



Фонд
интеллект

Аппроксимация пространственно-временной изменчивости городского острова тепла Москвы методами машинного обучения

Варенцов М.И.^{1-4*}, Криницкий М.А.^{4,5,1}, Степаненко В.М.^{1,2}

(1) МГУ имени М.В. Ломоносова, Научно-исследовательский вычислительный центр

(2) МГУ имени М.В. Ломоносова, Географический факультет

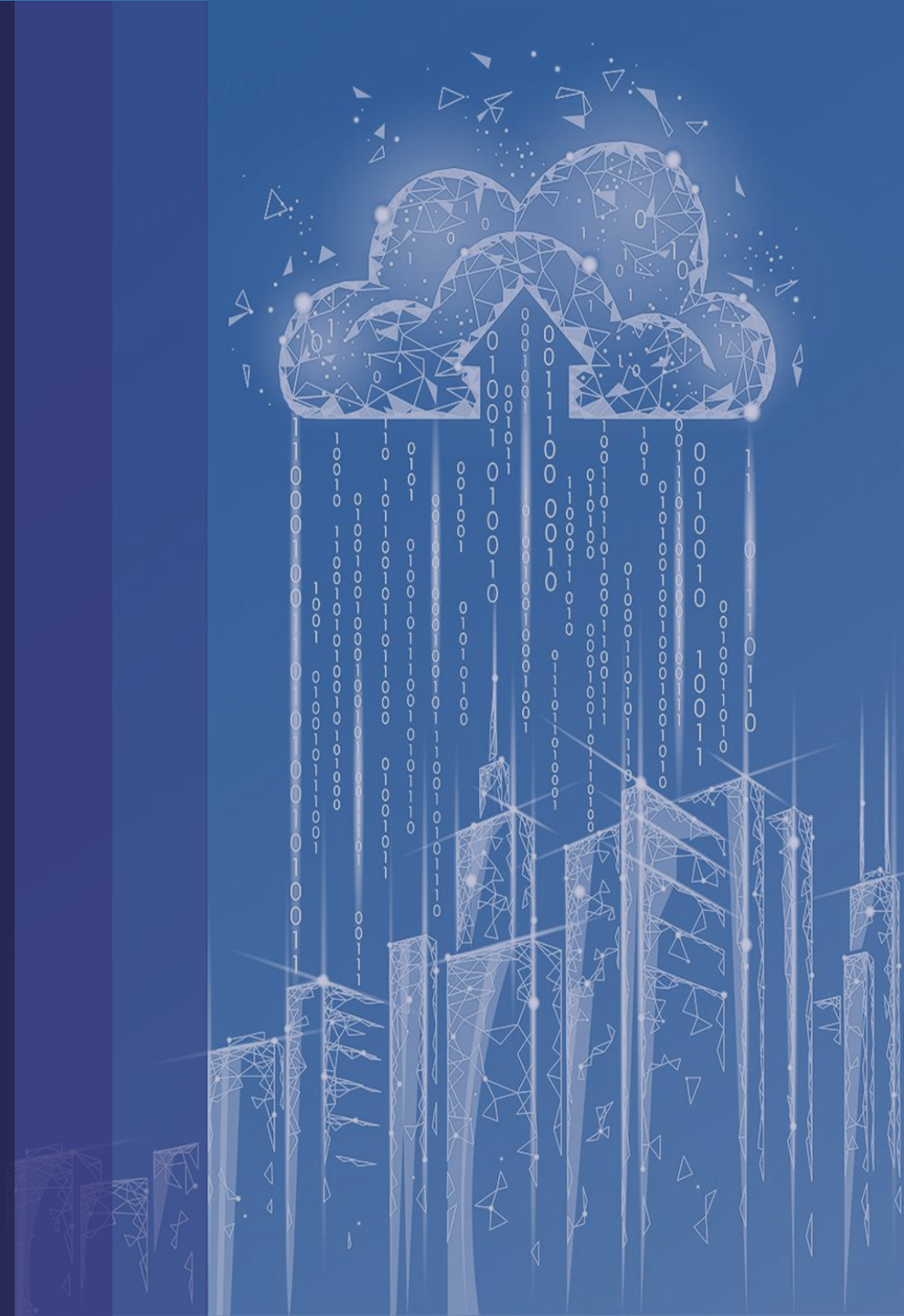
(3) Гидрометцентр России

(4) Институт физики атмосферы имени А.М. Обухова РАН,

(5) Институт океанологии им. П.П. Ширшова РАН

(6) Московский физико-технический институт

*mikhail.varentsov@srcc.msu.ru



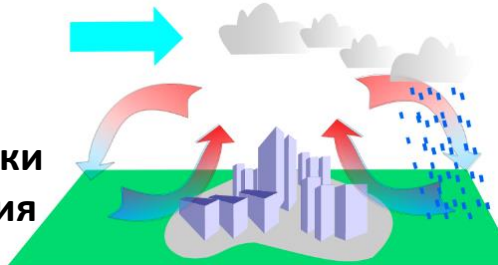
Городская метеорология?

Актуальность

Проблемы и вызовы для науки



Влияние на облачность, осадки и опасные явления погоды

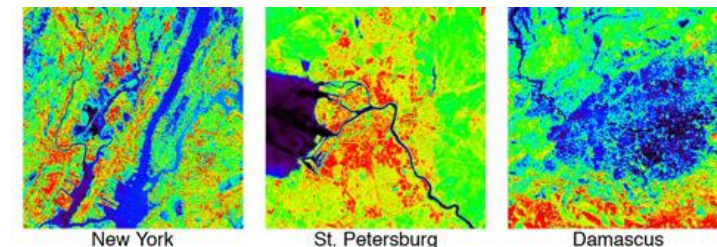
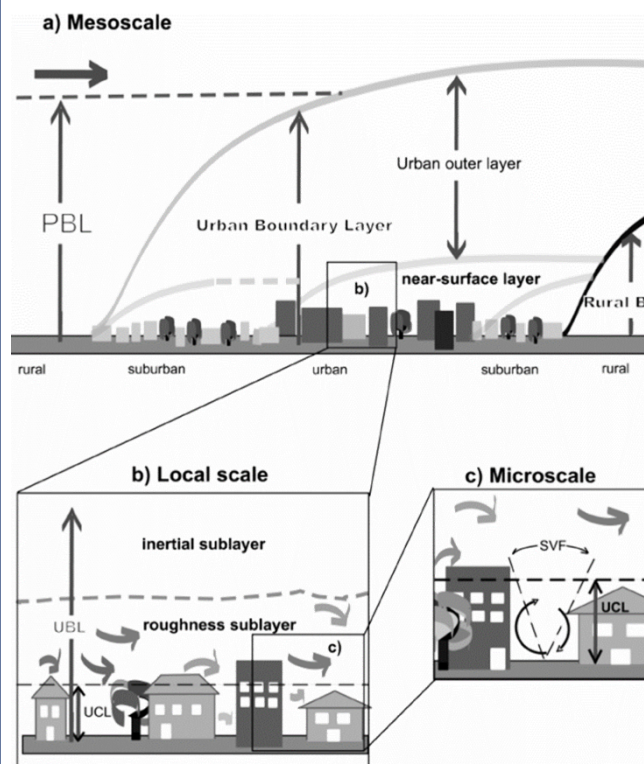


Влияние на термический комфорт населения



Сложность процессов

Пространственная неоднородность



Дефицит данных наблюдений



ИИ в метеорологии

Conservation of momentum:

$$\frac{\partial \vec{V}}{\partial t} = -(\vec{V} \cdot \nabla) \vec{V} - \frac{1}{\rho} \nabla p - \vec{g} - 2\vec{\Omega} \times \vec{V} + \nabla \cdot (\kappa \nabla \vec{V}) - \vec{F}_d$$

Conservation of energy:

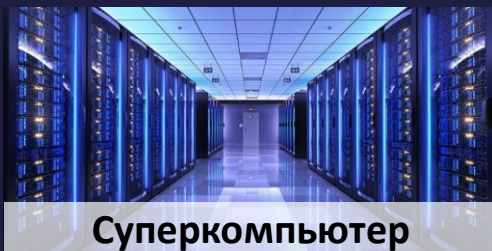
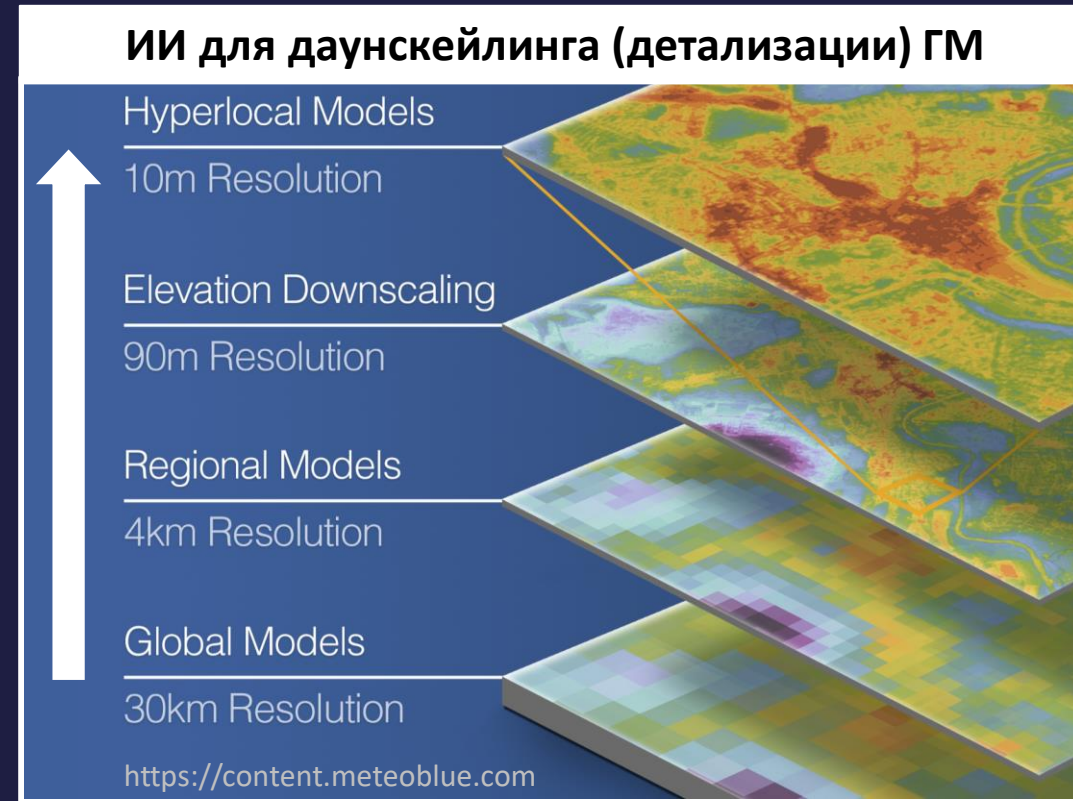
$$\rho c_p \frac{\partial T}{\partial t} = -\rho c_p (\vec{V} \cdot \nabla) T - \nabla \cdot \vec{R} + \nabla \cdot (\kappa \nabla T) + C + S$$

Conservation of mass:

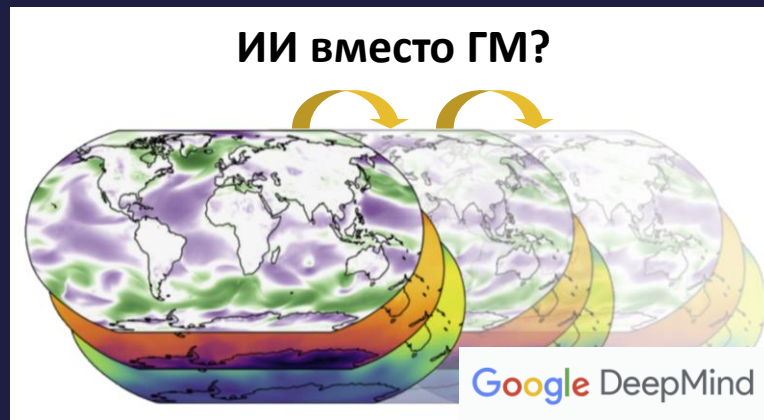
$$\frac{\partial \rho}{\partial t} = -(\vec{V} \cdot \nabla) \rho - \rho (\nabla \cdot \vec{V})$$

$$\frac{df_{1.5}}{dx} \approx \frac{f_2 - f_1}{\Delta x}$$

Гидродинамическая модель атмосферы (ГМ)



