Using GANs as multivariate weather generators

Clément Brochet [1], Laure Raynaud [2], Matthieu Plu [2],
Nicolas Thome [3], Clément Rambour [3]

[1]CNRM/ENPC [2]CNRM [3]CNAM

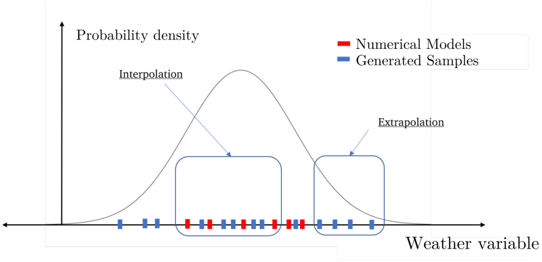
November 16th, 2022

1. Weather generators
2. Dataset and problem
3. Some results and metrics
4. Discussion

Weather generators

Main purpose : enrich ensemble approaches

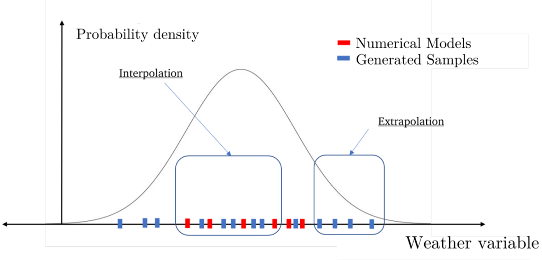
- Based on existing ensemble members...
- ... Generating new samples
- Ideas for ensemble forecast and DA



Weather generators

Main purpose : enrich ensemble approaches

- Based on existing ensemble members...
- ... Generating new samples
- Ideas for ensemble forecast and DA



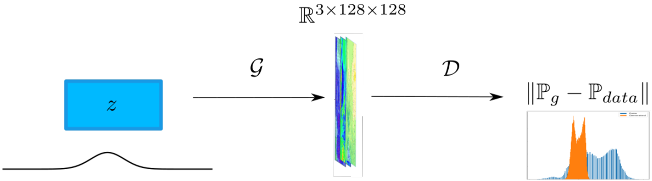
Weather generators

- Having ability to sample complex distributions
- In high-dimensional spaces

GANs as distribution learners

Adversarial training :

- "Optimal transport" of distributions (WGAN, [2])
- D acts as distribution distance estimator
- Allows working in high-dimensional spaces (maps)



High-resolution, Multivariate emulation with GANs

Need for proof-of-concept

1. Can we use GANs to sample the distribution of high-res, multivariate model outputs ?
2. How can we evaluate the quality of GANs samples ?

- Relatively few literature on high-resolution generation of atmospheric states ([10])
- We inspire from [3] generating global climate states

An AROME-EPS dataset

About AROME - EPS

- Convection-scale, NH-Model
- 1.3km grid-size resolution since 2022
- Dataset used :
  - 17 months of forecasts (1 per day)
  - 16 members for each forecast
  - Lead times up to 51h

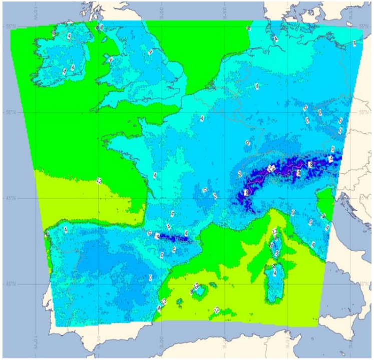


Figure 1: The AROME domain ([4])

An AROME-EPS dataset

About AROME - EPS

- Convection-scale, NH-Model
- 1.3km grid-size resolution since 2022
- Dataset used :
  - 17 months of forecasts (1 per day)
  - 16 members for each forecast
  - Lead times up to 51h

