**Авторска справка за научните приноси в списъка с публикации, представени като хабилитационен труд за участие в конкурс за доцент в област на висше образование 4. Природни науки, математика и информатика, професионално на-правление 4.6. Информатика и компютърни науки, научна специалност „Информатика“ (Моделиране на сложни системи с голяма размерност)**

**гл. ас. д-р Венелин Тодоров**

От списъка с публикации за участие в конкурса, 8 публикации са представени като хабилитационен труд. Представени са 5 публикации с импакт ранг (2 в престижната поредица American Institute of Physics и 3 в Springer поредицата Studies in Computational Inteligence), една в базата IEEE Xplore и видима в Scopus и още 2 статии, които не попадат в световноизвестните бази данни. Тематичното направление, в което попадат тези публикации е „Моделиране на сложни системи с голяма размерност “, а научните приноси в тях могат да се представят във вид на монография със следното заглавие

# Съвременни стохастичи подходи за моделиране на сложни системи с голяма размерност в екологията

## 1. Увод

Опазването на околната среда се нарежда сред водещите приоритети в цял свят. Необходимо е да се проведат задълбочени научни изследвания и да се анализират прецизно основните физични и химични процеси, протичащи при преноса на

замърсители, както и съпровождащите ги трансформации. Ефективното изпълнение на тази сложна задача изисква съвместни действия между специалисти в областта на моделирането на пренос на замърсители във въздушна среда, числения анализ и научните пресмятания. Има много предизвикателства в тази област на изследвания, тъй като това е болезнена тема за обществото, а също така и важен въпрос за системата на здравеопазването.

Съществен дял от стохастичните методи за големи задачи се пада на Монте Карло методите [10]. Монте Карло методите са методи за приближено пресмятане на решението на задачи от изчислителната математика чрез използване на случайни процеси, като параметрите на съответния процес съвпада с решението на задачата. Методът може да гарантира, че грешката при приближеното пресмятане на неизвестната величина е по-малка от зададена стойност с определена вероятност. Методите Моте Карло често се използват за решаването на физични и математически проблеми и са много полезни при решаването на големи задачи, когато е невъзможно или непрактично използването на други математически методи. Монте Карло методите са особено полезни за симулиране на явления със значителна несигурност във входната информация и системи с голям брой на -степени на свобода.

Монте Карло методите дават статистически оценки за функционал от решението при обработка на извадка на определена случайна величина, чието математическо очакване съвпада с търсения функционал. Монте Карло методите предлагат простота на конструкциите и често се използват за симулация на процеси, чието поведение може да се интерпретира само в статистически смисъл, каквито са входните параметри в математическия модел описващ далечен пренос на замърсители във въздуха, в който предстои да се приложи анализ на чувствителността.

Ефективните Монте Карло методи имат голямо значение за анализа на чувствителността на големи компютърни модели [11]. Настоящето изследване се фокусира върху много голям модел, описващ преноса на замърсители във въздуха, тъй като високите нива на замърсяване могат да доведат до разрушаване на еко-системите и да застрашат растенията, животните и хората. Тези задачи могат

успешно да се решат чрез разработването и използването на надежден апарат за анализа на чувствителността за изучаване на различни явления, свързани със замърсяването на околната среда. Обикновено изчислителната задача, възникващи при решаване на голям математически модел, се характеризира с много голяма размерност, като трудности от изчислителен характер се появяват, дори когато пресмятанията се извършват със съвременни мощних компютърни системи. За това е изключително важно модела да бъде опростен, ако е възможно, като са постигнати необходимото ниво на точност и надеждност на резултатите, получени от модела. Една възможност за постигането на тази цел е предварително да се проведе анализа на чувствителността [21]. По дефиниция анализа на чувствителността [22] е изследване как колебанието в изходните стойности на даден сложен модел или система може да се разпредели между различните източници на изменение сред входните параметри [23]. Анализа на чувствителността е съвременна перспективна техника за изследване на големи системи като екологичните системи. Основната идея на анализа на чувствителността е да оценява или прогнозира (чрез компютърни симулации върху големи математически модели) мярката за чувствителността на изходните резултати на модела към пертурбациите на някои входни параметри. Математическата формулировка на задачата за анализа на чувствителността се представя чрез множество от интеграли с висока размерност.

Унифицираният датски Ойлеров модел (UNI-DEM) [28] симулира процеси, свързани с преноса на замърсители във въздуха, и е разработен от Националния датски институт за изследване на околната среда с активното участие на учени от БАН. Пространствената област на модела включва цяла Европа, Средиземноморието, както и части от Азия и Африка. Моделът предоставя възможност за изследване във времето на концентрациите на основните типове замърсители (серни и азотни съединения, амоняк, амониеви йони, азот, свободни радикали, въглеводороди), което има значение за екологичната безопасност, селското стопанство и здравеопазването. Избраният математически модел описва по адекватен начин протичащите физични (адвекция, дифузия, депозиция, вредните емисии) и химични

процеси, които съпровождат преноса на замърсители във въздуха. UNI-DEM е мощно средство за моделиране на въздушното замърсяване в глобален мащаб, разработван и използван повече от две десетилетия за пресмятане на концентрациите на голям брой химически вещества (включвайки всички опасни замърсители на въздуха) над голяма географска област (4800 x 4800 km). В математическата формулировка на модела са отчетени всички основни физични процеси (адвекция, дифузия, депозиция, емисии и химични реакции). От друга страна, чрез този модел могат да се изследват основните видове химични замърсители (сулфидни, азотни, амонячни, озон, свободни радикали и въглеводороди). Химичните реакции играят изключително важна роля в математическия модел. Уравненията в модела се свързват в система именно чрез химичните реакции. От друга страна, нелинейността и свойството „твърдост“ на системата от уравнения се определя основно от присъствието на химичните процеси. Схемата, която се използва [29], е CBM IV (Carbon Bond Mechanism) и включва 35 замърсители и 116 химични реакции (69 зависят от времето и 47 не зависят от времето). Основната мотивация за избора на този модел е, че той е един от моделите, в които химичните процеси са зададени оптимално адекватно.

