Analysis of the Diurnal, Weekly, and Seasonal Cycles and Annual Trends in Atmospheric CO$_2$ and CH$_4$ at Tower Network in Siberia from 2005 to 2016

Dmitry Belikov $^{1,*}$, Mikhail Arshinov $^2$, Boris Belan $^2$, Denis Davydov $^2$, Aleksandr Fofonov $^2$, Motoki Sasakawa $^3$ and Toshinobu Machida $^3$

1 Center for Environmental Remote Sensing, Chiba University, 1–33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan
2 V.E. Zuev Institute of Atmospheric Optics, Russian Academy of Sciences, Tomsk 634055, Russia; michael@iao.ru (M.A.); bbd@iao.ru (B.B.); denis@iao.ru (D.D.); alenfo@iao.ru (A.F.)
3 Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Ibaraki 305-8506, Japan; sasakawa.motoki@nies.go.jp (M.S.); tmachida@nies.go.jp (T.M.)

Received: 27 August 2019; Accepted: 29 October 2019; Published: 8 November 2019

Abstract: We analyzed 12 years (2005–2016) of continuous measurements of atmospheric CO$_2$ and CH$_4$ concentrations made at nine tower observation sites in the Japan–Russia Siberian Tall Tower Inland Observation Network (JR-STATION), located in Siberia. Since the data are very noisy and have a low temporal resolution due to gaps in instrument operation, we used the recently developed Prophet model, which was designed to handle the common features of time series (multiple strong seasonalties, trend changes, outliers) and has a robust performance in the presence of missing data and trend shifts. By decomposing each sampled time-series into its major components (i.e., annual trend and seasonal, weekly, and hourly variation), we observed periodically changing patterns of tracer concentrations. Specifically, we detected multi-year variability of tracers and identified high-concentration events. The frequency of such events was found to vary throughout the year, reaching up to 20% of days for some months, while the number of such events was found to be different for CO$_2$ and CH$_4$. An analysis of weather conditions showed that, in most cases, high-concentration events were caused by a temperature inversion and low wind speed. Additionally, wind directions were found to be different for high- and low-concentration events. For some sites, the wind direction indicated the location of strong local sources of CO$_2$ and CH$_4$. As well as elucidating the seasonality of greenhouse gas concentrations, this study confirmed the potential of the Prophet model for detecting periodicity in environmental phenomena.

Keywords: methane; carbon dioxide; greenhouse gas emissions

1. Introduction

Detailed information on the distribution of sources and sinks of the atmospheric greenhouse gases (GHGs) CO$_2$ and CH$_4$ is a prerequisite for analyzing and understanding the role of the carbon cycle within the context of global climate change. Within the context of global climate change, accurate continuous long-term measurements in Siberia are particularly crucial for estimating global CO$_2$ and CH$_4$ budgets due to potential changes in natural emissions from the biosphere, wetlands, thawing lakes, and melting permafrost. Siberia is considered to be one of the world’s largest carbon reservoirs due to its large forest...
Therefore, it is important to accurately measure CO\textsubscript{2} and CH\textsubscript{4} concentrations in this region. In 2002, the National Institute for Environmental Studies (NIES) began a cooperative project for the continuous measurement of greenhouse gases in West Siberia called the Japan–Russia Siberian Tall Tower Inland Observation Network (JR-STATION). The number of observational sites has been progressively expanded, and the network now consists of nine towers, eight located in West Siberia and one located in Yakutsk, East Siberia [3].

Previous studies have performed detailed analysis of data subsets for individual towers in Siberia [3–8] with the highest density of observations. An analysis of the diurnal behavior of CO\textsubscript{2} and CH\textsubscript{4} concentration in West Siberia was described in Arshinov et al. [5]. A variation of CO\textsubscript{2} due to changes in the height of the planetary boundary layer (PBL) was studied in [7]. For a more complete understanding of concentration variations, it is necessary to perform an analysis over a long period for all JR-STATION sites simultaneously. However, such work required advanced statistical methods to handle noisy time-series with large gaps.

The establishment of JR-STATION improved the observation coverage providing a more accurate estimate of middle- and high-latitude fluxes and their inter-annual variability [9–11]. However, only a limited part of the measurements was used, as very strong data selection criteria were applied to avoid assimilating observations caused by large and not normally distributed values of the model representation errors. For example, in [11] all observations for times when the vertical temperature gradient from the lower to upper level was positive indicating atmospheric inversion conditions, and when the wind speed was below 3 m/s. Thus, a large percentage of the data was excluded. For better use of observations in inversions, it is necessary to develop more optimal criteria for filtering outliers, which requires an understanding of the processes of formation of temperature inversions and their influence on the concentration of tracers.

The commonly used time-series analysis methods (e.g., different types of moving average), usually work well but have a tendency to generate large errors when changes happen in the trends. A new time-series forecasting model Prophet, recently developed by Facebook, adopts a generalized additive model to fit the smoothing and forecasting functions [12]. The model released a open source package (available both for R and Python) providing a fitting procedure, interactive and flexible specifications, a robust forecasting function in the presence of outliers, missing data, and shifts in the trend and the ability to model multiple periods of seasonality simultaneously. In most of the experimental results, Prophet performs better than any other approach as shown by the authors. The Prophet model was previously used and evaluated in various analysis including atmospheric studies and air quality assessment [13,14].

The main objective of this study was to explore the daily, weekly, and yearly variability of CO\textsubscript{2} and CH\textsubscript{4} concentrations as observed at JR-STATION over a period of 12 years (2005–2016) using the Prophet model. The observational data and the Prophet model are described in Section 2, the results of the data analysis and a discussion are provided in Section 3, and conclusions are drawn in Section 4.

