

ОСОБЛИВОСТІ ФОРМУВАННЯ ПЕРВИННОГО ОПИСУ ОБ'ЄКТІВ МОНІТОРИНГУ ДОВКІЛЛЯ НА МІКРОРІВНІ

Голуб С.В.

Черкаський національний університет ім. Богдана Хмельницького

Absrtact

Golub S.V. Features of forming of primary description of objects of monitoring of environment on a microlevel. *Is offered technology of formation of the initial description for inductive modeling on microlevel of monitoring of objects of an environment. On an example voltamperometric measurements of zinc ions concentration in a multicomponent solution it is determined, that the offered technology is effective at application of inductive methods of modeling of objects of electrochemical measurements. The error of modeling decreases.*

Моделі мікрорівня є основою функціонування технічних засобів моніторингу — вимірювальних приладів та інформаційно-вимірювальних систем. На мікрорівні моделюється об'єкт вимірювання. Основною процедурою моделювання об'єктів мікрорівня є пошук аналітичних залежностей результатів вимірювання від значення аналітичного сигналу, параметрів режиму вимірювання та впливаючих факторів навколишнього середовища.

Модель об'єкта вимірювання (ОВ) описується функціями вигляду [1]:

$$Y=f(x, q_1, \dots, q_m, q_{m+1}, \dots, q_n), \quad (1)$$

де x — аналітичний сигнал,

q_1, \dots, q_m — конструктивні параметри, що характеризують засіб вимірювання (ЗВ);

q_{m+1}, \dots, q_n — дестабілізуючі фактори оточуючого середовища (збурення).

При цьому аналітичний сигнал є функцією вигляду:

$$x=F(t, z, w, \dots, A, B, C, \dots), \quad (2)$$

де: t, z, w — незалежні аргументи (час, параметри режиму вимірювання, частота),

A, B, C — параметри сигналу.

Основою для створення моделі ОВ є первинний опис об'єкта. Він містить модельовану функцію (результат вимірювання) та параметри моделювання — аналітичний сигнал, конструктивні параметри ЗВ, збурення навколишнього середовища, незалежні аргументи та інші фактори, які впливають на результат вимірювання.

Для аналізу первинного потоку інформації широко використовуються статистичні методи [2]. Вони застосовуються для оцінки вигляду функції розподілу та її параметрів, перевірки гіпотез незалежності даних, їх однорідності, стаціонарності та нормальності розподілу. За допомогою кореляційного аналізу оцінюються зв'язки між параметрами об'єкта спостереження а лінійний та нелінійний регресійний аналіз дозволяє визначити функціональну залежність між цими параметрами. Найбільш популярними є статистичні інформаційні системи Statistica, Statgraphics, DataDesk, Systat, TimeLab, Scenario та інші.

В електрохімічних методах аналізу, які широко використовуються в процесі моніторингу навколишнього середовища, найбільш застосованими є регресійні моделі, отримані за допомогою методу найменших квадратів.

Актуальність наукових досліджень процесу моделювання об'єктів моніторингу мікрорівня є низька точність моделей електрохімічних вимірювань. Основною вимогою до вольтамперометричних засобів вимірювання концентрацій важких металів у воді є похибка вимірювання, що не перевищує 25% при вимірюванні на рівні 0,5 гранично допустимих концентрацій (ГДК).

Для підвищення точності моделей автором запропоновано [3] розглядати об'єкт електрохімічного вимірювання як складну систему і для його моделювання застосувати індуктивні метод моделювання [4].

Метою даної роботи було дослідження особливостей формування первинного опису ОВ для застосування його в технологіях індуктивного моделювання.

Особливості застосування індуктивних методів при моделюванні об'єктів електрохімічних досліджень висуває велику кількість вимог до підсистем підготовки первинної інформації. Первинний опис об'єкта моделювання на мікрорівні являє собою таблицю значень параметрів моделювання, які зареєстровані при фіксованих значеннях модельованої функції. Значення первинного опису об'єкта вимірювання отримують шляхом проведення серії випробувань, які в традиційних методиках вимірювання описується як процес калібрування приладу, або отримання каліброваної залежності. Результати вимірювання фіксуються при послідовній зміні значення кожного із параметрів вимірювання при фіксованому значенні всіх інших параметрів. Таким чином реалізується методика багатофакторного експерименту.

Для створення первинного опису об'єкта моделювання мікрорівня необхідно звернутись до методик вимірювання, які в більшості випадків враховують основні впливаючі фактори.

