Mathematical Sciences for Development

Draft: 15th May 2018

Report following EPSRC ICMS-INI Workshop on Mathematical Sciences and GCRF
ICMS Edinburgh 26 February – 2 March 2018

Contents

1. Next steps and Recommendations

The meeting highlighted many examples of how mathematics is already helping in GCRF related research, and suggested new avenues and collaborations.

Following on from the meeting there are some immediate short term actions and some longer term activities for consideration. It is recommended that there is a meeting between the EPSRC Maths team and the report authors to discuss and agree which activities should be pursued, and by whom.

Actions following on from the meeting

- Generate a directory of GCRF interested parties. This will be generated originally from the meeting/advisory group participants and hosted, in the short term at least, on the ICMS website. A mechanism for signing up to the directory will be implemented
- Case Studies from the meeting have been collated and an overview provided in this document. The individual case studies can be made available via the ICMS website.
- A record (this report and appendices) of the meeting will be prepared and shared with participants and the extended distribution list
- An update email will be sent to the extended distribution list with copies of the report, links to the directory, sign up options, latest news etc.
- Information will also be distributed about research training materials available (e.g. INI library of talks, CDT course materials with open access).

Example longer term activities for consideration.

The aim should be to provide extra support and contact to mathematicians in DAC list countries (support they explicitly identify as important) and to ensure that the mathematics community is better able to respond to GCRF challenges in research.

- UK Universities could be asked to allow institutions based in DAC list countries to have access to, and/or participation in, Access Grid supported graduate training materials (e.g. MAGIC, SMSTC). Existing CDTs could be given extra funding to include a group of students from DAC list countries and to adapt training to promote interaction.
- ICMS-INI could provide the admin-support for a GCRF-related fund follow-on activity, small group work, strategic workshops etc. to support the development of detailed research agendas between partners. This would enhance capability and visibility of mathematics in broader projects involving GCRF funding (projects could include, for example, the Mekong River Dams modelling).
- ICMS-INI would like to enhance participation of delegates from DAC list countries. The contacts from this meeting will be valuable mechanisms to advertise appropriate activities.
- Consideration needs to be given to the sustainable funding for this at a reasonable level in the long term (e.g. charitable sources). Both ICMS and INI already support participants from DAC countries, where possible. However, due to funding constraints, this is limited to a small-scale activity.

2. Context and Consultation

2.1 Consultative Meeting Report

Despite the inclement weather the meeting was very positive and successful. It demonstrated that the UK Mathematical Sciences community has a strong desire to work with DAC List countries in disciplinary and interdisciplinary research, and that this is reciprocated by the representatives from Mathematical Sciences communities from DAC list countries represented at the meeting. Some 60 people participated in the event (most in person but some via electronic means).

Once the advisory committee had agreed the suggested structure and themes for the meeting, suggestions for DAC list country participants were invited. Initial plans had aimed for approximated 20 DAC list country participants across the week. Whilst well over 30 participants were identified and invited, a combination of factors (other commitments, diary clashes, visa issues) meant that, even allowing for the standard late notice visa/travel issues, there were fewer DAC country participants than we had hoped for. Accordingly, about 6-weeks before the meeting, it was agreed to bolster DAC country participation by inviting UK based researcher from those countries, or with extensive expertise in those countries.

DAC List country participants came from India, Brazil (2), Cambodia, Laos, Thailand (2), Zimbabwe, Colombia, Malaysia and Africa Institute for Mathematical Sciences (AIMS). UK based researcher from DAC list countries were able to provide additional insight for Botswana, Zimbabwe, Vietnam, India (3), China, Brazil. The meeting also looked to involve UK based researchers with experience of teaching and/or research in DAC listed countries. In addition to these there were UK based researchers with minimal experience of working in these areas, who were keen to find out more and establish how their work could fit into future GCRF initiatives. Accordingly, the meeting benefitted from a mix of insights, experience and viewpoints. Undoubtedly, increased DAC list country involvement would have been preferable. The 'directory' planned as an output from this meeting, will be a useful tool for achieving this in future meetings.

A full list of participants is provided inn Appendix D (it includes a handful of interested parties who were unable to attend in person or electronically in the end). Each themed day involved both presentations and group discussions.

The efficacy of the Mathematical Sciences and the GCRF agenda

Operational research techniques (hard and soft) play important roles in disaster management planning at all stages (before, during and after). It also helps inform the delivery of healthcare. Statistical methods are central to many issues such as healthcare, epidemiology and agriculture (e.g. using local data to predict positions of earthquake aftershocks, which informs direct aid delivery and determining model parameters for local crops in standard models of global crops). Modelling, often with uncertainty, makes it possible to plan for changing climates and environments (e.g. modelling the useful life of new arsenic filters, local implications of climate change, determining when there is enough information about a disease outbreak to be confident of intervention outcomes). Innovative approaches to data science and network analysis helps the understanding of city growth and the creation of sustainable cities, who to educate about medical interventions, how to create insurance premiums and credit ratings etc. In nearly all of these examples local data, infrastructure, climate and geography changes the nature of the solutions required and the design of the approaches taken. In most cases the central role of the mathematical sciences to create solutions is hidden behind the delivery of interventions.

There are many cutting-edge areas of interdisciplinary mathematical science research which are making positive contributions to the development and sustainability agenda. There is scope for much more to be done. We are assembling a Directory of interests and capability to ensure that partnerships which help unite the global and local problems can be made and maintained and so that the broad expertise and capability of the communities are more easily accessible.

What DAC list country participants want

This was investigated independent of the GCRF agenda – what would participants from the DAC list countries most like from the UK Mathematical Sciences community? Even in countries with a strong research culture, expertise is not uniformly distributed across the country and they would all benefit from partnership, mentoring and capacity building. There were broadly three different groups. For some participants (India, Brazil, AIMS Africa) research ties at the highest level are possible and they wish is to see more of them. For others (e.g. Malaysia) there is a growing ability to undertake research locally and a strong wish to produce results (e.g. climate change at a local level) for themselves and with minimal use of external facilities, such as HPC clusters. Here the wish is for greater technical partnership – developing smarter algorithms that make HPC use less necessary, and developing novel responses (e.g. agriculture, flooding) to the changing environment. For a last group (Botswana, Zimbabwe, Cambodia, Laos) applied mathematical research is in its infancy. Mathematicians are often perceived as irrelevant to real problems and there are cultural barriers to internal collaboration. Here a real need and enthusiasm was expressed for mentorship and enabling activities. Mathematicians often feel isolated (e.g. in Laos there is one university mathematics department in the whole country with a total of 5 PhDs on staff), and their role is seen as producing mathematics teachers. South Korea was suggested as a model where investment in mathematics rapidly changed a similar situation with huge economic benefit to the country, and there are isolated strong groups (e.g. statistics in Botswana).

There is no single model for collaborations between the UK and Mathematical Sciences community in DAC List Countries, we should be an enabler and have different expectations from countries with different levels of mathematical development. Interdisciplinarity is clearly a way around this in some places, but we must be careful not to exclude emerging mathematical scientists.

