**4. Використання фактору неідеальності та методології глибокого навчання для оцінки концентрації забрудюючого заліза в кремнієвих сонячних елементах.**

У попередньому розділі ми показали, що концентрацію заліза можна оцінити за допомогою фактора неідеальності КСЕ. Однак аналітичні залежності між та є складними, неуніверсальними й чутливими до багатьох фізичних параметрів, що ускладнює їхнє практичне застосування. У цьому розділі ми будемо розглядати 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.

У доповіді на конференції ECRES'21 ми вперше запропонували таку швидку та просту методику оцінки концентрації заліза в КСЕ [ECRES\_2021\_Istambul]. Згодом, у нашій статті [ProgressInPhotovoltaics\_2022\_Q1], були представлені результати тестів наших ГНМ на змодельованих та на реальних сонячних елементах, що містили домішкове залізо. У дослідженні використовувалася перша РМКСЕ.

**4.1 Розробка та навчання глибоких нейронних мереж**

Схематична методологія 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 включала в собі товщину бази (, логарифм концентрації бору , температуру та фактор неідеальності для стану, в якому співіснують комплекси та міжвузлові атоми ; друга включала в собі додатково до вже згаданих параметрів - фактор неідеальності для стану, в якому в КСЕ наявні тільки міжвузлові атоми заліза . Відповідно, розглядалися дві моделі ГНМ: та .

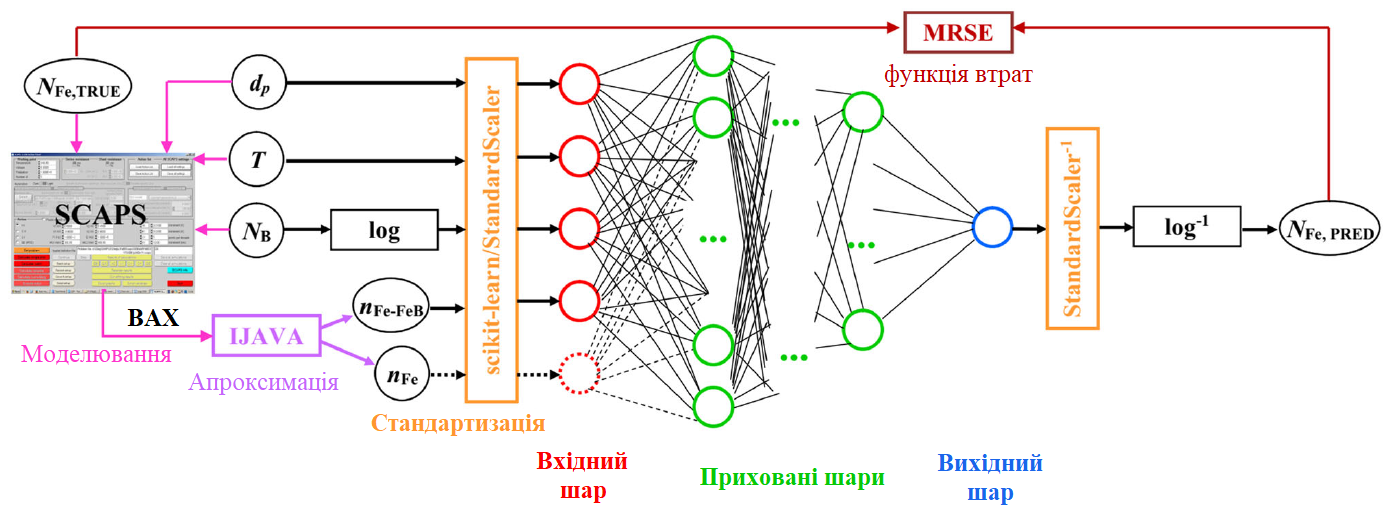


Рис. 4.1 Схема підходу на основі глибокого навчання для прогнозування .

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: «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. У всіх випадках для прихованих шарів використовувалася функція активації ReLU, а для оптимізації мереж - алгоритм Adam. Для запобігання перенавчання впроваджувалися регуляризаційні методи, зокрема Dropout.

Для кожної з ГНМ ми налаштовували 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*).

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*):

