npj | climate and atmospheric science

www.nature.com/npjclimatsci

ARTICLE OPEN

Check for updates

Where to place methane monitoring sites in China to better assist carbon management

Xiaorui Zhang[1], Chenhong Zhou[2], Yuzhong Zhang[3,4], Xiao Lu[5], Xiang Xiao[1], Fan Wang[1], Jun Song[1], Yike Guo[2], Kenneth K. M. Leung[6], Junji Cao[7] and Meng Gao[1]✉

Methane (CH_4) is the second most potent greenhouse gas (GHG), and China emerges as the largest anthropogenic CH_4 emitter by country. Current limited CH_4 monitoring systems in China are unfortunately inadequate to support carbon management. Here we use the Weather Research and Forecasting model (WRF) coupled with a GHG module and satellite constrained emissions to simulate the spatiotemporal distribution of CH_4 over East Asia in 2017. Model evaluations using both satellite retrievals and ground-based observations indicate reliable performance. We further inter-compare four proper orthogonal decomposition (POD)-based sensor placement algorithms and find they are able to capture main spatial features of surface CH_4 under an oversampled condition. The QR pivot algorithm exhibits superiority in capturing high CH_4, and it offers the best reconstruction with both high efficiency and accuracy. Areas with high CH_4 concentrations and intense anthropogenic activities remain underrepresented by current CH_4 sampling studies, leading to notable reconstruction error over central and eastern China. Optimal planning of 160 sensors guided by the QR pivot algorithm can yield reasonable reconstruction performance and costs of site construction. Our results can provide valuable references for future planning of CH_4 monitoring sites.

npj Climate and Atmospheric Science (2023)6:32 ; https://doi.org/10.1038/s41612-023-00359-6

INTRODUCTION

Atmospheric methane (CH_4) is the second most potent greenhouse gas (GHGs), trailing only carbon dioxide (CO_2) and responsible for more than one-quarter of the global radiative forcing of GHGs since pre-industrial times[1,2]. Although atmospheric residence time of CH_4 is relatively short (8–12 years), changes in CH_4 could have profound impacts on future climate and the oxidative capacity of the global atmosphere[1–3]. CH_4 mainly originates from natural and anthropogenic sources, including biomass burning, oil/gas industry, livestock, landfills, waste management, wetlands and rice cultivation[2,4,5]. Surface observations showed that atmospheric CH_4 rapidly increased from 1580 ppb in 1980s to 1910 ppb in 2022[2,6]. The specific reason for such increase remains unclear, yet it is most likely associated with increasing anthropogenic activities[4,7], which could further offset the climate benefits of carbon emission reductions significantly[8].

In 1978, Blake, et al.[9] began to measure tropospheric CH_4 worldwide and revealed a global increasing trend. Over time, several worldwide measurements of atmospheric CH_4 concentrations have been established (e.g., Global Atmosphere Watch (GAW) programme)[10–12], through ground-based instruments, tower, shipboard, and aircraft sampling[13,14]. Current monitoring networks are however unfortunately inadequate for sufficient spatial coverage. Although satellites instruments, such as Atmospheric Infrared Sounder (AIRS), Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), Greenhouse Gas Observation Satellite (GOSAT) and TROPOspheric Monitoring Instrument (TROPOMI), offer better data coverage, inconsistency is commonly found between satellite retrievals and in-situ observations due to the differences in observational density, accuracy and precision[15–17]. For example, the sensitivities of thermal emission instruments in the thermal infrared (TIR) are low at lower troposphere as they rely on thermal difference between surface and atmosphere[18,19]. In addition, small errors of satellite retrieval may also result in large errors in CH_4 emission estimation[20]. Lu, et al.[16] combined in-situ observations and satellite retrievals to quantify CH_4 emission comprehensively using an analytical inversion method. The surface observation can provide not only strong constraints for the inversion but also critical correlative components, such as methane isotopes and ethane[18]. In-situ observations with higher accuracy can also facilitate the evaluation of the ability of detecting CH_4 anomalies (e.g., large leaks from facilities) by satellites[18,21]. Therefore, it is of great need to build a comprehensive atmospheric CH_4 observing system.

China is the world-largest producer and consumer of coal[22,23], now emerging as the world's largest anthropogenic CH_4 emitter by country since the 2000s[24]. China also has the largest wetlands (rice paddies and natural wetlands) area in Asia, which produce additional amount of CH_4[25–27]. However, only three surface CH_4 observation sites have been established in China under the GAW of the World Meteorological Organization (WMO) since 1990s[28]. Several short-term CH_4 measurement campaigns were also organized[29,30], but measurements of CH_4 are still largely limited compared to ~1600 air quality monitoring sites across China. Air quality monitoring stations are mainly distributed in urban areas[31], while GAW sites focus on regional and continental background level of CH_4 concentration. High CH_4 from anthropogenic activities remain underrepresented. Both sufficient number and optimal locations[32] of CH_4 monitoring stations are essential for satellite

[1]Department of Geography, Faculty of Social Sciences, Hong Kong Baptist University, Hong Kong SAR, China. [2]Department of Computer Science, Faculty of Science, Hong Kong Baptist University, Hong Kong SAR, China. [3]School of Engineering, Westlake University, Hangzhou, Zhejiang Province, China. [4]Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang Province, China. [5]School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, Guangdong Province, China. [6]Air Science Group Environmental Protection Department, HKSAR, Hong Kong, China. [7]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China.
✉email: mmgao2@hkbu.edu.hk

data evaluation and assimilation, inversion of emission estimation and reconstruction of CH_4 concentration[21]. The optimal distribution of sensors refers to locations of limited sensors that could be used to derive the most accurate spatiotemporal distribution of ground-level CH_4 concentrations. However, identifying optimal locations by brute force search or exhaustive search is computationally expensive when the number of sensors and possible locations is large. Previous studies have implemented sparse reconstruction techniques to identify optimal locations of sparse monitoring sites for better reconstruction of spatiotemporal distributions. For instance, different sensor placement algorithms within the proper orthogonal decomposition (POD) framework were applied in the reconstruction of unsteady flow[33], ocean surface temperatures[34], and surface $PM_{2.5}$[32].

Given the essential role of CH_4 in carbon cycle and climate change mitigation, the significance of CH_4 monitoring has been recognized by China's government. To guide development of a CH_4 observing network in China, we infer spatiotemporal variations of CH_4 in China using the Weather Research and Forecasting model (WRF) with an updated GHG module (WRF-GHG)[35] and satellite constrained emission estimation[36]. We further investigate the reconstruction accuracy of four sensor placement algorithms for optimal planning of monitoring sites to better depict surface CH_4 concentration in China. The results can provide valuable references for future planning of CH_4 monitoring sites.

