УДК 519.853.6

ЗАСТОСУВАННЯ MLP-МЕТАМОДЕЛЕЙ В ЗАДАЧАХ СУРОГАТНОЇ ОПТИМІЗАЦІЇ

Трембовецька Р.В., Гальченко В.Я., Тичков В.В.
Черкаський державний технологічний університет

Розроблена обчислювальна технологія побудови метамоделей, що використовує сучасні досягнення в області теорії планування експерименту, інтелектуального аналізу даних і штучного інтелекту та визначені експериментальним шляхом закономірності, які дозволяють ефективно виконувати побудову метамоделей. Побудова метамоделей виконана на чисельних прикладах з функцією цілі, яка залежить від двох змінних. В якості плану експерименту використано генератор точок, що заповнює простір пошуку та в процесі реалізації якого використовуються ЛП$_\tau$-послідовності Соболя. Встановлено, що збільшенням кількості точок плану експерименту та кількості прихованих нейронів до певного рівня, параметри метамоделей покращуються; використання різних слабко корельованих ЛП$_\tau$-послідовностей з їх повного набору суттєво не впливають на показники якості кінцевої метамоделі. Результати чисельних експериментів свідчать про можливість використання запропонованої обчислювальної технології побудови MLP-метамоделей для апроксимації первинних функцій цілі з достатньо складною поверхнею відгуку.

Ключові слова: метамодель, план експерименту, ЛП$_\tau$-послідовність, поверхня відгуку, нейронна мережа, сурогатна оптимізація.

Постановка проблеми. В задачах оптимізації складних технічних систем з метою можливості їх реальної реалізації виникає необхідність замінити точну, проте обчислювально складну, математичну модель на значно більш просту – сурогатну [1, 2]. Сурогатна модель є апроксимацією з необхідною точністю первинної математичної моделі і будується на основі значень її вихідних характеристик.

Побудова метамоделі в загальному випадку передбачає вирішення двох взаємопов'язаних завдань: планування обчислювального експерименту і власне побудови регресійної моделі. Вирішення першої забезпечує мінімальну кількість точок в багатовимірному просторі пошуку, в яких потрібно обчислення цільової функції з використанням первинних моделей та які гарантують одночасно точне відтворення поверхні відгуку. Метамодель, отримана в результаті вирішення другого завдання, робить можливим ефективний пошук екстремуму вихідної сурогатної оптимізаційної задачі.

Аналіз останніх досліджень і публікацій. При виборі плану апроксимації серед можливого розмаїття варіантів перевагу слід віддати генераторам точок, що заповнюють простір пошуку та в процесі реалізації яких використовуються ЛП$_\tau$-послідовності Соболя. Визначальними для прийняття цього рішення є наступні властивості таких послідовностей, відмічені в [3]: висока ймовірність потрапляння зондувальних точок послідовності в просторі пошуку в околі точок екстремумів і перегинів поверхні відгуку цільової функції, яка набагато перевищує ймовірність відомих на даний час багатофакторних регулярних планів; слабо корельовані головні ефекти й ефекти взаємодії факторів; розподіл в евклідовому просторі точок, які є вже першими членами послідовності, характеризується унікальною властивістю однорідності та може бути отриманий за кінцевий час роботи генератора, а не в границі, як наприклад, у рівномірних випадкових розподілах при рівній кількості точок порівнюваних послідовностей. Таким чином, ЛП$_\tau$-послідовності Соболя мають кращі на даний час властивості рівномірного розподілу точок в одиничному гіперкубі ніж будь-які інші відомі науці послідовності. Тому, застосування ЛП$_\tau$-послідовностей при плануванні багатофакторного експерименту для отримання регресійних моделей є перспективним також для вирішення задач апроксимації з використанням нейронних мереж.

Крім того, дослідниками встановлено, що краща кількість точок для рівномірного заповнення багатовимірного простору пошуку може бути визначена з виразу $N = 2^Z - 1$, де $z = 2,3,4,...$ тобто при $N = 1...7,15,31,63,127,255,...$ спостерігається слабка корельованість головних ефектів та ефектів взаємодії факторів [4-5]. Бажаною також є процедура виявлення чи корелюють між собою послідовності з метою подальшого вибору серед них ортогонально близьких одна щодо іншої, проте доведеним вважається факт слабкої корельованості з поміж згенерованої групи ЛП$_\tau$-послідовностей і тому застосуванням цієї процедури можна в ряді випадків нехтувати.

Апроксимацію експериментальних даних можна виконати використовуючи штучні нейронні мережі: багатошаровий персептрон, мережу на радіально-базисних функціях, ймовірнісні мережі, узагальнено-регресійні мережі. В якості апроксиматора в даному дослідженні використовується нейронна MLP-мережа на багатошаровому персептроні [6]. Відомо, що будь-яка обмежена нелінійна багатовимірна функція може бути як завгодно точно наближена нейронною мережею з кінцевим числом нейронів.

MLP-мережа, якою обмежувались в цих дослідженнях, має в своєму складі: вхідний шар – призначений для вводу значень вхідних змінних; один прихований шар (або проміжний), який має гладкі функції активації (identity, logistic, tanh, exponential, sine); вихідний шар, який в результаті налаштування ваг, визначає вихід мережі і також має функції активації.

При використанні мережі на вхідні елементи подаються значення вхідних змінних, потім послідовно задіюються нейрони проміжного та вихідного шарів. Кожний із них розраховує своє значення активації за допомогою функції активації, беручи зважену суму виходів елементів попереднього шару та віднімаючи із неї граничне

значення (зміщення). Після того, як вся мережа відпрацює, вихідні значення елементів вихідного шару приймаються за вихід всієї мережі в цілому. Елементи організовані в пошарову структуру з прямою передачею сигналу. Така мережа загалом може моделювати функцію практично будь-якої складності, причому кількість шарів і кількість елементів в кожному із прихованих шарів визначається складністю функції.