2. Method

2.1. Data

This study used datasets which were measured at two different elevations at nine towers in JR-STATION, Siberia Azovo (AZV), Berezorechka (BRZ), Demyanskoe (DEM), Igrim (IGR), Karasevoe (KRS), Noyabrsk (NOY), Savvushka (SVV), Vaganovo (VGN), Yakutsk (YAK) as shown in Table 1. Although the BRZ tower was equipped for sampling at four elevations, only data from the two upper elevations were selected. The observation towers were distributed over a wide area encompassing different biome types. The BRZ tower is located in the middle of taiga (boreal forest); the DEM, KRS, and NOY towers are located in a forest zone and are surrounded by extensive wetlands; and the IGR tower is situated in the small town of Igrim (population of about 10,000), which is located next to the Ob River and is surrounded by extensive wetlands. Additionally, three towers—AZV, VGN, and SVV, are located in a steppe region.
The VGN tower is situated 100 km southeast of the city of Chelyabinsk, while the AZV tower is situated 30 km southwest of the city of Omsk; both cities have a population of over 1 million. The SVV tower is located 1 km south of a small village. The final tower, YAK, is located in East Siberia. The locations of the nine towers are depicted in Figure 1, which also shows natural emissions of CH$_4$ (mg/m$^2$/day) derived from CarbonTracker-CH$_4$ assimilation system v. 2016 [15,16]. Note, JR-STATION tower data are not assimilated by CarbonTracker.

Table 1. Tower sites in the Japan–Russia Siberian Tall Tower Inland Observation Network (JR-STATION), Siberia.

| #  | Identifying Code | Location | Lat., ° | Lon., ° | Air Inlet Heights, m | Period of Observations, Years (Data Coverage %) | CO$_2$ | CH$_4$ |
|----|------------------|----------|---------|---------|----------------------|-----------------------------------------------|--------|--------|
| 1  | AZV              | Azovo    | 54.71   | 73.03   | 29, 50               | 9 (79) / 4 (46)                               |        |        |
| 2  | BRZ              | Berezorechka | 56.15  | 84.33   | 40, 80               | 7 (19) / 4 (10)                               |        |        |
| 3  | DEM              | Demyanskoe  | 59.79  | 70.87   | 45, 63               | 12 (68) / 10 (70)                              |        |        |
| 4  | IGR              | Igrim     | 63.19   | 64.41   | 24, 47               | 9 (85) / 10 (82)                               |        |        |
| 5  | KRS              | Karasevoe | 58.25   | 82.42   | 35, 67               | 12 (70) / 10 (74)                              |        |        |
| 6  | NOY              | Noyabrsk  | 63.43   | 75.78   | 21, 43               | 12 (53) / 8 (37)                               |        |        |
| 7  | SVV              | Savvushka | 51.33   | 82.13   | 27, 52               | 9 (58) / 1 (68)                                |        |        |
| 8  | VGN              | Vaganovo  | 54.50   | 62.32   | 42, 85               | 9 (75) / 7 (56)                                |        |        |
| 9  | YAK              | Yakutsk   | 62.09   | 129.36  | 11, 77               | 8 (58) / 6 (28)                                |        |        |

Figure 1. The locations of the measurement towers in the Japan–Russia Siberian Tall Tower Inland Observation Network (JR-STATION), Siberia. Also shown are the natural emissions of CH$_4$ (mg/m$^2$/day) estimated by CarbonTracker-CH$_4$ [15] averaged from 2000 to 2010.
Emissions of CH$_4$ from wetlands have a significant impact on the greenhouse-gas budget of Siberia. As will be shown below, the location of the measurement station relative to the emission zones is the most important factor for determining the temporal pattern of CH$_4$ concentrations.

The detailed description of measurement systems applied for the observation sites is published in [3,17]. In the developed CO$_2$ measurement system the concentration is defined as the mole fraction in dried air using a nondispersive infrared analyzer (model LI-820, LI-COR, USA; a model LI-7000 was used until September 2008 at BRZ). For that purpose, a triple dehumidification system including adiabatic expansion in a glass water trap, a semipermeable membrane dryer (model PD-625–24SS, Permapure, USA), and a magnesium perchlorate trap was employed. With an established standard gas saving system the developed measurement system keeps the analysis precision within 0.3 ppm [17].

For CH$_4$ measurements, a tin dioxide sensor (TOS) was used. This instrument is precise, cost-effective, low power, low carrier-gas consumption, and high mobility. However, the TOS detects methane and other flammable gases (carbon monoxide, hydrogen, and alcohols) and is sensitive to temperature and water vapor in ambient air. In the developed measurement system the sensitivity to temperature and water vapor was reduced by implementing temperature-stabilization and dehydration of the atmospheric samples by a heater unit and low-pressure water trap with chemical desiccant made of P$_2$O$_5$. To avoid the interference of other combustible gases they are removed by an additional catalyst. The overall performance and stability of the TOS sensor for measurements of CH$_4$ in ambient air were validated using a gas chromatograph equipped with a flame ionization detector (GC/FID). The precision of the system is within 3 ppb [3,18].

In addition to CO$_2$ and CH$_4$ concentrations, wind speed and direction (at the high inlet), air temperature and humidity (at all levels), and solar radiation and precipitation (on top of the container laboratory) were measured at the nine JR-STATION towers.

Initially, we planned to analyze hourly averages of atmospheric CO$_2$ and CH$_4$ mixing ratios collected at the nine tower sites between 1 January 2005 and 31 December 2016 (12 years); however, such a long dataset was not available for all towers (Table 1). Although the equipment was set up to conduct observations every hour, this was not always technically possible, and the datasets therefore contain many gaps. Hourly measurements of CO$_2$ and CH$_4$ cover about 50–80% and 40–70% of the 12 year period, respectively. Significantly fewer measurements are available for the following sites: BRZ (CO$_2$ and CH$_4$), AZV (CH$_4$), SVV (CH$_4$), and YAK (CH$_4$) (Table 1). Such extremely low coverage could affect the analysis and distort the results. Therefore, to avoid this, we excluded from consideration the two datasets with the lowest coverage, namely BRZ (CH$_4$) and SVV (CH$_4$).

