

Ускорение численного моделирования в задаче рассеяния атмосферных примесей в городской застройке методами глубокого обучения

И. А. Герасимов¹, М. А. Криницкий^{1,2}, Е. В. Мортиков^{3,4}, А. В. Дебольский^{3,5}

1. Московский физико-технический институт (национальный исследовательский университет),
2. Институт океанологии им. П.П. Ширшова РАН,
3. Научно-исследовательский вычислительный центр Московского государственного университета имени М.В.Ломоносова,
4. Московский центр фундаментальной и прикладной математики, Россия,
5. Институт физики атмосферы им. А.М. Обухова РАН

Актуальность

1. Качество воздуха – значимый фактор здоровья людей. Ключевой вклад в загрязнение воздуха вносит транспорт
2. Достоверная пространственная оценка распределения загрязнений требует разрешение взаимодействий ветрового потока с городской застройкой
3. Прямой LES-расчет вычислительно затратен: выход на стационарный режим занимает время, исключающее оперативный анализ и изменчивость конфигураций застройки
4. Полуэмпирические Гауссовы модели никак не учитывают геометрию застроек
5. Машинное обучение может позволить приблизиться к точности LES с высокой скоростью оценки

Цель: Разработать вычислительно эффективную ML-модель, аппроксимирующую стационарное поле концентраций PM2.5 с учетом геометрии застройки, растительности и ветрового режима.

Задачи:

1. Численное LES моделирование
2. Подготовка данных: переход 3D $\rightarrow$ 2.5D
3. Настройка архитектуры
4. Обучение, аугментация, валидация
5. Анализ ошибок и физической согласованности

Физическая постановка

Метод:

Large Eddy Simulation (LES)

Домен:

256x256x64 метров, дискретизация на сетку
128x128x32 узлов.

Объекты:

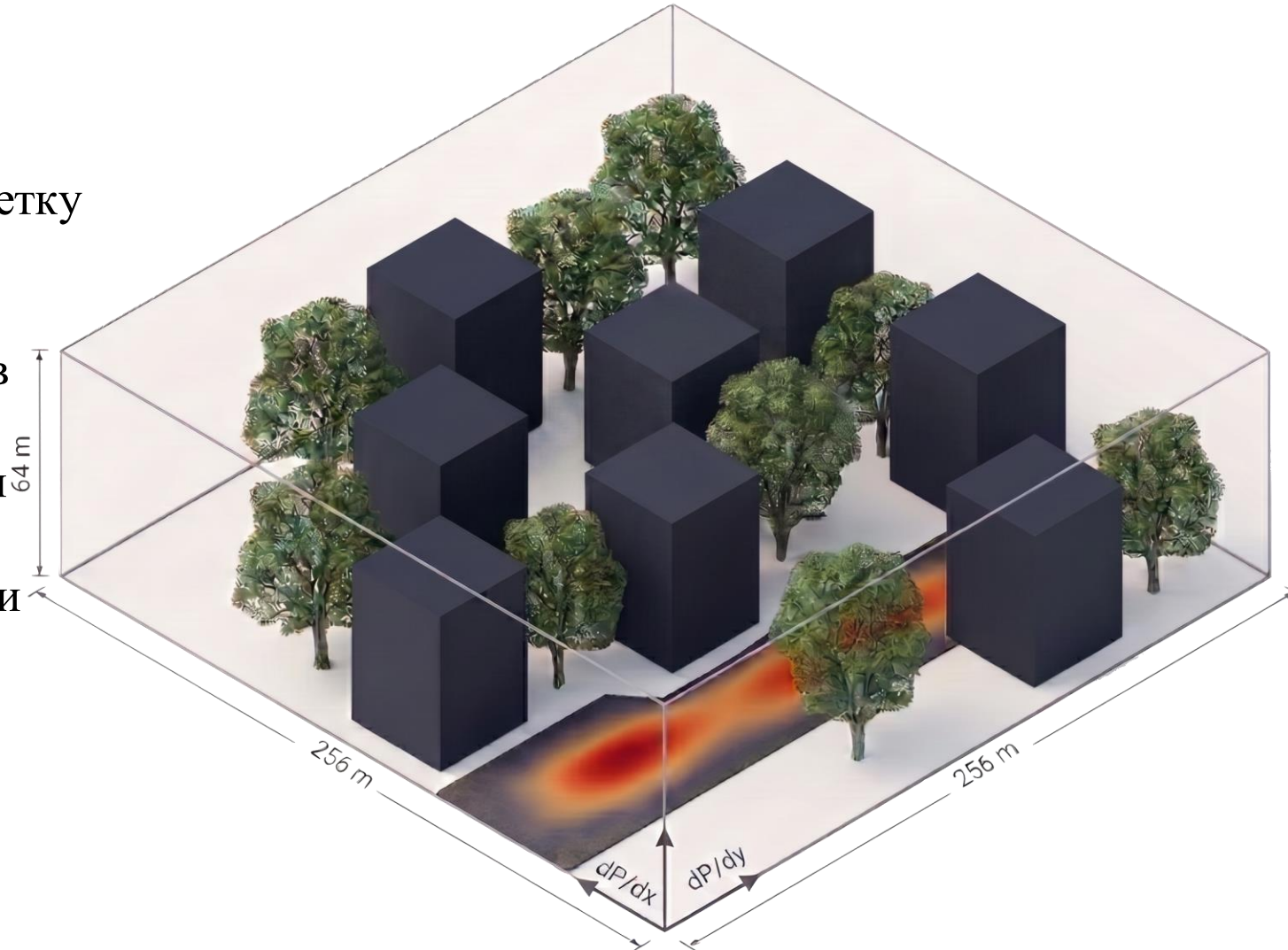
- Здания: фиксированная высота 16м (в шахматном порядке)
- Растительность: деревья с заданными LAD/LAI
- Дороги: объемные источники эмиссии (высота 8м, **PM2.5**)

