**4. Машино-орієнтований метод оцінки концентрації заліза з використанням темнових ВАХ.**

У попередньому розділі було показано, що існує взаємозв’язок між фактором неідеальності та концентрацією заліза в КСЕ. Однак аналітичні залежності між та є складними, неуніверсальними й чутливими до багатьох фізичних параметрів, що ускладнює їхнє практичне застосування. У цьому розділі буде розглянутий сучасний підхід до оцінювання в КСЕ на основі машинного навчання, а саме — глибоких нейронних мереж (ГНМ). Використання ГНМ дозволить враховувати нелінійні залежності та взаємозв’язки між фактор неідеальності та характеристиками КСЕ.

**4.1 Розробка та навчання глибоких нейронних мереж**

Схематична методика даного підходу наведена на Рис.4.1. Вона складалася з наступних етапів: спочатку були змодельовані темнові ВАХ для першої РМКСЕ; отримані криві апроксимувалися відповідно до дводіодної моделі з метою визначення фактору неідеальності; ГНМ була навчена оцінювати концентрацію забруднюючого заліза, використовуючи товщину бази, рівень легування бором, температуру та фактор неідеальності; ГНМ була протестована з використанням як змодельованих, так і експериментальних ВАХ.

Розглядалися дві конфігурації вхідних даних: перша містила в собі товщину бази (, логарифм концентрації бору , температуру та фактор неідеальності для стану, в якому співіснують комплекси та міжвузлові атоми ; в другій конфігурації, додатково до вже згаданих параметрів, - фактор неідеальності для стану, в якому в КСЕ наявні тільки міжвузлові атоми заліза . Відповідно, розглядалися дві моделі ГНМ: та .

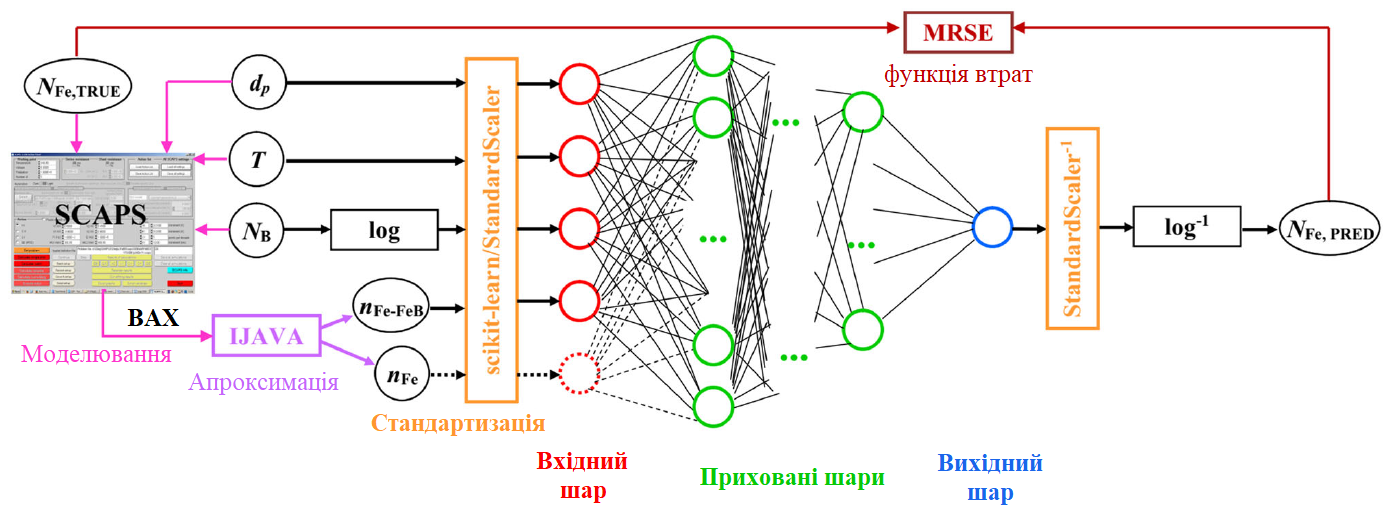


Рис. 4.1 Схема підходу на основі глибокого навчання для прогнозування .

Повнозв’язна ГНМ була реалізована за допомогою високорівневого API Keras за допомогою фреймворка TensorFlow. Вхідні шари мали або чотири або п’ять нейронів. Вихідний шар ГНМ мав лише один нейрон з лінійною функцією активації і передбачав логарифм концентрації заліза в КСЕ . Було розглянуто п'ять конфігурацій прихованих шарів (див. рис.4.2) : «pipe» - кожен прихований шар містив однакову кількість нейронів; «trapezium» - шість прихованих шарів, причому кількість нейронів лінійно зменшується від 100% (перший шар) до 50% (останній шар); «triangle» - десять прихованих шарів, причому кількість нейронів лінійно зменшується від 100% (перший шар) до 10% (останній шар); «butterfly» - дві послідовно з’єднанні «trapezium», причому друга відрезкалена відносно першої; «fir» - дві послідовно з’єднанні «trapezium».

Вибір оптимальної архітектури здійснювався емпірично на основі результатів навчання та тестування моделей. Для кожної з ГНМ потрібно було налаштовувати раціональні набори гіперпараметрів, що включали різні: кількісті нейронів для першого прихованого шару (), кількісті прихованих шарів (), розміри пакетів (*BS*), функції активації (*ActF*), оптимізатори (*Opt*), кількості епох (, темпи навчання (*LR*), методи попередньої обробки даних *(PreM)*, рівні відсічення (*DR*), функції регуляризації (*RegF*), коефіцієнти регуляризації (*RR*) та методи ініціалізації ваг (*WI*).

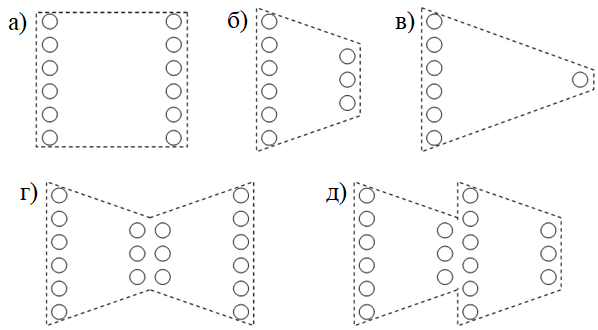


Рис. 4.2 Конфігурації схованих шарів ГНМ, що були використані в дослідженні: а - “pipe”, б - “trapezium”, в - “triangle”, г - “butterfly”, д – “fir”

Для оптимізації архітектури та гіперпараметрів ГНМ було застосовано поєднання методів грубого та точного налаштування. На першому етапі здійснювався грубий перебір (grid search), під час якого варіювалися значення одного з гіперпараметрів у фіксованому наборі, що дозволяло звузити область пошуку та визначити найбільш перспективні конфігурації. На другому етапі для більш точного налаштування використовували випадковий пошук (random search) у межах попередньо визначеного простору гіперпараметрів (див. Таблицю 4.1), що забезпечувало ефективний пошук раціонального поєднання параметрів моделі.

Для кількісної оцінки прогностичних можливостей та здатності узагальнювати використовували 10-кратну перехресну перевірку. Такий підхід передбачає розбиття навчальної вибірки на десять підмножин, дев’ять із яких використовуються для навчання, а одна - для тестування; процедура повторюється для кожної підмножини, а результати усереднюються. Це дозволяє мінімізувати ризик перенавчання та забезпечити об’єктивну оцінку якості моделі на нових даних.

Ефективність прогнозів побудованих ГНМ на тестових наборах оцінювалася за допомогою трьох основних метрик:

1) середнє значення відносної квадратичної похибки (MSRE):

|  |  |
| --- | --- |
|  | (4.1) |

де – кількість зразків у тренувальному (тестовому) наборі,  – істинне значення концентрації заліза для *і*-го зразка, – величина, передбачена ГНМ для даного зразка;

2) коефіцієнт детермінації ():

|  |  |
| --- | --- |
|  | (4.2) |

де - середнє з усіх істинних значень концентрацій заліза;

3) коефіцієнт кореляції Пірсона (*R*) (див. рівняння 3.9).

Додатково, після вибору оптимальної конфігурації гіперпараметрів, було проведено донавчання моделей на повному наборі даних (*Full*), який включав як навчальну, так і всі тестові підвибірки. Такий підхід забезпечив максимальне використання наявної інформації та дозволив досягти високої точності при прогнозуванні концентрації заліза у КСЕ.

Таблиця 4.1 − Початковий простір пошуку гіперпараметрів

|  |  |
| --- | --- |
| Гіперпараметр | Значення |
|  | 4, 5, 6, 8, 10, 15 |
|  | 30, 40, 50, 75, 100, 120, 150 |
| *BS* | 8, 16, 32, 64, 128 |
| *ActF* | ReLu, sigmoid, tanh, SELU, ELU |
| *Opt* | SGD, RMSprop, Adam, Adadelta, Adagrad, Adamax, Nadam, Ftrl |
| *LR* | , , , |
|  | 100, 300, 400, 600, 1000, 1500 |
| *PreM* | StandartScaler, MinMaxScaler |
| *RegF* | Без регуляризації, L2, L1, Dropout |
| *RR* | , , , |
| *DR* | 0.2, 0.3, 0.4, 0.5 |
| *WI* | Xavier Normal, Xavier Uniform, He Normal, He Uniform, Random Normal, Random Uniform, Ones |

**4.2 Результати моделювання та аналіз точності прогнозування**

Окрім навчальної вибірки, були сформовані незалежні тестові вибірки, у яких варіювалися або окремі параметри (, , , ) або всі параметри одночасно, причому використовувалися такі значення параметрів, які не зустрічалися в тренувальному наборі (див. Таблиця 4.2). Загалом, набір «Fe-altered» складався з 857 ВАХ, «d-altered» складався з 1189 ВАХ, «T-altered» складався з 832 ВАХ, «B-altered» складався з 514 ВАХ та «All-altered» складався з 684 ВАХ. В Таблиці 4.3 наведені найефективніші набори гіперпараметрів для наших ГНМ, що були налаштовані за допомогою автоматичного підбору параметрів Keras Tuner, а в Таблиці 4.4 результати 10-ти кратної перехресної перевірки.

Таблиця 4.2 Набори для тестування ГНМ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Назва набору | Варійовані параметри | Інтервали варіації | Параметри, що не змінювалися | Кількість ВАХ |
| Fe-altered (1) |  | {  - |  | 605 |
| Fe-altered (2) |  | {  - |  | 252 |
| d-altered |  | {180 - 260} | Інтервали з тренування | 1189 |
| T-altered |  | {290 - 340} | Інтервали з тренування | 832 |
| B-altered |  | { *-* | Інтервали з тренування | 514 |
| All-altered | *,T* | Комбінація всіх попередніх | - | 684 |

Таблиця 4.3 Найефективніші налаштування гіперпараметрів ГНМ

|  |  |  |
| --- | --- | --- |
| Мережа  Параметр |  |  |
| Значення | |
| конфігурація | 120-108-96-84-72-60 | 100-100-100-100 |
| *BS* | 32 | 32 |
| *ActF* | ReLu | ReLu |
| *Opt* | Adamax | Adamax |
| *LR* |  |  |
|  | 400 | 1500 |
| *WI* | Xavier Normal | Xavier Normal |
| *RF* | Без функції | Без функції |
| *PreM* | StandartScaler | StandartScaler |

Результати навчання та тестування представлені в Таблиці 4.5 та на рис.4.3. З отриманих результатів можемо бачити, що MSRE прогнозування для моделі є досить високою. Проте слід зазначити, що у більшості випадків частка передбачень з великими відхиленнями між істинним та передбаченим значенням концентрації заліза є невеликою. Зокрема, для 87%, 88% та 96% для T-altered, d-altered та Fe-altered відповідно, відносна квадратична похибка (SRE) не перевищує 0,05 (див. Рис. 4.3).

