

# Les modèles de diffusion et leurs applications scientifiques

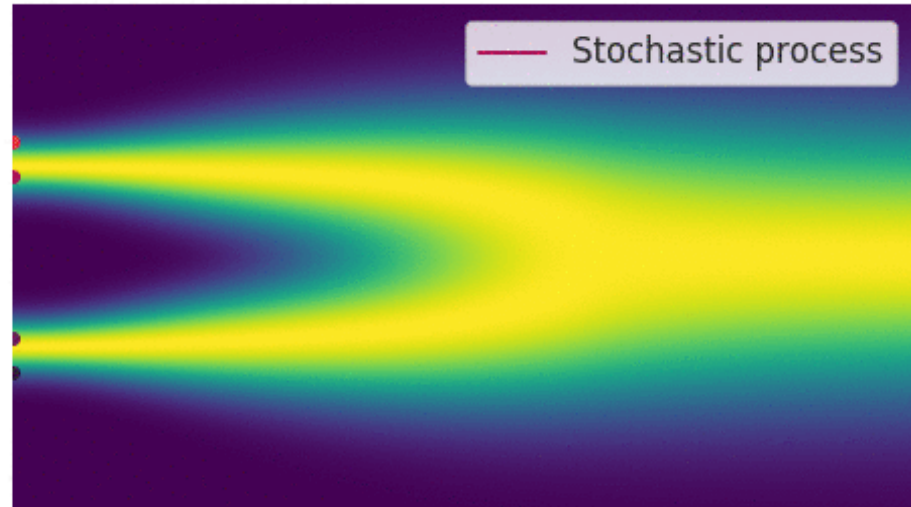
Café numérique  
3 mars 2025



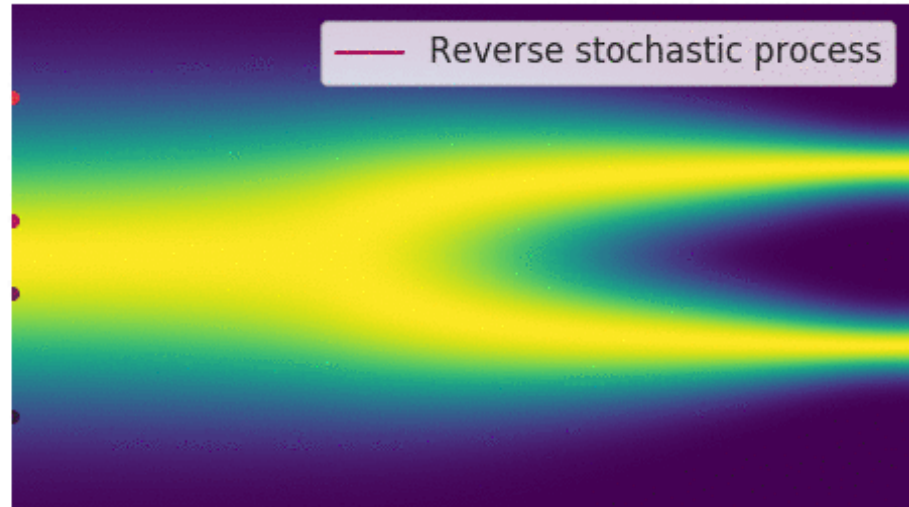
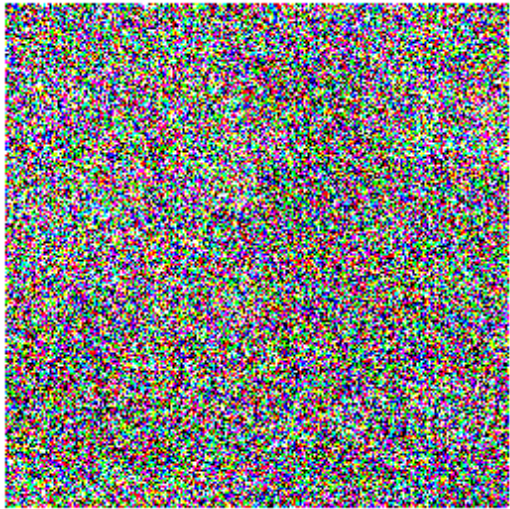
Gilles Louppe  
[g.louppe@uliege.be](mailto:g.louppe@uliege.be)



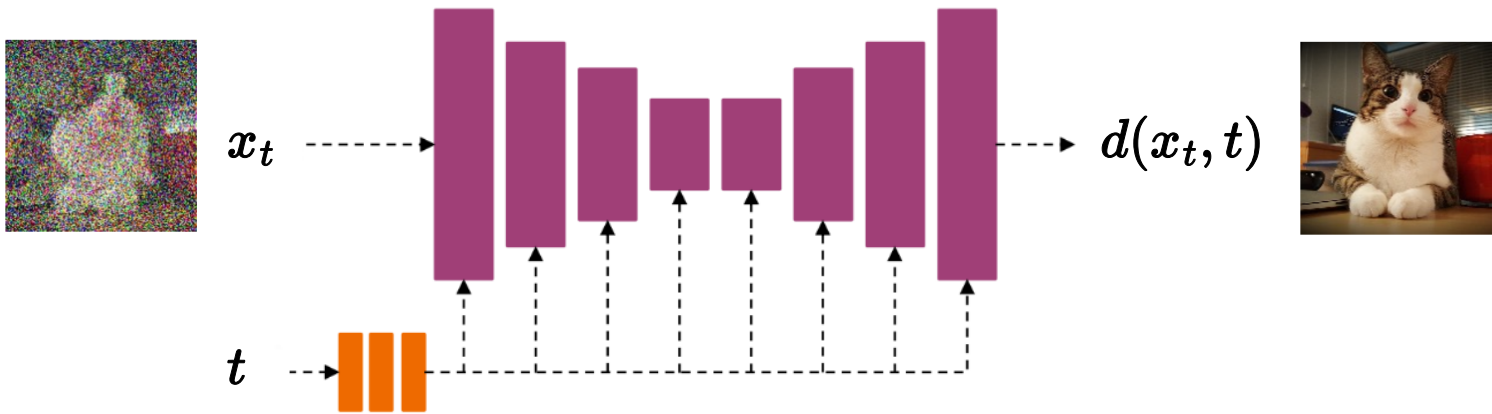
Ces images sont-elles réelles ou générées par une IA?



Un **processus de diffusion** est un processus stochastique  $x_t$  qui détériore progressivement un signal  $x$  en lui ajoutant du bruit.



En simulant le processus de **diffusion à l'envers**, on peut retirer progressivement le bruit et retrouver un signal  $x$ .



Mathématiquement, la diffusion inverse peut s'approximer par un **réseau de neurones**  $d_\theta(x_t, t)$  entraîné à retirer le bruit d'un signal perturbé  $x_t$ .