Завдання апроксимації MLP-мережею зводиться до оптимального вибору ваг вихідного шару для заданої функції активації, кількості прихованих функцій (нейронів) та функцій активації. З цією метою, мінімізується функціонал похибки, який представляє собою суму квадратів відхилень в точках заданого плану. Похибка для кожної мережі визначається шляхом прогону через мережу всіх наявних спостережень та порівняння реально отриманих вихідних значень із значеннями цільових функцій. Всі отримані різниці додаються в так звану «функцію похибок», значення якої і є похибкою мережі.

Виділення невирішених раніше частин загальної проблеми. При вирішенні задач оптимального синтезу різноманітних конструкцій в багатьох сферах науки та техніки є проблема високої вартості цільових функцій в сенсі великих витрат часових і обчислювальних ресурсів. В складних задачах оптимізації, в яких цільова функція задається алгоритмічно, як результат виконання чисельних розрахунків методами кінцевих елементів, граничних або просторових інтегральних рівнянь, реалізація навіть сучасних метаевристичних алгоритмів оптимізації, що характеризуються високою швидкістю збіжності, стає достатньо складною. В теорії оптимізації ця проблема вирішується або розпаралелюванням обчислень, що вимагає наявності спеціальних дороговартістних апаратних засобів і специфічних підходів до реалізації програмного забезпечення, або застосуванням технологій сурогатної оптимізації застосування якої передбачає побудову метамоделей.

Таким чином, **метою даних досліджень** є розробка обчислювальної технології побудови MLP-

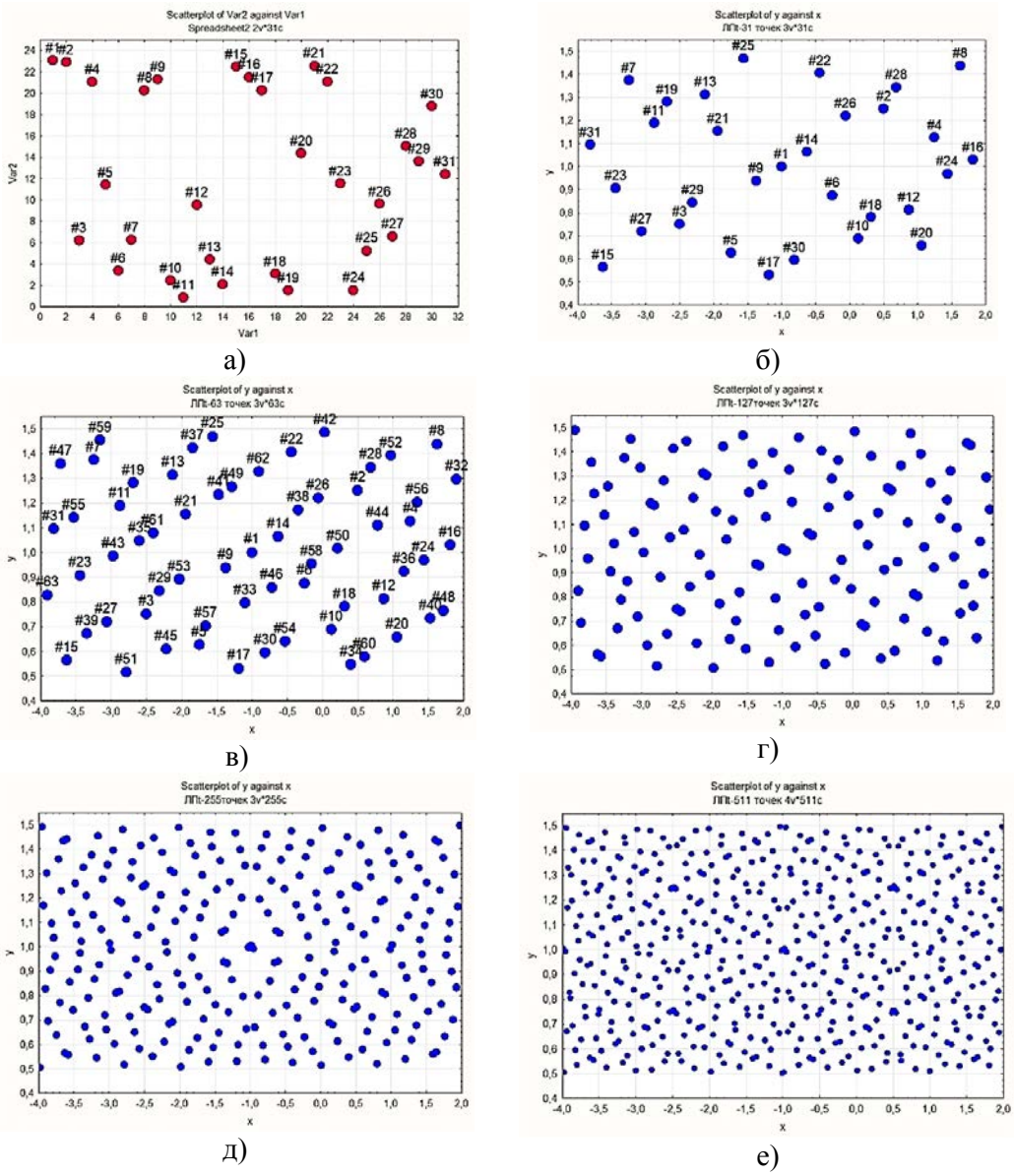


Рис. 1. Заповнення області пошуку точками зондування: а) рівномірна випадкова генерація; б)-е) генерація із застосуванням ЛП$_\tau$-послідовностей

метамоделей, що використовує сучасні досягнення в області теорії планування експерименту, інтелектуального аналізу даних і штучного інтелекту, та визначення експериментальним шляхом закономірностей, які дозволяють ефективно виконувати побудову метамоделей і їх подальше використання в задачах сурогатної оптимізації.