Another issue which could affect the results are the frequent and large fluctuations in concentration due to variations in emissions. The daily mean emissions of CH$_4$ from Siberian wetlands have been estimated at between 10 and 500 mg/m$^2$/day [19–22]. The daily variation in CH$_4$ concentration is much larger than the seasonal variation (e.g., for KRS, see Figure 2e).
2.2. The Prophet Model

The Prophet model was designed for the analysis and forecasting of time-series data based on an additive model with three main model components: trend, seasonality, and holidays [23]. These are combined in the following equation [12]:

$$y(t) = g(t) + s(t) + h(t) + e(t)$$

(1)

where $g(t)$ is the trend function, $s(t)$ represents periodic changes (e.g., weekly and yearly seasonality), and $h(t)$ represents the effects of holidays which occur on potentially irregular schedules over one or more days. The error term $e(t)$ represents any idiosyncratic changes which are not accommodated by the model. The parametric assumption of [12] shows that $e(t)$ is normally distributed.

Compared to traditional exponential smoothing models, the Prophet model can more easily handle temporal patterns with multiple periods and has no requirements regarding the regularity of measurement spacing. It works best with time-series that have strong seasonal effects and with several seasons of historical data. The Prophet model has a robust performance in the presence of missing data and trend shifts and typically handles outliers well [12]. The large number of gaps that exist in the JR-STATION measurements make the Prophet model particularly suitable to simultaneously reveal the daily, weekly, and monthly seasonality from the hourly time-series.

Prophet includes two trend models $g(t)$ that cover many applications: a saturating growth model, and a piecewise linear model. The logistic growth model in its most basic form is:

$$g(t) = \frac{C}{1 + \exp(-k(t-m))}$$

(2)
where \( C \) is the carrying capacity, \( k \) the growth rate, and \( m \) is an offset parameter.

There are two important aspects of growth that cannot be captured in Equation (2). First, the carrying capacity is not constant. Thus the fixed capacity \( C \) is replaced with a time-varying capacity \( C(t) \). Second, the growth rate is not constant. The model must be able to incorporate a varying rate in order to fit historical data. The piecewise logistic growth model is used to overcome those issues:

\[
g(t) = \frac{C(t)}{1 + \exp(-(k + a(t)T\delta)(t - (m + a(t)T\gamma)))},
\]

where the rate at any time \( t \) is the base rate \( k \), plus all of the adjustments up to that point: \( k + a(t)T\delta \). The correct adjustment at changepoint is easily computed as \( a(t)T\gamma \) [12].

To provide flexibility for periodic effects, the Prophet model uses the Fourier series [24]:

\[
s(t) = \sum_{n=1}^{N} \left( b_n \cos \left( \frac{2\pi nt}{P} \right) + b_n \sin \left( \frac{2\pi nt}{P} \right) \right)
\]

where \( t \) denotes the time and \( P \) represents the regular period that the time-series is expected to have (seven and 365.25 days for the weekly and yearly periods, respectively). Fitting seasonality requires the estimation of the \( 2N \) parameters \( \beta = [a_1, b_1, ..., a_N, b_N]^T \). This was done by constructing a matrix of seasonality vectors for each value of \( t \) in the historical data.

3. Results and Discussion

The Prophet model was implemented to analyze the time-series measured at towers in the JR-STATION, Siberia (Table 1). The trend and seasonal variation of the time-series analyzed in the present study have been analyzed for individual sites in past publications. Therefore, here we restricted the analysis to comparing \( \text{CO}_2 \) and \( \text{CH}_4 \) data between sites while paying more attention to short-term changes [3–8]. To analyze annual trends, and seasonal and weekly variation we use daytime (11:00–18:00 local time) observations only, while for daily variation all measurements were used.

The Prophet model is implemented through the Python package [25]. The flexibility of the trend is controlled by adjusting the number of changepoints automatically selected. The periods are set to 365.25 and 7 days. The model’s default values for \( N \) (i.e., \( N = 10 \) and \( N = 3 \) for modeling yearly and weekly components, respectively) were tested first. Though these values have been empirically demonstrated to work well for most practical situations [12], in this work \( N = 10 \) produces inexplicably large variability in seasonal cycles, so empirically we reduced \( N \) to 5.

3.1. Trend

The calculated \( \text{CO}_2 \) trends (Figure 3a) show a stable growth of about 2 ppm/yr, which corresponds to the global value [26]. The small differences in the \( \text{CO}_2 \) growth rate that are observed between sites are due to the different lengths of the datasets, the different heights of the sampling inlets, and the influence of local sources. For methane concentrations, the length of the dataset and the number of gaps are important. Therefore, the relatively short SVV dataset was excluded from the analysis, while the AZV, BRZ, and VGN datasets were interpreted with great care.

The observation towers can be grouped based on the average value of the \( \text{CH}_4 \) concentration (Figure 3b) measured at each one. The highest values of \( \text{CH}_4 \) concentration were measured at two towers located in the central part of the Vasyugan Swamp (IGR and NOY); medium values were measured at towers located on the outskirts of the Vasyugan Swamp (DEM, KRS); and low values were found at the towers located in West Siberia and furthest from the Vasyugan Swamp (AZV, BRZ, and VGN). The smallest value of \( \text{CH}_4 \) concentration was measured at YAK, in East Siberia. Measurements from
ground-based networks show that, between 1999 and 2006, atmospheric CH$_4$ concentrations were nearly constant [27]. From 2007 to 2013, the globally averaged mole fraction of CH$_4$ in the atmosphere increased by 5.7 ± 1.2 ppb/yr [28]. Most of the CH$_4$ time-series measured at the JR-STATION towers follow this tendency; however, VGN showed a faster increase after 2012 (as captured by observations Figure 2g), and YAK showed a slower increase for all years.

![Figure 3. Trends of concentrations of (a) CO$_2$ and (b) CH$_4$ for the JR-STATION tower sites used in this study for the period 2005–2015.](image_url)

3.2. Seasonal Variation

Measured concentration time-series often have multi-period seasonality as a result of natural and human activity. For instance, the biosphere produces effects that repeat every year, while industrial activities related to a five-day working week can produce effects that repeat every week. Periodic seasonality algorithms should be used to fit such effects. As stated in [12], unlike other statistical models, with the Prophet model, measurements do not need to be regularly spaced in time, and missing values do not need to be interpolated. To derive periodic signals the model uses all the observations for the study period (12 years), thereby increasing accuracy and reducing sensitivity to random outliers, for example generated by forest fires [29]. For some sites, such a strategy is the only one possible to identify the seasonal cycle if the data for different seasons refer to different years. Thus, the derived seasonal variation is generalized and different from that obtained in previous works (i.e., [5,7]).