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

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**

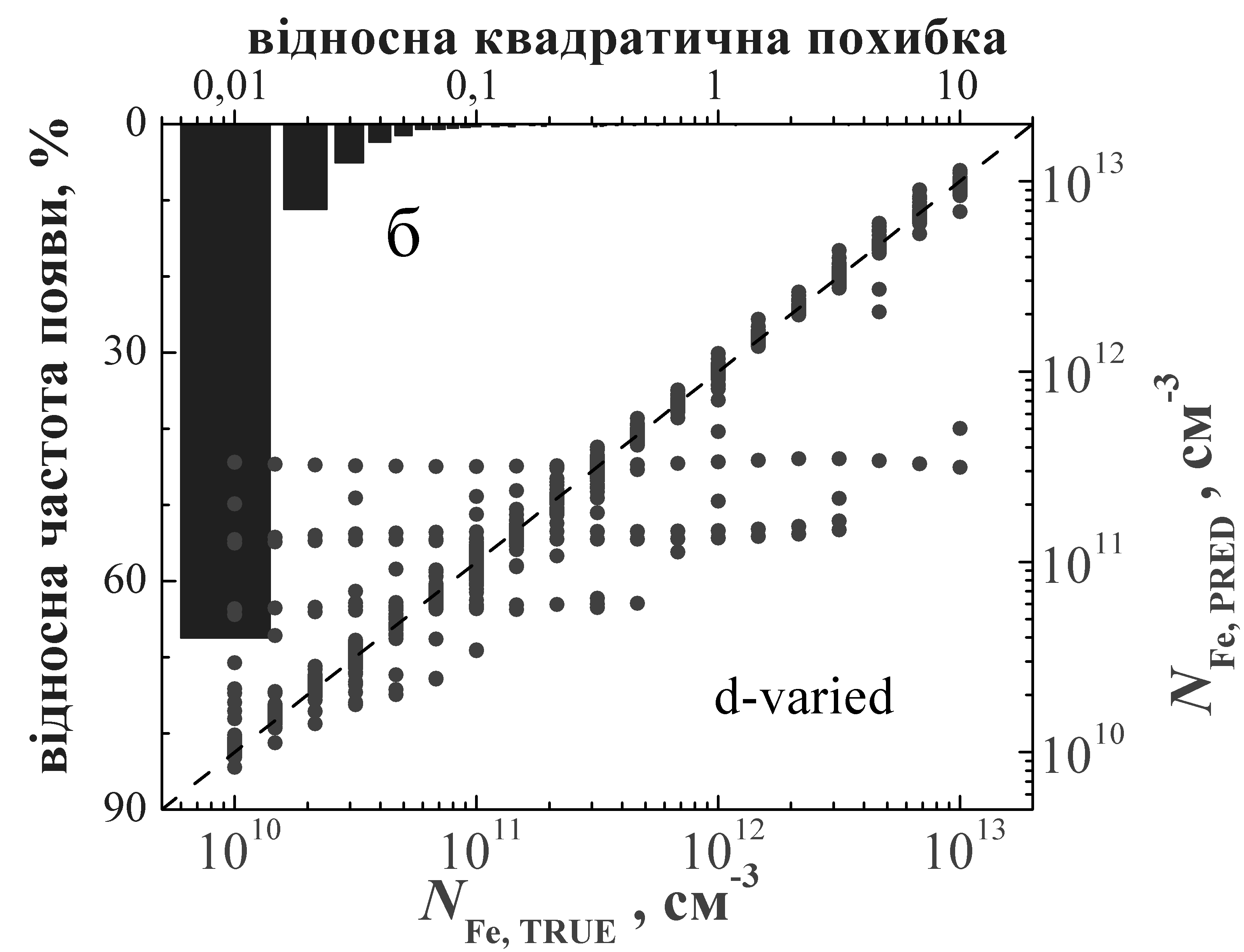
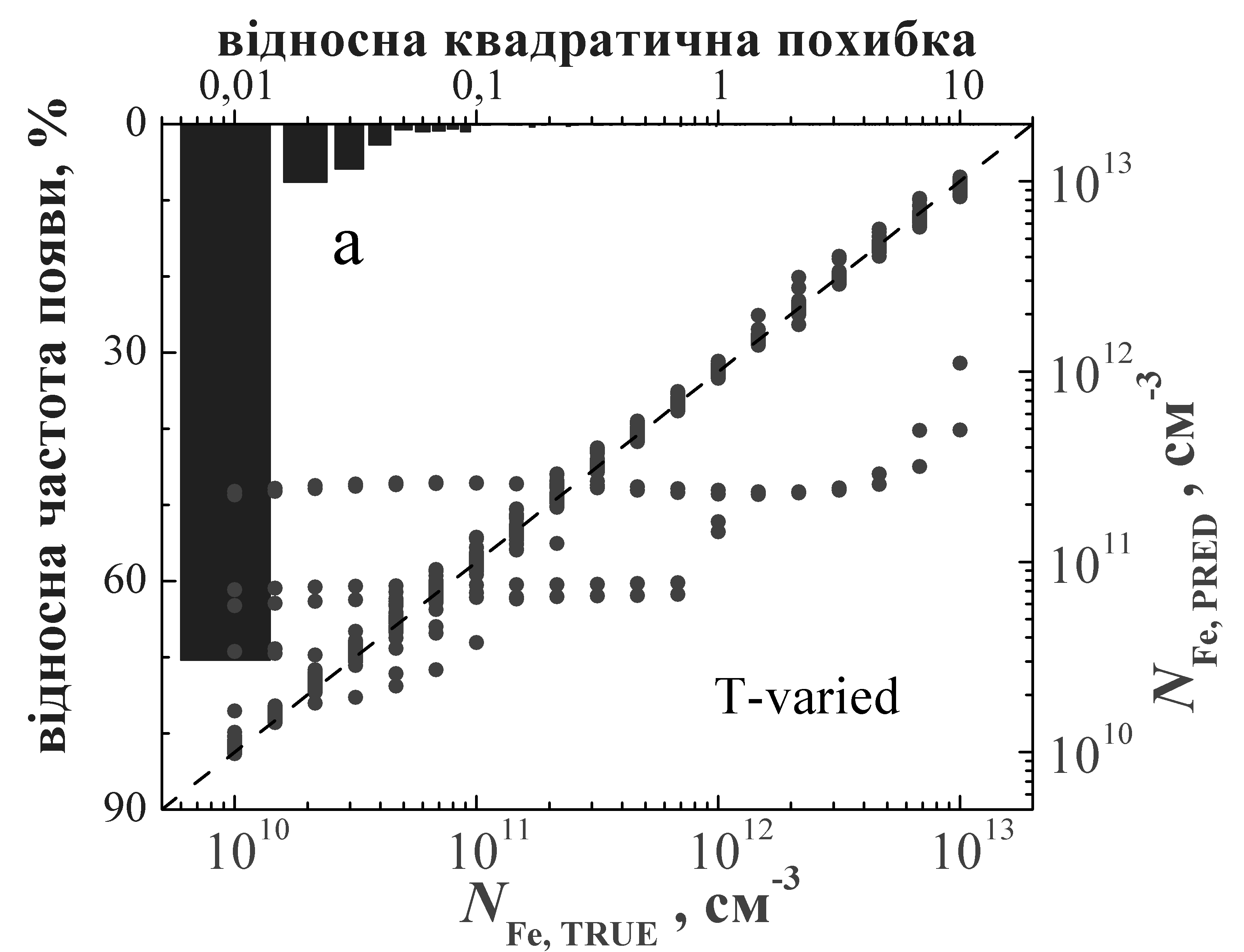
Окрім навчальної вибірки, були сформовані незалежні тестові вибірки, у яких варіювалися або окремі параметри (, , , ) або всі параметри одночасно, причому використовувалися такі значення параметрів, які не зустрічалися в тренувальному наборі. Набір «Fe-varied» складався з двох під наборів: при підготовці першого використовували концентрації заліза – {1., 2.471, 4.696, 8.927, 1.697, 3.225, 6.130, 1.165, 2.214, 4.209, 8.000} (не використовувалися в тренувальному наборі), температури *–* {290, 295, 300, 305, 310, 315, 320, 325, 330, 335, 340} K (використовувалися), товщину бази - 180 мкм (використовувалася) та концентрації бору –{, 1.778, 5.623, , 3.162, } (використовувалися); при підготовці другого використовували концентрації заліза – {1.2, 2.234, 4.160, 7.746, 1.442, 2.685, 5} (не зустрічалися у тренувальному наборі), температури {290, 300, 310, 320, 330, 340} K (зустрічалися), товщини бази {210, 240} мкм (зустрічалися) та концентрації бору {3.162, , } (зустрічалися). Загалом, набір «Fe-varied» мав 857 ВАХ. Аналогічно до «Fe-varied» ми створили тестові набори «d-varied» (1189 ВАХ), «T-varied» (832 ВАХ), «B-varied» (514 ВАХ) та All-varied (684 ВАХ).