We want our GAN to sample :

$$X_{GAN} = G(Z) \sim \mathbb{P}_{data} = \mathcal{U}(\mathcal{D})$$

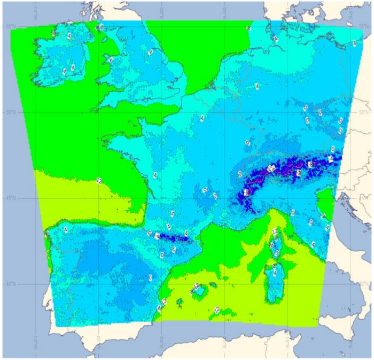
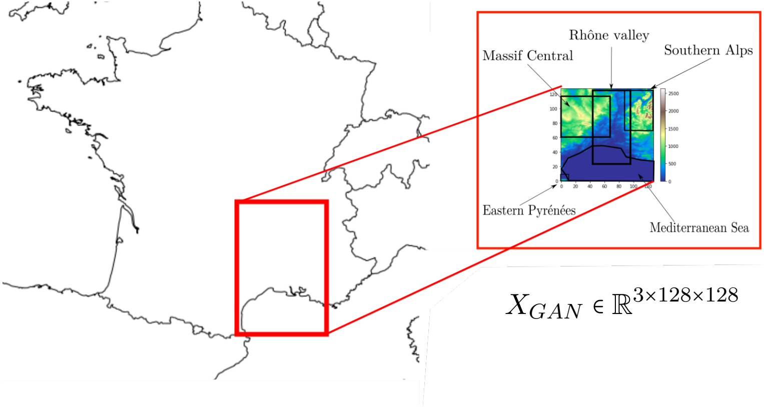


Figure 2: The AROME domain ([4])

Selected data



$$X_{GAN} \in \mathbb{R}^{3\times128\times128}$$

Figure 3: Selected subdomain and its geographical features

Learned and generated fields

- 10-m horizontal wind field : (u, v)
- 2-m temperature : t_{2m}

Joint generation !

GAN architecture and training

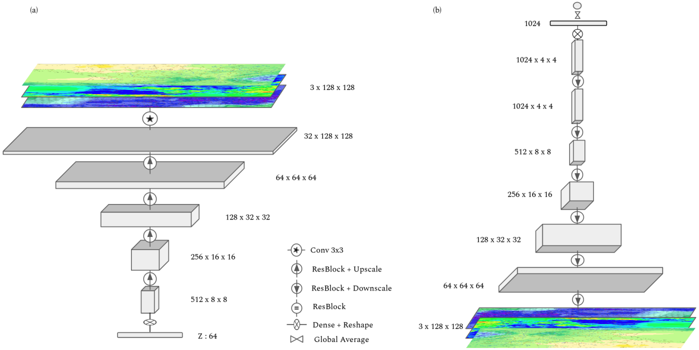


Figure 4: Generator (a) and Discriminator (b). From [8].

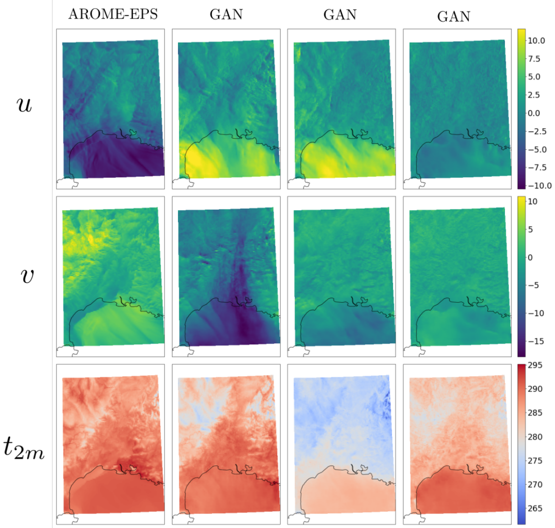
Hinge Loss :

$$\mathcal{L}_D = ReLU\,(1 - D(X)) + ReLU\,(1 + D(G(Z)))$$

Wasserstein GAN formulation

- Hinge loss is ok ([7])
- Spectral Normalization ([8])
- On **both G and D**

Results : samples



Consistency checks

Can the GAN reproduce simple statistical properties ?

