Estimation of inhaled ultrafine particle surface area dose for urban environments

S. Clifford[1] M. Mazaheri[2] E. R. Jayaratne[3]
M. A. Megat Mokhtar[4] F. Fuoco[5] G. Buonanno[6]
L. Morawska[7]

Abstract

There is considerable scientific interest in personal exposure to ultrafine particles. Owing to their small size, these particles are able to penetrate deep into the lung where they may cause adverse respiratory, pulmonary and cardiovascular health effects.

This article presents Bayesian hierarchical models for estimating and comparing inhaled particle surface area in the lung.

Contents

1 Introduction

Ultrafine particles (UFPs), particles less than 100 nm in diameter, account for the overwhelming majority of the number of outdoor airborne particles in Brisbane, Australia, the majority of which are generated by traffic [10]. While recent studies are inconclusive on the health effects of UFPs, the link between air quality and health is well known [5]. UFPs pose a unique risk,

compared to coarser particles, as they can be inhaled deep into the lungs where they may be absorbed into the bloodstream [11].

High frequency measurements of ultrafine particle number concentration (PNC) with mobile sampling devices such as the Philips Nanotracer [8] make it possible to estimate personal exposure to UFPs [3, 4, 9]. Such devices are able to sample from a location on the person's body, rather than a fixed location in an environment where the person spends a large portion of time. The total daily personal dose is apportioned to the various microenvironments which are determined from an activity diary or from directly measured spatial information such as GPS coordinates.

This article describes a modelling methodology for estimating the inhaled particle surface area dose deposited within the lung. The model is based on size-specific particle deposition within the different regions of the lung, measurements of PNC in the breathing zone and known and estimated parameters describing the particle size distribution.

2 Methodology

Bayesian regression models were developed to estimate inhaled doses in a number of microenvironments across a study region, the daily dose portion in each microenvironment, and the dose intensity relative to the time spent in each microenvironment. Here, all Bayesian models are estimated using R statistical software and Just Another Gibbs Sampler (JAGS) through the rjags package [12]. JAGS simulates the posterior distribution of the model parameters using Markov chain Monte Carlo (MCMC) techniques.

2.1 Inhaled dose

The model for inhaled particle surface area dose in the alveolar (AL) and tracheo-bronchial (TB) regions of the lung extends a previous model [3] to incorporate the particle size distribution. The efficiency of particle deposition in the lung, φ, in the lung varies with particle diameter, x_p, in micrometres, and region of the lung [6],

$$\begin{aligned}\varphi_{AL}\left(x_p\right) &= \frac{3.52\times10^{-3}}{x_p}\left(e^{-0.234(\log x_p+3.40)^2}+63.9e^{-0.819(\log x_p-1.61)^2}\right)\\ \varphi_{TB}\left(x_p\right) &= \frac{1.55\times10^{-2}}{x_p}\left(e^{-0.416(\log x_p+2.84)^2}+19.11e^{-0.482(\log x_p-1.362)^2}\right).\end{aligned} \tag{1}$$

Personal air quality samplers do not report the full particle size distribution for a sample of air, but is reconstructed by assuming a log-normal den-

sity, $f\left(x_p; \overline{x_p}(t), \sigma_k\right)$, with a location parameter based on the average particle diameter (measured by the Nanotracer), $\overline{x_p}(t)$, and some shape parameter, σ. The particle size distribution shape parameter varies across the different microenvironments and is estimated based on prior studies with a Scanning Mobility Particle Sizer ((SMPS)) system [4]. An individual's inhalation rate, IR_k, also varies with microenvironment, based on the type of activity engaged in at the time, as well as age. These inhalation rates are estimated from previous studies [4].

