**6. Оцінка концентрації заліза за кінетикою струму короткого замикання з використанням моделей комп’ютерного зору та трансферного навчання**

У попередніх розділах була продемонстрована можливість кількісної оцінки концентрації заліза в КСЕ за допомогою як темнових, так і світлових вольт-амперних характеристик з використанням алгоритмів машинного навчання. Проте, навіть за високої точності таких моделей, вони залишаються обмеженими у випадках, коли обсяг доступних навчальних даних є малим. Більше того, моделювання нерідко є дуже вимогливим з точки зору часових та розрахункових затрат, тому у цьому розділі буде розглядатися альтернативний підхід, що базується на використанні трансферного навчання з попередньо навченими моделями комп’ютерного зору (КЗ) для прогнозування концентрації заліза в КСЕ на основі кінетичних залежностей після дисоціації пар FeB.

**6.1 Збір та попередня обробка даних**

На рис.6.1 наведено схематичне зображення методики, що передбачає використання залежності для визначення концентрації домішкового заліза на основі трансферного навчання, моделей КЗ та ММН.

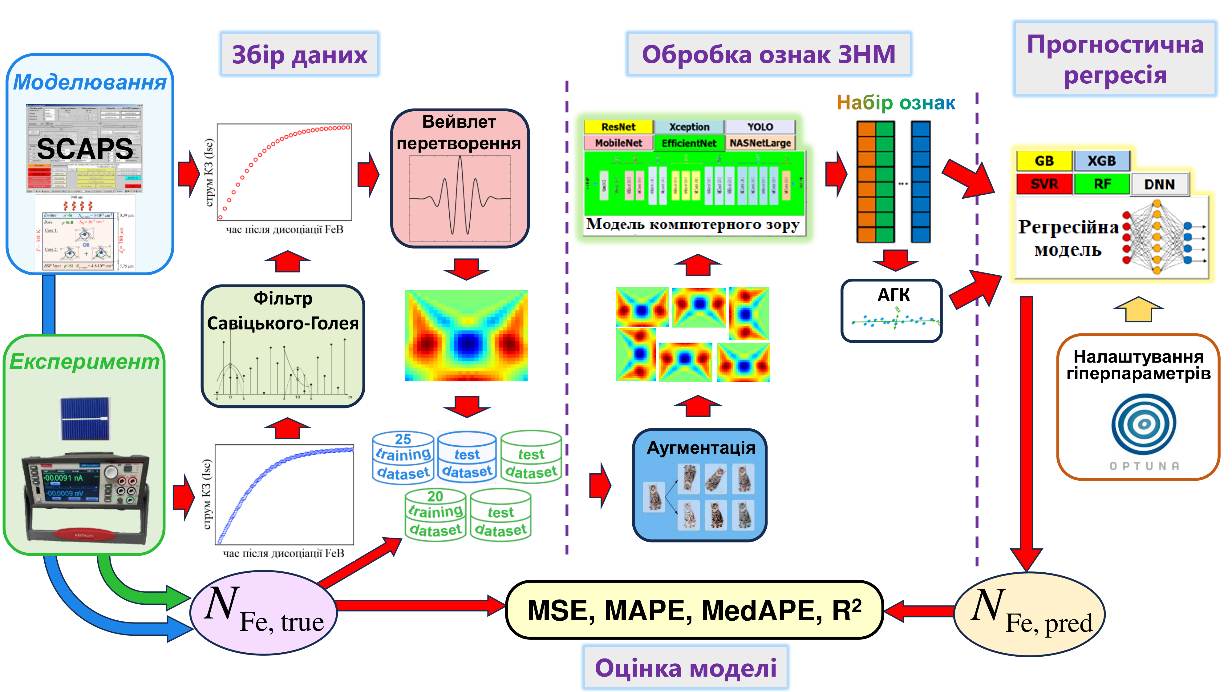


Рис. 6.1.Схематичне зображення методики визначення концентрації заліза в КСЕ на основі трансферного навчання, моделей КЗ та ММН

Схема складається з кількох основних етапів: збору даних, попередньої обробки ознак ЗНМ, використання алгоритмів машинного навчання для передбачення концентрації заліза (задача регресії) та оцінки ефективності ММН за допомогою метрик: MSE, MAPE, MedAPE та .

Не першому етапі відбувається моделювання або експериментальне вимірювання часових залежностей струму короткого замикання КСЕ після індукованого розпаду пар FeB. У випадку експериментальних кривих дані були згладжені за допомогою фільтра Савіцького-Голея [Krishnan2013]. Після цього проводилося неперервне вейвлет перетворення [Torrence1998] одномірних часових залежностей у двовимірні спектрограми, що представляються у вигляді зображень, кожна точка яких відповідає амплітуді вейвлет коефіцієнта для певних моментів часу та частоти. В дослідженні використовувався вейвлет Морле, процедура перетворення здійснювалася за допомогою Python toolkits PyWavelets, приклади зображень представлені на Рис.6.2. Отримане зображення було аугментовано: використовувалися перевертання вздовж вісей OX та OY, а також повороти зображень на 90°, 180°, та 270°. Така аугментація дозволяє збільшити точність прогнозів ММН, особливо у випадку невеликих за розміром наборів даних [Ahmad2020].

Під час обробки ознак ЗНМ всі зображення (вихідне та аугументовані) оброблювалися за допомогою однієї зі стандартних моделей КЗ, що дозволяло отримати набори ознак для кожного із зображень. Причому, жодного додаткового налаштування моделей ЗНМ не проводилося, вони використовувалися у незмінному вигляді після скачування. Загалом, розмірність ознак, які можна отримати на виході моделей ЗНМ суттєво перевищує кількість використаних зразків і тому, що всі вони не можуть бути незалежними. Тому для порівняння результатів у частині випадків застосовувався АГК.

Отримані набори ознак були використані як вхідні параметри регресійної моделі, яка базувалися на одному з стандартних алгоритмів машинного навчання (RF, GB, XGB, SVR, DNN), та мала на меті передбачити концентрацію заліза в сонячному елементі.

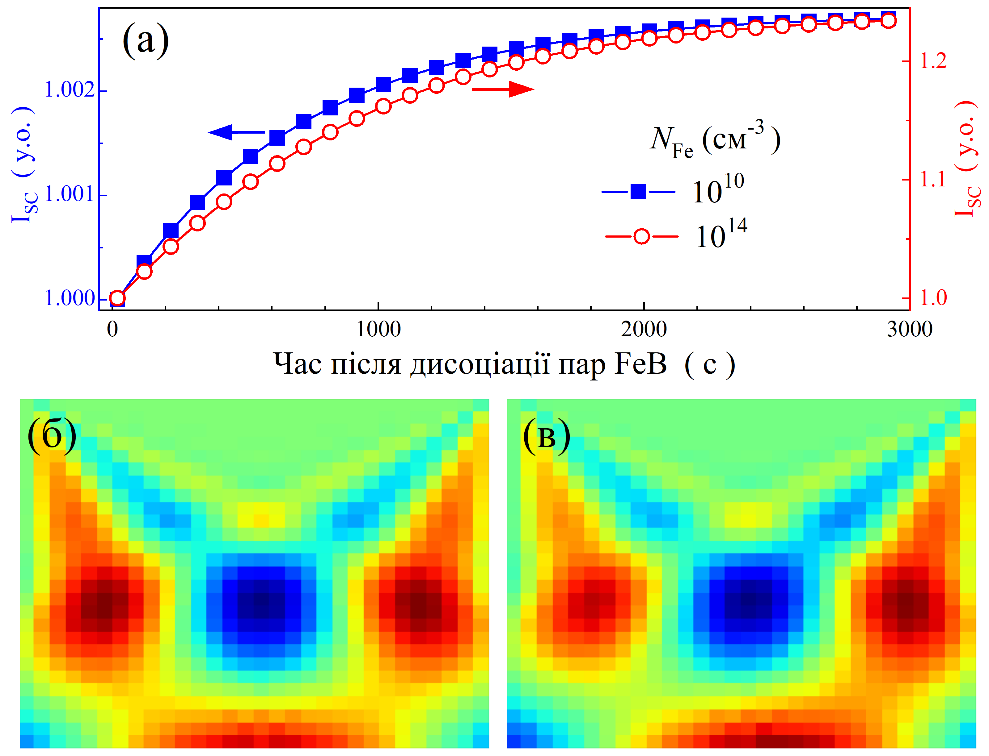


Рис.6.2. Змодельовані часові залежності струму короткого замикання (а) та відповідні вейвлет-спектрограми для концентрацій заліза (б) та (в). Дані на графіку (а) позначенні квадратами для , що відповідає спектрограмі (б), та колами для , що відповідає спектрограмі (в).

Під час тренування регресійних моделей набори ознак, що відповідали вихідній вейвлет спектрограмі та аугументованим зображенням розглядалися як окремі зразки; під час тестування як остаточний прогноз використовувалося медіанне значення з набору прогнозів значень для набору зображень, отриманих після впровадження методів аугментації. Зауважимо, що в одному випадку моделі навчалися на змодельованому наборі даних та тестувалися як на даних, отриманих в результаті моделювання, так і в результаті експериментальних вимірювань. В другому випадку частина експериментальних результатів використовувалася для навчання, а інша – для тестування відповідних моделей.

**6.2 Моделювання та експериментальні вимірювання**

Для отримання залежностей , було проведено моделювання ВАХ для кремнієвої структури під монохроматичним освітленням за допомогою пакету SCAPS 3.3.11. Під час моделювання вважалося, що товщина бази КСЕ становила 380 мкм, легуючою домішкою був бор з концентрацією , температура становила 340 К, для освітлення використовувалося світло з довжиною хвилі 940 нм та інтенсивністю 5 Вт/м2. Одним із параметрів моделювання була загальна концентрація домішкових атомів заліза . Вважалося, що атоми рівномірно розподілені по об’єму бази та шару сонячного елементу і можуть перебувати як в міжвузловому положенні (концентрація ), так і формувати пари FeB (концентрація ). Залежність від часу після дисоціації пари FeB описувалася рівнянням [MurphyJAP2011]:

|  |  |
| --- | --- |
|  | (6.1) |

де - концентрація міжвузлових атомів заліза, що утворилися внаслідок дисоціації пар FeB, ; - частка міжвузлових атомів заліза, що залишаються неспареними в рівноважному стані , згідно з [MurphyJAP2011]:

|  |  |
| --- | --- |
|  | (6.2) |

де - енергія зв'язку пар FeB (прийнята за 0,582 еВ [Wijaranakula]), залежить від кількості можливих орієнтацій пари та щільності гратки (прийнята як [MurphyJAP2011]), - рівень Фермі, - положення рівня донора відносно максимуму валентної зони (прийняте як 0,394 еВ [Moller2014]), - характерний час комплексної асоціації, згідно з [Moller2014]:

|  |  |
| --- | --- |
|  | (6.3) |

де — енергія міграції атомів (прийнята за 0,66 еВ [Moller2014]), - константа (прийнята за [Tan2011]). Концентрація пари залізо-бор була оцінена з рівняння . Загалом, концентрації дефектів, пов'язаних із залізом, залежали не тільки від часу, але й від їхнього просторового положення в структурі, що відображало неоднорідність різниці ().