Виклад основного матеріалу. Продемонструємо запропоновану техніку побудови метамоделі на чисельних прикладах. Для наочності обмежимося функцією цілі, яка залежить від двох змінних і має вигляд $f(x,y) = (0.5 + 0.5 \cdot x)^4 \cdot y^4 \cdot e^{\left[2-(0.5+0.5\cdot x)^4 - y^4\right]}$. З її використанням розраховувалися значення функції в точках плану апроксимації N = 1...31,63,127,255 в діапазоні зміни змінних $x \in [-4;2]$, $y \in [0,5;1,5]$ (рис. 1).

Рис. 1 а демонструє результати застосування рівномірного випадкового розподілу, а рис. 1 б, – ЛП$_\tau$-послідовностей Соболя за умови однакової кількості точок в співставних послідовностях, рис. 1 в, г, д, е – для планів на основі ЛП$_\tau$-

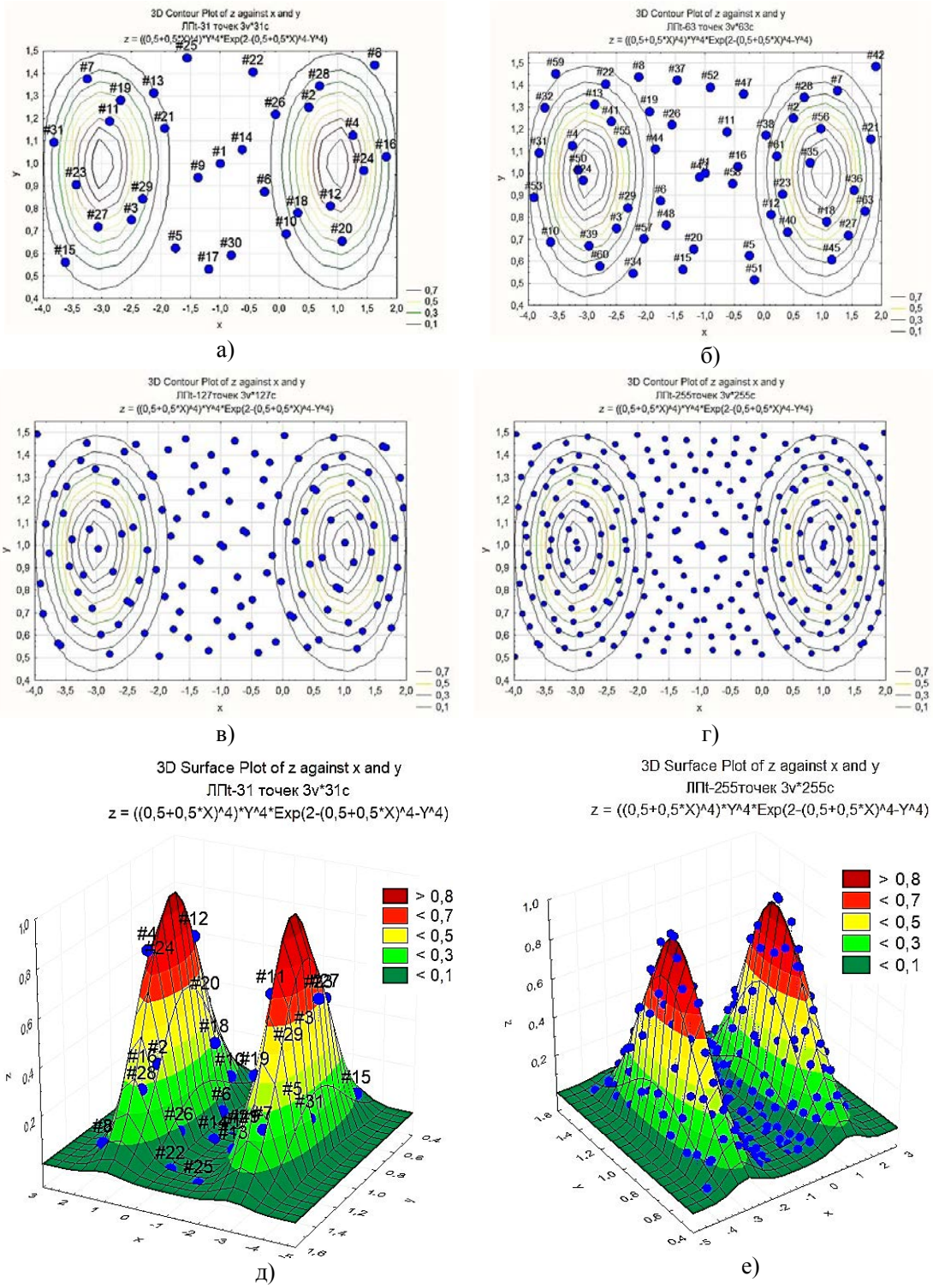


Рис. 2. а)-г) плани апроксимації N = 31, 63, 127, 255, нанесені на лінії рівня цільової функції; д), е) плани апроксимації N = 31, N = 255, нанесені на 3D поверхню цільової функції

послідовностей з різною кількістю точок. Порівняльний аналіз доводить достовірність наведених вище теоретичних положень. Особливо слід відзначити, що найбільші переваги використання $ЛП_т$-послідовностей спостерігаються при малій кількості зондувальних точок області.

У точках, згенерованого на першому етапі плану, виконуються обчислення з використанням первинної цільової функції, які в сукупності з координатами зондувальних точок в багатовимірному просторі, складають таблицю вихідних даних для виконання другого етапу – побудови метамоделі, а саме отримання в явному вигляді апроксимуючої функції поверхні відгуку первинної цільової функції (рис. 2).

На наступному етапі досліджень для моделювання використовувався проміжний шар із кількістю прихованих нейронів, який дорівнює півсумі числа вхідних та вихідних елементів. Для побудови мережі з m кількістю нейронів у прихованому шарі будь-якою стратегією (автоматичною, заданої користувачем, багатократних підвибірок) використовувалася вибірка з N точок, з яких 70% виконували функції навчальних, 15% – контрольних, 15% – тестових.

