Probabilistic bias adjustment of seasonal predictions of Arctic Sea Ice*

Parsa Gooya† and Reinel Sospedra-Alfonso

Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada

NEURAL INFORMATION PROCESSING SYSTEMS

Motivation

- Seasonal predictions of Arctic sea ice concentration (SIC) is key to mitigate the negative impact and assess potential opportunities posed by the rapid decline of sea ice coverage.
- Seasonal predictions produced with climate models have systematic biases and complex spatio-temporal errors.
- Operational forecasts are routinely bias corrected and calibrated.
- Arctic sea ice predictions are mainly corrected based on limited one-to-one deterministic post-processing methods.
- Decision-making requires proper quantification of uncertainty and likelihood of events, particularly of extremes.

We introduce a probabilistic bias correction scheme based on a conditional Variational Autoencoders (cVAE).

Problem Statement

Goal: t : initialization time, l : lead time

Learn probabilistic mapping from biased *ensemble mean predictions* $\bar{x}_{tl}$ to the *observational distribution* $p(y|\bar{x}_{tl})$ conditioned on $\bar{x}_{tl}$

Learning objective:

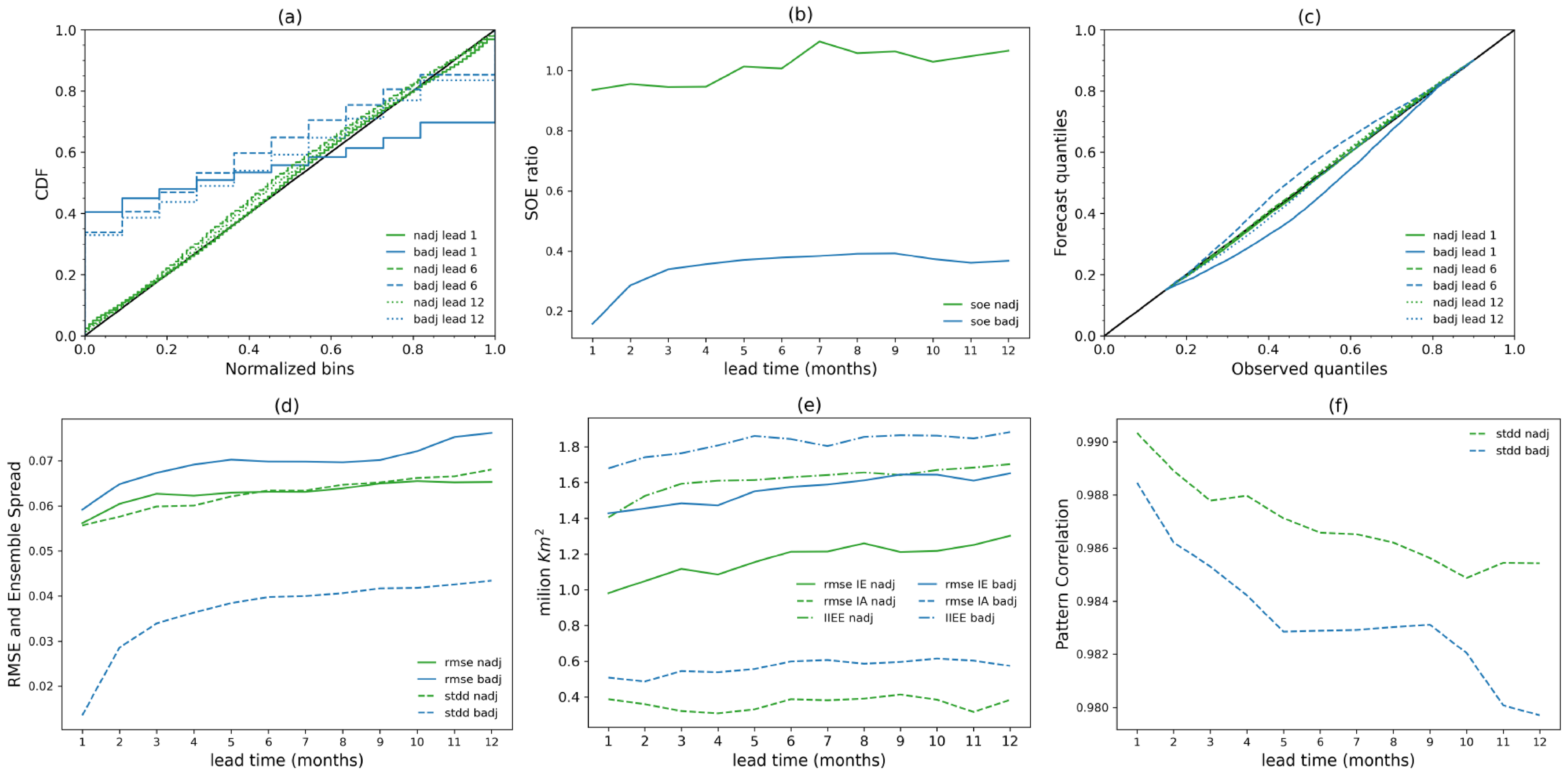
Maximize the likelihood of the observation y_{tl} from the distribution $p(y|\bar{x}_{tl})$

