

РП5: Монте Карло методи за анализ на чувствителността на големи математически модели

1. Отчет за изпълнение на задачите по пакета

Задача 5.1: Генериране на мрежови функции за различни климатични сценарии (химични замърсители, химични реакции), използвайки 2- и 3-D Унифициран датски Ойлеров модел (UNI-DEM).

Чрез поредица от тестови експерименти и анализ на скалируемостта върху различен брой процесори на суперкомпютъра IBM BlueGene/P е усъвършенствано изпълнението на пакета SA-DEM за анализ на чувствителността на Датския Ойлеров модел (във версията му с три йерархични нива на паралелизъм) за ефективно паралелно изпълнение върху IBM BlueGene/P. С помощта на тази версия на SA-DEM са генерирани голям брой данни за чувствителността на модела по отношение на скоростните коефициенти на част от химичните реакции, които бяха използвани за разработването на по-ефективни Монте Карло методи за анализ на чувствителността на DEM. Получените резултати са изложени детайлно в [DGOZ_13c, OGZ_12, ODGMZ_13b].

Направени са и нова серия от паралелни експерименти за уплътняване на мрежата от данни на мрежовите функции, използвани при пресмятането на глобалните индекси на чувствителност на Собол за DEM по отношение на скоростните константи. Резултатите са докладвани на 9th International Conference on „Large-Scale Scientific Computations“ (<http://parallel.bas.bg/Conferences/SciCom13/>), 3 - 7 юни, 2013, Созопол [6].

Инсталирана е и е оптимизирана за паралелно изпълнение специализираната версия за анализ на чувствителността на Датския Ойлеров модел SA-DEM с три йерархични нива на паралелизъм върху клъстъра IBM MareNostrum III на Суперкомпютърния център в Барселона, Испания (BSC). На първото, най-високо ниво се прави разпаралеляване на задачата по различните стойности на вектора от пертурбационни коефициенти, прилагани към анализирания параметри на модела. На средното ниво се прави разпаралеляване на базата на разделяне на изчислителната област. В тези две нива се използва модела на паралелизъм с разпределена памет (реализиран чрез библиотеката MPI от стандартни комуникационни процедури). На третото, най-ниско ниво се използва модела на паралелизъм с обща памет (реализиран с помощта на OpenMP директиви). Приложението му е хардуерно ограничено до наличните ядра в рамките на един възел от клъстъра (в случая – 16). IBM MareNostrum III е най-мощният за сега испански суперкомпютър, с 3028 16-просесорни възела и близо 1 PetaFLOPS пикова производителност. Освен по-големият брой и по-бързи процесори в сравнение с българския IBM BlueGene/P, тази машина има и значително повече оперативна памет (32 GB RAM на възел, при само 2 GB RAM на възел за BlueGene/P). Това даде възможност да се направят множество успешни експерименти върху различен брой процесори на MareNostrum и с версията на SA-DEM с най-фината мрежата за дискретизация на областта (480 x 480), която създаваше известни проблеми при изпълнение върху IBM BlueGene/P. Това, от своя страна, налагаше ограничения върху предишните експерименти с модела. Получените резултати от числените експерименти върху испанския суперкомпютър показват ефективността и добрата скалируемост на новата реализация на SA-DEM. Направени са и редица експерименти за получаване на необходимите данни за анализ на чувствителността на Датския Ойлеров модел по отношение на шест от скоростните коефициенти в химическия подмодел. Част от резултатите са представени в [ODGMZ_13a].

Получените нови резултати са докладвани на деветия международен семинар на IMACS по методи Монте Карло, проведен от 15 до 19 юли в Анси, Франция [8].

Задача 5.2: Изследване на различни техники за апроксимиране на мрежови функции, оценка на грешката.

Проведено е изследване на различни апроксимационни апарати за приближаване на данните от модела UNI-DEM при проучване на чувствителността на концентрациите:

- на амоняк, озон, азотен сулфат и азотен нитрат спрямо колебанията в нивата на четири основни групи вредни емисии с антропогенен произход, както и
- на озон спрямо вариациите в скоростните константи на шест химични реакции.

За изследването са взети полиноми от различни степени както по отделните променливи, така и спрямо общата степен на полинома. Проведен е експериментален анализ на надеждността на получените приближения, като е оценена средноквадратичната грешка. Първоначално, следвайки само този критерий, всички разгледани полиноми дават удовлетворителна точност. Някои допълнителни тестове дадоха възможност за прецизиране и уточняване на надежден апарат, а именно - сравнението с минималните и максималните стойности на данните от модела в съответните дефиниционни области.