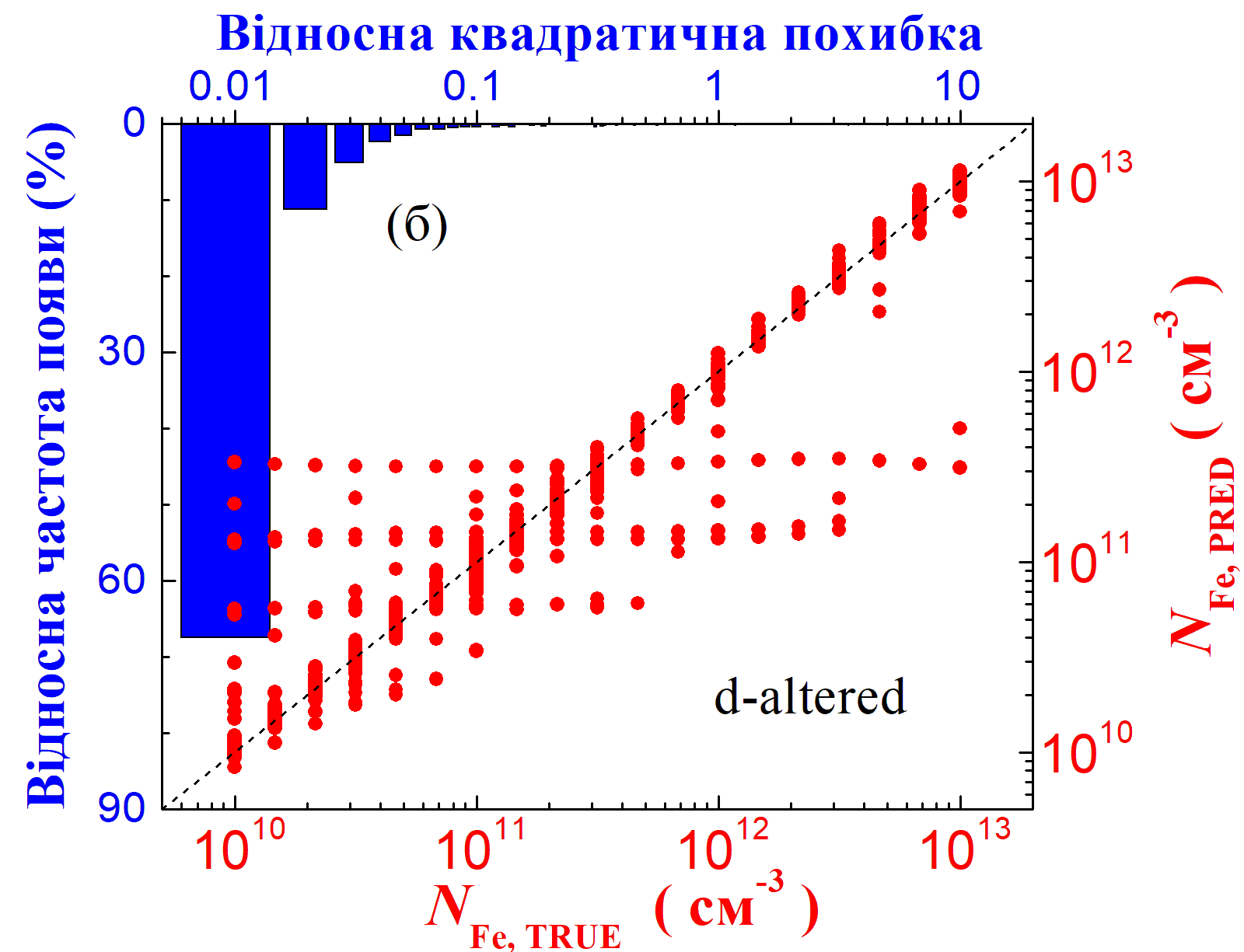
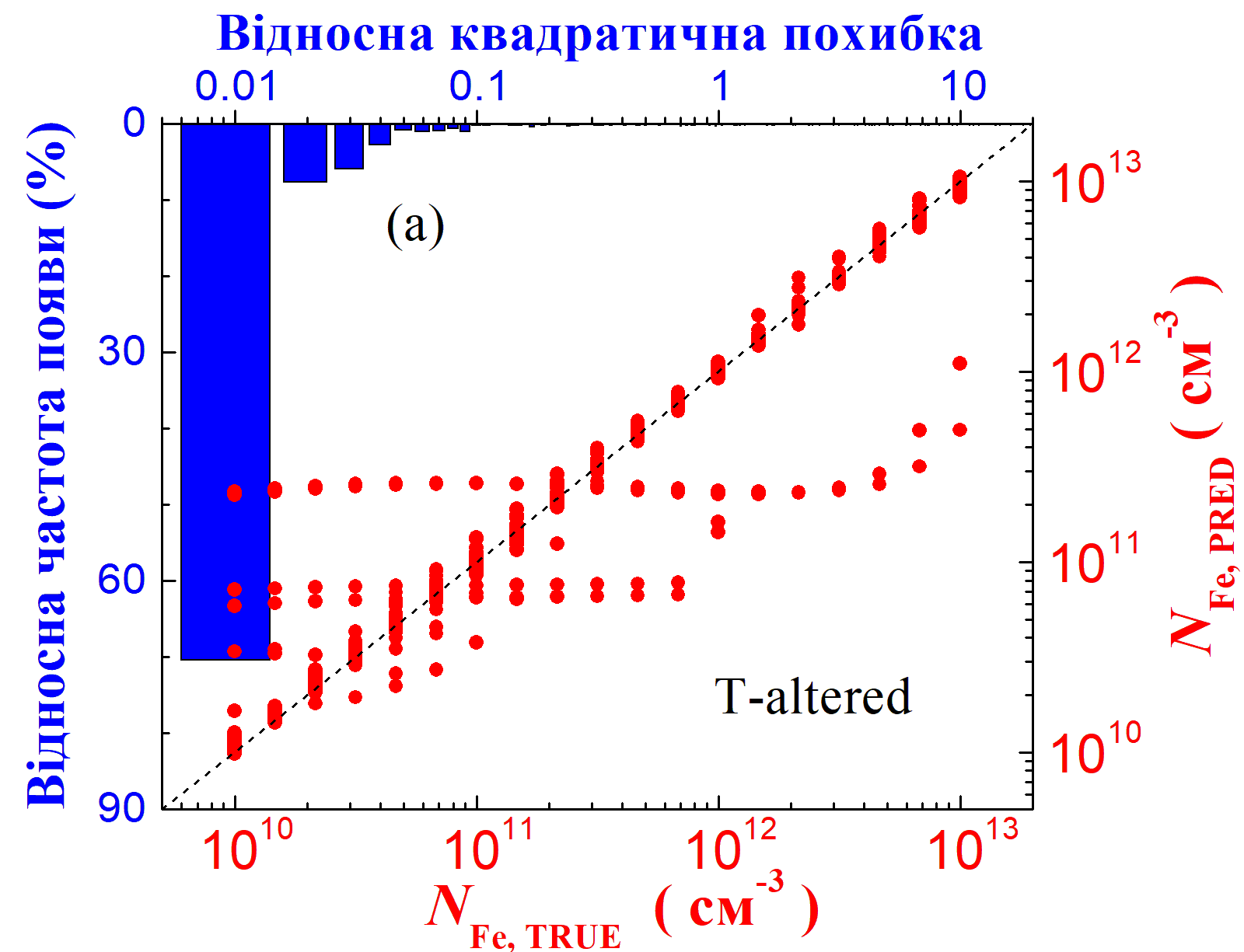
Таблиця 4.4 Результати 10-кратної перехресної перевірки

