



OPEN

DATA DESCRIPTOR

Greenhouse gas emissions from U.S. crude oil pipeline accidents: 1968 to 2020

Hongfang Lu^{1,9}, Zhao-Dong Xu¹✉, Kaihui Song^{2,9}, Y. Frank Cheng^{3,9}, Shaohua Dong⁴, Hongyuan Fang⁵, Haoyan Peng¹, Yun Fu⁴, Dongmin Xi¹, Zizhe Han¹, Xinmeng Jiang¹, Yao-Rong Dong⁶, Panpan Gai⁷, Zhiwei Shan¹ & Yuli Shan¹✉

Crude oil pipelines are considered as the lifelines of energy industry. However, accidents of the pipelines can lead to severe public health and environmental concerns, in which greenhouse gas (GHG) emissions, primarily methane, are frequently overlooked. While previous studies examined fugitive emissions in normal operation of crude oil pipelines, emissions resulting from accidents were typically managed separately and were therefore not included in the emission account of oil systems. To bridge this knowledge gap, we employed a bottom-up approach to conduct the first-ever inventory of GHG emissions resulting from crude oil pipeline accidents in the United States at the state level from 1968 to 2020, and leveraged Monte Carlo simulation to estimate the associated uncertainties. Our results reveal that GHG emissions from accidents in gathering pipelines (~720,000 tCO₂e) exceed those from transmission pipelines (~290,000 tCO₂e), although significantly more accidents have occurred in transmission pipelines (6883 cases) than gathering pipelines (773 cases). Texas accounted for over 40% of total accident-related GHG emissions nationwide. Our study contributes to enhanced accuracy of the GHG account associated with crude oil transport and implementing the data-driven climate mitigation strategies.

Background & Summary

The United States (U.S.) possesses an extensive network of crude oil pipelines, with the total length over 130,000 kilometers by the end of 2022^{1,2}. These pipelines, due to their immense scale and connectivity, play an essential role in facilitating energy supplies and economic growth of the country³. However, occurrence of accidents within this intricate pipeline system can pose significant environmental challenges, particularly cause release of substantial amount of greenhouse gases (GHGs)^{4,5}. Methane is the primary GHG released from crude oil pipelines when accidents occur. It has 28–36 times of global warming potential (GWP) compared to CO₂ for 100-year time span, and 84–87 times of GWP over a 20-year timeframe (GWP₂₀)⁶. Despite high GWP, these unintended emissions are usually neglected in official estimates of national GHG emissions from bottom-up approaches^{7,8}, as emissions resulting from accidents are typically managed separately and their impacts on climate change are usually ignored when public health risks present. The accidents, although occurring occasionally, can release tons of GHG emissions⁹, undermining the global climate mitigation efforts if not properly addressed^{10,11}. Quantifying emissions resulting from large-scale crude oil pipeline accidents is therefore of utmost importance, as it serves as a fundamental component for better understanding the broader implications

¹China-Pakistan Belt and Road Joint Laboratory on Smart Disaster Prevention of Major Infrastructures, Southeast University, Nanjing, 210096, China. ²Data-Driven EnviroLab, School of Public Policy, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA. ³Department of Mechanical Engineering, University of Calgary, Calgary, Alberta, T2N 1N4, Canada. ⁴School of Safety and Ocean Engineering, China University of Petroleum (Beijing), Beijing, 102249, China. ⁵Yellow River Laboratory, Zhengzhou University, Zhengzhou, 450001, China. ⁶School of Civil Engineering, Xi'an University of Architecture and Technology, Xi'an, 710055, China. ⁷School of Civil Engineering and Mechanics, Jiangsu University, Zhenjiang, 212013, China. ⁸School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, B15 2TT, UK. ⁹These authors contributed equally: Hongfang Lu, Kaihui Song, Y. Frank Cheng. ✉e-mail: zhdxu@163.com; y.shan@bham.ac.uk

of climate change mitigation efforts and developing robust response strategies to minimize the adverse impacts of these accidents for both researchers and policymakers^{12–14}.

Previous studies pertaining to quantifying upstream GHG emissions from crude oil have predominantly utilized lifecycle assessment (LCA) and centred on major procedures, including production, transportation, and refining processes^{15–19}. These “well-to-wheels” life-cycle GHG emissions from petroleum amount to ~1.7 gigatons of CO₂ equivalent (GtCO₂e), representing ~5% of total fuel combustion emissions¹⁵. In the refining process, well-to-refinery carbon intensity ranges from ~1.5 to 46.9 gCO₂e/MJ¹⁶. The carbon intensity range of refining is 13.9–62.1 kg CO₂ equivalent per barrel¹⁷. When studying the GHG emissions from each process, only macroscopic factors (such as distance travelled and throughput) under normal pipeline operations are carefully studied. As a common practice, studies employing LCA to analyze carbon intensity of crude oil production often ignore impacts that are difficult to measure or less well understood¹⁵. Despite the existing research on carbon emissions in the oil supply chain, the carbon emissions resulting from accidents have not been fully defined. It should be noted that fugitive emissions are fundamentally different from emissions caused by accidents. Fugitive emissions only consider leaks that occur during normal operations and do not include emissions resulting from unexpected accidents^{20,21}. For instance, valve leaks during the regular operation of pipelines would be classified as fugitive emissions, whereas the explosion of the Nord Stream pipeline in September 2022 was classified as accident-related emissions⁹.

Recent studies have recognized the necessity of considering natural gas leaks from oil and gas pipelines in city and national climate actions^{11,22}. However, these studies mainly rely on methane monitoring devices, and the process is labor-intensive and usually occurs in individual observation points. The recent movement of employing satellites to remote sense the methane content²³ offers a top-down approach to estimate and address the potential underestimation from traditional bottom-up approaches; however, it usually cannot identify the detailed sources of methane emissions and therefore becomes limited in developing effective strategies for policymakers. While understanding the implications for conducting climate change-focused integrity management is of paramount importance, the GHG emissions resulting from crude oil pipeline accidents have been largely unexplored in previous studies.

