An Intercomparison of BlueBird Methane Estimates with Coincident High-Accuracy Observations Obtained by an Air Quality Monitoring Site

J.A. Maslanik
Earthview, LLC.
18 Feb. 2021

1. Objectives

A key requirement for applications of gas sensor technology for well site monitoring is to assess and understand the performance of the instruments in outdoors field conditions, over a period of time, and in conjunction with reliable validation measurements. While testing in a lab setting is valuable, it is impossible to reproduce the range of conditions encountered in the actual environment.

To help address this, we are taking advantage of an opportunity to operate BlueBird instruments in an outdoors setting for an extended time period, coincident with comprehensive, high-accuracy air constituent measurements.

The main objectives of the resulting intercomparisons are to assess (a) the sensitivity of the BlueBird package to hydrocarbon gases in the presence of normal variability in atmospheric conditions, and (b) the ability to quantify the measurements in terms of estimated methane concentration.

2. Data

BlueBird data are compared to research-grade observations collected by the Boulder A.I.R. LLC air monitoring site at Soaring Eagle Park, at Anthem Ranch in Broomfield County, Colorado. Here, we use data collected from late October 2020 through January 2021. The Boulder A.I.R. instrumentation and sampling methodology are described at https://www.bouldair.com/broomfield.htm. Methane is measured by Boulder A.I.R. using a Picarro G-2401 Cavity Ring Down Spectrometer and reported at 1-minute intervals, while VOCs are sampled at 10-minute intervals, processed using a gas chromatograph, and reported hourly. The BSE air intakes are located at a 6-m height on a mast. We refer to these Boulder A.I.R. Soaring Eagle data as "BSE".

The BlueBird data used for comparison were acquired by a early-version BlueBird "primary" package installed at the top of the Boulder A.I.R. mast, with the BlueBird air intake located about 3 meters above the Boulder A.I.R. intakes. These BlueBird data were reported via cellular transmission to Earthview's processing center at approximately 70-second intervals. The raw sensor data from BlueBird were processed to compensate for environmental conditions.

To merge the BSE and BlueBird data, a time-matching procedure was written to find the BSE record closest in time to each of 54251 BlueBird records acquired over the study period. The mean absolute-value time difference between the resulting matched BSE data files is 16 seconds (maximum difference of 320 seconds). 54233 matches of the 54251 records differed in time by less than 30 seconds. Due to data gaps, the resulting matched data sets covered about 44% of the full time period from 23 Oct. 2020 through 31 Jan. 2021.

3. Methodology

For the analyses described here, two different relationships were used to estimate methane concentration from the BlueBird data. These relationships were derived by visually extracting subsets of the time series where there were overlapping peaks seen in the BSE data and the BlueBird sensor data. Using these subsets, we calculated best-fit equations where the dependent variable is the BSE methane concentration and the predictor variable is a parameter obtained from the BlueBird data. Two such relationships were estimated. One of these used data subsets where the gas sampled at a particular time was a mixture of methane and the heavier hydrocarbons (ethane and propane in particular). We refer to the resulting equation as a "natural gas" fit. The second equation was limited to subsets where the peak consisted of almost entirely methane. We refer to this equation as a "maximum methane" fit. The extracted subsets for the corresponding peaks amounted to 17% of the full time series, with the remaining 83% not contributing to the equation fit calculations. A third equation defining the relationship between pure methane and sensor readings, estimated from data provided in the sensor manufacturer's specification sheet, was also applied.

The matched time series were compared by analyzing corresponding time segments spanning 2000 records at a time. Individual segments within these periods were then studied in closer detail to visually assess correspondence between BSE methane peaks and BlueBird data variations. The data were also compared by calculating mean methane concentrations and by assessing the percentages of observations where potential leak events were either "missed" or "false" in the BlueBird data.

It should be noted that the BlueBird results presented here were obtained using Earthview's initial, relatively basic processing methodology, with minimal filtering and no "massaging" of the data. We anticipate significant improvements using the machine learning tools that are presently being implemented as part of our routine processing strategy.