Для створення тренувального набору було використано 25 значень , рівномірно розподілених у логарифмічному масштабі від до . Приклади отриманих залежностей представлені на рис.6.2. Там же показані відповідні вейвлет-спектрограми. Змодельований тестовий набір даних складався з 10 залежностей, розрахованих для значень , які не збігаються з тими, що входять до тренувального набору.

Методика експериментальної частини була описана в розділі 2. Кінетика струму короткого замикання вимірювалася в темряві при температурі 340 К на протязі 3000 с - відповідно до рівняння 6.3 цього інтервалу цілком достатньо для повного відновлення пар до рівноважної концентрації. Приклад виміряної залежності представлено на Рис.6.3а. Отриманий сигнал містить невеликий шум, який пов’язаний з тим, що незважаючи на використання термостату, температура LED коливалася в діапазоні шириною близько 0.4 К. Для згладжування сигналу було застосовано фільтр Савіцького-Голея, причому довжина вікна та порядок фільтра були обрані адаптивно відповідно до [Krishnan2013]. Для вейвлет-перетворення були збережені тільки поточні значення, що відповідають часовим точкам, використаним у моделюванні. Згладжена крива показана на рисунку 3а, а на інших панелях рисунка показані спектрограми, отримані з необробленої експериментальної кривої та обробленої залежності. Крім того, для вейвлет перетворення використовувалися значення струму лише в ті ж самі моменти часу, для яких проводилося моделювання. Приклад згладженої кривої представлений на Рис.6.3а, а на інших частинах цього рисунку показані спектрограми, отримані для необробленої та обробленої експериментальної кривої.

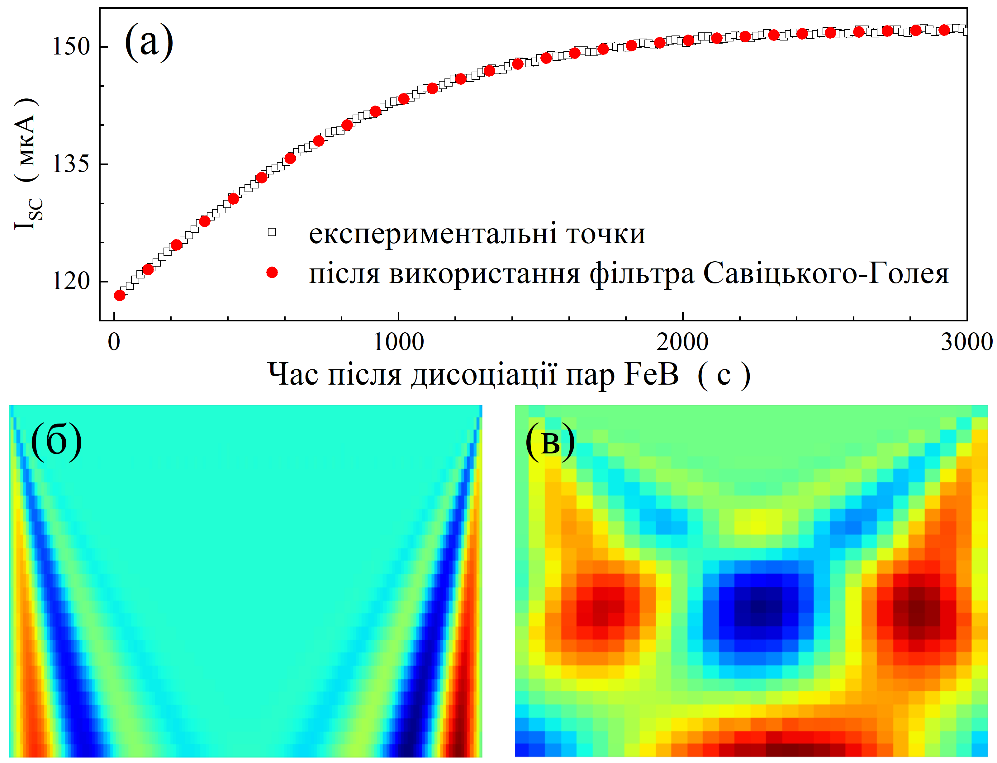


Рис.6.3. (a) Експериментально виміряна залежність струму короткого замикання від часу для зразка з (a, незафарбовані квадрати) і та ж залежність після застосування фільтра Савіцького-Голея (зафарбовані кола). Вейвлет спектрограми (б) і (в) відповідають кривим із зафарбованими квадратами та незафарбованими колами відповідно з рис.6.2.

Всього було розглянуто 28 зразків з концентрацією заліза від до . Для тестування моделей, натренованих на змодельованих даних використовувався набір, що містив всі експериментальні дані. У випадку, коли модель тренувалася на основі експериментальних результатів, тренувальний набір містив 20 випадково вибраних зразків, а тестовий, відповідно, 8.

**6.3 Моделі комп’ютерного зору та регресійні алгоритми**

Для отримання графічних ознак вейвлет спектрограм були використанні моделі комп’ютерного зору з пакету Keras, а саме EfficientNetB7, ResNet152V2, MobileNetV2, Xception та NASNetLarge. Хоча ці моделі мають різну архітектуру, всі вони належать до класу ЗНМ, які призначені для класифікації об'єктів і раніше успішно застосовувалися для обробки ЕЛ [Jia2024, tella2025]. Для всіх моделей розглядалися два варіанти виділення ознак: для першого розподіли ймовірностей для конкретних класів передавалися на наступний етап методики (див. рис.6.1), а в другій використовувалися необроблені вектори ознак, безпосередньо вилучені моделлю комп'ютерного зору. Крім того, було використано модель CSPDarknet53, яка слугувала основою ЗНМ для YOLOv4. Моделі цієї родини мають більш складну архітектуру ЗНМ, яка оптимізована не для класифікації окремих об'єктів, а для виявлення декількох об'єктів на зображеннях [Liu2024a]. Використана модель повертає три шари ознак, для подальшої обробки ми використовувати або лише найвищий рівень, або два останніх.

Для зменшення впливу надлишкових даних був застосований АГК. У цьому дослідженні були обрані головні компоненти, що пояснюють не менше 99,9% загальної дисперсії в оригінальних ознаках, що дозволило досягти істотного зменшення розмірності ознак. Ця процедура попередньої обробки була вибірково застосована до підмножини моделей комп'ютерного зору - зокрема, тих, що демонструють хорошу продуктивність на тестових наборах без АГК - з метою оцінки ефективності цього підходу. З огляду на надзвичайно високу розмірність ознак, що генеруються YOLOv4, було досліджено можливість застосування альтернативної техніки зменшення розмірності. Зокрема, до кожної згорткової карти ознак було застосовано глобальне усереднення, замінивши просторову карту її середнім значенням, що дало одне скалярне значення на канал. Конфігурації моделей комп'ютерного зору, використані в цьому дослідженні, наведено в таблиці 6.1. У таблиці також наведено позначення, які згодом використовуються для посилання на ці конфігурації.

Регресійні моделі, що використовувалися в дослідженні були такі самі як і в розділі 5. Кожна з цих моделей використовувалася для роботи з ознаками, отриманими за допомогою кожного з варіантів, вказаних в Табл.6.1. Виняток становили лише нестиснуті ознаки, отримані за допомогою YOLOv4, для яких використані розрахункові потужності дозволили використовувати лише SVR.

Таблиця 6.1. Використані моделі ЗНМ та варіанти вилучення ознак

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Базова модель | Тип моделі | Обробка | Вихідна розмірність | Маркування |
| EfficientNetB7 | Класифікатор | немає | 1000 | ENB7:CL |
|  | Витягувач ознак | немає | 2560 | ENB7:FE |
|  |  | АГК | 39 | ENB7:FE:P |
| MobileNetV2 | Класифікатор | немає | 1000 | MNV2:CL |
|  | Витягувач ознак | немає | 1280 | MNV2:FE |
|  |  | АГК | 124 | MNV2:FE:P |
| NASNetLarge | Класифікатор | немає | 1000 | NAS:CL |
|  |  | АГК | 30 | NAS:CL:P |
|  | Витягувач ознак | немає | 4032 | NAS:FE |
| ResNet152V2 | Класифікатор | немає | 1000 | R152:CL |
|  | Витягувач ознак | немає | 2048 | R152:FE |
| Xception | Класифікатор | немає | 1000 | XCP:CL |
|  | Витягувач ознак | немає | 2048 | XCP:FE |
| YOLOv4 (CSPDarknet53) | Витягувач ознак (верхній шар, необроблені дані) | немає | 86528 | YL:FE1 |
|  |  | АГК | 137 | YL:FE1:P |
| CSPDarknet53 (основа YOLO) | Витягувач ознак (верхній і передостанній шари, необроблені дані) | немає | 433640 | YL:FE2 |
|  |  | АГК | 142 | YL:FE2:P |
|  | Витягувач ознак (верхній шар, усереднені ознаки) | немає | 512 | YL:FP1 |
|  | Витягувач ознак (верхній і передостанній шари, усереднені ознаки) | немає | 1024 | YL:FP2 |

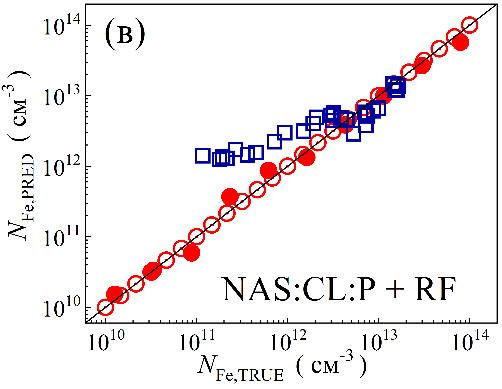
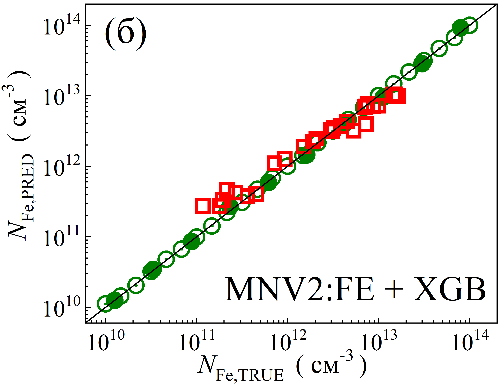
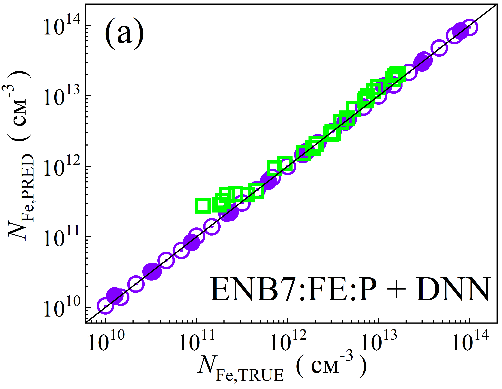
Цільовою змінною моделей був . Вхідні та цільові змінні були нормалізовані, щоб мати середнє значення нуль і стандартне відхилення один над навчальним набором. Для кожного сценарію регресійні моделі були оптимізовані з метою підвищення точності прогнозування. Налаштування гіперпараметрів виконувалося із використанням програмного пакета Optuna, у якому застосовано вибірку за алгоритмом TPE (Tree-structured Parzen Estimator) та метод обрізання Hyperband для забезпечення ефективного відбору оптимальних моделей. Повний перелік налаштованих гіперпараметрів та відповідні діапазони пошуку наведено в таблицях S6.1–S6.5 в додаткових матеріалах (див. Висновки до розділу 6). Під час налаштування моделей була реалізована п'ятикратна перехресна перевірка, причому 20% навчальних даних було використано як валідаційний набір для оцінки моделей, навчених на решті 80%. Отримані раціональні комбінації гіперпараметрів наведено в таблицях S6.6–S6.10.