For individual i in microenvironment k, the total inhaled surface area dose in the region of lung r is

$$d_{ikr} = \sum_{j=1}^{J} \mathrm{IR}_k \int_{T_{jk1}}^{T_{jk2}} \int_{0}^{\infty} N(t)\, f\left(x_p; \overline{x_p}(t), \sigma_k\right) \varphi_r\left(x_p\right) \pi x_p^2 \,\mathrm{d}x_p \,\mathrm{d}t \tag{2}$$

where T_{jk1} and T_{jk2} are, respectively, the start and end times of block j of J which the individual spends in microenvironment k and $N(t)$ are the PNCs as measured by the Nanotracer.

The model is implemented in R by looping over microenvironments within a loop for individuals (identified by a unique ID), selecting a block of contiguous time spent in that microenvironment and calculating the double integral in equation 2. Integration over the range of particle diameter, from 0 to ∞, is performed using adaptive quadrature in R's `integrate` function; time integration is performed using the trapezoid rule. Integration with an unbounded upper limit ensures that the entire mass of the log-normal density, f, is included even though it is almost zero beyond 3000 nm.

2.2 Spatial variation

To analyse the relationship between the inhaled doses within different study area subregions and microenvironments, Bayesian hierarchical linear models were developed. Bayesian hierarchical linear modelling casts the parameters as random variables described by a probability density. This is in contrast to classical ANOVA and linear regression models which assume that model parameters are unknown but fixed and rely on asymptotic properties of estimators. The incorporation of *a priori* beliefs about these parameters, whether assumed to be informative or not, is done by specifying a prior distribution for the parameters. These prior distributions are updated by the data toobtain a distribution representing *a posteriori* beliefs.

To estimate how microenvironment doses varies between subregions (or

cohorts), a model with exchangeable random effect means is formulated as

$$\begin{aligned}(d_{ik}|z_i = j) &\sim \log\mathcal{N}\left(\beta_{jk}, \tau_d\right)\\ \beta_{jk} &\sim \mathcal{N}\left(\alpha_k, \tau_k\right)\\ \alpha_k &\sim \mathcal{N}\left(0, 10^{-6}\right)\\ \tau_d, \tau_k &\sim \Gamma\left(0.001, 0.001\right).\end{aligned} \tag{3}$$

where β_{jk} is a location parameter for dose in microenvironment k within spatial subregion $z_i = j$ and α_k is the all-region location parameter for each microenvironment with a conjugate weakly informative prior. The normal distributions are parameterised in terms of a mean (location) and precision (shape; inverse of the variance) and weakly informative conjugate Gamma distributions are applied to the precision parameters. A comparison of the dose received in two different microenvironments (e.g. $k = 1, 2$) for any given spatial subregion is given by calculating the posterior difference in means, $\gamma_j = \beta_{j1} - \beta_{j2}$. This model allows comparisons to be made across subregions and across microenvironments similar to an ANOVA but without assuming independence of the levels of the grouping variables; individuals are assumed exchangeable within subregions and these subregions are also exchangeable within each microenvironment type.

2.3 Microenvironment proportions

To estimate the average proportion of each individual's total daily dose received in each of the microenvironments, a multinomial model was developed. The multinomial model ensures that credible intervals for all proportion parameters remain within the range $[0, 1]$. The model assumes exchangeability across individuals and the parameters of interest are the multinomial proportions, p_k, derived from the parameters of the multinomial distribution, θ_k. Each individual's dose in each microenvironment, d_{ik}, is some proportion of their 24 hour total inhaled dose, n_i,

$$\begin{aligned} y_{ik} &= \lfloor d_{ik}/n_i \times 100 \rceil & n_i &= \sum_{k=1}^{K} d_{ik}\\ y_{i1} &\sim \text{Bin}\left(p_1, n_i\right) & y_{ik}|k \neq 1 &\sim \text{Bin}\left(p_k, n_i - \sum_{m=1}^{k-1} d_{im}\right)\\ p_1 &= \theta_1 & p_k|k \neq 1 &= \theta_k \left(k - \sum_{m=1}^{k-1} \theta_m\right)^{-1}\\ \vec{\theta} &\sim \text{Dirichlet}\left(\vec{\alpha}\right) & \alpha_k &\sim \Gamma(2, 2). \end{aligned} \tag{4}$$

