



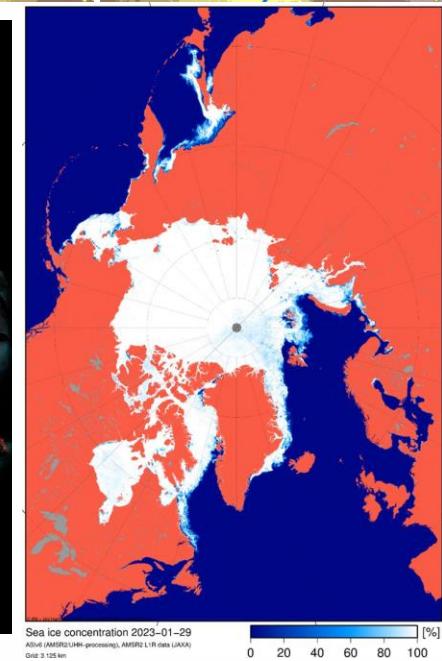
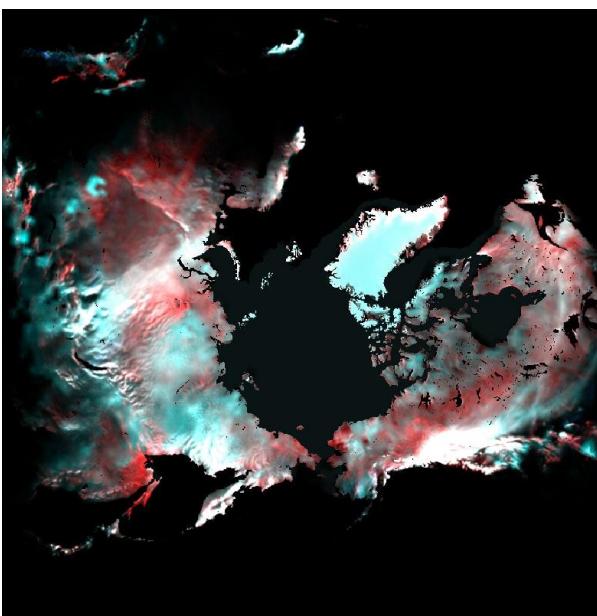
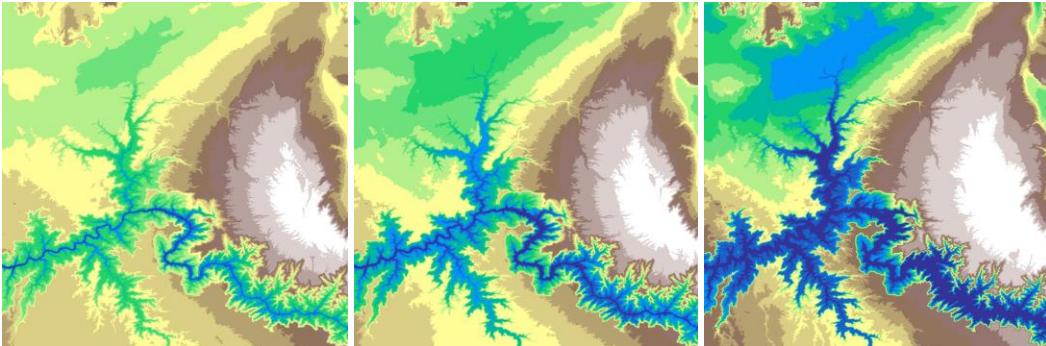
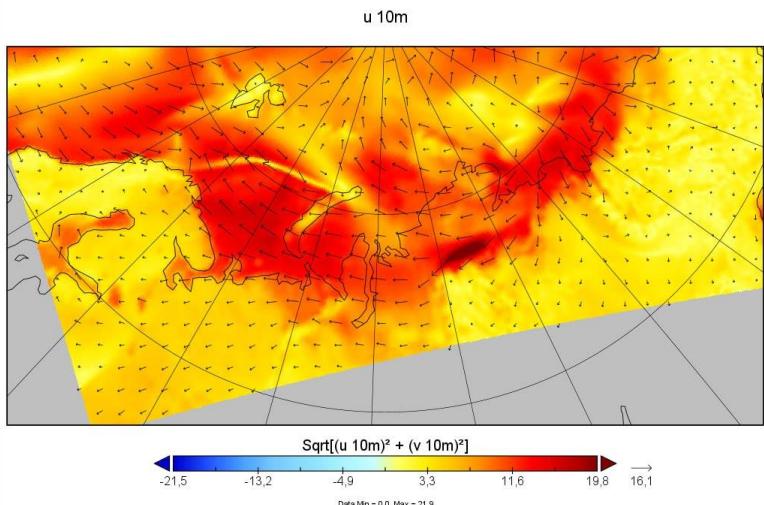
Гибридные модели глубокого обучения для повышения качества прогностического моделирования пространственно-временных процессов (на примере ледовой обстановки в Арктике)

Борисова Юлия Ивановна
м.н.с. Лабкии

Специальность «Искусственный интеллект и машинное обучение»
Научный руководитель: Никитин Николай Олегович, к.т.н., доцент ФЦТ

Постановка задачи

Целью работы является повышение качества прогностического моделирования пространственных явлений реального мира посредством разработки методов и алгоритмов гибридного моделирования облегченной структуры.



Имеющиеся решения (1/3)

Физические модели на основе
дифференциальных уравнений (NEMO - LIM):

- ресурсоемкость;
- обилие входных данных в виде граничных и начальных условий;
- параметризация под каждую отдельную тестовую территорию.

Решения на основе глубокого обучения:

- ресурсоемкость;
- большое количество дополнительных предикторов;
- ограниченность моментов запуска в году.

Предлагаемое решение:

- базируется только на предыстории прогнозируемого параметра;
- легковесное как на этапе обучения так и предсказания;
- унифицированный запуск для любой тестовой территории.

Имеющиеся решения (2/3)

Polar Modelling
Standard Python

Sea ice forecasting (IceNet)



Date = September 2020 | Lead time = 1 months
Lead time = 1 months
Observation
Predicted
Accuracy acc.: 91.7%
SE error: 0.207 ml km²

Coca-Castro (2022)

license MIT launch binder
render passing
DOI 10.24424/m8ew-pg51

IceNet получает в качестве входных данных 50 усредненных климатических переменных за месяц.

Результаты IceNet представляют собой прогнозы трех классов концентрации морского льда (SIC) ($SIC \leq 15\%$, $15\% < SIC < 80\%$ и $SIC \geq 80\%$) на следующие 6 месяцев в виде дискретных распределений вероятностей в каждой ячейке сетки.

Две последние вероятности ледового класса суммируются для получения вероятности морского льда $p = P(SIC > 15\%)$.

Input variable name	Lead or lag (months)	Source
Linear trend forecast	1-6	Computed from OSI-SAF
SIC	1-12	OSI-SAF
2-metre air temperature anomaly	1-3	ERA5
500 hPa air temperature anomaly	1-3	ERA5
Sea surface temperature anomaly	1-3	ERA5
Downwards surface solar radiation anomaly	1-3	ERA5
Upwards surface solar radiation anomaly	1-3	ERA5
Sea level pressure anomaly	1-3	ERA5
500 hPa geopotential height anomaly	1-3	ERA5
250 hPa geopotential height anomaly	1-3	ERA5
10 hPa zonal wind speed	1-3	ERA5
10-metre X-direction wind speed	1	ERA5
10-metre Y-direction wind speed	1	ERA5
Land mask	N/A	OSI-SAF
Cosine of initialisation month index	N/A	N/A
Sine of initialisation month index	N/A	N/A

Supplementary Table 2 | **IceNet's input variables and their sources for the observational training dataset.**

Clone the IceNet GitHub repo

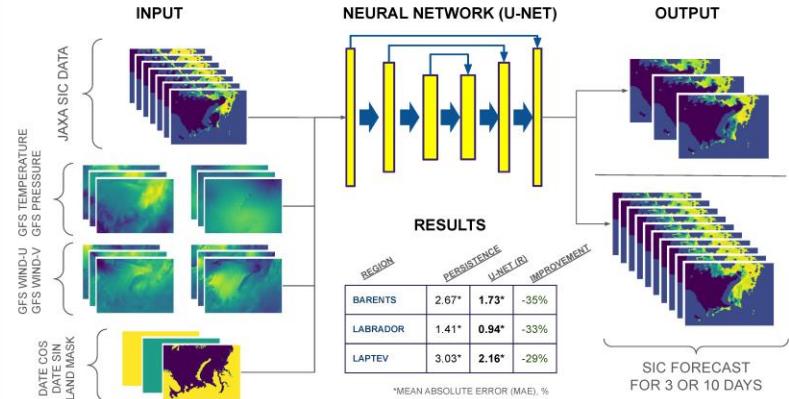
▼ Click to hide

```
git clone -q https://github.com/tom-andersson/icenet-paper.git notebook
```

<https://edsbook.org/notebooks/gallery/ac327c3a-5264-40a2-8c6e-1e8d7c4b37ef/notebook>

Andersson T. R. et al. Seasonal Arctic sea ice forecasting with probabilistic deep learning //Nature communications. – 2021. – Т. 12. – №. 1. – С. 5124.

Имеющиеся решения (3/3)



Grigoryev T. et al. Data-driven short-term daily operational sea ice regional forecasting //Remote Sensing. – 2022. – Т. 14. – №. 22. – С. 5837.