RESULTS

Model evaluation

Daily surface CH_4 from ground-based instruments and column mixing ratios of CH_4 from the GOSAT satellite were used to evaluate the performance of WRF-GHG in simulating CH_4. As shown in Fig. 1, high values of CH_4 dry column mixing ratios are found in Sichuan Basin (SCB), especially in summer (approximate 1930 ppb) due to emissions from paddy field, livestock and energy activities[37,38], and unfavorable dispersion conditions[39]. Relatively large CH_4 column can be seen over eastern China driven by coal mining (northern part)[22] and rice paddy fields (southern part)[26,38], reaching a maximum in summer (1910 ppb) and a minimum (1870 ppb) in winter. The WRF-GHG can reproduce the distribution pattern of CH_4 column, with the difference lower than 10 ppb during autumn and winter (Fig. 1). Although the simulation may overestimate CH_4 column in spring and summer, locations of CH_4 hot spots, such as eastern and southern China, are generally consistent. The hydroxyl radical (OH) is the main CH_4 sink in the troposphere, reaching high value in India and western China[40], where simulated CH_4 column is 10 to 30 ppb higher than observations (Supplementary Fig. 1). As the CH_4 is treated in a passive way in WRF-GHG without chemical reactions, the overestimation of CH_4 over western China might be mainly caused by missing chemical loss of CH_4. Compared with surface observations, the WRF-GHG model can capture the daily variation of CH_4 with a correlation coefficient of 0.67 (Fig. 2). The simulated values are generally higher than observation, especially during summer and autumn, which are the two main periods of CH_4 emissions from livestock and vegetation in Qinghai province[38]. Considering that the Mt. Waliguan (WLG) station is isolated from anthropogenic activities, it provides background CH_4 within the Eurasian continent rather than local conditions[41]. The locations of current sites are not in high CH_4 centers either[28]. The performance of model, as well as satellite retrievals, cannot be evaluated comprehensively with limited monitoring sites in China, requiring more CH_4 observation stations to support relevant research activities.

Reconstruction accuracy of four sensor placement algorithms

Four sensor placement algorithms were employed to estimate the distribution of surface CH_4 across China. The results generated by different algorithms using 10 POD modes ($n = 10$) are shown in Table 1. Under the condition that number of sensor quantity equals to POD modes ($P = n$), reconstructed results generated by matrix condition number (MCN) and Extrema exhibit relatively poor performance with seriously high mean percentage error (MPE) and root-mean-square error (RMSE) values. Their reconstruction performances are improved significantly when sensor quantity exceeds the number of modes ($P = 1.5n$ and $P = 2n$). Compared with other algorithms, Extrema exhibits the poorest ability with correlation of determination (R^2) of 0.34 and 0.44 under $P = 1.5n$ and $2n$ conditions, respectively. The evaluation metrics with DEIM and QR pivot under $P = n$ condition are

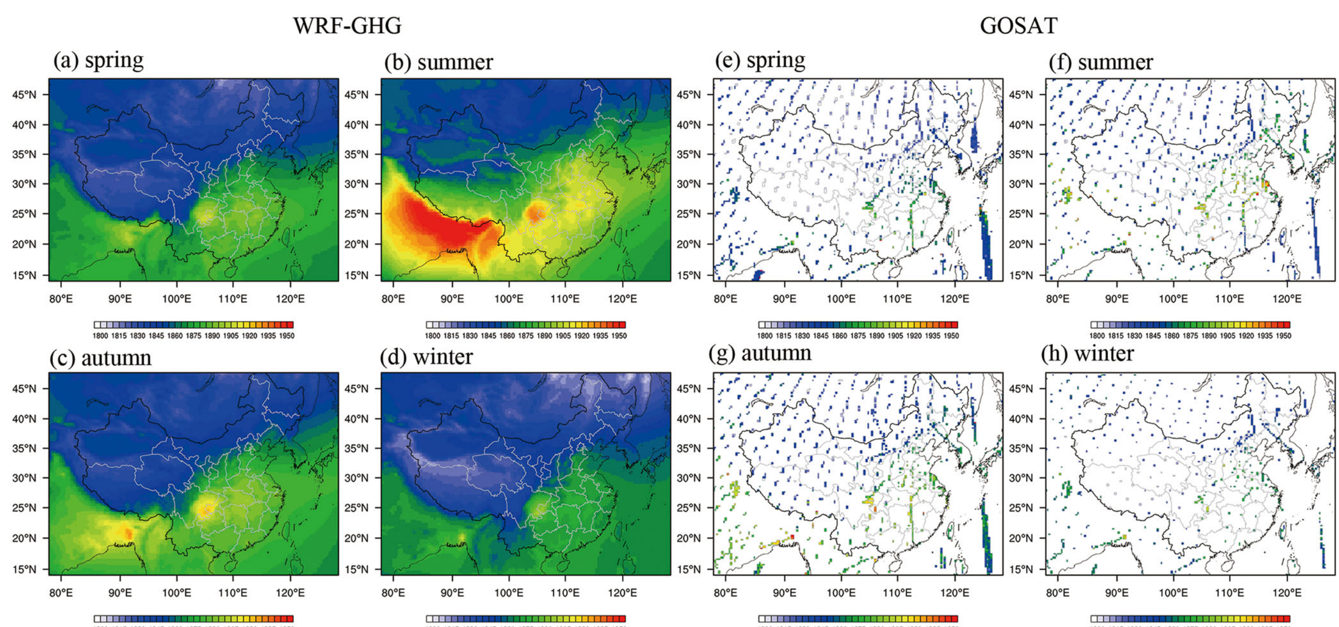

Fig. 1 Model evaluation of CH_4 distribution. Seasonal distributions of dry column mixing ratios of CH_4 (Unit: ppb) from WRF-GHG (**a**–**d**) and GOSAT (**e**–**h**) in 2017.

