



PhD opportunities in Statistics at St Andrews, 2023-2024

(Updated 16th December 2022.)

Applications are welcomed for students wishing to undertake a PhD in Statistics at St Andrews. Full funding (fees, plus stipend of approx. £17,668) is available for well-qualified students; we encourage applications as soon as possible to maximize your chances of being funded. UK, EU and other overseas students are all encouraged to apply. New PhD students would typically start in September 2023, but this is flexible. In particular, we are actively looking for students in ecological statistics who may wish to start in January 2023 (see below).

Some general information about the division of statistics is given below, followed by a list of specific topics that are on offer this year. Finally, more information is given about how to apply.

Statistics at St Andrews

Statistics is a lively area of research at St Andrews. The [Division of Statistics](#) is one of three divisions within the [School of Mathematics and Statistics](#), and currently consists of 15 members of academic staff, 15 research staff and 17 PhD students (plus PhD students co-supervised in other Schools). Our research is consistently rated highly in research assessments – for example 96% of our research outputs submitted to the 2021 [UK Research Excellence Framework](#) (REF) were assessed as being either world-leading or internationally-excellent. Our research environment was rated similarly well in REF2021, while our impact (i.e., the real-world effect of our research on wider society) rated particularly highly with 75% being judged to be world leading.

One major research strength is in the area of statistical ecology: contained within the School is the world-leading [Centre for Research into Ecological and Environmental Modelling](#) (CREEM), which is housed in tailor-made facilities at the St Andrews Observatory on the edge of the town. We are a founding member of the [National Centre for Statistical Ecology](#), a multi-institution consortium that ensures regular intellectual exchange between researchers worldwide with similar interests. Several members of CREEM are also part of the university's multi-school [Centre for Biological Diversity](#).

A second more recent and rapidly developing research focus is the [Statistical Medicine and Molecular Biology](#) group, led by a joint professorial appointment (Prof. Andy Lynch) between the Schools of Mathematics and Statistics and the School of Medicine, three recent appointments (Drs. Giorgos Minas, Nicolò Margaritella and Rachel Sippy) in this broad area, and research interests from several other staff members.

Many staff members are also active more generally in the field of [Statistical Methodology](#). Research areas include Bayesian statistical inference and, relatedly, computer-intensive inference, data mining, data smoothing, latent state models and experimental design.

A brief summary of the research interest of each member of staff is given at the bottom of this section; more details can be found by following links to [staff members' web pages on the Statistics Division web site](#).

New PhD students join a high calibre but friendly research environment. Training is provided in the first year in the as part of St Andrews' participation in the [Scottish Mathematical Sciences Training Centre](#) and [Academy for Postgraduate Training in Statistics](#), the latter consisting of four one-week residential courses. Students may get the opportunity to become involved in externally-funded research as part of [CREEM's consultancy group](#); they may also be able to assist on statistics training workshops delivered to professional scientists both in the UK and abroad. Some PhDs are supervised jointly with scientists from other institutions, and there may be opportunities for study at those places. PhD studies are expected to last approximately 3.5 years.

St Andrews is a small, vibrant university town. It is situated on the east coast of Scotland and framed by countryside, beaches and cliffs. The town has a rich cultural heritage, having once been at the centre of Scotland's political and religious life. Today it is known around the world as the Home of Golf and a bustling student town with a distinctively cosmopolitan feel, where students and university staff account for more than 30% of the local population. The university is the oldest in Scotland and third oldest in the English-speaking world. It is the top-rated university in Scotland for teaching quality and student satisfaction, and among the top rated in the UK for overall research; it regularly comes in the top few places in UK league tables compiled, for example, by broadsheet newspapers (e.g., 1st place 2023 The Guardian; 1st place 2022 Times and Sunday Times) and specialist bodies (1st place 2022 for student experience in the National Student Survey; 4th place 2023 Complete University Guide – 1st for student satisfaction). Its international reputation for delivering high quality teaching and research and student satisfaction make it one of the most sought-after destinations for prospective students from the UK, Europe and overseas.

