Information Entropy Initialized Concrete Autoencoder for Optimal Sensor Placement and Reconstruction of Geophysical Fields

Nikita A. Turko[1], Alexander A. Lobashev[2], Konstantim V. Ushakov[3,1], Maxim N. Kaurkin[3], and Rashit A. Ibrayev[3,4]

[1] *Moscow Institute of Physics and Technology*
[2] *Skolkovo Institute of Science and Technology*
[3] *Shirshov Institute of Oceanology, Russian Academy of Sciences*
[4] *Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences*

RuSCDays22

Overview

Introduction

Introduction

Operational ocean forecasting based on:

- Observation data
- Ocean circulation model
- Large CPU cluster

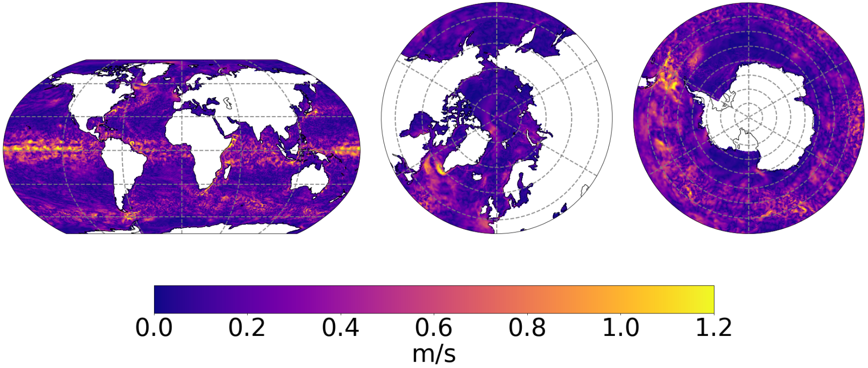


Figure: Example of ocean speed forecast

Introduction

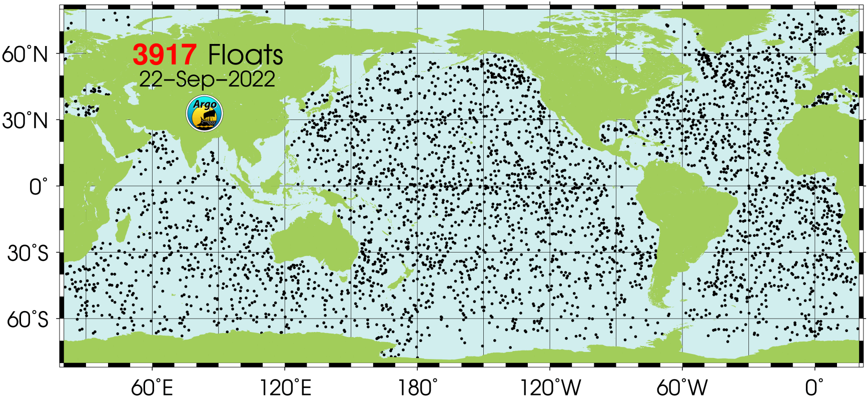


Figure: Locations of ARGO drifters, measuring temperature and salinity profiles

Motivation

Motivation

- Optimize ocean forecast operational systems
- Find locations to place new sensors
  - To select $k \in [k_{\min}, k_{\max}]$ sensors from n grid nodes search space grows exponentially as $\sum_{k=k_{\min}}^{k_{\max}} C_n^k$
  - Direct combinatorial search is impossible

- Most of common approximate methods use the singular value decomposition (SVD) which scales as $O(n^3)$ or $O(n^2)$ in different implementations.
- They cannot be applied to large grids of size $1440x720 \approx 10^6$ $(0.25^o x 0.25^o)$

Statistical approach

Statistical approach

1. Physical field as random variable realization
2. Informational entropy calculation
3. Proposition of optimal sensor locations
4. Sensor coordinates optimisation

Estimating uncertainty of a physical field

Idea:

