

Regional Climate Model Emulation Using Diffusion models

Auteur : Temoschenko, Maxime

Promoteur(s) : Louppe, Gilles

Faculté : Faculté des Sciences appliquées

Diplôme : Master : ingénieur civil électricien, à finalité spécialisée en "electronic systems and devices"

Année académique : 2024-2025

URI/URL : <http://hdl.handle.net/2268.2/24763>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Abstract

Regional Climate Models (RCMs) are essential for producing high-resolution climate simulations tailored to specific regions. However, their reliance on discretized physical laws and boundary conditions from Global Climate Models (GCMs) limits flexibility for assimilating new observations and makes long-term simulations computationally expensive, reducing the ability to explore multiple possible climate states from the same GCM.

This thesis investigates the application of *Score-Based Data Assimilation* (SDA), relying on diffusion models, in the context of the *Modèle Atmosphérique Régional* (MAR) over Belgium. The problem is framed from a statistical perspective as an inverse problem.

The approach consists of (1) Training a diffusion model to approximate the prior distribution of MAR outputs, (2) Exploiting zero-shot conditional generation to incorporate synthetic and real-world observations.

A proof-of-concept is demonstrated training on two variables : the temperature and wind speed, presenting the capabilities of the model to generate samples aligned with the observations in various contexts including incorporating IRM weather stations and generate samples from the GCM ERA-5 land-surface variables given at the boundaries, replicating the simulation of MAR but for two variables.

Several limitations are acknowledged: the learned prior contains artefacts, posteriors conditioned on part of the observations fail to predict other unassimilated parts, the model has not yet been successfully scaled to all MAR variables or to a full simulation with all prognostic variables across pressure levels, the generated ensembles are not calibrated, and computational gains over traditional simulations are limited.