10-20-1998

Flux Measurements and Sampling Strategies: Applications to Methane Emissions from Rice Fields

M. A. K. Khalil
aslamk@pdx.edu

R. A. Rasmussen
Oregon Graduate Institute

Martha J. Shearer
Portland State University

Citation Details

Khalil, M. A. K., R. A. Rasmussen, and M. J. Shearer (1998), Flux measurements and sampling strategies: Applications to methane emissions from rice fields, J. Geophys. Res., 103(D19), 25,211–25,218, doi:10.1029/98JD00690.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. D19, PAGES 25,211-25,218, OCTOBER 20, 1998

Flux measurements and sampling strategies: Applications to methane emissions from rice fields

M.A.K. Khalil
Department of Physics, Portland State University, Portland, Oregon

R.A. Rasmussen
Department of Environmental Science and Engineering, Oregon Graduate Institute, Portland

M. J. Shearer
Department of Physics, Portland State University, Portland, Oregon

Abstract. The emissions of methane from rice fields and other sources are often measured by placing chambers on the surface and taking sequential samples. Although static chambers pose several problems that affect the accuracy of the data, there are a few parameters that, if carefully chosen, can improve the reliability of the data and reduce the uncertainties. These parameters are the length of time the chamber is kept on the rice plants, the number of samples that are drawn to estimate the flux, the basal area and height of the chamber, the frequency of measurements during the growing season, and the number of plots sampled. In this paper we analyze a large data set to determine how these parameters can be chosen to improve data quality. The results show that, for individual flux measurements, extending the time the chambers are left on the plots improves precision more effectively than taking more sequential samples for each flux measurement. The exposure time cannot be extended too far, however, as it leads to a saturation effect so that the rate of accumulation in the chamber slows down. This can compromise the accuracy of the measurement. There is an optimum exposure time that balances these two effects. Many individual measurements are needed to characterize the emissions from the larger area of the fields and the seasonal patterns. For methane emissions from rice fields, the amplitude of the systematic seasonal cycle is usually large compared to the variability on shorter timescales. Consequently, reducing the sampling frequency increases the uncertainty of the seasonal flux very slowly. The spatial variability is large and random on the small scales of the basal area of the chambers. Reducing the number of plots sampled, therefore, has a major effect on the uncertainty of the seasonal average flux.

1. Introduction

In recent years there have been many studies of methane emissions from rice fields and other environments on the Earth's surface. Most of these studies use static chambers to enclose a small area of the field with rice plants. The accumulation of methane in the chamber over some time is directly proportional to the flux of methane. Data from such field studies are the foundation for estimating the global, country by country, or regional emissions of methane from rice fields and wetlands. For experiments of this type there are some key questions that determine whether the results are accurate or whether they can be reliably extrapolated to larger regions: How long should chambers be left on the plots? How many samples should be drawn to get a single flux measurement? How often should samples be taken during the year to delineate the diurnal cycles, the seasonal cycle, or the seasonally averaged emissions? How much of a field, or how many individual plots must be sampled to get a reliable measure of the large scale emissions?

Paper number 98JD00690.
0148-0227/98/98JD-00690$09.00

There has been very little research on these questions, although there is a clear need for answers. When we started our experiments, we did not have enough information to take these issues into account. To be on the safe side, we designed a sampling strategy that we felt would be more rigorous than was needed to obtain the average methane emissions from the region of our experiments and the factors that control the emission rates. In this paper we have analyzed our 7 -year data set from China, consisting of some 5000 individual flux measurements, to pose some answers to these questions. The results define the design parameters for field experiments of this type and also provide measures of reliability of existing data.

2. Database

Between 1988 and 1994 we took systematic measurements of methane emissions from rice fields at Tu Zu in the Sichuan Province of China. Tu Zu is a village about 20 km east of Leshan City and 100 km south of Chengdu (29.5°N, 106.7°E). A single crop of high-yielding natural and hybrid rice varieties is grown in this region every year. The fields are inundated throughout the growing season, and heavy use is made of organic fertilizers.

We chose four fields adjacent to each other and six plots in each field. Not all 24 plots were sampled every year. In the last year of the program, large chambers were used in addition to the smaller chambers that had been used in previous years. Samples were taken every other day and included measures of diurnal variability. Over the 7 years of the experiment, some 5000 valid flux measurements were taken, spanning the entire growing seasons, which lasted between 100 and 120 days. Valid fluxes were defined to be cases in which a linear accumulation was observed ($r^2>0.9$) and the initial concentration was comparable to the methane concentrations in the field. These criteria are part of the quality assurance that detect several types of disturbances due to the experimental procedures, including possible disturbance of the soils while placing the chambers on the plots. (For details, see *Khalil et al.* [this issue]).

In addition to the direct flux measurements, meta data were also gathered related to the environmental conditions, agricultural practices, and soil properties. These data include soil and air temperatures, wind, cloudiness, water level in the fields, planting density, height of the plants and agricultural records of type, amount and time of fertilizer applications, yield, and rice cultivars planted. Our experiments were done on farmers' fields under prevailing agricultural practices. A detailed discussion of the data set, its main features, and the conditions under which the measurements were taken are given in our earlier paper [*Khalil et al.*, this issue].