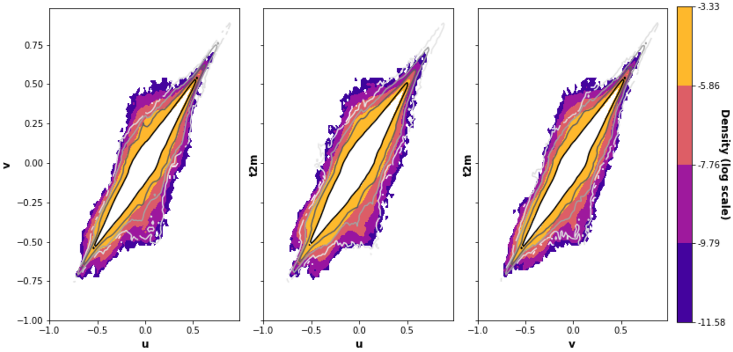


Figure 6: Bivariate histograms for GAN (contours) and AROME-EPS (heatmap)

Consistency checks, 2 : spectra

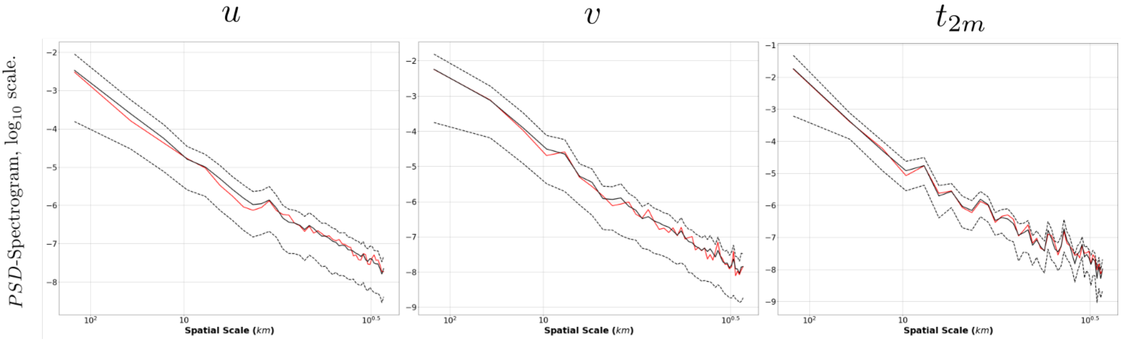


Figure 7: Black : AROME-EPS mean spectrum, Dashed : AROME-EPS Q10-Q90, Red : GAN

Fit of spectra is excellent

Mean Power Spectral Density (PSD) error of order of 1dB.

Evaluation metrics, 1 : 1D-Wasserstein distances

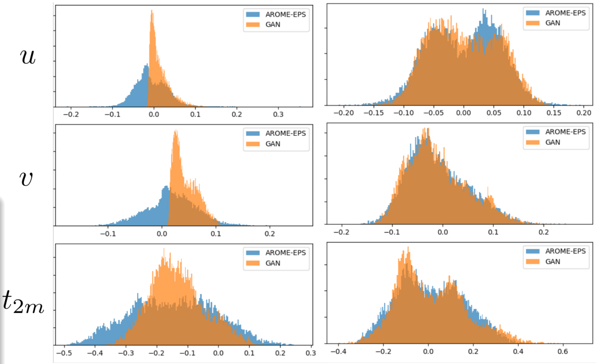
Effect on pixels :

Similar to [3].

$$W_1(\mathbb{P},\mathbb{Q}) = \int_0^1 |F_{\mathbb{P}}^{-1}(t) - F_{\mathbb{Q}}^{-1}(t)| dt$$