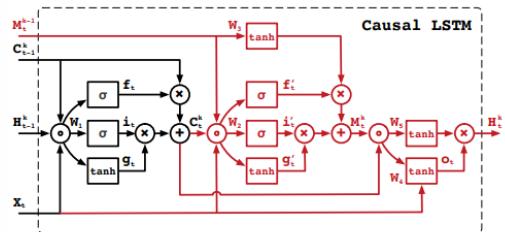
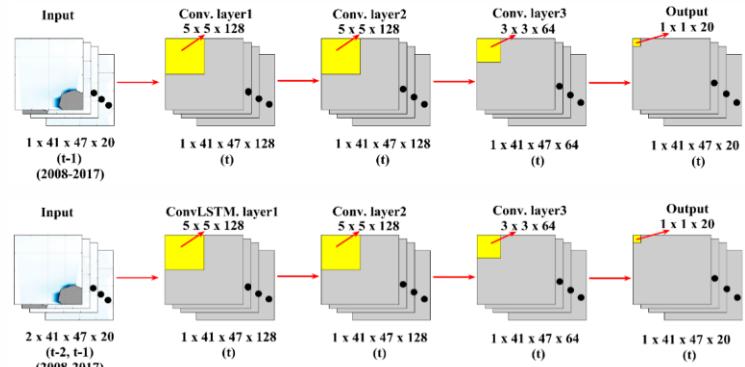
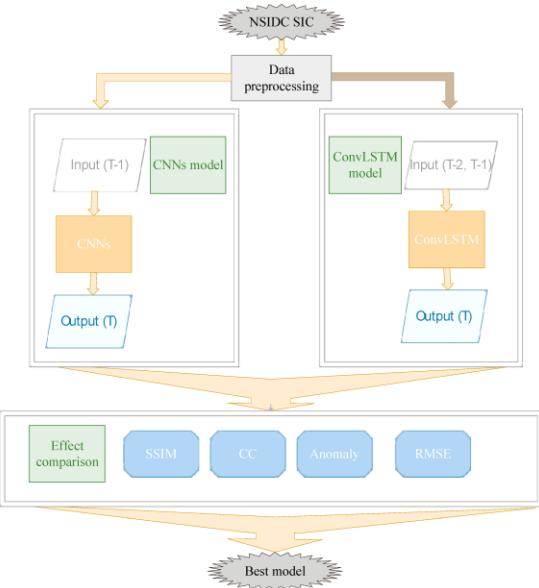


Figure 2. Causal LSTM, in which the temporal and spatial memories are connected in a cascaded way through gated structures. Colored parts are newly designed operations, concentric circles denote concatenation, and σ is the element-wise Sigmoid function.

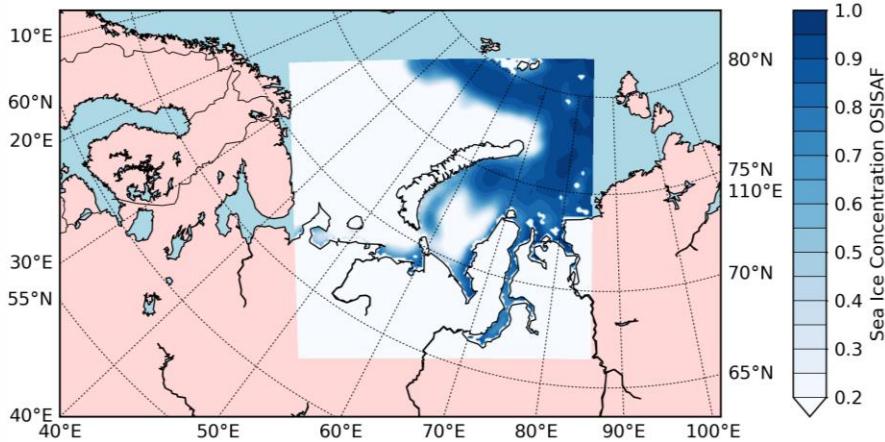
Wang Y. et al. Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning //International Conference on Machine Learning. – PMLR, 2018. – С. 5123-5132.



Liu Q. et al. Daily prediction of the arctic sea ice concentration using reanalysis data based on a convolutional LSTM network //Journal of Marine Science and Engineering. – 2021. – Т. 9. – №. 3. – С. 330.

Тестовые данные

Kara Sea - extent of zone of interest



- **Источники данных** - OSI SAF Global Sea Ice Concentration (SSMIS), предрассчитанная модель NEMO
- **Пространственное разрешение** – 14x14 км
- **Временное разрешение** – 1 день, 7 дней
- **Размер изображения** - 125x125 ячеек
- **Длина прогноза** – краткосрочный прогноз (60 дней), долгосрочный прогноз (1 год)

Моря российской Арктики:

- Навигация;
- Нефтедобыча на шельфе;
- Биоразнообразие.

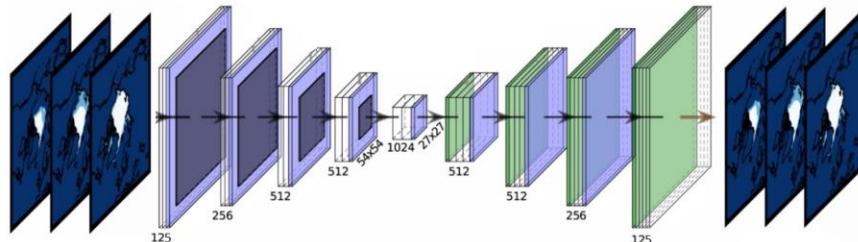


Базовый подход

ВХОД – многоканальное изображение:

- Предыстория параметра – n шагов

CNN структура



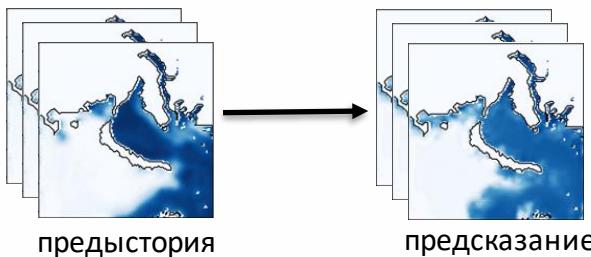
Функции потерь:

- 1) *L1Loss* (средняя абсолютная ошибка)
- 2) *SSIM* (structural similarity index measure – индекс структурной схожести)

ВЫХОД – многоканальное изображение:

- Предсказание параметра – k шагов вперед

Предсказание пространственно-временных данных



Вероятности наличия льда в каждый день периода становления льда(за всю выборку)

Описание метрик качества

$$\text{SSIM}(x, y) = \frac{(2\mu_x \cdot \mu_y + c_1)(2\delta_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\delta_x^2 + \delta_y^2 + c_2)},$$

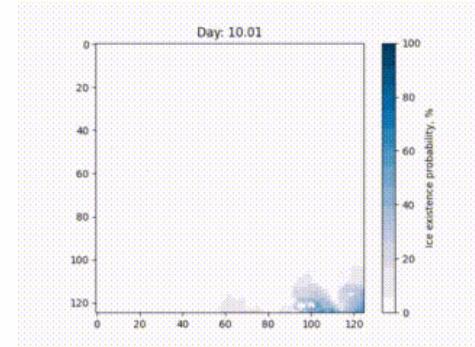
μ_x – среднее значение предсказания,

μ_y – среднее значение целевого параметра,

δ_x – среднеквадратичное отклонение предсказания,

δ_y – среднеквадратичное отклонение целевого параметра.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad \begin{aligned} y_i & \text{ – целевое значение,} \\ \hat{y}_i & \text{ – предсказанное значение,} \\ N & \text{ – число временных отсчетов} \end{aligned}$$



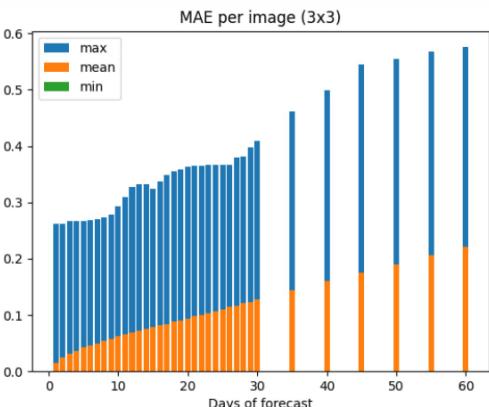
Краткосрочные прогнозы

Для краткосрочных прогнозов

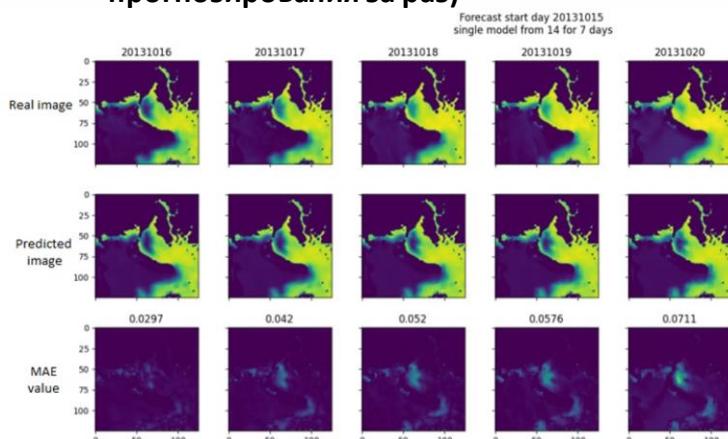
- С ростом горизонта падает качество;
- Качество падает сильнее при поочередном предсказании каждого шага.