Для плану з N=31 створено майже 250 MLP-метамоделей в автоматичному режимі, з яких відібрано відповідно найкращі: MLP-2-20-1 з показниками коефіцієнта детермінації навчальної вибірки – 0,957, S.D.ratio – 0,57, в якій було застосовано алгоритм навчання BFGS, функція активації прихованого шару – tang, вихідного шару – exponential; MLP-2-23-1 з показниками коефіцієнта детермінації навчальної вибірки – 0,957, S.D.ratio – 0,254, де також було застосовано алгоритм навчання BFGS, функція

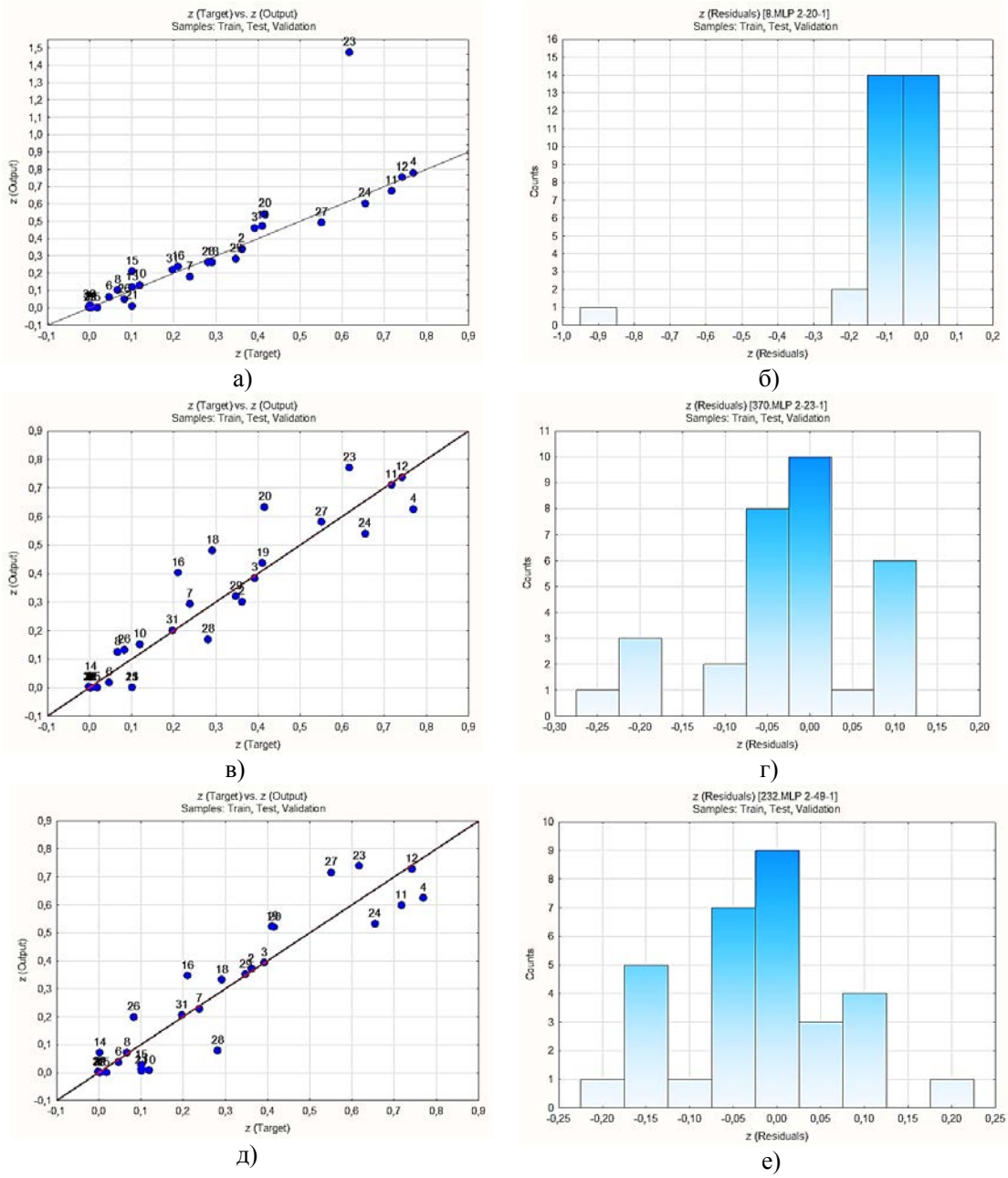
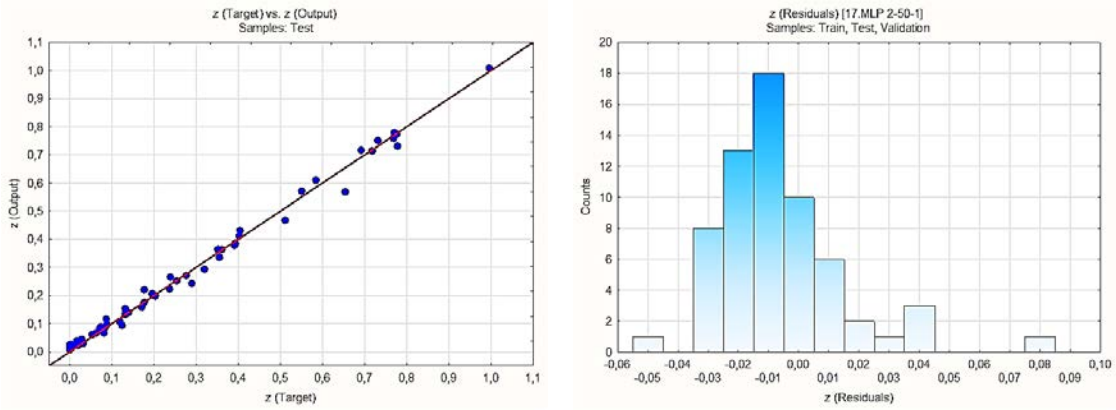