3. Exposure Times and Number of Samples for Individual Measurements

Fluxes of methane and other trace gases are measured by enclosing an area inside a chamber and drawing sequential samples over some length of time ΔT. The interval between each sample is $\delta T = \Delta T/(N-1)$ where N is the number of samples drawn during the exposure time. The flux is

$$F = \gamma \left[\frac{M}{N_o} \rho \frac{V}{A}\right] \frac{dC}{dt} \qquad (1)$$

where C is the measured concentrations in the chamber and dC/dt is the rate of accumulation in ppbv/min. A is the area from which methane is emitted into the chamber (m^2), V is the volume of the chamber (m^3), N_o is Avogadro's number, ρ is the density of air (molecules/m^3), γ is a unit conversion factor equal to 6.0×10^{-5} mg min^{-1} g hr^{-1} $ppbv^{-1}$, and M is the molecular weight of methane (g/mol). In most cases, V/A is the same as the height of the chamber, but for small chambers and certain configurations, it may not be.

To make a flux measurement we have to choose two parameters: the length of time the chamber is to be left on the source and the number of samples that are to be drawn during this period. If the chamber is left on for a short time, it reduces the ability to detect the flux or it increases the uncertainty in the calculated flux. If the chamber is left for a long time, it can cause saturation and feedback effects because of high accumulations.

The measurement of flux is more or less equivalent to measuring dC/dt in equation (1). There are several ways to estimate dC/dt from the data. We will use the linear regression method for this discussion because it is the most common. In this method the concentration in the chamber is written as $C = C_o+bt$, and the slope "b" is the estimate of the dC/dt, and its standard error "S_b" is the estimate of uncertainty in the flux. For evenly spaced measurements in time, S_b is given by

$$S_b = \left[\frac{\Sigma d_{c,t}^2/(N-2)}{\Sigma t^2}\right]^{1/2} \qquad (2)$$

$$\sqrt{\Sigma t^2} = \delta T \sqrt{\frac{(N-1)N(N+1)}{12}} = \Delta T \sqrt{\frac{N(N+1)}{12(N-1)}} \qquad (3)$$

where $d_{c,t}$ is the difference between the measured concentration and the concentration "predicted" by the regression equation for time t, and t is the difference between the real time and the mean time. The numerator of equation (2) is approximately the fundamental root-mean-square (rms) variability of the system and is not dependent on N (the number of samples collected for an individual measurement of flux). The denominator contains the effect of the number of samples. It has two aspects. First, it depends linearly on the total length of time over which samples are taken "Δ T," and second, it depends on the number of measurements taken. For large N, the S_b goes down approximately as $N^{-1/2}$, as is commonly known [Snedecor and Cochran, 1989]. Therefore, improvements in the precision and detectability of the flux (slope) can be made either by increasing the number of measurements or, more efficiently, by increasing the total length of time (of exposure in this case). If there were no other considerations, a few measurements spread out over a long exposure time would be the preferred sampling method.

Long exposure times cause significant disturbances to biological and chemical systems. There are two processes: the disturbance of the system under observation and instrumental nonlinearities or detector saturation effects. While the chambers are on the plot, the environment inside can become substantially different from the outside, causing potential feedbacks that affect the measurement of flux. The variables that change are temperature, solar radiation, levels of CO_2, and other gases that can affect the plants, and the concentration of methane itself can become very large compared to the conditions outside the chamber. As the methane concentration gets large, it may be taken back by the water and soil, causing a feedback that compromises the flux measurement. In addition, if the concentration are large, even small nonlinearities in the performance of the detectors used to measure methane may affect the measurement. In our case, as in most methane flux measurements, a gas chromatograph is used with a flame ionization detector (GC/FID). The concentrations that are being measured range from about 2 to 500 ppmv, while the detectors and the systems are often optimized for the lower part of this range where most of the measurements fall. For methane emissions from rice fields it seems that the saturation effect, whether it is caused by the feedbacks in the environment inside the chamber or the response of the detectors, has the effect of reducing the measured concentration of methane. This is reflected in a slowdown in the trend dC/dt measured at the end of the exposure time compared to the beginning. Because neither very long nor short exposure times are desirable, an optimal exposure time can exist.

In the case of our studies, we used 9-min exposure times and four measurements during this time (0, 3, 6, and 9 min). The true exposure time may be a few minutes longer since it takes time to install the chamber before a sample is drawn. We found that during this time there is little disturbance of the environmental conditions inside the chamber compared to the outside [*Khalil et al.*, this issue]. Shorter exposure times of 3 min or 6 min compromise the precision of the measured flux and create a more uncertain data set.

Longer exposures have too many cases where the saturation effect is observed. The sampling time we selected for the size of the chamber is therefore close to an optimal balance between these opposing factors. Next, we will demonstrate the observations of these opposing factors.