La génération par diffusion inverse peut être télécommandée par un signal auxiliaire  $y$  pour produire des données conditionnelles.



Prompt: Cinematic photography, close up, a banana laying on a table, slanted shadows wash across it, soft, diffused lighting.



Prompt: Cinematic photography, close up, grapes laying on a table, slanted shadows wash across it, soft, diffused lighting.



Prompt: Cinematic photography, close up, mango laying on a table, slanted shadows wash across it, soft, diffused lighting.

Text-to-image generation  
( $x$  = image,  $y$  = texte)

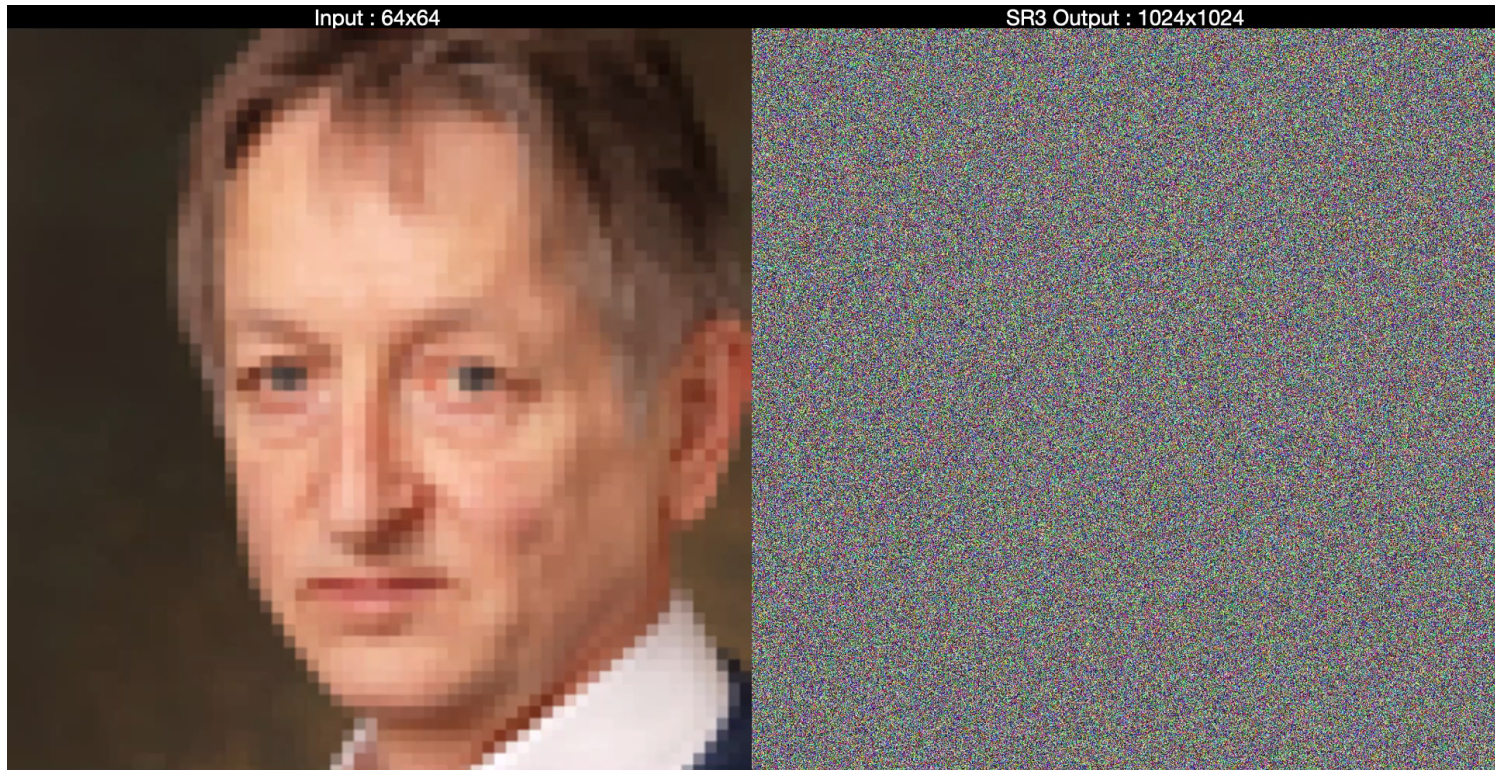
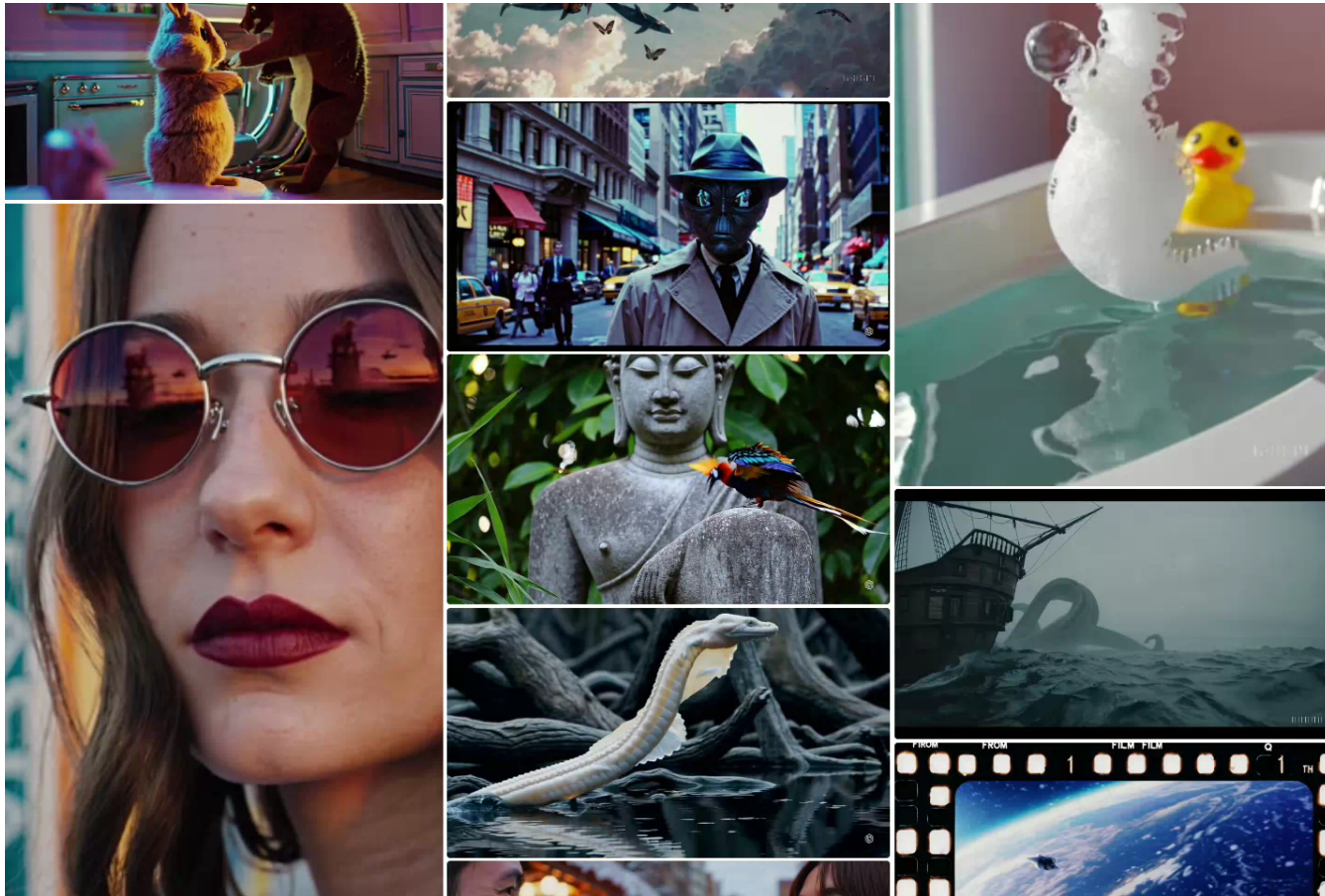