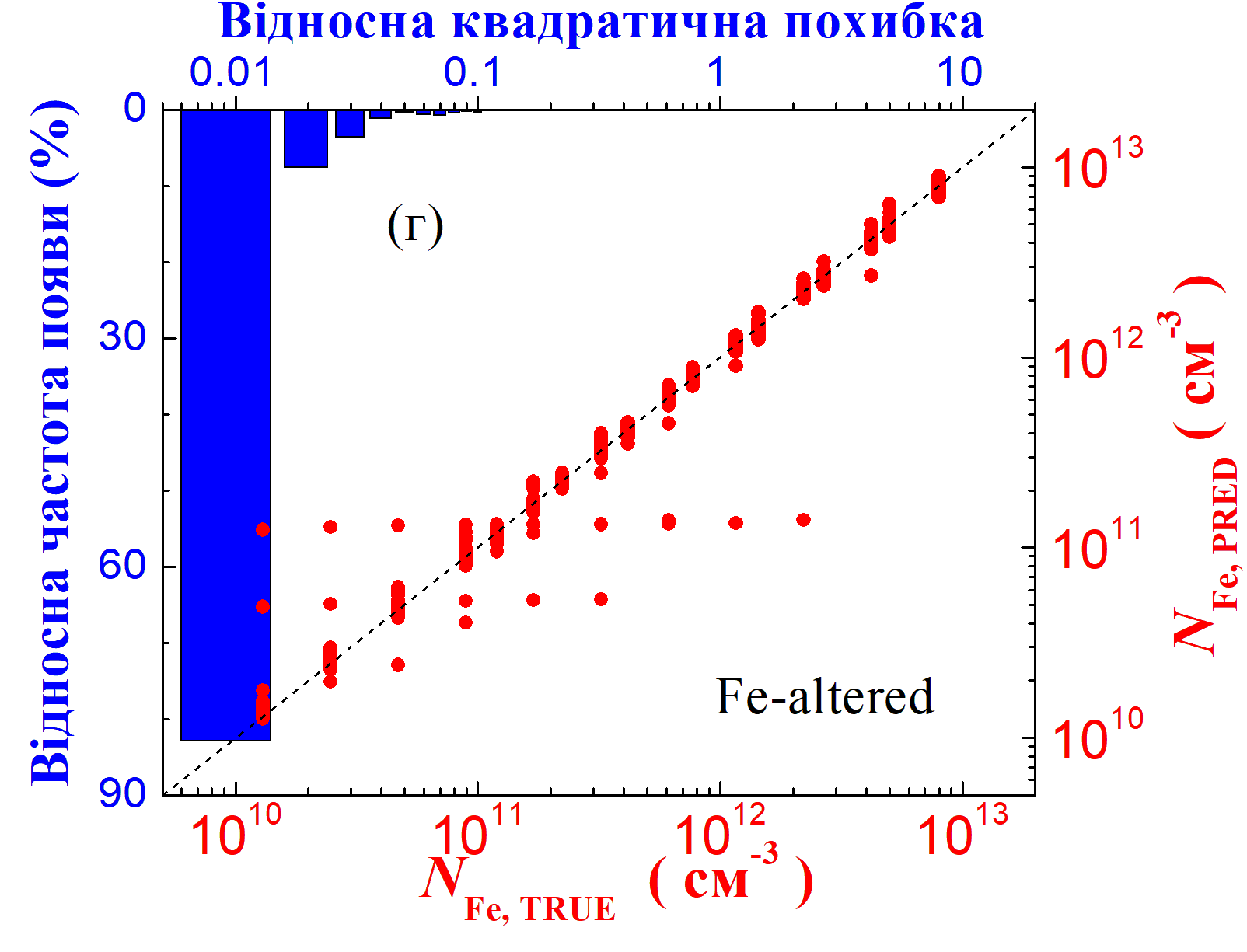
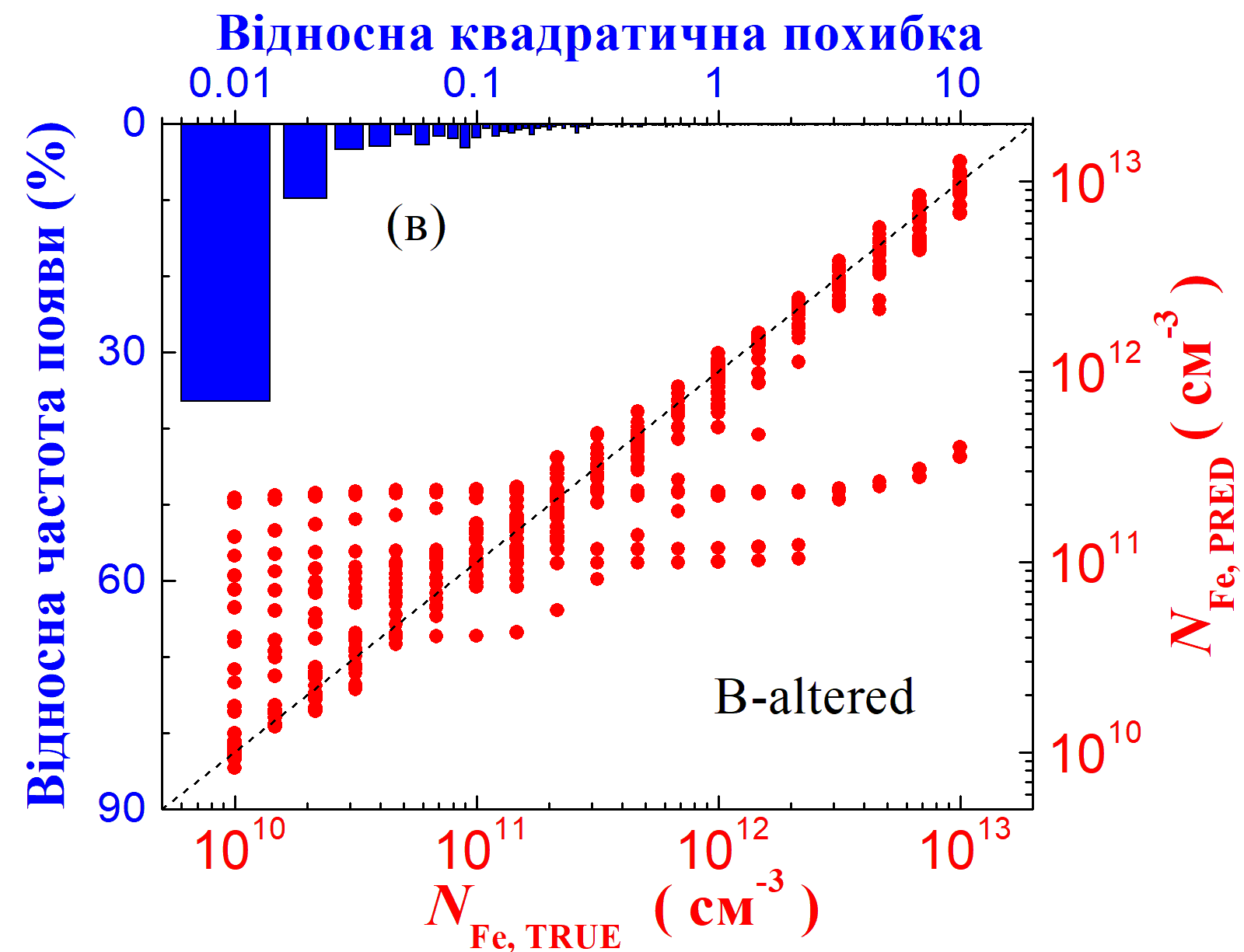
|  |  |  |
| --- | --- | --- |
| Набір | MSRE | |
| Мережа |  |  |
| тренувальний | 0,31±0,07 | 0,03±0,01 |
| повний | 0,28±0,05 | 0,03±0,01 |

Таблиця 4.5 Результати тестування ГНМ

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Набір |  | | |  | | |
| MSRE |  |  | MSRE |  |  |
| T-altered | 0,41 | 0,936 | 0,967 | 0,020 | 0,994 | 0,997 |
| d-altered | 0,37 | 0,961 | 0,980 | 0,018 | 0,996 | 0,998 |
| B-altered | 1,06 | 0,881 | 0,939 | 0,084 | 0,991 | 0,995 |
| Fe-altered | 0,06 | 0,991 | 0,996 | 0,005 | 0,996 | 0,999 |
| All-altered | 0,54 | 0,813 | 0,901 | 0,138 | 0,948 | 0,974 |

Для тестової вибірки B-altered, найбільше значення пов’язане переважно з невеликою кількістю зразків, для яких , тоді як для 54% зразків . Найгірші результати прогнозування, як і очікувалося, спостерігаються для тестової вибірки All-altered: коефіцієнт становить лише 0,813, а спостерігається лише для 18% зразків. Водночас, для вибірки Fe-altered, яка найбільше наближена до реальних умов експлуатації, значення коефіцієнтів та залишаються високими (0,991 та 0,996 відповідно). Також було проведено аналіз залежності похибки прогнозування ГНМ від параметрів КСЕ (див. Рис. 4.4–4.7).





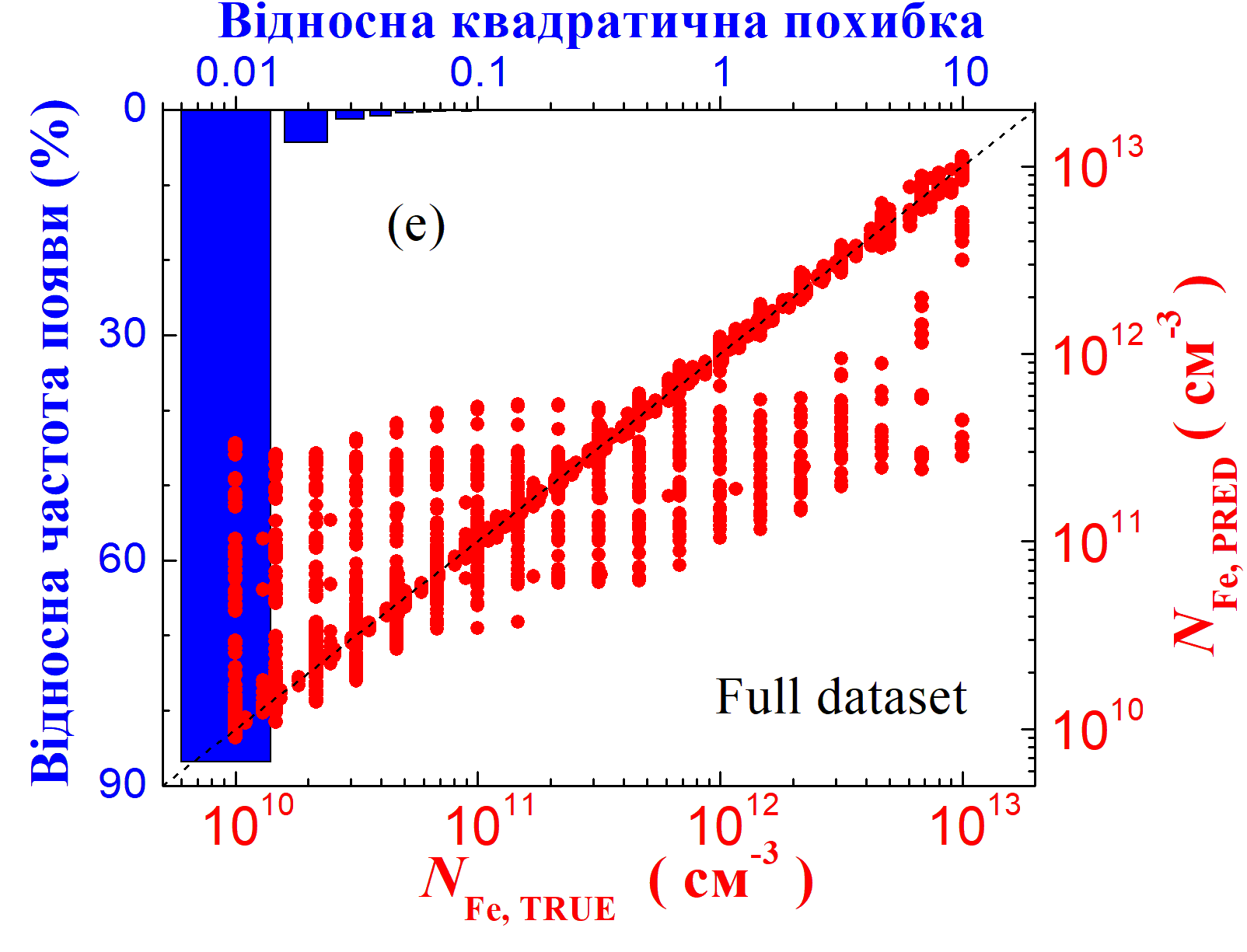
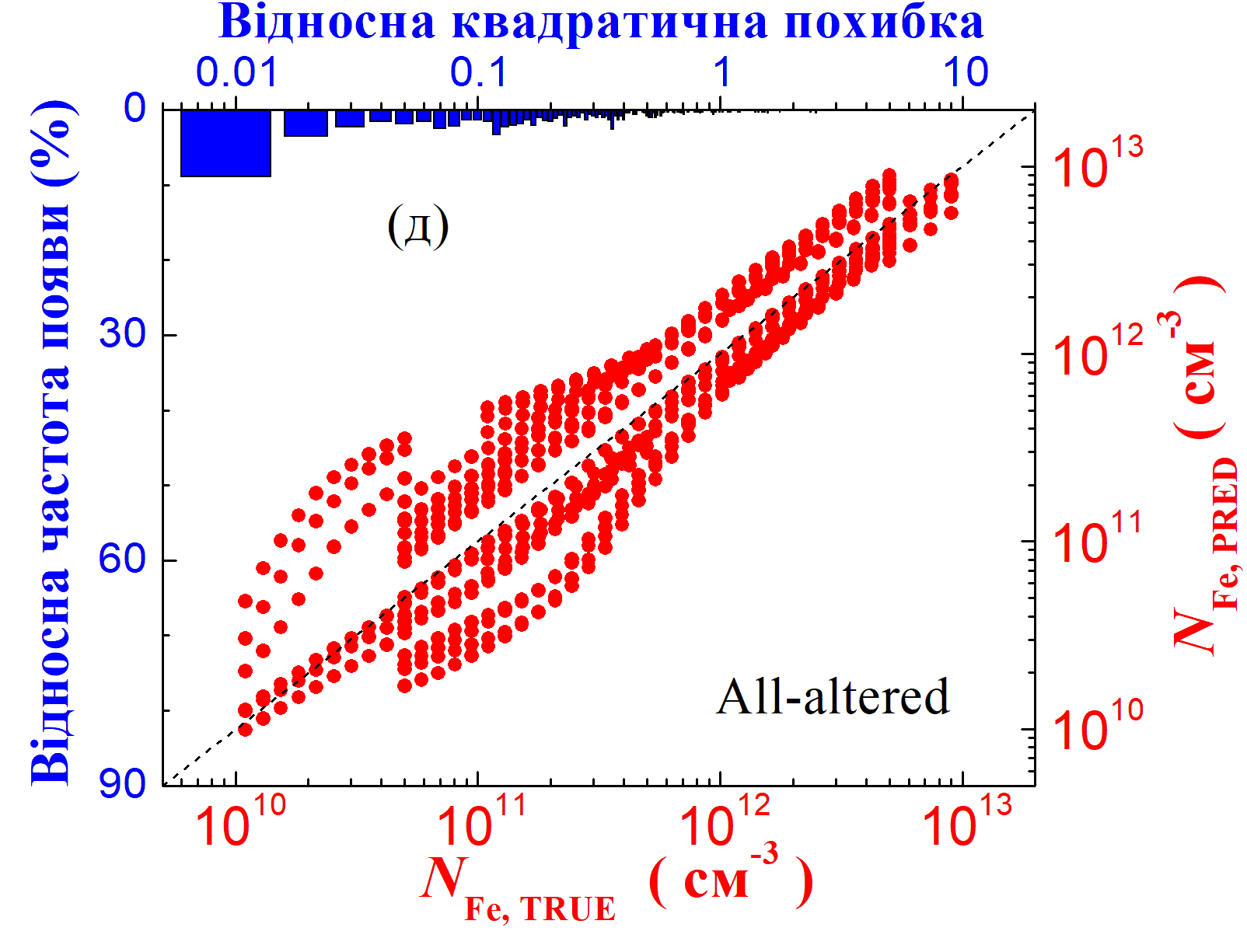