Предыстория - 1 изображение

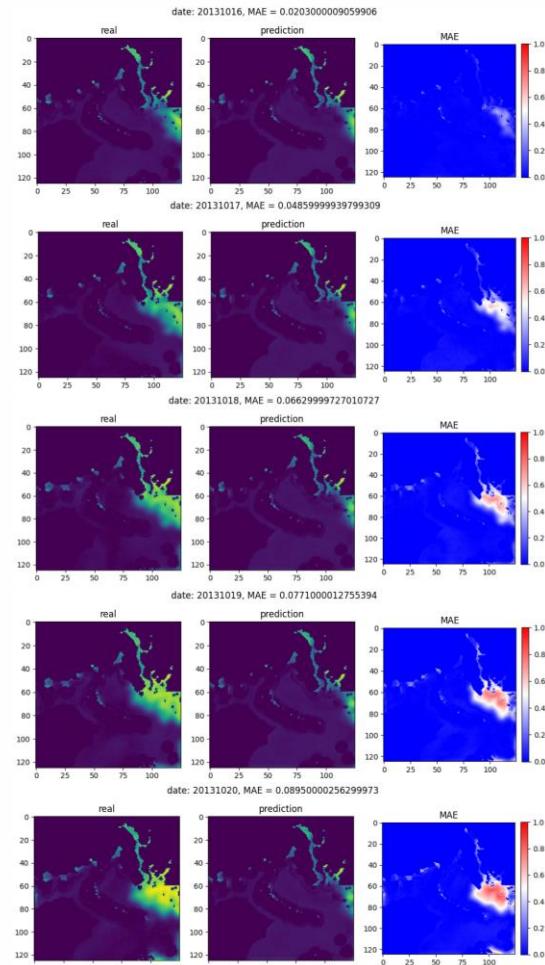
Выход - 1 изображение через N дней



In-sample prediction
(прогноз на весь горизонт прогнозирования за раз)

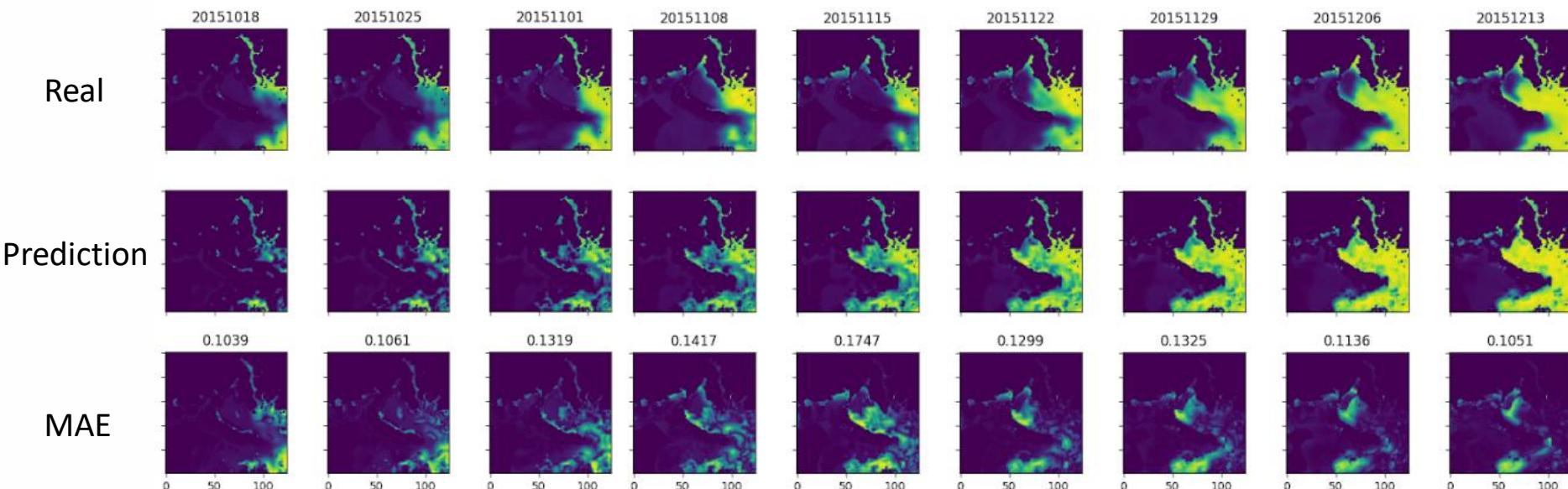


Out-of-sample
(каждый
последующий
прогноз базируется на
предыдущем)



Долгосрочные прогнозы (1/2)

- Временное разрешение приведено к 7 дням;
- Предыстория 104 шага (~ 2 года);
- Горизонт прогноза 52 шага (~ 1 год).

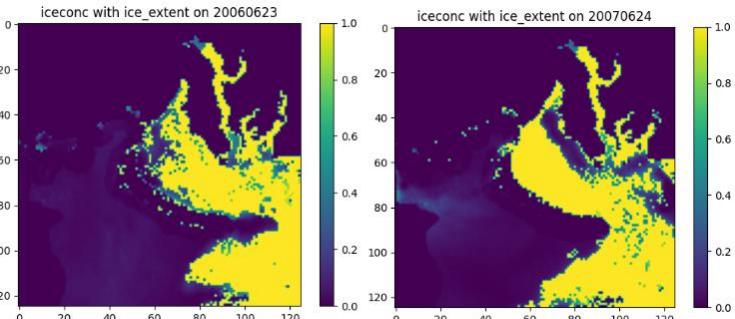
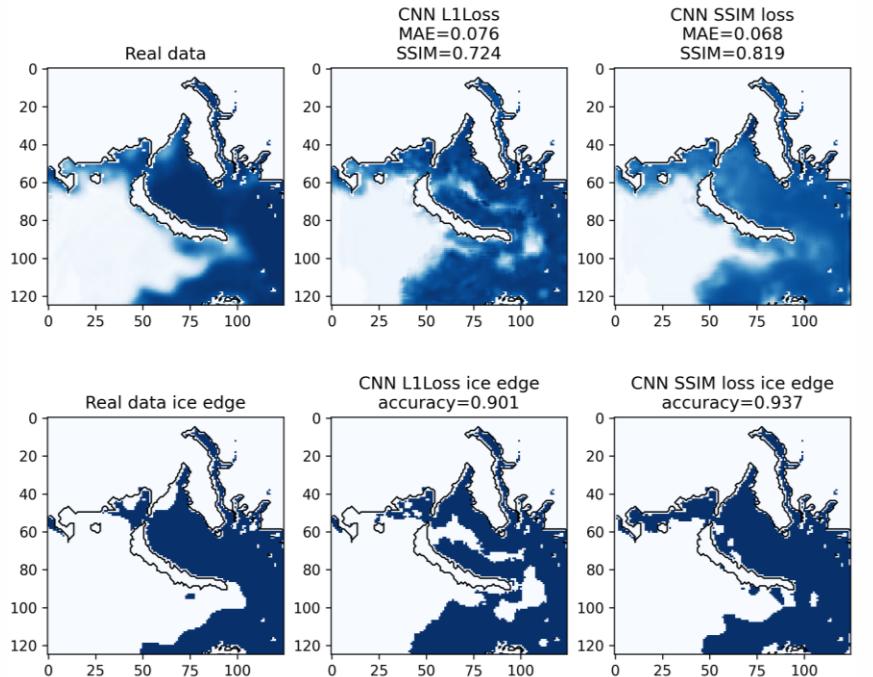


Долгосрочные прогнозы (2/2)

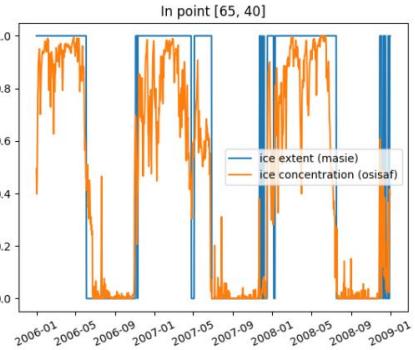
Постановка 1. Предсказание концентрации льда

Постановка 2. Восстановление кромки льда на основе прогноза концентрации (бинаризация)

