions spatiales et temporelles, qui ne sont pas uniformes.

D'autre part, nous avons à notre disposition les modèles climatiques. Ce qui est appelé modèle est en réalité le couplage de plusieurs modèles simulant les différents éléments du système climatique, et leurs interactions. Par exemple, les modèles atmosphériques et océaniques sont dérivés des équations de Navier Stokes. Ils peuvent être à l'échelle de la planète, comme les GCM (General Climate Model, e.g., TAYLOR et al., 2012), ou à des échelles régionales, comme les RCM (Regional Climate Model, e.g. le projet CORDEX). Ces modèles tentent de prédire une trajectoire du système climatique pour un (ou plusieurs) jeu (de données) de conditions initiales. Pourtant, leurs valeurs à un lieu / temps précis ne peuvent pas être interprétées comme une *prédiction*. Ces modèles cherchent en fait à simuler le climat, dont la définition, tirée du glossaire du GIEC (IPCC, 2014) est

Au sens étroit du terme, le climat désigne en général le temps moyen ou, plus précisément, se réfère à une description statistique fondée sur les moyennes et la variabilité de grandeurs pertinentes sur des périodes variant de quelques mois à des milliers, voire à des millions d'années (la période type, définie par l'Organisation météorologique mondiale, est de 30 ans). Ces grandeurs sont le plus souvent des variables de surface telles que la température, la hauteur de précipitation et le vent. Dans un sens plus large, le climat désigne l'état du système climatique, y compris sa description statistique.

Ceci explique que les valeurs en elles-mêmes, données par les modèles, ne sont pas importantes. Ce sont leurs statistiques, et l'évolution de ces statistiques qui déterminent une simulation du climat. En particulier, je vais m'intéresser à la *distribution de probabilité* des variables climatiques (dont la moyenne et la variance ne sont que des indicateurs). Tout forçage externe entraînant l'altération des éléments du système climatique, ou de leurs dynamiques, engendre des modifications de cette distribution, et est donc quantifiable.

Prenons comme exemple une simulation de la température en Ile-de-France par le modèle de l'IPSL (DUFRESNE et al., 2013). Cette simulation est dite de *contrôle*, car aucun forçage (autre que les forçages naturels) n'est appliqué. J'ai représenté sur la Figure 2 l'histogramme de 100 ans de données quotidiennes d'une telle simulation. Cet histogramme est une *estimation de la distribution de probabilité du système climatique* (pour la température, en Ile-de-France, sur une période de 100 ans, et pour une simulation). Une question naturelle est :

Cette distribution de probabilité représente-t-elle le système climatique ?

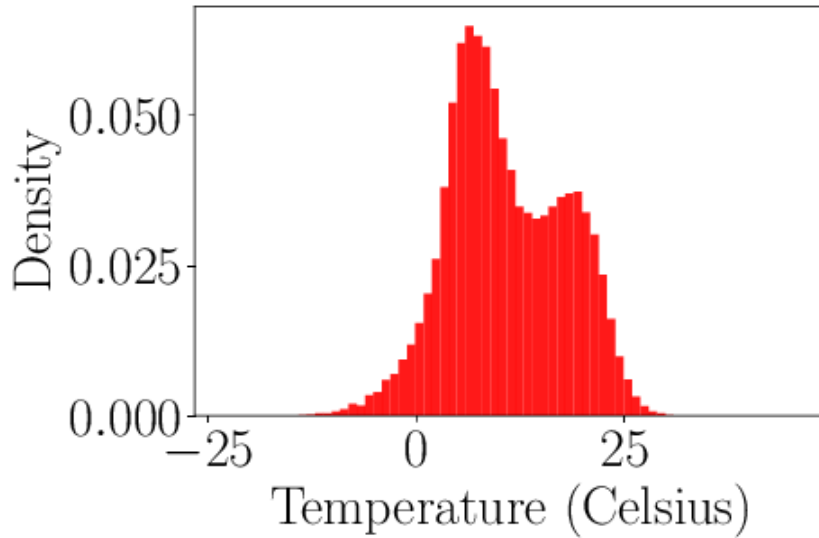


FIGURE 2 – Histogramme des températures quotidiennes en Ile-de-France issues d’une simulation de contrôle du modèle de l’IPSL (DUFRESNE et al., 2013). Le forçage radiatif est fixé à celui de l’ère pré-industrielle, i.e. avant 1850. La simulation fait 250 ans, correspondant à la période 1850-2100. Seules les valeurs de la période 2000-2100 ont été utilisées pour générer cet histogramme. Les deux modes (les bosses) correspondent à l’hiver et l’été.

La théorie des systèmes dynamiques fournit une réponse à cette question, au travers d’une notion mathématique nommée la *mesure SRB* (Sinai-Ruelle-Bowen). Cette mesure contient toute l’information du système sous jacent, et s’estime avec un histogramme. Pour la définir, j’ai eu besoin d’outils supplémentaires, que sont les attracteurs étranges.

Une fois les statistiques du climat définies, le problème est que cette mesure est soumise à une déformation en temps. Ceci peut être dû à des événements naturels (variations du rayonnement solaire, volcans... etc), ou d’origine anthropogénique (i.e. le changement climatique). Il faut donc pouvoir *mesurer* cette déformation. Classiquement, des métriques basées sur la distance Euclidienne sont utilisées, mais représentent mal la dynamique sous jacente. Le problème qui m’a occupé a donc été :

Comment mesurer la déformation d’une distribution de probabilité en tenant compte de sa dynamique ?

De nombreux problèmes compliquent la situation. La quantité de données disponibles pour estimer une distribution de probabilité peut ne pas être suffisante. Les données peuvent être entachées de bruit. Les modèles sont biaisés. Ce dernier problème peut théo-

riquement être résolu par un outil capable de mesurer finement les différences entre distributions de probabilités. La mesure de la différence offre la possibilité de la *corriger*. Ce domaine du climat se nomme la *correction de biais*. Je propose une réponse à ce problème dans ce mémoire.

La structure de ce mémoire est la suivante. Dans le Chapitre 1, je rappellerai la description du climat comme un système dynamique, ainsi que les outils de cette théorie, dont j'aurai besoin. Ce chapitre est construit autour du modèle de LORENZ (1963), afin de guider l'intuition. Dans le Chapitre 2, je présenterai la théorie du transport optimal, ainsi que la distance de Wasserstein qui en provient. Cette métrique tient naturellement compte de la dynamique, et sera l'outil utilisé pour mesurer la déformation du climat. Le Chapitre 3 est un article publié, qui montre comment utiliser la distance de Wasserstein pour quantifier la déformation du climat. La méthodologie est testée sur un modèle similaire au LORENZ (1963), dû aussi à LORENZ (1984), mais forcé par un cycle saisonnier et un équivalent du changement climatique. Le Chapitre 4 est un article en cours d'écriture qui présente une méthode de correction de biais stochastique, basée sur le transport optimal.

Table des matières

1	Description du climat comme un système dynamique	13
1.1	L'attracteur de Lorenz	13
1.1.1	Histoire et définition	14
1.1.2	Sensibilité aux conditions initiales : la source du chaos	16
1.1.3	Caractérisation du chaos par la dimension de l'attracteur	18
1.1.4	Les statistiques du modèle de LORENZ (1963)	18
1.1.5	Résumé	19
1.2	Systèmes dynamiques autonomes	20
1.2.1	Définition des systèmes dynamiques	20
1.2.2	Déterminer la dimension d'un attracteur	22
1.2.3	Caractérisation du comportement local de la dynamique : les expo- sants de Lyapunov	25
1.2.4	Statistiques d'un système dynamique : la mesure SRB	26
1.2.5	Résumé	28
1.3	Attracteur climatique	29
1.3.1	Existence d'un attracteur étrange climatique	29
1.3.2	Déformation de l'attracteur climatique	29
1.4	Dynamique non autonome	30
1.4.1	Système de LORENZ (1984)	30
1.4.2	Vision Lagrangienne et Eulerienne : des équations différentielles aux équations de transports	33
1.4.3	L'équivalent des orbites en dynamique non autonome : les snapshots attractors	34
1.5	Résumé et mise en place du problème	37
2	Transport Optimal	39
2.1	Introduction au transport optimal : le problème de Monge	40
2.2	Transport optimal de mesures empiriques	43
2.2.1	Formulation en terme de mesure de probabilité	43
2.2.2	Problème de Monge-Kantorovich	45

2.3	Distance de Wasserstein	46
2.4	Lien entre les systèmes dynamiques et le transport optimal	46
2.5	Intuitions des mécanismes d'action du transport optimal	48
2.5.1	Représentation du plan de transport comme mesure de probabilité	49
2.5.2	Visualisation dynamique du plan de transport	50
2.5.3	Résistance au bruit et influence de l'histogramme	50
2.6	Exemple d'utilisation de la distance de Wasserstein dans le climat	53
2.7	Résumé	57
3	Application à des systèmes dynamiques chaotiques non autonomes	59
3.1	Abstract (ROBIN et al., 2017)	59
3.2	Detecting changes in forced climate attractors with Wasserstein distance	60
3.3	Compléments : Détection des scénarios RCP dans CMIP5	74
3.3.1	Application aux températures	75
3.3.2	Application aux précipitations	75
3.4	Résumé	76
4	Correction de biais multivarié	79
4.1	Abstract	79
4.2	Multivariate stochastic bias corrections with optimal transport	80
4.3	Résumé	112
	Conclusion et perspectives	113
A	Utilisation de la distance de Wasserstein pour valider un indicateur	117
A.1	Abstract (RODRIGUES et al., 2017, accepted in J. Clim.)	117
A.2	Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis	118
	Bibliographie	160
	Liste des figures	167
	Index	171

Chapitre 1

Description du climat comme un système dynamique

Le but de ce chapitre est d'introduire les notions sur lesquelles je me suis appuyé tout au long de cette thèse, et qui m'ont permis d'envisager le climat en termes statistiques. La section 1.1 présentera le système de LORENZ (1963), qui est un modèle jouet du climat donnant une première intuition des notions guidant ma pensée : *chaos*, *système dynamique*, *attracteur étrange* et *mesure SRB*. La mesure SRB est l'outil central modélisant le climat. La section 1.2 formalisera ces concepts, pour l'instant abscons. Je discuterai ensuite dans la section 1.3 de l'existence d'un attracteur climatique, et des *forçages* qui affectent le système climatique, comme le forçage anthropogénique. Je présenterai finalement dans la section 1.4 les systèmes dynamiques dit *non-autonomes*, au travers du modèle de LORENZ (1984). Ces systèmes tiennent justement compte des forçages. Je présenterai la distribution de probabilité appropriée qu'il faut considérer dans ce cas.

1.1 L'attracteur de Lorenz

Le chaos va être expliqué dans cette section à l'aide du système de LORENZ (1963), qui est un exemple de système dynamique. Un système est chaotique si deux états initiaux aussi proches que nous le voulons se comportent différemment après un laps de temps court, mais restent confinés dans un intervalle fixe de valeurs. Ce faisant, les trajectoires d'un tel système dessinent un objet dans cet intervalle de valeurs nommé un *attracteur étrange*. Un excellent résumé de l'état de l'art sur le système de LORENZ (1963) a été écrit par VIANA (2000).

1.1.1 Histoire et définition

Bien que le chaos ait été entrevu par POINCARÉ (1892), il est tombé dans l'oubli pendant un demi-siècle. Le météorologue E. N. Lorenz a refait sa découverte lors de l'étude de la convection. Il cherchait un modèle simplifié qui décrit l'échauffement d'un liquide par le bas. Le liquide chauffé monte, se refroidit, et replonge. Partant des équations de SALTZMAN (1962), il en a dérivé un système tri dimensionnel simplifié, aujourd'hui appelé système de LORENZ (1963). Les équations simplifiées de Lorenz se présentent comme une équation différentielle à trois inconnues, avec des relations non linéaires quadratiques.

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}_{l63}(\mathbf{x}) := \begin{pmatrix} \sigma(y - x) \\ -xz + rx - y \\ xy - bz \end{pmatrix}, \quad \mathbf{x} = (x, y, z). \quad (1.1)$$

Les variables x , y et z sont respectivement proportionnelles à l'intensité du mouvement convectif, la variation de température horizontale et la variation de température verticale. De plus, trois paramètres peuvent être modifiés, σ , r et b . Le nombre σ est le nombre de Prandtl, r est une normalisation du nombre de RAYLEIGH (1916) par une valeur critique. Ce système ayant trois variables, ses solutions vivent dans $\mathbb{R}^3$, nommé *l'espace des phases*. Une solution de ce système est représentée sur la Figure 1.1. Nous pouvons voir que cette solution s'enroule sur un objet en forme de Papillon. En fait, *toutes* les solutions, après un laps de temps très court, viennent s'enrouler sur cet objet et sont « piégées » dedans. C'est cet objet qui est appelé un *attracteur*.

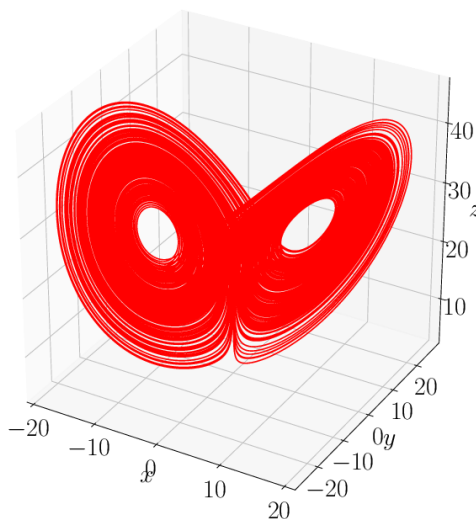


FIGURE 1.1 – Exemple d'une solution du système de LORENZ (1963).

Lorenz a calculé une solution de ce système pour les paramètres $\sigma = 10$, $r = 28$ et $b = 8/3$. Ce sont les paramètres couramment utilisés dans la littérature (et que j’ai utilisé tout au long de cette thèse). Lors du calcul d’une trajectoire, i.e. pour une condition initiale (x_0, y_0, z_0) , le calcul d’une approximation de la solution $\mathbf{x}(t)$, le calculateur s’est interrompu. En effet, à l’époque la sortie avait lieu sur un rouleau de papier, qu’il fallait changer. Repartant d’une des dernières solutions calculées, il a continué l’intégration. À sa surprise, les nouvelles valeurs étaient différentes des anciennes valeurs ! Il a constaté que les valeurs imprimées (qu’il avait utilisé pour continuer l’intégration) étaient des valeurs tronquées de celles utilisées par le calculateur. Autrement dit, il a redémarré l’intégration non pas à la valeur calculée par la machine, mais à une valeur très voisine. Ce faisant, il a mis en évidence l’une des propriétés élémentaires du chaos : deux trajectoires ayant des conditions initiales très proches divergent exponentiellement vite. Ce phénomène est illustré sur la Figure 1.2. Les deux courbes bleu et rouge sont deux trajectoires partant de deux points proches : $(5.16, -0.87, 30.64)$ et $(4.99, -0.53, 30.02)$. Nous pouvons voir que ces trajectoires se séparent après un temps court, rendant impossible toute prédiction à moyen terme. En effet, les conditions initiales d’un modèle ne pouvant être connues qu’avec une précision finie, elles sont différentes des conditions initiales de la réalité. Si le système est chaotique, alors la simulation s’écartera rapidement de la réalité. Ce phénomène avait déjà été compris par POINCARÉ (1892), qui l’avait formulé ainsi

Si nous connaissions exactement les lois de la nature et la situation de l’univers à l’instant initial, nous pourrions prédire la situation de ce même univers à un instant ultérieur. Mais, lors même que les lois naturelles n’auraient plus de secret pour nous, nous ne pourrions connaître la situation initiale qu’approximativement (...). Il peut arriver que des petites différences dans les conditions initiales en engendrent de très grandes dans les phénomènes finaux ; une petite erreur sur les premières produirait une erreur énorme sur les derniers. La prédiction devient impossible et nous avons le phénomène fortuit

POINCARÉ (1892)

Cependant, jusqu’aux travaux de Lorenz, ce phénomène n’était considéré possible que pour des systèmes hautement complexes, ayant un nombre important de degrés de liberté, tel que le problème à N corps. L’attracteur de Lorenz est le premier système simple (la non linéarité est uniquement quadratique) exhibant du chaos. Je propose de décrire d’où vient cette sensibilité aux conditions initiales.

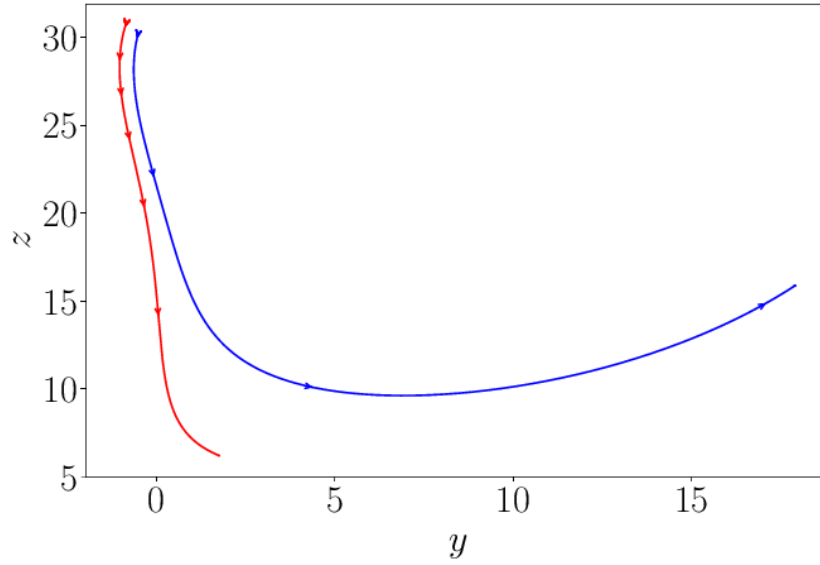


FIGURE 1.2 – Deux trajectoires de l’attracteur de LORENZ (1963) prises aux conditions initiales suivantes : $(5.16, -0.87, 30.64)$ (courbe rouge) et $(4.99, -0.53, 30.02)$ (courbe bleue), et intégrées pendant 0.6 unité de temps. Les trajectoires ont été projetées dans le plan (y, z) .

1.1.2 Sensibilité aux conditions initiales : la source du chaos

Commençons par considérer les points fixes du système. Rappelons qu’un point fixe est une solution constante au cours du temps, i.e. $\mathbf{x}_0$ est un point fixe si $\mathbf{x}(t) \equiv \mathbf{x}_0$ pour tout temps t . Les points fixes peuvent être calculés en résolvant le système $\mathbf{v}_{l63}(\mathbf{x}) = 0$. Dans le cas de l’attracteur de LORENZ (1963), nous trouvons trois points fixes, l’origine, et deux points symétriques situés au centre des ailes

$$\left(\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1 \right).$$

Le comportement de $\mathbf{v}_{l63}$ au voisinage de l’origine est donné par sa différentielle :

$$d\mathbf{v} = \begin{pmatrix} -\sigma & \sigma & 0 \\ -z+r & -1 & -x \\ y & x & -b \end{pmatrix}.$$

En prenant $(x, y, z) = (0, 0, 0)$, les trois valeurs propres de la différentielle à l’origine sont $\kappa_1 = -22.8$, $\kappa_2 = 11.8$ et $\kappa_3 = -2.66$. En linéarisant $\mathbf{v}_{l63}$ dans les directions propres de chaque valeur propre, nous trouvons comme estimation que $\mathbf{x}(t) \simeq \exp(\kappa_i t)$. Pour $\kappa_i < 0$, $\mathbf{x}(t)$ converge exponentiellement vite vers l’origine. Cette direction est dite *stable*. Pour $\kappa_i > 0$, $\mathbf{x}(t)$ s’éloigne exponentiellement vite de l’origine. Cette direction est dite *instable*.

Donc l'origine peut se décomposer en un plan stable qui attire $\mathbf{x}(t)$, et une droite instable qui repousse $\mathbf{x}(t)$. Ces notions de directions stables et instables sont la source du chaos. Pour se représenter le phénomène, examinons la Figure 1.3.

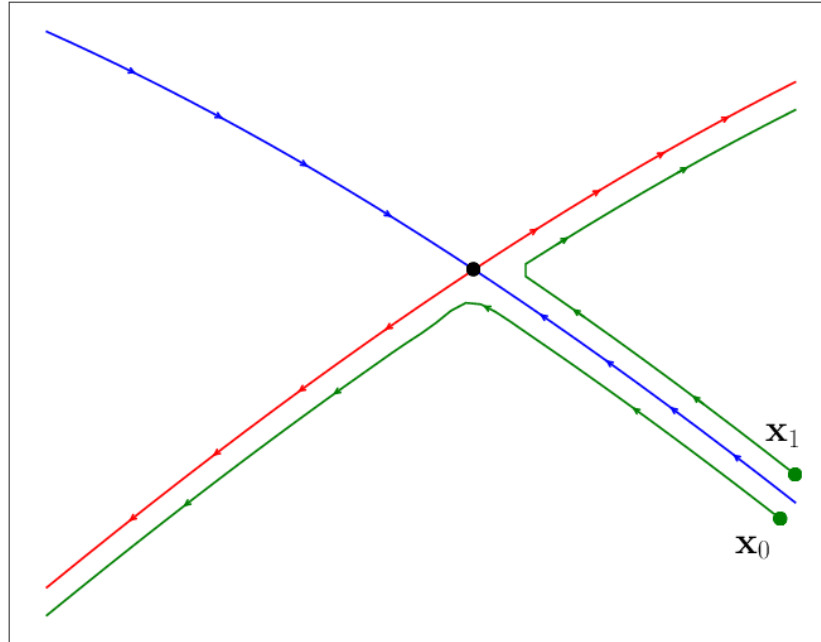


FIGURE 1.3 – Représentation d'un point fixe (en noir) d'un système dynamique. La courbe bleue représente sa direction stable, qui attire toutes les trajectoires. La courbe rouge représente sa direction instable, qui repousse toutes les trajectoires. Les courbes vertes sont les trajectoires des points $\mathbf{x}_0$ et $\mathbf{x}_1$.

La courbe rouge représente une direction instable, la bleue une direction stable. Le point d'intersection noir est un point fixe. Imaginons deux points $\mathbf{x}_0$ et $\mathbf{x}_1$ de part et d'autre de la direction stable (i.e. de la courbe rouge). Dans un premier temps, ils restent proches en suivant la direction stable. Mais arrivés à proximité du point fixe, la direction instable les repousse exponentiellement vite dans des directions opposées. C'est le phénomène de *sensibilité aux conditions initiales*. Les notions de directions stables et instables se généralisent aux points périodiques (i.e. $\mathbf{x}(t + \tau) = \mathbf{x}(t)$ pour tout temps t), voire n'importe quel point (HIRSCH et al., 2006 ; ROBINSON, 1999).

Maintenant dans le cas du chaos, et en particulier du modèle de Lorenz, les points périodiques sont en infinité et forment un sous ensemble dense (ARROYO et PUJALS, 2007). Leurs variétés stables et instables se recoupent. Cette « zone » de chaos est dans une boîte, située au centre du papillon. Ce résultat a d'abord été démontré dans un cas simplifié par WILLIAMS (1979), et étendu à l'attracteur de LORENZ (1963) par TUCKER

(1999).

1.1.3 Caractérisation du chaos par la dimension de l'attracteur

Pour étudier la dimension de l'attracteur, je propose d'examiner comment le volume d'un cube se comporte sous l'action de la dynamique du système de LORENZ (1963). Notons V_0 le volume d'un cube contenant tout l'attracteur, et $V(t)$ ce volume au temps t sous l'action du système de LORENZ (1963). Nous avons alors l'estimation suivante

$$V(t) \approx V_0 e^{\text{div}(\mathbf{v}_{l63})t}.$$

La divergence se calcule avec la différentielle obtenue précédemment :

$$\text{div}(\mathbf{v}_{l63}) = \frac{\partial \mathbf{v}_{l63}}{\partial x} + \frac{\partial \mathbf{v}_{l63}}{\partial y} + \frac{\partial \mathbf{v}_{l63}}{\partial z} = -(\sigma + b + 1) \approx -13.66.$$

Nous obtenons alors l'estimation suivante :

$$V(t) \approx e^{-13.66t} \times V_0 \approx 10^{-6t} \times V_0.$$

Autrement dit, le volume est contracté par un facteur 10^6 à chaque unité de temps. Le volume de notre cube est donc complètement écrasé, et le cube se retrouve confiné sur l'attracteur. La dimension de l'attracteur ne peut donc pas être égale à 3, mais l'objet dessiné est « plus gros » qu'une courbe (dimension 1). Il est assez semblable à un plan (dimension 2), comme nous pouvons le voir sur la Figure 1.4. Sa dimension réelle (dans un sens que nous précisons dans la Section 1.2.2) a été calculée par GRASSBERGER et PROCACCIA (1983), et est égale à environ 2.05. Autrement dit, un objet qui occupe légèrement plus d'espace qu'un objet de dimension 2. Cette dimension non entière est caractéristique des systèmes chaotiques. Les attracteurs ayant une dimension non entière sont appelés *attracteurs étranges* par RUELLE et TAKENS (1971).

1.1.4 Les statistiques du modèle de LORENZ (1963)

Si les trajectoires du modèle de LORENZ (1963) sont imprévisibles, quelle information fiable peut en être extraite? Mais même si la trajectoire est imprévisible, l'attracteur dessiné ne dépend pas des conditions initiales. A fortiori, la probabilité que la trajectoire se situe dans une zone de $\mathbb{R}^3$ ne dépend pas non plus de la condition initiale. Dit autrement, *l'histogramme* des valeurs prises par une longue trajectoire *ne dépend pas* des conditions initiales. Par exemple, j'ai représenté sur la Figure 1.5 deux histogrammes de la variable x . Nous pouvons voir qu'ils sont presque exactement identiques, les différences étant un mélange d'erreurs numériques, de trajectoires finies et des paramètres de la discrétisation utilisée. Ce sont ces statistiques, au travers des histogrammes, qui constituent le climat.

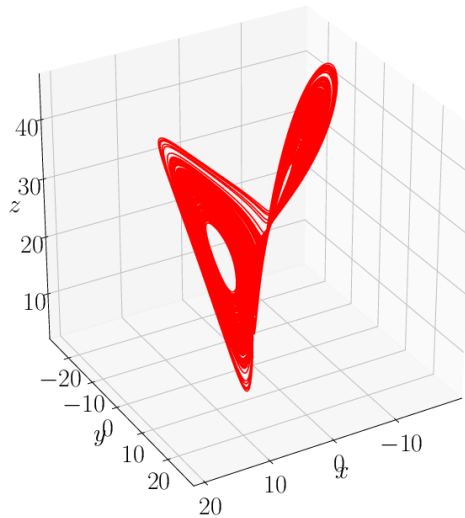


FIGURE 1.4 – Attracteur de LORENZ (1963). La projection est choisie de manière à mettre en valeur que l’attracteur a une dimension proche de 2.

Pour définir correctement cet objet, je vais avoir besoin de concepts provenant de la théorie des systèmes dynamiques, que je vais maintenant introduire.

1.1.5 Résumé

Cette présentation du modèle de LORENZ (1963) a permis de dégager plusieurs concepts, qui sont la base de ma représentation du climat :

1. L’équation différentielle (1.1) engendre un *système dynamique continu* dans $\mathbb{R}^3$, nommé *l’espace des phases*.
2. Deux trajectoires commençant très près l’une de l’autre divergent exponentiellement vite, mais dessinent le même objet en forme de papillon, nommée un *attracteur*. Nous avons un système *chaotique*.
3. La dimension de cet attracteur est plus petite que celle de l’espace des phases, mais semble légèrement plus grande que 2, l’attracteur est dit *étrange*.
4. Les statistiques de ce système ne dépendent pas non plus des conditions initiales, et semblent dessiner une mesure de probabilité particulière, nommée une *mesure SRB* (Sinai-Ruelle-Bowen).

Tous ces concepts ont été présentés sur un exemple particulier, je vais maintenant présenter la théorie qui permet de les formaliser.

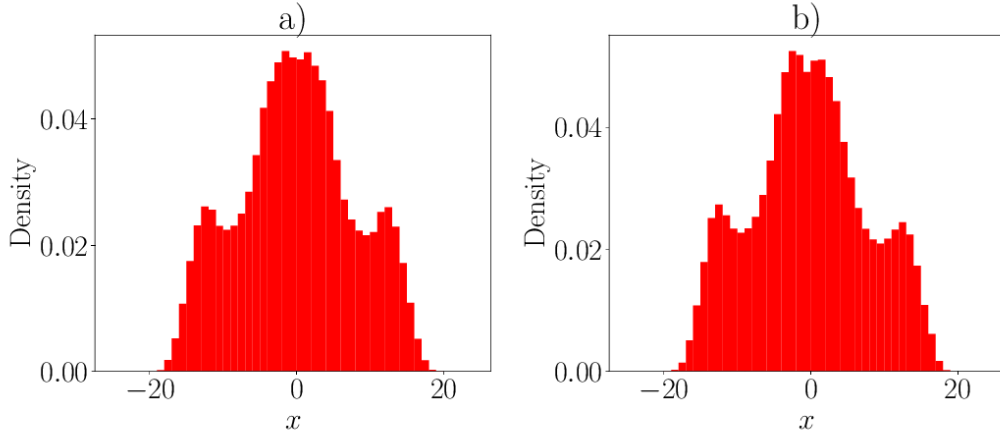


FIGURE 1.5 – Histogramme de la coordonnée x d’une trajectoire du modèle de LORENZ (1963). **a)** La condition initiale de la trajectoire est $(1, 1, 1)$. **b)** La condition initiale de la trajectoire est $(1, -1, -1)$.

1.2 Systèmes dynamiques autonomes

1.2.1 Définition des systèmes dynamiques

Le système de LORENZ (1963) considéré ci-dessus appartient à une classe plus générale d’objets, nommée les *systèmes dynamiques continus*. Ils ont été formalisés par SMALE (1967). Partant d’un champ de vecteurs $\mathbf{v} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ (qui est supposé suffisamment dérivable), une équation différentielle y est associée :

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}).$$

Remarquons que $\mathbf{v}$ ne dépend pas du temps. Ce type d’équation est dite *autonome*. Par exemple pour le système de LORENZ (1963), nous avons $\mathbf{v} = \mathbf{v}_{l63}$. Historiquement, l’étude des équations différentielles se fait par l’étude d’une solution $\mathbf{x}(t)$ pour une condition initiale $\mathbf{x}_0$ donnée. Afin d’étudier l’état de *toutes* les solutions à un instant t , POINCARÉ (1892) a introduit le *flot* $\varphi^t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ d’une équation différentielle. Cette fonction associe à *chaque* condition initiale $\mathbf{x}_0$ la solution au temps t , i.e. $\varphi^t(\mathbf{x}_0) := \mathbf{x}(t)$. Au lieu d’étudier l’évolution d’un point de l’espace le long du temps, c’est l’état de l’espace tout entier au temps t qui est considéré. Comme $\mathbf{v}$ ne dépend pas du temps, nous avons l’invariance temporelle suivante pour tout $t_0, t_1 \in \mathbb{R}$:

$$\varphi^{t_0} \circ \varphi^{t_1} = \varphi^{t_0+t_1}.$$

C’est cette propriété qui définit les *systèmes dynamiques continus*. Si maintenant nous fixons un $t_0 \in \mathbb{R}$, et que nous fixons $f := \varphi^{t_0}$, nous obtenons un *système dynamique discret*.

L'action de ce système est donnée en appliquant f de manière répétée, i.e. en partant de $\mathbf{x} \in \mathbb{R}^d$, et en considérant les éléments de la forme

$$f^n(\mathbf{x}) := \underbrace{(f \circ \dots \circ f)}_{n \text{ fois}}(\mathbf{x}) = \left(= \varphi^{nt_0}(\mathbf{x}_0) \right).$$

L'étude de ces deux types de systèmes dynamiques peut être différente, mais les concepts et propriétés présentés ici sont les mêmes. L'exposé peut être plus simple en utilisant un système continu, ou l'inverse. Dans les deux cas, un système dynamique agit sur un espace global (ici $\mathbb{R}^d$), cet espace est nommé *l'espace des phases*. Cet espace représente tous les états possibles que peut prendre le système dynamique.

Intéressons nous maintenant aux trajectoires d'un système dynamique. Nous avons vu que dans le cas du modèle de LORENZ (1963), elles sont dessinées par une solution de l'Équation (1.1). Ces solutions sont de la forme $\varphi^t(\mathbf{x}_0)$ (avec $\mathbf{x}_0$ une condition initiale). La trajectoire ainsi formée se nomme une *orbite* d'un système dynamique. Dans le cas des systèmes dynamiques discrets, les orbites sont de la forme $f^n(\mathbf{x}_0)$. La littérature note couramment $\mathcal{O}(\mathbf{x}_0)$ l'orbite d'un point $\mathbf{x}_0$. Nous pouvons maintenant définir formellement un attracteur d'un système dynamique φ^t , noté Λ , qui est caractérisé par les propriétés suivantes :

1. L'ensemble Λ est invariant, i.e. $\varphi^t(\Lambda) = \Lambda$,
2. Il existe $\mathbf{x} \in \Lambda$ tel que l'orbite de $\mathbf{x}$ passe aussi près que l'on veut de tous points de Λ (i.e. Λ est l'adhérence de $\mathcal{O}(\mathbf{x})$).
3. Le bassin d'attraction de Λ (l'ensemble des points dont l'orbite est piégée dans Λ) est de volume non nul dans $\mathbb{R}^d$.

Un attracteur est donc une « zone de l'espace des phases qui attire toutes les orbites voisines ». Remettons toutes ces définitions dans le contexte du système de LORENZ (1963). Sur la Figure 1.6, nous pouvons voir deux orbites du LORENZ (1963), avec deux conditions initiales très différentes. Pour la Figure 1.6a, nous avons pris $(80, 50, 50)$, et $(-50, -50, -50)$ pour la Figure 1.6b. Ces deux orbites tournent dans $\mathbb{R}^3$ avant de plonger dans le Papillon, noté Λ . Bien que les deux conditions initiales soient très différentes, après un certain temps elles sont piégées dans Λ . Elles sont donc dans le bassin d'attraction de Λ , qui est en fait tout $\mathbb{R}^3$. Le point 3. de la définition d'un attracteur est donc vérifié. De plus, les deux trajectoires parcourent tout Λ , le point 2. est donc vérifié. Finalement, si nous appliquons le flot de l'équation (1.1) à Λ , Λ reste invariant, validant le point 1. L'ensemble Λ (i.e. l'attracteur de Lorenz) est donc un attracteur au sens de la définition précédente.

Je viens de définir la notion d'attracteur, mais pour définir la notion d'attracteur

étrange, il faut pouvoir parler de la dimension d'un attracteur. Je vais donc parler de la dimension de *box-counting* et de la dimension de *corrélation*.

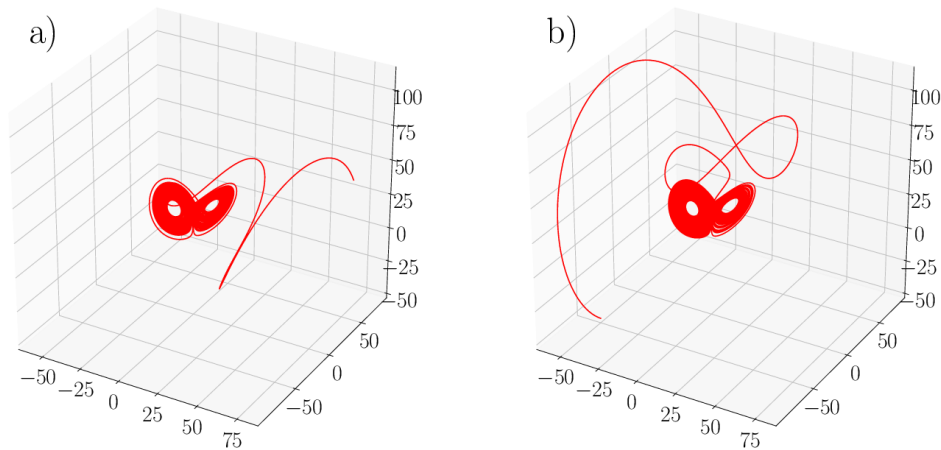


FIGURE 1.6 – Deux orbites du LORENZ (1963). Sur la Figure a), la condition initiale est $(80, 50, 50)$, tandis que c'est $(-50, -50, -50)$ sur la Figure b). Les deux orbites génèrent bien le même objet.

1.2.2 Déterminer la dimension d'un attracteur

Je vais chercher ici à calculer la dimension d'un objet $\Lambda \subset \mathbb{R}^d$, qui sera typiquement un attracteur. Les deux notions de dimension que je veux introduire ici reposent sur la même intuition :

Quel espace occupe Λ dans $\mathbb{R}^d$?

Commençons avec un exemple simple. Plaçons nous en dimension 2, et Λ sera le carré unité, i.e. $\Lambda = [0, 1] \times [0, 1]$. Comme Λ est le produit de $[0, 1]$ par $[0, 1]$, il *occupe deux dimensions*. Le résultat attendu est donc que la dimension de Λ est égale à 2. Pour mesurer l'espace occupé par ce carré, nous pouvons le recouvrir par des carrés plus petits, de taille par exemple ε . Si nous notons s la dimension (supposée inconnue) de Λ , nous pouvons quand même écrire qu'il faut $1/\varepsilon^s$ petits carrés pour recouvrir Λ . Nous avons représenté ce procédé sur la Figure 1.7. Ici, $\varepsilon = 0.25$, et il faut $1/0.25^2 = 16$ carrés pour recouvrir Λ . En notant $N(\varepsilon)$ le nombre de carrés nécessaires pour recouvrir Λ , nous avons l'égalité

$$N(\varepsilon) = \frac{1}{\varepsilon^s}, \quad (1.2)$$

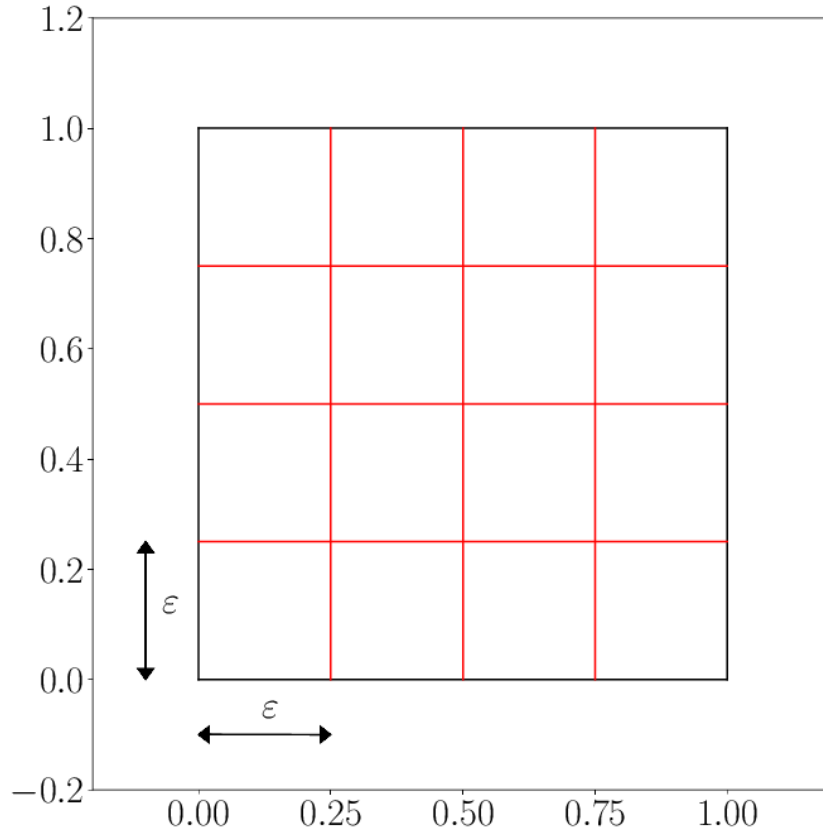


FIGURE 1.7 – Recouvrement du carré $[0, 1] \times [0, 1]$ par des petits carrés de taille $\varepsilon = 0.25$. Il faut $1/\varepsilon^2 = 16$ carrés pour le recouvrir.

et en renversant cette équation :

$$s = \frac{\log(N(\varepsilon))}{\log(1/\varepsilon)}.$$

Nous trouvons bien $s = 2$ pour tout ε . La valeur de ε ne joue pas de rôle, à cause de l'égalité de l'Equation (1.2). Sur des objets plus complexes cette égalité devient une relation de proportionnalité. La dimension de *box-counting*, aussi appelée *capacité* par YOUNG (1981), est définie par

$$s = \lim_{\varepsilon \rightarrow 0} \frac{\log(N(\varepsilon))}{\log(1/\varepsilon)}.$$

Cette définition, bien qu'intuitive, est impraticable dès que $d > 2$ (GREENSIDE et al., 1982). Le calcul est en fait impossible sur des données climatiques. En effet, la quantité de données nécessaires pour estimer la dimension d'une série temporelle est colossale. Pour obtenir la dimension avec une confiance à au moins 95%, il faut au minimum 42^{2d+1}

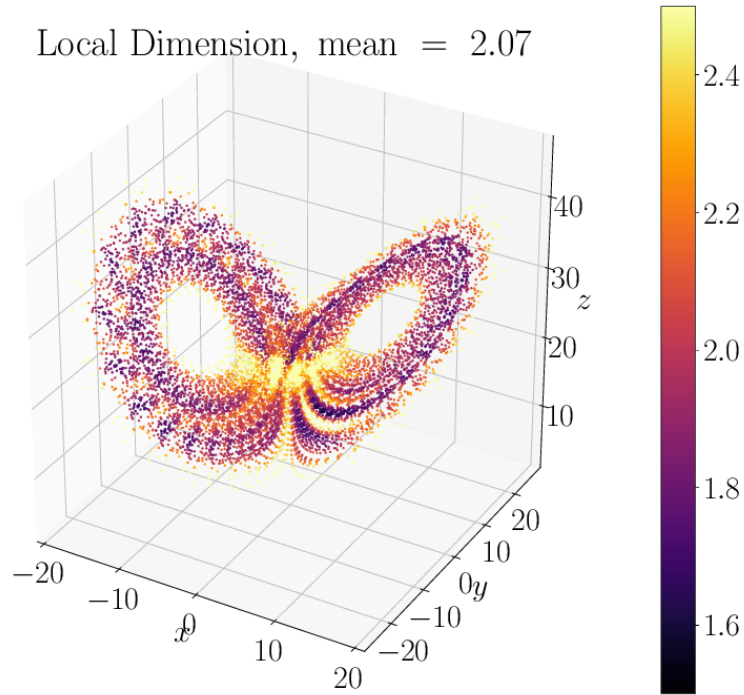


FIGURE 1.8 – Dimension locale de l’attracteur de LORENZ (1963)

points (SMITH, 1988), où d est la dimension de l’espace des phases. Ceci rend le calcul de dimension (par cette méthode) impraticable sur des données climatiques. SMITH (1988) proposa comme alternative le calcul de la *dimension locale*.

La dimension locale autour d’un point est définie comme la dimension de l’attracteur dans un petit voisinage de ce point. Une estimation de la dimension de l’attracteur est alors donnée par la moyenne des dimensions locales. Des travaux récents (FREITAS et al., 2010 ; FARANDA et al., 2011 ; FARANDA et al., 2013) permettent de calculer cette dimension, en reliant la théorie des valeurs extrêmes avec la théorie des systèmes dynamiques. Par exemple, j’ai représenté sur la Figure 1.8 les dimensions locales de l’attracteur de LORENZ (1963). Nous pouvons voir que la dimension est plus élevée au centre, dans la « boîte » de chaos, et diminue sur les ailes. La valeur moyenne, ici calculée avec 10000 points, est de 2.07, très proche de la valeur estimée par GRASSBERGER et PROCACCIA (1983), qui était de 2.05. En appliquant cette théorie, la dimension de la pression au niveau de la mer de l’atmosphère dans l’hémisphère nord est située entre 12 et 20 (FARANDA et al., 2017b ; FARANDA et al., 2017a).

1.2.3 Caractérisation du comportement local de la dynamique : les exposants de Lyapunov

Nous avons vu que les trajectoires des systèmes chaotiques divergent exponentiellement à partir de deux conditions initiales proches. Je voudrais quantifier l'évolution de la différence entre ces deux trajectoires. Afin de faciliter l'exposé, je vais considérer un système discret f^n . Partant d'une condition initiale x_0 , les points de l'orbite $\mathcal{O}(x_0)$ sont reliés par la relation :

$$x_{n+1} = f(x_n) = f^n(x_0).$$

Prenons maintenant une deuxième orbite $\mathcal{O}(y_0)$, qui commence proche de x_0 . La suite des différences ε_n entre x_n et y_n s'écrit $y_n = x_n + \varepsilon_n$. En linéarisant au voisinage de x_n , nous pouvons alors écrire

$$x_{n+1} + \varepsilon_{n+1} = f(x_n + \varepsilon_n) \sim f(x_n) + \varepsilon_n f'(x_n).$$

Ce qui entraîne l'équivalence suivante

$$\varepsilon_{n+1} \sim \varepsilon_n \times f'(x_n).$$

L'amplification de la différence entre x_n et y_n est donnée par le rapport de la différence à l'étape n et la différence au temps 0, autrement dit :

$$\left| \frac{\varepsilon_n}{\varepsilon_0} \right| = \left| \frac{\varepsilon_{n-1} f'(x_{n-1})}{\varepsilon_0} \right| = \left| \prod_{i=0}^{n-1} f'(x_i) \right|.$$

Comme l'orbite est expulsée exponentiellement vite le long de la direction instable, nous pouvons écrire que $\varepsilon_n \sim e^{\lambda n} \varepsilon_0$. En prenant le logarithme, nous obtenons

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \log |f'(x_i)|.$$

Ce coefficient λ est appelé *exposant de Lyapunov*. Si $\lambda < 0$, deux trajectoires se rapprochent exponentiellement vite. Si $\lambda > 0$, alors les deux trajectoires s'éloignent exponentiellement vite. Cette valeur permet donc de caractériser le comportement de la dynamique au voisinage d'un point. Pour un système de dimension $d > 1$, il faut choisir une direction $\mathbf{u} \in \mathbb{R}^d$ dans laquelle nous cherchons le comportement du système (en dimension 1, nous avons choisi $\mathbf{u} = 1$). La définition des *exposants de LYAPUNOV* (1992) $\lambda(\mathbf{x}, \mathbf{u})$ est alors donnée par

$$\lambda(\mathbf{x}, \mathbf{u}) = \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \|d_{\mathbf{x}}(f^n) \cdot \mathbf{u}\|. \quad (1.3)$$

Nous retrouvons bien la formule précédente grâce à la relation de récurrence suivante :

$$\log \|d_{\mathbf{x}}(f^n) \cdot \mathbf{u}\| = \log \|d_{f^{n-1}(\mathbf{x})}f \circ d_{\mathbf{x}}f^{n-1} \cdot \mathbf{u}\|.$$

Le théorème de OSELEDEC (1968), garantit que pour un $\mathbf{x}$ fixé, $\lambda(\mathbf{x}, \cdot)$ prend un nombre fini de valeurs $\lambda_1(\mathbf{x}) > \dots > \lambda_k(\mathbf{x})(\mathbf{x})$ avec $k(\mathbf{x}) \leq d$. Remarquons que comme l'Équation (1.3) est une limite sur l'orbite de $\mathbf{x}$, les exposants de Lyapunov sont constants le long de cette orbite.

Prenons l'exemple du LORENZ (1963). Dans ce cas, toutes les orbites convergent vers un attracteur étrange, le papillon. De plus, elles remplissent tout l'attracteur (c.f. la définition d'un attracteur étrange dans la Section 1.2.1). Donc les exposants de Lyapunov sont constants (ce sont les mêmes pour n'importe quel point de l'espace des phases). Ils sont au nombre de trois, 0.9056, 0 et -14.5723, i.e. une combinaison d'exposants positifs et négatifs. Le plus grand des exposants de Lyapunov donne une caractérisation du chaos, vu qu'il entraîne la séparation exponentielle de conditions initiales proches.

1.2.4 Statistiques d'un système dynamique : la mesure SRB

Les outils présentés dans les sections précédentes vont nous permettre de donner une définition des « statistiques d'un attracteur », déjà abordées avec la Figure 1.5 dans le cas du modèle de LORENZ (1963). Ces statistiques seront une *mesure de probabilité* supportées par l'attracteur (i.e. les probabilités sont nulles en dehors de l'attracteur). La mesure qui nous intéresse, et qui jouit de bonnes propriétés est la *mesure SRB* (SINAI, 1972 ; RUELLE, 1976 ; BOWEN et CHAZOTTES, 1975). Cette mesure s'estime en construisant *l'histogramme* de l'attracteur dans l'espace des phases. La procédure est la suivante :

1. Construire une longue orbite du système dynamique (il faut que l'orbite parcoure plusieurs fois l'attracteur pour le remplir),
2. Découper l'attracteur en petits cubes,
3. Compter le nombre de points dans chacun des petite cubes. Ces nombres, normalisés par la longueur de la trajectoire, sont une estimation de la valeur de la mesure SRB dans chacun des petits cubes.

J'ai représenté sur la Figure 1.9 une estimation de la mesure SRB du système de LORENZ (1963). Une orbite contenant 590000 points avec un pas de temps de 0.005 a été généré (les 10000 premiers ont été ignoré pour ne garder que les points sur l'attracteur). Ensuite, j'ai découpé $\mathbb{R}^3$ en cube de taille $0.5 \times 0.5 \times 0.5$. J'ai compté combien de points de l'orbite étaient dans chaque cube. Le centre de chaque cube est représenté par un point de couleur. La couleur indique le nombre de points comptés, normalisé par la taille de

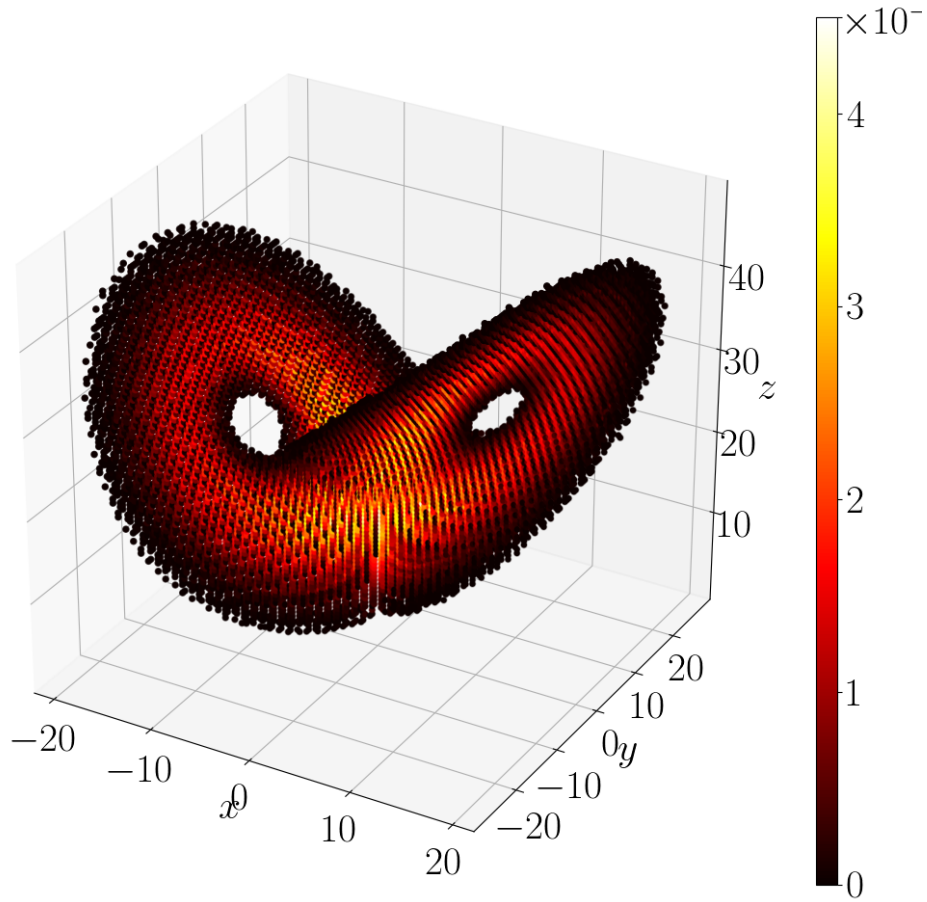


FIGURE 1.9 – Estimation de la mesure SRB du système de LORENZ (1963) à partir d’une orbite contenant 590000 points. 600000 points sont générés par pas de temps de 0.005 à l’aide de l’Équation (1.1). Les 10000 premiers sont ignorés pour ne conserver que des points sur l’attracteur. Chaque points de couleur est le centre d’un petit cube de taille $0.5 \times 0.5 \times 0.5$. Il y a 14575 cubes sur cette figure.

l’orbite. Nous pouvons voir que la masse est plus forte au centre de l’attracteur, et sur des cercles autour des ailes. Ceci correspond aux zones de l’espace des phases où l’orbite revient fréquemment.

Cette mesure de probabilité vérifie deux propriétés importantes, qui caractérisent la dynamique sous jacente : elle est *invariante* et *ergodique*. Je vais expliciter ces deux notions.

Une mesure μ est dite *invariante* pour un système dynamique f si pour tout sous ensemble A de l’espace de phase, $\mu(A) = \mu(f^{-1}(A))$, (i.e. les probabilités sont les mêmes). Cette propriété est naturelle, car la probabilité de « visiter » A en partant de $f^{-1}(A)$ est

égale à 1. L'invariance exprime donc une *compatibilité* de la mesure μ avec le système dynamique f .