Резултатите са докладвани на Fifth Conference of the Euro-American Consortium for Promoting the Application of Mathematics in Technical and Natural Sciences (<http://2013.eac4amitans.eu/index.html>), Албена, България, 24 - 29 юни, 2013 [7].

Задача 5.3: Анализ на чувствителността на изходните параметри на UNI-DEM към нивата на емисиите с цел подпомагане на вземащите решения при изработване на екологичните директиви; визуализация на резултатите.

Основни резултати от изследванията са представени в поканен пленарен доклад на 9-тия IMACS Seminar on Monte Carlo Methods, Annecy-le-Vieux, France, July 15–19, 2013 [2].

Изследвана е чувствителността на концентрациите на озон, симулирани от модела UNI-DEM, спрямо скоростните константи на шест химични реакции. Числените експерименти са направени за четири европейски града (Генуа, Милано, Манчестър и Единбург) и са осъществени със софтуерни пакети за научни пресмятания (Matlab и Mathematica). За пресмятането на индексите на чувствителността е приложен подходът *корелирана извадка*, предложен от Салтели и Кучеренко. В сравнение със известния подход на Соболев, този метод дава надеждни стойности за индексите и преодолява загубата на точност в случая на малки по стойност индекси на чувствителността. Получените резултати за индексите на чувствителността показват: а) адитивност на модела спрямо разглежданите входни параметри; б) потвърждаване на изводите, направени за по-малък брой химични реакции; в) установяване на още един *важен* входен параметър (скоростната константа на химична реакция #1, зависеща от времето) за модела; г) несъществено влияние на географското разположение върху стойностите на оценяваните величини. Резултатите са описани в [DGOZ_13a].

Проведено е експериментално проучване на алгоритъм Монте Карло, основан на квазислучайни редици на Соболев, алгоритъм квази-Монте Карло и алгоритъм с разбъркани квазислучайни редици на Соболев. Всеки от разгледаните алгоритми води до достатъчно ефективно пресмятане на неизвестните величини – стойности на многомерни интегрални с гладка и негладка подинтегрална функция, както и с приближаващата функция на данните, генерирани с модела UNI-DEM и по-точно със специализираната за анализ на чувствителността версия на модела (SA-DEM). Алгоритъмът Монте Карло, основан на модифицирани редици на Соболев, дава подобни резултати за порядъка на относителната грешка в сравнение с алгоритмите с разбъркани квазислучайни редици. Но има и редица случаи, в които този алгоритъм има съществено предимство, особено за малки по стойност индекси на чувствителността. Резултатите са описани в [DGOZ_13b].

Проведен е анализ на чувствителността на концентрациите на озон, амоняк, амониев сулфат и

амониев нитрат, симулирани от модела UNI-DEM, спрямо нивата на четири групи от вредни емисии - азотен оксид и диоксид, серен диоксид, амоняк, антропогенни въглеродороди, в четири европейски градове с различно географско разположение (Милано, Манчестър и Единбург). Проведено е задълбочено изследване на ефективността за техники за намаляване на дисперсията при пресмятането на малки по стойност индекси на чувствителността. Въпреки че някои от индексите от първи ред са малки по стойност, опростявания на модела не могат да се направят на този етап, защото всяка група от емисии влияе съществено върху някои от избраните химични съединения. Резултатите са описани в [DGOZ_13c].

Изследвани са подходи за анализ на чувствителността в случай на зависими променливи. За пресмятането на индексите от първи ред и пълните индекси на чувствителността е реализиран алгоритъм, предложен от Кучеренко, Тарантола и Аньони. Алгоритъмът представлява обобщение на стандартния подход на Собол за независими променливи. Като предварителна информация е необходимо да се знае разпределението на променливите, всяка от които е случайна величина. В случая на нормално разпределение, което се наблюдава в конкретния модел, прилагането на алгоритъма е директно. Резултатите са представени на седмото издание на международната конференция по анализ на чувствителността на математически модели SAMO 2013 (<http://www.gdr-mascotnum.fr/mascot13.html/>), която се проведе в Университета на Ница [3].