Рис. 4.3 Результати прогнозування моделі на наборах даних T-altered (а), d-altered (б), B-altered (в), Fe-altered (г), All-altered (д) та Full dataset (е) (червоні точки). Сині стовпчики представляють гістограми SRE. Чорні пунктирні лінії є еталонними прогнозами.

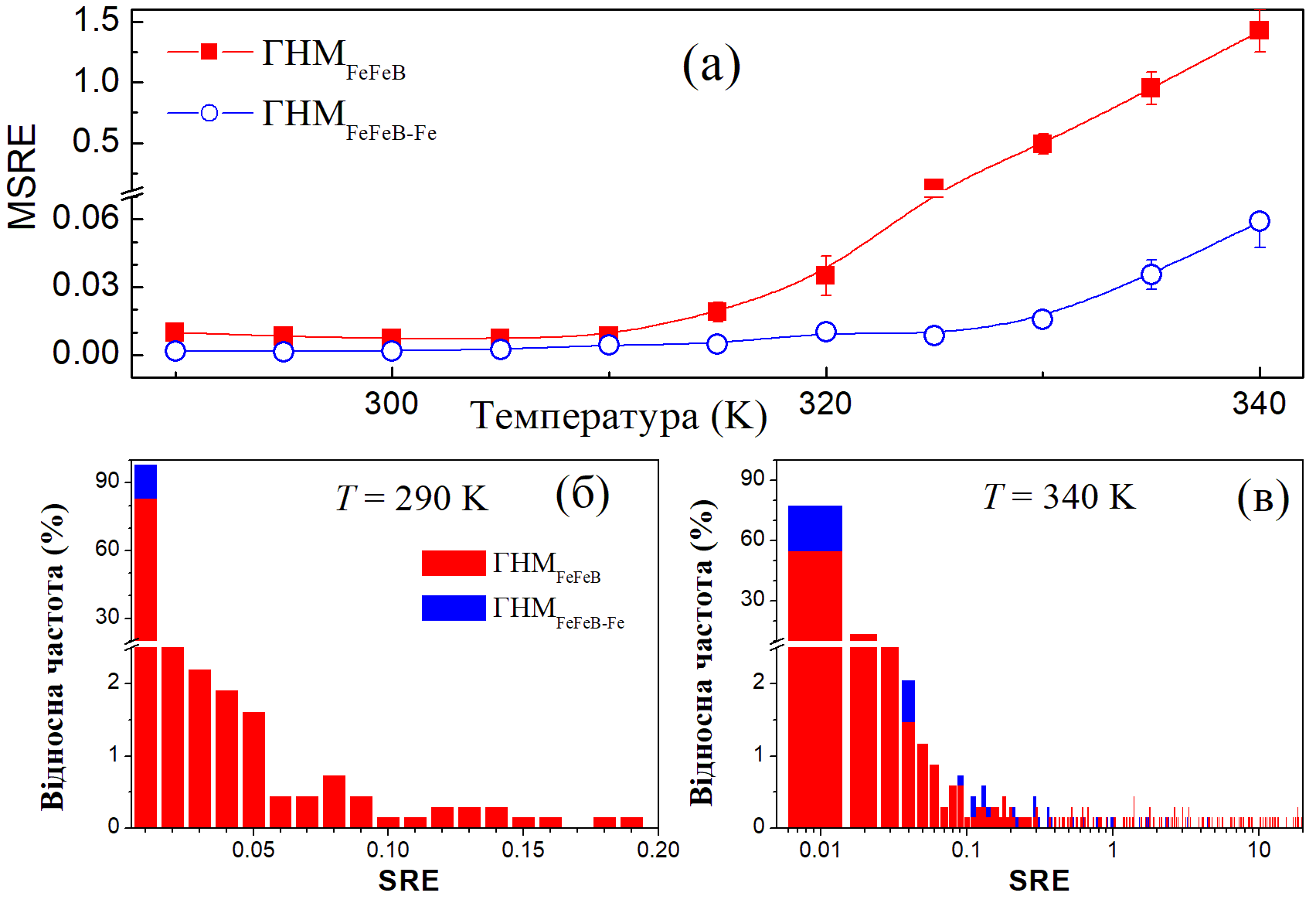


Рис. 4.4 – Залежність MSRE від температури (а) для тренувального набору. Гістограми частоти появи SRE для *T* = 290 К (б) та *T* = 340 К (в). Червоні стовпчики - ; сині стовпчики ‑ .

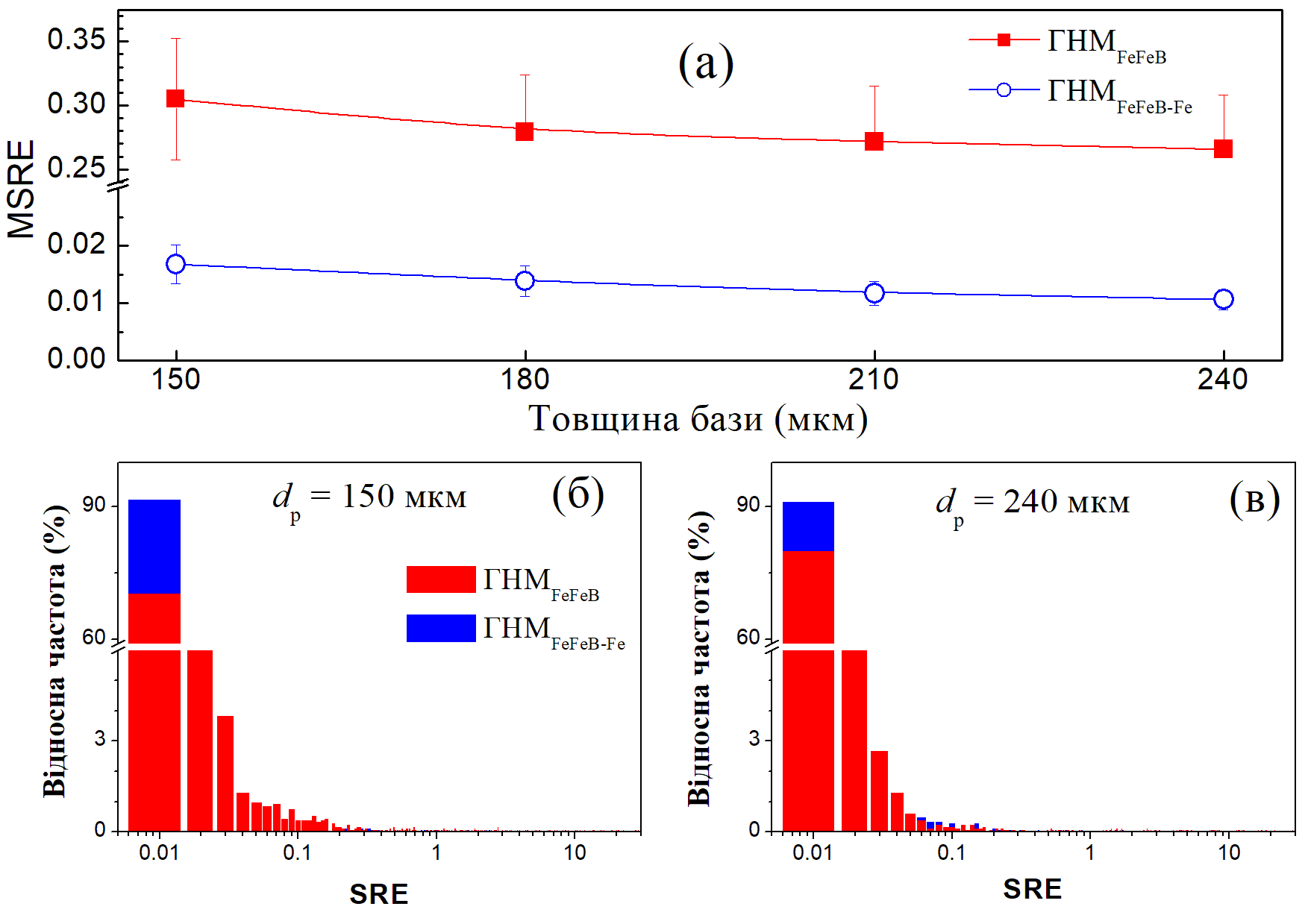


Рис. 4.5 – Залежність MSRE від товщини бази (а) для тренувального набору. Гістограми частоти появи SRE для  = 150 мкм (б) та  = 240 мкм (в). Червоні стовпчики - ; сині стовпчики ‑ .

