Autoregressive Conditional Neural Processes

Wessel P. Bruinsma*[12], Stratis Markou*[2], James Requiema*[2], Andrew Y. K. Foong*[1],
Tom R. Andersson[3], Anna Vaughan[2], Anthony Buonomo[2], J. Scott Hosking[34], Richard E. Turner[12]
*Equal contribution, [1]Microsoft Research AI4Science, [2]University of Cambridge, [3]British Antarctic Survey, [4]The Alan Turing Institute

- AR CNPs: take a CNP and feed predictions back into the model in an autoregressive fashion.
- ✓ A "free" performance boost: no modifications to model or training procedure!
- ✓ Produces **correlated** and **non-Gaussian** predictions without requiring approximations.
- ✗ However, depends on number and order of data, and requires multiple forward passes.

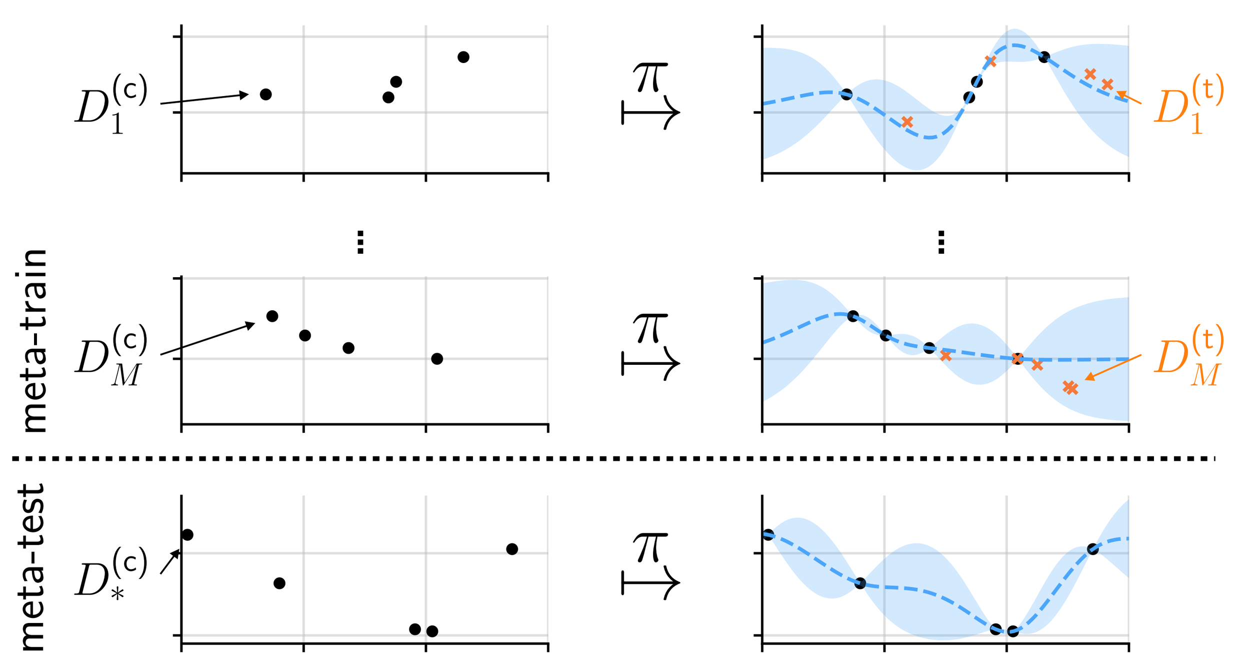
Meta-Learning

- In meta-learning, we have a **meta–data set**:

$$(D_m)_{m=1}^M \text{ with } D_m = \overbrace{D_m^{(c)}}^{\text{context set}} \cup \overbrace{D_m^{(t)}}^{\text{target set}}$$

- **Learn to predict** $D_m^{(t)}$ (×) given $D_m^{(c)}$ (•):

$$\pi: \quad \text{data sets } \mathcal{D} \quad \to \quad \text{predictions } \mathcal{P}$$



Neural Processes

- A **neural process** is a parametrisation of

$$\pi_\theta: \overbrace{\mathcal{D}}^{\text{data sets}} \to \underbrace{\mathcal{P}}_{\text{stochastic processes}}$$

using neural networks.

- $q_\theta(\mathbf{y} \mid \mathbf{x}, D)$: the density of $\pi_\theta(D)$ at $\mathbf{x}$.
- Training:

$$\theta \in \arg\max_{\theta \in \Theta} \sum_{m=1}^{M} \log q_\theta(\mathbf{y}_m^{(t)} \mid \mathbf{x}_m^{(t)}, D_m^{(c)})$$

✓ Extremely flexible and versatile
✓ Fast, probabilistic predictions
✓ Simple to train
✓ Work well in practice

Autoregressive Conditional Neural Processes (AR CNPs)

- Let $q(\mathbf{y} \mid \mathbf{x}, D)$ be a Conditional Neural Process [CNP; 1], possibly an existing/pretrained one!
- **AR CNPs:** feed predictions of CNP autoregressively back into the model:

$$q^{(\text{AR CNP})}(\mathbf{y}_{1:3} \mid D) = q(y_1 \mid x_1, D)\, q(y_2 \mid x_2, D \cup (x_1, y_1))\, q(y_3 \mid x_3, D \cup (\mathbf{x}_{1:2}, \mathbf{y}_{1:2})).$$

- Neural AR models certainly not new. Running neural processes in AR mode not yet explored!

