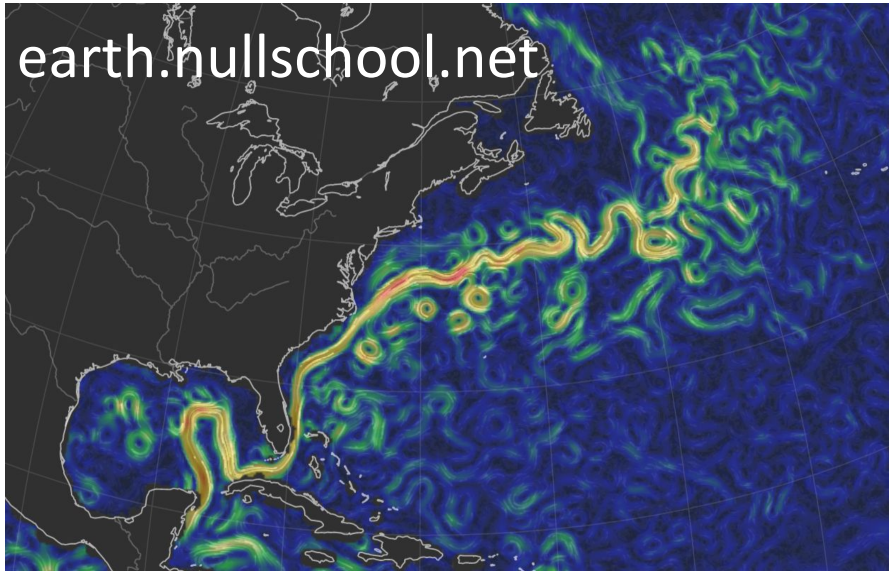
Data-driven Stochastic Parameterizations of Subgrid Mesoscale Eddies in Idealized Ocean Model

Pavel Perezhogin[1], Laure Zanna[1], Carlos Fernandez-Granda[1]
(1) New York University, Courant Institute of Mathematical Sciences

10 February, CESM 2023

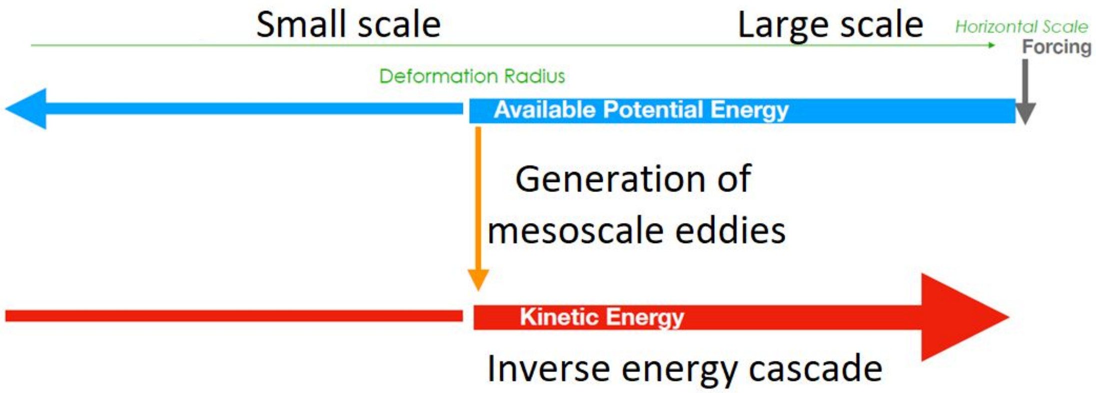
https://m2lines.github.io/

Parameterization of mesoscale eddies



- Eddies of scale 100-300km
- Carry most of kinetic energy

Zanna, Bachman & Jansen (2020). Sources and sinks of ocean mesoscale eddy energy.



- Existing parameterizations:
  - Jansen et al. (2019). Toward an energetically consistent, resolution aware parameterization of ocean mesoscale eddies.
  - Guillaumin and Zanna (2021). Stochastic-deep learning parameterization of ocean momentum forcing.

Plan of the talk

- Generative approach (ML) to build a stochastic parameterization
- Offline/online tests in 2-layer QG model

Filtered equations and subgrid forcing

(Large eddy simulation, LES)

- $$\frac{dq}{dt} = F(q)$$

Define:

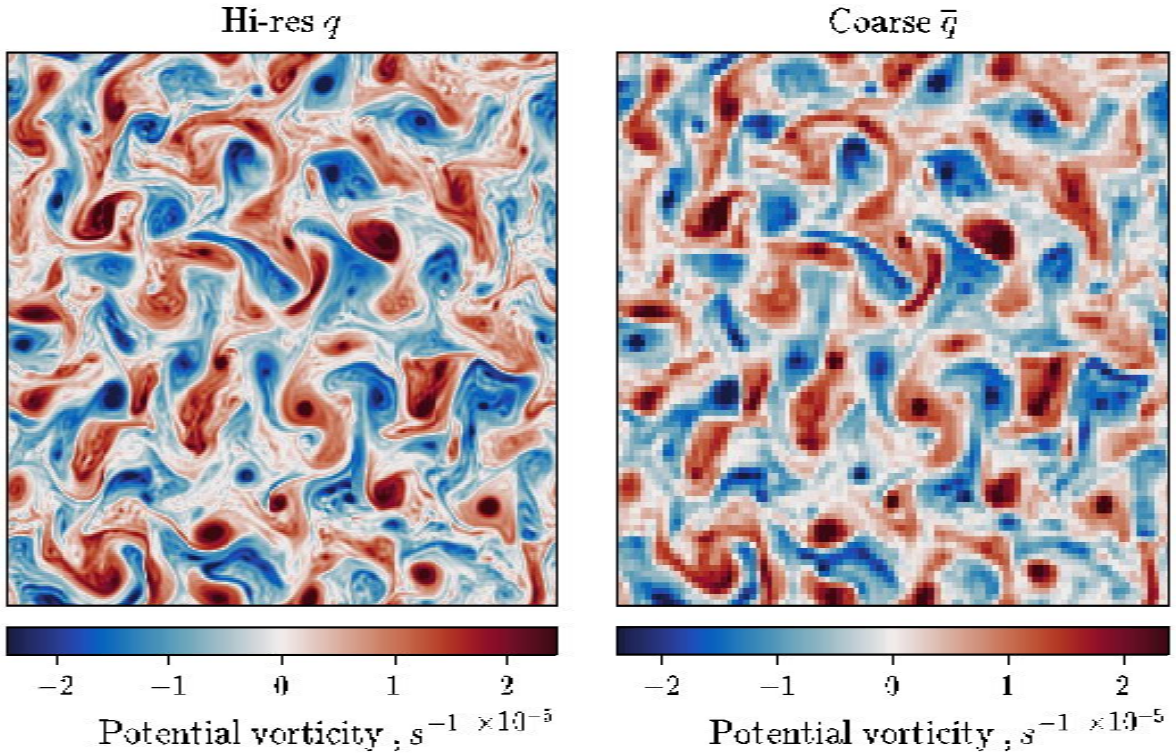
$$q \rightarrow \bar{q}$$

Apply filter to Eq.:

$$\frac{d\bar{q}}{dt} = F(\bar{q}) + S$$

Subgrid forcing:

$$\boxed{S = \overline{F(q)} - F(\bar{q})}$$



Filtered equations and subgrid forcing

(Large eddy simulation, LES)

- $$\frac{dq}{dt} = F(q)$$

Define:

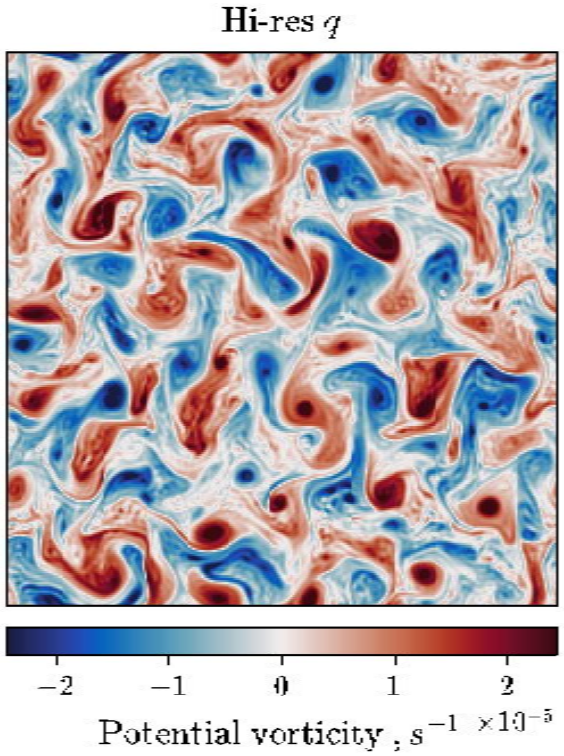
$$q \rightarrow \overline{q}$$

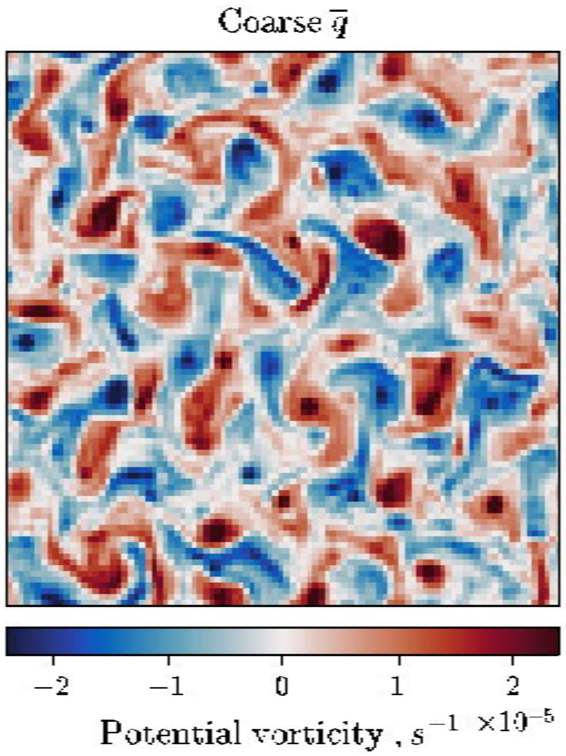
Apply filter to Eq.:

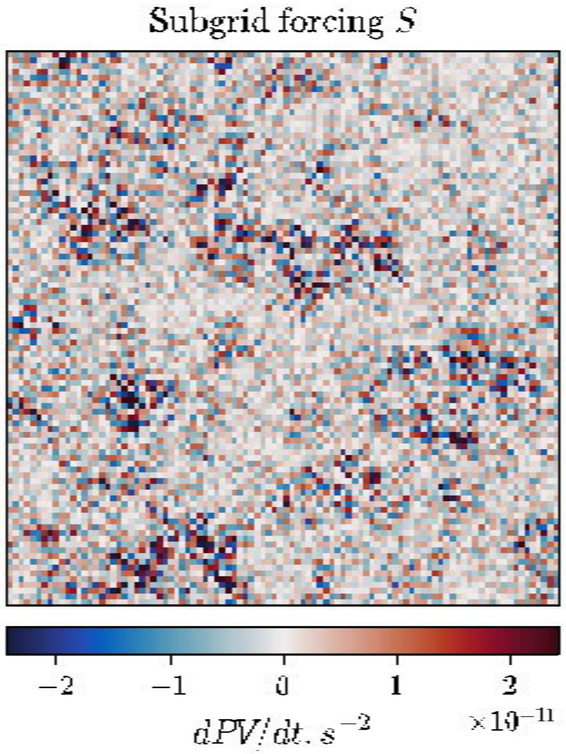
$$\frac{d\overline{q}}{dt} = F(\,\overline{q}\,) + S$$

PV flux:

$$\boxed{S = \nabla \cdot (\overline{\boldsymbol{u}}\ \overline{q} - \overline{\boldsymbol{u}q})}$$