The Mathematical Sciences - Development Interface

There was a strong level of agreement that the Mathematical Sciences-Development Interface is growing rapidly and has many similarities with the development of Maths & Industry from its early years. Many examples of good practice in knowledge exchange of the mathematical sciences would translate across (e.g. Modelling Camps and Study Groups for Development Problems, internships etc.) Many exciting areas of current research are represented e.g. data science, sometimes with novel aspects (data sources are mobile phones), and the joining up of statistical/modelling/data science predictions with operational research implementation. Some topics involve developing new classes of models to describe human interactions and interventions and to optimize them under different circumstances (cities, international agreements), ethics and policy (e.g. Mekong valley dams).

Ideas coming out of discussions at the meeting on how to exploit these opportunities included a MathSci-Development coordinating centre overseeing a network of relevant projects with UK-based and DAC list country partners, a distributed International CDT with PhD student training providing new expertise in the DAC list countries but with regular linking meetings and shared supervision), a growing group of MathSci for Development UK post-docs to help deliver results and coordinate interdisciplinary activity, and stronger links with NGOs and institutes (e.g. Agricultural Centres) who have a bank of relevant problems from DAC list countries but do not have the mathematical backgrounds to deliver these. The worthy efforts of individual universities should also be more visible, and Knowledge Exchange, Public Engagement and core mathematical activity within DAC list countries should be strengthened as part of a supporting system.

3. Challenges and Opportunities

During the meeting, the participants were asked to provide details of case studies and precedents illustrating the breadth and impact of the mathematical sciences in this field. Although the emphasis is on projects with UK participants some interesting examples from elsewhere are included. This section looks to link details of some of the case studies with the relevant sustainable development goals (SDGs). The case studies provided had differing levels of detail and maturity, so for some topics more information is required at this stage.

It is planned to make the individual case studies available on the ICMS website, and selected case studies will be developed into two-pagers similar to the IMA's Mathematics Matters series[1].

3.1. Sustaining Life

SDG 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture
SDG 6: Ensure availability and sustainable management of water and sanitation for all
SDG 7: Ensure access to affordable, reliable, sustainable and modern energy for all

It has been estimated that 13% of the world population are undernourished, significant proportions do not have reliable access to safe water supplies, and are still cooking over wood, charcoal or dung fires. Food, Water and Energy, along with Health (see 3.2), receive the most ODA support. However measures to address food, water and energy shortages are complicated by the effects of environmental, demographic and societal change, the need for appropriate governance, and the complex web of relationships that they have with each other. This section provides diverse examples of how the mathematical sciences are being, or could be, harnessed to increase our understanding of food, water and energy challenges, improve the effectiveness of interventions in these fields, and help build reliable, sustainable infrastructure to provide these resources.

3.1.1 Research with Small-Scale Farmers

Most[2] agriculture in low and middle income countries is undertaken by small-scale farmers with limited resources. With growing populations and changing environments there is an urgent need to increase the productivity of such farming, and to undertake research into how this may be achieved. Participatory research projects, working directly with the farmers, are believed to benefit from their local knowledge and experiences and to encourage the adoption of resulting advances by the farmers. They also produce different types of data than controlled experiments designed for research stations, and require different statistical methods for analysing them.

The *Collaborative Crop Research Program*[3] (CCRP) is a 15 year old participatory research programme which focuses on agroecological intensification in Africa and South America. It is led from Cornell University and

[1] https://ima.org.uk/case-studies/mathematics-matters/
[2] Quantify
[3] http://www.ccrp.org/

funded by the McKnight and Gates Foundations. Research methods support is provided by the UK social enterprise *Statistics for Sustainable Development*[4] (Stats4SD). In CCRP "large N" trials, large numbers of farmers grow different combinations of crop varieties and data is collected on yields and environmental and social factors. This very different from the controlled balanced block designs traditionally executed at agricultural research stations, but is potentially much more effective at exploring the full range of real life environments and diverse practices adopted by farmers. The approach recognizes that "farmers are not replicates" and aims to provide them with optimal "options by context". The statistical methods used include well-established multi-level regression and mixed methods, but there is considerable scope for developing new methods better adapted to the structure of the data related to those needed by other citizen science projects and by big data applications (see section 4.2).

Stats4SD also contributes to the *Participatory Integrated Climate Services for Agriculture*[5] (PICSA) programme, led by the Walker Institute at the University of Reading. PICSA works with farmers in sub-Saharan Africa where two thirds of the population depends on small-scale rain-fed agriculture. It provides farmers with information and tools to inform their decisions about what crops to plant and when they should be planted. This includes collaborative analysis of historical climate data and seasonal forecasts and a risk analysis of planting options. Improved seasonal forecasts and better understanding of longer term climate trends at the local scale would result in significantly higher yields and lower risk. To be able to provide these national meteorological services need the mathematical sciences to help provide better data and smarter algorithms for use in low resource environments.

3.1.2 Flood and Water Resource Management

The Mekong Valley dam projects provide an excellent example of the complex human and environmental issues that could be approached using mathematical techniques, see 3.7.1. There is also potential for collaboration in taking analysis of local features into account,

3.1.3 Virtual Testing of an Efficient Arsenic Filter for Water Purification

It has been estimated that around 150 million people in 50 countries have been drinking water with arsenic levels above WHO recommended limits.[6] In Bangladesh alone the figure is believed to be over 35 million, and in one area over 20% of deaths have been attributed to high arsenic levels. In 2000 this was described as the "largest poisoning of a population in history".[7] There is still a clear and urgent need for effective low-cost filters that can be fitted to the wells used by the populations at risk.

A simple arsenic filter in which the water percolates through a multi-layer bed made from readily available laterite soil is being developed by Prof Sirshendu De at the Indian Institute of Technology (IIT) in Kharagpur. Physical trials are being accompanied by virtual tests using a porous adsorption medium model developed by Sourav Mondal, Raka Mondal and Ian Griffiths at the Mathematics Institute in Oxford.[8] The model couples the fluid flow through the medium to the convective, diffusive and adsorptive transport of the arsenic and other potential contaminants. The lifetime of the filter is determined by the rate at which these accumulate in the filter and the model has been used to estimate this much more quickly and for a much wider range of environments than would be possible with physical experiments.[9] The filter is currently being trialled on water sources serving around 5000 people in India.[10]

4 https://stats4sd.org/

5 https://ccafs.cgiar.org/participatory-integrated-climate-services-agriculture-picsa/

6 http://www.who.int/mediacentre/factsheets/fs372/en/

7 A.H. Smith, E.O. Linga & M. Rahman, Contamination of drinking-water by arsenic in Bangladesh: a public health emergency, *Bulletin of the World Health Organization*, **78** (2000),1093-1103.

8 https://ecmiindmath.org/2016/10/19/using-mathematics-to-realise-the-potential-of-a-novel-soil-based-filter/

9 http://meetings.aps.org/link/BAPS.2017.DFD.KP1.104

10 S. Mondal, A. Roy, R. Mukherjee, M. Mondal, S. Karmakar, S. Chatterjee, M. Mukherjee, S. Bhattacharjee & S. De, A socio-economic study along with impact assessment for laterite based technology demonstration for arsenic mitigation, *Science of the Total Environment*, **583** (2017), 142-152, DOI: 10.1016/j.scitotenv.2017.01.042.

This work was funded by EPSRC GCRF Institutional Sponsorship, followed by the Royal Society, and was awarded a Mathematical, Physical and Life Sciences (MPLS) Impact Award by the University of Oxford in February 2018.