More general information about postgraduate student life at St Andrews is given at the [University postgraduate study web page](#) and in the [postgraduate prospectus](#). School-specific information about applying is given at the [School's postgraduate research page](#).

Brief summary of academic staff interests in the Division of Statistics

- David Borchers – spatial capture-recapture, camera trap surveys, terrestrial acoustic surveys
- Alison Johnston – monitoring biodiversity, citizen science data, species distribution models, data integration
- Andy Lynch – design or analysis of molecular biology experiments, especially applications of DNA/RNA sequencing to cancer research
- Nicolò Margaritella – Bayesian inference, functional data analysis and large-scale inference with application to neuroscience and other applied fields
- Giorgos Minas – stochastic processes and multivariate statistics in molecular biology and medicine
- Michail Papatomas – Bayesian methods with application to genetics and biostatistics
- Valentin Popov – time series and hidden process models
- Rachel Sippy – statistical modelling of infectious diseases, particularly dengue virus.

- Chris Sutherland – statistical ecology: spatial capture-recapture, spatial occupancy models, multi-species occupancy modelling, optimal survey design
- Ben Swallow - Bayesian statistical inference; stochastic systems of biological processes; continuous zero-inflated data; applications in ecology
- Len Thomas – wildlife (particularly acoustic) surveys, population dynamics modelling
- Hannah Worthington – hidden Markov models for mark-recapture analyses, links between ecological and epidemiological/medical methods

Academic staff not taking PhD students in the coming academic year:

- Rosemary Bailey – design of experiments in agriculture, horticulture, ecology and medicine
- Steve Buckland – biodiversity, sampling methods, computer-intensive methods
- Carl Donovan – data mining; commercial statistics; multivariate statistics
- Monique Mackenzie – random effects models, smoothing methods

Specific projects offered for 2023-24

We are currently looking for candidates for the following projects. In addition, prospective candidates with general interests related to those of staff members (see above) are welcome to contact them to discuss other possible projects.

Projects with a * after them are statistical ecology projects where, in addition to the usual August/September 2023 intake, we are seeking students for a January 2023 intake.

Time-to-event statistical models for ecological surveys

Supervisor: David Borchers

Ecological surveys are the foundation on which evidence-based conservation of the planet's biodiversity and wildlife resources is built. Wildlife surveys that provide the evidence base have traditionally been conducted by humans and the resulting data treated as "snapshots" in time. However, digital survey devices like camera traps, acoustic arrays or aerial video platforms generate streams of data, not snapshots. These data are more appropriately viewed as time-to-event data, with the events being detections of the species of interest. Using the event times has the potential to yield richer inferences about the populations under study than snapshot data can, but not much work has been done in this area. If you are interested in developing and applying new statistical methods to exploit the power of digital survey devices, this may be the PhD for you. Specific applications include camera trap surveys, acoustic surveys of vocalizing animals, drone-borne aerial surveys, walked transect surveys, and many others.

Estimating species density with acoustic data

Supervisors: David Borchers and Alison Johnston

There is a growing variety of acoustic data from terrestrial ecosystems. These data offer huge opportunities for more precise and real-time monitoring of wildlife populations and there is already a foundation of work which uses acoustic data to estimate species density. However, these data also present critical statistical challenges. To fully take advantage of these data, there is a need to develop statistical methods that accommodate the challenges. Potential areas of exploration for this PhD include:

- Developing continuous time models that can take advantage of the precise timing of animal sounds.
- Creating models that explicitly incorporate the uncertainty in acoustic species identification, which is likely correlated within individuals and across time and space.

This PhD will involve developing complex statistical models that will be able to help answer some of the pressing needs for biodiversity monitoring in the current ecological crisis.