Une mesure μ est dite *ergodique* pour un système dynamique f si pour tout sous ensemble A invariant (i.e. $f^{-1}(A) = A$), la probabilité $\mu(A)$ est soit nulle, soit égale à 1. Cette propriété garantit qu'il n'y a qu'un seul attracteur. En effet, si nous avons deux attracteurs distincts, ils seraient d'une part invariant (par définition). D'autre part, μ serait non nulle sur chacun d'eux, et donc forcément inférieure à 1. Ceci contredit la définition de l'ergodicité. Dans le cas du système de (LORENZ, 1963), il n'y a que trois ensembles invariants : $\mathbb{R}^3$, l'attracteur, et l'ensemble vide, qui ont respectivement comme probabilité 1, 1 et 0. Ce système est donc ergodique. Un corollaire important de cette propriété est une loi des grands nombres. Si nous prenons une fonction $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$ intégrable par rapport à μ (i.e. $\psi \in L^1(\mu)$), alors la moyenne *temporelle* de ψ le long d'une orbite et la moyenne *spatiale* coïncide, i.e.

$$\underbrace{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \psi \circ f^i(\mathbf{x})}_{\text{Moyenne temporelle}} = \underbrace{\int_{\mathbb{R}^d} \psi \cdot d\mu}_{\text{Moyenne spatiale}}.$$

Ce résultat est nommé le Théorème de BIRKHOFF (1927). Dans toute la suite, je supposerai implicitement que la mesure de probabilité sur laquelle je travaille est une mesure SRB. Ceci est justifié par la conjecture de PALIS (2005) et PALIS (2008). Cette conjecture affirme que l'existence d'une mesure SRB est une propriété générique dans les systèmes dynamiques (i.e. à une perturbation près aussi petite que désirée, il existe une mesure SRB). Le modèle de LORENZ (1963) admet par exemple une unique mesure SRB (BUNIMOVICH, 1983).

1.2.5 Résumé

Dans cette section j'ai vulgarisé les concepts mathématiques qui décrivent la théorie des systèmes dynamiques. Le point important qu'il faut retenir ici est la définition de la mesure SRB, qui a nécessité une quantité importante de concepts mathématiques pour être explicitée. Ses propriétés principales sont :

1. Son invariance par la dynamique,
2. Son adéquation à la loi des grands nombres,
3. Sa facilité d'estimation à l'aide d'un histogramme.

Avec ces propriétés, si les statistiques du climat sont une mesure SRB, alors effectivement quantifier cette mesure permet de d'étudier le système climatique. Dans toute la suite, je supposerai toujours implicitement que la mesure étudiée est une mesure SRB. Je vais

maintenant revenir sur le climat, et parler des travaux autour de l’existence d’un attracteur climatique, ainsi que du problème du forçage anthropogénique.

1.3 Attracteur climatique

1.3.1 Existence d’un attracteur étrange climatique

LIONS et al. (1992) et LIONS et al. (1993) (voir aussi TEMAM (2012)) ont donné ce qui est le plus proche d’une démonstration de l’existence d’un attracteur climatique. Ils ont travaillé directement sur les équations de Navier-Stokes, ou sur les équations primitives de l’atmosphère de RICHARDSON (1922) qui ont été reformulées par LIONS et al. (1992). Ils ont démontré l’existence d’un attracteur, et ont donné une borne (non explicite) à sa dimension, mais une borne finie. Leurs travaux, combinés à ceux de LORENZ (1969), GHIL et al. (2008) et VANNITSEM (2014) justifient l’existence d’un tel attracteur.

1.3.2 Déformation de l’attracteur climatique

La définition du climat laisse ouverte la possibilité à une altération dans le temps du système climatique. Par exemple, l’attracteur des températures du climat durant l’hiver est différent de celui de l’été. Au 21ème siècle, le forçage anthropogénique à l’oeuvre depuis la révolution industrielle en 1850, altère petit à petit le climat. Détecter et attribuer les événements dues à cette altération est l’un des enjeux actuels des sciences climatiques (BINDOFF et al., 2013). Prenons par exemple deux simulations du modèle de l’IPSL (DUFRESNE et al., 2013) des températures en Ile-de-France. Une simulation du scénario RCP85 (MOSS et al., 2010), et une simulation d’un climat figé au stade préindustriel (déjà présenté en introduction). Nous avons représenté sur la Figure 1.10 leurs distributions sur des périodes de 20 ans. La Figure 1.10a correspond à une simulation pré-industrielle, tandis que la Figure 1.10b correspond à une simulation du scénario RCP85. Du point de vue des systèmes dynamiques, ces distributions s’interprètent comme des projections de la mesure SRB. Nous voyons que les distributions sont constantes pour la simulation pré-industrielle, montrant la stationnarité du climat. Nous sommes dans un cas similaire au modèle de LORENZ (1963). Pour la simulation du scénario RCP85, les distributions se décalent vers des températures de plus en plus chaudes, indiquant une modification de la mesure SRB. Autrement dit, si nous reprenions la définition des système dynamiques donnée dans la section 1.2.1, il faudrait l’étendre en ajoutant la variable temps :

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t).$$

Les systèmes de ce type sont dit *non-autonomes*. Je présenterai dans la prochaine

section comment traiter ces systèmes. Je reviendrai sur l'exemple des températures en Ile-de-France dans le chapitre suivant.

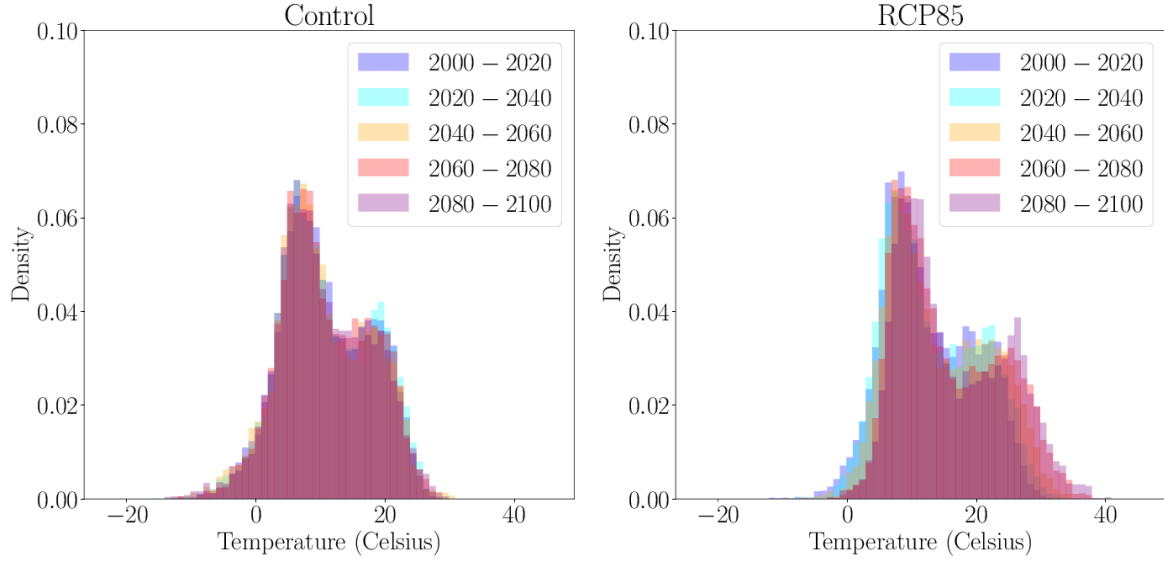


FIGURE 1.10 – Distribution des températures en Ile-de-France (France) pour les simulations CMIP5, modèle de l'IPSL. À gauche, une simulation de contrôle (pré-industrielle). À droite, une simulation du scénario RCP85.

1.4 Dynamique non autonome

1.4.1 Système de LORENZ (1984)

Ayant introduit le problème de la dépendance temporelle, nous devons utiliser des équations différentielles *non autonomes*, i.e. dépendantes du temps. Je vais les illustrer avec un nouveau modèle, l'attracteur de LORENZ (1984), décrivant la circulation de HADLEY (1735). Pour $\mathbf{x} = (x, y, z)$, ses équations sont données par

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t) = \begin{pmatrix} -y^2 - z^2 - ax + aF(t) \\ xy - bxz - y + G \\ xz + bxy - z \end{pmatrix}. \quad (1.4)$$

Suivant LORENZ (1984), les valeurs de a , b et G sont fixées à 0.25, 4 et 1. Dans un premier temps, le comportement de ce système a été étudié pour des valeurs constantes de F . MASOLLER et al. (1992) ont montré que si $F \in [1, 8]$, cet intervalle se décompose en sous intervalles où le système est soit stable soit chaotique. En particulier pour $F = 8$, ils ont calculé la valeur des exposants de Lyapunov : $\lambda_1 = 0.23$, $\lambda_2 = 0$ et $\lambda_3 = -0.53$ ainsi que

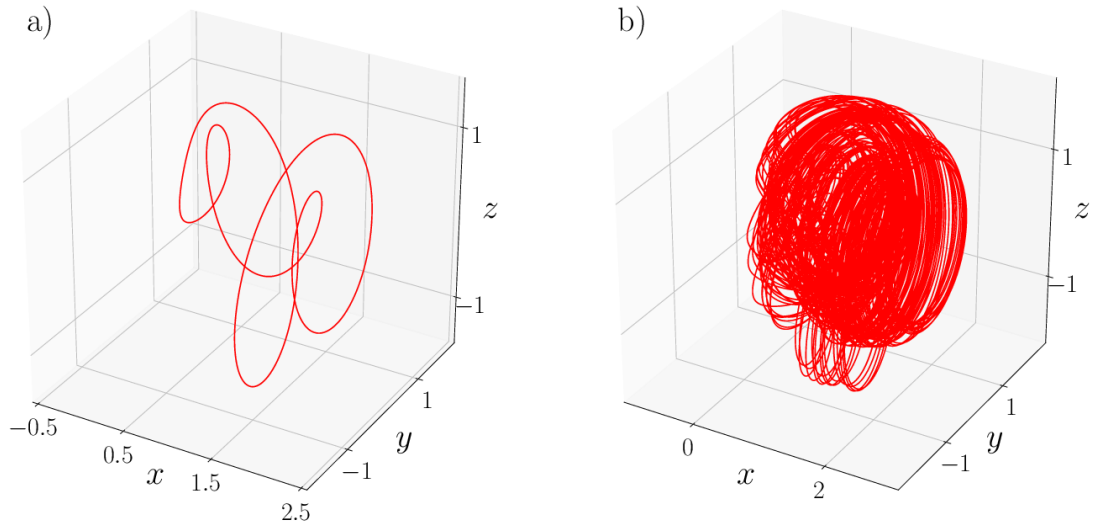


FIGURE 1.11 – Attracteurs de LORENZ (1984) pour deux valeurs de F . **a)** $F = 6$ (été permanent, système non chaotique bloqué sur un cycle limite). **b)** $F = 8$ (hiver permanent, système chaotique)

la dimension de corrélation, égale à 2.23, démontrant le comportement chaotique de ce système. Deux autres valeurs critiques de chaos ont été identifiées par MASOLLER et al. (1995), $F = 4.56$ et $F = 5.198$. Enfin, NICOLIS et al. (1995) ont utilisé ce système comme exemple de l'amplification de l'erreur. Ils ont mis en évidence qu'elle est due à deux mécanismes : la variabilité des exposants de Lyapunov locaux autour de valeur moyenne, et la non orthogonalité des vecteurs propres du linéarisé du système.

Nous conserverons les valeurs a , b et G à 0.25, 4 et 1. Pour F , deux valeurs nous intéresseront particulièrement, $F \equiv 6$ et $F \equiv 8$. Le premier cas correspond à un été permanent, et le deuxième à un hiver permanent. Nous avons représenté ces deux attracteurs sur la Figure 1.11. La Figure 1.11a correspond à $F = 6$, et n'est pas chaotique, le système s'enroule autour d'un cycle limite. Sur la Figure 1.11b nous sommes sur une des valeurs critiques ($F = 8$) décrites par MASOLLER et al. (1992), entraînant un comportement chaotique du système.

LORENZ (1990) a ajouté un équivalent du cycle saisonnier. La fonction $F(t)$ devient une sinusoïdale, d'équation

$$F(t) = F_0 + A \sin(\omega t). \quad (1.5)$$

Pour ces paramètres, une unité de temps revient à 5 jours, ou une année est équiva-

lente à 73 unités de temps. Ceci fixe $\omega = \frac{2\pi}{73}$. Les valeurs F_0 et A sont fixées à 9.5 et 2, ce qui donne des oscillations de F entre 7.5 (proche de l'été) et 11.5 (hivernal et chaotique). Des équivalents de saisons peuvent donc être définis. Suivant la Fig. 1.12a, l'équinoxe de l'automne est définie à $t = 0 \bmod 73$, le solstice d'hiver à $t = 73/4 \bmod 73$, puis l'équinoxe de printemps à $t = 73/2 \bmod 73$, et enfin $t = 3 \times 73/4 \bmod 73$ sera le solstice d'été. Ce type de modèle est particulièrement utile pour du « proof of concept », car nous obtenons un modèle chaotique de basse dimension simulant un cycle saisonnier mais avec de la variabilité interne. Une étude à long terme de la variabilité a été réalisée par PIELKE et ZENG (1994) (1100 ans) en comparant le cas $F = 8$ (stationnaire) et F le cycle saisonnier.

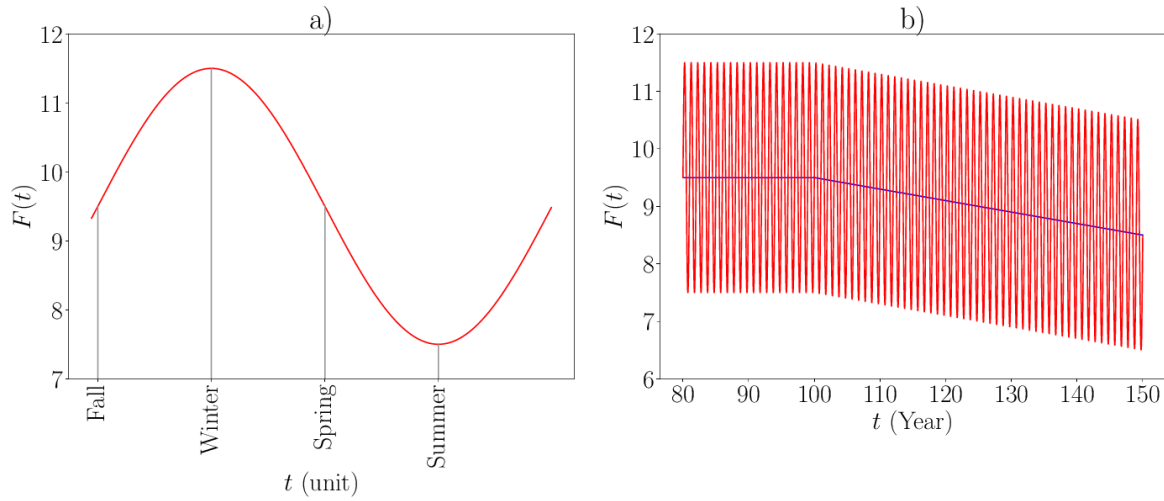


FIGURE 1.12 – **a)** Cycle saisonnier du LORENZ (1984). **b)** Rouge : Cycle saisonnier et forçage linéaire après 100 ans.

Affinant ce modèle, DRÓTOS et al. (2015) ont modifié F_0 en intégrant un forçage linéaire. Pendant les 100 premières années, F_0 est supposé constante et égale à 9.5, puis décroît. Notons $T_{cc} = 100 \times 73$ (*cc* pour *climate change*) l'instant où démarre le forçage linéaire. L'équation est donnée par :

$$F_0(t) = 9.5 + \begin{cases} 0, & t \leq T_{cc}, \\ -2 \frac{t - T_{cc}}{T_{cc}}, & t > T_{cc}. \end{cases} \quad (1.6)$$

Comme nous pouvons le voir sur la Figure 1.12b, les valeurs moyennes de $F(t)$ varient donc entre 7.5 et 11.5 avant 100 ans. Ensuite, les valeurs moyennes du cycle décroissent jusqu'à varier entre 4.5 et 8.5. Ceci revient à forcer l'attracteur vers un été permanent. Ce forcing permet de mettre à l'épreuve des méthodes de détection / quantification de la modification du climat sur un système simple et contrôlé.

Malheureusement les techniques d'études d'attracteurs présentées dans la section précédente ne s'appliquent pas à ce type de système. Pour deux conditions initiales différentes les attracteurs générés sont différents les uns des autres. En effet, l'ajout du temps rend l'attracteur généré par une orbite dépendant de la condition initiale. Ceci nous amène aux équations de transport.

1.4.2 Vision Lagrangienne et Eulerienne : des équations différentielles aux équations de transports

Les équations différentielles sont une manière de voir la dynamique, nommée la *vision Lagrangienne*, où nous suivons une trajectoire. Elles sont en réalité équivalentes aux équations de transport, qui représentent la *vision Eulerienne*, où nous analysons l'état de toutes les trajectoires à un temps donné. Rappelons ce qu'est une équation de transport. Étant donné un champ de vecteurs (dépendant du temps) $\mathbf{v} : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$, une *équation de transport* est une équation aux dérivées partielles de la forme

$$\partial_t u + \langle \mathbf{v}, \nabla u \rangle = 0. \quad (1.7)$$

Les notations ∂_t , ∇ et $\langle, \rangle$ désignent respectivement la dérivée partielle en temps, le gradient en espace et le produit scalaire. L'élément u est une fonction de $\mathbb{R}^d \times \mathbb{R}$ dans $\mathbb{R}$, et est l'inconnue. Résoudre une équation de transport est équivalent à résoudre l'équation différentielle donnée par $\mathbf{v}$. Premièrement, les équations différentielles non autonomes peuvent être ramenées à un système autonome en introduisant une dimension supplémentaire définie par une nouvelle variable $\tau \in \mathbb{R}$:

$$\begin{cases} \frac{d\mathbf{x}}{d\tau} = \mathbf{v}(\mathbf{x}, \tau), \\ \frac{dt}{d\tau} = 1. \end{cases} \quad (1.8)$$

Notons que cette astuce ne résout pas le problème du LORENZ (1984), car les orbites ainsi définies ne sont pas bornées pour la variable t . Maintenant, si u est solution de l'Équation (1.7), nous pouvons calculer sa dérivée le long d'une solution $\mathbf{x}(t)$ définie par l'Équation (1.8). Ceci revient à étudier les caractéristiques du système (EVANS, 2010, p. 100-120) :

$$\frac{d}{d\tau} u(\mathbf{x}, \tau) = \partial_\tau u(\mathbf{x}, \tau) + \left\langle \frac{d}{d\tau} \mathbf{x}, \nabla u \right\rangle = \partial_\tau u + \langle \mathbf{v}, \nabla u \rangle = 0.$$

Nous en concluons donc que u est constante le long des orbites définies par $\mathbf{v}$. Pour trouver $u(\mathbf{x}, t)$, il suffit de « remonter » le temps jusqu'à une valeur connue, i.e. une

condition initiale de l'Équation (1.7). Par exemple, si $u_0 = u(\cdot, 0)$ est une condition initiale de l'Équation (1.7) et φ^t est le flot de l'Équation (1.8), alors

$$u(\mathbf{x}, t) = u_0(\varphi^{-t}(\mathbf{x})).$$

Le point central ici est que la fonction u , à un temps t_0 fixé, définit un objet unique. Les orbites dépendent de leurs conditions initiales, mais pas *l'ensemble des orbites à un temps donné*. Dans le cas où $\mathbf{v}$ génère un attracteur chaotique, ceci implique que pour une collection de N conditions initiales $\mathbf{x}_0^1, \dots, \mathbf{x}_0^N$, les solutions au temps t_0 donné par $\varphi^{t_0}(\mathbf{x}_0^1), \dots, \varphi^{t_0}(\mathbf{x}_0^N)$ génèrent un objet qui ne dépend plus des $\mathbf{x}_0^i$. Cet objet se nomme un *snapshot attractor*, et est l'objet de la section suivante.

1.4.3 L'équivalent des orbites en dynamique non autonome : les snapshots attractors

Les snapshot attractors sont l'outil qui va remplacer les orbites. Ils ont été introduits par ROMEIRAS et al. (1990) pour les systèmes dynamiques ergodiques. Examinons dans un premier temps le cas du LORENZ (1963). Prenons $N = 1000$ conditions initiales aléatoirement dans un cube de $\mathbb{R}^3$ défini par $[-20, 20] \times [-20, 20] \times [0, 40]$. Ceci est représenté sur la Figure 1.13a. Ensuite nous itérons toutes les conditions initiales sur la période de temps $[0, 0.25]$ par pas de 0.005. Ceci revient à appliquer le flot φ^t du système de LORENZ (1963) à chaque condition initiale.

Le résultat est montré pour les temps $t = 0.05$ (Figure 1.13b), $t = 0.1$ (Figure 1.13c) et $t = 0.25$ (Figure 1.13d). Nous pouvons voir que les trajectoires finissent par dessiner l'attracteur de LORENZ (1963), et de manière extrêmement rapide (rappelons que la divergence vaut environ -13). Nous avons de plus représenté sur la Figure 1.14 deux snapshots aux mêmes temps $t = 0.25$ avec deux jeux de 1000 conditions initiales, en bleu et rouge. L'attracteur ainsi généré ne dépend plus du jeu de conditions initiales. Les snapshots attracteurs se construisent ainsi de la manière suivante :

1. Choisir aléatoirement N conditions initiales,
2. Itérer toutes ces conditions initiales jusqu'à un temps t ,
3. L'ensemble des N valeurs au temps t est un snapshot attracteur.

Dans le cas du système de LORENZ (1963), après un temps court, les snapshots attracteurs sont tous identiques. Examinons les snapshots attracteurs du système de LORENZ (1984), qui dépendent du temps. Tirons 1000 conditions initiales uniformément dans le cube $[-1, 3] \times [-3, 3] \times [-3, 3]$, et itérons les par l'Équation (1.4). Le forcing $F(t)$ est la somme du cycle saisonnier (Équation. (1.5)) et du changement climatique (Équa-

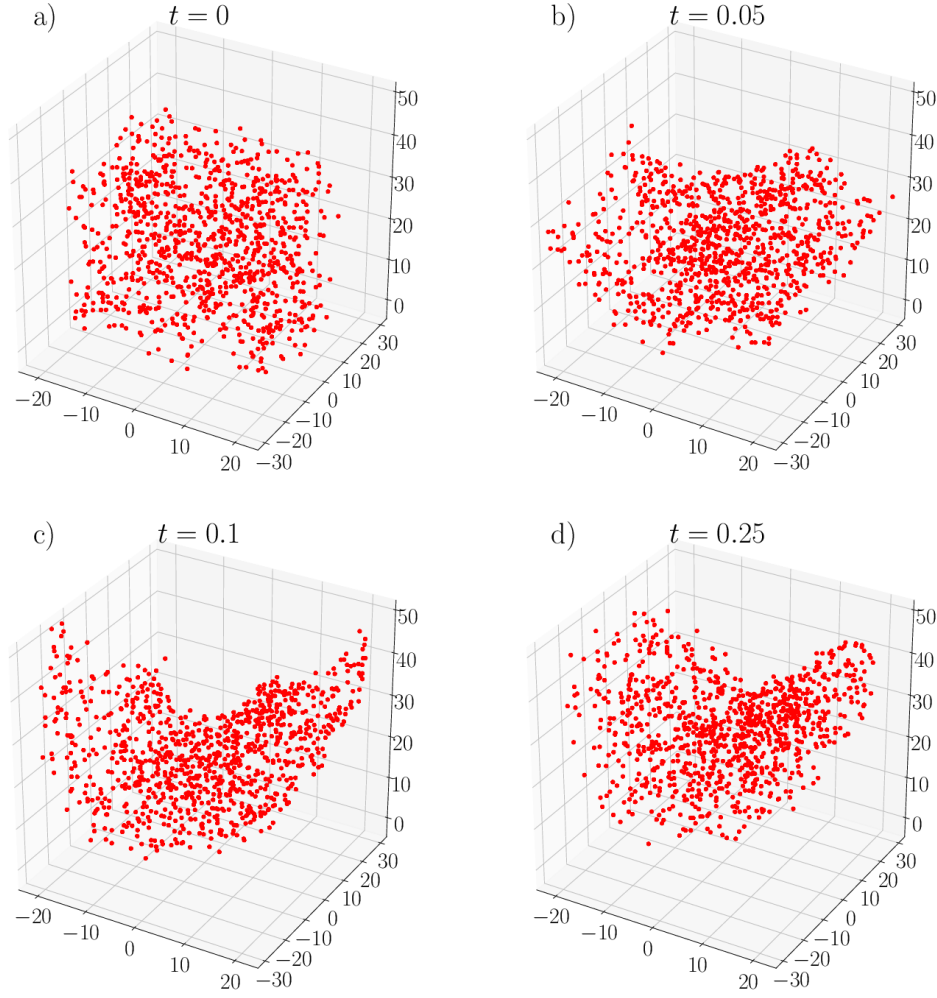


FIGURE 1.13 – Snapshot de l’attracteur de LORENZ (1963) construit avec 1000 conditions initiales aléatoires. La figure a) est l’état de toutes les conditions initiales au temps 0. La figure b) est l’état de toutes les conditions initiales au temps 0.05. La figure c) est l’état de toutes les conditions initiales au temps 0.1. La figure d) est l’état de toutes les conditions initiales au temps 0.25.

tion (1.6)). L’intégration est réalisée durant 200 ans (soit 200×73 unité de temps) par pas de 0.005 unité de temps, i.e. 0.025 jours.

Nous avons représenté sur la Figure 1.15 les snapshots des 4 saisons durant la première année, la dixième année et l’année 180. Nous pouvons voir que les attracteurs n’ont que très peu été modifiés entre l’année 0 et l’année 10 pour chaque saison. La séquence de snapshots attracteurs est donc bien périodique, et les attracteurs à un instant t et $t+k \times 73$ sont semblables. Pour l’année 180, les attracteurs ont été fortement altérés par rapport à leurs homologues d’avant $T_{cc} = 100 \times 73$. Visuellement, le nouvel hiver semble proche

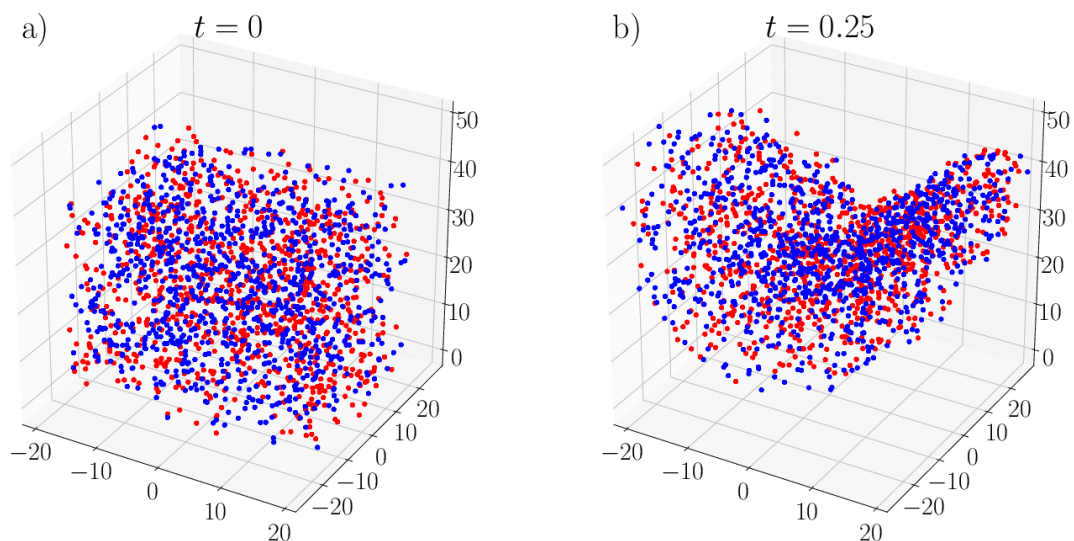


FIGURE 1.14 – Deux snapshot de l’attracteur de LORENZ (1963) construits avec deux jeux de 1000 conditions initiales en bleu et rouge. La figure a) est l’état de toutes les conditions initiales au temps 0. La figure b) est l’état de toutes les conditions initiales au temps 0.25.

des anciens automnes, ce qui est compatible avec un glissement des attracteurs vers des valeurs plus « chaudes ». Pour les autres saisons, l’interprétation visuelle est difficile.

À l’inverse de la dynamique autonome, où nous étudions des orbites qui génèrent un attracteur, la dynamique non autonome exhibe des *séquences d’attracteur*. Ces snapshots attracteurs ne sont que la solution de l’équation de transport associée à l’équation différentielle. C’est leur étude qui permet d’étudier le comportement d’un système. Dans l’objectif d’utiliser des attracteurs chaotiques de basses dimensions pour établir des propriétés ou valider des nouvelles méthodologies, les snapshots attracteurs sont des outils plus réalistes que les orbites des systèmes autonomes.

En dynamique autonome, les statistiques de l’attracteur ont été données par la mesure SRB, invariante par la dynamique. Nous allons maintenant nous intéresser à son équivalent en dynamique non autonome. Remarquons déjà que chaque snapshot attracteur étant unique à un pas de temps donné, il définit une mesure elle-même unique. Ceci donne une séquence μ_t de mesures associées à *chaque* snapshot attracteur. De manière générale, nous pouvons nous intéresser à la version « mesure » de l’Équation (1.7). Une équation de transport pour les mesures est un système de la forme

$$\begin{cases} \partial_t \mu_t + \langle \nabla, \mathbf{v} \mu_t \rangle = 0, \\ \mu|_{t=0} = \mu_0. \end{cases}$$

Il a été démontré (voir par exemple VILLANI (2003, Chap. 8)) que ce système admet une unique solution μ_t , et surtout que pour tout sous ensemble A de l'espace des phases, $\mu_t(A) = \mu_0(\varphi^{-t}(A))$. Autrement dit, c'est la mesure μ_0 , transporté par le flot φ^t . Cette séquence de mesure est l'équivalent de la mesure SRB (CHEKROUN et al., 2011).

1.5 Résumé et mise en place du problème

Le climat est défini comme les statistiques sur une échelle de temps donnée. Il a été mis en évidence dans la littérature l'existence d'un attracteur chaotique généré par le climat, au travers de la théorie des système dynamiques. Cette théorie permet de plus de donner un sens précis à ces statistiques, c'est la mesure SRB, qui contient toute l'information de la dynamique du climat. Nous savons de plus que le climat (et donc la mesure SRB) se modifie au cours du temps. Le problème qui va nous occuper est donc :

Comment quantifier la déformation d'un attracteur non autonome ?

Mesurer cette déformation soulève deux problèmes : d'une part nous devons inférer cette mesure (ou une de ses projections), et d'autre part nous devons quantifier la modification de cette mesure. Le premier point se heurte à plusieurs obstacles. La quantité de données dont nous disposons est finie, que ce soit les observations ou les modèles, et peuvent donc être en quantité insuffisante. De plus, ces données peuvent être entachées de bruit (pour les observations), ou biaisées (pour les modèles). Nous avons donc besoin d'un outil capable de résoudre ou contourner ces difficultés. Ce sera l'objet du chapitre suivant.

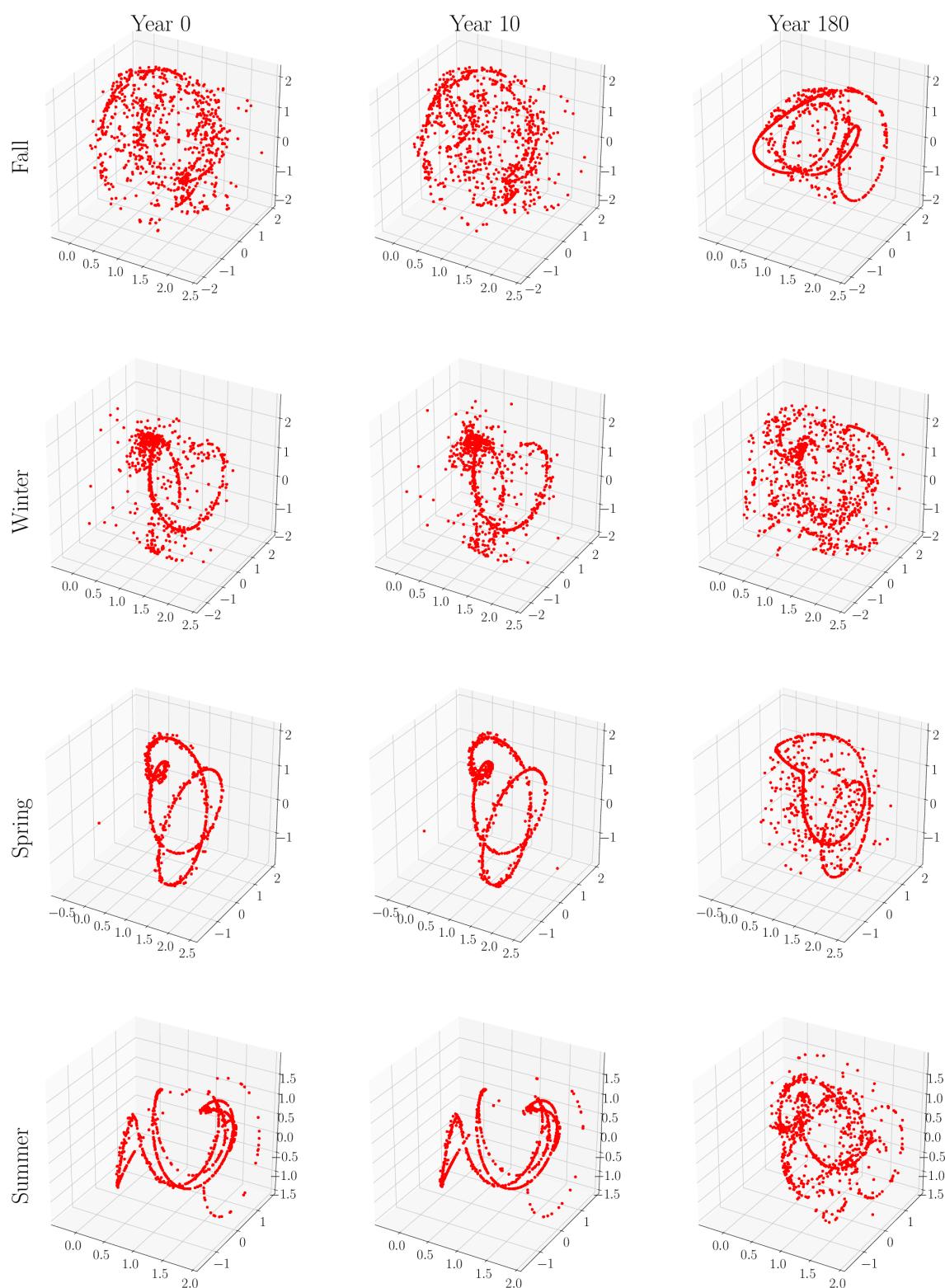


FIGURE 1.15 – Snapshots de l'attracteur de LORENZ (1984) généré à partir de 1000 conditions initiales tirées uniformément dans le cube $[-1, 3] \times [-3, 3] \times [-3, 3]$. En colonne : état des snapshots pendant le premier cycle saisonnier (année 0), puis le dixième cycle (année 10) et le cycle 180. Ce dernier est donc après que le forcing ait commencé à agir à l'année 0. En colonne : snapshots pendant l'automne, puis l'hiver, puis le printemps et enfin l'été.

Chapitre 2

Transport Optimal

Imaginons un système dynamique dont la mesure SRB, notée μ_t , suit une loi normale réduite. Elle est forcée de telle manière que au temps t , la moyenne est égale à $2t$. En notant ρ_t la densité de μ_t au temps t , μ_t s'écrit :

$$\mu_t(A) := \int_A \rho_t \cdot dx = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{1}{2}(x-2t)^2} \cdot dx, \quad A \subset \mathbb{R}.$$

Les densités de ρ_t ont été représentées aux temps 0, 2, 4, 6 et 8 sur la Figure 2.1a. Pour comparer les mesures μ_t et μ_0 , prenons la distance Euclidienne, donnée par la formule :

$$d_2(\mu_0, \mu_t) := \sqrt{\int_{\mathbb{R}} |\rho_t - \rho_0|^2 \cdot dx}.$$

La fonction $x \mapsto e^{-x^2/2}$ est quasiment nulle si $|x| > 4$. Nous pouvons donc écrire si $t > 4$:

$$d_2(\mu_0, \mu_t) \simeq \sqrt{2 \int_{\mathbb{R}} |\rho_0|^2 \cdot dx} = \text{constante}$$

J'ai représenté cette distance sur la Figure 2.1b. Entre $t = 0$ et $t = 2$, la distance Euclidienne augmente, puis se stabilise. Comme μ_t et μ_0 ne partagent que des points de masses nulles pour t assez grand, elle restera identique. Cette distance ne tient donc pas compte de la dynamique intrinsèque, qui applique une translation à μ_0 . La différence entre les moyennes en tiendrait compte, mais des variations des moments d'ordres supérieurs seraient invisibles. Notre but est d'avoir un outil qui tient compte de la dynamique qui agit entre deux distributions. Le transport optimal, que nous allons présenter dans ce chapitre, fournit une réponse à ce problème. La combinaison des outils de ce chapitre et du chapitre précédent a abouti à un papier publié (ROBIN et al., 2017).

Le transport optimal n'a été que très rarement employé jusqu'ici en climatologie. Il a été suggéré par GHIL (2015) pour caractériser la variabilité climatique. CULLEN et MAROOFI (2003) l'ont exploité pour résoudre les équations quasi-géostrophiques compressibles, tandis que FARCHI et al. (2016) l'ont appliqué à la comparaison de champs de

polluants. Le transport optimal a par contre été énormément étudié d'un point de vue théorique, ou utilisé pour des équations aux dérivées partielles en mécanique des fluides. Je renvoie aux deux livres de VILLANI (2003) et VILLANI (2008) qui décrivent l'état de l'art sur ce sujet (le premier livre étant plus digeste que le second). Mais de manière générale, sa complexité de calcul et le manque de librairies informatiques pouvant le résoudre en a limité l'adoption. Récemment, FLAMARY et COURTY (2017) ont développé une librairie python rendant les algorithmes de résolution facilement accessibles. Cette librairie implémente un solveur complet basé sur les algorithmes de simplex (BONNEEL et al., 2011), et un solveur par régularisation (CUTURI, 2013; SCHMITZER, 2016). Le transport optimal a aussi servi à la reconnaissance de forme et d'images (RUBNER et al., 2000), l'analyse de séries temporelles (MUSKULUS et VERDUYN-LUNEL, 2011) et des algorithmes de classification type K -means sont en développement (CUTURI et DOUCET, 2014). Je vais ici l'utiliser comme une mesure naturelle de la déformation de la dynamique.

Ce chapitre est organisé de la manière suivante. Dans la section 2.1, le transport optimal est introduit d'un point de vue historique, par le problème de Monge. Dans la section 2.2, les outils du transport optimal pour les mesures empiriques sont présentés. Dans la section 2.3, plusieurs façons de se représenter le plan de transport sont explicités, afin de se construire une intuition du mode d'action du transport optimal. Enfin, dans la section 2.5, un exemple d'utilisation est donné.

2.1 Introduction au transport optimal : le problème de Monge

Le problème du transport optimal a été d'abord étudié par MONGE (1781) dans un mémoire de l'académie royale de sciences. Sa formulation du problème était donnée dans son introduction, que je reproduis ici.

Lorsqu'on doit transporter des terres d'un lieu dans un autre, on a coutume de donner le nom de Déblai au volume des terres que l'on doit transporter, & le nom de Remblai à l'espace qu'elles doivent occuper après le transport.

Le prix du transport d'une molécule étant, toutes choses d'ailleurs égales, proportionnel à son poids & à l'espace qu'on lui fait parcourir, & par conséquent le prix du transport total devant être proportionnel à la somme des produits des molécules multipliés chacune

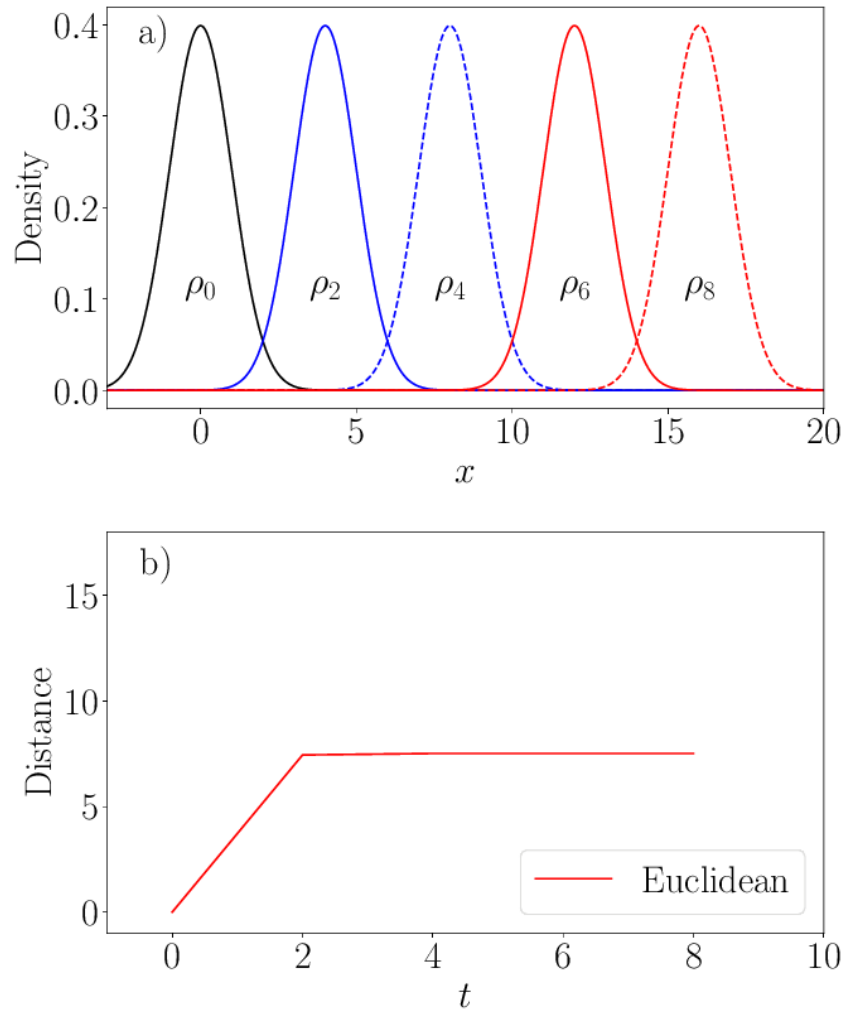


FIGURE 2.1 – Les mesures μ_t sont des lois normales centrées en $2t$, d'écart type égal à 1. **a)** Densités de μ_t , notées ρ_t , aux temps 0, 2, 4, 6 et 8. **b)** Distance Euclidienne entre μ_0 et μ_t .

par l'espace parcouru, il s'ensuit que le déblai & le remblai étant donnés de figure & de position, il n'est pas indifférent que telle molécule du déblai soit transportée dans tel ou tel autre endroit du remblai, mais qu'il y a une certaine distribution à faire des molécules du premier dans le second, d'après laquelle la somme de ces produits fera la moindre possible, & le prix du transport total fera un minimum.

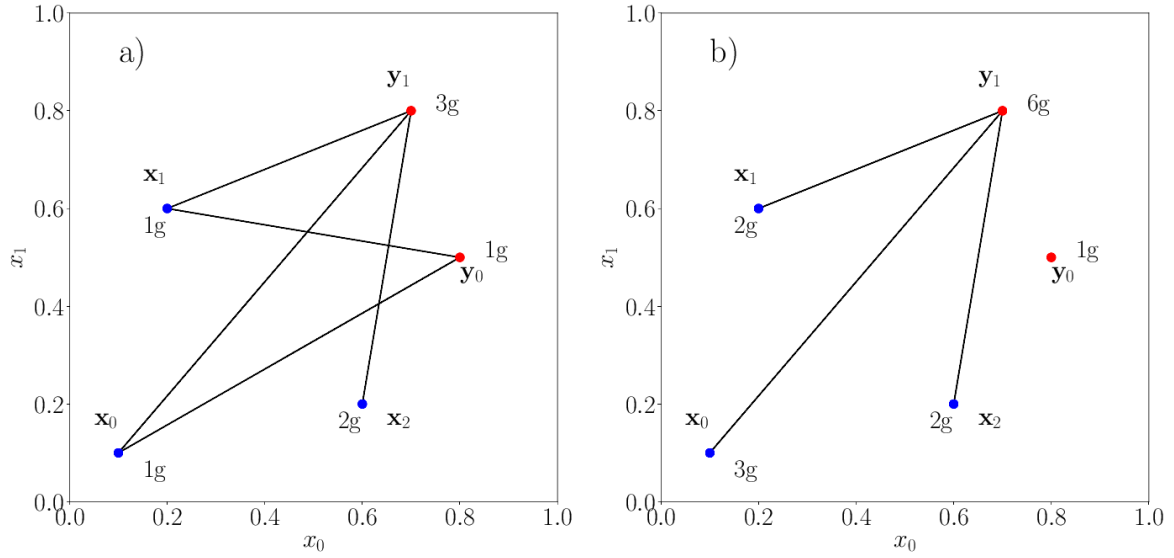


FIGURE 2.2 – Deux exemples du problème de Monge. Les points bleus $\mathbf{x}_i$ sont les déblais, les rouges $\mathbf{y}_j$ les remblais. Les segments noirs sont des déplacements possibles entre $\mathbf{x}_i$ et $\mathbf{y}_j$.

Illustrons son problème par la Figure 2.2a. Les déblais sont des cailloux, représentés par les marqueurs bleus à leur position d'origine, avec leur poids exprimé en gramme. Les remblais sont représentés en rouge, à la position où nous souhaitons les transporter, avec le poids que nous souhaitons obtenir à chaque position finale. Les droites noires sont les trajectoires possibles du transport de chaque pierre. À chaque droite noire est associée une distance, dont la pondération par le poids de la pierre à transporter donne le coût de transport le long de cette trajectoire. Notons μ_i les masses de départ aux positions $\mathbf{x}_i$, ν_j les masses d'arrivées aux positions $\mathbf{y}_j$, et γ_{ij} la masse transportée de $\mathbf{x}_i$ à $\mathbf{y}_j$. La masse γ_{ij} vérifie l'égalité suivante :

$$\nu_j = \sum_i \gamma_{ij}. \quad (2.1)$$

Du point de vue de Monge, chaque ν_j peut recevoir de la masse de plusieurs μ_i , mais

les μ_i ne peuvent pas se scinder. Cela revient à dire que les ν_j sont une fonction des μ_i . Le coût C s'écrit :

$$C(\gamma) = \sum_{ij} \gamma_{ij} \|\mathbf{x}_i - \mathbf{y}_j\|^2. \quad (2.2)$$

Le *problème de Monge* consiste à minimiser la fonction C selon γ sous la contrainte définie par l'Éq. (2.1). Résolvons le problème de la Figure 2.2a. Le point $\mathbf{x}_2$ ne peut être envoyé que sur $\mathbf{y}_1$. Nous avons alors deux possibilités : soit $\mathbf{x}_0$ est envoyé sur $\mathbf{y}_0$ (et $\mathbf{x}_1$ sur $\mathbf{y}_1$), soit l'inverse. Dans le premier cas, le coût est égal à 1.68, et dans le deuxième à 1.96. Le premier plan est donc le meilleur, et est en l'occurrence l'optimal.

Ce problème est la première occurrence du transport optimal, mais est limité : les déblais μ_i ne peuvent pas se séparer en sous masses qui approvisionneraient chaque remblai. Ce problème peut donc se révéler insoluble, comme il est représenté sur la Figure 2.2b. Dans ce cas, la masse de 1g au point $\mathbf{y}_0$ ne peut recevoir aucun des $\mathbf{x}_i$, qui ont tous une masse d'au moins 2g. De plus, ce transport n'est pas symétrique. En effet, pour l'exemple de la Figure 2.2a, μ peut-être déplacé sur ν , mais le contraire n'est pas possible. La masse de 3 au point $\mathbf{y}_1$ ne peut se scinder pour approvisionner les $\mathbf{x}_i$. Cette difficulté a été résolue par KANTOROVITCH (1958) dans le cadre des mesures de probabilités. Je vais le présenter dans le cas des mesures de probabilités empiriques.

2.2 Transport optimal de mesures empiriques

2.2.1 Formulation en terme de mesure de probabilité

Dans l'exemple de Monge sur les déblais et les remblais, la masse totale des cailloux ne varie pas lors du déplacement. Nous pouvons donc re-normaliser pour que la masse totale soit égale à 1. Les cailloux à déplacer sont donc la donnée d'une collection de positions, ayant chacune une masse associée. Nous pouvons donc écrire μ sous la forme d'une mesure de probabilité sur $\mathbb{R}^2$:

$$\mu = \sum_{i=0}^{I-1} \mu_i \delta_{\mathbf{x}_i}.$$

Les μ_i sont les masses à chaque point $\mathbf{x}_i$, et les $\delta_{\mathbf{x}_i}$ sont des masses de Dirac. L'objet μ est une *mesure de probabilité*, et est la mesure empirique que nous estimerons des données. L'objet d'arrivée ν est aussi une mesure de probabilité. Si nous reprenons l'exemple de la Figure 2.2b, nous aurions :

$$\mu = \frac{3}{7} \delta_{\mathbf{x}_0} + \frac{2}{7} \delta_{\mathbf{x}_1} + \frac{2}{7} \delta_{\mathbf{x}_2},$$

$$\nu = \frac{1}{7}\delta_{\mathbf{y}_0} + \frac{6}{7}\delta_{\mathbf{y}_1}.$$

Si nous acceptons maintenant que les masses μ_i peuvent se scinder, nous pouvons les décomposer en sous masses γ_{ij} approvisionnant chaque ν_j :

$$\mu_i = \sum_j \gamma_{ij}.$$

De la même manière, les ν_j se décomposent en sous masses venant de chaque μ_i :

$$\nu_j = \sum_i \gamma_{ij}.$$

Les coefficients γ_{ij} définissent un *plan de transport* entre μ et ν , en déplaçant une fraction de μ_i , localisé à $\mathbf{x}_i$ vers ν_j , localisé à $\mathbf{y}_j$. Le problème présenté à la Figure 2.2b admet maintenant plusieurs solutions. Nous avons représenté deux d'entre elles sur la Figure 2.3.

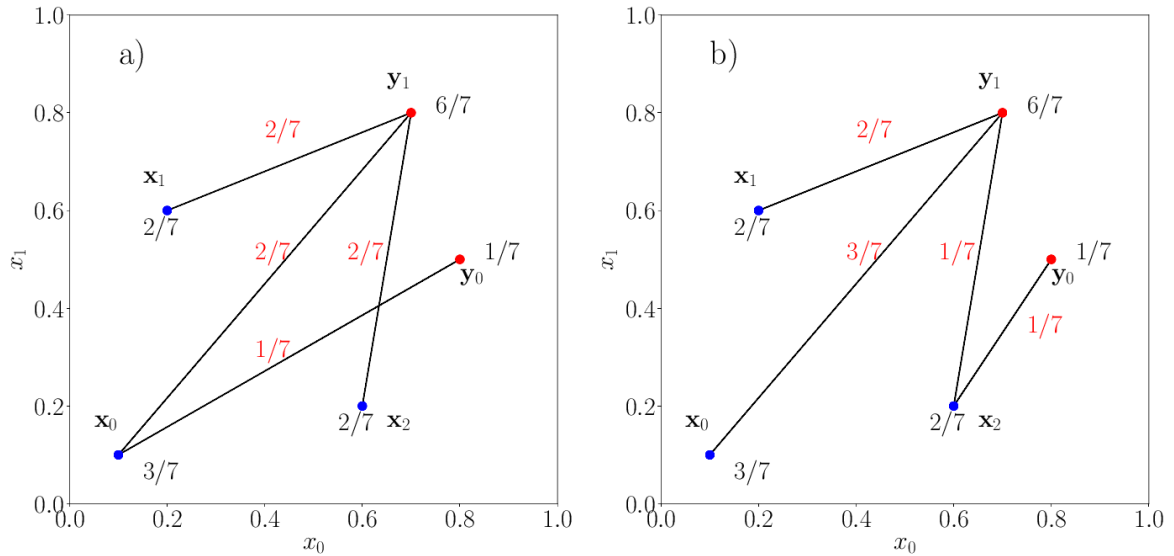


FIGURE 2.3 – Deux plans de transport entre les mesures empiriques μ (points bleus en $\mathbf{x}_i$) et ν (points rouges en $\mathbf{y}_j$). Les segments noirs sont les trajectoires de déplacements. Les valeurs rouges sont les masses déplacées.

Sur la Figure 2.3a, le plan γ_a envoie $\mathbf{x}_1$ et $\mathbf{x}_2$ sur $\mathbf{y}_1$. Par contre, $\mathbf{x}_0$ est scindé en deux sous masses de $1/7$ et $2/7$, qui sont respectivement envoyées sur $\mathbf{y}_0$ et $\mathbf{y}_1$. Sur la Figure 2.3b, avec le plan γ_b , c'est $\mathbf{x}_2$ qui est scindé, tandis que $\mathbf{x}_0$ approvisionne totalement $\mathbf{y}_1$. Calculons les coûts des plans γ_a et γ_b :

$$C(\gamma_a) = \frac{1}{7} \left(2\|\mathbf{x}_1 - \mathbf{y}_1\|^2 + 2\|\mathbf{x}_0 - \mathbf{y}_1\|^2 + \|\mathbf{x}_0 - \mathbf{y}_0\|^2 + 2\|\mathbf{x}_2 - \mathbf{y}_1\|^2 \right) = 0.724,$$

$$C(\gamma_b) = \frac{1}{7} \left(2\|\mathbf{x}_1 - \mathbf{y}_1\|^2 + 3\|\mathbf{x}_0 - \mathbf{y}_1\|^2 + \|\mathbf{x}_2 - \mathbf{y}_1\|^2 + \|\mathbf{x}_2 - \mathbf{y}_0\|^2 \right) = 0.72.$$

Le plan γ_b est donc meilleur. Il est en fait optimal, et minimise la fonctionnelle C .