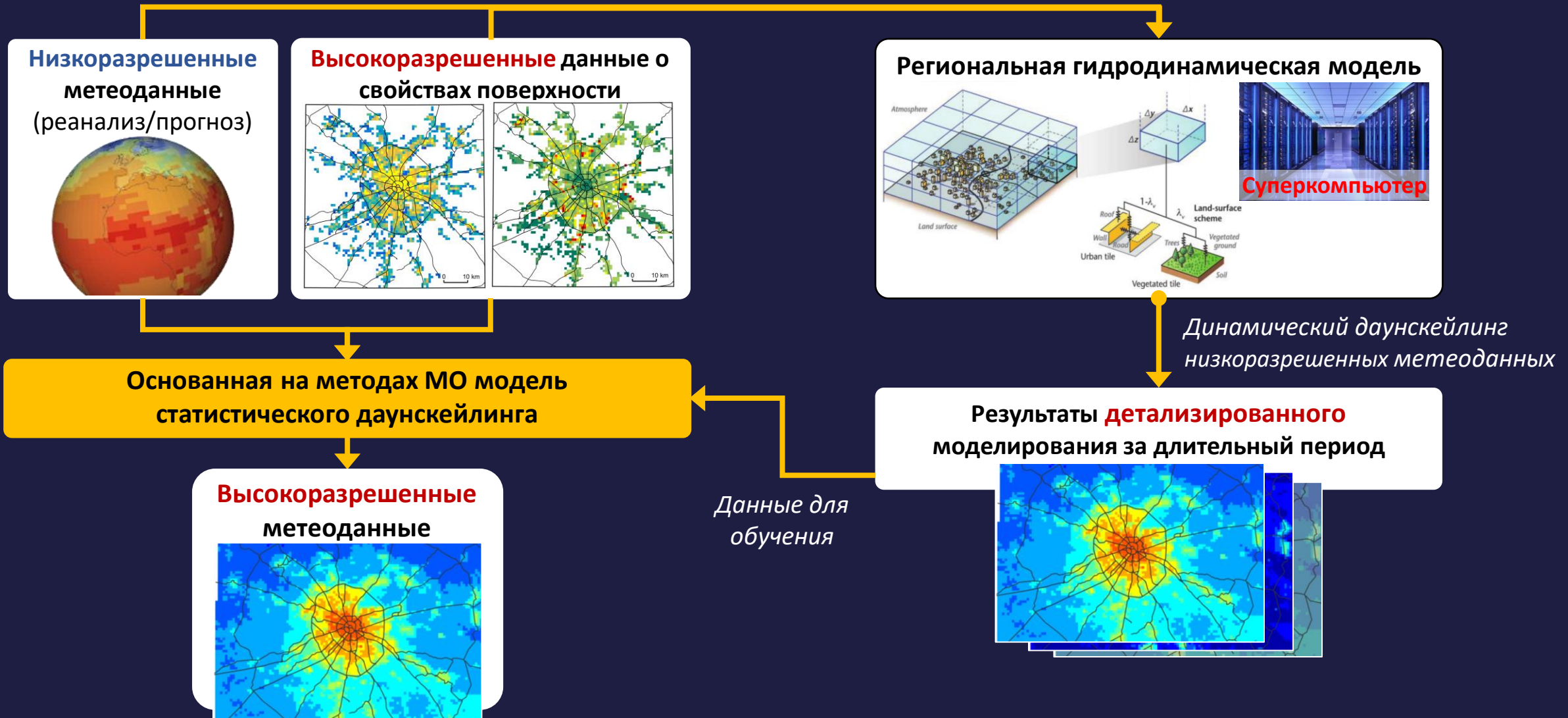
Суперкомпьютер



Прогнозы погоды
и климата

Про это наш доклад

Идея исследования



Этап №1

Аппроксимация временной динамики острова тепла в точке

Целевая переменная:

Разность температуры между центром Москвы (Балчуг) и средней температурой по 9 загородным метеостанциям, $\Delta T = T_{urb} - T_{rur}$

Формализация задачи МО:

$$\Delta T_i = f(x_{1,i}, x_{1,i-1}, \dots, x_{1,i-m}, \dots, x_{n,i}, x_{n,i-1}, \dots, x_{n,i-m})$$

Данные за предшествующие моменты времени

Предикторы:

- характеристики крупномасштабных атмосферных в Московском регионе по данным наблюдений на загородных метеостанциях и реанализа
- признаки, характеризующие временную связность рядов (тенденции и скользящие средние)
- высота солнца, время суток и года

Исходные данные:

Временные ряды с дискретность 3 ч за 2001-2021 гг.

- 5-блоковая кросс-валидация с отношением тренировочной и тестовой выборок 4:1

| Метеорологические предикторы | |
|------------------------------|---|
| t2m | Температура воздуха на 2 м, °C |
| rh2m | Отн. влажность на 2 м, % |
| vel10m | Скорость ветра на 10 м, м/с |
| tcc | Доля общей облачности |
| lcc | Доля нижней облачности |
| WF | Фактор погоды, эмпирическая функция скорости ветра и облачности [Oke, 1998] |
| sp | Атмосферное давление, гПа |
| blh | Высота погранслоя, м |
| str | Длинноволновый баланс, Вт/м ² |
| ssr | Коротковолновый баланс, Вт/м ² |
| strd | Приходящая коротковолновая радиация, Вт/м ² |
| ssrd | Приходящая длинноволновая радиация, Вт/м ² |
| tp | Сумма осадков за 3 ч, мм |

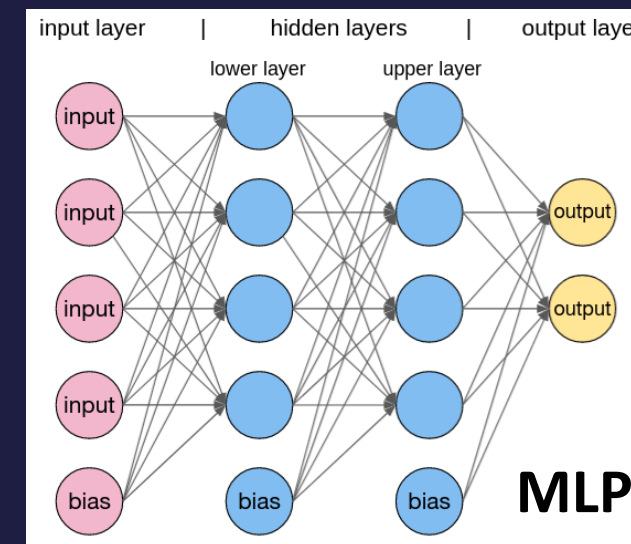
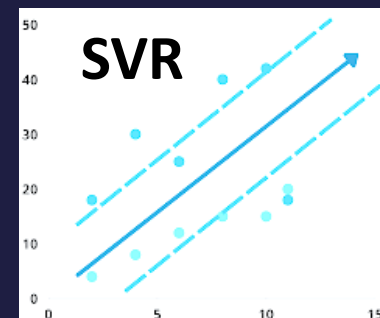
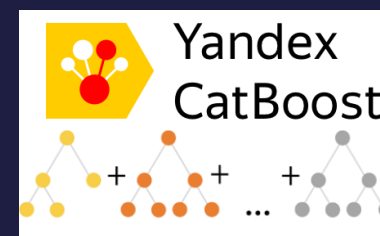
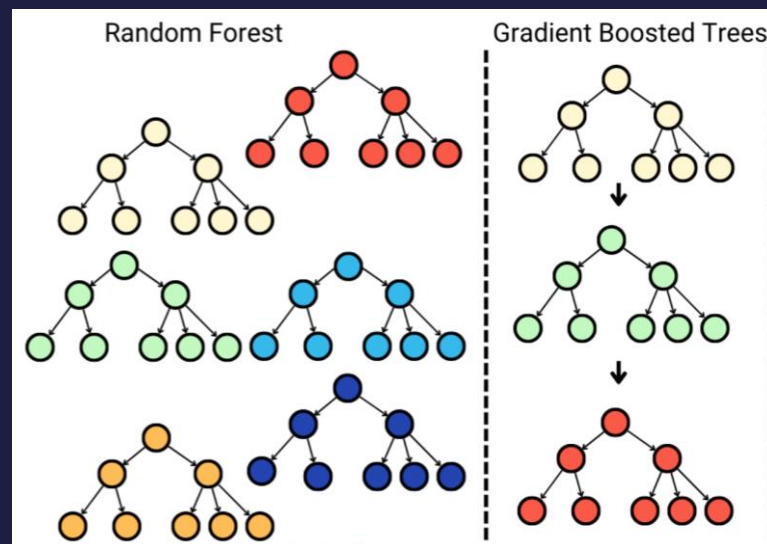
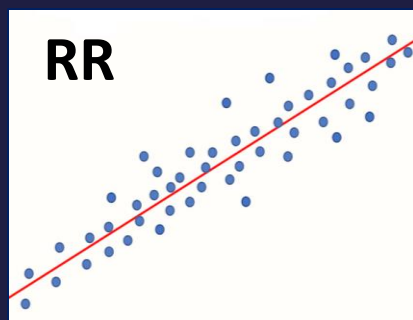
Наблюдения
и реанализ