Usage

- Easy to compute in 1D (not feasible in higher dimensions)
- Pixel-wise average of 1D-W_1 : a measure of distributional proximity



"Bad" gridpoint
W1 = 31.9e-3

"Good" gridpoint
W1 = 5.9e-3

Average W1 = 14.6e-3

Evaluation metrics, 2 : multi-scale Wasserstein Distances

From [6] and [9] : (Sliced Wasserstein Distance - SWD)

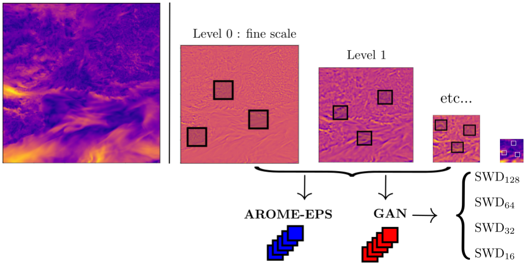


Figure 8: Multi-scale Wasserstein Distance

Metric	SWD_{128}	SWD_{64}	SWD_{32}	SWD_{16}
GAN	5.7	7.3	12	39
AROME*	1.5	1.5	1.6	4.6

*: Using a bootstrap technique

Pattern diversity

- Fine scales (SWD_{128}) better than Large scales (SWD_{16})

Evaluation metrics, 3 : correlation lengths

Fully local metric, following [11].

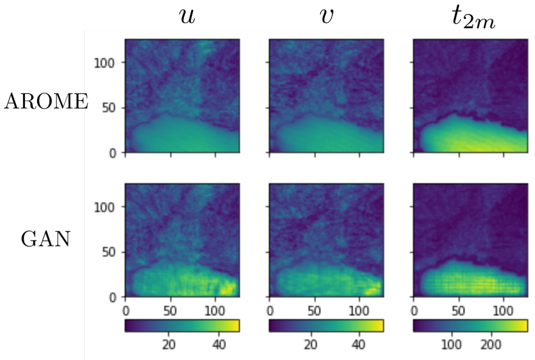


Figure 9: Length scales (color scale in km)

$$\rho(x, \delta x) = \langle X(x) X(x + \delta x) \rangle$$

$$L_{corr}(x) = \sqrt{\frac{1}{-\nabla^2 \rho(x, 0)}}$$

Quality of correlations

- Correct over land
- Degraded over sea
- Border effects
- Checkerboard patterns

Results from a parameter sweep

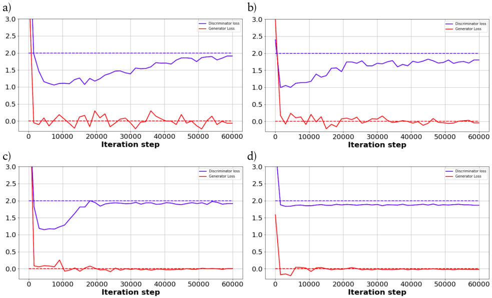


Figure 10: Training curves with increasing batch size, 32 (a) 64 (b) 128 (c) 256 (d). Blue : Discriminator loss; Red : Generator loss

Table 1: Average for 3 runs (lower is better)

Batch	W_1	PSD-error	SWD(avg)	MAE(L_{corr})
32	**12**	**0.9**	**18**	**5.3**
64	13	0.9	21	5.9
128	17	1.5	37	8.1
256	12	2.1	20	9.8
512	21	10.5	55	19.1

Increasing batch size

- Small effect on W_1
- Degrades spectra/SWD/L_{corr} severely
- Freezes training

What are we learning ?

1. Correlation maps show **border effects** and **worse scores** on sea
2. We are able to learn temperature correlation with altitude
3. ... And we learn on a fixed domain

What are we learning ?

