WAVELET-BASED VERIFICATION OF DETERMINISTIC AND ENSEMBLE FORECASTS

2026-06-24 I SEBASTIAN BUSCHOW

Member of the Helmholtz Association

Why wavelets?

Wavelet scores and diagnostics

Generalization to ensembles

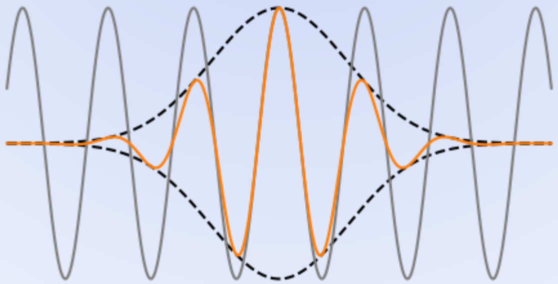
Summary & Outlook

continuous wavelet transform

Mother wavelet: localized waveform, for example Morlet wavelet

$$\psi(t) \propto e^{i\sigma t} \times e^{-0.5t^2}$$

Generate a set of daughters via
- Shift $\psi_u(t) = \psi(t-u)$
- Scaling $\psi_s(t) = \frac{1}{\sqrt{s}}\psi(t/s)$
- Rotation (*in higher dims*)



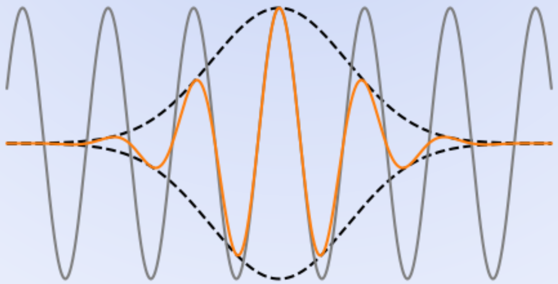
continuous wavelet transform

Mother wavelet: localized waveform, for example Morlet wavelet

$$\psi(t) \propto e^{i\sigma t} \times e^{-0.5t^2}$$

Generate a set of daughters via

- Shift $\psi_u(t) = \psi(t-u)$
- Scaling $\psi_s(t) = \frac{1}{\sqrt{s}}\psi(t/s)$
- Rotation (*in higher dims*)

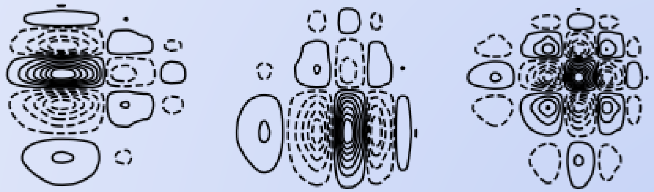


discrete wavelet transform

- Restrict scales to
  $$s = 2^0, 2^1, \ldots, 2^J$$
- Restrict shifts at scale s to
  $$u = s, 2 \cdot s, 3 \cdot s, \ldots$$

→ Orthogonal bases of $\mathbb{R}^N$
→ Fast algorithms via filterbanks

2D wavelets: alternating filters on rows and columns, resulting in three directional wavelets



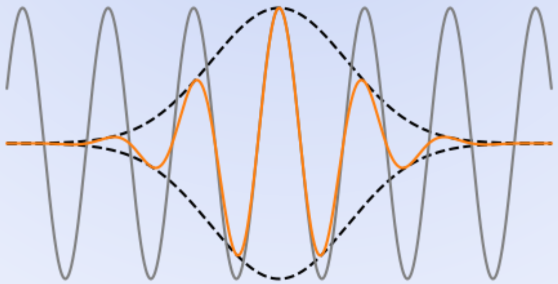
continuous wavelet transform

Mother wavelet: localized waveform, for example Morlet wavelet

$$\psi(t) \propto e^{i\sigma t} \times e^{-0.5t^2}$$

Generate a set of daughters via

- Shift $\psi_u(t) = \psi(t-u)$
- Scaling $\psi_s(t) = \frac{1}{\sqrt{s}}\psi(t/s)$
- Rotation (*in higher dims*)

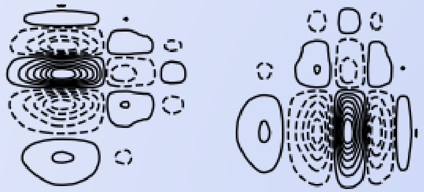


discrete wavelet transform

- Restrict scales to $$s = 2^0, 2^1, \ldots, 2^J$$
- Restrict shifts at scale s to $$u = s, 2 \cdot s, 3 \cdot s, \ldots$$

→ Orthogonal bases of $\mathbb{R}^N$
→ Fast algorithms via filterbanks

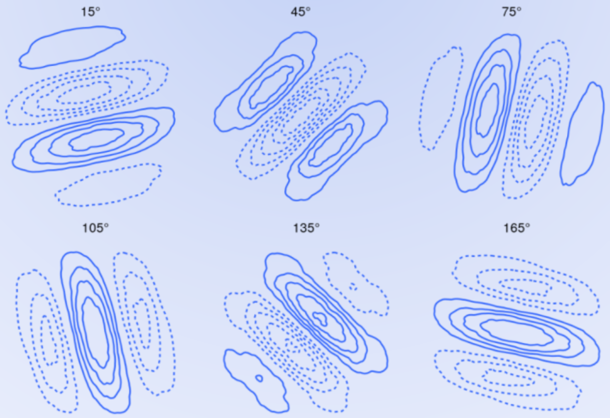
2D wavelets: alternating filters on rows and columns, resulting in three directional wavelets