3.1.4 The Water-Energy-Food Nexus

The concept of a Water-Energy-Food (WEF) Nexus, stressing the interdependencies between its constituent sectors, came to prominence at a conference in Berlin in 2011 on *The Water, Energy and Food Security Nexus: Solutions for the Green Economy.*[11] The WEF Nexus approach focuses particularly on the synergies and trade-offs between individual sector objectives and places them on the context of environmental, demographic and social change. The Nexus has been recognized by the UN Food and Agriculture Organization,[12] the World Economic Forum,[13] the Belmont Forum,[14] the US National Science Foundation,[15] and the UK Economic and Social Research Council.[16] Despite this attention most publications on the WEF Nexus have focused only on high-level qualitative understanding of inter-sectoral interactions and a central aim of current research is to develop the data and modelling tools needed to operationalise Nexus thinking.[17,18]

Because of its size and complexity, approaches to modelling the Nexus need to be very explicit about the purpose, scale and boundaries of proposed models, and the lens through which the Nexus is being viewed. In particular Carole Dalin (UCL) and co-workers have analysed food trading networks from the perspective of 'virtual water', ie the amount of water that is needed to produce the food. This is particularly significant in regions, such as sub-Saharan Africa, where water resources are critical. On a global scale the degrees of network nodes and the 'strengths' of links between nodes, ie the volumes of trade between them, follow exponential and stretched exponential distributions, respectively.[19] These distributions can be modelled in terms of country GDPs, populations and rainfall-on-agricultural-land measures, and forecasts of the future evolution of these quantities used to predict changes in network connectivity.[20] The results suggest that the network will become more heterogeneous, with fewer countries dominating most of the trade. Models of this type have considerable potential across the Nexus for assessing risks associated with environmental change and population growth.[21]

Other areas considered under this section included crop and soil modelling.

3.2. Promoting Well-Being

SDG 3: Ensure healthy lives and promote well-being for all at all ages

3.2.1 Disease Mapping in Low Resource Countries

If disease are to be managed effectively health services must know where they occurs and how intense they are. Combined with other data such information can be used to reveal possible causes for variations in

---

[11] https://www.water-energy-food.org
[12] FAO, *The Water-Energy-Food Nexus,* 2014, http://www.fao.org/3/a-bl496e.pdf .
[13] World Economic Forum, *Global Risks Report 2015,* World Economic Forum, 2015, http://www3.weforum.org/docs/WEF_Global_Risks_2015_Report15.pdf
[14] https://jpi-urbaneurope.eu/news/the-15-projects-that-will-take-on-the-food-water-energy-nexus/
[15] https://www.nsf.gov/pubs/2018/nsf18545/nsf18545.htm
[16] www.thenexusnetwork.org
[17] J. Liu *et al,* Challenges in operationalizing the water–energy–food nexus, *Hydrological Sciences Journal,* **62** (2017), 1714-1720, DOI: 10.1080/02626667.2017.1353695.
[18] Y. Chang *et al,* Quantifying the water-energy-food nexus: current status and trends, *Energies,* **9** (2016), 65, DOI:10.3390/en9020065
[19] C. Dalin *et al,* Evolution of the global virtual water trade network, *Proc. Natl. Acad. Sci.,* **109** (2012), 5989–5994, doi:10.1073/pnas.1203176109.
[20] C. Dalin *et al,* Modeling past and future structure of the global virtual water trade network, *Geophysical Research Letters,* **39** (2012), L24402, doi:10.1029/2012GL053871.
[21] C. Dalin *et al,* Groundwater depletion embedded in international food trade, *Nature,* **543** (2017), 700-704.

prevalence, predict changes in prevalence patterns, plan intervention strategies, and target resources more effectively. Where resources are not available to establish routine collection of health data from across a country more opportunistic methods must be used. These may include incidence data from clinics, national Demographic and Health Surveys, local long-term demographic surveillance sites, and surveys of 'convenient' locations such as schools. These data sources vary considerably in their spatial and temporal reaches and in the extent to which the data is likely to be biased. Advanced geostatistical and computational techniques are needed to optimize the information that can be extracted from these sources and to convert it into knowledge that can be used by health services.

Emanuele Giorgi and Peter Diggle at the Centre for Health Informatics, Computing, and Statistics (CHICAS)[22] at the University of Lancaster have been developing and applying these geostatistical techniques with a wide range of collaborators from both the UK and Africa. Theoretical advances have include methods for combining data from biased sources with those from an unbiased source, incorporating data from different points in space at different times,[23,24] and incorporating data from locations which are not known exactly[25]. These have been incorporated into a new software package for R.[26] Applications have included malnutrition among children in Ghana,[27] putative links between malaria and stunting across Africa,[28] and the distribution and intensity of *Loa loa.*[29]

3.2.2 Modelling Surveillance and Control Strategies for Avian Influenza

Between 2004 and 2016 the H5N1 strain of avian influenza affected over 60 countries, led to the culling of more than 400 million domestic poultry, and resulted in over 450 human deaths and economic damage valued at around $20 billion. The losses were particularly heavy in the parts of East, South-East and South Asia from where the epizootics originate. In Bangladesh incidences of H5N1 infection were reported from 554 poultry premises between March 2007 and April 2012. The high human population density and intensifying agricultural systems of Bangladesh make it a likely source for future new strains of avian influenza that could cross the poultry-human interface. Detailed studies of the propagation of H5N1 in Bangladesh can not only inform national surveillance and control strategies, but also lead to better understanding of how new highly pathogenic strains emerge and become epizootics.

Edward Hill, Michael Tildesley (both Zeeman Institute, Warwick) and their colleagues have been working with stakeholders in Bangladesh to develop a nested set of models of increasing complexity and fit these to data observed at two different administrative scales (division and district) from successive waves of disease outbreaks.[30] The parameter estimation was carried out using Bayesian methods and goodness-of-fit verified by stochastic simulations using individual-based models. It was found that a better fit could be obtained for the divisional level models than the district level models, suggesting that additional factors need to be considered for the latter. There were also significant differences between the parameters for different waves, most notably in the distance over which infection was passed from one premise to another

---

22 http://chicas.lancaster-university.uk/projects/

23 E. Giorgi, S.S.S. Sesay, D.J. Terlouw & P.J. Diggle, Combining data from multiple spatially referenced prevalence surveys using generalized linear geostatistical models, *Journal of the Royal Statistical Society, Series A,* **178** (2015), 445-464.

24 E. Giorgi, P. J. Diggle, R.W. Snow & M.N. Abdisalan, Geostatistical methods for disease mapping and visualization using data from spatio-temporally referenced prevalence surveys, preprint, arXiv: 1802.06359.

25 C. Fronterrè, E. Giorgi & P.J. Diggle, Geostatistical inference in the presence of geomasking: a composite-likelihood approach, preprint, arXiv: 1711.00437.

26 E. Giorgi, P. J. Diggle, PrevMap: an R package for prevalence mapping, *J. Statistical Software,* **78** (2017), DOI: 10.18637/jss.v078.i08

27 J.M.K. Aheto, B.M. Taylor, P.J. Diggle & T.J. Keegan, Modelling and forecasting spatio-temporal variation in the risk of chronic malnutrition among under-five children in Ghana, *Spatial and Spatio-temporal Epidemiology, 2017, DOI:* 10.1016/j.sste.2017.02.003.