In our work we take the best estimate of the flux to be given by a value of "b" or dC/dt obtained from the regression of C_i with time t_i = 0, 3, 6, 9 min for i = 0, 1, 2, 3. The 3-min interval is sufficiently long that samples can be taken precisely at this interval. While there may be small deviations from 3 min in the field, these are too small to affect our results. There are also several other calculations of flux possible from these same data. These calculations reflect what would happen if we had adopted a different sampling strategy. First, we can calculate a flux that would have been observed if we took only two measurements (N=2) during each experiment, instead of four measurements. For this case with N=2, we could have exposure times of 3, 6, or 9 min. The dC/dt in equation (1) is calculated by

$$\frac{dC_i}{dt} = \begin{cases} (C_{i+1}-C_0)/(t_{i+1}-t_0) & i = 0,1,2 \\ (C_i-C_{i-1})/(t_i-t_{i-1}) & i = 3 \end{cases} \quad (4)$$

Each of these are two-point calculations of flux and therefore use only half the data we actually obtained. The first, i = 0, is the flux we would have measured if we had left the chambers on for only 3 min. The second, i = 1, is the flux that would have been observed if we left the chamber on for 6 min. In this case the flux is not affected by an intermediate measurement at 3 min, so the result would be the same whether we used a regression of three measurements over 6 min (0, 3, 6 min) or just the two-point estimate as in equation (4). We will discuss this matter again later on. The third, i = 2, is the flux we would have measured if we kept the exposure time the same at 9 min but used only two points to estimate the flux. The three cases illustrate the effect of keeping N = 2 and changing ΔT in equations (3) and (4), while the comparison of the last case (i = 2) with the regression calculation using all four measurements illustrates the effect of keeping ΔT the same and changing N. The fourth case (i = 3) is the two-point flux calculated for the last 3 min of the exposure. This is used to evaluate saturation effects because it represents an accumulation in the presence of a large amount of methane in the chamber.

The relationship between the regression flux using all four measurements for each flux experiment and the two-point estimates for various exposure times is shown in Figure 1. This figure contains about 4500 valid flux measurements for 1988-1994 (except 1991, for which such calculations cannot be done because of different sampling procedures). Each of the two-point calculations represents only half the data, but the longer exposure times give a much better estimate of flux. A measure of the agreement between the different estimates of flux can be calculated as the root-mean-square variance between the two-point estimates and the flux based on the regression estimate for all four points.

$$\Delta \text{rms}_i = \left\{ \Sigma [F_{ji}(\text{reg}) - F_{ji}(2\text{-point})]^2 / n \right\}^{1/2} \quad (5)$$

where "j" is the individual measurement over a single plot and n is the total number of such measurements during the year "i." The results of these calculations are shown in Figure 2. There is a remarkable similarity of the results for each year. The variance decreases rapidly with longer exposure times and is only between 1-2 mg m^{-2} h^{-1}, or about 5% of the mean flux, for the two-point estimate over the 9-min exposure time. It is 12-16 mg m^{-2} h^{-1}, or

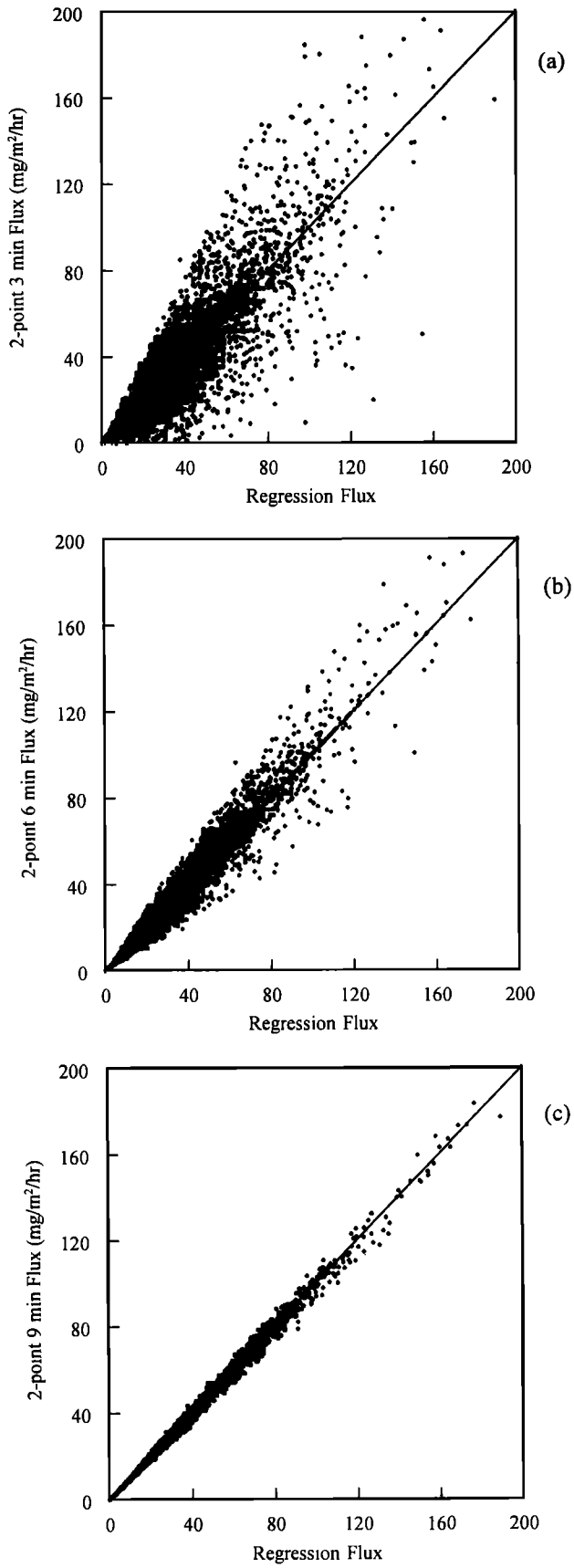

Figure 1. A comparison of estimates of flux using only two data out of four and the estimate based on the regression model using all four samples for each measurement. For an exposure time of a) 3 min, b) 6 min, and c) 9 min.