Image-to-image generation  
( $x$  = image haute résolution,  $y$  = image basse résolution)



Text-to-video generation  
( $x$  = vidéo,  $y$  = texte)

# L'IA générative, un réel progrès?

TECH-MÉDIAS

## L'IA générative nous rend paresseux et moins critiques : une étude le prouve



ACCUEIL • LÉNA

### Le « slop », ce raz-de-marée de contenus générés par l'intelligence artificielle qui menace Internet

Les IA génératives ont facilité la production automatique sans vérification d'images, de vidéos ou de sites web cherchant à capter le maximum d'audience sur YouTube ou Google. Un appât du gain qui trouble le rapport entre le vrai et le faux. Avec LÉNA, découvrez le meilleur du journalisme européen.

Illustration © ChatGPT Article réservé aux abonnés



Christophe Journaliste • 21- Mise à jour le: 21-02

LE FIGARO



L'intelligence artificielle nous rend plus en plus de paresseux et moins critiques. Elle altère d'ores et déjà nos capacités de réflexion.

Next Tech News

## Media company Thomson Reuters wins AI copyright case



Le Monde

DÉBATS • CINÉMA

### Axelle Ropert, réalisatrice : « On veut nous faire croire que l'IA est un progrès pour les cinéastes »

TRIBUNE

Axelle Ropert Cinéaste

A rebours des « gains de productivité » promis par les tenants de l'IA, la cinéaste appelle la profession, dans une tribune au « Monde », à réfléchir collectivement au bouleversement ontologique qui s'annonce.

Publié le 16 février 2025 à 10h30 | Lecture 6 min.

Offrir l'article

Article réservé aux abonnés

Alors que s'est tenu à Paris en grande pompe le Sommet mondial pour l'action sur l'intelligence artificielle (IA), sans aucune invitation officielle pour les grandes organisations de cinéastes, alors que la course mondiale à l'IA tente de faire croire que l'urgence doit servir de boussole

## Artists release silent album in protest against AI using their work

3 days ago

Paul Glynn Culture reporter

Share Save



Le Monde

ÉCONOMIE • INTELLIGENCES ARTIFICIELLES GÉNÉRATIVES

### « Montrer l'envers du décor » : au contre-sommet de l'IA, les conséquences déjà visibles sur de nombreux métiers

Cet événement a réuni différentes professions – traducteurs, scénaristes, doubleurs, journalistes... – déjà touchées par l'implantation de systèmes d'intelligence artificielle générative dans leurs secteurs d'activité. Le but : témoigner et débattre des conséquences de ces technologies sur le travail.

Par Jules Thomas

Publié le 11 février 2025 à 09h13, modifié le 11 février 2025 à 12h23

Lecture 4 min.