In the Northern Hemisphere, CO$_2$ has been shown to exhibit a prominent seasonal cycle with a minimum in July–August and a maximum in November–December, reflecting the metabolic cycles of the land biota. In the present study, the seasonal difference in CO$_2$ concentration was found to reach up to 30 ppm, which is consistent with previous estimations [4,5,7].

Unlike CO$_2$, CH$_4$ concentrations in the Northern Hemisphere are characterized by peaks not only in winter (predominantly due to anthropogenic sources [30]) but also in late summer, caused by emissions from wetlands [5,31]. The CH$_4$ maximum that was observed in July–August in the present study is consistent with the majority of land surface models used in the inter-comparison study of [21].

The two stations that are furthest from the wetland region (VGN, YAK) showed the lowest summer concentrations of CH$_4$. A summer peak in CH$_4$ concentration was measured at KRS, which is located at the border of the lowlands on the downwind side. The amplitude and phase of the CH$_4$ concentrations measured at DEM, IGR, and KRS are in good accordance with the results of [3].
3.2.1. Case Study: Winter Vacation

For most of the measurement sites, a marked decrease in CO$_2$ and CH$_4$ concentrations was observed in late December and the beginning of January (Figure 4a,b). In winter, anthropogenic emissions are the predominant factor behind CO$_2$ and CH$_4$ emissions. Therefore, we propose that the decline in concentration in winter is due to reduced industrial and transport activity caused by winter holidays (which usually last about one week [32]).

![Figure 4.](image)

Figure 4. Multi-year (2005–2015) seasonal variation of concentrations of (a) CO$_2$ and (b) CH$_4$ for the considered sites.

To test this hypothesis, we analyzed the concentrations of tracers in the first 10 day period (1–10 January) and second 10 day period (11–20 January) of January for all available years. We assumed that the first period, which we termed “Holiday”, was associated with low human (industrial and transport) activity, while the second period, which we termed “Working”, was associated with normal human activity. However, this division is rather arbitrary, since the length of the holidays varies from year to year. Moreover, some production cycles of enterprises do not always strictly follow the official working calendar. Some industrial processes include preliminary stages with a low load, and therefore produce lower emissions. Additionally, the schedules for holidays and working days may shift during this period.

After removing the trends, the average values of CO$_2$ and CH$_4$ concentration were calculated for the "Holiday" and “Working” periods, respectively (Table 2). Student’s t-test was used to estimate the statistical significance of the difference between "Holiday" and “Working” datasets. For only a few cases did the test indicate a non-significant difference: NOY and SVV for CO$_2$, AZV, IGR, and YAK for CH$_4$. The results suggest that, relative to CH$_4$, the CO$_2$ concentration is less sensitive to changes in industrial activity; the difference in CO$_2$ concentration between the two periods does not exceed 1%, while for two stations the difference is negative. This could be due to the intensive heating of private countryside houses or outdoor activities near to the measurement sites.
Table 2. The average difference in CO$_2$ and CH$_4$ concentrations observed at the JR-STATION tower sites between the “Holiday” period (1–10 January) and the “Working” period (11–20 January), “Diff.” is difference between the “Holiday” and “Working”. For BRZ and SVV, CH$_4$ time-series were not long enough for analysis. Bold font was used to mark statistically insignificant (estimated using Student’s t-test) values of difference.

| #  | Identifying Code | Mean CO$_2$, ppm | Mean CH$_4$, ppb | Diff. (%) | Working | Holiday | Diff. (%) | Working | Holiday | Diff. (%) |
|----|------------------|------------------|------------------|-----------|---------|---------|-----------|---------|---------|-----------|
| 1  | AZV              | 10.52            | 8.86             | 1.67 (15.85) | 113.84  | 120.09  | -6.25 (−5.49) |
| 2  | BRZ              | 13.71            | 5.89             | 7.81 (57.00) | -       | -       | - ( - ) |
| 3  | DEM              | 8.41             | 6.83             | 1.59 (18.87) | 60.76   | 10.76   | 50.00 (82.29) |
| 4  | IGR              | 9.07             | 7.54             | 1.53 (16.91) | 31.91   | 21.45   | 10.46 (32.80) |
| 5  | KRS              | 8.84             | 6.35             | 2.49 (28.14) | 42.90   | 6.60    | 36.30 (84.62) |
| 6  | NOY              | 8.78             | 9.05             | -0.28 (−3.15) | 67.43   | 20.46   | 46.98 (69.66) |
| 7  | SVV              | 6.87             | 7.27             | -0.39 (−5.76) | -       | -       | - ( - ) |
| 8  | VGN              | 7.52             | 8.16             | -0.65 (−8.58) | 54.63   | 41.24   | 13.39 (24.51) |
| 9  | YAK              | 8.17             | 6.71             | 1.46 (17.91) | 52.49   | 46.93   | 5.56 (10.59) |

Figure 5. Concentrations of CO$_2$ and CH$_4$ versus air temperature (a, b) and atmospheric pressure (c, d) for measurement stations DEM and KRS.

Strong high-pressure systems with low air temperature have been observed over Siberia in winter [3], which could potentially affect the measured greenhouse gas concentrations. To exclude the influence of such systems, the distribution of tracers with respect to air temperature (Figure 5a, b) and atmospheric pressure (Figure 5c, d) was considered. The 12 year datasets used for the present study were sufficiently large to analyze greenhouse gas concentrations over a wide range of temperature and pressure. For some sites with decreasing temperature (e.g., DEM), the difference in CO$_2$ concentration between the “Holiday” and “Working” periods was markedly higher, as more fuel is consumed for heating and transport in the surrounding areas. Therefore, strong high-pressure systems with low air temperature can enhance the effect, but are not the main factors.
3.3. Weekly Variation

The analysis of weekly patterns is useful to identify the tracer emissions associated with industry or transport operating on a five-day working week. Our analysis did not reveal any noticeable features in the distribution of CO$_2$ concentrations at a weekly scale (Figure 6a). For all sites, the amplitude of the change in CO$_2$ concentration was less than 1.5 ppm.

