

## ***Interactive comment on “Methane leak detection and sizing over long distances using dual frequency comb laser spectroscopy and a bootstrap inversion technique” by Caroline B. Alden et al.***

**Anonymous Referee #2**

Received and published: 20 October 2017

### General Comments

In atmospheric “inverse dispersion” problems, we attempt to deduce something about gas transfer to or from the atmosphere by (i) sampling the gas concentration field, (ii) simultaneously gathering sufficient meteorological data to characterize atmospheric transport (mean wind and turbulence), then (iii) invoking an atmospheric dispersion model to make inferences about the source(s). Within that overall scheme there are many variants. Such problems arise and can be solved on every meteorological scale

C1

of atmospheric (or, for that matter, oceanic) motion. Sometimes one knows the exact location of the source(s) and the object is their quantification; sometimes deducing source location (in space and or time) is the most important aim; and sometimes one seeks to deduce source space-time location *and* strength.

This paper is highly relevant to the problem of determining by inverse dispersion the spatial location(s) of point- or near-point sources in the context of gas leak detection (specifically methane) from industrial sites, and an element that is stressed is the avoidance of false positives (for potential leak locations) while not failing to find actual (non-zero) source locations. The authors briefly describe what appears to be a highly capable and useful instrument for their purpose (though describing the instrument is not the key aspect of this paper), but mostly focus on their statistical strategy for optimizing the useful information that can be extracted from their “measurements” (the quotes, because the paper invokes measurements both synthetic and real). I have not been able to fully comprehend that statistical strategy, and I feel the recipe for it can and should be clearer. Some readers may wonder why a Bayesian framework has not been adopted, as perhaps the most rational way to make use of prior information and account for uncertainties.

Setting aside that aspect of the paper, to my mind the most vulnerable aspect of the authors’ methodology is their use of the Gaussian plume model as their atmospheric transport paradigm, which treats the turbulent surface layer wind as if it were a regime of unshaped homogeneous turbulence. The authors do recognise that the Gaussian plume model is highly simplistic, and I can accept their argument that its use in the context of their paper is acceptable. But to drive home the importance of the choice of wind model, I must stress that the mean wind speed and the effective eddy diffusivity vary radically with height across the atmospheric surface layer (ASL), in a manner that is well described by Monin-Obukhov similarity theory. There are much higher fidelity models available, and one of these could be substituted without great penalty in terms of computational burden. It is true that at some sites, obstacles or topography may

C2

disturb the transporting wind field such that it is more complex than envisaged even by the better models (some of which are listed below), but they cannot be a worse choice than the Gaussian plume model, which is in effect a mental straitjacket.

### Specific Comments

1. The authors' admit that the Gaussian plume model (GPM) is highly simplistic, and I accept their argument that its use in the context of their paper is acceptable. I would disagree with their *categorical* (i.e. no exceptions) assertion that the GPM is "more suitable" if the inversion is based on line-averaged concentration data. The GPM is unnecessarily simple. There are much better analytical solutions that could and should be used to evaluate the **H** matrix. Whereas the GPM entirely neglects mean wind shear and treats the eddy diffusivity as constant (the atmospheric surface layer being represented as a regime of unsheared homogeneous turbulence), better models represent the ASL with a mean wind shear and a height-dependent eddy diffusivity that are consistent with the state of the ASL as parameterised by the Obukhov length and friction velocity. These solutions can be rapidly calculated; and where they provide (only) cross-wind integrated concentrations, the authors could easily introduce Gaussian *crosswind* spread (note: it is hard to improve on the assumption of Gaussian crosswind spread without using measured data on wind direction fluctuations).

A better analytical (or semi-empirical) eddy diffusion model could be sought out from the following references: Philip (1959), Ermak (1977), van Ulden (1978), Nieuwstadt and van Ulden (1978), Huang (1979), Lupini et al. (1981), Wilson (1982), Kormann and Meixner (2001), Sharan & Kumar (2009), Wilson (2015).

2. p5, line 30: Is there an easy argument that the optical detector's response is to line-averaged methane mole fraction as opposed to line averaged mass concentration [ $\text{kg m}^{-3}$ ]?

C3

3. At p6 line 5, why "attempts to solve"?
4. I don't understand why the authors contend (p6, line 19) that (in general, with  $m$  source locations and  $n$  concentrations) their "problem is overdetermined" — do they assume  $n > m$ ?
5. Why (p6 lines 22-23) should it be the case that (or why is it a safe assumption that) "model-data mismatch uncertainty has an un-biased Gaussian distribution"?
6. I find the derivation (p6 line 31 to p7 line 2) hard to follow, yet I suspect it has to be very simple. For instance we have the three symbols  $\epsilon_{Ri}$ ,  $\epsilon_R$  and  $\epsilon_{bi}$ : is the last just the first, in alternative guise?
7. If I have understood correctly, the pool of residual values is a set containing only  $n$  members, where  $n$  is the number of concentration measurements ( $n = 3$  for the field test). Then, for each detector ( $i = 1 \dots n$ ) one randomly draws 1000 samples from that set, with replacement, thereby obtaining a set of 1000 alternative model predictions  $y_{bi}$  for (each) source location. The logic for this is not very sound, it seems to me, because all observations are given equal status, irrespectively of their distance from the source(s). In real world cases there could be order-of-magnitude differences in the measured mean concentrations — and indeed in concentration variance and higher moments that, although irrelevant here, surely relate to the trustworthiness or representativeness of a measurement — and in the level of uncertainty in the modelling.
8. At p7, line 13, it might be helpful to be more specific as to what "law of large numbers" means in this context. I expect it amounts to an assumption that the distribution of some mean value (or sum) is Gaussian, even if the numbers being summed do not have a Gaussian distribution (central limit theorem)?
9. At p9 line 21, the authors allude to "model-data mismatch noise." It seems to me that "noise" type errors (which can be dealt with by averaging) are in practice likely