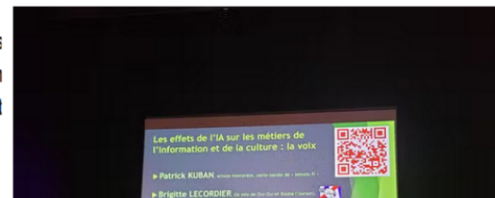
Offrir l'article

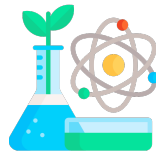
Article réservé aux abonnés

+1

ents

## Thomson Reuters, silent album content to train an AI against US copyright

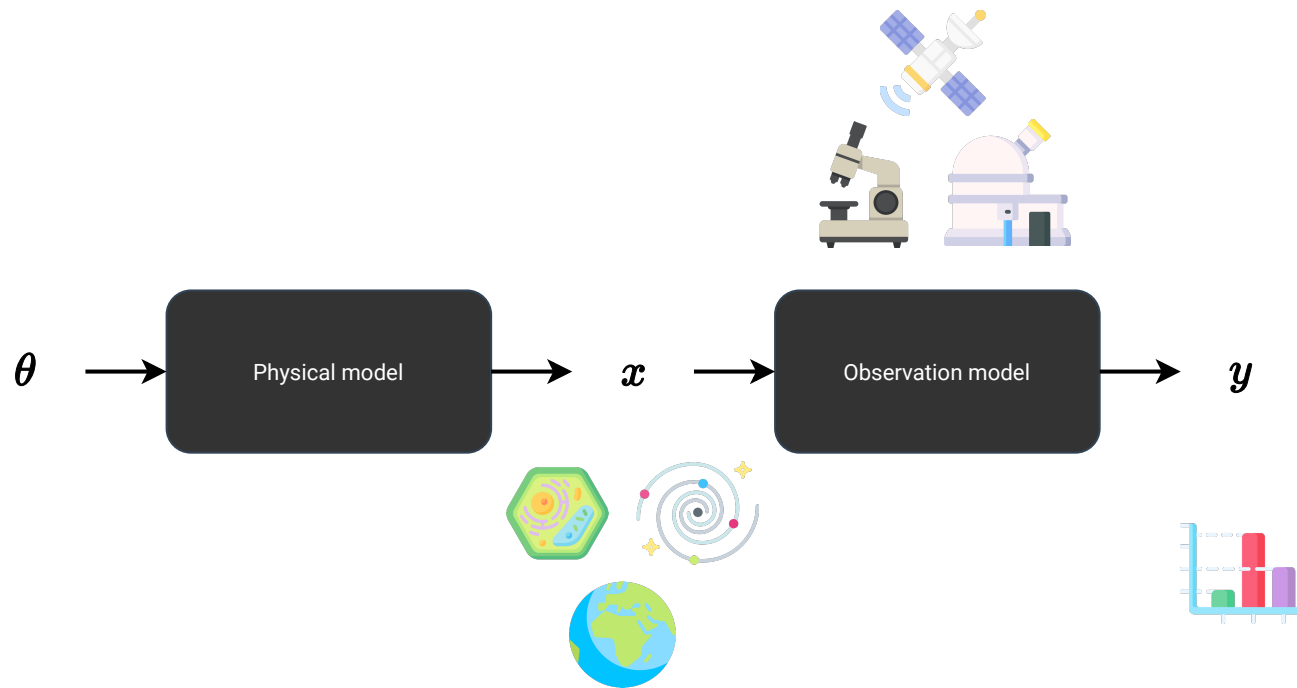




L'IA est capable de résoudre des problèmes scientifiques que personne ne peut résoudre.

**D'observations bruitées  $y$ ...**

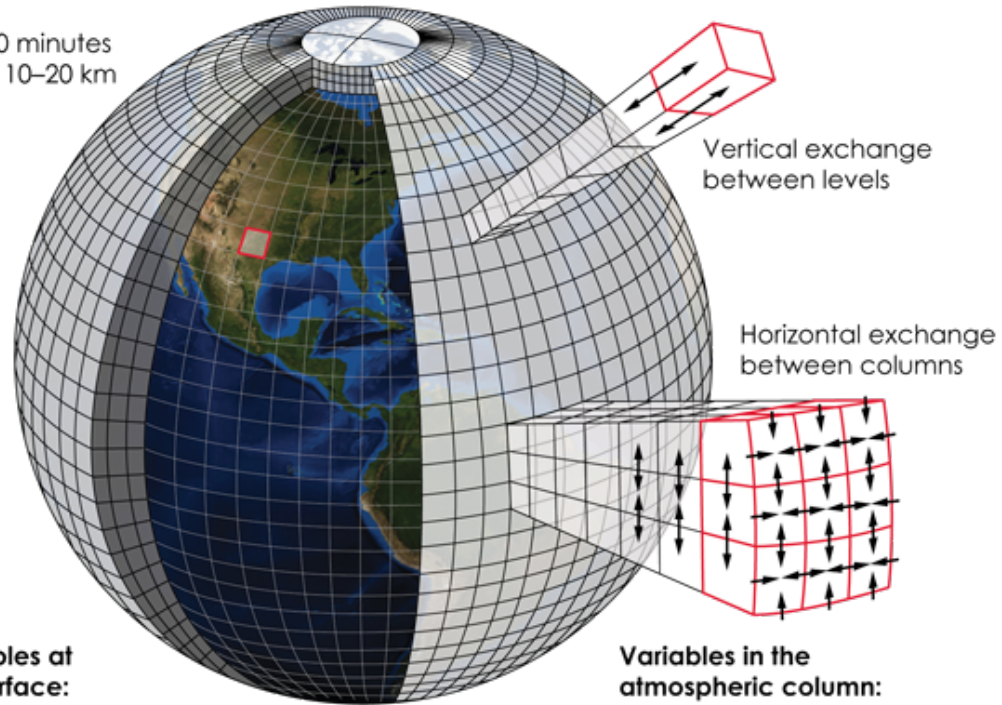
**... peut-on reconstruire  
un signal  $x$ ?**





# Weather forecast modeling

Timestep 5–10 minutes  
Grid spacing 10–20 km



## Variables at the surface:

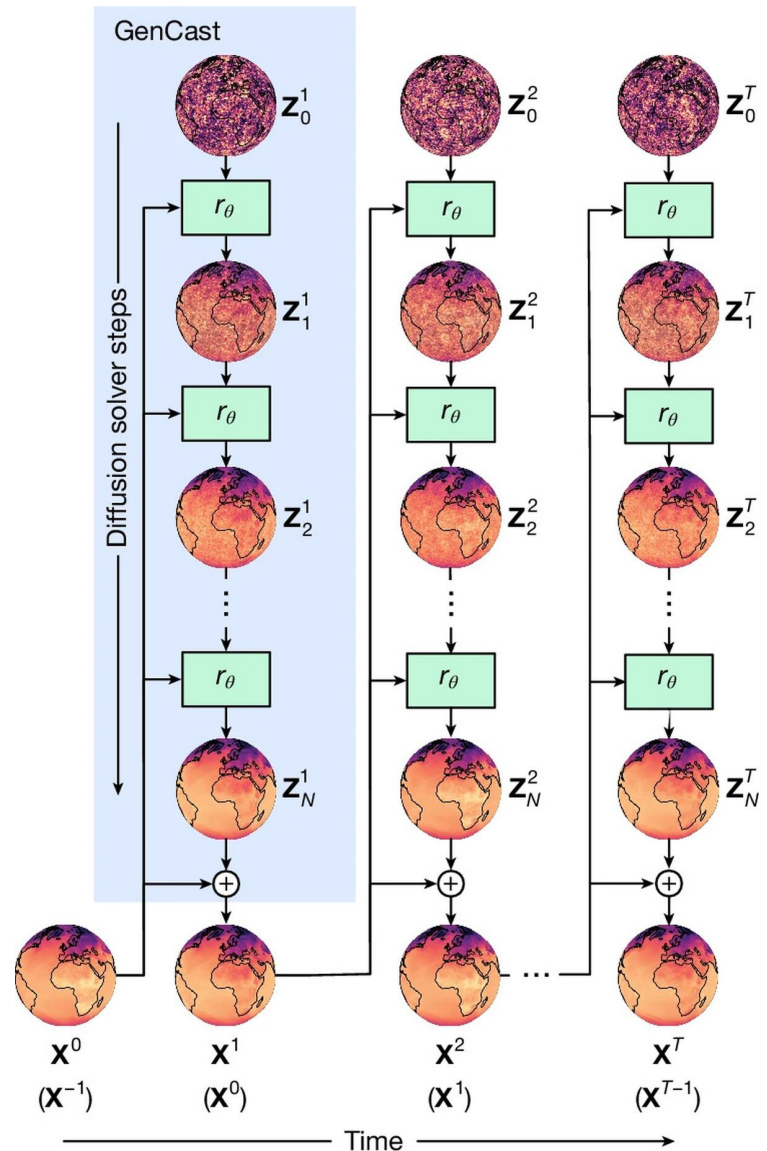
- Temperature
- Humidity
- Pressure
- Moisture fluxes
- Heat fluxes
- Radiation fluxes

