

## СЪПОСТАВИТЕЛЕН АНАЛИЗ НА МОДЕЛИРАНЕТО НА ДАННИ И СИМУЛАЦИОННОТО МОДЕЛИРАНЕ. МЕТОДИ, ПРИЛОЖЕНИЯ И ПРЕДИЗВИКАТЕЛСТВА

Светлана Сярова, Самуел Петков, Александър Кирков

*Университет по библиотекознание и информационни технологии*

**Резюме:** В съвременната информационна ера моделирането на данни и симулационното моделиране оформят бъдещето на науката, индустрията и обществото. Тези методологии променят начина, по който се създава, обработва и използва информацията.

Тази научна статия разглежда значението на моделирането на данни и симулационното моделиране в съвременния информационен контекст, както и техните приложения и перспективи. Разгледани са основните концепции и методи на моделирането на данни и симулационното моделиране. Подчертани са ключовите разлики между тях, както и тяхната роля в различни области на науката и бизнеса.

**Ключови думи:** моделиране на данни, симулационно моделиране

### Въведение

Съществуват много методологии за анализ на данни. В таблица 1 са представени най-известните. В зависимост от целите на анализа и характеристиките на данните във всяка конкретна ситуация, може да се използват комбинации от различни методи [1].

*Таблица 1. Методи за анализ на данни*

Методология	Описание	Приложения
<b>Моделиране на данни</b>	Изграждане на математически модели, които извличат информация от данните и прогнозираят тенденции и стойности.	Прогнозиране, класификация, регресия, клъстерен анализ.
<b>Симулационно моделиране</b>	Създаване на виртуални модели, които симулират динамиката на реални процеси или системи, позволявайки да се извършват експерименти.	Симулация на икономически модели, изследване на разпространението на вируси, анализ на трафика и др.
<b>Дескриптивна статистика</b>	Обобщаване и описване на основните характеристики и тенденции в данните.	Визуализация на данни, анализ на разпределения и тенденции.

Методология	Описание	Приложения
<b>Инферентна статистика</b>	Извличане на заключения и вземане на решения на базата на данни, като се използват статистически методи и техники.	Сравнителен анализ, оценка на ефективността, хипотезни тестове.
<b>Клъстерен анализ</b>	Групиране на данните в различни категории или клъстери, базирани на техните сходства.	Сегментация на пазара, клиентски анализ, обработка на изображения.
<b>Асоциативен анализ</b>	Откриване на взаимодействия и връзки между различни променливи в данните.	Маркетингов анализ, анализ на търговски данни, корелационен анализ.
<b>Времеви редове</b>	Анализ на данни, които са организирани по време, за предвиждане на бъдещи стойности или разпознаване на времеви тенденции.	Прогнозиране на спортни резултати, анализ на финансови времеви редове.

### Методология на изследването

**Моделирането на данни** е методология за анализ, обработка и интерпретация на данни с цел извличане на полезна информация и вземане на решения. Тази методология използва различни статистически техники, машинно самообучение и други алгоритми, за да изгражда модели, способни да прогнозират или класифицират данни, да откриват закономерности и тенденции в данните и други анализи. В таблица 2 е описан процесът на моделиране на данни.

Таблица 2. Процес на моделиране на данни

№	Наименование	Описание
1	Събиране на данни	Събиране на реални данни от различни източници (бази данни, файлове, сензори и др.).
2	Подготовка на данните за анализ	Почистване на данните от непълноти и грешки, преобразуването им във формат, подходящ за анализ и извличане на характеристики или атрибути, които са важни за моделирането.
3	Избор на подходящ модел за конкретната задача	В зависимост от характеристиките на данните и целта на моделирането се използват линейни модели, дървета на решения, невронни мрежи или други алгоритми.

<b>№</b>	<b>Наименование</b>	<b>Описание</b>
4	Трениране на модела с помощта на обучаващи данни	Това включва подкарване на данните през модела и коригиране на параметрите му, така че да минимизира грешката между реалните и предвидените стойности.
5	Оценка на модела с помощта на тестови данни	Определя се колко добре моделът се справя с нови, непознати данни, и при необходимост се правят корекции.
6	Използване на модела	За прогнозиране на нови данни, класифициране на обекти, откриване на аномалии и други приложения в зависимост от конкретната задача.

