

Impact case study (REF3b)

Institution: Imperial College London
Unit of Assessment: 10 Mathematical Sciences
Title of case study: C4 - BUGS (Bayesian inference using Gibbs sampling)
<p>1. Summary of the impact (indicative maximum 100 words)</p> <p>The WinBUGS software (and now OpenBUGS software), developed initially at Cambridge from 1989-1996 and then further at Imperial from 1996–2007, has made practical MCMC Bayesian methods readily available to applied statisticians and data analysts. The software has been instrumental in facilitating routine Bayesian analysis of a vast range of complex statistical problems covering a wide spectrum of application areas, and over 20 years after its inception, it remains the leading software tool for applied Bayesian analysis among both academic and non-academic communities internationally. WinBUGS had over 30,000 registered users as of 2009 (the software is now open-source and users are no longer required to register) and a Google search on the term ‘WinBUGS’ returns over 205,000 hits (over 42,000 of which are since 2008) with applications as diverse as astrostatistics, solar radiation modelling, fish stock assessments, credit risk assessment, production of disease maps and atlases, drug development and healthcare provider profiling.</p>
<p>2. Underpinning research (indicative maximum 500 words)</p> <p>Bayesian statistical approaches have several advantages over conventional statistical inference methods, particularly in situations with sparse data, complex hierarchical structure, missing information and multiple comparisons, and can result in substantial gains in efficiency through the formal inclusion of all relevant prior information. Unlike conventional statistical analyses, Bayesian methods also provide direct probability statements about quantities of interest, which enables results of complex statistical modelling to be more easily communicated to policy makers and end users. They also offer a method for formally combining prior information with current data to allow learning from evidence as it accumulates. However, application of Bayesian methods to real-world problems was delayed by several decades due to computational difficulties, until the development of Markov Chain Monte Carlo (MCMC) computational methods in the early 1990s. Even then, applied scientists were still constrained by the need for purpose-written computer code to implement the MCMC algorithms for each particular problem. This changed with the WinBUGS software, developed initially in Cambridge from 1989-1996 and then greatly expanded at Imperial from 1996 onwards, which aimed to make practical MCMC methods available to applied statisticians.</p> <p>In 1996, Nicky Best, Andrew Thomas and the project moved from Cambridge to Imperial, and work began under Best’s direction on expanding the software’s capabilities [1]. In particular, Jon Wakefield and Dave Lunn joined the project at this stage to work on implementing non-linear models, and development of WinBUGS gained momentum. In subsequent years, a number of other challenging model types were tackled and targeted to application areas with many extensions to the basic package to ensure wide dissemination [e.g. 1-6, G1-G4]. These include (i) GeoBUGS that fits spatial models and produces a range of maps as output [6, G2-G3], (ii) PKBUGS that fits pharmacokinetic/dynamic models [2, G1], (iii) JumpBUGS that implements variable-dimension models fitted using reversible jump MCMC [3, G4], (iv) WBDiff which allows the numerical solution of arbitrary systems of ordinary differential equations (ODEs) within the fitted models and (v) WBDev which enables users to implement their own specialized functions and (univariate) distributions. The development has been underpinned by the theoretical work of Best and her group, who have actively implemented in WinBUGS new analysis techniques which are at the forefront of biostatistical research. This work included research on:</p> <ul style="list-style-type: none"> • disease mapping and spatial regression, implemented in GeoBUGS [6] – Bayesian spatial and spatio-temporal hierarchical models are now widely used for smoothing small area disease rates based on sparse data, in order to identify disease clusters or general (spatial and/or temporal) trends in disease risk related to possible variations in risk factors or provision/access/uptake of health services, and for spatial prediction of health outcomes [5]. • population pharmacokinetic/dynamic (PK/PD) models, implemented in PKBUGS – PK/PD models estimate the relationship between a drug dosing regimen, the body’s exposure to the drug as measured by the nonlinear concentration time curve, and the drug’s efficacy. Such analyses are often based on combining a limited number of measurements from several individuals, and are naturally estimated using Bayesian non-linear hierarchical models to characterize inter and intra-individual variation, and to enable the inclusion of prior information

Impact case study (REF3b)

based on experience with similar compounds, and for predicting the effects of different schedules, doses and infusion times [2].