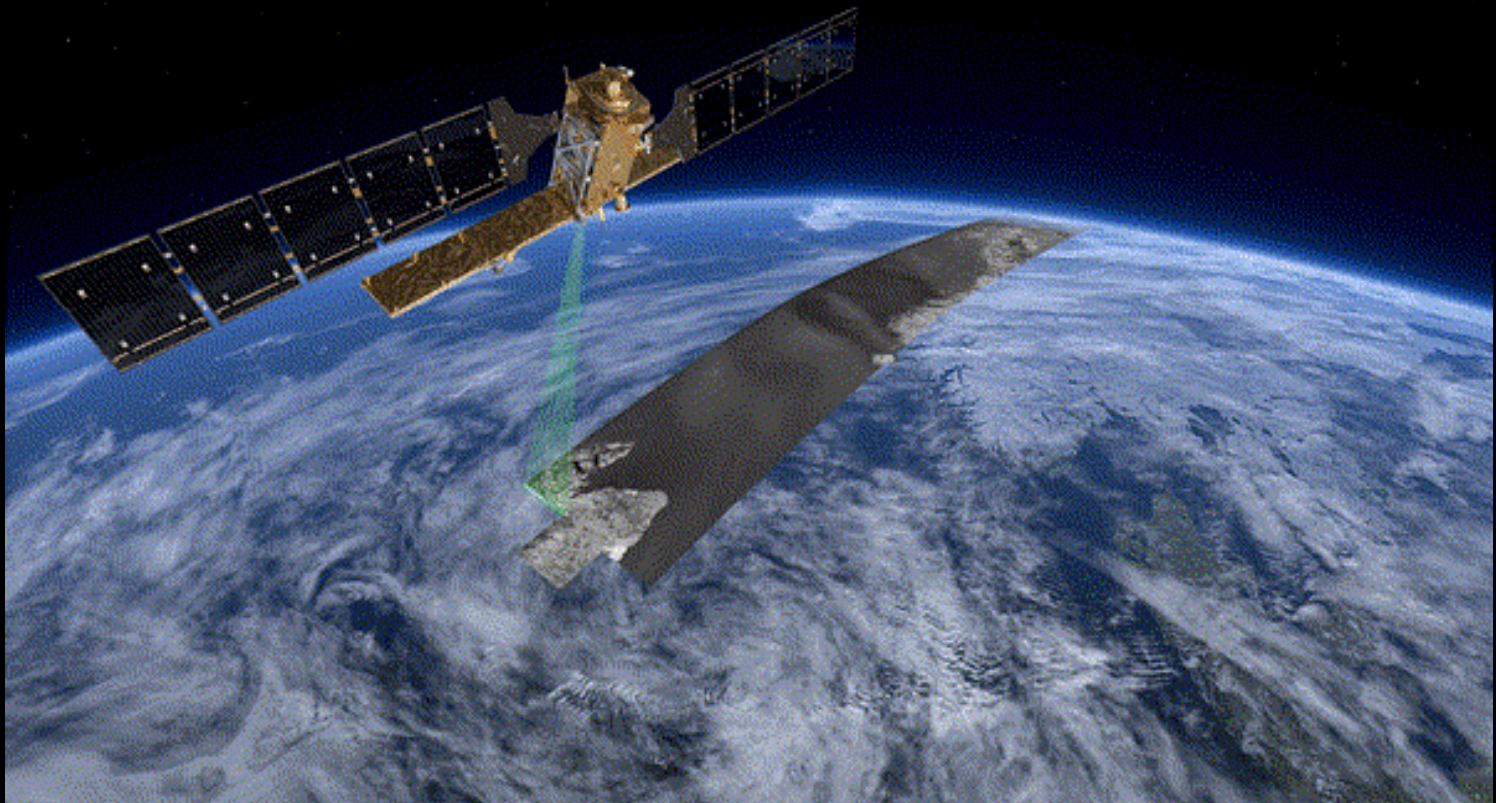
## Variables in the atmospheric column:

- Wind vectors
- Humidity
- Clouds
- Temperature
- Height
- Precipitation
- Aerosols

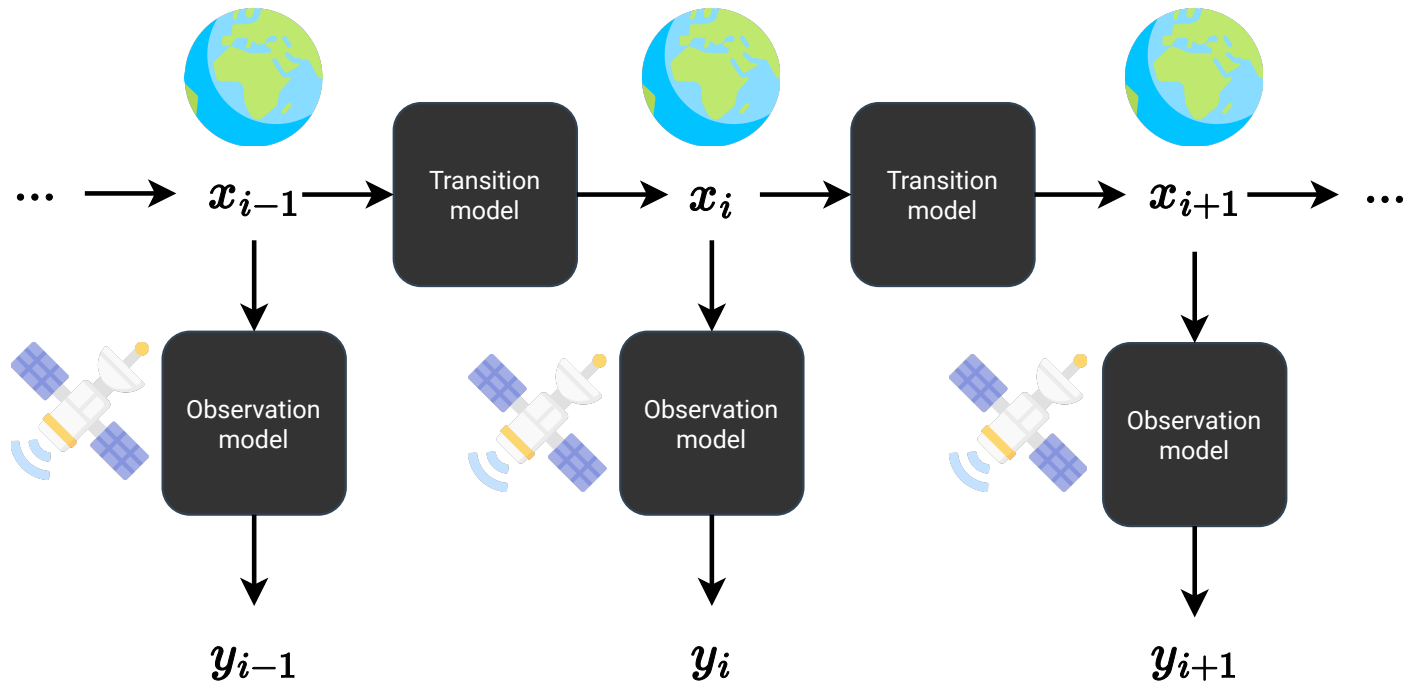
$x$  = état météorologique

**GenCast** (Google Deepmind) est un modèle de diffusion de type image-to-image pour la prévision météorologique du prochain état  $x_{t+1}$  à partir de l'état actuel  $x_t$ .