Моделирането на данни се използва в области, където данните играят ключова роля в процеса на вземане на решения: маркетинг, финанси, здравеопазване, транспорт, енергетика и др. [2, 3].

**Симулационното моделиране** е методология, при която се използват компютърни модели за създаване на виртуална репликация или симулация на реални процеси, системи или сценарии. Чрез него се изследва поведението на системите или процесите в различни условия и се правят прогнози за тяхното бъдещо развитие. Симулационното моделиране включва следните стъпки (таблица 3) [4, 5].

*Таблица 3. Процес на симулационното моделиране*

<b>№</b>	<b>Наименование</b>	<b>Описание</b>
1	Дефиниране на модела	Представя реална система или процес. Това може да бъде математически модел, компютърна програма или комбинация от двете.
2	Събиране на данни и параметризиране	Събират се необходимите данни и се параметризира моделът с входни параметри, описващи началното състояние на системата или процеса.
3	Изпълнение на симулацията	Симулация, базирана на модела. Програмата използва входните параметри, за да създаде виртуална среда и да симулира динамиката на системата или процеса във времето.
4	Анализ на резултатите	Анализират се данните, събрани по време на симулацията, за да се разбере поведението на системата или процеса и да се направят заключения за техните характеристики и прогнози.

Симулационното моделиране се използва в икономиката, биологията, климатологията и други области. То позволява на изследователите да извършват експерименти върху виртуални модели, което може да бъде по-

евтино, по-бързо и по-безопасно от изпитването на реални системи или процеси [6, 7].

Въз основа на посочените по-горе характеристики таблица 4 илюстрира обобщените характеристики на двата модела в различни перспективи. От гледна точка на структурата на моделиране подходът за моделиране на данни свързва входната променлива с изходната променлива под формата на статична карта [8, 9]. Така може да се направи модел на данните без никаква информация.

*Таблица 4. Моделиране на данни и симулационно моделиране – сравнение*

<b>Ключови аспекти</b>	<b>Моделиране на данни</b>	<b>Симулационно моделиране</b>
Цел и приложение	Използва се за анализ и извличане на информация от данни, за да се направят прогнози, класификации и други заключения, които да помогнат за вземането на решения.	Използва модели и алгоритми за създаването на виртуална репликация, която симулира динамиката на реалния свят и позволява да се извършват експерименти и да се предвидят резултати.
Използвани данни	Използва се съществуваща информация и данни за анализ и обработка.	Може да се използват както реални данни, така и генерирани входни данни (математически модели или стохастични процеси), описващи началното състояние на системата.
Методи и техники	Включва използването на статистически методи, машинно самообучение, регресионни модели, класификационни алгоритми и др.	Създават се компютърни модели, които симулират динамиката на системата с помощта на различни алгоритми, диференциални уравнения, агент базирани модели или други математически модели.
Динамика на модела	Статична карта на входната променлива към изходната променлива. Не са необходими познания за системата. (Няма състояние в модела.)	Динамична карта (вход, състояние) за изход. Необходими са познания за системата.  (Състояние Q в модела)
Гъвкавост и адаптивност	Моделирането на данни е по-гъвкаво и адаптивно, тъй като данните може да бъдат актуализирани и моделите да бъдат преобучени с нова информация.	Симулационното моделиране изисква предварително дефинирани модели и параметри и е по-трудно да се адаптира към променящите се условия.

<b>Ключови аспекти</b>	<b>Моделiranje на данни</b>	<b>Симулационно моделиране</b>
Нива на анализ	Описателен – Прогнозиращ	Прогнозиращ – Предписателен – Когнитивен
Време за симулация (решение)	Много кратко – в реално време.	Сравнително дълго – трудно се осъществява в реално време.
Приложения и области на приложение	Моделирването на данни се използва за анализ на пазарните тенденции, прогнозиране на бъдещи събития, персонализиране на услуги и др.	Симулационното моделиране се прилага в инженерството, науката за компютрите, биологията, за прогнозиране на икономически и социални сценарии и в други области, където е важно да се изследва поведението на системи или процеси в различни сценарии.