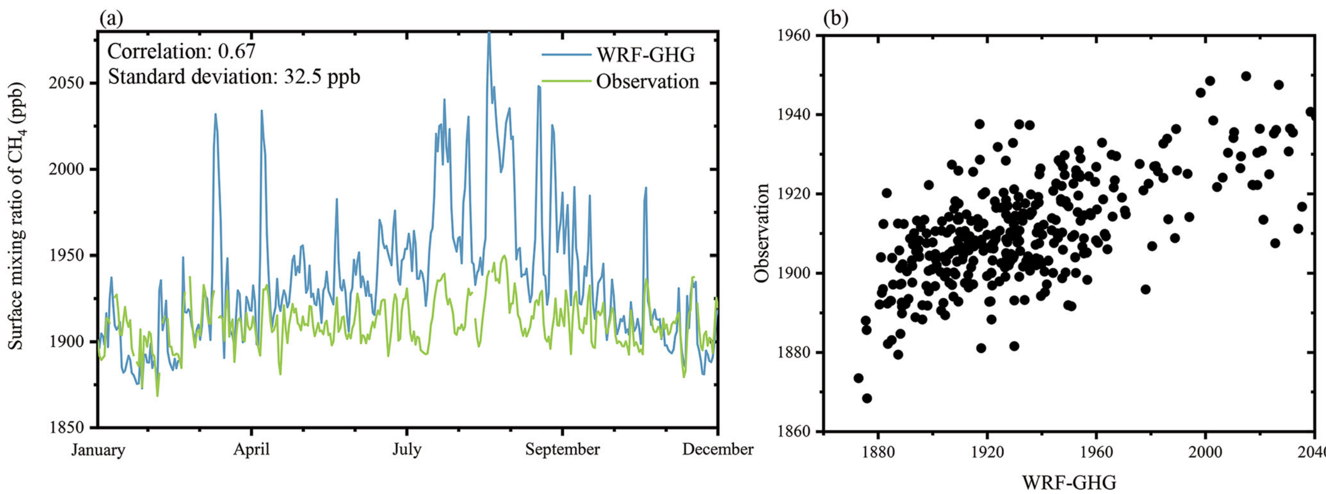

Fig. 2 Model evaluation of surface CH_4. **a** Daily variations of surface CH_4 (unit: ppb) from observation (green) and WRF-GHG (blue) in 2017. **b** Scatter plot of simulated and observed surface CH_4.

Table 1. 10-fold cross validation of reconstruction accuracy of four sensor placement algorithms (*n* and *P* denote the number of POD modes and sensors, unit of MPE: %, unit of RMSE: ppb).

$n = 10$	$P = n$			$P = 1.5n$			$P = 2n$		
	MPE	R^2	RMSE	MPE	R2	RMSE	MPE	R^2	RMSE
MCN	23.43	0.27	752.89	4.69	0.56	156.55	3.61	0.67	119.16
Extrema	31.47	0.18	1010.64	7.65	0.34	266.25	5.44	0.44	188.41
DEIM	5.01	0.71	124.92						
QR pivot	4.94	0.74	122.35	3.88	0.79	99.47	3.46	0.81	90.14

substantially better than those with MCN and Extrema. Given that the sensor quantity should equal to the dimension of POD basis[42], the DEIM algorithm only produces CH_4 reconstruction under $P = n$ condition. QR pivot shows the best performance among four algorithms with the largest R^2 of 0.86.

Figure 3 shows the spatial distributions of surface CH_4 concentrations over China from four algorithms and their associated locations of sensors under an oversampled case ($P = 2n$). The distributions of surface CH_4 are closely related with emission sources (Fig. 3a). The intense coal production activities in Shanxi and Guizhou provinces account for 35% and 28% of national coal mine CH_4 emissions, respectively (Sheng et al., 2019), resulting in surface CH_4 as high as 2700 ppb. All four algorithms can generally capture main spatial features of surface CH_4. Although the computational costs of Extrema algorithm are the lowest[34], the Extrema algorithm overestimates CH_4 over eastern China, with the largest MPE of 5.44%. Owing to proximity or coincidence of the locations of Extrema POD modes[43], their sensors are mainly densely located (Fig. 3g). Unlike Extrema, MCN algorithm evenly distributes sensor across China (Fig. 3f) and shows a better performance with R^2 of 0.75. The distributions of sensors from QR pivot algorithm are mainly concentrated in high CH_4 regions, providing the best CH_4 reconstruction with MPE and RMSE as low as 3.46% and 90.14 ppb.

As noted in Luo, et al.[31], the number of monitoring sites can influence the assessments of distribution of air pollutants. As shown in Fig. 4, reconstruction performance improves with increasing sensor quantity except for DEIM method. The accuracy of DEIM is limited by the condition that sensor quantity equals to mode quantity. The reconstruction accuracy of Extrema increases rapidly, reaching up to R^2 of 0.72 with 300 sensors. The reconstructed results using MCN are more consistent with simulation in low CH_4 regions due to evenly distributed sensors (Fig. 3f). Thus, MCN has the smallest MPE among four algorithms when the number of sensors is higher than 20, and MPE reaches 1.62% with 300 sensors. QR pivot in general offers the best reconstruction, which is competitive in both efficiency and accuracy. The RMSE and R^2 of QR pivot with 100 sensors are 69.82 ppb and 0.86, approximately equal to those of MCN with 300 sensors. Therefore, the QR pivot algorithm is regarded as a more reliable method.

Optimal planning of sensor locations

The locations of CH_4 sampling sites documented in previous observation studies are summarized in Table 2, which are regarded as places of potential stations in this study. The differences between simulated and reconstructed surface CH_4 using these potential stations and QR pivot-guided stations are shown in Fig. 5. In addition to GAW stations that focus on background CH_4 level, other field measurements were mainly conducted in megacities with larger influences of anthropogenic activities[44–46], or in southern China to capture CH_4 released from rice fields[29,37]. Thus, high CH_4 concentration over southern China are well reconstructed by using the potential stations located in urban and farmland areas. Performances of reconstruction using locations of potential sites are even slightly better than MCN with MPE of 3.14%, R^2 of 0.69 and RMSE of 104.16 ppb (Supplementary Table 1). However, potential stations are likely to miss sources of coal mining, leading to notable discrepancies in central and eastern China (Fig. 5a). MPE and RMSE values of reconstruction based on potential sites are 0.32% and 14.01 ppb larger than those of QR pivot.

Although QR pivot algorithm with 20 sensors can reproduce surface CH_4 reasonably, the results may be unacceptable with RMSE up to 90 ppb. As shown in Fig. 4, the performances of reconstruction