For CH$_4$, the amplitude of the observed change in concentration was larger. However, no clear weekly variation was observed (Figure 6b). This suggests that there were no large anthropogenic sources of CH$_4$ that constantly affected the concentrations measured at the stations at intervals corresponding to the five-day working week.

![Figure 6. Multi-year (2005–2015) weekly variations of concentrations of (a) CO$_2$ and (b) CH$_4$ derived using the Prophet model for the considered measurement sites.](image)

3.4. Daily Variation

The considered measurement sites are located in the interior of the Asian continent. At such locations, a strong daily change in the height of the planetary boundary layer (PBL) is common. The daily cycle of tracers is most pronounced in summer during dry anticyclonic weather conditions. In this case, frequent surface temperature inversions develop at night and last until the early morning, which impedes vertical atmospheric mixing and traps surface gas emissions in the stable shallow PBL [33]. Concentrations reach maxima in the early morning between 03:00 and 06:00 local time. Minimum concentrations of trace gases emitted from the soil are normally observed between 12:00 and 18:00, when solar heating of the ground induces vertical air turbulence; such turbulence leads to mixing in a deep layer of the atmosphere (approximately 200–600 m deep in winter and up to 2800 m deep in summer [4]).

The magnitude of the nocturnal increase in greenhouse-gas concentrations is variable and depends on the strength of the atmospheric inversion and on the intensity of the regional tracer source. Overcast conditions appear to slow the morning breakup of this inversion so that tracers do not move from the surface layer to high altitudes until 07:00–09:00 [34]. Since precipitation data are also available for the measurement stations used in the present study, it would be interesting to analyze this effect, however, it was not considered in this work.

As shown in Figure 7, concentrations of CO$_2$ and CH$_4$ clearly changed throughout the day at all measurement sites. For example, the time-series for NOY is characterized by an evening peak of CH$_4$ concentration (Figure 7b).

The minimum amplitude of CO$_2$ and CH$_4$ daily concentration change (maximum minus minimum concentration) are observed in winter (December–February) while the maximum amplitudes are observed
in summer (June–August), as shown in Table 3. Autumn amplitude values (September–November) are larger than the spring ones (March–May). The mean diurnal variation in CH$_4$ concentration for each month at the KRS site is shown in [3].

![Image of Figure 7](image_url)

**Figure 7.** Multi-year (2005–2015) daily variation of the concentrations of (a) CO$_2$ and (b) CH$_4$ derived by the Prophet model for the considered measurement stations.

**Table 3.** Multi-year (2005–2015) amplitudes of the daily variation of CO$_2$ and CH$_4$ concentrations for different seasons for JR-STATION sites in Siberia derived by the Prophet model. Methane data for Yakutsk were not sufficient to assess variations for December-February (DJF), March–May (MAM), June–August (JJA), and September–November (SON).

| # | Identifying Code | Amplitude for CO$_2$, ppm | | | Amplitude for CH$_4$, ppb | | |
|---|---|---|---|---|---|---|---|
| 1 | AZV | 1.04 | 5.46 | 26.12 | 5.20 | 21.19 | 50.68 |
| 2 | BRZ | 1.24 | 4.76 | 23.37 | 4.54 | - | - |
| 3 | DEM | 0.61 | 3.31 | 20.45 | 6.63 | 10.73 | 21.19 |
| 4 | IGR | 1.01 | 3.66 | 17.90 | 9.74 | 31.27 | 36.46 |
| 5 | KRS | 0.85 | 3.29 | 22.07 | 7.09 | 11.11 | 20.06 |
| 6 | NOY | 0.24 | 1.48 | 10.53 | 3.48 | - | - |
| 7 | SVV | 0.95 | 4.12 | 11.89 | 3.48 | - | - |
| 8 | VGN | 0.29 | 5.45 | 17.56 | 3.87 | 6.39 | 17.68 |
| 9 | YAK | 6.73 | 6.25 | 12.46 | 17.92 | - | 35.04 |

### 3.5. Diagnostic

Prophet includes functionality of time series nested cross-validation to measure the fitting error using historical data. This is done by selecting cutoff points in the history, and for each of them, fitting the model using data only up to that cutoff point. We can then compare the fitted values to the actual values. This method, called the simulated historical forecasts (SHFs), is based on classical “rolling origin” forecast evaluation procedures [35]. We specify the forecast horizon (365 days), the size of the initial training period (730 days) and the spacing between cutoff dates—period (180 days). Due to the limited number of observations the initial training period for CH$_4$ for AZV and YAK was reduced to 365 days. Table 4 shows the mean statistics of the fitting performance. The statistics computed are root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percent error (MAPE).

The average error of the model for CO$_2$ is fairly uniform and does not exceed one percent. For CH$_4$, the error strongly depends on the location relative to the emission regions (swamps), the length of the time series and the number of observations. Values are in the range of 2–5% (Figure 8). By default Prophet will
only return uncertainty in the trend and observation noise. To get uncertainty in seasonality, it needs to do full Bayesian sampling. We tried different options of the model tuning, but could not determine a method to convert data-density into a calculation of the uncertainty.

![Figure 8. Time-series of CO$_2$ (a–d) and CH$_4$ (e–h) for the KRS, NOY, VGN, and YAK towers for 2011. Black symbols represent daytime averaged measurements, the blue line is the Prophet model fitting, and the shaded area shows the 80% confidence interval.](image)

**Figure 8.** Time-series of CO$_2$ (a–d) and CH$_4$ (e–h) for the KRS, NOY, VGN, and YAK towers for 2011. Black symbols represent daytime averaged measurements, the blue line is the Prophet model fitting, and the shaded area shows the 80% confidence interval.