В результаті було досліджено 87 різних комбінацій моделей комп'ютерного зору та регресійних моделей. Кожна комбінація була згодом навчена та оцінена з використанням як змодельованих, так і експериментальних даних. Для ідентифікації результатів у кожному випадку було використано комбіновану мітку, отриману з останнього стовпця таблиці 6.1 та скороченої назви алгоритму регресії.

**6.4 Апробація моделей на змодельованих залежностях**

Для побудови надійних регресійних моделей були використанні метрики ефективності, що використовувалися в розділі 5: середня квадратична похибка (MSE), середня абсолютна відсоткова похибка (MAPE), медіана абсолютної відсоткової похибки (MedAPE) та коефіцієнт детермінації (R2).

На рис.6.4 показані типові результати прогнозування, отримані для моделей, навчених на основі змодельованого навчального набору даних. Повний набір результатів, що охоплює всі 87 досліджених конфігурацій, наведено на рис.S6.1 у додаткових матеріалах. Рис.6.5 відображає результати MAPE та R2, які були отримані при застосуванні моделей до тренувального набору даних.



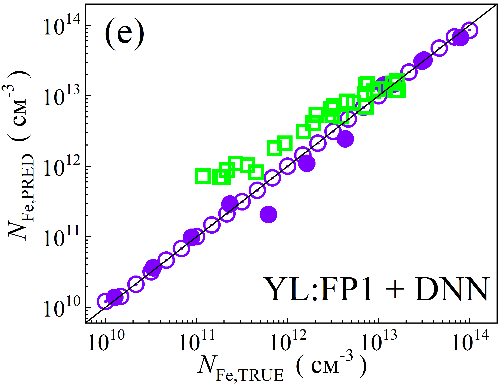
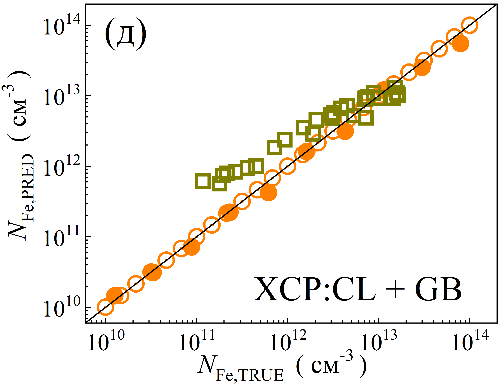
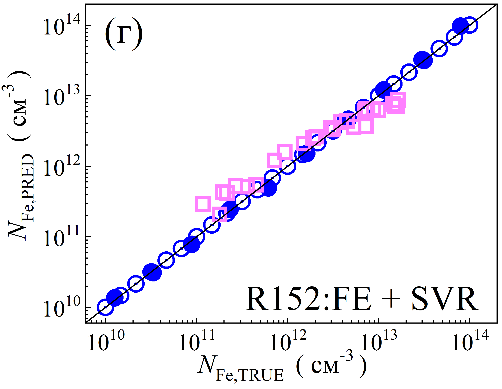


Рис.6.4. Точкові діаграми, що порівнюють еталонні концентрації заліза з прогнозованими значеннями , отриманими за допомогою векторів ознак, витягнутих із різних моделей КЗ у поєднанні з різними алгоритмами регресії. ММН були навчені на змодельованому наборі даних. Незафарбовані кола відповідають фазі навчання, тоді як зафарбовані кола та незафарбовані квадрати - фазі тестування для змодельованих та експериментальних наборів даних. Чорні лінії - слугують еталонами.

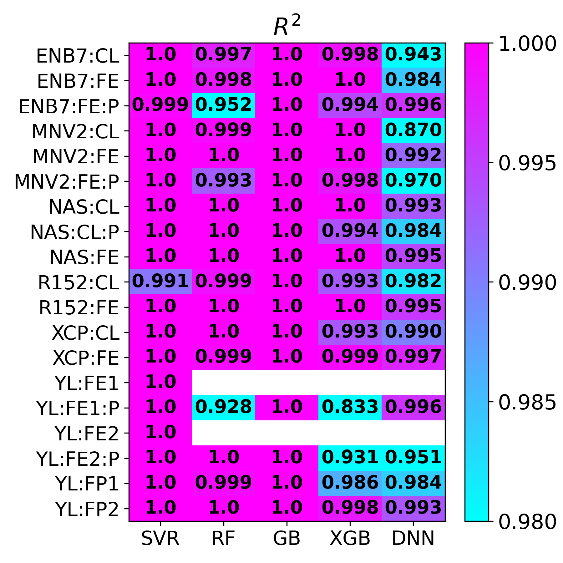
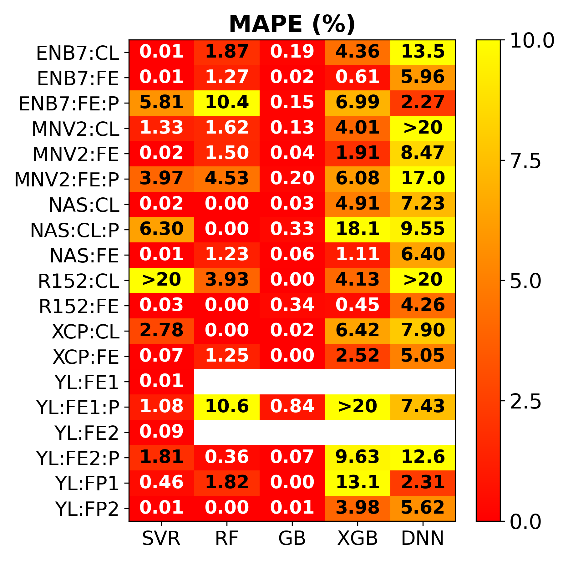


Рис.6.5. Середня абсолютна відсоткова похибка (лівий рисунок) та коефіцієнт детермінації (правий рисунок) для різних комбінацій моделей КЗ (вертикальна вісь) та регресійних моделей (горизонтальна вісь) під час етапу навчання. Моделі були навчені з використанням змодельованого набору даних.

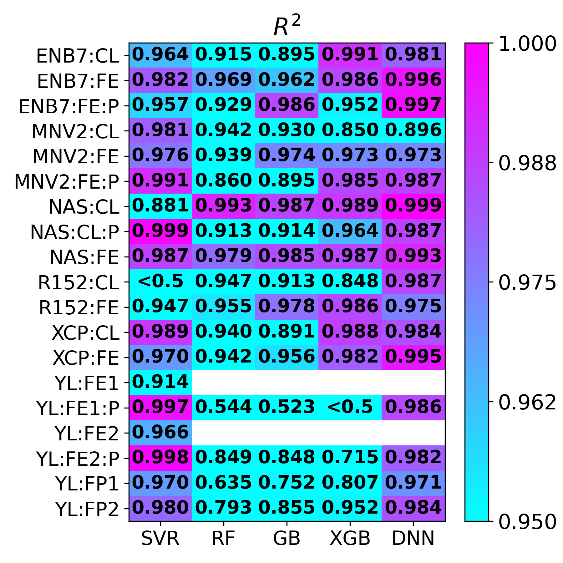
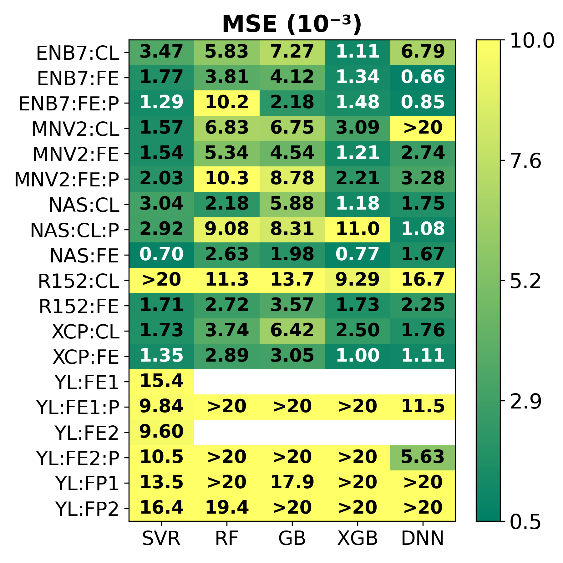
З величинами MedAPE та MSE можна ознайомитися в додаткових матеріалах (див. рис.S6.2). На рис.6.4 та рис.6.5 насамперед можна побачити, що незважаючи на дуже малий розмір тренувального набору (25 зразків), більшість моделей навчається достатньо якісно: у багатьох випадках середня відносна похибка менше або в околі 1% і дуже в небагатьох випадках вона перевищує 10%, тоді як R2 менший за 0,980 лише у 8 випадках з 87. Значення MedAPE як правило не перевищує MAPE, а в більшості випадків є навіть меншою (див. Рис.S6.2). Такі високі показники метрик свідчать, що 1) вейвлет-перетворення дозволили отримати дійсно інформативні зображення, в яких дійсно закодована інформація про концентрацію ; 2) Моделі КЗ змогли витягнути ознаки, які пов’язані з концентрацією .

Серед регресійних моделей найкращі показники у GB та SVR, найгірші – при використанні DNN. Це цілком очікувано, так як GB та SVRдобре працюють у випадках із обмеженим числом прикладів та ознаками з низьким рівним шуму (що характерно для змодельованих залежностей). Водночас нейронні мережі містять велику кількість параметрів і тому не демонструють стабільного узагальнення.