- Place sensors in locations where a physical field has high uncertainty

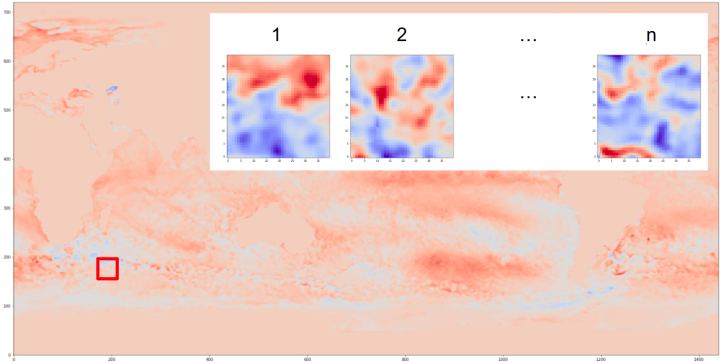


Figure: Example of patches extracted from sea the surface temperature anomaly field. Indices $1, 2, \ldots, n$ correspond to patches takes in the same spatial location at different time moments

Informational entropy

- Uncertainty of a physical field could be estimated using the information entropy

The informational entropy of the physical field as a function of spatial coordinates x, y can be estimated as:

$$H(x,y) = -\int \mathbb{P}(\xi|x,y) \log \mathbb{P}(\xi|x,y) d\xi \\ = -\frac{1}{N}\sum_{i=1}^{N} \log \mathbb{P}(\xi_i|x,y),\ \xi_i \sim \mathbb{P}(\xi_i|x,y) \quad (1)$$

ξ - values of the physical field, taken along the temporal dimension

Density estimation via autoregressive generative modeling

Conditional PixelCNN

Suppose our set of physical fields is encoded as a set of $L \times L$ images or patches s_i cropped from the domain of interest $\mathcal{D}$, each of which is labeled with spatial coordinates of the patch center $\mathbf{r} = \{r_1, \dots, r_K\} \in \mathcal{D}$

$$\mathcal{S}^{(\mathbf{r}_i \in \mathcal{D})} = \{s_1^{\mathbf{r}_1}, \dots, s_N^{\mathbf{r}_N}\}. \tag{2}$$

Joint density of all pixels could be expanded as a product of conditional densities

$$\mathbb{P}(s^{\mathbf{r}}|\mathbf{r}) = \prod_{i=1}^{L \times L} \mathbb{P}([s^{\mathbf{r}}]_i | [s^{\mathbf{r}}]_1, \dots, [s^{\mathbf{r}}]_{i-1}, \mathbf{r}), \tag{3}$$

where $[s^{\mathbf{r}}]_i$ stands for the i-th pixel of the image $s^{\mathbf{r}}$ with respect to the chosen ordering

Conditional PixelCNN

The network is trained on the dataset $\mathcal{S}^{(\mathbf{r}_i \in \mathcal{D})}$ maximizing probability of observed physical fields or equivalently by minimizing negative log-likelihood

$$L(\theta) = -\sum_{i=1}^{N} \log \mathbb{P}_\theta(s_i^{\mathbf{r}_i} | \mathbf{r}_i), \tag{4}$$

where θ is the vector of parameters of Conditional PixelCNN.
Then we compute entropy as

$$H(\mathbf{r}) = -\mathbb{E}_s \log \mathbb{P}_\theta(s|\mathbf{r}) = -\frac{1}{\#\{\mathbf{r}^i : |\mathbf{r}^i - \mathbf{r}| < \varepsilon\}} \sum_{|\mathbf{r}^i - \mathbf{r}| < \varepsilon} \log \mathbb{P}_\theta(s_i^{\mathbf{r}_i} | \mathbf{r}) \tag{5}$$

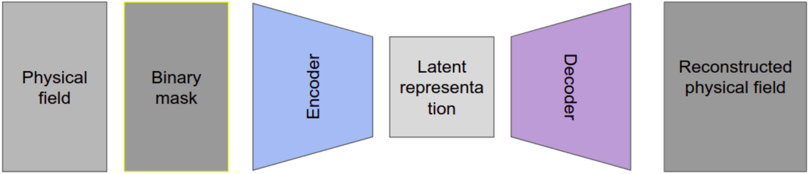
Proposed optimal sensor locations

Computed information entropy field can be used to propose optimal sensor locations by sampling from the distribution

$$\mathbb{P}(\mathbf{r}) = \frac{e^{\frac{1}{\tau}H(\mathbf{r})}}{\int_{\mathcal{D}} e^{\frac{1}{\tau}H(\mathbf{r})} d\mathbf{r}} \tag{6}$$

where we set hyperparameter $\tau = 0.2$ for the entropy field measured in nats per computational grid cell