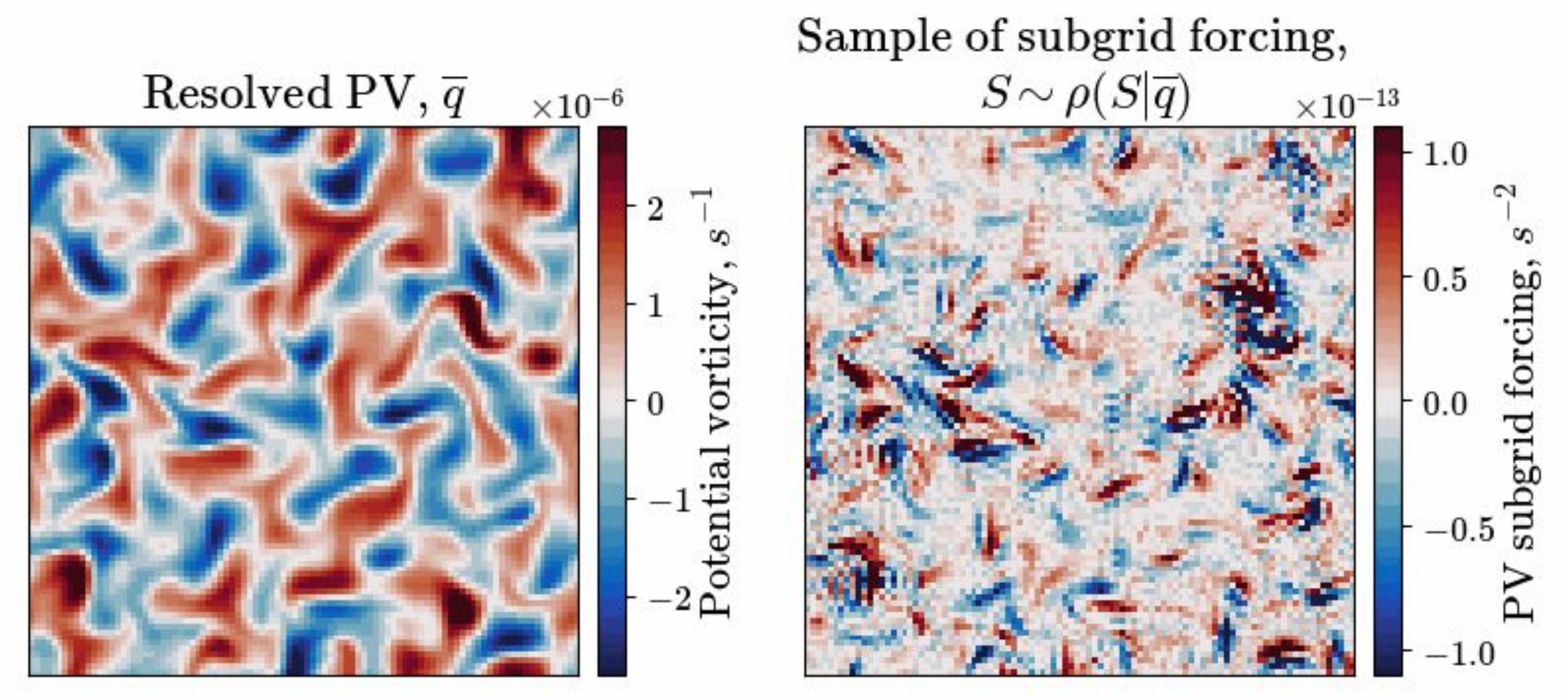






Stochastic parameterization

- To sample from the distribution of **all possible subgrid forcings conditioned on a resolved flow**



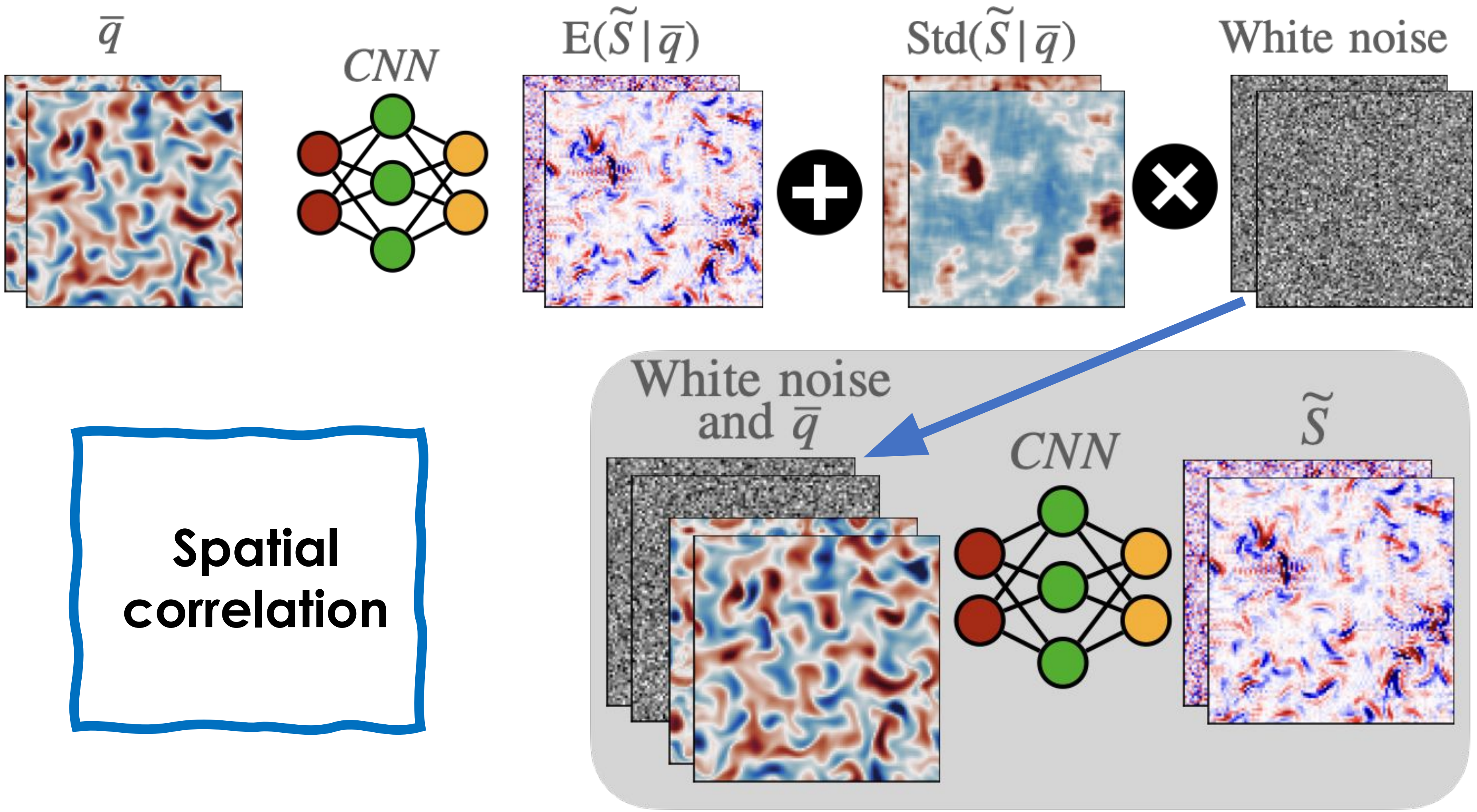
Generative approach for stochastic parameterization

Guillaumin Zanna 2021 (GZ)

No spatial correlation

We transform noise with CNN

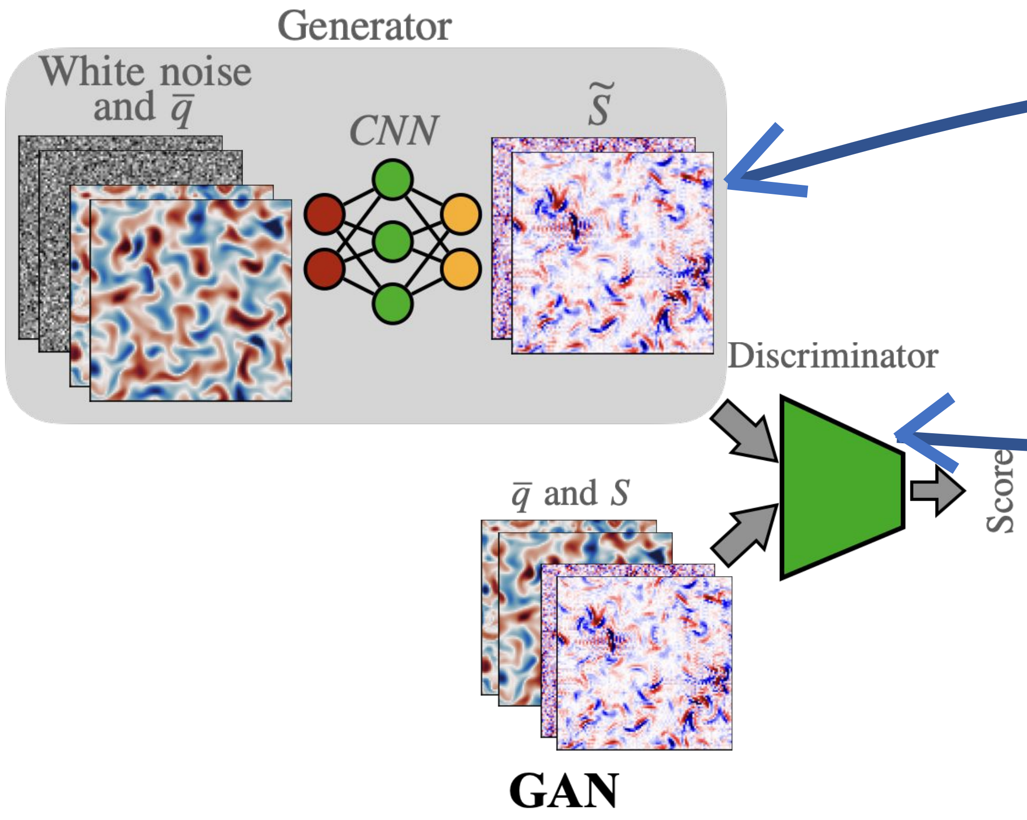
Spatial correlation



Generative adversarial network (**GAN**)

*For training was used conditional WGAN-GP + minibatch discrimination loss. *Adler Oktem. Deep Bayesian Inversion. 2018*

Simplified explanation for **unconditional WGAN***



We learn the true distribution from data

$$\min_G \max_D \mathbb{E}[D(S) - D(G(z))]$$

Adversarial game:

- Discriminator tries to detect fake images
- Generator tries to fool Discriminator

Variational autoencoder (**VAE**)

*For training was used CVAE loss with calibrated decoder. Rybkin, Daniilidis 2021.