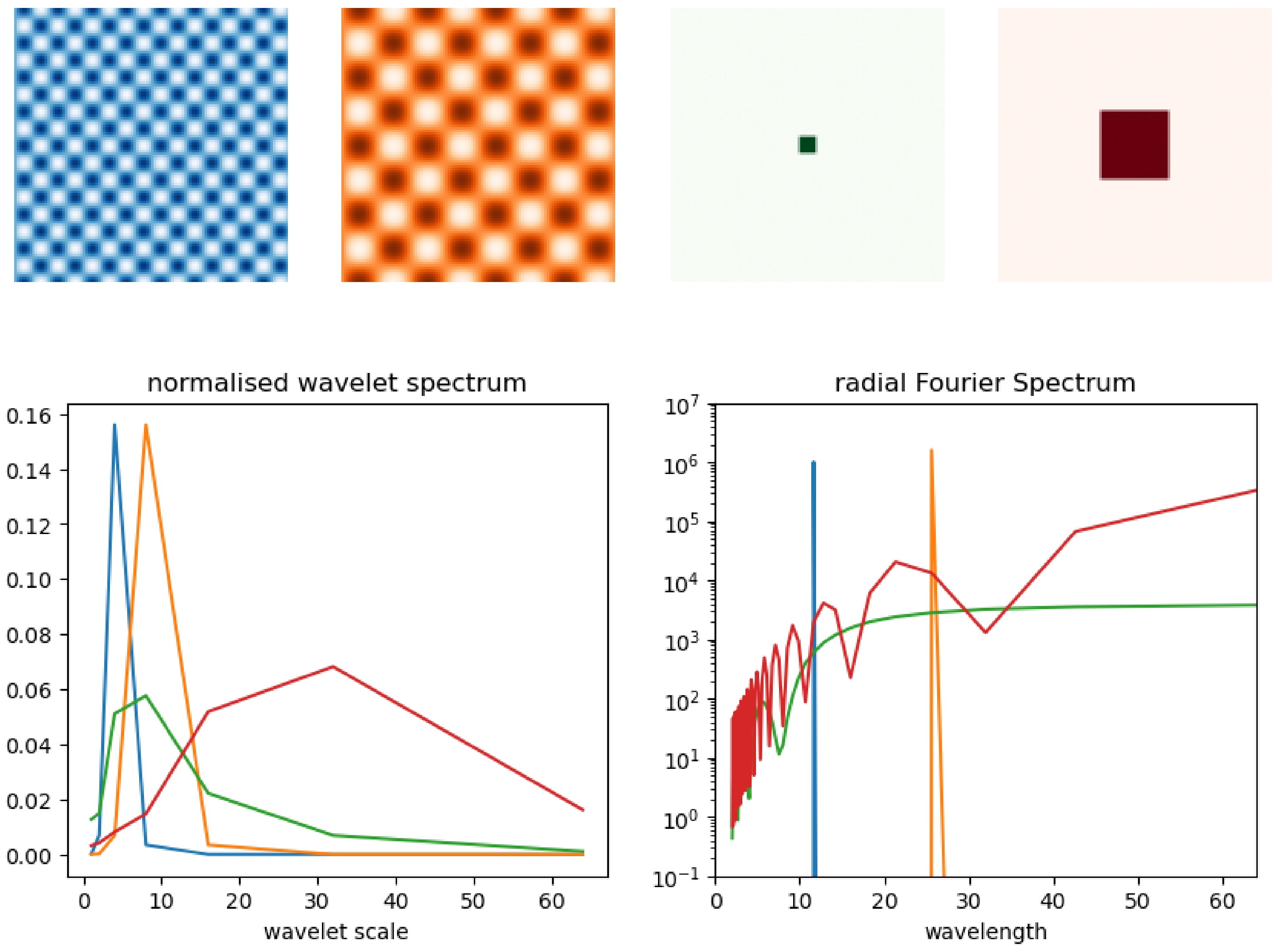
Dual-tree complex wavelets

Combine 4 classic DWTs into a single complex wavelet analysis with six directions

- Better directional selection
- More robust against shifts of the data
- Encode phase information