1. Correlation maps show **border effects** and **worse scores** on sea
2. We are able to learn temperature correlation with altitude
3. ... And we learn on a fixed domain

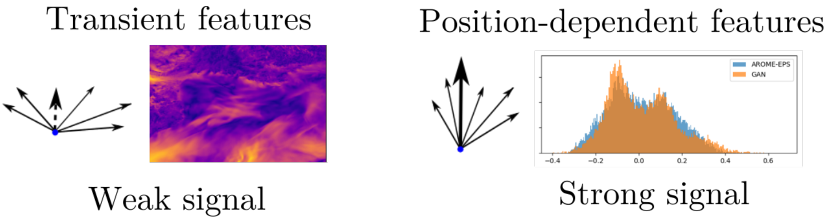
Absolute gridpoint position learning

1. Cf. literature ([1], [13], [12])
2. **Padding** in convolutions + **surface fields**

Position learning : consequences

Absolute gridpoint position learning

1. Learning signal is mainly position-related
2. Transient features are smoothed out



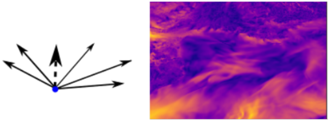
Position learning : consequences

Absolute gridpoint position learning

1. Learning signal is mainly position-related
2. Transient features are smoothed out

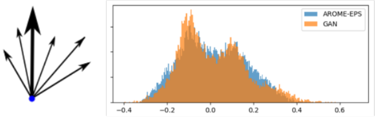
1. Increasing batch size causes training curves to **plateau**
2. Increasing batch size **degrades spatial correlations** but not grid-point error

Transient features



Weak signal

Position-dependent features



Strong signal

Conclusions

1. GANs **can generate high quality outputs** of AROME-EPS model.
2. **Multiplying meaningful metrics** deepens diagnosis
3. **Position learning** acts as strong inductive bias

PERSPECTIVES :

1. **Conditioning is the next step** : Emulate ensemble diversity with situation-aware GANs
2. **Add precipitations** : a non-gaussian irregular field with rare events.
3. **Towards SotA architectures** : ProGAN, StyleGAN-2 are promising candidates to enchance quality

Thank you !

Backup

References I

[1] Bilal Alsallakh et al. “Mind the Pad – {CNN}s Can Develop Blind Spots”. In: *International Conference on Learning Representations*. 2021. URL: https://openreview.net/forum?id=m1CD7tPubNy.

[2] Martin Arjovsky, Soumith Chintala, and Léon Bottou. “Wasserstein Generative Adversarial Networks”. In: *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. Sydney, NSW, Australia: JMLR.org, 2017, pp. 214–223.

[3] C. Besombes et al. “Producing realistic climate data with generative adversarial networks”. In: *Nonlinear Processes in Geophysics* 28.3 (2021), pp. 347–370. DOI: 10.5194/npg-28-347-2021. URL: https://npg.copernicus.org/articles/28/347/2021/.

References II

[4] François Bouttier et al. "Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX". In: *Quarterly Journal of the Royal Meteorological Society* 142.S1 (2016), pp. 390–403. DOI: https://doi.org/10.1002/qj.2622. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2622. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2622.

[5] Sihao Cheng and Brice Ménard. *How to quantify fields or textures? A guide to the scattering transform*. 2021. DOI: 10.48550/ARXIV.2112.01288. URL: https://arxiv.org/abs/2112.01288.

[6] Tero Karras et al. *Progressive Growing of GANs for Improved Quality, Stability, and Variation*. 2018. arXiv: 1710.10196 [cs.NE].

[7] Jae Hyun Lim and Jong Chul Ye. *Geometric GAN*. 2017. arXiv: 1705.02894 [stat.ML].

References III

[8] Takeru Miyato et al. "Spectral Normalization for Generative Adversarial Networks". In: *CoRR* abs/1802.05957 (2018). arXiv: 1802.05957. URL: http://arxiv.org/abs/1802.05957.

[9] Julien Rabin et al. "Wasserstein Barycenter and its Application to Texture Mixing". In: *SSVM'11*. Israel: Springer, 2011, pp. 435–446. URL: https://hal.archives-ouvertes.fr/hal-00476064.