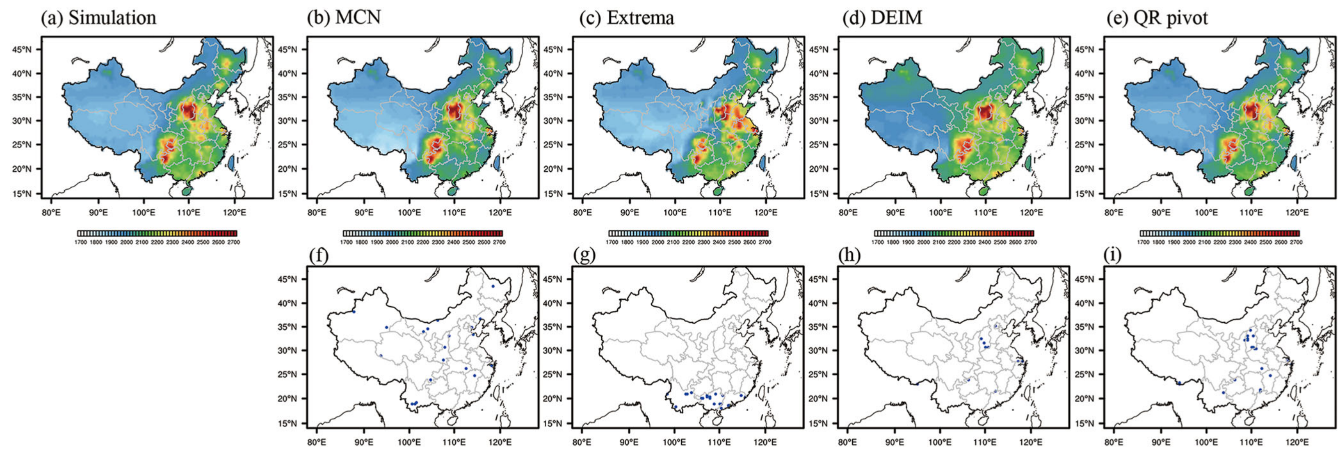

Fig. 3 Reconstructed surface CH_4 concentrations and locations of sensors. Spatial distributions of mean surface CH_4 concentrations across China from (**a**) WRF-GHG and reconstructed results with (**b**) MCN, (**c**) Extrema and (**e**) QR pivot algorithms under the condition of 10 POD modes and 20 sensors ($P = 2n$); (**d**) DEIM under $P = n$; associated locations of sensors (**f**–**i**).

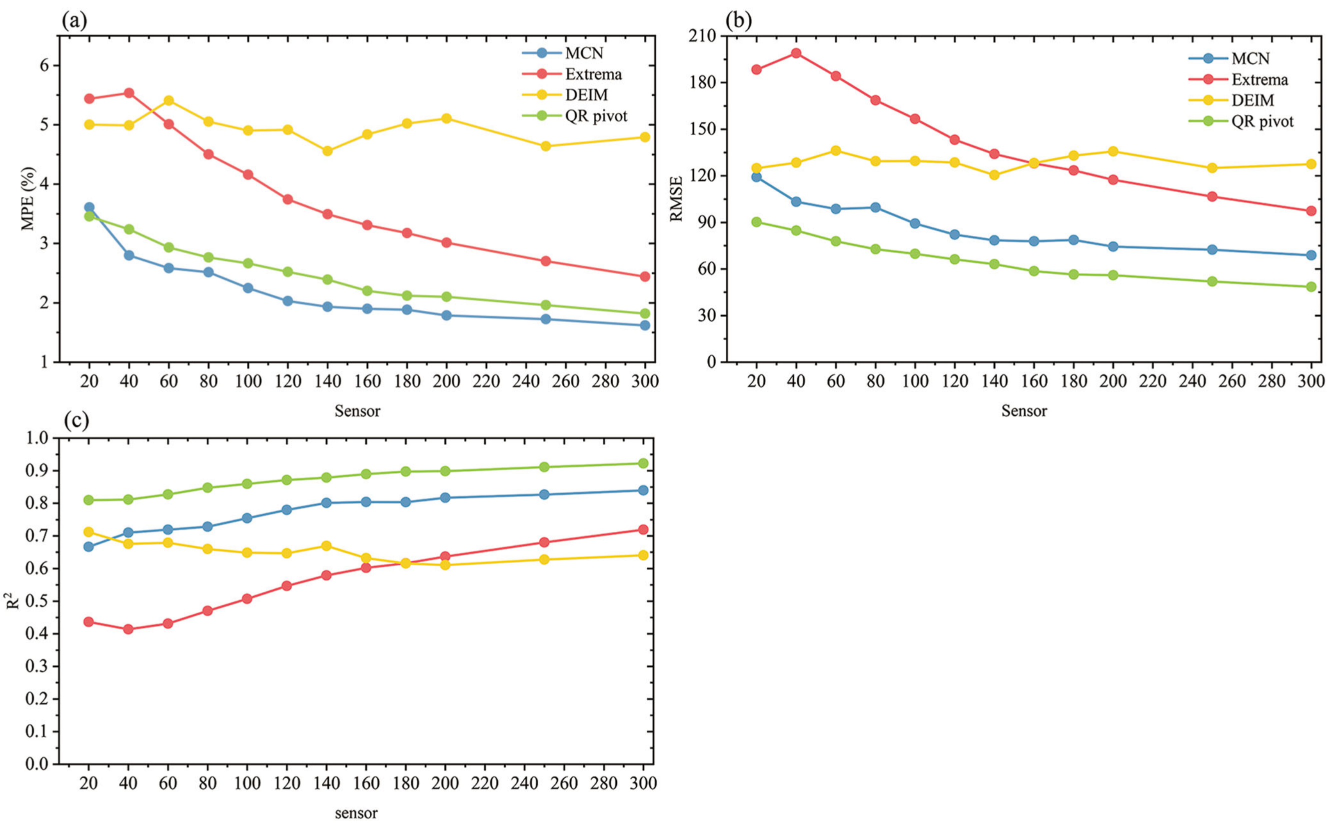

Fig. 4 Reconstruction errors with different sensor quantity. (**a**) MPE, (**b**) RMSE and (**c**) R^2 of MCN, Extrema and QR pivot with increasing numbers of sensors under the condition of $P = 2n$ and DEIM algorithm under the condition of $P = n$.

using QR pivot are improved significantly when sensor quantity increases from 20 to 160. To provide implications for optimal planning of CH_4 monitoring stations, the reconstruction performances of QR pivot algorithm with 40, 100, 160, 200 and 300 sensors under $P = 2n$ condition are further investigated (Fig. 6). When sensor quantity increases to 100, the reconstruction shows more reliable performance with RMSE of 69.82 ppb but underestimates CH_4 concentration by up to 60 ppb over southern China. Such underestimation does not exist in the reconstruction with 160 sensors. QR pivot with 160 sensors has exceptional reconstruction performance with RMSE of 58.56 ppb, and notable bias cannot be seen over western China. With further growth of sensor quantity, slight improvements are found in reconstructions generated by 200 and 300 sensors with RMSE of 55.96 and 48.46 ppb. It indicates that QR pivot with 160 sensors is suitable considering both reconstruction performance and costs of site construction.

DISCUSSION

Increasing atmospheric CH_4 concentration is of global concern with respect to climate change mitigation. China emerges as the largest anthropogenic CH_4 emission country and accounts for larger than one-quarter of the increase in global anthropogenic CH_4 emissions. However, current CH_4 monitoring networks are

Table 2. The locations of sampling sites from previous CH_4 observation studies.