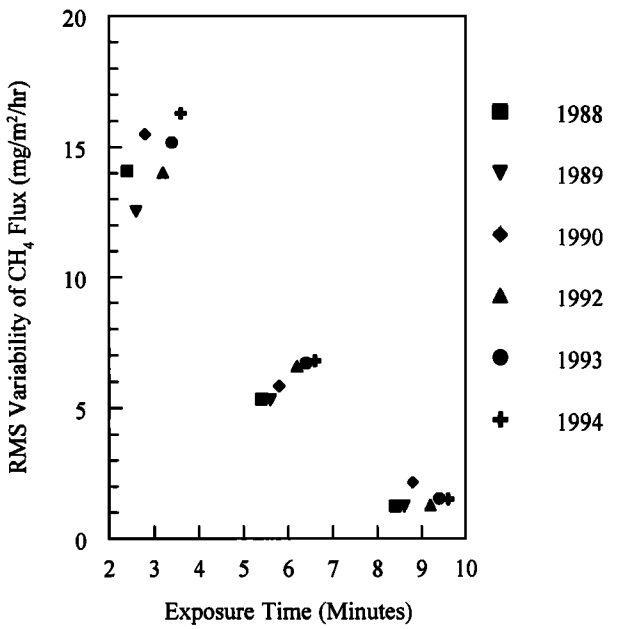

Figure 2. The root mean square variability of individual measurements during each year as a function of exposure time. Short exposures increase the uncertainty of the flux. The data are for 3-, 6-, and 9-min exposure times. The points have been shifted slightly to the left or right to avoid overlap.

about 50% of the mean flux, for 3-min exposure times. These calculations reflect the variability of single measurements for each year of data. This rms variability is small enough that most calculations, such as diurnal variability or the relationship between fluxes and other environmental factors, that can be done with the full data, can also be done with the reduced data sets, especially for the 6-min and 9-min exposure times. The diurnal variability is about the same order as the rms variability of the 3-min exposures, which may make it difficult to detect the diurnal cycle with such a data set.

Often we are interested only in the average flux for the whole season. The effect of two-point sampling relative to the four-point regression method is extremely small for the seasonally averaged flux. The calculations are shown in Figure 3. For 3-min exposures the differences are between +7 and -3%; for 6-min exposures the differences are between +5 and -1%; and for 9-min exposures the calculated seasonally averaged fluxes are between -1 and +0.2%, or almost the same as for the regression method using twice as much data.

These results show that sufficient exposure times have to be allotted to reduce the uncertainty of the measured flux and that the length of the exposure time is more important than the number of measurements done to estimate a single flux. The latter conclusion is based on the result that two-point measurements over a 9-min exposure time give nearly the same results as the four measurements over this same time. Issues related to these conclusions will be discussed later in the paper.

These results raise questions as to whether it is necessary to collect more than two samples per flux measurement. If only three samples are collected, regularly spaced in time, then the middle measurement has no effect on the measured slope or trend and hence does not affect the calculated flux. The only purpose this middle measurement could serve is to determine how linear was the accumulation of the trace gas in the chamber. If it is not linear, the measurement may be affected by sampling artifacts and could be unreliable. Thus the intermediate measurements can serve to validate the data and are often used for this purpose [*Sass et al.*, 1992; *Khalil et al.*, this issue]. Aside from this use, additional measurements between the beginning and end of the sampling process for $N > 3$ add some precision to the estimated flux, but this can be easily achieved by slightly longer total exposure times. When samples are manually analyzed in the laboratory, the number of measurements taken per flux measurement is much more time consuming than collecting fewer samples over a slightly longer exposure time. Since most such experiments, at least in our experience, are limited by the time available for laboratory analyses and sampling, and by the number of sampling containers available, fewer samples per flux measurement are desirable if the quality of the data can be maintained. Then the effort can be put into more replicates, which, we will show later, is important for estimating the emissions over large areas.

Exposures cannot be made too long, however, as saturation effects set in. To study the saturation effects in our data, we defined the difference of the "first" and "last" fluxes as ΔF. The first and last fluxes are given by equation (4) for $i = 0$ and $i = 3$, respectively. A positive ΔF means that the accumulation rate slows down towards the end of the exposure time, which we take to be an indication of

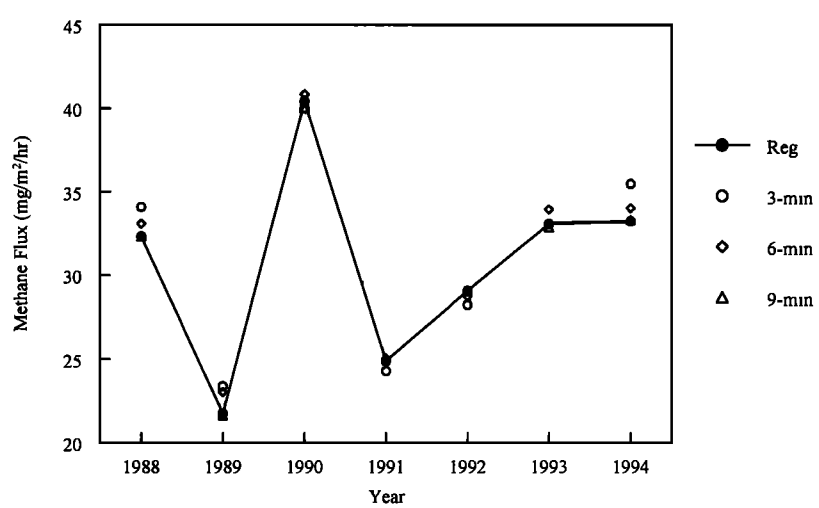

Figure 3. Seasonally averaged flux using the regression method with all four samples for each flux measurement and using only 2 points 3, 6, and 9 min apart.