Серед моделей КЗ найкращі показники у EfficientNetB7 та NASNetLarge, найгірші – у ResNet152V2 та YOLOv4. Це може бути пов’язано з тим, що перші дві достатньо сучасні архітектури орієнтовані на вилучення універсальних ознак і добре пристосовані до вейвлет-спектрограм. ResNet152V2 більш орієнтована саме на класифікацію об’єктів, тоді як YOLO менш пристосована до регресії за глобальним патерном (закономірністю) зображення, та зосереджує зусилля на окремих елементах.

Також можна бачити, що використання у якості дескрипторів ймовірностей класів погіршує якість прогнозів порівняно з випадками, коли використовуються безпосередньо ознаки рисунків. Більше того, застосування АГК, навіть незважаючи на те, що воно зберігає 99,9% дисперсії, призводить до зменшення точності прогнозування. Це спостереження свідчить про те, що патерни зображень, пов'язані з коливаннями значення концентрації заліза, можуть становити лише незначну частину загальної дисперсії даних.

Можливість моделей добре навчатися на тренувальному наборі, звичайно, є необхідною передумовою для ефективної їхньої роботи з незнайомими даними, проте не може гарантувати високої точності передбачень. Тому перевірка моделей на тестових даних є необхідністю. Цей крок стає особливо важливим, коли навчальний набір є невеликим, оскільки результати тестування надають основні підтвердження загальності моделі. На рис.6.4 та рис.6.6 представлені результати передбачень, отримані для тестового набору, що створювався на основі змодельованих даних (більш повні версії доступні на Рис.S6.1 та Рис.S6.3).



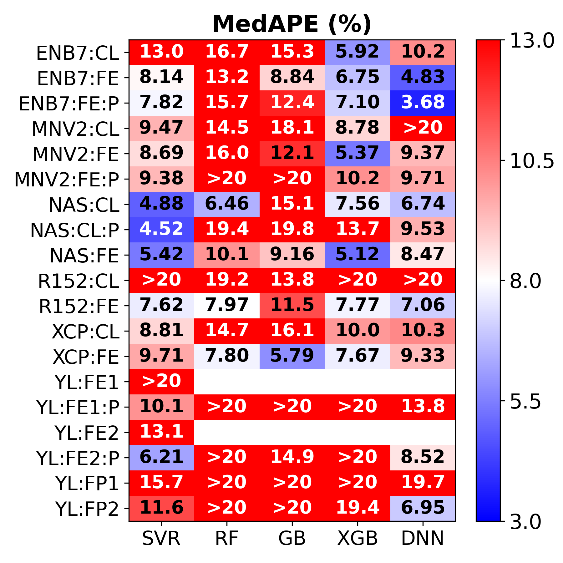
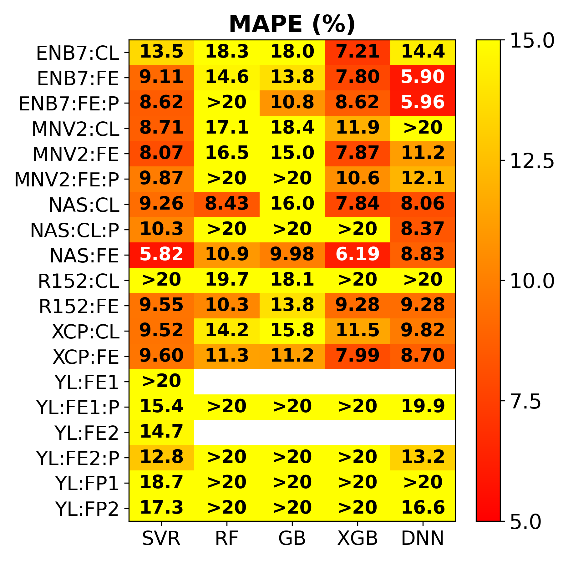


Рис.6.6. Середньоквадратична похибка, коефіцієнт детермінації, середня абсолютна відсоткова похибка та медіанна абсолютна відсоткова похибка для різних комбінацій моделей КЗ (вертикальна вісь) та регресійних моделей (горизонтальна вісь) під час тестової фази з використанням змодельованого набору даних. Моделі були навчені з використанням змодельованого навчального набору даних.

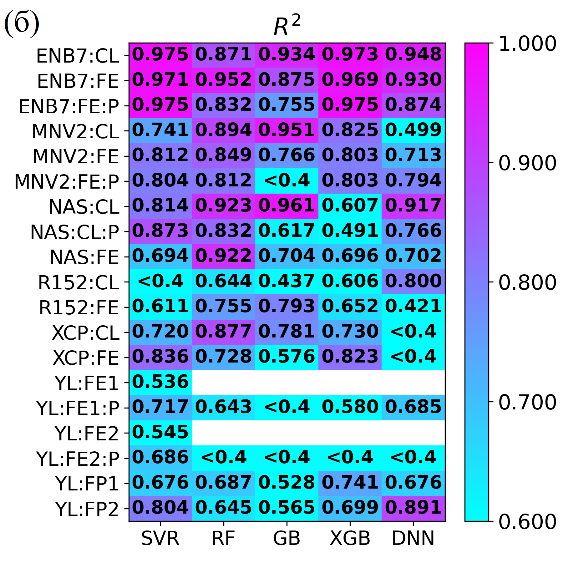
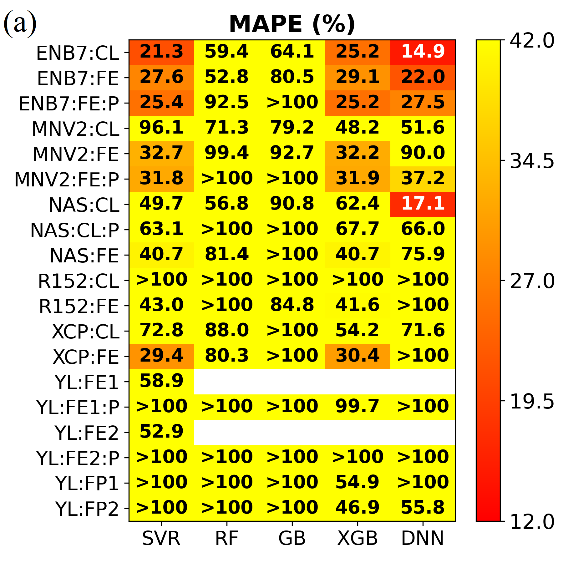
Як це цілком очікувалося, показники передбачень погіршилися. Проте зменшення точності було неоднаковим для різних регресійних алгоритмів. Зокрема, у випадку DNN це зменшення було мінімальним, в результаті чого для цього алгоритму спостерігаються найкращі результати. Найгірші показники були отримані для RF і GB, тоді як SVR і XGB показали результати дещо гірші, ніж DNN, але з відносно невеликим відривом. Таку поведінку можна пояснити здатністю DNN плавно апроксимувати неперервні залежності. На відміну від GB та RF, які створюють набір локальних правил, алгоритм нейронних мереж передбачає формування більш гладкої поверхні у просторі ознак, що сприяє інтерполяції для значень концентрацій заліза, які відсутні у тренувальному наборі. З цієї точки зору XGB та SVR займають проміжну позицію, що і зумовлює їхнє незначне відставання від DNN.

Серед моделей КЗ, архітектури, що мали найвищу та найнижчу продуктивність, залишилися такими: застосування EfficientNetB7 та NASNetLarge дозволяє отримати найкращі результати (тобто ці моделі дійсно виявляють характеристики вейвлет-зображень, що пов’язані з ), ResNet152V2 та YOLOv4 - найгірші. Виняток становлять конфігурації, в яких особливості YOLOv4 з двох шарів поєднуються з DNN або SVR. У цих випадках показники продуктивності значно покращилися. Це покращення свідчить про те, що використання більшої кількості ознак дозволило зафіксувати більш різноманітні закономірності, що в поєднанні з гнучкими регресорами частково пом'якшило обмеження, які спостерігалися в окремій моделі YOLO.

Цікаво, що застосування АГК у випадку тестового набору не суттєво погіршувало результати (а в окремих випадках і поліпшувало). Тобто, АГК не завжди шкідливий, але виграш від нього невеликий і непередбачуваний. Відмінності у показниках моделей, які використовувати ознаки класів та ознаки зображень теж не настільки великі, як у випадку тренувального набору даних, хоча використання нестиснених представлень продовжує мати перевагу.

Найкращими комбінаціями є ENB7:FE+DNN (MAPE = 5.90 %, MedAPE = 4.83 %, R2 = 0.996), ENB7:FE:P+DNN (MAPE = 5.96 %, MedAPE = 3.68 %, R2 = 0.997), NAS:FE+SVR (MAPE = 5.82 %, MedAPE = 5.42 %, R2 = 0.987), NAS:FE+XGB (MAPE = 6.19 %, MedAPE = 5.12 %, R2 = 0.987) та NAS:CL:P+SVR (MAPE = 10.3 %, MedAPE = 4.52 %, R2 = 0.999). Це достатньо високі абсолютні показники, що підтверджують зв’язок між вейвлет-спектрограмами та концентрацією заліза.

На рис.6.7а та рис.6.7б наведено метрики результатів застосування моделей, навчених на змодельованих даних, до результатів експериментальних вимірювань.



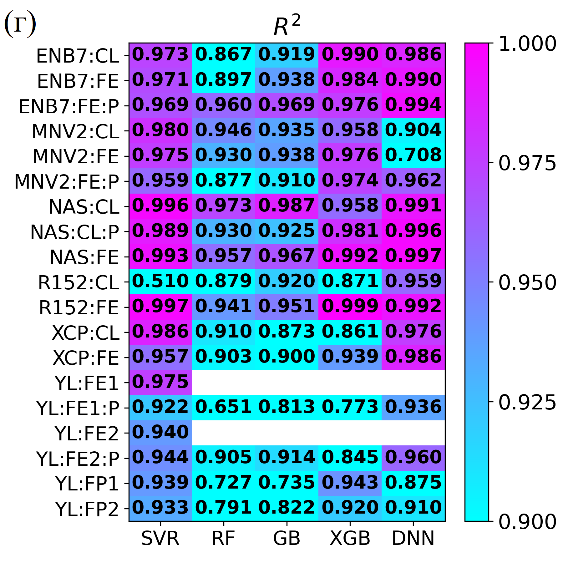
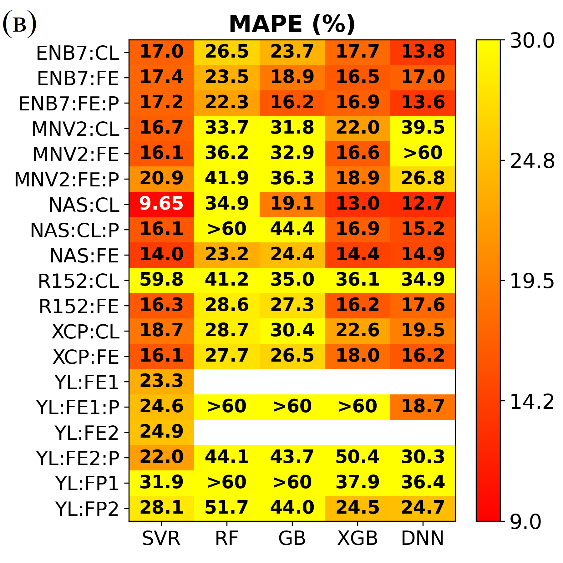


Рис.6.7. Середня абсолютна відсоткова похибка (а, в) та коефіцієнт детермінації (б, г) для різних комбінацій моделей КЗ (вертикальна вісь) та регресійних моделей (горизонтальна вісь) під час тестової фази з експериментальним набором даних без (а, б) та з (в, г) калібруванням після навчання (post-hoc) відповідно до рівняння 6.4. Моделі були навчені з використанням змодельованого набору даних.