To fill this gap, this study quantifies the GHG emissions resulting from crude oil pipeline accidents in the U.S. from 1968 to 2020. The accidents include both gathering and transmission pipeline systems, considering the combustion/explosion conditions associated with the accidents. In comparison to the accident dataset previously provided by the Pipeline and Hazardous Materials Safety Administration (PHMSA), the dataset proposed in this study (1) quantifies previously unaccounted GHG emissions in accident records, and (2) incorporates uncertainties involved in GHG emission assessment. Consequently, this study develops an accurate inventory of GHG emissions resulting from crude oil pipeline accidents. This work contributes to rectifying the omissions within the official audits of the oil system, thereby enhancing the accuracy of GHG emissions accounting. Furthermore, by attaining the GHG emissions associated with each crude oil pipeline accident, operators can utilize this information to conduct more profound climate change-oriented risk assessments across various regions, design parameters, and operational conditions of pipelines. This, in turn, facilitates the formulation of pragmatic and actionable measures to mitigate the occurrence of accidents and potentially offers guidance for future pipeline designs. Additionally, the proposed inventory methodology may serve as a reference for investigating GHG emissions in similar accidents.

This study yields three distinct datasets, including GHG emissions resulting from oil pipeline accidents assessed by year (1968 to 2020), GHG emissions resulting from oil pipeline accidents assessed by state, and GHG emissions from single-point accidents. The specific descriptions of the three datasets are presented in the section entitled “Data Records”. Each dataset carries significant implications: (1) The accurate assessment of GHG emissions resulting from crude oil pipeline accidents enables decision-makers to acquire a more precise understanding of the consequential impact on climate change. Consequently, this comprehensive knowledge fosters informed decision-making in the realm of subnational climate actions and policy formulation. (2) The datasets we provide will serve as a valuable resource for pipeline operators, providing them enhanced insight into the potential GHG emission risks arising from accidents. By making them access to the valuable information, operators are able to develop proactive measures and effective counterstrategies, thereby bolstering their social standing and reputation¹⁹.

Methods

Data collection. We collected data required to quantify GHG emissions resulting from crude oil pipeline accidents, such as the volume of oil spills and accident conditions. The PHMSA has documented data of pipeline accidents involving hazardous liquid in the U.S. since 1968. Hazardous liquids include biofuels, carbon dioxide, crude oil, highly volatile liquids, and refined petroleum products²⁴. The PHMSA requires all operators to report the occurrence of an accident within 30 days in accordance with Title 49 of the Code of Federal Regulations (49 CFR Parts 191, 195)^{25,26} (see Supplementary Note 1 for the definition of crude oil pipeline accidents). We extracted accident information from all crude oil pipelines, including gathering pipelines and transmission pipelines. However, various changes and updates in the PHMSA’s data management pose challenges to time-series analysis. Different levels of details are documented in four distinct time periods: 1968 to 1985, 1986 to 2001, 2002 to 2009, and 2010 to the present, with the most comprehensive information available for the period from 2010 to the present. Table 1 presents detailed information on accidents contained in each period of the dataset.

Regardless of the period, the PHMSA provides detailed information about when and where accidents happened and the volume of crude oil leaked. However, it is still insufficient for assessing the impact of the accident on climate change. First, while the PHMSA recorded the volume of oil spills for each accident, it did not provide information regarding GHG emissions associated with each incident. Second, crucial information on

Category	1968 to 1985	1986 to 2001	2002 to 2009	2010 to the present
Operator information	Included	Included	Included	Included
Accident information	Included	Included	Included	Included
Release information	Included	Included	Included	Included
Consequence information	Included	Included	Included	Included
Pipeline design information	Included	Included	Included	Included
Pipeline maintenance information	Included but limited	Included but limited	Included	Included
Accident cause	Included	Included	Included	Included
Drug and alcohol testing information	Not included	Not included	Included	Included
Number of items	61	63	259	654
Number of accidents recorded	2488 (transmission pipelines); 276 (gathering pipelines)	1093 (transmission pipelines); 234 (gathering pipelines)	1270 (transmission pipelines); 105 (gathering pipelines)	Continuously updated (2032 accidents as of 2020 for transmission pipelines, 158 accidents as of 2020 for gathering pipelines)

Table 1. Information on crude oil pipeline accidents contained in the PHMSA dataset.

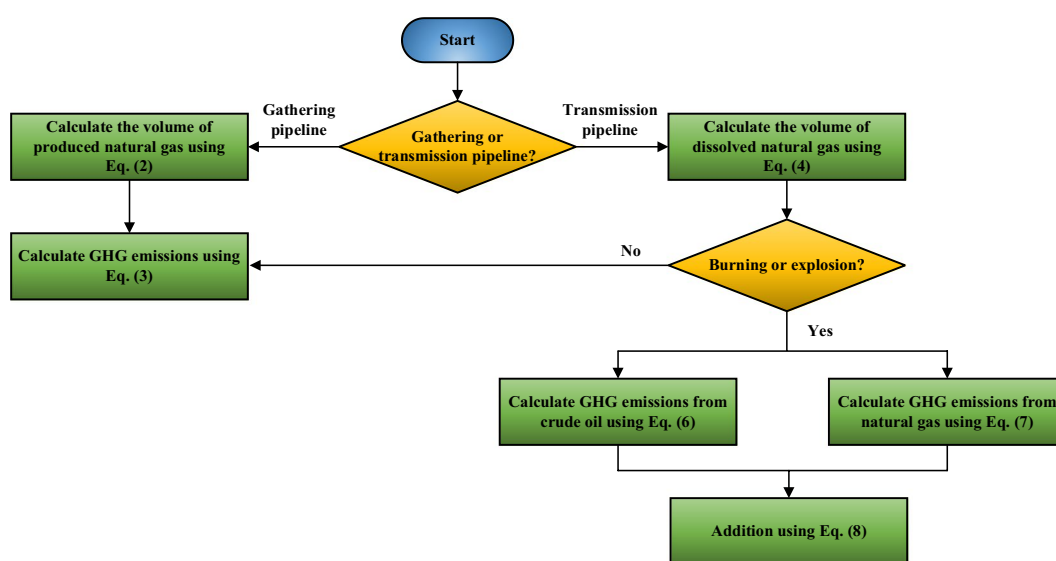


Fig. 1 Flow chart for calculation of GHG emissions from crude oil pipeline accidents.

oil recovery for transmission pipelines is missing from 1968 to 1985, which is essential for accounting for GHG emissions in both spill events and emergent responses.