От друга страна, подходът на симулационното моделиране е съставяне на динамична карта на входа и състоянието към изходната променлива въз основа на предизвикан ефект. Този подход изисква познаване на целевата система. Първият подход е проучен в изследователски области на компютърните науки, като извличане на данни и машинно обучение, а вторият е разгледан в изследвания, насочени към ясна причинно-следствена връзка, като физически и оперативни закони. От таблица 4 става ясно, че моделирането на данни се използва по-широко от симулационното моделиране в различни области на науката, бизнеса и технологиите [10, 11].

### **Резултати**

Различните методологии за анализ на данни – моделиране на данни и симулационно моделиране – може да бъдат използвани за различни типове задачи и ситуации. В таблица 5 са представени два примера за сценарии, в които може да се приложат моделиране на данни и симулационно моделиране. Важно е да се има предвид, че изборът на подход зависи от конкретната задача, наличните данни и целите на анализа [12, 13].

В областта на изкуствения интелект (ИИ) се използват различни техники, алгоритми и методи, които имитират човешкия интелект и поведение, като целта е създаването на системи и програми, които могат да изпълняват разнообразни задачи и да решават проблеми. Моделирането на данни и симулационното моделиране са важни инструменти в областта на ИИ. И двете методологии имат своите приложения и са важни за развитието на ИИ. В зависимост от конкретната задача или проблем, които трябва

да се решат, е възможно да се използват и двете методологии, интегрирани в една система или приложение [14].

Таблица 5. Примери

	<b>Пример за моделиране на данни</b>	<b>Пример за симулационно моделиране</b>
Възможна задача	Прогнозиране на температурните промени в даден регион.	Изследване на разпространението на вирус в градската общност.
Метод	Използване на статистически модел за прогнозиране на температурния растеж въз основа на исторически данни за температурата, влажността, атмосферното налягане и други фактори.	Създаване на компютърен модел, симулиращ различни сценарии за контакт и разпространение на вируса между индивидите в общността, като взема предвид фактори като плътността на населението, мобилността, вероятността за инфекция и т.н.
Резултат	Моделът може да прогнозира температурните промени в региона в следващите дни или седмици, което е полезно за планиране на дейности и реакция на потенциални опасности, като топлинни вълни.	Симулационният модел може да предостави информация за потенциалната разпространеност на вируса в общността и да помогне за планиране на мерки за превенция и контрол на заразата.

### Заклучение

Сравнението между моделирането на данни и симулационното моделиране подчертава техните различни цели, методи и приложения. Моделирането на данни се фокусира върху анализа на налични данни с цел извличане на информация и прогнозиране на бъдещи събития, като използва различни статистически методи и алгоритми за машинно самообучение. То е насочено към разбиране и интерпретация на реални данни. Симулационното моделиране създава виртуални модели, които симулират динамиката на реалния свят, позволявайки изпълнение на експерименти и прогнозиране на резултати в различни сценарии. Подходът за моделиране на данни има предимство, тъй като анализът е възможен независимо от познанията на системата по отношение на дълбочината на анализа: може да се използва само при повърхностен анализ. Подходът за моделиране на симулация изисква знания, но увеличава дълбочината на анализа. Обобщено, моделирането на данни е по-широко приложимо и се използва активно в множество области, докато симулационното моделиране има по-специализирано приложение и се използва за създаване на виртуални модели за анализ на системи и процеси.

