

The 10th International Conference on
Deep Learning in Computational Physics

July 7, 2026, Moscow, Russia

Transfer Learning via Source-Target Domain Data Mixing for Solving Multi-Parameter Inverse Problems in Optical Spectroscopy

Mardanov A.P.¹, Guskov A.A.^{1,2}, Isaev I.V.², Laptinskiy K.A.²,
Dolenko T.A.^{1,2}, Dolenko S.A.²

¹ Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia

² D.V. Skobeltsyn Institute of Nuclear Physics,
M.V. Lomonosov Moscow State University, Moscow, Russia

** This study has been conducted at the expense of the Russian Science Foundation
(grant No. 24-11-00266-P)*

Обратные задачи

Обратные задачи (ОЗ) – определение параметров системы по данным наблюдения за состоянием системы.

Обратная задача спектроскопии – определение характеристик объекта по его спектру.

В данной работе рассматривается задача **определения концентраций ионов** различных металлов в водных средах по **матрицам возбуждения-испускания фотолюминесценции углеродных точек (УТ)**, добавляемых к исследуемому раствору.

Актуальность

Оперативный анализ состава водных растворов востребован для:

- ✓ Экологического мониторинга природных вод
- ✓ Осуществления контроля за сточными водами и техническими водами промышленных предприятий
- ✓ Проверки качества пищевых продуктов

Современные требования к решению исследуемой задачи:

- **Бесконтактность** при использовании
- Определение в **экспресс-режиме** или в режиме реального времени
- Приемлемая точность

Актуальность

Традиционные химические методы:

- Высокая точность определения концентраций ионов
- **Контактные**
- **Требуют значительное количество времени**
- Каждый тест индивидуален для каждого иона
- Требуют специальное лабораторное оборудование и реактивы
- Требуют высокую квалификацию персонала

В качестве альтернативы рассматриваются
методы оптической спектроскопии (ОС).

Актуальность

Метод оптической спектроскопии:

- Фотолюминесцентная (ФЛ) спектроскопия

Особенность метода:

- Быстрый
- Бесконтактный
- Формы спектров ФЛ УТ в растворах чувствительны к изменениям ионного состава и концентраций ионов
- Отсутствует адекватная математическая модель,
описывающая изменения формы спектров ФЛ УТ

Актуальность

Для ОЗ ФЛ спектроскопии характерны следующие проблемы:

- возможная некорректность или плохая обусловленность
- **высокая размерность** как по входу
(спектры содержат сотни или тысячи каналов),
так и по выходу (десятки определяемых параметров)
- **трудоемкость и финансовая затратность сбора
большого количества экспериментальных данных**

Помимо этого, в многокомпонентных средах интенсивность в каналах спектра нелинейным образом зависит от концентраций компонентов.

Итог: многие стандартные методы решения ОЗ
недостаточно эффективны.

Актуальность

Одни из наиболее перспективных методов решения ОЗ – **методы машинного обучения (МО)**, а именно, нейронные сети (НС).

Достоинства:

- ✓ Не требуют априорной информации о виде искомой зависимости
- ✓ Устойчивы к шуму в данных

Недостаток: **чувствительны к недостаточному количеству данных**

Итог: необходимо разрабатывать различные методы компенсации этого недостатка.

Перенос обучения

Обучение алгоритмов МО – вычислительно дорогая процедура.

Требуются **большие массивы данных**, которые не всегда доступны.

Идея: для решения интересующей (целевой) задачи с небольшим массивом данных использовать **знания**, накопленные при решении другой подобной задачи с большим набором данных

Такой метод называется **Перенос обучения** (**Transfer Learning**)

Цель работы

Цель работы – исследование эффективности различных стратегий переноса обучения при решении обратной 7-параметрической задачи фотолюминесцентной спектроскопии.

Задача работы – сравнение эффективности различных стратегий комбинирования данных исходного и целевого доменов при увеличении числа ионов в многокомпонентном растворе.

Наносенсор на базе углеродных точек

В качестве базы для наносенсора использовались **УТ**, синтезированные **гидротермальным** методом из прекурсоров этилендиамина (ЭДА) и лимонной кислоты (ЛК).

Условия синтеза УТ:

- Соотношение прекурсоров ЭДА:ЛК: 2:1
- Температура: 200°C
- Время синтеза: 3 часа

Характеристики синтезированных УТ:

- Размер: 10 нм
- Квантовый выход: 96%
- Время жизни ФЛ люминофоров: ~3.5 нс и ~12.5 нс

Образцы

Массивы данных:

- ✓ 6-параметрическая задача (6П): Cu^{2+} , Ni^{2+} , Al^{3+} , Co^{2+} , Cr^{3+} , NO_3^-
- ✓ 7-параметрическая задача (7П): Cu^{2+} , Ni^{2+} , Al^{3+} , Co^{2+} , Cr^{3+} , NO_3^- , Pb^{2+}

Состав растворов:

- Бидистиллированная деионизованная вода
- Нитраты соответствующих катионов металлов
- Углеродные точки

Концентрации катионов: от 0 до 6 мМ с шагом 1.5 мМ

Концентрация УТ: 5 мг/л (фиксированная)

Данные

Для каждого приготовленного раствора были получены
матрицы возбуждения-испускания ФЛ УТ

Количество примеров в наборах данных:

- 6П: 3 125
- 7П: 7 813

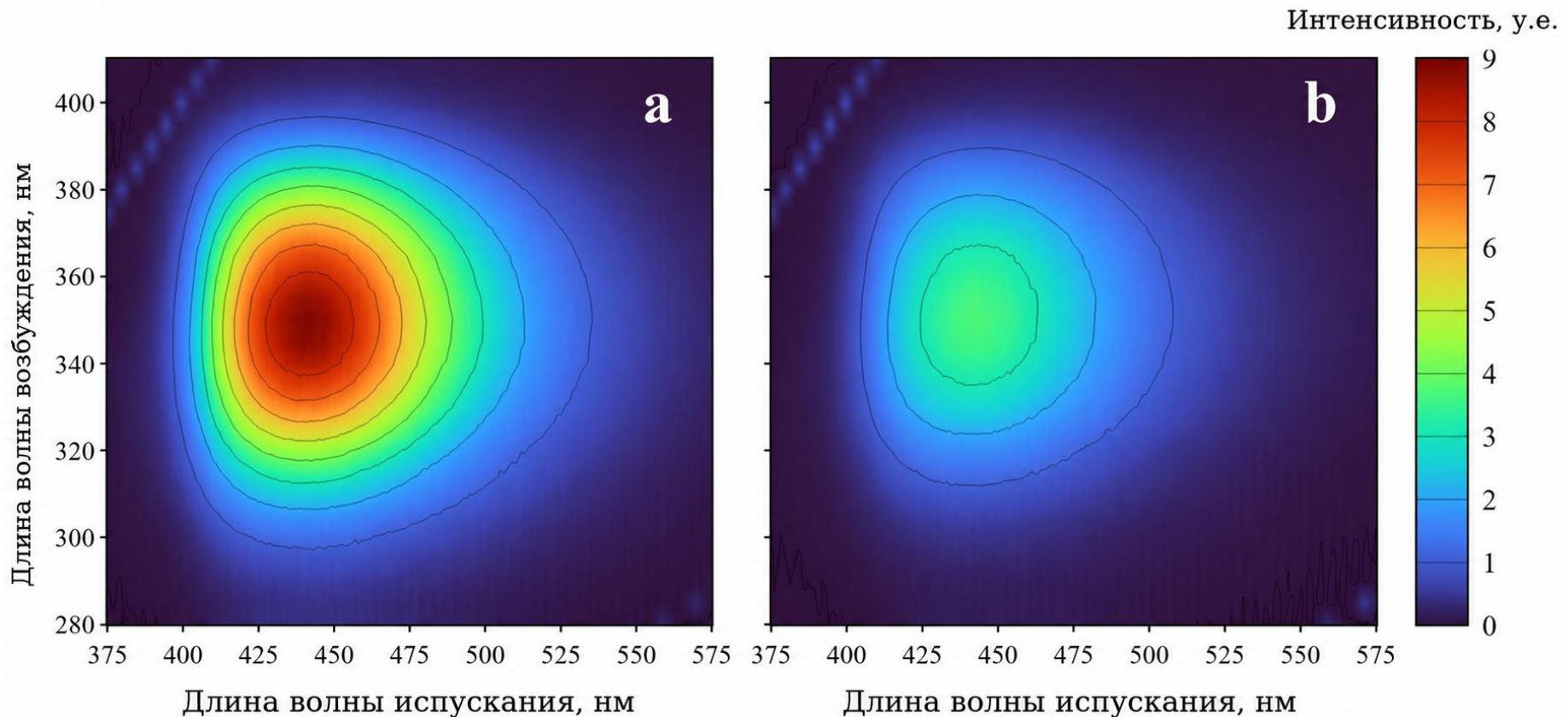
Разбиение на поднаборы:

- Тренировочный набор ~ 70%
- Валидационный набор ~ 20%
- Тестовый набор ~ 10%

Данные

Водные растворы УТ:

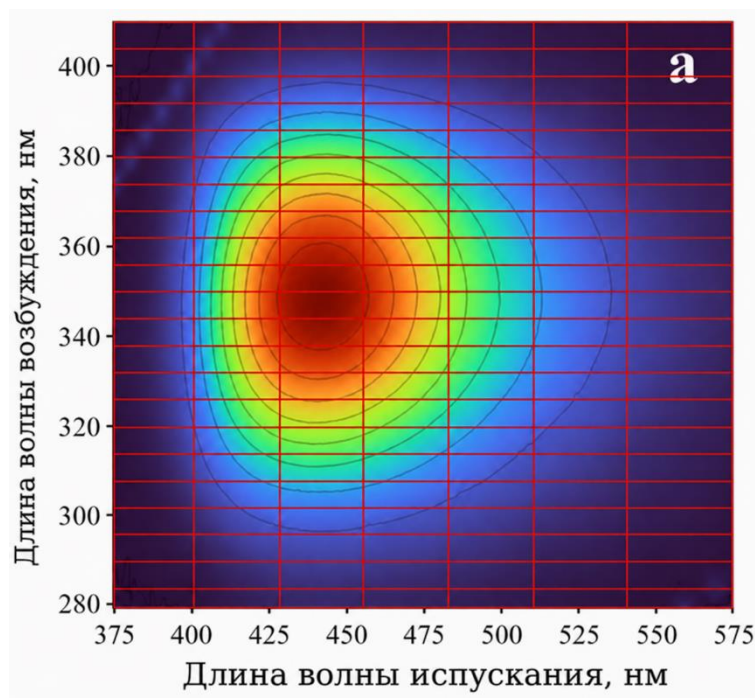
- Слева – без солей
- Справа – с солями $\text{Co}(\text{NO}_3)_2$ и $\text{Cu}(\text{NO}_3)_2$ по 6 мМ каждая



Входная размерность данных

Исходная входная размерность данных: 5 427 признаков
(27 длин волны возбуждения × 201 длину волны испускания)

Для уменьшения входной размерности данных использовалась
равномерная агрегация (суммирование) прямоугольниками
размером 3 на 7 (3 длины волны возбуждения × 7 длин волн испускания)



Параметры нейронной сети

Параметры обучения НС:

- **Архитектура:**
 - ✓ персептрон с 3 скрытыми слоями
 - ✓ 64+32+16 нейронов в скрытых слоях
 - ✓ передаточная функция для скрытых слоев – сигмоидальная, для выходного слоя – линейная
 - ✓ алгоритм оптимизации – стохастический градиентный спуск
 - ✓ скорость обучения от первого скрытого слоя к выходному – $0.1 \rightarrow 0.05 \rightarrow 0.02 \rightarrow 0.01$ с плавным уменьшением по косинусному закону.
 - ✓ момент – 0.9
- **Базовый критерий остановки обучения:**
 - ✓ 500 эпох после минимума ошибки на валидационном наборе

Параметры нейронной сети

Для каждого определяемого параметра строится отдельная модель с одним выходом (**автономное определение**)

Обучение моделей внутри каждого разбиения производилось с тремя разными случайными инициализациями весов **НС**

Второй валидационный набор

Критерий выбора состояния базовой модели:

- Независимый контроль пригодности модели для целевой задачи:
 - ✓ используется второй независимый валидационный набор
 - ✓ набор соответствует валидационному набору целевой 7П задачи – ВЛД_7П
 - ✓ ошибка ВЛД_7П контролируется параллельно с базовой валидацией
 - ✓ критерий останова ВЛД_7П – 2000 эпох после минимума ошибки

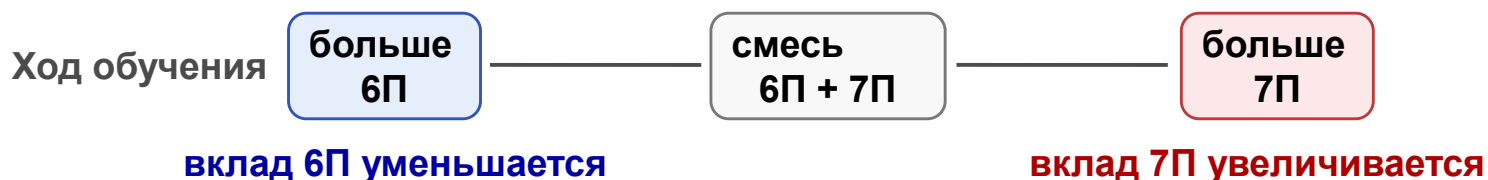
Идея: выбирать базовую модель не только по текущему этапу, но и по её поведению на целевой задаче