saturation. We took the composite data set from all years (except for 1991, for which this calculation cannot be done). These data were then ranked according to the final concentrations measured in the chambers. Generally high fluxes result in high final concentrations. We then calculated the average difference of flux ΔF and the average final concentration in the chamber for 200 data points at a time. The choice of 200 is somewhat arbitrary but inconsequential to our arguments. The results are shown in Figure 4. When the chamber concentrations are below 30 ppmv, we do not see any evidence of saturation as measured by ΔF. For concentrations greater than 30 ppmv, saturation effects appear but are not large until concentrations exceed 60 ppmv. The average difference of fluxes for this range of chamber concentrations is 10% or less, and for the worst case of the highest concentrations, it is 27%. The highest concentrations observed in the chambers were about 500 ppmv, but such values were exceedingly rare in the valid data. There were only about 150 data exceeding 100 ppmv out of about 5500 points.

For our studies, the 9-min exposure time is about optimum since it is the longest exposure time for which we have only a few percent of the data affected by saturation, and that too leads to relatively small (<10%) potential errors in the flux. At the same time, it gives repeatable measurements of the flux.

The selection of number of samples per experiment, or the exposure time, depends fundamentally on the concentrations inside the chamber, so the conclusions are the same for environmentally important gases other than methane, and for sources where chambers are used. For rectangular chambers, according to equation (1), for a given flux, the measured dC/dt (slope) is inversely proportional to the height of the chamber and independent of the surface area covered (we will return to this point later). Increasing the height of the chamber raises the detection limit for the flux. To detect the flux, $dC/dt \geq z_\alpha S_b$, where z_α is a criterion value that depends on the probability (α) and may be calculated by either a t distribution or nonparametric statistical methods. From equations (1)-(4) the minimum detectable flux is

$$F_{min} = \frac{(\sqrt{12}\, S_{c\,t}\, \beta z_{\alpha/2}) h N^{-1/2}}{\Delta T} \qquad (6)$$

$$\beta = \gamma \frac{M}{N_0} \rho$$

The term $\sqrt{12}\, N^{-1/2}/\Delta T$ is actually an approximation for $\sqrt{\Sigma t^2}$ in equation (3). If N is small, equation (3) should be used instead of the approximation for $\sqrt{t^2}$. The factors inside the parentheses are properties of the gas, the detection criterion, and the existing variability of the measured concentration (which includes both environmental and instrumental variabilities). These factors are not controllable in the experiment. In many cases, especially for methane emissions from rice fields, achieving the minimum detectable flux is not a problem with the chamber sizes commonly available. When the fluxes are very small, as for N_2O in certain agricultural systems, and the detection limits are high (large $S_{c\,t}$), then the height of the chamber can be reduced, and the exposure time can be increased to achieve a detection of flux [see also *Khalil and Rasmussen*, 1995].

To guard against saturation effects, the chamber height has to be increased, which necessarily increases the minimum flux that can be detected. In the case of methane emissions from rice fields, the small fluxes are not as important as the larger fluxes, and this compromise is acceptable within some prescribed criteria or acceptable minimum flux detection limit.

Once the chamber size, exposure time, and number of measurements to be taken to estimate the flux are selected, the method for a single measurement is established. The questions then arise as to how many such measurements are needed over the area of interest and how often the measurements should be taken during the growing season. These will be addressed next.

4. Sampling Frequency and Spatial Replicates

If the flux is spatially homogeneous, then one chamber would be sufficient to estimate the emissions from the entire region where this assumed homogeneity holds. In practice, the emissions are almost never homogeneous enough over the regions of interest. If we adopt a modest goal to estimate the emissions from the local rice fields in our area of interest, how many plots should we sample and how large should each plot be?

In our experiments we sampled up to 24 plots in four adjacent fields, twice a week during the growing season. This basic sampling strategy was modified from year to year as is discussed in detail in our earlier paper [*Khalil et al.*, this issue]. The data set