Calculating GHG emissions from single-point accidents. Leaks in oil pipelines release GHGs (methane). However, the GHG emissions associated with oil spills vary depending on oil products being transported in pipelines and accident conditions (such as non-combustion, combustion, and explosion). We designed a workflow to illustrate how the GHG emissions for single-point accidents were calculated, as shown in Fig. 1. We further analyzed GHG emissions from crude oil pipelines in two categories according to their functions^{27,28}, namely gathering pipelines and transmission pipelines (details about the two types of pipelines are shown in Supplementary Table 1).

In gathering pipelines, a considerable amount of natural gas is often present as a byproduct during the process of crude oil extraction at the wellhead. It is imperative to calculate the GHG emissions resulting from pipeline accidents by considering the production gas-oil ratio (PGOR) and conditions whether combustion/explosion occurs. PGOR refers to the ratio of natural gas to crude oil produced from oil wells in the process of oil field exploitation²⁹. It can be derived from historical production data in the U.S., as shown in Eq. (1).

$$PGOR = \frac{p_g}{p_{co}} \quad (1)$$

where, PGOR is production gas-oil ratio, cubic feet per barrel (ft³/bbl); p_g is production rate of natural gas, ft³/day; and p_{co} is production rate of crude oil, bbls/day. The volume of natural gas released can be deduced by applying Eq. (2).

$$V_{gas,gathering} = V_{crude\ oil,gathering} \times PGOR \quad (2)$$

where $V_{gas,gathering}$ is the volume of natural gas released from gathering pipelines, ft³; and $V_{crude\ oil,gathering}$ is the volume of oil spills from gathering pipelines, bbls.

From accident data recorded by the PHMSA spanning from 1968 to 2020, it was found that gathering pipelines have not experienced combustion/explosion accidents. This suggests that the GHG emissions from gathering pipeline accidents are solely due to the release of natural gas resulting from crude oil spills. If combustion or explosion is not involved in the accident, the GHG emissions from the release of natural gas can be calculated by Eq. (3)³⁰.

$$GHG_{gas,unburned} = 2.8 \times 10^{-5} \times GWP_{CH_4} \times \rho_{gas} \times V_{gas,gathering} \quad (3)$$

where $GHG_{gas,unburned}$ is GHG emissions from the release of unburned natural gas, metric tons of carbon dioxide equivalent (tCO₂e); ρ_{gas} is natural gas density, kg/m³; and GWP_{CH_4} is the global warming potential of methane, $GWP_{CH_4} = 27.9$ under the 100-year timeframe from IPCC the Sixth Assessment Report (AR6) from Working Group I (WRI)³¹.

Removing natural gas from the oil-gas fluid during the processing stage yields higher-quality crude oil, which is transported to refineries via transmission pipelines. However, natural gas could still present in transmission pipelines, albeit in small quantities³². In the case of pipeline accidents, pipeline pressure rapidly decreases, leading to escape of a limited amount of dissolved natural gas from the crude oil. The quantity of natural gas emissions can be calculated by considering the dissolved gas-oil ratio (DGOR)³², as shown by Eq. (4). The DGOR refers to the ratio between the volume of dissolved natural gas in crude oil under a specific pressure condition, as described by Eq. (5).

$$V_{gas,transmission} = V_{crude\ oil,transmission} \times DGOR \quad (4)$$

where V_{gt} is release volume of natural gas from transmission pipelines, ft³; V_{co} is oil spills from transmission pipelines, bbls; and DGOR is dissolved gas-oil ratio, ft³/bbl.

$$DGOR = \gamma_{gas} \left[\frac{P}{18} \times \frac{10^{0.0125API}}{10^{0.00091T}} \right]^{1.2048} \quad (5)$$

where γ_{gas} is gas-specific gravity; API is oil-API gravity; P is absolute pressure, psi; and T is temperature, °F. P is determined according to the accident pressure provided in the PHMSA database. T is uniformly taken as 60°F³². API values vary widely across the U.S.

According to data provided by PHMSA, transmission pipelines are susceptible to combustion or explosion accidents. If there is combustion or explosion in the crude oil transmission pipeline accident, the GHG emissions caused by the crude oil can be calculated according to Eq. (6).

$$GHG_{crude\ oil,burned} = 0.43 \times V_{crude\ oil,transmission} \times f_u \quad (6)$$

where $GHG_{crude\ oil,burned}$ is GHG emissions produced by the accident (with combustion or explosion) for crude oil transmission pipeline, tCO₂e; $V_{crude\ oil,transmission}$ is the volume of crude oil spilled from the transmission pipeline, bbls; and f_u is the unrecovered fraction following crude oil leakage, percent. Note that after crude oil spills, the escaped natural gas is proportional to the amount of oil spilled. However, a portion of the crude oil can be recovered through various methods^{33–39}. If combustion or explosion occurs during the accident, the recovered crude oil is not involved in the combustion accident and does not contribute to GHG emissions. Moreover, in the scenario where transmission pipelines are involved in combustion or explosion accidents, the GHG emissions generated from the combustion of the released natural gas can be computed utilizing Eq. (7)³⁰.

$$GHG_{gas,burned} = 5.48 \times 10^{-5} \times V_{gas,transmission,burned} \times f_c \times GWP_{CO_2} \quad (7)$$

where $GHG_{gas,burned}$ is GHG emissions from natural gas combustion or explosion, tCO₂e; $V_{gas,transmission,burned}$ is the volume of natural gas that burns or explodes from transmission pipelines, ft³; f_c is the fraction of natural gas oxidized into carbon dioxide⁴⁰; and GWP_{CO_2} is the global warming potential of carbon dioxide, $GWP_{CO_2} = 1$ under the 100-year timeframe from IPCC AR6 WGI report. Therefore, the total GHG emissions resulting from a transmission pipeline accident involving combustion or explosion are:

$$GHG_{total,burned} = GHG_{gas,burned} + GHG_{crudeoil,burned} \quad (8)$$

where $GHG_{total,burned}$ is the total GHG emissions from transmission pipeline accident (with combustion or explosion).