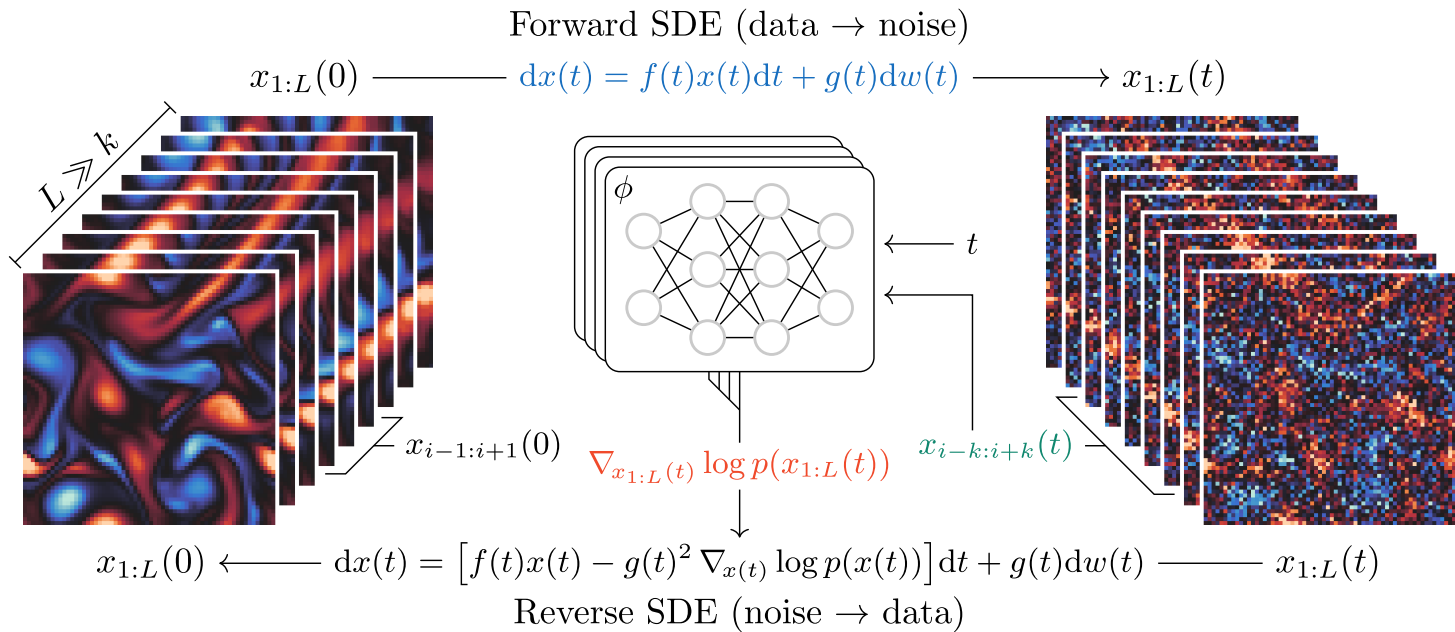




Comment les conditions initiales  $\mathbf{x}_0$  sont-elles connues?  
Seules des observations partielles  $\mathbf{y}$  sont disponibles!

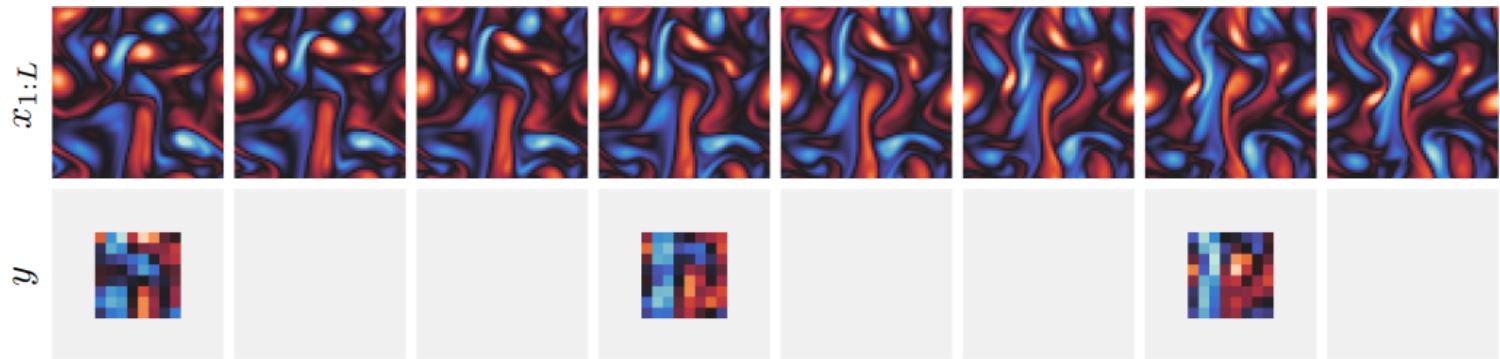


L'estimation de trajectoires plausibles  $x_{1:L}$  à partir d'observations bruitées  $y$  est un problème d'**assimilation de données**.