## 2. Научни приноси

**Научните приноси в публикациите, заместващи хабилитационен труд, могат да се групират в следните направления:**

**• Разработване на нови стохастични методи (точкови множества тип решетка с различни генериращи вектори базирани на оптимален генериращ вектор, вектор от полиномиална трансформираща функция, вектор базиран на обощената редица на Фибоначи; модифициран адаптивен алгоритъм и модифицирана редица на ван дер Корпут) за анализ на чувствителността на сложна екологична система с голяма размерност [B1-B4];**

**• Прилагане и сравнение за първи път на съществуващи стохастични методи (квазислучайни редици на Фор, Холтън и ван дер Корпут и извадка латински хиперкуб) за анализ на чувствителността на еколгоичния модел [B7, B8];**

**• Разработване на нов метод Монте Карло за големи линейни системи, сравнения с досега съществуващите стохастични подходи, и приложение за матрица с важно значение при опазването на околната среда [B5, B6].**

## 2.1 Разработване на нови стохастични методи за анализ на чувствителността на сложна екологична система с голяма размерност

Основната цел на изследванията, чиито резултати са описани по-долу, е разработване и изследване на нов (количествен) механизъм за анализ на чувствителността на сложна система с голяма размрност, описваща далечен пренос на замърсители във въздуха, като се използва описаният в увода математически модел UNI-DEM. Най-широко използваните способи за анализа на чувствителността са методите, основани на дисперсията, тъй като те предоставят възможност не само за качествена, но и количествена класификация на входните параметри на даден математически модел в зависимост от степента им на влияние върху изходния резултат. Един от приложените подходи е методът на Собол [11,24], чиято идея се състои в пресмятането на индикатори за чувствителността, представляващи частно на съответните частични и пълни дисперсии. От математическа гледна точка, пресмятането на числените индикатори на чувствителността (индекси на чувствителност) се свежда до приближено пресмятане на многомерни интеграли [11]. С цел надеждна интерпретация на получените резултати за анализ на чувствителността на математическите модели се налага разработване и прилагане на ефективни стохастични алгоритми за многомерно числено интегриране, съобразено с изчислителните особености на подинтегралните функции, описващи разглежданите математически модели, и характеристиките на избраните подходи за анализ на чувствителността.

Предполага се [22], че математическият модел може да се представи чрез моделна функция $f_0$, която в случая описва разпределението на концентрацията на замърсителите в модела. Пълната дисперсия D на изходния параметър на модела се разлага на частичните дисперсии по аналогичен начин на разлагането на моделната функция [22]. Пълният индекс на чувствителността $S_i^{tot}$ представлява мярка за цялостното влияние (пълния ефект) на даден входен параметър върху колебанията в изходния резултат и се дефинират чрез дисперсиите на условните очаквания. $S_i$ се нарича основен ефект (индекс на чувствителността от първи ред) и $S_{ij}$ е индекс на чувствителността от 2-ри ред (съответно двустранни взаимодействия между съответните параметри).

В публикациите са описани резултатите от проведения анализ на чувствителността на концентрациите на основни замърсители във въздуха (озон, амоняк, амониев сулфат и амониев нитрат), симулирани от модел за далечен пренос на замърсители във въздуха UNI-DEM, спрямо нивата на четири групи от вредни емисии в резултат от човешка дейност – азотен оксид и диоксид, серен диоксид, амоняк, антропогенни въглеводороди, в три европейски градове с различно географско разположение. Също така в публикациите са описани резултатите от численото изследване на чувствителността на концентрацията на важен замърсител като озона спрямо измененията на скоростните константи на шест протичащи химични реакции, описани в математическия модел UNI-DEM, с различни подходи за анализ на чувствителността и алгоритми за числено интегриране. От предварителни изследвания е установено, че само три химични реакции (две - зависещи от времето и една – независеща) оказват по-съществено влияние върху концентрациите на озон. Тези данни са използвани за настоящото изследване. В първия случай на вредни емисии от четири групи замърсители моделната функция $f_0$ се представя от 4-мерен интеграл, а пълните индекси на чувствителността и индексите на чувствителност от първи и втори ред от 8-мерни интеграли. Във втория случай на шест химични реакции, моделната функция $f_0$ се представя от 6-мерен интеграл, а пълните индекси на чувствителността и

индексите на чувствителност от първи и втори ред от 12-мерни интеграли. В първия случай в числените експерименти е изследвана концентрацията на амоняк в Милано, а във втория случай – концентрацията на озон в Генуа.

В [**B1**] за първи път за анализ на чувствителността е разработено и приложено точково множество тип решетка с генериращ вектор обобщената редица на Фибоначи.