Monte Carlo simulation for GHG emission inventory. Monte Carlo simulation constitutes a numerical methodology grounded in stochastic sampling and statistical analysis, finding extensive utility in addressing the intricacies of uncertainty analysis within complex systems. Within the context of GHG emission estimation, the

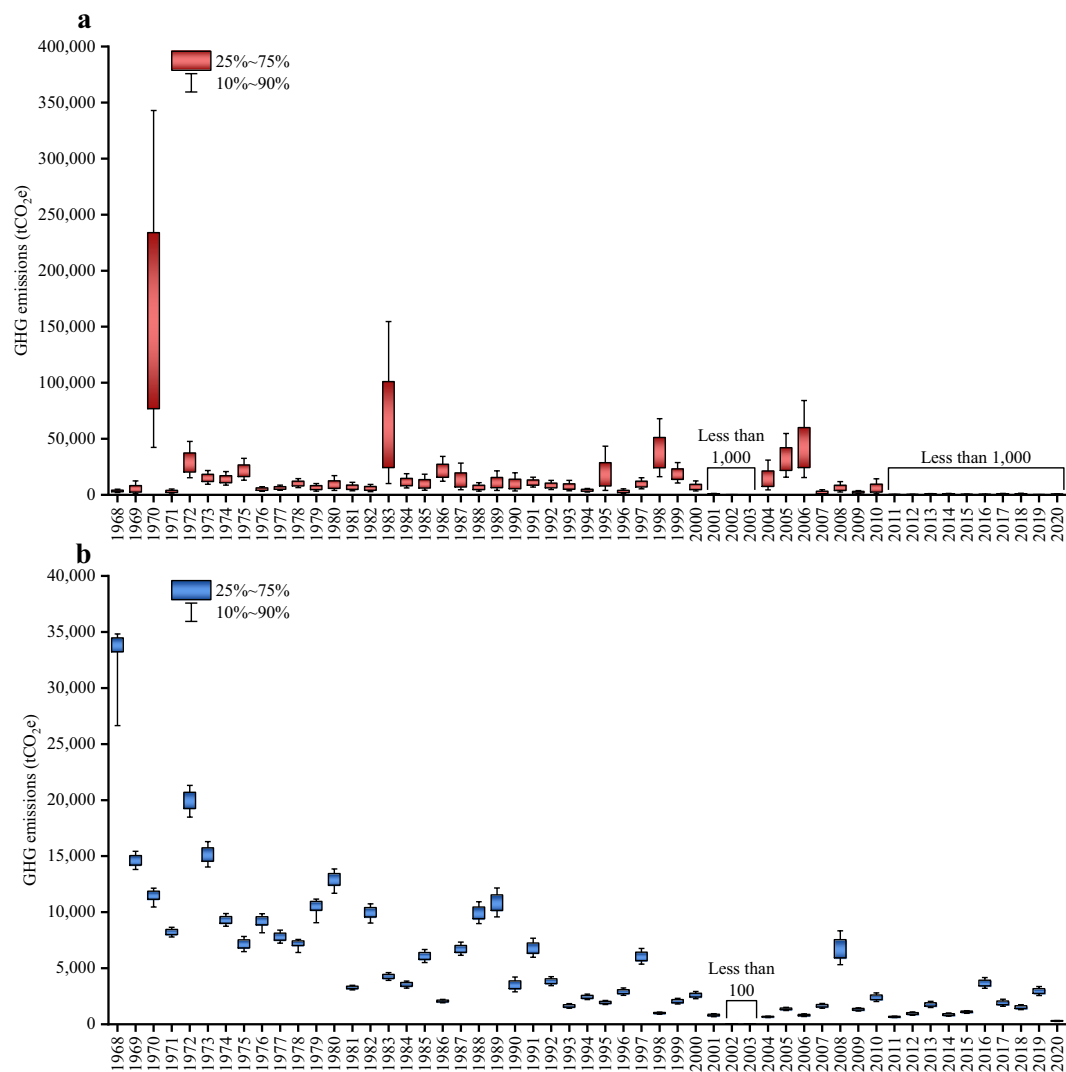


Fig. 2 Annual GHG emissions caused by crude oil pipeline accidents from 1968 to 2020. **(a)** Gathering pipelines. **(b)** Transmission pipelines. Note that the annotations “less than 1,000” and “less than 100” in subfigures a and b refer to the average values derived from Monte Carlo simulations. Specific data simulated by Monte Carlo are available at Figshare⁴³.

salient advantage of Monte Carlo simulation resides in its capacity to adeptly capture interdependencies among multiple uncertain parameters, thereby affording a comprehensive and precise assessment of uncertainties. By iteratively sampling from the probability distributions of each parameter, Monte Carlo simulation engenders the generation of a profusion of potential scenarios, thereby encompassing a spectrum of prospective outcomes inherent to the emissions process^{41,42}. In this study, the GHG emissions from gathering and transmission pipeline accidents are investigated separately. For the gathering pipeline system, a total of 773 accidents from 1968 to 2020 are recorded, and the spill volumes of these accidents can be obtained from PHMSA. Based on the probability density function (PDF) of PGOR, 773 data points are generated to correspond to the PGOR values of these accidents, and the gas volume is calculated based on Eq. (2). Additionally, when estimating GHG emissions according to Eq. (3), the density of natural gas is randomly selected between 0.66 and 1.05 kg/m³. The entire process is repeated 200,000 times to obtain the uncertainty of GHG emission inventory for gathering pipeline accidents.

From 1968 to 2020, a total of 6,883 transmission pipeline accidents were recorded. When estimating the dissolved natural gas associated with these accidents based on Eqs. (4) and (5), the values of γ_{gas} and API are uncertain. γ_{gas} is randomly selected between 0.55 and 0.87, and a PDF of [23 + WEIB(10.8, 6.35)] is used to generate 6,883 data points corresponding to the API values of these accidents. Additionally, the involvement of combustion and explosion in accidents needs to be considered. If combustion or explosion occurs, the dissolved natural gas and GHG emissions from the burned crude oil are estimated according to Eq. (7) and Eq. (6) respectively. It is important to note that, of the 222 accidents involving combustion or explosion, only 97 accidents record volumes of recovered crude oil, while the volumes of burned crude oil remain unknown for the remaining 125 accidents. Therefore, a PDF of [−0.001 + EXPO(0.417)] fitted based on the unrecovered ratio from these 97 accidents is used to generate 125 data points corresponding to the f_u values of these 125 accidents. The entire