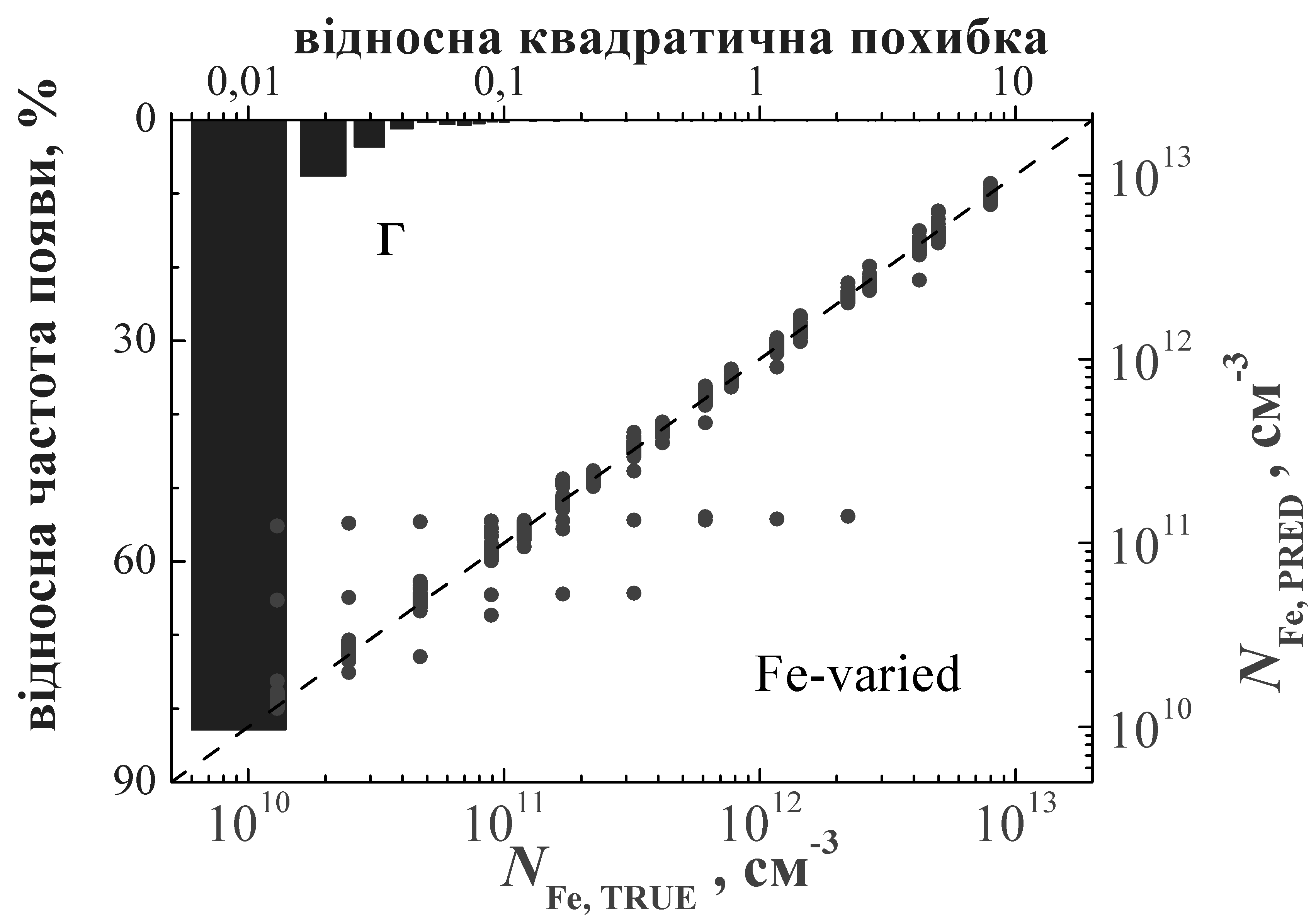
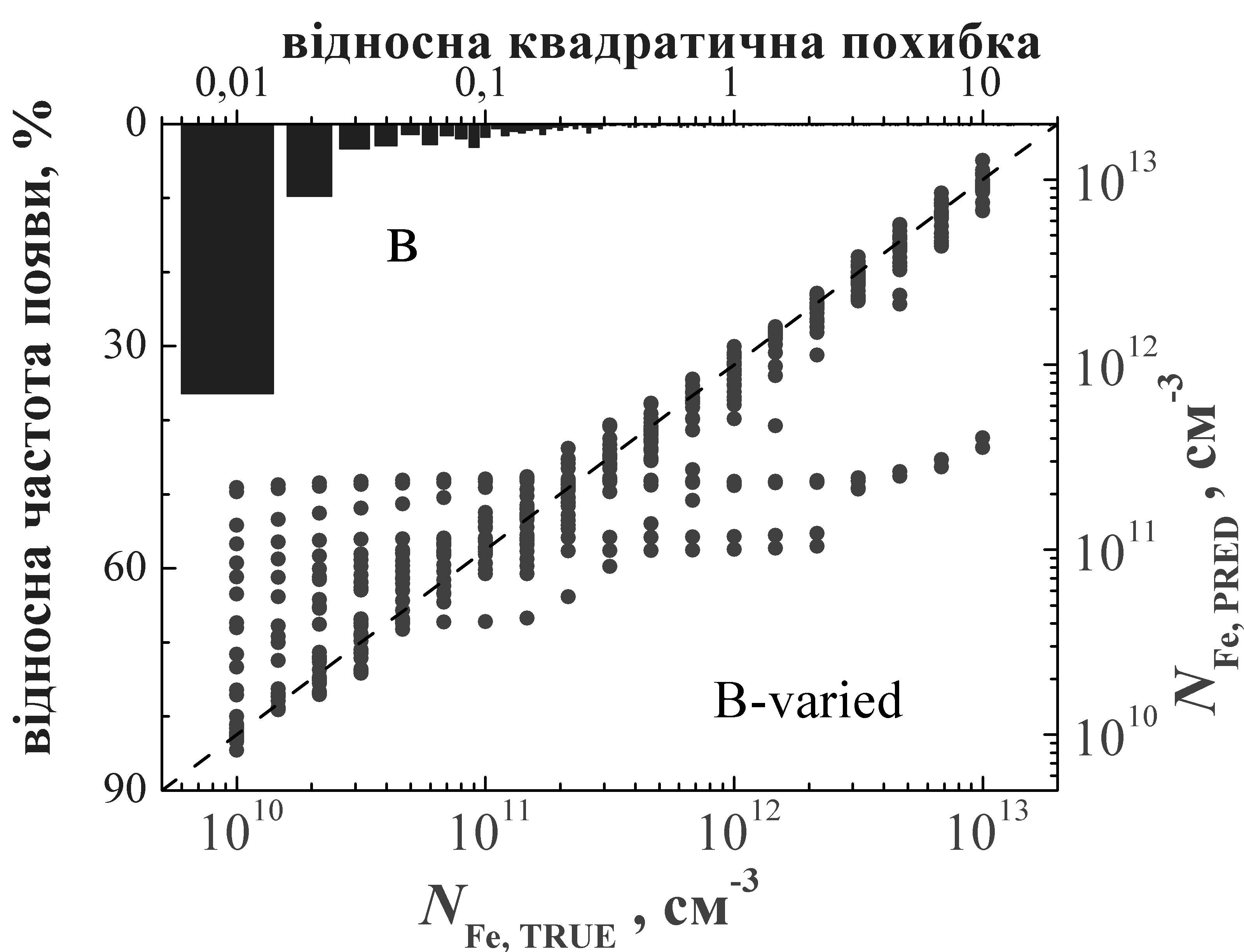
В Таблиці 4.2 наведені найефективніші набори гіперпараметрів для наших ГНМ, які ми визначили за допомогою Keras Tuner, а в Таблиці 4.3 результати 10-ти кратної перехресної перевірки.