Перераспределение весов примеров

Общий принцип:

- Совместное использование исходного и целевого наборов данных:
 - ✓ модель обучается на общей выборке 6П и 7П
 - ✓ каждый пример вносит свой вклад в изменение параметров сети
 - ✓ в процессе обучения вклад разных частей выборки может изменяться
 - ✓ акцент постепенно смещается от исходной задачи к целевой 7П задаче

Идея: сохранить информацию из 6П, но направить обучение к решению целевой 7П задачи



Задачи вычислительного эксперимента

1. Референсное решение для 7-параметрической задачи
2. Исследование эффективности дообучения с 6-параметрической задачи на 7-параметрическую задачу
3. Исследование эффективности дообучения с объединённой выборки 6П+7П на 7-параметрическую задачу

Задачи вычислительного эксперимента

4. Исследование эффективности поэтапного дообучения
 $6П \rightarrow 6П+7П \rightarrow 7П$
5. Исследование эффективности дообучения с $6П$ на $7П$
с использованием двух валидационных наборов
6. Исследование эффективности дообучения с $6П+7П$ на $7П$
с использованием двух валидационных наборов
7. Исследование эффективности поэтапного дообучения
 $6П \rightarrow 6П+7П \rightarrow 7П$
с использованием двух валидационных наборов

Задачи вычислительного эксперимента

8. Исследование эффективности перераспределения весов примеров при обучении на совместной выборке **6П** и **7П**

8.1. Коэффициент в функции потерь

$$L = \alpha \cdot L_{\text{6П}} + (1 - \alpha) \cdot L_{\text{7П}}$$

Ошибки на примерах из **6П** и **7П** входят в функцию потерь **с фиксированными весами**, при разных α меняется постоянное соотношение вкладов двух наборов

Задачи вычислительного эксперимента

8. Исследование эффективности перераспределения весов примеров при обучении на совместной выборке 6П и 7П

8.2. Доли наборов в пакетах

$$\text{пакет} = p(t) \cdot \text{пакет_7П} + (1 - p(t)) \cdot \text{пакет_6П}$$

В каждом **обучающем пакете** заранее задаётся **доля примеров** из 6П и 7П, влияние наборов регулируется через частоту попадания примеров в обучение

Задачи вычислительного эксперимента

8. Исследование эффективности перераспределения весов примеров при обучении на совместной выборке 6П и 7П

8.3. Отбор примеров исходной задачи

6П_отобр. = ближайшие примеры 6П к данным 7П

Модель сравнивает примеры 6П и 7П через своё внутреннее представление спектров, для дообучения используются только **наиболее близкие примеры** исходной задачи

Задачи вычислительного эксперимента

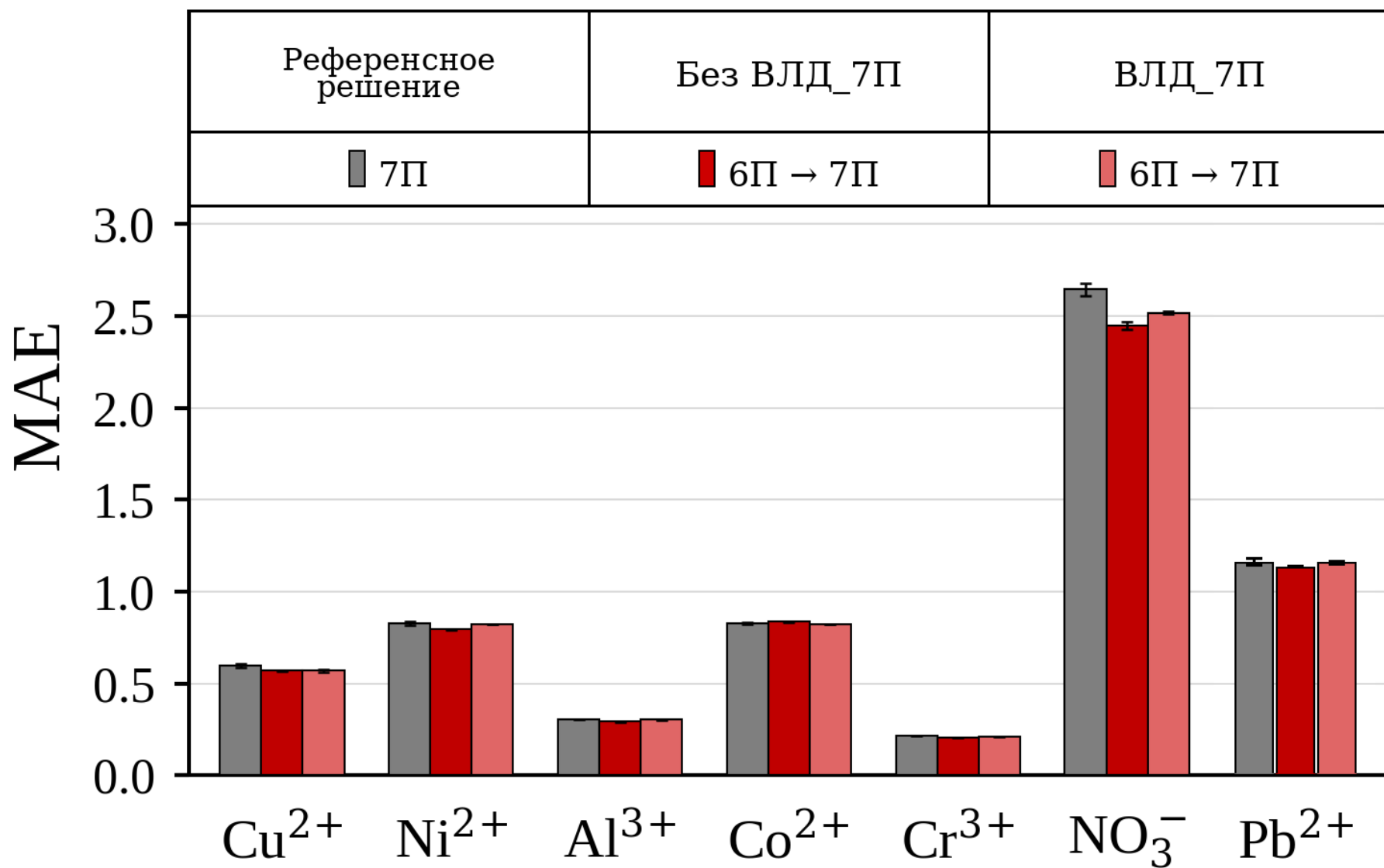
8. Исследование эффективности перераспределения весов примеров при обучении на совместной выборке **6П** и **7П**