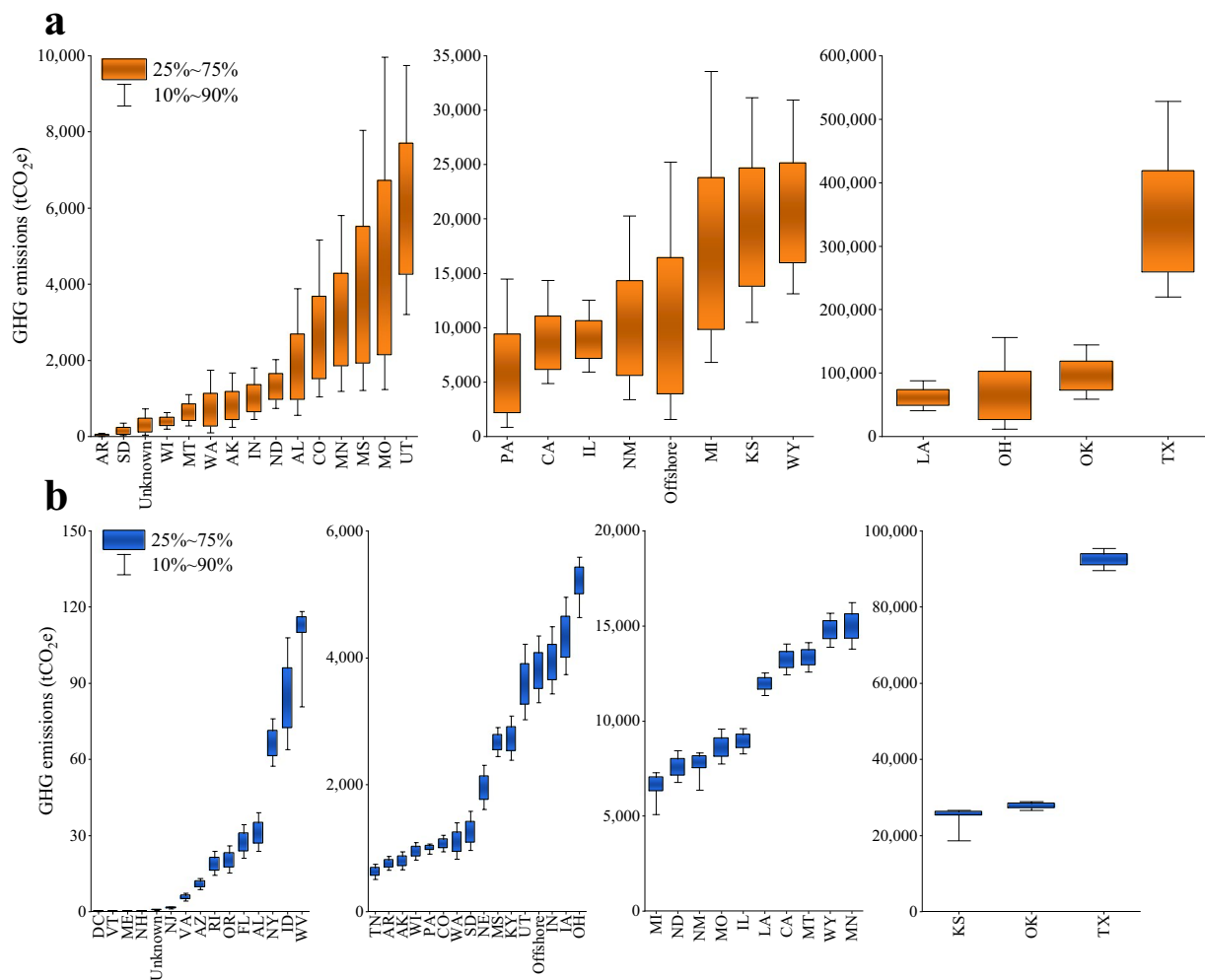


Fig. 3 GHG emissions caused by crude oil pipeline accidents in various states. **(a)** Gathering pipelines. **(b)** Transmission pipelines. Specific data simulated by Monte Carlo are available at Figshare⁴³. The heat map of GHG emissions is presented in Supplementary Fig. 1. “Unknown” indicates that the PHMSA dataset does not specify the state where the incident occurred. The full names of the state abbreviations are shown in ref. ⁴⁶.

process is repeated 200,000 times to estimate the uncertainty of GHG emission inventory for transmission pipeline accidents. Furthermore, the data utilized in the Monte Carlo simulation are sourced from Supplementary Table 2, as indicated.

Data Records

Three datasets were generated in this study:

- 1) GHG emissions resulting from gathering and transmission pipeline accidents annually from 1968 to 2020, as presented in Figshare⁴³.
- 2) GHG emissions resulting from gathering and transmission pipeline accidents in different states during the period from 1968 to 2020, as shown in Figshare⁴³. The gathering pipeline dataset covers a total of 27 states, while the transmission pipeline dataset includes 44 states.
- 3) GHG emissions from a total of 7,656 individual accidents that occurred from 1968 to 2020, as provided in Figshare⁴³.

It is important to note that the first two datasets display the mean values and uncertainties obtained from 200,000 Monte Carlo simulations, whereas the third dataset presents the mean value along with the upper and lower bounds of the 90% confidence interval derived from the same 200,000 Monte Carlo simulations.

Figure 2 illustrates that gathering pipelines were associated with the highest GHG emissions in the year 1970, amounting to $(170,119 \pm 123,398)$ tCO₂e. The subsequent years with notable emissions were 1983 and 2006, with values of $(70,491 \pm 59,660)$ and $(45,179 \pm 27,998)$ tCO₂e, respectively. Additionally, the GHG emissions (average values from Monte Carlo simulations) for the years 2001 to 2003 and 2011 to 2020 remained below 1,000 tCO₂e. For transmission pipelines, the year 1968 exhibited the highest GHG emissions, amounting to