Таблиця 4.2 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 4.3, таблиці 4.4 та на рис.4.2. З отриманих результатів можемо бачити, що відносна квадратична похибка (ВКП) 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.





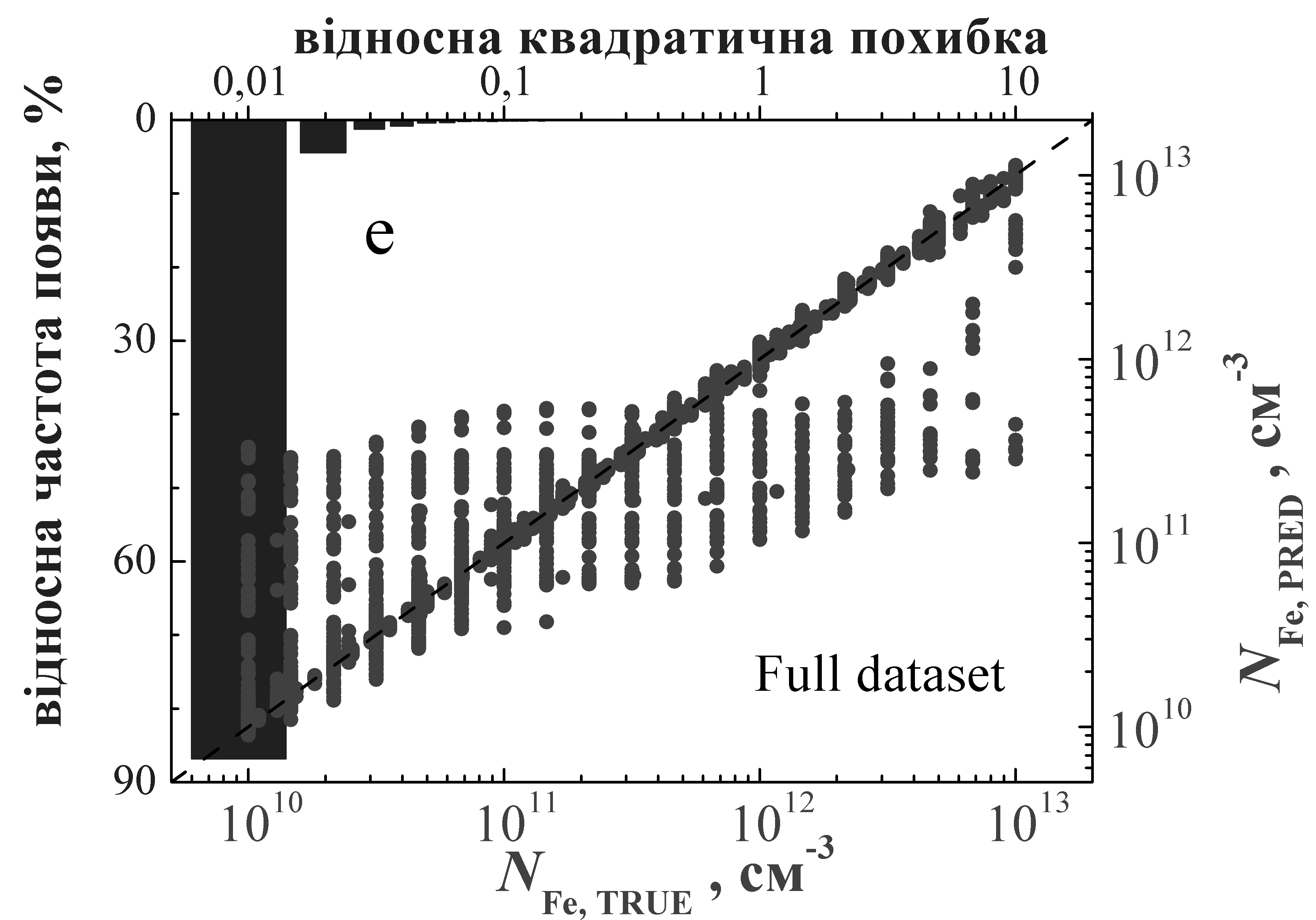
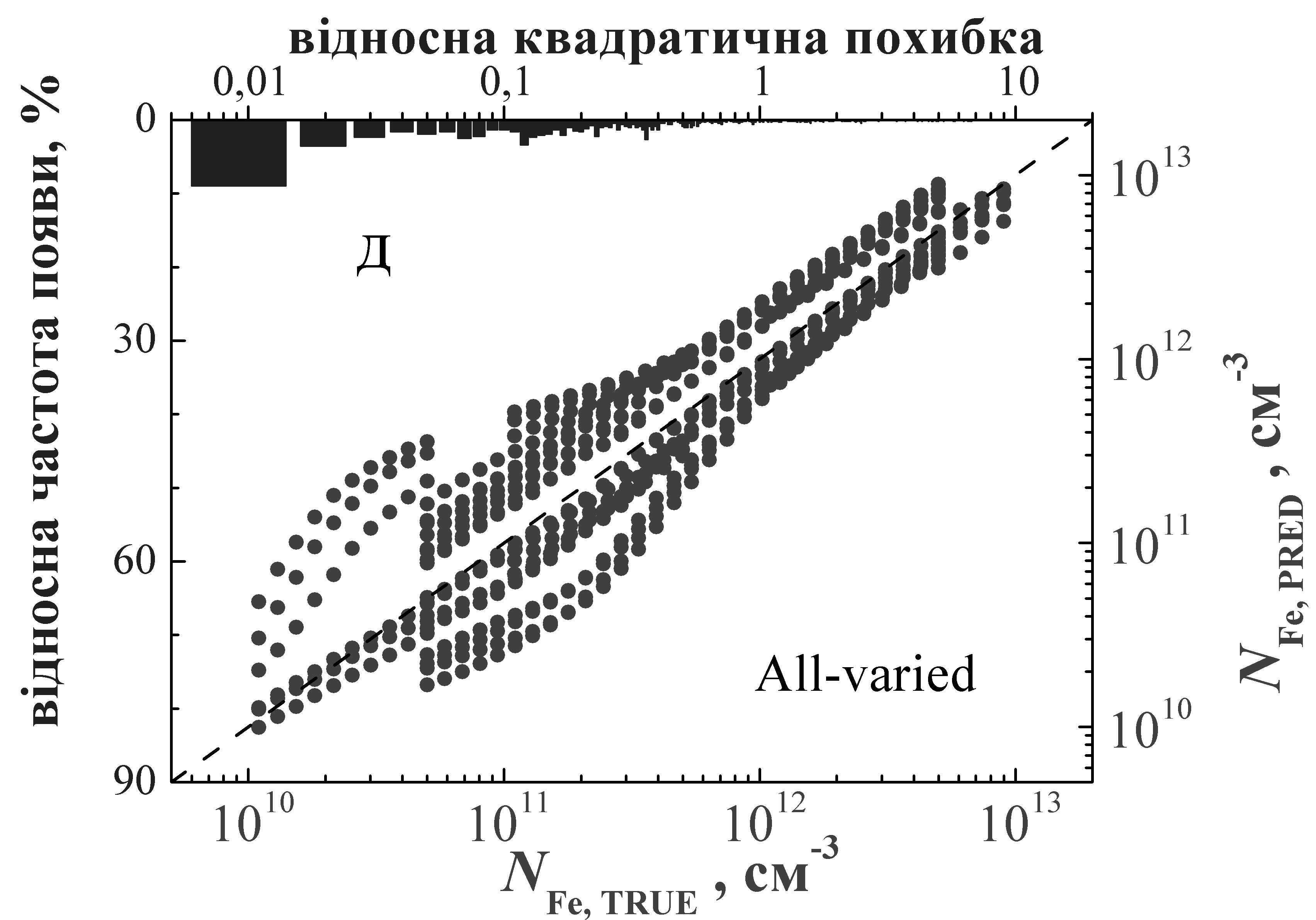


Рис. 4.2 Результати прогнозування моделі на наборах даних T-varied (A), d-varied (B), B-varied (C), Fe-varied (D), All-varied (E) та Full (F) (червоні точки). Стовпчики представляють гістограми відносної квадратичної похибки. Чорні пунктирні лінії є еталонними прогнозами.