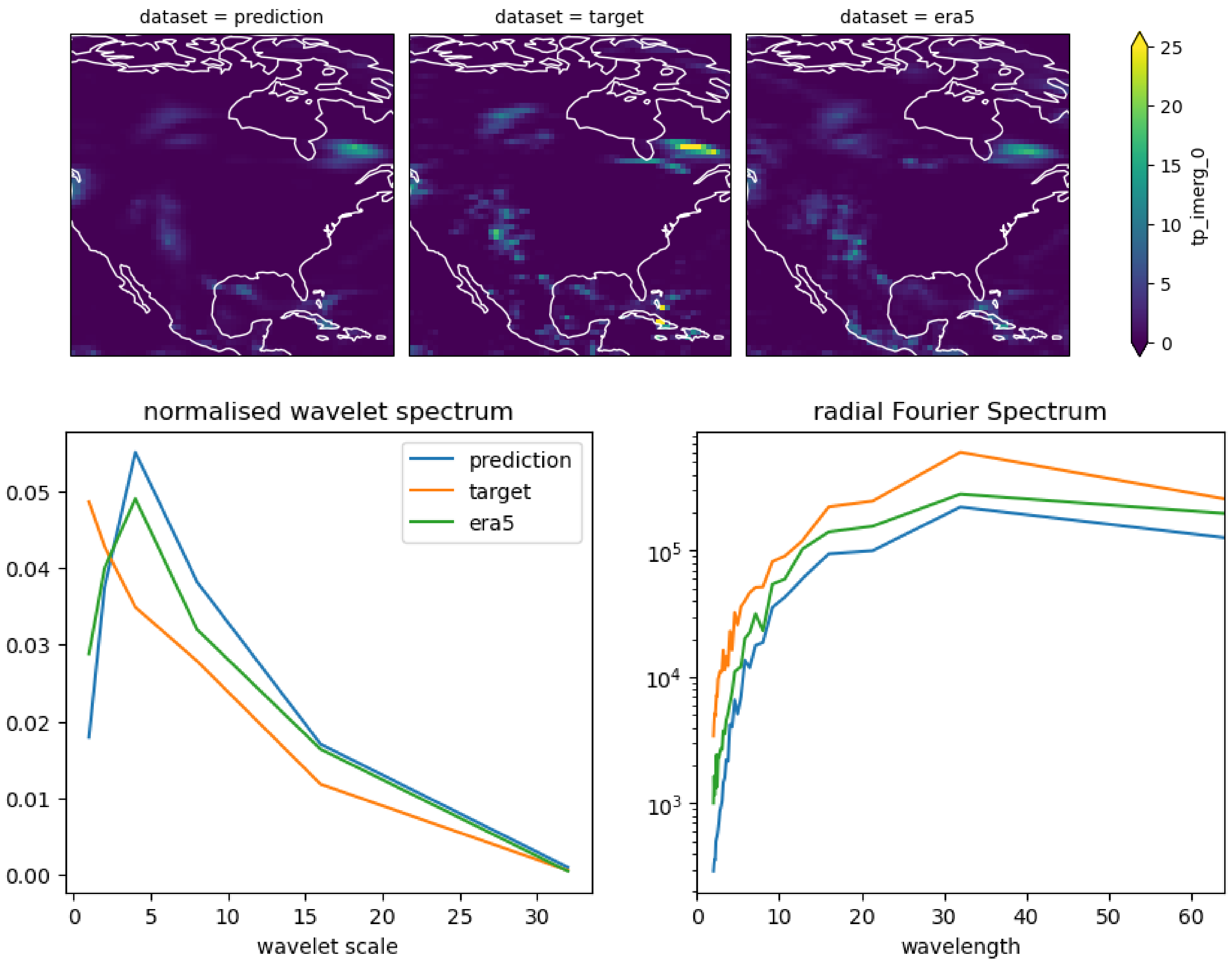


Comparison with Fourier



- Fourier transform isolates periodic components exactly, aliasing for the box signal
- Wavelets do a decent job for both, lower spectral resolution
- Wavelet scale does not equal wavelength

Comparison with Fourier



Example data: 6h precipitation sum predicted by our forecast model (WeatherGen), in the target (IMERG) and in ERA5

- Lacking small scale variance
- Wavelets see well-defined peak at small scales, Fourier broader

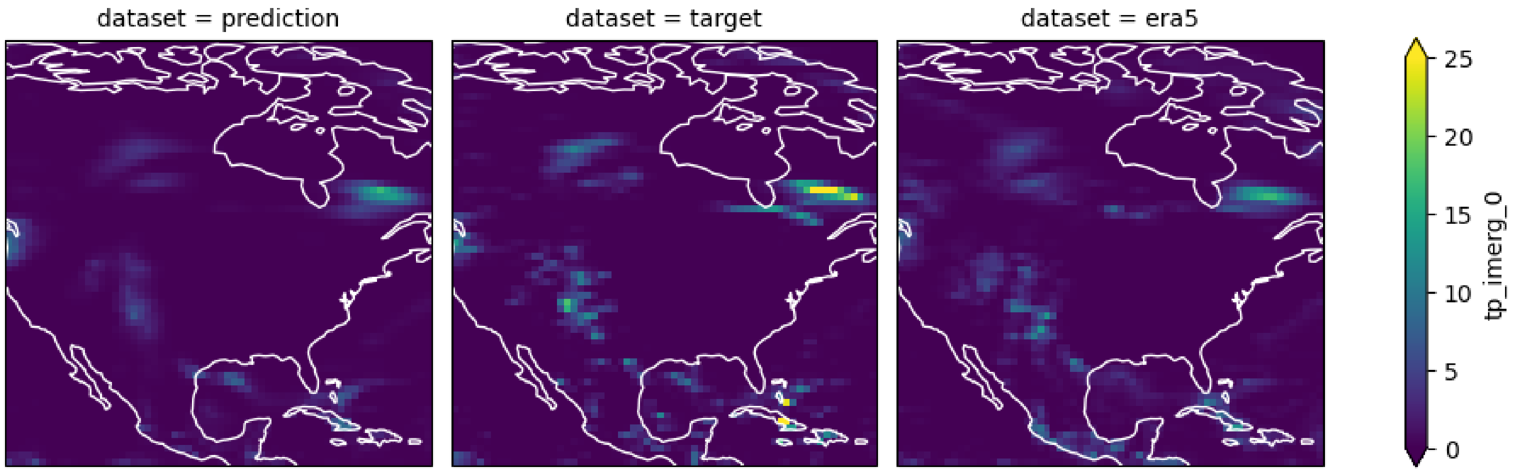
Why wavelets?