Атмосфера:

Нейтральная стратификация (Москва)

Ветер:

Задается градиентом давления: $\frac{dP}{dx}$, $\frac{dP}{dy}$

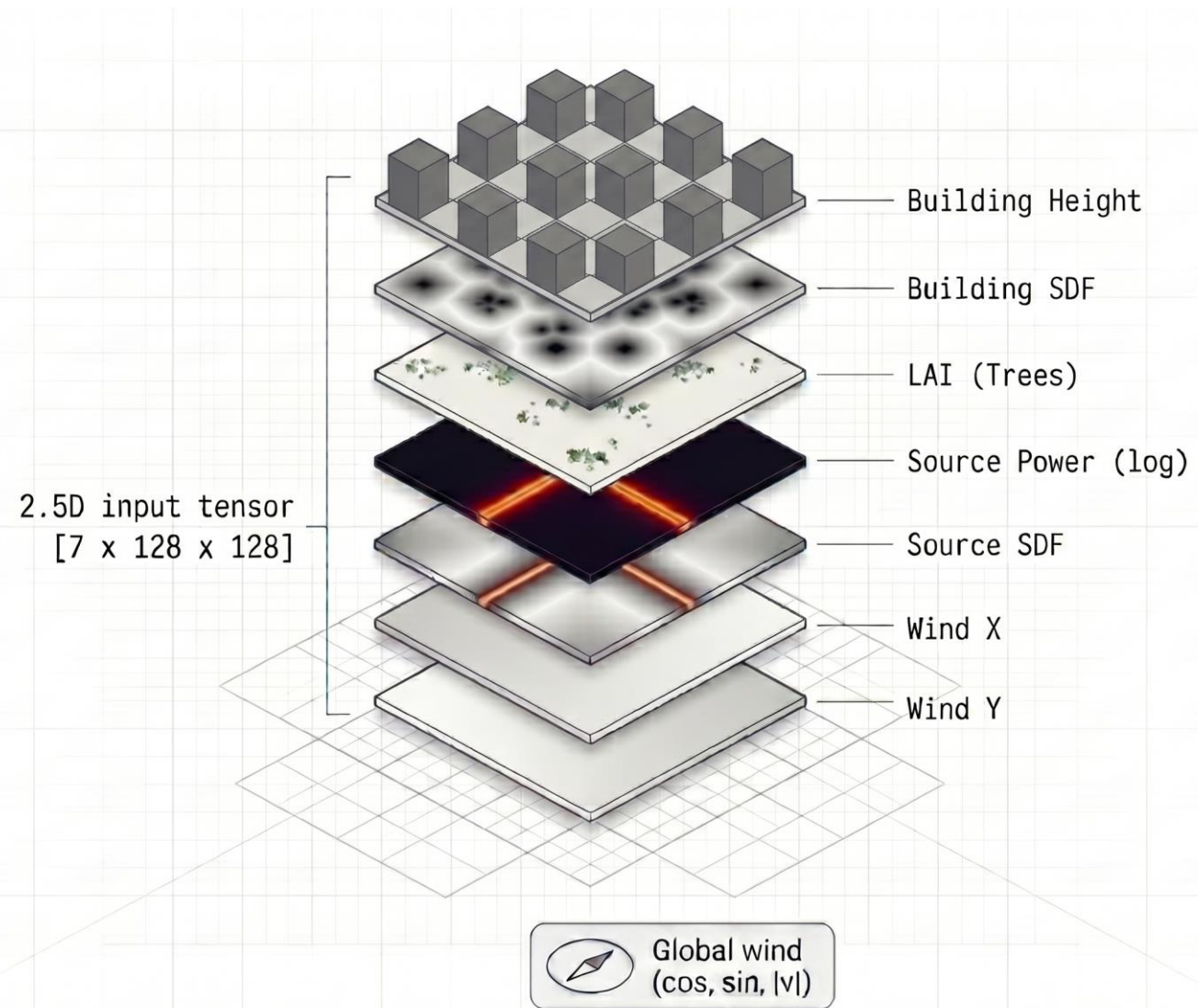


Входные данные

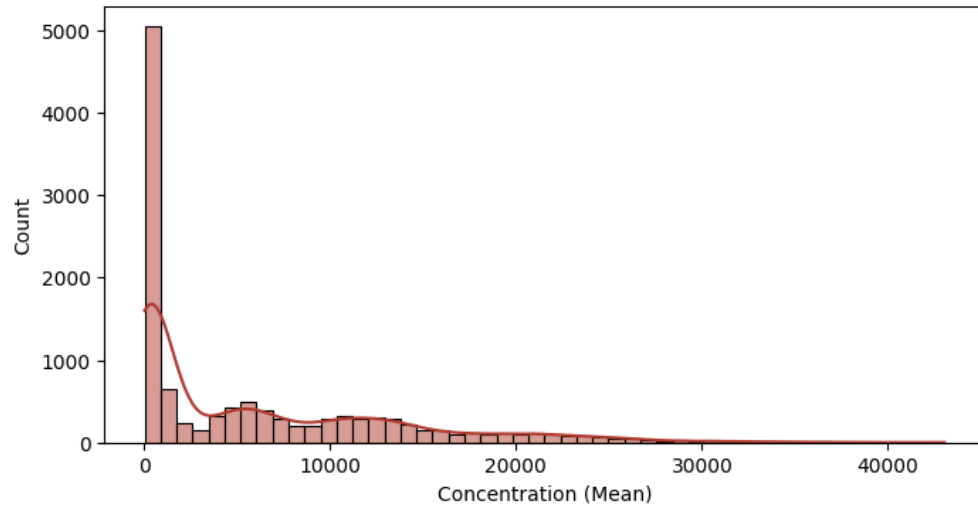
1. Каналы:

- **Building height:** нормированная высота зданий
- **Building SDF:** карта минимальных расстояний до зданий
- **LAI:** Карта индекса листовой поверхности
- **Source Power:** Мощность эмиссии дорог
- **Source SDF:** Карта минимальных расстояний до дорог
- **Wind X:** X-компонента ветра
- **Wind Y:** Y-компонента ветра