Приложени са ефективни алгоритми Монте Карло за многомерно интегриране при решаването на интегрални уравнения. Използва се подход, при който интегралното уравнение се свежда до пресмятането на краен брой интегрални. Алгоритмите са разработени и изследвани за многомерни интегрални при предишни наши изследвания, докато целта на настоящите изследвания е изследване на поведението им при решаването на интегрални уравнения. Алгоритмите използват псевдослучайни числа, които представляват „размърдани“ или „отместени“ квазислучайни числа на Собол. В единия случай новата псевдослучайна точка се генерира случайно върху сферата с център съответната квазислучайна точка, а във втория – освен първата псевдослучайна точка се генерира и втора, която е централно-симетрична на първата. Резултатите за относителната грешка и изчислителното време, получени с тези алгоритми, са сравнени с резултатите, получени с известен подход на Собол за интегрални уравнения. Основните предположения са, че поставената задача се решава с предварително зададена точност, както и че различните типове грешки при апроксимацията са балансирани. Направените числени експерименти демонстрират надеждност на получените резултати, както и известно предимство на изследваните алгоритми по отношение на изчислителната сложност. Резултатите са представени на деветото издание на международния семинар по методи Монте Карло IMACS MCM 2013 (<http://www.lama.univ-savoie.fr/IMACS2013/>), който се проведе в Университета на Савоя [4].

Задача 5.4: Компютърно моделиране и изследване на ефекта от реалните глобални климатични промени и изменения в количествата на вредните емисии върху нивата на замърсяване на въздуха в Югоизточна Европа, и в частност върху територията на България при отчитане на резултатите, получени за чувствителността на изходните резултати на UNI-DEM.

Направени са нови тестови експерименти с паралелния код на Датския Ойлеров модел за изследване на далечен пренос на замърсители във въздуха (UNI-DEM) на суперкомпютъра IBM Blue Gene/P с цел усъвършенстване на компютърния модел и настройка на параметри в модела, както и изследване на влиянието на глобалните промени в климата върху нивата на замърсяване на атмосферата.

В [ZGD_13a] са представени резултати от направените компютърни експерименти с реални метеорологични и емисионни данни с паралелната версия на UNI-DEM за завършване на изследването на влиянието на глобалните климатични промени върху нивата на замърсяване на въздуха в страните от Балканския полуостров. Дискутирани са и са направени изводи относно

следните важни за обществото въпроси: (i) влиянието на вредните емисии произведени в страни от Европа извън Балканския регион върху нивата на концентрациите на опасни замърсители на въздуха в страните от Балканския полуостров; (ii) влиянието на промените в нивата на замърсяване на въздуха върху добиви в селското стопанство и върху състоянието на горските масиви; (iii) измененията в броя на „лошите дни“ (дните, в които замърсяването с озон надхвърля максимално допустимия безопасен праг) вследствие на промените в климатичната обстановка; (iv) влиянието на повишаването на температурата отделно или в съчетание с други фактори върху нивата на някои опасни за здравето на хората замърсители.

В [ZDFG_13] са обобщени в многомерен аспект някои резултати получени във връзка с моделирането на процесите на адвекция в Датския Ойлеров модел за далечен пренос на замърсители във въздуха. Показано е, че висока степен на точност може да се постигне и в многомерния случай при комбинация на числената схема на Кранк-Никълсън с екстраполация на Ричардсън. Приложенията на разработената техника надхвърлят рамките на задачите за пренос на замърсители и имат приложение при решаване на задачи с голяма размерност в други клонове на науката и инженерните науки.

В [ZG_13] са представени и дискутирани извършените компютърни експерименти с цел изследване на използването на преобусловители, построени на основата на приближена LU факторизация с използване на всички преимущества на характерните свойства на коефициентните матрици на системите, в т. ч. тяхната разреденост. Дадени са важни изводи относно компютърната ефективност на различни алгоритми и избора на критерии за край на изчисленията.

В [ZGD_13b] са представени получените резултати от изследването на областите на абсолютна устойчивост при решаване на обикновени диференциални уравнения с методите на Рунге – Кута от ред от едно до четири, включително комбинирано с използването на екстраполация на Ричардсън. Това води до възможност на по-големи стъпки по времето при компютърната реализация на съответните алгоритми и съответно значително съкращаване на времето за изпълнение на съответните компютърни кодове.