4. Results

4.1 Time series comparisons with BSE data

Plots of different time segments illustrate the degree of correspondence between methane concentrations measured by Boulder A.I.R. system and those estimated from the BlueBird data. The time period in Figure 1 captures several aspects typical of these time-segment comparisons.

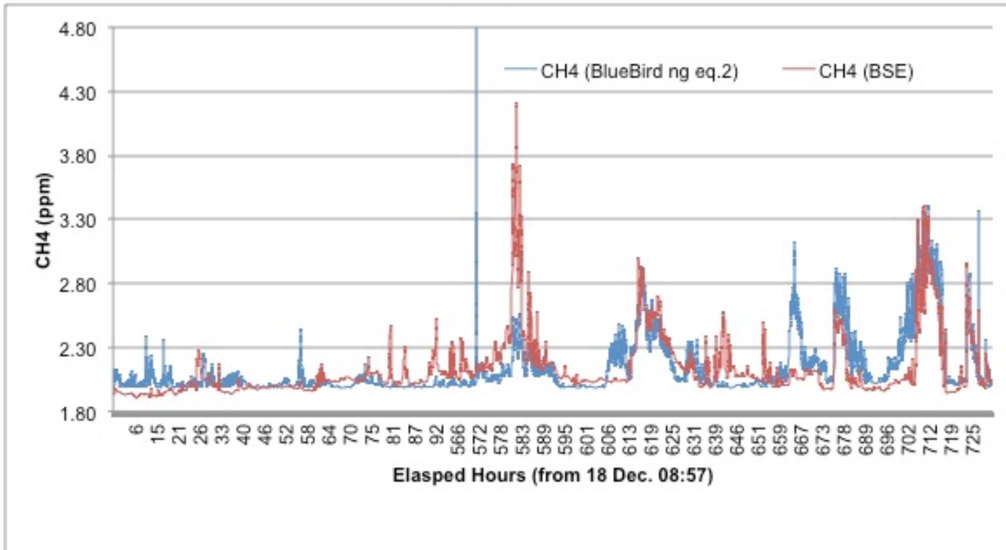

Figure 1. BSE-measured methane concentration (red) and methane concentration estimated from BlueBird data using the "natural gas" equation (blue). The data span a 30-day period, with a sampling interval of about 70 seconds.

One can identify 6 individual, substantial peaks in the BSE data (at hours 583, 619, 639, 678, 712, and 725). The BlueBird data show corresponding patterns for 5 of these peaks.

There are two locations in Figure 1 where BlueBird indicates noticeably higher methane concentrations than recorded in the BSE data. At around hour 570, a large spike is seen in the BlueBird data, with no matching spike in the BSE data. This portion is examined more closely in Figure 2 using data from two of the BlueBird sensors. In the top panel, the spike near observation number 1126 corresponds to the spike in estimated methane concentration at hour 570 in Figure 1. The second panel shows data from a second sensor in the BlueBird package, with a corresponding spike. This sensor has greater sensitivity to VOCs than does the first sensor, which is the primary methane sensor in our system. This sensor also responds to some VOCs but with less sensitivity than the second sensor. The third panel is a plot of BSE-measured acetylene, suggesting that gases other than methane contributed at least in part to this outlier.