Спільною вимогою до первинного опису ОВ є визначення ефективного способу застосування принципу зовнішнього доповнення [4]. Із спостережень первинного опису виділяється їх частина, яка використовується для визначення характеристик вже створених моделей і зветься перевіряючою послідовністю.

Для розбиття таблиці експериментальних вхідних даних на навчаючу та перевіряючу послідовності використовують кілька способів [4]. Таблиця вхідних даних ділиться на дві частини або рівні, або обсяг навчаючої послідовності більша за обсяг перевіряючої. В деяких випадках рядки з непарним номером складають навчаючу послідовність, з парним — перевіряючу.

Перед розбиттям застосовують процедуру ранжування — розміщують рядки таблиці за величиною дисперсії модельованої функції від значення її середнього арифметичного [5].

Пропонується формувати послідовності за результатами моделювання [6], організовуючи перебір способів формування цих послідовностей за критерієм незміщеності, враховуючи результати аналізу функції розподілу зовнішнього критерію [7].

В [4] описаний спосіб розділення вхідних даних на навчаючою та перевіряючою послідовністю за ознакою мінімуму рядів селекції багаторядного алгоритму МГВА. При ранжуванні вхідних даних за дисперсією було експериментально виявлено існування мінімуму рядів селекції при зміні співвідношення кількості точок навчаючої та перевіряючої послідовності. Вважається, що чим менше рядів селекції та простіша модель, тим вона достовірніша.

Ефективним вважається розбиття за критерієм мінімізації дисперсії помилки прогнозу за моделлю оптимальної складності [8].

Проблема застосування способу виділення зовнішнього доповнення в кожному конкретному випадку вирішується евристично залежно від об'єкту моделювання. При застосуванні індуктивних методів при електрохімічному аналізі не завжди є ефективним застосування існуючих способів виділення перевіряючої послідовності.

Основною проблемою, яка виникає при застосуванні індуктивних методів при моделюванні ОВ є те, що як правило методики вимірювання враховують від 2 до 5 впливаючих факторів, які використовуються як параметри моделі об'єкта вимірювання. Зрідка кількість одночасно врахованих параметрів доходить до 8–10. Для створення адекватних моделей об'єкту на основі методів індуктивного моделювання часто інформації

від цих параметрів недостатньо. Постає завдання здобування додаткової інформації. Для цього розробляються технології формування первинного опису ОВ для їх адаптації до індуктивних методів моделювання.

Для подальшого використання пропонується застосувати принцип групування спостережень за ознакою найсильніших зв'язків. Процеси навчання інформаційних моделей об'єктів електрохімічних вимірювань та визначення їхньої точності повинні проводитись на послідовностях, які максимально незалежні одна від одної. Розбиття спостережень на навчаючу та перевіряючу послідовності необхідно проводити за ознакою корельованості рядів цих спостережень між собою. Ця технологія застосування зовнішнього критерію отримала назву „Формування найближчих сусідів”.

Як об'єкт моделювання використаємо залежність концентрації цинку від параметрів методики інверсійного вольтамперометричного вимірювання. Для формування первинного опису об'єкту моделювання було проведено 50 спостережень, в процесі яких фіксувались значення модельованої функції — концентрації цинку, та 5 параметрів моделювання: потенціалу та часу концентрування, значення аналітичних сигналів окислювання міді, цинку та свинцю. Опрацювання первинного опису починалось із визначення коефіцієнтів кореляції між спостереженнями, тобто між рядками матриці. Визначався рядок із найбільшим коефіцієнтом кореляції по відношенню до першого рядка. Після цього рядок розміщувався після першого. Цю процедуру проробили зі всіма спостереженнями. У випадку, коли останній перенесений рядок мав найбільший коефіцієнт кореляції із рядком, який вже переміщений, вибирався наступний по спаданню коефіцієнт кореляції до тих пір, поки не визначався рядок, який іще не переміщався.

Для моделювання використовувався багаторядний алгоритм МГУА, [4]. На наступний ряд селекції переводилась кількість моделей, що рівнялась кількості початкових змінних. Для підготовки первинного опису використана технологія „Інформаційний прес” [9]. Використано 20 додаткових параметрів — коефіцієнти кореляції з 20 першими рядами селекції. Оцінювались абсолютна похибка моделі, як загальна характеристика та сума квадратів похибок, що відображає якість застосування моделі для кожного спостереження. На відміну від попереднього дослідження зупинявся процес генерації наступних рядів селекції за ознакою середньої квадратичної похибки ряду та за середнім квадратичним відхиленням моделі.