Concrete Autoencoder

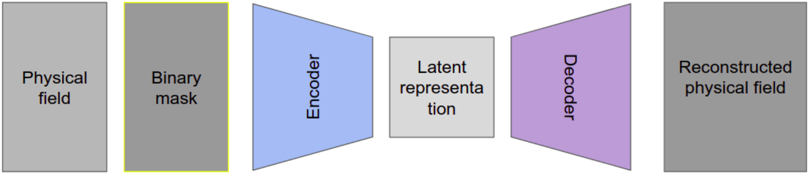


Minimizing the loss function

$$\mathcal{L}_G = \mathbb{E}_{S_{\text{full}}} ||G(S_{\text{full}} \cdot \text{mask}, w) - S_{\text{full}}||_{L_2} + \lambda \cdot \mathbb{E}|\text{mask}| \tag{7}$$

where the function G takes as input the physical field S_{full} in the entire simulation area, multiplies it component by the binary mask and tries to restore the original field

Concrete Autoencoder with Least Square GAN loss



$$\mathcal{L}_G = \lambda_1 \mathcal{L}_{\text{LSGAN}} + \lambda_2 \mathcal{L}_{\text{pixel-wise}} + \lambda_3 \mathcal{L}_{\text{sensors}} \quad (8)$$

where we use $\lambda_1 = 10^{-4}$, $\lambda_2 = 1$ and λ_3 dynamically changes during training from 0 to 1.

- We add adversarial term with a discriminator D which tries to distinguish real and reconstructed physical fields:

$$\mathcal{L}_{\text{LSGAN}} = \text{MSE}(D(G(\hat{M})), \mathbb{I}) \equiv ||D(G(\hat{M})) - \mathbb{I}||_{L_2}$$

$$\mathcal{L}_{\text{pixel-wise}} = \mathbb{E}_{S_{\text{full}}} ||G(S_{\text{full}} \cdot \text{mask}, w) - S_{\text{full}}||_{L_2}$$

$$\mathcal{L}_{\text{sensors}} = \mathbb{E}|\text{mask}|$$

Baselines

Baselines

Climate

$$S^{climate}(i,j,d) = \frac{1}{N^{years}} \sum_{y=1}^{N^{years}} S(i,j,y,d) \quad (9)$$

where $S(i,j,y,d)$ - the value of physical field with coordinates (i,j) at day number $d = \{1, 2, ..., 365\}$ in year y from train set, N^{years} - number of years with day d in train set

PCA-QR

Principal Component Analysis (Proper Orthogonal Decomposition or the method of Empirical Orthogonal Functions) with pivoted QR decomposition

Data

Experimental data

Global coupled ocean-ice model INMIO Compass-CICE-ERA5 with resolution 0.25 x 0.25, 17 model years from 2004 to 2020

