VAE-Var: Variational Autoencoder-Enhanced Variational Methods for Data Assimilation in Meteorology

Yi Xiao, Qilong Jia, Kun Chen, Lei Bai, Wei Xue

Tsinghua University

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

Background

Data Assimilation

prior predictions ($\mathbf{x}_b$) + observations ($\mathbf{y}$) → analysis ($\mathbf{x}_a$)

Variational Methods: $\mathbf{x}_a = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{x}_b, \mathbf{y})$

$$\mathbf{x}_a = \arg\min_{\mathbf{x}} \big(\mathcal{L}_o(\mathbf{x}, \mathbf{y}) + \mathcal{L}_b(\mathbf{x}, \mathbf{x}_b)\big)$$

observation term	background term
$\mathcal{L}_o(\mathbf{x}, \mathbf{y}) = -\log p(\mathbf{y}|\mathbf{x})$	$\mathcal{L}_b(\mathbf{x}, \mathbf{x}_b) = -\log p(\mathbf{x}|\mathbf{x}_b)$
$\mathbf{y}|\mathbf{x} \sim \mathcal{N}(\mathcal{H}(\mathbf{x}), \mathrm{R})$	High dimensional Complicated

$$\mathcal{L}_o(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\big(\mathbf{y} - \mathcal{H}(\mathbf{x})\big)^{\mathrm{T}} \mathrm{R}^{-1} \big(\mathbf{y} - \mathcal{H}(\mathbf{x})\big)$$

Traditional Algorithm (3DVar)

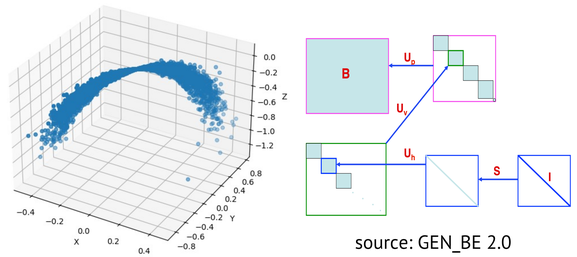
Assumption 1: $\mathbf{x}|\mathbf{x}_b \sim \mathcal{N}(\mathbf{x}_b, \mathrm{B})$

Assumption 2: $\mathrm{B} = \mathrm{U}\mathrm{U}^{\mathrm{T}}, \mathrm{U} = \mathrm{U_p U_v U_h S}$

VAE-Var

Purpose: improve the formulation of the background term with the help of neural networks.

Motivation



source: GEN_BE 2.0

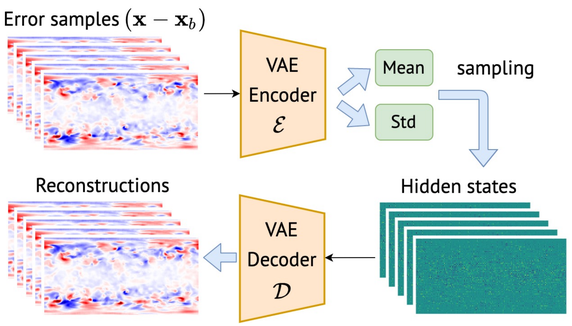
Gaussian assumption may not be correct.

Heavily reply on expert knowledge

Learning $p(\mathbf{x}|\mathbf{x}_b)$ with a generative neural network

Why VAE (instead of diffusion)?
Straightforward formulation, easy to train.