## References/Литература

1. **Velten, K., D. Schmidt, K. Kahlen.** Mathematical modeling and simulation: introduction for scientists and engineers. John Wiley & Sons, 2024.
2. **Carson, J. S.** Introduction to modeling and simulation. – In: *Proceedings of the Winter Simulation Conference*, 2005. IEEE, 2005, p. 8.
3. **Zeigler, B. P., H. Praehofer, T. G. Kim.** Theory of modeling and simulation. Academic press, 2000.
4. **Law, A. M., W. D. Kelton.** Simulation modeling and analysis. New York: Mcgraw-hill, 2007.
5. **Bossel, H.** Modeling and simulation. AK Peters/CRC Press, 2018.
6. **Sokolowski, J. A., C. M. Banks.** Modeling and simulation fundamentals: theoretical underpinnings and practical domains. John Wiley & Sons, 2010.
7. **Rossetti, M. D.** Simulation modeling and Arena. John Wiley & Sons, 2015.
8. **Thomopoylos, N. T.** Essentials of Monte Carlo simulation: Statistical methods for building simulation models. Springer Science & Business Media, 2012.
9. **Birta, L. G., G. Arbez.** Modelling and simulation. London: Springer, 2013.
10. **Sokolowski, J. A., C. M. Banks** (ed.). Principles of modeling and simulation: a multidisciplinary approach. John Wiley & Sons, 2011.
11. **Lukas, T. W., et al.** Changing the paradigm: Simulation, now a method of first resort. Naval Research Logistics (NRL), 2015, 62.4, pp. 293 – 303.
12. **Rubinstein, R. Y.** Optimization of computer simulation models with rare events. – In: *European Journal of Operational Research*, 1997, 99.1, pp. 89 – 112.
13. **De Paula Ferreira, W., et al.** Simulation in industry 4.0: A state-of-the-art review. – In: *Computers & Industrial Engineering*, 2020, 149: 106868.
14. **Choi S. H., S. J., T. G. Lee Kim.** Multi-fidelity modeling & simulation methodology for simulation speed up. – In: *Proceedings of the 2<sup>nd</sup> ACM SIGSIM/PADS conference on principles of advanced discrete simulation*, Denver, CO, USA, May 18 – 21, 2014, pp.139 – 150. New York: ACM.

## За авторите

**Светлана Сярова** е главен асистент в катедра „Компютърни науки“ в УниБИТ. През 2022 г. придобива докторска степен в УниБИТ по Автоматизирани системи за обработка и управление на информация. Нейните публикации и основни научни интереси са в областта на киберсигурността, информационните науки и науката за данните.

**За контакт с автора:** [s.syarova@unibit.bg](mailto:s.syarova@unibit.bg)

**Александър Кирков** е главен асистент в катедра „Компютърни науки“ в УниБИТ. Неговите публикации и основни научни интереси са в областта на киберсигурността, информационните науки и науката за данните.

**За контакт с автора:** [a.kirkov@unibit.bg](mailto:a.kirkov@unibit.bg)

**Самуел Петков** е студент трета година в специалност „Информационна сигурност“ в УниБИТ. Завършва средно образование през 2021 г. в 1. СУ „Пенчо

Славейков“, София, с профил английски език. Работил е като техническа поддръжка в УниБИТ. Интересува се от изкуствен интелект и киберсигурност.

**За контакт с автора:** 46732r@unibit.bg

## COMPARATIVE ANALYSIS OF DATA MODELING AND SIMULATION MODELING: METHODS, APPLICATIONS, AND CHALLENGES

**Svetlana Syarova, Samuel Petkov, Aleksandar Kirkov**  
*University of Library Studies and Information Technologies*

**Abstract:** In today's information age, data modeling and simulation modeling are shaping the future of science, industry, and society. These methodologies change the way information is created, processed and used.

This research paper examines the importance of data modeling and simulation modeling in the modern information context, as well as their applications and perspectives. Basic concepts and methods of data modeling and simulation modeling are reviewed. The key differences between them are highlighted, as well as their role in various fields of science and business.

**Keywords:** data modeling, simulation modeling.

### About the Authors

**Svetlana Syarova** is a Chief Assistant Professor in the Computer Science Department at the University of Library Studies and Information Technologies. In 2022, she obtained her PhD degree from ULSIT in Automated Systems for Information Processing and Management. Her publications and main research interests are in the field of Cybersecurity, Informing Science, and Data Science.

**To contact the Author:** s.syarova@unibit.bg

**Aleksandar Kirkov** is a Chief Assistant Professor in the Computer Science Department at the University of Library Studies and Information Technologies. His publications and main research interests are in the field of Cybersecurity, Informing Science, and Data Science.

**To contact the Author:** a.kirkov@unibit.bg

**Samuel Petkov** is a third-year bachelor's student specializing in information security at ULSIT. He graduated from secondary education in 2021 in 1 SU "Pencho Slaveykov" Sofia with English language profile. He worked as technical support at ULSIT. Interested in Artificial Intelligence and Cyber Security.

**To contact the Author:** 46732r@unibit.bg