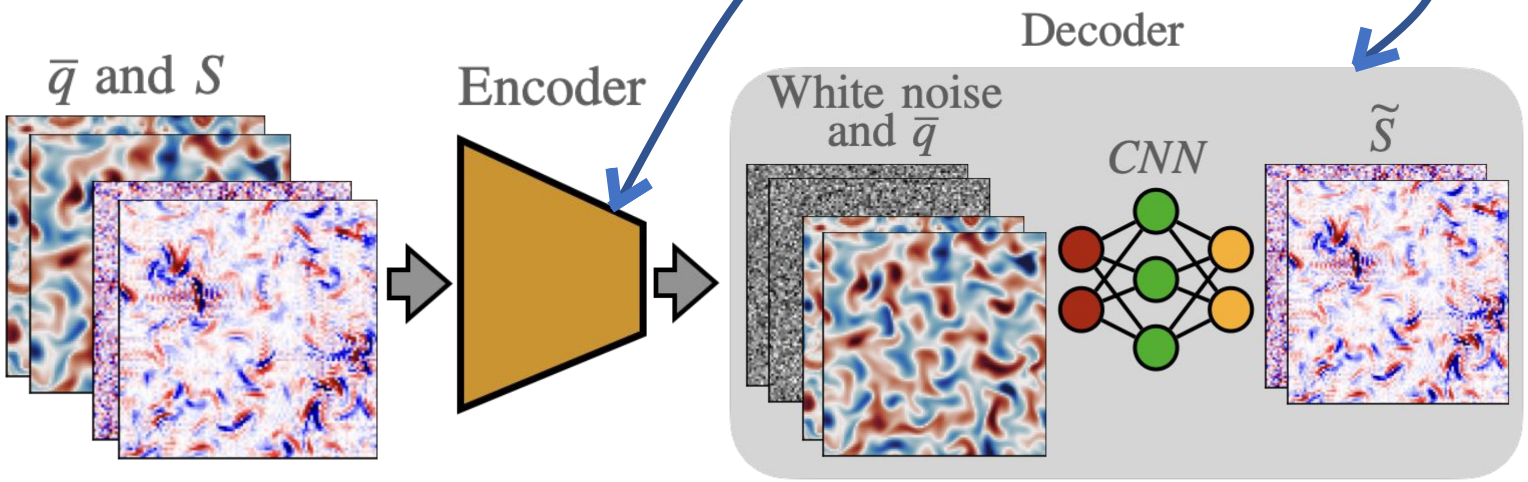
Simplified explanation for **unconditional VAE***

Variational approximation:

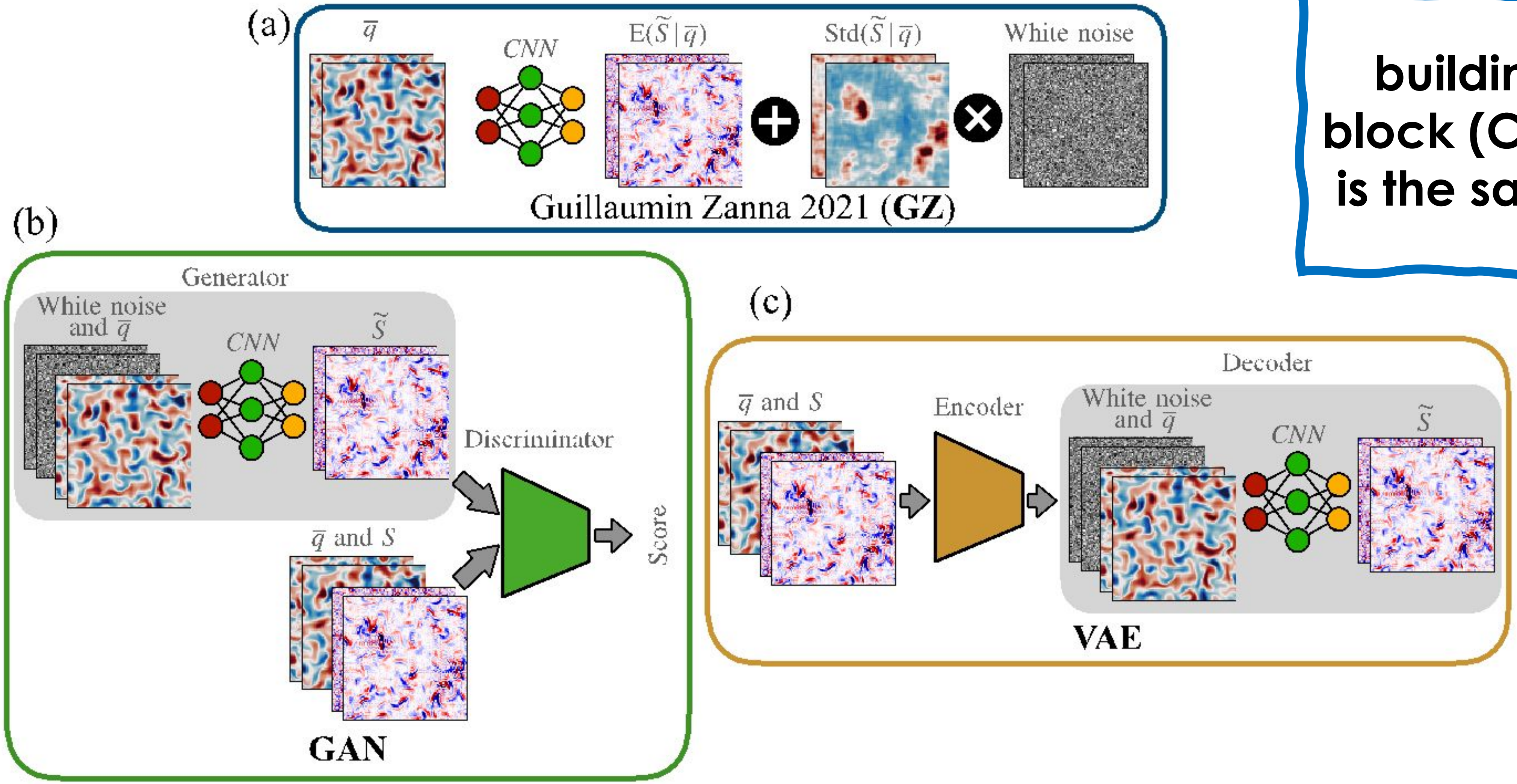
- Decoder fits data
- Encoder provides latent codes
- Regularization D_{KL} constrains latent codes

Evidence Lower Bound (**ELBO**):

$$\log \rho(S) \geq \mathbb{E}_{z \sim \rho(z|S)} \log \rho(S|z) + \mathrm{D}_{KL}(\rho(z|S), \mathcal{N}(0,1))$$