- genetic association studies, implemented in JumpBUGS – such studies involve selecting which combination of genotypes out of a typically very large set of candidates best predict a given phenotype. Standard hypothesis tests and regression methods have high error rates and low power in such settings, and approaches based on Bayesian model averaging (implemented using reversible jump MCMC) are a popular alternative that can overcome problems of multiple hypothesis testing and estimate the probabilities of association averaged over a number of different model structures [3, 4].

Since 2005, development of the BUGS software has focussed on the [OpenBUGS project](#), which is an open-source version of the core BUGS code with a variety of interfaces. It can run under Windows with a very similar graphical interface to WinBUGS, run on Linux with a plain-text interface, or embedded in R as BRugs. The OpenBUGS project is supported by a formal Collaboration Agreement between Imperial, MRC Biostatistics Unit and Dr Andrew Thomas.

Key contributors:

- Nicky Best, Professor of Statistics and Epidemiology, Imperial College London (1996-present).
- Jon Wakefield, Reader in Statistics, Imperial (1990-1999), now Prof at U. Washington.
- David Lunn, Research Fellow, Imperial (1996-2007), now at MRC Biostatistics Unit, Cambridge.
- Andrew Thomas, Senior Computing Officer, Department of Epidemiology and Biostatistics, Imperial (1996-2004), now at MRC Biostatistics Unit, Cambridge.

3. References to the research (indicative maximum of six references)

- [1] *Lunn, D.J., Thomas, A., Best, N. and Spiegelhalter, D., 'WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility', *Statistics and Computing*, 10, 325-337 (2000). [DOI](#).
- [2] *Lunn, D.J., Best, N., Thomas, A, Wakefield, J. and Spiegelhalter, D., 'Bayesian Analysis of Population PK/PD Models: General Concepts and Software', *Journal of pharmacokinetics and pharmacodynamics*, 29, 271-307 (2002). [DOI](#).
- [3] Lunn, D.J., Best, N. and Whittaker, J., 'Generic reversible jump MCMC using graphical models', *Statistics and Computing*, 19, 395-408 (2009). [DOI](#).
- [4] Lunn, D.J., Whittaker, J. C. and Best, N., 'A Bayesian toolkit for genetic association studies', *Genetic Epidemiology*, 30, 231-247 (2006). [DOI](#).
- [5] *Best N, Richardson S and Thomson A., 'A comparison of Bayesian spatial models for disease mapping', *Stat Methods Med Res*, 14(1), 35-59 (2005). [DOI](#).
- [6] Thomas, A., Best, N., Arnold, R.A., and Spiegelhalter, D.J., "GeoBUGS User Manual, Demonstration Version 1.2" Imperial College and MRC Biostatistics Unit, 2004, available from <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/geobugs12manual.pdf>, and also [here](#).

Selected Grants:

- [G1] EPSRC, [GR/L10437/01](#), 'Bayesian Population Pharmacokinetic & Pharmacodynamic modelling: Implementation and model selection', PI: J Wakefield, co-Is: N Best, D Spiegelhalter, Project partners: GlaxoSmithkline, Pfizer Global R&D, 01/10/96-31/03/99, £123,811
- [G2] ESRC, [H519255036](#), 'Statistical Analysis of Large Geographical Health and Environmental Databases: Methodology Software & Application', PI: N Best, 01/02/98-31/01/00, £155,550
- [G3] MRC, 'Modelling complexity in biomedical research', PI: N Best, 01/04/99-31/03/04, £473,917
- [G4] MRC, 'Computational Tools for Bayesian Bioinformatics', PI: Lunn, 6/10/03-30/9/06, £135,825

4. Details of the impact (indicative maximum 750 words)

WinBUGS is an established and stable, stand-alone version of the BUGS software, which remains available [A] but is no longer being further developed. WinBUGS is still used extensively (searching for 'WinBUGS' on Google returns over 205,000 hits, with 42,400 since 2008 and 11,300 in 2013) and has been described as "*the most widely accepted Bayesian modelling package*" [B]. Since 2005, development of the BUGS project has focussed on the **OpenBUGS** project [C], the open-source version of the BUGS code. OpenBUGS was first released in 2005 and the latest versions of OpenBUGS (from v3.0.7 onwards) have been designed to be at least as efficient and reliable as WinBUGS over a wide range of test applications, but with greater flexibility and extensibility [A].