C4

to be far less serious than *systematic* errors arising from the imperfect modelling of atmospheric transport.

10. How is gas-gas interference dealt with? What about detection of “stray” infra-red radiation emitted from the environment and/or from within the telescope?
11. Inversion of the **H** matrix can entail severe error if the matrix is “ill-conditioned” as a result of the relative positioning of the sources and detectors: for example if sources are aligned along the wind direction (see Crenna et al. 2008; Flesch et al. 2003).
12. Using the GPM entails selection of appropriate  $\sigma$ -curves: the choice should be documented.
13. The Conclusion does not, to my mind, sufficiently recognise the potential accuracy gain from using a more sophisticated and realistic atmospheric transport model.
14. I wondered whether the authors are familiar with normal practice in regard to averaging, in micrometeorology. To make sense of surface layer winds of course we necessarily must use statistics, and it is usual to base those on an averaging interval of at least about 15 minutes. However the discussion caused me to suspect that the authors envisage using nearly-instantaneous wind measurements (or, say, 1 min averages) to deduce leak rates from concentration averages over (say) one or two minutes. This may be feasible, indeed it may be a good idea (at the least, it would speed up a search across many potential leak sites). The thing is, however, that one has to recognize that existing, documented, tested, trusted surface layer dispersion models exploit a statistically stable estimate of surface layer state, and that statistical stability demands averaging intervals much longer than a minute (or two). To adapt existing atmospheric transport modes to very

C5

sort averaging intervals one will at the least have to adjust the parameterisation of the eddy diffusivity.

#### Technical corrections

1. First paragraph of Sec. 2.6.2 uses mixed tenses (present then past).
2. Figure 3 shows only two beams for the field test, whereas the discussion of the experiment alludes to three.
3. The number of panels on Figures (5,6) seems excessive — is it necessary to cover the range in MDM (model-data-mismatch) with such fine steps.
4. The structuring of this paper results in a degree of repetition.

#### References

- Crenna, B.P., Flesch, T.K., & Wilson, J.D. 2008. Influence of Source-Sensor Geometry on Multi-Source Emission Rate Estimates. *Atmos. Environ.*, **42**, 7373–7383. DOI: doi:10.1016/j.atmosenv.2008.06.019.
- Ermak, Donald L. 1977. An Analytical Model for Air Pollutant Transport and Deposition from a Point Source. *Atmos. Environ.*, Vol. **11**, 231–237.
- Flesch, T.K., Harper, L.A., Desjardins, R.L., Gao, Z., & Crenna, B.P. 2009. Multi-source emission determination using an inverse-dispersion technique. *Boundary-Layer Meteorol.*, **132**, 11–30.
- Huang, C.H. 1979. A Theory of Dispersion in Turbulent Shear Flow. *Atmos. Environ.*, **13**, 453–463.
- Kormann, R., & Meixner, F.X. 2001. An Analytical Footprint Model for Non-Neutral Stratification. *Boundary-Layer Meteorol.*, **99**, 207–224.
- Lupini, Renzo, & Tirabassi, Tiziano. 1981. A Simple Analytic Approximation of the Ground-Level Concentration for Elevated Line Sources. *J Appl Meteorol*, **20**, 565–570.

C6

- Nieuwstadt, F. T. M., & Ulden, A. P. Van. 1978. A Numerical Study on the Vertical Dispersion of Passive Contaminants from a Continuous Source in the Atmospheric Surface Layer. *atmos. envt*, **12**, 2119–2124.
- Philip, J. R. 1959. The Theory of Local Advection. *J. Meteorol.*, **16**, 535–547. (I have clarifying correspondence with JRP in a folder.)
- Sharan, M., & Kumar, P. 2009. An Analytical Model for Crosswind Integrated Concentrations Released from a Continuous Source in a Finite Atmospheric Boundary Layer. *Atmos. Environ.*, **43**, 2268–2277.
- van Ulden, A. P. 1978. Simple Estimates for Vertical Diffusion From Sources Near the Ground. *Atmos. Environ.*, **Vol. 12**, 2125–2129.
- Wilson, J.D. 1982. An Approximate Analytical Solution to the Diffusion Equation for Short-Range Dispersion from a Continuous Ground-Level Source. *Boundary-Layer Meteorol*, **23**, 85–103.
- Wilson, J.D. 2015. Dispersion from an area source in the unstable surface layer: an approximate analytical solution. *Q.J.R. Meteorol. Soc.*, **141**(693), 3285–3296.

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-262, 2017.