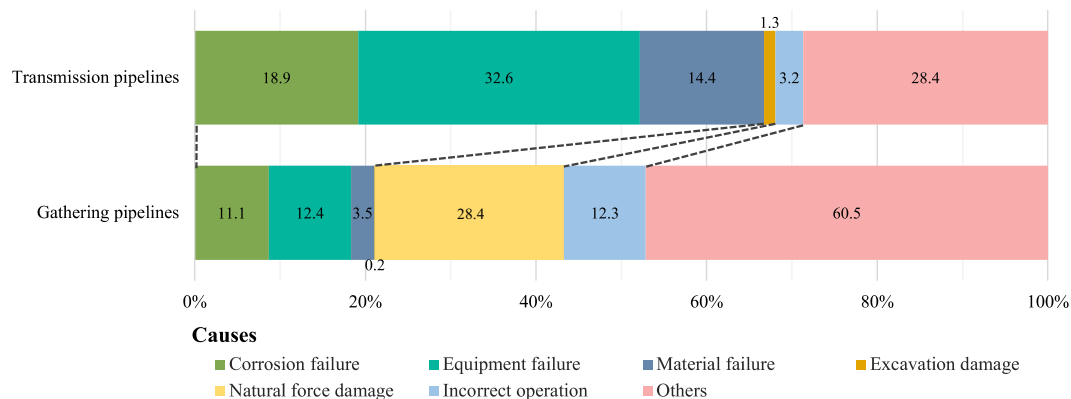


Fig. 4 Percentage of GHG emissions from various causes. (a) Gathering pipelines. (b) Transmission pipelines. The proportions are calculated based on the average values obtained from Monte Carlo simulations. The category “Others” mainly includes the following ten specific reasons: (1) destruction caused by boats, barges, drilling rigs, or other maritime equipment or vessels that are set adrift or have lost their mooring; (2) damage caused by cars, trucks, or other motorized vehicles or equipment not involved in excavation; (3) electrical arcing originating from other equipment or facilities; (4) intentional acts of damage; (5) accidents categorized as miscellaneous; (6) occurrence of nearby industrial, man-made, or other fire/explosion as the primary cause of the accident; (7) damage resulting from other external forces; (8) pre-existing mechanical damage unrelated to excavation; (9) accidents arising from routine or normal fishing or other maritime activities not related to excavation; (10) occurrences with unknown causes.

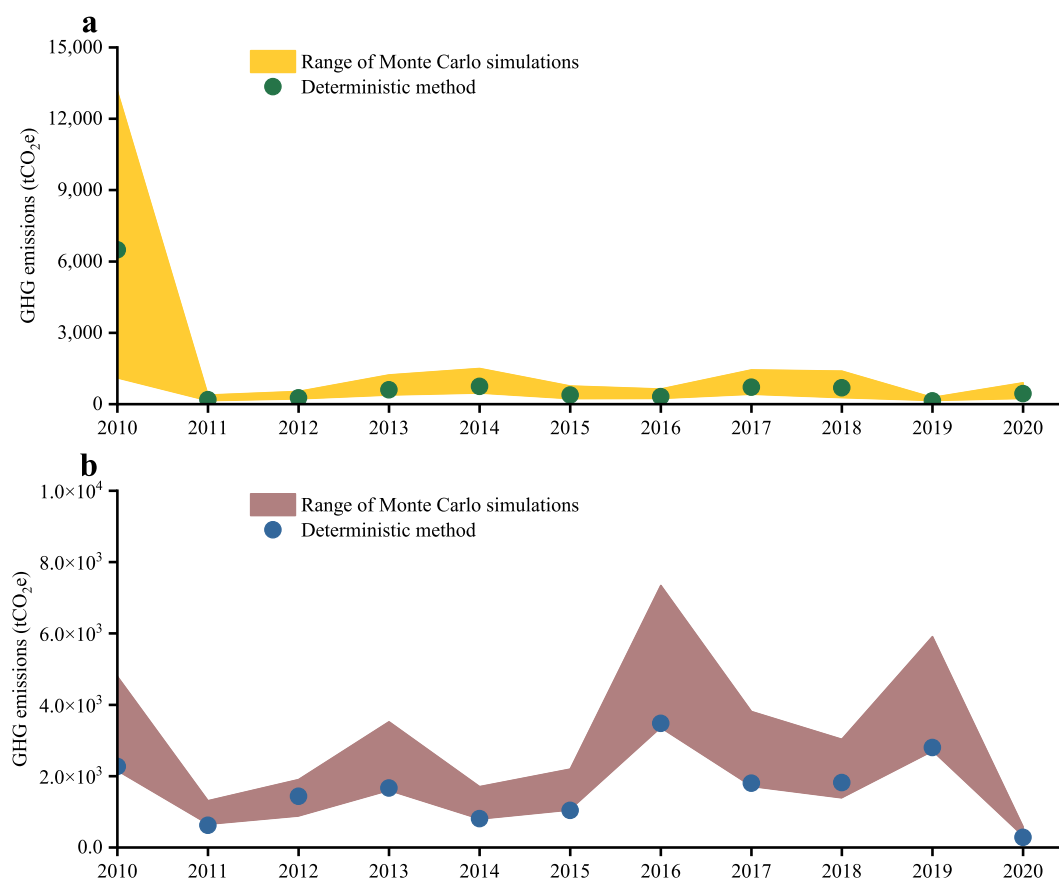


Fig. 5 Results comparison between Monte Carlo simulation and deterministic methods. (a) Gathering pipelines. (b) Transmission pipelines.

(32,464 ± 4,252) tCO₂e. This was followed by the years 1972 and 1969, with emissions of (19,928 ± 1,128) and (14,610 ± 636) tCO₂e, respectively. In comparison between the gathering and transmission systems, although the number of accidents in the gathering pipelines (773) is significantly lower than that in the transmission

pipelines (6,883), the GHG emissions resulting from accidents (~720,000 tCO₂e) are higher in the gathering pipelines than in the transmission pipelines (~290,000 tCO₂e).

Figure 3 highlights that, in the gathering pipeline system, Texas, Oklahoma, and Ohio have the highest GHG emissions. Texas, in particular, has the GHG emissions up to (352,675 ± 126,164) tCO₂e, accounting for nearly half of the emissions within the gathering pipeline system. In the transmission pipeline system, Texas again has the highest GHG emissions, amounting to (92,496 ± 2,247) tCO₂e. Following Texas, Oklahoma and Kansas rank the second and the third, with emissions of (27,795 ± 937) and (24,298 ± 4,243) tCO₂e, respectively.

The PHMSA categorizes the causes of pipeline accidents as corrosion, equipment failure, material failure, excavation damage, incorrect operation, natural force damage, and others^{44,45}. Figure 4 shows that equipment failure and incorrect operation are the primary factors contributing to GHG emissions from gathering pipeline accidents, excluding the “others” category. In the transmission pipeline system, the equipment failure is the primary contributor to GHG emissions, accounting for 32.6% of the total.