По дефиниция когато точките от едно множество са възлите на интеграционната решетка, попадащи в областта на интегриране, получаваме метод за апроксимиране на интеграла, наречен теоретико-числов метод за интегриране на възлите на интеграционната решетка (lattice rule). Сходимостта на решетката зависи от избора на генериращия вектор z. Векторът z, за който неравенството на Бахвалов [9] е изпълнено се нарича добър или оптимален генериращ вектор, а неговите компоненти- оптимални коефициенти в смисъла на Коробов [18]. Известни са редица теореми, доказващи съществуването на оптимални генериращи вектори. Трудността е в констуирането на тези оптимални вектори, особено в задачи от голяма размерност. Основната идея за получаването на оптимален генериращ вектор z е свързана с търсенето на минимума на най-лошата грешка, получена при интегрирането на тестови функции, принадлежащи на предварително дефиниран клас. За всички функции от даден клас се прави оценка на грешката от численото интегриране. Намира се "най-лошата" функция за класа, в смисъл на функцията, за която грешката от численото интегриране е максимална. Тогава под добър генериращ вектор се разбира векторът, минимизиращ грешката получена при интегрирането на най-лошата функция в класа. Следователно минимизирането на максимума на грешката от численото интегриране води до намиране на оптимален генериращ вектор z. Търсенето на оптимални генериращи вектори при висока размерност има висока изчислителна трудност. Изборът на добър генериращ вектор, който до малки грешки при интегриране, не е тривиален. Използват се сложни методи от теория на числата, основани на различни критерии като индекс на Zaremba и грешка на най-лошата функция, описани по горе. Установено е, че за добър генериращ вектор е необходимо индекса на Zaremba да бъде максимален

[27]. В двумерния случай оптимална конструкция съществува. Разработения тук метод за прозиволна размерност използва за генериращ вектор конструкцията, предложен от Hua и Wang през 1981 г. [14]. Предимството на решетката, базирана на обобщената редица на Фибоначи, е нейната бързина и високата точност при многомерни интеграли от гладки подинтегрални функции със сравнително ниска размерност.

В същата статия е приложена за първи път ,за анализ на чувствителността, и извадка латински хипекуб (LHS). В общия случай, при реализациите на равномерно разпределена случайна величина може да се наблюдава струпване в определени подобласти и това да доведе до повишаване на дисперсията, тъй като реализациите не са толкова добре равномерно разпределени в областта, колкото се очаква. Затова изследванията за намаляване на дисперсията са насочени към контролирано разпределяне на реализациите в съответната област. Конструирането на метода на ,,слоистата`` извадка (,,stratified sampling``) е резултат от тези усилия. Съществен недостатък при прилагане на този подход, е разделянето на подобласти. Ако една s-мерна област се разбие на равномерни подобласти чрез разделяне на две по всяко направление, се получават $2^s$ подобласти. В случая на големи s броят на подобластите рязко нараства, което ограничава силно избора на размерност на случайната извадка. Този проблем се преодолявам от извадката латински хиперкуб, използвана за първи път в [20]. Извадката "латински хиперкуб" е специален вид "слоиста извадка". В двумерния случай квадратна мрежа е "латински" квадрат, тогава и само тогава, когато във всеки ред и стълб на квадрат попада само една реализация на случайната величина. Тази конфигурация съответства на шахматна дъска, в която топовете са разположени така, че нито един не застрашава друг. "Латинки" хиперкуб е обобщение на тази концепция за произволна размерност. Тази техника изисква точно по една случайна точка във всяка подобласт, това свойство е едно от основните предимства на извадката "латиски" хиперкуб.

Редиците на Собол (1967) [25] са може би най-популярните редици в квази-Монте Карло методите. В редица сравнения на качествата на равномерно разпределените редици те се оказват най-добри или измежду най-добрите. Затова в тази и някои от

следващите публикации те са широко използвани и е направено сравнение с останалите стохастични подходи, повечето от които се прилагат за първи път към анализа на чувствителността на модела.
В изследването са приложени и три оптимални за определени класове от функции алгоритми Монте Карло за многомерно числено интегриране, използващи "изместени" квазислучайни редици на Собол. Идеята на алгоритмите е оригиналните квазислучайни точки на Собол да бъдат „изместени" на достатъчно малко разстояние (радиус), така че новата точка (която е псевдослучайна) да лежи в същия елементарен подинтервал, както и оригиналната квазислучайна точка, за да се запази равномерната разпределеност на точките. „Изместването" се реализира като псевдослучайната точка е избрана (равномерно разпределена) върху сферата с център съответната квазислучайна точка от редицата на Собол. Направено е сравнение с резултатите, получени с вече описаните Монте Карло алгоритми, използващи специално избрани точкови множества - извадка "латински хиперкуб" и обобщена редица на Фибоначи, които са приложени за първи път за екологичния модел. Алгоритмите са успешно приложени за приближеното пресмятане на глобалните индекси [24] на чувствителността на Собол за конкретната задача.