28 B. Amoah, E. Giorgi, D.J. Heyes, S. van Burren & P. J. Diggle, Geostatistical modelling of the association between malaria and child growth in Africa, *International Journal of Health Geographics,* **17** (2018), DOI: 10.1186/s12942-018-0127-y

29 E. Giorgi, D.K. Schlüter & P.J. Diggle, Bivariate geostatistical modelling of the relationship between *Loa loa* prevalence and intensity of infection, *Environmetrics,* 2017, DOI:10.1002/env.2447

30 E.M. Hill *et al,* Modelling H5N1 in Bangladesh across spatial scales: Model complexity and zoonotic transmission risk, *Epidemics* **20** (2017) 37–55, DOI: 10.1016/j.epidem.2017.02.007

and the sizes of the premise passed on the most infection. The results for all the models suggested that delays in reporting infections to authorities were significant, and that larger premises might have been more likely to report infections.

The models developed for H5N1 in Bangladesh were then used to compare a number of different intervention strategies incorporating one or more of active surveillance, vaccination and ring culling.[31] Different epizootic scenarios were considered, informed by the differences inferred for the different waves of H5N1 infection. Resource capacity and control objective uncertainty were also taken into account. The resulting optimal strategies tended to favour culling over vaccination, except when the objective was to minimize the number of culled birds rather than shorten the time span of the outbreak. Proactive surveillance, ie ongoing routine checks, was strongly preferred to reactive surveillance which is only introduced when an outbreak is identified. The models also provided criteria for which premises should be checked in surveillance regimes.

Other areas considered were ***Evaluating HIV/AIDS Strategies in India*** (Maini, Rao *et al*[32,33,34]), ***The 'One Health' Nexus*** (ZED[35]), ***Antimicrobial Resistance*** (Antimicrobial Resistance Centre[36]), ***Neglected Tropical Diseases*** (NTD Modelling Consortium[37]) and ***Public Health Logistics*** (CORMSIS[38,39,40,41]).

3.3. Empowering Inclusion

SDG 1: End poverty in all its forms everywhere
SDG 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
SDG 5: Achieve gender equality and empower all women and girls
SDG 8: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all
SDG 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels

Use of official statistics and other data to identify 'at risk' groups. Modelling and analysis of 'wicked' developmental problems.

3.3.1 Measuring Multidimensional Poverty

Poverty is traditionally measured in monetary terms using personal or household incomes. For example the World Bank defines extreme poverty in absolute terms as living on less than US$ 1.90 per day, while the OECD uses a relative measure based on a percentage of the median household income. More nuanced, and perhaps more significant, criteria are based on access to needs such as food, water, health and education. Such measures are necessarily 'multidimensional'.

---

31 E.M. Hill *et al,*The impact of surveillance and control on highly pathogenic avian influenza outbreaks in poultry in Dhaka division, Bangladesh, *Biorxiv* (2017), DOI: http://dx.doi.org/10.1101/193177
32 http://www.ox.ac.uk/sites/files/oxford/field/field_document/Influencing%20HIV%3AAIDS%20policy.pdf
33 http://impact.ref.ac.uk/casestudies2/refservice.svc/GetCaseStudyPDF/20178
34 http://www.ukcds.org.uk/mathematical-modelling-guiding-hivaids-policy-in-india
35 http://www.zoonotic-diseases.org/
36 http://amr.lshtm.ac.uk/epidemiology-modelling/
37 https://www.ntdmodelling.org/
38 https://www.southampton.ac.uk/cormsis/research/indexlisting.page
39 P. Chaiwuttisak *et al,* Location of low-cost blood collection and distribution centres in Thailand, *Operations Research for Health Care* **9** (2015) 7–15, DOI: 10.1016/j.orhc.2016.02.001
40 H.K. Smith *et al,* Location of a hierarchy of HIV/AIDS test laboratories in an inbound hub network: case study in South Africa, *Journal of the Operational Research Society* **68** (2017) 1068–1081, DOI: 10.1057/s41274-017-0240-5
41 H.K. Smith *et al,* Siting of HIV/AIDS diagnostic equipment in South Africa: a case study in locational analysis, *Intl. Trans. in Op. Res.* **25** (2017) 319–336, DOI: 10.1111/itor.12366

Sabina Alkire from the Oxford Poverty and Human Development Initiative[42] (OPHI) and her colleagues have developed a Multidimensional Poverty Index (MPI) that is now widely used by national governments and international agencies such as the World Bank and United Nations[43,44]. The MPI divides poverty into three dimensions, Health, Living Standards and Education, each of which is further split into a number of indicators. Individuals are regarded as poor if the weighted sum of indicators in which they fall below set levels is itself below a set level. The MPI itself is the product of the *incidence* of poverty, ie the number of poor people, and the *intensity* of their poverty, ie the average of the weighted sum of indicators over all people classified as poor. Apart from being able to discriminate between different aspects of poverty the MPI also behaves well when restricted to subpopulations and is relatively stable to variations in indicators, weights and thresholds[45].

A study of MPI changes in 34 LMICs over the period 1998 – 2012 showed that MPI poverty decreased significantly in most countries, with only one increasing significantly[46]. There was very little correlation between the changes in the incidence of poverty and those measured by the US$ 1.90 criterion, showing that the two measures are complementary. The highest MPI reductions were obtained by reducing both incidence and intensity, but other countries showed significant reductions in intensity without reductions in incidence. There were also significant differences between countries in the contributions to the MPI of the different poverty dimensions. A separate study[47] suggested that economic growth (as measured by GDP changes) contributes much more significantly to income poverty reduction than it does to MPI poverty reduction.

3.3.2 Understanding Disability

Disability, Employment and Poverty[48,49,50]
Demographics of Albinism in Africa[51]

3.3.3 Counting the Missing

In times of conflict, civil unrest or government repression data about human rights violations is open to widespread misunderstanding, misinterpretation and misuse. Data about attacks and murder may be collected by multiple stakeholders with different motivations and perspectives and the stories told by their data sets are typically incomplete and can be in conflict with each other. The problems are particularly serious in low-resource environments where transport and communications infrastructure are not well developed. Since 2002 the Human Rights Data Analysis Group[52] (HRDAG) has been using statistical techniques to compare and combine different data sets from conflict and post-conflict zones and use them to produce best estimates of numbers of human rights violations.

---

42 http://ophi.org.uk/

43 S. Alkire & J. Foster, Counting and multidimensional poverty measurement, *Journal of Public Economics*, **95** (2011), 476–487.

44 S. Alkire *et al*, *Multidimensional Poverty Measurement and Analysis,* Oxford University Press, 2015.

45 S. Alkire & M.E. Santos, Measuring acute poverty in the developing world: robustness and scope of the multidimensional poverty index, *World Development,* **52** (2014), 71–91.

46 S. Alkire, J.M. Roche & A. Vaz, Changes over time in multidimensional poverty: methodology and results for 34 countries, *World Development,* **94** (2017), 232-249.

47 M.E. Santos, C. Dabu & F. Delbianco, Growth and poverty revisited from a multidimensional perspective, Working Paper No. 105. Oxford Poverty and Human Development Initiative, 2016.

48 S. Mitra, A. Posarac & B. Vick, Disability and poverty in developing countries: a multidimensional study, *World Development,* **41** (2013), 1-18.

49 S. Mizunoya & S. Mitra, Is there a disability gap in employment rates in developing countries?, *World Development,* **42** (2013), 28-43.