[10] S. Ravuri et al. "Skilful precipitation nowcasting using deep generative models of radar". In: *Nature* 597, 672–677 (2021) (2021), pp. 672–677. DOI: 10.1038/s41586-021-03854-z. URL: https://www.nature.com/articles/s41586-021-03854-z.

References IV

[11] A. T. Weaver and I. Mirouze. "On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 139.670 (2013), pp. 242–260. DOI: https://doi.org/10.1002/qj.1955. eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.1955. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.1955.

[12] Rui Xu et al. *Positional Encoding as Spatial Inductive Bias in GANs*. 2020. DOI: 10.48550/ARXIV.2012.05217. URL: https://arxiv.org/abs/2012.05217.

[13] Richard Zhang. "Making Convolutional Networks Shift-Invariant Again". In: *ICML*. 2019.

And now for an open question...

Should GANs learn absolute position in weather datasets ?

PROS :

1. It seems useful and natural on aligned learning
2. It definitely helps converging at first

CONS :

1. It probably hampers convergence after a while
2. It fails grasping transient structures

Does the GAN copy Dataset ?

"Raw MSE test" on normalized data.

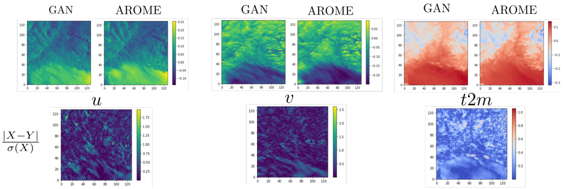


Figure 11: Closest GAN vs AROME samples RMSE-wise

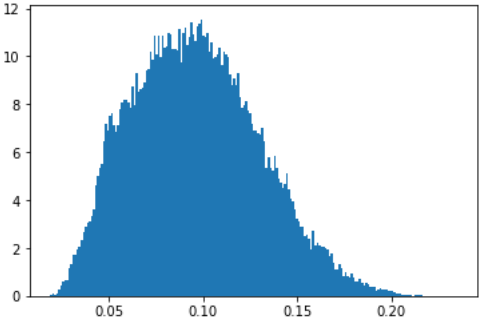


Figure 12: RMSE distribution to AROME samples

Evaluation metrics, 1 : 1D- Wasserstein distances

"Wasserstein Maps" during training (log scale)

Similar to Besombes et al., 2021.

$$W_1(\mathbb{P},\mathbb{Q}) = \int_0^1 |F_{\mathbb{P}}^{-1}(t) - F_{\mathbb{Q}}^{-1}(t)|dt$$

Takeaways

- GAN optimizes Wasserstein distance
- Dataset gridpoint variance acts as a **strong error signal**

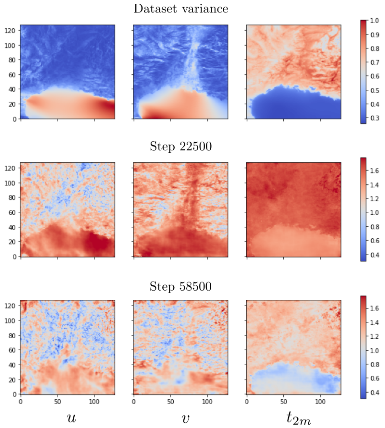


Figure 13: Left : u, Center : v, Right : t_{2m}.

Multivariate experiments : "ablation" study

Configuration	Baseline := (u, v, t_{2m})	t_{2m}	(u, v)

Table 2: Sets of variables (fine-tuned parameters for each)

How does multivariate generation impact performance ?