Validating species distribution models

Supervisors: David Borchers and Alison Johnston

Methods and data to estimate species distribution models have exploded in recent years, however, there is less information on the statistical validation of these models. To validate these species distributions, typically a random subset of the data or a spatially blocked subset of the data are held out for validation. However, validation based on these data only explore how good the model is at fitting to data with the same idiosyncrasies and observation process. Additionally, typically a very limited set of validation metrics are reported, which may not be suitable for different targets of ecological inference. These issues are particularly problematic for species distribution models calculated with citizen science data, due to the additional challenges of data errors and strong biases. There is a need to investigate this problem further in four key directions:

- 1) Explore how data selection impacts the validation of various species distribution models.
- 2) Assess spatial variation in validation metrics and how this affects model suitability for different purposes.
- 3) Align validation metrics with the ecological purposes.
- 4) Develop validation metrics for models of species abundance.

This PhD will make use of simulation and existing datasets to explore the statistical properties of model validation in species distribution models.

From citizen science data to demographic information, exploring the potential of the global eBird project

Supervisors: Alison Johnston and Hannah Worthington

The eBird project (ebird.org/home) is a global initiative bringing together the power of citizen science with the academic community. The freely available app gives participants of any experience

level the opportunity to add to the data collected on the world's bird species. The growing contribution rate to the database, over 100 million bird sightings are added every year, offers incredible statistical opportunities to offer data-driven progress in science, conservation and education.

The scale and volume of the data available offer many exciting statistical opportunities. This PhD project would look to explore and develop analytical approaches to produce detailed and reliable ecological information. Potential avenues of research may include:

- Considering statistical tools and techniques that fully describe all sources of error and variability from raw data collection, machine learning and data cleaning, through to quantifying and visualising the uncertainty from statistical analyses;
- Integrating eBird data with more structured demographic data to estimate population dynamics and make inference from large-scale patterns to mechanistic processes;
- Exploring changes in the observation process over time and potential drivers behind those changes, for example, the impact of COVID on the sampling distribution, or improvements to identification tools;
- Investigating the impact of survey effort and urban bias on the coverage and availability of data across the whole geographic area of interest;
- Identifying long-term patterns or evolving trends in population behaviour, e.g. breeding, phenology, environmental factors, changes in location preferences.

Suitable applicants should have a strong statistical background. Desirable experience would include: experience with programming (R/python); experience with advanced statistical methods or machine learning tools; experience dealing with large datasets; an interest in ecology or birds.

Methods for and applications of meta-analysis of gene-expression data

Supervisor: Andy Lynch

Where the same disease has been studied in multiple different populations, there is a desire to combine those data sets, but this tends to be done (at some expense), by going back to the raw data. Where multiple labs around the world have considered the same question, nominally similar experiments will have been published and there should be a desire to combine them, but differences in technologies, treatments, timings of measurements etc. can hinder this process. In this project you would look at a particular type of comparison and develop a framework for meta-analysis in that context, establish the costs of working with summary and raw data so as to inform future analysts, and perform meta-analyses where this adds to our understanding of disease.

Prof Lynch's interests mainly lie in prostate cancer – both at the population level (e.g. Sequencing of prostate cancers identifies new cancer genes, routes of progression and drug targets, *Nature Genetics* 50 682–692) and the experimental (e.g. "The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis", *EMBO Journal* 30:2719-2733) and there are

certainly opportunities to consider problems of this nature in that disease, but other diseases could be considered also.

Approaches for the design of prostate proteomic projects and their analysis integrated with other molecular data

Supervisor: Andy Lynch

The CRUK-funded International Cancer Genome Consortium prostate has generated whole-genome sequencing data, DNA-methylation sequencing data, mRNA expression data and small RNA data for a cohort of prostate cancer patients. For a subset of these patients protein samples have been banked, and some data have been generated in two 'pilot' mass-spectrometry experiments. While standard analysis methods could be applied, the design of our experiments invites the development of new methods, especially when one considers the additional molecular and clinical data we have available to us.

This project will look to perform an analysis of those proteomics data, developing methods as appropriate. It may also move into considering which additional data should be generated and designing a further experiment.