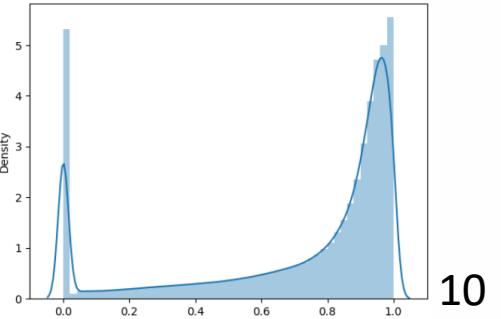
Prediction date - 20150226



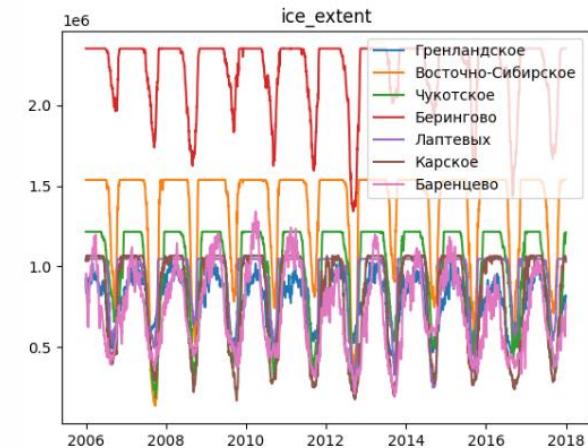
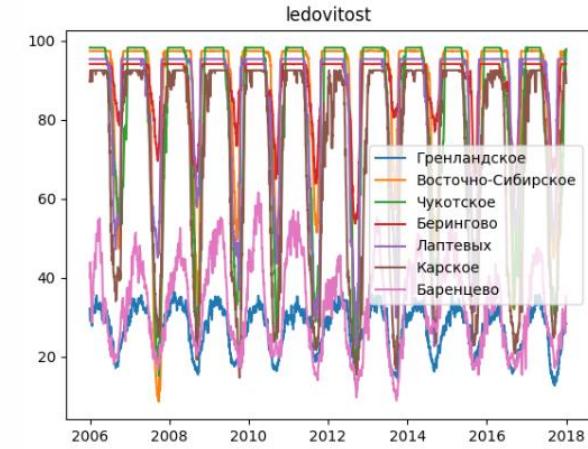
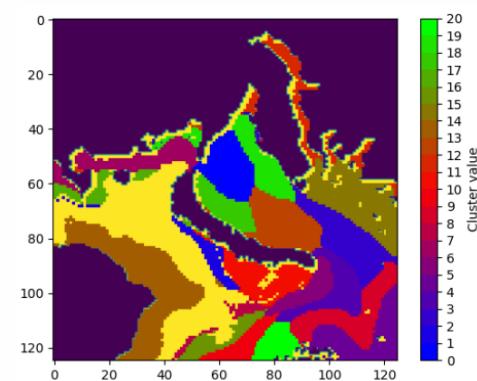
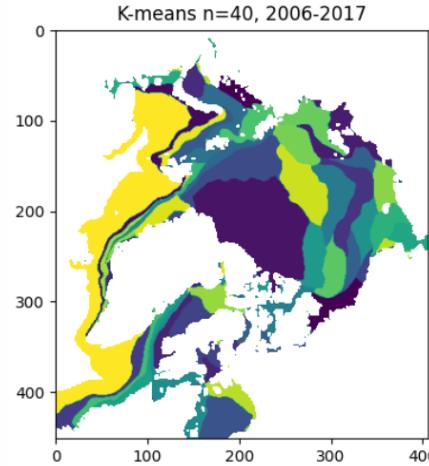
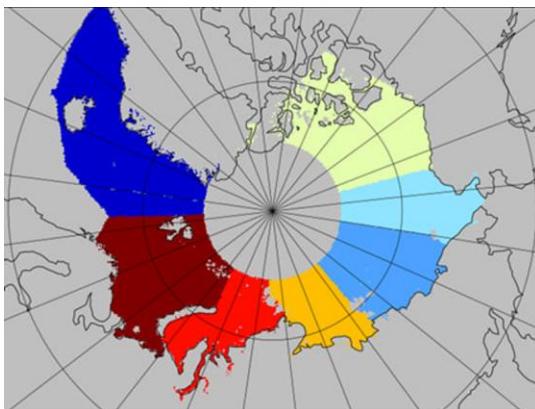
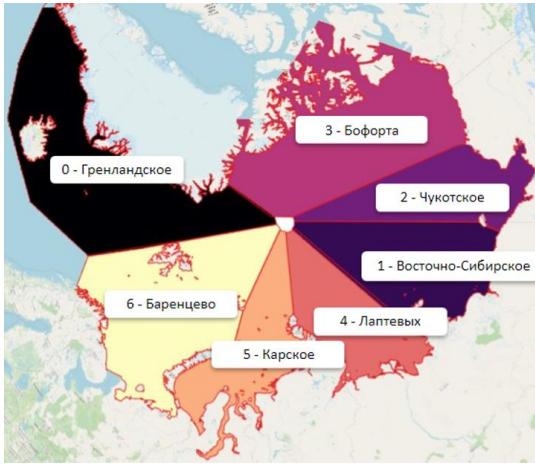
Бинарная кромка льда
(MASIE) наложена на
данные о
концентрации



За отсечку для
бинаризации принято
значение 0.8 (источники
рассматриваются)



Интегральные характеристики

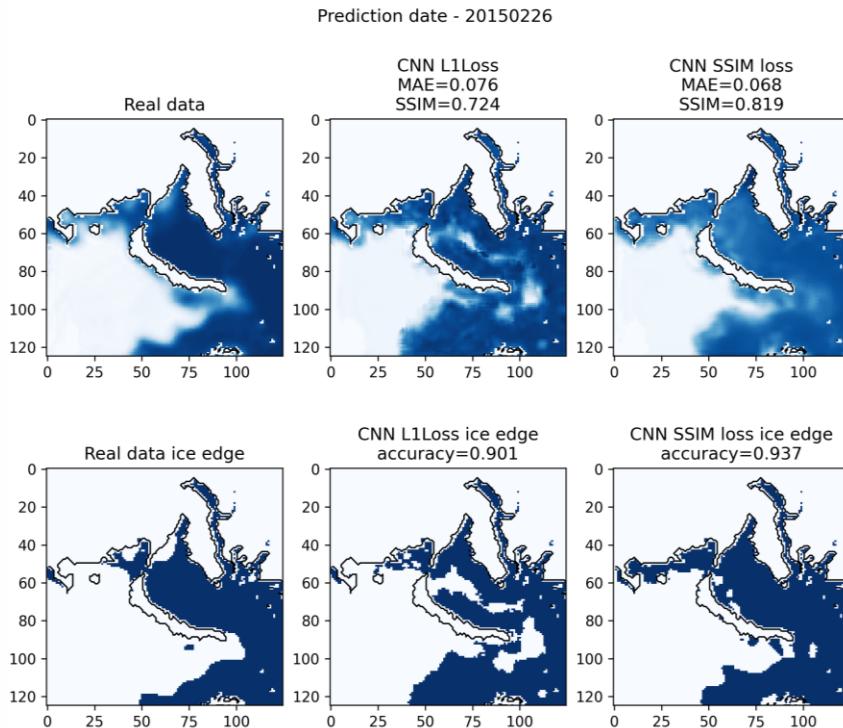


Примеры территориального деления на сектора для расчета интегральных характеристик

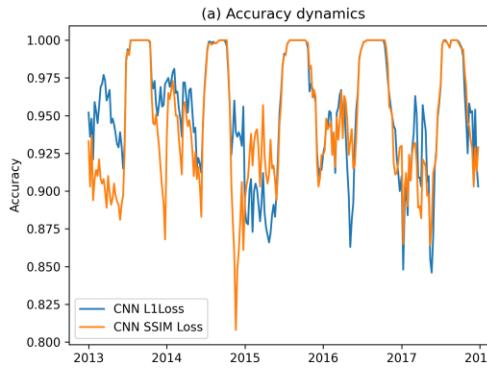
Интегральные характеристики

Постановка 1. Предсказание концентрации льда

Постановка 2. Восстановление кромки льда на основе прогноза концентрации (бинаризация)

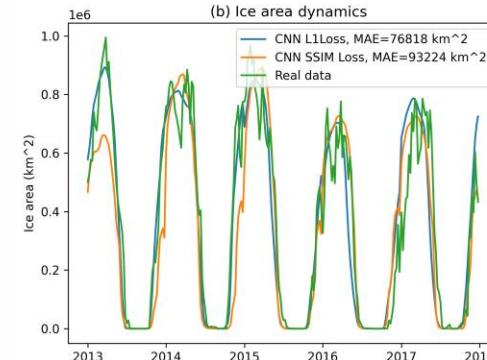


Интегральные характеристики



$$\text{Accuracy} = \frac{\text{len}(x = y)}{\text{len}(y)},$$

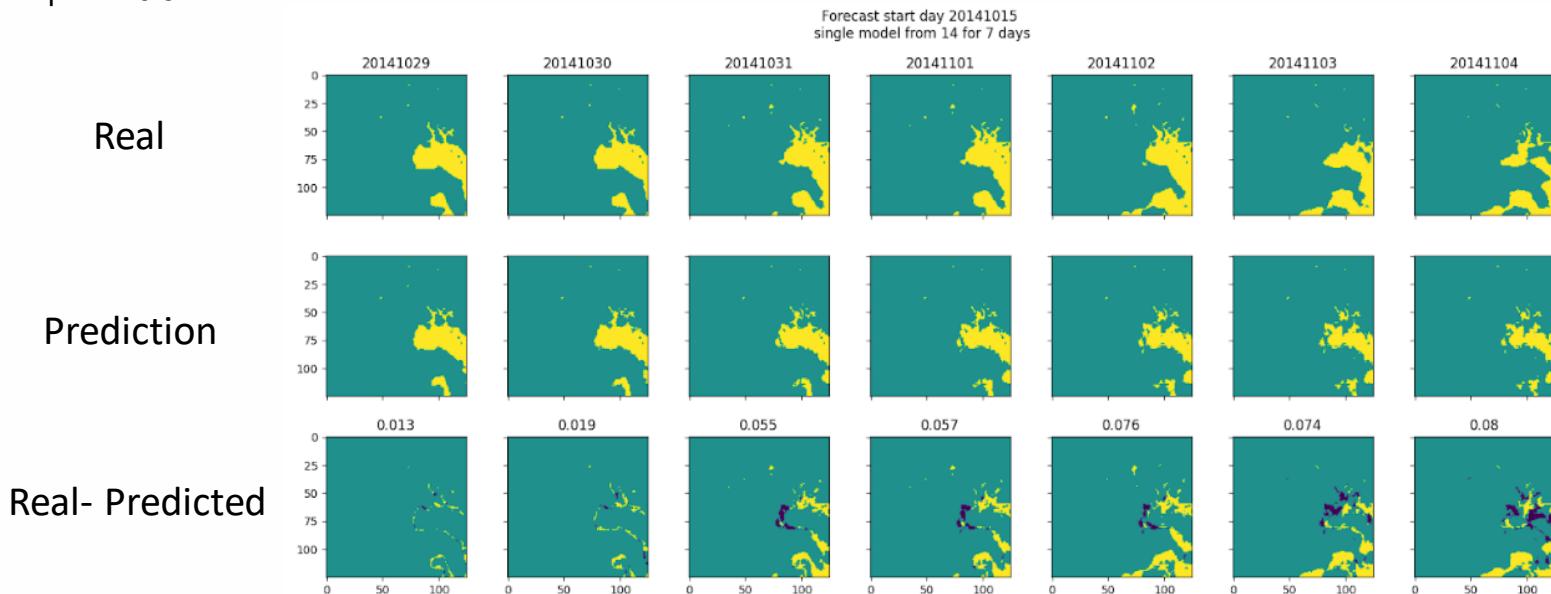
х – предсказанные значения,
у – целевые значения,
len – количество пикселей



Площадь была рассчитана с учетом пространственного разрешения 14 км на основе параметра кромки льда.