Observation site	Province	Location	Reference
Mt. Waliguan	Qinghai	36.29°N, 100.90°E	Fang et al.[28]
Longfengshan	Heilongjiang	44.73°N, 127.6°E	
Lin'an	Zhejiang	30.18°N, 119.44°E	
Shangdianzi	Beijing	36.29°N, 100.92°E	Fang et al.[45]
Xinglong station	Heibei	40.4°N, 117.5°E	Wang et al.[44]
Shanghai	Shanghai	31.24°N, 121.49°E	Wei and Wang[46]
Jingdezhen	Jiangxi	29.37°N, 117.22°E	Xia et al.[30]
Nanchang	Jiangxi	28.68°N, 115.85°E	
Ganzhou	Jiangxi	25.83°N, 114.93°E	
Guangzhou	Guangdong	23.25°N, 113.1°E	Cai et al.[29]
Yingtan	Jiangxi	28.3°N, 117.1°E	
Changsha	Hunan	28.15°N, 113.1°E	
Southwest University	Chongqing	29.8°N, 106.3°E	
Nanjing	Jiangsu	31.97°N, 118.8°E	
Jurong	Jiangsu	31.8°N, 119.15°E	
Suzhou	Jiangsu	31.3°N, 121.2°E	
Fengqiu	Henan	35.4°N, 114.4°E	
Beibei district	Chongqing	30.43°N, 106.43°E	Hao et al.[37]
Dunhuang	Gansu	40.08° N, 94.40° E	Wei et al.[53]
Hefei	Anhui	31.9°N, 117.17°E	Wang et al.[54]

unfortunately inadequate to offer sufficient spatial coverage, limiting the satellite evaluation, inversion of CH_4 emission estimation, data assimilation and reconstruction of CH_4 concentration. To build a comprehensive atmospheric CH_4 observing system, locations of monitoring sites should be considered. In this study, we simulated the spatial distribution of CH_4 over East Asia in 2017 using the WRF-GHG model and identified optimal site locations with four sensor placement algorithms. The influences of sensor quantity and locations on reconstruction accuracy of different algorithms were investigated systematically.

Model evaluations using satellite retrieval and surface observation indicated reliable performance of WRF-GHG in simulating spatial and temporal distributions of CH_4 in 2017. High surface CH_4 centers were mainly located over Shanxi and Guizhou province driven by coal mine CH_4 emissions, and eastern China due to emissions from paddy field and energy activities. Four POD-based sensor placement algorithms could capture main spatial distribution features under an oversampled condition. When sensor quantity equals to POD modes, the reconstructed results from DEIM and QR pivot were substantially better than those from MCN and Extrema. Owing to evenly distributed sensors, the reconstructed result from MCN was more consistent with simulation in low CH_4 regions with the smallest MPE when the number of sensors is higher than 20. The QR pivot showed the best performance in selecting optimal monitoring site locations with both high efficiency and accuracy. Using the locations of CH_4 sampling sites from previous observation studies as potential stations, reconstruction performance using potential sites had 0.32% and 14.01 ppb larger MPE and RMSE values than those of QR pivot. Notable errors were found over central and eastern China. The reconstruction performance could be significantly improved by increasing the number of sensors until the sensor quantity reached 160. QR pivot with 160 sensors exhibited exceptional reconstruction performance with RMSE of 58.56 ppb and overestimation only over low CH_4 concentration regions (western China). Therefore, QR pivot with 160 sensors can provide an optimal planning of CH_4 monitoring sites in China considering both reconstruction performance and costs of site construction. Given that sensor placement algorithms are data-driven methods, quality of input data has critical impacts on accuracy of algorithms. The WRF-GHG overestimated the CH_4 concentration in western China, which can be partly attributed to no chemical loss of CH_4 in the model. Future work to advance modeling and combine modeling and observations to derive better dynamical evolution of CH_4 would be helpful. Our results can provide valuable references for future planning of CH_4 monitoring sites.

METHODS

WRF-GHG model

WRF-Chem Version 3.9.1[47], enhanced with a GHG module[48], was used to simulate CH_4 in China. WRF-Chem is a mesoscale coupled meteorology-chemistry model, and WRF-GHG is now a module in WRF-Chem for transport of CO_2 and CH_4 tracers[47,49]. CH_4 in WRF-GHG, treated in a passive way, was transported online without atmospheric chemical reactions. The CH_4 emission inventories were taken from Zhang, et al.[36] with a spatial resolution of 0.5° × 0.625°, including emissions from biomass burning, coal, gas, landfills, livestock, oil, rice, geological seeps, termites, wastewater and wetlands.

The simulation domain covered the East Asia region with 115 × 164 grid points at a horizontal resolution of 36 km × 36 km. We used 29 vertical layers up to 50 hPa. The National Center for Environmental Prediction Final Analysis (NCEP FNL) dataset at a 6 hourly temporal interval and 1° × 1° horizontal resolution was used as meteorological initial and boundary conditions. The initial and lateral boundary conditions for CH_4 were implemented using GEOS-Chem simulations with 4° × 5° resolution from[16].

Observations

Both surface and GOSAT satellite observations were used to validate model performance. Daily surface CH_4 concentrations at the WLG station were provided by the World Data Centre for Greenhouse Gases (https://gaw.kishou.go.jp/). The WLG baseline observatory is located in western China (36.29°N, 100.90°E), isolated from industrial and populated regions. Other sites located in China, i.e., Shangdianzi and Lulin stations, were not used in this study due to missing data for 2017 or low temporal resolution. GOSAT, launched in 2009, measures column-averaged dry CH_4 mixing ratios with high precision of 0.7% from a polar sun-synchronous orbit at about 13:00 local time[50]. The University of Leicester version 9.0 Proxy XCH4 retrieval was used in this study, which has a global precision of 9 ppb[51]. It offers column-averaged dry-air mole fraction of CH_4 based on GOSAT Level 1B data[51].

Sensor placement algorithms within the POD framework

POD is a commonly used dimensionality reduction technique, which produces low-dimensional dynamical systems that could accurately model spatiotemporal variations of dominant structures of data. The gappy POD is a modification of POD to handle incomplete data, developed by Everson and Sirovich[52]. Gappy POD can be used in reconstruction of data based on sparse networks[33,34]. Suppose a snapshot $\theta(x,t)$ along the domain at time step t, contains K elements. It can be expressed as linear combination of K time-invariant patterns:

$$\theta(x,t) = a_1(t)\phi_1(x) + a_2(t)\phi_2(x) + \ldots a_K(t)\phi_K(x) \quad (1)$$

where $\{\varphi_i(x)\}_{i=1}^{K}$ are space-varying basis functions, which are orthogonal to each other, and $\{a_i(t)\}_{i=1}^{K}$ are corresponding time-varying coefficients. Equation (1) can be also approximated as a truncated expansion:

$$\theta(x,t) \approx a_1(t)\phi_1(x) + a_2(t)\phi_2(x) + \ldots a_n(t)\phi_n(x) \quad (2)$$