Identifying complex spatio-temporal biomarkers of brain diseases

Supervisor: Nicolò Margaritella and Michail Papathomas

Bayesian models today are providing the tools to explore the complexity of brain architecture. Therefore, there is a crucial need for leading researchers with an in-depth comprehension of the current challenges in neuroscience and the quantitative skills to develop cutting-edge solutions.

The aim of this project is the development of a modelling framework for the identification of new, complex spatio-temporal brain patterns which can improve our understanding of the functional activity of the brain, our ability to identify early signs of brain diseases and the prediction of their prognosis. In addition, further timely neuroscientific challenges such as the identification of inter-individual variations in brain responses and the inclusion of multiple covariates (e.g. laboratory and clinical) in the identification of complex biomarkers of brain diseases will be researched during the project.

The methodologies developed in the project will provide neuroscientists with innovative analytical tools that will contribute to neuroscientific research on a wide range of brain conditions, from developmental to neurodegenerative diseases, which affect millions of people in the UK and worldwide.

The student will acquire advanced modelling skills in the research areas of Bayesian nonparametrics and functional data analysis which will be essential to develop the innovative modelling framework. New methods will be tested on well-known publicly available neuroscientific datasets and results presented at international conferences in both statistics and neuroscience. The student will be also

involved in the development of R packages that will allow immediate access to all methods developed in this project to the wider scientific community.

Incorporating gene network dynamics to cell population models

Supervisors: Giorgos Minas (in collaboration with Tommaso Lorenzi, Politecnico di Torino)

This project will attempt to merge two fascinating areas of mathematical modelling, simulation, and statistical inference in cell biology. The first area studies the evolution of cell populations that undergo changes such as cell births and deaths, cell-to-cell communications, and interactions with their environment, using computational and partial differential equation models. The second area uses stochastic and ordinary differential equations to describe the evolution of intracellular gene networks that control almost every cell function, including responses to signals such as light, temperature, and inflammation. Despite strong biological relations between these two areas, the mathematical literature lacks approaches for addressing them. Therefore the opportunity for progress in this direction is promising, and progress will provide more complete models and powerful tools for understanding phenomena such as cell differentiation, immune system response, and disorders like cancer. Recently, we formally derived the basic structure of these models. This project will look into a substantial development of these models in modelling, simulation, and/or parameter estimation, depending on the student's interests. A strong interest in cell biology will be important, but no background on this topic is assumed. Some background in applied mathematics and/or probability or statistics will be beneficial, as well as some experience with programming.

Stochastic simulation, analysis, and inference of non-linear dynamical systems

Supervisor: Giorgos Minas

This project will develop a novel framework for stochastic modelling, simulation, analysis, and statistical inference for dynamical systems. It will focus on non-linear dynamical systems producing oscillations and multi-stabilities. These dynamical behaviours are abundant in many fields and especially in molecular biology, epidemiology, and ecology. The developed methods will apply to gene expression oscillations for biological time-keeping and cell-to-cell communication, multi-stabilities in cell development, epidemic oscillations driven by public awareness, ecological oscillations driven by species competitions, and many other settings.

To build this framework, we will use powerful results from the theory of dynamical systems to decompose large, non-linear dynamical systems. This decomposition will allow us to break down systems into their simpler components. For the non-linear components presenting oscillations or multi-stabilities, we will now be able to study their dynamics in detail and build suitable stochastic models that are equally accurate compared to much slower models. We will combine this modelling approach with stochastic models that are extremely fast and accurate for linear dynamical systems. The composition will provide accurate stochastic models, which are computationally faster in application to a much wider range of problems than ever before.