The multinomial distribution generalises the binomial distribution to more than two categories and is parameterised as a series of binomials. This is required as JAGS is not able to fit a multinomial model with zero counts [1], such as when an individual does not spend any time within a particular microenvironment. The Dirichlet prior for the multinomial proportions ($\vec{\theta}$) represents a belief about the relative proportions in each microenvironment in the absence of data. The Gamma prior for the Dirichlet parameters, $\vec{\alpha}$, results in a weakly informative conjugate prior which assumes that the proportions are equal *a priori*.

This model is also applied to the proportion of time spent in each microenvironment. Each individual's dose intensity for each microenvironment, ρ_{ik}, the ratio of daily dose portion to daily time portion, indicates which microenvironments pose the greatest risk of an adverse health outcome. That is,

$$\rho_{ik} = \frac{d_{ik}}{\sum_{k=1}^{K} d_{ik}} \Bigg/ \frac{\sum_{j=1}^{J} T_{jk2} - T_{jk1}}{\sum_{k=1}^{K} \sum_{j=1}^{J} T_{jk2} - T_{jk1}} \tag{5}$$

which is a dimensionless number greater than 0, a ratio of dose proportionto time proportion. A dose intensity greater than 1 indicates that the individual received a higher portion of their total daily dose in that microenvironment than the portion of the day that they spent there. Inference on average dose intensities for each microenvironment, $\widehat{\rho}_k$, were obtained by fitting

$$\begin{aligned} \rho_{ik}, \widehat{\rho}_k &\sim \log \mathcal{N}\left(\beta_k, \tau_k\right) \\ \beta_k &\sim \mathcal{N}\left(0, 10^{-4}\right) \\ \tau_k &\sim \Gamma\left(0.001, 0.001\right) \end{aligned} \tag{6}$$

where the dose intensities, ρ, are assumed to be independent across the microenvironments and are distributed log-normally with location parameter β and shape parameter τ.

3 Discussion

The Bayesian hierarchical linear models developed here allow analysis which is more flexible than classical ANOVA approaches. By including spatial variation with a structured random effect mean, which is assumed exchangeable rather than independent across spatial locations, small area estimates of mean inhaled doses are calculated which are related to each other and informed by partial pooling of the data through the hierarchical model.

The multinomial model with MCMC sampling provides credible intervals for the proportion parameters which are strictly between 0 and 1 and do not

rely on assumptions of asymptotic behaviour. These are desirable properties for intervals expressing uncertainty in parameters which may be close to zero and are based on a small number of observations.

The dose model presented here extended previous work on dose calculation [3] by including particle size distribution estimation based on shape parameters [4]. This modelling approach was used to calculate the inhaled surface area dose in the alveolar and tracheo-bronchial regions of the lung for children attending 25 primary schools in the Brisbane Metropolitan Area [9].

The biggest outstanding issue in this modelling is the estimation of the shape parameter for each microenvironment. Point estimates result in underestimating the variability of the inhaled doses. Replacing the point estimates with a prior distribution $\sigma_{jk}^2 \sim p\left(\sigma_{jk}^2\right)$ gives a more physically realistic estimate. While portable size distribution analysers have been in existence for approximately 40 years [7], the technology is not yet at a point where a particle size distribution analyser can be worn comfortably without interfering with normal movement (the Grimm Model 1.109 aerosol spectrometer has 32 size channels in which it measures PNC and can be carried in a backpack but weighs 2.5 kg [2]).

Because personal samplers such as the Philips Nanotracer do not currently measure the size distribution, the shape parameters must be derived from nearby stationary monitors such as an SMPS system. As such, the measurement of particle size distributions (and calculation of the geometric standard deviation of particles) makes dose calculations across a large spatial domain very difficult. For well-mixed air in microenvironments where individuals spend much of their time, the models provided in this article can be used with stationary samplers such as Condensation Particle Counters and SMPS which more fully characterise the particle number concentration and size distribution as they have larger ranges of measurement than personal samplers.