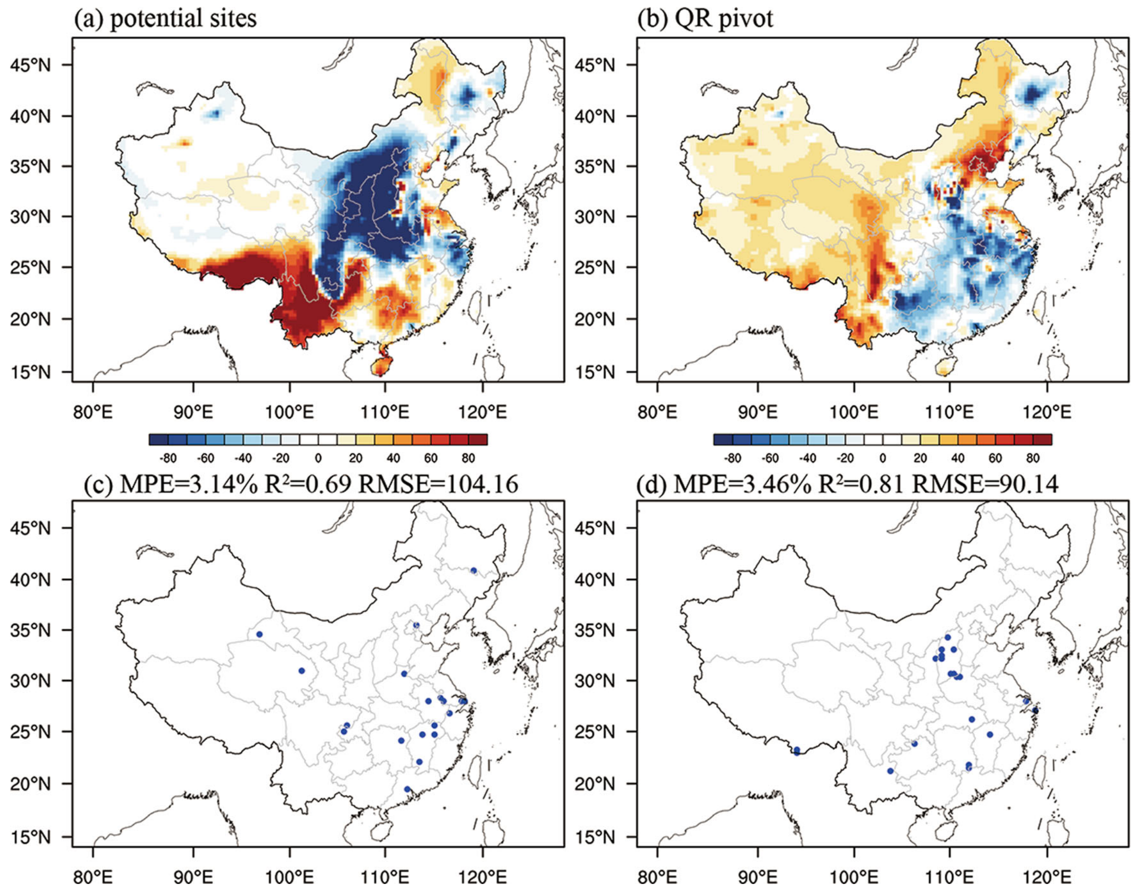

Fig. 5 Difference in reconstruction ability using potential sites and QR pivot algorithm. Differences between simulated and reconstructed surface CH_4 concentration using (**a**) potential sites and (**b**) QR pivot algorithm under the condition of $P = 2n$ ($n = 10$); associated locations of sensors (**c**, **d**).

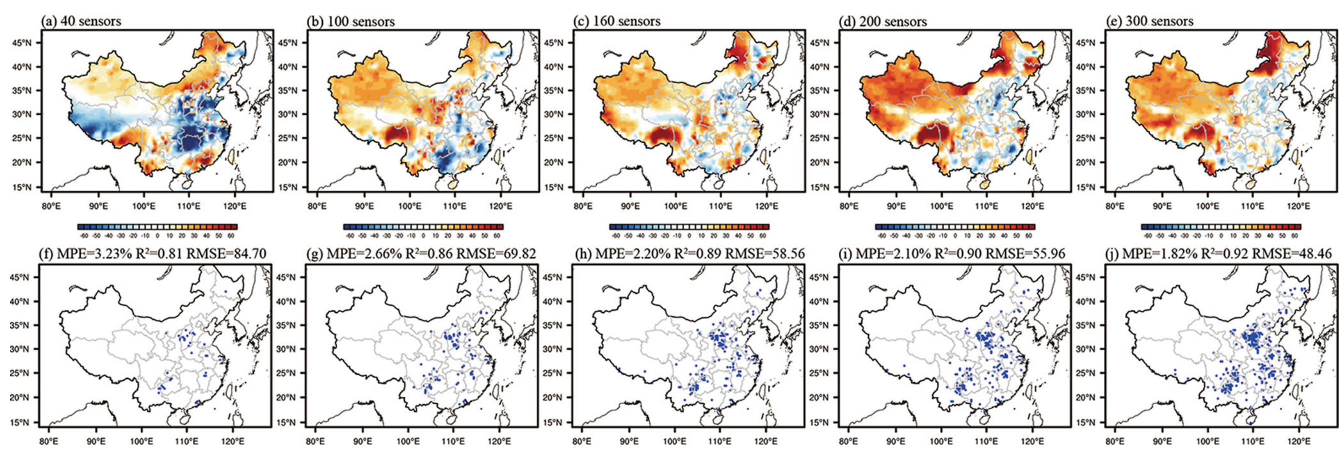

Fig. 6 Reconstruction ability of QR pivot with increasing sensor quantity. Differences between simulated and reconstructed surface CH_4 concentration using QR pivot with (**a**) 40, (**b**) 100, (**c**) 160, (**d**) 200 and (**e**) 300 sensors under the condition of $P = 2n$, and associated locations of sensors (**f**–**j**).

where n denotes the number of POD modes, which is fewer than K in Eq. (1). The gappy POD further defines a mask vector m(x,t) to describe where data are missing (m(x,t)=0) or available (m(x,t)=1). Pointwise multiplication is defined as $\Theta(x,t) = m(x,t) \cdot \theta(x,t)$. Assuming a "repaired" vector θ_R from the incomplete θ, it can be represented as follows:

$$\theta_R(x,t) \approx \beta_1(t)\phi_1(x) + \beta_2(t)\phi_2(x) + \ldots\beta_n(t)\phi_n(x) \quad (3)$$

where coefficient β is computed by minimizing the difference between θ_R and Θ. It can be differentiated with respect to β(t) and yielded the linear equation system: Mβ = f[32,52], where $M_{ij} = (\phi_i, \phi_j)_n$ and $f = (\Theta, \phi_i)_n$.