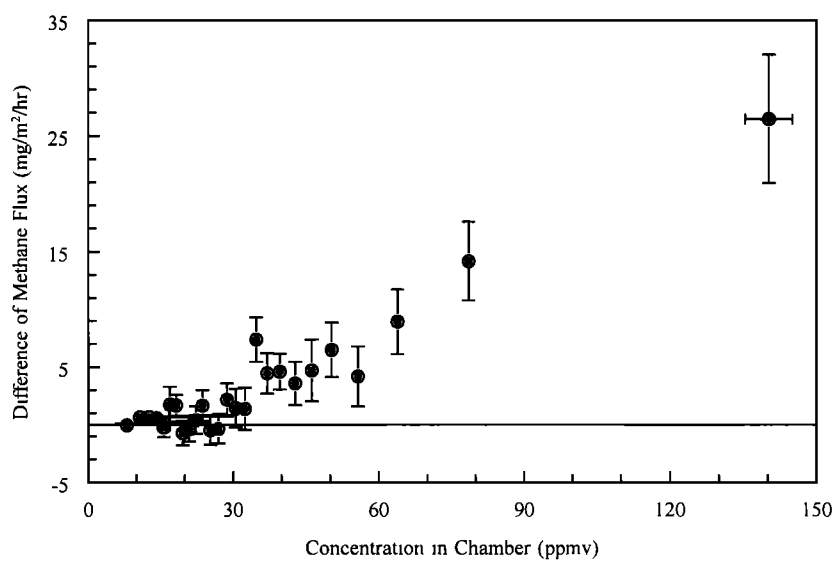

Figure 4. The difference between the emission rate at the beginning of the sampling period and at the end as a function of the methane concentration inside the chamber at the end of the experiment. Saturation effects appear when the difference significantly exceeds zero.

allows us to evaluate the errors that could have occurred if we used fewer plots or sampled the fields less frequently. From these results we will be able to draw some general conclusions about the frequency of sampling and the number of replicates needed to obtain accurate area wide emissions estimates.

Our data are in pairs of sampling days. One day the sampling is in the morning, and the next time it is in the afternoon. This was done to obtain an estimate of the diurnal variability of the emissions. We retained the diurnal cycle in our calculations by taking the following sub-sets of data for our analysis here: We started by deleting every fourth pair of sampling days leaving 75% of the data; we deleted every third pair of sampling days (67% of the data left); then we deleted every other pair (50% of the data left). These operations created three new data sets with reduced sampling frequency but with all the spatial replicates. Then we created more such data sets by deleting every fourth and fifth pairs of sampling days, then every third and fourth pairs, then every second and third pairs. This process was continued by deleting every fourth, fifth, and sixth pairs, every third, fourth, and fifth pairs, and so on, until we reached the case that had one pair of days in the beginning of the season, one in the middle, and one at the end, or just three pairs of days on which samples were collected, simulating a very sparse sampling strategy and constituting around 10% of the data actually collected.

From each of these data sets we calculated the seasonally averaged emission rates and compared them to the calculations based on all the data, using the same integration methods as described in our earlier paper [*Khalil et al.*, this issue]. The results are shown in Figure 5. It is remarkable that the error in seasonally averaged flux is within ±10% for most of these cases until we go down to using 20% or less of the data. Then the errors increase rapidly but are still within +20 and -40% even for a very sparse temporal sampling strategy.

We also evaluated the effect of having fewer plots each year but keeping the original full twice-weekly sampling strategy. We started with the calculated seasonally averaged flux for each plot during each year. These were ranked from lowest to highest. We calculated two indices of the maximum effect of using fewer plots compared to many plots. If we had sampled only one plot, we assumed that the lowest seasonally averaged flux we would have obtained would be equal to the minimum flux observed among the 24 plots. And the maximum flux we may have found is the maximum flux observed. If we had sampled two plots, the lowest flux we would calculate is the average of the fluxes from the two lowest emitting plots out of 24 plots, and similarly for the maximum we would have observed. This process is continued by averaging the emissions from the three lowest and three highest emitting plots and so on. These calculated average emissions from two, three, or more plots were then compared to the average of all the plots from which we calculated the percent error. Another index is the ratio of the maximum and minimum average emissions for two plots, three plots, etc. The results are plotted in Figure 6. When there were fewer than 24 plots sampled, we added the variability that is found for plots up to that number based on data from 1988 and 1989. We see from Figure 6a that if we sampled only one or two plots, we could have an error in the seasonally averaged flux of up to ±50%, which is quite large. If we sampled 12 plots, the error could be as large as ±20%. The calculations of Figure 6b show that sampling one or two plots could result in the seasonally averaged flux being wrong by factors of 2.5-3, and sampling 12 plots still leaves errors of a factor of 1.5. For the entire range of number of plots from 1 to 24, the errors in the ratio go down approximately as $N^{-1/3}$ but go more slowly when we go from very few plots to more plots, ~ $N^{-0.2}$, and faster as we add more plots after 12, ~ $N^{-0.7}$.

These calculations are based on the assumption that sampling 24 plots gives us the emission rate for the field, or equivalently that if we had sampled more plots, the maximum and minimum concentrations would not change (much). Although this is the only assumption we can make, it has no significant effect on the conclusions here, which are qualitative in nature. Increasing the number of plots is equivalent to covering more and more of the surface area of the field. Even with 24 plots, the fraction of the area of the fields covered by chambers is only 0.5%.