Acknowledgements This work was supported by the Australian Research Council (ARC), QLD Department of Transport and Main Roads (DTMR) and QLD Department of Education, Training and Employment (DETE) through Linkage Grant LP0990134. Sam Clifford is supported by a postdoctoral fellowship from the Centre for Air quality health Research and evaluation. The authors thank R. Fletcher (DTMR) and B. Robertson (DETE), G. Marks, P. Robinson, K. Mengersen, Z. Ristovski, G. Ayoko, C. He, G. Johnson, S. Low Choy, G. Williams, W. Ezz, F. Salimi, L. Crilley, N. Mishra, R. Laiman, L. Guo, C. Duchaine, H. Salonen, X. Ling, J. Davies, L. Leontjew

Toms, A. Cortes, B. Toelle, A. Quinones, P. Kidd, E. Belousova, M. Falk, F. Fatokun, J. Mejia, D. Keogh, T. Salthammer, R. Appleby and C. Labbe.

References

[1] Isabelle Albert and Jean-Baptiste Denis. Dirichlet and multinomial distributions: properties and uses in JAGS. Technical Report 2012-5, Unité Mathématiques et Informatique Appliquées, INRA, 2012. URL `http://w3.jouy.inra.fr/unites/miaj/public/nosdoc/rap2012-5.pdf`.

[2] Anna V. Broich, Lydia E. Gerharz, and Otto Klemm. Personal monitoring of exposure to particulate matter with a high temporal resolution. *Environmental Science and Pollution Research*, 19:2959–2972, 2012. `http://dx.doi.org/10.1007/s11356-012-0806-3`.

[3] Giorgio Buonanno, Gaspare Giovinco, Lidia Morawska, and Luca Stabile. Tracheobronchial and alveolar dose of submicrometer particles for different population age groups in Italy. *Atmospheric Environment*, 45(34):6216–6224, 2011. `http://dx.doi.org/10.1016/j.atmosenv.2011.07.066`.

[4] Giorgio Buonanno, Lidia Morawska, Luca Stabile Lina Wang, and Gaspare Giovinco. A comparison of submicrometer particle dose between Australian and Italian people. *Environmental Pollution*, 169:183–189, 2012. `http://dx.doi.org/10.1016/j.envpol.2012.03.002`.

[5] HEI. Understanding the Health Effects of Ambient Ultrafine Particles: HEI Review Panel on Ultrafine Particles. Technical Report HEI Perspectives 3, Health Effects Institute, Boston, MA, 2013. URL `http://pubs.healtheffects.org/getfile.php?u=893`.

[6] ICRP. *Human Respiratory Tract Model for Radiological Protection: A Report of a Task Group of the International Commission on Radiological Protection.* Elsevier Science Ltd., Oxford, U.K., 1994. URL `http://www.icrp.org/publication.asp?id=ICRP%20Publication%2066`.

[7] Benjamin Y.H. Liu, Kenneth T. Whitby, and David Y.H. Pui. A portable electrical analyzer for size distribution measurement of submicron aerosols. *Journal of the Air Pollution Control Association*, 24 (11):1067–1072, 1974. `http://dx.doi.org/10.1080/00022470.1974.10470016`.

[8] Johan Marra, Matthias Voetz, and Heinz-Jügen Kiesling. Monitor for detecting and assessing exposure to airborne nanoparticles. *Journal of Nanoparticle Research*, 12(1):21–37, 2010. http://dx.doi.org/10.1007/s11051-009-9695-x.

[9] Mandana Mazaheri, Sam Clifford, Rohan Jayaratne, Megat Azman Megat Mokhtar, Fernanda Fuoco, Giorgio Buonanno, and Lidia Morawska. School children's personal exposure to ultrafine particles in the urban environment. *Environmental Science and Technology*, 48(1): 113–120, 2014. http://dx.doi.org/10.1021/es403721w.