Stochastic modelling and inference for live-cell gene expression data to unravel the mechanisms of stem cell differentiation

Supervisors: Giorgos Minas, Jochen Kursawe, in collaboration with Dr Cerys Manning (University of Manchester)

This project will develop statistical methodology and computational algorithms to estimate model parameters using live-cell imaging data provided by the lab of our collaborator Dr Cerys Manning at the University of Manchester. Live-cell imaging is a powerful technique for real-time observation of the expression of targeted genes in single cells. These observations are important in understanding cellular processes, such as stem cell differentiation, which strongly depend on dynamic gene expression. Stem cell differentiation is a critical biological process for embryonic development, regeneration, and regenerative therapy approaches. Dr Cerys Manning has previously shown that gene expression oscillations are observed in stem cells of the central nervous system, and these are important for regulating the differentiation process. We now wish to unravel the mechanisms driving these oscillations. We also wish to examine the role of stochasticity in stem cell differentiation and its interplay with oscillations. For this purpose, we will develop stochastic models that will be fitted to the highly variable live-cell imaging data. Bayesian statistical methodology will be employed to estimate model parameters, quantify model uncertainty, perform model comparisons, and derive predictions. The ideal candidate will be interested in Bayesian statistics, dynamical systems, and stem cell differentiation. Background in at least one of the above subjects will be beneficial, but candidates with other backgrounds will be considered.

Strategies for detecting high probability dependence structures

Supervisor: Michail Papathomas

The aim of the project is to investigate and infer the complex dependence structure between a large number of categorical and continuous variables. This can also be in the context of an outcome that is explained by a number of covariates. It has been shown that non-parametric Bayesian modelling such as the Dirichlet Process Mixture Model can inform such an investigation, due to the fact that variable selection within mixture modelling selects covariates that combine to form homogeneous groups of subjects, rather than covariates with a strong marginal signal (Papathomas et al. 2012, Papathomas and Richardson, 2016). A model search algorithm informed by results from a procedure that combines Bayesian mixture modelling and variable selection could potentially detect relatively low probability areas in the space of possible models, thus enabling a quicker determination of high probability areas. This project will extend the results in Papathomas and Richardson (2016), by utilizing results in Jing, Papathomas and Liverani (2021), considering more complex mixture models as well as the effect of including both continuous and categorical variables within the analysis. This research has a wide range of potential applications, but the focus will be on applications in Biostatistics, in particular the detection of gene-gene and gene-environment interactions.

References

Jing, W. and Papathomas M. (2021) Challenges and proposals for Dirichlet process mixture models with Gaussian kernels. In preparation.

Papathomas, M. and Richardson, S. (2016): Exploring dependence between categorical variables benefits and limitations of using variable selection within Bayesian clustering in relation to log-linear modelling with interaction terms. *Journal of Statistical Planning and Inference*. 173, 47-63

Liverani, S., Hastie, D. I., Azizi, L., Papathomas, M. and Richardson, S. (2015) PReMiuM: An R package for Profile Regression Mixture Models using Dirichlet Processes. *Journal of Statistical Software*. 64, Issue 7, pp 1-30.

Papathomas, M., Molitor, J., Hoggart, C., Hastie, D. and Richardson, S. (2012) Exploring data from genetic association studies using Bayesian variable selection and the Dirichlet process: application to searching for gene-gene patterns. *Genetic Epidemiology*. 36, 663-674

Interdisciplinary approaches for improving biodiversity assessments using remote sensors

Supervisors: Chris Sutherland, Alison Johnston, Kasim Terzic (Computer Science), Ognjen Arandjelovic (Computer Science)

Global scale declines in and extinctions of populations across a wide range of taxa, ie, the biodiversity crisis, has resulted in the development of global biodiversity targets that shape local, national and international conservation priorities and strategies. Central to determining whether biodiversity targets are being met are the following:

- 1) the development of scalable biodiversity monitoring programs that take advantage of rapidly advancing remote sensors technology
- 2) the development of computer intensive methods for processing 'big data' to convert audio-visual data streams into usable biodiversity data
- 3) the development of statistical methods for analysing such data to test hypotheses about drivers of spatiotemporal biodiversity patterns

This clearly requires an interdisciplinary approach, and specifically research at the interface of ecology, statistics, and computer science.