To identify the optimal sensor locations, several sensor placement algorithms within the POD framework were proposed. Willcox[33] developed a minimization of the MCN algorithm. The condition number of M is used to evaluate the reconstruction performance, which becomes larger than 1 for gappy data because the orthogonality of M is lost. The sensors are placed at the grids that minimizes the condition number to reserve orthogonality. Yildirim, et al.[34] proposed a method to select extrema of the POD modes (Extrema algorithm), which would maximally capture the variance. The sensors are selected at the location that are the maximum and minimum of each the POD modes. To improve the dimension reduction efficiency, Chaturantabut and Sorensen[42] proposed the simplified discrete empirical interpolation method (DEIM) to approximate the nonlinearity by discretely sampling and evaluating the nonlinearity. DEIM recursively learns the interpolation points (sensor locations) according to the maximum linear dependence error. A column permutation matrix D is introduced by the QR with column pivoting. It contains ones and zeros to make the diagonal values of A in a decreasing order: AD = QR, to maximize the absolute value of M. The sensor locations can be obtained from D.

Four sensor placement algorithms mentioned above were applied to identify locations of sensors and reconstruct the simulated surface CH_4. Therefore, the data matrix provided by WRF-GHG is composed of 365 snapshots. Each snapshot has 115 × 164 pixels. 10-fold cross validation was used to ensure the reliability of our results. Three evaluation metrics were used to assess the reconstruction ability of POD-based sensor placement algorithms, including MPE, R^2, and RMSE, as follows:

$$MPE = \frac{100\%}{N}\sum_{i=1}^{N}\frac{\left|Con_{s,i} - Con_{e,i}\right|}{Con_{s,i}} \tag{4}$$

$$R^2 = \frac{\sum_{i=1}^{N}\left(Con_{s,i} - \overline{Con_s}\right)^2\left(Con_{e,i} - \overline{Con_e}\right)^2}{\sum_{i=1}^{N}\left(Con_{s,i} - \overline{Con_s}\right)^2\sum_{i=1}^{N}\left(Con_{e,i} - \overline{Con_e}\right)^2} \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Con_{s,i} - Con_{e,i}\right)^2}{N}} \tag{6}$$

where Con_s and Con_e represent simulated and estimated surface CH_4 concentrations, respectively. $\overline{Con_s}$ and $\overline{Con_e}$ are the mean value of Con_s and Con_e. N stands for the total number of grids within China. To better compare the capabilities of different algorithms, we initially used only 10 POD modes, and further explored the influences of the number of POD modes and sensors on reconstruction accuracy, with sensor quantity ranging from 20 to 300.

DATA AVAILABILITY

All model output data can be obtained from the authors upon request to mmgao2@hkbu.edu.hk. The CH_4 emission datasets are available through a public repository (https://doi.org/10.57760/sciencedb.02269).

CODE AVAILABILITY

All codes can be obtained from the authors upon request to mmgao2@hkbu.edu.hk.

Received: 4 November 2022; Accepted: 28 March 2023;
Published online: 21 April 2023

REFERENCES

1. Stocker, T. *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2014).
2. Kirschke, S. et al. Three decades of global methane sources and sinks. *Nat. Geosci.* **6**, 813–823 (2013).
3. Shindell, D. et al. Simultaneously mitigating near-term climate change and improving human health and food security. *Science* **335**, 183–189 (2012).
4. Saunois, M. et al. The global methane budget 2000–2012. *Earth Syst. Sci. Data* **8**, 697–751 (2016).
5. Zhang, Y. et al. Quantifying methane emissions from the largest oil-producing basin in the United States from space. *Sci. Adv.* **6**, eaaz5120 (2020).
6. Tollefson, J. Scientists raise alarm over'dangerously fast'growth in atmospheric methane. *Nature* https://doi.org/10.1038/d41586-022-00312-2 (2022).
7. Turner, A. J., Frankenberg, C., Wennberg, P. O. & Jacob, D. J. Ambiguity in the causes for decadal trends in atmospheric methane and hydroxyl. *Proc. Natl Acad. Sci.* **114**, 5367–5372 (2017).
8. Nisbet, E. G. et al. Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris Agreement. *Glob. Biogeochem. Cycles* **33**, 318–342 (2019).
9. Blake, D. R. et al. Global increase in atmospheric methane concentrations between 1978 and 1980. *Geophys. Res. Lett.* **9**, 477–480 (1982).
10. Dlugokencky, E. J., Steele, L. P., Lang, P. M. & Masarie, K. A. The growth rate and distribution of atmospheric methane. *J. Geophys. Res.: Atmospheres* **99**, 17021–17043 (1994).
11. Steele, L. et al. in *Scientific Application of Baseline Observations of Atmospheric Composition (SABOAC)* 417–463 (Springer, 1987).
12. Wunch, D. et al. The total carbon column observing network. *Philos. Trans. R. Soc. A: Math., Phys. Eng. Sci.* **369**, 2087–2112 (2011).
13. Houweling, S. et al. Global inverse modeling of CH 4 sources and sinks: an overview of methods. *Atmos. Chem. Phys.* **17**, 235–256 (2017).
14. Wu, X., Zhang, X., Chuai, X., Huang, X. & Wang, Z. Long-term trends of atmospheric CH4 concentration across China from 2002 to 2016. *Remote Sens.* **11**, 538 (2019).
15. Turner, A. J. et al. Assessing the capability of different satellite observing configurations to resolve the distribution of methane emissions at kilometer scales. *Atmos. Chem. Phys.* **18**, 8265–8278 (2018).
16. Lu, X. et al. Global methane budget and trend, 2010–2017: complementarity of inverse analyses using in situ (GLOBALVIEWplus CH 4 ObsPack) and satellite (GOSAT) observations. *Atmos. Chem. Phys.* **21**, 4637–4657 (2021).
17. Cressot, C. et al. On the consistency between global and regional methane emissions inferred from SCIAMACHY, TANSO-FTS, IASI and surface measurements. *Atmos. Chem. Phys.* **14**, 577–592 (2014).
18. Jacob, D. J. et al. Satellite observations of atmospheric methane and their value for quantifying methane emissions. *Atmos. Chem. Phys.* **16**, 14371–14396 (2016).
19. Worden, J. et al. Quantifying lower tropospheric methane concentrations using GOSAT near-IR and TES thermal IR measurements. *Atmos. Meas. Tech.* **8**, 3433–3445 (2015).
20. Buchwitz, M. et al. The Greenhouse Gas Climate Change Initiative (GHG-CCI): Comparison and quality assessment of near-surface-sensitive satellite-derived CO2 and CH4 global data sets. *Remote Sens. Environ.* **162**, 344–362 (2015).
21. Fraser, A. et al. Estimating regional methane surface fluxes: the relative importance of surface and GOSAT mole fraction measurements. *Atmos. Chem. Phys.* **13**, 5697–5713 (2013).
22. Sheng, J., Song, S., Zhang, Y., Prinn, R. G. & Janssens-Maenhout, G. Bottom-up estimates of coal mine methane emissions in China: a gridded inventory, emission factors, and trends. *Environ. Sci. Technol. Lett.* **6**, 473–478 (2019).
23. Thompson, R. L. et al. Methane emissions in East Asia for 2000–2011 estimated using an atmospheric Bayesian inversion. *J. Geophys. Res.: Atmospheres* **120**, 4352–4369 (2015).
24. Crippa, M. et al. *Fossil CO_2 and Ghg Emissions of All World Countries*. (Publication Office of the European Union, 2019).
25. Li, T. et al. Performance of CH4MODwetland for the case study of different regions of natural Chinese wetland. *J. Environ. Sci.* **57**, 356–369 (2017).
26. Chen, H. et al. Methane emissions from rice paddies natural wetlands, lakes in China: synthesis new estimate. *Glob. change Biol.* **19**, 19–32 (2013).
27. Wang, Z., Wu, J., Madden, M. & Mao, D. China's wetlands: conservation plans and policy impacts. *Ambio* **41**, 782–786 (2012).
28. Fang, S. X., Zhou, L. X., Masarie, K. A., Xu, L. & Rella, C. W. Study of atmospheric CH4 mole fractions at three WMO/GAW stations in China. *J. Geophys. Res.: Atmos.* **118**, 4874–4886 (2013).
29. Cai, Z., Tsuruta, H. & Minami, K. Methane emission from rice fields in China: measurements and influencing factors. *J. Geophys. Res.: Atmos.* **105**, 17231–17242 (2000).
30. Xia, L., Zhang, G., Zhan, M., Li, B. & Kong, P. Seasonal variations of atmospheric CH4 at Jingdezhen station in Central China: Understanding the regional transport and its correlation with CO2 and CO. *Atmos. Res.* **241**, 104982 (2020).
31. Luo, H. et al. The Impact of the Numbers of Monitoring Stations on the National and Regional Air Quality Assessment in China During 2013–18. *Adv. Atmos. Sci.* **39**, 1709–1720 (2022).
32. Zhou, C. et al. Optimal planning of air quality-monitoring sites for better depiction of PM2. 5 pollution across China. *ACS Environmental Au* **2**, 314–323 (2022).