**Table 4.** The Prophet model cross-validation performance metrics at tower network sites in Siberia (JR-STATION). For BRZ and SVV CH$_4$ data series are not sufficient for analysis. Root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percent error (MAPE).

| # | Identifying Code | CO$_2$ | CH$_4$ |
|---|------------------|--------|--------|
|   |                  | RMSE   | MAE    | MAPE (%) | RMSE   | MAE    | MAPE (%) |
| 1 | AZV              | 4.97   | 3.84   | 0.96     | 48.67  | 39.96  | 2.03     |
| 2 | BRZ              | 4.43   | 3.36   | 0.85     | -      | -      | -        |
| 3 | DEM              | 3.97   | 3.00   | 0.76     | 108.77 | 65.83  | 3.32     |
| 4 | IGR              | 4.92   | 3.63   | 0.92     | 120.91 | 93.78  | 4.57     |
| 5 | KRS              | 4.36   | 3.33   | 0.84     | 47.64  | 35.35  | 1.79     |
| 6 | NOY              | 3.76   | 2.85   | 0.72     | 98.66  | 74.19  | 3.64     |
| 7 | SVV              | 3.91   | 2.98   | 0.76     | -      | -      | -        |
| 8 | VGN              | 4.25   | 3.34   | 0.84     | 57.10  | 42.32  | 2.10     |
| 9 | YAK              | 4.76   | 3.63   | 0.91     | 97.49  | 73.06  | 3.79     |
3.6. Temperature Inversion

In order to analyze the effect of temperature inversions in more detail, we examined the difference in the concentrations of CO$_2$ and CH$_4$ measured from the lower and upper inlets of the measurement towers for cases with and without inversions (Figure 9). The results show that, for all sites except YAK, inversions have a strong effect on CO$_2$ concentrations in summer and a relatively weak effect in other seasons (Figure 9d). In the YAK area, temperature inversions consistently form in winter due to the extremely cold and dense air of the Siberian High pooling in deep hollows. In this area, the average minimum temperatures for January, February, and December occasionally fall below $-50$ °C.

The influence of inversions on CH$_4$ concentrations is more diverse: For some sites (e.g., KRS; Figure 9e), inversions have as prominent an effect as they do for CO$_2$ concentrations, while for others (VGN and YAK; Figure 9g,h), concentrations are more uniform throughout the year. However, for two stations located in the very center of the Vasyugan Swamp (IGR and NOY), we did not observe large differences in CO$_2$ or CH$_4$ concentrations during temperature inversion (Figure 9f). CH$_4$ emission is strong throughout the year around NOY.

![Figure 9](image_url)  
*Figure 9. Monthly mean CO$_2$ (a–d) and CH$_4$ (e–h) concentration difference between lower and upper inlets for events with (Positive) and without (Negative) temperature inversion for KRS, NOY, VGN, and YAK sites, respectively.*

Carbon monoxide (CO) emissions usually result from inefficient combustion, and have often been used as a tracer for CO$_2$ [36]. The ratios of the emissions of CO$_2$ to CO are different for various combustion and biological processes. Therefore, the simultaneous observation of CO$_2$ and CO may help to distinguish between fossil fuel sources and terrestrial ecosystem sources of CO$_2$ [37].

3.7. Elevated Concentration Events

For many types of study (e.g., inverse modeling [9,11,38]), it is important to understand the reason for elevated greenhouse gas concentrations and filter out concentrations caused by local sources or temperature inversions. For a high concentration, we consider the events when the threshold value of concentration is exceeded for three hours or more. The threshold values are selected to be equal to the magnitude of the uncertainty in the concentration calculated by the Prophet model. The threshold values for the considered measurement sites are given in Table 5.

During the study period, the frequency of high-concentration events was 15–20%. The seasonal distribution of these events was not uniform, and is likely to correspond to seasonal variations in
emissions. For CO\textsubscript{2} concentrations, peaks are observed in summer, while for CH\textsubscript{4}, summer and winter peaks are observed (Figure 10). For late spring and late autumn, the frequency of events is low. For high-concentration events, the difference in temperature between the upper and lower inlets was analyzed. Since the locations of the inlets are different for different towers, we considered the scaled temperature difference, $\Delta T_{\text{scaled}}$, scaled by the height between inlets, $\Delta H$, using the formula:

$$
\Delta T_{\text{scaled}} = \frac{\Delta T}{\Delta H} \times 10.
$$

For the IGR, NOY, and VGN towers, in most cases, elevated concentrations are accompanied by a temperature inversion of 0.1–0.3 °C/m. This is also sometimes observed at the KRS tower. At the YAK tower, temperature inversions reached a maximum of 0.5 °C/m during December–January. The strongest summer inversion was observed at the DEM tower.

![Figure 10. Number of days (%) with elevated concentrations of CO\textsubscript{2} (a–d) and CH\textsubscript{4} (e–h) for the KRS, NOY, VGN, and YAK measurement sites.](image)

**Table 5.** The threshold concentration values for high-concentration events for the considered measurement sites.

| #  | Identifying Code | Threshold CO\textsubscript{2}, ppm | Threshold CH\textsubscript{4}, ppb |
|----|-----------------|-------------------------------|----------------------------------|
| 1  | AZV             | 9.33                          | 45.00                            |
| 2  | BRZ             | 7.81                          | -                                |
| 3  | DEM             | 7.47                          | 88.78                            |
| 4  | IGR             | 8.45                          | 160.51                           |
| 5  | KRS             | 7.87                          | 98.08                            |
| 6  | NOY             | 5.57                          | 141.93                           |
| 7  | SVV             | 5.98                          | -                                |
| 8  | VGN             | 7.04                          | 53.49                            |
| 9  | YAK             | 6.32                          | 39.05                            |

In addition to temperature inversions, low wind speeds have been shown to significantly affect the concentration of tracers in Siberia [39]. Additionally, wind direction also greatly affects the concentration of tracers in Siberia if local sources are located near to the observation site. To identify the effect of wind speed and direction on concentrations of CO\textsubscript{2} and CH\textsubscript{4}, we also analyzed events with low tracer concentrations (i.e., concentrations less than the threshold value). A comparison of wind roses for high- and low-concentration events is presented in Figure 11.
From the figure, it is clear that the proposed approach allows wind speeds and directions to be determined for situations with high and low concentrations of CO$_2$ and CH$_4$. Since the towers’ measurement inlets are located fairly high above the ground, the frequency of calm conditions was very low.