Краткосрочные прогнозы (бинарные) (1/2)

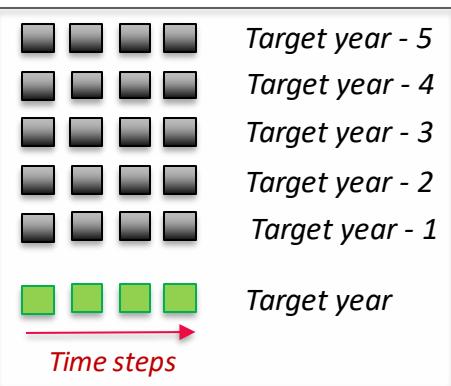
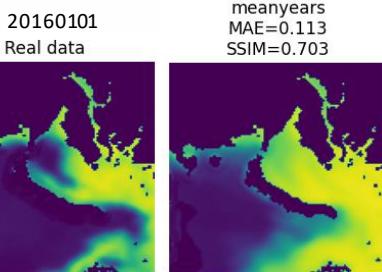
- **Входные данные** - OSISAF, переведенные в бинарную маску льда по порогу в 0.8
- В **структуре нейронной сети** функция активации для выходного слоя заменена на сигмоидальную (nn.Sigmoid)
- **Выход нейронной сети** - поле вероятностей отнесения писеля к классу 1, отсечение вероятности для разделения классов для расчета метрики - 0.5



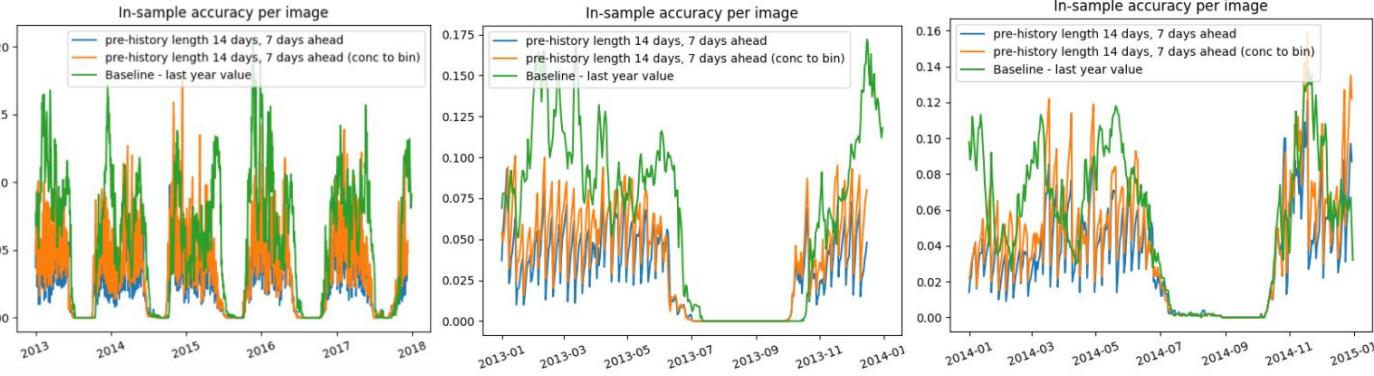
Метрики качества – бинарные прогнозы

*Baseline –
базовая модель –
инерционный прогноз –
среднегодичное*

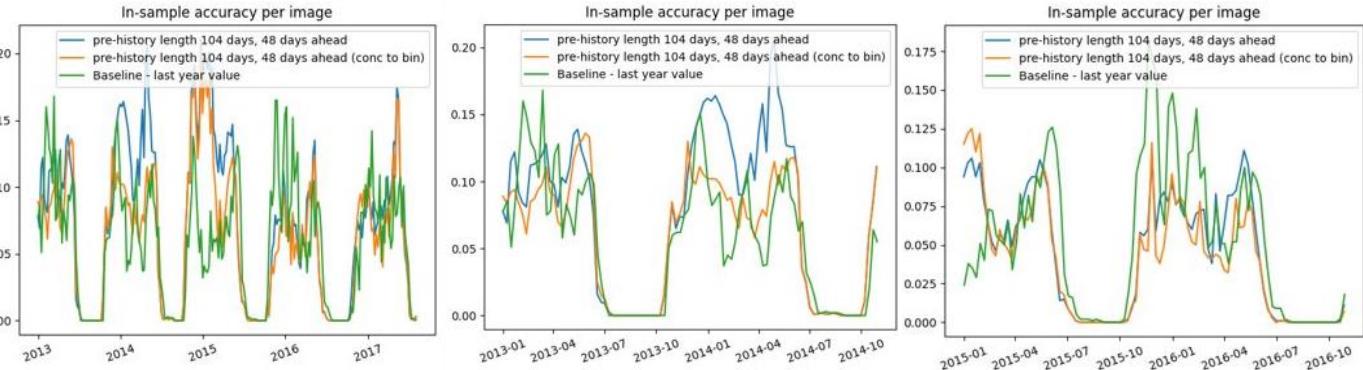
*На каждую дату рассчитывается
среднее изображение из 5
предыдущих лет*



Краткосрочные прогнозы



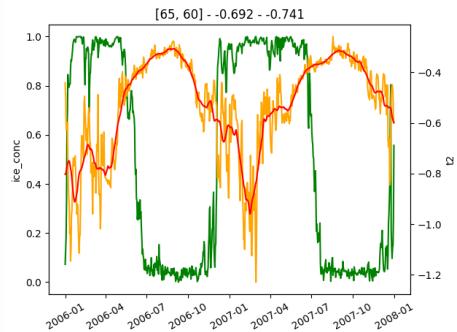
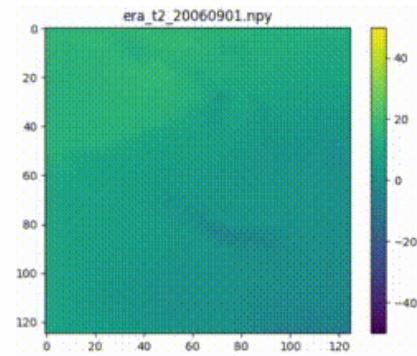
Долгосрочные прогнозы



*Обучение на бинарных данных с предсказанием бинарных данных дает лучшее качество,
чем обучение на концентрации с предсказанием бинарных данных.*

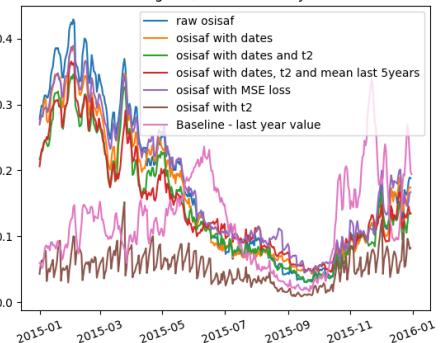
Дополнительные предикторы (1/2)

Температура воздуха
на 2м

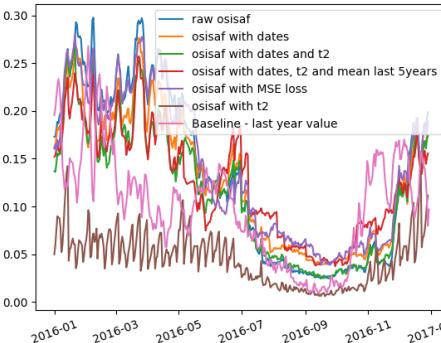


Краткосрочные прогнозы

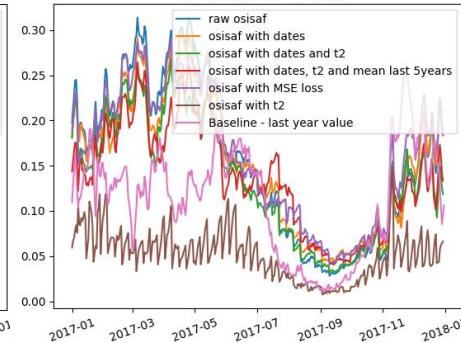
In-sample MAE per image
lag=14, forecast=7 days