Зокрема, для 87%, 88% та 96% для T-varied, d-varied та Fe-varied відповідно, ВКП не перевищує 0,05 (див. Рис. 4.2).

Таблиця 4.3 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 4.4 Результати 10-кратної перехресної перевірки

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

Для тестової вибірки B-varied, найбільше значення MSRE = 1,06 пов’язане переважно з невеликою кількістю зразків, для яких ВКП>20, тоді як для 54% зразків ВКП<0,05. Найгірші результати прогнозування, як і очікувалося, спостерігаються для тестової вибірки All-varied: коефіцієнт становить лише 0,813, а ВКП<0,05 спостерігається лише для 18% зразків. Водночас, для вибірки Fe-varied, 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. 4.3–4.6).

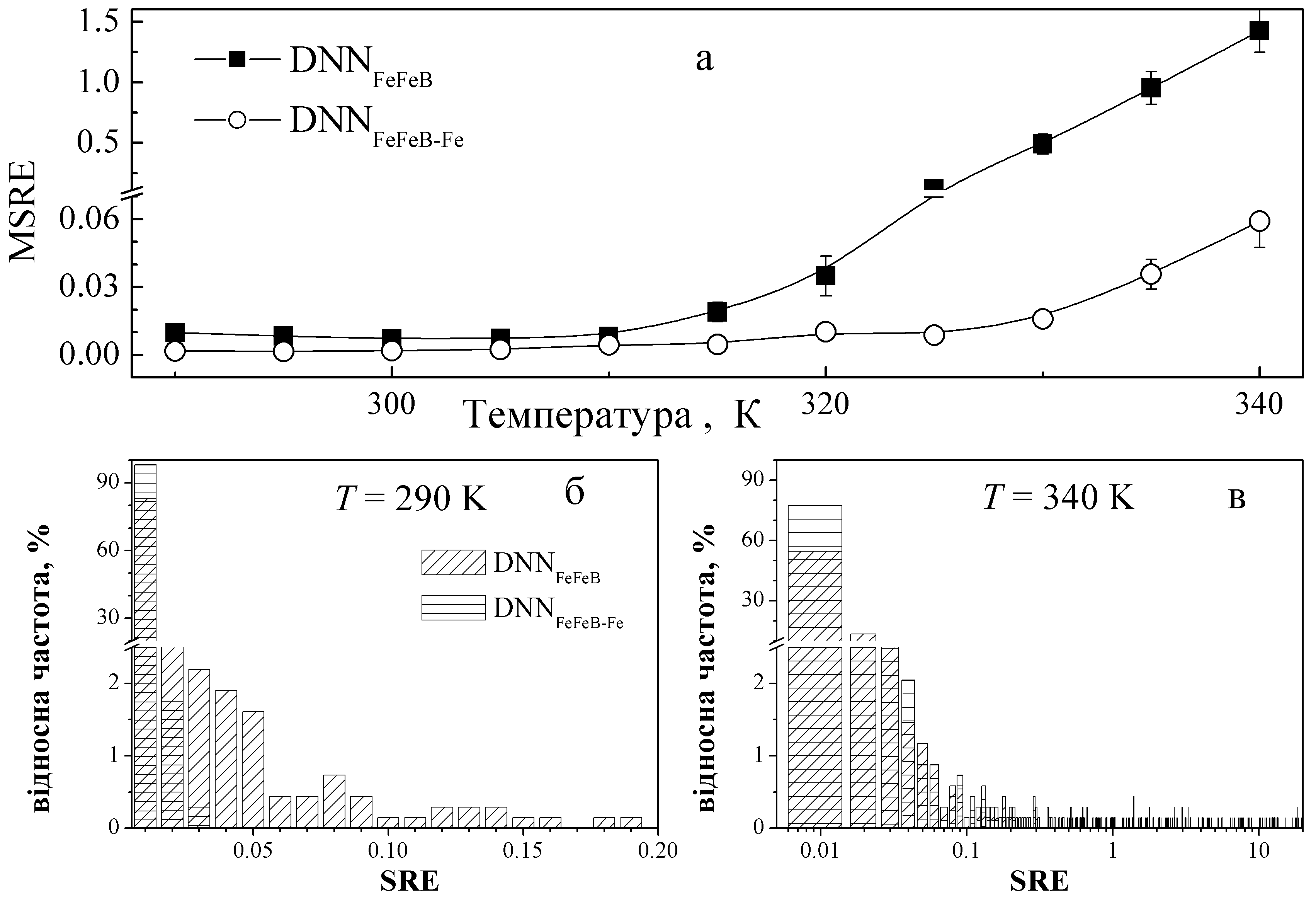


Рисунок 4.3 – Залежність MSRE від температури (а) для тренувального набору. Гістограми частоти появи ВКП для *T* = 290 К (б) та *T* = 340 К (в). Похиле штрихування - ; горизонтальне ‑ .