В [**B2**] за анализа на чувствителността са разработени точкови множества тип решетки с различни генериращи вектори. Разработени са 3 решетки с 3 различни генериращи вектора. Първата решетка използва генериращ вектор обобщената редица на Фибоначи, описан по-горе. Втората решетка използва многомерен генериращ вектор получен по метода на бързата конструкция, разработен в [17]. За този вектор е доказано в [17], че е оптимален гененриращ вектор за разглеждания клас задачи. Третата решетка използва генериращ вектор от полиномиална функция с биективни свойства. Направено е сравнение между редицата на Холтън и квази-Монте Карло методи, базирани на точкови множеста от тип решетка с различни генериращи вектори за пресмятане на пълните индекси на чувствителността. Това е един от първите пъти, в които редицата на Холтън е приложена за сложната система с голяма размерност, заедно с още две публикации

от Авторската справка [**Г3 и Г4**]. Редицата на Холтън е една от най-популярните редици с нисък дискрепанс. Тя е въведена за първи път през 1960 г. [16] и е пример за квазислучайна последователност от числа. Редиците на Холтън обобщават едномерните последователности на ван дер Корпут [26]. Редицата на Холтън е конструирана съгласно метод, който използва като основи относително прости (coprime) числа. Според очакването за индексите на чувствитеност, точковото множество, базирано на обобщената редица на Фибоначи от съответната размерност дава най-добри резултати за 4 и 6-мерните интеграли. За 8 и 12-мерните интеграли най-добри резултати дава стохастичния метод с оптимален генериращ вектор получен по метода на бързата конструкция, който има оптимален порядък на сходимост за разглеждания клас функции. Третия стохастичен метод, основан на специален генериращ вектор, базиран на трансформираща полиномиална функция, се оказва по неточен от редицата на Холтън за високи размерности на подинтегралната функция и се влияе от гладкостта на функцията.

В [**В3**] е направено приближено пресмятане на пълните индекси на чувствителността посредством модифициран адаптивен алгоритъм. Известно е, че адаптивните алгоритми са особено надеждни и ефективни при интегрирането на подинтегрални функции с изчислителни особености, като съществуват различни адаптивни алгоритми Монте Карло в зависимост от стратегията за адаптиране. Разработена е модификация на адаптивния алгоритъм, използван в [12,17]. Съществуват различни подходи при конструирането на адаптивни алгоритми Монте Карло. Реализираният тук адаптивен алгоритъм не използва никаква предварителна информация за гладкостта на подинтегралната функция, но използва апостериорна информация за дисперсията. Основната идея е концентриране на случайни точки в подобластите, в които дисперсията е голяма (по отношение на предварително зададена точност), т.е. подходът се основава на рекурсивно разделяне на областта, използвайки апостериорна информация за грешката при текущото разделяне. Алгоритъмът започва с разделяне на интервалите по всички направления на M подинтервала, като M е зададено като входящ параметър. При различни стойности на параметъра M се получава

модификация на основния алгоритъм. Във всяка подобласт се пресмята стойността на съответния интеграл и дисперсията. След това получената дисперсия се сравнява с предварително зададена стойност. Получената информация се използва за следващото сгъстяване, тъй като подобластта с най-голямо стандартно отклонение се разделя на $2^s$ нови подобласти. Алгоритъмът спира, когато стандартното отклонение във всички, получени след делението подобласти, е достигната предварително зададената точност (или когато е достигнат предварително зададен максимален брой на подобластите, в които не е изпълнен стоп-критерият, или на нивата на разделяне). Така се получава едно приближение на интеграла с подход Монте Карло. Направени са сравнения с точковото множество тип решетка, базирано на обобщената редица на Фибоначи и извадката латински хиперкуб. За ниски размерности най- ефективно е точковото множество с генериращ вектор обобщената редица на Фибоначи, а за високи размерности адаптивния подход показва предимство пред решетката на Фибоначи с 3-4 порядъка по-добра относителна грешка и извадката латински хиперкуб с 1-2 порядъка по-добра относителна грешка за индексите на чувствителност.

В [**B4**] за първи път за анализ на чувствителноста на модела е приложена редицата на ван дер Корпут. За първи път е направено сравнение на тази редица с точково множество тип решетка, базирано на обобщената редица на Фибоначи и извадка латински хиперкуб. За първ път за анализ на чувствителността на модела е разработена модифицирана редица на ван дер Корпут при основа 3. Методите са използвани в проучването за анализ на чувствителността на резултатите от изхода на Датския Ойлеров модел за някои замърсители на въздуха по отношение на нивата на вредни емисии и някои скоростни константи на химични реакции. Редицата на ван дер Корпут е пример за най-простата едномерна редица с нисък дискрепанс, описана за пръв път през 1935 [26] от холандския математик ван дер Корпут. Редицата на ван дер Корпут често се използва за генериране на последователност от точки, които имат по-добра сходимост отколкото псевдослучайните точки. Последователността на ван дер Корпут генерира

последователност от точки в [0,1] което никога не се повтаря. Елементите на последователността на ван дер Корпут са строго по-малко от 1. Първоначалната редица на ван дер Корпут записва цяло число в основа 2, но могат да бъдат получени модификации на редицата при по-голяма основа. Редицата е успешно приложена за изчисляване на глобалните индекси на чувствителността на Собол, съответстващи на влиянието на няколко входни параметъра (шест скоростни константи на химични реакции и четири различни групи замърсители) върху концентрациите на важни замърсители на въздуха. Числените експерименти показват, че разглежданите редици на Ван дер Корпут са ефективни за разглежданите многомерни интеграли и особено за изчисляване на малки по стойност индекси на чувствителност. Докато за ниски размерности точковото множество с генерирщ вектор обощената редица на Фибоначи има предимство, за високи размерности редицата на ван дер Корпут постига по-добра относителна грешка и от точковото множество, и близки резултати с извадката латински хиперкуб.