✓ A "free" performance boost: **no modifications to model or training procedure!**

	Correlated preds	Non-Gaussian preds	Exact training	Consistent preds
Conditional NPs [1]	✗	✓	✓	✓
Gaussian NPs [2]	✓	✗	✓	✓
Latent-variable NPs [3]	✓	✓	✗	✓
Autoregressive CNPs	✓	✓	✓	✗

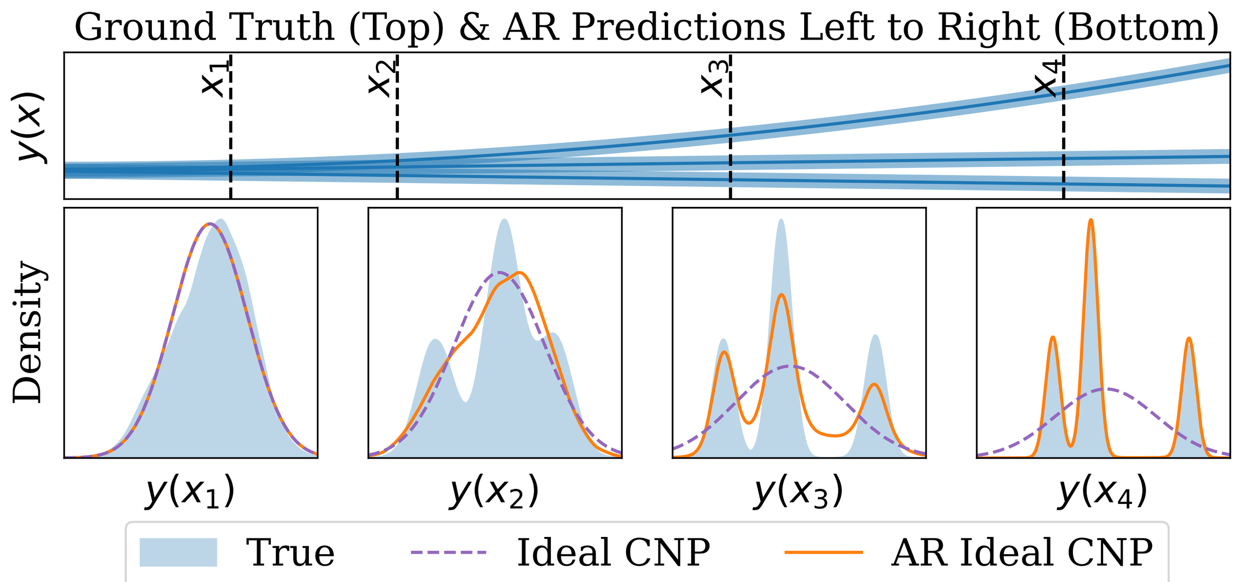
✗ However, depends on number and order of data, and requires multiple forward passes of the CNP.

Theoretical Analysis of Idealised Setting

- Analyse idealised setting of **infinite data** and **infinite capacity**.

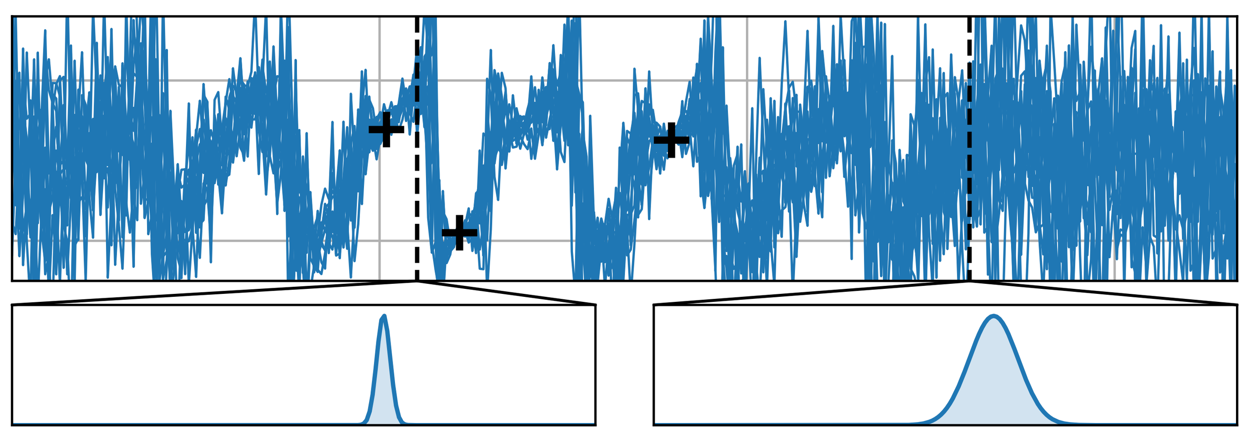
⇒ moment-matched posterior prediction map!