Generative Framework:

- cVAE learns the conditional distribution of data using latent variable z with prior distribution $p_\omega(z|\bar{x}_{tl}) : \mathrm{N}(\mu_{\mathrm{NN}_\omega}(\bar{x}_{tl}),\ \sigma_{\mathrm{NN}_\omega}(\bar{x}_{tl}))$
- Optimizes the **E**vidence **L**ower **Bo**und as the surrogate objective:

$$\log p(y = y_{tl}|\bar{x}_{tl}) \geq$$
$$-KL\left(q_\phi(z\,|y_{tl}, \bar{x}_{tl})\,||p_\omega(z|\bar{x}_{tl})\right) + \mathbb{E}_{q_\phi(z\,|y_{tl}, \bar{x}_{tl})}[\log p_\theta(y = y_{tl}|z, \bar{x}_{tl})]$$

- KL is the Kullback–Leibler Divergence regularizing the latent space and the expectation term is proportional to **M**ean **S**quare **E**rror (when,)
- Distributions are parametrized as Gaussians using neural networks with parameters θ for the output, ϕ for the posterior and ω for the prior.



nadj : bias adjusted outputs from cVAE (100 members)

badj : benchmark lead time dependent climatological mean adjustment (10 members)

IIEE : Integrated Ice Edge Error

IE : Integrated Ice Extent

IA : Integrated Ice Area

CDF : Cumulative Density Function

SOE : Spread over Error score

QQ : Quantile-Quantile plots

(a) CDF of rank histograms of the nadj/badj versus lead times measured at marginal ice grid cells. Only three lead times are plotted for visibility. **(b)** SOE versus lead time showing reliability. **(c)** QQ plots at three lead times comparing the distribution of SIC at marginal ice grid cells with obs. **(d)** RMSE (solid) over initialization time between the ensemble mean nadj/badj compared to obs at grid cells level averaged over the entire region. The dashed line shows the global mean ensemble spread averaged over initialization time. **(e)** For each lead time, RMSE of SIA (solid line) and SIE (dashed line) over initialization time, and average IIEE (dotted line) over initialization time is compared between ensemble mean nadj/badj and obs. **(f)** same as **(e)** but for pattern correlation relative to obs. Note: CDF and QQ are reported at critical marginal ice grid cells ($0.15 \leq \mathrm{SIC} \leq 0.90$) to avoid heavily weighting for fully covered or open ocean regions.

