VAE as a Stochastic Multidimensional Extension to Gaussian Anamorphosis

Daisuke Hotta[1], in collaboration with Bûn-Kim San[2], Loïk Berre[3] and Vincent Chabot[3]

[1] Meteorological Reserch Instutite, Japan Meteorological Agency

[2] École Nationale de la Météorologie, Météo-France

[3] Centre National de Recherches Météorologiques, Météo-France

1. Motivation

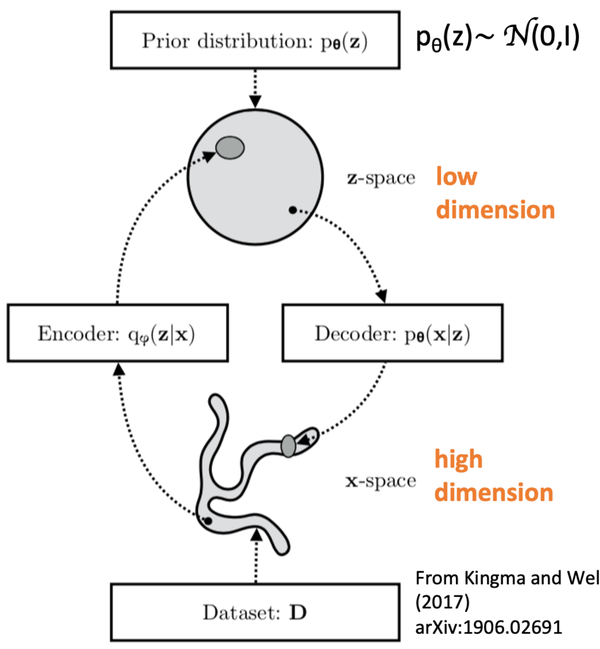
- **Image data** (e.g. satellite radiances, radar reflectivity) are increasing becoming important for NWP.
- Challenges in assimilating image data:
  - (1) non-Gaussian error distribution
  - (2) dimensional redundancy
  - (3) inter-pixel correlation

- **Gaussian Anamorphosis (GA)** is effective in accounting for non-Gaussianity, but in practice, can be done only in a univariate manner

- 💡 **Variational AutoEncoder (VAE)** solves all of these three challenges simultaneously!

- **Objective of this study:**
  - to formulate VAE as a multivariate GA
  - to assess effectiveness of its application to R-tuning