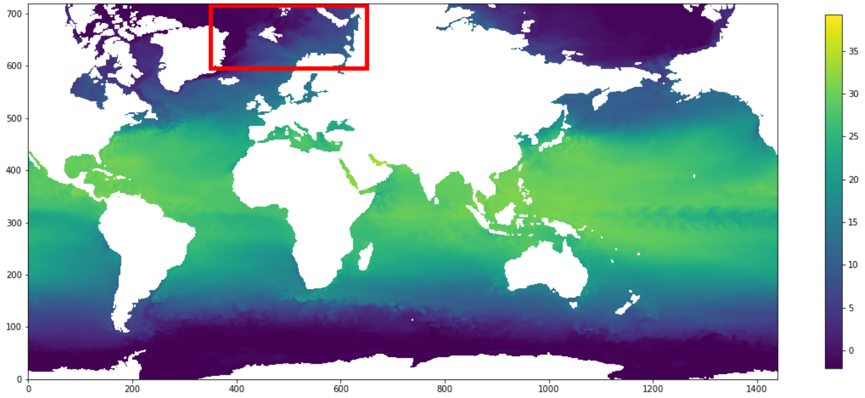


Figure: Temperature at 3 m

Results

Approximation of informational entropy

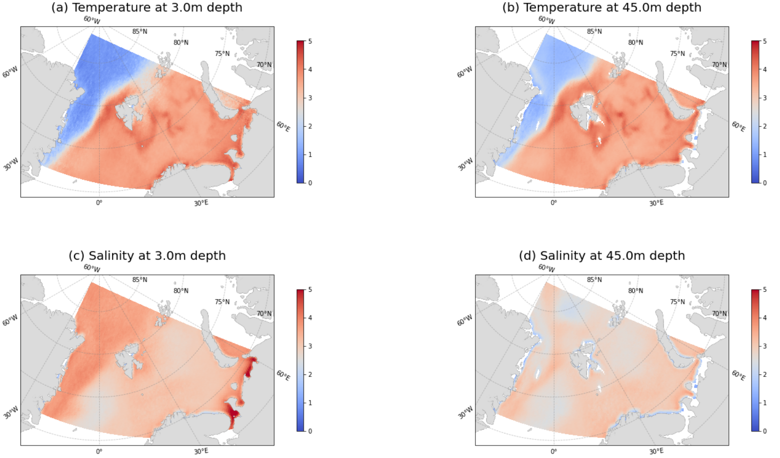


Figure: Smoothed ensemble mean information entropy of geophysical fields: temperature at (a) 3 meter and (b) 45 m depth; salinity at (c) 3 meter and (d) 45 m depth

Initialised mask

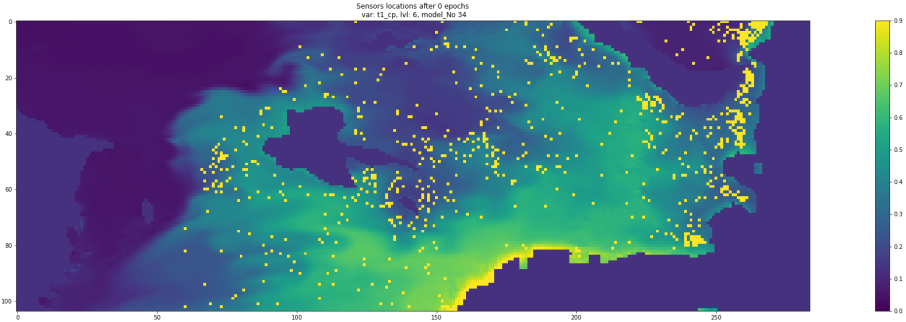


Figure: Proposed initial sensor locations based on Information entropy field for temperature at 45m depth

Optimizing the mask, 30 epochs

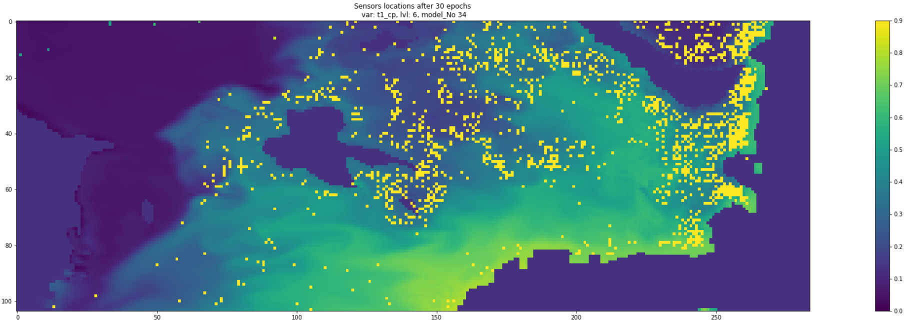


Figure: Proposed initial sensor locations based on Information entropy field for temperature at 45m depth, after 30 epochs

Optimizing the mask, 60 epochs

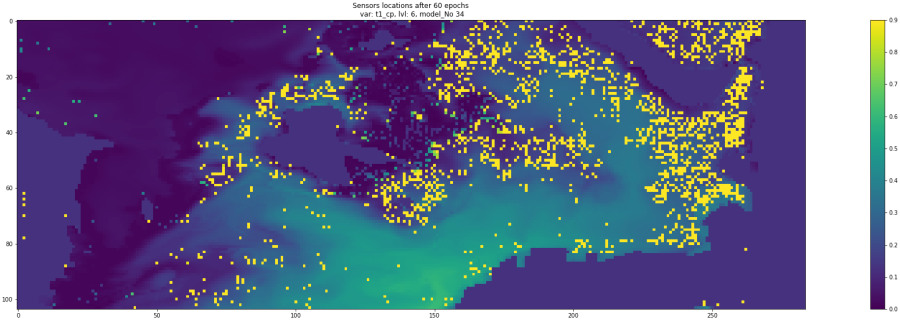


Figure: Proposed initial sensor locations based on Information entropy field for temperature at 45m depth, after 60 epochs