Рис. 3. MLP-нейронна мережа (вибірка N = 31): а), в), д) діаграми розсіювання значень цільової та вихідної функції для мереж MLP-2-20-1, MLP-2-23-1 та MLP-2-49-1 відповідно; б), г), е) гістограми залишків



а) б)

Рис. 4. MLP-нейронна мережа (вибірка N = 63): а) діаграма розсіювання значень цільової та вихідної функції для мережі MLP-2-50-1; б) гістограма залишків

активації прихованого шару – tang, вихідного шару – logistic; MLP-2-49-1 з показниками коефіцієнта детермінації навчальної вибірки – 0,92, S.D.ratio – 0,243, алгоритм навчання BFGS, функція активації прихованого шару – logistic, вихідного шару – logistic.

За результатами, що містяться на рис. 3, оцінювалася ефективність розробленої обчислювальної технології.

Гістограми (рис. 3 б, г, е) демонструють, що кожна модель має значні абсолютні похибки (залишки), які відповідно відображаються на діаграмах розсіювання значень цільової та вихідної функції нейронної мережі (рис. 3 а, в, д). Цілком зрозуміло, що така кількість точок плану експерименту і відповідно побудовані на їх основі MLP-метамоделі не відтворюють адекватно поведінку первинної цільової функції (рис. 2 а, д). Мережі, створені в режимі «задання користувачем» із вибором алгоритму навчання, із заданою кількістю епох та багатократних підвибірок не дали суттєвого покращення результату.

Для плану з N=63 з ЛП$_\tau$-послідовностями створено 60 MLP-метамоделей в автоматичному режимі, проте отримати принаймні одну метамодель з прийнятною похибкою не вдалося. Тому було змінено план експерименту, а саме застосовано інші слабко корельовані ЛП$_\tau$-послідовності з можливого набору для N=63 (рис. 1 в). Отримана найкраща мережа MLP-2-50-1 з показниками коефіцієнта детермінації навчальної вибірки – 0,997; S.D.ratio – 0,057, в якій було застосовано алгоритм навчання BFGS, функція активації прихованого шару – tang, вихідного шару – exponential. Рис. 4 ілюструє оцінку ефективності розробленої обчислювальної технології.

Гістограма (рис. 4 б) демонструє, що модель має незначні абсолютні похибки (залишки) у порівнянні із попередніми моделями (рис. 3), які відповідно відображаються на діаграмах розсіювання значень цільової та вихідної функції нейронної мережі (рис. 4 а). Із аналізу значень відносної похибки $\bar{\delta},\%$ в кожній точці плану видно, що максимальний вклад в похибку вносять точки в околі мінімуму цільової функції.

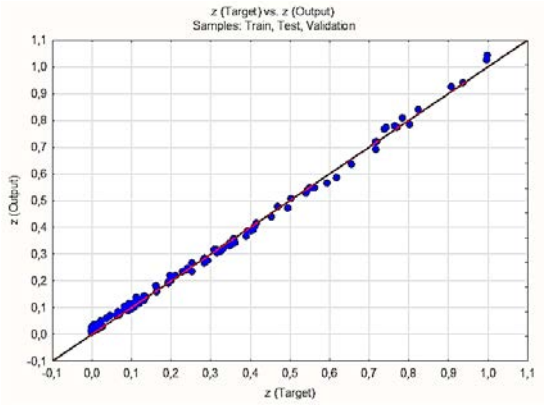
Для плану з N=127 з ЛП$_\tau$-послідовностями створено 50 MLP – метамоделей в автоматичному режимі з алгоритмом навчання BFGS. З яких відібрано відповідно найкращі: MLP-2-7-1 з показниками коефіцієнта детермінації навчальної вибірки – 0,998, S.D.ratio – 0,031, функція активації прихованого шару – tang, вихідного шару – exponential (рис. 5 а, б); MLP-2-9-1 з показниками коефіцієнта детермінації навчальної вибірки – 0,9999, S.D.ratio – 0,008, функція активації прихованого шару – logistic, вихідного шару – exponential (рис. 5 в, г); MLP-2-19-1 з коефіцієнтом детермінації навчальної вибірки – 0,99998, S.D.ratio – 0,0037, функція активації прихованого шару – tang, вихідного шару – exponential (рис. 5 д, е). Отримані результати відносної похибки $\bar{\delta},\%$ мало відрізняються від попередніх дослідів і збільшення кількості нейронів прихованого шару не впливає суттєво на значення цієї похибки. Також залишається тенденція максимального вкладу у відносну похибку точок в околі мінімуму цільової функції (рис. 5 ж, з).

Для плану з N=255 з ЛП$_\tau$-послідовностями створено MLP-метамоделі в автоматичному режимі з алгоритмом навчання BFGS. Отримано найкращі: MLP-2-8-1 з коефіцієнтом детермінації навчальної вибірки – 0,9999, S.D.ratio – 0,0063, функція активації прихованого шару – logistic, вихідного шару – exponential (рис. 6 а, б); MLP-2-13-1 з коефіцієнтом детермінації навчальної вибірки – 0,9999, S.D.ratio – 0,0067, функція активації прихованого шару – tang, вихідного шару – exponential (рис. 6 в, г); MLP-2-15-1 з коефіцієнтом детермінації навчальної вибірки – 0,9999, S.D.ratio – 0,006, функція активації прихованого шару – logistic, вихідного шару – exponential (рис. 6 д, е). Як видно із гістограм розподілу залишків (рис. 6 б, г, е), отримані метамоделі мають досить незначні абсолютні похибки.