## 2.2 Прилагане и сравнение за първи път на съществуващи стохастични методи за анализ на чувствителността;

В [**В7**] за първи път е направено е приближено пресмятане на пълните индекси на чувствителността посредством трите най широко използвани квазислучайната рецици на Собол, Холтън и Фор. Тези три редици никога досега не са били сравнявани при анализ на чувствителност на конкретния модел и това мотивира настоящото изследване. Редицата на Фор е конструирана за първи път през 1982 [15] и за първи път се прилага за анализ на чувствителността на модела, заедно с публикация [**Г2**] от списъка по конкурса. Редицата на Фор е като редицата на Холтън, но използва само една основа за всички размерности, като основата е най-малкото просто число, което е по-голямо или равно на броя на измеренията (или 2 за едномерна задача). За анализа на чувствителността е използвано и точковото множество с оптимален генерираща вектор, описан в [19]. В някои случаи на малки

по стойност индекси на чувствителност точковото множество с оптимален генериращ вектор дава по-добри резултати от квазислучайната редица на Собол. За големите по-стойност индекси на чувствителност точковото множество показва почти винаги по-добра относителна грешка от редицата на Собол. Редицата на Холтън дава с един порядък по-лоша относителна грешка от редицата на Собол. Редицата на Фор дава ненадеждни резултати за малките по-стойност индекси на чувствителността, като резултатите са лоши дори в случая на големи по стойност индекси на чувствителност, представени от 12-мерни интеграли. Проведения анализа на чувствителността се очаква да има изключително важна двустранна роля: за тестване и подобряване на математическите модели, и за надеждна интерпретация на числените резултати от страна на съответните специалистите.

В [**B8**] е направено сравнение между точково множества тип решетки с оптималния генериращ вектор, точково множество тип решетка с генериращ вектор обобщената редица на Фибоначи и квазислучайната редица на Собол. Показано е, че генериращия вектор получен по метода на бързата конструкция има предимство спрямо генериращия вектор базиран на обобщената редица на Фибоначи от съответната размерност. Направени са сравнения и с квазислучайната редица на Собол. Разгледани са многомерни интеграли с ниска и висока размерност, съответно 4,10 и 25-мерни, които намират приложение при пресмятането на числените индикатори на чувствителността на сложни екологични системи с голяма размерност. Докато за 4-мерния интеграл точковото множество с генериращия вектор обощената редица на Фибоначи постига малко по-добра относителна грешка от тази на оптималния генериращ вектор, за 10 и 25-мерния интеграл относителната грешка, получена с точковото множество базирано на оптималния генерирщ вектор превъзхожда относителната грешка получена с точковото множество базирано на обобщената редица на Фибоначи с 2 порядъка и квазислучайната редица на Собол с 1 порядък, което демонстрира предимствата на оптималния генериращ вектор за сложни системи с голяма рзмерност.

## 2.3 Разработване на нов метод Монте Карло за линейни системи с приложения в екологията

Разработен е нов метод Монте Карло за решаване на системи от линейни алгебрични уравнения (IWE). Новия алгоритъм Монте Карло за решаване на линейни системи от вида *Bx=f*, е реализиран на базата на метода ,,случайно блуждаене по уравненията‘‘ (WE) на лиейната система, разработен в [13] с използване на дискретна верига на Марков. Матрицата $B$ и дясната част $f$ са нормирани, за да се ускори сходимостта на стохастичния процес. Дефинираме итерационната матрица на алгоритъма като $A$=I-Db, и системата се представя във вида x=Ax+b, където b=Df и D е диагоналнта матрица. Тогава дефинираме релаксационния параметър като $g=d_i \times b_{ii}$, където $d_i$ са елементите на диагоналната матрица D и $b_{ii}$ са диагоналните елементи на матрицата B. Направен е оптимален избор на релаксационния параметър, което води до балансиране на итерационната матрица и до повишаване на точността на алгоритъма. Поради реализацията си, методът е в пъти по-бърз от Монте Карло метода WE, като експериментите показват и предимство по отношение на точността, което се вижда особено за матрици с голяма размерност. Новият метод постига много по-добра точност за по-малко време и по-малък брой отнапред зададени итерации. Методът може успешно да се конкурира с градиентните методи, като съществуват матрици като NOS4, за които методът Монте Карло е по-добър от метода на спрегнатия градиент. Направено е сравнение със стандартния рафиниран метод Монте Карло [10] за линейни системи и с оригиналния метод WE, като получените резултати утвърждават новия метод Монте Карло като един от най-добре известните методи Монте Карло за линейни системи. Благодарение на направените оптимизации новият метод дава много по-добри резултати и от рафинирания [10] метод Монте Карло и е установен като един от най-бързите и точни алгоритми Монте Карло за решаване на системи от линейни алгебрични уравнения. Матрицата NOS4 [30] играе важна роля за моделирането на преноси на замърсители във въздуха и

екологичната безопасност. Тя е свързана с анализа на чувствителността на сложна екологична система, описваща модел на далечен пренос на замърсители във въздуха.