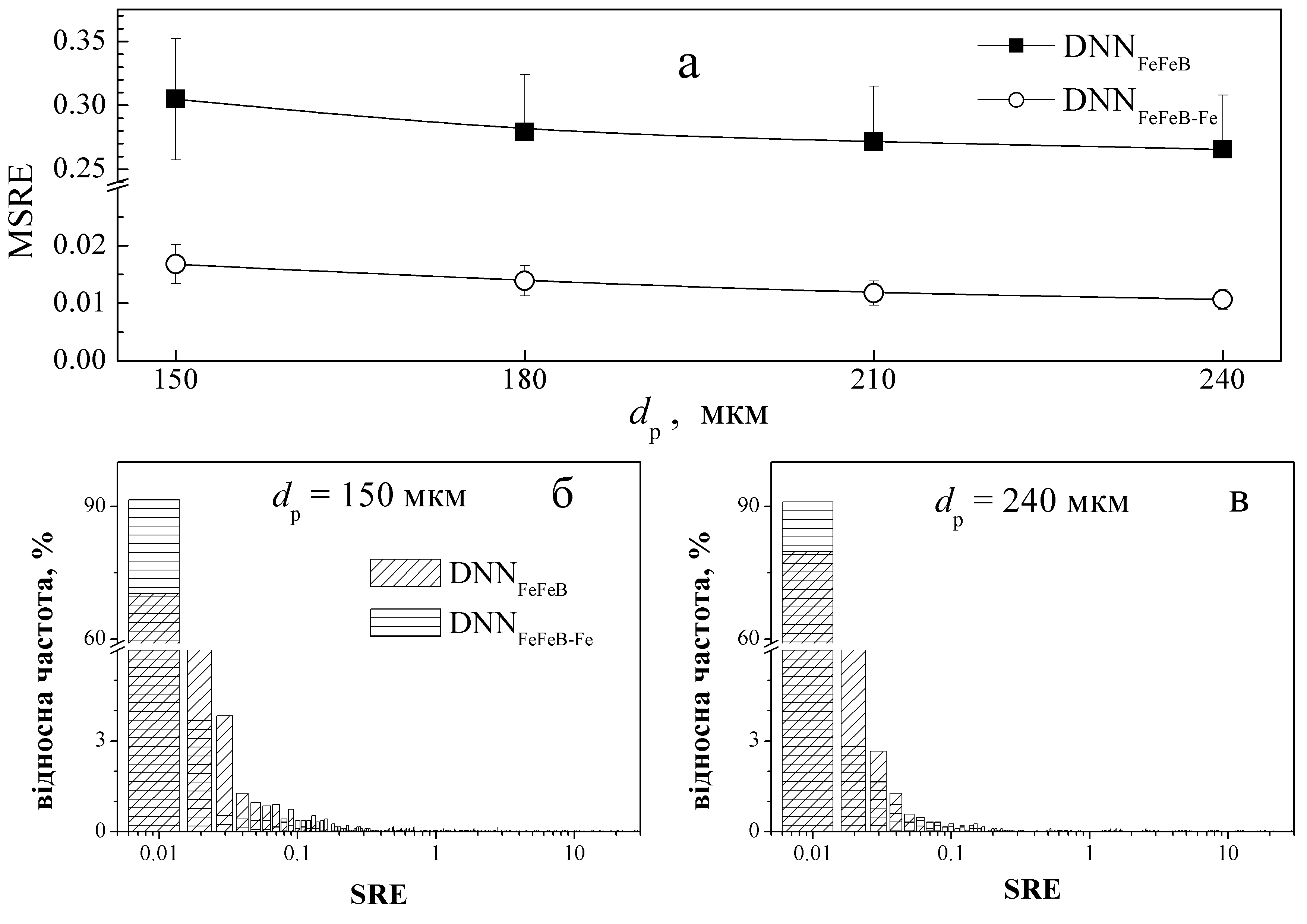


Рисунок 4.4 – Залежність MSRE від товщини бази (а) для тренувального набору. Гістограми частоти появи ВКП для  = 150 мкм (б) та  = 240 мкм (в). Похиле штрихування - ; горизонтальне ‑ .