- More variables add useful correlations
- Less variables simplify the dataset

Multivariate experiments : results

Config	W_1	Fine SWD	Large SWD	Spectra	L_{corr}
t_{2m}	$\uparrow$	$\sim$	$\uparrow$	$\uparrow$	$\downarrow$
t_{2m},orog	$\uparrow$	$\downarrow$	$\uparrow$	$\downarrow$	$\downarrow$
(u, v)	$\sim$	$\sim$	$\downarrow\downarrow$	$\uparrow$	$\uparrow\uparrow$

Table 3: Effects on scores

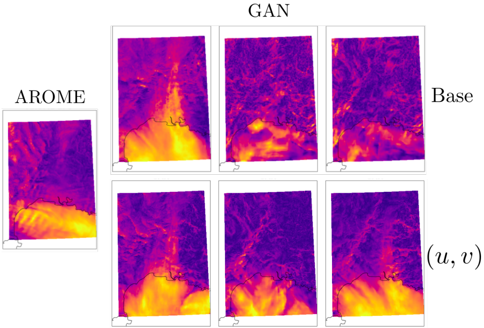
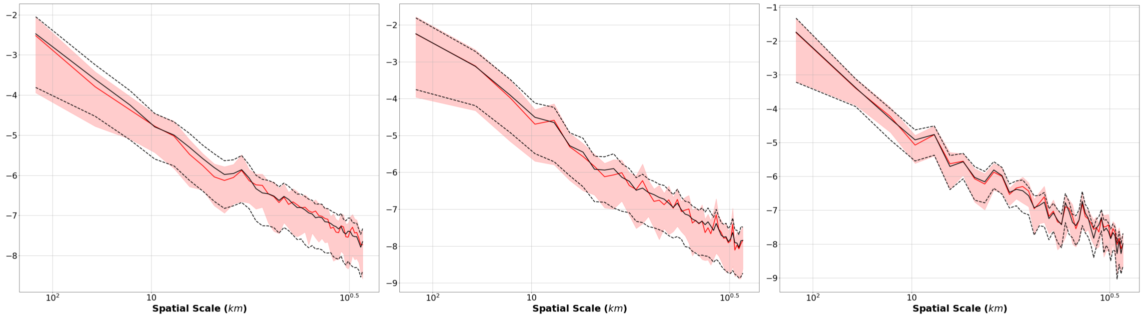


Figure 14: Effect on wind module.

Spectral distributions



Spatial consistency : scattering analysis

Wavelet basis decomposition :

$$\psi_\lambda(\mathbf{x}) \propto \exp\left(-\frac{\lambda^{-2}\mathbf{x}^2}{2\sigma_x}\right)\left(e^{i\frac{x}{\lambda}} - \beta\right)$$

Scattering coefficients :

$$S_1(\lambda_1) = \langle |X \star \psi_{\lambda_1}| \rangle\,,$$
$$S_2(\lambda_1,\lambda_2) = \langle ||X \star \psi_{\lambda_1}| \star \psi_{\lambda_2}| \rangle$$

