Bayesian methods in neural networks for inverse atmospheric modelling

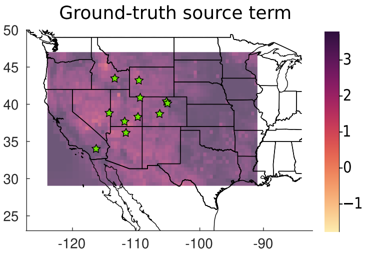
Antonie Brožová[1,2], Václav Šmídl[2], Ondřej Tichý[2]

[1]Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague
[2]Institute of Information Theory and Automation, CAS

21.6.2024

Source term estimation problem

- Data = amount of some substance (for example microplastics) at measuring stations (green stars in the figure)
- Task = estimate the source of the measured substance in space and time



Source term estimation problem

- $\mathbf{y} \in \mathbb{R}^{N}_{+,0}$ is the measured amount of the substance,
- $\mathbf{x}_{\text{lon,lat}} \in \mathbb{R}^{T}_{+,0}$ is the hypothetical emission over T time periods at a given longitude and latitude,
- $\mathbf{M}_{\text{lon,lat}} \in \mathbb{R}^{N \times T}_{+,0}$ is the source-receptor-sensitivity matrix for given longitude and latitude (provided by Nikolaos Evangeliou from NILU).

Then

$$\mathbf{y} = \sum_{\text{lon}} \sum_{\text{lat}} \mathbf{M}_{\text{lon,lat}} \mathbf{x}_{\text{lon,lat}}.$$

Challenge

This problem is **severaly ill-posed**, especially when the number of measurements is low

$\Rightarrow$ it is crucial to estimate the **uncertainty** of $\mathbf{x}$ as well.

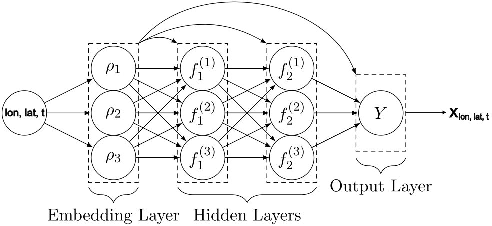
Spatial Bayesian neural networks (SBNN)

The l-th hidden layer is given by

$$f_l(.;\mathbf{W}_l,\mathbf{b}_l) = \mathbf{W}_l \tanh(.) + \mathbf{b}_l,$$

where

$$\mathbf{W}_l \sim \mathcal{N}(\mu_l^W, \sigma_l^W), \qquad \text{and} \qquad \mathbf{b}_l \sim \mathcal{N}(\mu_l^b, \sigma_l^b).$$



Spatial embedding layer

- Embedding of 3D coordinates by radial function on grid denoted by $\rho(\mathrm{lon}, \mathrm{lat}, \mathrm{t})$.
- Skip connections incorporate the spatial information directly into the estimation of the network parameters as

$$\mu_l^{W,b} = \alpha_l^{W,b} \rho(\mathrm{lon}, \mathrm{lat}, \mathrm{t}),$$
$$\sigma_l^{W,b} = \mathrm{softplus}\left(\beta_l^{W,b} \rho\left(\mathrm{lon}, \mathrm{lat}, \mathrm{t}\right)\right).$$

Pretraining to spatial lognormal process

- According to Zammit-Mangion et al., 2024, the SBNN can be trained to mimic lognormal process by optimization of Wasserstein distance.
- It proved necessary in order to obtain a reasonable estimate in the inverse modelling.

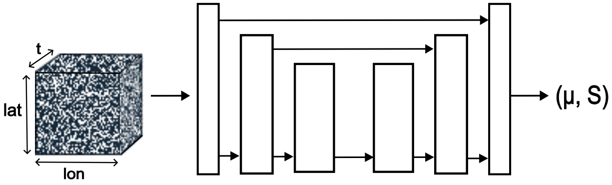
Inverse modelling

- $\mathbf{x}$ is obtained by sampling $\Rightarrow$ sample mean is the estimate of $\mathbf{x}$ and sample standard deviation is the estimate of uncertainty.
- MSE loss between prediction and measurements.

Deep Image Prior

- Convolutional neural networks are a good prior for image reconstruction problems (Ulyanov et al., 2018).
- Similarly to correlation of neighbouring pixels in images, closer spatial locations in source term should also correlate.

$\Rightarrow$ $\mathbf{x}$ will be estimated by U-net.