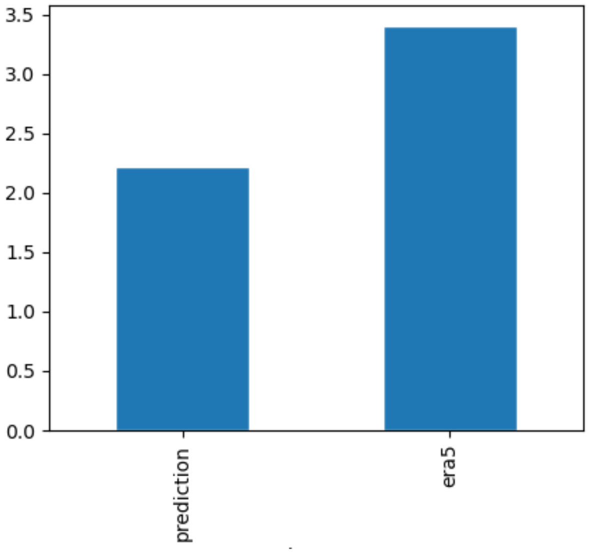
Wavelet scores and diagnostics

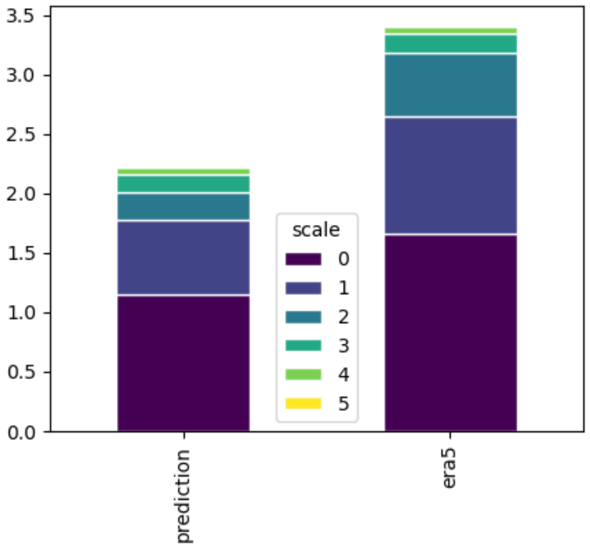
Generalizations to ensembles

Summary & Outlook

Error decomposition







Basic Idea: decompose forecast and target, compute for example MSE per scale

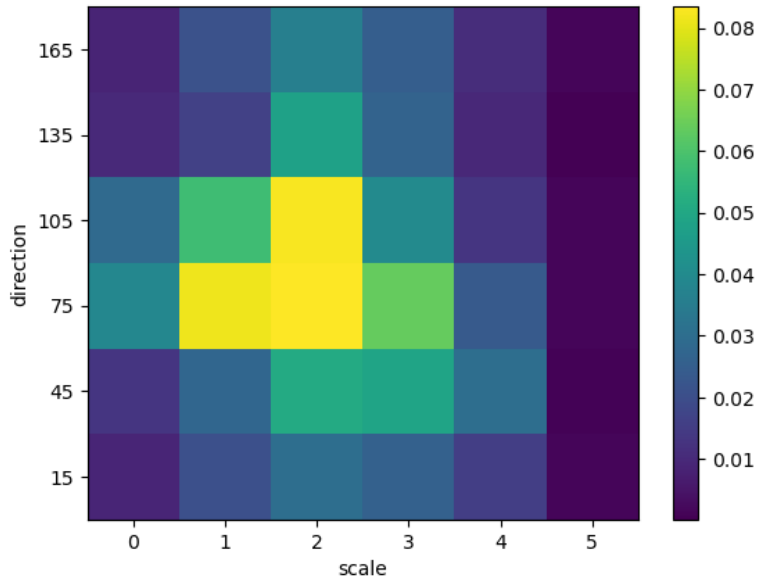
→ Parseval theorem guarantees

$$MSE = MSE_0 + MSE_1 + \ldots + MSE_J$$

Casati et al. (2004): intensity-scale scores.

Diagnosing Scale, Anisotropy, Direction (SAD)

Spatial mean spectrum of each field:



3D center of mass

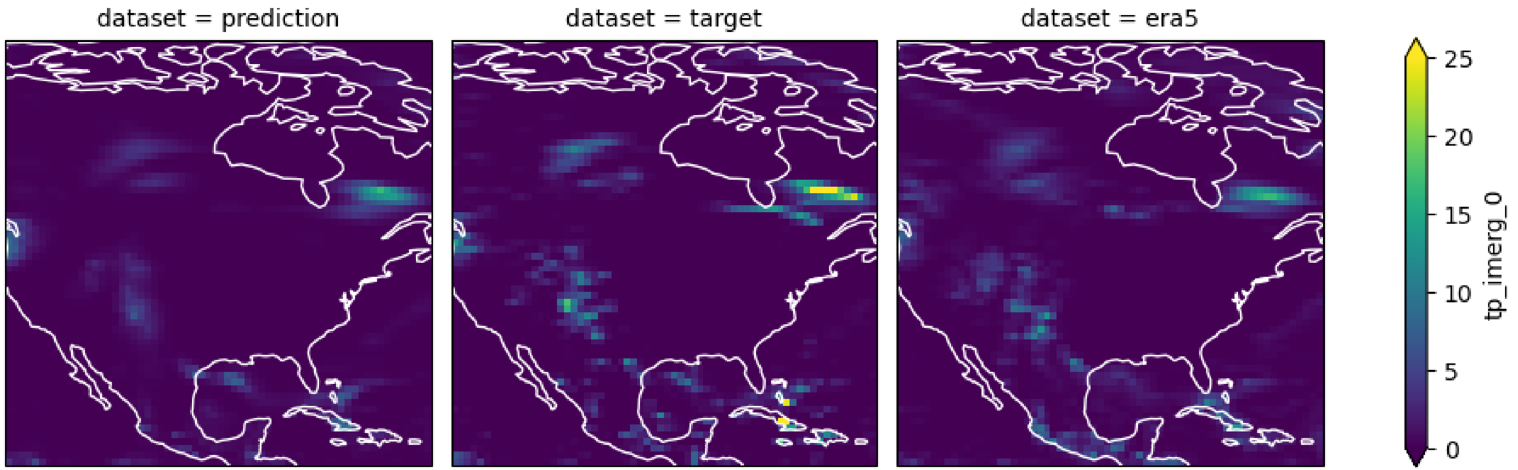
Scale $S \in [1, J]$
Anisotropy $A \in [0, \sqrt{2}]$
Direction $D \in [0°, 180°]$

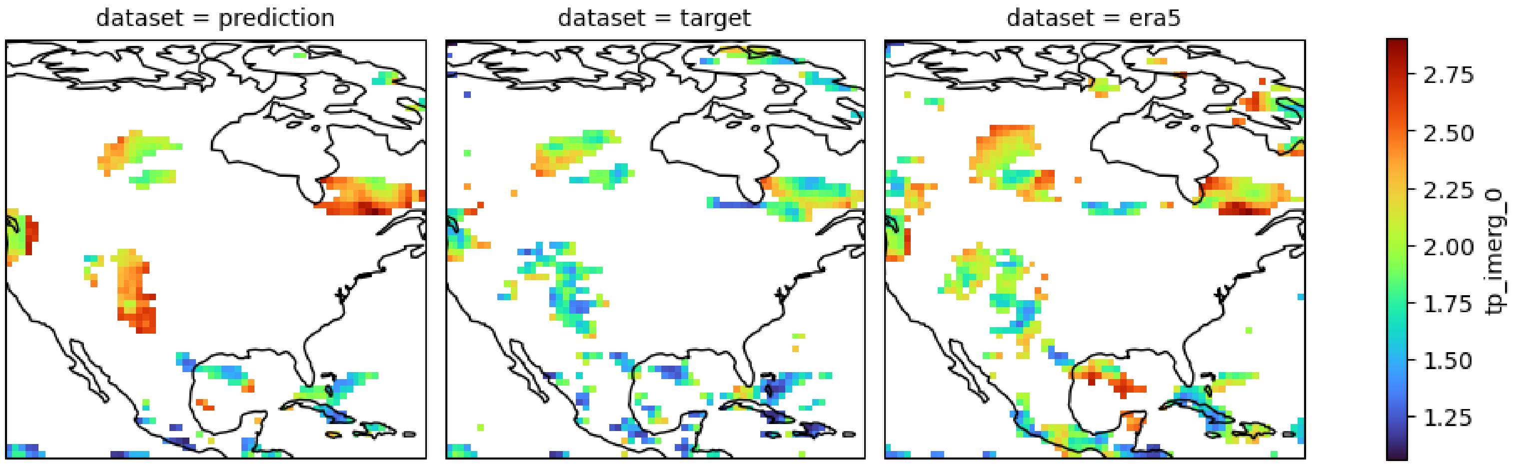
Compare forecast vs target

```python
from sadpy import wv_analysis, sad
an_tar = wv_analysis(realtest.sel(dataset="target"))
an_pr = wv_analysis(realtest.sel(dataset=["prediction","era5"]))
sad(an_pr, an_tar)
```

```
computing the wavelet transform
computing the wavelet transform
{'scale': array([0.53239862, 0.33222881]),
 'anisotropy': array([-0.07851568, -0.04675099]),
 'direction': array([1.90358978, 5.18958509])}
```