Low wind speeds clearly contribute to the increase of tracer concentrations. However, on the contrary, high wind speeds lead to the faster mixing of air volumes and thus lower tracer concentrations. Importantly, we successfully identified the wind direction which brought more tracer-contaminated air. For the YAK and NOY towers, this corresponded to a fairly narrow range of wind directions (Figure 11). It is very likely that the concentrations at the NOY tower were influenced by CO$_2$ and CH$_4$ emissions from the oil and natural gas pipelines which are located several km northwest of the tower [3]. Similarly, the concentrations at the VGN tower were likely influenced by Chelyabinsk, one of the major industrial centers of Russia.

![Wind roses](image)

**Figure 11.** Wind roses depicting wind speed (m/s) and wind direction for high- and low-concentration CH$_4$ events for the considered measurement sites.
4. Conclusions

In this study, we present a time-series analysis of composite CO\textsubscript{2} and CH\textsubscript{4} concentration records from nine tower measurement sites in Siberia. The use of the Prophet model facilitates the analysis of a time-series of tracer concentrations and meteorological parameters. It helps to extract trends and periodic signals with different time scales and to filter out noise and outliers from datasets. Although it is challenging to quantify local sources and sinks of CO\textsubscript{2} and CH\textsubscript{4}, we detected multi-year variability of tracers and identified events of elevated concentration. The periodicity of such events was found to vary throughout the year and to be different for CO\textsubscript{2} and CH\textsubscript{4}, reaching up to 20\% of days in some months. Furthermore, the weather conditions associated with elevated concentrations were analyzed. In most cases, elevated concentration events were caused by a temperature inversion and low wind speed. It was revealed that certain wind directions were correlated with high- and low-concentration events. For some sites, the wind direction was found to indicate the location of the sources of CO\textsubscript{2} and CH\textsubscript{4}. To identify the type of source (natural or anthropogenic), additional information (e.g., CO mixing ratio) is needed. Currently, fairly simple criteria are used to filter observational data, such as excluding nighttime data and data obtained during low wind speeds and temperature inversions [9,11]. This analysis can be used to develop advanced criteria to filter observations used in inverse modeling.

Author Contributions: Conceptualization, methodology, formal analysis, and original draft preparation: D.B.; discussion, writing review and editing: M.A., B.B., D.D., A.F., M.S., and T.M.; preparation of measurement equipment, observation sampling and data primary processing: M.A., B.B., D.D., A.F., M.S. and T.M.

Funding: This research was supported by the Environment Research and Technology Development Fund (2-1802) of the Environmental Restoration and Conservation Agency of Japan. M.A., B.B., D.D., and A.F. are partly supported by the Ministry of Education and Science of the Russian Federation under State Contract No 14.613.21.0082 (ID No RFMEFI61317X0082).

Acknowledgments: The used data are available from the Global Environmental Database, hosted by Center for Global Environmental Research (CGER), National Institute for Environmental Studies (NIES) (http://db.cger.nies.go.jp/portal/geds/index). CarbonTracker CT2016 results provided by National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL), Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- CO\textsubscript{2} carbon dioxide
- CH\textsubscript{4} methane

References

1. Kurganova, I.; Kudeyarov, V.; Gerenyu, L.D. Updated estimate of carbon balance on Russian territory. Tellus Chem. Phys. Meteorol. 2010, 62, 497–505. [CrossRef]
2. Le Quéré, C.; Andrew, R.M.; Friedlingstein, P.; Sitch, S.; Hauck, J.; Pongratz, J.; Pickers, P.A.; Korsbakken, J.I.; Peters, G.P.; Canadell, J.G.; et al. Global Carbon Budget 2018. Earth Syst. Sci. Data 2018, 10, 2141–2194. [CrossRef]
3. Sasakawa, M.; Shimoyama, K.; Machida, T.; Tsuda, N.; Suto, H.; Arshinov, M.; Davydov, D.; Fofonov, A.; Krasnov, O.; Saeki, T.; et al. Continuous measurements of methane from a tower network over Siberia. Tellus Chem. Phys. Meteorol. 2010, 62, 403–416. [CrossRef]
4. Lloyd, J.; Langenfelds, R.L.; Francey, R.J.; Gloer, M.; Tchebakova, N.M.; Zolotoukhine, D.; Brand, W.A.; Werner, R.A.; Jordan, A.; Allison, C.A.; et al. A trace-gas climatology above Zotino, central Siberia. Tellus Chem. Phys. Meteorol. 2002, 54, 749–767. [CrossRef]
5. Arshinov, M.Y.; Belan, B.; Davydov, D.; Inoue, G. Spatio-temporal variability of CO\textsubscript{2} and CH\textsubscript{4} concentration in the surface atmospheric layer over West Siberia. Opt. Atmos. Okeana 2009, 22, 183–192.