В таблиці поданий план модельного експерименту та його результати.

Із таблиці видно, що ефективність технології „Формування найближчого сусіди” залежить від вигляду опорної моделі, критерію селекції та способу виділення навчаючої послідовності. При формуванні первинного опису об'єкта електрохімічних досліджень параметри генератора інформаційної моделі необхідно підбирати за результатами моделювання. Таким чином технологія формування первинного опису дозволяє знизити похибку моделювання. В деяких випадках (дослід № 3, таблиці) при селекції моделей за показником якості ряду похибка знижується на порядок. Особливо ефект покращення точності моделі спостерігається при порівнянні показника суми квадратів похибок.

Висновок.

Запропонована технологія є ефективною при формуванні первинного опису об'єктів електрохімічних досліджень для їх індуктивного моделювання. Об'єктом для наступних досліджень є процес застосування запропонованої методики для моделювання об'єктів моніторингу макрорівня.

Література

1. Алиев Т.М., Тер-Хачатуров А.А. Измерительная техника: Учеб. Пособие для техн. вузов. — М.: Высш. шк., 1991. — 384 с.

Таблиця — План та результати досліджень технології „Формування найближчих сусідів”

№	Вигляд опорної моделі	Спосіб виділення послідовностей спостережень	Абсолютна похибка моделі, мг/дм ³		Сума квадратів похибок, мг/дм ³		Відносна похибка моделі, %			
			Селекція моделей за середньою квадратичною похибкою ряду						Селекція моделей за середнім квадратичним відхиленням моделі	
			Традиційна технологія	Нова технологія	Традиційна технологія	Нова технологія	Традиційна технологія	Нова технологія		
1	$Y=a_0+a_1x_1+a_2x_2$	Непарні — навчаюча, парні — перевіряюча	0,22	0,26	8,25	3,98	25.86	28.43		
2	$Y=a_0+a_1x_1+a_2x_2$	Перша половина спостережень — навчаюча, друга — перевіряюча	0,32	0,37	30,68	8,10	40.64	36.35		
3	$Y=a_0+a_1x_1+a_2x_2+a_3x_1x_2$	Непарні — навчаюча, парні — перевіряюча	0,001	8,33E-10	0,0002	3,02E-16	4.66	6.97		
4	$Y=a_0+a_1x_1+a_2x_2+a_3x_1x_2$	Перша половина спостережень — навчаюча, друга — перевіряюча	0,16	0,38	5,002	9,38	43.96	25.44		
5	$Y=a_0+a_1x_1+a_2x_2+a_4x_1^2+a_5x_2^2$	Непарні — навчаюча, парні — перевіряюча	0,29	0,05	4,748	0,24	23.84	18.46		
6	$Y=a_0+a_1x_1+a_2x_2+a_4x_1^2+a_5x_2^2$	Перша половина спостережень — навчаюча, друга — перевіряюча	0,24	0,26	7,96	5,59	55.75	42.25		

- Киселев М., Соломатин Е. Средства добычи знаний в бизнесе и финансах. — Открытые системы, 1997, №4. — С. 41–44.
- Голуб С.В. Моделирование мониторинговых процессов эвристической системой наблюдений. Электронное моделирование. — Киев. — Т. 26. — № 5. — С. 55–65.
- Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем. — Киев: Наук. думка, 1981. — 296 с. — С. 38.
- Павлов О.А. Критерий ранжирования для порогового самовідбору змінних в алгоритмах МГУА // Автоматика. — 1969. — № 4. — С. 89–91.
- Висоцький В.М. Про найкращий поділ вхідних даних в алгоритмах МГУА // Автоматика. — 1976. — № 3. — С. 71–74.
- Юрачковський Ю.П., Горшков А.Н. Оптимальное разбиение исходных данных на обучающую и проверочную последовательности на основе анализа функции распределения критерия. // Автоматика. — 1980. — № 2. — С. 5–9.
- Степашко В.С., Кондрашова Н.В. Исследование способов генерации вариантов разбиения выборки в алгоритмах МГУА. Праці I Міжнародної конференції з індуктивного моделювання, Львів 20–25 травня 2002. — Т.1, Ч.1. — С. 90–94.
- Голуб С.В. Використання технології „Інформаційний прес” при індуктивному моделюванні об’єктів оточуючого середовища. // Автоматика. Автоматизация. Электротехнические комплексы и системы. Видавництво ХДТУ. — 2006.