50 S. Mitra & U. Sambamoorthi, Disability prevalence among adults: estimates for 54 countries and progress towards a global estimate, *Disability and Rehabilitation,* **36** (2014), 940-947.

51 P.M. Lund & R.M. Roberts, Prevalence and population genetics of albinism: surveys in Zimbabwe, Namibia, and Tanzania, in *Albinism in Africa,* (eds: J. Kromberg & P. Manga), pp 81-98, Academic Press, 2018, DOI: http://dx.doi.org/10.1016/B978-0-12-813316-3.00004-0

52 https://hrdag.org/

The *multiple systems estimation* techniques used by HRDAG are based on *capture-recapture* methods used in epidemiology and public health applications[53] and in ecology to estimate numbers of wild animals.[54] These methods use multiple overlapping samples to estimate a total population. In their simplest forms they assume that the population is closed, each sample is unbiased, the samples are independent of each other, and the numbers of individuals 'captured' by multiple samples are known. These conditions are rarely satisfied in the context of human rights violations and so the methods are adapted by using the available data to produce statistical models for how the samples depart from ideal cases. These models are then incorporated into capture-recapture estimates to obtain overall numbers of victims.[55]

Consider, for example, the unbiased samples assumption. In practice the probability of an individual being included in a particular sample is not constant – some individuals are more likely to be captured than others. This is modelled by analysing the inclusion patterns determined by the data to construct a capture probability distribution.[56] Similar techniques can be used to explore and model dependencies between samples, and in both cases any additional information that is available on individuals, such as geographic location, can be used to stratify the samples. There are always choices of models to be made in these circumstances and so model selection (eg Bayesian information criteria) and/or averaging (Bayesian model averaging) tools are useful.[57]

HRDAG has applied Multiple Systems Estimation to counting human rights violations in more than 10 LMICs in Africa, Asia, and Central and South America. There are also potential applications in other areas of development, including health and conservation, and for further theoretical advances (see section 4).

3.3.4 Revealing Corruption

Corruption risk indices and applications[58,59,60,61]

3.4. Protecting our Environment

SDG 13: Take urgent action to combat climate change and its impacts
SDG 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development
SDG 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss

3.4.1 Tipping Points

Many of the processes that underpin the Earth's environmental and ecological systems are highly nonlinear and generate models demonstrating multiple steady states and complex dynamics. Gradual changes in the parameters of these systems can lead to abrupt catastrophic changes in the state of the environment[62]. Timothy Lenton (Exeter) and colleagues have formalised the idea of *tipping points* for such systems and

---

[53] D. Böhning, P.G.M. van der Heijden & J. Bunge, *Capture-Recapture Methods for the Social and Medical Sciences,* CRC Press, 2017.
[54] R.S. McCrea & B.J.T. Morgan, *Analysis of Capture-Recapture Data,* Chapman-Hall, 2014.
[55] K. Lum, M.E. Price & D. Banks, Applications of multiple systems estimation in human rights research, *The American Statistician,* **67** (2013), 191-200, DOI: 10.1080/00031305.2013.821093.
[56] D. Manrique-Vallier, M. E. Price & A. Gohdes, Multiple systems estimation techniques for estimating casualties in armed conflicts, in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict,* T.B. Seybolt, J.D. Aronson & B. Fischhoff (eds.), Oxford University Press, 2013.
[57] K. Lum, M.E. Price, T. Guberek & P. Ball, Measuring elusive populations with Bayesian model averaging for multiple systems estimation: a case study on lethal violations in Casanare, 1998-2007, *Statistics, Politics and Policy,* **1** (2010), 2151-7509.
[58] M. Fazekas, I.T. Tóth & L.P. King, An objective corruption risk index using public procurement data, *European Journal on Criminal Policy and Research,* **22** (2016), 368-397.
[59] https://www.maths.ox.ac.uk/node/25583
[60] https://scscsussex.wordpress.com/2017/05/
[61] http://www.govtransparency.eu/wp-content/uploads/2017/08/Czibik_et_al_TanzanianProcurementData_2017.pdf
[62] M. Scheffer *et al*, Catastrophic shifts in ecosystems. *Nature*, 413 (2001), 591–596.

identified a number of systems (*tipping elements*) that are candidates for tipping points as a result of anthropogenic climate forcing[63]. Among them are several that have a direct and significant impact on LMICs, including the El Niño–Southern Oscillation (ENSO), the West African Monsoon (WAM), the Indian Summer Monsoon (ISM) and the Amazon Rainforest. Interactions between tipping elements may broaden this list[64]. Tipping point ideas for systems forced by oscillations such as the Earth's diurnal and annual cycles have been applied to produce conceptual models for the abrupt onset and failure of summer monsoons[65].

A number of early warning signals for tipping points have been proposed[66]. These include monitoring time series to detect the critical slowing down and increase in perturbation amplitudes that typically occur just before bifurcations. Lenton and his Exeter colleague Mark Williamson have emphasised the roles played by different time scales in these schemes – the natural time scale(s) of the system, the time scale of the forcing, and the time scale of noise-induced perturbations[67]. In particular they studied the behaviour of early warning signals for systems with oscillatory forcing in the cases when the forcing time scale is much less than that of the system, when it is same, and when it is much greater. The methods were applied to satellite observation of arctic sea ice, a proposed tipping element, but no early warning of tipping was detected.

3.4.2 Safe and Just Operating Spaces

The concept of *safe operating spaces* that keep Earth systems within *planetary boundaries*[68,69] has been highly influential in recent ecosystems theory and has been extended to include development and social justice concerns[70]. The emerging picture is that Earth systems need to be steered to lie in *safe and just operating spaces* that lie between ceilings provided by environmental constraints and the social foundations that are needed to achieve the Sustainable Development Goals. To do this it is necessary to understand the complex social-ecological systems that underpin development issues.

John Dearing (Southampton) and colleagues have studied how the safe and just operating spaces ideas might be used concretely within specific regions[71]. For the Ganges delta in Bangladesh[72,73] they first constructed a conceptual model of the complex interactions between climate, hydrological, agricultural and economic variables and then quantified the interactions using both a statistical approach and an approach that focused on modelling the nonlinearities in the system. The models were constructed and validated within a participatory framework that included stakeholder workshops and were then compared with historical crop production data and subjected to parameter sensitivity tests. The safe operating space was defined to be the region where key indicators such as crop production and household income remained with the envelope of variability defined by recent historical data. Test scenarios that assumed different combinations of changes to climatic, hydrological and economic boundary conditions were run to determine whether, when and how operating space boundaries would be crossed. The results suggested that even the 2°C increase in temperature allowed for in the 2015 Paris Agreement would lead to drops in agricultural output and household income, and that these drops would cross into the dangerous region if

---

63 T.M. Lenton *et al,* Tipping elements in the Earth's climate system, *PNAS,* **105** (2008), 1786-1793.

64 Y. Cai, T.M. Lenton & T.S. Lontzek, Risk of multiple interacting tipping points should encourage rapid CO_2 emission reduction, *Nature Climate Change,* **6** (2016), 520-525.

65 S. Bathiany *et al,* Abrupt climate change in an oscillating world, *Scientific Reports,* (2018), 5040.

66 T.M. Lenton, Early warning of climate tipping points, *Nature Climate Change,* **1** (2011), 201-209.

67 M.S. Williamson *et al*, Early warning signals of tipping points in periodically forced systems, *Earth Syst. Dynam.,* **7** (2016), 313-326.