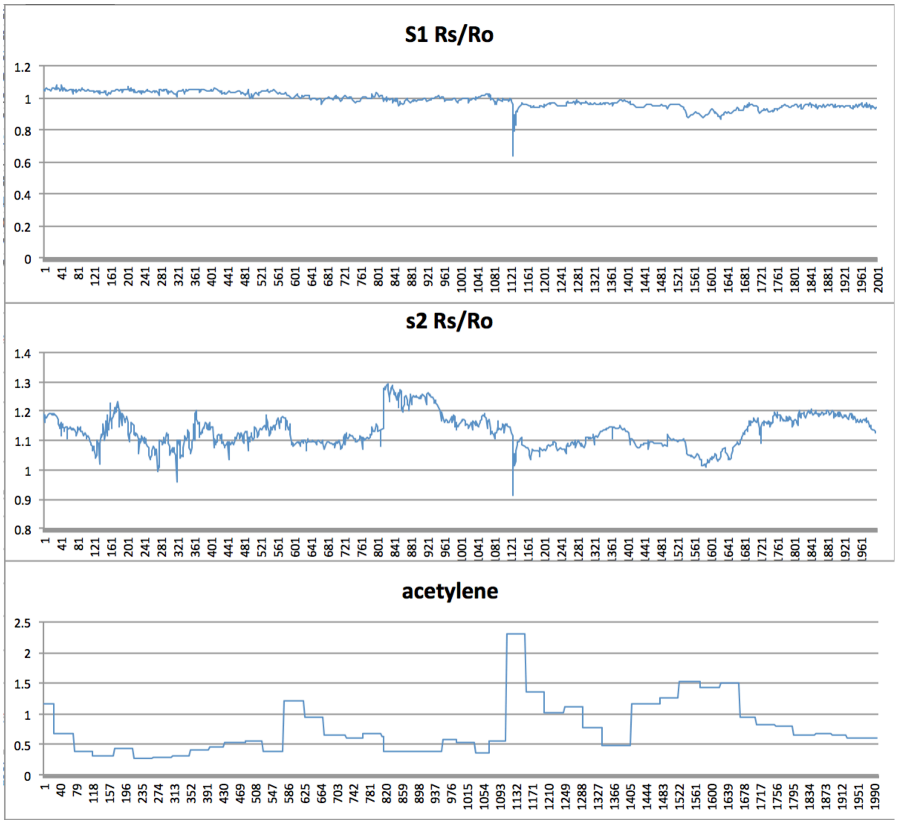

Figure 2. BlueBird sensor 1 and sensor 2 readings (top two panels) and near-coincident BSE-measured acetylene concentration (bottom panel).

Additional time-series examples are shown in Figures 3-6. The results seen are typical in terms of the tendency for BlueBird to respond to spikes in BSE-measured methane of about 1 ppm and greater. Methane estimated using the two different BlueBird equations ("natural gas" ["ng"] and "maximum methane" ["mm"]) are presented in Figures 4 and 5 along with the Boulder AIR BSE ethane and propane measurements. In Figure 4, the concentrations of these heavier natural gases versus methane are relatively low (i.e., a relatively "dry" natural gas mix) during the methane peak during hours 7-9. As a result, the BlueBird "maximum methane" equation provides a better fit, albeit still an underestimate, to the BSE-measured methane during this period. In Figure 5, the "maximum methane" estimate tracks the BSE concentration well for the peak at hour 4.8 but then overestimates concentration at hour 7.8. The opposite is the case for the "natural gas" equation result. The apparent contribution of the heavier gases is also seen in Figure 6, where the correspondence between BlueBird and BSE methane concentrations improves during the latter part of the time-series subset.

The sensor-to-methane-concentration equation derived from the sensor manufacturer's data consistently and substantially overestimated the actual methane concentration, presumably due to the presence of heavier hydrocarbons not accounted for by the manufacturer's use of pure methane. We therefore limit the analyses here to the ng and mm equation results.