Только реанализ



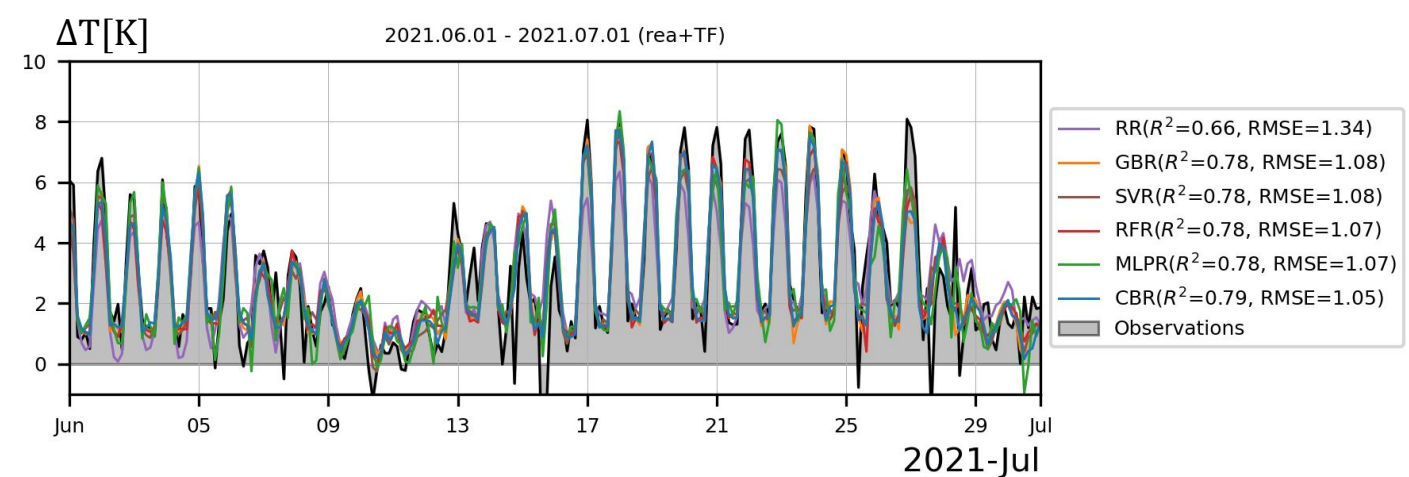
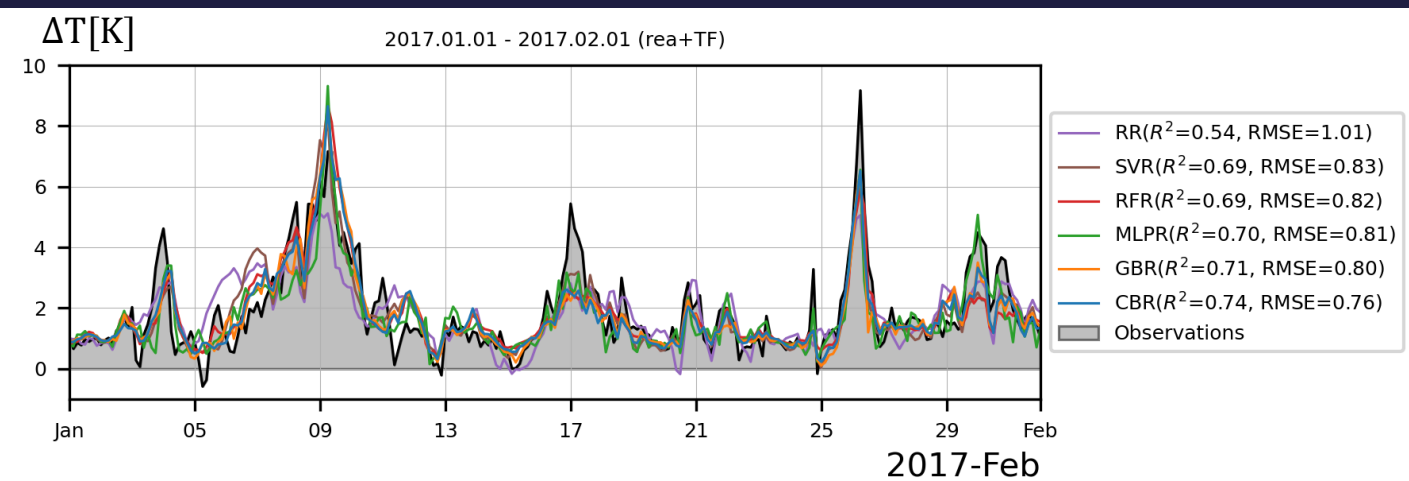
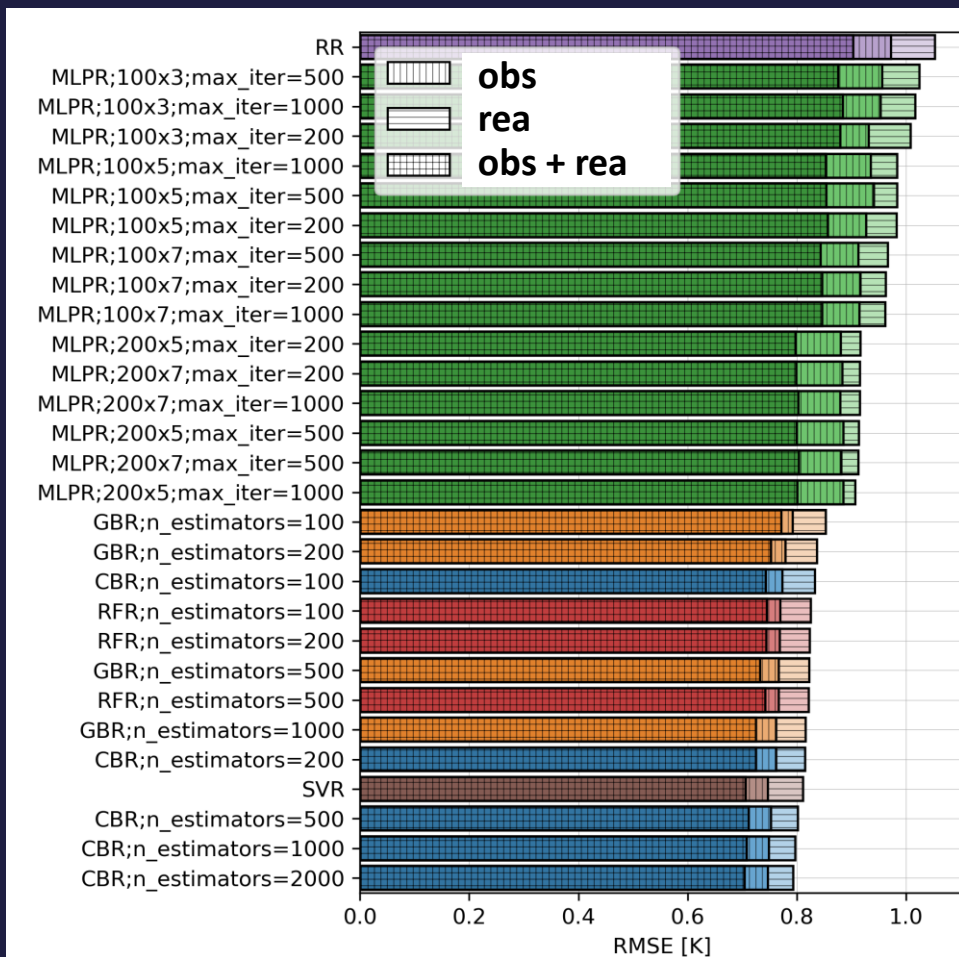
Сравнение моделей МО

| Модель | Сокращение | Настраиваемые гиперпараметры |
|----------------------------------|------------|------------------------------|
| Ridge Regression, baseline | RR | - |
| Random Forest Regression | RFR | n_estimators |
| Gradient Boosting Regression | RBR | n_estimators |
| Cat Boost Regression | CBR | n_estimators |
| Multi-layer Perceptron regressor | MLPR | hidden_layer_sizes, max_iter |
| Support Vector Regression | SVR | - |



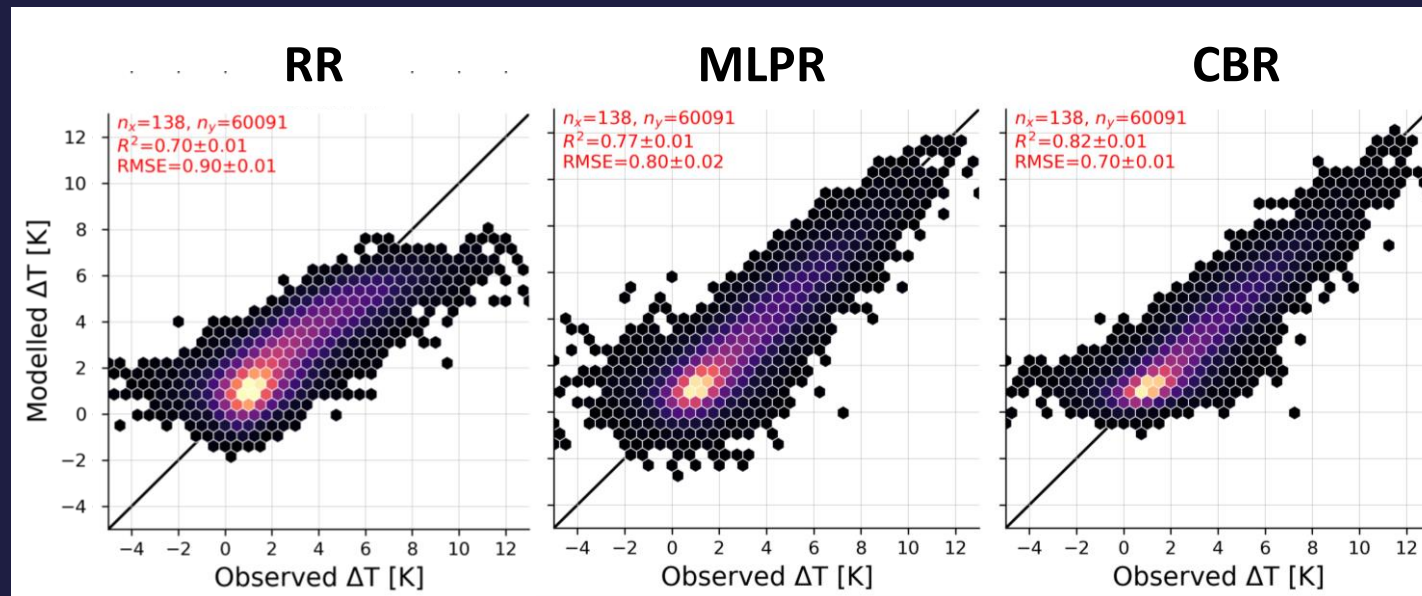
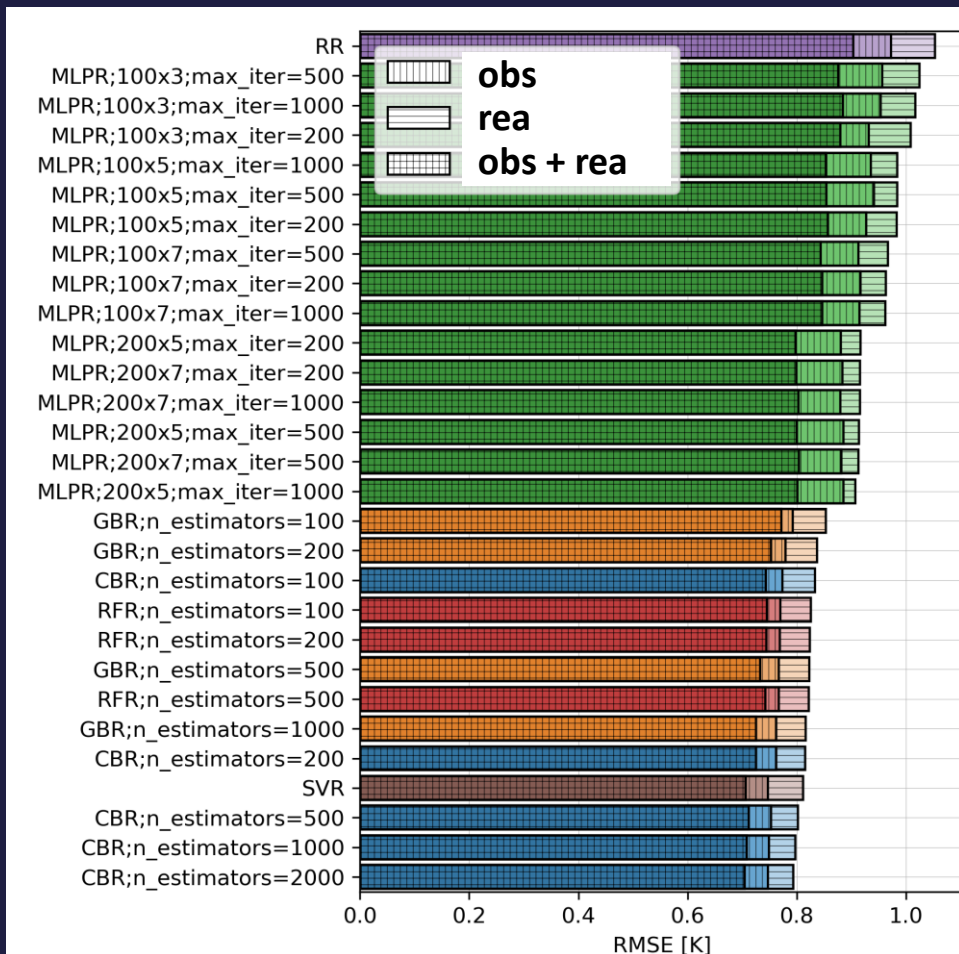