Optimizing the mask, 90 epochs

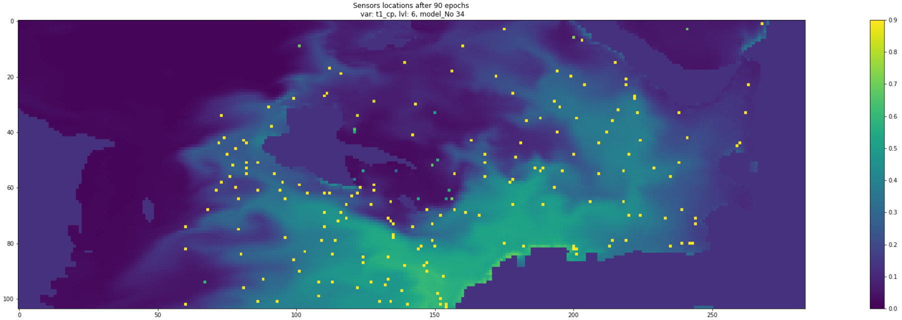


Figure: Proposed initial sensor locations based on Information entropy field for temperature at 45m depth, after 90 epochs

Optimizing the mask, 180 epochs

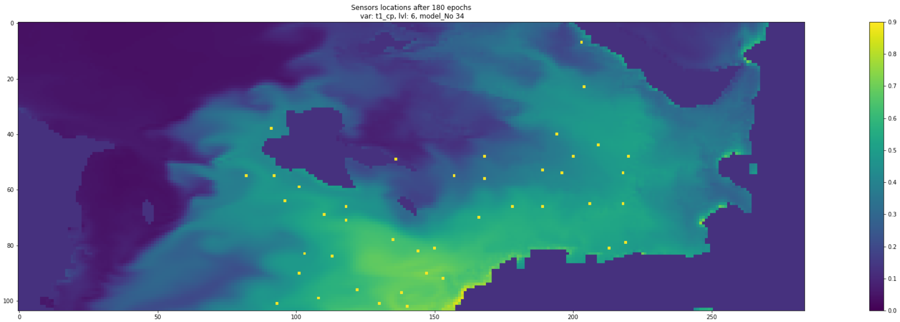


Figure: Proposed initial sensor locations based on Information entropy field for temperature at 45m depth, after 180 epochs

Reconstructed field

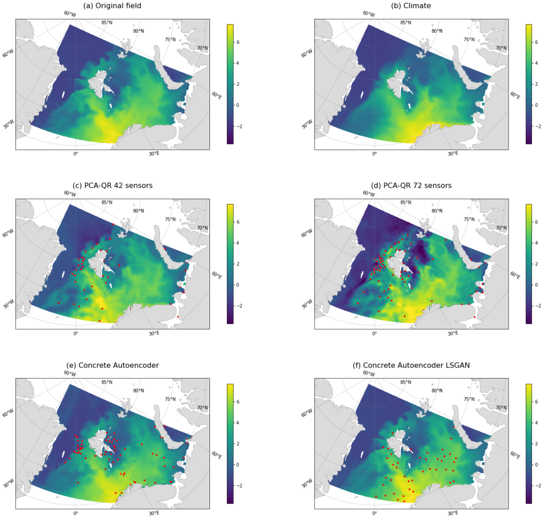


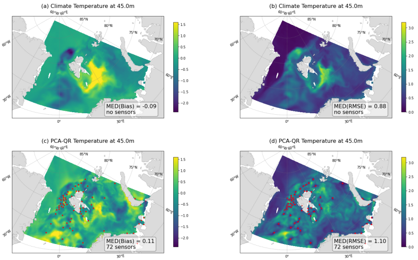
Figure: Temperature at 45.0 m depth, 2017-08-07