- **Prop. 2.1:** In this setting, CNPs perform better than GNPs.

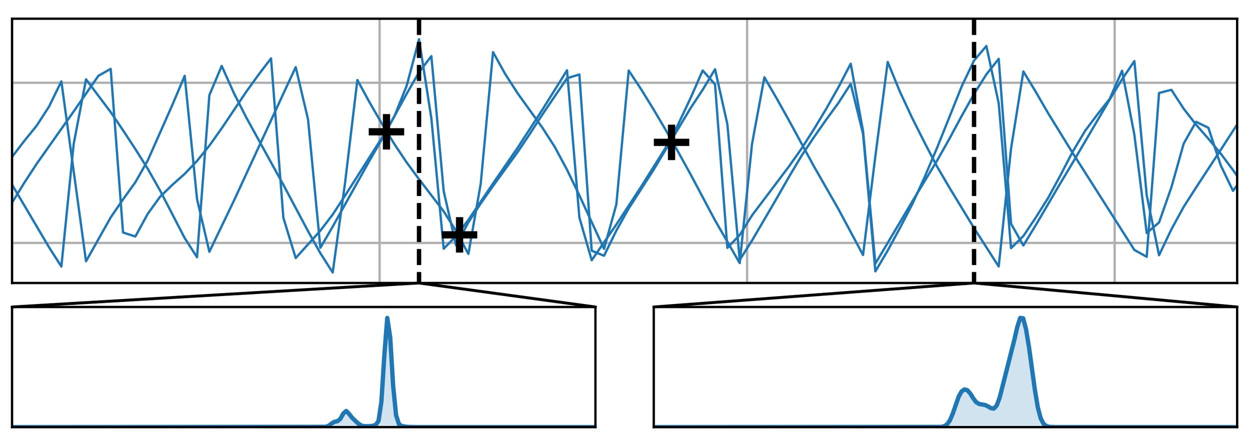


Synthetic Experiment on Sawtooth Data

- Sample of ConvCNP [4]:

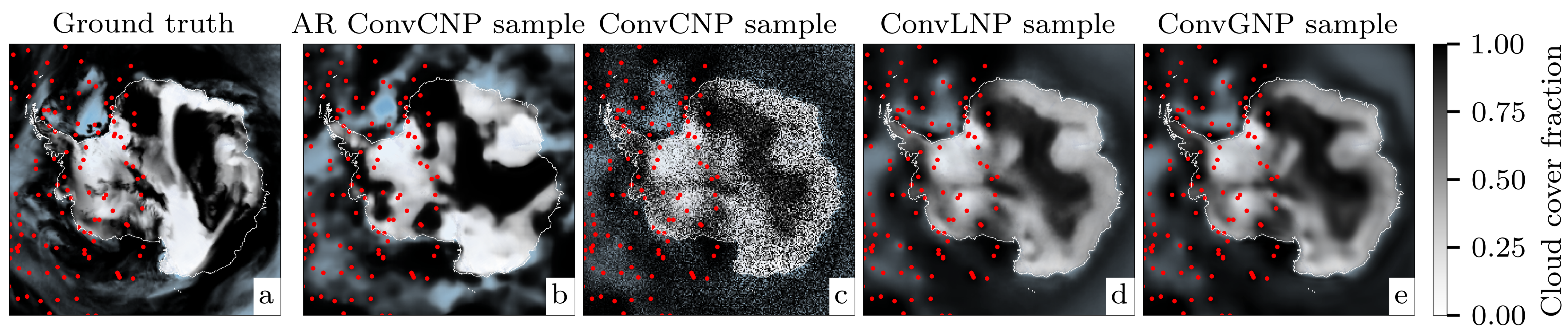


- **Same ConvCNP**, but run in AR mode:



Real-World Experiment on Cloud Coverage Data

- AR ConvCNP spatially interpolates simulated cloud cover observations over Antarctica:



References

[1] M. Garnelo, D. Rosenbaum, C. J. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. J. Rezende, and S. M. A. Eslami, "Conditional neural processes," in *Proceedings of 35th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 80, PMLR, 2018.

[2] S. Markou, J. Requeima, W. P. Bruinsma, A. Vaughan, and R. E. Turner, "Practical conditional neural processes via tractable dependent predictions," in *Proceedings of the 10th International Conference on Learning Representations*, 2022.

[3] M. Garnelo, J. Schwarz, D. Rosenbaum, F. Viola, D. J. Rezende, S. M. A. Eslami, and Y. W. Teh, "Neural processes," in *Theoretical Foundations and Applications of Deep Generative Models Workshop, 35th International Conference on Machine Learning*, 2018.

[4] J. Gordon, W. P. Bruinsma, A. Y. K. Foong, J. Requeima, Y. Dubois, and R. E. Turner, "Convolutional conditional neural processes," in *Proceedings of the 8th International Conference on Learning Representations*, 2020.

Paper: `https://wessel.ai/pdf/arcnps`   Code: `https://github.com/wesselb/neuralprocesses`