Сравнение моделей МО





Сравнение моделей МО



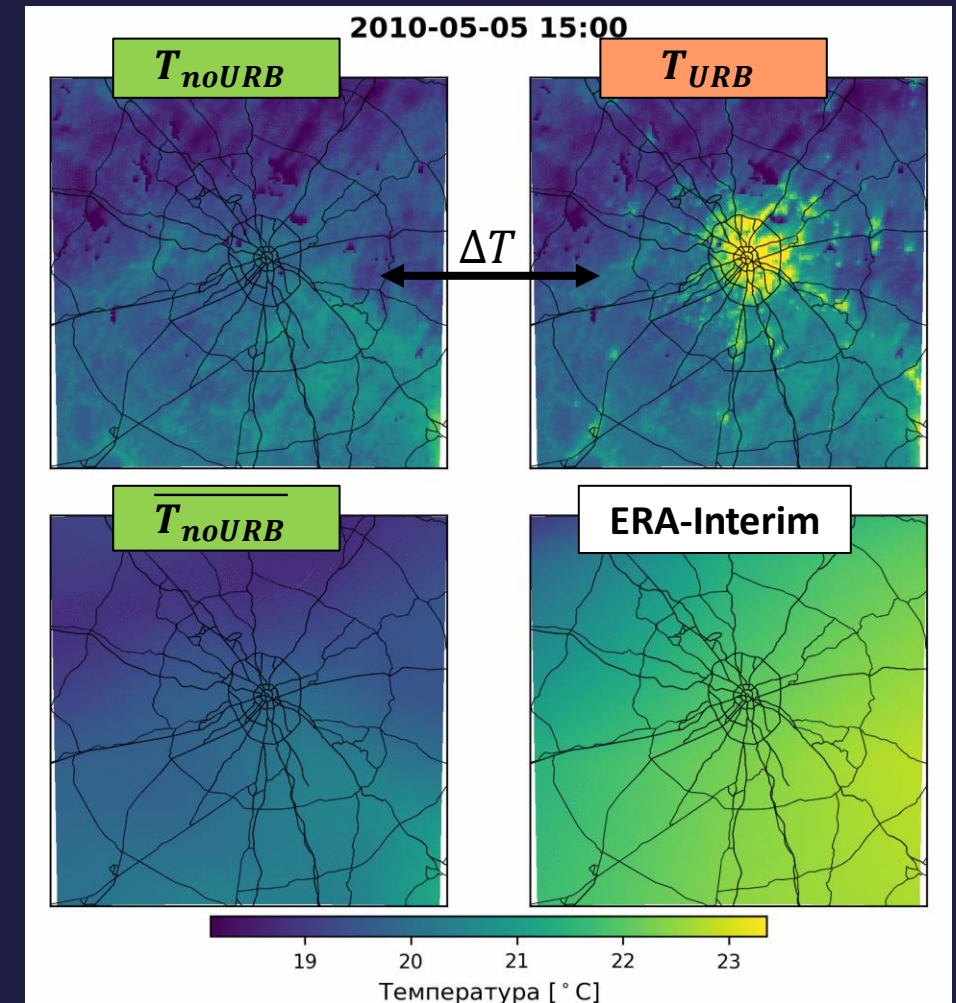
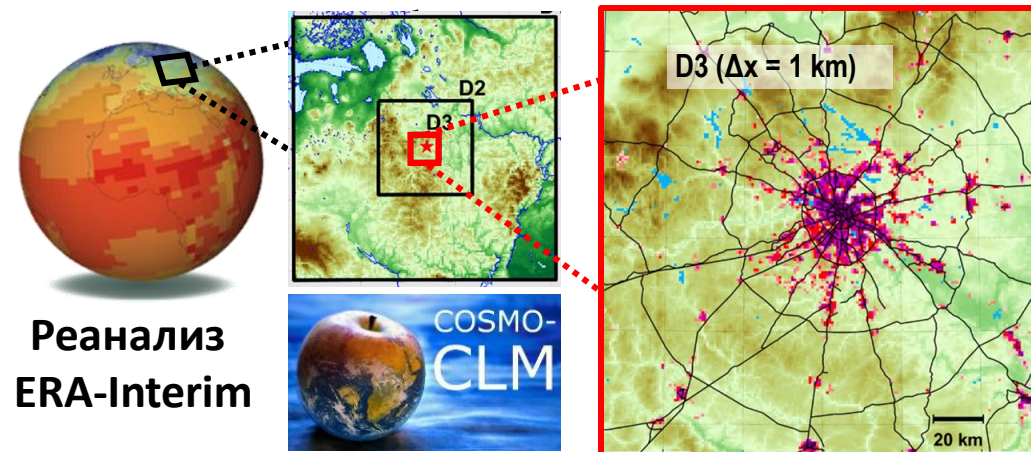
Этап №2

Аппроксимация острова тепла как пространственно-временного поля

Исходные данные: результаты численных экспериментов с региональной климатической моделью COSMO-CLM

- ✓ Динамический даунскейлинг реанализа ERA-Interim
- ✓ Целевая область моделирования с шагом сетки 1 км (160x160 км)*
- ✓ Эксперименты с учетом влияния города и без него (URB/noURB)* для 10 летних сезонов, 2007-16 гг. (Varentsov et al., 2018)
- ✓ Данные с дискретностью 1 час по времени

* пока используется только 5% от доступного массива данных



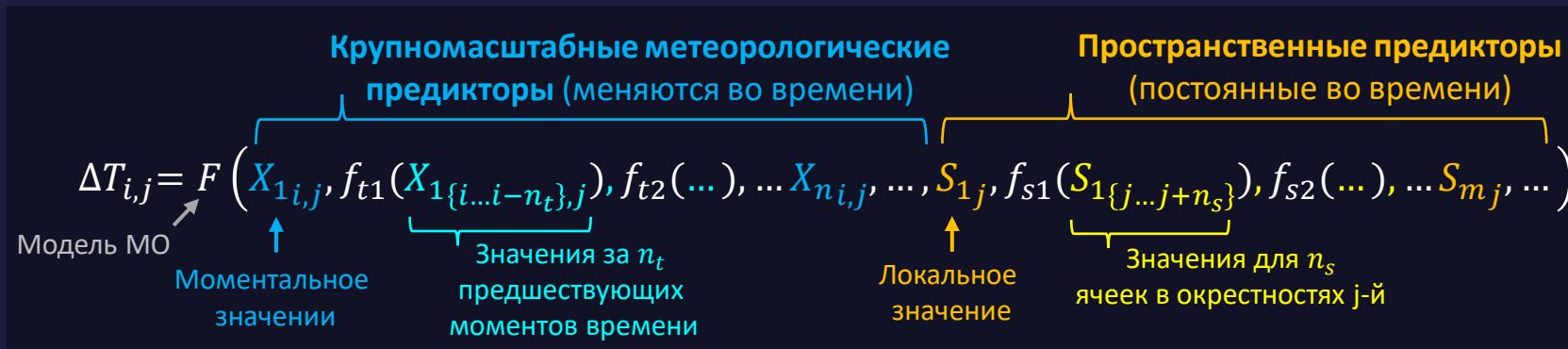
Этап №2

Аппроксимация острова тепла как пространственно-временного поля

- Целевая переменная: городская аномалия температуры, ΔT , в i -й момент времени и j -й точек пространства

- $\Delta T_{URB} = T_{URB} - T_{noURB}$

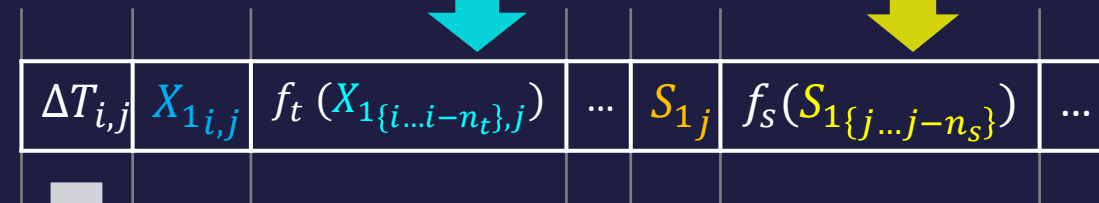
- Формулировка квазилокального подхода признакового описания:



- Метеорологические предикторы:** сглаженные до разрешения реанализа ERA-Interim результаты эксперимента noURB

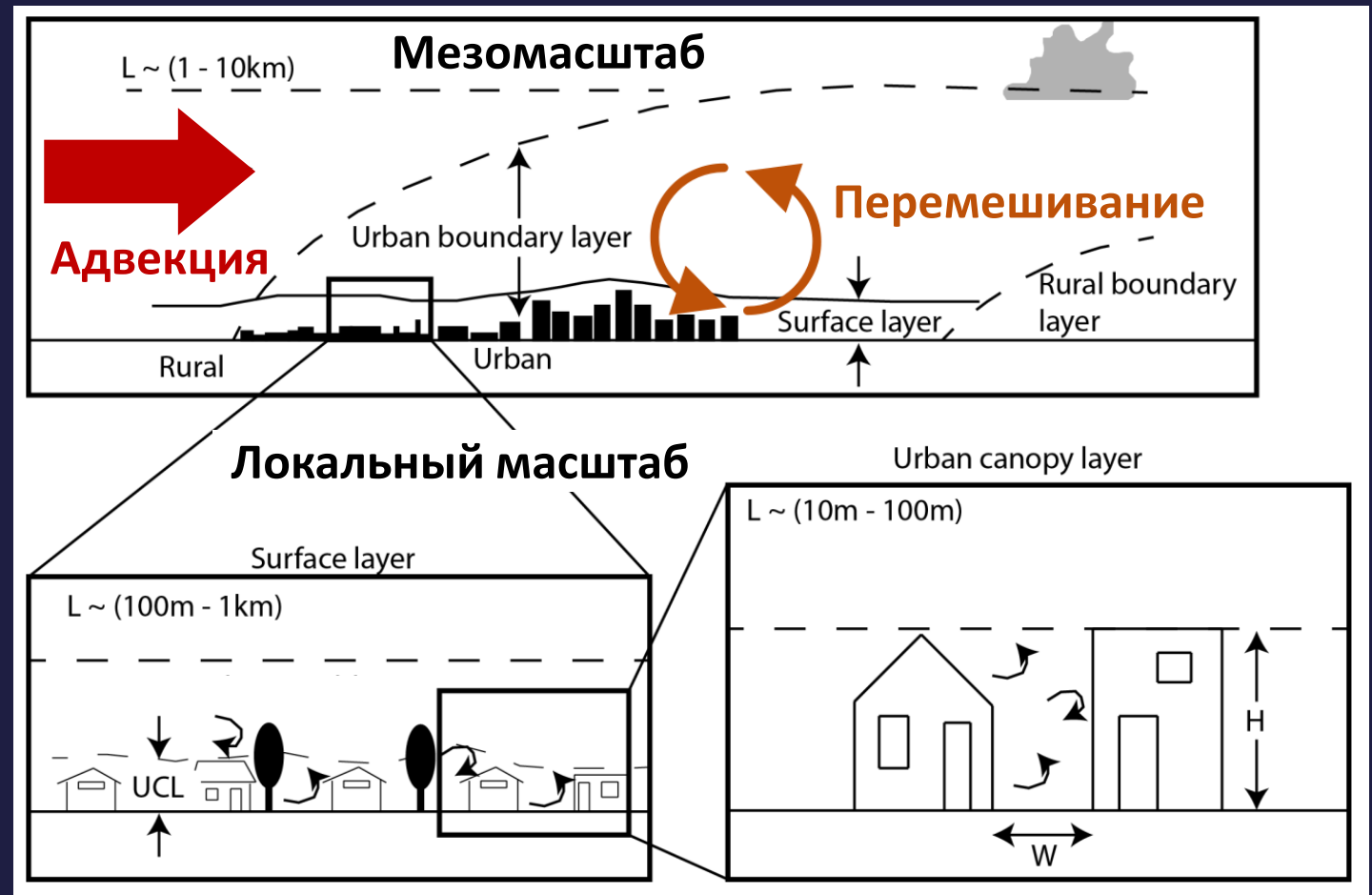
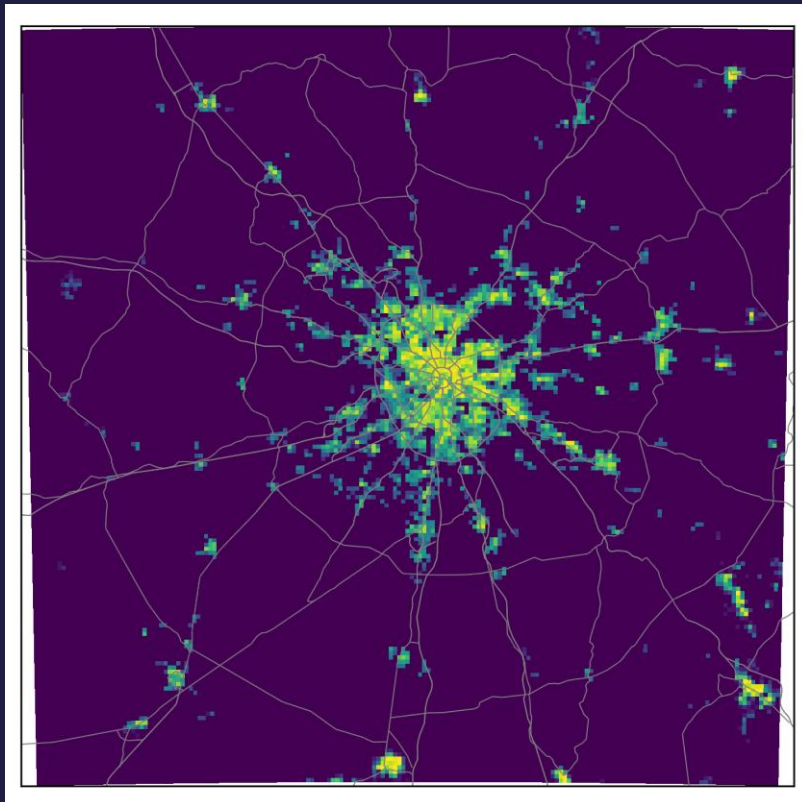
- Пространственные предикторы:** задаваемые в COSMO свойства поверхности и городской среды