[**B5**] е разширение на статия [**B6**] и е описан по-подробно новият метод Монте Карло за линейни системи и са разширени неговите приложения за моделиране на сложни екологични системи. Много научни и инженерни приложения се основават на на решаването на големи системи от линейни алгебрични уравнения. Времето за изчисляване на задачи от голяма размерност или за намиране на решения в реално време, може да бъде ограничено и това прави използването на много от установените алгоритми нецелесъобразно. Подобреният алгоритъм се основава на наскоро разработения метод Монте Карло ,,случайно блуждаене по уравненията“, (,,walk on equations“ (WE)), базиран на недисконтирана сума на случайни траектории с поглъщане. За усъвършенстване на основния алгоритъм се прилагат няколко техники, като броене на посещенията на уравненията на линейната система и последователния метод Монте Карло на Джон Холтън [13]. Алгоритъмът е оптимизиран чрез избиране на подходящи стойности на релаксационния параметър, което води до намаляване на изчислителната сложност и на повишаване на точността за даден брой итерации. Предимствата на алгоритъма могат да бъдат наблюдавани особено за системи с много голяма размерност. Поведението на алгоритъма Монте Карло за линейни системи не зависи от плътността на матрицата. Матрицата NOS4 има само 5.9 средно ненулеви елемента в колона и ред. Предимството на алгоритъма е в сила както за разредени, така и за плътни матрици. Извършени са числени експерименти за примери с разредени и плътни матрици с различна размерност и за матрицата NOS4 с фундаментално приложение, взета от приложения свързани с апроксимация по матода на крайните елементи на модел, описващ гредова структура в конструктивната механика. Матрица NOS4 се използва при направата и обдухването на крилата на самолетите и има важно значение при опазването на

околната среда и намира приложения в моделирането на сложни екологични системи.

За матрицата NOS4 е установено, че разработения алгоритъм (IWE) дава по-добри резултати от широко използвания метод на спрегнатия градиент с преобуславяне (PCG). Резултатите, представени в статията, показват, че сходимоста за IWE е по-добра: кривата, представяща относителната грешка с броя итерации, е по-гладка и постига точност $10^{-6}$ или $10^{-7}$, докато кривата, получена с PCG, постига точност $10^{-3}$, като в експериментите сме поискали точност $10^{-8}$. Този резултат може да бъде обяснен по следния начин. Успехът на метода PCG се основава на факта, че той работи в оптимални подпространства на Крилов. Нашият IWE алгоритъм също работи в оптимални подпространства на Крилов. Разликата е, че докато в PCG всъщност се решава оптимизационна задача и ако има локални минимуми, близки до глобалния, сходимостта на процеса може да попадне в локалния минимум, какъвто е и случаят с PCG. Нашият алгоритъм е независим от такава оптимизационна процедура, поради което резултатите са по-добри. Конкретната матрица (NOS4) изглежда е труден тест за PCG. Този факт не може да се обобщи, защото успехът на метода зависи от конкретния функционал, който трябва да се минимизира при метода PCG. Това е въпрос на допълнителен анализ, който не е тривиален. Не може да се гарантира, че подобен ефект се случва за всяка матрица, но има много важни случаи като матрицата NOS4, в които разработеният оптимизиран Монте Карло метод IWE се представят по-добре от един от най-добрите и широко използвани методи PCG. За матрици като NOS4, които имат локални минимуми, близо до глобалния е известно, че когато детерминистичните алгоритми попаднат в локален минимум, който е съседен на глобален минимум, не могат да излязат от него, алгоритъмът остава в този локален минимум и зацикля, докато методите Монте Карло, поради стохастичната си природа, дори да попаднат в локални минимуми, лесно излизат от тях. Затова за матрицата $\mathrm{NOS4}$ сходимостта на Монте Карло метода IWE е по-добра от тази на метода на PCG.

В [**B6**] е разработен нов подобрен алгоритъм Монте Карло (IWE) за линейни системи на базата на Монте Карло метода ,,случайно блуждаене по уравненията‘‘

(WE). Направено е сравнение и с метода на спрегнатия градиент и с рафинирания [10] метод Монте Карло за матрицата NOS4. Алгоритъмът е използван за намиране на всички компоненти на решението на реални линейни системи с голяма размерност, а именно 7,100 и 5000-мерни матрици. За 5000-мерната матрица нашият подобрен метод дава чувствително по-добри резултати от оригиналния метод. За 5000-мерната плътна матрица, след 30 итерации на алгоритъма, WE дава относителна грешка с порядък $10^{-9}$, докато подобреният алгоритъм IWE дава относителна грешка $10^{-16}$ и също така оптимизирания IWE е 15 пъти по бърз от предшественика си WE. Поведението на предложения алгоритъм Монте Карло не зависи от плътността на матрицата. Благодарение на направените оптимизации новият метод дава много по-добри резултати от стандартния метод Монте Карло и е установен като един от най-бързите и точни алгоритми Монте Карло за решаване на системи от линейни алгебрични уравнения с много голяма разменрост.

## 3. Заключение

Настоящото изследване е насочено към екологична безопасност. Разработен е нов стохастичен подход за решаване на линейни системи с голяма размерност, които имат важно значение за екологичното равновесие. Приложена и изследвана е систематизирана схема за анализ на чувствителноста към голям модел на замърсители във въздуха-UNI-DEM, за да се проучи чувствителността на концентрациите на важни замърсители във въздъха спрямо вредни емисии и спрямо скоростни константи на някои избрани химични реакции. Повечето от резултатите могат да бъдат приложени и за други големи математически модели. Анализът на чувствителността и в частност получените резултати имат изключително важна двустранна роля: за тестване и подобряване на математическите модели, и от друга страна, за надеждна интерпретация на числените резултати от страна на съответните специалистите. Анализът, основан на дисперсията, е много ефективен апарат за задълбочено проучване на зависимостта между отделните параметри, изходните резултати и вътрешните механизми, регулиращи разглежданата система. Определяйки основните химични