Співвідношення прогнозованих та реальних концентрацій показані на рис.6.4, додаткова інформація представлена на рис.S6.1 та рис.S6.4. Як видно з наведених даних, лише в деяких випадках (певних комбінацій EfficientNetB7 та NASNetLarge і DNN та SVR) середня та медіанна похибка знаходяться в діапазоні (15 - 25) %. З одного боку, це не поганий результат, враховуючи, що похибка визначення експериментальним шляхом в околі 10 %, проте це не зовсім відповідає очікуванням. Водночас, значення R2 залишається достатньо високим. Крім того, як видно з рис.6.4 похибка залежить від величини концентрації заліза: залежність від у логарифмічному масштабі є прямою лінією, проте нахил такої прямої не збігається з лінією істинності. Це спостереження вказує на систематичне відхилення прогнозів, а не на повну втрату кореляції, що означає, що моделі здатні фіксувати відносні відмінності в концентрації, але не можуть точно прогнозувати абсолютні значення.

Хоча наявність залишкових шумів на експериментальних кривих, що не були повністю видалені фільтрацією, могла сприяти спостережуваним розбіжностям, більш правдоподібним поясненням є неповна відповідність між фізичною моделлю, використаною для синтезу даних, та фактичною поведінкою сонячних елементів.

Ця невідповідність, ймовірно, пов'язана з числовими параметрами, що використовуються в основних рівняннях моделі (рівняння 6.1-6.3). Наприклад, для розрахунку характерного часу асоціації FeB (рівняння 6.3) були прийняті значення і , які найбільш часто зустрічаються в літературі, проте загалом існує певний розкид цих значень. Крім того, при розрахунках вважалося, що у всіх точках сонячного елементу значення незмінне, тоді як вказані величини дифузійного бар’єру справедливі лише в -Si, при іншому розташуванні рівня Фермі (що в нашому випадку спостерігається в області просторового заряду) ця енергія модифікується [Murphy2014].

Подібний розкид спостерігається і для інших параметрів. Відмінність будь-якого використаного значення параметра від реального може привести до спостережених відхилень в оцінці моделей. Крім того, в літературі також розглядаються ефекти впливу на оцінку концентрації заліза його неоднорідного розподілу по товщині структури, а також самої величини товщини, які також не брались до уваги під час моделювання.

Загальна стратегія підвищення точності прогнозування передбачає постфактумну (post-hoc) калібровку, при якій до результатів моделювання застосовується коригувальна функція з використанням параметрів, отриманих з обмеженої підмножини експериментальних даних. У даному випадку аналіз показав, що найдоцільнішим коригуванням є квадратична корекція цільової змінної, що виражається наступним чином:

|  |  |
| --- | --- |
|  | (6.4) |

де позначає безпосередній прогноз моделі. Показники ефективності після постфактумного калібрування представлені на рис.6.7в і рис.6.7г. Як показують дані, застосування цієї корекції суттєво зменшило похибки прогнозування. Зокрема, середня відносна похибка в експериментальному наборі даних із 28 зразками тепер становить (13 - 17) % для моделей із найкращими показниками (приблизно 20 із 87 конфігурацій) і залишається нижчою за 25 % для більшості інших. Медіанна похибка є ще нижчою і досягає (7 - 10) % у найсприятливіших випадках. Відповідно до результатів, отриманих для змодельованого тестового набору даних, найточнішими конфігураціями є EfficientNetB7, NASNetLarge, DNN, SVR та XGB. Цікаво, що MobileNetV2 також був серед найефективніших комбінацій, маючи найнижче значення MedAPE (7,64 %) для конфігурації MNV2:CL. Цей результат може свідчити про те, що ознаки, витягнуті MobileNetV2, хоча і менш інформативні для навчальних та змодельованих тестових наборів даних, фіксують специфічні закономірності в спектрограмах, які більш релевантні для експериментальних вимірювань, а застосована корекція могла б ще більше підвищити ефективність цих ознак.