Technical Validation

Uncertainties. For gathering pipelines, uncertainties in estimating GHG emissions from accidents primarily stem from the values of PGOR and natural gas density. In transmission pipelines, uncertainties arise from DGOR, natural gas density, API of crude oil, the combustion ratio of crude oil in combustion accidents, and the conversion ratio of natural gas. We employed Monte Carlo simulations (n = 200,000 times) to obtain the uncertainties using standard deviation. Ref. ⁴³ presents the uncertainties in GHG emissions inventory on an annual basis, indicating that the uncertainty range for gathering pipelines varies from ± 28.9% to ± 84.6%, while for transmission pipelines, it ranges from ± 4.1% to ± 16.8%. The uncertainties in GHG emissions inventory for gathering pipelines in different states, ranging from ± 30.1% to ± 86.5%. The uncertainties in GHG emissions inventory for transmission pipelines in various states, ranging from ± 2.4% to ± 19.7%.

Validation of the estimation method. To validate the reliability of the proposed methodology and generated datasets, a deterministic approach was employed to estimate GHG emissions resulting from accidents. The results were then compared with those obtained through Monte Carlo simulations. In the deterministic approach, all parameters were set as fixed values, as indicated in Supplementary Table 3. The years 2010 to 2020 were selected for validation purposes since PHMSA provides more detailed records of accidents during this period. Furthermore, Eq. (9) was employed to assess the deviation between the deterministic method and the average values obtained from Monte Carlo simulations.

$$E = \frac{V_D - V_M}{V_D} \quad (9)$$

where E is the deviation; V_D is GHG emissions from the deterministic methods, tCO₂e; and V_M is the average GHG emissions from Monte Carlo simulations, tCO₂e.

Figure 5 and Supplementary Tables 4, 5 indicate that the GHG emission estimates obtained through deterministic methods fall within the upper and lower bounds of the Monte Carlo simulation. The estimated biases for gathering pipelines range from −1.37% to −0.90%, while the estimated biases for the transmission pipeline range from −5.64% to 33.71%. It demonstrates the high reliability of the proposed GHG emission estimation method and the generated datasets in this study. Furthermore, after 2010, there has been a decrease in GHG emissions resulting from crude oil pipeline accidents. It is primarily attributed to an improved response speed to accidents and implementation of effective emergency measures and rapid repairs⁴⁴, leading to reduced duration and smaller scale of leaks after accidents, and thus, lower GHG emissions.

Code availability

The code utilized for the Monte Carlo simulations in this study is provided in the Supplementary Codes 1 and 2.

Received: 19 June 2023; Accepted: 15 August 2023;

Published online: 24 August 2023

References

1. U.S. Department of Transportation. Annual Report Mileage for Hazardous Liquid or Carbon Dioxide Systems. *Pipeline and Hazardous Materials Safety Administration (PHMSA)* <https://www.phmsa.dot.gov/data-and-statistics/pipeline/annual-report-mileage-hazardous-liquid-or-carbon-dioxide-systems> (2023).
2. Lu, H., Xu, Z. D., Iseley, T., Peng, H., & Fu, L. *Pipeline Inspection and Health Monitoring Technology: The Key to Integrity Management* 1st edn (Springer Nature Press, 2023).
3. Erickson, P. & Lazarus, M. Impact of the Keystone XL pipeline on global oil markets and greenhouse gas emissions. *Nat. Clim. Change* **4**, 778–781 (2014).
4. Siler-Evans, K., Hanson, A., Sunday, C., Leonard, N. & Tumminello, M. Analysis of pipeline accidents in the United States from 1968 to 2009. *Int. J. Crit. Infra. Prot.* **7**, 257–269 (2014).
5. Lu, H. *et al.* An inventory of greenhouse gas emissions due to natural gas pipeline incidents in the United States and Canada from 1980s to 2021. *Sci. Data* **10**, 282 (2023).
6. Intergovernmental Panel on Climate Change. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). https://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm (2007).
7. Guan, D., Liu, Z., Geng, Y., Lindner, S. & Hubacek, K. The gigatonne gap in China's carbon dioxide inventories. *Nat. Clim. Change* **2**, 672–675 (2012).
8. Alvarez, R. A. *et al.* Assessment of methane emissions from the US oil and gas supply chain. *Science* **361**, 186–188 (2018).