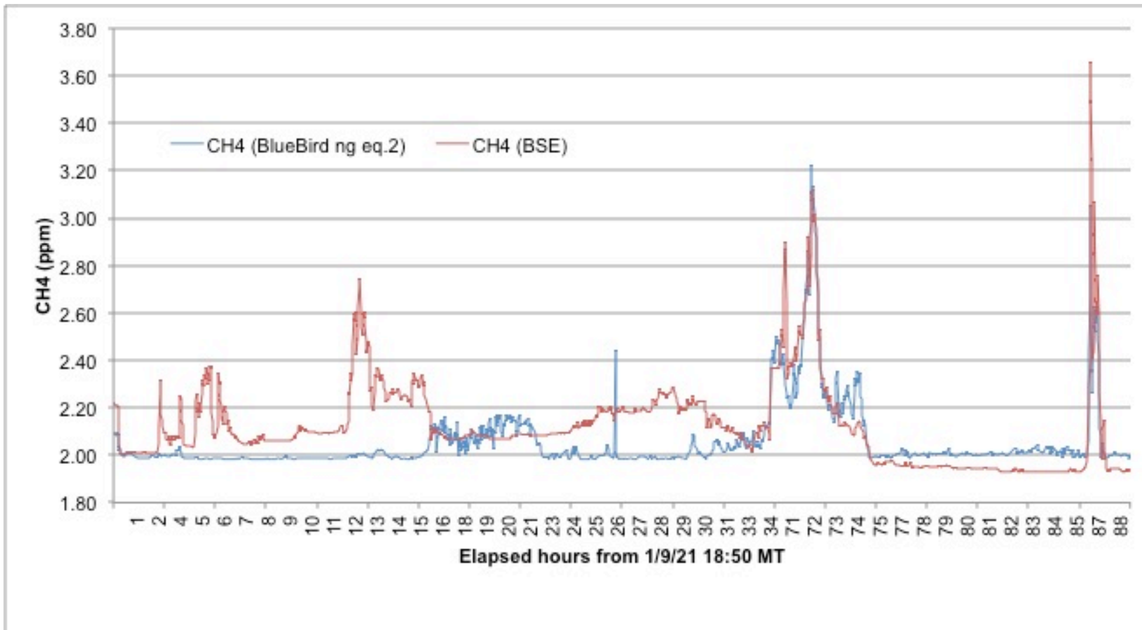

Figure 3. BSE-measured methane concentration (red) and methane concentration estimated from BlueBird data using the "natural gas" equation (blue).