8.4. Плавный переход к целевой задаче

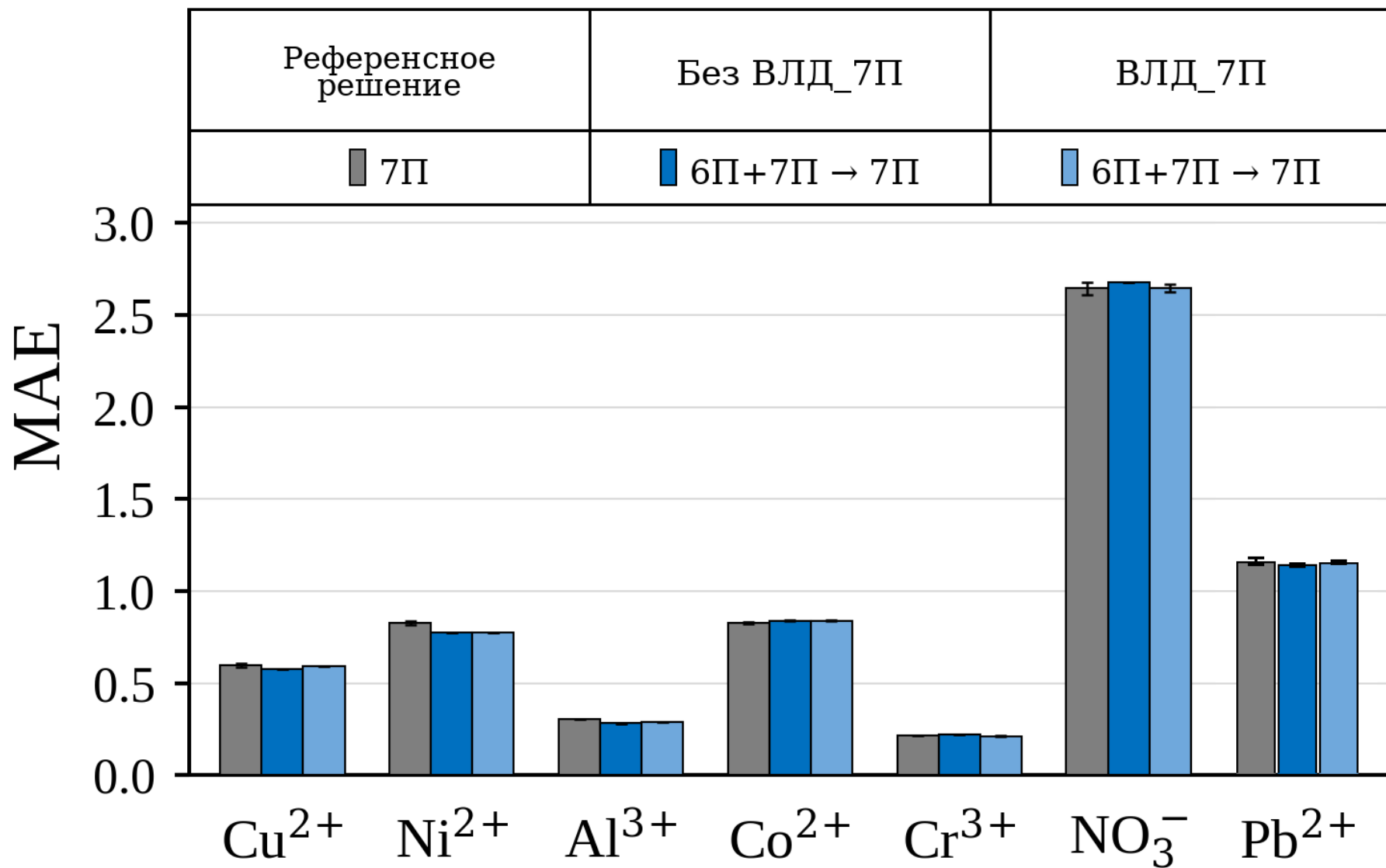
$$L(t) = \alpha(t) \cdot L_{\text{6П}} + (1 - \alpha(t)) \cdot L_{\text{7П}}, \quad \alpha(t) \downarrow$$

В начале обучения больше учитывается исходный набор **6П**, затем его **вклад уменьшается**, модель **постепенно** переходит к настройке на **целевую** задачу **7П**