Spatial distribution of reconstruction error

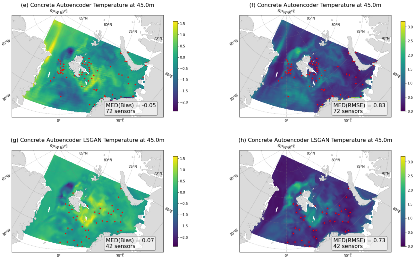
$$Bias(i,j) = \frac{1}{\#\{\tau \in TestSet\}} \sum_{\tau \in TestSet} (S^{recon}(i,j,\tau) - S^{ref}(i,j,\tau)) \quad (10)$$

$$RMSE(i,j) = \sqrt{\frac{1}{\#\{\tau \in TestSet\}} \sum_{\tau \in TestSet} (S^{recon}(i,j,\tau) - S^{ref}(i,j,\tau))^2} \quad (11)$$

Spatial distribution of reconstruction error



(a) Baselines



(b) Concrete Autoencoder

Figure: Bias/RMSE temperature reconstruction at depth 45m

Error on all test set

$$Bias(\tau) = \frac{1}{N^i}\frac{1}{N^i}\sum_{i=1}^{N^i}\sum_{j=1}^{N^j}(S^{recon}(i,j,\tau) - S^{ref}(i,j,\tau)) \tag{12}$$

$$RMSE(\tau) = \sqrt{\frac{1}{N^i}\frac{1}{N^i}\sum_{i=1}^{N^i}\sum_{j=1}^{N^j}(S^{recon}(i,j,\tau) - S^{ref}(i,j,\tau))^2} \tag{13}$$

Error on all test set

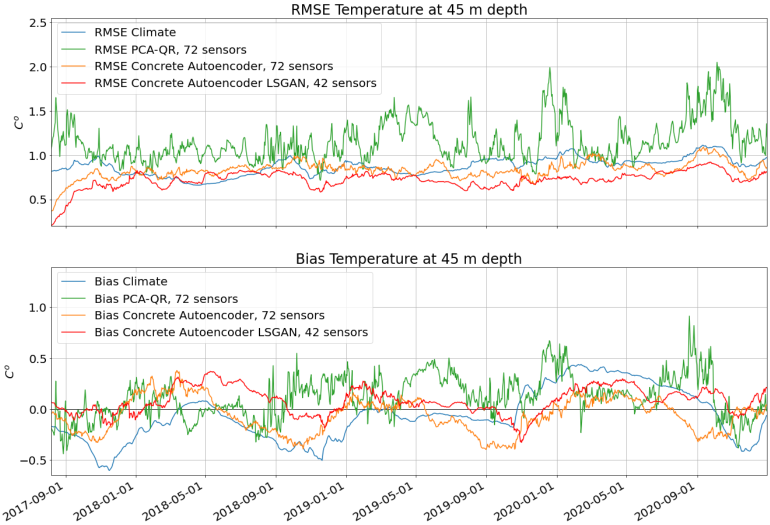


Figure: Temperature field reconstruction accuracy against original model data

Test set reconstruction errors

Temperature 3m			
Method	Number of sensors	MED(Bias)	MED(RMSE)
Climate	0	-0.19	0.98
PCA with QR	77	0.13	1.03
Concrete Autoencoder	77	-0.07	**0.73**

Temperature 45m			
Method	Number of sensors	MED(Bias)	MED(RMSE)
Climate	0	-0.09	0.88
PCA with QR	72	0.11	1.10
Concrete Autoencoder	72	-0.05	0.83
Concrete Autoencoder LSGAN	42	0.07	**0.73**

Salinity 3m			
Method	Number of sensors	MED(Bias)	MED(RMSE)
Climate	0	0.58	0.84
PCA with QR	57	-0.03	0.66
Concrete Autoencoder	57	0.05	**0.53**

Salinity 45m			
Method	Number of sensors	MED(Bias)	MED(RMSE)
Climate	0	0.59	0.72
PCA with QR	61	0.02	**0.30**
Concrete Autoencoder	61	0.26	0.41

Conclusion

Conclusion

- Proposed a method for optimal sensor placement and reconstruction of geophysical fields
- Proposed method outperforms baselines
- The addition of LSGAN loss improves reconstruction accuracy

Thank you for attention!