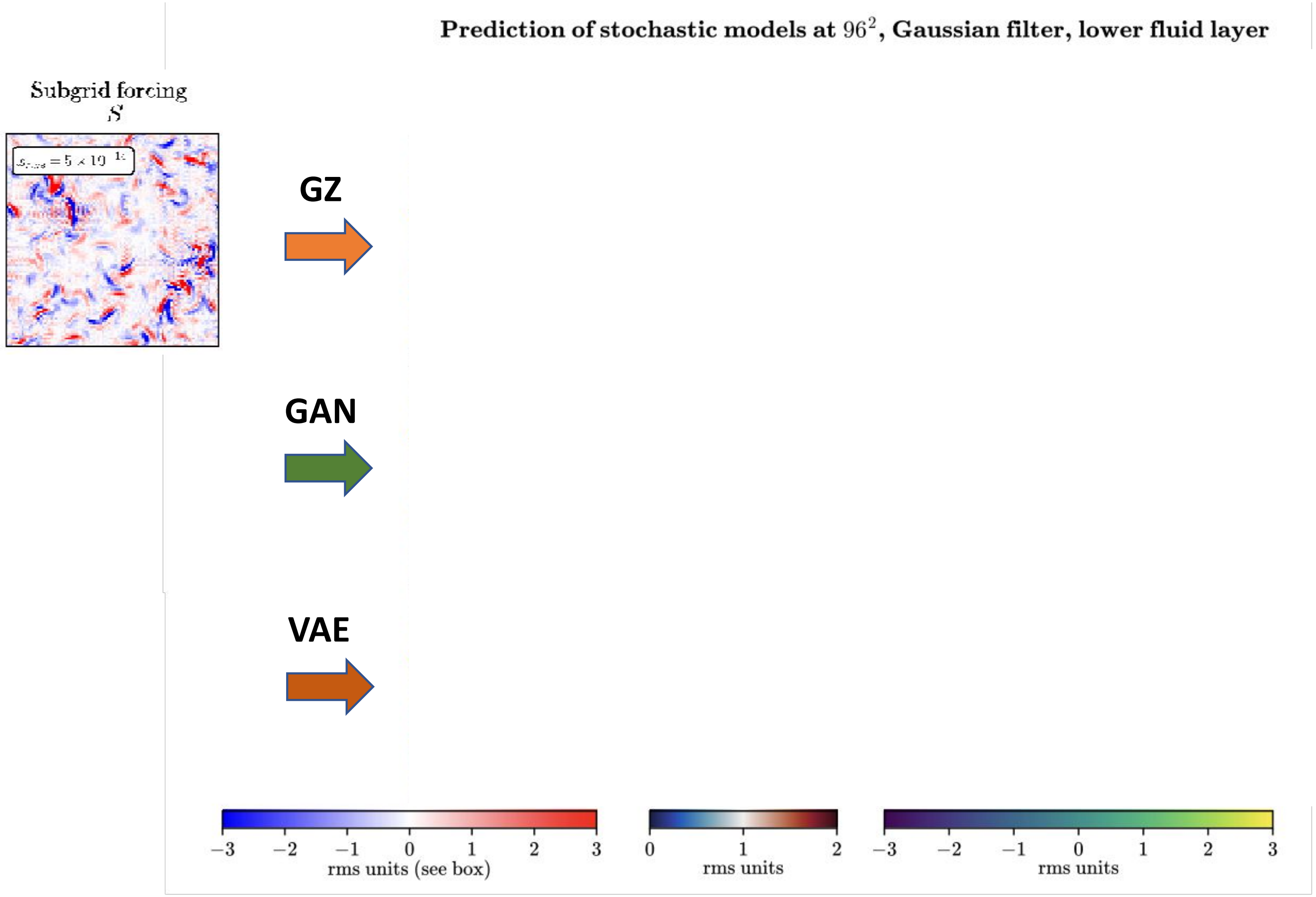


Three stochastic models



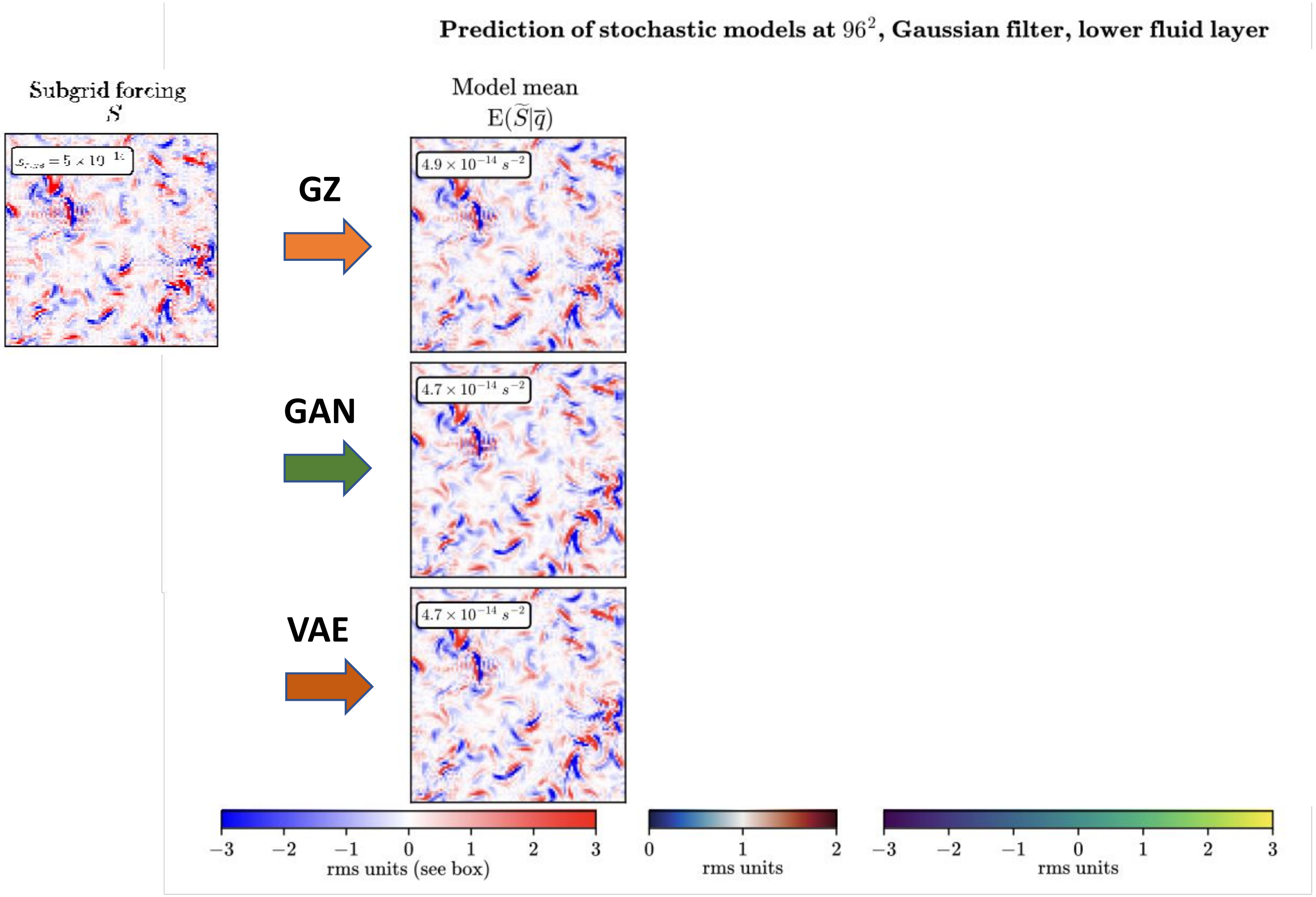
building block (CNN) is the same

Example of probabilistic prediction



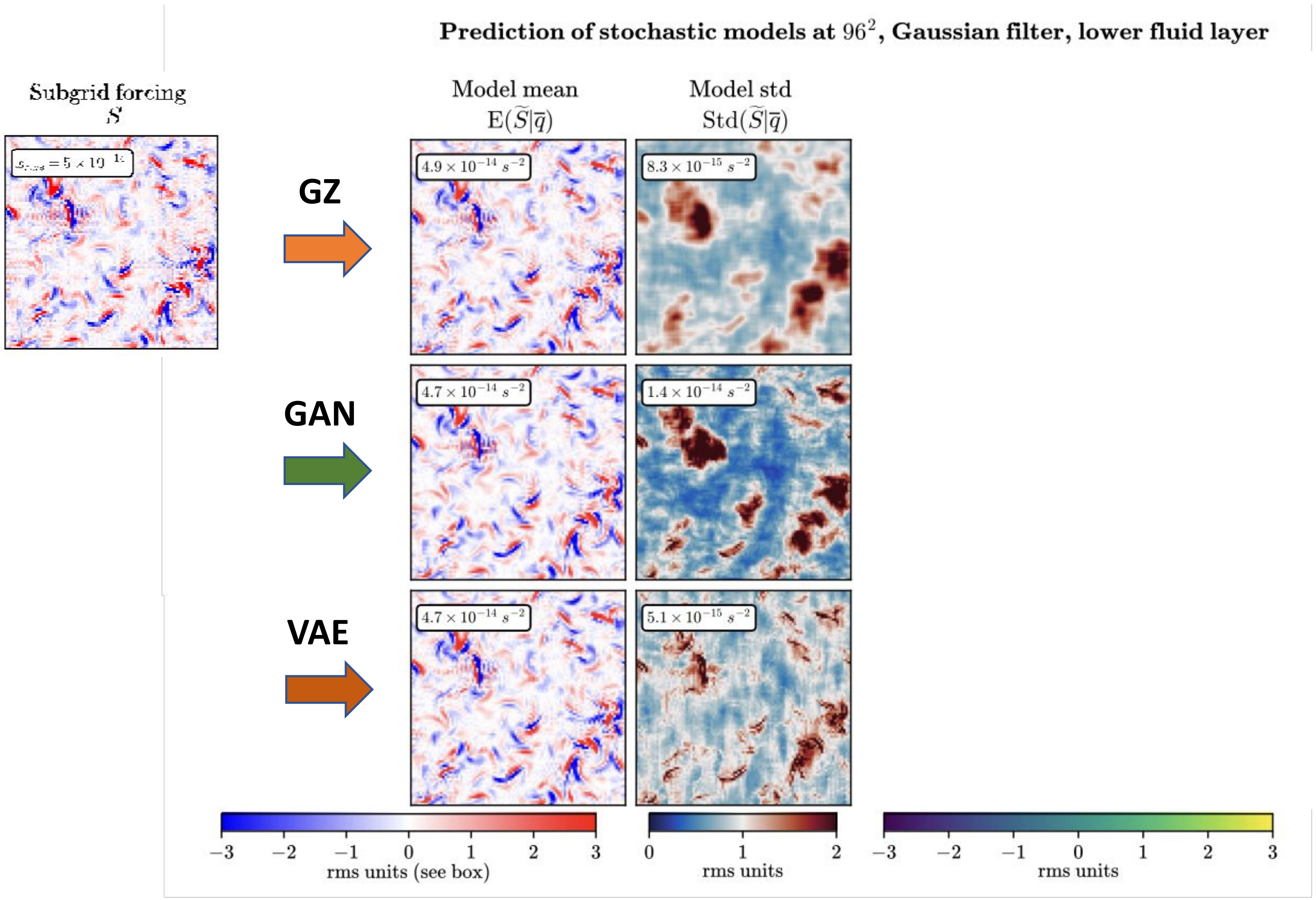
Example of probabilistic prediction

- Deterministic part



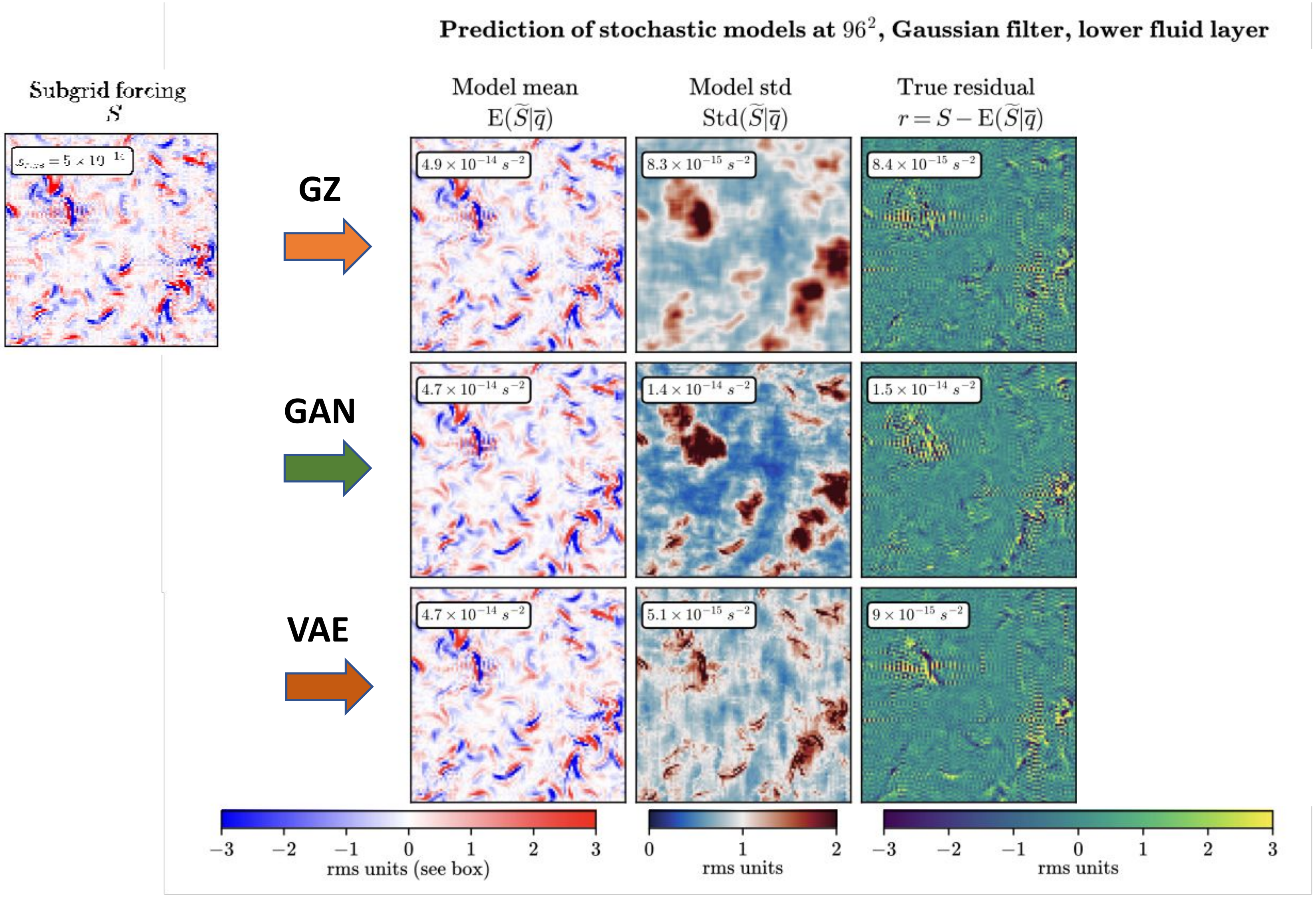
Example of probabilistic prediction

- Deterministic part
- Local std. Spots are in same locations



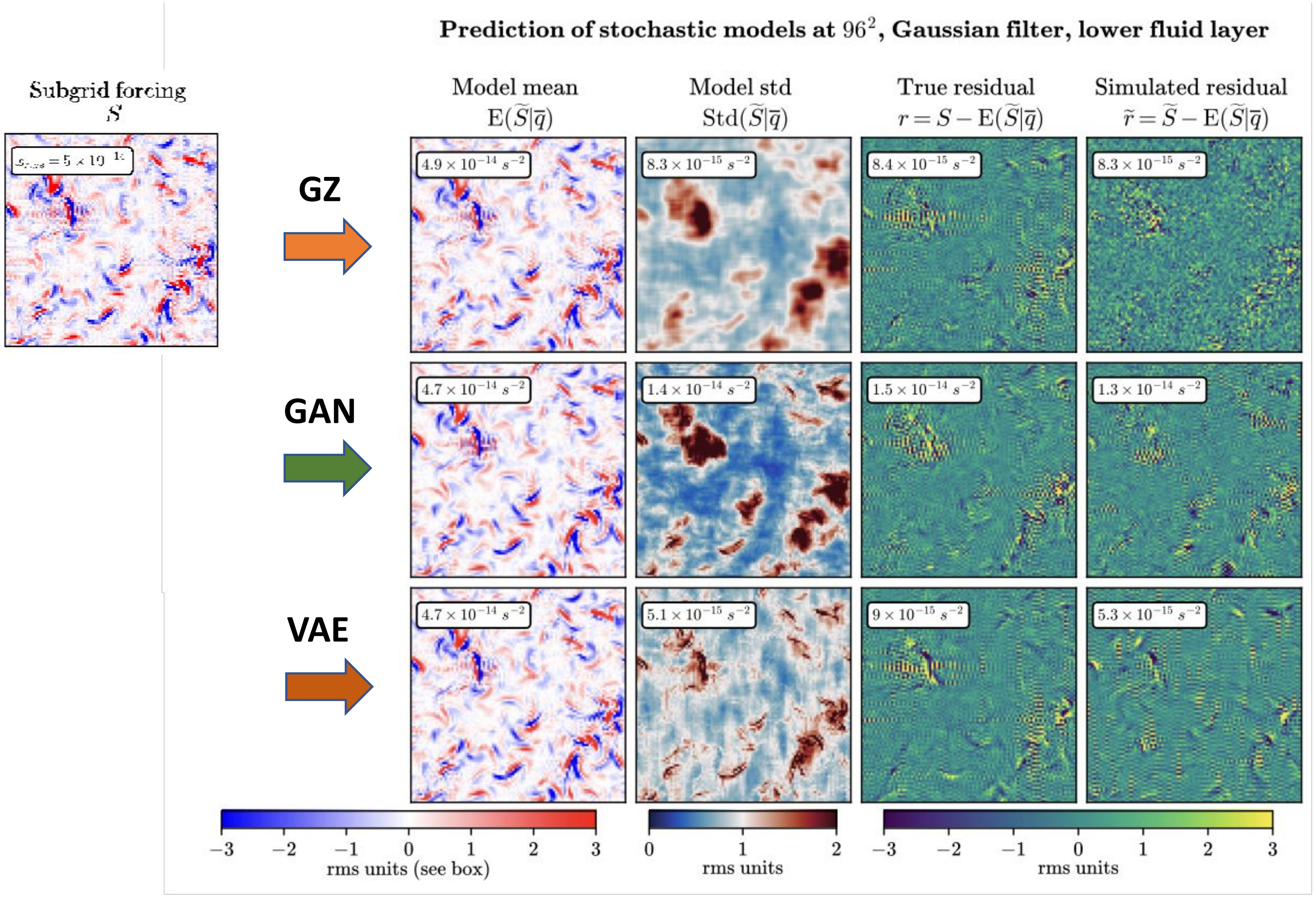
Example of probabilistic prediction

- Deterministic part
- Local variance. Spots are in same locations
- Unpredictable part of subgrid forcing