2.2.2 Problème de Monge-Kantorovich

Nous allons nous placer dans un contexte plus général. Les mesures μ et ν seront toujours des mesures empiriques, mais sur $\mathbb{R}^d$. Elles s'écrivent sous la forme suivante

$$\mu = \sum_{i=0}^{I-1} \mu_i \delta_{\mathbf{x}_i}, \quad \nu = \sum_{j=0}^{J-1} \nu_j \delta_{\mathbf{y}_j}.$$

Un plan de transport (optimal ou non), est une collection de masses γ_{ij} connectant $\mathbf{x}_i$ à $\mathbf{y}_j$, vérifiant

$$\sum_i \gamma_{ij} = \nu_j, \quad \sum_j \gamma_{ij} = \mu_i, \quad \gamma_{ij} \geq 0. \quad (2.3)$$

Les γ_{ij} induisent une mesure sur l'espace $\mathbb{R}^d \times \mathbb{R}^d$:

$$\gamma = \sum_{ij} \gamma_{ij} \delta_{(\mathbf{x}_i, \mathbf{y}_j)}.$$

Cette mesure γ est une mesure de probabilité, ce qui peut se vérifier en calculant la somme :

$$\sum_{ij} \gamma_{ij} = \sum_i \sum_j \gamma_{ij} = \sum_i \mu_i = 1.$$

Les égalités (2.3) s'interprètent comme une projection sur la seconde et la première marge (i.e. sur $\mathbb{R}^d$ et sur $\mathbb{R}^d$ dans le produit $\mathbb{R}^d \times \mathbb{R}^d$). Nous pouvons ainsi reformuler le problème de la recherche d'un plan optimal de la manière suivante. Notons π_0 (resp. π_1) la projection sur la première (resp. seconde) marge. Alors γ est un plan de transport optimal si il vérifie

$$\left\{ \begin{array}{l} \inf_{\gamma} C(\gamma) = \inf_{\gamma} \sum_{ij} \gamma_{ij} \|\mathbf{x}_i - \mathbf{y}_j\|^2, \\ \pi_0(\gamma) = \mu, \\ \pi_1(\gamma) = \nu. \end{array} \right. \quad (2.4)$$

Plus généralement, prenons μ et ν deux mesures de probabilités quelconques sur $\mathbb{R}^d$. Notons $\Gamma(\mu, \nu)$ l'ensemble des mesures γ admettant μ (resp. ν) comme première (resp.

seconde) marge. Alors γ est optimal si il minimise la fonctionnelle suivante :

$$\inf_{\tilde{\gamma} \in \Gamma} C(\tilde{\gamma}) = \inf_{\tilde{\gamma} \in \Gamma} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|\mathbf{x} - \mathbf{y}\|^2 . d\tilde{\gamma}(\mathbf{x}, \mathbf{y}).$$

Ce problème s'appelle le *problème de Monge-Kantorovich*. Dans le cas qui nous occupe (i.e. μ et ν sont des mesures empiriques), ce problème admet toujours une solution (SANTAMBROGIO, 2015 ; BAZARAA et al., 2009 ; VILLANI, 2008). Nous allons maintenant nous focaliser sur le coût en lui même.

2.3 Distance de Wasserstein

Nous avons utilisé jusqu'ici la fonctionnelle de coût C comme un critère de minimisation. Mais la valeur de son minimum est elle même importante. Elle définit une distance, nommée la *distance de WASSERSTEIN* (1969). Nous la noterons $\mathcal{W}$, et est définie par

$$\mathcal{W}(\mu, \nu) = \sqrt{\inf_{\tilde{\gamma} \in \Gamma} C(\tilde{\gamma})}. \quad (2.5)$$

Cette fonction est une distance dans le sens mathématique du terme, i.e. elle vérifie les propriétés suivantes

1. *Symétrie* : $\mathcal{W}(\mu, \nu) = \mathcal{W}(\nu, \mu)$,
2. *Séparation* : $\mathcal{W}(\mu, \nu) = 0$ est équivalent à $\mu = \nu$,
3. *Inégalité triangulaire* $\mathcal{W}(\mu, \nu) \leq \mathcal{W}(\mu, \xi) + \mathcal{W}(\xi, \nu)$.

Appliquons cette distance au problème présenté sur la Figure 2.1. Nous avons une Gaussienne μ_0 qui était translatée le long de l'axe réel, définie par $\mu_t = \mu_0 + 2t$. Nous avons reproduit sur la Figure 2.4a les Gaussiennes au temps 0, 2, 4, 6 et 8. Sur la Figure 2.4b, nous avons représenté la distance Euclidienne en rouge entre μ_0 et μ_t , et la distance de Wasserstein en bleu. La distance Euclidienne stagne à partir de $t = 2$, et n'est pas reliée à la dynamique de translation. La distance de Wasserstein est égale à la moyenne de μ_t à l'instant t , correspondant à un déplacement de μ_0 de $2t$. Ceci correspond au déplacement de chaque point de la densité de μ_0 : ils sont translatés de $2t$. Cette distance représente bien la dynamique sous jacente.

2.4 Lien entre les systèmes dynamiques et le transport optimal

Le plan de transport γ entre deux mesures μ et ν donne des directions de déplacements des masses. Dans ce sens, il est une estimation d'une fonction qui transforme μ en ν .

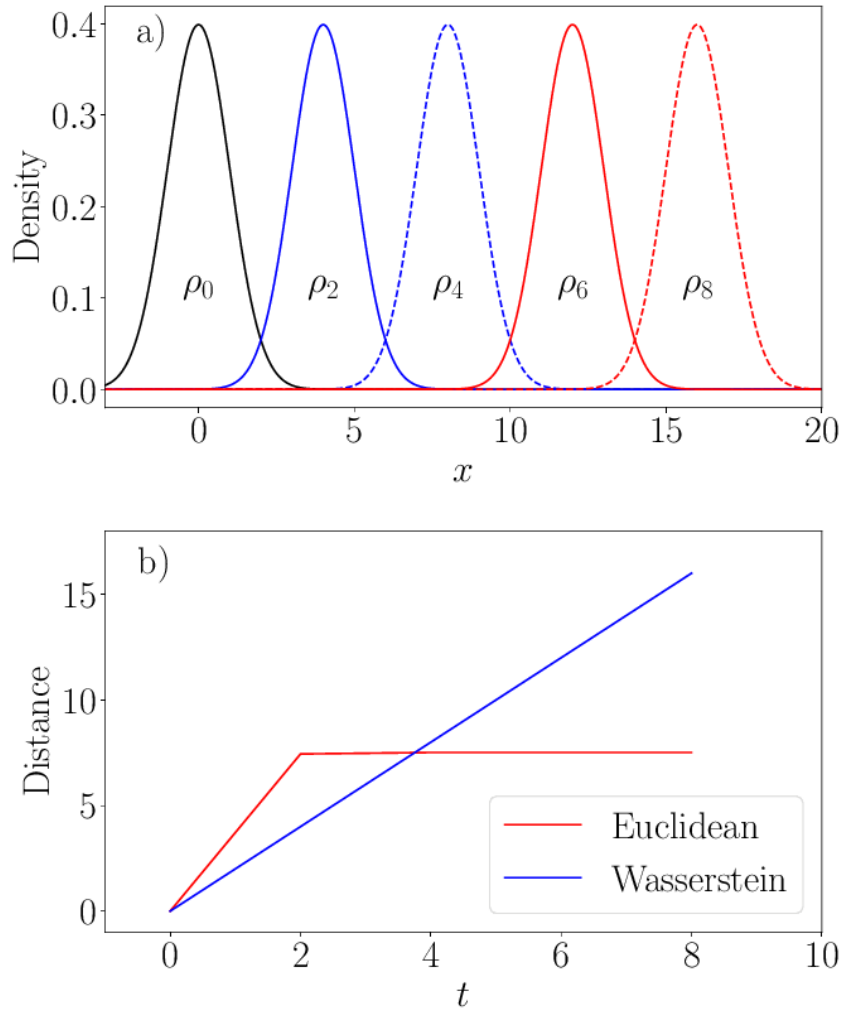


FIGURE 2.4 – Les mesures μ_t sont des lois normales centrées en $2t$, d'écart type égal à 1. **a)** Densités de $\mu_t = \mu_0 + 2t$ aux temps 0, 2, 4, 6 et 8. **b)** Distance entre μ_0 et μ_t . La distance Euclidienne est donnée en rouge, la distance de Wasserstein en bleu.

Replaçons nous dans le contexte des systèmes dynamiques. Pour un système dynamique non autonome, notons μ_t , $t \in [0, 1]$ sa séquence de mesures SRB entre un temps $t = 0$ et $t = 1$ (à une re-paramétrisation près, nous pouvons toujours nous ramener à l'intervalle $[0, 1]$ pour le temps). Nous avons vu au Chap. 1 que la mesure SRB d'un système dynamique non autonome $\mathbf{v}$ suivait l'équation de transport suivante :

$$\begin{cases} \partial_t \mu_t + \langle \nabla, \mathbf{v} \mu_t \rangle = 0 \\ \mu|_{t=0} = \mu_0 \end{cases}$$

Prenons le problème à l'envers, et cherchons maintenant à quantifier la différence entre μ_0 et μ_1 . À deux mesures μ_0 et μ_1 fixées, nous pouvons chercher un champ de vecteurs $\mathbf{v}$

et une séquence de mesures μ_t vérifiant l'équation de transport suivante :

$$\begin{cases} \partial_t \mu_t + \langle \nabla, \mathbf{v} \mu_t \rangle = 0 \\ \mu_{t=0} = \mu_0, \quad \mu_{t=1} = \mu_1. \end{cases} \quad (2.6)$$

Le théorème de BENAMOU et BRENIER (2000) nous assure alors le lien suivant avec la dynamique sous jacente. La distance de Wasserstein entre μ_0 et μ_1 est donnée par :

$$\mathcal{W}(\mu_0, \mu_1)^2 = \inf_{(\tilde{\mathbf{v}}, \tilde{\rho}_t)} \int_{\mathbb{R}^d} \int_0^1 \tilde{\rho}_t(\mathbf{x}) |\tilde{\mathbf{v}}(t, \mathbf{x})|^2 dt d\mathbf{x} \quad (2.7)$$

où $\tilde{\rho}_t$ est la densité de $\tilde{\mu}_t$, i.e. pour $A \subset \mathbb{R}^d$, $\tilde{\mu}_t(A) = \int_A \tilde{\rho}_t dx$, et où les couples $(\tilde{\mathbf{v}}, \tilde{\mu}_t)$ vérifient l'Éq. (2.6).

Ce résultat assure que la distance de Wasserstein (i.e. le coût minimum) représente l'action d'un champ de vecteurs $\mathbf{v}$ suivant une équation de transport. Dans le cas où nous disposons uniquement d'observations discrètes de μ_0 et μ_1 , le plan de transport *infère* la dynamique sous jacente. Dans le cas des snapshots attracteurs, cela revient à une inférence du champ de vecteurs. Si nous l'utilisons dans le cadre de données plus complexes (i.e. des observations, une simulation d'un modèle climatique), cela revient à supposer qu'une équation de transport agit entre les objets considérés.

Remarquons aussi que la formulation de Benamou-Brenier fait apparaître des mesures intermédiaires entre μ et ν . Ces mesures sont appelées les *interpolées*. L'interpolation au sens du transport optimal a d'abord été définie sur $\mathbb{R}$ par MCCANN (1997), avant de s'étendre naturellement à toutes dimensions au travers du théorème de Benamou-Brenier. L'intérêt de cette interpolation est de pouvoir reconstruire des temps intermédiaires avec une estimation de la dynamique agissant entre μ et ν .

2.5 Intuitions des mécanismes d'action du transport optimal

Je propose dans cette partie de se familiariser avec le transport optimal via un exemple unidimensionnel. Le transport optimal s'étend naturellement en dimensions quelconques, mais nous pourrons ainsi visualiser le plan de transport (qui sera sur $\mathbb{R}^2$). La mesure μ sera une Gaussienne, centrée en 0 et de variance égale à 1. La mesure ν sera bi-modale. Elle sera constituée de deux Gaussiennes, centrées en -3 et 3 , toujours de variance égale à 1. Les histogrammes de ces mesures sont représentés sur la Fig. 2.5.

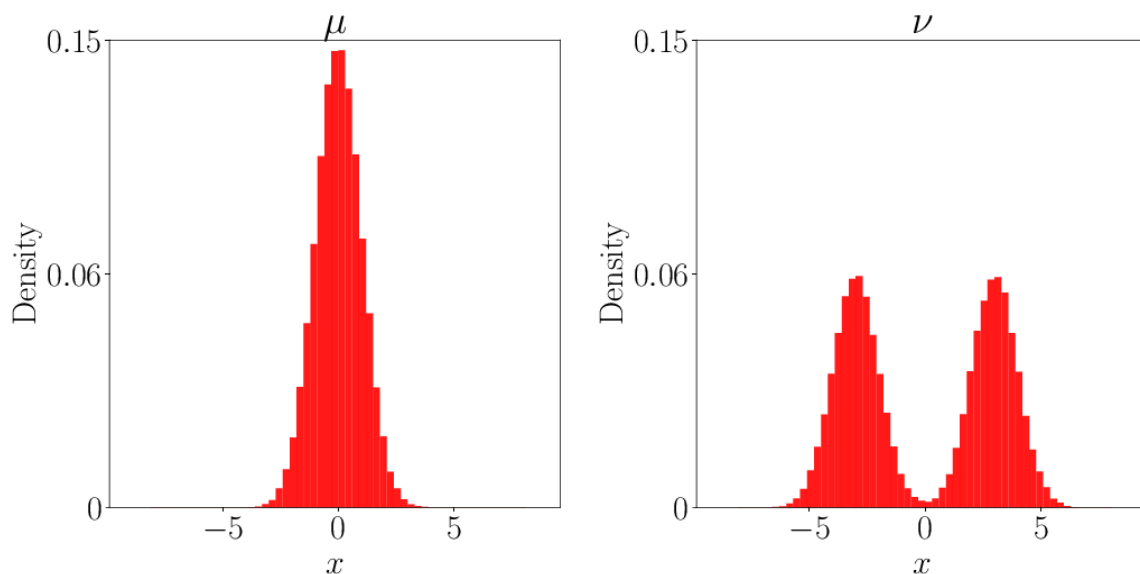


FIGURE 2.5 – Histogrammes d’un exemple de deux mesures μ et ν . La mesure μ est une Gaussienne centrée en 0, de variance 1. La mesure ν est bi-modale, constituée de deux Gaussiennes centrées respectivement en -3 et 3 , toutes deux de variances 1. Les bins sont uniformes, de longueur égale à 0.3. Enfin, 10000 points ont été utilisés pour générer ces histogrammes.

Je propose dans un premier temps de visualiser le plan de transport. Ensuite, nous regarderons comment le transport optimal se comporte lors d’une perturbation par un bruit blanc de μ et ν . Nous étudierons aussi l’influence du choix de la longueur des bins sur le transport.

2.5.1 Représentation du plan de transport comme mesure de probabilité

Je propose deux manières de visualiser le plan optimal qui permet de transformer μ en ν comme un histogramme sur $\mathbb{R}^2$. Les histogrammes de μ , ν et du plan de transport sont visibles sur la Figure 2.6. Les histogrammes de μ et ν sont respectivement sur les panneaux de droite et du haut. En dessous de chacun d’eux, la masse dans chaque bin est convertie en couleur. Une couleur bleu indique une masse proche de 0, une couleur rouge une masse importante. Le plan de transport est au centre avec le même code couleur. Dans l’alignement de chaque case de couleur, à droite se trouve l’origine de la masse, en haut la cible. Nous pouvons voir que les masses de μ se décomposent pour approvisionner plusieurs masses de ν . En particulier, le centre de μ contient une masse importante, qui doit être déplacée vers les points -3 et 3 de ν , expliquant l’étalement du plan au centre. Le

plan de transport a une structure monotone croissante, conservant les quantiles de chaque distribution. Cette représentation permet de visualiser le plan de transport comme mesure de probabilité sur l'espace $\mathbb{R}^2$, nous allons maintenant voir son action dynamique.

2.5.2 Visualisation dynamique du plan de transport

Les histogrammes μ_t des interpolés (c.f. Eq. (2.7)) de μ et ν sont visibles sur les figures 2.7a-f, respectivement aux temps $t = 0.2, 0.4, 0.6$ et 0.8 . Les mesures μ et ν sont aussi représentées comme des interpolées aux temps $t = 0$ et $t = 1$. Ces histogrammes ont été obtenus à partir du plan de transport γ de la manière suivante :

1. Tirer une réalisation x selon la mesure μ ,
2. Calculer y tel que x soit envoyé sur y ,
3. Le nombre $(1 - t)x + ty$ est une réalisation de μ_t .

Avec les histogrammes μ_t , nous pouvons visualiser pas à pas la façon dont le transport opère entre μ et ν . Pour $t = 0.2$, l'interpolé ressemble à une Gaussienne coupée en deux, dont chaque partie aurait été légèrement translatée de part et d'autre. À $t = 0.4$, la masse commence à s'étaler, diminuant les extrêmes proches de 0. À $t = 0.6$, les deux Gaussiennes de ν apparaissent distinctement, même si la partie centrale est encore très abrupte. À $t = 0.8$, les Gaussiennes sont presque formées, il reste un simple décalage pour finir le transfert.

L'intuition qu'il faut avoir en tête (y compris en dimension supérieure) est que le transport optimal décale continument les points de masses.

2.5.3 Résistance au bruit et influence de l'histogramme

Je propose dans cette section d'étudier la robustesse de la distance de Wasserstein. Pour cela, je vais d'une part perturber les mesures μ et ν , et d'autre part faire varier la longueur des bins permettant de les estimer.

Pour perturber μ et ν , je vais ajouter un bruit blanc Gaussien, centré en 0 et d'écart type $\sigma_0, \sigma_1 \in [0, 10]$. Ceci permet de définir les mesures μ_{σ_0} et ν_{σ_1} comme étant des réalisations de μ et ν , en ajoutant un bruit. Je calculerai ensuite les distances $\mathcal{W}(\mu_{\sigma_0}, \nu_{\sigma_1})$, pour $\sigma_0, \sigma_1 \in [0, 10]$, et une longueur de bin prenant les valeurs suivantes : 1, 0.5, 0.1 et 0.05. De plus, ce protocole sera relancé 200 fois, afin d'inférer les distances $\mathcal{W}(\mu_{\sigma_0}, \nu_{\sigma_1})$.

Sur les figures 2.8-2.9 sont représentés les moyennes et écart types des distances $\mathcal{W}(\mu_{\sigma_0}, \nu_{\sigma_1})$. Pour chaque figure, l'abscisse est la force du bruit σ_0 ajouté à μ , l'ordonnée

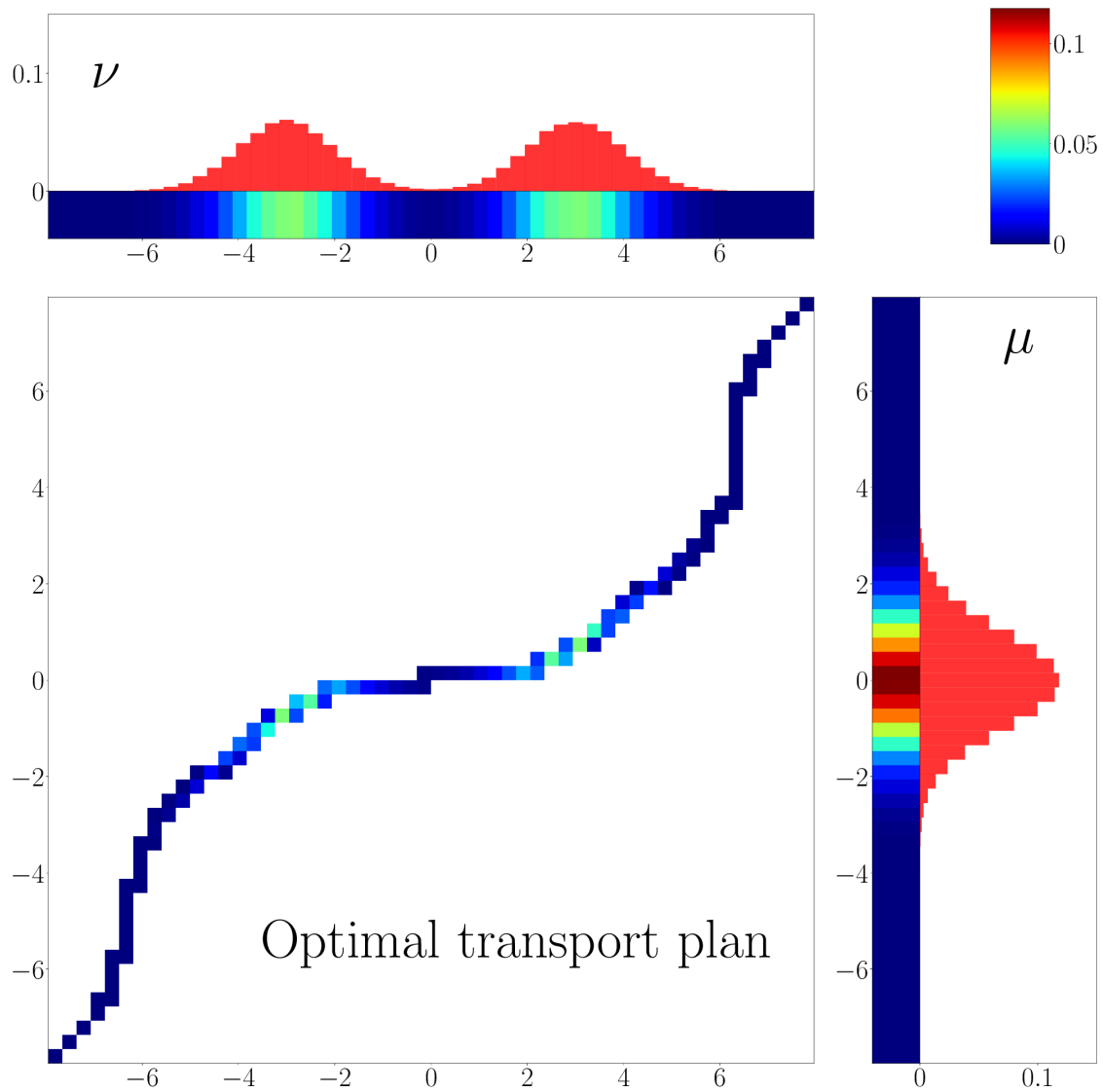


FIGURE 2.6 – Représentation du plan de transport entre μ et ν . La mesure μ est une Gaussienne centrée réduite. Son histogramme (Figure de droite) a été estimé à partir de 10000 points, avec une taille de bin constante égale à 0.3. La mesure ν est une bimodale composée de deux Gaussiennes centrées respectivement en -3 et 3 et de variance 1. Son histogramme (Figure du haut) a été estimé à partir de 10000 points, avec une taille de bin constante égale à 0.3. En dessous de chaque histogramme est représenté la quantité de masse dans chaque bin à l'aide d'une colormap. La figure centrale est le plan de transport optimal qui associe μ et ν . Les masses de couleur sont la quantité de masses déplacées de chaque bin de μ vers les bins de ν .

la force du bruit σ_1 ajoutée à ν . Le code couleur donne la valeur moyenne et l'écart type des distances. Les quatre panneaux a-d de chaque figure représentent respectivement les

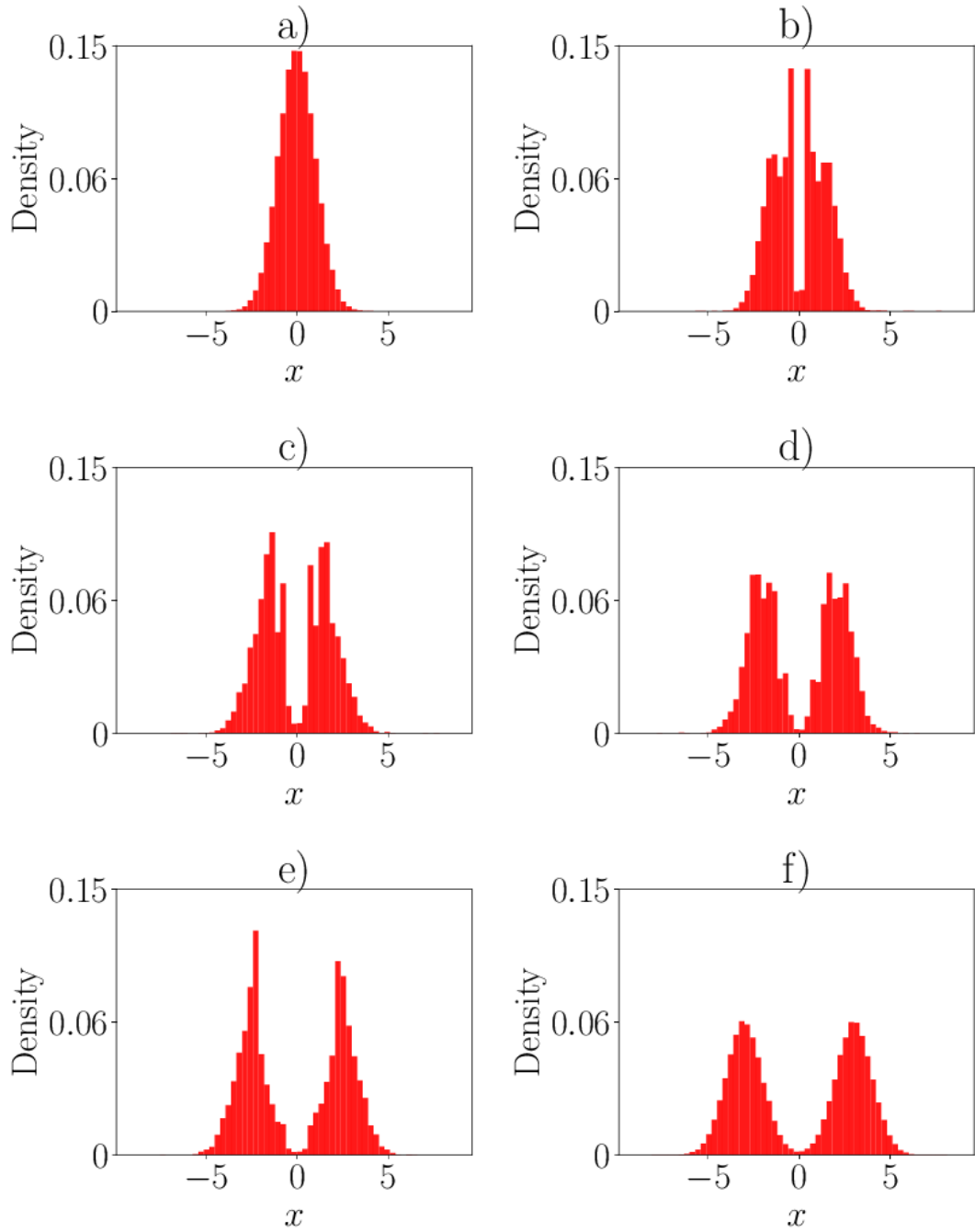


FIGURE 2.7 – Visualisation de la séquence μ_t , obtenue par interpolation de μ et ν (c.f. Eq. (2.7)) Les panneaux **a-f** correspondent respectivement à une interpolation aux temps $t = 0, 0.2, 0.4, 0.6, 0.8$ et 1 . En particulier, pour $t = 0$ et $t = 1$, les mesures obtenues sont μ et ν .

distances pour les quatre tailles de bins 1, 0.5, 0.1 et 0.05. Nous pouvons déjà voir que la taille du bin n'a quasiment aucune influence sur la valeur de la distance. La règle de STURGES (1926) pour ces données suggère une longueur d'histogramme d'environ 0.05,

mais une longueur 20 fois plus élevée ne dégrade pas les distances. Bien évidemment, ces valeurs restent dans un intervalle "raisonnable" par rapport aux données, i.e. supérieur à l'inverse de la taille des données et inférieur à leur étendue. Une longueur de bin tendant vers l'infini rendrait la distance de Wasserstein systématiquement nulle. Une longueur de bin tendant vers 0 rendrait le calcul trop coûteux, pour un faible gain d'information.

Analysons en détail les moyennes des distances sur la Figure 2.8. Remarquons d'abord que les distances sont continues avec σ_0 et σ_1 , i.e. quand le bruit change la distance varie continument avec le bruit. Suivons maintenant l'axe des abscisses, i.e. les distances de la forme $\mathcal{W}(\mu_{\sigma_0}, \nu)$. Les distances décroissent jusqu'à $\sigma_0 = 3.5$, puis augmentent. Lorsque nous perturbons μ , elle "s'étale" de plus en plus, i.e. nous augmentons son écart type. Ce faisant, nous augmentons les masses localisées autour de -3 et 3 , réduisant la distance entre μ_{σ_0} et ν . Les bruits de force supérieure envoient les masses au delà du support de ν , augmentant mécaniquement le coût de transport. Avec cette idée en tête, quand nous perturbons ν , nous envoyons de la masse loin dans le positif/négatif, augmentant la distance avec μ . Inversement, quand nous perturbons μ , nous rattrapons les perturbations de ν . Ceci explique la trajectoire de distance quasiment nulle, autour de laquelle se répartissent les distances plus fortes, quand un σ_0 prend l'avantage sur un σ_1 .

Analysons maintenant l'écart type des distances, sur la Figure 2.9. Là encore, la longueur du bin n'a quasiment aucune influence. Nous pouvons seulement relever que la diagonale pour un bin de 1 est légèrement plus élevée (nous passons de 0.03 à 0.05). Sinon, la valeur maximale tout confondu est à 0.1. Comparée à la taille des bins, et aux valeurs prises par les distances, cette valeur est extrêmement faible, montrant une robustesse face au bruit de la distance de Wasserstein.

Une étude plus approfondie sur des mesures plus complexes, issues de dynamiques chaotiques sera l'objet du chapitre suivant.

2.6 Exemple d'utilisation de la distance de Wasserstein dans le climat

Je vais revenir dans cette section sur l'exemple de la Figure 1.10, dans le chapitre 1. J'avais estimé les histogrammes de la température en Ile-de-France sur la période 2000-2100, découpée en 5 sous périodes de 20 ans. Les données proviennent de deux simulations du modèle de l'IPSL (DUFRESNE et al., 2013). La première est une simulation de contrôle, i.e. le forçage est celui de l'ère pré-industrielle, la deuxième suit le scénario RCP85 (MOSS

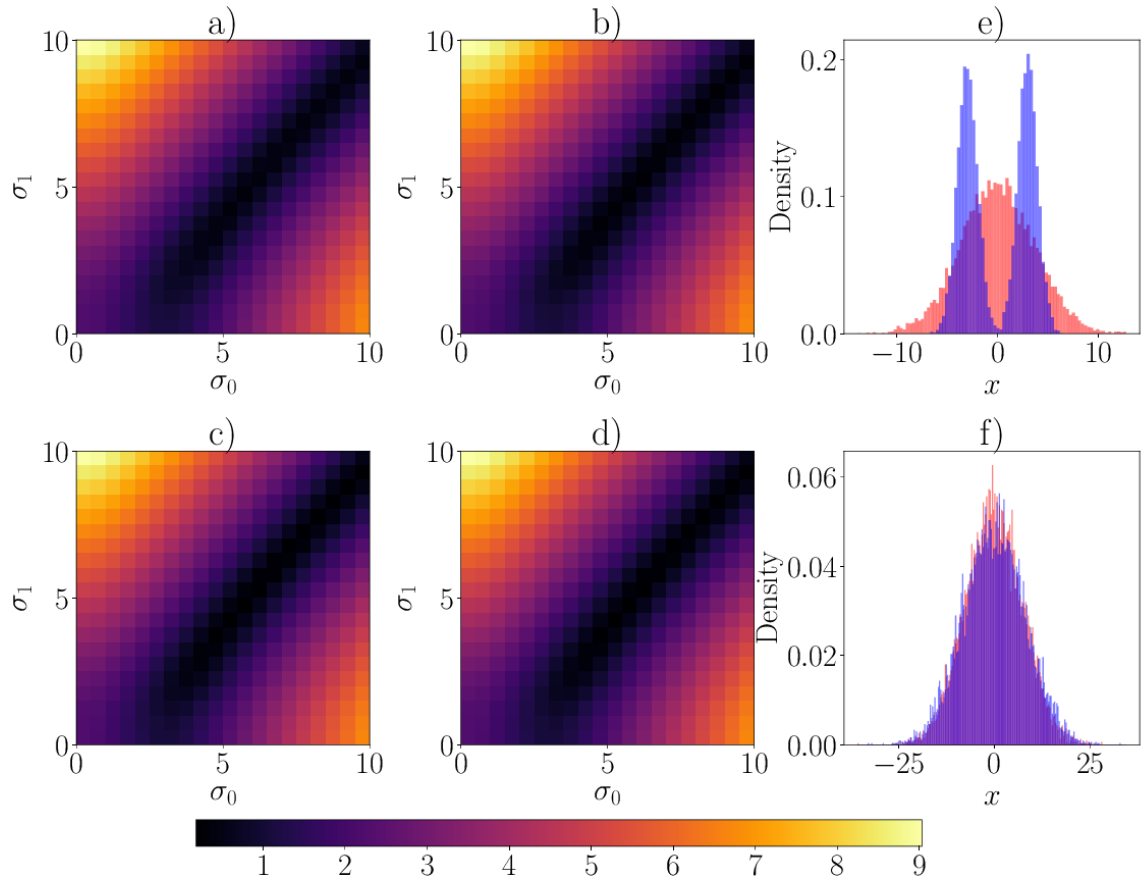


FIGURE 2.8 – Moyenne des 200 distances de Wasserstein entre une loi normale centrée réduite, bruitée par une loi normale centrée d'écart type σ_0 , et une mixture de deux lois normales centrées en -3 et 3 , bruitées par une loi normale d'écart type σ_1 . Sur les panneaux **a-d**, la longueur des bins est respectivement de 1, 0.5, 0.1 et 0.005. **e)** Histogramme de μ_{σ_0} (rouge) et ν_{σ_1} (bleu) pour $\sigma_0 = 3.5$ et $\sigma_1 = 0$. **f)** Histogramme de μ_{σ_0} (rouge) et ν_{σ_1} (bleu) pour $\sigma_0 = 7.5$ et $\sigma_1 = 7.5$.

et al., 2010). Visuellement. La Figure 1.10 mettait visuellement en évidence la stabilité de la simulation de contrôle, et montrait des modifications des histogrammes de température entre les différentes sous périodes de la simulation du scénario RCP85. Je propose d'utiliser la distance de Wasserstein pour quantifier ces déformations, et de la comparer à la moyenne et l'écart type des distributions, qui sont des indicateurs statistiques usuellement utilisés.

Notons $\mu_0, \mu_{20}, \mu_{40}, \mu_{60}, \mu_{80}$ les histogrammes de la simulation pré-industrielle sur les périodes 2000-2020, 2020-2040, 2040-2060 et 2080-2100, et utilisons $\nu_0, \nu_{20}, \nu_{40}, \nu_{60}$ et ν_{80} pour la simulation du scénario RCP85. Pour effectuer cette comparaison, j'ai considéré les histogrammes μ_0 et ν_0 comme des références, et calculé les distances :

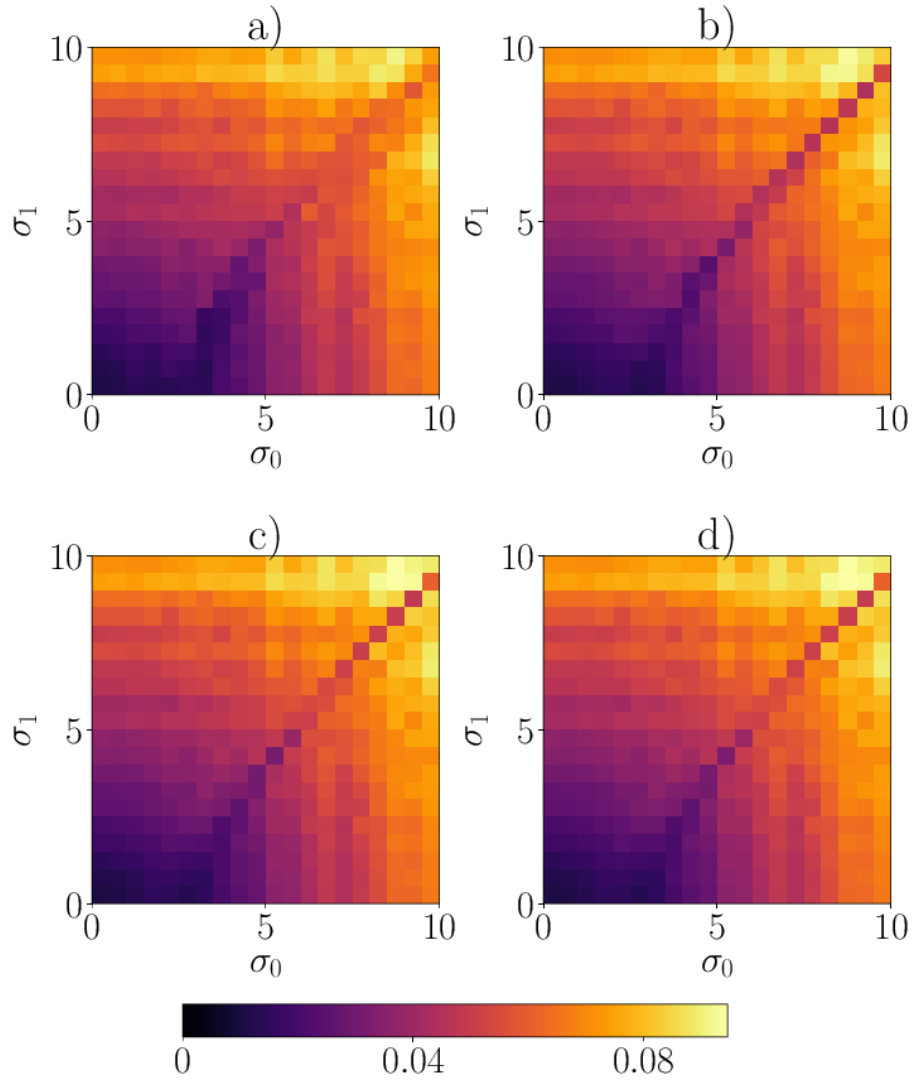


FIGURE 2.9 – Écart type des 200 distances de Wasserstein entre une loi normale centrée réduite, bruitée par une loi normale centrée d'écart type σ_0 , et une mixture de deux lois normales centrées en -3 et 3 , bruitées par une loi normale d'écart type σ_1 . Sur les panneaux **a-d**, la longueur des bins est respectivement de 1, 0.5, 0.1 et 0.005.

$$\mathcal{W}(\mu_0, \mu_i), \quad \mathcal{W}(\nu_0, \nu_i), \quad i \in \{20, 40, 60, 80\}.$$

Ces distances sont représentées en rouge (resp. bleu) pour μ (resp. ν) sur la Figure 2.10a. J'ai aussi calculé la différence entre la moyenne de μ_i (resp. ν_i) et μ_0 (resp. ν_0). Ces indicateurs sont la ligne rouge (resp. bleu) en pointillés de la Figure 2.10a. Les courbes de la différence de moyenne et celle de la distance de Wasserstein sont extrêmement similaires. Effectivement, la différence de moyenne correspond au transport moyen de tous les éléments des distributions, information immédiatement capturée par le transport optimal. Nous voyons aussi que la simulation RCP85 s'éloigne énormément de la

simulation pré-industrielle, traduisant un forçage dans la série temporelle, ce qui est cohérent avec ce scénario. Le forçage par le scénario RCP85 entraînant une modification forte de la moyenne, les modifications des moments d'ordre supérieur peuvent être masquées.

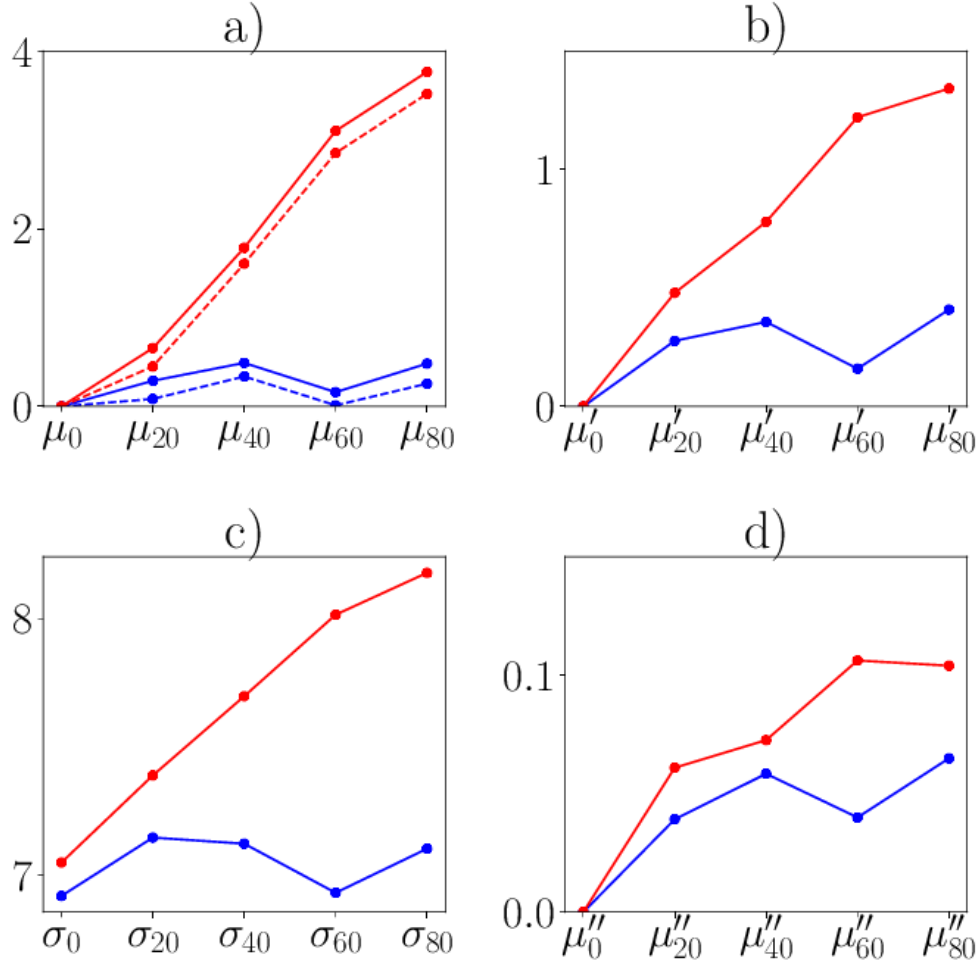


FIGURE 2.10 – **a)** Distance de Wasserstein entre μ_i et μ_0 . Les mesures μ_i sont la distribution de probabilités des températures en Ile-de-France pour le modèle de l'IPSL sur les périodes 2000-2020, 2020-2040, 2040-2060, 2060-2080 et 2080-2100. **b)** Idem mais entre les histogrammes centrés. **c)** Écart types pour les mêmes périodes et variables. **d)** Idem que a) et b) mais pour des distributions centrées et réduites.

Pour cette raison, j'ai calculé les mêmes distances mais entre les histogrammes μ'_i (resp. ν'_i), où $\mu'_\bullet$ (resp. $\nu'_\bullet$) est l'histogramme de la simulation pré-industrielle (resp. RCP85) recentrée (i.e. la moyenne a été soustraite). Ces distances ont été représentées sur la Figure 2.10b, en rouge pour $\mu'_\bullet$ et bleu pour $\nu'_\bullet$. La déviation par la moyenne n'influencera donc plus le calcul du transport. Nous pouvons voir que la trajectoire du scé-

nario RCP85 est là encore très différente du scénario historique, et traduit bien l’existence d’un forçage qui a altéré le système climatique. Nous pouvons en déduire que le forçage a aussi fortement influencé les moments d’ordre au moins égal à 2. Sur la Figure 2.10c, j’ai aussi représenté l’écart type de chaque sous période. Nous pouvons voir que les courbes sont extrêmement similaires à celles représentées sur la Figure 2.10, indiquant une influence importante de l’écart type sur la distance de Wasserstein.

Vérifions si le scénario RCP85 a aussi altéré les moments d’ordres supérieurs à 3. Je propose de recalculer les mêmes distances, mais en centrant et réduisant les distributions (i.e. retirer la moyenne et diviser par l’écart type). Ces distances sont représentées sur la Figure 2.10d. Nous pouvons voir que les distances des simulations pré-industrielle et du scénario RCP85 commencent à se rapprocher, même si ce dernier reste plus élevé. Une altération des moments d’ordre au moins 3 est donc possible.

Évidemment, les conclusions de cet exemple sont à relativiser (taille des données, uniquement un point de grille pour une simulation de un modèle... etc), et avait uniquement comme but de montrer comment la distance de Wasserstein peut-être utilisée pour manipuler des données climatiques. Le prochain chapitre sera consacré à l’utilisation de la distance de Wasserstein pour quantifier la déformation du modèle de LORENZ (1984) par un forçage type changement climatique.

2.7 Résumé

En travaillant uniquement avec des statistiques d’un système (comme le climat), il devient nécessaire de pouvoir quantifier finement les différences entre mesures de probabilités. Classiquement, ces différences sont calculées avec des distances type Euclidienne. Ces métriques ne prennent pas en compte la dynamique qui agit entre deux périodes temporelles.

Une métrique telle que définie par la distance de Wasserstein offre de multiples avantages. Une quantification naturelle de la dynamique (théorème de Benamou Brenier), une résistance au bruit et une prise en compte de toutes les caractéristiques des distributions. En contrepartie, son temps de calcul est bien plus élevé (d’ordre cubique), mais devient acceptable avec la puissance de calcul actuellement disponible.

Un exemple d’application à des données climatiques a été présenté, la quantification de la différence entre une simulation de contrôle pré-industrielle et une simulation du scénario RCP85. Poursuivant dans cette voie, le chapitre 3 présente l’étude d’un forçage type

changement climatique du modèle de LORENZ (1984). L'accent sera mis sur la capacité de la distance de Wasserstein à distinguer des dynamiques différentes, et sur la quantité de données nécessaires pour étudier un système dynamique complexe. A l'inverse, le chapitre 4 sera consacré à la correction de biais multivariés. La distance de Wasserstein mesurant la différence entre distributions en multivarié, il est possible d'utiliser son plan de transport pour corriger des données.

Chapitre 3

Application à des systèmes dynamiques chaotiques non autonomes

Pour illustrer la pertinence de l'utilisation de la distance de Wasserstein dans le climat, des tests ont été réalisés sur le modèle de LORENZ (1984). Dans un premier temps, la capacité de distinguer deux attracteurs a été vérifiée en analysant la distance de Wasserstein entre deux snapshots attracteurs de l'été et l'hiver de ce modèle. Partant de l'idée que les données disponibles en climat sont limitées, le nombre de points de grille nécessaires pour faire la distinction a été étudié. Une comparaison avec la distance Euclidienne a aussi été effectuée. Dans un second temps, la distance de Wasserstein a été utilisée pour analyser l'effet du forçage de DRÓTOS et al. (2015), similaire au changement climatique. Là aussi, le nombre de points nécessaires pour quantifier le forçage a été étudié. Ces résultats ont été publiés dans ROBIN et al. (2017).

Dans le prolongement de cet article, des tests ont été effectués sur le modèle de l'IPSL DUFRESNE et al. (2013) dans l'hémisphère nord. Une première analyse de l'influence des scénarios RCP sur la température et la précipitation a été effectuée. Ces résultats sont présentés dans la Section 3.3.

3.1 Abstract (ROBIN et al., 2017)

Le système climatique peut être décrit par un système dynamique et l'attracteur associé. La dynamique de cet attracteur dépend de forçages extérieurs qui peuvent influencer le climat. Ces forçages peuvent affecter les valeurs moyennes et les variances, mais les régions de l'attracteur qui sont rarement visitées peuvent aussi être affectées. Il est difficile de mesurer la réponse de l'attracteur climatique à différents forçages. Actuellement, la distance euclidienne ou

des métriques similaires comme la distance de Mahalanobis ont été favorisées pour mesurer la différence entre deux situations climatiques. Ces distances n'ont pas de mécanisme interne tenant compte de la dynamique de l'attracteur. Dans ce papier, nous montrons que la distance de Wasserstein, venant du transport optimal, offre un moyen efficace et applicable de discriminer plusieurs dynamiques. Après avoir traité un modèle jouet, nous explorons comment la distance de Wasserstein peut être appliquée et interprétée pour détecter des dynamiques non autonomes provenant d'un système de Lorenz forcé par un cycle saisonnier et un réchauffement.

3.2 Detecting changes in forced climate attractors with Wasserstein distance



Detecting changes in forced climate attractors with Wasserstein distance

Yoann Robin, Pascal Yiou, and Philippe Naveau

Laboratoire des Sciences du Climat et de l'Environnement, UMR8212 CEA-CNRS-UVSQ, Institut Pierre-Simon Laplace & Université Paris-Saclay, 91191 Gif-sur-Yvette Cedex, France

Correspondence to: Yoann Robin (yoann.robin@lsce.ipsl.fr)

Received: 3 February 2017 – Discussion started: 20 February 2017

Revised: 8 June 2017 – Accepted: 12 June 2017 – Published: 31 July 2017

Abstract. The climate system can be described by a dynamical system and its associated attractor. The dynamics of this attractor depends on the external forcings that influence the climate. Such forcings can affect the mean values or variances, but regions of the attractor that are seldom visited can also be affected. It is an important challenge to measure how the climate attractor responds to different forcings. Currently, the Euclidean distance or similar measures like the Mahalanobis distance have been favored to measure discrepancies between two climatic situations. Those distances do not have a natural building mechanism to take into account the attractor dynamics. In this paper, we argue that a Wasserstein distance, stemming from optimal transport theory, offers an efficient and practical way to discriminate between dynamical systems. After treating a toy example, we explore how the Wasserstein distance can be applied and interpreted to detect non-autonomous dynamics from a Lorenz system driven by seasonal cycles and a warming trend.

In addition to climate internal variability, external forcings (either natural or anthropogenic) perturb the climate system dynamics by introducing a time dependence of the attractor. This is the cause of non-stationary behavior of the climate system. At first order, this can translate into a general shift of the underlying attractor (Corti et al., 1999). At second order, interactions between a seasonal cycle and a slow forcing can even lead to trends in subtle quantities (e.g., Cassou and Cattiaux, 2016; Vrac et al., 2014). A few properties of the climate attractor due to external forcings (anthropogenic or not) have been treated by Pierini et al. (2016) and Drótos et al. (2015), who focused on low-dimensional strange attractors and investigated qualitative changes of the attractors, although all those studies are quantitative in many aspects. Lucarini et al. (2017) have recently used response theory (Ruelle, 2009) to quantify the modification of the dynamics submitted to a forcing.

Classical distances, like the Euclidean distance are often used to measure attractor differences. The goal of our paper is to present a framework, embedded in optimal transport theory (e.g., Villani, 2003), to measure the distance between strange attractors, and make a statistical inference of this tool on well-documented dynamical systems. To do this we exploit the fact that the attractor of the system defines an invariant measure, which is the multivariate probability distribution of all trajectories of the system. The distance between attractors is then computed through the cost to transform one invariant measure into another. A similar idea was already proposed in Ghil (2015) to characterize the climate variability. In particular, we assess that it is possible to discriminate between attractors, given a relatively low number of sampling points, in order to ensure the applicability of this

1 Introduction

If the climate system is viewed as a complex dynamical system yielding a strange attractor, i.e., a highly complicated object around which all trajectories wind up (Lorenz, 1963), then, climate variability is linked to the statistical properties of such an attractor (Ghil and Childress, 1987). Those statistical properties refer to the probability that trajectories visit each region of phase space (Mané, 2012; Eckmann and Ruelle, 1985). Mathematical concepts to describe those properties on rather simple dynamical systems have been investigated by Chekroun et al. (2011).

methodology. We test this method on a time-varying dynamical system in order to illustrate how the dynamics of a system can be affected by a constant forcing interacting with seasonality.

The paper is organized as follows. In Sect. 2, we recall some basic concepts used in optimal transport theory and recall the definition of the Wasserstein distance. In Sect. 3, we investigate the performance of the Wasserstein distance to discriminate between two “simple” autonomous systems (winter against summer of Lorenz, 1984 model). Section 4 explores how forcing can impact the Wasserstein distance capability at detecting changes in a non-stationarity context. Section 5 concludes and proposes some future research directions.

2 Distance between measures

To characterize changes in the properties of the attractor of a dynamical system, the first step of our methodology is to determine how two measures (or distributions of mass) differ. The idea is to derive a cost function for transporting one mass distribution onto the other. As a simple example, we consider the three mass distributions shown in Fig. 1, noted μ , ν and ξ . The distributions are on a grid of size $10 \times 10 = 100$, with mass positions located on pixels $\mathbf{x}_i$, $i = 1, \dots, 100$. ν is constructed to be a one pixel left shift of μ . The distribution ξ is a 90° rotation and a mirror image of μ , and we move one square to have a common point with μ . The distribution μ (resp., ν and ξ) can be written as

$$\mu = \sum_{i=1}^{100} \mu_i \delta_{\mathbf{x}_i},$$

where $\delta_{\mathbf{x}}$ is the Dirac mass at pixel $\mathbf{x}_i$, and $\mu_i = 1$ on the black boxes in Fig. 1 and $\mu_i = 0$ on the grey boxes. The Euclidean distance d between μ and ν is defined by

$$d(\mu, \nu)^2 = \sum_{i=1}^{100} |\mu_i - \nu_i|^2.$$

Panels a and b in Fig. 1 are visually very similar, whereas panel c cannot be deduced from a trivial transformation of the first panels. Therefore, it is expected that μ is “closer” to ν than ξ . We find that the Euclidean distance from μ to ν is 3.74, and the distance from μ to ξ is 3.46 (the example was constructed to show this). Thus, the Euclidean distance does not capture the structural proximity between μ and ν . The explanation is the following: among the squares that have no common mass, the value of the Euclidean distance is independent of the position of squares. This highlights the need of a distance that can take into account how masses should be moved, say, from the left panel to the middle panel of Fig. 1.

This mathematical problem traces back to Monge (1781) and is the basis of optimal transport theory (see, e.g., Villani, 2003). To transport the mass distribution μ contained in the boxes at $\mathbf{x}_i$ to the distribution ν in the boxes at $\mathbf{x}_j$, a total cost of the transport has to be defined. We note $\gamma_{ij} > 0$ the fraction of the mass transported from the boxes $\mathbf{x}_i$ to $\mathbf{x}_j$. The cost of the transport is defined by $\gamma_{ij} d(\mathbf{x}_i, \mathbf{x}_j)^2$. Consequently, the total transport cost from μ to ν is

$$\sum_{ij} \gamma_{ij} d(\mathbf{x}_i, \mathbf{x}_j)^2,$$

where d is the usual Euclidean distance between the location $\mathbf{x}_i$ and $\mathbf{x}_j$. The set of γ_{ij} coefficients is called the transport plan. It is a measure on product space of measures admitting μ and ν as margins. The optimal transport cost is obtained by minimizing this sum over all possible transport plans, i.e., all possible $\gamma_{ij} > 0$. This produces the Wasserstein distance

$$\mathcal{W}(\mu, \nu) = \left(\inf_{\gamma_{ij}} \sum_{ij} \gamma_{ij} d(\mathbf{x}_i, \mathbf{x}_j)^2 \right)^{1/2}. \quad (1)$$

Computing the right-hand side of Eq. (1) is a problem of minimization under constraints on the γ_{ij} coefficients, which have to be positive, and whose marginal sums equal μ_i and ν_j . This distance can be numerically computed by network simplex algorithms, coming from linear programming theory (see, e.g., Bazaraa et al., 2009). We refer to Appendix A for a general idea of the algorithm. Equation (1) is the discrete version of a more general formulation of the Wasserstein distance, whose properties are detailed by Villani (2003).