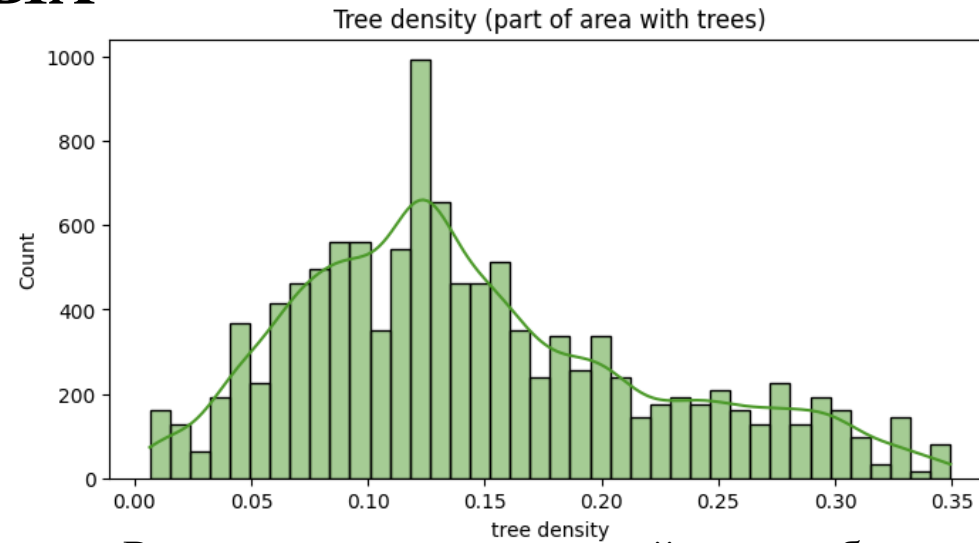
2. Global wind: Вектор компонент ветра



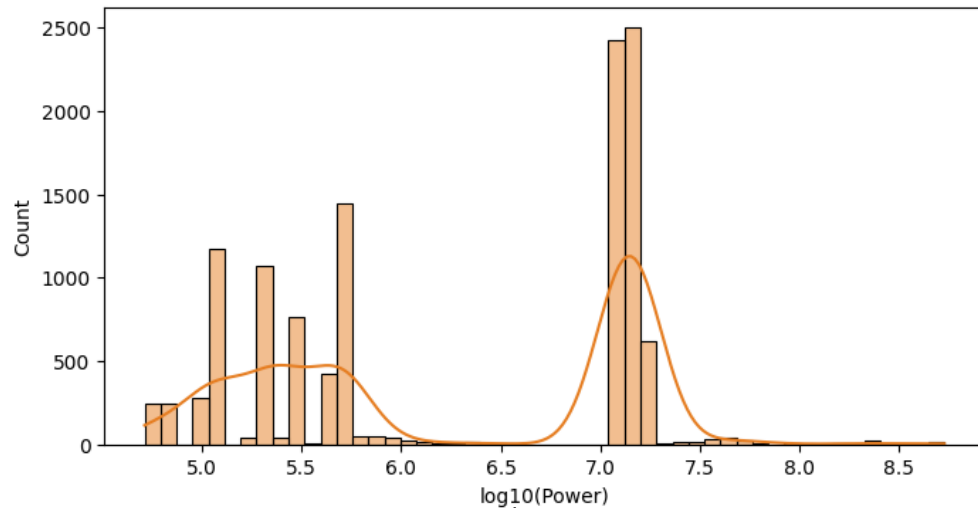
Описание входных данных



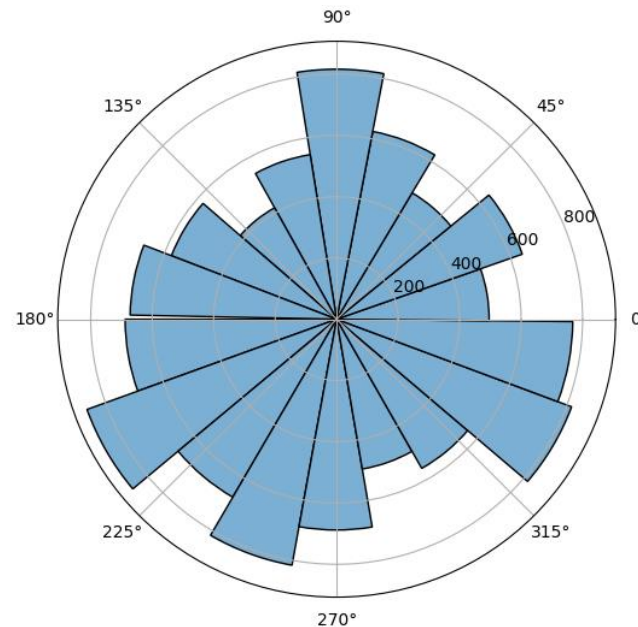
Распределение средних концентраций для слоя среднего