In-sample MAE per image
lag=14, forecast=7 days

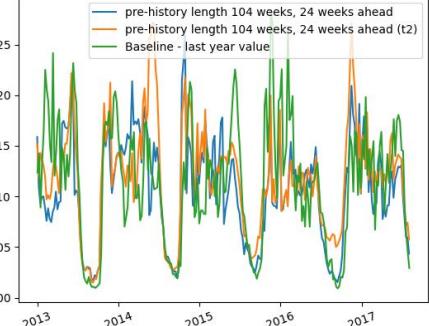


In-sample MAE per image
lag=14, forecast=7 days

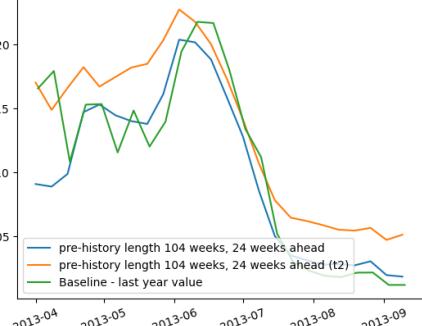


Долгосрочные прогнозы

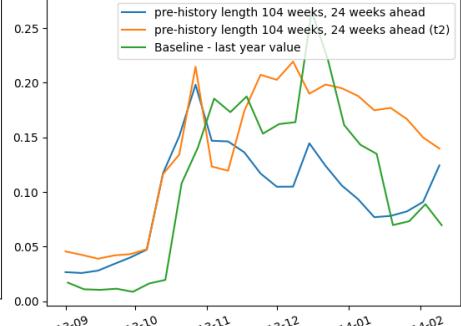
In-sample MAE per image



In-sample MAE per image



In-sample MAE per image

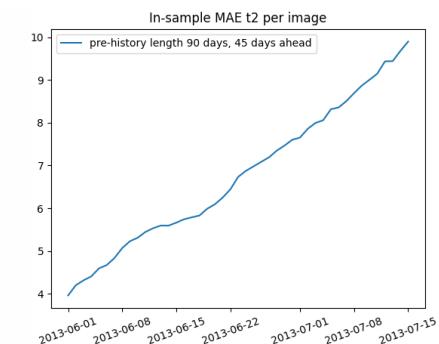
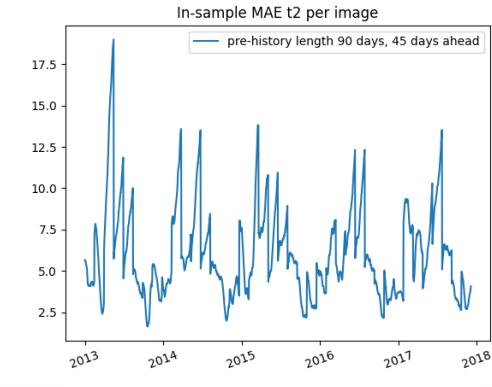
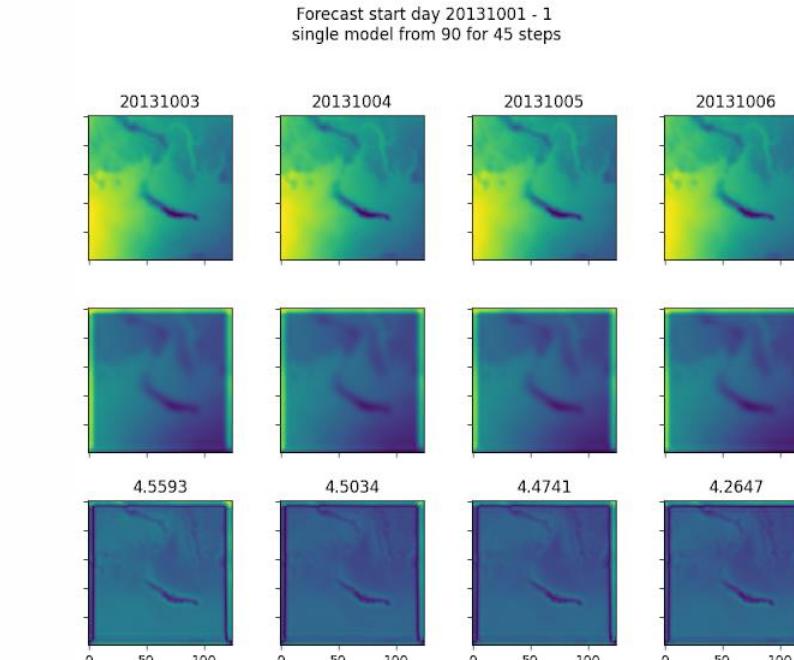
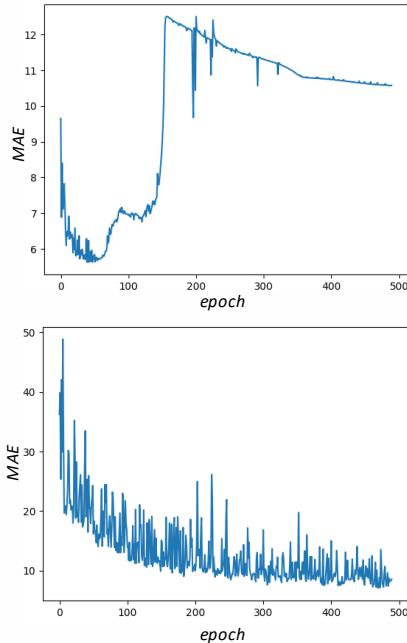


На краткосрочных прогнозах прирост качества присутствует, на долгосрочных нет.

Дополнительные предикторы (2/2)

Предсказание полей температуры как части метамодели

Графики сходимости нейронных сетей



- Специфичная динамика параметра не обеспечивает детерминированность запусков обучения.
- Модель сводится к воспроизведению предыдущего поля за невозможностью найти лучшие закономерности.
- На валидации видны артефакты несмотря на сходимость по метрике.

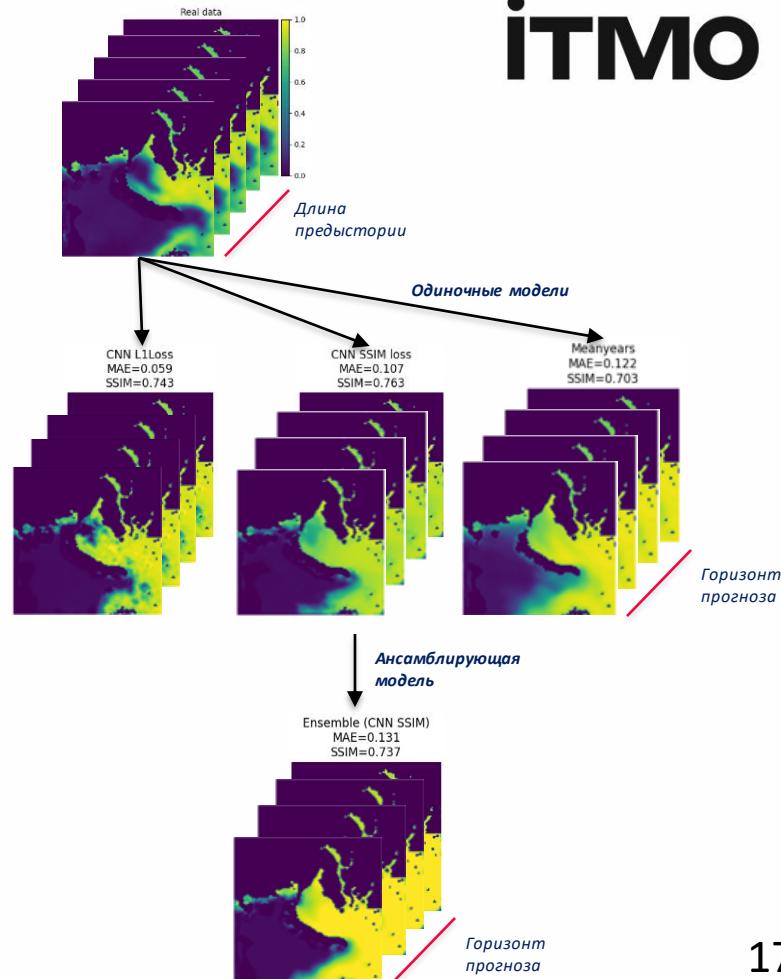
Ансамблевые методы

Испробованные виды ансамблей:

- Линейная регрессия (простой взвешенный ансамбль);
- Линейная регрессия для каждого дня в году (уникальные веса для каждого дня);
- Линейная регрессия для каждого пикселя изображения (уникальные веса для каждого пикселя);
- Сверточная нейронная сеть на основе SSIM функции потерь;
- Сверточная нейронная сеть на основе L1 функции потерь;
- Полносвязная нейронная сеть на основе L1 функции потерь.

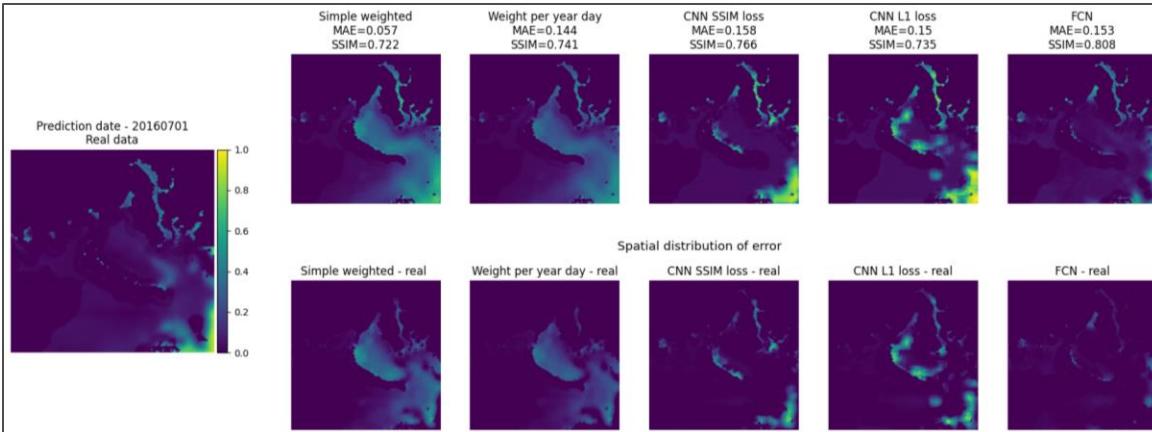
Элементы ансамбля:

- Сверточная нейронная сеть на основе L1 функции потерь;
- Сверточная нейронная сеть на основе SSIM функции потерь;
- Базовая модель (среднее за 5 лет).