Localized structure properties





Idea: use only scales 2^j, but all shifts for all scales

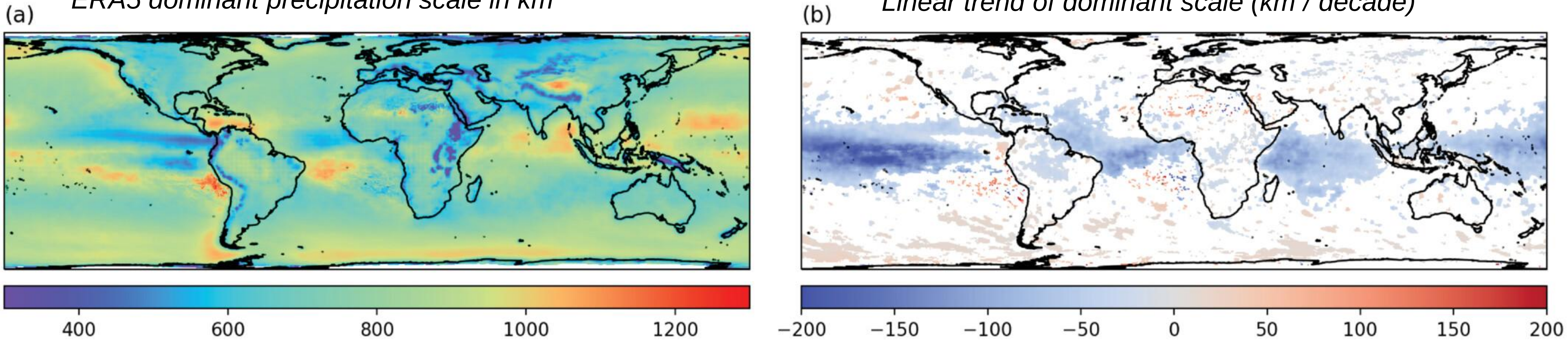
→ redundant discrete wavelet transform

→ Scale, anisotropy, direction at every grid-point

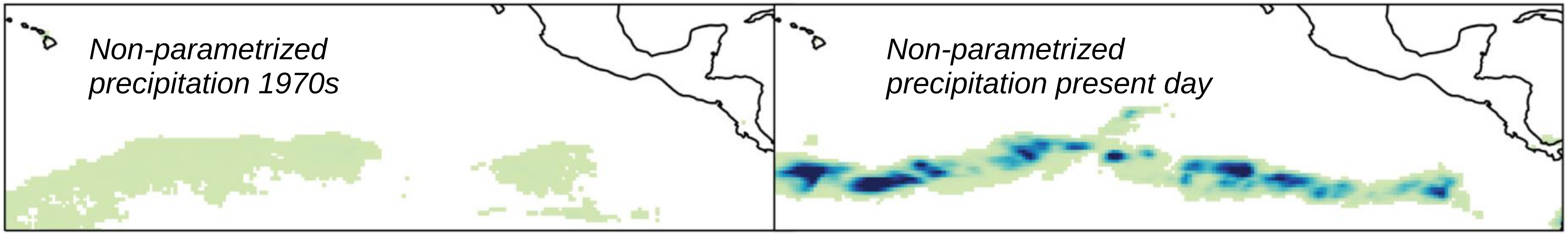
Artificial variance at large scales, can be corrected Nelson et al. (2018)

Localized structure properties

ERA5 dominant precipitation scale in km *Linear trend of dominant scale (km / decade)*



Benestad et al. (2022): ERA precipitation has shifted to smaller scales. Buschow (2024): phenomenon localised over tropical oceans, coincides with the emergence of grid-cell convection
→ signal is likely an artifact triggered by Satellite data assimilation.



Displacement errors from wavelet phase

Idea: for the Fourier transform of a shifted signal, we have

$$\mathcal{F}\{x(\cdot - \tau)\}(\omega) = \exp(-2\pi i \omega \tau) \cdot \mathcal{F}\{x\}(\omega)$$

Complex dual-tree wavelets have similar reaction to shifts → estimate displacement from wavelet phases

1. Take the dtcwt of prediction and target
2. For each location l, scale s and orientation d: compute phase distance $\Delta\Phi_{\mathrm{l,s,d}}$
3. Average over locations and directions, weighted by amplitude → per-scale phase error $\overline{\Delta\Phi_{\mathrm{s}}}$
4. (optional) compute $\max_s(2^s \times \overline{\Delta\Phi_{\mathrm{s}}})$ as the overall displacement estimate

Displacement errors from wavelet phase

Idea: for the Fourier transform of a shifted signal, we have

$$\mathcal{F}\{x(\cdot - \tau)\}(\omega) = \exp(-2\pi i \omega \tau) \cdot \mathcal{F}\{x\}(\omega)$$

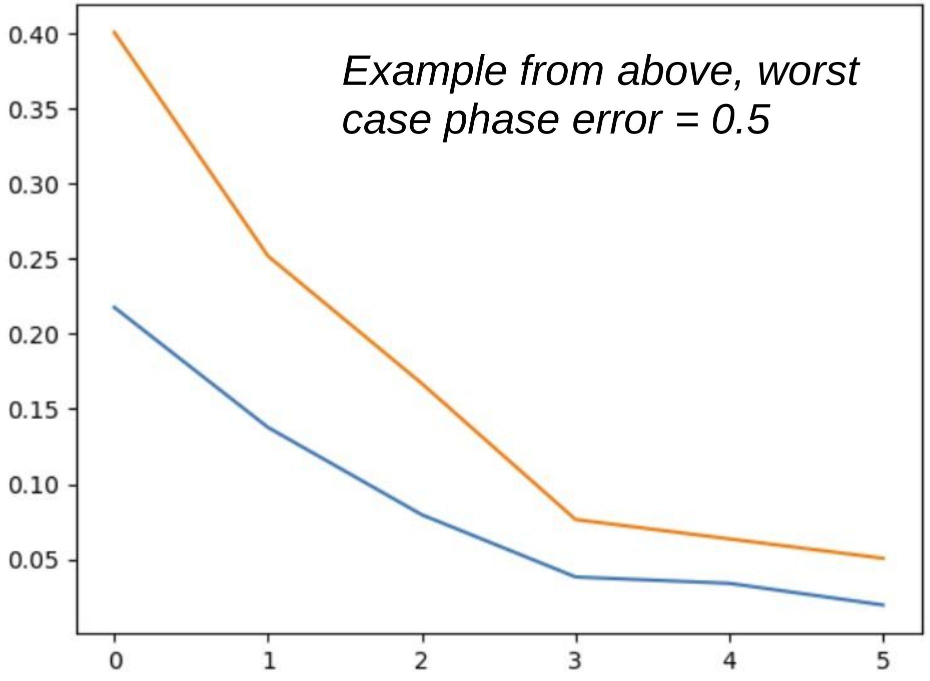
Complex dual-tree wavelets have similar reaction to shifts → estimate displacement from wavelet phases