- Модель МО:** CatBoost Regression – лучшая этапа №1



Пространственные предикторы

Локальные значения
пространственного предиктора

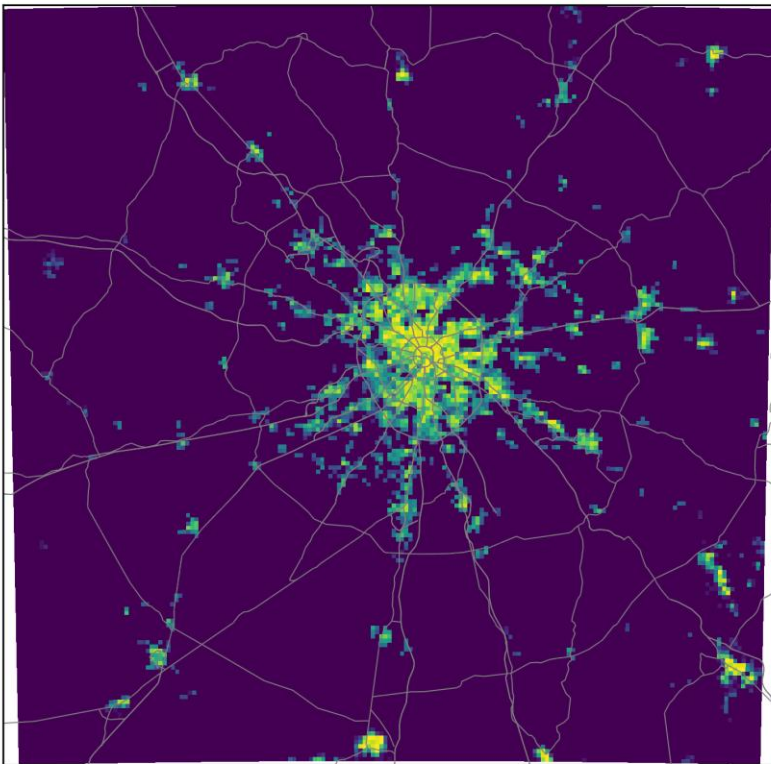


Пространственные предикторы

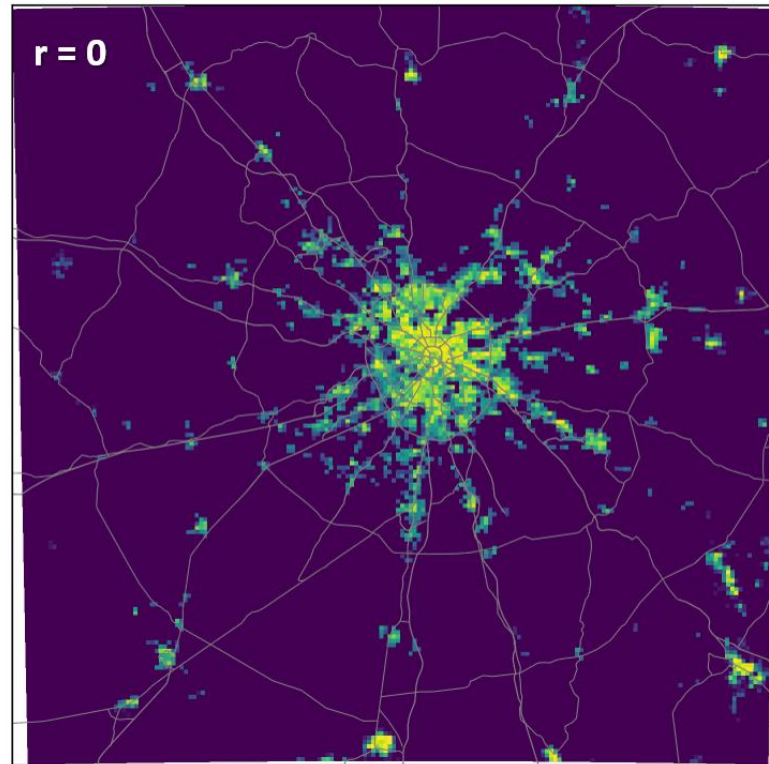
Локальные значения
пространственного предиктора

Учет перемешивания:
статический изотропный фильтр

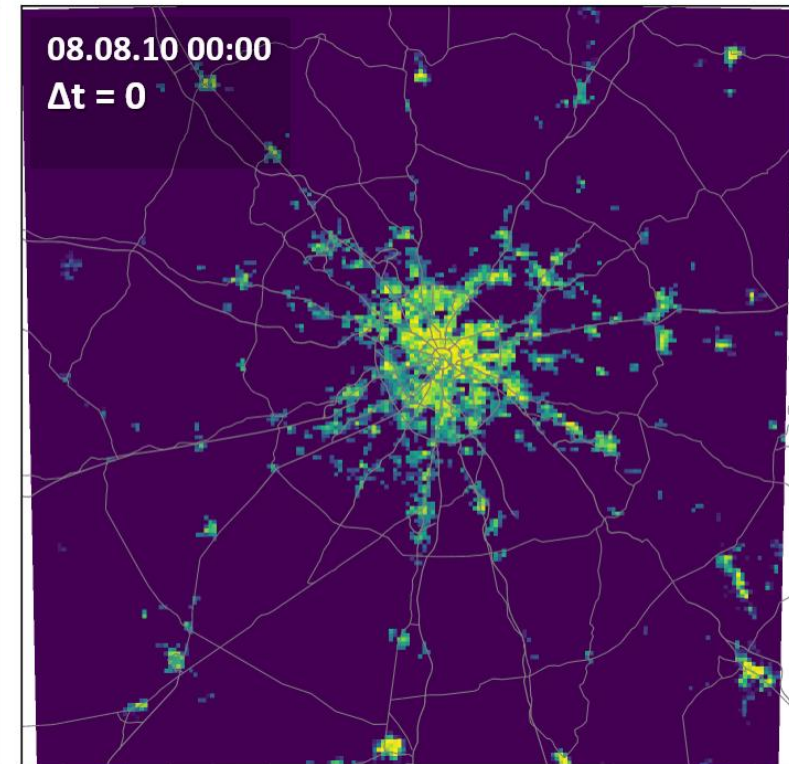
Учет адвекции:
эластичная трансформация +
динамический направленный фильтр



Эксперимент SF_loc



Эксперимент SF_dif



Эксперимент SF_adv

Аппроксимация ΔT_{URB}

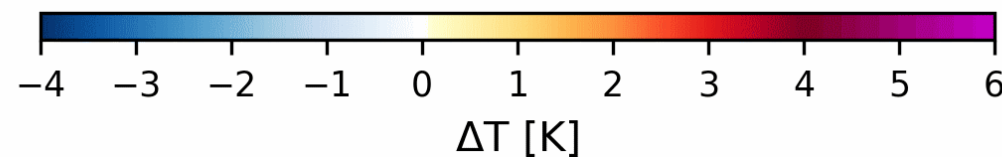
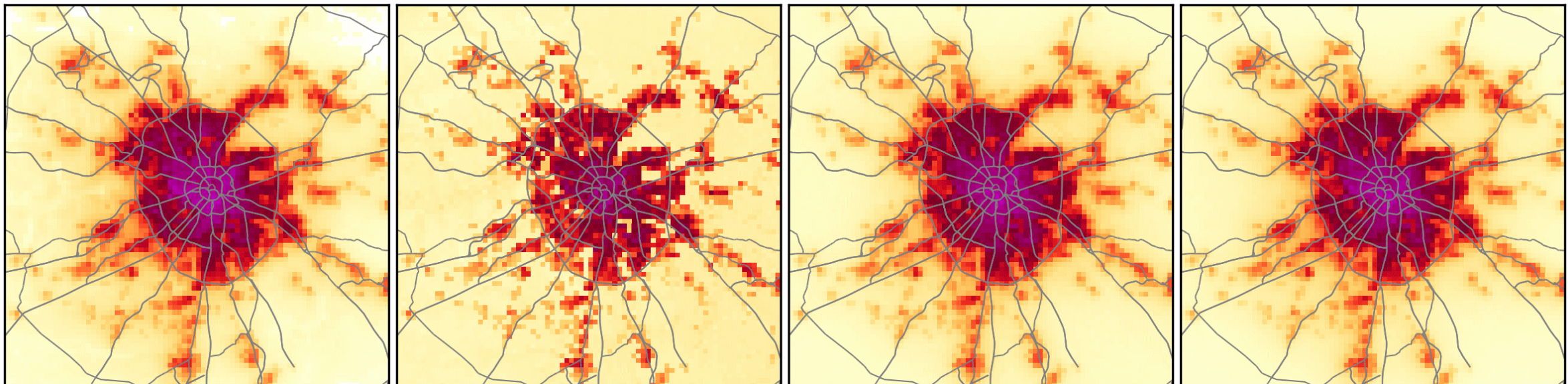
Средняя аномалия температуры за лето 2010 г.
0 MSK

COSMO (ground truth)

ML (SF_loc)

ML (SF_dif)

ML (SF_adv)



Аппроксимация ΔT_{URB}

Моменты времени с ярко выраженной адвекцией острова тепла

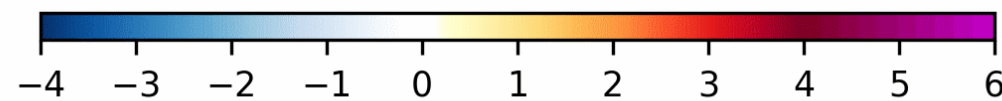
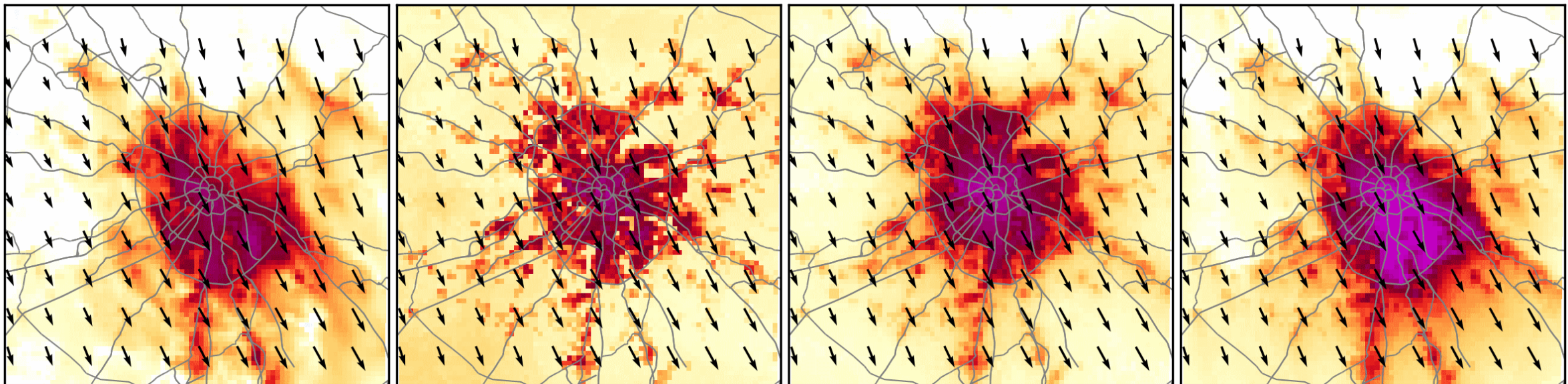
2010-08-12 03:00

COSMO (ground truth)

ML (SF_loc)

ML (SF_dif)