Сравнение ансамблей (1/3)

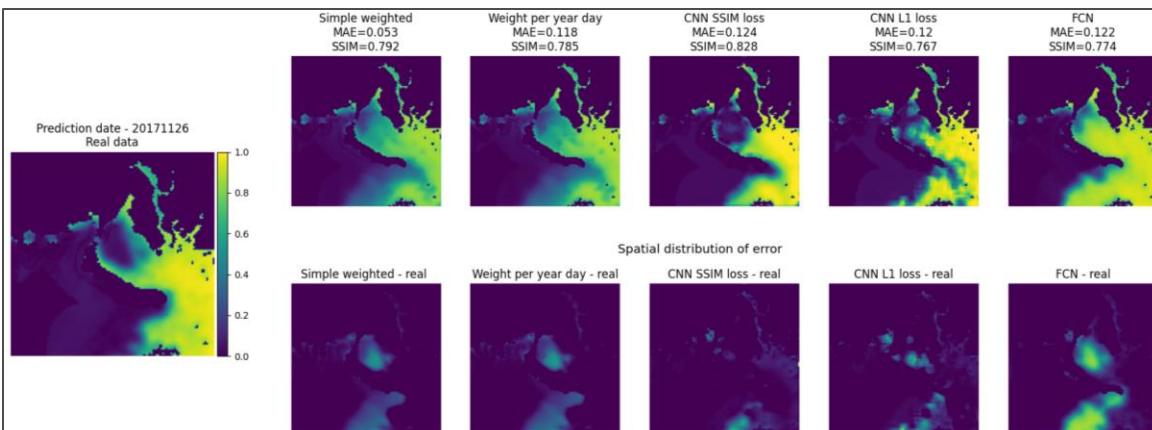
Примеры каждого из видов ансамблей по сравнению с реальными данными



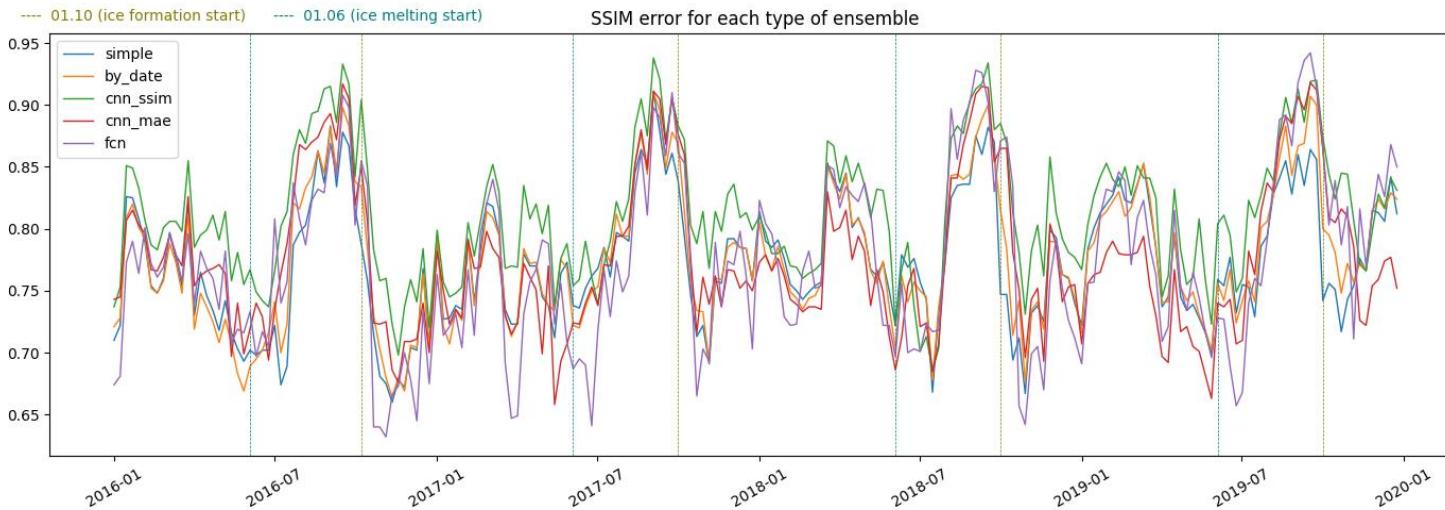
- Линейные ансамбли «размазывают» значения по всей картинке – метрика низкая, но пространственно картину не отражает.

- Полносвязная нейронная сеть отличается высокой ресурсоемкостью и обладает недостатками линейных ансамблей.

- Сверточная сеть с L1 функцией потерь визуально отличается сильной зернистостью и выглядит ненатурально.

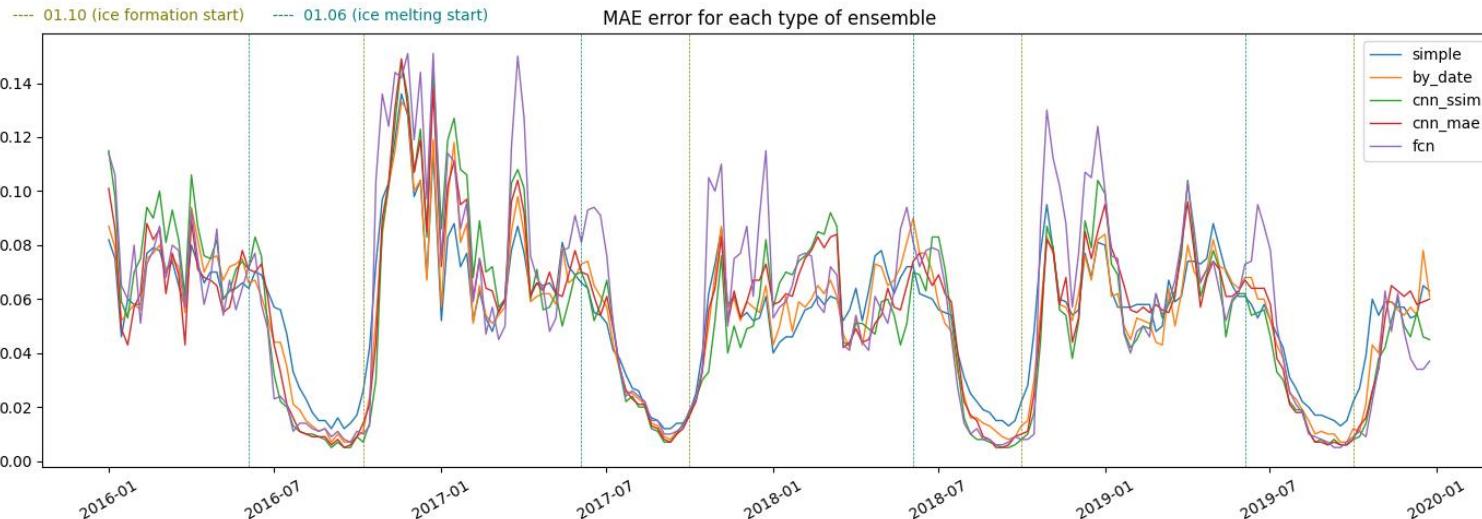


Сравнение ансамблей (2/3)



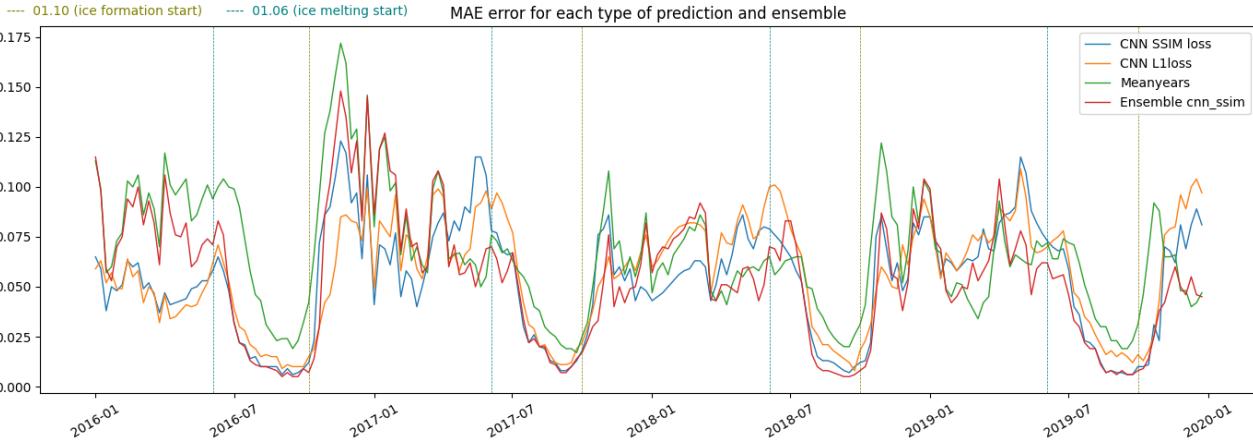
	ssim	2016	2017	2018	2019
single models	l1	0,726	0,701	0,690	0,689
	ssim	0,796	0,784	0,776	0,782
	mean	0,723	0,758	0,768	0,780
	simple	0,754	0,776	0,778	0,786
ensembles	by_date	0,758	0,777	0,779	0,793
	cnn_ssim	0,807	0,810	0,808	0,823
	cnn_mae	0,778	0,773	0,777	0,777
	fcn	0,758	0,761	0,781	0,794

Сравнение ансамблей (3/3)