9. Jia, M. *et al.* The Nord Stream pipeline gas leaks released approximately 220,000 tonnes of methane into the atmosphere. *Environ. Sci. Ecotechnology* **12**, 100210 (2022).
10. Saint-Vincent, P. M. & Pekney, N. J. Beyond-the-meter: Unaccounted sources of methane emissions in the natural gas distribution sector. *Environ. Sci. Technol.* **54**, 39–49 (2019).
11. Sargent, M. R. *et al.* Majority of US urban natural gas emissions unaccounted for in inventories. *P. Natl. Acad. Sci. USA* **118**, e2105804118 (2021).
12. Kemfert, C., Präger, F., Braunger, I., Hoffart, F. M. & Brauers, H. The expansion of natural gas infrastructure puts energy transitions at risk. *Nat. Energy* **7**, 582–587 (2022).
13. Gan, Y. *et al.* Carbon footprint of global natural gas supplies to China. *Nat. Commun.* **11**, 824 (2020).
14. Zhu, Y., Qian, X. M., Liu, Z. Y., Huang, P. & Yuan, M. Q. Analysis and assessment of the Qingdao crude oil vapor explosion accident: Lessons learnt. *J. Loss Prevent. Proc.* **33**, 289–303 (2015).
15. Masnadi, M. S. *et al.* Global carbon intensity of crude oil production. *Science* **361**, 851–853 (2018).
16. Masnadi, M. S. *et al.* Well-to-refinery emissions and net-energy analysis of China's crude-oil supply. *Nat. Energy* **3**, 220–226 (2018).
17. Jing, L. *et al.* Carbon intensity of global crude oil refining and mitigation potential. *Nat. Clim. Change* **10**, 526–532 (2020).
18. Greene, S., Jia, H. & Rubio-Domingo, G. Well-to-tank carbon emissions from crude oil maritime transportation. *Transport Res. D-Tr. E.* **88**, 102587 (2020).
19. Huang, L., Liao, Q., Yan, J., Liang, Y. & Zhang, H. Carbon footprint of oil products pipeline transportation. *Sci. Total Environ.* **783**, 146906 (2021).
20. Hendrick, M. F., Ackley, R., Sanaie-Movahed, B., Tang, X. & Phillips, N. G. Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments. *Environ. Pollut.* **213**, 710–716 (2016).
21. Joye, S. & MacDonald, I. Offshore oceanic impacts from the BP oil spill. *Nat. Geosci.* **3**, 446–446 (2010).
22. Song, K. *et al.* Can US multi-state climate mitigation agreements work? A perspective from embedded emission flows. *Global Environ. Chang.* **77**, 102596 (2022).
23. Saunio, M. *et al.* The global methane budget 2000–2017. *Earth Syst. Sci. Data* **12**, 1561–1623 (2020).
24. Pipeline and Hazardous Materials Safety Administration (PHMSA). Distribution, Transmission & Gathering, LNG, and Liquid Accident and Incident Data. <https://www.phmsa.dot.gov/data-and-statistics/pipeline/distribution-transmission-gathering-lng-and-liquid-accident-and-incident-data#:~:text=Incident%20reporting%20requirements%20%20%20Pipeline%20System%20Type,CFR%20191.3%20%202049%20CFR%20191.15> (2023).
25. US Government Publishing Office. Electronic Code of Federal Regulations. <https://www.ecfr.gov> (2023).
26. Lam, C. & Zhou, W. Statistical analyses of incidents on onshore gas transmission pipelines based on PHMSA database. *Int. J. Pres. Ves. Pip.* **145**, 29–40 (2016).
27. Biezma, M. V., Andrés, M. A., Agudo, D. & Briz, E. Most fatal oil & gas pipeline accidents through history: A lessons learned approach. *Eng. Fail. Anal.* **110**, 104446 (2020).
28. Lu, H., Xi, D. & Qin, G. Environmental risk of oil pipeline accidents. *Sci. Total Environ.* **874**, 162386 (2023).
29. Jones, R. S. Jr Producing-gas/oil-ratio behavior of multifractured horizontal wells in tight oil reservoirs. *SPE Reserv. Eval. Eng.* **20**, 589–601 (2017).
30. Intergovernmental Panel on Climate Change (IPCC). *2006 IPCC Guidelines for National Greenhouse Gas Inventories* 1st edn (Institute for Global Environmental Strategies, 2006).
31. Forster, P. *et al.* The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V. *et al.* (eds.)] Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 923–1054, <https://doi.org/10.1017/9781009157896.009> (2021).
32. Price, L. C., Wenger, L. M., Ging, T. & Blount, C. W. Solubility of crude oil in methane as a function of pressure and temperature. *Org. Geochem.* **4**, 201–221 (1983).
33. Ge, J. *et al.* Joule-heated graphene-wrapped sponge enables fast clean-up of viscous crude-oil spill. *Nat. Nanotechnol.* **12**, 434–440 (2017).
34. Bullock, R. J., Perkins, R. A. & Aggarwal, S. *In-situ* burning with chemical herders for Arctic oil spill response: Meta-analysis and review. *Sci. Total Environ.* **675**, 705–716 (2019).
35. Fritt-Rasmussen, J., Wegeberg, S. & Gustavson, K. Review on burn residues from *in situ* burning of oil spills in relation to Arctic waters. *Water Air Soil Poll.* **226**, 1–12 (2015).
36. Lei, W., Portehault, D., Liu, D., Qin, S. & Chen, Y. Porous boron nitride nanosheets for effective water cleaning. *Nat. Commun.* **4**, 1–7 (2013).
37. Silvani, L. *et al.* Characterizing biochar as alternative sorbent for oil spill remediation. *Sci. Rep.-UK* **7**, 1–10 (2017).
38. Wu, Z. Y. *et al.* Carbon nanofiber aerogels for emergent cleanup of oil spillage and chemical leakage under harsh conditions. *Sci. Rep.-UK* **4**, 1–6 (2014).
39. Reed, M. *et al.* Numerical model for estimation of pipeline oil spill volumes. *Environ. Modell. Softw.* **21**, 178–189 (2006).
40. Plant, G. *et al.* Inefficient and unlit natural gas flares both emit large quantities of methane. *Science* **377**, 1566–1571 (2022).
41. Louwen, A., Van Sark, W. G., Faaij, A. P. & Schropp, R. E. Re-assessment of net energy production and greenhouse gas emissions avoidance after 40 years of photovoltaics development. *Nat. Commun.* **7**, 13728 (2016).
42. Meinshausen, M. *et al.* Greenhouse-gas emission targets for limiting global warming to 2°C. *Nature* **458**, 1158–1162 (2009).
43. Lu, H. *et al.* Source code for: Greenhouse gas emissions from U.S. crude oil pipeline accidents: 1968 to 2020, *Figshare*, <https://doi.org/10.6084/m9.figshare.c.6702060.v1> (2023).
44. Belvederesi, C., Thompson, M. S. & Komers, P. E. Statistical analysis of environmental consequences of hazardous liquid pipeline accidents. *Heliyon* **4**, e00901 (2018).
45. Belvederesi, C., Thompson, M. S. & Komers, P. E. Canada's federal database is inadequate for the assessment of environmental consequences of oil and gas pipeline failures. *Environ. Rev.* **25**, 415–422 (2017).
46. U.S. Social Security Administration. Two-Letter State Abbreviations. <https://www.ssa.gov/international/coc-docs/states.html> (2023)

Acknowledgements

This study was financially supported by the Major Project of Fundamental Research on Frontier Leading Technology of Jiangsu Province (Grant No. BK20222006); and Natural Science Foundation of Jiangsu Province (Grant No. BK20220848).

Author contributions

Hongfang Lu contributed to research plan determination, acquisition of data, data analysis, result discussion, and manuscript writing. Zhao-Dong Xu contributed to research plan determination. Kaihui Song contributed to manuscript writing, data analysis, data processing, manuscript revising, and result discussion. Y. Frank Cheng contributed to data analysis, result discussion, and manuscript writing. Shaohua Dong contributed to data analysis and result discussion. Hongyuan Fang contributed to result discussion and manuscript writing. Haoyan

Peng contributed to Monte Carlo simulation. Yun Fu contributed to data processing. Dongmin Xi contributed to data processing. Zizhe Han contributed to data processing. Ximmeng Jiang contributed to data acquisition and data analysis. Yao-Rong Dong contributed to result discussion. Panpan Gai contributed to data analysis and manuscript reviewing. Yuli Shan contributed to research plan determination, data analysis, result discussion, and manuscript writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02478-4>.

Correspondence and requests for materials should be addressed to Z.-D.X. or Y.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023