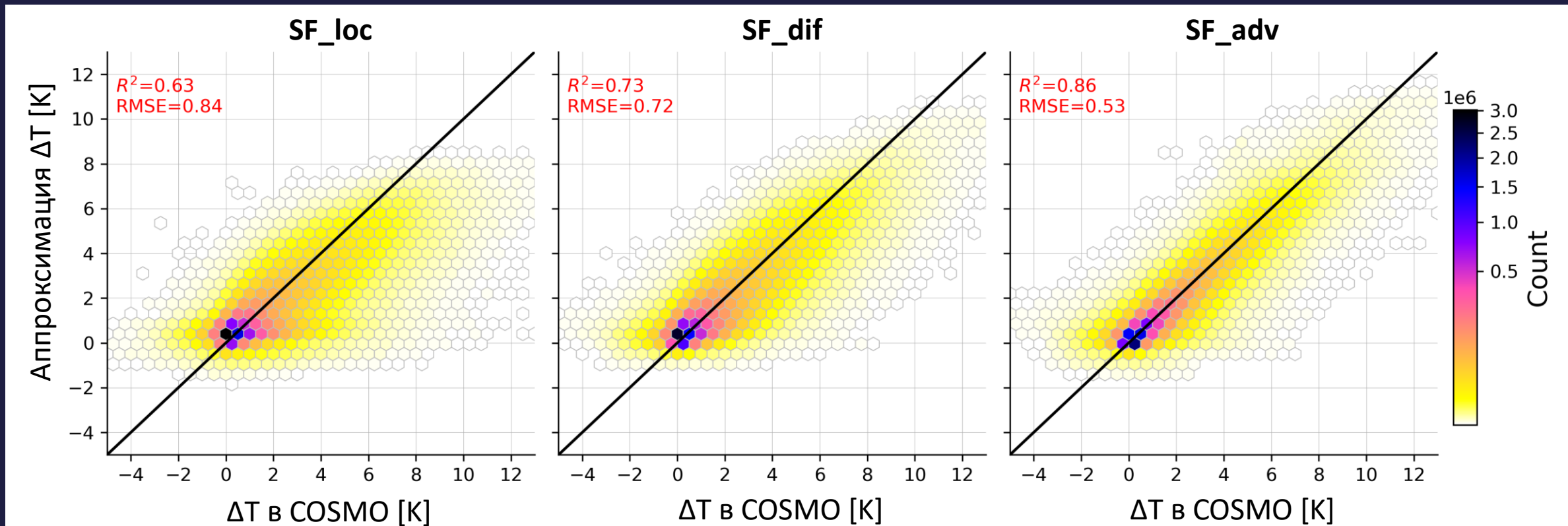
ML (SF_adv)



ΔT [K]



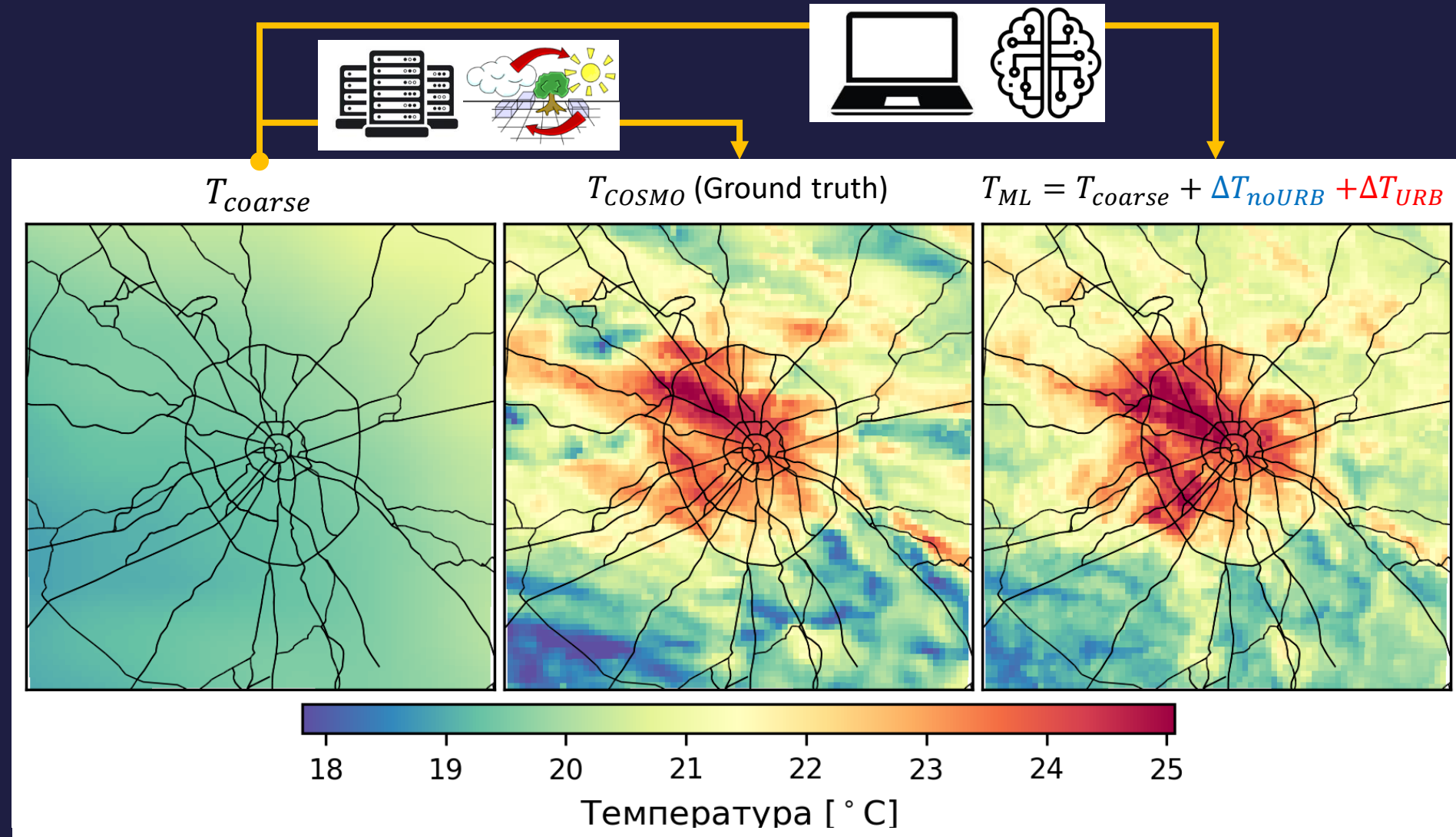
Аппроксимация ΔT_{URB}



Сравнение для всех точек по пространству
и только ночных сроков

Пример работы даунскейлинга

«Тропическая ночь»
накануне самого
жаркого дня в истории
Москвы, 29.07.2010





Результаты и обсуждение

I этап:

- ✓ Статистические модели основе методов МО успешно воспроизводят временную динамику интенсивности острова тепла Москвы на основе крупномасштабных метеорологических предикторов.
- ✓ Лучшие результаты достигнуты моделями градиентного бустинга, в первую очередь CatBoost.

II этап:

- ✓ Впервые предложена статистическая модель для совместной аппроксимации временной динамики и детализированной пространственной структуру городского острова тепла на основе крупномасштабных метеорологических предикторов и детализированных данных о подстилающей поверхности
- ✓ Предложен новый «квазилокальный» подход, позволяющий учитывать нелокальные взаимодействия между предикторами, в т.ч. перемешивание и адвекцию, в рамках «классических» моделей МО путем порождения дополнительных признаков.

Перспективы: сравнение «квазилокального» подхода с «нелокальным» подходом на основе свёрточных ИНС, анализ плюсов и минусов обоих подходов.

Open Access Article

Machine Learning for Simulation of Urban Heat Island Dynamics Based on Large-Scale Meteorological Conditions

Climate 2023, 11(10), 200;

<https://doi.org/10.3390/cli11100200>

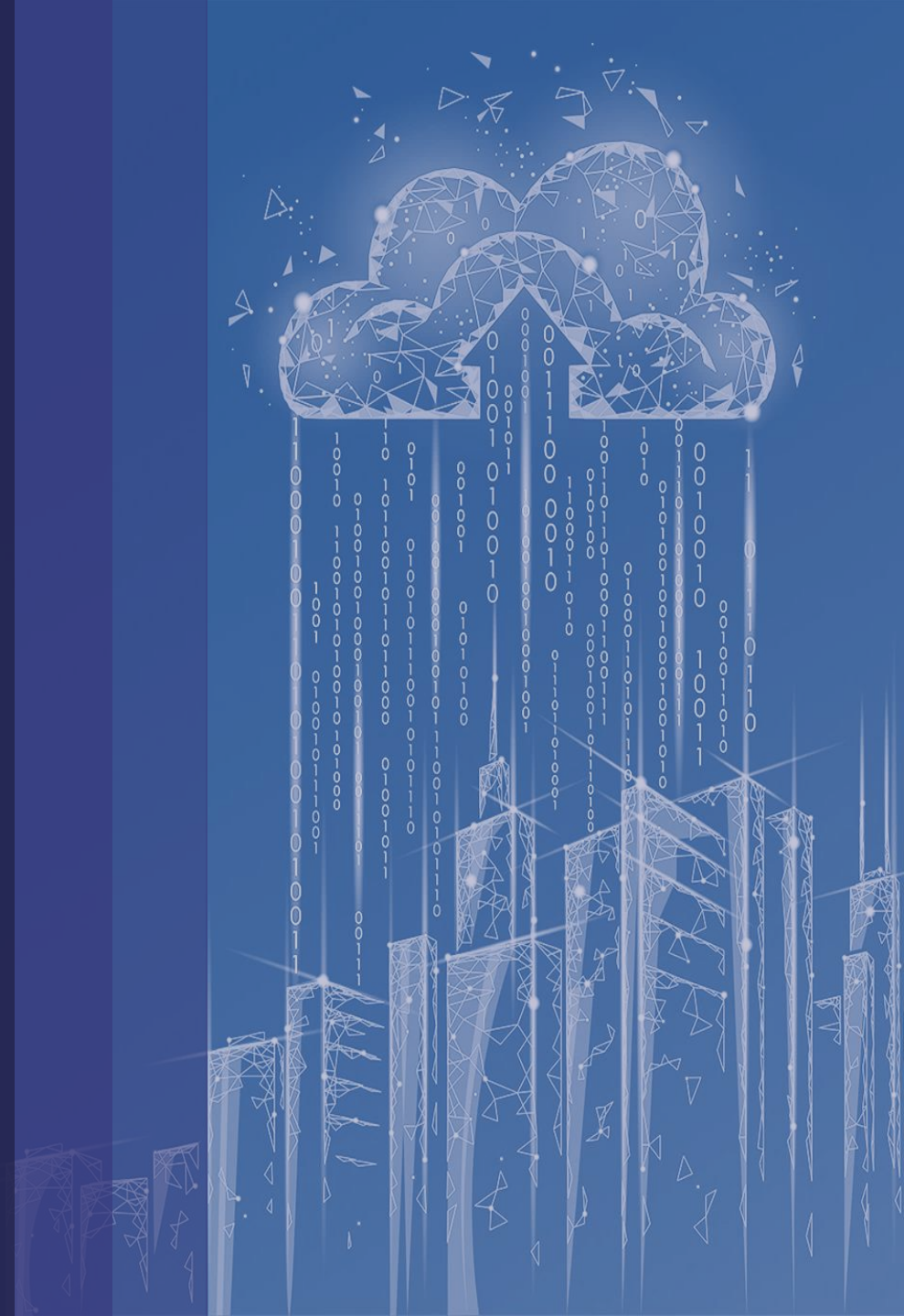




Фонд
интеллект

Спасибо за внимание! Вопросы?