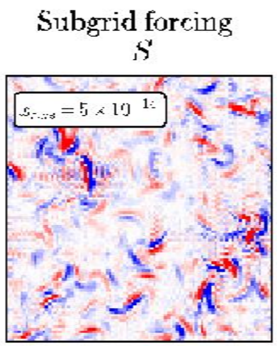
Example of probabilistic prediction

- Deterministic part
- Local variance. Spots are in same locations
- Unpredictable part of subgrid forcing
- Generation of residuals



Example of probabilistic prediction

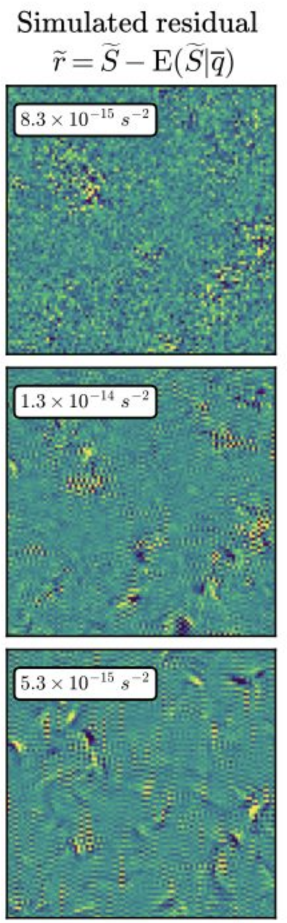
- Deterministic part
- Local variance. Spots are in same locations
- Unpredictable part of subgrid forcing
- Generation of residuals



GZ → **Uncorrelated noise**

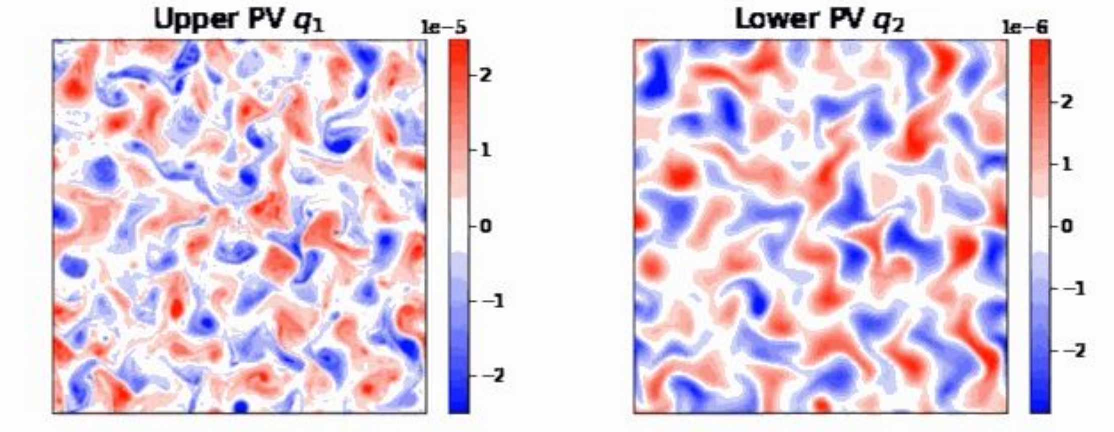
GAN → **Correlated noise**

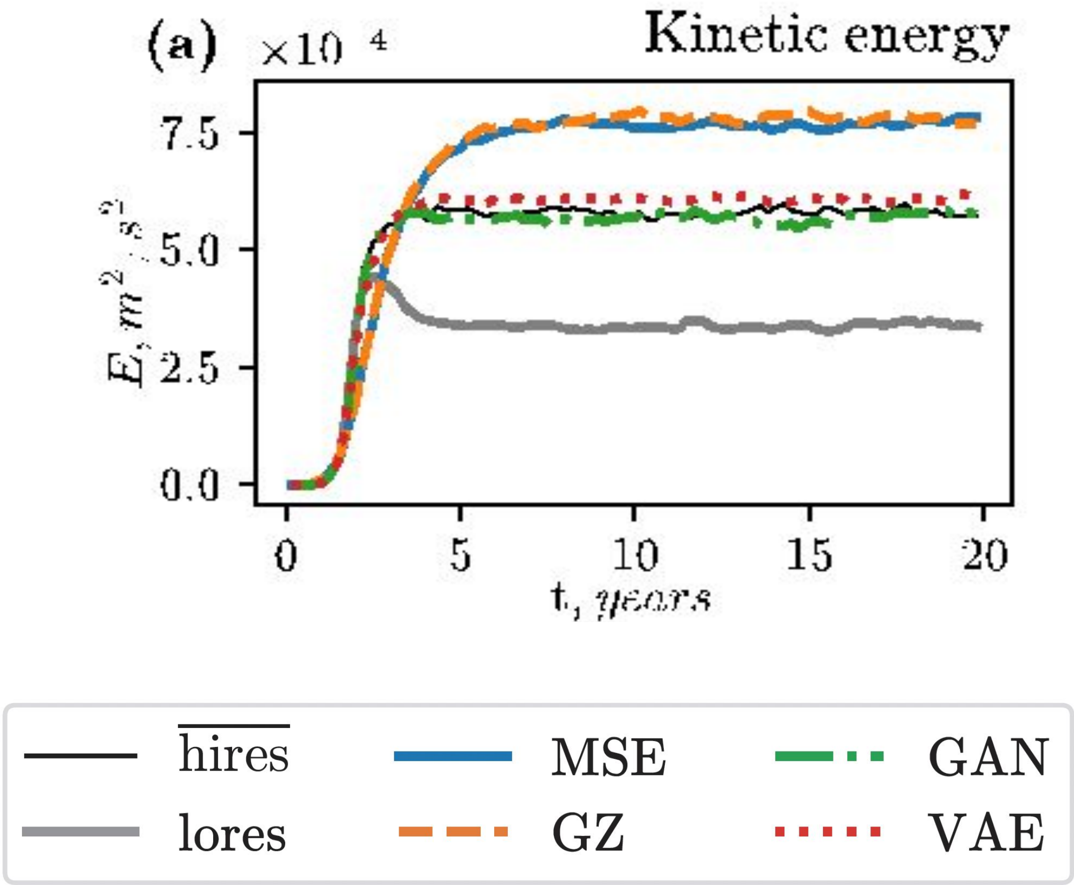
VAE → **Correlated noise**



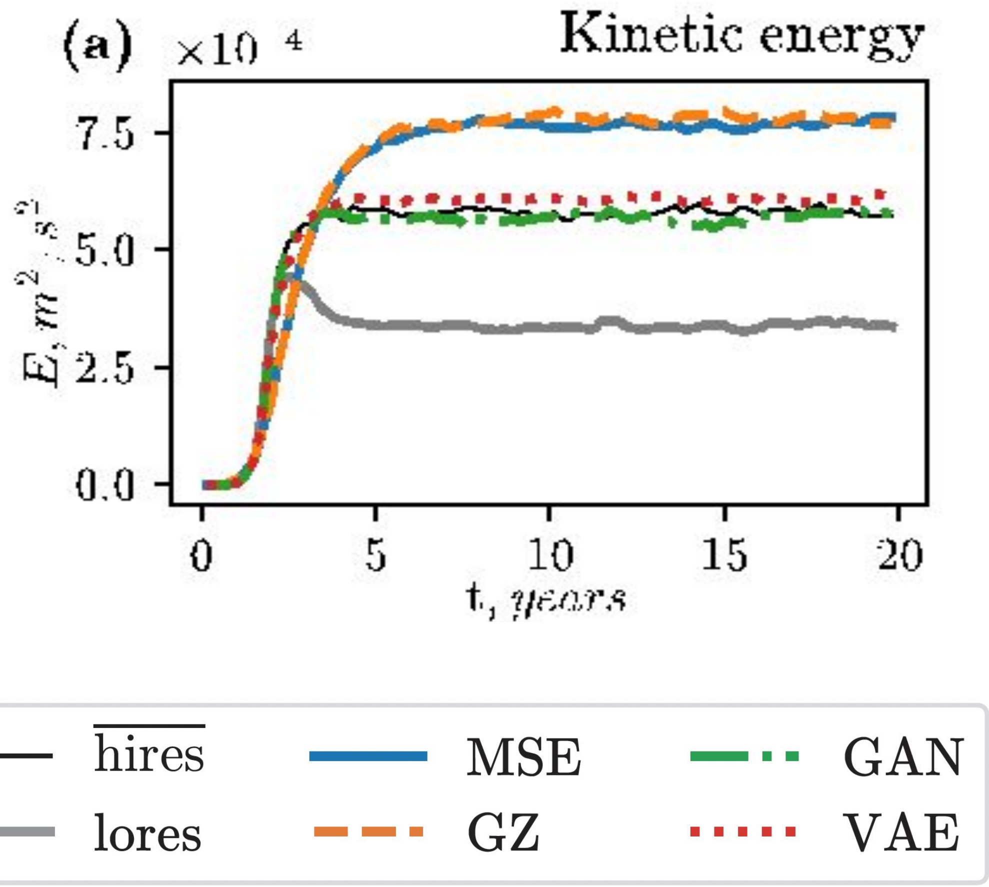
Online simulations with stochastic models