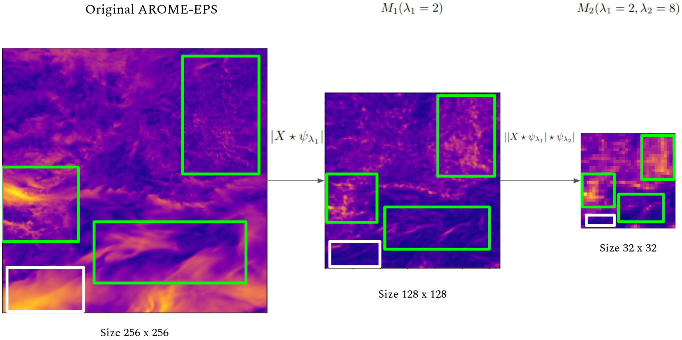
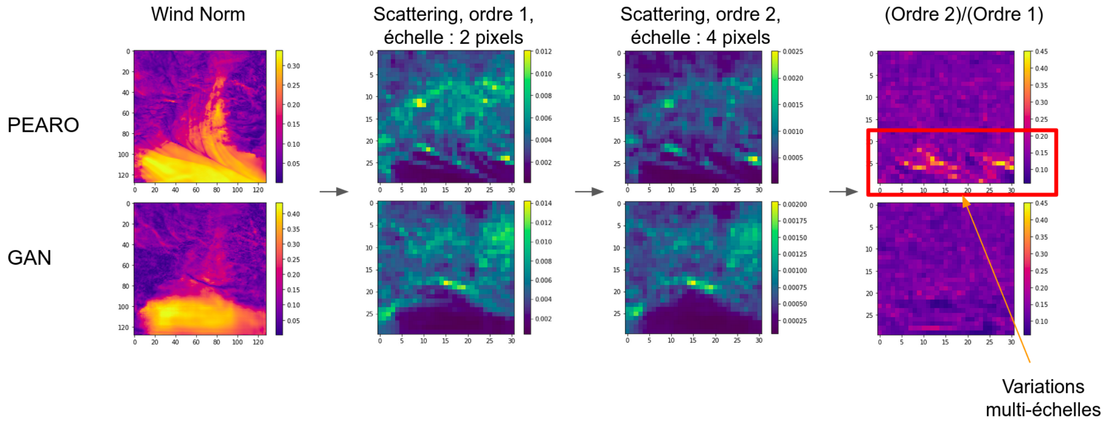


Figure 15: Scattering transform, order 2.

Offline analysis



Offline analysis

We set (cf. [5]) :

$$s_{21} = \left\langle \frac{S_2(\lambda_1, \lambda_2)}{S_1(\lambda_1)} \right\rangle_{\theta_1,\theta_2}$$

(Averaging on directions)

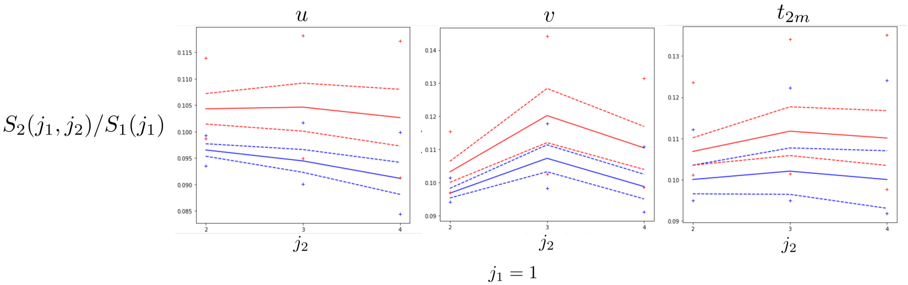


Figure 16: Comparing AROME (red) and GAN (blue) - dashed lines are Q10-Q90.

Scattering : scores

Table 4: Average for 3 runs (lower is better)

Batch	W_1	PSD-error	SWD(avg)	MAE(L_{corr})	Scattering
32	**12**	**0.9**	**18**	**5.3**	**3.9**
64	13	0.9	21	5.9	**3.9**
128	17	1.5	37	8.1	4.2
256	12	2.1	20	9.8	5.2
512	21	10.5	55	19.1	8.2

Scattering scores evolve consistently with others.

A word about "Bootstrap"

$$W_1(\mathbb{P},\mathbb{Q}) = \inf_{\gamma \in \Pi(\mathbb{P},\mathbb{Q})} \int \|x - y\| \gamma(x,y) dx dy$$

Why bother ?

- Need for a lower bound on distance estimates
- Finite sampling effects must be quantified

Process :

- Select two random Batches of size B from $\mathbb{P}_{data}$
- Compute an estimate $\hat{W}_1(\mathbb{P}_{1\sharp B}, \mathbb{P}_{2\sharp B})$
- Repeat N times with replacement and average over N

Parameters used : $B = 16384,\ N = 32$

Thanks for your attention

www.umr-cnrm.fr

www.meteofrance.com