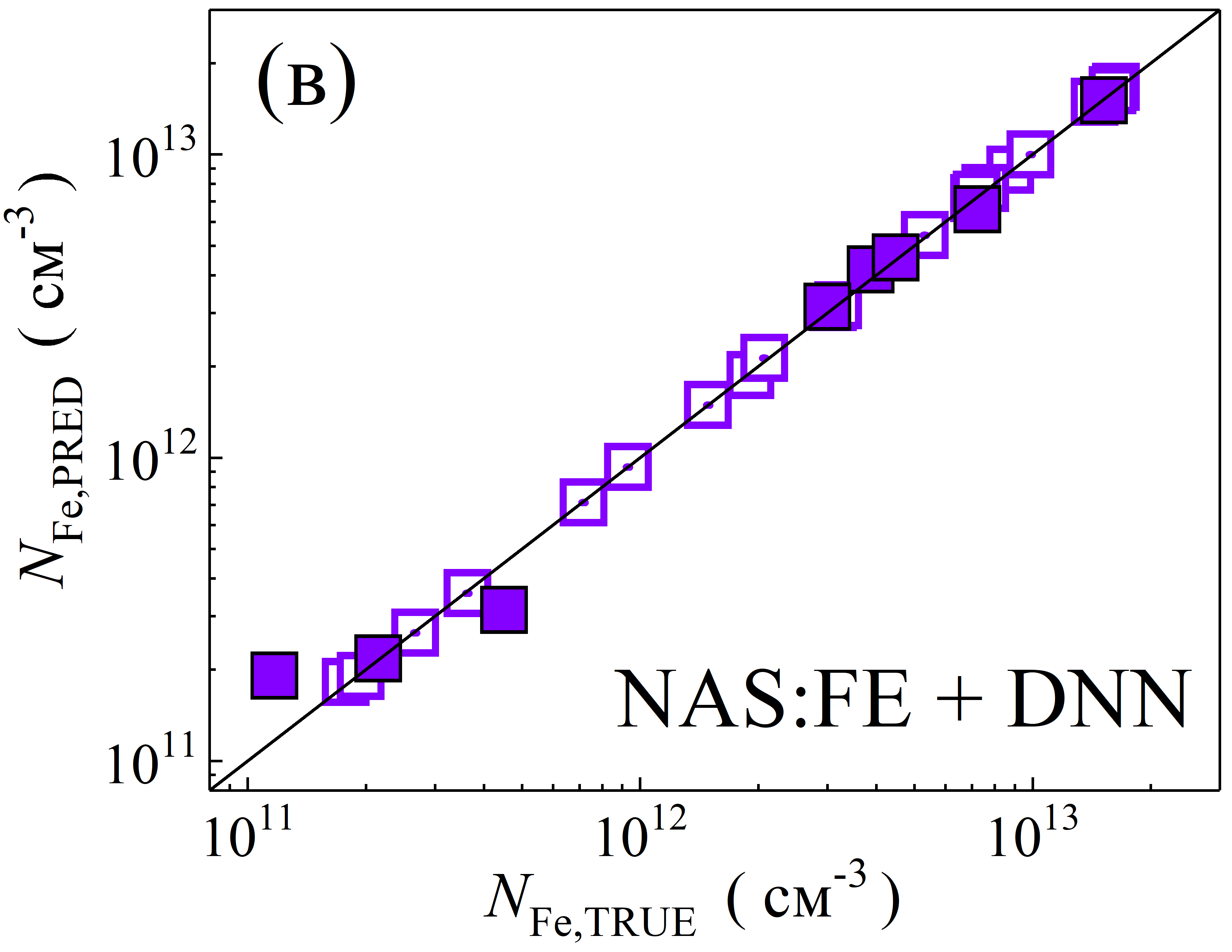
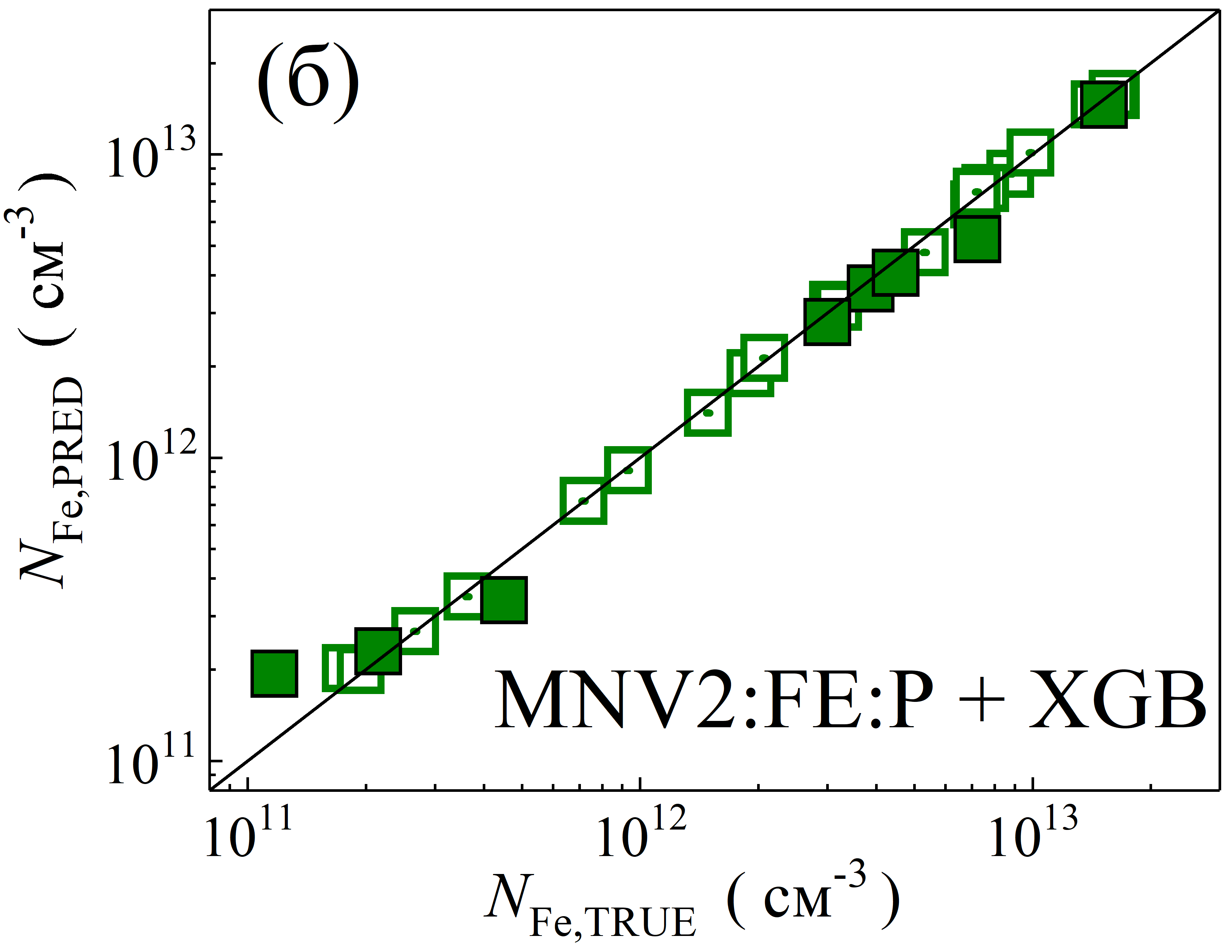
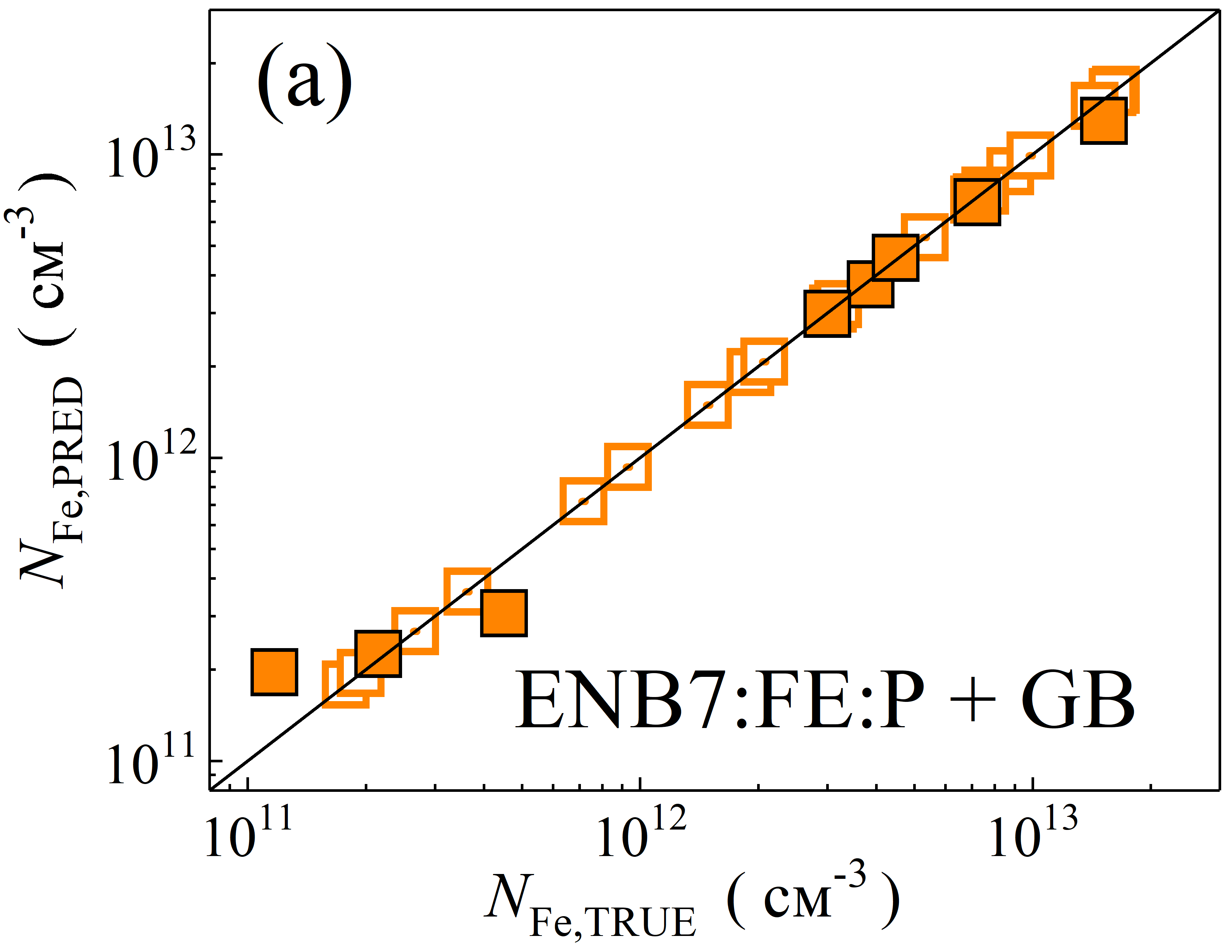
Варто зазначити, що застосування корекції збільшило розрив між середньою та медіанною абсолютними відсотковими похибками. Це спостереження свідчить про те, що, хоча корекція поліпшила загальну узгодженість між прогнозованими та істинними значеннями, у декількох зразках все ще спостерігалися відносно великі залишкові похибки. Ця корекція також призвела до зменшення значення R2. Такий результат є очікуваним, оскільки пост-обробка може послабити лінійну відповідність між початковими прогнозами та експериментальними значеннями, навіть якщо загальні похибки прогнозування одночасно зменшуються. Тому зменшення R2 слід розглядати як побічний ефект підвищення точності моделі, а не як доказ її погіршення.

Хоча квадратична корекція виявилася ефективною, вона не є універсальним рішенням для перенесення моделей із змодельованих даних на експериментальні. Більш надійний підхід передбачає інтеграцію експериментальних даних у процес навчання, як це реалізовано в наступному підрозділі.

**6.5 Апробація моделей на експериментальних залежностях**

На рис.6.8 представлені співвідношення між концентраціями заліза, передбаченими моделями, навченими на експериментальних даних, та відповідними величинами, отриманими використовуючи методологію, описану в [Olikh2022:JMatSci]. На рисунку показані типові результати, отримані для деяких комбінацій моделей КЗ та регресійних моделей під час тренування та тестування. Більш повна інформація міститься на рис.S6.6 в додаткових матеріалах.

Порівняно з випадком тренування на змодельованих даних в цьому випадку використовувалося ще менше зразків (20 проти 25), проте вони відповідали вужчому діапазону концентрацій заліза ( проти (). Рис.6.9 відображає частину метрик, отриманих під час фази навчання (більш повна картина на рис.S6.7). Загалом, картина схожа на ту, що спостерігалася на рис.6.5. Зокрема, в багатьох випадках спостерігаються надзвичайно низькі помилки (менше 0.5%) та високі значення коефіцієнта детермінації (близько 1). Особливо це характерно для GB, RF, SVR. Подібно до попереднього випадку DNN демонструє гірші ніж інші алгоритми результати, проте вони дещо кращі ніж для змодельованого тренувального набору даних.