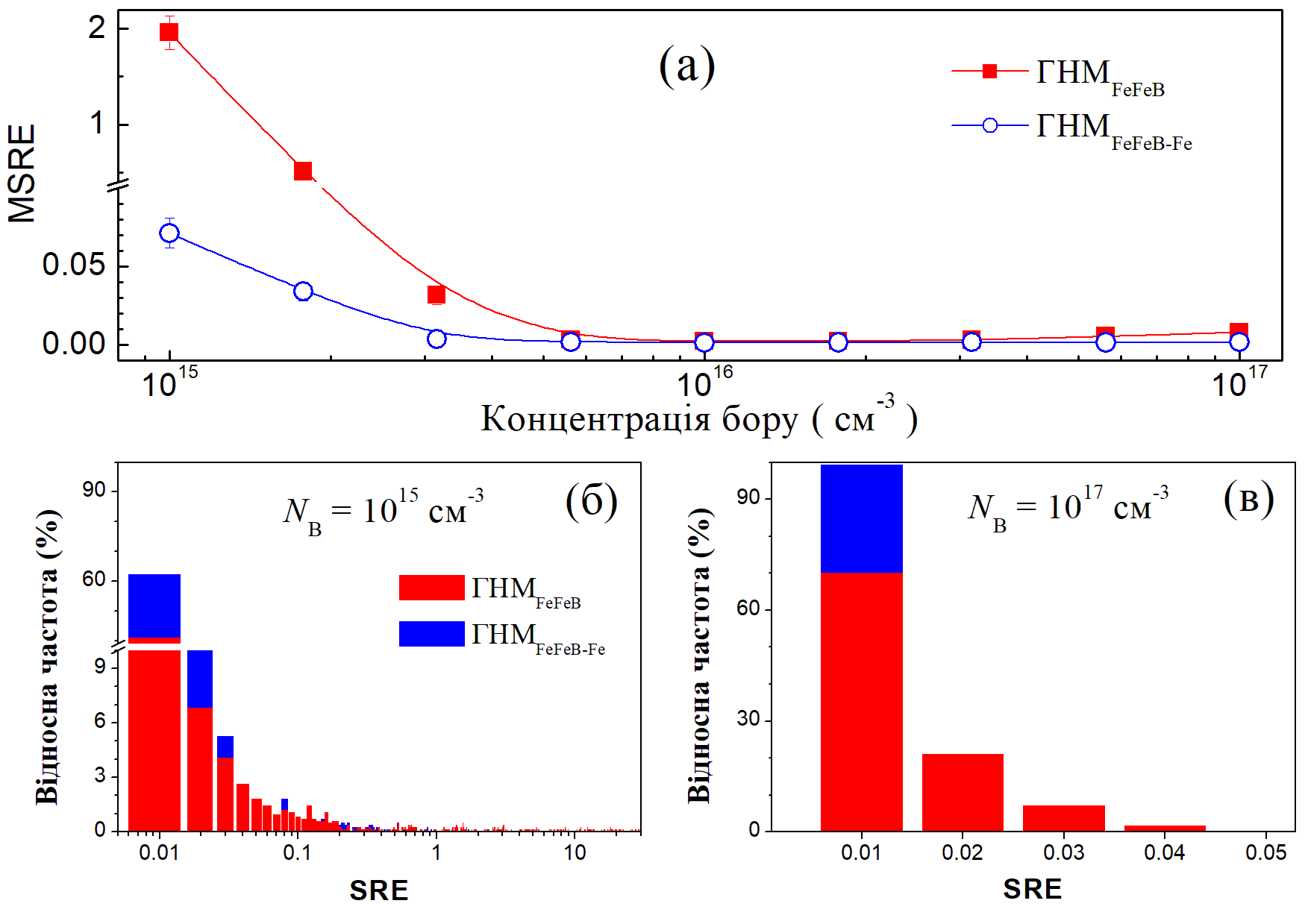


Рис. 4.6 – Залежність MSRE від концентрації бору (а) для тренувального набору. Гістограми частоти появи SRE для (б) та  (в). Червоні стовпчики - ; сині стовпчики ‑ .

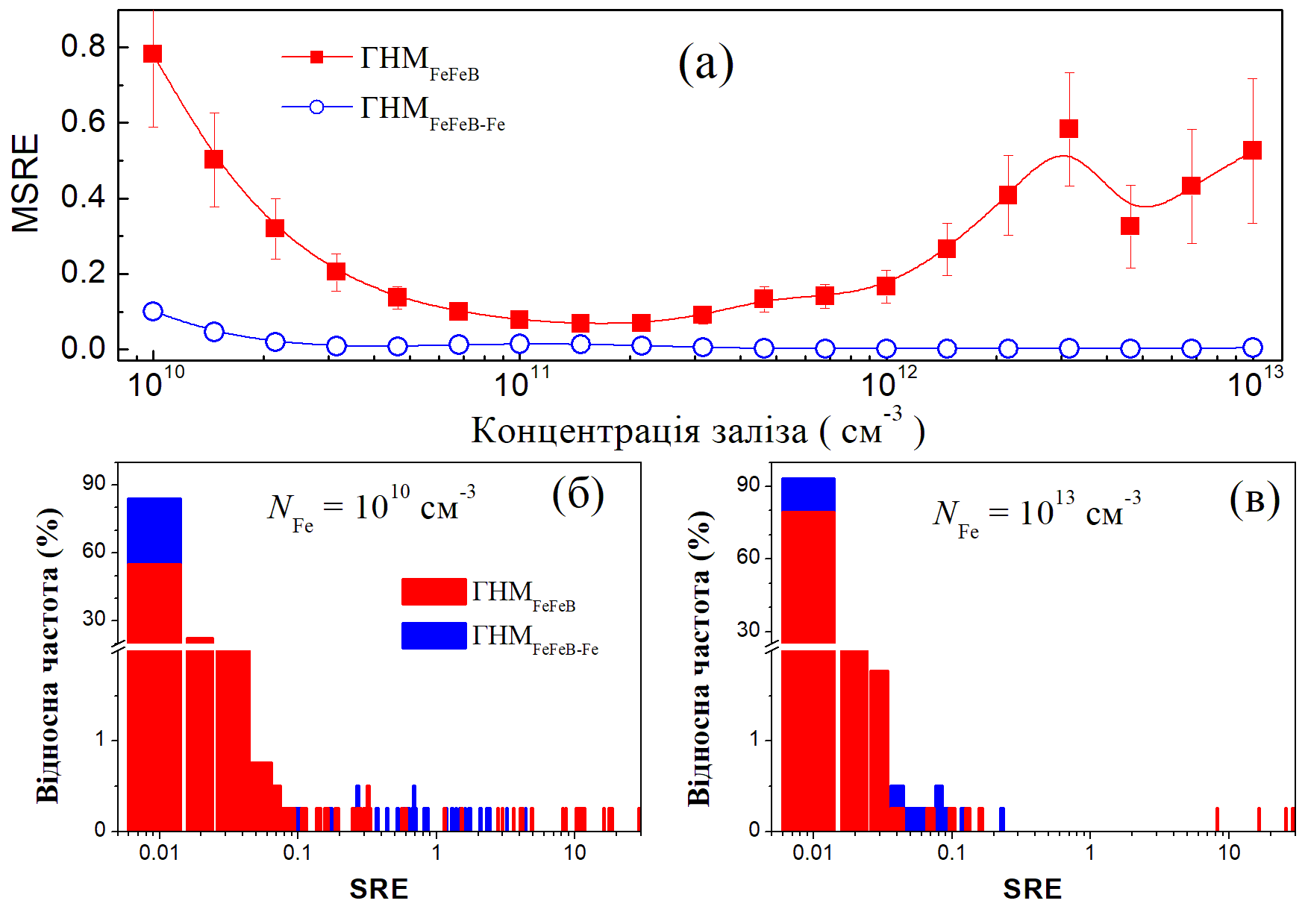


Рис. 4.7 – Залежність MSRE від концентрації заліза (а) для тренувального набору. Гістограми частоти появи SRE для (б) та  (в). Червоні стовпчики - ; сині стовпчики ‑ .

На рис.4.4-4.7 представлені дані для тренувального набору даних; результати для тестових наборів даних є подібними. Зокрема, на рис.4а можна побачити значне збільшення похибки прогнозування, яке спостерігається при для . Як видно з рис.4в , при максимальне SRE становить близько 20, а SRE нижче 0,01 спостерігається для 55% зразків, водночас ці значення дорівнюють 0,02 і 83% коли (див. рис.4б). Як було зазначено в попередньому розділі, підвищення температури спричиняє збільшення впливу власної рекомбінації на коефіцієнт неідеальності. В результаті вплив рекомбінації ШРХ на фактор неідеальності стає менш помітним, а прогнозна здатність ГНМ зменшується.

Як показано на рис.4.5, товщина бази КСЕ практично не впливає на похибку прогнозування (як на середнє значення SRE, так і на відносну частоту, з якою зустрічається помилка певної величини). Однак, коефіцієнт неідеальності залежить від при постійній . Тому є важливим параметром для навчання ГНМ.

Похибка прогнозування різко зростає із зменшенням рівня легування (див. рис. 4.6а). Зокрема, максимальне значення SRE становить приблизно 0,05 для (див. рис.6в), тоді як SRE нижче 0,05 спостерігається лише для 56 % зразків при . Відомо, що зайняття дірками рівня, пов'язаного з залізом, визначає ефективність рекомбінації ШРХ. Згідно зі статистикою Фермі-Дірака, ймовірність заповнення рівня діркою в невиродженому напівпровіднику p-типу з повністю іонізованими акцепторами можна виразити як:

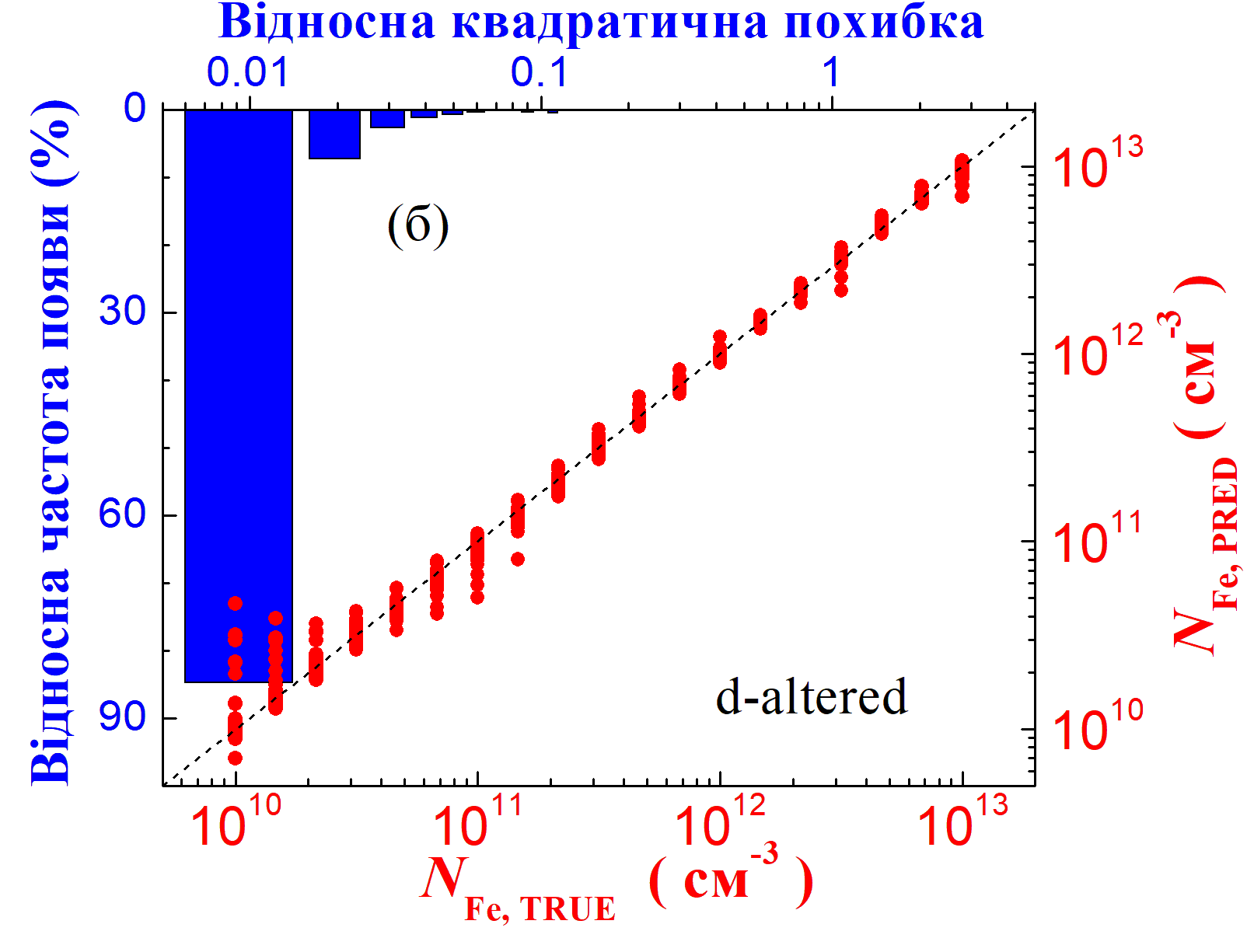
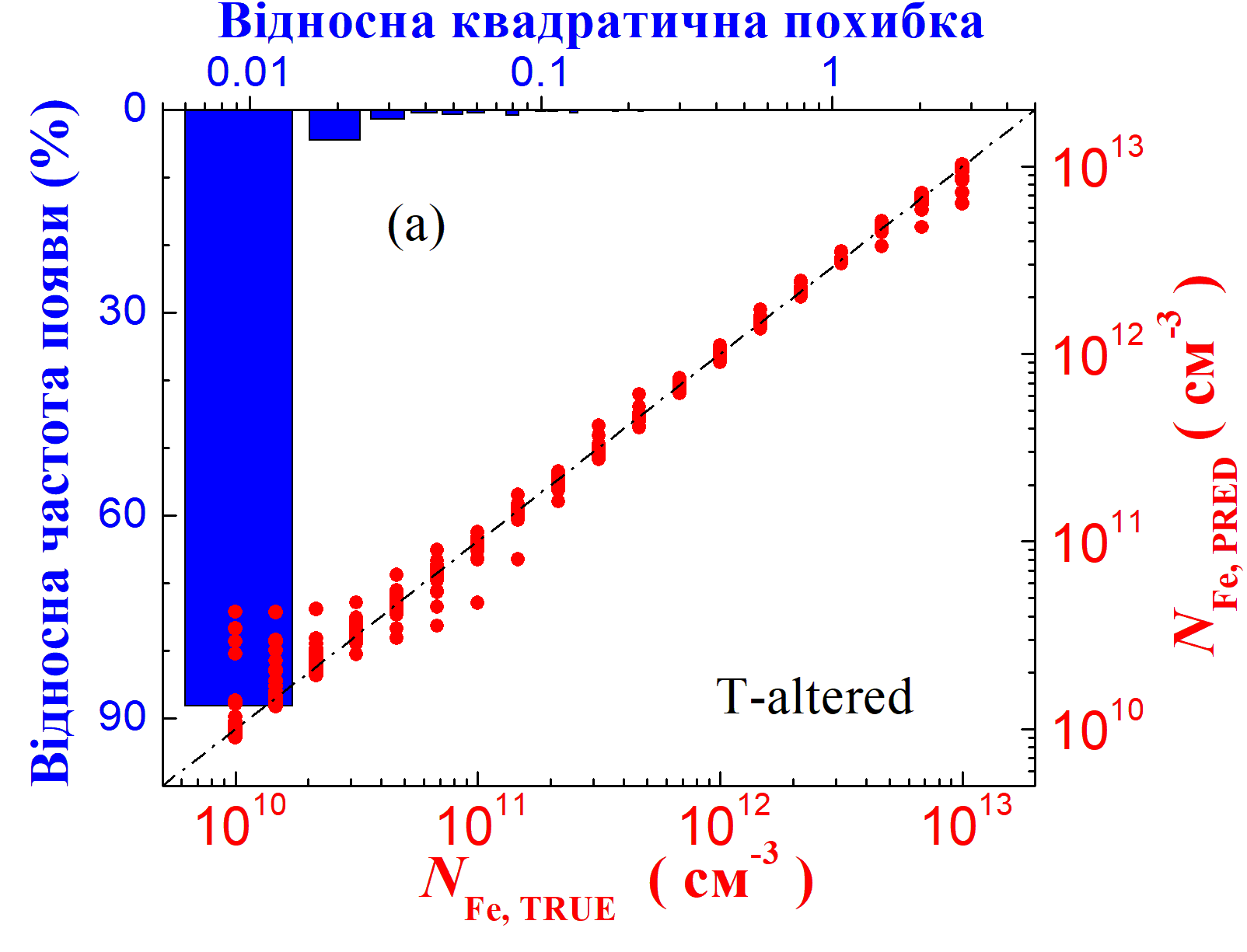
|  |  |
| --- | --- |
|  | (4.3) |

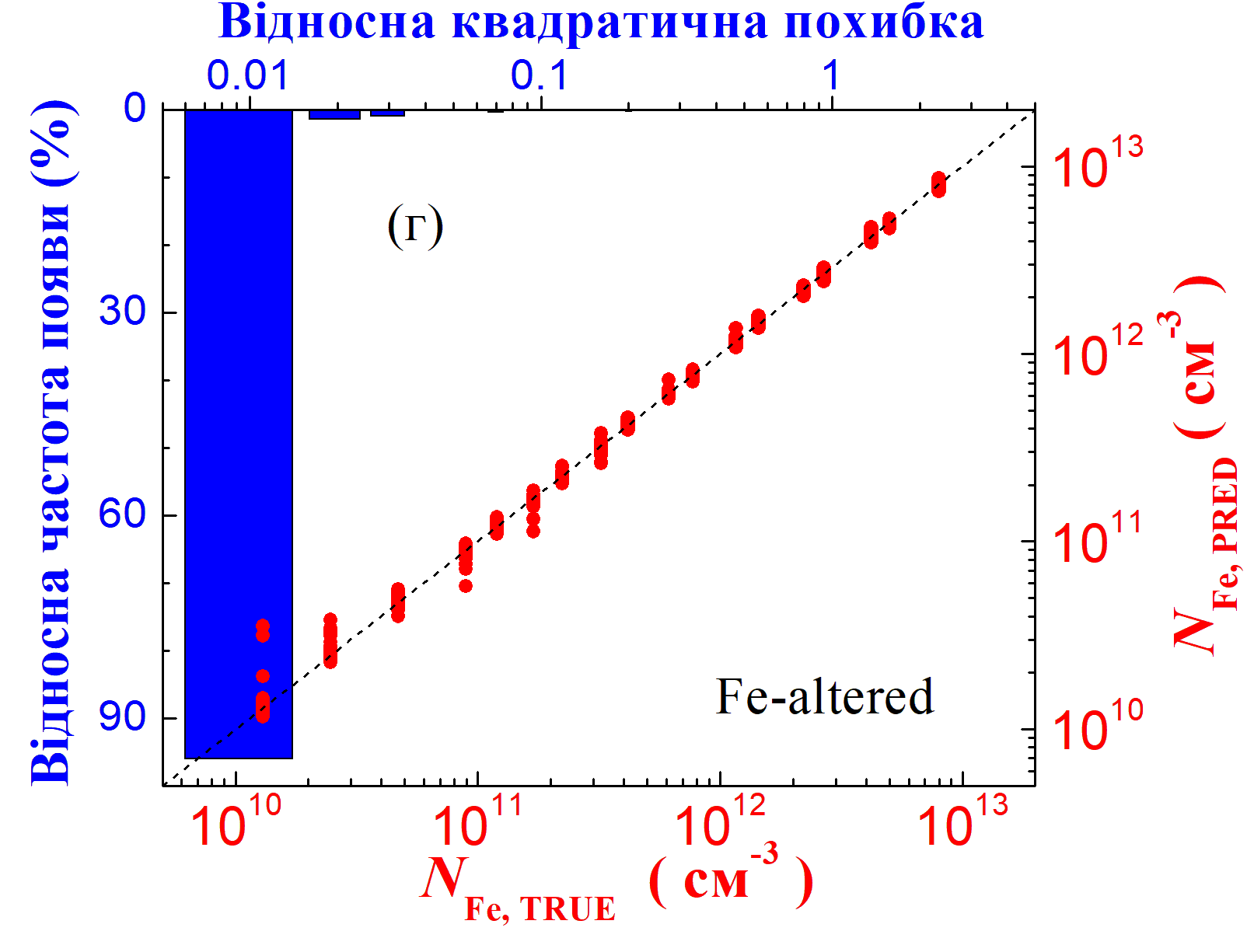
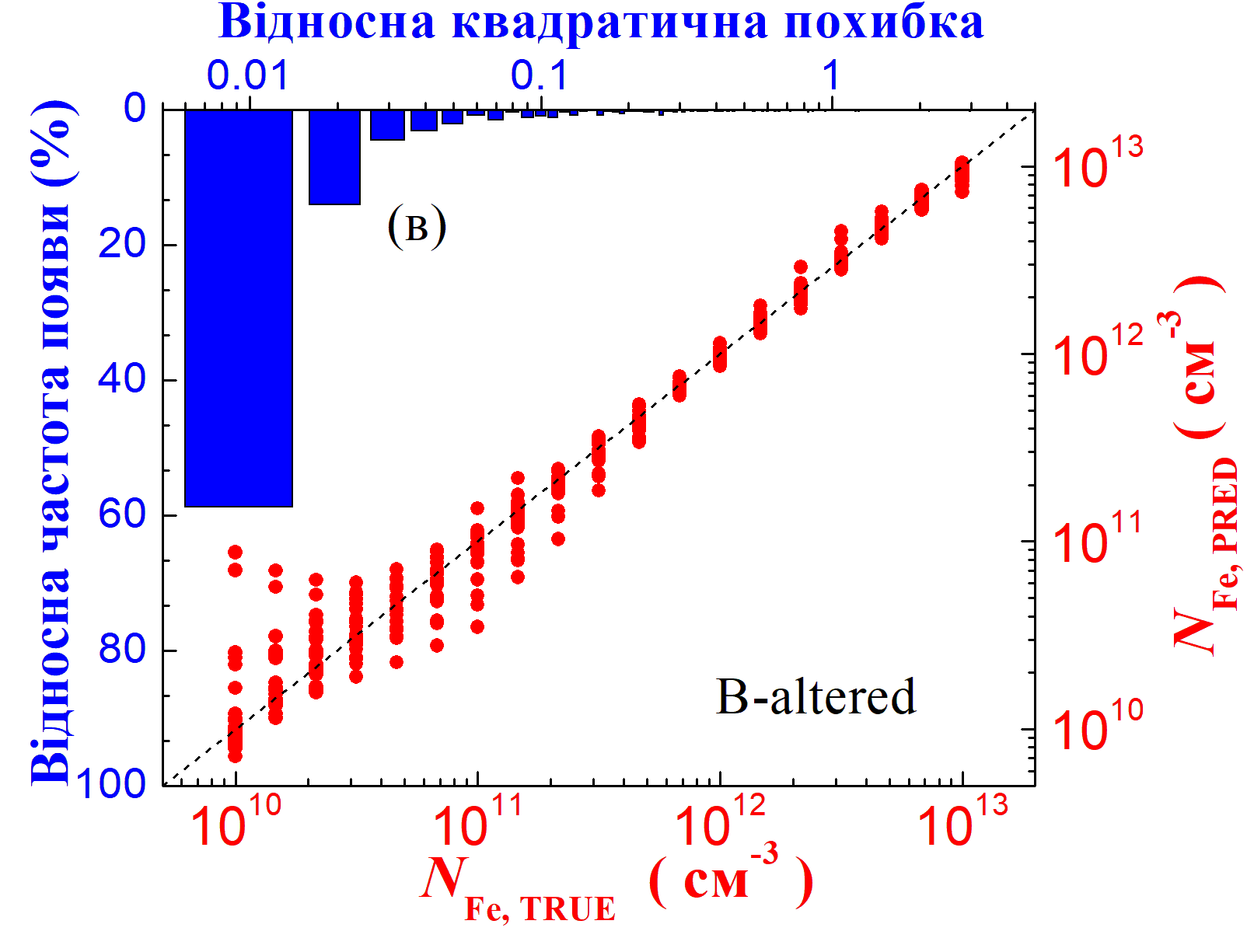
Якщо зменшується, рівень заповнюється електроном, рекомбінація ШРХ припиняється, а значення фактора неідеальності різко зменшується. Більше того, у випадку низького рівня легування домішки мають слабкий вплив на фактор неідеальності, тому спостерігається збільшення MSRE. Додатковим фактором, який спричиняє збільшення похибки передбачень при високих температурах, є заповнення рівня .