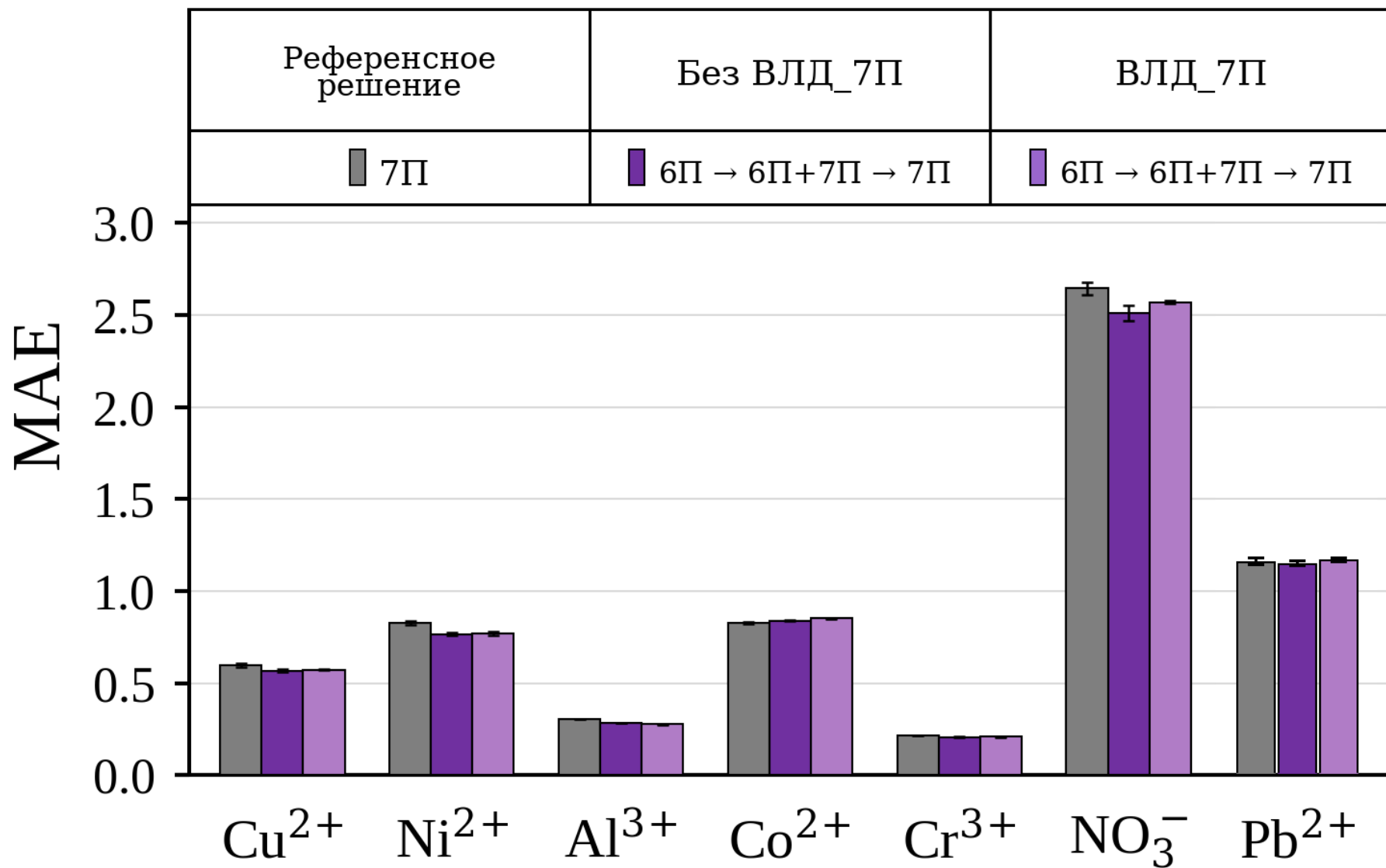
Результаты MAE: 6П → 7П



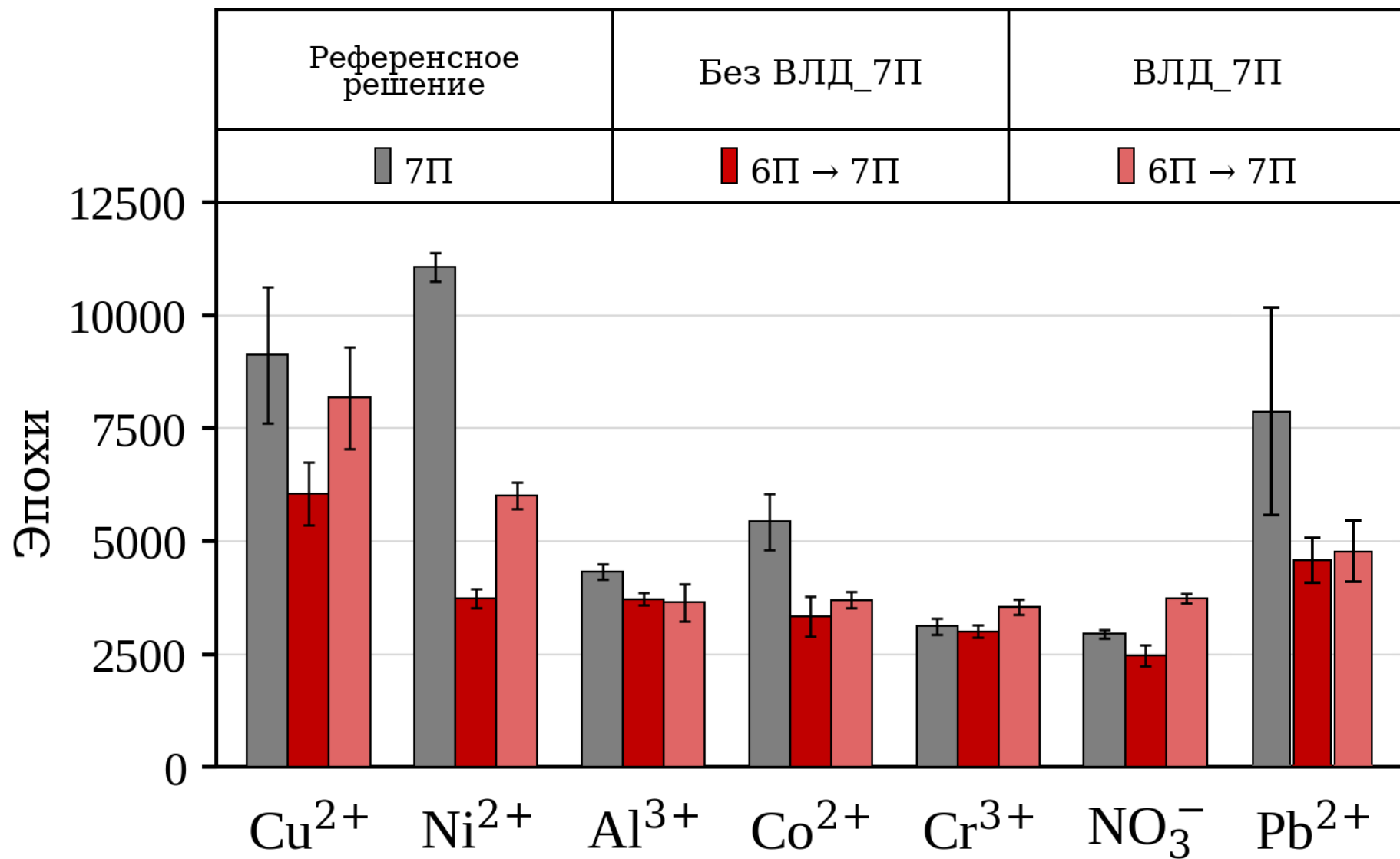
Результаты MAE: 6П+7П → 7П



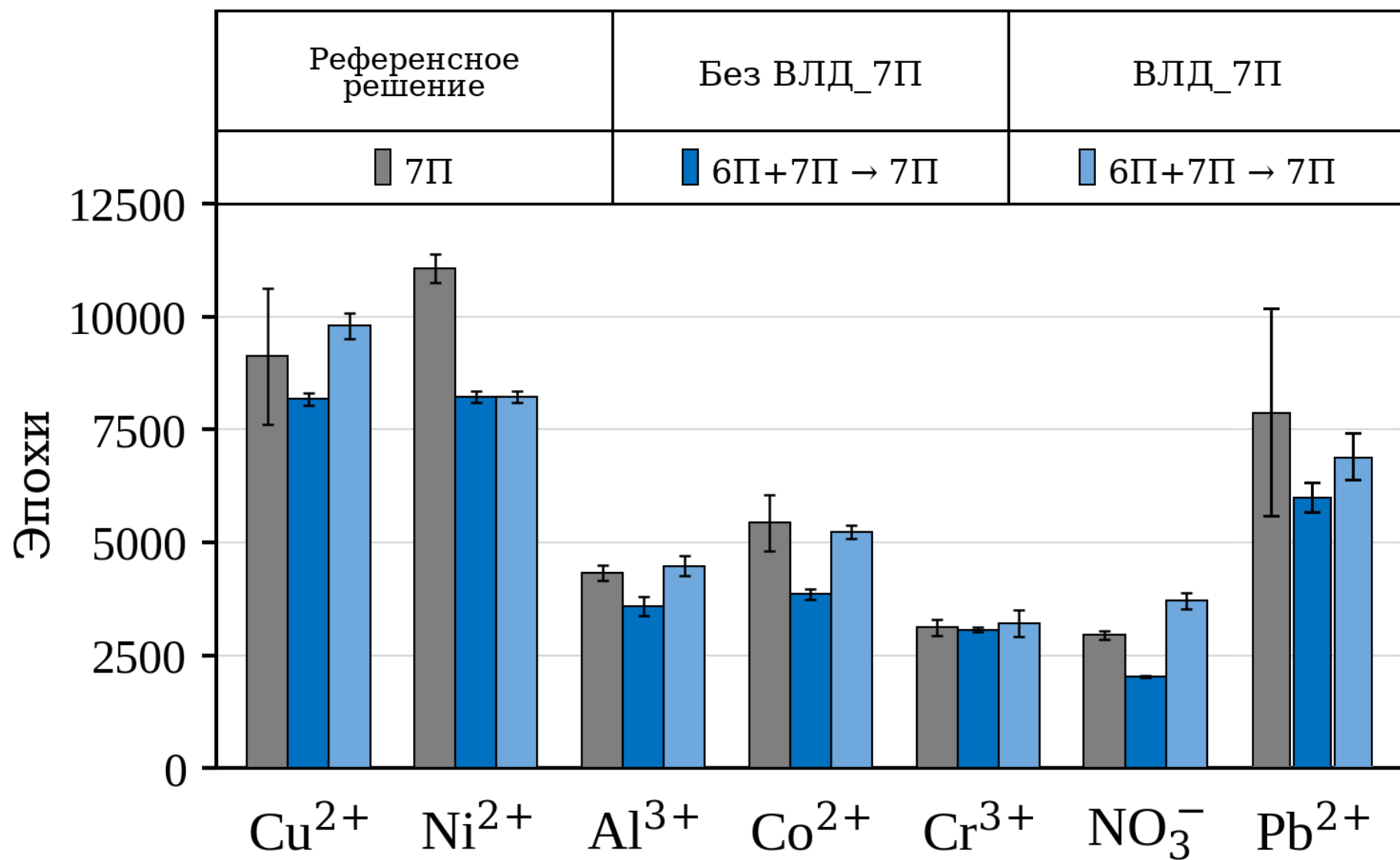
Результаты MAE: 6П → 6П+7П → 7П



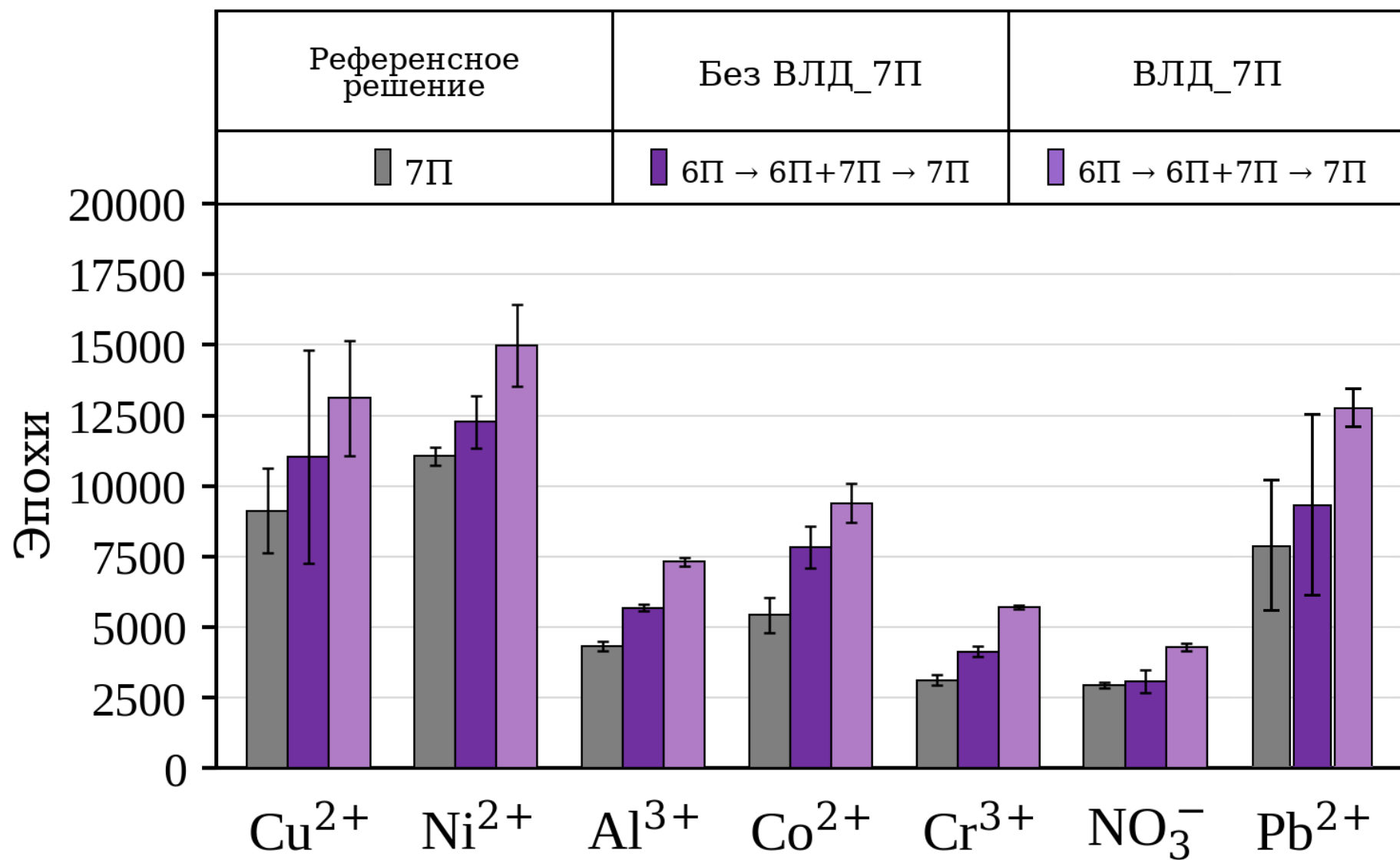
Результаты по эпохам: 6П → 7П



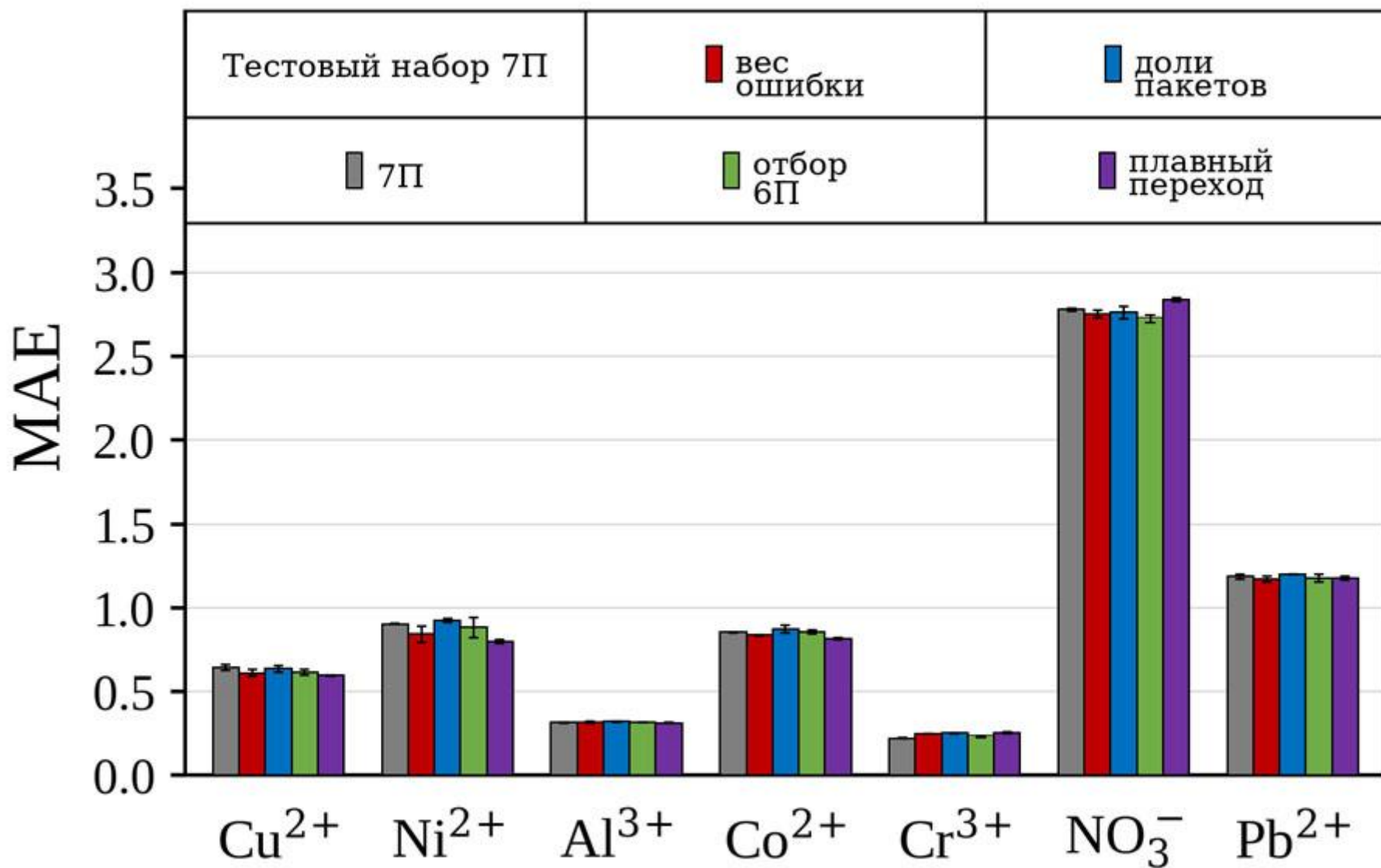
Результаты по эпохам: 6П+7П → 7П



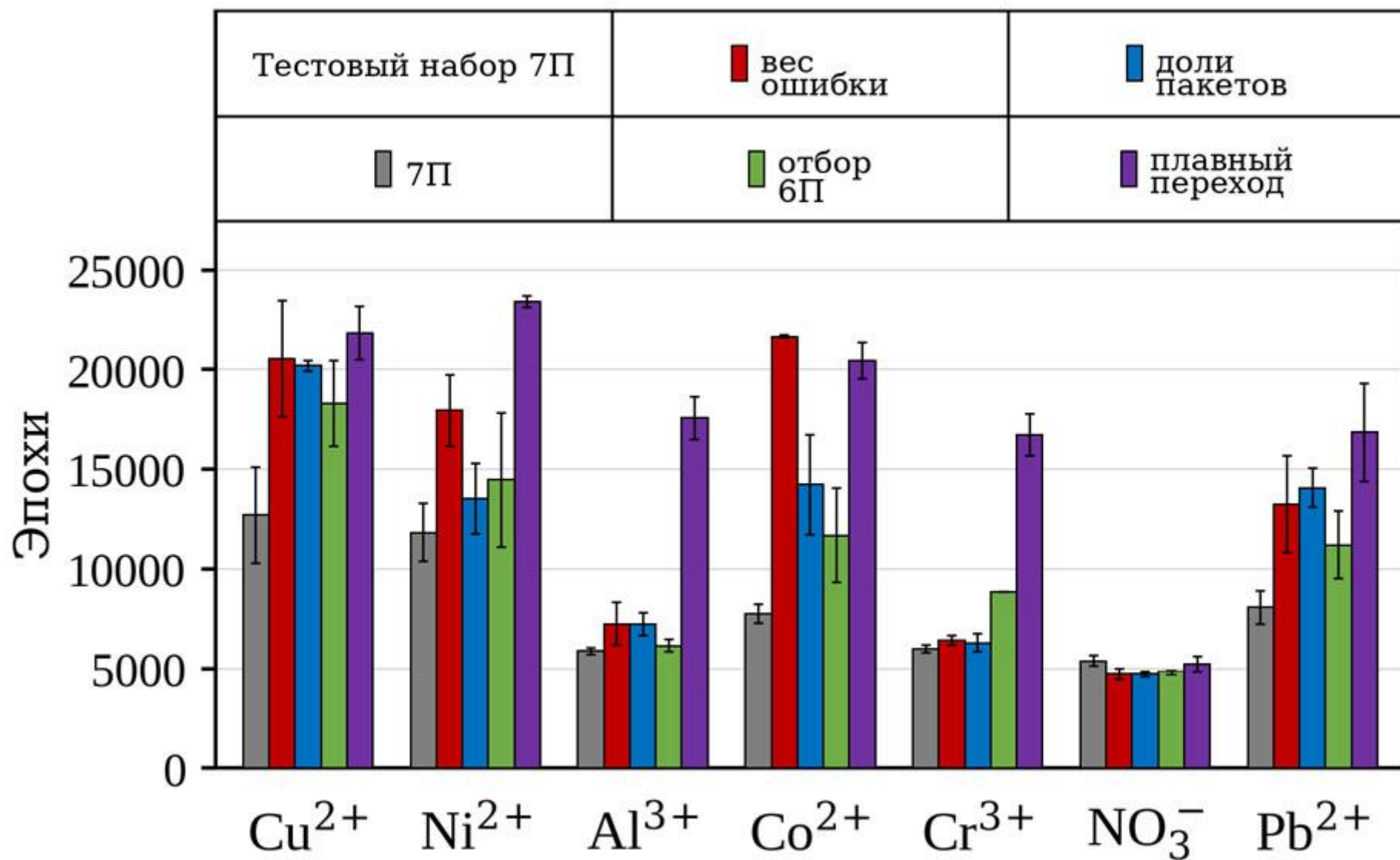