This project will focus on camera trap and acoustic data (audio-visual media), perhaps the most commonly used remote sensors in ecology at present. The student will have access to several existing sensor data sets and will also have the opportunity to conduct field experiments. Some potential, and not mutually exclusive, research directions the student can take are: understanding data-requirements and monitoring designs for machine- learning assisted inference from remote sensor data; developing new algorithms for species and individual identification from audio-visual media; statistical inference for quantifying spatiotemporal variation in biodiversity from sensor data; and developing tools for non-technical user-communities deploying remote sensors.

Purposefully broadly defined, the project will be tailored to the student specific skills and interests, and as such, we are seeking a student with a genuine interest in pursuing a career at the interface of ecology, statistics, and computer science.

Bayesian causal inference and trial emulation for ecological observational data

Supervisor: Ben Swallow

Conducting formal causal inference for ecological data is challenging due to the complex observational processes that are usually involved. The propensity score method allows the estimation of causal effects in non-experimental studies, however is dependent on constructing emulated experiments to ensure independence between observation process and treatment allocation. This PhD will develop and apply novel methodology in Bayesian causal inference for ecological observational studies, in areas such as citizen science studies, movement modelling and marine surveys, dependent on the interests of the student.

Approximate Bayesian computation using policy driven summary statistics

Supervisor: Ben Swallow

Approximate Bayesian inference are a subgroup of methods commonly employed when inference in mathematical mechanistic models is the focus, as these often do not have a probabilistic or direct likelihood that can be used for inference. Examples include stochastic compartmental models of epidemics and population dynamic models in ecology. The choice of summary statistics to calibrate these models is challenging particularly in high-dimensional settings. This project will study the choice of summary statistics in these models and particularly focus on cases where policy questions are of interest.

Wildlife risk assessment for offshore wind energy

Supervisors: Len Thomas, Saana Isojunno, Lindesay Scott-Hayward

Note that this is a fully-funded project and has a closing date for applications of 6th February 2023.

Many nations are developing offshore wind farms as part of the drive towards net zero carbon emissions. Installation and operation of turbines can affect marine fauna, and quantifying the potential risk through environmental impact assessments is an essential part of the permitting process. Risk assessment can also help to guide decisions about where and when to undertake installations, and where further research investment may be warranted to reduce uncertainties. In this PhD project, the student will work with their supervisors and as part of a large research team to compare different approaches to risk assessment (qualitative, semi-quantitative and fully quantitative) and undertake case studies on avian and marine mammal examples.

Potential research questions:

- What is the optimal risk assessment framework in different contexts (e.g., species, region)?
- How do we quantify different risks and what heuristics are robust in data-poor contexts?

- How do we assess aggregate risk from multiple installations?
- How should uncertainty be quantified and propagated in risk assessments? How should research efforts be targeted to reduce uncertainty?

The PhD student will drive the emphasis and structure of their thesis, but potential topics could include:

- Comparing existing approaches to risk assessment in a data-rich case study (e.g., northern gannet, harbour porpoise) to demonstrate the consequences of different model assumptions and identify areas of methodological improvement.
- Developing and testing risk assessment frameworks for contexts with little or no empirical data, e.g., no baseline monitoring. Investigating applicability of transferring risk assessments between different wind technologies, e.g., fixed versus floating wind farms.
- Undertaking sensitivity analysis to identify parameters leading to the greatest uncertainty in assessment endpoints.
- Developing and testing meta-analytical methods to support risk assessments, e.g., a method to estimate a dose-response function, with uncertainty, from multiple studies.
- Using baseline assessments to inform/evidence post-construction mitigation/compensation.
- Software development - further developing bespoke impact assessment software developed within CREEM, or developing new software for presenting assessment outputs.

For a review of risk assessment in the context of wind energy development, see May et al. (2018).

For a more general review of approaches see Linkov et al. (2009), and for a review of a quantitative approach we have pursued in a different context, see Pirotta et al. (2018).

This project is fully funded for 3.5 years via a scholarship from the University of St Andrews covering tuition fees, and from the WOW project covering living expenses at the UKRI stipend rate (level for 2023-24 set at time of writing, but 2022-23 rate was £17,668; rate increases annually). UK and overseas applicants are equally welcome and eligible.