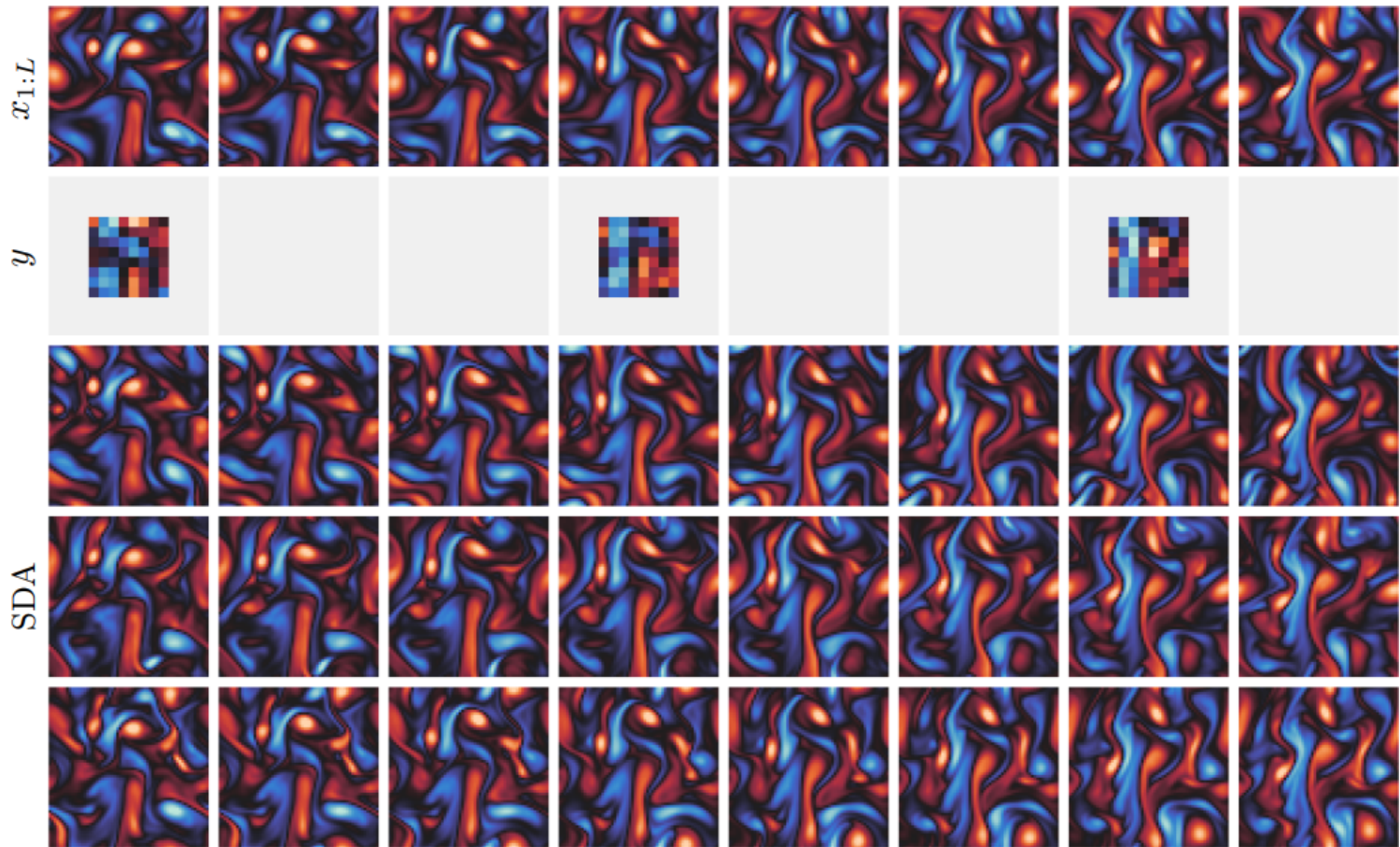


Notre approche (SDA):

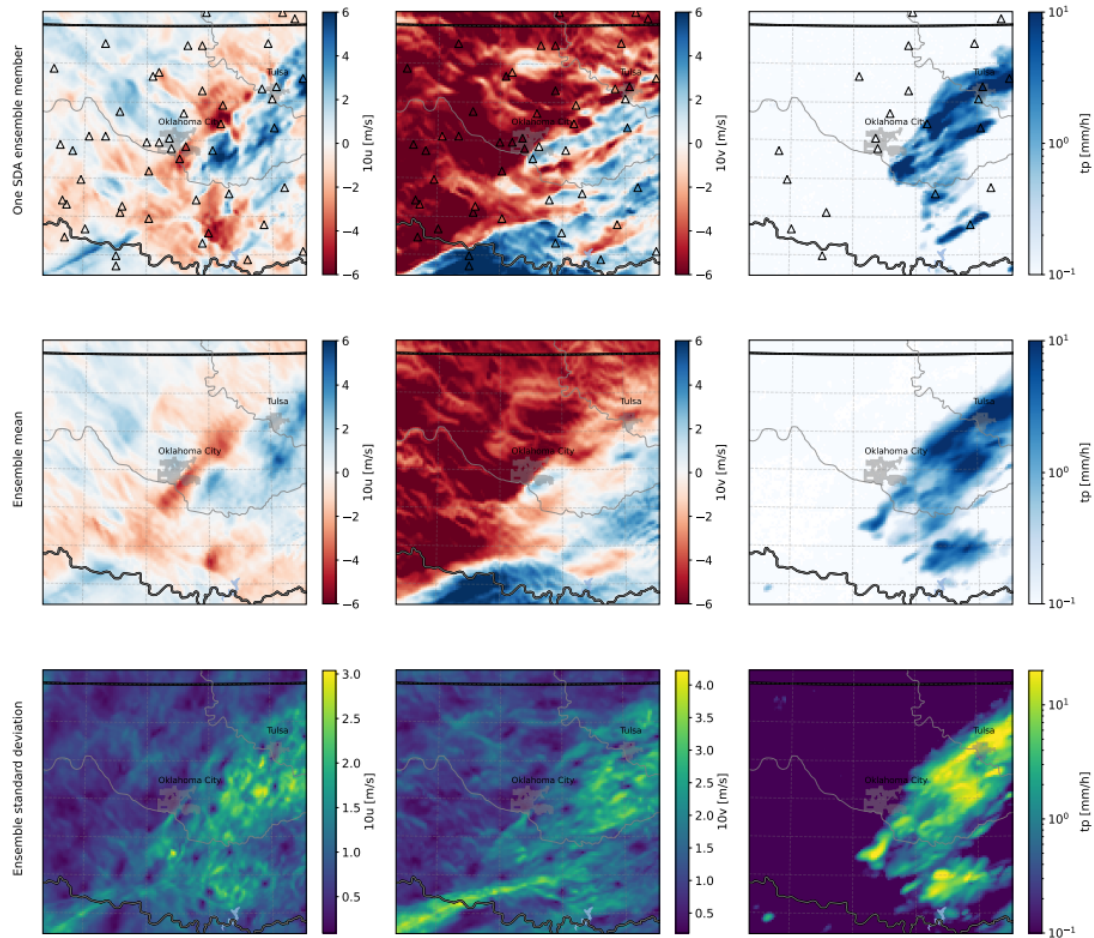
- Construire un modèle de diffusion pour générer des trajectoires  $x_{1:L}$ .
- Conditionner la génération à des observations bruitées  $y$ .



Exemple de trajectoires  $x_{1:L}$  générées à partir d'observations bruitées  $y$ .