68 J. Rockström *et al,* A safe operating space for humanity, *Nature,* **461** (2009), 472–475.

69 W. Steffen *et al*, Planetary boundaries: Guiding human development on a changing planet, *Science* **347** (2015). DOI: 10.1126/science.1259855

70 K. Raworth, A safe and just space for humanity, *Oxfam Policy and Practice: Climate Change and Resilience,* **8** (2012), 1-26.

71 J.A. Dearing *et al,* Safe and just operating spaces for regional social-ecological system, *Glob. Environ. Chang.*, **28** (2014), 227–238. DOI:10.1016/j.gloenvcha.2014.06.012.

72 M.S. Hossain *et al*, Unravelling the interrelationships between ecosystem services and human wellbeing in the Bangladesh delta, *Int. J. Sustain. Dev. World Ecol.*, **24** (2016), 120-134. DOI: 10.1080/13504509.2016.1182087.

73 M.S. Hossain *et al,* Operationalizing safe operating space for regional social-ecological systems, *Sci. Total Environ.,* (2017), DOI:10.1016/j.scitotenv.2017.01.095.

the temperature rise is accompanied by a sea level rise, increased upstream fresh water withdrawals from the Ganges, or decreases in agricultural subsidies. Even if the temperature rise is gradual the drops would occur suddenly and without obvious warning signs.

3.5. Enabling Resilient Infrastructure

SDG 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation

SDG 11: Make cities and human settlements inclusive, safe, resilient and sustainable

SDG 12: Ensure sustainable consumption and production patterns

3.5.1 Climate Adapted Building Design

Zero Peak Energy Design for India (ZED-I)[74]

3.6. Mobilising the Data Revolution

New sources of data. Data innovation. Transforming data into information, knowledge and services.

3.6.1 Agricultural Micro-Insurance

Agricultural insurance is widely considered to be a promising mechanism for improving food security and reducing poverty across a broad range of LMICs. In sub-Saharan Africa small-scale farmers are at considerable risk from crop failure as a result of rain failures and/or flooding at critical times of the year. A bad year results in both food shortages and financial hardship for the individual farmers and their families and can have significant consequences for food security across whole regions[75]. The cost of assessing losses incurred by farmers means that traditional indemnity-based crop insurance is too expensive. Instead Weather Index Insurance (WII) is being piloted by a number of agencies in countries across Africa. WII pays out automatically to a policy holder whenever an appropriately defined index crosses a critical level. Key issues that need to be resolved to make WII work efficiently on a large scale include the design of indices that accurately reflect the farmers' losses and for which accurate data can be obtained at a sufficiently high resolution.

Emily Black and her colleagues at the University of Reading worked with the insurance company MicroEnsure to investigate the feasibility of using satellite remote sensing to provide the meteorological data needed to construct an index suitable for cotton farming in Zambia[76]. Reading's TAMSAT system provides daily rainfall estimates for the whole of Africa at a resolution of 4km[77]. A well-established land surface model was used to convert this to estimates of soil moisture. The effects of aggregating rainfall over both time and space were explored and optimal levels of aggregation, balancing accuracy and resolution, were derived. The resulting index was used in a pilot project in Zambia providing insurance to 7,000 farmers of whom around 3,700 received payouts. There was a strong correlation between the payouts and the losses experienced by farmers[78].

3.6.2 Precision Health

Use of genetic data in medicine: 'precision/personalized medicine' / 'medicine by context'[79,80,81]

---

74 http://gtr.ukri.org/projects?ref=EP%2FR008612%2F1

75 J. Adegoke *et al,* Improving climate risk transfer and management for Climate-Smart agriculture: a review of existing examples of successful index-based insurance for scaling up, Global Alliance for Climate-Smart Agriculture, 2017. http://www.fao.org/3/a-bu216e.pdf

76 E. Black *et al,* The use of remotely sensed rainfall for managing drought risk: a case study of weather index insurance in Zambia, *Remote Sensing,* **8** (2016), 342. DOI: 10.3390/rs8040342.

77 https://www.tamsat.org.uk/

78 https://connect.innovateuk.org/documents/2906271/3749832/PA13-013_MicroEnsure_Reading_CaseStudy

79 N. Mulder, Development to enable precision medicine in Africa, *Personalized Medicine,* **14** (2017), 467-470.

MalariaGen[82]

3.6.3 Smart Communities

At all scales, from villages[83] to (mega-)cities[84]

3.7. Quantifying Uncertainty

Probabilistic Estimation/Forecasting. Risk. Moral obligation to include error bounds in forecasts.

3.7.1 Uncertainty and Risk for Natural Hazards

In 2012–16 NERC funded a programme of research on Probability, Uncertainty and Risk in the Environment (PURE)[85]. As part of this the Smith Institute coordinated a series of short projects, each of which was developed into a case study, and several of which were directly relevant to development in LMICs[86,87]. One, on the use of satellite data for weather index-based insurance, is described in 3.6.1. The others included the following[88]:

Volcanic Risk Reduction: Anna Hicks and colleagues at the University of East Anglia combined volcanology and probabilistic and spatio-economic modelling to provide risk assessments for a series of eruption scenarios on Tristan da Cunha and evaluate possible mitigation measures[89,90].

Hurricane Risk Modelling: Richard Chandler and Ken Liang from University College London worked with Steve Jewson from Risk Management Solutions (RMS) to incorporate parameter uncertainty into RMS hurricane models and hence improve hurricane loss predictions[91].

Probabilistic Flood Forecasting: Paul Smith and colleagues at Lancaster University worked with partners from Practical Action Consulting to develop a community-based early warning system for flooding in Nepal that combines mechanistic and probabilistic modelling with data assimilation[92,93].

Earthquake Aftershock Forecasting: The University of Ulster Geophysics Group and Concern Worldwide jointly developed a set of tools to forecast earthquake aftershocks that can be used on the ground by humanitarian agencies[94].

Assessing and Visualising Flood Vulnerability: Teams from the UCL Hazard Centre and CAFOD collaborated on methods that combined and visualised population and flood risk data to determine the vulnerability of communities along the Mekong River in Cambodia[95,96].

---

80 L. Nordling, Putting genomes to work in Africa, *Nature*, **544** (2017), 20-22.
81 J. Makani *et al,* Sickle cell disease: tipping the balance of genomic research to catalyse discoveries in Africa, *Lancet,* **389** (2017), 2355-2358.
82 https://www.malariagen.net/
83 https://e4sv.org/
84 http://www.mitec.gov.rw/fileadmin/Documents/Strategy/SMART_AFRICA_Sustainable-Cities_A_Blueprint_for_Africa.pdf.pdf
85 https://nerc.ukri.org/research/funded/programmes/pure/
86 http://www.smithinst.co.uk/case-studies/uncertainty-natural-hazards/
87 https://connect.innovateuk.org/web/pure/article-view/-/blogs/pure-associate-case-studies
88 Some of which could be developed into independent case studies.
89 https://connect.innovateuk.org/documents/2906271/3749832/PA13-045FCO_UEA_CaseStudy
90 A. Hicks *et al,* An interdisciplinary approach to volcanic risk reduction under conditions of uncertainty: a case study of Tristan da Cunha, *Natural Hazards and Earth System Sciences*, **14** (2014), 1871–1887.
91 https://connect.innovateuk.org/documents/2906271/3749832/PA13-040_RMS_UCL_CaseStudy
92 https://connect.innovateuk.org/documents/2906271/3749832/PA13-032_PracticalActionConsulting_Lancaster_CaseStudy
93 P.J. Smith, S. Brown & S. Dugar, Community-based early warning systems for flood risk mitigation in Nepal, *Natural Hazards and Earth System Sciences*, **17** (2017), 423–437.
94 https://connect.innovateuk.org/documents/2906271/3749832/ConcernWorldwide_Ulster
95 https://connect.innovateuk.org/documents/2906271/3749832/PA13-010_CAFOD_UCL_CaseStudy
96 pp 102-106 in: https://www.gfdrr.org/sites/default/files/solving-the-puzzle-contributions.pdf