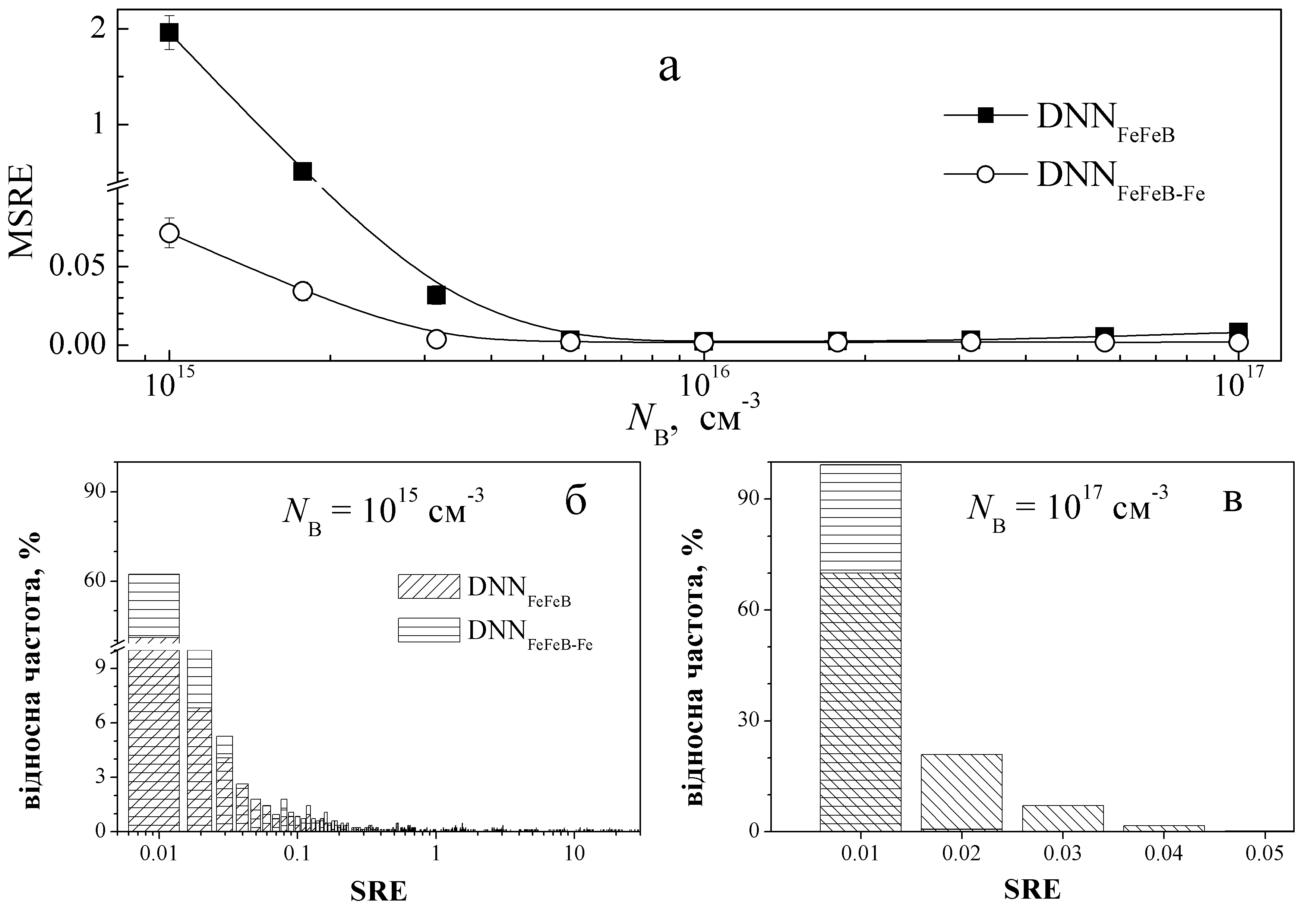


Рисунок 4.5 – Залежність MSRE від концентрації бору (а) для тренувального набору. Гістограми частоти появи ВКП для (б) та  (в). Похиле штрихування - ; горизонтальне ‑ .

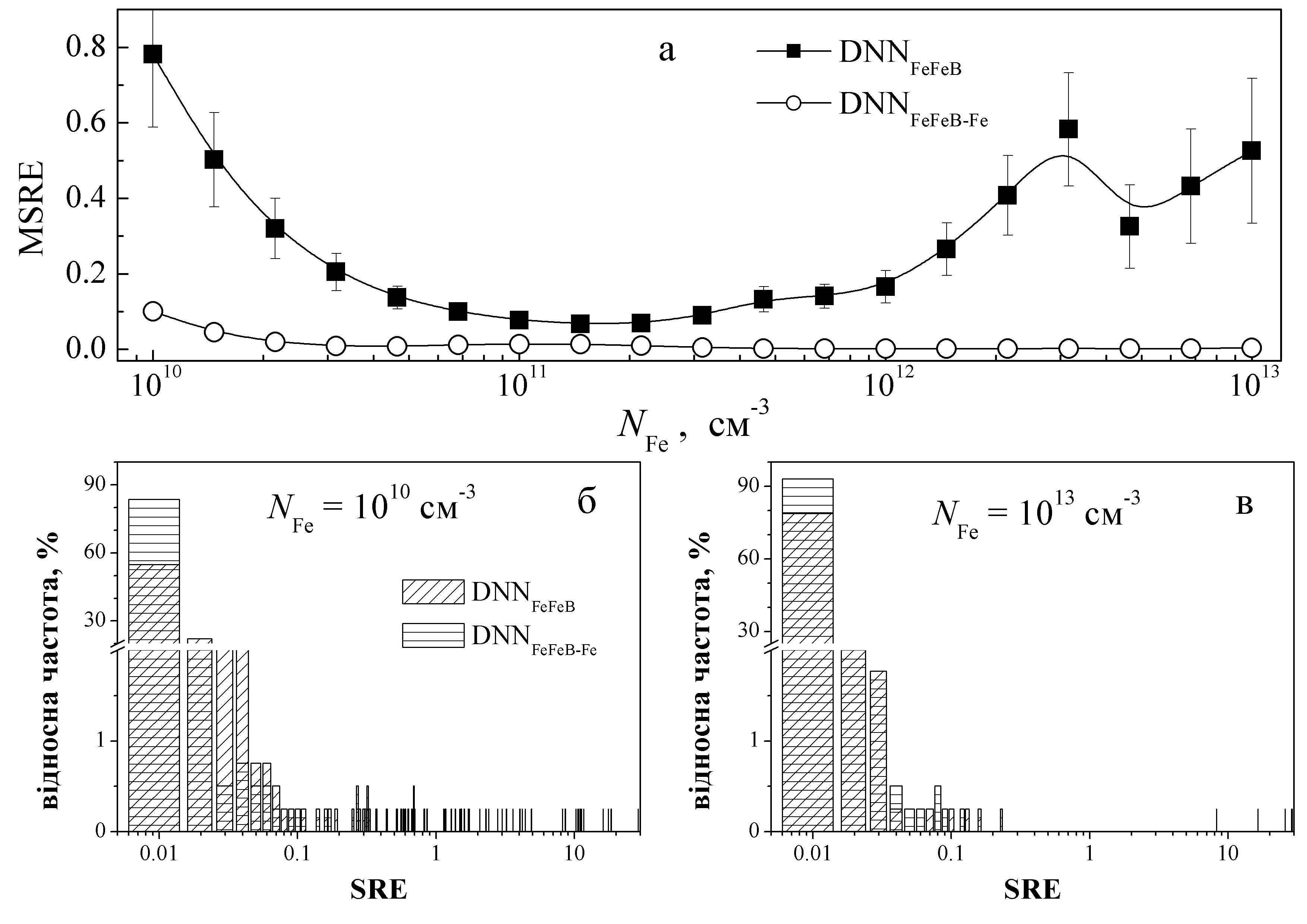


Рисунок 4.6 – Залежність MSRE від концентрації заліза (а) для тренувального набору. Гістограми частоти появи ВКП для (б) та  (в). Похиле штрихування - ; горизонтальне ‑ .