1. Take the dtcwt of prediction and target
2. For each location l, scale s and orientation d: compute phase distance $\Delta\Phi_{\mathrm{l,s,d}}$
3. Average over locations and directions, weighted by amplitude → per-scale phase error $\overline{\Delta\Phi_{\mathrm{s}}}$
4. (optional) compute $\max_s(2^s \times \overline{\Delta\Phi_{\mathrm{s}}})$ as the overall displacement estimate

```
from sadpy import wv_analysis, sad_phase
an_tar = wv_analysis(realtest.sel(dataset="target"), Nx=120, Ny=120)
an_pr = wv_analysis(realtest.sel(dataset=["prediction","era5"]), Nx=120, Ny=120)
[ plt.plot(x) for x in sad_phase(an_pr, an_tar, return_maxdist=False) ]
```

```
computing the wavelet transform
computing the wavelet transform
```

```
[[<matplotlib.lines.Line2D at 0x7f6b0ce9c830>],
 [<matplotlib.lines.Line2D at 0x7f6b0ce9c980>]]
```



Why wavelets?

Wavelet scores and diagnostics

Generalizations to ensembles

Summary & Outlook

Transformation-based proper scores (Pic et al. 2025)

Proposition 1. *Let $\mathcal{F} \subset \mathcal{P}(\mathbb{R}^d)$ be a class of Borel probability measure on $\mathbb{R}^d$ and let $F \in \mathcal{F}$ be a forecast and $\boldsymbol{y} \in \mathbb{R}^d$ an observation. Let $T : \mathbb{R}^d \to \mathbb{R}^k$ be a transformation and let S be a scoring rule on $\mathbb{R}^k$ that is proper relative to $T(\mathcal{F}) = \{\mathcal{L}(T(\boldsymbol{X})), \boldsymbol{X} \sim F \in \mathcal{F}\}$. Then, the scoring rule*

$$\mathrm{S}_T(F, \boldsymbol{y}) = \mathrm{S}(T(F), T(\boldsymbol{y}))$$

is proper relative to $\mathcal{F}$. If S is strictly proper relative to $T(\mathcal{F})$ and T is injective, then the resulting scoring rule S_T is strictly proper relative to $\mathcal{F}$.

Transformation-based proper scores (Pic et al. 2025)

Proposition 1. *Let $\mathcal{F} \subset \mathcal{P}(\mathbb{R}^d)$ be a class of Borel probability measure on $\mathbb{R}^d$ and let $F \in \mathcal{F}$ be a forecast and $\boldsymbol{y} \in \mathbb{R}^d$ an observation. Let $T : \mathbb{R}^d \to \mathbb{R}^k$ be a transformation and let S be a scoring rule on $\mathbb{R}^k$ that is proper relative to $T(\mathcal{F}) = \{\mathcal{L}(T(\boldsymbol{X})), \boldsymbol{X} \sim F \in \mathcal{F}\}$. Then, the scoring rule*

$$\mathrm{S}_T(F, \boldsymbol{y}) = \mathrm{S}(T(F), T(\boldsymbol{y}))$$

is proper relative to $\mathcal{F}$. If S is strictly proper relative to $T(\mathcal{F})$ and T is injective, then the resulting scoring rule S_T is strictly proper relative to $\mathcal{F}$.

Idea 1: wavelet transform T, followed by any proper score on the coefficients

- Could be aggregated to per-scale CRPS, ES, …
- **No** obvious addition theorem like

$$CRPS = \sum_j CRPS_j$$

Idea 2: proper score on the spatial mean spectrum

- Would reward spread in the spatial structure
- I suspect that ensemble members show little variety in this respect …

JÜLICH Forschungszentrum

Transformation-based proper scores (Pic et al. 2025)

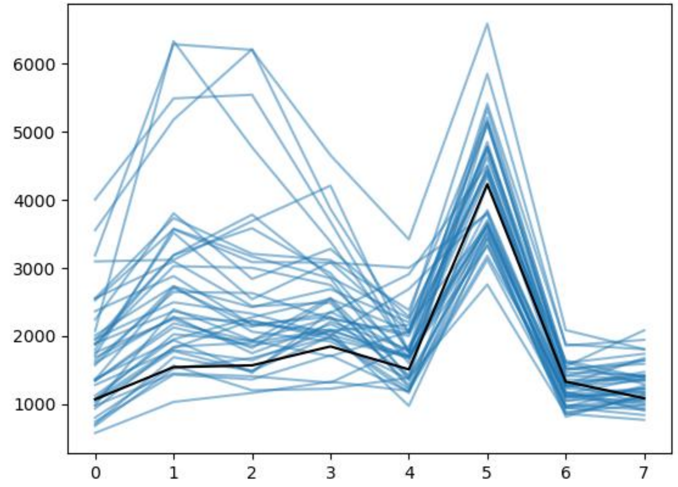
Proposition 1. *Let $\mathcal{F} \subset \mathcal{P}(\mathbb{R}^d)$ be a class of Borel probability measure on $\mathbb{R}^d$ and let $F \in \mathcal{F}$ be a forecast and $\boldsymbol{y} \in \mathbb{R}^d$ an observation. Let $T : \mathbb{R}^d \to \mathbb{R}^k$ be a transformation and let* S *be a scoring rule on $\mathbb{R}^k$ that is proper relative to $T(\mathcal{F}) = \{\mathcal{L}(T(\boldsymbol{X})), \boldsymbol{X} \sim F \in \mathcal{F}\}$. Then, the scoring rule*