В [ZGD_13c] са представени резултати от направените компютърни експерименти с реални метеорологични и емисионни данни с паралелната версия на UNI-DEM за изследване на влиянието на систематичните изменения на замърсяванията, причинени от човешки дейности в различни данни на изчислителната област, която обхваща главно територията на Европа. Установено е, че докато при сулфатните замърсители има пряка и почти линейна зависимост между тяхното изменение и това на замърсяванията от човешки дейности, то при болшинството вредни замърсявания това не е така. Показано е, че въпреки че изменението на замърсяванията от човешки дейности да е едно и също за цялата изчислителна област изменението в количеството на едни и същи замърсители е различно в различните подобласти на модела.

В [ZGD_13d] са представени резултати относно направените компютърни симулации с цел изучаване на паралелна ефективност на Датския Ойлеров модел за изследване на далечен пренос на замърсители във въздуха (UNI-DEM). Изследвано е влиянието на несигурността на входните данни (метеорологични данни, емисионни данни и скоростите на химичните реакции, които са реализирани в модела). Изследвана е организацията на паралелните изчисления и скалируемостта на използвания алгоритъм.

2. Списък на публикации по проекта, в които е цитиран проект ДЦВП 02/1

(i) *излезли от печат*

[DGOZ_12]

I. T. Dimov, R. Georgieva, Tz. Ostromsky, Z. Zlatev, *Variance-based Sensitivity Analysis of the Unified Danish Eulerian Model According to Variations of Chemical Rates*. - In: I. Dimov, I. Faragó, and L. Vulkov (Eds.) Proceedings of NAA 2012, LNCS 8236, Springer, 2013, 247 – 254. ISSN: 0302-9743. SJR (2012):0.517.

[DGOZ_13b].

I. T. Dimov, R. Georgieva, Tz. Ostromsky, Z. Zlatev, *Advanced Algorithms for Multidimensional Sensitivity Studies of Large-scale Air Pollution Models based on Sobol Sequences*. Special issue of Computers and Mathematics with Applications 65 (3), „Efficient Numerical Methods for Scientific Applications“. Elsevier, 2013, 338 - 351. ISSN: 0898–1221. IF (2012): 2.069. 5-year IF: 1.894.

[DGOZ_13c].

I. T. Dimov, R. Georgieva, Tz. Ostromsky, Z. Zlatev. *Sensitivity Studies of Pollutant Concentrations Calculated by UNI-DEM with Respect to the Input Emissions*. Central European Journal of Mathematics, "Numerical Methods for Large Scale Scientific Computing" 11 (8), 2013, 1531 – 1545. Doi: 10.2478/s11533-013-0256-2. IF (2012): 0.405.

[ODGMZ_13b]

Tz. Ostromsky, I. T. Dimov, R. Georgieva, P. Marinov, Z. Zlatev. High Performance Computing of Data for a New Sensitivity Analysis Algorithm, Applied in an Air Pollution Model. - In: I. Dimov, I. Faragó, and L. Vulkov (Eds.) Proceedings of NAA 2012, LNCS 8236, Springer, 2013, 428–436. ISSN: 0302-9743. SJR (2012): 0.517.

[OGZ_12]

Tz. Ostromsky, K. Georgiev, Z. Zlatev, An Efficient Highly Parallel Implementation of a Large Air Pollution Model on an IBM Blue Gene Supercomputer. In: AIP Conf. Proc. 1487, 2012, 135 - 142. ISSN: 0094-243X. ISBN: 978-0-7354-1099-2.

[ZGD_13a]

Z. Zlatev, K. Georgiev, I. Dimov, *Influence of Climatic Changes on Pollution Levels in the Balkan Peninsula*, Computers and Mathematics with Applications, **65**, 2013, 544–562, Elsevier, ISSN 0898–1221, IF (2012): 2.069. 5-year IF: 1.894.

(ii) *приети за печат*

[ODGMZ_13a]

Tz. Ostromsky, I. T. Dimov, R. Georgieva, P. Marinov, Z. Zlatev, *Sensitivity Study of a Large-Scale Air Pollution Model by Using High-Performance Computations and Monte Carlo Algorithms*. AIP Conference Proceedings, ISSN: 0094-243X.

[ZDFG_13]

Z. Zlatev, I. Dimov, I. Farago, K. Georgiev, A. Havasi, Tz. Ostromsky, *Application of Richardson Extrapolation with the Crank–Nicolson Scheme for Multi–dimensional Advection*.