Распределение засаженной части области

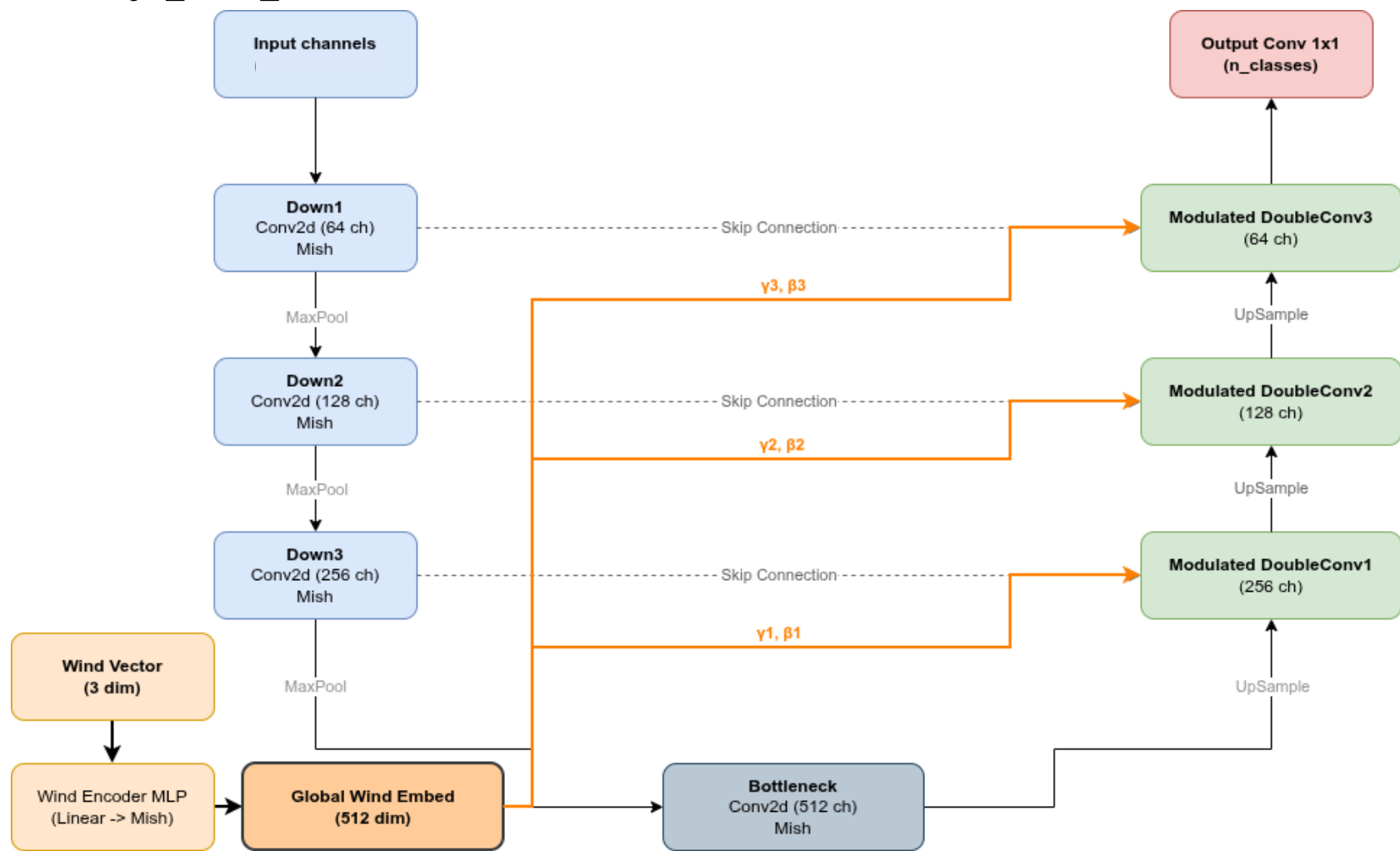


Распределение логарифма максимума мощности
источников



Распределение
направлений
ветра

Архитектура решения



Архитектура решения: UNet с FIMM вставкой

Механизм учета ветра: FILM

Проблема:

Необходим более строгий учет направления ветра

Решение:

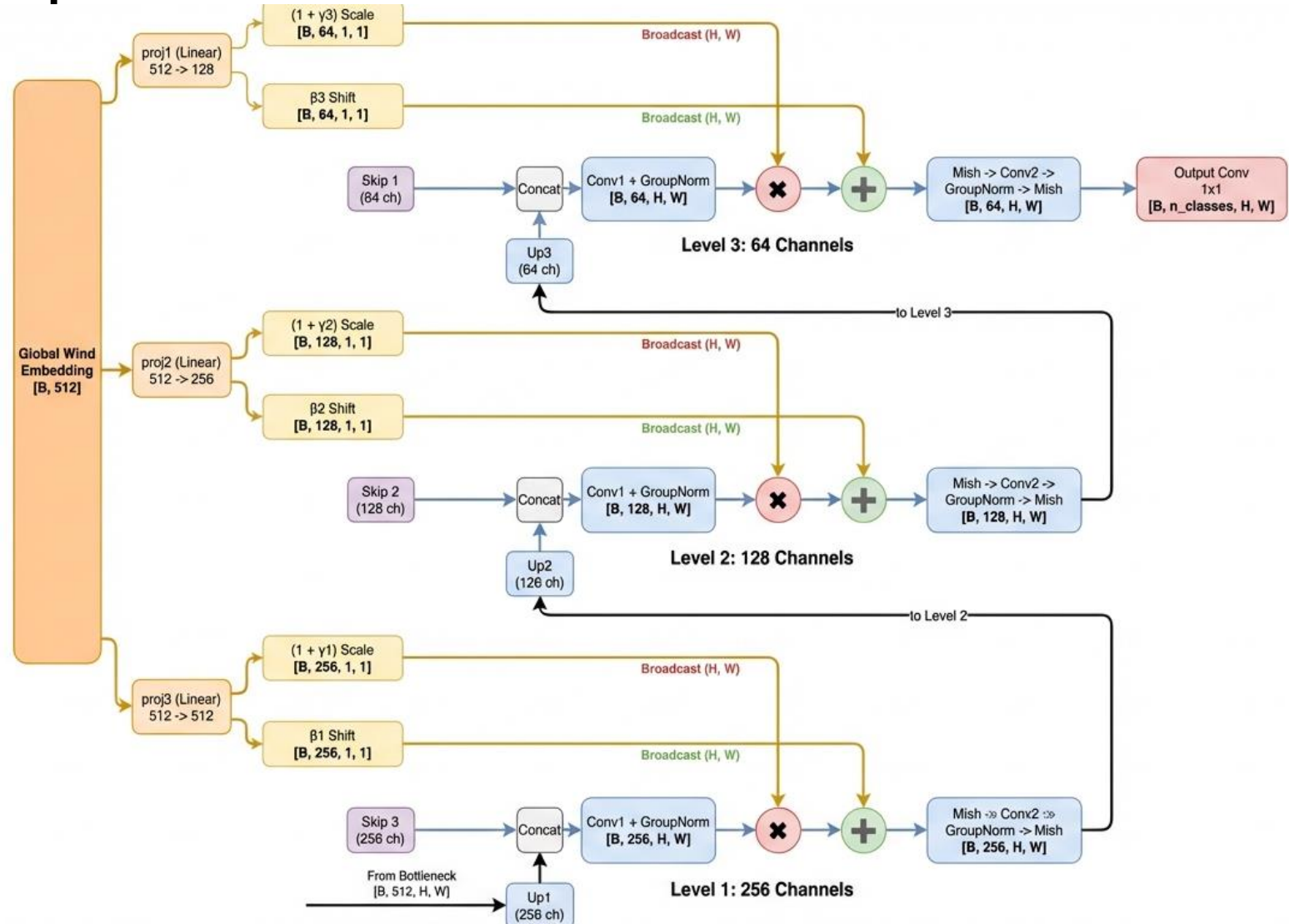
Feature-wise Linear Modulation

Вектор ветра $\rightarrow$ MLP $\rightarrow \gamma, \beta$ для каждого уровня декодера

Алгоритм:

Для каждого канала:

$$\begin{aligned} Channel_{new} \\ = Channel_{old} \cdot (1 + \gamma) + \beta \end{aligned}$$



Функция ошибки

Все вычисления — в лог-пространстве:

$$\hat{y} = \log(1 + \hat{C}), \quad y = \log(1 + C).$$

Функция ошибки:

$$L = MSE_w + \lambda_1 \cdot MAE_w + \lambda_2 \cdot MSE \quad (\lambda_1 = 0.05, \quad \lambda_2 = 1.0)$$

Веса (усиление шлейфа):

$$w = \frac{\min(\exp(\beta \cdot y), c_{max})}{mean_цена(\exp(\beta \cdot y))}, \quad \beta = 1, \quad c_{max} = 50$$

- MSE_w — основной член: веса усиливают влияние шлейфа (высокие концентрации)
- MAE_w — устойчивость к выбросам ($\lambda_1 = 0.05$)
- MSE — невзвешенная ошибка: смотрит на фон наравне со шлейфом, убирает завышение ($FB \rightarrow 0$)

Результаты: метрики

Уровень	Высота, м	R ²	RMSE	MAE	FB	FAC2	NMSE
z1	3	0.84	3038	1070	0.03	0.756	0.46
z5	10	0.88	2189	797	0.04	0.755	0.34
z8	16	0.89	1900	670	0.02	0.749	0.32
z25	50	0.89	1636	491	0.05	0.717	0.39
Mean		0.90	1735	559	0.03	0.776	0.31
Среднее		0.87	2150	714	0.03	0.751	0.39

Таблица 1. Финальные метрики архитектуры