In our example (Fig. 1), we have $\mathcal{W}(\mu, \nu) = 1 \ll 3.27 = \mathcal{W}(\mu, \xi)$. Therefore, we can quantify with the Wasserstein distance that the cost of transforming μ into ν is lower than transforming μ into ξ . This result is closer to the physical intuition that a small shift is less costly than a mirror image and a rotation. Our next step is to apply the Wasserstein distance to differentiate between dynamical systems.

3 Inference on simple dynamical systems

3.1 Attractors and measure of a dynamical system

A dynamical system can be defined by the action of an ordinary differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}),$$

on a set of initial conditions (see, e.g., Guckenheimer and Holmes, 1983; Katok and Hasselblatt, 1997). Here $\mathbf{x}$ is a multivariate vector in the phase space and $\mathbf{v}(\mathbf{x})$ is a vector field that acts on $\mathbf{x}$. The properties of the ensemble of trajectories from all initial conditions define the dynamics of the system. They are entirely determined by $\mathbf{v}$.

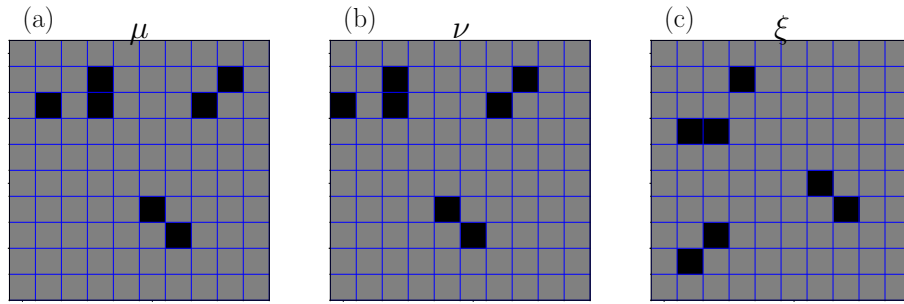


Figure 1. μ , ν and ξ are three examples of density of attractors. The black boxes have a measure of 1 and the grey a measure of 0. ν is a shift of μ , but ξ is very different to μ and ν . Finally, ν has no common point with μ , whereas ξ has one common point with μ . Finally, ν (resp., ξ) have no common (resp., one common) point with μ . The Euclidean distance between μ and ν (resp., ξ) is equal 3.74 (resp., 3.46), whereas the Wasserstein distance is equal to 1 between μ and ν , and 3.27 between μ and ξ .

For chaotic dynamical systems, trajectories $\mathbf{x}(t)$ emerging from almost all initial conditions converge to a unique object called an attractor, embedded in the phase space. Attractors define an invariant measure in phase space, which quantifies the weight of all trajectories of the dynamical system in sub-regions of the phase space. The measure of a sub-region of phase space is the probability of a trajectory of the system to go through the region. The invariance is characterized by the conservation of the volume by the dynamics of the system (Ruelle, 1989). The goal of this section is to estimate the distance between the empirical invariant measure of attractors.

We now focus on the Lorenz (1984) model, which is an idealized model of the Hadley circulation and its seasonality. The dynamics of this system is noted $\mathbf{v}(\mathbf{x})$, and, for a vector $\mathbf{x} = (x_1, x_2, x_3)$ given by

$$\mathbf{v}(\mathbf{x}) = \begin{pmatrix} -x_2^2 - x_3^2 - (x_1 - F)/4 \\ x_1x_2 - 4x_1x_3 - x_2 + 1 \\ x_1x_3 + 4x_1x_2 - x_3 \end{pmatrix}. \quad (2)$$

We propose to discriminate two attractors based on Eq. (2). A first attractor is generated with $F \equiv 11.5$ (noted Wi, for winter). A second attractor is generated with $F \equiv 7.5$ (noted Su, for summer). We choose those values and this terminology because F is interpreted as a seasonal cycle in Sect. 4, of length $\tau = 73$ units. Both systems have three variables (so the phase space is $\mathbb{R}^3$), are chaotic and yield a strange attractor. They are illustrated by two long trajectories in Fig. 2. To quantify the difference between the two attractors, it is first necessary to estimate the invariant measure of both attractors. We use the method of snapshot attractors (e.g., Romeiras et al., 1990; Chekroun et al., 2011) rather than considering one single long trajectory that could bias the sampling of some regions of the attractors, and requires the system to be ergodic. In the snapshot attractors, we draw N random initial conditions following a uniform distribution. All margins are independent. This approximates a Lebesgue measure in a cube that includes the attractors. We iterate the dynamics of the systems between $t_0 = 0$ and a long time multiple of τ .

Consistently with Drótos et al. (2015), we take $5\tau = 5 \times 73$ (i.e., 5 cycles, but we have checked that τ is enough). Both systems are dissipative outside of the attractors neighborhood, therefore all N trajectories collapse to the attractors after time 5τ and provide an efficient sampling of the invariant measure (Romeiras et al., 1990). After time 5τ , the set of N final points emerging of N initial conditions is called a snapshot attractor (see Algorithm 2). Snapshot attractors are special cases of pullback attractors (Chekroun et al., 2011). The latter class requires an integration between $-\infty$ and a desired final time. Equation (2) does not depend on time; therefore, the integration into Sect. 3.1 can be performed on any length intervals.

Algorithm 1 Simulation of a snapshot attractor with N initial conditions from the Lorenz 84 system.

Require: 5τ ($= 5 \times 73$) iteration time for convergence towards the attractor,
 N ($= 50, 100, 1000$) the number of points in the snapshot,
 $C = [-1, 3] \times [-3, 3] \times [-3, 3]$ a box that contains the attractor

Ensure: One snapshot with N points denoted $\{\mathbf{y}^i\} \in \mathbb{R}^3$ with $i = 1, \dots, N$

- 1: Draw uniformly N points $\mathbf{x}^1, \dots, \mathbf{x}^N$ in C
 - 2: **for** $i = 1, \dots, N$ **do**
 - 3: Integrate Eq. (2) between 0 and 5τ starting to $\mathbf{x}^i$. The ending point is $\mathbf{y}^i$. Integration is performed using the RK4 scheme with a time step of 0.005.
 - 4: **end for**
-

Then we compute the empirical measures associated with the snapshot attractors by discretizing the phase space (approximated by the box $[-1, 3] \times [-3, 3] \times [-3, 3]$) into cells of size $0.1 \times 0.1 \times 0.1$ (so $40 \times 60 \times 60 = A$ cells), and by counting the number of points of a snapshot attractor in each cell (see Algorithm 2). The empirical measure of the winter attractor (resp., summer) is noted μ^{Wi} (resp., μ^{Su}). They are sums of Dirac measures at each discrete cell. It is the equivalent of a multi-dimensional histogram of the attractor. We chose a bin length of 0.1 for the Lorenz attractor,

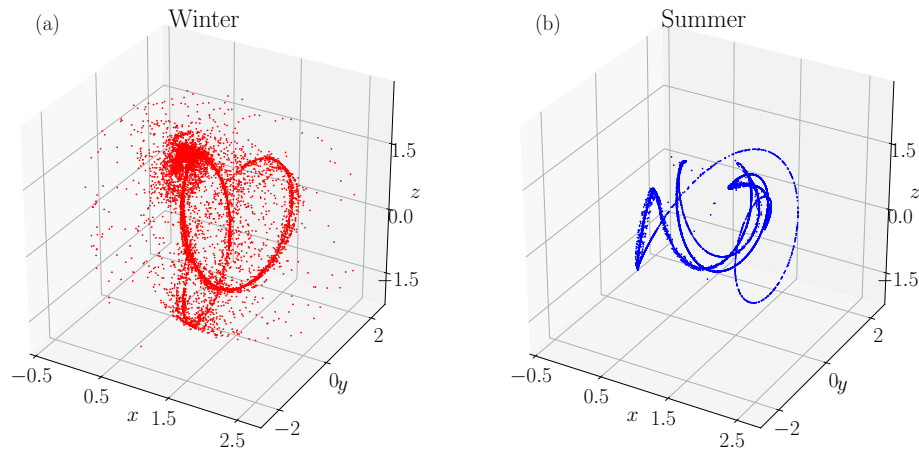


Figure 2. (a) Winter snapshot attractor of the Lorenz84 model. (b) Summer snapshot attractor of the Lorenz 84 model. Each of 10 000 points is the solution at time $5\tau = 5 \times 73$ of the Lorenz 84 equation (see Eq. 2), and constructed with a time step of integration of 0.005 using RK4 scheme.

which remains in a $[-1; 3] \times [-3; 3] \times [-3; 3]$ box. Therefore, $40 \times 60 \times 60$ bins cover the attractor. This number of bins is comparable to the number of grid cells that cover the North Atlantic region in the NCEP reanalysis (or most CMIP5 model simulations). This example refers to a few papers dealing with climate attractor properties (e.g., Corti et al., 1999; Faranda et al., 2017).

Algorithm 2 Determining the empirical invariant measure from simulated snapshot attractors.

Require: One snapshot attractor, $\{\mathbf{y}^i\}_{i=1,\dots,N}$, obtained from Algorithm 1, $[-1, 3] \times [-3, 3] \times [-3, 3]$, a large box containing the attractor, 0.1, the length of the edge of each cells to compute the histogram (so $40 \times 60 \times 60 = A$ cells)

Ensure: An approximated density measure, i.e. a sum of Dirac masses estimated from the number of points in each cell B_a with $a = 1, \dots, A$

$$\mu = \frac{1}{N} \sum_{a=1}^A \mu_a \delta_{B_a}$$

where δ_{B_a} is the Dirac measure around the cell B_a (equal to one if $\mathbf{x} \in B_a$ and zero otherwise) and $\mu_a \geq 0$ is the inferred mass. μ_a is not equal to 0 for a small numbers of boxes.

- 1: Divide the space into small gridded cell B_a of size $0.1 \times 0.1 \times 0.1$.
 - 2: **for all** cells B_a **do**
 - 3: $\mu_a \leftarrow (\text{number of } \mathbf{y}^i \text{ in } B_a)/N$
 - 4: **end for**
-

3.2 Protocol

The difference between the summer and winter attractors is evaluated by $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Su}})$ for different sample of size N . The probability distribution of Wasserstein distances is not known a priori for random measures. We first estimate the typical value of Wasserstein distances between identical at-

tractors in order to build a null hypothesis to be rejected if the distance is larger to some threshold. Therefore, we construct 50 winter (resp., summer) Lorenz 84 snapshot attractors, with empirical measure $\mu_k^{\text{Wi}}, k = 1, \dots, 50$ (resp., μ_k^{Su}), by drawing 50 sets of N random initial conditions, and applying Algorithms 2 and 2 between 0 and 5τ . By construction, $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{k'}^{\text{Wi}})$ should tend to 0 when N increase.

We detect a difference between the winter and summer of Lorenz 84 systems if

$$\begin{aligned} \mathcal{W}(\mu_k^{\text{Wi}}, \mu_{k'}^{\text{Wi}}) &\ll \mathcal{W}(\mu_k^{\text{Wi}}, \mu_{k'}^{\text{Su}}) \text{ and } \mathcal{W}(\mu_k^{\text{Su}}, \mu_{k'}^{\text{Su}}) \\ &\ll \mathcal{W}(\mu_k^{\text{Wi}}, \mu_{k'}^{\text{Su}}). \end{aligned}$$

This is quantified by a Kolmogorov–Smirnov (KS) test (Durbin, 1973; von Storch and Zwiers, 2001) between the distributions of $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{k'}^{\text{Wi}})$ (resp., $\mathcal{W}(\mu_k^{\text{Su}}, \mu_{k'}^{\text{Su}})$) and $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{k'}^{\text{Su}})$, in order to reject the null hypothesis that the probability distributions are equal. The KS test gives two values, the maximal difference between the cumulative distribution function of measures, and the “ p value”, which quantifies the probability to accept the null hypothesis. It is estimated by the Kolmogorov distribution (see Marsaglia et al., 2003). We choose to simulate 50 attractors of winter and 50 attractors of summer. We have $50 \times 50 = 2500$ different pairs between summer and winter. For the distances between the 50 attractors of the same season (summers or winters), we only consider $1 \leq (k, k') \leq 50$ pairs with $k < k'$. This means that we have 1225 distances for the winter or the summer. So we have at least 1000 distances per distribution. This is a reasonable sample size for a representative Kolmogorov–Smirnov test.

The estimation of the Wasserstein distance between attractors obviously depends on the number of available samples N of the dynamical systems on which the empirical measures are constructed. To explore the variability in the es-

timation of Wasserstein distances from finite observational sets, we sample its distances for three different values of N : $N = 50, 100$ and 1000 . We compute also one of each distance for $N = 10^6$. This later case represents a quasi-perfect estimation of the distance and we consider it as our benchmark for comparison.

The complete procedure to obtain an empirical probability distribution of Wasserstein distances, depending on the sample size N , is summarized in Algorithm 3.2.

Algorithm 3 Estimation of $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{\tilde{k}}^{\text{Wi}})$, $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{\tilde{k}}^{\text{Su}})$ and $\mathcal{W}(\mu_k^{\text{Su}}, \mu_{\tilde{k}}^{\text{Su}})$.

Require: N ($= 50, 100, 1000$) the number of points in snapshots

Ensure: 1225 independent estimates of the two Wasserstein distances $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Wi}})$ and $\mathcal{W}(\mu^{\text{Su}}, \mu^{\text{Su}})$ where the first differentiates two winter and the second two summer of Lorenz 84 snapshot attractors.

2500 independent estimates of the Wasserstein distances $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Su}})$ which compares winter and summer of Lorenz84 snapshot attractors.

- 1: Use Algorithm 1 to simulate fifty winter and fifty summer Lorenz 84 snapshot attractors, denoted Wi_k and Su_k . Each snapshot attractor has N points.
- 2: Use Algorithm 2 to transform each Wi_k (resp. Su_k) into measures, noted μ_k^{Wi} (resp. μ_k^{Su}).
- 3: Compute the Wasserstein distances $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{\tilde{k}}^{\text{Wi}})$ (resp. $\mathcal{W}(\mu_k^{\text{Su}}, \mu_{\tilde{k}}^{\text{Su}})$) for $k \neq \tilde{k}$ (see Appendix A). Thus, $\frac{50 \times (50+1)}{2} = 1225$ distances are stored.
- 4: Compute the Wasserstein distances $\mathcal{W}(\mu_k^{\text{Wi}}, \mu_{\tilde{k}}^{\text{Su}})$ for all $k, \tilde{k}$. Thus, $50^2 = 2500$ distances are stored.

3.3 Estimation

The probability distributions of the Wasserstein distance for $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Wi}})$ (resp., Su) and $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Su}})$ are summarized in Fig. 3a by box-and-whisker plots (boxplots: Chambers et al., 1983)

The distribution of the distances between winter (resp., summer) snapshot attractors decreases to 0 (the expected asymptotic value) when N increases (white and grey box plots). We explain the relatively high values of the distance when $N = 50$ by the fact that few cells of the discrete measure are filled when N is low so that the transport plan is not zero. By increasing N , all cells tend to be sampled so that the transport plans are less affected by sampling issues, and the cost of the transport decreases on average.

The distance between winter and summer attractors (black box plots) decrease with N and converge to the “true” value that is estimated with $N = 10^6$. The explanation is similar: if the measures of the snapshot attractors are estimated with low N , the “circles” composing the attractors are akin. Increasing the number of initial conditions N essentially allows one to differentiate the dynamics of the two attractors. We note that the distribution of distances

Table 1. Kolmogorov–Smirnov test applied between distribution $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Wi}})$ (resp., $\mathcal{W}(\mu^{\text{Su}}, \mu^{\text{Su}})$) and $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Su}})$ in left (resp., right) box on snapshots of size N .

N		50		100		1000	
Wasserstein distance	KS test	0.98	0.99	1	1	1	1
	p value	0	0	0	0	0	0
Euclidean distance	KS test	0.71	0.97	0.77	1	0.47	1
	p value	0	0	0	0	0	0

for $N = 1000$ is very close to the one with 10^6 . This highlights a rather quick convergence of the Wasserstein distance $\mathcal{W}$. Figure 3a shows a good discrimination between null hypothesis distances $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Wi}})$ and the distances $\mathcal{W}(\mu^{\text{Wi}}, \mu^{\text{Su}})$, even for $N = 50$. This discrimination is confirmed by Kolmogorov–Smirnov tests reported in Table 1. The null hypothesis of identical attractors is rejected with probability one, even for $N = 50$. Finally, the variance of distance decrease with N . Indeed, attractors are independent of the initial condition; thus, the variability is due to a low N . This property is shown by the Wasserstein distance.

This protocol was also applied for bin sizes of 0.05, 0.2 and 1.0. For 0.05 and 0.2, the maximal variation of median (resp., standard deviation) of Wasserstein distances is 0.03 (resp., 0.01); therefore, we have the distributions are indistinguishable in practice. For a bin size of 1.0, the maximal increase of the median is 0.22, but the difference with the median of winter against summer is at least equal to 0.3.

For illustration purposes, we compute Euclidean distances between the same snapshot attractors (Fig. 3b). The distances are normalized by $\sqrt{2}$, the maximum value being reached for two measures without common points. The distances between winter (resp., summer) snapshot attractors decrease as N increase (white and grey box plots). The distances between winter and summer snapshot attractors also decrease to the “perfect” estimate with $N = 10^6$ (black box plots), but the convergence to the limit is far from being reached with $N = 1000$. The difference between winter and summer cannot be detected for all N . For $N = 50$ and 100 , the distances between winter are greater than distances between winter and summer. For $N = 1000$, the Kolmogorov–Smirnov test (Table 1) shows the impossibility to reject the null hypothesis without ambiguity. Moreover, the variability is small and constant with N , which is incompatible with the high variability due to a low N . Therefore, the Euclidean distance might not be very useful to distinguish the dynamics in real world systems with few observations.

3.4 Inference with reduced information

In this section, we test whether it is possible to differentiate attractors if only partial information is available. This can happen if one or more variables of the system are

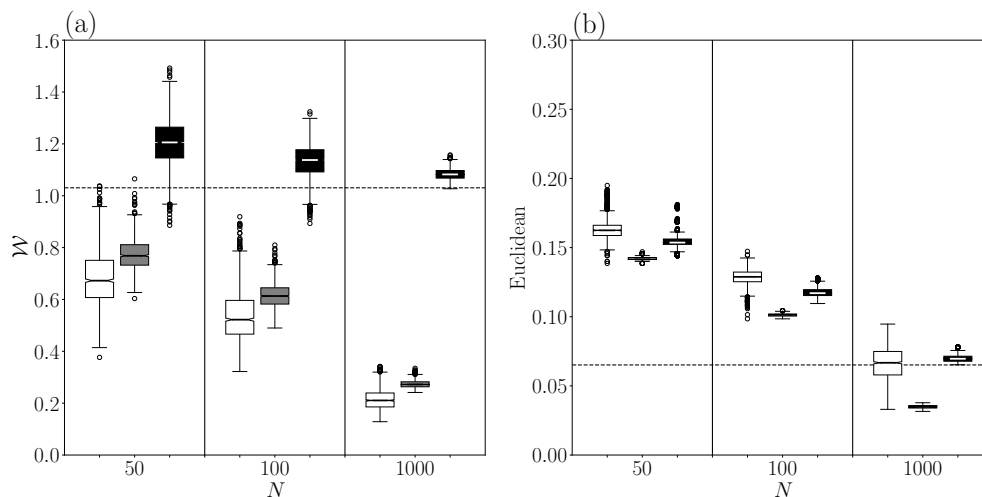


Figure 3. box plots of distances computed using the Wasserstein distance (a) and the Euclidean one (b). White box plots differentiate between two winter snapshots. Grey box plots differentiates two summer snapshots. Black box plots compare winter and summer snapshots. Dotted lines represent the distance between winter and summer attractors with $N = 10^6$ points.

omitted (projection onto the remaining variables) or if variables are censored (truncation of the values of a variable), or a combination of both. The motivation in atmospheric sciences is that the underlying dynamical system is defined in three spatial dimensions (on the sphere), and that observables of the attractor of this system are generally obtained over a limited area (censoring of the rest of the globe) and a fixed pressure level (projection).

It has been proven that a sequence of observables of a dynamical system convey the same dynamics as the whole system (Packard et al., 1980; Takens, 1981; Mañé, 1981). Therefore, it is meaningful to compare the distances between projected or truncated attractors.

For the Lorenz 84 attractors, a first reduction of information is performed by projecting the systems onto their (x_1, x_2) variables (design P). We hence compute the distances (Wasserstein and Euclidean) between attractors from the variables (x_1, x_2) and discard the information on x_3 . The second reduction consists in truncating negative values of the variable x_1 (design T). Thus, we only consider the values of (x_1, x_2, x_3) when $x_1 \geq 1$. The third reduction of information is a combination of projection onto the (x_1, x_2) variables and truncating negative values of x_1 (design $T + P$). These transformations are illustrated in Fig. 4a–c. Those three transformations create observables of the underlying attractors. We shall call them “observed attractors”, with designs P (projection), T (truncation) and $P + T$ (both). The distribution of the distance between observed attractors is shown in Fig. 4d–i, for the two distances and each of the information reduction design (P , T and $P + T$).

The Wasserstein distance distribution (Fig. 4a–c) shows a clear discrimination between winter and summer observed attractors, for all values of N . This is reflected in the Kolmogorov–Smirnov test: all test values are greater than

0.97, except for winter with $N = 50$, this is 0.84. All p value are equal to 0. The estimated distances between winter and summer observed attractors is always smaller than the idealized one (obtained on the full attractors) and shows little dependence on the number of points N . This is explained by the fact that the projection on a subspace of dimension 2 implies a reduction of transportation cost. Moreover, some points that are very far in the winter attractor become close to each other in the projection P . Overall, the reduction of information decreases the discriminating power of the Wasserstein distance, but the results are still significant for number of points N as small as 50.

The same experiment is conducted with the Euclidean distance (Fig. 4g–i). Contradicting the intuition, it clearly discriminates between winter and summer for all designs P , T and $P + T$. Comparing the full attractors (Fig. 2) and Fig. 4a–c, we see that some points very far, become close in the same boxes of the estimated measure. This is reflected by a gain of variance, which decreases with N . Finally, we need $N = 1000$ to have the distribution between winters (resp., summers) lower than the idealized distance.

We conclude that the Wasserstein distance has a high capacity of discriminating different attractors coming from this dynamical system, even with a partial information. It is particularly promising in atmospheric sciences, where analyses are performed on truncated variables (e.g., a surface field on a limited area: transformation T) and/or on only one atmospheric field (e.g., geopotential height, omitting other variables: transformation P).

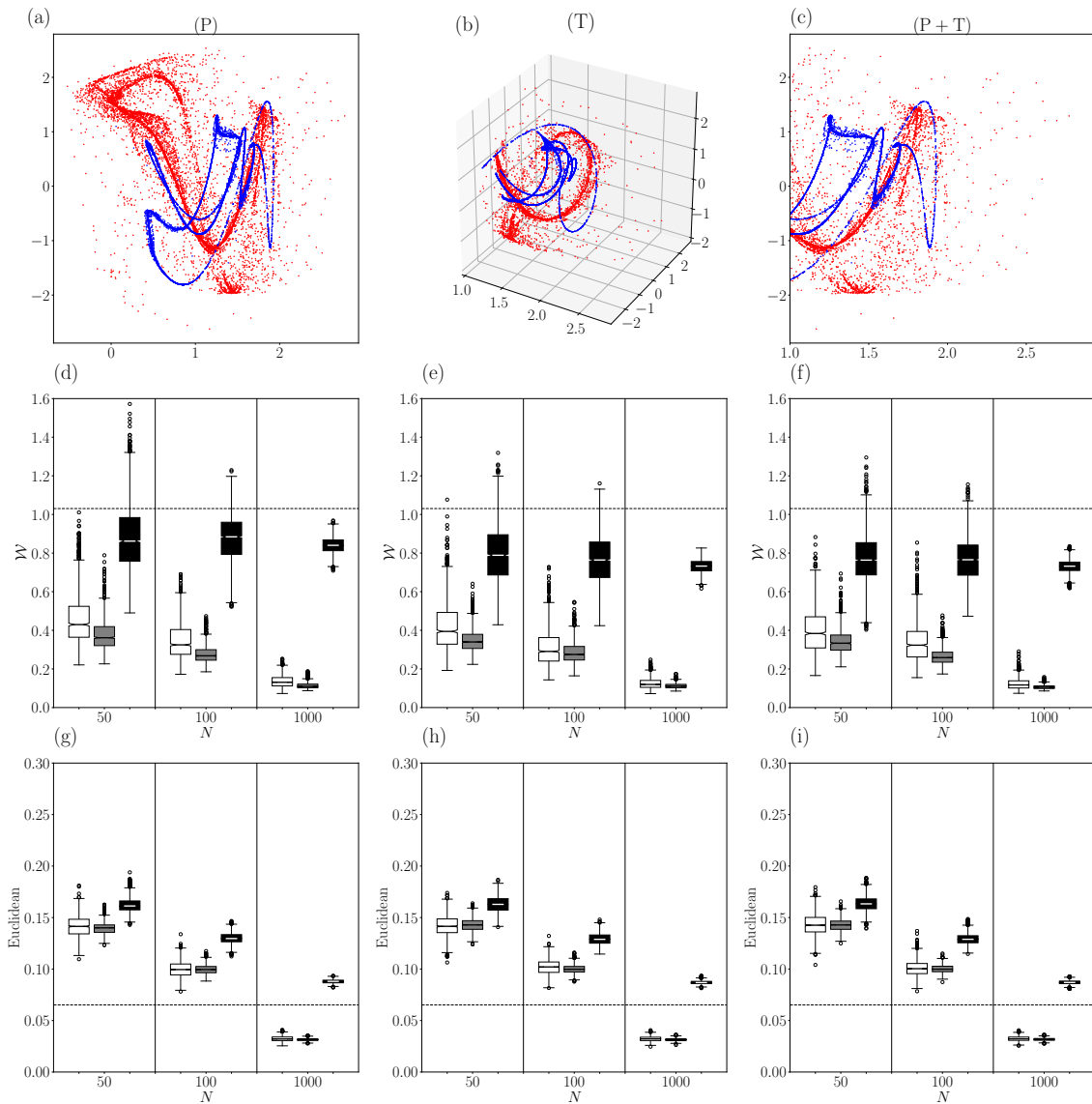


Figure 4. (a) Projection on axes (x, y) of the winter snapshot (red) and summer snapshot (blue). (b) Truncature at $x_1 \geq 1$ of the winter and summer. (c) Combination of (a) and (b) (see designs (T) , (P) and $(T + P)$ in Sect. 3.4). (d–f) box plots of distances between design T , P and $T + P$ computed with the Wasserstein distance. (g–i) box plots of distances between design T , P and $T + P$ computed with the Euclidean distance. (d–i) White box plots differentiate between two winter snapshots. Grey box plots differentiates two summer snapshots. Black box plots compare winter and summer snapshots. Dotted lines represent the distance between winter and summer attractors (without designs) with $N = 10^6$ points.

4 Time-varying dynamical system

We now focus on a time-varying dynamical system that mimics variability around a seasonal cycle, and a monotonic forcing that plays after a triggering time T . Such a system defines a snapshot attractor at all times t . We want to measure how snapshot attractors evolve after time T , when the forcing increases (we mean the forcing modifies more and more the attractors). The constant F in the system (2) is now a function of time, and include a seasonal cycle and a forcing.

$$F(t) = \underbrace{9.5 + 2 \sin\left(\frac{2\pi t}{73}\right)}_{\text{seasonality}} - \underbrace{2 \frac{t-T}{T} \mathbf{1}_{\{t>T\}}}_{\text{monotonic forcing}},$$

$$T = 100 \times 73. \quad (3)$$

The snapshot attractors of this system were investigated by Drótos et al. (2015), who performed an analysis of the mean and variance of each coordinate to detect the forcing F after time $t > T$.

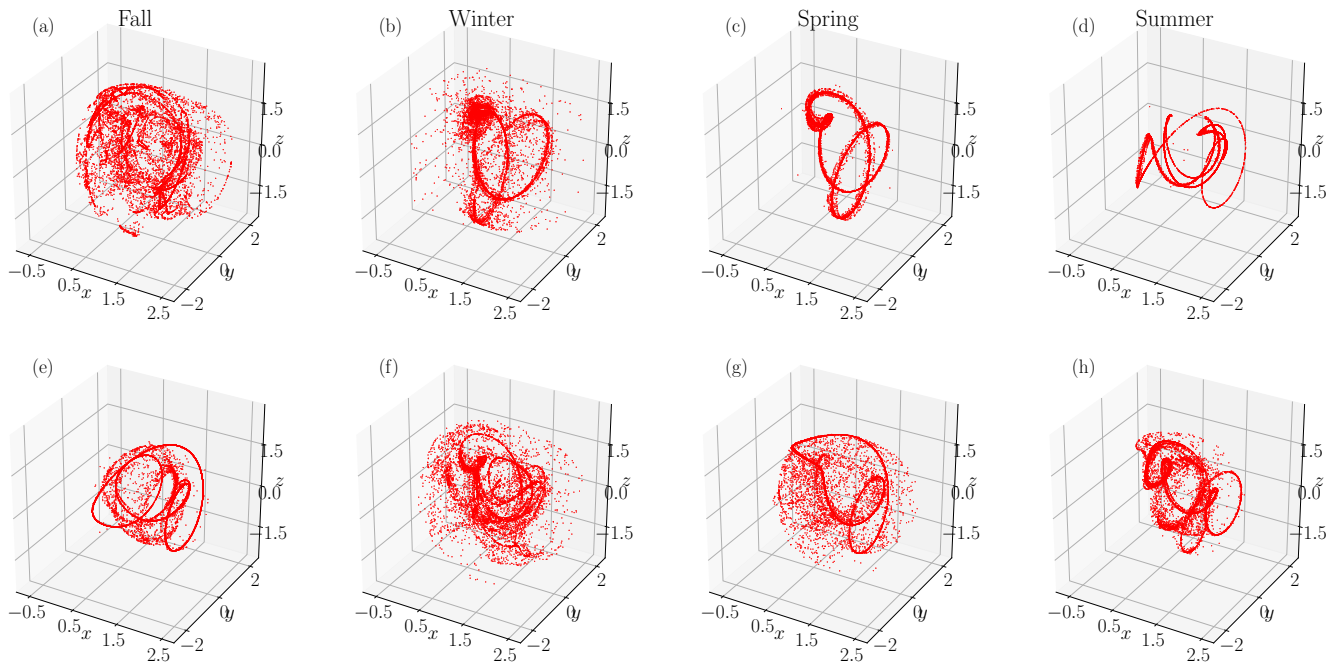


Figure 5. Snapshots of 10 000 points from the Lorenz 84 defined by Eq. (2). **(a–d)** The four seasons at time $t = 0$ years, $t = 0.25$ years, $t = 0.5$ years and $t = 0.75$ years. **(e–h)** The same seasons, but after the triggering of the linear forcing, during year 180.

Such a coupled behavior is present in most regional temperature time series at the decadal or centennial scales. The periodic part of the forcing F in Eq. (3) allows one to divide the year into seasons of the system (Lorenz, 1984; Drótos et al., 2015). To be consistent with Lorenz (1984) and Drótos et al. (2015), we consider that there are $73 = \tau$ time units in 1 year. We emphasize that a time unit is not analogous to a “day”, but corresponds to a typical variability timescale in the non-forced chaotic system in Eq. (2). We follow Drótos et al. (2015) and define the fall equinox at $t = 0$ year or $t \bmod 73 = 0$ year. Then, winter solstices correspond to $t \bmod 73 = 0.25$ year, spring equinoxes correspond to $t \bmod 73 = 0.5$ year and summer solstices correspond to $t \bmod 73 = 0.75$ year. This time-dependent system produces a different snapshot attractor at each time step. We focus on the snapshot attractors that occur at each equinox/solstice. These parameters are coherent with winter and summer defined in Sect. 3.

In this section, we want to quantify the change of the whole dynamics of the ensemble of snapshot attractors with the Wasserstein distance, and assess the detectability of changes from small numbers of observations.

4.1 Protocol

We compute snapshot attractors for each time step, for t between 0 to 200 years. To have the convergence of trajectories on attractor, we draw N initial conditions in a cube (see Algorithm 2) and perform a first integration during $5\tau = 5 \times 73$ time unit (i.e., 5 years). The attractor obtained is considered

at $t = 0$. As previous Section, $N = 50, 100, 1000$. We generate also a sequence with $N = 10000$ as benchmark. The N trajectories of the system in Eq. (2) are computed with a Runge–Kutta scheme of order 4 (RK4).

The empirical measure of the snapshot attractors is estimated at each time step t with the Algorithm 2. We then compute the Wasserstein distance between those time-varying snapshot attractors, and four reference seasonal snapshot attractors obtained for $t = 0, 18.25, 36.5$ and 54.75 , during the first year. The four reference seasonal snapshot attractors are shown in Fig. 5a–d, with $N = 10\,000$ points. For illustration purposes, the snapshot attractors corresponding to the same seasons, but at year 180, after the monotonous forcing is triggered. It is obvious from Fig. 5e–f that the forcing affects each of the seasonal attractors.

The yearly averages of the distances to the four reference attractors are shown in Fig. 6. We detect the change point, with a trend, after $t = 100$ years. Therefore, the detection of the forcing effect on the dynamics of the Lorenz84 system is rather immediate, with a lag < 10 years.

We find that the variability of the distance variations decrease with the number N of points to estimate the snapshot attractors, although it does not seem to affect the detection of the change point. Relatively low values of N show a bias of the distance, which is even higher for lower values of N . The mean values of the attractor distance distributions is similar if $N \geq 100$. This sets a lower bound for the number of points to estimate the measure of snapshot attractors.

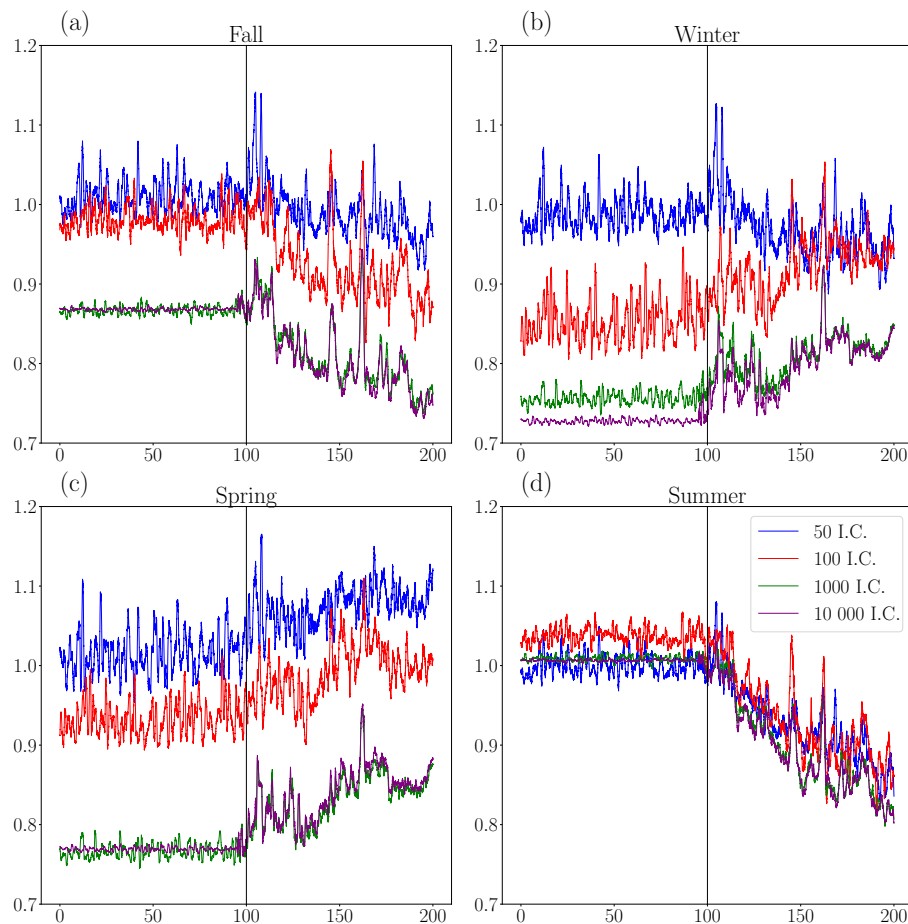


Figure 6. Yearly averages Wasserstein distance between the reference attractor before forcing, and all other attractors. The x axis is the time, the y axis the estimated Wasserstein distance. The blue (resp., red, green and purple) is the numbers of initial conditions (I.C.) for $N = 50$ (resp., 100, 1000 and 10 000). The vertical black line represents the instant when the linear trend is triggered in the forcing $F(t)$.

In this example, the distances of the snapshot attractors to winter and spring reference attractors increase with time after $t = 100$ years. Conversely, the distance to fall and summer reference attractors decrease with time. We interpret this as a shift of all snapshot attractors toward “hot” conditions.

Those results are consistent with those of Drótos et al. (2015). The main practical value of our approach is that the number of points that is needed to sample snapshot attractors can be as low as $N = 100$, rather than $N = 10^6$, which is generally not available.

The same experiment is conducted with the Euclidean distance. For $N = 50$ and 100, the maximal difference of the mean (resp., standard deviation) between the period before and after the forcing is 0.002 (resp., 0.002), whereas the mean is 0.2 (resp., 0.002). The distance distribution is approximately Gaussian, which implies that $\approx 70\%$ of values are within one standard deviation around the mean. Therefore, with the values we obtain, we cannot detect the forcing. For $N = 1000$ and 10000 the mean is 0.08, and its maximal modification is 0.004. The standard deviation is mul-

tiplied by a factor 20 (0.0002 becomes 0.005). Even if the forcing is detected, the trajectories of distances are not representative of a linear increasing forcing.

5 Conclusions

The Wasserstein distance appears to be efficient to measure changes in the dynamics in time evolving systems even with a relatively low number of points (e.g., $N = 100$). This discrimination is still powerful when only partial information on the attractor is available (truncation and/or projection). We made the assumption that the system we investigate yields an attractor, and the Wasserstein distance determines changes in the invariant measure of the attractors. This builds an interesting bridge between dynamical systems and optimal transport. A theoretical justification for this bridge is recalled in Appendix B. A caveat of the approach we present here is that we do not give an interpretation of the Wasserstein distance in terms of qualitative dynamical changes (e.g., changes in local dimensions Faranda et al., 2017). Villani (2003, Chap-

ter 9) provides links between the Wasserstein distance and entropy, but they are hard to interpret and infer for the problem we tried to tackle (measure a change in a strange attractor).

The other caveat of this approach is its computational cost. The minimization of the cost function, constrained by the estimated measures, has to be implemented by network simplex algorithms (Bazaraa et al., 2009; Boyd and Vandenberghe, 2004; Dantzig et al., 1955; Gottschlich and Dominic, 2014). Those algorithms are computationally expensive, but applicable, as shown with the Lorenz 84 model (200 000 distances computed in 60 h on a 12 core computer server).

The next research challenge is to adapt this method to climate model simulations (e.g. from the Coupled Model Intercomparison Project (Taylor et al., 2012)). The Wasserstein distance could be computed to discriminate between the atmospheric attractors from control, historical and scenario runs.

Data availability. The Lorenz model simulations data files can be obtained upon request to the first author.

Appendix A: Computation of the Wasserstein distance

We just give here the general idea to compute Wasserstein distance with the network simplex algorithm. We want to transport the measure μ to ν , can be written

$$\mu = \sum_{i=1}^n \mu_i \delta_{\mathbf{x}_i}, \quad \nu = \sum_{j=1}^p \nu_j \delta_{\mathbf{y}_j}.$$

The Wasserstein distance is given by minimizing over γ_{ij} (the mass transported from $\mathbf{x}_i$ to $\mathbf{y}_j$) the cost function

$$\sum_{ij} \gamma_{ij} d(\mathbf{x}_i, \mathbf{y}_j)^2.$$

Consequently, we have the following linear constraints:

$$\mu_i = \sum_{j=1}^p \gamma_{ij}, \quad \nu_j = \sum_{i=1}^n \gamma_{ij}, \quad \gamma_{ij} \geq 0$$

These constraints define a polyhedral convex set in the space of γ_{ij} . The solutions of all constraints are the extremal point of the polyhedra, and the $\mathcal{W}$ distance is one of its minima. The network simplex algorithm runs in two part:

1. Finding a first extremal point.
2. Iterate over the face of polyhedra (the simplex) until the minimal solution is reached.

Because the number of extremal points increases exponentially with the size of data, this algorithm has an exponential complexity. But, in practice the iterations over simplex are made in the direction of an optimal solution. Thus, it has been found that the complexity of the algorithm is polynomial in practice. Currently, we use a C++ implementation of the R-package transport (Baehre et al., 2016), using the methodology described in Gottschlich and Dominic (2014). We have also tested entropy regularization (Cuturi, 2013). This algorithm crosses the polyhedra until the optimal solution, but it requires a parameter changing for each distance. We preferred to use the network simplex method, which works all the time.

Appendix B: Theoretical justification

Besides the simulations studied in the previous sections, it is possible to theoretically justify the use of the Wasserstein distance for non-autonomous dynamical systems. Any dynamical system defined from an ordinary differential equation, say $\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t)$, is formally equivalent (e.g., see Villani, 2003; Evans, 2010) to the partial differential equation of a transport of the density of trajectories of the associated dynamical system, say $\frac{\partial \rho_t}{\partial t} + \langle \nabla, \rho_t \mathbf{v} \rangle = 0$. In other words, the variations between t_0 and t_1 of the time-varying attractor in $\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t)$ can be determined by the transport of the measure of the attractors by the dynamics $\mathbf{v}$. If μ_t denotes the density distribution of ρ_t (i.e., $\mu_t(A) = \int_A \rho_t(\mathbf{x}) d\mathbf{x}$), then the Wasserstein distance between μ_{t_0} and μ_{t_1} for attractors in dimension d is given by the Benamou–Brenier theorem (Benamou and Brenier, 1998)

$$\mathcal{W}(\mu_{t_0}, \mu_{t_1})^2 = \frac{1}{t_1 - t_0} \inf_{(\tilde{\rho}_t, \tilde{\mathbf{v}})} \int_{t_0}^{t_1} \int_{\mathbb{R}^d} \tilde{\rho}_t(\mathbf{x}) |\tilde{\mathbf{v}}(\mathbf{x}, t)|^2 dt d\mathbf{x}.$$

The minimization is done over all vector fields $\tilde{\mathbf{v}}$ and all sequences of density $\tilde{\rho}_t$ following $\tilde{\mathbf{v}}$ such that $\tilde{\rho}_{t_0} = \rho_{t_0}$ and $\tilde{\rho}_{t_1} = \rho_{t_1}$. This theorem connects the dynamical systems theory with the optimal transport theory. Therefore, the Wasserstein distance between two snapshot attractors of a time-varying dynamical system is linked to the energy ($\mathbf{v}$ is homogeneous to a velocity) of the system that transforms one attractor into the other. If the dynamics $\mathbf{v}$ is unknown and only simulations are available, this theorem allows (in principle) to infer $\mathbf{v}$ from the simulations because the optimum path going from the snapshot attractor at t_0 to t_1 is achieved by the actual dynamics $\mathbf{v}$.

Author contributions. YR conducted the computations and produced the figures. The idea of the experiments were formulated by the three authors. The three authors contributed to writing the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work is supported by ERC grant no. 338965-A2C2. We thank Davide Faranda and Ara Arakelian for useful discussions.

Edited by: Vicente Perez-Munuzuri

Reviewed by: Valerio Lucarini and one anonymous referee

References

- Baehre, D. S., Bjoern, C., and Gottschlich, C.: transport: Optimal Transport in Various Forms, R package version 0.7-3, available at: <https://CRAN.R-project.org/package=transport> (last access: 20 July 2017), 2016.
- Bazaraa, M. S., Jarvis, J. J., and Sherali, H. D.: Linear Programming and Network Flows, John Wiley & Sons, 4 Edn., 2009.
- Benamou, J.-D. and Brenier, Y.: The Optimal Time-Continuous Mass Transport Problem and its Augmented Lagrangian Numerical Resolution, Research Report RR-3356, INRIA, available at: <https://hal.inria.fr/inria-00073333> (last access: 20 July 2017), 1998.
- Boyd, S. and Vandenberghe, L.: Convex Optimization, Cambridge University Press, 2004.
- Cassou, C. and Cattiaux, J.: Disruption of the European climate seasonal clock in a warming world, *Nature Climate Change*, 6, 589–594, <https://doi.org/10.1038/nclimate2969>, 2016.
- Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A.: Graphical Methods for Data Analysis, Wadsworth, 1983.
- Chekroun, M. D., Simonnet, E., and Ghil, M.: Stochastic climate dynamics: Random attractors and time-dependent invariant measures, *Physica D*, 240, 1685–1700, <https://doi.org/10.1016/j.physd.2011.06.005>, 2011.
- Corti S., Molteni F., and Palmer T. N.: Signature of recent climate change in frequencies of natural atmospheric circulation regimes, *Nature*, 398, 799–802, <https://doi.org/10.1038/19745>, 1999.
- Cuturi, M.: Sinkhorn Distances: Lightspeed Computation of Optimal Transport, in: *Advances in Neural Information Processing Systems 26*, edited by: Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., 2292–2300, Curran Associates, Inc., available at: <http://papers.nips.cc/paper/4927-sinkhorn-distances-lightspeed-computation-of-optimal-transport.pdf> (last access: 20 July 2017), 2013.
- Dantzig, G. B., Orden, A., and Wolfe, P.: The generalized simplex method for minimizing a linear form under linear inequality restraints, *Pac. J. Math.*, 5, 183–195, 1955.
- Drótos, G., Bódi, T., and Tél, T.: Probabilistic concepts in a changing climate: a snapshot attractor picture, *J. Climate*, 28, 3275–3288, <https://doi.org/10.1175/JCLI-D-14-00459.1>, 2015.
- Durbin, J.: *Distribution Theory for Tests Based on Sample Distribution Function*, 9, Siam, 1973.
- Eckmann, J.-P. and Ruelle, D.: Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.*, 57, 617–656, 1985.
- Evans, L.: *Partial Differential Equations*, 19, Graduate studies in mathematics, Am. Math. Soc., 2010.
- Faranda D., Messori G., and Yiou P.: Dynamical proxies of North Atlantic predictability and extremes, *Sci. Rep.*, 7, 41278, <https://doi.org/10.1038/srep41278>, 2017.
- Ghil, M.: Climate Change: Multidecadal and Beyond, *A Mathematical Theory of Climate Sensitivity or, How to Deal With Both Anthropogenic Forcing and Natural Variability?*, World Scientific Publishing Co. Pte. Ltd., 6, 31–52, 2015.
- Ghil, M. and Childress, S.: *Topics in geophysical fluid dynamics: atmospheric dynamics, dynamo theory, and climate dynamics*, Springer-Verlag, New York, 1987.
- Gottschlich, C. and Dominic, S.: The shortlist method for fast computation of the Earth Mover's Distance and finding optimal solutions to transportation problems, *PLoS ONE*, 9, 1–10, <https://doi.org/10.1371/journal.pone.0110214>, 2014.
- Guckenheimer, J. and Holmes, P. J.: *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, 42, Applied Mathematical Sciences, Springer-Verlag, New York, 1983.
- Katok, A. and Hasselblatt, B.: *Introduction to the Modern Theory of Dynamical Systems*, 54, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1997.
- Lorenz, E. N.: Deterministic nonperiodic flow, *J. Atmos. Sci.*, 20, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2), 1963.
- Lorenz, E. N.: Irregularity: a fundamental property of the atmosphere, *Tellus A*, 36, 98–110, 1984.
- Lucarini, V., Ragone, F., and Lunkeit, F.: Predicting Climate Change Using Response Theory: Global Averages and Spatial Patterns, *J. Stat. Phys.*, 166, 1036–1064, <https://doi.org/10.1007/s10955-016-1506-z>, 2017.
- Mañé, R.: On the dimension of the compact invariant sets of certain non-linear maps, in: *Dynamical Systems and Turbulence*, Warwick 1980, v 898, Lecture Notes in Mathematics, Springer, New York, 230–242, 1981.
- Mané, R.: *Ergodic Theory and Differentiable Dynamics*, 8, Springer Science & Business Media, 2012.
- Marsaglia, G., Tsang, W. W., and Wang, J.: Evaluating Kolmogorov's Distribution, *J. Stat. Softw.*, 8, 1–4, <https://doi.org/10.18637/jss.v008.i18>, 2003.
- Monge, G.: Mémoire sur la théorie des déblais et des remblais, *Histoire de l'Académie Royale des Sciences*, 666–704, 1781.
- Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S.: Geometry from a time series, *Phys. Rev. Lett.*, 45, 712–716, 1980.
- Pierini, S., Ghil, M., and Chekroun, M. D.: Exploring the pullback attractors of a low-order quasigeostrophic ocean model: the deterministic case, *J. Climate*, 29, 4185–4202, <https://doi.org/10.1175/JCLI-D-15-0848.1>, 2016.
- Romeiras, F. J., Grebogi, C., and Ott, E.: Multifractal properties of snapshot attractors of random maps, *Phys. Rev. A*, 41, 784–799, <https://doi.org/10.1103/PhysRevA.41.784>, 1990.
- Ruelle, D.: *Chaotic Evolution and Strange Attractors*, Lezioni Lincee, Cambridge University Press, <https://doi.org/10.1017/CBO9780511608773>, 1989.

- Ruelle, D.: A review of linear response theory for general differentiable dynamical systems, *Nonlinearity*, 22, 855–870, 2009.
- Takens, F.: Detecting strange attractors in turbulence, in: *Dynamical Systems and Turbulence*, edited by: Rand, D. A. and Young, L.-S., 898, *Lecture Notes in Mathematics*, Springer, Berlin, 366–381, 1981.
- Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, *B. Am. Meteorol. Soc.*, 93, 485–498, 2012.
- Villani, C.: *Topics in Optimal Transportation*, 58, American Mathematical Society, 2003.
- von Storch, H. and Zwiers, F. W.: *Statistical Analysis in Climate Research*, Cambridge University Press, Cambridge, 2001.
- Vrac, M., Vaittinada Ayar, P., and Yiou, P.: Trends and variability of seasonal weather regimes, *Int. J. Climatol.*, 34, 472–480, <https://doi.org/10.1002/joc.3700>, 2014.

3.3 Compléments : Détection des scénarios RCP dans CMIP5

Je vais appliquer la méthodologie de ROBIN et al. (2017) au modèle de l'IPSL (DUFRESNE et al., 2013), pour les températures et les précipitations dans l'hémisphère nord (i.e. 30 à 90°N) sur la période 1850-2100. Je propose de mesurer l'évolution d'une période de référence, la période 1850-1855. En prenant la distance de Wasserstein entre cette période et toutes les fenêtres de 5 ans, l'altération du système climatique est mesurée. Pour cela, je vais travailler avec plusieurs simulations.

La première simulation est pré-industrielle, dite de *contrôle*. Dans cette simulation, la révolution industrielle n'a pas eu lieu, et le modèle est stationnaire. Nous nous attendons donc à trouver des distances de Wasserstein qui ne changent pas plus que la variabilité du système climatique.

Les quatre scénarios suivants seront aussi examinés : RCP26 (VUUREN et al., 2007), RCP45 (CLARKE et al., 2007 ; SMITH et WIGLEY, 2006 ; WISE et al., 2009), RCP60 (FUJINO et al., 2006 ; HIJIOKA et al., 2008) et RCP85 (RIAHI et al., 2007) (pour les précipitations, seul les RCP45 et RCP85 sont disponibles). Ces scénarios sont des projections du forçage radiatif, exprimé en W/m^2 sur la période 2005-2100. Ils sont complétés par des données historiques sur la période 1850-2005. La Figure 3.1 montre les valeurs correspondantes entre 1850 et 2100, tirées de la base de données des RCPs. Nous voyons que ces scénarios sont un forçage de plus en plus fort du système climatique. L'un des objectifs est de retrouver ces courbes dans la température et la précipitation avec la distance de Wasserstein.

Pour ces deux variables, je vais appliquer des variantes de la méthodologie présentée dans (ROBIN et al., 2017). Chaque point de grille sera traité indépendamment, et la moyenne des distances de Wasserstein sur tous les points de grille entre la fenêtre glissante et la période 1850-1855 va être calculée. La simulation de contrôle va permettre de mesurer la variabilité climatique et de vérifier que la période de référence est raisonnable. Je vais utiliser trois références différentes :

1. La première référence contiendra tous les jours de 1850-1855,
2. La seconde référence contiendra uniquement les étés,
3. La troisième référence contiendra uniquement les hivers.

Pour la fenêtre glissante de 5 ans, tous les jours des cinq années seront utilisés.

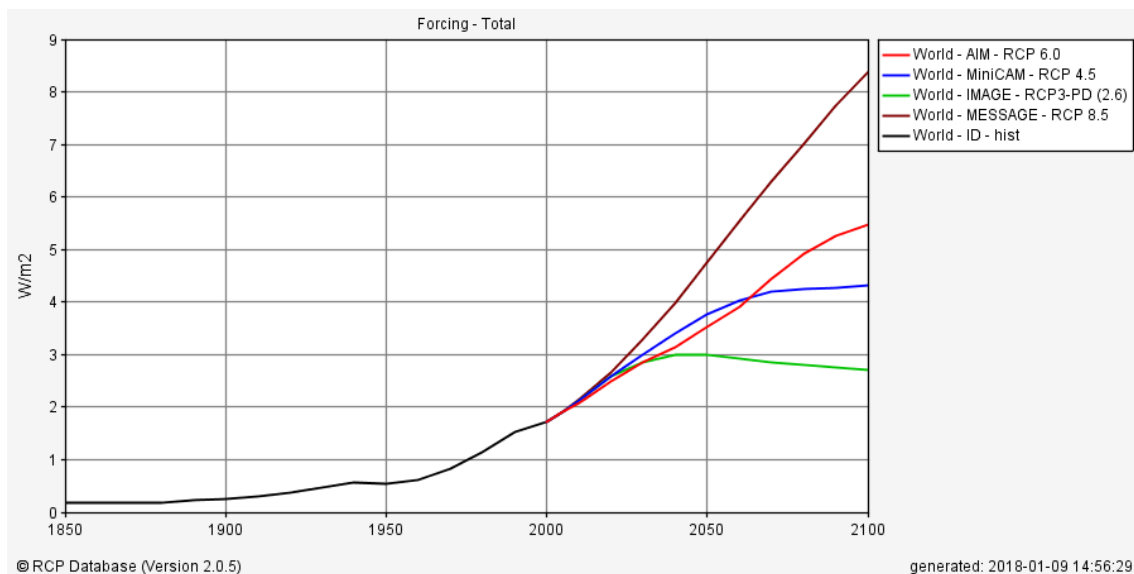


FIGURE 3.1 – Scénario RCP des forçages radiatifs entre 1850 et 2100.

