Institution: University of Southampton

Unit of Assessment: 10 Mathematical Sciences

Title of case study: 10-08 Using Novel Statistical Modelling Techniques to Deliver More Accurate Air Pollution Forecasts

1. Summary of the impact

Working closely with scientists at the United States Environmental Protection Agency (USEPA), the University of Southampton has developed new methods for space-time modelling that have trebled the accuracy of air pollution forecasts. The USEPA has adopted the research as its official forecasting method to protect the American public and agriculture. More than 19 million children and 16 million adult Americans suffering from respiratory conditions such as asthma now benefit by being able to adjust their outdoor activities based on the forecasts, and improved data has fed into policy debates on carbon emission regulations. Success in the USA has led the EPSRC to fund a similar project in the UK and Australia's national science agency is using Southampton-developed software for its air pollution forecasts.

2. Underpinning research

Even only moderate levels of air pollution can damage the health of vulnerable people with lung disease – in particular sufferers of asthma, chronic bronchitis and emphysema - and heart disease. Air pollutants can also adversely affect healthy people who regularly exercise outdoors. Children are at greater risk because they often play outdoors in warmer weather when ozone levels are higher, their lungs are still developing, and they are more likely to have asthma which is aggravated by ozone exposure. The resulting healthcare costs are high; aggravations of lung diseases lead to increased medication use, GP and A&E visits, and hospital admissions.

Professor Sujit Sahu (1999-present) has spent the last decade conducting research in spatio-temporal statistical modelling and applications to atmospheric processes that are harmful to human health, publishing more than 25 peer-reviewed papers in this period. Having established an ongoing collaboration with Dr David Holland, a senior scientist at the United States Environmental Protection Agency (USEPA), and Professor Alan Gelfand at Duke University, North Carolina, in 2003, Sahu has focused on developing accurate methods for forecasting air pollution levels at any given location, even where there is no air pollution monitoring station nearby.

This need to spatially interpolate and forecast at unmonitored sites over vast regions spanning thousands of square kilometres, led to the development of Bayesian hierarchical models for spatio-temporal processes that accurately describe air pollution levels over space and time. This research revealed that statistical models that *combine* past observations obtained from monitoring sites with the output of numerical air-quality models are far superior to either standalone empirical observations or numerical models **[3.1, 3.2]**. This insight led to the development of a coherent Bayesian forecasting framework, which was shown to be substantially faster and three times more accurate than previous models. The framework was able to instantly update the air pollution map for the current hour as soon as the monitor data was received for that hour, and forecast the map for several hours ahead.

Once the key modelling and forecasting techniques proved superior in principle, the natural next step was to develop a software package that would be capable of delivering the forecasts in production mode in real time for use by forecasters and environmental agencies. Sahu worked with the USEPA to develop this software in 2010-2011. The result was a robust software system that is able to produce the forecasts and their statistical uncertainties for the entire United States. An educational and research version of the software package, spTimer, which can be used under a freely available GNU public license, is available through the online Comprehensive R Archive Network, opening up the improved forecasting techniques for use by any agency anywhere in the world.

Additional research has developed methods to estimate location-specific, long-term trends in air pollution **[3.3]**. As in the case of forecasting, air pollution is often only monitored at a handful of locations. Hence, accurate spatial interpolations along with their uncertainties must be obtained from data through statistical modelling. A high-resolution, space-time model was developed for this purpose that has been further adapted to model several million sets of space-time data.

3. References to the research

Publications:

3.1 (*) Sahu, S. K., Gelfand, A. E. and Holland, D. M. (2010) Fusing point and areal level space-time data with application to wet deposition. J*ournal of the Royal Statistical Society*, Series C, **59**, 77-103.

3.2 (*) Sahu, S. K., Yip, S., and Holland, D. M. (2009) A fast Bayesian method for updating and forecasting hourly ozone levels. *Environmental and Ecological Statistics,* **18**, 185-207.

3.3 Sahu, S. K. and Bakar, K. S. (2012) Hierarchical Bayesian auto-regressive models for large space time data with applications to ozone concentration modelling, with discussion. *Applied Stochastic Models in Business and Industry*, **28**, 395-415.

(*) These references best indicate the quality of the underpinning research.

Grants:

3.G1 Sahu, S. Modelling large space-time data sets, The National Academies, USA, August 2007-January 2008, US$45,000.

3.G2 A grant in the form a Vice Chancellor's scholarship won from a university wide competition in Southampton. Sahu. Forecasting Air pollution Levels, University of Southampton, October 2012 – September 2015.

3.G3 Sahu S. A rigorous statistical framework for estimating the long-term health effects of air pollution. EPSRC, 2013-16, £356,643.