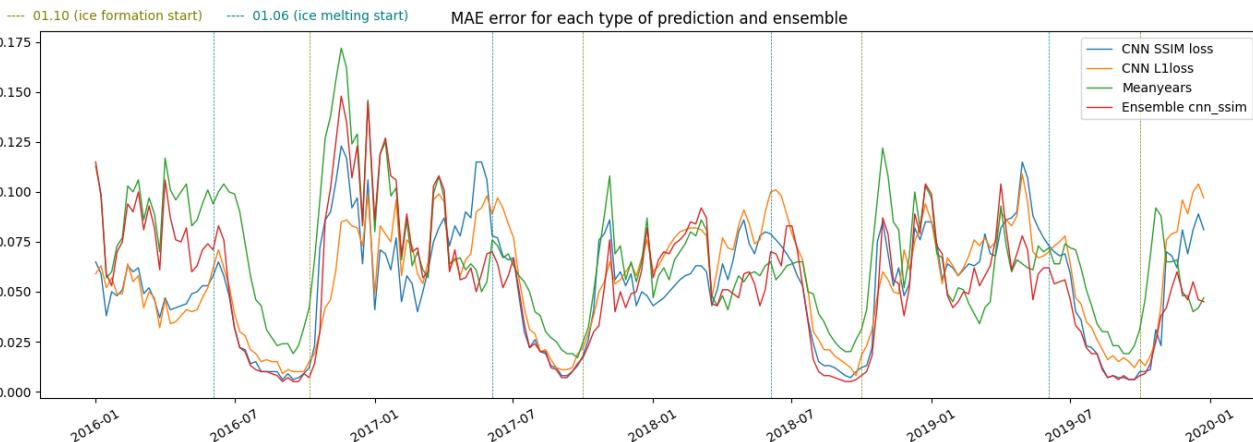


	mae	2016	2017	2018	2019
single models	l1	0,044	0,059	0,061	0,061
	ssim	0,048	0,056	0,052	0,056
	mean	0,086	0,064	0,060	0,055
ensembles	simple	0,062	0,053	0,052	0,051
	by_date	0,060	0,054	0,051	0,048
	cnn_ssimm	0,064	0,056	0,052	0,045
	cnn_mae	0,060	0,056	0,052	0,049
	fcn	0,065	0,065	0,057	0,048

Сравнение ансамблей с одиночными моделями (1/2)

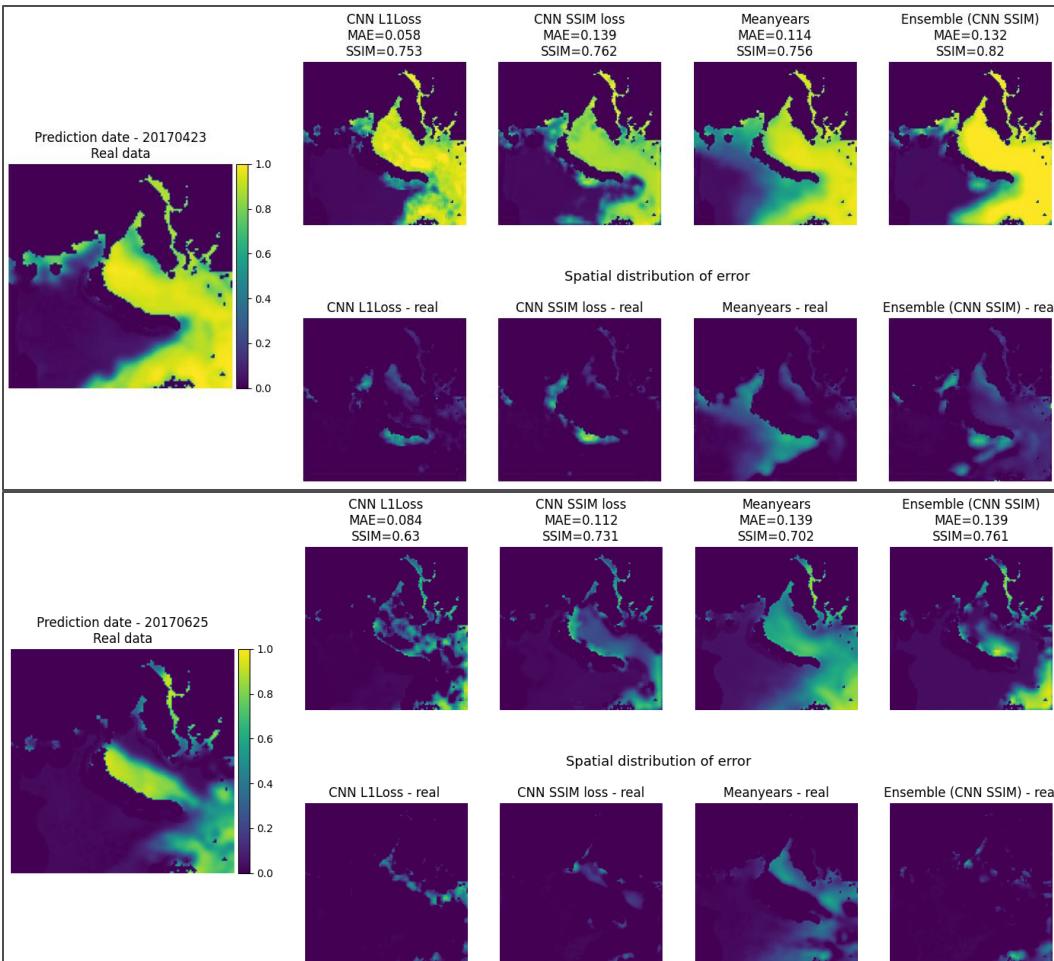


	2016	2017	2018	2019
ssim	0,726	0,701	0,690	0,689
l1	0,796	0,784	0,776	0,782
mean	0,723	0,758	0,768	0,780
cnn_ssim	0,807	0,810	0,808	0,823



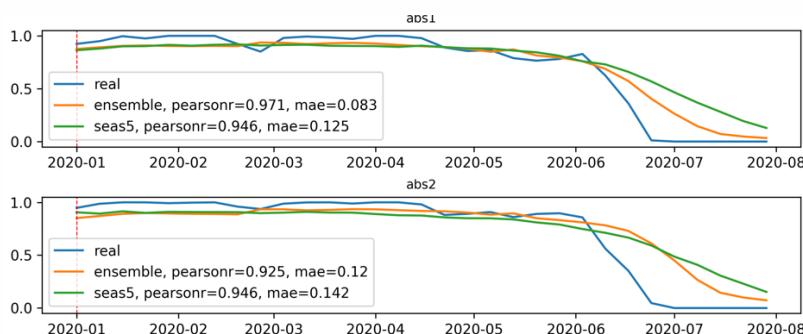
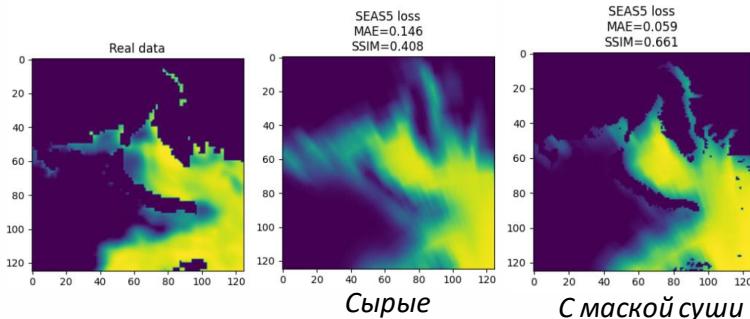
	2016	2017	2018	2019
mae	0,044	0,059	0,061	0,061
l1	0,048	0,056	0,052	0,056
ssim	0,086	0,064	0,060	0,055
cnn_ssim	0,064	0,056	0,052	0,045

Сравнение ансамблей с одиночными моделями (2/2)



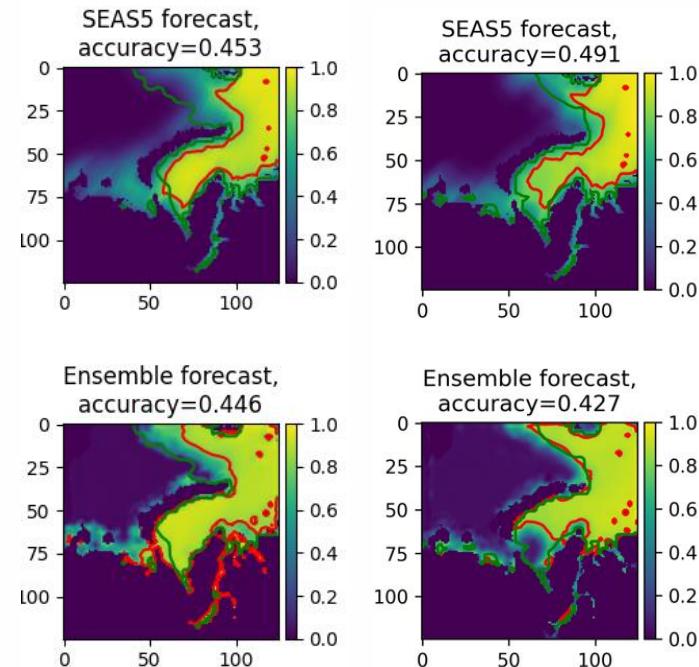
- *Лучшая ансамблевая модель не всегда показывает лучшее качество по сравнению с отдельно взятыми элементами ансамбля;*
- *Значимую роль играет разбиение на обучающие выборки для основных моделей и ансамбля, а также их дообучение.*

Сравнение с аналогами – SEAS5



Валидация в тестовых точках показывает лучшее качества ансамбля по сравнению с SEAS5.

Красный – предсказанная ледовая кромка;
Зеленый – фактическая ледовая кромка.



Проблема – используемые метрики (SSIM, MAE, accuracy)
не отражают фактическое качество прогноза,
оцениваемое экспертом.

- Реализована облегченная предсказательная модель на основе сверточных нейронных сетей;
- Реализованная структура модели позволяет осуществлять как долгосрочное так и краткосрочное прогнозирование;
- Применение нескольких функций потерь позволяет как воспроизводить пространственное распределение параметра внутри исследуемой территории (*SSIM*), так и его абсолютные значения;
- Произведена оценка качества ансамблевых моделей различной структуры, лучшее качество по двум метрикам (*SSIM* и *MAE*) достигнуто при использовании в качестве объединяющей модели сверточной нейронной сети с *SSIM* функцией потерь;
- Введение функции потерь, более конкретно отражающей положение кромки, позволит как сделать оценку и сравнение прогнозов более корректными, так возможно и улучшить качество прогнозирования бинарных данных за счет внесения модифицированной функции в расчет *Loss* прогностической модели.

Спасибо за внимание!

ITMO *more than a*
UNIVERSITY