



# Optimal experimental designs for estimating Henry's law constants via the method of phase ratio variation



Adam Kapelner<sup>a,\*</sup>, Abba Krieger<sup>b,\*\*</sup>, William J. Blanford<sup>c,\*\*</sup>

<sup>a</sup> Department of Mathematics, Queens College, The City University of New York, 65-30 Kissena Blvd., Kiely Hall Suite 605, Flushing, NY 11367, United States

<sup>b</sup> Department of Statistics, The Wharton School of the University of Pennsylvania, 3730 Walnut St, Jon M. Huntsman Hall Suite 442, Philadelphia, PA 19104, United States

<sup>c</sup> School of Earth and Environmental Sciences, Queens College, The City University of New York, 65-30 Kissena Blvd., Science Building Suite D202, Flushing, NY 11367, United States

## ARTICLE INFO

### Article history:

Received 12 April 2016

Received in revised form 15 August 2016

Accepted 22 August 2016

Available online 27 August 2016

### Keywords:

Henry's coefficient

Henry's law constant

Headspace gas chromatography

Phase ratio variation method

Slope divided by intercept

Optimal experimental design

Linear regression

Volatile organic compounds

## ABSTRACT

When measuring Henry's law constants ( $k_H$ ) using the phase ratio variation (PRV) method via headspace gas chromatography ( $G_C$ ), the value of  $k_H$  of the compound under investigation is calculated from the ratio of the slope to the intercept of a linear regression of the inverse  $G_C$  response versus the ratio of gas to liquid volumes of a series of vials drawn from the same parent solution. Thus, an experimenter collects measurements consisting of the independent variable (the gas/liquid volume ratio) and dependent variable (the  $G_C^{-1}$  peak area). A review of the literature found that the common design is a simple uniform spacing of liquid volumes. We present an optimal experimental design which estimates  $k_H$  with minimum error and provides multiple means for building confidence intervals for such estimates. We illustrate performance improvements of our design with an example measuring the  $k_H$  for Naphthalene in aqueous solution as well as simulations on previous studies. Our designs are most applicable after a trial run defines the linear  $G_C$  response and the linear phase ratio to the  $G_C^{-1}$  region (where the PRV method is suitable) after which a practitioner can collect measurements in bulk. The designs can be easily computed using our open source software `optDesignSlopeInt`, an R package on CRAN.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Henry [1] observed that the amount of gases such as carbon dioxide, hydrogen sulfide, and others were taken up by water at a particular temperature were proportional to their partial pressures. Subsequently referred to as Henry's Law Constant or Henry's Coefficient ( $k_H$ ), it is a ubiquitous metric especially critical in chemical processing and environmental sciences. Mackay et al. [2] among others has compiled measurements of Henry's Constant for over a thousand compounds for pure water under standard temperature and pressure as well as additional salinities and temperatures using a variety of methods. One measurement method is called the "Phase Ratio Variation Method" (PRV), and was developed by Ettre et al. [3]. Its development marked a substantial improvement over previous

methods which required external standards of known concentration.

In brief, the PRV method involves first subsampling from a stock liquid containing a dilute concentration of a volatile organic compound (VOC) and adding varying amounts of that liquid to a series of standard headspace vials thus creating an array of samples with a range of gas to liquid volume ratios. After equilibration under the same temperature and pressure conditions, the headspace of the vials is sampled and analyzed by gas chromatography (GC). The dimensionless form of the Henry's Constant can then be determined from the slope divided by the intercept of the best fit line through the inverse of the  $G_C$  peak area ( $G_C^{-1}$ ) measurements versus the gas to liquid volume ratios:

$$G_C^{-1} = \frac{\gamma}{mk_H} + \frac{\gamma}{m} \frac{V_{\text{gas}}}{V_{\text{liquid}}} \Rightarrow \frac{\gamma}{m} / \frac{\gamma}{mk_H} = k_H \quad (1)$$

where  $m$  is the total solute mass between the gas and the liquid phases and  $\gamma$  is the proportionality constant between  $G_C$  peak area and gas phase concentration (for a derivation of this relationship, see Eqs. (2)–(8) in Ramachandran et al. [4]).

\* Principal corresponding author.

\*\* Corresponding author.

E-mail addresses: [kapelner@qc.cuny.edu](mailto:kapelner@qc.cuny.edu) (A. Kapelner), [krieger@wharton.upenn.edu](mailto:krieger@wharton.upenn.edu) (A. Krieger), [william.blanford@qc.cuny.edu](mailto:william.blanford@qc.cuny.edu) (W.J. Blanford).

Note that the PRV method determines the dimensionless form of Henry's constant. It is a simple exercise to convert this form to other forms e.g. pressure as a function of mole fraction in solution, pressure vs. molality, pressure vs. molarity, etc. [5, page 79]. As with most analytical instruments in this field, it is commonly necessary to analyze samples of known composition to determine standard parameters (such as aqueous solubility or vapor pressure). However, this calibration is minimal with the PRV method (see Chaintreau et al. [6] for more discussion) and mostly done to ensure proper operation of the equipment.

It should be noted that the PRV method as developed by Ettre et al. [3] assumes a linear response for the  $G_C$  detector. This is a justified assumption; most detectors having a linear dynamic range of three to seven orders of magnitude in concentration. At the upper and lower ends of the detectable range, non-linear behavior is commonly observed. Atlan et al. [7] developed a modified version of the PRV method for use in this non-linear range of  $G_C$  detectors. They found the best results featured repeated analysis of an analyte over a range of concentration and for several different detectors with different relative response functions. Herein, we consider measurements only within this linear range with solutes that are not extremely volatile (Ettre et al. [3] observed non-linearity in solutes with high Henry's Constant e.g.  $k_H > 144$ ).