3.8. Modelling Society

Links to social sciences, economic complexity, etc

3.8.1 Modelling Social Networks for Public Health Interventions

Social networks have been widely recognized as having key roles to play in modelling both the spread of diseases and the optimisation of public health and social interventions. However different networks and different network properties are important in different contexts. Many diseases are passed from person to person by a single contact when they are close to each other, and so may be regarded as diffusing through proximity networks along shortest path routes. In contrast, willingness to participate in a vaccination or contraception programme may depend on receiving positive reinforcement from a number of different members of a friendship or acquaintance network and so depend on different network connectivity properties than simple epidemic diffusion in a proximity network.

Goylette Chami (Wellcome-Cambridge Centre for Global Health Research) and her collaborators have been analysing the social connectivity of community members who provide health advice and/or medicines. They mapped 'close friend' and 'trusted health adviser' networks in 17 villages in Uganda, identifying the nodes that corresponded to village members appointed as Community Medicine Distributors (CMDs). The proportion of each village to which the CMDs were able offer medicines, and the speed at which they did it, was compared statistically with their connectivity properties in the network.[97] It was found that high levels of performance were significantly correlated with measures of 'clustering' in the 'close friend' network, ie the extent to which the nodes connected to the CMD where themselves connected, but not to centrality measures such as the degree of the CMD node or the average degree of its neighbours. However, the likelihood that a village member was offered medicines by the CMD was significantly linked to the friendship degree of the recipient, whether the CMD was a trusted health adviser, and to indicators of the social status of the recipient. [98]

Drawing on their Ugandan village experiments, Chami and colleagues also considered efficient ways of removing nodes to disrupt social networks without having to fully map them in advance. For simple epidemics node removal might be achieved by vaccinating the corresponding member of the community. For health and social interventions the targeted network would be one which propagates negative information about the campaign and node removal correspond to providing health education to the corresponding person. When tested on the Ugandan data, removal algorithms based on iteratively removing high degree neighbours of randomly selected nodes were shown to significantly outperform algorithms that first targeted health workers, village leaders and school teachers.[99]

Neave O'Clery (Mathematics Institute, Oxford) leads a research group focused on studying cities as complex social, spatial and economic systems, with a particular interest in modelling the emergence of economic complexity through knowledge diffusion, and the development of innovative approaches to study the economic and spatial forces underlying agglomeration economies." [100,101,102] Peak Urban.[103]

---

[97] G.F Chami *et al,* Diffusion of treatment in social networks and mass drug administration, *Nature Communications* **8** (2017) 1929, DOI: 10.1038/s41467-017-01499-z

[98] G.F Chami *et al,* Community-directed mass drug administration is undermined by status seeking in friendship networks and inadequate trust in health advice networks, *Social Science & Medicine* **183** (2017) 37 – 47, DOI: 10.1016/j.socscimed.2017.04.009

[99] G.F Chami *et al,* Social network fragmentation and community health, *PNAS* **114** (2017) E7425-E7431, DOI: 10.1073/pnas.1700166114

[100] http://neaveoclery.com

[101] N. O'Clery & E. Lora, City size, distance and formal employment creation, *Development Bank of Latin America (CAF) Working Paper*, 2016

[102] N. O'Clery *et al,* The path to labor formality: urban agglomeration and the emergence of complex industries, *CID Working Paper*, 2016

[103] http://www.peak-urban.org/

3.9. Building Capacity

The case studies outlined above show that the mathematical sciences are already being used widely in development projects. Section 4 gives examples of the ways in which the mathematical sciences need to be advanced to increase their impact on development challenges. To achieve significant impact with these advances mathematical scientists must collaborate with development experts from other disciplines. The impact on LMICs will also be magnified many times if the research is conducted in partnership with stakeholders, including mathematical scientists, from those countries. To establish these partnerships and collaborations several different types of capacity building are needed. Several different potential collaborators were highlighted during discussions.

Collaborative partners could include:

- ***National Capacity Building Programmes***
  - Recognition of importance of Mathematical Sciences in South Korea and Rwanda
- ***African Institute for Mathematical Sciences (AIMS)***
- ***Schemes for Mathematical Sciences Development***
  - ***e.g.*** CDC / LMS / RSS / TWAS / ICTP . . . . . .
- ***Africa Programmes of UK Universities***
  - Cambridge-Africa, Oxford-Africa, Warwick in Africa, ...
- ***GCRF Programme Intitiatives***
  - Potential for capacity building aspects

3.10. Ensuring Ethical Engagement and Evaluation

3.10.1 Human Heredity and Health in Africa

H3Africa[104] as a precedent for research frameworks that recognize the need for ethical engagement[105]

3.10.2 Monitoring and Evaluating Progress Towards SDGs

"A World that Counts"[106]

4 Underpinning Mathematics

Drawing evidence from the case studies outlined above, this section provides a brief overview of some areas of mathematical sciences that are having, or are likely to have, most impact on development challenges in LMICs. It particularly highlights areas where new mathematical concepts and methods are needed or where their application to development challenges is novel. To maximise the effectiveness and impact of work in these areas mathematical scientists need to collaborate closely with experts from other disciplines and engage directly with the motivating development challenges.

The outlines below focus particularly on data, modelling and their interactions. These are necessarily the basis for any mathematical work that contributes to our understanding and ability to control 'real world'

[104] https://h3africa.org/
[105] https://www.nature.com/articles/d41586-018-04589-0
[106] www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf

phenomena. However the scope of these areas is very broad and can incorporate ideas from many other areas of mathematics. There is also considerable scope for work at interfaces between mathematical sciences and development to feed back into novel research in more theoretical areas of the mathematical sciences.

4.1 Data Collection and Curation

The adoption of the Sustainable Development Goals was accompanied by a call for a Data Revolution[107108]: new technologies should be exploited to produce more and higher quality data that can be used to both monitor and assist progress towards the SDGs. Mathematical scientists have major rolls to play in mobilising this data revolution at all stages from data collection through to effective applications of data to monitoring, forecasting and controlling the physical, biological and human systems that underpin the SDGs. Major challenges at the collection and curation end of the 'data journey' that cut across many different areas of application include developing new methods for:

- Assessing the quality of data
- Designing data collection systems
- Bridging data gaps
- Combining data from different sources
- Upscaling and downscaling data
- Quantifying uncertainty in both raw and derived data sets
- Designing accurate and meaningful indices that summarise data sets.

Mathematical scientists are well placed to assess and improve the quality of data sets and their applications and to build on best practice in individual areas of application and transfer it to other areas.