VAE Approximation

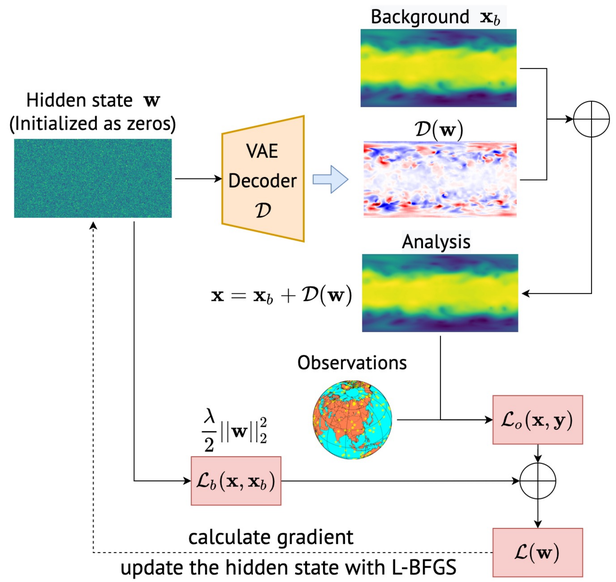


$$\log p(\mathbf{x}|\mathbf{x}_b) = \log p_\delta(\mathbf{x} - \mathbf{x}_b)$$
$$= \log \int p(\mathbf{x} - \mathbf{x}_b|\mathbf{z}) p(\mathbf{z}) \mathrm{d}\mathbf{z}$$
$$= \log \int \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_b - \mathcal{D}(\mathbf{z})\|_2^2}{2\sigma^2}\right) \exp\left(-\frac{\|\mathbf{z}\|_2^2}{2}\right) \mathrm{d}\mathbf{z}$$

Approximation: let $\mathbf{x} = \mathbf{x}_b + \mathcal{D}(\mathbf{w})$, then

$$-\log p_\delta\big(\mathcal{D}(\mathbf{w})\big) \le \frac{\lambda}{2}\|\mathbf{w}\|_2^2 + C$$

VAE-Var Framework



Results

Dynamical System

WeatherBench (a global weather system).
Variables: **4 surface var.** + **5 upper-air var**.
Horizontal resolution: **0.25° (1440×721)**
Vertical resolution: **13 layers**

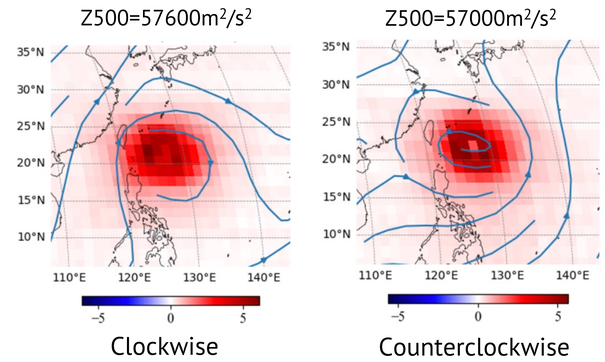
Forecasting Model

FengWu (a prominent global AI weather forecasting model)

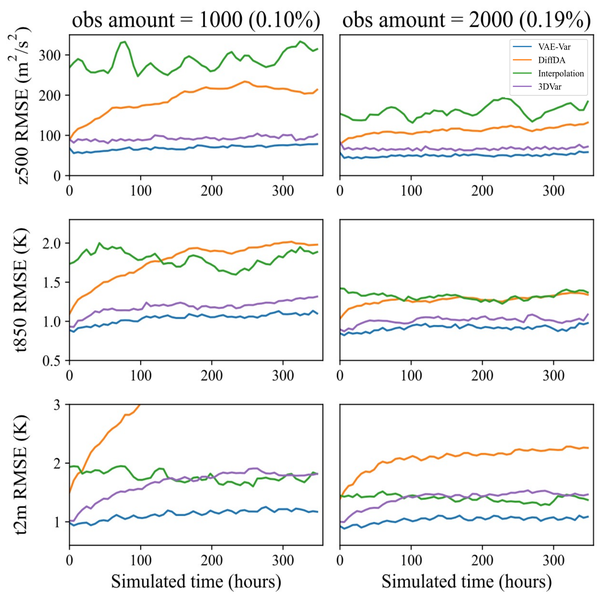
Single-Point Observation

How does an observation of Z500 at 21°N, 129°E affect the estimate of wind fields?

Analysis increment of the wind fields



Cyclic Forecasting



VAE-Var outperforms both 3DVar and DiffDA, especially under sparse observational settings.