mikhail.varentsov@srcc.msu.ru



Дополнительные слайды

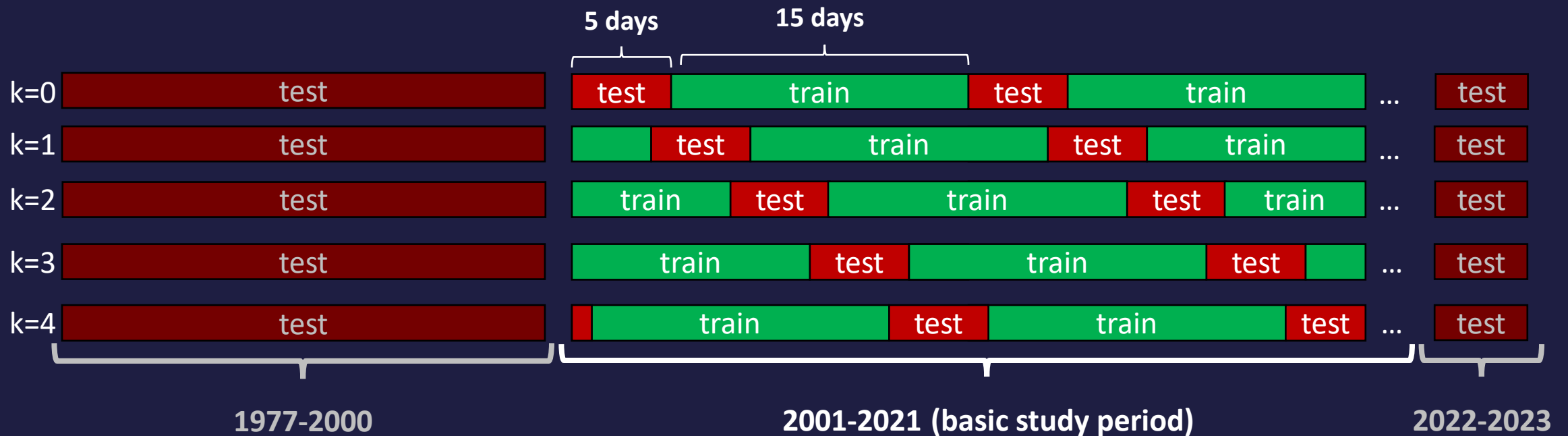
Этап №1: кросс-валидация

❑ The dataset:

- time series with 3-hour step for 47 years (1977-2021)
- 21 years (2001-2021) are used to train the models and for the major part of analysis

❑ Train-test split using blocked k-fold method with 5-days blocks and 4:1 train-to-test ratio

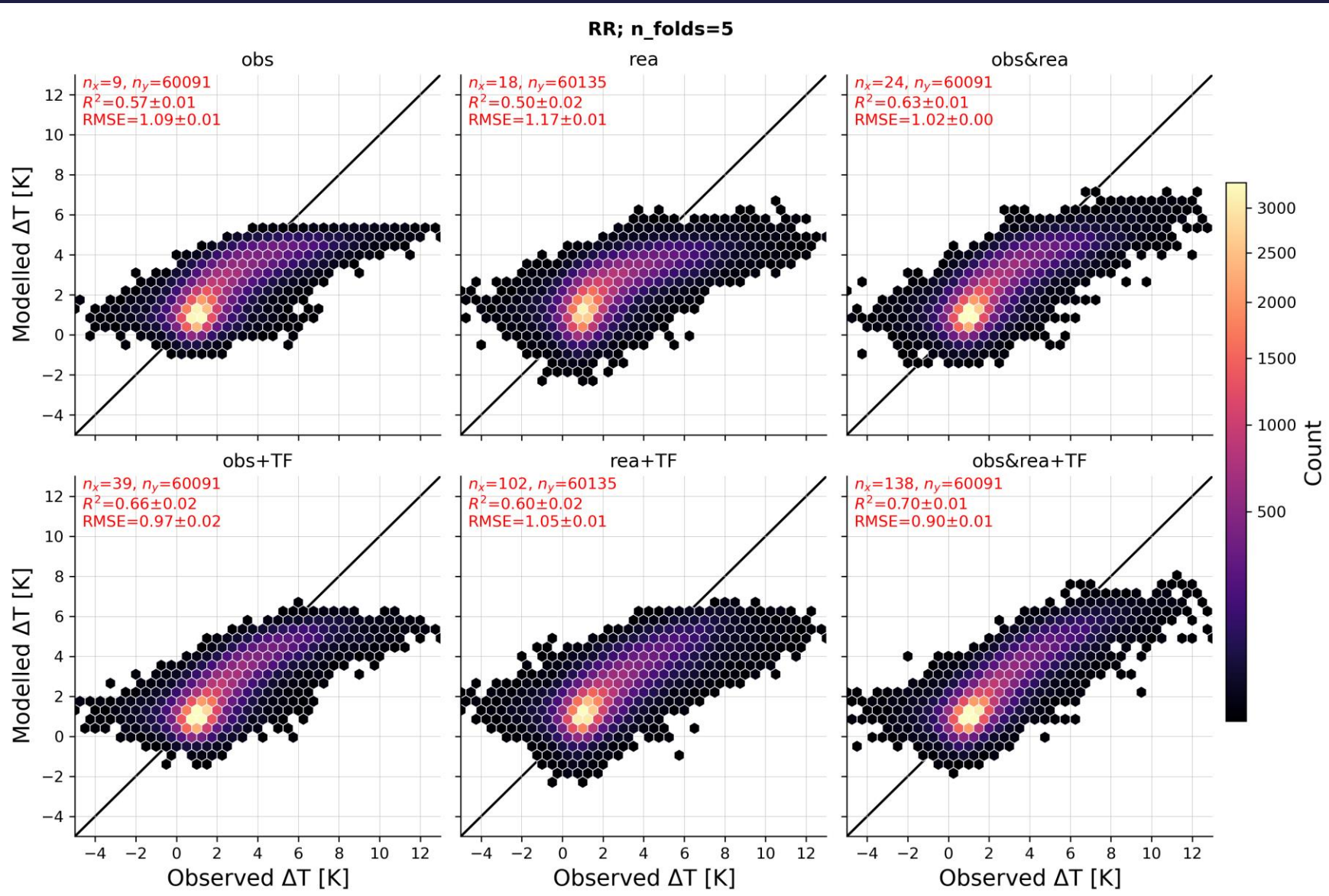
❑ Model quality metrics: RMSE, ME, R^2



Этап №1: результаты (RR)

Предикторы только
за текущий i-й
момент времени

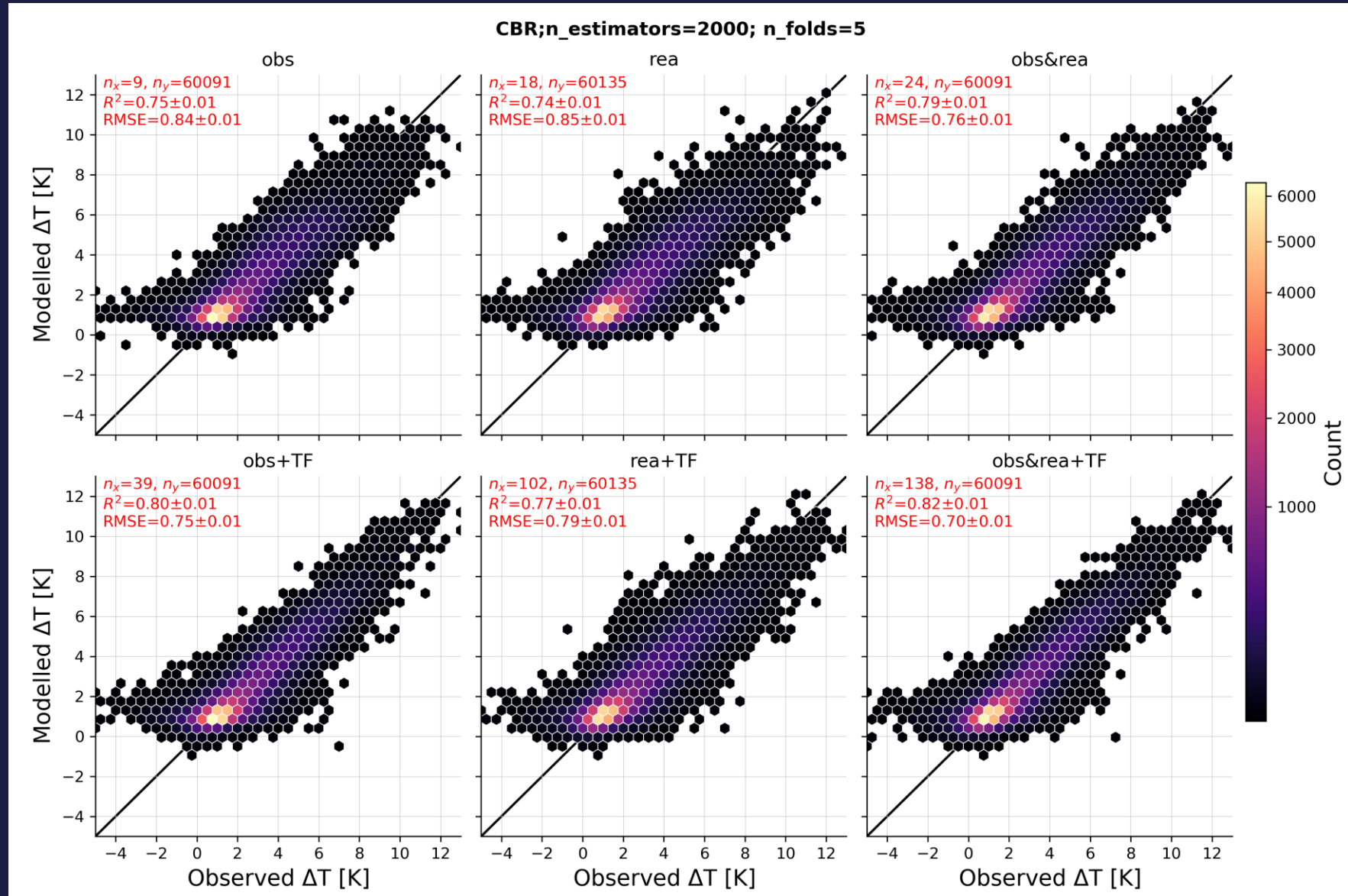
С дополнительными
признаками на
основе предикторов
за предшествующие
моменты времени



Этап №1: результаты (CVR)

Предикторы только
за текущий i-й
момент времени

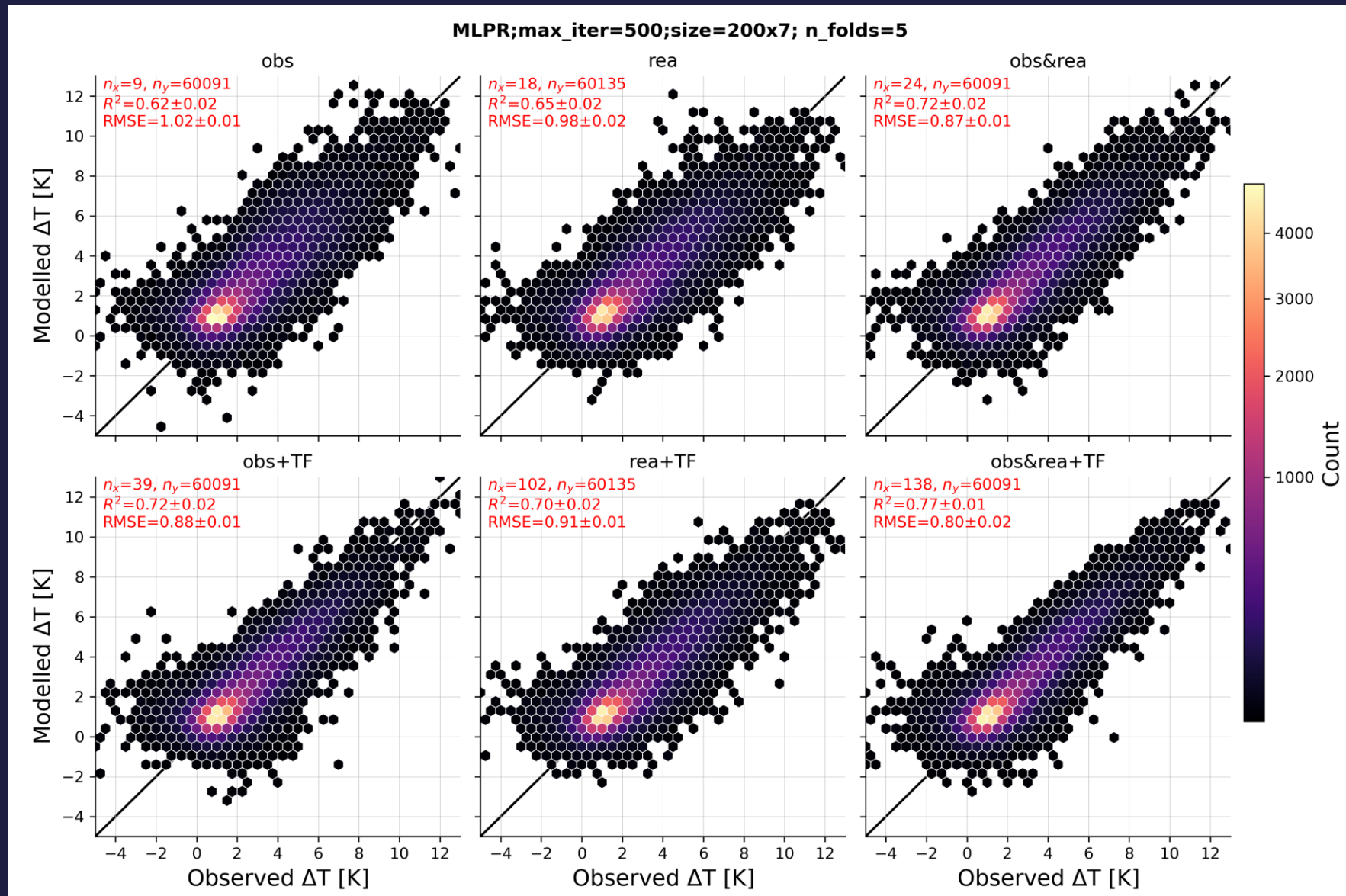
С дополнительными
признаками на
основе предикторов
за предшествующие
моменты времени