33. Willcox, K. Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. *Comput. fluids* **35**, 208–226 (2006).
34. Yildirim, B., Chryssostomidis, C. & Karniadakis, G. Efficient sensor placement for ocean measurements using low-dimensional concepts. *Ocean Model.* **27**, 160–173 (2009).
35. Beck, V. et al. WRF-Chem simulations in the Amazon region during wet and dry season transitions: evaluation of methane models and wetland inundation maps. *Atmos. Chem. Phys.* **13**, 7961–7982 (2013).
36. Zhang, Y. et al. Observed changes in China's methane emissions linked to policy drivers. *Proc. Natl Acad. Sci.* **119**, e2202742119 (2022).
37. Hao, Q. et al. Drainage, no-tillage and crop rotation decreases annual cumulative emissions of methane and nitrous oxide from a rice field in Southwest China. *Agric. Ecosyst. Environ.* **233**, 270–281 (2016).
38. Gong, S. & Shi, Y. Evaluation of comprehensive monthly-gridded methane emissions from natural and anthropogenic sources in China. *Sci. Total Environ.* **784**, 147116 (2021).
39. Zhang, X. et al. Observed sensitivities of PM2. 5 and O_3 extremes to meteorological conditions in China. *Environ. Int.* **168**, 107428 (2022).
40. Zhao, Y. et al. Reconciling the bottom-up and top-down estimates of the methane chemical sink using multiple observations. *Atmos. Chem. Phys.* **23**, 789–807 (2023).
41. Zhou, L. et al. Ten years of atmospheric methane observations at a high elevation site in Western China. *Atmos. Environ.* **38**, 7041–7054 (2004).
42. Chaturantabut, S. & Sorensen, D. C. Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput.* **32**, 2737–2764 (2010).
43. Yang, X., Venturi, D., Chen, C., Chryssostomidis, C. & Karniadakis, G. E. EOF-based constrained sensor placement and field reconstruction from noisy ocean measurements: Application to Nantucket Sound. *J. Geophys. Res.: Oceans* **115** (2010).
44. Wang, Y.-S., Zhou, L., Wang, M.-X. & Zheng, X.-H. Trends of atmospheric methane in Beijing. *Chemosphere-Glob. Change Sci.* **3**, 65–71 (2001).
45. Fang, S.-x, Tans, P. P., Dong, F., Zhou, H. & Luan, T. Characteristics of atmospheric CO_2 and CH_4 at the Shangdianzi regional background station in China. *Atmos. Environ.* **131**, 1–8 (2016).
46. Wei, C. & Wang, M. Spatial distribution of greenhouse gases (CO_2 and CH_4) on expressways in the megacity Shanghai, China. *Environ. Sci. Pollut. Res.* **27**, 31143–31152 (2020).
47. Skamarock, W. C. et al. A description of the Advanced Research WRF version 3. NCAR Technical note-475+ STR (2008).
48. Beck, V. et al. Jena, Germany, https://www.bgc-jena.mpg.de/bgcsystems/pmwiki2/uploads (2011).
49. Zhao, X. et al. Analysis of total column CO_2 and CH_4 measurements in Berlin with WRF-GHG. *Atmos. Chem. Phys.* **19**, 11279–11302 (2019).
50. Kuze, A. et al. Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space. *Atmos. Meas. Tech.* **9**, 2445–2461 (2016).
51. Parker, R. J. et al. A decade of GOSAT Proxy satellite CH_4 observations. *Earth Syst. Sci. Data* **12**, 3383–3412 (2020).
52. Everson, R. & Sirovich, L. Karhunen–Loeve procedure for gappy data. *JOSA A* **12**, 1657–1664 (1995).
53. Wei, C., Lyu, Z., Bu, L. & Liu, J. Occurrence and discrepancy of surface and column mole fractions of CO_2 and CH_4 at a desert site in Dunhuang, Western China. *Atmosphere* **13**, 571 (2022).
54. Wang, W. et al. Investigating the performance of a greenhouse gas observatory in Hefei, China. *Atmos. Meas. Tech.* **10**, 2627–2643 (2017).

ACKNOWLEDGEMENTS

This study was supported by grants from Research Grants Council of the Hong Kong Special Administrative Region, China (project no. HKBU22201820 and HKBU12202021) and National Natural Science Foundation of China (No. 42005084).

AUTHOR CONTRIBUTIONS

This study was conceived by M.G., and X.Z., X.L., and F.W. conducted simulations. X.Z., C.Z., J.S., and Y.G. conducted proper orthogonal decomposition. Y.Z. provided emission data. X.Z. and M.G. wrote the paper with inputs from X.X., K.K.L. and J.C. All authors reviewed and commented on the paper.

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/s41612-023-00359-6.

Correspondence and requests for materials should be addressed to Meng Gao.

Reprints and permission information is available at http://www.nature.com/reprints

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