For rice fields there is a fundamental spatial heterogeneity in the emission rates. It comes about because rice plants are grown at a regular distance from each other and most of the emissions occur through the rice plants, with little emission from the water in between. Also the emissions from the water, which are mostly by ebullition, are quite sporadic compared to the more steady emissions from the plants. There are other factors that are also

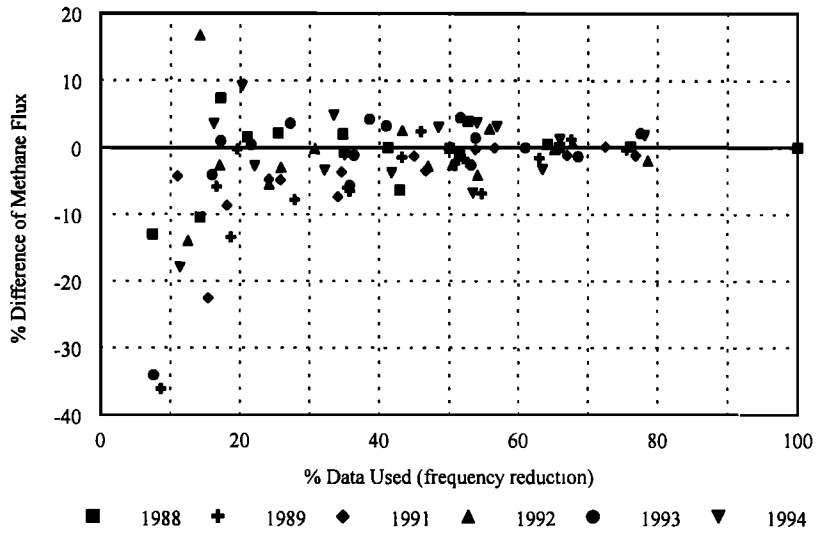

Figure 5. The effect of reduced sampling frequency on the seasonally averaged emission rate of methane from rice fields.

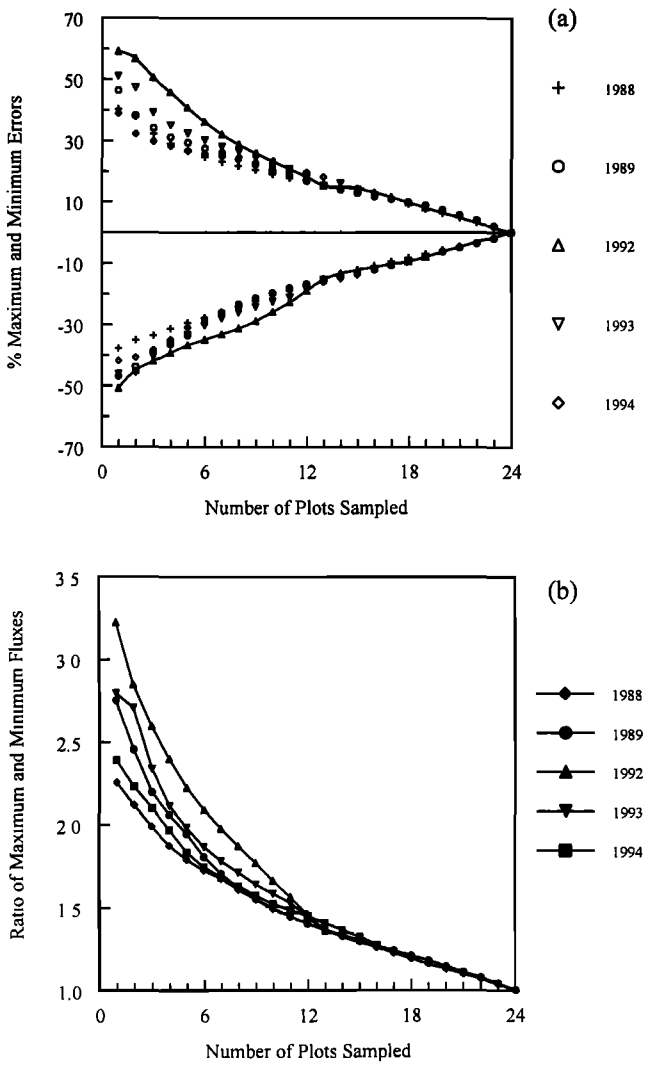

Figure 6. The maximum errors introduced by taking samples over fewer plots. (a) Difference between seasonally averaged fluxes if fewer than the full 24 plots are sampled. (b) The ratio of the maximum to minimum fluxes expected if fewer than 24 plots are sampled. Here 24 plots is the standard because that is the maximum number of plots that were sampled in our experiments. The figure shows that if only one or two plots are sampled, the results could be in error by up to a factor of 3.

distributed unevenly over the fields including soil texture, chemistry and fertilizer applications, but it is less likely that these affect the spatial homogeneity of emissions as much.

The regular planting creates a fundamental length for the rice field, which is the distance between plants. In our studies this distance is about 20 cm. To properly take into account the heterogeneity, sampling should cover an area several times the square of this fundamental length. In our past studies we used chambers that were close to this length, which caused some problems in evaluating the emissions from the field as a whole and required corrections that were experimentally determined over several years. Our present chambers have basal lengths about 7-8 times this fundamental length. The results seen here may be scaled for other sources if a fundamental length can be defined on the surface being sampled. Here the number of plots used can be regarded as representing multiples of the fundamental length. Although the rice fields do appear to have substantial spatial variability of emissions on small spatial scales covered by our chambers (about 0.56 m^2), there are other sources, such as landfills or wetlands, that could have much more spatial variability of emissions. In the case of landfills, the production, oxidation, and emission processes are all unevenly distributed over the surface, and it is difficult to define a length that can be used to determine the area of the surface that needs to be covered.

For spatially homogeneous emissions the basal area of the chamber does not affect the measured flux, but for rice fields the area covered by the chamber is an important factor in reducing the uncertainty of the seasonally averaged or daily methane emission rate. This issue is discussed in our earlier paper with the conclusion that a large base area (several times the fundamental length) is highly desirable for sampling emissions from rice fields [*Khalil et al.*, this issue]. The desired upper limit of the area covered is limited only by the logistics of constructing and setting up large chambers.