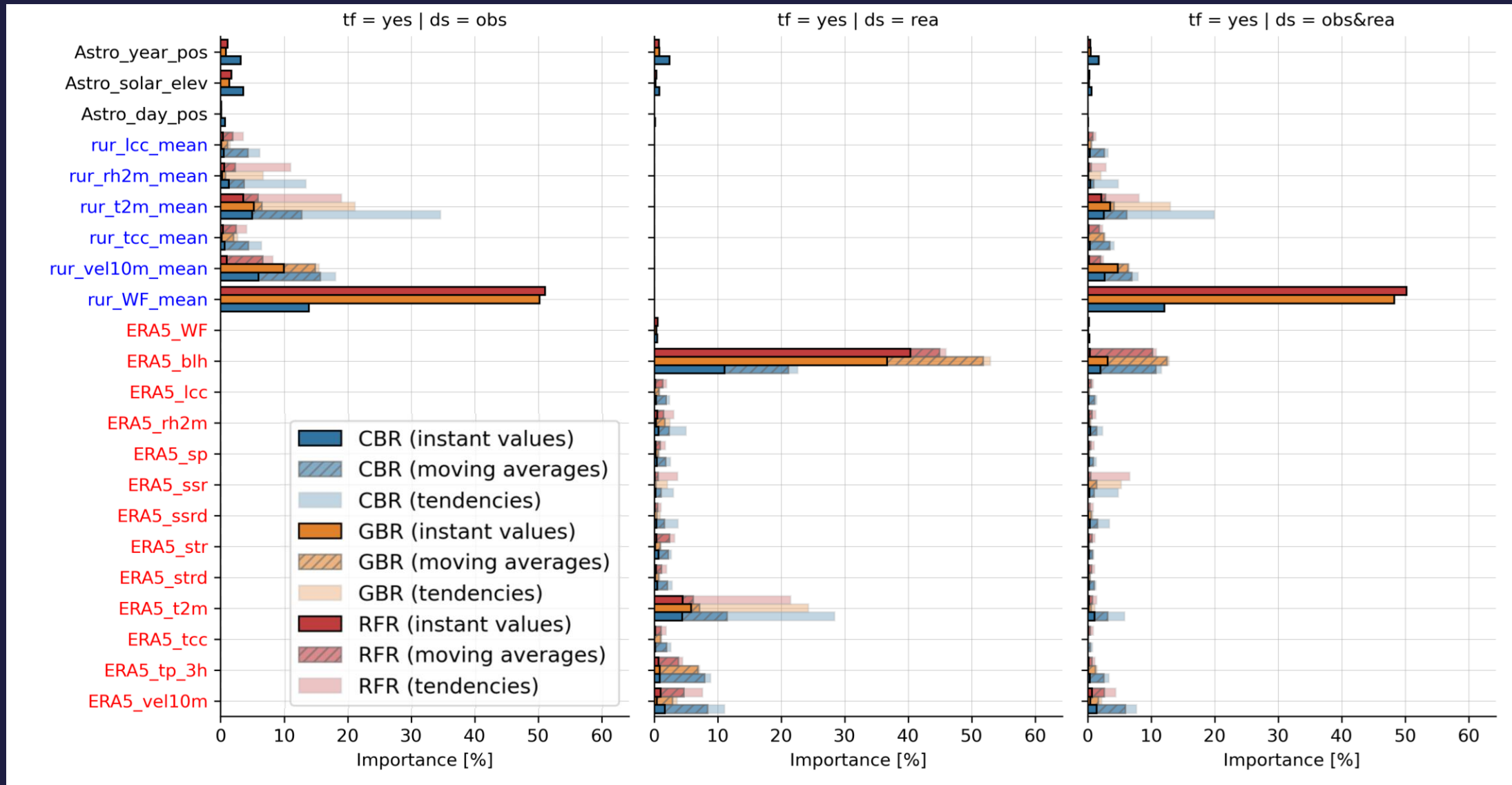
Этап №1: результаты (MLPR)

Предикторы только
за текущий i-й
момент времени

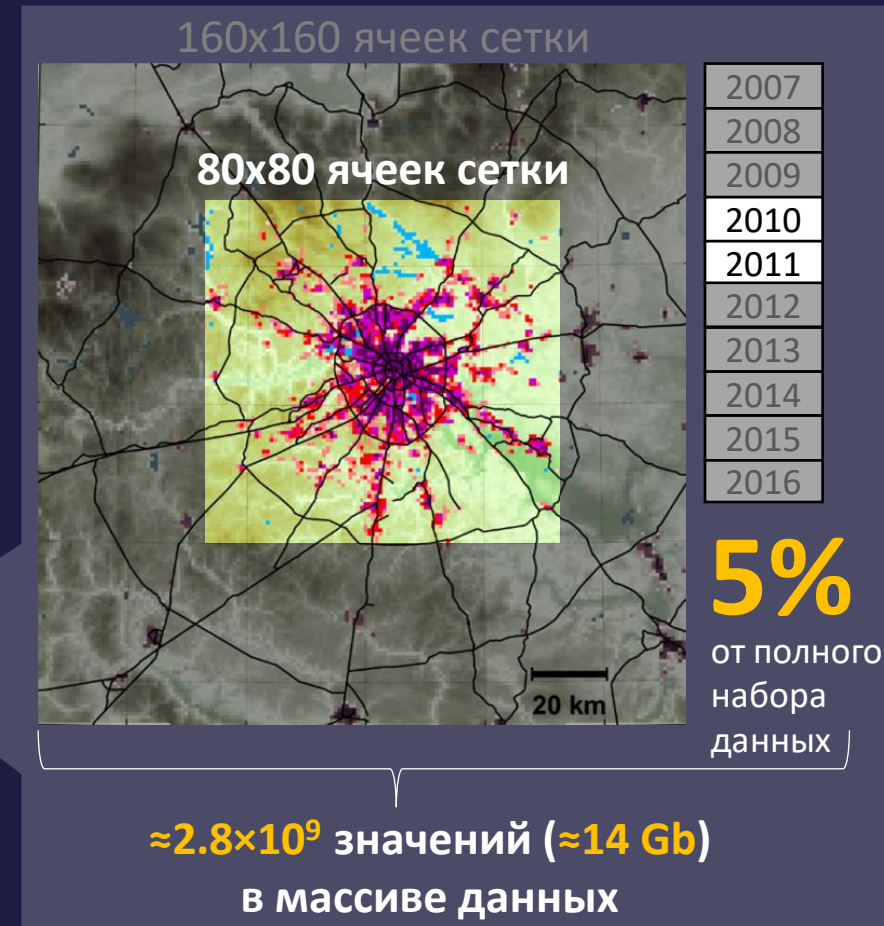
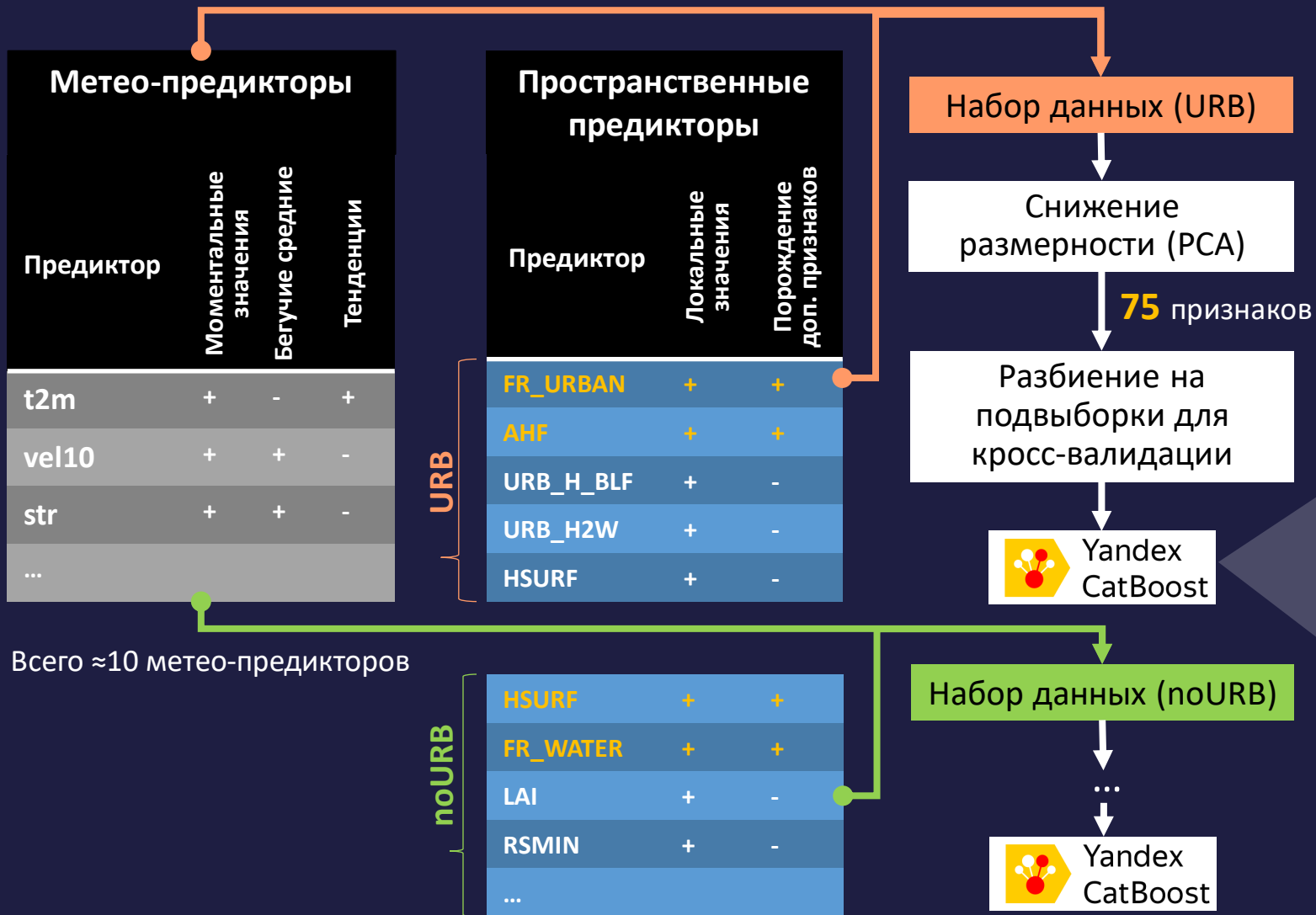
С дополнительными
признаками на
основе предикторов
за предшествующие
моменты времени



Этап №1: значимость предикторов



Этап №2: данные для обучения



«Квазилокальный» и «нелокальный» подходы