Exemple de trajectoires  $x_{1:L}$  générées à partir d'observations bruitées  $y$ .

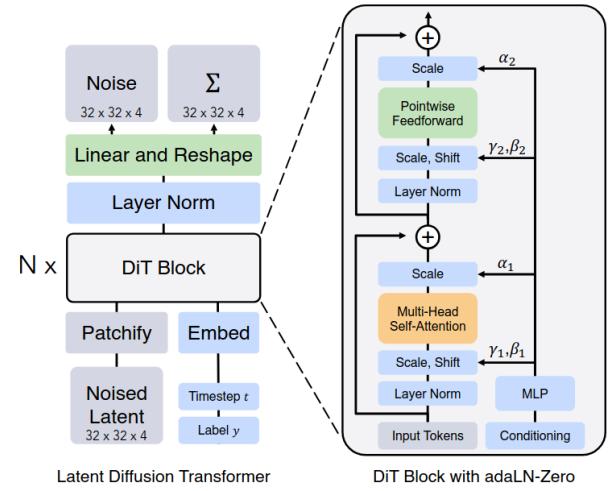
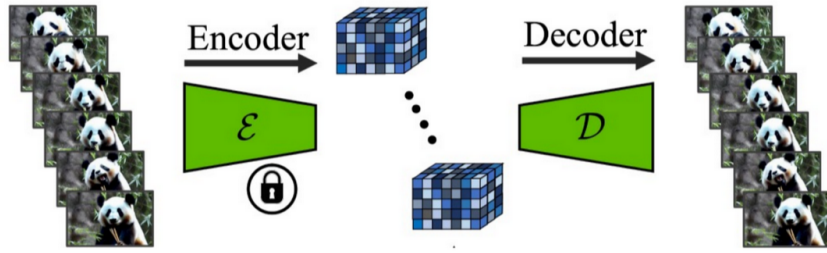


SDA peut assimiler des observations météorologiques bruitées.  
 $x$  = état météorologique régional,  $y$  = observations ponctuelles



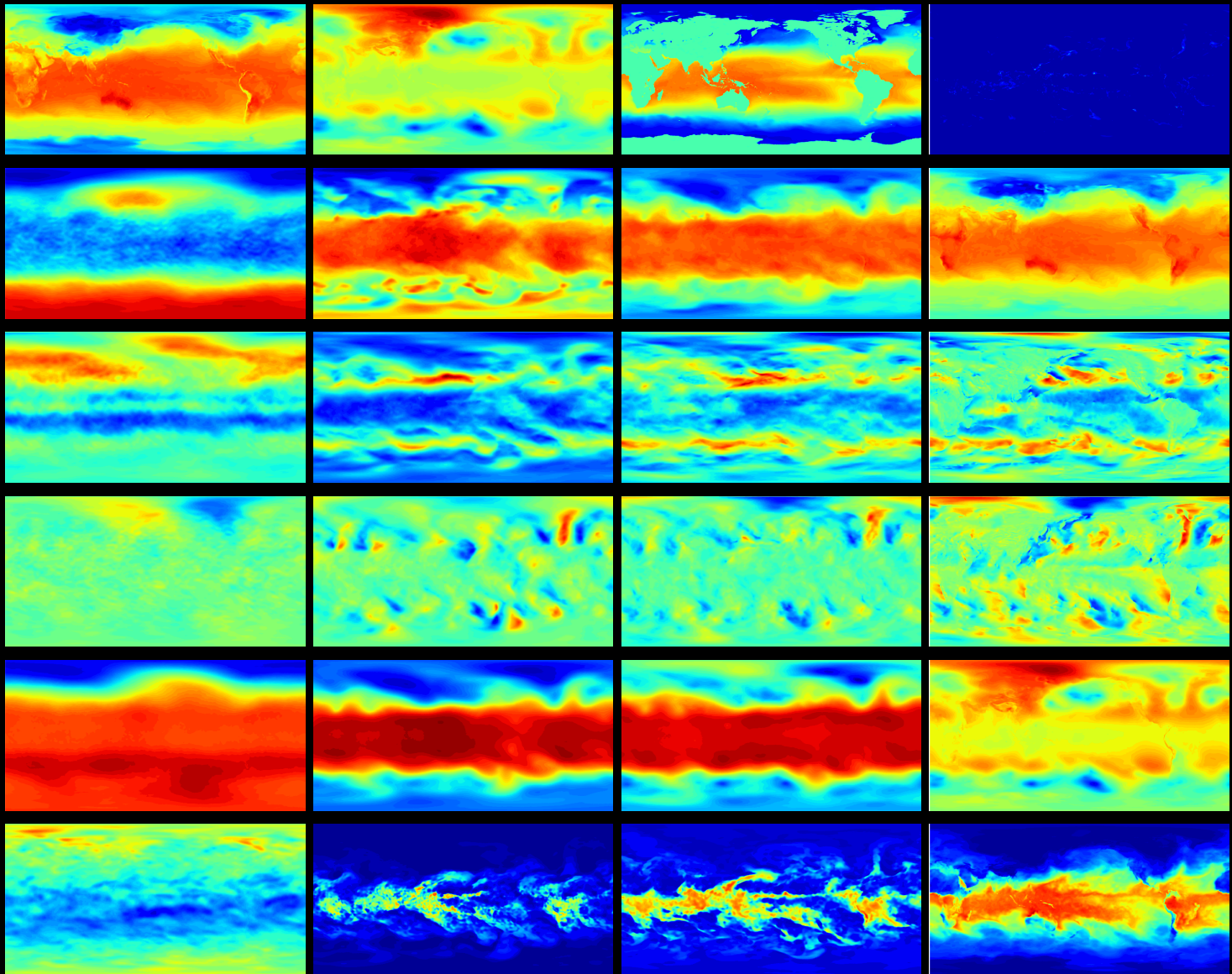
À l'échelle de la Terre (pour une résolution de  $0.25^\circ$ ), **une trajectoire  $x_{1:L}$  contient trop de variables.**

- $1440 \times 720 \times 6 \times 37 \times 24 \times 14 = 77 \times 10^9$  variables pour 14 jours.
- Un modèle de diffusion pour la Terre nécessiterait **des ressources colossales.**



Notre approche (en cours):

- Un **autoencoder** pour encoder et décoder des trajectoires  $x_{1:L}$  vers un espace latent de dimension réduite.
- Un **video diffusion model** (basé sur une variante spatiotemporelle du DiT) produisant des trajectoires  $z_{1:L}$  latentes.





## Conclusions

- Les modèles de diffusion sont pas que des générateurs d'images.
- Ils sont un outil pour traiter des problèmes inverses de haute dimension en science.



(Gérome, François, Victor, Omer, Sacha, Matthias, Elise)