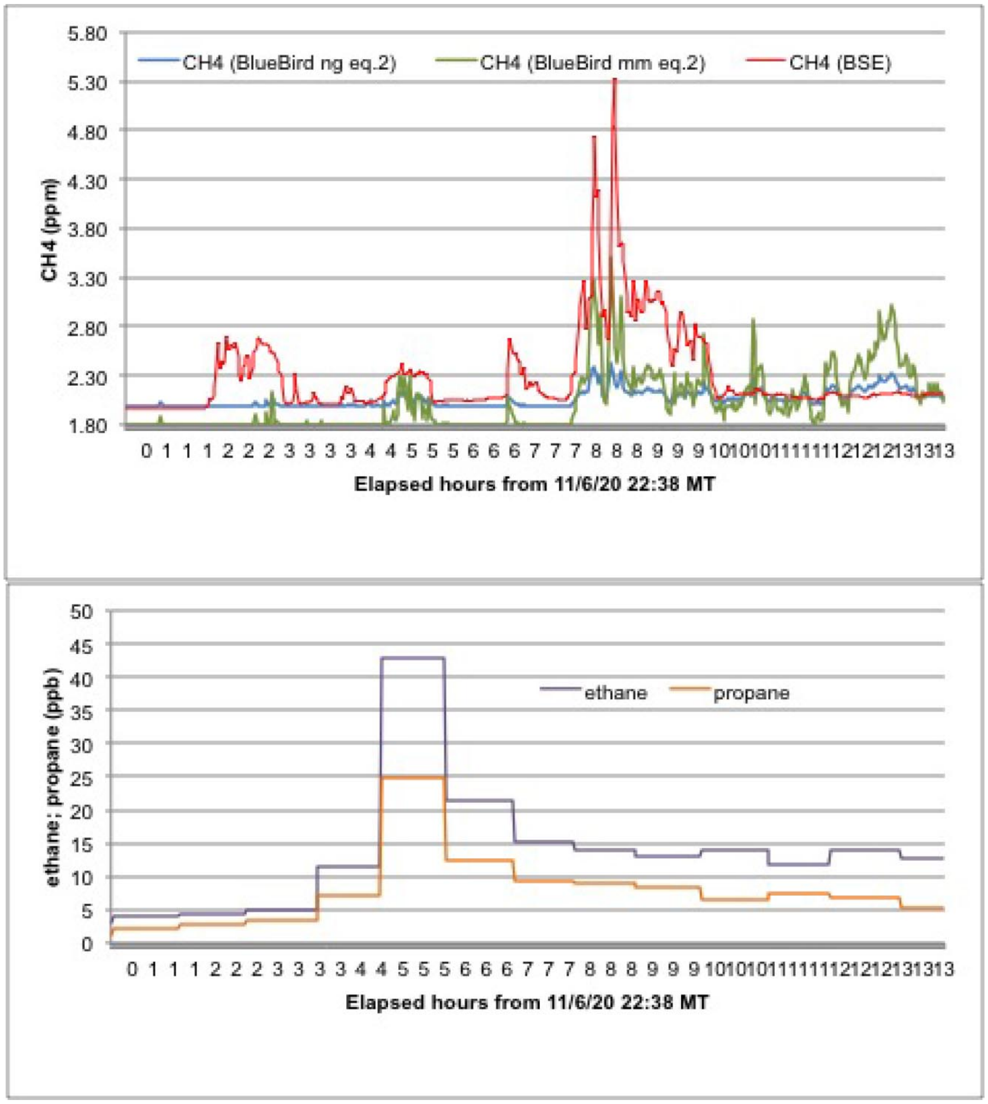

Figure 4. Top panel: BSE-measured methane concentration (red) and methane concentration estimated from BlueBird data using the "natural gas" equation (blue) and the "maximum methane" equation (green). Bottom panel: corresponding BSE-measured ethane and propane concentrations.

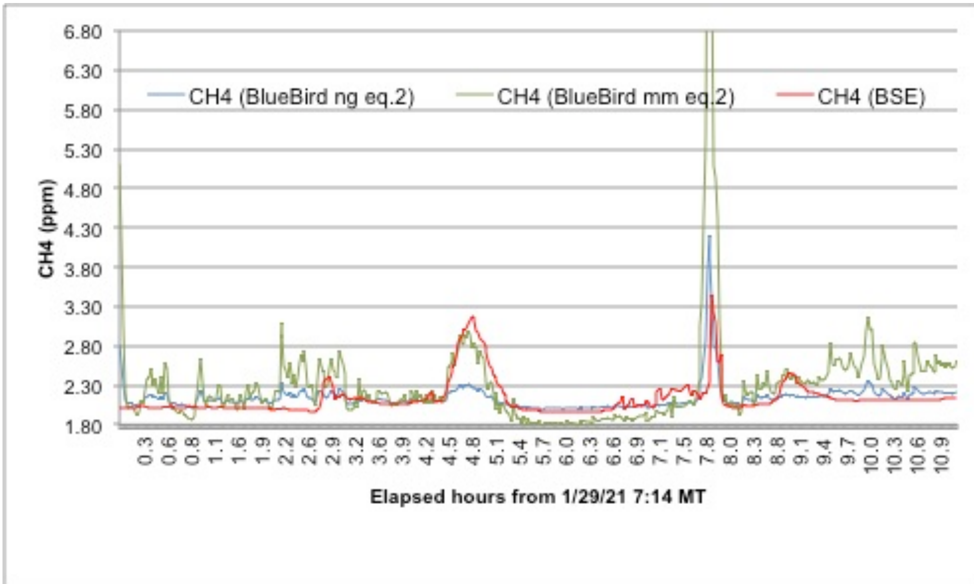

Figure 5. BSE-measured methane concentration (red) and methane concentration estimated from BlueBird data using the "natural gas" equation (blue) and the "maximum methane" equation (green).