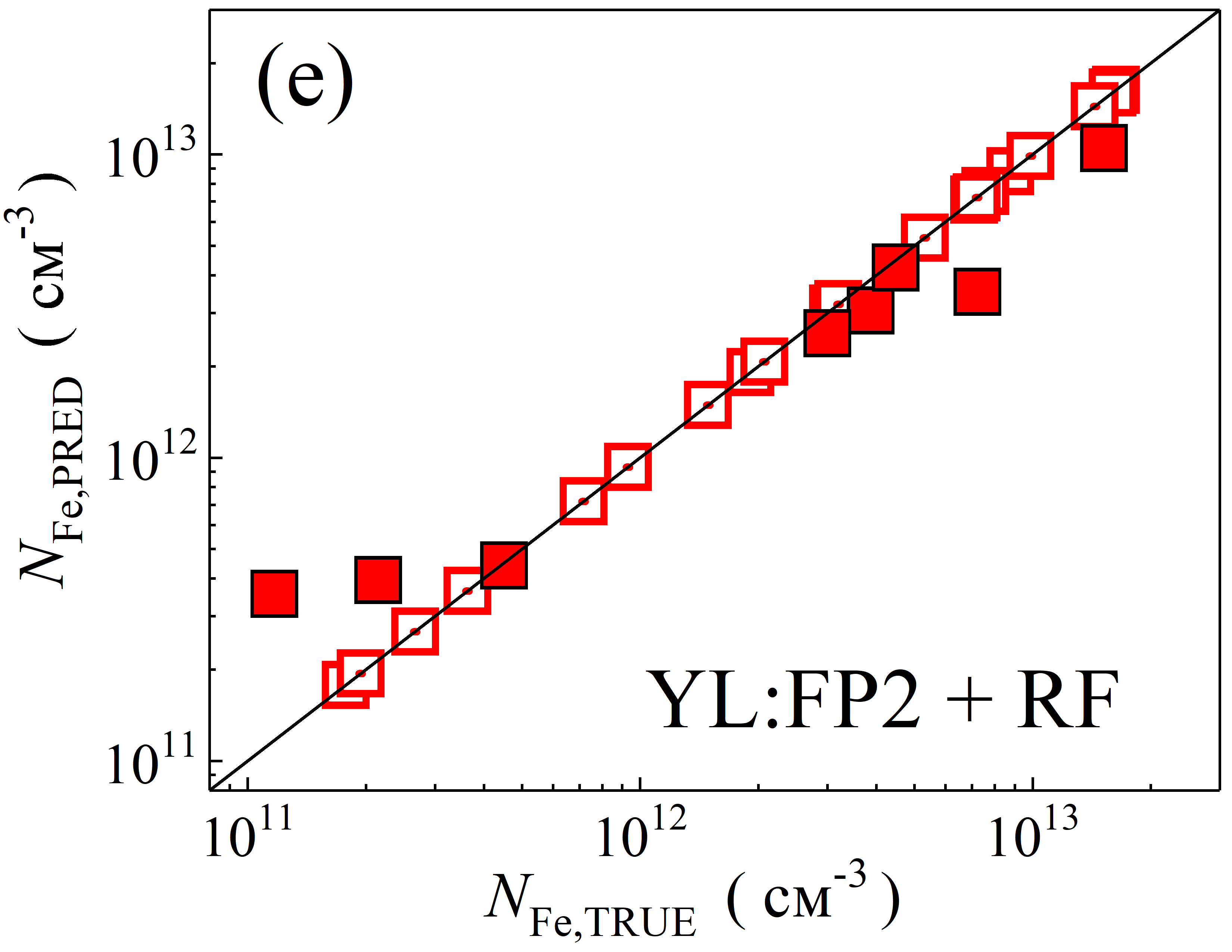
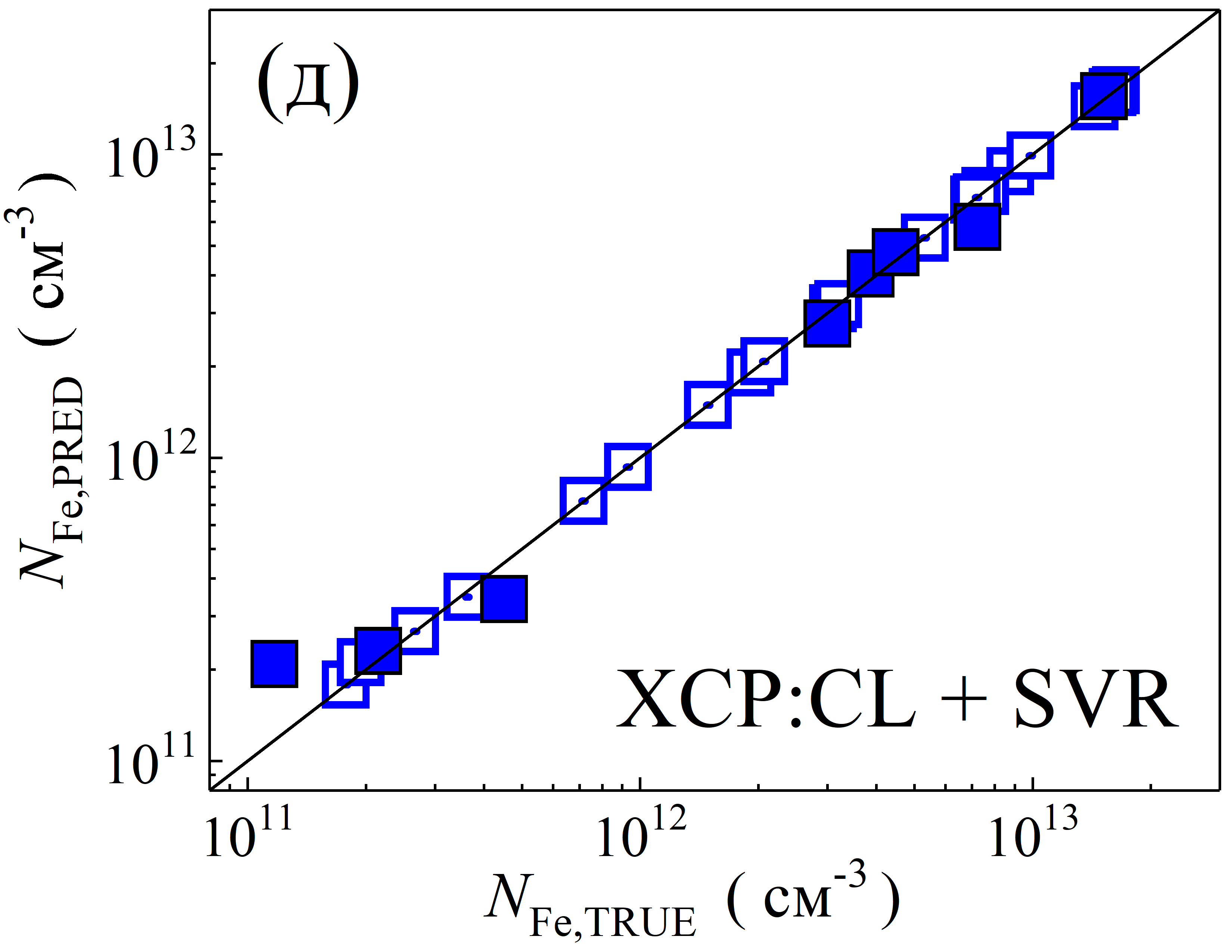
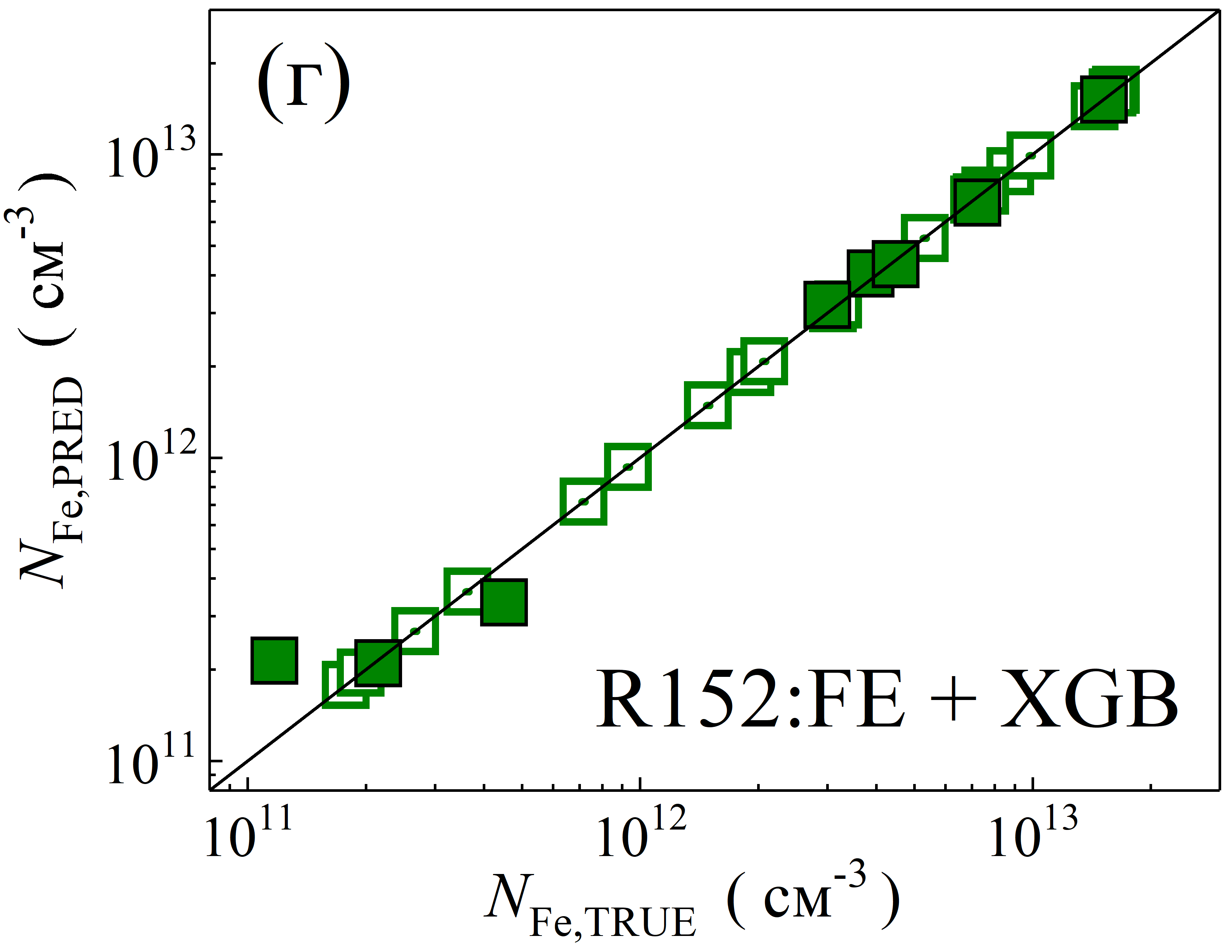


Рис.6.8. Точкові діаграми, що порівнюють еталонні концентрації заліза з прогнозованими значеннями , отриманими за допомогою векторів ознак, витягнутих із різних моделей КЗ у поєднанні з різними алгоритмами регресії. ММН були навчені з використанням набору даних, отриманого в результаті експериментальних вимірювань. Незафарбовані та зафарбовані квадрати відповідають етапам навчання та тестування. Чорні лінії - слугують еталонами.

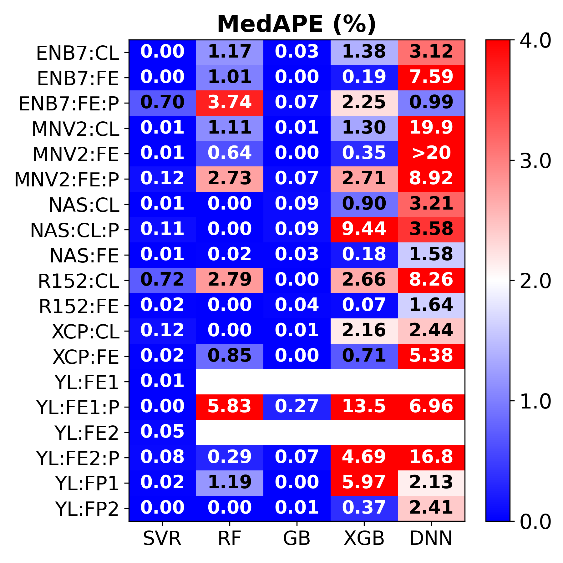
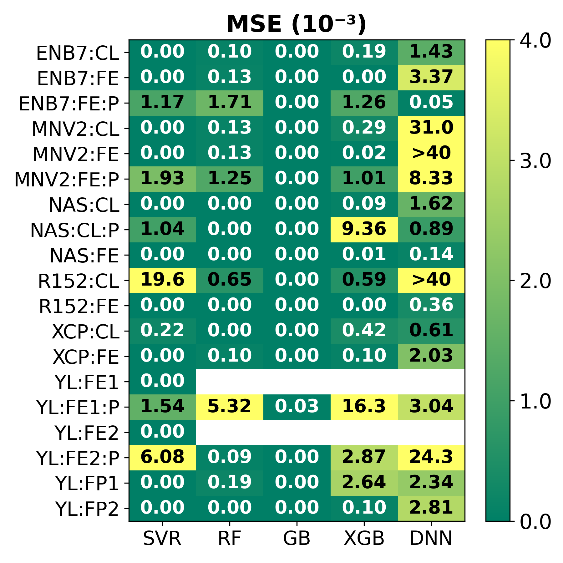