статья: Chang, J. C., & Hanna, S. R. (2004). Air quality model performance evaluation. Meteorology and Atmospheric Physics, 87(1–3), 167–196. DOI: 10.1007/s00703-003-0070-7

Результаты: сравнение и вариации

Модель	R^2	RMSE	MAE	FB	FAC2
UNet 2.5D	0.87	2150	714	0.03	0.75
Аппроксимация по среднему полю	-0.31	6934	3483	1.83	0.19
Аппроксимация по ветру	-0.72	7963	4708	0.32	0.15

Таблица 2. Финальная модель vs baseline`ы

Модель	R^2	FB	FAC2	mass_ratio
Полная	0.394	-0.16	0.535	1.17
Без FILM	0.514	+0.25	0.479	0.78
Без аугментаций	0.485	-0.06	0.596	1.06

Таблица 3. Метрики при вариациях архитектуры, обучение на 50 эпохах

Выводы

1. Разработана вычислительно эффективная суррогатная модель для стационарного поля PM2.5 в городской застройке.
2. Достигнута приемлемая по стандартам дисперсионного моделирования точность: физический $R^2 = 0.87$, $FAC2 = 0.75$, $NMSE = 0.39$ при почти нулевом смещении ($FB = 0.03$). По критериям Chang & Hanna ($FAC2 \geq 0.5$, $NMSE \leq 1.5$) результат допустим.
3. Модель превосходит наивные baseline'ы
4. Вариации подтверждают важность вклада FiLM компоненты модели. Без FiLM формально выше R^2 , но модель физически хуже — недооценка массы на 22% ($mass_ratio\ 0.78$) и рост смещения ($FB +0.25$)

Спасибо за внимание