На рис.4.7а показано, що MSRE збільшується як при низьких, так і при високих . Перша область з низьким рівнем передбачень ГНМ є цілком прогнозованою, друга ж здається досить несподіваною. Але як показано на рис.4.7в, збільшення MSRE, найімовірніше, пов'язане з тим, що лише кілька зразків прогнозуються з великим SRE при .

Можна зробити висновок, що значення фактора неідеальності для випадку, коли присутнє тільки міжвузольне залізо (), дає додаткову інформацію про дефекти в порівнянні з . Прогнози ГНМ покращуються: MSRE зменшується, немає великої різниці між значеннями і , діапазон SRE стає вужчим (рис.4.4–4.8). Як показано на рис.4.8, максимальне SRE не перевищує одиниці навіть для набору даних All-altered, а SRE нижче 0,02 для 93%, 92%, 73% і 97% зразків у наборах даних T-altered, d-altered, B-altered та Fe-altered відповідно. Слід зазначити, що для тестового набору даних Fe-altered як , так і R дорівнюють 0,999.

Результати навчання та на повному наборі даних (full dataset) представлені в Таблиці 4.4 та на рис.4.3 та рис.4.8е. Бачимо, що в нашому випадку розширення розміченого набору даних практично не покращує результат ГНМ. Це свідчить про: добре налаштування ГНМ; обмежену здатність до прогнозування, що спричинено неоднозначністю залежності .





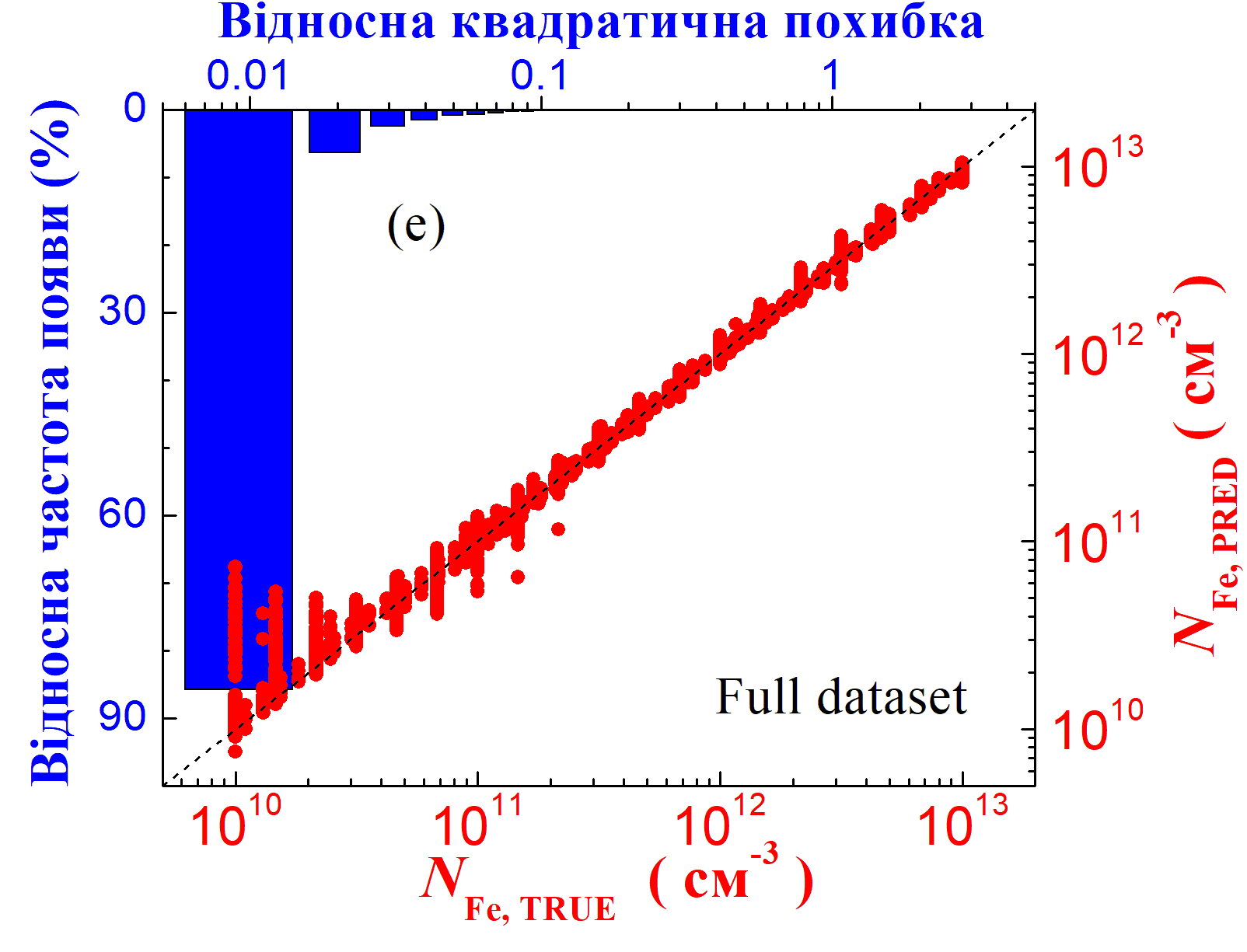
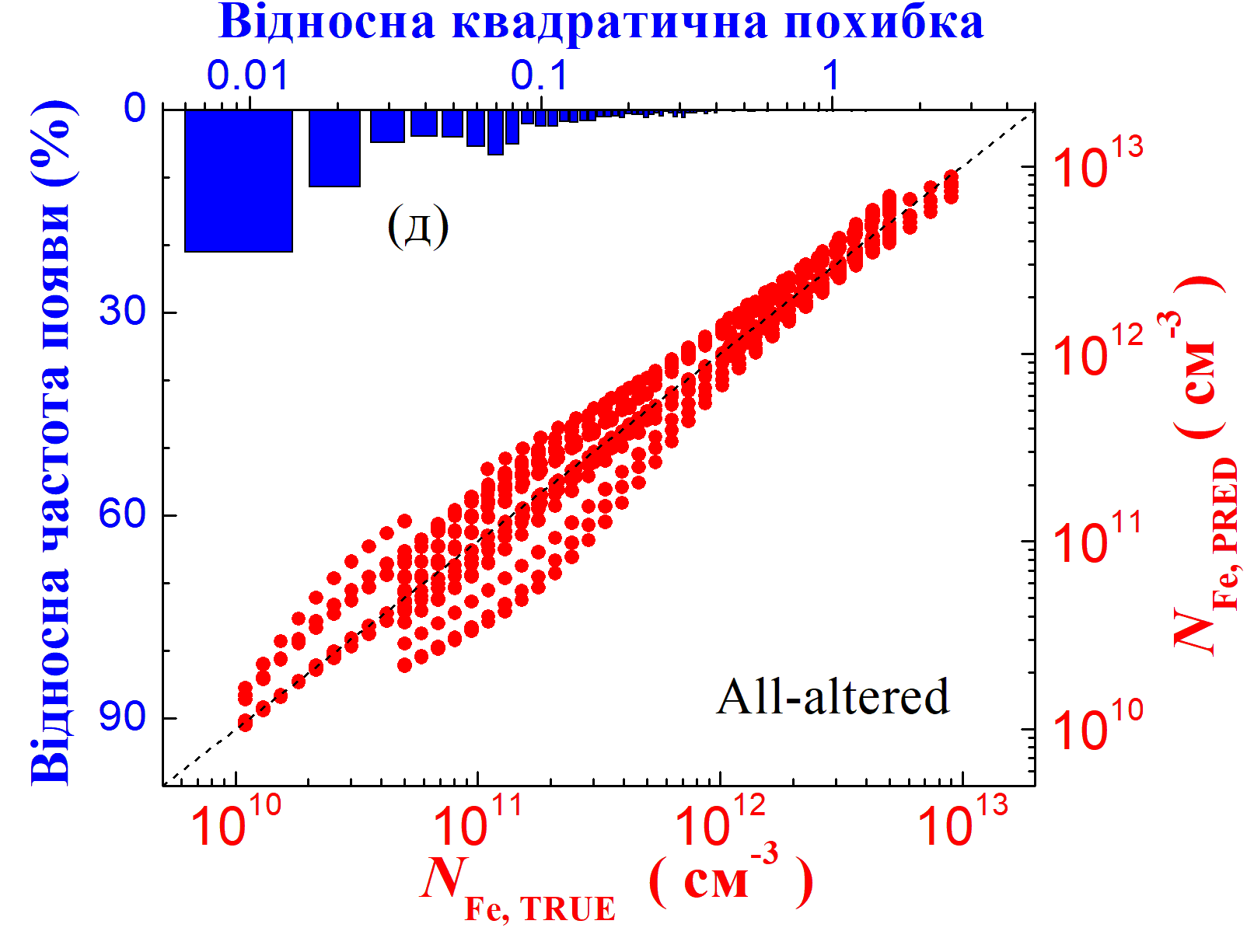


Рис. 4.8 Результати прогнозування моделі на наборах даних T-altered (а), d-altered (б), B-altered (в), Fe-altered (г), All-altered (д) та Full dataset (е) (червоні точки). Сині стовпчики представляють гістограми SRE. Чорні пунктирні лінії є еталонними прогнозами.