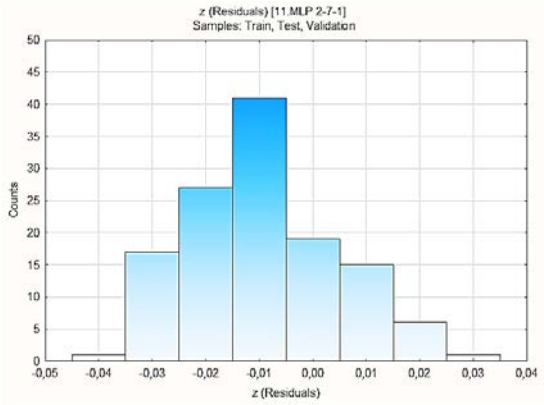
Висновки і пропозиції. Таким чином, при створенні MLP-метамоделей на основі планів із застосуванням ЛП$_\tau$-послідовностей з N = 1...31,63,127,255 отримано наступні результати:

· при створенні мережі на результат суттєво не впливає вибрана стратегія побудови – автоматична, користувача чи багатократних підвибірок;

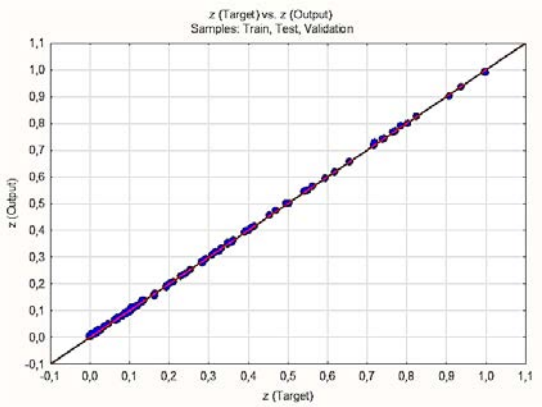
· для апроксимації поверхні відгуку складної багатоекстремальної первинної цільової функції недоцільно застосовувати плани з малою кількіс-



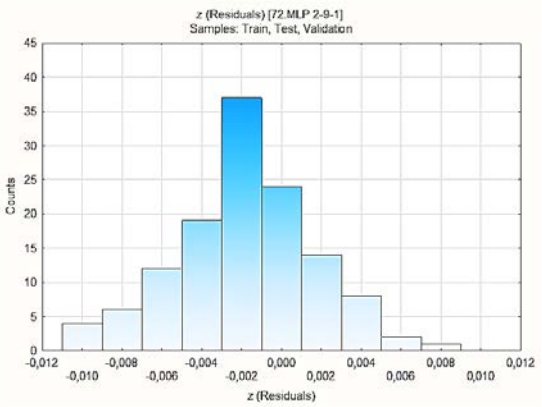
а)



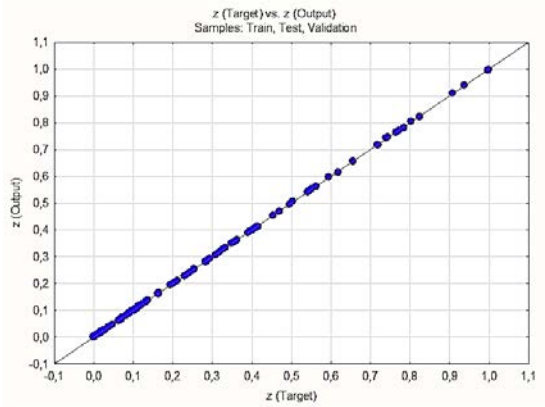
б)



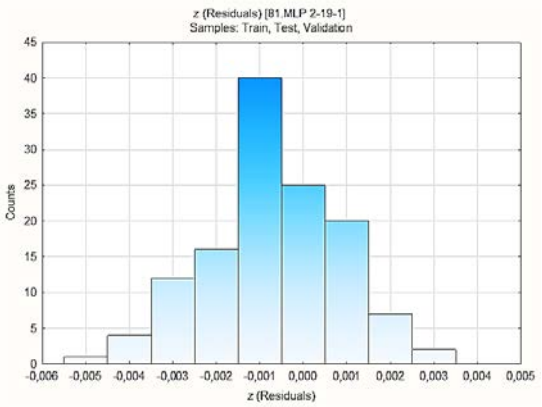
в)



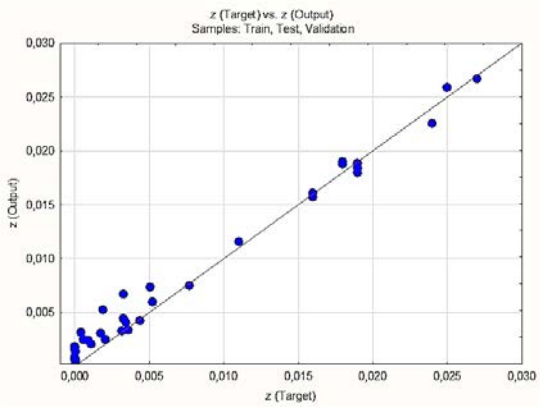
г)



д)



е)



ж)