реакции, влияещи върху поведението на системата, специалистите от различни приложни области (физика, химия) ще могат да получат ценна информация за подобряването на модела, а това от своя страна ще доведе до повишаване на надеждността и устойчивостта на прогнозите. Така чрез анализ на чувствителността математическият модел ще подпомогне за изготвянето на по-точни прогнози за въздействието на вредните емисии върху човешкото здраве и загубите в селското стопанство. Предимството на избрания метод за анализ на чувствителността е възможността да се пресметнат не само индексите от първи ред, но и индексите от по-висок ред, което носи информация за съвместния принос на входните параметри върху измененията на изходния резултат, а освен това избрания подход се характеризира с оптимална изчислителна сложност, която е пропорционална на размерността на случайната извадка и на броя на параметрите. Числените експерименти потвърждават, че разгледаните стохастични алгоритми са ефективни за многомерно числено интегриране и по-специално за пресмятането на малки по стойност индекси на чувствителността. За първи път за задачата са приложени точкови множества тип решетка с генериращ вектор обобщената редица на Фибоначи, генериращ вектор от олиномиална функция и оптимален генериращ вектор получен по методът на бързата конструкция. За първи път за задачата са разработени модификации на адаптивния алгоритъм и редицата на ван дер Корпут. За първи път са приложени и сравнени квазислучайната рецица на Холтън, Фор и Собол, и нейни модификации, както и извадката латински хиперкуб. Това е изключително важно за цялостната процедура на анализ на чувствителността на математически модели, както и за надеждната интерпретация на резултатите от тези модели. Получените резултати показват, че разработените и използвани за пръв път подходи за анализ на чувствиелността са един от най-добрите методи, от гледна точка на изчислителна ефективност и точност.  Стохастичните методи, основани на оптимални генериращи вектори, използвани за първи път за анализ на чувствителност на сложни системи с голяма размерност, отварят нови възможности за по прецизен и надежден анализ на чувствителност. Като бъдеща работа се очаква за моделиране на сложната система с голяма размерност да се разработят

и приложат за първи път свръх сходящи стохастични методи, използващи вместо оптимални генериращи вектори, оптимални генериращи матрици.

## ИЗПОЛЗВАНА ЛИТЕРАТУРА


1. **B1** I. Dimov, R. Georgieva, V. Todorov, Tz. Ostromsky. (2017) Efficient stochastic approaches for sensitivity studies of an Eulerian large-scale air pollution model, AIP Conference Proceedings 1895(1):050004,· October 2017 DOI: 10.1063/1.5007376, ISBN 978-0-7354-1579-9, ISSN 0094-243X, **SJR(2017):0.165**
   https://doi.org/10.1063/1.5007376
   AIP Conference Proceedings (scimagojr.com)

2. **B2** Venelin Todorov, Ivan Dimov, Tzvetan Ostromsky, and Zahari Zlatev. (2021) Advanced Quasi-Monte Carlo Algorithms for Multidimensional Integrals in Air Pollution Modelling. In: Dimov I., Fidanova S. (eds) Advances in High Performance Computing. HPC 2019. Studies in Computational Intelligence, vol 902. Springer, SCI 902, pp. 155–167, 2021. Print ISBN978-3-030-55346-3 Online ISBN 978-3-030-55347-0,
   **SJR(2019):0.215 Q4**
   https://doi.org/10.1007/978-3-030-55347-0_14
   Studies in Computational Intelligence (scimagojr.com)

3. **B3** Tzvetan Ostromsky, Venelin Todorov, Ivan Dimov, and Zahari Zlatev. (2021) Sensitivity Studies of an Air Pollution Model by Using Efficient Stochastic Algorithms for Multidimensional Numerical Integration. In: Dimov I., Fidanova S. (eds) Advances in High Performance Computing. HPC 2019. Studies in Computational Intelligence, vol 902. Springer, SCI 902, pp. 184-195, 2021. Print ISBN978-3-030-55346-3, Online ISBN 978-3-030-55347-0,

**SJR(2019):0.215 Q4**

https://doi.org/10.1007/978-3-030-55347-0_16

Studies in Computational Intelligence (scimagojr.com)

4. **B4** V. Todorov, I. Dimov and Tz. Ostromsky. (2020) A comparison of advanced quasi Monte Carlo methods for multidimensional integrals in air pollution modeling, AIP Conference Proceedings 2302, 030005, 2020,

   **SJR(2019):0.19**

   https://doi.org/10.1063/5.0034850

   AIP Conference Proceedings (scimagojr.com)

5. **B5** Todorov V., Ikonomov N., Apostolov S., Dimov I., Georgieva R., Dimitrov Y. (2020) An Improved "Walk on Equations"Monte Carlo Algorithm for Linear Algebraic Systems, Recent Advances in Computational Optimization, In: Fidanova S. (eds) Recent Advances in Computational Optimization. Studies in Computational Intelligence, vol 838. Springer, Cham, Print ISBN978-3-030-22722-7, Online ISBN978-3-030-22723-4,

   **SJR(2019):0.215 Q4**

   https://doi.org/10.1007/978-3-030-22723-4_14

   Studies in Computational Intelligence (scimagojr.com)