6. Arshinov, M.Y.; Belan, B.; Davydov, D.; Inouye, G.; Maksyutov, S.; Machida, T.; Fofonov, A. Vertical distribution of greenhouse gases above Western Siberia by the long-term measurement data. Atmos. Ocean. Opt. 2009, 22, 316–324. [CrossRef]

7. Sasakawa, M.; Machida, T.; Tsuda, N.; Arshinov, M.; Davydov, D.; Fofonov, A.; Krasnov, O. Aircraft and tower measurements of CO\textsubscript{2} concentration in the planetary boundary layer and the lower free troposphere over southern taiga in West Siberia: Long-term records from 2002 to 2011. J. Geophys. Res. Atmos. 2013, 118, 9489–9498. [CrossRef]

8. Sasakawa, M.; Machida, T.; Ishijima, K.; Arshinov, M.; Patra, P.; Ito, A.; Aoki, S.; Petrov, V. Temporal Characteristics of CH\textsubscript{4} Vertical Profiles Observed in the West Siberian Lowland Over Surgut From 1993 to 2015 and Novosibirsk From 1997 to 2015. J. Geophys. Res. Atmos. 2017, 122, 11–261. [CrossRef]

9. Saeki, T.; Maksyutov, S.; Sasakawa, M.; Machida, T.; Arshinov, M.; Tans, P.; Conway, T.; Saito, M.; Valsala, V.; Oda, T.; et al. Carbon flux estimation for Siberia by inverse modeling constrained by aircraft and tower CO\textsubscript{2} measurements. J. Geophys. Res. Atmos. 2013, 118, 1100–1122. [CrossRef]

10. Berchet, A.; Pison, I.; Chevallier, F.; Paris, J.D.; Bousquet, P.; Bonne, J.L.; Arshinov, M.Y.; Belan, B.; Cressot, C.; Davydov, D.; et al. Natural and anthropogenic methane fluxes in Eurasia: A mesoscale quantification by generalized atmospheric inversion. Biogeosciences 2015, 12, 5393–5414. [CrossRef]

11. Thompson, R.L.; Sasakawa, M.; Machida, T.; Aalto, T.; Worthy, D.; Lavric, J.V.; Lund Myhre, C.; Stohl, A. Methane fluxes in the high northern latitudes for 2005–2013 estimated using a Bayesian atmospheric inversion. Atmos. Chem. Phys. 2017, 17, 3553–3572. [CrossRef]

12. Taylor, S.J.; Letham, B. Forecasting at scale. Am. Stat. 2018, 72, 37–45. [CrossRef]

13. Papacharalampous, G.; Tyralis, H.; Koutsoyiannis, D. Predictability of monthly temperature and precipitation using automatic time series forecasting methods. Acta Geophys. 2018, 66, 807–831. [CrossRef]

14. Zhao, N.; Liu, Y.; Vanos, J.K.; Cao, G. Day-of-week and seasonal patterns of PM2.5 concentrations over the United States: Time-series analyses using the Prophet procedure. Atmos. Environ. 2018, 192, 116–127. [CrossRef]

15. Carbontracker-ch4. Available online: http://www.esrl.noaa.gov/gmd/ccgg/carbontracker-ch4/ (accessed on 1 November 2019).
24. Harvey, A.C.; Shephard, N. 10 Structural time series models. In Econometrics; Handbook of Statistics; Elsevier: Amsterdam, The Netherlands, 1993; Volume 11, pp. 261–302. [CrossRef]
25. The Prophet Package. Available online: https://facebook.github.io/prophet/ (accessed on 1 November 2019).
26. Annual Mean Growth Rate for Mauna Loa, Hawaii. Available online: https://www.esrl.noaa.gov/gmd/ccgg/trends/gr.html (accessed on 1 November 2019).
27. Dlugokencky, E.; Bruhwiler, L.; White, J.; Emmons, L.; Novelli, P.C.; Montzka, S.A.; Masarie, K.A.; Lang, P.M.; Crotwell, A.; Miller, J.B.; et al. Observational constraints on recent increases in the atmospheric CH$_4$ burden. Geophys. Res. Lett. 2009, 36. [CrossRef]
28. Nisbet, E.; Dlugokencky, E.; Manning, M.; Lowry, D.; Fisher, R.; France, J.; Michel, S.; Miller, J.; White, J.; Vaughn, B.; et al. Rising atmospheric methane: 2007–2014 growth and isotopic shift. Glob. Biogeochem. Cycles 2016, 30, 1356–1370. [CrossRef]
29. Arshinov, M.; Afonin, S.; Belan, B.; Belov, V.; Gridnev, Y.; Davydov, D.; Nedelec, P.; Paris, J.D.; Fofonov, A. Comparison between satellite spectrometric and aircraft measurements of the gaseous composition of the troposphere over Siberia during the forest fires of 2012. Izv. Atmos. Ocean. Phys. 2015, 50, 916–928. [CrossRef]
30. Umezawa, T.; Machida, T.; Aoki, S.; Nakazawa, T. Contributions of natural and anthropogenic sources to atmospheric methane variations over western Siberia estimated from its carbon and hydrogen isotopes. Glob. Biogeochem. Cycles 2012, 26. [CrossRef]
31. Worthy, D.; Levin, I.; Trivett, N.; Kuhlmann, A.; Hopper, J.; Ernst, M. Seven years of continuous methane observations at a remote boreal site in Ontario, Canada. J. Geophys. Res. Atmos. 1998, 103, 15995–16007. [CrossRef]
32. Oberlander, E.A.; Brenninkmeijer, C.; Crutzen, P.; Elansky, N.; Golitsyn, G.; Granberg, I.; Scharffe, D.; Hofmann, R.; Belikov, I.; Paretzke, H.; et al. Trace gas measurements along the Trans-Siberian railroad: The TROICA 5 expedition. J. Geophys. Res. Atmos. 2002, 107. [CrossRef]
33. Tashman, L.J. Out-of-sample tests of forecasting accuracy: An analysis and review. Int. J. Forecast. 2000, 16, 437–450. [CrossRef]
34. Turnbull, J.; Miller, J.; Lehman, S.; Tans, P.; Sparks, R.; Southon, J. Comparison of 14CO$_2$, CO, and SF$_6$ as tracers for recently added fossil fuel CO$_2$ in the atmosphere and implications for biological CO$_2$ exchange. Geophys. Res. Lett. 2006, 33, L01817. [CrossRef]
35. Baklanov, A.A.; Penenko, V.V.; Mahura, A.G.; Vinogradova, A.A.; Elansky, N.F.; Tsvetova, E.A.; Rigina, O.Y.; Maksimenkov, L.O.; Nuterman, R.B.; Pogarskii, F.A.; et al. Aspects of atmospheric pollution in Siberia. In Regional Environmental Changes in Siberia and Their Global Consequences; Springer: Berlin/Heidelberg, Germany, 2013; pp. 303–346.