3.3.1 Application aux températures

Les distances sont représentées sur la Figure 3.2. Pour les trois tests, la simulation de contrôle est stationnaire, avec une faible variabilité se confondant avec la variabilité climatique. Pour la Figure 3.2a, les distances de Wasserstein sont très semblables aux scénarios RCP. L'ordre et l'intensité sont bien conservés. La partie changement climatique des données historiques est aussi bien détectée, et nettement supérieure à la simulation de contrôle. Ces constatations sont aussi valables pour les figures 3.2b-c. De plus, pour la Figure 3.2b, le système climatique s'approche des étés, indiquant un réchauffement. Pour la Figure 3.2b, le système climatique s'éloigne des hivers, indiquant là aussi un réchauffement.

3.3.2 Application aux précipitations

Pour les précipitations, seuls les scénarios RCP45 et RCP85 étaient disponibles. Comme pour les températures, la simulation de contrôle est stationnaire, indiquant une référence suffisamment représentative. Sur la Figure 3.3a, nous pouvons voir que les scénarios sont bien détectés, et que les précipitations annuelles sont bien altérées par le changement climatique. Par contre, l'information donnée par ce test ne permet pas d'analyser l'altération en cours. Sur la Figure 3.3b, nous pouvons voir que l'altération est très faible pour le RCP45, quasiment confondue avec la simulation de contrôle. La simulation RCP85 ne se détache qu'après 2050, indiquant une très faible modification de la distance entre l'attracteur climatique et les étés. Sur la Figure 3.3c, nous retrouvons le forçage du changement climatique. Ceci indique que les conditions d'hivers sont de plus en plus rarement visitées

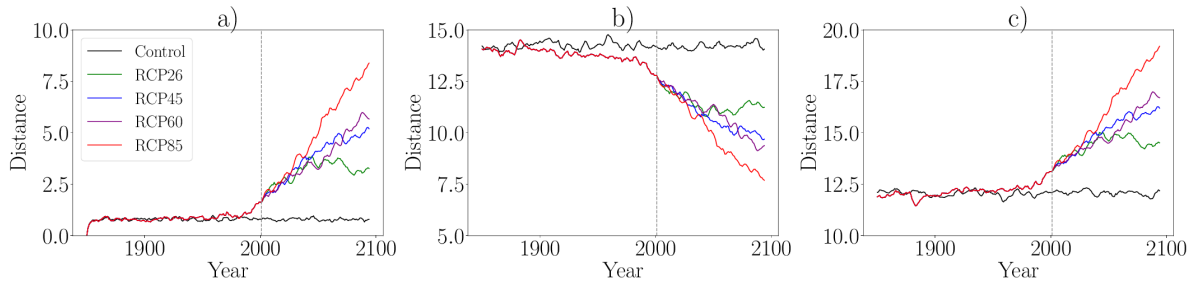


FIGURE 3.2 – Moyenne sur l’hémisphère nord des distances de Wasserstein entre une fenêtre glissante de 5 ans et une référence sur la période 1850-1855. La variable analysée est la température. La date sur l’axe des x correspond au début de la fenêtre. Le modèle utilisé est celui de l’IPSL (DUFRESNE et al., 2013). **Gris** : limite entre la partie historique et les scénarios RCP. **Noir** : simulation préindustrielle. **Vert** : scénario RCP26. **Bleu** : scénario RCP45. **Violet** : scénario RCP60. **Rouge** : scénario RCP85. **a)** Les fenêtres et la référence contiennent l’année complète. **b)** La référence est restreinte aux étés. **c)** La référence est restreinte aux hivers.

sur l’attracteur.

3.4 Résumé

Dans cette section, j’ai utilisé la distance de Wasserstein sur le modèle de LORENZ (1984), et sur plusieurs simulations d’un modèle climatique. Les résultats sur l’attracteur de LORENZ (1984) montrent la bonne capacité de la distance de Wasserstein à distinguer des situations différentes. Ceci, même quand la quantité de données est très faible. Ces résultats ont abouti à un article publié (ROBIN et al., 2017). Ces résultats sont un résumé de plusieurs tests effectués auparavant :

1. Un test de détection de la différence entre l’attracteur de LORENZ (1963) et l’attracteur de RÖSSLER (1976). Ce dernier attracteur a été représenté sur la Figure (3.4). D’un point de vue géométrique, il a aussi deux « ailes », et j’ai comparé la capacité de la distance de Wasserstein et Euclidienne à les distinguer. Les résultats obtenus sont équivalents à ceux de ROBIN et al. (2017).
2. Un test où une équation hybride non autonome entre le système de LORENZ (1963) et de RÖSSLER (1976) a été défini. Une séquence de snapshots a été construite sur l’intervalle de temps $[0, 150]$. À $t = 0$, le snapshot est celui de l’attracteur de LORENZ (1963). À $t = 150$, le snapshot est celui de l’attracteur de RÖSSLER (1976). Entre les deux, les équations différentielles se transforment linéairement de

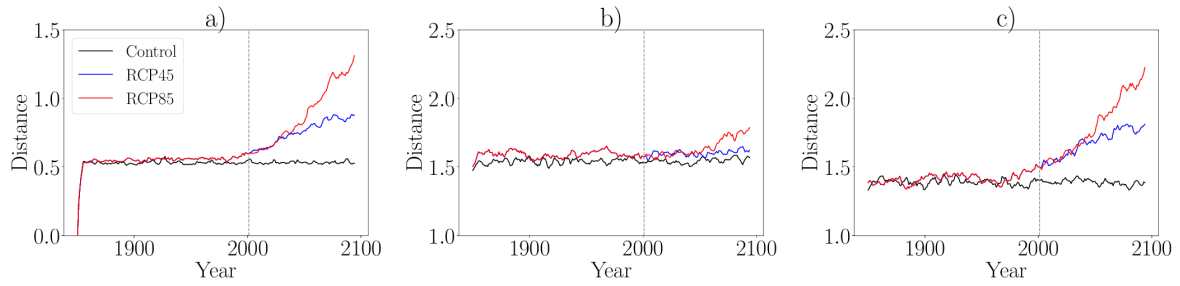


FIGURE 3.3 – Moyenne sur l’hémisphère nord des distances de Wasserstein entre une fenêtre glissante de 5 ans et une référence sur la période 1850-1855. La variable analysée est la précipitation. La date sur l’axe des x correspond au début de la fenêtre. Le modèle utilisé est celui de l’IPSL (DUFRESNE et al., 2013). **Gris** : limite entre la partie historique et les scénarios RCP. **Noir** : simulation pré-industrielle. **Vert** : scénario RCP26. **Bleu** : scénario RCP45. **Violet** : scénario RCP60. **Rouge** : scénario RCP85. **a)** Les fenêtres et la référence contiennent l’année complète. **b)** La référence est restreinte aux étés. **c)** La référence est restreinte aux hivers.

celle de LORENZ (1963) à celle de RÖSSLER (1976). La distance de Wasserstein entre les deux snapshots extrémaux et la séquence permet de bien visualiser la transformation.

3. Un test a été réalisé où la distance de Wasserstein (ou Euclidienne) devait retrouver un attracteur de LORENZ (1963) parmi un millier d’autres objets semblables. Ces autres objets étaient construits en tirant des points sur l’attracteur, mais selon une loi uniforme ou Gaussienne. La distance de Wasserstein permettait de retrouver le bon attracteur.

Une analyse sur le modèle de l’IPSL a aussi été effectuée. Une simulation pré-industrielle a servi de référence. Les scénarios RCP ont été étudiés. Sur deux variables, la température et la précipitation, j’ai retrouvé les forçages appliqués par ces scénarios, validant la distance de Wasserstein comme mesure de l’altération de l’attracteur climatique. Pour mesurer plus finement quelles sont ces altérations, le plan de transport peut permettre de suivre comment l’attracteur est modifié. De plus, le protocole utilisé ici a traité indépendamment chaque point de grille. Une amélioration notable serait de prendre en compte la dépendance spatiale.

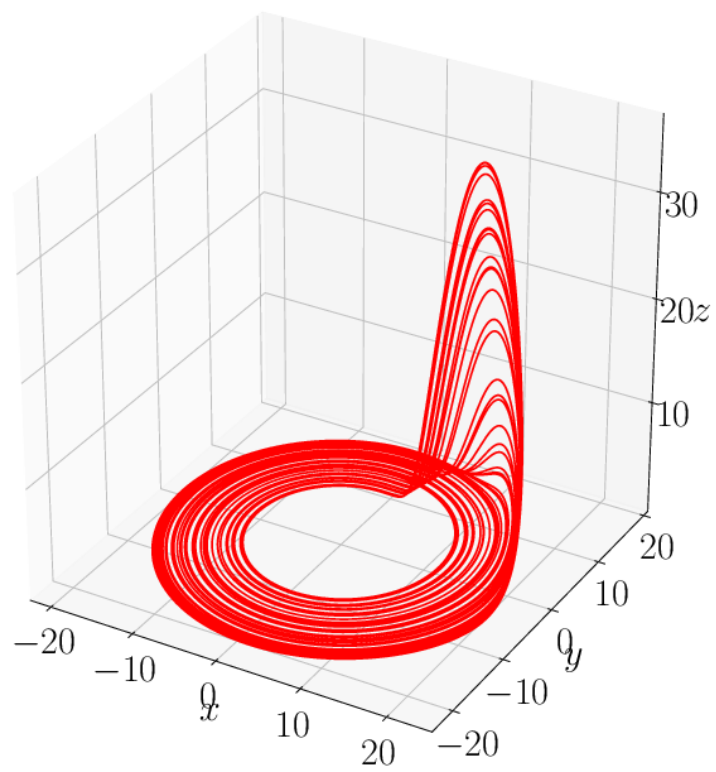


FIGURE 3.4 – Attracteur de (RÖSSLER, 1976).

Chapitre 4

Correction de biais multivarié

Dans ce chapitre, j’ai travaillé sur les liens entre la correction de biais, et le transport optimal. En effet, en climatologie, les modèles numériques sont en général biaisés par rapport aux observations. Par exemple, la température moyenne est trop basse, ou trop élevée. Ceci pose problème lorsque les sorties des GCM (Global Climate Model) ou RCM (Regional Climate Model) sont utilisées dans des études d’impacts (typiquement en agronomie ou hydrologie, CHEN et al., 2013).

Classiquement, ces modèles sont corrigés avec le *quantile mapping* (e.g., PIANI et al., 2010). Cette méthode applique la fonction $\mathcal{T} = F^{-1} \circ G$ à chaque point de grille et pour chaque variable indépendamment, où F (resp. G) est la CDF (Cumulative Distribution Function) des observations (resp. du modèle). Les corrélations inter-variables et spatiales ne sont donc pas corrigées, ce qui altère la correction (WILCKE et al., 2013).

Je propose dans ce chapitre une solution à ce problème via le transport optimal. Le plan de transport indique « comment » déplacer une mesure de probabilité (même multivarié) sur une autre. Ceci permet d’effectuer la correction. J’ai tout d’abord testé cette méthodologie sur un modèle idéalisé basé sur le LORENZ (1984). Puis je l’ai aussi vérifié en corrigeant une simulation de WRF (Weather Research and Forecasting, SKAMAROCK et al., 2008) avec comme références les ré-analyses Safran (Système d’Analyse Fournissant des Renseignements Atmosphériques à la Neige, VIDAL et al., 2010). Ces travaux ont abouti à un article. La version actuelle de cet article est donnée ci dessous.

4.1 Abstract

La correction de biais est souvent utilisée pour calibrer les sorties de modèles numériques avec des enregistrements d’observation. L’objectif principal est de s’assurer que les caractéristiques statistiques (telles que la

moyenne et la variance) des modèles numériques sont conformes aux observations. Dans cet article, une méthode de correction de biais stochastique multivariée est développée, basée sur des techniques de transport optimal. La correction de biais, généralement définie comme une fonction de transfert, est vue comme une loi de probabilité jointe entre une variable aléatoire biaisée et sa correction. La théorie du transport optimal permet de construire une loi jointe particulière, minimisant l'énergie. Ceci étend le quantile mapping dans un contexte multivarié. Une définition d'une correction de biais non stationnaire est proposée, comme le transfert du modèle dans le monde des observations, et nous étendons notre méthode à ce contexte. La méthodologie est d'abord validée sur un modèle chaotique idéalisé à 3 variables. Les corrélations entre variables semblent presque parfaitement corrigées par notre méthode dans cet environnement contrôlé, contrairement à une correction univariée. Notre méthodologie est aussi testée sur des précipitations et températures à 12 points de grille dans le sud de la France. La correction inter-variables et spatiales des températures et précipitations apparaît en accord avec l'évolution multi-dimensionnelle du modèle, satisfaisant notre définition suggérée de la non stationnarité.

4.2 Multivariate stochastic bias corrections with optimal transport

Multivariate stochastic bias corrections with optimal transport

Yoann Robin¹, Mathieu Vrac¹, Philippe Naveau¹, and Pascal Yiou¹

¹Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL & U Paris-Saclay, Gif-sur-Yvette, France

Correspondence: Yoann Robin (yoann.robin@lsce.ipsl.fr)

Abstract. Bias correction methods are used to calibrate climate model outputs with respect to observational records. The goal is to ensure that statistical features (such as means and variances) of climate simulations are coherent with observations. In this article, a multivariate stochastic bias correction method is developed based on optimal transport. Bias correction methods are usually defined as transfer functions between random variables. We show that such transfer functions induce a joint probability distribution between the biased random variable and its correction. The optimal transport theory allows us constructing a joint distribution that minimizes an energy spent in bias correction. This extends the classical univariate quantile mapping techniques in the multivariate case. We also propose a definition of non-stationary bias correction as a transfer of the model to the observational world, and we extend our method in this context. Those methodologies are first tested on an idealized chaotic system with three variables. In those controlled experiments, the correlations between variables appear almost perfectly corrected by our method, as opposed to a univariate correction. Our methodology is also tested on daily precipitation and temperatures over 12 locations in southern France. The correction of the inter-variable and inter-site structures of temperatures and precipitation appears in agreement with the multi-dimensional evolution of the model, hence satisfying our suggested definition of non-stationarity.

Copyright statement. TEXT

1 Introduction

Global Climate Model (GCM) and Regional Climate Model (RCM) are used to study the climate system. However, their outputs often appear biased compared to observational references (e.g., Randall et al., 2007). For example, the temperature means can be shifted. Thus removing this bias is often necessary to drive impact studies such as based on crop or hydrological models (Chen et al., 2013). The main goal of bias correction (BC) is to match the statistical features of climate models outputs with observations (see, e.g. Ehret et al., 2012; Gudmundsson et al., 2012). The most used method is the quantile mapping (Panofsky and Brier, 1958; Wood et al., 2004; Déqué, 2007), which adjusts the quantiles of the variables of interest in the stationary case (Shrestha et al., 2014). The stationarity has been questioned by Christensen et al. (2008); Maraun (2012); Nahar et al. (2017). Some extensions, like CDF- t (Cumulative Distribution Function transfer, Michelangeli et al., 2009), can take into account

some of the non-stationarity in GCM or RCM.

Most of those methods are univariate, and do not take into account the spatial and inter-variable correlations, which may alter the quality of the corrections (e.g., Wilcke et al., 2013; Maraun, 2016). This shortcoming has led to the recent development of multivariate techniques. As mentioned by Vrac (2018), two kinds of methods are currently available. The first type corrects separately each marginal, and apply afterwards a correction of the dependence structure (e.g., Vrac and Friederichs, 2015; Vrac, 2018; Nahar et al., 2018; Cannon, 2018). The second kind performs recursive corrections: each variable is corrected conditionally on the previously already corrected variables (Bárdossy and Pegram, 2012; Dekens et al., 2017). These last methods have two main limitations. First, the correction depends on the ordering of the marginals. Second, each marginal is adjusted conditionally on previously corrected marginals, which reduces the number of data at each step. Furthermore, the variability of observations is generally greater than that of the climate models. To increase the variability, von Storch (1999), Wong et al. (2014) and Mao et al. (2015) suggested to introduce a stochastic component in the bias correction procedure. In this paper, we develop a multivariate and stochastic bias correction method, different from the two categories presented, based on elements from the optimal transport theory.

Optimal transport theory is a natural way to measure the dissimilarity between multivariate probability distributions (Villani, 2008; Muskulus and Verduyn-Lunel, 2011; Robin et al., 2017), especially in a multivariate case. For example, this has been already successfully applied in image processing to transfer colors between images (Rubner et al., 2000; Ferradans et al., 2013). Here, our goal is to apply optimal transport techniques to perform bias correction in estimating a particular *joint law* (called a *transport plan*) that links the probability distributions of a biased random variable and its correction. This joint law minimizes a cost function, representing the *energy* needed to transform a multivariate probability distribution to another. In this optimal transport context, any realization of the biased random variable induces a conditional law of the transport plan, associating the realization and its correction. As the corrections are randomly drawn from these conditional laws, the suggested method is stochastic by construction.

This paper is organized as follows. In Section 2, the developed theoretical framework to perform bias correction is presented. In particular, the classical definition of bias correction as transfer function is generalized with optimal transport theory. Two methods are presented : Optimal Transport Correction (OTC, stationary case) and dynamical Optimal Transport Correction (dOTC, non-stationary case). In Section 3, the proposed methodology is tested on an idealized non-stationary case based on chaotic attractors. In Section 4, a multivariate bias correction is performed on a regional climate model (RCM) simulation of temperatures and precipitation, in a cross-validation case. Section 5 provides conclusions and perspectives.

2 Theoretical framework

2.1 Bias correction as a joint distribution

The general goal of this paper is the correction of a random variable, noted $\mathbf{X}$ (e.g., a biased climate model output) with respect to a reference random variable, noted $\mathbf{Y}$. The random variables $\mathbf{X}$ and $\mathbf{Y}$ live in dimension d . If $d = 1$, we note them X and Y . The probability law of $\mathbf{X}$ (resp. $\mathbf{Y}$) is noted $\mathbb{P}_{\mathbf{X}}$ (resp. $\mathbb{P}_{\mathbf{Y}}$).

5

Following Piani et al. (2010), a bias correction method of $\mathbf{X}$ with respect to $\mathbf{Y}$ is a *transfer function* $\mathcal{T} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that the random variable $\mathcal{T}(\mathbf{X})$ follows the same law as $\mathbf{Y}$, i.e. $\mathbb{P}_{\mathcal{T}(\mathbf{X})} = \mathbb{P}_{\mathbf{Y}}$. This definition covers most of the practical cases, but we can construct random variables where no transfer function exists, e.g. if $\mathbf{X}$ is constant and $\mathbf{Y}$ is not. Thus, beyond a multivariate transfer function, it is necessary to extend the definition of bias correction.

10

We assume momentarily the existence of a transfer function $\mathcal{T}$. By construction, the random variables $\mathbf{X}$ and $\mathcal{T}(\mathbf{X})$ are dependent, and their associated joint law can be summarized by the function $\kappa : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$

$$\kappa(\mathbf{x}) := (\mathbf{x}, \mathcal{T}(\mathbf{x})) \in \mathbb{R}^d \times \mathbb{R}^d.$$

The map κ connects the random variable $\mathbf{X}$ with its correction $\mathcal{T}(\mathbf{X})$ on the space $\mathbb{R}^d \times \mathbb{R}^d$. Furthermore, the map κ induces a probability law on $\mathbb{R}^d \times \mathbb{R}^d$, noted $\mathbb{P}_{\mathcal{T}}$, and given for all measurable sets $A \subset \mathbb{R}^d \times \mathbb{R}^d$ by

$$\mathbb{P}_{\mathcal{T}}(A) := \mathbb{P}_{\mathbf{X}}(\kappa^{-1}(A)) = \mathbb{P}_{\mathbf{X}}(\{\mathbf{x} \in \mathbb{R}^d \text{ such that } \kappa(\mathbf{x}) \in A\}) = \mathbb{P}((\mathbf{x}, \mathcal{T}(\mathbf{x})) \in A).$$

The critical property here concerns the margins of $\mathbb{P}_{\mathcal{T}}$: *the first (resp. second) margin of $\mathbb{P}_{\mathcal{T}}$ is $\mathbb{P}_{\mathbf{X}}$ (resp. $\mathbb{P}_{\mathbf{Y}}$)*. To understand why it is critical, we note $\Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$ the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ such that $\mathbb{P}_{\mathbf{X}}$ is the first margin and $\mathbb{P}_{\mathbf{Y}}$ the second one. By definition, $\mathbb{P}_{\mathcal{T}} \in \Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$. Thus, any bias correction method defined by a transfer function induces an element of $\Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$.

We argue that *any probability distribution in $\Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$ induces a bias correction method*. For $\gamma \in \Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$, $\gamma(\mathbf{x}, \mathbf{y})$ can be interpreted as *the probability that $\mathbf{y}$ is the correction of $\mathbf{x}$* . Formally, thanks to the Jirina theorem (see e.g., Strook, 1995, chap. 5), there exists a collection of probability laws $\gamma_{\mathbf{x}}$, $\mathbf{x} \in \mathbb{R}^d$, such that $\gamma_{\mathbf{x}}$ are the conditional laws of $\mathbf{Y}$ *given $\mathbf{X}$* . In other words, for $B \subset \mathbb{R}^d$, $\gamma_{\mathbf{x}}(B)$ is the probability that the correction $\mathbf{y} \in B$, given $\mathbf{X} = \mathbf{x}$. The correction of $\mathbf{x}$ is then sampled from the law $\gamma_{\mathbf{x}}$. Thus, any $\gamma \in \Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$ defines a bias correction method, through the conditional laws $\gamma_{\mathbf{x}}$. Here, the stochastic part of this approach is highlighted. All corrections are sampled from the laws $\gamma_{\mathbf{x}}$, and the corrected values follow the law $\mathbb{P}_{\mathbf{Y}}$ (by definition of a conditional law).

30 We note that the problem where $\mathbf{X}$ is constant is easily solved with this approach. The set $\Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$ is reduced to one element: the independent law $\delta_{\mathbf{x}} \times \mathbb{P}_{\mathbf{Y}}$, where $\delta_{\mathbf{x}}$ is the Dirac mass in $\mathbf{x}$. Thus, $\gamma_{\mathbf{x}} = \mathbb{P}_{\mathbf{Y}}$, and the correction of $\mathbf{X}$ is given by sampling each correction with the law $\mathbb{P}_{\mathbf{Y}}$.

We have defined a bias correction method as an element of $\Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$. However, this set can be very large. The goal of the
 5 next section is to present a *criterion* to select an element of $\Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$.

2.2 Selection of a joint law with optimal transport theory

To select a probability law $\gamma \in \Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$, we propose to use a cost function on this set. The minimum of this cost function corresponds to an optimal bias correction method. We propose to minimize the energy needed to transform a realization $\mathbf{x}$ of $\mathbf{X}$, to its correction $\mathbf{y}$, i.e. minimize $\|\mathbf{x} - \mathbf{y}\|^2$, weighted by $\gamma(\mathbf{x}, \mathbf{y})$. Thus, the cost function C is given by :

$$10 \quad C : \begin{cases} \Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}}) & \rightarrow \mathbb{R}_+, \\ \gamma & \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^d} \|\mathbf{x} - \mathbf{y}\|^2 \cdot d\gamma(\mathbf{x}, \mathbf{y}). \end{cases} \quad (1)$$

This cost function minimizes the square of the distance between $\mathbf{x}$ and its correction $\mathbf{y}$. Our bias correction method is associated with the law γ that minimizes C . This cost function stems from optimal transport theory (Villani, 2008). To understand this choice, we propose to examine the univariate case. We note F (resp. G) the Cumulative Distribution Function (CDF) of X (resp. Y), defined by $F(x) := \mathbb{P}_X(X \leq x)$ (resp. $G(y) := \mathbb{P}_Y(Y \leq y)$). We assume that F and G are continuous. The transfer
 15 function of the quantile mapping method is given by $\mathcal{T} = G^{-1} \circ F$, where G^{-1} corresponds to the quantile function of Y , i.e. the inverse of G . To understand this definition, we can write that

$$\begin{aligned} \mathbb{P}_{\mathcal{T}(X)}(\mathcal{T}(X) \leq y) &= \mathbb{P}_{\mathcal{T}(X)}(F(X) \leq G(y)), && \text{by definition of } \mathcal{T}, \\ &= G(y), && \text{because } F(X) \text{ follows the uniform law,} \\ &= \mathbb{P}_Y(Y \leq y), && \text{by definition of } G. \end{aligned}$$

Hence, the random variable $\mathcal{T}(X)$ follows the same law as Y . In the univariate case, it can be shown that the joint law defined by $\mathcal{T} = G^{-1} \circ F$ minimizes the cost function C of Eq. (1). Proofs of this statement can be found in Farchi et al. (2016, Appendix
 20 A) and Santambrogio (2015, Chap. 2).

To summarize the univariate case, we have seen that the classical quantile mapping can be viewed as the solution of a minimization problem based on the cost function of Eq. (1). This solution can be derived from optimal transport theory. Our next step is to explain how this minimization strategy can be extended in the multivariate case.

2.3 Multivariate bias correction with optimal transport selection : the stationary case

Given $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ and $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ two independent and identically distributed (i.i.d.) samples of the random variables $\mathbf{X}$ and $\mathbf{Y}$. A first step is to estimate the empirical distributions, $\hat{\mathbb{P}}_{\mathbf{X}}$ and $\hat{\mathbb{P}}_{\mathbf{Y}}$. We note $\mathbf{c}_i$ a collection of regularly spaced cells that

partition $\mathbb{R}^d$, and cover $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ and $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$. The center of each cell is also noted $\mathbf{c}_i$. With this notation, $\hat{\mathbb{P}}_{\mathbf{X}}$ and $\hat{\mathbb{P}}_{\mathbf{Y}}$ can be written as a sum of I and J Dirac masses:

$$5 \quad \hat{\mathbb{P}}_{\mathbf{X}}(A) = \sum_{i=1}^I p_{\mathbf{X},i} \delta_{\mathbf{c}_i}(A), \text{ where } p_{\mathbf{X},i} = \frac{1}{n} \sum_{l=1}^n \mathbf{1}(\mathbf{X}_l \in \mathbf{c}_i), \text{ and } A \subset \mathbb{R}^d,$$

$$\hat{\mathbb{P}}_{\mathbf{Y}}(B) = \sum_{j=1}^J p_{\mathbf{Y},j} \delta_{\mathbf{c}_j}(B), B \subset \mathbb{R}^d.$$

The scalar $p_{\mathbf{X},i}$ (resp. $p_{\mathbf{Y},j}$) is the empirical weight around $\mathbf{c}_i$ (resp. $\mathbf{c}_j$) and induced from the sampling of $\mathbf{X}$ (resp. $\mathbf{Y}$). A natural estimator of $\gamma \in \Gamma(\mathbb{P}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}})$ can be written as

$$\hat{\gamma}(A \times B) = \sum_{i,j=1}^{I,J} \gamma_{i,j} \delta_{(\mathbf{c}_i, \mathbf{c}_j)}(A \times B).$$

- 10 The coefficients γ_{ij} are the probabilities to transform $\mathbf{c}_i$ (i.e. a $\mathbf{x} \in \mathbf{c}_i$) to $\mathbf{c}_j$ (i.e. a $\mathbf{y} \in \mathbf{c}_j$). They are unknown, and they have to obey the marginal properties:

$$\sum_{j=1}^J \gamma_{ij} = p_{\mathbf{X},i}, \tag{2}$$

$$\sum_{i=1}^I \gamma_{ij} = p_{\mathbf{Y},j}. \tag{3}$$

Finally, the cost function defined in Eq. (1) can be approximated by

$$15 \quad \hat{C}(\hat{\gamma}) = \sum_{i,j=1}^{I,J} \|\mathbf{c}_i - \mathbf{c}_j\|^2 \gamma_{ij} \tag{4}$$

- Finding γ_{ij} , i.e. solving the problem defined by constraints of Eqs. (2-3) and minimization of Eq. (4), is called a *linear programming problem*. It can be solved (for example) by the network simplex algorithm (see, e.g. Bazaraa et al., 2009). We use the python implementation of Flamary and Courty (2017). To correct $\mathbf{X}$, we have to follow the plan of γ_{ij} . For a realization $\mathbf{X}_l$ of $\mathbf{X}$, we take the cell $\mathbf{c}_i$ that contains $\mathbf{X}_l$. Following $\hat{\gamma}$, $\mathbf{c}_i$ is moved to $\mathbf{c}_j$ with probability $\gamma_{ij}/p_{\mathbf{X},i}$ (applying Eq. (2), the sum over j is 1). To find which $\mathbf{c}_j$, we draw it according to the conditional law $\hat{\gamma}_{\mathbf{X}_l} = (\gamma_{i1}, \dots, \gamma_{iJ})/p_{\mathbf{X},i}$. Finally, we draw uniformly $\mathbf{y}$ in $\mathbf{c}_j$. This methodology is summarized in Algorithm 1, and we refer to it as *Optimal Transport Correction (OTC)*.
- 20

Algorithm 1 Optimal Transport Correction (OTC)

Require: $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ a sample i.i.d. of the random variable $\mathbf{X}$

$(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ a sample i.i.d. of the random variable $\mathbf{Y}$

Ensure: $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ a sample i.i.d., correction of $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ with respect to the estimation of the law of $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$

- 1: Estimate the law $\hat{\mathbb{P}}_{\mathbf{X}}$, from $(\mathbf{X}_1, \dots, \mathbf{X}_n)$
 - 2: Estimate the law $\hat{\mathbb{P}}_{\mathbf{Y}}$, from $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$
 - 3: Compute the optimal plan γ_{ij} between $\hat{\mathbb{P}}_{\mathbf{X}}$ and $\hat{\mathbb{P}}_{\mathbf{Y}}$ (see, e.g. Flamary and Courty, 2017)
 - 4: **for all** $\mathbf{X}_l$ **do**
 - 5: Find the cell $\mathbf{c}_i$ containing $\mathbf{X}_l$
 - 6: Construct the vector $\hat{\gamma}_{\mathbf{X}_l} = (\gamma_{i,1} \dots \gamma_{i,J}) / p_{\mathbf{X},i}$ (The conditional law)
 - 7: Draw $j \in \{1, \dots, J\}$ according to probability vector $\hat{\gamma}_{\mathbf{X}_l}$.
 - 8: Draw uniformly $\mathbf{Z}_l$ a realization of $\mathbf{Y}$ in cell $\mathbf{c}_j$.
 - 9: $\mathbf{Z}_l$ is a realization of $\mathbf{Z}$, corresponding to a correction of $\mathbf{X}_l$
 - 10: **end for**
-

We propose an example, where $\mathbb{P}_{\mathbf{X}}$ is a standardized normal distribution and $\mathbb{P}_{\mathbf{Y}}$ a mixture of two normal distributions centered at 3 and -3 , respectively. This example is univariate, and the corresponding plan is a histogram in dimension $d + d = 2$, see Fig. 1. The panels of Fig. 1 show the empirical histograms of $\hat{\mathbb{P}}_{\mathbf{X}}$ and $\hat{\mathbb{P}}_{\mathbf{Y}}$, with a bin size of 0.3. Under each histogram, the color scale represents the mass at each location. The central panel shows γ_{ij} in $\mathbb{R}^2$. The black arrow coming from $\hat{\mathbb{P}}_{\mathbf{X}}$ splits the mass located in bins $[-0.3, 0)$ into eleven γ_{ij} . These γ_{ij} are now moved into the segment $[-2, 1]$, following the grey arrows. To correct a point $x \in [-0.3, 0)$, we have to follow the black arrow, and draw a grey arrow, according to $\hat{\gamma}$.

Note that the traditional one-dimensional quantile mapping preserves the ordering of quantiles. In the multivariate case, this type of property can be viewed as the *Monge-Mather (1991) shortening principle* (see e.g. Villani, 2008, chap. 8). The idea is that the extremes of a multivariate distribution is moved to extremes, the boundary to the boundary, the level lines to level lines, etc.

2.4 Non stationary bias correction

Climate models offer a valuable tool to study future realistic climate trajectories. Climate model outputs of the present period need to be bias corrected with respect to current observations. Future climate simulations also need to be adjusted. However, no observation is available for the future and clear assumptions have to be made to correct simulations for future periods. Table 1 displays the basic framework of bias correction. Future unobserved data, say $\mathbf{Y}^1$, should be inferred from the current reference vector, $\mathbf{Y}^0$, and two numerical runs, one in the present, say $\mathbf{X}^0$, and one in the future, say $\mathbf{X}^1$. The period 0 is called the

calibration period, the period 1 the *projection period*. In the univariate case, noting F^i (resp. G^i) the CDF of X^i (resp. Y^i), the CDF- t (CDF transform) method of Michelangeli et al. (2009) assumes that

$$(G^1)^{-1} \circ G^0 = \mathcal{T}_{Y^0, Y^1} = \mathcal{T}_{X^0, X^1} = (F^1)^{-1} \circ F^0. \quad (5)$$

- 5 Recombining Eq. (5), the CDF of Y^1 is given by $G^1 = G^0 \circ (F^0)^{-1} \circ F^1$, and can be used to perform a quantile mapping correction. Here, the fundamental hypothesis $\mathcal{T}_{Y^0, Y^1} = \mathcal{T}_{X^0, X^1}$ means that the transfer functions to capture the temporal changes are identical in the model and observational worlds.

CDF- t learns the change between X^0 and X^1 , and transfers it to Y^0 to estimate Y^1 . In the multivariate case, following CDF-
10 t , we want to learn the evolution (i.e. the change or the temporal evolution) between $\mathbf{X}^0$ and $\mathbf{X}^1$, and apply it to $\mathbf{Y}^0$. This generates $\mathbf{Y}^1$, and OTC can then be applied between $\mathbf{X}^1$ and $\mathbf{Y}^1$. Our definition of non-stationary bias correction assumes a transfer of the evolution of the model to observational world. But the evolution of observation can be different, and the resulting correction can be also different from observations. This methodology is justified because we want to keep the evolution of the model, even if the dynamic of the model is different of the dynamic of the observations.

15

Using OTC, we define two optimal plans. The optimal plan γ , between $\mathbf{X}^0$ and $\mathbf{Y}^0$, and the optimal plan φ , between $\mathbf{X}^0$ and $\mathbf{X}^1$. The law γ is the *bias* between $\mathbf{X}^0$ and $\mathbf{Y}^0$, whereas φ is the *evolution* between $\mathbf{X}^0$ and $\mathbf{X}^1$. Our goal is to move φ *along* γ , defining a plan $\tilde{\varphi}$, to estimate $\mathbf{Y}^1$ as the evolution of $\mathbf{Y}^0$, i.e. $\mathbf{Y}^1 = \tilde{\varphi}(\mathbf{Y}^0)$. Then, we correct $\mathbf{X}^1$ with respect to $\mathbf{Y}^1 = \tilde{\varphi}(\mathbf{Y}^0)$, with OTC method. This is summarized in Fig. 2.

20

The estimation of $\tilde{\varphi}$ is performed in three steps:

1. transformation of φ into a collection of vectors,
2. transfer of these vectors along γ and
3. adaptation of these vectors to $\mathbf{Y}^0$.

- 25 To illustrate our methodology, Fig. 3 shows an example where the random variables $\mathbf{X}^0$, $\mathbf{X}^1$ and $\mathbf{Y}^0$ follow a bivariate Gaussian law. They are respectively centered at $(0, 0)$, $(10, 0)$ and $(0, 10)$, with covariance matrices $4 \times \mathbf{Id}_2$, $\mathbf{Id}_2/4$ and $\mathbf{Id}_2/4$ (the matrix $\mathbf{Id}_d$ is the d -dimensional identity matrix). Without loss of generality, we write the empirical distribution of $\mathbf{X}^0$, $\mathbf{X}^1$ and $\mathbf{Y}^0$ as a sum of Dirac masses,

$$\hat{\mathbb{P}}_{\mathbf{X}^0} = \sum_{i=1}^I p_{\mathbf{X}^0, i} \delta_{\mathbf{c}_i},$$

$$\hat{\mathbb{P}}_{\mathbf{Y}^0} = \sum_{j=1}^J p_{\mathbf{Y}^0, j} \delta_{\mathbf{c}_j},$$

$$\hat{\mathbb{P}}_{\mathbf{X}^1} = \sum_{k=1}^K p_{\mathbf{X}^1, k} \delta_{\mathbf{c}_k}.$$

- 5 **Step 1 : transformation of φ .** Using OTC method, φ moves the bin $\mathbf{c}_i$ of $\hat{\mathbb{P}}_{\mathbf{X}^0}$ to the bin $\mathbf{c}_k$ of $\hat{\mathbb{P}}_{\mathbf{X}^1}$. The vector $\mathbf{v}_{ik} := \mathbf{c}_k - \mathbf{c}_i$ represents the evolution from $\mathbf{c}_i$ to $\mathbf{c}_k$ (i.e. the local evolution between $\mathbf{X}^0$ and $\mathbf{X}^1$). The collection of vectors $\mathbf{v}_{ik}$ is an estimation of the process between $\mathbf{X}^0$ and $\mathbf{X}^1$. In Fig. 3, the red arrow is an example of vector $\mathbf{v}_{ik}$.

- Step 2 : transfer along γ .** Using OTC method, γ moves the bin $\mathbf{c}_i$ of $\hat{\mathbb{P}}_{\mathbf{X}^0}$ to the bin $\mathbf{c}_j$ of $\hat{\mathbb{P}}_{\mathbf{Y}^0}$. Thus, the estimation of $\tilde{\varphi}$ could
 10 be defined by the vector $\mathbf{v}_{ik}$ applied to $\mathbf{c}_j$, i.e. a realization of $\mathbf{Y}^1$ is given by $\mathbf{c}_j + \mathbf{v}_{ik}$. The grey arrow in Fig. 3a depicts this operation. But the $\mathbf{v}_{ik}$ can cross, and the correction is not coherent. This is due to normalizing issues and because the collection of vectors $\mathbf{v}_{ik}$ applied to $\mathbf{Y}^0$ do not define an optimal transport plan. The standard deviation decreases between $\mathbf{X}^0$ and $\mathbf{X}^1$, whereas it increases between $\mathbf{Y}^0$ and $\mathbf{Y}^1$. Consequently, we have to adapt the vectors $\mathbf{v}_{ik}$ to $\hat{\mathbb{P}}_{\mathbf{Y}^0}$.

- 15 **Step 3 : adaptation of $\mathbf{v}_{ik}$.** To solve this problem, we introduce a matrix factor $\mathbf{D}$, which rescales the collection of vectors $\mathbf{v}_{ik}$. In the univariate case, Bürger et al. (2011) proposed a factor $\sigma_{\mathbf{Y}^0} \sigma_{\mathbf{X}^0}^{-1}$, where $\sigma_{\bullet}$ is the standard deviation. The idea is to remove the scale of $\mathbf{X}^0$, and to replace it by the scale of $\mathbf{Y}^0$. Bárdossy and Pegram (2012); Cannon (2016) proposed a multivariate equivalent that uses the Cholesky decomposition of the covariance matrix. Noting Σ the covariance matrix, and $\text{Cho}(\Sigma)$ its Cholesky decomposition, we multiply (in matrix sense) $\mathbf{v}_{ik}$ by the following matrix:

$$20 \quad \mathbf{D} := \text{Cho}(\Sigma_{\mathbf{Y}^0}) \cdot \text{Cho}(\Sigma_{\mathbf{X}^0})^{-1}. \quad (6)$$

- The Cholesky decomposition only exists if Σ is symmetric, positive and definite. Some covariance matrices do not have this property, e.g. highly correlated random variables. In such a case, Σ must be lightly perturbed to be definite (see, e.g. Higham, 1988; Knol and ten Berge, 1989). Furthermore, the Cholesky decomposition can be poorly estimated if the number of available data is too small compared to the dimension. Indeed, the inverse of a covariance matrix is highly biased. In this case,
 25 a pragmatic solution is to replace the matrix $\mathbf{D}$ by the diagonal matrix of standard deviation, i.e. $\mathbf{D} = \text{diag}(\sigma_{\mathbf{Y}^0} \sigma_{\mathbf{X}^0}^{-1})$.

- Finally, a realization of $\mathbf{Y}^1$ is given by $\mathbf{c}_j + \mathbf{D} \cdot \mathbf{v}_{ik}$. Figure 3b shows an estimation of $\mathbf{Y}^1$. Visually, the shape of $\mathbf{Y}^1$ appears coherent with the evolution between $\mathbf{X}^0$ and $\mathbf{X}^1$. The mean of $\mathbf{Y}^1$ is (2.53, 10). The standard deviation between $\mathbf{X}^0$ and $\mathbf{X}^1$ is divided by 4. The mean shift between $\mathbf{X}^0$ and $\mathbf{X}^1$ is (10, 0). This shift of 10 units is correctly taken into account in the rescaling
 30 of $\mathbf{Y}^0$ by the standard deviation (equal to 4) between $\mathbf{X}^0$ and $\mathbf{X}^1$:

$$\underbrace{(2.53, 10)}_{\mathbf{Y}^1 \text{ mean}} = \underbrace{(10, 0)}_{\text{mean shift between } \mathbf{X}^0 \text{ and } \mathbf{X}^1} / \underbrace{4}_{\text{Rescaling}} + \underbrace{(0, 10)}_{\mathbf{Y}^0 \text{ mean}}.$$

The value of the covariance matrix of $\mathbf{Y}^1$ is $\Sigma_{\mathbf{Y}^1} \simeq 0.018 \times \mathbf{Id}_2$. It is close of the expected value $(1/4)/16 \times \mathbf{Id}_2 \simeq 0.015 \times \mathbf{Id}_2$. The shift of 10 units of the model is not respected. It is interpreted as a correction of the bias into the evolution of the model. However, depending on the hypotheses desired by the user, the dOTC method can easily provide corrections whose mean evolutions and trends are in agreement with those given by the simulations to be corrected, like in the EDQM bias correction method (Li et al., 2010). The complete method of correction is summarized in Algorithm 2. We refer to it by *dOTC* (dynamical Optimal Transport Correction).

Algorithm 2 dynamical Optimal Transport Correction (dOTC)

Require: $(\mathbf{X}_1^0, \dots, \mathbf{X}_n^0)$ a sample i.i.d. of the random variable $\mathbf{X}^0$

$(\mathbf{X}_1^1, \dots, \mathbf{X}_n^1)$ a sample i.i.d. of the random variable $\mathbf{X}^1$

$(\mathbf{Y}_1^0, \dots, \mathbf{Y}_n^0)$ a sample i.i.d. of the random variable $\mathbf{Y}^0$

Ensure: $(\mathbf{Z}_1^1, \dots, \mathbf{Z}_n^1)$ a sample i.i.d. of the random variable $\mathbf{Z}^1$

- 1: Estimate the law $\hat{\mathbb{P}}_{\mathbf{X}^0}$, from $(\mathbf{X}_1^0, \dots, \mathbf{X}_n^0)$
 - 2: Estimate the law $\hat{\mathbb{P}}_{\mathbf{X}^1}$, from $(\mathbf{X}_1^1, \dots, \mathbf{X}_n^1)$
 - 3: Estimate the law $\hat{\mathbb{P}}_{\mathbf{Y}^0}$, from $(\mathbf{Y}_1^0, \dots, \mathbf{Y}_n^0)$
 - 4: Compute the optimal plan γ_{ij} between $\hat{\mathbb{P}}_{\mathbf{X}^0}$ and $\hat{\mathbb{P}}_{\mathbf{Y}^0}$ (see, e.g. Flamary and Courty, 2017)
 - 5: Compute the optimal plan φ_{ik} between $\hat{\mathbb{P}}_{\mathbf{X}^0}$ and $\hat{\mathbb{P}}_{\mathbf{X}^1}$
 - 6: Compute the Cholesky factor $\mathbf{D}$ between $(\mathbf{X}_1^0, \dots, \mathbf{X}_n^0)$ and $(\mathbf{Y}_1^0, \dots, \mathbf{Y}_n^0)$, given by Eq. (6).
 - 7: **for all** $\mathbf{Y}_l^0$ **do**
 - 8: Find the cell $\mathbf{c}_j$ containing $\mathbf{Y}_l^0$
 - 9: Using the plan γ_{ij} (see Alg. 1), find a cell $\mathbf{c}_i$ of $\hat{\mathbb{P}}_{\mathbf{X}^0}$
 - 10: Using the plan φ_{ik} , find a cell $\mathbf{c}_k$ of $\hat{\mathbb{P}}_{\mathbf{X}^1}$
 - 11: Compute the vector $\mathbf{v}_{ik} := \mathbf{c}_k - \mathbf{c}_i$
 - 12: $\mathbf{Y}_l^1 = \mathbf{Y}_l^0 + \mathbf{D} \cdot \mathbf{v}_{ik}$ is a realization of $\mathbf{Y}^1$
 - 13: **end for**
 - 14: Estimate the law $\hat{\mathbb{P}}_{\mathbf{Y}^1}$, from $(\mathbf{Y}_1^1, \dots, \mathbf{Y}_n^1)$
 - 15: Apply OTC (see Alg. 1) between $(\mathbf{X}_1^1, \dots, \mathbf{X}_n^1)$ and $(\mathbf{Y}_1^1, \dots, \mathbf{Y}_n^1)$ to generate $(\mathbf{Z}_1^1, \dots, \mathbf{Z}_n^1)$
-

We first propose to evaluate OTC and dOTC on an idealized case.

3 Bias correction on an idealized case

3.1 Model and methodology

To evaluate our bias correction method, we construct an idealized biased case, based on the Lorenz (1984) model. This three dimensional system is generated by the differential equations

5

$$\frac{d\mathbf{x}}{dt} = \begin{pmatrix} -x_2^2 - x_3^2 - (x_1 - \psi(t))/4 \\ x_1x_2 - 4x_1x_3 - x_2 + 1 \\ x_1x_3 + 4x_1x_2 - x_3 \end{pmatrix}. \quad (7)$$

The function $\psi(t)$ is a linear forcing proposed by Drótos et al. (2015). Classically, ψ contains also a seasonal cycle (Lorenz, 1990), where the length of a “year” is fixed at $t = 73$ time units. Here we integrate this equation for the following forcing between 0 and 7×73 (i.e. 7 “years” of integration):

$$10 \quad \psi(t) = 9.5 - 20 \frac{t-T}{T} \mathbf{1}_{\{t>T\}}, \quad T = 6 \times 73. \quad (8)$$

The integration is performed with a Runge-Kutta (order 4) scheme with a time step of size 0.005. All trajectories of the Lorenz (1984) model converge on a unique subset of $\mathbb{R}^3$ (called an attractor), and remain trapped on it. According to Drótos et al. (2015), the five first “years” correspond to the time required to trap the trajectories.

- 15 One realization of random variable $\mathbf{Y}^0$ (resp. $\mathbf{Y}^1$) is the year 6 (resp. year 7). Each year contains 14600 ($= 73/0.005$) elements. According to Eq. (8), the linear forcing is applied during the year 7. The non-stationarity is induced by the change between the two time periods.

We introduce a bias by multiplying each point of the trajectories by a triangular matrix $\mathbf{S}$, and add a vector $\mathbf{m}$, i.e. $\mathbf{X} = \mathbf{S}\mathbf{Y} + \mathbf{m}$.

- 20 The addition changes the mean, whereas the multiplication alters the covariances. The matrix $\mathbf{S}$ is chosen empirically such that the covariance matrices of $\mathbf{X}^0$, $\mathbf{X}^1$, $\mathbf{Y}^0$ and $\mathbf{Y}^1$ differ. We fix:

$$\mathbf{S} = \begin{pmatrix} 1.22 & 0 & 0 \\ -0.41 & 1.04 & 0 \\ -0.41 & 0.56 & 0.52 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

The random variables $\mathbf{X}$ and $\mathbf{Y}$ are plotted in Fig. 4a and Fig. 4d. The blue (resp. red) curve of Fig. 4a is the trajectory of $\mathbf{Y}^0$ (resp. $\mathbf{X}^0$). The mean is largely altered. We estimate the covariance matrices as

$$\hat{\text{Cov}}_{\mathbf{Y}^0} = \begin{pmatrix} 0.43 & -0.37 & -0.24 \\ -0.37 & 0.93 & 0.17 \\ -0.24 & 0.17 & 0.69 \end{pmatrix}, \quad \hat{\text{Cov}}_{\mathbf{X}^0} = \begin{pmatrix} 0.64 & -0.68 & -0.62 \\ -0.68 & 1.39 & 1.0 \\ -0.62 & 1.0 & 0.92 \end{pmatrix}$$

Similarly to Fig. 4a, Fig. 4d depicts in blue $\mathbf{Y}^1$, and in red $\mathbf{X}^1$. The forcing of Eq. (8) has changed the properties of the trajectories, and they became chaotic. It is worthwhile to notice that the dynamic of $\mathbf{Y}$ is comparable to the one of $\mathbf{X}$. The covariance matrices are largely affected:

$$\hat{\text{Cov}}_{\mathbf{Y}^1} = \begin{pmatrix} 0.27 & -0.09 & -0.14 \\ -0.09 & 0.81 & 0.08 \\ -0.14 & 0.08 & 0.73 \end{pmatrix}, \quad \hat{\text{Cov}}_{\mathbf{X}^1} = \begin{pmatrix} 0.4 & -0.25 & -0.29 \\ -0.25 & 1.0 & 0.65 \\ -0.29 & 0.65 & 0.64 \end{pmatrix}$$

We estimate the empirical distributions $\mathbb{P}_{\mathbf{Y}^0}$, $\mathbb{P}_{\mathbf{Y}^1}$, $\mathbb{P}_{\mathbf{X}^0}$ and $\mathbb{P}_{\mathbf{X}^1}$ with a 3 dimensional histogram. We cut a large cube around the trajectories into cells of size $0.2 \times 0.2 \times 0.2$. Then we count the number of points in each cell.

- 10 Finally, we evaluate the quality of the correction by comparing the covariance matrices of $\mathbf{Y}^0$ and $\mathbf{X}^0$, and the covariance matrices of $\mathbf{Y}^1$ and $\mathbf{X}^1$.

3.2 Correction of biased Lorenz (1984) model

- We apply our method to correct $\mathbf{X}^0$ and $\mathbf{X}^1$. The random variable $\mathbf{X}^0$ is corrected with respect to $\mathbf{Y}^0$ and using the OTC method. The random variable $\mathbf{X}^1$ is corrected with respect to the estimation of $\mathbf{Y}^1$, coming from the dOTC method. The resulting random variables $\mathbf{Z}^0$ and $\mathbf{Z}^1$ are given in green on panels **b** and **e** of Fig. 4. We depict also on Fig. 4c and Fig. 4f a univariate correction with quantile mapping (resp. CDF- t) for the period 0 (resp. 1), generating the random variables $\mathbf{Q}^0$ (resp. $\mathbf{Q}^1$).

- The correction $\mathbf{Z}^0$ is visually very similar to the reference in blue in Fig. 4a. The covariance matrix is almost perfectly reproduced

$$\hat{\text{Cov}}_{\mathbf{Z}^0} = \begin{pmatrix} 0.42 & -0.36 & -0.24 \\ -0.36 & 0.93 & 0.17 \\ -0.24 & 0.17 & 0.69 \end{pmatrix}, \quad \sup |\hat{\text{Cov}}_{\mathbf{Z}^0} - \hat{\text{Cov}}_{\mathbf{Y}^0}| = 0.004.$$

The correction $\mathbf{Z}^1$ is depicted in green in Fig. 4d. It is visually hard to compare to Fig. 4b, but we recognize $\mathbf{Y}^1$. The covariance matrix is correctly rectified

$$\hat{\text{Cov}}_{\mathbf{Z}^1} = \begin{pmatrix} 0.26 & -0.11 & -0.11 \\ -0.11 & 0.82 & 0.08 \\ -0.11 & 0.08 & 0.71 \end{pmatrix}, \quad \sup |\hat{\text{Cov}}_{\mathbf{Z}^1} - \hat{\text{Cov}}_{\mathbf{Y}^1}| = 0.03.$$

25 Finally, the cost of transformation (given by Eq. (1)) of $\mathbf{Z}^1$ into $\mathbf{Y}^1$ is 93% smaller than the cost between $\mathbf{Y}^1$ and $\mathbf{X}^1$, i.e. $\mathbb{P}_{\mathbf{Z}^1}$ is more similar to $\mathbb{P}_{\mathbf{Y}^1}$ than $\mathbb{P}_{\mathbf{X}^1}$. Furthermore, if we replace the Cholesky matrix of dOTC by the matrix of standard deviation, the maximum difference between covariance matrices increases to 0.22, but the cost is 85% smaller. Thus, using the standard deviation slightly degrades the correction. However, visually, it is very hard to distinguish the corrections with the Cholesky matrix or the standard deviation matrix. The figure corresponding to Fig. 4 with standard deviation matrix is given
5 in the supplementary material.

On the contrary, $\mathbf{Q}^0$ and $\mathbf{Q}^1$, depicted respectively in Fig. 4c and 4f, do not reproduce $\mathbf{Y}^0$ and $\mathbf{Y}^1$. Thus, the multivariate correction is largely better than the univariate correction, and $\mathbf{Y}^1$. It is confirmed by the covariance matrices, and $\mathbf{Y}^1$, which reproduce exactly the covariances of $\mathbf{X}^0$ and $\mathbf{X}^1$:

$$10 \quad \hat{\text{Cov}}_{\mathbf{Q}^0} = \begin{pmatrix} 0.42 & -0.42 & -0.42 \\ -0.42 & 0.95 & 0.68 \\ -0.42 & 0.68 & 0.69 \end{pmatrix}, \quad \sup |\hat{\text{Cov}}_{\mathbf{Q}^0} - \hat{\text{Cov}}_{\mathbf{Y}^0}| = 0.51,$$

$$\hat{\text{Cov}}_{\mathbf{Q}^1} = \begin{pmatrix} 0.13 & -0.1 & -0.14 \\ -0.1 & 0.59 & 0.39 \\ -0.14 & 0.39 & 0.46 \end{pmatrix}, \quad \sup |\hat{\text{Cov}}_{\mathbf{Q}^1} - \hat{\text{Cov}}_{\mathbf{Y}^1}| = 0.31.$$

We have performed a tri-variate correction on a non linear system exhibiting non standard probability measures (i.e. non Gaussian, non exponential, etc.). In the stationary case, the OTC method works almost perfectly. In the non-stationary case, the dOTC method produces a probability distribution closed to the expected result. We propose now to apply OTC and dOTC
15 on climate model simulations.