We have seen that a 50% reduction of the data by lower sampling frequency has less than a ±8% effect on the seasonal average emissions (Figure 6a), but a 50% reduction of data by reducing the number of plots has an effect of up to ±16-20%. The situation is worse as we reduce the data further. One reason for the difference of the sensitivities to frequency relative to spatial extent is that methane flux follows a systematic seasonal cycle, but the spatial variability is random as far as we can tell. The amplitude of the systematic seasonal cycle is larger than the variability on smaller timescales. So sampling at a few times the seasonal frequency captures the whole cycle, giving us good estimates of seasonal flux with infrequent measurements. As long as there is a systematic seasonal cycle that exceeds variability on other timescales, the sampling frequency can be adjusted to this fundamental frequency of the seasonal cycle. This is not so for the spatial variability. In that case, the estimates are likely to continue improving as more plots are sampled or larger chambers are used.

5. Conclusions

In designing field studies there are a number of logistical and scientific considerations that need to be balanced. The static chamber experiments discussed here are quite common, and most current data on the emissions of methane from rice fields are obtained from such studies. It consists of trapping the emissions from the rice fields inside an enclosed chamber and then measuring the buildup of the gas from a series of sequential samples. With automated measurements, large amounts of data can be obtained without substantial additional costs, but for most studies the number of measurements that can be done during the growing season are limited by the time available for collecting and analyzing the samples. In such cases, samples have to be collected in a manner that provides the most reliable data from the limited numbers of analyses. We have shown here that having many spatial replicates, or equivalently covering a larger area of the field, produces more benefit than collecting the same number of samples with a higher frequency. It is difficult to derive quantitative relationships for this finding.

Regardless of the frequency and spatial coverage, each flux measurement has to be reliable. There are two factors: exposure time and number of samples per flux measurement. Here, more quantitative measures can be derived. The exposure time has to be kept short to avoid feedbacks and saturation effects, both of which can occur with high concentrations of methane inside the chambers.

Feedbacks can also occur even when the concentrations of the trace gas being measured are not high. This is because other environmental conditions change inside the chamber if it is left on too long. If the exposures are made too short, the uncertainty of the flux increases. So there has to be a balance between variability and saturation effects.

In our study we used chambers with different heights during the growing season. This strategy lowers the minimum detection limit for the flux in the beginning of the growing season when the emissions are low, and raises the detection limit, but guards against saturation, in the middle of the season when the emissions are high. Also, as the plants grow, taller chambers are needed to enclose them. There is perhaps never a good scientific reason to use chambers with small basal areas comparable to some measure of the spatial scale of flux variability. In addition, it seems that increased spatial coverage is always desirable in such experiments. The frequency of sampling can be optimized by using the seasonal cycle of the emissions as a guide. Proper adjustments of the height and basal areas of the chambers can be an important tool for reducing uncertainties in the estimated fluxes of methane from rice fields. If a repeating seasonal cycle or basic length can be defined, an efficient sampling strategy can be based on these parameters.

The basic principals derived here are applicable to the measurement of fluxes of other gases and emissions from other sources. The values of the fundamental lengths for homogeneity, chamber dimensions, and lengths of exposure may have to be changed to obtain optimal sampling strategies, but the variables are the same. From the findings discussed here and in our earlier paper, more reliable measurements can be taken, and loss of data can be avoided.

Acknowledgments. This project was supported by the U.S. Department of Energy (DE-FG06-85ER60313). We thank Michael Riches for his many suggestions and his help with the logistics of carrying out the experiments. We thank R.W. Dalluge and D. Stearns at OGI, Ren Lixin, and Wang Ming-Xing at the Institute of Atmospheric Physics (CAS) in Beijing, and Duan Chang-Lin at the Institute of Mountain Disasters and Environment in Chengdu for their contributions to this project. Additional support was provided by the Biospherics Research Corporation and the Andarz Co.

References

Khalil, M.A.K., and R.A. Rasmussen, The changing composition of the earth's atmosphere, in *Composition, Chemistry, and Climate of the Atmosphere*, edited by H.B. Singh, chap. 3, pp. 50-87, Van Nostrand Reinhold, New York, 1995.

Khalil, M.A.K., R.A. Rasmussen, M..J. Shearer, R.W. Dalluge, Ren L.X., and Duan C.L., Measurements of methane emissions from rice fields in China, J. Geophys. Res., this issue.

Sass, R.L., F.M. Fisher, Y.B. Wang, F.T. Turner, and M.F. Jund, Methane emission from rice fields: The effect of floodwater management, *Global Biogeochem. Cycles, 6*(3), 249-262, 1992.

Snedecor, G.W., and W.G. Cochran, Statistical Methods, 8th ed., Iowa State Univ. Press, Ames, 1989.

M.A.K. Khalil and M.J. Shearer, Department of Physics, Portland State University, P.O. Box 751, Portland, OR 97207 (email: aslam@atmos.phy.pdx.edu)

R.A. Rasmussen, Department of Environmental Science and Engineering, Oregon Graduate Institute, P.O. Box 91000, Portland, OR 97291.

(Received July 29, 1997; revised February 17, 1998;
accepted February 24, 1998.)