4. Details of the impact

Some 107 million Americans live in areas that violate health standards for ozone, the USEPA reports. Like the weather, air quality can change from day to day, hour to hour, and national and local media provide daily air quality reports as part of weather forecasts. It is the responsibility of government agencies, principally the USEPA, to continuously monitor and provide air quality information to the public and media. According to the 2009 USEPA report on Air Quality Index (AQI): '*In large US cities (more than 350,000 people), state and local agencies are required to report the AQI to the public daily. When the AQI is above 100, agencies must also report which groups, such as children or people with asthma or heart disease, may be sensitive to that pollutant and would alert groups about how to protect their health. These forecasts help local residents protect their health by alerting them to plan their strenuous outdoor activities for a time when air quality is better*.' **[5.C1]**

From 2010, the USEPA has used Southampton's forecasting methods and its associated software to continuously increase the quality of its air pollution data in its short-term forecasts for the whole of the USA. More accurate and reliable forecasts have had an immediate positive impact by enabling those at risk to plan their outdoor activities to minimise any negative effect on their health. The current USEPA training document aimed at doctors 'Ozone and your Patients Health' **[5.C2]** states: '*Healthcare providers should recommend that patients reduce their ozone exposure on days when air quality is bad, especially people with asthma, who are more susceptible to the effects of exposure*.' It gives detailed advice on how those at risk should modify their behaviour when air quality (and ozone in particular) is predicted to be unhealthy and provides evidence that such advice has a positive impact on health. Commenting on the impact of Southampton's research, the USEPA said: '*The forecast methods of Sahu et al (2010) were adapted to forecast spatial patterns of current 8 hour average ozone concentrations in real-time. This fusion model combines both real-time ozone monitoring data with numerical model output to achieve precise and accurate forecasts*' **[5.1]**.

The general public, including schools and local authorities, are among the direct beneficiaries. According to the Center for Disease Control and Prevention, the existence of accurate and reliable air pollution forecasts bring significant health benefits to 7 million children and 18.7 million adult Americans who suffer from asthma and millions of others who use healthcare facilities to alleviate their respiratory illnesses **[5.C3]**. The WHO advises that limiting exposure to high levels of air pollution impacts positively on long-term health **[5.C4]**. A reduction in GP and hospital visits also delivers substantial economic impact.

The improved forecasts benefit the US agricultural economy. Prolonged exposure to high levels of ozone is harmful to crops. It damages materials such as rubber, paint and textiles. Further, it reduces growth and survivability of tree seedlings, and increases susceptibility to diseases, pests and harsh weather conditions. In the USA, ground-level ozone is responsible for an estimated $500 million in reduced crop production each year. This emphasises the need for a long-term dimension to the modelling of air pollution. Southampton's models have provided accurate estimates of trends in air pollution for the eastern US on which local authorities and regulators, including the USEPA, can base long-term emission control strategies and assess compliance to the regulatory standards of areas that fall outside the sparse ozone-monitoring network in the US. If national standards are not met, local authorities can introduce more stringent emission measures.

Southampton's research has fed into the policy debate in the US over regulations to limit air pollution. It contributed directly to the formation of USEPA's influential 'NOx Budget Trading Program: 2005 Program Compliance and Environmental Results' **[5.2]**, which set out new measures that it said, by 2015, would secure *'$85 to $100bn in annual health benefits, annually preventing 17,000 premature deaths, millions of lost work and school days and ten of thousands of non-fatal heart attacks and hospital admissions.'* The 2005 report – combined with the use of Sahu's statistical spatio-temporal modelling to provide air quality information at unmonitored sites – formed the foundations of the current environmental regulatory policies in the USA, notably the 2011 Integrated Review Plan for the Ozone National Ambient Air Quality Standards **[5.3]**. These reports help develop programs aimed at reducing pollution by reformulating fuels and consumer/commercial products that contain harmful chemicals, and voluntary programs that encourage communities to adopt practices, such as carpooling, to reduce harmful emissions.

The impact of Southampton's work in the USA is being felt further afield. It prompted the EPSRC to provide a grant of £365,643 in 2013 for Sahu to produce a new statistical framework for estimating the long-term health effects of air pollution in the UK **[3.G3]**. The Australia's national science agency, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), has used the spTimer software developed by Sahu **[5.4]** to model large scale environmental data **[5.5]**.

5. Sources to corroborate the impact

Contextual References:

5.C1 'Air Quality: A Guide to Air Quality and Your Health' US Environment Protection Agency Report (2009) EPA-456/F-09-002 [This demonstrates the importance US EPA attaches to predicting AQI]

5.C2 'Ozone and your Patients Health: Patient Exposure and the Air Quality Index': http://www.epa.gov/o3healthtraining/aqi.html [This demonstrates that clinicians actively promote the use of AQI forecasts and this has a significant positive impact on health]

5.C3 'Asthma: Centers for Disease Control and Prevention' http://www.cdc.gov/asthma/default.htm [This demonstrates the scale of the problem and the impact of reliable forecasts of air quality]

5.C4 'WHO Report: Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide (2003) [This details the long-term health impacts of exposure to air pollution]

Sources to corroborate impact:

5.1 Senior Statistician, National Exposure Research Laboratory, United States Environmental Protection Agency (USEPA).

5.2 'NOx Budget Trading Program 2005 Program Compliance and Environmental Results', US Environment Protection Agency Report (2006) EPA430-R-06-013 [**Senior Statistician, USEPA in 5.1 will confirm the input of the work of Sahu into this document]**

5.3 'Integrated Review Plan for the Ozone National Ambient Air Quality Standards', US Environment Protection Agency Report (2011) EPA 452/R-11-006 [**This drew directly on the work of 5.2 above]**

5.4 spTimer: available through Comprehensive R Archive Network http://cran.r-project.org/web/packages/spTimer/index.htm

5.5 http://www.eric-lehmann.com/Presentations/MODSIM_pres_ChiuLehmann.pdf [**For example this work on water resources done in conjunction with the Australian Bureau of Meteorology makes direct use of the spTimer software developed by Sahu]**