The impact of the BUGS software is summed up by Prof Brad Carlin, Head of Biostatistics

Impact case study (REF3b)

University of Minnesota: "MCMC freed Bayes from the shackles of conjugate priors and the curse of dimensionality; BUGS then brought MCMC-Bayes to the masses, yielding an astonishing explosion in the number, quality, and complexity of Bayesian inference over a vast array of application areas, from finance to medicine to data mining" [D].

Books and Training:

A demonstration of the popularity and wide applicability of the BUGS, WinBUGS and OpenBUGS software has been the wide number of books published on them since their launch. Since 2008 there have been over 10 dedicated books published about WinBUGS and OpenBUGS, including one by Nicky Best and colleagues. Some examples are:

- 1) [Bayesian Modeling Using WinBUGS](#) (2009), Wiley, I Ntzoufras
- 2) [Introduction to WinBUGS for Ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses](#) (2010), Academic Press, M Kery ([text removed for publication] copies sold, 21/11/13, [E]).
- 3) [Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective](#) (2011), Academic Press, Mark Kery ([text removed for publication], 21/11/13, [E]).
- 4) [Bayesian Analysis Made Simple: An Excel GUI for WinBUGS](#) (2011), Chapman & Hall/CRC Biostatistics Series, Phil Woodward ([text removed for publication] as at 21/11/13, [F]).
- 5) [Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support](#) (2011), Springer Texts in Statistics, B Vidakovic
- 6) [The BUGS Book: A Practical Introduction to Bayesian Analysis](#) D Lunn, C Jackson, N Best, A Thomas, and D Spiegelhalter. (2012), Chapman and Hall. ([text removed for publication] copies sold, 21/11/13, [F]).
- 7) [Applied Bayesian Statistics: With R and OpenBUGS Examples](#) (2013), Springer, MK Cowles
- 8) [R Tutorial with Bayesian Statistics Using OpenBUGS](#), Amazon Media EU (2012), Chi Yau

Sales for four of the books above have been made available to us [E, F] totalling [text removed for publication] copies sold worldwide. The BUGS Book, 6), co-authored by Best, has sold [text removed for publication] copies in its first year since publication, and in November 2013 was ranked #42 on the Amazon "best sellers in Mathematical Probability and Statistics" list (this was the top ranked Bayesian text book) and ranked #28,188 overall (out of 39,995,344) in their books bestsellers. The descriptions for books 2) and 3) state "*Bayesian statistics has exploded into biology and its sub-disciplines [...]. WinBUGS and its open-source sister OpenBUGS is currently the only flexible and general-purpose program available with which the average ecologist can conduct standard and non-standard Bayesian statistics*" [G].

The BUGS software is also used widely for the teaching of Bayesian modelling ideas to students and researchers the world over, and several texts (as demonstrated by the list above) use WinBUGS and OpenBUGS extensively for illustrating the Bayesian approach across both distinct and general application areas. For example, 'The BUGS Book' (book 6) has been adopted as material for Bayesian courses at 19 universities across the world in its first year since publication, with a further 34 universities reviewing the book for courses beginning in 2014. Countries include the UK, Ireland, USA, Canada, Germany, Norway, Finland, the Netherlands and Singapore [F].

Selected Applications of WinBUGS/OpenBUGS:

To provide a flavour of the impact that the BUGS software has had on the practice of Bayesian statistics outside of the academic arena, three exemplars are provided below:

- **Pharmaceutical Industry:** For Pfizer Neusentis WinBUGS has been the software of choice when undertaking Bayesian analyses and has been used in numerous Phase 2 and 3 studies to analyse data, adopting informative prior distributions for the placebo and dose response, and sometimes the standard of care response, saving time, money and, more importantly, unnecessary patient recruitment. The most notable example is one in which Pfizer reduced the placebo arm of the trial by 100 patients, from 300 to 200. As a result the duration of the trial was reduced by approximately 12 months and saved about \$7.5M. In a further exemplar, the results of a BUGS-based meta-analysis of dose response integrating seven phase 2 and 3 studies proved central in a recent discussion supporting dose selection that resulted in approval of a new compound for the treatment of rheumaty arthritis [H].

More generally, Pfizer have stated that "*Bayesian methods have contributed a great deal to the*

efforts being made by Pfizer to improve the efficiency of drug discovery and development. It is only due to the invention of MCMC methods, and their practical implementation in the BUGS software, that we have been able to apply Bayesian methods as widely as we now do. By making MCMC methods for Bayesian analysis both free and relatively easy to program, BUGS is a major factor in overcoming the inertia that exists in the adoption of new methodologies” [H].