[10] Lidia Morawska, Steven Thomas, Neville Bofinger, David Wainwright, and Don Neale. Comprehensive characterization of aerosols in a subtropical urban atmosphere: particle size distribution and correlation with gaseous pollutants. *Atmospheric Environment*, 32:2467–2478, 1998. http://dx.doi.org/10.1016/S1352-2310(98)00023-5.

[11] Günter Oberdörster. Pulmonary effects of inhaled ultrafine particles. *International Archives of Occupational and Environmental Health*, 74(1): 1–8, 2001. URL http://www.ncbi.nlm.nih.gov/pubmed/11196075.

[12] Martyn Plummer. *rjags: Bayesian graphical models using MCMC*, 2012. URL http://cran.r-project.org/web/packages/rjags/index.html. R package version 3-9.

Author addresses

1. **S. Clifford**, Science and Engineering Faculty, Queensland University of Technology, Brisbane Queensland 4000, Australia.
mailto:samuel.clifford@qut.edu.au

2. **M. Mazaheri**, Science and Engineering Faculty, Queensland University of Technology, Brisbane Queensland 4000, Australia.

3. **E. R. Jayaratne**, Science and Engineering Faculty, Queensland University of Technology, Brisbane Queensland 4000, Australia.

4. **M. A. Megat Mokhtar**, Science and Engineering Faculty, Queensland University of Technology, Brisbane Queensland 4000, Australia.

5. **F. Fuoco**, Department of Civil and Mechanical Engineering, University of Cassino and Southern Lazio, via Di Biasio 43, 03043 Cassino, Italy

6. **G. Buonanno**, Department of Civil and Mechanical Engineering, University of Cassino and Southern Lazio, via Di Biasio 43, 03043 Cassino, Italy

7. **L. Morawska**, Science and Engineering Faculty, Queensland University of Technology, Brisbane Queensland 4000, Australia.

Estimation of inhaled ultrafine particle surface area dose for urban environments

S. Clifford[1] M. Mazaheri[2] E. R. Jayaratne[3]
M. A. Megat Mokhtar[4] F. Fuoco[5] G. Buonanno[6]
L. Morawska[7]

A Model code

Table 1: JAGS code for dose comparison across subregions.

```
# i indexes each calculated dose
model {
  for (i in 1:N) {
    # data model
    y[i] ~ dlnorm(mu[i], tau.y)
    # random effects mean
    mu[i] <- beta[subregion[i], microenvironment[i]]
  }

  # for each subregion
  for (j in 1:n.subregions){
    for (k in 1:K){
      beta[j,k] ~ dnorm(alpha[k], tau[k])
      # predict at each subregion
      y.pred[j,k] ~ dlnorm(beta[j,k], tau.y) # indoors
    }
  }

  # hyper-parameters
  for (k in 1:K){
    alpha[k] ~ dnorm(0,1e-6)
    tau[k]   ~ dgamma(0.001,0.001)
  }

  tau.y ~ dgamma(0.001,0.001) # precision for data model
}
```

Table 2: JAGS code for Multinomial model.

```
# i indexes each calculated dose
# there are K microenvironments
model{
  theta[1:K] ~ ddirch(alpha);
  d.pred[1:K] ~ dmulti(theta[1:K],100);

  for (k in 1:K){
    alpha[k] ~ dgamma(2,2);
  }

  # the first microenvironment
  p[1] <- theta[1]
  # subsequent microenvironments
  for (k in 2:K){
    p[k] <- theta[k]/(1-sum(theta[1:(k-1)]))
  }

  for (i in 1:N){
    d[i,1] ~ dbin(pp[1], n[i])
    for (k in 2:K){
      d[i,k] ~ dbin(p[k], n[i] - sum(d[i,1:(k-1)]))
    }
  }
} # end model
```