4.2 Statistics

The advent of 'big data' and 'data science' has caused mathematical scientists, and particularly statisticians, to reflect on the roles of their disciplines and the opportunities offered by the data revolution[109,110]. The natures of many new data sources (whether 'big' or not) and their potential applications mean that existing statistical methods need to be extended and new methods developed to ensure that the data are used effectively and rigorously. All the bullet points listed in 4.1 need state-of-the-art statistical thinking applied to them. Statistical inference theory and methods need to fully embrace the differences between the data collection and curation processes they have traditionally dealt with and those that are emerging in the data revolution. In particular this means adapting to new data structures and biased data as well as the sheer size of the data sets that are becoming available.

4.3 Modelling

Modelling is an essential component of the process of turning data into information, knowledge, and tools that can impact on real world problems. They range from the conceptual models that all disciplines use to structure the domains in which they work, via qualitative models that describe networks of interactions between key components of complex systems and simple quantitative models that can be investigated using analytical tools, through to large simulation systems that aim for accurate quantitative representations of complex systems. All these levels of modelling are needed, as are methods for passing between them.

---

[107] http://www.undatarevolution.org/

[108] CSIS, *Harnessing the Data Revolution to Achieve the Sustainable Development Goals: Enabling Frogs to Leap,* Rowman and Littlefield, 2017. https://www.csis.org/analysis/harnessing-data-revolution-achieve-sustainable-development-goals.

[109] P. Diggle, Statistics: a data science for the 21st Century, *Journal of the Royal Statistical Society* **178** (2015), 783-813.

[110] E.M. Scott, The role of statistics in the era of big data: crucial, critical and under-valued, *Statistics and Probability Letters,* in press, available online (2018). DOI: https://doi.org/10.1016/j.spl.2018.02.050.

Mathematical representations that are commonly used in models include functional relationships (linear or nonlinear), dynamical systems (discrete and continuous; deterministic and stochastic), networks and the 'games' of game theory. A universal challenge is balancing the accuracy or usefulness of a model[111] with its analytical tractability. Full understanding of a system and ability to impact on it are likely to draw on a suite of interrelated, synergistic models that provide different perspectives. As an example consider the challenge of incorporating human decision-making into models, especially when the behaviour of the system depends on decisions made by large numbers of participants. Modelling components that might be included in such a project could include the initial construction of a qualitative model identifying the relevant factors and links between them, perhaps via a participatory approach[112], followed by the exploration of an agent based model with a systematic programme of simulations. To understand the phenomena that emerge from the simulations it may then be appropriate to use methods such as homogenisation to obtain reduced models that can be analysed analytically, or to apply 'equation-free' methods to characterise the different types of behaviour that are observed[113].

Effective implementation of a modelling programme such as that described in the previous paragraph needs a multidisciplinary team of investigators that brings together knowledge of the relevant area(s) of development, the ability to engage stakeholders, and modelling, simulation and analysis skills. Models play an essential role at the interfaces between different disciplines and modellers can play leading roles at the centre of such projects[114]. To do this modellers should have a wide knowledge of the mathematical theories and techniques that might be relevant to the project. This approach is very different from one that relies on the use of 'black box' toolkits that may have been developed in a different context and for different purposes.

4.4 Data Assimilation

To have impact on the real world all modelling should be driven by data, whether qualitative or quantitative. The usefulness of a model depends on how much it contributes to our understanding of observations made of the processes or issues under investigation. Modelling should also respond to new observations. This puts the concept of *data assimilation* into a very broad framework that embraces all links between data, statistics and modelling. In particular it incorporates the idea of fitting models to data. Mathematical state-of-the-art paradigms for data assimilation include the Kalman filter and its relations, and methods used for weather and climate forecasting. The introduction of, for example, Bayesian methodology and ensemble forecasting provide rigorous approaches to quantifying uncertainty in predictions. These concepts and methods ideas are being applied in other contexts, including examples relevant to development, but their use needs to be expanded to provide the rigorous assessments of model accuracy that are required. Expansion to other contexts will require the methods to be adapted to a broader range of model types (eg networks, individual based[115]) and data types (eg sparse, non-normal/Poisson, biased).

4.5 Genetics and Genomics

- Non-equilibrium population genetics
- Bioinformatics
- Statistical inference (eg for health applications)

[111] "All models are wrong, but some are useful": G.E.P. Box, Robustness in the strategy of scientific model building, pp 201-236 in R.L. Launer & G.N. Wilkinson, *Robustness in Statistics,* Academic Press, 1979.

[112] A. Voinov *et al,* Modelling with stakeholders – next generation, *Environmental Modelling and Software,* **77** (2016), 196-220.

[113] S.A. Thomas, D.J.B. Lloyd & A.C. Skeldon, Equation-free analysis of agent-based models and systematic parameter determination: A NetLogo Implementation, *Physica A,* **464** (2016), 27-53.

[114] M.E. Kragt, B.J. Robson & C.J.A Macleod, Modellers' roles in structuring integrative research projects, *Environmental Modelling and Software,* **39** (2013), 322-330.

[115] D.J.B. Lloyd, N. Santitissadeekorn & M.B. Short, Exploring data assimilation and forecasting issues for an urban crime model, *European Journal of Applied Mathematics,* **27** (2016), 451-478.

4.6 Uncertainty and Risk

- Extreme events: "Multivariate extremes are not well understood"

4.7 Computational Methods

- Modelling & DA in low resource environments (smart algorithms, data pre-processing, . . .)

5 Vision and Strategy

The meeting collated perspectives from DAC list country participants (see Section 2 – Context and Consultation), highlighted the research in the mathematical sciences already contributing to the GCRF agenda (see Section 3 – Challenges and Opportunities), as well as identifying areas of mathematics which could be appropriate for GCRF research (see Section 4 – Underpinning mathematics. Discussions highlighted that many SDG-related projects with mathematical sciences content are funded via other research councils. However EPSRC and maths institutions need to be involved to ensure that mathematical methods are used appropriately and efficiently and that the necessary new methods are being developed. Additionally, whilst the primary driver is the impact of the mathematical methods there is still considerable scope for developing novelty.

To make best use of the output of this meeting, and to integrate appropriately with other initiatives/programmes it is suggested that precedents for Maths/EPSRC-led high-impact interdisciplinary programmes are considered as methods to progress these discussions, e.g. Maths in Industry movement, EPSRC's Life Sciences Interface. Initial and longer –term actions are highlighted in Section 1, and these should be the considered the first step in establishing the vision/strategy for GCRF and the UK Mathematical Sciences.

6 Building a Community

This meeting brought together representatives from across the UK Mathematical Sciences community with researchers from DAC list countries. There was a strong desire to be involved and recognition that a co-ordinated approach across the community would be beneficial. It was suggested Mathematics in Industry community could be considered as a model of the type of activity that should be considered

It was recognised that this meeting represented the first step towards building a community. Initially there is a requirement for co-ordination activity, to information share, provide a mechanism for sharing contacts, highlighting examples of best practise etc. Accordingly, the initial actions from the meeting relate to providing the short term co-ordination effort required.

Beyond the immediate future, there needs to be consideration regarding how to best achieve the community building aspiration. Recommendation 4 of *The Era of Mathematics,* specifically mentions raising awareness of wider research challenges, including the sustainable development goals addressed in the GCRF). Therefore, the efforts to build a community should be co-ordinated with the outcome of the Bond Review. The most practical way to initiate this will be for the report authors to have discussions with EPSRC and then report back to meeting participants.