4 Bias correction of an RCM simulation

4.1 Data

The dataset used as reference for the bias correction (BC) is the reanalysis "Systeme d'Analyse Fournissant des Renseignements Atmospheriques a la Neige" (SAFRAN, Vidal et al., 2010). SAFRAN is a hourly reanalysis over France between 1958
20 and present, with a horizontal resolution of $8 \text{ km} \times 8 \text{ km}$. Quintana-Seguí et al. (2008) claimed that daily mean of the Surface Atmospheric Temperature (tas) and Precipitation (pr) present no bias compared to observations from the climatological

database of Météo-France. This justifies the use of SAFRAN as a reference.

We test our multivariate BC method on a simulation of the Weather Research and Forecast (WRF) atmospheric model (Skamarock et al., 2008) performed within the EURO-CORDEX initiative (Vautard et al., 2013; Jacob et al., 2014) with a $0.11^\circ \times 0.11^\circ$ horizontal resolution. The boundaries of the simulation were forced by a historical simulation of the Institut Pierre-Simon Laplace (IPSL) coupled model (Marti et al., 2010; Dufresne et al., 2013). This EURO-CORDEX historical simulation will be called “WRF” in the following.

SAFRAN and WRF data are re-mapped onto the same grid, with a spatial resolution of $0.11^\circ \times 0.11^\circ$ (i.e. $\sim 12\text{km} \times 12\text{km}$). The closest neighborhood method is used. We only keep the land region comprised in $1.8 - 7.85^\circ E \times 41.8 - 45.2^\circ N$ i.e. covering the south-east of France. This region is characterized by a complex topography, which creates a strong spatial heterogeneity, especially for precipitation. For the present application, we extract 12 grid points regularly spaced, see Fig. 5a, with a one to one spatial correspondence between SAFRAN and WRF.

In both datasets, we will consider daily surface air temperatures and precipitation. The goal of this section is to correct the bias in *tas* and *pr* in the WRF data with respect to SAFRAN.

4.2 Cross-validation protocol

We focus on the daily time scale over the 1970-2000 period. We correct the warm season (May-September). The analysis and conclusions are available for the cold season, and the corresponding figure (ie. Fig.5) is given in the supplementary material. We split that period into two sub-periods, 1970-1985 (2295 days), and 1985-2000 (2295 days) to perform a cross-validation. The SAFRAN (resp. WRF) values over the first time period correspond to the random variable $\mathbf{Y}^0$ (resp. $\mathbf{X}^0$), and is called the *calibration period*. The SAFRAN (resp. WRF) values over the second time period correspond to $\mathbf{Y}^1$ (resp. $\mathbf{X}^1$), and is called the *projection period*. SAFRAN during 1985-2000 (i.e. $\mathbf{Y}^1$) is assumed to be unknown, and is used for cross-validation.

We perform two bias corrections: univariate and 24-variate (12 grid points and 2 variables).

1. For univariate correction, quantile mapping is used for the calibration period, and CDF-*t* for the projection period.
 2. For 24-variate correction, OTC is used for the calibration period, and dOTC for the projection period. The spatial structure and the dependence between the two variables are used. Due to the dimension, the Cholesky matrix is poorly estimated. We replace it by the matrix of standard deviation in the rescaling step.
- We estimate the empirical distributions by computing histograms with bins of size 0.1 in each dimension. Furthermore, CDF-*t* and dOTC can shift close to 0 values to negative values for precipitation. Thus, negative precipitation values are replaced by 0 after correction. We test the quality of the correction by plotting the evolution of the mean, of the standard deviation, and the spatial and inter-variables covariance, i.e. the difference between projection and calibration period. These indicators are

summarized in Fig. 5. During calibration period, the goal is that the probability distribution of correction of the WRF simulation is the probability distribution of SAFRAN. By construction of OTC, the correction is almost perfect, and we focus on projection period. In projection period, the goal is that the evolution of corrections is close to the evolution of WRF simulation.

4.3 Evolution analysis

As we have seen in previous section, the correction of $\mathbf{X}^1$ and $\mathbf{Y}^1$ are identical only if the evolution of SAFRAN is identical to the evolution of WRF. To analyze the evolution of WRF, SAFRAN and the corrections, we compute the difference of statistical indicators between the projection and the calibration period at each grid points. The indicators are the mean (Fig. 5b,f), the variance (Fig. 5c,g), the covariance between pr and tas (Fig. 5e) and the spatial covariance for each variables (Fig. 5d,h).

5

The x axis of Figs. 5a-h is the evolution of the correction (i.e. $\mathbb{E}(\mathbf{Z}^1) - \mathbb{E}(\mathbf{Z}^0), \dots$). The y axis of Figs. 5a-h is the evolution of WRF in red (i.e. $\mathbb{E}(\mathbf{X}^1) - \mathbb{E}(\mathbf{X}^0), \dots$), and the evolution of SAFRAN in blue (i.e. $\mathbb{E}(\mathbf{Y}^1) - \mathbb{E}(\mathbf{Y}^0), \dots$). Furthermore, the red line is the linear regression between the evolution of the 24-variate correction and the evolution of WRF. The correlation (r -value), p -value and standard error of each linear regression are summarized in Tab. 2.

10

The linear regression between evolution of 24-variate correction and evolution of WRF (red line) shows a strong statistical link for all statistical indicators. The evolution of the mean is almost perfectly reproduced for the two variables (r -values is at least equal to 0.98, with a maximal p -value at 10^{-9}). The evolution of variance of WRF is also reproduced, the linear regression being significative (maximal p -value is 5×10^{-2}).

15

The evolution of dependence structure is given by the evolution of spatial and inter-variables covariance. The minimal r -value for linear regression is equal to 0.59 with a maximal p -value equal to 2×10^{-3} . This means that dOTC reproduces the evolution of WRF between calibration and projection period. Because the calibration period is perfectly corrected, the correction during projection period appears as the evolution of WRF, applied to SAFRAN.

20

A linear regression, the Spearman rank correlation between the evolution of SAFRAN, and the evolution of the correction with WRF do not show a significant statistical link (not shown). We conclude that the evolution of WRF is different of the evolution of SAFRAN. This indicates it is not possible to reproduce SAFRAN during projection period using dOTC and WRF.

25 The correction with CDF- t appears to be satisfactory for the temperatures, and very similar to the correction with dOTC. But for the precipitation, the structure is not coherent with WRF or SAFRAN. This dissimilarity is due to the difference between the probability distribution of temperatures (quasi-Gaussian) and precipitations (exponential/Gamma laws).

We conclude that the evolution of the 24-variate correction with dOTC between calibration and projection period is close to the evolution of WRF. Furthermore, the evolution of SAFRAN is very different from the evolution of WRF.

30

5 Conclusions

We have developed a new method for multivariate bias correction, generalizing the quantile mapping in the multivariate case. To do so, we have developed a new theoretical framework to understand any bias correction (BC) method: any BC method is here characterized by a joint law between the biased dataset and the correction. This joint probability distribution is estimated based on optimal transport techniques, and the BC method is then referred to as "Optimal Transport Correction" (OTC). A definition of non-stationary bias correction is also proposed: the evolution of the model is learned, and transferred to the reference world. An extension of OTC called dynamical OTC (dOTC) has been developed to account for temporal non-stationarities.

OTC and dOTC methods have been tested on an idealized 3 dimensional case based on Lorenz (1984) time-dependent attractors, which induced changes in the correlation between variables. The bias correction appeared to perform very well in those idealized experiments.

- 10 Then, 12 grid points of a WRF simulation have been corrected with respect to SAFRAN reanalyses for precipitation and temperature in Southern France. A 24-variate correction was performed. The correction in stationary context was almost perfect. In the non-stationary case, the evolutions of WRF and SAFRAN were different, and, as expected, the correction with dOTC differed from SAFRAN. However, the correction presented a multidimensional evolution similar to that of WRF. We can therefore conclude that the correction is consistent with the definition proposed for the non-stationary case.
- 15 Furthermore, although the number of available data is very small compared to the dimension (2295 days and 24 dimensions), the OTC and dOTC performed a correction without numerical problems, and, moreover, only in a few minutes on a personal computer.

- As a perspective of improvement of the method, we note that the optimal plan can only be used to correct data points that
- 20 are already known. If a new data point is obtained, and alters the estimate of the probability density function, then the plan needs to be recomputed. However, such a situation is relatively rare in bias correction. Indeed, the corrections usually have to be performed on climate model simulations that cover many years and decades. This means that the whole time series are available at once and are not continuously updated.

Code availability. OTC and dOTC are implemented in two packages: ARyga (R) and Apyga (python3). These packages are available upon request.

Data availability. TEXT

Code and data availability. TEXT

Author contributions. YR performed the analyses. The experiments were co-designed by YR and MV. All the authors contributed to writing
5 the manuscript

Competing interests. The authors declare no competing interest.

Acknowledgements. This work was supported by ERC Grant No. 338965-A2C2.

References

- Bárdossy, A. and Pegram, G.: Multiscale spatial recorrelation of RCM precipitation to produce unbiased climate change scenarios over large areas and small, *Water Resour. Res.*, 48, <https://doi.org/10.1029/2011WR011524>, <http://dx.doi.org/10.1029/2011WR011524>, w09502, 2012.
- Bazaraa, M. S., Jarvis, J. J., and Sherali, H. D.: *Linear Programming and Network Flows*, John Wiley & Sons, 4 edn., 2009.
- Bürger, G., Schulla, J., and Werner, A. T.: Estimates of future flow, including extremes, of the Columbia River headwaters, *Water Resour. Res.*, 47, <https://doi.org/10.1029/2010WR009716>, <http://dx.doi.org/10.1029/2010WR009716>, w10520, 2011.
- 15 Cannon, A. J.: Multivariate Bias Correction of Climate Model Output: Matching Marginal Distributions and Intervariable Dependence Structure, *J. Clim.*, 29, 7045–7064, <https://doi.org/10.1175/JCLI-D-15-0679.1>, <https://doi.org/10.1175/JCLI-D-15-0679.1>, 2016.
- Cannon, A. J.: Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables, *Clim. Dyn.*, 50, 31–49, <https://doi.org/10.1007/s00382-017-3580-6>, <https://doi.org/10.1007/s00382-017-3580-6>, 2018.
- 20 Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, *Water Resour. Res.*, 49, 4187–4205, <https://doi.org/10.1002/wrcr.20331>, <http://dx.doi.org/10.1002/wrcr.20331>, 2013.
- Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophys. Res. Lett.*, 35, <https://doi.org/10.1029/2008GL035694>, <http://dx.doi.org/10.1029/2008GL035694>, 120709, 2008.
- 25 Dekens, L., Parey, S., Grandjacques, M., and Dacunha-Castelle, D.: Multivariate distribution correction of climate model outputs: A generalization of quantile mapping approaches, *Environmetrics*, 28, e2454, <https://doi.org/10.1002/env.2454>, <http://dx.doi.org/10.1002/env.2454>, e2454 env.2454, 2017.
- Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, *Glob. Planet. Change*, 57, 16 – 26, <https://doi.org/https://doi.org/10.1016/j.gloplacha.2006.11.030>, <http://www.sciencedirect.com/science/article/pii/S0921818106002748>, extreme Climatic Events, 2007.
- 30 Drótos, G., Bódai, T., and Tél, T.: Probabilistic concepts in a changing climate: a snapshot attractor picture, *J. Climate*, 28, 3275–3288, <https://doi.org/10.1175/JCLI-D-14-00459.1>, <http://dx.doi.org/10.1175/JCLI-D-14-00459.1>, 2015.
- Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, *Clim. Dyn.*, 40, 2123–2165, <https://doi.org/10.1007/s00382-012-1636-1>, <https://doi.org/10.1007/s00382-012-1636-1>, 2013.
- 35 Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", *Hydrol. Earth Syst. Sci.*, 16, 3391–3404, <https://doi.org/10.5194/hess-16-3391-2012>, <http://www.hydrol-earth-syst-sci.net/16/3391/2012>, 2012.

- Farchi, A., Bocquet, M., Roustan, Y., Mathieu, A., and Qu  rel, A.: Using the Wasserstein distance to compare fields of pollutants: application to the radionuclide atmospheric dispersion of the Fukushima-Daiichi accident, *Tellus B*, 68, 31682, <https://doi.org/10.3402/tellusb.v68.31682>, <https://doi.org/10.3402/tellusb.v68.31682>, 2016.
- Ferradans, S., Papadakis, N., Rabin, J., Peyr  , G., and Aujol, J.-F.: Regularized Discrete Optimal Transport, pp. 428–439, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-38267-3_36, https://doi.org/10.1007/978-3-642-38267-3_36, 2013.
- Flamary, R. and Courty, N.: POT Python Optimal Transport library, 2017.
- Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods, *Hydrol. Earth Syst. Sci.*, 16, 3383–3390, <https://doi.org/10.5194/hess-16-3383-2012>, <https://www.hydrol-earth-syst-sci.net/16/3383/2012/>, 2012.
- Higham, N. J.: Computing a nearest symmetric positive semidefinite matrix, *Linear Algebra Appl.*, 103, 103 – 118, [https://doi.org/https://doi.org/10.1016/0024-3795\(88\)90223-6](https://doi.org/https://doi.org/10.1016/0024-3795(88)90223-6), <http://www.sciencedirect.com/science/article/pii/0024379588902236>, 1988.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., D  qu  , M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kr  ner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, *Reg. Environ. Change*, 14, 563–578, <https://doi.org/10.1007/s10113-013-0499-2>, <https://doi.org/10.1007/s10113-013-0499-2>, 2014.
- Knol, D. L. and ten Berge, J. M. F.: Least-squares approximation of an improper correlation matrix by a proper one, *Psychometrika*, 54, 53–61, <https://doi.org/10.1007/BF02294448>, <https://doi.org/10.1007/BF02294448>, 1989.
- Li, H., Sheffield, J., and Wood, E. F.: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, *J. Geophys. Res. Atmos.*, 115, n/a–n/a, <https://doi.org/10.1029/2009JD012882>, <http://dx.doi.org/10.1029/2009JD012882>, d10101, 2010.
- Lorenz, E. N.: Irregularity: a fundamental property of the atmosphere, *Tellus A*, 36, 98–110, 1984.
- Lorenz, E. N.: Can chaos and intransitivity lead to interannual variability?, *Tellus A*, 42, 378–389, 1990.
- Mao, G., Vogl, S., Laux, P., Wagner, S., and Kunstmann, H.: Stochastic bias correction of dynamically downscaled precipitation fields for Germany through Copula-based integration of gridded observation data, *Hydrol. Earth Syst. Sci.*, 19, 1787–1806, <https://doi.org/10.5194/hess-19-1787-2015>, <https://www.hydrol-earth-syst-sci.net/19/1787/2015/>, 2015.
- Maraun, D.: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, *Geophys. Res. Lett.*, 39, <https://doi.org/10.1029/2012GL051210>, <http://dx.doi.org/10.1029/2012GL051210>, l06706, 2012.
- Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, *Curr. Clim. Change Rep.*, 2, 211–220, <https://doi.org/10.1007/s40641-016-0050-x>, <https://doi.org/10.1007/s40641-016-0050-x>, 2016.
- Marti, O., Braconnot, P., Dufresne, J.-L., Bellier, J., Benshila, R., Bony, S., Brockmann, P., Cadule, P., Caubel, A., Codron, F., de Noblet, N., Denvil, S., Fairhead, L., Fich  fet, T., Foujols, M.-A., Friedlingstein, P., Goosse, H., Grandpeix, J.-Y., Guilyardi, E., Hourdin, F., Idelkadi, A., Kageyama, M., Krinner, G., L  vy, C., Madec, G., Mignot, J., Musat, I., Swingedouw, D., and Talandier, C.: Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution, *Clim. Dyn.*, 34, 1–26, <https://doi.org/10.1007/s00382-009-0640-6>, <https://doi.org/10.1007/s00382-009-0640-6>, 2010.

- Mather, J. N.: Action minimizing invariant measures for positive definite Lagrangian systems, *Math. Z.*, 207, 169–207, <https://doi.org/10.1007/BF02571383>, <https://doi.org/10.1007/BF02571383>, 1991.
- 10 Michelangeli, P.-A., Vrac, M., and Loukos, H.: Probabilistic downscaling approaches: Application to wind cumulative distribution functions, *Geophys. Res. Lett.*, 36, <https://doi.org/10.1029/2009GL038401>, <http://dx.doi.org/10.1029/2009GL038401>, 111708, 2009.
- Muskulus, M. and Verduyn-Lunel, S.: Wasserstein distances in the analysis of time series and dynamical systems, *Physica D*, 240, 45 – 58, <https://doi.org/https://doi.org/10.1016/j.physd.2010.08.005>, <http://www.sciencedirect.com/science/article/pii/S0167278910002368>, 2011.
- Nahar, J., Johnson, F., and Sharma, A.: Assessing the extent of non-stationary biases in GCMs, *J. Hydrol.*, 549, 148 –
15 162, <https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.03.045>, <http://www.sciencedirect.com/science/article/pii/S0022169417301919>, 2017.
- Nahar, J., Johnson, F., and Sharma, A.: Addressing Spatial Dependence Bias in Climate Model Simulations: An Independent Component Analysis Approach, *Water Resour. Res.*, <https://doi.org/10.1002/2017WR021293>, <http://dx.doi.org/10.1002/2017WR021293>, 2018.
- Panofsky, H. A. and Brier, G. W.: Some applications of statistics to meteorology, Mineral Industries Extension Services, College of Mineral
20 Industries, Pennsylvania State University, 1958.
- Piani, C., Weedon, G., Best, M., Gomes, S., Viterbo, P., Hagemann, S., and Haerter, J.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, *J. Hydrol.*, 395, 199–215, <https://doi.org/10.1016/j.jhydrol.2010.10.024>, <http://www.sciencedirect.com/science/article/pii/S0022169410006475>, 2010.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., and Morel, S.: Analysis
25 of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France, *J. Appl. Meteorol. Climatol.*, 47, 92–107, <https://doi.org/10.1175/2007JAMC1636.1>, <https://doi.org/10.1175/2007JAMC1636.1>, 2008.
- Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., et al.: Climate models and their evaluation, in: *Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR)*, pp. 589–662, Cambridge University Press, 2007.
- 30 Robin, Y., Yiou, P., and Naveau, P.: Detecting changes in forced climate attractors with Wasserstein distance, *Nonlinear Process. Geophys.*, 24, 393, 2017.
- Rubner, Y., Tomasi, C., and Guibas, L. J.: The Earth Mover’s Distance as a Metric for Image Retrieval, *Int. J. Comput. Vis.*, 40, 99–121, <https://doi.org/10.1023/A:1026543900054>, <https://doi.org/10.1023/A:1026543900054>, 2000.
- Santambrogio, F.: *Optimal Transport for Applied Mathematicians*, vol. 87, Birkhäuser Basel, 2015.
- 35 Shrestha, R. R., Schnorbus, M. A., Werner, A. T., and Zwiers, F. W.: Evaluating Hydroclimatic Change Signals from Statistically and Dynamically Downscaled GCMs and Hydrologic Models, *J. Hydrometeorol.*, 15, 844–860, <https://doi.org/10.1175/JHM-D-13-030.1>, <https://doi.org/10.1175/JHM-D-13-030.1>, 2014.
- Skamarock, W., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., Huang, X., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, in: *NCAR Technical Note, NCAR/TN-475+STR*, <https://doi.org/10.5065/D68S4MVH>, 2008.
- Strook, D. W.: *Probability Theory, an Analytic View*, *J. Royal Stat. Soc. Series A*, 158, 339, 1995.
- Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., Fernández, J., García-Díez, M., Goergen, K., Güttler, I., Halenka, T., Karacostas, T., Katragkou, E., Keuler, K., Kotlarski, S., Mayer, S., van Meijgaard, E., Nikulin, G., Patarčić, M., Scinocca, J., Sobolowski,
5 S., Suklitsch, M., Teichmann, C., Warrach-Sagi, K., Wulfmeyer, V., and Yiou, P.: The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project, *Clim. Dyn.*, 41, 2555–2575, <https://doi.org/10.1007/s00382-013-1714-z>, <https://doi.org/10.1007/s00382-013-1714-z>, 2013.

- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, *Int. J. Climatol.*, 30, 1627–1644, <https://doi.org/10.1002/joc.2003>, <http://dx.doi.org/10.1002/joc.2003>, 2010.
- 10 Villani, C.: Optimal Transport: Old and New, vol. 338 of *Grundlehren der mathematischen Wissenschaften*, Springer Science & Business Media, 1 edn., 2008.
- von Storch, H.: On the Use of “Inflation” in Statistical Downscaling, *J. Clim.*, 12, 3505–3506, [https://doi.org/10.1175/1520-0442\(1999\)012<3505:OTUOII>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<3505:OTUOII>2.0.CO;2), [https://doi.org/10.1175/1520-0442\(1999\)012<3505:OTUOII>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<3505:OTUOII>2.0.CO;2), 1999.
- Vrac, M.: Multivariate bias adjustment of high-dimensional climate simulations: The “Rank Resampling for Distributions and Dependences” (R^2D^2) Bias Correction, *Hydrol. Earth Syst. Sci. Disc.*, 2018, 1–33, <https://doi.org/10.5194/hess-2017-747>, <https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-747/>, 2018.
- 15 Vrac, M. and Friederichs, P.: Multivariate—Intervariable, Spatial, and Temporal—Bias Correction, *J. Clim.*, 28, 218–237, <https://doi.org/10.1175/JCLI-D-14-00059.1>, <https://doi.org/10.1175/JCLI-D-14-00059.1>, 2015.
- Wilcke, R. A. I., Mendlik, T., and Gobiet, A.: Multi-variable error correction of regional climate models, *Clim. Change*, 120, 871–887, <https://doi.org/10.1007/s10584-013-0845-x>, <https://doi.org/10.1007/s10584-013-0845-x>, 2013.
- Wong, G., Maraun, D., Vrac, M., Widmann, M., Eden, J. M., and Kent, T.: Stochastic Model Output Statistics for Bias Correcting and
- 510 Downscaling Precipitation Including Extremes, *J. Clim.*, 27, 6940–6959, <https://doi.org/10.1175/JCLI-D-13-00604.1>, <https://doi.org/10.1175/JCLI-D-13-00604.1>, 2014.
- Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, *Clim. Change*, 62, 189–216, <https://doi.org/10.1023/B:CLIM.0000013685.99609.9e>, <http://dx.doi.org/10.1023/B:CLIM.0000013685.99609.9e>, 2004.

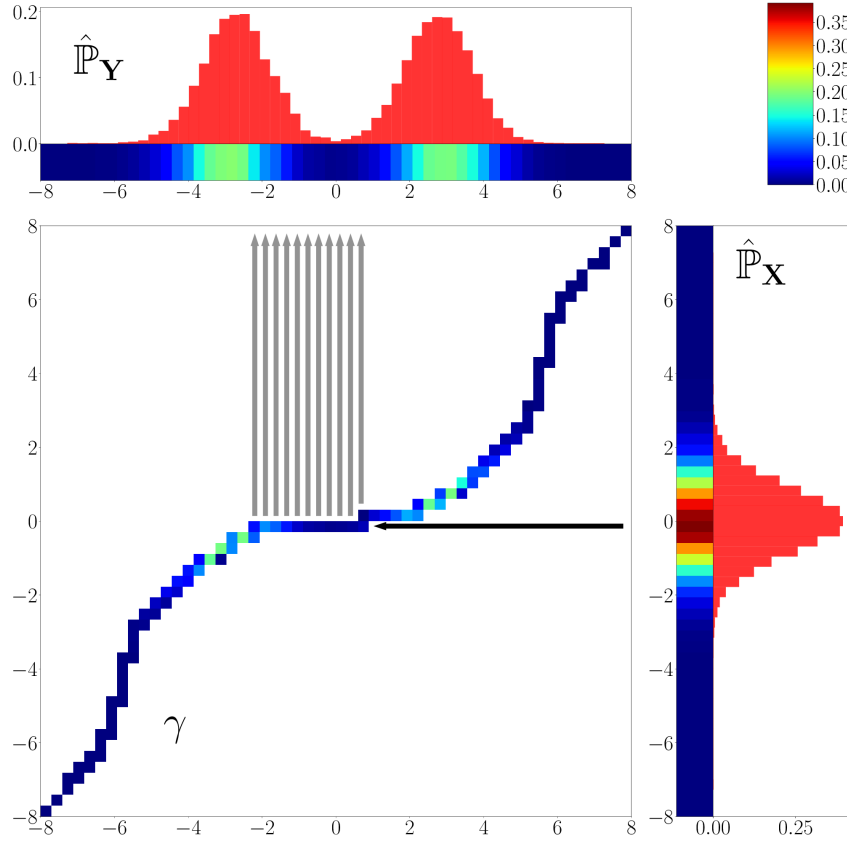


Figure 1. The right panel of $\hat{\mathbb{P}}_{\mathbf{X}}$ is the histogram of the standardized normal law. The upper panel of $\hat{\mathbb{P}}_{\mathbf{Y}}$ is the histogram of a mixture of two normal laws, centered in -3 and 3 . The size of bins is 0.3 . Under each histogram, the color scale represents the mass at each location. The central panel is the optimal joint law, minimizing the Eq. (4). The black and grey arrow represents the correction with Alg. 1.

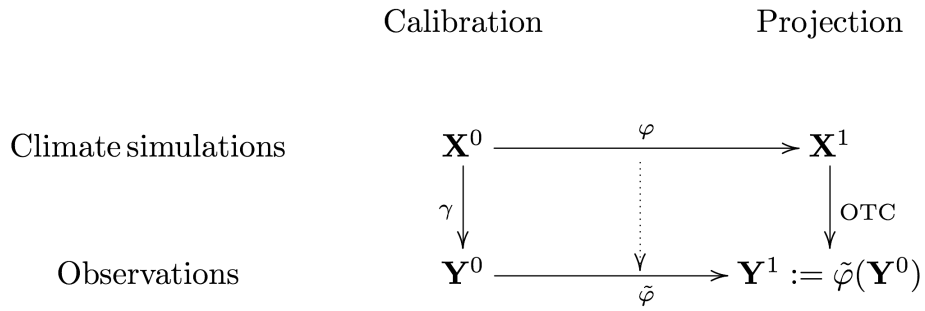


Figure 2. Estimation of the unobserved random variable $\mathbf{Y}^1$. The random variables $\mathbf{X}^0$, $\mathbf{X}^1$ and $\mathbf{Y}^0$ are known. The plans γ and φ are the optimal joint laws in the sense of equations (2-4). $\tilde{\varphi}$ is the evolution of $\mathbf{Y}^0$ estimated from γ and φ . OTC is used to correct $\mathbf{X}^1$ with respect to the estimation of $\mathbf{Y}^1$.

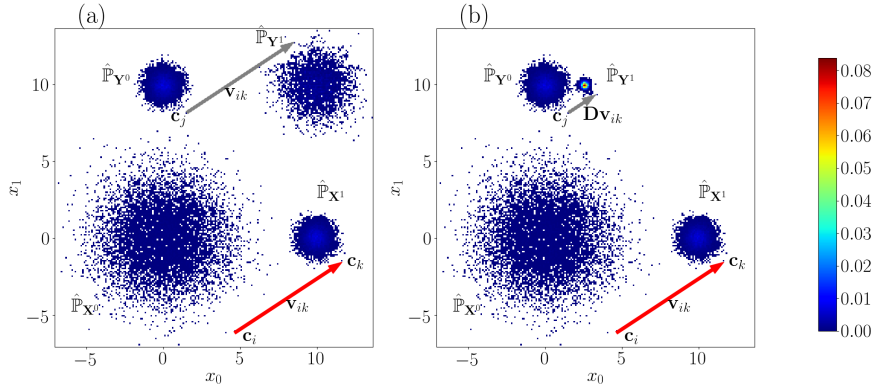


Figure 3. Bivariate histogram with bin size equal to 0.1. On each Panel we have: a Gaussian law centered in $(0,0)$ with covariance $4\mathbf{Id}_2$ ($\hat{\mathbb{P}}_{\mathbf{X}^0}$), a Gaussian law centered in $(10,0)$ with covariance $1/4\mathbf{Id}_2$ ($\hat{\mathbb{P}}_{\mathbf{X}^1}$) and a Gaussian law centered in $(0,10)$ with covariance $1/4\mathbf{Id}_2$ ($\hat{\mathbb{P}}_{\mathbf{Y}^0}$). The red arrow is the local evolution between $\hat{\mathbb{P}}_{\mathbf{X}^0}$ and $\hat{\mathbb{P}}_{\mathbf{X}^1}$. (a) The probability distribution $\hat{\mathbb{P}}_{\mathbf{Y}^1}$ is the correction with OTC- t and $\mathbf{D} = \mathbf{Id}_2$. The grey arrow is the estimation of the evolution of $\hat{\mathbb{P}}_{\mathbf{Y}^0}$. (b) The probability distribution $\hat{\mathbb{P}}_{\mathbf{Y}^1}$ is the correction with dOTC and $\mathbf{D}$ given by Eq. (6). The grey arrow is the estimation of the evolution of $\hat{\mathbb{P}}_{\mathbf{Y}^0}$.

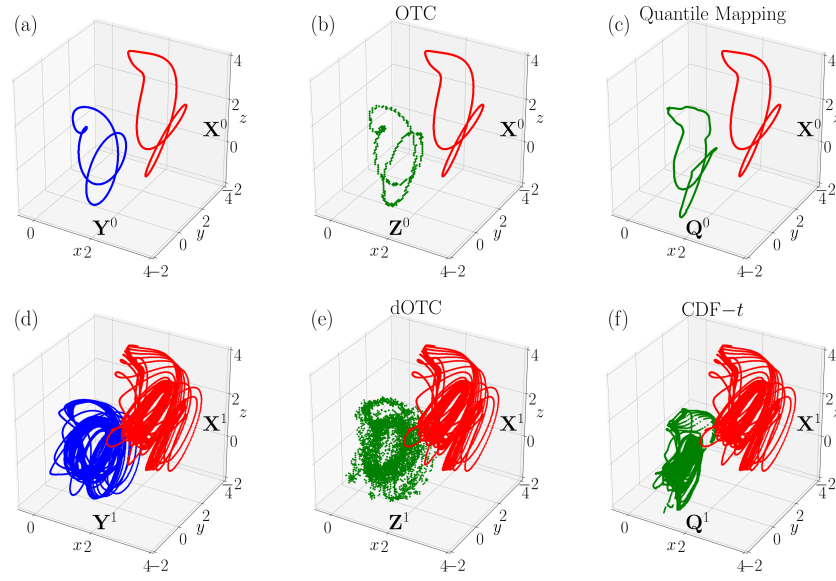


Figure 4. Random variables generated by Lorenz (1984) model, OTC, dOTC, quantile mapping and CDF- t . (a) Biased random variable X^0 (red) and references Y^0 (blue) for time period 0. (b) Biased random variable X^0 (red) and correction Z^0 with OTC (green). (c) Biased random variable X^0 (red) and correction Q^0 with quantile mapping (green). (d) Biased random variable X^1 (red) and references Y^1 (blue) for time period 1. (e) Biased random variable X^1 (red) and correction Z^1 with dOTC (green). (f) Biased random variable X^1 (red) and correction Q^1 with CDF- t (green).

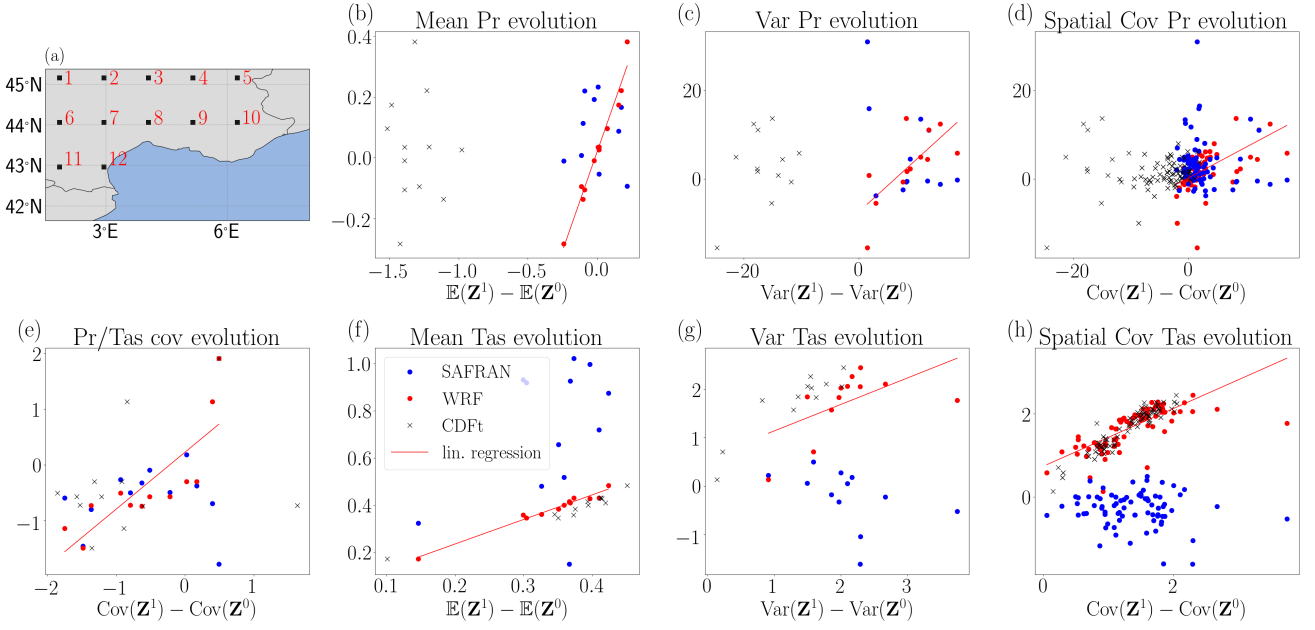


Figure 5. (a) Map of the south east of the France. The 12 black squares are the locations where corrections are performed. (b-h) The x axis of panels a-h is the evolution of the correction with dOTC. The y axis of panels a-h is the evolution of WRF in red and the evolution of SAFRAN in blue. The red line is the linear regression between the evolution of correction and the evolution of WRF. The black cross marker are the scatter plot between the evolution of correction with CDF- t and evolution of WRF. (b) Evolution of mean precipitation, i.e. difference between the projection period and the calibration period. (c) Evolution of variance of precipitation. (d) Evolution of spatial covariance of precipitation. (e) Evolution of covariance between precipitation and temperatures. (f) Evolution of mean temperatures. (g) Evolution of variance of temperatures. (h) Evolution of spatial covariance of temperatures.

Table 1. Representation of bias correction in context of climate change.

	Present	Future
Numerical model	$\mathbf{X}^0$	$\mathbf{X}^1$
Observations	$\mathbf{Y}^0$	unknown ($\mathbf{Y}^1$)

Table 2. r -value, p -value and standard error of linear regression between evolution of correction and evolution of WRF.

	r -value	p -value	standard error
Mean evolution Pr	0.98	10^{-8}	0.08
Mean evolution Tas	0.99	10^{-9}	0.05
Variance evolution Pr	0.71	10^{-2}	0.37
Variance evolution Tas	0.57	5×10^{-2}	0.25
Covariance Pr/Tas evolution	0.81	2×10^{-3}	0.24
Spatial covariance Pr	0.59	10^{-14}	0.08
Spatial covariance Tas	0.76	10^{-29}	0.05

Multivariate stochastic bias corrections with optimal transport: supplementary materials

Yoann Robin¹, Mathieu Vrac¹, Philippe Naveau¹, and Pascal Yiou¹

¹Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL & U Paris-Saclay, Gif-sur-Yvette, France

Correspondence: Yoann Robin (yoann.robin@lsce.ipsl.fr)

References

Lorenz, E. N.: Irregularity: a fundamental property of the atmosphere, *Tellus A*, 36, 98–110, 1984.

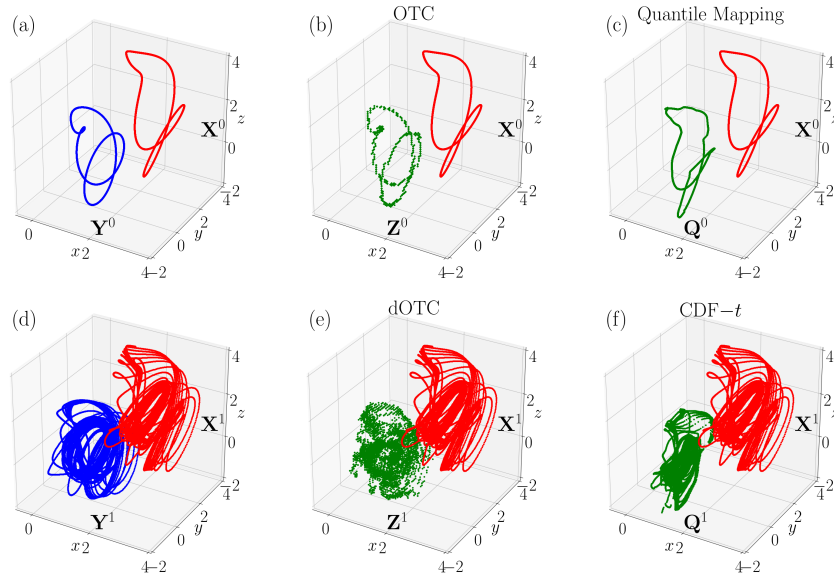


Figure 1. Trajectories generated by Lorenz (1984) model, OTC and dOTC, quantile mapping and CDF- t . The method dOTC uses the standard deviation matrix instead of the Cholesky matrix. (a) Biased random variable $\mathbf{X}^0$ (red) and references $\mathbf{Y}^0$ (blue) for time period 0; (b) Biased random variable $\mathbf{X}^0$ (red) and correction $\mathbf{Z}^0$ with OTC (green); (c) Biased random variable $\mathbf{X}^0$ (red) and correction $\mathbf{Q}^0$ with quantile mapping (green); (d) Biased random variable $\mathbf{X}^1$ (red) and references $\mathbf{Y}^1$ (blue) for time period 1; (e) Biased random variable $\mathbf{X}^1$ (red) and correction $\mathbf{Z}^1$ with dOTC (green); (f) Biased random variable $\mathbf{X}^1$ (red) and correction $\mathbf{Q}^1$ with CDF- t (green).

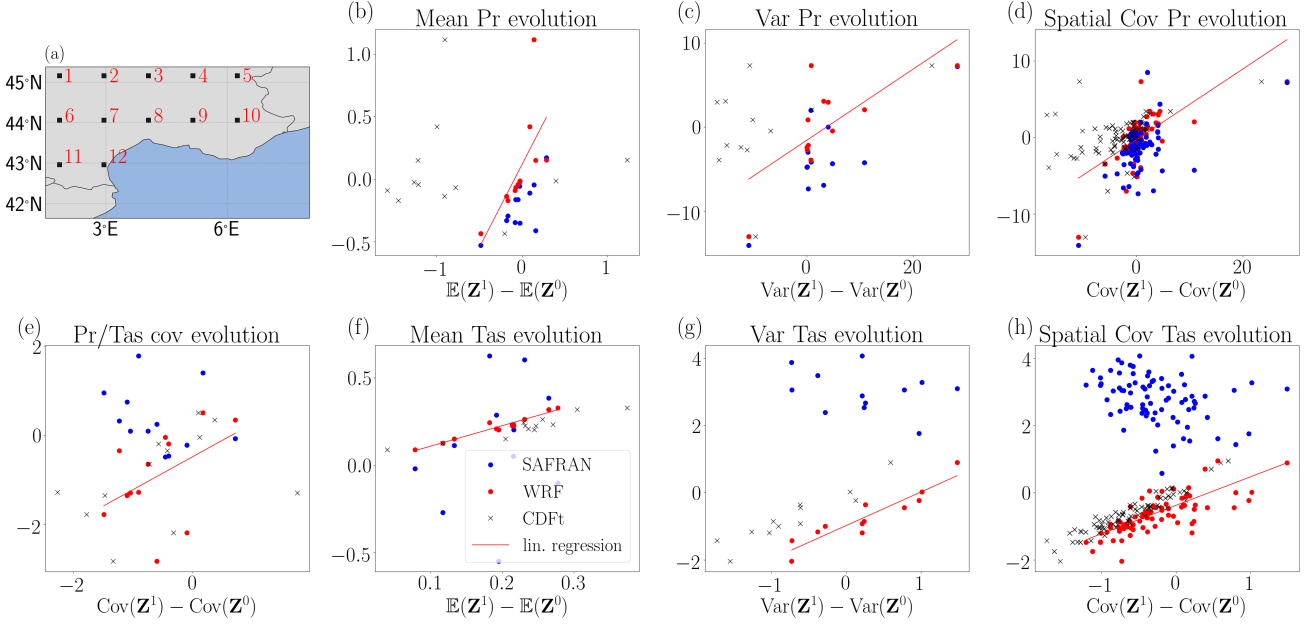


Figure 2. Evolution of correction with the dOTC method during winter. (a) Map of the south east of France. The 12 black squares are the locations where corrections are performed. (b-h) The x -axis of panels (a-h) is the evolution of the correction with dOTC. The y -axis of panels (a-h) is the evolution of WRF in red and the evolution of SAFRAN in blue. The red line is the linear regression between the evolution of correction and the evolution of WRF. The black cross markers are the scatter plot between the evolution of the correction with CDF- t and the evolution of WRF; (b) Evolution of mean precipitation, i.e. difference (in terms of mean) between the projection period and the calibration period; (c) Evolution of variance of precipitation; (d) Evolution of spatial covariance of precipitation; (e) Evolution of covariance between precipitation and temperatures; (f) Evolution of mean temperatures; (g) Evolution of variance of temperatures; (h) Evolution of spatial covariance of temperatures.

Table 1. r -values, p -values and standard errors of linear regressions between evolution of corrections and evolution of WRF.

	r -value	p -value	standard error
Mean evolution Pr	0.68	10^{-2}	0.45
Mean evolution Tas	0.97	10^{-7}	0.09
Variance evolution Pr	0.71	10^{-2}	0.13
Variance evolution Tas	0.93	10^{-5}	0.13
Covariance Pr/Tas evolution	0.45	10^{-1}	0.46
Spatial covariance Pr	0.63	10^{-17}	0.04
Spatial covariance Tas	0.77	10^{-29}	0.05

4.3 Résumé

Dans cette section, la définition d'une méthode de correction de biais a été étendue. Auparavant, elle était définie comme une fonction de transfert entre les données à corriger, et les données de références. J'ai proposé de définir une méthode de correction de biais par une loi jointe entre les données biaisées, et leurs corrections. Le transport optimal a été utilisé comme méthode de sélection d'une loi jointe, ce qui inclut le cas multivariée, et ajoute une partie stochastique à la correction. Deux méthodes ont été développées, une dans le cas stationnaire (OTC), une dans le cas de projections non stationnaires (dOTC). Elles ont été testées sur deux exemples :

1. Un modèle idéalisé, basé sur le système de LORENZ (1984). Les corrections sont presque parfaites dans ce contexte.
2. La correction d'une simulation de WRF avec les ré-analyses Safran.

Dans ce dernier cas, les corrélations spatiales des températures sont parfaitement corrigées. Les corrélations spatiales des précipitations ne sont pas corrigées, mais les corrélations entre les précipitations et les températures sont bien corrigées. Des travaux supplémentaires sont nécessaires pour tenir compte du cas particulier des précipitations.

Conclusion et perspectives

Le but original du projet A2C2 (qui a financé ma thèse) était d'utiliser des méthodes d'analogues pour mesurer des déformation d'attracteurs climatiques à la suite de forçages. Nous avons réalisé que nous pouvions répondre à cette question de manière directe avec des résultats de la théorie du transport optimal (VILLANI, 2008), et sans utiliser d'analogues. Lors de ces travaux de thèse, je me suis particulièrement intéressé à la quantification de la déformation de l'attracteur du climat, pour des systèmes idéalisés (de E.N. Lorenz) et des données de simulations numériques du climat.

Pour cela, la théorie du transport optimal a été l'objet central permettant une quantification naturelle de la dynamique sous jacente. Cette théorie, bien que très ancienne a été relancée récemment par les travaux de BENAMOU et BRENIER (2000). Son application a d'autant plus tardé que des algorithmes pouvant le résoudre en temps raisonnable (et facilement utilisables) n'ont été implémentés que récemment. La librairie Python de FLAMARY et COURTY (2017), en développement, fournit des algorithmes de résolution. J'ai moi-même écrit ma propre librairie, incluant leur algorithme de résolution, qui permet d'utiliser la distance de Wasserstein et d'appliquer une correction de biais de manière aisée.

L'utilisation du transport optimal en climatologie m'a permis de mettre au point une méthode d'analyse de données tenant compte de la dynamique et mathématiquement justifiée au travers de la théorie des systèmes dynamiques. Cette nouvelle méthode, décrite au Chapitre 3, permet de quantifier les changements de dynamique dans un système au cours du temps, même en présence d'une faible quantité de données ou de données partielles. Ces résultats ont été publiés dans la revue *Nonlinear Processes in Geophysics* (ROBIN et al., 2017). La principale limite est son temps de calcul, particulièrement élevé par rapport aux indicateurs usuels. Ce temps reste néanmoins raisonnable, et peut être amélioré par les développements de futurs algorithmes ou l'utilisation de GPU (Graphics Processing Unit) (CUTURI, 2013).

Dans un second temps, en quantifiant la différence entre deux attracteurs, j'ai aussi pu mettre au point une méthode de correction de biais multivariée, pouvant tenir compte de

la non stationnarité du climat. Là où actuellement les méthodes de correction multivariées sont inopérantes au-delà de 10 dimensions, la correction par transport optimal a permis de corriger de la température en dimension élevée (au delà de 1000), en corrigeant de manière quasi-parfaite les corrélations spatiales. La limitation principale est que la correction d’une nouvelle donnée oblige à recalculer complètement le plan de transport. Des travaux préliminaires pour contourner cette limitation ont été effectués (non affichés dans cette thèse). L’apprentissage d’un plan de transport par un réseau de neurones a fourni des résultats équivalents à une correction directe sur le modèle de LORENZ (1984). Une deuxième approche a aussi été testée : régulariser la mesure de probabilité induite par le plan de transport par un noyau Gaussien. Des contraintes algorithmiques m’ont bloqué pour exploiter cette idée, et des travaux supplémentaires sont nécessaires pour la rendre praticable. Mais ma thèse a jeté les bases de la correction de biais par transport optimal.

Les algorithmes générant les attracteurs, résolvant le transport optimal, permettant une correction de biais et calculant les dimensions locales ont été développés et placés sous licence CeCILL-C, favorisant leurs utilisations par la communauté scientifique. Ils sont regroupés au sein d’une librairie python, Apyga, qui devrait bientôt être disponible sur le site web du projet A2C2 (Atmospheric flow Analogues for Climate Change).

L’utilisation du transport optimal dans les sciences du climat ouvre un nombre important de perspectives. Plusieurs pistes sont envisagées. La théorie de FREITAS et al. (2010), qui permet de calculer les dimensions locales d’un attracteur, est limitée aux systèmes autonomes. Son utilisation dans le climat est donc insuffisamment justifiée. Un premier test pourrait être fait en utilisant la distance de Wasserstein entre des snapshots attracteurs du système non autonome de LORENZ (1984).

Ensuite, le plan de transport entre deux états du climat fournit la probabilité qu’un évènement se transforme en un autre évènement. Le plan pourrait donc être utilisé pour analyser les modifications d’évènements particuliers sous la contrainte du changement climatique. Les modifications des sous ensembles d’une distribution de probabilité peuvent être suivies au cours du temps, comme par exemple les extrêmes.

LUCARINI et al. (2017) ont utilisé la « théorie de la réponse linéaire » (linear response theory) de RUELLE (2009) pour détecter et quantifier l’impact du changement climatique sur les températures et les précipitations dans le modèle PLASIM (FRAEDRICH et al., 2005). La méthodologie présentée dans (ROBIN et al., 2017) pourrait être appliquée à des simulations de ce modèle, qui est plus proche du système climatique que le système de LORENZ (1984). Les résultats pourraient être comparés à ceux de LUCARINI et al. (2017).

Il est possible de faire un lien entre le transport optimal et l'utilisation d'analogues (tels qu'envisagés à l'origine de la thèse). Le transport optimal pourrait être utilisé pour calculer des analogues climatiques. Deux applications sont envisagées. La première est de l'utiliser directement comme distance entre cartes climatiques. La fonction de coût devra être modifiée, pour intégrer la distance sur une sphère. La seconde est de justifier l'utilisation d'un petit nombre d'analogues pour caractériser les événements extrêmes. Des travaux préliminaires en collaboration avec Aglaé Jézéquel, s'appuyant sur la théorie de FREITAS et al. (2010), ont permis de corroborer l'utilisation de 20 analogues pour le modèle de LORENZ (1984) et sur des données NCEP de pression au niveau de la mer sur l'Atlantique Nord. Des travaux supplémentaires sont nécessaires pour valider cette argumentation.

Dans le domaine de la correction de biais, le plan de transport permet de définir finement quelles sont les valeurs biaisées d'un modèle climatique, et quelles sont celles qui n'ont pas nécessité de correction. Étudier si les valeurs biaisées ont une structure particulière (spatiale, temporelle, corrélées à d'autres variables) permettrait de comprendre les sources des biais, et éventuellement les corriger en amont dans les modèles.

Deux points n'ont pas été abordés. Les méthodes de corrections de biais non stationnaires appliquent une estimation de la dynamique du modèle aux observations, afin de corriger les projections. Pour effectuer cela, l'estimation de la dynamique du modèle *est modifiée* pour correspondre aux observations. L'impact de cette modification sur les corrections des projections est peu discuté dans la littérature. C'est un point important pour les modèles d'impacts (comme les modèles agronomiques ou hydrologiques), qui peuvent utiliser des données corrigées. De plus, les méthodes de corrections non stationnaires déplacent la dynamique du modèle sur des observations, mais l'idée de déplacer le biais sur les projections du modèle n'est pas abordée. Ces deux approches ne donnent pas le même résultat si la variabilité du modèle climatique est significativement plus grande que celle des observations. Il est nécessaire d'identifier (par exemple avec une validation croisée) laquelle de ces deux approches est la plus pertinente, et dans quel contexte. Un projet d'inter-comparaison de méthodes de corrections multivariées est en discussion.

Annexe A

Utilisation de la distance de Wasserstein pour valider un indicateur

A.1 Abstract (RODRIGUES et al., 2017, accepted in J. Clim.)

Il est d'une importance fondamentale d'évaluer la capacité des modèles climatiques à capturer les schémas de circulation atmosphérique à grande échelle et, dans le contexte d'un forçage dû à l'effet de serre en croissance rapide, la robustesse des changements simulés dans ces modèles au cours du temps. Ici, nous abordons ce problème d'un point de vue novateur basé sur la théorie des systèmes dynamiques. Nous caractérisons la circulation atmosphérique sur l'Atlantique Nord dans les simulations historiques CMIP5 (1851 à 2000) en termes de deux métriques instantanées : la dimension locale de l'attracteur et la stabilité des trajectoires de l'espace des phases. Nous utilisons ensuite ces mesures pour comparer les modèles aux réanalyses 20CRv2c au cours de la même période historique. La comparaison suggère que : i) la plupart des modèles capturent dans une certaine mesure les propriétés de l'attracteur médian et les modèles avec des grilles plus fines donnent généralement de meilleurs résultats ; ii) dans la plupart des modèles, les extrêmes des systèmes dynamiques correspondent à des schémas à grande échelle similaires à ceux trouvés dans les réanalyses ; iii) les modifications des propriétés de l'attracteur observées pour les réanalyses de l'ensemble moyen 20CRv2c pourraient être des artefacts dus à des inhomogénéités dans l'écart-type de l'ensemble au cours

du temps; iv) les tendances à long terme de la dimension locale observées chez les 56 membres des ensembles 20-CR ont le même signe que celles observées dans la moyenne multimodèle CMIP5.

A.2 Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis



AMERICAN METEOROLOGICAL SOCIETY

Journal of Climate

EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/JCLI-D-17-0176.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Rodrigues, D., M. Alvarez-Castro, G. Messori, P. Yiou, Y. Robin, and D. Faranda, 2018: Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis. *J. Climate*. doi:10.1175/JCLI-D-17-0176.1, in press.



Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis

David Rodrigues¹, M. Carmen Alvarez-Castro¹, Gabriele Messori²,

Pascal Yiou¹, Yoann Robin¹, Davide Faranda^{1,3*}

¹*Laboratoire de Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL, Université Paris-Saclay, 91191 Gif-sur-Yvette, France.*

²*Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden.*

³*London Mathematical Laboratory, 14 Buckingham Street, London WC2N 6DF, UK*

*Corresponding author address: Laboratoire de Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL, Université Paris-Saclay, 91191 Gif-sur-Yvette, France.

E-mail: davide.faranda@lsce.ipsl.fr

ABSTRACT

13 It is of fundamental importance to evaluate the ability of climate models to
14 capture the large-scale atmospheric circulation patterns and, in the context of
15 a rapidly increasing greenhouse forcing, the robustness of the changes sim-
16 ulated in these patterns over time. Here we approach this problem from an
17 innovative point of view based on dynamical systems theory. We characterize
18 the atmospheric circulation over the North Atlantic in the CMIP5 historical
19 simulations (1851 to 2000) in terms of two instantaneous metrics: local di-
20 mension of the attractor and stability of phase-space trajectories. We then use
21 these metrics to compare the models to the 20CRv2c reanalysis over the same
22 historical period. The comparison suggests that: i) most models capture to
23 some degree the median attractor properties and models with finer grids gen-
24 erally perform better; ii) in most models the extremes in the dynamical sys-
25 tems metrics match large-scale patterns similar to those found in the reanal-
26 ysis; iii) changes in the attractor properties observed for the ensemble-mean
27 20CRv2c reanalysis are artifacts due to inhomogeneities in the standard devia-
28 tion of the ensemble over time; iv) the long-term trends in local dimension
29 observed among the 56 members of the 20-CR ensemble have the same sign
30 as those observed in the CMIP5 multimodel mean, although the multi-model
31 trend is much weaker.