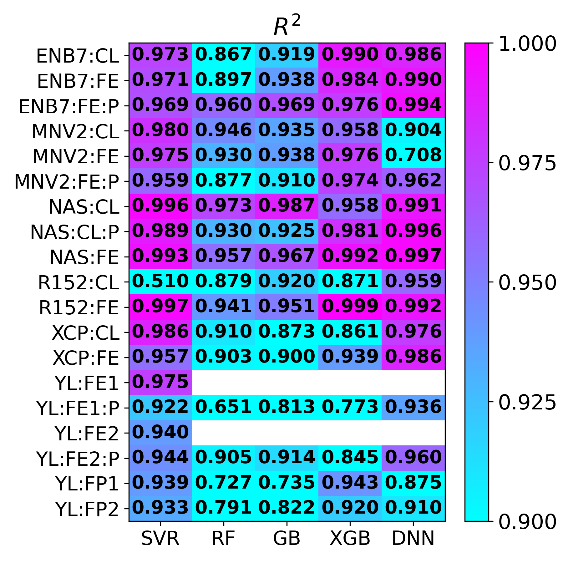
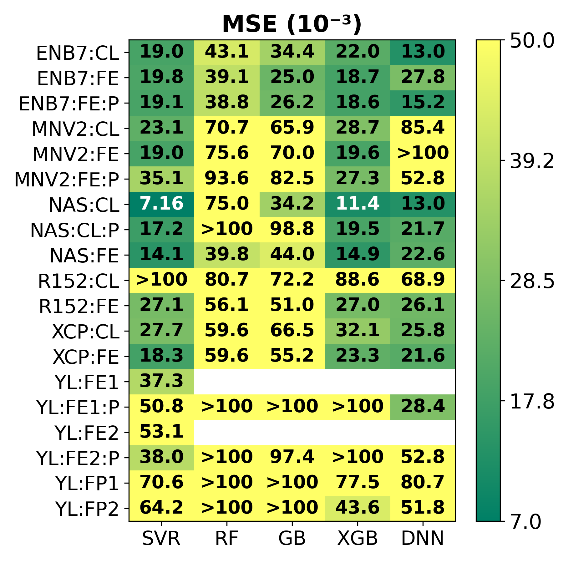
Рис.6.9. Середньоквадратична похибка (лівий рисунок) та медіанна абсолютна відсоткова похибка (правий рисунок) для різних комбінацій моделей КЗ (вертикальна вісь) та регресійних моделей (горизонтальна вісь) під час етапу навчання. Моделі були навчені з використанням експериментального набору даних.

Це може бути пов’язане з більшою однорідність експериментальних даних. Лідерство EfficientNetB7 та NASNetLarge серед моделей КЗ зберіглося, що підтверджує те, що ці архітектури формують найбільш релевантні ознаки для даної задачі, проте різниця показників з іншими моделями комп’ютерного зору зменшилася.

Цікаво, що застосування АГК часто покращує результати для DNN (наприклад, для ENB7:FE та ENB7:FE:P значення MAPE складає 7.3 % та 0.97 %, відповідно), зменшуючи вплив великої розмірності ознак при малій кількості зразків. Для інших регресорів подібного ефекту не спостерігається. Крім того, для експериментального навчального набору даних менша різниця між MAPE та MedAPE, що свідчить про зменшення асиметричності розподілу похибок та кількості катастрофічних відхилень.

На рис.6.10 та рис.S6.8 наведено теплові карти метрик передбачень для тестового експериментального набору, здійснених моделями, що тренувалися на інших, але також експериментальних даних. Якість передбачень суттєво покращилася порівняно з випадком, коли моделі навчалися на змодельованих даних і тестувалися на експериментальних. Зокрема, для низки найкращих комбінацій моделей КЗ та регресійних моделей MAPE та MedAPE знаходяться в діапазоні (7 - 17) %, що значно краще за (20 - 30) % у попередньому випадку. Крім того, значення R2 здебільшого більше 0.97, що вказує на гарне узгодження із реальною залежністю. Фактично, точність, досягнута завдяки безпосередньому навчанню на експериментальних даних, є порівнянною з точністю, отриманою за допомогою постфактумної корекції, проте коефіцієнт кореляції суттєво вищий, що свідчить про більш точне відображення як варіацій, так і масштабу залежності.

Як і раніше, найкращі результати спостерігаються при використанні у регресійній частині SVR, DNN та XGB. Щодо моделей КЗ, то найменш ефективними є YOLOv4 та MobileNetV2, найефективнішими EfficientNetB7, NASNetLarge та, що є певним сюрпризом, ResNet152V2 (у варіанті використання ознак зображень).



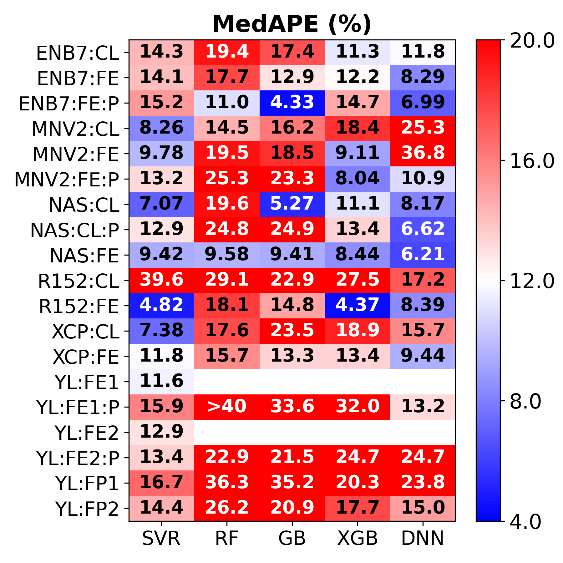
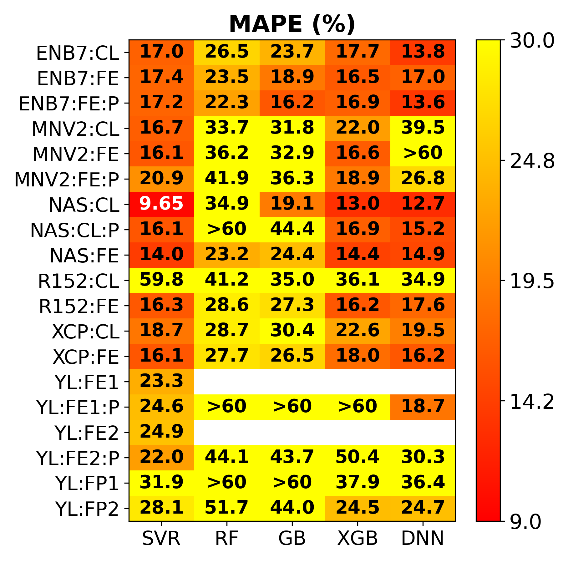


Рис.6.10. Середньоквадратична похибка, коефіцієнт детермінації, середня абсолютна відсоткова похибка та медіанна абсолютна відсоткова похибка для різних комбінацій моделей КЗ (вертикальна вісь) та регресійних моделей (горизонтальна вісь) під час тестової фази. Моделі були навчені з використанням експериментального набору даних.

Цікаво, що для експериментальних даних використання ймовірностей класів як дескрипторів (:CL) дає результати, порівнянні з прямими характеристиками зображення (:FE), а в деяких випадках навіть дещо кращі за них, особливо коли потужні архітектури КЗ поєднуються з гнучкими регресорами. Це, ймовірно, відображає більш компактний та агрегований характер характеристик класів, що робить їх менш чутливими до експериментального шуму. І навпаки, для слабших моделей КЗ конфігурації :FE зберігають свою перевагу, що узгоджується з попередніми висновками. Нарешті, застосування АГК до моделей :FE покращило точність прогнозування, тоді як це не дало переваг моделям, заснованим на ознаках класів. Це свідчить про те, що зменшення розмірності ознак з високим рівнем шуму сприяє покращенню точності прогнозування на експериментальних даних.

Підсумовуючи, було оцінено три підходи до прогнозування концентрацій заліза на основі експериментальних даних: 1) моделі, навчені на змодельованих даних; 2) моделі, навчені на змодельованих даних з постфактумною корекцією; та 3) моделі, навчені безпосередньо на експериментальних даних. Навчання лише на змодельованих даних дало незначну збіжність з експериментальними вимірами; однак було виявлено систематичні похибки через відмінності між змодельованими та реальними системами. Постфактумна корекція ефективно зменшила ці метрики, знизивши середню та медіанну похибки, проте кореляція з фактичними варіаціями залишилася обмеженою. Навчання безпосередньо на експериментальних даних дало найбільш збалансований результат, досягнувши як низьких похибок прогнозування, так і високих коефіцієнтів кореляції, тим самим продемонструвавши важливість включення реальних вимірювань під час розробки моделі. Таким чином, навчання безпосередньо на експериментальних даних покращує точність прогнозування та усуває систематичні похибки, що спостерігаються в моделях, перенесених із змодельованих даних. Водночас для досягнення оптимальної продуктивності необхідний як ретельний підбір комбінації моделі КЗ та регресора, так і включення реальних експериментальних даних.

**Висновки до розділу 6**

1. Показана ефективність запропонованого підходу на прикладі визначення концентрації домішкового заліза з кінетичних залежностей струму короткого замикання після індукованої дисоціації пар FeB з використанням тренувального набору, що складався лише з 20-25 зразків.

2. Розглянуто особливості функціонування моделей, які навчалися як на синтетичних, так і на експериментальних залежностях. Показано, що в обох випадках серед використаних моделей комп’ютерного зору EfficientNetB7 і NASNetLarge дозволяють отримати найбільш релевантні ознаки для регресора. У якості останнього доцільно використовувати алгоритми SVR чи DNN, які перевершують інші підходи, досягаючи значень MSE, MAPE, MedAPE та R2 до 0,001, 6%, 4% та 0,999 для змодельованих даних, та 0,008, 10%, 5% та 0,996 для експериментальних даних.

3. Встановлено, що при тренуванні моделей на штучних даних як дескриптори регресійної моделі доцільно використовувати вектор ознак зображення, тоді як при використанні експериментальних даних зменшити шуми можна завдяки використанню ймовірностей класів або застосуванню АГК.

4. Поєднання трансферного навчання з ЗНМ із відповідним вибором типу дескрипторів та регресора є перспективним підходом для дослідження матеріалів у випадках, коли накопичення великих наборів даних є складним або неможливим.

Основні результати даного розділу представленні в роботі [USCPS\_2025\_Uzhgorod]. Додаткові матеріали до цього розділу, що включають: таблиці інтервалів пошуку раціональних гіперпараметрів, набори раціональних гіперпараметрів, точкові діаграми залежностей еталонних концентрацій заліза від прогнозованих та теплові карти метрик ефективності для 87 різних комбінацій моделей комп'ютерного зору та регресійних моделей можна знайти за посиланням [https://github.com/Zavhorodnii-Oleksii/supplementary\_materials\_for\_thesis/blob/main/Додаткові%20матеріали%20до%20Розділу%206.pdf].