For more information about this project, including desired qualifications and experience, see www.findaphd.com/phds/projectdetails.aspx?PJID=153230

Literature cited:

Linkov, I., D. Loney, S. Cormier, F.K. Satterstrom and T. Bridges. (2009) Weight-of-evidence evaluation in environmental assessment: review of qualitative and quantitative approaches. *Science of the Total Environment* 407: 5199-5205. <https://doi.org/10.1016/j.scitotenv.2009.05.004>

May, R., E.A. Masden, F. Bennet and M. Perron. 2019. Considerations for upscaling individual effects of wind energy development towards population-level impacts on wildlife. *Journal of Environmental Management* 230: 84-93. <https://doi.org/10.1016/j.jenvman.2018.09.062>

Pirotta, E., C. Booth, D. Costa, E. Fleishman, S. Kraus, D. Lusseau, D. Moretti, L.F. New, R. Schick, L. Schwarz, S. Simmons, L. Thomas, P. Tyack, M. Weise, R. Wells and J. Harwood. 2018. Understanding the population consequences of disturbance. *Ecology and Evolution* 8: 9934-9946. <https://doi.org/10.1002/ece3.4458>

Application procedure

Although there is no fixed deadline (unless noted otherwise for a particular topic), you are strongly encouraged to make your application as early as possible!

Many details of the general requirements and admissions procedure are given at the [University postgraduate research application web page](#).

Applicants should have a good first degree in mathematics, statistics or another discipline (e.g., biology, computer science), with substantial statistical component. A masters' level degree (MSc, etc.) is an advantage, as is any other relevant professional experience. A major criterion for selection is academic excellence: most successful applicants (particularly those who are awarded scholarships) have a good to very good 1st class undergraduate degree and/or a distinction at MSc level. Those who do not have English as a first language, and who have not undertaken an undergraduate or graduate degree taught in English, should provide evidence of English proficiency (minimum IELTS 6.5 or equivalent).

A full list of the criteria we look at when assessing candidates is as follows:

- Academic merit (degree level, classification/grade, and relevance)
- Research potential and engagement (employment history, papers, etc)
- Alignment of research interests with topic applied for
- Relevant personal and professional development

In awarding scholarship funding, we also take account of factors internal to the School, such as whether the proposed supervisor is a new supervisor (which is weighted positively) and whether the research group applied to has relatively few PhD students. Lastly, for applicants and projects rated similarly by other criteria, we may give preference applicants from historically underrepresented groups in Mathematics and Statistics PhD programs at the University of St Andrews.

Potential applicants are encouraged to contact the Postgraduate Officer responsible for PhDs in Statistics, in advance of making a formal application. He is: Len Thomas, email len.thomas@st-andrews.ac.uk, tel. 01334 461801.

To make a formal application, complete the appropriate online form at <https://www.st-andrews.ac.uk/study/pg/apply/research/> (click on "Apply Now" on that page). You also need to provide the following supporting documentation: CV, evidence of qualifications and evidence of English language (if applicable). You are welcome to include a covering letter. You don't need to provide a research proposal unless you are proposing your own project, or sample of academic written work. You will need to ask two referees to provide academic references for you – once you fill in their name on the form, they will be sent emails asking them to upload their references. Please note that we give serious consideration to both the stature of your referees and the remarks that they make about you. More details about the application procedure are given at <https://www.st-andrews.ac.uk/study/pg/apply/research/>

Further School-specific information is on this page

<https://www.st-andrews.ac.uk/mathematics-statistics/prospective/pgr/>

and links from that page.

In addition to the scholarships mentioned on those pages:

- The Centre of Research into Ecological and Environmental Modelling has a small scholarship fund; all students applying for School funding with an intended PhD topic in the field of statistical ecology are automatically considered.
- An up-to-date list of external scholarships is given at <https://www.st-andrews.ac.uk/study/fees-and-funding/postgraduate/scholarships/research-scholarships/>.

We look forward to hearing from you!