1. Introduction

One of the main sources of uncertainty in determining the impact of climate change on extreme events is the forced response of atmospheric dynamics (Shepherd 2014; Field 2012). While changes in observables such as surface temperature are easily diagnosed, shifts in the mid-latitude atmospheric patterns have proved very difficult to quantify. Some advances have been made by focussing on specific features such as atmospheric blocking (Kay et al. 2015; Cassou and Cattiaux 2016; Faranda et al. 2016b), which in turn influence the occurrence of European cold spells and heat waves, but the broader appreciation of circulation changes is still unsatisfactory. Here we shed some light on this knowledge gap by using a dynamical systems framework. We illustrate the power of such an approach by considering the well-known Lorenz (1963) system, a conceptual model of atmospheric convection consisting of three differential equations¹

$$\dot{x} = \sigma(y - x), \quad \dot{y} = rx - y - xz, \quad \dot{z} = xy - bz, \quad (1)$$

where x , y , z represent respectively the convection strength, the temperature difference between the surface and the top of the troposphere and the asymmetry of the convection cells. The parameters σ , r are the Prandtl and the Rayleigh numbers, while b is a ratio of critical parameters. A trajectory of the Lorenz (1963) attractor is shown in blue in Figure 1. The figure consists of 2000 points obtained by iterating the Lorenz equations with $dt \simeq 0.035$, $\sigma = 28$, $r = 10$, $b = 8/3$ with a Runge Kutta scheme of order 4. These are the classical parameters used by Lorenz (1963).

To study the effects of an external forcing, we now increase σ by $\sim 2\%$ with respect to the classical value, such that $\sigma = 28.5$ (magenta trajectory in Fig. 1). This new trajectory on average favors higher values of z , but the changes relative to the original trajectory depend on the point

¹a list of symbols used in the manuscript is provided in the supplementary material Table S1.

being considered: some points are not displaced, while some others are mapped elsewhere. Assuming no knowledge of the system other than the trajectories' paths, is there a way to determine that they belong to Lorenz attractors with different forcings? To answer this question we would need: i) to measure the dynamical properties of an ensemble of trajectories representing the two configurations; and ii) to estimate the distance between these trajectories and determine if the shift has changed the properties of the points in a detectable way. As a further complicating factor, we note that changes in natural systems are often gradual, and a sudden jump from one forcing to a different one is unlikely. A more realistic example discussing a Lorenz attractor with a smoothly varying forcing is presented in Section 2.

The atmospheric equivalent of a point on the Lorenz attractor is the ensemble of instantaneous fields describing the atmosphere at a time t . Here we study the atmospheric circulation over the North Atlantic and focus on a single field: the sea-level pressure (SLP). The SLP field reflects the major modes of variability affecting the North Atlantic (Hurrell 1995; Moore et al. 2013) and can further be used to diagnose a wealth of other atmospheric features, ranging from teleconnection patterns to storm track activity to atmospheric blocking e.g. (Rogers 1997; Comas-Bru and McDermott 2014). The trajectories of our dynamical systems are the successions of daily SLP fields from 26 CMIP5 models (Taylor et al. 2012) and the 20CRv2c reanalysis (Compo et al. 2011) over the period 1851 to 2000. The choice of the North Atlantic domain is motivated by the better observational coverage over the region in the first part of the analysis period compared to other regions of the globe (Krueger et al. 2013). In order to measure changes in the systems, one must be able to specify at each point (each day) the local (in phase-space, daily in time) dynamical properties and track their evolution. Recent contributions to dynamical systems analysis have proven that local properties of the trajectories are characterized by two quantities:

the local dimension and the stability of the field considered (Lucarini et al. 2016; Faranda et al. 2017b). They correspond respectively to the rarity and the typical persistence of the configuration. Faranda et al. (2017b) and Messori et al. (2017) have also shown that these two metrics can be connected to the predictability of a given atmospheric state and that their extremes match climate extremes.

In this work we will first assess whether the models and reanalysis present similar attractor properties over the full time period considered. To do this, we compute daily values for the dimension and stability of the SLP fields and study their average and extreme properties. Next, we quantify their changes over the analysis period. We then compare the changes seen in the models to those observed in the reanalysis.

2. Data & Methods

We use daily SLP model output from the historical simulations of 26 CMIP5 models (Table 1) available from the CMIP5 archive (Taylor et al. 2012). We then compare these simulations to the 20th Century Reanalysis (20CRv2c) dataset (Compo et al. 2011), studying both the ensemble mean and the 56 individual members. The analysis focuses on the region $22.5^{\circ}N - 70^{\circ}N$ and $80^{\circ}W - 50^{\circ}E$ over the period 1851-2000. SLP anomalies are defined as deviations from a daily climatology. For example, the SLP anomaly at a given geographical point on the 5th December 2000 is computed relative to the mean SLP value at that same location for all 5th Decembers over the analysis period.

In order to compute the dynamical systems metrics, we combine the statistical tools of extreme value theory with the results obtained by Freitas et al. (2010) for Poincaré recurrences. The pa-

rameters mentioned in the introduction (local dimension d and stability $1/\theta$) are computed for the points ζ on the attractor obtained as a sequence of states of the system. The dynamical indicators are linked to the probability p that a trajectory $x(t)$ explores a ball centered on ζ with diameter 2ε , i.e. the recurrence rate of the configuration ζ . We briefly outline the physical meaning of these quantities and the way they are computed below.

(i) *Local Dimension:* The Freitas et al. (2010) theorem and its modification in Lucarini et al. (2012) states that the probability p of entering a ball centered on ζ with a radius ε for chaotic attractors obeys a generalized Pareto distribution (Pickands III 1975). In order to compute such probability, we first calculate the series of distances $\delta(x(t), \zeta)$ between the point on the attractor ζ and all other points $x(t)$ on the trajectory. We then put a logarithmic weight on the time series of the distance:

$$g(x(t)) = -\log(\delta(x(t), \zeta)).$$

The reason for taking the logarithm is explained by Collet and Eckmann (2009): in the dynamical system set-up the negative logarithm increases the discrimination of small values of $\delta(x, y)$, which correspond to large values of $g(x(t))$. The probability of entering a ball of radius ε centered on ζ is now translated into the probability p of exceeding a threshold q . In the limit of an infinitely long trajectory, such probability is the exponential member of the generalized Pareto distribution:

$$p = \Pr(g(x(t)) > q, \zeta) \simeq \exp(-[x - \mu(\zeta)]/\beta(\zeta))$$

whose parameters μ and β depend on the point ζ chosen on the attractor. Remarkably, $\beta(\zeta) = 1/d(\zeta)$, where $d(\zeta)$ is the local dimension around the point ζ . This result has been applied to SLP fields in Faranda et al. (2017b) and also to surface temperature and precipitations

fields in Faranda et al. (2017a). In this paper we use the quantile 0.975 of the series $g(x(t))$ to determine q . We have further checked the stability of the results against reasonable changes in the quantile.

(ii) *Local Stability:* The probability p contains information about the geometry of the ball around ζ but provides no insight on the temporal evolution of the dynamics around ζ . In particular, it is interesting to know the mean residence time of the trajectory within the neighborhood of ζ . To measure this quantity, we employ the extremal index θ (Freitas et al. 2012; Faranda et al. 2016a), namely the inverse of this mean residence time. Heuristically, if the trajectory enters the neighborhood of ζ N times and at each time i the length of the cluster (that is, the number of successive timesteps when the trajectory is within the neighborhood) is τ_i , a simple estimate is: $\theta^{-1} = (1/N) \sum_i \tau_i$, such that θ varies between 0 and 1. The value $\theta = 0$ corresponds to a stable fixed point of the dynamics where the observation ζ is repeated infinite times (as for a pendulum left in its equilibrium position). A value of $\theta = 1$ indicates a point immediately leaving the neighborhood of ζ . Since θ is the inverse of a persistence time, it depends on the dt used. If dt is too large, the time dependence structure is hidden. The system will rarely persist more than one time-step within the neighbourhood of ζ , and θ will therefore be close to 1 at all points because of the cluster under-sampling. In Faranda et al. (2017b) it has been observed that θ for SLP fields over the North Atlantic varies between 0.3 and 0.5, when $dt = 1$ day. In this work we use the same dt . The extremal index is estimated using the likelihood estimator developed by Süveges (2007).

Figure 2 illustrates the meaning of the indicators for atmospheric flows: the local dimension d is the number of degrees of freedom needed to describe the dynamics of the system linearized

around the state ζ and is therefore proportional to the number of degrees of freedom of ζ . In the pictogram, $d = 2$ since we only consider two possible precursors to and evolutions of the state ζ . The inverse of the persistence time θ is linked to the probability that the trajectory follows a path where each field resembles those of the previous and subsequent days. We present two possibilities: i) along the green trajectory, the pattern changes every day so that $\theta(t) = 1$; ii) along the red trajectory, the same pattern is observed for three consecutive days so that $\theta = 1/3$.

In order to evaluate how close the d and θ of the different models are to those of the reanalysis, we adopt a number of distance metrics. The simplest metrics that can be defined are: $R(d) = \Delta(d)/\max(\Delta(d))$ and $R(\theta) = \Delta(\theta)/\max(\Delta(\theta))$, where Δ represents the difference between the median values of d or θ of each model and the 20CRv2c values. We further compute a global score:

$$R_{tot} = (R(d) + R(\theta))/2. \quad (2)$$

To check the validity of the global score, we compare R_{tot} with the Wasserstein distance $\mathcal{W}$, computed as described in Robin et al. (2017).

The Wasserstein distance $\mathcal{W}$ measures how two multivariate probability density functions μ and η differ from each other. Probability distributions are like distributions of mass (normalized to 1) across space. The Wasserstein distance is proportional to the minimal work that is needed to transport one distribution of mass into another. It is therefore rooted in optimal transport theory (Villani 2008; Santambrogio 2015). In practice, the two measures μ and η are discretized into multivariate histograms, which are sums of Dirac masses:

$$\mu = \sum_i \mu_i \delta_{x_i}, \quad \eta = \sum_j \eta_j \delta_{x_j}.$$

The distribution μ is transformed into η by the transport plan γ such that γ_{ij} transports the mass μ_i at x_i to η_j at x_j . Therefore the $\gamma_{ij} \geq 0$ verify:

$$\mu_i = \sum_j \gamma_{ij}, \quad \eta_j = \sum_i \gamma_{ij}, \quad \sum_{ij} \gamma_{ij} = 1.$$

If Γ is the set of all possible transport plans of μ to η , then the Wasserstein distance is:

$$\mathcal{W}(\mu, \eta) = \sqrt{\inf_{\gamma \in \Gamma} \sum_{ij} \gamma_{ij} \|x_i - x_j\|^2},$$

where $\|\cdot\|$ is the usual Euclidean distance.

Before beginning the analysis of the model and reanalysis data, it is necessary to determine if the dynamical properties are sensitive to the changes of the attractor due to continuous modifications of the underlying forcing and if these changes are statistically detectable. There are few theoretical results on non-stationary statistics of dynamical systems, as well as on non-stationary extreme value theory. Luckily, the recurrence approach we use here to estimate d and θ allows to bypass most of the technical difficulties linked to non-stationarity, because the dynamical properties are measured with respect to each individual state ζ of the attractor. If the change affects the neighborhood of a state, it will change its dynamical properties. If most of the states are affected by the changes in the dynamics, then the average dimension of the attractor and the average persistence will change accordingly.

In order to test this idea, we consider the Lorenz (1963) systems discussed in the Introduction and perform a set of 30 realizations (trajectories), where σ for each realization varies continuously over $28 < \sigma < 28.54$ according to: $\sigma(t) = \sigma(t - dt) + \delta\sigma$, with $\delta\sigma = 10^{-5}$. Each realization consists of 54000 iterations with time-step $dt = 0.02$, starting from random initial conditions in the basin of attraction. The number of time-steps and the dt are chosen to mimic the persistence properties of the SLP field over the North Atlantic, which displays a median θ of around 0.5 (see Section 3). If our methodology can indeed detect the gradual change from $\sigma = 28$ to $\sigma = 28.54$, then the d and θ distributions for the first and second half of the simulations should be significantly different. In order to provide a visually immediate picture, we show distributions of the medians of (d, θ) for each half of each simulation (Figure 3). It is straightforward to verify that the data forms two clouds of well-separated median centroids. This analysis therefore confirms that the indicators are sensitive to small changes in the attractor properties and that we can attempt to use them to detect long term changes in the dynamical properties of reanalysis and CMIP5 data.

3. Aggregate analysis of model and reanalysis attractors

We begin the analysis of the daily SLP fields from 1851 to 2000 by presenting the scatterplot of d versus θ for the ensemble mean of the 20CRv2c reanalysis (Figure 4). Hereafter we will call this run 20CR-EM. With respect to the computations done for the Lorenz (1963) system, ζ is now a daily SLP map and distances are computed using the euclidean metric at each grid point as in Faranda et al. (2017b). The average of d is proportional to the number of degrees of freedom needed to represent the systems' dynamics (this quantity is called attractor dimension in dynamical systems theory) while the average of θ is the inverse of the mean persistence time of a given SLP configuration. Maxima (minima) of d therefore correspond to the most complex (simple) SLP configurations. Maxima (minima) of θ correspond to the most unsteady (stable)

configurations (Messori et al. 2017; Faranda et al. 2017b). Panels A-D in Figure 4 show the
 composite SLP anomalies for dynamical extremes - namely days beyond the 0.975 and 0.025
 quantiles of the d and θ distributions. These closely – albeit not exactly – resemble the canonical
 North Atlantic weather regimes (Vautard 1990). In particular, maxima of θ (A) reproduce an
 Atlantic Ridge pattern, while minima of θ (B) correspond to a negative North Atlantic Oscillation
 (NAO) phase. Similarly, maxima of d (C) correspond to a blocking pattern and minima of d (D)
 to a positive NAO. This is in agreement with previous results from Faranda et al. (2017b). The
 patterns are stable for definitions of dynamical extremes up to the 20th and 80th percentiles of the
 relevant distributions, although the magnitude of the composite anomalies reduces when including
 days corresponding to lower percentiles (not shown). We note that the values of θ should not be
 compared directly to the persistence of the traditional weather regimes defined using clustering
 algorithms, as the requirement that the flow does not leave the neighborhood of the state ζ is a
 more restrictive condition than continued permanence within a given cluster. For example, intense
 or frequent mobile synoptic systems leading to substantial day-to-day fluctuations in sea-level
 pressure could cause the dynamics to leave the ball centered on the state ζ in phase space while
 remaining within a weather regime cluster. Indeed, if one considers the typical partition of the
 North Atlantic atmospheric variability into 4 weather regimes, the probability of being in one of
 them is of order 0.25, whereas the probability of being close to ζ is set by the threshold q , in our
 case 0.025 (see Section 2).

The 20CR-EM data analyzed above is constructed by averaging instantaneous fields from a 56-
 member ensemble of simulations. The ensemble is less constrained at the beginning of the period,
 when surface observations were scarce, than at the end. This implies that the 20CR-EM fields are
 smoother at the beginning of the period than at the end, because the latter are obtained averaging

over an ensemble with smaller differences between individual members. This may affect the effective number of degrees of freedom as measured by d . For this reason, we also measure d and θ for the 56 individual members and then average them to obtain a single daily value. We will refer to this quantity as the 20CRv2c mean of the ensemble (20CR-ME). Schematically:

- 20CR-EM indicates the daily dynamical properties computed for the 20CRv2c ensemble mean.
- 20CR-ME indicates the average of daily dynamical properties computed for each single ensemble member.

We can now compare the distributions of d, θ for 20CR-EM to d, θ for 20CR-ME. The scatterplot of d versus θ for the 20CR-ME is shown in Figure 5. A number of differences relative to the 20CR-EM appear (cf. Figure 4). First, the median value of d is lower for 20CR-EM than for 20CR-ME (Table 2), which indicates that the averaging of SLP fields in the Ensemble Mean has suppressed some degrees of freedom. Similarly, the ensemble mean has higher persistence (lower θ) because smoother fields tend to have slower variations. Although the numerical values of d and θ differ, the cross-correlation coefficient between the time series for the two data-sets are 0.93 and 0.97, respectively. This suggests that features such as the seasonality and interannual variability of the d and θ time series are preserved with the EM averaging. We next look at the SLP anomaly fields corresponding to the dynamical extremes of the 20CR-ME (Panels A-D in Fig. 5). We find similar patterns to those observed for 20CR-EM (Panels A-D in Fig. 5)). Indeed, 70% of d maxima from 20CR-EM match those from 20CR-ME, while this percentage is 61% for the minima. For θ maxima and minima we find values of 81% and 55%, respectively. We further investigate the differences between 20CR-EM and 20CR-ME by looking at the changes of the

dynamical properties over time in Section 4.

We next compare the (d, θ) bivariate histograms obtained for the 20CR-EM (Figure 6a) with those computed for the CMIP5 models. The (d, θ) histograms and the composite SLP anomalies corresponding to the d and θ extremes for all CMIP5 models analysed here are shown in Figures S1-S26 in the supplementary material. For conciseness, we limit our discussion to the models shown in (Figure 6c,e). Two different behaviors emerge: some of the models (e.g. CMCC-CMS, Figure 6c) yield a unimodal distribution resembling that obtained for the 20CR-EM; other models (e.g. the IPSL-CM5A, Figure 6e) show bimodal distributions. We find these different behaviors to be related to different seasonal cycles. In Figure 6 b,d,f, we plot the (d, θ) scatterplots for the same models by coloring each point according to the month of the year it occurs in. In the 20CR-EM and the CMCC-CMS model, the different seasons are spread across the cloud, although maxima of θ mostly occur in winter and the summer season is biased towards low d and θ . The IPSL-CM5A displays a much stronger seasonal discrimination, with two distinct (d, θ) clouds for the winter and for the summer seasons corresponding to the different modes of the bivariate histograms. This implies that both the bulk statistics and the extremes are modified by the seasonal cycle. For a more detailed discussion of the seasonality of the dynamical extremes, we refer the reader to Faranda et al. (2017b).

Given the variety of the possible behaviors, we will analyze separately the mean and the extreme behavior of the CMIP5 dynamical properties. We report the aggregate analysis in Table 2 and in Figure 7. In the latter figure, the colored numbers correspond to the median values for each model (numbered as in Table 1), MM (in black) to the CMIP5 multi-model mean, the number 1 to the 20CR-EM median and the blue dots to the medians of each of the 56 members of the 20CRv2c

ensemble. For 20CR-EM and 20CR-ME we also draw the ellipses whose semiaxes correspond to the standard deviation of the mean. Panel a) refers to the whole period whereas panels b) and c) show the analysis averaging over two different periods (1851-1978 and 1979-2000 respectively). We note that the median values of the 20CR-ME are so close to each other as to be almost indistinguishable, meaning that the individual members provide a coherent representation of the SLP field. As noted above, the 20CR-EM d and θ median values are both smaller than those of 20CR-ME. All of the models fall within the ellipse of 20CR-EM and most of the models within that of 20CR-ME. By inspecting Figure 7b,c it is evident that the distance between 20CR-EM and 20CR-ME reduces in the second period. However, the distance between 20CR-ME and models remains about constant. In Table 2 we provide the median values and standard deviations of d and θ for all the models and also present the distance metrics $R(d)$, $R(\theta)$ and R_{tot} between models and reanalysis as introduced in Section 2. In order to verify their robustness, we further compare R_{tot} to the Wasserstein distance $\mathcal{W}_2$, also defined in Section 2 (Figure 8). The two indicators provide very similar information (Pearson coefficient: $r_{pear} = 0.90$ and Spearman coefficient $r_{spear} = 0.85$ (von Storch and Zwiers 2002)); from now on we will therefore use the simpler R_{tot} when discussing our results. Indeed, the Wasserstein distance would be particularly complex to compute for the dataset analyzed here.

All the models are within one standard deviation of the 20CR-EM (Figure 7), while 5 are not within the 20CR-ME ellipse. Both Figure 7 and Table 2 further indicate that in many cases models with a higher horizontal resolution have median values closer to those of the reanalysis. We note, however, that the median values in d and θ are statistically different from those found in 20CR-EM for all models except models 24 and 26 for d and model 23 for θ when we consider the full 1851-2000 span. For the period 1851-1978, medians are statistically different except model

21 for d and model 12 for θ . For the second period (1979-2000) all medians are different except for models 13, 15, 25 for d and models 2, 6, 10, 14, 20, 22, 24, 25, 26 for θ .

For 20CR-ME, the medians are different for all models except model 13 in the period 1851-2000. For the period 1851-1978, all the medians are different. For the period 1979-2000 all the medians are different except for models 8 and 14 for d . The statistical significance is determined using the Wilcoxon ranksum test (von Storch and Zwiers 2002). The null hypothesis is that the variables are samples from continuous distributions with equal medians. A rejection of the null hypothesis therefore indicates a significant difference between the medians.

We also find that different models have different spreads in d and θ (Table 2), suggesting investigating the extremes of these quantities and their relation with the weather regimes found in 20CR-ME and 20CR-EM. A quantitative analysis is reported in Table S2 using the Root Mean Square Error (RMSE) between 20CR-EM, 20CR-ME and the CMIP5 SLP composite anomalies. In general, we find the B and D (NAO-like) patterns to have a higher RMSE whereas A and C are better represented. This is counterintuitive, since one might naively expect models to better reproduce low-dimensional rather than high-dimensional patterns. At the same time, the NAO is the dominant mode of variability in the North Atlantic region, which could explain why a higher RMSE is found in that region of the phase space. The B and D patterns also display stronger gradients than the A and C ones, suggesting that small displacement errors might lead to larger RMSE values for the former than for the latter.

4. Changes in the attractor properties

We now investigate whether the SLP's dynamical indicators have changed over time by computing 30-year moving time averages (denoted by $\langle \cdot \rangle$) for 20CR-EM, 20CR-ME (and separately for the 56 members) and the 26 CMIP5 models. These results are presented in Figure 9. The abscissa show the final year of each averaging window. In the left panels, we consider the full 150 years and subtract from each model/reanalysis the respective mean values of the indicators in 1880 (that is, a constant value for each timeseries given by the average over 1851-1880); in b,d) we focus on the last 35 years of data and subtract the values of 1979 (that is, averaged over 1950-1979). All timeseries therefore have a value of 0 on the year the baseline mean values are computed for. In panels e,f), we display the absolute values of d and θ , without subtracting any baseline values.

20CR-EM (red) shows an increase in d over time which is opposite in sign and an order of magnitude larger than the changes observed in 20CR-ME (blue), the single members (yellow), the CMIP5 multimodel mean (black) and the single CMIP5 models (color-scale, ordered by R_{tot} in Figure 9a). When focusing on the last 25 years of the datasets, the curves scale similarly and most of them show a moderate decrease of the dimension (Figure 9b). Figure 9e,f) shows that the spread of the 20CRv2C ensemble reduces with time, contrary to that of the CMIP5 models. The rapid increase of the dimension for 20CR-EM during the pre-satellite era may then be caused by changes in the variance of the ensemble members which, as discussed in Section 3 above, become more closely constrained as we approach the present day. In fact, the standard deviation of the ensemble is closely linked to the number of observations incorporated in the analysis (e.g. Krueger et al. (2013)). To test this hypothesis, we compute the daily values of the standard deviation of the 56 SLP fields and plot the moving average of such quantity against that of the

dimension of 20CR-EM (Figure 10). The expected near-linear relationship is found.

This points to the increase in $\langle d \rangle$ 20CR-EM not being a physical trend, but rather an artifact due to the scarcity of observations during the first part of the reanalysis period. However, from 1979 onwards, the variance stabilizes and both the 20CR-EM and 20CR-ME show a decrease in local dimension. This decrease is also observed in most of CMIP5 models, as reflected by the multi-model mean (Fig. 9b). Even though the models are ordered by increasing R_{tot} , it is evident that the *best* models for the overall period do not correspond to the models showing a decrease in time in the local dimension. It should be further noted that, when computing the R_{tot} statistics for different periods, the best models change and there are no clear *winners*.

The inverse persistence θ of 20CR-EM shows a similar behavior (Figure 9c). θ increases up to 1950 and then levels off. In terms of persistence time, this increase is small compared to the resolution of the datasets used for the analysis (1 day). 20CR-ME shows a similar, albeit weaker, trend, while the CMIP5 models mostly oscillate around zero. No systematic trends can be identified after 1970 (Figure 9d).

We conclude by noting that the changes found in CMIP5 models are mostly of the same sign as those found for 20CR-ME, but systematically smaller in magnitude. This is associated with a much larger spread between the different models than between the different members of the reanalysis, with a number of models showing trends of opposite sign to those of the multi-model mean.

5. Discussion and Conclusions

We have computed the instantaneous dynamical properties of the SLP fields over the North Atlantic Region for the 20CRv2c and the CMIP5 historical runs, over the period 1851-2000. The goal of our analysis was to assess whether different models with different physics and spatial resolutions quantitatively represent the *same* dynamical system and therefore possess attractors with similar characteristics. The metrics we use are the local dimension d and the inverse of the persistence time θ . As described in Faranda et al. (2017b), these two quantities give a complete characterization of the attractor of the system. To take into account the possibility that inhomogeneities in the assimilated data create artifacts in the 20CRv2c ensemble-mean SLP fields, we have computed two sets of dynamical properties. The first is composed of the dynamical properties of the ensemble-mean field (20CR-EM). The second is computed as the average of the instantaneous dynamical properties of each of the 56 ensemble members (20CR-ME).

When the whole analysis period is considered, we find that the models successfully capture many of the dynamical-system features identified in the reanalysis, such as the range and variability of the d and θ metrics. The SLP fields corresponding to extremes in d and θ are also similar across the models and reanalysis. At the same time, some models exaggerate the effects of the seasonal cycle on the dynamical indicators, and the statistical agreement in the median values of the metrics is generally poor. Models with higher horizontal resolutions mostly perform better.

To detect the changes in the attractor properties with time, we have analysed the 30-year moving averages of d and θ for the models and reanalysis. We find a number of interesting behaviors: up to 1950, the 20CR-EM shows an increase in d and θ due to the averaging among ensemble

members, whereas 20CR-ME shows a decrease in d and a slower increase in θ . After 1950, there is no trend in θ and a weakly decreasing trend in d for both datasets. The CMIP5 multi-model mean shows a weakly decreasing trend in d throughout the analysis period, in agreement with 20CR-ME, and no trend in θ . These results suggest that one should be very careful in using ensemble means for studying the atmospheric dynamics of the late 19th and beginning of the 20th century, even over the North Atlantic (Krueger et al. 2013; Ferguson and Villarini 2012, 2014; Alvarez-Castro et al. 2018). The 20CRv2c ensemble members are increasingly constrained by a growing number of SLP observations as one approaches the present day. This causes a decrease of the ensemble spread with time, since the system is more closely pinned to a specific manifold (the observations) without the possibility of exploring the full phase space. This is the root cause of the upward trend in the 20CR-EM local dimension, which therefore reflects a characteristic of the dataset as opposed to a physical change in the climate system. Our results are therefore in agreement with (Wang et al. 2013; Krueger et al. 2014; Dell'Aquila et al. 2016), who identified spurious trends in the 20CRv2C datasets in cyclone activity, storminess and synoptic variability. Finally, while in the latter part of the analysis period the 20CR-EM and 20CR-ME values of d converge (consistently with the fact that ensemble spread decreases, Fig. 10), this is not the case for θ (cf. Fig. 9e,f). This may be a signature of recurrent wave activity. Roughly periodic fluctuations do not change d , but will increase θ . In the 20CR-ME, atmospheric waves would therefore lower the persistence (and increase θ) more than in 20CR-EM, where they would be smoothed out. This would however have a relatively small impact on d . Preliminary analysis on a quasi-geostrophic 2-layer model supports this interpretation, although further analysis would be necessary to demonstrate this is the case for the 20CR dataset.

The next natural question is whether we can trust the results obtained for single reanalysis ensemble members. We believe that the answer is affirmative, as: i) the dataset has a sufficiently high horizontal resolution to obtain a good estimate of the local dimension distribution (Faranda et al. 2017b) and ii) we focus here on the North Atlantic sector, which can be expected to perform better than elsewhere since most of the observations used to constrain 20CRv2c in the first part of the dataset are located in Europe or eastern North America (Cram et al. 2015). We therefore argue that the results obtained for the 20CR-ME and the multimodel ensemble are valuable, and that the decrease in dimension with time is a real and interesting feature of the atmospheric dynamics which merits a more detailed analysis in further studies. In fact, a decreasing dimension implies a more predictable atmosphere. This conclusion needs to be corroborated by looking at other reanalyses of the modern period that use a fully 3-d set of observations, and will be the subject of a forthcoming study.

We also stress that there is a much stronger dynamical coherence within the 20CRv2C ensemble than among CMIP5 models. In fact, the spread in d and θ among 20CR members throughout the analysis period is significantly smaller than the differences among models (Figure 9). Considering a single member of the 20CRv2C ensemble would still provide good estimates of d and θ , but the same would not always be true of a randomly chosen model. The CMIP5 model analysis further suggests that differences in model set-up, resolution and physics affect the trends in dynamical systems metrics more than the average properties, so that d increases for some models contrary to the 20CR-ME.

As a final caveat we note that our analysis does not attempt to separate the forced variability from natural low-frequency oscillations and that, especially during the first part of the analysis

period, it is unclear whether the greenhouse forcing can be clearly discerned above the background
"climate noise" (Paeth et al. 1999; Lyu et al. 2015). We must therefore take into account the
possibility that the data's internal variability dominates over the long-term forcing trends for the
time period considered.

Acknowledgments. P.Yiou and D. Faranda were supported by ERC grant No. 338965, M.C.
Alvarez-Castro was supported by Swedish Research Council grant No. C0629701. G. Messori was
supported by Swedish Research Council grant No. 2016-03724 and a grant from the Department
of Meteorology of Stockholm University.

References

- Alvarez-Castro, M. C., D. Faranda, and P. Yiou, 2018: Atmospheric dynamics leading to west
european summer hot temperatures since 1851. *Complexity*, **2018**, doi:10.1155/2018/2494509.
- Cassou, C., and J. Cattiaux, 2016: Disruption of the european climate seasonal clock in a warming
world. *Nature Climate Change*, **6 (6)**, 589–594.
- Collet, P., and J.-P. Eckmann, 2009: *Iterated maps on the interval as dynamical systems*. Springer
Science & Business Media.
- Comas-Bru, L., and F. McDermott, 2014: Impacts of the ea and sca patterns on the european
twentieth century nao–winter climate relationship. *Quarterly Journal of the Royal Meteorolog-
ical Society*, **140 (679)**, 354–363.
- Compo, G. P., and Coauthors, 2011: The twentieth century reanalysis project. *Quarterly Journal
of the Royal Meteorological Society*, **137 (654)**, 1–28.

- Cram, T. A., and Coauthors, 2015: The international surface pressure databank version 2. *Geoscience Data Journal*, **2** (1), 31–46, doi:10.1002/gdj3.25, URL <http://dx.doi.org/10.1002/gdj3.25>.
- Dell’Aquila, A., and Coauthors, 2016: Benchmarking northern hemisphere midlatitude atmospheric synoptic variability in centennial reanalysis and numerical simulations. *Geophysical Research Letters*, **43** (10), 5442–5449.
- Faranda, D., M. C. Alvarez-Castro, and P. Yiou, 2016a: Return times of hot and cold days via recurrences and extreme value theory. *Climate Dynamics*, 1–13, doi:10.1007/s00382-016-3042-6.
- Faranda, D., G. Masato, N. Moloney, Y. Sato, F. Daviaud, B. Dubrulle, and P. Yiou, 2016b: The switching between zonal and blocked mid-latitude atmospheric circulation: a dynamical system perspective. *Climate Dynamics*, **47** (5-6), 1587–1599.
- Faranda, D., G. Messori, M. C. Alvarez-Castro, and P. Yiou, 2017a: Dynamical properties and extremes of northern hemisphere climate fields over the past 60 years. *Nonlinear Processes in Geophysics*, **24** (4), 713.
- Faranda, D., G. Messori, and P. Yiou, 2017b: Dynamical proxies of north atlantic predictability and extremes. *Scientific Reports*, **7**, 41 278.
- Ferguson, C. R., and G. Villarini, 2012: Detecting inhomogeneities in the twentieth century reanalysis over the central united states. *Journal of Geophysical Research: Atmospheres*, **117** (D5), n/a–n/a, doi:10.1029/2011JD016988, URL <http://dx.doi.org/10.1029/2011JD016988>, d05123.
- Ferguson, C. R., and G. Villarini, 2014: An evaluation of the statistical homogeneity of the twentieth century reanalysis. *Climate Dynamics*, **42** (11), 2841–2866, doi:10.1007/s00382-013-1996-1, URL <http://dx.doi.org/10.1007/s00382-013-1996-1>.

- 475 Field, C. B., 2012: *Managing the risks of extreme events and disasters to advance climate change*
 476 *adaptation: special report of the intergovernmental panel on climate change*. Cambridge Uni-
 477 versity Press.
- 478 Freitas, A. C. M., J. M. Freitas, and M. Todd, 2010: Hitting time statistics and extreme value
 479 theory. *Probability Theory and Related Fields*, **147** (3-4), 675–710.
- 480 Freitas, A. C. M., J. M. Freitas, and M. Todd, 2012: The extremal index, hitting time statistics and
 481 periodicity. *Advances in Mathematics*, **231** (5), 2626–2665.
- 482 Hurrell, J. W., 1995: Decadal trends in the north atlantic oscillation: Regional tem-
 483 peratures and precipitation. *Science*, **269** (5224), 676–679, doi:10.1126/science.269.5224.
 484 676, URL <http://science.sciencemag.org/content/269/5224/676>, <http://science.sciencemag.org/content/269/5224/676.full.pdf>.
- 486 Kay, J., and Coauthors, 2015: The community earth system model (cesm) large ensemble project:
 487 A community resource for studying climate change in the presence of internal climate variabil-
 488 ity. *Bulletin of the American Meteorological Society*, **96** (8), 1333–1349.
- 489 Krueger, O., F. Feser, L. Bärring, E. Kaas, T. Schmith, H. Tuomenvirta, and H. von Storch, 2014:
 490 Comment on trends and low frequency variability of extra-tropical cyclone activity in the en-
 491 semble of twentieth century reanalysis by xiaolan l. wang, y. feng, gp compo, vr swail, fw
 492 zwiers, rj allan, and pd sardeshmukh, climate dynamics, 2012. *Climate dynamics*, **42** (3-4),
 493 1127–1128.
- 494 Krueger, O., F. Schenk, F. Feser, and R. Weisse, 2013: Inconsistencies between long-term trends
 495 in storminess derived from the 20cr reanalysis and observations. *Journal of Climate*, **26** (3),
 496 868–874, doi:10.1175/JCLI-D-12-00309.1.

- 497 Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20** (2),
498 130–141.
- 499 Lucarini, V., D. Faranda, G. Turchetti, and S. Vaienti, 2012: Extreme value theory for singular
500 measures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **22** (2), 023 135.
- 501 Lucarini, V., and Coauthors, 2016: *Extremes and recurrence in dynamical systems*. John Wiley &
502 Sons.
- 503 Lyu, K., X. Zhang, J. A. Church, and J. Hu, 2015: Quantifying internally generated and ex-
504 ternally forced climate signals at regional scales in cmip5 models. *Geophysical Research*
505 *Letters*, **42** (21), 9394–9403, doi:10.1002/2015GL065508, URL [http://dx.doi.org/10.1002/](http://dx.doi.org/10.1002/2015GL065508)
506 2015GL065508, 2015GL065508.
- 507 Messori, G., R. Caballero, and D. Faranda, 2017: A dynamical systems approach to studying
508 midlatitude weather extremes. *Geophysical Research Letters*, **44** (7), 3346–3354.
- 509 Moore, G., I. A. Renfrew, and R. S. Pickart, 2013: Multidecadal mobility of the north atlantic
510 oscillation. *Journal of Climate*, **26** (8), 2453–2466.
- 511 Paeth, H., A. Hense, R. Glowienka-Hense, S. Voss, and U. Cubasch, 1999: The north at-
512 lantic oscillation as an indicator for greenhouse-gas induced regional climate change. *Cli-*
513 *mate Dynamics*, **15** (12), 953–960, doi:10.1007/s003820050324, URL [http://dx.doi.org/10.](http://dx.doi.org/10.1007/s003820050324)
514 1007/s003820050324.
- 515 Pickands III, J., 1975: Statistical inference using extreme order statistics. *the Annals of Statistics*,
516 119–131.
- 517 Robin, Y., P. Yiou, and P. Naveau, 2017: Detecting changes in forced climate attractors with
518 wasserstein distance. *Nonlinear Processes in Geophysics*, **24** (3), 393.

- 519 Rogers, J. C., 1997: North atlantic storm track variability and its association to the north atlantic
520 oscillation and climate variability of northern europe. *Journal of Climate*, **10** (7), 1635–1647,
521 doi:10.1175/1520-0442(1997)010<1635:NASTVA>2.0.CO;2.
- 522 Santambrogio, F., 2015: *Optimal Transport for Applied Mathematicians*, Vol. 87. Birkhäuser
523 Basel, XXVII, 353 pp.
- 524 Shepherd, T. G., 2014: Atmospheric circulation as a source of uncertainty in climate change
525 projections. *Nature Geoscience*, **7** (10), 703–708.
- 526 Süveges, M., 2007: Likelihood estimation of the extremal index. *Extremes*, **10** (1), 41–55.
- 527 Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment
528 design. *Bulletin of the American Meteorological Society*, **93** (4), 485–498.
- 529 Vautard, R., 1990: Multiple weather regimes over the north atlantic: Analysis of precursors and
530 successors. *Monthly Weather Review*, **118** (10), 2056–2081.
- 531 Villani, C., 2008: *Optimal transport: old and new*, Vol. 338. Springer Science & Business Media.
- 532 von Storch, H., and F. W. Zwiers, 2002: Statistical analysis in climate research. Citeseer.
- 533 Wang, X. L., Y. Feng, G. Compo, V. Swail, F. Zwiers, R. Allan, and P. Sardeshmukh, 2013: Trends
534 and low frequency variability of extra-tropical cyclone activity in the ensemble of twentieth
535 century reanalysis. *Climate Dynamics*, **40** (11-12), 2775–2800.

LIST OF TABLES

Table 1.	List of CMIP5 Models analysed and 20CRv2c reanalysis from 1851 to 2001. Models are approximately ordered by increasing horizontal resolution.	27
Table 2.	List of median values and standard deviations of dimension d and inverse persistence θ for 20CR-EM, 20CR-ME and the 26 CMIP5 models. The table further presents the relative distances $R(d) = \delta(d)/\max(\delta(d))$ and $R(\theta) = \delta(\theta)/\max(\delta(\theta))$ normalized with respect to the farthest model and the global score $R_{tot} = (R(d) + R(\theta))/2$	28

TABLE 1. List of CMIP5 Models analysed and 20CRv2c reanalysis from 1851 to 2001. Models are approximately ordered by increasing horizontal resolution.

No. ¹	Model	Institution/ID	Country	Resolution ²
0	20CRv2c ME	NOAA-CIRES	USA	2 x 2
1	20CRv2c EM	NOAA-CIRES	USA	2 x 2
2	CMCC-CESM	Centro Euro-Mediterraneo sui Cambiamenti Climatici	Italy	3.75x3.75
3	CanESM2	Canadian Centre for Climate Modelling and Analysis, CCCMa	Canada	2.81x 2.79
4	MIROC-ESM-CHEM	MIROC ⁴	Japan	2.81x 2.79
5	MIROC-ESM	MIROC ⁴	Japan	2.81x 2.79
6	BCC-CSM1-1	Beijing Climate Center	China	2.81x 2.79
7	IPSL-CM5B-LR	Institute Pierre Simon Laplace, IPSL	France	3.75x1.89
8	NorESM1-M	Norwegian Climate Center	Norway	2.5x1.89
9	FGOALS-2	Institute of Atmospheric Physics, Chinese Academy of Sciences	China	2.81x2.81
10	MPI-ESM-P	Max Planck Institute for Meteorology, MPI	Germany	1.87x1.87
11	MPI-ESM-LR	Max Planck Institute for Meteorology, MPI	Germany	1.87x1.87
12	CSIRO-MK3-6-0	CSIRO-BOM ⁵	Australia	1.87x1.87
13	CMCC-CMS	Centro Euro-Mediterraneo sui Cambiamenti Climatici	Italy	1.87x1.87
14	MPI-ESM-MR	Max Planck Institute for Meteorology, MPI	Germany	1.87x1.87
15	IPSL-CM5A-MR	Institute Pierre Simon Laplace, IPSL	France	2.5x1.26
16	INM-CM4	Institute for Numerical Mathematics, INM	Russia	2x1.5
17	ACCESS 1-0	CSIRO-BOM ⁵	Australia	1.87x1.25
18	MIROC5	MIROC ⁴	Japan	1.40x1.40
19	CNRM-CM5	CNRM-CERFACS ³	France	1.40x1.40
20	MRI-ESM1	Meteorological Research Institute, MRI	Japan	1.125x1.125
21	BCC-CSM1-M	Beijing Climate Center	China	1.125x1.125
22	MRI-CGCM3	Meteorological Research Institute, MRI	Japan	1.125x1.125
23	EC-EARTH	Danish Meteorological Institute, DMI	Denmark	1.125x1.125
24	CESM1-FASTCHEM	Community Earth System Model Contributors, NCAR	USA	1.25x0.94
25	CESM1-CAM5	Community Earth System Model Contributors, NCAR	USA	1.25x0.94
26	CESM1-BGC	Community Earth System Model Contributors, NCAR	USA	1.25x0.94
27	CCSM4	National Center for Atmospheric Research, NCAR	USA	1.25x0.94

¹ Order by increasing horizontal resolution

² Longitude x Latitude (°)

³ Centre National de Recherches Meteorologiques - Centre Europeen de Recherche et de Formation Avancee en Calcul Scientifique

⁴ Atmosphere and Ocean Research Institute (University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology

⁵ Commonwealth Scientific and Industrial Research Organisation(CSIRO), Bureau of Meteorology(BOM)

TABLE 2. List of median values and standard deviations of dimension d and inverse persistence θ for 20CR-EM, 20CR-ME and the 26 CMIP5 models. The table further presents the relative distances $R(d) = \delta(d)/\max(\delta(d))$ and $R(\theta) = \delta(\theta)/\max(\delta(\theta))$ normalized with respect to the farthest model and the global score $R_{tot} = (R(d) + R(\theta))/2$.

Model	median(d)	median(θ)	std(d)	std(θ)	$R(d)$ EM	$R(\theta)$ EM	R_{tot} EM	$R(d)$ ME	$R(\theta)$ ME	R_{tot} ME
00-20CR-ME	12.26	0.53	1.85	0.05						
01-20CR-EM	11.56	0.50	2.06	0.06						
02-CMCC-CESM	12.22	0.51	1.90	0.06	0.54	0.13	0.33	0.04	0.23	0.14
03-CanESM2	11.99	0.51	2.06	0.07	0.35	0.02	0.19	0.34	0.30	0.32
04-MIROC-ESM-CHEM	12.54	0.47	1.87	0.06	0.80	0.75	0.77	0.35	0.83	0.59
05-MIROC-ESM	12.48	0.47	1.95	0.05	0.75	0.76	0.76	0.27	0.84	0.56
06-BCC-CSM1	12.12	0.51	1.95	0.05	0.46	0.12	0.29	0.16	0.24	0.20
07-IPSL-CM5B	12.73	0.46	1.79	0.06	0.95	0.93	0.94	0.58	0.95	0.77
08-NorESM1-M	12.12	0.48	2.27	0.06	0.46	0.44	0.45	0.18	0.62	0.40
09-FGOALS-S2	11.63	0.45	1.89	0.06	0.06	1.00	0.53	0.78	1.00	0.89
10-MPI-ESM-P	12.17	0.51	1.82	0.07	0.50	0.06	0.28	0.11	0.28	0.19
11-MPI-ESM-LR	12.13	0.51	1.91	0.05	0.47	0.14	0.30	0.16	0.23	0.19
12-CSIRO-MK3-6-0	12.66	0.50	1.91	0.05	0.90	0.02	0.46	0.50	0.34	0.42
13-CMCC-CMS	11.95	0.52	2.14	0.06	0.32	0.22	0.27	0.38	0.17	0.28
14-MPI-ESM-MR	12.09	0.51	1.86	0.06	0.43	0.09	0.26	0.21	0.26	0.24
15-IPSL-CM5A	11.86	0.48	1.90	0.06	0.25	0.51	0.38	0.49	0.67	0.58
16-INM-CM4	12.79	0.47	2.14	0.07	1.00	0.70	0.85	0.65	0.80	0.72
17-ACCESS '1-0	11.74	0.49	1.92	0.06	0.15	0.31	0.23	0.64	0.53	0.58
18-MIROC5	12.58	0.49	2.00	0.07	0.83	0.33	0.58	0.39	0.54	0.47
19-CNRM-CM5	12.36	0.47	2.08	0.05	0.65	0.59	0.62	0.12	0.72	0.42
20-MRI-ESM1	11.72	0.51	2.09	0.05	0.13	0.12	0.13	0.67	0.23	0.45
21-BCC-CSM1-M	11.45	0.55	2.00	0.07	0.09	0.91	0.50	1.00	0.30	0.65
22-MRI-CGCM3	11.74	0.51	1.66	0.07	0.15	0.07	0.11	0.64	0.27	0.45
23-EC-EARTH	11.87	0.50	1.99	0.07	0.26	0.01	0.13	0.47	0.33	0.40
24-CESM1-FASTCHEM	11.56	0.51	1.90	0.06	0.00	0.18	0.09	0.86	0.19	0.53
25-CESM1-CAM5	11.88	0.51	1.83	0.07	0.26	0.13	0.20	0.47	0.23	0.35
26-CESM1-BGC	11.53	0.51	2.01	0.06	0.02	0.11	0.06	0.90	0.24	0.57
27-CCSM4	11.57	0.51	1.84	0.07	0.01	0.12	0.07	0.85	0.24	0.54

LIST OF FIGURES

- Fig. 1.** Two realizations of the Lorenz attractor. Blue: classic attractor with: $dt \simeq 0.035$, $\sigma = 28$, $r = 10$, $b = 8/3$; Violet: $\sigma = 28.5$ 31
- Fig. 2.** Schematic representation of the dynamical indicators on fictive SLP maps. H indicates high and L low pressure systems; the arrows mimic horizontal wind fields. The local dimension d is proportional to the number of possible configurations preceding and following the day being analyzed (here $d = 2$). θ is the inverse of the persistence time of a given configuration. If the pattern persist for three days (red path), then $\theta = 1/3$. If the patterns change every day (green path), then $\theta = 1$ 32
- Fig. 3.** Medians of d and θ (dots) and standard deviation (semiaxes of the ellipses) for 30 realizations of the Lorenz attractor with σ varying continuously over $28 < \sigma < 28.54$. The blue markers correspond to the first half of each simulation σ_1 ; the red markers to the second half σ_2 . The asterisks mark the medians of the ensemble of realizations. 33
- Fig. 4.** Scatterplot of the daily values of instantaneous dimension d and inverse persistence θ for the 20CRv2c ensemble-mean SLP (20CR-EM). The straight black lines mark the 0.025 and 0.975 quantiles of d and θ . The composite anomalies in SLP obtained by averaging the days beyond the quantiles can be associated with known weather regimes: A) Atlantic Ridge (maxima of θ), B) NAO- (minima of θ), C) Blocking (maxima of d), D) NAO+ (minima of d). The black contours in panels A-D indicate regions where at least 2/3 of the composite members display SLP anomalies of the same sign. 34
- Fig. 5.** Scatterplot of the daily values of instantaneous dimension d and inverse persistence θ for the SLP fields of the 56 individual 20CRv2c members (20CR-ME). The straight black lines mark the 0.025 and 0.975 quantiles of d and θ . The composite anomalies in SLP obtained by averaging the days beyond the quantiles can be associated with known weather regimes: A) Atlantic Ridge (maxima of θ), B) NAO- (minima of θ), C) Blocking (maxima of d), D) NAO+ (minima of d). The black contours in panels A-D indicate regions where at least 2/3 of the composite members display SLP anomalies of the same sign. 35
- Fig. 6.** (d, θ) bivariate histograms (a,c,e) and scatter-plots (b,d,f) for the 20CR-EM reanalysis (a,b), CMCC-CMS (c,d) and IPSL-CM5A models (e,f). The color scales in (a,c,e) indicate the frequency of observations in number of days. The color scales in (b,d,f) indicate the month of the observation and show the dependence of the (d, θ) diagrams on the seasonal cycle. 36
- Fig. 7.** Comparison between the 56 20CR-ME median values of (d, θ) (blue points whose average is denoted by 0), the 20CR-EM (in red and numbered by 1) and all the CMIP5 models (progressive numbers 2-27, see table 1 for details). The CMIP5 multi-model mean is marked by MM (black). The semiaxes of the two ellipses represent one standard deviation of d and θ for 20CR-EM (red) and 20CR-ME (blue). (a) 1851-2000 (b) 1851-1978 and (c) 1979-2000 37
- Fig. 8.** Comparison between R_{tot} values (black) and Wasserstein distances $\mathscr{W}$ (blue) between the (d, θ) of 20CR-EM (a) and 20CR-ME (b) and of the CMIP5 models. The distances are normalized by the maximum value. 38
- Fig. 9.** 30-year moving averages $\langle \cdot \rangle$ of instantaneous dimension d and inverse persistence θ minus the respective mean values over 1851-1880 (a,c) and 1950-1979 (b,d). Note that labels on the abscissa mark the last year in each 30-year averaging window. (a,b): local dimension d ; (c,d): inverse persistence θ . Red: 20CR-EM; Blue: 20CR-ME; yellow: single 20CRv2c members; Black: CMIP5 multimodel mean; Color scale: single CMIP5 models ordered

by R_{tot} . Panels (e) and (f) shows the absolute values of the 30-year moving averages of instantaneous dimension d (e) and inverse persistence θ (f). 39

Fig. 10. 30-year moving average of the mean standard deviation of SLP fields across the 20CRv2c ensemble $\langle std(SLP) \rangle$ versus $\langle d \rangle$ for 20CR-EM. The color scale shows time. Values in Pa. . . . 40

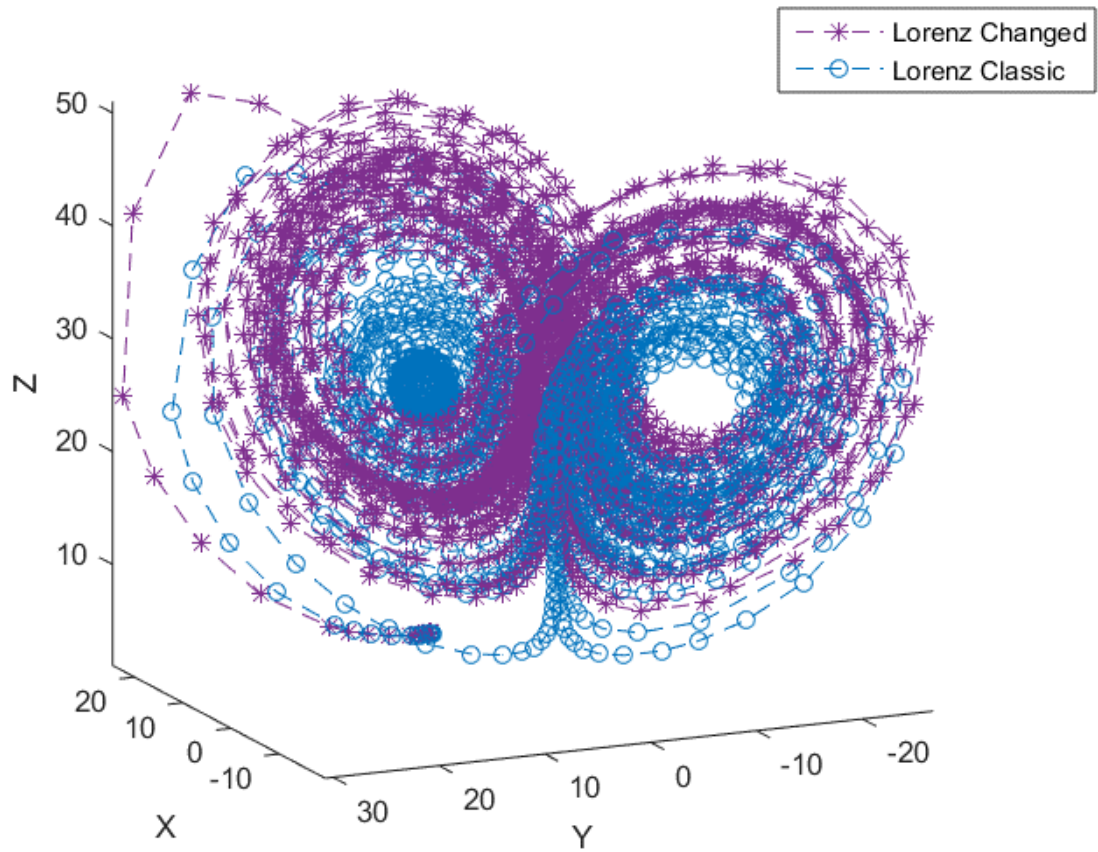


FIG. 1. Two realizations of the Lorenz attractor. Blue: classic attractor with: $dt \simeq 0.035$, $\sigma = 28$, $r = 10$,
 $b = 8/3$; Violet: $\sigma = 28.5$

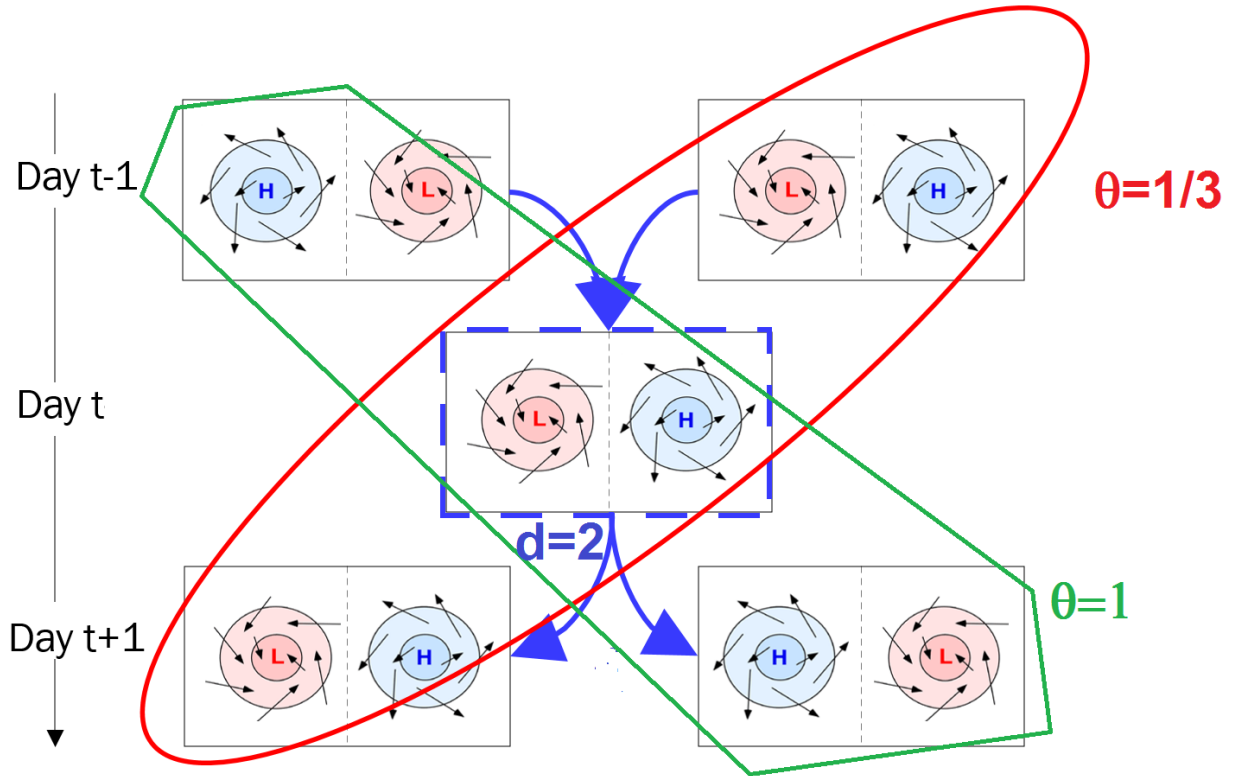


FIG. 2. Schematic representation of the dynamical indicators on fictive SLP maps. H indicates high and L low pressure systems; the arrows mimic horizontal wind fields. The local dimension d is proportional to the number of possible configurations preceding and following the day being analyzed (here $d = 2$). θ is the inverse of the persistence time of a given configuration. If the pattern persists for three days (red path), then $\theta = 1/3$. If the patterns change every day (green path), then $\theta = 1$.