In addition, to use the PRV method to measure Henry's coefficient, one must operate within this *linear PRV range*. At small liquid volumes, solute and solvent loss during sample preparation renders the  $V_{\text{gas}}/V_{\text{liquid}}$  measurement too high, yielding a non-linear PRV relationship in Eq. (1). At relatively large liquid volumes and high analysis temperatures, solvent volatilization and resulting elevated gas phase pressurization can result in significant vapor loss from the vial and possibly violate the underlying assumption that all the test vials are being measured at atmospheric pressures, ultimately resulting in a non-linear PRV relationship. In addition, limiting measurements to the linear PRV range will also keep measurement errors relatively constant; taming possible "heteroskedasticity", a phenomenon we will discuss at length.

Thus, we limit the scope of our designs to gas–liquid volume ratios which are neither too large or too small to guarantee a linear PRV relationship. Of course, this range is not known a priori for every solute–solvent–temperature scenario. Thus, we recommend that the practitioner investigate the linear range by performing a trial run measuring  $G_C$  across a wide range of  $V_{\text{gas}}/V_{\text{liquid}}$  ratios. Once the range is established the practitioner can compute our proposed design and proceed to take  $k_H$  measurements in bulk.

Once both ranges are guaranteed linear ( $G_C$  response and PRV), we can measure  $k_H$  by computing a best-fit line. To do so, we employ the method most commonly in use, ordinary least squares linear regression. As an example of the widespread use of this method, Table 1 shows example measurements by previous studies, examples which we will return to when we discuss our performance gains in Section 4.

In essence, the volume ratio is the independent variable in the PRV method whose values are free to be chosen by the experimenter within the linear PRV range. Thus, our problem is one of "optimal design:" which volume ratios should be chosen to measure  $k_H$  with the least error?

Optimal design is of enormous importance and has been studied for nearly 100 years since Smith's [8] paper on designs for polynomial regression and Fisher's [9] advice for treatment comparison experimentation. The field was formalized beginning with Kiefer's [10] paper which laid the foundations for further work. Equivalence theorems for optimality criteria were to follow. Modern goals include variable screening, response surface exploration, system optimization and optimal inference [11] which is our focus here. There are many good textbooks written on the subject for the interested reader e.g. Pukelsheim [12].

Given the heterogeneity in experimental conditions, covariate domains, parameters of interest and error structures, it is difficult to provide universal optimal designs. Thus, some work in this field focuses on tailoring designs to specific applications. An optimal design for the specific application of estimating the slope-to-intercept ratio and inference (confidence interval construction) for such estimates to our knowledge has not been studied in detail (especially in the application settings of the PRV method) nor has specialized software been developed for this application. This is the modest goal herein.

Optimization specific to PRV is of great interest [13]. We would like to stress that naively estimating the slope-to-intercept ratio is dangerous: the sample slope divided by the sample intercept estimator has infinite moments and therefore can vary wildly; estimates far away from the true value are all too common in the low sample laboratory setting with realistic measurement error. Simulations demonstrate our design can achieve gains of orders of magnitude smaller in standard error of the estimator. Since  $k_H$  values are widely used in vital calculations of the phase distribution and total levels of volatile solutes at hazardous wastes (e.g. for projection of the performance of air strippers, studying urban air pollution estimation, landfill gas generation, waste water treatment among other applications), improvements in accuracy for their determination offers substantial societal benefits.

The paper's outline is as follows. We describe our improved design for homoskedastic and heteroskedastic data in Section 2. We discuss many strategies for building confidence interval at the end of this section. We then illustrate an application of the optimal design by employing it to estimate the  $k_H$  of a common persistent organic pollutant, naphthalene, in an aqueous solution in Section 3 along with simulations of interval performance and robustness to a priori parameter decisions. This section also demonstrates how the software, written in R [14], is used by an experimenter in this laboratory setting. In Section 4, we use simulation to estimate performance gains of our improved design in the previous studies listed in Table 1. We conclude and discuss the wider applicability of the design in Section 5.

**Table 1**  
Information about 10 previous studies who use the PRV method to measure  $k_H$  for a variety of VOC's in reverse chronological order.

#	Authors	Solutes	Solvent
1	van Durme and Werbrouck [11]	Fifty indoor VOCs	Nalophan
2	Benjamin et al. [2]	Five VOCs	Oil-in-water emulsions
3	Gao et al. [5]	BTEX and chlorinated solvents	Cyclodextrin aqueous solutions
4	Kechagia et al. [7]	Two monomers	Water
5	Lei et al. [8]	Alkanols	Water
6	Atlan et al. [1]	Twelve VOCs	Propylene glycol
7	Jouquand et al. [6]	Eight VOCs	Cyclodextrin aqueous solutions
8	Chai and Zhu [3]	Methanol	Water
9	Peng and Wan [9,10]	BTEX and chlorinated solvents	Water and saline waters
10	Ettre et al. [4]	Four VOCs	Water

Download English Version:

<https://daneshyari.com/en/article/5135880>

Download Persian Version:

<https://daneshyari.com/article/5135880>

[Daneshyari.com](https://daneshyari.com)