[ZG_13]

Z. Zlatev, K. Georgiev, *Applying Approximate LU-factorizations as Preconditioners in Eight Iterative Methods for Solving Systems of Linear Algebraic Equations*, Central European Journal of Mathematics, Springer, IF: 0581.

[ZGD_13c]

Z. Zlatev, K. Georgiev, I. Dimov, *Sensitivity of European Pollution Levels to Changes of Human-Made Emissions*, глава от електронна книга, озаглавена „Advanced Numerical Methods for Complex Environmental Models: Needs and Availability“.

[ZGD_13d]

Z. Zlatev, K. Georgiev, I. Dimov. *Parallel Computations in a Large-Scale Air Pollution Model*, глава от електронна книга, озаглавена „Advanced Numerical Methods for Complex Environmental Models: Needs and Availability“.

(iii) изпратени за публикуване

[ZGD_13b]

Z. Zlatev, K. Georgiev, I. Dimov, *Absolute Stability Properties of the Richardson Extrapolation Combined with Explicit Runge-Kutta Methods*.

3. Презентации и доклади по темата на проекта

- [1] I. T. Dimov, Sensitivity analysis for large-scale problems, Workshop on Supercomputer Applications, 31 март – 2 април 2013 г., Трявна.
- [2] I. T. Dimov, *Efficient Monte Carlo Algorithms with Applications to Sensitivity Analysis*, Ninth IMACS Seminar on Monte Carlo Methods, July 15 - July 19, 2013, Annecy-le-Vieux, France.
- [3] I. T. Dimov, R. Georgieva, *Global Sensitivity Study of Compact Models in Nanodevices with Correlated Inputs*, 7th International Conference on Sensitivity Analysis of Model Output, Nice, France, July 1-4, 2013.
- [4] I. T. Dimov, R. Georgieva, *Advanced Monte Carlo Algorithms for Solving Integral Equations*, Ninth IMACS Seminar on Monte Carlo Methods, July 15 - July 19, 2013, Annecy-le-Vieux, France.
- [5] I. T. Dimov, R. Georgieva, P. Marinov, Tz. Ostromsky, Z. Zlatev, *Sensitivity Studies of a Large-Scale Air Pollution Model: Approximation Techniques and Monte Carlo Algorithms*, 9th International Conference on „Large-Scale Scientific Computations“, June 3-7, 2013, Sozopol.
- [6] Tz. Ostromsky, I. T. Dimov, R. Georgieva, P. Marinov, Z. Zlatev, *Sensitivity Study of a Large-Scale Air Pollution Model with Respect to Various Input Data Sets, Based on High-Performance Computations*, 9th International Conference on „Large-Scale Scientific Computations“, June 3-7, 2013, Sozopol.
- [7] Tz. Ostromsky, I. T. Dimov, R. Georgieva, P. Marinov, Z. Zlatev, *Sensitivity Study of a Large-Scale Air Pollution Model by Using High-Performance Computations and Monte Carlo Algorithms*, Fifth Conference of the Euro-American Consortium for Promoting the Application of Mathematics in Technical and Natural Sciences, Albena, Bulgaria, June 24-29, 2013.
- [8] Tz. Ostromsky, I. T. Dimov, R. Georgieva, P. Marinov, Z. Zlatev, *A Monte Carlo Method for Sensitivity analysis of an Air Pollution Model - Implementation, Performance and Results*, July 15 - July 19, 2013, Annecy-le-Vieux, France.

4. Други

[1] Специфични дейности от организационно-технически характер:

- а) Участие в организацията и провеждането на семинара „Суперкомпютърни приложения“ в рамките на проекта SuperCA++, 31.03 – 02.04.2013 в гр. Трявна.
- б) Участие в организацията и провеждането на първата Европа – Китай конференция в областта на суперкомпютърните системи “Partnership for supercomputing applications”, проведена в София в периода 08 – 10.04.2013 г.

[2] Организационно-финансови дейности:

- а) Участие в периодични сбирки на оперативна група за организация на работата по РП1, РП4 и РП5 в ИИКТ-БАН (С. Маргенов, Е. Атанасов, И. Димов, К. Георгиев).

[3] Допълнителни дейности по разпространяване и популяризиране на резултатите в рамките на:

- а) Участие в организацията на Conference on Large Scale Scientific Computing, Sozopol, June 3-7, 2013;
- б) Ninth IMACS Seminar on Monte Carlo Methods, July 15 - July 19, 2013, Annecy-le-Vieux, France.