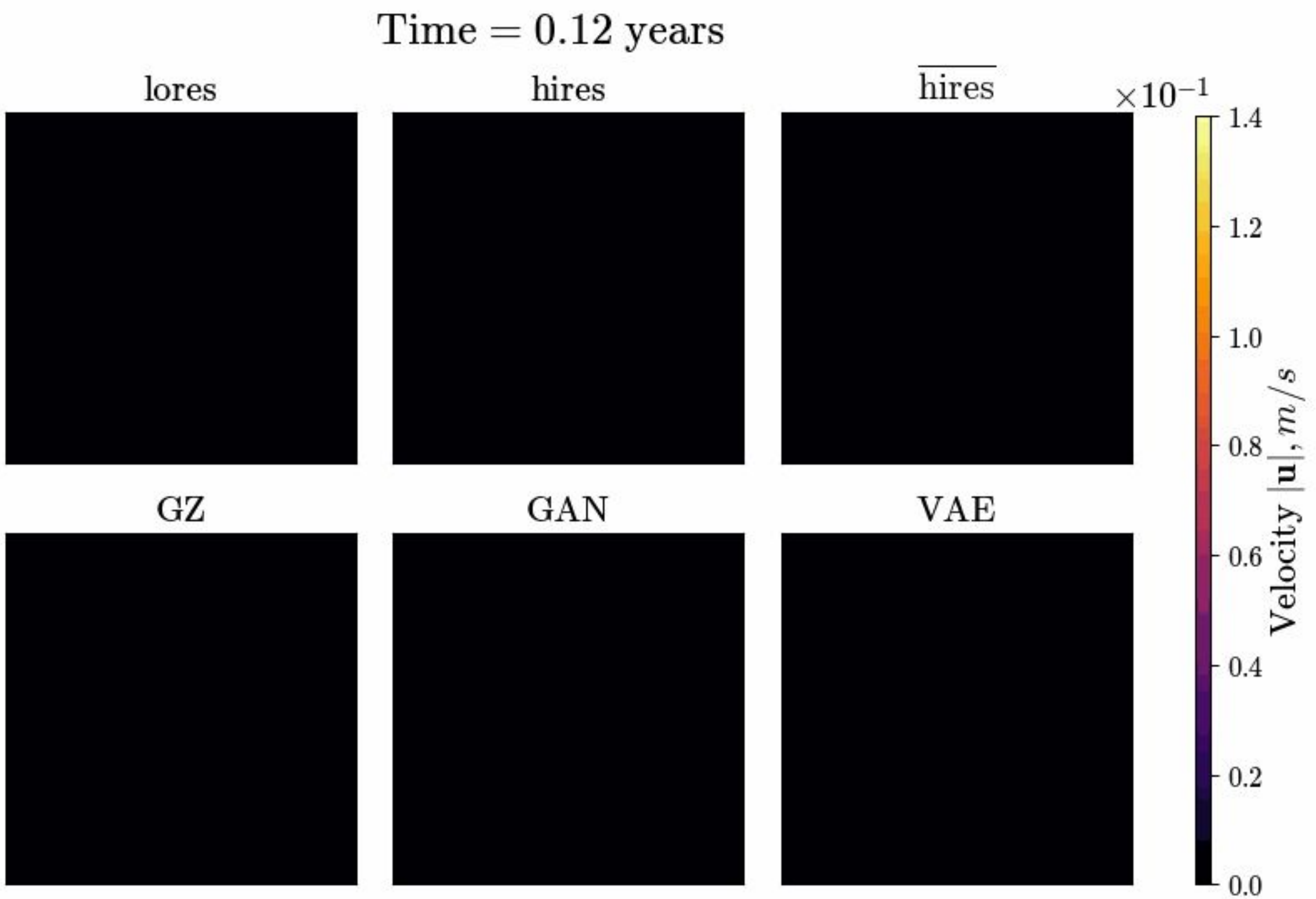
Ross 2022. JAMES.





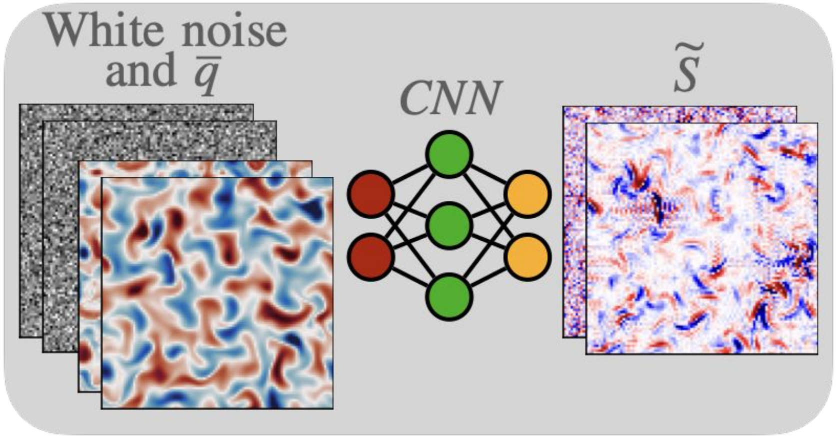
Online simulations
with stochastic models

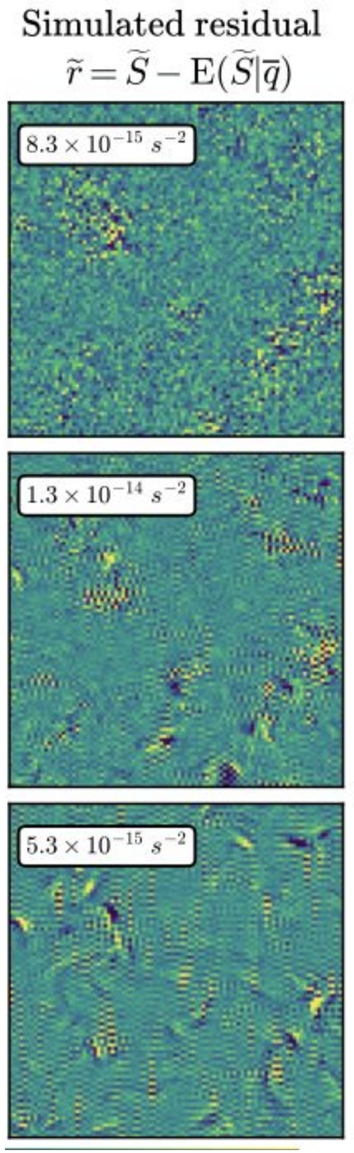


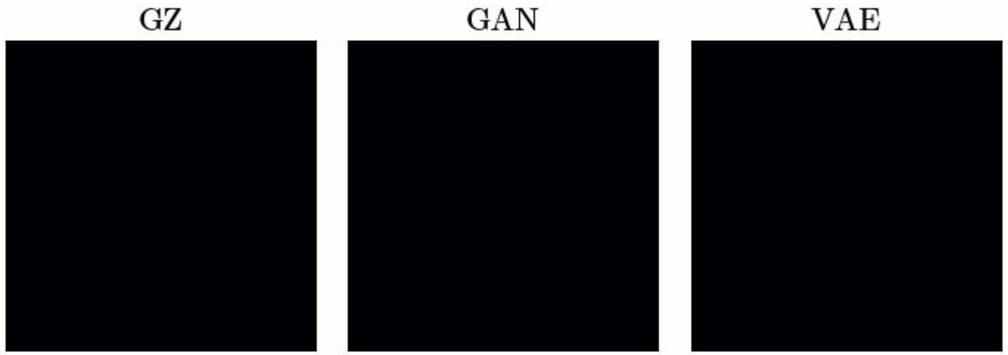


Conclusions on stochastic models

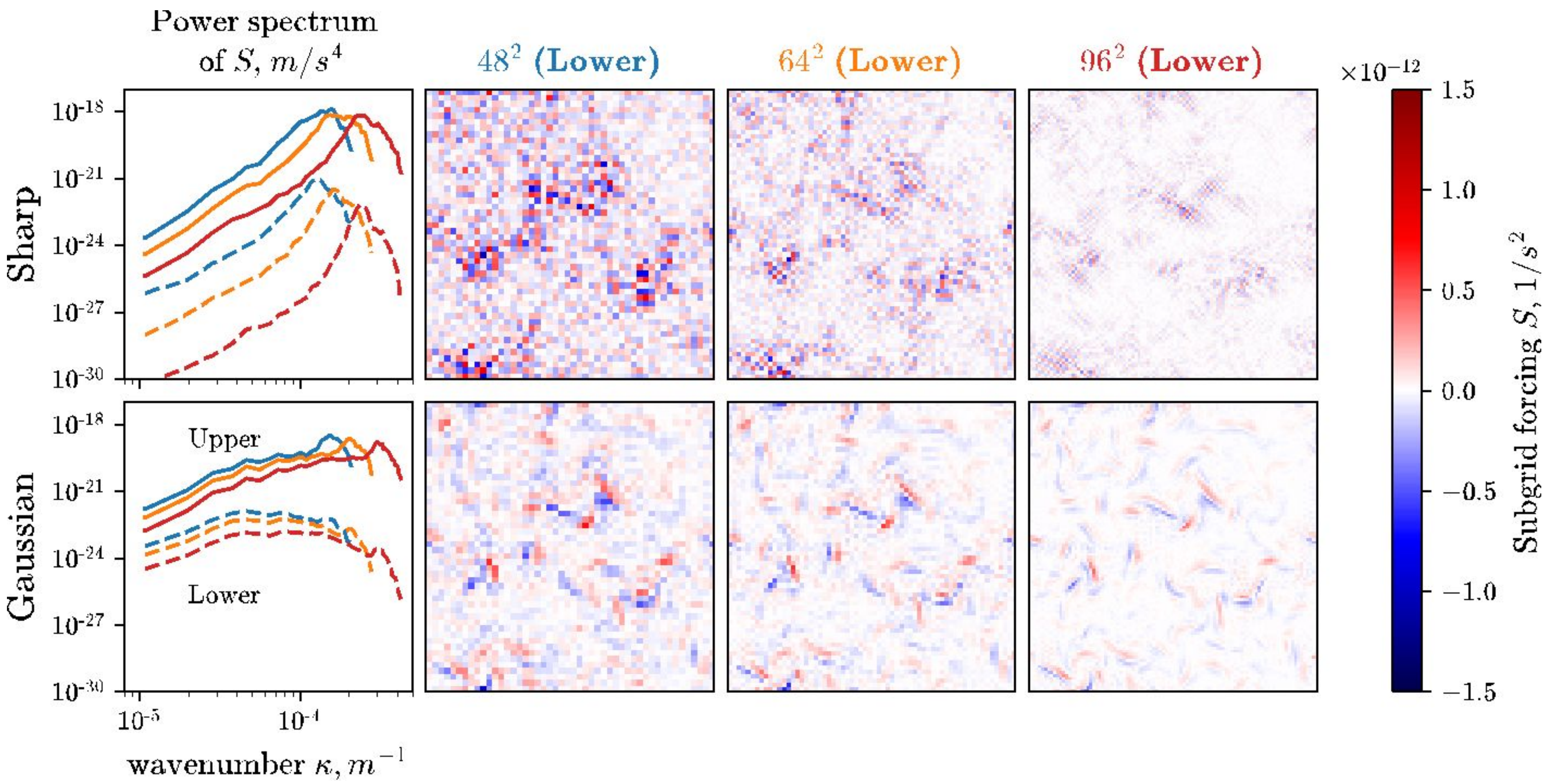
- We leverage generative models (**GAN/VAE**) to stochastic parameterization
- New models account for spatial correlation of stochastic residuals
- Improved performance in online simulations
- Improved numerical stability