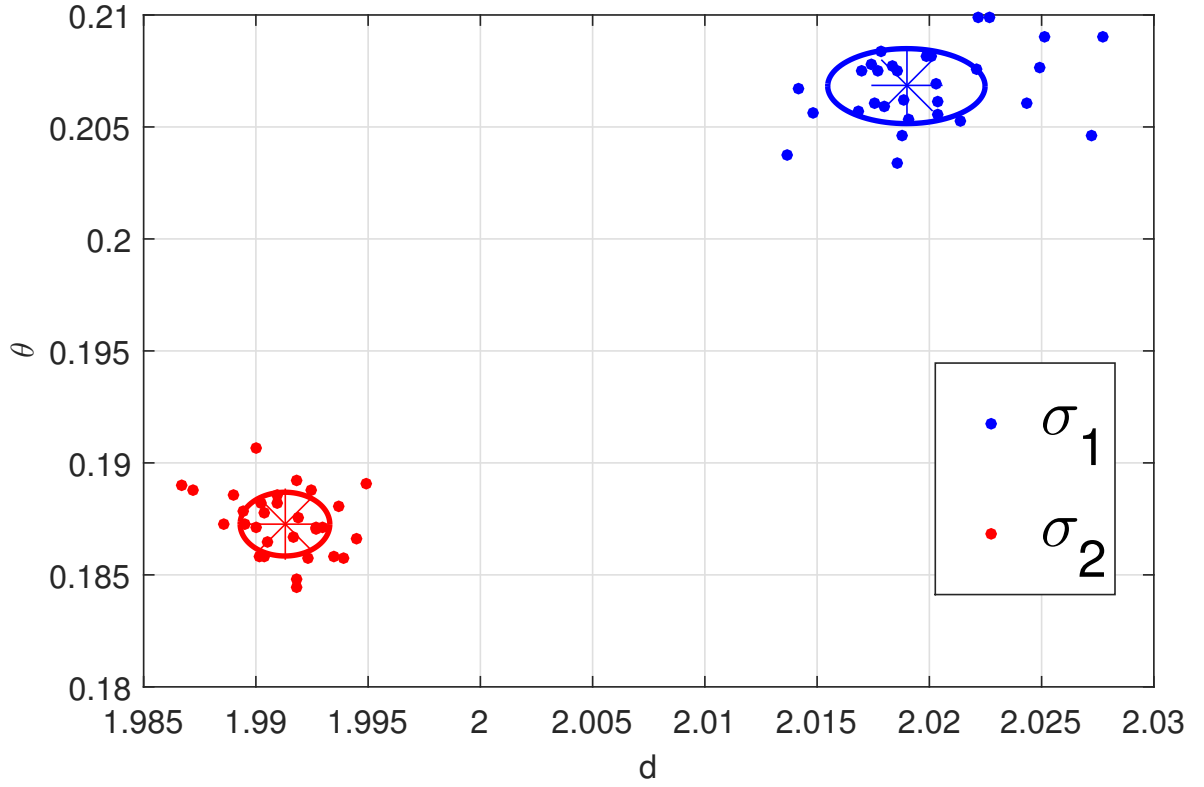


FIG. 3. Medians of d and θ (dots) and standard deviation (semiaxes of the ellipses) for 30 realizations of the Lorenz attractor with σ varying continuously over $28 < \sigma < 28.54$. The blue markers correspond to the first half of each simulation σ_1 ; the red markers to the second half σ_2 . The asterisks mark the medians of the ensemble of realizations.

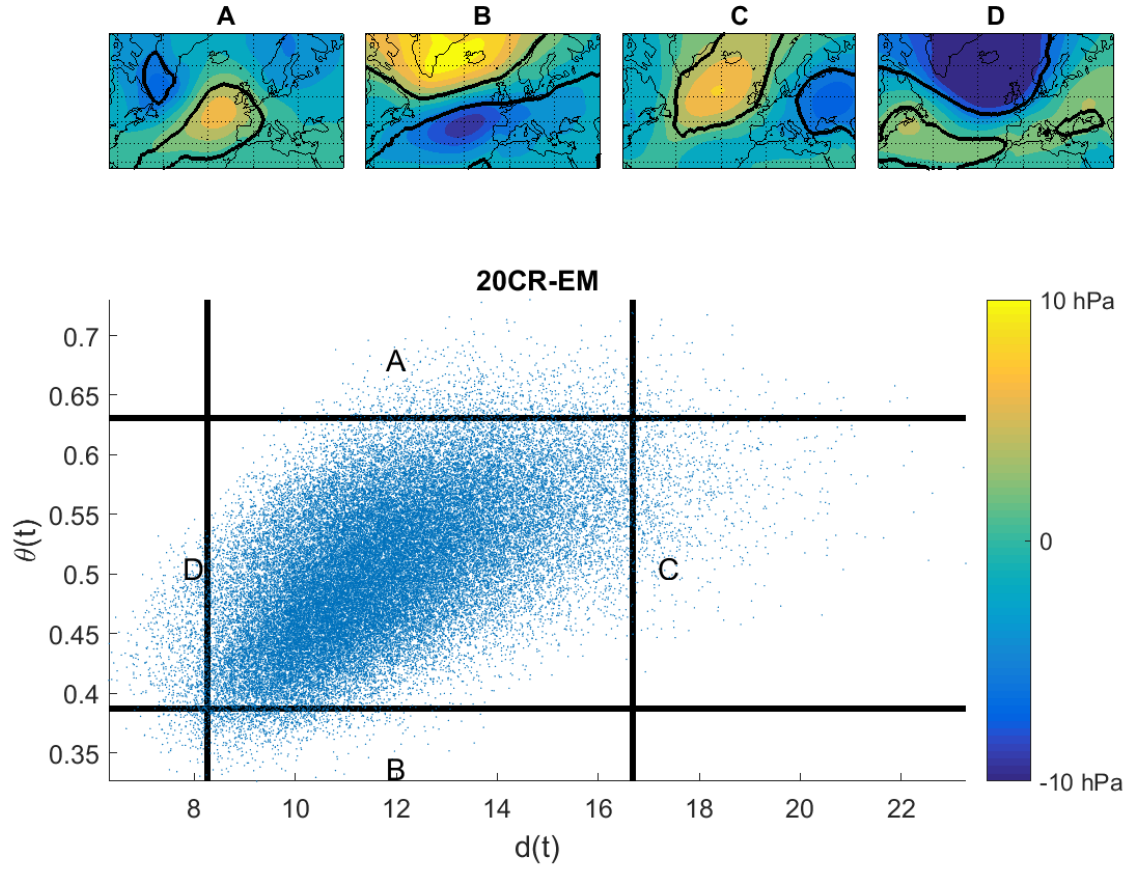
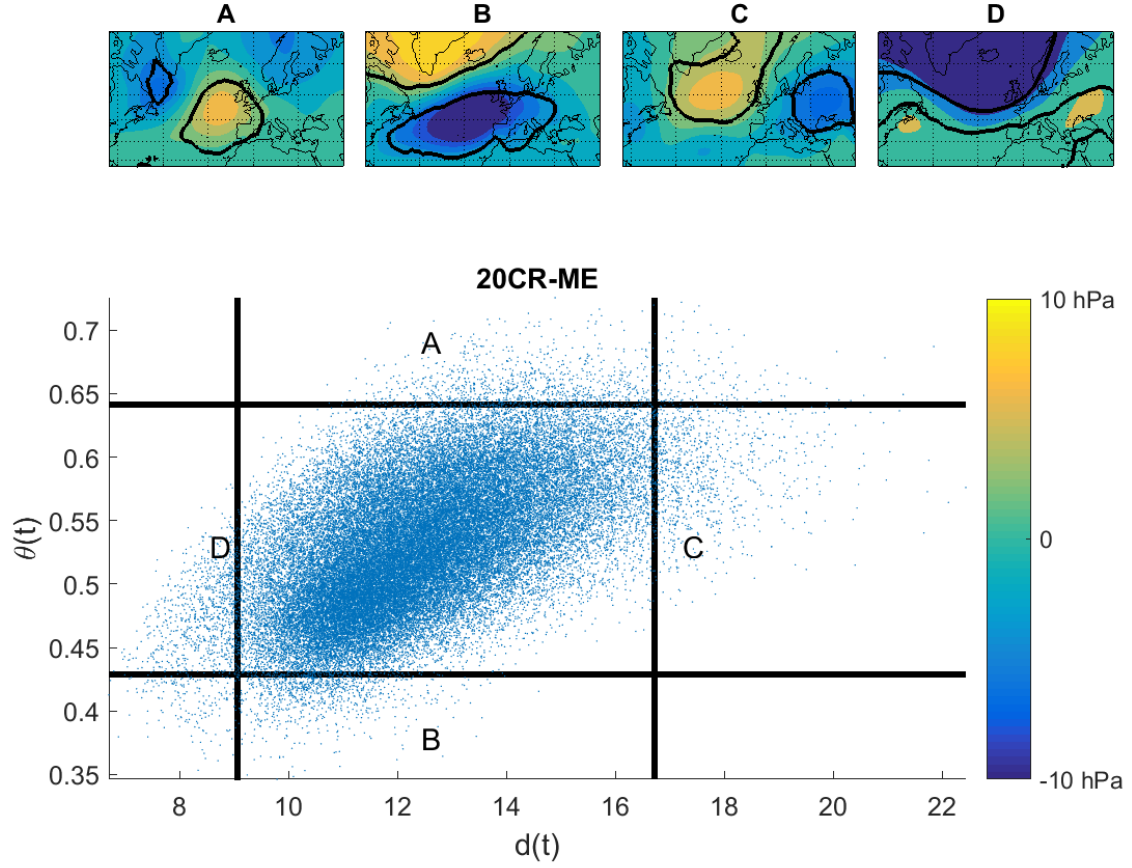


FIG. 4. Scatterplot of the daily values of instantaneous dimension d and inverse persistence θ for the 20CRv2c ensemble-mean SLP (20CR-EM). The straight black lines mark the 0.025 and 0.975 quantiles of d and θ . The composite anomalies in SLP obtained by averaging the days beyond the quantiles can be associated with known weather regimes: A) Atlantic Ridge (maxima of θ), B) NAO- (minima of θ), C) Blocking (maxima of d), D) NAO+ (minima of d). The black contours in panels A-D indicate regions where at least 2/3 of the composite members display SLP anomalies of the same sign.



613 FIG. 5. Scatterplot of the daily values of instantaneous dimension d and inverse persistence θ for the SLP
 614 fields of the 56 individual 20CRv2c members (20CR-ME). The straight black lines mark the 0.025 and 0.975
 615 quantiles of d and θ . The composite anomalies in SLP obtained by averaging the days beyond the quantiles
 616 can be associated with known weather regimes: A) Atlantic Ridge (maxima of θ), B) NAO- (minima of θ), C)
 617 Blocking (maxima of d), D) NAO+ (minima of d). The black contours in panels A-D indicate regions where at
 618 least 2/3 of the composite members display SLP anomalies of the same sign.

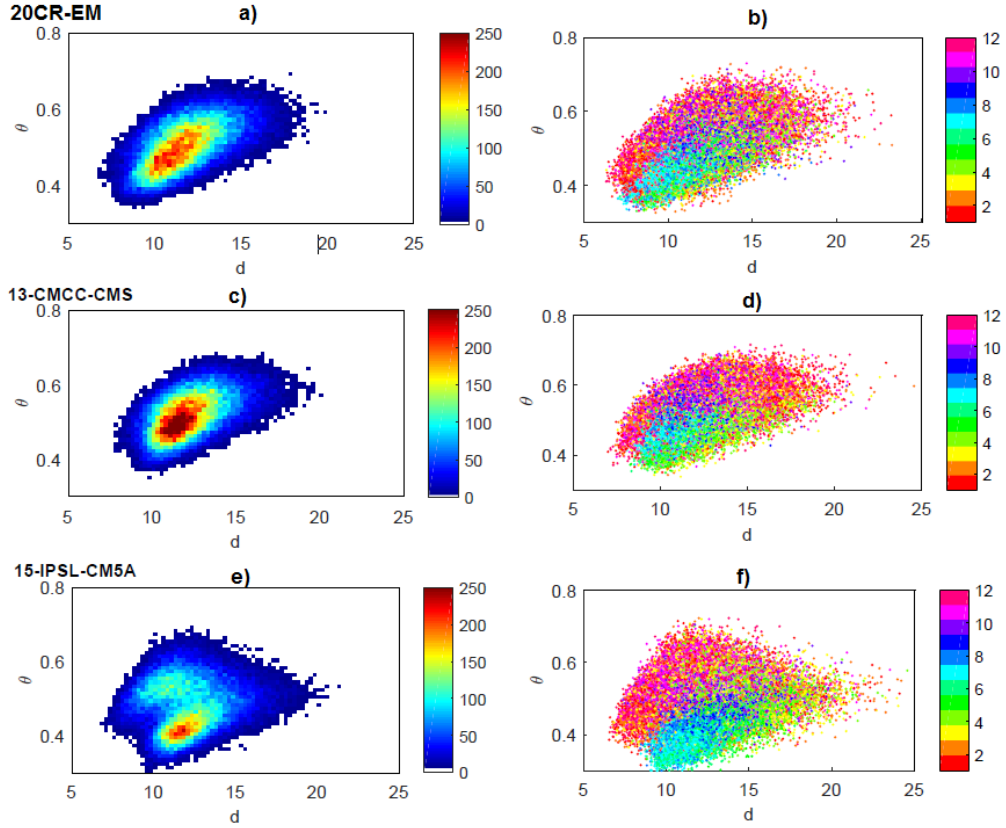


FIG. 6. (d, θ) bivariate histograms (a,c,e) and scatter-plots (b,d,f) for the 20CR-EM reanalysis (a,b), CMCC-
CMS (c,d) and IPSL-CM5A models (e,f). The color scales in (a,c,e) indicate the frequency of observations in
number of days. The color scales in (b,d,f) indicate the month of the observation and show the dependence of
the (d, θ) diagrams on the seasonal cycle.

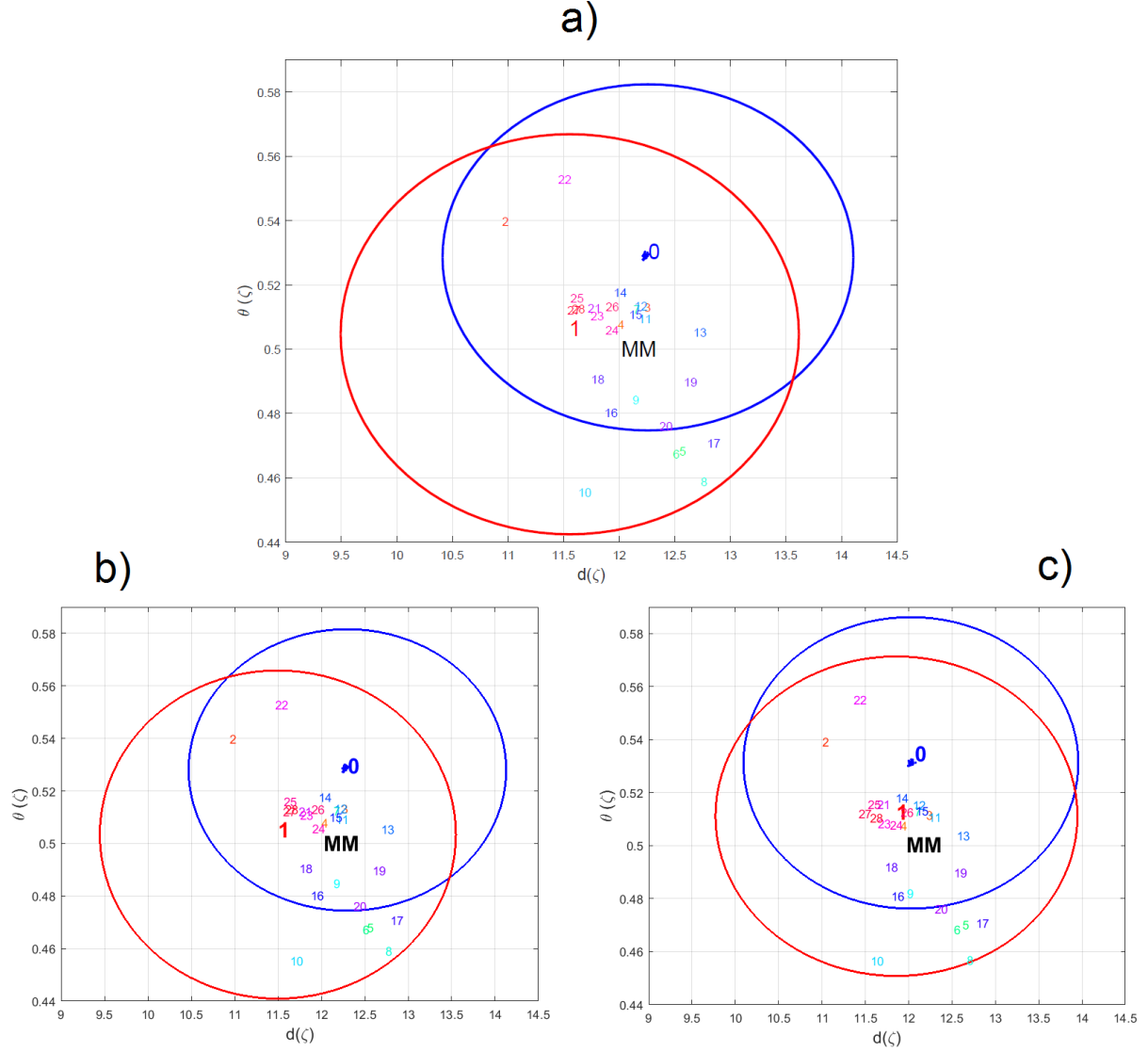


FIG. 7. Comparison between the 56 20CR-ME median values of (d, θ) (blue points whose average is denoted by 0), the 20CR-EM (in red and numbered by 1) and all the CMIP5 models (progressive numbers 2-27, see table 1 for details). The CMIP5 multi-model mean is marked by MM (black). The semiaxes of the two ellipses represent one standard deviation of d and θ for 20CR-EM (red) and 20CR-ME (blue). (a) 1851-2000 (b) 1851-1978 and (c) 1979-2000

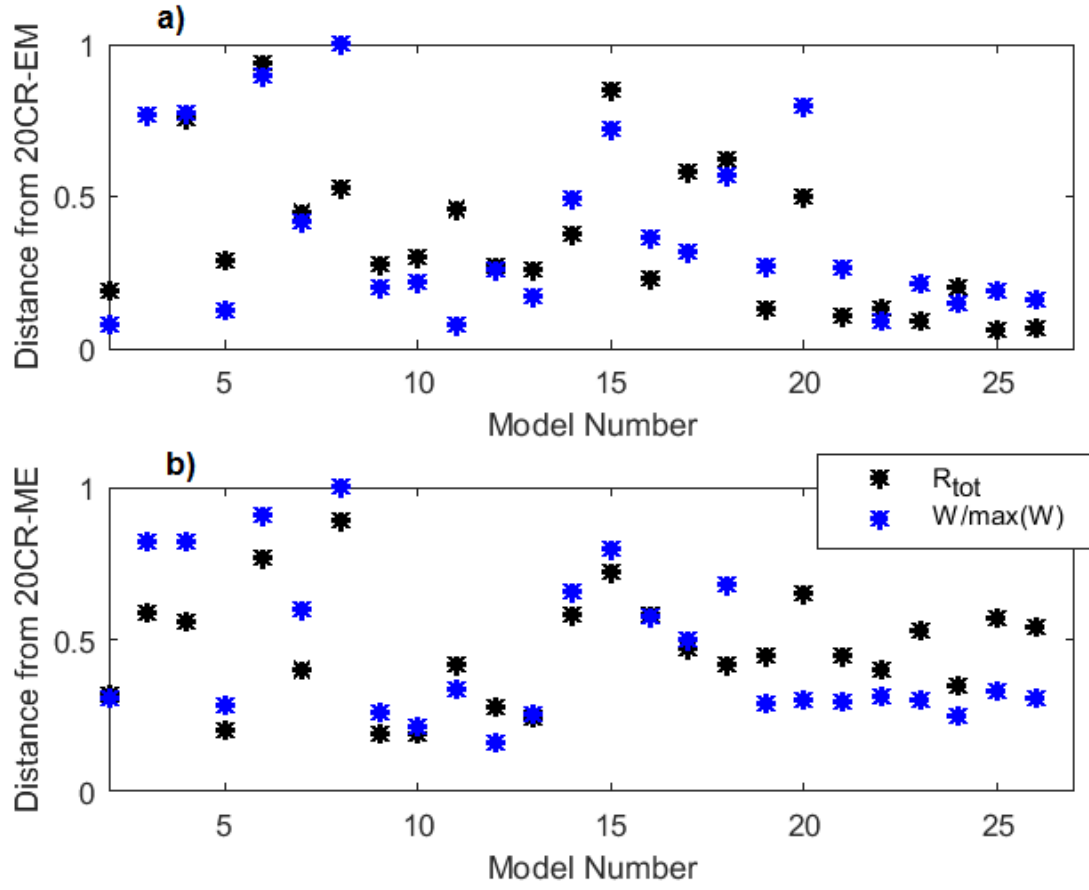


FIG. 8. Comparison between R_{tot} values (black) and Wasserstein distances $\mathcal{W}$ (blue) between the (d, θ) of 20CR-EM (a) and 20CR-ME (b) and of the CMIP5 models. The distances are normalized by the maximum value.

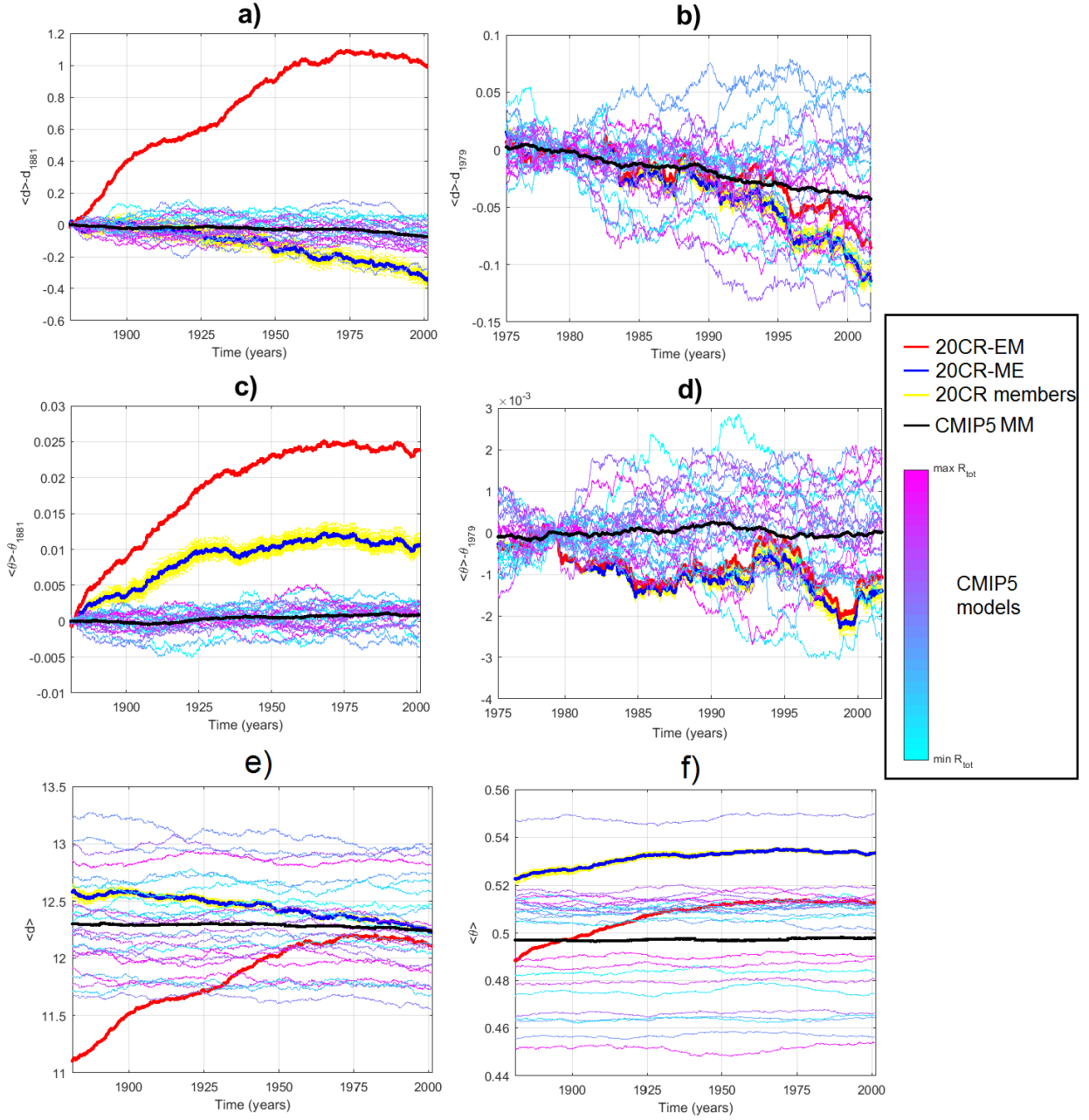


FIG. 9. 30-year moving averages $\langle \cdot \rangle$ of instantaneous dimension d and inverse persistence θ minus the respective mean values over 1851-1880 (a,c) and 1950-1979 (b,d). Note that labels on the abscissa mark the last year in each 30-year averaging window. (a,b): local dimension d ; (c,d): inverse persistence θ . Red: 20CR-EM; Blue: 20CR-ME; yellow: single 20CRv2c members; Black: CMIP5 multimodel mean; Color scale: single CMIP5 models ordered by R_{tot} . Panels (e) and (f) shows the absolute values of the 30-year moving averages of instantaneous dimension d (e) and inverse persistence θ (f).

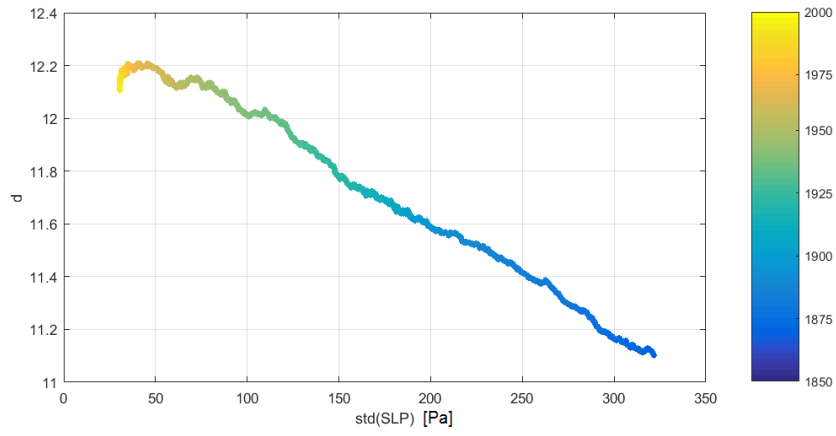


FIG. 10. 30-year moving average of the mean standard deviation of SLP fields across the 20CRv2c ensemble
 $\langle std(SLP) \rangle$ versus $\langle d \rangle$ for 20CR-EM. The color scale shows time. Values in Pa.

Bibliographie

- ARROYO, A. et E. R. PUJALS (2007). “Dynamical properties of singular-hyperbolic attractors”. In : *Discrete Contin. Dyn. Syst*, p. 67–87.
- BAZARAA, M. S. et al. (2009). *Linear Programming and Network Flows*. 4^e éd. John Wiley & Sons.
- BENAMOU, J.-D. et Y. BRENIER (2000). “A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem”. In : *Numerische Mathematik* 84.3, p. 375–393.
- BINDOFF, N. et al. (2013). “Detection and Attribution of Climate Change : from Global to Regional.” In : *Climate Change 2013 : The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Sous la dir. de T. STOCKER et al. Cambridge University Press, Cambridge, United Kingdom et New York, NY, USA, 867–952.
- BIRKHOFF, G. D. (1927). *Dynamical Systems*. Amer. Math. Soc., New York.
- BONNEEL, N. et al. (2011). “Displacement Interpolation Using Lagrangian Mass Transport”. In : *ACM Trans. Graph.* 30.6, 158 :1–158 :12.
- BOWEN, R. et J.-R. CHAZOTTES (1975). *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. T. 470. Springer.
- BUNIMOVICH, L. (1983). “Statistical properties of Lorenz attractors”. In : *Nonlinear dynamics and turbulence*, p. 71–92.
- CHEKROUN, M. D. et al. (2011). “Stochastic climate dynamics : Random attractors and time-dependent invariant measures”. In : *Physica D* 240.21, 1685–1700.
- CHEN, J. et al. (2013). “Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America”. In : *Water Resources Research* 49.7, p. 4187–4205.
- CLARKE, L. et al. (2007). “Scenarios of greenhouse gas emissions and atmospheric concentrations”. In : *US Department of Energy Publications*, p. 6.
- CULLEN, M. et H. MAROOFI (2003). “The Fully Compressible Semi-Geostrophic System from Meteorology”. In : *Archive for Rational Mechanics and Analysis* 167.4, p. 309–336.

- CUTURI, M. (2013). “Sinkhorn Distances : Lightspeed Computation of Optimal Transport”. In : *Advances in Neural Information Processing Systems 26*. Sous la dir. de C. J. C. BURGESS et al. Curran Associates, Inc., p. 2292–2300.
- CUTURI, M. et A. DOUCET (2014). “Fast computation of Wasserstein barycenters”. In : *International Conference on Machine Learning*, p. 685–693.
- DRÓTOS, G. et al. (2015). “Probabilistic concepts in a changing climate : a snapshot attractor picture”. In : *J. Climate* 28.8, p. 3275–3288.
- DUFRESNE, J.-L. et al. (2013). “Climate change projections using the IPSL-CM5 Earth System Model : from CMIP3 to CMIP5”. In : *Climate Dynamics* 40.9, p. 2123–2165.
- EVANS, L. (2010). *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society.
- FARANDA, D. et al. (2017a). “Dynamical properties and extremes of Northern Hemisphere climate fields over the past 60 years”. In : *Nonlinear Processes in Geophysics* 24.4, p. 713–725.
- FARANDA, D. et al. (2011). “Numerical convergence of the Block-Maxima approach to the generalized extreme value distribution”. In : *Journal of Statistical Physics* 145.5, p. 1156–1180.
- FARANDA, D. et al. (2013). “Extreme value statistics for dynamical systems with noise”. In : *Nonlinearity* 26.9, p. 2597–2622.
- FARANDA, D. et al. (2017b). “Dynamical proxies of North Atlantic predictability and extremes”. In : *Scientific Reports* 7.
- FARCHI, A. et al. (2016). “Using the Wasserstein distance to compare fields of pollutants : application to the radionuclide atmospheric dispersion of the Fukushima-Daiichi accident”. In : *Tellus B : Chemical and Physical Meteorology* 68.1, p. 31682.
- FLAMARY, R. et N. COURTY (2017). “POT Python Optimal Transport library”. In :
- FRAEDRICH, K. et al. (2005). “The Planet Simulator : Towards a user friendly model”. In : *Meteorologische Zeitschrift* 14.3, p. 299–304.
- FREITAS, A. C. M. et al. (2010). “Hitting time statistics and extreme value theory”. In : *Probability Theory and Related Fields* 147.3, p. 675–710.
- FUJINO, J. et al. (2006). “Multi-gas Mitigation Analysis on Stabilization Scenarios Using Aim Global Model”. In : *The Energy Journal* 27, p. 343–353.
- GHIL, M. (2015). “A Mathematical Theory of Climate Sensitivity or, How to Deal With Both Anthropogenic Forcing and Natural Variability?” In : *Climate Change : Multi-decadal and Beyond*. WORLD SCIENTIFIC. Chap. Chapter 2, p. 31–51.
- GHIL, M. et al. (2008). “Climate dynamics and fluid mechanics : Natural variability and related uncertainties”. In : *Physica D : Nonlinear Phenomena* 237.14, p. 2111–2126.
- GRASSBERGER, P. et I. PROCACCIA (1983). “Measuring the strangeness of strange attractors”. In : *Physica D : Nonlinear Phenomena* 9.1, p. 189–208.

- GREENSIDE, H. S. et al. (1982). “Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors”. In : *Phys. Rev. A* 25 (6), p. 3453–3456.
- HADLEY, G. (1735). “Concerning the cause of the general trade-winds”. In : *Phil. Trans.* 39.436–444, p. 58–62.
- HIJIOKA, Y. et al. (2008). “Global GHG emission scenarios under GHG concentration stabilization targets”. In : *Journal of Global Environment Engineering* 13, p. 97–108.
- HIRSCH, M. W. et al. (2006). *Invariant manifolds*. T. 583. Springer.
- IPCC (2014). “Annex II : Glossary”. In : *Climate Change 2014 : Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Sous la dir. de K. MACH et al., p. 117–130.
- KANTOROVITCH, L. (1958). “On the Translocation of Masses”. In : *Management Science* 5.1, p. 1–4.
- LIONS, J. L. et al. (1992). “New formulations of the primitive equations of atmosphere and applications”. In : *Nonlinearity* 5.2, p. 237–288.
- LIONS, J. et al. (1993). *Models for the Coupled Atmosphere and Ocean*. Computational mechanics advances : an official publication of the International Association for Computational Mechanics. North-Holland.
- LORENZ, E. N. (1969). “Atmospheric Predictability as Revealed by Naturally Occurring Analogues”. In : *Journal of the Atmospheric Sciences* 26.4, p. 636–646.
- (1990). “Can chaos and intransitivity lead to interannual variability?” In : *Tellus A* 42.3, p. 378–389.
- LORENZ, E. N. (1963). “Deterministic nonperiodic flow”. In : *J. Atmos. Sci.* 20.2, p. 130–141.
- LORENZ, E. N. (1984). “Irregularity : a fundamental property of the atmosphere”. In : *Tellus A* 36.2, p. 98–110.
- LUCARINI, V. et al. (2017). “Predicting Climate Change Using Response Theory : Global Averages and Spatial Patterns”. In : *Journal of Statistical Physics* 166.3, p. 1036–1064.
- LYAPUNOV, A. (1992). *General Problem of the Stability Of Motion*. CRC Press.
- MASOLLER, C. et al. (1992). “Regular and chaotic behavior in the new Lorenz system”. In : *Physics Letters A* 167.2, p. 185–190.
- MASOLLER, C. et al. (1995). “Characterization of strange attractors of lorenz model of general circulation of the atmosphere”. In : *Chaos, Solitons & Fractals* 6.Supplement C. Complex Systems in Computational Physics, p. 357–366.
- MCCANN, R. J. (1997). “A Convexity Principle for Interacting Gases”. In : *Advances in Mathematics* 128.1, p. 153–179.
- MONGE, G. (1781). “Mémoire sur la théorie des déblais et des remblais”. In : *Histoire de l’Académie Royale des Sciences*, p. 666–704.

- MOSS, R. H. et al. (2010). “The next generation of scenarios for climate change research and assessment”. In : *Nature* 463, p. 747–756.
- MUSKULUS, M. et S. VERDUYN-LUNEL (2011). “Wasserstein distances in the analysis of time series and dynamical systems”. In : *Physica D : Nonlinear Phenomena* 240.1, p. 45–58.
- NICOLIS, C et al. (1995). “Short-range predictability of the atmosphere : Mechanisms for superexponential error growth”. In : *Quarterly Journal of the Royal Meteorological Society* 121.523, p. 705–722.
- OSELEDEC, V. I. (1968). “A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems”. In : *Trans. Moscow Math. Soc* 19.2, p. 197–231.
- PALIS, J. (2005). “A global perspective for non-conservative dynamics”. In : *Annales de l’Institut Henri Poincaré (C) Non Linear Analysis* 22.4, p. 485–507.
- (2008). “Open questions leading to a global perspective in dynamics”. In : *Nonlinearity* 21.4, T37.
- PIANI, C. et al. (2010). “Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models”. In : *Journal of Hydrology* 395.3–4, p. 199–215.
- PIELKE, R. A. et X. ZENG (1994). “Long-term variability of climate”. In : *Journal of the atmospheric sciences* 51.1, p. 155–159.
- POINCARÉ, H. (1892). *Les Méthodes Nouvelles de la Mécanique Céleste*. T. 1–3. Gauthier-Villars, Paris.
- RAYLEIGH, L. (1916). “LIX. On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side”. In : *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 32.192, p. 529–546.
- RIAH, K. et al. (2007). “Scenarios of long-term socio-economic and environmental development under climate stabilization”. In : *Technological Forecasting and Social Change* 74.7. Greenhouse Gases - Integrated Assessment, p. 887–935.
- RICHARDSON, L. F. (1922). “Weather prediction by numerical process”. In :
- ROBIN, Y. et al. (2017). “Detecting changes in forced climate attractors with Wasserstein distance”. In : *Nonlinear Processes in Geophysics* 24.3, p. 393–405.
- ROBINSON, C. R. (1999). *Dynamical systems*. Studies in Advanced Mathematics. CRC press.
- RODRIGUES, D. et al. (2017). “Dynamical properties of the North Atlantic atmospheric circulation in the past 150 years in CMIP5 models and the 20CRv2c Reanalysis”. (accepted in J. Clim.)
- ROMEIRAS, F. J. et al. (1990). “Multifractal properties of snapshot attractors of random maps”. In : *Phys. Rev A* 41 (2), p. 784–799.

- RUBNER, Y. et al. (2000). “The Earth Mover’s Distance as a Metric for Image Retrieval”. In : *International Journal of Computer Vision* 40.2, p. 99–121.
- RUELLE, D. (1976). “A measure associated with Axiom-A attractors”. In : *Amer. J. Math.* 98.3, 619—654.
- RUELLE, D. (2009). “A review of linear response theory for general differentiable dynamical systems”. In : *Nonlinearity* 22, p. 855–870.
- RUELLE, D. et F. TAKENS (1971). “On the nature of turbulence”. In : *Commun. Math. Phys.* 20, p. 167–192.
- RÖSSLER, O. (1976). “An equation for continuous chaos”. In : *Physics Letters A* 57.5, p. 397–398.
- SALTZMAN, B. (1962). “Finite amplitude free convection as an initial value problem—I”. In : *Journal of the Atmospheric Sciences* 19.4, p. 329–341.
- SANTAMBROGIO, F. (2015). *Optimal Transport for Applied Mathematicians*. T. 87. Birkhäuser Basel, p. XXVII, 353.
- SCHMITZER, B. (2016). “Stabilized Sparse Scaling Algorithms for Entropy Regularized Transport Problems”. In : *ArXiv e-prints*.
- SINAI, Y. G. (1972). “Gibbs Measures In Ergodic Theory”. In : *Russian Mathematical Surveys* 27.4, p. 21–69.
- SKAMAROCK, W. et al. (2008). *A Description of the Advanced Research WRF Version 3*. Rapp. tech.
- SMALE, S. (1967). “Differential Dynamical Systems”. In : *Bull. Amer. Math. Soc* 73, p. 747–817.
- SMITH, L. A. (1988). “Intrinsic limits on dimension calculations”. In : *Physics Letters A* 133.6, p. 283–288.
- SMITH, S. J. et T. M. L. WIGLEY (2006). “Multi-Gas Forcing Stabilization with Minicam”. In : *The Energy Journal* 27, p. 373–391.
- STURGES, H. A. (1926). “The choice of a class interval”. In : *Journal of the american statistical association* 21.153, p. 65–66.
- TAYLOR, K. E. et al. (2012). “An overview of CMIP5 and the experiment design”. In : *Bulletin of the American Meteorological Society* 93.4, p. 485–498.
- TEMAM, R. (2012). *Infinite-dimensional dynamical systems in mechanics and physics*. T. 68. Springer Science & Business Media.
- TUCKER, W. (1999). “The Lorenz attractor exists”. In : *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics* 328.12, p. 1197–1202.
- VANNITSEM, S. (2014). “Dynamics and predictability of a low-order wind-driven ocean–atmosphere coupled model”. In : *Climate Dynamics* 42.7, p. 1981–1998.
- VIANA, M. (2000). “What’s new on lorenz strange attractors?” In : *The Mathematical Intelligencer* 22.3, p. 6–19.

- VIDAL, J.-P. et al. (2010). “A 50-year high-resolution atmospheric reanalysis over France with the Safran system”. In : *International Journal of Climatology* 30.11, p. 1627–1644.
- VILLANI, C. (2008). *Optimal Transport : Old and New*. 1^{re} éd. T. 338. Grundlehren der mathematischen Wissenschaften. Springer Science & Business Media.
- VILLANI, C. (2003). *Topics in Optimal Transportation*. 58. American Mathematical Society.
- VUUREN Detlef P. and den Elzen, M. G. J. van et al. (2007). “Stabilizing greenhouse gas concentrations at low levels : an assessment of reduction strategies and costs”. In : *Climatic Change* 81.2, p. 119–159.
- WASSERSTEIN, L. (1969). “Markov processes on a countable product space, describing large systems of automata”. In : *Problemy Peredachi Infomatsii* 5 3, p. 64–73.
- WILCKE, R. A. I. et al. (2013). “Multi-variable error correction of regional climate models”. In : *Climatic Change* 120.4, p. 871–887.
- WILLIAMS, R. F. (1979). “The structure of Lorenz attractors”. In : *Publ. Math. IHES* 50, p. 73–99.
- WISE, M. et al. (2009). “Implications of Limiting CO₂ Concentrations for Land Use and Energy”. In : *Science* 324.5931, p. 1183–1186.
- YOUNG, L.-S. (1981). “Capacity of attractors”. In : *Ergodic Theory and Dynamical Systems* 1.3, p. 381–388.

Table des figures

1	a) Représentation schématique du forçage solaire. b) Représentation schématique de la dynamique atmosphérique.	6
2	Histogramme des températures quotidiennes en Ile-de-France issues d'une simulation de contrôle du modèle de l'IPSL (DUFRESNE et al., 2013). Le forçage radiatif est fixé à celui de l'ère pré-industrielle, i.e. avant 1850. La simulation fait 250 ans, correspondant à la période 1850-2100. Seules les valeurs de la période 2000-2100 ont été utilisées pour générer cet histogramme. Les deux modes (les bosses) correspondent à l'hiver et l'été. . . .	8
1.1	Exemple d'une solution du système de LORENZ (1963).	14
1.2	Deux trajectoires de l'attracteur de LORENZ (1963) prises aux conditions initiales suivantes : $(5.16, -0.87, 30.64)$ (courbe rouge) et $(4.99, -0.53, 30.02)$ (courbe bleue), et intégrées pendant 0.6 unité de temps. Les trajectoires ont été projetées dans le plan (y, z)	16
1.3	Représentation d'un point fixe (en noir) d'un système dynamique. La courbe bleue représente sa direction stable, qui attire toutes les trajectoires. La courbe rouge représente sa direction instable, qui repousse toutes les trajectoires. Les courbes vertes sont les trajectoires des points $\mathbf{x}_0$ et $\mathbf{x}_1$	17
1.4	Attracteur de LORENZ (1963). La projection est choisie de manière à mettre en valeur que l'attracteur a une dimension proche de 2.	19
1.5	Histogramme de la coordonnée x d'une trajectoire du modèle de LORENZ (1963). a) La condition initiale de la trajectoire est $(1, 1, 1)$. b) La condition initiale de la trajectoire est $(1, -1, -1)$	20
1.6	Deux orbites du LORENZ (1963). Sur la Figure a), la condition initiale est $(80, 50, 50)$, tandis que c'est $(-50, -50, -50)$ sur la Figure b). Les deux orbites génèrent bien le même objet.	22
1.7	Recouvrement du carré $[0, 1] \times [0, 1]$ par des petits carrés de taille $\varepsilon = 0.25$. Il faut $1/\varepsilon^2 = 16$ carrés pour le recouvrir.	23
1.8	Dimension locale de l'attracteur de LORENZ (1963)	24

1.9	Estimation de la mesure SRB du système de LORENZ (1963) à partir d'une orbite contenant 590000 points. 600000 points sont générés par pas de temps de 0.005 à l'aide de l'Équation (1.1). Les 10000 premiers sont ignorés pour ne conserver que des points sur l'attracteur. Chaque points de couleur est le centre d'un petit cube de taille $0.5 \times 0.5 \times 0.5$. Il y a 14575 cubes sur cette figure.	27
1.10	Distribution des températures en Ile-de-France (France) pour les simulations CMIP5, modèle de l'IPSL. À gauche, une simulation de contrôle (pré-industrielle). À droite, une simulation du scénario RCP85.	30
1.11	Attracteurs de LORENZ (1984) pour deux valeurs de F . a) $F = 6$ (été permanent, système non chaotique bloqué sur un cycle limite). b) $F = 8$ (hiver permanent, système chaotique)	31
1.12	a) Cycle saisonnier du LORENZ (1984). b) Rouge : Cycle saisonnier et forçage linéaire après 100 ans.	32
1.13	Snapshot de l'attracteur de LORENZ (1963) construit avec 1000 conditions initiales aléatoires. La figure a) est l'état de toutes les conditions initiales au temps 0. La figure b) est l'état de toutes les conditions initiales au temps 0.05 .La figure c) est l'état de toutes les conditions initiales au temps 0.1. La figure d) est l'état de toutes les conditions initiales au temps 0.25. . . .	35
1.14	Deux snapshot de l'attracteur de LORENZ (1963) construis avec deux jeux de 1000 conditions initiales en bleu et rouge. La figure a) est l'état de toutes les conditions initiales au temps 0. La figure b) est l'état de toutes les conditions initiales au temps 0.25.	36
1.15	Snapshots de l'attracteur de LORENZ (1984) généré à partir de 1000 conditions initiales tirées uniformément dans le cube $[-1, 3] \times [-3, 3] \times [-3, 3]$. En colonne : état des snapshots pendant le premier cycle saisonnier (année 0), puis le dixième cycle (année 10) et le cycle 180. Ce dernier est donc après que le forcing ait commencé à agir à l'année 0. En colonne : snapshots pendant l'automne, puis l'hiver, puis le printemps et enfin l'été.	38
2.1	Les mesures μ_t sont des lois normales centrées en $2t$, d'écart type égal à 1. a) Densités de μ_t , notées ρ_t , aux temps 0, 2, 4, 6 et 8. b) Distance Euclidienne entre μ_0 et μ_t	41
2.2	Deux exemples du problème de Monge. Les points bleus $\mathbf{x}_i$ sont les déblais, les rouges $\mathbf{y}_j$ les remblais. Les segments noirs sont des déplacements possibles entre $\mathbf{x}_i$ et $\mathbf{y}_j$	42
2.3	Deux plans de transport entre les mesures empiriques μ (points bleus en $\mathbf{x}_i$) et ν (points rouges en $\mathbf{y}_j$). Les segments noirs sont les trajectoires de déplacements. Les valeurs rouges sont les masses déplacées.	44

2.4	Les mesures μ_t sont des lois normales centrées en $2t$, d'écart type égal à 1. a) Densités de $\mu_t = \mu_0 + 2t$ aux temps 0, 2, 4, 6 et 8. b) Distance entre μ_0 et μ_t . La distance Euclidienne est donnée en rouge, la distance de Wasserstein en bleu.	47
2.5	Histogrammes d'un exemple de deux mesures μ et ν . La mesure μ est une Gaussienne centrée en 0, de variance 1. La mesure ν est bi-modale, constituée de deux Gaussiennes centrées respectivement en -3 et 3 , toutes deux de variances 1. Les bins sont uniformes, de longueur égale à 0.3. Enfin, 10000 points ont été utilisés pour générer ces histogrammes.	49
2.6	Représentation du plan de transport entre μ et ν . La mesure μ est une Gaussienne centrée réduite. Son histogramme (Figure de droite) a été estimé à partir de 10000 points, avec une taille de bin constante égale à 0.3. La mesure ν est une bimodale composée de deux Gaussiennes centrées respectivement en -3 et 3 et de variance 1. Son histogramme (Figure du haut) a été estimé à partir de 10000 points, avec une taille de bin constante égale à 0.3. En dessous de chaque histogramme est représenté la quantité de masse dans chaque bin à l'aide d'une colormap. La figure centrale est le plan de transport optimal qui associe μ et ν . Les masses de couleur sont la quantité de masses déplacées de chaque bin de μ vers les bins de ν	51
2.7	Visualisation de la séquence μ_t , obtenue par interpolation de μ et ν (c.f. Eq. (2.7)) Les panneaux a-f correspondent respectivement à une interpolation aux temps $t = 0, 0.2, 0.4, 0.6, 0.8$ et 1 . En particulier, pour $t = 0$ et $t = 1$, les mesures obtenues sont μ et ν	52
2.8	Moyenne des 200 distances de Wasserstein entre une loi normale centrée réduite, bruitée par une loi normale centrée d'écart type σ_0 , et une mixture de deux lois normales centrées en -3 et 3 , bruitées par une loi normale d'écart type σ_1 . Sur les panneaux a-d , la longueur des bins est respectivement de 1, 0.5, 0.1 et 0.005. e) Histogramme de μ_{σ_0} (rouge) et ν_{σ_1} (bleu) pour $\sigma_0 = 3.5$ et $\sigma_1 = 0$. f) Histogramme de μ_{σ_0} (rouge) et ν_{σ_1} (bleu) pour $\sigma_0 = 7.5$ et $\sigma_1 = 7.5$	54
2.9	Écart type des 200 distances de Wasserstein entre une loi normale centrée réduite, bruitée par une loi normale centrée d'écart type σ_0 , et une mixture de deux lois normales centrées en -3 et 3 , bruitées par une loi normale d'écart type σ_1 . Sur les panneaux a-d , la longueur des bins est respectivement de 1, 0.5, 0.1 et 0.005.	55

2.10	a) Distance de Wasserstein entre μ_i et μ_0 . Les mesures μ_i sont la distribution de probabilités des températures en Ile-de-France pour le modèle de l'IPSL sur les périodes 2000-2020, 2020-2040, 2040-2060, 2060-2080 et 2080-2100. b) Idem mais entre les histogrammes centrés. c) Écarts types pour les mêmes périodes et variables. d) Idem que a) et b) mais pour des distributions centrées et réduites.	56
3.1	Scénario RCP des forçages radiatifs entre 1850 et 2100.	75
3.2	Moyenne sur l'hémisphère nord des distances de Wasserstein entre une fenêtre glissante de 5 ans et une référence sur la période 1850-1855. La variable analysée est la température. La date sur l'axe des x correspond au début de la fenêtre. Le modèle utilisé est celui de l'IPSL (DUFRESNE et al., 2013). Gris : limite entre la partie historique et les scénarios RCP. Noir : simulation préindustrielle. Vert : scénario RCP26. Bleu : scénario RCP45. Violet : scénario RCP60. Rouge : scénario RCP85. a) Les fenêtres et la référence contiennent l'année complète. b) La référence est restreinte aux étés. c) La référence est restreinte aux hivers.	76
3.3	Moyenne sur l'hémisphère nord des distances de Wasserstein entre une fenêtre glissante de 5 ans et une référence sur la période 1850-1855. La variable analysée est la précipitation. La date sur l'axe des x correspond au début de la fenêtre. Le modèle utilisé est celui de l'IPSL (DUFRESNE et al., 2013). Gris : limite entre la partie historique et les scénarios RCP. Noir : simulation pré-industrielle. Vert : scénario RCP26. Bleu : scénario RCP45. Violet : scénario RCP60. Rouge : scénario RCP85. a) Les fenêtres et la référence contiennent l'année complète. b) La référence est restreinte aux étés. c) La référence est restreinte aux hivers.	77
3.4	Attracteur de (RÖSSLER, 1976).	78

Index

Équation de Transport, 33

Attracteur Étrange, 18

Dimension locale, 24

Direction instable, 16

Direction stable, 16

Distance de Wasserstein, 46

Espace des Phases, 21

Exposants de Lyapunov, 25

Flot (d'une équation différentielle), 20

Mesure Ergodique, 28

Mesure invariante, 27

Problème de Monge, 43

Problème de Monge-Kantorovich, 46

Sensibilité aux conditions initiales, 17

Système Autonome, 20

Système Climatique, 5

Système Dynamique Continu, 20

Système Dynamique Discret, 20

Vision Eulerienne, 33

Vision Lagrangienne, 33