Results

Probabilistic metrics:

- **Fig 1a:** Cumulative Distribution Function (**CDF**) curves for nadj show close to **uniform rank histograms,** indicating observations have similar probability of falling at every rank, confirming a **well-calibrated** ensemble. Though degrading, CDFs remain close to uniform at longer lead times.
- **Fig 1b:** Spread-over-Error (**SOE**) ratio consistently close to 1 for nadj, indicating ensemble members **indistinguishable** from observation. badj ensemble predictions **overconfident** with a heavy-tailed CDF and SOE<1.
- **Fig 1c:** Quantile-Quantile (**QQ**) plots indicating agreement between SIC distribution of nadj forecast ensemble and observational distribution.

Deterministic metrics:

- **Fig d-f:** Ensemble mean nadj is more accurate than badj compared to observations at grid level **(panel d)**, pattern correlation **(panel f)**, and for bulk measures **(panel e)** of integrated ice coverage (**SIA**), extent (**SIE**) and the boundaries of the edges of ice **(IIEE)**

- **nadj outperforms badj for all metrics.**
- **Errors increase with lead time as uncertainty grows expectedly (panel d).**
- **nadj SOE ratio remains close to 1 for all lead times (panel b).**
- **cVAE maps raw biased forecasts into skillful, reliable, well-calibrated forecasts.**

Architecture and Inference

Encoder/Prior:

- Input (5) → 3x3 partial convolution (16) → Layer normalization (16) → DoubleConvNeXt (32) → MaxPool (32) → DoubleConvNeXt → MaxPool (64) → DoubleConvNeXt (128) → MaxPool (128) → DoubleConvNeXt (256) → MaxPool (256) → DoubleConvNeXt (256) → Layer normalization (256) → Dense (2 x 1000)

Decoder:

- Latent samples (1000) → Dense (256) → Upsampling (256) → DoubleConvNeXt (128) → Upsampling (128) → DoubleConvNeXt (64) → Upsampling (64) → DoubleConvNeXt (32) → Upsampling (32) → DoubleConvNeXt (16) → Layer normalization (16) → ReLu (16) } → 1x1 partial convolution (1)

Inference:

Find scaling factor for prior standard deviation based on SOE over validation set → sample $z^l\ (l = 1, .., n)$ from scaled prior $p_\omega(z|\bar{x}_{tl})$ → generate outputs using decoder $p_\theta\ (y|z^l, \bar{x}_{tl}) : \mathrm{N}\left(\mu_{\mathrm{NN}_\theta}(z^l, \bar{x}_{tl}),\ \sigma^2 \mathrm{I}\right)$

Data

- **Model:** Retrospective 12-months seasonal forecasts of Arctic SIC from CanSIPSv3's CanESM5 during 1980 to 2021.
- **Observation:** Satellite-based NOAA/NSIDC Climate Data Record of passive microwave SIC v4 during 1980 to 2021.

Future work

- Improve spectral bias (smooth images)
- Explore the source of skill and generated variability
- Running the model autoregressively to incorporate time dependence

References

- Cavalieri, D. J. and C. L.: Parkinson. Arctic sea ice variability and trends, 1979–2010. The Cryosphere, 6(4):881–889, 2012.
- Sohn, K., H. Lee, and X. Yan.: Learning structured output representation using deep conditional generative models. 28, 2015.
- Lin, H., W. J. Merryfield, R. Muncaster, et al. The Canadian Seasonal to Interannual Prediction System version 2 (CanSIPSv2). Weather and Forecasting, 35(4):1317 – 1343, 2020.
- Swart, N., C., J. N. S. Cole, V. V. Kharin, et al. The Canadian Earth System model version 5 (CanESM5.0.3). Geoscientific Model Development,12(11):4823–4873, 2019.
- Meier, W. N., F. Fetterer, A. Windnagel, and J. S. Stewart. NOAA/NSIDC Climate Data Record of passive microwave Sea Ice Concentration, version 4, 2021.

[*] https://arxiv.org/abs/2510.09891

† **Corresponding Author:** Parsa Gooya (parsa.gooya@ec.gc.ca)