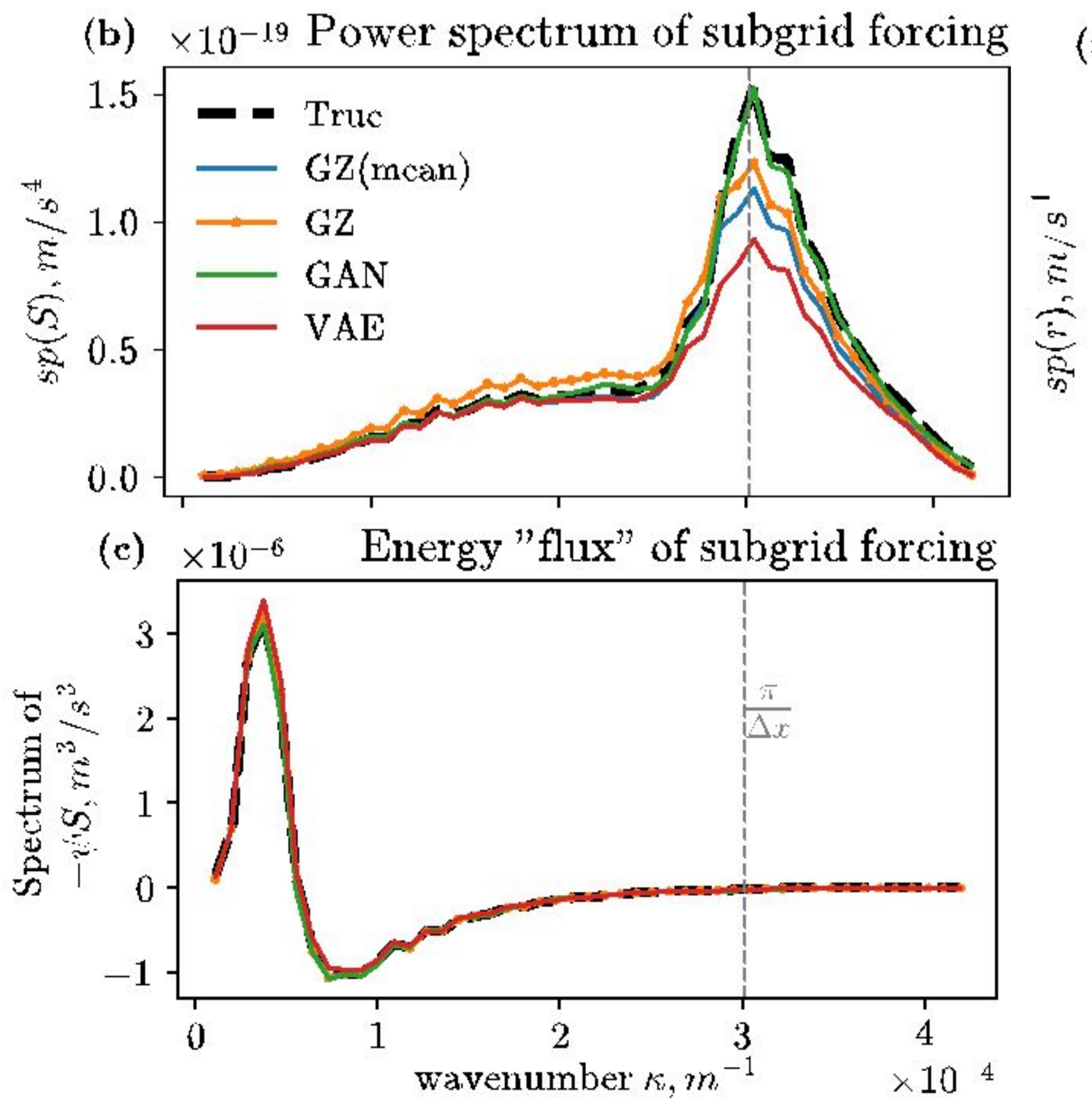


Dataset



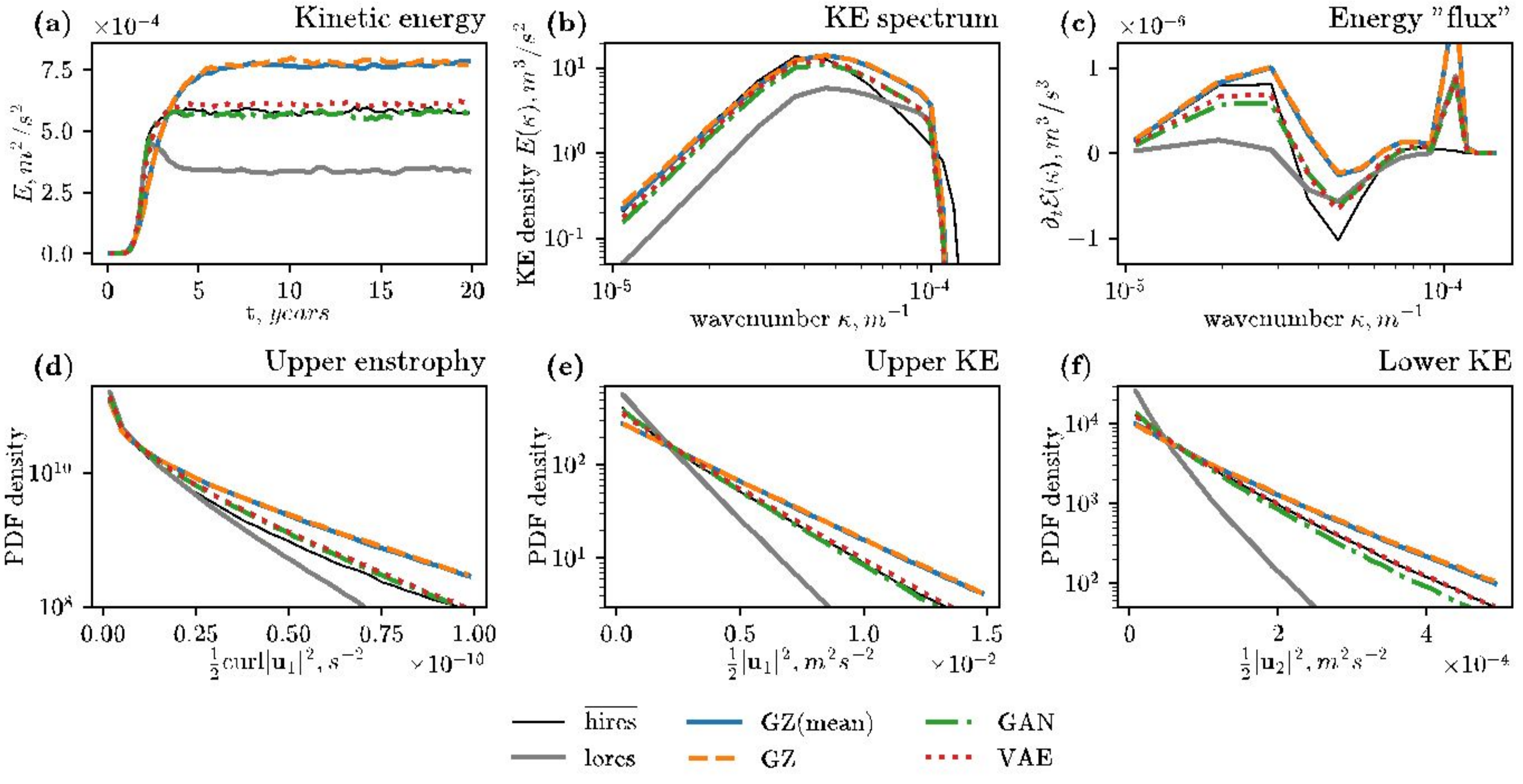
Power spectrum of S, m/s^4
48^2 (Lower)
64^2 (Lower)
96^2 (Lower)
Sharp
Gaussian
Upper
Lower
10^{-18}
10^{-21}
10^{-24}
10^{-27}
10^{-30}
10^{-5}
10^{-4}
wavenumber κ, m^{-1}
$\times 10^{-12}$
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
Subgrid forcing S, $1/s^2$

Physical properties of stochastic models



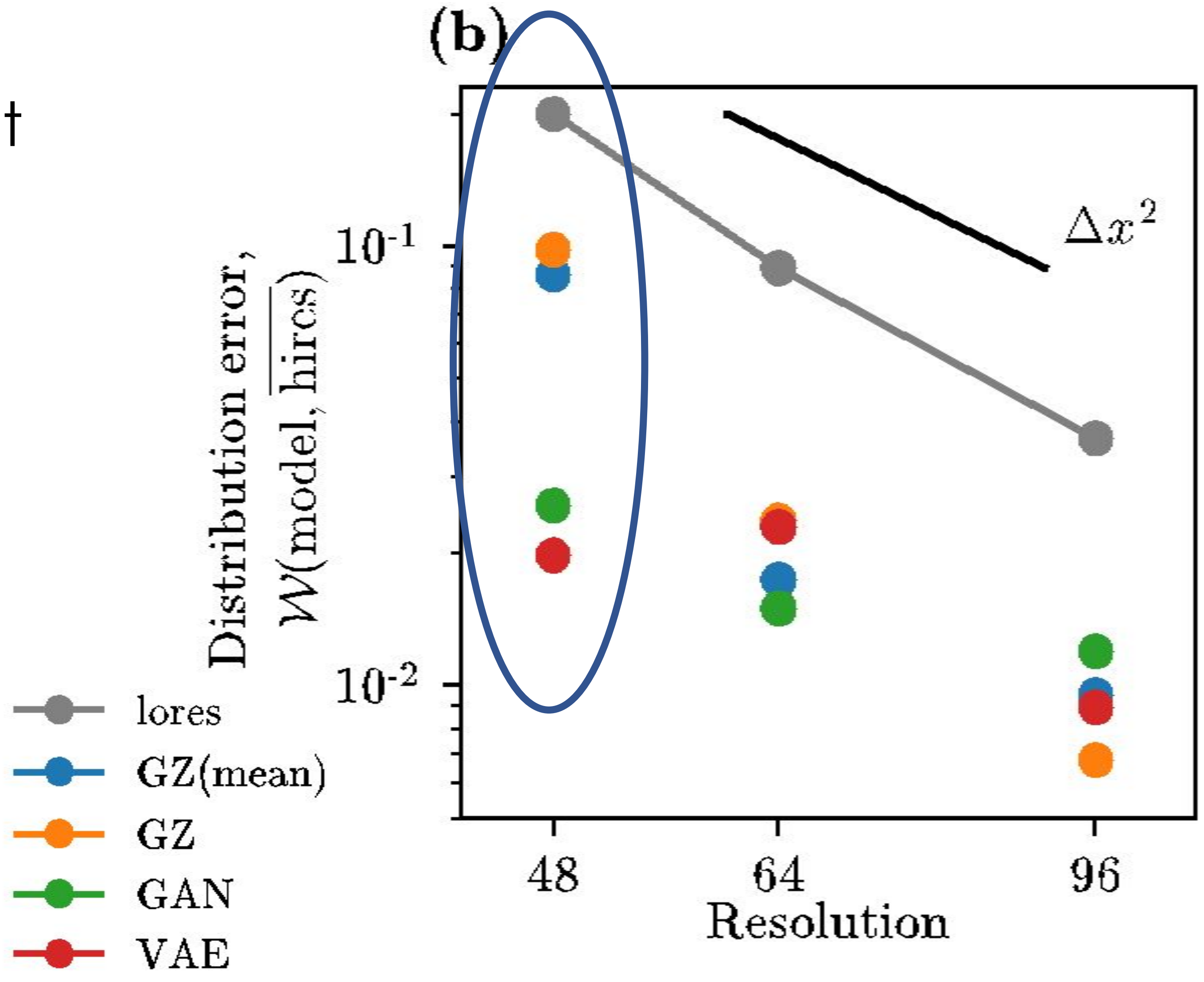
- Stochastic backscatter near the grid scale
- Large-scale energy backscatter

Online simulation

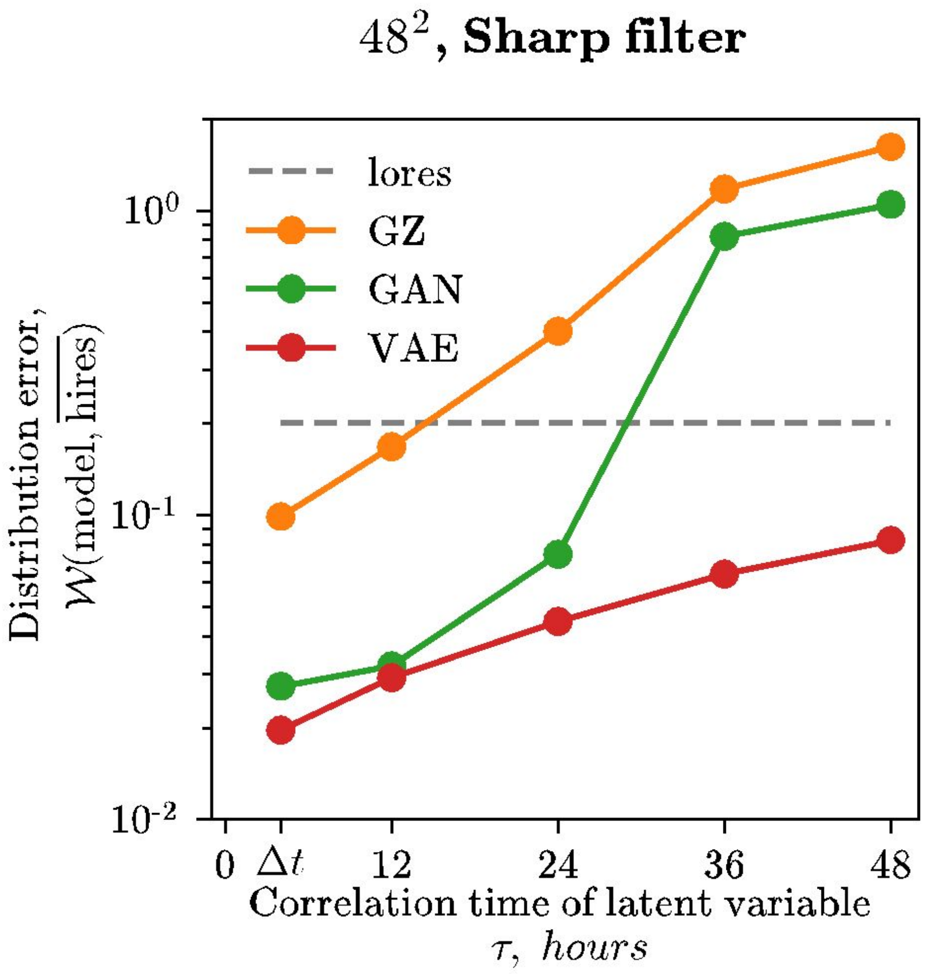


Online distributional error

- Improvement is evident for the coarsest resolution
- At finer resolutions (64,96) generative models perform as good as baselines