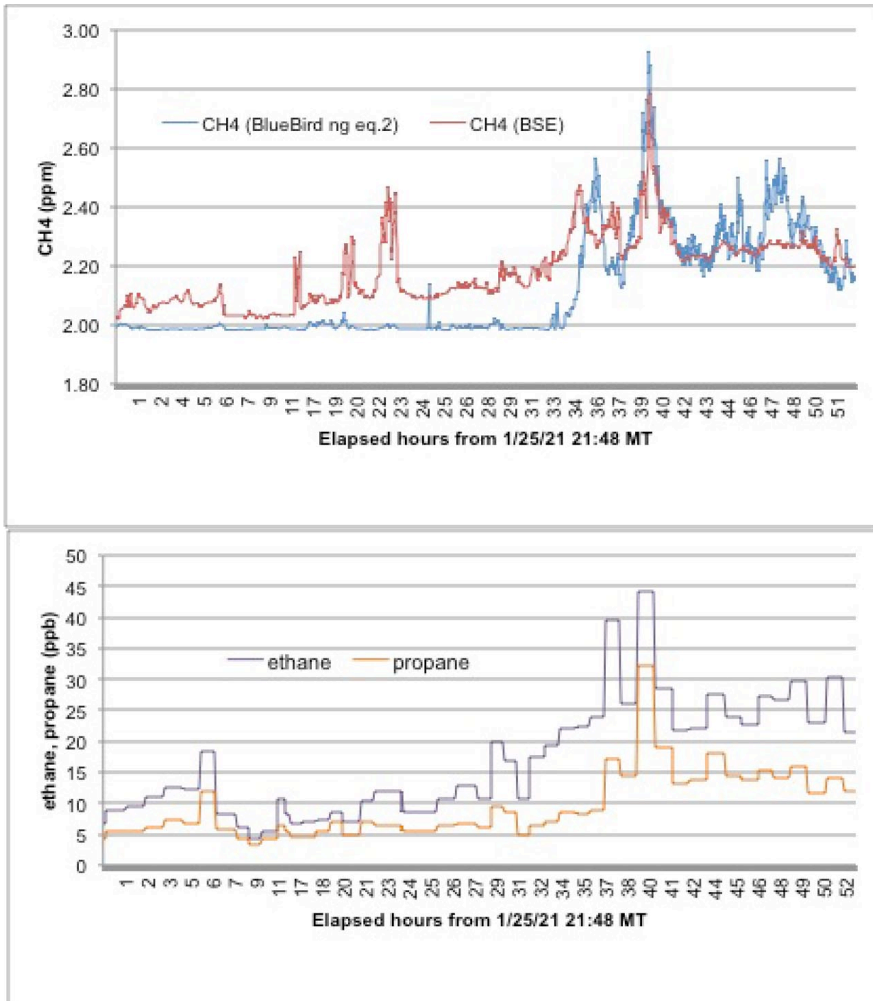

Figure 6. Top panel: BSE-measured methane concentration (red) and methane concentration estimated from BlueBird data using the "natural gas" equation (blue). Bottom panel: corresponding BSE-measured ethane and propane concentrations.

4.2 Detection of relatively large methane peaks

For the study period, there were 20 separate events (as defined visually) where methane levels exceed 3.0 ppm in the BSE data and 10 such events where 3.5 ppm was exceeded. For these events, a corresponding peak was observed in the BlueBird data for 19 of the 20 3.0 ppm events and 10 of the 10 3.5 ppm events, based on a visual comparison of plots.

4.3 Overall correspondence of peaks

An additional visual comparison was done to determine correspondence of peaks with smaller variations in methane. The criterion used was a BSE methane spike of 2.5 ppm or above, along with a corresponding visually noticeable spike in BlueBird-derived concentrations.

Using this visual analysis we found that there were corresponding peaks 67% of the time (57 of 84 cases) using the "natural gas" BlueBird equation and 76% of the time using the "maximum methane" equation. There were 15 and 17 "false" peaks for the two equations, respectively. As referred to earlier, many of these false peaks could have been filtered out using readings from other sensor(s) in the BlueBird package, where the combination of readings from the multiple sensors is inconsistent with a natural gas-dominant signal.

4.4 Means and occurrences

Another approach to assess the overall agreement between the BSE methane observations and the BlueBird estimates is to consider mean values and the number of cases where both data types agree in showing high readings. Also significant is the number of cases where the BlueBird data might fail to detect leaks or indicate leaks that do not exist (i.e., where BlueBird either "misses" a high reading or indicates "false" high readings).