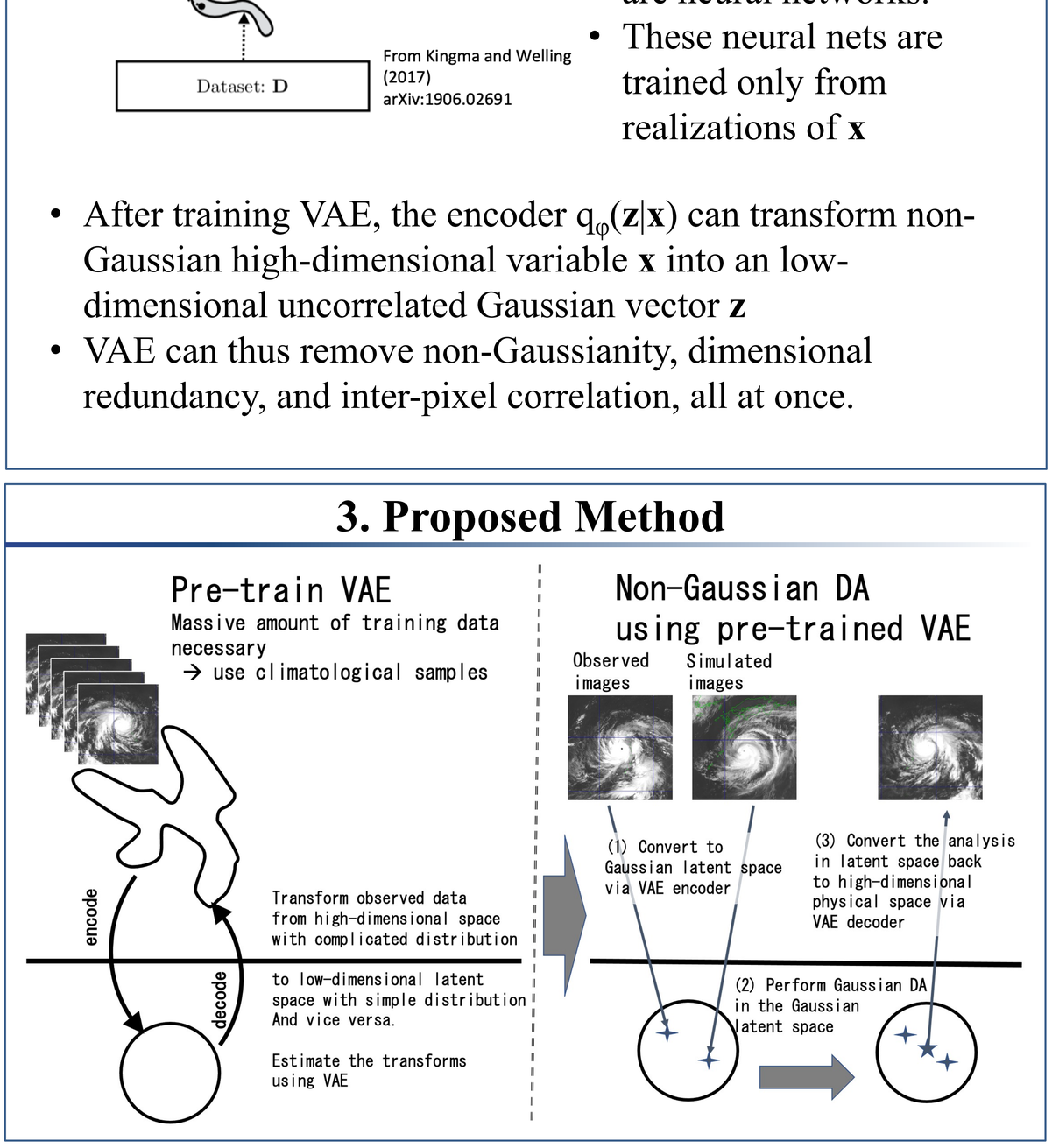
2. VAE: Schematic illustration



- High-dim image **x** is generated from low-dim uncorrelated Gaussian vector **z** that follow $\mathcal{N}(0,\mathbf{I})$
- Mapping from **z** to **x**, $p(\mathbf{x}|\mathbf{z})$, and **x** to **z**, $p(\mathbf{z}|\mathbf{x})$, are approximated by Gaussian distributions whose mean and variances are neural networks.
- These neural nets are trained only from realizations of **x**

- After training VAE, the encoder $q_\varphi(\mathbf{z}|\mathbf{x})$ can transform non-Gaussian high-dimensional variable **x** into an low-dimensional uncorrelated Gaussian vector **z**
- VAE can thus remove non-Gaussianity, dimensional redundancy, and inter-pixel correlation, all at once.

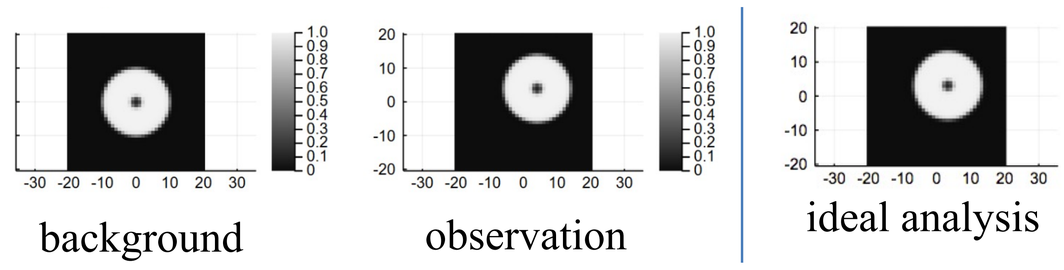
3. Proposed Method



4. Idealized experiments with a toy system

Problem

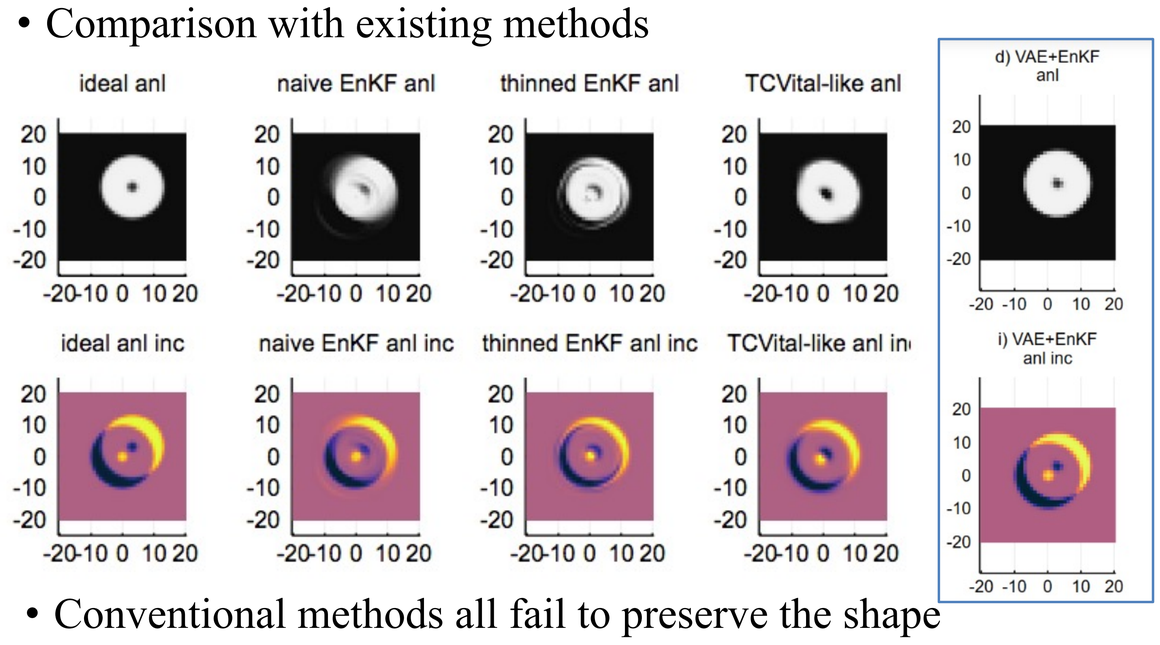
Given this background and obs, what the analysis should be like?