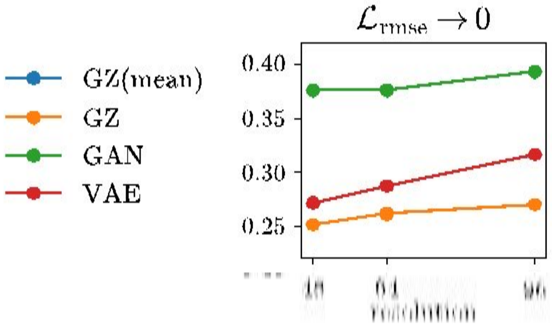


Autoregressive process (AR1) for samping of latent variable



Offline analysis

Deterministic prediction
$$\|E(\tilde{S}|\overline{q}) - S\|$$



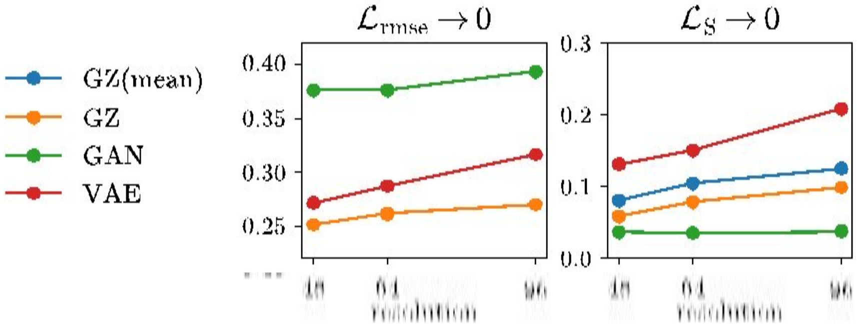
Offline analysis

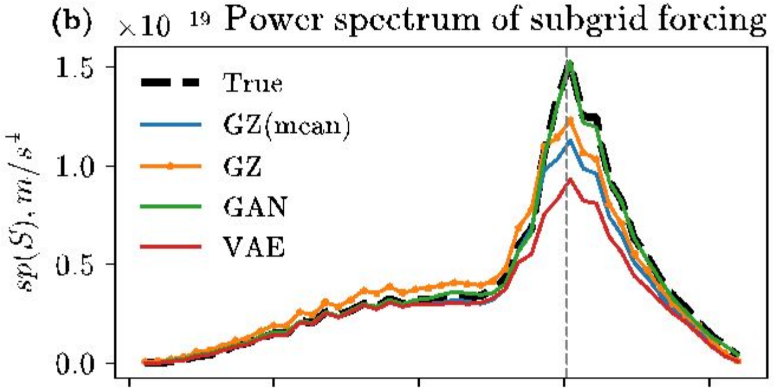
Deterministic prediction

$$\|E(\tilde{S}|\overline{q}) - S\|$$

Quality of individual samples

$$\|sp(S) - sp(\tilde{S})\|$$



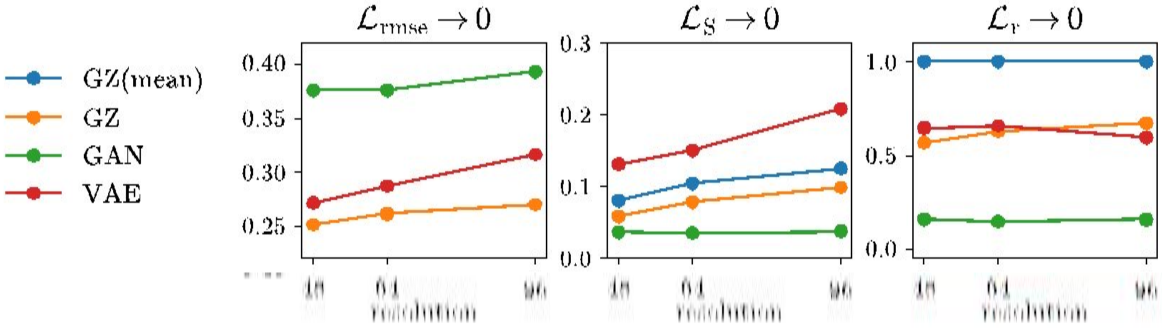


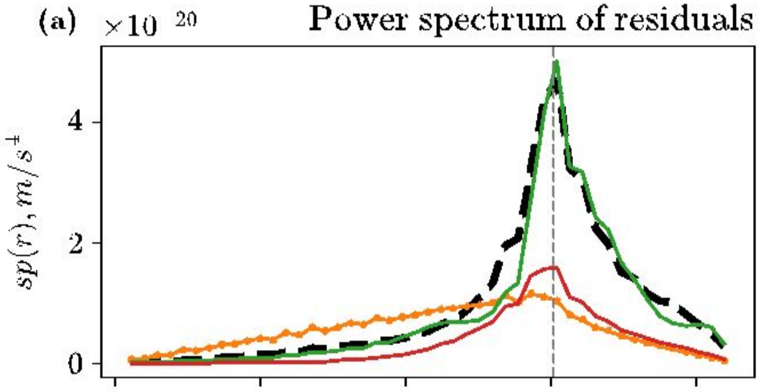
Offline analysis

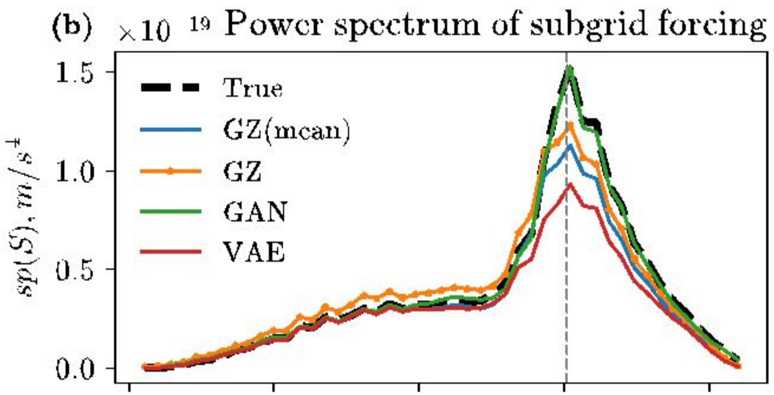
Deterministic prediction
$\|E(\tilde{S}|\overline{q}) - S\|$

Quality of individual samples
$\|sp(S) - sp(\tilde{S})\|$

Sprectrum of residuals
$\|sp(r) - sp(\tilde{r})\|$







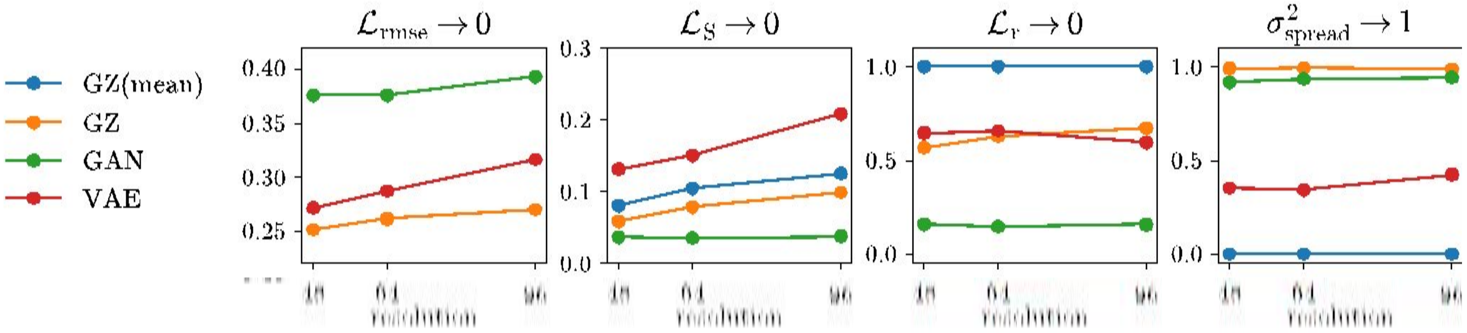
Offline analysis

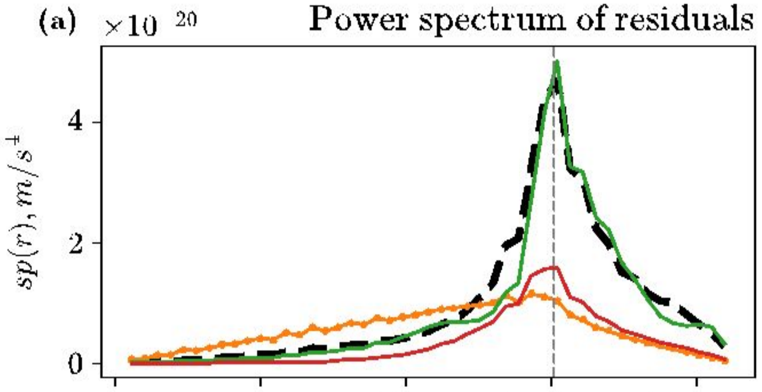
Deterministic prediction
$\|E(\tilde{S}|\overline{q}) - S\|$

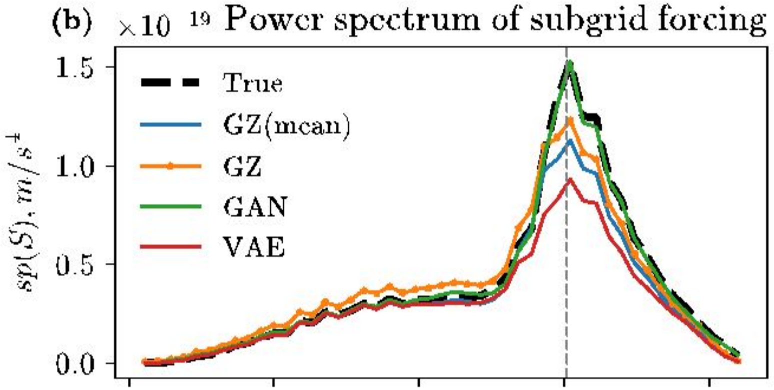
Quality of individual samples
$\|sp(S) - sp(\tilde{S})\|$

Sprectrum of residuals
$\|sp(r) - sp(\tilde{r})\|$

Spread
$\langle\tilde{r}^2\rangle / \langle r^2\rangle$







Offline analysis

Deterministic prediction
$$\|E(\tilde{S}|\overline{q}) - S\|$$

Quality of individual samples
$$\|sp(S) - sp(\tilde{S})\|$$

Sprectrum of residuals
$$\|sp(r) - sp(\tilde{r})\|$$

Spread
$$\langle \tilde{r}^2 \rangle / \langle r^2 \rangle$$