Результаты по эпохам: 6П → 6П+7П → 7П



Результаты: перераспределение весов



Результаты: перераспределение весов



Выводы

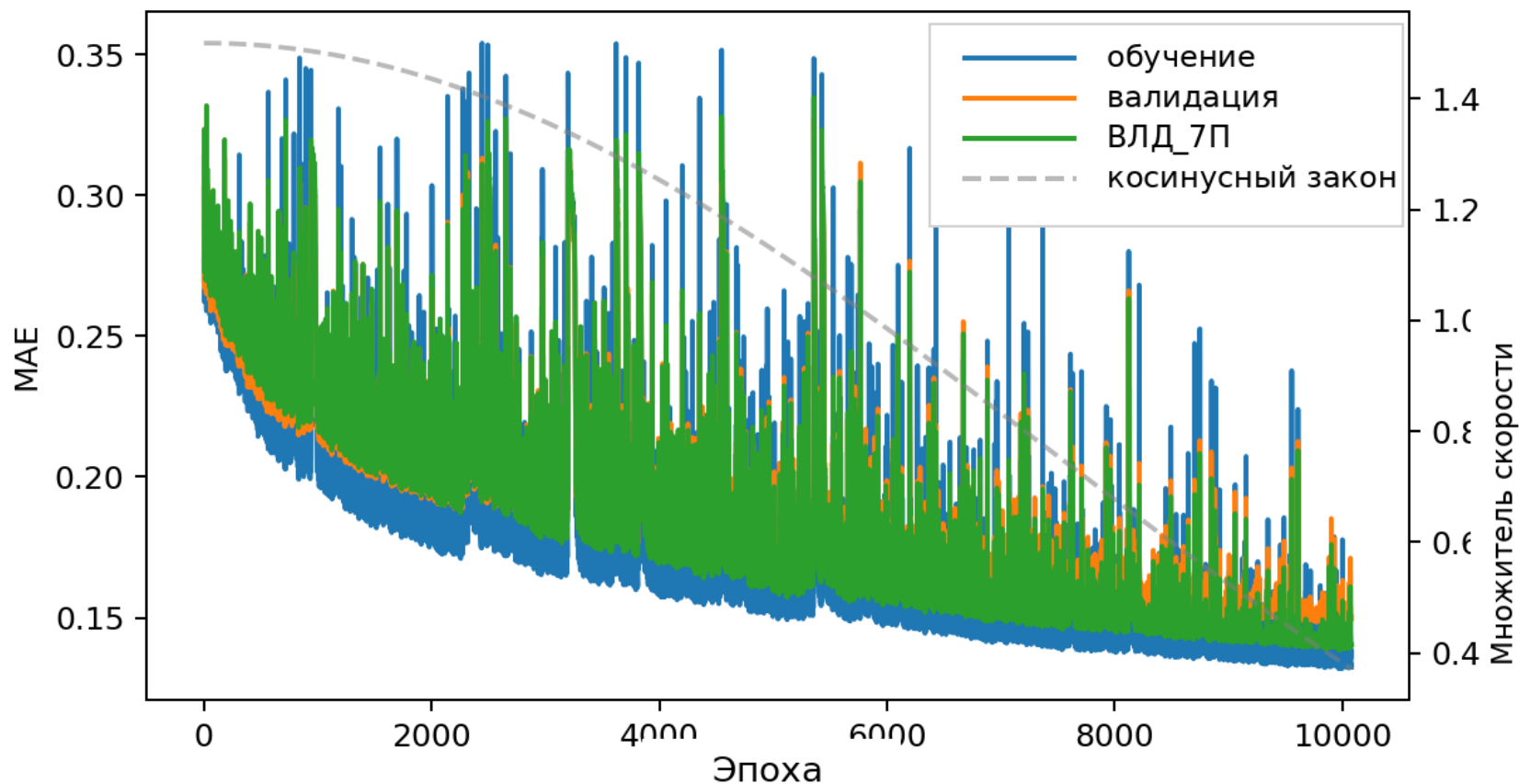
1. Совместное использование данных 6П и 7П при переносе обучения улучшает MAE для большинства исходных параметров 7П задачи
2. Для нового параметра Pb^{2+} результат сильно зависит от выбранного исходного иона: лучшие источники — Cu^{2+} и Ni^{2+}
3. Второй валидационный набор не даёт универсального выигрыша по качеству, однако обычно увеличивает длительность обучения моделей.

Выводы

4. Поэтапное дообучение $6П \rightarrow 6П+7П \rightarrow 7П$ даёт хорошее качество, но требует больше эпох из-за суммы этапов
5. Перераспределение весов примеров позволяет использовать $6П$ не только как предварительный этап, но и прямо при обучении на смеси $6П+7П$
6. Отбор примеров $6П$ показал лучший компромисс между качеством и числом эпох среди методов перераспределения весов

СПАСИБО ЗА ВНИМАНИЕ!

Косинусный закон изменения скорости обучения



$$m(t) = m_{\min} + 0.5 \cdot (m_{\max} - m_{\min}) \cdot (1 + \cos(\pi t/T))$$

$m(t)$: $1.5 \rightarrow 0.01$; скорость каждого слоя = базовая скорость слоя $\times m(t)$