$$\mathrm{S}_T(F, \boldsymbol{y}) = \mathrm{S}(T(F), T(\boldsymbol{y}))$$

is proper relative to $\mathcal{F}$. If S *is strictly proper relative to $T(\mathcal{F})$ and T is injective, then the resulting scoring rule S_T is strictly proper relative to $\mathcal{F}$.*

Idea 2: proper score on the spatial mean spectrum

- Would reward spread in the spatial structure
- I suspect that ensemble members show little variety in this respect …



Idea 3: a proper displacement score based on wavelet phases?

Transformation T would map the field to the phases, then we face some problems …

Problem 1 : Not every score is appropriate for circular quantities

→ use circular CRPS (Grimit et al. 2006)?

Problem 2 : I originally clipped phase errors to $\frac{\pi}{2}$ (score for a random phase), that probably makes it improper → skip this?

Problem 3 : In fields with many zeros (precipitation), I weighted with the amplitude to ignore the random noise, that destroys propriety (Gneiting & Ranjan 2011)?

→ Censored or weighted scoring rules? Add artificial patterns?

Why wavelets?

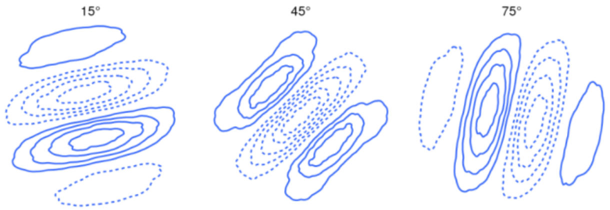
Wavelet scores and diagnostics

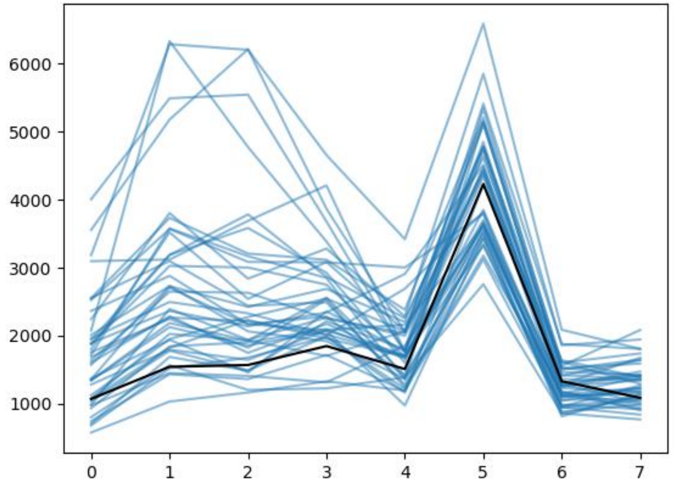
Generalizations to ensembles

Summary & Outlook

Summary

try out sadpy:





wavelets

- Localised basis functions for specific scales
- Fast algorithms for discrete wavelets

Deterministic verification

- Mean spectrum encodes information on Scale, Anisotropy, Direction
- Redundant transform allows for structure at every grid point
- Location scores based on complex phase

Ensemble verification

- Proper scores via “Transform + Aggregate” (Pic et al.)
- Open questions remain

JÜLICH Forschungszentrum

References

Buschow, S. (2022). Measuring Displacement Errors with Complex Wavelets. *Weather and Forecasting*, *37*(6), 953–970. https://doi.org/10.1175/WAF-D-21-0180.1

Buschow, S., & Friederichs, P. (2021). SAD: Verifying the scale, anisotropy and direction of precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, *147*(735), 1150–1169. https://doi.org/10.1002/qj.3964

Casati, B., Ross, G., & Stephenson, D. B. (2004). A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorological Applications*, *11*(2), 141–154. https://doi.org/10.1017/S1350482704001239

Gneiting, T., & Ranjan, R. (2011). Comparing Density Forecasts Using Threshold-and Quantile-Weighted Scoring Rules. *Journal of Business & Economic Statistics*, *29*(3), 411–422.

Nelson, J. D. B., Gibberd, A. J., Nafornita, C., & Kingsbury, N. (2018). The locally stationary dual-tree complex wavelet model. *Statistics and Computing*, *28*(6), 1139–1154. https://doi.org/10.1007/s11222-017-9784-0

Pic, R., Dombry, C., Naveau, P., & Taillardat, M. (2025). Proper scoring rules for multivariate probabilistic forecasts based on aggregation and transformation. *Advances in Statistical Climatology, Meteorology and Oceanography*, *11*(1), 23–58. https://doi.org/10.5194/ascmo-11-23-2025

Selesnick, I. W., Baraniuk, R. G., & Kingsbury, N. C. (2005). The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, *22*(6), 123–151. https://doi.org/10.1109/MSP.2005.1550194