6. **B6** Venelin Todorov, Ikonomov Nikolay, Dimov Ivan, Georgieva Rayna. (2018) A New Monte Carlo Algorithm for Linear Algebraic Systems Based on the "Walk on Equations"Algorithm. Proceedings of the 2018 Federated Conference on Computer Science and Information Systems, Annals of Computer Science and Information Systems, 15, **IEEE**, 2018, ISSN:2300- 5963, pp. 257-260, DOI: 10.15439/2018F121

   http://dx.doi.org/10.15439/2018F121

   https://ieeexplore.ieee.org/document/8511284

7. **B7** Dimov I., **V. Todorov**, Tz. Ostromsky. (2020) Advanced Quasi Monte Carlo Methods for Sensitivity Analysis in Air Pollution Modeling. Knowledge, Language and Models, Volume in Honour of Prof. Galia Angelova on Her 65th Birthday. Milena Slavcheva, editor. INCOMA Ltd. Shoumen, BULGARIA. INCOMA (IIKT), 2020, ISSN:978-954-452-062-5, 74-90
8. **B8 Todorov V.**, Dimitrov V., Tsvetkov I.. (2018) A NUMERICAL STUDY ON QUASI MONTE CARLO METHODS BASED ON RANDOMLY SHIFTED LATTICE RULE FOR COMPUTATION OF MULTIPLE INTEGRALS. Journal Scientific and Applied Research, Konstantin Preslavsky Publishing House, vol. 13, 2018, ISSN:1314-6289, 21-30.
   journal-scientific-and-applied-research-vol-13.pdf (rst-tto.com)

9. Bahvalov, N.: On the Approximate Computation of Multiple Integrals. In Vestnik Moscow State University, Ser. Mat., Mech., volume 4, pages 3–18, 1959.
10. Dimov, I.: Monte Carlo Methods for Applied Scientists, New Jersey, London, Singapore, World Scientific (2008), 291 p.
11. I. T. Dimov, R. Georgieva. Monte Carlo Algorithms for Evaluating Sobol' Sensitivity Indices. Math. Comput. Simul. (2009).
12. Dimov, I., Karaivanova, A., Georgieva, R., & Ivanovska, S. (2002, August). Parallel importance separation and adaptive Monte Carlo algorithms for multiple integrals. In International Conference on Numerical Methods and Applications (pp. 99-107). Springer, Berlin, Heidelberg.
13. Dimov, I. T., Maire, S., Sellier, J. M.. A New Walk on Equations Monte Carlo Method for Solving Systems of Linear Algebraic Equations. Applied Mathematical Modelling, 39, 15, Elsevier, 2015, ISSN:0307-904X, DOI:10.1016/j.apm.2014.12.018, pp. 4494-4510.
14. Hua, L.K., Wang, Y.: Applications of Number Theory to Numerical analysis, 1981
15. Henri Faure, Discrepance de suites associees a un systeme de numeration (en dimension s), Acta Arithmetica, Volume 41, 1982, pages 337-351.

16. Halton, J. (1964), "Algorithm 247: Radical-inverse quasi-random point sequence", Communications of the ACM, **7**: 701-701.
17. Karaivanova, Aneta, and Ivan Dimov. "Error analysis of an adaptive Monte Carlo method for numerical integration." *Mathematics and Computers in simulation* 47.2-5 (1998): 201-213.
18. Korobov, N.M.: Properties and calculation of optimal coefficients. Doklady Akademii Nauk SSSR, 132:1009–1012 (Russian), 1959. English Translation: Soviet Mathematics Doklady, 1, 696–700.
19. F.Y. Kuo and D. Nuyens Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients – a survey of analysis and implementation, Foundations of Computational Mathematics 16(6), 1631–1696, 2016.
20. McKay, M.D.; Beckman, R.J.; Conover, W.J. (May 1979). "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code". Technometrics. American Statistical Association. 21 (2): 239–245.
21. Saltelli A.: Making Best Use of Model Valuations to Compute Sensitivity Indices. Computer Physics Communications, 145:280–297, 2002.
22. A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo. Sensitivity Analysis for Chemical Models. Chemical Reviews, 105:2811–2828, 2005.
23. Saltelli A, Tarantola S., Campolongo F, and Ratto M: Sensitivity Analysis in Practice: a Guide to Assessing Scientific Models, 2004.
24. I.M. Sobo`l. Sensitivity Estimates for Nonlinear Mathematical Models. *Mathematical Modeling and Computational Experiment*, 1:407–414, 1993.
25. Sobol,I.M. (1967), "Distribution of points in a cube and approximate evaluation of integrals". *Zh. Vych. Mat. Mat. Fiz.* **7**: 784–802 (in Russian); *U.S.S.R Comput. Maths. Math. Phys.* **7**: 86–112 (in English).
26. van der Corput, J.G. (1935), "Verteilungsfunktionen (Erste Mitteilung)" (PDF), Proceedings of the Koninklijke Akademie van Wetenschappen te Amsterdam (in German), 38: 813–821, Zbl 0012.34705
27. Wang, Y., Hickernell, F.J.: An historical overview of lattice point sets, (2002).

28. Zlatev, Z.: Computer Treatment of Large Air Pollution Models (Kluwer Acad. Publishers, 1995).
29. Zlatev, Z., Dimov, I.: Computational and Numerical Challenges in Environmental Modelling (Elsevier, Amsterdam, 2006).
30. Website: Matrix market, NOS4: Lanczos with partial reorthogonalization. Finite element approximation to a beam structure, http://math.nist.gov/MatrixMarket/data/Harwell-Boeing/lanpro/nos4.html

Дата: 22.3.2021 Подпис:

/Венелин Тодоров/