Predictions spreadsheet for z (ЛПt-127точек)
Samples: Train, Test, Validation

Case nan	Sample	x Input	y Input	z Target	z - Output 81. MLP 2-19-1	z - Abs. Res. 81. MLP 2-19-1	NewVar =V6/Abs(V
7	Train	-3,25000	1,375000	0,239000	0,237431	0,001569	0,00656489
8	Test	1,62500	1,438000	0,067000	0,065725	0,001275	0,01902806
9	Validation	-1,37500	0,938000	0,003254	0,006647	0,003393	1,0426114
10	Test	0,12500	0,688000	0,120000	0,120068	0,000068	0,00056551
11	Test	-2,87500	1,188000	0,718000	0,717712	0,000288	0,00040147
12	Validation	0,87500	0,813000	0,743000	0,745688	0,002688	0,00361811
13	Validation	-2,12500	1,313000	0,102000	0,103222	0,001222	0,01197872
14	Train	-0,62500	1,063000	0,003250	0,004407	0,001157	0,35598645
15	Train	-3,62500	0,563000	0,102000	0,099943	0,002057	0,02016389
16	Train	1,81300	1,031000	0,211000	0,209627	0,001373	0,00650845
17	Validation	-1,18800	0,531000	0,000042	0,000568	0,000526	12,5338048
18	Train	0,31300	0,781000	0,292000	0,292952	0,000952	0,00325891
19	Validation	-2,68800	1,281000	0,411000	0,413775	0,002775	0,00675251
20	Train	1,06300	0,656000	0,416000	0,413023	0,002977	0,00715658
21	Validation	-1,93800	1,156000	0,102000	0,100299	0,001701	0,01667746
22	Train	-0,43800	1,406000	0,003598	0,003321	0,000277	0,07709916
23	Train	-3,43800	0,906000	0,617000	0,615428	0,001572	0,00254733
24	Train	1,43800	0,969000	0,655000	0,654582	0,000418	0,00063878
25	Train	-1,56300	1,469000	0,002037	0,002479	0,000442	0,21699937
26	Train	-0,06300	1,219000	0,083000	0,082628	0,000372	0,00448601
27	Test	-3,06300	0,719000	0,551000	0,552029	0,001029	0,00186792
28	Validation	0,68800	1,344000	0,282000	0,281085	0,000915	0,00324514
29	Train	-2,31300	0,844000	0,348000	0,349601	0,001601	0,00460077

з)

Рис. 5. MLP-нейрона мережа (вибірка N=127): а) діаграма розсіювання значень цільової та вихідної функції для мережі MLP-2-7-1; б) гістограма залишків; в), д) діаграми розсіювання значень цільової та вихідної функції для мереж MLP-2-9-1, MLP-2-19-1; г), е) гістограми залишків; ж) діаграма розсіювання функцій в околі мінімуму цільової функції; з) розрахунок похибок апроксимації

тю точок, оскільки отримані метамоделі мають значний розкид абсолютної похибки;

· із збільшенням кількості точок плану експерименту та кількості прихованих нейронів до певного рівня параметри метамоделей покращуються; використання різних слабко корельованих $ЛП_τ$-послідовностей з їх повного набору суттєво не впливають на показники якості кінцевої метамоделі;

· при оцінюванні якості створених метамоделей, окрім коефіцієнту детермінації навчальної вибірки, S.D.ratio необхідно також звертати увагу на продуктивність метамоделей для контрольної і тестової вибірок та одночасно аналізувати гістограми залишків і діаграми розсіювання. Також, для кожної створеної моделі, необхідно окремо визначати значення відносної похибки для всіх точок плану експерименту та оцінювати їх вплив в сумарну похибку.

Результати чисельних експериментів свідчать про можливість використання запропонованої обчислювальної технології побудови MLP-метамоделей для апроксимації первинних функцій цілі з достатньо складною поверхнею відгуку.