- WinBUGS is also the only software package discussed by name in the report by the FDA on Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials (2010) U.S. Department of Health and Human Services, Food and Drug Administration [I].
- Informing national disease control programmes in developing countries: The geostatistical model functionality in add on package GeoBUGS has been used to produce spatial predictive infectious disease risk maps to aid implementation of national disease control programmes in developing countries across Africa and the Asia-Pacific region. These maps, modelled by Prof Clements (U. Queensland), have informed the allocation of resources for various diseases including schistosomiasis, soil-transmitted helminth infections, malaria and rift valley fever. Examples include generating maps for the planning of mass drug administration campaigns to control schistosomiasis in Africa, and maps of malaria risk at the baseline stage of a malaria elimination programme in Vanuatu, forming the basis of a decision to limit indoor residual spraying of insecticide to within 2km of the coastline of Tanna Island. Work on Rift Valley Fever (RVF) was done in collaboration with Prof Best [J], and created maps to support the planning of the siting of sentinel surveillance sites for RVF activity in northern Senegal. Clements states that WinBUGs “overcomes a number of limitations associated with traditional geostatistics allowing for robust spatial predictions that incorporate information from a range of sources” [K].
- Fisheries stock assessments: Fisheries stock assessments are conducted to evaluate the consequences of different management actions. In 1999, a seminal paper by Meyer and Millar [L] recognised the potential of WinBUGS for Bayesian fish stock assessments: “we report on significant progress made in facilitating the routine implementation that may have a revolutionary effect on Bayesian stock assessment in everyday practice. This is achieved through BUGS, a recently developed software package” [p. 1078]. They conclude their article with the prediction that “the routine implementation of Bayesian inference that is now possible will ‘almost surely’ have an impact on fisheries stock assessment” [p. 1084]. Fourteen years later, a Google search on the terms “winbugs + fish + stock + assessment” yields over 1,700 hits since 2008. These include stock assessments of sword fish for the Western and Central Pacific Fisheries Commission [M], Chinook salmon for the Alaska Department of Fish and Game [N] and Bottomfish for the NOAA Pacific Islands Fisheries Science Center [O].

5. Sources to corroborate the impact (indicative maximum of 10 references)

- [A] The BUGS Project, <http://www.mrc-bsu.cam.ac.uk/bugs/> (archived [here](#) on 19/11/13)
- [B] Reuters article 27/4/11, <http://www.reuters.com/article/2011/04/27/idUS152367+27-Apr-2011+BW20110427> (archived at <https://www.imperial.ac.uk/ref/webarchive/t8f> on 19/11/13)
- [C] OpenBUGS webpage, <http://www.openbugs.info/w/> (archived [here](#) on 19/11/13)
- [D] Review from Prof Carlin, <http://statistics.crcpress.com/reviews/the-bugs-book/> (archived [here](#))
- [E] Elsevier, Customer Service [statement received, available from Imperial on request]
- [F] Senior Acquisitions Editor, Statistics, CRC [statement available from Imperial on request]
- [G] http://store.elsevier.com/product.jsp?isbn=9780123870209&locale=en_UK(archived [here](#)).
- [H] Email from VP Head of Pharma Therapeutics Statistics, Pfizer Neusentis, Nov 2013 (available from Imperial on request)
- [I] <http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071072.htm> (archived [here](#) on 25/11/2013).
- [J] Clements ACA, Pfeiffer DU, Martin V, Pittliglio C, Best N, Thiongane Y. “Spatial risk assessment of Rift Valley fever in Senegal”. *Vector-Borne Zoonot* 7(2), (2007), 203-216, [DOI](#).
- [K] Email from Head of Infectious Disease Epidemiology Unit, University of Queensland, November 2013 (available from Imperial on request)
- [L] Meyer R and Millar M. “[BUGS in Bayesian stock assessments](#)”. *Can. J. Fish. Aquat. Sci.* 56: 1078–1086 (1999), also available [here](#).
- [M] Sam McKechnie, Simon Hoyle (2013). Western and Central Pacific Fisheries Commission Report <http://www.wcpfc.int/system/files/SA-IP-08-SWO-CPUE-NZ.pdf>, also available [here](#).
- [N] Alaska Department of Fish and Game, [Fishery Manuscript Series No. 13-02](#), also [here](#).
- [O] NOAA Pacific Islands Fisheries Science Center, [Bottomfish stock assessment](#) 2012, also [here](#).