**4.3 Перевірка методики на експериментальних ВАХ**

Здатність ГНМ передбачати була перевірена на реальних КСЕ. Зразки, використані в експерименті, мали структуру . Вони були виготовлені з кремнієвої пластини Чохральського p-типу, легованої бором, з орієнтацією [100] і питомим опором 10 Ом·см (). Емітерний шар з питомим опором близько (20-30) мав товщину 0,7 мкм та був сформований шляхом дифузії фосфору при температурі . Товщина бази елементів становила 350 мкм, а їх площа . ПЗП-шар мав поверхневий опір 10-20 та товщину 0,6 мкм. Концентрація заліза в базі КСЕ була визначена за кінетикою струму короткого замикання під дією монохроматичним освітленням після дисоціації пар FeB. Для тестування методики були використані два зразки, що були позначенні як SC320 і SC349 з - (2,0 ± 0,4) і (6,7 ± 0,7) відповідно.

З наведених результатів бачимо, що ГНМ зіткнулася з досить складним завданням, коли складність була пов'язана з певною невідповідністю між параметрами реальних структур і тими, що використовувалися під час моделювання. Проте головним критерієм вибору саме цих зразків була необхідність того, щоб рекомбінація ШРХ визначалась саме дефектами, пов’язаними із залізом.

Темнові ВАХ зразків вимірювалися при температурах 300, 320 і 340 К. Вимірювання були проведені після 48-годинної витримки структур в темряві при кімнатній температурі (випадок «FeFeB»), а також відразу після інтенсивного освітлення структур галогенною лампою (випадок «Fe»). Після цього проводилася апроксимація отриманих кривих відповідно до рівняння 2.12 та визначили з них , , . Типові результати вимірювань та наближень наведено на рисунку 4.9 та в Таблиці 4.6. Варто зазначити, що для реальних ВАХ, на відміну від змодельованих, впливом послідовного і шунтуючого опорів не можна знехтувати.

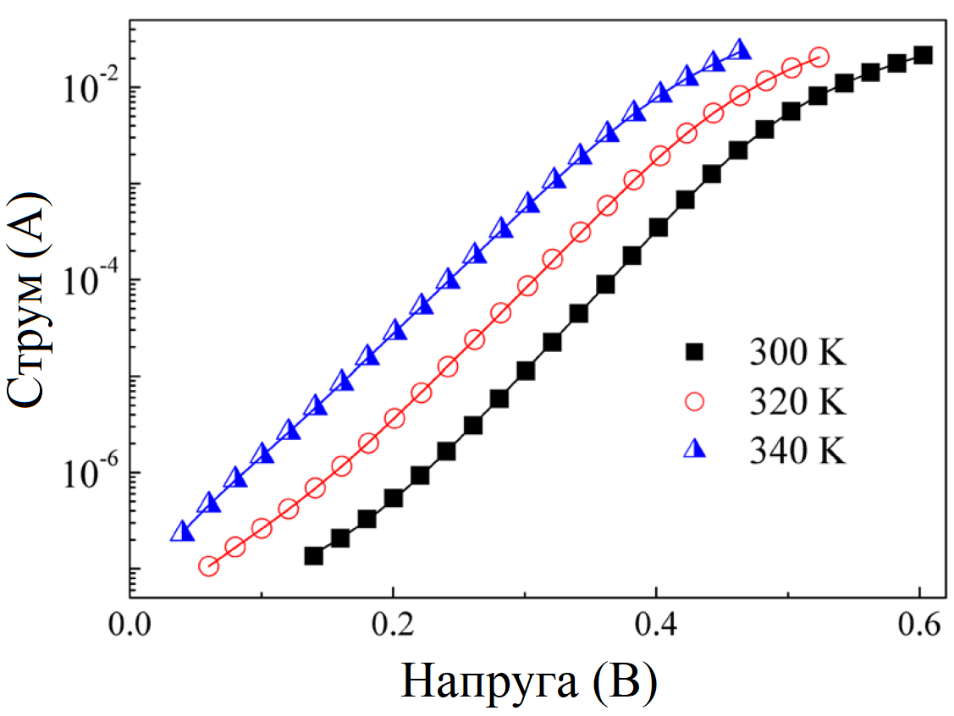


Рис. 4.9 ВАХ, виміряні при 300 К, 320 К і 340 К для зразка SC320. Позначки — це експериментальні результати, а суцільні лінії — криві, апроксимовані згідно з рівнянням (2.12).

Таблиця 4.6 Результати апроксимації реальних ВАХ та передбачень ГНМ

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Зразок | *N*Fe,MEAS, 1012 cм-3 | *Т*, К | *n*FeFeB | RSH, FeFeB,Ом | *n*Fe | RSH, Fe,Ом | *N*Fe,PRED, 1012 cм-3 | | | |
| ГНМFeFeB | | ГНМFeFeB-Fe | |
| трен. | повн. | трен. | повн. |
| SC320 | 2,0±0,4 | 300 | 1,214 | 1,6⋅106 | 1,195 | 1,4⋅106 | 3,9 | 2,8 | 3,0 | 2,0 |
| 320 | 1,204 | 8,6⋅105 | 1,148 | 8,0⋅105 | 6,6 | 1,9 | 16 | 19 |
| 340 | 1,118 | 4,3⋅105 | 1,111 | 4,3⋅105 | 3,8 | 1,2 | 89 | 574 |
| SC349 | 6,7±0,7 | 300 | 1,223 | 2,9⋅106 | 1,222 | 2,6⋅106 | 8,9 | 5,6 | 15 | 11 |
| 320 | 1,183 | 1,7⋅106 | 1,182 | 1,7⋅106 | 1,2 | 0,4 | 10 | 32 |
| 340 | 1,138 | 1,3⋅106 | 1,173 | 1,3⋅106 | 9,8 | 1,7 | 26 | 411 |

Величини факторів неідеальності, що були визначенні з експериментальних кривих та параметрів зразків, були використані як вхідні дані для та , які були попередньо натреновані або на тренувальному або на повному наборі даних. Результати передбачень наведені в Таблиці 4.6.

Перш за все, слід зазначити, що, незважаючи на використання не складної методики моделювання, результати перевершили всі очікування. Зокрема, прогнози моделі відрізнялися від виміряних лише в декілька разів. Для зразка SC320, використовуючи модель , що була натренована на повному наборі даних, похибка передбачень не перевищувала 40%.

Також слід зазначити, що результати, отримані під час експериментальної перевірки, підтверджують тенденції, виявлені в результаті аналізу змодельованих ВАХ. Зокрема, точність прогнозування зменшується при і концентраціях заліза, близьких до верхньої межі досліджуваного діапазону (). Крім того, величина рівня легування бази КСЕ () не використовувалася при створенні тренувального набору розмічених даних, однак використовувалася в тестовому наборі B-altered. З результатів, що наведені в Таблиці 4.6, можна зробити висновок про гарні передбачення , яка була натренована на повному наборі даних, особливо якщо порівнювати з випадком використання тільки тренувального набору для навчання моделі. Цей факт підтверджує зроблений раніше висновок про важливість тренування ГНМ з тими значеннями , які будуть використані під час тестування моделей.

З іншого боку, модель у більшості випадків демонструвала гірші результати, ніж . Існує кілька можливих причин такого результату:

1. використання двох значень фактора неідеальності підвищує чутливість моделі до спрощень, закладених у процесі моделювання (зокрема, до впливу процесів, які не враховувалися під час моделювання, наприклад, появи послідовного та шунтуючого опорів).

2. з експериментальної точки зору, визначення є більш складним завданням, ніж визначення . Зокрема, для проведення вимірювань ВАХ після інтенсивного освітлення було необхідно витримати зразок протягом близько 100 секунд. Цей інтервал забезпечував стабілізацію температури зразка після нагріву, індукованого світлом, та дозволяв виконати точні вимірювання струму і напруги.

З іншої сторони, відповідно до літературних даних, характерний час утворення пари FeB при *T* = 340 К та становить близько 600 с. Отже, за таких умов неможливо вважати, що отримане значення фактора неідеальності відповідає повністю дисоційованому стану пари FeB. Таким чином, попри потенційно вищу точність прогнозування моделі її практичне використання ускладнюється низкою обмежень.

**Висновки до розділу 4**

1. Був розроблений машинно-орієнтований підхід для швидкої оцінки концентрації заліза у КСЕ на основі фактора неідеальності, температури, товщини бази та рівня легування. В рамках дослідження проводилося моделювання структури з метою створення тренувального та тестового наборів розмічених даних. Загалом, було змодельовано близько 7524 ВАХ для тренування та 3896 ВАХ для тестування ГНМ. Додатково, моделі були випробувані з використанням параметрів реальних КСЕ.

2. Було налаштовано дві моделі ГНМ: , що мала 4 вхідні параметри - , , , та , яка мала 5 вхідних параметрів - , , , , . Для навчання була використання 10-ти кратна перехресна перевірка та комбінований підхід до підбору гіперпараметрів (grid search і random search).

3. Показано, що найнижчі похибки прогнозування досягаються при високому рівні легування, низькій температурі та використанні обох факторів неідеальності. Середня відносна квадратична похибка не перевищує 10% для більшості тестових вибірок, а коефіцієнт детермінації перевищує 0,98 у найбільш репрезентативних випадках. Отримано, що найбільші похибки спостерігаються у випадках низьких концентрацій бору та високих температур (), що пов’язано зі зменшенням внеску власної рекомбінації за механізмом ШРХ у фактор неідеальності.

4. Під час перевірки методики на реальних ВАХ було показано, що тренувати ГНМ треба на значеннях , що відповідають рівню легування досліджуваних структур.

Матеріали цього розділу були опубліковані в журналі [ProgressInPhotovoltaics\_2022\_Q1], а також були представлені на конференції [ECRES\_2021\_Istambul]. Треба зауважити, що в розділі наведена не повна інформація про вигляд altered-наборів та лише найбільш інформативні графіки залежностей. Більш детально з матеріалами розділу можна ознайомитися за посиланням [https://github.com/Zavhorodnii-Oleksii/supplementary\_materials\_for\_thesis/blob/main/Розділ%204.pdf] .