Easy problem (to human eyes)
Solution: Simply shift the entire picture by a fixed amount.

Results

- Comparison with existing methods

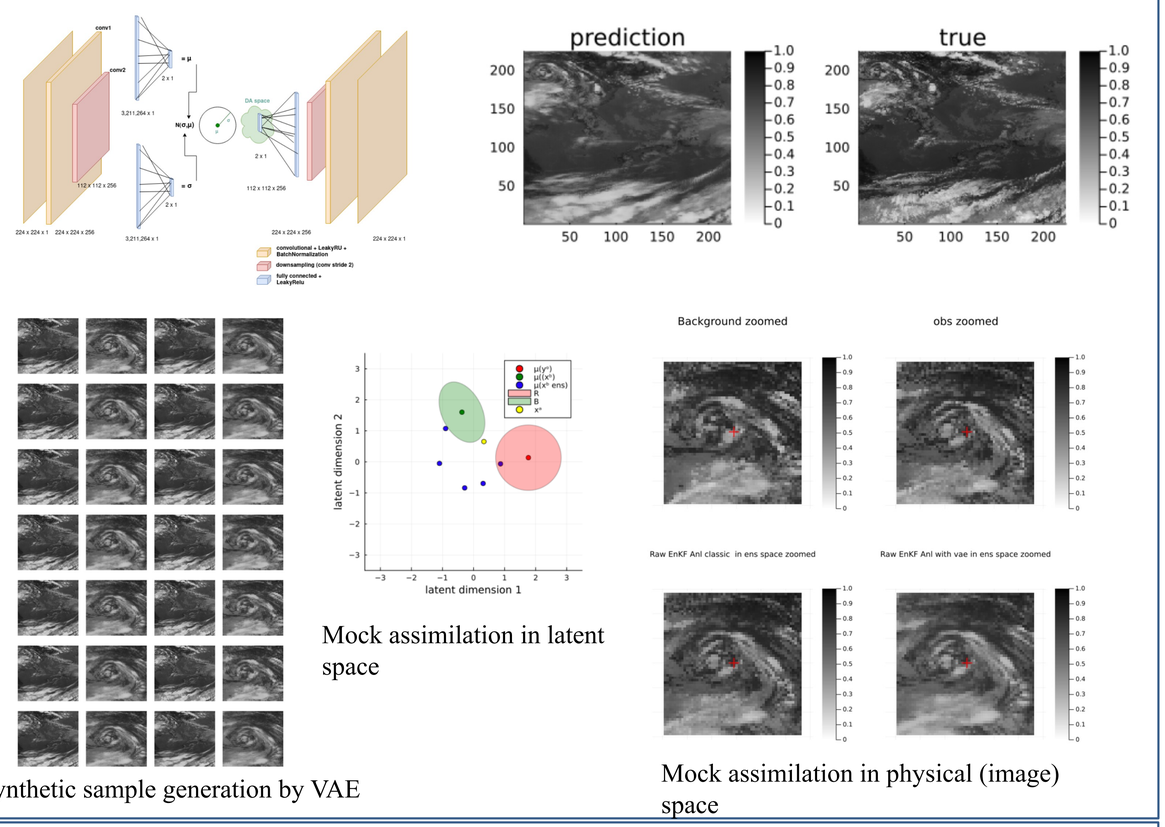


- Conventional methods all fail to preserve the shape
- With VAE, no deformation to the shape

5. VAE training on real satellite data

- Work by Bûn-Kim San (ENM, Météo-France) co-supervised with Loïk Berre and Vincent Chabot (CNRM, Météo-France)

- Train VAE on 7 sequences of Himawari IR images sampled at 10-minute intervals, with **the latent space having only 2 dimensions**



Synthetic sample generation by VAE

Mock assimilation in latent space

Mock assimilation in physical (image) space

6. Summary

- VAE enables multi-dimensional Gaussian anamorphosis that is applicable to complicated 2D images like satellite radiances
- VAE also achieves dimension reduction and removal of correlation
- Gaussian assimilation in the latent space of VAE preserves features of the data
- VAE is powerful enough to represent data as complex as real infrared satellite radiance (work by Bûn-Kim San)