Variational U-net

Hierarchical model

- Likelihood: $\mathbf{y}|\mathbf{x} \sim \mathcal{N}\left(\mathbf{Mx}, 10^{-5}\right)$
- Prior:
  $\mathbf{x} \sim \mathcal{N}\left(\boldsymbol{\mu}_0, \mathrm{diag}\left(\boldsymbol{\tau}^{-1}\right)\right),$ where $\boldsymbol{\tau} \sim \mathcal{G}\left(\alpha_0, \beta_0\right),$
- Posterior: $\mathbf{x}|\mathbf{y} \sim \mathcal{N}\left(\boldsymbol{\mu}, \mathrm{diag}(\mathbf{S}^2)\right)$ and $\boldsymbol{\tau}|\mathbf{y} \sim \delta\left(\hat{\boldsymbol{\tau}}\right),$

$\Rightarrow$ estimation of $\boldsymbol{\mu}, \boldsymbol{S}$ and $\boldsymbol{\tau}$ by minimization of evidence lower bound.

Inverse modelling

- $\boldsymbol{\mu}$ is the estimate of $\mathbf{x}$ and $\boldsymbol{S}$ is its standard deviation quantifying uncertainty.

Data simulation

Real SRS matrix and simulated ground truth

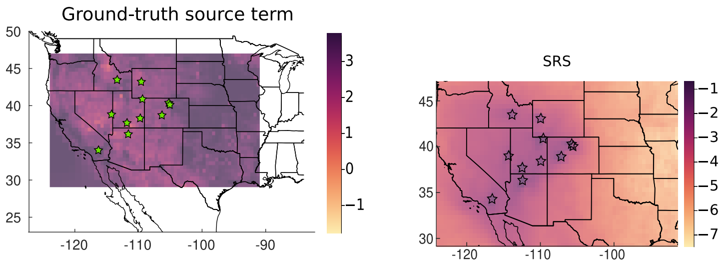
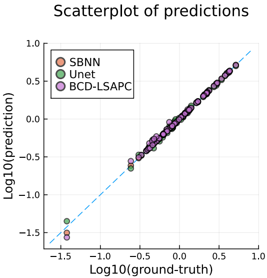
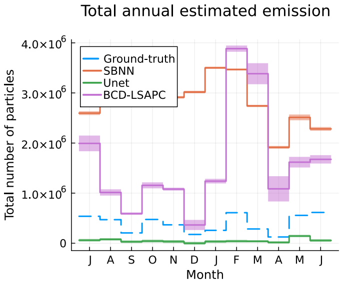


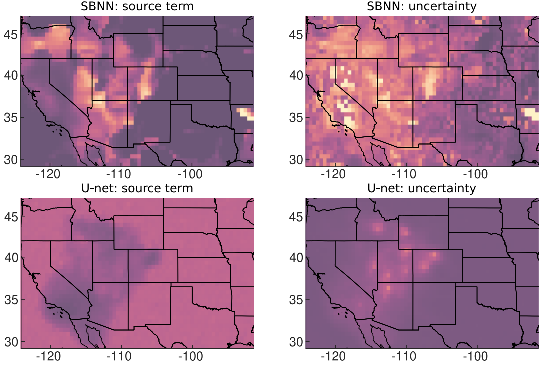
Figure: Log10(values)

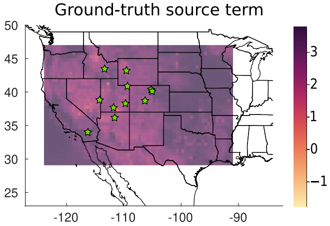
Quality of solution



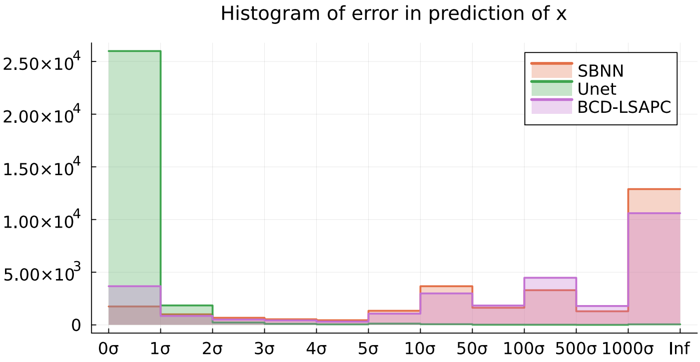


Estimates and uncertainty





Error of prediction



Conclusions

- SBNN strongly overestimates the deposition, especially in areas of low SRS, and underestimates its uncertainty.
- Variational U-net estimates the uncertainty better than the other two methods, with most error covered by standard deviation. It also reflects the SRS in its estimate, which makes it the most suitable model for the problem.

References:

- Zammit-Mangion, A., Kaminski, M. D., Tran, B.-H., Filippone, M., Cressie, N., 2024. Spatial Bayesian neural networks. Spatial Statistics, vol. 60, pp. 100825.
- Evangeliou, N., Tichy, O., Eckhardt, S., Zwaaftink, C. G., Brahney, J. , 2022. Sources and fate of atmospheric microplastics revealed from inverse and dispersion modelling: From global emissions to deposition. Journal of Hazardous Materials, 432, 128585.
- Ulyanov, D., Vedaldi, A., Lempitsky, V., 2018. Deep Image Prior. CVPR 2018.