For the full set of 54259 matched observations, the mean methane concentrations are 2.09 ppm, 2.08 ppm, and 2.20 ppm for the BSE data, BlueBird "natural gas (ng) equation" estimates, and BlueBIrd "maximum methane (mm) equation" estimates, respectively. For only those BSE observations where methane concentration was greater than 2.5 ppm, the mean values are 2.85 ppm, 2.41 ppm, and 3.47 ppm for BSE, BlueBird ng and Bluebird mm data, respectively. Of the full set of observations, 3.3% exceeded 2.5 ppm of methane in the BSE data. In the BlueBird methane data, 3.9% of the concentrations derived using the ng equation and 16.5% of the mm equation concentrations exceeded 2.5 ppm.

High concentrations (observations where BSE methane exceeds 2.5 ppm) are "missed" in the BlueBird data (i.e., BSE > 2.5 ppm and BlueBird <= 2.5 ppm) for 2.1% of the observations, or 1.5% of the observations if the BlueBird threshold is set to 2.3 ppm. "False" peaks, e.g., observations where BlueBird ng equation data indicate high concentrations but BSE does not (BlueBird > 2.5 ppm and BSE <= 2.5 ppm), make up 2.7% of the total. Increasing the BlueBird threshold necessarily decreases this "false" percentage. Using a threshold of 2.8 ppm reduces the number of "false" observations to less than 1% of the total observations.

For these BSE observations with concentrations greater than 2.5 ppm, the BlueBird concentrations are also above 2.5 ppm for 36% of the corresponding ng observations and 63% of the mm observations. In other words, for the cases where BSE-measured methane exceeds 2.5 ppm, BlueBird yields high readings for one-third to two-thirds of those cases, depending on methane algorithm used. If the threshold used for the BlueBird data is lowered to 2.3 ppm, then the percentages increase to 54% and 68%, respectively.

5. Summary

Near-coincident data collected by a BlueBird system co-located with the Boulder A.I.R. air monitoring site (BSE) at Soaring Eagle Park at Anthem Ranch in Broomfield County, Colorado are used to assess the ability of the BlueBird instrument to detect and quantify changes in methane versus background conditions. The data considered here were collected at 70-second intervals for 44 days spanning a 3-month period, under a variety of weather conditions. Methane concentrations were estimated from the BlueBird measurements using a pre-determined equation and two equations calculated from a subset of the Boulder A.I.R. BSE data.

The analyses described here suggest that the BlueBird system realistically captured the minute-to-minute variability in methane concentration as measured by the Boulder A.I.R. instrument at the BSE site for the field conditions encountered. The BlueBird data typically show a visually detectable response (i.e., provide a signal above the general noise level) to increases of 1 ppm or greater versus background concentrations.

"False" peaks occur but amount to a relatively small percent of observed peaks and, in some cases, could be filtered out using data from other sensors in the BlueBird package. Estimated concentration amounts vary depending on which conversion equation is used, which in turn reflects differences in natural gas composition. Specifically, accuracy of estimated concentrations depends in part of how well the particular equation used represents the "wetness" of the natural gas mixture being sampled. Even with this uncertainly in play, the estimates would likely be useful for calculating ranges of concentration and for estimating leak rates using an approach like inverse Gaussian plume modeling. Given that the nearest known potential source of natural gas emissions at the BSE location is about 700 m from the monitoring instruments, the level of sensitivity demonstrated here suggests that the BlueBird system could detect relatively small leaks on a well pad but also have the ability to characterize general "background" atmospheric conditions by monitoring changes over an extended time period. This would help in discriminating between small, local leaks versus a larger leak that might be originating off site.

Acknowledgments We would like to thank the City and County of Broomfield, Colorado, and Boulder A.I.R. Ltd. for providing access to the data collected by Boulder A.I.R. and for allowing installation of three BlueBird units on site.