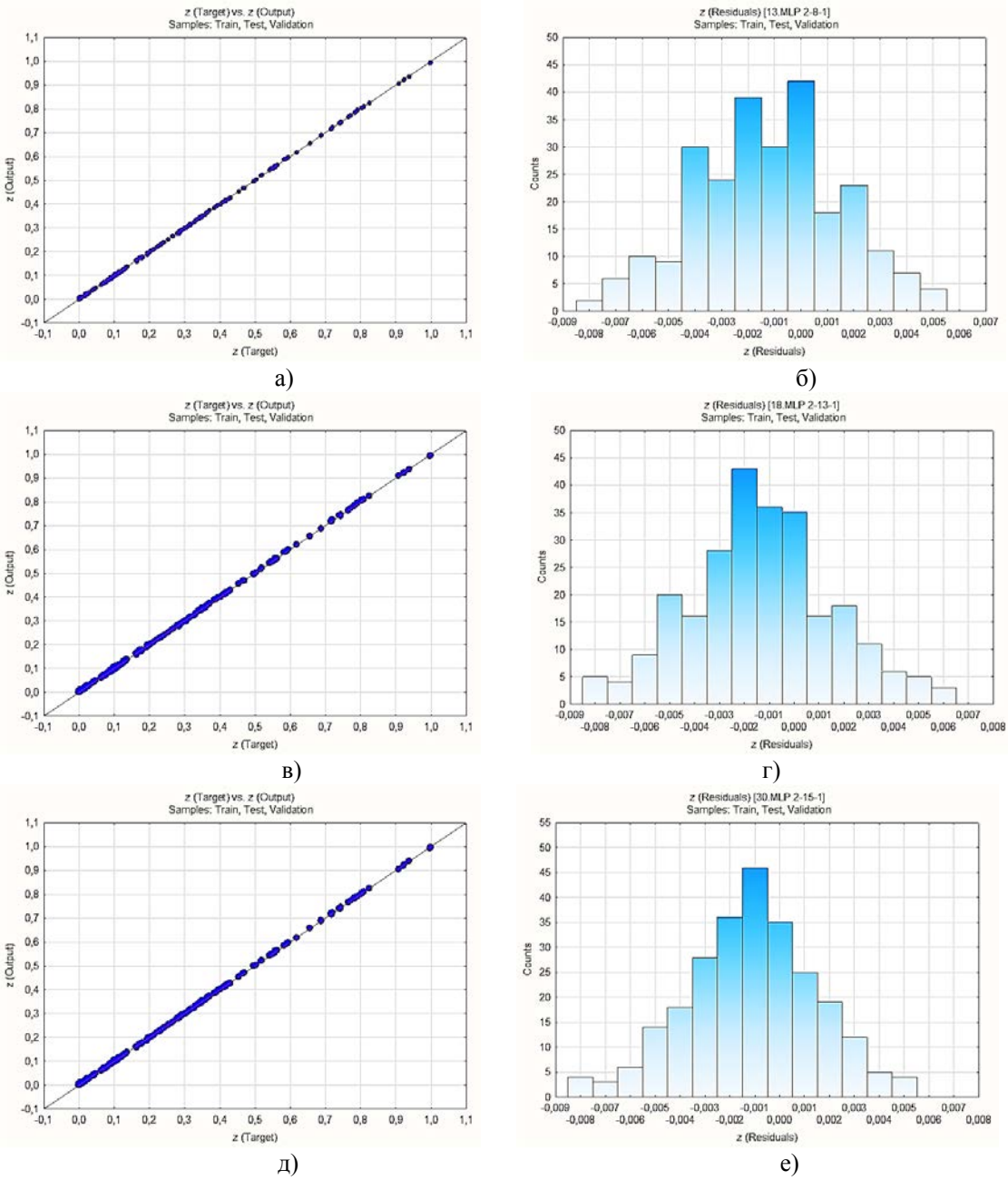


Рис. 6. MLP-нейрона мережа (вибірка N=255): а), в) діаграми розсіювання значень цільової та вихідної функції для мереж MLP-2-8-1, MLP-2-13-1, MLP-2-15-1; б), г) гістограми залишків; д) діаграма розсіювання значень цільової та вихідної функції для мережі MLP-2-15-1; е) гістограма залишків

Список літератури:

1. Gal'chenko V.Ya. Pareto-Optimal Parametric Synthesis of Axisymmetric Magnetic Systems with Allowance for Nonlinear Properties of the Ferromagnet / V.Ya. Gal'chenko, A.N. Yakimov, D.L. Ostapushchenko // Technical Physics. – 2012. – Vol. 57. – No. 7. – Pp. 893-899.
2. Galchenko V.Ya. A Turmitobionic Method for the Solution of Magnetic Defectometry Problems in Structural-Parametric Optimization Formulation / V.Ya. Galchenko, A.N. Yakimov // Russian Journal of Nondestructive Testing. – 2014 – Vol. 50. – No. 2. – Pp. 59-71.
3. Соболь И.М. Выбор оптимальных параметров в задачах со многими критериями / Соболь И.М., Статников Р.Б. [2-е изд., перераб. и доп.]. – Москва: Дрофа, 2006. – 175 с.
4. Радченко С.Г. Планы экспериментов для получения моделей высокой точности / С.Г. Радченко, О.В. Козырь // Математичні машини і системи. – 2014. – № 2. – С. 117-127.
5. Радченко С.Г. Применение ЛП$_\tau$ равномерно распределенных последовательностей для решения прикладних задач моделирования / С.Г. Радченко, О.В. Козырь // Математичні машини і системи. – 2014. – № 1. – С. 151-158.
6. Хайкин Саймон. Нейронные сети: полный курс, 2-е изд.: Пер. с англ. – Москва: Издательский дом «Вильямс», 2006. – 1104 с.
7. Гальченко В.Я. Побудова RBF-метамоделей в задачах сурогатної оптимізації / Гальченко В.Я., Трембовецька Р.В.: Збірник тез доповідей II Всеукраїнської науково-технічної конференції «Теоретико-практичні проблеми математичних методів і комп'ютерно-орієнтованих технологій в освіті та науці»; 2018 Бер 28; Київ, Київ: Київ.ун-т ім. Б. Грінченка. (Друк).

Трембовецкая Р.В., Гальченко В.Я., Тычков В.В.
Черкасский государственный технологический университет

ПРИМЕНЕНИЕ MLP-МЕТАМОДЕЛЕЙ В ЗАДАЧАХ СУРОГАТНОЙ ОПТИМИЗАЦИИ

Аннотация
Разработана вычислительная технология построения метамоделей, использующая современные достижения в области теории планирования эксперимента, интеллектуального анализа данных и искусственного интеллекта и определенные экспериментальным путем закономерности, которые позволяют эффективно выполнять построение метамоделей. Построение метамоделей выполнена на многочисленных примерах с функцией цели, которая зависит от двух переменных. В качестве плана эксперимента использован генератор точек, заполняющий пространство поиска и в процессе реализации которого используются ЛП$_\tau$-последовательности Соболя. Установлено, что увеличением количества точек плана эксперимента и количества скрытых нейронов до определенного уровня, параметры метамоделей улучшаются; использование различных слабо коррелированных ЛП$_\tau$-последовательностей из их полного набора существенно не влияет на показатели качества конечной метамодели. Результаты многочисленных экспериментов свидетельствуют о возможности использования предложенной вычислительной технологии построения MLP-метамоделей для аппроксимации первичных функций цели с достаточно сложной поверхностью отклика.
Ключевые слова: метамодель, план эксперимента, ЛП$_\tau$-последовательность, поверхность отклика, нейронная сеть, суррогатная оптимизация.

Trembovetska R.V., Halchenko V.Ya., Tychkov V.V.
Cherkassy State Technological University

THE MLP-METAMODELS APPLICATION IN THE SURROGATE OPTIMIZATION TASKS

Summary
A computational technology for constructing metamodels was developed using modern achievements in the field of the experimental planning theory, intellectual data analysis and artificial intelligence, and experimentally determined patterns that allow metamodels efficient construction. The metamodels construction is performed on numerous examples with a goal function, which depends on two variables. As an experiment plan, a point generator is used, which fills the search space and during the implementation of which the Sobol LP$_\tau$-sequences are used. It has been established that by increasing the points number in the experiment plan and the hidden neurons number to a certain level, the metamodel parameters are improving; the use of various weakly correlated LP$_\tau$-sequences from their complete set does not significantly affect the quality indicators of the final metamodel. The results of numerous experiments testify to the possibility of using the proposed computational technology for constructing MLP-metamodels for approximating the primary goal functions with a rather complex response surface.
Keywords: metamodel, design of experiment, LP$_\tau$-sequence, response surface, neural network, surrogate optimization.